A multiscale method for semi-linear elliptic equations with localized uncertainties and non-linearities

Anthony Nouy, Florent Pled

Mathematical Modelling and Numerical Analysis Will be set by the publisher
Modélisation Mathématique et Analyse Numérique

A MULTISCALE METHOD FOR SEMI-LINEAR ELLIPTIC EQUATIONS WITH LOCALIZED UNCERTAINTIES AND NON-LINEARITIES *

Anthony Nouy[1] and Florent Pled[2]

Abstract. A multiscale numerical method is proposed for the solution of semi-linear elliptic stochastic partial differential equations with localized uncertainties and non-linearities, the uncertainties being modeled by a set of random parameters. It relies on a domain decomposition method which introduces several subdomains of interest (called patches) containing the different sources of uncertainties and non-linearities. An iterative algorithm is then introduced, which requires the solution of a sequence of linear global problems (with deterministic operators and uncertain right-hand sides), and non-linear local problems (with uncertain operators and/or right-hand sides) over the patches. Non-linear local problems are solved using an adaptive sampling-based least-squares method for the construction of sparse polynomial approximations of local solutions as functions of the random parameters. Consistency, convergence and robustness of the algorithm are proved under general assumptions on the semi-linear elliptic operator. A convergence acceleration technique (Aitken's dynamic relaxation) is also introduced to speed up the convergence of the algorithm. The performances of the proposed method are illustrated through numerical experiments carried out on a stationary non-linear diffusion-reaction problem.

1991 Mathematics Subject Classification. 35R60, 60H15, 65N30, 65N55, 65D15.

May 8, 2018.

Uncertainty quantification has become a topical issue in computational sciences and engineering. Numerous methods have been proposed to propagate uncertainties through models governed by partial differential equations (see e.g. [1–3]). While these methods have reached a certain degree of maturity and become nowadays widespread, a major concern has emerged for multiscale models where uncertainties occur at various scales.

Several numerical methods dedicated to deterministic multiscale models have been extended to the stochastic framework. For multiscale problems with global sources of uncertainties, spectral stochastic methods have been combined with deterministic multiscale methods, *e.g.* the multiscale finite element method (FEM) [4], the variational multiscale method [5] or the heterogeneous multiscale method [6], leading to the so-called multiscale stochastic FEM [7] and its variants [8–15]. These methods are well adapted to global uncertainties and are proved to be efficient when assuming small random fluctuations and scale separation. Meanwhile,

Keywords and phrases: Uncertainty quantification, Non-linear elliptic stochastic partial differential equation, Multiscale, Domain decomposition, Sparse approximation

* *The financial support of the French National Research Agency - Agence Nationale de la Recherche (ANR) - under Grant ICARE ANR-12-MONU-0002 is acknowledged by the authors.*

[1] Centrale Nantes, LMJL UMR CNRS 6629, 1 rue de la Noë, BP 92101, 44321 Nantes Cedex 3, France ; e-mail: anthony.nouy@ec-nantes.fr

[2] Université Paris-Est, Laboratoire Modélisation et Simulation Multi Echelle, MSME UMR 8208 CNRS, 5 bd Descartes, 77454 Marne-la-Vallée, France ; e-mail: florent.pled@univ-paris-est.fr

traditional substructuring and domain decomposition methods have been introduced for stochastic monoscale models [16–18] and recently extended to multiscale models [19, 20] in order to benefit from scalable parallel algorithms available in the deterministic framework. These methods are also well adapted to problems where the uncertainties are scattered in the whole domain.

The present work focuses on non-linear stochastic multiscale models where localized sources of uncertainties and non-linearities may occur in some regions of interest. Concurrent approaches, initially developed in the deterministic framework, have been proposed to couple numerical models. First of all, mono-model approaches currently rely on adaptive remeshing strategies [21,22] or enrichment techniques (*e.g.* the eXtended FEM [23,24] or the Generalized FEM [25]) and generally require high computational resources or specific (intrusive) implementations. Conversely, multi-model approaches based on patches have a high potential to manage complex multiscale problems by operating a separation of scales. The separation of scales allows to capture the local features of multiscale solutions at a micro scale (local level) while keeping a simplified global description at a macro scale (global level). Several multiscale coupling methods have been developed within the deterministic framework and some have been extended to the stochastic framework. They distinguish themselves by the way of defining and constructing the coupling operator between global and local models. First, superposition methods, such as the method of finite element patches [26,27] and the method of harmonic patches [28], consist in adding a fine local correction to a coarse global solution. Second, surface coupling methods include the Chimera-Schwarz method [29,30], the Semi-Schwarz method [31], the Semi-Schwarz-Lagrange method [32–35] and the local multigrid method [36–39]. Both multiscale superposition and surface coupling methods are based on global-local iterative algorithms originally developed for domain decomposition methods or multigrid methods. Nevertheless, the former can be interpreted as a local model refinement technique, while the latter can be seen as a local model substitution technique. Third, volume coupling methods, such as the Arlequin method [40], belong to the class of overlapping domain decomposition methods and require the definition of a coupling zone between the different models. Among all these multi-model approaches, few have been explored in the stochastic framework. The Arlequin (volume coupling) method has been applied to deterministic-stochastic coupling in [41, 42] for homogenization purposes. Besides, the Semi-Schwarz-Lagrange (surface coupling) method has been recently extended to linear stochastic multiscale models with localized sources of uncertainties in [43].

This work extends [43] to a class of non-linear stochastic multiscale models. Alternative multiscale approaches have been recently proposed to handle non-linear elliptic problems. For semi-linear elliptic equations, we refer to the variational multiscale method proposed in [44], while for other non-linear elliptic equations, we refer to the multiscale FEM presented in [45,46], the variational multiscale method developed in [47] or the heterogeneous multiscale method proposed in [48]. In the present work, a dedicated multiscale method based on a domain decomposition is proposed to exploit the localized side of uncertainties and non-linearities. It relies on a global-local iterative algorithm which requires the solution of a sequence of linear global problems (with deterministic operators and uncertain right-hand sides) at a macro scale and non-linear local problems (with uncertain operators and right-hand sides) at a micro scale (over patches). Appropriate approximation spaces and solvers can be considered to solve both types of problems efficiently. This multiscale approach then appears to be flexible and non-intrusive in the sense that it requires no modification of both global and local models and solvers, which makes possible the use of stand-alone codes. The main motivation is the deployment and transfer of methods towards complex large-scale industrial applications [49]. Besides, different types of uncertainties can be considered in the non-linear local models. They may be associated with some variabilities of the operator but also of the geometry, the source terms or the boundary conditions.

The remainder of the paper is structured as follows. Section 1 introduces the initial formulation of the semi-linear elliptic stochastic partial differential equation with localized uncertainties and non-linearities and states suitable assumptions. Section 2 presents the global-local (two-scale) formulation with patches containing localized variabilities and non-linearities. A global-local iterative algorithm is then introduced and analyzed in Section 3. Consistency, convergence and robustness properties are deduced from the assumptions introduced in Section 1. Subsequently, the computational aspects associated with the solution of both global and local problems are detailed in Section 4. In particular, the stochastic local problems are solved using sampling-based

(non-intrusive) approaches and working sets algorithms proposed in [50] for the adaptive construction of sparse polynomial approximations of local solutions. Finally, the proposed method is illustrated through numerical examples in Section 5.

1. Problem statement

Let ξ denote a set of real-valued random variables modeling the different sources of uncertainties (on the operator, geometry, source terms and boundary conditions). We assume that ξ takes values in a set Ξ and we let μ be the probability law of ξ. We consider the following semi-linear second-order stochastic partial differential equation

$$-\nabla \cdot B(u, \nabla u; x, \xi) + C(u, \nabla u; x, \xi) = f(x, \xi) \quad \text{for } x \in \Omega(\xi), \tag{1a}$$

where $\Omega(\xi)$ is an uncertain domain of $\mathbb{R}^d$ with sufficiently smooth (*e.g.* Lipschitz) boundary $\partial\Omega(\xi)$. Here $B(\cdot,\cdot;x,\xi)\colon \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^d$ and $C(\cdot,\cdot;x,\xi)\colon \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}$, and $f(\cdot,\xi)\colon \Omega(\xi) \to \mathbb{R}$ is a given source term. For a given value of ξ, the solution $u(\cdot,\xi)$ is a function from $\Omega(\xi)$ to $\mathbb{R}$. We supply equation (1a) with the following Dirichlet and Neumann boundary conditions

$$u = 0 \qquad \text{on } \Gamma_D(\xi), \tag{1b}$$

$$B(u, \nabla u; x, \xi) \cdot n = g(x, \xi) \quad \text{on } \Gamma_N(\xi), \tag{1c}$$

where $\Gamma_D(\xi)$ and $\Gamma_N(\xi)$ are disjoint and complementary parts of $\partial\Omega(\xi)$ such that $\overline{\Gamma_D(\xi) \cup \Gamma_N(\xi)} = \partial\Omega(\xi)$ and $\Gamma_D(\xi) \cap \Gamma_N(\xi) = \emptyset$, and $\mathrm{meas}(\Gamma_D(\xi)) \neq 0$. $g(\cdot,\xi)\colon \Gamma_N(\xi) \to \mathbb{R}$ is a prescribed normal flux on $\Gamma_N(\xi)$, and n is the unit outward normal to $\Gamma_N(\xi)$.

Example 1.1 (Non-linear diffusion-reaction equation)**.** As a model example, we consider a non-linear diffusion-reaction equation (1a) in dimension $d \leqslant 3$, with

$$B(u, \nabla u; x, \xi) = K(x, \xi)\nabla u \quad \text{and} \quad C(u, \nabla u; x, \xi) = R(x, \xi)u^3,$$

where K and R are respectively the diffusion and reaction coefficients. This example will serve as a guideline. Such a semi-linear second-order stochastic partial differential equation describes transport phenomena at equilibrium such as steady-state diffusion-reaction processes arising from mathematical models in population dynamics [51–53] as well as in chemical kinetics (kinetics/dynamics of autocatalytic chemical reactions) [54–61].

1.1. Localized uncertainties and non-linearities

We consider that non-linearities and uncertainties on operator and geometry only affect a given subdomain of interest $\Lambda_\star \subset \Omega$. For the sake of simplicity, we also consider that uncertainties on the right-hand side are localized in $\Lambda_\star$.

First, the subdomain $\Lambda_\star$ may depend on ξ while the complementary subdomain $\Omega\backslash\Lambda_\star$ is supposed independent of ξ, which means that geometrical uncertainties are contained in $\Lambda_\star$. The boundary $\partial\Lambda_\star$ of $\Lambda_\star$ contains the possible uncertainties of the boundary $\partial\Omega$ of domain Ω.

Also, B and C are supposed linear and independent of ξ outside $\Lambda_\star$. More precisely, we suppose that B can be split into a linear part B_L (such that $u \mapsto B_L(u, \nabla u; x, \xi)$ is linear) and a non-linear part B_N, such that $B = B_L + B_N$. The same decomposition is introduced for $C = C_L + C_N$. Then, we consider that B and C are such that

$$B_L(\cdot,\cdot;x,\xi) = B_L(\cdot,\cdot;x) \quad \text{and} \quad C_L(\cdot,\cdot;x,\xi) = C_L(\cdot,\cdot;x) \quad \text{for } x \in \Omega \setminus \Lambda_\star, \tag{2}$$

and

$$B_N(\cdot,\cdot;x,\xi) = C_N(\cdot,\cdot;x,\xi) = 0 \quad \text{for } x \in \Omega \setminus \Lambda_\star. \tag{3}$$

Also, the prescribed source term f and normal flux g are such that

$$f(x,\xi) = f(x) \quad \text{and} \quad g(x,\xi) = g(x) \quad \text{for } x \in \Omega \setminus \Lambda_\star. \tag{4}$$

Example 1.2. In Example 1.1, we consider that diffusion coefficient K and reaction parameter R are such that $K(x,\xi) = K(x)$ and $R(x,\xi) = 0$ for $x \in \Omega \setminus \Lambda_\star$. A practical application is the chlorite-thiosulfate autocatalytic reaction [60] or the iodate-arsenous acid autocatalytic reaction [54–56, 58, 59, 61] described by a third-order reaction-diffusion equation for the iodide concentration u in the absence of convection, with localized and random molecular diffusion coefficient K and reaction rate kinetic coefficient R.

In the following, a function $v(x,\xi)$ of two variables defined for $\xi \in \Xi$ and $x \in D(\xi)$, with $D(\xi)$ a parametrized domain of $\mathbb{R}^d$, will be equivalently considered as a function $v(\xi)$ defined on $D(\xi)$. For the sake of readability, we will often omit the dependence on ξ for geometrical domains and for function spaces defined on these domains.

1.2. Assumptions

Here $\mathcal{O}$ denotes a subset of Ω. For a function $v \in \mathrm{H}^1(\mathcal{O})$, we denote

$$|v|_{\mathrm{H}^1(\mathcal{O})} = \|\nabla v\|_{\mathrm{L}^2(\mathcal{O})} \quad \text{and} \quad \|v\|^2_{\mathrm{H}^1(\mathcal{O})} = |v|^2_{\mathrm{H}^1(\mathcal{O})} + \|v\|^2_{\mathrm{L}^2(\mathcal{O})}.$$

1.2.1. *Assumptions on the source terms*

For a function v defined on $\mathcal{O}$, we introduce the linear form

$$\ell_{\mathcal{O}}(v;\xi) = \int_{\mathcal{O}} f(\cdot,\xi)v + \int_{\Gamma_N \cap \partial\mathcal{O}} g(\cdot,\xi)v, \tag{5}$$

and we assume that f and g are such that the following assumption holds.

Assumption 1.3 (Properties of linear form $\ell_{\mathcal{O}}$)**.** *We assume that the linear form $\ell_{\mathcal{O}}(\cdot;\xi)\colon \mathrm{H}^1(\mathcal{O}) \to \mathbb{R}$ is almost surely continuous, that means there exists a random variable $\kappa(\xi) > 0$ such that it holds*

$$|\ell_{\mathcal{O}}(v;\xi)| \leqslant \kappa(\xi)\|v\|_{\mathrm{H}^1(\mathcal{O})} \quad \forall v \in \mathrm{H}^1(\mathcal{O}), \tag{6}$$

and we further assume that $\kappa \in \mathrm{L}^p_\mu(\Xi)$ for some $2 \leqslant p \leqslant +\infty$.

1.2.2. *Assumptions on the differential operator*

For functions u, v defined on $\mathcal{O}$, we introduce the semi-linear form

$$d_{\mathcal{O}}(u,v;\xi) = \int_{\mathcal{O}} B(u,\nabla u;\cdot,\xi)\cdot\nabla v + \int_{\mathcal{O}} C(u,\nabla u;\cdot,\xi)v, \tag{7}$$

which can be written as

$$d_{\mathcal{O}}(u,v;\xi) = a_{\mathcal{O}}(u,v;\xi) + n_{\mathcal{O}}(u,v;\xi),$$

where $a_{\mathcal{O}}(\cdot,\cdot;\xi)$ is a bilinear form and $n_{\mathcal{O}}(\cdot,\cdot;\xi)$ is a semi-linear form, respectively defined by

$$a_{\mathcal{O}}(u,v;\xi) = \int_{\mathcal{O}} B_L(u,\nabla u;\cdot,\xi)\cdot\nabla v + \int_{\mathcal{O}} C_L(u,\nabla u;\cdot,\xi)v,$$
$$n_{\mathcal{O}}(u,v;\xi) = \int_{\mathcal{O}} B_N(u,\nabla u;\cdot,\xi)\cdot\nabla v + \int_{\mathcal{O}} C_N(u,\nabla u;\cdot,\xi)v.$$

We make the following assumptions.

Assumption 1.4 (Properties of bilinear form $a_{\mathcal{O}}$)**.** *We assume that the bilinear form $a_{\mathcal{O}}(\cdot,\cdot;\xi)\colon \mathrm{H}^1(\mathcal{O}) \times \mathrm{H}^1(\mathcal{O}) \to \mathbb{R}$ is such that there exist constants $0 < \alpha_a \leqslant \beta_a < +\infty$ such that it holds almost surely*

$$a_{\mathcal{O}}(v,v;\xi) \geqslant \alpha_a |v|^2_{\mathrm{H}^1(\mathcal{O})} \qquad \forall v \in \mathrm{H}^1(\mathcal{O}), \tag{8}$$
$$|a_{\mathcal{O}}(u,v;\xi)| \leqslant \beta_a \|u\|_{\mathrm{H}^1(\mathcal{O})}\|v\|_{\mathrm{H}^1(\mathcal{O})} \quad \forall u, v \in \mathrm{H}^1(\mathcal{O}), \tag{9}$$

and we further assume that α_a and β_a are independent of ξ and $\mathcal{O}$.

Assumption 1.5 (Properties of semi-linear form $n_\mathcal{O}$)**.** *We assume that the semi-linear form $n_\mathcal{O}(\cdot,\cdot;\xi)\colon \mathrm{H}^1(\mathcal{O})\times \mathrm{H}^1(\mathcal{O}) \to \mathbb{R}$ is almost surely continuous with respect to the second variable and radially continuous with respect to the first variable, that means for all $u, v \in \mathrm{H}^1(\mathcal{O})$, the map $t \mapsto n_\mathcal{O}(u+tv, v;\xi)$ is almost surely continuous. We also assume that $n_\mathcal{O}(\cdot,\cdot;\xi)$ is almost surely monotone in the first variable, that means*

$$n_\mathcal{O}(u, u-v;\xi) - n_\mathcal{O}(v, u-v;\xi) \geqslant 0 \quad \forall u, v \in \mathrm{H}^1(\mathcal{O}) \tag{10}$$

holds almost surely. Finally, we assume that $n_\mathcal{O}(\cdot,\cdot;\xi)$ satisfies almost surely

$$n_\mathcal{O}(0, v;\xi) = 0 \quad \forall v \in \mathrm{H}^1(\mathcal{O}). \tag{11}$$

Example 1.6. Concerning Example 1.1, assumption 1.4 on $a_\mathcal{O}$ is satisfied if the diffusion coefficient K is such that $0 < K_{\inf}|\zeta|^2 \leqslant K(x,\xi)\zeta\cdot\zeta \leqslant K_{\sup}|\zeta|^2 < +\infty$ for all $\zeta \in \mathbb{R}^d$ holds almost surely and almost everywhere on $\mathcal{O}$, where $K_{\inf}$ and $K_{\sup}$ are some strictly positive constants independent of ξ and independent of the considered subdomain $\mathcal{O} \subset \Omega$. Also, assumption 1.5 on $n_\mathcal{O}$ is satisfied if the reaction coefficient R is such that $0 \leqslant R(x,\xi) \leqslant R_{\sup} < +\infty$ holds almost surely and almost everywhere on $\mathcal{O}$, where $R_{\sup}$ is a strictly positive constant independent of ξ and independent of the considered subdomain $\mathcal{O} \subset \Omega$. That means assumptions 1.4 and 1.5 on $a_\mathcal{O}$ and $n_\mathcal{O}$ are satisfied if the diffusion coefficient K is almost surely uniformly bounded and elliptic and the reaction parameter R is almost surely non negative and uniformly bounded.

1.2.3. *Assumption on the geometry*

We suppose that the considered domains have sufficiently smooth boundary (*e.g.* Lipschitz). For a subset $\mathcal{E} \subset \partial\mathcal{O}$ with non zero measure, we denote by $\mathrm{H}^{1/2}(\mathcal{E})$ the space of traces on $\mathcal{E}$ of functions in $\mathrm{H}^1(\mathcal{O})$. We recall that we have

$$\|v\|_{\mathrm{H}^1(\mathcal{O})} \leqslant C_{\mathcal{O},\mathcal{E}} \left(|v|_{\mathrm{H}^1(\mathcal{O})} + \|v\|_{\mathrm{H}^{1/2}(\mathcal{E})} \right), \tag{12}$$

for all $v \in \mathrm{H}^1(\mathcal{O})$, with a constant $C_{\mathcal{O},\mathcal{E}}$ depending only on $\mathcal{O}$ and $\mathcal{E}$ (see [62, Theorem 7.3.13]).

Assumption 1.7. *For any considered domains $\mathcal{O}$ and $\mathcal{E} \subset \partial\mathcal{O}$, we assume that the constant $C_{\mathcal{O},\mathcal{E}}$ involved in (12) is independent of ξ.*

Assumption 1.7 is obviously satisfied if the domains $\mathcal{O}$ and $\mathcal{E}$ are independent of ξ. In the case of uncertain domains $\mathcal{O}(\xi)$ and $\mathcal{E}(\xi)$, assumption 1.7 implies some restrictions on the dependence of the geometry on the parameters ξ. Let us describe a typical situation where the uncertain domain is described through a parametrized mapping defined on a fixed domain. Assume that there exist domains $\mathcal{O}_0$ and $\mathcal{E}_0 \subset \partial\mathcal{O}_0$ independent of ξ, and a parametrized diffeomorphism $\phi(\cdot;\xi)\colon \overline{\mathcal{O}_0} \to \overline{\mathcal{O}(\xi)}$ such that $\phi(\mathcal{O}_0;\xi) = \mathcal{O}(\xi)$ and $\phi(\mathcal{E}_0;\xi) = \mathcal{E}(\xi)$. Then it can be proved that assumption 1.7 is satisfied if the mapping $\phi(\cdot;\xi)$ satisfies almost surely

$$\alpha_\phi|\zeta| \leqslant |\nabla\phi(x_0;\xi)\zeta| \leqslant \beta_\phi|\zeta| \quad \forall \zeta \in \mathbb{R}^d, \quad \forall x_0 \in \mathcal{O}_0,$$

with constants α_ϕ and β_ϕ independent of ξ. That means assumption 1.7 on the geometry is satisfied if we consider a fixed (deterministic) domain $\mathcal{O}_0$ and a random geometrical transformation $\phi(\cdot;\xi)$ mapping the fixed domain $\overline{\mathcal{O}_0}$ to the parameter-dependent domain $\overline{\mathcal{O}(\xi)}$ such that the singular values of the jacobian matrix $\nabla\phi(\cdot;\xi)$ are almost surely uniformly bounded (from above and from below).

2. Global-local formulation with patch

2.1. Domain decomposition: introduction of a patch

We introduce a subdomain $\Lambda \subset \Omega$, hereafter called a patch, such that $\Lambda_\star \subset \Lambda$, and such that $\Omega \setminus \Lambda$ is independent of ξ. This yields the following partition of domain $\Omega(\xi)$:

$$\Omega(\xi) = (\Omega \setminus \Lambda) \cup \Lambda(\xi).$$

The patch Λ is chosen such that

$$\mathrm{dist}(\Lambda_\star, \Omega \setminus \Lambda) > \delta, \tag{13}$$

that means uncertainties on operator, geometry and right-hand side affect a region in Λ whose distance to $\Omega \setminus \Lambda$ is greater than δ. We assume that the patch Λ has a sufficiently smooth boundary (*e.g.* Lipschitz). We denote by

$$\Gamma = \partial\Lambda \cap \partial(\Omega \setminus \Lambda)$$

the deterministic interface between the patch Λ and the exterior subdomain $\Omega \setminus \Lambda$ (see Figure 1).

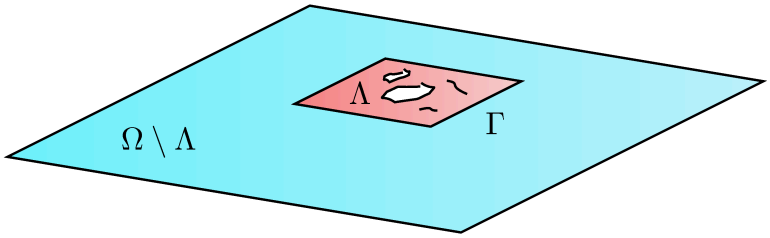


FIGURE 1. Representation of interface Γ between patch Λ and complementary subdomain $\Omega \setminus \Lambda$

We denote by $U(\xi)$ and $w(\xi)$ the restrictions of $u(\xi)$ to subdomains $\Omega \setminus \Lambda$ and Λ, respectively. For $U(\xi) \in \mathrm{H}^1(\Omega \setminus \Lambda)$ and $w(\xi) \in \mathrm{H}^1(\Lambda)$, we denote by $U(\xi)_{|\Gamma}$ and $w(\xi)_{|\Gamma}$ in $\mathrm{H}^{1/2}(\Gamma)$ the traces on Γ of $U(\xi)$ and $w(\xi)$, respectively. A weak continuity condition is enforced on interface Γ by imposing

$$b_\Gamma(\delta\lambda, U(\xi)_{|\Gamma}) = b_\Gamma(\delta\lambda, w(\xi)_{|\Gamma}) \quad \forall \delta\lambda \in \mathrm{H}^{1/2}(\Gamma)^*, \tag{14}$$

where b_Γ denotes the duality pairing between $\mathrm{H}^{1/2}(\Gamma)$ and its topological dual $\mathrm{H}^{1/2}(\Gamma)^*$. In the following, we denote by $\mathcal{M} = \mathrm{H}^{1/2}(\Gamma)^*$ and $\|\cdot\|_{\mathcal{M}} = \|\cdot\|_{\mathrm{H}^{1/2}(\Gamma)^*}$. We use the same notation b_Γ for the bilinear form $b_\Gamma \colon \mathcal{M} \times \mathrm{H}^{1/2}(\Gamma) \to \mathbb{R}$ and its extension to $\mathcal{M} \times \mathrm{H}^1(\Omega \setminus \Lambda)$ (resp. $\mathcal{M} \times \mathrm{H}^1(\Lambda)$) defined using the trace operator from $\mathrm{H}^1(\Omega \setminus \Lambda)$ (resp. $\mathrm{H}^1(\Lambda)$) to $\mathrm{H}^{1/2}(\Gamma)$.

Remark 2.1. The proposed multiscale approach can be naturally extended to the case where the sources of uncertainties and possible non-linearities are localized in several non-overlapping local subdomains of interest (or patches). The patch Λ and the interface Γ can then be respectively interpreted as the disjoint union of Q patches $\{\Lambda_q\}_{q=1}^Q$ and Q interfaces $\{\Gamma_q\}_{q=1}^Q$, where $\Gamma_q = \partial\Lambda_q \cap \partial(\Omega \setminus \Lambda)$ is the deterministic interface between the patch Λ_q and the exterior subdomain $\Omega \setminus \Lambda$.

2.1.1. *Weak formulation*

We introduce the Hilbert spaces

$$\begin{aligned}\mathcal{U} &= \{U \in \mathrm{H}^1(\Omega \setminus \Lambda) : U = 0 \text{ on } \Gamma_D \cap \partial(\Omega \setminus \Lambda)\},\\ \mathcal{W} &= \{w \in \mathrm{H}^1(\Lambda) : w = 0 \text{ on } \Gamma_D \cap \partial\Lambda\},\end{aligned}$$

equipped with norms $\|\cdot\|_{\mathcal{U}} = \|\cdot\|_{\mathrm{H}^1(\Omega\setminus\Lambda)}$ and $\|\cdot\|_{\mathcal{W}} = \|\cdot\|_{\mathrm{H}^1(\Lambda)}$, respectively. Also, we introduce the Hilbert space

$$\widehat{\mathcal{V}} = \{u \colon \Omega \to \mathbb{R} : u_{|\Omega\setminus\Lambda} \in \mathcal{U} \text{ and } u_{|\Lambda} \in \mathcal{W}\}$$

equipped with the norm $\|\cdot\|_{\mathcal{V}}$ defined by

$$\|u\|_{\mathcal{V}}^2 = \|u_{|\Omega\setminus\Lambda}\|_{\mathrm{H}^1(\Omega\setminus\Lambda)}^2 + \|u_{|\Lambda}\|_{\mathrm{H}^1(\Lambda)}^2,$$

and the closed linear subspace

$$\mathcal{V} = \{u \in \widehat{\mathcal{V}} : b_\Gamma(\delta\lambda, u_{|\Omega\setminus\Lambda}) = b_\Gamma(\delta\lambda, u_{|\Lambda}) \text{ for all } \delta\lambda \in \mathcal{M}\},$$

which is a Hilbert space when equipped with norm $\|\cdot\|_{\mathcal{V}}$.

Lemma 2.2. *There exists a constant $C_{\mathcal{V}}$ such that $|v|_{\mathcal{V}} \leqslant \|v\|_{\mathcal{V}} \leqslant C_{\mathcal{V}}|v|_{\mathcal{V}} \quad \forall v \in \mathcal{V}$, with*

$$|v|^2_{\mathcal{V}} = |v_{|\Omega\setminus\Lambda}|^2_{\mathrm{H}^1(\Omega\setminus\Lambda)} + |v_{|\Lambda}|^2_{\mathrm{H}^1(\Lambda)}.$$

Under assumption 1.7, $C_{\mathcal{V}}$ is independent of ξ.

Proof. See Section 7.3 in Appendix 7. □

In the following, for a given Hilbert space H (possibly dependent on ξ) equipped with a norm $\|\cdot\|_H$, we denote by H^Ξ the space $H^\Xi := \{v\colon \xi \in \Xi \mapsto v(\xi) \in H(\xi)\}$, and we identify functions in H^Ξ that are equal almost surely. We denote by $\mathrm{L}^p_\mu(\Xi; H) = \{v \in H^\Xi : \mathbb{E}(\|v(\xi)\|^p_{H(\xi)}) < +\infty\}$, where $\mathbb{E}$ denotes the mathematical expectation defined by $\mathbb{E}(a(\xi)) = \int_\Xi a(\xi)\mu(\mathrm{d}\xi)$.

We consider the following weak formulation of the problem: find $u \in \mathcal{V}^\Xi$ such that it holds almost surely

$$d_\Omega(u(\xi), \delta u; \xi) = \ell_\Omega(\delta u; \xi) \quad \forall \delta u \in \mathcal{V}. \tag{15}$$

Theorem 2.3. *Under assumptions 1.3, 1.4 and 1.5, problem (15) is well-posed, that means for almost all $\xi \in \Xi$, it admits a unique solution $u(\xi) \in \mathcal{V}$ and the application that maps $\ell_\Omega(\cdot;\xi)$ to this solution $u(\xi)$ is Lipschitz continuous with Lipschitz constant $C^2_{\mathcal{V}}/\alpha_a$. Moreover, under assumption 1.7, the solution $u \in \mathrm{L}^p_\mu(\Xi; \mathcal{V})$, with exponent p defined in assumption 1.3.*

Proof. See Section 7.4 in Appendix 7. □

2.1.2. *Reformulation using a Lagrange multiplier*

From (2) and (3), we have that

$$\begin{aligned} d_\Omega(u(\xi), \delta u; \xi) &= a_{\Omega\setminus\Lambda}(U(\xi), \delta U) + a_\Lambda(w(\xi), \delta w; \xi) + n_\Lambda(w(\xi), \delta w; \xi), \\ \ell_\Omega(\delta u; \xi) &= \ell_{\Omega\setminus\Lambda}(\delta U) + \ell_\Lambda(\delta w; \xi), \end{aligned}$$

for all $\delta u\colon \Omega \to \mathbb{R}$ such that $\delta u_{|\Omega\setminus\Lambda} = \delta U$ and $\delta u_{|\Lambda} = \delta w$. A formulation equivalent to (15) can be written as follows: find $(U, w, \lambda) \in \mathcal{U}^\Xi \times \mathcal{W}^\Xi \times \mathcal{M}^\Xi$ such that it satisfies almost surely

$$a_{\Omega\setminus\Lambda}(U(\xi), \delta U) + b_\Gamma(\lambda(\xi), \delta U) = \ell_{\Omega\setminus\Lambda}(\delta U), \tag{16a}$$

$$a_\Lambda(w(\xi), \delta w; \xi) + n_\Lambda(w(\xi), \delta w; \xi) - b_\Gamma(\lambda(\xi), \delta w) = \ell_\Lambda(\delta w; \xi), \tag{16b}$$

$$b_\Gamma(\delta\lambda, U(\xi)) - b_\Gamma(\delta\lambda, w(\xi)) = 0, \tag{16c}$$

for all $(\delta U, \delta w, \delta\lambda) \in \mathcal{U} \times \mathcal{W} \times \mathcal{M}$, where λ represents the Lagrange multiplier allowing to ensure the weak continuity condition (14) of solution u across interface Γ.

Theorem 2.4. *Under assumptions 1.3, 1.4 and 1.5, problem (16) admits a unique solution $(U(\xi), w(\xi), \lambda(\xi)) \in \mathcal{U} \times \mathcal{W} \times \mathcal{M}$ for almost all $\xi \in \Xi$. Moreover, under assumption 1.7, $U \in \mathrm{L}^p_\mu(\Xi; \mathcal{U})$, $w \in \mathrm{L}^p_\mu(\Xi; \mathcal{W})$ and $\lambda \in \mathrm{L}^p_\mu(\Xi; \mathcal{M})$, with exponent p defined in assumption 1.3.*

Proof. See Section 7.5 in Appendix 7. □

2.2. **Reformulation with extended domain: introduction of a fictitious patch**

Let us now introduce a deterministic fictitious patch $\widetilde{\Lambda} \supset \Lambda$ such that $\Gamma \subset \partial\widetilde{\Lambda}$ and define the corresponding deterministic fictitious domain $\widetilde{\Omega} \supset \Omega$ such that $\widetilde{\Omega} = (\Omega \setminus \Lambda) \cup \widetilde{\Lambda}$ and $\widetilde{\Omega} \setminus \widetilde{\Lambda} = \Omega \setminus \Lambda$ (see Figure 2). Note that in the case where the patch Λ does not contain any geometrical details (*i.e.* no internal boundary such as holes, cracks, *etc*), we simply have $\widetilde{\Lambda} = \Lambda$ and $\widetilde{\Omega} = \Omega$.

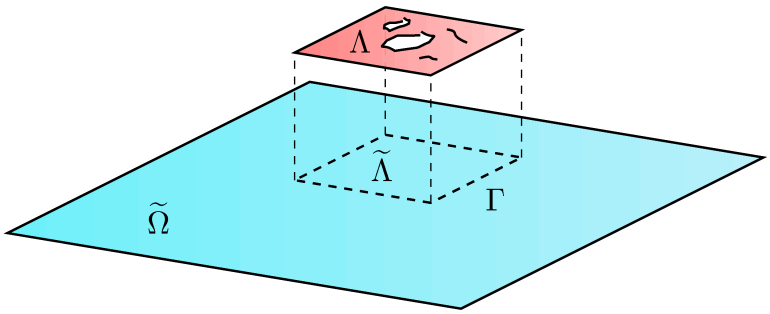


FIGURE 2. Representation of fictitious domain $\widetilde{\Omega}$, fictitious patch $\widetilde{\Lambda}$, real patch Λ and interface Γ

We now consider an extension of global solution U from subdomain $\Omega\setminus\Lambda$ to fictitious domain $\widetilde{\Omega}$. We introduce the new Hilbert space $\widetilde{\mathcal{U}} = \{U \in \mathrm{H}^1(\widetilde{\Omega}) : U = 0 \text{ on } \Gamma_D \cap \partial\widetilde{\Omega}\}$ equipped with the norm $\|\cdot\|_{\widetilde{\mathcal{U}}} = \|\cdot\|_{\mathrm{H}^1(\widetilde{\Omega})}$. We then define a new bilinear form $c_{\widetilde{\Omega}} \colon \widetilde{\mathcal{U}}\times\widetilde{\mathcal{U}} \to \mathbb{R}$ as the following extension of $a_{\Omega\setminus\Lambda} \colon \mathcal{U}\times\mathcal{U} \to \mathbb{R}$ to $\widetilde{\mathcal{U}}\times\widetilde{\mathcal{U}}$: for all $U, V \in \widetilde{\mathcal{U}}$,

$$c_{\widetilde{\Omega}}(U,V) = a_{\Omega\setminus\Lambda}(U,V) + c_{\widetilde{\Lambda}}(U,V), \tag{17}$$

where, for a subdomain $\mathcal{O} \subset \widetilde{\Omega}$, $c_{\mathcal{O}}$ is a bilinear form defined by

$$c_{\mathcal{O}}(U,V) = \int_{\mathcal{O}} \widetilde{B}_L(U,\nabla U;\cdot)\cdot\nabla V + \int_{\mathcal{O}} \widetilde{C}_L(U,\nabla U;\cdot)V,$$

where $\widetilde{B}_L(\cdot,\cdot;x,\xi) \colon \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}^d$ and $\widetilde{C}_L(\cdot,\cdot;x,\xi) \colon \mathbb{R}\times\mathbb{R}^d \to \mathbb{R}$ are such that

$$\widetilde{B}_L(\cdot,\cdot;x) = B_L(\cdot,\cdot;x) \quad \text{and} \quad \widetilde{C}_L(\cdot,\cdot;x) = C_L(\cdot,\cdot;x) \quad \text{for } x \in \Omega \setminus \Lambda.$$

We make the following assumption. Here $\mathcal{O}$ denotes a subset of $\widetilde{\Omega}$.

Assumption 2.5 (Properties of bilinear form $c_{\mathcal{O}}$)**.** *We assume that the bilinear form $c_{\mathcal{O}} \colon \mathrm{H}^1(\mathcal{O})\times\mathrm{H}^1(\mathcal{O}) \to \mathbb{R}$ is symmetric and such that there exist constants $0 < \alpha_c \leqslant \beta_c < +\infty$ such that*

$$c_{\mathcal{O}}(V,V) \geqslant \alpha_c |V|^2_{\mathrm{H}^1(\mathcal{O})} \qquad \forall V \in \mathrm{H}^1(\mathcal{O}), \tag{18}$$

$$|c_{\mathcal{O}}(U,V)| \leqslant \beta_c \|U\|_{\mathrm{H}^1(\mathcal{O})}\|V\|_{\mathrm{H}^1(\mathcal{O})} \quad \forall U, V \in \mathrm{H}^1(\mathcal{O}), \tag{19}$$

and we further assume that α_c and β_c are independent of ξ and $\mathcal{O}$.

Example 2.6. In Example 1.1, $\widetilde{B}_L$ and $\widetilde{C}_L$ can be respectively defined by

$$\widetilde{B}_L(U,\nabla U;x) = \widetilde{K}(x)\nabla U \quad \text{and} \quad \widetilde{C}_L(U,\nabla U;x) = 0,$$

where $\widetilde{K}$ is a fictitious diffusion coefficient such that $\widetilde{K}(x) = K(x)$ for $x \in \Omega \setminus \Lambda$. Assumption 2.5 on $c_{\widetilde{\Omega}}$ is satisfied if the fictitious diffusion coefficient $\widetilde{K}$ is uniformly bounded and elliptic on $\widetilde{\Omega}$, that means condition

$0 < \widetilde{K}_{\inf}|\zeta|^2 \leqslant \widetilde{K}(x)\zeta\cdot\zeta \leqslant \widetilde{K}_{\sup}|\zeta|^2 < +\infty$ for all $\zeta \in \mathbb{R}^d$ holds almost everywhere on $\widetilde{\Omega}$, where $\widetilde{K}_{\inf}$ and $\widetilde{K}_{\sup}$ are some strictly positive constants.

Afterwards, a reformulation of the global-local problem (16) reads: find $(U, w, \lambda) \in \widetilde{\mathcal{U}}^{\Xi}\times\mathcal{W}^{\Xi}\times\mathcal{M}^{\Xi}$ such that it satisfies almost surely

$$c_{\widetilde{\Omega}}(U(\xi),\delta U) - c_{\widetilde{\Lambda}}(U(\xi),\delta U) + b_{\Gamma}(\lambda(\xi),\delta U) = \ell_{\Omega\setminus\Lambda}(\delta U), \tag{20a}$$

$$a_{\Lambda}(w(\xi),\delta w;\xi) + n_{\Lambda}(w(\xi),\delta w;\xi) - b_{\Gamma}(\lambda(\xi),\delta w) = \ell_{\Lambda}(\delta w;\xi), \tag{20b}$$

$$b_{\Gamma}(\delta\lambda, U(\xi)) - b_{\Gamma}(\delta\lambda, w(\xi)) = 0, \tag{20c}$$

for all $(\delta U, \delta w, \delta\lambda) \in \widetilde{\mathcal{U}}\times\mathcal{W}\times\mathcal{M}$. Let us here mention that problem (20) admits infinitely many solutions (U, w, λ) that only differ by the value of global solution U in fictitious patch $\widetilde{\Lambda}$. A particular solution can be uniquely defined by defining the value of U in $\widetilde{\Lambda}$ as a particular extension of the value of U on interface Γ. The global-local iterative algorithm presented in the next section will be proven to converge to a solution (U, w, λ) in a subspace of $\widetilde{\mathcal{U}}^{\Xi}\times\mathcal{W}^{\Xi}\times\mathcal{M}^{\Xi}$ corresponding to a particular definition of the extension.

3. Global-local iterative algorithm

We now introduce and analyze an iterative algorithm to solve problem (20).

3.1. Description of the algorithm

We initialize the algorithm with $U^0 = w^0 = \lambda^0 = 0$. Then, at iteration $k \geqslant 1$, $(U^k, w^k, \lambda^k) \in \widetilde{\mathcal{U}}^{\Xi}\times\mathcal{W}^{\Xi}\times\mathcal{M}^{\Xi}$ is defined by three steps (global step, relaxation step and local step), described below.

3.1.1. *Global step*

We first define $\widehat{U}^k \in \widetilde{\mathcal{U}}^{\Xi}$ such that it satisfies almost surely

$$c_{\widetilde{\Omega}}(\widehat{U}^k(\xi),\delta U) = c_{\widetilde{\Lambda}}(U^{k-1}(\xi),\delta U) - b_{\Gamma}(\lambda^{k-1}(\xi),\delta U) + \ell_{\Omega\setminus\Lambda}(\delta U) \tag{21}$$

for all $\delta U \in \widetilde{\mathcal{U}}$. The computation of $\widehat{U}^k \in \widetilde{\mathcal{U}}^{\Xi}$ thus requires the solution of a linear problem defined on fictitious domain $\widetilde{\Omega}$ with a deterministic operator and an uncertain right-hand side (involving Lagrange multiplier λ^{k-1} on interface Γ and global iterate U^{k-1} in fictitious patch $\widetilde{\Lambda}$ at previous iteration $k-1$).

Remark 3.1. Although $\widetilde{B}_L$ and $\widetilde{C}_L$ could *a priori* be chosen arbitrarily (uncertain or deterministic) on $\widetilde{\Lambda}$, a convenient choice consists in taking for $\widetilde{B}_L$ and $\widetilde{C}_L$ parameter-independent functions, *i.e.* $\widetilde{B}_L(\cdot,\cdot;x,\xi) = \widetilde{B}_L(\cdot,\cdot;x)$ and $\widetilde{C}_L(\cdot,\cdot;x,\xi) = \widetilde{C}_L(\cdot,\cdot;x)$ for $x \in \widetilde{\Lambda}$, which allows to preserve a linear global problem with deterministic linear operator throughout iterations. Also, a natural choice consists in taking for $\widetilde{B}_L$ and $\widetilde{C}_L$ over $\widetilde{\Lambda}$ the mean value of the corresponding linear functions B_L and C_L over Λ, *i.e.* $\widetilde{B}_L(\cdot,\cdot;x) = \mathbb{E}\left(B_L(\cdot,\cdot;x,\xi)\right)$ and $\widetilde{C}_L(\cdot,\cdot;x) = \mathbb{E}\left(C_L(\cdot,\cdot;x,\xi)\right)$ for $x \in \widetilde{\Lambda}$. Besides, choosing parameter-dependent functions $\widetilde{B}_L$ and $\widetilde{C}_L$ on $\widetilde{\Lambda}$ could allow to accelerate the convergence of the algorithm (see Remark 3.14). Another possible choice would consist in taking for $\widetilde{B}_L$ and $\widetilde{C}_L$ over $\widetilde{\Lambda}$ the tangent linear functions to the corresponding semi-linear functions $B = B_L + B_N$ and $C = C_L + C_N$ over Λ.

Remark 3.2. Assume that $\partial\widetilde{\Lambda} = \Gamma \cup (\partial\widetilde{\Lambda}\cap\Gamma_D)$. By using Green's formula in the definition of $c_{\widetilde{\Lambda}}$, the global problem (21) can be reformulated as

$$c_{\widetilde{\Omega}}(\widehat{U}^k(\xi),\delta U) = -b_{\Gamma}(\mu^{k-1}(\xi) + \lambda^{k-1}(\xi),\delta U) + \ell_{\Omega\setminus\Lambda}(\delta U) + \ell_{\widetilde{\Lambda}}(\delta U;\xi) \tag{22}$$

for all $\delta U \in \widetilde{\mathcal{U}}$, where $\mu^{k-1}(\xi) \in \mathcal{M}$ is defined by the following expression (interpreted in a weak sense): $\mu^{k-1}(\xi) = -\widetilde{B}_L(U^{k-1}(\xi), \nabla U^{k-1}(\xi); \cdot) \cdot n$ on Γ, with n the unit normal to Γ pointing outward $\widetilde{\Lambda}$, and where $\ell_{\widetilde{\Lambda}}(\cdot;\xi)$ is a linear form defined by $\ell_{\widetilde{\Lambda}}(V;\xi) = -\int_{\widetilde{\Lambda}} \nabla \cdot \widetilde{B}_L(U^{k-1}(\xi), \nabla U^{k-1}(\xi); \cdot)V + \int_{\widetilde{\Lambda}} \widetilde{C}_L(U^{k-1}(\xi), \nabla U^{k-1}(\xi); \cdot)V$. The quantity $\mu^{k-1} + \lambda^{k-1}$ is seen as a flux discontinuity on the interface Γ between global and local models. The iterative algorithm can then be interpreted as a modified Newton method (with constant linear global operator) formulated on the flux equilibrium over interface Γ (interpreted in a weak sense) [32,34].

3.1.2. *Relaxation step*

We then define $U^k \in \widetilde{\mathcal{U}}^{\Xi}$ by

$$U^k(\xi) = \rho_k \widehat{U}^k(\xi) + (1-\rho_k) U^{k-1}(\xi), \tag{23}$$

where $\rho_k > 0$ is a relaxation parameter (possibly depending on ξ) chosen sufficiently small to ensure convergence (see convergence analysis in Section 3.2.2). Relaxation parameter ρ_k may have a significant impact on the convergence and stability properties of the algorithm. Practical choices for ρ_k will be discussed in Section 4.4.

3.1.3. *Local step*

We finally define $(w^k, \lambda^k) \in \mathcal{W}^{\Xi} \times \mathcal{M}^{\Xi}$ such that it satisfies almost surely

$$a_\Lambda(w^k(\xi), \delta w; \xi) + n_\Lambda(w^k(\xi), \delta w; \xi) - b_\Gamma(\lambda^k(\xi), \delta w) = \ell_\Lambda(\delta w; \xi), \tag{24a}$$

$$b_\Gamma(\delta\lambda, w^k(\xi)) = b_\Gamma(\delta\lambda, U^k(\xi)), \tag{24b}$$

for all $(\delta w, \delta\lambda) \in \mathcal{W} \times \mathcal{M}$. The computation of $(w^k, \lambda^k) \in \mathcal{W}^{\Xi} \times \mathcal{M}^{\Xi}$ thus requires the solution of a non-linear problem defined on patch Λ with uncertain operator and right-hand side (involving global iterate U^k at current iteration k as a boundary data). The Lagrange multiplier λ^k allows to enforce the weak continuity conditions on interface Γ between local iterate w^k and global iterate U^k, which corresponds to non-homogeneous Dirichlet boundary conditions imposed on an external boundary Γ of patch Λ in the local computation. Recall that, contrary to the global step, the local step takes into account possible non-linearities and uncertainties in the operator, as well as possible uncertainties in the geometry of the domain.

Remark 3.3. The local problem (24) can be reformulated as a single-field problem by noting $w^k(\xi) = \widetilde{w}^k(\xi) + z^k(\xi)$, where $\widetilde{w}^k(\xi) \in \mathcal{W}$ is an extension of global iterate $U^k(\xi)$ from interface Γ to patch Λ such that $b_\Gamma(\delta\lambda, \widetilde{w}^k(\xi)) = b_\Gamma(\delta\lambda, U^k(\xi))$ for all $\delta\lambda \in \mathcal{M}$, and $z^k(\xi) \in \mathcal{W}_0 = \{z \in \mathcal{W} : z = 0 \text{ on } \Gamma\}$. Local problem then consists in computing $z^k \in \mathcal{W}_0^{\Xi}$ such that it satisfies almost surely $a_\Lambda(\widetilde{w}^k(\xi)+z^k(\xi), \delta z; \xi)+n_\Lambda(\widetilde{w}^k(\xi)+z^k(\xi), \delta z; \xi) = \ell_\Lambda(\delta z; \xi)$ for all $\delta z \in \mathcal{W}_0$. The Lagrange multiplier $\lambda^k \in \mathcal{M}^{\Xi}$ is then determined *a posteriori* from (24a).

Remark 3.4. When the patch Λ contains geometrical variabilities, the local stochastic problem (24) can be reformulated on a fixed (deterministic) domain by using either a random mapping technique [63,64] or a fictitious domain method [43,65–67]. Both techniques do not require any remeshing procedure and allow to handle complex geometries. The former approach consists in introducing a suitable random mapping between a random domain and a reference deterministic domain, thus transforming a (deterministic or stochastic) partial differential equation defined on a random domain into a stochastic partial differential equation (with uncertain operator and right-hand side depending on the mapping and its derivatives) defined on a deterministic domain. The latter approach consists in embedding a random domain into a deterministic fictitious domain and considering a prolongation of the solution on the fictitious domain. In the present context of multiscale problems with localized geometrical variabilities, different fictitious domain formulations (depending on the type of boundary conditions) have been proposed in [43].

Remark 3.5. Following Remark 2.1, in the case of Q non-overlapping patches $\{\Lambda_q\}_{q=1}^Q$, the local step consists in solving Q independent non-linear local problems defined on each of the patches Λ_q. The solution of such uncoupled problems can be performed independently on each patch Λ_q in a fully parallel way.

3.2. Analysis of the algorithm

3.2.1. *Consistency*

Let $(U, w, \lambda) \in \mathcal{U}^{\Xi} \times \mathcal{W}^{\Xi} \times \mathcal{M}^{\Xi}$ denote the solution of the initial problem (16). We now introduce the closed linear subspace $\widetilde{\mathcal{U}}_\star$ of $\widetilde{\mathcal{U}}$ defined by

$$\widetilde{\mathcal{U}}_\star = \{V \in \widetilde{\mathcal{U}} : c_{\widetilde{\Lambda}}(V, \delta U) = 0 \text{ for all } \delta U \in \mathrm{H}^1_0(\widetilde{\Lambda})\},$$

where $\mathrm{H}^1_0(\widetilde{\Lambda})$ is considered as the subset of functions of $\widetilde{\mathcal{U}}$ which are zero on $\widetilde{\Omega} \setminus \widetilde{\Lambda}$. For any function $V \in \mathcal{U}$, there exists a unique extension $\widetilde{V} \in \widetilde{\mathcal{U}}_\star$ such that $\widetilde{V} = V$ on $\Omega \setminus \Lambda$ and the restriction of $\widetilde{V}$ to $\widetilde{\Lambda}$ is uniquely defined by the trace of V on the interface Γ. Then, we also denote by $U(\xi) \in \widetilde{\mathcal{U}}_\star$ the unique extension to $\widetilde{\Omega}$ of the global solution $U(\xi) \in \mathcal{U}$.

Lemma 3.6. *All global iterates $U^k(\xi)$ belong to the subspace $\widetilde{\mathcal{U}}_\star$.*

Proof. Considering test functions $\delta U \in \mathrm{H}^1_0(\widetilde{\Lambda})$ in global problem (21), we obtain that for all $k \geqslant 1$, $c_{\widetilde{\Lambda}}(\widehat{U}^k(\xi), \delta U) = c_{\widetilde{\Lambda}}(U^{k-1}(\xi), \delta U)$ for all $\delta U \in \mathrm{H}^1_0(\widetilde{\Lambda})$. Then, using (23), we have that $c_{\widetilde{\Lambda}}(U^k(\xi), \delta U) = c_{\widetilde{\Lambda}}(U^{k-1}(\xi), \delta U)$ for all $\delta U \in \mathrm{H}^1_0(\widetilde{\Lambda})$. Since $U^0 = 0$, we obtain by induction that all global iterates $U^k(\xi)$ belong to $\widetilde{\mathcal{U}}_\star$. □

We then derive the following consistency result.

Theorem 3.7 (Consistency)**.** *If the sequence $\{(U^k(\xi), w^k(\xi), \lambda^k(\xi))\}_{k\in\mathbb{N}}$ strongly converges to an element $(\widetilde{U}(\xi), w(\xi), \lambda(\xi))$ in $\widetilde{\mathcal{U}}\times\mathcal{W}\times\mathcal{M}$, then $(\widetilde{U}(\xi)_{|\Omega\setminus\Lambda}, w(\xi), \lambda(\xi)) \in \mathcal{U}\times\mathcal{W}\times\mathcal{M}$ is the unique solution $(U(\xi), w(\xi), \lambda(\xi))$ of problem (16). Also, the limit $\widetilde{U}(\xi)$ is the unique extension of $U(\xi)$ to $\widetilde{\mathcal{U}}_\star$.*

Proof. Taking the limit with k in (21), (23) and (24), we obtain that $(\widetilde{U}(\xi), w(\xi), \lambda(\xi))$ satisfies problem (20), and therefore $(\widetilde{U}(\xi)_{|\Omega\setminus\Lambda}, w(\xi), \lambda(\xi)) \in \mathcal{U}\times\mathcal{W}\times\mathcal{M}$ is the unique solution of problem (16). Then, as all global iterates $U^k(\xi)$ belong to the closed linear subspace $\widetilde{\mathcal{U}}_\star$ of $\widetilde{\mathcal{U}}$ (see Lemma 3.6), the limit $\widetilde{U}(\xi)$ also belongs to $\widetilde{\mathcal{U}}_\star$. □

Note that problem (20) is well-posed in $\widetilde{\mathcal{U}}_\star\times\mathcal{W}\times\mathcal{M}$ and admits $(U(\xi), w(\xi), \lambda(\xi)) \in \widetilde{\mathcal{U}}_\star\times\mathcal{W}\times\mathcal{M}$ as its unique solution. The algorithm can then be analyzed in the subspace $\widetilde{\mathcal{U}}_\star$ of $\widetilde{\mathcal{U}}$ and we have the following useful result which proves that $\|\cdot\|_{\widetilde{\mathcal{U}}}$ defines a norm equivalent to $\|\cdot\|_{\mathcal{U}}$ on $\widetilde{\mathcal{U}}_\star$.

Lemma 3.8. *The norms $\|\cdot\|_{\mathcal{U}}$ and $\|\cdot\|_{\widetilde{\mathcal{U}}}$ are equivalent on $\widetilde{\mathcal{U}}_\star$, with $\|V\|_{\mathcal{U}} \leqslant \|V\|_{\widetilde{\mathcal{U}}} \leqslant C_{\widetilde{\mathcal{U}}}\|V\|_{\mathcal{U}}$ for all $V \in \widetilde{\mathcal{U}}_\star$, with a constant $C_{\widetilde{\mathcal{U}}}$ independent of ξ.*

Proof. See Section 7.6 in Appendix 7. □

From Theorem 2.4 and Lemma 3.8, we directly deduce the following property.

Corollary 3.9. *The extended global solution U is in $\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})$, with exponent p defined in assumption 1.3.*

3.2.2. *Convergence*

We now prove the convergence of the sequence $\{(U^k(\xi), w^k(\xi), \lambda^k(\xi))\}_{k\in\mathbb{N}}$ to the exact solution $(U(\xi), w(\xi), \lambda(\xi))$ in $\widetilde{\mathcal{U}}_\star\times\mathcal{W}\times\mathcal{M}$. The global problem (21) being linear, the solution $\widehat{U}^k \in \widetilde{\mathcal{U}}^{\Xi}$ can be written as

$$\widehat{U}^k(\xi) = \overline{U} + \Upsilon(U^{k-1}(\xi)) + \Phi(\lambda^{k-1}(\xi)),$$

where $\Upsilon\colon \widetilde{\mathcal{U}} \to \widetilde{\mathcal{U}}$ and $\Phi\colon \mathcal{M} \to \widetilde{\mathcal{U}}$ are linear mappings. Mapping Υ is such that for $V \in \widetilde{\mathcal{U}}$, $\Upsilon(V) \in \widetilde{\mathcal{U}}$ is the unique solution of

$$c_{\widetilde{\Omega}}(\Upsilon(V), \delta U) = c_{\widetilde{\Lambda}}(V, \delta U) \quad \forall \delta U \in \widetilde{\mathcal{U}}. \tag{25a}$$

Similarly, mapping Φ is such that for $\beta \in \mathcal{M}$, $\Phi(\beta) \in \widetilde{\mathcal{U}}$ is the unique solution of

$$c_{\widetilde{\Omega}}(\Phi(\beta), \delta U) = -b_\Gamma(\beta, \delta U) \quad \forall \delta U \in \widetilde{\mathcal{U}}. \tag{25b}$$

Lastly, $\overline{U} \in \widetilde{\mathcal{U}}$ is the unique solution of

$$c_{\widetilde{\Omega}}(\overline{U}, \delta U) = \ell_{\Omega\setminus\Lambda}(\delta U) \quad \forall \delta U \in \widetilde{\mathcal{U}}.$$

The solution $(w^k, \lambda^k) \in \mathcal{W}^\Xi \times \mathcal{M}^\Xi$ of the local problem (24) can be written as

$$w^k(\xi) = \Theta(U^k(\xi); \xi) \quad \text{and} \quad \lambda^k(\xi) = \Psi(U^k(\xi); \xi),$$

where $\Theta(\cdot;\xi)\colon \widetilde{\mathcal{U}} \to \mathcal{W}$ and $\Psi(\cdot;\xi)\colon \widetilde{\mathcal{U}} \to \mathcal{M}$ are non-linear mappings. Mappings Θ and Ψ are such that for $V \in \widetilde{\mathcal{U}}$, $(\Theta(V;\xi), \Psi(V;\xi)) \in \mathcal{W}\times\mathcal{M}$ is the solution of

$$a_\Lambda(\Theta(V;\xi), \delta w; \xi) + n_\Lambda(\Theta(V;\xi), \delta w; \xi) - b_\Gamma(\Psi(V;\xi), \delta w) = \ell_\Lambda(\delta w; \xi) \qquad \forall \delta w \in \mathcal{W}, \tag{26a}$$

$$b_\Gamma(\delta\lambda, \Theta(V;\xi)) = b_\Gamma(\delta\lambda, V) \qquad \forall \delta\lambda \in \mathcal{M}. \tag{26b}$$

Consequently, the algorithm generates a sequence $\{(U^k, w^k, \lambda^k)\}_{k\in\mathbb{N}}$ by applying the following iterative scheme:

$$U^k(\xi) = \rho_k \left(\overline{U} + \Upsilon(U^{k-1}(\xi)) + \Phi(\lambda^{k-1}(\xi))\right) + (1-\rho_k)U^{k-1}(\xi), \tag{27a}$$

$$w^k(\xi) = \Theta(U^k(\xi); \xi), \tag{27b}$$

$$\lambda^k(\xi) = \Psi(U^k(\xi); \xi). \tag{27c}$$

Lemma 3.10. *The linear mappings $\Upsilon\colon \widetilde{\mathcal{U}} \to \widetilde{\mathcal{U}}$ and $\Phi\colon \mathcal{M} \to \widetilde{\mathcal{U}}$ defined in (25) are continuous, with respective continuity constants β_Υ and β_Φ independent of ξ.*

Proof. See Section 7.7 in Appendix 7. □

Lemma 3.11. *The non-linear mappings $\Theta(\cdot;\xi)\colon \widetilde{\mathcal{U}} \to \mathcal{W}$ and $\Psi(\cdot;\xi)\colon \widetilde{\mathcal{U}} \to \mathcal{M}$ defined in (26) are Lipschitz continuous, with respective Lipschitz constants β_Θ and β_Ψ independent of ξ.*

Proof. See Section 7.8 in Appendix 7. □

Let us now define the errors at a given iteration of the algorithm. At the global level, the error at iteration k is

$$\begin{aligned}\widehat{U}^k(\xi) - U(\xi) &= \Upsilon(U^{k-1}(\xi)) - \Upsilon(U(\xi)) + \Phi(\lambda^{k-1}(\xi)) - \Phi(\lambda(\xi)) \\ &= \Upsilon(U^{k-1}(\xi) - U(\xi)) + \Phi(\Psi(U^{k-1}(\xi);\xi) - \Psi(U(\xi);\xi)), \\ &= U^{k-1}(\xi) - U(\xi) - (A(U^{k-1}(\xi);\xi) - A(U(\xi);\xi)),\end{aligned}$$

and

$$\begin{aligned}U^k(\xi) - U(\xi) &= \rho_k(\widehat{U}^k(\xi) - U(\xi)) + (1-\rho_k)(U^{k-1}(\xi) - U(\xi)) \\ &= U^{k-1}(\xi) - U(\xi) - \rho_k(A(U^{k-1}(\xi);\xi) - A(U(\xi);\xi)),\end{aligned}$$

where $A(\cdot;\xi)\colon \widetilde{\mathcal{U}} \to \widetilde{\mathcal{U}}$ is the non-linear mapping defined by

$$A(V;\xi) = V - \Upsilon(V) - \Phi(\Psi(V;\xi)). \tag{28}$$

From the definitions of $\Upsilon(V) \in \widetilde{\mathcal{U}}$ and $\Phi(\Psi(V;\xi)) \in \widetilde{\mathcal{U}}$ and from (17), we deduce that mapping A is such that for $V \in \widetilde{\mathcal{U}}$, $A(V;\xi) \in \widetilde{\mathcal{U}}$ is the solution of

$$c_{\widetilde{\Omega}}(A(V;\xi),\delta U) = a_{\Omega\setminus\Lambda}(V,\delta U) + b_{\Gamma}(\Psi(V;\xi),\delta U) \quad \forall \delta U \in \widetilde{\mathcal{U}}. \tag{29}$$

Note that the non-linear nature of map A is inherited from that of map Ψ. At the local level, the error at iteration k writes

$$w^k(\xi) - w(\xi) = \Theta(U^k(\xi);\xi) - \Theta(U(\xi);\xi), \tag{30a}$$
$$\lambda^k(\xi) - \lambda(\xi) = \Psi(U^k(\xi);\xi) - \Psi(U(\xi);\xi). \tag{30b}$$

Given that non-linear map $\Theta(\cdot;\xi)$ (resp. $\Psi(\cdot;\xi)$) is Lipschitz continuous, the almost sure convergence of the local sequence $\{w^k(\xi)\}_{k\in\mathbb{N}}$ (resp. $\{\lambda^k(\xi)\}_{k\in\mathbb{N}}$) to $w(\xi)$ (resp. $\lambda(\xi)$) in $\mathcal{W}$ (resp. $\mathcal{M}$) can be directly obtained from that of the global sequence $\{U^k(\xi)\}_{k\in\mathbb{N}}$ to $U(\xi)$ in $\widetilde{\mathcal{U}}$. Recalling that the exact global solution $U(\xi)$ as well as all global iterates $U^k(\xi)$ belong to subspace $\widetilde{\mathcal{U}}_\star \subset \widetilde{\mathcal{U}}$, one can restrict the convergence analysis to that of the sequence $\{U^k(\xi)\}_{k\in\mathbb{N}}$ to $U(\xi)$ in the subspace $\widetilde{\mathcal{U}}_\star$. Let $C_{\widetilde{\Omega}}\colon \widetilde{\mathcal{U}} \to \widetilde{\mathcal{U}}$ be the linear map defined for all $U, V \in \widetilde{\mathcal{U}}$ by $\langle C_{\widetilde{\Omega}}(U), V\rangle_{\widetilde{\mathcal{U}}} = c_{\widetilde{\Omega}}(U,V)$.

Lemma 3.12. *The non-linear mapping $D(\cdot;\xi)\colon \widetilde{\mathcal{U}}_\star \to \widetilde{\mathcal{U}}$ defined by $D(\cdot;\xi) = C_{\widetilde{\Omega}}(A(\cdot;\xi))$ with $A(\cdot;\xi)$ defined in (28), is Lipschitz continuous and strongly monotone, with Lipschitz constant β_D and strong monotonicity constant α_D both independent of ξ.*

Proof. See Section 7.9 in Appendix 7. □

One iteration of the algorithm can be written as

$$U^k(\xi) = B_{\rho_k}(U^{k-1}(\xi);\xi),$$

where $B_{\rho_k}(\cdot;\xi)\colon \widetilde{\mathcal{U}} \to \widetilde{\mathcal{U}}$ is the non-linear iteration map defined by $B_{\rho_k}(V;\xi) = \rho_k \overline{U} + V - \rho_k A(V;\xi)$. We finally derive the following convergence result.

Theorem 3.13 (Convergence)**.** *Assume that the sequence of relaxation parameters $\{\rho_k\}_{k\in\mathbb{N}}$ is such that*

$$0 < \rho_{\inf} \leqslant \rho_k \leqslant \rho_{\sup} < +\infty, \tag{31}$$

for some strictly positive constants $\rho_{\inf}$ and $\rho_{\sup}$ independent of ξ and k. Then, for $\rho_{\sup}$ sufficiently small, the sequence $\{(U^k(\xi), w^k(\xi), \lambda^k(\xi))\}_{k\in\mathbb{N}}$ converges almost surely to the unique solution $(U(\xi), w(\xi), \lambda(\xi))$ of problem (20) in $\widetilde{\mathcal{U}}_\star \times \mathcal{W} \times \mathcal{M}$. Also, the sequence $\{U^k\}_{k\in\mathbb{N}}$ (resp. $\{w^k\}_{k\in\mathbb{N}}$ and $\{\lambda^k\}_{k\in\mathbb{N}}$) converges to U (resp. w and λ) in $\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})$ (resp. $\mathrm{L}^p_\mu(\Xi;\mathcal{W})$ and $\mathrm{L}^p_\mu(\Xi;\mathcal{M})$).

Proof. See Section 7.10 in Appendix 7. □

Remark 3.14. As long as (31) is satisfied, the algorithm converges to the exact solution whatever the choice of relaxation parameter ρ_k (bounded from above by a sufficiently small $\rho_{\sup}$ and from below by any value $\rho_{\inf} > 0$ to ensure the convergence) and fictitious operators $\widetilde{B}_L$ and $\widetilde{C}_L$ (satisfying assumption 2.5 on $c_{\widetilde{\Omega}}$). Nevertheless, these choices may have a significant influence on the convergence properties of the algorithm. Note that $\widetilde{B}_L$ and $\widetilde{C}_L$ play the role of preconditioners for the iterative algorithm.

Remark 3.15. If $\Lambda = \widetilde{\Lambda}$ and if $B = B_L = \widetilde{B}_L$ and $C = C_L = \widetilde{C}_L$, that means if $n_\Lambda = 0$ and $c_{\widetilde{\Lambda}} = a_\Lambda$ (*i.e.* in the linear case with the same geometry and bilinear form over Λ and $\widetilde{\Lambda}$), then $A(\cdot;\xi)$ is such that $A(U;\xi) - A(V;\xi) = U - V$ for all $U, V \in \widetilde{\mathcal{U}}_\star$ and $B_{\rho_k}(\cdot;\xi)$ is such that $B_{\rho_k}(U;\xi) - B_{\rho_k}(V;\xi) = (1-\rho_k)(U-V)$ for all $U, V \in \widetilde{\mathcal{U}}_\star$, so that the convergence of the algorithm is achieved if the condition $0 < \rho_{\inf} \leqslant \rho_k \leqslant \rho_{\sup} < 2$ is fulfilled, with a convergence in two iterations for a fixed relaxation parameter $\rho_k = 1$.

3.2.3. *Robustness with respect to approximations*

Let us now consider that some approximations are introduced in the different steps of the global-local iterative algorithm.

Such approximations arise when using finite element approximations of spatial functions (see Section 4.1), approximations of parameter-dependent functions (see Sections 4.2 and 4.3), or when resorting to the use of unconverged iterative solvers for the approximate solution of either global or local problems with a certain prescribed accuracy. For example, the solution of non-linear local problems may be performed by means of classical non-linear solvers, such as Newton-type iterative solvers, leading to approximate local solutions.

We now analyze the sensitivity of the global-local iterative algorithm with respect to these approximations at the different steps of the algorithm. Due to these approximations, the algorithm, initially defined by the unperturbed iterative scheme (27), generates a sequence $\{(U_\varepsilon^k, w_\varepsilon^k, \lambda_\varepsilon^k)\}_{k\in\mathbb{N}}$ defined by the following perturbed iterative scheme:

$$U_\varepsilon^k(\xi) = \rho_k \left(\overline{U}_\varepsilon + \Upsilon_\varepsilon(U_\varepsilon^{k-1}(\xi)) + \Phi_\varepsilon(\lambda_\varepsilon^{k-1}(\xi))\right) + (1-\rho_k)U_\varepsilon^{k-1}(\xi), \tag{32a}$$

$$w_\varepsilon^k(\xi) = \Theta_\varepsilon(U_\varepsilon^k(\xi);\xi), \quad \lambda_\varepsilon^k(\xi) = \Psi_\varepsilon(U_\varepsilon^k(\xi);\xi), \tag{32b}$$

where Υ_ε and Φ_ε (resp. Θ_ε and Ψ_ε) are approximations of linear maps Υ and Φ (resp. non-linear maps Θ and Ψ). Similarly, $\overline{U}_\varepsilon$ represents an approximation of $\overline{U}$. At the global level, the unperturbed global iterate $U^k \in \widetilde{\mathcal{U}}_\star^\Xi$ at iteration k satisfies $U^k(\xi) = B_{\rho_k}(U_\varepsilon^{k-1}(\xi);\xi)$. The approximate (or perturbed) global iterate U_ε^k at iteration k is assumed to belong to $\widetilde{\mathcal{U}}_\star^\Xi$ and satisfies $U_\varepsilon^k(\xi) = B_{\rho_k}^\varepsilon(U_\varepsilon^{k-1}(\xi);\xi)$, where $B_{\rho_k}^\varepsilon(\cdot;\xi)$ denotes an approximation of iteration map $B_{\rho_k}(\cdot;\xi)$ defined by $B_{\rho_k}^\varepsilon(V;\xi) = \rho_k\overline{U}_\varepsilon + V - \rho_k A_\varepsilon(V;\xi)$, with $A_\varepsilon(V;\xi) = V - \Upsilon_\varepsilon(V;\xi) - \Phi_\varepsilon(\Psi_\varepsilon(V;\xi);\xi)$.

We assume that approximations are controlled in a L_μ^p-norm (typically $p=2$ or $p=\infty$), that means the approximation error at iteration k, $U_\varepsilon^k(\xi) - U^k(\xi) = B_{\rho_k}^\varepsilon(U_\varepsilon^{k-1}(\xi);\xi) - B_{\rho_k}(U_\varepsilon^{k-1}(\xi);\xi)$, should satisfy

$$\|U_\varepsilon^k - U^k\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})} \leqslant \varepsilon\|U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})} + \varepsilon^*\|U_\varepsilon^{k-1} - U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})},$$

where ε conveys an absolute error with respect to the solution norm $\|U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})}$, while ε^* conveys an approximation error controlled relatively to the solution error in L_μ^p-norm $\|U_\varepsilon^{k-1} - U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})}$ at previous iteration $k-1$. In practice, ε (resp. ε^*) is related to the user-specified tolerance (prescribed to the iterative solver) for the precision of the residual norms associated with global problem (21) and local problem (24) formulated on the current iterates U^k and (w^k, λ^k) (resp. on the current increments $\delta U^k = U^k - U_\varepsilon^{k-1}$ and $(\delta w^k, \delta\lambda^k) = (w^k - w_\varepsilon^{k-1}, \lambda^k - \lambda_\varepsilon^{k-1})$) (see [43, Section 3.5] for further details). We then provide for a robustness result relative to both types of errors.

Theorem 3.16 (Robustness)**.** *Suppose that the set of iteration maps $\{B_{\rho_k}(\cdot;\xi)\}_{k\geqslant 1}$ is uniformly contractive on $\widetilde{\mathcal{U}}_\star$, that means*

$$\|B_{\rho_k}(V;\xi) - B_{\rho_k}(W;\xi)\|_{\widetilde{\mathcal{U}}} \leqslant \rho_B\|V - W\|_{\widetilde{\mathcal{U}}},$$

for all $V, W \in \widetilde{\mathcal{U}}_\star$, with a contractivity constant $\rho_B < 1$ independent of ξ. Further assume that the set of perturbed iteration maps $\{B_{\rho_k}^\varepsilon(\cdot;\xi)\}_{k\geqslant 1}$ is such that for all V in a δ-neighborhood $\mathcal{V}_\delta$ of the exact global solution U, defined by $\mathcal{V}_\delta = \{V \in \mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}}_\star) : \|V - U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})} < \delta\|U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})}\}$, we have

$$\|B_{\rho_k}^\varepsilon(V) - B_{\rho_k}(V)\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})} \leqslant \varepsilon\|U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})} + \varepsilon^*\|V - U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})},$$

for some given tolerances $0 \leqslant \varepsilon^* < 1 - \rho_B$ and $0 \leqslant \varepsilon \leqslant \delta(1 - \rho_B - \varepsilon^*)$. Then, if the initial iterate $U_\varepsilon^0 = 0 \in \mathcal{V}_\delta$, the approximate sequence $\{U_\varepsilon^k\}_{k\in\mathbb{N}}$ is such that

$$\limsup_{k\to+\infty} \|U_\varepsilon^k - U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})} \leqslant \gamma(\varepsilon,\varepsilon^*)\|U\|_{\mathrm{L}_\mu^p(\Xi;\widetilde{\mathcal{U}})}, \tag{33}$$

with $\gamma(\varepsilon,\varepsilon^*)=\frac{\varepsilon}{1-(\rho_B+\varepsilon^*)}\to 0$ *as* $\varepsilon\to 0$*, and tends to a neighborhood of* U *in* $\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})$ *whose size is proportional to* $\gamma(\varepsilon,\varepsilon^*)$*.*

Proof. See Section 7.11 in Appendix 7. □

Finally, the approximate sequence $\{U^k_\varepsilon\}_{k\in\mathbb{N}}$ (resp. $\{w^k_\varepsilon\}_{k\in\mathbb{N}}$ and $\{\lambda^k_\varepsilon\}_{k\in\mathbb{N}}$) generated by the perturbed iterative scheme (32) converges in $\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})$ (resp. $\mathrm{L}^p_\mu(\Xi;\mathcal{W})$ and $\mathrm{L}^p_\mu(\Xi;\mathcal{M})$) to a neighborhood of the exact solution U (resp. w and λ). Therefore, the proposed global-local iterative algorithm exhibits robustness properties with respect to possible perturbations such as numerical approximations, either controlled with relative precision ε^* or absolute precision ε, which is an essential feature from a numerical point of view.

Remark 3.17. The convergence rate of the perturbed algorithm (32), which is $\rho_B+\varepsilon^*$ (see the proof of Theorem 3.16 in Section 7.11 in Appendix 7), depends on the perturbations with controlled relative precision ε^* but not on perturbations with controlled absolute precision ε.

Remark 3.18. Under the more restrictive assumption that the set of perturbed iteration maps $\{B^\varepsilon_{\rho_k}(\cdot;\xi)\}_{k\geqslant 1}$ is such that for all V in a δ-neighborhood $\mathcal{V}_\delta(\xi)$ of the exact global solution $U(\xi)$, defined by $\mathcal{V}_\delta(\xi)=\{V\in\widetilde{\mathcal{U}}_\star : \|V-U(\xi)\|_{\widetilde{\mathcal{U}}}<\delta\|U(\xi)\|_{\widetilde{\mathcal{U}}}\}$, we have almost surely

$$\|B^\varepsilon_{\rho_k}(V;\xi)-B_{\rho_k}(V;\xi)\|_{\widetilde{\mathcal{U}}}\leqslant\varepsilon\|U(\xi)\|_{\widetilde{\mathcal{U}}}+\varepsilon^*\|V-U(\xi)\|_{\widetilde{\mathcal{U}}},$$

with $0\leqslant\varepsilon^*<1-\rho_B$ and $0\leqslant\varepsilon\leqslant\delta(1-\rho_B-\varepsilon^*)$, then we can prove that the approximate sequence $\{(U^k_\varepsilon(\xi),w^k_\varepsilon(\xi),\lambda^k_\varepsilon(\xi)\}_{k\in\mathbb{N}}$ generated by the perturbed iterative scheme (32) converges almost surely to a neighborhood of the exact solution $(U(\xi),w(\xi),\lambda(\xi))$.

4. Computational aspects

In this section, we address computational aspects related to the proposed global-local iterative algorithm.

4.1. Finite element approximations at spatial level

At the spatial level, we employ a standard Galerkin finite element method by introducing finite-dimensional approximation spaces $\widetilde{\mathcal{U}}_H\subset\widetilde{\mathcal{U}}$ and $\mathcal{W}_h\subset\mathcal{W}$ with dimensions n_U and n_w, respectively. We denote by $\mathcal{T}_H(\widetilde{\Omega})$ (resp. $\mathcal{T}_h(\Lambda)$) the finite element mesh of fictitious domain $\widetilde{\Omega}$ (resp. patch Λ) composed of elements of maximum size H (resp. h). For the sake of simplicity, we further make the following assumptions:

- domain $\widetilde{\Omega}$ and patch Λ are exactly covered by global mesh $\mathcal{T}_H(\widetilde{\Omega})$ and local mesh $\mathcal{T}_h(\Lambda)$, respectively;
- global mesh $\mathcal{T}_H(\widetilde{\Omega})$ is partitioned into two submeshes $\mathcal{T}_H(\Omega\setminus\Lambda)$ and $\mathcal{T}_H(\widetilde{\Lambda})$ (associated with subdomain $\Omega\setminus\Lambda$ and fictitious patch $\widetilde{\Lambda}$, respectively) such that $\mathcal{T}_H(\widetilde{\Omega})=\mathcal{T}_H(\Omega\setminus\Lambda)\cup\mathcal{T}_H(\widetilde{\Lambda})$, that means interface Γ coincides with the intersection of boundaries of both submeshes $\mathcal{T}_H(\Omega\setminus\Lambda)$ and $\mathcal{T}_H(\widetilde{\Lambda})$ and therefore does not cut any element of global mesh $\mathcal{T}_H(\widetilde{\Omega})$.

Both meshes $\mathcal{T}_H(\widetilde{\Omega})$ and $\mathcal{T}_h(\Lambda)$ are *a priori* not compatible at interface Γ, that means they may not match on interface Γ. Note that interface Γ is part of the boundary of meshes $\mathcal{T}_h(\Lambda)$, $\mathcal{T}_H(\Omega\setminus\Lambda)$ and $\mathcal{T}_H(\widetilde{\Lambda})$. We now introduce an approximation space $\mathcal{M}_h\subset\mathcal{M}$ with dimension n_λ. In the general case of non-matching meshes and geometries, where both meshes $\mathcal{T}_H(\widetilde{\Omega})$ and $\mathcal{T}_h(\Lambda)$ do not match and even do not share a common interface Γ, one should pay attention to the construction of a suitable approximation space $\mathcal{M}_h$ of Lagrange multipliers. The interested reader can refer to [68–70] for further information on the construction of appropriate Lagrange multiplier spaces using mortar (non-conforming) finite elements. In our particular case where both meshes $\mathcal{T}_H(\Omega\setminus\Lambda)$ and $\mathcal{T}_h(\Lambda)$ share a common interface Γ, a natural choice consists in taking for $\mathcal{M}_h$ a finite-dimensional subspace of trace space $\mathrm{H}^{1/2}(\Gamma)$, so that $\mathcal{M}_h\subset\mathrm{H}^{1/2}(\Gamma)\subset\mathrm{L}^2(\Gamma)\subset\mathrm{H}^{1/2}(\Gamma)^*=\mathcal{M}$. If interface Γ does not present any boundary, a convenient choice consists in taking the trace of $\mathcal{W}_h$ on interface Γ for the practical construction of $\mathcal{M}_h$. Otherwise, if interface Γ has a boundary, an alternative choice consists in taking

a subspace of the trace of $\mathcal{W}_h$ on interface Γ (see [69, 70]). The interested reader can refer to [35, 68, 69] for details about the properties of trace spaces and mortar projection operators.

For a given Hilbert space H (possibly dependent on ξ), we denote by H_n a finite element approximation subspace of H spanned by basis functions $\{\varphi_i\}_{i\in\mathcal{I}}$ and with dimension $n = \#\mathcal{I}$. A function $v \in H_n$ can then be identified with a vector $\mathbf{v} = (v_i)_{i\in\mathcal{I}} \in \mathbb{R}^n$ such that $v = \sum_{i\in\mathcal{I}} v_i\varphi_i$. Similarly, an element $v \in (H_n)^\Xi$ can be identified with a random vector $\mathbf{v} = (v_i)_{i\in\mathcal{I}} \in (\mathbb{R}^n)^\Xi$ such that $v(\xi) = \sum_{i\in\mathcal{I}} v_i(\xi)\varphi_i$. For the bilinear forms $c_\mathcal{O}$ and $a_\mathcal{O}$, semi-linear form $n_\mathcal{O}$ and linear form $\ell_\mathcal{O}$, we introduce the finite element matrices $\mathbf{C}_\mathcal{O}$ and $\mathbf{A}_\mathcal{O}(\xi)$, discretized random non-linear map $\mathbf{N}_\mathcal{O}(\cdot;\xi)$ and finite element random vector $\mathbf{l}_\mathcal{O}(\xi)$, respectively defined for a subdomain $\mathcal{O} \subset \widetilde{\Omega}$ by

$$\begin{aligned} c_\mathcal{O}(u,v) &= \mathbf{u}^T\mathbf{C}_\mathcal{O}\mathbf{v}, \quad a_\mathcal{O}(u,v;\xi) &&= \mathbf{u}^T\mathbf{A}_\mathcal{O}(\xi)\mathbf{v}, \\ n_\mathcal{O}(u,v;\xi) &= \mathbf{u}^T\mathbf{N}_\mathcal{O}(\mathbf{v};\xi), \quad \ell_\mathcal{O}(v;\xi) &&= \mathbf{v}^T\mathbf{l}_\mathcal{O}(\xi). \end{aligned}$$

As the coupling bilinear form b_Γ is defined on the two distinct subspaces $\mathcal{M}_h\times\widetilde{\mathcal{U}}_H$ and $\mathcal{M}_h\times\mathcal{W}_h$, we introduce two finite element matrices $\widetilde{\mathbf{B}}_\Gamma$ and $\mathbf{B}_\Gamma$ defined by

$$b_\Gamma(\lambda,v) = \mathbf{v}^T\widetilde{\mathbf{B}}_\Gamma\boldsymbol{\lambda} \text{ for } v \in \widetilde{\mathcal{U}}_H, \quad \text{and} \quad b_\Gamma(\lambda,v) = \mathbf{v}^T\mathbf{B}_\Gamma\boldsymbol{\lambda} \text{ for } v \in \mathcal{W}_h.$$

In the discrete setting, an approximation of global problem (21) reads: find $\widehat{U}^k \in \widetilde{\mathcal{U}}_H^\Xi$ satisfying (21) for all $\delta U \in \widetilde{\mathcal{U}}_H$. In an algebraic setting, it boils down to solving the following system of linear algebraic equations:

$$\mathbf{C}_{\widetilde{\Omega}}\widehat{\mathbf{U}}^k(\xi) = \mathbf{C}_{\widetilde{\Lambda}}\mathbf{U}^{k-1}(\xi) - \widetilde{\mathbf{B}}_\Gamma\boldsymbol{\lambda}^{k-1}(\xi) + \mathbf{l}_{\Omega\setminus\Lambda}. \tag{34}$$

An approximation of local problem (24) reads: find $(w^k,\lambda^k) \in \mathcal{W}_h^\Xi\times\mathcal{M}_h^\Xi$ such that it satisfies almost surely (24) for all $(\delta w,\delta\lambda) \in \mathcal{W}_h\times\mathcal{M}_h$. In an algebraic setting, it comes down to solving the following system of non-linear algebraic equations:

$$\mathbf{A}_\Lambda(\xi)\mathbf{w}^k(\xi) + \mathbf{N}_\Lambda(\mathbf{w}^k(\xi);\xi) - \mathbf{B}_\Gamma\boldsymbol{\lambda}^k(\xi) = \mathbf{l}_\Lambda(\xi), \tag{35a}$$

$$\mathbf{B}_\Gamma^T\mathbf{w}^k(\xi) = \widetilde{\mathbf{B}}_\Gamma^T\mathbf{U}^k(\xi). \tag{35b}$$

Remark 4.1. The convergence properties of the algorithm may be affected by the choice of spatial approximation spaces. The discretization errors can be viewed as additional perturbations occurring at both global and local steps of the algorithm. The impact of these perturbations on the behavior of the algorithm is addressed through Theorem 3.16.

4.2. **Approximations at stochastic level**

At the stochastic level, we introduce a basis $\{\psi_\alpha\}_{\alpha\in\mathcal{F}}$ of $\mathrm{L}^2_\mu(\Xi)$ (typically a polynomial basis) and we consider approximation spaces $\mathcal{S}_\mathcal{A} = \mathrm{span}\{\psi_\alpha\}_{\alpha\in\mathcal{A}}$, where $\mathcal{A}$ is a finite subset of $\mathcal{F} = \mathbb{N}^m$ which is a partially ordered set such that for $\alpha,\beta \in \mathcal{F}$, $\alpha \leqslant \beta \iff \alpha_i \leqslant \beta_i$ for all $i \in \{1,\dots,m\}$. Then a function $v = \sum_{\alpha\in\mathcal{A}} v_\alpha\psi_\alpha \in \mathcal{S}_\mathcal{A}$ is identified with the vector of its coefficients $(v_\alpha)_{\alpha\in\mathcal{A}} \in \mathbb{R}^{\#\mathcal{A}}$ on the basis of $\mathcal{S}_\mathcal{A}$.

At the global level, suppose that finite element random vectors associated with U^{k-1} and λ^{k-1} are respectively given by $\mathbf{U}^{k-1}(\xi) = \sum_{\alpha\in\mathcal{A}} \mathbf{U}_\alpha^{k-1}\psi_\alpha(\xi)$ and $\boldsymbol{\lambda}^{k-1}(\xi) = \sum_{\alpha\in\mathcal{A}} \boldsymbol{\lambda}_\alpha^{k-1}\psi_\alpha(\xi)$. Then, finite element random vectors associated with $\widehat{U}^k$ and U^k admit the expansions $\widehat{\mathbf{U}}^k(\xi) = \sum_{\alpha\in\mathcal{A}} \widehat{\mathbf{U}}_\alpha^k\psi_\alpha(\xi)$ and $\mathbf{U}^k(\xi) = \sum_{\alpha\in\mathcal{A}} \mathbf{U}_\alpha^k\psi_\alpha(\xi)$, respectively, with

$$\widehat{\mathbf{U}}_\alpha^k = \mathbf{C}_{\widetilde{\Omega}}^{-1}\left(\mathbf{C}_{\widetilde{\Lambda}}\mathbf{U}_\alpha^{k-1} - \widetilde{\mathbf{B}}_\Gamma\boldsymbol{\lambda}_\alpha^{k-1} + \mathbf{l}_{\Omega\setminus\Lambda}\mathbb{E}(\psi_\alpha(\xi))\right) \quad \text{and} \quad \mathbf{U}_\alpha^k = \rho_k\widehat{\mathbf{U}}_\alpha^k + (1-\rho_k)\mathbf{U}_\alpha^{k-1}.$$

It is worthy noticing that since $\widetilde{B}_L$ and $\widetilde{C}_L$ are chosen deterministic on $\widetilde{\Lambda}$ (see Remark 3.1), then $\mathbf{C}_{\widetilde{\Omega}}$ and $\mathbf{C}_{\widetilde{\Lambda}}$ are independent of ξ, and the set of expansion coefficients $\{\widehat{\mathbf{U}}_\alpha^k\}_{\alpha\in\mathcal{A}}$ are simply obtained by solving a system of only $\#\mathcal{A}$ uncoupled linear algebraic equations with the same deterministic global finite element matrix $\mathbf{C}_{\widetilde{\Omega}}$, which can be factorized only once for all at initialization of the iterative procedure. The solution of such uncoupled global problems can be performed in parallel using traditional solvers available in standard deterministic finite element codes.

Remark 4.2. Convergence acceleration techniques based on Quasi-Newton or Newton update formulas have been proposed in [32, 71] within the deterministic framework and rely on successive corrections of the global finite element matrix $\mathbf{C}_{\widetilde{\Omega}}$ at each global step of the iterative procedure in order to improve the convergence rate of the algorithm. In the present stochastic framework, this would yield to a parameter-dependent matrix $\mathbf{C}_{\widetilde{\Omega}}$, unless using deterministic approximations of the successive corrections of $\mathbf{C}_{\widetilde{\Omega}}$.

At the local level, approximations of finite element random vectors $\mathbf{w}^k(\xi)$ and $\boldsymbol{\lambda}^k(\xi)$ associated with local iterates w^k and λ^k, respectively, are searched under the form $\mathbf{w}^k(\xi) \approx \sum_{\alpha\in\mathcal{A}} \mathbf{w}_\alpha^k \psi_\alpha(\xi)$ and $\boldsymbol{\lambda}^k(\xi) \approx \sum_{\alpha\in\mathcal{A}} \boldsymbol{\lambda}_\alpha^k \psi_\alpha(\xi)$. The determination of these approximations through Galerkin projection methods [1, 72] requires the solution of a large system of $\#\mathcal{A}$ coupled non-linear algebraic equations whose computational cost and memory storage requirements may be prohibitive and whose implementation may be cumbersome as it usually requires a modification (or at least an adaptation) of existing deterministic codes. Note however that non intrusive (or weakly intrusive) implementations of Galerkin methods can be introduced [73, 74].

Here, for computing approximations of local iterates, we rather rely on an adaptive least-squares method which uses evaluations of the solution of (24) at some samples $\{\xi^l\}_{l=1}^N$ of random variables ξ. These evaluations are obtained by N calls to an existing non-linear deterministic solver, *i.e.* without requiring any modification of the underlying deterministic computer code. For computing the solution $(\mathbf{w}^k(\xi^l), \boldsymbol{\lambda}^k(\xi^l))$ of (35) for $\xi = \xi^l$, we employ a Newton-type iterative algorithm with some prescribed tolerance. Note that the resulting numerical error can be viewed as an additional perturbation occurring at each local step of the iterative algorithm, whose impact on the behavior of the algorithm is analyzed in Section 3.2.3. The adaptive least-squares method is described in Section 4.3.

4.3. Adaptive least-squares method for sparse polynomial approximation

Here we describe an adaptive least-squares method for sparse approximation of a random vector $\mathbf{u} \in \mathbb{R}^n \otimes \mathrm{L}^2_\mu(\Xi)$. We assume $\Xi \subset \mathbb{R}^m$ $(m < \infty)$ and we consider an orthonormal tensor product basis $\{\psi_\alpha(\xi) = \prod_{i=1}^m \psi_{\alpha_i}^{(i)}(\xi_i)\}_{\alpha\in\mathcal{F}}$ of $\mathrm{L}^2_\mu(\Xi)$, where $\psi_k^{(i)}$ is a univariate polynomial of degree k. For a given subset $\mathcal{A} \subset \mathcal{F}$, we define the corresponding polynomial space $\mathcal{S}_\mathcal{A} = \mathrm{span}\{\psi_\alpha\}_{\alpha\in\mathcal{A}}$. A subset $\mathcal{A}$ is called monotone (or lower or downward closed) if it is such that $(\beta \in \mathcal{A} \quad \text{and} \quad \alpha \leqslant \beta) \implies \alpha \in \mathcal{A}$. If $\mathcal{A}$ is monotone, the subspace $\mathcal{S}_\mathcal{A}$ coincides with the polynomial space $\mathbb{P}_\mathcal{A} = \mathrm{span}\{\xi_1^{\alpha_1} \dots \xi_m^{\alpha_m} : \alpha \in \mathcal{A}\}$ whatever the choice of univariate polynomial bases. Note that univariate polynomial bases could be replaced by other hierarchical bases (such as wavelet bases) with which we can expect accurate sparse approximations of the random vector.

4.3.1. *Approximation in a given subspace*

For a given subset $\mathcal{A}$, a least-squares approximation $\mathbf{v}$ of $\mathbf{u}$ in $\mathbb{R}^n \otimes \mathcal{S}_\mathcal{A}$ can be written as $\mathbf{v}(\xi) = \sum_{\alpha\in\mathcal{A}} \mathbf{v}_\alpha \psi_\alpha(\xi)$, where the set of coefficients $\mathbf{V} = (\mathbf{v}_\alpha)_{\alpha\in\mathcal{A}} \in \mathbb{R}^{n\times\#\mathcal{A}}$ is solution of

$$\min_{(\mathbf{v}_\alpha)_{\alpha\in\mathcal{A}}} \sum_{l=1}^N \|\mathbf{u}(\xi^l) - \sum_{\alpha\in\mathcal{A}} \mathbf{v}_\alpha \psi_\alpha(\xi^l)\|_2^2. \tag{36}$$

Assuming $N \geqslant \#\mathcal{A}$ and $\boldsymbol{\Psi}^T\boldsymbol{\Psi}$ invertible, this yields $\mathbf{V}^T = (\boldsymbol{\Psi}^T\boldsymbol{\Psi})^{-1}\boldsymbol{\Psi}^T\mathbf{Y}$, where $\boldsymbol{\Psi} = (\psi_\alpha(\xi^l))_{1\leqslant l\leqslant N, \alpha\in\mathcal{A}} \in \mathbb{R}^{N\times\#\mathcal{A}}$ and $\mathbf{Y} = (\mathbf{u}(\xi^l))_{1\leqslant l\leqslant N} \in \mathbb{R}^{N\times n}$. The stability of the least-squares approximation is related to the properties of the random matrix $\boldsymbol{\Psi}^T\boldsymbol{\Psi}$. Some theoretical results can be found in [75] and the references therein.

In practice, for a given set $\mathcal{A}$, the stability of the least-squares approximation can be improved by increasing the number of samples.

The approximation error can be estimated *a posteriori* using cross-validation techniques which are classical statistical methods for computing error estimates based on a random partitioning of the available sample set into two subsets, the training set (or learning set) and the test set (or validation set). In the k-fold cross-validation procedure, the sample set $\xi = \{\xi^l\}_{l=1}^N$ is randomly partitioned into k disjoint and complementary sample subsets $\{\xi_s\}_{s=1}^k$ of nearly equal size. Each subset ξ_s is in turn retained as the test set, while the remaining $k-1$ subsets gathered in $\chi_s = \xi \setminus \xi_s$ are used as the training set. An approximation $\mathbf{v}$ is computed independently for each training set χ_s and tested against the corresponding validation set ξ_s in order to assess its accuracy. The cross-validation error is estimated for each of the k training sets χ_s and then averaged over the k sets. Such a cross-validation technique requires k additional calls to the least-squares solver and thus may be computationally demanding. In practice, the vector of cross-validation error estimates $\boldsymbol{\varepsilon} = (\varepsilon_i)_{i\in\mathcal{I}}$ can be directly obtained from the approximate random vector $\mathbf{v} = (v_i)_{i\in\mathcal{I}}$ (computed using the available sample set $\xi = \{\xi^l\}_{l=1}^N$) using the Bartlett matrix inversion formula [76] (a special case of well-known Sherman-Morrison-Woodbury formula) without any additional call to the least-squares solver. The leave-one-out cross-validation procedure is a special case of k-fold cross-validation procedure where the number of folds k is equal to the number of samples N. Note that the k-fold cross-validation technique depends on the chosen partition, contrary to the leave-one-out cross-validation technique. In the present work, we use the fast leave-one-out cross-validation procedure presented in [77] and summarized in Algorithm 7.1 (see Section 7.1 in Appendix 7) to assess the accuracy of $\mathbf{v}$.

4.3.2. *Working set strategy for adaptive approximation*

Now, we introduce a working set strategy for the construction of a sequence of approximation spaces $(\mathcal{S}_{\mathcal{A}_j})_{j\geqslant 1}$, where $(\mathcal{A}_j)_{j\geqslant 1}$ is an increasing sequence of monotone sets. Given $\mathcal{A}_j$, we define $\mathcal{A}_{j+1} = \mathcal{A}_j \cup \mathcal{N}_j$, where $\mathcal{N}_j$ is selected in a set of candidate multi-indices in $\mathcal{F} \setminus \mathcal{A}_j$. A natural approach consists in choosing for $\mathcal{N}_j$ a subset of the margin $\mathcal{M}_j = \mathcal{M}(\mathcal{A}_j)$ of $\mathcal{A}_j$, where the margin of a monotone set $\mathcal{A}$ is defined by

$$\mathcal{M}(\mathcal{A}) = \{\alpha \notin \mathcal{A} : \exists i \in \{1,\dots,m\} \text{ such that } \alpha_i \neq 0 \implies \alpha - e_i \in \mathcal{A}\},$$

where $(e_i)_j = \delta_{ij}$ is the Kronecker delta for $i,j \in \mathbb{N}^*$. A strategy for the selection of $\mathcal{N}_j$ (referred to as bulk search strategy in [50]) consists in computing a least-squares approximation $\mathbf{v}(\xi) = \sum_{\alpha\in\mathcal{A}_j\cup\mathcal{M}_j} \mathbf{v}_\alpha \psi_\alpha(\xi)$ associated with the augmented approximation space $\mathcal{S}_{\mathcal{A}_j\cup\mathcal{M}_j}$, and then in selecting a subset $\mathcal{N}_j$ such that $e(\mathcal{N}_j) \geqslant \theta e(\mathcal{M}_j)$, where $\theta \in [0\,,1]$ is a parameter and where for a given set $\mathcal{N}$, $e(\mathcal{N}) = \sum_{\alpha\in\mathcal{N}} \|\mathbf{v}_\alpha\|_2^2$ corresponds to the contribution of coefficients $(\mathbf{v}_\alpha)_{\alpha\in\mathcal{N}}$ to the L^2_μ-norm of $\mathbf{v}$. Note that the construction of an optimal (smallest) monotone subset $\mathcal{N}_j$ in the margin $\mathcal{M}_j$ of $\mathcal{A}_j$ by a fast algorithm is still an open question.

Remark 4.3. A practical choice for constructing a monotone set $\mathcal{N}_j$ is to consider the smallest subset in the margin $\mathcal{M}_j$ such that $e(\mathcal{N}_j) \geqslant \theta e(\mathcal{M}_j)$ and which contains the multi-indices α corresponding to the largest elements in the monotone envelope[1] $(\mathfrak{v}_\alpha)_{\alpha\in\mathcal{M}_j}$ of the bounded sequence $(\|\mathbf{v}_\alpha\|_2)_{\alpha\in\mathcal{M}_j}$.

Also, as the cardinality of the margin $\mathcal{M}(\mathcal{A}_j)$ may become prohibitively large in high dimension m, an alternative strategy consists in considering for $\mathcal{M}_j$ the reduced margin $\mathcal{M}_{\mathrm{red}}(\mathcal{A}_j)$ of $\mathcal{A}_j$, where

$$\mathcal{M}_{\mathrm{red}}(\mathcal{A}) = \{\alpha \notin \mathcal{A} : \forall i \in \{1,\dots,m\} \text{ such that } \alpha_i \neq 0 \implies \alpha - e_i \in \mathcal{A}\}.$$

The additional set $\mathcal{N}_j$ is then defined as the smallest non-empty subset of the reduced margin $\mathcal{M}_j$ of $\mathcal{A}_j$ such that $e(\mathcal{N}_j) \geqslant \theta e(\mathcal{M}_j)$, which is a monotone set by construction. Therefore, $\mathcal{A}_{j+1} = \mathcal{A}_j \cup \mathcal{N}_j$, as a union of monotone sets, is a monotone set. For $\theta = 1$, the selected subset $\mathcal{N}_j = \{\alpha \in \mathcal{M}_j : \|\mathbf{v}_\alpha\|_2 \neq 0\}$. For $\theta = 0$, $\mathcal{N}_j$ is one arbitrary element of $\{\alpha \in \mathcal{M}_j : \alpha = \arg\max_{\alpha\in\mathcal{M}_j} \|\mathbf{v}_\alpha\|_2\}$, and the strategy corresponds to the largest

[1]For a given set $\mathcal{A}$, the monotone envelope (also called monotone majorant) $(\mathfrak{v}_\alpha)_{\alpha\in\mathcal{A}}$ of a bounded sequence $(\|\mathbf{v}_\alpha\|_2)_{\alpha\in\mathcal{A}}$ is defined by $\mathfrak{v}_\alpha = \max_{\beta\in\mathcal{A},\beta\geqslant\alpha} \|\mathbf{v}_\beta\|_2$ for $\alpha \in \mathcal{A}$.

neighbor strategy proposed in [50]. In the numerical experiments, we will consider the strategy with a parameter $\theta = 0.5$.

4.3.3. *Adaptive strategy*

In order to reach a desired accuracy, we finally propose an algorithm with adaptive random sampling and an adaptive selection of the approximation space $\mathcal{S}_\mathcal{A}$ with the working set strategy presented above. The algorithm is summarized in Algorithm 7.2 (see Section 7.2 in Appendix 7). The convergence, stagnation and overfitting criteria in Algorithm 7.2 are respectively defined by

$$\|\boldsymbol{\varepsilon}\|_2 \leqslant \varepsilon_{\text{cv}}, \quad \frac{\|\boldsymbol{\varepsilon} - \boldsymbol{\varepsilon}^{\text{prev}}\|_2}{\|\boldsymbol{\varepsilon}\|_2} \leqslant \varepsilon_{\text{stagn}} \quad \text{and} \quad \frac{\|\boldsymbol{\varepsilon}\|_2}{\|\boldsymbol{\varepsilon}^{\text{prev}}\|_2} > 1 + \varepsilon_{\text{overfit}},$$

where $\boldsymbol{\varepsilon}$ (resp. $\boldsymbol{\varepsilon}^{\text{prev}}$) is the vector of cross-validation error estimates computed at current (resp. previous) iteration, and ε_{cv} (resp. $\varepsilon_{\text{stagn}}$ and $\varepsilon_{\text{overfit}}$) is the convergence (resp. stagnation and overfitting) threshold.

For computing sparse approximations of the solution $(\mathbf{w}^k(\xi), \boldsymbol{\lambda}^k(\xi))$ of non-linear local problem (35) at iteration k of the algorithm, we apply the adaptive least-squares method to $\mathbf{u}(\xi) = (\mathbf{w}^k(\xi), \boldsymbol{\lambda}^k(\xi))$. In the adaptive sparse least-squares solver presented in Algorithm 7.2, the adaptive random sampling step is performed by computing N_{add} additional samples $(\mathbf{w}^k(\xi^{N+l}), \boldsymbol{\lambda}^k(\xi^{N+l}))_{1 \leqslant l \leqslant N_{\text{add}}}$ of random vectors $(\mathbf{w}^k(\xi), \boldsymbol{\lambda}^k(\xi))$ (thus solving N_{add} deterministic non-linear local problems) until the vectors of cross-validation error estimates $\boldsymbol{\varepsilon}^{\mathbf{w}}$ and $\boldsymbol{\varepsilon}^{\boldsymbol{\lambda}}$ for both random vectors $\mathbf{w}^k(\xi)$ and $\boldsymbol{\lambda}^k(\xi)$ satisfy a stagnation criterium. Then, the working set strategy is applied to $\mathbf{w}^k(\xi)$ and $\boldsymbol{\lambda}^k(\xi)$ separately to construct a sparse polynomial approximation space $\mathcal{S}_\mathcal{A}$ dedicated to each random vector.

Remark 4.4. Following Remark 3.4, in the case of a patch Λ containing geometrical variabilities, remeshing techniques could also be used in conjunction with sampling-based approaches. Indeed, the algorithm requires the computation of local iterates λ^k (defined on a deterministic interface Γ) but not of local iterates w^k (defined on the uncertain domain $\Lambda(\xi)$). An approximation of λ^k can therefore be computed from samples of the local iterate $\lambda^k(\xi)$, where each new sample involves a specific mesh of $\Lambda(\xi)$. After convergence of the algorithm (at final iteration k), samples of the local solution $w^k(\xi)$ can be obtained by using remeshing techniques for each sample, from which we can deduce samples of quantities of interest expressed as functionals of $w^k(\xi)$. An explicit approximation of w^k in terms of ξ can also be obtained by using approaches based on reformulations of the local problem on a fixed deterministic domain (such as the random mapping techniques or the fictitious domain approaches briefly described in Remark 3.4).

4.4. **Relaxation step**

The relaxation step may affect the convergence rate of the iterative algorithm as it can be interpreted as a line-search step of a non-linear solver.

The simplest method is to choose a fixed relaxation parameter ρ throughout iterations. A large relaxation parameter may speed up the convergence but can lead to a divergence of the algorithm, while a small relaxation parameter ensures the convergence but triggers more iterations in return. The computation of an optimal relaxation parameter leading to an optimal convergence rate of the algorithm is not obvious in the non-linear framework. Even in the linear case, the optimal fixed value of relaxation parameter ρ is problem-dependent and not known *a priori*.

The Aitken's Delta-Squared method [78, 79] is a convergence acceleration technique which allows improving the current solution by using information gained at two previous iterations. The current global iterate U^k is then obtained from the two pairs $(\widehat{U}^k, U^{k-1})$ and $(\widehat{U}^{k-1}, U^{k-2})$ and defined as

$$U^k(\xi) = \widehat{U}^k(\xi) - \frac{\langle \delta^k(\xi) - \delta^{k-1}(\xi), \widehat{U}^k(\xi) - \widehat{U}^{k-1}(\xi) \rangle_{\widetilde{\mathcal{U}}}}{\|\delta^k(\xi) - \delta^{k-1}(\xi)\|^2_{\widetilde{\mathcal{U}}}} \delta^k(\xi),$$

where $\delta^k(\xi) = \widehat{U}^k(\xi) - U^{k-1}(\xi)$ is the difference between the current global solution $\widehat{U}^k(\xi)$ and the previous global iterate $U^{k-1}(\xi)$. Then, using (23), the relaxation parameter ρ_k is dynamically updated and defined by

$$\rho_k = -\rho_{k-1} \frac{\langle \delta^k(\xi) - \delta^{k-1}(\xi), \delta^{k-1}(\xi) \rangle_{\widetilde{\mathcal{U}}}}{\|\delta^k(\xi) - \delta^{k-1}(\xi)\|^2_{\widetilde{\mathcal{U}}}}. \tag{37}$$

As the Aitken's recursive formula (37) requires two iterations of the algorithm, the first two values ρ_1 and ρ_2 are commonly set to 1. For the subsequent iterations, the relaxation parameter ρ_k can de defined as

$$\rho_k = T_{[\rho_{\inf}, \rho_{\sup}]} \left(-\rho_{k-1} \frac{\langle \delta^k(\xi) - \delta^{k-1}(\xi), \delta^{k-1}(\xi) \rangle_{\widetilde{\mathcal{U}}}}{\|\delta^k(\xi) - \delta^{k-1}(\xi)\|^2_{\widetilde{\mathcal{U}}}} \right)$$

where $T_{[\rho_{\inf}, \rho_{\sup}]}(\rho)$ is the projection of ρ on the interval $[\rho_{\inf}, \rho_{\sup}]$, which allows to ensure the convergence of the algorithm (see convergence condition (31)).

Such a convergence acceleration technique is very simple to implement and computationally cheap. Also, the Aitken's acceleration method has been successfully applied to relaxation-based fixed-point algorithms in the context of fluid-structure interaction [80] and multiscale coupling [71, 81] problems. It has been proved to be particularly efficient with good convergence properties at low cost, compared with other relaxation methods such as the steepest descent method.

5. Numerical results

In order to demonstrate the efficiency and the robustness of the proposed method, we present different numerical experiments for a stationary non-linear diffusion-reaction equation defined on a deterministic rectangular (two-dimensional) domain $\Omega = (0, 2) \times (0, 16) \subset \mathbb{R}^2$. This equation is complemented with deterministic homogeneous Dirichlet boundary conditions $u = 0$ applied on the entire boundary $\Gamma_D = \partial\Omega$. A deterministic volumetric source term $f = 1$ is imposed on the whole domain Ω. The only sources of uncertainties come from diffusion coefficient $K(x, \xi)$ and reaction parameter $R(x, \xi)$ which are input random fields depending on a set of random variables $\xi \in \Xi$. The variabilities are assumed to be confined in $Q = 8$ patches $\{\Lambda_q\}_{q=1}^Q$ distributed along the y (vertical) axis. Each patch Λ_q is a square subdomain $\Lambda_q = (0.5, 1.5) \times (2q - 1.5, 2q - 0.5)$. None of the patches present geometrical details. Domain Ω and the set of Q patches $\{\Lambda_q\}_{q=1}^Q$ are illustrated on Figure 3a.

The solution u satisfies almost surely

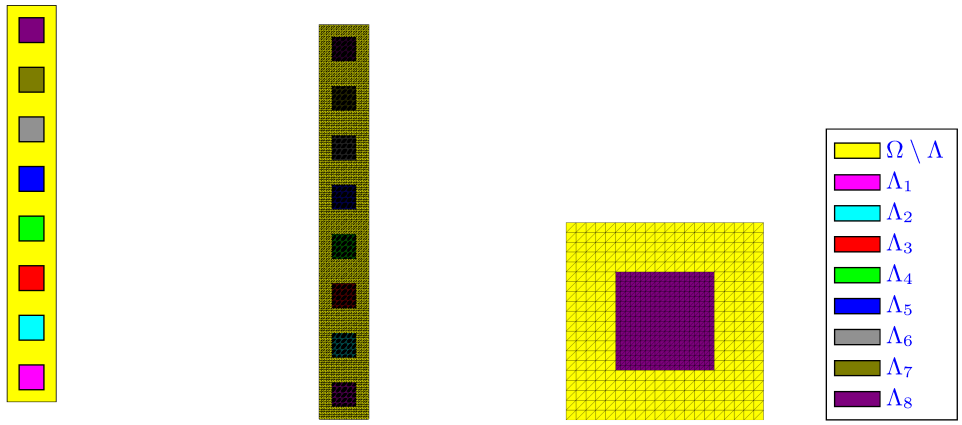
$$-\nabla \cdot (K(x, \xi) \nabla u) + R(x, \xi) u^3 = f \quad \text{on } \Omega, \quad u = 0 \quad \text{on } \Gamma_D = \partial\Omega,$$

where random diffusion coefficient K and random reaction parameter R are such that

$$K(x, \xi) = \begin{cases} K_0 = 1 & \text{for } x \in \Omega \setminus \Lambda \\ K_q(x, \xi_{2q-1}) = 1 + \gamma_q \xi_{2q-1} \chi_q(x) & \text{for } x \in \Lambda_q, \text{ for all } q \in \{1, \dots, Q\} \end{cases},$$

$$R(x, \xi) = \begin{cases} 0 & \text{for } x \in \Omega \setminus \Lambda \\ R_q(x, \xi_{2q}) = \gamma_q \xi_{2q} \chi_q(x) & \text{for } x \in \Lambda_q, \text{ for all } q \in \{1, \dots, Q\} \end{cases},$$

with $\chi_q(x)$ the indicator function of subdomain $\Lambda_q^\star = (0.75, 1.25) \times (2q - 1.25, 2q - 0.75) \subset \Lambda_q$ for all $q \in \{1, \dots, Q\}$, and where the weights γ_q are real coefficients in $(0, 1)$ whose values define a level of uncertainty in the patch Λ_q. We consider two different situations: (i) an isotropic case, for which all weights $\gamma_q = 1$; (ii) an anisotropic case, for which the weights $\gamma_q = 1 - 0.1(q + 1)$. Random diffusion coefficient K and reaction parameter R are parametrized by a set $\xi = (\xi_i)_{i=1}^m$ of $m = 2Q = 16$ real-valued random variables ξ_i assumed to be mutually independent and uniformly distributed on $(0, 1)$. The parameter space is then the hypercube $\Xi = (0, 1)^m \subset \mathbb{R}^m$



(A) Domain and patches (B) Nested triangulations (C) Zoom around patch Λ_1

FIGURE 3. (a) Domain Ω partitioned into $Q = 8$ patches $\{\Lambda_q\}_{q=1}^{Q}$ and the complementary subdomain $\Omega \setminus \Lambda$ with $\Lambda = \bigcup_{q=1}^{Q} \Lambda_q$, and (b) associated nested global and local finite element meshes $\mathcal{T}_h(\Lambda)$ and $\mathcal{T}_H(\Omega \setminus \Lambda)$ with (c) zoom around patch Λ_1

.

endowed with the uniform probability measure. Each patch Λ_q is characterized by a random diffusion coefficient K_q (parametrized by ξ_{2q-1}) and a random reaction parameter R_q (parametrized by ξ_{2q}). The material properties of patch Λ_q therefore depends on 2 real-valued random variables ξ_{2q-1} and ξ_{2q}. The ranges of variations of K_q and R_q in each patch Λ_q are respectively $(1, 2)$ and $(0, 1)$ for the isotropic case, and $(1, 1 + \gamma_q)$ and $(0, \gamma_q)$ for the anisotropic case. In this example, all uncertain material parameters K_q and R_q depend on the random variables ξ in an affine manner, so do the bilinear forms a_{Λ_q} and semi-linear forms n_{Λ_q} for all $q \in \{1, \ldots, Q\}$. As the Hilbert space $\mathcal{V}$ is assumed to be deterministic, the solution $u \in \mathrm{L}^p_\mu(\Xi; \mathcal{V})$ belongs to the tensor product vector space $\mathcal{V} \otimes \mathrm{L}^p_\mu(\Xi)$. Given that Dirichlet boundary conditions are applied on the whole boundary $\partial\Omega$, and the source term f is deterministic on domain Ω, the linear form ℓ_Ω is independent of ξ. Lastly, as domain Ω does not contain any geometrical defects, $\widetilde{\Omega} = \Omega$ in this example.

5.1. **Approximation spaces**

At the spatial level, we introduce nested finite element approximation spaces $\widetilde{\mathcal{U}}_H \subset \widetilde{\mathcal{U}}$ and $\mathcal{W}_h^q \subset \mathcal{W}^q$ for all $q \in \{1, \ldots, Q\}$ (see Figures 3b and 3c). The coarse global mesh $\mathcal{T}_H(\widetilde{\Omega})$ is a regular triangulation of fictitious domain $\widetilde{\Omega}$ which is composed of 3-nodes linear triangular elements with uniform element size $H = 0.1$. It thus comprises 3 381 nodes and 6 400 elements. For every $q \in \{1, \ldots, Q\}$, the fine local mesh $\mathcal{T}_h(\Lambda_q)$ is a regular triangulation of patch Λ_q which is composed of 3-nodes linear triangular elements with uniform element size $h_q = 0.05$. It is thus made of 441 nodes and 800 elements. Every fine local mesh $\mathcal{T}_h(\Lambda_q)$ corresponds to a uniform refinement of the corresponding coarse local mesh $\mathcal{T}_H(\widetilde{\Lambda}_q)$. The resulting spatial approximation spaces $\widetilde{\mathcal{U}}_H$ and $\mathcal{W}_h^q$ for all $q \in \{1, \ldots, Q\}$, have dimensions $n_U = \dim(\widetilde{\mathcal{U}}_H) = 3\,381$ and $n_{w_q} = \dim(\mathcal{W}_h^q) = 441$, respectively.

At the stochastic level, we adaptively build a multidimensional polynomial approximation space $\mathcal{S}_\mathcal{A}$ spanned by generalized polynomial chaos basis $\{\psi_\alpha\}_{\alpha \in \mathcal{A}}$ (multidimensional Legendre polynomials) by using the adaptive sparse least-squares solver described in Algorithm 7.2. At each iteration of the iterative algorithm, the linear global problem (21) defined on fictitious domain $\widetilde{\Omega}$ is solved exactly (at the machine precision using a direct solver) as it involves a deterministic operator, while the Q non-linear local problems (24) defined on the Q patches $\{\Lambda_q\}_{q=1}^{Q}$ are solved using the adaptive sampling-based least-squares method described in Section 4.3. For each patch Λ_q, the N_q deterministic non-linear local problems (35) associated with the N_q samples $\{\xi^l\}_{l=1}^{N_q}$

are partially solved using a tangent-Newton iterative algorithm with a prescribed tolerance set to $\epsilon = 10^{-12}$. In our application case, one deterministic non-linear local problem typically requires only few iterations (less than 5) to reach this stopping criterion. All the local computations have been performed in parallel on 32 cores of a single computer node. The sample set $\{\xi^l\}_{l=1}^{N_q}$ and the approximation spaces $\mathcal{S}_{\mathcal{A}_q}$ are sequentially enriched (independently for each patch) in order to control the accuracy of the local approximations (w_q^k, λ_q^k). The stagnation and overfitting thresholds in Algorithm 7.2 are both set to $\varepsilon_{\text{stagn}} = \varepsilon_{\text{overfit}} = 10^{-1}$. An initial sample set of size $N = 1$ is used with a sampling factor $p_{\text{add}} = 0.1$ (percentage of additional samples) and a parameter $\theta = 0.5$ for a good trade-off between computational efficiency and stability of local solutions (w_q^k, λ_q^k).

5.2. **Convergence analysis**

The accuracy of global approximations U^k is measured in L^2_μ-norm with respect to a global reference solution U^{ref} using the relative error indicator $\varepsilon_{\Xi,\Omega\setminus\Lambda}$ defined as

$$\varepsilon_{\Xi,\Omega\setminus\Lambda}(U^k; U^{\text{ref}}) = \frac{\|U^k - U^{\text{ref}}\|_{\mathrm{L}^2_\mu(\Xi;\mathrm{L}^2(\Omega\setminus\Lambda))}}{\|U^{\text{ref}}\|_{\mathrm{L}^2_\mu(\Xi;\mathrm{L}^2(\Omega\setminus\Lambda))}}, \quad \text{with} \quad \|U\|^2_{\mathrm{L}^2_\mu(\Xi;\mathrm{L}^2(\Omega\setminus\Lambda))} = \mathbb{E}(\|U(\xi)\|^2_{\mathrm{L}^2(\Omega\setminus\Lambda)}).$$

The reference solution $(U^{\text{ref}}, w^{\text{ref}}, \lambda^{\text{ref}})$ is obtained by directly solving the full-scale coupled problem (16) using the adaptive sparse least-squares method described in Section 4.3. Following Theorem 3.7, the global reference solution U^{ref} is the restriction to subdomain $\Omega \setminus \Lambda$ of the limit U of the sequence of global iterates U^k. At the spatial level, the global (resp. local) reference solution U^{ref} (resp. $(w_q^{\text{ref}}, \lambda_q^{\text{ref}})$) is discretized using the same finite element mesh as the global (resp. local) approximations U^k (resp. (w_q^k, λ_q^k)). At the stochastic level, the number of samples N^{ref} and the approximation spaces $\mathcal{S}_{\mathcal{A}^{\text{ref}}}$ are controlled by using the leave-one-out cross-validation procedure presented in Algorithm 7.1. The prescribed tolerance for the convergence of Algorithm 7.2 is set to $\varepsilon_{\text{cv}}^{\text{ref}} = 10^{-6}$. The resulting sample size is $N^{\text{ref}} = 795$ for the isotropic case and $N^{\text{ref}} = 491$ for the anisotropic case. The partial polynomial degrees p_i^{ref} (in each random variable ξ_i) and the dimension $\#\mathcal{A}^{\text{ref}}$ of approximation spaces $\mathcal{S}_{\mathcal{A}^{\text{ref}}}$ for global and local reference solutions U^{ref} and $(w_q^{\text{ref}}, \lambda_q^{\text{ref}})$ are reported in Table 1 for both isotropic and anisotropic cases. The local reference solution $(w_q^{\text{ref}}, \lambda_q^{\text{ref}})$ mainly depends on the random variables ξ_{2q-1} and ξ_{2q} associated with the corresponding patch Λ_q, as well as on the random variables confined in the surrounding patches. Note that, in the anisotropic setting, the stronger the variabilities in the material properties within a patch Λ_q are, the more the reference local solution $(w_q^{\text{ref}}, \lambda_q^{\text{ref}})$ is sensitive to the random variables ξ_{2q-1} and ξ_{2q} associated with Λ_q.

We first consider a fixed cross-validation tolerance $\varepsilon_{\text{cv}} = 10^{-3}$ in Algorithm 7.2 for the accuracy of local solutions (w_q^k, λ_q^k) and we study the influence of relaxation parameter ρ_k on the convergence of the global-local iterative algorithm. Figure 4 represents the evolution of relative error indicator $\varepsilon_{\Xi,\Omega\setminus\Lambda}$ with respect to the number of iterations k for different fixed values of relaxation parameter $\rho \in \{0.2, 0.4, 0.6, 0.8, 1, 1.2, 1.4, 1.6, 1.8\}$ and for a relaxation parameter ρ_k dynamically updated through the Aitken's Delta-Squared method presented in Section 4.4 in the isotropic case. As expected, the relaxation parameter has a strong influence on the convergence rate of the iterative algorithm. The Aitken's Delta-Squared acceleration technique provides similar results as those obtained with an optimal fixed relaxation parameter without any additional computational cost. The relative error indicator decreases sharply ($\varepsilon_{\Xi,\Omega\setminus\Lambda} = 5.10^{-5}$ after only $k = 3$ iterations) and reaches a plateau $\varepsilon_{\Xi,\Omega\setminus\Lambda} = 3.10^{-6}$ for $k \geqslant 5$. Similar results can be obtained for the anisotropic case.

We now use the Aitken's dynamic relaxation and we investigate the influence of the prescribed tolerance ε_{cv} for cross-validation in Algorithm 7.2 on the performances of the global-local iterative algorithm in terms of accuracy and computational efficiency. Figure 5 shows the evolution of relative error indicator $\varepsilon_{\Xi,\Omega\setminus\Lambda}$ and computational cost per iteration as functions of the number of iterations k of the global-local iterative algorithm for different cross-validation tolerances $\varepsilon_{\text{cv}} \in \{10^{-2}, 10^{-3}, 10^{-4}, 10^{-5}\}$ for both problems. The iterative algorithm converges quite fast until the relative error indicator $\varepsilon_{\Xi,\Omega\setminus\Lambda}$ stabilizes around a value smaller than the tolerance ε_{cv} imposed to Algorithm 7.2 for cross-validation. The cross-validation threshold ε_{cv} can then be seen as the level

	p_1^{ref}	p_2^{ref}	p_3^{ref}	p_4^{ref}	p_5^{ref}	p_6^{ref}	p_7^{ref}	p_8^{ref}	p_9^{ref}	p_{10}^{ref}	p_{11}^{ref}	p_{12}^{ref}	p_{13}^{ref}	p_{14}^{ref}	p_{15}^{ref}	p_{16}^{ref}	$\#\mathcal{A}^{\mathrm{ref}}$
U^{ref}	7	3	6	4	6	4	6	4	6	4	6	4	6	4	7	3	384
w_1^{ref}	7	3	5	3	3	2	2	2	1	1	0	0	0	0	0	0	85
λ_1^{ref}	8	4	6	4	5	3	3	2	2	1	0	1	0	0	0	0	186
w_2^{ref}	6	3	7	4	5	3	3	2	2	2	1	1	0	0	0	0	142
λ_2^{ref}	6	3	7	4	5	3	4	2	2	2	1	1	0	0	0	0	170
w_3^{ref}	4	2	5	3	7	4	5	3	3	2	2	1	1	1	0	0	150
λ_3^{ref}	5	2	5	3	7	4	5	3	4	3	2	2	1	1	1	0	215
w_4^{ref}	3	1	3	2	5	3	7	4	5	3	3	2	2	2	1	1	168
λ_4^{ref}	3	2	4	3	5	3	7	4	5	3	4	3	3	2	2	1	235
w_5^{ref}	1	1	2	1	3	2	5	3	7	4	5	3	3	2	3	1	167
λ_5^{ref}	2	1	3	2	4	3	5	3	7	4	5	3	4	2	3	2	234
w_6^{ref}	1	0	1	1	2	1	3	2	5	3	7	3	5	3	4	2	149
λ_6^{ref}	1	0	1	1	3	2	4	3	6	3	7	4	6	3	5	2	239
w_7^{ref}	0	0	0	0	0	1	2	2	3	2	5	3	8	4	5	3	164
λ_7^{ref}	0	0	0	0	1	1	2	2	4	2	5	3	7	4	7	3	174
w_8^{ref}	0	0	0	0	0	1	1	1	2	1	3	2	5	3	8	3	91
λ_8^{ref}	0	1	0	0	1	1	2	2	3	2	5	3	6	4	8	4	202

(A) Isotropic case

	p_1^{ref}	p_2^{ref}	p_3^{ref}	p_4^{ref}	p_5^{ref}	p_6^{ref}	p_7^{ref}	p_8^{ref}	p_9^{ref}	p_{10}^{ref}	p_{11}^{ref}	p_{12}^{ref}	p_{13}^{ref}	p_{14}^{ref}	p_{15}^{ref}	p_{16}^{ref}	$\#\mathcal{A}^{\mathrm{ref}}$
U^{ref}	6	3	5	3	5	3	4	3	4	3	4	3	3	2	3	2	173
w_1^{ref}	7	3	4	3	2	2	1	1	0	1	0	0	0	0	0	0	57
λ_1^{ref}	7	3	5	3	4	2	2	2	1	1	0	0	0	0	0	0	129
w_2^{ref}	5	2	6	3	4	3	3	2	1	1	0	1	0	0	0	0	95
λ_2^{ref}	5	3	6	3	4	3	3	2	2	1	1	1	0	0	0	0	115
w_3^{ref}	4	2	4	3	6	3	4	2	2	2	1	1	0	0	0	0	106
λ_3^{ref}	4	2	5	3	6	3	4	3	3	2	1	1	0	1	0	0	127
w_4^{ref}	2	1	3	2	4	2	5	3	3	2	2	2	1	1	0	0	84
λ_4^{ref}	3	1	3	2	4	3	5	3	4	3	2	2	1	1	1	0	112
w_5^{ref}	1	1	1	1	2	2	3	2	5	3	3	2	2	1	1	1	68
λ_5^{ref}	2	1	2	1	3	2	4	3	5	3	3	2	2	2	1	1	94
w_6^{ref}	0	0	0	1	1	1	2	2	3	2	4	3	3	2	2	1	54
λ_6^{ref}	0	0	1	1	2	2	3	2	4	3	5	3	3	2	2	1	84
w_7^{ref}	0	0	0	0	0	0	1	1	2	2	3	2	4	2	2	1	40
λ_7^{ref}	0	0	0	0	1	1	2	1	3	2	3	2	4	3	3	2	55
w_8^{ref}	0	0	0	0	0	0	0	0	1	1	2	1	2	2	3	2	23
λ_8^{ref}	0	0	0	0	0	0	1	1	2	1	3	2	3	2	3	2	41

(B) Anisotropic case

TABLE 1. Partial polynomial degrees p_i^{ref} (in each random variable ξ_i), $i \in \{1, \ldots, 16\}$, and dimension $\#\mathcal{A}^{\mathrm{ref}}$ of approximation spaces $\mathcal{S}_{\mathcal{A}^{\mathrm{ref}}}$ for global and local reference solutions U^{ref} and $(w_q^{\mathrm{ref}}, \lambda_q^{\mathrm{ref}})$, $q \in \{1, \ldots, 8\}$, where values displayed in red correspond to random variables ξ_{2q-1} and ξ_{2q} associated with patch Λ_q

of a perturbation occurring at each local step of the iterative algorithm and having an impact on its convergence

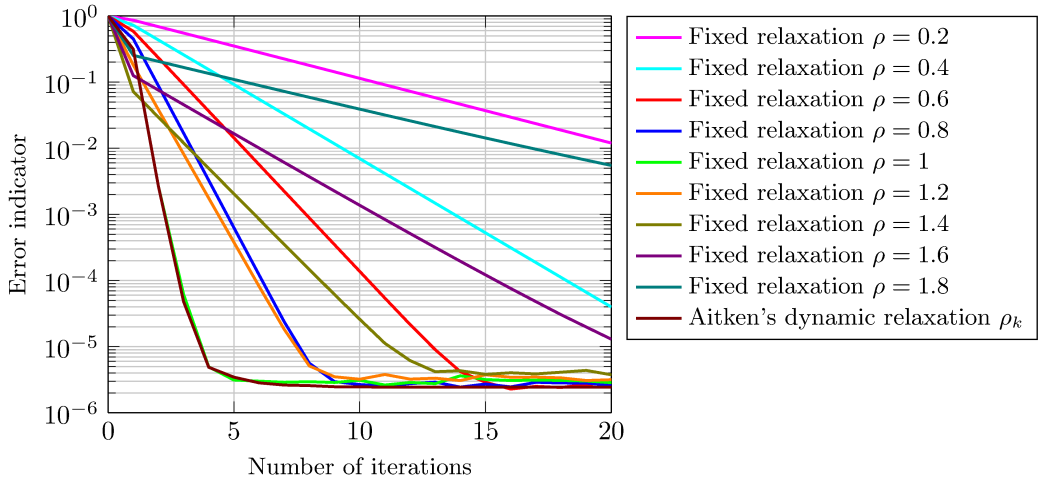
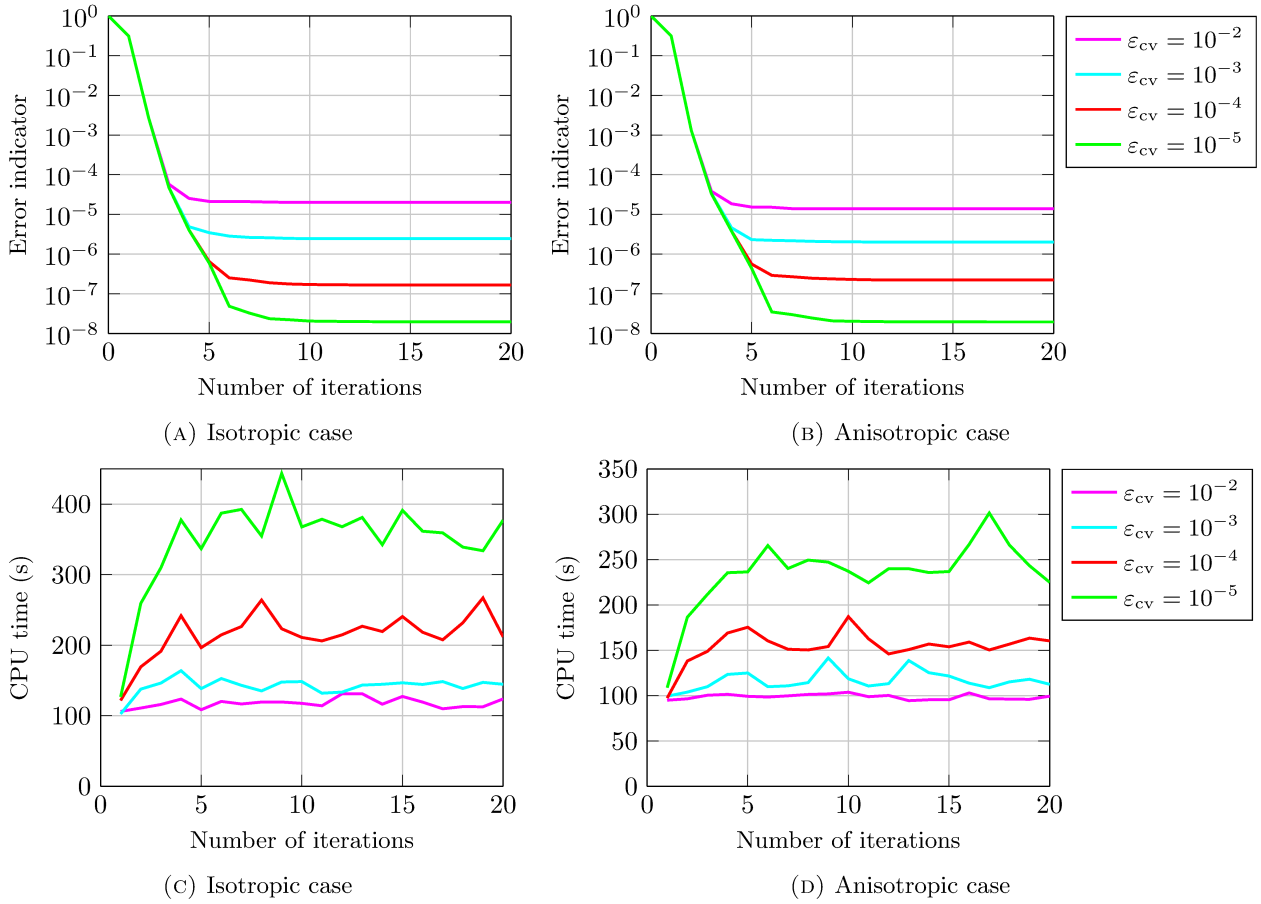


FIGURE 4. Isotropic problem: evolution of error indicator $\varepsilon_{\Xi,\Omega\setminus\Lambda}$ with respect to iteration number k for different fixed relaxation parameters ρ and for Aitken's dynamic relaxation parameter ρ_k

properties. The accuracy of multiscale solution $u = (U, w_1, \ldots, w_Q)$ obtained at convergence of the global-local iterative algorithm is then controlled by the cross-validation tolerance ε_{cv} prescribed to Algorithm 7.2 at each local stage. Note that using a relatively high cross-validation tolerance $\varepsilon_{\text{cv}} = 10^{-2}$ allows to reach a rather small precision $\varepsilon_{\Xi,\Omega\setminus\Lambda} = 6.10^{-5}$ (resp. 4.10^{-5}) after only $k = 3$ iterations for the isotropic (resp. anisotropic) case. Figure 6 displays the evolution of Aitken's dynamic relaxation parameter ρ_k with respect to iteration number k for the aforementioned cross-validation tolerances $\varepsilon_{\text{cv}} = 10^{-2}, \ldots, 10^{-5}$ for both isotropic and anisotropic cases. As mentioned in [80], relaxation parameter ρ_k varies but does not seem to follow a definite pattern during the iterations of the algorithm. In our numerical experiments, ρ_k varies slightly around the value 1 during the first iterations, then decreases toward zero as soon as the relative error indicator $\varepsilon_{\Xi,\Omega\setminus\Lambda}$ stagnates around a certain value depending on the prescribed tolerance ε_{cv} for cross-validation. Figure 7 shows the evolutions of the sample size N_q and the dimension $\#\mathcal{A}_q$ of the approximation space for local solution w_q^k and Lagrange multiplier λ_q^k within patch Λ_q, for $q = 4$, as functions of the number of iterations k for cross-validation threshold values ε_{cv} varying from 10^{-2} to 10^{-5} for both isotropic and anisotropic cases. The number of samples N_q and the dimension $\#\mathcal{A}_q$ of the polynomial spaces increase during the first iterations and then stagnate around a certain value which is higher as the prescribed tolerance ε_{cv} for cross-validation is lower. The sample sizes and the dimensions of approximation spaces are higher for the isotropic case than for the anisotropic case. Note that the dimension of the approximation space obtained for Lagrange multiplier λ_q is higher than the one for local solution w_q.

In order to illustrate the capability of the adaptive least-squares solver given in Algorithm 7.2 to capture sparse high-dimensional polynomial approximations of local solutions, Table 2 shows the partial polynomial degrees p_i with respect to each random variable ξ_i, $i \in \{1, \ldots, m\}$, and the dimension $\#\mathcal{A}$ of approximation spaces $\mathcal{S}_{\mathcal{A}}$ for global and local solutions U and (w^q, λ^q), $q \in \{1, \ldots, Q\}$, obtained at convergence of the global-local iterative algorithm and using a fixed cross-validation tolerance $\varepsilon_{\text{cv}} = 10^{-3}$ for the convergence of Algorithm 7.2. We observe that the use of the adaptive sparse least-squares solver allows to detect sparsity in local solutions (w^q, λ^q). Indeed, Algorithm 7.2 gives local solutions (w^q, λ^q) with a very low effective dimensionality in so far as they are mainly dependent on only few random variables, especially the random variables ξ_{2q-1} and ξ_{2q} associated with patch Λ_q.



(A) Isotropic case (B) Anisotropic case

(C) Isotropic case (D) Anisotropic case

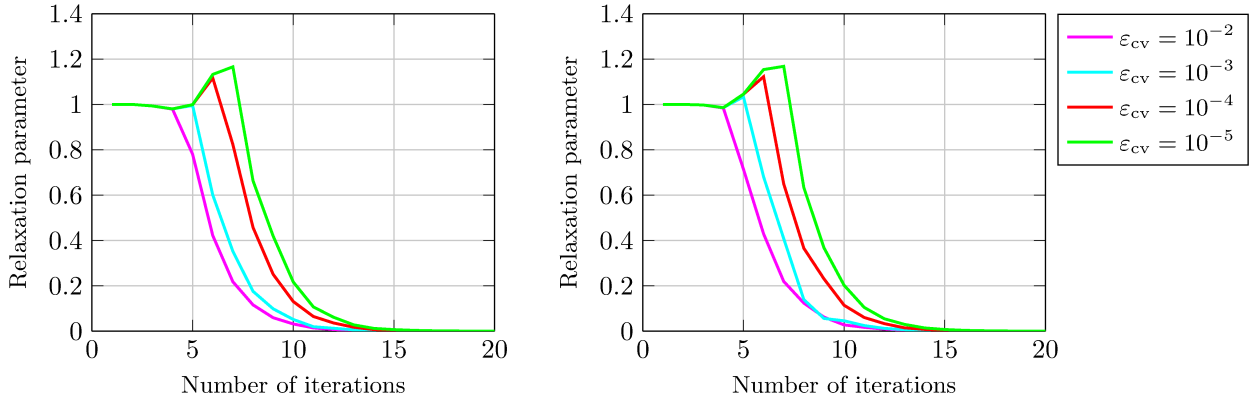
FIGURE 5. Evolutions of (a)-(b) error indicator $\varepsilon_{\Xi,\Omega\setminus\Lambda}$ and (c)-(d) CPU time per iteration with respect to iteration number k for different cross-validation tolerances ε_{cv}

5.3. **Illustration of quantities of interest**

We now look at the effect of input uncertainties in material properties, namely diffusion coefficient K and reaction parameter R, on the variability of the solution. We apply the Aitken's acceleration technique to the relaxation step of the global-local iterative algorithm and we set the cross-validation tolerance to $\varepsilon_{\text{cv}} = 10^{-3}$ in Algorithm 7.2. We then consider the multiscale solution $u = (U, w_1, \ldots, w_Q)$ obtained at final iteration of the algorithm. Figures 8a and 8b show the mean and variance of global solution U and local solutions w_q as well as that of multiscale solution u, for isotropic and anisotropic cases, respectively. In the anisotropic case, the variability in the material properties of patch Λ_q decreases with q due to the anisotropy introduced in the weights γ_q. The highest spatial contributions to the variance $\mathbb{V}(u)$ of multiscale solution u are fully captured by the local solution w_q within every patch Λ_q and localized in the first patches in the anisotropic case.

In order to quantify the relative impact of each input random variable ξ_i on the variability of solution u, we introduce the following global sensitivity indices:

$$\widetilde{S}_i(u) = \frac{\mathbb{V}(\mathbb{E}(u(x,\xi)|\xi_i))}{\max_{x\in\Omega}(\mathbb{V}(u(x,\xi)))},$$



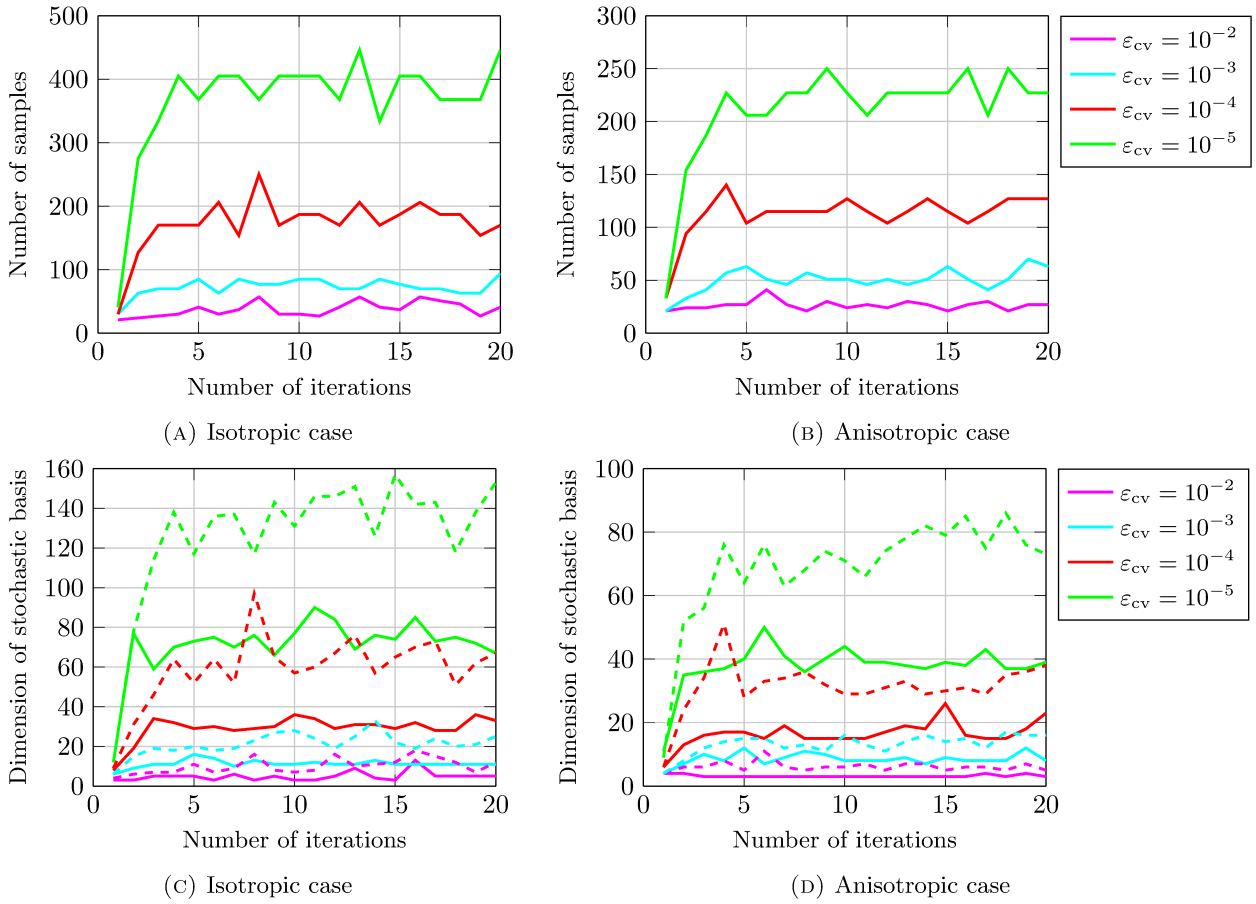
(A) Isotropic case (B) Anisotropic case

FIGURE 6. Evolutions of Aitken's dynamic relaxation parameter ρ_k with respect to iteration number k for different cross-validation tolerances ε_{cv}

where $\mathbb{E}(u(x,\xi)|\xi_i)$ is the conditional expectation of solution u with respect to random variable ξ_i. $\widetilde{S}_i(u)$ is a sensitivity index which reflects the zone of influence of a random variable ξ_i (associated with patch Λ_q for $i \in \{2q-1, 2q\}$) on the variability of solution u. Note that global sensitivity indices $\widetilde{S}_i(u)$ can be straightforwardly computed from the expansion of $u(x,\xi)$ on an orthonormal polynomial basis (see [82]). Figure 9 shows the spatial distributions of sensitivity indices $\widetilde{S}_i(u)$, computed at final iteration of the global-local iterative algorithm for all $i \in \{1,\dots,m\}$. We observe that random variables ξ_{2q-1} and ξ_{2q} have only a local influence within the corresponding patch Λ_q on the variance of solution u. In the anisotropic case, the magnitude of sensitivity indices $\widetilde{S}_i(u)$ again reflects the highest input variabilities in the material properties within the first patches. Note that the patches Λ_q are sufficiently large to capture the main effects of the input uncertainties on local solutions w_q, $q \in \{1,\dots,Q\}$.

6. Conclusion

A global-local iterative method has been proposed for the solution of non-linear stochastic multiscale problems with localized sources of uncertainties and non-linearities. The proposed multiscale approach relies on a domain decomposition method with patches. The iterative coupling strategy is performed by sequentially solving a linear global problem (with deterministic operator and uncertain right-hand side) and a set of independent non-linear local problems (with uncertain operators and right-hand sides) defined on patches. The global-local iterative coupling algorithm is said non-intrusive in the sense that it does not require any modification of both models and solvers during iterations. The local problems can thus be easily handled using dedicated approximation methods and specific solvers. The consistency, convergence and robustness of the proposed algorithm have been analyzed. Numerical results demonstrate the high potential and relevance of the stochastic global-local multiscale approach for dealing with models involving localized uncertainties and possible non-linearities. Several perspectives could be addressed in forthcoming works. First, the modularity of the multiscale approach and in particular its solver coupling capabilities could be exploited in order to take advantage of both commercial software packages available in the industry and in-house research codes, as it was done in recent works [32, 34, 49, 71]. Second, the approach could be extended to more complex non-linear models at multiple scales (e.g. plasticity, damage or fracture in solid mechanics). Also, in the context of heat and mass transfer in fluid mechanics, localized input uncertainties may result in uncertainties in the solution that are scattered in the whole domain (and not confined within patches) due to high convective terms (even in the linear setting), thus requiring an adaptation



(A) Isotropic case (B) Anisotropic case

(C) Isotropic case (D) Anisotropic case

FIGURE 7. Evolutions of (a)-(b) the number of samples N_4 and (c)-(d) the dimension $\#\mathcal{A}_4$ of the approximation basis for local solution w_4^k (solid lines) and Lagrange multiplier λ_4^k (dashed lines) within patch Λ_4 with respect to iteration number k for different cross-validation tolerances ε_{cv}

of the size of the patches during iterations. Quantifying the effects of localized uncertainties in such multiscale stochastic models is currently one of the appealing engineering and scientific challenges.

7. Appendix

The algorithms and the proofs of some technical results are collected in this appendix.

7.1. Leave-one-out cross-validation procedure

Algorithm 7.1 (Leave-one-out cross-validation procedure)**.**

Input: *Coefficients* $\mathbf{V} = (\mathbf{v}_\alpha)_{\alpha\in\mathcal{A}}$ *of the approximation* $\mathbf{v}(\xi) = \sum_{\alpha\in\mathcal{A}} \mathbf{v}_\alpha\psi_\alpha(\xi)$ *of* $\mathbf{u}(\xi)$*, and matrices* $\mathbf{\Psi} = (\psi_\alpha(\xi^l))_{1\leqslant l\leqslant N,\alpha\in\mathcal{A}}$ *and* $\mathbf{Y}^T = (\mathbf{u}(\xi^l))_{1\leqslant l\leqslant N}$ *containing the evaluations of* $(\psi_\alpha(\xi))_{\alpha\in\mathcal{A}}$ *and* $\mathbf{u}(\xi) = (u_i(\xi))_{i\in\mathcal{I}}$

Output: *Vector* $\boldsymbol{\varepsilon} = (\varepsilon_i)_{i\in\mathcal{I}}$*, where* ε_i *is an estimation of the error* $\mathbb{E}((u_i(\xi) - v_i(\xi))^2)$

1: ***for*** $i = 1, \ldots, n$ ***do***

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}	$\#\mathcal{A}$
U	5	2	4	3	4	2	4	3	4	3	4	3	4	2	5	2	192
w_1	3	1	1	1	1	0	0	0	0	0	0	0	0	0	0	0	9
λ_1	4	2	2	2	1	1	0	0	0	0	1	0	0	0	0	0	21
w_2	2	1	3	2	1	1	0	0	0	0	0	0	0	0	0	0	13
λ_2	3	1	3	2	2	2	0	1	0	0	0	0	0	0	0	0	19
w_3	0	0	1	1	3	2	1	1	0	0	0	0	0	0	0	0	11
λ_3	1	1	2	1	3	2	2	1	0	1	0	0	0	0	1	1	21
w_4	0	0	0	0	1	1	3	2	1	1	0	0	0	0	0	0	11
λ_4	0	0	0	1	2	2	4	2	2	1	1	1	0	1	1	0	25
w_5	0	0	0	0	0	0	1	1	3	2	1	1	0	0	0	0	11
λ_5	0	0	0	0	0	1	1	2	3	2	2	1	1	1	0	0	20
w_6	0	0	0	0	0	0	0	0	1	1	3	2	1	1	0	0	11
λ_6	0	0	1	0	0	0	1	1	2	2	4	2	2	2	1	1	28
w_7	0	0	0	0	0	0	0	0	0	0	1	1	3	2	2	1	13
λ_7	0	0	0	0	0	0	0	0	1	1	2	2	3	2	3	1	20
w_8	0	1	1	0	0	0	0	0	1	0	1	0	1	1	3	1	12
λ_8	0	1	1	0	0	0	0	0	1	1	1	1	3	2	5	2	26

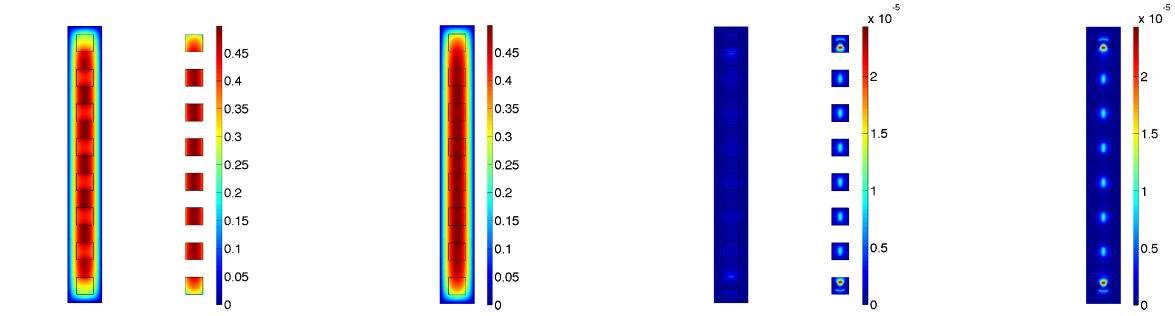
(A) Isotropic case

	p_1	p_2	p_3	p_4	p_5	p_6	p_7	p_8	p_9	p_{10}	p_{11}	p_{12}	p_{13}	p_{14}	p_{15}	p_{16}	$\#\mathcal{A}$
U	5	3	3	2	3	2	3	2	3	2	3	1	2	3	2	1	112
w_1	3	1	1	1	0	0	0	0	0	0	0	0	0	0	0	0	8
λ_1	4	2	2	1	1	1	1	0	0	0	0	0	0	0	0	0	18
w_2	1	1	2	1	1	1	0	0	0	0	0	0	0	0	0	0	9
λ_2	2	1	3	2	1	1	0	1	0	0	0	0	0	0	0	0	15
w_3	1	0	1	1	2	1	1	2	0	0	0	0	0	0	0	0	13
λ_3	1	1	2	1	3	2	1	2	0	1	0	0	0	0	0	0	18
w_4	0	0	0	0	1	1	2	1	1	1	0	0	0	0	0	0	8
λ_4	0	0	0	1	1	1	2	2	1	2	0	1	0	0	1	0	16
w_5	0	0	0	0	0	0	1	1	2	1	0	1	0	0	0	0	7
λ_5	0	0	0	0	0	1	1	1	2	2	1	1	0	0	0	0	11
w_6	0	0	0	0	0	0	0	0	1	1	2	1	0	1	0	0	7
λ_6	0	0	0	0	0	0	0	0	1	1	2	1	1	1	0	0	8
w_7	0	0	0	0	0	0	0	0	0	0	0	1	1	1	1	0	5
λ_7	0	0	0	0	0	0	0	0	0	0	1	1	2	1	1	1	8
w_8	0	0	0	0	0	0	0	0	0	0	0	0	0	1	1	1	4
λ_8	0	0	0	0	0	0	0	0	0	0	0	1	1	1	2	1	7

(B) Anisotropic case

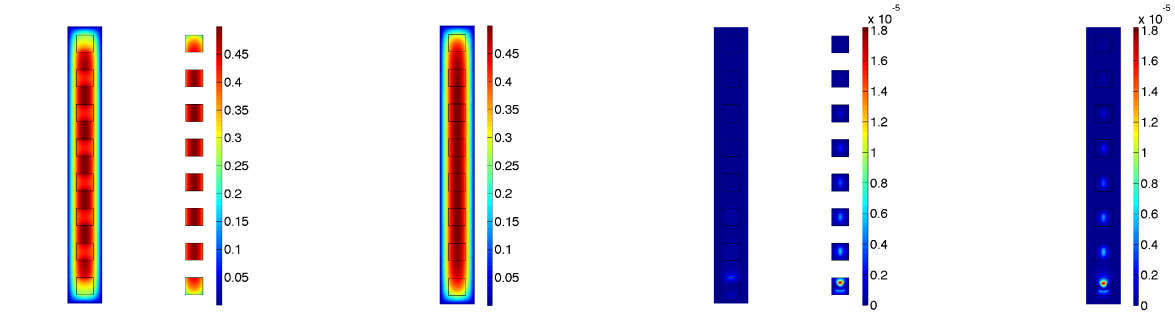
TABLE 2. Partial polynomial degrees p_i (in each random variable ξ_i), $i \in \{1, \ldots, 16\}$, and dimension $\#\mathcal{A}$ of approximation spaces $\mathcal{S}_{\mathcal{A}}$ for global and local solutions U and (w_q, λ_q), $q \in \{1, \ldots, 8\}$, where values displayed in red correspond to random variables ξ_{2q-1} and ξ_{2q} associated with patch Λ_q

2: *Compute the set of predicted residuals* $\{\Delta_l^i\}_{l=1}^N$ *associated with sample set* $\{\xi^l\}_{l=1}^N$ *using Sherman-Morrison-Woodbury formula:* $\Delta_l^i = \frac{r_{li}}{1-h_l}$ *with* r_{li} *the* (l,i)*-th entry of matrix* $\mathbf{R} = \mathbf{\Psi}\mathbf{V}^T - \mathbf{Y}$*, and* h_l *the* l*-th diagonal term in matrix* $\mathbf{H} = \mathbf{\Psi}(\mathbf{\Psi}^T\mathbf{\Psi})^{-1}\mathbf{\Psi}^T$



$\mathbb{E}(U)$ (left) and $\mathbb{E}(w_q)$ (right) $\mathbb{E}(u)$ $\mathbb{V}(U)$ (left) and $\mathbb{V}(w_q)$ (right) $\mathbb{V}(u)$

(A) Isotropic case



$\mathbb{E}(U)$ (left) and $\mathbb{E}(w_q)$ (right) $\mathbb{E}(u)$ $\mathbb{V}(U)$ (left) and $\mathbb{V}(w_q)$ (right) $\mathbb{V}(u)$

(B) Anisotropic case

FIGURE 8. Mean $\mathbb{E}$ and variance $\mathbb{V}$ of global solution U, local solutions w_q and multiscale solution u

Compute the leave-one-out error $e_i = \frac{E_i}{\widehat{m}_2(\mathbf{Y}_i)}$, *where* $E_i = \frac{1}{N}\sum_{l=1}^{N}(\Delta_l^i)^2$ *and* $\widehat{m}_2(\mathbf{Y}_i)$ *is the empirical second moment of the* i*-th column* $\mathbf{Y}_i = (u_i(\xi^l))_{l=1}^N$ *of matrix* $\mathbf{Y}$

Compute the corrected leave-one-out error $\varepsilon_i = e_i \times T(\mathcal{A}, N)$, *where* $T(\mathcal{A}, N) = \left(1 - \frac{\#\mathcal{A}}{N}\right)^{-1}\left(1 + \frac{\mathrm{tr}\left(\widehat{\mathbf{C}}^{-1}\right)}{N}\right)$ *is a correction factor allowing to reduce the sensitivity to overfitting [83, 84] and where* $\widehat{\mathbf{C}} = \frac{1}{N}\mathbf{\Psi}^T\mathbf{\Psi}$ *is the empirical covariance matrix of* $(\psi_\alpha(\xi))_{\alpha\in\mathcal{A}}$

end for

7.2. Adaptive sparse least-squares solver with random sampling and working set strategy

Algorithm 7.2 (Adaptive sparse least-squares solver)**.**

Input: *Initial number of samples* $N \geqslant 1$, *sampling factor* $p_{add} > 0$, *parameter* $\theta \in [0\,,1]$

Output: *Monotone set* $\mathcal{A}$ *and coefficients* $\mathbf{V} = (\mathbf{v}_\alpha)_{\alpha\in\mathcal{A}}$ *of the least-squares approximation* $\mathbf{v}(\xi) = \sum_{\alpha\in\mathcal{A}}\mathbf{v}_\alpha\psi_\alpha(\xi)$ *of* $\mathbf{u}(\xi)$

Start with null initial set $\mathcal{A} = \{0_\mathcal{F}\}$

Generate the initial sample set $\{\xi^l\}_{l=1}^N$ *randomly*

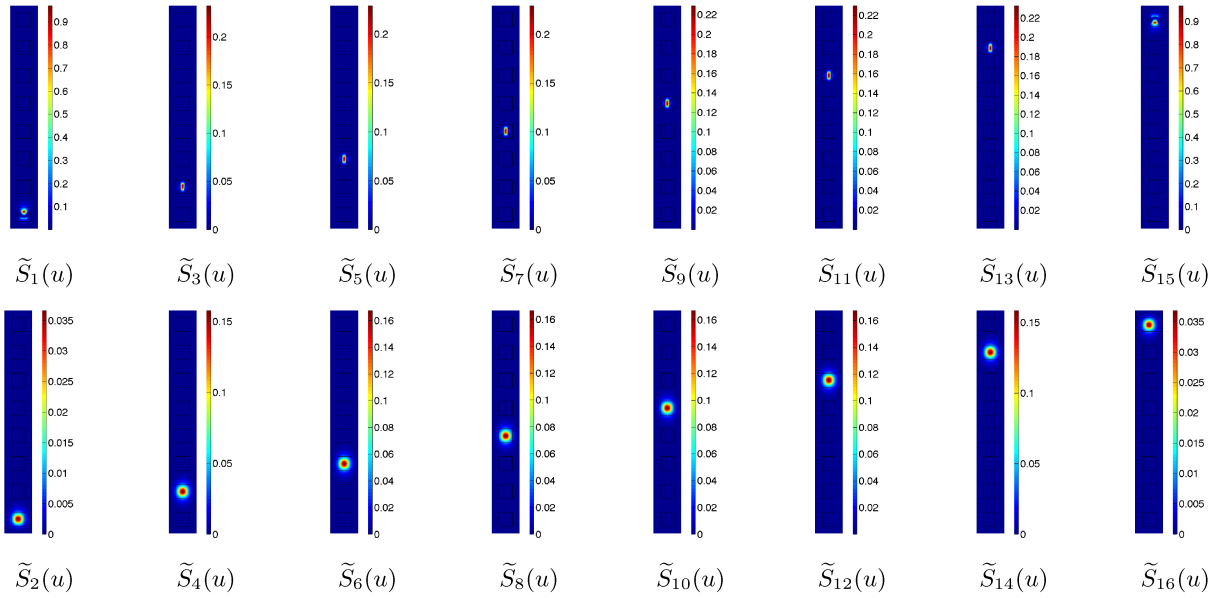
Compute the matrices $\mathbf{\Psi}_\mathcal{A} = (\psi_\alpha(\xi^l))_{1\leqslant l\leqslant N,\alpha\in\mathcal{A}}$ *and* $\mathbf{Y}^T = (\mathbf{u}(\xi^l))_{1\leqslant l\leqslant N}$

while *no convergence* ***do***

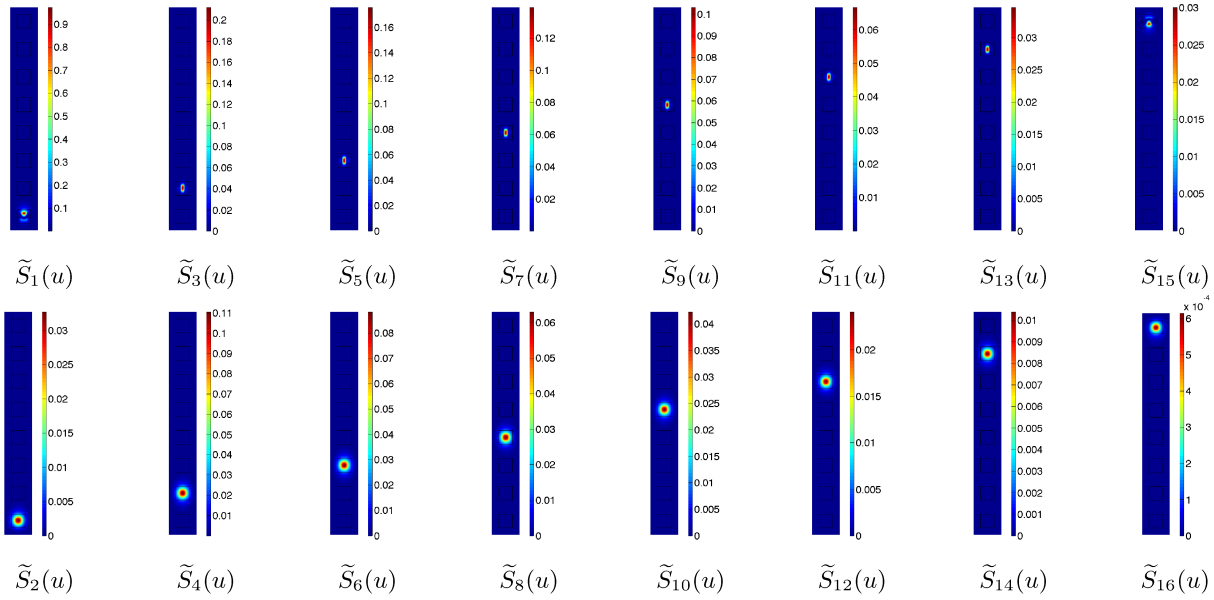
  // Adaptive random sampling

  while *no convergence and no stagnation* ***do***

    Generate the additional sample set $\{\xi^{N+l}\}_{l=1}^{N_{add}}$ *randomly, with* $N_{add} = \mathrm{ceil}(p_{add}N)$



(A) Isotropic case



(B) Anisotropic case

FIGURE 9. Sensitivity indices $\widetilde{S}_i(u)$ of multiscale solution u with respect to each random variable ξ_i, $i \in \{1, \ldots, 16\}$, where the first row gathers the sensitivity indices $\widetilde{S}_{2q-1}(u)$ associated with random diffusion coefficient K_q (parameterized by ξ_{2q-1}) and the second row gathers the sensitivity indices $\widetilde{S}_{2q}(u)$ associated with random reaction parameter R_q (parameterized by ξ_{2q}), $q \in \{1, \ldots, 8\}$

Compute the matrices $\mathbf{\Psi}_{\mathcal{A},add} = (\psi(\xi^{N+l}))_{1\leqslant l\leqslant N_{add},\alpha\in\mathcal{A}}$ *and* $\mathbf{Y}^T_{add} = (\mathbf{u}(\xi^{N+l}))_{1\leqslant l\leqslant N_{add}}$
Update the number of samples $N \leftarrow N + N_{add}$ *and the matrices* $\mathbf{\Psi}^T_{\mathcal{A}} \leftarrow (\mathbf{\Psi}^T_{\mathcal{A}}, \mathbf{\Psi}^T_{\mathcal{A},add})$ *and* $\mathbf{Y}^T \leftarrow (\mathbf{Y}^T, \mathbf{Y}^T_{add})$
Compute the coefficients $\mathbf{V} = (\mathbf{v}_\alpha)_{\alpha\in\mathcal{A}}$ *such that* $\mathbf{V}^T = (\mathbf{\Psi}^T_{\mathcal{A}}\mathbf{\Psi}_{\mathcal{A}})^{-1}\mathbf{\Psi}^T_{\mathcal{A}}\mathbf{Y}$
end while
// Working set strategy
while *no convergence and no overfitting* **do**
Compute the reduced margin $\mathcal{M} = \mathcal{M}_{red}(\mathcal{A})$ *of monotone set* $\mathcal{A}$ *and the set* $\mathcal{T} = \mathcal{A}\cup\mathcal{M}$
if $\#\mathcal{T} > N$ **then**
break
end if
Compute the matrices $\mathbf{\Psi}_{\mathcal{M}} = (\psi_\alpha(\xi^l))_{1\leqslant l\leqslant N,\alpha\in\mathcal{M}}$ *and* $\mathbf{\Psi}_{\mathcal{T}} = (\mathbf{\Psi}_{\mathcal{A}}, \mathbf{\Psi}_{\mathcal{M}})$
Compute the coefficients $\mathbf{V} = (\mathbf{v}_\alpha)_{\alpha\in\mathcal{T}}$ *such that* $\mathbf{V}^T = (\mathbf{\Psi}^T_{\mathcal{T}}\mathbf{\Psi}_{\mathcal{T}})^{-1}\mathbf{\Psi}^T_{\mathcal{T}}\mathbf{Y}$
Compute the vector $(\|\mathbf{v}_\alpha\|_2)_{\alpha\in\mathcal{M}}$
Define the smallest (monotone) subset $\mathcal{N}$ *of* $\mathcal{M}$ *such that* $e(\mathcal{N}) \geqslant \theta e(\mathcal{M})$*, with* $e(\mathcal{N}) = \sum_{\alpha\in\mathcal{N}}\|\mathbf{v}_\alpha\|_2^2$ *and* $e(\mathcal{M}) = \sum_{\alpha\in\mathcal{M}}\|\mathbf{v}_\alpha\|_2^2$
Update the multi-index set $\mathcal{A} \leftarrow \mathcal{A}\cup\mathcal{N}$ *and the matrix* $\mathbf{\Psi}_{\mathcal{A}} \leftarrow (\mathbf{\Psi}_{\mathcal{A}}, \mathbf{\Psi}_{\mathcal{N}})$*, where* $\mathbf{\Psi}_{\mathcal{N}}$ *is the submatrix of* $\mathbf{\Psi}_{\mathcal{M}}$ *whose columns correspond to multi-indices* $\alpha\in\mathcal{N}$
Compute the coefficients $\mathbf{V} = (\mathbf{v}_\alpha)_{\alpha\in\mathcal{A}}$ *such that* $\mathbf{V}^T = (\mathbf{\Psi}^T_{\mathcal{A}}\mathbf{\Psi}_{\mathcal{A}})^{-1}\mathbf{\Psi}^T_{\mathcal{A}}\mathbf{Y}$
end while
end while

7.3. Proof of Lemma 2.2

The first inequality $|v|_{\mathcal{V}} \leqslant \|v\|_{\mathcal{V}}$ is obvious. Using (12) with $(\mathcal{O},\mathcal{E}) = (\Lambda,\Gamma)$, we obtain $\|v_{|\Lambda}\|_{\mathrm{H}^1(\Lambda)} \leqslant C_{\Lambda,\Gamma}(|v_{|\Lambda}|_{\mathrm{H}^1(\Lambda)} + \|v_{|\Lambda}\|_{\mathrm{H}^{1/2}(\Gamma)})$. Then, using the weak continuity of v on Γ, we have $\|v_{|\Lambda}\|_{\mathrm{H}^{1/2}(\Gamma)} = \|v_{|\Omega\setminus\Lambda}\|_{\mathrm{H}^{1/2}(\Gamma)} \leqslant \beta_\tau\|v_{|\Omega\setminus\Lambda}\|_{\mathrm{H}^1(\Omega\setminus\Lambda)}$, where β_τ is the norm of the trace operator $\tau\colon \mathrm{H}^1(\Omega\setminus\Lambda) \to \mathrm{H}^{1/2}(\Gamma)$. Now, using (12) with $(\mathcal{O},\mathcal{E}) = (\Omega\setminus\Lambda, \Gamma_D\cap\partial(\Omega\setminus\Lambda))$, we obtain $\|v_{|\Omega\setminus\Lambda}\|_{\mathrm{H}^1(\Omega\setminus\Lambda)} \leqslant \widehat{C}|v_{|\Omega\setminus\Lambda}|_{\mathrm{H}^1(\Omega\setminus\Lambda)}$, with $\widehat{C} = C_{\Omega\setminus\Lambda,\Gamma_D\cap\partial(\Omega\setminus\Lambda)}$. Then, we deduce that

$$\begin{aligned}\|v\|^2_{\mathcal{V}} &\leqslant \widehat{C}^2|v_{|\Omega\setminus\Lambda}|^2_{\mathrm{H}^1(\Omega\setminus\Lambda)} + C^2_{\Lambda,\Gamma}(|v_{|\Lambda}|_{\mathrm{H}^1(\Lambda)} + \beta_\tau\widehat{C}|v_{|\Omega\setminus\Lambda}|_{\mathrm{H}^1(\Omega\setminus\Lambda)})^2\\ &\leqslant \widehat{C}^2(1+2C^2_{\Lambda,\Gamma}\beta^2_\tau)|v_{|\Omega\setminus\Lambda}|^2_{\mathrm{H}^1(\Omega\setminus\Lambda)} + 2C^2_{\Lambda,\Gamma}|v_{|\Lambda}|^2_{\mathrm{H}^1(\Lambda)} \leqslant C^2_{\mathcal{V}}|v|^2_{\mathcal{V}},\end{aligned}$$

where $C^2_{\mathcal{V}} = \max\{\widehat{C}^2(1+2C^2_{\Lambda,\Gamma}\beta^2_\tau), 2C^2_{\Lambda,\Gamma}\}$. $C_{\mathcal{V}}$ is independent of ξ since $\widehat{C}$ and $C_{\Lambda,\Gamma}$ are independent of ξ (assumption 1.7) and β_τ is independent of ξ.

7.4. Proof of Theorem 2.3

Let $\mathcal{V}^*$ be the topological dual space to $\mathcal{V}$ and let $\langle\cdot,\cdot\rangle$ denote the duality pairing between $\mathcal{V}$ and $\mathcal{V}^*$. For $u\in\mathcal{V}$, $v\mapsto d_\Omega(u,v;\xi)$ is linear and continuous. Then, there exists a unique non-linear map $\mathcal{D}(\xi)\colon \mathcal{V}\to\mathcal{V}^*$ such that $d_\Omega(u,v;\xi) = \langle\mathcal{D}(\xi)(u), v\rangle$ for all $u,v\in\mathcal{V}$. As $\ell_\Omega(\cdot;\xi)$ is linear and continuous on $\mathcal{V}$, there exists a unique $L(\xi)\in\mathcal{V}^*$ such that $\ell_\Omega(v;\xi) = \langle L(\xi), v\rangle$ for all $v\in\mathcal{V}$. Problem (15) can then be written as $\mathcal{D}(\xi)(u(\xi)) = L(\xi)$. First, we have that for all $u,v\in\mathcal{V}$, the map $t\mapsto\langle\mathcal{D}(\xi)(u+tv), v\rangle = a_\Omega(u,v;\xi) + ta_\Omega(v,v;\xi) + n_\Lambda(u+tv,v;\xi)$ is continuous, which implies that $\mathcal{D}(\xi)$ is radially continuous. Then, assumption (8) on $a_{\Omega\setminus\Lambda}$ and a_Λ and assumption (10) on n_Λ imply that for all $u,v\in\mathcal{V}$,

$$\begin{aligned}\langle\mathcal{D}(\xi)(u) - \mathcal{D}(\xi)(v), u-v\rangle &= a_{\Omega\setminus\Lambda}(u-v,u-v) + a_\Lambda(u-v,u-v) + n_\Lambda(u,u-v) - n_\Lambda(v,u-v)\\ &\geqslant \alpha_a(|u-v|^2_{\mathrm{H}^1(\Omega\setminus\Lambda)} + |u-v|^2_{\mathrm{H}^1(\Lambda)}) = \alpha_a|u-v|^2_{\mathcal{V}} \geqslant \frac{\alpha_a}{C^2_{\mathcal{V}}}\|u-v\|^2_{\mathcal{V}} := \alpha_{\mathcal{D}}\|u-v\|^2_{\mathcal{V}},\end{aligned}$$

where the last inequality comes from Lemma 2.2. Then $\mathcal{D}(\xi)$ is strongly monotone with monotonicity constant $\alpha_{\mathcal{D}} = \frac{\alpha_a}{C_{\mathcal{V}}^2}$. Also, assumption (11) on n_Λ implies $\mathcal{D}(\xi)(0) = 0$. From this latter condition and from the strong monotonicity of $\mathcal{D}(\xi)$, we obtain that $\langle \mathcal{D}(\xi)(u), u\rangle \geqslant \alpha_{\mathcal{D}}\|u\|_{\mathcal{V}}^2$ for all $u \in \mathcal{V}$, and therefore $\mathcal{D}(\xi)$ is coercive. Accordingly, $\mathcal{D}(\xi)$ being radially continuous, monotone and coercive, the Browder-Minty theorem [85, Theorem 2.18] ensures that $\mathcal{D}(\xi)$ is surjective, and therefore there exists a solution $u(\xi) \in \mathcal{V}$ to problem (15). The strict monotonicity of $\mathcal{D}(\xi)$ ensures that this solution is unique, so that we can define an inverse map $\mathcal{D}(\xi)^{-1}\colon \mathcal{V}^* \to \mathcal{V}$. The strong monotonicity of $\mathcal{D}(\xi)$ then implies that $\mathcal{D}(\xi)^{-1}$ is Lipschitz continuous, with $\|\mathcal{D}(\xi)^{-1}(L) - \mathcal{D}(\xi)^{-1}(\tilde{L})\|_{\mathcal{V}} \leqslant \frac{1}{\alpha_{\mathcal{D}}}\|L - \tilde{L}\|_{\mathcal{V}^*}$ for all $L, \tilde{L} \in \mathcal{V}^*$. Finally, from the strong monotonicity and from assumption 1.3, we have that $\|u(\xi)\|_{\mathcal{V}} \leqslant \frac{1}{\alpha_{\mathcal{D}}}\|\mathcal{D}(\xi)(u(\xi))\|_{\mathcal{V}^*} = \frac{1}{\alpha_{\mathcal{D}}}\|L(\xi)\|_{\mathcal{V}^*} = \frac{1}{\alpha_{\mathcal{D}}}\sup_{\|v\|_{\mathcal{V}}=1}\langle L(\xi), v\rangle = \frac{1}{\alpha_{\mathcal{D}}}\sup_{\|v\|_{\mathcal{V}}=1}\ell_\Omega(v;\xi) \leqslant \frac{1}{\alpha_{\mathcal{D}}}\kappa(\xi)$, with $\kappa \in \mathrm{L}^p_\mu(\Xi)$. Since α_a and $C_{\mathcal{V}}$ are independent of ξ (see Lemma 2.2), $\alpha_{\mathcal{D}}$ is independent of ξ and we deduce that $u \in \mathrm{L}^p_\mu(\Xi;\mathcal{V})$.

7.5. Proof of Theorem 2.4

Let $u(\xi) \in \widehat{\mathcal{V}}$ such that $u(\xi)_{|\Omega\setminus\Lambda} = U(\xi) \in \mathcal{U}$ and $u(\xi)_{|\Lambda} = w(\xi) \in \mathcal{W}$. Equation (16c) implies that $u(\xi) \in \mathcal{V}$. Then, considering a test function $\delta u \in \mathcal{V}$ and summing (16a) and (16b), we obtain that $u(\xi)$ verifies (15). From Theorem 2.3, we deduce that problem (16) admits a unique solution $(U(\xi), w(\xi)) \in \mathcal{U}\times\mathcal{W}$ which coincides with the solution of (15). Moreover, since $u \in \mathrm{L}^p_\mu(\Xi;\mathcal{V})$, we deduce that $U \in \mathrm{L}^p_\mu(\Xi;\mathcal{U})$ and $w \in \mathrm{L}^p_\mu(\Xi;\mathcal{W})$. Then, let $R\colon \mathrm{H}^{1/2}(\Gamma) \to \mathrm{H}^1(\Omega\setminus\Lambda)$ denote a linear continuous extension operator with continuity constant β_R. Equation (16a) yields

$$b_\Gamma(\lambda(\xi), v) = b_\Gamma(\lambda(\xi), R(v)) = \ell_{\Omega\setminus\Lambda}(R(v)) - a_{\Omega\setminus\Lambda}(U(\xi), R(v))$$

for all $v \in \mathrm{H}^{1/2}(\Gamma)$. The right-hand side being a continuous linear form on $\mathrm{H}^{1/2}(\Gamma)$, we obtain the existence and uniqueness of a solution $\lambda(\xi) \in \mathcal{M}$. Also, $\|\lambda(\xi)\|_{\mathcal{M}} = \sup_{\|v\|_{\mathrm{H}^{1/2}(\Gamma)}=1} b_\Gamma(\lambda(\xi), v) = \sup_{\|v\|_{\mathrm{H}^{1/2}(\Gamma)}=1} \ell_{\Omega\setminus\Lambda}(R(v)) - a_{\Omega\setminus\Lambda}(U(\xi), R(v)) \leqslant \beta_R(\kappa(\xi) + \beta_a\|U(\xi)\|_{\mathcal{U}})$. Since both $\kappa(\xi)$ and $\|U(\xi)\|_{\mathcal{U}}$ belong to $\mathrm{L}^p_\mu(\Xi)$, λ belongs to $\mathrm{L}^p_\mu(\Xi;\mathcal{M})$.

7.6. Proof of Lemma 3.8

Let $\tau\colon \mathrm{H}^1(\Omega\setminus\Lambda) \to \mathrm{H}^{1/2}(\Gamma)$ denote the trace operator which is linear and continuous with norm β_τ independent of ξ. Let $R\colon \mathrm{H}^{1/2}(\Gamma) \to \mathrm{H}^1(\widetilde{\Lambda})$ denote a linear continuous extension operator with norm β_R independent of ξ. For $V \in \widetilde{\mathcal{U}}_\star$, we write $V_{|\widetilde{\Lambda}} = R(V_{|\Gamma}) + Z$, where $Z \in \mathrm{H}^1_0(\widetilde{\Lambda})$ is such that $c_{\widetilde{\Lambda}}(R(V_{|\Gamma}) + Z, \delta U) = 0$ for all $\delta U \in \mathrm{H}^1_0(\widetilde{\Lambda})$. From assumption 2.5 on $c_{\widetilde{\Lambda}}$ and using (12) with $(\mathcal{O},\mathcal{E}) = (\widetilde{\Lambda},\Gamma)$, we obtain that $\|Z\|^2_{\mathrm{H}^1(\widetilde{\Lambda})} \leqslant C^2_{\widetilde{\Lambda},\Gamma}|Z|^2_{\mathrm{H}^1(\widetilde{\Lambda})} \leqslant \frac{C^2_{\widetilde{\Lambda},\Gamma}}{\alpha_c}c_{\widetilde{\Lambda}}(Z,Z) = \frac{C^2_{\widetilde{\Lambda},\Gamma}}{\alpha_c}c_{\widetilde{\Lambda}}(-R(V_{|\Gamma}), Z) \leqslant \frac{C^2_{\widetilde{\Lambda},\Gamma}\beta_c}{\alpha_c}\|R(V_{|\Gamma})\|_{\mathrm{H}^1(\widetilde{\Lambda})}\|Z\|_{\mathrm{H}^1(\widetilde{\Lambda})}$. Then, from the continuity of τ and R, we deduce that $\|V\|_{\mathrm{H}^1(\widetilde{\Lambda})} \leqslant (1+\frac{\beta_c C^2_{\widetilde{\Lambda},\Gamma}}{\alpha_c})\beta_R\|V\|_{\mathrm{H}^{1/2}(\Gamma)} \leqslant (1+\frac{\beta_c C^2_{\widetilde{\Lambda},\Gamma}}{\alpha_c})\beta_R\beta_\tau\|V\|_{\mathrm{H}^1(\Omega\setminus\Lambda)}$. Finally, since $\|V\|^2_{\widetilde{\mathcal{U}}} = \|V\|^2_{\mathcal{U}} + \|V\|^2_{\mathrm{H}^1(\widetilde{\Lambda})}$, we obtain that $\|V\|_{\mathcal{U}} \leqslant \|V\|_{\widetilde{\mathcal{U}}} \leqslant C_{\widetilde{\mathcal{U}}}\|V\|_{\mathcal{U}}$ for all $V \in \widetilde{\mathcal{U}}_\star$, with $C_{\widetilde{\mathcal{U}}} = (1 + (1 + \frac{\beta_c C^2_{\widetilde{\Lambda},\Gamma}}{\alpha_c})^2\beta_R^2\beta_\tau^2)^{1/2}$ independent of ξ.

7.7. Proof of Lemma 3.10

First, using property (18) for $c_{\widetilde{\Omega}}$, property (19) for $c_{\widetilde{\Lambda}}$ and relation (12) for $(\mathcal{O},\mathcal{E}) = (\widetilde{\Omega}, \Gamma_D \cap \partial\widetilde{\Omega})$ (with constant $\widetilde{C} = C_{\widetilde{\Omega},\Gamma_D\cap\partial\widetilde{\Omega}}$), we obtain that $\|\Upsilon(V;\xi)\|_{\widetilde{\mathcal{U}}} \leqslant \beta_\Upsilon\|V\|_{\widetilde{\mathcal{U}}}$ for all $V \in \widetilde{\mathcal{U}}$, with $\beta_\Upsilon = \frac{\beta_c\widetilde{C}^2}{\alpha_c}$ independent of ξ. Then, using again property (18) for $c_{\widetilde{\Omega}}$ and relation (12) for $(\mathcal{O},\mathcal{E}) = (\widetilde{\Omega}, \Gamma_D \cap \partial\widetilde{\Omega})$, we have that $\|\Phi(\beta;\xi)\|^2_{\widetilde{\mathcal{U}}} \leqslant \frac{\widetilde{C}^2}{\alpha_c}|b_\Gamma(\beta, \Phi(\beta;\xi))| \leqslant \frac{\widetilde{C}^2}{\alpha_c}\|\beta\|_{\mathcal{M}}\|\Phi(\beta;\xi)_{|\Gamma}\|_{\mathrm{H}^{1/2}(\Gamma)} \leqslant \frac{\beta_\tau\widetilde{C}^2}{\alpha_c}\|\beta\|_{\mathcal{M}}\|\Phi(\beta;\xi)\|_{\widetilde{\mathcal{U}}}$, where β_τ is the norm of the trace operator $\tau\colon \mathrm{H}^1(\Omega\setminus\Lambda) \to \mathrm{H}^{1/2}(\Gamma)$. That proves $\|\Phi(\beta;\xi)\|_{\widetilde{\mathcal{U}}} \leqslant \beta_\Phi\|\beta\|_{\mathcal{M}}$ for all $\beta \in \mathcal{M}$, with $\beta_\Phi = \frac{\beta_\tau\widetilde{C}^2}{\alpha_c}$ independent of ξ.

7.8. **Proof of Lemma 3.11**

Given the property (13) for the definition of the patch, we can introduce a partition $\overline{\Lambda} = \overline{\Lambda_i \cup \Lambda_e}$ with $\Lambda_i \cap \Lambda_e = \emptyset$ and $\mathrm{dist}(\Lambda_i, \Gamma) = \delta$, for which $\Lambda_\star \subset \Lambda_i$. That means Λ_e is a band of width δ around Γ where the differential operator is linear. Let $\Gamma_e = \partial\Lambda_e \cap \partial\Lambda_i$ and $\Gamma_e^D = \partial\Lambda_e \cap \Gamma_D$. The restriction $w_e(\xi)$ of the local solution $w(\xi) = \Theta(U(\xi);\xi)$ to Λ_e is such that $w_e(\xi) \in \mathrm{H}^1(\Lambda_e)$, $w_e(\xi)_{|\Gamma} = U(\xi)_{|\Gamma}$, $w_e(\xi)_{|\Gamma_e} = w(\xi)_{|\Gamma_e}$, $w_e(\xi)_{|\Gamma_e^D} = 0$ and

$$a_{\Lambda_e}(w_e(\xi), \delta w_e) = \ell_{\Lambda_e}(\delta w_e; \xi) \tag{38}$$

for all $\delta w_e \in \mathrm{H}^1(\Lambda_e)$ such that $\delta w_e = 0$ on $\Gamma \cup \Gamma_e \cup \Gamma_e^D$. Using the linearity of this problem and introducing linear continuous extension operators from $\mathrm{H}^{1/2}(\Gamma)$ and $\mathrm{H}^{1/2}(\Gamma_e)$ to $\mathrm{H}^1(\Lambda_e)$, it can be easily proved that $w_e(\xi)$ can be written as

$$w_e(\xi) = F_\Gamma(U(\xi)_{|\Gamma}) + F_{\Gamma_e}(w(\xi)_{|\Gamma_e}) + \overline{w}_e(\xi), \tag{39}$$

where $\overline{w}_e(\xi) = 0$ on $\Gamma \cup \Gamma_e \cup \Gamma_e^D$, and where $F_\Gamma \colon \mathrm{H}^{1/2}(\Gamma) \to \mathrm{H}^1(\Lambda_e)$ and $F_{\Gamma_e} \colon \mathrm{H}^{1/2}(\Gamma_e) \to \mathrm{H}^1(\Lambda_e)$ are linear continuous extension operators with respective norms $\|F_\Gamma\|$ and $\|F_{\Gamma_e}\|$ independent of ξ, such that $F_\Gamma(v) = v$ on Γ, $F_\Gamma(v) = 0$ on $\Gamma_e \cup \Gamma_e^D$, $F_{\Gamma_e}(v) = v$ on Γ_e, and $F_{\Gamma_e}(v) = 0$ on $\Gamma \cup \Gamma_e^D$. From (38), we obtain that the normal flux on Γ_e, denoted $\lambda_e(\xi) = -B_L(w(\xi), \nabla w(\xi); \cdot) \cdot n \in \mathrm{H}^{1/2}(\Gamma_e)^*$ (with n the unit normal to Γ_e pointing outward Λ_e), is such that $b_{\Gamma_e}(\lambda_e(\xi), \delta w_{e|\Gamma_e}) = \ell_{\Lambda_e}(\delta w_e; \xi) - a_{\Lambda_e}(w_e(\xi), \delta w_e)$ for all $\delta w_e \in \mathrm{H}^1(\Lambda_e)$ such that $\delta w_e = 0$ on $\partial\Lambda_e \setminus \Gamma_e$, where b_{Γ_e} denotes the duality pairing between $\mathrm{H}^{1/2}(\Gamma_e)$ and $\mathrm{H}^{1/2}(\Gamma_e)^*$. From (39), we deduce that

$$b_{\Gamma_e}(\lambda_e(\xi), \delta v) = -g_\Gamma(U(\xi)_{|\Gamma}, \delta v) - g_{\Gamma_e}(w(\xi)_{|\Gamma_e}, \delta v) + b_{\Gamma_e}(\overline{\lambda}_e(\xi), \delta v)$$

for all $\delta v \in \mathrm{H}^{1/2}(\Gamma_e)$, where $g_\Gamma \colon \mathrm{H}^{1/2}(\Gamma) \times \mathrm{H}^{1/2}(\Gamma_e) \to \mathbb{R}$ and $g_{\Gamma_e} \colon \mathrm{H}^{1/2}(\Gamma_e) \times \mathrm{H}^{1/2}(\Gamma_e) \to \mathbb{R}$ are continuous bilinear forms with respective norms $\|g_\Gamma\|$ and $\|g_{\Gamma_e}\|$ independent of ξ, and $\overline{\lambda}_e(\xi) \in \mathrm{H}^{1/2}(\Gamma_e)^*$. By definition, g_{Γ_e} is such that $g_{\Gamma_e}(v, v) = a_{\Lambda_e}(F_{\Gamma_e}(v), \delta w)$ for all $v \in \mathrm{H}^{1/2}(\Gamma_e)$ and $\delta w \in \mathrm{H}^1(\Lambda_e)$ such that $\delta w = v$ on Γ_e and $\delta w = 0$ on $\Gamma \cup \Gamma_e^D$. Choosing $\delta w = F_{\Gamma_e}(v)$ and using property (8) for a_{Λ_e} and relation (12) for $(\mathcal{O}, \mathcal{E}) = (\Lambda_e, \Gamma)$, we obtain that for all $v \in \mathrm{H}^{1/2}(\Gamma_e)$,

$$g_{\Gamma_e}(v, v) = a_{\Lambda_e}(F_{\Gamma_e}(v), F_{\Gamma_e}(v)) \geqslant \alpha_a |F_{\Gamma_e}(v)|^2_{\mathrm{H}^1(\Lambda_e)} \geqslant \frac{\alpha_a}{C^2_{\Lambda_e,\Gamma}} \|F_{\Gamma_e}(v)\|^2_{\mathrm{H}^1(\Lambda_e)} \geqslant \frac{\alpha_a}{C^2_{\Lambda_e,\Gamma} \beta^2_{\tau_{\Lambda_e,\Gamma_e}}} \|v\|^2_{\mathrm{H}^{1/2}(\Gamma_e)},$$

where $\beta_{\tau_{\Lambda_e,\Gamma_e}}$ is the norm of the trace operator from $\mathrm{H}^1(\Lambda_e)$ to $\mathrm{H}^{1/2}(\Gamma_e)$. Then, let $\Gamma_i^D = \partial\Lambda_i \cap \Gamma_D$. The restriction $w_i(\xi)$ of the local solution $w(\xi) = \Theta(U(\xi);\xi)$ to Λ_i is such that $w_i(\xi) \in \mathrm{H}^1(\Lambda_i)$, $w_i(\xi)_{|\Gamma_e} = w(\xi)_{|\Gamma_e}$, $w_i(\xi)_{|\Gamma_i^D} = 0$ and

$$d_{\Lambda_i}(w_i(\xi), \delta w_i; \xi) = \ell_{\Lambda_i}(\delta w_i; \xi) - g_\Gamma(U(\xi)_{|\Gamma}, \delta w_{i|\Gamma_e}) + b_{\Gamma_e}(\overline{\lambda}_e(\xi), \delta w_{i|\Gamma_e}),$$

for all $\delta w_i \in \mathrm{H}^1(\Lambda_i)$ such that $\delta w_i = 0$ on Γ_i^D, where $d_{\Lambda_i}(\cdot, \cdot; \xi)$ is a semi-linear form defined by $d_{\Lambda_i}(u, v; \xi) = a_{\Lambda_i}(u, v; \xi) + n_{\Lambda_i}(u, v; \xi) + g_{\Gamma_e}(u_{|\Gamma_e}, v_{|\Gamma_e})$ for $u, v \in \mathrm{H}^1(\Lambda_i)$. Now, let $\langle \cdot, \cdot \rangle$ denote the duality pairing between $\mathrm{H}^1(\Lambda_i)$ and $\mathrm{H}^1(\Lambda_i)^*$. From properties of a_{Λ_i}, n_{Λ_i} and g_{Γ_e}, we deduce that $d_{\Lambda_i}(u, v; \xi) = \langle \hat{\mathcal{D}}(\xi)(u), v \rangle$ for all $u, v \in \mathrm{H}^1(\Lambda_i)$, where $\hat{\mathcal{D}}(\xi) \colon \mathrm{H}^1(\Lambda_i) \to \mathrm{H}^1(\Lambda_i)^*$ is a radially continuous, coercive and strongly monotone map such that for all $u, v \in \mathrm{H}^1(\Lambda_i)$,

$$\begin{aligned} \langle \hat{\mathcal{D}}(\xi)(u) - \hat{\mathcal{D}}(\xi)(v), u - v \rangle &= a_{\Lambda_i}(u - v, u - v) + n_{\Lambda_i}(u, u - v) - n_{\Lambda_i}(v, u - v) + g_{\Gamma_e}(u_{|\Gamma_e} - v_{|\Gamma_e}, u_{|\Gamma_e} - v_{|\Gamma_e}) \\ &\geqslant \alpha_a |u - v|^2_{\mathrm{H}^1(\Lambda_i)} + \frac{\alpha_a}{C^2_{\Lambda_e,\Gamma} \beta^2_{\tau_{\Lambda_e,\Gamma_e}}} \|u - v\|^2_{\mathrm{H}^{1/2}(\Gamma_e)} \geqslant \alpha_{\hat{\mathcal{D}}} \|u - v\|^2_{\mathrm{H}^1(\Lambda_i)}, \end{aligned}$$

with $\alpha_{\hat{\mathcal{D}}} = \frac{\alpha_a}{2C^2_{\Lambda_i,\Gamma_e}} \min\{1, \frac{1}{C^2_{\Lambda_e,\Gamma} \beta^2_{\tau_{\Lambda_e,\Gamma_e}}}\}$ independent of ξ. Following the proof of Theorem 2.3 in Section 7.4, we obtain that the solution $w_i(\xi)$ is unique and can be written as $w_i(\xi) = G(U(\xi)_{|\Gamma}; \xi) + \overline{w}_i(\xi)$, with $\overline{w}_i(\xi) = 0$ on

$\Gamma \cup \Gamma_i^D$, and where $G(\cdot;\xi)\colon \mathrm{H}^{1/2}(\Gamma) \to \mathrm{H}^1(\Lambda_i)$ is a Lipschitz continuous map with $\beta_G = \frac{\|g_\Gamma\|\beta_{\tau_{\Lambda_i,\Gamma_e}}}{\alpha_{\hat{\mathcal{D}}}}$ independent of ξ, where $\beta_{\tau_{\Lambda_i,\Gamma_e}}$ is the norm of the trace operator from $\mathrm{H}^1(\Lambda_i)$ to $\mathrm{H}^{1/2}(\Gamma_e)$. Finally, we deduce that for all $U, V \in \widetilde{\mathcal{U}}$,

$$
\begin{aligned}
&\|\Theta(U;\xi) - \Theta(V;\xi)\|_{\mathcal{W}}^2 = \|\Theta(U;\xi)_{|\Lambda_e} - \Theta(V;\xi)_{|\Lambda_e}\|_{\mathrm{H}^1(\Lambda_e)}^2 + \|\Theta(U;\xi)_{|\Lambda_i} - \Theta(V;\xi)_{|\Lambda_i}\|_{\mathrm{H}^1(\Lambda_i)}^2 \\
&= \|F_\Gamma(U_{|\Gamma} - V_{|\Gamma})\|_{\mathrm{H}^1(\Lambda_e)}^2 + \|F_{\Gamma_e}(\Theta(U;\xi)_{|\Gamma_e} - \Theta(V;\xi)_{|\Gamma_e})\|_{\mathrm{H}^1(\Lambda_e)}^2 + \|G(U_{|\Gamma};\xi) - G(V_{|\Gamma};\xi)\|_{\mathrm{H}^1(\Lambda_i)}^2 \\
&\leqslant (\|F_\Gamma\|^2 + \beta_G^2)\|U - V\|_{\mathrm{H}^{1/2}(\Gamma)}^2 + \|F_{\Gamma_e}\|^2\|\Theta(U;\xi)_{|\Gamma_e} - \Theta(V;\xi)_{|\Gamma_e}\|_{\mathrm{H}^{1/2}(\Gamma_e)}^2 \\
&\leqslant (\|F_\Gamma\|^2 + \beta_G^2 + \|F_{\Gamma_e}\|^2\beta_{\tau_{\Lambda_i,\Gamma_e}}^2\beta_G^2)\|U - V\|_{\mathrm{H}^{1/2}(\Gamma)}^2 \\
&\leqslant \beta_\Theta^2\|U - V\|_{\widetilde{\mathcal{U}}}^2,
\end{aligned}
$$

with $\beta_\Theta^2 = (\|F_\Gamma\|^2 + \beta_G^2 + \|F_{\Gamma_e}\|^2\beta_{\tau_{\Lambda_i,\Gamma_e}}^2\beta_G^2)\beta_{\tau_{\Omega\setminus\Lambda,\Gamma}}^2$ independent of ξ, where $\beta_{\tau_{\Omega\setminus\Lambda,\Gamma}}$ is the norm of the trace operator from $\mathrm{H}^1(\Omega \setminus \Lambda)$ to $\mathrm{H}^{1/2}(\Gamma)$.

Using (26a) with test functions $\delta w = 0$ on Λ_i and introducing a linear continuous extension operator R_Γ from $\mathrm{H}^{1/2}(\Gamma)$ to $\mathrm{H}^1(\Lambda_e)$ (with norm $\|R_\Gamma\|$ independent of ξ) such that $R_\Gamma(v) = 0$ on $\Gamma_e \cup \Gamma_e^D$, we have that for all $V \in \widetilde{\mathcal{U}}$,

$$
b_\Gamma(\Psi(V;\xi), \delta w_e) = a_{\Lambda_e}(\Theta(V;\xi)_{|\Lambda_e}, \delta w_e) - \ell_{\Lambda_e}(\delta w_e;\xi)
$$

for all $\delta w_e \in \mathrm{H}^1(\Lambda_e)$ such that $\delta w_e = 0$ on $\Gamma_e \cup \Gamma_e^D$. Then, we obtain that for all $U, V \in \widetilde{\mathcal{U}}$,

$$
\begin{aligned}
\|\Psi(U;\xi) - \Psi(V;\xi)\|_{\mathcal{M}} &= \sup_{\substack{v\in\mathrm{H}^{1/2}(\Gamma)\\ \|v\|_{\mathrm{H}^{1/2}(\Gamma)}=1}} b_\Gamma(\Psi(U;\xi) - \Psi(V;\xi), v) \\
&= \sup_{\substack{v\in\mathrm{H}^{1/2}(\Gamma)\\ \|v\|_{\mathrm{H}^{1/2}(\Gamma)}=1}} a_{\Lambda_e}(\Theta(U;\xi)_{|\Lambda_e} - \Theta(V;\xi)_{|\Lambda_e}, R_\Gamma(v)) \\
&\leqslant \beta_a\|\Theta(U;\xi)_{|\Lambda_e} - \Theta(V;\xi)_{|\Lambda_e}\|_{\mathrm{H}^1(\Lambda_e)}\|R_\Gamma\| \\
&\leqslant \beta_\Psi\|U - V\|_{\widetilde{\mathcal{U}}},
\end{aligned}
$$

with $\beta_\Psi = \beta_a\beta_\Theta\|R_\Gamma\|$ independent of ξ.

7.9. **Proof of Lemma 3.12**

From the Lipschitz continuity of mappings Υ, Φ and Ψ (see Lemmas 3.10 and 3.11) and from the continuity of the linear map $C_{\widetilde{\Omega}}$, we deduce the Lipschitz continuity of $D(\cdot;\xi)$ on $\widetilde{\mathcal{U}}$ and *a fortiori* on $\widetilde{\mathcal{U}}_\star$, with Lipschitz constant $\beta_D = \beta_c(1 + \beta_\Upsilon + \beta_\Phi\beta_\Psi)$ independent of ξ. From (29), we have that for all $U, V \in \widetilde{\mathcal{U}}$,

$$
\begin{aligned}
\langle D(U;\xi) - D(V;\xi), U - V\rangle_{\widetilde{\mathcal{U}}} &= c_{\widetilde{\Omega}}(A(U;\xi) - A(V;\xi), U - V) \\
&= a_{\Omega\setminus\Lambda}(U - V, U - V) + b_\Gamma(\Psi(U;\xi) - \Psi(V;\xi), U - V) \\
&= a_{\Omega\setminus\Lambda}(U - V, U - V) + b_\Gamma(\Psi(U;\xi) - \Psi(V;\xi), \Theta(U;\xi) - \Theta(V;\xi)) \\
&= a_{\Omega\setminus\Lambda}(U - V, U - V) + a_\Lambda(\Theta(U;\xi) - \Theta(V;\xi), \Theta(U;\xi) - \Theta(V;\xi);\xi) \\
&\quad + n_\Lambda(\Theta(U;\xi), \Theta(U;\xi) - \Theta(V;\xi);\xi) - n_\Lambda(\Theta(V;\xi), \Theta(U;\xi) - \Theta(V;\xi);\xi).
\end{aligned}
$$

From assumptions 1.4 and 1.5, we have $a_\Lambda(w, w;\xi) \geqslant 0$ and $n_\Lambda(w, w - w';\xi) - n_\Lambda(w', w - w';\xi) \geqslant 0$ for all $w, w' \in \mathcal{W}$. Therefore, using property (8) for $a_{\Omega\setminus\Lambda}$ and relation (12) for $(\mathcal{O}, \mathcal{E}) = (\Omega \setminus \Lambda, \Gamma_D \cap \partial(\Omega \setminus \Lambda))$, we have that for all $U, V \in \widetilde{\mathcal{U}}$,

$$
\langle D(U;\xi) - D(V;\xi), U - V\rangle_{\widetilde{\mathcal{U}}} \geqslant a_{\Omega\setminus\Lambda}(U - V, U - V) \geqslant \alpha_a|U - V|_{\mathrm{H}^1(\Omega\setminus\Lambda)}^2 \geqslant \frac{\alpha_a}{\widehat{C}^2}\|U - V\|_{\widetilde{\mathcal{U}}}^2,
$$

where $\widehat{C} = C_{\Omega\setminus\Lambda,\Gamma_D\cap\partial(\Omega\setminus\Lambda)}$. Finally, using Lemma 3.8, we obtain that for all $U, V \in \widetilde{\mathcal{U}}_\star$,

$$\langle D(U;\xi) - D(V;\xi), U - V\rangle_{\widetilde{\mathcal{U}}} \geqslant \alpha_D \|U - V\|^2_{\widetilde{\mathcal{U}}},$$

with $\alpha_D = \frac{\alpha_a}{\widehat{C}^2 C^2_{\widetilde{\mathcal{U}}}}$ independent of ξ. That proves the strong monotonicity of $D(\cdot;\xi)$ on $\widetilde{\mathcal{U}}_\star$.

7.10. **Proof of Theorem 3.13**

First recall that the global solution $U(\xi)$ and all global iterates $U^k(\xi)$ are in $\widetilde{\mathcal{U}}_\star$ (see Lemma 3.6 and Theorem 3.7). Using relation (12) for $(\mathcal{O},\mathcal{E}) = (\widetilde{\Omega}, \Gamma_D \cap \partial\widetilde{\Omega})$, the symmetry and property (18) for $c_{\widetilde{\Omega}}$, as well as the Lipschitz continuity and the strong monotonicity of map $D(\cdot;\xi)$ on $\widetilde{\mathcal{U}}_\star$ (see Lemma 3.12), we obtain that for all $U, V \in \widetilde{\mathcal{U}}_\star$,

$$\begin{aligned}
&\|B_{\rho_k}(U;\xi) - B_{\rho_k}(V;\xi)\|^2_{\widetilde{\mathcal{U}}} \leqslant \widetilde{C}^2 |B_{\rho_k}(U;\xi) - B_{\rho_k}(V;\xi)|^2_{\mathrm{H}^1(\widetilde{\Omega})} \\
&\leqslant \frac{\widetilde{C}^2}{\alpha_c} c_{\widetilde{\Omega}}(B_{\rho_k}(U) - B_{\rho_k}(V;\xi), B_{\rho_k}(U;\xi) - B_{\rho_k}(V;\xi);\xi) \\
&= \frac{\widetilde{C}^2}{\alpha_c} \big(c_{\widetilde{\Omega}}(U - V, U - V) - 2\rho_k c_{\widetilde{\Omega}}(A(U;\xi) - A(V;\xi), U - V) + \rho_k^2 c_{\widetilde{\Omega}}(A(U;\xi) - A(V;\xi), A(U;\xi) - A(V;\xi))\big) \\
&\leqslant \frac{\widetilde{C}^2}{\alpha_c} \big(\langle C_{\widetilde{\Omega}}(U - V), U - V\rangle_{\widetilde{\mathcal{U}}} - 2\rho_k \langle D(U;\xi) - D(V;\xi), U - V\rangle_{\widetilde{\mathcal{U}}} + \rho_k^2 \|D(U;\xi) - D(V;\xi)\|_{\widetilde{\mathcal{U}}} \|A(U;\xi) - A(V;\xi)\|_{\widetilde{\mathcal{U}}}\big) \\
&\leqslant \frac{\widetilde{C}^2}{\alpha_c} \left(\beta_c - 2\rho_k \alpha_D + \rho_k^2 \frac{\beta_D^2}{\beta_c}\right) \|U - V\|^2_{\widetilde{\mathcal{U}}},
\end{aligned}$$

where $\widetilde{C} = C_{\widetilde{\Omega},\Gamma_D\cap\partial\widetilde{\Omega}}$. If $\{\rho_k\}_{k\in\mathbb{N}}$ satisfies (31) with $\rho_{\sup} < \frac{\beta_c}{\beta_D^2}\left(2\alpha_D - \frac{1}{\rho_{\inf}}\left(\beta_c - \frac{\alpha_c}{\widetilde{C}^2}\right)\right) := \rho^*_{\sup}$, then the set of mappings $\{B_{\rho_k}(\cdot;\xi)\}_{k\in\mathbb{N}}$ is uniformly contractive on $\widetilde{\mathcal{U}}_\star$, with a contractivity constant $\rho_B = \left(\frac{\widetilde{C}^2}{\alpha_c}(\beta_c - \rho_{\inf}(2\alpha_D - \rho_{\sup}\frac{\beta_D^2}{\beta_c}))\right)^{1/2} < 1$ independent of ξ. Then, we obtain that

$$\|U^k(\xi) - U(\xi)\|_{\widetilde{\mathcal{U}}} = \|B_{\rho_k}(U^{k-1}(\xi);\xi) - B_{\rho_k}(U(\xi);\xi)\|_{\widetilde{\mathcal{U}}} \leqslant \rho_B^k \|U(\xi) - U^0(\xi)\|_{\widetilde{\mathcal{U}}},$$

from which we deduce that the sequence $\{U^k(\xi)\}_{k\in\mathbb{N}}$ converges almost surely to $U(\xi)$ in $\widetilde{\mathcal{U}}_\star$. Also, with $U^0 = 0$, we obtain that $\|U^k(\xi) - U(\xi)\|_{\widetilde{\mathcal{U}}} \leqslant \|U(\xi)\|_{\widetilde{\mathcal{U}}}$. Since $\|U(\xi)\|_{\widetilde{\mathcal{U}}} \in \mathrm{L}^p_\mu(\Xi)$ (see Corollary 3.9), the dominated convergence theorem gives that the sequence $\{U^k\}_{k\in\mathbb{N}}$ converges to U in $\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})$. Finally, from (30) and using the Lipschitz continuity of mappings Θ and Ψ (see Lemma 3.11), we directly obtain that the sequence $\{w^k\}_{k\in\mathbb{N}}$ converges to w almost surely and in $\mathrm{L}^p_\mu(\Xi;\mathcal{W})$, and the sequence $\{\lambda^k\}_{k\in\mathbb{N}}$ converges to λ almost surely and in $\mathrm{L}^p_\mu(\Xi;\mathcal{M})$.

7.11. **Proof of Theorem 3.16**

First, if initial guess $U^0_\varepsilon \in \mathcal{V}_\delta$ and if $\varepsilon^* < 1 - \rho_B$ and $\varepsilon \leqslant \delta(1 - \rho_B - \varepsilon^*)$, we can prove by induction that all approximate global iterates U^k_ε belong to $\mathcal{V}_\delta$. Indeed, suppose that $U^j_\varepsilon \in \mathcal{V}_\delta$ for all $j < k$. Then, the error at iteration k is such that

$$\begin{aligned}
\|U^k_\varepsilon - U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} &\leqslant \|U^k_\varepsilon - U^k\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} + \|U^k - U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} \\
&\leqslant \|B^\varepsilon_{\rho_k}(U^{k-1}_\varepsilon) - B_{\rho_k}(U^{k-1}_\varepsilon)\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} + \|B_{\rho_k}(U^{k-1}_\varepsilon) - B_{\rho_k}(U)\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} \\
&\leqslant \varepsilon \|U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} + (\rho_B + \varepsilon^*)\|U^{k-1}_\varepsilon - U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})}.
\end{aligned}$$

As $U_\varepsilon^{k-1} \in \mathcal{V}_\delta$ and $\varepsilon \leqslant \delta(1-\rho_B-\varepsilon^*)$, $U_\varepsilon^k \in \mathcal{V}_\delta$. By induction, we finally prove that $U_\varepsilon^k \in \mathcal{V}_\delta$ for all $k \in \mathbb{N}$. It means that if initial global iterate U_ε^0 is contained in the open ball $\mathcal{V}_\delta$ of radius $\delta\|U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})}$ centered at the exact global solution U, then all global iterates U_ε^k remain in this ball. Subsequently, we obtain that

$$\begin{aligned}\|U_\varepsilon^k - U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} &\leqslant \varepsilon\|U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})}\sum_{j=0}^{k-1}(\rho_B+\varepsilon^*)^j + (\rho_B+\varepsilon^*)^k\|U_\varepsilon^0 - U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})}\\ &\leqslant \frac{\varepsilon\left(1-(\rho_B+\varepsilon^*)^k\right)}{1-(\rho_B+\varepsilon^*)}\|U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} + (\rho_B+\varepsilon^*)^k\|U_\varepsilon^0 - U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})}\\ &\leqslant \frac{\varepsilon}{1-(\rho_B+\varepsilon^*)}\|U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})} + (\rho_B+\varepsilon^*)^k\|U_\varepsilon^0 - U\|_{\mathrm{L}^p_\mu(\Xi;\widetilde{\mathcal{U}})}\end{aligned}$$

and therefore, as $0 < \rho_B+\varepsilon^* < 1$, the approximate sequence $\{U_\varepsilon^k\}_{k\in\mathbb{N}}$ satisfies (33), with $\gamma(\varepsilon,\varepsilon^*) = \frac{\varepsilon}{1-(\rho_B+\varepsilon^*)} \to 0$ as $\varepsilon \to 0$.

References

[1] Anthony Nouy. Recent developments in spectral stochastic methods for the numerical solution of stochastic partial differential equations. *Archives of Computational Methods in Engineering*, 16:251–285, 2009.

[2] D. Xiu. Fast numerical methods for stochastic computations: a review. *Communications in computational physics*, 5(2-4):242–272, 2009.

[3] O. P. Le Maître and O. M. Knio. *Spectral Methods for Uncertainty Quantification With Applications to Computational Fluid Dynamics*. Springer Netherlands, 2010.

[4] Thomas Y. Hou and Xiao-Hui Wu. A Multiscale Finite Element Method for Elliptic Problems in Composite Materials and Porous Media. *Journal of Computational Physics*, 134(1):169–189, 1997.

[5] Thomas J.R. Hughes, Gonzalo R. Feijóo, Luca Mazzei, and Jean-Baptiste Quincy. The variational multiscale method—a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166(1-2):3–24, 1998. Advances in Stabilized Methods in Computational Mechanics.

[6] Weinan E and Björn Engquist. The heterogeneous multiscale methods. *Communications in Mathematical Sciences*, 1(1):87–132, 2003.

[7] X. Frank Xu. A multiscale stochastic finite element method on elliptic problems involving uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 196(25-28):2723–2736, 2007.

[8] Velamur Asokan Badri Narayanan and Nicholas Zabaras. Variational multiscale stabilized FEM formulations for transport equations: stochastic advection–diffusion and incompressible stochastic Navier–Stokes equations. *Journal of Computational Physics*, 202(1):94–133, 2005.

[9] Badrinarayanan Velamur Asokan and Nicholas Zabaras. A stochastic variational multiscale method for diffusion in heterogeneous random media. *Journal of Computational Physics*, 218(2):654–676, 2006.

[10] Baskar Ganapathysubramanian and Nicholas Zabaras. Modeling diffusion in random heterogeneous media: Data-driven models, stochastic collocation and the variational multiscale method. *Journal of Computational Physics*, 226(1):326–353, 2007.

[11] P. Dostert, Y. Efendiev, and T.Y. Hou. Multiscale finite element methods for stochastic porous media flow equations and application to uncertainty quantification. *Computer Methods in Applied Mechanics and Engineering*, 197(43-44):3445–3455, 2008.

[12] X. Frank Xu, Xi Chen, and Lihua Shen. A Green-function-based multiscale method for uncertainty quantification of finite body random heterogeneous materials. *Computers & Structures*, 87(21-22):1416–1426, 2009.

[13] B. Ganapathysubramanian and N. Zabaras. A stochastic multiscale framework for modeling flow through random heterogeneous porous media. *Journal of Computational Physics*, 228(2):591–618, 2009.

[14] V. Ginting, A. Målqvist, and M. Presho. A Novel Method for Solving Multiscale Elliptic Problems with Randomly Perturbed Data. *Multiscale Modeling & Simulation*, 8(3):977–996, 2010.

[15] Le Bris, Claude, Legoll, Frédéric, and Thomines, Florian. Multiscale finite element approach for "weakly" random problems and related issues. *ESAIM: M2AN*, 48(3):815–858, 2014.

[16] C. Jin, X. Cai, and C. Li. Parallel Domain Decomposition Methods for Stochastic Elliptic Equations. *SIAM Journal on Scientific Computing*, 29(5):2096–2114, 2007.

[17] Kai Zhang, Ran Zhang, Yunguang Yin, and Shi Yu. Domain decomposition methods for linear and semilinear elliptic stochastic partial differential equations. *Applied Mathematics and Computation*, 195(2):630–640, 2008.

[18] Abhijit Sarkar, Nabil Benabbou, and Roger Ghanem. Domain decomposition of stochastic PDEs: Theoretical formulations. *International Journal for Numerical Methods in Engineering*, 77(5):689–701, 2009.

[19] B. Ganis, I. Yotov, and M. Zhong. A Stochastic Mortar Mixed Finite Element Method for Flow in Porous Media with Multiple Rock Types. *SIAM Journal on Scientific Computing*, 33(3):1439–1474, 2011.

[20] Mary F. Wheeler, Tim Wildey, and Ivan Yotov. A multiscale preconditioner for stochastic mortar mixed finite elements. *Computer Methods in Applied Mechanics and Engineering*, 200(9-12):1251–1262, 2011.

[21] R. Verfürth. *A Review of A Posteriori Error Estimation and Adaptive Mesh-refinement Techniques*. Wiley-Teubner, Stuttgart, 1996.

[22] E. Stein and S. Ohnimus. Coupled model- and solution-adaptivity in the finite-element method. *Computer Methods in Applied Mechanics and Engineering*, 150(1-4):327–350, 1997.

[23] T. Belytschko and T. Black. Elastic crack growth in finite elements with minimal remeshing. *International Journal for Numerical Methods in Engineering*, 45(5):601–620, 1999.

[24] N. Moës, J. Dolbow, and T. Belytschko. A finite element method for crack growth without remeshing. *International Journal for Numerical Methods in Engineering*, 46(1):131–150, 1999.

[25] T. Strouboulis, I. Babuška, and K. Copps. The design and analysis of the Generalized Finite Element Method. *Computer Methods in Applied Mechanics and Engineering*, 181(1-3):43–69, 2000.

[26] Jacques-Louis Lions and Olivier Pironneau. Domain decomposition methods for CAD. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 328(1):73–80, 1999.

[27] Roland Glowinski, Jiwen He, Alexei Lozinski, Jacques Rappaz, and Joël Wagner. Finite element approximation of multi-scale elliptic problems using patches of elements. *Numerische Mathematik*, 101(4):663–687, 2005.

[28] Jiwen He, Alexei Lozinski, and Jacques Rappaz. Accelerating the method of finite element patches using approximately harmonic functions. *Comptes Rendus Mathematique*, 345(2):107–112, 2007.

[29] Joseph L Steger, F Carroll Dougherty, and John A Benek. A Chimera grid scheme. In K.N. Ghia and U. Ghia, editors, *Advances in Grid Generation*, volume 5, pages 59–69, New York, 1983. American Society of Mechanical Engineers, FED.

[30] Franco Brezzi, Jacques-Louis Lions, and Olivier Pironneau. Analysis of a Chimera method. *Comptes Rendus de l'Académie des Sciences - Series I - Mathematics*, 332(7):655–660, 2001.

[31] Olivier Pironneau, Alexei Lozinski, Alain Perronnet, and Frédéric Hecht. Numerical zoom for multiscale problems with an application to flows through porous media. *Discrete & Continuous Dynamical Systems - A*, 23(1/2):265–280, 2009.

[32] Lionel Gendre, Olivier Allix, Pierre Gosselet, and François Comte. Non-intrusive and exact global/local techniques for structural problems with local plasticity. *Computational Mechanics*, 44:233–245, 2009.

[33] Alexei Lozinski. *Méthodes numériques et modélisation pour certains problèmes multi-échelles*. Habilitation à diriger des recherches, Université Paul Sabatier, Toulouse 3, France, 2010.

[34] L. Gendre, O. Allix, and P. Gosselet. A two-scale approximation of the Schur complement and its use for non-intrusive coupling. *International Journal for Numerical Methods in Engineering*, 87(9):889–905, 2011.

[35] Corinna Hager, Patrice Hauret, Patrick Le Tallec, and Barbara I. Wohlmuth. Solving dynamic contact problems with local refinement in space and time. *Computer Methods in Applied Mechanics and Engineering*, 201-204(0):25–41, 2012.

[36] A. Gravouil, J. Rannou, and M.-C. Baïetto. A Local Multi-grid X-FEM approach for 3D fatigue crack growth. *International Journal of Material Forming*, 1(1):1103–1106, 2008.

[37] J. Rannou, A. Gravouil, and M. C. Baïetto-Dubourg. A local multigrid X-FEM strategy for 3-D crack propagation. *International Journal for Numerical Methods in Engineering*, 77(4):581–600, 2009.

[38] J. C. Passieux, A. Gravouil, J. Réthoré, and M. C. Baïetto. Direct estimation of generalized stress intensity factors using a three-scale concurrent multigrid X-FEM. *International Journal for Numerical Methods in Engineering*, 85(13):1648–1666, 2011.

[39] Jean-Charles Passieux, Julien Réthoré, Anthony Gravouil, and Marie-Christine Baïetto. Local/global non-intrusive crack propagation simulation using a multigrid X-FEM solver. *Computational Mechanics*, 52(6):1381–1393, 2013.

[40] Hachmi Ben Dhia. Problèmes mécaniques multi-échelles: la méthode Arlequin - Multiscale mechanical problems: the Arlequin method. *Comptes Rendus de l'Académie des Sciences - Series IIB - Mechanics-Physics-Astronomy*, 326(12):899–904, 1998.

[41] Ludovic Chamoin, J.T. Oden, and Serge Prudhomme. A stochastic coupling method for atomic-to-continuum Monte-Carlo simulations. *Computer Methods in Applied Mechanics and Engineering*, 197(43-44):3530–3546, 2008.

[42] R. Cottereau, D. Clouteau, H. Ben Dhia, and C. Zaccardi. A stochastic-deterministic coupling method for continuum mechanics. *Computer Methods in Applied Mechanics and Engineering*, 200(47-48):3280–3288, 2011.

[43] M. Chevreuil, A. Nouy, and E. Safatly. A multiscale method with patch for the solution of stochastic partial differential equations with localized uncertainties. *Computer Methods in Applied Mechanics and Engineering*, 255:255–274, 2013.

[44] Henning, Patrick, Målqvist, Axel, and Peterseim, Daniel. A localized orthogonal decomposition method for semi-linear elliptic problems. *ESAIM: M2AN*, 48(5):1331–1349, 2014.

[45] Y. Efendiev, T. Y. Hou, and V. Ginting. Multiscale Finite Element Methods for Nonlinear Problems and Their Applications. *Commun. Math. Sci.*, 2(4):553–589, 12 2004.

[46] Yalchin Efendiev and Thomas Y Hou. *Multiscale Finite Element Methods: Theory and Applications*, volume 4 of *Surveys and Tutorials in the Applied Mathematical Sciences*. Springer-Verlag New York, 2009.

[47] Jan Martin Nordbotten. *Variational and Heterogeneous Multiscale Methods*, pages 713–720. Springer Berlin Heidelberg, Berlin, Heidelberg, 2010.

[48] Patrick Henning and Mario Ohlberger. Error control and adaptivity for heterogeneous multiscale approximations of nonlinear monotone problems. *Discrete and Continuous Dynamical Systems - Series S*, 8(1):119–150, 2015.
[49] Olivier Allix, Lionel Gendre, Pierre Gosselet, and Guillaume Guguin. Non-intrusive Coupling: An Attempt to Merge Industrial and Research Software Capabilities. In Dana Mueller-Hoeppe, Stefan Loehnert, and Stefanie Reese, editors, *Recent Developments and Innovative Applications in Computational Mechanics*, chapter 15, pages 125–133. Springer Berlin Heidelberg, 2011.
[50] Abdellah Chkifa, Albert Cohen, Ronald DeVore, and Christoph Schwab. Sparse adaptive Taylor approximation algorithms for parametric and stochastic elliptic PDEs. *ESAIM: M2AN*, 47(1):253–280, 2013.
[51] A. D. Bazykin. Hypothetical Mechanism of Speciaton. *Evolution*, 23(4):685–687, 1969.
[52] D. G. Aronson and H. F. Weinberger. *Nonlinear diffusion in population genetics, combustion, and nerve pulse propagation*, pages 5–49. Springer Berlin Heidelberg, Berlin, Heidelberg, 1975.
[53] B.H Bradshaw-Hajek and P Broadbridge. A robust cubic reaction-diffusion system for gene propagation. *Mathematical and Computer Modelling*, 39(9):1151–1163, 2004.
[54] Adel Hanna, Alan Saul, and Kenneth Showalter. Detailed studies of propagating fronts in the iodate oxidation of arsenous acid. *Journal of the American Chemical Society*, 104(14):3838–3844, 1982.
[55] Jie Huang and Boyd F. Edwards. Pattern formation and evolution near autocatalytic reaction fronts in a narrow vertical slab. *Phys. Rev. E*, 54:2620–2627, Sep 1996.
[56] Boyd F. Edwards. Poiseuille Advection of Chemical Reaction Fronts. *Phys. Rev. Lett.*, 89:104501, Aug 2002.
[57] Mads Kærn and Michael Menzinger. Propagation of Excitation Pulses and Autocatalytic Fronts in Packed-Bed Reactors. *The Journal of Physical Chemistry B*, 106(14):3751–3758, 2002.
[58] Robert S. Spangler and Boyd F. Edwards. Poiseuille advection of chemical reaction fronts: Eikonal approximation. *The Journal of Chemical Physics*, 118(13):5911–5915, 2003.
[59] M. Leconte, J. Martin, N. Rakotomalala, and D. Salin. Pattern of Reaction Diffusion Fronts in Laminar Flows. *Phys. Rev. Lett.*, 90:128302, Mar 2003.
[60] Igor V. Koptyug, Vladimir V. Zhivonitko, and Renad Z. Sagdeev. Advection of Chemical Reaction Fronts in a Porous Medium. *The Journal of Physical Chemistry B*, 112(4):1170–1176, 2008. PMID: 18173259.
[61] Sandeep Saha, Severine Atis, Dominique Salin, and Laurent Talon. Phase diagram of sustained wave fronts opposing the flow in disordered porous media. *EPL (Europhysics Letters)*, 101(3):38003, 2013.
[62] Kendall Atkinson and Weimin Han. *Theoretical Numerical Analysis: A Functional Analysis Framework*, volume 39. Springer, 2009.
[63] Dongbin Xiu and Daniel M. Tartakovsky. Numerical Methods for Differential Equations in Random Domains. *SIAM Journal on Scientific Computing*, 28(3):1167–1185, 2006.
[64] Daniel M. Tartakovsky and Dongbin Xiu. Stochastic analysis of transport in tubes with rough walls. *Journal of Computational Physics*, 217(1):248–259, 2006. Uncertainty Quantification in Simulation Science.
[65] Claudio Canuto and Tomas Kozubek. A fictitious domain approach to the numerical solution of PDEs in stochastic domains. *Numerische Mathematik*, 107(2):257, May 2007.
[66] A. Nouy, A. Clément, F. Schoefs, and N. Moës. An extended stochastic finite element method for solving stochastic partial differential equations on random domains. *Computer Methods in Applied Mechanics and Engineering*, 197(51):4663–4682, 2008.
[67] Anthony Nouy, Mathilde Chevreuil, and Elias Safatly. Fictitious domain method and separated representations for the solution of boundary value problems on uncertain parameterized domains. *Computer Methods in Applied Mechanics and Engineering*, 200(45–46):3066–3082, 2011.
[68] Faker Ben Belgacem. The Mortar finite element method with Lagrange multipliers. *Numerische Mathematik*, 84(2):173–197, 1999.
[69] Barbara I. Wohlmuth. *Discretization Methods and Iterative Solvers Based on Domain Decomposition*, volume 17 of *Lecture Notes in Computational Science and Engineering*. Springer, Berlin, New York, 2001.
[70] C. Kim, R. Lazarov, J. Pasciak, and P. Vassilevski. Multiplier Spaces for the Mortar Finite Element Method in Three Dimensions. *SIAM Journal on Numerical Analysis*, 39(2):519–538, 2001.
[71] Mickaël Duval, Jean-Charles Passieux, Michel Salaün, and Stéphane Guinard. Non-intrusive Coupling: Recent Advances and Scalable Nonlinear Domain Decomposition. *Archives of Computational Methods in Engineering*, 23(1):17–38, Mar 2016.
[72] Hermann G. Matthies and Andreas Keese. Galerkin methods for linear and nonlinear elliptic stochastic partial differential equations. *Computer Methods in Applied Mechanics and Engineering*, 194(12-16):1295–1331, 2005.
[73] L. Giraldi, A. Litvinenko, D. Liu, H. Matthies, and A. Nouy. To Be or Not to Be Intrusive? The Solution of Parametric and Stochastic Equations—the "Plain Vanilla" Galerkin Case. *SIAM Journal on Scientific Computing*, 36(6):A2720–A2744, 2014.
[74] L. Giraldi, D. Liu, H. G. Matthies, and A. Nouy. To Be or Not to be Intrusive? The Solution of Parametric and Stochastic Equations—Proper Generalized Decomposition. *SIAM Journal on Scientific Computing*, 37(1):A347–A368, 2015.
[75] Abdellah Chkifa, Albert Cohen, Giovanni Migliorati, Fabio Nobile, and Raúl Tempone. Discrete least squares polynomial approximation with random evaluations - application to parametric and stochastic elliptic PDEs. *ESAIM: M2AN*, 49(3):815–837, 2015.

[76] Maurice S. Bartlett. An Inverse Matrix Adjustment Arising in Discriminant Analysis. *The Annals of Mathematical Statistics*, 22(1):107–111, 1951.

[77] Gavin C. Cawley and Nicola L.C. Talbot. Fast exact leave-one-out cross-validation of sparse least-squares support vector machines. *Neural Networks*, 17(10):1467–1475, 2004.

[78] Bruce M Irons and Robert C Tuck. A version of the Aitken accelerator for computer iteration. *International Journal for Numerical Methods in Engineering*, 1(3):275–277, 1969.

[79] Allan J. Macleod. Acceleration of vector sequences by multi-dimensional Δ^2 methods. *Communications in Applied Numerical Methods*, 2(4):385–392, 1986.

[80] Ulrich Küttler and Wolfgang A. Wall. Fixed-point fluid–structure interaction solvers with dynamic relaxation. *Computational Mechanics*, 43(1):61–72, 2008.

[81] Y.J. Liu, Q. Sun, and X.L. Fan. A non-intrusive global/local algorithm with non-matching interface: Derivation and numerical validation. *Computer Methods in Applied Mechanics and Engineering*, 277(0):81–103, 2014.

[82] Bruno Sudret. Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety*, 93(7):964–979, 2008.

[83] Olivier Chapelle, Vladimir Vapnik, and Yoshua Bengio. Model Selection for Small Sample Regression. *Machine Learning*, 48(1-3):9–23, 2002.

[84] Géraud Blatman and Bruno Sudret. Adaptive sparse polynomial chaos expansion based on least angle regression. *Journal of Computational Physics*, 230(6):2345–2367, 2011.

[85] Tomáš Roubíček. *Nonlinear Partial Differential Equations with Applications*, volume 153. Springer, 2005.