SIMULATION OF STOCHASTIC PARTIAL DIFFERENTIAL EQUATIONS USING FINITE ELEMENT METHODS

ANDREA BARTH AND ANNIKA LANG

Abstract. These notes describe numerical issues that may arise when implementing a simulation method for a stochastic partial differential equation. It is shown that an additional approximation of the noise does not necessarily affect the order of convergence of a discretization method for a stochastic partial differential equation driven by Lévy noise. Furthermore finite element methods are explicitly given and simulations are done. In statistical tests, it is shown that the simulations obey the theoretical orders of convergence.

1. Introduction

The numerical study and simulation of Hilbert space valued stochastic differential equation (SPDEs) is a fairly new topic. Within the last years the extension of PDEs to SPDEs has become more and more important in applications especially in engineering such as image analysis, surface analysis, filtering [17, 19, 25, 27, 32]. On the other hand side, in finance, people extend finite dimensional systems of stochastic differential equations (SDEs) to infinite dimensional ones [5], i.e. to SPDEs. Explicit solutions to most of the problems do not exist. Therefore it is natural to simulate a discrete version of these SPDEs. The theory of approximating the mild solution of SPDEs with colored noise has been done e.g. in [2, 4, 9, 10, 11, 12, 15, 16, 18, 19, 22, 24, 31, 33] and references therein. When these approximations are simulated, some more problems arise especially with the approximation of the noise. These will be faced in this paper. Approximations of stochastic processes have been done e.g. in [1, 7, 29, 30] while the literature on simulations of Hilbert space valued stochastic processes is relatively rare (see e.g. [23, 26]).

In this paper we discuss issues that arise when simulating SPDEs which will be seen in the more general framework of Hilbert space valued SDEs. The equations that we want to simulate are of the form

$$dX(t) = AX(t)\,dt + G(X(t))\,dL(t) \tag{1.1}$$

with initial condition $X(0) = X_0$, where A is a differential operator and L is a Hilbert space valued Lévy process.

In [2, 4, 21, 22] discretization schemes have been introduced and it has been shown that these approximations converge almost surely resp. in mean square to the mild resp. weak solution of the SDE. The goal of this paper is to implement examples of these equations and to give explicit simulation schemes. It turns out that in general cases it is easier to implement the approximation scheme, if we do an additional approximation of the noise

Date: March 1, 2010.

Key words and phrases. Finite Element Method, Stochastic Partial Differential Equation, Lévy Process, Galerkin Method.

The authors wish to express many thanks to Pao-Liu Chow, José-Luis Menaldi, and Jürgen Potthoff for fruitful discussions and helpful comments.

because otherwise either an infinite number of independent Lévy processes or a finite number of Lévy processes with given correlation might have to be simulated. This is motivated in Section 3. We show that, depending on the covariance of the Lévy process, the order of convergence of the approximation scheme that was shown in [2] and [4] is not affected.

The work is organized as follows: Section 2 introduces the framework including the SPDE properties and the known approximation results. In Section 3, the driving noise is approximated by the truncated Karhunen–Loeve expansion. It is shown which properties of the eigenvalues of the covariance of the stochastic process imply that the overall order of convergence of the approximation scheme for an SPDE is not affected. Section 4 is devoted to explicitly give a finite element method (FEM) for the approximation. The simulation results including paths and statistics are displayed in Section 5. Finally, Section 6 concludes the results of the notes.

2. Framework

Let $(H,(\cdot,\cdot)_H)$ and $(U,(\cdot,\cdot)_U)$ denote separable Hilbert spaces and $(\Omega,\mathcal{F},(\mathcal{F}_t),\mathbb{P})$ a filtered probability space satisfying the "usual conditions". We are interested in simulating the solution of an SPDE driven by a U–valued Lévy process L and start with an introduction to the noise.

Definition 2.1. A stochastic process $(L = L(t), t \geq 0)$ with values in U is called a *Lévy process*, if

(a) L has independent increments, i.e. if for any $0 \leq t_0 < t_1 < \cdots < t_n$, the U–valued random variables $L(t_1) - L(t_0)$, $L(t_2) - L(t_1)$, $\ldots$, $L(t_n) - L(t_{n-1})$ are independent,
(b) L has stationary increments, i.e. the law $\mathcal{L}(L(t) - L(s))$ of $L(t) - L(s)$ depends only on the difference $t - s$ for $t > s$,
(c) $L(0) = 0$,
(d) L is stochastically continuous, i.e. $\lim_{s\to t} P(\|L(t) - L(s)\|_U > \epsilon) = 0$ for every $\epsilon > 0$, $s, t \geq 0$.

For the covariance of a centered, square integrable Lévy process, we get that there exists Q in the set of all symmetric, non-negative and nuclear operators, $L_1^+(U)$, such that for all $t, s \geq 0$ and $x, y \in U$

$$\mathbb{E}\big((L(t),x)_U(L(s),y)_U\big) = (t \wedge s)(Qx,y)_U.$$

Since $Q \in L_1^+(U)$, there exists an orthonormal basis $(e_i, i \in \mathbb{N})$ of U consisting of eigenvectors of Q. Therefore we have the representation $Qe_i = \gamma_i e_i$, where $\gamma_i \geq 0$ is the eigenvalue corresponding to e_i. The square root of Q is defined as

$$Q^{1/2}\varphi := \sum_i (\varphi, e_i)_U \, \gamma_i^{1/2} \, e_i, \qquad \varphi \in U$$

and $Q^{-1/2}$ is the pseudo inverse of $Q^{1/2}$. Let us denote by $(\mathcal{H},(\cdot,\cdot)_{\mathcal{H}})$ the Hilbert space defined by $\mathcal{H} = Q^{1/2}(U)$ endowed with the inner product $(x,y)_{\mathcal{H}} = (Q^{-1/2}x, Q^{-1/2}y)_U$ for $x, y \in \mathcal{H}$.

The Itô isometry holds for Lévy process L with covariance Q that are càdlàg square integrable martingales. In the following proposition $L_{HS}(\mathcal{H},H)$ refers to the space of all Hilbert–Schmidt operators from $\mathcal{H}$ to H and $\|\cdot\|_{L_{HS}(\mathcal{H},H)}$ denotes the corresponding norm.

Proposition 2.2 ([28, Cor. 8.17])**.** *Let* $\mathbb{L}^2_{\mathcal{H},T}(H) := L^2(\Omega\times[0,T],\mathcal{P}_{[0,T]},\mathbb{P}\otimes d\lambda;\, L_{HS}(\mathcal{H},H))$ *be the space of integrands, where* $\mathcal{P}_{[0,T]}$ *denotes the* σ*–field of predictable sets in* $\Omega\times[0,T]$

and $d\lambda$ is the Lebesgue measure, then for every $\Psi \in \mathbb{L}^2_{\mathcal{H},T}(H)$

$$\mathbb{E}(\| \int_0^t \Psi(s)\, dL(s)\|_H^2) = \mathbb{E}\big(\int_0^t \|\Psi(s)\|^2_{L_{HS}(\mathcal{H},H)}\, ds\big).$$

From here on we will assume that L always refers to a Lévy process that is a càdlàg square integrable martingale. Stochastic integrals with respect to this type of Lévy processes also satisfy a Burkholder–Davis–Gundy type inequality for $\Psi \in \mathbb{L}^2_{\mathcal{H},T}(H)$ and $p \in (0,2]$ [13]:

$$\mathbb{E}(\sup_{t\in[0,T]} \| \int_0^t \Psi(s)\, dL(s)\|_H^p) \le C\, \mathbb{E}\big((\int_0^T \|\Psi(s)\|^2_{L_{HS}(\mathcal{H},H)}\, ds)^{p/2}\big) \tag{2.1}$$

for a constant C. This inequality stays true for $p \in (2,+\infty)$, if L is a continuous square integrable martingale, i.e. in our approach a Wiener process. Similar estimates also apply to stochastic convolutional integrals with respect to a semigroup (see e.g. [13]).

Following [28], a square integrable Lévy process L has the series representation

$$L(t) = \sum_{i=1}^{\infty} L_i(t) a_i,$$

where the elements L_i are real-valued Lévy processes and $a := (a_i, i \in \mathbb{N})$ is a basis of U. Furthermore, we already know that there exists an eigenbasis $(e_i, i \in \mathbb{N})$ and eigenvalues $(\gamma_i, i \in \mathbb{N})$ of the covariance Q. Then L can be represented by

$$L(t) = \sum_{i=1}^{\infty} \sqrt{\gamma_i}\, L_i(t)\, e_i, \tag{2.2}$$

where the elements L_i are again real-valued Lévy processes that are not necessarily independent. On the other hand side, if we have a sequence of independent, identically distributed, one dimensional real-valued Lévy processes $(L^s_i, i \in \mathbb{N})$ and set

$$L^s(t) := \sum_{i=1}^{\infty} \sqrt{\gamma_i}\, L^s_i(t)\, e_i,$$

for $t \in [0,T]$, then by Theorem 4.40 in [28] the series on the right hand side converges $\mathbb{P}$–a.s., uniformly in t on any compact interval and defines a Lévy process L^s on the Hilbert space U. This series also converges in mean square, if Q is nuclear as $(e_i, i \in \mathbb{N})$ is an orthonormal basis and the elements L_i are i.i.d. square integrable Lévy processes. In this paper we will limit ourselves to Lévy processes that have a series representation consisting of a sequence of independent, identically distributed, real-valued Lévy processes $(L_i, i \in \mathbb{N})$. Examples of such processes are Wiener processes and random fields of correlated NIG processes.

After having introduced the framework of the noise that we are interested in, we introduce the corresponding SPDEs. We are interested in simulating the solution of equation

$$dX(t) = AX(t)\, dt + G(X(t))\, dL(t) \tag{2.3}$$

on the finite time interval $[0,T]$ with initial condition $X(0) = X_0$, where A is generator of a C_0–semigroup $S = (S(t), t \ge 0)$ and L is a Lévy process as introduced before. Let G be a mapping from H to $L(\mathcal{H}, H)$. If G is Lipschitz continuous and of linear growth, i.e. there exists a constant C such that for $u, v \in H$

$$\|G(u)\|_{L_{HS}(\mathcal{H},H)} \le C(1 + \|u\|_H),$$
$$\|G(u) - G(v)\|_{L_{HS}(\mathcal{H},H)} \le C\|u - v\|_H,$$

and if X_0 is $\mathcal{F}_0$–measurable, then by results in Chapter 9 of [28], Equation (2.3) has a unique mild solution, i.e. $\sup_{t\in[0,T]} \mathbb{E}(\|X(t)\|_H^2) < +\infty$ and $X(t)$ can be written as

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

For a full introduction to Hilbert space valued stochastic differential equations we refer the reader to [6, 8, 28].

The authors approximated equations similar to Equation (2.3) in [2, 3, 4, 20] using Galerkin methods. Similarly to these approaches, let $\mathcal{S}_h$ be a family of finite element spaces, consisting of piecewise continuous polynomials with respect to the triangulation $\mathcal{T}_h$ of $D \subset \mathbb{R}^d$ with piecewise smooth boundary. We assume $\mathcal{S}_h \subset H_0^1(D)$. Furthermore let P_h denote the L^2–Projection of $L^2(D)$ on $\mathcal{S}_h$. We set $A_h := P_h A P_h$. The operator $S_h(t)$ refers to the discrete analog of $S(t)$, formally introduced by $S_h(t) = e^{-tA_h}$. The rational approximation of the semigroup is given by $r(\lambda) := (1+\lambda/2)/(1-\lambda/2)$. In [4] it was shown that the Galerkin–Milstein–Crank–Nicolson scheme with equidistant time discretization $0 = t_0 < t_1 < \cdots < t_n = T$

$$\begin{aligned} X^n = {} & r(kA_h)X^{n-1} + \int_{t_{n-1}}^{t_n} \tfrac{1}{2}\,(r(kA_h)+\mathbb{1})P_hBP_hX^{n-1}\,ds \\ & + \int_{t_{n-1}}^{t_n} \Big(\tfrac{1}{2}\,(r(kA_h)+\mathbb{1})P_hBP_h \int_{t_{n-1}}^{s} P_hG(X^{n-1})\,dM(r)\Big)\,ds \\ & + \int_{t_{n-1}}^{t_n} \tfrac{1}{2}\,(r(kA_h)+\mathbb{1})P_hG(X^{n-1})\,dM(s) \\ & + \int_{t_{n-1}}^{t_n} \Big(\tfrac{1}{2}\,(r(kA_h)+\mathbb{1})P_hG\Big(\int_{t_{n-1}}^{s} P_hG(X^{n-1})\,dM(r)\Big)\Big)\,dM(s). \end{aligned} \tag{2.4}$$

leads to almost surely convergence of the approximated solution to the mild solution of

$$dX(t) = (A+B)X(t)\,dt + G(X(t))\,dM(t),$$

where A is a dissipative second order differential operator, B is a first order differential operator, G is linear, and M is a continuous square integrable martingale. The order of convergence is $\mathrm{O}((h^2+k+k^{1/2}h)^{1-\epsilon})$, where h denotes the size of the space discretization and k is the step size in time.

In [2] a stochastic heat equation with Dirichlet boundary conditions on the Hilbert space $L^2(D)$ with $D \subset \mathbb{R}^d$

$$dX(t) = \Delta X(t)\,dt + G(X(t))\,dM(t)$$

with a not necessarily continuous square integrable martingale was approximated by a Galerkin method and a backward Euler scheme given by the recursive scheme

$$X^n = r(kA_h)X^{n-1} + \int_{t_{n-1}}^{t_n} r(kA_h)P_hG(X^{n-1})\,dM(s),$$

where $A_h = P_h\Delta P_h$ and $r(\lambda) = (1+\lambda)^{-1}$. It was shown that the approximate solution converges to the mild solution in mean square of order $\mathrm{O}(\sqrt{k}+h)$ for an appropriate initial condition X_0. The same order of convergence is achieved for the weak solution of a hyperbolic problem of type

$$dX(t) = BX(t)\,dt + G(X(t))\,dM(t)$$

with initial condition $X(0) = X_0$ and B is a first order differential operator on $D \subset \mathbb{R}^d$. The backward Euler approximation scheme that has these properties is given in weak form by

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

where G^* denotes the adjoint of G and $a_1(\phi, \psi) = (B\phi, \psi)_H$ for ϕ, ψ in the finite element space $\mathcal{S}_h$.

In these schemes, the noise is not approximated separately. An extra approximation can simplify the simulations a lot. This extra approximation of the noise can preserve the overall order of convergence, which is shown in the next section.

3. Approximation of the Noise

In this section we introduce an extra approximation of the U–valued driving noise of Equation (2.3) which might lead to easier simulations. This can be seen in the following. Therefore let us denote the Itô integral to be simulated by

$$\int_a^b P_h \Psi(s)\, dL(s)$$

with $a < b$ and $\Psi \in \mathbb{L}^2_{\mathcal{H},T}(H)$. This expression can be rewritten using the series representation of L , Equation (2.2), to

$$\sum_{i=1}^{\infty} \sqrt{\gamma_i} \int_a^b P_h \Psi(s) e_i\, dL_i(s).$$

If $\mathcal{S}_h$ is generated by $(\phi_k, k = 1, \dots m)$, we have that

$$\int_a^b P_h \Psi(s)\, dL(s) = \sum_{i=1}^{\infty} \sum_{k=1}^{m} \sqrt{\gamma_i} \int_a^b (\Psi(s) e_i, \phi_k)_H\, dL_i(s)\, \phi_k.$$

Without further assumptions, nothing is known about $(\Psi(s)e_i, \phi_k)_H$ and therefore although the result is finite dimensional it might be necessary to simulate an infinite number of Lévy processes. One possible way out of this problem is to approximate the Lévy process by a truncation of the series expansion, i.e. set

$$L^\kappa(t) = \sum_{i=1}^{\kappa} \sqrt{\gamma_i}\, L_i(t)\, e_i.$$

The goal of this section is to show which conditions imply that the overall order of convergence is not affected by the additional approximation.

Let $(L^\kappa, \kappa \in \mathbb{N})$ be the sequence of truncated series expansions with covariance Q^κ that converges almost surely to the Lévy process L with covariance Q. We set

$$L^{c\kappa}(t) = L(t) - L^\kappa(t) = \sum_{i=\kappa+1}^{\infty} \sqrt{\gamma_i}\, L_i(t)\, e_i.$$

with covariance $Q^{c\kappa} = Q - Q^\kappa$, which converges almost surely to zero. This implies for the Itô integral with respect to $\Psi \in \mathbb{L}^2_{\mathcal{H},T}(H)$ that

$$\int_a^b \Psi(s)\, dL(s) - \int_a^b \Psi(s)\, dL^\kappa(s) = \int_a^b \Psi(s)\, dL^{c\kappa}(s). \tag{3.1}$$

This difference converges to zero depending on the decay of the eigenvalues $(\gamma_i, i \in \mathbb{N})$, which is shown in the following lemma.

Lemma 3.1. *If* $\mathbb{E}(\sup_{t\in[a,b]} \|\Psi(t)\|^p_{L(U,H)}) < +\infty$ *and there exists a constant* $C > 0$ *such that the eigenvalues satisfy* $\gamma_i \le C\, i^{-\alpha}$ *for* $\alpha > 1$ *and* $\kappa \ge C\, h^{-1}$*, then*

$$\mathbb{E}(\sup_{t\in[a,b]} \| \int_a^t \Psi(s)\, dL(s) - \int_a^t \Psi(s)\, dL^\kappa(s)\|_H^p) \le C_p\, \mathbb{E}(\sup_{t\in[a,b]} \|\Psi(t)\|^p_{L(U,H)})\, h^{\frac{(\alpha-1)p}{2}},$$

for a constant C_p *and* $p \in (0, 2]$*. If, in addition,* L *is continuous, the inequality also holds for* $p > 2$*.*

Proof. Let us choose $p \in (0, 2]$ or $p > 0$ depending on the properties of L, i.e. if it is continuous or not. We first observe that

$$\begin{aligned}\mathbb{E}(\sup_{t\in[a,b]} \| \int_a^t \Psi(s)\, dL(s) - \int_a^t \Psi(s)\, dL^\kappa(s)\|_H^p) &= \mathbb{E}(\sup_{t\in[a,b]} \| \int_a^t \Psi(s)\, dL^{c\kappa}(s)\|_H^p) \\ &\le C_p\, \mathbb{E}((\int_a^b \|\Psi(s)\|^2_{L_{HS}(Q^{c\kappa}U,H)}\, ds)^{p/2})\end{aligned}$$

by Equations (3.1) and (2.1). Next we calculate the Hilbert–Schmidt norm. We have that

$$\mathbb{E}((\int_a^b \|\Psi(s)\|^2_{L_{HS}(Q^{c\kappa}U,H)}\, ds)^{p/2}) = \mathbb{E}((\int_a^b \sum_{i=\kappa+1}^{\infty} \gamma_i \|\Psi(s)e_i\|_H^2\, ds)^{p/2}).$$

With the properties of Ψ it holds

$$\mathbb{E}((\int_a^b \sum_{i=\kappa+1}^{\infty} \gamma_i \|\Psi(s)e_i\|_H^2\, ds)^{p/2}) \le C_p\, \mathbb{E}(\sup_{t\in[a,b]} \|\Psi(t)\|^p_{L(U,H)})(\sum_{i=\kappa+1}^{\infty} \gamma_i)^{p/2},$$

and the decay of the eigenvalues and the assumptions on κ imply

$$\sum_{i=\kappa+1}^{\infty} \gamma_i \le C \sum_{i=\kappa+1}^{\infty} i^{-\alpha} = C \sum_{i=1}^{\infty} (i+\kappa)^{-\alpha} \le C \int_0^{\infty} (x+\kappa)^{-\alpha} dx = C\, (\alpha-1)^{-1} h^{(\alpha-1)}.$$

This proves the lemma. □

We remark that the estimates stay true, if Ψ also depends on the upper integration limit, i.e. the stochastic integral is a convolutional integral with respect to a semigroup, since Equation (2.1) also holds for this type of integrands (see [13]).

In [4] it was shown that the Galerkin–Milstein scheme given by Equation (2.4) converges almost surely of order $\mathrm{O}((h^2 + k + k^{1/2}h)^{1-\epsilon})$ by a Borel–Cantelli argument. In the proof it was especially shown that for an equidistant time discretization $0 = t_0 < t_1 < \cdots < t_n = T$ and a constant $C_p(T)$

$$\mathbb{E}(\sup_{0\le j\le n} \|X(t_j) - X^j\|_H^p) \le C_p(T)\, (h^2 + k + k^{1/2}h)^p.$$

So the extra approximation of the noise preserves the order of convergence, if the eigenvalues $(\gamma_i, i \in \mathbb{N})$ decay with order $i^{-\alpha}$ and $\alpha \ge 5$.

The order of convergence in mean square for the approximation of the heat equation as well as of a hyperbolic problem in [2] is $\mathrm{O}(\sqrt{k} + h)$. This implies that the order is preserved with an extra approximation of the noise, if the eigenvalues $(\gamma_i, i \in \mathbb{N})$ decay with order $i^{-\alpha}$ and $\alpha \ge 3$.

4. A Finite Element Method for SPDEs

A Galerkin approximation for Equation (1.1) was introduced in [2] and [4], where convergence was proved in mean square and almost sure senses. Here we are more concerned with the issues which arise when actually implementing the FEM for those equations.

We work on the same approximation in space as introduced in [4]. For the approximation in time we present a Crank–Nicolson scheme with an Euler–Maruyama approximation. This converges in general slower than the corresponding Milstein scheme but for simplicity reasons it is more convenient. Furthermore, in the case of additive noise as presented in the next section, the Euler–Maruyama and the Milstein scheme have the same order of convergence.

First we give an introduction to the approximation in time using a model problem, afterwards we address specific problems that occur in the implementation of SPDEs.

Let us start with a benchmark model problem: the heat equation on $[0,T]\times[0,\delta] = [0,T]\times D \in \mathbb{R}^2$, $\delta > 0$, with a source term f given by

$$\frac{dX(t)}{dt} = \Delta X(t) + \frac{f(t)}{dt}$$

with Dirichlet boundary condition and a smooth initial condition given by some function X_0. Thus we are facing a parabolic stochastic partial differential equation. The source term f will later transform into the noise term.

The Galerkin approximation for this problem is given by

$$\int_D \dot{X}(t,x)\psi(x)\,dx + \int_D X_x(t,x)\psi_x(x)\,dx = \int_D \dot{f}(t,x)\psi(x)\,dx,$$

where we multiplied the equation above with a *test function* $\psi \in \mathcal{S}_h$, which is an element of the finite dimensional space, where the solution to the approximated problem lies, and then we integrated by parts. $\dot{X}$ denotes the derivative with respect to the first argument and X_x the derivative with respect to the second one. The projection of X onto $\mathcal{S}_h$ denoted by X_h has the unique representation

$$X_h(t,x) = \sum_{j=1}^{n} \xi_j(t)\phi_j(x),$$

where $\phi = (\phi_j, j = 1,\dots,n)$ is a basis of $\mathcal{S}_h$ and the elements $\xi_j(t)$ are the corresponding coefficient functions. The elements ϕ_j are called the *trial functions*. If we combine this representation with the Galerkin approximation above, we obtain the system

$$R\dot{\xi}(t) - Z\xi(t) = \dot{b}(t),$$

where $R = (r_{i,j}, i,j = 1,\dots,n)$, $Z = (z_{i,j}, i,j = 1,\dots,n)$, and $\dot{b} = (\dot{b}_j, j = 1,\dots,n)$ with $r_{i,j} = \int_D \phi_j(x)\phi_i(x)\,dx$, $z_{i,j} = \int_D (\phi_x)_j(x)(\phi_x)_i(x)\,dx$, and $\dot{b}_j(t) = \int_D \dot{f}(t,x)\phi_j(x)\,dx$. One possible choice for ϕ is the hat function basis. The dimension of the finite element space gives directly the number of approximation points, here $x_0,\dots,x_n$, and as well the number of finite elements $([x_{i-1},x_i], i = 1,\dots,n)$. The matrices R and Z can be assembeled from submatrices. For the hat function basis, the submatrices R_k and Z_k, $k = 1,\dots,n$ are given by

$$R_k = \Big(\int_{x_{k-1}}^{x_k} \phi_j(x)\phi_i(x)\,dx,\ i,j = k-1,k\Big),$$
$$Z_k = \Big(\int_{x_{k-1}}^{x_k} (\phi_x)_j(x)(\phi_x)_i(x)\,dx,\ i,j = k-1,k\Big).$$

These are calculated for every $k = 1, \ldots, n$ separately and then the matrices are assembled in the end. If the hat functions are all symmetric, i.e. on an equidistant grid $(x_0, \ldots, x_n)$ with the same scaling, those submatrices are identical for all $k = 2, \ldots, n-1$. Only the elements subject to the boundary conditions have to be updated in the very end. For the model problem, if we take for simplicity equidistant nodes $x_0, \ldots, x_n$ and a hat function basis given by

$$\phi_j(x_i) = \begin{cases} 1, & \text{for } j = i, \\ 0, & \text{otherwise,} \end{cases}$$

then for each x

$$\phi_j(x) = \begin{cases} \frac{x - x_{j-1}}{h}, & \text{for } x_{j-1} \le x \le x_j, \\ \frac{x_{j+1} - x}{h}, & \text{for } x_j \le x \le x_{j+1}, \\ 0, & \text{otherwise.} \end{cases}$$

Every matrix Z_k, for $k = 2, \ldots, n-1$, is then of the form

$$Z_k = \frac{1}{h} \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix},$$

while R_k, for $k = 2, \ldots, n-1$, becomes

$$R_k = \frac{h}{3} \begin{pmatrix} 2 & \frac{1}{2} \\ \frac{1}{2} & 2 \end{pmatrix}.$$

The calculation of R and Z does not depend on the discretization of the time. So the matrices can be computed in advance. Let us introduce the equidistant time discretization $0 = t_0 < t_1 < \ldots < t_N = T$ with step size $k = T/N$. Since R and Z are constant, we get

$$R(\xi(t_i) - \xi(t_{i-1})) - k\, Z\xi(t) = b(t_i) - b(t_{i-1}), \qquad t \in [t_i, t_{i+1}].$$

Here the time derivatives are approximated by

$$\dot{\xi}(t) \approx \frac{\xi(t_i) - \xi(t_{i-1})}{k}, \qquad t \in [t_i, t_{i+1}],$$

and

$$\dot{b}(t) \approx \frac{b(t_i) - b(t_{i-1})}{k}, \qquad t \in [t_i, t_{i+1}].$$

For $\xi(t)$ in the second term of the linear system of equations we apply the Crank–Nicolson approximation, given by

$$\xi(t) \approx \frac{\xi(t_i) + \xi(t_{i-1})}{2}, \qquad t \in [t_i, t_{i+1}].$$

This leads to the following fully discrete equation

$$R(\xi(t_i) - \xi(t_{i-1})) - kZ\big(\tfrac{1}{2}\xi(t_i) + \tfrac{1}{2}\xi(t_{i-1})\big) = b(t_i) - b(t_{i-1}),$$

which is clearly equal to

$$(R - \tfrac{k}{2}Z)\,\xi(t_i) = (R + \tfrac{k}{2}Z)\,\xi(t_{i-1}) + (b(t_i) - b(t_{i-1})).$$

Each component of this equation is known, except for $\xi(t_i)$. Therefore we can calculate the approximation for each time step by solving the system of linear equations. The initial condition is X_0 is approximated as follows:

$$X(0, x) \approx \sum_{j=1}^{n} \xi(0)\phi_j(x).$$

If we set for the abstract source term f the additive noise term $dL(t)$, we get for each entry $j = 1, \dots, n$

$$(b(t_i) - b(t_{i-1}))_j = \frac{1}{2}h(\Delta_i L(x_j) - \Delta_i L(x_{j-1})), \tag{4.1}$$

where $\Delta_i L(x) = L(t_i, x) - L(t_{i-1}, x)$.

The situation is slightly more delicate when we wish to simulate a first order hyperbolic equation. The Galerkin approximation will only lead to stable results, if the solution is smooth. One approach to a stable numerical scheme is to introduce an artificial diffusion along the characteristic lines. These approximation schemes are called Petrov–Galerkin methods.

The main difference between a Petrov–Galerkin method and a Galerkin method is the set of used test functions. In a Galerkin approximation, one uses the same functions as test and trial functions, as we did above. In contrast, to implement our Petrov–Galerkin method we choose as test functions $\phi + \gamma\hat{\phi}$, where ϕ is an element of the hat function basis, $\hat{\phi}$ denotes the (e.g. weak) derivative of the basis function, and $\gamma \in \mathbb{R}$. The trial functions remain ϕ. The parameter γ has to be chosen according to the problem and depends on h. For first order hyperbolic equations we only have essential boundary conditions at the inflow boundary ∂D^-, which leads to the following definition for the finite element space:

$$\mathcal{S}_h^- = \{\chi \in \mathcal{S}_h : \chi = c \text{ at } \partial D^-\},$$

for some real function c.

Consider the initial boundary value problem

$$\frac{dX(t)}{dt} = \nabla X(t) + \frac{f(t)}{dt}$$

or with the previous notations

$$\dot{X} = X_x + \dot{f}$$

in a domain $D \subset \mathbb{R}_+$ with some boundary condition on the inflow boundary and initial condition $u(0) = v$ equal to the boundary condition at the inflow boundary ∂D^-. The semidiscrete problem is now to find a function $X_h(t) \in \mathcal{S}_h^-$ such that

$$(\dot{X}_h, \phi + \gamma\hat{\phi}) = ((X_h)_x, \phi + \gamma\hat{\phi}) + (\dot{f}, \phi + \gamma\hat{\phi}) \qquad \forall \phi \in \mathcal{S}_h^-,\ t > 0,$$

where $(\cdot,\cdot)$ is the abbreviation of $(\cdot,\cdot)_H$. In terms of the hat function basis $(\phi_j,\ j = 1, \dots, n)$, we write

$$(R + \hat{R})\,\dot{u} + (Z + \hat{Z})\,u = \dot{b},$$

where $R = (r_{i,j},\ i,j = 1, \dots, n)$ has elements $r_{i,j} = (\phi_i, \phi_j)$, $\hat{R} = (\hat{r}_{i,j},\ i,j = 1, \dots, n)$ has elements $\hat{r}_{i,j} = (\phi_i, \hat{\phi}_j)$, $Z = (z_{i,j},\ i,j = 1, \dots, n)$ consists of $z_{i,j} = (\phi_i, (\phi_x)_j)$ and $\hat{Z} = (\hat{z}_{i,j},\ i,j = 1, \dots, n)$ consists of $\hat{z}_{i,j} = (\phi_i, (\hat{\phi}_x)_j)$. The vector b denotes the source term evaluated according to the test function $\dot{b}_j(t) = \int_D \dot{f}(t,x)\phi_j(x)\,dx + \gamma \int_D \dot{f}(t,x)\hat{\phi}_j(x)\,dx$. One has to take into account the essential boundary conditions on the inflow boundary here as well.

For the fully discrete problem we will use a Crank–Nicolson approach. This leads to the equation

$$\big(R + \hat{R} - \tfrac{k}{2}(Z + \hat{Z})\big)\xi(t_i) = \big(R + \hat{R} + \tfrac{k}{2}(Z + \hat{Z})\big)\xi(t_{i-1}) + (b(t_i) - b(t_{i-1})).$$

The system of linear equations can be solved recursively, where $(b(t_i) - b(t_{i-1}))$ is defined for the random field f as before in the parabolic case.

5. Simulations

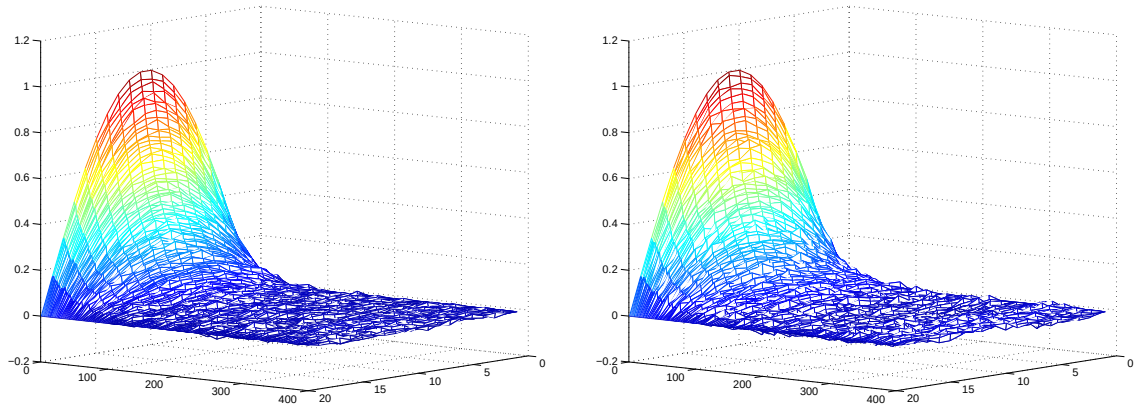
In many models for finance we face SPDEs with additive noise, see [3]. We consider the following setup for the heat equation with additive noise:

$$dX(t) = \Delta X(t)\,dt + dM(t), \qquad t \in [0,1],$$

subject to the initial condition $X(0,x) = \sin(\pi x)$ for $x \in [0,1]$, and Dirichlet boundary conditions $X(t,0) = X(t,1) = 0$. The stochastic process M is either of Wiener type W or a Lévy type martingale L. The solution to the corresponding deterministic system

$$du(t) = \Delta u(t)\,dt$$

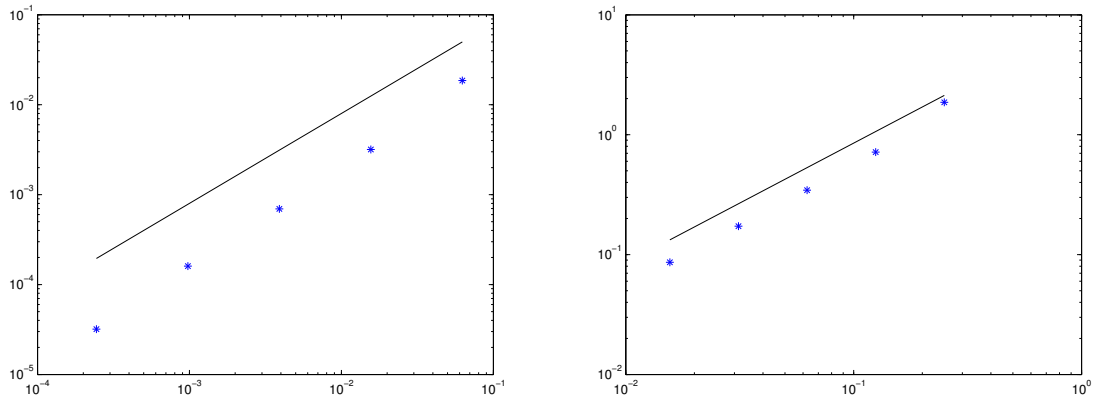
is in this case $u(t,x) = \exp(-\pi^2 t)\sin(\pi x)$. As trial and test functions we take the hat functions introduced in Section 4. As a kernel function for the covariance operator Q we set $q(x,y) = \exp(-\kappa|x-y|^2)$ with range parameter $\kappa > 0$ (in the example we take $\kappa = 2$). In Figure 1, one path is displayed on a grid of 400×20, i.e. we have 400 discretization points in time and 20 in the space domain. In this example the kernel function of the covariance operator is $q(x,y) = \exp(-\kappa|x-y|)$. As driving noise process we have a Gaussian random field on the left



(a) Gaussian random field (b) NIG random field

Figure 1. Simulation of one path of a heat equation with additive noise on a grid with 400×20 points.

hand side and a random field of correlated NIG processes on the right. The random fields have the same dimension as the finite element space. These approximations converge almost surely to the mild solution by Theorem 5.1 in [4] adapted to an Euler–Maruyama scheme and Lemma 3.1. The smoothing effect of the Laplace operator is evident. For small t the effect of the noise is hardly visible, whereas towards the end of the time interval, where the change of the slope of the solution is very small, the noise becomes observable. Convergence results for the Wiener case are illustrated in Figure 2(a) on a logarithmic and in Figure 3(a) on a linear scale. Clearly, there is no exact solution to the stochastic problem available. We take a finite element approximation on a very fine grid as a substitute to an exact solution. In the parabolic case this means 2^7 points in space (leading to 2^{14} points in time). On the x-axis we



(a) Parabolic equation (b) Hyperbolic equation

FIGURE 2. Simulated errors with respect to the time discretization and a line as reference slope on a logarithmic scale.

have the grid size in time and on the y-axis the error which is the approximation of

$$e = \mathbb{E}(\sup_{j=0,\dots,n} \|X(t_j) - X^j\|_H^2)^{1/2}$$

and which was calculated by taking the square root of the average over the realizations of the maximum over all time steps of the normalized sum of the squared difference of the solution on the finest grid and the solution on the coarser grids, i.e. let $(Y_i, i = 1,\dots,N)$ be the set of simulated paths on a time grid $(t_j, j = 0,\dots,n)$ and a space grid $(x_k, k = 0,\dots,l)$, and $(\hat{Y}_i, i = 1,\dots,m)$ the set of the simulated paths of the "exact" solution on the coarser grid in time and space, then the error is given by

$$e_N = \sqrt{\frac{1}{N}\sum_{i=1}^{N} \max_{j=0,\dots,n} \frac{1}{m}\sum_{k=1}^{m} (\hat{Y}_i(t_j, x_k) - Y_i(t_j, x_k))^2}.$$

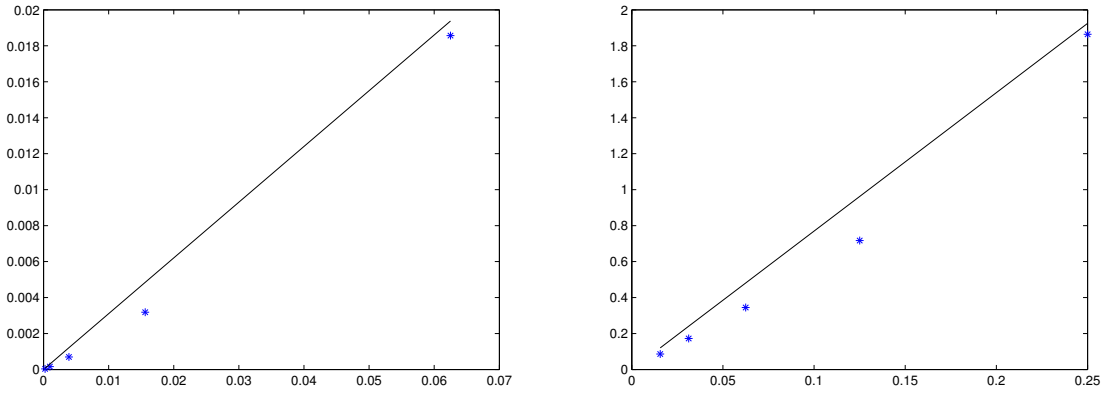
Note that the noise for Y_i and $\hat{Y}_i$, for $i = 1,\dots,m$, consists of the same random numbers. For the coarser grids the increments are added according to the grid size. The number of realizations was $N = 100$ and therefore quite small but the convergence is already visible. The discretization was done starting on a grid of size $2^4 \times 2^2$ up to a grid of $2^{12} \times 2^6$ points. The reference slope is $c\,x$ for a constant $c > 0$ because the Euler–Maruyama scheme converges of order $\mathrm{O}(k)$ and $\mathrm{O}(h^2)$ respectively in the case of additive noise.

The second example that we discuss is a first order hyperbolic stochastic partial differential equation. Here we choose as a model problem for M Wiener noise W or Lévy martingale noise L respectively on the interval $[0,1]$

$$dX(t) = \nabla X(t)\,dt + dM(t), \qquad t \in [0,1],$$

with initial condition $X(0,x) = \sin(2\pi x)$ and inflow boundary condition $X(t,0) = -\sin(2\pi t)$. The solution to the corresponding deterministic system

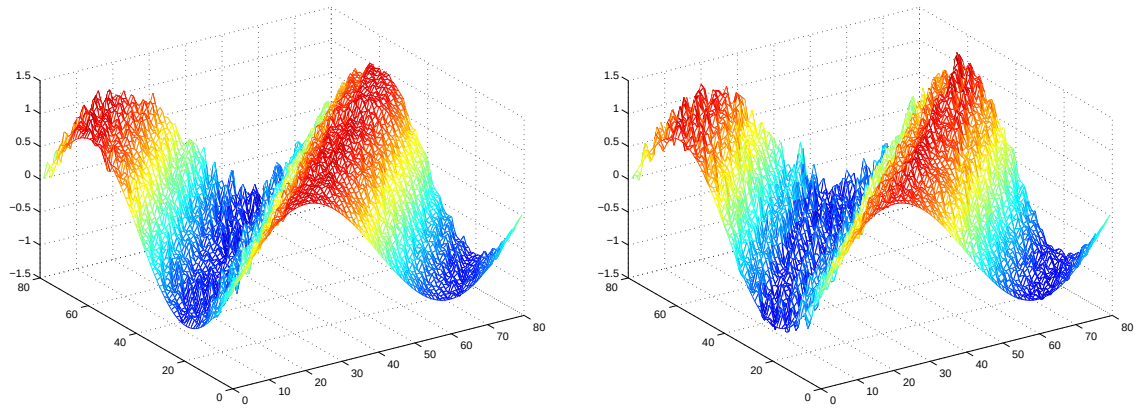
$$du(t) = \nabla u(t)\,dt$$



(a) Parabolic equation (b) Hyperbolic equation

FIGURE 3. Simulated errors with respect to the time discretization and a line as reference slope on a linear scale.

is $u(t,x) = \sin(2\pi(x-t))$, obtained by dropping the noise term. In Figure 4 we have one realization on a grid of 80×80 points with a Gaussian random field on the left hand side and correspondingly with a random field of correlated NIG processes on the right. One can clearly



(a) Gaussian random field (b) NIG random field

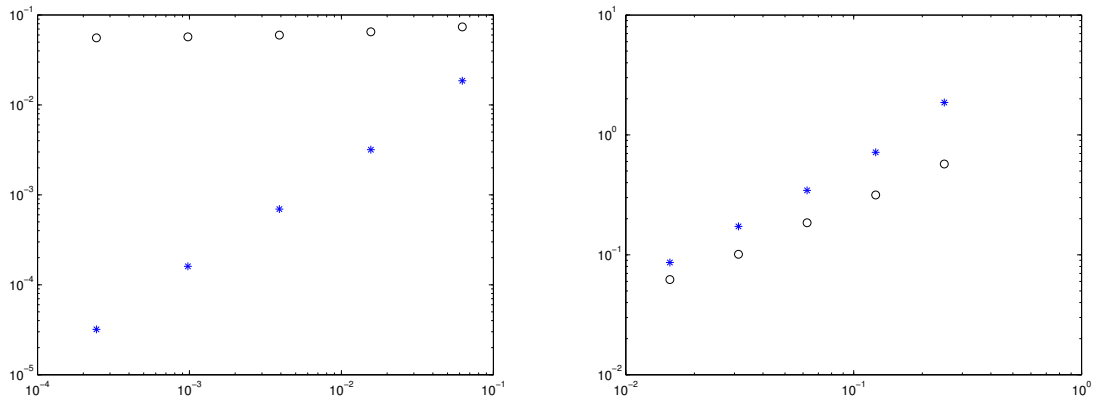
FIGURE 4. Simulation of one path of a hyperbolic equation with additive noise on a grid with 80×80 points.

see the characteristic lines along which the information in the system is transported. The first order differential operator has no smoothing effect which leads to a "rougher" realization than in the parabolic case. The effect of the noise is much stronger (also due to a different scaling in time).

Since the solution is not sufficiently smooth, a Galerkin approximation would lead to oscillations along the inflow boundary, which would spread over the whole domain over time. This

oscillations would be bounded but the result would still be unsatisfactory. With a Petrov–Galerkin approximation those oscillations are smoothed by the diffusion introduced by the test functions. The convergence of this method is shown in Figure 2(b) on a logarithmic scale and in Figure 3(b) on a linear scale, where we have the step size in time on the x-axis and the error in the y-direction. The reference slope is cx for a constant $c > 0$. The error was calculated in the same way as in the parabolic case with 100 paths for each discretization. The discretization grids were of size $2^2 \times 2^2$ up to $2^6 \times 2^6$ points, the "exact" solution was calculated with 2^9 approximation points in space and time. Although the number of simulations was quite small, the convergence results are visible.

Lemma 3.1 states the necessary convergence speed of the eigenvalues of the covariance operator, so that the overall convergence is not dominated by the convergence of the noise. The kernel function in the examples above provides this. If we take $q(x, y) = \exp(-\kappa|x - y|)$ instead of $q(x, y) = \exp(-\kappa|x - y|^2)$ as the kernel function of the covariance operator the condition of Lemma 3.1 on the decay of the eigenvalues would not be fulfilled. Figure 5 displays the errors calculated in the parabolic and hyperbolic case with the different kernel functions on a logarithmic scale. The difference is significant in the parabolic case since the Lemma states that the decay has to be much faster than in the hyperbolic case. The blue stars refer to the error calculated with the kernel function that satisfies Lemma 3.1 and the black circles to a covariance function with slower decay of the eignevalues. Figure 6 shows the same errors on a linear scale.

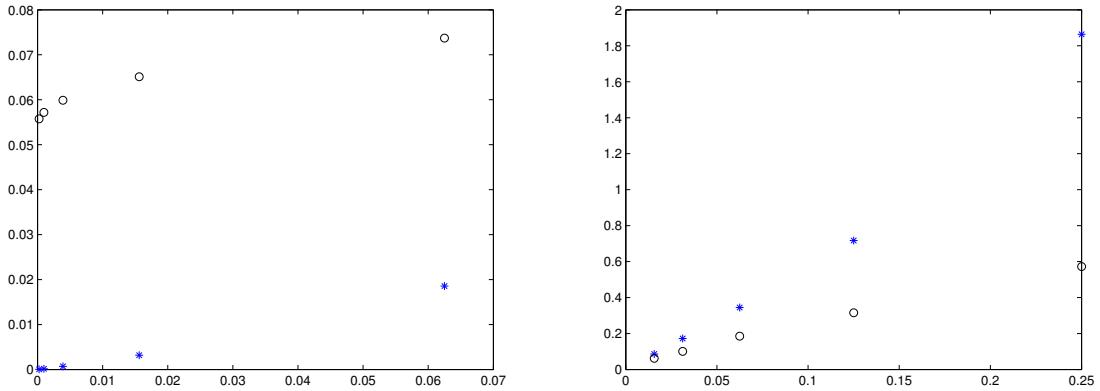


(a) Parabolic equation (b) Hyperbolic equation

FIGURE 5. Simulated errors with respect to the time discretization for different kernel functions on a logarithmic scale.

We conclude this section with a remark on the law of large numbers for Hilbert space valued processes, which is the essence of a Monte Carlo simulation.

Remark. A completely different question about the viability of the *law of large numbers* arises in the numerical implementation. Our approximations are pathwise. For practical reasons, people are often interested in the law of the solution of an SPDE and not in single paths. To ensure that the presented methods also approximate the law of the equations, we need a law of large numbers for Hilbert space valued processes. In [14] a result of this nature is given.



(a) Parabolic equation

(b) Hyperbolic equation

FIGURE 6. Simulated errors with respect to the time discretization for different kernel functions on a linear scale.

In Theorem 2.1 the authors show that if we are given a sequence of centered, independent H–valued random variables $\{Y_n\}$, that fulfill Chung's condition, i.e. $\sum_{n=1}^{\infty} n^{-2}\,\mathbb{E}(\|Y_n\|^2)$, then the law of large numbers holds. In our approximations we assume independently simulated paths and therefore we have independent random variables. These are all equally distributed with finite second moment and therefore Chung's condition holds, so we can apply the law of large numbers from [14]. Practically this will be done by the simulation of many paths which are used to do statistics.

6. Conclusions

In this paper we verified the theoretical convergence rates of a Galerkin–Crank–Nicolson approximation of a parabolic SPDE and a Petrov–Galerkin–Crank—Nicolson approximation of a hyperbolic SPDE. The results resemble the expected convergence rates. We addressed many practical problems arising with the actual implementation, like the approximation of the infinite dimensional noise and the convergence of the Monte Carlo method. An example showed that the convergence of the noise can dominate the overall convergence if the covariance function is not chosen appropriate. The sample sizes used here were quite small, however convergence rates are visible. The Euler–Maruyama scheme used here is in the case of a multiplicative equation converging much slower than the Milstein scheme, so further work might be the implementation of a Milstein scheme for a multiplicative equation.

References

[1] S. r. Asmussen and P. W. Glynn, *Stochastic Simulation: Algorithms and Analysis*, Stochastic Modelling and Applied Probability 57. New York, NY: Springer, 2007.

[2] A. Barth, *A finite element method for martingale-driven stochastic partial differential equations.* To appear in Comm. Stoch. Anal., January 2009.

[3] A. Barth and F. E. Benth, *Forward dynamics in energy markets – an infinite dimensional framework.* Submitted, 2010.

[4] A. Barth and A. Lang, *Almost sure convergence of a Galerkin–Milstein approximation for stochastic partial differential equations.* Submitted, June 2009.

[5] F. E. Benth, J. Šaltyte Benth, and S. Koekebakker, *Stochastic Modelling of Electricity and Related Markets*, World Scientific, Advanced Series on Statistical Science and Applied Probability - Vol. 11, 2008.
[6] P.-L. Chow, *Stochastic Partial Differential Equations*, Chapman & Hall/CRC Applied Mathematics and Nonlinear Science Series. Boca Raton, FL: Chapman & Hall/CRC, 2007.
[7] R. Cont and P. Tankov, *Financial Modelling with Jump Processes*, Chapman & Hall/CRC Financial Mathematics Series. Boca Raton, FL: Chapman and Hall/CRC, 2004.
[8] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and Its Applications. 44. Cambridge: Cambridge University Press, 1992.
[9] M. Geissert, M. Kovács, and S. Larsson, *Rate of weak convergence of the finite element method for the stochastic heat equation with additive noise*, BIT, 49 (2009), pp. 343–356.
[10] E. Hausenblas, *Approximation for semilinear stochastic evolution equations*, Potential Anal., 18 (2003), pp. 141–186.
[11] ———, *Wong–Zakai type approximation of spdes of Lévy noise*, Acta Appl. Math., 98 (2007), pp. 99–134.
[12] E. Hausenblas and I. Marchis, *A numerical approximation of parabolic stochastic partial differential equations driven by a Poisson random measure*, BIT, 46 (2006), pp. 773–811.
[13] E. Hausenblas and J. Seidler, *Stochastic convolutions driven by martingales: Maximal inequalities and exponential integrability*, Stochastic Anal. Appl., 26 (2008), pp. 98–119.
[14] J. Hoffmann-Jørgensen and G. Pisier, *The law of large numbers and the central limit theorem in banach spaces*, Ann. Probability, 4 (1976), pp. 587–599.
[15] A. Jentzen, *Taylor Expansions for Stochastic Partial Differential Equations*, PhD thesis, Johann Wolfgang Goethe Universität Frankfurt am Main, 2009.
[16] A. Jentzen and P. E. Kloeden, *The numerical approximation of stochastic partial differential equations*, Milan Journal of Mathematics, 77 (2009), pp. 205–244.
[17] O. Juan, R. Keriven, and G. Postelnicu, *Stochastic motion and the level set method in computer vision: Stochastic active contours*, International Journal of Computer Vision, 69 (2006), pp. 7–25.
[18] M. Kovács, S. Larsson, and F. Lindgren, *Strong convergence of the finite element method with truncated noise for semilinear parabolic stochastic equations with additive noise*, Numerical Algorithms, 53 (2010), pp. 309–320.
[19] A. Lang, *Simulation of Stochastic Partial Differential Equations and Stochastic Active Contours*, PhD thesis, Universität Mannheim, 2007.
[20] ———, *A Lax equivalence theorem for stochastic differential equations*. Submitted, 2009.
[21] ———, *Mean square convergence of a semidiscrete scheme for SPDEs of Zakai type driven by square integrable martingales*. To appear in Proc. of ICCS 2010, Proc. Comp. Sci., 2010.
[22] A. Lang, P.-L. Chow, and J. Potthoff, *Almost sure convergence of a semidiscrete Milstein scheme for SPDE's of Zakai type*. Submitted, April 2009.
[23] A. Lang and J. Potthoff, *Fast simulation of Gaussian random fields*. Submitted, 2009.
[24] G. J. Lord and J. Rougemont, *A numerical scheme for stochastic PDEs with Gevrey regularity*, IMA J. Numer. Anal., 24 (2004), pp. 587–604.
[25] Y. Lou and P. Christofides, *Nonlinear feedback control of surface roughness using a stochastic PDE*, in Decision and Control, 2006 45th IEEE Conference on, Dec. 2006, pp. 936–943.
[26] O. Moklyachuk, *Simulation of random processes with known correlation function with the help of karhunen–loeve decomposition*, Theory Stoch. Process., 13 (2007), pp. 163–169.
[27] D. Ni and P. D. Christofide, *Construction of stochastic PDEs for feedback control of surface roughness in thin film deposition*, in Proceedings of the 2005 American Control Conference, vol. 4, June 2005, pp. 2540–2547.
[28] S. Peszat and J. Zabczyk, *Stochastic Partial Differential Equations with Lévy Noise. An Evolution Equation Approach*, Encyclopedia of Mathematics and Its Applications 113. Cambridge: Cambridge University Press, 2007.
[29] T. H. Rydberg, *The normal inverse Gaussian Lévy process: Simulation and approximation*, Commun. Stat., Stochastic Models, 13 (1997), pp. 887–910.
[30] W. Schoutens, *Lévy Processes in Finance: Pricing Financial Derivatives*, Wiley, 2003.
[31] T. Shardlow, *Weak convergence of a numerical method for a stochastic heat equation*, BIT, 43 (2003), pp. 179–193.
[32] ———, *Numerical simulation of stochastic PDEs for excitable media*, J. Comput. Appl. Math., 175 (2005), pp. 429–446.

[33] Y. Yan, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal., 43 (2005), pp. 1363–1384.

(Andrea Barth)
Centre of Mathematics for Applications (CMA)
University of Oslo
P.O. Box 1053, Blindern
N–0316 Oslo, Norway
E-mail address: andrea.barth@cma.uio.no

(Annika Lang)
Fakultät für Mathematik und Informatik
Universität Mannheim
D–68131 Mannheim, Germany
E-mail address: lang@math.uni-mannheim.de