SPATIAL APPROXIMATION OF STOCHASTIC CONVOLUTIONS

MIHÁLY KOVÁCS[1], FREDRIK LINDGREN[1], AND STIG LARSSON[1,2]

Abstract. We study linear stochastic evolution partial differential equations driven by additive noise. We present a general and flexible framework for representing the infinite dimensional Wiener process which is driving the equation. Since the eigenfunctions and eigenvalues of the covariance operator of the process are usually not available for computations, we propose an expansion in an arbitrary frame. We show how to obtain error estimates when the truncated expansion is used in the equation. For the stochastic heat and wave equations we combine the truncated expansion with a standard finite element method and derive a priori bounds for the mean square error. Specializing the frame to biorthogonal wavelets in one variable, we show how the hierarchical structure, support and cancellation properties of the primal and dual bases lead to near sparsity and can be used to simplify the simulation of the noise and its update when new terms are added to the expansion.

1. Introduction

We study linear stochastic evolution problems of the form

$$\mathrm{d}X(t) = AX(t)\,\mathrm{d}t + B\,\mathrm{d}W(t),\ t > 0; \quad X(0) = 0, \tag{1.1}$$

where $X(t)$ is a stochastic process on a probability space $(\Omega, \mathcal{F}, \mathbf{P})$ with values in a separable Hilbert space H. The operator A is the infinitesimal generator of a strongly continuous semigroup e^{tA} of bounded linear operators on H, $W(t)$ is a Q-Wiener process on a Hilbert space U, and $B : U \to H$ is a bounded linear operator. The covariance operator Q of $W(t)$ is a self-adjoint, positive semidefinite, bounded linear operator on U.

Under appropriate assumptions, (1.1) has a unique weak solution which is given by the stochastic convolution (see Subsection 3.2 below),

$$X(t) = W_A(t) := \int_0^t \mathrm{e}^{(t-s)A} B\,\mathrm{d}W(s).$$

The motivation for studying the stochastic convolution W_A is that this is the first step towards studying more general evolution problems driven by additive noise of the form

$$\mathrm{d}X(t) = \big(AX(t) + f(X(t))\big)\,\mathrm{d}t + B\,\mathrm{d}W(t),\ t > 0; \quad X(0) = X_0.$$

Date: July 28, 2010.

1991 *Mathematics Subject Classification.* 65M60, 65T60, 60H15, 60H35, 65C30.

Key words and phrases. finite element, wavelet, stochastic heat equation, stochastic wave equation, Wiener process, additive noise, error estimate.

[1]Supported by the Swedish Research Council (VR).

[2]Supported by the Swedish Foundation for Strategic Research (SSF) through GMMC, the Gothenburg Mathematical Modelling Centre.

This can be given a rigorous meaning as an integral equation,

$$\begin{aligned}X(t) &= \mathrm{e}^{tA}X_0 + \int_0^t \mathrm{e}^{(t-s)A} f(X(s))\,\mathrm{d}s + \int_0^t \mathrm{e}^{(t-s)A} B\,\mathrm{d}W(s)\\ &= Y(t) + W_A(t),\end{aligned}$$

where Y satisfies

$$Y'(t) = AY(t) + f\big(Y(t) + W_A(t)\big),\ t > 0; \quad Y(0) = X_0.$$

Thus, once W_A is known, we may study Y by means of methods for evolution differential equations with random data. This abstract framework is sufficiently general to include the stochastic heat equation, the stochastic wave equation, and the stochastic Cahn-Hilliard equation. The above program; that is, splitting the solution of a semilinear problem into the stochastic convolution and the solution of a random PDE, is carried out, for example, for the stochastic Cahn-Hilliard equation in [5, 20, 22]. The analysis methods for W_A and Y are usually quite different, both on the PDE level and on the numerical level, and the present work is focused on the numerical approximation of the stochastic convolution W_A.

The Q-Wiener process is often represented as an orthogonal series,

$$W(t) = \sum_{k=1}^{\infty} \gamma_k^{1/2} \beta_k(t) f_k,$$

where $\{\gamma_k\}_{k=1}^{\infty}$ are the eigenvalues and $\{f_k\}_{k=1}^{\infty}$ an orthonormal basis of eigenvectors of the covariance operator Q and $\{\beta_k\}_{k=1}^{\infty}$ are independent real-valued Brownian motions. However, these eigenvectors are not always available for computations. We therefore propose an expansion in terms of an arbitrary frame which is not related to Q.

Let thus $\{\phi_j\}_{j\in\mathcal{J}}$, with countable index set $\mathcal{J}$, be a frame for U with corresponding dual frame $\{\tilde{\phi}_j\}_{j\in\mathcal{J}}$, so that $\langle \phi_j, \tilde{\phi}_j\rangle = \delta_{ij}$ and

$$f = \sum_{j\in\mathcal{J}} \langle f, \tilde{\phi}_j\rangle \phi_j, \quad f \in U,$$

see [9]. Let $J \subset \mathcal{J}$ be a finite set and define a projector P_J by

$$P_J f := \sum_{j\in J} \langle f, \tilde{\phi}_j\rangle \phi_j, \quad f \in U.$$

Define the truncated finite dimensional process

$$W^J(t) := \sum_{j\in J} \langle W(t), \tilde{\phi}_j\rangle \phi_j,\ t \geq 0.$$

and the corresponding stochastic convolution

$$W_A^J(t) := \int_0^t \mathrm{e}^{(t-s)A} B\,\mathrm{d}W^J(s).$$

In Theorem 3.2 we prove a formula for the mean square of the truncation error,

$$\mathbf{E}\Big(\|W_A(t) - W_A^J(t)\|^2\Big) = \int_0^t \|\mathrm{e}^{sA} B(I - P_J) Q^{1/2}\|_{\mathrm{HS}}^2\,\mathrm{d}s,$$

which is the basis for our further analysis. Here, $\|T\|_{\mathrm{HS}}$ denotes the Hilbert-Schmidt norm of a bounded linear operator $T\colon U \to H$ given by

$$\|T\|_{\mathrm{HS}}^2 = \sum_{k=1}^{\infty} \|Tf_k\|^2 \tag{1.2}$$

for some and, hence, for any orthonormal basis $\{f_k\}_{k=1}^{\infty}$ of U.

In Section 4 we introduce the deterministic heat and wave equations and their spatial approximation by a standard Galerkin finite element method. In particular, we consider the elliptic operator $\Lambda u = -\nabla \cdot (a\nabla u) + cu$ in a spatial domain $\mathcal{D}$ with boundary condition $u = 0$ on $\partial\mathcal{D}$ as an unbounded linear operator on the Hilbert space $H = L_2(\mathcal{D})$. Its finite element approximation is denoted Λ_h.

The stochastic heat equation is then of the form (1.1) with $A = -\Lambda$, $B = I$, $H = U = L_2(\mathcal{D})$ and the spatial finite element discretization leads to the truncated stochastic convolution,

$$W_{A_h}^J(t) := \int_0^t \mathrm{e}^{(t-s)A_h} P_h P_J \,\mathrm{d}W(s) = \int_0^t \mathrm{e}^{-(t-s)\Lambda_h} P_h P_J \,\mathrm{d}W(s),$$

where $A_h = -\Lambda_h$ and P_h is the orthogonal projector onto the finite element function space.

For the discretization error we prove in Theorem 5.1 the convergence estimate

$$\begin{aligned}\mathbf{E}\Big(\|W_A^J(t) - W_{A_h}^J(t)\|^2\Big) &\le Ch^{2\beta}\|\Lambda^{\frac{\beta-1}{2}} P_J Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \\ &= Ch^{2\beta} \sum_{j,k\in J} \langle \Lambda^{\frac{\beta-1}{2}}\phi_j, \Lambda^{\frac{\beta-1}{2}}\phi_k\rangle\langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle, \quad \beta \in [0, r],\end{aligned}$$

where h is the mesh size and $r \ge 2$ is the order of the finite element method. Similarly, for the truncation error we show in Theorem 5.2 that

$$\begin{aligned}\mathbf{E}\Big(\|W_A(t) - W_A^J(t)\|^2\Big) &\le \frac{1}{2}\|\Lambda^{-\frac{1}{2}}(I - P_J)Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \\ &= \frac{1}{2}\sum_{j,k\in\mathcal{J}\setminus J} \langle \Lambda^{-1}\phi_j, \phi_k\rangle\langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle.\end{aligned}$$

Analogous convergence estimates are proved for the stochastic wave equation in Section 6.

The first form of the above convergence estimates, expressed in terms of the Hilbert-Schmidt norm, can be evaluated easily by (1.2) if Λ and Q have a common eigenbasis; that is, if Λ and Q commute, and if W is expanded in the common eigenbasis. This approach is taken in several papers on numerical methods for stochastic partial differential equations, for example, [23], [30], and [31].

However, it is often not realistic to assume that Λ and Q commute. Then the latter form of the estimates is useful if the frames $\{\phi_j\}_{j\in\mathcal{J}}$, $\{\tilde{\phi}_j\}_{j\in\mathcal{J}}$ are chosen so that we can exploit decay properties and near sparsity of $\langle \Lambda^{\frac{\beta-1}{2}}\phi_j, \Lambda^{\frac{\beta-1}{2}}\phi_k\rangle$ and $\langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle$.

This is exemplified in Section 7, where we specialize to biorthogonal wavelets in one variable. Assuming that the covariance operator Q is an integral operator with smooth kernel, we show in Theorem 7.1 how to balance the discretization and truncation error so that the total error convergences with rate $O(h^2)$.

We also demonstrate how the hierarchical structure of the wavelet basis can be exploited to simplify the simulation of the Wiener process and its update when new terms are added to the expansion.

The study of numerical methods for evolution partial differential equations driven by noise started with the works of Grecksch and Kloeden [12] and Gyöngy and Nualart [15]. Further contributions include Allen, Novosel, and Zhang [1], Davie and Gaines [10], Du and Zhang [11], Gyöngy [13, 14], Hausenblas [17, 18], Shardlow [26], Müller-Gronbach and Ritter [23], Yan [31, 30], Quer-Sardanyons and Sanz-Solé [25], and Walsh [28, 29].

The present work was inspired by [1], where the noise is viewed as a martingale measure on space-time, which is approximated by a random function, piecewise constant in space-time. The resulting differential equation is then solved by the finite element method. This method does not generalize to spatially correlated noise and is therefore limited to one spatial dimension. This is because the solution of the stochastic heat equation with uncorrelated noise in multiple dimensions is not smooth enough to admit convergence estimates. The Haar wavelet was used in [11] together with correlated noise, but this work is also limited to one spatial dimension.

2. Preliminaries

Let H and U denote two separable real Hilbert spaces. We denote both their scalar products and norms by $\langle\cdot,\cdot\rangle$ and $\|\cdot\|$; they are distinguished by the context. The space of bounded linear operators from U to H is denoted by $\mathcal{B}(U,H)$ with standard norm also denoted $\|\cdot\|$. We write $T\geq 0$ if $T\in\mathcal{B}(H,H)$ is selfadjoint, positive semidefinite.

A countable subset $\{\phi_j\}_{j\in\mathcal{J}}\subset H$ is a frame for H if there exist $a,b>0$ such that

$$a\|f\|^2\leq\sum_{j\in\mathcal{J}}|\langle f,\phi_j\rangle|^2\leq b\|f\|^2,\quad f\in H. \tag{2.1}$$

The numbers a and b are called frame constants. Then there exists a frame $\{\tilde\phi_j\}_{j\in\mathcal{J}}$ with $\langle\phi_j,\tilde\phi_j\rangle=\delta_{ij}$ and

$$b^{-1}\|f\|^2\leq\sum_{j\in\mathcal{J}}|\langle f,\tilde\phi_j\rangle|^2\leq a^{-1}\|f\|^2,\quad f\in H.$$

The frame $\{\tilde\phi_j\}_{j\in\mathcal{J}}$ is called the dual frame of $\{\phi_j\}_{j\in\mathcal{J}}$, see, for example, [9]. We may now write

$$f=\sum_{j\in\mathcal{J}}\langle f,\tilde\phi_j\rangle\phi_j,\quad f\in H.$$

Let $\mathcal{L}_1(U,H)$ denote the set of nuclear operators from U to H; that is, $T\in\mathcal{L}_1(U,H)$ if $T\in\mathcal{B}(U,H)$ and there are sequences $\{a_j\}_{j=1}^\infty\subset H$, $\{b_j\}_{j=1}^\infty\subset U$ with $\sum_{j=1}^\infty\|a_j\|\|b_j\|<\infty$ and such that

$$Tf=\sum_{j=1}^\infty\langle f,b_j\rangle a_j,\quad f\in U. \tag{2.2}$$

These operators are also referred to as trace class operators from U to H. Clearly, trace class operators are compact. It is well known that $\mathcal{L}_1(U,H)$ is a Banach space with the norm

$$\|T\|_1 = \inf\Big\{ \sum_{j=1}^{\infty} \|a_j\|\|b_j\| : Tf = \sum_{j=1}^{\infty} \langle f, b_j\rangle a_j \Big\}.$$

Below we collect some facts about trace class operators.

Lemma 2.1. *Let $T \in \mathcal{L}_1(H,H)$ and $\{\phi_j\}_{j\in\mathcal{J}}$ be a frame with corresponding frame constants a and b. Then the trace of T,*

$$\mathrm{Tr}(T) = \sum_{j\in\mathcal{J}} \langle T\phi_j, \tilde{\phi}_j\rangle, \tag{2.3}$$

is well defined and is independent of the choice of frame. If, in addition, $T \geq 0$, then

$$a\,\mathrm{Tr}(T) \leq \sum_{j\in\mathcal{J}} \langle T\phi_j, \phi_j\rangle \leq b\,\mathrm{Tr}(T). \tag{2.4}$$

Proof. Since $T \in \mathcal{L}_1(H,H)$ we have (2.2). Then $\langle T\phi_k, \tilde{\phi}_k\rangle = \sum_{j=1}^{\infty} \langle \phi_k, b_j\rangle\langle a_j, \tilde{\phi}_k\rangle$ and hence

$$\begin{aligned}\sum_{k\in\mathcal{J}} |\langle T\phi_k, \tilde{\phi}_k\rangle| &\leq \sum_{k\in\mathcal{J}} \sum_{j=1}^{\infty} |\langle \phi_k, b_j\rangle\langle a_j, \tilde{\phi}_k\rangle| = \sum_{j=1}^{\infty} \sum_{k\in\mathcal{J}} |\langle \phi_k, b_j\rangle\langle a_j, \tilde{\phi}_k\rangle| \\ &\leq \sum_{j=1}^{\infty} \Big(\sum_{k\in\mathcal{J}} |\langle \phi_k, b_j\rangle|^2 \Big)^{\frac{1}{2}} \Big(\sum_{k\in\mathcal{J}} |\langle a_j, \tilde{\phi}_k\rangle|^2 \Big)^{\frac{1}{2}} \leq \sqrt{\frac{b}{a}} \sum_{j=1}^{\infty} \|a_j\|\|b_j\| < \infty.\end{aligned}$$

Therefore, the series in (2.3) converges absolutely and by Fubini's theorem

$$\sum_{k\in\mathcal{J}} \langle T\phi_k, \tilde{\phi}_k\rangle = \sum_{k\in\mathcal{J}} \sum_{j=1}^{\infty} \langle \phi_k, b_j\rangle\langle a_j, \tilde{\phi}_k\rangle = \sum_{j=1}^{\infty} \sum_{k\in\mathcal{J}} \langle \phi_k, b_j\rangle\langle a_j, \tilde{\phi}_k\rangle = \sum_{j=1}^{\infty} \langle a_j, b_j\rangle$$

is independent of the frame. This proves the first statement; for the second we refer to [9, p. 64]. ☐

Lemma 2.2. *If $T \in \mathcal{L}_1(H_1,H_2)$, $S_1 \in \mathcal{B}(H_2,H_3)$, and $S_2 \in \mathcal{B}(H_3,H_1)$, then $S_1T \in \mathcal{L}_1(H_1,H_3)$ and $TS_2 \in \mathcal{L}_1(H_3,H_2)$. Moreover, if $T \in \mathcal{L}_1(H_1,H_2)$, $S \in \mathcal{B}(H_2,H_1)$, then*

$$\mathrm{Tr}(ST) = \mathrm{Tr}(TS) \leq \|S\|\|T\|_1. \tag{2.5}$$

If $T \geq 0$, then $T \in \mathcal{L}_1(H,H)$ if and only if the series in (2.3) *converges for some orthonormal basis $\{\phi_j\}_{j\in\mathcal{J}}$ and in this case $\|T\|_1 = \mathrm{Tr}(T)$.*

Proof. The proofs for $H_i = H$ are given in [6, Appendix C]. The general case is proved in the same way. ☐

Lemma 2.3. *Let $T \in \mathcal{B}(U,H)$ and assume that $TT^* \in \mathcal{L}_1(H,H)$. Then $T^*T \in \mathcal{L}_1(U,U)$ and $\mathrm{Tr}(TT^*) = \mathrm{Tr}(T^*T)$.*

Proof. Since $TT^* \geq 0$, it follows from the spectral theorem and Lemma 2.2 that $\|TT^*\|_1 = \mathrm{Tr}(TT^*) = \sum_{i=1}^{\infty} \lambda_i$, where $\{\lambda_i\} \subset \mathbb{R}_+$ are the eigenvalues of TT^*. Let $\{e_i\} \subset H$ be corresponding orthonormal eigenvectors. Since $(T^*T)T^*e_i = T^*(TT^*)e_i = \lambda_i T^*e_i$, λ_i are eigenvalues of T^*T. By assumption TT^* is compact and, since $(T^*T)^2 = T^*TT^*T = T^*(TT^*)T$, it follows that $(T^*T)^2$ is compact and, hence, so is T^*T. Finally, as above, eigenvalues of T^*T are eigenvalues of TT^* and thus their eigenvalues coincide. Hence, $\mathrm{Tr}(TT^*) = \mathrm{Tr}(T^*T)$ by the last statement of Lemma 2.2. □

Finally, we recall that $T \in \mathcal{B}(U,H)$ is a Hilbert-Schmidt operator if

$$\|T\|_{\mathrm{HS}}^2 = \sum_{k=1}^{\infty} \|Tf_k\|^2 = \sum_{k=1}^{\infty} \langle T^*Tf_k, f_k\rangle < \infty, \tag{2.6}$$

for some and hence, for any orthonormal basis $\{f_k\}_{k=1}^{\infty}$ of U. in U. It is well known that the set of Hilbert-Schmidt operators, denoted by $\mathcal{L}_2(U,H)$, becomes a separable Hilbert space under the usual addition and scalar multiplication and with scalar product $\langle S,T\rangle = \sum_{k=1}^{\infty}\langle Sf_k, Tf_k\rangle$, where $\{f_k\}_{k=1}^{\infty}$ is any orthonormal basis of U. It is clear from the above that

$$\|T\|_{\mathrm{HS}}^2 = \mathrm{Tr}(T^*T) = \mathrm{Tr}(TT^*) = \|T^*\|_{\mathrm{HS}}^2$$

and, by (2.4), we have the norm equivalence

$$b^{-1} \sum_{j\in\mathcal{J}} \|T\phi_j\|^2 \le \|T\|_{\mathrm{HS}}^2 = \mathrm{Tr}(T^*T) \le a^{-1} \sum_{j\in\mathcal{J}} \|T\phi_j\|^2 \tag{2.7}$$

for any frame $\{\phi_j\}_{j\in\mathcal{J}}$ in U. This makes it possible to estimate the trace, or Hilbert-Schmidt norm, by using an arbitrary frame instead of an orthonormal basis, which will be crucial in the following. More generally, we have the following result for a product of operators.

Lemma 2.4. *Let $Q \in \mathcal{B}(H)$ with $Q \ge 0$ and with an orthonormal basis of eigenvectors. Let $T \in \mathcal{B}(H)$ and let $\{\phi_j\}_{j\in\mathcal{J}}$ be a frame for H. If $QT^*T \in \mathcal{L}_1(H,H)$, then $\|TQ^{\frac12}\|_{\mathrm{HS}} < \infty$ and*

$$\mathrm{Tr}(TQT^*) = \|TQ^{\frac12}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(QT^*T) = \sum_{j,k\in\mathcal{J}} \langle T\phi_j, T\phi_k\rangle \langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle.$$

Proof. Let $\{(\gamma_k, f_k)\}_{k=1}^{\infty}$ be eigenpairs of Q, cf. Remark 3.1. Since QT^*T is trace class, we may use (2.3) to expand $\mathrm{Tr}(QT^*T)$ in $\{f_k\}_{k=1}^{\infty}$:

$$\begin{aligned}\mathrm{Tr}(QT^*T) &= \sum_{k=1}^{\infty}\langle QT^*Tf_k, f_k\rangle = \sum_{k=1}^{\infty}\langle Tf_k, TQf_k\rangle = \sum_{k=1}^{\infty}\gamma_k\langle Tf_k, Tf_k\rangle \\ &= \sum_{k=1}^{\infty}\|TQ^{\frac12}f_k\|^2 = \sum_{k=1}^{\infty}\|TQ^{\frac12}f_k\|^2 = \|TQ^{\frac12}\|_{\mathrm{HS}}^2 = \mathrm{Tr}(TQT^*),\end{aligned}$$

where (2.6) was finally used. On the other hand, by expanding in $\{\phi_j\}_{j\in\mathcal{J}}$ and using $Q\tilde{\phi}_j = \sum_{k\in\mathcal{J}}\langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle\phi_k$, we conclude

$$\mathrm{Tr}(QT^*T) = \sum_{j\in\mathcal{J}}\langle QT^*T\phi_j, \tilde{\phi}_j\rangle = \sum_{j\in\mathcal{J}}\langle T\phi_j, TQ\tilde{\phi}_j\rangle = \sum_{j,k\in\mathcal{J}}\langle T\phi_j, T\phi_k\rangle\langle Q\tilde{\phi}_j, \tilde{\phi}_k\rangle.$$

□

3. Approximation of the stochastic convolution

3.1. **Wiener process.** Let $(\Omega, \mathcal{F}, \mathbf{P}, \{\mathcal{F}_t\}_{t\ge0})$ be a filtered probability space. Let U be a separable Hilbert space and $Q \in \mathcal{B}(U,U)$ with $Q \ge 0$ (selfadjoint, positive semidefinite). Let $\{W(t)\}_{t\ge0}$ be a U-valued stochastic process on $(\Omega, \mathcal{F}, \mathbf{P})$ which is adapted; that is, $W(t)$ is $\mathcal{F}_t$-measurable. We say that W is a *Q-Wiener process in U* if

(i) $W(0) = 0$,
(ii) W has continuous trajectories (almost surely),
(iii) W has independent increments,
(iv) $W(t) - W(s)$ is a U-valued Gaussian random variable with zero mean and covariance operator $(t-s)Q$ for $0 \le s \le t$.

The last statement means that Q is the unique operator defined by

$$\mathbf{E}\Big(\big\langle (W(t) - W(s)), x\big\rangle\big\langle (W(t) - W(s)), y\big\rangle\Big) = (t-s)\langle Qx, y\rangle, \quad x, y \in U. \tag{3.1}$$

Condition (iv) implies that $\mathrm{Tr}(Q) < \infty$ because the covariance operator of a Gaussian random variable is necessarily of trace class, see [6, Proposition 2.15]. Therefore, W is also called a nuclear Wiener process.

A nuclear Wiener process can be constructed starting from its covariance operator Q and the construction extends to the case when $\mathrm{Tr}(Q) = \infty$ in the following way. Let $Q \in \mathcal{B}(U, U)$ with $Q \ge 0$. The *Cameron-Martin space* is defined as $U_0 := Q^{\frac{1}{2}} U$ endowed with the scalar product $\langle x, y\rangle_0 := \langle Q^{-\frac{1}{2}} x, Q^{-\frac{1}{2}} y\rangle$, where Q^{-1} is understood as the pseudo-inverse if Q is not injective. Let $\{e_j\}_{j=1}^{\infty}$ be an orthonormal basis for U_0, let $\{\beta_j\}_{j=1}^{\infty}$ be mutually independent real-valued Brownian motions on $(\Omega, \mathcal{F}, \mathbf{P})$.

Let $L_2(\Omega, U)$ denote the space of square integrable U-valued random variables endowed with the usual norm

$$\|X\|_{L_2(\Omega,U)} = \Big(\mathbf{E}\big(\|X\|_U^2\big)\Big)^{1/2} = \Big(\int_\Omega \|X(\omega)\|_U^2 \, \mathrm{d}\mathbf{P}(\omega)\Big)^{1/2}.$$

If $\mathrm{Tr}(Q) < \infty$, then the series

$$W(t) := \sum_{k=1}^{\infty} \beta_k(t) e_k \tag{3.2}$$

converges in $L_2(\Omega, U)$ to a U-valued stochastic process, which has a version that is a nuclear Q-Wiener process, see [6, Section 4] and [24, Section 2].

If $\mathrm{Tr}(Q) = \infty$, then the series (3.2) does not converge in $L_2(\Omega, U)$. However, it converges in $L_2(\Omega, U_1)$ for a suitable (usually larger) space U_1 (see [6, Section 4.3.1]) to a U_1-valued stochastic process, which has a version that is a U_1-valued nuclear Wiener process. The constructed process, still denoted by $W(t)$, is called a *cylindrical Q-Wiener process in U*. Also, it is easy to see that

$$W_x(t) = \sum_{k=1}^{\infty} \beta_k(t) \langle e_k, x\rangle, \quad x \in U, \tag{3.3}$$

exists in $L_2(\Omega, \mathbb{R})$ and defines a real-valued Wiener process (Brownian motion) satisfying

$$\mathbf{E}\Big(W_x(t) W_y(t)\Big) = t\langle Qx, y\rangle, \quad x, y \in U, \tag{3.4}$$

cf. (3.1). Hence, we may write formally $\langle W(t), x\rangle = W_x(t)$ although the process $W(t)$ constructed from (3.2) takes values in U_1.

In either case, $\mathrm{Tr}(Q) < \infty$ or $\mathrm{Tr}(Q) = \infty$, we denote by $W(t)$ the series in (3.2), which is formal in case $\mathrm{Tr}(Q) = \infty$, and call it a *Q-Wiener process* in U.

Remark 3.1. It is often the case that there is an orthonormal basis $\{f_k\}_{k=1}^{\infty}$ in U consisting of eigenvectors of Q with corresponding non-negative eigenvalues $\{\gamma_k\}_{k=1}^{\infty}$.

Then $e_k = Q^{1/2} f_k = \gamma_k^{1/2} f_k$ is an orthonormal basis for U_0 and, in particular, (3.2) becomes

$$W(t) = \sum_{k=1}^{\infty} \gamma_k^{1/2} \beta_k(t) f_k.$$

However, we prefer to avoid the eigenvector expansion of $W(t)$.

3.2. **Stochastic convolution.** In what follows we need a simplified case of the stochastic integral, namely where the integrand is deterministic. In this case the class of integrands can be easily described. Let $F\colon [0,\infty) \to \mathcal{L}_2(U_0, H)$ be a measurable function, where $\mathcal{L}_2(U_0, H)$ is regarded as a Hilbert space endowed with its Borel sigma algebra, and assume that F is square integrable,

$$\int_0^t \|F(s)\|^2_{\mathcal{L}_2(U_0,H)} \, \mathrm{d}s = \int_0^t \|F(s) Q^{\frac{1}{2}}\|^2_{\mathrm{HS}} \, \mathrm{d}s < \infty.$$

Then the stochastic integral $\int_0^t F(s) \, \mathrm{d}W(s)$ is a well defined Gaussian random variable with covariance operator

$$Q_F(t)x = \int_0^t F(s) Q F^*(s) x \, \mathrm{d}s, \quad x \in H,$$

and the Itô isometry,

$$\Big\| \int_0^t F(s) \, \mathrm{d}W(s) \Big\|^2_{L_2(\Omega,H)} = \int_0^t \|F(s) Q^{\frac{1}{2}}\|^2_{\mathrm{HS}} \, \mathrm{d}s = \mathrm{Tr}(Q_F(t)) \tag{3.5}$$

holds, see [6, Chapter 4] and [24, Chapter 2]. In particular, let A generate a C_0-semigroup e^{tA} on H and let $B \in \mathcal{B}(U, H)$. Assume that the operator $Q_A(t)$, defined by

$$Q_A(t)x = \int_0^t \mathrm{e}^{sA} B Q B^* \mathrm{e}^{sA^*} x \, \mathrm{d}s, \tag{3.6}$$

has finite trace for all $t \geq 0$. Note that, by (2.5), the latter always holds in case $\mathrm{Tr}(Q) < \infty$. Then the stochastic convolution,

$$W_A(t) = \int_0^t \mathrm{e}^{(t-s)A} B \, \mathrm{d}W(s), \tag{3.7}$$

exists and defines an H-valued a Gaussian random variable with covariance operator $Q_A(t)$. Furthermore, W_A is the unique weak solution of

$$\mathrm{d}X(t) = AX(t) \, \mathrm{d}t + B \, \mathrm{d}W(t), \ t > 0; \quad X(0) = 0.$$

More precisely, this means that W_A is the unique (up to modification) solution of

$$\langle X(t), \eta \rangle = \int_0^t \langle X(s), A^* \eta \rangle \, \mathrm{d}s + \int_0^t l_\eta B \, \mathrm{d}W(s), \quad t \geq 0, \quad \forall \eta \in D(A^*), \tag{3.8}$$

where $l_\eta \colon H \to \mathbb{R}$ is given by $l_\eta x = \langle x, \eta \rangle$ (see [6, Theorem 5.4]).

3.3. **Truncation of the Wiener process.** We now approximate the stochastic convolution by truncating the expansion of $W(t)$ in an arbitrary frame. Thus, let $\{\phi_j\}_{j\in\mathcal{J}} \subset U$ be a frame for U with frame constants a, b and dual frame $\{\tilde{\phi}_j\}_{j\in\mathcal{J}}$. Let $J \subset \mathcal{J}$ be a finite set and define a projection onto $S_J := \mathrm{span}(\phi_j)_{j\in J}$ by

$$P_J x := \sum_{j\in J} \langle x, \tilde{\phi}_j\rangle \phi_j, \quad x \in U.$$

The adjoint P_J^* of P_J is given by $P_J^* x := \sum_{j\in J}\langle x, \phi_j\rangle \tilde{\phi}_j,\ x \in U$. Set

$$W^J(t) := \sum_{j\in J} \langle W(t), \tilde{\phi}_j\rangle \phi_j, \tag{3.9}$$

where $\langle W(t), \tilde{\phi}_j\rangle = W_{\tilde{\phi}_j}(t) = \sum_{k=1}^\infty \beta_k(t)\langle e_k, \tilde{\phi}_j\rangle$ is well defined by (3.3) even if $\mathrm{Tr}(Q) = \infty$ and thus W is not necessarily U-valued.

Lemma 3.1. *If $\{W(t)\}_{t\geq 0}$ is a Q-Wiener process in U given formally by* (3.2), *then the process $\{W^J(t)\}_{t\geq 0}$ in* (3.9) *is a nuclear $Q_J = P_J Q P_J^*$-Wiener process in U.*

Proof. We have that

$$\begin{aligned} W^J(t) &= \sum_{j\in J} W_{\tilde{\phi}_j}(t)\phi_j = \sum_{j\in J}\sum_{k=1}^\infty \beta_k(t)\langle e_k, \tilde{\phi}_j\rangle\phi_j \\ &= \sum_{k=1}^\infty \beta_k(t) \sum_{j\in J}\langle e_k, \tilde{\phi}_j\rangle \phi_j = \sum_{k=1}^\infty \beta_k(t) P_J e_k, \end{aligned} \tag{3.10}$$

where the latter series converges in $L_2(\Omega, U)$. The continuity of the paths follows from the fact that the processes $\{W_{\tilde{\phi}_j}(t)\}$, $j \in J$, are real-valued Brownian motions and that the index set J is finite. That the increments are independent and have a Gaussian law with the proper covariance operator can be verified from (3.10). □

Note that if $\mathrm{Tr}(Q) < \infty$, then $W^J(t) = P_J W(t)$ and the lemma above is even more straightforward. We define the corresponding stochastic convolution

$$W_A^J(t) := \int_0^t \mathrm{e}^{(t-s)A} B \, \mathrm{d}W^J(s), \tag{3.11}$$

which exists as $\mathrm{Tr}(Q_J) < \infty$. Next we provide a formula for the truncation error.

Theorem 3.2. *Let $Q \in \mathcal{B}(U, U)$ with $Q \geq 0$. Let A generate a C_0-semigroup e^{tA} on H, let $B \in \mathcal{B}(U, H)$, and let $W(t)$ be a Q-Wiener process in U. Assume that $\mathrm{Tr}(Q_A(t)) < \infty$, $t \geq 0$, where $Q_A(t)$ is defined in* (3.6). *Then the stochastic convolutions in* (3.7) *and* (3.11) *are well defined and*

$$W_A(t) - W_A^J(t) = \int_0^t \mathrm{e}^{(t-s)A} B(I - P_J)\, \mathrm{d}W(s)$$

and

$$\begin{aligned} \mathbf{E}\Big(\|W_A(t) - W_A^J(t)\|^2\Big) &= \int_0^t \mathrm{Tr}\left(\mathrm{e}^{sA}B(I-P_J)Q(I-P_J)^*B^*\mathrm{e}^{sA^*}\right)\mathrm{d}s \\ &= \int_0^t \|\mathrm{e}^{sA}B(I-P_J)Q^{1/2}\|_{\mathrm{HS}}^2 \,\mathrm{d}s. \end{aligned} \tag{3.12}$$

Proof. We first show that $W_A^J(t) = Z(t)$ a.s., where

$$Z(t) = \int_0^t \mathrm{e}^{(t-s)A} B P_J \, \mathrm{d}W(s).$$

We have that Z and W_A^J are, respectively, the unique solutions of

$$\langle X(t), \eta \rangle = \int_0^t \langle X(s), A^* \eta \rangle \, \mathrm{d}s + \int_0^t l_\eta B P_J \, \mathrm{d}W(s), \ t \geq 0, \ \forall \eta \in D(A^*),$$

and

$$\langle X(t), \eta \rangle = \int_0^t \langle X(s), A^* \eta \rangle \, \mathrm{d}s + \int_0^t l_\eta B \, \mathrm{d}W^J(s), \ t \geq 0, \ \forall \eta \in D(A^*),$$

where $l_\eta \colon H \to \mathbb{R}$ is given by $l_\eta x = \langle x, \eta \rangle$; cf. (3.8). Since $\mathrm{Tr}(Q_J) < \infty$, it follows that W^J is U-valued and $\int_0^t l_\eta B \, \mathrm{d}W^J(s) = l_\eta B W^J(t)$. A simple calculation, similar to that in the proof of Lemma 3.1, shows that

$$l_\eta B W^J(t) = \sum_{k=1}^\infty \beta_k(t) l_\eta B P_J e_k.$$

Now it is not hard to see that the latter equals to $\int_0^t l_\eta B P_J \, \mathrm{d}W(s)$, almost surely and hence the claim is proved. Therefore, $W_A(t) - W_A^J(t) = \int_0^t \mathrm{e}^{(t-s)A} B(I - P_J) \, \mathrm{d}W(s)$ and thus (3.12) follows by Itô's isometry (3.5). □

4. The finite element method for the deterministic problem

In this section we set the deterministic heat and wave equations in the form

$$X'(t) = AX(t), \ t > 0; \quad X(0) = X_0.$$

We also consider spatial approximation by the finite element method and recall some error estimates.

4.1. **An elliptic operator.**

Let $\mathcal{D} \subset \mathrm{R}^d$, $d = 1, 2, 3$, be a bounded spatial domain with sufficiently smooth boundary $\partial\mathcal{D}$. We introduce the elliptic operator

$$\Lambda u := -\nabla \cdot (a \nabla u) + cu, \quad \text{in } \mathcal{D},$$

where a, c are smooth coefficients with $a(x) \geq a_0 > 0$ and $c(x) \geq 0$ for all $x \in \mathcal{D}$. Together with the boundary condition $u = 0$ on $\partial\mathcal{D}$ this defines an unbounded operator Λ in $L_2(\mathcal{D})$ with domain of definition $D(\Lambda) = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$.

In order to describe regularity of fractional order we introduce the norms

$$\|v\|_{\dot{H}^\beta} = \|\Lambda^{\beta/2} v\| = \Big(\sum_{j=1}^\infty \lambda_j^\beta \langle v, \varphi_j \rangle^2 \Big)^{1/2}, \quad \beta \in \mathrm{R}, \tag{4.1}$$

where $\langle \cdot, \cdot \rangle$, $\|\cdot\|$ are the scalar product and norm in $L_2(\mathcal{D})$ and λ_j, φ_j denote the eigenvalues and corresponding orthonormal eigenvectors of Λ. The corresponding spaces are

$$\dot{H}^\beta = D(A^{\beta/2}), \quad \beta \geq 0,$$

and, for $\beta < 0$, $\dot{H}^\beta$ is the closure of $L_2(\mathcal{D})$ with respect to the norm in (4.1).

Clearly $\dot{H}^0 = L_2(\mathcal{D})$, and it is known that, for integer $\beta > 0$, these spaces can be described in terms of standard Sobolev spaces and that the norms are equivalent to the standard Sobolev norms. For example, $\dot H^1 = H_0^1(\mathcal{D})$ and $\dot H^2 = H^2(\mathcal{D}) \cap H_0^1(\mathcal{D})$ with

$$\|v\|_{\dot H^1} \equiv \|v\|_{H^1},\ v \in \dot H^1; \quad \|v\|_{\dot H^2} \equiv \|v\|_{H^2},\ v \in \dot H^2,$$

see [27, Lemma 3.1]. The spaces of negative order can be identified with dual spaces, $\dot H^{-\beta} = (\dot H^\beta)^*$ with $\|f\|_{\dot H^{-\beta}} = \sup_v \langle f, v\rangle / \|v\|_{\dot H^\beta}$.

We now introduce the standard finite element method. For this purpose we consider the equation $\Lambda u = f$. Its weak formulation is: find $u \in H_0^1(\mathcal{D}) = \dot H^1$ such that

$$a(u, v) = \langle f, v\rangle, \quad \forall v \in \dot H^1, \tag{4.2}$$

where $a(u, v) = \langle a\nabla u, \nabla v\rangle + \langle cu, v\rangle$ is the bilinear form associated with Λ.

Let $\{\mathcal{T}_h\}$ be a regular family of triangulations of $\mathcal{D}$ with meshsize h. Let $\{V_h\}_{0<h<1}$ be a family of finite dimensional subspaces of $\dot H^1$, where each V_h consists of continuous piecewise polynomials of degree $\le r-1$ ($r \ge 2$) with respect to a triangulation $\mathcal{T}_h$.

The approximate solution $u_h \in V_h$ of (4.2) is defined by

$$a(u_h, \chi) = \langle f, \chi\rangle, \quad \forall \chi \in V_h. \tag{4.3}$$

We define orthogonal projectors $P_h \colon \dot H^0 \to V_h$ and $R_h \colon \dot H^1 \to V_h$ by

$$\langle P_h f, \chi\rangle = \langle f, \chi\rangle, \quad a(R_h v, \chi) = \langle v, \chi\rangle, \quad \forall f \in \dot H^0,\ \forall v \in \dot H^1,\ \forall \chi \in V_h.$$

We also define the linear operator $\Lambda_h \colon V_h \to V_h$ by

$$\langle \Lambda_h \psi, \chi\rangle = a(\psi, \chi), \quad \forall \psi, \chi \in V_h,$$

so that equation (4.3) can be written $\Lambda_h u_h = P_h f$.

Our assumptions about the finite element method are summarized in the following error estimate:

$$\|R_h v - v\| \le C h^r \|v\|_{\dot H^r}, \quad \forall v \in \dot H^r. \tag{4.4}$$

For $d = 1$ this holds in great generality. For $d = 2, 3$ this holds for piecewise linear finite elements (with $r = 2$) in convex polygonal domains $\mathcal{D}$. For domains with curved boundary, and for higher order elements, there are additional difficulties concerning the approximation near the boundary, which we do not address here, see [27]. Actually, $v \in H^r(\mathcal{D}) \cap H_0^1(\mathcal{D})$ would be sufficient for the error estimate in (4.4) but the present formulation is more convenient.

4.2. **The deterministic heat equation.** We now consider the parabolic problem

$$u'(t) + \Lambda u(t) = 0,\ t > 0; \quad u(0) = v,$$

and its spatially semidiscrete finite element approximation

$$u_h'(t) + \Lambda_h u_h(t) = 0,\ t > 0; \quad u_h(0) = P_h v.$$

Their solutions given by the analytic semigroups on $H = \dot H^0$ generated by $A = -\Lambda$ and $A_h = -\Lambda_h$, respectively,

$$u(t) = \mathrm{e}^{-t\Lambda} v = \sum_{j=1}^{\infty} \mathrm{e}^{-t\lambda_j} \langle v, \varphi_j\rangle \varphi_j, \quad u_h(t) = \mathrm{e}^{-t\Lambda_h} P_h v = \sum_{j=1}^{N_h} \mathrm{e}^{-t\lambda_{h,j}} \langle v, \varphi_{h,j}\rangle \varphi_{h,j}.$$

Here $\{(\lambda_j,\varphi_j)\}_{j=1}^{\infty}$, $\{(\lambda_{h,j},\varphi_{h,j})\}_{j=1}^{N_h}$ are orthonormal eigenpairs of Λ and Λ_h, respectively. We will use the smoothing property

$$\int_0^t \|\mathrm{e}^{-s\Lambda}v\|^2\,\mathrm{d}s \le \frac{1}{2}\|v\|^2. \tag{4.5}$$

Finally, we introduce the error operator

$$F_h(t)v = \mathrm{e}^{-t\Lambda_h}P_h v - \mathrm{e}^{-t\Lambda}v. \tag{4.6}$$

Under the above assumptions we have the following error estimate, where $0 \le \beta \le r$,

$$\Big(\int_0^t \|F_h(s)v\|^2\,\mathrm{d}s\Big)^{1/2} \le Ch^\beta \|v\|_{\dot H^{\beta-1}}, \quad t \ge 0. \tag{4.7}$$

This follows from [27, Theorem 2.5].

4.3. **The deterministic wave equation.** We now consider the wave equation,

$$u''(t) + \Lambda u(t) = 0,\ t>0; \quad u(0)=v_1,\ u'(0)=v_2, \tag{4.8}$$

and its spatially semidiscrete finite element approximation,

$$u_h''(t) + \Lambda_h u_h(t) = 0,\ t>0; \quad u_h(0)=P_h v,\ u_h'(0)=P_h w. \tag{4.9}$$

In the standard way we set

$$U = \begin{bmatrix} u \\ u' \end{bmatrix}, \quad V = \begin{bmatrix} v_1 \\ v_2 \end{bmatrix}, \quad A = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix}.$$

Then A is an unbounded operator on $H = \dot H^0 \times \dot H^{-1}$ with

$$D(A) = \Big\{ v \in H : Av = \begin{bmatrix} v_2 \\ -\Lambda v_1 \end{bmatrix} \in H = \dot H^0 \times \dot H^{-1} \Big\} = \dot H^1 \times \dot H^0.$$

Here Λ is regarded as a bounded linear operator $\dot H^1 \to \dot H^{-1}$. The operator A is the generator of a strongly continuous semigroup (C_0-semigroup) e^{tA} on H and

$$\mathrm{e}^{tA} = \begin{bmatrix} C(t) & \Lambda^{-1/2}S(t) \\ -\Lambda^{1/2}S(t) & C(t) \end{bmatrix},$$

where $C(t) = \cos(t\Lambda^{1/2})$ and $S(t) = \sin(t\Lambda^{1/2})$ are the cosine and sine operators. For example, using $\{(\lambda_j,\varphi_j)\}_{j=1}^{\infty}$, orthonormal eigenpairs of Λ, we have

$$\Lambda^{-1/2}S(t)v = \Lambda^{-1/2}\sin(t\Lambda^{1/2})v = \sum_{j=1}^{\infty} \lambda_j^{-1/2}\sin(t\lambda_j^{1/2})(v,\varphi_j)\varphi_j.$$

Defining A_h and e^{tA_h} in the analogous way,

$$A_h = \begin{bmatrix} 0 & I \\ -\Lambda_h & 0 \end{bmatrix}, \quad \mathrm{e}^{tA_h} = \begin{bmatrix} C_h(t) & \Lambda_h^{-1/2}S_h(t) \\ -\Lambda_h^{1/2}S_h(t) & C_h(t) \end{bmatrix},$$

where $C_h(t) = \cos(t\Lambda_h^{1/2})$ and $S_h(t) = \sin(t\Lambda_h^{1/2})$, we may write the solutions of (4.8) and (4.9) as

$$U(t) = \mathrm{e}^{tA}V, \quad U_h(t) = \mathrm{e}^{tA_h}P_h V.$$

We will find that it is relevant to focus on the error in the first component $u_h = U_{1,h}$ with initial-values $v_1 = 0$, $v_2 = v$ and define an error operator by

$$F_h(t)v = \Lambda_h^{-1/2}S_h(t)P_h v - \Lambda^{-1/2}S(t)v. \tag{4.10}$$

Under the above assumptions we have the error estimate, where $0 \le \beta \le r+1$,

$$\|F_h(t)v\| \le C(1+t)h^{\frac{r}{r+1}\beta}\|v\|_{\dot H^\beta}, \quad t \ge 0. \tag{4.11}$$

This follows from [21, Corollary 4.3, Theorem 5.3].

5. Application to the stochastic heat equation

We now consider the stochastic heat equation

$$\mathrm{d}X(t) + \Lambda X(t)\,\mathrm{d}t = \mathrm{d}W(t),\ t>0; \quad X(0)=0,$$

which is of the form (1.1) with $H = U = \dot H^0$, $A=-\Lambda$, W a Q-Wiener process on $U = \dot H^0$, and $B = I$. We thus study the stochastic convolutions

$$\begin{aligned} X(t) = W_A(t) &= \int_0^t \mathrm{e}^{(t-s)A}\,\mathrm{d}W(s) = \int_0^t \mathrm{e}^{-(t-s)\Lambda}\,\mathrm{d}W(s),\\ X^J(t) = W_A^J(t) &= \int_0^t \mathrm{e}^{(t-s)A}P_J\,\mathrm{d}W(s) = \int_0^t \mathrm{e}^{-(t-s)\Lambda}P_J\,\mathrm{d}W(s),\\ X_h^J(t) = W_{A_h}^J(t) &= \int_0^t \mathrm{e}^{(t-s)A_h}P_hP_J\,\mathrm{d}W(s) = \int_0^t \mathrm{e}^{-(t-s)\Lambda_h}P_hP_J\,\mathrm{d}W(s). \end{aligned}$$

The condition $\mathrm{Tr}(Q_A(t))<\infty$ now becomes, see (3.5), (3.6), (2.6), and (4.5),

$$\begin{aligned} \mathrm{Tr}(Q_A(t)) &= \int_0^t \|\mathrm{e}^{-s\Lambda}Q^{\frac12}\|_{\mathrm{HS}}^2\,\mathrm{d}s = \int_0^t \sum_{j=1}^\infty \|\mathrm{e}^{-s\Lambda}Q^{\frac12}\phi_j\|^2\,\mathrm{d}s\\ &= \sum_{j=1}^\infty \int_0^t \|\Lambda^{\frac12}\mathrm{e}^{-s\Lambda}\Lambda^{-\frac12}Q^{\frac12}\phi_j\|^2\,\mathrm{d}s \le \frac12\sum_{j=1}^\infty \|\Lambda^{-\frac12}Q^{\frac12}\phi_j\|^2\\ &= \frac12\|\Lambda^{-\frac12}Q^{\frac12}\|_{\mathrm{HS}}^2<\infty. \end{aligned} \tag{5.1}$$

This guarantees the existence of the stochastic convolutions, see Theorem 3.2. We begin with the discretization error.

Theorem 5.1. *Let $A=-\Lambda$, $A_h=-\Lambda_h$, and let W a Q-Wiener process in $\dot H^0$. Assume $\|\Lambda^{\frac{\beta-1}{2}}Q^{\frac12}\|_{\mathrm{HS}}<\infty$ for some $\beta\in[0,r]$. If $\{\phi_j\}_{j\in\mathcal J}$ is a frame for $\dot H^0$ with $\phi_j\in\dot H^{\beta-1}$, then*

$$\mathbf{E}\Big(\|W_A^J(t)-W_{A_h}^J(t)\|^2\Big) \le Ch^{2\beta}\|\Lambda^{\frac{\beta-1}{2}}P_JQ^{\frac12}\|_{\mathrm{HS}}^2. \tag{5.2}$$

If, in addition, Q has an orthonormal basis of eigenvectors and $\phi_j\in\dot H^{\beta-1}$, then

$$\mathbf{E}\Big(\|W_A^J(t)-W_{A_h}^J(t)\|^2\Big) \le Ch^{2\beta}\sum_{j,k\in J}\langle\Lambda^{\frac{\beta-1}{2}}\phi_j,\Lambda^{\frac{\beta-1}{2}}\phi_k\rangle\langle Q\tilde\phi_j,\tilde\phi_k\rangle. \tag{5.3}$$

Proof. With F_h as in (4.6), we have

$$W_A^J(t)-W_{A_h}^J(t) = \int_0^t F_h(t-s)P_J\,\mathrm{d}W(s)$$

and hence, by using (3.5), (2.6), (4.7), and an orthonormal basis,

$$\mathbf{E}\Big(\|W_A^J(t)-W_{A_h}^J(t)\|^2\Big)=\sum_{j=1}^{\infty}\int_0^t\|F_h(s)P_JQ^{\frac12}f_j\|^2\,\mathrm{d}s$$
$$\le Ch^{2\beta}\sum_{j=1}^{\infty}\|\Lambda^{\frac{\beta-1}{2}}P_JQ^{\frac12}f_j\|^2=Ch^{2\beta}\|\Lambda^{\frac{\beta-1}{2}}P_JQ^{\frac12}\|_{\mathrm{HS}}^2.$$

This proves (5.2). Using Lemma 2.4 with $T=\Lambda^{\frac{\beta-1}{2}}P_J$, we obtain

$$\|\Lambda^{\frac{\beta-1}{2}}P_JQ^{\frac12}\|_{\mathrm{HS}}^2=\sum_{j,k\in\mathcal{J}}\langle\Lambda^{\frac{\beta-1}{2}}P_J\phi_j,\Lambda^{\frac{\beta-1}{2}}P_J\phi_k\rangle\langle Q\tilde\phi_j,\tilde\phi_k\rangle$$
$$=\sum_{j,k\in J}\langle\Lambda^{\frac{\beta-1}{2}}\phi_j,\Lambda^{\frac{\beta-1}{2}}\phi_k\rangle\langle Q\tilde\phi_j,\tilde\phi_k\rangle,$$

which proves (5.3). □

We now consider the truncation error. We assume that $Q\Lambda^{-1}$ is trace class. By Lemma 2.4 with $T=\Lambda^{-\frac12}$ this implies that $\Lambda^{-\frac12}Q^{\frac12}$ is Hilbert-Schmidt as required in (5.1). Clearly, the two assumptions coincide when Λ and Q commute, in particular, when $Q=I$.

Theorem 5.2. *Let $A=-\Lambda$ and let W be a Q-Wiener process in $\dot H^0$, where Q has an orthonormal basis of eigenvectors. Assume $Q\Lambda^{-1}\in\mathcal{L}_1(\dot H^0,\dot H^0)$. If $\{\phi_j\}_{j\in\mathcal{J}}$ is a frame for $\dot H^0$, then*

$$\mathbf{E}\Big(\|W_A(t)-W_A^J(t)\|^2\Big)\le\frac12\|\Lambda^{-\frac12}(I-P_J)Q^{\frac12}\|_{\mathrm{HS}}^2$$
$$=\frac12\sum_{j,k\in\mathcal{J}\setminus J}\langle\Lambda^{-1}\phi_j,\phi_k\rangle\langle Q\tilde\phi_j,\tilde\phi_k\rangle.$$

Proof. By using (3.12), (4.5), and an orthonormal basis, we get

$$\mathbf{E}\Big(\|W_A(t)-W_A^J(t)\|^2\Big)=\sum_{k=1}^{\infty}\int_0^t\|\mathrm{e}^{-s\Lambda}(I-P_J)Q^{\frac12}e_k\|^2\,\mathrm{d}s$$
$$\le\frac12\sum_{k=1}^{\infty}\|\Lambda^{-\frac12}(I-P_J)Q^{\frac12}e_k\|^2=\frac12\|\Lambda^{-\frac12}(I-P_J)Q^{\frac12}\|_{\mathrm{HS}}^2.$$

Lemma 2.4 with $T=\Lambda^{-\frac12}(I-P_J)$ now gives

$$\|\Lambda^{-\frac12}(I-P_J)Q^{\frac12}\|_{\mathrm{HS}}^2=\sum_{j,k\in\mathcal{J}}\langle\Lambda^{-\frac12}(I-P_J)\phi_j,\Lambda^{-\frac12}(I-P_J)\phi_k\rangle\langle Q\tilde\phi_j,\tilde\phi_k\rangle$$
$$=\sum_{j,k\in\mathcal{J}\setminus J}\langle\Lambda^{-1}\phi_j,\phi_k\rangle\langle Q\tilde\phi_j,\tilde\phi_k\rangle.$$

□

The same framework, with $A=-\Lambda^2$, $A_h=-\Lambda_h^2$, applies to the linear stochastic Cahn-Hilliard equation,

$$\mathrm{d}X(t)+\Lambda^2X(t)\,\mathrm{d}t=\mathrm{d}W(t),\ t>0;\quad X(0)=0,$$

see [22], where an analog of (4.7) and error estimates for $W_A(t)-W_{A_h}(t)$ without truncation are proved. Theorems, analogous to Theorems 5.1–5.2, may then be proved but we refrain from giving the details. The importance of studying the W_A and its numerical approximation when handling the full semilinear stochastic Cahn-Hilliard equation was explained in the Introduction. We would like to mention that the error analysis of the finite element method for the stochastic Cahn-Hilliard equation is difficult not only because of the nonlinear term but also because the finite element method involves $A_h = -\Lambda_h^2$. This makes the error analysis for $F_h(t) = \mathrm{e}^{-t\Lambda_h^2} P_h - \mathrm{e}^{-t\Lambda^2}$ associated with the Cahn-Hilliard equation significantly more complicated than the one for the heat equation in (4.7); compare [22, Theorem 2.1] and [27, Theorem 2.5].

6. Application to the stochastic wave equation

We now consider the stochastic wave equation

$$\begin{aligned} \mathrm{d}X_1(t) &= X_2(t)\,\mathrm{d}t, \\ \mathrm{d}X_2(t) &= -\Lambda X_1(t) + \mathrm{d}W(t), \end{aligned} \qquad t > 0; \quad X_1(0) = X_2(0) = 0,$$

which is of the form (1.1) with $H = \dot{H}^0 \times \dot{H}^{-1}$,

$$A = \begin{bmatrix} 0 & I \\ -\Lambda & 0 \end{bmatrix}, \quad B = \begin{bmatrix} 0 \\ I \end{bmatrix},$$

as in Subsection 4.3, and W a Q-Wiener process on $U = \dot{H}^0$. We thus study the stochastic convolutions

$$\begin{aligned} X(t) = W_A(t) &= \int_0^t \mathrm{e}^{(t-s)A}\,\mathrm{d}W(s) = \int_0^t \begin{bmatrix} \Lambda^{-\frac12} S(t-s) \\ C(t-s) \end{bmatrix} \mathrm{d}W(s), \\ X^J(t) = W_A^J(t) &= \int_0^t \mathrm{e}^{(t-s)A} P_J\,\mathrm{d}W(s) = \int_0^t \begin{bmatrix} \Lambda^{-\frac12} S(t-s) \\ C(t-s) \end{bmatrix} P_J\,\mathrm{d}W(s), \\ X_h^J(t) = W_{A_h}^J(t) &= \int_0^t \mathrm{e}^{(t-s)A_h} P_h P_J\,\mathrm{d}W(s) = \int_0^t \begin{bmatrix} \Lambda_h^{-\frac12} S_h(t-s) P_h \\ C_h(t-s) P_h \end{bmatrix} P_J\,\mathrm{d}W(s). \end{aligned}$$

Estimating $\mathrm{Tr}(Q_A(t))$ by means of (3.5), an orthonormal basis, and the boundedness of the sine and cosine operators, we get

$$\begin{aligned} \mathrm{Tr}(Q_A(t)) &= \int_0^t \|\mathrm{e}^{sA} B Q^{\frac12}\|^2_{\mathcal{L}_2(U,H)}\,\mathrm{d}s = \int_0^t \sum_{j=1}^\infty \|\mathrm{e}^{sA} B Q^{\frac12}\phi_j\|_H^2\,\mathrm{d}s \\ &= \int_0^t \sum_{j=1}^\infty \big(\|\Lambda^{-\frac12} S(s) Q^{\frac12}\phi_j\|^2 + \|\Lambda^{-\frac12} C(s) Q^{\frac12}\phi_j\|^2 \big)\,\mathrm{d}s \\ &\le 2t \sum_{j=1}^\infty \|\Lambda^{-\frac12} Q^{\frac12}\phi_j\|^2 = 2t \|\Lambda^{-\frac12} Q^{\frac12}\|^2_{\mathrm{HS}}. \end{aligned} \tag{6.1}$$

We thus have $\mathrm{Tr}(Q_A(t)) < \infty$, and existence of the stochastic convolutions, under the same condition as for the heat equation, namely, $\|\Lambda^{-\frac12} Q^{\frac12}\|_{\mathrm{HS}} < \infty$, see (5.1).

We begin with the discretization error. We restrict the analysis to the first component $X_1^J = W_{A,1}^J$ in order to shorten the presentation.

Theorem 6.1. *Let W_A^J and $W_{A_h}^J$ be as above. Assume $\|\Lambda^{\frac{\beta-1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}} < \infty$ for some $\beta \in [0, r+1]$. If $\{\phi_j\}_{j\in\mathcal{J}}$ is a frame for $\dot H^0$ with $\phi_j \in \dot H^{\beta-1}$, then*

$$\mathbf{E}\Big(\|W_{A,1}^J(t) - W_{A_h,1}^J(t)\|^2\Big) \le C(t) h^{\frac{2r}{r+1}\beta} \|\Lambda^{\frac{\beta-1}{2}} P_J Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2. \tag{6.2}$$

If, in addition, Q has an orthonormal basis of eigenvectors, then

$$\mathbf{E}\Big(\|W_{A,1}^J(t) - W_{A_h,1}^J(t)\|^2\Big) \le C(t) h^{\frac{2r}{r+1}\beta} \sum_{j,k\in J} \langle \Lambda^{\frac{\beta-1}{2}} \phi_j, \Lambda^{\frac{\beta-1}{2}} \phi_k\rangle \langle Q\tilde\phi_j, \tilde\phi_k\rangle. \tag{6.3}$$

Proof. With F_h as in (4.10) we have

$$W_A^J(t) - W_{A_h}^J(t) = \int_0^t F_h(t-s) P_J \,\mathrm{d}W(s)$$

and hence, by using (3.5), (4.11), and an orthonormal basis,

$$\begin{aligned}\mathbf{E}\Big(\|W_{A,1}^J(t) - W_{A_h,1}^J(t)\|^2\Big) &= \int_0^t \sum_{j=1}^\infty \|F_h(s) P_J Q^{\frac{1}{2}} e_j\|^2 \,\mathrm{d}s \\ &\le C(t) h^{\frac{2r}{r+1}\beta} \sum_{j=1}^\infty \|\Lambda^{\frac{\beta-1}{2}} P_J Q^{\frac{1}{2}} e_j\|^2 = C(t) h^{\frac{2r}{r+1}\beta} \|\Lambda^{\frac{\beta-1}{2}} P_J Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2.\end{aligned}$$

This proves (6.2). The bound (6.3) is then obtained in the same was as (5.3). □

We now consider the truncation error. Recall from the discussion before Theorem 5.2 that $Q\Lambda^{-1} \in \mathcal{L}_1(\dot H^0, \dot H^0)$ implies $\|\Lambda^{-\frac{1}{2}} Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 < \infty$ as required in (6.1).

Theorem 6.2. *Let W_A^J and $W_{A_h}^J$ be as above. Assume that W is Q-Wiener process in $\dot H^0$ where Q has an orthonormal basis of eigenvectors. Assume $Q\Lambda^{-1} \in \mathcal{L}_1(\dot H^0, \dot H^0)$. If $\{\phi_j\}_{j\in\mathcal{J}}$ is a frame for $\dot H^0$, then*

$$\begin{aligned}\mathbf{E}\Big(\|W_A(t) - W_A^J(t)\|^2\Big) &\le 2t \|\Lambda^{-\frac{1}{2}} (I - P_J) Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2 \\ &= 2t \sum_{j,k\in\mathcal{J}\setminus J} \langle \Lambda^{-1}\phi_j, \phi_k\rangle \langle Q\tilde\phi_j, \tilde\phi_k\rangle.\end{aligned}$$

Proof. By using (3.12) and an orthonormal basis we get

$$\begin{aligned}\mathbf{E}\Big(\|W_A(t) - W_A^J(t)\|^2\Big) &= \int_0^t \sum_{k=1}^\infty \|\mathrm{e}^{sA}(I-P_J) Q^{\frac{1}{2}} e_k\|_H^2 \,\mathrm{d}s \\ &= \int_0^t \sum_{k=1}^\infty \big(\|\Lambda^{-\frac{1}{2}} S(s)(I-P_J) Q^{\frac{1}{2}} e_k\|^2 + \|\Lambda^{-\frac{1}{2}} C(s)(I-P_J) Q^{\frac{1}{2}} e_k\|^2\big) \,\mathrm{d}s \\ &\le 2t \sum_{k=1}^\infty \|\Lambda^{-\frac{1}{2}} (I-P_J) Q^{\frac{1}{2}} e_k\|^2 = 2t \|\Lambda^{-\frac{1}{2}} (I-P_J) Q^{\frac{1}{2}}\|_{\mathrm{HS}}^2.\end{aligned}$$

The proof is now completed in the same way as the proof of Theorem 5.2. □

7. Application to wavelets

In this section we investigate the error bounds for the heat equation in Section 5 when $d = 1$ and various assumptions on Q and choices of the frame $\{\phi_j\}_{j\in\mathcal{J}}$. The error bounds for the wave equation in Section 6 can be dealt with in a similar way.

7.1. **White noise.** Let $Q = I$ and $\{\phi_j\}_{j\in\mathcal{J}} \subset \dot{H}^{\beta-1}$ ($\beta \le r$) be a frame for H. Theorems 5.2 and 5.1 then yield

$$\begin{aligned}\mathbf{E}\Big(\|W_A(t) - W_{A_h}^J(t)\|^2\Big) &\le 2\mathbf{E}\Big(\|W_A(t) - W_A^J(t)\|^2\Big) + 2\mathbf{E}\Big(\|W_A^J(t) - W_{A_h}^J(t)\|^2\Big)\\ &\le \|\Lambda^{-\frac{1}{2}}(I-P_J)\|_{\mathrm{HS}}^2 + Ch^{2\beta}\|\Lambda^{\frac{\beta-1}{2}}P_J\|_{\mathrm{HS}}^2\\ &\le C\sum_{j\in\mathcal{J}\setminus J}\|\Lambda^{-\frac{1}{2}}\phi_j\|^2 + Ch^{2\beta}\sum_{j\in J}\|\Lambda^{\frac{\beta-1}{2}}\phi_j\|^2,\end{aligned}$$

where we used (2.7) to evaluate the Hilbert-Schmidt norm. Let now $d = 1$, $\mathcal{D} = (0,1)$, $\Lambda = -\frac{d^2}{dx^2}$ and $\phi_j = \psi_{l,k}$, $\mathcal{J} = \{j = (l,k) : k = 0,\dots,2^l-1,\ l = 0,1,\dots\}$, where $\{\{\psi_{l,k}\}_{k=0}^{2^l-1}\}_{l=0}^{\infty}$ is the Haar wavelet basis for $L_2(\mathcal{D})$. Then, with $\beta = 1$, $J = \{(l,k) \in \mathcal{J} : l \le N\}$, and anticipating the bound for $\langle\Lambda^{-1}\psi_{l,k},\psi_{l,k}\rangle$ in (7.12), we have

$$\begin{aligned}\mathbf{E}\Big(\|W_A(t) - W_{A_h}^J(t)\|^2\Big) &\le C\sum_{l=N+1}^{\infty}\sum_{k=0}^{2^l-1}\langle\Lambda^{-1}\psi_{l,k},\psi_{l,k}\rangle + Ch^2\sum_{l=0}^{N}\sum_{k=0}^{2^l-1}1\\ &\le C\sum_{l=N+1}^{\infty}\sum_{k=0}^{2^l-1}2^{-2l} + Ch^2\sum_{l=0}^{N}2^l \le C2^{-N} + Ch^2 2^N.\end{aligned}$$

To optimize the error estimate choose $h = 2^{-N}$ and obtain

$$\mathbf{E}\Big(\|W_A(t) - W_{A_h}^J(t)\|^2\Big) \le Ch.$$

If instead we choose $\phi_j(x) = \varphi_j(x) = \sqrt{2}\sin(\pi jx)$, $j = 1,2,\dots$, the orthonormal eigenfunctions of Λ, and again $\beta = 1$, we get

$$\mathbf{E}\Big(\|W_A(t) - W_{A_h}^N(t)\|^2\Big) \le C\sum_{j=N+1}^{\infty}\frac{\pi^2}{j^2} + Ch^2\sum_{j=1}^{N}1 \le C\frac{1}{N} + Ch^2N.$$

Optimizing by setting $h = \frac{1}{N}$, we obtain

$$\mathbf{E}\Big(\|W_A(t) - W_{A_h}^N(t)\|^2\Big) \le Ch.$$

Thus, in both cases we obtain the mean square rate of convergence $O(h^{\frac{1}{2}})$, which is optimal for $Q = I$. Note that without truncation we would have (cf. [30])

$$\mathbf{E}\Big(\|W_A(t) - W_{A_h}(t)\|^2\Big) \le Ch^{2\beta}\|\Lambda^{\frac{\beta-1}{2}}\|_{\mathrm{HS}}^2,$$

where $\|\Lambda^{\frac{\beta-1}{2}}\|_{\mathrm{HS}}^2 = \pi^2\sum_{j=0}^{\infty}j^{2(\beta-1)} < \infty$ if and only if $\beta < \frac{1}{2}$.

7.2. **Smoother noise.** When turning to concrete examples one usually assumes that the frame and its dual satisfy support and cancelation conditions. To make this more precise we assume that there is a levelwise organization of the frame; that is, $\mathcal{J} = \{(j,k) : j \in \mathbb{N},\ j \ge j_0,\ k \in J_j\}$, where J_j is an index set whose size depends on j and the spatial dimension d. Then the support and the cancelation conditions can be written as

(H1) $\mathrm{diam}(\mathrm{supp}\,\phi_{j,k}) \sim \mathrm{diam}(\mathrm{supp}\,\tilde{\phi}_{j,k}) \sim 2^{-j}$, $j \ge j_0$, (H2) for $f \in W^{\tilde{m},\infty}(\mathcal{D})$ we have

$$|\langle f, \phi_{j,k}\rangle| \le C2^{-j(s+d/2)}|f|_{W^{s,\infty}(\operatorname{supp}\phi_{j,k})}, \quad s \le \tilde{m},\ j \ge j_0,$$

and for $f \in W^{m,\infty}(\mathcal{D})$ we have

$$|\langle f, \tilde{\phi}_{j,k}\rangle| \le C2^{-j(s+d/2)}|f|_{W^{s,\infty}(\operatorname{supp}\tilde{\phi}_{j,k})}, \quad s \le m,\ j \ge j_0.$$

Here $\mathcal{D} \subset \mathbb{R}^d$ with polygonal or smooth boundary and $|f|_{W^{s,\infty}(\cdot)}$ denotes the usual seminorm. We remark that in the wavelet literature condition (2.1) is often referred to as H-stability or stability. For example, the Haar basis in one dimension satisfies the above conditions with $m = \tilde{m} = 1$. In multiple dimensions for nontrivial domains it is highly complicated to construct an explicit basis together with a dual basis satisfying these conditions. Even in one dimension for an interval the construction is tedious, but there are explicit wavelet bases (with explicit dual bases) satisfying (H1) and (H2) for all $m \le \tilde{m}$ with $m + \tilde{m}$ even, see [8]. Assuming a frame with properties (H1) and (H2) and enough regularity, one obtains decay estimates for scalar products like $\langle Q\tilde{\phi}_{i,l}, \tilde{\phi}_{j,k}\rangle$ and $\langle \Lambda^{\frac{\beta-1}{2}}\phi_{i,l}, \Lambda^{\frac{\beta-1}{2}}\phi_{j,k}\rangle$ needed for the error estimates in Theorems 5.1 and 5.2; see [7].

Finally, we demonstrate in a simple concrete example how to get optimal error estimates by choosing an appropriate frame if the noise is smooth enough. Let $d = 1$, $\mathcal{D} = (0,1)$, $U = H = L_2(\mathcal{D})$, $B = I$, and $\Lambda u = -(au')' + cu$ with smooth coefficients $a \ge a_0 > 0,\ c \ge 0$. Let Q be given as an integral operator $(Qf)(x) := \int_0^1 q(x,y)f(y)\,\mathrm{d}y$. Unless the functions a, c, q are very special, Λ and Q do not commute and their eigenfunctions are not known explicitly. Since Q is assumed to be given, one can simulate the truncated noise W^J efficiently, see Subsection 7.3.

We will use the wavelet basis constructed in [8]. It satisfies (H1) and (H2) with $m \le \tilde{m}$ and $m + \tilde{m}$ even. Moreover, for $j \in \mathbb{N}$, one obtains inverse estimates

$$\begin{aligned} &\|\phi_{j,k}\|_{H^s(\mathcal{D})} \le C2^{sj}\|\phi_{j,k}\|_{L_2(\mathcal{D})}, \quad 0 \le s \le \gamma, \\ &\|\tilde{\phi}_{j,k}\|_{H^s(\mathcal{D})} \le C2^{sj}\|\tilde{\phi}_{j,k}\|_{L_2(\mathcal{D})}, \quad 0 \le s \le \tilde{\gamma}, \end{aligned} \tag{7.1}$$

where $\gamma = m - \frac{1}{2}$ and $\tilde{\gamma}$ can be chosen as large as we want by using $\tilde{m}$ large in the construction (see also [4]). Further, the number of frame elements on level i; that is, $\#J_i$ (this index set is the same for the primal and dual frames), satisfies

$$\#J_i \le C2^i. \tag{7.2}$$

We also have a bound on the number of basis functions that have intersecting supports. For this purpose, let

$$\Delta_{jkil} := \operatorname{supp}\phi_{j,k} \cap \operatorname{supp}\phi_{i,l}. \tag{7.3}$$

Then, for $j \ge i$, the number of $\phi_{j,k}$ whose supports intersect the support of a fixed $\phi_{i,l}$ is given by

$$\#\{k \in J_j : \Delta_{jkil} \ne \emptyset\} \le C2^{j-i}. \tag{7.4}$$

Taking also into account the number of $\phi_{i,l}$ given by (7.2), and finally interchanging the roles of i and j, we conclude

$$\#\{l \in J_i,\ k \in J_j : \Delta_{jkil} \ne \emptyset\} \le C2^{\max(i,j)}. \tag{7.5}$$

The reason why (7.4) holds is that the construction in [8] may be performed so that, except for some boundary functions whose number is uniformly bounded in j, the $\phi_{j,k}$ are linear combinations of a uniformly bounded number of translates and dilates $\theta_{j+1,k}$ of a function θ with compact support. This is done in such a way that the supports of the $\phi_{j,k}$ move equally fast as the $\theta_{j+1,k}$ when k grows. More precisely, $\theta_{j,k}(x) = \theta(2^j x - k)$, and there exists a non-positive integer N and a non-negative integer M such that for all $j \geq j_0$ and $k \in J_j \setminus J_j^B$ (where J_j^B refer to the boundary functions), $\phi_{j,k}$ may be written as

$$\phi_{j,k} = \sum_{l=N}^{M} a_{j+1,k+l}\theta_{j+1,k+l}$$

for some real numbers $a_{j+1,k+l}$. To show (7.4) is then a matter of computing bounds on the number of k for which Δ_{jkil} is nonempty.

Theorem 7.1. *Let $d = 1$, $\mathcal{D} = (0,1)$, $U = H = L_2(\mathcal{D})$, $B = I$, and $A = -\Lambda u = -(-(au')' + cu)$ with smooth coefficients $a \geq a_0 > 0$, $c \geq 0$, and W_A as in Section 5. Let $(Qf)(x) := \int_0^1 q(x,y)f(y)\,\mathrm{d}y$ with $q \in W^{3,\infty}(\mathcal{D}\times\mathcal{D})$. Let $\{\phi_{j,k}\}$ be a frame with dual frame $\{\tilde{\phi}_{j,k}\}$ as constructed in* [8] *with properties (H1) and (H2) with $m \geq 2$ and $\tilde{m} \geq 2$ so large that* (7.1) *holds with $\gamma = \tilde{\gamma} = 1$. Then, for $J = \{(j,k) \in \mathcal{J} : j \leq N\}$ and $h = 2^{-N}$, we have*

$$\mathbf{E}\Big(\|W_A(t) - W_{A_h}^J(t)\|^2\Big) \leq Ch^4.$$

Proof. We use Theorems 5.1 and 5.2 with $\beta = 2$. We must bound $\langle Q\tilde{\phi}_{i,l}, \tilde{\phi}_{j,k}\rangle$, $\langle \Lambda^{-1}\phi_{i,l}, \phi_{j,k}\rangle$, and $\langle \Lambda^{\frac{1}{2}}\phi_{i,l}, \Lambda^{\frac{1}{2}}\phi_{j,k}\rangle$.

By using (H2), first with $s = 2$, then with $s = 1$, we obtain

$$\begin{aligned}
|\langle Q\tilde{\phi}_{i,l}, \tilde{\phi}_{j,k}\rangle| &\leq C2^{-j(2+1/2)}|Q\tilde{\phi}_{i,l}|_{W^{2,\infty}(\operatorname{supp}\tilde{\phi}_{j,k})} \\
&= C2^{-\frac{5}{2}j} \operatorname*{ess-sup}_{x\in\operatorname{supp}\tilde{\phi}_{j,k}} \big|\langle q''_{xx}(x,\cdot), \tilde{\phi}_{i,l}\rangle\big| \\
&\leq C2^{-\frac{5}{2}j}2^{-\frac{3}{2}i} \operatorname*{ess-sup}_{x\in\operatorname{supp}\tilde{\phi}_{j,k},\, y\in\operatorname{supp}\tilde{\phi}_{i,l}} \big|q'''_{xxy}(x,y)\big| \qquad (7.6)\\
&\leq C2^{-\frac{5}{2}j-\frac{3}{2}i}\|q\|_{W^{3,\infty}(\mathcal{D}\times\mathcal{D})} \leq C2^{-\frac{5}{2}j-\frac{3}{2}i}. \qquad (7.7)
\end{aligned}$$

Since Q is symmetric we have the same estimate with i and j interchanged, so that

$$|\langle Q\tilde{\phi}_{i,l}, \tilde{\phi}_{j,k}\rangle| \leq C2^{-\frac{5}{2}\max(i,j)-\frac{3}{2}\min(i,j)}, \qquad (7.8)$$

and, alternatively,

$$|\langle Q\tilde{\phi}_{i,l}, \tilde{\phi}_{j,k}\rangle| = \sqrt{\langle Q\tilde{\phi}_{i,l}, \tilde{\phi}_{j,k}\rangle\langle Q\tilde{\phi}_{j,k}, \tilde{\phi}_{i,l}\rangle} \leq C2^{-2(i+j)}. \qquad (7.9)$$

By our assumption on Λ we have $(\Lambda^{-1}u)(x) = \int_0^1 g(x,y)u(y)\,\mathrm{d}y$, where Green's function $g \in W^{1,\infty}(\mathcal{D}\times\mathcal{D})$. Thus, by (H2) with $s=1$ and (H1), we get

$$\begin{aligned}
|\langle \Lambda^{-1}\phi_{i,l},\phi_{j,k}\rangle| &\le C2^{-\frac32 j}|\Lambda^{-1}\phi_{i,l}|_{W^{1,\infty}(\operatorname{supp}\phi_{j,k})}\\
&= C2^{-\frac32 j}\operatorname*{ess-sup}_{x\in\operatorname{supp}\phi_{j,k}}\Big|\int_0^1 g_x'(x,y)\phi_{i,l}(y)\,\mathrm{d}y\Big|\\
&\le C2^{-\frac32 j}\operatorname*{ess-sup}_{x\in\operatorname{supp}\phi_{j,k},\,y\in\operatorname{supp}\phi_{i,l}}|g_x'(x,y)|\int_0^1|\phi_{i,l}(y)|\,\mathrm{d}y \qquad (7.10)\\
&\le C2^{-\frac32 j}\|g\|_{W^{1,\infty}(\mathcal{D}\times\mathcal{D})}\int_0^1|\phi_{i,l}(y)|\,\mathrm{d}y \qquad (7.11)\\
&\le C2^{-\frac32 j}\|g\|_{W^{1,\infty}(\mathcal{D}\times\mathcal{D})}|\operatorname{supp}\phi_{i,l}|^{\frac12}\|\phi_{i,l}\|_{L_2(\mathcal{D})} \le C2^{-\frac32 j-\frac12 i}.
\end{aligned}$$

By the symmetry of Λ^{-1}, we conclude

$$|\langle \Lambda^{-1}\phi_{i,l},\phi_{j,k}\rangle| \le C2^{-(i+j)}. \qquad (7.12)$$

This also holds for the Haar basis used in Subsection 7.1 because it has $m=\tilde{m}=1$.

Since $\|\Lambda^{\frac12}u\|_{L_2(\mathcal{D})}\le C\|u\|_{H^1(\mathcal{D})}$ for $u\in H_0^1(\mathcal{D})$, we have, by (7.1) with $s=1$,

$$|\langle \Lambda^{\frac12}\phi_{i,l},\Lambda^{\frac12}\phi_{j,k}\rangle| \le C2^{i+j} = C2^{\max(i,j)+\min(i,j)}. \qquad (7.13)$$

If $\Delta_{jkil} := \operatorname{supp}\phi_{j,k}\cap\operatorname{supp}\phi_{i,l} = \emptyset$ (cf. (7.3)), then the left hand side of (7.13) vanishes. This is because Λ is a local operator. More precisely, if $\Delta_{jkil}=\emptyset$, then

$$\langle \Lambda^{\frac12}\phi_{i,l},\Lambda^{\frac12}\phi_{j,k}\rangle = \langle a\phi_{i,l}',\phi_{j,k}'\rangle + \langle c\phi_{i,l},\phi_{j,k}\rangle = 0.$$

To finish the proof we use Theorem 5.2, (7.2), (7.9), (7.12), and $h=2^{-N}$, to get

$$\begin{aligned}
&\mathbf{E}\Big(\|W_A(t)-W_A^J(t)\|^2\Big) \le \sum_{i=N+1}^\infty\sum_{l\in J_i}\sum_{j=N+1}^\infty\sum_{k\in J_j}\langle\Lambda^{-1}\phi_{i,l},\phi_{j,k}\rangle\langle Q\tilde{\phi}_{i,l},\tilde{\phi}_{j,k}\rangle\\
&\le C\sum_{i=N+1}^\infty\sum_{j=N+1}^\infty 2^i2^j2^{-(i+j)}2^{-2(i+j)} = C\Big(\sum_{i=N+1}^\infty 2^{-2i}\Big)^2 \le C2^{-4N} = Ch^4.
\end{aligned}$$

Finally, by Theorem 5.1 with $\beta=2$, (7.5) (7.8), and (7.13), we get

$$\begin{aligned}
&\mathbf{E}\Big(\|W_A^J(t)-W_{A_h}^J(t)\|^2\Big) \le Ch^4\sum_{i=j_0}^N\sum_{l\in J_i}\sum_{j=j_0}^N\sum_{k\in J_j}\langle\Lambda^{\frac12}\phi_{i,l},\Lambda^{\frac12}\phi_{j,k}\rangle\langle Q\tilde{\phi}_{i,l},\tilde{\phi}_{j,k}\rangle\\
&\le Ch^4\sum_{i=j_0}^N\sum_{j=j_0}^N 2^{\max(i,j)}2^{\max(i,j)+\min(i,j)}2^{-\frac52\max(i,j)-\frac32\min(i,j)}\\
&= Ch^4\sum_{i=j_0}^N\sum_{j=j_0}^N 2^{-\frac12(\max(i,j)+\min(i,j))} = Ch^4\sum_{i=j_0}^N\sum_{j=j_0}^N 2^{-\frac12(i+j)}\\
&= Ch^4\Big(\sum_{i=j_0}^N 2^{-\frac12 i}\Big)^2 \le Ch^4.
\end{aligned}$$

This completes the proof. □

Remark 7.1. In applications the kernel q and its derivatives up to a certain degree often exhibit a decay; that is, $D^\alpha q(x,y) \to 0$ as $|x-y| \to \infty$ for $|\alpha| \leq M$. This decay can be taken into account when estimating (7.6) and instead of using the uniform estimate that leads to (7.7) one obtains additional decay for terms involving basis functions with disjoint supports based on the decay of the appropriate derivative of q. The same applies when estimating (7.10) by (7.11) in case the differential operator is of higher order with corresponding decay of its Green's function. This additional decay results in a lower truncation level N than the $N = -\log_2(h)$ required in Theorem 7.1 to balance the order of the truncation error and the discretization error. This is also the case when using a wavelet (dual wavelet) basis with higher order of cancelation and smoothness provided the noise is more smooth; that is, if $m > 2$ we may obtain a higher rate in (7.7).

Remark 7.2. If the noise is less smooth but still trace class, say $q \in W^{1,\infty}(\mathcal{D} \times \mathcal{D})$, then the convergence rate in Theorem 7.1 reduces to $O(h^2)$, but no smoothness of the wavelets is needed and lower order cancelation property suffices; that is, the simple Haar basis can be used. For example, the case $Q := \Lambda^{-1}$ is covered here, corresponding to an SPDE arising in path sampling problems for SDE's [16].

7.3. **Computational considerations.** The key to the approximation of the noise is the ability to simulate the truncated process $W^J(t)$, or for practical purposes, in the presence of time discretization, its increments $\Delta W^J(t) = W^J(t+\Delta t) - W^J(t)$. In order to do this, one needs to generate $\Delta \vec{W}^J(t)$, an $\mathbb{R}^{\#J}$-valued Gaussian random variable with covariance matrix $(Q_{\Delta \vec{W}^J}(t))_{jk} = \Delta t \langle Q\tilde{\phi}_j, \tilde{\phi}_k \rangle$. This can be achieved in the following way. First generate an $\mathbb{R}^{\#J}$-valued random variable $\vec{Z}$ with independent standard Gaussian components. Then compute the Cholesky factorization $Q_{\Delta \vec{W}^J}(t) = L_J(t) L_J(t)^*$ of the covariance matrix of $\Delta \vec{W}^J(t)$. Finally, $\Delta \vec{W}^J(t) = L_J(t)\vec{Z}$ (in distribution) and $\Delta W^J(t) = \sum_{j \in J} (\Delta \vec{W}^J(t))_j \phi_j$

In [2] it is shown that the Cholesky factorization can be obtained by successively updating an initial factorization by adding rows successively to the initial Cholesky factor. It is also shown there that adding one row and column results in roughly $\log^2(\#J)$ operations (as $\#J \to \infty$) when updating the Cholesky factorization for a nearly sparse matrix and that the Cholesky factor of such a matrix remains nearly sparse. This implies that the cost of computing the Cholesky factorization is $O(\#J \log^2(\#J))$ and the matrix vector multiplication with $L_J(t)$ can be achieved in $O(\#J \log(\#J))$ operations. If the kernel q of the covariance operator Q exhibits decay as discussed in Remark 7.1, then the matrix $Q_{\Delta \vec{W}^J(t)}$ will be nearly sparse (see [2] and [3]). Thus the above computational complexity applies. If one wants to refine the finite element mesh, then one needs to truncate the process on a higher level $J' \supset J$ in order to preserve the order of the finite element method according to Theorem 7.1. However, since the approximation of the process is independent of the finite element method and it is expanded in a hierarchical basis, there is no need to simulate the process, or its increments, from scratch. The new covariance matrix $Q_{\Delta \vec{W}^{J'}}(t)$ is obtained from $Q_{\Delta \vec{W}^J}(t)$ by adding $\#J' - \#J$ new rows and columns. If one stores the initial random variable $\vec{Z}$ one just generates $\#J' - \#J$ additional independent standard Gaussian random variables, updates the Cholesky factor $L_J(t)$ to $L_{J'}(t)$ by computing $\#J' - \#J$ new rows and then updates the $\Delta W^J(t)$ to $\Delta W^{J'}(t)$ by computing the last $\#J' - \#J$ components.

Remark 7.3. Since the decay in the error estimate in Theorem 7.1 comes mainly from the smoothness (and decay) of q it might be worthwhile to interchange the roles of the primal and dual basis. Usually the primal basis is easier to work with and if the dual basis does not have some of the desired properties (cancelation, small support), then the loss in the error estimate can be compensated, if the kernel q is smooth (or decays) enough, by using the good properties of the primal basis. This will also make the computation of the elements of the relevant matrix $(Q_{\Delta\vec{W}^J}(t))_{jk} = \Delta t\langle Q\phi_j, \phi_k\rangle$ simpler.

Remark 7.4. The finite dimensional process $P_hW(t)$ could be simulated directly via a finite eigenfunction expansion which is very expensive in general as it requires the diagonalization of P_hQP_h in V_h. The other main drawback of this approach in contrast to the biorthogonal wavelet expansion, is that it does not allow updates. That is, when using a different mesh, the eigenvalue computation has to be done from scratch, while for the wavelet expansion of W the existing computations can be updated. Nevertheless, this direct approach is quite feasible in the case of stationary kernels analytic at 0, such as the Gauss kernel, and the orthogonal expansion of $P_hW(t)$ can be truncated severely without losing the asymptotic order of the finite element method [19].

References

[1] E. J. Allen, S. J. Novosel, and Z. Zhang, *Finite element and difference approximation of some linear stochastic partial differential equations*, Stochastics Stochastics Rep. **64** (1998), 117–142.

[2] N. N. Bennett and A. Malinverno, *Fast model updates using wavelets*, Multiscale Model. Simul. **3** (2004/05), 106–130 (electronic).

[3] G. Beylkin, R. Coifman, and V. Rokhlin, *Fast wavelet transforms and numerical algorithms. I*, Comm. Pure Appl. Math. **44** (1991), 141–183.

[4] A. Cohen, I. Daubechies, and J.-C. Feauveau, *Biorthogonal bases of compactly supported wavelets*, Comm. Pure Appl. Math. **45** (1992), 485–560.

[5] G. Da Prato and A. Debussche, *Stochastic Cahn-Hilliard equation*, Nonlinear Anal. **26** (1996), 241–263.

[6] G. Da Prato and J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Encyclopedia of Mathematics and its Applications, vol. 44, Cambridge University Press, Cambridge, 1992.

[7] W. Dahmen, *Wavelet methods for PDEs—some recent developments*, J. Comput. Appl. Math. **128** (2001), 133–185, Numerical analysis 2000, Vol. VII, Partial differential equations.

[8] W. Dahmen, A. Kunoth, and K. Urban, *Biorthogonal spline wavelets on the interval—stability and moment conditions*, Appl. Comput. Harmon. Anal. **6** (1999), 132–196.

[9] I. Daubechies, *Ten Lectures on Wavelets*, CBMS-NSF Regional Conference Series in Applied Mathematics, vol. 61, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1992.

[10] A. M. Davie and J. G. Gaines, *Convergence of numerical schemes for the solution of parabolic stochastic partial differential equations*, Math. Comp. **70** (2001), 121–134 (electronic).

[11] Q. Du and T. Zhang, *Numerical approximation of some linear stochastic partial differential equations driven by special additive noises*, SIAM J. Numer. Anal. **40** (2002), 1421–1445 (electronic).

[12] W. Grecksch and P. E. Kloeden, *Time-discretised Galerkin approximations of parabolic stochastic PDEs*, Bull. Austral. Math. Soc. **54** (1996), 79–85.

[13] I. Gyöngy, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. I*, Potential Anal. **9** (1998), 1–25.

[14] ———, *Lattice approximations for stochastic quasi-linear parabolic partial differential equations driven by space-time white noise. II*, Potential Anal. **11** (1999), 1–37.

[15] I. Gyöngy and D. Nualart, *Implicit scheme for quasi-linear parabolic partial differential equations perturbed by space-time white noise*, Stochastic Process. Appl. **58** (1995), 57–72.

[16] M. Hairer, A. M. Stuart, J. Voss, and P. Wiberg, *Analysis of SPDEs arising in path sampling. I. The Gaussian case*, Commun. Math. Sci. **3** (2005), 587–603.
[17] E. Hausenblas, *Numerical analysis of semilinear stochastic evolution equations in Banach spaces*, J. Comput. Appl. Math. **147** (2002), 485–516.
[18] ______, *Approximation for semilinear stochastic evolution equations*, Potential Anal. **18** (2003), 141–186.
[19] 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*, Numer. Algorithms **53** (2010), 309–320.
[20] M. Kovács, S. Larsson, and A. Mesforush, *Finite element approximation of the Cahn-Hilliard-Cook equation*, preprint 2010:18, Department of Mathematical Sciences, Chalmers University of Technology.
[21] M. Kovács, S. Larsson, and F. Saedpanah, *Finite element approximation of the linear stochastic wave equation with additive noise*, SIAM J. Numer. Anal. **48** (2010), 408–427.
[22] S. Larsson and A. Mesforush, *Finite element approximation of the linearized Cahn-Hilliard-Cook equation*, preprint 2009:20, Department of Mathematical Sciences, Chalmers University of Technology.
[23] T. Müller-Gronbach and K. Ritter, *Lower bounds and nonuniform time discretization for approximation of stochastic heat equations*, Found. Comput. Math. **7** (2007), 135–181.
[24] C. Prévôt and M. Röckner, *A Concise Course on Stochastic Partial Differential Equations*, Lecture Notes in Mathematics, vol. 1905, Springer, Berlin, 2007.
[25] L. Quer-Sardanyons and M. Sanz-Solé, *Space semi-discretisations for a stochastic wave equation*, Potential Anal. **24** (2006), 303–332.
[26] T. Shardlow, *Numerical methods for stochastic parabolic PDEs*, Numer. Funct. Anal. Optim. **20** (1999), 121–145.
[27] V. Thomée, *Galerkin Finite Element Methods for Parabolic Problems*, second ed., Springer Series in Computational Mathematics, vol. 25, Springer-Verlag, Berlin, 2006.
[28] J. B. Walsh, *Finite element methods for parabolic stochastic PDE's*, Potential Anal. **23** (2005), 1–43.
[29] ______, *On numerical solutions of the stochastic wave equation*, Illinois J. Math. **50** (2006), 991–1018 (electronic).
[30] Y. Yan, *Semidiscrete Galerkin approximation for a linear stochastic parabolic partial differential equation driven by an additive noise*, BIT **44** (2004), 829–847.
[31] ______, *Galerkin finite element methods for stochastic parabolic partial differential equations*, SIAM J. Numer. Anal. **43** (2005), 1363–1384 (electronic).

Department of Mathematics and Statistics, University of Otago, P.O. Box 56, Dunedin, New Zealand
E-mail address: `mkovacs@maths.otago.ac.nz`

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: `fredlind@chalmers.se`

Department of Mathematical Sciences, Chalmers University of Technology and University of Gothenburg, SE–412 96 Göteborg, Sweden
E-mail address: `stig@chalmers.se`