International Journal of Computer Mathematics

Publication details, including instructions for authors and subscription information:
http://www.tandfonline.com/loi/gcom20

Approximation of stochastic partial differential equations by a kernel-based collocation method

Igor Cialenco a, Gregory E. Fasshauer a & Qi Ye a

a Department of Applied Mathematics, Illinois Institute of Technology, 10 West 32nd Str, Bld E1, Room 208, Chicago, IL, 60616, USA

Available online: 28 May 2012


To link to this article: http://dx.doi.org/10.1080/00207160.2012.688111

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Approximation of stochastic partial differential equations by a kernel-based collocation method

Igor Cialenco, Gregory E. Fasshauer and Qi Ye*

Department of Applied Mathematics, Illinois Institute of Technology, 10 West 32nd Str, Bld E1, Room 208, Chicago, IL 60616, USA

(Received 21 August 2011; revised version received 14 April 2012; accepted 19 April 2012)

In this paper we present the theoretical framework needed to justify the use of a kernel-based collocation method (meshfree approximation method) to estimate the solution of high-dimensional stochastic partial differential equations (SPDEs). Using an implicit time-stepping scheme, we transform stochastic parabolic equations into stochastic elliptic equations. Our main attention is concentrated on the numerical solution of the elliptic equations at each time step. The estimator of the solution of the elliptic equations is given as a linear combination of reproducing kernels derived from the differential and boundary operators of the SPDE centred at collocation points to be chosen by the user. The random expansion coefficients are computed by solving a random system of linear equations. Numerical experiments demonstrate the feasibility of the method.

Keywords: kernel-based collocation; numerical solutions; stochastic partial differential equation; reproducing kernel; Matérn function; Gaussian process

2010 AMS Subject Classifications: 46E22; 65D05; 60G15; 60H15; 65N35

1. Introduction

Stochastic partial differential equations (SPDEs) frequently arise from applications in areas such as physics, engineering and finance. However, in many cases it is difficult to derive an explicit form of their solutions. Moreover, current numerical algorithms often show limited success for high-dimensional problems or in complex domains – even for deterministic partial differential equations (PDEs). The kernel-based approximation method (meshfree approximation method [4,8,21]) is a relatively new numerical tool for the solution of high-dimensional problems. In this paper we apply – to our knowledge for the first time – such a kernel-based collocation method to construct numerical estimators for SPDEs. Galerkin-type approximation methods based on the eigenvalues and eigenfunctions of the underlying differential operator are currently very popular for the numerical solution of SPDEs [7,14,17]. With the kernel-based meshfree collocation method introduced here, explicit knowledge of these eigenvalues and eigenfunctions is not required since the kernels can be directly obtained as Green kernels of the differential operators [10,11].

*Corresponding author. Email: qye3@iit.edu

ISSN 0020-7160 print/ISSN 1029-0265 online
© 2012 Taylor & Francis
http://dx.doi.org/10.1080/00207160.2012.688111
http://www.tandfonline.com
collocation methods using a polynomial basis are also frequently found in the literature [2,18]. These methods usually require the collocation points to lie on a regular grid. In our method, the collocation points can be placed at rather arbitrarily scattered locations. This allows for the use of either deterministic or random designs such as, e.g. uniform or Sobol points, but also for the use of adaptively chosen locations. In this paper, we do not study the design aspect of our algorithm, but reserve this important aspect for future work. Another advantage of using a meshfree method is its ability – also not explicitly pursued here – to deal with problems on a complex domain 

Another advantage of using a meshfree method is its ability to deal with problems on a complex domain. This allows for the use of collocation points at rather arbitrarily scattered locations. This allows for the use of collocation methods using a polynomial basis are also frequently found in the literature [2,18].

I. Cialenco et al.

Assume that \( D \) is a regular open bounded domain in \( \mathbb{R}^d \) (see Appendix 2), and let \( H \) be a separable Hilbert space of functions defined on \( D \). Also, let \((\Omega_W, \mathcal{F}_W, [\mathbb{F}_t], \mathbb{P}_W)\) be a stochastic basis with the usual assumptions. We consider the following parabolic Itô equation:

\[
\begin{aligned}
    dU_t &= AU_t \, dt + \sigma \, dW_t \quad \text{in } D, \\
    BU_t &= 0 \quad \text{on } \partial D, \\
    U_0 &= u_0,
\end{aligned}
\]

where \( A \) is a linear elliptic operator in \( H \), \( B \) is a boundary operator for Dirichlet or Neumann boundary conditions, \( u_0 \in H \), and \( W \) is a Wiener process in \( H \), with mean zero and spatial covariance function \( R \) given by \( \mathbb{E}(W(t,x)W(s,y)) = \min\{t,s\}R(x,y), x,y \in D, t,s > 0 \), and \( \sigma > 0 \) (see for instance [5]).

We assume that Equation (1) has a unique solution \( U \in L_2(\Omega_W \times (0,T); H) \).

The proposed numerical method for solving a general SPDE of the form (1) can be described as follows:

(S1) Discretize (1) in time by the implicit Euler scheme at equally spaced time points \( 0 = t_0 < t_1 < \cdots < t_n = T \),

\[
U_{t_j} - U_{t_{j-1}} = AU_{t_j} \delta t + \sigma \delta W_j, \quad j = 1, \ldots, n,
\]

where \( \delta t := t_j - t_{j-1} \) and \( \delta W_j := W_{t_j} - W_{t_{j-1}} \).

(S2) Since it follows from (2) and the definition of Brownian motion that the noise increment \( \delta W_j \) at each time instance \( t_j \) is independent of the solution \( U_{t_{j-1}} \) at the previous step, we simulate the Gaussian field with covariance structure \( R(x,y) \) at a finite collection of predetermined collocation points

\[
X_D := \{x_1, \ldots, x_N\} \subset D, \quad X_{\partial D} := \{x_{N+1}, \ldots, x_{N+M}\} \subset \partial D.
\]

(S3) Let the differential operator \( P := I - \delta t A \), and the noise term \( \xi := \sigma \delta W_j \). Since \( \xi \) is a Gaussian field with \( \mathbb{E}(\xi_x) = 0 \) and \( \text{Cov}(\xi_x \xi_y) = \sigma^2 \delta t R(x,y) \), Equation (2) together with the
corresponding boundary condition becomes an elliptic SPDE of the form
\[
\begin{align*}
Pu &= f + \xi & \text{in } \mathcal{D}, \\
Bu &= 0 & \text{on } \partial\mathcal{D},
\end{align*}
\]
(3)
where \( u := U_t \) is seen as an unknown part and \( f := U_{t-} \) and \( \xi \) are viewed as given parts. We will solve for \( u \) using a kernel-based collocation method written as
\[
u(x) \approx \hat{u}(x) := \sum_{k=1}^N c_k P_2^* K(x, x_k) + \sum_{k=1}^M c_{N+k} B_2^* K(x, x_{N+k}), \quad x \in \mathcal{D},
\]
(4)
where \( K \) is a reproducing kernel and the integral-type kernels \( P_2^*, B_2^* \) are defined in Lemmas 2.2 and A.7. The unknown random coefficients \( c := (c_1, \ldots, c_{N+M})^T \) are obtained by solving a random system of linear equations (with constant deterministic system matrix and different random right-hand side) at each time step. Details are provided in Section 3.

(S4) Repeat S2 and S3 for all \( j = 1, \ldots, n \).

Obviously, many other – potentially better – time-stepping schemes could be applied here. However, as mentioned earlier, we focus mainly on step S3 and are for the time being content with using the implicit Euler scheme. Naturally, the rate of convergence of the above numerical scheme depends on the size of the time step \( \delta t \), and the fill distance \( h_\mathcal{D} := \sup_{x \in D} \min_{x_k \in \partial\mathcal{D} \cup \mathcal{D}} \| x - x_k \|_2 \) of the collocation points. We support this statement empirically in Section 4. We should mention that even for deterministic time-dependent PDEs to find the exact rates of convergence of kernel-based methods is a delicate and nontrivial question, only recently solved in [12]. We will address this question in the case of SPDEs in future works.

The fundamental building block of our mesh-free method is the reproducing kernel \( K : \mathcal{D} \times \mathcal{D} \to \mathbb{R} \) and its reproducing-kernel Hilbert space \( H_K(\mathcal{D}) \) (see Appendix 1 for more details). By the very nature of such a kernel-based method, the approximate solution \( U_j, j = 1, \ldots, n \), must live in \( H_K(\mathcal{D}) \). Thus, we make the following standing assumption throughout the paper:

The solution \( U \) of the original equation (1) lies in a Hilbert space \( \mathcal{H} \) which can be embedded in the reproducing-kernel Hilbert space \( H_K(\mathcal{D}) \).

Usually, \( \mathcal{H} \) is a Sobolev space \( \mathcal{H}^m(\mathcal{D}) \), for some positive \( m \). In this case it is possible to choose an appropriate kernel \( K \) such that the above embedding assumption is satisfied. For a general discussion of existence and uniqueness of the solution of problem (1) see, e.g., [5,6,19].

2. Reproducing-kernel collocation method for Gaussian processes

In this section, we briefly review the standard kernel-based approximation method for high-dimensional interpolation problems. However, since we will later be interested in solving an SPDE, we present the following material mostly from the stochastic point of view. For further discussion of this method, we refer the reader to the recent survey papers [9,20] and references therein.

Assume that the function space \( H_K(\mathcal{D}) \) is a reproducing-kernel Hilbert space with norm \( \| \cdot \|_{K, \mathcal{D}} \) and its reproducing kernel \( K \in C(\mathcal{D} \times \mathcal{D}) \) is symmetric positive definite (see Appendix A.1). Given the data values \( \{ y_j \}_{j=1}^N \subset \mathbb{R} \) at the collocation points \( X_\mathcal{D} := \{ x_j \}_{j=1}^N \subset \mathcal{D} \) of an unknown function \( u \in H_K(\mathcal{D}) \), i.e.
\[
y_j = u(x_j), \quad x_j = (x_{1,j}, \ldots, x_{d,j}) \in \mathcal{D} \subset \mathbb{R}^d, \quad j = 1, \ldots, N,
\]
the goal is to find an optimal estimator from \( H_K(\mathcal{D}) \) that interpolates these data.
A stochastic process \( S : \Omega \times \Omega \to \mathbb{R} \) is said to be Gaussian with mean \( \mu : \Omega \to \mathbb{R} \) and covariance kernel \( \Phi : \Omega \times \Omega \to \mathbb{R} \) if, for any pairwise distinct points \( X_\Omega := \{ x_1, \ldots, x_N \} \subset \Omega \), the random vector \( S := (S_{x_1}, \ldots, S_{x_N})^T \) is a multi-normal random variable on \( \Omega, \mathcal{F}, \mathbb{P} \) with mean \( \mu \) and covariance matrix \( \Phi \), i.e. \( S \sim \mathcal{N}(\mu, \Phi) \), where \( \mu := (\mu(x_1), \ldots, \mu(x_N))^T \) and \( \Phi := (\Phi(x_j, x_k))_{j,k=1}^N \).

### 2.1 Data-fitting problems via deterministic interpolation and simple kriging

In the deterministic formulation of kernel interpolation, we solve an optimization problem by minimizing the reproducing-kernel norm subject to interpolation constraints, i.e.

\[
\hat{u}_K = \arg\min_{u \in H_K(D)} \| u \|_{K,D} \text{ s.t. } u(x_j) = y_j, \; j = 1, \ldots, N.
\]

In this case, the minimum norm interpolant (also called the collocation solution) \( \hat{u}_K(x) \) is a linear combination of ‘shifts’ of the reproducing kernel \( K \),

\[
\hat{u}_K(x) := \sum_{k=1}^N c_k K(x, x_k), \quad x \in \mathcal{D},
\]

where the coefficients \( c := (c_1, \ldots, c_N)^T \) are obtained by solving the following system of linear equations:

\[
Kc = y_0,
\]

with \( K := (K(x_j, x_k))_{j,k=1}^{N,N} \) and \( y_0 := (y_1, \ldots, y_N)^T \).

For simple kriging, i.e. in the stochastic formulation, we let \( S \) be a Gaussian process with mean 0 and covariance kernel \( K \) on some probability space \( \mathcal{D} \). Kriging is based on the modelling assumption that \( u \) is a realization of the Gaussian field \( S \). The data values \( y_1, \ldots, y_N \) are then realizations of the random variables \( S_{x_1}, \ldots, S_{x_N} \). The optimal unbiased predictor of \( S_x \) based on \( S \) is equal to

\[
\hat{U}_x := \sum_{k=1}^N c(x) S_{x_k} = \arg\min_{U \in \text{span}(S_{x_k})} \mathbb{E}|U - S_x|^2,
\]

where the coefficients \( \mathbf{c}(x) := (c_1(x), \ldots, c_N(x))^T \) are given by

\[
\mathbf{c}(x) = K^{-1} \mathbf{k}(x),
\]

with \( \mathbf{k}(x) := (K(x, x_1), \ldots, K(x, x_N))^T \) and the same matrix \( K \) as above. We can also compute that

\[
\mathbb{E}(\hat{U}_x|S_{x_1} = y_1, \ldots, S_{x_N} = y_N) = \hat{u}_K(x).
\]

Note that, in the kriging approach we consider only the values of the stochastic process \( S \) at the collocation points, and view the obtained vector as a random variable. However, if we view \( S \) as a real function, then \( \mathbb{P}(S \in H_K(D)) = 0 \) by [16, Theorem 7.3]. A simple example for this fact is given by the scalar Brownian motion defined in the domain \( \mathcal{D} := (0,1) \) (see, e.g. [11, Example 5.1]). This means that it is difficult to apply the kriging formulation to PDE problems. Next we will introduce a new stochastic data-fitting approach that will subsequently allow us to perform kernel-based collocation for SPDEs.
2.2 Data-fitting problems via a new stochastic approach

From now on we will view the separable reproducing-kernel Hilbert space $H_k(D)$ as a sample space and its Borel $\sigma$-field $\mathcal{B}(H_k(D))$ as a $\sigma$-algebra to set up the probability spaces so that the stochastic process $S_\omega(x) := \omega(x)$ is Gaussian. We use the techniques of [13,16] to verify Lemma 2.2, which is a restatement of [16, Theorem 7.2]. This theoretical result is a generalized form of Wiener measure defined on the measurable space $(\mathbb{C}[0, \infty), \mathcal{B}(\mathbb{C}[0, \infty)))$, called canonical space, such that the coordinate mapping process $W_\omega(x) := \omega(x)$ is a Brownian motion (see, for instance, [15, Chapter 2]).

Lemma 2.2 Let the positive-definite kernel $K \in \mathcal{C}(\Omega \times \Omega)$ be the reproducing kernel of the reproducing-kernel Hilbert space $H_k(D)$. Given a function $\mu \in H_k(D)$ there exists a probability measure $\mathbb{P}^\mu$ defined on $(\Omega_k, \mathcal{F}_\xi, \mathbb{P}_\xi)$, i.e. $\xi_k \sim \text{i.i.d. } \mathcal{N}(0, 1)$, $k \in \mathbb{N}$, $(\lambda_k)^\infty_{k=1}$ and $(\xi_k)^\infty_{k=1}$ are independent Gaussian random variables with mean $\mu_k := \langle \mu, \sqrt{\lambda_k} \xi_k \rangle_{D, k}$ and variance $\lambda_k, k \in \mathbb{N}$.

Before we prove Lemma 2.2 we remark that we have introduced the integral-type kernel $K^*$ for convenience only. As seen later, in order to ‘match the spaces’, any other kernel that ‘dominates’ $K$ (in the sense of [16]) could play the role of the integral-type kernel $K^*$.

Proof We first consider the case when $\mu = 0$. There exist countably many independent standard normal random variables $\{\xi_k\}^\infty_{k=1}$ on a probability space $(\Omega_k, \mathcal{F}_k, \mathbb{P}_k)$, i.e. $\xi_k \sim \text{i.i.d. } \mathcal{N}(0, 1)$, $k \in \mathbb{N}$. Let $\{\lambda_k\}^\infty_{k=1}$ and $(\xi_k)^\infty_{k=1}$ be the eigenvalues and eigenfunctions of the reproducing kernel $K$ as in Theorem A.3. We define $S := \sum_{k=1}^\infty \xi_k \lambda_k e_k \mathbb{P}_k$-a.s. Note that $S$ is Gaussian with mean 0 and covariance kernel $K$. Since $\mathbb{E}(\sum_{k=1}^\infty \xi_k^2 \lambda_k) \leq \sum_{k=1}^\infty \text{Var}(\xi_k) \lambda_k = \sum_{k=1}^\infty \lambda_k < \infty$ indicates that $\sum_{k=1}^\infty |\xi_k| \lambda_k^2 < \infty \mathbb{P}_k$-a.s., Theorem A.3 shows that $S(\cdot, \omega) \in H_k(D)$ $\mathbb{P}_k$-a.s. Therefore, $S$ is a measurable map from $(\Omega_k, \mathcal{F}_k)$ into $(\Omega_k, \mathcal{F}_k)$ by [3, Chapter 4.3.1] and [16, Lemma 2.1]. On the other hand, the probability measure $\mathbb{P}^0 := \mathbb{P}_k \circ S^{-1}$ (also called the image measure of $\mathbb{P}_k$ under $S$) is well defined on $(\Omega_k, \mathcal{F}_k)$, i.e. $\mathbb{P}^0(A) := \mathbb{P}_k(S^{-1}(A))$ for each $A \in \mathcal{F}_k$. Hence, $S$ is also a Gaussian process with mean $\mu$ and covariance kernel $K^*$ on $(\Omega_k, \mathcal{F}_k, \mathbb{P}^0)$.

Let $S^\mu := S + \mu$ on $(\Omega_k, \mathcal{F}_k, \mathbb{P}^0)$. Then $\mathbb{E}(S_\omega^\mu) = \mathbb{E}(S_\omega) + \mu(x)$ and $\text{Cov}(S_\omega^\mu, S_y^\mu) = \text{Cov}(S_\omega, S_y)$ with respect to $\mathbb{P}^0$. We define a new probability measure $\mathbb{P}^\mu$ by $\mathbb{P}^\mu (A) := \mathbb{P}^0 (A - \mu)$ for each $A \in \mathcal{F}_k$. It is easy to check that $\Omega_k + \mu = H_k(\Omega) = \Omega_k$ and $\{\mu + A : A \in \mathcal{F}_k\} = \mathcal{B}(H_k(\Omega)) = \mathcal{F}_k$. Thus, $S$ is Gaussian with mean $\mu$ and covariance kernel $K^*$ on $(\Omega_k, \mathcal{F}_k, \mathbb{P}^\mu)$.

Moreover, since $\mu \in H_k(\Omega)$, it can be expanded in the form $\mu = \sum_{k=1}^\infty \mu_k \sqrt{\lambda_k} e_k$, where $\mu_k := \langle \mu, \sqrt{\lambda_k} \xi_k \rangle_{k, D}$, so that $S_\omega^\mu = \sum_{k=1}^\infty (\mu_k + \sqrt{\lambda_k} \xi_k) \sqrt{\lambda_k} e_k$. But then $\xi_k \sim \mu_k + \sqrt{\lambda_k} \xi_k \sim \mathcal{N}(\mu_k, \lambda_k)$ are independent on $(\Omega_k, \mathcal{F}_k, \mathbb{P}^\mu)$. •
According to [3, Theorem 4.91], we can also verify that the random variable $V_f(\omega) := \langle \omega, f \rangle_{K,D}, f \in H_K(D)$, is a scalar normal variable on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$, i.e.,

$$V_f(\omega) \sim \mathcal{N}(m_f, \sigma_f^2), \quad \omega \in \Omega_K = H_K(D),$$

where $m_f := \langle \mu, f \rangle_{K,D}$ and $\sigma_f := \|f\|_{L_2(D)}$. Therefore, the probability measure $\mathbb{P}^\mu$ defined in Lemma 2.2 is Gaussian.

Let $p^K_X : \mathbb{R}^N \to \mathbb{R}$ be the joint probability density function of $S_{x_1}, \ldots, S_{x_N}$ defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$. Then it is a normal density function with mean $\mu := (\mu(x_1), \ldots, \mu(x_N))^T$ and covariance matrix $K := (K(x_j, x_k))_{1 \leq j, k \leq N}$. In analogy to the kriging formulation, we can find the optimal mean function $\hat{\mu} = \arg\max_{\mu \in \mathcal{H}_K(D)} \mathbb{P}^\mu(S = y_0)$, i.e.,

$$\hat{\mu} := \hat{k}^T K^{-1} y_0 = \sup_{\mu \in \mathcal{H}_K(D)} p^K_X(y_0) = \sup_{\mu \in \mathcal{H}_K(D)} \mathbb{P}^\mu(S = y_0),$$

where $\hat{k}(x) := (\hat{k}(x, x_1), \ldots, \hat{k}(x, x_N))^T$.

We now fix any $x \in D$. Straightforward calculation shows that the random variable $S_x$, given $S_{x_1}, \ldots, S_{x_N}$, defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$ has a conditional probability density function

$$p^K_y(v|y) := \frac{1}{\sigma(x) \sqrt{2\pi}} \exp\left(-\frac{(v - m^K_y(v))^2}{2\sigma(x)^2}\right), \quad v \in \mathbb{R}, \quad y \in \mathbb{R}^N,$$

where $m^K_y(v) := \mu(x) + \hat{k}(x)^T K^{-1} (v - m^K)$, $m^K := (\mu(x_1), \ldots, \mu(x_N))^T$, and $\sigma(x)^2 := K(x, x) - \hat{k}(x)^T K^{-1} \hat{k}(x)$. Then the optimal estimator that maximizes the probability

$$\max_{v \in \mathbb{R}} \mathbb{P}^{\hat{\mu}}(S_x = v | S_{x_1} = y_1, \ldots, S_{x_N} = y_N)$$

is given by

$$\hat{u}(x) := \hat{\mu}(x) = \arg\max_{v \in \mathbb{R}} p^K_y(v|y_0).$$

**Proposition 2.3** With the above notations, the following equality holds true:

$$p^K_y(\hat{u}(x)|y_0) = \sup_{v \in \mathbb{R}} p^K_y(v|y_0). \quad (7)$$

Moreover, for any $\epsilon > 0$,

$$\sup_{\mu \in \mathcal{H}_K(D)} \mathbb{P}^\mu(|\hat{u}(x) - u(x)| \geq \epsilon) \leq \sup_{\mu \in \mathcal{H}_K(D)} \mathbb{P}^\mu(\mathcal{E}^x_\epsilon) = \text{erfc}\left(\frac{\epsilon}{\sqrt{2\sigma(x)}}\right), \quad (8)$$

where

$$\mathcal{E}^x_\epsilon := \left\{ \omega \in \Omega_K : |\omega(x) - \hat{u}(x)| \geq \epsilon \text{ s.t. } \omega(x_1) = y_1, \ldots, \omega(x_N) = y_N \right\}.$$  

Identity (7) follows by direct evaluations. Consequently, taking into account that $S$ is Gaussian, inequality (8) follows also immediately.

**Remark 1** Instead of giving a deterministic (or strong) error bound for the proposed numerical scheme, we provide a weak type convergence of the approximated solution $\hat{u}$ to the true solution $u$, as stated in Proposition 2.3. In fact, inequality (8) can be seen as a confidence interval for the estimator $\hat{u}$ with respect to the probability measure $\mathbb{P}^\mu$.

In the next section, we generalize this stochastic approach to solve elliptic PDEs and SPDEs.
3. Collocation method for elliptic PDEs and SPDEs

We begin by setting up Gaussian processes via reproducing kernels with differential and boundary operators. Suppose that the reproducing-kernel Hilbert space $\mathcal{H}_K(D)$ is embedded into the Sobolev space $\mathcal{H}^m(D)$ with $m > d/2$. Let $n := [m - d/2] - 1$. By the Sobolev embedding theorem $\mathcal{H}^m(D) \subset C^n(D)$. When the differential operator $P$ and the boundary operator $B$ have the orders $\mathcal{O}(P) < m - d/2$ and $\mathcal{O}(B) < m - d/2$, then the stochastic processes $PS_x(\omega) := (P\omega)(x)$ and $BS_x(\omega) := (B\omega)(x)$ are well defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$. According to Lemma A.6, we have $PS = \sum_{k=1}^\infty \xi_k \sqrt{\lambda_k} P e_k$ and $BS = \sum_{k=1}^\infty \xi_k \sqrt{\lambda_k} B e_k$. If $\mu \in \mathcal{H}_K(D) \subset \mathcal{H}^m(D)$, then $P\mu \in C(\partial D)$ and $B\mu \in C(\partial D)$. Lemma A.7 implies that $P_1 P_2^* K(x, y) = \sum_{k=1}^\infty \lambda_k^2 P e_k(x) P e_k(y)$ and $B_1 B_2^* K(x, y) = \sum_{k=1}^\infty \lambda_k^2 B e_k(x) B e_k(y)$ (here we can use the fact that $\text{Cov}(PS_x, PS_y) = P_2 P_1 \text{Cov}(S_x, S_y)$ and $\text{Cov}(BS_x, BS_y) = B_2 B_1 \text{Cov}(S_x, S_y)$). Applying Lemma 2.2, we can obtain the main theorem for the construction of Gaussian processes via reproducing kernels coupled with differential or boundary operators.

**Theorem 3.1** Suppose that the reproducing-kernel Hilbert space $\mathcal{H}_K(D)$ is embedded into the Sobolev space $\mathcal{H}^m(D)$ with $m > d/2$. Further assume that the differential operator $P$ and the boundary operator $B$ have the orders $\mathcal{O}(P) < m - d/2$ and $\mathcal{O}(B) < m - d/2$. Given a function $\mu \in \mathcal{H}_K(D)$ there exists a probability measure $\mathbb{P}^\mu$ defined on $(\Omega_K, \mathcal{F}_K) = (H_K(D), \mathcal{B}(H_K(D)))$ (as in Lemma 2.2) such that the stochastic processes $PS, BS$ given by

$$PS_x(\omega) = PS(x, \omega) := (P\omega)(x), \quad x \in D \subset \mathbb{R}^d, \quad \omega \in \Omega_K = H_K(D),$$

$$BS_x(\omega) = BS(x, \omega) := (B\omega)(x), \quad x \in \partial D, \quad \omega \in \Omega_K = H_K(D),$$

are jointly Gaussian processes with means $P\mu, B\mu$ and covariance kernels $P_1 P_2^*, B_1 B_2^*$ defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$, respectively. In particular, they can be expanded as

$$PS_x = \sum_{k=1}^\infty \xi_k \sqrt{\lambda_k} P e_k(x), \quad x \in D \quad \text{and} \quad BS_x = \sum_{k=1}^\infty \xi_k \sqrt{\lambda_k} B e_k(x), \quad x \in \partial D, \quad \mathbb{P}^\mu\text{-a.s.},$$

where $\{\lambda_k\}_{k=1}^\infty$ and $\{e_k\}_{k=1}^\infty$ are the eigenvalues and eigenfunctions of the reproducing kernel $K$ and their related Fourier coefficients are the independent normal variables $\xi_k \sim N(\mu_k, \lambda_k)$ and $\hat{\mu}_k := \langle \mu, \sqrt{\lambda_k} e_k \rangle_{K, D}$, $k \in \mathbb{N}$.

**Corollary 3.2** Suppose all notations and conditions are as in Theorem 3.1. Given collocation points $X_D := \{x_j\}_{j=1}^N \subset D$ and $X_{\partial D} := \{x_{N+j}\}_{j=1}^M \subset \partial D$, the random vector $S_{PB} := (PS_{x_1}, \ldots, PS_{x_N}, BS_{x_{N+1}}, \ldots, BS_{x_{N+M}})^\top$ defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$ has a multi-normal distribution with mean $m_{PB}^\mu$ and covariance matrix $K_{PB}$, i.e.

$$S_{PB} \sim N(m_{PB}^\mu, K_{PB}),$$

where $m_{PB}^\mu := (P\mu(x_1), \ldots, P\mu(x_N), B\mu(x_{N+1}), \ldots, B\mu(x_{N+M}))^\top$ and

$$K_{PB} := \begin{bmatrix}
(P_1 P_2^* K(x_j, x_k))_{j,k=1}^{N,N} & (P_1 B_2^* K(x_j, x_{N+k}))_{j,k=1}^{N,M} \\
(B_1 P_2^* K(x_{N+j}, x_k))_{j,k=1}^{M,N} & (B_1 B_2^* K(x_{N+j}, x_{N+k}))_{j,k=1}^{M,M}
\end{bmatrix}.$$
Remark 2 While the covariance matrix $K_{PB}$ may be singular, it is always positive semi-definite and therefore always has a pseudo-inverse $K_{PB}^+$ denoted with a superscript “$^+$” instead of with a superscript “$-$”. 

Using Corollary 3.2, we can compute the joint probability density function $p^\mu_x$ of $S_{PB}$ defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$. In the same way, we can also get the joint density function $p^\mu_j$ of $(S_x, S_{PB})$ defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$. By Bayes’ rule, we can obtain the conditional probability density function of the random variable $S_x$ given $S_{PB}$.

Corollary 3.3 We follow the notations of Corollary 3.2. For any fixed $x \in \mathcal{D}$, the random variable $S_x$ given $S_{PB}$ defined on $(\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu)$ has a conditional probability density function

$$p^\mu_x(v|x) := \frac{p^\mu_j(v, y)}{p^\mu_x(v)} = \frac{1}{\sigma(x)\sqrt{2\pi}} \exp \left(-\frac{(v - m^\mu_x(v))^2}{2\sigma^2(x)}\right), \quad v \in \mathbb{R}, \quad v \in \mathbb{R}^{N+M},$$

where

$$m^\mu_x(v) := \mu(x) + k_{PB}(x)^T K_{PB}^+ (v - m^\mu_{PB}),$$

$$\sigma(x)^2 := K(x, x) - k_{PB}(x)^T K_{PB}^+ k_{PB}(x),$$

$$k_{PB}(x) := (P_2K(x, x_1), \ldots, P_2K(x, x_N), P_2K(x, x_{N+1}), \ldots, B_2K(x, x_{N+M}))^T.$$

In particular, given the real observation $y := (y_1, y_{N+M})^T$, $S_x$ conditioned on $S_{PB} = y$ has the probability density $p^\mu_x(\cdot|y)$.

This corollary is similar to the features of Gaussian conditional distributions (see [13, Theorem 9.9]).

3.1 Elliptic deterministic PDEs

Suppose that $u \in H_K(\mathcal{D})$ is the unique solution of the deterministic elliptic PDE

$$
\begin{align*}
Pu &= f \quad \text{in} \; \mathcal{D} \subset \mathbb{R}^d, \\
Bu &= g \quad \text{on} \; \partial \mathcal{D},
\end{align*}
$$

(9)

where $f : \mathcal{D} \to \mathbb{R}$ and $g : \partial \mathcal{D} \to \mathbb{R}$. Denote by $y_j = (y_j)_{j=1}^N$ and $(y_{N+k})_{k=1}^M$ the values of $f$ and $g$ at the collocation points $X_\mathcal{D}$ and $X_{3\mathcal{D}}$, respectively:

$$y_j := f(x_j), \quad y_{N+k} := g(x_{N+k}), \quad j = 1, \ldots, N, \quad k = 1, \ldots, M.$$

From now on we assume that the covariance matrix $K_{PB}$ defined in Corollary 3.2 is nonsingular and we therefore can replace pseudo-inverses with inverses.

Let $y_0 := (y_1, \ldots, y_N, y_{N+1}, \ldots, y_{N+M})^T$, and denote by $p^\mu_x(\cdot|\cdot)$ the conditional density function defined in Corollary 3.3. We approximate the solution $u$ of (9) by the optimal estimator $\hat{u}(x)$ derived in the previous section, i.e. we maximize the conditional probability given the data values $y_0$:

$$u(x) \approx \hat{u}(x) = \arg\max_{\mu \in H_\mu(\mathcal{D})} \sup_{\nu \in \mathbb{R}} p^\mu_x(\nu|y_0), \quad x \in \mathcal{D}.$$
By direct evaluation as in Section 2.2 one finds that
\[
\hat{u}(x) := k_{PB}(x)^T K_{PB}^{-1} y_0, \quad x \in \mathcal{D},
\]
where the basis functions \( k_{PB}(x) \) are defined in Corollary 3.3. Moreover, the estimator \( \hat{u} \in H_K(\mathcal{D}) \) fits all the data values: \( \hat{P}u(x_1) = y_1, \ldots, \hat{P}u(x_N) = y_N \) and \( \hat{B}u(x_{N+1}) = y_{N+1}, \ldots, \hat{B}u(x_{N+M}) = y_{N+M} \). This means that we have computed a collocation solution of the PDE \((9)\). Also note that \( \hat{u} \) can be written as a linear combination of the kernels as in \((4)\), i.e.
\[
\hat{u}(x) = \sum_{k=1}^N c_k P_2^*(x,x_k) + \sum_{k=1}^M c_{N+k} B_2^*(x,x_{N+k}), \quad x \in \mathcal{D},
\]
where \( \epsilon := (c_1, \ldots, c_{N+M})^T = K_{PB}^{-1} y_0 \in \mathbb{R}^{N+M} \).

Finally, we can perform a weak error analysis for \( |u(x) - \hat{u}(x)| \) as in Proposition 2.3, and deduce that
\[
\mathbb{P}^\mu(\mathcal{E}_x^\epsilon) = \mathbb{P}^\mu(|S_x - \hat{u}(x)| \geq \epsilon | S_{PB} = y_0) = \text{erfc} \left( \frac{\epsilon}{\sqrt{2} \sigma(x)} \right),
\]
where \( \sigma(x)^2 \) is defined in Corollary 3.3, and
\[
\mathcal{E}_x^\epsilon := \{ \omega \in \Omega_K : |\omega(x) - \hat{u}(x)| \geq \epsilon \text{ s.t. } P\omega(x_1) = y_1, \ldots, B\omega(x_{N+M}) = y_{N+M} \}.
\]

Because the form of the expression for the variance \( \sigma(x)^2 \) is analogous to that of the power function, we can use the same techniques as in the proofs from [8,21] to obtain a formula for the order of \( \sigma(x) \).

**Lemma 3.4**
\[
\sigma(x) = \mathcal{O}(h_X^{m-\rho-d/2}), \quad x \in \mathcal{D},
\]
where \( \rho := \max\{\|P\|, \|B\|\} \) and \( h_X \) is the fill distance of \( X_\mathcal{D} \) and \( x_\partial \).

**Proof** Since there is at least one collocation point \( x_j \in X_\mathcal{D} \cup X_\partial \) such that \( \|x - x_j\|_2 \leq h_X \) we can use the multivariate Taylor expansion of \( K(x,x_j) \) to introduce the order of \( \sigma(x) \), i.e.
\[
\hat{K}(x,x_j) = \sum_{|\alpha|,|\beta|<n} \frac{1}{\alpha!\beta!} D_1^{\alpha} D_2^{\beta} K(x,x_j)(x-x_j)^{\alpha+\beta} + R(x,x_j), \quad \alpha, \beta \in \mathbb{N}_0^d,
\]
where \( R(x,x_j) := \sum_{|\alpha|,|\beta|=n} (1/\alpha!\beta!) D_1^{\alpha} D_2^{\beta} K(z_1,z_2)(x-x_j)^{\alpha+\beta} \) for some \( z_1, z_2 \in \mathcal{D} \) and \( n := \lfloor m - d/2 \rfloor - 1 \). The rest of the proof proceeds as in [8, Chapter 14.5] and [21, Chapters 11.3, 16.3].

Using Lemma 3.4 we can deduce the following proposition.

**Proposition 3.5** For any \( \epsilon > 0 \),
\[
\sup_{\mu \in H_K(\mathcal{D})} \mathbb{P}^\mu(\mathcal{E}_x^\epsilon) = \mathcal{O}\left( \frac{h_X^{m-\rho-d/2}}{\epsilon} \right), \quad x \in \mathcal{D},
\]
which indicates that
\[
\sup_{\mu \in H_K(\mathcal{D})} \mathbb{P}^\mu(|u - \hat{u}|_{L^\infty(\mathcal{D})} \geq \epsilon) \leq \sup_{\mu \in H_K(\mathcal{D}), x \in \mathcal{D}} \mathbb{P}^\mu(\mathcal{E}_x^\epsilon) \to 0, \text{ when } h_X \to 0.
\]
Therefore, we say that the estimator \( \hat{u} \) converges to the exact solution \( u \) of the PDE (9) in all probabilities \( \mathbb{P}^\mu \) when \( h_X \) goes to 0.

Sometimes we know only that the solution \( u \in \mathcal{H}^m(\mathcal{D}) \). In this case, as long as the reproducing-kernel Hilbert space is dense in the Sobolev space \( \mathcal{H}^m(\mathcal{D}) \) with respect to its Sobolev norm, we can still say that \( \hat{u} \) converges to \( u \) in probability.

### 3.2 Elliptic SPDEs

Let \( \xi : \mathcal{D} \times \Omega_W \to \mathbb{R} \) be Gaussian with mean 0 and covariance kernel \( \Psi : \mathcal{D} \times \mathcal{D} \to \mathbb{R} \) on the probability space \( (\Omega_W, \mathcal{F}_W, \mathbb{P}_W) \). We consider an elliptic PDE driven by a Gaussian additive noise \( \xi \)

\[
\begin{cases}
Pu = f + \xi & \text{in } \mathcal{D} \subset \mathbb{R}^d, \\
Bu = g & \text{on } \partial \mathcal{D},
\end{cases}
\] (11)

and suppose its solution \( u \in L_2(\Omega_W; \mathcal{H}^m(\mathcal{D})) \).

Since \( \xi \) is a Gaussian process, on some underlying probability space \( (\Omega_W, \mathcal{F}_W, \mathbb{P}_W) \) with a known correlation structure, we can simulate the values of \( \xi \) at \( x_j, j = 1, \ldots, N \). Consequently, we assume that the values \( \{y_j\}_{j=1}^N \) and \( \{y_{N+k}\}_{k=1}^M \) defined by

\[ y_j := f(x_j) + \xi, \quad y_{N+k} := g(x_{N+k}), \quad j = 1, \ldots, N, \quad k = 1, \ldots, M, \]

are known.

In order to apply the general interpolation framework developed in Section 2.2, we consider the product space

\[ \Omega_{KW} := \Omega_K \times \Omega_W, \quad \mathcal{F}_{KW} := \mathcal{F}_K \otimes \mathcal{F}_W, \quad \mathbb{P}^\mu_W := \mathbb{P}^\mu \otimes \mathbb{P}_W. \]

We assume that the random variables defined on the original probability spaces are extended to random variables on the new probability space in the natural way: if random variables \( V_1 : \Omega_K \to \mathbb{R} \) and \( V_2 : \Omega_W \to \mathbb{R} \) are defined on \( (\Omega_K, \mathcal{F}_K, \mathbb{P}^\mu) \) and \( (\Omega_W, \mathcal{F}_W, \mathbb{P}_W) \), respectively, then

\[ V_1(\omega, \tilde{\omega}) := V_1(\omega), \quad V_2(\omega, \tilde{\omega}) := V_2(\tilde{\omega}) \quad \text{for each } \omega \in \Omega_K \text{ and } \tilde{\omega} \in \Omega_W. \]

Note that in this case the random variables have the same probability distributional properties, and they are independent on \( (\Omega_{KW}, \mathcal{F}_{KW}, \mathbb{P}^\mu_W) \). This implies that the stochastic processes \( S, PS, BS \) and \( \xi \) can be extended to the product space \( (\Omega_{KW}, \mathcal{F}_{KW}, \mathbb{P}^\mu_W) \) while preserving the original probability distributional properties. In this case, \( (y_1, \ldots, y_N) \sim \mathcal{N}(f, \Psi) \), where \( f := (f(x_1), \ldots, f(x_N))^T \) and \( \Psi := (\Psi(x_j, x_k))_{j,k=1}^{N,N} \) with \( \Psi \) being the covariance kernel of \( \xi \). According to Corollary 3.3 we can find the conditional probability density function \( p^\mu_\xi(\cdot|\cdot) \), and the optimal estimator \( \hat{u}(x) \)

\[ \hat{u}(x) = \arg\max_{v \in \mathbb{R}} \sup_{\mu \in \mathcal{H}_K(\mathcal{D})} p^\mu_\xi(v|y_\xi)p_W(y_\xi), \quad x \in \mathcal{D}, \]

where

\[ y_\xi := (y_1, \ldots, y_{N+M})^T, \quad m_\xi := \left( \begin{array}{c} f \\ 0 \end{array} \right) \in \mathbb{R}^{N+M}, \quad \Sigma := \left( \begin{array}{cc} \Psi & 0 \\ 0 & 0 \end{array} \right) \in \mathbb{R}^{(N+M) \times (N+M)}, \]

\[ p_W(v) := \frac{1}{\sqrt{\det(\Sigma)(2\pi)^{(N+M)/2}}} \exp\left( -\frac{1}{2}(v - m_\xi)^T \Sigma^+ (v - m_\xi) \right), \quad v \in \mathbb{R}^{N+M}. \]
We consider the following stochastic heat equation with zero boundary condition:

\[ W_t \]

driven by two types of space–time white noise (colored in space)

\[ \text{if} \]

Note that choosing the larger value of \( i \)

Similar to the analysis of the error bounds from Section 3.1, we also deduce the following proposition (for more details see [23]).

**Proposition 3.6** Assume \( \Psi \in C^1(\mathcal{D} \times \mathcal{D}) \). Then

\[
\lim_{E \to 0} \sup_{x \in \mathcal{D}} \mathbb{P}_W (|u - \hat{u}| \leq \epsilon) = 0 \quad \text{for any } \epsilon > 0.
\]

**4. Numerical experiments**

We consider the following stochastic heat equation with zero boundary condition:

\[ \begin{align*}
\frac{dU_t}{dt} &= \frac{d^2}{dx^2} U_t + \sigma \, dW_{t,i} \quad \text{in } \mathcal{D} := (0, 1) \subset \mathbb{R}, \quad 0 < t < T := 1, \\
U_t &= 0 \quad \text{on } \partial \mathcal{D}, \\
U_0(x) &= u_0(x) := \sqrt{2} (\sin(\pi x) + \sin(2\pi x) + \sin(3\pi x)),
\end{align*} \tag{12} \]

driven by two types of space–time white noise (colored in space) \( W_t \) of the form

\[ W_{t,i} := \sum_{k=1}^{\infty} W^k \phi(t, \phi_k), \quad q_k := \frac{1}{k^2}, \quad \phi_k(x) := \sqrt{2} \sin(k\pi x), \]

where \( W^k, k \in \mathbb{N} \), is a sequence of independent one-dimensional Brownian motions, and \( i = 1, 2 \). Note that choosing the larger value of \( i \) corresponds to a noise that is smoother in space.

The spatial covariance function \( R^i(x, y) = \sum_{k=1}^{\infty} q^2_k \phi_k(x) \phi_k(y), \quad i = 1, 2 \), takes the specific forms

\[ R^1(x, y) = \min\{x, y\} - xy, \quad 0 < x, y < 1, \]

and

\[ R^2(x, y) = \begin{cases} 
- \frac{1}{6} x^3 + \frac{1}{6} x^3 y + \frac{1}{6} x y^3 - \frac{1}{2} x y^2 + \frac{1}{3} x y, & 0 < x < y < 1, \\
- \frac{1}{6} x^3 + \frac{1}{6} x^3 y + \frac{1}{6} x y^3 - \frac{1}{2} x^2 y + \frac{1}{3} x y, & 0 < y < x < 1.
\end{cases} \]

The solution of SPDE (12) is given by (for more details see, for instance, [5])

\[ U_t(x) = \sum_{k=1}^{\infty} \xi^k_t \phi_k(x), \quad x \in \mathcal{D} := (0, 1), \quad 0 < t < T := 1, \]

where

\[ \xi^k_0 := \int_{\mathcal{D}} u_0(x) \phi_k(x) \, dx, \quad \xi^k_t := \xi^k_0 e^{-\lambda^2 t} + \frac{\sigma}{q^2_k} \int_0^t e^{-\lambda^2 s} \, dW^k_s. \]

From this explicit solution we can get that

\[ \mathbb{E}(U_t(x)) = \sum_{k=1}^{\infty} \xi^k_0 e^{-\lambda^2 t} \phi_k(x), \quad \text{Var}(U_t(x)) = \sum_{k=1}^{\infty} \frac{\sigma^2}{2k^2 \pi^2 q^2_k} (1 - e^{-2\lambda^2 t}) |\phi_k(x)|^2. \]
We discretize the time interval $[0, T]$ with $n$ equal time steps so that $\delta t := T/n$. We also choose the reproducing kernel $K(x, y) := g_{3.2\theta}(x - y)$, where $g_{3.2\theta}$ is the Matérn function with degree $m := 3$ and shape parameter $\theta > 0$ (see Example A.4). As collocation points we select uniform grid points $X_{\mathcal{D}} \subset (0, 1)$ and $X_{\mathcal{D}} := [0, 1]$. Let $P := I - \delta \frac{d^2}{dx^2}$ and $B := |I_{[0,1]}$. Using our kernel-based collocation method, we can perform the following computations to numerically estimate the sample paths $\hat{u}_j^0 \approx U_n(x_j)$. Algorithm to solve SPDE (12):

(1) Initialize

- $\hat{u}^0 := (u_0(x_1), \ldots, u_0(x_N))^T$,
- $K_{PB} := \left( (P_1 P_2 K(x_j, x_k))_{j,k=1}^{N,N}, (P_1 B_2 K(x_j, x_{N+k}))_{j,k=1}^{N,M} \right)$,
- $B := \left( (P_2 K(x_j, x_k))_{j,k=1}^{N,N}, (B_2 K(x_j, x_{N+k}))_{j,k=1}^{N,M} \right)$,
- $\Psi := \sigma^2 \delta t (R(x_j, x_k))_{j,k=1}^{N,N}$, $\Sigma := \begin{pmatrix} \Psi & 0 \\ 0 & 0 \end{pmatrix}$,
- $A := B(K_{PB} + \Sigma)^{-1}$.

(2) Repeat for $j = 1, 2, \ldots, n$, i.e. for $t_1, t_2, \ldots, t_n = T$

- Simulate $\xi \sim \mathcal{N}(0, \Psi)$.
- $\hat{u}^j := B(K_{PB} + \Sigma)^{-1} \begin{pmatrix} \hat{u}^{j-1} + \xi \\ 0 \end{pmatrix} = A \begin{pmatrix} \hat{u}^{j-1} + \xi \\ 0 \end{pmatrix}$.

Note that in the very last step the matrix $A$ is pre-computed and can be used for all time steps, and for different sample paths; that makes the proposed algorithm to be quite efficient.

We approximate the mean and variance of $U_n(x)$ by sample mean and sample variance from $s := 10,000$ simulated sample paths using the above algorithm, i.e.

$$\mathbb{E}(U_n(x_j)) \approx \frac{1}{s} \sum_{k=1}^{s} \hat{u}_j^n(\omega_k), \quad \text{Var}(U_n(x_j)) \approx \frac{1}{s} \sum_{i=1}^{s} \left( \hat{u}_j^n(\omega_i) - \frac{1}{s} \sum_{k=1}^{s} \hat{u}_j^n(\omega_k) \right)^2.$$ 

Figure 1 shows that the histograms at different values of $t$ and $x$ resemble the theoretical normal distributions. We notice a small shift in the probability distribution function of the solution $U$, at

![Graphs showing empirical and theoretical probability distribution](image_url)
times closer to zero, and when the noise is equal to \( W_1 \) (Figure 1, left panel). This shift is due to the fact that \( W_1 \) is rougher in space than \( W_2 \).

Our use of an implicit time-stepping scheme reduces the frequency of the white noise, i.e. \( \lim_{\delta t \to 0} \delta W / \delta t \sim \delta_0 \). Consequently, Figure 2 shows that the approximate mean is well behaved but the approximate variance is a little smaller than the exact variance. According to Figure 3 we find that this numerical method is convergent as both \( \delta t \) and \( h_X \) are refined. Finally, we want to mention that the distribution of collocation points, the shape parameter, and the kernel itself were chosen empirically and based on the authors’ experience. As mentioned before, more precise methods are currently not available. A rigorous investigation of these questions, as well as determination of precise rates of convergence, is reserved for future work.

5. Final remarks

This new numerical approach can also be used to approximate systems of elliptic PDEs with vector Gaussian noises \( \xi_1 \) and \( \xi_2 \) or nonlinear PDEs with Gaussian noise \( \xi \), i.e.

\[
\begin{cases}
\mathbf{Pu} = f + \xi_1 & \text{in } D \subset \mathbb{R}^d, \\
\mathbf{Bu} = g + \xi_2 & \text{on } \partial D,
\end{cases}
\quad \text{or} \quad \begin{cases}
F(\mathbf{Pu}) = \psi(f, \xi) & \text{in } D \subset \mathbb{R}^d, \\
G(\mathbf{Bu}) = g & \text{on } \partial D,
\end{cases}
\]

where \( \mathbf{P} := (P^1, \ldots, P^{n_p})^T \) is a vector differential operator and \( \mathbf{B} := (B^1, \ldots, B^{n_b})^T \) is a vector boundary operator, and \( F : \mathbb{R}^{n_p} \to \mathbb{R} \) and \( G : \mathbb{R}^{n_b} \to \mathbb{R} \) (see [23]).

In addition to the additive noise case discussed here, we can also use the kernel-based collocation method to approximate other well-posed stochastic parabolic equations with multiplicative noise, e.g.

\[
\begin{cases}
\frac{dU_t}{dt} = A U_t + \psi(U_t) \, dW_t & \text{in } D \subset \mathbb{R}^d, \quad 0 < t < T, \\
B U_t = 0 & \text{on } \partial D, \\
U_0 = u_0,
\end{cases}
\]

(13)

where \( \psi : \mathbb{R} \to \mathbb{R} \). Since \( \int_{t_{i-1}}^{t_i} \psi(U_s) \, dW_s \approx \psi(U_{t_{i-1}}) \delta W_i \), the algorithm for SPDE (13) is similar to before:
Figure 3. Convergence of mean and variance with respect to refinement of points and time steps for $\sigma := 1$. (The relative RMSE of exact $U$ and approximate $\hat{U}$ is defined by $\text{RMSE}(U, \hat{U}) := (1/nN) \sum_{j=1}^{n} \sum_{k=1}^{N} (U(t_j, x_k) - \hat{U}(t_j, x_k))^2 / \|U(t_k, \cdot)\|_2^2$. (a) With spatial covariance $R_1$ and (b) with spatial covariance $R_2$.)

1. Initialize
   - $\hat{u}_0 \hat{u}_0 := (u_0(x_1), \ldots, u_0(x_N))^T$,
   - $\mathbf{K}_{PB} := \begin{pmatrix} (P_1 P_2 \hat{K}(x_j, x_k))_{j,k=1}^{N,N} & (P_1 B_2 \hat{K}(x_j, x_{N+k}))_{j,k=1}^{N,M} \\ (B_1 P_2 \hat{K}(x_{N+j}, x_k))_{j,k=1}^{M,N} & (B_1 B_2 \hat{K}(x_{N+j}, x_{N+k}))_{j,k=1}^{M,M} \end{pmatrix}$,
We thank Phaedon-Stelios Koutsourelakis for the inspiration to solve SPDEs with a maximum likelihood-based approach. The work of Igor Cialenco was partially supported by the National Science Foundation (NSF) grant DMS-0908099. T. Müller-Gronbach, K. Ritter, and T.Wagner, 2001, Amer. Math. Soc. 353 (2001), pp. 3945–3969.

(2) Repeat for $j = 1, 2, \ldots, n$, i.e. for $t_1, t_2, \ldots, t_n = T$

- $V_1 := \text{diag}(\psi(u^{-1}_1), \ldots, \psi(u^{-1}_N))$, $V_2 := \begin{pmatrix} V_1 & 0 \\ 0 & 0 \end{pmatrix}$.
- $\Psi := V_1 \Psi_0 V_1$, $\Sigma := V_2 \Sigma_0 V_2$.
- Simulate $\xi \sim \mathcal{N}(0, \Sigma)$, $A := B(K_{PB} + \Sigma)^{-1}$.
- $u^j := B(K_{PB} + \Sigma)^{-1} \begin{pmatrix} u^{-1}_j + \xi \\ 0 \end{pmatrix} = A \begin{pmatrix} u^{-1}_j + \xi \\ 0 \end{pmatrix}$.

Of course, now the matrix $A$ needs to be updated for each time step and for each sample path so that the algorithm is much costlier.

Acknowledgements

We thank Phaedon-Stelios Koutsourelakis for the inspiration to solve SPDEs with a maximum likelihood-based approach. The work of Igor Cialenco was partially supported by the National Science Foundation (NSF) grant DMS-0908099. Gregory E. Fasshauer and Qi Ye acknowledge support from NSF grants DMS-0713848 and DMS-1115392. The authors also thank the anonymous referee and the editors for their helpful comments and suggestions which improved greatly the final manuscript.

References

Appendix 1. Reproducing-kernel Hilbert spaces

DEFINITION A.1 [21, Definition 10.1] A Hilbert space $H_K(D)$ of functions $f : D \rightarrow \mathbb{R}$ is called a reproducing-kernel Hilbert space with a reproducing kernel $K : D \times D \rightarrow \mathbb{R}$ if

(i) $K(\cdot, y) \in H_K(D)$ and (ii) $f(y) = \langle K(\cdot, y), f \rangle_{K, D}$,

for all $f \in H_K(D)$ and all $y \in D$. Here $\langle \cdot, \cdot \rangle_{K, D}$ is the inner product of $H_K(D)$.

According to [21, Theorem 10.4] all reproducing kernels are positive semi-definite. Wendland [21, Theorem 10.10] showed that a symmetric positive semi-definite kernel $K$ is always a reproducing kernel of a reproducing-kernel Hilbert space $H_K(D)$.

If $D$ is open and bounded (pre-compact) and $K \in L_2(D \times D)$ is symmetric positive definite, then Mercer’s theorem [8, Theorem 13.5] guarantees the existence of a countable set of positive values $\lambda_1 \geq \lambda_2 \geq \cdots > 0$ with $\sum_{k=1}^{\infty} \lambda_k < \infty$ and an orthonormal basis $\{e_k\}_{k=1}^{\infty}$ of $L_2(D)$ such that $K$ possesses the absolutely and uniformly convergent representation

$$K(x, y) = \sum_{k=1}^{\infty} \lambda_k e_k(x) e_k(y), \quad x, y \in D.$$ 

This Mercer series of $K$ implies that

$$\lambda_k e_k(y) = \int_D K(x, y) e_k(x) \, dx, \quad y \in D, \quad k \in \mathbb{N}.$$ 

DEFINITION A.2 The sequences $\{\lambda_k\}_{k=1}^{\infty}$ and $\{e_k\}_{k=1}^{\infty}$ given above are called the eigenvalues and eigenfunctions of the reproducing kernel $K$.

Since $\{e_k\}_{k=1}^{\infty}$ is orthonormal in $L_2(D)$ we can compute the series expansion of the integral-type kernel $K$ defined in Lemma 2.2, i.e.

$$\hat{K}(x, y) = \sum_{j=1}^{\infty} \sum_{k=1}^{\infty} \int_D \lambda_j e_j(x) e_j(z) \lambda_k e_k(z) e_k(y) \, dz = \sum_{k=1}^{\infty} \lambda_k^2 e_k(x) e_k(y).$$

It is easy to check that $\hat{K} \in L_2(D \times D)$ is symmetric positive definite and $\{\lambda_k^2\}_{k=1}^{\infty}$ and $\{e_k\}_{k=1}^{\infty}$ are the eigenvalues and eigenfunctions of $K$.

THEOREM A.3 [21, Theorem 10.29] Suppose that $K \in L_2(D \times D)$ is a symmetric positive-definite kernel on a pre-compact $D \subset \mathbb{R}^d$. Then its reproducing-kernel Hilbert space is given by

$$H_K(D) = \left\{ f \in L_2(D) : \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \left| \int_D f(x) e_k(x) \, dx \right|^2 < \infty \right\}$$

and the inner product has the representation

$$\langle f, g \rangle_{K, D} = \sum_{k=1}^{\infty} \frac{1}{\lambda_k} \int_D f(x) e_k(x) \, dx \int_D g(x) e_k(x) \, dx, \quad f, g \in H_K(D),$$

where $\{\lambda_k\}_{k=1}^{\infty}$ and $\{e_k\}_{k=1}^{\infty}$ are the eigenvalues and eigenfunctions of $K$. 
We can verify that \( \{ \sqrt{\lambda_k} e_k \}_{k=1}^\infty \) is an orthonormal base of \( H_K(D) \).

**Example A.4** The papers [10, 22] show that the Sobolev spline (Matérn function) of degree \( m > d/2 \),

\[
g_{m,\theta}(x) := \frac{2^{1-m-d/2}}{\pi^{d/2} \Gamma(m) \theta^{2m-d}} (\theta |x|)^{m-d/2} K_{d/2-m}(\theta |x|), \quad x \in \mathbb{R}^d, \quad \theta > 0,
\]

is a full-space Green function of the differential operator \( L := (\theta^2 I - \Delta)^m \), i.e. \( L g_{m,\theta} = \delta_0 \), where \( t \mapsto K_\nu(t) \) is the modified Bessel function of the second kind of order \( \nu \). The kernel function

\[
K_{m,\theta}(x, y) := g_{m,\theta}(x - y), \quad x, y \in \mathbb{R}^d,
\]

is a positive-definite kernel whose reproducing-kernel Hilbert space is equivalent to the \( L_2 \)-based Sobolev space of degree \( m \), i.e. \( H_{K_{m,\theta}}(\mathbb{R}^d) \simeq \mathcal{H}^m(\mathbb{R}^d) \). Its inner product has the explicit form

\[
(f, g)_{K_{m,\theta}, \mathbb{R}^d} = \int_{\mathbb{R}^d} P_\theta f(x)^T P_\theta g(x) \, dx, \quad f, g \in H_{K_{m,\theta}}(\mathbb{R}^d),
\]

where \( P_\theta := (Q_0, Q_1, \ldots, Q_m) \) and

\[
Q_j := \begin{cases} a_j \Delta_j^k, & j = 2k, \\ a_j \Delta_j \nabla^k, & j = 2k + 1, \\ a_j := \sqrt{k!(2m-2j)}, & k \in \mathbb{N}_0, \quad j = 0, 1, \ldots, m. \end{cases}
\]

According to [3, Theorem 1.4.6], the reproducing-kernel Hilbert space \( H_{K_{m,\theta}}(D) \) is endowed with the reproducing-kernel norm

\[
\|f\|_{K_{m,\theta}, D} = \inf_{\tilde{f} \in H_{K_{m,\theta}}(\mathbb{R}^d)} \{ \|\tilde{f}\|_{K_{m,\theta}, \mathbb{R}^d} : \tilde{f}|_D = f \}, \quad f \in H_{K_{m,\theta}}(D).
\]

Moreover, if the open bounded domain \( D \subset \mathbb{R}^d \) is regular then \( H_{K_{m,\theta}}(D) \) is equivalent to the \( L_2 \)-based Sobolev space of degree \( m \), i.e. \( H_{K_{m,\theta}}(D) \simeq \mathcal{H}^m(D) \).

### Appendix 2. Differential and boundary operators

In this section we define differential and boundary operators on Sobolev spaces \( \mathcal{H}^m(D) \). The differential and boundary operators in this paper are well defined since we assume that the open and bounded domain \( D \) is regular, i.e. it satisfies a strong local Lipschitz condition which implies a uniform cone condition (see [1, Chapter 4.1]). This means that \( D \) has a regular boundary \( \partial D \).

Let the notation for typical derivatives be

\[
D^\alpha := \prod_{k=1}^d \frac{\partial^{\alpha_k}}{\partial x_k^{\alpha_k}}, \quad |\alpha| := \sum_{k=1}^d \alpha_k, \quad \alpha := (\alpha_1, \ldots, \alpha_d) \in \mathbb{N}_0^d,
\]

We extend these derivatives to weak derivatives (see [1, Chapter 1.5]) using the same symbol \( D^\alpha \). Using these weak derivatives, the classical \( L_2 \)-based Sobolev space \( \mathcal{H}^m(D) \) is given by

\[
\mathcal{H}^m(D) := \{ f \in L_2^{loc}(D) : D^\alpha f \in L_2(D), |\alpha| \leq m, \alpha \in \mathbb{N}_0^d \}, \quad m \in \mathbb{N}_0,
\]

equipped with the natural inner product

\[
(f, g)_{\mathcal{H}^m(D)} := \sum_{|\alpha| \leq m} \int_D D^\alpha f(x) D^\alpha g(x) \, dx, \quad f, g \in \mathcal{H}^m(D).
\]

The weak derivative \( D^\alpha \) is a bounded linear operator from \( \mathcal{H}^m(D) \) into \( L_2(D) \) when \( |\alpha| \leq m \). Moreover, \( D^\alpha f \) is well posed on the boundary \( \partial D \) when \( f \in \mathcal{H}^m(D) \) and \( |\alpha| \leq m - 1 \) and we denote that \( D^\alpha |_{\partial D} := D^\alpha f |_{\partial D} \). The book [1] and the paper [11] show that \( D^\alpha |_{\partial D} \) can be extended to a bounded linear operator from \( \mathcal{H}^m(D) \) into \( L_2(\partial D) \) when \( |\alpha| \leq m - 1 \) because \( D \) has a regular boundary \( \partial D \). The \( L_2(\partial D) \)-inner product is denoted by

\[
(f, g)_{\partial D} = \int_{\partial D} f(x) g(x) \, d\mathcal{S}(x) \quad \text{when } d \geq 2 \text{ and } \partial D \text{ is the boundary manifold,}
\]

and

\[
(f, g)_{\partial D} = f(b)g(b) + f(a)g(a) \quad \text{when } d = 1 \text{ and } \partial D = \{a, b\}.
\]
Definition A.5 A differential operator $P : \mathcal{H}^m(\mathcal{D}) \to L_2(\mathcal{D})$ is well defined by

$$P = \sum_{|\alpha| \leq m} c_{\alpha} D^\alpha$$

where $c_{\alpha} \in C^\infty(\bar{\mathcal{D}})$ and $\alpha \in \mathbb{N}_0^d$, $m \in \mathbb{N}$,

and its order is given by $\mathcal{O}(P) := \max\{|\alpha| : c_{\alpha} \neq 0, |\alpha| \leq m, \alpha \in \mathbb{N}_0^d\}$. A boundary operator $B : \mathcal{H}^m(\mathcal{D}) \to L_2(\partial \mathcal{D})$ is well defined by

$$B = \sum_{|\alpha| \leq m-1} b_{\alpha} D^\alpha |_{\partial \mathcal{D}}$$

where $b_{\alpha} \in C^\infty(\partial \mathcal{D})$ and $\alpha \in \mathbb{N}_0^d$, $m \in \mathbb{N}$,

and its order is given by $\mathcal{O}(B) := \max\{|\alpha| : b_{\alpha} \neq 0, |\alpha| \leq m-1, \alpha \in \mathbb{N}_0^d\}$.

It is obvious that the differential operator $P$ and the boundary operator $B$ are bounded (continuous) linear operators on $\mathcal{H}^m(\mathcal{D})$ with values in $L_2$ whenever $\mathcal{O}(P) \leq m$ and $\mathcal{O}(B) \leq m - 1$. Much more detail on differential and boundary operators can be found in [1,11].

If $\mathcal{H}_K(\mathcal{D})$ is embedded into $\mathcal{H}^m(\mathcal{D})$ then the eigenvalues $\{\lambda_k\}_{k=1}^{\infty}$ and eigenfunctions $\{e_k\}_{k=1}^{\infty}$ of the reproducing kernel $K$ satisfy

$$\lambda_k |e_k|_{\mathcal{H}_K}^2 \leq C^2 \sqrt{\lambda_k} |e_k|_{\mathcal{H}}^2 = C^2, \quad k \in \mathbb{N},$$

because $\{\sqrt{\lambda_k} e_k\}_{k=1}^{\infty}$ is an orthonormal base of $\mathcal{H}_K(\mathcal{D})$. When $m > d/2$ then $\mathcal{H}^m(\mathcal{D})$ is embedded into $C(\bar{\mathcal{D}})$ by the Sobolev embedding theorem. The recent papers [10,11,22] show that there exist kernels $K$ and its order is given by $\mathcal{O}(K)$. Then we can obtain the following lemma.

Lemma A.6 Consider a differential operator $P$ with order $\mathcal{O}(P) \leq m$ and a boundary operator $B$ with order $\mathcal{O}(B) \leq m - 1$, where $m > d/2$. If the reproducing-kernel Hilbert space $\mathcal{H}_K(\mathcal{D})$ is embedded into the Sobolev space $\mathcal{H}^m(\mathcal{D})$, then

$$Pf = \sum_{k=1}^{\infty} \hat{f}_k \sqrt{\lambda_k} P e_k, \quad Bf = \sum_{k=1}^{\infty} \hat{f}_k \sqrt{\lambda_k} B e_k, \quad f \in \mathcal{H}_K(\mathcal{D}),$$

where $\hat{f}_k = \langle f, e_k \rangle_{\mathcal{H}_K, \mathcal{D}}$ for each $k \in \mathbb{N}$ and $\{\lambda_k\}_{k=1}^{\infty}$ and $\{e_k\}_{k=1}^{\infty}$ are the eigenvalues and eigenfunctions of the reproducing kernel $K$.

Proof According to Theorem A.3 each $f \in \mathcal{H}_K(\mathcal{D})$ can be expanded as $f = \sum_{k=1}^{\infty} \hat{f}_k \sqrt{\lambda_k} e_k$. Since $\{\sqrt{\lambda_k} e_k\}_{k=1}^{\infty}$ is an orthonormal basis we have $\sum_{k=1}^{\infty} |\hat{f}_k|^2 < \infty$. Let $f_n := \sum_{k=1}^{n} \hat{f}_k \sqrt{\lambda_k} e_k$ for each $n \in \mathbb{N}$. Then

$$|f_n - f|_{\mathcal{H}^m}^2 \leq C^2 |f_n - f|_{\mathcal{H}_K, \mathcal{D}}^2 \leq C^2 \sum_{k=n+1}^{\infty} |\hat{f}_k|^2 \to 0 \text{ when } n \to \infty.$$

The proof is completed by remembering that $P$ and $B$ are bounded linear operators on $\mathcal{H}^m(\mathcal{D})$.

If $\mathcal{H}_K(\mathcal{D})$ is embedded into $\mathcal{H}^m(\mathcal{D})$, then for each $|\alpha| \leq m$, $|\beta| \leq m$ and $\alpha, \beta \in \mathbb{N}_0^d$, we have

$$\left( \int_{\mathcal{D}} \int_{\mathcal{D}} |\sum_{k=1}^{\infty} \lambda_k^2 |D^\alpha e_k(x)D^\beta e_k(y)|^2 \, dx \, dy \right)^{1/2} \leq \sum_{k=1}^{\infty} |\lambda_k|^2 |D^\alpha e_k|_{L_2(\mathcal{D})} |D^\beta e_k|_{L_2(\mathcal{D})} \leq \sum_{k=1}^{\infty} |\lambda_k|^2 |e_k|_{\mathcal{H}^m}^2 \leq C^2 \sum_{k=1}^{\infty} \lambda_k < \infty,$$

which implies that $\mathcal{K} \in \mathcal{H}^{m,m}(\mathcal{D} \times \mathcal{D})$. Let $n := [m - d/2] - 1$. The Sobolev embedding theorem shows that $\mathcal{H}_m(\mathcal{D} \times \mathcal{D}) \subset C^{n,n}(\mathcal{D} \times \mathcal{D})$. Then we can obtain the following lemma.
Lemma A.7 Consider a differential operator $P$ with order $O(P) < m - d/2$ and a boundary operator $B$ with order $O(B) < m - d/2$, where $m > d/2$. If the reproducing-kernel Hilbert space $H_K(D)$ is embedded into the Sobolev space $H^m(D)$, then

\[
P_1P_2^*K(x,y) := P_{z_1}P_{z_2}^*K(z_1,z_2)|_{z_1=x,z_2=y} = \sum_{k=1}^{\infty} \lambda_k^2 P e_k(x) P e_k(y),
\]

\[
B_1B_2^*K(x,y) := B_{z_1}B_{z_2}^*K(z_1,z_2)|_{z_1=x,z_2=y} = \sum_{k=1}^{\infty} \lambda_k^2 B e_k(x) B e_k(y),
\]

\[
P_1B_2^*K(x,y) := P_{z_1}B_{z_2}^*K(z_1,z_2)|_{z_1=x,z_2=y} = \sum_{k=1}^{\infty} \lambda_k^2 P e_k(x) B e_k(y),
\]

where $\{\lambda_k\}_{k=1}^{\infty}$ and $\{e_k\}_{k=1}^{\infty}$ are the eigenvalues and eigenfunctions of $K$. Moreover, $P_1P_2^* \in C(\bar{D} \times \bar{D})$, $B_1B_2^* \in C(\partial D \times \partial D)$ and $P_1B_2^* \in C(\bar{D} \times \partial D)$.