Multiscale numerical methods for the simulation of diffusion processes in random heterogeneous media with guaranteed accuracy

Daniel Alves Paladim

Supervisors: Prof. Stéphane Bordas

Dr. Pierre Kerfriden

A thesis submitted to the graduate school

in fulfilment of the requirements for the degree of

Doctor of Philosophy

Cardiff School of Engineering

June 2016

Summary

The possibility of combining several constituents to obtain properties that cannot be obtained with any of them alone, explains the growing proliferation of composites in mechanical structures. However, the modelling of such heterogeneous systems poses extreme challenges to computational mechanics. The direct simulation of the aforementioned gives rise to computational models that are extremely expensive if not impossible to solve.

Through homogenisation, the excessive computational burden is eliminated by separating the two scales (the scale of the constituents and the scale of the structure). Nonetheless, the hypotheses under which homogenisation applies are usually violated. Traditional homogenisation schemes provide no means to quantify this error.

The first contribution of this thesis is the development of a method to quantify the homogenisation error. In this method, the heterogeneous medium is represented by a stochastic partial differential equation where each realisation corresponds to a particle layout. This representation allows us to derive guaranteed error estimates with a low computational cost. The effectivity (ratio between true error and estimate) is characterised and a relation is established between the error estimates and classical results in micromechanics. Moreover, a strategy to reduce the homogenisation error is presented.

The second contribution of this thesis is made by developing a numerical method with guaranteed error bounds that directly approximates the solution of heterogeneous models by using shape functions that incorporate information of the microscale. The construction of those shape functions resembles the methods of computational homogenisation where microscale boundary value problems are solved to obtain homogenised properties.

DECLARATION

This work has not previously been accepted in substance for any degree and is not concurrently submitted in candidature for any degree.

Signed.....(candidate) Date:.....

STATEMENT 1

This thesis is being submitted in partial fulfillment of the requirements for the degree of PhD.

Signed.....(candidate) Date:.....

STATEMENT 2

This work is the result of my own independent work/investigation, except where otherwise stated. Other sources are acknowledged by explicit references.

Signed.....(candidate) Date:.....

DECLARATION

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed.....(candidate) Date:.....

Contents

Introduction

The recent Airbus 380 and Boeing 787 Dreamliner [59] are just two of the success stories of the use of composite materials. The possibility of combining several constituents to obtain properties that cannot be obtained with any of them alone, explains their growing proliferation in mechanical structures. However, the modelling of such systems poses extreme challenges to computational mechanics due to the fast spatial variation of the material properties. The direct simulation of the aforementioned gives rise to computational models that are extremely expensive if not impossible to solve.

Through an additional step, homogenisation, the excessive computational burden is eliminated. Homogenisation allows us to represent the multiphase medium by an equivalent single phase medium which is then used to model the structure. In other words, the original problem where two scales were present and tightly linked (the scale of the constituents and the scale of the structure), is split into two tractable problems. Homogenisation only applies when the scales are separable *i.e.* the smallest volume of the microstructure that captures its behaviour as a continuum has a size that is negligible when compared to the structure. This material unit is called representative volume element (RVE) and in practice, this hypothesis is violated in regions with high variation of the gradients (namely cracks, notches, sharp corners...). In these regions, homogenisation must be abandoned in favour of the microstructure, otherwise a significant error is committed. Therefore, the importance of identifying those regions is paramount.

In [7,53,56,81,90,91], different heuristic methods, usually based on the second derivatives of the field variables, were developed to identify those critical regions. In contrast, a few others have taken a different approach and focused on developing guaranteed measures of the error, meaning that those measures/estimates strictly lower and upper bound the error. The work of Oden et al. in [61, 65–68, 86, 87, 93, 103] is particularly relevant in this regard. Oden et al. considered two models, the homogenised model (or surrogate model) of the structure and the intractable heterogeneous model (or reference/"true" model). An approximate solution of the latter is obtained by solving the former. The error introduced by homogenisation is estimated by adapting traditional a posteriori error estimates (see [5, 6, 18, 34, 51, 89, 94] for a review) that have been previously used for

estimating the discretisation error in global (*i.e.* error in energy norm) and local quantities (*i.e.* the so called quantities of interest). Yet, the main drawback of this approach is that the calculation of error estimates strongly couples the two scales again and therefore is unaffordable for large-scale composites.

In this thesis, we continue and extend this work with the aim of developing robust numerical models for the modelling of heterogeneous materials. We aim to emulate the main advantage of homogenisation (scale separation) at the analysis and at the postprocessing stages where error is estimated. We present two strategies.

In the first strategy, we also adopt the traditional framework of two models. However, we represent the heterogeneous problem by a stochastic PDE. Each realisation of this PDE represents a different layout of particles/heterogeneities. This allows for the modelling of problems where a complete description of the structure is not available (*i.e.* the position of every particle is not known) but instead a partial description is on hand (*i.e.* the volume fraction, the physical properties of the constituents and their shape is available, but not the exact layout). The second key ingredient is to derive the error estimates through the Prager-Synge hypercircle method [78] also known as the constitutive relation error (CRE) [50]. The errors are derived fundamentally to measure error in terms the average response of the structure without paying special attention to any particular layout. As we will show, this choice of "true" model and "average measure" will allow us also to achieve scale separation in the estimation of the error. There are significant differences between these error estimates and the ones from the preceding work which we now highlight. Besides the model error, the estimates also account for the discretisation error. More importantly, their effectivity (the ratio between the exact error and the estimate) can be characterised and optimised analytically. The optima turn out to be linked to classical results of micromechanics.

This summarises our efforts in this first strategy regarding error quantification. Regarding error reduction, we introduce error indicators which allows us to identify the regions that contribute most to the error. In those regions, homogenisation is substituted by the microstructure. The fact that the microscale is only introduced locally prevents the

computational cost from growing out of control.

In the second strategy, we abandon the homogenised model and try directly to approach the solution of the true model. The problem is rewritten to allow for a new class of shape functions which incorporate information of the microscale. Those shape functions are added to our approximation through partition of unity, the same method used in the extended/generalised finite element method [12,60] to enrich the solution with new functions. Those shape functions represent the solution of an RVE and their construction resembles the methods of computational homogenisation where the effective (homogenised) properties are extracted. Indeed, from the RVEs, several constants (microstructure parameters) are extracted and are the only means through which the micro and macroscale communicate. Consequently, we retain scale separation in the determination of the response of the structure and quantification of error. In comparison to model adaptivity, this approach has two key advantages. Firstly, it eliminates the need of picking the "right" homogenisation scheme whereas in model adaptivity the approximation quality and error estimates are sensitive to this choice. Secondly, certain geometric constraints may complicate the application of model adaptivity in certain regions whereas here the method is always applied globally and is indifferent to the geometry.

In summary, the novel contributions of this thesis are:

- a framework to estimate the modelling error due to homogenisation where the reference model is stochastic to account for all the possible particle layouts. This allows the development of guaranteed error estimates for the average error in local and global quantities where the scale separation is retained. Furthermore, the analytical characterisation and optimisation of the estimates is given. Finally, a strategy to reduce error based on the local solution of the microstructure is introduced.

- a numerical method that allows the use of shape functions that incorporate information about the microstructure while at the same time retaining the separation of scales. The means to construct such shape functions, the model approximation and the estimation of error are detailed. In addition, a formulation of the theory suitable for

computer implementation is presented.

The thesis is divided in five chapters. In the first three, the relevant literature is reviewed.

- In chapter 1, the classical theory of a posteriori error estimation is reviewed. Special attention is paid to the Prager-Synge theorem since it is the basis for the bounds later derived.
- Chapter 2 is dedicated to micromechanics and multiscale methods. We focus on homogenisation methods based on the Hill-Mandel approach.
- In chapter 3, the theory and implementation of stress finite elements is reviewed. Stress finite elements are a key ingredient in our approach to error estimation.

In the remaining chapters, we present our numerical models for the robust modelling of heterogeneous materials.

- In chapter 4, the novel error bounds for homogenisation are presented. The theory is complemented with numerical examples.
- In chapter 5, the strategy for direct approximation of the solution of the reference model is presented. As in the previous chapter, the theory is illustrated with examples. In addition, possible extensions of the method are discussed.

Papers and presentations

Part of the work contained in this thesis has also appeared in the following papers and presentations:

International journals

- Daniel Paladim, José de Almeida, Stéphane Bordas, and Pierre Kerfriden. Guaranteed error bounds in homogenisation: an optimum stochastic approach to preserve the numerical separation of scales. *International Journal for Numerical Methods in Engineering*, 2016, Accepted for publication

- Daniel Paladim, Stéphane Bordas, and Pierre Kerfriden. A two scale error bound for the modelling random particulate materials. In preparation

Conference papers and presentations

- Daniel Paladim, Pierre Kerfriden, and Stéphane Bordas. Efficient modelling of random heterogenenous materials with an uniform probability density function. Presented in 11th World Congress on Computational Mechanics, 2014
- Pierre Kerfriden, Daniel Paladim, and Stéphane Bordas. Homogenisation methods with guaranteed accuracy: quantifying the scale separability. Presented in 5th International Conference on Computational Methods (ICCM2014), 2014
- Daniel Paladim, Pierre Kerfriden, José de Almeida, and Stéphane Bordas. An adaptive scheme for homogenised domains. Presented in International Conference on Adaptive Modelling and Simulation, 2015
- Daniel Paladim, Pierre Kerfriden, José de Almeida, Mathilde Chevreuil, and Stéphane Bordas. Advances in error estimation for homogenisation. Presented in 13th US National Congress on Computational Mechanics, 2015
- Daniel Paladim, José de Almeida, Stéphane Bordas, and Pierre Kerfriden. Guaranteed error bounds for the homogenisation of random materials. Presented in ACME-UK, 2016
- Daniel Paladim, Pierre Kerfriden, and Stéphane Bordas. Guaranteed error bounds for rve based homogenisation. Presented in VII European Congress on Computational Methods in Applied Sciences and Engineering, 2016
- Pierre Kerfriden and Daniel Paladim. A new homogenisation scheme with certified accuracy for random media. Presented in VII European Congress on Computational Methods in Applied Sciences and Engineering, 2016

Chapter 1

A posteriori error bounds for approximation methods applied to elliptic PDEs

1.1 Introduction

In the general case, the finite element approximation does not coincide with the exact solution of the problem. The assessment of this discrepancy is the objective of error estimation. Error estimates are classified as a priori and posteriori. The former describes the convergence properties of a finite element and as suggested by the name are available before the analysis. A posteriori error estimates aim to quantify the discretisation error and/or guide the mesh refinement and, are determined for a particular approximation. For instance, for a linear elliptic problem solved using linear triangular elements, an a priori estimate establishes that the error in energy-norm converges linearly with the element size h,

$$\|u - u^h\| \leq Ch \tag{1.1}$$

where u is the exact solution, u^h is an approximation with element size h and C is a constant independent of the element size. In contrast, an a posteriori estimate aims to quantify or bound the value of the error in energy norm, *i.e.*

$$\|u - u^h\| \approx \eta \quad \text{or} \quad \|u - u^h\| \leq \eta \tag{1.2}$$

where η is a computable constant; and/or also indicate the areas of the domain that contribute most to the error.

In this chapter, we focus on a posteriori error estimates. Those are usually further classified in three categories [18]: recovery, explicit and implicit methods. The recovery methods are covered in section 1.2. Explicit methods are covered in the following section. Implicit methods are covered in section 1.4 with special emphasis since they are used extensively in chapters 4 and 5. The chapter closes with an overview of goal oriented error estimation.

The discussion in this chapter is far from complete, we direct the interested reader to Ainsworth and Oden [5, 6], Verfürth [94], Grätsch and Bathe [34], Ladevèze [51], Stein and Rüter [89], and Chamoin and Diez [18].

1.1.1 Model problem

Before describing the different error estimators, we introduce the notation and the model problem that is going to be used throughout the chapter.

We consider the problem of stationary heat conduction in a body Ω defined in a subset of $\mathbb{R}^d$ $(d = 1, 2, 3)$. The boundary of this domain is denoted by Γ whilst its outward unit normal is denoted by $\boldsymbol{n}$. The boundary can be further divided in two parts Γ_N and Γ_D, Neumann and Dirichlet boundary respectively, such that $\Gamma_D \neq \emptyset$, $\Gamma_D \cup \Gamma_N = \Gamma$ and $\Gamma_D \cap \Gamma_N = \emptyset$. Fluxes g are prescribed on Γ_N, temperatures D are prescribed on Γ_D and the source term f accounts for the internal heat generation in the interior of the domain. The conductivity is denoted by k. The strong form of the problem reads

Find a temperature field $u \in \mathscr{C}^2(\Omega)$ and a flux field $\boldsymbol{q} \in \left[\mathscr{C}^1(\Omega)\right]^d$ such that

$$\nabla \cdot \boldsymbol{q}(\boldsymbol{x}) = f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Omega \qquad \text{conservation eq.} \tag{1.3}$$

$$\boldsymbol{q}(\boldsymbol{x}) \cdot \boldsymbol{n} = g(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_N \qquad \text{prescribed fluxes} \tag{1.4}$$

$$u(\boldsymbol{x}) = D(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_D \qquad \text{prescribed temperatures} \tag{1.5}$$

$$-k\nabla u(\boldsymbol{x}) = \boldsymbol{q}(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Omega \qquad \text{constitutive relation} \tag{1.6}$$

where $\mathscr{C}^n(\Omega)$ denotes the space of functions with n continuous derivatives. The corresponding weak formulation of this problem for the temperature field u reads[1]

For all $v \in \mathcal{U}_0(\Omega)$, find $u \in \mathcal{U}(\Omega)$ such that

$$a(u, v) := \int_\Omega k(\boldsymbol{x}) \nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \, d\Omega = \int_\Omega f(\boldsymbol{x}) v(\boldsymbol{x}) \, d\Omega - \int_{\Gamma_N} g(\boldsymbol{x}) v(\boldsymbol{x}) \, d\Gamma =: l(v) \tag{1.7}$$

[1] The weak form of the problem for the flux is discussed in chapter 3.

where

$$\mathcal{U}(\Omega) = \{u \in H^1(\Omega) : u(\boldsymbol{x}) = D(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_D\}$$
$$\mathcal{U}_0(\Omega) = \{u \in H^1(\Omega) : u(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Gamma_D\}$$

and $H^1(\Omega)$ is the Sobolev space of square integrable functions with square integrable generalised first derivatives on Ω. The energy norm mentioned earlier is defined by

$$\|u\| = \sqrt{a(u,u)} \tag{1.8}$$

A conforming approximation of the solution is denoted by u^h and the error field is denoted by $e = u - u^h$.

1.2 Recovery methods

The recovery methods are based on the original work of Zienkiewicz and Zhu [100]. The premise which was followed in all their subsequent work was to postprocess the flux/stress field of the approximation u^h in order to obtain a better approximation $\boldsymbol{q}^*$ and then error is estimated by

$$\|e\| \approx \gamma := \sqrt{\int_\Omega k^{-1}(\boldsymbol{q}^* + k\nabla u^h)^2 \, d\Omega}. \tag{1.9}$$

In [100] the postprocessing technique used is called nodal averaging. The main idea is to reconstruct a continuous flux/stress field from the discontinuous approximated flux $-k\nabla u^h$ by averaging its value at each node and then using the finite element shape functions to interpolate.

This technique was further improved in [101, 102] where the super-convergent patch recovery (SPR) was introduced. In the SPR method, the flux field is reconstructed patch by patch. In a first step, a surface is fitted through the super-convergent points of the patch using least squares. Let $\boldsymbol{q}_i^*$ be the value of this surface at the node in the centre of the patch. Then, the recovery field $\boldsymbol{q}^*$ is built by FE interpolation, *i.e.*

$$\boldsymbol{q}^*(\boldsymbol{x}) = \sum_{i=1}^{n_{\text{nodes}}} N_i(\boldsymbol{x}) \boldsymbol{q}_i^*. \tag{1.10}$$

For certain classes of problems, the SPR technique is asymptotically exact [98] meaning that as the element size tends to 0 ($h \to 0$), the effectivity tends to 1 ($\frac{\gamma}{\|e\|} \to 1$).

More recently, Ródenas et al. developed the constrained SPR (SPR-C) method [84] where the quality of the recovered field is improved by ensuring the local satisfaction of the equilibrium equations. In [33], the SPR-C is used to estimate error in quantities of interest.

The main attractiveness of this method is its simplicity and easiness of implementation, especially in its original version. However, the error estimates produced by the recovery method are not guaranteed.

1.3 Explicit methods

The naming of these methods as explicit is related to the fact that the computation of these error bounds does not involve the solution of any additional problem. The bases of those methods were established in the works of Babuska and Rheinboldt [9], Babuska and Miller [8] and Kelly [43]. The error bounds obtained usually depend on constants that are hard to estimate and/or the estimation of those constants result in bounds that are very pessimistic.

For instance, the following estimate can be derived for the error in the energy norm using results of interpolation theory (see [6] or [58] for the derivation),

$$\|u - u^h\| \leq C \left(\sum_{e=1}^{n_{\text{elm}}} h_e^2 \|f + k\Delta u^h\|_{L_2(\Omega_e)}^2 + \sum_{e=1}^{n_{\text{edges}}} h_e \|R\|_{L_2(\gamma_e)}^2 \right)^{1/2} \tag{1.11}$$

where C is a constant independent of the element size, h_i is the element size, γ_e is an element edge,

$$R(\boldsymbol{x}) = \begin{cases} g(\boldsymbol{x}) + k\nabla u^h(\boldsymbol{x}) \cdot \boldsymbol{n} & \forall \boldsymbol{x} \in \Gamma_N \\ [\![k\nabla u^h]\!](\boldsymbol{x}) \cdot \boldsymbol{n} & \text{otherwise} \end{cases}, \tag{1.12}$$

and $[\![\cdot]\!]$ denotes the jump. In eq. (1.11), we see that the error was decomposed in two terms, one term that accounts for how our approximation fails to satisfy the conservation equation and a term that accounts for the discontinuity on the element edges and non satisfaction of the Neumann boundary conditions. Even if the bound is not computable

due to the constant, it is still useful to guide the mesh refinement process. The areas of the domain that contribute most to the error indicator should be refined.

1.4 Implicit methods

The computation of implicit error estimates requires the solution of an additional problem. The reward for this additional computational cost is a computable bound (it does not depend on constants) that is usually guaranteed to hold. We review in great detail the Prager-Synge hypercircle theorem which is the base for the error estimates presented in chapters 4 and 5.

1.4.1 The Prager-Synge hypercircle theorem

The Prager-Synge hypercircle theorem [78] introduces an equality which can be used to calculate guaranteed error bounds. Before introducing the theorem, we present two definitions which classify our approximations.

Loosely speaking, a temperature field u^h is kinematically admissible (KA) if it fulfils the prescribed Dirichlet boundary conditions and it is also continuous. We will also say that a flux field $\boldsymbol{q}^h$ is statically admissible (SA) if it fulfils

$$\nabla \cdot \boldsymbol{q}^h(\boldsymbol{x}) = f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Omega \tag{1.13}$$

$$\boldsymbol{q}^h(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x}) = g(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_N \tag{1.14}$$

More accurately and in the language of functional analysis, KA fields are elements of $\mathcal{U}(\Omega)$[2] whilst SA fields are elements of

$$\mathcal{S}(\Omega) = \left\{ \boldsymbol{q} \in H^{\text{div}}(\Omega) | \nabla \cdot \boldsymbol{q} = f(\boldsymbol{x}) \, \forall \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{q}(\boldsymbol{x}) \cdot \boldsymbol{n} = g(\boldsymbol{x}) \, \forall \boldsymbol{x} \in \Gamma_N \right\} \tag{1.15}$$

[2]Strictly speaking, these functions are not necessarily continuous. Elements of $H^1(\Omega)$ can be be discontinuous in a point, but cannot have more severe discontinuities such as a jump over a line/surface. The theorems that guarantee existence and uniqueness of solutions of problems like 1.7, requires us to work in such spaces. A more in depth discussion of the topic can be found in [14].

where $H^{\text{div}}(\Omega)$ is the Sobolev space of vectorial square integrable functions and with generalised square integrable divergence, *i.e.*

$$H^{\text{div}}(\Omega) = \left\{ \boldsymbol{q} | \int_\Omega \boldsymbol{q} \cdot \boldsymbol{q} \, d\Omega < \infty \text{ and } \int_\Omega (\nabla \cdot \boldsymbol{q})^2 \, d\Omega < \infty \right\}. \tag{1.16}$$

The naming originates from the original work of Veubeke [99] in linear elasticity. Kinematically admissible fields correspond to the displacement fields that fulfil the prescribed displacements whilst the statically admissible fields correspond to stress fields that fulfil the equilibrium equations.

With those definitions, the theorem reads

Theorem 1.1. *If u^h is a kinematically admissible field and $\boldsymbol{q}^h$ is a statically admissible field then*

$$\|u - u^h\|^2 + \| - k\nabla u - \boldsymbol{q}^h\|^2_{k^{-1}} = \|\boldsymbol{q}^h + k\nabla u^h\|^2_{k^{-1}}. \tag{1.17}$$

where

$$\|\boldsymbol{q}\|_{k^{-1}} := \sqrt{\int_\Omega k^{-1}\boldsymbol{q} \cdot \boldsymbol{q} \, d\Omega}. \tag{1.18}$$

Also, the following identity holds for the arithmetic average of the flux fields, $\frac{\boldsymbol{q}^h - k\nabla u^h}{2}$,

$$\left\| \boldsymbol{q} - \frac{\boldsymbol{q}^h - k\nabla u^h}{2} \right\|_{k^{-1}} = \frac{1}{2}\|\boldsymbol{q}^h + k\nabla u^h\|_{k^{-1}}. \tag{1.19}$$

Proof. The proof is done in 3 steps. Firstly, we are going to prove that the error of a equilibrated flux field is orthogonal to the flux of the error of a compatible solution,

$$-k\nabla u - \boldsymbol{q}^h \perp -k\nabla u + k\nabla u^h, \tag{1.20}$$

in the k^{-1}-inner product

$$\langle \boldsymbol{v}, \boldsymbol{w} \rangle_{k^{-1}} := \int_\Omega k^{-1}\boldsymbol{v} \cdot \boldsymbol{w} \, d\Omega \tag{1.21}$$

Since $\boldsymbol{q}^h$ is SA,

$$\nabla \cdot \boldsymbol{q}^h = f. \tag{1.22}$$

Multiplying by $v \in \mathcal{U}_0(\Omega)$ and integrating over the domain,

$$\int_\Omega \nabla \cdot \boldsymbol{q}^h v \, d\Omega = \int_\Omega v f \, d\Omega \qquad \forall v \in \mathcal{U}_0(\Omega) \tag{1.23}$$

And now integrating the left hand side by parts,

$$\int_\Gamma v\boldsymbol{q}^h\cdot\boldsymbol{n}\,d\Gamma-\int_\Omega \boldsymbol{q}^h\cdot\nabla v\,d\Omega=\int_\Omega vf\,d\Omega\qquad\forall v\in\mathcal{U}_0(\Omega)\tag{1.24}$$

Since $v=0$ on Γ_D and $\boldsymbol{q}^h\cdot\boldsymbol{n}=g$ on Γ_N, we obtain that

$$-\int_\Omega \boldsymbol{q}^h\cdot\nabla v\,d\Omega=\int_\Omega vf\,d\Omega-\int_\Gamma vg\cdot\boldsymbol{n}\,d\Gamma=l(v)\qquad\forall v\in\mathcal{U}_0(\Omega)\tag{1.25}$$

In addition, since u is the solution of eq. (1.7),

$$\int_\Omega k\nabla u\cdot\nabla v\,d\Omega=l(v)\qquad\forall v\in\mathcal{U}_0(\Omega)\tag{1.26}$$

Subtracting both equations, we obtain

$$\int_\Omega(\boldsymbol{q}^h+k\nabla u)\cdot\nabla v\,d\Omega=0\qquad\forall v\in\mathcal{U}_0(\Omega)$$

and the result follows from setting $v=-u+u^h$,

$$\langle\boldsymbol{q}^h+k\nabla u,-k\nabla u+k\nabla u^h\rangle_{k^{-1}}=0$$

Secondly, we observe that,

$$\|-k\nabla u+k\nabla u^h\|_{k^{-1}}=\|e\|\tag{1.27}$$

Due to the orthogonality of the functions, we use the Pythagoras theorem to prove eq. (1.17)

$$\begin{aligned}\|\boldsymbol{q}^h+k\nabla u^h\|^2_{k^{-1}}&=\|\boldsymbol{q}^h+k\nabla u-k\nabla u+k\nabla u^h\|^2_{k^{-1}}=\\&=\|\boldsymbol{q}^h+k\nabla u\|^2_{k^{-1}}+\|-k\nabla u+k\nabla u^h\|^2_{k^{-1}}=\|\boldsymbol{q}^h+k\nabla u\|^2_{k^{-1}}+\|e\|^2\end{aligned}$$

The average result in eq. (1.19) follows from the identity just proved,

$$\begin{aligned}\left\|\boldsymbol{q}-\frac{\boldsymbol{q}^h-k\nabla u^h}{2}\right\|^2_{k^{-1}}&=\left\|\frac{1}{2}(-k\nabla u-\boldsymbol{q}^h)+\frac{1}{2}(-k\nabla u+k\nabla u^h)\right\|\\&=\frac{1}{4}\|-k\nabla u-\boldsymbol{q}^h\|^2_{k^{-1}}+\frac{1}{4}\|u-u^h\|^2=\frac{1}{4}\|\boldsymbol{q}^h+k\nabla u^h\|^2_{k^{-1}}\end{aligned}$$

□

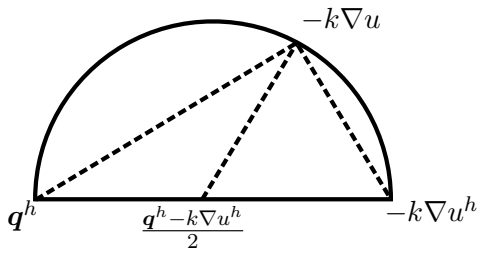


Figure 1.1: Prager-Synge hypercircle

From now on, we will denote the term $\|\boldsymbol{q}^h + k\nabla u^h\|_{k^{-1}}$ by η. This term is usually called the constitutive relation error [50] (CRE) since the pair of fields $(u^h, \boldsymbol{q}^h)$ fulfil all the equations

$$\nabla \cdot \boldsymbol{q}^h(\boldsymbol{x}) = f(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Omega \tag{1.28}$$

$$\boldsymbol{q}^h(\boldsymbol{x}) \cdot \boldsymbol{n} = g(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_N \tag{1.29}$$

$$u^h(\boldsymbol{x}) = D(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_D \tag{1.30}$$

except the constitutive relation (unless the pair is the exact solution),

$$\boldsymbol{q}^h \neq -k\nabla u^h \quad \forall \boldsymbol{x} \in \Omega. \tag{1.31}$$

Loosely speaking, the theorem states that sum of the squares two errors, the error in the gradient field and the error in the flux field, equals the square of the error in the constitutive relation.

It is also worth mentioning the geometrical interpretation of the theorem which gives the theorem its name. Thales' theorem [1] (not to be confused with Thales theorem for similar triangles) states that A, B and C are points on a circle and $\overline{BC}$ is a diameter if and only if the segments $\overline{AB}$ and $\overline{AC}$ form a right-angle. Since $-k\nabla u - \boldsymbol{q}^h$ is orthogonal to $-k\nabla u + k\nabla u^h$, we can identify A with $-k\nabla u$ and B and C with $\boldsymbol{q}^h$ and $-k\nabla u^h$ (see fig. 1.1). The average approximation coincides with the centre of the circle and its distance to $-k\nabla u$, the radius, is half of the distance between $\boldsymbol{q}^h$ and $-k\nabla u^h$, the diameter.

With those comments, we proceed to derive bounds using the Prager-Synge theorem:

Corollary. *If u^h is a kinematically admissible field and $\boldsymbol{q}^h$ is a statically admissible field, then*

$$\|u - u^h\| \leq \eta \tag{1.32}$$

and

$$\| - k\nabla u - \boldsymbol{q}^h\|_{k^{-1}} \leq \eta \tag{1.33}$$

Proof. Both inequalities follow from the non-negativity of the terms in eq. (1.17). □

Hence, the CRE is an error bound for the error in energy norm. Its effectivity is going to be controlled by the quality of flux field,

$$\text{Effectivity} = \frac{\text{Upper bound}}{\text{Exact error}} = \frac{\eta}{\|e\|} = \sqrt{1 + \frac{\| - k\nabla u - \boldsymbol{q}^h\|^2_{k^{-1}}}{\|e\|^2}}. \tag{1.34}$$

Assuming that both error norms are of the same orders, effectivities of $\sqrt{2}$ are to be expected.

1.4.2 Construction of the SA fields. The EET method

We have seen that the error bound presented so far requires the construction of a kinematically admissible field and a statically admissible field. The displacement formulation of the finite element method generates approximations of the field u that are KA whilst the stress formulations results in SA approximations. Due to widespread popularity of displacement formulations, normally we would like to asses the quality of KA field u^h which requires an SA field $\boldsymbol{q}^h$.

The field $\boldsymbol{q}^h$ can be built independently of u^h by using stress finite elements as proposed in [22, 23, 44, 99]. We describe in detail the approach introduced [23] in the chapter 3. The main disadvantage of these approaches is that they involve the solution of a new problem with a cost comparable to the obtention of u^h.

On the other hand, several methods exist where the field $\boldsymbol{q}^h$ is built from the KA approximation. In [50], Ladeveze et al. proposed the element equilibration technique. In this method, the SA field is constructed in two steps. In a first step, an equilibrated field of fluxes (tractions) is defined on the element boundaries. The construction of this field is

done patch by patch and it never involves a global problem. In the second step, using the fluxes previously determined as boundary conditions, the fluxes on the interior of each of the elements are determined. Both fields are linked by the so called *prolongation condition* which requires

$$\int_{\Omega_e} \boldsymbol{q}^h \nabla N_i = -\int_{\Omega_e} k \nabla u^h \nabla N_i \qquad i = 1...n_{\text{nodes}} \tag{1.35}$$

where N_i denotes the shape function associated with node i. This holds for each element e. The prolongation condition requires equality of the virtual work done by both fields for all the shape functions restricted to an element and it ensures self equilibration of element problems. In the paper, it is also proved that the bound produced lower bounds the error,

$$\eta \leq C\|e\| \tag{1.36}$$

where C is a constant independent of the mesh size. This means that the bound decreases when the solution is refined mirroring the actual behaviour of the error.

1.4.3 Flux-free approximation

Another strategy to estimate the error is the so called flux free approximation proposed in [76] and inspired by the work developed in [16, 57, 63]. The main idea is to transform the problem that characterises the error

$$a(e, v) = R(v) := l(v) - a(u^h, v) \quad \forall v \in \mathcal{U}_0(\Omega) \tag{1.37}$$

in a collection of local problems defined in each of the patches ω_i

$$a_{\omega_i}(e_{\omega_i}, v) = R(N_i v) \quad \forall v \in \mathcal{U}_0(\Omega) \cap H^1(\Omega_{\omega_i}) \tag{1.38}$$

such that $e = \sum_{i=1}^{n_{\text{patches}}} e_{\omega_i}$ and where a_{ω_i} is the restriction of the bilinear form to the patch ω_i and N_i is the shape function whose support coincides with the patch. Each of the fields e_{ω_i} is approximated by mesh refinement of the patch. It is worth noting that in the context of linear elasticity, in the case that linear elements are used to approximate u^h, the right hand side eq. (1.38) must be modified to ensure the solvability of the local problems. The

solvability is recovered by subtracting from the test functions, their nodal projection over the patch, *i.e.*

$$R(N_i(v - \pi^h v)) \tag{1.39}$$

where π^h denotes the nodal projection.

The main advantage of the flux free method is the easiness of its implementation. The reason being that the operations involved in its computation only involve the interior of elements and the data structures for implementation of FEM can be reused. In contrast, the implementation of the EET or stress FEM require new data structures to deal with the edges/facets. On the other hand, it is shown in [77] that the computational cost of the flux-free approximation in 3D is significantly greater than the computational cost of EET. Yet, we expect its computational cost to be smaller than the cost of hybrid stress FEM since the latter involves the solution of a full problem.

1.5 Goal oriented error estimation

Up to this point, all the error estimators and bounds presented were for the error in energy-norm. Nonetheless, engineers are often interested in local quantities (*e.g.* the traction on part of the boundary, average temperature in a subdomain...). This apparent contradiction is justified by the fact that the problem of estimating error in quantities of interest can be reduced through the introduction of the adjoint problem to a matter of estimating error in energy norm. The main ideas of those methods were set out by Becker and Rannacher in [11], by Cirak and Ramm in [19] and by Oden and Prudhomme in [79]. Let s be a linear functional that extracts the desired quantity of interest, *i.e.* we are interested in estimating $s(u)$. We start by introducing the adjoint problem,

For all $v(\boldsymbol{x}) \in \mathcal{U}_0(\Omega)$, find $\phi(\boldsymbol{x}) \in \mathcal{U}_0(\Omega)$ such that,

$$a(\phi, v) = s(v) \tag{1.40}$$

Using the linearity of the functional and the definition of the adjoint problem, it follows that

$$s(u) - s(u^h) = s(e) = a(\phi, e) \tag{1.41}$$

since $e \in \mathcal{U}_0(\Omega)$ because $u^h \in \mathcal{U}(\Omega)$. Now let ϕ^h be an approximation to the solution of the adjoint problem. Adding and subtracting $a(\phi^h, e)$, we obtain

$$s(u) - s(u^h) = a(\phi - \phi^h, e) + a(\phi^h, e) = a(\phi - \phi^h, e) + R(\phi^h). \tag{1.42}$$

By using the Cauchy-Schwarz inequality,

$$a(\phi - \phi^h, e) \leq \|\phi - \phi^h\|\|e\|, \tag{1.43}$$

or the polarisation identity,

$$a(\phi - \phi^h, e) = \frac{\|e + (\phi - \phi^h)\|^2 - \|e - (\phi - \phi^h)\|^2}{4}, \tag{1.44}$$

the bilinear form is bounded by/transformed into a pair of norms. These norms can be bounded by bounds for the error in the energy norm, meaning that error in QoIs can be estimated by means of estimates for the error in the energy norm. In practice, we recommend the second approach since Cauchy-Schwarz introduces an additional inequality and this results in less effective bounds. A detailed analysis is carried out in section 4.3.3.1 and appendix A.

We note that nonlinear quantities can be considered through linearisation of the functional and neglecting the higher order terms. This approach results in non guaranteed error bounds. Guaranteed error bounds for nonlinear quantities can be obtained by other means. In [48], Ladevèze et al. describe a method to construct guaranteed error bounds for nonlinear pointwise quantities, through the introduction of additional functions using partition of unity.

1.6 Conclusion

In this chapter, a short review of the main types of a posteriori error estimates was presented, namely recovery, explicit and implicit methods. The recovery methods are based

on smoothing the approximation and they are generally easy to implement. Nonetheless, the resulting error estimates are not strictly guaranteed. In contrast, explicit error estimates are guaranteed and their computational cost is low. However, they usually depend on constants that are not easily computable. Therefore, their use is normally restricted to guide the mesh refinement process. Finally, implicit estimates require the solution of a local or global problem, resulting in a higher computational cost. The bounds are fully computable and often guaranteed.

All the bounds and estimates discussed in sections 1.2 to 1.4 apply to the error in the energy norm. The relevance of the error in energy norm was shown at the end of the chapter. Even though this quantity is a global measure of the error and not always meaningful to the analyst, it was shown in section 1.5 that the estimation of error in local quantities can be transformed in a problem of estimation of the former.

Chapter 2

Multiscale methods for diffusion problems in heterogeneous media

2.1 Introduction

Depending on the scale of observation, we may see in a material different phases and not the continuum that we often use to study them. The broad area of mechanics which studies the laws that relate the layout of those phases (the microstructure) to the mechanical response at the macroscale is called micromechanics. Our practical interest in micromechanics is twofold. Firstly, it simplifies our models, in many cases intractable if we were to try to incorporate all the scales. Secondly, understanding the link between the microscale and the macroscale response permits us to optimise the microstructure to take advantage of each constituent.

One of the fundamental tools in micromechanics is homogenisation. Homogenisation seeks to find an equivalent homogeneous (continuous) representation of a heterogeneous material. In sections 2.2 and 2.3, we review two of the main approaches to homogenisation.

The theory of homogenisation is built on certain hypotheses that are frequently violated: separation of scales, periodicity... In those circumstances homogenisation does not apply and we must adopt strategies where the microstructure is part of our model. Here, we review two of those approaches, the multiscale finite element method and the variational multiscale method in sections 2.6 and 2.7 respectively. Unless stated otherwise, throughout the chapter, we use the same notation introduced in section 1.1.1.

The discussion in this chapter can be further complemented by [13, 42, 64, 104].

2.2 The Hill-Mandel approach

As already mentioned in the introduction, the excessive detail required and the fast oscillation of physical properties make the direct modelling of systems composed of heterogeneous materials intractable. Instead, to reduce the computational cost, a macroscopic model which contains a sufficient amount of information of the microscale is built. In the Hill-Mandel approach, the macroscopic variables $\nabla\bar{u}$ and $\bar{\boldsymbol{q}}$ are related to the microscopic

variables ∇u and $\boldsymbol{q}$ through spatial averages,

$$\nabla \bar{u} := \langle \nabla u \rangle := \frac{1}{|\Omega|} \int_{\Omega} \nabla u \, d\Omega$$
$$\bar{\boldsymbol{q}} := \langle \boldsymbol{q} \rangle = \frac{1}{|\Omega|} \int_{\Omega} \boldsymbol{q} \, d\Omega$$

Our aim is to find an effective conductivity $\bar{k}$ that links the macroscopic variables, *i.e.*

$$\bar{\boldsymbol{q}} = -\bar{k} \nabla \bar{u}. \tag{2.1}$$

We note that such a conductivity depends on the volume Ω where the spatial average is performed and hence this conductivity $\bar{k}$ is more appropriately called apparent conductivity. Yet, we are able to choose the volume Ω in such a manner that if another volume Ω' is chosen with same dimensions, the variation of $\bar{k}$ is negligible.

A volume that fulfils such properties is called representative volume element (RVE). It is usually cubic-shaped (or square-shaped) and it statistically represents the material, meaning that extracting this volume from other parts of the domains will similarly characterise the microstructure. We require the size of this sample, the size of the RVE, to be larger than that of the heterogeneities, so that the obtained characterisation is not significantly influenced by the particular particle layout. At the same time, its dimension should be small when compared to the size of the structure that we aim to study, so that the macroscale gradient varies slowly inside the RVE, *i.e.* the macroscale gradient is close to constant inside the RVEs.

Furthermore, besides the choice of the representative volume, we still need to specify the boundary conditions that are applied to it. The Hill's macrohomogeneity condition [37] specifies that boundary conditions that produce a pair of flux field and gradient field that satisfy

$$-\langle \boldsymbol{q} \cdot \nabla u \rangle = -\langle \boldsymbol{q} \rangle \cdot \langle \nabla u \rangle \tag{2.2}$$

are suitable for the estimation of $\bar{k}$. This condition requires that the work estimated using the macroscopic variables equals the average work performed by the microscopic variables. Furthermore, it implies that

$$\langle k \nabla u \cdot \nabla u \rangle = \bar{k} \langle \nabla u \rangle \cdot \langle \nabla u \rangle \tag{2.3}$$

which is the energetic definition of the effective properties.

In the remainder of the section, we introduce a series of results that allows us to establish the boundary conditions that fulfil the Hill's macrohomogeneity condition. We also present two analytical approaches for bounding the values of the effective conductivity.

2.2.1 Average stress and strain theorems

Two fundamental theorems in micromechanics are the average stress and average strain theorems. They relate the average stress and strain in a heterogeneous material to the tractions and displacements applied on the boundary. In this section, we present their application the heat equation, which relates the average flux and average gradient with the prescribed fluxes/temperatures on the boundary. Later, those results will allow us to establish precisely when the Hill macrohomogeneity condition holds.

Theorem 2.1. Average gradient *If the temperature field u is continuous, the average gradient can be expressed as a boundary integral, namely*

$$\langle \nabla u \rangle = \frac{1}{|\Omega|} \int_\Gamma u \mathbf{n} \, d\Gamma. \tag{2.4}$$

Proof. The result follows immediately from an application of the divergence theorem

$$\int_\Omega \frac{\partial u}{\partial x_i} \, d\Omega = \int_\Gamma u n_i \, d\Gamma \tag{2.5}$$

and noting that,

$$|\Omega| \langle \nabla u \rangle_i = \int_\Omega \frac{\partial u}{\partial x_i} \, d\Omega. \tag{2.6}$$

□

Theorem 2.2. Average flux *The average flux in a domain can be written as*

$$\langle \boldsymbol{q} \rangle = \frac{1}{|\Omega|} \int_\Gamma \boldsymbol{x} \boldsymbol{q}^T \boldsymbol{n} \, d\Gamma - \frac{1}{|\Omega|} \int_\Omega (\nabla \cdot \boldsymbol{q}) \boldsymbol{x} \, d\Omega. \tag{2.7}$$

Proof. We start by observing that

$$\nabla \cdot (\boldsymbol{q} \otimes \boldsymbol{x})_i = \frac{\partial q_j x_i}{\partial x_j} = \frac{\partial q_j}{\partial x_j} + \underbrace{q_j \frac{\partial x_i}{\partial x_j}}_{= q_j \delta_{ij}} = \frac{\partial q_j}{\partial x_j} + q_i. \tag{2.8}$$

In matrix notation,

$$\nabla \cdot (\boldsymbol{q} \otimes \boldsymbol{x}) = (\nabla \cdot \boldsymbol{q})\boldsymbol{x} + \boldsymbol{q}. \tag{2.9}$$

Taking the average in the previous identity,

$$\langle \boldsymbol{q} \rangle = \frac{1}{|\Omega|} \int_\Omega \nabla \cdot (\boldsymbol{q} \otimes \boldsymbol{x})\, d\Omega - \frac{1}{|\Omega|} \int_\Omega (\nabla \cdot \boldsymbol{q})\boldsymbol{x}\, d\Omega \tag{2.10}$$

and applying the divergence theorem to the first term on the right hand side[1]

$$\langle \boldsymbol{q} \rangle = \frac{1}{|\Omega|} \int_\Gamma \boldsymbol{x}\boldsymbol{q}^T \boldsymbol{n}\, d\Gamma - \frac{1}{|\Omega|} \int_\Omega (\nabla \cdot \boldsymbol{q})\boldsymbol{x}\, d\Omega \tag{2.11}$$

□

We can eliminate the domain integral in the average flux theorem by assuming that the source term is 0, *i.e.* $\nabla \cdot \boldsymbol{q} = 0$, giving an expression that only depends on the boundary,

$$\langle \boldsymbol{q} \rangle = \frac{1}{|\Omega|} \int_\Gamma \boldsymbol{x}\boldsymbol{q}^T \boldsymbol{n}\, d\Gamma. \tag{2.12}$$

For this reason, on the remainder of the section, we prescribe $\nabla \cdot \boldsymbol{q} = 0$.

2.2.2 Boundary conditions on RVEs

We review in this section three types of boundary conditions that guarantee that Hill-Mandel criterion holds.

A RVE is under kinematic uniform boundary conditions (KUBC) if we prescribe on its boundary,

$$u(\boldsymbol{x}) = \boldsymbol{U}\boldsymbol{x} \quad \forall \boldsymbol{x} \in \Gamma \tag{2.13}$$

with $\boldsymbol{U}$ constant. The following theorem shows that this constant is the average gradient.

Theorem 2.3. *If an RVE is under kinematically uniform boundary conditions (eq.* (2.13)*), then*

$$\langle \nabla u \rangle = \boldsymbol{U}. \tag{2.14}$$

[1] $\int_\Omega \nabla \cdot \boldsymbol{A}\, d\Omega = \int_\Gamma \boldsymbol{A}^T \boldsymbol{n}\, d\Gamma$.

Proof. Using the average temperature theorem,

$$|\Omega|\langle\nabla u\rangle_i = \int_\Gamma U_j x_j n_i \, d\Gamma = \int_\Omega U_j \frac{\partial x_j}{\partial x_i} \, d\Omega = \int_\Omega U_j \delta_{ij} \, d\Omega = |\Omega| U_i. \tag{2.15}$$

□

A boundary is under static uniform boundary conditions if we prescribe the following flux on its boundary,

$$\boldsymbol{q}(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x}) = \boldsymbol{Q} \cdot \boldsymbol{n}(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma \tag{2.16}$$

with $\boldsymbol{Q}$ constant. Similarly to $\boldsymbol{U}$, $\boldsymbol{Q}$ is the average flux:

Theorem 2.4. *If an RVE is under statically uniform boundary conditions (eq.* (2.16)*), then*

$$\langle \boldsymbol{q} \rangle = \boldsymbol{Q}. \tag{2.17}$$

Proof. The proof is very similar to the proof for KUBC. Applying the average flux theorem,

$$|\Omega|\langle \boldsymbol{q} \rangle_i = \int_\Gamma x_i n_j Q_j \, d\Gamma = Q_j \int_\Omega \frac{\partial x_i}{\partial x_j} \, d\Omega = |\Omega| Q_i. \tag{2.18}$$

□

Finally, a square/cubic RVE with the origin at its centre is under periodic boundary conditions (PBC) if we prescribe a temperature field in the form

$$u(\boldsymbol{x}) = \boldsymbol{U} \cdot \boldsymbol{x} + v(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Omega \tag{2.19}$$

with $\boldsymbol{U}$ constant, $v(\boldsymbol{x})$ is a function that take the same values on opposing faces of the domain and we also require $\boldsymbol{q} \cdot \boldsymbol{n}$ to take opposite values on opposing faces. It can be shown that $\boldsymbol{U}$ is the average gradient. The PBC can also be formulated in order to prescribe the average flux.

We now show the Hill-Mandel lemma, which allows us to prove that the boundary conditions described above are suitable for estimating the effective conductivity.

Theorem 2.5. Hill-Mandel lemma *If the BCs are SUBC then*

$$\langle \boldsymbol{q} \cdot \nabla \delta u \rangle = \langle \boldsymbol{q} \rangle \cdot \langle \nabla \delta u \rangle \tag{2.20}$$

where δu is any KA temperature field. Also, if the BCs are KUBC and $\delta \boldsymbol{q}$ is a given SA flux field, then

$$\langle \delta \boldsymbol{q} \cdot \nabla u \rangle = \langle \delta \boldsymbol{q} \rangle \cdot \langle \nabla u \rangle. \tag{2.21}$$

In words, the microscale virtual work equals the virtual work of the macroscale variables.

Proof. To prove the first part, we apply integration by parts followed by the average gradient theorem,

$$\int_\Omega \boldsymbol{q} \cdot \nabla \delta u \, d\Omega = \int_\Gamma \delta u \boldsymbol{q} \cdot \boldsymbol{n} \, d\Gamma - \underbrace{\int_\Omega \delta u \nabla \cdot \boldsymbol{q} \, d\Omega}_{=0} = \boldsymbol{Q} \cdot \int_\Gamma \delta u \boldsymbol{n} \, d\Gamma = \boldsymbol{Q} \cdot \langle \delta u \rangle |\Omega| = \langle \boldsymbol{q} \rangle \cdot \langle \delta u \rangle |\Omega| \tag{2.22}$$

In the proof for the KUBC, we use average flux theorem instead. □

The main consequence of the Hill-Mandel lemma is that fields under SUBC and KUBC fulfil the Hill macrohomogeneity condition, *i.e.*

$$\langle \boldsymbol{q} \cdot \nabla \bar{u} \rangle = \langle \boldsymbol{q} \rangle \cdot \langle \nabla \bar{u} \rangle \tag{2.23}$$

which follows immediately from setting $\delta u = u$ or $\delta \boldsymbol{q} = \boldsymbol{q}$ respectively. This also holds for fields under PBC but we omit the proof. This makes these three types of boundary conditions suitable for the estimation of the effective conductivity of the domain. In the next section, we use them to derive guaranteed bounds.

2.2.3 The Reuss and Voigt bounds

The Reuss and Voigt bounds establish a lower and an upper bound on the value of the effective conductivity $\bar{k}$. These bounds are not very sharp except for the case of small contrast between the different phases, however their computation only requires the knowledge of the material properties of the different phases and their respective volume fractions.

Theorem 2.6. Reuss and Voigt bounds *The following inequality holds for the effective conductivity*

$$\frac{1}{\langle k^{-1}\rangle} \leq \bar{k} \leq \langle k\rangle \tag{2.24}$$

Proof. The upper bound is derived by computing the work performed by the field $\nabla u - \langle \nabla u\rangle$,

$$0 \leq \int_\Omega k(\nabla u - \langle \nabla u\rangle)^2 \, d\Omega = \int_\Omega k\nabla u^2 \, d\Omega - 2\int_\Omega k\nabla u \cdot \langle \nabla u\rangle \, d\Omega + \int_\Omega k\langle \nabla u\rangle^2 \, d\Omega \tag{2.25}$$

We manipulate the first term to obtain,

$$\int_\Omega k\nabla u^2 \, d\Omega = -\int_\Omega \boldsymbol{q} \cdot \nabla u \, d\Omega = |\Omega|\langle \boldsymbol{q} \cdot \nabla u\rangle = -|\Omega|\langle \boldsymbol{q}\rangle \cdot \langle \nabla u\rangle = |\Omega|\bar{k}\langle \nabla u\rangle^2 \tag{2.26}$$

where we applied the definition of flux, followed by Hill's criterion and the definition of the effective conductivity. By applying again the definition of the effective conductivity, the second term becomes,

$$-2\int_\Omega k\nabla u \cdot \langle \nabla u\rangle \, d\Omega = -2|\Omega|\bar{k}\langle \nabla u\rangle^2 \tag{2.27}$$

We also rewrite the third integral in terms of the spatial average to obtain

$$\int_\Omega k\langle \nabla u\rangle^2 \, d\Omega = |\Omega|\langle k\rangle\langle \nabla u\rangle^2 \tag{2.28}$$

Introducing the three terms back into eq. (2.25)

$$0 \leq |\Omega|\langle k\rangle\langle \nabla u\rangle^2 - |\Omega|\bar{k}\langle \nabla u\rangle^2 \tag{2.29}$$

upond which the upper bound follows. The proof of the lower bound is similar, where we instead compute the work of the field $\boldsymbol{q} - \langle \boldsymbol{q}\rangle$. □

The upper bound $\langle k\rangle$ is usually called rule of mixtures or Voigt's bound. It is a weighted arithmetic average of the conductivity of the different phases. The weights correspond to their volume fractions. Similarly, the lower bound is called inverse rule of mixtures or Reuss' bound and it coincides with the weighted harmonic average of the conductivity of the different phases, *i.e.*

$$\langle k^{-1}\rangle^{-1} = \frac{1}{\sum_{i=1}^{n_{\text{phases}}} \frac{\nu_i}{k_i}} \tag{2.30}$$

where k_i and ν_i are the conductivity and volume fraction of phase i[2].

These bounds are named after Voigt and Reuss and their early works in the area of homogenisation. Voigt in [95] proved that the effective tensor results from the inverse rule of mixtures under the assumption that the strain in a heterogeneous material was constant. Similarly, under the assumption that the stress was constant, Reuss showed in [83] that the effective tensor results from rule of mixtures.

2.3 Asymptotic homogenisation of periodic media

In the theory of asymptotic homogenisation of periodic media [88], we assume that our domain is constructed by the repetition of a cell of dimensions $\epsilon L_x \times \epsilon L_y$(see fig. 2.1). We study the limit case when the dimension of the cell tends to 0, *i.e.* $\epsilon \to 0$ and we start by rewriting the temperature field u as a function of two variables $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$,

$$u(\boldsymbol{x}) = u(\boldsymbol{x}_0, \boldsymbol{x}_1) \tag{2.31}$$

which are defined by

$$\boldsymbol{x}_0 := \boldsymbol{x} \quad \boldsymbol{x}_1 := \frac{\boldsymbol{x}}{\epsilon} \tag{2.32}$$

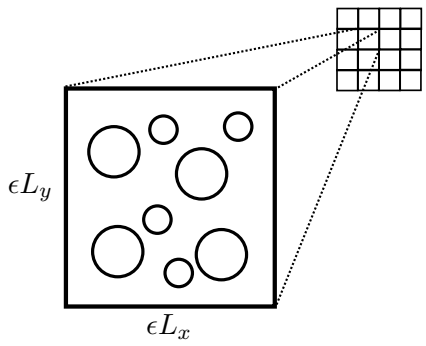


Figure 2.1: Unit cell and tiling of the domain.

[2]The relation of the bounds with arithmetic and harmonic average only apply to materials with clearly distinguished phases, *i.e.* k is piecewise constant in the RVE.

The variable $\boldsymbol{x}_0$ describes the macroscopic scale while the variable $\boldsymbol{x}_1$ describes the microscopic scale (the cell). We seek a temperature field that is periodic with the variable $\boldsymbol{x}_1$, *i.e.*

$$u(\boldsymbol{x}_0, \boldsymbol{x}_1) = u\left(\boldsymbol{x}_0, \boldsymbol{x}_1 + \begin{pmatrix} \epsilon L_x \\ 0 \end{pmatrix}\right) = u\left(\boldsymbol{x}_0, \boldsymbol{x}_1 + \begin{pmatrix} 0 \\ \epsilon L_y \end{pmatrix}\right) \tag{2.33}$$

In addition, since ϵ tends to 0, $\boldsymbol{x}_1$ varies a lot faster than $\boldsymbol{x}_0$. For this reason, even if $\boldsymbol{x}_0$ and $\boldsymbol{x}_1$ are related, we treat them as independent variables. Expanding asymptotically u in terms of the parameter ϵ

$$u(\boldsymbol{x}_0, \boldsymbol{x}_1) = u_0(\boldsymbol{x}_0, \boldsymbol{x}_1) + u_1(\boldsymbol{x}_0, \boldsymbol{x}_1)\epsilon + u_2(\boldsymbol{x}, \boldsymbol{x}_1)\epsilon^2 + ... \tag{2.34}$$

The derivative of u can be computed using the chain rule, *i.e.*

$$\frac{du}{d\boldsymbol{x}} = \frac{\partial u}{\partial \boldsymbol{x}_0} + \frac{1}{\epsilon}\frac{\partial u}{\partial \boldsymbol{x}_1} \tag{2.35}$$

This suggests the introduction of the notation

$$\nabla_0 := \frac{\partial}{\partial \boldsymbol{x}_0} \quad \nabla_1 := \frac{\partial}{\partial \boldsymbol{x}_1} \tag{2.36}$$

for convenience. Expanding the flux asymptotically,

$$\boldsymbol{q}(\boldsymbol{x}_0, \boldsymbol{x}_y) = \boldsymbol{q}_0(\boldsymbol{x}_0, \boldsymbol{x}_1) + \boldsymbol{q}_1(\boldsymbol{x}_0, \boldsymbol{x}_1)\epsilon + ... \tag{2.37}$$

we obtain by using the constitutive relation and identifying the terms of the same order that

$$0 = \nabla_1 u_0 \tag{2.38}$$

$$\boldsymbol{q}_0 = -k(\nabla_0 u_0 + \nabla_1 u_1) \tag{2.39}$$

$$\boldsymbol{q}_1 = -k(\nabla_0 u_1 + \nabla_1 u_2) \tag{2.40}$$

By substituting the flux into the conservation equation, we now obtain

$$\epsilon^{-1}\nabla_1 \cdot \boldsymbol{q}_0 + (\nabla_0 \cdot \boldsymbol{q}_0 + \nabla_1 \cdot \boldsymbol{q}_1 - f) + \text{ higher order terms} = 0 \tag{2.41}$$

Due to the presence of the parameter ϵ each of terms must cancel independently. Ignoring higher order terms, we obtain three problems. Equation (2.38) is a macroscopic problem

and allows us to conclude that u_0 is independent of $\boldsymbol{x}_1$, meaning that u_0 is a macroscopic smooth field. By considering the term of order ϵ^{-1} in eq. (2.41), we obtain a microscopic problem,

$$\nabla_1 \cdot \boldsymbol{q}_0 = \nabla_1 \cdot [-k(\nabla_0 u_0 + \nabla_1 u_1)] = 0 \tag{2.42}$$

from which we can determine the field u_1. Finally, we consider the term of order 1 in eq. (2.41). Taking the average in a cell,

$$\langle \nabla_0 \cdot \boldsymbol{q}_0 + \nabla_1 \cdot \boldsymbol{q}_1 - f \rangle := \frac{1}{|\Omega_e|} \int_{\Omega_e} (\nabla_0 \cdot \boldsymbol{q}_0 + \nabla_1 \cdot \boldsymbol{q}_1 - f)\, d\Omega = \nabla_0 \cdot \langle \boldsymbol{q}_0 \rangle - \langle f \rangle = 0 \tag{2.43}$$

we obtain an equation for the macroscopic flux. The term $\langle \nabla_1 \cdot \boldsymbol{q}_1 \rangle$ is cancelled due to periodicity. By finding $\langle \boldsymbol{q}_0 \rangle$, we can find the homogenised constitutive relation $\bar{k}$ by using

$$\langle \boldsymbol{q}_0 \rangle = -\bar{k} \nabla_0 u_0. \tag{2.44}$$

As we have just seen, the advantage of this approach is that it allows us to derive the form of the macroscopic constitutive equations only using microscopic constitutive relations.

2.4 $\mathbf{FE}^2$

In [25, 26] Feyel established the fundamentals of the FE^2 method. In this method, the constitutive relation in the macroscale is determined point by point by solving an RVE of the microscale. For this reason, it is considered a computational homogenisation method [64].

Given the macroscopic strain $\bar{\boldsymbol{\epsilon}}$ in a point of the domain, the macroscopic stress $\bar{\boldsymbol{\sigma}}$ is determined in two steps. Firstly, the microscopic strain and stress are determined by solving an RVE with PBC where the average strain is equal to the macroscopic strain. Secondly, the macroscopic stress is determined by averaging the microscopic stress.

Having established the constitutive relation, the macroscopic problem is solved by means of an iterative algorithm. Let us assume that a macroscopic strain field is given on the quadrature points (Gauss points) of the domain. Using the algorithm described above, the macroscopic stress field can be determined and the residue can be computed using the quadrature rule. If the residue is below a certain tolerance, the problem is considered to be solved, otherwise the strain field is updated.

The application of PBC is a severe constraint on the RVE mesh since it requires the discretisations on opposing sides to coincide. In [54], Larsson et al. reformulates the variational form with a a weaker form of the PBCs.

The main advantage of FE^2 is that the macroscopic response can be determined even for very complex microscopic constitutive relations. The main drawback is the computational cost due to the significant amount of RVE problems that must be solved, specially for nonlinear constitutive relations. In [52], error estimates are introduced to guide the process of the macro mesh refinement and reduce the total computational cost. The error is decomposed additively into discretisation and modelling error. The former is measured following the techniques presented in chapter 1. An hierarchy of models is introduced to account for the modelling error. At the top of this hierarchy lies the heterogeneous model, at the bottom its homogenised counterpart and in between models with different levels of detail (*e.g.* each level in this hierarchy has a smaller tolerance for the discretisation error in RVE computations). Error is measured by comparing two models in this hierarchy. With the use of those estimates to guide the mesh refinement process, a significant reduction in the computational cost is achieved.

2.5 The coupled-volume method

Methods such as FE^2 which link microscale to a single point in macroscale show a strong dependency of the macroscale mesh on the softening regime. The material becomes more brittle as the mesh is refined. In [32], Gitman et al. demonstrated this issue and proposed the coupled-volume method, where now each macro element corresponds to an equally shaped microscale cell. With this approach the dependency on the macroscale mesh disappears.

2.6 Multiscale FEM

The multiscale finite element was introduced by Hou and Wu [38] for composite materials and porous media. In this method specific shape functions are constructed to represent

the microscale phenomena. Immediately after the construction of the shape functions, an approximate solution is sought through the Galerkin method. The construction of these specific shape functions greatly reduces the amount of degrees of freedom required to approximate the exact solution.

In more detail, the shape functions are constructed element by element. Let us assume a coarse discretisation of the domain where N_i is the classical FE hat function associated with node i. Then, for each element, we seek functions ϕ_i such that $\phi_i(\boldsymbol{x}) = N_i(\boldsymbol{x})$ on the boundary of the element and that fulfil

$$a(\phi_i, v) = 0 \quad \forall v \in \mathcal{U}_0(\Omega_e). \tag{2.45}$$

With the microscale shape functions, our conforming approximation reads

$$u^h(\boldsymbol{x}) = \sum_{i=1}^{n_{\text{nodes}}} u_i \phi_i(\boldsymbol{x}). \tag{2.46}$$

The coefficients u_i are usually obtained by means of Bubnov-Galerkin, *i.e.* the test space coincides with the trial space or by means of Petrov-Galerkin where the test space is span $\{N_i\}_{i=1}^{n_{\text{nodes}}}$ in order to reduce the computational cost associated with integration. A good account of more recent developments of this method can be found in [24].

2.7 The variational multiscale method

The variational multiscale method (VMS) was introduced by Hughes in [39] and further advanced in [40, 41] by Hughes et al. to develop numerical methods for PDEs that exhibit a multiscale behaviour. Let us consider a generic problem with the following weak form,

$$\langle \mathcal{L}u, v \rangle = \langle f, v \rangle \quad \forall v \in V \tag{2.47}$$

where $\mathcal{L}$ is a differential operator, f is a given function, V is the test space and $\langle \cdot \rangle$ denotes the L_2 inner product, *i.e.*

$$\langle u, v \rangle = \int_\Omega uv \, d\Omega. \tag{2.48}$$

The main idea of VMS is to decompose the solution u additively in two terms: $\bar{u}$ which aims to represent the coarse scale phenomena and u' which represents the fine scale phenomena.

Then, an expression for u' in terms of the macroscale residue and a Green function $g(\boldsymbol{x}, \boldsymbol{y})$ is derived, *i.e.*

$$u'(\boldsymbol{y}) = -\int_{\Omega} g(\boldsymbol{x}, \boldsymbol{y})(\mathcal{L}\bar{u} - f)(\boldsymbol{x})\, d\Omega_{\boldsymbol{x}} \tag{2.49}$$

The expression in eq. (2.49) can be introduced in eq. (2.47) to obtain a new weak form that depends only on $\bar{u}$. We note that the Green function $g(\boldsymbol{x}, \boldsymbol{y})$ is not a classical Green function and it depends not only on the operator, but also on the discretisation. A closed-form expression for $g(\boldsymbol{x}, \boldsymbol{y})$ is not available. Therefore, the application of VMS requires the approximation of this function or the elementwise approximation of u'. In both cases, bubble functions are used to approximate them. This local approximation of the fine-scale functions resembles the MsFEM method.

2.8 Heterogeneous multiscale FEM and computational continua

The heterogeneous multiscale FE method (HMFEM) was introduced in [97]. We discuss here the formulation of this method presented in [2]. Given a macroscale discretisation of the domain, we associate a microscale problem to each pair made of a quadrature point and a macroscale shape function[3]. However, in contrast to FE^2 and similar to the coupled-volume FEM, each microscale problem is defined on a finite region surrounding the quadrature point. These regions are called sampling domains and we denote them by Ω_δ. The microscale problems read

Find $u^i_{\Omega_\delta}$ such that $u^i_{\Omega_\delta} - u^i \in U$ such that $\forall v \in V$

$$\int_{\Omega_\delta} k\nabla u^i_{\Omega_\delta} \cdot \nabla v\, d\Omega = 0 \tag{2.50}$$

where u^i is the linearisation of the corresponding macroscale function. Once the

[3]The quadrature point must be included in the shape function's support.

microscale functions are defined, the macroscale bilinear form is given by

$$a(u^i, u^j) = \sum_{\Omega_\delta} \omega_{\Omega_\delta} \int_{\Omega_\delta} k \nabla u^i_{\Omega_\delta} \cdot \nabla u^j_{\Omega_\delta} \, \Omega \tag{2.51}$$

where the sum is taken over the sampling domains and ω_{Ω_δ} is a weight which accounts for the area of the sampling region.

A priori error estimates which characterise the convergence properties of HMFEM, were derived in [3] by Abdulle. The same author with Nonnenmacher in [2], developed explicit a posteriori estimates which were used to guide mesh refinement process and control the total computational cost.

In [27, 28], Fish et al. introduced the computational continua. This method resembles HMFEM in that a finite microscale problem is associated to each quadrature point and that a modified macroscale bilinear form is adopted. This bilinear form is also reduced to a sum of integrals over the sampling domains (here called computational unit cells). Nonetheless, the quadrature points do not necessarily coincide with the Gauss points and are chosen accordingly to reproduce the weak form of the governing equations of the microscale up to a prespecified polynomial degree. In this approach, the macroscale stress (or coarse-scale stress) is prescribed to be linear while its microscale counterpart is periodic. As a result of this, the computational continua can be understood as form of second order homogenisation in the sense that the unit cells are under linear varying stresses.

2.9 Hierarchical modelling of heterogeneous bodies

In [103], Oden et al. introduced a technique to model heterogeneous materials with a posteriori error bounds. In this method, it is assumed that an accurate description of the microstructure is available. To avoid the excessive computational cost of this problem, an homogenised surrogate is solved and the quality of this solution is quantified through a posteriori error estimates. The error estimates used were implicit such as those described in chapter 1 and measured error in the energy norm. This work was expanded in series of papers [61, 65–68, 86, 87, 93]. These extensions include measures of the error in quantities

of interest, techniques to reduce the error through the local solution of the microstructure and the application of the method to nonlinear and random problems.

The main advantage of this method is the great reduction of computational cost by means of homogenisation and at the same time, the error introduced by this simplification can be quantified through fully computable error bounds (as opposed to the error bounds for HMFEM). Nonetheless, the main shortcoming of this method is that the quantification of the error requires the computation of an integral over the heterogeneous domain, *i.e.* an integration mesh that adjusts to microstructure is required. As already advanced in the introduction, in chapter 4, we present an extension of this method that alleviates several of the shortcomings of this approach.

2.10 Conclusion

In this chapter, we have reviewed the Hill-Mandel approach, the asymptotic theory of homogenisation, FE^2, the coupled-volume method, the multiscale FEM, the variational multiscale method and HMFEM. In the first three methods, the microscale is solved (almost) independently of the macroscale. Such methods are classified as nested methods. In contrast, the remaining approaches, the concurrent methods, the microscale is solved in the macroscale model. In those methods, since the microscale is part of the variational formulation, quantification of error is more straight forward than in the nested approaches.

In the next two chapters, we develop methods which resemble the nested approaches, while at the same time, these methods allow us to quantify the error.

Chapter 3

Equilibrated finite element method formulation

3.1 Introduction

We describe in this chapter the numerical approach to obtain a statically admissible field (section 1.4.1) by means of a stress/equilibrated finite element formulation. We recall that a statically admissible (SA) field is an element of the set

$$\mathcal{S}(\Omega) = \left\{ \boldsymbol{q} \in H^{\text{div}}(\Omega) | \nabla \cdot \boldsymbol{q} = f(\boldsymbol{x}) \, \forall \boldsymbol{x} \in \Omega \text{ and } \boldsymbol{q}(\boldsymbol{x}) \cdot \boldsymbol{n} = g(\boldsymbol{x}) \, \forall \boldsymbol{x} \in \Gamma_N \right\} \tag{3.1}$$

where the notation introduced in section 1.1.1 was used.

We seek to construct such a field to be able to compute the CRE based error bound described in chapter 1. The advantage of the CRE over the flux-free approach is that it is based on an identity which simplifies the characterisation of the effectivity of the error estimate and as we shall see, this will also allow us to characterise our error bounds for homogenisation in chapter 4. Moreover, this approach is favoured over the element equilibration technique since it minimises the error in the energy norm, *i.e.* $\|\boldsymbol{q} - \boldsymbol{q}^h\|_{k^{-1}}$, meaning that the effectivity of the estimate is maximised.

We start the chapter by reviewing the weak formulation of the heat equation in terms of fluxes before proceeding to the presentation of its finite element formulation.

3.2 Flux formulation

The weak form of the heat equation in terms of fluxes reads:

For all $\boldsymbol{p} \in \mathcal{S}_0(\Omega)$, find $\boldsymbol{q} \in \mathcal{S}(\Omega)$ such that

$$b(\boldsymbol{p}, \boldsymbol{q}) := \int_\Omega k^{-1}(\boldsymbol{x}) \, \boldsymbol{q}(\boldsymbol{x}) \cdot \boldsymbol{p}(\boldsymbol{x}) \, d\Omega = - \int_{\Gamma_D} \boldsymbol{p}(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x}) \, D(\boldsymbol{x}) \, d\Gamma =: p(\boldsymbol{p}) \tag{3.2}$$

where the following space was introduced

$$\mathcal{S}_0(\Omega) = \{ \boldsymbol{q} \in H^{\text{div}}(\Omega) \, | \, \boldsymbol{q}(\boldsymbol{x}) \cdot n(\boldsymbol{x}) = 0 \; \forall \boldsymbol{x} \in \Gamma_N \text{ and } \nabla \cdot \boldsymbol{q}(\boldsymbol{x}) = 0 \; \forall \boldsymbol{x} \in \Omega \}. \tag{3.3}$$

We note that the solution of this problem and u, the solution of eq. (1.7), are related by the constitutive relation, *i.e.* $\boldsymbol{q} = -k\nabla u$. To show this, let $\boldsymbol{q}_u$ be the flux field of u, $\boldsymbol{q}_u = -k\nabla u$. We multiply the constitutive relation by any $\boldsymbol{p} \in \mathcal{S}_0(\Omega)$ and integrate over the domain,

$$\int_\Omega k^{-1}\boldsymbol{q}_u \cdot \boldsymbol{p}\, d\Omega = -\int_\Omega \boldsymbol{p} \cdot \nabla u\, d\Omega. \tag{3.4}$$

Now, integrating the right-hand side by parts, it follows that

$$-\int_\Omega \boldsymbol{p} \cdot \nabla u\, d\Omega = -\int_\Gamma u\boldsymbol{p} \cdot \boldsymbol{n} + \int_\Omega u\nabla \cdot \boldsymbol{p}\, d\Omega. \tag{3.5}$$

The boundary integral can be decomposed additively in an integral over the Dirichlet boundary and an integral over the Neumann boundary. The latter and the domain integral are both cancelled since $\boldsymbol{p}$ is an element of $\mathcal{S}_0(\Omega)$. In addition, $u = D$ on the Dirichlet boundary, therefore,

$$\int_\Omega k^{-1}\boldsymbol{q}_u \cdot \boldsymbol{p} = -\int_{\Gamma_D} u\boldsymbol{p} \cdot \boldsymbol{n}\, d\Gamma = -\int_{\Gamma_D} D\boldsymbol{p} \cdot \boldsymbol{n}\, d\Gamma \quad \forall \boldsymbol{p} \in \mathcal{S}_0(\Omega). \tag{3.6}$$

Since this boundary value problem has a unique solution, $\boldsymbol{q} = \boldsymbol{q}_u$.

3.3 Hybrid stress FEM

The hybrid stress FEM was introduced [23] and a more recent and detailed account of this method can be found in [62]. We assume that the domain Ω is discretised in disjoint sets (elements) $\{\Omega_i\}_{i=1}^{n_{\text{elm}}}$ with triangular/tetrahedral shapes and that f and g are polynomials in each of the elements (otherwise the approximation obtained will be nonconforming).

We seek an approximation $\boldsymbol{q}^h$ made of piecewise polynomials. More precisely, a set of independent polynomials is defined for each element, *i.e.*

$$\boldsymbol{q}^h \in \mathcal{S}^h(\Omega) := \left\{ \boldsymbol{q} \in \mathcal{S}(\Omega) | \quad \boldsymbol{q}\big|_{\Omega_i} \in (\mathbb{P}_N(\Omega_i))^d \, i = 1, 2 ... n_{\text{elm}} \right\} \tag{3.7}$$

where $(\mathbb{P}_N(\Omega_i))^d$ is the space of d-tuples of polynomials with degree $\leq N$ defined over Ω_i. Here $d = 1, 2, 3$ is the number of space dimensions.

The approximation is decomposed additively into a homogeneous term $\boldsymbol{q}^h_{\mathrm{H}}$ and a nonhomogeneous term $\boldsymbol{q}^h_{\mathrm{NH}}$ such that

$$\nabla \cdot \boldsymbol{q}^h_{\mathrm{H}} = 0$$
$$\nabla \cdot \boldsymbol{q}^h_{\mathrm{NH}} = f$$

This decomposition suggests that we further decompose each $(\mathbb{P}_N(\Omega_i))^d$ in two spaces, a space of divergence free and a space of nondivergence free polynomials. For the moment, we assume that a basis $\{\boldsymbol{p}\}_{i=1}^{n_p}$ and $\{\boldsymbol{m}\}_{i=1}^{n_m}$ of the space of divergence free and nondivergence free polynomials respectively are available. We postpone the construction of the bases of those subspaces to section 3.4.

We proceed now with the construction of the field $\boldsymbol{q}^h_{\mathrm{NH}}$. The field $\boldsymbol{q}^h_{\mathrm{NH}}$ is constructed element-wise by projecting f on the transformed basis $\{\nabla \cdot \boldsymbol{m}\}_{i=1}^{n_m}$, *i.e.* the divergence of the basis of the space of nondivergence free polynomials. More accurately, if $\boldsymbol{q}^h_{\mathrm{NH}}$ in the element e with domain Ω_e reads $\boldsymbol{q}^h_{\mathrm{NH}}\big|_{\Omega_e} = \alpha^e_1 \boldsymbol{m}_1 + \alpha^e_2 \boldsymbol{m}_2 + ... + \alpha^e_n \boldsymbol{m}_{n_m}$, then we can determine the coefficients $\{\alpha^e_i\}_{i=1}^{n_m}$ in this element by solving the following system of equations

$$\int_{\Omega_e} \nabla \cdot \boldsymbol{m}_j \nabla \cdot \boldsymbol{q}^h_{\mathrm{NH}} \, d\Omega = \int_{\Omega_e} f \nabla \cdot \boldsymbol{m}_j \, d\Omega \quad j = 1, 2..., n_m. \tag{3.8}$$

In matrix form,

$$\boldsymbol{K}\boldsymbol{u} = \boldsymbol{f}$$

with

$$K_{ij} = \int_{\Omega_e} \nabla \cdot \boldsymbol{m}_i \, \nabla \cdot \boldsymbol{m}_j \, d\Omega$$
$$u_i = \alpha^e_i$$
$$f_i = \int_{\Omega_e} f \, \nabla \cdot \boldsymbol{m}_i \, d\Omega \qquad i, j = 1, 2..., n_m.$$

Owning to the fact that the polynomials $\{\nabla \cdot \boldsymbol{m}\}_{i=1}^{n_m}$ are linearly independent (otherwise $\{\boldsymbol{m}\}_{i=1}^{n_m}$ would not constitute a basis), $\boldsymbol{K}$ is positive definite. Since this matrix is also symmetric, it is invertible. Moreover, the approximation $\boldsymbol{q}^h_{\mathrm{NH}}$ is optimal in the sense that

if f is non representable by the basis $\{\boldsymbol{m}\}_{i=1}^{n'}$ in an element e, the resulting approximation $\boldsymbol{q}_{\mathrm{NH}}^h$ minimises

$$\int_{\Omega_e} (f - \nabla \cdot \boldsymbol{q}_{\mathrm{NH}}^h)^2 \, d\Omega. \tag{3.9}$$

This is a consequence of the best approximation property.

Now, we proceed to calculate the field $\boldsymbol{q}_{\mathrm{H}}^h$. The field $\boldsymbol{q}_{\mathrm{H}}^h$ must be such that the resulting field $\boldsymbol{q}^h$:

- optimally approximates $\boldsymbol{q}$ in the weak form's energy norm,

$$\min_{\boldsymbol{q}^h \in \mathcal{S}^h(\Omega)} b(\boldsymbol{q} - \boldsymbol{q}^h, \boldsymbol{q} - \boldsymbol{q}^h) = \|\boldsymbol{q} - \boldsymbol{q}^h\|_{k^{-1}}^2. \tag{3.10}$$

  This choice, besides allowing us to use the Galerkin method, has an additional advantage. If we use the Prager-Synge theorem (section 1.4.1) to bound $\|u - u^h\|$, the bound effectivity will be optimal since it is controlled by $\|\boldsymbol{q} - \boldsymbol{q}^h\|_{k^{-1}}^2$.

- has its normal continuous along the element edges (or facets) and fulfils the Neumann boundary conditions, *i.e.* $\boldsymbol{q}^h \cdot \boldsymbol{n} = g \, \forall \boldsymbol{x} \in \Gamma_N$. Otherwise, the resulting approximation is nonconforming, *i.e.* $\boldsymbol{q}^h \notin \mathcal{S}(\Omega)$.

We build $\boldsymbol{q}_{\mathrm{H}}^h$ through the Galerkin-Bubnov method combined with Lagrange multipliers. The Lagrange multipliers are used to enforce the second condition. We briefly recall the method of Lagrange multipliers for the minimisation (or maximisation) of functionals (for further details see [29]). This method states that the minimisation of a functional $\mathbb{F} : V \to \mathbb{R}$ with an additional constraint on the elements of V,

$$g(f, \boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Omega \tag{3.11}$$

is equivalent to finding a function $f \in V$ and a function λ[1] such that they minimise the functional

$$\mathbb{F}(f) + \int_{\Omega} \lambda(\boldsymbol{x}) g(f, \boldsymbol{x}) \, d\Omega. \tag{3.12}$$

The function λ is the so called Lagrange multiplier.

[1] The space of the function λ is dependent on the constraint.

In our particular case, the functional to minimise is

$$\frac{1}{2}b(\boldsymbol{q}^h, \boldsymbol{q}^h) + p(\boldsymbol{q}^h) \tag{3.13}$$

fulfilling the constraints

$$\boldsymbol{q}_1^h(\boldsymbol{x}) \cdot \boldsymbol{n} - \boldsymbol{q}_2^h(\boldsymbol{x}) \cdot \boldsymbol{n} = 0 \tag{3.14}$$

for each interior edge with the indices indicating the restriction of $\boldsymbol{q}^h$ to each of the elements that meet in the specific edge; and also fulfilling

$$\boldsymbol{q}^h(\boldsymbol{x}) \cdot \boldsymbol{n} = g(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_N. \tag{3.15}$$

In this case, since the set of interior edges and the set of edges on the Neumann boundary are disjoint, we can define the Lagrange multiplier as a single function defined over the element edges. The minimisation problem now reads,

$$\frac{1}{2}b(\boldsymbol{q}^h, \boldsymbol{q}^h) + p(\boldsymbol{q}^h) + \int_{\Gamma_I} \lambda \boldsymbol{q}^h \cdot \boldsymbol{n}\, d\Gamma + \int_{\Gamma_N} \lambda(\boldsymbol{q}^h \cdot \boldsymbol{n} - g)\, d\Gamma \tag{3.16}$$

where Γ_I are the interior edges. It is important to note that the outward and inward direction for the normal over element edges is not defined *a priori* and this must be defined *ad hoc*. Taking the first variation of this functional, we obtain its weak form which reads

For all $\boldsymbol{p}(\boldsymbol{x}) \in \mathcal{S}_0^h(\Omega)$ and $\mu \in V$, find $\boldsymbol{q}^h(\boldsymbol{x}) \in \mathcal{S}^h(\Omega)$ and λ in V such that

$$\begin{aligned} b(\boldsymbol{p}, \boldsymbol{q}^h) + L(\lambda, \boldsymbol{q}^h) &= p(\boldsymbol{p}) \\ L(\mu, \boldsymbol{q}^h) &= \int_{\Gamma_N} g\mu. \end{aligned}$$

with $L(\mu, \boldsymbol{p}) = \int_{\Gamma_I \cup \Gamma_j} \mu \boldsymbol{p} \cdot \boldsymbol{n}\, d\Gamma$ introduced for compactness. This is a mixed problem and we forward the reader to chapter 11 of [14] for a rigorous discussion on the well-posedness of these kind of problems. We only point out that since $\mathcal{S}^h(\Omega)$ is a finite dimensional space, we can guarantee that the constraints are fulfilled by choosing

$$V = \left\{ \lambda | \quad \lambda\big|_{\gamma_j} \in \mathbb{P}_N(\gamma_j) \quad \forall \gamma_j \in \gamma \right\} \tag{3.17}$$

where γ is the set of all edges lying in the interior of the domain and in the Neumann boundary. In other words, λ is a polynomial of degree N on each element edge that lies in $\Gamma_I \cup \Gamma_N$.

The resulting system of equations takes the form

$$\begin{pmatrix} \boldsymbol{F} & \boldsymbol{D}^T \\ \boldsymbol{D} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \end{pmatrix} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \end{pmatrix} \tag{3.18}$$

with

$$\boldsymbol{F} = \begin{pmatrix} \boldsymbol{F}^1 & \boldsymbol{0} & \dots & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{F}^2 & & \vdots \\ \vdots & & \ddots & \boldsymbol{0} \\ \boldsymbol{0} & \dots & \boldsymbol{0} & \boldsymbol{F}^{n_{\text{elm}}} \end{pmatrix} \quad F^e_{ij} = b_{\Omega_e}(\boldsymbol{p}_i, \boldsymbol{p}_j) := \int_{\Omega_e} k^{-1} \boldsymbol{p}_i \cdot \boldsymbol{p}_j \, d\Omega$$

$$\boldsymbol{u} = \begin{pmatrix} \boldsymbol{u}^1 \\ \vdots \\ \boldsymbol{u}^{n_{\text{elm}}} \end{pmatrix} \quad \boldsymbol{u}^e_i = -\int_{\Gamma_D \cap \Omega_e} D\boldsymbol{p}_i \cdot \boldsymbol{n} \, d\Gamma - \int_{\Omega_e} \boldsymbol{q}^h_{\text{NH}} \cdot \boldsymbol{p}_i \, d\Omega$$

$$i, j = 1...n_p \quad e = 1...n_{\text{elm}}$$

and

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{D}^{11} & \boldsymbol{D}^{12} & \dots & \boldsymbol{D}^{1 n_{\text{elm}}} \\ \boldsymbol{D}^{21} & \boldsymbol{D}^{22} & & \vdots \\ \vdots & & \ddots & \\ \boldsymbol{D}^{n_{\text{edges}} 1} & \dots & & \boldsymbol{D}^{n_{\text{edges}} n_{\text{elm}}} \end{pmatrix} \quad D^{ke}_{ij} = \int_{\gamma_k \cap \partial\Omega_e} s^{\gamma_k}_i \boldsymbol{p}_j \, d\Gamma$$

$$\boldsymbol{v} = \begin{pmatrix} \boldsymbol{v}^1 \\ \vdots \\ \boldsymbol{v}^{n_{\text{edges}}} \end{pmatrix} \quad v^k_i = -\int_{\gamma_k \cap \partial\Omega_e} s^{\gamma_k}_i \boldsymbol{q}^h_{\text{NH}} \cdot \boldsymbol{n} \, d\Gamma + \int_{\gamma_i \cap \partial\Omega_e \cap \Gamma_N} s^{\gamma_k}_i g \, d\Gamma$$

$$i = 1...N \; j = 1...n_p \; e = 1...n_{\text{elm}} \; k = 1...n_{\text{edges}}.$$

The functions $\{s^{\gamma_k}_i\}^N_{i=1}$ are any polynomial basis of $\mathbb{P}_N(\gamma_j)$. We remark that matrix D is not dense since for many combinations of i and e, $\gamma_i \cap \partial\Omega_e = \emptyset$. We also note that the

matrix $\boldsymbol{F}$ is block diagonal and that each block is invertible which means that $\boldsymbol{F}^{-1}$ can be computed by inverting each block. This suggests a more tractable approach for solving eq. (3.18). By expanding eq. (3.18)

$$\boldsymbol{F}\boldsymbol{\alpha} + \boldsymbol{D}^T\boldsymbol{\beta} = \boldsymbol{u}$$
$$\boldsymbol{D}\boldsymbol{\alpha} = \boldsymbol{v}$$

and writing $\boldsymbol{\alpha}$ in terms of $\boldsymbol{\beta}$, in the first equation

$$\boldsymbol{\alpha} = \boldsymbol{F}^{-1}(\boldsymbol{u} - \boldsymbol{D}^T\boldsymbol{\beta}) \tag{3.19}$$

and introducing this back into the second equation, we can solve the following smaller problem for $\boldsymbol{\beta}$

$$\boldsymbol{D}\boldsymbol{F}^{-1}\boldsymbol{D}^T\boldsymbol{\beta} = \boldsymbol{D}\boldsymbol{F}^{-1}\boldsymbol{u} - \boldsymbol{v} \tag{3.20}$$

The matrix $\boldsymbol{D}\boldsymbol{F}^{-1}\boldsymbol{D}^T$ is symmetric and it can be also shown that it is positive definite, meaning that it is invertible. See [62] for a proof of the latter. Once we determine the vector $\boldsymbol{\beta}$, we can obtain $\boldsymbol{\alpha}$ by substituting $\boldsymbol{\beta}$ in eq. (3.19).

3.4 Construction of the approximation polynomials

We complete the description of the hybrid stress FEM by describing the procedure to construct divergence free polynomials. We start by introducing $\{\boldsymbol{b}_i\}_{i=1}^m$, the canonical polynomial basis of $(\mathbb{P}_N(\mathbb{R}^d))^d$, for instance, the canonical basis for $(\mathbb{P}_2(\mathbb{R}^2))^2$ is

$$\begin{pmatrix}1\\0\end{pmatrix}, \begin{pmatrix}x\\0\end{pmatrix}, \begin{pmatrix}y\\0\end{pmatrix}, \begin{pmatrix}x^2\\0\end{pmatrix}, \begin{pmatrix}xy\\0\end{pmatrix}, \begin{pmatrix}y^2\\0\end{pmatrix}, \begin{pmatrix}0\\1\end{pmatrix}, \begin{pmatrix}0\\x\end{pmatrix}, \begin{pmatrix}0\\y\end{pmatrix}, \begin{pmatrix}0\\x^2\end{pmatrix}, \begin{pmatrix}0\\xy\end{pmatrix}, \begin{pmatrix}0\\y^2\end{pmatrix}. \tag{3.21}$$

We can construct a basis of the divergence free polynomials by solving

$$(\nabla \cdot \boldsymbol{b}(\boldsymbol{x})) \cdot \boldsymbol{\alpha} := \sum_{i=1}^{m} \alpha_i \nabla \cdot b_i(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^n. \tag{3.22}$$

If $\{\boldsymbol{\beta}_i\}_{i=1}^k$, $\boldsymbol{\beta}_i = (\alpha_1^i\, \alpha_2^i \ldots \alpha_m^i)^T$ is a basis of the solutions of eq. (3.22), *i.e.*

$$(\nabla \cdot \boldsymbol{b}(\boldsymbol{x})) \cdot \boldsymbol{\beta}_i = 0 \quad \forall \boldsymbol{x} \in \mathbb{R}^d \quad i = 1 \ldots k \tag{3.23}$$

then $\{\boldsymbol{b}\cdot\boldsymbol{\beta}_i\}_{i=1}^{k}$ is basis of the divergence free polynomials. Let $\{\boldsymbol{\delta}_i\}_{i=1}^{n-k}$ be a basis orthogonal to the space spanned by $\{\boldsymbol{\beta}_i\}_{i=1}^{k}$, *i.e.*

$$\boldsymbol{\beta}_i\cdot\boldsymbol{\delta}_j=0\quad i=1....k,\ j=1....(n-k). \tag{3.24}$$

then, $\{\boldsymbol{b}\cdot\boldsymbol{\delta}_i\}_{i=1}^{n-k}$ is a basis of the space of non divergence free polynomials.

Most numerical computing software is unable to solve eq. (3.22) in that specific form. By rewriting eq. (3.22) in matrix form, where we identify each row with a monomial, the bases $\{\boldsymbol{\beta}_i\}_{i=1}^{k}$ and $\{\boldsymbol{\delta}_i\}_{i=1}^{n-k}$ that we seek are the bases of the kernel and image of this matrix respectively. A singular value decomposition of this matrix will reveal a basis of its kernel and image.

To clarify the previous development, we consider the earlier example, the canonical basis of $(\mathbb{P}_2(\mathbb{R}^2))^2$. Equation (3.22) for these polynomials has the form

$$\alpha_2+2\alpha_4x+\alpha_5y+\alpha_9+\alpha_{11}x+2\alpha_{12}y=0\quad\forall\boldsymbol{x}\in\mathbb{R}^2. \tag{3.25}$$

The associated matrix reads

$$\boldsymbol{M}=\begin{pmatrix}0&1&0&0&0&0&0&0&1&0&0&0\\0&0&0&2&0&0&0&0&0&0&1&0\\0&0&0&0&1&0&0&0&0&0&0&2\end{pmatrix} \tag{3.26}$$

where we identified the first row with 1 (the constant monomial), the second row with x and the third row with y. Applying a singular value decomposition, we obtain the basis of the kernel, $\{\boldsymbol{\beta}_i\}_{i=1}^{k}$, and the basis of the image, $\{\boldsymbol{\delta}_i\}_{i=1}^{n-k}$ which can then be used to compute the polynomials that we seek.

3.5 $H^{\mathbf{div}}$ elements

The hybrid stress FEM is not the only method to obtain SA fields. Such approximations can also be obtained by using any of the elements of the H^{div} family such as the Raviart-Thomas elements [82] and the Brezzi-Douglas-Marini elements [15].

In the hybrid approach, we constrained a priori the elements to fulfil the conservation equation, *i.e.* $\nabla\cdot\boldsymbol{q}=f$ and we forced a posteriori the elements to have their normals

continuous along the element's edges. The approach is reversed when H^{div} elements are used. The normal continuity along the element edges is enforced by construction whilst the conservation equation is satisfied a posteriori by a reformulation of the problem (*e.g.* through Lagrange multipliers). However, provided that the same polynomial space is used, the fields resulting from the hybrid and the H^{div} approach will coincide and therefore, no change will be observed in the error bounds computed.

For this reason, we do not explore here this alternative into more depth. Nonetheless, the interested reader can find more details about the efficient implementation of such elements in [10, 85].

3.6 Conclusion

In this chapter, we have reviewed the details of the implementation of hybrid stress finite element method. Our main interest in this method is its use in the computation of guaranteed error bounds.

The implementation has two main steps. Firstly, a function that fulfils the energy equation element by element is constructed. In a second step, through the Galerkin method combined with Lagrange multipliers we find the best approximation in a finite dimensional subspace. The Lagrange multipliers were used to enforce the normal continuity of the fluxes.

In comparison with the element equilibration technique (described in chapter 1), the hybrid stress FEM is more computationally expensive since a full problem is solved instead of a collection of local patch problems. In exchange for the added computational cost, if the approximation is sought in the same finite dimensional space, the hybrid stress FEM always minimises the error in the energy norm. This error controls the effectivity and therefore, hybrid stress always results in better effectivities.

Chapter 4

Scale separated error bounds for stochastic homogenisation

4.1 Introduction

Based on the author's work presented in [71], we address the problem of error estimation for homogenisation in this chapter.

Our final objective is to model a structure where two scales are present, the microscale, the scale of the heterogeneities that compose the material, and the macroscale, the scale of the structure itself. The direct solution of such models is usually out of computational reach due to the required meshing effort combined with increased degrees of freedom. The burden of the heterogeneities and the two scales can be overcome through homogenisation of the heterogeneities as seen in chapter 2. In the homogenised model, also called surrogate model, only one scale is present and the computational cost is greatly reduced. However, homogenisation applies only when the scales are separated. This condition only holds in ideal cases and for this reason the surrogate model incorporates a model error. This error should be quantified (i) to establish the quality of the approximation, (ii) and to help analyst decide when and where homogenisation should be abandoned and more advanced strategies (such as hybrid nested-concurrent methods) should be used.

We recall that in the work of Oden et al. (described in section 2.9) , this question was answered by adapting implicit error estimates to asses whether the solution of the tractable model represents well the solution of the reference model. The main issue of this approach is that, even though an analysis suitable mesh of the heterogeneous domain is not required, the calculation of the error bounds involves an integral over this domain and at least a mesh suitable for integration is required. In other words, even though the scales are decoupled in the analysis through homogenisation, they are coupled back by the error estimation procedure. The practical consequence is that the approach is limited to small multiphase problems.

The first key contribution of this thesis is to represent the heterogeneous problem by a stochastic PDE where each realisation represents a different layout of the particles whilst the surrogate model remains the same. Due to the stochastic nature of heterogeneous model, we estimate the average error. Besides allowing the application of the method to

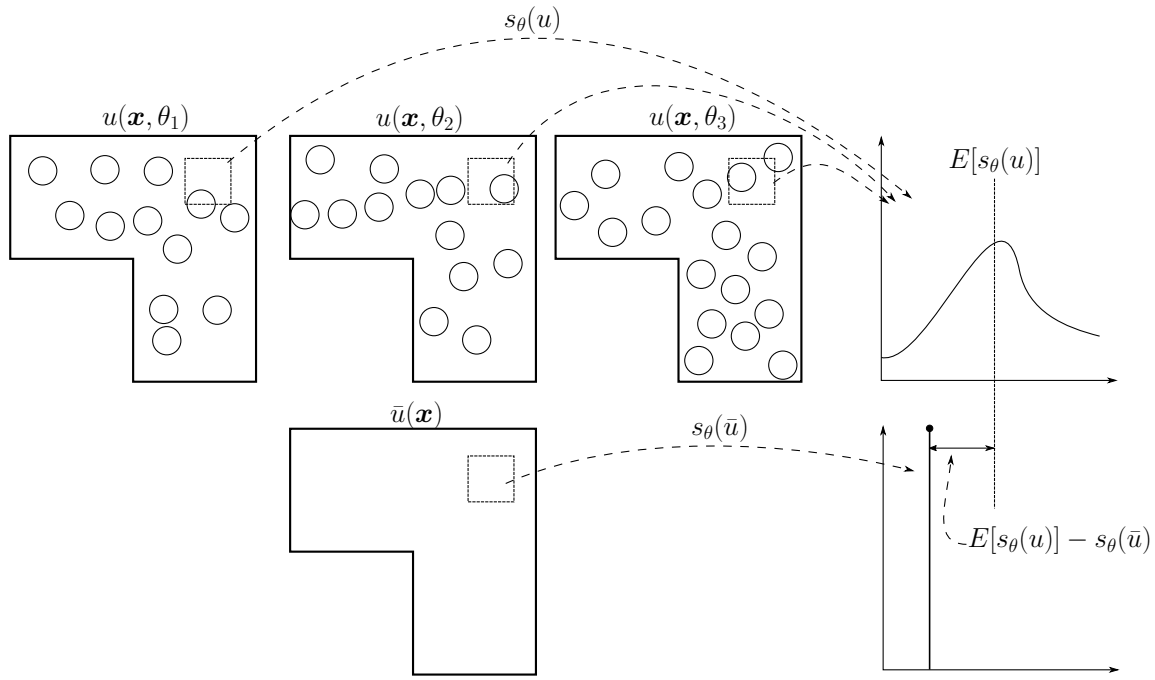


Figure 4.1: In the framework presented in this chapter, the accuracy of homogenisation, $\bar{u}$, is measured in terms of the "true" intractable stochastic model u. Direct approximation of a quantity of interest $s_\theta(u)$, a random variable, is not possible. Instead, we estimate $s_\theta(\bar{u})$ (a deterministic quantity) and bound the statistical moments of the difference, $s_\theta(u) - s_\theta(\bar{u})$.

a wider class of models in which the precise particle layout is not known, the proposed framework allows us to retain scale separation in the computation of error bounds. The choice to measure the average error smooths the fast varying diffusion fields and as a consequence, integration meshes that adjust to the constituents are not required.

In addition, the error estimates also account for the discretisation error (which was neglected in the preceding work), are fully computable, guaranteed. Furthermore, we are able to characterise and optimise the effectivity of the error bounds. These optima are related to the classical Reuss and Voigt bounds already discussed in chapter 2.

The approach is completed with a method to reduce error through model adaptivity. With the aid of error indicators, the regions that contribute most to the error are identified. By solving the microstructure in those regions, sharper error estimates are obtained.

The chapter is organised as follows. In section 4.2, the heterogeneous or reference model and the surrogate model are introduced. This section is concluded with a discussion of the different measures of the error. In section 4.3, a detailed derivation and characterisation of the error bounds for the error in energy norm and for the error in quantities of interest is given. Model adaptivity is presented at the end of the section. In section 4.4, the theory is complemented with five numerical examples. Finally, the conclusions are presented in section 4.5.

4.2 Reference problem

In this section, we introduce the reference problem and the notation. We recall the notation used throughout chapter 1, 2 and 3 and introduce new notation to describe the stochastic nature of variables functions and forms.

We consider the problem of stationary heat conduction in a body Ω defined in a subset of $\mathbb{R}^d$ $(d = 1, 2, 3)$. The boundary of this domain is denoted by Γ whilst its outward unit normal is denoted by $\boldsymbol{n}$. The boundary can be further divided in two parts Γ_N and Γ_D, Neumann and Dirichlet boundary respectively, such that $\Gamma_D \neq \emptyset$, $\Gamma_D \cup \Gamma_N = \Gamma$ and $\Gamma_D \cap \Gamma_N = \emptyset$. Deterministic fluxes g are prescribed on Γ_N, deterministic temperatures D are prescribed on Γ_D and a deterministic source term f that accounts for the internal heat generation is applied over the interior of the domain. The different phases in the domain are distributed according to a random algorithm. An example of such, would be the following algorithm:

A point in the domain is chosen according to a certain distribution and then a particle is placed in this point if and only if it does not intersect any other existing particles, otherwise a new point is generated. This process is repeated until a certain volume fraction is reached.

We note that the resulting conductivity k is a function of the spatial domain Ω and the stochastic domain Θ, that accounts for all the possible heterogeneities layouts. It is assumed that the expectation of the conductivity, and of its reciprocal $E[k^{-1}](\boldsymbol{x})$ are

known for all $\boldsymbol{x} \in \Omega$ or easily computable. We recall that the expectation is defined by

$$E[X] := \int_{\Theta} X \, d\Theta. \tag{4.1}$$

With the notation introduced, the stochastic heat conduction problem for the temperature field u reads:

For each $\theta \in \Theta$ and for all $v(\boldsymbol{x}) \in \mathcal{U}_0(\Omega)$, find $u(\boldsymbol{x}, \theta) \in \mathcal{U}(\Omega)$ such that

$$\int_{\Omega} k(\boldsymbol{x}, \theta) \, \nabla u(\boldsymbol{x}, \theta) \cdot \nabla v(\boldsymbol{x}) \, d\Omega = \int_{\Omega} f(\boldsymbol{x}) v(\boldsymbol{x}) \, d\Omega - \int_{\Gamma_N} g(\boldsymbol{x}) v(\boldsymbol{x}) \, d\Gamma \tag{4.2}$$

where

$$\mathcal{U}(\Omega) = \{u \in H^1(\Omega) : u(\boldsymbol{x}) = D(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_D\}$$
$$\mathcal{U}_0(\Omega) = \{u \in H^1(\Omega) : u(\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Gamma_D\}.$$

and $H^1(\Omega)$ is the Sobolev space of square integrable functions with square integrable first generalised derivatives on Ω. The elements of $\mathcal{U}(\Omega)$ are called kinematically admissible (KA) recalling the nomenclature of chapter 1. The left hand side of eq. (4.2) is a positive (semi-) definite bilinear form that will be denoted by

$$a_\theta(u, v) = \int_{\Omega} k(\boldsymbol{x}, \theta) \nabla u(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \, d\Omega$$

while its induced (semi-) norm will be denoted by

$$\|v\|_\theta = \sqrt{a_\theta(v, v)}.$$

The right hand side of eq. (4.2) is a linear form that will be denoted by

$$l(v) = \int_{\Omega} f(\boldsymbol{x}) v(\boldsymbol{x}) d\Omega - \int_{\Gamma_N} v(\boldsymbol{x}) g(\boldsymbol{x}) d\Gamma.$$

The choice of constraining the source term, f, the prescribed fluxes, g, and the prescribed temperature, D, to be deterministic functions while allowing the conductivity to be stochastic is justified by the fact that we assume that the boundary conditions are

known and that uncertainty lies on the domain. This work could be trivially extended to allow for uncertainty on f, g and D, provided we assume that the random variables that regulate the latter and k are independent *i.e.* the domain composition is not linked to the prescribed temperature, fluxes and sources. In this case, classical methods for stochastic problems such as spectral stochastic finite element methods [31] could be combined with the methods described here.

4.2.1 Homogenisation surrogate

Though eq. (4.2) describes in a realistic way a heterogeneous material, several difficulties arise in its solution. Firstly, the heterogeneities are usually small compared to the domain size which make difficult the generation of meshes. Secondly, an elevated amount of realisations might be required to produce a meaningful approximation.

For these reasons, we seek to solve a surrogate model that is tractable. This model is obtained by homogenising the conductivity field $k(\boldsymbol{x}, \theta)$. The resulting conductivity field $\bar{k}(\boldsymbol{x})$ is smooth and deterministic. The resulting problem is tractable and reads:

For all $v \in \mathcal{U}_0(\Omega)$, find $\bar{u} \in \mathcal{U}(\Omega)$

$$\int_\Omega \bar{k}(\boldsymbol{x}) \nabla \bar{u}(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) \, d\Omega = l(v). \tag{4.3}$$

The bilinear form of the left hand side of eq. (4.3) will be referred as

$$\bar{a}(u, v) := \int_\Omega \bar{k}(\boldsymbol{x}) \nabla \bar{u}(\boldsymbol{x}) \cdot \nabla v(\boldsymbol{x}) d\Omega. \tag{4.4}$$

We will use the finite element method (FEM) to solve this problem. In this method, a finite dimensional subspace $\bar{\mathcal{U}}(\Omega) \subseteq \mathcal{U}$ is defined and a projection of the solution is sought in this subspace. The approximated solution $\bar{u}^h$ will also be KA.

4.2.2 Error field and error measures

Having introduced the reference and the surrogate model, we introduce the error field,

$$e(\boldsymbol{x}, \theta) := u(\boldsymbol{x}, \theta) - \bar{u}^h(\boldsymbol{x}) \tag{4.5}$$

which describes the discrepancy between both approximations. Naturally, the approximation of this field is as intractable as the approximation of the reference model. This field is governed by the following problem

For each $\theta \in \Theta$ and for all $v(\boldsymbol{x}) \in \mathcal{U}_0(\Omega)$, find $e(\boldsymbol{x}, \theta) \in \mathcal{U}_0(\Omega)$ such that

$$a_\theta(e, v) = R_\theta(v) \tag{4.6}$$

where the residue was introduced for convenience and is defined as

$$R_\theta(v) := l(v) - a_\theta(\bar{u}^h, v) \tag{4.7}$$

We further note that the error field can be further decomposed in two fields,

$$e(\boldsymbol{x}, \theta) = [u(\boldsymbol{x}, \theta) - \bar{u}(\boldsymbol{x})] + [\bar{u}(\boldsymbol{x}) - \bar{u}^h(\boldsymbol{x})] \tag{4.8}$$

where the second term inside square brackets measures the discretisation error due to the use of FEM and any of the strategies described in chapter 1 can be applied to its estimation; whilst the first term accounts for the model error which is hard to measure directly. For this reason, the methods developed seek to measure the contributions of both errors jointly. More specifically, we will develop estimates for the expectation of the error in energy norm,

$$\|e\| := \sqrt{a(e, e)} \tag{4.9}$$

where

$$a(u, v) = E[a_\theta(u, v)] \quad \text{with } u, v \in \mathcal{U}_0^\theta,$$

and

$$\mathcal{U}_0^\theta = \{v \,|\, \forall \theta \in \Theta \quad v(\boldsymbol{x}, \theta) \in \mathcal{U}_0 \text{ and } \|v\| < \infty\}.$$

Choosing to measure an expectation of the error will result in deterministic error bounds, *i.e.* bounds that do not depend on stochastic quantities. The bounds will not depend on particular realisations of the conductivity field. Rather, they will depend on the expected conductivity field $E[k](\boldsymbol{x})$. This means that we will achieve a full separation of scales. Neither the computation of the approximation $\bar{u}^h$, nor the computation of the bounds will require operations on the microscale, meaning that integration meshes that adjust to heterogeneities will not be needed.

Two alternative natural measures are $\|e\|_\theta$ and $E\left[\|e\|_\theta\right]$, but we delay their discussion to the next section.

Additionally, the expected error in energy norm is favoured since it is useful in the construction of bounds for quantities of interest. We recall from section 1.5, that quantities of interest are outputs of the analysis, namely, the flux exiting part of the boundary or the average temperature in a subdomain. The bounds that will be developed are only valid for quantities of interest that can be expressed as a linear functional $s_\theta(u)$. In particular, deterministic bounds will be developed for the expectation and the second moment of $s_\theta(u) - s_\theta(\bar{u}^h)$.

4.3 Guaranteed modelling error bounds

In this section, we derive bounds for the error in energy norm and quantities of interest. The error bounds presented rely on the Prager-Synge hypercircle theorem (theorem 1.1) which relies on a statically admissible approximation. Hence, we start this section by rewriting eq. (4.2) in terms of fluxes and defining its tractable homogenised surrogate.

4.3.1 Complementary formulation

The formulation of the problem in eq. (4.3) in terms of fluxes, also called complementary formulation, reads:

For each $\theta \in \Theta$ and for all $\boldsymbol{p} \in \mathcal{S}_0(\Omega)$, find $\boldsymbol{q}(\boldsymbol{x}, \theta) \in \mathcal{S}(\Omega)$ such that

$$\int_\Omega k^{-1}(\boldsymbol{x}, \theta)\, \boldsymbol{q}(\boldsymbol{x}, \theta) \cdot \boldsymbol{p}(\boldsymbol{x})\, d\Omega = -\int_{\Gamma_D} \boldsymbol{p}(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x})\, D(\boldsymbol{x})\, d\Gamma \tag{4.10}$$

where

$$\mathcal{S}(\Omega) = \{\boldsymbol{q} \in H^{\mathrm{div}}(\Omega) \,:\, \boldsymbol{q}(\boldsymbol{x}) \cdot n(\boldsymbol{x}) = g(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \Gamma_N \text{ and } \nabla \cdot \boldsymbol{q}(\boldsymbol{x}) = f(\boldsymbol{x}) \qquad \forall \boldsymbol{x} \in \Omega\},$$
$$\mathcal{S}_0(\Omega) = \{\boldsymbol{q} \in H^{\mathrm{div}}(\Omega) \,:\, \boldsymbol{q}(\boldsymbol{x}) \cdot n(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in \Gamma_N \text{ and } \nabla \cdot \boldsymbol{q}(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in \Omega\},$$

$H^{\mathrm{div}}(\Omega)$ is the Sobolev space of square integrable functions with square integrable divergences on Ω. We remember from chapter 1 that the elements of $\mathcal{S}(\Omega)$ are said to be statically admissible (SA).

The solution of the boundary value problems in eq. (4.10) and eq. (4.2) are related by the constitutive relation, u *i.e.* $\boldsymbol{q} = -k\nabla u$. In words, $\boldsymbol{q}$ is the flux field of the temperature field u. A proof of this fact was given in section 3.2.

Problem (4.10) is also intractable. Accordingly, we also introduce the following surrogate model,

For all $\boldsymbol{p} \in \mathcal{S}_0$, find $\hat{\boldsymbol{q}} \in \mathcal{S}$ such that

$$\int_\Omega \hat{k}^{-1}(\boldsymbol{x})\, \hat{\boldsymbol{q}}(\boldsymbol{x}) \cdot \boldsymbol{p}(\boldsymbol{x})\, d\Omega = -\int_{\Gamma_D} \boldsymbol{p}(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x})\, D(\boldsymbol{x})\, d\Gamma \tag{4.11}$$

where $\hat{k}$ is again a homogenised, deterministic conductivity field, possibly different from $\bar{k}$. Similarly, if $\hat{k} = \bar{k}$, the homogenised problems will be related too by the constitutive relation, $\hat{\boldsymbol{q}} = -\bar{k}\nabla\bar{u}$.

We will seek approximations of eq. (4.11) through the hybrid equilibrium finite element formulation described in chapter 3. Alternatively, if $\bar{k} = \hat{k}$, a conforming approximation can be constructed from $\bar{u}^h$ by using the element equilibration technique described in section 1.4.2.

4.3.2 Estimates for the error in the energy norm

4.3.2.1 Upper bound for the energy-norm of the error

In this section, we aim to obtain an upper bound to the quantity $\|e\|$. Since the source term and the boundary conditions are deterministic, our pair of approximations $\bar{u}^h$ and $\hat{\boldsymbol{q}}^h$ are KA and SA for each realisation θ and the Prager-Synge hypercircle theorem can be applied. This results in

$$\|\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h\|^2_{k^{-1},\theta} = \|\mathbf{e}^q\|^2_{k^{-1},\theta} + \|e\|^2_\theta \tag{4.12}$$

where

$$\|\boldsymbol{v}\|^2_{k^{-1},\theta} = \int_\Omega k^{-1}(\boldsymbol{x},\theta)\boldsymbol{v}(\boldsymbol{x})\cdot\boldsymbol{v}(\boldsymbol{x})\,d\Omega$$
$$\mathbf{e}^q = -k\nabla u - \hat{\boldsymbol{q}}^h,$$

the former being the norm induced by the bilinear form of eq. (4.10) and the latter being its associated error field. By taking expectations on both sides and by the linearity of the operator, we obtain

$$E[\|\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h\|^2_{k^{-1},\theta}] = E[\|\mathbf{e}^q\|^2_{k^{-1},\theta}] + \|e\|^2. \tag{4.13}$$

The bounds result from the non-negativity of the terms in eq. (4.13),

$$\|e\|^2 \leq E[\|\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h\|^2_{k^{-1},\theta}] =: \eta^2 \tag{4.14}$$

Since this result is an extension of theorem 1.1, we will also name η as the constitutive relation error (CRE). This result is of utmost importance since the local error estimates are also based on this result. For this reason, we further characterise this estimate.

Firstly, we show its computation is fully deterministic and does not involve the generation of any realisation. By expanding η^2, and applying Fubini's theorem [21] to change the order of the expectation operator and the domain integral, we obtain

$$\begin{aligned}\eta^2 = \int_\Omega E[k^{-1}](\boldsymbol{x})\hat{\boldsymbol{q}}^h(\boldsymbol{x})\cdot\hat{\boldsymbol{q}}^h(\boldsymbol{x})\,d\Omega + \int_\Omega E[k](\boldsymbol{x})\nabla\bar{u}^h(\boldsymbol{x})\cdot\nabla\bar{u}^h(\boldsymbol{x})\,d\Omega \\ +2\int_\Omega \hat{\boldsymbol{q}}^h(\boldsymbol{x})\cdot\nabla\bar{u}^h(\boldsymbol{x})\,d\Omega\end{aligned} \tag{4.15}$$

This makes the bound fully computable and guaranteed. Furthermore, as announced in the introduction, if the fields $E[k]$ and $E[k^{-1}]$ are of slow variation, then the bounds separate the scales and a specific integration mesh is not needed. This contrasts with earlier work [61, 65–68, 86, 87, 93, 103] where an integration mesh is needed which made the bound impractical for small particle-domain sizes ratio. In addition, the discretisation error was neglected (except for [61]) and due to the method used to derive the bounds, $\hat{\boldsymbol{q}}^h$ was prescribed implicitly to be equal to $-\bar{k}\nabla\bar{u}$, in other words, the same homogenised conductivity was used in both surrogate problems, $\bar{k} = \hat{k}$. This has a negative impact on the effectivity of the bound, since it depends on how well $\hat{\boldsymbol{q}}^h$ approximates the exact flux,

$$\frac{\eta}{\|e\|} = \sqrt{\frac{\|e\|^2 + E[\|\mathbf{e}^q\|^2_{k^{-1},\theta}]}{\|e\|^2}} = \sqrt{1 + \frac{E[\|\mathbf{e}^q\|^2_{k^{-1},\theta}]}{\|e\|^2}}. \tag{4.16}$$

We will later show in section 4.3.2.3 that the optimal effectivity is attained when $\hat{k} = E[k^{-1}]$.

We now return to the alternative measures of the error mentioned in section 4.2.2, namely $\|e\|_\theta$ and $E[\|e\|_\theta]$. By considering the realisations in isolation and direct application of theorem 1.1 a bound for $\|e\|_\theta$ follows

$$\|e\|_\theta \leq \|\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h\|_{k^{-1},\theta}. \tag{4.17}$$

Taking expectations on this equation, a bound for $E[\|e\|_\theta]$ is also obtained

$$E[\|e\|_\theta] \leq E[\|\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h\|_{k^{-1},\theta}]. \tag{4.18}$$

The requirement of the generation of integration meshes for each of the realisations limits their practical application. Nonetheless, from the definition of variance, we obtain

$$\mathrm{Var}[\|e\|_\theta] := E[(\|e\|_\theta - E[\|e\|_\theta])^2] = \|e\|^2 - E[\|e\|_\theta]^2 \Rightarrow E[\|e\|_\theta]^2 \leq \|e\|^2 \tag{4.19}$$

This means that η also upper bounds $E\left[\|e\|_\theta\right]$ with effectivity

$$\frac{\eta}{E[\|e\|_\theta]} = \sqrt{1 + \frac{\mathrm{Var}[\|e\|_\theta]}{E[\|e\|_\theta]^2} + \frac{E[\|\mathbf{e}^q\|^2_{k^{-1},\theta}]}{E[\|e\|_\theta]^2}}. \tag{4.20}$$

Loosely speaking, if the variance is small, the difference between $E[\|e\|_\theta]$ and $\|e\|$ for each particular realisation is expected to be small. Hence, we may also expect η to be a practical

upper bound for each of the realisations independently, *i.e.* an upper bound for $\|e\|_\theta$. In conclusion, under those assumptions, in the case that an accurate description of the microstructure was available on the entire domain, we still could apply the techniques here described and avoid generating a integration mesh which adjusts to the phases.

Before moving onto the development of lower bounds, we mention that it is also possible to derive bounds for $E[\|\mathbf{e}^q\|^2_{k^{-1},\theta}]$, however we did not explore this possibility in the present work.

4.3.2.2 A lower bound for the error in energy norm

The derivation of the lower bound starts with the following identity,

$$\|e\| = \sup_{v \in \mathcal{U}_0^\theta \setminus \{0\}} \frac{|a(e,v)|}{\|v\|}. \tag{4.21}$$

which is a direct consequence of the Cauchy-Schwarz inequality. Combining the identity with eq. (4.6),

$$\sup_{v \in \mathcal{U}_0^\theta \setminus \{0\}} \frac{|a(e,v)|}{\|v\|} = \sup_{v \in \mathcal{U}_0^\theta \setminus \{0\}} \frac{|R(v)|}{\|v\|} \geq \frac{|R(w)|}{\|w\|} \quad \forall w \in \mathcal{U}_0^\theta(\Omega) \setminus \{0\} \tag{4.22}$$

where

$$R(v) := E[R_\theta(v)]. \tag{4.23}$$

Therefore, any element of $\mathcal{U}_0^\theta(\Omega)$ will result in a lower bound.

To simplify the computation of the bound, it is desirable to choose w in $\bar{\mathcal{U}_0}(\Omega)$. The element in $\bar{\mathcal{U}_0}$ that maximises the lower bound is Πe where Π is the orthogonal projection operator on $\bar{\mathcal{U}_0}$ with respect to $a(\cdot,\cdot)$. Indeed, by using the defining properties of orthogonal projection and the Cauchy-Schwarz inequality, we obtain the following

$$\frac{|R(w)|}{\|w\|} = \frac{|a(e,w)|}{\|w\|} = \frac{|a(e,\Pi w)|}{\|w\|} = \frac{|a(\Pi e,w)|}{\|w\|} \leq \|\Pi e\| \quad \forall w \in \bar{\mathcal{U}_0}(\Omega). \tag{4.24}$$

This proves that $\|\Pi e\|$ is an upper bound for the lower bound. The equality is achieved by setting $w = \Pi e$. Hence, the optimal lower bound in $\bar{\mathcal{U}_0}(\Omega)$ takes the form $\|\Pi e\|$.

4.3.2.3 Influence of the surrogate conductivities on the error bounds

In this section, we characterise and optimise the error in the energy norm and its estimates in terms of the homogenised conductivities $\bar{k}$ and $\hat{k}$. As we shall see, the optimisation can be always carried out analytically and the optima coincide with classical homogenisation bounds. We will also make the relation between the lower and upper bound explicit.

4.3.2.3.1 Optimum result for the upper bound. We show below how to choose $\hat{k}$ in order to maximise the effectivity of the upper bound, and how to choose $\bar{k}$ in order to minimise the error itself. These two optimisation processes are analytical and independent.

We start by noting that if we set $\bar{k}$, the conductivity of the temperature problem to $E[k]$, then the error in energy norm, $\|e\|$, is minimised. This is a direct outcome of the minimisation of $\|e\| = \|u - \bar{u}^h\|$ in the finite dimensional space of deterministic, kinematically admissible fields $\bar{\mathcal{U}}(\Omega)$. This is shown by taking the first variation of $\|u-\bar{u}^h\|$ to obtain that in the stationary point,

$$a(\bar{u}^h, v) = a(u, v) = l(v) \quad \forall v \in \bar{\mathcal{U}}_0(\Omega). \tag{4.25}$$

Expanding $a(\bar{u}^h, v)$, and using the fact that both $\bar{u}^h$ and v are deterministic, we obtain

$$E\left[\int_\Omega k(\boldsymbol{x},\theta)\,\nabla\bar{u}^h(\boldsymbol{x})\cdot\nabla v(\boldsymbol{x})\,d\Omega\right] = \int_\Omega E[k](\boldsymbol{x})\,\nabla\bar{u}^h(\boldsymbol{x})\cdot\nabla v(\boldsymbol{x})\,d\Omega \tag{4.26}$$

Therefore, if $\bar{k} = E[k]$, it follows that $a(\cdot,\cdot) = \bar{a}(\cdot,\cdot)$ and that $\|e\|$ will be minimised.

The field $E[k]$ is also a rule of mixtures because it is also a weighted average of the properties of the constituents (see section 2.2.3). This allows to understand the optima as an extension of the Voigt assumption. Indeed, in order to obtain homogenised properties, Voigt constrains the strain field in a heterogeneous material to be constant. The resulting homogenised tensor is obtained by applying the spatial rule of mixture. In our context, under the assumption that the temperature field is constant in the stochastic domain (deterministic), the reference temperature field u is best approached in norm $\|\cdot\|$ when using the stochastic rule of mixture.

A similar argument can be used to show that setting $\hat{k} = 1/E[k^{-1}]$, will optimise the effectivity of the upper bound and that it is an inverse rule of mixture. This is an

extension of Reuss' theory. Reuss showed that by assuming that the stress field is constant in a heterogeneous material, the resulting approximate effective tensor is the inverse rule of mixture. In our case, under the assumption that the flux field is deterministic, the solution of eq. (4.11) is best approximated using the inverse rule of mixture as the effective conductivity.

4.3.2.3.2 Remark concerning the lower bound and the rule of mixture. Choosing the rule of mixture as our surrogate conductivity, which minimises the error in energy norm, leads to a trivial lower bound. This can be shown by noting that in this case,

$$\bar{a}(v, w) = a(v, w) \tag{4.27}$$

for any deterministic fields v and w. Hence, $R(v) = l(v) - a(\bar{u}^h, v) = 0$ for any deterministic v, which leads to the lower bound $\|\Pi e\| = 0$. This can be understood as a result of Galerkin orthogonality. Despite of this result and the result above, we will see that when model adaptivity is used $E[k]$ is not optimal and the lower bound will be essential for the obtention of sharp estimates.

4.3.2.3.3 Relation between the lower and upper bound. We show now how the lower and upper bound are related. This result will be crucial to characterise the estimates for quantities of interest.

More specifically, we prove that the distance between the square of upper bound and the square of the lower bound,

$$\eta^2 - \|\Pi e\|^2 \tag{4.28}$$

does not depend on the conductivity field $\bar{k}$. Note that the previous expression is the uncertainty interval length for $\|e\|^2$, *i.e.*

$$\|\Pi e\|^2 \leq \|e\|^2 \leq \eta^2. \tag{4.29}$$

In order to show this, we observe that

$$\|e\|^2 = \|e - \Pi e\|^2 + \|\Pi e\|^2 \tag{4.30}$$

due to the orthogonality of $e - \Pi e$ and Πe, $a(e - \Pi e, \Pi e) = 0$. Hence, by expanding eq. (4.28)

$$\|\mathbf{e}^q\|^2_{k^{-1},\theta} + \|e\|^2 - \|\Pi e\|^2 = \|\mathbf{e}^q\|^2_{k^{-1},\theta} + \|e - \Pi e\|^2. \tag{4.31}$$

The proof is concluded by showing that that $\|e - \Pi e\|$ does not depend on $\bar{k}$. By writing e in terms of u and $\bar{u}^h$, we obtain

$$e - \Pi e = u - \bar{u}^h + \Pi(u - \bar{u}^h) = u - \Pi u \tag{4.32}$$

where neither u, nor Πu depend on the field $\bar{k}$, proving the desired result. More precisely, $\bar{u}^h = \Pi u$ if we set $\bar{k} = E\,[k]$. We note that the field Πu coincides with $\bar{u}^h$ if we choose $\bar{k} = E[k]$.

4.3.3 Error bounds for quantities of interest

4.3.3.1 Error bounds for the expectation

In this section, we expand the discussion started in section 1.5 to derive error bounds for quantities of interest. The techniques presented here are standard in the literature (see [11, 79] for discretisation error and [61, 65–68, 86, 87, 93, 103] for model error).

We start by introducing the dual (or adjoint) problem,

For each $\theta \in \Theta$ and for all $v(\boldsymbol{x}) \in \mathcal{U}_0(\Omega)$, find $\phi(\boldsymbol{x}, \theta) \in \mathcal{U}_0(\Omega)$ such that,

$$a_\theta(\phi, v) = s_\theta(v). \tag{4.33}$$

and its homogenised counterpart

For all $v \in \mathcal{U}_0(\Omega)$, find $\bar{\phi} \in \mathcal{U}_0(\Omega)$

$$\bar{a}(\bar{\phi}, v) = s(v) \tag{4.34}$$

where $s(v) = E[s_\theta(v)]$ is the QoI and $\bar{a}$ is the bilinear form defined in eq. (4.4) with possibly a different homogenisation scheme. These problems are called dual problems, in contrast to eqs. (4.2) and (4.3) that are also called primal problems. A deterministic KA approximation of $\bar{\phi}$ will be denoted by $\bar{\phi}^h$, while the SA approximation of its flux will be denoted by $\hat{\boldsymbol{q}}^h_\phi$. Using a standard procedure, we can show that,

$$\begin{aligned} s(u) - s(\bar{u}^h) &= s(e) && (\textit{linearity of } q) \\ &= a(\phi, e) && (\textit{eq. } (4.33)) \\ &= a(\phi - \bar{\phi}^h, e) + a(\bar{\phi}^h, e) && (\textit{adding and subtracting } a(\bar{\phi}^h, e)) \\ &= a(e_\phi, e) + R(\bar{\phi}^h) && (\textit{eq. } (4.6)) \end{aligned}$$

where $e_\phi = \phi - \bar{\phi}^h$. Bounds for the term $a(e_\phi, e)$ can be derived in at least two different ways. We review them both. Firstly, by applying Cauchy-Schwarz to the product of the errors, we obtain

$$R(\bar{\phi}^h) - \|e\|\|e_\phi\| \leq s(u) - s(\bar{u}^h) \leq R(\bar{\phi}^h) + \|e\|\|e_\phi\| \tag{4.35}$$

and using the bounds in energy-norm developed in section 4.3.2,

$$R(\bar{\phi}^h) - \eta\eta_\phi \leq s(u) - s(\bar{u}^h) \leq R(\bar{\phi}^h) + \eta\eta_\phi \quad \text{with } \|e_\phi\| \leq \eta_\phi. \tag{4.36}$$

Alternatively, bounds can be derived by noting that

$$a(e, e_\phi) = a(\alpha e,\, \alpha^{-1} e_\phi) \quad \forall \alpha \in \mathbb{R} \setminus \{0\} \tag{4.37}$$

and using the polarisation identity (as in [80]) to conclude that

$$s(u) - s(\bar{u}^h) = R(\bar{\phi}^h) + \frac{\|\alpha e + \alpha^{-1} e_\phi\|^2 - \|\alpha e - \alpha^{-1} e_\phi\|^2}{4}. \tag{4.38}$$

Each of the terms $\|\alpha e \pm \alpha^{-1} e_\phi\|$ can be upper and lower bounded. Due to the bilinearity of $a(\cdot, \cdot)$, the bounds from section 4.3.2 can be applied to obtain,

$$(\eta_L^\pm)^2 \leq \|\alpha e \pm \alpha^{-1} e_\phi\|^2 \leq (\eta_U^\pm)^2 \tag{4.39}$$

where the following notation was used for compactness,

$$\begin{aligned}\eta_L^{\pm} &:= \|\alpha\Pi e \pm \alpha^{-1}\Pi e_\phi\| \\ \eta_U^{\pm} &:= \sqrt{E[\|\hat{\boldsymbol{q}}^h_{\alpha^\pm} + k\nabla \bar{u}^h_{\alpha^\pm}\|^2_{k^{-1},\theta}]} \\ \bar{u}^h_{\alpha^\pm} &:= \alpha\bar{u}^h \pm \alpha^{-1}\bar{\phi}^h \\ \hat{\boldsymbol{q}}^h_{\alpha^\pm} &:= \alpha\hat{\boldsymbol{q}}^h \pm \alpha^{-1}\hat{\boldsymbol{q}}^h_\phi\end{aligned}$$

The results can be summarised in the following equation,

$$\underbrace{R(\bar{\phi}^h) - \frac{1}{4}(\eta_U^-)^2 + \frac{1}{4}(\eta_L^+)^2}_{=:\eta^{\text{QoI}}_{\text{Low}}} \leq s(u) - s(\bar{u}^h) \leq \underbrace{R(\bar{\phi}^h) + \frac{1}{4}(\eta_U^+)^2 - \frac{1}{4}(\eta_L^-)^2}_{=:\eta^{\text{QoI}}_{\text{Upp}}} \tag{4.40}$$

We set

$$\alpha = \sqrt[4]{\frac{\eta_\phi^2 - \|\Pi e_\phi\|^2}{\eta^2 - \|\Pi e\|^2}}. \tag{4.41}$$

With this choice, the bound is optimal. The upper bound, $\eta^{\text{QoI}}_{\text{Upp}}$, is minimised and the lower bound, $\eta^{\text{QoI}}_{\text{Low}}$, is maximised. Moreover, this bound is sharper than eq. (4.36) and the resulting interval length, $\eta^{\text{QoI}}_{\text{Upp}} - \eta^{\text{QoI}}_{\text{Low}}$, is at most half of it. These results are proven in detail in appendix A.

Remark 1. *We note that the bound in eq.* (4.36) *could be made sharper by considering it for a single realisation and then taking expectations. Indeed,*

$$s_\theta(u) - s_\theta(\bar{u}^h) \leq R_\theta(\bar{\phi}^h) + \|e\|_\theta\|e_\phi\|_\theta \leq \|\hat{\boldsymbol{q}}^h + k\bar{u}^h\|_{k^{-1},\theta}\|\hat{\boldsymbol{q}}^h_\phi + k\bar{\phi}^h\|_{k^{-1},\theta} \tag{4.42}$$

$$s(u) - s(\bar{u}^h) \leq R(\bar{\phi}^h) + E\left[\|\hat{\boldsymbol{q}}^h + k\bar{u}^h\|_{k^{-1},\theta}\|\hat{\boldsymbol{q}}^h_\phi + k\bar{\phi}^h\|_{k^{-1},\theta}\right] \tag{4.43}$$

It follows from the application of the Cauchy-Schwarz inequality to $E[\cdot]$ *that,*

$$E\left[\|\hat{\boldsymbol{q}}^h + k\bar{u}^h\|_{k^{-1},\theta}\|\hat{\boldsymbol{q}}^h_\phi + k\bar{\phi}^h\|_{k^{-1},\theta}\right] \leq \eta\eta_\phi. \tag{4.44}$$

However, this discussion does not extend to eq. (4.40)*. This follows from the fact that* $(\eta_L^\pm)^2$ *and* $(\eta_U^\pm)^2$ *are left unchanged if independently computed for each realisation, followed by the computation of the expectation.*

4.3.3.2 Choice of the homogenised conductivity

We readily observe that the uncertainty length for $s(e)$ in eq. (4.40),

$$\eta_{\text{Upp}}^{\text{QoI}} - \eta_{\text{Low}}^{\text{QoI}} = \frac{1}{4}\left[(\eta_U^+)^2 - (\eta_L^+)^2 + (\eta_U^-)^2 - (\eta_L^-)^2\right] = \sqrt{(\eta^2 - \|\Pi e\|^2)(\eta_\phi^2 - \|\Pi e_\phi\|^2)} \quad (4.45)$$

is independent of the field $\bar{k}$. This is a consequence of the consideration in section 4.3.2.3 which shows that $\eta^2 - \|\Pi e\|^2$ and by extension $\eta_\phi^2 - \|\Pi e_\phi\|^2$ are independent of the surrogate conductivity field $\bar{k}$. In fact, the lower and upper bounds for $s(u)$,

$$s(\bar{u}^h) + \eta_{\text{Low}}^{\text{QoI}} \leq s(u) \leq s(\bar{u}^h) + \eta_{\text{Upp}}^{\text{QoI}} \quad (4.46)$$

are also independent of the field $\bar{k}$ (see appendix B for the details). In conclusion, the choice of the $\bar{k}$ has no influence in the estimation of error in QoIs. This means that we cannot obtain a better estimation of $s(u)$ by considering different $\bar{k}$ and combining all the bounds. All possible choices of the field $\bar{k}$ will result in the same estimate.

Regarding $\hat{k}$, the same consideration presented in section 4.3.2.3 applies and hence we set $\hat{k} = 1/E[k^{-1}]$.

Remark 2. *If $\bar{k} = E[k]$ in the primal problem and the same discretisation is used to solve eq.* (4.34) *(though not necessarily the same homogenisation scheme), then*

$$R(\bar{\phi}^h) = l(\bar{\phi}^h) - a(\bar{u}^h, \bar{\phi}^h) = l(\bar{\phi}^h) - \bar{a}(\bar{u}^h, \bar{\phi}^h) = 0 \quad (4.47)$$

since $\bar{\phi}^h \in \bar{\mathcal{U}}_0(\Omega)$.

4.3.3.3 Error bound for the second moment

Following a similar methodology to the one presented in section 4.3.3.1, it can be shown that the second moment of the quantity of interest,

$$m_2 = E[s_\theta(u)^2], \quad (4.48)$$

is bounded by

$$m_2 \leq s(\bar{u}^h)^2 + (\gamma + \beta)^2 + 2\Delta[l(\bar{\phi}^h) + s(\bar{u}^h)] + l(\bar{\phi}^h)[R(\bar{\phi}^h) - a(\bar{u}^h, \bar{\phi}^h)] =: \zeta_2 \quad (4.49)$$

under the non-restrictive assumption that $s_\theta(\bar{u}^h) > 0$ and $l(\bar{\phi}^h) > 0$ (the expression is slightly different if either of those terms is negative) and where

$$\gamma = \sqrt{E\left[\int_\Omega k^{-1}(\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h)\,d\Omega \int_\Omega k^{-1}(\hat{\boldsymbol{q}}^h_\phi + k\nabla\bar{\phi}^h)\,d\Omega\right]} \tag{4.50}$$

$$\beta = \sqrt{E\left[\int_\Omega k\nabla\bar{u}^h \cdot \nabla\bar{\phi}^h\right]^2} \tag{4.51}$$

and Δ is an upper estimate for $s(e)$ from section 4.3.3.1. The complete result and a proof are given in appendix D.

The quantities β and γ are also deterministic. In other words, their computation does not involve the generation of realisations of the domain. This is shown by transforming the product of domain integrals into a double domain integral (see appendix D) by grouping together the stochastic terms. For instance, β^2 can be expanded to

$$\int_\Omega \int_\Omega E\left[k(\boldsymbol{x})k(\boldsymbol{x}')\right](\nabla\bar{u}^h \cdot \nabla\bar{\phi}^h)(\boldsymbol{x})(\nabla\bar{u}^h \cdot \nabla\bar{\phi}^h)(\boldsymbol{x}')d\Omega\,d\Omega \tag{4.52}$$

The term $E\left[k(\boldsymbol{x})k(\boldsymbol{x}')\right]$ is directly related to the covariance function $\mathrm{Cov}(k(\boldsymbol{x}), k(\boldsymbol{x}')) = E[k(\boldsymbol{x})k(\boldsymbol{x}')] - E[k(\boldsymbol{x})]E[k(\boldsymbol{x}')]$ and loosely speaking, it incorporates in the bound information regarding the shape of the particles whilst the expectation includes information of the volume fraction. More accurately, $E\left[k(\boldsymbol{x})\right]$ can be constructed from functions $p_i : \Omega \to [0, 1]$ which give the probability that a point lies inside a certain heterogeneity indexed by i. Then

$$E\left[k\right](\boldsymbol{x}) = \sum_i k_i p_i(\boldsymbol{x}) \tag{4.53}$$

In many practical cases, those functions are constants whose value coincide with the volume fraction of the heterogeneity i. Similarly, $E[k(\boldsymbol{x})k(\boldsymbol{x}')]$ can be constructed from functions $p_{ij} : \Omega \times \Omega \to [0, 1]$ which given a pair of points $(\boldsymbol{x}, \boldsymbol{x}')$ returns the probability that point $\boldsymbol{x}$ lies in heterogeneity i and $\boldsymbol{x}'$ lies in heterogeneity j in the same realisation. These functions, due to their dependence of two points of the domain, will depend on the specific shape of the particles.

We conclude this section extending the bound for the second moment to the variance. From the definition of the variance,

$$\mathrm{Var}[s_\theta(u)] = m_2 - s(u)^2 \tag{4.54}$$

it follows that a bound for the variance can be obtained by combining the bound for the second moment and a lower bound for $s(u)^2$. Let L and U be any of the lower and upper bounds for $s(u)$ developed in section 4.3.3.1, then

$$s(u)^2 \geq \begin{cases} L^2 & \text{if } L \geq 0 \\ U^2 & \text{if } U < 0 \\ 0 & \text{if } L < 0 \text{ and } 0 < U. \end{cases} \tag{4.55}$$

Then, an upper bound for the variance reads

$$\mathrm{Var}[s_\theta(u)] = m_2 - s(u)^2 \leq \begin{cases} \zeta_2 - L^2 & \text{if } L \geq 0 \\ \zeta_2 - U^2 & \text{if } U < 0 \\ \zeta_2 & \text{if } L < 0 \text{ and } 0 < U. \end{cases} \tag{4.56}$$

Due to the inequalities used in the derivation, we expect the second moment to not be as efficient as the bounds for the expectation. In addition, since the bound for the variance results from the difference of two bounds, we expect it to be less efficient than the bound for the second moment.

4.3.4 Adaptive modelling

Tighter error estimates can be obtained by solving the true model only in a subset of the domain. The resulting surrogate model is called adaptive model. In this section, we describe the obtention of the adaptive model and we introduce error indicators. Error indicators guide the construction of the adaptive model by signalling where the true model should be solved.

The adaptive methodology followed here is similar to [86], though the proposed error estimates and error indicators differ. The error estimates proposed in [86] do not converge

to 0 with model refinement and the error indicator selects regions for model refinement whose resolution might not necessarily improve the bounds. Our error estimates and indicators address those issues. A more technical discussion of those two issues is given in the remarks at the end of this section.

We start by dividing the domain in N disjoint subdomains Ω_i, *i.e.*

$$\cup_{i=1}^{N} \Omega_i = \Omega \qquad \Omega_i \cap \Omega_j = \emptyset \quad (i \neq j). \tag{4.57}$$

This is followed by the computation of the error indicators for each of the subdomains. For the error in energy norm,

$$\eta_{\Omega_i}^2 = E\left[\int_{\Omega_i} k^{-1}(\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h)^2 \, d\Omega\right]. \tag{4.58}$$

Notice that the sum of the error indicators is equal to the upper bound,

$$\eta^2 = \sum_{i=1}^{N} \eta_{\Omega_i}^2 \tag{4.59}$$

The error indicator for the quantity of interest is given by

$$\begin{aligned}\beta_{\Omega_i}^2 = \frac{\alpha^2}{2} E\left[\int_{\Omega_i} k^{-1}(\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h)^2 \, d\Omega - \int_{\Omega_i} (\Pi e)^2 \, d\Omega\right] + \\ \frac{\alpha^{-2}}{2} E\left[\int_{\Omega_i} k^{-1}(\hat{\boldsymbol{q}}_\phi^h + k\nabla\bar{\phi}^h)^2 \, d\Omega - \int_{\Omega_i} (\Pi e_\phi)^2 \, d\Omega\right].\end{aligned} \tag{4.60}$$

In this case, their sum,

$$\eta_{\text{Upp}}^{\text{QoI}} - \eta_{\text{Low}}^{\text{QoI}} = \sum_{i=1}^{N} \beta_{\Omega_i}^2, \tag{4.61}$$

is equal to the interval length of eq. (4.40).

The subdomains Ω_i whose error indicators are above a certain user predefined tolerance are selected to be part of the adaptive model. Both error indicators, η_{Ω_i} and β_{Ω_i}, depend on the choice of the conductivity field. Nevertheless, in our numerical experiments the dependence was small enough to not change the domains selected for refinement.

Let W be the set of indexes of subdomains that are selected for refinement. We define the following pair of problems for each $w \in W$:

For each $\theta \in \Theta$ and for all $v \in \mathcal{U}_0(\Omega_w)$, find $\tilde{u}_{\Omega_w}(\boldsymbol{x}, \theta) \in \mathcal{U}(\Omega_w)$ such that

$$\int_{\Omega_w} k(\boldsymbol{x}, \theta)\, \nabla \tilde{u}_{\Omega_w}(\boldsymbol{x}, \theta) \cdot \nabla v(\boldsymbol{x}, \theta) = \int_{\Omega_w} f(\boldsymbol{x}) v(\boldsymbol{x})\, d\Omega - \int_{\partial\Omega_w \cap \Gamma_N} g(\boldsymbol{x}) v(\boldsymbol{x})\, d\Gamma. \quad (4.62)$$

where

$$\begin{aligned} \mathcal{U}(\Omega_w) &= \{v \in H^1(\Omega_w) : v(\boldsymbol{x}) = \bar{u}^h(\boldsymbol{x}) \qquad \forall \boldsymbol{x} \in \partial\Omega_w \backslash \Gamma_N\} \\ \mathcal{U}_0(\Omega_w) &= \{v \in H^1(\Omega_w) : v(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in \partial\Omega_w \backslash \Gamma_N\}. \end{aligned}$$

For each $\theta \in \Theta$ and for all $p \in \mathcal{S}_0(\Omega_w)$, find $\tilde{\boldsymbol{q}}_{\Omega_w} \in \mathcal{S}(\Omega_w)$

$$\int_{\Omega_w} k^{-1}(\boldsymbol{x}, \theta) \tilde{\boldsymbol{q}}_{\Omega_w}(\boldsymbol{x}, \theta) \cdot \boldsymbol{p}(\boldsymbol{x}) d\Omega = \int_{\partial\Omega_w \cap \Gamma_D} \boldsymbol{p}(\boldsymbol{x}) \cdot \boldsymbol{n}(\boldsymbol{x})\, D(\boldsymbol{x}) d\Gamma \quad (4.63)$$

where

$$\begin{aligned} \mathcal{S}(\Omega_w) = \{\boldsymbol{q} \in H^{\mathrm{div}}(\Omega_w)\ &: \boldsymbol{q}(\boldsymbol{x}) \cdot n(\boldsymbol{x}) = \hat{\boldsymbol{q}}^h(\boldsymbol{x}) \quad \forall \boldsymbol{x} \in \partial\Omega_w - \Gamma_D \\ &\text{and} \quad \nabla \cdot \boldsymbol{q}(\boldsymbol{x}) = f(\boldsymbol{x}) \qquad \forall \boldsymbol{x} \in \Omega_w\}, \\ \mathcal{S}_0(\Omega_w) = \{\boldsymbol{q} \in H^{\mathrm{div}}(\Omega_w)\ &: \boldsymbol{q}(\boldsymbol{x}) \cdot n(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in \partial\Omega_w - \Gamma_D \\ &\text{and} \quad \nabla \cdot \boldsymbol{q}(\boldsymbol{x}) = 0 \qquad \forall \boldsymbol{x} \in \Omega_w\}. \end{aligned}$$

In words, we seek to improve our temperature and flux approximation by solving only local stochastic problems to keep them tractable. In the temperature case, we prescribe our local approximation to coincide with our deterministic approximation. Equivalently, we prescribe our local flux approximation to coincide with $\hat{\boldsymbol{q}}^h$. Owing to the fact $\hat{\boldsymbol{q}}^h$ is SA, the problem in eq. (4.63) is well-posed.

Discrete KA and SA approximations $\tilde{u}^h_{\Omega_w}$ and $\tilde{\boldsymbol{q}}^h_{\Omega_w}$ are sought to the problems in combination with the Monte Carlo method. It is assumed that sufficient realisations are considered and for practical reasons, we have used a very fine KA approximation as $\tilde{\boldsymbol{q}}^h_{\Omega_w}$. As it will be seen in the examples, for fine meshes, the discretisation error is negligible

when compared with the homogenisation error and consequently, the impact on the error bounds is negligible.

Using χ, the set indicator function, the approximations to the local problems can be combined to form the adaptive solution,

$$\tilde{u}^h = \bar{u}^h + \sum_{w \in W} \chi_{\Omega_w} (\tilde{u}^h_{\Omega_w} - \bar{u}^h)$$
$$\tilde{\boldsymbol{q}}^h = \hat{\boldsymbol{q}}^h + \sum_{w \in W} \chi_{\Omega_w} (\tilde{\boldsymbol{q}}^h_{\Omega_w} - \hat{\boldsymbol{q}}^h).$$

Similarly, on the same subdomains, local approximations $\tilde{\phi}^h_{\Omega_w}$ and $\tilde{\boldsymbol{q}}^h_{\phi\Omega_w}$ of the dual true model are sought and used to build $\tilde{\phi}^h$ and $\tilde{\boldsymbol{q}}^h_\phi$. Note that $\tilde{u}^h$ and $\tilde{\phi}^h$ are kinematically admissible and $\tilde{\boldsymbol{q}}^h$ and $\tilde{\boldsymbol{q}}^h_\phi$ are statically admissible. Hence, the bounds presented in sections 4.3.2 and 4.3.3 can be immediately applied.

We note that the resulting bounds do depend on the homogenisation scheme in contrast to the fully deterministic case. This follows from the fact that the we use $\bar{u}^h, \bar{\phi}^h, \hat{\boldsymbol{q}}^h$ and $\tilde{\boldsymbol{q}}^h_\phi$ to define the boundary conditions of the BVPs that the define the local approximations. Moreover, the introduction of the local approximations may result in a less sharp lower or upper bound for QoIs. In these regions, the integrand of $\eta^\pm_L$ reduces to 0. This reduction may not be compensated by the increase of sharpness of $\eta^\pm_U$. The example in section 4.4.5 illustrates both effects.

From an implementation point of view, we compute and store the restriction to the subdomains Ω_i of each of the terms that appear in the expression of the error bound, eq. (4.40). In other words, we calculate the quantities $R(\bar{\phi}^h)_{\Omega_i}$, $(\eta^\pm_L)^2_{\Omega_i}$ and $(\eta^\pm_U)^2_{\Omega_i}$ such that,

$$R(\bar{\phi}^h) = \sum_{i=1}^{N} R(\bar{\phi}^h)_{\Omega_i}, \quad (\eta^\pm_L)^2 = \sum_{i=1}^{N} (\eta^\pm_L)^2_{\Omega_i}, \quad (\eta^\pm_U)^2 = \sum_{i=1}^{N} (\eta^\pm_U)^2_{\Omega_i}. \tag{4.64}$$

Those quantities can be combined to obtain the $\eta^{\text{QoI}}_{\text{Low}}$, $\eta^{\text{QoI}}_{\text{Upp}}$ and the error indicators. The refined error bounds can be obtained by recomputing $R(\bar{\phi}^h)_{\Omega_i}$, $(\eta^\pm_L)^2_{\Omega_i}$, and $(\eta^\pm_U)^2_{\Omega_i}$ only in

the refined subdomains and using

$$R(\tilde{\phi}^h) = R(\bar{\phi}^h) - \sum_{w\in W}\left(R(\bar{\phi}^h)_{\Omega_w} - R(\tilde{\phi}^h)_{\Omega_w}\right)$$
$$(\eta_L^{\pm})^{2\prime} = (\eta_L^{\pm})^2 - \sum_{w\in W}\left((\eta_L^{\pm})^2_{\Omega_w} - (\eta_L^{\pm})^{2\prime}_{\Omega_w}\right)$$
$$(\eta_U^{\pm})^{2\prime} = (\eta_U^{\pm})^2 - \sum_{w\in W}\left((\eta_U^{\pm})^2_{\Omega_w} - (\eta_U^{\pm})^{2\prime}_{\Omega_w}\right)$$

where the prime indicates that quantity is related to the adaptive model.

Remark 3. *In [68], the authors prove the following alternative error bound (which is extended to QoIs in [87]),*

$$\|u - \tilde{u}\|^2 \leq 2[J(\tilde{u}) - J(\bar{u})] + \eta^2 \tag{4.65}$$

where

$$J(v) = \frac{1}{2}a(v,v) - l(v), \tag{4.66}$$

under the assumption that the discretisation errors are negligible. The drawback of this error bound is that it does not converge to 0 *with model refinement. Indeed, by adding and subtracting* $J(u)$ *and using the identity,* $\|u - v\|^2 = 2J(v) - 2J(u)$ *(see [68]) for* v *KA,*

$$\|u - \tilde{u}\|^2 \leq \|u - \tilde{u}\|^2 + E\left[\|\hat{\boldsymbol{q}} + k\nabla u\|^2_{k^{-1}}\right] \tag{4.67}$$

In other words, the effectivity of this bound is controlled by $\hat{\boldsymbol{q}}$*, a field which is not affected by model refinement.*

Remark 4. *An alternative error indicator for QoIs,*

$$\kappa_{\Omega_i} = \eta_{\Omega_i}\|\bar{\phi}^h\|_{\Omega_i} + \eta_{\Omega_i}\,\eta_{\phi\,\Omega_i} \tag{4.68}$$

where

$$\|\bar{\phi}^h\|_{\Omega_i} = E\left[\int_{\Omega_i} k\nabla\bar{\phi}^h\cdot\nabla\bar{\phi}^h\,d\Omega\right]$$
$$\eta_{\phi\,\Omega_i} = E\left[\int_{\Omega_i} k^{-1}(\hat{\boldsymbol{q}}^h_\phi + k\nabla\bar{\phi}^h)^2\,d\Omega\right].$$

was proposed in [67, 87]. Recalling eq. (4.36) *and that* $R(\bar{\phi}^h) = a(e, \bar{\phi}^h)$*, it follows*

$$s(u) - s(\bar{u}^h) \leq a(e, \bar{\phi}^h) + \eta\eta_\phi \leq \eta\|\bar{\phi}^h\| + \eta\eta_\phi. \tag{4.69}$$

The error indicator follows from computing this upper bound for each subdomain.

However, we noticed that $R(\bar{\phi}^h)$ *contributes to this indicator whilst this term does not affect the interval length as seen in eq.* (4.36)*. As a consequence, the decrease of this term may not lead to a sharpening of the bound on the QoI (although the error itself would be decreased).*

The error indicator presented in section 4.3.4 was built with the interval length in mind. It does not suffer from this conceptual problem. Another advantage of this error estimate is that its value is the maximum possible reduction of the error interval length when the respective subdomain is refined. This follows by considering that the all the subdomains not selected for refinement will contribute with the same quantity, while the subdomains refined will contribute with a non negative quantity.

4.4 Numerical examples

In this section five numerical examples are presented. In the first example, the bounds are compared against a Monte Carlo reference solution of the heterogeneous problem. In the second example, we examine the effect of the contrast of the material properties and the volume fraction in the bounds. The next example applies the bounds on a complex 3D geometry. This example is followed by a 1D numerical simulation where the bound for the second moment is applied. The section is concluded with an application of model adaptivity.

The meshes used were generated using Gmsh [30]. The code made extensive use of the library Eigen [35] and the post processing was done in ParaView [36].

4.4.1 Validation of the bounds

We compare the solution of a random heterogeneous problem with the bounds obtained from the homogenised problem. Since the random heterogeneous problem cannot be solved

exactly, its solution is going to be approximated by an overkill Monte Carlo solution.

We consider the domain shown in figs. 4.2a and 4.2b. The source term, the prescribed fluxes and temperatures read

$$
\begin{aligned}
f(\boldsymbol{x}) &= 0 \qquad \forall \boldsymbol{x} \in \Omega, \\
g(\boldsymbol{x}) &= \begin{cases} -10(y+1) & \forall \boldsymbol{x} \in \{1\} \times [-1,1] \\ -10(x+1) & \forall \boldsymbol{x} \in [-1,1] \times \{1\} \\ 0 & \forall \boldsymbol{x} \in \{-1\} \times [0,1] \cup [0,1] \times \{-1\} \end{cases}, \\
D(\boldsymbol{x}) &= 0 \qquad \forall \boldsymbol{x} \in [-1,0] \times \{0\} \cup \{0\} \times [-1,0].
\end{aligned}
$$

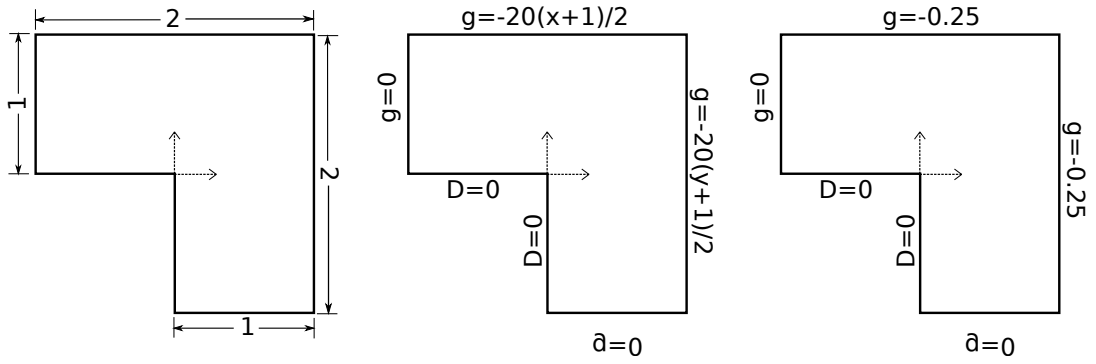
The quantity of interest is the average temperature on the external boundaries, $\omega = \{1\} \times [-1,1] \cup [-1,1] \times \{1\}$,

$$
s(u) = \frac{1}{|\omega|} E\left[\int_\omega u d\Gamma\right].
$$

In the domain, there are 75 circular particles of radius 0.05 and conductivity $k_p = 0.5$. The matrix has a conductivity of $k_m = 1$. Consequentially, the volume fraction is approximately $\nu = 0.196$. The centres of the particles are placed following a uniform distribution over the domain and not allowing them to intersect with other particles or the boundaries of the domain. An approximation to the solution of this problem is built by using the finite element method with 512 Monte Carlo realisations. The resulting histogram of the quantity of interest is shown in fig. 4.3.

To compute bounds, we approximate the probability of a point being inside a particle for any point of the domain by the volume fraction[1], hence $E[k] \approx k_p \nu + (1-\nu) k_m$. Our approximations will be compared with the a posteriori bounds from the section 4.3. The kinematically admissible approximations are obtained using $\bar{k} = E[k]$ as homogenised conductivity, while we use $\hat{k} = 1/E[k^{-1}]$ as the homogenised conductivity for the statically admissible solution schemes. The same mesh is used for all approximations and it is shown in fig. 4.4a.

[1] As the domain increases and the no. of particles increase accordingly while their size remain constant, the error becomes negligible.



(a) (b) (c)

Figure 4.2: (a) Geometrical description of the domain, (b) boundary conditions for the primal and (c) dual problem in section 4.4.1.

We note that in this case, the adjoint problem has a physical interpretation which is shown in fig. 4.2c.

The results are summarised in table 4.1 and figs. 4.4 and 4.5. We see that the bounds hold and that they are sharp.The ratio between error interval length and exact QoI oscillates between 0.08 and 0.16. The alternative measure, the ratio between error interval and the error in the QoI is not a good choice to evaluate the sharpness of the bounds, since the interval length of *eq.* (4.40) is independent of $\bar{k}$ whilst $s(e)$ depends on it. We also mention that in this specific example, the effect of neglecting the discretisation errors and using a KA field as an SA field does not have an important influence on the bounds as we can see on the results. In fact, the discretisation error can be computed by considering eq. (4.3) as the true model. By applying the bound in eq. (4.36), we obtain,

$$|s(\bar{u}) - s(\bar{u}^h)| \leq 0.016. \tag{4.70}$$

which is negligible when compared to the total error.

4.4.2 Effects of the volume fraction and material contrast

We seek in this numerical example to illustrate the effect of the volume fraction and contrast between the material properties in the error bounds. The domain considered for this study

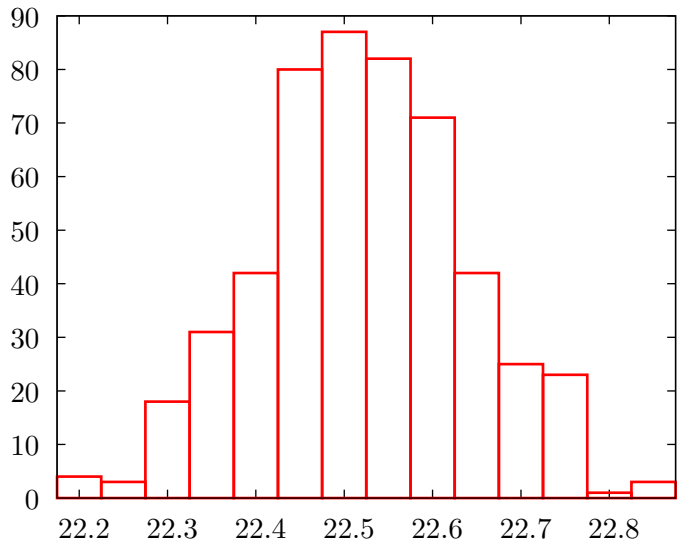
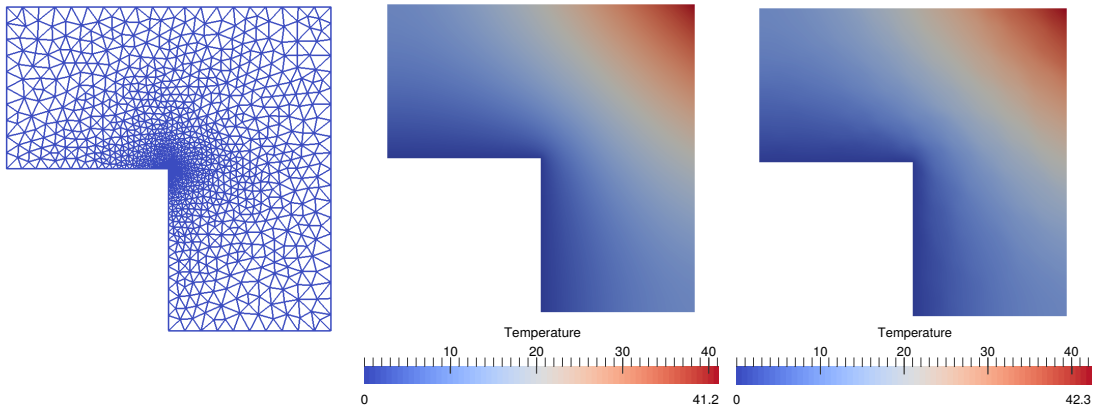


Figure 4.3: Histogram of the quantity of interest from section 4.4.1. Mean: 22.55 Std. deviation: 0.12.

$s(\bar{u}^h)$	ζ_l	$s(u) - s(\bar{u}^h)$	ζ_u	$\zeta_l + s(\bar{u}^h)$	$s(u)$	$\zeta_u + s(\bar{u}^h)$
21.92	-1.842	0.63	1.842	20.08	22.55	23.76
	-0.0483		1.794	21.87		23.71
	-1.822		1.822	20.1		23.74

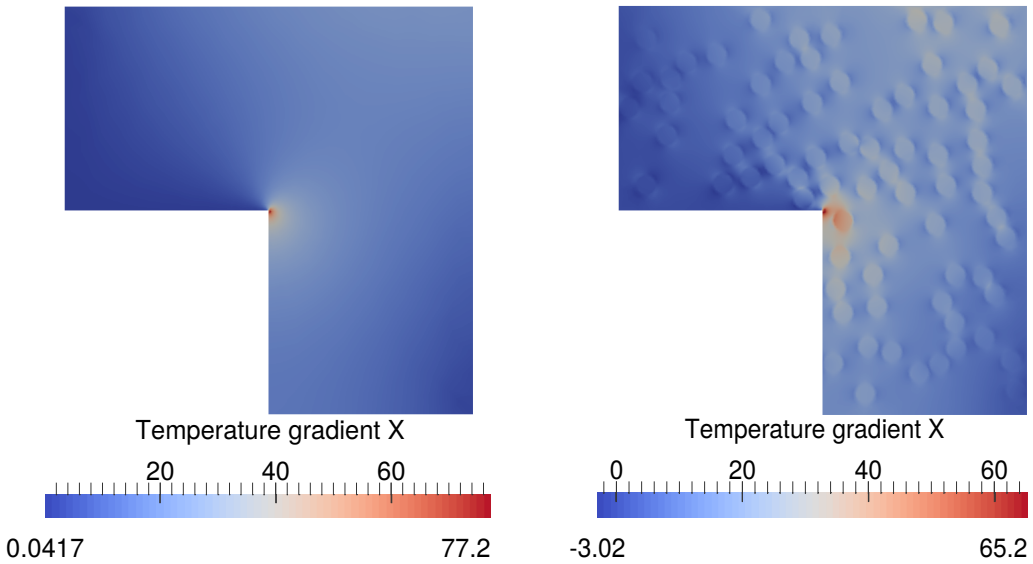
Table 4.1: Results from section 4.4.1. ζ_l and ζ_u represent both lower and upper bounds respectively. The first row shows the results for the error bound in eq. (4.36) and the second row for the error bound in eq. (4.40). In the third row, it was assumed that the discretisation error is negligible and a KA approximation was used as a SA field. The bound applied was the one in eq. (4.36).

is shown in fig. 4.6. The conductivity of the matrix is fixed to the value $k_m = 1$, while the conductivity of the inclusions, k_p will take the values 0.25, 0.5 and 0.75. These problems are solved for volume fractions ν equal to 0.1, 0.2 and 0.3. We assume that the probability of being inside a particle is a constant and coincides with the volume fraction. Under this



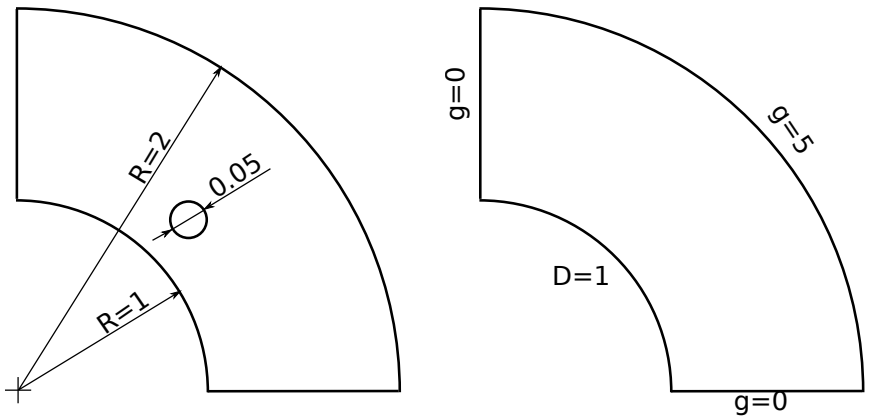
(a) (b) (c)

Figure 4.4: (a)Mesh used for the homogeneous problem. 2066 linear triangular elements (b) The temperature field of the compatible solution (c) The temperature field of one realisation (section 4.4.1).



(a) (b)

Figure 4.5: (a) The temperature gradient in X direction of the compatible solution (b) The temperature gradient in X direction of a realisation (section 4.4.1).

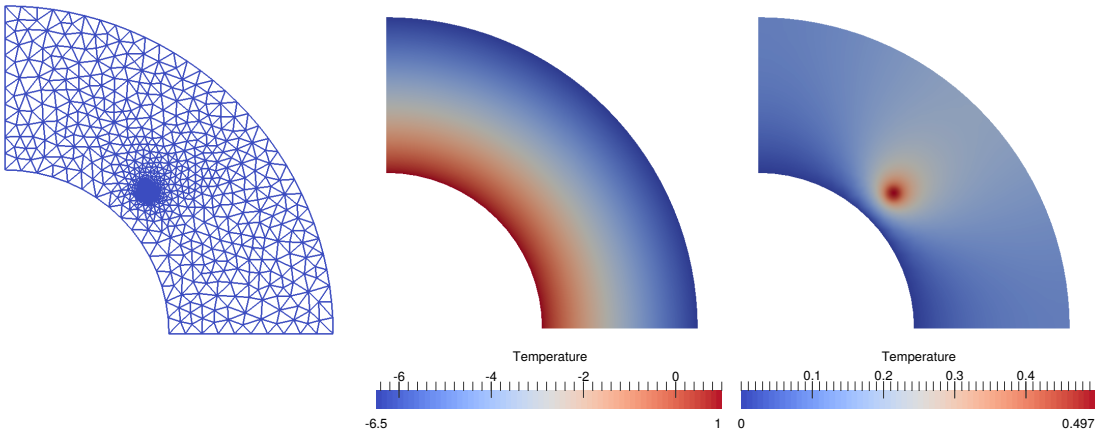


(a) (b)

Figure 4.6: Description of the problem from section 4.4.2 (a) Geometry of the domain. The coordinates of the centre of the circle are $(\sqrt{0.75}, \sqrt{0.75})$. The quantity of interest is the average temperature in this circle. (b) Boundary conditions.

assumption $E[k] = \nu k_p + (1-\nu)k_m$ and $E[k^{-1}] = \nu/k_p + (1-\nu)/k_m$. The results were restricted to the bounds in eq. (4.40). The KA approximations are obtained using rule of mixture $E[k]$, while the SA approximations are obtained using inverse rule of mixture $1/E[k^{-1}]$. In fig. 4.7, the mesh used in the analysis is shown together with the resulting primal and dual temperature fields. The primal homogenised solution is radial in its nature while the dual homogenised solution peaks in the region of interest due to the induced source term. The bounds are plotted in fig. 4.8 and their numerical values are summarised in table 4.2.

The results show that as we increase the material contrast and/or the volume fraction, the interval defined by the error bounds grows in length. This example highlights one of the limitations of this work, high material contrasts can increase drastically the interval defined by the bounds. In the appendix C, under certain assumptions, we present an analytical discussion regarding this behaviour.



(a) (b) (c)

Figure 4.7: Results of the problem from section 4.4.2 (a) Linear triangular mesh composed of 1412 elements used for the simulation (b) Temperature field of the primal problem for a conductivity of 0.25 and a volume fraction of 0.1 (c) Temperature field of the dual problem for a conductivity of 0.25 and a volume fraction of 0.1

4.4.3 Complex 3D example

We consider in this example a complex 3D domain. The methodology is applied to the cylinder head of an engine (fig. 4.9). At the bottom of the cylinder a temperature of $460K$ is prescribed. At the fins and the upper face it is assumed that a flux of $200W \cdot m^{-2}$ exits the body, while it is assumed that there is no heat exchange in the hole and lateral surfaces. The body is made of matrix enriched with particles. The thermal conductivity of the matrix is $460W/(m \cdot K)$, while the conductivity of the inclusions is $230W/(m \cdot K)$. The inclusions add up to 20% of the volume of the domain. Again, we assume that $E[k] = \nu k_p + (1-\nu)k_m$ and $E[k^{-1}] = \nu/k_p + (1-\nu)/k_m$. The quantity of interest is the average temperature on the upper face. The domain was discretised with roughly 1.5 million linear tetrahedrons. The KA approximations were obtained using rule of mixture $E[k]$, while the SA approximations were obtained using inverse rule of mixture $1/E[k^{-1}]$.

The resulting temperature field can be seen in fig. 4.10 while the bounds can be found in table 4.3. The bounds are sharp when compared to the temperatures present in the

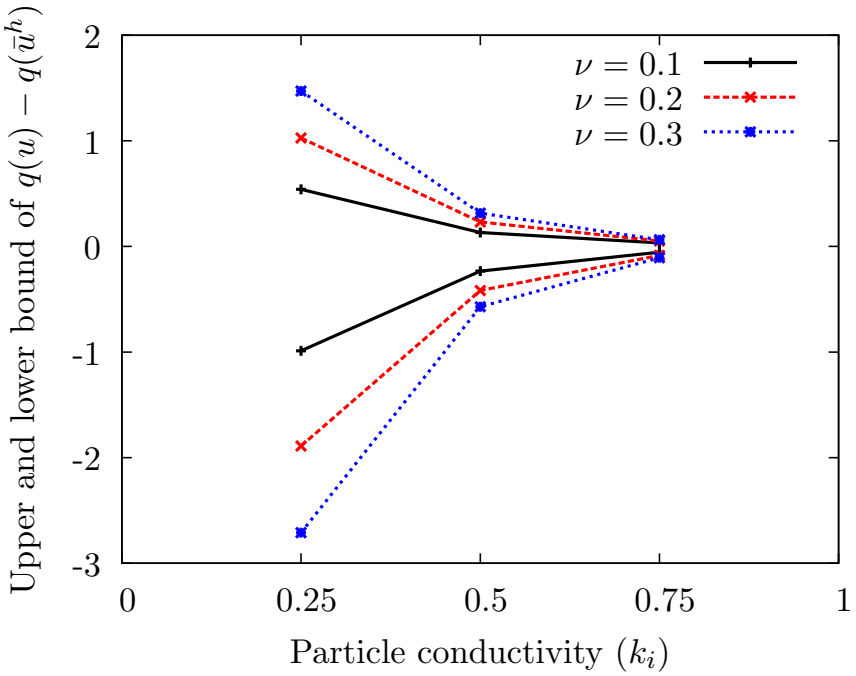


Figure 4.8: Effect of the contrast of conductivities in the lower and upper bounds. From section 4.4.2.

problem (≈ 450). Furthermore, the resulting interval length of the bound in eq. (4.40), 1.5, is half of the interval length resulting from eq. (4.36), 3. This relation between both bounds is not a coincidence and it is explained in appendix A. We also note that in this specific example, there is a large discrepancy between the bounds when the discretisation error is neglected, *i.e.* the KA approximation is used as an SA approximation.

4.4.4 Validation of the bound for the second moment

The purpose of this numerical example is to validate the bound for the second moment presented in section 4.3.3.3. We consider a unidimensional domain defined on the interval $[0, 8]$. In the domain, there are 4 particles of length of 0.5 with unknown position. Their conductivity is 0.5 and the conductivity of the rest of the domain is 1. At both ends of the domain, the temperature is set to 0. The source term is $f(x) = -1$. The quantity of interest is the average temperature in the interval $[0.45, 0.5]$.

ν	k_i	$s(\bar{u}^h)$	$\zeta_l + s(\bar{u}^h)$	$\zeta_u + s(\bar{u}^h)$	$\zeta_u - \zeta_l$
0.1	0.25	-1.18672	-2.17531	-0.646484	1.52882
0.1	0.5	-1.12935	-1.36253	-0.997478	0.365049
0.1	0.75	-1.07491	-1.1297	-1.04107	0.0886278
0.2	0.25	-1.3791	-3.26921	-0.351572	2.91764
0.2	0.5	-1.24729	-1.66483	-1.01533	0.6495
0.2	0.75	-1.12935	-1.21403	-1.07858	0.135452
0.3	0.25	-1.60872	-4.3196	-0.137407	4.1822
0.3	0.5	-1.3791	-1.95062	-1.06355	0.887065
0.3	0.75	-1.18672	-1.29409	-1.12329	0.170796

Table 4.2: Results from the section 4.4.2. ζ_u and ζ_l are the lower and upper bounds in eq. (4.40). The last column presents the interval length.

$s(\bar{u}^h)$	ζ_l	$s(u) - s(\bar{u}^h)$	ζ_u	$\zeta_l + s(\bar{u}^h)$	$s(u)$	$\zeta_u + s(\bar{u}^h)$
445.8	-1.506	?	1.506	444.3	?	447.3
	-1.503		0.002728	444.3		445.8
	-1.085		1.085	444.7		446.9

Table 4.3: Results from section 4.4.3. ζ_l and ζ_u represent both lower and upper bounds respectively. The first row corresponds to the error bound in eq. (4.36) and the second row corresponds to the error bound in eq. (4.40). In the third row, it was assumed that the discretisation error is negligible and a KA approximation was used as a SA field. The bound applied was the one in eq. (4.36).

The centres of the particles are placed following an uniform distribution in the domain. The particles are not allowed to intersect with each other, however, when they intersect with the boundary, the exceeding part of the particle is placed on the other end of the domain. This results in a probability of being inside of a particle of 0.25 for all points of

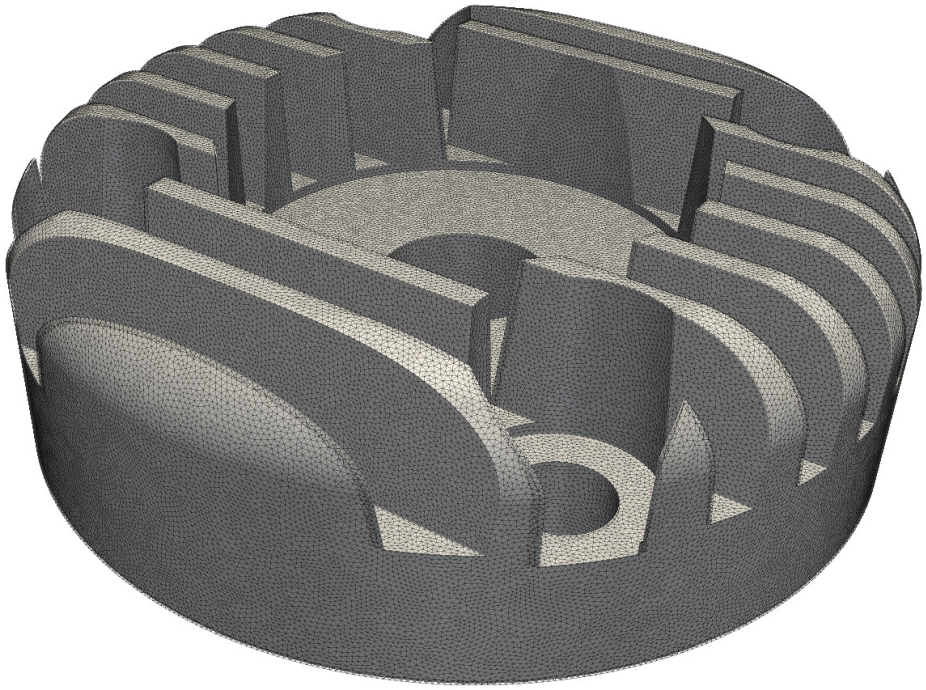


Figure 4.9: Cylinder head (section 4.4.3)

the domain and that coincides with the volume fraction.

The KA and SA surrogate models are built by setting $\bar{k} = E[k]$ and $\hat{k} = 1/E[k^{-1}]$. The surrogate models and several realisations are shown in fig. 4.11. Bounds for the expectation and the second moment of the quantity of interest are computed using eqs. (4.40) and (4.49).

To compute the bounds for the second moment, we need the fields $E[k(x)k(y)]$, $E[k^{-1}(x)k(y)]$ and $E[k^{-1}(x)k^{-1}(y)]$. Those functions are computed using the auxiliary probability functions $p_i(d)$, $p_o(d)$ and $p_d(d)$. Given two points separated by a distance d: $p_i(d)$ is the probability that both points are inside a particle, $p_o(d)$ is the probability of both points being in the matrix, and $p_d(d)$ is the probability of one point being inside a particle and the other being in the matrix. These functions were computed numerically by generating 50000 realisations of the domain. In fig. 4.12 their numerical approximation is shown. Once those functions are available, we can compute

$$E[k(x)k(y)] = k_i^2 p_i(|x-y|) + k_m^2 p_o(|x-y|) + k_i k_m p_d(|x-y|) \tag{4.71}$$

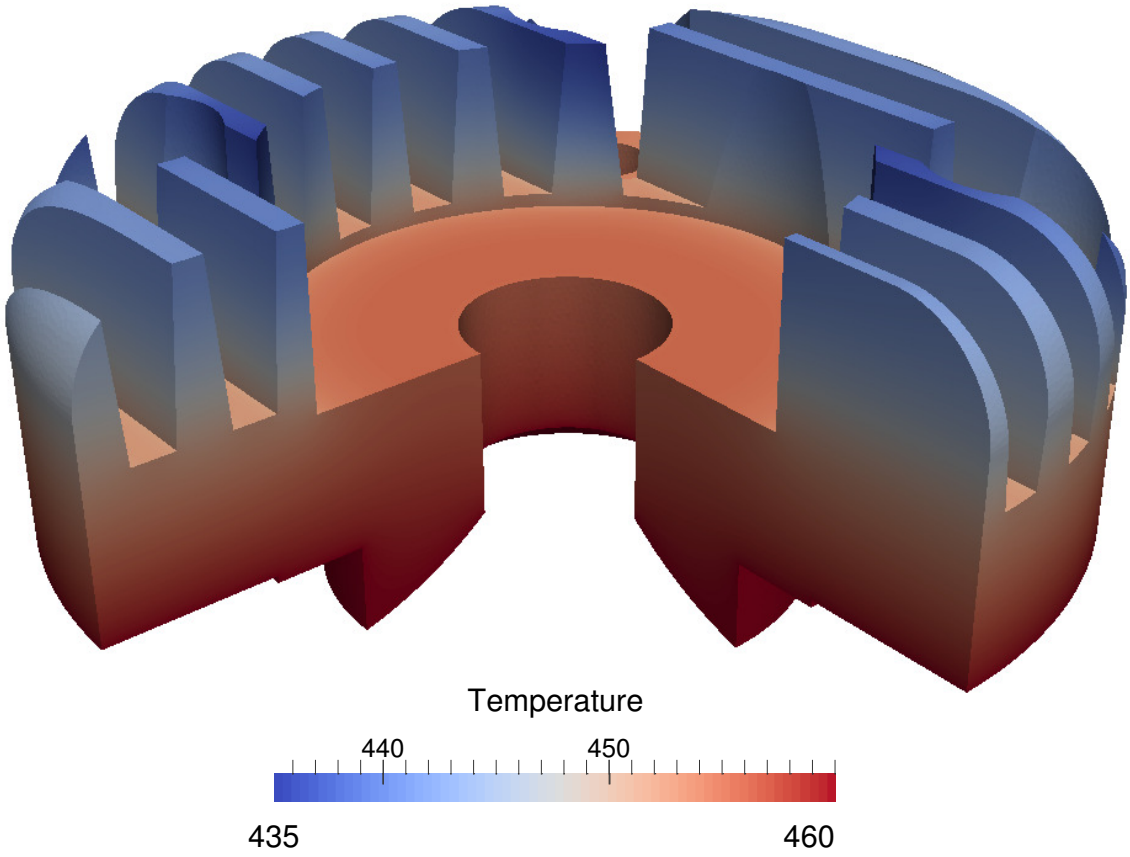
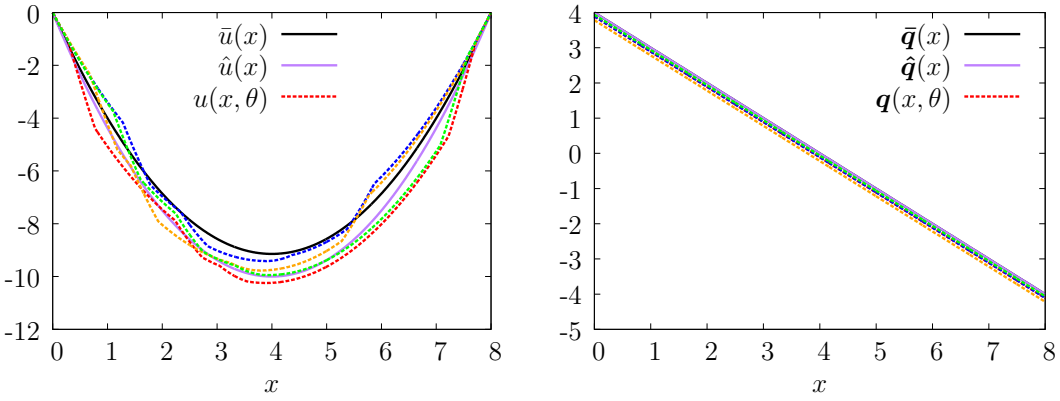


Figure 4.10: Temperature field of the cylinder head (section 4.4.3)

and similarly for $E[k^{-1}(x)k^{-1}(y)]$ and $E[k(x)k^{-1}(y)]$.

All the results are summarised in table 4.4. The bounds for the expectation hold and they were computed using eq. (4.40). The ratio between the interval length and $s(u)$ is 0.15. The bound for the second moment also holds but it is not very sharp. In this case the ratio between the bound and the exact second moment m_2 is 1.87. This is expected due to the inequalities used in the derivation of the bound that rely on the effectivity of



(a) (b)

Figure 4.11: (a) The temperature field of the homogenised solution and four realisations. Notice its independence of the homogenisation scheme (b) The flux field of the homogenised solution and of the exact solution.

the bound for the expectation. However, the bound for the variance has an even lower effectivity. The variance of the QoI is 0.34 while the upper bound is 5.81. The effectivity is 17.23. This difference in effectivity is explained by the fact that bound results from the difference of two bounds, *i.e.* the overestimation of each of the bounds is not cancelled, but added. Further work is required to improve the effectivity of the bound for the variance.

$s(u)$	$s(\bar{u}^h)$	ζ_l	ζ_u	$s(\bar{u}^h)+\zeta_l$	$s(\bar{u}^h)+\zeta_u$	m_2	$s(\bar{u}^h)^2$	ζ_2
-2.23	-2.04	-0.33	0.013	-2.37	-2.027	5.31	4.16	9.92

Table 4.4: Summary of the results of section 4.4.4. The exact quantity of interest was approximated by the solution of 2096 realisations of the problem. ζ_l and ζ_u are lower and upper bounds for $s(e)$ respectively. This implies that $s(u)$ must lie in the interval $[s(\bar{u}^h)+\zeta_l, s(\bar{u}^h)+\zeta_u]$. ζ_2 is the upper bound for m_2 (see section 4.3.3.3) .

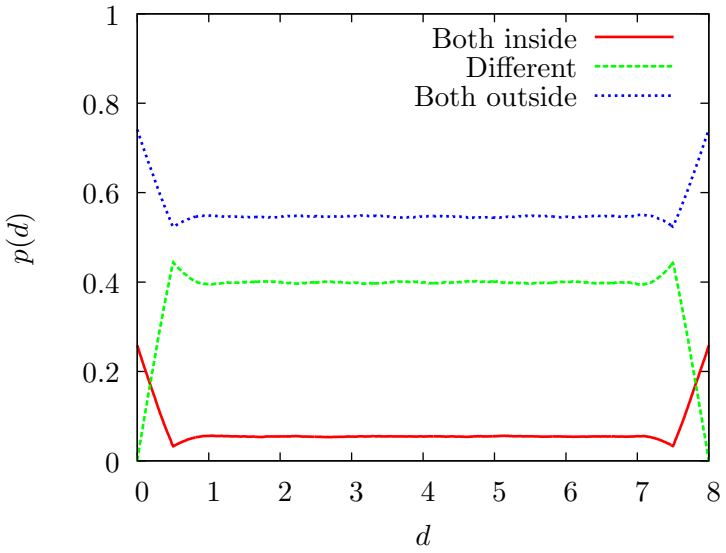


Figure 4.12: Probability of two points separated a distance x being both inside a particle (Both inside), both in the matrix (Both outside) and one being inside a particle and one inside the matrix (Different). Computed using 50000 realizations of the domain.

4.4.5 Adaptive modelling example

In this example, we apply the ideas presented in section 4.3.4. The domain is described in fig. 4.13. The domain is an aggregate of two constituents, round particles (radius 0.04, conductivity 5) surrounded by a matrix (conductivity 1). There are 740 particles and the centre of the particles behaves like a uniform random variable inside the domain. They are not allowed to intersect each other, but are allowed to intersect the boundaries. The average volume fraction is $\nu = 0.189$ and we assume that $E[k] = \nu k_p + (1-\nu)k_m$ and $E[k^{-1}] = \nu/k_p + (1-\nu)/k_m$. We set $\hat{k} = 1/E[k^{-1}]$, while we will set $\bar{k}$ to 11 equally spaced values in the interval $[1/E[k^{-1}], E[k]]$.

The domain is further subdivided in 8 squares which naturally align with the edges of the interior void. The corresponding error estimators are shown in table 4.5. We note that the difference between the estimators for different $\bar{k}$ is small. Subdomain 2, the area below the void, is selected for refinement. The bounds are computed using the technique described in section 4.3.4. Figure 4.15 remarks the importance of the choice of the conductivity in

Ω_i	β_{Ω_i}	Ω_i	β_{Ω_i}
1	0.4405	5	0.4973
2	0.6359	6	0.2852
3	0.3646	7	0.3872
4	0.5809	8	0.1568

Table 4.5: Local error indicators for the QoI ($\bar{k} = 1.47$). Subdomain 2 contributes most to the error in the QoI and it is selected for model refinement.

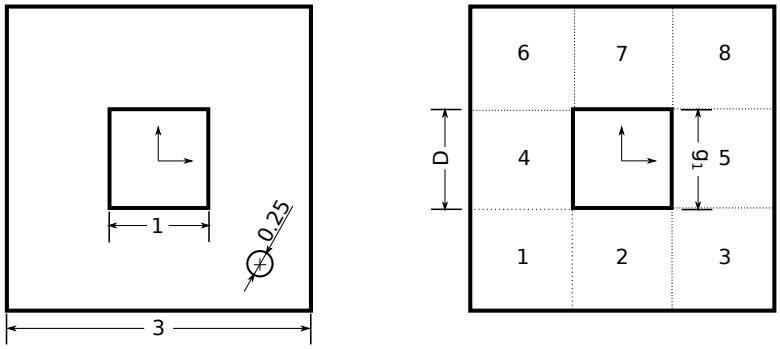
order to optimise the uncertainty gap. Without model refinement, the uncertainty gap is

$$-7.60118 \leq s(u) \leq -4.25279 \tag{4.72}$$

This is independent of $\bar{k}$. With model refinement, the smallest uncertainty gap for $s(u)$,

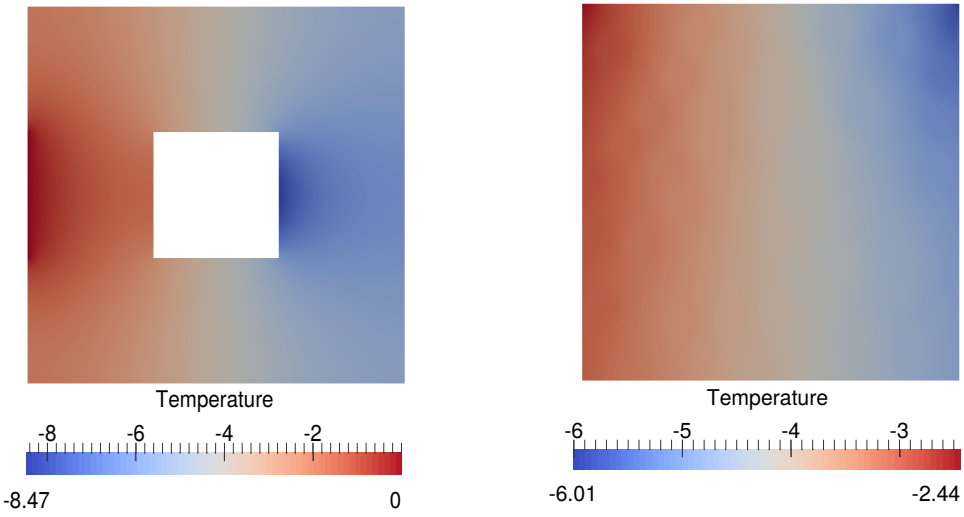
$$-7.26408 \leq s(u) \leq -4.48026 \tag{4.73}$$

is attained when $\bar{k} = 1.47$. The homogenised temperature field and a realisation of the window for this conductivity are shown in fig. 4.14. In this case, the interval length reduction is equal 0.56, value close to the error indicator in the subdomain, 0.64. This means that, as mentioned in remark 4, we have reduced the uncertainty in the QoI by an amount close to what we consider to be a "theoretical maximum". We also note that for small $\bar{k}$, the lower bound for the adaptive model is worst than the deterministic. An explanation for this deterioration was given in section 4.3.4. Finally, the upper bound for error in energy norm η ($\bar{k} = 1.47$), takes the value 4.938 before model refinement and 4.576 after model refinement. Using the bound mentioned in remark 3, the value is 4.669. The difference between both bounds is small, though, it could be magnified in other problems.



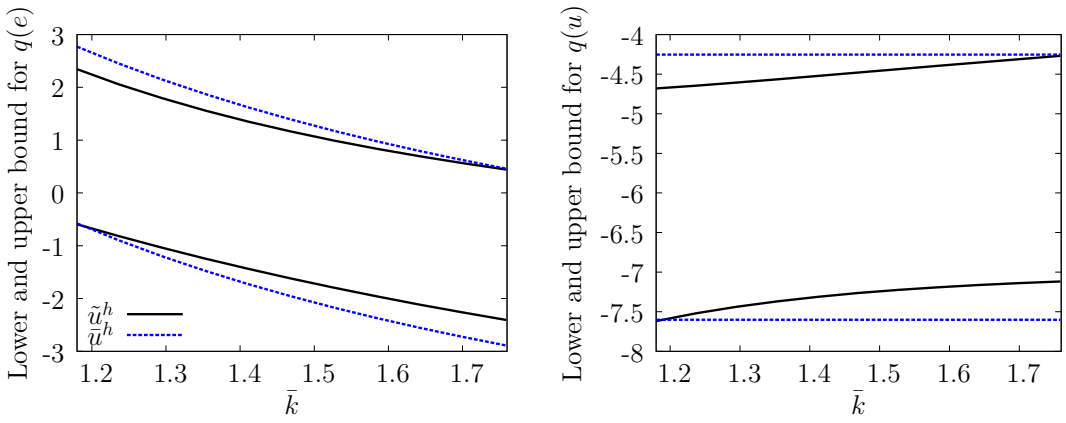
(a) (b)

Figure 4.13: Description of the problem from section 4.4.5 (a) Geometry of the domain. The coordinates of the centre of the circle are $(1, -1)$. The quantity of interest is the average temperature in this circle. (b) Boundary conditions and subdomain numbering. $D = 0$ and $g_1 = 40(0.5 + y)(0.5 - y)$ on the indicated regions. $g = 0$ in the rest of the boundary.



(a) (b)

Figure 4.14: (a) Temperature field of the surrogate model for $\bar{k} = 1.47$. (b) A realisation of the local approximation.



(a) (b)

Figure 4.15: Lower and upper bounds for the solution $\bar{u}^h$ to the purely macroscopic problem and for the solution $\tilde{u}^h$ for the adaptive model as a function of the surrogate conductivity $\bar{k}$. The graph on the left-hand side presents the bounds for the error in QoI, while the graph on the right-hand side presents the bound for the QoI itself.

4.5 Conclusion

In this chapter, we have presented a method to quantify the modelling error introduced by homogenisation. The approach introduces two novelties: the representation of the heterogeneous model through an stochastic PDE and the derivation of the bounds through the Prager-Synge hypercircle theorem. Those ideas combined allowed us to retain scale separation in the computation of error bounds. This combination also allowed us to characterise and optimise the effectivity of the bounds. It was shown that one of the bounds for quantities of interest is independent of the homogenised conductivity field. Furthermore, the theory was validated in the numerical examples.

The numerical examples and the theory (appendix C) have also shown the main limitation of this work, the sensitivity of the bounds to the contrast between the material properties of the different phases. This sensitivity is explained by the close relation between the derived bounds and the classical "Reuss-Voigt" bounds. A solution to this issue was proposed through model adaptivity, *i.e.* to solve locally the reference problem to construct

a better approximation to the exact solution. This approach has some disadvantages as well. It is sensitive to the homogenisation scheme used and the geometry makes it difficult to introduce the window where the reference model is solved. In the next chapter, we propose another solution which does not suffer of those disadvantages.

Lastly, taking into account the complete characterisation of the estimates given by the theory developed, their simple implementation (excluding model adaptivity) and their sharpness, we consider that the method is ready for use for contrasts in the range $[0.5, 2]$. For higher contrasts, we conclude that the method is not yet ready for use and we believe that a possible solution might be the derivation of new bounds that can incorporate more information of the conductivity field such as its covariance.

Chapter 5

Enriched multiscale method with guaranteed accuracy for the modelling of random heterogeneneous media

5.1 Introduction

In the previous chapter, we have seen that a high contrast between the material properties has a great impact on the error bounds, which makes them valueless in some instances. To alleviate this problem, error indicators were introduced to mark the regions that contribute the most to this uncertainty and the microstructure was fully resolved in those areas. The effectivity of this method was shown in section 4.4.5. Complex geometries configurations or the need for very narrow bounds may prevent the application of this approach. In this chapter, we look at a second approach that mitigates the geometrical constraints.

Owing to the fact that mesh refinement will only reduce the discretisation error, but won't affect the model error, the improvement of our approximation must include stochastic terms. The difficulty lies in the fact that a parametrisation of the random variables is not available since the microstructure is generated through an algorithm. Furthermore, the impenetrability condition between constituents poses severe difficulties in the application of the Karhunen–Loève decomposition to obtain a parametrisation [20].

In the present approach, we will rewrite the reference problem to allow for a wider class of shape functions and approach its solution directly, instead of solving a surrogate problem and assessing the quality of its solution. In this new form, the stiffness matrix reads,

$$\boldsymbol{K}_{ij} = a(\Phi_i, \Phi_j) = E\left[\int_\Omega k \nabla\Phi_i \cdot \nabla\Phi_j \, d\Omega\right]. \tag{5.1}$$

Furthermore, we rewrite our approximation as the sum of two terms, namely a deterministic and a stochastic term,

$$u(\boldsymbol{x}, \theta) \approx u^h(\boldsymbol{x}, \theta) = \bar{u}^h(\boldsymbol{x}) + \sum_{i=1}^{M} \tilde{u}_i f_i(\boldsymbol{x}, \theta). \tag{5.2}$$

with f_i to be described later. Since the expectation operator appears in the bilinear form, the computation of the best approximation does not require the pointwise value of f_i for every realisation θ. In fact, only the pointwise values of the expectations of f_i (and related terms such as its gradients) are needed. For instance, the product of a certain deterministic

shape function N_i and a stochastic function f_j takes the form

$$a(N_i, f_j) = E\left[\int_\Omega k\nabla N_i \cdot \nabla f_j \, d\Omega\right] = \int_\Omega \nabla N_i \cdot E\left[k\nabla f_j\right] \, d\Omega. \tag{5.3}$$

Hence, this demonstrates that this block of the matrix can be computed only with a partial description of f_i. In other words, we only need $E[k\nabla f_j]$ and not the value of f_j for every single realisation of the domain. Similar terms appear in the construction of the other blocks and the generalised force vector.

The function f_i still needs to be defined. We choose to define f_i as the temperature field of a RVE-type problem defined in a domain that encompasses the domain of interest multiplied by a classical FE hat function. Hence, the strategy followed in this chapter is to solve an RVE and not only extract the homogenised conductivity, but also the fields such as $E[k\nabla f_i]$ that appear in the construction of $\boldsymbol{K}$ and $\boldsymbol{f}$. With the appropriate filters, the fields can be transformed in constants which can be understood as characteristics of the microstructure.

Under certain assumptions, the bounds will be also guaranteed and sharper than the ones for the deterministic solution. Their application is easier for complex geometries and for a given material, the computation of material constants is the same, hence done in advance for any structure.

The chapter is organised as follows. Firstly, in section 5.2, the problem is reformulated. Then, in section 5.3 the approximation is described in detail: the expressions for the construction of the system of equations and for the error bounds are given. In section 5.4, the filters that are used to extract the quantities are described. In the following section, we express the theory in a form that is suitable for computer implementation. The theory is illustrated with numerical examples in section 5.6. The chapter is concluded with a discussion on the extension of the method for a statically admissible approximation.

5.2 Reference problem

We use the same notation introduced in the previous chapter, f, g and D are still deterministic functions that represent the source term, the prescribed fluxes and the prescribed

temperatures respectively. We also recall the following bilinear form and its induced (semi-)norm,

$$a(u,v) = E\left[\int_\Omega k(\boldsymbol{x},\theta)\,\nabla u(\boldsymbol{x},\theta)\cdot\nabla v(\boldsymbol{x},\theta)\,d\Omega\right]$$
$$\|v\| = \sqrt{a(v,v)}$$

we recall and also introduce the following spaces

$$\mathcal{U}^\theta = \{v \,|\, \forall\theta\in\Theta \quad v(\boldsymbol{x},\theta)\in\mathcal{U} \text{ and } \|v\|<\infty\}$$
$$\mathcal{U}_0^\theta = \{v \,|\, \forall\theta\in\Theta \quad v(\boldsymbol{x},\theta)\in\mathcal{U}_0 \text{ and } \|v\|<\infty\}$$

and we redefine the linear form

$$l(v) := E\left[\int_\Omega f(\boldsymbol{x})v(\boldsymbol{x},\theta)\,d\Omega - \int_{\Gamma_N} g(\boldsymbol{x})v(\boldsymbol{x},\theta)\,d\Gamma\right] \tag{5.4}$$

With this notation, the reference problem reads

For all $v(\boldsymbol{x},\theta)\in\mathcal{U}_0^\theta(\Omega)$, find $u(\boldsymbol{x},\theta)\in\mathcal{U}^\theta(\Omega)$ such that

$$a(u,v) = l(v) \tag{5.5}$$

We emphasize that if u is a solution of eq. (4.2) and $\|u\|$ is bounded, then u is also a solution of eq. (5.5). As mentioned in the introduction, the problem is reformulated in order to allow the use of shape functions for which only a partial description is available. More precisely, the solution of the problem can be approximated only with the knowledge of certain expectations of the shape functions, instead of requiring their pointwise value for each realisation.

5.3 Approximation

Our approximation takes the form

$$u(\boldsymbol{x},\theta) \approx u^h(\boldsymbol{x},\theta) = \bar{u}^h(\boldsymbol{x}) + \tilde{u}^h(\boldsymbol{x},\theta) = \sum_{i=1}^{N} N_i(\boldsymbol{x})u_i + \sum_{d=1}^{n_d}\sum_{i=1}^{N'} a_i^d\phi_i(x)\mathbb{U}^d(\boldsymbol{x},\theta) \tag{5.6}$$

where

- $\bar{u}^h$ is a linear combination of shape functions N_i that form a partition of unity. The degrees of freedom are introduced through the scalars u_i. The resulting function is constant in the stochastic domain.

- $\tilde{u}^h$ is a linear combination of shape functions ϕ_i that form a partition of unity, times a stochastic function $\mathbb{U}^d$. The discretisation associated with ϕ_i does not necessarily coincide with the discretisation of N_i. $\mathbb{U}^d$ is the solution, possibly filtered, of a RVE of the material that composes domain. We have defined one function per spatial direction, corresponding to the BVP with kinematically or statically uniform boundary conditions on that direction. Nonetheless, other boundary conditions could be chosen and more functions could be incorporated. a_i are scalars that represent the degrees of freedom.

The function $\mathbb{U}$ is introduced through a partition of unity in the same way as additional functions are introduced in extended/generalised finite element method [12, 60] and for this reason we may call $\mathbb{U}$ an enrichment function. The use of partition of unity in this particular case relates to the construction of kinematically admissible approximations. If the discretisation can capture exactly the boundary of the domain, the shape functions N_i and ϕ_j have the delta-Kronecker property and N_i can interpolate exactly $D(\boldsymbol{x})$, then by setting

$$u_i = D(\boldsymbol{x}_i)$$
$$a_j = 0$$

for each of the nodes over the Dirichlet boundary, we will obtain a KA approximation.

5.3.1 Construction of the system of equations

In this section, we examine the construction of the stiffness matrix and the generalised force vector with the aim of establishing the minimal information of the enrichment functions

required to obtain the best approximation of u in the $\|\cdot\|$ norm. The stiffness matrix is appropriately divided in four blocks,

$$\boldsymbol{K} = \begin{pmatrix} \boldsymbol{K}^{NN} & \boldsymbol{K}^{N\phi_1} & \boldsymbol{K}^{N\phi_2} & \dots & \boldsymbol{K}^{N\phi_{n_d}} \\ \boldsymbol{K}^{\phi_1 N} & \boldsymbol{K}^{\phi_1\phi_1} & & & \vdots \\ \boldsymbol{K}^{\phi_2 N} & & \ddots & & \vdots \\ \vdots & & & \ddots & \vdots \\ \boldsymbol{K}^{\phi_{n_d} N} & \dots & \dots & \dots & \boldsymbol{K}^{\phi_{n_d}\phi_{n_d}} \end{pmatrix} \tag{5.7}$$

with

$$\boldsymbol{K}^{NN}_{ij} = a(N_i, N_j) = \int_\Omega E[k]\nabla N_i \nabla N_j \, d\Omega \tag{5.8}$$

$$\boldsymbol{K}^{N\phi_d}_{ij} = a(N_i, \phi_j \mathbb{U}^d) = \int_\Omega E[k\mathbb{U}^d]\nabla N_i \nabla \phi_j \, d\Omega + \int_\Omega E[k\nabla\mathbb{U}^d]\nabla N_i \phi_j \, d\Omega \tag{5.9}$$

$$\boldsymbol{K}^{\phi_d N} = (\boldsymbol{K}^{N\phi_d})^T \tag{5.10}$$

$$\begin{aligned} \boldsymbol{K}^{\phi_d\phi_{d'}}_{ij} = a(\phi_i\mathbb{U}^d, \phi_j\mathbb{U}^{d'}) = &\int_\Omega E[k\mathbb{U}^d\mathbb{U}^{d'}]\nabla\phi_i\nabla\phi_j \, d\Omega + \int_\Omega E[k\nabla\mathbb{U}^d\nabla\mathbb{U}^{d'}]\phi_i\phi_j \, d\Omega + \\ &+ \int_\Omega E[k\nabla\mathbb{U}^d\mathbb{U}^{d'}]\phi_i\nabla\phi_j \, d\Omega + \int_\Omega E[k\mathbb{U}^d\nabla\mathbb{U}^{d'}]\nabla\phi_i\phi_j \, d\Omega \\ d, d' = 1...n_d. & \end{aligned}$$

On the other hand, the generalised force vector is divided in two blocks

$$\boldsymbol{f} = \begin{pmatrix} \boldsymbol{f}^N \\ \boldsymbol{f}^\phi_1 \\ \vdots \\ \boldsymbol{f}^{\phi_{n_d}} \end{pmatrix} \tag{5.11}$$

with

$$\boldsymbol{f}^N_i = l(N_i) = \int_\Omega f N_i \, d\Omega - \int_{\Gamma_N} g N_i \, d\Gamma \tag{5.12}$$

$$\boldsymbol{f}^{\phi_d}_i = l(\phi_i\mathbb{U}^d) = \int_\Omega E[\mathbb{U}^d] f \phi_i \, d\Omega - \int_{\Gamma_N} E[\mathbb{U}^d] g \phi_i \, d\Gamma \tag{5.13}$$

$$d = 1, ..., n_d. \tag{5.14}$$

We note that in the derivation of these expressions, we had to apply Fubini's theorem [21] (as in chapter 4) to change the order of application of the expectation and the domain integral, *i.e.*

$$E\left[\int_\Omega \cdot \, d\Omega\right] = \int_\Omega E\left[\cdot\right] \, d\Omega. \tag{5.15}$$

In conclusion, the u^h that best approximates u in the norm $\|\cdot\|$ can be obtained only with partial knowledge of $\mathbb{U}$. We will see that the same holds for the computation of the error bounds and we will exploit this property so the construction of the approximation remains tractable.

5.3.2 Computation of the error bounds

The derivation of the error bounds does not require any additional results, we can apply directly the Prager-Synge theorem derived in section 4.3. The result is applied to the pair $(u^h, \hat{\boldsymbol{q}}^h)$, u^h is kinematically admissible since we constrain the stochastic degrees of freedom to be 0 on the Dirichlet boundary (see section 5.3); $\hat{\boldsymbol{q}}^h$ is SA and deterministic, it is constructed as discussed in section 4.3.1 (the discussion of stochastic approximations of the flux field is delayed to section 5.7). Therefore, the constitutive relation error is an upper bound for the error in energy norm,

$$\|u - u^h\| \leq \|\hat{\boldsymbol{q}}^h + k\nabla u^h\|_{k^{-1}} \tag{5.16}$$

Again, we will show that only a partial description of $\mathbb{U}$ is required to compute the bounds. We start by expanding the CRE,

$$\begin{aligned}\|\hat{\boldsymbol{q}}^h + k\nabla u^h\|^2_{k^{-1}} = E\left[\int_\Omega k^{-1}\hat{\boldsymbol{q}}^h \cdot \hat{\boldsymbol{q}}^h \, d\Omega + \int_\Omega k\nabla u^h \cdot \nabla u^h \, d\Omega\right] \\ +2E\left[\int_\Omega \hat{\boldsymbol{q}}^h \cdot \nabla u^h \, d\Omega\right]\end{aligned}$$

Switching the order of the expectation and the domain integral,

$$\begin{aligned}\|\hat{\boldsymbol{q}}^h + k\nabla u^h\|^2_{k^{-1}} = \int_\Omega E\left[k^{-1}\right]\hat{\boldsymbol{q}}^h \cdot \hat{\boldsymbol{q}}^h \, d\Omega + \\ + \int_\Omega E\left[k\nabla u^h \cdot \nabla u^h\right] \, d\Omega + 2\int_\Omega \hat{\boldsymbol{q}}^h \cdot E\left[\nabla u^h\right] \, d\Omega.\end{aligned}$$

Now, we have to expand ∇u^h. We start by the third term,

$$\begin{aligned}\int_\Omega \hat{\boldsymbol{q}}^h \cdot E\left[\nabla u^h\right]\, d\Omega = &\int_\Omega \hat{\boldsymbol{q}}^h \cdot \nabla \bar{u}^h \, d\Omega \\ &+ \sum_{i=1}^{N}\sum_{d=1}^{n_d} \int_\Omega a_i E\left[\mathbb{U}^d\right] \hat{\boldsymbol{q}}^h \cdot \nabla \phi_i \, d\Omega \\ &+ \sum_{i=1}^{N}\sum_{d=1}^{n_d} \int_\Omega \phi_i a_i \hat{\boldsymbol{q}}^h \cdot E\left[\nabla \mathbb{U}^d\right]\, d\Omega\end{aligned}$$

The expansion of the second term results in

$$\int_\Omega E\left[k\nabla u^h \cdot \nabla u^h \, d\Omega\right] = \int_\Omega E\left[k\right] \nabla \bar{u}^h \cdot \nabla \bar{u}^h \, d\Omega + 2\int_\Omega E\left[k\nabla \tilde{u}^h\right] \cdot \nabla \bar{u}^h \, d\Omega +$$
$$\int_\Omega E\left[k\nabla \tilde{u}^h \cdot \nabla \tilde{u}^h\right]\, d\Omega = \int_\Omega E\left[k\mathbb{U}^d\mathbb{U}^{d'}\right] \nabla\phi_i \cdot \nabla\phi_j \, d\Omega + \int_\Omega E\left[k\mathbb{U}^d\nabla\mathbb{U}^{d'}\right] \cdot \nabla\phi_i\phi_j \, d\Omega +$$
$$\int_\Omega E\left[k\mathbb{U}^{d'}\nabla\mathbb{U}^{d}\right] \cdot \nabla\phi_i\phi_j \, d\Omega + \int_\Omega E\left[k\nabla\mathbb{U}^{d'} \cdot \nabla\mathbb{U}^{d}\right] \cdot \nabla\phi_i\phi_j \, d\Omega$$

Summarising the two last sections, the terms needed to characterise our approximation, in order to compute the best approximation and to estimate error, are

$E[k]$	$E[k^{-1}]$	$E[\mathbb{U}^d]$
$E[k\mathbb{U}^d]$	$E[\nabla\mathbb{U}^d]$	$E[k\nabla\mathbb{U}^d]$
$E[k\mathbb{U}^d\mathbb{U}^{d'}]$	$E[k\mathbb{U}^d\nabla\mathbb{U}^{d'}]$	$E[k\nabla\mathbb{U}^d \cdot \nabla\mathbb{U}^{d'}]$

for $d, d' = 1, ..., n_d$.

In section 5.5, those expressions are re-expressed in a form that is suitable for computer implementation.

5.4 Computation of the homogenised quantities

In this section, we seek to define the fields $\mathbb{U}^d$ used in our approximation. We begin by considering a finite RVE that contains the domain Ω, *i.e.* $\Omega \subseteq \Omega_{\text{RVE}}$ on which we require that every realisation θ of the RVE matches the realisation θ of the domain,

$$k_{\text{RVE}}(\boldsymbol{x}, \theta) = k_\Omega(\boldsymbol{x}, \theta) \quad \forall (\boldsymbol{x}, \theta) \in \Omega \times \Theta. \tag{5.17}$$

In other words, each realisation θ of the domain is cut from the RVE as in fig. 5.1.

Let U^d be the temperature field on the RVE for some boundary conditions that we will later specify. If we were to define $\mathbb{U}^d = U^d$, the bounds presented above are guaranteed. However, due to the dimensions of the RVE, the computation of U^d is as costly as approximating u directly. Nevertheless, if we apply a certain filter on U^d such that all the expectations required for the computation of bounds and error estimates ($E\left[k\mathbb{U}^d\right], E\left[k\nabla\mathbb{U}^d\right]$ ...) are independent of $\boldsymbol{x}$, *i.e.* constants, we could then consider a smaller RVE with a tractable dimension and extract those constants from it.

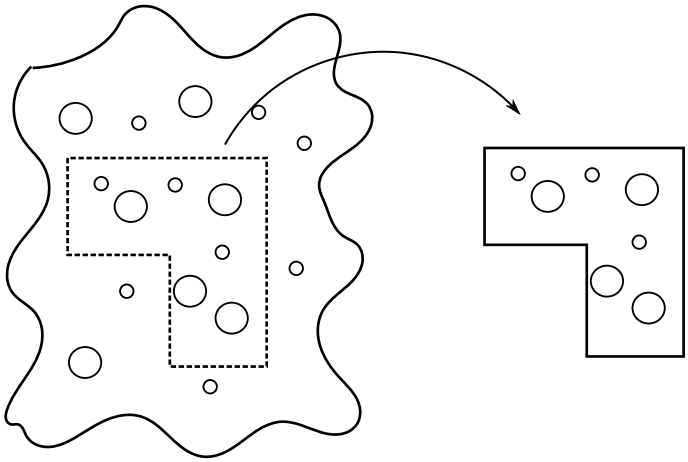


Figure 5.1: Each realisation of the domain (right) corresponds to a realisation of an RVE (left).

Our numerical experiments have shown that the fields $\mathbb{U}^d$ with the desired properties, can be obtained by defining

$$\mathbb{U}^d(\boldsymbol{x},\theta) = U^d(\boldsymbol{x},\theta) - (\chi_{\Omega_F} * U^d)(\boldsymbol{x},\theta), \tag{5.18}$$

where

$$\chi_{\Omega_F}(\boldsymbol{x}) = \begin{cases} 1 & \boldsymbol{x} \in \Omega_F \\ 0 & \text{otherwise} \end{cases} \qquad \text{(set indicator function)}$$

$$(u * v)(\boldsymbol{x}) = \int_\Omega u(\boldsymbol{x}-\boldsymbol{\tau})v(\boldsymbol{\tau})\,d\boldsymbol{\tau}, \qquad \text{(convolution operator)}$$

and combined with either kinematically uniform boundary conditions,

$$U^d(\boldsymbol{x},\theta) = \beta\boldsymbol{x}_d \quad \forall\boldsymbol{x}\in\Gamma, \quad \forall\theta\in\Theta, \quad \beta\in\mathbb{R}/\{0\}, \quad d = 1,...,n_d \tag{5.19}$$

or either statically uniform boundary conditions,

$$-k\nabla U^d(\boldsymbol{x},\theta)\cdot\boldsymbol{n}(\boldsymbol{x})=\lambda\boldsymbol{n}_d(\boldsymbol{x})\quad\forall\boldsymbol{x}\in\Gamma,\quad\forall\theta\in\Theta,\quad\lambda\in\mathbb{R}/\{0\},\quad d=1,...,n_d. \tag{5.20}$$

Basically, in eq. (5.18) a uniform filter is being applied to the field U^d meaning that from every point of the RVE, the local spatial average over Ω_F is subtracted. In $\mathbb{U}^d$ only the oscillating part of the function U^d remains and we expect smoother part of U^d to be captured by $\bar{u}^h$. We have chosen the set Ω_F to be a square big enough to contain a particle.

Regarding the boundary conditions, in eq. (5.19), we recall from chapter 2 that we are prescribing the average temperature gradient to have its norm equal to β and to have its direction equal to each of the main axes. The same applies for the statically uniform boundary conditions, the average flux is prescribed in eq. (5.20) to have its norm equal to λ and to have its direction equal to each of the main axes. We have seen in our numerical experiments no meaningful difference between the application of either boundary conditions.

In practice, in order to accelerate convergence, for each parameter P, realisations are generated until the convergence of the spatial average,

$$\frac{1}{|\Omega|}\int_\Omega P\,d\Omega \tag{5.21}$$

instead of achieving pointwise convergence. The ergodic property exhibited by $\mathbb{U}^d$ (equality between the spatial and ensemble average) allows us to proceed in the aforementioned manner. Furthermore, near the boundaries of the RVE, the behaviour of $\mathbb{U}^d$ will differ from the behaviour far from them. This "boundary layer" will disturb the results and for this reason, it is important to reduce the integration domain to $\Omega_r\subset\Omega$ in order to not distort the results (see fig. 5.2).

Remark 5. *In the early stage of our work, we define* $\mathbb{U}^d$ *by*

$$\mathbb{U}^d(\boldsymbol{x})=U^d(\boldsymbol{x})-\beta\boldsymbol{x}_d \tag{5.22}$$

with kinematic uniform boundary conditions,

$$U^d(\boldsymbol{x})=\beta\boldsymbol{x}_d\quad\forall\boldsymbol{x}\in\Gamma,\,m\in\mathbb{R}/\{0\}\,d=1,...,n_{sd} \tag{5.23}$$

In other words, we remove the average gradient from the solution. However, our numerical experiments have shown that quantities such as $E\left[k\nabla\mathbb{U}^d \cdot \nabla\mathbb{U}^d\right]$ *will depend on* $\boldsymbol{x}$*. Therefore, the application of the method, would require RVEs bigger than the domain, making it impractical.*

5.5 Implementation details

In this section, the theory is reformulated to be suitable for implementation. More specifically, it is going to be assumed that the same shape functions are used for $\bar{u}^h$ and $\tilde{u}^h$, *i.e.* $\phi_i = N_i$.

The solution can be understood as a vectorial field with $n_d + 1$ components. We will consider here the bidimensional case with two enrichment functions $\mathbb{U}^x$ and $\mathbb{U}^y$. The presentation can be trivially extended to the three dimensional case or to the case where more enrichment functions are used.

Our approximation reads

$$u^h(\boldsymbol{x},\theta) = \sum_{i=1}^{N}[u_i + a_i\mathbb{U}^x(\boldsymbol{x},\theta) + b_i\mathbb{U}^y(\boldsymbol{x},\theta)]N_i(\boldsymbol{x}) \tag{5.24}$$

Due to our approach, we can only compute expectations of the solutions, hence we express u^h product of a deterministic and a stochastic vectorial field

$$\begin{aligned} u^h(\boldsymbol{x},\theta) &= S(\boldsymbol{x},\theta)\cdot F(\boldsymbol{x}) \\ S(\boldsymbol{x},\theta) &:= (1 \quad \mathbb{U}^x(\boldsymbol{x},\theta) \quad \mathbb{U}^y(\boldsymbol{x},\theta)\,)^T \\ F(\boldsymbol{x}) &:= \sum_{i=1}^{N}(u_i \quad a_i \quad b_i)^T N_i(\boldsymbol{x}) \end{aligned}$$

The field F can be implemented using classes/routines already available to handle vectorial fields. On the other hand, storing certain expectations of S, we can easily compute the expected temperature, gradient, flux and energy of u^h. More specifically, we need to store the following matrices

$$E\left[S\right] = (1 \quad E\left[\mathbb{U}^x\right] \quad E\left[\mathbb{U}^y\right])^T$$

$$E\left[kS\right] = \left(E[k] \quad E\left[k\mathbb{U}^x\right] \quad E\left[k\mathbb{U}^y\right]\right)^T$$

$$E\left[k\nabla S\right] = \begin{pmatrix} 0 & 0 \\ E\left[k\frac{\partial \mathbb{U}^x}{\partial x}\right] & E\left[k\frac{\partial \mathbb{U}^x}{\partial y}\right] \\ E\left[k\frac{\partial \mathbb{U}^y}{\partial x}\right] & E\left[k\frac{\partial \mathbb{U}^y}{\partial y}\right] \end{pmatrix}$$

$$E\left[kSS^T\right] = \begin{pmatrix} 1 & E\left[\mathbb{U}^x\right] & E\left[\mathbb{U}^y\right] \\ E\left[\mathbb{U}^x\right] & E\left[(\mathbb{U}^x)^2\right] & E\left[\mathbb{U}^x\mathbb{U}^y\right] \\ E\left[\mathbb{U}^y\right] & E\left[\mathbb{U}^x\mathbb{U}^y\right] & E\left[(\mathbb{U}^y)^2\right] \end{pmatrix}$$

$$E\left[k\nabla S\nabla S^T\right] = \begin{pmatrix} 0 & 0 & 0 \\ 0 & E\left[k\frac{\partial \mathbb{U}^x}{\partial x}^2 + k\frac{\partial \mathbb{U}^x}{\partial y}^2\right] & E\left[k\frac{\partial \mathbb{U}^x}{\partial x}\frac{\partial \mathbb{U}^y}{\partial x} + k\frac{\partial \mathbb{U}^x}{\partial y}\frac{\partial \mathbb{U}^y}{\partial y}\right] \\ 0 & E\left[k\frac{\partial \mathbb{U}^x}{\partial x}\frac{\partial \mathbb{U}^y}{\partial x} + k\frac{\partial \mathbb{U}^x}{\partial y}\frac{\partial \mathbb{U}^y}{\partial y}\right] & E\left[k\frac{\partial \mathbb{U}^y}{\partial y}^2 + k\frac{\partial \mathbb{U}^y}{\partial x}^2\right] \end{pmatrix}$$

and the following order 3 tensor

$$E\left[kS \otimes \nabla S\right]_{ijk} =$$

$$\begin{pmatrix} 0 & 0 \\ E\left[k\frac{\partial \mathbb{U}^x}{\partial x}\right] & E\left[k\frac{\partial \mathbb{U}^x}{\partial y}\right] \\ E\left[k\frac{\partial \mathbb{U}^y}{\partial x}\right] & E\left[k\frac{\partial \mathbb{U}^y}{\partial y}\right] \end{pmatrix}_{1jk} \begin{pmatrix} 0 & 0 \\ E\left[k\mathbb{U}^x\frac{\partial \mathbb{U}^x}{\partial x}\right] & E\left[k\mathbb{U}^x\frac{\partial \mathbb{U}^x}{\partial y}\right] \\ E\left[k\mathbb{U}^x\frac{\partial \mathbb{U}^y}{\partial x}\right] & E\left[k\mathbb{U}^x\frac{\partial \mathbb{U}^y}{\partial y}\right] \end{pmatrix}_{2jk} \begin{pmatrix} 0 & 0 \\ E\left[k\mathbb{U}^y\frac{\partial \mathbb{U}^x}{\partial x}\right] & E\left[k\mathbb{U}^y\frac{\partial \mathbb{U}^x}{\partial y}\right] \\ E\left[k\mathbb{U}^y\frac{\partial \mathbb{U}^y}{\partial x}\right] & E\left[k\mathbb{U}^y\frac{\partial \mathbb{U}^y}{\partial y}\right] \end{pmatrix}_{3jk}$$

In the last expression, the operator $\otimes$ (tensor product) was introduced for convenience. We briefly introduce this operator and the colon product which will be soon required. The tensor product transforms two tensors A and B of order n and m and entries $A_{i_1 i_2 \ldots i_n}$ and $B_{j_1 j_2 \ldots j_m}$ in the tensor $A_{i_1 i_2 \ldots i_n} B_{j_1 j_2 \ldots j_m}$ of order $n+m$. Whilst the colon product is denoted by : and transforms the tensors A and B of order n and m with $n \geq m$ in the tensor

$$C_{i_1 i_2 \ldots i_{n-m}} = \sum_{i_{n-m+1}} \sum_{i_{n-m+2}} \ldots \sum_{i_n} A_{i_1 i_2 \ldots i_n} B_{i_{n-m+1} i_{n-m+2} \ldots i_n} \tag{5.25}$$

With the matrices and tensor introduced, the expression for the expected temperature reads

$$E\left[u^h\right](\boldsymbol{x}) = E\left[S\right] \cdot F(\boldsymbol{x}). \tag{5.26}$$

The expression for the expected gradient reads

$$E\left[\nabla u^h\right](\boldsymbol{x}) = E\left[\nabla S\right]^T F(\boldsymbol{x}) + \nabla F(\boldsymbol{x})^T E\left[S\right], \tag{5.27}$$

the expression for the expected flux reads

$$E\left[-k\nabla u^h\right](\boldsymbol{x}) = -E\left[k\nabla S\right]^T F(\boldsymbol{x}) - \nabla F(\boldsymbol{x})^T E\left[kS\right], \tag{5.28}$$

and the expression for the energy reads

$$\begin{aligned} E\left[k\nabla u^h \cdot \nabla u^h\right](\boldsymbol{x}) =& F(\boldsymbol{x})^T E\left[k\nabla S \nabla S^T\right] F(\boldsymbol{x}) + E\left[kSS^T\right] : \nabla F(\boldsymbol{x}) \nabla F^T(\boldsymbol{x}) \\ &+ 2E[kS \otimes \nabla S] : F(\boldsymbol{x}) \otimes \nabla F(\boldsymbol{x}) \end{aligned}$$

Using this representation the computation of the constitutive relation error is trivial,

$$\|\hat{\boldsymbol{q}}^h + k\nabla u^h\|^2_{k^{-1}} = \int_\Omega E\left[k^{-1}\right] \hat{\boldsymbol{q}}^h \cdot \hat{\boldsymbol{q}}^h \, d\Omega + \int_\Omega E\left[k\nabla u^h \cdot \nabla u^h\right] \, d\Omega + 2\int_\Omega \hat{\boldsymbol{q}}^h \cdot E\left[\nabla u^h\right] \, d\Omega \tag{5.29}$$

We can also use the expectations of S to compute the element stiffness matrix. We construct the element matrix by considering triplets of functions

$$SN_i = (1 \quad \mathbb{U}^x \quad \mathbb{U}^y)^T N_i \tag{5.30}$$

and the blocks associated to them

$$\boldsymbol{k}_{ij} := \begin{pmatrix} a(N_i, N_j) & a(N_i, N_j\mathbb{U}^x) & a(N_i, N_j\mathbb{U}^y) \\ a(N_i\mathbb{U}^x, N_j) & a(N_i\mathbb{U}^x, N_j\mathbb{U}^x) & a(N_i\mathbb{U}^x, N_j\mathbb{U}^y) \\ a(N_i\mathbb{U}^y, N_j) & a(N_i\mathbb{U}^y, N_j\mathbb{U}^x) & a(N_i\mathbb{U}^y, N_j\mathbb{U}^y) \end{pmatrix} \tag{5.31}$$

By algebraic manipulation, the block $\boldsymbol{k}_{ij}$ in terms of the matrices already defined reads

$$\begin{aligned} \boldsymbol{k}_{ij} = \int_{\Omega_e} E\left[k(\nabla(SN_i))^T \nabla(SN_j)\right] \, d\Omega = \\ \int_{\Omega_e} E\left[kSS^T\right] \nabla N_i \cdot \nabla N_j \, d\Omega + \int_{\Omega_e} E\left[k\nabla S \nabla S^T\right] N_i N_j \, d\Omega + \\ \int_{\Omega_e} E\left[kS \otimes \nabla S\right] : \nabla N_i N_j \, d\Omega + \int_{\Omega_e} (E\left[kS \otimes \nabla S\right] : \nabla N_j)^T N_i \, d\Omega \end{aligned}$$

Similarly, the block of the element vector associated with the triplet SN_i reads

$$\boldsymbol{f}_i = \begin{pmatrix} l(N_i) \\ l(N_i\mathbb{U}^x) \\ l(N_i\mathbb{U}^y) \end{pmatrix} = \int_{\Omega_e} f N_i E\left[S\right]\, d\Omega - \int_{\Gamma_e} g N_i E\left[S\right]\, d\Gamma \tag{5.32}$$

5.6 Numerical examples

5.6.1 Validation

In this first example, we validate the error bound. More accurately, we will check that the identity in eq. (4.13) holds, i.e.

$$E_{\text{Abs}} := \| - k\nabla u^h - \hat{\boldsymbol{q}}^h\|^2_{k^{-1}} - \|u - u^h\|^2 - \| - k\nabla u - \hat{\boldsymbol{q}}^h\|^2_{k^{-1}} = 0 \tag{5.33}$$

It is hard to construct a random particulate problem such that closed expression for u is available. Hence, u will be constructed from 100 realisations using very fine FE discretisations. More precisely, the discretisations have around 130000 linear triangular elements. Consequently, we will verify that

$$E_{\text{Rel}} := \left| \frac{\| - k\nabla u^h - \hat{\boldsymbol{q}}^h\|^2_{k^{-1}} - \|u - u^h\|^2 - \| - k\nabla u - \hat{\boldsymbol{q}}^h\|^2_{k^{-1}}}{\|u - u^h\|^2 + \| - k\nabla u - \hat{\boldsymbol{q}}^h\|^2_{k^{-1}}} \right| \leq \text{Tolerance.} \tag{5.34}$$

The domain is the square $[-1, 1] \times [-1, 1]$. We set $\Gamma_D = \Gamma$ and prescribe the temperature on it to be $D(x, y) = (x + 1)(y + 1)$. The domain contains 200 circular particles of radius 0.04 that are placed according to a uniform distribution and are not allowed to intersect with each other and with the boundary. The volume fraction is approximately 0.25. The conductivity of the matrix is 1 whilst the conductivity of the particles is 2.

We will exploit the fact that the identity holds even if u^h and $\hat{\boldsymbol{q}}^h$ are not the best approximations of u in their respective space. The identity only requires u^h to be KA and $\hat{\boldsymbol{q}}^h$ to be SA. We set the flux arbitrarily to

$$\hat{\boldsymbol{q}}^h(\boldsymbol{x}) = -1.5 \begin{pmatrix} y + 1 \\ x + 1 \end{pmatrix} \tag{5.35}$$

The temperature approximation is decomposed in two parts, $u^h = \bar{u}^h + \tilde{u}^h$ and we set,

$$\bar{u}^h = (x+1)(y+1) \quad \tilde{u}^h = 0.2(x+1)\mathbb{U}^x(\boldsymbol{x},\theta) + 0.2(y+1)\mathbb{U}^y(\boldsymbol{x},\theta) \tag{5.36}$$

where $\mathbb{U}^x(\boldsymbol{x},\theta)$ and $\mathbb{U}^y(\boldsymbol{x},\theta)$ are the solutions of realisation θ with kinematically uniform boundary conditions. In order to make this approximation KA, we set $\tilde{u}^h(\boldsymbol{x},\theta) = 0$ near the boundary by setting to zero the degrees of freedom of the shape functions that lie on the boundary.

The results are summarised in table 5.1. We see that E_{Abs} is greater than 0. This small discrepancy is due to the use of discrete approximation of u. We also note that the number of realisations does not affect the accuracy of the approximation, since the identity also holds for each realisation independently.

$\|u-u^h\|^2$	$\|-k\nabla u - \hat{\boldsymbol{q}}^h\|^2_{k^{-1}}$	η^2	E_{Abs}	E_{Rel}
147.56	1.65	149.21	$-1.40\cdot 10^{-4}$	$9.35\cdot 10^{-7}$

Table 5.1: Results of section 5.6.1.

5.6.2 Parameter extraction

In this example, we detail the extraction of the parameters that describe a microstructure. We consider that our domain is made of a microstructure of circular particles of radius 0.08 and conductivity 2 embedded in matrix of conductivity 1. The volume fraction is 0.25.

Realisations of the microstructure are generated in the domain $[-1,1]\times[-1,1]$. For each realisation and relating to the enrichments, we define two problems with the following kinematically uniform boundary conditions

$$L^x(\boldsymbol{x},\theta) = x \quad \forall \boldsymbol{x}\in\Gamma \quad \forall\theta\in\Theta \tag{5.37}$$

and

$$L^y(\boldsymbol{x},\theta) = y \quad \forall \boldsymbol{x}\in\Gamma \quad \forall\theta\in\Theta \tag{5.38}$$

Approximations of L^x and L^y are obtained using the finite element method. These fields are then filtered using the uniform filter to obtain $\mathbb{U}^x$ and $\mathbb{U}^y$. We set $\Omega_F = [-0.15, 0.15] \times [-0.15, 0.15]$.

Realisations are generated until convergence of the expectation of all quantities $k, k\mathbb{U}^x ...$ To decide whether a quantity X has converged, the spatial average is computed for each realisation generated so far,

$$X_i = \frac{1}{|\Omega_r|} \int_{\Omega_r} X \, d\Omega \tag{5.39}$$

in the domain $\Omega_r := [-0.6, 0.6] \times [-0.6, 0.6]$. As already explained, the domain is reduced to prevent the boundary layer to disturb the results (see fig. 5.2). The averages X_i are combined into a vector $\boldsymbol{X}$ and the mean μ_X and standard deviation σ_X of the vector are computed. If

$$\left| \frac{3\sigma_X}{\mu_X \sqrt{n}} \right| \leq \text{Tolerance} \tag{5.40}$$

we consider that $E[X]$ converges to μ_X. If

$$\left| \frac{\mu}{C} \right| \leq \text{Tolerance} \tag{5.41}$$

where C is another quantity that has already converged, we consider that $E[X]$ converges to 0. If neither of those two conditions hold, more realisations are needed. Briefly, the first criterion assumes that the sample average, μ_X, behaves like a normal variable with mean μ_X and standard deviation $\sigma_X/\sqrt{n}$. Under this reasonable assumption due to the central limit theorem, the probability of committing an error greater than tolerance in the estimation of μ_X is roughly 0.03%. In this particular example, we set the tolerance to 0.005.

The rationale for the second criterion is that if it holds and even if $E[X]$ does not converge to zero, it is too small to affect the construction of the approximation and computation of the error bounds. We set the tolerance for this criterion to 10^{-4}.

In table 5.2, the expectation of the quantities are summarised. A total of 392 realisations were needed. We note, in particular, that $E[k\mathbb{U}^d\mathbb{U}^d]$ which is non negative, converges to 0 too. This is due to the criterion eq. (5.41) and decreasing the tolerance will result in a value different than 0. However, the same was not observed for the remaining quantities

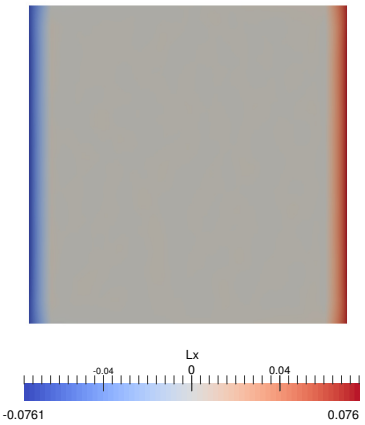


Figure 5.2: Field $E[\mathbb{U}^x]$. The spatial average must be computed on a subdomain otherwise the field near the boundary may distort the estimation of the quantity.

that converge to 0. This is not specific of this example and the same was observed for different volume fractions, different contrasts and for elliptical particles.

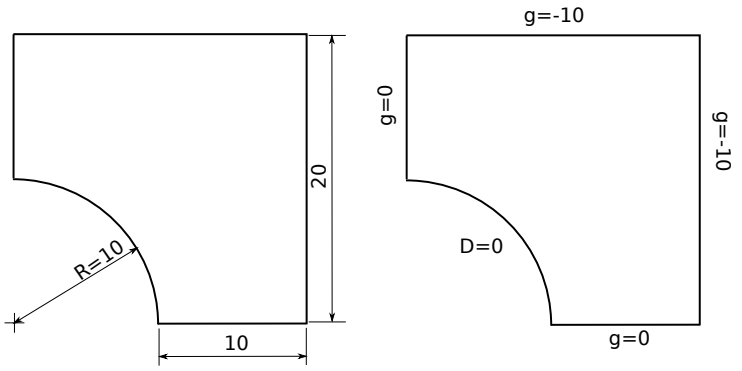
$E[k]$	$E[k\nabla\mathbb{U}^x]$	$E[k\nabla\mathbb{U}^x \cdot \mathbb{U}^x]$	$E[k\nabla\mathbb{U}^y]$	$E[k\nabla\mathbb{U}^y \cdot \mathbb{U}^y]$
1.28	$(-0.055,\ 0)$	0.049	$(0,\ -0.055)$	0.049

Table 5.2: Parameters of the example in section 5.6.2. All the quantities that converge to zero are not shown.

5.6.3 Enriched approximation

In this numerical example, we build a stochastic approximation as described in this chapter and compute the error bound. In addition, we compare it to the deterministic approximation presented in chapter 4. The domain considered and the boundary conditions are described in fig. 5.3.

The matrix has conductivity 1 and the particles have conductivity 2. The particles are circular and the particle radius is 0.08. The resulting volume fraction is approximately 0.26.

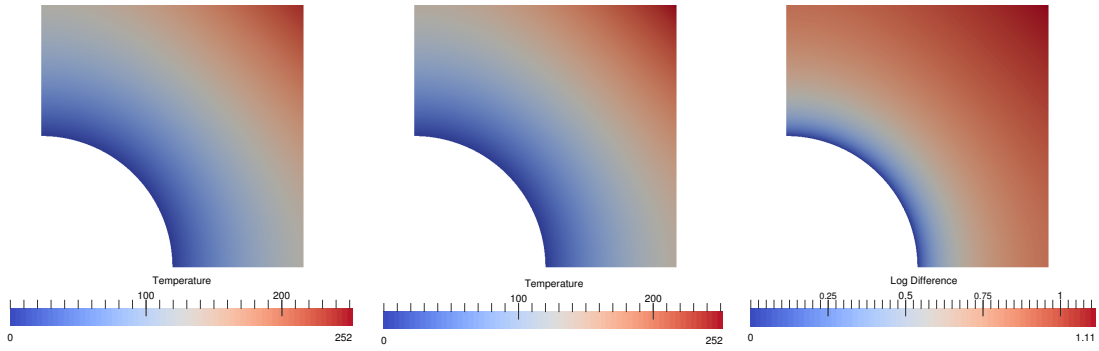


(a) (b)

Figure 5.3: (a) Geometrical description of the domain and (b) boundary conditions for the problem in section 5.6.3.

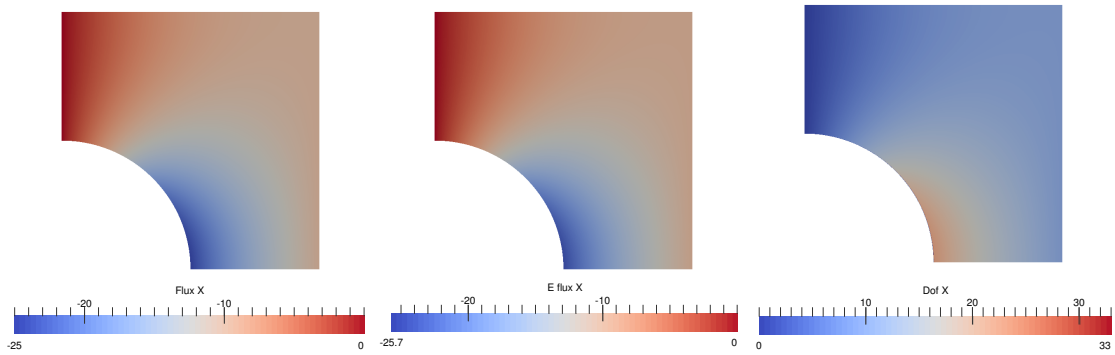
We construct an approximation u^h as described in section 5.3. The characterisation of $\mathbb{U}^x$ and $\mathbb{U}^y$ is obtained from the solution of RVEs until convergence. The flux field $\hat{\boldsymbol{q}}^h$ is approximated using the hybrid finite element method as described in chapter 3 using $\hat{k} = 1/E[k^-1]$. We also construct for the purpose of comparison a deterministic KA approximation using $\bar{k} = E[k]$ and which we denote by $\bar{u}^h_C$.

In fig. 5.4, we see $E\left[u^h\right]$, $\bar{u}^h_C$ and the logarithm of the difference of both fields. Both fields are very similar and their difference can only be seen in the logarithm plot. Though we note that the maximum temperature in $\bar{u}^h_C$ is 240 whilst the maximum temperature in $E\left[u^h\right]$ is 252. This is due the fact that rule of mixture usually overestimates the conductivity of the domain. In fig. 5.5, we plot the degrees of freedom a^x_i that multiply the enrichment function $\mathbb{U}^x$ along the x-component of the flux fields $\hat{\boldsymbol{q}}^h$ and $E[-k\nabla u^h]$. We see that the regions where the enrichment function is active coincide with the areas where the flux is highest in the x-direction. This is a result of the construction of the enrichment function as the solution of a problem with a mean flux on the x-direction. The symmetric nature of the problem allows us to make the same observations regarding $\mathbb{U}^y$ and the y-component of the flux.



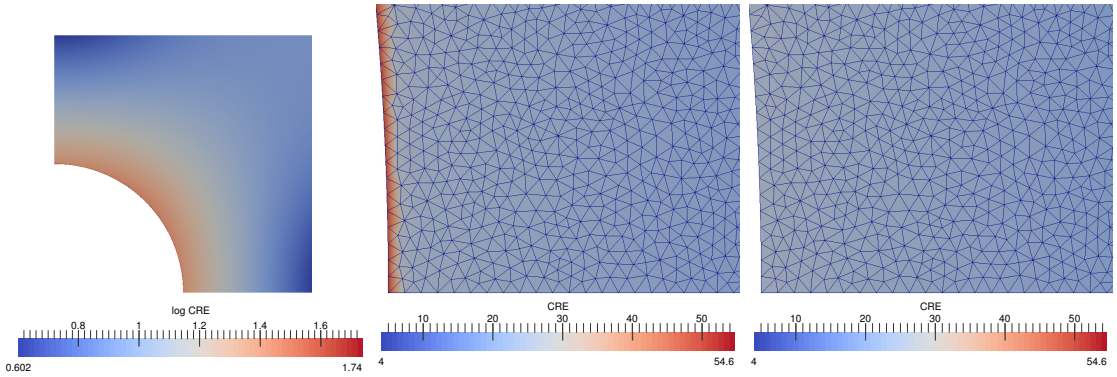
(a) (b) (c)

Figure 5.4: (a) Temperature field without enrichment $\bar{u}_C^h$, (b) expectation of the enriched field $E\left[u^h\right]$ and (c) logarithm of the difference $\log(1 + E\left[u^h\right] - \bar{u}_C^h)$ for $R = 0.08$.



(a) (b) (c)

Figure 5.5: (a) x-component of $\hat{\boldsymbol{q}}^h$ (b) x-component of $E\left[-k\nabla u^h\right]$ (c) Degrees of freedom of the enrichment function $\mathbb{U}^x$, *i.e.* a_i^x.



(a) (b) (c)

Figure 5.6: (a) Logarithm of the integrand of the CRE (b) and (c) Integrand of the CRE amplified in part of the Dirichlet boundary for strict and weak enforcement of Dirichlet boundary conditions.

In fig. 5.6, we plot the integrand of the CRE,

$$E\left[k^{-1}(\hat{\boldsymbol{q}}^h + k\nabla u^h)^2\right](\boldsymbol{x}) \tag{5.42}$$

This function behaves as an error indicator as it exposes the regions of the domain that contribute most to the error bound. We note that the error is concentrated around the Dirichlet boundary and that it is specially high in the first layer of elements. This is expected since the approximation u^h is constrained in that region. The high error layer disappears if we enforce the less restrictive constraint $E\left[u^h\right](\boldsymbol{x}) = 0$ instead of $u^h(\boldsymbol{x}, \theta) = 0$ on the Dirichlet boundary. This is shown fig. 5.6.c. However, the error bounds do not necessarily hold for this approximation.

The weaker constraint $E\left[u^h\right](\boldsymbol{x}) = 0 \quad \forall \boldsymbol{x} \in \Gamma_D$ was implemented by constraining $\bar{u}^h(\boldsymbol{x}) = 0$ on the Dirichlet boundary and not enforcing any constraint on $\tilde{u}^h$. Indeed, owing to the fact that $E\left[\mathbb{U}^x\right] = E\left[\mathbb{U}^y\right] = 0$ due to the filtering, it follows that

$$E[u^h](\boldsymbol{x}) = \bar{u}^h(\boldsymbol{x}) \tag{5.43}$$

The error bounds are summarised in table 5.3. The results show that the enrichment greatly reduces the error in energy norm. The square of the relative reduction is also

reported since is relevant for the estimation of error for QoIs[1]. Finally, we also report the non-guaranteed upper bound for the approximation built by weakly enforcing the Dirichlet boundary conditions. The difference is negligible since the additional error is only present in a single layer of elements.

η	η'	η_C	$\frac{\eta_C^2-\eta^2}{\eta_C^2}$
59.37	59.33	89.33	0.49

Table 5.3: η, η' are the upper bounds for the error in energy for the approximations u^h with strict and weak enforcement of the Dirichlet boundary conditions whilst η_C is the upper bound for $\bar{u}_C^h$.

We conclude this section by repeating the same problem whilst varying the size of the particles (0.16, 0.32, 0.64) and keeping the volume fraction constant. The results are shown in table 5.4 and fig. 5.7. We notice that as the particle size increases the enrichment is less effective due to the probably increased variance of the solution. Through the constants, the fact that scales are less separated is incorporated.

Radius	0.08	0.16	0.32	0.64
$\frac{\eta_C^2-\eta^2}{\eta_C^2}$	0.49	0.47	0.46	0.38

Table 5.4: Effect of the particle size in the upper bound.

5.6.4 Error contribution

In this example, we seek to estimate the error of the flux approximation

$$\| - k\nabla u - \hat{\boldsymbol{q}}^h\|_{k^{-1}}^2 = \|\mathbf{e}^q\|_{k^{-1}}^2 \tag{5.44}$$

[1]The sharpness of the error estimates is proportional to $\eta\eta_\phi$(see appendix A) and for this reason, we consider the impact of the enrichment on η^2.

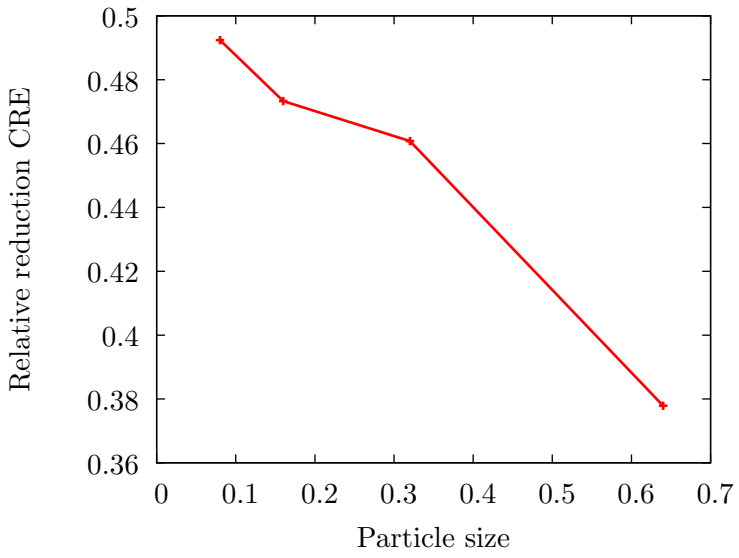


Figure 5.7: Effect of the particle size in the upper bound.

and the error of both temperature approximations

$$\|u - u^h\|^2 \quad \text{and} \quad \|u - \bar{u}^h\|^2. \tag{5.45}$$

and determine their relative contributions to the error bound.

The domain, the boundary conditions and material properties are the same as the problem in section 5.6.1. We compute the deterministic approximations $\hat{\boldsymbol{q}}^h$ and $\bar{u}^h$ using the techniques described in chapter 4 whilst the enriched stochastic approximation u^h is built as described in this chapter. An approximation of the field u is obtained through a combination of FEM and Monte Carlo. A total of $N = 200$ realisations of the domain are generated and solved using very fine discretisations. The flux error is estimated using Monte Carlo integration, *i.e.*

$$\|\mathbf{e}^q\|^2_{k^{-1}} \approx \frac{1}{N} \sum_{i=1}^{N} \| - k\nabla u_{\theta_i} - \hat{\boldsymbol{q}}^h\|^2_{k^{-1}}. \tag{5.46}$$

The error on the temperature field is estimated from the difference between the CRE and

the norm of the flux error, *i.e.*

$$\|u - u^h\|^2 = \|\hat{\boldsymbol{q}}^h + k\nabla u^h\|^2_{k-1} - \|\mathbf{e}^q\|^2_{k^{-1}},$$
$$\|u - \bar{u}^h\|^2 = \|\hat{\boldsymbol{q}}^h + k\nabla \bar{u}^h\|^2_{k-1} - \|\mathbf{e}^q\|^2_{k^{-1}}$$

We proceed in this manner for practical reasons. This approach only involves computing the difference of a deterministic field $\hat{\boldsymbol{q}}^h$ and a stochastic field u, whilst the other involves the comparison of two stochastic fields u and u^h.

The results are summarised in table 5.5. We see that in the deterministic case, the error introduced by the temperature approximation dominates over the flux error. In contrast, when the temperature field is enriched with RVE approximations, the opposite holds. Furthermore, we see that the error was greatly reduced, nonetheless, due to the reduced effectivity of the estimates, this is not reflected on the bounds. This accentuates the need of a stochastic approximation of the flux field as well if we want to quantify the error. In the next session, we describe a first attempt to solve this matter.

$\|\mathbf{e}^q\|^2_{k-1}$	$\|u - \bar{u}^h\|^2$	$\bar{\eta}^2$	$\frac{\|\mathbf{e}^q\|^2_{k-1}}{\bar{\eta}^2}$	$\frac{\|u-\bar{u}^h\|^2}{\bar{\eta}^2}$
0.44	0.75	1.19	0.37	0.63
$\|\mathbf{e}^q\|^2_{k-1}$	$\|u - u^h\|^2$	η^2	$\frac{\|\mathbf{e}^q\|^2_{k-1}}{\eta^2}$	$\frac{\|u-u^h\|^2}{\eta^2}$
0.44	0.17	0.61	0.72	0.28

Table 5.5: Results of section 5.6.4. $\bar{\eta}$ is the CRE for the deterministic approximations whilst η denotes the CRE for the enriched approximation.

5.7 Flux approximation

The construction of an enriched flux approximation is very challenging. The authors proposed solution suffers from locking, *i.e.* the enriched degrees of freedom are null and only deterministic part of the solution is active. Still, we describe in this section the approach attempted. We restrict our presentation to the bidimensional case, since some of the arguments are not immediately extended to the tridimensional case.

We also reformulate the flux problem to allow for a partial description of the shape functions. The problem reads:

For all $\boldsymbol{p}(\boldsymbol{x},\theta) \in \mathcal{S}_0^\theta(\Omega)$, find $\boldsymbol{q}(\boldsymbol{x},\theta) \in \mathcal{S}^\theta(\Omega)$ such that

$$E\left[\int_\Omega k^{-1}(\boldsymbol{x},\theta)\,\boldsymbol{q}(\boldsymbol{x},\theta)\cdot\boldsymbol{p}(\boldsymbol{x})\,d\Omega\right] = -E\left[\int_{\Gamma_D}\boldsymbol{p}(\boldsymbol{x},\theta)\cdot\boldsymbol{n}(\boldsymbol{x})\,D(\boldsymbol{x})\,d\Gamma\right] \tag{5.47}$$

where,

$$\mathcal{S}_0^\theta := \{\boldsymbol{q}\,|\,\forall\theta\in\Theta \quad \boldsymbol{q}(\boldsymbol{x},\theta)\in\mathcal{S}_0 \text{ and } \|\boldsymbol{q}\|_{k^{-1}}<\infty\}$$
$$\mathcal{S}^\theta := \{\boldsymbol{q}\,|\,\forall\theta\in\Theta \quad \boldsymbol{q}(\boldsymbol{x},\theta)\in\mathcal{S} \text{ and } \|\boldsymbol{q}\|_{k^{-1}}<\infty\}$$

Provided the solution of eq. (4.10) is bounded in the norm $\|\cdot\|_{k^{-1}}$, both formulations result in problems with identical solution.

Our approximation reads

$$\boldsymbol{q}(\boldsymbol{x},\theta)\approx\boldsymbol{q}^h(\boldsymbol{x},\theta)=\hat{\boldsymbol{q}}^h(\boldsymbol{x})+\underbrace{\sum_i\alpha_i^x\boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^x(\boldsymbol{x},\theta)+\sum_i\alpha_i^y\boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^y(\boldsymbol{x},\theta)}_{=:\tilde{\boldsymbol{q}}(\boldsymbol{x},\theta)} \tag{5.48}$$

where $\hat{\boldsymbol{q}}^h$ is the deterministic part of approximation as described in chapter 3, a_i^x, a_i^y are coefficients to be determined, $\boldsymbol{H}_i$ are 2×2 matrices whose entries are polynomials defined over the entire domain and $\mathbb{Q}^x$, $\mathbb{Q}^y$ are solutions of RVEs. More specifically, we have considered kinematically and statically uniform boundary conditions and the same filters described in section 5.4 were used.

We now focus our attention on the constraints on the matrices $\boldsymbol{H}$. We will require that our approximation fulfils

$$\forall\boldsymbol{x}\in\Omega \quad \forall\theta\in\Theta \begin{cases} \nabla\cdot\bar{\boldsymbol{q}}^h(\boldsymbol{x}) & = f(\boldsymbol{x}) \\ \nabla\cdot\tilde{\boldsymbol{q}}(\boldsymbol{x},\theta) & = 0 \end{cases}$$

irrespectively of the values of the degrees of freedom. The construction of $\bar{\boldsymbol{q}}^h$ follows the same procedure described in chapter 3. The satisfaction of the second equation implies

that the $\nabla \cdot \boldsymbol{H}\mathbb{Q}^d = 0$,

$$\nabla \cdot [\boldsymbol{H}(\boldsymbol{x})\mathbb{Q}^d(\boldsymbol{x},\theta)] = \boldsymbol{H}(\boldsymbol{x}) : [\nabla \mathbb{Q}^d(\boldsymbol{x},\theta)]^T + \nabla \cdot \boldsymbol{H}(\boldsymbol{x}) \cdot \mathbb{Q}^d(\boldsymbol{x},\theta) = 0 \tag{5.49}$$

where the divergence of a matrix means

$$\nabla \cdot (H_{ij})_j = \left(\sum_{i=1}^{n} \frac{\partial a_{ij}}{\partial x_i} \right)_j \tag{5.50}$$

In other words, the divergence of a matrix is a vector whose rows are the divergence of each of the columns.

Each of the terms in eq. (5.49) must be cancelled. We start with the first term on the right hand side. By observing that $\mathbb{Q}^d_{2,1} = \mathbb{Q}^d_{1,2}$ ($\nabla\mathbb{Q}^d$ is symmetric since it is the gradient of a field[2]) and $\mathbb{Q}^d_{1,1} = -\mathbb{Q}^d_{2,2}$ (since $\mathbb{Q}^d$ is divergence free), we obtain that

$$\boldsymbol{H}(\boldsymbol{x}) : [\nabla \mathbb{Q}^d(\boldsymbol{x},\theta)]^T = (H_{11} - H_{22})\mathbb{Q}^d_{1,1} + (H_{12} + H_{21})\mathbb{Q}^d_{1,2} = 0 \tag{5.51}$$

Furthermore, the equality holds for all realisations, hence

$$H_{11} - H_{22} = 0 \tag{5.52}$$

$$H_{12} + H_{21} = 0 \tag{5.53}$$

Denoting H_{11} by a and H_{12} by b, the matrices must have the form

$$\boldsymbol{H} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \tag{5.54}$$

The cancellation of the second term in eq. (5.49) for every realisation implies that $\nabla \cdot \boldsymbol{H} = 0$. Expanding this equation, we attain

$$\begin{aligned} \frac{\partial a}{\partial x} &= \frac{\partial b}{\partial y} \\ \frac{\partial a}{\partial y} &= -\frac{\partial b}{\partial x} \end{aligned} \tag{5.55}$$

If we define $f(x,y) = a(x,y) + b(x,y)i$ (i imaginary unit), the equations above are the Cauchy-Riemann equations of the complex function f [4]. A function f is holomorphic

[2] under the additional assumption that the conductivity is a constant inside each phase

(*i.e.* differentiable in the complex sense) in Ω if and only f satisfies the Cauchy-Riemann equations in Ω. This suggests that we can build pairs of polynomials (a, b) that fulfil the Cauchy-Riemann equations by taking respectively the real and imaginary part of an holomorph complex polynomial. A basis of the pairs polynomials up to degree n that fulfil eq. (5.55) can be built from $1, i, z, zi, z^2, z^2i, ..., z^n, z^ni$. For example, the basis for $n = 2$ is

$p(z)$	$a(x,y)$	$b(x,y)$
1	1	0
i	0	1
z	x	y
zi	$-y$	x
z^2	$x^2 - y^2$	$2xy$
z^2i	$-2xy$	$x^2 - y^2$

Since $\boldsymbol{H}$ is continuous, $\boldsymbol{H}\mathbb{Q}$ is also continuous and continuity along the element edges does not need to be enforced through Lagrange multipliers as in chapter 3.

The matrix system of equations has now the form

$$\begin{pmatrix} \boldsymbol{F} & \boldsymbol{D}^T & \boldsymbol{S}^T \\ \boldsymbol{D} & \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{S} & \boldsymbol{0} & \boldsymbol{P} \end{pmatrix} \begin{pmatrix} \boldsymbol{\alpha} \\ \boldsymbol{\beta} \\ \boldsymbol{\gamma} \end{pmatrix} = \begin{pmatrix} \boldsymbol{u} \\ \boldsymbol{v} \\ \boldsymbol{t} \end{pmatrix} \tag{5.56}$$

where the matrices $\boldsymbol{F}$ and $\boldsymbol{D}$, and the vectors $\boldsymbol{u}$ and $\boldsymbol{v}$ have the same form as in chapter 3 with $\hat{k} = 1/E[k^{-1}]$,the matrices $\boldsymbol{P}$ and $\boldsymbol{S}$ are defined as follows,

$$P = \begin{pmatrix} \boldsymbol{P^{xx}} & \boldsymbol{P^{xy}} \\ \boldsymbol{P^{yx}} & \boldsymbol{P^{yy}} \end{pmatrix}$$

$$P_{ij}^{dd'} = b(\boldsymbol{H}_i\mathbb{Q}^d, \boldsymbol{H}_j\mathbb{Q}^{d'}) = E\left[\int_\Omega k^{-1}\boldsymbol{H}_i\mathbb{Q}^d \cdot \boldsymbol{H}_j\mathbb{Q}^{d'}\, d\Omega\right] = \int_\Omega \boldsymbol{H}_i\boldsymbol{H}_j^T E[k^{-1}\mathbb{Q}^d{\mathbb{Q}^{d'}}^T]\, d\Omega$$

$$S = \begin{pmatrix} \boldsymbol{S}_x^1 & \boldsymbol{S}_x^2 \dots \boldsymbol{S}_x^{n_{\text{elm}}} \\ \boldsymbol{S}_y^1 & \boldsymbol{S}_y^2 \dots \boldsymbol{S}_y^{n_{\text{elm}}} \end{pmatrix}$$

$$(S_d^e)_{ik} = b_{\Omega_e}(\boldsymbol{H}_i\mathbb{Q}^d, \boldsymbol{p}_k) = E\left[\int_{\Omega_e} k^{-1}\boldsymbol{H}_i\mathbb{Q}^d \cdot \boldsymbol{p}_k\, d\Omega\right] = \int_{\Omega_e} \boldsymbol{H}_i E\left[k^{-1}\mathbb{Q}^d\right] \cdot \boldsymbol{p}_k\, d\Omega$$

$$i, j = 1, ..., n_{\text{mat}} \quad k = 1, ..., n_{\text{pol}} \quad e = 1, ..., n_{\text{elm}}$$

$\{\boldsymbol{p}_i\}_{i=1}^{n_{\text{pol}}}$ denotes the basis of divergence free polynomials used to construct the deterministic part. The vector $\boldsymbol{t}$ on the right hand side reads

$$\begin{aligned}
\boldsymbol{v} &= \begin{pmatrix} \boldsymbol{v}^x \\ \boldsymbol{v}^y \end{pmatrix} \\
v_i^d = p(\boldsymbol{H}_i \mathbb{Q}^d) &= -E\left[\int_{\Gamma_D} D\boldsymbol{H}_i \mathbb{Q}^d \, d\Gamma\right] + E\left[\int_{\Omega} \boldsymbol{H}_i \mathbb{Q}^d \cdot \bar{\boldsymbol{q}}^h_{\text{NH}} \, d\Gamma \, d\Omega\right] \\
&= -\int_{\Gamma_D} D\boldsymbol{H}_i E[\mathbb{Q}^d] \, d\Gamma + \int_{\Omega} \boldsymbol{H}_i E[\mathbb{Q}^d] \cdot \bar{\boldsymbol{q}}^h_{\text{NH}} \, d\Gamma \, d\Omega \\
i &= 1, ..., n_{\text{mat}}
\end{aligned}$$

$\bar{\boldsymbol{q}}^h_{\text{NH}}$ is the deterministic flux field that fulfils the conservation equation. Nonetheless, as already mentioned in the beginning of this section, the approximation $\boldsymbol{q}^h$ locks, $\boldsymbol{\gamma} = 0$, and the deterministic approximation is recovered. This means that the stochastic part of the solution is overconstrained. Further research is required for the construction SA stochastic field.

The reader might wonder why it was not chosen to proceed analogously to the SA deterministic setting where the polynomials were defined elementwise and continuity of the normals along the element edges was enforced. It is shown in appendix E that this seemingly more general approach where independent polynomials are defined inside each element and normal continuity along the element edges is enforced *a posteriori*, reduces this approach to the one presented in this section.

5.8 Conclusion

In this chapter, a strategy to model intractable heterogeneous problems was presented. The novel idea is to enrich the approximation with information of the microstructure. This information is incorporated through functions that represent the exact solution of representative volume elements of the microstructure. The computation of the "best approximation" and the estimation of error are tractable operations due to a suitable rewriting of the problem that only require a partial description of the enrichment functions *i.e.* only the expectation is required instead of their response for each realisation.

With this approach, two of the drawbacks of model adaptivity (described in section 4.3.4), the difficulty of application for complex geometries and sensitivity to the homogenisation scheme, were overcome.

Several numerical examples complemented the theory and have shown a great reduction of the actual error and the estimates. The examples have also revealed that due to the dual nature of the error bounds, further improvement of them will require the construction of an enriched flux approximation. That is, being the error estimates the sum of two errors, the contribution of the flux error dominates due its deterministic nature. However, in section 5.7 it was shown that the construction of an enriched flux approximation is not a trivial task. The proposed approach locks and a deterministic flux approximation is obtained.

Future work will aim to solve this matter. A possible strategy is the use of a different error bound which poses less constraints on the flux approximation. Repin in [55] presents an error estimate which generalises the Prager-Synge bound and allows the flux to be nonconforming.

We also seek to extend this work to cover other quantities such as the variance and the other higher order moments allowing us to not only characterise the average behaviour of the structure but also how particular layouts may deviate from it. Finally, we also plan to explore the use of different boundary conditions in the definition of the enrichment functions. This would allow for even further reduction of the error.

Chapter 6

Conclusion

The main purpose of the work presented in this thesis was the development of numerical methods to model heterogeneous structures. It was a requirement that those numerical methods incorporate means to quantify the error and to have a low computational cost. Two approaches were proposed. Both approaches share the reference/heterogeneous model and the error estimates whilst they differ in the way that the reference problem is approximated.

Expanding on their common characteristics, the reference model was chosen to be stochastic where each realisation represented a different particle layout of the heterogeneous structure. The error estimates were constructed to measure error in terms of the average response of the structure. This particular combination of model and measure allowed us to decouple the two scales of the problem and achieve the requirements of low computational cost while still being able to gauge the error. The error bounds presented are fully computable, guaranteed and can be applied to global (error in energy norm) and local quantities (error in quantities of interest).

Regarding the differences between both approaches, in chapter 4, the approximation is built by means of a surrogate problem. The surrogate problem results from the homogenisation of the heterogeneous/reference model. The effectivity of the error bounds were characterised and optimised. We have seen that in this approach that the error bounds are sensitive to the contrast of the material properties between different phases. To mitigate this effect, error indicators were introduced to indicate the regions of the domain where the reference model must be solved locally. In addition to those contributions, an error bound for the second moment of quantities of interest was presented.

In chapter 5, the solution of the reference model is approximated directly. Still, the computational cost remains under control since both scales are decoupled. The construction of the approximation was decomposed in two steps. In a first step, boundary value problems are solved on RVEs of the microstructure with the aim of estimating certain quantities/constants that characterise the average of behaviour of the microstructure. In the second step, those constants are introduced in a modified weak form to obtain an approximation of the solution. These constants are the only means by which both scales communicate.

The numerical examples have also identified an area for improvement of the second approach. By construction, the error bounds are the sum of errors of two independent approximations, an approximation of the temperature field and an approximation of the flux field. It was shown that flux error contributes most to the error since it incorporates less information of the microscale. Moreover, section 5.7 showed that the construction of the flux approximation enriched with information of the microscale must fulfil very restrictive constraints which pose additional difficulties.

Future extensions of this work will seek to solve this issue. A possible solution is to rederive the error estimates. Repin in [55] generalised the Prager-Synge theorem to allow for nonconforming fluxes. The use of this bound would pose fewer constraints on the enriched flux field.

The same issue could be addressed by dropping the requirement of obtaining guaranteed error bounds. We could aim to obtain approximate estimates of the error, by using fluxes fields whose average is statically admissible (SA) instead of requiring each realisation to be SA. This weaker condition may prevent the occurrence of locking.

In the following, we discuss other possible extensions of the work presented in this thesis.

- We believe that randomness is an essential attribute of composites and for this reason we should aim to characterise how the variations of the microstructure affect the response. This characterisation can be carried out by estimating the variance and other higher order moments of the response. We contemplate two strategies to estimate those quantities. Firstly, by improving the bound for the second moment and the variance developed in chapter 4. This could be done by adapting the more recent work on bounds for the effective properties [92] where tighter bounds were derived by using two and three point correlation functions. Secondly, we could approach this problem by extending the approach developed in chapter 5. Ideally, by further characterising the microstructure through the extraction of additional constants, would enable us to estimate higher order moments in the same way that we estimate the average response.

- In [17, 47, 96], the constitutive relation error was extended to provide guaranteed error bounds to problems of plasticity, viscoelasticity and nonlinear quantities. In future works, we could investigate the applicability of these generalisations to the development of macroscopic bounds for microstructures with more complex constitutive relations.

- In future work, we could focus on the obtention of confidence intervals as in [49] instead of error bounds. While the latter provides guarantees of the average response, confidence intervals describe the behaviour of a fraction of the realisations. A possible approach with low computational cost is to assume that the quantity of interest behaves like a normal variable. Under this assumption, the mean and the variance of this random variable could be approximated by the bounds for the expectation and the variance. Nonetheless, unless further developments are made for the bounds for the second moment developed in chapter 4, we expect the resulting confidence intervals to be very pessimistic.

Appendix A

Optimal value of α

We optimise the bound given in eq. (4.40) in terms of α. We seek α that minimises $f(\alpha) := \eta_{\text{Upp}}^{\text{QoI}} - \eta_{\text{Low}}^{\text{QoI}}$, the interval length. We later show that the optimal value also minimises the upper bound, $\eta_{\text{Upp}}^{\text{QoI}}$, and maximises the lower bound, $\eta_{\text{Low}}^{\text{QoI}}$.

We proceed with the minimisation of f. Through algebraic manipulations, we obtain

$$\begin{aligned} f(\alpha) &= \frac{1}{4}\left[(\eta_U^+)^2 - (\eta_L^+)^2 + (\eta_U^-)^2 - (\eta_L^-)^2\right] \\ (\eta_U^\pm)^2 &= \|\hat{\boldsymbol{q}} + k\nabla\bar{u}^h\|_{k^{-1}}^2\alpha^2 + \|\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h\|_{k^{-1}}^2\alpha^{-2} \\ &\quad \pm 2E\left[\int_\Omega k^{-1}(\hat{\boldsymbol{q}} + k\nabla\bar{u}^h)(\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h)\,d\Omega\right] \\ (\eta_L^\pm)^2 &= \alpha^2\|\Pi e\| + \alpha^{-2}\|\Pi e_\phi\| \pm 2a(\Pi e, \Pi e_\phi). \end{aligned}$$

Hence,

$$f(\alpha) = \frac{1}{4}[2\alpha^2(\eta^2 - \|\Pi e\|^2) + 2\alpha^{-2}(\eta_\phi^2 - \|\Pi e_\phi\|^2)]. \tag{A.1}$$

Since f is even, we can assume $\alpha > 0$. Then, the minimum is attained when the first derivative is cancelled

$$\begin{aligned} f'(\alpha_{\text{opt}}) = \alpha_{\text{opt}}(\eta^2 - \|\Pi e\|^2) + \alpha_{\text{opt}}^{-3}(\eta_\phi^2 - \|\Pi e_\phi\|^2) = 0 \\ \Leftrightarrow \alpha_{\text{opt}} = \sqrt[4]{\frac{\eta_\phi^2 - \|\Pi e_\phi\|^2}{\eta^2 - \|\Pi e\|^2}}. \end{aligned}$$

Finally, by computing the interval length for the optimum value of α, we attain

$$f(\alpha_{\text{opt}}) = \sqrt{(\eta^2 - \|\Pi e\|^2)(\eta_\phi^2 - \|\Pi e_\phi\|^2)} \leq \sqrt{\eta^2\eta_\phi^2} = \eta\eta_\phi \tag{A.2}$$

which is exactly the half of the size of the interval in eq. (4.36).

Now, we carry out the proof that α_{opt} minimises $\eta^{\text{QoI}}_{\text{Upp}}$. By considering its derivative, it follows that

$$\frac{d\eta^{\text{QoI}}_{\text{Upp}}}{d\alpha} = \frac{1}{2}[\alpha(\eta^2 - \|\Pi e\|^2) + \alpha^{-3}(\eta_\phi^2 - \|\Pi e_\phi\|^2)] = \frac{1}{2}f'(\alpha). \tag{A.3}$$

This implies that it attains its minimum also in α_{opt}. Similarly, we can prove that $\eta^{\text{QoI}}_{\text{Low}}$ attains its minimum in α_{opt}.

We conclude this section by proving that the bounds obtained through the polarisation identity (eq. (4.40)) are sharper than the bounds obtained through Cauchy-Schwarz (eq. (4.36)). Even though, we have just proved the interval is smaller, it could still be the case that the upper bound from the former is smaller than the upper bound from the latter, in other words, we need to show that

$$\eta^{\text{QoI}}_{\text{Upp}} \leq R(\bar{\phi}^h) + \eta\eta_\phi \tag{A.4}$$

By expanding $\eta^{\text{QoI}}_{\text{Upp}}$,

$$\begin{aligned}\eta^{\text{QoI}}_{\text{Upp}} = R(\bar{\phi}^h) + \frac{1}{2}\Bigg[&\sqrt{(\eta^2 - \|\Pi e\|^2)(\eta_\phi^2 - \|\Pi e_\phi\|^2)} \\ &+ E\left[\int_\Omega k^{-1}(\hat{\boldsymbol{q}} + k\nabla\bar{u}^h)(\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h)\,d\Omega\right] + a(\Pi e, \Pi e_\phi)\Bigg]\end{aligned} \tag{A.5}$$

We can bound the first term inside square brackets as in eq. (A.2),

$$\sqrt{(\eta^2 - \|\Pi e\|^2)(\eta_\phi^2 - \|\Pi e_\phi\|^2)} \leq \sqrt{\eta^2\eta_\phi^2} = \eta\eta_\phi \tag{A.6}$$

The second term inside brackets can be bounded using the Cauchy-Schwarz inequality twice,

$$\begin{aligned}&E\left[\int_\Omega k^{-1}(\hat{\boldsymbol{q}} + k\nabla\bar{u}^h)(\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h)\,d\Omega\right] \leq \\ &\leq E\left[\sqrt{\int_\Omega k^{-1}(\hat{\boldsymbol{q}} + k\nabla\bar{u}^h)^2\,d\Omega}\sqrt{\int_\Omega k^{-1}(\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h)^2\,d\Omega}\right] \leq \\ &\leq \sqrt{E\left[\int_\Omega k^{-1}(\hat{\boldsymbol{q}} + k\nabla\bar{u}^h)^2\,d\Omega\right]}\sqrt{E\left[\int_\Omega k^{-1}(\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h)^2\,d\Omega\right]} = \eta\eta_\phi\end{aligned}$$

Provided the same discretization is used in the primal and dual problem, $a(\Pi e, \Pi e_\phi) = 0$. Hence, by introducing the inequalities back in eq. (A.5), we obtain eq. (A.4). By proving a similar inequality for the lower bound, we conclude that the obtained using the polarisation identity is always sharper than the bound obtained using Cauchy-Schwarz.

Appendix B

Independence of the homogenised conductivity field

We prove that

$$\eta_{\mathrm{U_{pp}}}^{\mathrm{QoI}} + s(\bar{u}^h), \tag{B.1}$$

the upper bound for $s(u)$ is independent of the conductivity field. The proof of the lower bound is similar. Expanding and rearranging the terms,

$$\begin{aligned}
&R(\bar{\phi}^h) + s(\bar{u}^h) + \underbrace{\frac{1}{2}\sqrt{(\eta^2 - \|\Pi e\|^2)(\eta_\phi^2 - \|\Pi e_\phi\|^2)}}_{=:C} + \\
&+\frac{1}{2}E\left[\int_\Omega k^{-1}(\hat{\boldsymbol{q}}^h + k\nabla\bar{u}^h)(\hat{\boldsymbol{q}}_\phi^h + k\nabla\bar{\phi}^h)\right] + \frac{1}{2}a(\Pi e, \Pi e_\phi) = \\
&= l(\bar{\phi}^h) - a(\bar{u}^h, \bar{\phi}^h) + s(\bar{u}^h) + C + \\
&\frac{1}{2}\int_\Omega (\hat{\boldsymbol{q}}^h\nabla\bar{\phi}^h + E\left[k^{-1}\right]\hat{\boldsymbol{q}}^h\hat{\boldsymbol{q}}_\phi^h + \hat{\boldsymbol{q}}_\phi^h\nabla\bar{u}^h + E\left[k\right]\nabla\bar{u}^h\nabla\bar{\phi}^h)\,d\Omega + \\
&+\frac{1}{2}[a(\Pi u, \Pi\phi) - a(\Pi u, \bar{\phi}^h) - a(\bar{u}^h, \Pi\phi) + a(\bar{u}^h, \bar{\phi}^h)]
\end{aligned}$$

Grouping in C' the terms that do not depend on $\bar{k}$,

$$C' = C + \frac{1}{2}a(\Pi u, \Pi\phi) + \frac{1}{2}\int_\Omega E\left[k^{-1}\right]\hat{\boldsymbol{q}}^h\hat{\boldsymbol{q}}_\phi^h\,d\Omega, \tag{B.2}$$

and noting that,

$$a(\Pi u, \bar{\phi}^h) = -\int_\Omega \hat{\boldsymbol{q}}^h \nabla \bar{\phi}^h = l(\bar{\phi}^h)$$
$$a(\bar{u}^h, \bar{\phi}^h) = \int_\Omega E\left[k\right] \nabla \bar{u}^h \nabla \bar{\phi}^h \, d\Omega$$

we attain,

$$\eta_{\text{Upp}}^{\text{QoI}} + s(\bar{u}^h) = C' + s(\bar{u}^h) + \frac{1}{2}\int_\Omega \hat{\boldsymbol{q}}_\phi^h \nabla \bar{u}^h \, d\Omega - \frac{1}{2} a(\bar{u}^h, \Pi\phi) \tag{B.3}$$

The proof is concluded by considering the upper bound of a temperature field $\bar{v}^h$, resulting from another conductivity field, although using the same discretisation. Subtracting both quantities

$$\eta_{\text{Upp}}^{\text{QoI}} - \eta_{\text{Upp}}^{\text{QoI}\,\prime} + q(\bar{u}^h - \bar{v}^h) = q(\bar{u}^h - \bar{v}^h) + \frac{1}{2}\int_\Omega \hat{\boldsymbol{q}}_\phi^h \nabla(\bar{u}^h - \bar{v}^h) \, d\Omega - \frac{1}{2} a(\bar{u}^h - \bar{v}^h, \Pi\phi) = 0 \tag{B.4}$$

since $\bar{u}^h - \bar{v}^h \in \bar{\mathcal{U}}_0$ and

$$a(v, \Pi\phi) = -\int_\Omega \hat{\boldsymbol{q}}_\phi^h \nabla v = s(v) \quad \forall v \in \bar{\mathcal{U}}_0. \tag{B.5}$$

Appendix C

The effect of the contrast of the material properties

We show under certain assumptions that as the contrast between the conductivities of the constituents is increased, the intervals defined by the error estimates also grow. More precisely, we will show that η grows with the increase in the contrast and we will assume that p_i the probability of being inside the constituent i is the same for every point of the domain, hence $E[k]$ and $E[k^{-1}]$ are not a function of $\boldsymbol{x}$.

We start by decomposing the problem in eq. (4.3) in the following two problems,

For all $v \in \mathcal{U}_0$, find $\bar{u}_N \in \mathcal{U}_0$

$$\bar{a}(\bar{u}_N, v) = l(v). \tag{C.1}$$

and

For all $v \in \mathcal{U}_0$, find $\bar{u}_D \in \mathcal{U}$

$$\bar{a}(\bar{u}_D, v) = 0. \tag{C.2}$$

We note that for $\bar{k}$ constant over the domain:

- $\bar{u}_N + \bar{u}_D$ is a solution to eq. (4.3). Hence, $\bar{u}_N + \bar{u}_D$ is KA and SA.
- $\bar{u}_D$ and $\nabla\bar{u}_D$ are independent of the value of $\bar{k}$.
- $\bar{\boldsymbol{q}} := -\bar{k}\nabla\bar{u}_N$ is independent of the value of $\bar{k}$.
- $a(\bar{u}_D, \bar{u}_N) = 0$.

We will denote the solutions of those two problems for $\bar{k} = E[k]$ by $\bar{u}_N$ and $\bar{u}_D$ respectively. From the observations, the solutions for $\bar{k} = 1/E[k^{-1}]$ can be expressed as $\bar{u}'_D := \bar{u}_D$ and $\bar{u}'_D := E[k]E[k^{-1}]\bar{u}_N$.

Next, we calculate η^2 with $\hat{\boldsymbol{q}}^h = -1/E[k^{-1}]\nabla(\bar{u}'_D + \bar{u}'_N)$ and with $\bar{u}^h = \bar{u}_D + \bar{u}_N$ and obtain,

$$\eta^2 = \left(E[k] - \frac{1}{E[k^{-1}]}\right)\int_\Omega \nabla\bar{u}_D \cdot \nabla\bar{u}_D \, d\Omega + \left(E[k^{-1}] - \frac{1}{E[k]}\right)\int_\Omega \bar{\boldsymbol{q}} \cdot \bar{\boldsymbol{q}} \, d\Omega. \tag{C.3}$$

We remark that in eq. (C.3), only the expression inside the parenthesis depends on the contrast between the particles. The integrals are independent of $\bar{k}$ and hence independent of the conductivity of the constituents. Expressing $E[k]$ and $E[k^{-1}]$ as,

$$E[k] = \sum_{i=1}^{N} p_i k_i \tag{C.4}$$

$$E[k^{-1}] = \sum_{i=1}^{N} \frac{p_i}{k_i} \quad \text{where} \quad \sum_{i=1}^{N} p_i = 1 \tag{C.5}$$

where k_i is the conductivity and p_i probability of being inside constituent i, we observe that $E[k]$ is the weighted arithmetic mean of the conductivities, while $1/E[k^{-1}]$ is the weighted harmonic mean of the conductivities. Hence, the two expressions inside parenthesis in eq. (C.3) are non-negative due to the inequality between the arithmetic and harmonic mean. In addition, these expressions are equal to 0 only and only if $k_1 = k_2 = ... = k_N$. Furthermore, as the contrast between conductivities increases, the quantities increase without bound.

Appendix D

Derivation of a deterministic bound for the second moment

The proof is done in 2 parts. In the first part, we find an upper bound for m_2 that only depend on computable quantities. Secondly, we show that the computation of this bound is deterministic. Firstly, through algebraic manipulations, we can show that

$$m_2 - s(\bar{u}^h)^2 = 2s(e)s(\bar{u}^h) + E\left[s_\theta(e)^2\right] \tag{D.1}$$

The first term can be upper bounded by

$$2s(e)s(\bar{u}^h) \leq 2\Delta s(\bar{u}^h) \tag{D.2}$$

where Δ is one of the upper estimates from eqs. (4.36) and (4.40) if $s(\bar{u}^h)$ is positive, and a lower estimate otherwise. Regarding the second term, it is easy to prove that

$$s_\theta(e) = a_\theta(e, e_\phi) + R_\theta(\bar{\phi}^h) \tag{D.3}$$

By squaring both sides and taking the expectation, we obtain

$$E\left[s_\theta(e)^2\right] = E\left[a_\theta(e, e_\phi)^2 + R_\theta(\bar{\phi}^h)^2 + 2a_\theta(e, e_\phi)R_\theta(\bar{\phi}^h)\right] \tag{D.4}$$

Firstly, the term $E\left[a_\theta(e, e_\phi)^2\right]$ is bounded. By applying the Cauchy-Schwarz inequality

$$a_\theta(e, e_\phi) \leq \|e\|_\theta \|e_\phi\|_\theta \tag{D.5}$$

followed by the Prager-Synge hypercircle theorem, we obtain

$$\|e\|_\theta \|e_\phi\|_\theta \leq \|\hat{\boldsymbol{q}} + k\nabla\bar{u}^h\|_{\theta,k^{-1}} \|\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h\|_{\theta,k^{-1}} \tag{D.6}$$

This allows us to conclude that

$$E\left[a_\theta(e,e_\phi)^2\right] \leq \underbrace{E\left[\|\hat{\boldsymbol{q}} + k\nabla\bar{u}^h\|^2_{\theta,k^{-1}} \|\hat{\boldsymbol{q}}_\phi + k\nabla\bar{\phi}^h\|^2_{\theta,k^{-1}}\right]}_{=:\gamma^2} \tag{D.7}$$

where

$$\|v\|_{\theta,k^{-1}} := \sqrt{\int_\Omega k(\boldsymbol{x},\theta)^{-1} v(\boldsymbol{x})^2 \, d\Omega}. \tag{D.8}$$

Next, we bound $E\left[R_\theta(\bar{\phi}^h)^2\right]$. By expanding it, we obtain,

$$E\left[R_\theta(\bar{\phi}^h)^2\right] = l(\bar{\phi}^h)^2 - 2l(\bar{\phi}^h)a(\bar{u}^h,\bar{\phi}^h) + \underbrace{E\left[a_\theta(\bar{u}^h,\bar{\phi}^h)^2\right]}_{=:\beta^2}$$

Finally, we have to bound the cross term in equation eq. (D.4)

$$\begin{aligned}
E\left[2a_\theta(e,e_\phi)R_\theta(\bar{\phi}^h)\right] &= 2l(\bar{\phi}^h)a(e,e_\phi) - 2E\left[a_\theta(e,e_\phi)a_\theta(\bar{u}^h,\bar{\phi}^h)\right] \\
&\leq 2\Delta' l(\bar{\phi}^h) - 2E\left[a_\theta(e,e_\phi)a_\theta(\bar{u}^h,\bar{\phi}^h)\right] \\
&\leq 2\Delta' l(\bar{\phi}^h) + 2\sqrt{E\left[a_\theta(e,e_\phi)^2\right]}\sqrt{E\left[a_\theta(\bar{u}^h,\bar{\phi}^h)^2\right]} \\
&\leq 2\Delta' l(\bar{\phi}^h) + 2\gamma\beta
\end{aligned}$$

where again, Δ' is a is one of the upper estimates from eqs. (4.36) and (4.40) if $l(\bar{\phi}^h)$ is positive, and a lower estimate otherwise. Combining those results, we obtain that

$$m_2 - s(\bar{u}^h)^2 \leq (\gamma+\beta)^2 + 2[\Delta s(\bar{u}^h) + \Delta' l(\bar{\phi}^h)] + l(\bar{\phi}^h)[R(\bar{\phi}^h) - a(\bar{u}^h,\bar{\phi}^h)] \tag{D.9}$$

Now, we have to show that β and γ are deterministic quantities. We start by expanding β^2

$$\beta^2 = E\left[a_\theta(\bar{u}^h,\bar{\phi}^h)^2\right] = E\left[\int_\Omega k\nabla\bar{u}^h\cdot\nabla\bar{\phi}^h\,d\Omega \int_\Omega k\nabla\bar{u}^h\cdot\nabla\bar{\phi}^h\,d\Omega\right] \tag{D.10}$$

By combining the two domain integrals into one and switching the order of the integrals (the prime indicates whether a term belongs to the first or the second domain integral),

$$\begin{aligned}
\beta^2 &= E\left[\int_\Omega\int_{\Omega'} kk'(\nabla\bar{u}^h\cdot\nabla\bar{\phi}^h)(\nabla u'^h\cdot\nabla\phi'^h)\,d\Omega\,d\Omega\right] \\
&= \int_\Omega\int_{\Omega'} E\left[kk'\right](\nabla\bar{u}^h\cdot\nabla\bar{\phi}^h)(\nabla u'^h\cdot\nabla\phi'^h)d\Omega\,d\Omega
\end{aligned}$$

which is a deterministic function. The spatial function $E[kk']$ is actually the covariance of the conductivity plus a constant,

$$Cov(k(\boldsymbol{x}), k(\boldsymbol{x}')) = E[k(\boldsymbol{x})k(\boldsymbol{x}')] - E[k(\boldsymbol{x})]E[k(\boldsymbol{x}')] \tag{D.11}$$

Proceeding in an analogue manner, we can show that γ is also a deterministic quantity,

$$\begin{aligned}
\gamma^2 = \int_\Omega \int_{\Omega'} \{ & E[k^{-1}k'^{-1}](\hat{\boldsymbol{q}} \cdot \hat{\boldsymbol{q}})(\hat{\boldsymbol{q}}'_\phi \cdot \hat{\boldsymbol{q}}'_\phi) + E[kk'](\nabla \bar{u}^h \cdot \nabla \bar{u}^h)(\nabla \phi'^h \cdot \nabla \phi'^h) + \\
& E[k^{-1}k'](\hat{\boldsymbol{q}} \cdot \hat{\boldsymbol{q}})(\nabla \phi'^h \cdot \nabla \phi'^h) + E[kk'^{-1}](\nabla \bar{u}^h \cdot \nabla \bar{u}^h)(\hat{\boldsymbol{q}}'_\phi \cdot \hat{\boldsymbol{q}}'_\phi) \} \, d\Omega \, d\Omega \\
& + 2 \int_\Omega \hat{\boldsymbol{q}}_\phi \cdot \nabla \bar{\phi}^h \, d\Omega \left[\int_\Omega \{ E[k^{-1}](\hat{\boldsymbol{q}} \cdot \hat{\boldsymbol{q}}) + E[k](\nabla \bar{u}^h \cdot \nabla \bar{u}^h) \} \, d\Omega \right] \\
& + 2 \int_\Omega \hat{\boldsymbol{q}} \cdot \nabla \bar{u}^h \, d\Omega \left[\int_\Omega \{ E[k^{-1}](\hat{\boldsymbol{q}}_\phi \cdot \hat{\boldsymbol{q}}_\phi) + E[k](\nabla \phi^h \cdot \nabla \phi^h) \} \, d\Omega \right] \\
& + 4 \int_\Omega \hat{\boldsymbol{q}} \cdot \nabla \bar{u}^h \, d\Omega \int_\Omega \hat{\boldsymbol{q}}_\phi \cdot \nabla \bar{\phi}^h \, d\Omega
\end{aligned}$$

Appendix E

Alternative enriched flux construction

We consider in this section a generalisation of the approximation described in section 5.7. Instead of defining the shape functions $\boldsymbol{H}$ as global shape functions, we define them independently on each of the elements and we enforce normal continuity of the fluxes along the element edges (otherwise the approximation obtained is nonconforming). We will show that this approach reduces to the former. The proof will invoke a few theorems from complex analysis, namely Morera's theorem and Cauchy's integral theorem ([4]).

Firstly, we note that using the same considerations of section 5.7, we can show that each of shape functions $\boldsymbol{H}$ has the form

$$\boldsymbol{H} = \begin{pmatrix} a & b \\ -b & a \end{pmatrix} \tag{E.1}$$

with

$$\begin{aligned} \frac{\partial a}{\partial x} &= \frac{\partial b}{\partial y} \\ \frac{\partial a}{\partial y} &= -\frac{\partial b}{\partial x} \end{aligned} \quad \text{(Cauchy-Rienmann eqns.)} \tag{E.2}$$

We proceed the proof by showing that the imposition of normal continuity of $\boldsymbol{q}^h$ along the element edges is equivalent to require $\boldsymbol{H}$ to be continuous across the element edges.

The normal continuity constraint reads

$$\left(\boldsymbol{q}^h(\boldsymbol{x},\theta)\Big|_{a,\gamma} - \boldsymbol{q}^h(\boldsymbol{x},\theta)\Big|_{b,\gamma}\right)\cdot\boldsymbol{n}(\boldsymbol{x}) = 0 \tag{E.3}$$

where a, b indicate two elements that share an edge γ and $\boldsymbol{n}$ denotes the normal unit vector along the edge and oriented according to some predefined criteria. The expansion of this equation reads

$$\left([\hat{\boldsymbol{q}}^h(\boldsymbol{x}) + \sum_{d,i} a_i^d \boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^d(\boldsymbol{x},\theta)]\Big|_{a,\gamma} - [\hat{\boldsymbol{q}}^h(\boldsymbol{x}) + \sum_{d,i} a_i^{d'} \boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^d(\boldsymbol{x},\theta)]\Big|_{b,\gamma}\right)\cdot\boldsymbol{n}(\boldsymbol{x}) = 0. \tag{E.4}$$

Since this must hold true for every single realisation, the constraint can be rewritten as the following three constraints

$$\left(\hat{\boldsymbol{q}}^h(\boldsymbol{x})\Big|_{a,\gamma} - \hat{\boldsymbol{q}}^h(\boldsymbol{x})\Big|_{b,\gamma}\right)\cdot\boldsymbol{n}(\boldsymbol{x}) = 0 \tag{E.5}$$

$$\left(\sum_i a_i^x \boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^x(\boldsymbol{x},\theta)\Big|_{a,\gamma} - \sum_i a_i^{x\prime} \boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^x(\boldsymbol{x},\theta)\Big|_{b,\gamma}\right)\cdot\boldsymbol{n}(\boldsymbol{x}) = 0 \tag{E.6}$$

$$\left(\sum_i a_i^y \boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^y(\boldsymbol{x},\theta)\Big|_{a,\gamma} - \sum_i a_i^{y\prime} \boldsymbol{H}_i(\boldsymbol{x})\mathbb{Q}^y(\boldsymbol{x},\theta)\Big|_{b,\gamma}\right)\cdot\boldsymbol{n}(\boldsymbol{x}) = 0 \tag{E.7}$$

Taking common factor $\mathbb{Q}$ in the second and third equation gives

$$\boldsymbol{n}(\boldsymbol{x})^T\left(\sum_i a_i^d \boldsymbol{H}_i(\boldsymbol{x})\Big|_{a,\gamma} - \sum_i a_i^{d'} \boldsymbol{H}_i(x)\Big|_{b,\gamma}\right)\mathbb{Q}^d(x,\theta) = 0. \tag{E.8}$$

Since this must hold for any realisation,

$$\boldsymbol{n}(\boldsymbol{x})^T\left(\sum_i a_d^i \boldsymbol{H}_i(\boldsymbol{x})\Big|_{a,\gamma} - \sum_i a_i^{d'} \boldsymbol{H}_i(x)\Big|_{b,\gamma}\right) = 0. \tag{E.9}$$

Since $\boldsymbol{H}_i$ is 2×2 matrix, the coefficients must satisfy two equations for each element edge. Noting that the first and the second column of matrix $\boldsymbol{H}$ are orthogonal vectors (recall eq. (E.1)), eq. (E.9) prescribes the vector function defined by the first column to be normally and tangentially continuous along the element edge. In other words, the vector function $(a\ -b)^T$ is continuous in the entire domain. This also implies that a complex function f defined by $f(x,y) = a(x,y) + ib(x,y)$ is continuous over the entire domain.

We recall now Morera's theorem [4]

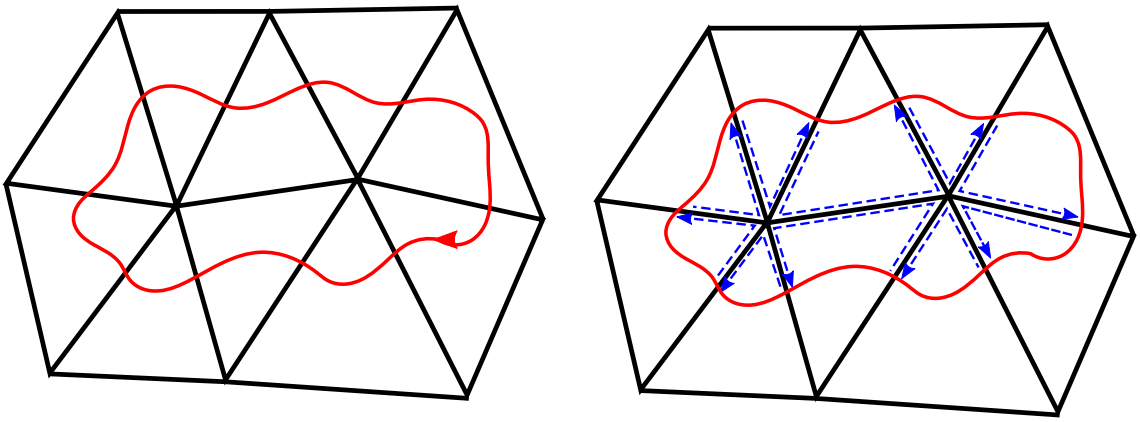


Figure E.1: Any path γ can be rewritten as a sum of paths that are contained inside each element.

Theorem E.1. *If $f(z)$ is defined and continuous in a region Ω, and if $\int_\gamma f\,dz = 0$ for all closed curves γ in Ω, then $f(z)$ is analytic in Ω.*

Since f is holomorphic inside each element (it fulfils the Cauchy-Rienmann equations, eq. (E.2)), due to Cauchy's integral theorem, it follows that $\int_\gamma f\,dz = 0$ for closed curves γ fully contained in an element. Every path γ can be rewritten as the sum of paths γ_i fully contained in a single element (see fig. E.1) and owing to the fact that $f(z)$ is continuous along the element edges,

$$\int_\gamma f(z)\,dz = \sum_i \int_{\gamma_i} f(z)\,dz = 0 \tag{E.10}$$

This allows us to conclude that $f(z)$ is analytic and in particular that the functions a and b are $C^\infty(\Omega)$. This reduces this approach to the one presented in section 5.7.

Bibliography

[1] Wikipedia - Thales' theorem. `https://en.wikipedia.org/wiki/Thales'_theorem`. Accessed: 2016-06-15.

[2] A. Abdulle and A. Nonnenmacher. Adaptive finite element heterogeneous multiscale method for homogenization problems. *Computer Methods in Applied Mechanics and Engineering*, 200(37-40):2710–2726, 2011.

[3] Assyr Abdulle. On a priori error analysis of fully discrete heterogeneous multiscale fem. *Multiscale Modeling & Simulation*, 4(2):447–459, 2005.

[4] Lars V. Ahlfors. *Complex analysis: an introduction to the theory of analytic functions of one complex variable*. 1979.

[5] Mark Ainsworth and J Tinsley Oden. A posteriori error estimation in finite element analysis. *Computer Methods in Applied Mechanics and Engineering*, 142(1-2):1–88, 1997.

[6] Mark Ainsworth and J Tinsley Oden. *A posteriori error estimation in finite element analysis*, volume 37. John Wiley & Sons, 2011.

[7] Ahmad Akbari Rahimabadi, Pierre Kerfriden, and Stéphane Bordas. Scale selection in nonlinear fracture mechanics of heterogeneous materials. *Philosophical Magazine*, 95(28-30):3328–3347, oct 2015.

[8] Ivo Babuška and A Miller. A feedback finite element method with a posteriori error estimation: Part i. the finite element method and some basic properties of the a

posteriori error estimator. *Computer Methods in Applied Mechanics and Engineering*, 61(1):1–40, 1987.

[9] Ivo Babuška and Werner C Rheinboldt. A-posteriori error estimates for the finite element method. *International Journal for Numerical Methods in Engineering*, 12(10):1597–1615, 1978.

[10] Corinna Bahriawati and Carsten Carstensen. Three matlab implementations of the lowest-order raviart-thomas mfem with a posteriori error control. *Computational Methods in Applied Mathematics Comput. Methods Appl. Math.*, 5(4):333–361, 2005.

[11] Roland Becker and Rolf Rannacher. *A feed-back approach to error control in finite element methods: Basic analysis and examples.* Citeseer, 1996.

[12] Ted Belytschko and Tom Black. Elastic crack growth in finite elements with minimal remeshing. *International journal for numerical methods in engineering*, 45(5):601–620, 1999.

[13] Jacques Besson, Georges Cailletaud, Jean-Louis Chaboche, and Samuel Forest. *Non-linear mechanics of materials*, volume 167. Springer Science & Business Media, 2009.

[14] Susanne Brenner and Ridgway Scott. *The mathematical theory of finite element methods*, volume 15. Springer Science & Business Media, 2007.

[15] Franco Brezzi, Jim Douglas, and L Donatella Marini. Two families of mixed finite elements for second order elliptic problems. *Numerische Mathematik*, 47(2):217–235, 1985.

[16] Carsten Carstensen and Stefan A Funken. Fully reliable localized error control in the fem. *SIAM Journal on Scientific Computing*, 21(4):1465–1484, 1999.

[17] L Chamoin and P Ladeveze. Bounds on history-dependent or independent local quantities in viscoelasticity problems solved by approximate methods. *International Journal for Numerical Methods in Engineering*, 71(12):1387–1411, 2007.

[18] Ludovic Chamoin and Pedro Diez. *Verifying Calculations-Forty Years On: An Overview of Classical Verification Techniques for FEM Simulations*. Springer, 2015.

[19] Fehmi Cirak and Ekkehard Ramm. A posteriori error estimation and adaptivity for linear elasticity using the reciprocal theorem. *Computer Methods in Applied Mechanics and Engineering*, 156(1):351–362, 1998.

[20] A Clément, Christian Soize, and Julien Yvonnet. Computational nonlinear stochastic homogenization using a nonconcurrent multiscale approach for hyperelastic heterogeneous microstructures analysis. *International Journal for Numerical Methods in Engineering*, 91(8):799–824, 2012.

[21] Donald L. Cohn. *Measure Theory*. Birkhäuser Advanced Texts Basler Lehrbücher. Springer New York, New York, NY, 2013.

[22] JP Moitinho De Almeida and OJB Almeida Pereira. A set of hybrid equilibrium finite element models for the analysis of three-dimensional solids. *International Journal for Numerical Methods in Engineering*, 39(16):2789–2802, 1996.

[23] J.P.Moitinho de Almeida and J.A.Teixeira de Freitas. Alternative approach to the formulation of hybrid equilibrium finite elements. *Computers & Structures*, 40(4):1043–1047, jan 1991.

[24] Yalchin Efendiev and Thomas Y Hou. *Multiscale Finite Element Methods*. Springer New York, New York, NY, 2009.

[25] Frédéric Feyel. Multiscale FE2 elastoviscoplastic analysis of composite structures. *Computational Materials Science*, 16(1–4):344–354, 1999.

[26] Frédéric Feyel. A multilevel finite element method (FE2) to describe the response of highly non-linear structures using generalized continua. *Computer Methods in Applied Mechanics and Engineering*, 192(28-30):3233–3244, 2003.

[27] Jacob Fish, Vasilina Filonova, and Dimitrios Fafalis. Computational continua revisited. *International Journal for Numerical Methods in Engineering*, 102(3-4):332–378, 2015.

[28] Jacob Fish and Sergey Kuznetsov. Computational continua. *International Journal for Numerical Methods in Engineering*, 84(7):774–802, 2010.

[29] Izrail Moiseevitch Gelfand, Sergei Fomin, and Richard A Silverman. *Calculus of variations*. Courier Corporation, 2000.

[30] Christophe Geuzaine and Jean-François Remacle. Gmsh: A 3-d finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79(11):1309–1331, 2009.

[31] Roger G Ghanem and Pol D Spanos. *Stochastic finite elements: a spectral approach*. Courier Corporation, 2003.

[32] IM Gitman, H Askes, and LJ Sluys. Coupled-volume multi-scale modelling of quasi-brittle material. *European Journal of Mechanics-A/Solids*, 27(3):302–327, 2008.

[33] Octavio Andres González-Estrada, E Nadal, JJ Ródenas, Pierre Kerfriden, Stéphane Pierre-Alain Bordas, and FJ Fuenmayor. Mesh adaptivity driven by goal-oriented locally equilibrated superconvergent patch recovery. *Computational Mechanics*, 53(5):957–976, 2014.

[34] Thomas Grätsch and Klaus-Jürgen Bathe. A posteriori error estimation techniques in practical finite element analysis. *Computers & Structures*, 83(4-5):235–265, 2005.

[35] Gaël Guennebaud, Benoît Jacob, et al. Eigen v3. http://eigen.tuxfamily.org, 2010.

[36] Amy Henderson, Jim Ahrens, Charles Law, et al. *The ParaView Guide*. Kitware Clifton Park, NY, 2004.

[37] Richard Hill. The elastic behaviour of a crystalline aggregate. *Proceedings of the Physical Society. Section A*, 65(5):349, 1952.

[38] 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, jun 1997.

[39] Thomas J R Hughes. Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. *Computer Methods in Applied Mechanics and Engineering*, 127(1-4):387–401, 1995.

[40] 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, nov 1998.

[41] Thomas J.R. Hughes and James R. Stewart. A space-time formulation for multiscale phenomena. *Journal of Computational and Applied Mathematics*, 74(1-2):217–229, 1996.

[42] P. Kanouté, D. P. Boso, J. L. Chaboche, and B. A. Schrefler. Multiscale Methods for Composites: A Review. *Archives of Computational Methods in Engineering*, 16(1):31–75, mar 2009.

[43] DW Kelly, De SR Gago, OC Zienkiewicz, I Babuska, et al. A posteriori error analysis and adaptive processes in the finite element method: Part i—error analysis. *International journal for numerical methods in engineering*, 19(11):1593–1619, 1983.

[44] Martin Kempeneers, Jean-François Debongnie, and Pierre Beckers. Pure equilibrium tetrahedral finite elements for global error estimation by dual analysis. *International Journal for Numerical Methods in Engineering*, 81(4):513–536, 2010.

[45] Pierre Kerfriden and Daniel Paladim. A new homogenisation scheme with certified accuracy for random media. Presented in VII European Congress on Computational Methods in Applied Sciences and Engineering, 2016.

[46] Pierre Kerfriden, Daniel Paladim, and Stéphane Bordas. Homogenisation methods with guaranteed accuracy: quantifying the scale separability. Presented in 5th International Conference on Computational Methods (ICCM2014), 2014.

[47] Pierre Ladevèze, Benoît Blaysat, and Éric Florentin. Strict upper bounds of the error in calculated outputs of interest for plasticity problems. *Computer Methods in Applied Mechanics and Engineering*, 245:194–205, 2012.

[48] Pierre Ladevèze and Ludovic Chamoin. Calculation of strict error bounds for finite element approximations of non-linear pointwise quantities of interest. *International Journal for Numerical Methods in Engineering*, 84(13):1638–1664, dec 2010.

[49] Pierre Ladevèze and Eric Florentin. Verification of stochastic models in uncertain environments using the constitutive relation error method. *Computer Methods in Applied Mechanics and Engineering*, 196(1-3):225–234, dec 2006.

[50] Pierre Ladevèze and Dominique Leguillon. Error estimate procedure in the finite element method and applications. *SIAM Journal on Numerical Analysis*, 20(3):485–509, 1983.

[51] Pierre Ladevèze and Jean-Pierre Pelle. *Mastering calculations in linear and nonlinear mechanics*. Springer, 2005.

[52] Fredrik Larsson and Kenneth Runesson. Adaptive computational meso–macro-scale modeling of elastic composites. *Computer methods in applied mechanics and engineering*, 195(4):324–338, 2006.

[53] Fredrik Larsson and Kenneth Runesson. On two-scale adaptive FE analysis of micro-heterogeneous media with seamless scale-bridging. *Computer Methods in Applied Mechanics and Engineering*, 200(37-40):2662–2674, 2011.

[54] Fredrik Larsson, Kenneth Runesson, Sepehr Saroukhani, and Reza Vafadari. Computational homogenization based on a weak format of micro-periodicity for rve-problems. *Computer Methods in Applied Mechanics and Engineering*, 200(1):11–26, 2011.

[55] Raytcho Lazarov, Sergey Repin, and SATYENDRA Tomar. Functional a posteriori error estimates for discontinuous galerkin approximations of elliptic problems. *Numer. Methods Partial Differential Equations*, 25:952–971, 2006.

[56] S. Loehnert, C. Prange, and P. Wriggers. Error controlled adaptive multiscale XFEM simulation of cracks. *International Journal of Fracture*, 178(1-2):147–156, 2012.

[57] Luc Machiels, Yvon Maday, and Anthony T Patera. A "flux-free" nodal neumann subproblem approach to output bounds for partial differential equations. *Comptes Rendus de l'Académie des Sciences-Series I-Mathematics*, 330(3):249–254, 2000.

[58] Yvon Maday. Explicit residual methods. In *Verifying Calculations-Forty Years On*, pages 1–18. Springer, 2016.

[59] George Marsh. Composites flying high (Part 1). `http://www.materialstoday.com/composite-applications/features/composites-flying-high-part-1/`. Accessed: 2016-06-15.

[60] Jens M Melenk and Ivo Babuška. The partition of unity finite element method: basic theory and applications. *Computer methods in applied mechanics and engineering*, 139(1):289–314, 1996.

[61] Nicolas Moës, J.Tinsley Oden, and Tarek I. Zohdi. Investigation of the interactions between the numerical and the modeling errors in the Homogenized Dirichlet Projection Method. *Computer Methods in Applied Mechanics and Engineering*, 159(1-2):79–101, jul 1998.

[62] José Paulo Moitinho de Almeida and Edward Maunder. *Equilibrium finite element formulations*. In print, 2016.

[63] Pedro Morin, Ricardo Nochetto, and Kunibert Siebert. Local problems on stars: a posteriori error estimators, convergence, and performance. *Mathematics of Computation*, 72(243):1067–1097, 2003.

[64] Vinh Phu Nguyen, Martijn Stroeven, and Lambertus Johannes Sluys. Multiscale Continuous and Discontinuous Modeling of Heterogeneous Materials :. 3(4):1–42, 2011.

[65] J. Tinsley Oden, Kumar Vemaganti, and Nicolas Moës. Hierarchical modeling of heterogeneous solids. *Computer Methods in Applied Mechanics and Engineering*, 172(98):3–25, 1999.

[66] J.Tinsley Oden and Serge Prudhomme. Estimation of Modeling Error in Computational Mechanics. *Journal of Computational Physics*, 182(2):496–515, nov 2002.

[67] J.Tinsley Oden and Kumar S. Vemaganti. Estimation of Local Modeling Error and Goal-Oriented Adaptive Modeling of Heterogeneous Materials. *Journal of Computational Physics*, 164(1):22–47, oct 2000.

[68] J.Tinsley Oden and Tarek I. Zohdi. Analysis and adaptive modeling of highly heterogeneous elastic structures. *Computer Methods in Applied Mechanics and Engineering*, 148(3-4):367–391, sep 1997.

[69] Daniel Paladim, Stéphane Bordas, and Pierre Kerfriden. A two scale error bound for the modelling random particulate materials. In preparation.

[70] Daniel Paladim, José de Almeida, Stéphane Bordas, and Pierre Kerfriden. Guaranteed error bounds for the homogenisation of random materials. Presented in ACME-UK, 2016.

[71] Daniel Paladim, José de Almeida, Stéphane Bordas, and Pierre Kerfriden. Guaranteed error bounds in homogenisation: an optimum stochastic approach to preserve the numerical separation of scales. *International Journal for Numerical Methods in Engineering*, 2016, Accepted for publication.

[72] Daniel Paladim, Pierre Kerfriden, and Stéphane Bordas. Efficient modelling of random heterogenenous materials with an uniform probability density function. Presented in 11th World Congress on Computational Mechanics, 2014.

[73] Daniel Paladim, Pierre Kerfriden, and Stéphane Bordas. Guaranteed error bounds for rve based homogenisation. Presented in VII European Congress on Computational Methods in Applied Sciences and Engineering, 2016.

[74] Daniel Paladim, Pierre Kerfriden, José de Almeida, and Stéphane Bordas. An adaptive scheme for homogenised domains. Presented in International Conference on Adaptive Modelling and Simulation, 2015.

[75] Daniel Paladim, Pierre Kerfriden, José de Almeida, Mathilde Chevreuil, and Stéphane Bordas. Advances in error estimation for homogenisation. Presented in 13th US National Congress on Computational Mechanics, 2015.

[76] Núria Parés, Pedro Díez, and Antonio Huerta. Subdomain-based flux-free a posteriori error estimators. *Computer Methods in Applied Mechanics and Engineering*, 195(4-6):297–323, jan 2006.

[77] F Pled, Ludovic Chamoin, and Pierre Ladevèze. On the techniques for constructing admissible stress fields in model verification: Performances on engineering examples. *International Journal for Numerical Methods in Engineering*, 88(5):409–441, nov 2011.

[78] William Prager and John L Synge. Approximations in elasticity based on the concept of function space. *Quart. Appl. Math*, 5(3):241–269, 1947.

[79] S. Prudhomme and J.T. Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Computer Methods in Applied Mechanics and Engineering*, 176(1-4):313–331, jul 1999.

[80] Serge Prudhomme and J Tinsley Oden. On goal-oriented error estimation for elliptic problems: application to the control of pointwise errors. *Computer Methods in Applied Mechanics and Engineering*, 176(1):313–331, 1999.

[81] P. Raghavan and S. Ghosh. *Concurrent multi-scale analysis of elastic composites by a multi-level computational model*, volume 193. 2004.

[82] P Raviart and J Thomas. A mixed finite element method for 2-nd order elliptic problems. *Mathematical aspects of finite element methods*, pages 292–315, 1977.

[83] A Reuss. Berechnung der fließgrenze von mischkristallen auf grund der plastizitätsbedingung für einkristalle. *ZAMM-Journal of Applied Mathematics and Mechanics/Zeitschrift für Angewandte Mathematik und Mechanik*, 9(1):49–58, 1929.

[84] JJ Ródenas, M Tur, FJ Fuenmayor, and A Vercher. Improvement of the superconvergent patch recovery technique by the use of constraint equations: the spr-c technique. *International Journal for Numerical Methods in Engineering*, 70(6):705–727, 2007.

[85] Marie E Rognes, Robert C Kirby, and Anders Logg. Efficient assembly of h(div) and h(curl) conforming finite elements. *SIAM Journal on Scientific Computing*, 31(6):4130–4151, 2009.

[86] Albert Romkes, J. Tinsley Oden, and Kumar Vemaganti. Multi-scale goal-oriented adaptive modeling of random heterogeneous materials. *Mechanics of Materials*, 38(8-10):859–872, aug 2006.

[87] Albert Romkes, Kumar Vemaganti, and JT Oden. The extension of the GOALS algorithm to the analysis of elastostatics problems of random heterogeneous materials. *ICES Report*, (September), 2004.

[88] Enrique Sanchez-Palencia and André Zaoui. Homogenization techniques for composite media. In *Homogenization Techniques for Composite Media*, volume 272, 1987.

[89] Erwin Stein and Marcus Rüter. Finite Element Methods for Elasticity with Error-controlled Discretization and Model Adaptivity. In *Encyclopedia of Computational Mechanics*. John Wiley & Sons, Ltd, Chichester, UK, oct 2007.

[90] M. Steven Greene, Yu Liu, Wei Chen, and Wing Kam Liu. Computational uncertainty analysis in multiresolution materials via stochastic constitutive theory. *Computer Methods in Applied Mechanics and Engineering*, 200(1-4):309–325, jan 2011.

[91] İ. Temizer, T. Wu, and P. Wriggers. On the optimality of the window method in computational homogenization. *International Journal of Engineering Science*, 64:66–73, mar 2013.

[92] Salvatore Torquato. *Random heterogeneous materials: microstructure and macroscopic properties*, volume 16. Springer Science & Business Media, 2013.

[93] Kumar S. Vemaganti and J.Tinsley Oden. Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. *Computer Methods in Applied Mechanics and Engineering*, 190(46-47):6089–6124, sep 2001.

[94] R. Verfürth. A review of a posteriori error estimation techniques for elasticity problems. *Computer Methods in Applied Mechanics and Engineering*, 176(1-4):419–440, jul 1999.

[95] Woldemar Voigt. Ueber die beziehung zwischen den beiden elasticitätsconstanten isotroper körper. *Annalen der Physik*, 274(12):573–587, 1889.

[96] Julien Waeytens, Ludovic Chamoin, and Pierre Ladevéze. Guaranteed error bounds on pointwise quantities of interest for transient viscodynamics problems. *Computational Mechanics*, 49(3):291–307, 2012.

[97] E Weinan, Bjorn Engquist, and Zhongyi Huang. Heterogeneous multiscale method: a general methodology for multiscale modeling. *Physical Review B*, 67(9):092101, 2003.

[98] Zhimin Zhang and Jianzhong Zhu. Analysis of the superconvergent patch recovery technique and a posteriori error estimator in the finite element method (i). *Computer methods in applied mechanics and engineering*, 123(1):173–187, 1995.

[99] O. C. Zienkiewicz. Displacement and equilibrium models in the finite element method by B. Fraeijs de Veubeke, Chapter 9, Pages 145–197 of Stress Analysis, Edited by O. C. Zienkiewicz and G. S. Holister, Published by John Wiley & Sons, 1965. *International Journal for Numerical Methods in Engineering*, 52(3):287–342, sep 2001.

[100] Olgierd C Zienkiewicz and Jian Z Zhu. A simple error estimator and adaptive procedure for practical engineerng analysis. *International Journal for Numerical Methods in Engineering*, 24(2):337–357, 1987.

[101] Olgierd Cecil Zienkiewicz and Jian Zhong Zhu. The superconvergent patch recovery and a posteriori error estimates. part 1: The recovery technique. *International Journal for Numerical Methods in Engineering*, 33(7):1331–1364, 1992.

[102] Olgierd Cecil Zienkiewicz and Jian Zhong Zhu. The superconvergent patch recovery and a posteriori error estimates. part 2: Error estimates and adaptivity. *International Journal for Numerical Methods in Engineering*, 33(7):1365–1382, 1992.

[103] Tarek I. Zohdi, J.Tinsley Oden, and Gregory J. Rodin. Hierarchical modeling of heterogeneous bodies. *Computer Methods in Applied Mechanics and Engineering*, 138(1-4):273–298, dec 1996.

[104] Tarek I. Zohdi and Peter Wriggers. *An introduction to computational micromechanics*. Springer Science & Business Media, 2008.