

# Error analysis for a statistical finite element method

Toni Karvonen<sup>a,b,\*</sup>, Fehmi Cirak<sup>c</sup>, Mark Girolami<sup>c</sup>

<sup>a</sup>*School of Engineering Sciences, Lappeenranta–Lahti University of Technology LUT, Lappeenranta, Finland*

<sup>b</sup>*Department of Mathematics and Statistics, University of Helsinki, Finland*

<sup>c</sup>*Department of Engineering, University of Cambridge, United Kingdom*

---

## Abstract

The recently proposed statistical finite element (statFEM) approach synthesises measurement data with finite element models and allows for making predictions about the unknown true system response. We provide a probabilistic error analysis for a prototypical statFEM setup based on a Gaussian process prior under the assumption that the noisy measurement data are generated by a deterministic true system response function that satisfies a second-order elliptic partial differential equation for an unknown true source term. In certain cases, properties such as the smoothness of the source term may be misspecified by the Gaussian process model. The error estimates we derive are for the expectation with respect to the measurement noise of the  $L^2$ -norm of the difference between the true system response and the mean of the statFEM posterior. The estimates imply polynomial rates of convergence in the numbers of measurement points and finite element basis functions and depend on the Sobolev smoothness of the true source term and the Gaussian process model. A numerical example for Poisson’s equation is used to illustrate these theoretical results.

**Keywords:** finite element methods, Gaussian processes, reproducing kernel Hilbert spaces

**2020 MSC:** 65N30, Secondary 46E22, 60G15, 65D05

---

## 1. Introduction

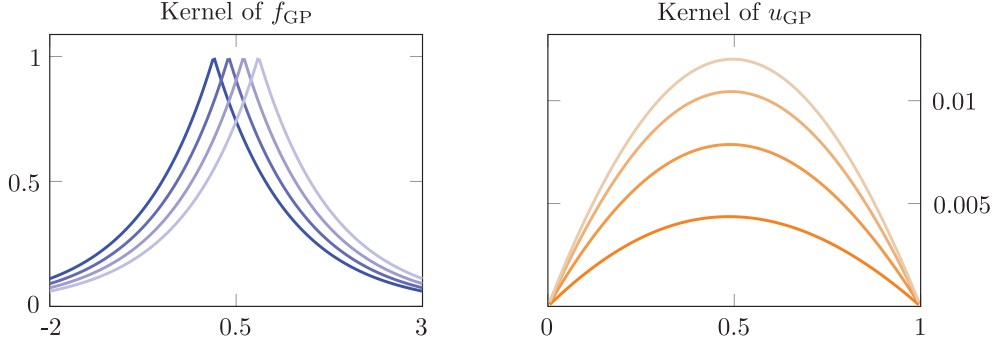
The finite element method has become an indispensable tool for solving partial differential equations in engineering and applied sciences. Today, the design, manufacture and maintenance of most engineering products rely on mathematical models based on finite element discretised partial differential equations (PDEs). These models depend on a wide range of parameters, including material, geometry, and loading, which are inevitably subject to both epistemic and aleatoric uncertainties. Consequently, the response of the actual engineering product and the inevitably misspecified mathematical model often bear little resemblance to each other, resulting in inefficient designs and overtly cautious operational decisions. Fortunately, more and more engineering products are equipped with sensor networks that provide operational measurement data [e.g., 14]. The statistical finite element method (statFEM) allows us to infer the true system response by synthesising limited measurement data with finite element models [19]. By adopting a Bayesian approach, the prior probability measure of the finite element solution is obtained from the misspecified finite element model by solving a probabilistic forward problem. Although any parameters of the finite element model can be random, in this article only the source term of the respective PDE is random and Gaussian, so that the finite element solution is Gaussian. The assumed data-generating process for determining the likelihood of the measured data is additively composed of the random finite element solution, the known random measurement noise, and, possibly, an unknown random discrepancy component. The chosen prior and the likelihood ensure that the posterior finite element probability density conditioned on the measurement data is Gaussian and easily computable.

More concretely, we consider the following mathematical problem. Let  $\mathcal{L}$  be a differential operator and suppose that the system response  $u$  that generated the measurement data is given by the solution of

$$\mathcal{L}u = f \tag{1}$$

---

\*Corresponding author. Email address: [toni.karvonen@lut.fi](mailto:toni.karvonen@lut.fi)



**Fig. 1:** Left: Four translates  $K(\cdot, x)$  of the Matérn covariance kernel in (7) with  $\nu = 1/2$  and  $\sigma = \ell = 1$ . Right: Four translates  $K_u(\cdot, x)$  of the corresponding kernel of the process  $u_{\text{GP}}$  for Poisson's equation with zero boundary conditions.

on a bounded domain  $\Omega$  with zero Dirichlet boundary conditions. The statistical component of the statFEM solution arises from the placement of a stochastic process prior on the source term  $f$  and, possibly, the differential operator  $\mathcal{L}$  or some of its parameters. Doing this induces a stochastic process prior over the solution  $u$ . After hyperparameter estimation and inclusion of additional levels of statistical modelling [25], which may account for various modelling discrepancies, one uses Bayesian inference to obtain a posterior distribution over the PDE solution given the measurement data. The posterior can then be used to predict the system behaviour at previously unseen data locations and provide associated uncertainty quantification. See [8, 19, 27] for applications of this methodology to different types of PDEs and [1] for a somewhat different approach focusing on random meshes. In any non-trivial setting, computation of the prior for  $u$  from that placed on  $f$  requires solving the PDE (1). In statFEM, the PDE is solved using finite elements. Due to their tractability, Gaussian processes (GPs) are often the method of choice for modelling physical phenomena. In the PDE setting that we consider Gaussian processes are particularly convenient because a GP prior,  $f_{\text{GP}}$ , on  $f$  induces a GP prior,  $u_{\text{GP}}$ , on  $u$  if the PDE is linear (see Fig. 1). The induced prior  $u_{\text{GP}}$  has been studied in [32, 35] and Section 3.1.2 of [7]. Although  $u_{\text{GP}}$  is generally not available in closed form, it is straightforward to approximate its mean and covariance functions from those of  $f_{\text{GP}}$  by using finite elements.

In this article, we provide estimates of the predictive error for the GP-based statFEM when the data are noisy evaluations of some deterministic true system response function  $u_t$  which is assumed to be the solution of (1) for an unknown—but deterministic—true source term  $f_t$ . Due to the complexity and difficulty of analysing a full statFEM approach, we consider a prototypical version that consists of a GP prior on  $f$  and, possibly, a GP discrepancy term. Scaling and other parameters these processes may have are assumed fixed. Despite recent advances in understanding the behaviour of GP hyperparameters and their effect on the convergence of GP approximation [24, 41, 49], these results are either not directly applicable in our setting or too generic, in that they assume that the parameter estimates remain in some compact sets, which has not been verified for commonly used parameter estimation methods, such as maximum likelihood.

As mentioned, finite elements are needed for computation of the induced prior  $u_{\text{GP}}$  and the associated posterior. But why not simply use a readily available and explicit GP prior for  $u$ , such as one with a Matérn covariance kernel, instead of something that requires finite element approximations? The main reason (besides this being the first step towards analysing the full statFEM) is that a prior  $u_{\text{GP}}$ , for which  $\mathcal{L}u_{\text{GP}} = f_{\text{GP}}$ , satisfies the structural constraints imposed by the PDE model and can therefore be expected to yield more accurate point estimates and more reliable uncertainty quantification than a more arbitrary prior if the data are generated by a solution of (1) for some source term. We give a detailed description of the considered method in Section 2.

### 1.1. Contributions

Our contribution consists of a number of error estimates for the statFEM approach sketched above. Suppose that the measurements are  $y_i = u_t(\mathbf{x}_i) + \varepsilon_i$  for  $n$  locations  $\mathbf{x}_i \in \Omega \subset \mathbb{R}^d$  and independent Gaussian noises  $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$ . The

regression error estimates we prove are of the form

$$\mathbb{E} \left[ \|u_t - \tilde{m}\|_{L^2(\Omega)} \right] \leq C_1 n^{-1/2+a} + C_2 n_{\text{FE}}^{-q} n^{3/2}, \quad (2)$$

where  $\tilde{m}$  is a posterior mean function obtained from statFEM and the expectation is with respect to the measurement noise. The constant  $a \in (0, 1/2)$  depends on the smoothness of  $f_t$  and  $q > 0$  is the dimension  $d$  dependent characteristic rate of convergence of the finite element approximation with  $n_{\text{FE}}$  elements. In (2) it is assumed that the points  $\mathbf{x}_i$  cover  $\Omega$  sufficiently uniformly. In Section 6.5 we present error estimates for four different variants of statFEM, each of which corresponds to a different  $\tilde{m}$ :

- Theorem 1 assumes that no finite element discretisation is required for computation of  $\tilde{m}$ . In this case  $C_1 > 0$  and  $C_2 = 0$ . It is required that  $f_t$  be at least as smooth as the prior  $f_{\text{GP}}$ .
- In Theorem 3, the more realistic assumption that  $\tilde{m}$  is constructed via a finite element approximation is used. In this case  $C_1, C_2 > 0$ . It is required that  $f_t$  be at least as smooth as the prior  $f_{\text{GP}}$ .
- Theorems 4 and 5 concern versions which include a GP discrepancy term  $v_{\text{GP}}$  (i.e., the prior for  $u$  is  $u_{\text{GP}} + v_{\text{GP}}$ ) and do not use or use, respectively, finite element discretisation to compute  $\tilde{m}$ . These theorems allow the priors to misspecify the source term and system response smoothness as it is not required that  $f_t$  be at least as smooth as  $f_{\text{GP}}$  or that  $u_t$  be at least as smooth as  $v_{\text{GP}}$  or  $u_{\text{GP}}$ . Such misspecification results are valuable because one often lacks precise information about the smoothness (or other properties) of  $u_t$ .

As discussed in Remark 1, these rates are likely slightly sub-optimal. Some numerical examples for one-dimensional Poisson’s equation are given in Section 4.

The proofs of these results are based on reproducing kernel Hilbert space (RKHS) techniques which are commonly used to analyse approximation properties of GPs [6, 7, 24, 41, 44, 47, 49]. Our central tool is Theorem 8, which describes the RKHS associated to the prior  $u_{\text{GP}}$  under the assumptions that the RKHS for  $f_{\text{GP}}$  is a Sobolev space and  $\mathcal{L}$  is a second-order elliptic differential operator. This result is used to “export” (a) regression error estimates in some of the aforementioned references and (b) bounds on the concentration function [29, 44] from a “standard” GP prior  $f_{\text{GP}}$  (e.g., one specified by a Matérn covariance kernel) to the transformed prior  $u_{\text{GP}}$ . When a finite element approximation is used, the regression error estimates are combined with a simple result (Proposition 4) which bounds the difference between GP posterior means for two different kernels in terms of the maximal difference of the kernels.

## 1.2. Related work

Solving PDEs with kernel-based methods goes back at least to the work of Kansa [22]; see [11, 15] as well as Chapter 16 in [48] for a more general treatment. In the language of GPs, this radial basis function collocation approach is essentially based on modelling  $u$  as a GP with a given covariance kernel and conditioning on the derivative observations  $\mathcal{L}u(\mathbf{x}_i) = f(\mathbf{x}_i)$ . Typically no synthesis of actual measurement data is present (though this could be easily included). For convergence results in a well-specified setting, see for example Theorem 16.15 in [48]. In a GP setting similar methods have been proposed and analysed by Cialenco et al. [6], Cockayne et al. [7], Graepel [20], and Raissi et al. [35]. For some error estimates, see Lemma 3.4 and Proposition 3.5 in [6]. Priors and covariance kernels derived from Green’s function have been considered by Fasshauer and Ye [12, 13] and Owhadi [32]. Furthermore, Papandreou et al. [33] have recently derived bounds on the Wasserstein distance  $W_2$  between the ideal prior and posterior (see Section 2.1 in the present article) and their finite element approximations.

## 2. Statistical finite element methods

This section describes the statFEM approach that is analysed in Section 6.5 and discusses some extensions that are not covered by our analysis. We begin by defining the class of second-order elliptic PDE problems that are considered in this article.

Let  $d \in \mathbb{N}$  and  $\Omega \subset \mathbb{R}^d$  be an open and bounded set which satisfies an interior cone condition [e.g., 48, Definition 3.6] and has a Lipschitz boundary  $\partial\Omega$  (i.e., the boundary is locally the graph of a Lipschitz function). Occasionally we also require an assumption that  $\partial\Omega$  be  $C^k$  or  $C^{k,\alpha}$ , which means that its boundary can be interpreted locally as the graph of a

function in  $C^k(\mathbb{R}^{d-1})$  or in the Hölder space  $C^{k,\alpha}(\mathbb{R}^d)$ . Definitions of Hölder spaces and the smoothness of the domain can be found in Sections 3.1 and 6.1.

Let  $\mathcal{L}$  be a second-order partial differential operator of the form

$$\mathcal{L}u = - \sum_{i=1}^d \sum_{j=1}^d \partial_i(a_{ij}\partial_j u) + \sum_{i=1}^d b_i \partial_i u + cu \quad (3)$$

for coefficient functions  $a_{ij}$ ,  $b_i$  and  $c$  which are bounded on the closure  $\bar{\Omega}$ . We further assume that  $a_{ij} \in C^1(\Omega)$  and  $a_{ij} = a_{ji}$  for all  $i, j \in \{1, \dots, d\}$ . The differential operator is assumed to be uniformly elliptic, which is to say that there is a positive constant  $\lambda$  such that  $\sum_{i=1}^d \sum_{j=1}^d a_{ij}(\mathbf{x})z_i z_j \geq \lambda \|\mathbf{z}\|_2^2$  for any  $\mathbf{x} \in \Omega$  and  $\mathbf{z} \in \mathbb{R}^d$ . Moreover, our results use the following regularity assumption.

**Assumption 1** (Regularity). For a given  $k \in \mathbb{N}_0$ , the boundary  $\partial\Omega$  is  $C^{k+2}$  and  $a_{ij}, b_i, c \in C^{k+1}(\bar{\Omega})$  for all  $i, j \in \{1, \dots, d\}$ .

Given a known differential operator  $\mathcal{L}$ , we consider the elliptic PDE

$$\begin{aligned} \mathcal{L}u &= f & \text{in } \Omega, \\ u &= 0 & \text{on } \partial\Omega \end{aligned} \quad (4)$$

for a source term  $f: \Omega \rightarrow \mathbb{R}$ . Let  $\mathcal{H}_{\mathcal{L}}(\Omega)$  be some space of functions defined on  $\Omega$  such that the above PDE admits a unique classical (i.e., pointwise) solution  $u: \bar{\Omega} \rightarrow \mathbb{R}$  for every  $f \in \mathcal{H}_{\mathcal{L}}(\Omega)$ . Therefore there is a linear operator  $\mathcal{L}^{-1}: \mathcal{H}_{\mathcal{L}}(\Omega) \rightarrow C^2(\Omega)$  such that  $u = \mathcal{L}^{-1}f$  is the unique solution of (4) for any  $f \in \mathcal{H}_{\mathcal{L}}(\Omega)$ . Suppose that there is a true system response  $u_t$  which is the unique solution of

$$\begin{aligned} \mathcal{L}u_t &= f_t & \text{in } \Omega, \\ u_t &= 0 & \text{on } \partial\Omega \end{aligned} \quad (5)$$

for a certain true source term  $f_t \in \mathcal{H}_{\mathcal{L}}(\Omega)$ , which may be unknown, and that one has access to  $n \in \mathbb{N}$  noisy observations  $\mathbf{Y} = (y_1, \dots, y_n) \in \mathbb{R}^n$  of  $u_t$  at distinct data locations  $X = \{\mathbf{x}_1, \dots, \mathbf{x}_n\} \subset \Omega$ : for  $i \in \{1, \dots, n\}$ ,

$$y_i = u_t(\mathbf{x}_i) + \varepsilon_i \quad (6)$$

where  $\sigma_\varepsilon > 0$  and  $\varepsilon_i \sim \mathcal{N}(0, \sigma_\varepsilon^2)$  are i.i.d. The statFEM approach provides a means for predicting the value of the true system response,  $u_t(\mathbf{x})$ , at any point,  $\mathbf{x}$ , in the domain. It also provides uncertainty estimates for these predictions that are based on the observations, the differential operator  $\mathcal{L}$ , and a prior which encodes, for example, assumptions on the smoothness of the true source term. We emphasise that here  $u_t$  and  $f_t$  are always assumed to be some fixed and deterministic functions. Although we use GPs to model them, the functions themselves are not considered stochastic processes. All expectations that occur in this article are with respect to the noise variables  $\varepsilon_i$  alone. The locations  $\mathbf{x}_i$  are not assumed random in this article.

### 2.1. Gaussian process inference

Let  $K: \Omega \times \Omega \rightarrow \mathbb{R}$  be a positive-semidefinite kernel. That is, for any  $n \in \mathbb{N}$ ,  $\alpha_1, \dots, \alpha_n \in \mathbb{R}$ , and  $\mathbf{z}_1, \dots, \mathbf{z}_n \in \Omega$  it holds that  $\sum_{i=1}^n \sum_{j=1}^n \alpha_i \alpha_j K(\mathbf{z}_i, \mathbf{z}_j) \geq 0$ . One of the most ubiquitous—as well as one which appears repeatedly in this article—classes of positive-semidefinite kernels is that of the Matérn kernels

$$K(\mathbf{x}, \mathbf{y}) = \sigma^2 \frac{2^{1-\nu}}{\Gamma(\nu)} \left( \frac{\sqrt{2\nu} \|\mathbf{x} - \mathbf{y}\|_2}{\ell} \right)^\nu K_\nu \left( \frac{\sqrt{2\nu} \|\mathbf{x} - \mathbf{y}\|_2}{\ell} \right), \quad (7)$$

where  $\nu > 0$  is a smoothness parameter,  $\ell > 0$  a length-scale parameter,  $\sigma > 0$  a scaling parameter,  $\Gamma$  the gamma function, and  $K_\nu$  the modified Bessel function of the second kind of order  $\nu$ . These kernels are important because they induce Sobolev spaces (see Section 3.1). We model  $f_t$  as a Gaussian process  $f_{\text{GP}} \sim \text{GP}(m, K)$  and assume that

$$(i) \ m \in \mathcal{H}_{\mathcal{L}}(\Omega), \quad (ii) \ K(\cdot, \mathbf{x}) \in \mathcal{H}_{\mathcal{L}}(\Omega), \quad \mathcal{L}_{\mathbf{x}}^{-1} K(\cdot, \mathbf{x}) \in \mathcal{H}_{\mathcal{L}}(\Omega), \quad \mathbf{x} \in \Omega, \quad (8)$$

where the subscript denotes the variable with respect to which the linear operator is applied. These assumptions ensure that various functions that we are about to introduce are unique and pointwise well-defined. Because  $\mathcal{L}$  is a linear differential operator, the above GP prior over  $f_t$  induces the prior  $u_{\text{GP}} \sim \text{GP}(m_u, K_u)$  over  $u_t$  with mean function  $m_u = \mathcal{L}^{-1}m$  and covariance kernel  $K_u$ , which satisfies

$$\mathcal{L}_x \mathcal{L}_y K_u(\mathbf{x}, \mathbf{y}) = K(\mathbf{x}, \mathbf{y}) \quad (9)$$

for all  $\mathbf{x}, \mathbf{y} \in \Omega$ , as well as  $\mathcal{L}K_u(\cdot, \mathbf{y}) = 0$  on  $\partial\Omega$  for every  $\mathbf{y} \in \bar{\Omega}$ . The existence and uniqueness of the mean and covariance are guaranteed by the assumptions in (8).

To arrive at an ideal version of the GP-based statFEM we condition the GP  $u_{\text{GP}}$  on the measurement data in (6). This yields the conditional process

$$u_{\text{GP}} | \mathbf{Y} \sim \text{GP}(m_{u|\mathbf{Y}}, K_{u|\mathbf{Y}}),$$

whose mean and covariance are

$$m_{u|\mathbf{Y}}(\mathbf{x}) = m_u(\mathbf{x}) + \mathbf{K}_u(\mathbf{x}, X)^\top (\mathbf{K}_u(X, X) + \sigma_\varepsilon^2 \mathbf{I}_n)^{-1} (\mathbf{Y} - \mathbf{m}_u(X)), \quad (10a)$$

$$K_{u|\mathbf{Y}}(\mathbf{x}, \mathbf{y}) = K_u(\mathbf{x}, \mathbf{y}) - \mathbf{K}_u(\mathbf{x}, X)^\top (\mathbf{K}_u(X, X) + \sigma_\varepsilon^2 \mathbf{I}_n)^{-1} \mathbf{K}_u(X, \mathbf{y}), \quad (10b)$$

where  $\mathbf{K}_u(X, X)$  is the  $n \times n$  kernel matrix with elements  $K_u(\mathbf{x}_i, \mathbf{x}_j)$ ,  $\mathbf{K}_u(\mathbf{x}, X)$  and  $\mathbf{m}_u(X)$  are  $n$ -vectors with elements  $K_u(\mathbf{x}, \mathbf{x}_i)$  and  $m_u(\mathbf{x}_i)$ , respectively, and  $\mathbf{I}_n$  is the  $n \times n$  identity matrix. However, the mean function  $m_u = \mathcal{L}^{-1}m$  and the covariance kernel  $K_u$  in (9) cannot be solved in closed form in all but the simplest of cases. This necessitates replacing their occurrences in (10) with finite element approximations.

## 2.2. Finite Element methods

Let  $H_0^1(\Omega)$  denote the Sobolev space of order one and zero trace. Informally, this space consists of once weakly differentiable functions that vanish on  $\partial\Omega$ . A function  $u \in H_0^1(\Omega)$  is a weak solution to the PDE in (4) if

$$B(u, v) = \int_{\Omega} f(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x}, \quad v \in H_0^1(\Omega), \quad (11)$$

where

$$B(u, v) = \int_{\Omega} \left( \sum_{i=1}^d \sum_{j=1}^d a_{ij}(\mathbf{x}) [\partial_i u(\mathbf{x})] [\partial_j v(\mathbf{x})] + \sum_{i=1}^d b_i(\mathbf{x}) [\partial_i u(\mathbf{x})] v(\mathbf{x}) + c(\mathbf{x}) u(\mathbf{x}) v(\mathbf{x}) \right) d\mathbf{x}$$

is the bilinear form associated with the elliptic differential operator  $\mathcal{L}$  in (3). The bilinear form is obtained from  $\int_{\Omega} [\mathcal{L}u(\mathbf{x})]v(\mathbf{x}) \, d\mathbf{x} = \int_{\Omega} f(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x}$  via integration by parts.

Finite element methods construct approximations to the weak solution by replacing the infinite-dimensional Sobolev space in (11) with a finite-dimensional trial space  $V \subset H_0^1(\Omega)$  and finding  $u^{\text{FE}} \approx u$  in  $V$  such that

$$B(u^{\text{FE}}, v) = \int_{\Omega} f(\mathbf{x})v(\mathbf{x}) \, d\mathbf{x}, \quad v \in V.$$

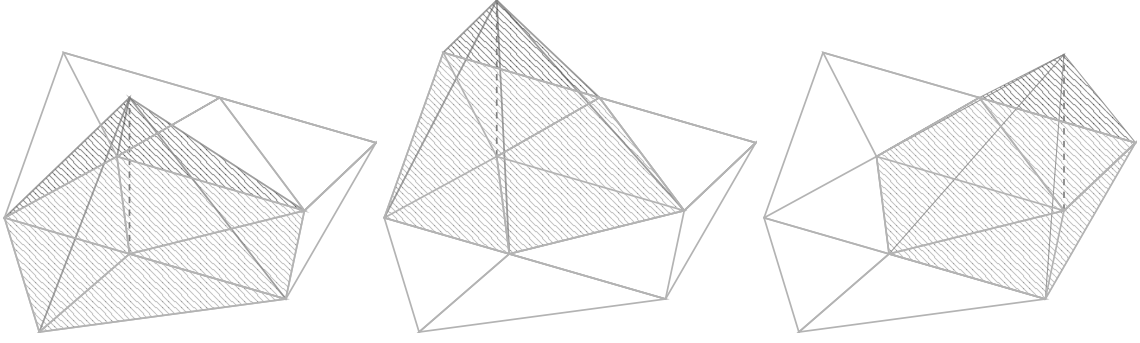
If  $V$  is an  $n_{\text{FE}}$ -dimensional vector space with basis  $\phi_1, \dots, \phi_{n_{\text{FE}}}$ , it follows from the bilinearity of  $B$  and linearity of integration that

$$u^{\text{FE}}(\mathbf{x}) = \sum_{i=1}^{n_{\text{FE}}} u_i \phi_i(\mathbf{x}), \quad (12)$$

where the coefficient vector  $\mathbf{u} = (u_1, \dots, u_{n_{\text{FE}}})$  is the solution of the linear system

$$\begin{pmatrix} B(\phi_1, \phi_1) & \cdots & B(\phi_{n_{\text{FE}}}, \phi_1) \\ \vdots & \ddots & \vdots \\ B(\phi_{n_{\text{FE}}}, \phi_1) & \cdots & B(\phi_{n_{\text{FE}}}, \phi_{n_{\text{FE}}}) \end{pmatrix} \begin{pmatrix} u_1 \\ \vdots \\ u_{n_{\text{FE}}} \end{pmatrix} = \begin{pmatrix} \int_{\Omega} f(\mathbf{x})\phi_1(\mathbf{x}) \, d\mathbf{x} \\ \vdots \\ \int_{\Omega} f(\mathbf{x})\phi_{n_{\text{FE}}}(\mathbf{x}) \, d\mathbf{x} \end{pmatrix}, \quad (13)$$

where the  $n_{\text{FE}} \times n_{\text{FE}}$  matrix on the left-hand side is called the stiffness matrix and denoted by  $\mathbf{A}$ .



**Fig. 2:** Triangulation of a polyhedral domain in  $\mathbb{R}^2$  and three piecewise linear hat functions that vanish on the boundary.

The choice of the trial space  $V \subset H_0^1(\Omega)$  determines how well  $u^{\text{FE}}$  approximates  $u$ . The simplest and most prevalent approach is to use piecewise linear basis functions. Assume for a moment that  $\Omega$  is a polyhedron. One first constructs a triangular partition of  $\Omega$  (or mesh) that consists of simplices  $\Omega_1, \dots, \Omega_{\tilde{n}}$  defined by  $n_{\text{FE}} < \tilde{n}$  vertices  $\mathbf{x}_1^{\text{FE}}, \dots, \mathbf{x}_{n_{\text{FE}}}^{\text{FE}}$  in the interior of  $\Omega$  plus additional vertices on the boundary and then selects  $V$  as the space of continuous functions that are piecewise linear in each simplex and zero on the boundary. This space is spanned by piecewise linear hat functions  $\phi_1, \dots, \phi_{n_{\text{FE}}}$  with the property  $\phi_i(\mathbf{x}_j^{\text{FE}}) = \delta_{ij}$ . The finite element approximation is then given by (12), where the coefficients  $u_i$  are obtained from the linear system (13). See [4, 26, 30] for detailed reviews of finite element methods. Fig. 2 shows a triangular mesh for a two-dimensional domain and piecewise linear hat functions.

### 2.3. Finite element posterior

One can approximate the mean  $m_u$  and covariance  $K_u$  of  $u_{\text{GP}} \sim \text{GP}(m_u, K_u)$  with the finite element method. Because  $\mathcal{L}m_u = m$ , by solving  $\mathbf{u}$  from (13) with  $f = m$  we obtain the mean approximation

$$m_u^{\text{FE}}(\mathbf{x}) = \boldsymbol{\mu}^\top \mathbf{A}^{-1} \boldsymbol{\phi}(\mathbf{x}) \approx m_u(\mathbf{x}),$$

where  $\boldsymbol{\phi}(\mathbf{x}) = (\phi_1(\mathbf{x}), \dots, \phi_{n_{\text{FE}}}(\mathbf{x}))$  and  $\boldsymbol{\mu} = (\int_{\Omega} m(\mathbf{x})\phi_1(\mathbf{x}) d\mathbf{x}, \dots, \int_{\Omega} m(\mathbf{x})\phi_{n_{\text{FE}}}(\mathbf{x}) d\mathbf{x})$ . Because  $\mathcal{L}_x \mathcal{L}_y K_u(\mathbf{x}, \mathbf{y}) = K(\mathbf{x}, \mathbf{y})$ , we may approximate  $K_u$  by first forming an approximation with  $f = K(\cdot, \mathbf{y})$  in (13) and subsequently forming a second approximation with the first approximation as  $f$  in (13). From this we obtain the covariance approximation

$$K_u^{\text{FE}}(\mathbf{x}, \mathbf{y}) = \boldsymbol{\phi}(\mathbf{x})^\top \mathbf{A}^{-1} \mathbf{M} \mathbf{A}^{-1} \boldsymbol{\phi}(\mathbf{y}) \approx K_u(\mathbf{x}, \mathbf{y}),$$

where the matrix  $\mathbf{M} \in \mathbb{R}^{n_{\text{FE}} \times n_{\text{FE}}}$  has the elements

$$(\mathbf{M})_{ij} = \int_{\Omega} \int_{\Omega} \phi_i(\mathbf{x}') K(\mathbf{x}', \mathbf{y}') \phi_j(\mathbf{y}') d\mathbf{x}' d\mathbf{y}'.$$

Substituting  $m_u^{\text{FE}}$  for  $m_u$  and  $K_u^{\text{FE}}$  for  $K_u$  yields the approximations

$$m_{u|\mathbf{Y}}^{\text{FE}}(\mathbf{x}) = m_u^{\text{FE}}(\mathbf{x}) + \mathbf{K}_u^{\text{FE}}(\mathbf{x}, X)^\top (\mathbf{K}_u^{\text{FE}}(X, X) + \sigma_\varepsilon^2 \mathbf{I}_n)^{-1} (\mathbf{Y} - \mathbf{m}_u^{\text{FE}}(X)), \quad (14a)$$

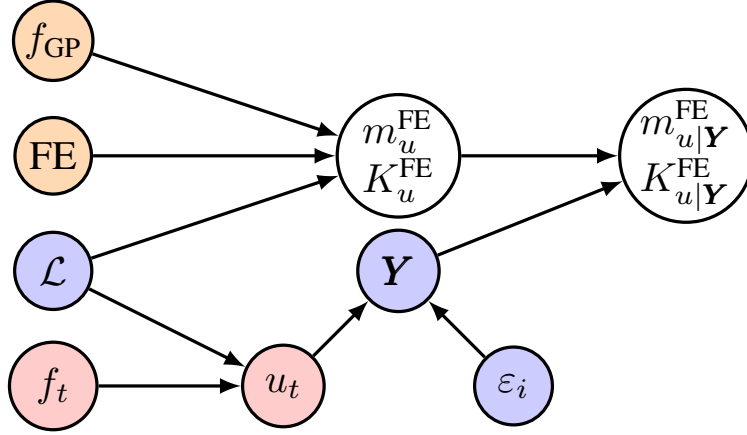
$$K_{u|\mathbf{Y}}^{\text{FE}}(\mathbf{x}, \mathbf{y}) = K_u(\mathbf{x}, \mathbf{y}) - \mathbf{K}_u^{\text{FE}}(\mathbf{x}, X)^\top (\mathbf{K}_u^{\text{FE}}(X, X) + \sigma_\varepsilon^2 \mathbf{I}_n)^{-1} \mathbf{K}_u^{\text{FE}}(\mathbf{y}, X) \quad (14b)$$

to the conditional moments in (10). In practice, it may be tedious or impossible to compute the elements of  $\mathbf{M}$  in closed form. When the supports of  $\phi_i$  are contained within small neighbourhoods of some nodes  $\mathbf{x}_i^{\text{FE}} \in \Omega$ , one may treat the kernel as constant within these supports and employ the approximations

$$\int_{\Omega} \int_{\Omega} \phi_i(\mathbf{x}') K(\mathbf{x}', \mathbf{y}') \phi_j(\mathbf{y}') d\mathbf{x}' d\mathbf{y}' \approx \left( \int_{\Omega} \phi_i(\mathbf{x}') d\mathbf{x}' \right) K(\mathbf{x}_i^{\text{FE}}, \mathbf{x}_j^{\text{FE}}) \left( \int_{\Omega} \phi_j(\mathbf{y}') d\mathbf{y}' \right). \quad (15)$$

The structure of the resulting statFEM approach is displayed in Fig. 3.

Later we will use the following generic assumption on the error of the finite element approximations.



**Fig. 3:** Graphical model of the statFEM approach considered in this article. Blue variables are known to the user while red variables are unknown. Orange variables represent parameters that the user can choose.

**Assumption 2** (Finite element error). Let  $q_1, q_2 > 0$  and  $q = \min\{q_1, q_2\}$ . There exist positive constants  $C_1$  and  $C_2$ , which do not depend on  $n_{\text{FE}}$ , such that

$$\sup_{\mathbf{x} \in \Omega} |m_u(\mathbf{x}) - m_u^{\text{FE}}(\mathbf{x})| \leq C_1 n_{\text{FE}}^{-q_1}, \quad \sup_{\mathbf{x}, \mathbf{y} \in \Omega} |K_u(\mathbf{x}, \mathbf{y}) - K_u^{\text{FE}}(\mathbf{x}, \mathbf{y})| \leq C_2 n_{\text{FE}}^{-q_2}$$

for all  $n_{\text{FE}} \in \mathbb{N}$ .

One should expect that  $q_2 \leq q_1$  because the approximation to the covariance  $K_u$  is obtained via two-fold finite element approximation. In the error analysis of finite element methods it is typical to express error estimates in terms of a characteristic simplex size defined as

$$h = \max_{i \in \{1, \dots, \tilde{n}\}} \min_{\mathbf{x}, \mathbf{y} \in \Omega_i} \|\mathbf{x} - \mathbf{y}\|_2$$

instead of the number of elements as we have done in Assumption 2. If the mesh is quasi-uniform, in that  $h$  tends to zero as  $n_{\text{FE}}^{-1/d}$ , rates expressed in terms of  $h$  are easily converted to rates in terms of  $n_{\text{FE}}$ . Our focus is on the statistical component of the statFEM approach and we do not wish to introduce the machinery that is necessary for precise presentation of estimates on the error of finite elements in the uniform norm. Under certain regularity assumptions one can show that piecewise linear finite elements (see Section 2.2 and Fig. 2) for Poisson's equation  $-\Delta u = f$  in dimension  $d$  converge with the rate  $h^2 |\ln h|$  in the uniform norm, which translates to the rate  $n_{\text{FE}}^{-2/d} \ln n$  in the quasi-uniform case [31]. See [37–39] for additional estimates in the uniform norm. For more casual introductions to error estimates we refer to [4, 30]. An additional motivation for Assumption 2 is the presence of integral approximations, such as that in (15), which are not incorporated into many of the existing error estimates for finite elements.

#### 2.4. The discrepancy term

It is often desirable to include a discrepancy term  $v_{\text{GP}} \sim \text{GP}(m_d, K_d)$  to account for modelling errors. We do this by replacing the induced GP model  $u_{\text{GP}} \sim \text{GP}(m_u, K_u)$  for  $u_t$  with  $u_{\text{GP}} + v_{\text{GP}}$ , so that the full GP model for  $u_t$  is

$$u_{\text{GP}} + v_{\text{GP}} \sim \text{GP}(m_u + m_d, K_u + K_d).$$

Unlike  $u_{\text{GP}}$ , which is induced by the GP prior  $f_{\text{GP}}$  over  $f_t$  and is thus accessible only by solving the PDE (4), the discrepancy term is typically taken to be a GP with some standard covariance kernel, such as a Matérn in (7). Denote  $m_{ud} = m_u + m_d$  and  $K_{ud} = K_u + K_d$ . When the discrepancy term is included, the exact conditional moments in (10) become

$$m_{d;u|Y}(\mathbf{x}) = m_{ud}(\mathbf{x}) + \mathbf{K}_{ud}(\mathbf{x}, X)^\top (\mathbf{K}_{ud}(X, X) + \sigma_\varepsilon^2 \mathbf{I}_n)^{-1} (\mathbf{Y} - \mathbf{m}_{ud}(X)), \quad (16a)$$

$$K_{d;u|Y}(\mathbf{x}, \mathbf{y}) = K_{ud}(\mathbf{x}, \mathbf{y}) - \mathbf{K}_{ud}(\mathbf{x}, X)^\top (\mathbf{K}_{ud}(X, X) + \sigma_\varepsilon^2 \mathbf{I}_n)^{-1} \mathbf{K}_{ud}(X, \mathbf{y}). \quad (16b)$$

When a finite element approximation is employed, we get

$$m_{d;u}^{\text{FE}}(\mathbf{x}) = m_{ud}^{\text{FE}}(\mathbf{x}) + \mathbf{K}_{ud}^{\text{FE}}(\mathbf{x}, X)^{\top} \left( \mathbf{K}_{ud}^{\text{FE}}(X, X) + \sigma_{\varepsilon}^2 \mathbf{I}_n \right)^{-1} (\mathbf{Y} - \mathbf{m}_{ud}^{\text{FE}}(X)), \quad (17a)$$

$$K_{d;u}^{\text{FE}}(\mathbf{x}, \mathbf{y}) = K_{ud}^{\text{FE}}(\mathbf{x}, \mathbf{y}) - \mathbf{K}_{ud}^{\text{FE}}(\mathbf{x}, X)^{\top} \left( \mathbf{K}_{ud}^{\text{FE}}(X, X) + \sigma_{\varepsilon}^2 \mathbf{I}_n \right)^{-1} \mathbf{K}_{ud}^{\text{FE}}(\mathbf{y}, X). \quad (17b)$$

where  $m_{ud}^{\text{FE}} = m_u^{\text{FE}} + m_d$  and  $K_{ud}^{\text{FE}} = K_u^{\text{FE}} + K_d$ .

### 2.5. Extensions

In practice, a variety of additional levels of statistical modelling, or altogether a more complex PDE model, are typically used in statFEM [19]. These can include an additional factor with GP models placed on them on the left-hand side of (4). The standard example being is Poisson's equation

$$-\nabla(e^{\mu}\nabla u) = f \quad (18)$$

in which a GP prior is placed on  $\mu$  (and the exponential ensures positivity of the diffusion coefficient,  $e^{\mu}$ ) in addition to  $f$ . Estimation of various parameters present in model, such as parameters of the covariance kernel  $K$  (e.g.,  $\sigma$ ,  $\ell$  and  $\nu$  of a Matérn kernel), is also common. If GP priors are placed on  $\mu$  and  $f$  in the model (18) or its generalisation of some form, the prior induced on  $u$  is no longer a GP. This will render most of the theoretical tools that we use inoperative, and this generalisation is not accordingly pursued here. While there is some recent theoretical work on parameter estimation in Gaussian process regression for deterministic data-generating functions and its effect on posterior rates of convergence and reliability [23, 24, 41, 46], the results that have been obtained are not yet sufficiently general or flexible to be useful in the PDE setting considered here. Reliability refers to the quality of the posterior standard deviation as a measure of approximation error. Informally speaking a GP model is reliable if the posterior standard deviation decays with the same rate as the approximation error as  $n \rightarrow \infty$ . See [40] for some work on reliability and its connection to model parameter estimation in the context of the Gaussian sequence model.

## 3. Main Results

This section contains the main results of this article. The results provide rates of contraction, as  $n \rightarrow \infty$ , of the expectation (with respect to the observation noise distribution) of the  $L^2(\Omega)$ -norm between the true source term  $u_t$  and the GP conditional means in (10a), (14a), and (17a). All proofs are deferred to Section 6. The results are expressed in terms of the fill-distance

$$h_{X,\Omega} = \sup_{\mathbf{x} \in \Omega} \min_{i \in \{1, \dots, n\}} \|\mathbf{x} - \mathbf{x}_i\|_2 \quad (19)$$

of the set of points  $X \subset \Omega$ . The fill-distance cannot tend to zero with a rate faster than  $n^{-1/d}$ , a rate which is achieved by, for example, uniform Cartesian grids.

### 3.1. Function spaces

Let  $D^{\alpha}f$  denote the weak derivative of order  $\alpha \in \mathbb{N}_0^d$  of any sufficiently regular function  $f: \Omega \rightarrow \mathbb{R}$ . Let  $k \in \mathbb{N}_0$ . The Sobolev space  $H^k(\Omega)$  consists of functions for which  $D^{\alpha}f$  exists for all  $|\alpha| \leq k$  and the norm

$$\|f\|_{H^k(\Omega)} = \left( \sum_{|\alpha| \leq k} \|D^{\alpha}f\|_{L^2(\Omega)}^2 \right)^{1/2}$$

is finite. The Hölder space  $C^{k,\alpha}(\Omega)$  consists of functions which are  $k \in \mathbb{N}_0$  times differentiable on  $\Omega$  and whose derivatives of order  $k$  are Hölder continuous with exponent  $\alpha \in (0, 1]$ .

Some of our assumptions are expressed in terms of reproducing kernel Hilbert spaces (RKHSs). By the classical Moore–Aronszajn theorem [3, p. 19] every positive-semidefinite kernel  $K: \Omega \times \Omega \rightarrow \mathbb{R}$  induces a unique RKHS,  $\mathcal{H}(K, \Omega)$ , which consists of functions  $f: \Omega \rightarrow \mathbb{R}$  and is equipped with an inner product  $\langle \cdot, \cdot \rangle_K$  and norm  $\|\cdot\|_K$ . Two fundamental properties of this space are that (i)  $K(\cdot, \mathbf{x})$  is an element of  $\mathcal{H}(K, \Omega)$  for every  $\mathbf{x} \in \Omega$  and (ii) the reproducing property

$$\langle f, K(\cdot, \mathbf{x}) \rangle_K = f(\mathbf{x}), \quad f \in \mathcal{H}(K, \Omega) \text{ and } \mathbf{x} \in \Omega. \quad (20)$$

Our results use an assumption that  $\mathcal{H}(K, \Omega)$  is norm-equivalent to a Sobolev space.



**Definition 1** (Norm-equivalence). The RKHS  $\mathcal{H}(K, \Omega)$  is norm-equivalent to the Sobolev space  $H^k(\Omega)$ , denoted  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$ , if  $\mathcal{H}(K, \Omega) = H^k(\Omega)$  as sets and there exist positive constants  $C_K$  and  $C'_K$  such that

$$C_K \|f\|_{H^k(\Omega)} \leq \|f\|_K \leq C'_K \|f\|_{H^k(\Omega)} \quad (21)$$

for all  $f \in \mathcal{H}(K, \Omega)$ .

The RKHS of a Matérn kernel of smoothness  $\nu$  in (7) is norm-equivalent to  $H^{\nu+d/2}(\Omega)$ . If  $k > d/2$ , the Sobolev embedding theorem ensures that any kernel which is norm-equivalent to  $H^k(\Omega)$  is continuous and that all functions in its RKHS are continuous. From now on we assume that  $\mathcal{H}(K, \Omega) \subset \mathcal{H}_L(\Omega)$ , which is to say that the PDE in (4) admits a unique classical solution for every  $f \in \mathcal{H}(K, \Omega)$ .

### 3.2. Exact posterior

Our first result concerns an ideal statFEM that uses no finite element discretisation. The relevant posterior moments are given in (10).

**Theorem 1.** *Let  $k > d/2$  and suppose that Assumption 1 holds and  $c \leq 0$ . If  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$ ,  $m \in H^k(\Omega)$ , and  $f_i \in H^k(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$ , then there is a positive constant  $C_1$  that is independent of  $X$  such that*

$$\mathbb{E} \left[ \|u_t - m_{u|Y}\|_{L^2(\Omega)} \right] \leq C_1 \left( h_{X,\Omega}^{k+2} \sqrt{n} + h_{X,\Omega}^{d/2} n^{d/(4k)} \right). \quad (22)$$

If  $h_{X,\Omega} = O(n^{-1/d})$ , then there is a positive constant  $C_2$  that is independent of  $X$  such that

$$\mathbb{E} \left[ \|u_t - m_{u|Y}\|_{L^2(\Omega)} \right] \leq C_2 n^{-1/2+d/(4k)} \quad (23)$$

for all  $n \geq 1$ .

**Proof sketch.** We first establish in Theorem 8 that the RKHS of  $K_u$  is a sub-space of  $H^{k+2}(\Omega)$  with an equivalent norm. The rest of the proof mirrors that of Corollary 3 in [49] with minor modifications. The main ingredients are a sampling inequality from [2] and bounds on the concentration function similar to those in [44].  $\square$

**Remark 1.** The mini-max optimal rate for regression in  $H^k(\Omega)$  is  $n^{-k/(2k+d)}$  [43, Chapter 2]. Since

$$\frac{1}{2} - \frac{d}{4k} \leq \frac{1}{2} - \frac{d}{4k+2d} = \frac{k}{2k+d},$$

the rate (23) that we have proved is slightly sub-optimal.

Minor modifications to the proof of Theorem 2 yield a variant in which it is assumed that the observations are noiseless. That is,  $y_i = u_t(\mathbf{x}_i)$  for  $i \in \{1, \dots, n\}$ . One can informally interpret this as the case  $\sigma_\varepsilon = 0$ .

**Theorem 2.** *Consider the setting of Theorem 1 but assume that the observations are noiseless. Then there is a positive constant  $C_1$  that is independent of  $X$  such that*

$$\mathbb{E} \left[ \|u_t - m_{u|Y}\|_{L^2(\Omega)} \right] \leq C_1 h_{X,\Omega}^{k+2} \left( \|u_t\|_{H^{k+2}(\Omega)} + \|m_u\|_{H^{k+2}(\Omega)} \right).$$

If  $h_{X,\Omega} = O(n^{-1/d})$ , then there is a positive constant  $C_2$  that is independent of  $X$  such that

$$\mathbb{E} \left[ \|u_t - m_{u|Y}\|_{L^2(\Omega)} \right] \leq C_2 n^{-(k+2)/d} \left( \|u_t\|_{H^{k+2}(\Omega)} + \|m_u\|_{H^{k+2}(\Omega)} \right)$$

for all  $n \geq 1$ .

It is straightforward to prove a similar noiseless variant of Theorem 4. However, Theorem 2 does not extend to posteriors obtained via finite element discretisation because in the noiseless case it is difficult to bound the difference of posterior means corresponding to different kernels in terms of some distance between the said kernels.

### 3.3. Finite element posterior

Next we turn to the analysis of the effect of finite element discretisation and consider the posterior mean in (14a). A straightforward combination of Theorem 1 and Proposition 4 yields an error estimate that combines the errors from GP regression and finite element discretisation.

**Theorem 3.** *Let  $k > d/2$ . Suppose that Assumptions 1 and 2 hold and that  $c \leq 0$ . If  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$ ,  $m \in H^k(\Omega)$ ,  $f_t \in H^k(\mathbb{R}^d) \cap C^k(\mathbb{R}^d)$ , and  $h_{X,\Omega} = O(n^{-1/d})$ , then there is a positive constant  $C$  that is independent of  $X$  such that*

$$\mathbb{E} \left[ \|u_t - m_{u|\mathbf{Y}}^{\text{FE}}\|_{L^2(\Omega)} \right] \leq C \left( n^{-1/2+d/(4k)} + (n_{\text{FE}}^{-q} + \sigma_\varepsilon^2) \sigma_\varepsilon^{-2} (\|f_t\|_{L^\infty(\Omega)} + \sigma_\varepsilon) n_{\text{FE}}^{-q} n^{3/2} \right) \quad (24)$$

for all  $n \geq 1$ .

**Remark 2.** To obtain the best possible rate of convergence in terms of  $n$  in (24), we could set

$$n_{\text{FE}} = n^{(2-d/(4k))/q}. \quad (25)$$

By incorporating all other terms in the constant  $C$ , we then obtain the error estimate

$$\mathbb{E} \left[ \|u_t - m_{u|\mathbf{Y}}^{\text{FE}}\|_{L^2(\Omega)} \right] \leq C \left( n^{-1/2+d/(4k)} + n_{\text{FE}}^{-q} n^{3/2} \right) = 2C n^{-1/2+d/(4k)},$$

which is equal to the bound in (23) up to a constant factor.

Practical application of Remark 2 is difficult because, while (25) yields the best possible polynomial rate in (24), what one would actually like to obtain is the smallest possible right-hand side in (24). But finding  $n_{\text{FE}}$  that minimises the right-hand side is difficult because the constant factors involved are rarely, if ever, available.

### 3.4. Inclusion of a discrepancy term

Finally, we consider inclusion of a discrepancy term as described in Section 2.4. The following two theorems concern the posterior means in (16a) and (17a). In these theorems it is assumed that the points are quasi-uniform, which means that there is  $C_{\text{qu}} > 0$  such that

$$q_X \leq h_{X,\Omega} \leq C_{\text{qu}} q_X,$$

where  $h_{X,\Omega}$  is the fill-distance in (19) and  $q_X$  is the separation radius

$$q_X = \frac{1}{2} \min_{i \neq j} \|\mathbf{x}_i - \mathbf{x}_j\|_2.$$

Quasi-uniformity implies that the mesh ratio  $\rho_{X,\Omega} = h_{X,\Omega}/q_X$  is uniformly bounded from above and that  $h_{X,\Omega} = O(n^{-1/d})$ ; see Chapter 14 in [48].

**Theorem 4.** *Let  $k_1 \geq r-2 \geq k_2 > d/2$  and suppose that Assumption 1 holds with  $k = k_1$  and  $c \leq 0$ . If  $\mathcal{H}(K, \Omega) \simeq H^{k_1}(\Omega)$ ,  $\mathcal{H}(K_d, \Omega) \simeq H^r(\Omega)$ ,  $m_d \in H^{k_2+2}(\Omega)$ ,  $m \in H^{k_2}(\Omega)$ , and  $f_t \in H^{k_2}(\mathbb{R}^d) \cap C^{k_2}(\mathbb{R}^d)$ , then there is a positive constant  $C_1$  that is independent of  $X$  such that*

$$\mathbb{E} \left[ \|u_t - m_{d;u|\mathbf{Y}}\|_{L^2(\Omega)} \right] \leq C_1 \left( h_{X,\Omega}^{k_2+2} \rho_{X,\Omega}^{r-k_2-2} + \sqrt{n} h_{X,\Omega}^r + n^{\kappa(k_2,r)} h_{X,\Omega}^{d/2} \right). \quad (26)$$

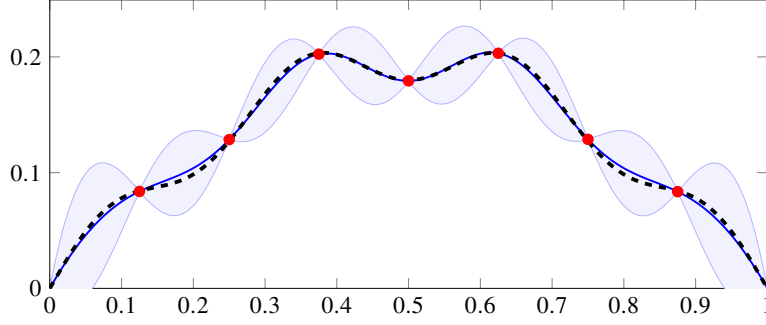
The constant  $\kappa(k_2, r) \leq 1/2$  is given in (53). If the points are quasi-uniform, there is a positive constant  $C_2$  independent of  $X$  such that

$$\mathbb{E} \left[ \|u_t - m_{d;u|\mathbf{Y}}\|_{L^2(\Omega)} \right] \leq C_2 \left( n^{-(k_2+2)/d} + n^{-r/d+1/2} + n^{-1/2+\kappa(k_2,r)} \right) \quad (27)$$

for all  $n \geq 1$ .

**Proof sketch.** We proceed as in the proof of Theorem 1. However, now the assumption  $k_1 \geq r-2$  ensures that the  $K_u$  is not rougher than  $K_d$ , which means that the regularity of the RKHS of  $K_{ud} = K_u + K_d$  is determined by the regularity of  $K_d$ . In this case we can directly apply Theorem 2 in [49].  $\square$

Combining Theorem 4 with Proposition 4 allows including the effect of finite element discretisation.



**Fig. 4:** The statFEM conditional mean (blue) and 95% credible interval given some noisy data (red points). The dashed black line is the true system response.

**Theorem 5.** Let  $k_1 \geq r - 2 \geq k_2 > d/2$ . Suppose that Assumptions 1 (with  $k = k_1$ ) and 2 hold and that  $c \leq 0$ . If  $\mathcal{H}(K, \Omega) \simeq H^{k_1}(\Omega)$ ,  $\mathcal{H}(K_d, \Omega) \simeq H^r(\Omega)$ ,  $m_d \in H^{k_2+2}(\Omega)$ ,  $m \in H^{k_2}(\Omega)$ ,  $f_t \in H^{k_2}(\mathbb{R}^d) \cap C^{k_2}(\mathbb{R}^d)$ , and the points are quasi-uniform, then there is a positive constant  $C$  independent of  $X$  such that

$$\mathbb{E} \left[ \|u_t - m_{d;u|Y}^{\text{FE}}\|_{L^2(\Omega)} \right] \leq C \left( n^{-(k_2+2)/d} + n^{-r/d+1/2} + n^{-1/2+\kappa(k_2,r)} + n_{\text{FE}}^{-q} n^{3/2} \right) \quad (28)$$

for all  $n \geq 1$ . The constant  $\kappa(k_2, r) \leq 1/2$  is given in (53).

Unlike the results in Sections 3.2 and 3.3, these theorems are valid also when the smoothness of the source term is misspecified. That is, in Theorems 4 and 5 it is possible that  $k_1$ , the smoothness of the kernel  $K$  which specifies the prior for the source term, is larger than the smoothness,  $k_2$ , of the true source term  $f_t$ . Such misspecification results for GP regression in different settings can be found in, for example, [21, 24, 41, 49]. The effectiveness of the discrepancy term in Theorems 4 and 5 is due to the condition  $k_1 \geq r - 2 \geq k_2$ . Because the prior on  $u_t$  has regularity  $k_1 + 2 \geq r$ , the first inequality states that the discrepancy term is rougher than the  $u_{\text{GP}}$ . It is a general phenomenon in approximation and regression that the roughest component dominates, which is precisely what one can observe in the bounds (26)–(28). The inequality  $r - 2 \geq k_2$  states that  $u_t$  is rougher than the discrepancy term, which is to say that the discrepancy term oversmooths the truth. Oversmoothing is the scenario in which convergence rates tend to adapt to misspecification (see the references above).

#### 4. Numerical example

In this section we investigate the convergence of the posterior mean  $m_{u|Y}^{\text{FE}}$  to the true system response  $u_t$  for different values of the kernel smoothness parameter  $k$ . We consider the one-dimensional Poisson's equation

$$-u'' = f \quad \text{in } \Omega = (0, 1), \quad u(0) = u(1) = 0. \quad (29)$$

The true source term is set as the function

$$f_t(x) = \frac{\pi^2}{5} \sin(\pi x) + \frac{49\pi^2}{50} \sin(7\pi x). \quad (30)$$

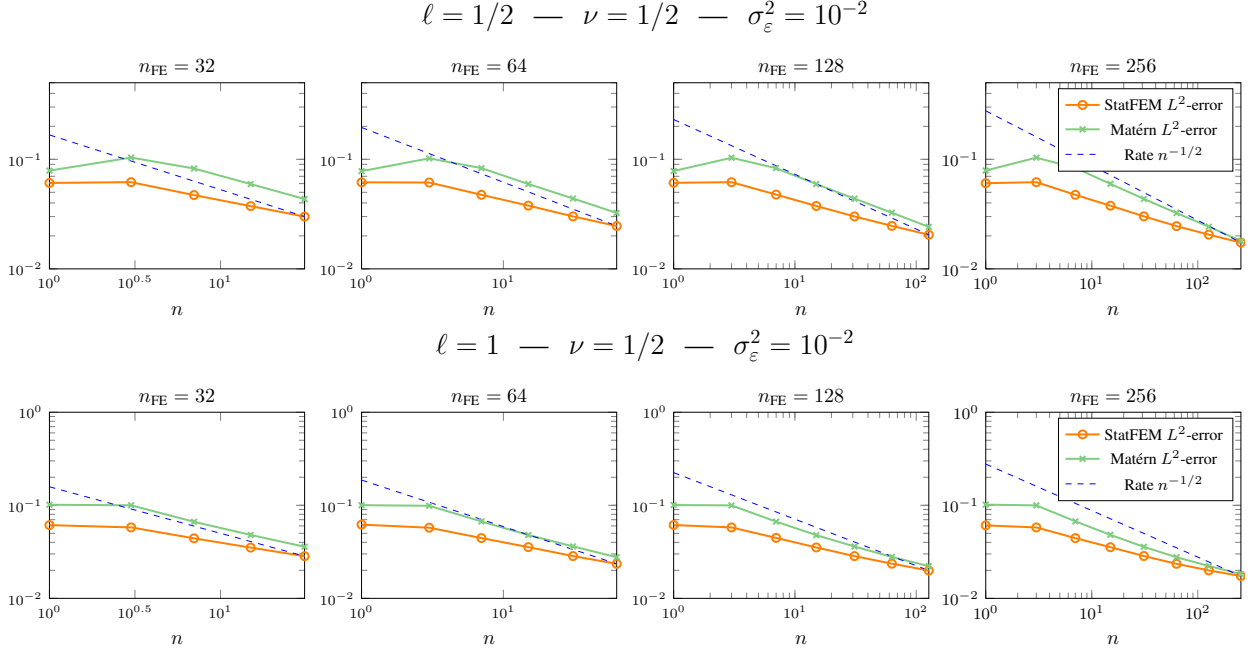
The respective true system response is given in closed form by

$$u_t(x) = \frac{1}{5} \sin(\pi x) + \frac{1}{50} \sin(7\pi x). \quad (31)$$

A similar example was used in [19].

For the source term, we use a zero-mean Gaussian prior  $f_{\text{GP}} \sim \text{GP}(0, K)$  with the Matérn covariance kernel (7). We use the kernel hyperparameters

$$\ell \in \left\{ \frac{1}{2}, 1 \right\}, \quad \nu \in \left\{ \frac{1}{2}, \frac{5}{2} \right\}.$$

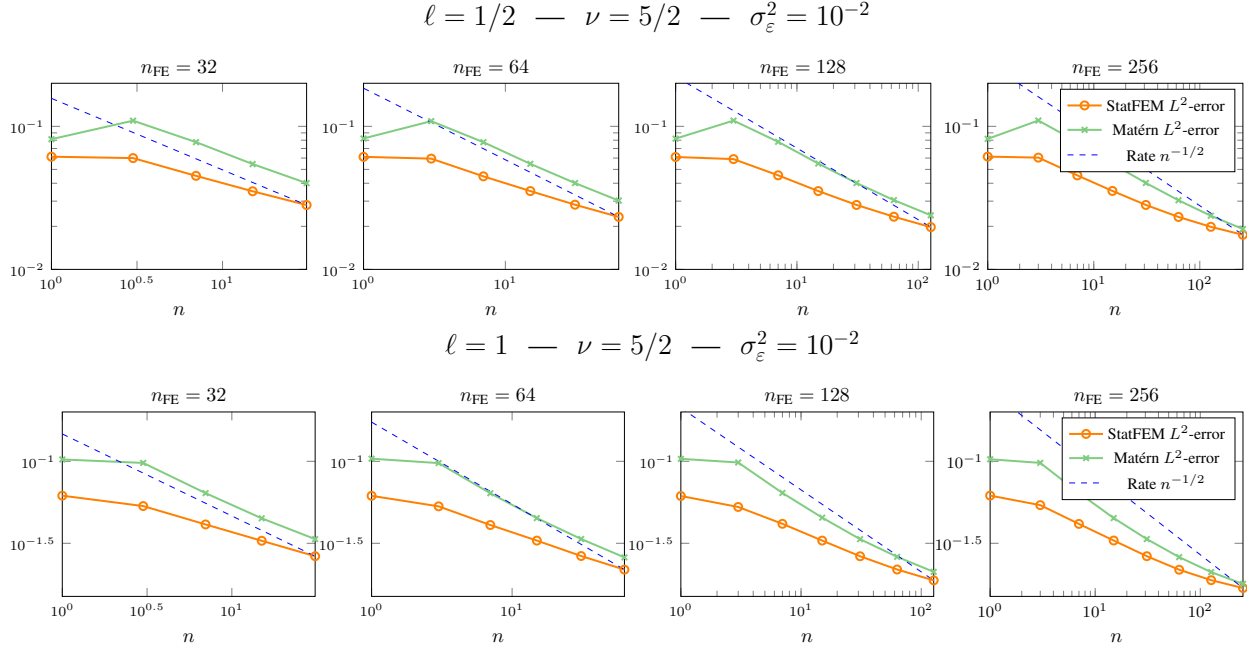


**Fig. 5:** Convergence of the statFEM conditional mean (StatFEM  $L^2$ -error)  $m_{u|Y}^{\text{FE}}$  to  $u_t$ , true system response (31) of the Poisson's equation (29) with the source term (30). For comparison we also plot the convergence of the conditional mean when  $u_t$  is modelled as a GP with a Matérn kernel (Matérn  $L^2$ -error). The plots show empirical approximations to expected  $L^2$ -errors over 10 000 noise realisations. The statFEM prior for the source term  $f_t$  is a Matérn with parameters  $\ell \in \{1/2, 1\}$  and  $\nu = 1/2$ . The Matérn prior placed on  $u_t$  has the same length-scales and smoothness  $\nu + 2$ , which ensures that both the induced statFEM prior and the Matérn prior for  $u_t$  have the same regularity. Noise variance is  $\sigma_\varepsilon^2 = 10^{-2}$ .

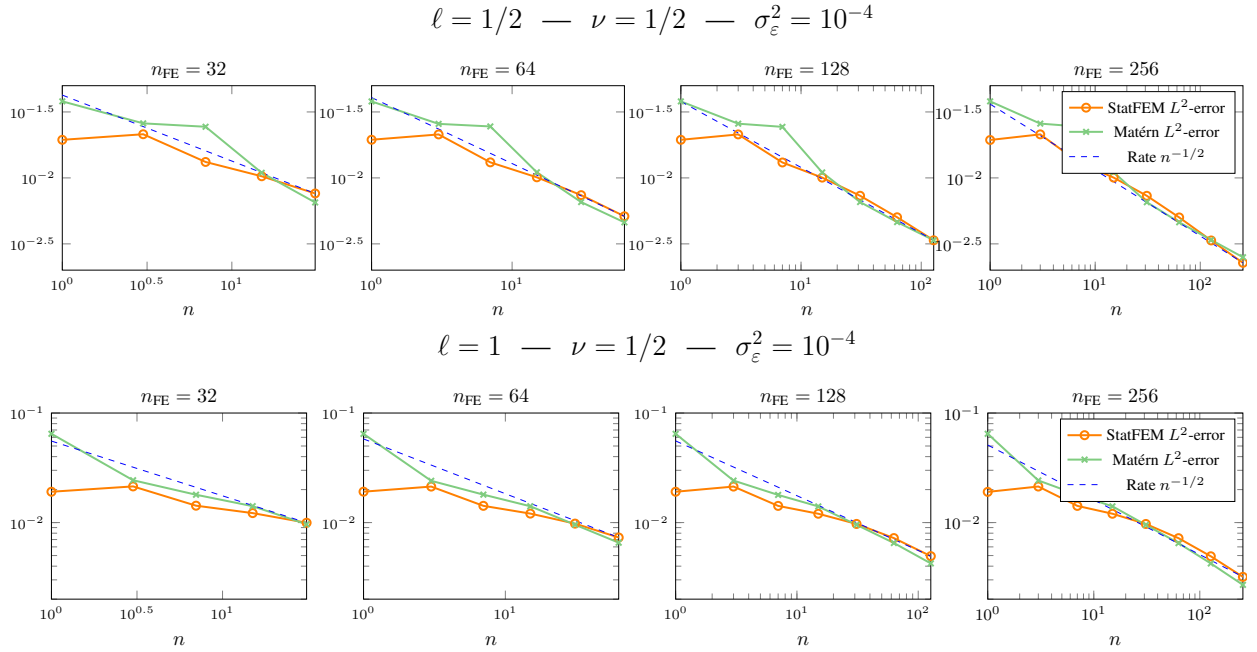
As explained in Section 3.1, the values of  $\nu$  correspond to the values  $k \in \{1, 3\}$  of the RKHS smoothness parameter. In order to facilitate comparison with standard GP regression based on a Matérn kernel, we set scaling parameter  $\sigma$  such that the maximum of  $K_u$  equals one (see Fig. 1). For each  $k$ , the true source term in (30) is an element of the RKHS and of  $C^k(\Omega)$ . The selection of the hyperparameters could be automated by considering the marginal likelihood or cross-validation [36, Chapter 5]. For finite element analysis we use the standard piecewise linear basis functions (see Fig. 2 for their bivariate versions) centered at  $n_{\text{FE}} \in \{32, 64, 128, 256\}$  uniformly placed points on  $\Omega = (0, 1)$ . To compute the conditional mean  $m_{u|Y}^{\text{FE}}$ , we use the approximation in (15); see Section 2.2 in [19] for more details. Observations of  $u_t$  at  $n \in \{2^1 - 1, 2^2 - 1, \dots, n_{\text{FE}} - 1\}$  uniformly placed points are corrupted by Gaussian noise with variances  $\sigma_\varepsilon^2 \in \{10^{-2}, 10^{-4}\}$ . For illustration, the true system response and the finite element approximation to the conditional mean  $m_{u|Y}^{\text{FE}}$  and the corresponding 95% credible interval are shown in Fig. 4 for  $\nu = 1/2$ ,  $\ell = 1/2$ ,  $n_{\text{FE}} = 2048$ , and  $n = 7$ .

Convergence results are depicted in Figs. 5–8. In these results the  $L^2$ -norm is approximated by numerical quadrature and the expectation by an average over 10 000 independent observation noise realisations. For each  $\nu$  and  $n_{\text{FE}}$  we also plot the  $L^2$ -error of standard GP regression when  $u_t$  is directly modelled as a purely data-driven GP whose kernel is a Matérn with smoothness  $\nu + 2$  and parameters  $\sigma_{\text{Matérn}} = 1$  and  $\ell_{\text{Matérn}} = \ell$ . The selection of the smoothness parameter of the Matérn prior for  $u_t$  corresponds to the smoothness of the induced prior  $u_{\text{GP}}$  in statFEM. Being purely data-driven, this Matérn model for  $u_t$  does not incorporate the boundary conditions or other structural characteristics.

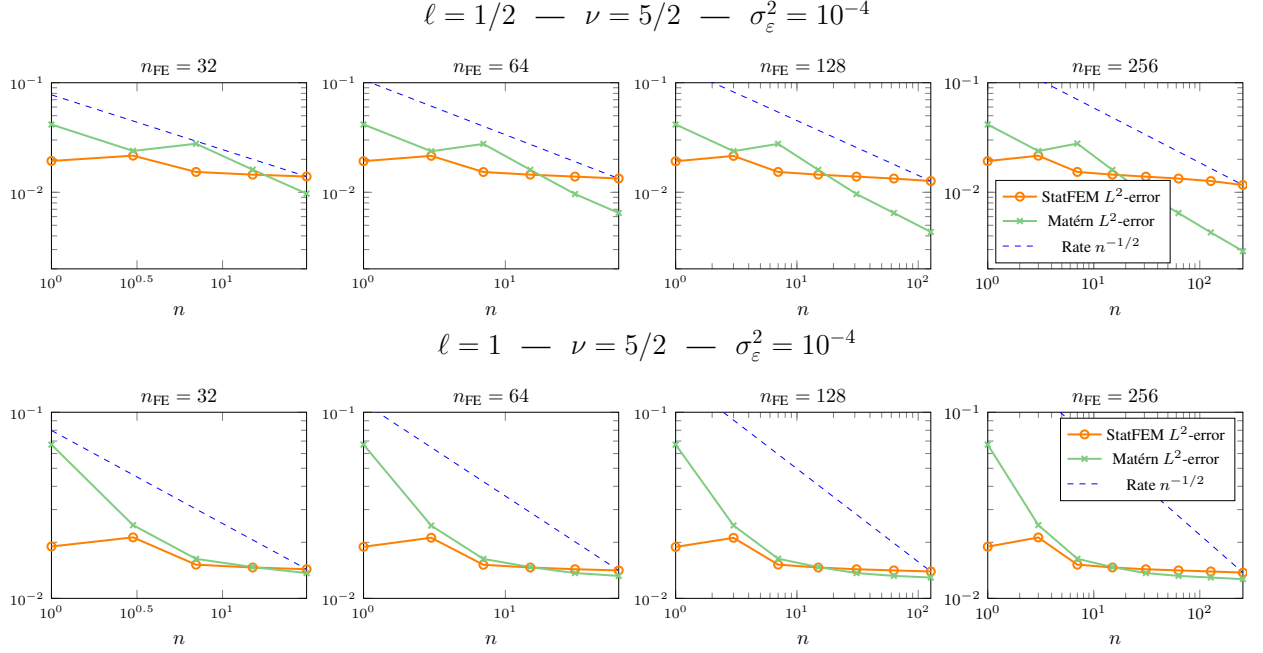
We see that statFEM outperforms the Matérn model in Figs. 5 to 7, particularly for small  $n$ . This is to be expected as the prior dominates when there is little data. In Fig. 8 ( $\ell \in \{1/2, 1\}$ ,  $\nu = 5/2$  and  $\sigma_\varepsilon^2 = 10^{-4}$ ), statFEM exhibits clear saturation. However, as the Matérn model behaves similarly when  $\ell = 1$ ,  $\nu = 5/2$  and  $\sigma_\varepsilon^2 = 10^{-4}$ , it seems that the saturation effect is not specific to statFEM in this example. The plots also show that statFEM works well even when the number of finite element nodes,  $n_{\text{FE}}$ , is small. This is important because, especially in higher dimensions, the number of data points will be significantly smaller than the number of finite element nodes so that it will become even more important to encode the PDE and its boundary conditions.



**Fig. 6:** Convergence of the statFEM conditional mean (StatFEM  $L^2$ -error)  $m_{u|Y}^{\text{FE}}$  to  $u_t$ , true system response (31) of the Poisson's equation (29) with the source term (30). For comparison we also plot the convergence of the conditional mean when  $u_t$  is modelled as a GP with a Matérn kernel (Matérn  $L^2$ -error). The plots show empirical approximations to expected  $L^2$ -errors over 10 000 noise realisations. The statFEM prior for the source term  $f_t$  is a Matérn with parameters  $\ell \in \{1/2, 1\}$  and  $\nu = 5/2$ . The Matérn prior placed on  $u_t$  has the same length-scales and smoothness  $\nu + 2$ , which ensures that both the induced statFEM prior and the Matérn prior for  $u_t$  have the same regularity. Noise variance is  $\sigma_\varepsilon^2 = 10^{-2}$ .



**Fig. 7:** Convergence of the statFEM conditional mean (StatFEM  $L^2$ -error)  $m_{u|Y}^{\text{FE}}$  to  $u_t$ , true system response (31) of the Poisson's equation (29) with the source term (30). For comparison we also plot the convergence of the conditional mean when  $u_t$  is modelled as a GP with a Matérn kernel (Matérn  $L^2$ -error). The plots show empirical approximations to expected  $L^2$ -errors over 10 000 noise realisations. The statFEM prior for the source term  $f_t$  is a Matérn with parameters  $\ell \in \{1/2, 1\}$  and  $\nu = 1/2$ . The Matérn prior placed on  $u_t$  has the same length-scales and smoothness  $\nu + 2$ , which ensures that both the induced statFEM prior and the Matérn prior for  $u_t$  have the same regularity. Noise variance is  $\sigma_\varepsilon^2 = 10^{-4}$ .



**Fig. 8:** Convergence of the statFEM conditional mean (StatFEM  $L^2$ -error)  $m_{u|Y}^{\text{FE}}$  to  $u_t$ , true system response (31) of the Poisson’s equation (29) with the source term (30). For comparison we also plot the convergence of the conditional mean when  $u_t$  is modelled as a GP with a Matérn kernel (Matérn  $L^2$ -error). The plots show empirical approximations to expected  $L^2$ -errors over 10 000 noise realisations. The statFEM prior for the source term  $f_t$  is a Matérn with parameters  $\ell \in \{1/2, 1\}$  and  $\nu = 5/2$ . The Matérn prior placed on  $u_t$  has the same length-scales and smoothness  $\nu + 2$ , which ensures that both the induced statFEM prior and the Matérn prior for  $u_t$  have the same regularity. Noise variance is  $\sigma_\varepsilon^2 = 10^{-4}$ .

## 5. Concluding Remarks

We have analysed a particular formulation of the statFEM approach in a deterministic setting with generic points and finite elements. A different set of assumptions could be used—the practical relevance of these assumptions would likely depend much on the application and whether or not the user views the data-generating process as an actual Gaussian process or some unknown deterministic function. Settings that we believe could or could not be analysed using similar or related techniques as those in this article include the following:

- We have considered a “mixed” case in which a GP is used to model a deterministic function. But one could alternatively assume that  $f_t$ , and consequently  $u_t$ , is a GP (or some other stochastic process) and proceed from there. This is how statFEM is formulated in [19].
- Distribution of the points  $\mathbf{x}_1, \dots, \mathbf{x}_n$  where the measurement data are obtained is quite generic in this article in that no reference is made to how these points might be selected or sampled and all results are formulated in terms of the fill-distance  $h_{X,\Omega}$  or, in the quasi-uniform case,  $n$ . In applications the points may be sampled randomly from some distribution on  $\Omega$ . We refer to [5, Theorem 1] and [28] for related results concerning random points.
- To remove the assumption that the measurement data are noisy is likely to be challenging if one is interested in including the effect of the finite element discretisation. It is straightforward to derive versions of Theorems 1 and 4 in the noiseless case (see Theorem 2), but no other result in Section 6.5 generalises readily. The reason is the presence of the factor  $\sigma^{-2}$  in (49) which is used to prove all theorems concerning finite element discretisations: if  $\sigma = 0$ , this bound is rendered meaningless.
- As already mentioned in Section 3.3, the bounds include a variety of non-explicit constants. We do not believe that the constants are computable in all but perhaps the simplest of special cases.

- Although desirable for computational reasons, the Gaussian process framework that we have used is restrictive. It may be possible to derive rates under more general assumptions [16]. However, the proof technique would have to be different as our proof relies on sampling inequalities and a connection between the GP posterior and approximation in an RKHS.
- An interesting problem that we have not encountered in the GP or kernel literature would be to recover  $f_t$  from observations of  $u_t$ . Because  $f_t = \mathcal{L}^{-1}u_t$ , this is a generalised regression problem in which one tries to recover a function from linear information. Alternatively, this is a linear inverse problem in which one recovers the source from the response. If a GP prior with covariance  $K$  is placed on  $f_t$ , conditioning on evaluations of  $u_t$  at  $X$  gives rise to a posterior that involves the kernel matrix  $\mathbf{K}_u(X, X)$ ; see [42]. Some of the results derived in this article could therefore be useful.

## 6. Proofs

This section contains the proofs of the theorems in Section 3. We first collect and derive a number of auxiliary results that we use to prove the error estimates. These results are of four types: (i) standard regularity results for solutions of elliptic PDEs; (ii) results on the RKHS of the kernel  $K_u$ ; (iii) sampling inequalities; and (iv) results related to the concentration function and small ball probabilities of Gaussian measures. We use the notation  $C = C(\theta_1, \dots, \theta_p)$  to indicate that a constant  $C$  depends only on some parameters  $\theta_1, \dots, \theta_p$ .

### 6.1. Regularity results for elliptic pdes

Recall the function spaces from Section 3.1. In Assumption 1 and Theorem 7, the boundary of  $\Omega$  is required to have certain smoothness. The boundary  $\partial\Omega$  is said to be  $C^k$  (or  $C^{k,\alpha}$ ) if for each  $\mathbf{x}_0 \in \partial\Omega$  there is an open ball  $B$  centered at  $\mathbf{x}_0$  and a surjection  $\psi$  from  $B$  to some  $D \subset \mathbb{R}^d$  such that (a)  $\psi(B \cap \Omega) \subset \mathbb{R}_+^d$ , (b)  $\psi(B \cap \partial\Omega) \subset \partial\mathbb{R}_+^d$ , (c)  $\psi \in C^k(\Omega)$  [resp.  $\psi \in C^{k,\alpha}(\Omega)$ ], and (d)  $\psi^{-1} \in C^k(D)$  [resp.  $\psi \in C^{k,\alpha}(D)$ ]. Certain standard regularity results and estimates play a crucial role in the derivation of our results. The following regularity theorem can be found in, for example, [10, Theorem 5 in Section 6.3].

**Theorem 6.** *Consider the elliptic PDE given in (4). Let  $k \in \mathbb{N}_0$  and suppose that Assumption 1 holds. If  $f \in H^k(\Omega) \cap \mathcal{H}_{\mathcal{L}}(\Omega)$ , then  $u \in H^{k+2}(\Omega)$  and there is a constant  $C = C(k, \Omega, \mathcal{L})$  such that*

$$\|u\|_{H^{k+2}(\Omega)} \leq C \|f\|_{H^k(\Omega)}.$$

The following boundedness result is a combination of the a priori bound in Theorem 3.7 and the Schauder regularity result in Theorem 6.14 of [17].

**Theorem 7.** *Consider the elliptic PDE in (4). Suppose that  $\partial\Omega$  is  $C^{2,\alpha}$ ,  $a_{ij}, b_i, c \in C^{0,\alpha}(\bar{\Omega})$  for all  $i, j \in \{1, \dots, d\}$  and some  $\alpha \in (0, 1)$  and  $c \leq 0$ . If  $f \in C^{0,\alpha}(\bar{\Omega})$ , then  $u \in C^{2,\alpha}(\bar{\Omega})$  and there is a constant  $C = C(\Omega, \mathcal{L})$  such that*

$$\|u\|_{L^\infty(\Omega)} \leq C \|f\|_{L^\infty(\Omega)}.$$

Note that Assumption 1 implies the regularity assumptions in Theorem 7.

### 6.2. Transformed reproducing kernel Hilbert spaces

The following lemma justifies the assumption in (8) that  $\mathcal{L}_{\mathbf{x}}^{-1}K(\cdot, \mathbf{x})$  is an element of  $\mathcal{H}(K, \Omega) \subset \mathcal{H}_{\mathcal{L}}(\Omega)$  for every  $\mathbf{x} \in \Omega$ . Let  $\delta_{\mathbf{x}}$  be the point evaluation functional at  $\mathbf{x} \in \Omega$ , which is to say that  $\delta_{\mathbf{x}}(f) = f(\mathbf{x})$ . As earlier, whenever there is a risk of ambiguity we use subscripts to denote the variable to which a functional or an operator applies to. That is,

$$\mathcal{A}f(\mathbf{x}) = \mathcal{A}_{\mathbf{x}}f(\mathbf{x}) = (\delta_{\mathbf{x}} \circ \mathcal{A})_{\mathbf{x}'}f(\mathbf{x}')$$

for any operator  $\mathcal{A}$ .

**Lemma 1.** *Let  $k > d/2$ . Suppose that  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$  and that Assumption 1 holds. Then the functional  $\delta_{\mathbf{x}} \circ \mathcal{L}^{-1}$  is bounded on  $\mathcal{H}(K, \Omega)$  and  $\mathcal{L}_{\mathbf{x}}^{-1}K(\cdot, \mathbf{x}) \in \mathcal{H}(K, \Omega)$  for every  $\mathbf{x} \in \Omega$ .*

**Proof.** Let  $\mathbf{x} \in \Omega$  and set  $\ell_{\mathbf{x}} = \delta_{\mathbf{x}} \circ \mathcal{L}^{-1}$ . Since  $\mathcal{H}(K, \Omega)$  is norm-equivalent to  $H^k(\Omega)$  and  $H^{k+2}(\Omega)$  is continuously embedded in  $C(\Omega)$ , it follows from Theorem 6 that

$$|\ell_{\mathbf{x}}(f)| = |u(\mathbf{x})| \leq \|u\|_{L^\infty(\Omega)} \leq C_1 \|u\|_{H^{k+2}(\Omega)} \leq C_1 C_2 \|f\|_{H^k(\Omega)} \leq C_1 C_2 C_K^{-1} \|f\|_K$$

for any  $f \in \mathcal{H}(K, \Omega)$  and constants  $C_1 = C(k, \Omega)$  and  $C_2 = C(k, \Omega, \mathcal{L})$ . This proves that  $\ell_{\mathbf{x}}$  is bounded on  $\mathcal{H}(K, \Omega)$ . Because  $\ell_{\mathbf{x}}$  is bounded, it follows from the Riesz representation theorem that there exists a unique function  $l_{\mathbf{x}} \in \mathcal{H}(K, \Omega)$  such that  $\ell_{\mathbf{x}}(f) = \langle f, l_{\mathbf{x}} \rangle_K$  for every  $f \in \mathcal{H}(K, \Omega)$ . Setting  $f = K(\cdot, \mathbf{y})$  and using the reproducing property (20) we get, for any  $\mathbf{y} \in \Omega$ ,

$$\mathcal{L}_{\mathbf{x}}^{-1} K(\mathbf{y}, \mathbf{x}) = \ell_{\mathbf{x}}(K(\mathbf{y}, \mathbf{x})) = \langle K(\mathbf{y}, \cdot), l_{\mathbf{x}} \rangle_K = l_{\mathbf{x}}(\mathbf{y}).$$

That is,  $\mathcal{L}_{\mathbf{x}}^{-1} K(\cdot, \mathbf{x}) = l_{\mathbf{x}} \in \mathcal{H}(K, \Omega)$ .  $\square$

Next we want to understand how the RKHS of the kernel  $K_u$  defined in (9) relates to that of  $K$ . We use the following general proposition about transformations of RKHSs. See Theorems 16.8 and 16.9 in [48], Section 5.4 in [34], and Section 7 in [45] for similar results. A proof is included here for completeness and because our formulation differs slightly from those that we have found in the literature.

**Proposition 1.** *Let  $K$  be a positive-semidefinite kernel on  $\Omega$  and  $\mathcal{A}$  an invertible linear operator on  $\mathcal{H}(K, \Omega)$  such that the functional  $\delta_{\mathbf{x}} \circ \mathcal{A}$  is bounded on  $\mathcal{H}(K, \Omega)$  for every  $\mathbf{x} \in \Omega$ . Then*

$$R(\mathbf{x}, \mathbf{y}) = \mathcal{A}_{\mathbf{x}} \mathcal{A}_{\mathbf{y}} K(\mathbf{x}, \mathbf{y}) = (\delta_{\mathbf{x}} \circ \mathcal{A})_{\mathbf{x}'} (\delta_{\mathbf{y}} \circ \mathcal{A})_{\mathbf{y}'} K(\mathbf{x}', \mathbf{y}')$$

*defines a positive-semidefinite kernel on  $\Omega$ . Furthermore,*

$$\mathcal{H}(R, \Omega) = \mathcal{A}(\mathcal{H}(K, \Omega)) = \{\mathcal{A}f : f \in \mathcal{H}(K, \Omega)\} \quad \text{and} \quad \|\mathcal{A}f\|_R = \|f\|_K \quad \text{for every } f \in \mathcal{H}(K, \Omega).$$

**Proof.** Because the functional  $\ell_{\mathbf{y}} = \delta_{\mathbf{y}} \circ \mathcal{A}$  is bounded on  $\mathcal{H}(K, \Omega)$ , the Riesz representation theorem implies that there exists a unique representer  $l_{\mathbf{y}} \in \mathcal{H}(K, \Omega)$  such that  $\ell_{\mathbf{y}}(f) = \langle f, l_{\mathbf{y}} \rangle_K$  for every  $f \in \mathcal{H}(K, \Omega)$ . Therefore, by the reproducing property,

$$l_{\mathbf{y}}(\mathbf{x}) = \langle K(\mathbf{x}, \cdot), l_{\mathbf{y}} \rangle_K = \ell_{\mathbf{y}}(K(\mathbf{x}, \cdot)) = (\delta_{\mathbf{y}} \circ \mathcal{A})_{\mathbf{u}} K(\mathbf{x}, \mathbf{u})$$

for any  $\mathbf{x}, \mathbf{y} \in \Omega$ . Since  $l_{\mathbf{y}}$  is an element of  $\mathcal{H}(K, \Omega)$ ,  $\ell_{\mathbf{x}}(l_{\mathbf{y}}) = \langle l_{\mathbf{x}}, l_{\mathbf{y}} \rangle_K$  and

$$\ell_{\mathbf{x}}(l_{\mathbf{y}}) = (\delta_{\mathbf{x}} \circ \mathcal{A})(l_{\mathbf{y}}) = (\delta_{\mathbf{x}} \circ \mathcal{A})_{\mathbf{v}} (\delta_{\mathbf{y}} \circ \mathcal{A})_{\mathbf{u}} K(\mathbf{v}, \mathbf{u}) = R(\mathbf{x}, \mathbf{y}),$$

from which it follows that  $R$  is a well-defined kernel. To verify that  $R$  is positive-semidefinite, compute

$$\sum_{i=1}^N \sum_{j=1}^N a_i a_j R(\mathbf{x}_i, \mathbf{x}_j) = \sum_{i=1}^N \sum_{j=1}^N a_i a_j \langle l_{\mathbf{x}_i}, l_{\mathbf{x}_j} \rangle_K = \left\langle \sum_{i=1}^N a_i l_{\mathbf{x}_i}, \sum_{i=1}^N a_i l_{\mathbf{x}_i} \right\rangle_K = \left\| \sum_{i=1}^N a_i l_{\mathbf{x}_i} \right\|_K^2 \geq 0 \quad (32)$$

for any  $N \in \mathbb{N}$ ,  $a_i \in \mathbb{R}$ , and  $\mathbf{x}_i \in \Omega$ . To prove the claims related to  $\mathcal{H}(R, \Omega)$  we use a classical characterisation [e.g., 34, Section 3.4] which states that  $f \in \mathcal{H}(K, \Omega)$  if and only if there is  $c > 0$  such that

$$K_c(\mathbf{x}, \mathbf{y}) = c^2 K(\mathbf{x}, \mathbf{y}) - f(\mathbf{x})f(\mathbf{y}) \quad (33)$$

defines a positive-semidefinite kernel. The smallest constant for which  $K_c$  is positive-semidefinite equals the RKHS norm of  $f$ . Now, assuming that  $f \in \mathcal{H}(K, \Omega)$  and applying  $\mathcal{A}$  twice on (33) yields the kernel

$$R_c(\mathbf{x}, \mathbf{y}) = c^2 R(\mathbf{x}, \mathbf{y}) - \mathcal{A}f(\mathbf{x})\mathcal{A}f(\mathbf{y}),$$

which, by the argument used in (32), is positive-semidefinite if  $K_c$  is. Hence  $\mathcal{A}(\mathcal{H}(K, \Omega)) \subset \mathcal{H}(R, \Omega)$  and  $\|\mathcal{A}f\|_R \leq \|f\|_K$ . That these are equalities follows from the invertibility of  $\mathcal{A}$ :  $\mathcal{H}(K, \Omega) \rightarrow \mathcal{A}(\mathcal{H}(K, \Omega))$ .  $\square$

Applying Proposition 1 to  $\mathcal{A} = \mathcal{L}^{-1}$  yields the following theorem.

**Theorem 8.** *Let  $k > d/2$  and consider the kernel  $K_u$  in (9). If  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$  and Assumption 1 holds, then*



(i) The kernel  $K_u$  is positive-semidefinite on  $\Omega$  and its RKHS is

$$\mathcal{H}(K_u, \Omega) = \{u : u \text{ is a solution of (4) for some } f \in \mathcal{H}(K, \Omega)\}. \quad (34)$$

Furthermore,  $\|u\|_{K_u} = \|f\|_K$ .

(ii) It holds that  $\mathcal{H}(K_u, \Omega) \subset H^{k+2}(\Omega)$  and there are constants  $C_u = C(K, k, \Omega, \mathcal{L})$  and  $C'_u = C(K, k, \Omega, \mathcal{L})$  such that

$$C_u \|u\|_{H^{k+2}(\Omega)} \leq \|u\|_{K_u} \leq C'_u \|u\|_{H^{k+2}(\Omega)} \quad (35)$$

for all  $u \in \mathcal{H}(K_u, \Omega)$ .

**Proof.** Because  $\mathcal{H}(K, \Omega) \subset \mathcal{H}_{\mathcal{L}}(\Omega)$ , the linear operator  $\mathcal{L}$  is invertible on  $\mathcal{H}(K, \Omega)$ . Furthermore, by Lemma 1 the functionals  $\delta_x \circ \mathcal{L}^{-1}$  are bounded on  $\mathcal{H}(K, \Omega)$ . Therefore the first claim follows by applying Proposition 1 to  $\mathcal{A} = \mathcal{L}^{-1}$ . To verify the second claim, observe that it now follows from the norm-equivalence  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$  and Theorem 6 that, for a constant  $C = C(k, \Omega, \mathcal{L})$ ,

$$\|u\|_{K_u} = \|f\|_K \geq C_K \|f\|_{H^k(\Omega)} \geq C_K C^{-1} \|u\|_{H^{k+2}(\Omega)}$$

and

$$\|u\|_{K_u} = \|f\|_K \leq C'_K \|f\|_{H^k(\Omega)} = C'_K \|\mathcal{L}u\|_{H^k(\Omega)} \leq C'_K C_{\mathcal{L}} \|u\|_{H^{k+2}(\Omega)},$$

where  $C_{\mathcal{L}} = C_{\mathcal{L}}(k, \Omega, \mathcal{L})$  and the last inequality follows from the fact that the differential operator  $\mathcal{L}$  is second-order and its coefficient functions are in  $C^{k+1}(\bar{\Omega})$  by Assumption 1.  $\square$

Theorem 8 is essential in what follows. We believe that there is room for significant generalisations by studying operators  $\mathcal{A}$  for which the transformed RKHS can be connected to a Sobolev (or other classical) function space. Finally, we need the following result [e.g., 3, p. 24] on the RKHS of a sum kernel to analyse statFEM when a discrepancy term is included (recall Section 2.4).

**Theorem 9.** Let  $K_1$  and  $K_2$  be two positive-semidefinite kernels on  $\Omega$ . Then  $R = K_1 + K_2$  is a positive-semidefinite kernel on  $\Omega$  and its RKHS consists of functions which can be written as  $f = f_1 + f_2$  for  $f_1 \in \mathcal{H}(K_1, \Omega)$  and  $f_2 \in \mathcal{H}(K_2, \Omega)$ . The RKHS norm is

$$\|f\|_R^2 = \min\{\|f_1\|_{K_1}^2 + \|f_2\|_{K_2}^2 : f = f_1 + f_2 \text{ s.t. } f_1 \in \mathcal{H}(K_1, \Omega), f_2 \in \mathcal{H}(K_2, \Omega)\}.$$

### 6.3. Sampling inequalities

Denote  $(x)_+ = \max\{x, 0\}$  for  $x \in \mathbb{R}$ . The following sampling inequality is the main building block of our error estimates.

**Theorem 10** (Theorem 4.1 in [2]). Let  $p \in [1, \infty]$ ,  $k > d/2$ , and  $\gamma = \max\{p, 2\}$ . If  $g \in H^k(\Omega)$ , then there is a constant  $C = C(k, p, \Omega)$  such that

$$\|g\|_{L^p(\Omega)} \leq C \left( h_{X, \Omega}^{k-d(1/2-1/p)_+} \|g\|_{H^k(\Omega)} + h_{X, \Omega}^{d/\gamma} \|\mathbf{g}(X)\|_2 \right). \quad (36)$$

Here  $\mathbf{g}(X) = (g(\mathbf{x}_1), \dots, g(\mathbf{x}_n)) \in \mathbb{R}^n$ .

See [24, 41, 46, 47, 49] for a variety of applications of this and related sampling inequalities to error analysis of GP regression. Note that it is often stated that (36) and similar bounds hold when  $h_{X, \Omega}$  is sufficiently small, a requirement that can be eliminated by enlarging the multiplicative constant  $C$  on the right-hand side.

#### 6.4. The concentration function and small ball probabilities

Final ingredients that we need are certain results on the concentration function and small ball probabilities of Gaussian measures. Given  $u_0 = \mathcal{L}^{-1} f_0$  for some  $f_0 \in \mathcal{H}_{\mathcal{L}}(\Omega)$ , define the concentration function

$$\phi_{u_0}(\varepsilon) = \gamma_{u_0}(\varepsilon) + \beta(\varepsilon),$$

where

$$\gamma_{u_0}(\varepsilon) = \inf_{u \in \mathcal{H}(K_u, \Omega)} \left\{ \|u\|_{K_u}^2 : \|u - u_0\|_{L^\infty(\Omega)} < \varepsilon \right\}, \quad \beta(\varepsilon) = -\ln \Pi_u(\{u : \|u\|_{L^\infty(\Omega)} < \varepsilon\}).$$

Here  $\Pi_u$  denotes the Gaussian probability measure associated to the zero-mean Gaussian process with covariance kernel  $K_u$ .

**Proposition 2.** *Let  $k \geq r > d/2$ . Suppose that  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$ , that Assumption 1 holds and that  $c \leq 0$ . If there exists  $f_0 \in H^r(\mathbb{R}^d) \cap C^r(\mathbb{R}^d)$  such that  $u_0 = \mathcal{L}^{-1} f_0|_\Omega$ , then there is a constant  $C = C(f_0, K, k, r, \Omega, \mathcal{L})$  such that*

$$\gamma_{u_0}(\varepsilon) \leq C \varepsilon^{-2(k-r)/r}$$

when  $\varepsilon$  is sufficiently small.

**Proof.** Theorem 8 implies that

$$\gamma_{u_0}(\varepsilon) = \inf_{f \in \mathcal{H}(K, \Omega)} \left\{ \|f\|_K : \|u - u_0\|_{L^\infty(\Omega)} < \varepsilon \text{ for } u = \mathcal{L}^{-1} f \right\}.$$

Since  $k > d/2$ , so that  $\mathcal{H}(K, \Omega)$  is embedded in a Hölder space, and  $r \geq 1$ , for any  $f \in \mathcal{H}(K, \Omega)$  the function  $f - f_0|_\Omega$  has a unique continuous extension to  $\bar{\Omega}$  that satisfies the assumptions of Theorem 7. Thus there is  $C_1 = C(\Omega, \mathcal{L})$  such that  $\|u - u_0\|_{L^\infty(\Omega)} \leq C_1 \|f - f_0|_\Omega\|_{L^\infty(\Omega)}$ . Therefore  $\|u - u_0\|_{L^\infty(\Omega)} < \varepsilon$  if  $\|f - f_0|_\Omega\|_{L^\infty(\Omega)} < \varepsilon C_1^{-1}$ , which implies that

$$\gamma_{u_0}(\varepsilon) \leq \inf_{f \in \mathcal{H}(K, \Omega)} \left\{ \|f\|_K : \|f - f_0|_\Omega\|_{L^\infty(\Omega)} < \varepsilon / C_1 \right\}.$$

Lemma 4 in [44] and Lemma 23 in [49] bound the right-hand side as

$$\inf_{f \in \mathcal{H}(K, \Omega)} \left\{ \|f\|_K : \|f - f_0|_\Omega\|_{L^\infty(\Omega)} < \varepsilon / C_1 \right\} \leq C_2 C_1^{2(k-r)/r} \varepsilon^{-2(k-r)/r}$$

for  $C_2 = C(f_0, K, k, r, \Omega)$  when  $\varepsilon$  is sufficiently small. This completes the proof.  $\square$

Let  $\varepsilon > 0$  and let  $(\mathcal{H}, d_{\mathcal{H}})$  be a metric space. The metric entropy of a compact subset  $A$  of  $(\mathcal{H}, d_{\mathcal{H}})$  is defined as  $H_{\text{ent}}(A, d_{\mathcal{H}}, \varepsilon) = \ln N(A, d_{\mathcal{H}}, \varepsilon)$ . Here  $N(A, d_{\mathcal{H}}, \varepsilon)$  denotes the minimum covering number:

$$N(A, d_{\mathcal{H}}, \varepsilon) = \min \left\{ n \geq 1 : \text{there exist } x_1, \dots, x_n \in A \text{ such that } A \subset \bigcup_{i=1}^n B_\varepsilon(x_i; \mathcal{H}, d_{\mathcal{H}}) \right\},$$

where  $B_\varepsilon(x; \mathcal{H}, d_{\mathcal{H}})$  is the  $x$ -centered  $\varepsilon$ -ball in  $(\mathcal{H}, d_{\mathcal{H}})$ . If  $(\mathcal{H}, d_{\mathcal{H}})$  is a normed space, we have the scaling identity

$$H_{\text{ent}}(\lambda A, d_{\mathcal{H}}, \varepsilon) = H_{\text{ent}}(A, d_{\mathcal{H}}, \varepsilon |\lambda|^{-1}) \quad (37)$$

for any  $\lambda \neq 0$ ; see, for example, Equation (4.171) in [18].

**Lemma 2.** *Let  $k > d/2$ . Suppose that  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$ , that Assumption 1 holds, and that  $c \leq 0$ . Let  $B_1^u$  and  $B_1^k$  denote the unit balls of  $(\mathcal{H}(K_u, \Omega), \|\cdot\|_{K_u})$  and  $(H^k(\Omega), \|\cdot\|_{H^k(\Omega)})$ , respectively. Then there is a constant  $C = C(K, \Omega, \mathcal{L})$  such that*

$$H_{\text{ent}}(B_1^u, \|\cdot\|_{L^\infty(\Omega)}, \varepsilon) \leq H_{\text{ent}}(B_1^k, \|\cdot\|_{L^\infty(\Omega)}, C\varepsilon). \quad (38)$$

**Proof.** Fix  $\varepsilon > 0$  and denote  $n_\varepsilon = N(B_1^k, \|\cdot\|_{L^\infty(\Omega)}, \varepsilon)$  and let  $f_1, \dots, f_{n_\varepsilon} \in B_1^k$  be such that

$$B_1^k = B_1(0; H^k(\Omega), \|\cdot\|_{H^k(\Omega)}) \subset \bigcup_{i=1}^{n_\varepsilon} B_\varepsilon(f_i; H^k(\Omega), \|\cdot\|_{L^\infty(\Omega)}).$$

Then

$$\mathcal{L}^{-1} B_1(0; H^k(\Omega), \|\cdot\|_{H^k(\Omega)}) \subset \bigcup_{i=1}^{n_\varepsilon} \mathcal{L}^{-1} B_\varepsilon(f_i; H^k(\Omega), \|\cdot\|_{L^\infty(\Omega)}). \quad (39)$$

We have  $\|u\|_{K_u} = \|f\|_K \geq C_K \|f\|_{H^k(\Omega)}$  by Theorem 8. The norm-equivalence  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$  and the estimate  $\|u\|_{L^\infty(\Omega)} \leq C_\infty \|f\|_{L^\infty(\Omega)}$  for  $C_\infty = C(\Omega, \mathcal{L})$  follow from the Sobolev embedding theorem and Theorem 7. Therefore

$$B_1^u = B_1(0; \mathcal{H}(K_u, \Omega), \|\cdot\|_{K_u}) = \mathcal{L}^{-1} B_1(0; \mathcal{H}(K, \Omega), \|\cdot\|_K) \subset \mathcal{L}^{-1} B_{1/C_K}(0; H^k(\Omega), \|\cdot\|_{H^k(\Omega)})$$

and

$$\mathcal{L}^{-1} B_\varepsilon(f_i; H^k(\Omega), \|\cdot\|_{L^\infty(\Omega)}) \subset B_{C_\infty \varepsilon}(u_i; \mathcal{H}(K_u, \Omega), \|\cdot\|_{L^\infty(\Omega)}),$$

where  $u_i = \mathcal{L}^{-1} f_i$ . Applying these two inclusion relations to (39) and using the definition of metric entropy, together with (37), yields the claim.  $\square$

**Proposition 3.** Let  $k > d/2$ . Suppose that  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$ , that Assumption 1 holds, and that  $c \leq 0$ . Let  $B_1^u$  denote the unit ball of  $(\mathcal{H}(K_u, \Omega), \|\cdot\|_{K_u})$ . Then there is a positive constant  $C$ , which does not depend on  $\varepsilon$ , such that

$$\beta(\varepsilon) \leq C \varepsilon^{-2d/(2k-d)}$$

for sufficiently small  $\varepsilon$ .

**Proof.** It is a standard result that the metric entropy of the unit ball of  $H^k(\Omega)$  in Lemma 2 satisfies

$$H_{\text{ent}}(B_1^k, \|\cdot\|_{L^\infty(\Omega)}, \varepsilon) \leq C_{\text{ent}} \varepsilon^{-d/k}$$

for a positive constant  $C_{\text{ent}} = C(k)$  and any  $\varepsilon < 1$ . See, for instance, Theorem 4.3.36 in [18], Theorem 3.3.2 in [9], the proof of Lemma 3 in [44], and Appendix F in [49]. It follows from Equation (38) that

$$H_{\text{ent}}(B_1^u, \|\cdot\|_{L^\infty(\Omega)}, \varepsilon) \leq C_{\text{ent}} C^{-d/k} \varepsilon^{-d/k} \quad (40)$$

for sufficiently small  $\varepsilon$ . According to Theorem 1.2 in [29], the estimate (40) implies that

$$\beta(\varepsilon) \leq C' \varepsilon^{-2d/(2k-d)}$$

for a positive constant  $C'$  which does not depend on  $\varepsilon$ .  $\square$

A combination of Propositions 2 and 3 yields an estimate for  $\phi_{u_0}(\varepsilon)$ . Define the function

$$\psi_{u_0}(\varepsilon) = \frac{\phi_{u_0}(\varepsilon)}{\varepsilon^2}$$

and let  $\psi_{u_0}^{-1}(n) = \sup\{\varepsilon > 0 : \psi_{u_0}(\varepsilon) \geq n\}$  denote its generalised inverse.

**Theorem 11.** Let  $k \geq r > d/2$ . Suppose that  $\mathcal{H}(K, \Omega) \simeq H^k(\Omega)$ , that Assumption 1 holds, and that  $c \leq 0$ . If there exists  $f_0 \in H^r(\mathbb{R}^d) \cap C^r(\mathbb{R}^d)$  such that  $u_0 = \mathcal{L}^{-1} f_0|_\Omega$ , then there is a positive constant  $C$ , which does not depend on  $\varepsilon$ , such that

$$\psi_{u_0}^{-1}(n) \leq C n^{-\min\{r, k-d/2\}/(2k)}$$

for all sufficiently large  $n \geq 1$ .

**Proof.** By Propositions 2 and 3,

$$\psi_{u_0}(\varepsilon) \leq C_0 \left( \varepsilon^{-2(k-r)/r-2} + \varepsilon^{-2d/(2k-d)-2} \right) \leq 2C_0 \varepsilon^{-2k/\min\{r, k-d/2\}}$$

whenever  $\varepsilon$  is sufficiently small, where the positive constant  $C_0$  does not depend on  $\varepsilon$ . It follows from the definition of  $\psi_{u_0}^{-1}$  that

$$\psi_{u_0}^{-1}(n) \leq C n^{-\min\{r, k-d/2\}/(2k)}$$

for  $C = (2C_0)^{\min\{r, k-d/2\}/(2k)}$ .  $\square$

### 6.5. Proofs of main results

We are now ready to prove the theorems in Section 3. Given an  $n$ -vector  $\mathbf{Z}$ , we employ the interpolation operator notation

$$I_X(\mathbf{Z})(\mathbf{x}) = \mathbf{K}_u(\mathbf{x}, X)^\top (\mathbf{K}_u(X, X) + \sigma_\varepsilon^2 \mathbf{I}_n)^{-1} \mathbf{Z}. \quad (41)$$

That is, the ideal conditional mean in (10a) can be written as

$$m_{u|Y} = m_u - I_X(\mathbf{m}_u(X)) + I_X(\mathbf{Y}). \quad (42)$$

**Proof of Theorem 1.** By Theorems 6 and 8,  $u_k, m_u \in \mathcal{H}(K_u, \Omega) \subset H^{k+2}(\Omega)$ . Therefore it follows from Theorem 10 with  $p = 2$  and  $g = u_t - m_{u|Y}$  that there is a constant  $C_1$  independent of  $X$  such that

$$\|u_t - m_{u|Y}\|_{L^2(\Omega)} \leq C_1 \left( h_{X,\Omega}^{k+2} \|u_t - m_{u|Y}\|_{H^{k+2}(\Omega)} + h_{X,\Omega}^{d/2} \|\mathbf{u}_t(X) - \mathbf{m}_{u|Y}(X)\|_2 \right). \quad (43)$$

The decomposition in (42) gives

$$\|u_t - m_{u|Y}\|_{H^{k+2}(\Omega)} \leq \|u_t - I_X(\mathbf{Y})\|_{H^{k+2}(\Omega)} + \|m_u - I_X(\mathbf{m}_u(X))\|_{H^{k+2}(\Omega)}. \quad (44)$$

The triangle inequality and Lemma 17 in [49], in combination with (35), yield

$$\begin{aligned} \|u_t - I_X(\mathbf{Y})\|_{H^{k+2}(\Omega)} &\leq \|u_t\|_{H^{k+2}(\Omega)} + \|I_X(\mathbf{Y})\|_{H^{k+2}(\Omega)} \leq \|u_t\|_{H^{k+2}(\Omega)} + C_u^{-1} \|I_X(\mathbf{Y})\|_{K_u} \\ &\leq (1 + C_u^{-1} C'_u) \|u_t\|_{H^{k+2}(\Omega)} + C_u^{-1} \sigma_\varepsilon^{-1} \|\boldsymbol{\varepsilon}\|_2 \\ &\leq 2C_u^{-1} C'_u \|u_t\|_{H^{k+2}(\Omega)} + C_u^{-1} \sigma_\varepsilon^{-1} \|\boldsymbol{\varepsilon}\|_2 \end{aligned} \quad (45)$$

where  $\boldsymbol{\varepsilon} = (\varepsilon_1, \dots, \varepsilon_n) \in \mathbb{R}^n$  is the noise vector and we used the fact that  $C_u \leq C'_u$ . The second term in (44) has the bound

$$\|m_u - I_X(\mathbf{m}_u(X))\|_{H^{k+2}(\Omega)} \leq 2C_u^{-1} C'_u \|m_u\|_{H^{k+2}(\Omega)}, \quad (46)$$

which is obtained in the same way as (45) but with  $\boldsymbol{\varepsilon}$  set as the zero vector. From Theorem 22 in [49] and Theorem 11 with  $r = k$  and  $f_0 = f_t$  (i.e.,  $u_0 = u_t$  and  $\min\{r, k - d/2\} = k - d/2$ ) we get

$$\mathbb{E}[\|\mathbf{u}_t(X) - \mathbf{m}_{u|Y}(X)\|_2] \leq C_2 \sqrt{n} \psi_{u_t}^{-1}(n) \leq C_2 C_3 n^{d/(4k)}, \quad (47)$$

for positive constants  $C_2$  and  $C_3$  which do not depend on  $X$ . Here we have enlarged the constant  $C_3$  to remove the requirement that  $n$  be sufficiently large. Inserting the estimates (45)–(47) into (43) and using the bound  $\sigma_\varepsilon^{-1} \mathbb{E}[\|\boldsymbol{\varepsilon}\|_2] \leq \sqrt{n}$ , which follows from the Gaussianity of the noise terms, yields

$$\begin{aligned} \mathbb{E}[\|u_t - m_{u|Y}\|_{L^2(\Omega)}] &\leq 2C_1 C_u^{-1} C'_u h_{X,\Omega}^{k+2} (\|u_t\|_{H^{k+2}(\Omega)} + \|m_u\|_{H^{k+2}(\Omega)}) + C_1 C_u^{-1} h_{X,\Omega}^{k+2} \sigma_\varepsilon^{-1} \mathbb{E}[\|\boldsymbol{\varepsilon}\|_2] + C_1 C_2 C_3 h_{X,\Omega}^{d/2} n^{d/(4k)} \\ &\leq 2C_1 C_u^{-1} C'_u h_{X,\Omega}^{k+2} (\|u_t\|_{H^{k+2}(\Omega)} + \|m_u\|_{H^{k+2}(\Omega)}) + C_1 C_u^{-1} h_{X,\Omega}^{k+2} \sqrt{n} + C_1 C_2 C_3 h_{X,\Omega}^{d/2} n^{d/(4k)}. \end{aligned} \quad (48)$$

This concludes the proof of (22).  $\square$

**Proof of Theorem 2.** The proof proceeds exactly as that of Theorem 1, except that now  $\boldsymbol{\varepsilon} = \mathbf{0}$  and  $\mathbf{m}_{u|Y}(X) = \mathbf{u}_t(X)$  because  $\mathbf{Y} = \mathbf{u}_t(X)$ . This means that the terms in (48) that arise from  $\|\boldsymbol{\varepsilon}\|_2$  and (47) are now zero, so that we are left with the claimed bound.  $\square$

The following proposition allows us to make use of Assumption 2 on the error of the finite element discretisation. Although this basic proposition must have appeared several times and in various forms in the literature on scalable approximations for GP regression, we have not been able to locate a convenient reference for it.

**Proposition 4.** Let  $R_1$  and  $R_2$  be any positive-semidefinite kernels,  $\sigma > 0$ , and  $\mathbf{Z} \in \mathbb{R}^n$ . If

$$\sup_{\mathbf{x}, \mathbf{y} \in \Omega} |R_1(\mathbf{x}, \mathbf{y}) - R_2(\mathbf{x}, \mathbf{y})| = \delta$$

for some  $\delta > 0$ , then the functions

$$m_{|\mathbf{Z}}^1(\mathbf{x}) = \mathbf{R}_1(\mathbf{x}, X) \left( \mathbf{R}_1(X, X) + \sigma^2 \mathbf{I}_n \right)^{-1} \mathbf{Z}, \quad m_{|\mathbf{Z}}^2(\mathbf{x}) = \mathbf{R}_2(\mathbf{x}, X) \left( \mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n \right)^{-1} \mathbf{Z}$$

satisfy

$$\sup_{\mathbf{x} \in \Omega} |m_{|\mathbf{Z}}^1(\mathbf{x}) - m_{|\mathbf{Z}}^2(\mathbf{x})| \leq \|\mathbf{Z}\|_2 (\delta + C \sigma^{-2}) \sigma^{-2} \delta n, \quad (49)$$

where  $C = \sup_{\mathbf{x}, \mathbf{y} \in \Omega} |R_2(\mathbf{x}, \mathbf{y})|$ .

**Proof.** Write

$$\begin{aligned} |m_{|\mathbf{Z}}^1(\mathbf{x}) - m_{|\mathbf{Z}}^2(\mathbf{x})| &= \left| \mathbf{Z}^\top [(\mathbf{R}_1(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}_1(\mathbf{x}, X) - (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}_2(\mathbf{x}, X)] \right| \\ &\leq \|\mathbf{Z}\|_2 \left\| (\mathbf{R}_1(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}_1(\mathbf{x}, X) - (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}_2(\mathbf{x}, X) \right\|_2. \end{aligned}$$

Let  $R(\mathbf{x}, \mathbf{y}) = R_1(\mathbf{x}, \mathbf{y}) - R_2(\mathbf{x}, \mathbf{y})$  so that  $\sup_{\mathbf{x}, \mathbf{y} \in \Omega} |R(\mathbf{x}, \mathbf{y})| = \delta$  and

$$\begin{aligned} &\left\| (\mathbf{R}_1(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}_1(\mathbf{x}, X) - (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}_2(\mathbf{x}, X) \right\|_2 \\ &= \left\| (\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} (\mathbf{R}_2(\mathbf{x}, X) + \mathbf{R}(\mathbf{x}, X)) - (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}_2(\mathbf{x}, X) \right\|_2 \\ &\leq \left\| [(\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} - (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1}] \mathbf{R}_2(\mathbf{x}, X) \right\|_2 + \left\| (\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}(\mathbf{x}, X) \right\|_2 \\ &\leq \sqrt{n} C \left\| [(\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} - (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1}] \right\|_2 \\ &\quad + \sqrt{n} \delta \left\| (\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \right\|_2. \end{aligned}$$

Because the matrix  $\mathbf{R}_1(X, X) = \mathbf{R}_2(X, X) + \mathbf{R}(X, X)$  is positive-semidefinite, the largest singular value of the matrix  $(\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1}$  is  $(\sigma^2 + \lambda_{\min}(\mathbf{R}_1(X, X)))^{-1}$ . Therefore

$$\sqrt{n} \delta \left\| (\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \right\|_2 = \sqrt{n} \delta [\sigma^2 + \lambda_{\min}(\mathbf{R}_1(X, X))]^{-1} \leq \sqrt{n} \sigma^{-2} \delta.$$

Finally,

$$\begin{aligned} &\left\| (\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} - (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \right\|_2 \\ &= \left\| (\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \mathbf{R}(X, X) (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \right\|_2 \\ &\leq \|\mathbf{R}(X, X)\|_2 \left\| (\mathbf{R}_2(X, X) + \mathbf{R}(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \right\|_2 \left\| (\mathbf{R}_2(X, X) + \sigma^2 \mathbf{I}_n)^{-1} \right\|_2 \\ &\leq \sqrt{n} \sigma^{-4} \delta. \end{aligned}$$

The claim follows by putting these estimates together.  $\square$

The proof of Theorem 3 is a straightforward combination of Theorem 1 and Proposition 4.

**Proof of Theorem 3.** The triangle inequality yields

$$\|u_t - m_{u|\mathbf{Y}}^{\text{FE}}\|_{L^2(\Omega)} \leq \|u_t - m_{u|\mathbf{Y}}\|_{L^2(\Omega)} + \|m_{u|\mathbf{Y}} - m_{u|\mathbf{Y}}^{\text{FE}}\|_{L^2(\Omega)}. \quad (50)$$

Theorem 1 bounds the expectation of the first term as

$$\mathbb{E} \left[ \|u_t - m_{u|\mathbf{Y}}\|_{L^2(\Omega)} \right] \leq C_1 n^{-1/2+d/(4k)} \quad (51)$$

for a constant  $C_1 > 0$  that is independent of  $X$ , while Proposition 4 and Assumption 2 give

$$\|m_{u|\mathbf{Y}} - m_{u|\mathbf{Y}}^{\text{FE}}\|_{L^2(\Omega)} \leq C_2 \|\mathbf{Y}\|_2 (n_{\text{FE}}^{-q} + \sigma_\varepsilon^{-2}) \sigma_\varepsilon^{-2} n_{\text{FE}}^{-q} n$$

for a constant  $C_2 > 0$  that is independent of  $X$ . Since

$$\mathbb{E}[\|\mathbf{Y}\|_2] \leq \|\mathbf{u}_t(X)\|_2 + \mathbb{E}[\|\boldsymbol{\varepsilon}\|_2] \leq (\|u_t\|_{L^\infty(\Omega)} + \sigma_\varepsilon) \sqrt{n},$$

from Theorem 7 we obtain

$$\mathbb{E} \left[ \|m_{u|Y} - m_{u|Y}^{\text{FE}}\|_{L^2(\Omega)} \right] \leq C_3(n_{\text{FE}}^{-q} + \sigma_\varepsilon^{-2})\sigma_\varepsilon^{-2}(\|f_t\|_{L^\infty(\Omega)} + \sigma_\varepsilon)n_{\text{FE}}^{-q}n^{3/2} \quad (52)$$

for a constant  $C_3 > 0$  that is independent of  $X$ . Taking expectation of (50) and using the bounds (51) and (52) concludes the proof.  $\square$

**Proof of Theorem 4.** By Theorem 8, the norm-equivalence assumption, and the inequality  $k_1 + 2 \geq r$ , it holds that  $\mathcal{H}(K_u, \Omega) \subset H(K_d, \Omega)$ . From this inclusion, Theorem 9 and (35) it follows that  $\mathcal{H}(K_{ud}, \Omega) \simeq H^r(\Omega)$ . By Theorem 8 and our assumptions, the functions  $m_d$ ,  $m_u$ , and  $u_t$  are in  $H^{k_2+2}(\Omega)$  and  $r \geq k_2 + 2$ . We can therefore apply Theorem 2 in [49] with

$$k = K_{ud}, \quad f = u_t, \quad \tau_k^- = \tau_k^+ = r, \quad \tau_f = k_2 + 2, \quad s = 0, \quad q = 2.$$

This yields the estimate

$$\mathbb{E} \left[ \|u_t - m_{d,u|Y}\|_{L^2(\Omega)} \right] \leq Ch_{X,\Omega}^{d/2} \left( h_{X,\Omega}^{k_2+2-d/2} \rho_{X,\Omega}^{r-k_2-2} + \sqrt{n} h_{X,\Omega}^{r-d/2} + n^\kappa \right) = C \left( h_{X,\Omega}^{k_2+2} \rho_{X,\Omega}^{r-k_2-2} + \sqrt{n} h_{X,\Omega}^r + n^{\kappa(k_2,r)} h_{X,\Omega}^{d/2} \right),$$

where

$$\kappa(k_2, r) = \max \left\{ \frac{1}{2} - \frac{k_2 + 2}{2r}, \frac{d}{4(k_2 + 2)} \right\} \leq \frac{1}{2}, \quad (53)$$

for a positive constant  $C$  that is independent of  $X$ . Theorem 2 in [49] requires that  $h_{X,\Omega}$  be sufficiently small. We eliminate this assumption by enlarging  $C$ . This proves (26) while (27) follows from  $h_{X,\Omega} = O(n^{-1/d})$  and the fact that the mesh ratio  $\rho_{X,\Omega}$  is bounded for quasi-uniform points.  $\square$

**Proof of Theorem 5.** The proof is identical to that of Theorem 3 except that the bound (27) is used in place of (23).  $\square$

## Acknowledgements

TK was supported by the Research Council of Finland grants 338567 (“Scalable, adaptive and reliable probabilistic integration”) and 359183 (“Flagship of Advanced Mathematics for Sensing, Imaging and Modelling”). We thank the associate editor and reviewers for comments and suggestions that improved the article.

## References

- [1] A. Abdulle, G. Garegnani, A probabilistic finite element method based on random meshes: Error estimators and Bayesian inverse problems, *Computer Methods in Applied Mechanics and Engineering* 384 (2021) 113961.
- [2] R. Arcangéli, M. C. L. de Silanes, J. J. Torrens, An extension of a bound for functions in Sobolev spaces, with applications to  $(m, s)$ -spline interpolation and smoothing, *Numerische Mathematik* 107 (2007) 181–211.
- [3] A. Berlinet, C. Thomas-Agnan, *Reproducing Kernel Hilbert Spaces in Probability and Statistics*, Springer, Berlin, 2004.
- [4] S. C. Brenner, L. R. Scott, *The Mathematical Theory of Finite Element Methods*, number 15 in *Texts in Applied Mathematics*, Springer, Berlin, 2008.
- [5] F.-X. Briol, C. J. Oates, M. Girolami, M. A. Osborne, D. Sejdinovic, Probabilistic integration: A role in statistical computation? (with discussion and rejoinder), *Statistical Science* 34 (2019) 1–22.
- [6] I. Cialenco, G. E. Fasshauer, Q. Ye, Approximation of stochastic partial differential equations by a kernel-based collocation method, *International Journal of Computer Mathematics* 89 (2012) 2543–2561.
- [7] J. Cockayne, C. J. Oates, T. Sullivan, M. Girolami, Probabilistic numerical methods for partial differential equations and Bayesian inverse problems, in: *Proceedings of the 36th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering*, volume 1853 of *AIP Conference Proceedings*.
- [8] C. Duffin, E. Cripps, T. Stemler, M. Girolami, Statistical finite elements for misspecified models, *Proceedings of the National Academy of Sciences* 118 (2021).
- [9] D. Edmunds, H. Triebel, *Function Spaces, Entropy Numbers, Differential Operators*, Cambridge University Press, 1996.
- [10] L. C. Evans, *Partial Differential Equations*, number 19 in *Graduate Studies in Mathematics*, American Mathematical Society, New York, 1998.
- [11] G. E. Fasshauer, Solving partial differential equations by collocation with radial basis functions, in: *Proceedings of Chamonix*, Vanderbilt University Press, 1996, pp. 1–8.
- [12] G. E. Fasshauer, Q. Ye, Reproducing kernels of generalized Sobolev spaces via a Green function approach with distributional operators, *Numerische Mathematik* 119 (2011) 585–611.
- [13] G. E. Fasshauer, Q. Ye, Reproducing kernels of Sobolev spaces via a Green kernel approach with differential operators and boundary operators, *Advances in Computational Mathematics* 38 (2013) 891–921.

- [14] E. Febrianto, L. Butler, M. Girolami, F. Cirak, Digital twinning of self-sensing structures using the statistical finite element method, *Data-Centric Engineering* 3 (2022) e31.
- [15] C. Franke, R. Schaback, Solving partial differential equations by collocation using radial basis functions, *Applied Mathematics and Computation* 93 (1998) 73–82.
- [16] S. Ghosal, J. L. Ghosh, A. van der Vaart, Convergence rates of posterior distributions, *The Annals of Statistics* 28 (2000) 500–531.
- [17] D. Gilbarg, N. S. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer, Berlin, 2nd edition, 1983.
- [18] E. Giné, R. Nickl, *Mathematical Foundations of Infinite-Dimensional Statistical Models*, number 40 in Cambridge Series in Statistical and Probabilistic Mathematics, Cambridge University Press, 2015.
- [19] M. Girolami, E. Febrianto, G. Yin, F. Cirak, The statistical finite element method (statFEM) for coherent synthesis of observation data and model predictions, *Computer Methods in Applied Mechanics and Engineering* 375 (2021) 113533.
- [20] T. Graepel, Solving noisy linear operator equations by Gaussian processes: application to ordinary and partial differential equations, in: *Proceedings of the 20th International Conference on Machine Learning*, pp. 234–241.
- [21] M. Kanagawa, B. K. Sriperumbudur, K. Fukumizu, Convergence analysis of deterministic kernel-based quadrature rules in misspecified settings, *Foundations of Computational Mathematics* 20 (2020) 155–194.
- [22] E. J. Kansa, Multiquadrics—A scattered data approximation scheme with applications to computational fluid-dynamics—II solutions to parabolic, hyperbolic and elliptic partial differential equations, *Computers & Mathematics with Applications* 19 (1990) 147–161.
- [23] T. Karvonen, Asymptotic bounds for smoothness parameter estimates in Gaussian process interpolation, *SIAM/ASA Journal on Uncertainty Quantification* 11 (2023) 1225–1257.
- [24] T. Karvonen, G. Wynne, F. Tronarp, C. J. Oates, S. Särkkä, Maximum likelihood estimation and uncertainty quantification for Gaussian process approximation of deterministic functions, *SIAM/ASA Journal on Uncertainty Quantification* 8 (2020) 926–958.
- [25] M. C. Kennedy, A. O’Hagan, Bayesian calibration of computer models, *Journal of the Royal Statistical Society: Series B* 63 (2002) 425–464.
- [26] P. Knabner, L. Angermann, *Numerical Methods for Elliptic and Parabolic Partial Differential Equations. With Contributions by Andreas Rupp*, number 44 in Texts in Applied Mathematics, Springer, Berlin, 2021.
- [27] K. J. Koh, F. Cirak, Stochastic PDE representation of random fields for large-scale Gaussian process regression and statistical finite element analysis, *Computer Methods in Applied Mechanics and Engineering* 417 (2023) 116358.
- [28] D. Krieg, M. Sonnleitner, Random points are optimal for the approximation of Sobolev functions, *IMA Journal of Numerical Analysis* 44 (2024) 1346–1371.
- [29] W. V. Li, W. Linde, Approximation, metric entropy and small ball estimates for Gaussian measures, *The Annals of Probability* 27 (1999) 1556–1578.
- [30] G. J. Lord, C. E. Powell, T. Shardlow, *An Introduction to Computational Stochastic PDEs*, Cambridge University Press, 2014.
- [31] J. Nitsche,  $L_\infty$ -convergence of finite element approximations, in: *Mathematical Aspects of Finite Element Methods*, number 606 in Lecture Notes in Mathematics, Springer-Verlag, 1977, pp. 261–274.
- [32] H. Owhadi, Bayesian numerical homogenization, *Multiscale Modeling & Simulation* 13 (2015) 812–828.
- [33] Y. Papandreou, J. Cockayne, M. Girolami, A. B. Duncan, Theoretical guarantees for the statistical finite element method, *SIAM/ASA Journal on Uncertainty Quantification* 11 (2023) 1278–1307.
- [34] V. I. Paulsen, M. Raghupathi, *An Introduction to the Theory of Reproducing Kernel Hilbert Spaces*, number 152 in Cambridge Studies in Advanced Mathematics, Cambridge University Press, 2016.
- [35] M. Raissi, P. Perdikaris, G. E. Karniadakis, Machine learning of linear differential equations using Gaussian processes, *Journal of Computational Physics* 348 (2017) 683–693.
- [36] C. E. Rasmussen, C. K. I. Williams, *Gaussian Processes for Machine Learning*, Adaptive Computation and Machine Learning, MIT Press, Boston, 2006.
- [37] A. H. Schatz, L. B. Wahlbin, Interior maximum norm estimates for finite element methods, *Mathematics of Computation* 31 (1977) 414–442.
- [38] A. H. Schatz, L. B. Wahlbin, Maximum norm estimates in the finite element method on plane polygonal domains. I, *Mathematics of Computation* 32 (1978) 73–109.
- [39] A. H. Schatz, L. B. Wahlbin, Interior maximum norm estimates for finite element methods. II, *Mathematics of Computation* 64 (1995) 907–928.
- [40] B. Szabó, A. W. van der Vaart, J. H. van Zanten, Frequentist coverage of adaptive nonparametric Bayesian credible sets, *The Annals of Statistics* 43 (2015) 1391–1428.
- [41] A. L. Teckentrup, Convergence of Gaussian process regression with estimated hyper-parameters and applications in Bayesian inverse problems, *SIAM/ASA Journal on Uncertainty Quantification* 8 (2020) 1310–1337.
- [42] C. Travelleri, D. Ginsbourger, Disintegration of Gaussian measures for sequential assimilation of linear operator data, *Electronic Journal of Statistics* 18 (2024) 3825–3857.
- [43] A. B. Tsybakov, *Introduction to Nonparametric Estimation*, Springer Series in Statistics, Springer, Berlin, 2009.
- [44] A. van der Vaart, H. van Zanten, Information rates of nonparametric Gaussian process methods, *Journal of Machine Learning Research* 12 (2011) 2095–2119.
- [45] A. W. van der Vaart, J. H. van Zanten, Reproducing kernel Hilbert spaces of Gaussian priors, volume 3 of *IMS Collections*, Institute of Mathematical Statistics, 2008, pp. 200–222.
- [46] W. Wang, On the inference of applying Gaussian process modeling to a deterministic function, *Electronic Journal of Statistics* 15 (2021) 5014–5066.
- [47] W. Wang, R. Tuo, C. F. J. Wu, On prediction properties of kriging: uniform error bounds and robustness, *Journal of the American Statistical Association* 115 (2020) 920–930.
- [48] H. Wendland, *Scattered Data Approximation*, number 17 in Cambridge Monographs on Applied and Computational Mathematics, Cambridge University Press, 2005.
- [49] G. Wynne, F.-X. Briol, M. Girolami, Convergence guarantees for Gaussian process means with misspecified likelihoods and smoothness, *Journal of Machine Learning Research* 22 (2021) 1–40.