

A FINITE ELEMENT METHOD FOR MARTINGALE-DRIVEN STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS

ANDREA BARTH

ABSTRACT. The main objective of this work is to describe a Galerkin approximation for stochastic partial differential equations driven by square-integrable martingales. Error estimates in the semidiscrete case, where discretization is only done in space, and in the fully discrete case are derived. Parabolic as well as transport equations are studied.

1. Introduction

Arguably the modern literature on Finite Elements Methods (FEM's from this point on) can be traced back to the 1956 paper of Turner, Clough, Martin and Topp [26]. Several contemporary publications followed [1] [17], mostly under an Engineering scope. This should come as no surprise given the aerospace background of the authors. It was not until the early 1970's that research on FEM's applied to the estimation of solutions to partial differential equations picked up steam. Here one finds, among others, the works of Fujita and Mizutani [11], Ushijima [27] and Zlámal [30] [31] on parabolic PDE's, together with those of Ciarlet [8] and Nedoma [19] for elliptic problems. From this point on the stream of related publications swelled in breadth and depth, from various forms of approximations (Galerkin, Riez-Galerkin, Lagrange-Galerkin, etc.) to different methodologies (energy methods, dynamic FEM's, etc.) and a more focused treatment of specific problems (notably Navier-Stokes equations). Although beyond the scope of this paper, it should be mentioned that the theory of convergence of finite-elements approximations developed in parallel fashion.

Unsurprisingly, the (numerical) study of stochastic partial differential equations (SPDE's from this point on) took more time to develop. It was not until the mid-to-late 1990's that work such as that of Gyöngy and Nualart [13], Yoo [29], Crisan, Gaines and Lyons [9], and Gaines [12] started breaking trail in this direction. However, none of these papers present a FEM approach. Albeit more complex to implement, FEM's do present the advantage of greater degrees of freedom when choosing a discretization of the space-time continuum. This can be particularly useful to design finite elements tailor made to the characteristics of a specific problem. To our knowledge, the 2003 paper by Yan [28] was the first to employ FEM's to estimate solutions to SPDE's, where error estimates for linear equations

Received 2009-10-12; Communicated by Hui-Hsiung Kuo.

2000 *Mathematics Subject Classification.* 60H15; 60H35.

Key words and phrases. Finite element method, stochastic partial differential equation, Lévy process, martingale.

with a Brownian motion as driving noise were studied. Similar estimates for some parabolic equations and Zakai's equation were presented by Chow and Jiang in [6] and Chow, Jiang and Menaldi in [7] respectively. The reader is directed to [5] [10] [21] [22] for a thorough treatment of Hilbert-space valued stochastic equations.

In this paper we use Galerkin approximations to estimate the solution of some SPDE's driven by square-integrable martingales. Our main aim is to provide error estimates in several cases. In a first instance we look into a parabolic equation. Namely, we analyze approximations of the solution to

$$dX(t) = AX(t)dt + G(X(t))dM(t), \quad X(t_0) = X_0, \quad (1.1)$$

where M is a square-integrable martingale defined on some probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ that takes values in a Hilbert space U ; A denotes a second order differential operator acting on a Hilbert space H , which is not necessarily equal to U . For $t \in [0, T]$, $X(t)$ is a H -valued random process. Infinite dimensional equations of this kind appear in financial Mathematics and Physics among other disciplines.

On the other hand we study some transport equations that appear, for instance, in the modeling of bond markets: The price of a zero-coupon bond at time t is given by $P(t, \theta)$, $0 \leq t \leq \theta$. This is an instrument that delivers 1 dollar at time θ . $P(t, \theta)$ is given by the equation

$$P(t, \theta) = \exp\left(-\int_t^\theta f(t, s)ds\right),$$

and $f(t, s)$, $0 \leq t \leq s$, is the so-called forward rate. The latter depends on two parameters: time and time to maturity. For a general introduction to interest rate models in infinite dimensions we refer the reader to [4]. The widely used Heath-Jarrow-Morton approach to interest rate modeling assumes the forward rate satisfies the equation

$$df(t, \theta) = \mu(t, \theta)dt + (\sigma(t, \theta), dZ(t))_U, \quad \theta \geq t.$$

Here Z denotes a Lévy process defined on a probability space $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ that takes values in a Hilbert space $(U, (\cdot, \cdot)_U)$, and $\mu(t, \theta)$ and $\sigma(t, \theta)$ are predictable processes for each $\theta \geq 0$. A common practice in interest rate theory is to parameterize forward rates by time to maturity. Arguably this approach was introduced by Musiela in [18]. This paper provides a link between stochastic partial differential equations and interest rate theory. To embed our setting into such framework we require the following definitions: For $t, \xi \geq 0$ and $u \in U$ let

$$r(t)(\xi) := f(t, t + \xi), \quad a(t)(\xi) := \mu(t, t + \xi) \text{ and } (g(t)u)(\xi) := (\sigma(t, t + \xi), u)_U.$$

We may then write

$$\begin{aligned} r(t)(\xi) &= r(0)(t + \xi) + \int_0^t a(s)(t - s + \xi)ds + \int_0^t g(s)(t - s + \xi)dZ(s) \\ &= S(t)r(0)(\xi) + \int_0^t S(t - s)a(s)(\xi)ds + \int_0^t S(t - s)g(s)(\xi)dZ(s), \end{aligned}$$

where S denotes the shift semigroup $S(t)\phi(\xi) = \phi(\xi + t)$. The equation above is the mild solution to the equation

$$dr(t) = \left(\frac{\partial}{\partial \xi}r(t) + a(t)\right)dt + g(t)dZ(t), \quad (1.2)$$

which is the well known Musiela parametrization. As is the case for models whose randomness stems from finite dimensional Wiener processes, a has to fulfill a HJM–drift condition (see for details [21] [16] [3]). The latter is required to have an arbitrage–free model. If the volatility g depends on r , we obtain the HJM–Musiela equation

$$dr(t)(\xi) = \left(\frac{\partial}{\partial \xi}r(t)(\xi) + F(t, r(t))(\xi)\right)dt + G(t, r(t))(\xi)dZ(t).$$

The existence of solutions to Equation (1.2) is studied in [21]. The same type of equation can be derived for energy forwards (see [2]), where the non–hedgeable underlying is the price of electricity instead of the interest rate.

These transport equations can also be stated in the terminology of differential operators as

$$dX(t) = BX(t)dt + G(X(t))dM(t), \quad X(t_0) = X_0. \quad (1.3)$$

The essential difference between Equation (1.1) and the one above is that in the former case the differential operator A generates an analytic semigroup, whereas in the latter B generates a C_0 –semigroup of contractions.

Given the rapid growth of the fixed income and electricity markets (among other possible applications), the simulation of SPDE's with Lévy–Noise is an interesting and current problem. We derive estimates for the mean–square error of a Galerkin approximation for parabolic and hyperbolic equations.

The remainder of this paper is structured as follows: Section 2 contains a detailed analysis of Equation (1.1) and Equation (1.3). In Section 3 we introduce the discretization schemes. Then we compute error estimates for the semidiscrete, i.e. discrete in space, and the discrete equations, i.e. discrete in time and space, for the parabolic equation. We borrow some results for the deterministic equations from [25]. Finally we derive error estimates for first order hyperbolic stochastic differential equations (stochastic transport equation), using results from [15].

2. Preliminaries

Throughout this paper $(U, (\cdot, \cdot)_U)$ will be a Hilbert space and $(\Omega, \mathcal{F}, (\mathcal{F}_t), \mathbb{P})$ will denote a filtered probability space satisfying the “usual conditions”. The space of all càdlàg square integrable (\mathcal{F}_t) –martingales taking values in U will be represented by $\mathcal{M}^2(U)$. The spatial covariance structure of $M \in \mathcal{M}^2(U)$ is implied by a covariance operator $Q \in L_1^+(U)$, the space of all symmetric, non-negative and nuclear operators. We restrict ourselves to the following class of square integrable martingales

$$\mathcal{C} := \{M \in \mathcal{M}^2(U) : \exists Q \in L_1^+ \text{ such that} \\ \forall t \geq s \geq 0, \langle \langle M, M \rangle \rangle_t - \langle \langle M, M \rangle \rangle_s \leq (t - s)Q\}.$$

It is a known result (see [24]) that if $Q \in L_1^+(U)$ there exists an orthonormal basis $\{e_n\}$ of U consisting of eigenvectors of Q . This leads to the representation $Qe_n = \gamma_n e_n$, where γ_n is the eigenvalue corresponding to e_n . The square root of Q is defined as

$$Q^{\frac{1}{2}}x := \sum_n (x, e_n)_U \gamma_n^{\frac{1}{2}} e_n \quad x \in U,$$

and $Q^{-\frac{1}{2}}$ will be the pseudo inverse to $Q^{\frac{1}{2}}$.

Let $(\mathcal{H}, (\cdot, \cdot)_{\mathcal{H}})$ be the Hilbert space defined by $\mathcal{H} = Q^{\frac{1}{2}}(U)$ endowed with the inner product $(x, y)_{\mathcal{H}} = (Q^{-\frac{1}{2}}x, Q^{-\frac{1}{2}}y)_U$ for $x, y \in \mathcal{H}$. We write $L_{HS}(\mathcal{H}, H)$ to refer to the space of all Hilbert-Schmidt operators from \mathcal{H} to some Hilbert space H . We use the following proposition repeatedly, its proof can be found in [21]. Let $\mathcal{P}_{[0, T]}$ denote the σ -field of predictable sets in $\Omega \times [0, T]$.

Proposition 2.1. *Let $\mathbb{L}_{\mathcal{H}, T}^2(H) = L^2(\Omega \times [0, T], \mathcal{P}_{[0, T]}, \mathbb{P}dt; L_{HS}(\mathcal{H}, H))$ be the space of integrands. Then for every $X \in \mathbb{L}_{\mathcal{H}, T}^2(H)$*

$$\mathbb{E} \left\| \int_0^t X(s) dM(s) \right\|_H^2 \leq \mathbb{E} \int_0^t \|X(s)\|_{L_{HS}(\mathcal{H}, H)}^2 ds. \tag{2.1}$$

As we mentioned previously, the two equations that concern us are

$$dX(t) = AX(t)dt + G(X(t))dM(t), \quad X(0) = X_0, \tag{2.2}$$

where A generates an analytic semigroup (parabolic case) S on the Hilbert space $H = L^2(\mathcal{D})$ and $\mathcal{D} \subset \mathbb{R}^d$ has a smooth boundary $\partial\mathcal{D}$ and

$$dX(t) = BX(t)dt + G(X(t))dM(t), \quad X(0) = X_0 \tag{2.3}$$

where B generates a C_0 -semigroup (hyperbolic case) S on H . We denote by $D(A)$ and $D(B)$ the domains of the differential operators, and by $D(G)$ the domain of G , which takes values in $L(\mathcal{H}, H)$. H^α denotes the Sobolev space of order α endowed with the corresponding norm $\|\cdot\|_{H^\alpha}$ for $\alpha > 0$.

Let $M \in \mathcal{M}^2(U)$, then Equations (2.2) and (2.3) are well defined if the following conditions are satisfied:

- (1) X_0 is an \mathcal{F}_0 -measurable random variable with values in H .
- (2) $D(G)$ is dense in H .
- (3) There exists a function $b : \mathbb{R}_+ \rightarrow \mathbb{R}_+$ with $\int_0^T b^2(t) dt < \infty$ for all $T < \infty$, such that:
 - (a) $\|S(t)G(x)\|_{L_{HS}(\mathcal{H}, H)} \leq b(t)(1 + \|x\|_H)$ for all $t > 0$ and $x \in D(G)$.
 - (b) $\|S(t)(G(x) - G(y))\|_{L_{HS}(\mathcal{H}, H)} \leq b(t)\|x - y\|_H$ for all $t > 0$ and $x, y \in D(G)$.
 - (c) $\|A^\alpha G(x)\|_{L_{HS}(\mathcal{H}, H)} \leq C(\|x\|_{H^{2\alpha}})$ for $\alpha \in \{0, 1/2, 1\}$ and $x \in H^{2\alpha}$.
 - (d) $\|B^\alpha G(x)\|_{L_{HS}(\mathcal{H}, H)} \leq C(\|x\|_{H^\alpha})$ for $\alpha \in \{0, 1\}$ and $x \in H^\alpha$.

Note that the properties above also hold if we take $S(t) = I$.

Definition 2.2. Let X_0 be a \mathcal{F}_0 -measurable, square integrable random variable with values in H . A predictable process $X : \mathbb{R}_+ \times \Omega \rightarrow H$ is called a *mild solution* to Equation (2.2) or (2.3) if

$$\sup_{t \in [0, T]} \mathbb{E}|X(t)|_H^2 < \infty \quad \forall T \in (0, \infty)$$

and

$$X(t) = S(t)X_0 + \int_0^t S(t-s)G(X(s))dM(s) \quad \forall t > 0. \tag{2.4}$$

The conditions above imply that the integral in Equation (2.4) is well defined. If $\mathbb{E}\|X_0\|_{H^2}^2 < \infty$, then the solution $X(t)$, for $t \in [0, T]$, will be element of H^2 as well.

Definition 2.3. Let X_0 be a \mathcal{F}_0 -measurable, square integrable random variable with values in H . A predictable process $X : \mathbb{R}_+ \times \Omega \rightarrow H$ is called a *weak solution* to (2.3) if

$$\sup_{t \in [0, T]} \mathbb{E}\|X(t)\|_H^2 < \infty \quad \forall T \in (0, \infty)$$

and for all $a \in D(B^*)$

$$(a, X(t))_H = (a, X_0)_H + \int_0^t (B^*a, X(s))_H ds + \int_0^t (G^*(X(s))a, dM(s))_{\mathcal{H}} \quad \forall t > 0. \tag{2.5}$$

Here $D(B^*)$ denotes the domain of the adjoint operator B^* . It can be shown that X is a mild solution if and only if it is a weak solution if the following conditions are fulfilled: For $a \in D(B^*)$ exists $c(a) < \infty$ such that

$$\|G^*(x)a\|_{\mathcal{H}} \leq c(a)(1 + \|x\|_H)$$

and

$$\|(G^*(x) - G^*(y))a\|_{\mathcal{H}} \leq c(a)\|x - y\|_H$$

for all $x, y \in D(G)$.

However, for the weak solution regularity can not be achieved so easily as for the mild solution. For the time being we assume $\sup_{t \in [0, T]} \mathbb{E}\|X(t)\|_{H^2}^2 < \infty$.

We quote below Theorem 6.13 from [20] since it is essential for most proofs in this section.

Theorem 2.4. *Let A be the infinitesimal generator of a semigroup $T(t)$. If $0 \in \rho(A)$ then*

- (1) $T(t) : X \rightarrow D(A^\alpha)$ for every $t > 0$ and $\alpha \geq 0$.
- (2) For every $x \in D(A^\alpha)$ we have $T(t)A^\alpha x = A^\alpha T(t)x$.
- (3) For every $t, \delta > 0$ the operator $A^\alpha T(t)$ is bounded and

$$\|A^\alpha T(t)\| < C_\alpha t^{-\alpha} e^{-\delta t}.$$

- (4) Let $0 < \alpha \leq 1$ and $x \in D(A^\alpha)$ then

$$\|T(t)x - x\| < C_\alpha t^\alpha \|A^\alpha x\|.$$

In the following sections we need a result for the regularity of the mild solution when it is discretized in time.

Lemma 2.5. *If X is the mild solution to (2.2) or (2.3), then we have for $0 \leq t_1 \leq t_2 \leq T$ and $\alpha \in \{1, 2\}$*

$$\mathbb{E}\|X(t_2) - X(t_1)\|_H^2 \leq C(t_2 - t_1)(\mathbb{E}\|X_0\|_{H^\alpha}^2 + \sup_{s \in [0, t_2]} \mathbb{E}\|X(s)\|_{H^\alpha}^2). \tag{2.6}$$

Proof. To ease notation we write R as a proxy for operators A and B , both of which generate a C_0 -semigroup. The general mild solution is given by

$$X(t) = S(t)X_0 + \int_0^t S(t-s)G(X(s))dM(s).$$

The regularity is provided by

$$\begin{aligned} \mathbb{E}\|X(t_2) - X(t_1)\|_H^2 &\leq C(\mathbb{E}\|(S(t_2) - S(t_1))X_0\|_H^2 \\ &\quad + \mathbb{E}\|\int_0^{t_1} (S(t_2 - s) - S(t_1 - s))G(X(s))dM(s)\|_H^2 \\ &\quad + \mathbb{E}\|\int_{t_1}^{t_2} S(t_2 - s)G(X(s))dM(s)\|_H^2). \end{aligned}$$

Theorem 2.4, resp. the basic properties of a C_0 -semigroup, implies for the first term that

$$\begin{aligned} \mathbb{E}\|(S(t_2) - S(t_1))X_0\|_H^2 &= \mathbb{E}\|S(t_1)(S(t_2 - t_1) - I)X_0\|_H^2 \\ &\leq C(t_2 - t_1) \mathbb{E}\|RX_0\|_H^2 \\ &\leq C(t_2 - t_1) \mathbb{E}\|X_0\|_{H^\alpha}^2. \end{aligned}$$

To bound the second term we use Theorem 2.4 and Equation (2.1):

$$\begin{aligned} \mathbb{E}\|\int_0^{t_1} (S(t_2 - s) - S(t_1 - s))G(X(s))dM(s)\|_H^2 &\leq \mathbb{E}\int_0^{t_1} \|R(S(t_2 - s) - S(t_1 - s))G(X(s))\|_{L_{HS}(\mathcal{H},H)}^2 ds \\ &\leq \mathbb{E}\int_0^{t_1} \|S(t_1 - s)R(S(t_2 - t_1) - I)G(X(s))\|_{L_{HS}(\mathcal{H},H)}^2 ds \\ &\leq C(t_2 - t_1) \mathbb{E}\int_0^{t_1} \|RG(X(s))\|_{L_{HS}(\mathcal{H},H)}^2 ds \\ &\leq C(t_2 - t_1) \mathbb{E}\int_0^{t_1} \|X(s)\|_{H^\alpha}^2 ds \\ &\leq C(t_2 - t_1) t_1 \sup_{s \in [0, t_1]} \mathbb{E}\|X(s)\|_{H^\alpha}^2. \end{aligned}$$

The last term is dominated by

$$\begin{aligned} \mathbb{E}\|\int_{t_1}^{t_2} S(t_2 - s)G(X(s))dM(s)\|_H^2 &\leq \mathbb{E}\int_{t_1}^{t_2} \|S(t_2 - s)RG(X(s))\|_{L_{HS}(\mathcal{H},H)}^2 ds \\ &\leq \mathbb{E}\int_{t_1}^{t_2} \|S(t_2 - s)\|^2 \|RG(X(s))\|_{L_{HS}(\mathcal{H},H)}^2 ds \\ &\leq C(t_2 - t_1) \sup_{s \in [t_1, t_2]} \mathbb{E}\|X(s)\|_{H^\alpha}^2. \end{aligned}$$

This concludes the proof, since

$$\begin{aligned} \mathbb{E}\|X(t_2) - X(t_1)\|_H^2 &\leq C(t_2 - t_1) \mathbb{E}\|X_0\|_{H^\alpha}^2 + C(t_2 - t_1) \sup_{s \in [0, t_1]} \mathbb{E}\|X(s)\|_{H^\alpha}^2 \\ &\quad + C(t_2 - t_1) \sup_{s \in [t_1, t_2]} \mathbb{E}\|X(s)\|_{H^\alpha}^2 \\ &\leq C(t_2 - t_1) (\mathbb{E}\|X_0\|_{H^\alpha}^2 + \sup_{s \in [0, t_2]} \mathbb{E}\|X(s)\|_{H^\alpha}^2). \end{aligned}$$

The case $\alpha = 1$ refers here to the hyperbolic equation and $\alpha = 2$ to the parabolic equation. □

3. Finite Element Method for Parabolic SPDE's

In this section we study the case $A = \Delta$, the Laplace operator with Dirichlet boundary conditions. This is done for notational simplicity. The results contained herein are valid for a broader set of second order differential operators. Namely, those that posses semidiscrete analogs of the form $A_h = P_h A P_h$. The notions of A_h and P_h are fully discussed below.

We follow the error estimates for the Galerkin approximation for non-smooth data described in [25]. The solution operator for the SPDE has a smoothing property that results in a regular solution even if the initial data is not so. We also work in the space $\dot{H}^s(\mathcal{D}) = D(A^{\frac{s}{2}})$ endowed with the norm $\|\cdot\|_{\dot{H}^s} = \|A^{\frac{s}{2}} \cdot\|$ for $s \in \mathbb{R}$ as introduced in [25].

3.1. Error estimates in the semidiscrete case. Let $\{\mathcal{T}_h\}$ be a family of triangulations of \mathcal{D} , indexed by the maximal edge length h . For each \mathcal{T}_h we construct a finite element space \mathcal{S}_h , consisting of piecewise continuous polynomials. Furthermore, we assume $\{\mathcal{S}_h\} \subset H_0^1(\mathcal{D})$. The semidiscrete problem to (2.2) is then to find $X_h(t) \in \mathcal{S}_h$ such that for $t \in [0, T]$

$$dX_h(t) = A_h X_h(t) dt + P_h G(X_h(t)) dM(t), \quad X_h(0) = P_h X_0. \tag{3.1}$$

Here $A_h : \mathcal{S}_h \rightarrow \mathcal{S}_h$ denotes the discrete Laplacian operator defined by

$$(A_h \phi, \psi) = -(\nabla \phi, \nabla \psi) \quad \forall \phi, \psi \in \mathcal{S}_h,$$

and P_h denotes the L^2 -projection on \mathcal{S}_h . We write $S_h(t)$ for the discrete analogue of the operator $S(t)$. This object is formally introduced via $S_h(t) = e^{-tA_h}$. The semidiscrete mild solution to Equation (3.1) is:

$$X_h(t) = S_h(t) P_h X_0 + \int_0^t S_h(t-s) P_h G(X_h(s)) dM(s). \tag{3.2}$$

In what follows we study the magnitude of the error that results from approximating Equation (2.4) by (3.1). For this we require the error estimates for the deterministic case. The deterministic equation is given by

$$Y(t) = S(t) Y_0$$

and the corresponding semidiscrete equation by

$$Y_h(t) = S_h(t) P_h Y_0.$$

The following error estimate is a consequence of Theorem 3.1 in [25].

Lemma 3.1. For $\|Y_0\|_{\dot{H}^2}^2 < \infty$

$$\|Y_h(t) - Y(t)\|_H \leq Ch^2 \|Y_0\|_{\dot{H}^2} \quad \forall t \geq 0. \quad (3.3)$$

For simplicity in what follows C absorbs any constants stemming from our computations. The theorem below provides a bound for the mean square error in the semidiscrete case.

Theorem 3.2. If $\mathbb{E}\|X_0\|_{\dot{H}^1}^2 < \infty$, then for $t \in [0, T]$ there exists $C(t) < \infty$ such that

$$\mathbb{E}\|X_h(t) - X(t)\|_H^2 \leq C(t)h^2(\mathbb{E}\|X_0\|_{\dot{H}^1}^2 + \sup_{s \in [0, t]} \mathbb{E}\|X(s)\|_H^2). \quad (3.4)$$

Proof. The mild and the semidiscrete mild solutions to (2.4) are respectively

$$X(t) = S(t)X_0 + \int_0^t S(t-s)G(X(s))dM(s)$$

and

$$X_h(t) = S_h(t)P_h X_0 + \int_0^t S_h(t-s)P_h G(X_h(s))dM(s).$$

The mean square error satisfies

$$\begin{aligned} \mathbb{E}\|X_h(t) - X(t)\|_H^2 &\leq C(\mathbb{E}\|(S_h(t)P_h - S(t))X_0\|_H^2 \\ &\quad + \mathbb{E}\|\int_0^t (S_h(t-s)P_h - S(t-s))G(X(s))dM(s)\|_H^2 \\ &\quad + \mathbb{E}\|\int_0^t S_h(t-s)P_h(G(X_h(s)) - G(X(s)))dM(s)\|_H^2). \end{aligned}$$

We analyze the terms on the right-hand side one by one. For the first one, it follows from Lemma 3.1 that

$$\mathbb{E}\|(S_h(t)P_h - S(t))X_0\|_H^2 \leq Ch^2 \mathbb{E}\|X_0\|_{\dot{H}^1}^2.$$

To approximate the second term we need (2.1), which yields

$$\begin{aligned} \mathbb{E}\|\int_0^t (S_h(t-s)P_h - S(t-s))G(X(s))dM(s)\|_H^2 &\leq \mathbb{E}\int_0^t \|(S_h(t-s)P_h - S(t-s))G(X(s))\|_{L_{HS}(\mathcal{H}, H)}^2 ds \\ &\leq \int_0^t \|(S_h(t-s)P_h - S(t-s))\|^2 ds \sup_{s \in [0, t]} \mathbb{E}\|G(X(s))\|_{L_{HS}(\mathcal{H}, H)}^2 \\ &\leq Ch^2 \sup_{s \in [0, t]} \mathbb{E}\|X(s)\|_H^2. \end{aligned}$$

We have used the fact that for $v \in H$

$$\begin{aligned} \int_0^t \|(S_h(t-s)P_h - S(t-s))v\|_H^2 ds &= \int_0^t \sup_{v \neq 0} \frac{\|(S_h(t-s)P_h - S(t-s))v\|_H^2}{\|v\|_H^2} ds \\ &\leq \sup_{v \neq 0} \frac{\int_0^t \|(S_h(t-s)P_h - S(t-s))v\|_H^2 ds}{\|v\|_H^2} \\ &\leq Ch^2. \end{aligned}$$

This is a consequence of Lemma 3.7 from [25].

Proposition 2.1 and the Lipschitz-type condition for G , imply for the third term

$$\begin{aligned} \mathbb{E} \left\| \int_0^t S_h(t-s)P_h(G(X_h(s)) - G(X(s)))dM(s) \right\|_H^2 \\ \leq \mathbb{E} \int_0^t \|S_h(t-s)P_h(G(X_h(s)) - G(X(s)))\|_{LHS(\mathcal{H},H)}^2 ds \\ \leq C \int_0^t \|S_h(t-s)P_h\|_H^2 \mathbb{E}\|(X_h(s) - X(s))\|_H^2 ds \\ \leq C \int_0^t \mathbb{E}\|(X_h(s) - X(s))\|_H^2 ds. \end{aligned}$$

Using Grönwall's inequality we get

$$\begin{aligned} \mathbb{E}\|X_h(t) - X(t)\|_H^2 &\leq C(Ch^2 \mathbb{E}\|X_0\|_{H^1}^2 + Ch^2 \sup_{s \in [0,t]} \mathbb{E}\|X(s)\|_H^2 \\ &\quad + C \int_0^t \mathbb{E}\|(X_h(s) - X(s))\|_H^2 ds) \\ &\leq C(t) h^2 (\mathbb{E}\|X_0\|_{H^1}^2 + \sup_{s \in [0,t]} \mathbb{E}\|X(s)\|_H^2). \end{aligned}$$

□

From the proof can be seen that a higher regularity of the initial condition, $\mathbb{E}\|X_0\|_{H^2}^2 < \infty$, leads to optimal convergence of order $O(h^2)$.

3.2. Error estimates in the fully discrete case. We use the linearized backward Euler–Galerkin Method described in [25]. If one were to implement a non-linearized backward Euler scheme the result would not be satisfactory. This is due to the fact that both backward and forward schemes are incompatible with the Itô-Integral. As to have a discretization of $[0, T]$ we introduce the time steps $t_n = nk$, with step size k . We also introduce

$$\frac{dX_h(t)}{dt} \approx \frac{X^n - X^{n-1}}{k},$$

as the backward difference for the time derivative. Formally the discrete SPDE is

$$X^n = r(kA_h)X^{n-1} + \int_{t_{n-1}}^{t_n} r(kA_h)P_h G(X^{n-1})dM(s) \tag{3.5}$$

with initial condition $X^0 = P_h X_0$. The rational function r , which is defined on the spectrum of kA_h , is used to approximate the semigroup $S_h(t)$. In the case of

a backward Euler scheme r is given by $r(\lambda) = (1 + \lambda)^{-1}$, and the approximation of X_h is

$$X_h(t_n) \approx X^n.$$

For the Crank–Nicolson scheme $r(\lambda) = (1 - \frac{\lambda}{2})(1 + \frac{\lambda}{2})^{-1}$ for the first term; however, in the stochastic integral $r(\lambda) = (1 + \frac{\lambda}{2})^{-1}$. When using this scheme the approximation of X_h is given by

$$X_h(t_n) \approx \frac{X^n + X^{n-1}}{2}.$$

For both schemes X_h has to be approximated via X^{n-1} in the stochastic integral, since in an Itô-integral the integrand can only depend on the lower bound of the domain of integration.

As before we borrow a result for the error approximation in the fully discrete case for deterministic equations (Theorem 7.8 in [25]):

The deterministic equation is given as before by

$$Y(t) = S(t)Y_0,$$

the semidiscrete equation by

$$Y_h(t) = S_h(t)P_h Y_0.$$

and the corresponding discrete equation by

$$Y^n = r(kA_h)Y^{n-1}$$

Lemma 3.3. For $\|Y_0\|_{\dot{H}^2}^2 < \infty$

$$\|Y^n - Y(t_n)\|_H \leq C(k + h^2)\|Y_0\|_{\dot{H}^2} \quad \forall t \geq 0. \quad (3.6)$$

With Lemma 2.5 in hand we can prove the main result in this section, namely the mean square error estimate for the fully discrete case:

Theorem 3.4. If $\mathbb{E}\|X_0\|_{\dot{H}^2}^2 < \infty$, then for $t_n \in [0, T]$ there exists $C(t_n) < \infty$ such that

$$\mathbb{E}\|X^n - X(t_n)\|_H^2 \leq C(t_n)\left(k + \frac{h^4}{k}\right)\left(\mathbb{E}\|X_0\|_{\dot{H}^2}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{\dot{H}^1}^2\right). \quad (3.7)$$

Proof. For all $t_n \in [0, T]$ the mild solution to Equation (2.2) is given by

$$X(t_n) = S(t_n)X_0 + \int_0^{t_n} S(t_n - s)G(X(s))dM(s).$$

We iterate (3.5) and obtain

$$\begin{aligned} X^n &= r(kA_h)X^{n-1} + \int_{t_{n-1}}^{t_n} r(kA_h)P_h G(X^{n-1})dM(s) \\ &= r(kA_h)^n P_h X_0 + \sum_{j=1}^n \int_{t_{j-1}}^{t_j} r(kA_h)^{n-j+1} P_h G(X^{j-1})dM(s). \end{aligned}$$

The error satisfies

$$\begin{aligned}
\mathbb{E}\|X^n - X(t_n)\|_H^2 &\leq C(\mathbb{E}\|(r(kA_h)^n P_h - S(t_n))X_0\|_H^2 \\
&+ \mathbb{E}\|\sum_{j=1}^n \int_{t_{j-1}}^{t_j} r(kA_h)^{n-j+1} P_h(G(X^{j-1}) - G(X(t_{j-1})))dM(s)\|_H^2 \\
&+ \mathbb{E}\|\sum_{j=1}^n \int_{t_{j-1}}^{t_j} r(kA_h)^{n-j+1} P_h(G(X(t_{j-1})) - G(X(s)))dM(s)\|_H^2 \\
&+ \mathbb{E}\|\sum_{j=1}^n \int_{t_{j-1}}^{t_j} (r(kA_h)^{n-j+1} P_h - S(t_n - t_{j-1}))G(X(s))dM(s)\|_H^2 \\
&+ \mathbb{E}\|\sum_{j=1}^n \int_{t_{j-1}}^{t_j} (S(t_n - t_{j-1}) - S(t_n - s))G(X(s))dM(s)\|_H^2).
\end{aligned}$$

Lemma 3.3 implies for the first term

$$\begin{aligned}
\mathbb{E}\|(r(kA_h)^n P_h - S(t_n))X_0\|_H^2 &\leq C(k + h^2)^2 \mathbb{E}\|X_0\|_{H^2}^2 \\
&\leq C(k^2 + h^4) \mathbb{E}\|X_0\|_{H^2}^2.
\end{aligned}$$

For the other four terms we need Proposition 2.1. Proposition 1.12 in [10] guarantees that we can interchange the sum with the norm, since all the cross terms are zero. It is here that it is crucial that the integrands depend exclusively on the lower end of the domain of integration. Dependence on the upper limit of integration would lead to non-convergence, since n would appear as a factor. To get an estimate for the second term we use the Lipschitz-type condition for G :

$$\begin{aligned}
\mathbb{E}\|\sum_{j=1}^n \int_{t_{j-1}}^{t_j} r(kA_h)^{n-j+1} P_h(G(X^{j-1}) - G(X(t_{j-1})))dM(s)\|_H^2 \\
&\leq \mathbb{E}\sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|r(kA_h)^{n-j+1} P_h(G(X^{j-1}) - G(X(t_{j-1})))\|_{L_{HS}(\mathcal{H},H)}^2 ds \\
&\leq k \mathbb{E}\sum_{j=1}^n \|r(kA_h)^{n-j+1} P_h(G(X^{j-1}) - G(X(t_{j-1})))\|_{L_{HS}(\mathcal{H},H)}^2 \\
&\leq k \sum_{j=1}^n \|r(kA_h)^{n-j+1} P_h\|^2 \mathbb{E}\|X^{j-1} - X(t_{j-1})\|_H^2 \\
&\leq Ck \sum_{j=1}^n \mathbb{E}\|X^{j-1} - X(t_{j-1})\|_H^2.
\end{aligned}$$

Here, as well as for the next term we need the stability of the approximated solution operator $E_k^n = r(kA_h)^n P_h$, i.e. $\|E_k^n\| \leq C$ for $n \geq 1$. Hence, we calculate using Lemma 2.5, on the regularity of the mild solution, and the Lipschitz-type

condition for G :

$$\begin{aligned}
& \mathbb{E} \left\| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} r(kA_h)^{n-j+1} P_h (G(X(t_{j-1})) - G(X(s))) dM(s) \right\|_H^2 \\
& \leq C \mathbb{E} \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|r(kA_h)^{n-j+1} P_h (G(X(t_{j-1})) - G(X(s)))\|_{L_{HS}(\mathcal{H}, H)}^2 ds \\
& \leq C \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|r(kA_h)^{n-j+1} P_h\|^2 \mathbb{E} \|X(t_{j-1}) - X(s)\|_H^2 ds \\
& \leq C \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (s - t_{j-1}) ds (\mathbb{E} \|X_0\|_{\dot{H}^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_{\dot{H}^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_H^2) \\
& \leq C(t_n) k (\mathbb{E} \|X_0\|_{\dot{H}^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_{\dot{H}^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_H^2).
\end{aligned}$$

Next we use Theorem 7.7 from [25] to obtain, in analogous fashion to the semidiscrete case,

$$\begin{aligned}
& k \sum_{j=1}^n \|r(kA_h)^{n-j+1} P_h - S(t_n - t_{j-1})\|^2 \\
& = k \sum_{j=1}^n \sup_{v \neq 0} \frac{\|(r(kA_h)^{n-j+1} P_h - S(t_n - t_{j-1}))v\|_H^2}{\|v\|_H^2} \\
& \leq k \sum_{j=1}^n \sup_{v \neq 0} \frac{C(h^2 t_{n-j+1}^{-1} + k t_{n-j+1}^{-1})^2 \|v\|_H^2}{\|v\|_H^2} \\
& \leq C(h^2 + k)^2 k \sum_{j=1}^n \frac{1}{((n-j+1)k)^2} \\
& \leq C(h^2 + k)^2 \frac{1}{k} \sum_{j=1}^n \frac{1}{j^2} \\
& \leq C(k + \frac{h^4}{k}).
\end{aligned}$$

It follows from the above that

$$\begin{aligned}
& \mathbb{E} \left\| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (r(kA_h)^{n-j+1} P_h - S(t_n - t_{j-1})) G(X(s)) dM(s) \right\|_H^2 \\
& \leq \mathbb{E} \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|(r(kA_h)^{n-j+1} P_h - S(t_n - t_{j-1})) G(X(s))\|_{L_{HS}(\mathcal{H}, H)}^2 ds \\
& \leq C k \sum_{j=1}^n \|(r(kA_h)^{n-j+1} P_h - S(t_n - t_{j-1}))\|^2 \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_H^2 \\
& \leq C(k + \frac{h^4}{k}) \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_H^2.
\end{aligned}$$

Finally, for the last term we use Theorem 2.4 to get

$$\begin{aligned}
& \mathbb{E} \left\| \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (S(t_n - t_{j-1}) - S(t_n - s)) G(X(s)) dM(s) \right\|_H^2 \\
& \leq \mathbb{E} \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \|S(t_n - s)(S(s - t_{j-1}) - I)G(X(s))\|_{L_{HS}(\mathcal{H}, H)}^2 ds \\
& \leq \mathbb{E} \sum_{j=1}^n \int_{t_{j-1}}^{t_j} C(s - t_{j-1})^{2\alpha} \|A^\alpha G(X(s))\|_{L_{HS}(\mathcal{H}, H)}^2 ds \\
& \leq C \sum_{j=1}^n \int_{t_{j-1}}^{t_j} (s - t_{j-1})^{2\alpha} ds \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_{\dot{H}^{2\alpha}}^2 \\
& \leq C \sum_{j=1}^n (t_j - t_{j-1})^{2\alpha} (t_j - t_{j-1}) \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_{\dot{H}^{2\alpha}}^2 \\
& \leq C n k k^{2\alpha} \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_{\dot{H}^{2\alpha}}^2.
\end{aligned}$$

We use the discrete form of Grönwall's inequality and set $\alpha = \frac{1}{2}$ to obtain

$$\begin{aligned}
\mathbb{E} \|X^n - X(t_n)\|_H^2 & \leq C(t_n)(k^2 + h^4) \mathbb{E} \|X_0\|_{\dot{H}^2}^2 + C(t_n) \left(k + \frac{h^4}{k}\right) \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_H^2 \\
& \quad + C(t_n) k (E \|X_0\|_{\dot{H}^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_{\dot{H}^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_H^2) \\
& \quad + C \sum_{j=1}^n \int_{t_{j-1}}^{t_j} \mathbb{E} \|X^j - X(t_j)\|_H^2 ds \\
& \leq C(t_n) \left(k + \frac{h^4}{k}\right) (\mathbb{E} \|X_0\|_{\dot{H}^2}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E} \|X(s)\|_{\dot{H}^1}^2).
\end{aligned}$$

□

4. Finite Element Method for Hyperbolic SPDE's

In this section we look into the approximation of the transport equation with multiplicative noise, which is of the form

$$dX(t) = BX(t)dt + G(X(t))dM(t), \quad X(0) = X_0. \quad (4.1)$$

Here the first order differential operator B generates the C_0 -semigroup of contractions S , $G : H \rightarrow L(U, H)$ is a linear map and M denotes a square integrable martingale as before. As mentioned in the introduction, S is usually a shift semigroup. For a given vector field b , the first order differential operator is defined as

$$B\phi(x) := \sum_{i=1}^d b_i(x) \phi_{x_i}(x).$$

We require for technical reasons that $\sup_{x \in \mathcal{D}} |\operatorname{div} b(x)| < \infty$. Let the domain \mathcal{D} possess a piecewise smooth boundary $\partial\mathcal{D}$ of class C^1 . The weak solution to

Equation (4.1) is a process X that satisfies

$$(\phi, X(t)) = (\phi, X(0)) + \int_0^t (B^* \phi, X(s)) ds + \int_0^t (G^*(X(s)) \phi, dM(s)), \quad (4.2)$$

for any $\phi \in D(B^*)$ and for all $t \in [0, T]$. We assume that the weak solution satisfies $\sup_{t \in [0, T]} \mathbb{E} \|X(t)\|_{H^2}^2 < \infty$. The inflow boundary is the set

$$\partial \mathcal{D}^- := \{x \in \partial \mathcal{D} : b(x) \cdot \mathbf{n}(x) > 0\},$$

where $\mathbf{n}(x)$ denotes the exterior normal to $\partial \mathcal{D}$ at x . For convenience we impose the Dirichlet boundary condition

$$X = 0 \text{ on } \partial \mathcal{D}^-.$$

This particular structure has to be taken into consideration when defining the finite dimensional spaces for the approximation: $\{\mathcal{S}_h^- \subset H^1(\mathcal{D})\}$ is a family of finite element spaces consisting of piecewise continuous polynomials with respect to the family of triangulations $\{\mathcal{T}_h\}$ of \mathcal{D} , which vanish on the inflow boundary.

In contrast to the parabolic case, in this section we derive error estimates by use of the weak solution. For all $\phi \in H^1(\mathcal{D})$ and $\psi \in H$ we define the bilinear form

$$a_1(\phi, \psi) := (B\phi, \psi).$$

Integrating by parts we obtain

$$a_1(\phi, \phi) = \frac{1}{2} \left(- \int_{\mathcal{D}} (\operatorname{div} b) \phi^2 dx + \int_{\partial \mathcal{D} \setminus \partial \mathcal{D}^-} b \cdot \mathbf{n} \phi^2 dx \right) \quad \forall \phi \in \mathcal{S}_h^-.$$

The integral over $\partial \mathcal{D} \setminus \partial \mathcal{D}^-$ is negative given the conditions on the inflow boundary. We let $\bar{\mathcal{D}} = \{x \in \mathcal{D} : \operatorname{div} b(x) < 0\}$ and we define $\mu := \sup_{\bar{\mathcal{D}}} |\operatorname{div} b|$.

4.1. Error estimates in the semidiscrete case. The semidiscrete problem is to find $X_h(t) \in \mathcal{S}_h^-$ such that for all $\phi \in \mathcal{S}_h^-$ and $t \in [0, T]$

$$d(X_h(t), \phi) = a_1(X_h(t), \phi) dt + (G^*(X_h(t)) \phi, dM(t))_{\mathcal{H}}. \quad (4.3)$$

To investigate the stability of this approximation we set $\phi = X_h$, which yields by use of the Itô formula for square integrable Hilbert space valued martingales

$$\begin{aligned} \mathbb{E}(X_h(t), X_h(t))_H &= \mathbb{E}(X_h(0), X_h(0))_H + \mathbb{E} \int_0^t a_1(X_h(s-), X_h(s)) ds \\ &\quad + \mathbb{E} \int_0^t (G^*(X_h(s-)) X_h(s), dM(s))_{\mathcal{H}} + \mathbb{E} \int_0^t d[M, M]_s. \end{aligned}$$

Here the last term is, by the assumptions on the covariance process of M , bounded. It follows from the fact that X_h is càdlàg that it has at most countably many jumps. The latter, together with our assumption on the divergence of b , implies that

$$\mathbb{E} \int_0^t a_1(X_h(s-), X_h(s)) ds \leq \frac{\mu}{2} \mathbb{E} \int_0^t \|X_h(s)\|_H^2 ds.$$

Using Proposition 2.1, Jensen's inequality and the fact that

$$\|G^*(X_h(s)) X_h(s-)\|_{L_{HS}(\mathcal{H}, \mathbb{R})} \leq \|G(X_h(s))\|_{L_{HS}(\mathcal{H}, H)} \|X_h(s-)\|_H$$

we may write

$$\mathbb{E} \int_0^t (G^*(X_h(s))X_h(s), dM(s))_{\mathcal{H}} \leq \hat{C} \mathbb{E} \int_0^t \|X_h(s)\|_H^2 ds.$$

The constant \hat{C} is the modulus of continuity of G . With the two previous inequalities in hand we get

$$\mathbb{E}\|X_h(t)\|_H^2 \leq \mathbb{E}\|X_h(0)\|_H^2 + (\hat{C} + \frac{\mu}{2}) \mathbb{E} \int_0^t \|X_h(s)\|_H^2 ds$$

so we obtain with Grönwall's inequality

$$\mathbb{E}\|X_h(t)\|_H^2 \leq C\mathbb{E}\|X_h(0)\|_H^2. \tag{4.4}$$

We now show an error estimate for the semidiscrete case. In order to do so we need an interpolation operator, which is defined on the space of continuous functions and takes values in the finite element space \mathcal{S}_h^- . For $v \in C^0(\mathcal{D})$ we define

$$I_h(v) := \sum_{i=1}^n v(x_i)\phi_i,$$

where n denotes the dimension of the finite dimensional space and $\{\phi_i\}_{i=1}^n$ is the corresponding shape function basis. Notice that $I_h(v)(x_i) = v(x_i)$ at all the nodes $x_i, i = 1, \dots, n$.

Theorem 4.1. *Let X and X_h be the solutions of Equation 4.2 and Equation 4.3. Then, if $\mathbb{E}\|X_0 - I_h(X_0)\| \leq Ch \mathbb{E}\|X_0\|_{H^1}$, we have*

$$\mathbb{E}\|X_h(t) - X(t)\|^2 \leq Ch^2(\mathbb{E}\|X(0)\|_{H^1}^2 + \sup_{s \in [0,t]} \mathbb{E}\|X(s)\|_{H^2}^2)$$

Proof. We follow closely the proof of the deterministic case as presented in [15] or [23]. Since the error estimate in the deterministic case does not depend on the source term, the same will be the case in the stochastic setting. We write

$$X_h - X = (X_h - I_h X) + (I_h X - X) =: \eta + \xi$$

We only concern ourselves with η , since by assumption

$$\mathbb{E}\|\xi(t)\|_H^2 \leq Ch^2 \mathbb{E}\|X(t)\|_{H^1}^2.$$

By the definitions of the weak solutions for X and X_h , we have that for all $\phi \in \mathcal{S}_h^-$, $\eta \in \mathcal{S}_h^-$ satisfies

$$(\eta(t), \phi) - (\eta(0), \phi) - a_1\left(\int_0^t \eta(s) ds, \phi\right) = -(w, \phi).$$

We have used the shorthand

$$w = \xi(t) - \xi(0) - \int_0^t B\xi(s) ds + \int_0^t (G(X(s)) - P_h G(X_h(s))) dM(s).$$

If $X_h(0) = I_h(X(0))$, then $\eta(0) = 0$ and Equation (4.4) implies

$$\begin{aligned} \mathbb{E}\|\eta(t)\|_H^2 &\leq C (\mathbb{E}\|\eta(0)\|_H^2 + \mathbb{E}\|\xi(t)\|_H^2 + \mathbb{E}\|\xi(0)\|_H^2 + \mathbb{E} \int_0^t \|B\xi(s)\|_H^2 ds \\ &\quad + \mathbb{E} \left\| \int_0^t (G(X(s)) - P_h G(X_h(s))) dM(s) \right\|_H^2. \end{aligned}$$

For the terms involving ξ we write

$$\begin{aligned} \mathbb{E}\|\xi(t)\|_H^2 + \mathbb{E}\|\xi(0)\|_H^2 + \mathbb{E} \int_0^t \|B\xi(s)\|_H^2 ds \\ \leq Ch^2 (\mathbb{E}\|X(0)\|_{H^1}^2 + \mathbb{E}\|X(t)\|_{H^1}^2 + \mathbb{E} \int_0^t \|X(s)\|_{H^2}^2 ds) \\ \leq C(t)h^2 (\mathbb{E}\|X(0)\|_{H^1}^2 + \sup_{s \in [0,t]} \mathbb{E}\|X(s)\|_{H^1}^2 + \sup_{s \in [0,t]} \mathbb{E}\|X(s)\|_{H^2}^2). \end{aligned}$$

Theorem 1.1 in [25] and the Lipschitz-type condition on G as well as Equation (2.1) give:

$$\begin{aligned} \mathbb{E} \left\| \int_0^t (G(X(s)) - P_h G(X_h(s))) dM(s) \right\|_H^2 \\ \leq \mathbb{E} \left\| \int_0^t (G(X(s)) - P_h G(X(s))) dM(s) \right\|_H^2 \\ \quad + \mathbb{E} \left\| \int_0^t (P_h G(X(s)) - P_h G(X_h(s))) dM(s) \right\|_H^2 \\ \leq \mathbb{E} \int_0^t \|(1 - P_h)\|^2 \|X(s)\|_H^2 ds + \int_0^t \mathbb{E}\|X(s) - X_h(s)\|_H^2 ds \\ \leq Ch^2 \sup_{s \in [0,t]} \mathbb{E}\|X(s)\|_{H^1}^2 + \int_0^t \mathbb{E}\|X(s) - X_h(s)\|_H^2 ds. \end{aligned}$$

We conclude by applying Grönwall's inequality, since

$$\begin{aligned} \mathbb{E}\|X_h(t) - X(t)\|^2 &\leq C (\mathbb{E}\|\eta(t)\|^2 + \mathbb{E}\|\xi(t)\|^2) \\ &\leq C(t)h^2 (\mathbb{E}\|X(0)\|_{H^1}^2 + \sup_{s \in [0,t]} \mathbb{E}\|X(t)\|_{H^1}^2 + \sup_{s \in [0,t]} \mathbb{E}\|X(s)\|_{H^2}^2). \end{aligned}$$

□

Theorem 4.1 does not provide a result of optimal convergence order. For smooth initial conditions and smooth solutions the Galerkin approximation might produce reasonable results. However, if we encounter non-smooth solutions oscillations may occur at the boundary (in both the deterministic and stochastic settings). A Petrov-Galerkin approximation, which broadly speaking consists of adding an artificial diffusion, may increase the order of convergence. In particular the *streamline diffusion method* converges with order $O(h^{3/2})$. The effect of the diffusion is to dampen the oscillations along the characteristics near the boundary, even in the presence of non-smooth solutions. The results concerning convergence of the streamline diffusion method presented in [14] [15] can be naturally extended to the stochastic setting by proceeding as we did above.

4.2. Error estimates in the fully discrete case. In order to show convergence in the fully discrete case we employ the linearized backward Euler scheme as introduced in Section 3. The fully discrete problem is to find X^i for $i = 1, \dots, n$ such that

$$(X^n, \phi)_H = (X^0, \phi)_H + k \sum_{i=1}^n a_1(X^{i-1}, \phi)_H + \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G^*(X^{i-1})\phi, dM(s))_{\mathcal{H}}. \quad (4.5)$$

To show stability we choose, as in the semidiscrete case, $\phi = X^*$ to obtain

$$\begin{aligned} \mathbb{E}(X^n, X^n)_H &\leq \mathbb{E}(X^0, X^0)_H + k \sum_{i=1}^n a_1(X^{i-1}, X^{i-1})_H \\ &\quad + \mathbb{E} \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G^*(X^{i-1})X^{i-1}, dM(s))_{\mathcal{H}}. \end{aligned}$$

In analogous fashion to Section 4.1 we get

$$\begin{aligned} \mathbb{E}\|X^n\|_H^2 &\leq \mathbb{E}\|X^0\|_H^2 + \left(\frac{\mu}{2} + \hat{C}\right)k \sum_{i=1}^n \mathbb{E}\|X^{i-1}\|_H^2 \\ &\leq C(t_n) \mathbb{E}\|X^0\|_H^2, \end{aligned}$$

where k stands for the time step. Note that the discrete version of Grönwall's inequality produces a factor n . The latter times k yields the dependence of C on t_n . The mean square error estimate in the fully discrete case reads:

Theorem 4.2. *Let X and X^* be the respective solutions to Equations 4.2 and 4.5. If $\mathbb{E}\|X_0 - I_h(X_0)\| \leq Ch\mathbb{E}\|X_0\|_{H^1}$, then*

$$\mathbb{E}\|X^n - X(t_n)\|_H^2 \leq C(t_n)(h^2 + k) (\mathbb{E}\|X(0)\|_{H^2}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{H^2}^2).$$

Proof. The proof follows the same train of thought as the one for the semidiscrete case. Once again we split $X^n - X(t_n)$ and treat the summands separately.

$$X^n - X(t_n) = (X^n - I_h X(t_n)) + (I_h X(t_n) - X(t_n)) =: \eta^n + \xi^n.$$

The fact that

$$\mathbb{E}\|\xi^n\|_H^2 \leq Ch^2 \mathbb{E}\|X(t_n)\|_{H^1}^2$$

follows by assumption. The definitions of the weak solutions to X^n and $X(t_n)$ allow us to write

$$(\eta^n, \phi) - (\eta^0, \phi) - k \sum_{i=1}^n a_1(\eta^{i-1}, \phi) = -(w, \phi)$$

for $\eta^n \in \mathcal{S}_h^-$ and $\phi \in \mathcal{S}_h^-$, and w is given by

$$\begin{aligned} w &= \xi^n - \xi^0 - k \sum_{i=1}^n B\xi^{i-1} - \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (BX(t_{i-1}) - BX(s))ds \\ &\quad - \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G(X^{i-1}) - G(X(t_{i-1})))dM(s) \\ &\quad - \sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G(X(t_{i-1})) - G(X(s)))dM(s). \end{aligned}$$

As before $\eta(0) = 0$, hence

$$\begin{aligned} \mathbb{E}\|\eta^n\|_H^2 &\leq \mathbb{E}\|\xi^n\|_H^2 + \mathbb{E}\|\xi^0\|_H^2 + \mathbb{E}\|k \sum_{i=1}^n B\xi^{i-1}\|_H^2 \\ &\quad + \mathbb{E}\left\|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (BX(t_{i-1}) - BX(s))ds\right\|_H^2 \\ &\quad + \mathbb{E}\left\|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G(X(t_{i-1})) - G(X(s)))dM(s)\right\|_H^2 \\ &\quad + \mathbb{E}\left\|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G(X^{i-1}) - G(X(t_{i-1})))dM(s)\right\|_H^2. \end{aligned}$$

The first three terms on the right hand side can be approximated as follows:

$$\begin{aligned} \mathbb{E}\|\xi^n\|_H^2 + \mathbb{E}\|\xi^0\|_H^2 + \mathbb{E}\|k \sum_{i=1}^n B\xi^{i-1}\|_H^2 \\ \leq Ch^2 \mathbb{E}\|X(0)\|_{H^1}^2 + Ch^2 \mathbb{E}\|X(t_n)\|_{H^1}^2 + C(t_n)h^2 \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{H^2}^2. \end{aligned}$$

Lemma 2.5 implies for the fourth term that

$$\begin{aligned} \mathbb{E}\left\|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (BX(t_{i-1}) - BX(s))ds\right\|_H^2 &\leq C \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \mathbb{E}\|X(t_{i-1}) - X(s)\|_{H^1}^2 ds \\ &\leq C(t_n)k (\mathbb{E}\|X(0)\|_{H^2}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{H^2}^2). \end{aligned}$$

We use the same Lemma, plus the Lipschitz condition for G and Equation (2.1) in the approximation of the fifth term:

$$\begin{aligned} \mathbb{E}\left\|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G(X(t_{i-1})) - G(X(s)))dM(s)\right\|_H^2 \\ \leq C \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \mathbb{E}\|X(t_{i-1}) - X(s)\|_{H^1}^2 ds \\ \leq C(t_n)k (\mathbb{E}\|X(0)\|_{H^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{H^1}^2). \end{aligned}$$

The Lipschitz condition for G and Equation (2.1) provide an approximation of the last term on the right hand side of the estimate for $\mathbb{E}\|\eta^n\|_H^2$, namely

$$\begin{aligned} \mathbb{E}\left\|\sum_{i=1}^n \int_{t_{i-1}}^{t_i} (G(X^{i-1}) - G(X(t_{i-1})))dM(s)\right\|_H^2 \\ \leq C \sum_{i=1}^n \int_{t_{i-1}}^{t_i} \mathbb{E}\|X^{i-1} - X(t_{i-1})\|_H^2 ds \\ \leq Ck \sum_{i=1}^n \mathbb{E}\|X^{i-1} - X(t_{i-1})\|_H^2. \end{aligned}$$

The proof concludes with the use of the discrete version of Grönwall's inequality, since

$$\begin{aligned} \mathbb{E}\|X^n - X(t_n)\|_H^2 &\leq \mathbb{E}\|\eta^n\|_H^2 + \mathbb{E}\|\xi^n\|_H^2 \\ &\leq C(t_n)h^2 (\mathbb{E}\|X(0)\|_{H^1}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{H^2}^2) \\ &\quad + Ck (\mathbb{E}\|X(0)\|_{H^2}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{H^2}^2) \\ &\quad + Ck \sum_{i=1}^n \mathbb{E}\|X^{i-1} - X(t_{i-1})\|_H^2 \\ &\leq C(t_n)(h^2 + k) (\mathbb{E}\|X(0)\|_{H^2}^2 + \sup_{s \in [t_0, t_n]} \mathbb{E}\|X(s)\|_{H^2}^2). \end{aligned}$$

□

Theorem 4.2 does not provide convergence of optimal order, which has the same cause as the sub-optimality of the semidiscrete approximation. In the deterministic case, the error approximations in space are of the same magnitude than those in time. However, the Euler–Maruyama scheme we used for the approximation of the SPDE has a maximum convergence of order $\Delta t = k$. For higher convergence one has to introduce a Milstein scheme, which also takes into account higher order terms of the Itô–Taylor approximation. Another option is to treat time with a Galerkin approximation as well.

In the hyperbolic case, convergence of the fully discrete equation is not coupled in time and space, i.e. convergence in space is independent of convergence in time, whereas in the parabolic case our results indicate that refinements should be made simultaneously in both variables. This is not surprising, since in contrast to a parabolic equation, in a transport equation the variables can be treated in a somehow indistinct manner.

Simulation of the discrete equations bears some different task. The square integrable martingale has to be discretized. One could approximate the noise by truncating the series

$$M(t) = \sum_{i=1}^{\infty} \sqrt{\lambda_i} m_i(t) e_i,$$

where e_i are the eigenvectors of the covariance operator with the corresponding eigenvalues λ_i and m_i a sequence of independent Lévy processes. If the eigenvalues

converge fast enough, the error by this approximation is much smaller than the overall order of convergence. However, this question is open and subject to further research.

Acknowledgment. The author wishes to express her thanks to Fred Espen Benth, Annika Lang, Santiago Moreno and Jürgen Potthoff for all the helpful comments and fruitful discussions. Further the author is very grateful to the referee whose comments led to the improvement of the paper.

References

1. Argyris, J. H. and Kelsey, S.: *Energy Theorems and Structural Analysis*, Butterworths, London-Toronto-Sydney-Wellington-Durban, 1960.
2. Barth, A. and Benth, F. E.: Forward dynamics in energy markets – an infinite dimensional approach, *preprint* (2009).
3. Björk, T., Kabanov, Y., and Runggaldier, W.: Bond market structure in the presence of marked point process, *Math. Finance* **7** (1997) 211–239.
4. Carmona, R. and Tehranchi, M.: *Interest Rate Models: An infinite dimensional stochastic analysis perspective*, Springer Finance, Springer, 2006.
5. Chow, P. L.: *Stochastic Partial Differential Equations*, Chapman and Hall/CRC Applied Mathematics and Nonlinear Science Series. Boca Raton, FL: Chapman and Hall/CRC, 2007.
6. Chow, P. L. and Jiang, J. L.: Almost sure convergence of some approximate solutions for random parabolic equations, in: *Random partial differential equations, Proc. Conf., Oberwolfach, Germany* **102** (1989) 45–54.
7. Chow, P. L., Jiang, J. L., and Menaldi, J. L.: Pathwise convergence of approximate solutions to Zakai’s equation in a bounded domain, in: *Stochastic partial differential equations and applications. Proceedings of the third meeting on stochastic partial differential equations and applications held at Villa Madruzzo, Trento, Italy* **268**, (1990) 111–123, Harlow: Longman Scientific and Technical. Pitman Res. Notes Math.
8. Ciarlet, P. G.: *The Finite Elements Method for Elliptic Problems*, North-Holland, 1977.
9. Crisan, D., Gaines, J. G., and Lyons, T.: Convergence of a branching particle method to the solution of the Zakai equation, *SIAM Journal on Applied Mathematics* **58** (1998) 1568–1590.
10. Da Prato, G. and Zabczyk, J.: *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, 1992.
11. Fujita, H. and Mizutani, A.: On the finite element method for parabolic equations, I ; Approximation of Holomorphic Semi-groups, *J. Math. Soc. Japan* **28** (1976) 93–107.
12. Gaines, J. G.: Numerical experiments with S(P)DE’s, *Stochastic Partial Differential Equations, London Mathematical Society Lecture Note Series* **216** (1995) 55–71.
13. Gyöngy, I. and Nualart, D.: Implicit scheme for quasi-linear parabolic partial differential equations perturbed by space-time white noise, *Stochastic Processes and Their Applications* **58** (1995) 57–72.
14. Hesthaven, J. S. and Warburton, T.: *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*, Springer, New York, 2008.
15. Larsson, S. and Thomee, V.: *Partial Differential Equations with Numerical Methods*, Texts in Applied Mathematics, Springer, 2003.
16. Di Masi, G., Björk, T., Runggaldier, W., and Kabanov, Y.: Towards a general theory of bond markets, *Finance and Stochastics* **2** (1997) 141–174.
17. Melosh, R. J.: Structural analysis of solids, *J. ASCE Struct. Div.* **ST4-89** (1963) 205–223.
18. Musiela, M.: Stochastic PDEs and the term structure model, *Journées Internationales de Finance IGR-AFFL. La Baule* **69** (1993).
19. Nedoma, J.: The finite elements solution to elliptic and parabolic equations using simplicial isoparametric elements, *R.A.I.R.O.* **13** (1979) 257–289.
20. Pazy, A.: *Semigroups of Linear Operators and Applications to Partial Differential Equations*, Applied Mathematical Sciences, 44, Springer, New York, 1983.

21. Peszat, S. and Zabczyk, J.: *Stochastic Partial Differential Equations with Lévy Noise*, Cambridge University Press, 2007.
22. Prévôt, C. and Röckner, M.: *A Concise Course on Stochastic Partial Differential Equations*, Springer, 2007.
23. Quarteroni, A. M. and Valli, A.: *Numerical Approximation of Partial Differential Equations*, Springer, 2008.
24. Riesz, F. and Sz.-Nagy, B.: *Functional Analysis*, Dover Books on Advanced Mathematics, Dover Publications Inc., New York, 1995. Translated from the second French edition by Leo F. Boron, Reprint of the 1955 original.
25. Thomée, V.: *Galerkin Finite Element Methods for Parabolic Problems*, Springer Series in Computational Mathematics 25. Berlin: Springer, 2006.
26. Turner, M. J., Clough, R. W., Martin, H. C., and Topp, L. J.: Stiffness and deflection analysis of complex structures, *Journal of Aeronautical Sciences* **23** (1956) 805-824.
27. Ushijima, T.: On the finite element approximation of parabolic equations – Consistency, Boundedness and Convergence, *Mem. Numer. Math.* **2** (1975) 21–33.
28. Yan, Y.: The finite element method for a linear stochastic parabolic partial differential equation driven by additive noise, *Preprint 2003-07, Chalmers University of Technology* (2003).
29. Yoo, H.: Semi-discretization of stochastic partial differential equations on \mathbb{R} by a finite-difference method, *Mathematics of Computation* **69** (2000) 653–666.
30. Zlámal, M.: Finite elements methods for parabolic equations, *Mathematics of Computation* **28** (1974) 393–409.
31. Zlámal, M.: Finite elements methods for non-linear parabolic equations *R.A.I.R.O.* **11** (1977) 93–107.

ANDREA BARTH: CENTRE OF MATHEMATICS FOR APPLICATIONS (CMA), UNIVERSITY OF OSLO,
N-0316 OSLO, NORWAY

E-mail address: andrea.barth@cma.uio.no