**IMA Journal of Numerical Analysis Advance Access published July 11, 2016**

*IMA Journal of Numerical Analysis* (2016) **00**, 1–36
doi: 10.1093/imanum/drw030

# Approximation of the invariant law of SPDEs: error analysis using a Poisson equation for a full-discretization scheme

CHARLES-EDOUARD BRÉHIER*
*Institut Camille Jordan, Université Claude Bernard Lyon 1, CNRS UMR 5208, 43 Blvd. du 11 Novembre 1918, F-69622 Villeurbanne cedex, France*
*Corresponding author: brehier@math.univ-lyon1.fr

AND

MARIE KOPEC
*ENS Rennes - IRMAR, Université Rennes 1, Avenue Robert Schumann, F-35170 Bruz, France*
marie.kopec@bretagne.ens-cachan.fr

[Received on 27 November 2014; revised on 23 May 2016]

We study the long-time behavior of fully discretized semilinear stochastic partial different equations (SPDEs) with additive space-time white noise, which admits a unique invariant probability measure $\mu$. We show that the average of (regular) test functions with respect to the (possibly nonunique) invariant laws of the approximations are close to the corresponding average with respect to $\mu$.

More precisely, we analyse the rate of convergence with respect to time and space discretization parameters. Here we focus on the discretization in time thanks to a scheme of Euler type, and on a finite element discretization in space. The main new contribution here is the treatment of the spatial error.

The technique of the proof is original in the SPDE context: we generalize the approach of Mattingly *et al.* (2010, Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.*, **48**, 552–577), which relies on the use of a Poisson equation, to an infinite-dimensional setting. We show that the rates of convergence for the invariant laws are given by the corresponding weak orders of the discretization on finite time intervals: order $1/2$ with respect to the time step and order 1 with respect to the mesh size.

*Keywords*: stochastic partial differential equations; invariant measures and ergodicity; weak approximation; Euler scheme; finite element method; Poisson equation.



## 1. Introduction

In this article, we want to analyse in a quantitative way the effect of time and space discretization schemes on the knowledge of the unique invariant law of a semilinear stochastic partial different equation (SPDE) of parabolic type, written in the abstract form of Da Prato & Zabczyck (1992):

$$\begin{aligned} \mathrm{d}X(t,x) &= \big(AX(t,x) + F(X(t,x))\big)\,\mathrm{d}t + \mathrm{d}W(t), \quad 0 < t \leq T, \\ X(0,x) &= x. \end{aligned} \tag{1.1}$$

This process takes values in an infinite-dimensional, separable Hilbert space $H$—typically $H = L^2(0,1)$—and $x \in H$ denotes an arbitrary initial condition; $A$ is a negative, self-adjoint, unbounded linear operator on $H$, with a compact inverse—for instance, $A = \partial^2/\partial\xi^2$, with domain $H^2(0,1) \cap H_0^1(0,1)$ when homogeneous Dirichlet boundary conditions are applied; see Assumptions 2.2 and Example 2.3. The coefficient

$F : H \to H$ is a nonlinear function, with appropriate regularity and growth conditions given by Assumptions 2.7 and 2.9. Finally, $\big(W(t)\big)_{t\in[0,T]}$ is a cylindrical Wiener process on $H$ (see Section 2.3): this means that in SPDE (1.1) the Gaussian noise is white in time and in space. We are interested in the regime where $T$ is arbitrarily large and goes to $+\infty$.

In this setting, thanks to dissipativity conditions stated in Assumptions 2.9, it is known that SPDE (1.1) admits a unique invariant probability measure $\mu$, and that convergence is exponentially fast; see Proposition 4.1. The main arguments that lead to ergodicity are recalled in Section 4. We refer for instance to Da Prato & Zabczyck (1996) and Debussche (2013), and references therein, for treatments of the asymptotic behavior of SPDEs.

In general, no expression of $\mu$ is available for practical use; moreover, the support of this measure is an infinite-dimensional space. The approximation of averages $\int_H \phi \,\mathrm{d}\mu$ for bounded test functions $\phi$ is therefore complicated. The exponential convergence ensures that $\mathbb{E}\phi(X(t))$ tends to $\int_H \phi \,\mathrm{d}\mu$ when $t$ tends to infinity, exponentially fast. Since it is not possible to simulate exactly the $H$-valued random variable $X(t)$ for every $t \geq 0$, two discretization schemes are introduced:

- a discretization in time, in order to get an approximation of the law of the random variables $X(t)$ for different, fixed values of $t$, using a finite number of calculations; here it is performed with a semiimplicit Euler scheme;
- a discretization in space, in order to sample finite-dimensional random variables; here it is performed with a finite element method.

The spatial approximation is specific to the case of infinite-dimensional processes, solutions of SPDEs (and PDEs); the temporal approximation has already been studied extensively in the case of SDEs, and more recently for SPDEs. One of the main original contributions of this article is the analysis of the long-time behavior of a numerical scheme when considering a spatial discretization of an SPDE.

Different techniques to control the error are available in the literature. A first method is presented in Talay (1990), where an estimate of the weak error introduced by the numerical scheme is proved, holding for any value of the finite time $T$. The idea here is to expand the error, thanks to the solution of the Kolmogorov equation associated with the diffusion, and to prove bounds on the spatial derivatives of this solution, with an exponential decrease with respect to the time variable.

This strategy has been generalized to the class of SPDEs (1.1) in Bréhier (2014), where a semiimplicit Euler scheme is used. The main additional difficulty, when compared with the SDE case, is the need for tools introduced in Debussche (2011), to estimate the weak error.

Using these tools aims at proving that at a given time $T \in (0, +\infty)$, the weak order of convergence is twice the strong one: in other words, laws at fixed times are approximated more accurately than the trajectories. These tools have also been used in Wang & Gan (2013) to treat the time discretization in a slightly more-general setting, and in Andersson & Larsson (2016) where discretization in space with a finite element method is studied. Basically, the two ingredients are the following:

- improved estimates on the derivatives of the solution of the Kolmogorov equations, with spatial regularization;
- an integration by parts formula using Malliavin calculus, in order to transform some stochastic expressions with insufficient spatial regularity into more suitable ones.

These tools are fundamental to treat equations with nonlinear terms; they are used again in the present work. Notice that for linear equations (i.e., $F = 0$ in (1.1)) a specific idea simplifies the proof—so that the second tool is not required—but cannot be adapted to nonlinear parabolic equations like (1.1) and in particular cannot be used in this article: see Debussche & Printems (2009), and De Bouard & Debussche (2006) in the case of the discretization of a stochastic Schrödinger equation.

Here, we apply another method to analyse the approximation of the invariant measure: we follow the approach of Mattingly *et al.* (2010). Here, the authors study the distance between time averages of a test function $\phi$ along the realization of the numerical scheme, and its average $\int \phi \, \mathrm{d}\mu$ with respect to the invariant law $\mu$. They introduce the solution $\Psi$ of the Poisson equation

$$\mathcal{L}\Psi = \phi - \int \phi \, \mathrm{d}\mu, \tag{1.2}$$

where $\mathcal{L}$ is the infinitesimal generator of the SDE; the solvability of this elliptic or hypoelliptic PDE is ensured by ergodic properties. Then they show how to expand the error for various numerical methods, using stochastic Taylor expansions, and they get convergence results.

Using a Poisson equation to prove convergence results of law of large numbers type is standard, as explained in Mattingly *et al.* (2010). In the context of SPDEs, it has been notably used in Bréhier (2012) and Cerrai & Freidlin (2009) to study the averaging principle for systems evolving with two separate timescales.

Note that even if the numerical scheme is not ergodic, and thus having possibly several invariant laws, the technique gives an approximation result for $\mu$. In the SDE case, the study of ergodicity for time-discretized processes is treated in Higham *et al.* (2002) where the authors use general results on Markov chains, like the Harris theorem. To our knowledge, there is no such general theory for the discretization of SPDEs.

Our main result is the generalization of SPDEs for the approach of Mattingly *et al.* (2010), with time and space approximation procedures: we prove the following result—a more precise statement is Theorem 5.1: for any function $\phi$ of class $\mathcal{C}_b^2(H)$, there exists a constant $C(\phi) > 0$ such that for any parameters $\tau \in (0,1)$ and $h \in (0,1)$, any time $N \geq 1$ and any initial condition $x \in H$,

$$\left| \frac{1}{N} \sum_{m=0}^{N-1} \left( \mathbb{E}\phi(X_m^h) - \int_H \phi(z)\mu(\mathrm{d}z) \right) \right| \leq C(\phi) \left( \tau^{1/2} + h + \frac{1}{N\tau} \right).$$

One of the key tools to prove Theorem 5.1 is the study of the Poisson equation (1.2) associated with the SPDE (1.1). More precisely, we work with Galerkin approximations and we prove bounds that are independent of the dimension $M$ of the approximating subspace; see Sections 3.2 and 6.2. We emphasize the necessity of regularization properties for the solutions of the Poisson equation to obtain the correct weak orders. Many arguments and error terms are reminiscent of Debussche (2011), Bréhier (2014) and Andersson & Larsson (2016), where similar regularization properties on the Kolmogorov equation are used: this means that the method used in this article is a variant of the previous methods but does not simplify the proof.

We refer to Section 5.2 for a discussion of cases which do not rigorously fit into the setting of Section 2, but for which Theorem 5.1 can be extended with only straightforward modifications of the technical arguments. In particular, some SPDEs with noise colored in space, in space dimension 2 or 3, can be considered. However, in order to avoid introducing additional notation (the orders of convergence depend on the regularity of the noise) and to emphasize the main arguments, we prove results for SPDEs in dimension 1, with space-time white noise.

The article is organized as follows. Section 2 is devoted to the statement of assumptions on the coefficients of (1.1). The space and time discretization schemes are presented in Section 3. In Section 4, we state results on the asymptotic behavior and invariant laws of the processes, with emphasis on the consequences of the dissipativity conditions. The main result of this article, Theorem 5.1, is given in Section 5; possible extensions and open questions are discussed in Section 5.2. The main ingredients of the proof, namely the Poisson equation (Section 6.2), the decomposition of the error (Section 6.3) and an integration by parts formula (Section 6.4), are given in Section 6. Finally, detailed proofs of error estimates are given in Section 7.

Note that in the search for conciseness we omit the details for (most of) the error terms due to the time discretization, since the arguments are not essentially different from those in Bréhier (2014). We refer to the Ph.D. thesis (Kopec, 2014) of the second named author for complete details. This allows us to focus on the main contribution of the article: the treatment of space discretization.

## 2. Notation and assumptions

Let $\mathcal{D} = (0,1)$. Let $H = L^2(\mathcal{D})$, with norm and inner product denoted by $|\cdot|_H$ and $\langle\cdot,\cdot\rangle_H$ or if no confusion is possible $|\cdot|$ and $\langle\cdot,\cdot\rangle$.

We consider equations in the abstract form

$$\begin{aligned} \mathrm{d}X(t,x) &= (AX(t,x) + F(X(t,x)))\,\mathrm{d}t + \mathrm{d}W(t), \\ X(0,x) &= x \in H. \end{aligned} \tag{2.1}$$

We now state the assumptions made on the coefficients $A$ and $F$ in (1.1). We also recall standard statements concerning the cylindrical Wiener process $W$, and on the mild solution of the SPDE. We refer to Da Prato & Zabczyck (1992) for more details.

### 2.1 *Test functions*

The test function $\phi$—which we refer to as *admissible* in the sequel—is assumed to belong to the space $\mathcal{C}_b^2(H,\mathbb{R})$ of twice continuously differentiable functions $\phi : H \to \mathbb{R}$, which are bounded, and have bounded first- and second-order derivatives.

REMARK 2.1 In the sequel, we often identify the first derivative $D\phi(x)$ with the gradient in the Hilbert space $H$, and the second derivative $D^2\phi(x)$ with a linear operator on $H$ via,

$$\begin{aligned} \langle D\phi(x), h\rangle &= D\phi(x)\cdot h \quad \text{for every } h \in H, \\ \langle D^2\phi(x)\cdot h, k\rangle &= D^2\phi(x)\cdot(h,k) \quad \text{for every } h,k \in H. \end{aligned}$$

For an admissible test function $\phi$ and $i \in \{1,2\}$, set $\| \phi \|_{i,\infty} = \sup_{0\le j\le i}(\| \phi \|_j)$ with

$$\|\phi\|_0 = \sup_{x\in H}|\phi(x)|_H, \quad \|\phi\|_1 = \sup_{x\in H}|D\phi(x)|_H, \quad \|\phi\|_2 = \sup_{x\in H}|D^2\phi(x)|_{\mathcal{L}(H)}.$$

### 2.2 *Assumptions on the coefficients*

2.2.1 *The linear operator.* We denote by $\mathbb{N} = \{0,1,2,\ldots\}$ the set of non-negative integers.

We assume that the following properties are satisfied.

ASSUMPTIONS 2.2 (1) There exists a complete orthonormal system $(e_k)_{k\in\mathbb{N}}$ of $H$ and a nondecreasing sequence $(\lambda_k)_{k\in\mathbb{N}}$ in $(0,+\infty)$ such that

$$Ae_k = -\lambda_k e_k \quad \text{for all } k \in \mathbb{N}.$$

(2) The sequence $(\lambda_k)_{k\in\mathbb{N}}$ goes to $+\infty$ and

$$\sum_{k=0}^{+\infty} \frac{1}{\lambda_k^\alpha} < +\infty \quad \Leftrightarrow \quad \alpha > \frac{1}{2}.$$

The smallest eigenvalue of $-A$ is then $\lambda_0$. Note that Assumptions 2.2 implies that $A$ is self-adjoint, with a compact resolvent.

EXAMPLE 2.3 We can choose $A = \mathrm{d}^2/\mathrm{d}x^2$, with the domain $H^2(0,1) \cap H_0^1(0,1) \subset L^2(0,1)$—corresponding to homogeneous Dirichlet boundary conditions. In this case, for all $k \in \mathbb{N}$, $\lambda_k = \pi^2(k+1)^2$, and $e_k(\xi) = \sqrt{2}\sin((k+1)\pi\xi)$—see for instance Brézis (1994).

Let us now introduce, for each $M \in \mathbb{N}$, a finite-dimensional subspace $H_M$ of $H$, with associated orthogonal projection $P_M$.

DEFINITION 2.4 For any $M \in \mathbb{N}$, we define $H_M$ the subspace of $H$ generated by $e_0, \ldots, e_M$, by

$$H_M = \text{Span}\,\{e_k; 0 \le k \le M\}$$

and $P_M \in \mathcal{L}(H)$ the orthogonal projection onto $H_M$: for any $x = \sum_{k=0}^{+\infty} x_k e_k \in H$,

$$P_M x = \sum_{k=0}^{M} x_k e_k.$$

The domain $D(A)$ of $A$ is equal to $D(A) = \left\{x = \sum_{k=0}^{+\infty} x_k e_k \in H, \sum_{k=0}^{+\infty} (\lambda_k)^2 |x_k|^2 < +\infty\right\}$. More generally, fractional powers of $-A$, are defined for $\alpha \in [0,1]$:

$$(-A)^\alpha x = \sum_{k=0}^{\infty} \lambda_k^\alpha x_k e_k \quad \in H,$$

with domains $D\left((-A)^\alpha\right) = \left\{x = \sum_{k=0}^{+\infty} x_k e_k \in H, |x|_\alpha^2 = \sum_{k=0}^{+\infty} (\lambda_k)^{2\alpha} |x_k|^2 < +\infty\right\}$. In particular, for $\alpha = 0$, $|\cdot|_0 = |\cdot|_H$ is the norm in the Hilbert space $H$.

EXAMPLE 2.5 In the case when $A$ is the Laplace operator with homogeneous Dirichlet boundary conditions on $H = L^2(\mathcal{D})$, then $D((-A)^{1/2}) = H_0^1(\mathcal{D})$ and $D(A) = H_0^1(\mathcal{D}) \cap H^2(\mathcal{D})$.

For $\alpha \in [0,1]$, it is also possible to define spaces $D\left((-A)^{-\alpha}\right)$ and operators $(-A)^{-\alpha}$, with norm denoted by $|\cdot|_{-\alpha}$; in particular, when $x = \sum_{k=0}^{+\infty} x_k e_k \in H$, then $(-A)^{-\alpha} x = \sum_{k=0}^{+\infty} \lambda_k^{-\alpha} x_k e_k$ and $|x|_{-\alpha}^2 = \sum_{k=0}^{+\infty} (\lambda_k)^{-2\alpha} |x_k|^2$.

The semigroup $\left(e^{tA}\right)_{t\geq 0}$ is defined by the Hille–Yosida theorem—see Brézis (1994). The following formula holds true: for all $x = \sum_{k=0}^{+\infty} x_k e_k \in H$ and all $t \geq 0$,

$$e^{tA}x = \sum_{k=0}^{+\infty} e^{-\lambda_k t} x_k e_k. \tag{2.2}$$

For any $t \geq 0$, $e^{tA}$ is a continuous linear operator in $H$, with operator norm $|e^{tA}|_{\mathcal{L}(H)} = e^{-\lambda_0 t}$. The semigroup $\left(e^{tA}\right)_{t\geq 0}$ is used to define the solution $Z(t) = e^{tA}z$ of the linear Cauchy problem

$$\frac{\mathrm{d}Z(t)}{\mathrm{d}t} = AZ(t) \quad \text{with} \quad Z(0) = z.$$

To define solutions of semilinear equations, we use a mild formulation (Duhamel principle).

The semigroup $(e^{tA})_{t\geq 0}$ enjoys smoothing properties that are often used in this work. Using (2.2), in particular one obtains the following results.

Proposition 2.6 Under Assumptions 2.2, for any $\sigma \in [0, 1]$, there exists $C_\sigma \in (0, +\infty)$ such that

(1) for all $t > 0$ and $x \in H$,

$$|e^{tA}x|_\sigma \leq C_\sigma t^{-\sigma} e^{-\lambda_0/2t} |x|_H;$$

(2) for all $0 < s < t$ and $x \in H$ (respectively, $x \in D\left((-A)^\sigma\right)$),

$$|e^{tA}x - e^{sA}x|_H \leq C_\sigma \frac{(t-s)^\sigma}{s^\sigma} e^{-\lambda_0/2s} |x|_H \quad \left(\text{respectively,} \quad |e^{tA}x - e^{sA}x|_H \leq C_\sigma (t-s)^\sigma e^{-\lambda_0/2s} |x|_\sigma\right).$$

2.2.2 *The nonlinear coefficient.* First, the nonlinear coefficient $F$ is assumed to satisfy regularity conditions (Assumptions 2.7). Examples of coefficients $F$ that satisfy these conditions are given in Example 2.8.

Assumptions 2.7 The function $F : H \to H$ is assumed to be Lipschitz continuous with Lipschitz constant denoted by $L_F$.

For all $M \in \mathbb{N}$, define $F_M : H_M \to H_M$, such that $F_M(x) = P_M\big(F(x)\big)$ for all $x \in H_M$. We assume that for all $M \in \mathbb{N}$ the function $F_M$ is twice continuously differentiable, with the following uniform bounds on the second-order derivative, uniformly with respect to $M \in \mathbb{N}$: there exists a parameter $\eta \in [0, 1[$ and $C_\eta \in (0, +\infty)$ such that for all $M \in \mathbb{N}$, $x \in H_M$ and $h, k \in H_M$,

$$|D^2F_M(x) \cdot (h, k)|_{-\eta} \leq C_\eta |h|_0 |k|_0, \quad |D^2F_M(x) \cdot (h, k)|_0 \leq C_\eta |h|_\eta |k|_0.$$

Observe that for all $M \in \mathbb{N}$, the Lipschitz constant of $F_M$ is bounded from above by $L_F$.

EXAMPLE 2.8 Functions $F$ defined below satisfy Assumptions 2.7.

(1) Function $F : H \to H$ is of class $\mathcal{C}^2$, with bounded first- and second-order derivatives (consider $\eta = 0$).
(2) Function $F$ is a **Nemytskii** operator, with $H = L^2(0,1)$: $F(x)(\cdot) = f\big(x(\cdot)\big)$, for some $f : \mathbb{R}^2 \to \mathbb{R}$ of class $\mathcal{C}^2$ with bounded first- and second-order derivatives. When $A$ is given as in Example 2.3, the conditions are satisfied for $\eta > 1/4$.

Let us now introduce two usual sufficient conditions that imply ergodicity of the SPDE (1.1); see Section 4 for details and references.

ASSUMPTIONS 2.9 (Dissipativity) Assume that at least one of the following conditions is satisfied:

(i) $F$ is bounded (weak dissipativity condition);
(ii) $L_F < \lambda_0$ (strict dissipativity condition).

Note that under Assumptions 2.9, there exists $c, C \in (0, +\infty)$ such that for any $x \in D(A)$,

$$\langle Ax + F(x), x\rangle \leq -c|x|^2 + C. \tag{2.3}$$

This inequality is sufficient to ensure the ergodicity result (see Section 4) for the continuous time process $X$ defined by (1.1). On the one hand, the strict dissipativity condition provides ergodicity for processes that are discrete in time; on the other hand, it is not known whether this ergodicity holds true under the weak dissipativity condition. Note that the approximation result, Theorem 5.1, holds true under both conditions.

Observe that due to Assumptions 2.7, $F_M$ satisfies Assumptions 2.9 for all $M \in \mathbb{N}$.

## 2.3 *The cylindrical Wiener process and stochastic integration in H*

We now recall the definition of the cylindrical Wiener process and of the stochastic integral on a separable infinite-dimensional Hilbert space $H$. For more details, see Da Prato & Zabczyck (1992).

Assume that a filtered probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t\geq 0}, \mathbb{P})$ is given. The definition of the cylindrical Wiener process $\big(W(t)\big)_{t\in\mathbb{R}^+}$ on $H$ requires that

- a complete orthonormal system of $H$, denoted by $(q_i)_{i\in\mathbb{N}}$, and
- a family $(\beta_i)_{i\in\mathbb{N}}$ of independent real Wiener processes with respect to the filtration $\big(\mathcal{F}_t\big)_{t\geq 0}$

are given. Then set

$$W(t) = \sum_{i\in\mathbb{N}} \beta_i(t) q_i. \tag{2.4}$$

This series does not converge in $H$; however, $W(t)$ is an element of $D\left((-A)^{-1/4-\kappa}\right)$ when $\kappa > 0$, for all $t \geq 0$. Note that the law of the process does not depend on the choice of $(q_i)_{i\in\mathbb{N}}$ and $(\beta_i)_{i\in\mathbb{N}}$.

A bounded linear operator $\Psi : H \to H$ is said to be Hilbert–Schmidt when

$$|\Psi|^2_{\mathcal{L}_2(H,H)} = \sum_{k=0}^{+\infty} |\Psi(q_k)|_H^2 < +\infty;$$

the definition of the norm $|\Psi|_{\mathcal{L}_2(H,H)}$ does not depend on the orthonormal basis $(q_k)$ of $H$.

The stochastic integral $\int_0^t \Psi(s)\,\mathrm{d}W(s)$ is defined in $H$ for predictable processes $\Psi$ with values in $\mathcal{L}_2(H,H)$ such that $\int_0^t |\Psi(s)|^2_{\mathcal{L}_2(H,H)}\,\mathrm{d}s < +\infty$ a.s. Moreover, when $\Psi \in L^2(\Omega \times [0,t]; \mathcal{L}_2(H,H))$, the following two properties hold:

$$\mathbb{E}\left|\int_0^t \Psi(s)\,\mathrm{d}W(s)\right|_H^2 = \mathbb{E}\int_0^t |\Psi(s)|^2_{\mathcal{L}_2(H,H)}\,\mathrm{d}s \quad \text{(Itô isometry)}, \qquad \mathbb{E}\int_0^t \Psi(s)\,\mathrm{d}W(s) = 0.$$

Under assumptions above, solutions to the equation (1.1) are well defined in a mild sense. The following result is standard—see Da Prato & Zabczyck (1992).

Proposition 2.10 For every $T > 0$, $x \in H$, equation (1.1) admits a unique mild solution $X \in \mathrm{L}^2(\Omega, \mathcal{C}([0,T], H))$:

$$X(t) = e^{tA}x + \int_0^t e^{(t-s)A}F(X(s))\,\mathrm{d}s + \int_0^t e^{(t-s)A}\,\mathrm{d}W(s), \tag{2.5}$$

where $W^A(t) = \int_0^t e^{(t-s)A}\,\mathrm{d}W(s)$ is the stochastic convolution.

## 3. Definition of the discretization schemes

We consider approximations in time and space of the process $X$. In this section, we introduce the corresponding schemes: a finite element approximation for discretization in space (Section 3.1) and a semiimplicit Euler scheme for discretization in time (Section 3.3). We also discuss a spectral Galerkin discretization (Section 3.2), which is an important tool in the analysis below.

### 3.1 *Discretization in space: finite element approximation*

We use the same framework as in Debussche & Printems (2009) and Andersson & Larsson (2016). For general references on finite element methods, see for instance Ciarlet (2002) and Ern & Guermond (2004).

Let $(V_h)_{h\in(0,1)}$ be a family of spaces of continuous piecewise linear functions corresponding to a (quasiuniform) family of meshes in $\mathcal{D} = (0,1)$ such that $V_h \subset H_0^1(\mathcal{D}) = D\left((-A)^{1/2}\right)$ – 0 and 1 should be included as nodes in the partition of $[0,1]$. The parameter $h$ denotes the mesh size, which is the length of the largest subinterval in the partition.

Let $P_h : H \to V_h$ denote the orthogonal projection onto the finite-dimensional space $V_h$. According to the context, we also consider $P_h$ as a linear operator in $\mathcal{L}(H)$, since $V_h \subset H$.

Finally we define the approximation of the operator $A$: it is a linear operator $A_h \in \mathcal{L}(V_h)$.

Definition 3.1 The linear operator $A_h : V_h \to V_h$ is defined such that the following variational equality holds: for any $x_h \in V_h$ and $y_h \in V_h$,

$$\langle A_h x_h, y_h\rangle = -\langle(-A)^{1/2}x_h, (-A)^{1/2}y_h\rangle.$$

We recall a few important properties of the operator $A_h$.

Proposition 3.2 For all $h \in (0,1)$, $-A_h$ is symmetric and positive definite. If $N_h$ is the dimension of $V_h$, we denote by $(e_i^h)_{i=0}^{N_h-1} \subset V_h$ an orthonormal eigenbasis corresponding to $-A_h$ with corresponding eigenvalues $0 < \lambda_0^h \leq \lambda_1^h \leq \cdots \leq \lambda_{N_h-1}^h$. Then for any $h \in (0,1)$, we have $\lambda_0^h \geq \lambda_0$.

Indeed, since $V_h \subset D\left((-A)^{1/2}\right)$,

$$\lambda_0 = \inf_{v,u \in D((-A)^{1/2}), |u|=|v|=1} \langle -Au, v\rangle \leq \inf_{u,v\in V_h, |u|=|v|=1} \langle -Au, v\rangle = \inf_{u,v\in V_h, |u|=|v|=1} \langle -A_h u, v\rangle = \lambda_0^h.$$

For any $h \in (0,1)$, $A_h$ generates a semigroup on $V_h$, which is denoted $(e^{tA_h})_{t\in\mathbb{R}^+}$. The definition of fractional powers $(-A_h)^\alpha$ of $-A_h$, for any $\alpha \in [-1,1]$, is straightforward: for all $x^h = \sum_{i=0}^{N_h-1} x_i^h e_i^h \in V_h$, we have

$$e^{tA_h} x^h = \sum_{i=0}^{N_h-1} e^{-\lambda_i^h t} x_i^h e_i^h, \quad (-A_h)^\alpha x^h = \sum_{i=0}^{N_h-1} (\lambda_i^h)^\alpha x_i^h e_i^h.$$

The regularization estimates of Proposition 2.6 are then easily generalized to these semigroups; moreover, bounds are uniform with respect to the mesh size $h \in (0,1)$.

We focus now on the approximations of SPDEs—seen as equations in the Hilbert space $H$—with equations in finite-dimensional spaces $V_h$.

We consider the spatially semidiscrete approximation of (1.1): $(X^h(t))_{t\in\mathbb{R}^+}$ is a process, taking values in $V_h$, such that

$$\mathrm{d}X^h(t) = A_h X^h(t)\,\mathrm{d}t + F^h(X^h(t))\,\mathrm{d}t + P_h \mathrm{d}W(t), \quad X^h(0) = P_h x = P_h X_0, \tag{3.1}$$

where the nonlinear coefficient $F^h : V_h \to V_h$ is defined by $F^h(x) = P_h(F(x))$ for all $x \in V_h$.

Note that the regularity properties of Assumptions 2.7 and the dissipativity inequality (2.3) are satisfied if $A$ (respectively, $F$) is replaced with $A_h$ (respectively, $F^h$).

Equation (3.1) admits a unique mild solution: for all $0 \leq t \leq T$,

$$X^h(t) = e^{tA_h} P_h x + \int_0^t e^{(t-s)A_h} F^h(X^h(s))\,\mathrm{d}s + \int_0^t e^{(t-s)A_h} P_h\,\mathrm{d}W(s). \tag{3.2}$$

Notice that the stochastic integral in (3.2) is always well defined, since for any $h \in (0,1)$ the linear operator $P_h$ has finite rank. The noise process $P_h W$ has covariance operator $P_h$ as an $H$-valued process; seen as a process in $V_h$, it is a standard $N_h$-dimensional Wiener process—as is easily seen by expanding $W$ in a complete orthonormal system $(q_i)_{i\in\mathbb{N}}$ with $q_i = e_i^h$ for $0 \leq i \leq N_h - 1$.

To be able to state a convergence result of $X^h$ to $X$, and to give an order of convergence, we now recall some important results—see Andersson & Larsson (2016) for details.

Proposition 3.3 (i) An equivalence of norms holds true: there exist two constants $c, C \in (0,+\infty)$, such that for any $h \in (0,1)$, any $\alpha \in [-1/2, 1/2]$ and any $x^h \in V_h$,

$$c|(-A_h)^\alpha x^h| \leq |(-A)^\alpha x^h| \leq C|(-A_h)^\alpha x^h|. \tag{3.3}$$

Moreover, for any $h \in (0,1)$, $\alpha \in [-1/2, 1/2]$ and $x \in D((-A)^\alpha)$,

$$|(-A_h)^\alpha P_h x| \le C|(-A)^\alpha x|. \tag{3.4}$$

(ii) Let us denote by $R_h$ the so-called Ritz projector, defined as the orthogonal projection onto $V_h$ in $D((-A)^{1/2})$. We have the identity $R_h = (-A_h)^{-1}P_h(-A)$ on $D(A)$, and

$$\left|(-A)^{s/2}(I-R_h)(-A)^{-r/2}\right|_{\mathcal{L}(H)} \le C_{r,s}h^{r-s} \quad \text{for all } 0 \le s \le 1 \le r \le 2. \tag{3.5}$$

(iii) For $P_h$, we have the following error estimate:

$$\left|(-A)^{s/2}(I-P_h)(-A)^{-r/2}\right|_{\mathcal{L}(H)} \le C_{r,s}h^{r-s} \quad \text{for all } 0 \le s \le 1 \text{ and } 0 \le s \le r \le 2. \tag{3.6}$$

The following result is a consequence of Proposition 3.3 (for an alternative proof, see Kovacs *et al.* (2012, equation (4.16))).

PROPOSITION 3.4 For all $\kappa \in (0, 1/2)$, the linear operator $P_h(-A_h)^{-1/2-\kappa}P_h$ is continuous from $H$ to $H$, self-adjoint and semidefinite positive. Moreover,

$$\sup_{0<h<1} \operatorname{Tr}\left(P_h(-A_h)^{-1/2-\kappa}P_h\right) < +\infty.$$

Recall that for $M, N \in \mathcal{L}(H)$, if $N$ is symmetric and semidefinite positive then

$$|\operatorname{Tr}(MN)| \le |M|_{\mathcal{L}(H)}\operatorname{Tr}(N).$$

*Proof.* The operator is well defined on $H$, and clearly self-adjoint, since $(-A_h)^{-1/2-\kappa} \in \mathcal{L}(V_h)$ is symmetric.

Now from Proposition 3.3 (i), the following linear operators are defined and continuous on $H$ (since $\kappa \le 1/2$): $(-A)^\kappa(-A_h)^{-\kappa}P_h$ and $(-A)^{1/2}P_h(-A_h)^{-1/2}P_h$; their norm are uniformly bounded with respect to $h$.

By duality, the operator $P_h(-A_h)^{-1/2}P_h(-A)^{1/2}$ is well defined on $H$—by unique continuous extension from the dense subspace $D((-A)^{1/2})$—and it has the same norm as $(-A)^{1/2}P_h(-A_h)^{-1/2}P_h$.

Finally, we write that for any $0 < h < 1$,

$$\begin{aligned}
\operatorname{Tr}\left(P_h(-A_h)^{-1/2-\kappa}P_h\right) &= \operatorname{Tr}\left((P_h(-A_h)^{-1/2}P_h(-A)^{1/2})(-A)^{-1/2-\kappa}((-A)^\kappa(-A_h)^{-\kappa}P_h)\right)\\
&\le \left|P_h(-A_h)^{-1/2}P_h(-A)^{1/2}\right|_{\mathcal{L}(H)}\operatorname{Tr}\left((-A)^{-1/2-\kappa}\right)\left|(-A)^\kappa(-A_h)^{-\kappa}P_h\right|_{\mathcal{L}(H)}\\
&\le C\operatorname{Tr}\left((-A)^{-1/2-\kappa}\right).
\end{aligned}$$

□

We recall that, for all $T \in (0, +\infty)$, $X^h(T)$ converges to $X(T)$, with *strong* order of convergence $1/2$, in $L^1(\Omega)$ (see for instance Kovacs *et al.*, 2010 and Kruse, 2013 for the semilinear case, and Yan, 2005 for the linear case) and *weak* order of convergence 1, in distribution when tested against admissible test functions $\phi$ (see Andersson & Larsson, 2016): for all $r \in (0, 1/2)$,

$$\mathbb{E}|X^h(T) - X(T)| = \mathcal{O}\left(h^{1/2-r}\right), \quad |\mathbb{E}\phi(X^h(T)) - \mathbb{E}\phi(X(T))| = \mathcal{O}\left(h^{1-r}\right).$$

To conclude this part, we introduce some nice notation.

Definition 3.5 For $h = 0$, set $X^0 = X$, as well as $V_0 = H$, $A_0 = A$, $P_0 = \mathrm{Id}_H$.

### 3.2 *Another discretization in space: spectral Galerkin projection*

A tool in our proof will be an additional finite-dimensional projection onto the subspaces $H_M$. This approximation allows us to justify rigorously the computations; even if the process $X^h$ takes values in a finite-dimensional subspace of $H$, it is convenient to prove some estimates with a process taking values in finite-dimensional subspaces that are left invariant by the action of $A$ and of the noise. We define here the corresponding approximating processes and give a few important convergence properties.

Let $M \in \mathbb{N}$. Following Definition 2.4, consider the equation

$$\mathrm{d}X^{(M)}(t) = AX^{(M)}(t)\,\mathrm{d}t + F_M(X^{(M)}(t))\,\mathrm{d}t + P_M\,\mathrm{d}W(t), \quad X^{(M)}(0) = P_M x \tag{3.7}$$

in the finite-dimensional subspace $H_M$, where $F_M = P_M \circ F$. The process $W^{(M)} = P_M W$ is a standard Wiener process with values in $H_M$.

For any $T \in (0, +\infty)$, there is a unique mild solution, taking values in $H_M \subset H$:

$$X^{(M)}(t) = e^{tA} P_M x + \int_0^t e^{(t-s)A} F_M(X^{(M)}(s))\,\mathrm{d}s + \int_0^t e^{(t-s)A} P_M\,\mathrm{d}W(s).$$

The proof of the following inequality is straightforward:

$$|(I - P_M)A^{-r}|_{\mathcal{L}(H)} \le C_r \lambda_{M+1}^{-r}, \quad 0 \le r \le 1. \tag{3.8}$$

Then, for all $T \in (0, +\infty)$, $X^{(M)}(T)$ converges to $X(T)$, such that for all $r \in (0, 1/4)$,

$$\mathbb{E}|X^{(M)}(T) - X(T)| = \mathcal{O}\big(\lambda_{M+1}^{-1/4+r}\big), \quad |\mathbb{E}\phi(X^{(M)}(T)) - \mathbb{E}\phi(X(T))| = \mathcal{O}\big(\lambda_{M+1}^{-1/2+r}\big).$$

Indeed, the projections $P_M$ satisfy the estimates of Proposition 3.3 with $h = \lambda_{M+1}^{-1/2}$; see Kruse (2013, Example 3.4).

It is useful to introduce notation for when $M = \infty$.

Definition 3.6 For $M = \infty$, we set $X^{(\infty)} = X$, as well as $H_\infty = H$ and $P_\infty = \mathrm{Id}r_H$.

### 3.3 *Discretization in time*

For each fixed mesh size $h \in (0, 1)$, and for $h = 0$, we now define a time approximation of the process $X^h$: denoting by $\tau > 0$ a time step, we use a semiimplicit Euler scheme to define, for $k \in \mathbb{N}$,

$$X_{k+1}^h(\tau, x) = X_k^h(\tau, x) + \tau A_h X_{k+1}^h(\tau, x) + \tau P_h F(X_k^h(\tau, x)) + \sqrt{\tau} P_h \chi_{k+1},$$
$$X_0^h(\tau, x) = P_h x,$$

where $\chi_{k+1} = \frac{1}{\sqrt{\tau}}(W((k+1)\tau) - W(k\tau))$. Note that even if $\chi_{k+1}$ does not take values in the Hilbert space $H$, the stochastic term $P_h \chi_{k+1}$ can be given a meaning in $H$ in a straightforward way: indeed, choose $q_i = e_i^h$ for $0 \le i \le N_h$ in the cylindrical Wiener process expansion (2.4).

We often omit the dependence of $X_k^h$ on the time step $\tau$ and on the initial condition $P_h x$.

A rigorous formulation is given by

$$X_{k+1}^h = S_{\tau,h} X_k^h + \tau S_{\tau,h} P_h F(X_k^h) + \sqrt{\tau} S_{\tau,h} P_h \chi_{k+1}, \tag{3.9}$$

where the linear operator $S_{\tau,h}$ on $V_h$ is defined by

$$S_{\tau,h} = (I - \tau A_h)^{-1}. \tag{3.10}$$

When $h = 0$, the process is well defined in $H$, since it is easily checked that $S_{\tau,0}$ is a Hilbert–Schmidt operator on $H$. When $h > 0$, it is well defined in the finite-dimensional space $V_h$.

For the analysis of the convergence of the scheme, we use regularization estimates on the discrete-time semigroup $(S_{\tau,h}^j)_{j\in\mathbb{N}}$ for $\tau > 0$ and $h \geq 0$—compare with Proposition 2.6.

LEMMA 3.7 For any $0 \leq \kappa \leq 1$, $h \in [0, 1)$ and $j \geq 1$,

$$|(-A_h)^{1-\kappa} S_{\tau,h}^j P_h|_{\mathcal{L}(H)} \leq \frac{1}{(j\tau)^{1-\kappa}} \frac{1}{(1+\lambda_0\tau)^{j\kappa}}.$$

Moreover,

$$|(-A_h)^{\beta} S_{\tau,h}^j P_h|_{\mathcal{L}(H)} \leq \frac{\beta^\beta}{(j\tau)^\beta}, \quad \beta \geq 1,\ j \geq \beta,$$

$$|(-A_h)^{-\beta} (S_{\tau,h} - I) P_h|_{\mathcal{L}(H)} \leq 2\tau^\beta, \quad 0 \leq \beta \leq 1.$$

For a proof of the first estimate, we refer to Thomée (2006, Lemma 7.3). The other estimates are obtained using similar arguments.

REMARK 3.8 We often use the following expression (discrete mild formulation) for $X_k^h$:

$$X_k^h = S_{\tau,h}^k P_h x + \tau \sum_{l=0}^{k-1} S_{\tau,h}^{k-l} P_h F(X_l^h) + \sqrt{\tau} \sum_{l=0}^{k-1} S_{\tau,h}^{k-l} P_h \chi_{l+1}. \tag{3.11}$$

The following expression is also useful: if $l_s = \lfloor \frac{s}{\tau} \rfloor$ (where $\lfloor \cdot \rfloor$ denotes the integer part function),

$$\sqrt{\tau} \sum_{l=0}^{k-1} S_{\tau,h}^{k-l} P_h \chi_{l+1} = \int_0^{t_k} S_{\tau,h}^{k-l_s} P_h \, \mathrm{d}W(s). \tag{3.12}$$

For $h \in (0, 1)$, we finally introduce the following processes: for $0 \leq k \leq m-1$ and $t_k \leq t \leq t_{k+1}$,

$$\tilde{X}^h(t) = X_k^h + \int_{t_k}^t [A_h S_{\tau,h} X_k^h + S_{\tau,h} P_h F(X_k^h)] \, \mathrm{d}s + \int_{t_k}^t S_{\tau,h} P_h \, \mathrm{d}W(s). \tag{3.13}$$

The process $(\tilde{X}^h(t))_{t\in\mathbb{R}^+}$ is an interpolation in time of the numerical solution $(X_k^h)_{k\in\mathbb{N}}$ defined by (3.9): $\tilde{X}^h(t_k) = X_k^h$.

### 3.4 *A priori bounds on moments*

We give bounds on moments of $(X(t))_{t\geq 0}$, $(X^h(t))_{t\in\mathbb{R}^+}$ and $(X_k^h)_{k\in\mathbb{N}}$.

Note that the constants are uniform with respect to $h \in [0,1)$ and $\tau \in (0,1)$.

LEMMA 3.9 For any $p \geq 1$, there exists a constant $C_p \in (0,+\infty)$ such that for every $h \in [0,1)$, $t \geq 0$ and $x \in H$,

$$\mathbb{E}|X^h(t,x)|^p \leq C_p(1+|x|^p).$$

LEMMA 3.10 For any $p \geq 1$, $\tau_0 > 0$, there exists a constant $C \in (0,+\infty)$ such that for every $h \in [0,1)$, $0 < \tau \leq \tau_0$, $k \in \mathbb{N}$, $t \geq 0$ and $x \in H$,

$$\mathbb{E}|X_k^h|^p \leq C(1+|x|^p) \quad \text{and} \quad \mathbb{E}|\tilde{X}^h(t)|^p \leq C(1+|x|^p).$$

Since the arguments are standard, in the interest of conciseness, we omit the proofs of Lemmas 3.9 and 3.10. For $h = 0$, we refer to Bréhier (2014, Lemma 4.2) under Assumptions 2.9(i), and to Bréhier & Vilmart (2015, Proposition 3.2) for a similar study under Assumptions 2.9(ii). For $h > 0$, we refer to the Ph.D. thesis of the second named author, Kopec (2014, Lemma 3.10).

## 4. Asymptotic behavior of the processes and invariant laws

The objective of this section is to present results about the behavior of the continuous- and discrete-time processes, either discretized in space by the finite element method, the spectral Galerkin method, or not (i.e., for $H$-valued processes). In Section 4.1, we prove the existence of invariant distributions for all of them, under Assumptions 2.9. Then in Section 4.2, we discuss the uniqueness (i.e., ergodicity), and we state the main differences between the different settings, and between the weak and the strict dissipativity conditions.

We state precise and general results, but we omit the proofs and give only the essential arguments. We refer to Da Prato & Zabczyck (1996) and Debussche (2013), and references therein, for treatments of the asymptotic behavior of SPDEs.

The main result of this section is the following.

PROPOSITION 4.1 Let $h \in [0,1)$, $M \in \mathbb{N}\cup\{\infty\}$ and $\tau \in (0,1)$.

(1) **Existence of invariant laws.** Under Assumptions 2.9, the processes $t \in \mathbb{R}^+ \mapsto X^h(t)$, $t \in \mathbb{R}^+ \mapsto X^{(M)}(t)$ and $k \in \mathbb{N} \mapsto X_k^h(\tau)$ admit (at least) an invariant distribution.

(2) **Uniqueness, continuous-time processes.** Under Assumptions 2.9, the process $t \in \mathbb{R}^+ \mapsto X^h(t)$ (respectively, $t \in \mathbb{R}^+ \mapsto X^{(M)}(t)$) admits a unique invariant law denoted by $\mu^h$ (respectively, $\mu^{(M)}$). Moreover, convergence is exponentially fast: there exist $c, C \in (0,+\infty)$ (independent of $h$ and $M$) such that for any bounded test function $\phi : H \to \mathbb{R}$, any $t \geq 0$ and any initial conditions $x^h \in V_h$ and $x^{(M)} \in H_M$,

$$\left|\mathbb{E}\phi(X^h(t,x^h)) - \int_{V_h} \phi \, \mathrm{d}\mu^h\right| \leq C\|\phi\|_\infty(1+|x^h|^2)e^{-ct},$$

$$\left|\mathbb{E}\phi(X^{(M)}(t,x^{(M)})) - \int_{H_M} \phi \, \mathrm{d}\mu^{(M)}\right| \leq C\|\phi\|_\infty(1+|x^{(M)}|^2)e^{-ct}.$$

(3) **Uniqueness, discrete-time processes.** Under the strict dissipativity condition, Assumptions 2.9(ii), the process $k \in \mathbb{N} \mapsto X_k^h(\tau)$ admits a unique invariant law.

Note that choosing $h = 0$ and $M = \infty$, the SPDE (1.1) is ergodic, and that $\mu = \mu^0 = \mu^{(\infty)}$.

### 4.1 *Existence of invariant distributions*

The main tool to prove existence of invariant laws is a compactness argument, namely the well-known Krylov–Bogoliubov criterion—see Da Prato & Zabczyck (1996, Section 3.1).

First, the semigroup associated with each Markov process we consider (denoted by $\mathcal{X}$) satisfies the Feller property at all (continuous or discrete) times $t$: if $Q_t\phi(x) = \mathbb{E}[\phi(\mathcal{X}(t))|\mathcal{X}(0) = x]$, then $Q_t\phi$ is continuous when $\phi$ is assumed bounded and continuous.

Second, the required tightness property (Da Prato & Zabczyck, 1996, Corollary 3.1.2) comes from two facts: $D\left((-A)^\gamma\right)$ is compactly embedded in $H$ when $\gamma > 0$; and when $\gamma < 1/4$ moments of the processes are controlled (the proof is standard and thus omitted).

LEMMA 4.2 For any $0 < \gamma < 1/4$, $\tau > 0$ and any $x \in H$, there exist $C(\gamma,\tau,x), C(\gamma,x) > 0$ such that for every $h \in [0,1)$, $m \geq 1$ and $t \geq 1$,

$$\mathbb{E}|X_m^h(\tau,x)|_\gamma^2 \leq C(\gamma,\tau,x) \quad \text{and} \quad \mathbb{E}|X^h(t,x)|_\gamma^2 \leq C(\gamma,x).$$

### 4.2 *Uniqueness of the invariant distribution*

Let us first assume that the strict dissipativity condition, Assumptions 2.9(ii), is satisfied. Consider then two initial conditions $x_1, x_2 \in V_h$; for each realization of $\omega \in \Omega$, define $X^h(\cdot,x_1)$ and $X^h(\cdot,x_2)$ (respectively, $X_\cdot^h(\tau,x_1)$ and $X_\cdot^h(\tau,x_2)$) with (3.1) (respectively, (3.9)), with the same driving Wiener process. Then for all discretization parameters $\tau > 0$, $h \geq 0$ and $t \geq 0$, $k \geq 0$, by straightforward computations (using Gronwall's lemma),

$$\begin{aligned}|X^h(t,x_1) - X^h(t,x_2)| &\leq \exp\big(-(\lambda_0 - L_F)t\big)|x_1 - x_2|,\\ |X_k^h(\tau,x_1) - X_k^h(\tau,x_2)| &\leq \exp\left(-\frac{\lambda_0 - L_F}{1+\lambda_0\tau}t\right)|x_1 - x_2|.\end{aligned}$$

Then the process $t \mapsto X^h(\cdot)$ (respectively, $k \mapsto X_k^h(\tau)$) admits a unique invariant law. The same reasoning applies to the process $X^{(M)}$ defined by (3.7), obtained by spectral Galerkin discretization.

Now consider that the weak dissipativity condition, Assumptions 2.9(i), is satisfied. Then the uniqueness of the invariant distribution holds true for continuous-time processes $X^h$ and $X^{(M)}$ for all $h \in [0,1)$ and $M \in \mathbb{N} \cup \{\infty\}$. The proof is based on Doob's theorem—see Da Prato & Zabczyck (1996, Proposition 4.1.1 and and Theorem 4.2.1)—which requires two arguments: a regularizing effect (the semigroup satisfies the strong Feller property: for all $t > 0$ and bounded measurable $\phi$, then $Q_t\phi$ is continuous) and an irreducibility property (support of invariant distributions).

Note that these two arguments heavily rely on the choice of space-time white noise. Moreover, the same reasoning is not sufficient to deal with discrete-time processes: to our knowledge, the uniqueness of the invariant distribution for $k \in \mathbb{N} \mapsto X_k^h(\tau)$, for arbitrary $\tau \in 0$ and $h \in [0,1)$, is not known.

Exponential convergence to equilibrium for the continuous-time processes $X^h$ and $X^{(M)}$ is proved using coupling arguments. We refer to Doeblin (1938), Lindvall (1992), Meyn & Tweedie (2009) for the description of this approach for Markov chains, and for instance to Kuksin & Shirikyan (2001), Mattingly (2002), Mueller (1993) for various applications to SPDEs.

The coupling approach yields the following result: for a proof, see Debussche *et al.* (2011, Section 6.1), and Debussche (2013).

PROPOSITION 4.3 There exist $c, C \in (0, +\infty)$ such that for any bounded test function $\phi : H \to \mathbb{R}$, any $t \geq 0$, any discretization parameters $h \in [0, 1)$ and $M \in \cup\{\infty\}$, and any initial conditions $x_1^h, x_2^h \in V_h$ and $x_1^{(M)}, x_2^{(M)} \in H_M$,

$$|\mathbb{E}\phi(X^h(t, x_1^h)) - \mathbb{E}\phi(X^h(t, x_2^h))| \leq C\|\phi\|_\infty(1 + |x_1^h|^2 + |x_2^h|^2)e^{-ct},$$
$$|\mathbb{E}\phi(X^{(M)}(t, x_1^{(M)})) - \mathbb{E}\phi(X^{(M)}(t, x_2^{(M)}))| \leq C\|\phi\|_\infty(1 + |x_1^{(M)}|^2 + |x_2^{(M)}|^2)e^{-ct}.$$

## 5. The convergence results

### 5.1 *Approximation of the invariant law $\mu$*

The main result of this article is the following Theorem 5.1. Definition and notation of admissible test functions are given in Section 2.1.

THEOREM 5.1 For any $0 < \kappa < 1/2$, $\tau_0$, there exists a constant $C > 0$ such that for any admissible test function $\phi$, $h \in (0, 1)$, $N \geq 1$, $x \in H$ and $0 < \tau \leq \tau_0$,

$$\left|\frac{1}{N}\sum_{m=0}^{N-1}\Big(\mathbb{E}\phi(X_m^h) - \overline{\phi}\Big)\right| \leq C \parallel \phi \parallel_{2,\infty} (1 + |x|^3)\left(1 + \frac{1}{(N\tau)^{1-\kappa}} + \frac{1}{N\tau}\right)\left(\tau^{1/2-\kappa} + h^{1-\kappa} + \frac{1}{N\tau}\right),$$

where $\overline{\phi} = \int_H \phi(z)\,\mathrm{d}\mu(z)$.

This result has a statistical interpretation: $\frac{1}{N}\sum_{m=0}^{N-1}\mathbb{E}\phi(X_m^h)$ is an *estimator* of the average $\overline{\phi} = \int_H \phi(z)\mu(\mathrm{d}z)$ of the admissible test function $\phi$ with respect to the invariant law $\mu$ of the SPDE. Theorem 5.1 gives an error bound on its *bias*.

Of the two factors in parentheses in the theorem, only the second one is important—the presence of the first one is for technical estimates which degenerate at time 0, whereas we are interested in the asymptotic behavior of the quantity. The main observation is that the orders of convergence with respect to $\tau$ and $h$ are given by the corresponding weak orders $1/2$ and 1 in the approximation of $X(T)$ for a fixed value of the final time $T < +\infty$—given in Debussche (2011) and Andersson & Larsson (2016). The aim of this article is to show how the corresponding error bounds are preserved asymptotically—under appropriate conditions. The additional term $\frac{1}{N\tau}$ corresponds to the bias introduced between the average in time and its limit when time increases.

Note that Theorem 5.1 yields approximation results when only one type (time or space) of discretization is applied.

As a fundamental consequence of Theorem 5.1, we obtain in Proposition 5.2 error bounds controlling the distance between the average of admissible test functions with respect to the (possibly nonunique) ergodic invariant laws of the discretized process and the invariant law of the SPDE.

PROPOSITION 5.2 For any $0 < \kappa < 1/2$, $\tau_0 > 0$, there exists a constant $C > 0$ such that the following holds: for any $0 < \tau < \tau_0$ and $h \in (0,1)$, assume that $\mu^{\tau,h}$ is an ergodic invariant law of $(X_k^h)_{k\in\mathbb{N}}$; then for any admissible test function $\phi$, we have

$$\left|\int_H \phi(z)\,\mathrm{d}\mu(z) - \int_{V_h} \phi(z)\,\mathrm{d}\mu^{\tau,h}(z)\right| \leq C \parallel \phi \parallel_{2,\infty} \left(\tau^{1/2-\kappa} + h^{1-\kappa}\right).$$

The proof of this result is straightforward: let $N$ go to $+\infty$ in Theorem 5.1, and use the convergence of the time average for the $\mu^{\tau,h}$ a.e. initial condition; see also Bréhier (2014). Instead of assuming that $\mu^{\tau,h}$ is ergodic, one can also assume that it has a finite third-order moment.

Recall that $\mu^{\tau,h}$ is unique under the strict dissipativity condition, Assumptions 2.9(ii), but may be nonunique under the weak dissipativity condition, Assumptions 2.9(i).

To conclude this section, note that Theorem 5.1 and Proposition 5.2 also hold for the process $X^{(M)}$ defined by (3.7) and its unique invariant law $\mu^{(M)}$. In particular, the convergence estimate of Proposition 5.2 then reads

$$\left|\int_H \phi\,\mathrm{d}\mu - \int_{H_M} \phi_M \mathrm{d}\mu^{(M)}\right| \leq C_\kappa \frac{1}{\lambda_{M+1}^{1/2-\kappa}}.$$

### 5.2 *Extensions*

The aim of this section is to discuss possible straighforward extensions of the results above under different settings and open questions that require further investigations.

First, grant the strict dissipativity condition, Assumptions 2.9(ii). In this situation, it is straightforward to extend the results when the additive Gaussian space-time white noise is replaced with additive Gaussian white noise in time but colored in space. The orders of convergence then depend on the properties of the covariance kernel. As a consequence, SPDEs with a higher-dimensional space variable $\xi \in \mathcal{D}$, $\mathcal{D} \subset \mathbb{R}^d$ with $d > 1$ can also be handled.

Note that the extension for colored noise under the weak dissipativity condition, Assumptions 2.9(i), is challenging: it is then unclear and difficult to give ergodicity results (nondegeneracy of the noise is required) and in particular, exponential convergence properties.

Second, we have restricted our study to the additive noise case, even if we acknowledge that from a practical point of view the case of finite element discretization of the stochastic part of the solution is not a trivial problem. So far, to our knowledge, there is no general satisfactory weak approximation result when the noise is multiplicative. Indeed, as soon as the equation is discretized either in time—Debussche (2011)—or in space—Andersson & Larsson (2016)—the diffusion coefficients must satisfy strict conditions if one wants to obtain the expected weak order of convergence: they should be decomposed as the sum of a continuous affine function, and another function such that the second-order derivative is controlled with respect to a very weak norm—namely, the norm associated with a negative power of the linear operator. Moreover, the treatment of such noise requires lengthier computations. We could do so here by adding our argument to those in Debussche (2011) and Andersson & Larsson (2016), but this would result only in hiding the main ideas of our work.

Finally, note that contrary to Mattingly *et al.* (2010), we have not studied the *statistical error*. In Mattingly *et al.* (2010), two more error bounds are proved: first in the mean-square sense, and then in an almost sure statement—thanks to a Borel–Cantelli-type argument. We have not been able to treat these questions in the SPDE context. We claim that it is for the following reason. The right order of convergence with respect to $\tau$ in Theorem 5.1 is obtained thanks to an appropriate integration by parts formula—as explained in Section 1; the study of the mean-square error—now in a stronger sense—implies that the use of such a technique seems impossible. To generalize the results of Mattingly *et al.* (2010) in the infinite-dimensional setting, new arguments should be found.

## 6. Description of the proof

The aim of this section is to introduce the main objects required to prove the error estimate in Theorem 5.1. This section is organized as follows. We first explain in Section 6.1 why estimates for projections onto $H_M$, uniform over $M \in \mathbb{N}$, are sufficient. We then present in Section 6.2 properties of the solution of Poisson equation in $H_M$. Section 6.3 presents the decomposition of the error; in particular, Lemmas 6.3, 6.4 and 6.5 are the main technical estimates. Note that in Section 7, Lemmas 6.3 and 6.5 are proved with full details, while we omit the proof of Lemma 6.4 since it would be redundant with Bréhier (2014); instead we refer to Kopec (2014) for a detailed proof. Finally in Section 6.4 we recall an integration by parts formula (Malliavin calculus), which has been already used in Debussche (2011) and Bréhier (2014).

Let $\tau_0 \in (0, +\infty)$, and $\tau \in (0, \tau_0)$; let $N \in \mathbb{N}$ and set $T = N\tau$. Introduce the notation, for $k \in \mathbb{N}$, $t_k = k\tau$. We denote by $\kappa > 0$ an arbitrarily small parameter. Also let $\phi$ be an admissible test function.

### 6.1 *Projection in finite dimension*

First, decompose the error using the orthogonal projection $P_M$ onto $H_M$:

$$\frac{1}{N}\sum_{m=0}^{N-1}\mathbb{E}\phi(X_m^h) - \overline{\phi} = \frac{1}{N}\sum_{m=0}^{N-1}\mathbb{E}\phi(P_M X_m^h) - \overline{\phi}_M + \left(\overline{\phi}_M - \overline{\phi}\right) + \frac{1}{N}\sum_{m=0}^{N-1}\Big(\mathbb{E}\phi(X_m^h) - \mathbb{E}\phi(P_M X_m^h)\Big),$$

where $\overline{\phi}_M = \int_{H_M}\phi\, \mathrm{d}\mu^{(M)}$. Note that $\overline{\phi}_M \underset{M\to\infty}{\to} \overline{\phi}$. Indeed, for all $x \in H$ and $t \geq 0$,

$$\int_{H_M}\phi(z)\,\mathrm{d}\mu^{(M)}(z) - \int_H \phi(z)\,\mathrm{d}\mu(z) = \mathbb{E}\phi(X(t)) - \int_H\phi(z)\,\mathrm{d}\mu(z) + \int_{H_M}\phi(z)\,\mathrm{d}\mu^{(M)}(z) - \mathbb{E}\phi(X^{(M)}(t)) + \mathbb{E}\phi(X^{(M)}(t)) - \mathbb{E}\phi(X(t)).$$

Then for any $t > 0$,

$$\limsup_{M\to+\infty}\left|\int_{H_M}\phi(z)\,\mathrm{d}\mu^{(M)}(z) - \int_H\phi(z)\,\mathrm{d}\mu(z)\right| \leq C\exp(-ct),$$

and it remains to take $t \to +\infty$. Note that the constant $c$ does not depend on dimension $M$.

As a consequence, the effort in the following is concentrated on the proof of error bounds on $\frac{1}{N}\sum_{m=0}^{N-1}\mathbb{E}\phi(P_M X_m^h) - \overline{\phi}_M$, uniform over $M \in \mathbb{N}$.

### 6.2 *Some results on the Poisson equation in finite dimensions*

Let $M \in \{1, 2, \ldots\}$. Let $\phi \in \mathcal{C}_b^2(H, \mathbb{R})$. The function $\Psi^{(M)} : H_M \to \mathbb{R}$ is defined as the unique solution of the Poisson equation on $H_M$:

$$\mathcal{L}^{(M)}\Psi^{(M)} = \phi \circ P_M - \overline{\phi}_M \quad \text{with} \quad \int_{H_M} \Psi^{(M)} \, \mathrm{d}\mu^{(M)} = 0, \tag{6.1}$$

where $\mathcal{L}^{(M)}$ is the infinitesimal generator of the SPDE (3.7): for $\psi : H \to \mathbb{R}$ of class $\mathcal{C}^2$ and $x \in H$,

$$\mathcal{L}^{(M)}\psi(x) = \langle AP_M x + P_M F(x), D\psi(x)\rangle + \frac{1}{2}\mathrm{Tr}(P_M D^2\psi(x)).$$

In the following, we will need to control the first and the second derivatives of $\Psi^{(M)}$. Proposition 6.1 is the key result needed to obtain the optimal orders of convergence.

PROPOSITION 6.1 Let $M \in \{1, 2, \ldots\}$. Let $\phi \in \mathcal{C}_b^2(H)$. The function $\Psi^{(M)}$ defined for $x \in H_M$ by

$$\Psi^{(M)}(x) = -\int_0^{+\infty} \mathbb{E}\Big(\phi(X^{(M)}(t, x)) - \overline{\phi}_M\Big)\, \mathrm{d}t$$

is of class $\mathcal{C}^2$ and the unique solution of (6.1). Moreover, we have the following estimates: for any $0 \le \beta, \gamma < 1/2$ there exists $C, C_\beta, C_{\beta,\gamma} \in (0, +\infty)$ (independent of $M$) such that for any $x \in H_M$,

$$\begin{aligned} |\Psi^{(M)}(x)| &\le C(1 + |x|^2) \parallel \phi \parallel_\infty, \\ |D\Psi^{(M)}(x)|_\beta &\le C_\beta(1 + |x|^2) \parallel \phi \parallel_{1,\infty} \end{aligned} \tag{6.2}$$

and

$$|(-A)^\beta D^2\Psi^{(M)}(x)(-A)^\gamma|_{\mathcal{L}(H_M)} \le C_{\beta,\gamma}(1 + |x|^2) \parallel \phi \parallel_{2,\infty} . \tag{6.3}$$

REMARK 6.2 In fact, the result on $D\Psi$ is also true for $\beta < 1$, and the result on $D^2\Psi$ is also true for $\beta, \gamma < 1$ such that $\beta + \gamma < 1$. Moreover, all the constants are uniform with respect to $M \in \{1, 2, \ldots\}$.

We refer to Kopec (2014, Chapter 4, Section 8) for a detailed proof of Proposition 6.1; see also Bréhier (2014, Section 5.2).

### 6.3 *Decomposition of the error*

Let $M \in \mathbb{N}$; define the auxiliary function $\tilde{\Psi}^{(M)}$ as follows: for $x \in H$,

$$\tilde{\Psi}^{(M)}(x) = \Psi^{(M)}(P_M x),$$

where $\Psi^{(M)}$ is the solution of the Poisson equation (6.1). Using the notation of Remark 2.1, for any $x \in H$,

$$\begin{aligned} D\tilde{\Psi}^{(M)}(x) &= P_M D\Psi^{(M)}(P_M x), \\ D^2\tilde{\Psi}^{(M)}(x) &= P_M D^2\Psi^{(M)}(P_M x)P_M. \end{aligned}$$

Then the estimates of Proposition 6.1 are satisfied if $\Psi^{(M)}$ is replaced with $\tilde{\Psi}^{(M)}$.

For all $m \in \mathbb{N}$, the process $\tilde{X}^h$ defined by (3.13) is related on $[t_m, t_{m+1}]$ with the generator $\mathcal{L}^{\tau,m,h}$ defined for $x \in V_h$ and $\phi \in \mathcal{L}(H)$ by

$$\mathcal{L}^{\tau,m,h}\phi(x) = \langle S_{\tau,h}A_hX_m^h + S_{\tau,h}P_hF(X_m^h), D\phi(x)\rangle + \frac{1}{2}\mathrm{Tr}(S_{\tau,h}S_{\tau,h}^*P_hD^2\phi(x)). \tag{6.4}$$

Thanks to Itô's formula and Proposition 6.1, for any $m \in \mathbb{N}$,

$$\mathbb{E}\tilde{\Psi}^{(M)}(X_{m+1}^h) - \mathbb{E}\tilde{\Psi}^{(M)}(X_m^h) = \int_{t_m}^{t_{m+1}} \mathbb{E}\mathcal{L}^{\tau,m,h}\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s.$$

Let us also define the generator $\mathcal{L}^h$ of the finite element solution $X^h$: for $x \in V_h$,

$$\mathcal{L}^h\phi(x) = \langle A_hx + P_hF(x), D_x\phi(x)\rangle + \frac{1}{2}\mathrm{Tr}(P_hD_{xx}^2\phi(x)).$$

Then introduce the following decomposition:

$$\begin{aligned}\mathbb{E}\tilde{\Psi}^{(M)}(X_{m+1}^h) - \mathbb{E}\tilde{\Psi}^{(M)}(X_m^h) &= \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^{\tau,m,h} - \mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\ &\quad+ \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^h - \mathcal{L}^{(M)}\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\ &\quad+ \int_{t_m}^{t_{m+1}} \mathbb{E}\mathcal{L}^{(M)}\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s.\end{aligned}$$

Note that $\mathcal{L}^{(M)}\tilde{\Psi}^{(M)}(x) = \mathcal{L}^{(M)}\Psi^{(M)}(P_Mx) + \langle P_MF(x) - P_MF(P_Mx), D\Psi^{(M)}(P_Mx)\rangle$, for $x \in H$, and due to definition (6.1) of $\Psi^{(M)}$ as the solution of the Poisson equation on $H_M$ associated with the generator $\mathcal{L}^{(M)}$, then

$$\begin{aligned}\mathbb{E}\tilde{\Psi}^{(M)}(X_{m+1}^h) - \mathbb{E}\tilde{\Psi}^{(M)}(X_m^h) &= \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^{\tau,m,h} - \mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\ &\quad+ \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^h - \mathcal{L}^{(M)}\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\ &\quad+ \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\phi(P_M\tilde{X}^h(s)) - \overline{\phi}_M\Big)\,\mathrm{d}s\\ &\quad+ \int_{t_m}^{t_{m+1}} \mathbb{E}\Big\langle P_M\Big(F(\tilde{X}^h(s)) - F(P_M\tilde{X}^h(s))\Big), D\Psi^{(M)}(P_M\tilde{X}^h(s))\Big\rangle\,\mathrm{d}s\\ &= \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^{\tau,m,h} - \mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\ &\quad+ \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^h - \mathcal{L}^{(M)}\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\ &\quad+ \tau\big(\mathbb{E}\phi(P_MX_m^h) - \overline{\phi}_M\big)\end{aligned}$$

$$+\int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\phi(P_M\tilde{X}^h(s))\Big) - \mathbb{E}\Big(\phi(P_M X_m^h)\Big)\,\mathrm{d}s$$

$$+\int_{t_m}^{t_{m+1}} \mathbb{E}\Big\langle P_M\Big(F(\tilde{X}^h(s)) - F(P_M\tilde{X}^h(s))\Big), D\Psi^{(M)}(P_M\tilde{X}^h(s))\Big\rangle\,\mathrm{d}s.$$

Then summing over $m = 1, \ldots, N-1$ and dividing by $N\tau$, the error is decomposed as

$$\begin{aligned}
\frac{1}{N}\sum_{m=0}^{N-1}\Big(\mathbb{E}\phi(P_M X_m^h) - \overline{\phi}_M\Big) &= \frac{1}{N\tau}\big(\mathbb{E}\Psi^{(M)}(P_M X_N^h) - \mathbb{E}\Psi^{(M)}(P_M X_1^h)\big)\\
&\quad + \frac{1}{N}\big(\phi(P_M x) - \overline{\phi}_M\big)\\
&\quad + \frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^{(M)} - \mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\
&\quad + \frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^h - \mathcal{L}^{\tau,m,h}\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s\\
&\quad - \frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}} \Big(\mathbb{E}\phi(P_M\tilde{X}^h(s)) - \mathbb{E}\phi(P_M X_m^h)\Big)\,\mathrm{d}s\\
&\quad - \frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}} \mathbb{E}\Big\langle P_M\Big(F(\tilde{X}^h(s)) - F(P_M\tilde{X}^h(s))\Big), D\Psi^{(M)}(P_M\tilde{X}^h(s))\Big\rangle\,\mathrm{d}s\\
&= I_1 + I_2 + I_3 + I_4 + I_5 + I_6.
\end{aligned} \tag{6.5}$$

Some of the terms $I_i$ are easily controlled. Indeed, using Proposition 6.1 and Lemma 3.10, for $0 < \tau < \tau_0$,

$$|I_1 + I_2| \le C(1 + |x|^2)\frac{1}{N\tau},$$

where $\tau_0$ is any fixed positive real number. Moreover, since $F$ is Lipschitz continuous, Proposition 6.1 and Lemma 3.10 yield

$$\lim_{M\to\infty} I_6 \to 0.$$

The control (uniform with respect to $M \in \mathbb{N}$) of the three other terms is performed in Section 7. First, in Section 7.1, the following estimate of $I_3$ is shown.

LEMMA 6.3 (Space-discretization error) For any $0 < \kappa < 1/2$ and $\tau_0$, there exists a constant $C > 0$ such that for any $\phi \in \mathcal{C}_b^2(H)$, $x \in H$ and $0 < \tau \le \tau_0$,

$$\limsup_{M\to\infty} \frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^{(M)} - \mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s \le C(1 + |x|^3)\|\phi\|_{2,\infty} h^{1-\kappa}(1 + (N\tau)^{-1}).$$

The term $I_4$ is controlled in a similar way.

LEMMA 6.4 (Time-discretization error) For any $0 < \kappa < 1/2$ and $\tau_0$, there exists a constant $C > 0$ such that for any $\phi \in \mathcal{C}_b^2(H)$, $M \in \{1, 2, \ldots\}$, $y \in H$ and $0 < \tau \leq \tau_0$,

$$\left| \frac{1}{N\tau} \sum_{m=1}^{N-1} \int_{t_m}^{t_{m+1}} \mathbb{E} \left( \mathcal{L}^h - \mathcal{L}^{\tau,m,h} \right) \tilde{\Psi}^{(M)}(\tilde{X}^h(t)) \, dt \right| \leq C \|\phi\|_{2,\infty} (1 + |x|^3) \tau^{1/2-\kappa} (1 + (N\tau)^{-1+\kappa} + (N\tau)^{-1}).$$

We omit the proof of this result in the interest of conciseness. Indeed, no original argument is required compared with Bréhier (2014). We refer to Kopec (2014) for a detailed proof. Note that the growth conditions on the second-order derivative of the nonlinearity $F$, Assumptions 2.7, are only explicitly used in the proof of Lemma 6.4 (they also appear in the proof of Proposition 6.1).

Finally, the control of $I_5$ is provided in Section 7.2. Precisely, we have the following lemma.

LEMMA 6.5 (Additional time-discretization error) For any $0 < \kappa < 1/4$ and $\tau_0$, there exists a constant $C > 0$ such that for any $\phi \in \mathcal{C}_b^2(H)$, $M \in \{1, 2, \ldots\}$, $y \in H$ and $0 < \tau \leq \tau_0$,

$$\left| \frac{1}{N\tau} \sum_{m=1}^{N-1} \int_{t_m}^{t_{m+1}} \left( \mathbb{E}\phi(P_M \tilde{X}^h(t)) - \mathbb{E}\phi(P_M X_m^h) \right) dt \right| \leq C \|\phi\|_{2,\infty} \tau^{1/2-2\kappa} \left( 1 + \frac{|x|}{(N\tau)^{1-\kappa}} \right).$$

A detailed proof for Lemma 6.5 is given since this term due to the discretization error is specific to the strategy to decompose the error (using the solution of the Poisson equation), and in particular no similar term appears in Debussche (2011) and Bréhier (2014).

## 6.4 *A Malliavin integration by parts formula*

As explained in Section 1, one of the key tools to obtain the right weak order is a transformation of some spatially irregular terms involving the stochastic integral with respect to the cylindrical Wiener process, into more suitable, deterministic ones, thanks to an integration by parts formula involving Malliavin calculus—we refer to the monographs Sanz-Solé (2005), Nualart (2006) for an extensive study of this object and for the definition of the important notation. In this article, the Malliavin derivative and the integration by parts formula are taken as (essential) tools, but it is not the intention to give more information on it.

The notation here is the same as in Debussche (2011), where the following useful integration by parts formula is given—see Lemma 2.1 therein.

LEMMA 6.6 For any $\mathrm{F} \in \mathbb{D}^{1,2}(V_h)$, $u \in \mathcal{C}_b^2(V_h)$ and $\Psi \in L^2(\Omega \times [0, T], \mathcal{L}_2(V_h))$ an adapted process,

$$\mathbb{E}\left[ Du(\mathrm{F}) \cdot \int_0^T \Psi(s) \, dW^{(M)}(s) \right] = \mathbb{E}\left[ \int_0^T \mathrm{Tr}(\Psi(s)^* D^2 u(\mathrm{F}) \mathcal{D}_s \mathrm{F}) \, ds \right], \tag{6.6}$$

where $\mathcal{D}_s \mathrm{F} : \ell \in H \mapsto \mathcal{D}_s^\ell \mathrm{F} \in V_h$ stands for the Malliavin derivative of F. The domain $\mathbb{D}^{1,2}(V_h)$ of $\mathcal{D}_s$ is the set of $H$-valued random variables $\mathrm{F} = \sum_{i \in \mathbb{N}, i \leq N_h} \mathrm{F}_i e_i$, such that, for all $i$, $\mathrm{F}_i$ belongs to the domain $\mathbb{D}^{1,2}$ of the Malliavin derivative for $\mathbb{R}$-valued random variables.

In the sequel, Lemma 6.6 is used with the Galerkin approximations; the components all belong to the domain $\mathbb{D}^{1,2}$, and all calculations are valid. In addition to Lemma 6.6, we use the chain rule to compute Malliavin derivatives.

Note that (6.6) holds true if $u$ is not assumed to be bounded but only $u \in \mathcal{C}^2(V_h)$ provided the expectations and the integral above are well defined. This is easily seen by approximation of $u$ by bounded functions.

Under Assumptions 2.9(i), it is not possible to get uniform-in-time estimates of the Malliavin derivative of $\tilde{X}^h$; we circumvent this problem below by using these derivatives only at times $t_k = k\tau$ and $s$ such that $t_{k-l_s} \leq 1$, where we recall that $l_s = \lfloor \frac{s}{\tau} \rfloor$ ($\lfloor \cdot \rfloor$ denoting the integer part function).

Lemma 6.7 For any $0 \leq \beta < 1$ and $\tau_0 > 0$, there exists a constant $C > 0$ such that for every $h \in (0,1)$, $k \geq 1$, $0 < \tau \leq \tau_0$ and $s \in [0, t_k]$,

$$|(-A_h)^\beta \mathcal{D}_s X_k^h|_{\mathcal{L}(V_h)} \leq C(1 + L_F\tau)^{k-l_s} \left(1 + \frac{1}{(1+\lambda_0\tau)^{(1-\beta)(k-l_s)} t_{k-l_s}^\beta}\right).$$

Moreover, if $t_k \leq t < t_{k+1}$, we have

$$|(-A_h)^\beta \mathcal{D}_s^\ell \tilde{X}^h(t)|_{\mathcal{L}(V_h)} \leq C |(-A_h)^\beta \mathcal{D}_s^\ell X_k^h|_{\mathcal{L}(V_h)}.$$

We want to emphasize that the constant in Lemma 6.7 is uniform with respect to $h \in (0,1)$.

*Proof.* According to the definition of the Malliavin derivative $\mathcal{D}_s^\ell \tilde{X}^h(t)$ as a linear operator in $V_h$, we need to control $|(-A_h)^\beta \mathcal{D}_s^\ell \tilde{X}^h(t)|$ and $|(-A_h)^\beta \mathcal{D}_s^\ell X_k^h|$, uniformly with respect to $\ell \in V_h$ with $|\ell| \leq 1$.

Let $h \in (0,1)$. For any $k \geq 1$, $\ell \in V_h$ and $s \in [0, t_k]$, using the chain rule for Malliavin calculus and expressions (3.11) and (3.12), we have

$$\mathcal{D}_s^\ell X_k^h = S_{\tau,h}^{k-l_s}\ell + \tau \sum_{i=l_s+1}^{k-1} S_{\tau,h}^{k-i} D(P_h F)(X_i^h) \cdot \mathcal{D}_s^\ell X_i^h.$$

We recall that $l_s$ denotes the integer part of $\frac{s}{\tau}$, so that when $i \leq l_s$ we have $\mathcal{D}_s^\ell X_i^h = 0$.

As a consequence, the discrete Gronwall lemma ensures that for $k \geq l_s + 1$,

$$|\mathcal{D}_s^\ell X_k^h| \leq (1 + L_F\tau)^{k-l_s} |\ell|.$$

Now using Lemma 3.7, we have

$$|(-A_h)^\beta \mathcal{D}_s^\ell X_k^h| \leq \frac{1}{(1+\lambda_0\tau)^{(1-\beta)(k-l_s)} t_{k-l_s}^\beta} |\ell| + L_F\tau \sum_{i=l_s+1}^{k-1} \frac{(1+L_F\tau)^{l-l_s}}{(1+\lambda_0\tau)^{(1-\beta)(k-i)} t_{k-i}^\beta} |\ell|.$$

To conclude, we see that when $0 < \tau \leq \tau_0$,

$$\tau \sum_{i=l_s+1}^{k-1} \frac{1}{(1+\lambda_0\tau)^{(1-\beta)(k-i)} t_{k-i}^\beta} \leq C \int_0^{+\infty} t^{-\beta} \frac{1}{(1+\lambda_0\tau)^{(1-\beta)t/\tau}} \,\mathrm{d}t \leq C < +\infty.$$

The second inequality is a consequence of the following equality for $s \le t_k \le t < t_{k+1}$, thanks to (3.13):

$$\mathcal{D}_s^\ell \tilde{X}^h(t) = \mathcal{D}_s^\ell X_k^h + (t - t_m)(A_h S_{\tau,h} \mathcal{D}_s^\ell X_k^h + S_{\tau,h} P_h DF(X_k^h)\mathcal{D}_s^\ell X_k^h),$$

and the conclusion follows since

$$\sup_{h\in(0,1)} |\tau A_h S_{\tau,h}|_{\mathcal{L}(V_h)} < +\infty. \qquad \square$$

## 7. Detailed proof of the estimates

We warn the reader that constants may vary from line to line during the proofs, and that in order to use lighter notation we usually forget to mention dependence on the parameters. We use the generic notation $C$ for such constants. All constants will depend on a parameter $\kappa > 0$, which can be chosen arbitrarily small; $\kappa$ may also change and be unified at the end of the computation.

To simplify the expressions, the dependence of the error with respect to the test function $\phi$ is not mentioned in the proof.

### 7.1 *Proof of Lemma 6.3: space discretization error*

7.1.1 *Strategy.* To control this term, we mix ideas described in Bréhier (2014) and Andersson & Larsson (2016), where the authors use estimates of the solution $u$ of the Kolmogorov equation. Here we use a similar approach, with $u$ replaced with the solution $\Psi^{(M)}$ of the Poisson equation (6.1), with $M \in \{1, 2, \ldots\}$.

Let $M \in \{1, 2, \ldots\}$ be fixed. First, the difference $\mathcal{L}^{(M)} - \mathcal{L}^h$ is decomposed into three terms. For any $x \in H$,

$$\begin{aligned}\Big(\mathcal{L}^{(M)} - \mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(x) &= \Big\langle \big(A_M - A_h\big)x, D\tilde{\Psi}^{(M)}(x)\Big\rangle \\ &+ \Big\langle \big(P_M - P_h\big)F(x), D\tilde{\Psi}^{(M)}(x)\Big\rangle \\ &+ \tfrac{1}{2}\mathrm{Tr}\Big(\big(P_M - P_h\big)D^2\tilde{\Psi}^{(M)}(x)\Big),\end{aligned}$$

where $A_M = AP_M$; thus

$$\frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}} \mathbb{E}\Big(\mathcal{L}^{(M)} - \mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\,\mathrm{d}t = \frac{1}{N\tau}\sum_{m=1}^{N-1}(a^m + b^m + c^m),$$

where for $1 \le m \le N-1$,

$$\begin{aligned}a^m &= \mathbb{E}\int_{t_m}^{t_{m+1}} \langle (A_M - A_h)\tilde{X}^h(t), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle\,\mathrm{d}t,\\ b^m &= \mathbb{E}\int_{t_m}^{t_{m+1}} \langle (P_M - P_h)F(\tilde{X}^h(t)), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle\,\mathrm{d}t,\\ c^m &= \frac{1}{2}\mathbb{E}\int_{t_m}^{t_{m+1}} \mathrm{Tr}\Big((P_M - P_h)D^2\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\Big)\,\mathrm{d}t.\end{aligned}$$

7.1.2 *Estimate of* $a^m$. The Ritz projection $R_h$ can be expressed in the form $R_h = A_h^{-1}P_hA$. Using this we can write

$$\begin{aligned}\langle (A_M - A_h)\tilde{X}^h(t), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle &= \langle (A_M P_h - A_h P_h)\tilde{X}^h(t), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \\ &= \langle \tilde{X}^h(t), (P_h A_M - A_h P_h) D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \\ &= \langle \tilde{X}^h(t), A_h P_h (R_h P_M - I) D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \\ &= \langle \tilde{X}^h(t), A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \\ &\quad + \langle \tilde{X}^h(t), A_h P_h (P_M - I) D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle.\end{aligned}$$

The idea of this decomposition is to apply the error estimates (3.5) and (3.8) for $R_h$ and $P_M$, respectively. We now use formula (3.13) on $\tilde{X}^h(t)$. We then need to estimate the following five terms:

$$\begin{aligned} a^m = &\ \mathbb{E}\int_{t_m}^{t_{m+1}} \langle X_m^h, A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \,\mathrm{d}t \\ &+ \mathbb{E}\int_{t_m}^{t_{m+1}} (t - t_m)\langle A_h S_{\tau,h} X_m^h, A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \,\mathrm{d}t \\ &+ \mathbb{E}\int_{t_m}^{t_{m+1}} (t - t_m)\langle S_{\tau,h} P_h F(X_m^h), A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \,\mathrm{d}t \\ &+ \mathbb{E}\int_{t_m}^{t_{m+1}} \left\langle \int_{t_m}^{t} S_{\tau,h} P_{ht}\,\mathrm{d}W(s), A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\right\rangle \mathrm{d}t \\ &+ \mathbb{E}\int_{t_m}^{t_{m+1}} \langle A_h \tilde{X}^h(t), (P_M - I) D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \,\mathrm{d}t \\ = &\ a_1^{m,h} + a_2^{m,h} + a_3^{m,h} + a_4^{m,h} + a^{m,M}. \end{aligned}$$

(1) **Estimate of** $a_1^{m,h}$**.** We use expressions (3.11) of $X_m^h$ and (3.12) to decompose $a_1^{m,h}$:

$$\begin{aligned} a_1^{m,h} = &\ \mathbb{E}\int_{t_m}^{t_{m+1}} \langle S_{\tau,h}^m P_h x, A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \,\mathrm{d}t \\ &+ \mathbb{E}\int_{t_m}^{t_{m+1}} \tau \sum_{\ell=0}^{k-1} \langle S_{\tau,h}^{m-\ell} P_h F(X_\ell^h), A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \,\mathrm{d}t \\ &+ \mathbb{E}\int_{t_m}^{t_{m+1}} \left\langle \int_0^{t_m} S_{\tau,h}^{m-l_s} P_h \mathrm{d}W(s), A_h P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\right\rangle \mathrm{d}t \\ = &\ a_{1,1}^{m,h} + a_{1,2}^{m,h} + a_{1,3}^{m,h}. \end{aligned}$$

- *Estimate of* $a_{1,1}^{m,h}$. The ideas are to write $(-A_h) = (-A_h)^\kappa(-A_h)^{1-\kappa}$ and to use regularization properties of the semigroup $(S_{\tau,h}^k)_{k\in\mathbb{N}}$. Thanks to Proposition 3.3, Proposition 6.1 for $\beta = 1/2$, Lemma 3.10 and Lemma 3.7, we get, for any small enough parameter $0 < \kappa < 1/2$,

$$\begin{aligned}|a_{1,1}^{m,h}| &= \left|\mathbb{E}\int_{t_m}^{t_{m+1}} \langle(-A_h)^{1-\kappa} S_{\tau,h}^m P_h x, (-A_h)^{\kappa} P_h(R_h - I)(-A)^{-1/2} P_M (-A)^{1/2} D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle \,\mathrm{d}t\right| \\ &\leq \mathbb{E}\int_{t_m}^{t_{m+1}} |(-A_h)^{1-\kappa} S_{\tau,h}^m P_h|_{\mathcal{L}(H)}|x|\, |(-A_h)^{\kappa} P_h(R_h - I)(-A)^{-1/2}|_{\mathcal{L}(H)} |P_M|_{\mathcal{L}(H)} \\ &\quad \times |(-A)^{1/2} D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))|\,\mathrm{d}t \\ &\leq C\frac{1}{(m\tau)^{1-\kappa}}\frac{1}{(1+\lambda_0\tau)^{m\kappa}}|x|\,|(-A)^{\kappa}(R_h - I)(-A)^{-1/2}|_{\mathcal{L}(H)} \int_{t_m}^{t_{m+1}} \mathbb{E}(1+|\tilde{X}^h(t)|^2)\,\mathrm{d}t \\ &\leq C\tau\frac{1}{(m\tau)^{1-\kappa}}\frac{1}{(1+\lambda_0\tau)^{m\kappa}}(1+|x|^3)h^{1-2\kappa}.\end{aligned}$$

We will now use the following useful inequality: for $\tau \leq \tau_0$ and any $N \geq 1$,

$$\tau\sum_{l=1}^{N}\frac{1}{(l\tau)^{1-\kappa}}\frac{1}{(1+\lambda_0\tau)^{l\kappa}} \leq C_\kappa. \tag{7.1}$$

Indeed,

$$\begin{aligned}\tau\sum_{l=1}^{N}\frac{1}{(l\tau)^{1-\kappa}}\frac{1}{(1+\lambda_0\tau)^{l\kappa}} &\leq C\int_0^{t_N}\frac{1}{t^{1-\kappa}}\frac{1}{(1+\lambda_0\tau)^{\kappa t/\tau}}\,\mathrm{d}t \\ &\leq \int_0^{\infty}\frac{1}{t^{1-\kappa}}\mathrm{th}\, e^{-t\kappa/\tau\ \log(1+\lambda_0\tau)}\,\mathrm{d}t \\ &\leq \int_0^{\infty}\frac{1}{s^{1-\kappa}}\, e^{-s}\,\mathrm{d}s\left(\frac{\tau}{\kappa\ \log(1+\lambda_0\tau)}\right)^{\kappa} \leq C_\kappa.\end{aligned}$$

Then, using (7.1), we get

$$\frac{1}{N\tau}\sum_{m=1}^{N-1}|a_{1,1}^{m,h}| \leq C\frac{1}{T}h^{1-2\kappa}(1+|x|^3). \tag{7.2}$$

- *Estimate of* $a_{1,2}^{m,h}$. Using the same ideas as for estimating $a_{1,1}^{m,h}$, we have

$$\begin{aligned}|a_{1,2}^{m,h}| \leq C\tau\mathbb{E}\int_{t_m}^{t_{m+1}}\sum_{l=0}^{m-1}|(-A_h)^{1-\kappa}S_{\tau,h}^{m-l}P_hF(X_l^h)|\,|(-A_h)^{\kappa}P_h(R_h - I)(-A)^{-1/2}|_{\mathcal{L}(H)} \\ \times |(-A)^{1/2}D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))|\,\mathrm{d}t.\end{aligned}$$

Since $F$ is Lipschitz continuous, using Lemma 3.10, estimate (7.1) yields

$$\mathbb{E}\left|\tau(-A_h)^{1-\kappa}\sum_{l=0}^{m-1}S_{\tau,h}^{m-l}F(X_l^h)\right| \leq C(1+|x|)\tau\sum_{l=1}^{m}\frac{1}{(l\tau)^{1-\kappa}}\frac{1}{(1+\lambda_0\tau)^{l\kappa}} \leq C_\kappa.$$

With Proposition 3.3, Lemma 3.10 and Proposition 6.1 for $\beta = 1/2$, we can now write

$$|a_{1,2}^{m,h}| \leq C(1+|x|^3)h^{1-2\kappa}\tau,$$

and we get

$$\frac{1}{N\tau}\sum_{m=0}^{N-1}|a_{1,2}^{m,h}| \leq C(1+|x|^3)h^{1-2\kappa}. \tag{7.3}$$

- *Estimate of* $a_{1,3}^{m,h}$. The analysis of this term is more complicated. We refer the reader to Bréhier (2014) for a discussion of the problem, and for detailed explanations of the strategy of the proof—following the original idea of Debussche (2011).
We recall that the problem lies in the regularity in space of the process due to the whiteness in space of the driving noise. The strategy used to control $a_{1,1}^{m,h}$ and $a_{2,1}^{m,h}$ would give an order of convergence of only $1/2$, instead of 1.
We decompose $a_{1,3}^{m,h}$ into two parts, corresponding to different intervals for the stochastic integration. We can work directly on one of these parts. On the other, a Malliavin integration by parts is performed: it allows us to use appropriate regularization properties and to obtain the correct order of convergence 1. We emphasize the length of the interval where this integration by parts is applied: its maximal size is independent of $\tau$ and $h$, and allows us to use Lemma 6.7 with controlled upper bounds.
By using (3.12), decompose

$$\begin{aligned}
a_{1,3}^{m,h} &= \mathbb{E}\int_{t_m}^{t_{m+1}}\left\langle \int_0^{t_m} S_{\tau,h}^{m-l_s}P_h\,\mathrm{d}W(s), (-A_h)P_h(R_h-I)P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\right\rangle \mathrm{d}t \\
&= \mathbb{E}\int_{t_m}^{t_{m+1}}\left\langle \int_0^{(t_m-3\tau_0)\vee 0} (-A_h)^{1-\kappa}S_{\tau,h}^{m-l_s}P_h\,\mathrm{d}W(s), (-A_h)^{\kappa}P_h(R_h-I)P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\right\rangle \mathrm{d}t \\
&\quad + \mathbb{E}\int_{t_m}^{t_{m+1}}\left\langle \int_{(t_m-3\tau_0)\vee 0}^{t_m} P_M(R_h-I)P_h(-A_h)S_{\tau,h}^{m-l_s}P_h\,\mathrm{d}W(s), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\right\rangle \mathrm{d}t.
\end{aligned}$$

For the first term—which is equal to 0 when $t_m < 3\tau_0$—thanks to the Cauchy–Schwarz inequality,

$$\begin{aligned}
&\left|\mathbb{E}\left\langle \int_0^{(t_m-3\tau_0)\vee 0}(-A_h)^{1-\kappa}S_{\tau,h}^{m-l_s}P_H\,\mathrm{d}W(s), (-A_h)^{\kappa}P_h(R_h-I)P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\right\rangle\right| \\
&\quad\leq (\mathbb{E}|(-A_h)^{\kappa}P_h(R_h-I)(-A)^{-1/2}P_M(-A)^{1/2}D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))|^2)^{1/2} \\
&\quad\quad\times\left(\mathbb{E}\left|\int_0^{(t_m-3\tau_0)\vee 0}(-A_h)^{1-\kappa}S_{\tau,h}^{m-l_s}P_h\,\mathrm{d}W(s)\right|^2\right)^{1/2}.
\end{aligned}$$

We have the following inequality—we remark that in the integral below $t_{m-l_s} \geq 1$:

$$\mathbb{E}\left|\int_0^{(t_m-3\tau_0)\vee 0}(-A_h)^{1-\kappa}S_{\tau,h}^{m-l_s}P_h\,\mathrm{d}W(s)\right|^2 = \int_0^{(t_m-3\tau_0)\vee 0}|(-(A_h)^{1-\kappa}S_{\tau,h}^{m-l_s}P_h|_{\mathcal{L}_2(H)}^2\,\mathrm{d}s$$

$$
\begin{aligned}
&= \int_0^{(t_m-3\tau_0)\vee 0} \mathrm{Tr}((-A_h)^{2-2\kappa} S_{\tau,h}^{2(m-l_s)} P_h)\,\mathrm{d}s \\
&\leq \int_0^{(t_m-3\tau_0)\vee 0} |S_{\tau,h}^{(m-l_s)} P_h|_{\mathcal{L}(H)} \\
&\quad \times |(-A_h)^{2+1/2+\kappa} S_{\tau,h}^{(m-l_s)} P_h|_{\mathcal{L}(H)}\,\mathrm{d}s \\
&\quad \times \mathrm{Tr}(P_h(-A_h)^{-1/2-\kappa} P_h) \\
&\leq C \int_0^{(t_m-3\tau_0)\vee 0} \frac{1}{(1+\lambda_0\tau)^{m-l_s} t_{m-l_s}^{2+1/2-\kappa}}\,\mathrm{d}s \\
&\leq C \int_0^{(t_m-3\tau_0)\vee 0} \frac{1}{(1+\lambda_0\tau)^{m-l_s}}\,\mathrm{d}s \\
&\leq C \int_0^{+\infty} \frac{1}{(1+\lambda_0\tau)^{s/\tau}}\,\mathrm{d}s \\
&\leq C,
\end{aligned}
$$

when $\tau \leq \tau_0$ and thanks to Proposition 3.4 and Lemma 3.7. Then, thanks to Proposition 3.3, Proposition 6.1 for $\beta = 1/2$ and Lemma 3.10, we get

$$
\begin{aligned}
&\left| \mathbb{E} \int_{t_m}^{t_{m+1}} \left\langle \int_0^{(t_m-3\tau_0)\vee 0} (-A_h)^{1-\kappa} S_{\tau,h}^{m-l_s} P_h\,\mathrm{d}W(s), (-A_h)^{\kappa} P_h (R_h - I) P_M D\tilde{\Psi}^{(M)}(\tilde{X}^h(t)) \right\rangle \mathrm{d}t \right| \\
&\qquad \leq C(1+|x|^2)\tau h^{1-2\kappa}.
\end{aligned}
$$

For the second term, we use the Malliavin integration by parts formula (Lemma 6.6) to get

$$
\begin{aligned}
&\mathbb{E} \int_{t_m}^{t_{m+1}} \left\langle \int_{(t_m-3\tau_0)\vee 0}^{t_m} P_M (R_h - I) P_h (-A_h) S_{\tau,h}^{m-l_s} P_h\,\mathrm{d}W(s), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t)) \right\rangle \mathrm{d}t \\
&= \mathbb{E} \int_{t_m}^{t_{m+1}} \int_{(t_m-3\tau_0)\vee 0}^{t_m} \mathrm{Tr}\left( S_{\tau,h}^{m-l_s} (-A_h) P_h (R_h - I) P_M D^2\tilde{\Psi}^{(M)}(\tilde{X}^h(t)) \mathcal{D}_s \tilde{X}^h(t) \right) \mathrm{d}s\,\mathrm{d}t.
\end{aligned}
$$

Thanks to both estimates of Lemma 6.7, we have for $(t_m - 3\tau_0) \vee 0 \leq s \leq t_m \leq t < t_{m+1}$,

$$
|(-A)^{\alpha} \mathcal{D}_s^{\ell} \tilde{X}^h(t)| \leq C(1+L_F\tau)^{m-l_s} \left( 1 + \frac{1}{(1+\lambda_0\tau)^{(1-\alpha)(m-l_s)} t_{m-l_s}^{\alpha}} \right),
$$

and we see that $(1+L_F\tau)^{m-l_s}$ is bounded by a constant.
We can then control the second term of $a_{3,1}^{m,h}$ with

$$
\begin{aligned}
&\mathbb{E} \int_{t_m}^{t_{m+1}} \int_{(t_m-3\tau_0)\vee 0}^{t_m} |(-A_h)^{1-3\kappa/2} S_{\tau,h}^{m-l_s}|_{\mathcal{L}(H)} |(-A_h)^{3\kappa/2} P_h (R_h - I)(-A)^{-1/2}|_{\mathcal{L}(H)} \\
&\quad \times |(-A)^{1/2} D^2\tilde{\Psi}^{(M)}(\tilde{X}^h(t))(-A)^{1/2-\kappa/2}|_{\mathcal{L}(H)} \mathrm{Tr}((-A)^{-1/2-\kappa/2}) |(-A)^{\kappa} \mathcal{D}_s \tilde{X}^h(t)|_{\mathcal{L}(H)}\,\mathrm{d}s\,\mathrm{d}t
\end{aligned}
$$

$$\leq C \int_{(t_m - 3\tau_0)\vee 0}^{t_m} t_{m-l_s}^{-1+3\kappa/2} \frac{1}{(1+\lambda_0\tau)^{(m-l_s)3\kappa/2}} \left(1 + t_{m-l_s}^{-\kappa} \frac{1}{(1+\lambda_0\tau)^{(m-l_s)(1-\kappa)}}\right) \mathrm{d}s\, \tau h^{1-3\kappa}(1+|x|^2),$$

using Proposition 6.1 and Lemmas 6.7 and 3.7.
We have

$$\int_{(t_m-3\tau_0)\vee 0}^{t_m} t_{m-l_s}^{-1+3\kappa/2} \frac{1}{(1+\lambda_0\tau)^{(m-l_s)3\kappa/2}} \mathrm{d}s \leq \int_0^{t_m} \frac{1}{s^{1-3\kappa/2}} \frac{1}{(1+\lambda_0\tau)^{3\kappa/2s/\tau}} \mathrm{d}s \leq C < +\infty,$$

for $\tau \leq \tau_0$, thanks to (7.1).
Therefore

$$\frac{1}{N\tau}\sum_{m=1}^{N-1} |a_{1,3}^{m,h}| \leq C(1+|x|^2)h^{1-3\kappa}. \tag{7.4}$$

Using (7.2–7.4), we have

$$\frac{1}{N\tau}\sum_{m=1}^{N-1} |a_1^{m,h}| \leq C\left(1+\frac{1}{T}\right) h^{1-3\kappa}(1+|x|^3). \tag{7.5}$$

(2) **Estimate of $a_2^{m,h}$.** Since $(t-t_m)|(-A_h)S_{\tau,h}|_{\mathcal{L}(H)} \leq C$, $a_2^{m,h}$ is bounded by the same expression as $a_1^{m,h}$: by (7.5), we have

$$\frac{1}{N\tau}\sum_{m=1}^{N-1} |a_2^{m,h}| \leq C\left(1+\frac{1}{T}\right) h^{1-3\kappa}(1+|x|^3). \tag{7.6}$$

(3) **Estimate of $a_3^{m,h}$.** We have

$$|a_3^{m,h}| \leq \mathbb{E}\int_{t_m}^{t_{m+1}} (t-t_m)|(-A_h)^{1-\kappa} S_{\tau,h} P_h|_{\mathcal{L}(H)} |F(X_m^h)|$$
$$\times |(-A_h)^{\kappa} P_h (R_h - I)(-A)^{-1/2} P_M (-A)^{1/2} D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))|\, \mathrm{d}t.$$

Since $(t-t_m)|(-A_h)^{1-\kappa}S_{\tau,h}P_h|_{\mathcal{L}(H)}$ is bounded, using Lipschitz continuity of $F$ and Lemma 3.10, following the proof of the bound on $a_{1,1}^{m,h}$ one gets

$$\frac{1}{N\tau}\sum_{m=1}^{N-1} |a_3^{m,h}| \leq \frac{C}{T} h^{1-2\kappa}(1+|x|^3). \tag{7.7}$$

(4) **Estimate of $a_4^{m,h}$.** We again use the integration by parts formula to rewrite $a_4^{m,h}$:

$$a_4^{m,h} = -\mathbb{E}\int_{t_m}^{t_{m+1}} \left\langle \int_{t_m}^{t} S_{\tau,h} P_h \,\mathrm{d}W(s), (-A_h)P_h(R_h - I) D\tilde{\Psi}^{(M)}(\tilde{X}^h(t)) \right\rangle \mathrm{d}t$$
$$= -\mathbb{E}\int_{t_m}^{t_{m+1}} \int_{t_m}^{t} \mathrm{Tr}(S_{\tau,h} P_h (-A_h) P_h (R_h - I) \tilde{\Psi}^{(M)}(\tilde{X}^h(t)) \mathcal{D}_s \tilde{X}^h(t))\, \mathrm{d}s\, \mathrm{d}t.$$

From (3.13), for $t_m \leq s \leq t \leq t_{m+1}$ we have $\mathcal{D}_s^{\ell}\tilde{X}^h(t) = S_{\tau,h}P_h\ell$; as a consequence, the situation is much simpler and we do not need to use the same trick as in the control of $a_{1,3}^{m,h}$.

Then, as previously, we have

$$\begin{aligned}
|a_4^{m,h}| &\leq \mathbb{E}\int_{t_m}^{t_{m+1}} (t-t_m)\mathrm{Tr}((-A_h)^{1-\kappa}S_{\tau,h}P_h(-A_h)^{\kappa}P_h(R_h-I)(-A)^{-1/2}P_M \\
&\qquad \times (-A)^{1/2}D^2\tilde{\Psi}^{(M)}(\tilde{X}^h(t))(-A)^{1/2-\kappa/2}(-A)^{-1/2-\kappa/2}(-A)^{\kappa}S_{\tau,h})\,\mathrm{d}t \\
&\leq c|(-A_h)^{1-\kappa}S_{\tau,h}P_h|_{\mathcal{L}(H)}\mathrm{Tr}((-A)^{-1/2-\kappa/2})|(-A_h)^{\kappa}P_h(R_h-I)(-A)^{-1/2}|_{\mathcal{L}(H)} \\
&\qquad \times \mathbb{E}\int_{t_m}^{t_{m+1}} |(-A)^{1/2}D^2\tilde{\Psi}^{(M)}(\tilde{X}^h(t))(-A)^{1/2-\kappa/2}|_{\mathcal{L}(H)}|(-A)^{\kappa}S_{\tau,h}P_h|_{\mathcal{L}(H)}\,\mathrm{d}t \\
&\leq c(1+|x|^2)\tau h^{1-2\kappa}.
\end{aligned}$$

Therefore

$$\frac{1}{N\tau}\sum_{m=1}^{N-1}|a_4^{m,h}| \leq C(1+|x|^2)h^{1-2\kappa}. \tag{7.8}$$

(5) **Estimate of $a^{m,M}$.** Using Proposition 6.1, Lemma 3.10 and estimate (3.8), we have

$$\begin{aligned}
|a^{m,M}| &\leq \int_{t_m}^{t_{m+1}} \mathbb{E}\Big(|(-A_h)P_h|_{\mathcal{L}(H)}|\tilde{X}^h(t)|\,|(P_M-I)(-A)^{-1/2+\kappa}|\,|(-A)^{1/2+\kappa}D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))|\Big)\,\mathrm{d}t \\
&\leq C_h \parallel \phi \parallel_{1,\infty} \lambda_M^{-1/2+\kappa}\int_{t_m}^{t_{m+1}} \mathbb{E}\Big(|\tilde{X}^h(t)|(1+|\tilde{X}^h(t)|^2)\Big)\,\mathrm{d}t \\
&\leq C_h \parallel \phi \parallel_{1,\infty} \lambda_M^{-1/2+\kappa}\tau(1+|x|^3).
\end{aligned}$$

Then, we get

$$\lim_{M\to\infty}\frac{1}{N\tau}\sum_{m=1}^{N-1}|a^{m,M}| = 0.$$

With the previous estimates, we get

$$\limsup_{M\to+\infty}\frac{1}{N\tau}\sum_{m=1}^{N-1}|a^m| \leq C \parallel \phi \parallel_{1,\infty} (1+|x|^3)(1+T^{-1})h^{1-3\kappa}. \tag{7.9}$$

7.1.3 *Estimate of $b^m$.* Writing $P_M - P_h = (P_M - I) + (I - P_h)$, we get the natural decomposition

$$b^m = \mathbb{E}\int_{t_m}^{t_{m+1}} \langle (P_M - I)F(P_M\tilde{X}^h(t)), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle\,\mathrm{d}t$$

$$+\,\mathbb{E}\int_{t_m}^{t_{m+1}} \langle (I-P_h)F(P_M\tilde{X}^h(t)), D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle\,\mathrm{d}t$$
$$= b^{m,M} + b^{m,h}.$$

Using Lemma 3.10 and the Lipschitz continuity of $F$, Proposition 6.1 and Lemma 3.9, one gets for $i \in \{h, M\}$,

$$\begin{aligned}
|b^{m,i}| &= \left|\mathbb{E}\int_{t_m}^{t_{m+1}} \langle F(P_M\tilde{X}^h(t)), (P_i - I)(-A)^{-1/2+\kappa}(-A)^{1/2-\kappa}D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))\rangle\,\mathrm{d}t\right| \\
&\le \int_{t_m}^{t_{m+1}} C(1+|x|)\,|(P_i-I)(-A)^{-1/2+\kappa}|_{\mathcal{L}(H)}\big(\mathbb{E}|(-A)^{1/2-\kappa}D\tilde{\Psi}^{(M)}(\tilde{X}^h(t))|^2\big)^{1/2}\,\mathrm{d}t \\
&\le C\tau(1+|x|^3)\,|(P_i-I)(-A)^{-1/2+\kappa}|_{\mathcal{L}(H)}.
\end{aligned}$$

Using (3.6) and (3.8), we get

$$|b^{m,h}| \le C\tau(1+|x|^3)h^{1-2\kappa}$$

and

$$|b^{m,M}| \le C\tau(1+|x|^3)\lambda_M^{-1/2+\kappa}.$$

Thus

$$\frac{1}{N\tau}\sum_{m=1}^{N-1}|b^m| \le C(1+|x|^3)(h^{1-2\kappa} + \lambda_M^{-1/2+\kappa})$$

and

$$\limsup_{M\to+\infty}\frac{1}{N\tau}\sum_{m=1}^{N-1}|b^m| \le C(1+|x|^3)h^{1-2\kappa}.$$

7.1.4 *Estimate of* $c^m$. Decompose $c^m$ like $b^m$, i.e., $c^m = c^{m,h} + c^{m,M}$, where for $i \in \{h, M\}$,

$$\begin{aligned}
2|c^{m,i}| &= |\mathbb{E}\int_{t_m}^{t_{m+1}} \mathrm{Tr}\Big((-A)^{2\kappa}(P_i-I)(-A)^{-1/2+\kappa}(-A)^{1/2-\kappa}D^2\tilde{\Psi}^{(M)}(\tilde{X}^h(t))(-A)^{1/2-\kappa}(-A)^{-1/2-\kappa}\Big)\,\mathrm{d}t| \\
&\le \mathrm{Tr}((-A)^{-1/2-\kappa})\,|(-A)^{2\kappa}(P_i-I)(-A)^{-1/2+\kappa}|_{\mathcal{L}(H)} \\
&\quad\times \int_{t_m}^{t_{m+1}} \mathbb{E}|(-A)^{1/2-\kappa}D^2\tilde{\Psi}^{(M)}(\tilde{X}^h(t))(-A)^{1/2-\kappa}|\,\mathrm{d}t.
\end{aligned}$$

Using Assumptions 2.2, Proposition 6.1, Lemma 3.9, commutativity of $A$ and $P_M$ and estimates (3.8) and (3.6), we get

$$2|c^{m,h}| \le C\tau(1+|x|^2)\lambda_M^{-1/2+3\kappa}$$

and

$$2|c^{m,M}| \le C\tau(1+|x|^2)h^{1-6\kappa}.$$

Then, we have

$$\frac{1}{N\tau}\sum_{m=1}^{N-1}|c^m| \le C(1+|x|^2)(h^{1-6\kappa}+\lambda_M^{-1/2+3\kappa})$$

and

$$\limsup_{M\to+\infty}\frac{1}{N\tau}\sum_{m=1}^{N-1}|c^m| \le C(1+|x|^2)h^{1-6\kappa}.$$

7.1.5 *Conclusion.* With the above estimation, we get

$$\limsup_{M\to\infty}\frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}}\mathbb{E}\Big(\mathcal{L}^{(M)}-\mathcal{L}^h\Big)\tilde{\Psi}^{(M)}(\tilde{X}^h(s))\,\mathrm{d}s \le C(1+|x|^3)h^{1-\kappa}(1+T^{-1}). \tag{7.10}$$

## 7.2 *Proof of Lemma 6.5 (additional time-discretization error)*

The error analysed in this section is due to the replacement of $\tilde{X}^h(t)$ with $X_m^h$, for $t_m \le t < t_{m+1}$.

Compared with other error terms, the expression involves the test function $\phi$, instead of $\tilde{\Psi}^{(M)}$. Since $\phi$ is assumed to be of class $\mathcal{C}_b^2$ only, its derivatives do not satisfy estimates like in Proposition 6.1. However, we are still able to distribute appropriately the powers of the operator $-A_h$ to obtain a good rate of convergence.

We define an auxiliary function $\tilde{\phi}_M : H \to \mathbb{R}$ with $\tilde{\phi}_M = \phi \circ P_M$. It is of class $\mathcal{C}_b^2$ and using the identifications introduced in Remark 2.1 we have for any $x \in H$,

$$D\tilde{\phi}_M(x) = P_M D\phi(P_M x),$$
$$D^2\tilde{\phi}_M(x) = P_M D^2\phi(P_M x)P_M.$$

Thanks to the Itô's formula, from (3.13) we get for $t_m \le t < t_{m+1}$,

$$\begin{aligned}\mathbb{E}\phi(P_M\tilde{X}^h(t)) - \mathbb{E}\phi(P_M X_m^h) &= \mathbb{E}\tilde{\phi}_M(\tilde{X}^h(t)) - \mathbb{E}\tilde{\phi}_M(\tilde{X}(t_m))\\ &= \mathbb{E}\int_{t_m}^{t}\langle S_{\tau,h}A_h X_m^h, D\tilde{\phi}_M(\tilde{X}^h(s))\rangle\,\mathrm{d}s\\ &\quad + \mathbb{E}\int_{t_m}^{t}\langle S_{\tau,h}P_h F(X_m^h), D\tilde{\phi}_M(\tilde{X}^h(s))\rangle\,\mathrm{d}s\\ &\quad + \mathbb{E}\int_{t_m}^{t}\frac{1}{2}\mathrm{Tr}((S_{\tau,h}P_h)(S_{\tau,h}P_h)^* D^2\tilde{\phi}_M(\tilde{X}^h(s))\,\mathrm{d}s\\ &= E_1(t)+E_2(t)+E_3(t).\end{aligned}$$

The error is naturally divided into three terms. We first treat the easiest ones: $E_2$ and $E_3$.

Using boundedness of the linear operator $S_{\tau,h}$, of the orthogonal projectors $P_M$ and $P_h$ and of the first-order derivative of $\phi$, and Lipschitz continuity of $F$ and Lemma 3.10, then for $t_m \le t < t_{m+1}$,

$$|E_2(t)| = \left|\mathbb{E}\int_{t_m}^{t}\left\langle S_{\tau,h}P_hF(X_m^h), D\tilde{\phi}_M(\tilde{X}^h(s))\right\rangle \mathrm{d}s\right| \le C(1+|x|)\tau.$$

We now control $E_3(t)$. Using the boundedness of the second-order derivative of $\tilde{\phi}_M$, uniformly with respect to $M$, we have

$$\begin{aligned}|E_3(t)| &\le C(t-t_m)\mathrm{Tr}\big((S_{\tau,h}P_h)(S_{\tau,h}P_h)^*\big)\\ &\le C\tau\mathrm{Tr}\big[((-A_h)^{1/2+\kappa}S_{\tau,h}^2P_h)P_h(-A_h)^{-1/2-\kappa}P_h\big]\\ &\le C\tau|(-A_h)^{1/2+\kappa}S_{\tau,h}^2P_h|_{\mathcal{L}(H)}\mathrm{Tr}\big(P_h(-A_h)^{-1/2-\kappa}P_h\big)\\ &\le C\tau^{1/2-\kappa},\end{aligned}$$

where $\kappa \in (0,1/2)$ is a small parameter, thanks to the first inequality of Lemma 3.7 and to Proposition 3.4.

The treatment of $E_1$ is the most complicated amongst the three terms, due to the presence of the unbounded operator $A_h$. We recall that moments of $|X_m^h|_\alpha$ are controlled uniformly in $h$, only for $\alpha < 1/4$ (Lemma 4.2); to obtain the correct weak order of convergence $1/2$ with respect to $\tau$, we need a careful control. One of the ingredients is the Malliavin integration by parts.

Thanks to (3.11) and (3.12), $E_1$ is divided into three parts: $E_1(t) = E_{1,1}(t) + E_{1,2}(t) + E_{1,3}(t)$, such that, for $t_m \le t < t_{m+1}$,

$$\begin{aligned}E_{1,1}(t) &= \mathbb{E}\int_{t_m}^{t}\langle S_{\tau,h}^{m+1}A_hP_hx, D\tilde{\phi}_M(\tilde{X}^h(s))\rangle\,\mathrm{d}s,\\ E_{1,2}(t) &= \mathbb{E}\int_{t_m}^{t}\left\langle \tau A_hS_{\tau,h}\sum_{k=0}^{m-1}S_{\tau,h}^{m-k}P_hF(X_m^h), D\tilde{\phi}_M(\tilde{X}^h(s))\right\rangle \mathrm{d}s,\\ E_{1,3}(t) &= \mathbb{E}\int_{t_m}^{t}\left\langle A_hS_{\tau,h}\int_0^{t_m}S_{\tau,h}^{m-l_r}P_h\,\mathrm{d}W(r), D\tilde{\phi}_M(\tilde{X}^h(s))\right\rangle \mathrm{d}s.\end{aligned}$$

We have isolated the stochastic part in $X_m^h$; then only the treatment of $E_{1,3}(t)$ is difficult.

First, using Lemma 3.7, we have if $m \ge 1$,

$$\begin{aligned}|A_hS_{\tau,h}^{m+1}P_hx| &\le |(-A_h)^\kappa S_{\tau,h}P_h|_{\mathcal{L}(H)}|(-A_h)^{1-\kappa}S_{\tau,h}^mP_h|_{\mathcal{L}(H)}|x|_H\\ &\le C|x|_H\tau^{-\kappa}t_m^{-1+\kappa}.\end{aligned}$$

As a consequence, for $t_m \le t < t_{m+1}$,

$$|E_{1,1}(t)| \le C|x|\frac{\tau^{1-\kappa}}{t_m^{1-\kappa}}.$$

The treatment of $E_{1,2}$ is similar: we have when $m \geq 1$,

$$\mathbb{E}\left|\tau A_h S_{\tau,h} \sum_{k=0}^{m-1} S_{\tau,h}^{m-k} P_h F(X_k^h)\right| \leq C\tau |(-A_h)^\kappa S_{\tau,h} P_h|_{\mathcal{L}(H)} \sum_{k=0}^{m-1} |(-A_h)^{1-\kappa} S_{\tau,h}^{m-k} P_h|_{\mathcal{L}(H)} \mathbb{E}|F(X_k^h)|_H$$
$$\leq C(1+|x|)\tau^{-\kappa}\tau \sum_{k=0}^{m-1} |(-A_h)^{1-\kappa} S_{\tau,h}^{m-k} P_h|_{\mathcal{L}(H)},$$

thanks to Lipschitz continuity of $F$ and Lemma 3.10. Then using Lemma 3.7 and inequality (7.1), we obtain for $m \geq 1$ and $t_m \leq t < t_{m+1}$,

$$|E_{1,2}(t)| \leq C\tau^{-\kappa}(t - t_m) \leq C\tau^{1-\kappa}.$$

It remains to control $E_{1,3}(t)$, which contains the stochastic term, with low regularity properties. We use a Malliavin integration by parts formula, with the same decomposition of the integral as for $a_{1,3}^{m,h}$: for any $t_m \leq s \leq t < t_{m+1}$,

$$\mathbb{E}\left\langle A_h S_{\tau,h} \int_0^{t_m} S_{\tau,h}^{m-l_r} \mathrm{d}W(r), D\tilde{\phi}_M(\tilde{X}^h(s))\right\rangle = \mathbb{E}\left\langle A_h S_{\tau,h} \int_0^{(t_m-3\tau_0)\vee 0} S_{\tau,h}^{m-l_r} P_h \,\mathrm{d}W(r), D\tilde{\phi}_M(\tilde{X}^h(s))\right\rangle$$
$$+ \mathbb{E}\left\langle A_h S_{\tau,h} \int_{(t_m-3\tau_0)\vee 0}^{t_m} S_{\tau,h}^{m-l_r} P_h \,\mathrm{d}W(r), D\tilde{\phi}_M(\tilde{X}^h(s))\right\rangle$$
$$= E_{1,3,1}(s,t) + E_{1,3,2}(s,t).$$

For the first error term, we directly use the Cauchy–Schwarz inequality and we have (see term $a_{1,3}^{m,h}$ of Section 7.1 for more details)

$$|E_{1,3,1}(s,t)|^2 \leq C\left(E\left|\int_0^{(t_m-3\tau_0)\vee 0} S_{\tau,h} A_h S_{\tau,h}^{m-l_r} P_h \,\mathrm{d}W(r)\right|^2\right)\left(\mathbb{E}\left|D\tilde{\phi}_M(\tilde{X}^h(s))\right|^2\right)$$
$$\leq C\int_0^{(t_m-3\tau_0)\vee 0} \mathrm{Tr}\big(P_h A_h S_{\tau,h}^{(m-l_r)+1} S_{\tau,h}^{(m-l_r)+1} A_h P_h\big)\,\mathrm{d}r$$
$$\leq C\int_0^{(t_m-3\tau_0)\vee 0} \mathrm{Tr}(P_h(-A_h)^{-1/2-\kappa}P_h)\,|(-A_h)^{5/2+\kappa} S_{\tau,h}^{2(m-l_r)+1} P_h|_{\mathcal{L}(H)}\,\mathrm{d}r$$
$$\leq C.$$

For the second error term, using the Malliavin integration by parts formula (Lemma 6.6), we get for any $t_m \leq s \leq t < t_{m+1}$,

$$\mathbb{E}\left\langle A_h S_{\tau,h} \int_{(t_m-3\tau_0)\vee 0}^{t_m} S_{\tau,h}^{m-l_r} P_h \,\mathrm{d}W(r), D\tilde{\phi}_M(\tilde{X}^h(s))\right\rangle$$
$$= \mathbb{E}\int_{(t_m-3\tau_0)\vee 0}^{t_m} \mathrm{Tr}\Big(S_{\tau,h}^{m-l_r} A_h S_{\tau,h} P_h D^2\tilde{\phi}_M(\tilde{X}^h(s))\mathcal{D}_r\tilde{X}^h(s)\Big)\,\mathrm{d}r.$$

Then

$$\left|\mathbb{E}\int_{(t_m-3\tau_0)\vee 0}^{t_m} \mathrm{Tr}\Big(S_{\tau,h}^{m-l_r}A_hS_{\tau,h}P_hD^2\tilde{\phi}_M(\tilde{X}^h(s))\mathcal{D}_r\tilde{X}^h(s)\Big)\,\mathrm{d}r\right|$$
$$\leq \int_{(t_m-3\tau_0)\vee 0}^{t_m} \mathrm{Tr}\Big((-A_h)S_{\tau,h}^{m-l_r}S_{\tau,h}P_h\Big)\mathbb{E}[|\mathcal{D}_r\tilde{X}^h(s)|_{\mathcal{L}(H)}|D^2\tilde{\phi}_M(\tilde{X}^h(s))|]\,\mathrm{d}r.$$

Since

$$\mathrm{Tr}((-A_h)S_{\tau,h}^{m-l_r}S_{\tau,h}P_h) \leq \mathrm{Tr}(P_h(-A_h)^{-1/2-\kappa}P_h)\,|S_{\tau,h}^{m-l_r}\big((-A_h)^{3/2+\kappa}S_{\tau,h}P_h\big)|_{\mathcal{L}(H)},$$

we have (see Section 7.1 for more details on a similar expression), using Lemma 3.7,

$$\left|\mathbb{E}\int_{(t_m-3\tau_0)\vee 0}^{t_m} \mathrm{Tr}(S_{\tau,h}^{m-l_r}A_hS_{\tau,h}P_h\mathcal{D}_r\tilde{X}^h(s)D^2\tilde{\phi}_M(\tilde{X}^h(s)))\,\mathrm{d}r\right|$$
$$\leq C\tau^{-1/2-2\kappa}\int_{(t_m-3\tau_0)\vee 0}^{t_m} \frac{1}{(1+\lambda_0\tau)^{m-l_r}}(1+L_F\tau)^{m-l_r}\left(1+\frac{1}{(1+\lambda_0\tau)^{\kappa(m-l_r)}t_{m-l_r}^{1-\kappa}}\right)\mathrm{d}r.$$

Using that $(1+L_F\tau)^{m-l_r} \leq C$ for the range of $r$ used to compute the integral, we see that

$$|E_{1,3,2}(s,t)| \leq C\tau^{-1/2-2\kappa}.$$

After integration with respect to $s$, we obtain

$$|E_{1,3}(t)| \leq \int_{t_m}^{t}(|E_{1,3,1}(s,t)|+|E_{1,3,2}(s,t)|)\,\mathrm{d}s \leq C(\tau+\tau^{1/2-2\kappa}),$$

and

$$|E_1(t)| \leq C\left(\tau^{1/2-2\kappa}+|x|\frac{\tau^{1-\kappa}}{t_m^{1-\kappa}}+\tau^{1-\kappa}\right).$$

Using the bounds on $E_2$ and $E_3$, we therefore obtain that when $m \geq 1$ and $t_m \leq t \leq t_{m+1}$,

$$|\mathbb{E}\phi(P_M\tilde{X}^h(t))-\mathbb{E}\phi(P_MX_m^h)| \leq C\tau^{1/2-2\kappa}\left(1+\frac{|x|}{(m\tau)^{1-\kappa}}\right).$$

As a consequence, we obtain

$$\left|\frac{1}{N\tau}\sum_{m=1}^{N-1}\int_{t_m}^{t_{m+1}}\Big(\mathbb{E}\phi(P_M\tilde{X}^h(t))-\mathbb{E}\phi(P_MX_m^h)\Big)\,\mathrm{d}t\right|$$
$$\leq C\tau^{1/2-2\kappa}\left(1+|x|\frac{1}{N\tau}\int_0^{N\tau}\frac{1}{t^{1-\kappa}}\,\mathrm{d}t\right)$$
$$\leq C\tau^{1/2-2\kappa}\left(1+\frac{|x|}{(N\tau)^{1-\kappa}}\right). \tag{7.11}$$

### 7.3 *Conclusion*

With (7.10) (which proves Lemma 6.3), (7.11) (which proves Lemma 6.5) and Lemma 6.4 (proof is omited), we get

$$\frac{1}{N}\sum_{m=0}^{N-1}\mathbb{E}\Big(\phi(X_m^h)-\overline{\phi}\Big) \leq C(1+|x|^3)\tau^{1/2-\kappa}(1+T^{-1+\kappa}+T^{-1})(1+h^{1-\kappa}),$$

where $C$ does not depend on $T$, $h$ and $M$.

**Acknowledgements**

The authors would like to thank the referees for their remarks which helped us improve the presentation of the article.



## References

Andersson, A. & Larsson, S. (2016) Weak convergence for a spatial approximation of the nonlinear stochastic heat equation. *Math. Comp.*, **85**, 1335–1358.

Bréhier, C.-E. (2012) Strong and weak orders in averaging for SPDEs. *Stochastic Processes Appl.*, **122**, 2553–2593.

Bréhier, C.-E. (2014) Approximation of the invariant measure via an Euler scheme for stochastic PDEs driven by space-time white noise. *Potential Anal.*, **40**, 1–40.

Bréhier, C.-E. & Vilmart, G. (2015) High-order integrator for sampling the invariant distribution of a class of parabolic SPDEs with additive space-time noise. To appear in Siam Journal on Scientific Computing, 2016+; preprint http://arxiv.org/abs/1505.05061

Brézis, H. (1994) *Functional Analysis. Theory and Applications. (Analyse Fonctionnelle. Théorie et Applications.) Collection Mathématiques Appliquées pour la Maîtrise*. Paris: Masson. 248p.

Cerrai, S. & Freidlin, M. (2009) Averaging principle for a class of stochastic reaction-diffusion equations. *Probab. Theory Relat. Fields*, **144**, 137–177.

Ciarlet, P. G. (2002) *The Finite Element Method for Elliptic Problems*. Classics in Applied Mathematics, vol. 40. Amsterdam: North-Holland. xxviii, 530p.

Da Prato, G. & Zabczyk, J. (1992) *Stochastic Equations in Infinite Dimensions*. Encyclopedia of Mathematics and Its Applications, vol. 44. Cambridge: Cambridge University Press. xviii, 454p.

Da Prato, G. & Zabczyk, J. (1996) *Ergodicity for Infinite Dimensional Systems*. London Mathematical Society Lecture Note Series, vol. 229. Cambridge: Cambridge University Press. xi, 339p.

De Bouard, A. & Debussche, A. (2006) Weak and strong order of convergence of a semidiscrete scheme for the stochastic nonlinear Schrödinger equation. *Appl. Math. Optim.*, **54**, 369–399.

Debussche, A. (2011) Weak approximation of stochastic partial differential equations: the nonlinear case. *Math. Comput.*, **80**, 89–117.

Debussche, A. (2013) Ergodicity results for the stochastic Navier–Stokes equations: an introduction. *Topics in Mathematical Fluid Mechanics*. Lecture Notes in Mathematics, vol. 2073. Heidelberg: Springer, pp. 23–108.

Debussche, A., Hu, Y. & Tessitore, G. (2011) Ergodic BSDEs under weak dissipative assumptions. *Stochastic Processes Appl.*, **121**, 407–426.

Debussche, A. & Printems, J. (2009) Weak order for the discretization of the stochastic heat equation. *Math. Comput.*, **78**, 845–863.

Doeblin, W. (1938) Exposé de la théorie des chaînes simples constantes de Markoff à un nombre fini détats. *Rev. Math. Union Interbalkan.*, **2**, 77–105.

Ern, A. & Guermond, J. L. (2004) *Theory and Practice of Finite Elements*. Applied Mathematical Sciences, vol. 159. New York: Springer. xiv, 524p.

HIGHAM, D. J., MATTINGLY, J. C. & STUART, A. M. (2002) Ergodicity for SDEs and approximations: locally Lipschitz vector fields and degenerate noise. *Stochastic Processes Appl.*, **101**, 185–232.

KOPEC, M. (2014) Quelques contributions à l'analyse numérique d'équations stochastiques. *Ph.D. Thesis*, ENS Rennes.

KOVACS, M., LARSSON, S. & LINDGREN, F. (2010) Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise. *Numer. Algorithms*, **53**, 309–320.

KOVACS, M., LARSSON, S. & LINDGREN, F. (2012) Weak convergence of finite element approximations of linear stochastic evolution equations with additive noise. *BIT Numer. Math.*, **52**, 85–108.

KRUSE, R. (2013) Optimal error estimates of Galerkin finite element methods for stochastic partial differential equations with multiplicative noise. *IMA J. Num. Anal.*, **34**, 217–251.

KUKSIN, S. & SHIRIKYAN, A. (2001) A coupling approach to randomly forced nonlinear PDE's. I. *Commun. Math. Phys.*, **221**, 351–366.

LINDVALL, E. T. (1992) *Lectures on the Coupling Method*. Wiley Series in Probability and Mathematical Statistics. Applied Probability and Statistics. New York: Wiley. 272p.

MATTINGLY, J. C. (2002) Exponential convergence for the stochastically forced Navier–Stokes equations and other partially dissipative dynamics. *Commun. Math. Phys.*, **230**, 421–462.

MATTINGLY, J. C., STUART, A. M. & TRETYAKOV, M. V. (2010) Convergence of numerical time-averaging and stationary measures via Poisson equations. *SIAM J. Numer. Anal.*, **48**, 552–577.

MEYN, S. & TWEEDIE, R. L. (2009) *Markov Chains and Stochastic Stability*. Cambridge: Cambridge University Press. xviii, 594p.

MUELLER, C. (1993) Coupling and invariant measures for the heat equation with noise. *Ann. Probab.*, **21**, 2189–2199.

NUALART, D. (2006) *The Malliavin Calculus and Related Topics*, 2nd edn. Probability and Its Applications. Berlin: Springer. xiv, 382p.

SANZ-SOLÉ, M. (2005) *Malliavin Calculus with Applications to Stochastic Partial Differential Equations*. Fundamental Sciences: Mathematics. Boca Raton, FL: CRC Press. viii, 162p.

TALAY, D. (1990) Second-order discretization schemes of stochastic differential systems for the computation of the invariant law. *Stochastics Rep.*, **29**, 13–36.

THOMÉE, V. (2006) *Galerkin Finite Element Methods for Parabolic Problems*, 2nd edn. Springer Series in Computational Mathematics, vol. 25. Berlin: Springer. xii, 370p.

WANG, X. & GAN, S. (2013) Weak convergence analysis of the linear implicit Euler method for semilinear stochastic partial differential equations with additive noise. *J. Math. Anal. Appl.*, **398**, 151–169.

YAN, Y. (2005) Galerkin finite element methods for stochastic parabolic partial differential equations. *SIAM J. Numer. Anal.*, **43**, 136–1384.