Probabilistic methods for elliptic partial differential equations

Assyr Abdulle* Giacomo Garegnani*

Abstract

AMS subject classification.

Keywords.

1 Introduction

TO DO [1, 3, 7, 12, 13]

$$\begin{aligned} -\nabla \cdot (\kappa \nabla u) &= f, \quad \text{in } D, \\ u &= g, \quad \text{on } \partial D. \end{aligned} \tag{1}$$

important: review of probabilistic methods for PDEs and ODEs. Have PDEs really been treated already? How? Inverse problems: what is the current state of things? Has anyone gone infinite dimensional?

2 Random mesh probabilistic Finite Elements

Weak formulation: bilinear form $a\colon V \times V \to \mathbb{R}$ and a linear functional $F\colon V \to \mathbb{R}$ satisfying the usual continuity and coercivity constraints, look for $u \in V$ satisfying

$$a(u, v) = F(v), \tag{2}$$

for all functions $v \in V$. Galerkin formulation: for $V_h \subset V$ such that $\dim V_h < \infty$, find $u_h \in V_h$ such that

$$a(u_h, v_h) = f(v_h), \quad \forall v_h \in V_h, \tag{3}$$

for all $v_h \in V_h$. Given a triangulation $\mathcal{T}_h$ of the domain D, we choose V_h to be the space of linear finite elements, i.e., $V_h = X_h^1 \cap V$, where

$$X_h^1 = \{v_h \in C^0(\overline{D})\colon v_h|_K \in \mathcal{P}_1, \text{ for all } K \in \mathcal{T}_h\}, \tag{4}$$

and where $\mathcal{P}_1$ is the space of polynomials of degree at most one. The finite element space can be written then as $V_h = \operatorname{span}\{\varphi_i\}_{i=1}^N$, where the basis $\{\varphi_i\}_{i=1}^N$ are the Lagrange basis functions. Hence, each $v_h \in V_h$ can be written as $v_h = \sum_{i=1}^N v_i \varphi_i$, where v_i are the coefficients of v_h on the basis $\{\varphi_i\}_{i=1}^N$. Our probabilistic method is based on a randomly perturbed mesh $\widetilde{\mathcal{T}}_h$ which is defined as follows.

*Institute of Mathematics, École Polytechnique Fédérale de Lausanne ({assyr.abdulle, giacomo.garegnani}@epfl.ch)

Definition 1. Let us consider a domain $D \subset \mathbb{R}^d$ and a triangulation $\mathcal{T}_h$ characterised by the maximum diameter $h > 0$ of its elements and by the set of vertices $\mathcal{N}_h = \{x_i\}_{i=1}^N$ such that $\mathcal{N}_h = \mathcal{N}_h^I \cup \mathcal{N}_h^B$, where $\mathcal{N}_h^I$ and $\mathcal{N}_h^B$ are the vertices in the interior of D and on ∂D respectively, and where we denote $N_I = |\mathcal{N}_h^I|$ and $N_B = |\mathcal{N}_h^B|$. Given a probability space (Ω, Σ, μ), the random mesh $\widetilde{\mathcal{T}}_h$ is defined by a sequence of random variables $\{\alpha_i\}_{i=1}^{N_I}$, $\alpha_i \colon \Omega \to \mathbb{R}^d$, which are used to perturb the internal nodes as

$$\tilde{x}_i = x_i + \bar{h}_i^p \alpha_i, \ x_i \in \mathcal{N}_h^I \tag{5}$$

where $p \geq 1$ and $\bar{h}_i$ is defined as the minimum diameter of the elements K having x_i as a vertex, i.e.

$$\bar{h}_i = \min_{K \in \Delta(x_i)} h_K, \tag{6}$$

where $\Delta(x_i)$ is such set of elements. The vertices laying on ∂D in $\mathcal{T}_h$ are unperturbed in $\widetilde{\mathcal{T}}_h$.

Once the perturbed mesh $\widetilde{\mathcal{T}}_h$ is obtained, let us denote by $\widetilde{V}_h$ the piecewise linear finite element space defined on $\widetilde{\mathcal{T}}_h$. Let us remark that the space $\widetilde{V}_h = \widetilde{V}_h(\omega)$ is random itself, i.e., for each realisation of the random variables $\{\alpha_i\}_{i=1}^{N_I}$ we obtain a different perturbed finite element space.

Definition 2. With the notation above, the probabilistic solution $\tilde{u}_h \colon \Omega \times D \to \mathbb{R}$ is a random field satisfying for all $\omega \in \Omega$

$$\tilde{u}_h(\omega, \cdot) \in \widetilde{V}_h(\omega), \text{ s.t. } a(\tilde{u}_h(\omega, \cdot), \tilde{v}_h) = F(\tilde{v}_h), \text{ for all } \tilde{v}_h \in \widetilde{V}_h(\omega). \tag{7}$$

Let us finally introduce the following assumption on the random variables defining the mesh perturbation.

Assumption 1. The random variables α_i are chosen such that the perturbed mesh $\widetilde{\mathcal{T}}_h$ has the same topology of the mesh $\mathcal{T}_h$ (e.g., no exchange of vertices in one-dimension and no crossing edges in two-dimensions) almost surely.

3 A priori error analysis

Close finite element spaces, sum space [2, 10]. Introduce the main ideas.

Definition 3. Given two possibly disjoint linear finite element spaces V_h and $\widetilde{V}_h$, we denote by $V_h^+ = V_h + \widetilde{V}_h$ the space of functions that can be written as the sum of a function in V_h and a function in $\widetilde{V}_h$, i.e., for any function $v_h^+ \in V_h^+$ there exists functions $v_h \in V_h$ and $\tilde{v}_h \in \widetilde{V}_h$ such that $v_h^+ = v_h + \tilde{v}_h$.

Remark 1. Clearly from the definition, we have the inclusion $(V_h \cup \widetilde{V}_h) \subseteq V_h^+$. Moreover, with homogeneous boundary conditions $V_h \cap \widetilde{V}_h = \{0\}$ almost surely. Hence, $V_h^+ = V_h \oplus \widetilde{V}_h$ and the decomposition of $v_h^+ \in V_h^+$ is unique.

Lemma 1. *The solutions $u_h \in V_h$ and $\tilde{u}_h \in \widetilde{V}_h$ satisfy almost surely in Ω*

$$\begin{aligned}\|u_h - \tilde{u}_h\|_V^2 \leq \left(\frac{M}{\alpha}\right)^2 \inf_{v_h, w_h \in V_h} \Bigg(\inf_{\tilde{v}_h, \tilde{w}_h \in \widetilde{V}_h} \Big(&\left\|u_h^+ - w_h\right\|_V \|\tilde{u}_h - v_h\|_V + \\ &\left\|u_h^+ - \tilde{w}_h\right\|_V \|u_h - \tilde{v}_h\|_V \Big) \Bigg),\end{aligned} \tag{8}$$

where $u_h^+ \in V_h^+$ is the solution of

$$a(u_h^+, v_h^+) = F(v_h^+), \tag{9}$$

for all $v_h^+ \in V_h^+$.

Proof. Since V_h and $\widetilde{V}_h$ are both subspaces of V_h^+, we have

$$\begin{aligned} a(u_h^+ - u_h, v_h) &= 0, \quad \forall v_h \in V_h, \\ a(u_h^+ - \tilde{u}_h, \tilde{v}_h) &= 0, \quad \forall \tilde{v}_h \in \widetilde{V}_h, \end{aligned} \tag{10}$$

which means that u_h and $\tilde{u}_h$ are the elliptic projection of u_h^+ onto V_h and $\widetilde{V}_h$ respectively. Hence, thanks to Cea's lemma

$$\begin{aligned} \left\|u_h^+ - u_h\right\|_V &\leq \frac{M}{\alpha} \left\|u_h^+ - w_h\right\|_V, \\ \left\|u_h^+ - \tilde{u}_h\right\|_V &\leq \frac{M}{\alpha} \left\|u_h^+ - \tilde{w}_h\right\|_V, \end{aligned} \tag{11}$$

for all $w_h \in V_h$ and $\tilde{w}_h \in \widetilde{V}_h$. Using the coercivity on V of $a(\cdot,\cdot)$, adding and subtracting $a(u_h^+, u_h - \tilde{u}_h)$ and thanks to the orthogonality property (10) we have for all $v_h \in V_h$ and $\tilde{v}_h \in \widetilde{V}_h$

$$\begin{aligned} \alpha \left\|u_h - \tilde{u}_h\right\|_V^2 &\leq a(u_h - \tilde{u}_h, u_h - \tilde{u}_h) \\ &= a(u_h - u_h^+, u_h - \tilde{u}_h) + a(u_h^+ - \tilde{u}_h, u_h - \tilde{u}_h) \\ &= a(u_h - u_h^+, v_h - \tilde{u}_h) + a(u_h^+ - \tilde{u}_h, u_h - \tilde{v}_h). \end{aligned} \tag{12}$$

Thanks to the continuity of the bilinear form we then have for all $w_h \in V_h$ and $\tilde{w}_h \in \widetilde{V}_h$

$$\begin{aligned} \alpha \left\|u_h - \tilde{u}_h\right\|_V^2 &\leq M\Big(\left\|u_h^+ - u_h\right\|_V \left\|\tilde{u}_h - v_h\right\|_V + \left\|u_h^+ - \tilde{u}_h\right\|_V \left\|u_h - \tilde{v}_h\right\|_V \Big) \\ &\leq \frac{M^2}{\alpha}\Big(\left\|u_h^+ - w_h\right\|_V \left\|\tilde{u}_h - v_h\right\|_V + \left\|u_h^+ - \tilde{w}_h\right\|_V \left\|u_h - \tilde{v}_h\right\|_V \Big), \end{aligned} \tag{13}$$

which implies the result taking the infimum at the right hand side. □

Let us remark that the lemma above is valid for any choice of V_h and $\widetilde{V}_h$. In order to study the approximation of a function v_h of V_h with functions of $\widetilde{V}_h$, we exploit the Legendre interpolation.

Definition 4. We denote by $\Pi_h \colon V \to V_h$ and $\widetilde{\Pi}_h \colon V \to \widetilde{V}_h$ the Legendre piecewise linear interpolation operators on V_h and $\widetilde{V}_h$ respectively, i.e.,

$$\begin{aligned} \Pi_h v(x) &= \sum_{x_j \in \mathcal{N}_h^I} v(x_j)\varphi_j(x), \\ \widetilde{\Pi}_h v(x) &= \sum_{\tilde{x}_j \in \widetilde{\mathcal{N}}_h^I} v(x_j)\tilde{\varphi}_j(x), \end{aligned} \tag{14}$$

where $\{\varphi_i\}_{i=1}^{N^I}$ and $\{\tilde{\varphi}_i\}_{i=1}^{N^I}$ are the basis functions of V_h and $\widetilde{V}_h$ respectively.

In the following lemma we characterise the Legendre interpolant $\widetilde{V}_h$ through its values on the internal nodes of the original mesh $\mathcal{T}_h$.

Lemma 2. *Let V_h and $\widetilde{V}_h$ be defined as in Definition 1. Then for all $v_h \in V_h$ and all $x_i \in \mathcal{N}_h^I$ it holds*

$$\widetilde{\Pi}_h v_h(x_i) - v_h(x_i) = \bar{h}_i^p \alpha_i^\top \Big(\nabla v_h(\tilde{x}_i) - \nabla \widetilde{\Pi}_h v_h(x_i)\Big). \tag{15}$$

Proof. Let us denote $e_h = \widetilde{\Pi}_h v_h - v_h$. An exact Taylor expansion of the linear basis function $\tilde{\varphi}_j$ gives

$$\begin{aligned} e_h(x_i) &= \sum_j v_h(\tilde{x}_j)\varphi_j(x_i) - v_h(x_i) \\ &= \sum_j v_h(\tilde{x}_j)\Big(\tilde{\varphi}_j(\tilde{x}_i) - \bar{h}_i^p \alpha_i^\top \nabla\tilde{\varphi}_j(x_i)\Big) - v_h(x_i) \\ &= v_h(\tilde{x}_i) - v_h(x_i) - \sum_j \bar{h}_i^p \alpha_i^\top v_h(\tilde{x}_j) \nabla\tilde{\varphi}_j(x_i). \end{aligned} \tag{16}$$

We can now expand the function v_h, which is linear on the segment connecting x_i and $\tilde{x}_i$, as

$$v_h(\tilde{x}_i) = v_h(x_i) + \bar{h}_i^p \alpha_i^\top \nabla v_h(\tilde{x}_i), \tag{17}$$

to obtain

$$\begin{aligned} e_h(x_i) &= \bar{h}_i^p \alpha_i^\top \nabla v_h(\tilde{x}_i) - \bar{h}_i^p \alpha_i^\top \sum_j v_h(\tilde{x}_j) \nabla \tilde{\varphi}_j(x_i) \\ &= \bar{h}_i^p \alpha_i^\top \Big(\nabla v_h(\tilde{x}_i) - \nabla \widetilde{\Pi}_h v_h(x_i) \Big), \end{aligned} \tag{18}$$

which is the desired result. □

Let us remark that the quality of the approximation given by the operator $\widetilde{\Pi}_h$ depends on the regularity of v_h, as it will be clarified in the following. In order to exploit Lemma 1, it is necessary to control the interpolation of a function in V_h^+ on both V_h and $\widetilde{V}_h$, which is explained in the following lemma.

Lemma 3. *For all $v_h^+ \in V_h^+$ it holds*

$$\begin{aligned} \Pi_h v_h^+ - v_h^+ &= \Pi_h \tilde{v}_h - \tilde{v}_h, \\ \widetilde{\Pi}_h v_h^+ - v_h^+ &= \Pi_h v_h - v_h, \end{aligned} \tag{19}$$

where $v_h \in V_h$ and $\tilde{v}_h \in \widetilde{V}_h$ are the components of v_h^+ in V_h and $\widetilde{V}_h$ respectively, i.e., $v_h^+ = v_h + \tilde{v}_h$.

Proof. The result can be obtained directly as $\Pi_h v_h = v_h$ for all $v_h \in V_h$ and $\widetilde{\Pi}_h \tilde{v}_h = \tilde{v}_h$ for all $\tilde{v}_h \in \widetilde{V}_h$. □

Let us now introduce the smoothness assumption we impose on the solution, which will constitute the framework for our analysis.

Assumption 2. The exact solution u of (1) is in $W^{2,\infty}(D)$.

Under this assumption, the following result of uniform convergence holds.

Theorem 1 (Theorem 3.3.7 of [4]). *Under Assumption 2, the Finite Element solution $u_h \in V_h$ satisfies*

$$\begin{aligned} \|u - u_h\|_{L^\infty(D)} &\leq Ch^2 \left|\log h\right|^{3/2} |u|_{W^{2,\infty}(D)}, \\ |u - u_h|_{W^{1,\infty}(D)} &\leq Ch \left|\log h\right| |u|_{W^{2,\infty}(D)}, \end{aligned} \tag{20}$$

where $C > 0$ is a constant independent of h.

3.1 One-dimensional case

Lemma 4. *Let $D = (0,1)$, V_h be the finite element space defined on the regular grid $\{x_i = ih\}_{i=0}^N$ for $h = N^{-1}$ and $v_h \in V_h$. If there exists $v \in W^{2,\infty}(D)$ such that*

$$\|v - v_h\|_{W^{1,\infty}(D)} \leq Ch \left|\log h\right| |v|_{W^{2,\infty}(D)}, \tag{21}$$

then almost surely in Ω

$$\begin{aligned} \left\| v_h - \widetilde{\Pi}_h v_h \right\|_{L^\infty(D)} &\leq Ch^{p+1} \left|\log h\right|, \\ \left| v_h - \widetilde{\Pi}_h v_h \right|_{H^1(D)} &\leq Ch^{(p+1)/2} \left|\log h\right|, \end{aligned} \tag{22}$$

where $C > 0$ is a constant independent of h.

Proof. Let us denote $e_h = \widetilde{\Pi}_h v_h - v_h$. By definition $e_h(\tilde{x}_i) = 0$ for all $i = 0, \dots, N$ and thanks to Lemma 2, in this setting

$$e_h(x_i) = h^p \alpha_i \big(v_h'(\tilde{x}_i) - (\widetilde{\Pi}_h v_h)'(x_i)\big). \tag{23}$$

In the following, we denote by ε_i the quantity

$$\varepsilon_i = v_h'(\tilde{x}_i) - (\widetilde{\Pi}_h v_h)'(x_i). \tag{24}$$

Before proceeding to the actual estimation of the error, we shall now bound $|\varepsilon_i|$. Let us consider without loss of generality $\alpha_i > 0$. In this case, it is clear that

$$\varepsilon_i = v_h'\big|_{K_{i+1}} - (\widetilde{\Pi}_h v_h)'\big|_{\widetilde{K}_i}. \tag{25}$$

Let us decompose ε_i in two terms $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$, defined as

$$\begin{aligned} \varepsilon_{i,1} &= v_h'\big|_{K_{i+1}} - v_h'\big|_{K_i}, \\ \varepsilon_{i,2} &= v_h'\big|_{K_i} - (\widetilde{\Pi}_h v_h)'\big|_{\widetilde{K}_i}, \end{aligned} \tag{26}$$

so that $|\varepsilon_i| \leq |\varepsilon_{i,1}| + |\varepsilon_{i,2}|$. Thanks to (21) we have

$$\begin{aligned} |\varepsilon_{i,1}| &\leq 2\, |v_h - v|_{W^{1,\infty}(D)} + |v'(\bar{x}_i) - v'(\bar{x}_{i+1})| \\ &\leq Ch\, |\log h|\, |v|_{W^{2,\infty}(D)}\,, \end{aligned} \tag{27}$$

where we denote by $\bar{x}_j$ the midpoint of element K_j. Let us now consider $\varepsilon_{i,2}$. First, thanks to (17) and as we assumed $\alpha_i > 0$

$$\begin{aligned} (\widetilde{\Pi}_h v_h)'\big|_{\widetilde{K}_i} &= \frac{1}{h + (\alpha_i - \alpha_{i-1})h^p} \big(v_h(\tilde{x}_i) - v_h(\tilde{x}_{i-1})\big) \\ &= \frac{1}{h + (\alpha_i - \alpha_{i-1})h^p} \Big(v_h(x_i) - v_h(x_{i-1}) + \big(\alpha_i v_h'(\tilde{x}_i) - \alpha_{i-1} v_h'(\tilde{x}_{i-1})\big) h^p\Big) \\ &= \frac{1}{1 + (\alpha_i - \alpha_{i-1})h^{p-1}} \Big(v_h'\big|_{K_i} + \big(\alpha_i v_h'\big|_{K_{i+1}} - \alpha_{i-1} v_h'(\tilde{x}_{i-1})\big) h^{p-1}\Big), \end{aligned} \tag{28}$$

where $v_h'(\tilde{x}_{i-1})$ could alternatively be equal to the derivative of v_h on K_{i-1} or K_i depending on the sign of α_{i-1}. Let us now remark that

$$\frac{1}{1 + (\alpha_i - \alpha_{i-1})h^{p-1}} = 1 + R, \tag{29}$$

where

$$R = \sum_{k=1}^{\infty} (-1)^k (\alpha_i - \alpha_{i-1})^k h^{k(p-1)}, \tag{30}$$

which implies

$$\begin{aligned} (\widetilde{\Pi}_h v_h)'\big|_{\widetilde{K}_i} = v_h'\big|_{K_i} &+ \big(\alpha_i v_h'\big|_{K_{i+1}} - \alpha_{i-1} v_h'(\tilde{x}_{i-1})\big) h^{p-1} \\ &+ R\Big(v_h'\big|_{K_i} + \big(\alpha_i v_h'\big|_{K_{i+1}} - \alpha_{i-1} v_h'(\tilde{x}_{i-1})\big) h^{p-1}\Big), \end{aligned} \tag{31}$$

and finally

$$\begin{aligned} |\varepsilon_{i,2}| = \Big|&\big(\alpha_i v_h'\big|_{K_{i+1}} - \alpha_{i-1} v_h'(\tilde{x}_{i-1})\big) h^{p-1} \\ &+ R\Big(v_h'\big|_{K_i} + \big(\alpha_i v_h'\big|_{K_{i+1}} - \alpha_{i-1} v_h'(\tilde{x}_{i-1})\big) h^{p-1}\Big)\Big|. \end{aligned} \tag{32}$$

Let us rewrite the expression above regrouping the terms by powers of h as

$$\begin{aligned} |\varepsilon_{i,2}| &= \left| \sum_{k=1}^{\infty} (-1)^k (\alpha_i - \alpha_{i-1})^{k-1} \Big((\alpha_i - \alpha_{i-1}) v_h'\big|_{K_i} - \big(\alpha_i v_h'\big|_{K_{i+1}} - \alpha_{i-1} v_h'(\tilde{x}_{i-1})\big)\Big) h^{k(p-1)} \right| \\ &= \left| \sum_{k=1}^{\infty} (-1)^k (\alpha_i - \alpha_{i-1})^{k-1} \Big(\alpha_i \big(v_h'\big|_{K_i} - v_h'\big|_{K_{i+1}}\big) + \alpha_{i-1}\big(v_h'(\tilde{x}_{i-1}) - v_h'\big|_{K_i}\big)\Big) h^{k(p-1)} \right| \\ &\leq \sum_{k=1}^{\infty} |\alpha_i - \alpha_{i-1}|^{k-1} \Big(|\alpha_i| \left| v_h'\big|_{K_i} - v_h'\big|_{K_{i+1}} \right| + |\alpha_{i-1}| \left| v_h'\big|_{K_i} - v_h'(\tilde{x}_{i-1}) \right| \Big) h^{k(p-1)}. \end{aligned} \tag{33}$$

Hence, thanks to (27) and to the hypotheses on the coefficients α_j, we have

$$|\varepsilon_{i,2}| \le Ch^p \left|\log h\right| |v|_{W^{2,\infty}(D)}\,, \tag{34}$$

which finally implies $|\varepsilon_i| \le Ch \left|\log h\right|$ as $p \ge 1$. This means that the interpolation error satisfies on the nodes of the original mesh

$$|e_h(x_i)| \le Ch^{p+1} \left|\log h\right| |v|_{W^{2,\infty}(D)}\,, \tag{35}$$

and on the nodes of the perturbed mesh by definition $e_h(\tilde{x}_i) = 0$. We can now proceed to the estimation of the error. Let us now remark that since $e_h \in V_h^+$, the maximum of e_h has to be realised on one of the nodes of the original mesh. Hence

$$\|e_h\|_{L^\infty(D)} = \max_{x_i \in \mathcal{N}_h^I} |e_h(x_i)| \le Ch^{p+1} \left|\log h\right| |v|_{W^{2,\infty}(D)}\,, \tag{36}$$

which implies the desired result for all L^q spaces with $1 \le q < \infty$. Let us now consider the H^1 semi-norm. We fix a single element of the modified grid $\widetilde{K}_i = (\tilde{x}_{i-1}, \tilde{x}_i)$. There are four distinct cases depending on the sign of the random variables α_{i-1} and α_i. First, if $\alpha_{i-1} > 0$ and $\alpha_i < 0$, i.e., $\widetilde{K}_i \subset K_i$, $e_h \equiv 0$ and we have trivially $|e_h|_{H^1(\tilde{K})} = 0$. We now fix $\alpha_{i-1} < 0$ and $\alpha_i < 0$, getting explicitly

$$\begin{aligned}\int_{\widetilde{K}_i} (e_h')^2 &= \alpha_{i-1}\varepsilon_{i-1}^2 h^p \left(1 + \frac{\alpha_{i-1}h^{p-1}}{1+\alpha_i h^{p-1}}\right) \\ &\le Ch^{p+2} \left|\log h\right|^2.\end{aligned} \tag{37}$$

The same result is obtained when both α_{i-1} and α_i are positive. For the last case, i.e. $K_i \subset \widetilde{K}_i$, we have

$$\begin{aligned}\int_{\widetilde{K}_i} (e_h')^2 &= h^p(\alpha_{i-1}\varepsilon_{i-1}^2 + \alpha_i\varepsilon_i^2) + h^{2p-1}(\alpha_i\varepsilon_i - \alpha_{i-1}\varepsilon_{i-1})^2 \\ &\le Ch^{p+2} \left|\log h\right|^2.\end{aligned} \tag{38}$$

Summing over the elements we have $|e_h|_{H^1(D)} \le Ch^{(p+1)/2} \left|\log h\right|$, which gives the desired result. □

Remark 2. The same result holds exchanging the roles of V_h and $\widetilde{V}_h$, i.e., considering the interpolant Π_h instead of $\widetilde{\Pi}_h$.

Theorem 2. *Under Assumption 2 and in the one-dimensional setting of Lemma 4, the probabilistic solution $\tilde{u}_h$ satisfies almost surely in Ω*

$$\|u_h - \tilde{u}_h\|_V \le Ch^{(p+1)/2} \left|\log h\right|, \tag{39}$$

where $C > 0$ is a constant independent of h. Moreover,

$$\|u - \tilde{u}_h\|_V \le Ch, \tag{40}$$

holds almost surely in Ω.

Proof. Under Assumption 2, condition (21) is verified for u_h and u the solution of (1) thanks to Theorem 1. Hence, choosing in Lemma 1 $v_h = \widetilde{\Pi}_h u_h$, $\tilde{v}_h = \Pi_h \tilde{u}_h$, $w_h = \Pi_h u_h^+$ and $\tilde{w}_h = \widetilde{\Pi}_h u_h^+$ and then applying Lemma 4 and Lemma 3 gives the result. The estimate involving the exact solution is then obtained with the triangular inequality and as $p \ge 1$. □

Remark 3. The convergence result in Theorem 2 could be seen as the PDE counterpart of the ODE convergences shown in [1, 6]. However, Theorem 2 holds almost surely, whereas for ODEs only mean-square and weak convergences have been taken into account.

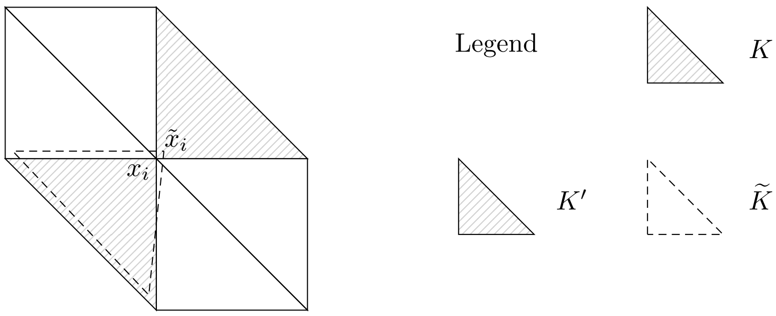

Figure 1: Scheme for error analysis.

3.2 Two-dimensional case

Establish Lemma 4 in the two-dimensional case

Lemma 5. *Restate Lemma 4 in the 2d case.*

Proof. Let us denote $e_h = \widetilde{\Pi}_h v_h - v_h$. By definition $e_h(\tilde{x}_i) = 0$ for all $i = 0, \dots, N$ and thanks to Lemma 2

$$e_h(x_i) = h^p \alpha_i \big(\nabla v_h(\tilde{x}_i) - \nabla \widetilde{\Pi}_h v_h(x_i)\big). \tag{41}$$

In the following, we denote by ε_i the quantity

$$\varepsilon_i = \nabla v_h(\tilde{x}_i) - \nabla \widetilde{\Pi}_h v_h(x_i). \tag{42}$$

Before proceeding to the actual estimation of the error, we shall now bound $|\varepsilon_i|$. Let us denote by K an arbitrary element which has x_i as a vertex and such that the corresponding element $\widetilde{K}$ in the perturbed mesh contains x_i. Furthermore, let us denote by K' the element in the original mesh containing $\tilde{x}_i$. We refer to Fig. 1 for a schematic representation of these elements. With this notation, we have $\nabla v_h(\tilde{x}_i) = \nabla v_h\big|_{K'}$, $\nabla \widetilde{\Pi}_h v_h(x_i) = \nabla \widetilde{\Pi}_h v_h\big|_{\widetilde{K}}$, and we can then decompose ε_i in two terms $\varepsilon_{i,1}$ and $\varepsilon_{i,2}$, defined as

$$\begin{aligned} \varepsilon_{i,1} &= \nabla v_h\big|_{K'} - \nabla v_h\big|_{K}, \\ \varepsilon_{i,2} &= \nabla v_h\big|_{K} - \nabla \widetilde{\Pi}_h v_h\big|_{\widetilde{K}}, \end{aligned} \tag{43}$$

so that $|\varepsilon_i| \leq |\varepsilon_{i,1}| + |\varepsilon_{i,2}|$. Thanks to Theorem 1 we have

$$\begin{aligned} |\varepsilon_{i,1}| &\leq 2\,|v_h - v|_{W^{1,\infty}(D)} + |\nabla v(\bar{x}_K) - \nabla v(\bar{x}_{K'})| \\ &\leq Ch\,|\log h|\,|v|_{W^{2,\infty}(D)}\,, \end{aligned} \tag{44}$$

where we denote by $\bar{x}_K$ the barycenter of element K. Let us now consider $\varepsilon_{i,2}$. We denote by B_K and $B_{\widetilde{K}}$ the matrices identifying the affine transformations mapping K and $\widetilde{K}$ to the reference triangle $\widehat{K}$, with vertices $\hat{x}_1 = (0,0)^\top$, $\hat{x}_2 = (1,0)^\top$, $\hat{x}_3 = (0,1)^\top$. In particular, ordering arbitrarily the vertices of K and $\widetilde{K}$ and calling them x_i, $\tilde{x}_i$, $i = 1,2,3$ respectively, we have

$$\begin{aligned} B_K &= \big(x_2 - x_1 \mid x_3 - x_1\big)\,, \\ B_{\widetilde{K}} &= \big(\tilde{x}_2 - \tilde{x}_1 \mid \tilde{x}_3 - \tilde{x}_1\big) \\ &= B_K + h^p \Lambda, \end{aligned} \tag{45}$$

where

$$\Lambda = \big(\alpha_2 - \alpha_1 \mid \alpha_3 - \alpha_1\big)\,. \tag{46}$$

Then (introduce notations)

$$\Big|\nabla v_h\big|_K - \nabla \widetilde{\Pi}_h v_h\big|_{\widetilde{K}}\Big| = \Big|B_K^{-\top} \nabla \widehat{v_h} - B_{\widetilde{K}}^{-\top} \nabla \widehat{\widetilde{\Pi}_h v_h}\Big|\,. \tag{47}$$

Thanks to (17) we can write

$$\nabla \widehat{\widetilde{\Pi}_h v_h} = \nabla \widehat{v_h} + h^p \gamma, \tag{48}$$

where the vector γ is given by

$$\gamma = \begin{pmatrix} \alpha_2^\top \nabla v_h(\tilde{x}_2) - \alpha_1^\top \nabla v_h(\tilde{x}_1) \\ \alpha_3^\top \nabla v_h(\tilde{x}_3) - \alpha_1^\top \nabla v_h(\tilde{x}_1) \end{pmatrix}. \tag{49}$$

With algebraic operations we can rewrite $B_{\widetilde{K}}^{-\top}$ as

$$\begin{aligned} B_{\widetilde{K}}^{-\top} &= (B_K + h^p \Lambda)^{-\top} \\ &= B_K^{-\top}(I + h^p B_K^{-1}\Lambda)^{-\top} \\ &= B_K^{-\top}(I - h^p \Gamma), \end{aligned} \tag{50}$$

where the matrix Γ is given by the series expansion $\Gamma = \sum_{j=0}^{\infty} \Gamma_j$, with

$$\Gamma_j = (-1)^j h^{pj} (\Lambda^\top B_K^{-\top})^{j+1}. \tag{51}$$

Since $\left|B_K^{-1}\right| \leq Ch$ [4, Theorem 3.1.3], the single addends Γ_j satisfy $|\Gamma_j| \leq Ch^{j(p-1)-1}$. Let us remark that the expression above implies $\Gamma_j = -h^p \Gamma_{j-1}\Gamma_0$. Rearranging the terms, we finally get

$$\begin{aligned} \left| \nabla v_h\big|_K - \nabla \widetilde{\Pi}_h v_h\big|_{\widetilde{K}} \right| &= h^p \left| B_K^{-\top}(-\gamma + \Gamma \nabla \widehat{v_h} + h^p \Gamma \gamma) \right| \\ &= h^p \left| B_K^{-\top} \Big(-\gamma + \Gamma_0 \nabla \widehat{v_h} + \sum_{j=1}^{\infty} (\Gamma_j \nabla \widehat{v_h} + h^p \Gamma_{j-1}\gamma) \Big) \right| \\ &= h^p \left| B_K^{-\top} h^p \sum_{j=0}^{\infty} \Gamma_{j-1}(\gamma - \Gamma_0 \nabla \widehat{v_h}) \right|, \end{aligned} \tag{52}$$

where we define $\Gamma_{-1} = -h^{-p} I$. We can now compute explicitly the difference $\gamma - \Gamma_0 \nabla \widehat{v_h}$ as

$$\begin{aligned} \gamma - \Gamma_0 \nabla \widehat{v_h} &= \gamma - \Lambda^\top B_K^{-\top} \nabla \widehat{v_h} \\ &= \gamma - \Lambda^\top \nabla v_h\big|_K \\ &= \begin{pmatrix} \alpha_2^\top \nabla v_h(\tilde{x}_2) - \alpha_1^\top \nabla v_h(\tilde{x}_1) \\ \alpha_3^\top \nabla v_h(\tilde{x}_3) - \alpha_1^\top \nabla v_h(\tilde{x}_1) \end{pmatrix} - \begin{pmatrix} (\alpha_2^\top - \alpha_1^\top) \nabla v_h\big|_K \\ (\alpha_3^\top - \alpha_1^\top) \nabla v_h\big|_K \end{pmatrix} \\ &= \begin{pmatrix} \alpha_2^\top \left(\nabla v_h(\tilde{x}_2) - \nabla v_h\big|_K\right) + \alpha_1^\top \left(\nabla v_h\big|_K - \nabla v_h(\tilde{x}_1)\right) \\ \alpha_3^\top \left(\nabla v_h(\tilde{x}_3) - \nabla v_h\big|_K\right) + \alpha_1^\top \left(\nabla v_h\big|_K - \nabla v_h(\tilde{x}_1)\right) \end{pmatrix}. \end{aligned} \tag{53}$$

Reasoning as in (44) we thus have

$$|\gamma - \Gamma_0 \nabla \widehat{v_h}| \leq Ch \left|\log h\right| |v|_{W^{2,\infty}(D)}, \tag{54}$$

which implies

$$\left| \nabla v_h\big|_K - \nabla \widetilde{\Pi}_h v_h\big|_{\widetilde{K}} \right| \leq h^{p+1} \left|\log h\right| |v|_{W^{2,\infty}(D)} \left|B_K^{-\top}\right| \sum_{j=0}^{\infty} h^p \left|\Gamma_{j-1}\right|. \tag{55}$$

From the definition of Γ_j, $j = -1, 0, \ldots, \infty$, we have $\sum_{j=0}^{\infty} h^p \left|\Gamma_{j-1}\right| \leq C$, which finally gives

$$\left| \nabla v_h\big|_K - \nabla \widetilde{\Pi}_h v_h\big|_{\widetilde{K}} \right| \leq Ch^p \left|\log h\right| |v|_{W^{2,\infty}(D)}. \tag{56}$$

Since $p \geq 1$, the triangular inequality yields $\varepsilon_i \leq Ch \left|\log h\right| |v|_{W^{2,\infty}(D)}$ for each node. Replacing this bound in (41), we get for each $x_i \in \mathcal{N}_h^I$

$$|e_h(x_i)| \leq Ch^{p+1} \alpha_i \left|\log h\right| |v|_{W^{2,\infty}(D)}. \tag{57}$$

Let us now remark that since $e_h \in V_h^+$ and by definition $e_h(\tilde{x}_i) = 0$ for all modified nodes, the maximum of e_h has to be realised on one of the nodes of the original mesh. Hence

$$\|e_h\|_{L^\infty(D)} = \max_{x_i \in \mathcal{N}_h^I} |e_h(x_i)| \leq Ch^{p+1} |\log h| \, |v|_{W^{2,\infty}(D)}, \tag{58}$$

which implies the desired result for all L^q spaces with $1 \leq q < \infty$. □

4 A probabilistic a posteriori error estimator

Several techniques exist for obtaining a posteriori error estimators in the framework of the FEM (see [17] for an overview), with the twofold goal of controlling the quality of numerical solutions and hence improve the meshing procedure to maximise efficiency. The main purpose of probabilistic numerical methods is to quantify the uncertainty introduced by approximate computations [11]. For the reasons above, we believe that deriving an error estimator from a family of numerical solutions fits perfectly in the probabilistic framework. In this section we present such a procedure for a probabilistic error estimation.

Assumption 3. Let $u_h^+ \in V_h^+$ be defined in (9). Then we assume there exists $0 \leq \beta < 1$ such that

$$\left\|u - u_h^+\right\|_a \leq \beta \left\|u - u_h\right\|_a, \tag{59}$$

where $\|u\|_a^2 = a(u,u)$. Moreover, there exists a constant $\gamma > 0$ such that

$$\left\|u_h - u_h^+\right\|_a \leq \gamma \left\|u_h - \tilde{u}_h\right\|_a, \tag{60}$$

almost surely, where $\tilde{u}_h$ is the probabilistic solution.

Let us remark that since $V_h \subset V_h^+$, we have $\beta \leq 1$ for the best approximation property of the Galerkin method and that Assumption 3 is often denoted in literature as the saturation assumption.

Lemma 6. *Let us denote by $w_h \in V_h$ and $\tilde{w}_h \in \widetilde{V}_h$ the components of u_h^+, i.e., $u_h^+ = w_h + \tilde{w}_h$. Then*

$$|w_h(x_i) - \tilde{w}_h(\tilde{x}_i)| \leq \tag{61}$$

and

$$\|w_h - \tilde{w}_h\|_a \leq \tag{62}$$

Proof. □

Lemma 7. *Let us denote by $z_h \in V_h$ the function $z_h = w_h - u_h/2$. Then*

$$\left\|z_h - \widetilde{\Pi}_h z_h\right\|_a \leq \dots \tag{63}$$

Proof.

$$\|z_h\|_a \leq \frac{1}{2} \|w_h - \tilde{w}_h\|_a. \tag{64}$$

□

Lemma 8. *Under ..., there exists $\gamma > 0$ independent of h and p such that*

$$\left\|u_h - u_h^+\right\|_a \leq \gamma \left\|u_h - \tilde{u}_h\right\|_a, \tag{65}$$

almost surely in Ω.

Proof. Let us write $u_h^+ = w_h + \tilde{w}_h$, where w_h and $\tilde{w}_h$ are the two components of u_h^+ in V_h and $\widetilde{V}_h$ respectively. For any $v_h^+ \in V_h^+$, $v_h^+ = v_h + \tilde{v}_h$, with $v_h \in V_h$ and $\tilde{v}_h \in \widetilde{V}_h$, by Galerkin orthogonality

$$\begin{aligned} a(u_h^+ - u_h, v_h^+) &= a(u_h^+ - u_h, \tilde{v}_h) - a(u_h^+ - \tilde{u}_h, \tilde{v}_h) \\ &= a(\tilde{u}_h - u_h, \tilde{v}_h). \end{aligned} \tag{66}$$

Choosing $v_h^+ = u_h^+ - u_h$, we have $\tilde{v}_h = \tilde{w}_h$ and

$$\left\|u_h^+ - u_h\right\|_a^2 = a(\tilde{u}_h - u_h, \tilde{w}_h). \tag{67}$$

The same procedure applied to $u_h^+ - \tilde{u}_h$ yields

$$\left\|u_h^+ - \tilde{u}_h\right\|_a^2 = a(u_h - \tilde{u}_h, w_h). \tag{68}$$

Hence

$$\left\|u_h^+ - u_h\right\|_a^2 + \left\|u_h^+ - \tilde{u}_h\right\|_a^2 = a(u_h - \tilde{u}_h, w_h - \tilde{w}_h). \tag{69}$$

Let us introduce the functions $z_h = w_h - u_h/2 \in V_h$ and $\tilde{z}_h = \tilde{w}_h - \tilde{u}_h/2 \in \widetilde{V}_h$. Then

$$\begin{aligned} \left\|u_h^+ - u_h\right\|_a^2 + \left\|u_h^+ - \tilde{u}_h\right\|_a^2 &= \frac{1}{2} a(u_h - \tilde{u}_h, u_h - \tilde{u}_h) + a(u_h - \tilde{u}_h, w_h - \frac{u_h}{2} - (\tilde{w}_h - \frac{\tilde{u}_h}{2})) \\ &= \frac{1}{2} \left\|u_h - \tilde{u}_h\right\|_a^2 + a(u_h - \tilde{u}_h, z_h - \tilde{z}_h). \end{aligned} \tag{70}$$

Consider now the second term in the sum. Adding and subtracting $a(u_h^+, z_h - \tilde{z}_h)$ and considering Galerkin orthogonality we obtain

$$a(u_h - \tilde{u}_h, z_h - \tilde{z}_h) = a(u_h - u_h^+, v_h - \tilde{z}_h) + a(u_h^+ - \tilde{u}_h, z_h - \tilde{v}_h), \tag{71}$$

for all $v_h \in V_h$ and $\tilde{v}_h \in \widetilde{V}_h$. Hence, applying Cauchy–Schwarz and Young's inequalities we obtain

$$\left\|u_h^+ - u_h\right\|_a^2 + \left\|u_h^+ - \tilde{u}_h\right\|_a^2 \leq \left\|u_h - \tilde{u}_h\right\|_a^2 + \inf_{v_h \in V_h} \left\|\tilde{z}_h - v_h\right\|_a^2 + \inf_{\tilde{v}_h \in \widetilde{V}_h} \left\|z_h - \tilde{v}_h\right\|_a^2. \tag{72}$$

□

Moreover, since the perturbed mesh and the original mesh could switch their roles by changing the sign to the random perturbations, the same assumption as (60) should be imposed for the probabilistic solution, i.e.

$$\left\|\tilde{u}_h - u_h^+\right\|_a \leq \widetilde{\gamma} \left\|u_h - \tilde{u}_h\right\|_a. \tag{73}$$

Applying the triangular inequality, we get

$$\begin{aligned} (\gamma + \widetilde{\gamma}) \left\|u_h - \tilde{u}_h\right\|_a &\geq \left\|\tilde{u}_h - u_h^+\right\|_a + \left\|u_h - u_h^+\right\|_a \\ &\geq \left\|u_h - \tilde{u}_h\right\|_a, \end{aligned} \tag{74}$$

which implies that $(\gamma + \widetilde{\gamma}) \geq 1$. The duality in the roles of deterministic and probabilistic meshes implies that γ and $\widetilde{\gamma}$ should be in general approximately equal, at least asymptotically. Hence, the lower bound above guarantees that neither γ nor $\widetilde{\gamma}$ should tend to zero with $h \to 0$.

It is known [2] that under Assumption 3 the estimate

$$\left\|u_h - u_h^+\right\|_a \leq \left\|u - u_h\right\|_a \leq \frac{1}{1-\beta} \left\|u_h - u_h^+\right\|_a, \tag{75}$$

holds almost surely. The quantity $\left\|u_h - u_h^+\right\|_a$ thus serves as an a posteriori error estimator for the error. However, computations involving the sum space V_h^+ are often intractable if the dimension $d > 1$. Hence, we further expand the upper bound thanks to (60) as

$$\left\|u - u_h\right\|_a \leq \frac{\gamma}{1-\beta} \left\|u_h - \tilde{u}_h\right\|_a, \tag{76}$$

which means that the difference between the deterministic and the probabilistic solutions can be employed as an a posteriori upper bound for the error.

Remark 4. Let us remark that the value of β is influenced by the choice of p in Assumption 1. Let us consider the limit case of $p \to \infty$. In this case, the spaces V_h and $\tilde{V}_h$ coincide, and in turn coincide both with V_h^+. Hence, the space V_h^+ is in the limit not wider than V_h and one expects $\beta \to 1$. We hence postulate that $\beta = \beta(h, p)$ takes the form

$$\beta(h, p) = 1 - \beta_1 h^{\beta_2(p-1)}, \tag{77}$$

for some $0 < \beta_1 \leq 1$ and $\beta_2 > 0$. This is motivated by the fact that the two terms in (76) converge with the same rate $\mathcal{O}(h)$ in case $p = 1$ due to a priori error results. Hence, in this case, $\beta(h, 1)$ is independent of h and equals a constant value β. Conversely, if $p > 1$, one gets on the right hand side a term of order $\mathcal{O}(h^{\beta_2(1-p)} h^{(p+1)/2})$, bounding a term of order $\mathcal{O}(h)$ on the left hand side. Hence, we impose

$$\beta_2(1 - p) + \frac{p + 1}{2} \leq 1, \tag{78}$$

which, since $p > 1$, gives $\beta_2 \geq 1/2$. Numerical experiments confirm the qualitative behaviour of the function $\beta(h, p)$ explained above, and lead to the good working practice of fixing $p = 1$.

A more robust estimator could be obtained by averaging a family of M probabilistic solutions $\tilde{u}_h^{(i)}$, $i = 1, \ldots, M$, obtained by M i.i.d. random perturbations of the original mesh. In particular, we have

$$\|u - u_h\|_a^2 \leq C\, \mathbb{E}\, \|u_h - \tilde{u}_h\|_a^2 =: C\eta_h^2, \tag{79}$$

where we approximate the estimator η_h via Monte Carlo sampling as

$$\eta_h \approx \sqrt{\frac{1}{M} \sum_{i=1}^{M} \left\| u_h - \tilde{u}_h^{(i)} \right\|_a^2}. \tag{80}$$

Taking the expectation over several realisations should in practice provide a sharper error estimator, as in case $p = 1$ a good portion of the domain D is explored by the vertices of several realisations of the random mesh. Let us consider for simplicity the case $\kappa \equiv 1$, so that $\|u\|_a = \|\nabla u\|_{L^2(D)}$ for all $u \in H_0^1(D)$. In this case, we have

$$\begin{aligned} \eta_h &= \int_K \mathbb{E}\, |\nabla(u_h - \tilde{u}_h)|^2 \,\mathrm{d}x \\ &\approx \int_K \mathbb{E}\, |\mathbb{E}(\nabla u_h) - \nabla \tilde{u}_h|^2 \,\mathrm{d}x \\ &= \int_K \operatorname{tr}(\operatorname{Var} \nabla u_h) \mathrm{d}x. \end{aligned} \tag{81}$$

Hence, following the probabilistic numerics canon, it is possible to interpret the error estimator as an integral measure of the statistical dispersion of numerical solutions over the domain.

We now consider the task of adapting the mesh. Given the error estimator derived above and a prescribed tolerance, we apply a standard technique for generating a sequence of meshes, which we briefly summarise in the following. Let us first split the estimator over the elements of the original mesh as

$$\begin{aligned} \eta_h^2 &= \sum_{K \in \mathcal{T}_h} \mathbb{E} \int_K \kappa \nabla(u_h - \tilde{u}_h) \cdot \nabla(u_h - \tilde{u}_h) \\ &= \sum_{K \in \mathcal{T}_h} \eta_K^2, \end{aligned} \tag{82}$$

where we consider η_K to be an indicator of the error at a local level. If we impose a tolerance level ϵ for the error, i.e.,

$$\|u - u_h\|_a \leq \epsilon \, \|u_h\|_a \,, \tag{83}$$

we obtain that a sufficient condition is given by

$$\eta_K \leq \frac{\epsilon \, \|u_h\|_a}{\sqrt{N}}, \tag{84}$$

Algorithm 1: Probabilistic mesh adaptivity.

Data: $\mathcal{T}_h^{(0)}$, tolerance ε, safety factors fac_1, fac_2, $M \in \mathbb{N}$.

Set $i = 0$;
while $\eta_h > \epsilon \left\| u_h \right\|_a$ **do**
  Compute u_h and $\left\| u_h \right\|_a$;
  Draw M random meshes and compute $\tilde{u}_h^{(j)}$ for $j = 1, \dots, M$;
  for $K \in \mathcal{T}_h^{(i)}$ **do**
    Compute η_K ;
    if $\eta_K > \mathrm{fac}_1\, \epsilon \left\| u_h \right\|_a / \sqrt{N}$ **then**
      Mark element K for refinement ;
    else if $\eta_K < \mathrm{fac}_2\, \epsilon \left\| u_h \right\|_a / \sqrt{N}$ **then**
      Mark element K for coarsening ;
  Build $\mathcal{T}^{(i+1)}$;
  Set $i \leftarrow i + 1$;

where N is the number of elements in $\mathcal{T}_h$. Hence, we proceed iteratively by refining the mesh around elements which do not fulfil the error requirement until the required tolerance is attained. Coarsening of elements where the error indicator is small could be as well employed for saving computational power. The algorithm for mesh adaptation is given in Algorithm 1, where safety factors fac_1 and fac_2 are introduced.

5 Inverse problems

Probabilistic numerical methods are particularly helpful when inserted in the framework of Bayesian inverse problems (BIPs) involving differential equations, as studied in [1,7] for ODEs, and in [3,5] for PDEs. Furthermore, in [14] a theoretical basis is laid for ensuring the well-posedness of probabilistic solutions to BIPs.

We consider the framework introduced in [15] and expanded in [8]. With the notation of (1), we consider the PDE

$$\begin{aligned} -\nabla \cdot (\exp(\vartheta) \nabla u) &= f, \quad \text{in } D, \\ u &= 0, \quad \text{on } \partial D, \end{aligned} \tag{85}$$

where the conductivity field κ is transformed through an exponential function $\kappa = \exp(\vartheta)$ in order to ensure positivity and hence well-posedness of the solution. Moreover, we suppose that $u \in W^{2,\infty}(D)$ and we let $\mathcal{U} = \text{addspace}$ be the space of admissible log-conductivity fields ϑ. The BIP consists in retrieving the true value $\vartheta^\dagger$ of the field ϑ given prior information and corrupted observations $z \in \mathbb{R}^m$ given by

$$z = \mathcal{G}(\vartheta^\dagger) + \varepsilon, \tag{86}$$

where we assume that $\varepsilon \sim \mathcal{N}(0, \Sigma_\varepsilon)$ is a Gaussian source of additive noise and $\mathcal{G} \colon \mathcal{U} \to \mathbb{R}^m$ is the forward operator. In particular, we can write $\mathcal{G} = \mathcal{O} \circ \mathcal{S}$, where $\mathcal{S} \colon \mathcal{U} \to W^{2,\infty}(D)$ is the solution operator, mapping any value of the field ϑ to the solution u of (85), and $\mathcal{O} \colon W^{2,\infty}(D) \to \mathbb{R}^m$ is the observation operator. In this work, we simply consider $\mathcal{O}$ to be defined by point-wise evaluations of the solution, i.e.,

$$\mathcal{O} \colon \vartheta \mapsto \begin{pmatrix} u(x_1) & u(x_2) & \dots & u(x_m) \end{pmatrix}^\top. \tag{87}$$

If the prior information is encoded by a prior measure μ_0 over the space $\mathcal{U}$, then the solution of the BIP is given by the posterior distribution μ such that its Radon–Nikodym derivative satisfies

$$\frac{\mathrm{d}\mu}{\mathrm{d}\mu_0}(\vartheta; z) = \frac{1}{Z} \exp(-\Phi(\vartheta; z)), \tag{88}$$

where $\Phi\colon (L^\infty)^d \times \mathbb{R}^m \to \mathbb{R}$ is referred to as the potential function and Z is a normalisation constant. Under the Gaussian assumption for the noise, we have

$$\Phi(\vartheta; z) = \frac{1}{2} \left\| z - \mathcal{G}(\vartheta) \right\|_{\Sigma_\varepsilon}^2, \tag{89}$$

where the norm $\|\cdot\|_{\Sigma_\varepsilon}$ is defined as

$$\|y\|_{\Sigma_\varepsilon} = \left\| \Sigma_\varepsilon^{-1/2} y \right\|_{\mathbb{R}^m}. \tag{90}$$

In the following, we will consider the observations z to be fixed and hence denote $\mu(\mathrm{d}\vartheta) = \mu(\mathrm{d}\vartheta; z)$ as well as $\Phi(\vartheta) = \Phi(\vartheta; z)$. Let us denote by $\mathcal{G}_h \colon \mathcal{U} \to \mathbb{R}^m$ the forward model obtained as $\mathcal{G}_h = \mathcal{O} \circ \mathcal{S}_h$, where $S_h \colon \mathcal{U} \to V_h$ is the solution operator given by the linear FEM and we still denote by $\mathcal{O}$ the restriction of $\mathcal{O}$ to V_h. Denoting by Φ_h the approximate potential, given by

$$\Phi_h(\vartheta) = \frac{1}{2} \left\| z - \mathcal{G}_h(\vartheta) \right\|_{\Sigma_\varepsilon}^2, \tag{91}$$

we obtain the approximate posterior measure μ_h as

$$\frac{\mathrm{d}\mu_h}{\mathrm{d}\mu_0}(\vartheta) = \frac{1}{Z_h} \exp(-\Phi_h(\vartheta)), \tag{92}$$

where Z_h is the normalisation constant. Stuart proved [15, Theorem 4.6] that under suitable assumptions $d_{\mathrm{H}}(\mu_h, \mu) \to 0$ for $h \to 0$, where $d_{\mathrm{H}}(\cdot, \cdot)$ is the Hellinger distance for probability measures. Hence, assuming an infinite computational budget is available it is possible to compute the posterior measure via approximate computations. This result has then been extended to more general priors than Gaussian [9, 16].

It has been shown empirically that under a fixed computational budget, employing a standard numerical method for the approximation of the solution operator $\mathcal{S}$ can lead to inaccurate results [1, 5, 7]. In particular, in case the variance Σ_ε of the observational noise is small with respect to the discretisation error, the posterior measure μ_h will be overconfident and peaked away from the true value of the unknown field. Probabilistic numerical methods can efficiently tackle this overconfidence issue thanks to the uncertainty quantification of numerical errors they naturally introduce. Given the probability space Ω on which the random variables defining the probabilistic scheme $\alpha_i \colon \Omega \to \mathbb{R}^d$ introduced in Assumption 1 are defined, let us denote by $\widetilde{\mathcal{G}}_h \colon \Omega \times \mathcal{U} \to \mathbb{R}^m$ the random forward model obtained as $\widetilde{\mathcal{G}}_h = \mathcal{O} \circ \widetilde{\mathcal{S}}_h$, where $\widetilde{\mathcal{S}}_h \colon \Omega \times \mathcal{U} \to \widetilde{V}_h$ is the solution operator corresponding to the random FEM introduced in this work. Replacing $\mathcal{G}_h$ with $\widetilde{\mathcal{G}}_h$ in (91) we get a random potential $\widetilde{\Phi}_h$ and eventually a random posterior measure $\tilde{\mu}_h$ defined by

$$\frac{\mathrm{d}\tilde{\mu}_h}{\mathrm{d}\mu_0}(\vartheta) = \frac{1}{\widetilde{Z}_h} \exp(-\widetilde{\Phi}_h(\vartheta)), \tag{93}$$

where $\widetilde{Z}_h$ is the normalisation constant. In order to obtain an approximation of μ through $\tilde{\mu}_h$, we need to take the expectation of the random probabilistic solution, which is viable in two different manners as explained in [14]. The first approach is to define the measure $\tilde{\mu}_h^{\mathrm{fix}} = \mathbb{E}\, \tilde{\mu}_h$. Otherwise, one could define a measure $\tilde{\mu}_h^{\mathrm{var}}$ through

$$\frac{\mathrm{d}\tilde{\mu}_h^{\mathrm{var}}}{\mathrm{d}\mu_0}(\vartheta) = \frac{1}{\mathbb{E}\, \widetilde{Z}_h} \mathbb{E} \exp(-\widetilde{\Phi}_h(\vartheta)), \tag{94}$$

which is already a deterministic measure. The choice of the names of these two approximation comes from their computation, which is in spirit slightly different. In the case of $\tilde{\mu}_h^{\mathrm{fix}}$, for each event ω one evaluates the forward model and computes the value of the posterior. The expectation is then taken with the respect to the posterior itself, in practice via averaging techniques. Hence, for each ω we fix a perturbed mesh $\widetilde{\mathcal{T}}_h(\omega)$ and compute the posterior for several values of ϑ. Conversely, in the case of the measure $\tilde{\mu}_h^{\mathrm{var}}$ the field ϑ is first fixed, and then the posterior is in practice obtained evaluating the forward model on several (variable) realisations of the random probabilistic solution.

We now need to prove the convergence of the posterior distributions $\tilde{\mu}_h$ and $\tilde{\mu}_h^{\mathrm{var}}$ towards the true posterior μ with respect to the mesh size, which is granted by the following result under three regularity assumptions.

Theorem 3 (Theorem 3.9 of [14]). *With the notation above, if*

(i) there exists $q > 0$ *such that* $\exp(\Phi) \in L^q_{\mu_0}(\mathcal{U})$,

(ii) there exists a constant $C > 0$ *such that*

$$\mathbb{E}_{\mu_0}[\widetilde{\Phi}_N] \leq C, \quad \textit{almost surely in } \Omega, \tag{95}$$

(iii) it holds

$$\lim_{h \to 0} \left\| \left(\mathbb{E} \left\| \widetilde{\mathcal{G}}_h - \mathcal{G} \right\|^2 \right)^{1/2} \right\|_{L^s_{\mu_0}(\mathcal{U})} = 0, \tag{96}$$

where $s = 2q/(q-1)$ *and* q *is given in (i),*

then

$$\begin{aligned} \mathbb{E}\left[d_{\mathrm{H}}(\mu, \widetilde{\mu}_h)^2\right]^{1/2} &\leq C \left\| \left(\mathbb{E} \left\| \widetilde{\mathcal{G}}_h - \mathcal{G} \right\|^4_{\mathbb{R}^m} \right)^{1/2} \right\|^{1/2}_{L^2_{\mu_0}(\mathcal{U})}, \\ d_{\mathrm{H}}(\mu, \widetilde{\mu}^{\mathrm{var}}_h) &\leq C \left\| \mathbb{E} \left\| \widetilde{\mathcal{G}}_h - \mathcal{G} \right\|^2_{\mathbb{R}^m} \right\|^{1/2}_{L^s_{\mu_0}(\mathcal{U})}. \end{aligned} \tag{97}$$

Let us remark that for a measure μ the spaces $L^q_\mu(\mathcal{U})$ are defined as

$$L^q_\mu(\mathcal{U}) = \left\{ f \colon \mathcal{U} \to \mathbb{R} : \int_{\mathcal{U}} f(\vartheta)^q \, \mu(\mathrm{d}\vartheta) < \infty \right\}, \tag{98}$$

with norm

$$\|f\|_{L^q_\mu(\mathcal{U})} = \left(\int_{\mathcal{U}} f(\vartheta)^q \, \mu(\mathrm{d}\vartheta) \right)^{1/q}. \tag{99}$$

Theorem 3 gives in a general framework the convergence of posterior measures defined through approximate random forward models The following result now guarantees the convergence of the posterior distributions

6 Numerical experiments

6.1 Convergence

One-dimensional case

We consider (1) with $\kappa \equiv 1$ on $D = (0,1)$ and $f(x) = (x - 1/2)\chi_{(1/2,1)}(x)$, so that the solution u satisfies Assumption 2. We verify the result of Theorem 2 by choosing $p \in \{1,2,3\}$ and by varying the mesh size h in the range $[9 \cdot 10^{-3}, 0.25]$. Moreover, we compute only one realisation of the random mesh for each couple $\{p, h\}$ as our bound holds almost surely. Results, shown in Fig. 2, confirm the validity of the convergence estimates.

6.2 Error estimators

One-dimensional example

Consider

$$\begin{aligned} -u'' &= f, \quad \text{in } (0,1), \\ u(0) &= u(1) = 0, \end{aligned} \tag{100}$$

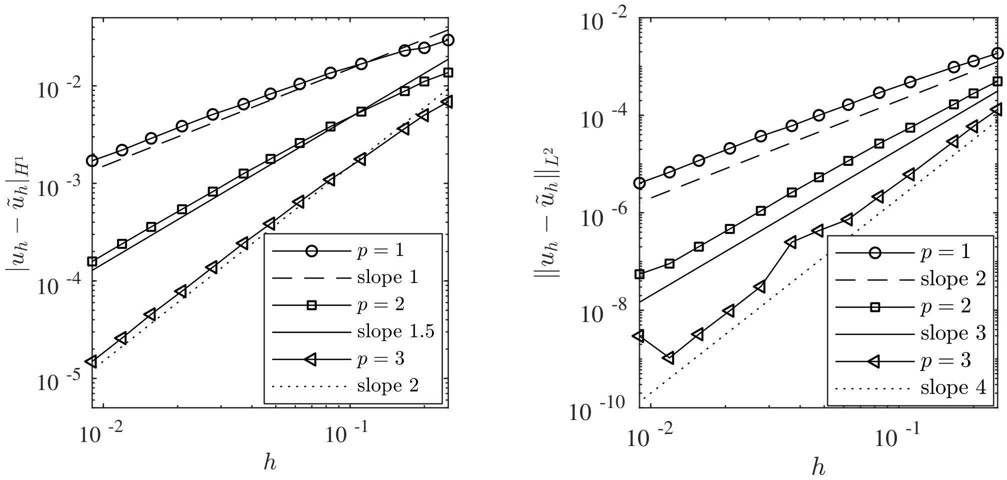

Figure 2: Convergence rates in the H^1 semi-norm and the L^2 norm for the one-dimensional Poisson equation.

with f chosen such that $u(x) = -\sin(12\pi x)\exp(-100(x-1/2)^2)$ is the true solution. We consider the error estimations of presented in Section 4, both in a local and global manner. Results, displayed in Figs. 3 and 4 show that the estimates hold in practice for this case. In particular, in Fig. 4 we can remark that the overall effectivity index $\eta_{\mathcal{X}}$, defined as

$$\eta_{\mathcal{X}} = \frac{\mathbb{E}\, \|u_h - \tilde{u}_h\|_{\mathcal{X}}}{\|u_h - u\|_{\mathcal{X}}}, \tag{101}$$

with $\mathcal{X} = H_0^1, L^2$, is in this case close to one for both norms. Errors are estimated employing $M = 10$ realisations of the probabilistic solution and with a Monte Carlo simulation.

Two-dimensional case

TO DO

6.3 Mesh adaptivity

Two-dimensional case

See results Fig. 5.

6.4 Bayesian inverse problems

Let us consider the following one-dimensional elliptic equation

$$\begin{aligned} -\frac{\mathrm{d}}{\mathrm{d}x}\big(e^{\kappa}\frac{\mathrm{d}u}{\mathrm{d}x}\big) &= f, \quad \text{in } (0,1), \\ u &= 0, \quad \text{on } \{0,1\}, \end{aligned} \tag{102}$$

and the inverse problem of retrieving the field $\kappa \in L^2(0,1)$ given synthetic noisy observations of the solution u corresponding to a true field κ^*. First, we consider a case where information on κ is

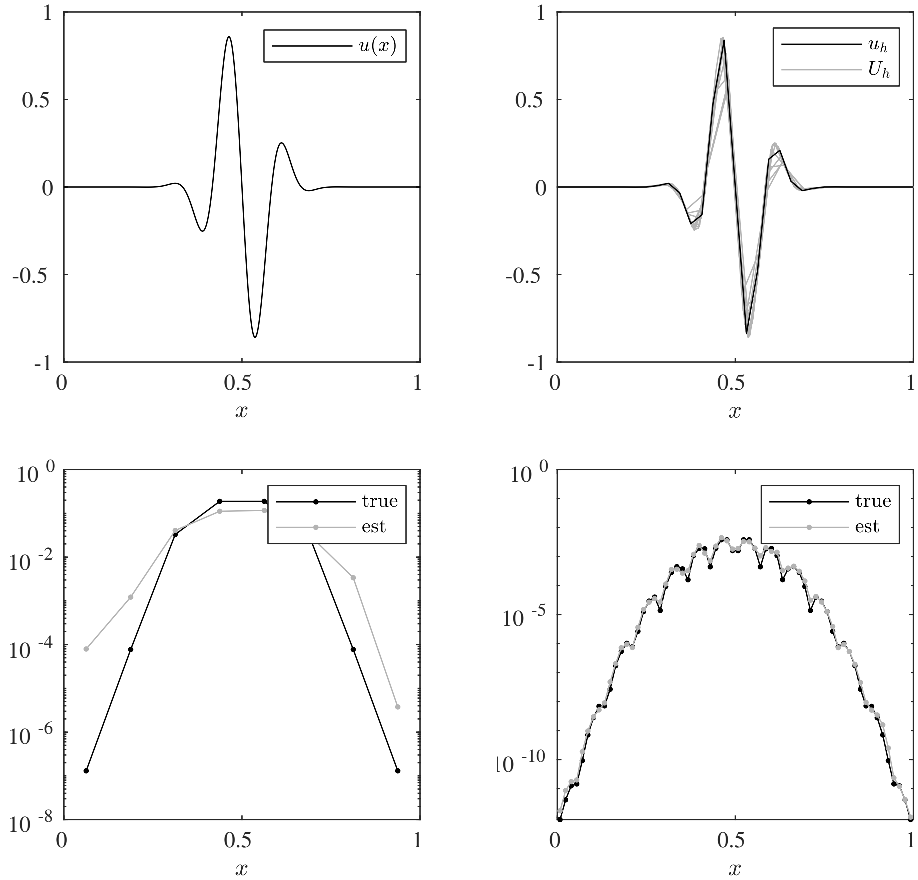

Figure 3: Error estimation for the 1D problem with two different values of h – error in each element.

available beforehand. In particular, we assume that κ has the form

$$\kappa(x) = \begin{cases} \log(1+\kappa_1), & \text{if } x \in I_1, \\ \log(1+\kappa_2), & \text{if } x \in I_2, \\ 0 & \text{otherwise,} \end{cases} \tag{103}$$

where κ_1, κ_2 are real scalars and I_1, I_2 are the intervals $(0.2, 0.4)$ and $(0.6, 0.8)$ respectively. Fixing a standard Gaussian prior on both parameters κ_1 and κ_2 we are able to compute the posterior distribution corresponding to both the deterministic and probabilistic forward models. In particular, we vary the number of elements N in the set $\{20, 40, 80, 160\}$, thus studying the effects of numerical errors on the numerical posterior distribution. Observations are obtained from a reference solution evaluated at four equispaced points in the interior of $(0, 1)$ each corrupted by an additive source of noise $\varepsilon \sim \mathcal{N}(0, 10^{-4})$. The posterior distributions are obtained with Metropolis–Hastings initialised near the true value of (κ_1, κ_2) and ran as explained in Section 5, with 200 parallel chains employed for the probabilistic forward model. Results are shown in Fig. 7, where TO DO

In a second experiment, we consider the same exact field κ^* and observation model, but without the additional information encoded in (103).

References

[1] A. Abdulle and G. Garegnani, *Random time step probabilistic methods for uncertainty*

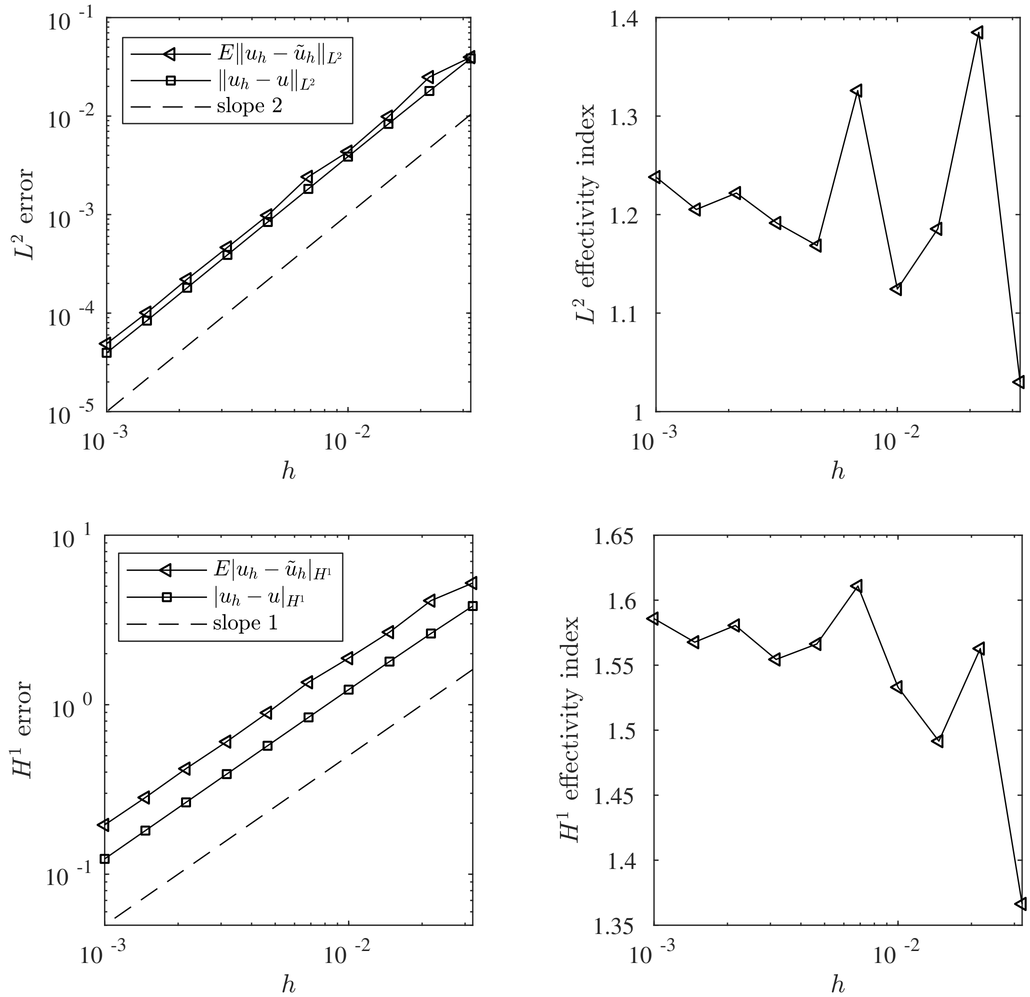

Figure 4: Error estimation for the 1D problem with two different values of h – convergence of error estimators and effectivity indices

quantification in chaotic and geometric numerical integration. arXiv preprint arXiv:1801.01340, 2018.

[2] R. E. Bank and R. K. Smith, *A posteriori error estimates based on hierarchical bases*, SIAM J. Numer. Anal., 30 (1993), pp. 921–935, https://doi.org/10.1137/0730048, https://doi.org/10.1137/0730048.

[3] O. A. Chkrebtii, D. A. Campbell, B. Calderhead, and M. A. Girolami, *Bayesian solution uncertainty quantification for differential equations*, Bayesian Anal., 11 (2016), pp. 1239–1267, https://doi.org/10.1214/16-BA1017.

[4] P. G. Ciarlet, *The finite element method for elliptic problems.*, vol. 40 of Classics Appl. Math., SIAM, Philadelphia, 2002.

[5] J. Cockayne, C. Oates, T. Sullivan, and M. Girolami, *Probabilistic numerical methods for PDE-constrained Bayesian inverse problems*, AIP Conference Proceedings, 1853 (2017), p. 060001.

[6] P. R. Conrad, M. Girolami, S. Särkkä, A. Stuart, and K. Zygalakis, *Statistical analysis of differential equations: introducing probability measures on numerical solutions*, Stat. Comput., (2016).

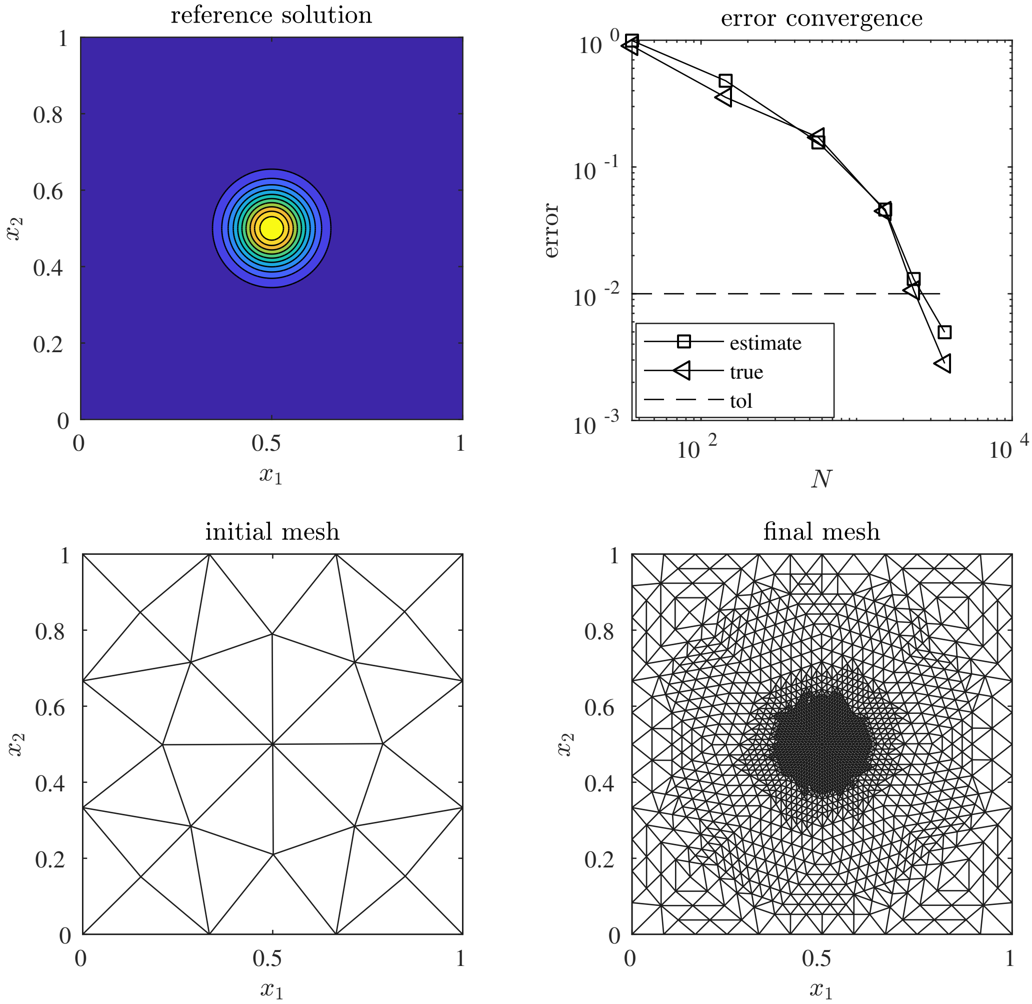

Figure 5: Mesh adaptivity – two-dimensional case

[7] P. R. Conrad, M. Girolami, S. Särkkä, A. Stuart, and K. Zygalakis, *Statistical analysis of differential equations: introducing probability measures on numerical solutions*, Stat. Comput., 27 (2017), pp. 1065–1082, https://doi.org/10.1007/s11222-016-9671-0, https://doi.org/10.1007/s11222-016-9671-0.

[8] M. Dashti and A. M. Stuart, *Uncertainty quantification and weak approximation of an elliptic inverse problem*, SIAM J. Numer. Anal., 49 (2011), pp. 2524–2542, https://doi.org/10.1137/100814664, https://doi.org/10.1137/100814664.

[9] M. Dashti and A. M. Stuart, *The Bayesian Approach to Inverse Problems*, in Handbook of Uncertainty Quantification, Springer, 2016, pp. 1–118.

[10] E. S. Gawlik and A. J. Lew, *Supercloseness of orthogonal projections onto nearby finite element spaces*, ESAIM Math. Model. Numer. Anal., 49 (2015), pp. 559–576, https://doi.org/10.1051/m2an/2014045, https://doi.org/10.1051/m2an/2014045.

[11] P. Hennig, M. A. Osborne, and M. Girolami, *Probabilistic numerics and uncertainty in computations*, Proc. A., 471 (2015), pp. 20150142, 17, https://doi.org/10.1098/rspa.2015.0142, https://doi.org/10.1098/rspa.2015.0142.

[12] H. Kersting and P. Hennig, *Active uncertainty calibration in Bayesian ODE solvers*, in Proceedings of the 32nd Conference on Uncertainty in Artificial Intelligence (UAI 2016), AUAI Press, 2016, pp. 309–318, http://www.auai.org/uai2016/proceedings/papers/163.pdf.

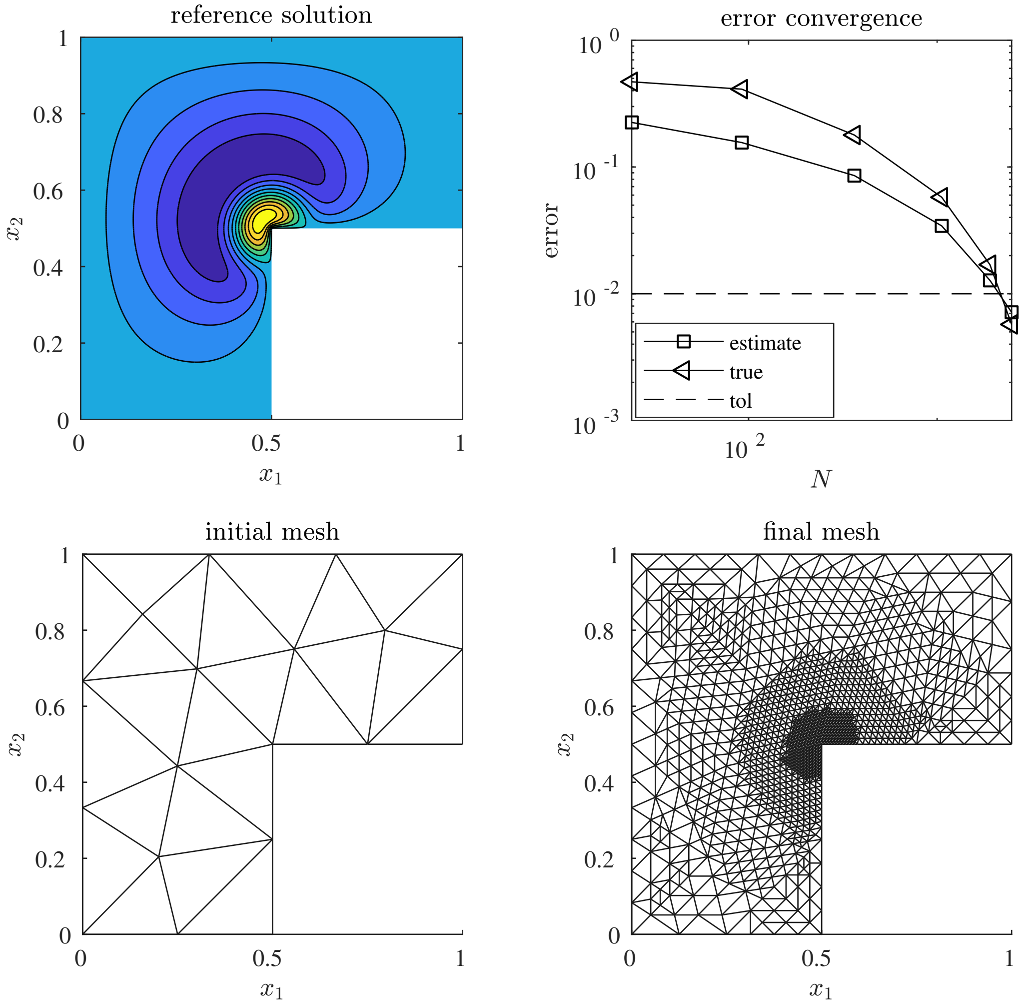

Figure 6: Mesh adaptivity – two-dimensional case

[13] H. Kersting, T. J. Sullivan, and P. Hennig, *Convergence rates of Gaussian ODE filters*. arXiv preprint arXiv:1807.09737, 2018.

[14] H. C. Lie, T. J. Sullivan, and A. L. Teckentrup, *Random Forward Models and Log-Likelihoods in Bayesian Inverse Problems*, SIAM/ASA J. Uncertain. Quantif., 6 (2018), pp. 1600–1629, https://doi.org/10.1137/18M1166523, https://doi.org/10.1137/18M1166523.

[15] A. M. Stuart, *Inverse problems: a Bayesian perspective*, Acta Numer., 19 (2010), pp. 451–559, https://doi.org/10.1017/S0962492910000061, https://doi.org/10.1017/S0962492910000061.

[16] T. J. Sullivan, *Well-posed Bayesian inverse problems and heavy-tailed stable quasi-Banach space priors*, Inverse Probl. Imaging, 11 (2017), pp. 857–874, https://doi.org/10.3934/ipi.2017040, https://doi.org/10.3934/ipi.2017040.

[17] R. Verfürth, *A posteriori error estimation techniques for finite element methods*, Numerical Mathematics and Scientific Computation, Oxford University Press, Oxford, 2013, https://doi.org/10.1093/acprof:oso/9780199679423.001.0001, http://dx.doi.org/10.1093/acprof:oso/9780199679423.001.0001.

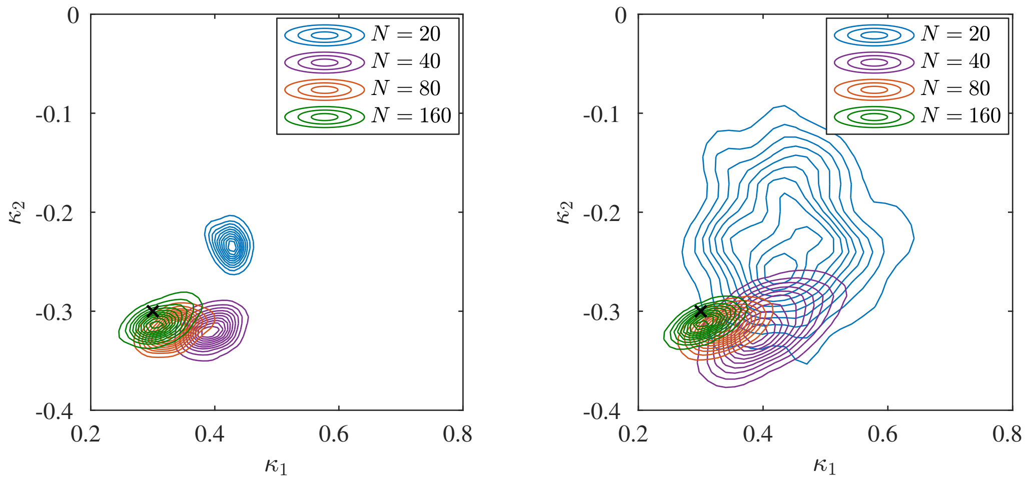

Figure 7: Bayesian inverse problem – finite dimensional case.