The Effect of Finite Element Discretisation on the Stationary Distribution of SPDEs

Jochen Voss

17th February 2012

Abstract

This article studies the effect of discretisation error on the stationary distribution of stochastic partial differential equations (SPDEs). We restrict the analysis to the effect of space discretisation, performed by finite element schemes. The main result is that under appropriate assumptions the stationary distribution of the finite element discretisation converges in total variation norm to the stationary distribution of the full SPDE.

SPDEs, finite element discretisation, stationary distribution

subject classifications. 60H35, 60H15, 65C30

Introduction

In this article we consider the finite element discretisation for stochastic partial differential equations (SPDEs) of the form

$$\partial_t u(t,x) = \partial_x^2 u(t,x) + f\big(u(t,x)\big) + \sqrt{2}\,\partial_t w(t,x) \qquad \forall (t,x) \in [0,\infty) \times [0,1], \tag{1}$$

where $\partial_t w$ is space-time white noise and $f\colon \mathbb{R} \to \mathbb{R}$ is a smooth function with bounded derivatives, and the differential operator ∂_x^2 is equipped with boundary conditions such that it is a negative operator on the space $\mathrm{L}^2\big([0,1],\mathbb{R}\big)$. More specifically, we are considering the effect that discretisation of the SPDE has on its stationary distribution.

Our motivation for studying this problem lies in a recently proposed, SPDE-based sampling technique: when trying to sample from a distribution on path-space, *e.g.* in filtering/smoothing problems to sample from the conditional distribution of a process given some observations, one can do so using a Markov chain Monte Carlo approach. Such MCMC methods requires a process with values in path-space and it transpires that in some situations SPDEs of the form (1) can be used, see *e.g.* Hairer et al. (2005, 2007) and Hairer et al. (2009) for a review. When implementing the resulting methods on a computer, the sampling SPDEs must be discretised and, because MCMC methods use the sampling process only as a source of samples from its stationary distribution, the effect of the discretisation error on an MCMC method depends on how well the stationary distribution of the SPDE is approximated. While there are many results of approximation of trajectories of SPDEs (*e.g.* Millet and Morien, 2005; Walsh, 2005; Hausenblas, 2008; Gyöngy and Millet, 2009; Jentzen, 2011), approximation of the stationary distribution seems not to be well-studied so far.

When discretising an SPDE, discretisation of space and time can be considered to be two independent problems. In cases where only the stationary distribution of the process is of interest, Metropolis sampling, using the next time step of the time discretisation as a proposal, can be used to completely eliminate the error introduced by time discretisation (Beskos et al., 2008). For this reason, in this article we restrict the analysis to the effect of space discretisation alone. The discretisation technique discussed here is a finite element discretisation, which is a much-studied technique for deterministic PDEs. The approximation problem for *stochastic* PDEs, as studied in this article, differs from the deterministic case significantly, since here we have to compare the full *distribution* of the solutions instead of considering the approximation of the solution as a function.

Finally, the error of the resulting sampling method is not only affected by the error in the stationary distribution, but also by the time it takes for the sampling equation to reach equilibrium. Here, we concentrate on the error in the equilibrium distribution itself and refer to Bou-Rabee and Hairer (2012) for discussion of the speed of convergence to equilibrium.

While the results of this article are formulated for SPDEs with values in $\mathbb{R}$, we expect the results and techniques to carry over to SPDEs with values in $\mathbb{R}^d, d > 1$ without significant changes. We only restrict discussion to the one-dimensional case to ease notation. This is in contrast to the domain of the SPDEs: we consider the case of one spacial dimension because this is the relevant case for the sampling techniques discussed above, but this choice significantly affects the proofs and a different approach would likely be required to study the case of higher-dimensional spatial domains.

The text is structured as follows: In section 1 present the required results to characterise the stationary distribution of the SPDE (1). In section 2 we introduce the finite element discretisation scheme for (1) and identify the stationary distribution of the discretised equation. Building on these results, in section 3, we state our main result about convergence of the discretised stationary distributions to the full stationary distribution. Finally, in section 4, we give two examples in order to illustrate the link to the MCMC methods discussed above and also to demonstrate that the considered finite element discretisation forms a concrete and easily implemented numerical scheme.

1 The Infinite-Dimensional Equation

In order to study the SPDE (1), it is convenient to rewrite the equation as an evolution equation on the Hilbert space $\mathcal{H} = \mathrm{L}^2\bigl([0,1], \mathbb{R}^d\bigr)$; For a description of the underlying theory we refer, for example, to the monograph of Da Prato and Zabczyk (1992). We consider

$$du(t) = \mathcal{L}u(t)\,dt + f\bigl(u(t)\bigr) + \sqrt{2}\,dw(t) \qquad \forall t \geq 0 \tag{2}$$

where the solution u takes values in $\mathcal{H}$ and f acts pointwise on u, *i.e.* $f(u)(x) = \tilde{f}\bigl(u(x)\bigr)$ for almost all $x \in [0,1]$ for some function $\tilde{f} \colon \mathbb{R}^d \to \mathbb{R}^d$, such that f maps $\mathcal{H}$ into itself. Furthermore, w is an L^2-cylindrical Wiener process and we equip the linear operator $\mathcal{L} = \partial_x^2$ with boundary conditions given by the domain

$$\mathcal{D}(\mathcal{L}) = \bigl\{ u \in \mathrm{H}^2([0,1], \mathbb{R}) \bigm| \alpha_0 u(0) - \beta_0 \partial_x u(0) = 0,\ \alpha_1 u(1) + \beta_1 \partial_x u(1) = 0 \bigr\} \tag{3}$$

where $\alpha_0, \alpha_1, \beta_0, \beta_1 \in \mathbb{R}$. The boundary conditions in (3) include the cases of Dirichlet ($\beta_i = 0$) and v. Neumann ($\alpha_i = 0$) boundary conditions. The general case of $\alpha_i, \beta_i \neq 0$ is known as Robin boundary conditions.

We start our analysis by considering the linear equation

$$du(t) = \mathcal{L}u(t)\,dt + \sqrt{2}\,dw(t) \qquad \forall t \geq 0. \tag{4}$$

For equation (4) to have a stationary distribution, we require $\mathcal{L}$ to be negative definite. The following lemma states necessary and sufficient conditions on α_i and β_i for this to be the case.

Lemma 1.1. *The operator $\mathcal{L}$ is a self-adjoint operator on the Hilbert space $\mathcal{H}$. The operator $\mathcal{L}$ is negative definite, if and only if $\alpha_0, \beta_0, \alpha_1, \beta_1$ are contained in the set*

$$\begin{aligned} A = \Bigl\{ &\beta_0(\alpha_0 + \beta_0) > 0, \beta_1(\alpha_1 + \beta_1) > 0, \bigl|(\alpha_0 + \beta_0)(\alpha_1 + \beta_1)\bigr| > |\beta_0 \beta_1| \Bigr\} \\ &\cup \Bigl\{ \beta_0 = 0, \alpha_0 \neq 0,\ \beta_1(\alpha_1 + \beta_1) > 0 \Bigr\} \cup \Bigl\{ \beta_0(\alpha_0 + \beta_0) > 0,\ \beta_1 = 0, \alpha_1 \neq 0 \Bigr\} \\ &\cup \Bigl\{ \beta_0 = 0, \alpha_0 \neq 0,\ \beta_1 = 0, \alpha_1 \neq 0 \Bigr\}. \end{aligned}$$

proof. From the definition of $\mathcal{L}$ it is easy to see that the operator is self-adjoint.

We have to show that $\mathcal{L}$ is negative if and only if $(\alpha_0, \beta_0, \alpha_1, \beta_1) \in A$. Without loss of generality we can assume $\beta_i \geq 0$ for $i = 1, 2$ and $\alpha_i \geq 0$ whenever $\beta_i = 0$ (since we can replace (α_i, β_i) by $(-\alpha_i, -\beta_i)$ if required). Assume that λ is an eigenvalue of $\mathcal{L}$. If $\lambda > 0$, the corresponding eigenfunctions are of the form

$$u(x) = c_1 \mathrm{e}^{\sqrt{\lambda} x} + c_2 \mathrm{e}^{-\sqrt{\lambda} x}$$

where c_1 and c_2 are given by the boundary conditions: For u to be in the domain of $\mathcal{L}$, the coefficients c_1 and c_2 need to satisfy

$$\begin{pmatrix} \alpha_0 - \beta_0\sqrt{\lambda} & \alpha_0 + \beta_0\sqrt{\lambda} \\ (\alpha_1 + \beta_1\sqrt{\lambda})\mathrm{e}^{\sqrt{\lambda}} & (\alpha_1 - \beta_1\sqrt{\lambda})\mathrm{e}^{-\sqrt{\lambda}} \end{pmatrix} \begin{pmatrix} c_1 \\ c_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

Non-trivial solutions exist only if the matrix is singular or, equivalently, if its determinant

$$f(\lambda) = \alpha_0\alpha_1 + \big(\alpha_0\beta_1 + \alpha_1\beta_0\big)\sqrt{\lambda}\coth\sqrt{\lambda} + \beta_0\beta_1\lambda \tag{5}$$

satisfies $f(\lambda) = 0$. For $\lambda = 0$ the eigenfunctions are of the form $u(x) = c_1 1 + c_2 x$ and an argument similar to the one above shows that the boundary conditions can be satisfied if and only if $\alpha_0\alpha_1 + \alpha_0\beta_1 + \alpha_1\beta_0 = 0$. Since $x\coth(x) \to 1$ as $x \to 0$, this condition can be written as $f(0) = 0$ where

$$f(0) = \lim_{\lambda \downarrow 0} f(\lambda) = \alpha_0\alpha_1 + \alpha_0\beta_1 + \alpha_1\beta_0.$$

This shows that $\mathcal{L}$ is negative whenever $f(\lambda) \neq 0$ for all $\lambda \geq 0$.

Let $(\alpha_0, \beta_0, \alpha_1, \beta_1) \in A$. Assume first $\beta_i \neq 0$ for $i = 1, 2$ and let $\xi_i = \alpha_i/\beta_i$. Then, by the first condition in A, we have $\xi_0, \xi_1 > -1$ and $(\xi_0 - 1)(\xi_1 - 1) > 1$, and for $\lambda \geq 0$ we get

$$f(\lambda) = \xi_0\xi_1 + \big(\xi_0 + \xi_1\big)\sqrt{\lambda}\coth\sqrt{\lambda} + \lambda \geq (\xi_0 + 1)(\xi_1 + 1) - 1 > 0.$$

The cases $\beta_0 = 0$ or $\beta_1 = 0$ can be treated similarly. Thus, for $(\alpha_0, \beta_0, \alpha_1, \beta_1) \in A$, there are no eigenvalues with $\lambda \geq 0$ and the operator is negative.

For the converse statement, assume that $(\alpha_0, \beta_0, \alpha_1, \beta_1) \notin A$. We then have to show that there is a $\lambda \geq 0$ with $f(\lambda) = 0$. Assume first $\beta_i > 0$ for $i = 1, 2$ and define ξ_i as above. If $(\xi_0 + 1)(\xi_1 + 1) = 1$, we have $f(0) = 0$. If $(\xi_0 + 1)(\xi_1 + 1) < 1$, the function f satisfies $f(0) < 0$ and $f(\lambda) \to \infty$ as $\lambda \to \infty$; by continuity there is a $\lambda > 0$ with $f(\lambda) = 0$. Finally, if $(\xi_0 + 1)(\xi_1 + 1) > 1$ but $\xi_0, \xi_1 \leq -1$, we have $f(0) > 0$ and for λ with $\sqrt{\lambda} = -(\xi_0 + \xi_1)/2 > 0$ we find $f(\lambda) < \xi_0\xi_1 + \big(\xi_0 + \xi_1\big)\sqrt{\lambda} + \lambda = \xi_0\xi_1 - (\xi_0 + \xi_1)^2/4 = -(\xi_0 - \xi_1)^2/4 \leq 0$ and by continuity there is a $\lambda > 0$ with $f(\lambda) = 0$. Again, the cases $\beta_0 = 0$ and $\beta_1 = 0$ can be treated similarly. ∎

A representation of the eigenvalues of $\mathcal{L}$ which is similar to the one in the proof of lemma 1.1 can be found in section 3 of Cacciapuoti and Finco (2010).

The statement from lemma 1.1 reproduces the well-known results that the Laplacian with Dirichlet boundary conditions ($\alpha_i = 1, \beta_i = 0$) is negative definite whereas the Laplacian with von Neumann boundary conditions ($\alpha_i = 0, \beta_i = 1$) is not (since constants are eigenfunctions with eigenvalue 0).

Lemma 1.2. Let $\mathcal{L}$ be negative definite. Then the following statements hold:

1. The linear SPDE (4) has global, continuous $\mathcal{H}$-valued solutions.
2. Equation (4) has a unique stationary distribution ν on $\mathcal{H}$. The measure ν is Gaussian with mean 0 and covariance function

$$C(x, y) = \frac{\beta_0\beta_1 + \alpha_0\beta_1\, xy + \beta_0\alpha_1(1 - x)(1 - y)}{\alpha_0\alpha_1 + \alpha_0\beta_1 + \beta_0\alpha_1} + x \wedge y - xy \tag{6}$$

where $x \wedge y$ denotes the minimum of x and y.

3. The measure ν coincides with the distribution of $U \in \mathrm{C}\big([0,1],\mathbb{R}\big)$ given by

$$U(x) = (1-x)L + xR + B(x) \qquad \forall x \in [0,1],$$

where $L \sim \mathcal{N}(0,\sigma_L^2)$, $R \sim \mathcal{N}(0,\sigma_R^2)$ with $\mathrm{Cov}(L,R) = \sigma_{LR}$, the process B is a Brownian bridge, independent of L and R, and

$$\sigma_L^2 = \frac{\beta_0(\alpha_1+\beta_1)}{\alpha_0\alpha_1+\alpha_0\beta_1+\beta_0\alpha_1}, \quad \sigma_R^2 = \frac{(\alpha_0+\beta_0)\beta_1}{\alpha_0\alpha_1+\alpha_0\beta_1+\beta_0\alpha_1},$$
$$\sigma_{LR} = \frac{\beta_0\beta_1}{\alpha_0\alpha_1+\alpha_0\beta_1+\beta_0\alpha_1}.$$

proof. From Iscoe et al. (1990) and Da Prato and Zabczyk (1996) (see Hairer et al., 2005, Lemma 2.2) we know that (4) has global, continuous $\mathcal{H}$-valued solutions as well as a unique stationary distribution given by $\nu = \mathcal{N}(0, -\mathcal{L}^{-1})$. An easy computation shows that C as given in equation (6) is a Green's function for the operator $-\mathcal{L}$, *i.e.* $-\partial_x^2 C(x,y) = \delta(x-y)$ and for every $y \in (0,1)$ the function $x \mapsto C(x,y)$ satisfies the boundary conditions (3). This completes the proof of the first two statements.

For the third statement we note that U is centred Gaussian with covariance function

$$\begin{aligned} C(x,y) &= \mathrm{Cov}\big(U(x),U(y)\big) \\ &= \mathrm{Cov}\big((1-x)L+xR+B(x),(1-y)L+yR+B(y)\big) \\ &= (1-x)(1-y)\,\sigma_L^2 + \big((1-x)y + x(1-y)\big)\,\sigma_{LR} + xy\,\sigma_R^2 \\ &\qquad + x\wedge y - xy. \end{aligned}$$

The fact that this covariance function can be written in the form (6) can be checked by a direct calculation. ∎

Using the results for the linear SPDE (4) we can now study the full SPDE (2). The result is given in the following lemma.

Lemma 1.3. Let $\mathcal{L}$ be negative definite. Furthermore, let $f = F'$ where $F \in \mathrm{C}^2(\mathbb{R},\mathbb{R})$ is bounded from above with bounded second derivative. Then the following statements hold:

1. The nonlinear SPDE (2) has global, continuous $\mathcal{H}$-valued solutions.
2. Equation (2) has a unique stationary distribution μ which is given by

$$\frac{d\mu}{d\nu}(u) = \frac{1}{Z}\exp\Big(\int_0^1 F\big(u(x)\big)\,dx\Big)$$

where ν is the stationary distribution of (4) from lemma 1.2 and Z is the normalisation constant.

proof. This result is well known, see *e.g.* Zabczyk (1989) or Hairer et al. (2007, corollary 4.5). ∎

2 Finite Element Approximation

In this section we consider finite dimensional approximations of the SPDE (2), obtained by discretising space using the finite element method. The approximation follows the same approach as for deterministic PDEs. For background on the deterministic case we refer to Brenner and Scott (2002) or Johnson (1990).

To discretise space, let $n \in \mathbb{N}$, $\Delta x = 1/n$ and consider x-values on the grid $k\Delta x$ for $k \in \mathbb{N}$. Since the differential operator $\mathcal{L}$ in (2) is a second order differential operator, we can choose a finite element basis consisting of "hat functions" φ_i for $i \in \mathbb{Z}$ which have $\varphi_i(i\,\Delta x) = 1$, $\varphi_i(j\,\Delta x) = 0$ for all $j \neq i$, and which are affine between the grid points. Formally, the weak (in the PDE-sense) formulation of SPDE (2) can be written as

$$\langle v, du(t)\rangle = B(v,u)\,dt + \langle v, f\big(u(t)\big)\rangle + \sqrt{2}\langle v, dw(t)\rangle$$

where $\langle \cdot, \cdot \rangle$ denotes the L^2-inner product and the bilinear form B is given by

$$B(u,v) = \langle v, \mathcal{L}u \rangle = u(1)v'(1) - u(0)v'(0) - \int_0^1 u'(x)v'(x)\,dx.$$

The discretised solution is found by taking u and v to be in the space spanned by the functions φ_i, *i.e.* by using the ansatz

$$u(t) = \sum_j U_j(t)\varphi_j$$

and then considering the following system of equations:

$$\langle \varphi_i, \sum_j dU_j\varphi_j \rangle = \langle \varphi_i, \partial_x^2 \sum_j U_j\varphi_j \rangle\,dt + \langle \varphi_i, f\big(\sum_j U_j\varphi_j\big) \rangle\,dt + \sqrt{2}\langle \varphi_i, dw \rangle. \tag{7}$$

The domain V of the bilinear form B depends on the boundary conditions of $\mathcal{L}$; there are four different cases:

1. If $\beta_0, \beta_1 \neq 0$ in (3), *i.e.* for von Neumann or Robin boundary conditions, we have $V = H^1\big([0,1],\mathbb{R}\big)$ and we consider the basis functions φ_i for $i \in I = \{0, 1, \ldots, n-1, n\}$.
2. If $\beta_0 = 0$ and $\beta_1 \neq 0$, *i.e.* for a Dirichlet boundary condition at the left boundary, we have $V = \big\{u \in H^1\big([0,1],\mathbb{R}\big) \bigm| u(0) = 0\big\}$ and we consider the basis functions φ_i for $i \in I = \{1, \ldots, n-1, n\}$.
3. If $\beta_0 \neq 0$ and $\beta_1 = 0$, *i.e.* for a Dirichlet boundary condition at the right boundary, we have $V = \big\{u \in H^1\big([0,1],\mathbb{R}\big) \bigm| u(1) = 0\big\}$ and we consider the basis functions φ_i for $i \in I = \{0, 1, \ldots, n-1\}$.
4. If $\beta_0, \beta_1 = 0$, *i.e.* for Dirichlet boundary conditions at both boundaries, we have $V = \big\{u \in H^1\big([0,1],\mathbb{R}\big) \bigm| u(0) = u(1) = 0\big\}$ and we consider the basis functions φ_i for $i \in I = \{1, \ldots, n-1\}$.

Throughout the rest of the text we will write I for the index set of the finite element discretisation as above, and the discretised solution $u = \sum_{j\in I} U_j\varphi_j$ will be described by the coefficient vector $U \in \mathbb{R}^I$. In all cases we define the "stiffness matrix" $L^{(n)} \in \mathbb{R}^{I\times I}$ by $L^{(n)}_{ij} = B(\varphi_i, \varphi_j)$ for all $i, j \in I$. For the given basis functions we get

$$L^{(n)}_{ij} = \begin{cases} -\frac{2}{\Delta x} & \text{if } i = j \notin \{0, n\}, \\ +\frac{1}{\Delta x} & \text{if } i \in \{j-1, j+1\}, \\ -\frac{1}{\Delta x} - \frac{\alpha_0}{\beta_0} & \text{if } i = j = 0, \\ -\frac{1}{\Delta x} - \frac{\alpha_1}{\beta_1} & \text{if } i = j = n, \\ 0 & \text{else,} \end{cases}$$

where the cases $i = j = 0$ and $i = j = n$ cannot occur for Dirichlet boundary conditions. The "mass matrix" $M \in \mathbb{R}^{I\times I}$ is defined by $M_{ij} = \langle \varphi_i, \varphi_j \rangle$ and for $i, j \in I$ we get

$$M_{ij} = \begin{cases} \frac{4}{6}\Delta x & \text{if } i = j \notin \{0, n\}, \\ \frac{1}{6}\Delta x & \text{if } i \in \{j-1, j+1\}, \\ \frac{2}{6}\Delta x & \text{if } i = j \in \{0, n\}, \\ 0 & \text{else,} \end{cases}$$

where, again, the cases $i = j = 0$ and $i = j = n$ don't occur for Dirichlet boundary conditions. We note that the matrix $L^{(n)}$ only has the prefactor $1/\Delta x$ instead of the $1/\Delta x^2$ one would expect for a second derivative. The "missing" Δx appears in the matrix M.

Since

$$\mathrm{Cov}(\langle \varphi_i, w \rangle, \langle \varphi_j, w \rangle) = \langle \varphi_i, \varphi_j \rangle = M_{ij},$$

equation (7) can be written as

$$M\,dU_t = L^{(n)}U_t\,dt + f_n(U_t)\,dt + \sqrt{2}M^{1/2}dW_t$$

where $f_n\colon \mathbb{R}^I \to \mathbb{R}^I$ is defined by

$$f_n(u)_i = \big\langle \varphi_i, f\big(\sum_{j\in I} u_j\varphi_j\big)\big\rangle \tag{8}$$

for all $u \in \mathbb{R}^I$ and $i \in I$. Multiplication with M^{-1} then yields the following SDE describing the evolution of the coefficients $(U_i)_{i\in I}$:

Definition 2.1. The *finite element discretisation* of SPDE (2) is given by

$$dU_t = M^{-1}L^{(n)}U_t\,dt + M^{-1}f_n(U_t)\,dt + \sqrt{2}M^{-1/2}\,dW_t \tag{9}$$

where W is an $|I|$-dimensional standard Brownian motion, $I \subseteq \{0, 1, \ldots, n\}$ is the index set of the finite element discretisation, and $L^{(n)}$ and M are as above.

Our aim is to show that the stationary distribution of (9) converges to the stationary distribution of the SPDE (2). We start our analysis by considering the linear case $f \equiv 0$. For this case the finite element discretisation simplifies to (10) below.

Lemma 2.2. Let $\mathcal{L}$ be negative definite. Then $\nu_n = \mathcal{N}\big(0, (-L^{(n)})^{-1}\big)$ is the unique stationary distribution of

$$dU_t = M^{-1}L^{(n)}U_t\,dt + \sqrt{2}M^{-1/2}\,dW_t. \tag{10}$$

proof. Since $\mathcal{L}$ is a negative operator, the matrix $L^{(n)}$ is a symmetric, negative definite matrix. As the product of a positive definite symmetric matrix and a negative definite symmetric matrix, $M^{-1}L^{(n)}$ is negative definite; its eigenvalues coincide with the eigenvalues of

$$M^{-1/2}L^{(n)}M^{-1/2} = -\big((-L^{(n)})^{1/2}M^{-1/2}\big)^\top \big((-L^{(n)})^{1/2}M^{-1/2}\big).$$

From Arnold (1974, theorem 8.2.12) we know that then the unique stationary distribution of the SDE (10) is $\mathcal{N}(0, C^{(n)})$ where $C^{(n)}$ solves the Lyapunov equation

$$M^{-1}L^{(n)}C^{(n)} + C^{(n)}L^{(n)}M^{-1} = -2M^{-1}.$$

By theorem 5.2.2 of Lancaster and Rodman (1995), this system of linear equations has a unique solution and it is easily verified that this solution is given by $C^{(n)} = (-L^{(n)})^{-1}$. ∎

The following lemma shows that for $f \equiv 0$ there is no discretisation error at all: the stationary distributions of the SPDE (4), projected to $\mathbb{R}^I$, and of the finite element discretisation (9) coincide.

Lemma 2.3. Define $\Pi\colon \mathrm{C}\big([0,1],\mathbb{R}\big) \to \mathbb{R}^I$ by

$$(\Pi u)_i = u(i\Delta x) \qquad \forall i \in I. \tag{11}$$

Let ν be the stationary distribution of the linear SPDE (4) on $\mathrm{C}\big([0,1],\mathbb{R}\big)$ and let ν_n be the stationary distribution of the linear finite element discretisation (10). Then we have

$$\nu_n = \nu \circ \Pi^{-1}$$

for every $n \in \mathbb{N}$.

proof. Let C^{exact} be the covariance matrix of $\nu \circ \Pi^{-1}$ and let $C^{(n)}$ be the covariance matrix of ν_n. Since both measures under consideration are centred Gaussian, if suffices to show $C^{\text{exact}} = C^{(n)}$. By lemma 1.2, the matrix C^{exact} satisfies

$$C^{\text{exact}}_{i,j} = C(i\Delta x, j\Delta x) \qquad \forall i, j \in I$$

where C is given by equation (6). By lemma 2.2 we have $C^{(n)} = (-L^{(n)})^{-1}$. A simple calculation, using the fact that both C^{exact} and $L^{(n)}$ are known explicitly, shows $C^{\text{exact}} L^{(n)} = -I$ and thus $C^{\text{exact}} = C^{(n)}$ (the four different cases for the boundary conditions need to be checked separately). This completes the proof. ∎

The preceding results only consider the linear case and for the general case, in the presence of the non-linearity f, we can of course no longer expect a similar result to hold. As a starting point for analysing this case, we reproduce a well-known result which allows to identify the stationary distribution of the discretised finite element equation.

Lemma 2.4. Let $F \in \mathrm{C}^2(\mathbb{R}^d, \mathbb{R})$ with bounded second derivatives and satisfying the condition $Z = \int_{\mathbb{R}^d} \mathrm{e}^{2F(x)}\, dx < \infty$. Furthermore, let $A \in \mathbb{R}^{d\times d}$ be invertible. Then the SDE

$$dX_t = AA^\top \nabla F(X_t)\, dt + A\, dW_t \tag{12}$$

has a unique stationary distribution which has density

$$\varphi(x) = \frac{1}{Z} \mathrm{e}^{2F(x)}$$

with respect to the Lebesgue measure on $\mathbb{R}^d$.

proof. Define $G(y) = F(Ay)$ for all $y \in \mathbb{R}^d$. By the assumptions on F we have $G \in \mathrm{C}^2(\mathbb{R}^d, \mathbb{R})$ with bounded second derivatives and $Z_G = \int_{\mathbb{R}^d} \mathrm{e}^{2G(y)}\, dy < \infty$. Therefore, the SDE

$$dY_t = \nabla G(Y_t)\, dt + dW_t$$

has a unique stationary distribution with density

$$\psi(y) = \frac{1}{Z_G} \mathrm{e}^{2G(y)}.$$

Since $\nabla G(y) = A^\top \nabla F(Ay)$, we have

$$dY_t = A^\top \nabla F(AY_t)\, dt + dW_t$$

and multiplying this equation by A gives

$$d(AY_t) = AA^\top \nabla F(AY_t)\, dt + A\, dW_t.$$

Consequently, $X_t = AY_t$ satisfies the SDE (12) and has a unique stationary distribution with density proportional to $\psi(A^{-1}x) \propto \mathrm{e}^{2G(A^{-1}x)} = \mathrm{e}^{2F(x)}$. Since this function, up to a multiplicative constant, coincides with φ, the process X has stationary density φ. ∎

Because the stationary distribution in the lemma does not depend on A, the stationary distribution of (12) does not change when we remove/add A from the equation. The process of introducing the matrix A is sometimes called "preconditioning the SDE".

In cases where we are only interested in the stationary distribution of a discretised SPDE, the argument from lemma 2.4 allows us to omit the mass matrix M from the finite element SDE (9). In particular we don't need to consider the potentially computationally expensive square root $M^{1/2}$ in numerical simulations.

Lemma 2.5. Let $\mathcal{L}$ be negative definite. Furthermore, let $f = F'$ where $F \in \mathrm{C}^2(\mathbb{R}, \mathbb{R})$ is bounded from above with bounded second derivative. Then the finite element SDE (9) has a unique stationary distribution μ_n given by

$$\frac{d\mu_n}{d\nu_n} = \frac{1}{Z_n} \exp\bigl(F_n\bigr) \tag{13}$$

where

$$F_n(u) = \int_0^1 F\Bigl(\sum_{j\in I} u_j \varphi_j(x)\Bigr)\, dx \qquad \forall u \in \mathbb{R}^I,$$

Z_n is the normalisation constant and ν_n is the stationary distribution of the linear equation from lemma 2.2.

proof. Let $\Phi(u) = \frac{1}{2}u^\top L^{(n)}u + F_n(u)$ for all $u \in \mathbb{R}^I$. Then

$$\partial_i \Phi(u) = (L^{(n)}u)_i + \big\langle \varphi_i, F'\big(\sum_{j\in I} u_j\varphi_j\big)\big\rangle = \big(L^{(n)}u + f_n(u)\big)_i$$

for all $i \in I$ and thus (9) can be written as

$$dU_t = M^{-1}\nabla\Phi(U_t)\,dt + \sqrt{2}M^{-1/2}\,dW_t.$$

By lemma 2.4, this SDE has a unique stationary distribution μ_n whose density w.r.t. the $|I|$-dimensional Lebesgue measure λ is given by

$$\frac{d\mu_n}{d\lambda}(u) = \frac{1}{\tilde{Z}_n}\mathrm{e}^{\Phi(u)} = \frac{1}{\tilde{Z}_n}\exp\Big(-\frac{1}{2}u^\top(-L^{(n)})u + F_n(u)\Big).$$

From lemma 2.2 we know that the density of ν_n w.r.t. λ is

$$\frac{d\nu_n}{d\lambda}(x) = \frac{1}{\big(2\pi\big)^{|I|/2}\big(\det(-L^{(n)})\big)^{\frac{1}{2}}}\exp\Big(-\frac{1}{2}x^{\mathrm{T}}(-L^{(n)})x\Big)$$

and consequently the distribution μ_n satisfies

$$\frac{d\mu_n}{d\nu_n}(x) = \frac{d\mu_n}{d\lambda}(x)/\frac{d\nu_n}{d\lambda}(x) \propto \exp\big(F_n\big).$$

Since the right-hand side, up to constants, coincides with the expression in (13), the proof is complete. ∎

3 Main Result

Now we have identified the stationary distribution of the SPDE (in section 1) and of the SDE (in section 2), we can compare the two stationary distributions. The result is given in the following theorem.

Theorem 3.1. Let μ be the stationary distribution of the SPDE (2) on $\mathrm{C}\big([0,1],\mathbb{R}\big)$. Let μ_n be the stationary distribution of the finite element equation (9) on $\mathbb{R}^I$. Let $\mathcal{L}$ be negative and assume $f = F'$ where $F \in \mathrm{C}^2(\mathbb{R})$ is bounded from above with bounded second derivative. Then

$$\big\|\mu \circ \Pi^{-1} - \mu_n\big\|_{\mathrm{TV}} = \mathcal{O}\big(\frac{1}{n}\big)$$

as $n \to \infty$, where $\|\cdot\|_{\mathrm{TV}}$ denotes total-variation distance between probability distributions on $\mathbb{R}^I$.

Before we prove this theorem, we first show some auxiliary results. The following lemma will be used to get rid of the (not explicitly known) normalisation constant Z_n.

Lemma 3.2. Let $(\Omega, \mathcal{F}, \mu)$ be a measure space and $f_1, f_2\colon \Omega \to [0,\infty]$ integrable with $Z_i = \int f_i\,d\mu > 0$ for $i = 1, 2$. Then

$$\int\Big|\frac{f_1}{Z_1} - \frac{f_2}{Z_2}\Big|\,d\mu \le \frac{2}{\max(Z_1, Z_2)}\int\big|f_1 - f_2\big|\,d\mu.$$

proof. Using the L^1-norm $\|f\| = \int |f|\,d\mu$ we can write

$$\Big\|\frac{f_1}{Z_1} - \frac{f_2}{Z_2}\Big\| \le \Big\|\frac{f_1}{Z_1} - \frac{f_2}{Z_1}\Big\| + \Big\|\frac{f_2}{Z_1} - \frac{f_2}{Z_2}\Big\| = \frac{1}{Z_1}\big\|f_1 - f_2\big\| + \frac{|Z_2 - Z_1|}{Z_1 Z_2}\big\|f_2\big\|.$$

Since $Z_i = \|f_i\|$ we can conclude

$$\Big\|\frac{f_1}{Z_1} - \frac{f_2}{Z_2}\Big\| \le \frac{1}{Z_1}\big\|f_1 - f_2\big\| + \frac{\big|\|f_2\| - \|f_1\|\big|}{Z_1} \le \frac{2}{Z_1}\big\|f_1 - f_2\big\|$$

where the second inequality comes from the inverse triangle inequality. Without loss of generality we can assume $Z_1 \geq Z_2$ (otherwise interchange f_1 and f_2 in the above argument) and thus the claim follows. ∎

Lemma 3.3. Let μ and ν be probability measures on $C\big([0,1],\mathbb{R}\big)$ with $\mu \ll \nu$ and let $\Pi\colon C\big([0,1],\mathbb{R}\big) \to \mathbb{R}^I$ be the projection from (11). Then $\mu\circ\Pi^{-1} \ll \nu\circ\Pi^{-1}$ and

$$\frac{d(\mu\circ\Pi^{-1})}{d(\nu\circ\Pi^{-1})}\circ\Pi = \mathbb{E}_\nu\big(\frac{d\mu}{d\nu}\mid \Pi\big).$$

proof. Let $\varphi = \frac{d\mu}{d\nu}$. Since $\mathbb{E}(\varphi|\Pi)$ is Π-measurable, there is a function $\psi\colon \mathbb{R}\to\mathbb{R}$ with $\mathbb{E}(\varphi|\Pi) = \psi\circ\Pi$. Let $A \subseteq \mathbb{R}^I$ be measurable. Then

$$\begin{aligned}\int_{\mathbb{R}} \psi\, 1_A \, d(\nu\circ\Pi^{-1}) &= \int_{C\left([0,1],\mathbb{R}\right)} \psi\circ\Pi\, 1_{\Pi^{-1}(A)}\, d\nu = \int_{C\left([0,1],\mathbb{R}\right)} \mathbb{E}(\varphi|\Pi)\, 1_{\Pi^{-1}(A)}\, d\nu \\ &= \int_{C\left([0,1],\mathbb{R}\right)} \varphi\, 1_{\Pi^{-1}(A)}\, d\nu = \mu\circ\Pi^{-1}(A)\end{aligned}$$

by the definition of conditional expectation. This shows that ψ is indeed the required density. ∎

proof (of theorem 3.1). Let Π, ν and ν_n as in lemma 2.3. Using lemmata 2.3 and 3.3 we find

$$\begin{aligned}\big\|\mu\circ\Pi^{-1} - \mu_n\big\|_{\mathrm{TV}} &= \mathbb{E}_{\nu_n}\Big|\frac{d\mu\circ\Pi^{-1}}{d\nu_n} - \frac{d\mu_n}{d\nu_n}\Big| \\ &= \mathbb{E}_{\nu}\Big|\frac{d\mu\circ\Pi^{-1}}{d\nu\circ\Pi^{-1}}\circ\Pi - \frac{d\mu_n}{d\nu_n}\circ\Pi\Big| \\ &= \mathbb{E}_{\nu}\Big|\mathbb{E}_\nu\big(\frac{d\mu}{d\nu} - \frac{d\mu_n}{d\nu_n}\circ\Pi \mid \Pi\big)\Big| \\ &\leq \mathbb{E}_{\nu}\Big|\frac{d\mu}{d\nu} - \frac{d\mu_n}{d\nu_n}\circ\Pi\Big|.\end{aligned} \tag{14}$$

From lemma 1.3 we know

$$\frac{d\mu}{d\nu}(U) = \frac{1}{Z}\exp\Big(\int_0^1 F\big(U_x\big)\, dx\Big).$$

Lemma 2.5 gives

$$\frac{d\mu_n}{d\nu_n} = \frac{1}{Z_n}\exp\big(F_n\big),$$

and by the definition of F_n we have

$$\frac{d\mu_n}{d\nu_n}\circ\Pi(U) = \frac{1}{Z_n}\exp\Big(\int_0^1 F\big(U_x^{(n)}\big)\, dx\Big)$$

where $U_x^{(n)} = \sum \Pi(U)_j \varphi_j(x)$ for all $U \in C\big([0,1],\mathbb{R}\big)$ and φ_j, $j \in I$ are the finite element basis functions. Using lemma 3.2 we get

$$\begin{aligned}\big\|\mu\circ\Pi^{-1} - \mu_n\big\|_{\mathrm{TV}} &\leq \mathbb{E}_{\nu}\Big|\frac{d\mu}{d\nu} - \frac{d\mu_n}{d\nu_n}\circ\Pi\Big| \\ &\leq \frac{2}{Z}\int\Big|\exp\Big(\int_0^1 F\big(U_x\big)\, dx\Big) - \exp\Big(\int_0^1 F\big(U_x^{(n)}\big)\, dx\Big)\Big|\, d\nu(U).\end{aligned}$$

Since the inequality $\big|e^x - 1\big| \leq |x|\exp\big(|x|\big)$ holds for all $x \in \mathbb{R}$, we conclude

$$\begin{aligned}&\big\|\mu\circ\Pi^{-1} - \mu_n\big\|_{\mathrm{TV}} \\ &\leq \frac{2}{Z}\mathbb{E}\left(\exp\Big(\int_0^1 F\big(U_x^{(n)}\big)\, dx\Big)\cdot\Big|\exp\Big(\int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big)\, dx\Big) - 1\Big|\right) \\ &\leq \frac{2}{Z}\mathbb{E}\left(\exp\Big(\int_0^1 F\big(U_x^{(n)}\big)\, dx\Big)\right. \\ &\qquad \left.\cdot\Big|\int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big)\, dx\Big|\cdot\exp\Big(\Big|\int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big)\, dx\Big|\Big)\right)\end{aligned}$$

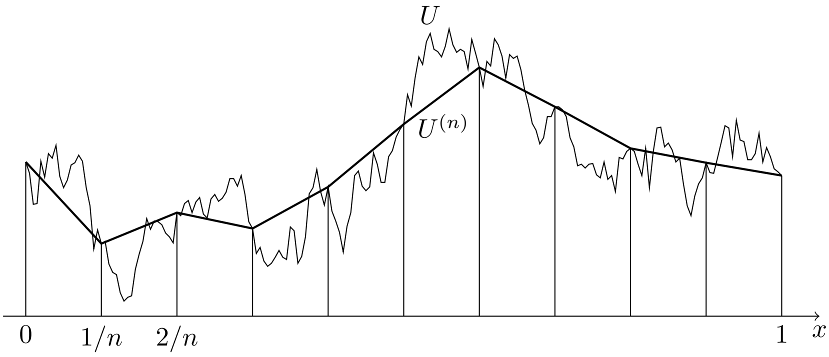

Figure 1: *Illustration of the convergence of $U^{(n)}$ to U. Under the distribution ν, the path U is a Brownian Bridge with random boundary conditions. Since $U^{(n)}$ is the linear interpolation of U between the grid points, the difference $U^{(n)} - U$ consists of a chain of n independent Brownian bridges.*

where U is distributed according to the Gaussian measure ν. Since F is bounded from above we can estimate the the first exponential in the expectation by a constant. Using the Cauchy-Schwarz inequality we get

$$\begin{aligned}&\big\|\mu \circ \Pi^{-1} - \mu_n\big\|_{\mathrm{TV}} \\ &\qquad \leq c_1 \Big\| \int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big)\, dx \Big\|_2 \cdot \Big\| \exp\Big(\big| \int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big)\, dx \big| \Big) \Big\|_2 \qquad (15)\end{aligned}$$

for some constant c_1.

The main step in the proof is to estimate the right-hand side of 15 by showing that $\big|\int_0^1 F(U_x) - F(U_x^{(n)})\, dx\big|$ gets small as $n \to \infty$. By lemma 1.2, the path U in stationarity is just a Brownian bridge (with random boundary values) and, by definition, $U^{(n)}$ is the linear interpolation of the values of U at the grid points (see figure 1 for illustration). Thus, the difference $U^{(n)} - U$ can be written as

$$\big(U - U^{(n)}\big)(x) = \sum_{i=1}^{n} 1_{[\frac{i-1}{n}, \frac{i}{n}]}(x)\, \frac{1}{\sqrt{n}} B^{(i)}_{nx-(i-1)}$$

where $B^{(1)}, \ldots, B^{(n)}$ are standard Brownian bridges, independent of each other and of $U^{(n)}$. Using Taylor approximation for F we find

$$\begin{aligned}&\Big| \int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big)\, dx \Big| \\ &\qquad \leq \Big| \int_0^1 f\big(U_x^{(n)}\big)\big(U_x - U_x^{(n)}\big)\, dx \Big| + \frac{1}{2}\|F''\|_\infty \int_0^1 \big(U_x - U_x^{(n)}\big)^2\, dx \\ &\qquad =: \big|P_n\big| + \frac{1}{2}\|F''\|_\infty Q_n. \qquad (16)\end{aligned}$$

For the term $|P_n|$ we find

$$P_n = \sum_{i=1}^{n} \int_0^{1/n} f\big(U^{(n)}_{\frac{i-1}{n}+x}\big) \frac{1}{\sqrt{n}} B^{(i)}_{nx}\, dx = n^{-3/2} \sum_{i=1}^{n} \int_0^1 f\big(U^{(n)}_{\frac{i-1}{n}+y/n}\big) B^{(i)}_y\, dy$$

where $B^{(i)}$ are the Brownian bridges defined above. As an abbreviation write $\bar f_i(y) = f\big(U^{(n)}_{\frac{i-1}{n}+y/n}\big)$. Conditioned on the value of $U^{(n)}$, the integrals $\int_0^1 \bar f_i(y) B^{(i)}_y\, dy$ are centred Gaussian with

$$\begin{aligned}\mathrm{Var}\Big(\int_0^1 \bar f_i(y) B^{(i)}_y\, dy \;\Big|\; U^{(n)} \Big) &= \int_0^1 \bar f_i(y) \int_0^1 (y \wedge z - yz) \bar f_i(z)\, dz\, dy \\ &\leq c_2 \|\bar f_i\|_\infty^2 \leq c_3^2 \big(\|U^{(n)}\|_\infty + 1\big)^2\end{aligned}$$

for some constants $c_2, c_3 > 0$ where the last inequality uses the fact the f is Lipschitz continuous. We get $\mathbb{E}\big(P_n \bigm| \Pi(U)\big) = 0$ and, since the $B^{(i)}$ are independent,

$$\mathbb{E}\big(P_n^2 \bigm| U^{(n)}\big) \leq c_3^2 \big(\|U^{(n)}\|_\infty + 1\big)^2 n^{-2}.$$

Using the tower property for conditional expectations, and using $\|U^{(n)}\|_\infty \leq \|U\|_\infty$, we get

$$\mathbb{E}\big(P_n^2\big) \leq c_3^2 \mathbb{E}\Big(\big(\|U\|_\infty + 1\big)^2\Big) n^{-2}.$$

Since U is a Gaussian process, $\big\|\|U\|_\infty + 1\big\|_2$ is finite, and we can conclude

$$\mathbb{E}\big(P_n^2\big) \leq c_4^2 n^{-2}$$

for some constant c_4. Similarly, for Q_n we find

$$Q_n = \sum_{i=1}^n \int_0^1 (B_y^{(i)})^2 \, dy \cdot n^{-2} \tag{17}$$

and thus, using independence of the $B^{(i)}$ again,

$$\mathbb{E}\big(Q_n^2\big) = c_5^2 n^{-3}$$

for some constant c_5. Combining these estimates we get

$$\Big\| \int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big) \, dx \Big\|_2 = \big\|P_n\big\|_2 + \frac{1}{2} \big\|F''\big\|_\infty \big\|Q_n\big\|_2 \leq c_4 n^{-1} + c_5 n^{-3/2} \tag{18}$$

and thus we have shown the required bound for the first factor of (15).

Finally, we have to show that the second factor in (15) is bounded, uniformly in n: From (16) we get

$$\begin{aligned} \Big\| \exp\Big(\Big| \int_0^1 F\big(U_x\big) - F\big(U_x^{(n)}\big) \, dx \Big| \Big) \Big\|_2 &\leq \mathbb{E}\Big(\exp\big(2|P_n| + \|F''\|_\infty Q_n\big) \Big)^{1/2} \\ &\leq \big\| \mathrm{e}^{2|P_n|} \big\|_2^{1/2} \cdot \big\| \mathrm{e}^{\|F''\|_\infty Q_n} \big\|_2^{1/2}. \end{aligned} \tag{19}$$

It is easy to check that, for all $\sigma > 0$, an $\mathcal{N}(0, \sigma^2)$-distributed random variable X satisfies the inequality $\mathbb{E}(\mathrm{e}^{|X|}) \leq 2\mathrm{e}^{\sigma^2/2}$ and thus we have

$$\big\| \mathrm{e}^{2|P_n|} \big\|_2^{1/2} = \mathbb{E}\big(\mathrm{e}^{4|P_n|}\big)^{1/4} \leq 2 \exp(8 \, c_4^2 n^{-2}) < 1 \tag{20}$$

for all sufficiently large n. Furthermore, using (17) and the fact that the $B^{(i)}$ are i.i.d., we find

$$\big\| \mathrm{e}^{\|F''\|_\infty Q_n} \big\|_2^{1/2} = \mathbb{E}\Big(\exp\big(2\|F''\|_\infty |B^{(1)}|_{\mathrm{L}^2}^2 \, n^{-2}\big) \Big)^{n/4} \tag{21}$$

where we write $|\cdot|_{\mathrm{L}^2}$ for the L^2-norm on the space $\mathrm{L}^2\big([0,1], \mathbb{R}\big)$. By Fernique's theorem (Fernique, 1970) there exists an $\varepsilon > 0$ with $\mathbb{E}\big(\exp(\varepsilon \, |B^{(1)}|_{\mathrm{L}^2}^2)\big) < \infty$. For all $\lambda > 0$ we have

$$\mathbb{E}\big(\mathrm{e}^{\lambda |B^{(1)}|_{\mathrm{L}^2}^2}\big) = \int_0^\infty \mathbb{P}\big(\mathrm{e}^{\lambda |B^{(1)}|_{\mathrm{L}^2}^2} \geq a\big) \, da \leq 1 + \int_0^\infty \mathbb{P}\big(|B^{(1)}|_{\mathrm{L}^2}^2 \geq b\big) \, \lambda \mathrm{e}^{\lambda b} \, db$$

and using Markov's inequality $\mathbb{P}\big(|B^{(1)}|_{\mathrm{L}^2}^2 \geq b\big) \leq \mathbb{E}\big(\mathrm{e}^{\varepsilon |B^{(1)}|_{\mathrm{L}^2}^2}\big) \mathrm{e}^{-\varepsilon b}$ we get

$$\mathbb{E}\big(\mathrm{e}^{\lambda |B^{(1)}|_{\mathrm{L}^2}^2}\big) \leq 1 + \int_0^\infty \mathbb{E}\big(\mathrm{e}^{\varepsilon |B^{(1)}|_{\mathrm{L}^2}^2}\big) \mathrm{e}^{-\varepsilon b} \, \lambda \mathrm{e}^{\lambda b} \, db = 1 + \frac{\lambda}{\varepsilon - \lambda} \mathbb{E}\big(\mathrm{e}^{\varepsilon |B^{(1)}|_{\mathrm{L}^2}^2}\big)$$

for all $\lambda < \varepsilon$. Substituting this bound into (21) we have

$$\big\| \mathrm{e}^{\|F''\|_\infty Q_n} \big\|_2^{1/2} \leq \big(1 + \frac{c_6}{n^2}\big)^n \leq 2 \tag{22}$$

for some constant c_6 and all sufficiently large n. From (20) and (22) we see that the right-hand side of (19) is bounded uniformly in n, *i.e.*

$$\Big\|\exp\Big(\big|\int_0^1 F(U_x) - F\big(U_x^{(n)}\big)\,dx\big|\Big)\Big\|_2 \leq c_7 \tag{23}$$

for all $n \in \mathbb{N}$ and some constant c_7.

Combining (18) and (23) we see that the right-hand side in (15) is of order $\mathcal{O}(n^{-1})$. This completes the proof. ∎

For deterministic problems, the interpolation error of finite element methods can often be reduced by considering basis functions composed of higher-order polynomials instead of the linear basis functions considered here. This is a consequence of Céa's lemma, which states that the rate of error of the finite element approximation is determined by how well the exact solution can be approximated by the basis functions (see *e.g.* theorem 2.8.1 in Brenner and Scott, 2002). In contrast, when approximating Brownian motion by splines, the approximation error is always of order $1/\sqrt{n}$ where n is the number of spline-nodes, independent of the polynomial order of the splines (see for example Kon and Plaskota, 2005; Creutzig et al., 2007). For this reason we would expect that, for the situation considered here, the order of the leading error term P_n in (16) cannot be improved by considering higher order polynomial basis functions.

In theorem 3.1 we compared the stationary distribution μ_n of the finite element SDE on $\mathbb{R}^I$ and the stationary distribution μ of the SPDE on $\mathrm{C}\big([0,1],\mathbb{R}\big)$ by projecting μ onto the finite dimensional space $\mathbb{R}^I$. An alternative approach is to embed $\mathbb{R}^I$ into $\mathrm{C}\big([0,1],\mathbb{R}\big)$ instead. A naïve implementation of this idea would be to extend vectors from $\mathbb{R}^I$ to continuous functions via linear interpolation. Unfortunately, the image of μ_n when projected to $\mathrm{C}\big([0,1],\mathbb{R}\big)$ in this way would be mutually singular with μ and thus the total variation norm would not provide a useful measure for the distance between the two distributions. For this reason, we choose here a different approach, described in the following definition.

Definition 3.4. Given a probability measure μ_n on $\mathbb{R}^I$, we define a distribution $\hat{\mu}_n$ as follows: Consider a random variable X which is distributed according to μ_n. Given X, construct $Y \in \mathrm{C}\big([0,1],\mathbb{R}\big)$ by setting $Y(k\,\Delta x) = X_k$ for $k = 0, 1, \ldots, n$ and filling the gaps between these points with n Brownian bridges, independent of X and of each other. Then we denote the distribution of Y by $\hat{\mu}_n$.

Lemma 3.5. Let μ_n and ν_n be probability measures on $\mathbb{R}^I$ with $\mu_n \ll \nu_n$. Then $\hat{\mu}_n \ll \hat{\nu}_n$ with

$$\frac{d\hat{\mu}_n}{d\hat{\nu}_n} = \frac{d\mu_n}{d\nu_n} \circ \Pi$$

on $\mathrm{C}\big([0,1],\mathbb{R}\big)$.

proof. Let $\psi = \frac{d\mu_n}{d\nu_n}$. Using substitution we get

$$\mathbb{E}_{\hat{\mu}_n}\big(f \circ \Pi\big) = \int_{\mathbb{R}^I} f\,d\mu_n = \int_{\mathbb{R}^I} f\psi\,d\nu_n = \mathbb{E}_{\hat{\nu}_n}\big(f \circ \Pi \cdot \psi \circ \Pi\big)$$

for all integrable $f\colon \mathbb{R}^I \to \mathbb{R}$. Since, conditioned on the value of Π, the distributions $\hat{\mu}_n$ and $\hat{\nu}_n$ are the same, we can use the tower property to get

$$\begin{aligned}\hat{\mu}(A) &= \mathbb{E}_{\hat{\mu}_n}\big(\mathbb{E}_{\hat{\mu}_n}(1_A|\Pi)\big) = \mathbb{E}_{\hat{\mu}_n}\big(\mathbb{E}_{\hat{\nu}_n}(1_A|\Pi)\big)\\ &= \mathbb{E}_{\hat{\nu}_n}\big(\mathbb{E}_{\hat{\nu}_n}(1_A|\Pi) \cdot \psi \circ \Pi\big) = \mathbb{E}_{\hat{\nu}_n}\big(1_A \psi \circ \Pi\big)\end{aligned}$$

for every measurable set A. This shows that $\psi \circ \Pi$ is the required density. ∎

Corollary 3.6. Let μ be the stationary distribution of the SPDE (2) on $\mathrm{C}\big([0,1],\mathbb{R}\big)$. Let μ_n be the stationary distribution of the finite element equation (9) on $\mathbb{R}^I$. Let

$\mathcal{L}$ be negative and assume $f = F'$ where $F \in \mathrm{C}^2(\mathbb{R})$ is bounded from above with bounded second derivative. Then

$$\|\mu - \hat{\mu}_n\|_{\mathrm{TV}} = \mathcal{O}\big(\frac{1}{n}\big)$$

as $n \to \infty$.

proof. Let ν be the stationary distribution of the linear SPDE (4) on $\mathrm{C}\big([0,1],\mathbb{R}\big)$ and let ν_n be the stationary distribution of the linear finite element equation (10) on $\mathbb{R}^I$. By construction of the process U in the third statement of lemma 1.2 and by the Markov property for Brownian bridges, the distribution of U between the grid points, conditioned on the values at the grid points, coincides with the distribution of n independent Brownian bridges. By lemma 2.3 the distribution of U on the grid points is given by ν_n. Thus we have $\nu = \hat{\nu}_n$. Using this equality and lemma 3.5 we find

$$\|\mu - \hat{\mu}_n\|_{\mathrm{TV}} = \mathbb{E}_\nu \Big| \frac{d\mu}{d\nu} - \frac{d\hat{\mu}_n}{d\nu} \Big| = \mathbb{E}_\nu \Big| \frac{d\mu}{d\nu} - \frac{d\mu_n}{d\nu_n} \circ \Pi \Big|.$$

Now we are in the situation of equation (14) and the proof of theorem 3.1 applies without further changes. ∎

4 Examples

To illustrate that the suggested finite element method is a concrete and implementable scheme, this section gives two examples for the finite element discretisation of SPDEs, both in the context of infinite dimensional sampling problems.

For the first example, for $c > 0$, consider the SPDE

$$\partial_t u(t,x) = \partial_x^2 u(t,x) - c^2 u(t,x) + \sqrt{2}\partial_t w(t,x) \qquad \forall (t,x) \in \mathbb{R}_+ \times (0,1)$$

with Robin boundary conditions

$$\partial_x u(t,0) = c u(t,0), \quad \partial_x u(t,1) = -c u(t,1) \qquad \forall t \in \mathbb{R}_+, \tag{24}$$

where $\partial_t w$ is space-time white noise. From Hairer et al. (2007) we know that the stationary distribution of this SPDE on $\mathrm{C}\big([0,1],\mathbb{R}\big)$ coincides with the distribution of the process X given by

$$\begin{aligned} dX_\tau &= -cX_\tau\, d\tau + dW_\tau \qquad \forall \tau \in [0,1] \\ X_0 &\sim \mathcal{N}\big(0, \frac{1}{2c}\big), \end{aligned} \tag{25}$$

where the time τ in the SDE plays the rôle of the space x in the SPDE. In the framework of section 1, the boundary conditions (24) correspond to the case $\alpha_0 = \alpha_1 = c$ and $\beta_0 = \beta_1 = 1$. Since $\beta_i \neq 0$, we need to include both boundary points in the finite element discretisation and thus have $I = \{0, 1, \ldots, n\}$ and $\mathbb{R}^I \cong \mathbb{R}^{n+1}$. The matrix $L^{(n)}$ is given by

$$L^{(n)} = \frac{1}{\Delta x} \begin{pmatrix} -1 - c\Delta x & 1 & & & \\ 1 & -2 & 1 & & \\ & 1 & -2 & 1 & \\ & & 1 & -2 & 1 \\ & & & 1 & -1 - c\Delta x \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)},$$

where the middle rows are repeated along the diagonal to obtain tridiagonal $(n+1) \times (n+1)$-matrices. Similarly, the mass matrix M is given by

$$M = \frac{\Delta x}{6} \begin{pmatrix} 2 & 1 & & & \\ 1 & 4 & 1 & & \\ & 1 & 4 & 1 & \\ & & 1 & 4 & 1 \\ & & & 1 & 2 \end{pmatrix} \in \mathbb{R}^{(n+1)\times(n+1)}$$

where, again, the middle rows are repeated along the diagonal. Finally, it transpires that the discretised drift for this example is given by $f_n(u) = -cMu$. By lemma 2.4 the $n+1$ dimensional SDEs

$$dU_t = M^{-1} L^{(n)} U_t \, dt - cU_t \, dt + \sqrt{2} M^{-1/2} \, dW_t$$

and

$$dU_t = L^{(n)} U_t \, dt - cMU_t \, dt + \sqrt{2} \, dW_t,$$

where W is an $(n+1)$-dimensional Brownian motion, both have the same stationary distribution and this stationary distribution converges to the distribution of the process X from (25) in the sense given in theorem 3.1 and corollary 3.6.

As a second example, consider the SPDE

$$\partial_t u(t,x) = \partial_x^2 u(t,x) - \big(gg' + \frac{1}{2} g''\big)(u) + \sqrt{2} \partial_t w(t,x) \qquad \forall (t,x) \in \mathbb{R}_+ \times (0,1)$$

with Dirichlet boundary conditions

$$u(t,0) = u(t,1) = 0 \qquad \forall t \in \mathbb{R}_+,$$

where $g \in \mathrm{C}^3\big(\mathbb{R}, \mathbb{R}\big)$ with bounded derivatives g', g'' and g'''. From Hairer et al. (2007) we know that the stationary distribution of this SPDE on $\mathrm{C}\big([0,1], \mathbb{R}\big)$ coincides with the conditional distribution of the process X given by

$$\begin{aligned} dX_\tau &= g(X_\tau) \, d\tau + dW_\tau \qquad \forall \tau \in [0,1] \\ X_0 &= 0, \end{aligned} \tag{26}$$

conditioned on $X_1 = 0$.

Since we have Dirichlet boundary conditions, the boundary points in the finite element discretisation are not included: we have $I = \{1, 2, \ldots, n-1\}$ and $\mathbb{R}^I \cong \mathbb{R}^{n-1}$. The matrices $L^{(n)}$ and M are given by

$$L^{(n)} = \frac{1}{\Delta x} \begin{pmatrix} -2 & 1 & \\ 1 & -2 & 1 \\ & 1 & -2 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)},$$

and

$$M = \frac{\Delta x}{6} \begin{pmatrix} 4 & 1 & \\ 1 & 4 & 1 \\ & 1 & 4 \end{pmatrix} \in \mathbb{R}^{(n-1)\times(n-1)}$$

where, again, the middle rows are repeated along the diagonal to obtain matrices of the required size. The discretised drift f_n can be computed from (8); if an analytical solution is not available, numerical integration can be used. By the assumptions on g, the function $F = -\frac{1}{2}(g^2 + g')$ satisfies the conditions of theorem 3.1. Thus, the stationary distributions of the $(n-1)$-dimensional SDEs

$$dU_t = M^{-1} L^{(n)} U_t \, dt + M^{-1} f_n(U_t) \, dt + \sqrt{2} M^{-1/2} \, dW_t$$

and

$$dU_t = L^{(n)} U_t \, dt + f_n(U_t) \, dt + \sqrt{2} \, dW_t$$

coincide and converge to the conditional distribution of X from (26), conditioned on $X_1 = 0$.

References

Ludwig Arnold. *Stochastic Differential Equations: Theory and Applications.* John Wiley & Sons, 1974.

Alexandros Beskos, Gareth O. Roberts, Andrew M. Stuart, and Jochen Voss. MCMC methods for diffusion bridges. *Stochastics and Dynamics*, 8(3):319–350, 2008. doi: 10.1142/S0219493708002378.

Nawaf Bou-Rabee and Martin Hairer. Non-asymptotic mixing of the MALA algorithm. To appear in IMA Journal of Numerical Analysis, 2012.

Susanne C. Brenner and L. Ridgway Scott. *The Mathematical Theory of Finite Element Methods*. Springer, second edition, 2002.

Claudio Cacciapuoti and Domenico Finco. Graph-like models for thin waveguides with Robin boundary conditions. *Asymptotic Analysis*, 70(3–4):199–230, 2010. doi: 10.3233/ASY-2010-1014.

Jakob Creutzig, Thomas Müller-Gronbach, and Klaus Ritter. Free-knot spline approximation of stochastic processes. *Journal of Complexity*, 23(4–6):867–889, 2007. doi: 10.1016/j.jco.2007.05.003.

Giuseppe Da Prato and Jerzy Zabczyk. *Stochastic Equations in Infinite Dimensions*, volume 44 of *Encyclopedia of Mathematics and its Applications*. Cambridge University Press, 1992. ISBN 0-521-38529-6.

Giuseppe Da Prato and Jerzy Zabczyk. *Ergodicity for Infinite-Dimensional Systems*, volume 229 of *London Mathematical Society Lecture Note Series*. Cambridge University Press, 1996.

Xavier Fernique. Intégrabilité des vecteurs gaussiens. *Comptes rendus hebdomadaires des séances de l'Académie des sciences, série A*, 270:1698–1699, 1970.

István Gyöngy and Annie Millet. Rate of convergence of space time approximations for stochastic evolution equations. *Potential Analysis*, 30(1):29–64, 2009. doi: 10.1007/s11118-008-9105-5.

Martin Hairer, Andrew M. Stuart, Jochen Voss, and Petter Wiberg. Analysis of SPDEs arising in path sampling, part I: The Gaussian case. *Communications in Mathematical Sciences*, 3(4):587–603, 2005.

Martin Hairer, Andrew M. Stuart, and Jochen Voss. Analysis of SPDEs arising in path sampling, part II: The nonlinear case. *Annals of Applied Probability*, 17(5): 1657–1706, 2007. doi: 10.1214/07-AAP441.

Martin Hairer, Andrew M. Stuart, and Jochen Voss. Sampling conditioned diffusions. In *Trends in Stochastic Analysis*, volume 353 of *London Mathematical Society Lecture Note Series*, pages 159–186. Cambridge University Press, 2009. ISBN 9780521718219.

Erika Hausenblas. Finite element approximation of stochastic partial differential equations driven by Poisson random measures of jump type. *SIAM Journal on Numerical Analysis*, 46(1):437–471, 2008. doi: 10.1137/050654141.

I. Iscoe, M. B. Marcus, D. McDonald, M. Talagrand, and J. Zinn. Continuity of l^2-valued Ornstein-Uhlenbeck processes. *The Annals of Probability*, 18(1):68–84, 1990.

Arnulf Jentzen. Higher order pathwise numerical approximations of SPDEs with additive noise. *SIAM Journal on Numerical Analysis*, 49(2):642–667, 2011. doi: 10.1137/080740714.

C. Johnson. *Numerical Solution of Partial Differential Equations by the Finite Element Method*. Cambridge University Press, 1990.

Mark Kon and Leszek Plaskota. Information-based nonlinear approximation: an average case setting. *Journal of Complexity*, 21(2):211–229, 2005. doi: 10.1016/j.jco.2004.11.001.

Peter Lancaster and Leiba Rodman. *Algebraic Riccati Equations*. Clarendon Press, Oxford, 1995.

Annie Millet and Pierre-Luc Morien. On implicit and explicit discretization schemes for parabolic SPDEs in any dimension. *Stochastic Processes and their Applications*, 115(7):1073–1106, 2005. doi: 10.1016/j.spa.2005.02.004.

J. B. Walsh. Finite element methods for parabolic stochastic PDE's. *Potential Analysis*, 23(1):1–43, 2005. doi: 10.1007/s11118-004-2950-y.

Jerzy Zabczyk. Symmetric solutions of semilinear stochastic equations. In Giuseppe Da Prato and L. Tubaro, editors, *Stochastic Partial Differential Equations and Applications II*, volume 1390 of *Lecture Notes in Mathematics*, pages 237–256. Springer, 1989. doi: 10.1007/BFb0083952.