Wong-Zakai Approximation and Support Theorem for SPDEs with Locally Monotone Coefficients

Ting Ma[a], Rongchan Zhu[b,c, *†‡1]

[a]College of Mathematics, Sichuan University, Chengdu 610065, China
[b]Department of Mathematics, Beijing Institute of Technology, Beijing 100081, China
[c]Department of Mathematics, University of Bielefeld, D-33615 Bielefeld, Germany

Abstract

In this paper we present the Wong-Zakai approximation results for a class of nonlinear SPDEs with locally monotone coefficients and driven by multiplicative Wiener noise. This model extends the classical monotone one and includes examples like stochastic 2d Navier-Stokes equations, stochastic porous medium equations, stochastic p-Laplace equations and stochastic reaction-diffusion equations. As a corollary, our approximation results also describe the support of the distribution of solutions.

Mathematics Subject Classification: 60H15, 76M35

Keywords: Wong-Zakai approximation, local monotonicity, coercivity, support theorem.

1 Introduction

The Wong-Zakai type approximation problem has been intensively studied, since it was first investigated by Wong and Zakai ([26]) for SDEs driven by one-dimensional Brownian motion. It states that when replacing the driven noise by a suitable

*Research supported in part by NSFC (No.11671035). Financial support by the DFG through the CRC 1283 Taming uncertainty and profiting from randomness and low regularity in analysis, stochastics and their applications is acknowledged.
†Corresponding author
‡Email address: matingting2008@yeah.net(T. Ma), zhurongchan@126.com(R. C. Zhu)

smooth approximation (for example, piecewise linear approximation or convolution with a mollifier) and making a drift correction in the equation, the solutions to the approximating equations converge to the solution to the original SDE. Similar results have been extended to the multidimensional case (see, e.g. [27]). In infinite-dimensional case Wong-Zakai type approximation has also attracted a lot of attention. In [25] K. Twardowska claimed the convergence of Wong-Zakai approximation for infinite-dimensional equations with usual monotone and coercive coefficients. I. Chueshov and A. Millet in [5] studied Wong-Zakai approximation for stochastic 2d hydrodynamical systems, which can be applied to stochastic 2d Navier-Stokes equations. In [9, 10], I. Gyöngy, A. Shmatkov and P. R. Stinga obtained the rate of convergence for the Wong-Zakai approximation by estimating the convergence rate of the corresponding approximating noise.

However, most papers in the literature consider Wong-Zakai approximation for semi-linear equations in infinite dimensional case. In [9, 10, 25, 27] only SPDEs with monotone coefficients were considered. M. Hairer and É. Pardoux in their recent work [14] mainly concerned semi-linear SPDEs driven by space-time white noise with spatial variable in one dimension. Many interesting nonlinear equations have been studied a lot recently, especially quasi-linear equations, including stochastic porous medium equations and stochastic p-Laplace equations. We would like to know whether Wong-Zakai approximation results hold for these equations. For this purpose, we extend the Wong-Zakai approximation theorem to a class of nonlinear SPDEs driven by trace-class noise, where the coefficients satisfy local monotonicity condition, which can cover all the above nonlinear SPDEs (see Theorem 2.6). In fact, our main results can cover the results in [5] and stochastic 2d Navier-Stokes equations of course, if we choose the Gelfand triple $V \subset H \subset V^*$ as in [5] and define a function $\rho(\cdot)$ on V appropriately (see Section 3.1).

For PDEs and SPDEs with monotone coefficients, the variational framework is a basic approach for studying existence and uniqueness of the solutions to the equations, especially for those that are not semi-linear. A big difference from the semigroup approach (see [6]) is that the variational approach has no need of the semigroup generated by a linear operator in the drift term. This variational framework was initiated in the pioneering work of É. Pardoux [22] and further developed in studying equations with martingales as integrators in the noise term (e.g. [8, 11, 13, 15, 21]). In all the papers mentioned above, the coefficients satisfy the standard monotonicity and coercivity conditions ([15, 21]). Recently, this framework has been substantially extended by W. Liu and M. Röckner [16, 17] for a more general class of SPDEs, for which the monotonicity condition holds locally. Hence some more interesting examples, e.g., stochastic 2d Navier-Stokes equations, stochastic Burgers type equations can be covered in this framework.

We prove the Wong-Zakai approximation theorem under their framework.

Furthermore, the approximation theorem provides a description for the Stroock-Varadhan characterization (see [23]) of the topological support of the solutions to SDEs and SPDEs. The support theorem has been studied for SDEs in [12, 27]. For mild solutions to semi-linear equations, by properties of the corresponding semigroup for the linear operator, similar results have been obtained in [20] for SDEs in Hilbert spaces and [2, 3] for parabolic SPDEs. By [12, 18, 19] it is standard to conclude the support theorem using Wong-Zakai approximation theorem. Since the Wong-Zakai approximation results have been extended to the local monotonicity framework, we can also describe the support of the solutions to SPDEs under this framework (see Theorem 4.4).

The paper is organized as follows. In Section 2 we describe our framework introduced in [16] and then obtain the main approximation results in Theorem 2.6. In Section 3 we discuss examples satisfying our assumptions. In Section 4 we characterize the support of the distribution for the solutions. Appendix contains the proof of Lemma 4.1, which can be applied to obtain Theorem 2.4 for the existence and uniqueness of solutions to the approximating equations.

2 Framework and main result

Let $(H, \langle \cdot, \cdot \rangle)$ be a separable Hilbert space and identified with its dual space H^* by the Riesz isomorphism, and let $(V, \langle \cdot, \cdot \rangle_V)$ be a Hilbert space which is continuously and densely embedded into H. Then we have the following Gelfand triple

$$V \subset H \equiv H^* \subset V^*,$$

where V^* is the dual space of V. It follows that

$${}_{V^*}\langle h, v \rangle_V = \langle h, v \rangle, \quad \text{for all } h \in H,\, v \in V. \tag{2.1}$$

Let $\{W(t)\}_{t \geq 0}$ be a cylindrical Wiener process in a separable Hilbert space $(U, \langle \cdot, \cdot \rangle_U)$ on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}; \mathbb{P})$, where $(\mathcal{F}_t)_{t \geq 0}$ is the normal filtration generated by W. Let $(L_2(U; H), \| \cdot \|_{L_2})$ denote the space of all Hilbert-Schmidt operators from U to H. We consider Wong-Zakai approximation for the following stochastic evolution equation on H:

$$dX(t) = A(t, X(t))dt + B(X(t))dW(t), \tag{2.2}$$

where for any fixed time $T > 0$, the maps

$$A : [0, T] \times V \times \Omega \to V^*, \quad B : H \times \Omega \to (L_2(U; H), \| \cdot \|_{L_2})$$

are progressively measurable.

We now define an adapted finite-dimensional approximation of the process W. In fact, we have the representation

$$W(t)=\sum_{j=1}^{\infty}\beta_j(t)e_j,\quad t\in[0,T], \tag{2.3}$$

where β_j's are standard independent real-valued Brownian motions and $\{e_j, j\geq 1\}$ is an orthonormal basis in U. For $n\in\mathbb{N}$, we set $\delta=\frac{T}{2^n}$ and define

$$\dot{W}^n(t)=\sum_{j=1}^{n}\delta^{-1}[\beta_j(\lfloor\frac{t}{\delta}\rfloor\delta)-\beta_j((\lfloor\frac{t}{\delta}\rfloor-1)\delta)]e_j=:\sum_{j=1}^{n}\dot{\beta}_j^n(t)e_j,\ t\in[0,T]. \tag{2.4}$$

Here and in the following for $s\in[0,T]$, $\lfloor s\rfloor$ denotes the largest integer which is no more than s, and $\lceil s\rceil$ denotes the smallest integer which is larger than s. We always set $\beta_j(t)=0$ for $t\leq 0$ and $\beta_j(t)=\beta_j(T)$ for $t\geq T$; so $\dot{\beta}_j(t)=0$ for $t>T$. Then $\dot{\beta}_j^n(t), j=1,\ldots,n$ are $\mathcal{F}_t$-adapted and so is $\dot{W}^n(t)$.

For $j=1,\ldots,n$, let $B_j: H\to H$ be defined by $B_j(u)=B(u)e_j,\ u\in H$. We assume that for each j, B_j is Fréchet differentiable with its derivative denoted by $DB_j: H\to L(H,H)$. Then we define the map

$$\widetilde{tr}_n: H\to H,\quad \widetilde{tr}_n(u)=\sum_{j=1}^{n}DB_j(u)B_j(u),\quad u\in H. \tag{2.5}$$

Consider the following approximating equations

$$\begin{cases} dX^n(t)=A(t,X^n(t))dt+B(X^n(t))\dot{W}^n(t)dt-\frac{1}{2}\widetilde{tr}_n(X^n(t))dt\\ X^n(0)=\xi, \end{cases} \tag{2.6}$$

with $\dot{W}^n$, $\widetilde{tr}_n$ defined in (2.4) and (2.5), and the initial value ξ, the condition of which will be given later.

Below we give the main assumptions and notations.

2.1 Assumptions and Notations

Assumption 1. *There exist constants $K>0$, $\alpha>1$, $\theta>0$, $\beta\geq 0$ and a nonnegative adapted process $f\in L^{\frac{p}{2}}([0,T]\times\Omega; dt\otimes\mathbb{P})$ with $p>\beta+2$ such that the following conditions hold for all $v_1,v_2,v\in V$, $u_1,u_2,u\in H$, $t\in[0,T]$.*

(H1) (Hemicontinuity) The map $\lambda\mapsto {}_{V^*}\langle A(t,v_1+\lambda v_2),v\rangle_V$ is continuous on $\mathbb{R}$.

(H2) (Local monotonicity)

$$2_{V^*}\langle A(t,v_1)-A(t,v_2),v_1-v_2\rangle_V\leq (f(t)+\rho(v_2))\|v_1-v_2\|_H^2,$$

$$\|B(u_1)-B(u_2)\|_{L_2}^2\leq \rho'(u_2)\|u_1-u_2\|_H^2,$$

where the functions ρ on V and ρ' on H are measurable.

(H3) (Coercivity)

$$2_{V^*}\langle A(t,v),v\rangle_V \le K\|v\|_H^2 - \theta\|v\|_V^\alpha + f(t).$$

(H4) (Growth)

$$\|A(t,v)\|_{V^*}^{\frac{\alpha}{\alpha-1}} \le (f(t)+K\|v\|_V^\alpha)(1+\|v\|_H^\beta),$$
$$\|B(u)\|_{L_2}^2 \le K(1+\|u\|_H^2),$$
$$\rho(v) \le K(1+\|v\|_V^\alpha)(1+\|v\|_H^\beta),\ \rho'(u) \le K(1+\|u\|_H^\beta).$$

Remark 2.1. As we have mentioned in Introduction, Assumption 1, which originated from [16], is a general framework for the existence and uniqueness of solution to equation (2.2). Under this framework, well-posedness of a lot of interesting semi-linear and quasi-linear SPDEs have been obtained. Here we consider the terms A and B separately in (H2), which is a little different from that in [16]. But the same examples as in [16] can still be covered. This is required for obtaining existence and uniqueness of solutions to approximating equations (2.6), since we consider (2.6) as deterministic equations (the diffusion coefficient is 0) and we need to estimate the drift parts A and $B\dot{W}^n$ separately. More details can be found in proving Theorem 2.4.

For our approximation results, similarly as in [5], we give some regularity assumptions on the diffusion coefficient B.

Assumption 2. *For each $j\in\mathbb{N}$, the map B_j is twice Fréchet differentiable with its second Fréchet derivative denoted by $D^2B_j : H \to L(H, L(H,H)) \simeq L(H\times H, H)$, and satisfies that*

(P1) for any $N>0$, there exists a positive constant $C(N)$ such that

$$\sup_{j\in\mathbb{N}}\sup_{\|u\|_H\le N}\{\|DB_j(u)\|_{L(H,H)} \vee \|B_j(u)\|_H \vee \|D^2B_j(u)\|_{L(H\times H,H)}\} \le C(N),$$
$$DB_j^*|_V : V\to V,\quad \sup_{j\in\mathbb{N}}\sup_{\|u\|_H\le N}\|DB_j(u)^*v\|_V \le C(N)\|v\|_V,\ v\in V,$$

and for $m\in\mathbb{N}$,

$$\lim_{m\to\infty}\sup_{\|u\|_H\le N}\|B(u)-B(u)\circ\Pi_m\|_{L_2}=0,$$

where Π_m denotes the orthogonal projection onto $U_m := span\{e_1,\cdots,e_m\}$ in U, i.e. $\Pi_m x = \sum_{i=1}^m \langle x,e_i\rangle_U e_i$, $x\in U$. $DB_j(\cdot)^*$ denotes the dual operator of $DB_j(\cdot)$.

(P2) there exists a constant $C>0$ such that for every $n\in\mathbb{N}$ and $u,u_1,u_2\in H$

$$\|\widetilde{tr}_n(u)\|_H^2 \le C(1+\|u\|_H^2),$$
$$\langle \widetilde{tr}_n(u_2)-\widetilde{tr}_n(u_1), u_1-u_2\rangle \le \rho'(u_2)\|u_1-u_2\|_H^2,$$

where ρ' is given by (H4).

Assumption 2 is used to obtain well-posedness of the approximating equations (2.6), which is similar as the conditions for Wong-Zakai approximation in the literature (e.g. [14, 25–27]). We will give examples for which Assumption 2 holds in Section 3.

We recall the following definition from [16].

Definition 2.2. (cf. [16, Definition 1.1]) A continuous H-valued $(\mathcal{F}_t)$-adapted process $(X(t))_{t\in[0,T]}$ is called a solution to equation (2.2), if for its $dt\otimes\mathbb{P}$-equivalence class we have

$$X \in L^{\alpha}([0,T]\times\Omega; dt\otimes\mathbb{P}; V),$$

with $\alpha > 1$ in (H4) and $\mathbb{P}$-$a.s.$

$$X(t) = X(0) + \int_0^t A(s, X(s))ds + \int_0^t B(X(s))dW(s)\ ,\ t \in [0,T]. \tag{2.7}$$

The following conclusion is proved in [16].

Theorem 2.3. *(cf. [16, Theorem 1.1]) Suppose that Assumption 1 holds, then if further $\xi \in L^p(\Omega,\mathcal{F}_0,\mathbb{P};H)$ with p in Assumption 1, there exists a unique solution $(X(t))_{t\in[0,T]}$ to equation (2.2) such that $X(0) = \xi$ $\mathbb{P}$-a.e. and*

$$\mathbb{E}\Big(\sup_{t\in[0,T]} \|X(t)\|_H^p + \int_0^T \|X(t)\|_V^{\alpha} dt\Big) < \infty. \tag{2.8}$$

Under Assumptions 1 and 2, we similarly obtain existence and uniqueness of solutions to approximating equations (2.6).

Theorem 2.4. *Suppose that $\xi \in L^p(\Omega,\mathcal{F}_0,\mathbb{P};H)$ with $p > \{\frac{\alpha}{\alpha-1} \vee (\beta+2)\}$ in Assumption 1, then under Assumptions 1 and 2, there exist unique solutions $(X^n(t))_{t\in[0,T]}$ to equations (2.6) satisfying $X^n(0) = \xi$ $\mathbb{P}$-a.e. and*

$$X^n(t) = \xi + \int_0^t A(s, X^n(s))ds + \int_0^t B(X^n(s))\dot{W}^n(s)ds - \frac{1}{2}\int_0^t \widetilde{tr}_n(X^n(s))ds.$$

Moreover,

$$\sup_{n\geq 1}\mathbb{E}\Big(\sup_{t\in[0,T]} \|X^n(t)\|_H^p + \int_0^T \|X^n(t)\|_V^{\alpha} dt\Big) < \infty. \tag{2.9}$$

This is a special case of Lemma 4.1. Comparing with Theorem 2.3, we add in Theorem 2.4 the assumption $p > \frac{\alpha}{\alpha-1}$, which ensures that the drift parts $B\dot{W}^n$, $\widetilde{tr}_n$ and A stay in the same space. More details can be seen in Appendix.

2.2 Main Result

We first recall the following result from [5, Lemma 2.1], which is prepared for our proof.

Lemma 2.5. *Let $T > 0$, then there exists a constant $\gamma_0 > 0$ such that for every $\gamma > \gamma_0/\sqrt{T},\ t \in [0, T]$,*

$$\begin{aligned}
&\lim_{n\to\infty} \mathbb{P}\big(\sup_{1\le j\le n}\sup_{s\le t} |\dot\beta_j^n(s)| > \gamma n^{1/2}2^{n/2}\big) = 0,\\
&\lim_{n\to\infty} \mathbb{P}\big(\sup_{s\le t} \|\dot W^n(s)\|_U > \gamma n 2^{n/2}\big) = 0.
\end{aligned} \tag{2.10}$$

Now for $N \ge 0$, $n \in \mathbb{N}$, $\gamma > \gamma_0/\sqrt{T}$, we define the stopping times:

$$\begin{aligned}
\tau_N^{(1)} &: = \inf\Big\{t \in [0, T] : \|X(t)\|_H + \int_0^t (f(s) + \|X(s)\|_V^\alpha)ds > N\Big\} \wedge T,\\
\tau_{n,N}^{(2)} &: = \inf\Big\{t \in [0, T] : \|X^n(t)\|_H + \int_0^t \|X^n(s)\|_V^\alpha ds > N\Big\} \wedge T,\\
\tau_n^{(3)} &: = \inf\Big\{t \in [0, T] : \big[\sup_{s\in[0,t]}\sup_{1\le j\le n} \big|\dot\beta_j^n(s)\big|\big]\\
&\qquad\qquad \vee \big[n^{-1/2}\sup_{s\in[0,t]} \|\dot W^n(s)\|_U\big] > \gamma n^{1/2}2^{n/2}\Big\} \wedge T,
\end{aligned}$$

and

$$\tau_{n,N} := \tau_N^{(1)} \wedge \tau_{n,N}^{(2)} \wedge \tau_n^{(3)}. \tag{2.11}$$

Then by (2.8), (2.9) and (2.10), it is obvious that

$$\begin{aligned}
&\lim_{N\to\infty} \mathbb{P}\big(\tau_N^{(1)} = T\big) = 1; \lim_{n\to\infty} \mathbb{P}\big(\tau_n^{(3)} = T\big) = 1;\\
&\lim_{N\to\infty} \mathbb{P}\big(\tau_{n,N}^{(2)} = T\big) = 1 \ \text{ uniformly for } n \in \mathbb{N}.
\end{aligned}$$

Below we state the Wong-Zakai approximation results.

Theorem 2.6. *Suppose that $\xi \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ with $p > \{\frac{\alpha}{\alpha-1} \vee (\beta + 2)\}$ in Assumption 1, and that Assumptions 1 and 2 hold. Let X and X^n be the solutions to equations* (2.2) *and* (2.6) *with the same initial condition ξ, respectively. Then*

$$\lim_{n\to\infty} \mathbb{E}\big(\sup_{t\in[0,T]} \|X(t) - X^n(t)\|_H^2\big) = 0. \tag{2.12}$$

Remark 2.7. When comparing X^n with X, the most significant difference is that the term $\int_0^\cdot B(X^n(s))\dot W^n(s)ds$ cannot be expressed as stochastic integral directly. Instead, we consider an additional term as follows. Using the identity $\int_0^t =$

$\sum_{k=0}^{\lfloor \frac{t}{\delta} \rfloor} \int_{k\delta}^{(k+1)\delta\wedge t}$ and (2.4), we plus and then minus for correctness on the right-hand side of (2.6)

$$\begin{aligned}
&\int_0^t B(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))\dot{W}^n(s)ds \\
&= \sum_{k=0}^{\lfloor \frac{t}{\delta} \rfloor} \int_{k\delta}^{(k+1)\delta\wedge t} ds \frac{1}{\delta} \int_{(k-1)\delta}^{k\delta} B(X^n((k-1)\delta))\Pi_n dW(u) \\
&= \sum_{k=0}^{\lfloor \frac{t}{\delta} \rfloor} \int_{(k-1)\delta}^{k\delta} \big(\frac{1}{\delta} \int_{k\delta}^{(k+1)\delta\wedge t} B(X^n((k-1)\delta))du\big)\Pi_n dW(s) \qquad (2.13) \\
&= \sum_{k=0}^{\lfloor \frac{t}{\delta} \rfloor} \int_{(k-1)\delta}^{k\delta} \big(\frac{1}{\delta} \int_{\lceil \frac{s}{\delta} \rceil\delta}^{(\lceil \frac{s}{\delta} \rceil+1)\delta} 1_{\{u\le t\}} B(X^n(\lfloor \frac{s}{\delta} \rfloor\delta))du\big)\Pi_n dW(s) \\
&= \int_0^t \big(\frac{1}{\delta} \int_{\lceil \frac{s}{\delta} \rceil\delta}^{(\lceil \frac{s}{\delta} \rceil+1)\delta} 1_{\{u\le t\}} du\big) B(X^n(\lfloor \frac{s}{\delta} \rfloor\delta))\Pi_n dW(s),
\end{aligned}$$

where we used stochastic Fubini's theorem in the second equality. To obtain the last equality we added an extra term in the third equality

$$\int_{\lfloor \frac{t}{\delta} \rfloor\delta}^t \big(\frac{1}{\delta} \int_{\lceil \frac{s}{\delta} \rceil\delta}^{(\lceil \frac{s}{\delta} \rceil+1)\delta} 1_{\{u\le t\}} du\big) B(X^n(\lfloor \frac{s}{\delta} \rfloor\delta))\Pi_n dW(s),$$

which is equal to 0 since $\lceil \frac{s}{\delta} \rceil\delta > t$ for any $s \in [\lfloor \frac{t}{\delta} \rfloor\delta, t]$. In our proof of Theorem 2.6 below, we actually use the term in (2.13) instead of $\int_0^t B(X^n(s))\dot{W}^n(s)ds$ to compare with the corresponding diffusion term $\int_0^t B(X(s))dW(s)$ in (2.2).

Proof of Theorem 2.6. It is sufficient to prove that for $N > 0$ large enough,

$$\lim_{n\to\infty} \mathbb{E}\big(\sup_{t\in[0,\tau_{n,N}]} \|X(t) - X^n(t)\|_H^2\big) = 0 \qquad (2.14)$$

with $\tau_{n,N}$ given in (2.11). In fact, set $\Omega_{n,N} := \{\omega \in \Omega : \tau_{n,N} = T\}, N > 0$, using Hölder's inequality

$$\mathbb{E}\big(\sup_{t\in[0,T]} 1_{\Omega_{n,N}^c} \|X(t) - X^n(t)\|_H^2\big) \le \mathbb{P}(\Omega_{n,N}^c)^{\frac{p-2}{p}} \mathbb{E}\big(\sup_{t\in[0,T]} \|X(t) - X^n(t)\|_H^p\big)^{\frac{2}{p}}.$$

Then (2.8), (2.9) and (2.11) imply that for any $\epsilon > 0$, there exists a constant N_ϵ (independent of n) large enough such that the term on the right-hand side of the above inequality is smaller than $\epsilon/2$ for n large enough. Fix such N_ϵ we choose as above, we still denote it by N and denote $\tau_{n,N}$ by τ_n for simplicity. Hence

$$\begin{aligned}
\mathbb{E}\Big(\sup_{t\in[0,T]} \|X(t) - X^n(t)\|_H^2\Big) =&\mathbb{E}\Big(\sup_{t\in[0,T]} [1_{\Omega_{n,N}^c} + 1_{\Omega_{n,N}}]\|X(t) - X^n(t)\|_H^2\Big) \\
\le&\epsilon/2 + \mathbb{E}\Big(\sup_{t\in[0,\tau_n]} \|X(t) - X^n(t)\|_H^2\Big).
\end{aligned}$$

Following the definition of τ_n, for some fixed constant γ with $\gamma > \gamma_0/\sqrt{T}$, there exists a constant $C(N)$ such that for all $t \in [0, \tau_n]$ and $j = 1, \ldots, n$

$$\|X(t)\|_H + \|X^n(t)\|_H \le C(N), \quad \int_0^t \big(\|X(s)\|_V^\alpha + \|X^n(s)\|_V^\alpha\big) ds \le C(N), \quad (2.15)$$

$$|\dot{\beta}_j^n(t)| + n^{-1/2}\|\dot{W}^n(t)\|_U \le 2\gamma n^{1/2} 2^{n/2}. \quad (2.16)$$

This property will be used repeatedly throughout this section. Constants may change from line to line, but we indicate their dependence on parameters when necessary. By (2.13) we obtain the following decomposition

$$\begin{aligned} X^n(t) - X(t) = &\int_0^t \Big(A(s, X^n(s)) - A(s, X(s))\Big) ds \\ &+ \int_0^t \Big(\big(\frac{1}{\delta} \int_{\lceil \frac{s}{\delta} \rceil \delta}^{(\lceil \frac{s}{\delta} \rceil + 1)\delta} 1_{\{u \le t\}} du\big) B(X^n(\lfloor \frac{s}{\delta} \rfloor \delta)) \Pi_n - B(X(s)) \Big) dW(s) \\ &+ \int_0^t \Big(\big[B(X^n(s)) - B(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))\big] \dot{W}^n(s) - \frac{1}{2} \widetilde{tr}_n(X^n(s)) \Big) ds. \end{aligned}$$

Now applying Itô's formula ([17, Theorem 4.2.5]) for $X^n(t) - X(t)$, we obtain

$$\begin{aligned} &\|X^n(t) - X(t)\|_H^2 \\ =& \int_0^t \Big(2_{V^*}\langle A(s, X^n(s)) - A(s, X(s)), X^n(s) - X(s) \rangle_V ds \\ &+ \|\big(\frac{1}{\delta} \int_{\lceil \frac{s}{\delta} \rceil \delta}^{(\lceil \frac{s}{\delta} \rceil + 1)\delta} 1_{\{u \le t\}} du\big) B(X^n(\lfloor \frac{s}{\delta} \rfloor \delta)) \Pi_n - B(X(s))\|_{L_2}^2 ds \\ &+ 2\langle X^n(s) - X(s), \big[\big(\frac{1}{\delta} \int_{\lceil \frac{s}{\delta} \rceil \delta}^{(\lceil \frac{s}{\delta} \rceil + 1)\delta} 1_{\{u \le t\}} du\big) B(X^n(\lfloor \frac{s}{\delta} \rfloor \delta)) \Pi_n - B(X(s))\big] dW(s) \rangle \\ &+ 2\langle [B(X^n(s)) - B(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))] \dot{W}^n(s) - \frac{1}{2} \widetilde{tr}_n(X^n(s)), X^n(s) - X(s) \rangle ds \Big). \end{aligned} \quad (2.17)$$

Then the procedure of estimate to (2.17) will be mainly divided into three steps.
Step 1: We will obtain that there exists $o(1) \stackrel{n\to\infty}{\longrightarrow} 0$ such that

$$\begin{aligned} &\mathbb{E}\Big(\sup_{t \in [0, \tau_n]} (\|X^n(t) - X(t)\|_H^2) \Big) \\ \le& \mathbb{E}\Big(\sup_{t \in [0, \tau_n]} \int_0^t 4\langle [B(X^n(s)) - B(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))] \dot{W}^n(s) \\ &\qquad - \frac{1}{2} \widetilde{tr}_n(X^n(s)), X^n(s) - X(s) \rangle ds \Big) + o(1) \\ &+ \mathbb{E}\Big(\int_0^{\tau_n} 2\big(f(s) + \rho(X(s)) + 72\rho'(X(s))\big) \|X^n(s) - X(s)\|_H^2 ds \Big). \end{aligned} \quad (2.18)$$

Step 2: We will prove the right-hand side in (2.18)

$$\begin{aligned} \lim_{n\to\infty} \mathbb{E}\Big(\sup_{t \in [0, \tau_n]} \int_0^t &\langle [B(X^n(s)) - B(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))] \dot{W}^n(s) \\ &- \frac{1}{2} \widetilde{tr}_n(X^n(s)), X^n(s) - X(s) \rangle ds \Big) = 0. \end{aligned} \quad (2.19)$$

Then inserting (2.19) into (2.18), there exist constants $A_n \stackrel{n\to\infty}{\longrightarrow} 0$ such that

$$\begin{aligned}&\mathbb{E}\Big(\sup_{s\in[0,t\wedge\tau_n]}\|X^n(s)-X(s)\|_H^2\Big)\\ \le& A_n+\mathbb{E}\Big(\int_0^{t\wedge\tau_n}2\big(f(s)+\rho(X(s))+72\rho'(X(s))\big)\|X^n(s)-X(s)\|_H^2ds\Big).\end{aligned} \tag{2.20}$$

Step 3: Set $F(t):=\sup_{s\in[0,t]}\|X^n(s)-X(s)\|_H^2,\ Z(t):=\int_0^t 2\big(f(s)+\rho(X(s))+72\rho'(X(s))\big)ds$ for $t\in[0,T]$. Then F and Z are adapted, nonnegative and continuous. By (2.11) and (H4) we have a constant $C'(N)$ such that $Z(t)\le C'(N)$ uniformly for $t\in[0,\tau_n]$. Then we rewrite (2.20) as

$$\mathbb{E}\Big(F(\tau_n)\Big)\le A_n+\mathbb{E}\Big(\int_0^{\tau_n}F(s)dZ(s)\Big).$$

By [7, Lemma 2.2] we have

$$\mathbb{E}\Big(\int_0^{\tau_n}F(s)dZ(s)\Big)\le A_ne^{C'(N)}\int_0^{C'(N)}e^{-y}dy,$$

and by (2.20) we obtain that $\mathbb{E}\big(\sup_{t\in[0,\tau_n]}F(t)\big)\stackrel{n\to\infty}{\longrightarrow}0$. The proof is completed. □

Below we prove (2.18) and (2.19) successively.

Proof of (2.18). We denote the four terms on the right-hand side of (2.17) by $D_n(t,i), i=1,\dots,4$ respectively, i.e.

$$\|X^n(t)-X(t)\|_H^2=\sum_{i=1}^4 D_n(t,i). \tag{2.21}$$

For $D_n(t,1)$, by (H2) we have

$$D_n(t,1)\le\int_0^t\big(f(s)+\rho(X(s))\big)\|X^n(s)-X(s)\|_H^2ds. \tag{2.22}$$

For $D_n(t,2)$, we first see that $D_n(s,2)$ is dominated for uniform $s\in[0,t\wedge\tau_n]$ by

$$\begin{aligned}&\int_0^{t\wedge\tau_n}2\|(\frac{1}{\delta}\int_{\lceil\frac{s}{\delta}\rceil\delta}^{(\lceil\frac{s}{\delta}\rceil+1)\delta}1_{\{u>t\wedge\tau_n\}}du)B(X^n(\lfloor\frac{s}{\delta}\rfloor\delta))\|_{L_2}^2ds\\&+\int_0^{t\wedge\tau_n}8\|B(X^n(s))\Pi_n-B(X^n(s))\|_{L_2}^2ds\\&+\int_0^{t\wedge\tau_n}4\|B(X^n(\lfloor\frac{s}{\delta}\rfloor\delta))-B(X^n(s))\|_{L_2}^2ds\\&+\int_0^{t\wedge\tau_n}8\|B(X^n(s))-B(X(s))\|_{L_2}^2ds.\end{aligned} \tag{2.23}$$

In (2.23), the first term is dominated by $2\int_{t\wedge\tau_n-2\delta}^{t\wedge\tau_n}\|B(X^n(\lfloor\frac{s}{\delta}\rfloor\delta))\|_{L_2}^2ds$, which by (H4) and (2.15) converges to zero. By (P1) and (2.15) we see that the second term also converges to zero. It is required that the third term converges to zero and that the convergence rate is $\delta^{1/2+\epsilon}$ for $\epsilon > 0$ (see Lemma 2.8 below). Then by Lemma 2.8, (H2), (H4) and (2.15), we have for the third term in (2.23)

$$\lim_{n\to\infty} \mathbb{E}\Big(\int_0^{\tau_n} \|B(X^n(\lfloor \frac{s}{\delta} \rfloor \delta)) - B(X^n(s))\|_{L_2}^2 ds \Big) = 0.$$

Using (H2) for the forth term in (2.23) and putting all these estimates together, we deduce that there exists $o(1) \overset{n\to\infty}{\longrightarrow} 0$ such that

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]} D_n(t,2) \Big) \le o(1) + \mathbb{E}\Big(\int_0^{\tau_n} 8\rho'(X(s)) \|X^n(t) - X(t)\|_H^2 dt \Big). \qquad (2.24)$$

For $D_n(t,3)$, using the B-D-G inequality, $\mathbb{E}(\sup_{t\in[0,\tau_n]} |D_n(t,3)|)$ is dominated by

$$\mathbb{E}\Big(4 \sup_{t\in[0,\tau_n]} \|X^n(t) - X(t)\|_H \cdot [\int_0^{\tau_n} \|(\frac{1}{\delta} \int_{\lceil \frac{s}{\delta} \rceil \delta}^{(\lceil \frac{s}{\delta} \rceil + 1)\delta} 1_{\{u\le\tau_n\}} du) B(X^n(\lfloor \frac{s}{\delta} \rfloor \delta)) \Pi_n - B(X(s))\|_{L_2}^2 ds]^{1/2} \Big).$$

By Young's inequality, we further obtain that

$$\mathbb{E}(\sup_{t\in[0,\tau_n]} |D_n(t,3)|) \le \frac{1}{2} \mathbb{E}\big(\sup_{t\in[0,\tau_n]} \|X^n(t) - X(t)\|_H^2 \big) + 8\mathbb{E}\big(\sup_{t\in[0,\tau_n]} |D_n(t,2)| \big). \quad (2.25)$$

Hence inserting (2.22), (2.24) and (2.25) into (2.21), we obtain (2.18). □

Lemma 2.8. *Let τ_n be defined by* (2.11)*, then under the assumptions in Theorem 2.6, there exists a constant $C(N, T, \|f\|_{L^{p/2}})$ such that*

$$\begin{aligned} &\mathbb{E}\Big(\int_0^{\tau_n} \|X(s) - X(\lfloor \frac{s}{\delta} \rfloor \delta)\|_H^2 ds \Big) \le C(N, T, \|f\|_{L^{p/2}}) 2^{-\frac{3}{4}n}, \\ &\mathbb{E}\Big(\int_0^{\tau_n} \|X^n(s) - X^n(\lfloor \frac{s}{\delta} \rfloor \delta)\|_H^2 ds \Big) \le C(N, T, \|f\|_{L^{p/2}}) 2^{-\frac{3}{4}n}. \end{aligned} \qquad (2.26)$$

Remark 2.9. Actually, in a similar way we obtain the results below:

$$\begin{aligned} &\mathbb{E}\Big(\int_0^{\tau_n} \|X(s) - X((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)\|_H^2 ds \Big) \le C(N, T, \|f\|_{L^{p/2}}) 2^{-\frac{3}{4}n}, \\ &\mathbb{E}\Big(\int_0^{\tau_n} \|X(s) - X(\lceil \frac{s}{\delta} \rceil \delta)\|_H^2 ds \Big) \le C(N, T, \|f\|_{L^{p/2}}) 2^{-\frac{3}{4}n}. \end{aligned} \qquad (2.27)$$

The results also hold when X is replaced by X^n in (2.27).

Proof of (2.19). The main idea is to find a suitable term from $[B(X^n(s)) - B(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))] \dot{W}^n(s)$ to compensate the correction term $-\frac{1}{2}\widetilde{tr}_n(X^n(s))$. For this purpose, by (2.4) and (2.5) we equivalently write

$$\widetilde{tr}_n(X^n) = \sum_{j=1}^n DB_j(X^n) B_j(X^n), \quad B(X^n)\dot{W}^n = \sum_{j=1}^n B_j(X^n) \dot{\beta}_j^n. \qquad (2.28)$$

Since for $j \in \mathbb{N}$, B_j is twice Fréchet differentiable, we apply the second-order Taylor's formula to B_j and have

$$\begin{aligned}
&B_j(X^n(s)) - B_j(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)) \\
=&DB_j(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))[X^n(s) - X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)] \\
&+ \int_0^1 (1-\mu) D^2 B_j(\mu X^n(s) + (1-\mu) X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)) d\mu \\
&\quad \{X^n(s) - X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta), X^n(s) - X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)\},
\end{aligned} \tag{2.29}$$

where $D^2 B_j(v)\{v_1, v_2\}$ denotes the value of the second Fréchet derivative $D^2 B_j(v)$ on elements v_1 and v_2. We rewrite the term $X^n(s) - X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)$ in (2.29) as $\int_{(\lfloor \frac{s}{\delta} \rfloor - 1)\delta}^s \dot{X}^n(u) du$, with $\dot{X}^n$ formulated by (2.6). Then $X^n(s) - X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)$ equals to

$$\int_{(\lfloor \frac{s}{\delta} \rfloor - 1)\delta}^s \big(A(u, X^n(u)) + B(X^n(u))\dot{W}^n(u) - \frac{1}{2}\widetilde{tr}_n(X^n(u))\big) du. \tag{2.30}$$

Using (2.4) and the second equality in (2.28), the term $\int_{(\lfloor \frac{s}{\delta} \rfloor - 1)\delta}^s B(X^n(u))\dot{W}^n(u) du$ in (2.30) equals to

$$\sum_{q=1}^n \big[\dot{\beta}_q^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta) \int_{(\lfloor \frac{s}{\delta} \rfloor - 1)\delta}^{\lfloor \frac{s}{\delta} \rfloor \delta} B_q(X^n(u)) du + \dot{\beta}_q^n(s) \int_{\lfloor \frac{s}{\delta} \rfloor \delta}^s B_q(X^n(u)) du\big]. \tag{2.31}$$

Then inserting (2.28)-(2.31) into (2.19), we have the following decomposition:

$$\begin{aligned}
&\int_0^t \langle [B(X^n(s)) - B(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta))]\dot{W}^n(s) - \frac{1}{2}\widetilde{tr}_n(X^n(s)), \\
&\qquad X^n(s) - X(s)\rangle ds =: \sum_{i=1}^6 J_n(t, i),
\end{aligned} \tag{2.32}$$

with

$$\begin{aligned}
J_n(t,1) :=& \sum_{j=1}^n \int_0^t \dot{\beta}_j^n(s) \langle DB_j(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)) \\
&\qquad \int_{(\lfloor \frac{s}{\delta} \rfloor - 1)\delta}^s A(u, X^n(u)), X^n(s) - X(s)\rangle du ds, \\
J_n(t,2) :=& \sum_{j=1}^n \sum_{q=1}^n \int_0^t \dot{\beta}_j^n(s) \dot{\beta}_q^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta) \langle DB_j(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)) \\
&\qquad \int_{(\lfloor \frac{s}{\delta} \rfloor - 1)\delta}^{\lfloor \frac{s}{\delta} \rfloor \delta} B_q(X^n(u)) du, X^n(s) - X(s)\rangle ds, \\
J_n(t,3) :=& \sum_{j=1}^n \sum_{1 \le q \le n, q \ne j} \int_0^t \dot{\beta}_j^n(s) \dot{\beta}_q^n(s) \langle DB_j(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)) \\
&\qquad \int_{\lfloor \frac{s}{\delta} \rfloor \delta}^s B_q(X^n(u)) du, X^n(s) - X(s)\rangle ds,
\end{aligned}$$

$$J_n(t,4) := \sum_{j=1}^n \int_0^t \langle \dot\beta_j^n(s)^2 DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)) \int_{\lfloor\frac{s}{\delta}\rfloor\delta}^s B_j(X^n(u))du$$
$$- \frac{1}{2} DB_j(X^n(s))B_j(X^n(s)), X^n(s)-X(s)\rangle ds,$$
$$J_n(t,5) := -\frac{1}{2}\sum_{j=1}^n \int_0^t \dot\beta_j^n(s)\langle DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))$$
$$\int_{(\lfloor\frac{s}{\delta}\rfloor-1)\delta}^s \widetilde{tr}_n(X^n(u))du, X^n(s)-X(s)\rangle ds,$$
$$J_n(t,6) := \sum_{j=1}^n \int_0^t \dot\beta_j^n(s)\langle \int_0^1 (1-\mu)D^2B_j(\mu X^n(s)+(1-\mu)X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))d\mu$$
$$\{X^n(s)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta), X^n(s)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\}, X^n(s)-X(s)\rangle ds.$$

In the following we estimate each term separately.

Estimate of $J_n(t,1)$: By (2.1) and (P1) we obtain an equivalent representation

$$J_n(t,1) = \sum_{j=1}^n \int_0^t \dot\beta_j^n(s) \int_{(\lfloor\frac{s}{\delta}\rfloor-1)\delta}^s {}_{V^*}\langle A(u,X^n(u)),$$
$$DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))^*[X^n(s)-X(s)]\rangle_V duds.$$

Then by (2.16) and (P1) we have

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|J_n(t,1)|\Big)\\
\le& C(N)\sum_{j=1}^n \mathbb{E}\Big(\int_0^{\tau_n} |\dot\beta_j^n(s)|[\int_{(\lfloor\frac{s}{\delta}\rfloor-1)\delta}^s \|A(u,X^n(u))\|_{V^*}^{\frac{\alpha}{\alpha-1}}du]^{\frac{\alpha-1}{\alpha}}\\
&\cdot[\int_{(\lfloor\frac{s}{\delta}\rfloor-1)\delta}^s \|X^n(s)-X(s)\|_V^\alpha du]^{\frac{1}{\alpha}}ds\Big)\\
\le& C(N)n^{3/2}2^{n/2}\Big(\mathbb{E}\int_0^{\tau_n} ds\int_{(\lfloor\frac{s}{\delta}\rfloor-1)\delta}^s \big(f(u)+K\|X^n(u)\|_V^\alpha\big)du\Big)^{\frac{\alpha-1}{\alpha}}\delta^{\frac{1}{\alpha}}\\
\le& C(N)n^{3/2}2^{n/2}\Big(\delta\cdot\mathbb{E}\int_0^{\tau_n}\big(f(u)+\|X^n(u)\|_V^\alpha\big)du\Big)^{\frac{\alpha-1}{\alpha}}\delta^{\frac{1}{\alpha}}\\
\le& C(N,T,\|f\|_{L^{p/2}})n^{3/2}2^{-n/2}.
\end{aligned}\tag{2.33}$$

Here we used (H4) and Hölder's inequality in the second inequality, and in the third inequality by stochastic Fubini's theorem we used

$$\begin{aligned}
&\mathbb{E}\Big(\int_0^{\tau_n} ds\int_{(\lfloor\frac{s}{\delta}\rfloor-1)\delta}^s \big(f(u)+\|X^n(u)\|_V^\alpha\big)du\Big)\\
\le&\mathbb{E}\Big(\sum_{k=0}^{\lfloor\frac{\tau_n}{\delta}\rfloor}\int_{(k-1)\delta}^{(k+1)\delta\wedge\tau_n}du\int_{k\delta}^{(k+1)\delta\wedge\tau_n}\big(f(u)+\|X^n(u)\|_V^\alpha\big)ds\Big)\\
\le&2\delta\mathbb{E}\Big(\int_0^{\tau_n}\big(f(u)+\|X^n(u)\|_V^\alpha\big)du\Big).
\end{aligned}\tag{2.34}$$

Estimate of $J_n(t,2)$: Using (2.15), (2.16) and (P1) we roughly see that $J_n(t,2)$ is dominated by $C(N)n^3$ uniformly for $t \in [0,\tau_n]$, which is useless for our proof. Fortunately, the Gaussian r.v.s $\dot\beta_j^n(s)$ and $\dot\beta_q^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)$ appeared in $J_n(t,2)$ are independent, it means that for any $s\in[0,\tau_n]$, $j,q=1,\dots,n$, $\mathbb{E}(\dot\beta_j^n(s)\dot\beta_q^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))=0$, by which we can give more precise calculations for $J_n(t,2)$. Using the first-order Taylor's formula for each B_q and we have for $u\in[0,\tau_n]$,

$$
\begin{aligned}
&B_q(X^n(u)) = B_q(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))\\
&+\int_0^1 DB_q(\nu X^n(u)+(1-\nu)X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))[X^n(u)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)]d\nu,
\end{aligned}
$$

which yields that $J_n(t,2)$ is split into the sum

$$
J_n(t,2) := \sum_{j=1}^{3} M_n(t,j),
$$

where

$$
\begin{aligned}
M_n(t,1) := \sum_{j,q=1}^{n}\delta\int_0^t &\dot\beta_j^n(s)\dot\beta_q^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\langle DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))B_q(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)),\\
&X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)-X((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\rangle ds,
\end{aligned}
$$

$$
\begin{aligned}
M_n(t,2) := \sum_{j,q=1}^{n}\delta\int_0^t &\dot\beta_j^n(s)\dot\beta_q^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\langle DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))B_q(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)),\\
&X^n(s)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)+X((\lfloor\frac{s}{\delta}\rfloor-1)\delta)-X(s)\rangle ds,
\end{aligned}
$$

$$
\begin{aligned}
M_n(t,3) := \sum_{j,q=1}^{n}\int_0^t &\dot\beta_j^n(s)\dot\beta_q^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\langle DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))\\
&\int_{(\lfloor\frac{s}{\delta}\rfloor-1)\delta}^{\lfloor\frac{s}{\delta}\rfloor\delta}du\int_0^1 DB_q(\nu X^n(u)+(1-\nu)X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))d\nu\\
&\qquad[X^n(u)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)],X^n(s)-X(s)\rangle ds.
\end{aligned}
$$

To estimate $J_n(t,2)$, we define a martingale in (2.35) below, and use its property and the independence of β_j.

For $M_n(t,1)$, since by (2.4) $\dot\beta_j^n(s)\equiv\dot\beta_j^n(k\delta)$ for $s\in[k\delta,(k+1)\delta)$, equivalently

$$
\begin{aligned}
M_n(t,1) =: &\sum_{k=1}^{\lfloor\frac{t}{\delta}\rfloor-1}\sum_{j,q=1}^{n}\delta^2\dot\beta_j^n(k\delta)\dot\beta_q^n((k-1)\delta)f_{k-1,j,q}\\
&+\sum_{j,q=1}^{n}\delta(t-\lfloor\frac{t}{\delta}\rfloor\delta)\dot\beta_j^n(\lfloor\frac{t}{\delta}\rfloor\delta)\dot\beta_q^n((\lfloor\frac{t}{\delta}\rfloor-1)\delta)g_{j,q}(t)\\
=:&H_n(t,1)+H_n(t,2),
\end{aligned}
$$

with the coefficients $f_{k-1,j,q}, g_{j,q}(t)$ given by

$$f_{k-1,j,q} := \langle DB_j(X^n((k-1)\delta))B_q(X^n((k-1)\delta)), X^n((k-1)\delta) - X((k-1)\delta)\rangle,$$

$$g_{j,q}(t) := \langle DB_j(X^n((\lfloor\frac{t}{\delta}\rfloor - 1)\delta))B_q((\lfloor\frac{t}{\delta}\rfloor - 1)\delta), X^n((\lfloor\frac{t}{\delta}\rfloor - 1)\delta) - X((\lfloor\frac{t}{\delta}\rfloor - 1)\delta)\rangle.$$

For $j, q, k \in \mathbb{N}$, the centered r.v $\dot\beta_j^n(k\delta)$ and the coefficient $f_{k,j,q}$ are $\mathcal{F}_{k\delta}$-measurable, $\dot\beta_j^n(k\delta)$ is $\mathcal{F}_{(k-1)\delta}$-independent. The local boundedness of DB_j and B_j yields that there exists a positive constant $C(N)$ such that for all $t \in [0, \tau_n]$, $k = 0, \dots, \lfloor\frac{\tau_n}{\delta}\rfloor$, $j = 1, \dots, n$, the coefficients $f_{k,j,q}$, $g_{j,q}(t)$ are bounded by $C(N)$. Thus for $l = 1, \dots, \lfloor\frac{\tau_n}{\delta}\rfloor$, the process defined by

$$\sum_{k=1}^{l}\sum_{j,q=1}^{n}\dot\beta_j^n(k\delta)\dot\beta_q^n((k-1)\delta)f_{k-1,j,q} \tag{2.35}$$

is an $(\mathcal{F}_{l\delta})$-martingale. Furthermore, for any $k_1 \neq k_2 \in \mathbb{N}$, $j_1, j_2, q_1, q_2 = 1, \dots, n$, $\dot\beta_{j_1}^n(k_1\delta)$ and $\dot\beta_{j_2}^n(k_2\delta)$ are independent, and satisfy

$$\mathbb{E}\big(\dot\beta_{j_1}^n(k_1\delta)\dot\beta_{q_1}^n((k_1-1)\delta)f_{k_1-1,j_1,q_1}\dot\beta_{j_2}^n(k_2\delta)\dot\beta_{q_2}^n((k_2-1)\delta)f_{k_2-1,j_2,q_2}\big) = 0. \tag{2.36}$$

Doob's inequality implies that

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|H_n(t,1)|\Big) &\le 2\delta^2\Big(\mathbb{E}\Big\{\sum_{k=1}^{\lfloor\frac{\tau_n}{\delta}\rfloor-1}\sum_{j,q=1}^{n}\dot\beta_j^n(k\delta)\dot\beta_q^n((k-1)\delta)f_{k-1,j,q}\Big\}^2\Big)^{1/2}\\ &= 2\delta^2\Big(\mathbb{E}\sum_{k=1}^{\lfloor\frac{\tau_n}{\delta}\rfloor-1}\Big\{\sum_{j,q=1}^{n}\dot\beta_j^n(k\delta)\dot\beta_q^n((k-1)\delta)f_{k-1,j,q}\Big\}^2\Big)^{1/2},\end{aligned}$$

where we used (2.36) in the equality and by (2.16) is further dominated by $C(N,T)n^3 2^{-n/2}$. Using (2.16) again we have

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|H_n(t,2)|\Big) \le C(N)T^2 n^3 2^{-n}.$$

Then

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|M_n(t,1)|\Big) \le C(N,T)n^3 2^{-n/2}. \tag{2.37}$$

For $M_n(t,2)$, the boundedness of DB_j, B_j and (2.16) imply that

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|M_n(t,2)|\Big) &\le C(N)Tn^3\mathbb{E}\Big(\int_0^{\tau_n}\|X^n(s) - X^n((\lfloor\frac{s}{\delta}\rfloor - 1)\delta)\|_H ds\Big)\\ &\quad + C(N)Tn^3\mathbb{E}\Big(\int_0^{\tau_n}\|X((\lfloor\frac{s}{\delta}\rfloor - 1)\delta) - X(s)\|_H ds\Big),\end{aligned}$$

which combined with Lemma 2.8 and Hölder's inequality, implies that

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|M_n(t,2)|\Big) \le C(N,T,\|f\|_{L^{p/2}})n^3 2^{-3n/8}. \tag{2.38}$$

Similarly, we also have

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|M_n(t,3)|\Big) \le C(N,T,\|f\|_{L^{p/2}})n^3 2^{-3n/8}. \tag{2.39}$$

Estimate of $J_n(t,3)$: Let $q\in\mathbb{N}$, we use the first-order Taylor's formula for B_q, and by a similar computation as that for $J_n(t,2)$ we deduce that

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|J_n(t,3)|\Big) \le C(N,T,\|f\|_{L^{p/2}})n^3 2^{-3n/8}. \tag{2.40}$$

Estimate of $J_n(t,5)$: (P1), (P2) and (2.16) yield that

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|J_n(t,5)|\Big) \le C(N,T)n^{3/2} 2^{-n/2}. \tag{2.41}$$

Estimate of $J_n(t,6)$: By Lemma 2.8, (2.16) and (P1) we have

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|J_n(t,6)|\Big) \le C(N,T,\|f\|_{L^{p/2}})n^{3/2} 2^{-n/4}. \tag{2.42}$$

Estimate of $J_n(t,4)$: We deal with it similarly as for $J_n(t,2)$. Let $j\in\mathbb{N}$, we use the first-order Taylor's formula for B_j and have

$$J_n(t,4) =: \sum_{i=1}^{6} Z_n(t,i),$$

where

$$\begin{aligned} Z_n(t,1) :=& \sum_{j=1}^{n}\int_0^t \dot{\beta}_j^n(s)^2\langle DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)) \\ &\qquad \int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\int_0^1 DB_j(\nu X^n(u)+(1-\nu)X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))d\nu \\ &\qquad\qquad [X^n(u)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)]du, X^n(s)-X(s)\rangle ds, \end{aligned}$$

$$\begin{aligned} Z_n(t,2) :=& \sum_{j=1}^{n}\int_0^t (s-\lfloor\frac{s}{\delta}\rfloor\delta)\dot{\beta}_j^n(s)^2\langle DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta))B_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)), \\ &\qquad X^n(s)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)+X((\lfloor\frac{s}{\delta}\rfloor-1)\delta)-X(s)\rangle ds, \\ Z_n(t,3) :=& \sum_{j=1}^{n}\int_0^t \langle[(s-\lfloor\frac{s}{\delta}\rfloor\delta)\dot{\beta}_j^n(s)^2-\frac{1}{2}]DB_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)) \\ &\qquad B_j(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)), X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)-X((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\rangle ds, \\ Z_n(t,4) :=& -\frac{1}{2}\int_0^t \langle\widetilde{tr}_n(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)), X^n(s)-X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\rangle ds, \\ Z_n(t,5) :=& \frac{1}{2}\int_0^t \langle\widetilde{tr}_n(X^n((\lfloor\frac{s}{\delta}\rfloor-1)\delta)), X(s)-X((\lfloor\frac{s}{\delta}\rfloor-1)\delta)\rangle ds, \end{aligned}$$

$$Z_n(t,6) := -\frac{1}{2}\int_0^t \langle \widetilde{tr}_n(X^n(s)) - \widetilde{tr}_n(X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)), X^n(s) - X(s)\rangle ds.$$

First, the boundedness of DB_j, B_j, D^2B_j yields that there exists $C(N) > 0$

$$\sup_{\|v\|_H, \|u\|_H \le N} \|\widetilde{tr}_n(u) - \widetilde{tr}_n(v)\|_H \le C(N)n\|u-v\|_H, \quad u, v \in H,$$

which combined with Remark 2.9 and (2.16), implies that up to a constant $C(N,T)$

(i) $\mathbb{E}(\sup_{t\in[0,\tau_n]}|Z_n(t,2)|)$ is dominated by $n^2 2^{-3n/8}$;

(ii) $\mathbb{E}(\sup_{t\in[0,\tau_n]}|Z_n(t,j)|), 4 \le j \le 6$ are all dominated by $n2^{-3n/8}$.

Next, for $Z_n(t,1)$, by (2.15), (2.16) and (P1)

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|Z_n(t,1)|\Big) \le C(N)n^2 2^n \mathbb{E}\Big(\int_0^{\tau_n} ds \int_{\lfloor \frac{s}{\delta}\rfloor \delta}^s \|X^n(u) - X^n((\lfloor \frac{s}{\delta} \rfloor - 1)\delta)\|_H du\Big),$$

where $\lfloor \frac{u}{\delta} \rfloor = \lfloor \frac{s}{\delta} \rfloor$ for any $u \in [\lfloor \frac{s}{\delta} \rfloor \delta, s)$. Similarly as in (2.34), by Fubini's theorem the term on the right-hand side of the above inequality is further dominated by

$$C(N)Tn^2\mathbb{E}(\int_0^{\tau_n} \|X^n(u) - X^n((\lfloor \frac{u}{\delta} \rfloor - 1)\delta)\|_H du).$$

Then using Hölder's inequality and Remark 2.9 we obtain that

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|Z_n(t,1)|\Big) \le C(N,T,\|f\|_{L^{p/2}})n^2 2^{-3n/8}. \tag{2.43}$$

Now we only have to consider $Z_n(t,3)$, in which the correction term appears. We rewrite it as

$$\begin{aligned}
Z_n(t,3) =& \sum_{k=1}^{\lfloor \frac{t}{\delta} \rfloor} \sum_{j=1}^n \int_{k\delta}^{(k+1)\delta \wedge t} \langle [(s-k\delta)\dot{\beta}_j^n(k\delta)^2 - \frac{1}{2}] \\
& DB_j(X^n((k-1)\delta))B_j(X^n((k-1)\delta)), X^n((k-1)\delta) - X((k-1)\delta)\rangle ds \\
=:& \frac{\delta}{2} \sum_{k=1}^{\lfloor \frac{t}{\delta} \rfloor - 1} \sum_{j=1}^n \big[(\sqrt{\delta}\dot{\beta}_j^n(k\delta))^2 - 1\big] g_{j,k-1}^1 \\
& + \frac{t - \lfloor \frac{t}{\delta} \rfloor \delta}{2} \sum_{j=1}^n \big[\frac{t - \lfloor \frac{t}{\delta} \rfloor \delta}{\delta}(\sqrt{\delta}\dot{\beta}_j^n(\lfloor \frac{t}{\delta} \rfloor \delta))^2 - 1\big] h_j(t) \\
=:& I_n(t,1) + I_n(t,2),
\end{aligned}$$

where the coefficients $g_{j,k-1}^1, h_j(t)$ are given by

$$g_{j,k-1}^1 := \langle DB_j(X^n((k-1)\delta))B_j(X^n((k-1)\delta)), X^n((k-1)\delta) - X((k-1)\delta)\rangle,$$

$$\begin{aligned}
h_j(t) :=& \langle DB_j(X^n((\lfloor \frac{t}{\delta} \rfloor - 1)\delta))B_j(X^n((\lfloor \frac{t}{\delta} \rfloor - 1)\delta)), \\
& X^n((\lfloor \frac{t}{\delta} \rfloor - 1)\delta) - X((\lfloor \frac{t}{\delta} \rfloor - 1)\delta)\rangle.
\end{aligned}$$

We see that for $j=1,\dots,n,k=1,\dots,\lfloor\frac{\tau_n}{\delta}\rfloor$, the coefficient $g^1_{j,k}$ is $\mathcal{F}_{k\delta}$-measurable and is bounded by some constant $C(N)$. The independence of $\dot\beta^n_j(k\delta)$ and $\mathcal{F}_{(k-1)\delta}$ yields that the r.v.s $\big(\sqrt{\delta}\dot\beta^n_j(k\delta)\big)_{k,j}$ are real-valued i.i.d standard Gaussian. Then for $l=1,\dots,\lfloor\frac{\tau_n}{\delta}\rfloor$, the process defined by

$$\sum_{k=1}^{l}\sum_{j=1}^{n}\big((\sqrt{\delta}\dot\beta^n_j(k\delta))^2-1\big)g^1_{j,k-1} \tag{2.44}$$

is an $\mathcal{F}_{l\delta}$-martingale. Furthermore, for $j_1,j_2=1,\dots,n$, $k_1\neq k_2$, we have

$$\mathbb{E}\Big(\big[(\sqrt{\delta}\dot\beta^n_{j_1}(k_1\delta))^2-1\big]g^1_{j_1,k_1-1}\big[(\sqrt{\delta}\dot\beta^n_{j_2}(k_2\delta))^2-1\big]g^1_{j_2,k_2-1}\Big)=0. \tag{2.45}$$

As for $J_n(t,2)$, by Doob's inequality, $\mathbb{E}\big(\sup_{t\in[0,\tau_n]}|I_n(t,1)|\big)$ is dominated by

$$\begin{aligned}
&\delta\Big(\mathbb{E}\big[\sum_{k=1}^{\lfloor\frac{\tau_n}{\delta}\rfloor-1}\sum_{j=1}^{n}((\sqrt{\delta}\dot\beta^n_j(k\delta))^2-1)g^1_{j,k-1}\big]^2\Big)^{1/2}\\
=&\delta\Big(\mathbb{E}\sum_{k=1}^{\lfloor\frac{\tau_n}{\delta}\rfloor-1}\big[\sum_{j=1}^{n}((\sqrt{\delta}\dot\beta^n_j(k\delta))^2-1)g^1_{j,k-1}\big]^2\Big)^{1/2}\\
\le&C(N)T2^{-n}\big(n^22^n\mathbb{E}(Z^2-1)^2\big)^{1/2}\le C(N)Tn2^{-n/2},
\end{aligned} \tag{2.46}$$

where we used (2.45) in the equality and Z is a standard Gaussian r.v., i.e.

$$\mathbb{E}Z=0,\quad \mathbb{E}Z^2=1,\quad \mathbb{E}Z^4=3.$$

For any $j=1,\dots,n,\ k=1,\dots,\lfloor\frac{\tau_n}{\delta}\rfloor,\ t\in[k\delta,(k+1)\delta)$, we have $h_j(t)=g^1_{j,k-1}$. Then $\mathbb{E}\big(\sup_{t\in[0,\tau_n]}|I_n(t,2)|\big)$ is dominated by

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{1\le k\le\lfloor\frac{\tau_n}{\delta}\rfloor}\sup_{k\delta\le t<(k+1)\delta\wedge\tau_n}\Big|\frac{t-k\delta}{2}\sum_{j=1}^{n}\big[\frac{t-k\delta}{\delta}(\sqrt{\delta}\dot\beta^n_j(k\delta))^2-1\big]h_j(t)\Big|\Big)\\
\le&C(N)\delta\Big(\sum_{k=1}^{2^n}\mathbb{E}\big(\sup_{k\delta\le t\le(k+1)\delta}\big[\sum_{j=1}^{n}\big|\frac{t-k\delta}{\delta}(\sqrt{\delta}\dot\beta^n_j(k\delta))^2-1\big|\big]^2\big)\Big)^{1/2}\\
\le&C(N)\delta\Big(n^22^n\mathbb{E}\big(Z^4+2Z^2+1\big)\Big)^{1/2}\le C(N,T)n2^{-n/2}.
\end{aligned} \tag{2.47}$$

Combining (2.46) and (2.47) we have

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}|Z_n(t,3)|\Big)\le C(N)Tn2^{-n/2}. \tag{2.48}$$

Putting all these estimates of $J_n(t,1)$-$J_n(t,6)$ together, we complete the proof. □

Below we complete the proof of Lemma 2.8, which is similar as in [5, Prop. 5.1].

Proof of Lemma 2.8. Itô's formula ([17, Theorem 4.2.5]) implies that

$$\begin{aligned}
&\|X(s)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\|_H^2\\
=&2\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s} {}_{V^*}\langle A(u,X(u)),X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\rangle_V du+\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\|B(X(u))\|_{L_2}^2du\\
&+2\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\langle X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta),B(X(u))dW(u)\rangle.
\end{aligned}\tag{2.49}$$

We take integral w.r.t s for three terms on the right-hand side of (2.49) and consider their expectations separately. By Hölder's inequality, (H4) and (2.8)

$$\begin{aligned}
&\mathbb{E}\Big(\int_0^{\tau_n}|\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s} {}_{V^*}\langle A(u,X(u)),X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\rangle_V du|ds\Big)\\
\le&\Big(\mathbb{E}\int_0^{\tau_n}\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\|X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\|_V^\alpha duds\Big)^{\frac{1}{\alpha}}\\
&\cdot\Big(\mathbb{E}\int_0^{\tau_n}\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\big(f(u)+K\|X(u)\|_V^\alpha\big)duds\Big)^{\frac{\alpha-1}{\alpha}}\\
\le&\Big(\delta\mathbb{E}\int_0^T\|X(s)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\|_V^\alpha ds\Big)^{\frac{1}{\alpha}}\cdot\Big(\delta\mathbb{E}\int_0^T\big(f(s)+\|X(s)\|_V^\alpha\big)ds\Big)^{\frac{\alpha-1}{\alpha}}\\
\le&C(N,T,\|f\|_{L^{p/2}})2^{-n},
\end{aligned}\tag{2.50}$$

where we used stochastic Fubini's theorem in the second inequality. By (H4) and (2.15) we easily have

$$\mathbb{E}\big(\int_0^{\tau_n}\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\|B(X(u))\|_{L_2}^2duds\big)\le C(N,T)2^{-n}.\tag{2.51}$$

The Burkholder-Davies-Gundy inequality implies

$$\begin{aligned}
&\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}\int_0^t\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\langle X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta),B(X(u))dW(u)\rangle ds\Big)\\
\le&C\mathbb{E}\Big(\int_0^{\tau_n}ds\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\|B(X(u))\|_{L_2}^2\|X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\|_H^2du\Big)^{1/2},
\end{aligned}\tag{2.52}$$

which by (H4) and (2.15) is dominated by $C(T,N)2^{-n/2}$. Then combining (2.50)-(2.52) with (2.49) we have

$$\mathbb{E}\Big(\int_0^{\tau_n}\|X(s)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\|_H^2ds\Big)\le C(N,T,\|f\|_{L^{p/2}})2^{-n/2}.\tag{2.53}$$

However, (2.53) is not enough for our use. Below we improve it by estimating the third term on the right-hand side of (2.49) with the help of (2.53). Using stochastic Fubini's theorem and (2.53)

$$\begin{aligned}
&\mathbb{E}\Big(\int_0^{\tau_n}ds\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^{s}\|X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta)\|_H^2du\Big)\\
\le&\delta\mathbb{E}\Big(\int_0^{\tau_n}\|X(u)-X(\lfloor\frac{u}{\delta}\rfloor\delta)\|_H^2du\Big)\le C(N,T,\|f\|_{L^{p/2}})2^{-3n/2}.
\end{aligned}$$

Hence inserting into (2.52) and by (H4) we obtain

$$\mathbb{E}\Big(\sup_{t\in[0,\tau_n]}\int_0^t\int_{\lfloor\frac{s}{\delta}\rfloor\delta}^s\langle X(u)-X(\lfloor\frac{s}{\delta}\rfloor\delta),B(X(u))dW(u)\rangle ds\Big)\leq C(N,T,\|f\|_{L^{p/2}})2^{-\frac{3n}{4}}.$$

Then together with (2.50), (2.51), we complete the proof. □

3 Application to Examples

Let $\mathbb{D}\subset\mathbb{R}^m$ be a bounded open domain with smooth boundary denoted by $\partial\mathbb{D}$. $C_0^\infty(\mathbb{D};\mathbb{R}^n)$ denotes the set of all smooth functions from $\mathbb{D}$ to $\mathbb{R}^n$ with compact support. $W_0^{1,p}(\mathbb{D};\mathbb{R}^n)$ is the standard Sobolev space, i.e. the closure of $C_0^\infty(\mathbb{D};\mathbb{R}^n)$ with respect to the norm:

$$\|u\|_{1,p}=\Big(\int_{\mathbb{D}}|u(x)|^p+|\nabla u(x)|^pdx\Big)^{\frac{1}{p}}.$$

Let $p=2$, since $\mathbb{D}$ is bounded, by Poincare's inequality there exists a constant c such that

$$c\int_{\mathbb{D}}|\nabla u(x)|^2dx\geq\int_{\mathbb{D}}|u(x)|^2dx,\quad u\in W_0^{1,2}(\mathbb{D};\mathbb{R}^n).\tag{3.1}$$

Then we can consider $W_0^{1,2}(\mathbb{D};\mathbb{R}^n)$ with the norm $\|\cdot\|_{W_0^{1,2}(\mathbb{D};\mathbb{R}^n)}$ and the corresponding scalar product given by

$$\langle u,v\rangle_{W_0^{1,2}(\mathbb{D};\mathbb{R}^n)}=\int_{\mathbb{D}}\langle\nabla u(x),\nabla v(x)\rangle dx,\quad u,v\in W_0^{1,2}(\mathbb{D};\mathbb{R}^n).$$

In the following we use the notations $|y|^2:=\sum_{k=1}^m|y^k|^2$, $y\cdot x:=\sum_{k=1}^m y^kx^k$ for $y=(y^k),x=(x^k)\in\mathbb{R}^m$ and $|z|^2:=\sum_{k=1}^m\sum_{j=1}^n|z^{k,j}|^2$ for $z=(z^{k,j})\in\mathbb{R}^{m\times n}$.

3.1 Stochastic 2d hydrodynamical type systems

Let $\mathbb{D}\subset\mathbb{R}^2$ be a bounded open domain with smooth boundary. We consider the Hilbert space $H=L^2(\mathbb{D};\mathbb{R}^2)$ with the inner product $\langle\cdot,\cdot\rangle$ given by $\langle f,g\rangle:=\int_{\mathbb{D}}f(x)\cdot g(x)dx$ for $f,g\in H$. Let A_0 be a self-adjoint positive linear operator on H. Set $V=Dom(A_0^{\frac{1}{2}})$ and $\|\cdot\|_V=\|A_0^{\frac{1}{2}}\cdot\|_H$. Let V^* denote the dual of V with respect to $\langle\cdot,\cdot\rangle$. Thus we have the Gelfand triple $V\subset H\subset V^*$ and we study the following equation (c.f [4, 5])

$$\partial_tu(t)+A_0u(t)+C(u(t),u(t))+Ru(t)=\sum_{k=1}^\infty h_k(u(t))dW^k(t),x\in\mathbb{D}\tag{3.2}$$

with the initial condition $u(0)=u_0\in L^4(\mathbb{D};\mathbb{R}^2)$. R is a continuous operator in H, the map $C:V\times V\to V^*$ satisfies the following conditions:

(Φ1) The map C is bilinear continuous;

(Φ2) For $v_i \in V$, $i = 1, 2, 3$,

$$_{V^*}\langle C(v_1, v_2), v_3\rangle_V = -_{V^*}\langle C(v_1, v_3), v_2\rangle_V;$$

(Φ3) There exists a Banach space $\mathcal{H}$ possessing the properties:

(i) $V \subset \mathcal{H} \subset H$;

(ii) there exists a constant $a_0 > 0$ such that

$$\|v\|^2_{\mathcal{H}} \le a_0\|v\|_H\|v\|_V, \quad v \in V;$$

(iii) for every $\eta > 0$ there exists $C_\eta > 0$ such that

$$|_{V^*}\langle C(v_1, v_2), v_3\rangle_V| \le \eta\|v_3\|^2_V + C_\eta\|v_1\|^2_{\mathcal{H}}\|v_2\|^2_{\mathcal{H}}, \quad v_i \in V, i = 1, 2, 3.$$

$\{W^k(t); t \ge 0, k = 1, 2, \ldots\}$ is a sequence of independent real-valued standard Brownian motions on a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}; \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t\ge0}$ generated by W. $h : \mathbb{D} \times \mathbb{R}^2 \to \mathbb{R}^2 \times l^2$ is measurable, where l^2 denotes the Hilbert space consisting of all sequences of square summable real numbers with standard norm $\|\cdot\|_{l^2}$. For additive noise it is obvious that (P1) and (P2) in Assumption 2 are fulfilled. We can also consider more general case and here for simplicity we check the conditions for linear multiplicative noise, i.e. $h = (h_k)$ is given by

$$h_k(x, u) = g_k(x)u, \quad (x, u) \in \mathbb{D} \times \mathbb{R}^2, k \in \mathbb{N},$$

where $g = (g_k) : \mathbb{D} \to l^2$ is differentiable on $\mathbb{D}$, and there exists a constant M such that for any $x \in \mathbb{D}, i = 1, 2$

$$\|\partial_{x^i} g(x)\|^2_{l^2} + \|g(x)\|^2_{l^2} \le M. \tag{3.3}$$

This model involves qualitative properties of stochastic models, which describe cooperative effects in fluids by taking into account macroscopic parameters such as temperature and magnetic field. The corresponding mathematical models consist in, for example, 2d Navier-Stokes equations and magneto-hydrodynamic equations described below:

(1) stochastic 2d Navier-Stokes equations

$$\begin{cases} \partial_t u = \nu\triangle u - (u\cdot\nabla)u - \nabla p + B(u_t)dW_t, & x \in \mathbb{D}, \\ div u = 0, \quad x \in \mathbb{D}; \quad u = 0 \text{ on } \partial\mathbb{D} \end{cases} \tag{3.4}$$

where $u = (u^1(t, x), u^2(t, x))$ is the velocity of a fluid, $p(t, x)$ is the pressure and ν is the kinematic viscosity; $\triangle u = \nabla \cdot \nabla u$.

(2) stochastic 2d magneto-hydrodynamic equations (see [24])

$$
\begin{cases}
\partial_t u =\dfrac{1}{Re}\triangle u-(u\cdot\nabla)u-\nabla p-\dfrac{M^2}{ReRm}\big(\dfrac{\nabla|b|^2}{2}-(b\cdot\nabla)b\big)+B(u_t)dW_t\\
\partial_t b =\dfrac{1}{Re}\triangle b-(u\cdot\nabla)b+(b\cdot\nabla)u,\quad x\in\mathbb{D},\\
divu =0,\quad divb=0,\quad x\in\mathbb{D},\\
\quad u =0,\quad b\cdot n=0,\quad \partial_1 b^2-\partial_2 b^1=0 \text{ on } \partial\mathbb{D}
\end{cases}
\tag{3.5}
$$

where $u=(u^1(t,x),u^2(t,x))$ and $b=(b^1(t,x),b^2(t,x))$ denote velocity and magnetic fields, $p(t,x)$ is a scalar pressure. n denotes the outward normal to ∂D and Re, Rm, M correspond to the Reynolds number, the magnetic Reynolds number and the Hartman number, respectively.

Equation (3.4): Define

$$
\begin{aligned}
V_1=\{v\in W_0^{1,2}(\mathbb{D};\mathbb{R}^2):\nabla\cdot v=0 \text{ in } \mathbb{D}, v\cdot n=0 \text{ on } \partial\mathbb{D}\},\\
\|v\|_{V_1}:=(\int_{\mathbb{D}}|\nabla v|^2dx)^{1/2},\ v\in V_1.
\end{aligned}
\tag{3.6}
$$

Let H_1 be the closure of V_1 in the norm $\|u\|_{H_1}:=\big(\int_{\mathbb{D}}|u|^2dx\big)^{1/2}$ and endowed with the L^2 scalar product. Set

$$
W^{k,2}(\mathbb{D};\mathbb{R}^2)=\{u\in L^2(\mathbb{D};\mathbb{R}^2):\ D^\alpha u\in L^2(\mathbb{D};\mathbb{R}^2),\ \forall|\alpha|\le k\},\ k\in\mathbb{N}.
$$

The linear operators A_0, P_H (Helmhotz-Hodge projection) and the map C are defined by

$$
\begin{aligned}
P_H:\ L^2(\mathbb{D};\mathbb{R}^2)\to H_1 \text{ orthogonal projection},\\
A_0:\ W^{2,2}(\mathbb{D};\mathbb{R}^2)\cap V_1\to H_1,\quad A_0u=\nu P_H\triangle u,\\
C:\ H_1\times V_1\to H_1,\ C(u,v)=-P_H\big[(u\cdot\nabla)v\big],\ C(u)=C(u,u).
\end{aligned}
\tag{3.7}
$$

Choosing the Gelfand triple

$$
V_1\subset H_1\equiv H_1^*\subset V_1^*,
$$

we know that the maps

$$
A_0:\ V_1\to V_1^*,\ C:V_1\times V_1\to V_1^*
$$

are well defined and satisfy the conditions (Φ1)-(Φ3) with $\mathcal{H}=L^4(\mathbb{D};\mathbb{R}^2)\cap H_1$ (see [4,2.1.1]).

Equation (3.5): Set

$$
\begin{aligned}
V_2=\{v\in W^{1,2}(\mathbb{D};\mathbb{R}^2):\nabla\cdot v=0 \text{ in } \mathbb{D}, v\cdot n=0 \text{ on } \partial\mathbb{D}\},\\
\|v\|_{V_2}:=\big(\int_{\mathbb{D}}|\nabla v|^2dx\big)^{1/2},\ v\in V_2.
\end{aligned}
\tag{3.8}
$$

First, without loss of generality we can assume that $\frac{M^2}{ReRm}=1$; Indeed, if $\frac{M^2}{ReRm}\neq 1$, we consider $\sqrt{\frac{M^2}{ReRm}}b$ instead. Then we can write (3.5) as (3.2) by choosing $H=H_1\times H_1$ with $\tilde{A}_0=A_0\times A_0$, $R\equiv 0$. We also set $V=V_1\times V_2$ and define $\tilde{C}:V\times V\to V^*$ by

$$\langle \tilde{C}(z_1,z_2),z_3\rangle=\langle C(u_1,u_2),u_3\rangle-\langle C(b_1,b_2),u_3\rangle+\langle C(u_1,b_2),u_3\rangle-\langle C(b_1,u_2),b_3\rangle$$

for $z_i=(u_i,b_i)\in V$. Here the spaces H_1,V_1,V_2 and operators C,A_0 have been defined in (3.6), (3.7) and (3.8). Then the conditions (Φ1)-(Φ3) hold with $\mathcal{H}=[L^4(\mathbb{D};\mathbb{R}^2)\times L^4(\mathbb{D};\mathbb{R}^2)]\cap H$ (see [4, 2.1.2]).

Verifying (H1)-(H4),(P1)-(P2): In (3.2), we consider the Gelfand triple $V:=W_0^{1,2}(\mathbb{D};\mathbb{R}^2)\subset H=L^2(\mathbb{D};\mathbb{R}^2)\simeq H^*\subset (W_0^{1,2}(\mathbb{D};\mathbb{R}^2))^*$ and define the coefficients A and B below: for $v\in V,u\in H$, $A(v):=-A_0v-C(v,v)-Rv\in V^*$ and $B(u):=h(\cdot,u(\cdot))\in L_2(l^2;H)$; set $C(v):=C(v,v)$, $v\in V$. We have $B_k(u):=g_ku\in H$ for $u\in H$, $k\in\mathbb{N}$.

Estimates of A and B:

(H1): The continuity of C and R implies that (H1) holds.

(H2)+(H3): For $v_1,v_2,v\in V$ we have

$$_{V^*}\langle -A_0v,v\rangle_V=-\|v\|_V^2,\quad -{}_{V^*}\langle A_0v_1-A_0v_2,v_1-v_2\rangle_V=-\|v_1-v_2\|_V^2.$$

(Φ2,ii,iii) imply that (see [5, (2.8)]) there exists $a_1>0$ such that

$$_{V^*}\langle C(v),v\rangle_V=0,$$

$$|_{V^*}\langle C(v_1)-C(v_2),v_1-v_2\rangle_V|\leq\|v_1-v_2\|_V^2+a_1\|v_1-v_2\|_H^2\|v_2\|_{\mathcal{H}}^4.$$

Since R is continuous and linear in H, by (2.1) there exists $a_2>0$ such that

$$|_{V^*}\langle Rv,v\rangle_V|=|\langle Rv,v\rangle|\leq a_2\|v\|_H^2,\ \ _{V^*}\langle Rv_1-Rv_2,v_1-v_2\rangle_V\leq a_2\|v_1-v_2\|_H^2.$$

Then by (Φ3,ii) we deduce that

$$_{V^*}\langle A(v),v\rangle_V\leq-\|v\|_V^2+a_2\|v\|_H^2,$$

$$_{V^*}\langle A(v_1)-A(v_2),v_1-v_2\rangle_V\leq(a_2+a_1a_0^2\|v_2\|_V^2\|v_2\|_H^2)\|v_1-v_2\|_H^2.$$

By (3.3) we have for any $u_1,u_2\in H$

$$\|B(u_1)-B(u_2)\|_{L_2(l^2;H)}^2=\int_{\mathbb{D}}\|g(x)\|_{l^2}^2|u_1(x)-u_2(x)|^2dx\leq M\|u_1-u_2\|_H^2.$$

Hence (H2) and (H3) hold with $\rho(\cdot)=a_1a_0^2\|\cdot\|_V^2\|\cdot\|_H^2$ on V, $\rho'\equiv M$ on H and $\alpha=2,\ \beta=2$.

(H4): For any $v_1,v_2\in V$, we have

$$|_{V^*}\langle A_0v_1,v_2\rangle_V|\leq\|v_1\|_V\|v_2\|_V,\ |_{V^*}\langle Rv_1,v_2\rangle_V|\leq a_2\|v_1\|_H\|v_2\|_H,$$

and by condition (Φ3,iii) and [5, (2.6)] there exists $a_3 > 0$ such that

$$|_{V^*}\langle C(v_1), v_2\rangle_V| \leq a_3\|v_1\|_{\mathcal{H}}^2\|v_2\|_{\mathcal{H}} \leq a_3\|v_1\|_V\|v_1\|_H\|v_2\|_V.$$

Considering that $\|\cdot\|_H \leq c\|\cdot\|_V$, then there exists $a > 0$ such that

$$\|A(v_1)\|_{V^*}^2 \leq a(1 + \|v_1\|_V^2)(1 + \|v_1\|_H^2).$$

Using (3.3) we have for any $u \in H$

$$\|B(u)\|_{L_2(l^2;H)}^2 = \int_{\mathbb{D}} \|g(x)u(x)\|_{l^2}^2 dx \leq M\|u\|_H^2.$$

Hence (H4) holds.
We continue to check (P1)-(P2). For $k \in \mathbb{N}$, $u_1, u_2 \in H$, $x \in \mathbb{D}$,

$$h_k(x, u_1(x) + u_2(x)) - h_k(x, u_1(x)) = g_k(x)u_2(x).$$

Then DB_k on H is given by

$$DB_k(u_1)u_2 = g_k u_2 \in H,\ \ u_1, u_2 \in H.$$

We also obtain the dual operator of DB_k given by

$$DB_k(u_1)^* u_2 = g_k u_2,\ \ u_1, u_2 \in H.$$

Similarly, the operator $D^2B_k = 0$ on H.
(P1): For any $N > 0, k \in \mathbb{N}, u \in H, v \in V$

$$\|B(u) - B(u)\Pi_n\|_{L_2(l^2;H)}^2 = \sum_{k=n+1}^{\infty} \int_{\mathbb{D}} |g_k(x)u(x)|^2 dx, \tag{3.9}$$

which according to (3.3), uniformly converges to 0 for all $u \in H$ with $\|u\|_H \leq N$.
Again by (3.3) and (3.1)

$$\begin{aligned} \|DB_k(u)^* v\|_V^2 &= \sum_{j=1,2} \int_{\mathbb{D}} |g_k(x)\partial_{x^j} v(x) + \partial_{x^j} g_k(x)v(x)|^2 dx \\ &\leq 2M \int_{\mathbb{D}} (\sum_{j=1,2} |\partial_{x^j} v(x)|^2 + |v(x)|^2) dx \leq 2M(c+1)\|v\|_V^2. \end{aligned} \tag{3.10}$$

(3.3), (3.9) and (3.10) imply that (P1) holds.
(P2): By (3.3) we have for every $u, u_1, u_2 \in H, n \in \mathbb{N}$

$$\|\widetilde{tr}_n(u)\|_H = \|\sum_{k=1}^{n} g_k^2 u\|_H \leq M\|u\|_H,$$

$$\langle \widetilde{tr}_n(u_1) - \widetilde{tr}_n(u_2), u_1 - u_2\rangle = \sum_{k=1}^{n} \|g_k^2[u_1 - u_2]\|_H^2 \leq M\|u_1 - u_2\|_H^2.$$

Hence (P2) holds with $\rho' \equiv M$ on H.

3.2 Stochastic porous medium equations

Porous medium equation is a model to describe the flow of an ideal gas in a homogeneous porous medium (e.g. beds of sand, ground). Forgetting about physical constants, it is given in one dimensional case by (c.f. [1, 21])

$$dX(t) = \triangle(|X(t)|^{m-2}X(t))dt + B(X(t))dW(t),\ x \in [0,1],$$

with $m \geq 2$, the initial condition $X(0,x) = X_0(x),\ x \in [0,1]$ and Dirichlet boundary condition $X(t,0) = X(t,1) = 0$.
$\{W^k(t), t \geq 0, k \geq 1\}$ is a sequence of independent real-valued standard Brownian motions on a complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}; \mathbb{P})$ with normal filtration $(\mathcal{F}_t)_{t\geq 0}$ generated by W. The map $B = (B_k) : H \to L_2(l^2; H)$ is $\mathcal{B}(H)/\mathcal{B}(L_2(l^2; H))$-measurable and satisfies Assumptions 1,2 in Section 2.
A more general form is the following quasi-linear stochastic equation

$$dX(t) = \triangle\Psi(X(t))dt + \sum_{k=1}^{\infty} B_k(X(t))dW^k(t),\ x \in \mathcal{O},$$

where $\mathcal{O} \subset \mathbb{R}$ is a bounded and open domain. $\Psi : \mathbb{R} \to \mathbb{R}$ (c.f. [21, 4.1.11]) is a function satisfying:

(Ψ1) Ψ is continuous;

(Ψ2) For all $s, t \in \mathbb{R}$

$$(t-s)(\Psi(t) - \Psi(s)) \geq 0;$$

(Ψ3) There exist $q \geq 2,\ a > 0,\ c \geq 0$ such that for all $s \in \mathbb{R}$,

$$s\Psi(s) \geq a|s|^q - c;$$

(Ψ4) There exist $c_3, c_4 > 0$ such that for all $s \in \mathbb{R}$,

$$|\Psi(s)| \leq c_4 + c_3|s|^{q-1},$$

where q is as in (Ψ3).

Let q be given in (Ψ3), we take $H := (W_0^{1,2}(\mathcal{O}))^*$ and identify H with its dual H^* and consider the Gelfand triple:

$$V := L^q(\mathcal{O}) \subset (W_0^{1,2}(\mathcal{O}))^* = H \simeq H^* \subset V^* = (L^q(\mathcal{O}))^*,$$

and define the porous medium operator $A : L^q(\mathcal{O}) \to V^*$

$$A(v) := \Lambda\Psi(v), \quad v \in V, \tag{3.11}$$

where by [21, Lemma 4.1.13], the Laplacian operator $\triangle$ defined on $W_0^{1,2}(\mathcal{O})$ extends to a linear isometry $\Lambda : L^{\frac{q}{q-1}}(\mathcal{O}) \to (L^q(\mathcal{O}))^* = V^*$ satisfying that for all $u \in L^{\frac{q}{q-1}}(\mathcal{O})$, $v \in L^q(\mathcal{O})$

$$ {}_{V^*}\langle -\Lambda u, v\rangle_V = {}_{L^{\frac{q}{q-1}}}\langle u, v\rangle_{L^q} = \int_{\mathcal{O}} u(x)v(x)dx. \tag{3.12}$$

(Ψ4) implies that $\Psi(v) \in L^{\frac{q}{q-1}}(\mathcal{O})$ for any $v \in L^q(\mathcal{O})$. Hence A is well-defined. More details about the above Gelfand triple can be seen in [21, Remark 4.1.14]. The conditions (H1)-(H4) for the coefficient A are satisfied with the related constants $\alpha = q,\ K = 0,\ \theta = a,\ f(t) = 2c \cdot vol(\mathcal{O})$ where $vol(\mathcal{O})$ denotes the volume of $\mathcal{O}$; see [21, Remark 4.1.14].

Remark 3.1. A typical example satisfying (Ψ1)-(Ψ4) is to set $\Psi(s) = s|s|^{q-2}$ for $q \geq 2$. We can also use the framework to other quasi-linear case, e.g. p-Laplace evolution equation. The equation becomes

$$dX(t) = div(|\nabla X(t)|^{p-2}\nabla X(t))dt + B(X(t))dW(t).$$

Again we take $p \in [2, \infty)$, $\mathbb{D} \subset \mathbb{R}^m$ open and bounded with smooth boundary. Then we take $V := W_0^{1,p}(\mathbb{D}; \mathbb{R}^n),\ H := L^2(\mathbb{D})$, so $V^* = (W_0^{1,p}(\mathbb{D}); \mathbb{R}^n)^*$. Define $A : V \to V^*$ by

$$A(u) := div(|\nabla u|^{p-2}\nabla u), \quad u \in V,$$

$${}_{V^*}\langle A(u), v\rangle_V := -\int_{\mathbb{D}} |\nabla u(x)|^{p-2}\langle \nabla u(x), \nabla v(x)\rangle dx, \quad v, u \in V.$$

A is called the p-Laplacian and $A = \triangle$ when $p = 2$. For any $u, v \in V$, using Hölder's inequality

$$\int_{\mathbb{D}} |\nabla u(x)|^{p-1}|\nabla v(x)|dx \leq \Big(\int_{\mathbb{D}} |\nabla u(x)|^p dx\Big)^{\frac{p-1}{p}} \Big(\int_{\mathbb{D}} |\nabla v(x)|^p dx\Big)^{\frac{1}{p}} \leq \|u\|_{1,p}^{p-1}\|v\|_{1,p},$$

which implies that the p-Laplacian operator A is well-defined. Under the Gelfand Triple $W_0^{1,p}(\mathbb{D}; \mathbb{R}^n) \subset L^2(\mathbb{D}) = H \simeq H^* \subset V^* = (W_0^{1,p}(\mathbb{D}; \mathbb{R}^n))^*$, A satisfies conditions (H1)-(H4) (see [21, Remark 4.1.9]).

4 Support Problem

In this section we describe the support of solutions with the help of Wong-Zakai approximation results. Let $T > 0$ and let W be a cylindrical Wiener process in U on some complete probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}; \mathbb{P})$, with $(\mathcal{F}_t)_{t\geq 0}$ being normal filtration generated by W. For $n \in \mathbb{N}$, $t \in [0, T]$, set

$$M^n(t) := \exp\Big(\int_0^t \dot{W}^n(s)dW(s) - \frac{1}{2}\int_0^t \|\dot{W}^n(s)\|_U^2 ds\Big)$$

with $\dot W^n$ given in (2.4), and

$$\widetilde{W}^n(t) := W(t) - \int_0^t \dot W^n(s)ds. \tag{4.1}$$

Since the real-valued r.v.s $\dot\beta_j(k\delta), j,k\in\mathbb{N}$ are independent and for each $j,k\in\mathbb{N}$, $\delta^{1/2}\dot\beta_j(k\delta)$ is standard Gaussian. So for every $n\in\mathbb{N}$

$$\sup_{t\in[0,T]}\mathbb{E}\big(e^{\lambda\|\dot W^n(t)\|_U^2}\big) = \sup_{t\in[0,T]}\prod_{1\le j\le n}\mathbb{E}\big(e^{\lambda|\dot\beta_j^n(t)|^2}\big) = \big(\mathbb{E}(e^{\lambda\frac{1}{\delta}|Z|^2})\big)^n < \infty$$

holds for some standard Gaussian random variable Z and $\lambda>0$ small enough. Thus by Girsanov's theorem ([6, Theorem 10.14 and Proposition 10.17]), the process $\{\widetilde{W}^n(t)\}_{t\in[0,T]}$ defined by (4.1) is a cylindrical Wiener process under $\mathbb{P}^n$ with the measure $\mathbb{P}^n\ll\mathbb{P}$ satisfying

$$\frac{d\mathbb{P}^n}{d\mathbb{P}}\Big|\mathcal{F}_t = M^n(t)\ ,\ \text{for } t\in[0,T].$$

Similarly, for arbitrary $h\in L^2([0,T];U)$, we define the process

$$M_h^n(t) := \exp\Big(-\int_0^t h(s)d\widetilde{W}^n(s) - \frac{1}{2}\int_0^t \|h(s)\|_U^2 ds\Big),$$

and

$$\widetilde{W}_h^n(t) := \widetilde{W}^n(t) + \int_0^t h(s)ds\ ,\ \text{for } t\in[0,T], n\in\mathbb{N}. \tag{4.2}$$

Again by Girsanov's theorem, we obtain another measure $\mathbb{P}_h^n\ll\mathbb{P}^n\ll\mathbb{P}$ such that

$$\frac{d\mathbb{P}_h^n}{d\mathbb{P}^n}\Big|\mathcal{F}_t = M_h^n(t)\ ,\ \text{for } t\in[0,T], \tag{4.3}$$

and $\widetilde{W}_h^n$ is a cylindrical Wiener process under $\mathbb{P}_h^n$. Consider the following equations

$$\begin{aligned} dY_h^n(t) =& A(t,Y_h^n(t))dt + B_1(Y_h^n(t))dW(t) + B_2(Y_h^n(t))\dot W^n(t)dt \\ &+ B_3(Y_h^n(t))h(t)dt - F(Y_h^n(t))dt, \end{aligned} \tag{4.4}$$

where for any fixed time $T>0$, the maps

$$A:[0,T]\times V\times\Omega\to V^*, F: H\times\Omega\to H,$$

$$B_1,B_2,B_3: H\times\Omega\to (L_2(U;H),\|\cdot\|_{L_2(U;H)})$$

are progressively measurable.

We note that (2.6) can be seen as a special case of (4.4) with $B_1=0,\ B_2=B,\ B_3=0$ and $F=\frac{1}{2}\tilde{tr}_n$. Then we obtain Theorem 2.4 by using Lemma 4.1. We can also write (2.6) as (2.2) with the drift coefficient $\tilde A = A + B\dot W^n - \frac{1}{2}\widetilde{tr}_n$ and the diffusion coefficient $\tilde B = 0$. However, we cannot use Theorem 2.3 directly to solve equations (2.6). By Assumptions 1 and 2 we can deduce that (H1), (H2) and (H4) hold for $\tilde{A}$. However, (H3) fails to hold, since by Assumption 1 we can only obtain that

$$\langle B(u)\dot{W}^n(t), u\rangle \leq \sqrt{K}\sqrt{1+\|u\|_H^2}\|u\|_H\|\dot{W}^n(t)\|_U,$$

where we cannot find a uniform bound of $\|\dot{W}^n(t)\|_U$ for all $\omega \in \Omega$ and $t \in [0,T]$.

Below we give existence and uniqueness of solutions to equations (4.4). The argument is similar as in the proof of [16, Theorem 1.1] and we put the proof in Appendix.

Lemma 4.1. *Let $T > 0, h \in L^2([0,T];U)$ and $\xi \in L^p(\Omega,\mathcal{F}_0,\mathbb{P};H)$ with $p > \{\frac{\alpha}{\alpha-1} \vee (\beta+2)\}$ in Assumption 1. Assume that the coefficients A satisfies Assumption 1, B_1, B_2, B_3 satisfy the conditions in Assumption 1 for B and F satisfies (P2). Then there exist unique solutions Y_h^n to equations (4.4) with initial value ξ. Moreover, $Y_h^n \in \mathcal{C}([0,T];H)$ $\mathbb{P}$-a.e. and*

$$\sup_{n\geq 1}\mathbb{E}\Big(\sup_{t\in[0,T]}\|Y_h^n(t)\|_H^p + \int_0^T \|Y_h^n(t)\|_V^\alpha dt\Big) < \infty. \tag{4.5}$$

Remark 4.2. For $h \in L^2([0,T];U)$, we consider the following two equations, which can be seen as two special cases of (4.4)

$$dZ_h(t) = A(t,Z_h(t))dt + B(Z_h(t))h(t)dt - \frac{1}{2}\widetilde{tr}_n(Z_h(t))dt, \tag{4.6}$$

$$\begin{aligned} dZ_h^n(t) =& A(t,Z_h^n(t))dt + B(Z_h^n(t))h(t)dt \\ &+ B(Z_h^n(t))dW(t) - B(Z_h^n(t))\dot{W}^n(t)dt, \end{aligned} \tag{4.7}$$

with $\widetilde{tr}_n$ given in (2.5). By Lemma 4.1 there exist unique solutions Z_h and Z_h^n to equations (4.6) and (4.7), respectively; $Z_h, Z_h^n \in \mathcal{C}([0,T];H)$, $\mathbb{P}$-a.e. and

$$\mathbb{E}\Big(\sup_{t\in[0,T]}\|Z_h(t)\|_H^p + \int_0^T \|Z_h(t)\|_V^\alpha dt\Big) < \infty,$$

$$\sup_{n\geq 1}\mathbb{E}\Big(\sup_{t\in[0,T]}\|Z_h^n(t)\|_H^p + \int_0^T \|Z_h^n(t)\|_V^\alpha dt\Big) < \infty.$$

By a similar computation as in Theorem 2.6 we obtain the following Wong-Zakai approximation results.

Lemma 4.3. *Suppose that $\xi \in L^p(\Omega,\mathcal{F}_0,\mathbb{P};H)$ with $p > \{\frac{\alpha}{\alpha-1} \vee (\beta+2)\}$ in Assumption 1, and that Assumptions 1 and 2 hold. For $h \in L^2([0,T];U)$, let Z_h, Z_h^n denote the solutions to equations* (4.6) *and* (4.7)*, respectively. Then*

$$\lim_{n\to\infty}\mathbb{E}\Big(\sup_{t\in[0,T]}\|Z_h^n(t) - Z_h(t)\|_H^2\Big) = 0. \tag{4.8}$$

Proof. Let $h \in L^2([0,T];U)$, for arbitrary $n \in \mathbb{N}$, we set the stopping times:

$$\sigma_N^{(1)} := \inf\Big\{t \in [0,T] : \|Z_h(t)\|_H + \int_0^t \big(f(s) + \|Z_h(s)\|_V^\alpha\big)ds > N\Big\} \wedge T,$$
$$\sigma_{n,N}^{(2)} := \inf\Big\{t \in [0,T] : \|Z_h^n(t)\|_H + \int_0^t \|Z_h^n(s)\|_V^\alpha ds > N\Big\} \wedge T,$$

and

$$\sigma_{n,N} := \sigma_N^{(1)} \wedge \sigma_{n,N}^{(2)} \wedge \tau_n^{(3)},$$

with $\tau_n^{(3)}$ in (2.11) and by Remark 4.2 we similarly choose $N > 0$ large enough as in the proof of Theorem 2.6 and denote $\sigma_{n,N}$ by σ_n for simplicity. Thus in order to obtain (4.8), it is sufficient to prove the following

$$\lim_{n\to\infty} \mathbb{E}\big(\sup_{t\in[0,\sigma_n]} \|Z_h(t) - Z_h^n(t)\|_H^2\big) = 0. \tag{4.9}$$

Using Itô's formula for $\|Z_h^n(t)-Z_h(t)\|_H^2$, and by comparison with (2.17), it suffices to control the following term: by (H2) and Young's inequality

$$\begin{aligned}&\int_0^t \langle [B(Z_h^n(s)) - B(Z_h(s))]h(s), Z_h^n(s) - Z_h(s)\rangle ds\\ \le &\int_0^t (\rho'(Z_h(s)) + \|h(s)\|_U^2)\|Z_h^n(s) - Z_h(s)\|_H^2 ds.\end{aligned}$$

Together with estimates of terms on the right-hand side of (2.17), there exist $C_n \to 0$ as $n \to \infty$ such that for all $t \in [0,T]$, $\mathbb{E}\big(\sup_{s\in[0,\sigma_n\wedge t]} \|Z_h^n(s) - Z_h(s)\|_H^2\big)$ is dominated by

$$C_n + \mathbb{E}\Big(\int_0^{\sigma_n\wedge t} C\big(f(s) + \rho(Z_h(s)) + \rho'(Z_h(s))\big) + \|h(s)\|_U^2)\|Z_h^n(s) - Z_h(s)\|_H^2 ds\Big).$$

By similar arguments as the estimate for (2.20), and using [7, Lemma 2.2], (4.5) and (H4) we complete the proof. □

Let $D = \mathcal{C}([0,T];H)$ denote the space of continuous functions, with the norm $\|\cdot\|_D = \sup_{t\in[0,T]} \|\cdot(t)\|_H$. We set $\mathcal{L} := \{Z_h, h \in L^2([0,T];U)\}$ and see that $\mathcal{L} \subset D$. Now we describe the support theorem.

Theorem 4.4. *Suppose that $\xi \in L^p(\Omega, \mathcal{F}_0, \mathbb{P}; H)$ with $p > \{\frac{\alpha}{\alpha-1} \vee (\beta+2)\}$ in Assumption 1, and that Assumptions 1 and 2 hold. Let X denote the solution to equation (2.2) with initial condition ξ. Then supp $(\mathbb{P}\circ X^{-1}) = \bar{\mathcal{L}}$, where $\bar{\mathcal{L}}$ denotes the closure of $\mathcal{L}$ in D and supp $(\mathbb{P}\circ X^{-1})$ denotes the support of the distribution $\mathbb{P}\circ X^{-1}$.*

Proof. Let $\{W(t)\}_{t\ge0}$ be a cylindrical Wiener process in $(U, \langle\cdot,\cdot\rangle_U)$ on a complete filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\ge0}; \mathbb{P})$ with $(\mathcal{F}_t)_{t\ge0}$ being normal filtration generated by W, and let X and X^n be the solutions to equations (2.2) and (2.6), respectively. Choose $h = \dot{W}^n$ in equations (4.6), existence and uniqueness of solutions in this case can also be obtained by Lemma 4.1. We denote the solutions to (4.6) by $Z_{\dot{W}^n}$ and have $Z_{\dot{W}^n} = X^n$, $\mathbb{P}$-a.e. Then Theorem 2.6 implies that for every $\lambda > 0$

$$\lim_{n\to\infty} \mathbb{P}\Big(\|Z_{\dot{W}^n} - X\|_D \geq \lambda\Big) = \lim_{n\to\infty} \mathbb{P}\Big(\|X^n - X\|_D \geq \lambda\Big) = 0.$$

Since $\mathbb{P}$-a.e. $\dot{W}^n \in L^2([0,T];U)$,

$$Supp(\mathbb{P}\circ X^{-1}) \subset \bar{\mathcal{L}}. \tag{4.10}$$

Conversely, by [17, Remark 2.5.1] we can always find another Hilbert space $\bar{U} \supset U$ such that there exists a Hilbert-Schmidt embedding from $(U, \langle\cdot,\cdot\rangle_U)$ to $(\bar{U}, \langle\cdot,\cdot\rangle_{\bar{U}})$. It follows that there exist $\{e_k, k\in\mathbb{N}\} \subseteq U, 0 < \lambda_k \uparrow \infty, k \in \mathbb{N}$ such that $\{e_k, k \in \mathbb{N}\}$ is an orthonormal basis in U and $\{\sqrt{\lambda_k} e_k, k \in \mathbb{N}\}$ is an orthonormal basis in $\bar{U}$. Fix such $\bar{U}$, set $\mathbf{W}^{\bar{U}} := \mathcal{C}([0,\infty);\bar{U})$ and $\mathbf{W}_0^{\bar{U}} := \{x \in \mathbf{W}^{\bar{U}} | x(0) = 0\}$. $\mathbf{W}_0^{\bar{U}}$ is equipped with metric

$$\rho(x_1,x_2) := \sum_{k=1}^{\infty} (\max_{0\leq t\leq k} \|x_1(t) - x_2(t)\|_{\bar{U}} \wedge 1),\ x_1, x_2 \in \mathbf{W}_0^{\bar{U}},$$

which makes it a Polish space. Its Borel σ-algebra is denoted by $\mathcal{B}(\mathbf{W}_0^{\bar{U}})$. Then $\mathbb{P}$-a.e. $W \in \mathbf{W}_0^{\bar{U}}$. Let $\{\mathcal{B}_t(\mathbf{W}_0^{\bar{U}})\}_{t\geq 0}$ be the normal filtration generated by the canonical process ω. We obtain another complete probability space

$$(\mathbf{W}_0^{\bar{U}}, \vee_{t\geq 0}\mathcal{B}_t(\mathbf{W}_0^{\bar{U}}), \mathcal{B}_t(\mathbf{W}_0^{\bar{U}}); \bar{\mathbb{P}}),$$

where $\bar{\mathbb{P}}$ denotes the distribution of ω in $C([0,\infty);\bar{U})$, i.e.

$$\bar{\mathbb{P}} \circ \omega^{-1} = \mathbb{P} \circ W^{-1}. \tag{4.11}$$

Let ξ be $\mathcal{F}_0/\mathcal{B}(H)$-measurable and satisfy the assumptions in Theorem 2.3. Then by Theorem 2.3 and the Yamada-Watanabe Theorem in [21, Theorem E.1.8], there exists a measurable map

$$\mathcal{S}_{\mathbb{P}\circ\xi^{-1}} : (H \times \mathbf{W}_0^{\bar{U}}, \mathcal{B}(H)\otimes\mathcal{B}(\mathbf{W}_0^{\bar{U}})) \to (\mathcal{C}([0,T];H), \mathcal{B}(\mathcal{C}([0,T];H)))$$

such that $X := \mathcal{S}_{\mathbb{P}\circ\xi^{-1}}(\xi, W)$ is the solution to equation (2.2) with the initial value $X(0) = \xi$ $\mathbb{P}$-a.e. For simplicity we always denote $\mathcal{S}_{\mathbb{P}\circ\xi^{-1}}$ by $\mathcal{S}$. For $h \in L^2([0,T];U)$, define maps T_h^n on $(\mathbf{W}_0^{\bar{U}}, \mathcal{B}(\mathbf{W}_0^{\bar{U}}))$

$$T_h^n(x) = x - \int_0^{\cdot} \dot{x}^n(s)ds + \int_0^{\cdot} h(s)ds, \quad x \in \mathbf{W}_0^{\bar{U}},$$

where

$$\dot{x}^n(t) := \sum_{k=1}^{n} \langle \delta^{-1}\lambda_k [x(\lfloor \frac{t}{\delta} \rfloor \delta) - x((\lfloor \frac{t}{\delta} \rfloor - 1)\delta)], e_k \rangle_{\bar{U}} e_k, \ t \in [0, T].$$

Then by (4.1)-(4.3), T_h^n can be seen as measurable transformations of Wiener space $\mathbf{W}_0^{\bar{U}}$. Choose a $\mathcal{B}_0(\mathbf{W}_0^{\bar{U}})/\mathcal{B}(H)$-measurable map $\xi_0 : \mathbf{W}_0^{\bar{U}} \to H$ such that $\bar{\mathbb{P}} \circ \xi_0^{-1} = \mathbb{P} \circ \xi^{-1}$. Then $\mathcal{S}(\xi_0(\omega), \omega)$ is also a solution to equation (2.2) with initial condition ξ_0 and noise ω. Since ξ_0 is $\mathcal{B}_0(\mathbf{W}_0^{\bar{U}})/\mathcal{B}(H)$-measurable, $\xi_0(T_h^n(\omega)) = \xi_0(\omega)$. By the Yamada-Watanabe theorem, pathwise uniqueness implies that for every $\lambda > 0, n \in \mathbb{N}$

$$\bar{\mathbb{P}}(\omega : \|\mathcal{S}(\xi_0(\omega), T_h^n(\omega)) - Z_h\|_D \geq \lambda) = \mathbb{P}(\|Z_h^n - Z_h\|_D \geq \lambda), \tag{4.12}$$

where by Lemma 4.3 we have for the above λ

$$\lim_{n\to\infty} \mathbb{P}(\|Z_h^n - Z_h\|_D \geq \lambda) = 0. \tag{4.13}$$

$\bar{\mathbb{P}}_h^n = \bar{\mathbb{P}} \circ T_h^{n-1}, n \in \mathbb{N}$ together with (4.12),(4.13) implies that we can find some $n_0 \in \mathbb{N}$ such that

$$\bar{\mathbb{P}}_h^{n_0}(\omega : \|\mathcal{S}(\xi_0(\omega), \omega) - Z_h\|_D < \lambda) = \bar{\mathbb{P}}(\omega : \|\mathcal{S}(\xi_0(\omega), T_h^{n_0}(\omega)) - Z_h\|_D < \lambda) > 0.$$

Then by (4.3) $\bar{\mathbb{P}}_h^{n_0} \ll \bar{\mathbb{P}}$ we have

$$\mathbb{P}(\|X - Z_h\|_D < \lambda) = \bar{\mathbb{P}}(\omega : \|\mathcal{S}(\xi_0(\omega), \omega) - Z_h\|_D < \lambda) > 0.$$

So

$$Supp(\mathbb{P} \circ X^{-1}) \supset \bar{\mathcal{L}}.$$

Together with (4.10) we complete the proof. □

A Proof of Lemma 4.1.

The proof follows by a similar argument as in [16]. We consider the standard Galerkin approximation to equations (4.4). Let $\{g_1, g_2, \cdots\}$ be an orthonormal basis of H and set $H_m := span\{g_1, \ldots, g_m\}$. Let $P_m : V^* \to H_m$ be defined by

$$P_m y = \sum_{j=1}^{m} {}_{V^*}\langle y, g_j \rangle_V g_j, \quad y \in V^*. \tag{A.1}$$

For $h \in L^2([0, T]; U)$, consider the following equations on H_m

$$\begin{cases} dY_{h,m}^n(t) = A^m(t, Y_{h,m}^n(t))dt + B_1^m(Y_{h,m}^n(t))\Pi_m dW(t) \\ \qquad + B_2^m(Y_{h,m}^n(t))\dot{W}^n(t)dt + B_3^m(Y_{h,m}^n(t))h(t)dt - F^m(Y_{h,m}^n(t))dt, \\ Y_{h,m}^n(0) = P_m \xi, \end{cases} \tag{A.2}$$

with Π_m, P_m defined in (P1) and (A.1), $A^m = P_m A$, $F^m = P_m F$, $B_i^m = P_m B_i, i = 1,2,3$. For any $t \in [0,T]$, $\mathbb{E}(\|\dot{W}^n(t)\|_U^2) = n\delta^{-1}$ and

$$\langle [B(v_1) - B(v_2)]\dot{W}^n(t), v_1 - v_2\rangle \leq \|B(v_1) - B(v_2)\|_{L_2}\|\dot{W}^n(t)\|_U\|v_1 - v_2\|_H^2,$$

which by (H2) is dominated by $\big(\rho'(v_2) + \|\dot{W}^n(t)\|_U^2\big)\|v_1 - v_2\|_H^2$, $v_1, v_2 \in H_m$. [17, Theorem 3.1.1], Assumptions 1 and 2 imply that there exist unique solutions $Y_{h,m}^n$ to equations (A.2). In order to construct the solutions to equations (4.4), we first need some a priori estimate of $Y_{h,m}^n$.

Lemma A.1. *Under the assumptions in Lemma 4.1, there exists $C > 0$*

$$\begin{aligned} \sup_{n,m\in\mathbb{N}} \mathbb{E}\Big(\sup_{t\in[0,T]} \|Y_{h,m}^n(t)\|_H^p + \int_0^T \|Y_{h,m}^n(t)\|_H^{p-2}\|Y_{h,m}^n(t)\|_V^\alpha dt\Big) \\ \leq Ce^{\int_0^T(1+\|h(s)\|_U^2)ds}\Big(\mathbb{E}\|\xi\|_H^p + \|f\|_{L^{p/2}}^{\frac{p}{2}} + 1\Big). \end{aligned} \tag{A.3}$$

Proof. First we see that

$${}_{V^*}\langle A^m(t,u), v\rangle_V = \langle A^m(t,u), v\rangle = {}_{V^*}\langle A(t,u), v\rangle_V, \quad u \in V, v \in H_m.$$

Similar equalities also hold for B_2^m, B_3^m and F^m. Then we use Itô's formula for $\|Y_{h,m}^n(t)\|_H^p$ and consider each term on the right-hand side separately:

$$\begin{aligned} \|Y_{h,m}^n(t)\|_H^p =& p\int_0^t \|Y_{h,m}^n(s)\|_H^{p-2}{}_{V^*}\langle A(s, Y_{h,m}^n(s)), Y_{h,m}^n(s)\rangle_V \\ &+ p\int_0^t \|Y_{h,m}^n(s)\|_H^{p-2}\langle Y_{h,m}^n(s), B_1^m(Y_{h,m}^n(s))\Pi_m dW(s)\rangle \\ &+ p\int_0^t \|Y_{h,m}^n(s)\|_H^{p-2}\langle B_2(Y_{h,m}^n(s))\dot{W}^n(s), Y_{h,m}^n(s)\rangle ds \\ &+ p\int_0^t \|Y_{h,m}^n(s)\|_H^{p-2}\langle B_3(Y_{h,m}^n(s))h(s), Y_{h,m}^n(s)\rangle ds \\ &- p\int_0^t \|Y_{h,m}^n(s)\|_H^{p-2}\langle F(Y_{h,m}^n(s)), Y_{h,m}^n(s)\rangle ds + \|P_m\xi\|_H^p \\ &+ \frac{p}{2}\int_0^t \|Y_{h,m}^n(s)\|_H^{p-2}\|B_1^m(Y_{h,m}^n(s))\Pi_m\|_{L_2}^2 ds \\ &+ p(p-2)\int_0^t \|Y_{h,m}^n(s)\|_H^{p-4}\|\big(B_1^m(Y_{h,m}^n(s))\Pi_m\big)^* Y_{h,m}^n(s)\|_H^2 ds \\ =:& \sum_{i=1}^7 I_{n,m}(t,i) + \|P_m\xi\|_H^p, \end{aligned} \tag{A.4}$$

where $\|P_m\xi\|_H$ is dominated by $\|\xi\|_H$. For $I_{n,m}(t,1)$, by (H3) it is dominated by

$$\frac{p}{2}\int_0^t \|Y_{h,m}^n(s)\|_H^{p-2}\big(f(s) + K\|Y_{h,m}^n(s)\|_H^2 - \theta\|Y_{h,m}^n(s)\|_V^\alpha\big)ds,$$

and by Young's inequality, is further dominated by

$$\int_0^t \big(|f(s)|^{p/2}+(Kp+p-2)/2\|Y^n_{h,m}(s)\|^p_H-p\theta/2\|Y^n_{h,m}(s)\|^{p-2}_H\|Y^n_{h,m}(s)\|^\alpha_V\big)ds. \quad \text{(A.5)}$$

For $I_{n,m}(t,2)$, by the B-D-G inequality, $\mathbb{E}\big(\sup_{s\in[0,t]}\big|I_{n,m}(s,2)\big|\big)$ is bounded by

$$\mathbb{E}\Big(\lambda\sup_{s\in[0,t]}\|Y^n_{h,m}(s)\|^p_H+C_\lambda C(T,p)(\int_0^t\|Y^n_{h,m}(s)\|^p_H ds+1)\Big), \quad \text{(A.6)}$$

with $\lambda>0$ small enough and C_λ defined by Young's inequality.
For $I_{n,m}(t,3)$, we replace $\dot{W}^n(t)$ by $\frac{1}{\delta}\int_{(\lfloor\frac{t}{\delta}\rfloor-1)\delta}^{\lfloor\frac{t}{\delta}\rfloor\delta}\Pi_n dW(u)$. Using Fubini's theorem we see that $I_{n,m}(t,3)$ equals to

$$\int_0^t\sum_{k=0}^{\lfloor\frac{t}{\delta}\rfloor}\frac{p}{\delta}\int_{k\delta}^{(k+1)\delta\wedge t}1_{\{u\in[(k-1)\delta\vee 0,k\delta]\}}\langle\|Y^n_{h,m}(s)\|^{p-2}_H B_2(Y^n_{h,m}(s))^*Y^n_{h,m}(s),\Pi_m dW(u)\rangle ds.$$

Then by the B-D-G inequality, $\mathbb{E}\big(\sup_{t\in[0,T]}I_{n,m}(t,3)\big)$ is dominated by

$$\mathbb{E}\Big(\int_0^T\sum_{k=0}^{2^n-1}\frac{p^2}{\delta}\int_{k\delta}^{(k+1)\delta}ds1_{\{u\in[(k-1)\delta\vee 0,k\delta]\}}\|Y^n_{h,m}(s)\|^{2p-2}_H\|B_2(Y^n_{h,m}(s))\|^2_{L_2}du\Big)^{1/2},$$

which by Fubini's theorem, equals to $\mathbb{E}\big(\int_0^T p^2\|Y^n_{h,m}(s)\|^{2p-2}_H\|B_2(Y^n_{h,m}(s))\|^2_{L_2}ds\big)^{1/2}$. Together with (H4) we deduce

$$\begin{aligned}\mathbb{E}\Big(\sup_{t\in[0,T]}I_{n,m}(t,3)\Big)\le&\sqrt{K}p\mathbb{E}\Big(\int_0^T\|Y^n_{h,m}(s)\|^{2p-2}_H\big(1+\|Y^n_{h,m}(s)\|^2_H\big)ds\Big)^{1/2}\\ \le&\mathbb{E}\Big(\mu\sup_{t\in[0,T]}\|Y^n_{h,m}(t)\|^p_H+C(T,p)C_\mu\int_0^T\big(\|Y^n_{h,m}(s)\|^p_H+1\big)ds\Big),\end{aligned} \quad \text{(A.7)}$$

where $\mu>0$ is a small constant and C_μ is defined by Young's inequality.
For $I_{n,m}(t,4)$, by Young's inequality and (H4) there exists $C>0$ such that

$$\begin{aligned}I_{n,m}(t,4)\le&\int_0^t p\|Y^n_{h,m}(s)\|^{p-1}_H\sqrt{K}\big(1+\|Y^n_{h,m}(s)\|_H\big)\|h(s)\|_U ds\\ \le&C\int_0^t\big(\|h(s)\|^2_U\|Y^n_{h,m}(s)\|^p_H+\|Y^n_{h,m}(s)\|^p_H+1\big)ds.\end{aligned} \quad \text{(A.8)}$$

For $I_{n,m}(t,5)$, by Young's inequality and (P2) there exists $\lambda>0$ small enough such that for all $t\in[0,T]$, $I_{n,m}(t,5)$ is dominated by

$$\lambda\sup_{t\in[0,T]}\|Y^n_{h,m}(t)\|^p_H+C(T,p)C_\lambda\int_0^T\big(\|Y^n_{h,m}(s)\|^p_H+1\big)ds. \quad \text{(A.9)}$$

By (H4), the sum of $I_{n,m}(t,6)$ and $I_{n,m}(t,7)$ is dominated by

$$K\big[\frac{p}{2}+p(p-2)\big]\int_0^t\|Y^n_{h,m}(s)\|^{p-2}_H(1+\|Y^n_{h,m}(s)\|^2_H)ds.$$

Again by Young's inequality,

$$I_{n,m}(t,6) + I_{n,m}(t,7) \le C_p \int_0^T (1 + \|Y^n_{h,m}(s)\|_H^p) ds. \tag{A.10}$$

Insert (A.5)-(A.10) into (A.4) and by Gronwall's inequality we obtain (A.3). □

The rest of the proof is similar to the argument in [16] and we give all the details here for completeness. We follow the notations that:

$$J = L^\alpha([0,T]\times\Omega; dt\otimes\mathbb{P}; V);\ \ J^* = L^{\frac{\alpha}{\alpha-1}}([0,T]\times\Omega; dt\otimes\mathbb{P}; V^*),$$

$$K = L^2([0,T]\times\Omega; dt\otimes\mathbb{P}; L_2(U;H)).$$

Then according to Lemma A.1 and (H4), for all $n, m \in \mathbb{N}$ we have

$$\|Y^n_{h,m}\|_J + \|A(\cdot, Y^n_{h,m})\|_{J^*} < \infty.$$

Let $\lambda = \frac{\alpha}{p(\alpha-1)}$, which by assumption $p > \frac{\alpha}{\alpha-1}$ is less than 1. Then by Young's inequality there exists a constant $C_\lambda > 0$ such that

$$\|B_2(u)\dot{W}^n(t)\|_H^{\frac{\alpha}{\alpha-1}} \le \big(\lambda\|B_2(u)\|_{L_2}^{\frac{1}{\lambda}} + C_\lambda\|\dot{W}^n(t)\|_U^{\frac{1}{1-\lambda}}\big)^{\frac{\alpha}{\alpha-1}},$$

which by the inequality $(a+b)^{\frac{\alpha}{\alpha-1}} \le 2^{\frac{1}{\alpha-1}}(a^{\frac{\alpha}{\alpha-1}} + b^{\frac{\alpha}{\alpha-1}})$ and (H4), is further dominated by

$$2^{\frac{1}{\alpha-1}}\Big(\lambda^{\frac{\alpha}{\alpha-1}} K^{\frac{\alpha}{2\lambda(\alpha-1)}}\big(1+\|u\|_H^2\big)^{\frac{\alpha}{2\lambda(\alpha-1)}} + C_\lambda^{\frac{\alpha}{\alpha-1}}\|\dot{W}^n(t)\|_U^{\frac{\alpha}{(\alpha-1)(1-\lambda)}}\Big).$$

It means that there exists $C(\lambda,\alpha,K) > 0$ such that for all $u \in H$, $t \in [0,T]$

$$\|B_2(u)\dot{W}^n(t)\|_H^{\frac{\alpha}{\alpha-1}} \le C(\lambda,\alpha,K)\big(1 + \|u\|_H^p + \|\dot{W}^n(t)\|_U^{\frac{\alpha p}{\alpha p-\alpha-p}}\big).$$

By (2.4), the r.v.s. $\|\dot{W}^n(k\delta)\|_{U, k=1,\dots,2^n}$ are independent centered Gaussian with $\mathbb{E}\|\dot{W}^n(k\delta)\|_U^2 = n\delta^{-1}$. Then there exists a constant $C_{\alpha,p}$ such that

$$\mathbb{E}\int_0^T \|\dot{W}^n(t)\|_U^{\frac{\alpha p}{\alpha p-\alpha-p}} dt = \sum_{k=1}^{2^n} \delta\mathbb{E}\|\dot{W}^n(k\delta)\|_U^{\frac{\alpha p}{\alpha p-\alpha-p}} \le \sum_{k=1}^{2^n} \delta C_{\alpha,p}\mathbb{E}\|\dot{W}^n(k\delta)\|_U^2 = C_{\alpha,p} n 2^n.$$

Also by (P2) there exists $C_\alpha > 0$ such that

$$\|F(v)\|_H^{\frac{\alpha}{\alpha-1}} \le C_\alpha(1+\|v\|_H^2)^{\frac{\alpha}{2(\alpha-1)}},\ v \in H.$$

Let c_1 be a constant such that $\|\cdot\|_{V^*} \le c_1\|\cdot\|_H$, again using the assumption $p > \frac{\alpha}{\alpha-1}$ and Lemma A.1 we deduce that for each $n \in \mathbb{N}$

$$\|B_2(Y^n_{h,m})\dot{W}^n + F(Y^n_{h,m})\|_{J^*} < \infty \text{ uniformly for } m \in \mathbb{N}.$$

Thus for each $n \in \mathbb{N}$, there exists a subsequence $m_k(n) \to \infty$ (which we still denote by m_k for simplicity):

1. $Y^n_{h,m_k} \to \check{Y}^n_h$ weakly in J and weakly star in $L^p(\Omega; L^\infty([0,T]; H))$.

2. $A^{m_k}(\cdot, Y^n_{h,m_k}) \to U^n_h$ weakly in J^*.

3. $B_2^{m_k}(Y^n_{h,m_k})\dot{W}^n + B_3^{m_k}(Y^n_{h,m_k})h - F^{m_k}(Y^n_{h,m_k}) \to Z^n_h + V^n_h - M^n_h$ weakly in J^*.

4. $B_1^{m_k}(Y^n_{h,m_k}) \to N^n_h$ weakly in K and hence

$$\int_0^\cdot B_1^{m_k}(Y^n_{h,m_k})\Pi_{m_k} dW \to \int_0^\cdot N^n_h dW \quad \text{weakly in } L^\infty([0,T]; dt; L^2(\Omega, \mathbb{P}; H)).$$

Now we define the process below: $t \in [0,T]$,

$$Y^n_h(t) := \xi + \int_0^t \big(U^n_h(s) + Z^n_h(s) + V^n_h(s) - M^n_h(s)\big) ds + \int_0^t N^n_h(s) dW(s). \quad \text{(A.11)}$$

Following the proof in [17, Theorem 4.2.4] we similarly show $Y^n_h = \check{Y}^n_h,\ dt \otimes \mathbb{P}$-a.e. Then together with [17, Theorem 4.2.5] and Lemma A.1 we know that Y^n_h is an H-valued continuous $(\mathcal{F}_t)$-adapted process. Therefore, for the existence of solutions to (4.4) it remains to verify that

$$\begin{aligned} A(\cdot, Y^n_h) + B_2(Y^n_h)\dot{W}^n + B_3(Y^n_h)h - F(Y^n_h) &= U^n_h + Z^n_h + V^n_h - M^n_h, \\ B_1(Y^n_h) &= N^n_h, \qquad dt \otimes \mathbb{P}. \end{aligned} \quad \text{(A.12)}$$

Let $\rho,\ \rho'$ be defined by (H4) and set

$$\mathcal{M} := \Big\{\phi : \phi \text{ is } V\text{-valued } \mathcal{F}_t \text{ -adapted process, } \mathbb{E}\int_0^T \rho(\phi_s) ds < \infty\Big\}.$$

For $\phi \in J \cap \mathcal{M} \cap L^p(\Omega; L^\infty([0,T]; H))$, we deduce that

$$\begin{aligned} &\mathbb{E}\Big(e^{-\int_0^t (f_s + \rho(\phi_s) + 4\rho'(\phi_s) + \|\dot{W}^n(s)\|_U^2 + \|h_s\|_U^2) ds} \|Y^n_{h,m_k}(t)\|_H^2\Big) - \mathbb{E}\Big(\|P_{m_k}\xi\|_H^2\Big) \\ =&\mathbb{E}\Big(\int_0^t e^{-\int_0^s (f_r + \rho(\phi_r) + 4\rho'(\phi_r) + \|\dot{W}^n(r)\|_U^2 + \|h_r\|_U^2) dr} \Big[2_{V^*}\langle A(s, Y^n_{h,m_k}(s)), Y^n_{h,m_k}(s)\rangle_V \\ &\quad + 2\langle B_2(Y^n_{h,m_k}(s))\dot{W}^n(s) + B_3(Y^n_{h,m_k}(s))h_s - F(Y^n_{h,m_k}(s)), Y^n_{h,m_k}(s)\rangle \\ &\quad + \|B_1^{m_k}(Y^n_{h,m_k}(s))\Pi_{m_k}\|_{L_2}^2 \\ &\quad - \big(f_s + \rho(\phi_s) + 4\rho'(\phi_s) + \|\dot{W}^n(s)\|_U^2 + \|h_s\|_U^2\big)\|Y^n_{h,m_k}(s)\|_H^2\Big] ds\Big), \end{aligned}$$

which is further dominated by

$$
\begin{aligned}
&\mathbb{E}\Big(\int_0^t e^{-\int_0^s (f_r+\rho(\phi_r)+4\rho'(\phi_r)+\|\dot W^n(r)\|_U^2+\|h_r\|_U^2)dr} \\
&\quad\Big[2_{V^*}\langle A(s,Y^n_{h,m_k}(s))-A(s,\phi_s),Y^n_{h,m_k}(s)-\phi_s\rangle_V+\|B_1(Y^n_{h,m_k}(s))-B_1(\phi_s)\|_{L_2}^2 \\
&\quad+2\langle[B_2(Y^n_{h,m_k}(s))-B_2(\phi_s)]\dot W^n(s)+[B_3(Y^n_{h,m_k}(s))-B_3(\phi_s)]h_s \\
&\qquad\quad-[F(Y^n_{h,m_k}(s))-F(\phi_s)],Y^n_{h,m_k}(s)-\phi_s\rangle \\
&\quad-\big(f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot W^n(s)\|_U^2+\|h_s\|_U^2\big)\|Y^n_{h,m_k}(s)-\phi_s\|_H^2\Big]ds\Big) \\
&\;+\mathbb{E}\Big(\int_0^t e^{-\int_0^s (f_r+\rho(\phi_r)+4\rho'(\phi_r)+\|\dot W^n(r)\|_U^2+\|h_r\|_U^2)dr} \\
&\quad\Big[2_{V^*}\langle A(s,Y^n_{h,m_k}(s))-A(s,\phi_s),\phi_s\rangle_V+2_{V^*}\langle A(s,\phi_s),Y^n_{h,m_k}(s)\rangle_V \\
&\quad+2\langle[B_2(Y^n_{h,m_k}(s))-B_2(\phi_s)]\dot W^n(s),\phi_s\rangle+2\langle B_2(\phi_s)\dot W^n(s),Y^n_{h,m_k}(s)\rangle \\
&\quad+2\langle[B_3(Y^n_{h,m_k}(s))-B_3(\phi_s)]h_s,\phi_s\rangle+2\langle B_3(\phi_s)h_s,Y^n_{h,m_k}(s)\rangle \\
&\quad-2\langle F(Y^n_{h,m_k}(s))-F(\phi_s),\phi_s\rangle-2\langle F(\phi_s),Y^n_{h,m_k}(s)\rangle \\
&\quad-\|B_1(\phi_s)\|_{L_2}^2+2\langle B_1(Y^n_{h,m_k}(s)),B_1(\phi_s)\rangle_{L_2} \\
&\quad-2(f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot W^n(s)\|_U^2+\|h_s\|_U^2)\langle Y^n_{h,m_k}(s),\phi_s\rangle \\
&\quad+(f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot W^n(s)\|_U^2+\|h_s\|_U^2)\|\phi_s\|_H^2\Big]ds\Big).
\end{aligned}
$$

We first note that for any nonnegative $\psi\in L^\infty([0,T];dt)$,

$$
\begin{aligned}
&\mathbb{E}\Big(\int_0^T \psi_t\|Y^n_h(t)\|_H^2dt\Big)=\lim_{k\to\infty}\mathbb{E}\Big(\int_0^T\langle\psi_tY^n_h(t),Y^n_{h,m_k}(t)\rangle dt\Big) \\
\le&\Big(\mathbb{E}\int_0^T\psi_t\|Y^n_h(t)\|_H^2dt\Big)^{1/2}\liminf_{k\to\infty}\Big(\mathbb{E}\int_0^T\psi_t\|Y^n_{h,m_k}(t)\|_H^2dt\Big)^{1/2}.
\end{aligned}
$$

Together with (H2) and (P2), this implies that

$$
\begin{aligned}
&\mathbb{E}\Big(\int_0^T \psi_t \Big[e^{-\int_0^t (f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot{W}^n(s)\|_U^2+\|h_s\|_U^2)ds} \|Y_h^n(t)\|_H^2 - \|\xi\|_H^2 \Big] dt \Big) \\
\le& \liminf_{k\to\infty} \mathbb{E}\Big(\int_0^T \psi_t \Big[e^{-\int_0^t (f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot{W}^n(s)\|_U^2+\|h_s\|_U^2)ds} \|Y_{h,m_k}^n(t)\|_H^2 - \|P_{m_k}\xi\|_H^2 \Big] dt \Big) \\
\le& \mathbb{E}\Big(\int_0^T \psi_t \Big[\int_0^t e^{-\int_0^s (f_r+\rho(\phi_r)+4\rho'(\phi_r)+\|\dot{W}^n(r)\|_U^2+\|h_r\|_U^2)dr} \\
&\quad \Big(2_{V^*}\langle U_h^n(s) - A(s,\phi_s), \phi_s\rangle_V + 2_{V^*}\langle A(s,\phi_s), Y_h^n(s)\rangle_V \\
&\quad + 2\langle Z_h^n(s) - B_2(\phi_s)\dot{W}^n(s), \phi_s\rangle + 2\langle B_2(\phi_s)\dot{W}^n(s), Y_h^n(s)\rangle \\
&\quad + 2\langle V_h^n(s) - B_3(\phi_s)h_s, \phi_s\rangle + 2\langle B_3(\phi_s)h_s, Y_h^n(s)\rangle \\
&\quad - 2\langle M_h^n(s) - F(\phi_s), \phi_s\rangle - 2\langle F(\phi_s), Y_h^n(s)\rangle \\
&\quad - \|B_1(\phi_s)\|_{L_2}^2 + 2\langle N_h^n(s), B_1(\phi_s)\rangle_{L_2} \\
&\quad - 2\big(f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot{W}^n(s)\|_U^2+\|h_s\|_U^2\big)\langle Y_h^n(s), \phi_s\rangle \\
&\quad + \big(f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot{W}^n(s)\|_U^2+\|h_s\|_U^2\big)\|\phi_s\|_H^2 \Big) ds \Big] dt \Big).
\end{aligned} \tag{A.13}
$$

We also have the following equality:

$$
\begin{aligned}
&\mathbb{E}\Big(e^{-\int_0^t (f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot{W}^n(s)\|_U^2+\|h_s\|_U^2)ds} \|Y_h^n(t)\|_H^2 - \|\xi\|_H^2 \Big) \\
=& \mathbb{E}\Big(\int_0^t e^{-\int_0^s (f_r+\rho(\phi_r)+4\rho'(\phi_r)+\|\dot{W}^n(r)\|_U^2+\|h_r\|_U^2)dr} \big[2_{V^*}\langle U_h^n(s), Y_h^n(s)\rangle_V \\
&\quad + 2\langle Z_h^n(s) + V_h^n(s) - M_h^n(s), Y_h^n(s)\rangle + \|N_h^n(s)\|_{L_2}^2 \\
&\quad - \big(f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot{W}^n(s)\|_U^2+\|h_s\|_U^2\big)\|Y_h^n(s)\|_H^2 \big] ds \Big).
\end{aligned} \tag{A.14}
$$

Combining (A.13) and (A.14) we obtain that

$$
\begin{aligned}
0 \ge& \mathbb{E}\Big(\int_0^T \psi_t \Big[\int_0^t e^{-\int_0^s (f_r+\rho(\phi_r)+4\rho'(\phi_r)+\|\dot{W}^n(r)\|_U^2+\|h_r\|_U^2)dr} \\
&\quad \Big(2_{V^*}\langle U_h^n(s) - A(s,\phi_s), Y_h^n(s) - \phi_s\rangle_V \\
&\quad + 2\langle Z_h^n(s) - B_2(\phi_s)\dot{W}^n(s) + V_h^n(s) - B_3(\phi_s)h_s \\
&\qquad - M_h^n(s) + F(\phi_s), Y_h^n(s) - \phi_s\rangle + \|B_1(\phi_s) - N_h^n(s)\|_{L_2}^2 \\
&\quad - \big(f_s+\rho(\phi_s)+4\rho'(\phi_s)+\|\dot{W}^n(s)\|_U^2+\|h_s\|_U^2\big)\|Y_h^n(s) - \phi_s\|_H^2 \Big) ds \Big] dt \Big).
\end{aligned} \tag{A.15}
$$

Note that Lemma $A.1$ and (H4) imply that

$$
Y_h^n \in J \cap \mathcal{M} \cap L^p(\Omega; L^\infty([0,T]; H)).
$$

Thus for (A.15) if we first take $\phi = Y_h^n - \epsilon\tilde{\phi}v$ for $\tilde{\phi} \in L^\infty([0,T]\times\Omega; dt\otimes\mathbb{P}; \mathbb{R})$ and

$v \in V$, then divide it by ϵ and let $\epsilon \to 0$, we finally have

$$0 \geq \mathbb{E}\Big(\int_0^T \psi_t \big[\int_0^t e^{-\int_0^s (f_r + \rho(Y_h^n(r)) + 4\rho'(Y_h^n(r)) + \|\dot{W}^n(r)\|_U^2 + \|h_r\|_U^2) dr} \tilde{\phi}_s$$
$$\big(2_{V^*}\langle U_h^n(s) - A(s,\phi_s), v\rangle_V + 2\langle Z_h^n(s) - B_2(\phi_s)\dot{W}^n(s)$$
$$+ V_h^n(s) - B_3(\phi_s)h_s - F(\phi_s) + M_h^n(s), v\rangle + \|B_1(\phi_s) - N_h^n(s)\|_{L_2}^2 \big) ds \big] dt \Big).$$

Because of the arbitrariness of ψ and $\tilde{\phi}$ we obtain (A.12). Therefore, Y_h^n are solutions to (2.6). For further estimate of $\|Y_h^n(t)\|_H^p$, we repeat the method used in the proof of Lemma A.1 and similarly obtain that there exists $C > 0$ such that $\mathbb{E}(\sup_{t\in[0,T]} \|Y_h^n(t)\|_H^p + \int_0^T \|Y_h^n(t)\|_H^{p-2} \|Y_h^n(t)\|_V^\alpha dt)$ is dominated by

$$Ce^{\int_0^T (1+\|h(s)\|_U^2) ds} \big(\mathbb{E}\|\xi\|_H^p + \|f\|_{L^{p/2}}^{\frac{p}{2}} + 1 \big) \text{ uniformly for all } n.$$

Uniqueness: For any $n \in \mathbb{N}$ given, let X_h^n, Y_h^n be the solutions to equations (2.6) with initial values X_0 and Y_0 respectively. Then by (H2), (H4) and (P2) we have the following estimate

$$\mathbb{E}\big(e^{-\int_0^t (f_s + \rho(Y_h^n(s)) + 4\rho'(Y_h^n(s)) + \|\dot{W}^n(s)\|_U^2 + \|h_s\|_U^2) ds} \|X_h^n(t) - Y_h^n(t)\|_H^2 \big) \leq \mathbb{E}\|X_0 - Y_0\|_H^2.$$

So if $X_0 = Y_0$, $\mathbb{P}$-a.s., we easily have

$$\mathbb{E}\big(e^{-\int_0^t (f_s + \rho(Y_h^n(s)) + 4\rho'(Y^n(s)) + \|\dot{W}^n(s)\|_U^2 + \|h_s\|_U^2) ds} \|Y_h^n(t) - X_h^n(t)\|_H^2 \big) = 0.$$

By (H4) and Lemma A.1 we have

$$\int_0^t \big(f_s + \rho(Y_h^n(s)) + 4\rho'(Y_h^n(s)) + \|\dot{W}^n(s)\|_U^2 + \|h_s\|_U^2 \big) ds < \infty, \ \mathbb{P}\text{-a.s.}, \quad t \in [0,T].$$

Then we obtain that

$$X_h^n(t) = Y_h^n(t), \ \mathbb{P}\text{-a.e.}, \quad t \in [0,T].$$

The pathwise uniqueness follows from the path continuity of X_h^n and Y_h^n in H.

Acknowledgement

We would like to thank Professor Zhiming Ma for his encouragement and suggestions for this work.

References

[1] D. G. Aronson. The Porous Medium Equation. *Nonlinear Diffusion Problems (A. Fasano and M. Primicerio, eds.), Lecture Notes in Mathematics 1224, Springer Verlag, New York*, (1986) 1-46.

[2] V. Bally, A. Millet, M. Sanz-Solé. Approximation and support theorem in Hölder norm for parabolic stochastic partial differential equations. *Annals of Probability 23*, (1995) 178-222.

[3] C. Cardon-Weber, A. Millet. A support theorem for a generalized Burgers equation. *Potential Analysis*, (2001) 361-408.

[4] I. Chueshov, A. Millet. Stochastic 2D hydrodynamical type systems: Well posedness and large deviations. *Applied Mathematics and Optimization*, (2010) 379-420.

[5] I. Chueshov, A. Millet. Stochastic 2d hydrodynamical systems:Wong-Zakai approximation and support theorem. *Stochastic Analysis and Applications*, (2011) 570-611.

[6] G. Da Prato, J. Zabczyk. Stochastic Equations in Infinite Dimensions. *Encyclopedia of Mathematics and its Applications, Cambridge University Press*, (1992).

[7] B. Grigelionis, R. Mikulevičius. Stochstic evolution equations and densities of the conditional distributions. *Theory and Application of Random Fields*, (1983) 49-88.

[8] I. Gyöngy, N. V. Krylov. On stochastic equations with respect to semimartingales. I. *Stochastics, 4*, (1980/81) 1-21.

[9] I. Gyöngy, A. Shmatkov. Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. *Appl Math Optim*, (2006) 315-341.

[10] I. Gyöngy, P. R. Stinga. Rate of convergence of Wong-Zakai approximations for stochastic partial differential equations. *Probability (math.PR); Numerical Analysis (math.NA)*, (2012).

[11] I. Gyöngy, N. V. Krylov. On stochastic equations with respect to semimartingales. II. *Itô formula in Banach spaces, Stochastics, 6*, (1981/82) 153-173.

[12] I. Gyöngy, T. Pröhle. On the approximation of stochastic differential equation and on Stroock-Varadhan's support theorem. *Computers Math. Appfic, Vol. 19, No. 1*, (1990) 65-70.

[13] I. Gyöngy. On stochastic equations with respect to semimartingales. III. *Stochastics, 7*, (1982) 231-254.

[14] M. Hairer, É. Pardoux. A Wong-Zakai theorem for stochastic PDEs. *J. Math. Soc. Japan (67), No. 4*, (2015) 1551-1604.

[15] N. V. Krylov, B. L. Rozovskiĭ. Stochastic evolution equations. *Current problems in mathematics, Vol. 14 (Russian), Akad. Nauk SSSR, Vsesoyuz. Inst. Nauchn. i Tekhn. Informatsii, Moscow*, (1979) 71-147.

[16] W. Liu, M. Röckner. SPDE in Hilbert space with locally monotone coefficients. *J. Funct. Anal*, (2010) 2902-2922.

[17] W. Liu, M. Röckner. Stochastic Partial Differential Equations: An Introduction. *Springer*, (2015).

[18] A. Millet, M. Sanz-Solé. A simple proof of the support theorem for diffusion processes. *Séminaire de probabilités, tome 28*, (1994) 36-48.

[19] V. Mackevičius. On the support of the solution of stochastic differential equations. *Lietuvos matematikow rinkings XXXVI(1)*, (1986) 91-98.

[20] T. Nakayama. Support theorem for mild solutions of SDEs in Hilbert spaces. *J. Math. Sci. Univ. Tokyo*, 11 (2004) 245-311.

[21] C. Prévôt, M. Röckner. A concise course on stochastic partial differential equations. *Lecture Notes in Math., vol. 1905, Springer, Berlin*, (2007).

[22] É. Pardoux. Sur des équations aux dérivées partielles stochastiques monotones. *C. R. Acad. Sci. Paris Sér. A-B*, (1972) 101-103.

[23] D. W. Stroock, S. R. S. Varadhan. On the support of diffusion processes with applications to the strong maximum principle. *Proc. of Sixth Berkeley Sym. Math. Stat.Prob., Univ. California Press, Berkeley*, (1972) 333-359.

[24] M. Sermange, R. Temam. Some mathematical questions related to MHD equations. *Communications in Pure and Applied Mathematics*, (1983) 635-664.

[25] K. Twardowska. Wong-Zakai approximations for stochastic differential equations. *Acta Applic. Math.* 43 (1996) 317-369.

[26] E. Wong, M. Zakai. On the convergence of ordinary integrals to stochastic integrals. *Ann. Math. Statist*, 36 (1965) 1560-1564.

[27] S. Watanabe, N. Ikeda. Stochastic differential equations and diffusion processes. *North Holland, Amsterdam*, (1981).