UNIVERSIDAD NACIONAL DE COLOMBIA

On generalized multiscale methods for flow in complex porous media and their applications

Luis Fernando Contreras Hernandez

Universidad Nacional de Colombia
Facultad de Ciencias, Departamento de matemáticas
Bogotá, Colombia
2022

On generalized multiscale methods for flow in complex porous media and their applications

Luis Fernando Contreras Hernandez

Tesis o trabajo de grado presentada(o) como requisito parcial para optar al título de:
Doctorado en ciencias matemáticas.

Director(a):
Ph.D. Juan Galvis

Línea de Investigación:
Numerical analysis, Partial differential equations.

Universidad Nacional de Colombia
Facultad de ciencias, Departamento de matemáticas.
Bogotá, Colombia
2022

Resumen

Sobre métodos multiescala generalizados para flujo en medios porosos complejos y sus aplicaciones

En este documento se estudia el Método de Elementos Finitos Multiescala Generalizados (GMsFEM), el cual trata de la construcción de funciones base espectrales multiescala que están diseñadas para problemas de alto contraste. Las funciones base multiescala se construyen a partir del producto entre los vectores propios, construidos a partir de un problema espectral local y una partición de la unidad sobre el dominio de estudio. Los valores propios detectan características importantes de las soluciones que no son capturadas por las funciones base multiescala iniciales. En este trabajo, se presenta un estudio de convergencia donde las estimaciones de error son generales, y están escritas en términos de los valores propios asociados a los vectores propios no utilizados en la construcción. El análisis de errores implica normas locales y globales que miden la descomposición de la expansión de la solución en términos de vectores propios locales, esto se logra con una elección cuidadosa de las funciones de base multiescala iniciales y la configuración de los problemas de valores propios. Se presentan dos aplicaciones numéricas importantes: la primera, es el problema de represa con frontera libre planteado sobre un medio heterogéneo de alto contraste, donde introducimos una variable de tiempo ficticia que motiva una discretización de tiempo adecuada que puede entenderse como una iteración de punto fijo a la solución de estado estacionario, y usamos el método de dualidad para tratar con los términos no lineales multivaluados involucrados; luego, se calculan aproximaciones eficientes de la presión y la saturación usando el método GMsFEM. La segunda aplicación es la solución de una ecuación parabólica donde al implementar discretizaciones de tiempo como diferencias finitas o integradores exponenciales sobre un coeficiente de alto contraste, puede no ser práctico porque cada iteración de tiempo necesita el cálculo de operadores matriciales que involucran matrices dispersas, muy grandes y mal condicionadas; es por esto que el GMsFEM es importante ya que permite la obtención de la solución del problema de una forma más sencilla, permitiendo combinar GMsFEM con el método de integradores exponenciales en el tiempo para obtener una buena aproximación de la solución temporal final.

Palabras clave: Métodos Multiescala Generalizados, problemas de alto contraste, problemas multiescala, Elementos Finitos. .

Abstract

In this document, the Generalized Multiscale Finite Element Method (GMsFEM) is studied, which deals with constructing multiscale spectral basis functions designed for high-contrast multiscale problems. The multiscale basis functions are built from the product of the eigenvectors, computed from a local spectral problem and a partition of unity over the study domain. The eigenvalues detect essential features of the solutions that are not captured by the initial multiscale basis functions. This document reviews the general convergence study where the error estimates are written in terms of the eigenvalues associated with the eigenvectors not used in the construction. Error analysis involves local and global norms that measure the convergence speed of the expansion of the solution in terms of local eigenvectors; this is achieved with a careful choice of the initial multiscale basis functions and the configuration of the eigenvalue problems. Two novel important numerical applications are presented: the first is the free-boundary dam problem posed on a heterogeneous high-contrast medium, where we introduce a fictitious time variable that motivates an adequate time discretization that can be understood as a fixed-point iteration. For the steady-state solution, we use the duality method to deal with the multivalued nonlinear terms involved; then, efficient approximations of pressure and saturation are calculated using the GMsFEM method. The second application is the solution of a parabolic equation. Here implementing time discretizations, such as finite differences or exponential integrators in the presence of a high contrast coefficient, it may not be practical in because each time iteration one needs the computation of matrix operators involving very large and extremely ill-conditioned sparse matrices. The GMsFEM is essential since it allows obtaining the solution of the problem more simply, allowing to combine the GMsFEM with the method of exponential integrators in time to get a good approximation of the final temporary solution.

Keywords: Generalize Multiscale Finite Elements Methods, High-contrast problems, Multiscale problem, Finite Elements Methods

Content

List of Figures

List of Tables

1 Introduction

Many applications, such as modeling environmental problems and groundwater flow, have become highly relevant in everyday life and academia; see [13, 4]. The primary modeling tool in these areas is the diffusion of substances in a heterogeneous porous medium with high-contrast multiscale permeability properties. Mathematical modeling and numerical simulation are relevant here to understand this problem but face several challenges regarding the accuracy and computational efficiency of the implemented numerical methods.

Numerical methods for porous media flow include many techniques, for example, the finite difference method, finite volume method, and finite element method. When high-contrast multiscale coefficients model the permeability of the porous media, the resolution needed to obtain good approximation results is impractical since any implementation needs to solve a very large and ill-conditioned problem. Several multiscale methods are proposed in this setting to efficiently approximate solutions to these problems (see [33, 28]).

In classical Multiscale Finite Element Methods (MsFEM), a coarse mesh is explored. There is a fine mesh that resolves all variations and discontinuities of the coefficients, but this mesh is too fine to allow efficient computations. Therefore, a coarse mesh is introduced and the main idea is to construct coarse basis functions in each coarse node neighborhood using the local mesh and local information of the coefficients. After the basis functions are constructed, a global formulation of the problem at the coarse scale is formulated and solved efficiently since the size of the coarse matrices is proportional to the number of coarse blocks (see [42, 1, 33]). One such methodology is known as the Generalized Multiscale Finite Element Method (GMsFEM), which is relatively new. The main goal of the GMsFEM is to build coarse spaces for the MsFEM that result in accurate coarse-scale solutions. This methodology was first developed in [25, 30] related to the robustness of iterative domain decomposition methods to solve the elliptic equation with heterogeneous coefficients. We consider the following problem

$$-\text{div}(\kappa(x)\nabla u) = f, \tag{1-1}$$

where $\kappa(x)$ is a multiscale high-contrast heterogeneous field. In particular, it is assumed that $\kappa(x) \geq c_0 > 0$ (bounded below), while $\kappa(x)$ can have very large values and local variations. A main ingredient in the construction was local generalized eigenvalue problems and partition of unity functions to construct the coarse spaces. In addition to using one coarse

function per coarse node, the GMsFEM proposed to use several multiscale basis functions per coarse node. These basis functions represent important solution features within a coarse grid block and are computed using eigenvectors from a local eigenvalue problem. Then, in the works [57, 14], some studies of the coarse approximation properties of the GMsFEM were carried out.

This work shows a convergence analysis for the GMsFEM, which is suitable for computational practice. We assume square integrability of the right side f. To obtain error bounds in terms of the decay of the eigenvalues used in the construction, we assume that the problem is regular in the sense that local eigenvectors can well approximate the solution well enough. The main difference between classical finite element analysis and the GMsFEM is when trying to write the interpolation error estimates, where the solution is assumed to be smooth enough in the classical Sobolev sense using Hilbert rules (at least for elliptic problems). In the case of discontinuous multiscale coefficients, it is known that solutions are not regular in the classical sense (see [1]). Therefore, the classic finite element analysis arguments are not suitable. In this document and based on [1], we can write interpolation error estimates using rules appropriate to the problem. In particular, to measure the "smoothness" of the solution, we use the decay of the solution expansion in terms of global eigenvectors. This is motivated by the fact that eigenvectors are a good model for smooth functions for a given elliptic operator. We then define global norms, using the decay of the expansion on global eigenvectors. We also define the local norms using the expansion decay in terms of local eigenvectors (computed locally in a coarse node neighborhood). A main result is that we can compare the local and global norms. Also, we test the error estimates in terms of the eigenvalues of the eigenvalue problem used in the construction.

This thesis's main contribution is applying the GMsFEM to two important non-linear problems. One of the applications studied in this document is the problem of heterogeneous dams initially raised in [49]. This model considers the numerical homogenization approximation of a free boundary dam problem in a multiscale heterogeneous environment. In the work presented here, we deal with the numerical expansion of a similar free boundary problem posed in a high-contrast multiscale medium, in this case, without the scale separation assumption. Therefore, although we can formulate the free boundary problem, due to the multiscale nature of porous media, very high resolution, e.g., in a finite element approximation, will be needed to obtain realistic results, leading impractical computations.

Following [9, 49], we first approximate the nonlinear steady-state dam problem: By using Darcy's law for porous media and the relation between pressure and water saturation, we obtain

$$-g\frac{\partial(\theta\kappa)}{\partial x_2} - \operatorname{div}(\kappa\nabla p) = 0, \quad p \geq 0, \quad \theta \in H(p), \tag{1-2}$$

where p is the pressure, θ the saturation, g denotes the gravity and $H(\cdot)$ denotes the multivalued Heaviside operator, in such a way that for positive pressure ($p > 0$) the porous media is fully saturated ($\theta = 1$) and $\theta \in [0, 1)$ when $p = 0$ in the non-saturated region. We introduce an artificial time variable; this procedure can be understood as a fixed-point iteration indexed by the dummy time variable. For time discretization, we consider a feature method based on the numerical approximation of the material (or total) derivative, which is a well-understood concept in continuum mechanics (see [13]).

In addition, a duality method is considered (see [10, 6]), which allows transforming the inequality with multivalued terms into an equation with single-valued terms with the help of Yosida's approximation (see [10, 23]). As in [9, 49], each iteration requires the spatial approximation of a resulting pressure equation after applying these techniques. The pressure equation mentioned here is posed in multiscale high-contrast porous media.

The main innovative achievement of the present work comes from the proposed method to address the spatial approximation of pressure in the case of no-scale separation in the heterogeneous porous medium. Solving the pressure equation with the resolution of the medium is not practical for this application, as is the case in various flow models in porous media involving multiple scales (see [31, 33, 25, 57, 36]). We propose to compute an efficient pressure approximation using the GMsFEM introduced in [28, 31, 15] and the references therein. The GMsFEM method provides numerical results that capture the solution behavior due to coefficient variations at fine resolution by solving linear systems with size proportional to the number of coarse blocks from a coarse resolution. The grid does not need to be adapted to the coefficient variations.

There we apply the GMsFEM methodology to the case of modeling the heterogeneous dam problem. However, as mentioned before, this method is used in partial differential equations models where the diffusion operator plays a leading role. For instance, another application studied in this document is a nonlinear parabolic problem on a multiscale high-contrast medium. A widely used model of diffusion is the following semilinear parabolic problem posed in a high-contrast multiscale media,

$$\begin{cases} \partial_t p - \operatorname{div}(\kappa(x)\nabla p) = f(p), & \text{in } \Omega \times I, \\ p = p_D, & \text{on } \partial\Omega \times I, \\ p(0, x) = \hat{p}(x), & \text{on } \Omega, \end{cases} \tag{1-3}$$

Here Ω is a two-dimensional convex domain with a boundary $\partial\Omega$ and $I = [0, T]$ is the time domain. The field $\kappa(x)$ is a multiscale high-contrast heterogeneous field. Additionally, p is an unknown pressure field satisfying the Dirichlet condition given by p_D and the initial condition given by $\hat{p}$. The constructions and methods developed here can be easily adapted to the three-dimensional domains.

Approximations of solutions of problem (1-3) and many other interesting questions have been considered in the literature. In particular we mention [24, 43, 2, 54] and references therein. We focus our discussion on the numerical computation of solutions of this problem. In the presence of high-contrast multiscale coefficients, classical methods for the numerical approximation of solutions need to be revisited due to the lack of robustness and efficiency, see [36, 37, 26, 27, 1]. In this thesis, we design robust numerical approximation procedures against the presence of multiscale variations and high-contrast in the coefficient κ. We call the attention to two important challenges in order to design efficient and robust numerical methods for equation (1-3):

1. In the presence of high-contrast multiscale coefficients, the spatial resolution needed to correctly approximate the solution of (1-3) (or its steady-state version) is related to the smallest scale at which we find variations of the coefficient κ. Additionally to the multiscale variations, the discontinuities and high-jumps of the coefficient bring additional difficulties to the numerical approximation of this time-dependent problem. Accuracy and efficiency can be negatively affected by solving large and ill-conditioned linear systems at each time step; See [28, 57, 1, 16].

2. The presence of high-contrast in the coefficients (even without complicated multiscale variations) reduces the stability region of time discretization methods such as Crank–Nicolson and similar time integrators; see Sections 5.3 and 5.4 below.

Let us mention first that challenge 1 above also affects time-independent problems. For time-independent problems classical multiscale methods provide good approximations only for moderated-to-low-contrast coefficients. However, Generalized Multiscale Finite Element Methods (GMsFEM) were designed to correctly handle problems with high-contrast in the coefficient where the main ingredient was to use local eigenvalue problems to construct appropriate coarse-mesh approximation spaces. For more details on the construction and analysis related to the approximation capabilities of the GMSFEM, see [28, 57, 1, 16, 58].

The GMsFEM has also been applied to time-dependent problems, linear and nonlinear parabolic, hyperbolic partial differential equations, sampling, and inverse problems. Sometimes, discretization or optimization iteration has to be added on top of the space approximation. We consider the case of time marching schemes where, in each iteration, a large ill-conditioned linear problem has to be solved. We mention the recent papers [19, 34, 20, 13, 2, 50, 46] where different time dependent problems have been considered within the GMsFEM framework. Unfortunately, the loss of stability (due to the high contrast in the coefficients) requires the reduction of the time step size (inversely proportional to the contrast in the coefficient), which ends up reducing the gain obtained by using the GMsFEM method to improve the overall computational time to obtain the final time solution. This brings us to face challenge 2 above, which is an important contribution of this thesis.

Our idea to gain stability and accuracy in time discretizations is to move to exponential integration; see for instance [44, 41, 3, 40, 52].
Exponential integration is a more efficient numerical method to overcome rigid problems and improve the accuracy of numerical computation. The main bottleneck of the calculations required by the exponential integrator is the calculation of matrix functions [39, 45], and this is even more critical for the finite element arrays associated with the problem (1-3) since these are huge and ill-conditioned sparse arrays, as mentioned before. In this section, we show that the function of the necessary matrices in the exponential integration can be well approximated using a GMsFEM approach, computed by projection onto the coarse scale space constructed using the GMsFEM approach. In our numerical experiments, we show that only by computing operators in approximate spaces can we advance significant time steps without losing stability and precision in the solution.

The document is organized as follows. Chapter 1 shows the finite element method (FEM) background and its convergence analysis. In chapter 2, we study the GMsFEM. In chapter 3, we show the first application, which is the problem of the dam on a high-contrast medium with a free boundary. Finally, in chapter 4, a parabolic problem is solved with the combination of the GMsFEM and the method of exponential integrators. For these last two chapters, we present numerical evidence of the good performance of the GMsFEM methodology. Finally, we present a future work based in [51], where we make a preliminary study on the free boundary problem.

2 Preliminaries

2.1 Space and problems

The mathematical treatment of several applications of the finite element method is based on the variational formulation of elliptic partial differential equations. The solutions of many important differential equations can be characterized by finding a minimum over an appropriate functional. The corresponding variational problems have solutions in spaces of specific functions such as Sobolev spaces. The numerical treatment involves minimization into appropriate finite-dimensional linear subspaces. The so-called finite element spaces are suitable for this finite-dimensional minimization, both from a practical and theoretical point of view. Here, we present a review of the main ingredients of the method. This section is based in [11, 47].

2.1.1 Sobolev spaces

Let Ω be an open subset of $\mathbb{R}^n$ with a piecewise smooth boundary. The Sobolev spaces, which play an important role in this document, are built on the function space $L^2(\Omega)$. Which consists of all functions, which are square integrable in the Lebesgue sense.

Definition 1. *Let Ω be a domain in $\mathbb{R}^n$ and let p be a positive real number. We denote by $L^p(\Omega)$ the space of functions u defined on Ω for which*

$$\int_\Omega |u(x)|^p \, dx < \infty.$$

We identify two functions u, $v \in L^2(\Omega)$ whenever $u(x) = v(x)$ for $x \in \Omega$, except in a subset of measure zero.

Definition 2. *Let $\Omega \subset \mathbb{R}^n$ be open and $1 \leq p \leq \infty$. The Sobolev space $W^{m,p}(\Omega)$ is defined by*

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega) : D^\alpha f \in L_p(\Omega), \text{ for all } \alpha \in \mathbb{N}^n : |\alpha| \leq m\},$$

where α is a multi-index of order $|\alpha| = k$ and $D^\alpha f = \frac{\partial^{|\alpha|}}{\partial x_1^{\alpha 1} \partial x_2^{\alpha 2} \ldots \partial x_n^{\alpha n}} f$ is the weak derivative of f. In particular, we write $H^m(\Omega) = W^{m,2}(\Omega)$.

We can define a scalar product on $H^m(\Omega)$ by

$$(u, v)_m = \sum_{|\alpha| \leq m} (D^\alpha u, D^\alpha v), \tag{2-1}$$

with the associated m-norm

$$\|u\|_m = \sqrt{(u,u)_m} = \sqrt{\sum_{|\alpha|\leq m} \|D^\alpha u\|^2_{L_2(\Omega)}}, \tag{2-2}$$

and seminorm

$$|u|_m = \sqrt{\sum_{|\alpha|=m} \|D^\alpha u\|^2_{L_2(\Omega)}}. \tag{2-3}$$

We define $H_0^m(\Omega)$ as the closure in $H^m(\Omega)$ of the space $C_c^\infty(\Omega)$ of infinitely differentiable compactly supported functions. Particularly, in the space $H_0^1(\Omega)$, the seminorm (2-3) is equivalent to the norm (2-2). This is due to the next result.

Theorem 1 (Poincaré-Friedrichs inequality, [11, Theorem 1.5])**.** *Suppose Ω is contained in an n-dimensional cube with side length s. Then*

$$\|v\|_0 \leq s|v|_1 \text{ for all } v \in H_0^1(\Omega). \tag{2-4}$$

Proof. Since C_0^∞ is dense in $H_0^1(\Omega)$, it suffices to establish the inequality for $v \in C_0^\infty$. We may assume that $\Omega \subset W = \{(x_1, x_2, ..., x_n) : 0 < x_i < s\}$. Then

$$v(x_1, x_2, ..., x_n) = v(0, x_2, ..., x_n) + \int_0^{x_1} \partial_1 v(t, x_2, ..., x_n) dt.$$

Given that $v \in C_0^\infty(\Omega)$, then the first term vanishes, and using the Cauchy-Schwarz inequality,

$$\begin{aligned} |v(x)|^2 &\leq \int_0^{x_1} 1^2 dt \int_0^{x_1} |\partial_1 v(t, x_2, ..., x_n)|^2 \, dt \\ &\leq s \int_0^s |\partial_1 v(t, x_2, ..., x_n)|^2 \, dt. \end{aligned}$$

Since the right-hand side is independent of x_1, we obtain

$$\int_0^s |v(x)|^2 \, dx_1 \leq s^2 \int_0^s |\partial_1 v(x)|^2 \, dx_1.$$

We integrate over the other coordinates to obtain

$$\int_W |v|^2 \, dx \leq s^2 \int_W |\partial_1 v|^2 \, dx \leq s^2 \, |v|_1^2 .$$

□

Remark 2. *The proof of the Poincare–Friedrichs inequality only requires zero boundary conditions on part of the boundary. If $\Gamma = \partial\Omega$ is piecewise smooth, the function vanishes on the part of the boundary ∂D, where ∂D is a set with positive $(n-1)$-dimensional measure.*

Moreover, if zero Dirichlet boundary conditions are on the whole boundary, without loss of generality, we can assume that Ω lies between two hyperplanes whose distance apart is s. See [35].

2.1.2 Variational formulations

This section shows the notation and main features of variational formulations for elliptic equations.

Theorem 3 (Characterization theorem, [11, Theorem 2.2])**.** *Let V be a linear space; suppose that*

$$a : V \times V \to \mathbb{R}$$

is a symmetric positive definite bilinear form[1]. In addition, let $\ell : V \to \mathbb{R}$ be a linear functional. Then, the functional

$$J(v) = \frac{1}{2}a(v, v) - \ell(v)$$

attains its minimum over V at u if and only if

$$a(u, v) = \ell(v) \text{ for all } v \in V. \tag{2-5}$$

Proof. For u, $v \in V$ and $t \in \mathbb{R}$, we have

$$\begin{aligned} J(u + tv) &= \frac{1}{2}a\left(u + tv, u + tv\right) - \ell(u + tv) \\ &= \frac{1}{2}\left(a\left(u, u\right) + ta\left(u, v\right) + ta\left(v, u\right) + t^2 a\left(v, v\right)\right) - \ell(u) - t\ell(v) \\ &= J(u) + t\left(a(u, v) - \ell(v)\right) + \frac{1}{2}t^2 a(v, v). \end{aligned} \tag{2-6}$$

If $u \in V$ satisfies (2-5), then (2-6) with $t = 1$ implies

$$\begin{aligned} J(u + v) &= J(u) + \frac{1}{2}a(v, v) \quad \text{for all } v \in V \\ &> J(u) \qquad \text{if } v \neq 0. \end{aligned} \tag{2-7}$$

Therefore, u is a unique minimal point. On the other side, if J has a minimum at u, then for every $v \in V$, the derivative of the function $t \to J(u + tv)$ must vanish at $t = 0$. By (2-6) the derivate is $a(u, v) = \langle \ell, v \rangle$.

□

Now, we associate Theorem 3 with a classic boundary-value problem such as the Poisson equation.

[1] A bilinear form a is definite positive if $a(v, v) > 0$ for all $v \in V$, $v \neq 0$.

Theorem 4 (Minimal property, [11, Theorem 2.3])**.** *Consider the following boundary value problem*

$$\begin{cases} -\Delta u + bu = f \ \text{in } \Omega, \\ u = 0 \ \text{on } \Gamma, \end{cases} \tag{2-8}$$

where Ω is a bounded open domain in the plane $\mathbb{R}^2$ with boundary Γ and b is a constant. Every solution of the problem (2-8) *is a solution of the variational problem,*

$$J(v) = \frac{1}{2}\int_\Omega \nabla v^2 + bv^2 \ dx - \int_\Omega fv \ dx \to \arg\min, \tag{2-9}$$

among all functions in $C^2(\Omega) \cap C^0(\bar{\Omega})$ with zero boundary values.

Proof. Let's start with the two-dimensional divergence theorem

$$\int_\Omega \mathrm{div} A \ dx = \int_\Gamma A \cdot \mathbf{n} \ ds,$$

where $A = (A_1, A_2)$ is a sufficiently smooth vector function defined on Ω, $\mathbf{n} = (\mathbf{n}_1, \mathbf{n}_2)$ is the outward unit normal to Γ, and dx represents the elements of Ω.

$$\mathrm{div} A = \frac{\partial A_1}{\partial x_1} + \frac{\partial A_2}{\partial x_2}.$$

If we apply the divergence theorem to $A_1 = (vw, 0)$ and $A_2 = (0, vw)$, we obtain

$$\int_\Omega \frac{\partial v}{\partial x_i} w dx + \int_\Omega v \frac{\partial w}{\partial x_i} dx = \int_\Gamma vw\mathbf{n}_i ds \quad i = 1, 2. \tag{2-10}$$

We can use this (Green's formula) to rewrite equation (2-8) as:

$$\begin{aligned} \int_\Omega \nabla v \nabla w dx &= \int_\Omega \left(\frac{\partial v}{\partial x_1}\frac{\partial w}{\partial x_1} + \frac{\partial v}{\partial x_2}\frac{\partial w}{\partial x_2} \right) dx \\ &= \int_\Gamma \left(v\frac{\partial w}{\partial x_1}\mathbf{n}_1 + v\frac{\partial w}{\partial x_2}\mathbf{n}_2 \right) ds - \int_\Omega v \left(\frac{\partial^2 w}{\partial x_1^2} + \frac{\partial^2 w}{\partial x_2^2} \right) dx \\ &= \int_\Gamma v\frac{\partial w}{\partial \mathbf{n}} ds - \int_\Omega v\Delta w dx, \end{aligned} \tag{2-11}$$

where

$$\frac{\partial w}{\partial \mathbf{n}} = \frac{\partial w}{\partial x_1}\mathbf{n}_1 + \frac{\partial w}{\partial x_2}\mathbf{n}_2.$$

If u is a solution of the problem (2-8), then u is also a solution of the following variational problem: Find $u \in V$ such that

$$a(u, v) = \ell(v) \text{ for all } v \in V, \tag{2-12}$$

where

$$\begin{aligned} a(v,w) &= \int_\Omega \nabla v \nabla w + bvw \; dx \\ \ell(v) &= \int_\Omega fv \; dx. \\ V &= \left\{ v : v \in C^1 \text{ and } v = 0 \text{ on } \Gamma \right\} \end{aligned} \tag{2-13}$$

Therefore,

$$\begin{aligned} a(v,w) - \ell(v) &= \int_\Omega \nabla v \nabla w + bvw \; dx - \int_\Omega fv \; dx \\ &= \int_\Omega v\Delta w + bvw \; dx - \int_\Omega fv \; dx \\ &= \int_\Omega v \left(\Delta w + bw - f \right) \; dx = 0. \end{aligned} \tag{2-14}$$

This last step is derived from equation (2-8), and if u is a classical solution[2], the characterization theorem implies the minimality property.

□

Definition 3. *Let H ba a Hilbert space. A bilinear form $a : H \times H \to \mathbb{R}$ is called continuous, provided there exists $C > 0$ such that*

$$|a(u,v)| \leq C\|u\|\|v\| \text{ for all } u, v \in H.$$

A symmetric continuous bilinear form a is called H-elliptic or for short elliptic or coercive, if for some $\alpha > 0$

$$a\left(v,v\right) \geq \alpha \|v\|^2 \text{ for all } v \in V, \tag{2-15}$$

and if a is an H-elliptic bilinear form, then it induces a norm via

$$\|v\|_a = \sqrt{a(v,v)}. \tag{2-16}$$

An important theorem in the theory of the Finite Elements Method (FEM) is the Lax-Milgram theorem, which directs to the existence of the solution of the equation (2-5).

Theorem 5 (Lax-Milgram theorem, [11, Theorem 2.5])**.** *Let V be a closed convex set in a Hilbert space H, and let $a : H \times H \to \mathbb{R}$ be an elliptic bilinear form. Then, there exists a unique v such that it minimizes the function*

$$J(v) = \frac{1}{2} a(v,v) - \ell(v).$$

[2] A function that satisfies a given second-order partial differential equation and assumes prescribed boundary values is called a classical solution.

Proof. Given that a is elliptic, we have the following result:

$$\begin{aligned} J(v) &= \frac{1}{2}a(v,v) - \ell(v), \\ &\geq \frac{1}{2}\alpha\|v\|^2 - \|\ell\|\|v\|, \\ &= \frac{1}{2\alpha}(\alpha\|v\| - \|\ell\|)^2 - \frac{\|\ell\|}{2\alpha}, \\ &\geq -\frac{\ell}{2\alpha}, \end{aligned} \tag{2-17}$$

and $J(v)$ is bounded below, then these exists c_1 such that is an infimum of the $J(v)$ and a minimizing sequence[3] $\{v_n\}$. Now, we prove that V_n is a Cauchy sequence in H,

$$\begin{aligned} \alpha\|v_n - v_m\|^2 &\leq a(v_n - v_m, v_n - v_m), \\ &= a(v_n, v_n) - 2a(v_n, v_m) + a(v_m, v_m), \\ &= 2a(v_n, v_n) + 2a(v_m, v_m) - \left(a(v_n, v_n) + 2a(v_n, v_m) + a(v_m, v_m)\right), \\ &= 2a(v_n, v_n) + 2a(v_m, v_m) - a(v_n + v_m, v_n + v_m), \\ &= 4J(v_n) + 4J(v_m) - 8J\left(\frac{v_n + v_m}{2}\right), \\ &\leq 4J(v_n) + 4J(v_m) - 8c_1, \end{aligned}$$

given that V is convex then $\frac{1}{2}(v_n + v_m) \in V$. Since $J(v_n)$, $J(v_m) \to c_1$, implies $\|v_n - v_m\| \to 0$ for $n, m \to \infty$. Therefore (v_n) is a Cauchy sequence in H and $u = \lim_{n\to\infty} v_n$ exists. Given that V is closed, then $u \in V$.
Now, we show that the solution is unique. Suppose u_1 and u_2 are both solutions. Clearly u_1, u_2, u_1, u_2,... is a minimizing sequence. As we saw above, every minimizing sequence is a Cauchy sequence. This is only possible if $u_1 = u_2$.

□

With the above preparations, we can now make the concept of a solution to the boundary-value problem more precise.

Definition 4. *A function $u \in H_0^1(\Omega)$ is called a weak solution of the second-order elliptic boundary-value problem*

$$\begin{cases} Lu = -\mathrm{div}(\kappa(x)\nabla u) + bu = f \ \text{in } \Omega, \\ u = 0 \ \text{on } \Gamma, \end{cases} \tag{2-18}$$

with homogeneous Dirichlet boundary conditions, provided that

[3] A sequence of elements y_n is called minimizing on a set M, if the corresponding sequence of function values $j(y_n)$ tends to the greatest lower bound of j on M, that is, $\lim_{n\to\infty} j(y_n) = \inf_{y\in M} j(y)$.

$$a(u,v) = \ell(v), \text{ for all } v \in H_0^1(\Omega). \tag{2-19}$$

where

$$a(u,v) = \int_\Omega \kappa(x)\nabla u \nabla v dx + \int_\Omega buv dx \quad \text{and} \quad \ell(v) = \int_\Omega fv dx. \tag{2-20}$$

Theorem 6 (Existence theorem. ([11], Theorem 2.9))**.** *Let L be a second-order uniformly elliptic partial differential operator. The Dirichlet problem* (2-18) *always has a weak solution in $H_0^1(\Omega)$. It is a minimum of the variational problem*

$$\frac{1}{2}a(v,v) - \ell(v) \to \min!$$

over $H_0^1(\Omega)$, where a and ℓ are defined in (2-20)*.*

2.2 The Finite Element Method (FEM)

The main idea of the FEM to solve problem (2-8) is to take a finite-dimensional subspace of $H^m(\Omega)$ or $H_0^m(\Omega)$, and in this subspace, we want to minimize the functional J defined in (3). The standard notation for the subspace is V^h, where h is a parameter of the discretization such that if $h \to 0$, the approximation solution will converge to the proper solution.
The solution to the variational problem

$$J(v) = \frac{1}{2}a(v,v) - \langle \ell, v\rangle \to \min_{V^h}, \tag{2-21}$$

in the subspace V^h can be computed using Theorem 3; in particular, we want $u^h \in V^h$ such that

$$a(u^h, v) = \langle \ell, v\rangle \text{ for all } v \in V^h. \tag{2-22}$$

Suppose $\{\varphi_1, \varphi_2, ..., \varphi_N\}$ is a basis for V^h. Then problem (2-22) is equivalent to

$$a(u^h, \varphi_i) = \langle \ell, \varphi_i\rangle \quad i = 1, 2, ..., N,$$

where

$$u^h = \sum_{k=1}^{N} \alpha_k \varphi_k, \tag{2-23}$$

Replacing (2-23), in (2-22) we get

$$\sum_{k=1}^{N} \alpha_k a(\varphi_k, \varphi_i) = \langle \ell, \varphi_i\rangle \quad i = 1, 2, ..., N, \tag{2-24}$$

which we can write in matrix form as

$$A\alpha = b, \tag{2-25}$$

where $A_{i,k} = a(\varphi_k, \varphi_i)$ and $b_i = \langle \ell, \varphi_i \rangle$. The matrix A is called the stiffness matrix.

Lemma 1 (Céa's Lemma, [11, Lemma 4.2])**.** *Suppose that the bilinear form a is continuous and $V-$elliptic with $H_0^m(\Omega) \subset V \subset H^m(\Omega)$. In addition, suppose that u and u^h are the solutions of the variational problem in V and $V^h \subset V$, respectively. Then*

$$\left\|u - u^h\right\|_m \leq \frac{C}{\alpha} \inf_{v^h \in V^h} \left\|u - v^h\right\|. \tag{2-26}$$

Proof. Given that u and u^h are solutions of the variational problem in V and V_h, respectively, we have that

$$\begin{aligned} a(u, v) &= \langle \ell, v \rangle \text{ for all } v \in V, \\ a(u^h, v) &= \langle \ell, v \rangle \text{ for all } v \in V^h. \end{aligned}$$

Since $V^h \subset V$, we have

$$a(u - u^h, v) = 0 \text{ for all } v \in V^h. \tag{2-27}$$

Now, let $v^h \in V^h$. Since $u^h - v^h \in V^h$ we obtain that $a(u - u^h, v^h - u^h) = 0$, and

$$\begin{aligned} \alpha \left\|u - u^h\right\|_m^2 &\leq a(u - u^h, u - u^h) \\ &= a(u - u^h, u - v^h) + a(u - u^h, v^h - u^h) \\ &\leq C \left\|u - u^h\right\|_m \left\|u - v^h\right\|_m. \end{aligned} \tag{2-28}$$

After dividing by $\alpha \left\|u - u^h\right\|_m$, we obtain the desired result.

□

With Cea's Lema and choosing an interpolation $u = I^h u \in V^h$, we can estimate the interpolation error $\|u - I^h u\|_V$ and obtain an estimate of the error $\|u - u^h\|_V$.

2.2.1 Error estimation for an elliptic problems

Let us now make a triangulation of Ω by subdividing Ω into a set $\mathcal{T}^h = \{K_1, ..., K_m\}$ of non-overlapping triangles or quadrilaterals K_i, where

1. $\overline{\Omega} = \bigcup_{K \in \mathcal{T}^h} K$.
2. If $K_i \cap K_j$ consists of exactly one point, then it is a common vertex of K_i and K_j.

3. If for $i \neq j$, $K_i \cap K_j$ consists of more than one point, then $K_i \cap K_j$ is a common edge of K_i and K_j.

With mesh parameter

$$h = \max_{K \in \mathcal{T}_h} (\operatorname{diam}(K)) .$$

From Cea's Lemma, we have

$$\|u - u^h\| \leq \frac{C}{\alpha} \|u - v\| \quad \text{for all } v \in V^h.$$

If we choose an interpolate $v = I^h u \in V^h$ and estimate the interpolation error $\|u - I^h u\|_m$, we can bound the error $\|u - u^h\|_m$.
Sean N_i, $i = 1, 2, 3, ..., M$, los nodos de $\mathcal{T}^h$. Asumiendo que $u \in H_0^1(\Omega) \cap H^2(\Omega)$, definimos el interpolante $I^h u \in V^h$ usando el enfoque descrito en [47].

$$I^h u(N_j) = u(N_j), \qquad j = 1, 2, 3, ..., M.$$

Therefore, the function I^h represents the piecewise linear interpolation that coincides with the function u at the nodes of the finite element mesh $\mathcal{T}^h$.
We have the following result.

Theorem 7 ([47, Theorem 2.3]). *Let $K \in \mathcal{T}^h$ be a triangle with vertices a_i, $i = 1, 2, 3$. Given $v \in H_0^1(K)$, let the interpolant $Iv \in P_1(K)$ be defined by*

$$Iv(a_i) = v(a_i), \qquad i = 1, 2, 3.$$

Then

$$\|u - Iv\|_{L_\infty(K)} \leq 2h_k^2 \max_{|\alpha|=2} \|D^\alpha v\|_{L_\infty(K)}.$$

The preceding theorem provides estimates of the interpolation error using the L_∞-norm. However, for estimating $\left\|u - I^h u\right\|_{H^1(\Omega)}$ with the L_2-norm, we employ the following theorem.

Theorem 8 ([47, Theorem 4.2]). *Under the assumption of Theorem 7, there exists a constant $C > 0$ such that*

$$\|u - Iv\|_{L_2(K)} \leq C h_k^2 |v|_{H^2(K)}.$$

Now, for an error estimate in the $L_2(\Omega)$-norm, we have seen that if we apply the FEM with the space V^h, with Ω a polygonal domain, then we have the following estimate for the error $u - u^h$ in the $H^1(\Omega)$-norm,

$$\|u - u^h\|_{H^1(\Omega)} \leq C h |u|_{H^2(\Omega)}.$$

2.2.2 FEM for Poisson equation

We consider the following boundary value problem for the Poisson equation,

$$\begin{cases} -\Delta u = f \text{ in } \Omega \\ \quad u = 0 \text{ on } \Gamma, \end{cases} \tag{2-29}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded open domain with boundary Γ. By Theorem 4, we have the following variational formulation: Find $u \in V$ such that

$$\begin{cases} a(u,v) = \langle f, v\rangle \text{ for all } v \in V \\ \quad u(x) = 0 \text{ for all } x \in \partial\Omega, \end{cases} \tag{2-30}$$

where

$$a(u,v) = \int_\Omega \nabla u \nabla v \quad dx = \int_\Omega \frac{\partial u}{\partial x_1}\frac{\partial v}{\partial x_1} + \frac{\partial u}{\partial x_2}\frac{\partial v}{\partial x_2} \quad dx,$$

and $V = H_0^1(\Omega)$. We now define V^h as follows

$$V^h = \left\{ v \in C(\bar{\Omega}) : v|_K \text{ is linear for } K \in \mathcal{T}^h,\ v = 0 \text{ on } \Gamma \right\}.$$

The space V^h consists of all continuous linear functions on each triangle K. Also, we define the following subspace of V^h,

$$V_0^h(\Omega) = \left\{ v \in V^h(\omega) : v|_\Gamma = 0 \right\}. \tag{2-31}$$

The basis functions $\{\varphi_j\} \in V^h$, for $j = 1, 2, .., N_v$ are defined by

$$\varphi_j(x_i) = \delta_{i,j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

with $i, j = 1, 2, ..., N_v$ and where x_i is a node of the triangulation $\mathcal{T}^h$.
The Galerkin formulation of (2-30) is to find $u \in V_0^h(\Omega)$ such that

$$\begin{cases} a(u,v) = \ell(v), & \text{for all } v \in V_0^h(\Omega), \\ u(x) = 0, & \text{for all } x \in \partial\Omega, \end{cases} \tag{2-32}$$

where $\ell(v) = \langle f, v\rangle$. We consider the following representation for the solution of (2-32)

$$u^h(x) = \sum_{i=1}^{N_v} \alpha_i \varphi_i(x), \tag{2-33}$$

where α_i are constants that we can determine. Using (2-32) taking $v = \varphi_j$ for $j = 1, 2, 3, ..., N_v$, we have

$$\sum_{i=1}^{N_v} \alpha_i a(\varphi_i, \varphi_j) = \ell(\varphi_j), \quad \text{for all } v \in V_0^h(\Omega). \tag{2-34}$$

The equation (2-34) can be equivalently written in matrix form as follows

$$Au = b, \tag{2-35}$$

where

$$u^T Av = \int_\Omega \kappa \nabla u \nabla v \quad \text{and} \quad v^T b = \int_\Omega fv, \qquad \text{for all } u, v \in V_0^h(\Omega). \tag{2-36}$$

Example

We consider the problem:

$$\begin{cases} -\text{div}(\nabla p) = 5\pi^2 \sin(2\pi x_1)\sin(\pi x_2), & \text{in} \quad \Omega = [0,1]^2, \\ p(x_1, x_2) = 0, & \text{on} \quad \partial\Omega, \end{cases} \tag{2-37}$$

where the exact solution is $p(x_1, x_2) = \sin(2\pi x_1)\sin(\pi x_2)$.

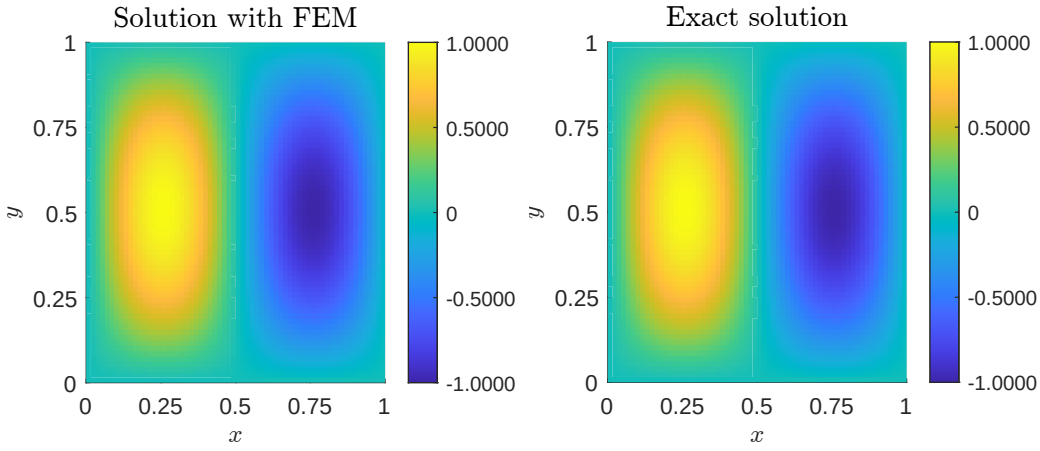

Figure 2-1: Numerical solution of problem (2-37) by FEM, with a mesh of 100×100.

The relative error is computed with the following norms:

$$\|u\|_{L_2}^2 = \int_\Omega u^2, \qquad \|u\|_{H^1}^2 = \int_\Omega (\nabla u)^2 + \int_\Omega u^2. \tag{2-38}$$

h	L_2 % Error	H^1 % Error
2^{-3}	$50.8 \cdot 10^{-4}$	$5.7 \cdot 10^{-3}$
2^{-4}	$12.1 \cdot 10^{-4}$	$1.7 \cdot 10^{-3}$
2^{-5}	$3.3 \cdot 10^{-4}$	$0.5 \cdot 10^{-3}$
2^{-6}	$0.7 \cdot 10^{-4}$	$0.1 \cdot 10^{-3}$

Table 2-1: The weighted L_2 and H^1 errors of the solution of problem (2-37) by FEM.

2.2.3 A problem with Neumann boundary condition

In this example, we study the following Neumann problem

$$\begin{cases} -\Delta u + u = f, & \text{in } \Omega, \\ \dfrac{\partial u}{\partial n} = g, & \text{on } \Gamma, \end{cases} \tag{2-39}$$

where $\frac{\partial}{\partial n}$ is the outward normal derivative to Γ. The variational formulation of the problem (2-39) is: Find $u \in H^1(\Omega)$ such that

$$a(u,v) = \ell(v) + \langle g, v\rangle \qquad \text{for all } v \in H^1(\Omega), \tag{2-40}$$

where $g = \frac{\partial u}{\partial n}$,

$$a(u,v) = \int_\Omega \nabla u \nabla v + uv dx,\ \ell(v) = \int_\Omega f v dx \text{ and } \langle g, v\rangle = \int_\Gamma \frac{\partial u}{\partial n} v ds. \tag{2-41}$$

This result is obtained using the method developed in the proof of Theorem (4).
We use the representation (2-33), and we obtain the following Galerking formulation: Find $u \in V^h$ such that

$$a(u^h, v) = \ell(v) + \langle g, v\rangle \quad \text{for all} \quad v \in V^h(\Omega), \tag{2-42}$$

and the matrix form

$$Au = b, \tag{2-43}$$

where

$$u^T A v = \int_\Omega \kappa \nabla u \nabla v + uv \quad \text{and} \quad v^T b = \int_\Omega fv + \int_\Gamma g v ds, \quad \text{for all } u, v \in V^h(\Omega). \tag{2-44}$$

2.2.4 FEM for the parabolic problem

An important problem to study is the determination of heat conduction in a body with conductivity κ in a region $\Omega \subset \mathbb{R}^d$. The associated equation is

$$\begin{cases} \dfrac{\partial u}{\partial t} - \operatorname{div}(\kappa \nabla u) = f, & \text{in } \Omega \times I, \\ u = 0, & \text{on } \Gamma \times I, \\ u(x,0) = u_0(x), & x \in \Omega, \end{cases} \tag{2-45}$$

where I is an interval contain in $\mathbb{R}$ and γ is the boundary of domain Ω. We obtain the following variational formulation by multiplying by a test function $v \in C_0^\infty(\Omega)$ and integrating by parts

$$\left(\frac{\partial u}{\partial t}, v\right) + a(u,v) = \ell(v) \quad \text{for all } v \in C_0^1(\Omega), \tag{2-46}$$

where

$$\left(\frac{\partial u}{\partial t}, v\right) = \int_\Omega \frac{\partial u}{\partial t} v dx, \quad a(u,v) = \int_\Omega \nabla u \nabla v dx \quad \text{and} \quad \ell(v) = \int_\Omega f v dx. \tag{2-47}$$

Space discretization

Let $\mathcal{T}^h$ be a triangulation of Ω, and V_0^h be a linear function space associated with $\mathcal{T}^h$. We define the semi-discrete approximation to the heat equation as

$$\int_\Omega \frac{\partial u^h}{\partial t} v dx + \int_\Omega \nabla u^h \nabla v dx = \int_\Omega f v dx \quad \text{for all } v \in V_0^h, \tag{2-48}$$

subject to the initial condition $u^h(x,0) = I^h(u_0)$.

Temporal discretization using Euler method

For the temporal discretization, let $u \in C^2((0,T)\times(a,b))$. First, we obtain the Taylor series of u; this is

$$u(x,t+k) = u(x,t) + k\partial_t u(x,t) + k^2 r(x,t),$$

where $r(x,t)$ is the residue such that is bounded and uniform for t and x. We then obtain

$$\partial_t u(x,t) = \frac{u(x,t+k) - u(x,t)}{k} + kr(x,t). \tag{2-49}$$

and therefore we can approximate $\partial_t(u)$ by

$$\partial_t u(x,t) \approx \frac{u(x,t+k) - u(x,t)}{k}.$$

We denote by $u^{n,h} \in V^h$ an approximation for the solution of the heat equation in the time $t = nk$. We take the weak formulation of the heat equation (2-48), and using the approximation of $\partial_t u$ described before; we have that

$$\int_\Omega \frac{u^{n+1,k} - u^{n,h}}{k} v dx + \int_\Omega \nabla u^{n,h} \nabla v dx = \int_\Omega f v dx \quad \text{for all } v \in V_0^h. \tag{2-50}$$

Given that $u^{n,h} \in V_0^h$, there exist α_i constants, such that

$$u^{n,h}(x) = \sum_{i=1}^{N_0^h} \alpha_i^n \phi_i(x). \tag{2-51}$$

If we sustitute (2-51) in (2-50), we obtain

$$\sum_{i=1}^{N_0^h}\left(\frac{\alpha_i^{n+1}-\alpha_i^n}{k}\right)\int_\Omega \phi_i(x)\phi_j(x)+\sum_{i=1}^{N_0^h}\alpha_i^n\int_\Omega \nabla\phi_i^{n,h}\nabla\phi_j dx=\int_\Omega f\phi_j dx, \tag{2-52}$$

and in matrix form,

$$\frac{\alpha^{n+1}-\alpha^n}{k}M+\alpha^n A=b. \tag{2-53}$$

Example

We consider the problem:

$$\begin{cases} \partial_t p-\operatorname{div}(\nabla p)=(5\pi^2-1)e^{-t}\sin(2\pi x_1)\sin(\pi x_2), & \text{in} \quad \Omega=[0,1]^2,\\ p(0,x_1,x_2)=\sin(2\pi x_1)\sin(\pi x_2), & \\ p(t,x_1,x_2)=0, & \text{on} \quad \partial\Omega, \end{cases} \tag{2-54}$$

where the exact solution is $p(x_1,x_2,t)=e^{-t}\sin(2\pi x_1)\sin(\pi x_2)$. For the solution, we use FEM with the Finite difference method.

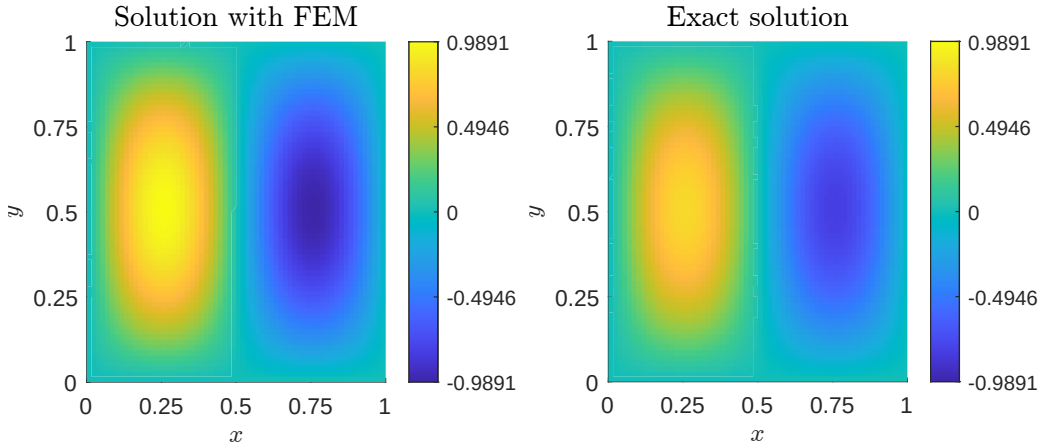

Figure 2-2: Numerical solution of problem (2-54) by FEM, with a mesh of 100×100 and time of $t=0.2$.

The solutions are computed until the time $T=0.2$, and relative errors in L_2 and H^1 described in (2-38) are show in Table **2-2**.

h	L_2 % Error	H^1 % Error
2^{-3}	57.52	54.55
2^{-4}	19.25	18.65
2^{-5}	5.20	5.12
2^{-6}	1.32	1.31

Table 2-2: The weighted L_2 and H^1 errors between the exact solution with time $t=0.2$.

2.3 Yosida operator

This section is based on [55]. We also drew upon results from [21, 22, 23, 38]. We consider H to be a Hilbert space equipped with the inner product denoted by $(\cdot, \cdot)$ and the associated norm $\|\cdot\|$. First, we present some definitions that are necessary for our study.

Definition 5. *Let G be an operator that maps H to the set of subsets of H, denoted by $\mathcal{P}(H)$.*

Definition 6. *An operator $G : D(G) \subset H \to \mathcal{P}(H)$ is monotone if it satisfies that for all $w_1 \in G(v_1)$ and for all $w_2 \in G(v_2)$,*

$$(w_1 - w_2, v_1 - v_2) \geq 0.$$

The domain of G is given by

$$D(G) = \{x \in H : G(x) \in H\}. \tag{2-55}$$

The range of G is given by

$$\text{rang}(G) = \bigcup_{x \in D(G)} G(x). \tag{2-56}$$

The graph of operator $G \in H \times H$ given by

$$\{(x, y) \in H \times H : y \in G(x)\}. \tag{2-57}$$

We define the following order relation in terms of the graphs inclusion:

$$G_1 \subset G_2 \Leftrightarrow G_1(x) \subset G_2(x) \quad \text{for all} \quad x \in H, \tag{2-58}$$

Definition 7. *The operator $G : H \to \mathcal{P}(H)$ is maximal monotone if and only if G is monotone and there is no monotone mapping $G_1 : H \to \mathcal{P}(H)$ such that $G \subset G_1$.*

The Yosida approximation is essential for transforming an multivalue operator to an unival-ued operator.

Definition 8. *Let $G : H \to \mathcal{P}(H)$ be a given maximal monotone operator, and $\lambda \geq 0$. The resolvent of G with parameter λ is the operator from H to H defined by*

$$J_\lambda = (I + \lambda G)^{-1}. \tag{2-59}$$

Definition 9. *Let $G : H \to \mathcal{P}(H)$ be a maximal monotone operator and $\lambda > 0$. The Yosida approximation of G with parameter λ is the unvalued operator defined by*

$$G_\lambda = \frac{I - J_\lambda}{\lambda}. \tag{2-60}$$

Proposition 1. *Let G be a maximal monotone operator. Then, for real numbers λ and ω satisfying $\lambda\omega < 1$, $\lambda > 0$, and $\omega \geq 0$, and for any $f \in H$, there exists $y \in H$ such that $f \in (1 - \lambda\omega)y + \lambda G(y)$.*

With this proposition, we can establish that any element in H can be expressed as a convex linear combination of an element y and its application through G. The following lemma plays a crucial role in justifying the duality algorithm proposed to solve the problem addressed in this thesis.

Lemma 2. *Let $G : H \to \mathcal{P}(H)$ be a maximal monotone operator in H. The following statements are equivalent*

1. $u \in G(y) - wy$,
2. $u = G_\lambda^\omega(y + \lambda u)$, *for* $\lambda > 0$,

Here, G_λ^ω denotes the Yosida approximation with real parameter λ of the operator $G - \omega I$.

If we denote by J_λ the resolvent operator of G and by J_λ^ω the resolvent operator of $G - \omega I$, i.e.,

$$J_\lambda = (I + \lambda G)^{-1}, \qquad J_\lambda^\omega = (I + \lambda(G - \omega I))^{-1}. \tag{2-61}$$

We have the following equality

$$\begin{aligned} J_\lambda^\omega &= (I + \lambda(G - \omega I))^{-1} \\ &= (I + \lambda G - \lambda\omega I)^{-1} \\ &= ((1 - \lambda\omega)I + \lambda G)^{-1} \\ &= (1 - \lambda\omega)^{-1} \left(I + \frac{\lambda}{1 - \lambda\omega} G\right)^{-1} \\ &= \frac{1}{1 - \lambda\omega} J_{\frac{\lambda}{1-\lambda\omega}}. \end{aligned}$$

By the Yosida approximation of $G - \omega I$, we obtain

$$\begin{aligned} G_\lambda^\omega &= \frac{I - J_\lambda^\omega}{\lambda} \\ &= \frac{I - \frac{1}{1-\lambda\omega} J_{\frac{\lambda}{1-\lambda\omega}}}{\lambda} \\ &= \frac{(1 - \lambda\omega)\, I - J_{\frac{\lambda}{1-\lambda\omega}}}{\lambda\,(1 - \lambda\omega)} \\ &= \frac{I - J_{\frac{\lambda}{1-\lambda\omega}}}{\lambda\,(1 - \lambda\omega)} - \frac{\omega I}{1 - \lambda\omega} \\ &= \frac{1}{(1 - \lambda\omega)^2} G_{\frac{\lambda}{1-\lambda\omega}} - \frac{\omega}{1 - \lambda\omega} I, \end{aligned} \tag{2-62}$$

where $G_{\frac{\lambda}{1-\lambda\omega}}$ is the Yosida approximation of G, with parameter $\frac{\lambda}{1-\lambda\omega}$.

Two important operators for our investigation are the Heaviside operator and the subdifferential operator associated with the indicatrix function, as well as their respective Yosida approximations.

2.3.1 Yosida approximation for the Heaviside operator

An important example in our study is the Heaviside operator:

Definition 10 (The Heaviside operator)**.** *Let $He : H \to \mathcal{P}(H)$ be an operator defined as*

$$He(u) = \begin{cases} 0, & \text{if } u < 0, \\ [0,1], & \text{if } u = 0, \\ 1, & \text{if } u > 0, \end{cases} \tag{2-63}$$

which is known as the Heaviside operator.

If we take $u_1 \leq u_2$ then for $v_1 \in He(u_1)$, $v_2 \in He(u_2)$, we have,

$$\begin{aligned}(v_1 - v_2, u_1 - u_2) &= \int_\Omega (v_1(x) - v_2(x))(u_1(x) - u_2(x))dx \\ &= \int_{\{f_1>0\}\cup\{f_2>0\}} f_1(x) - f_2(x)dx \\ &\geq 0.\end{aligned}$$

The resolvent of the Heaviside operator is

$$J_\lambda = (I + \lambda He(u))^{-1} = \begin{cases} u, & \text{if } \quad u < 0, \\ 0, & \text{if } \quad 0 \leq u \leq \lambda, \\ u - \lambda, & \text{if } \quad u > \lambda, \end{cases} \tag{2-64}$$

and the Yosida approximation is

$$G_\lambda(u) = \begin{cases} 0, & \text{if } \quad u < 0, \\ \frac{u}{\lambda}, & \text{if } \quad 0 \leq u \leq \lambda, \\ 1, & \text{if } \quad u > \lambda. \end{cases} \tag{2-65}$$

For clarification and illustration, let's consider the Heaviside function $He : \mathbb{R} \to \mathcal{P}(\mathbb{R})$ (see Figure **2-3**), defined by

$$He(t) = \begin{cases} 0, & \text{if} \quad t < 0, \\ [0,1], & \text{if} \quad t = 0, \\ 1, & \text{if} \quad t > 0. \end{cases} \tag{2-66}$$

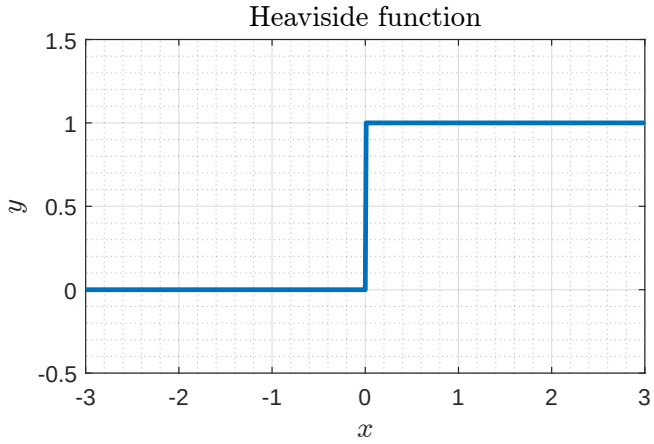

Figure 2-3: Heaviside operator $He : \mathbb{R} \to \mathcal{P}(\mathbb{R})$, this is called the Heaviside function.

see also Figures **2-4** and **2-5** for an illustration of the resolvent and Yosida approximation of the Heaviside function.

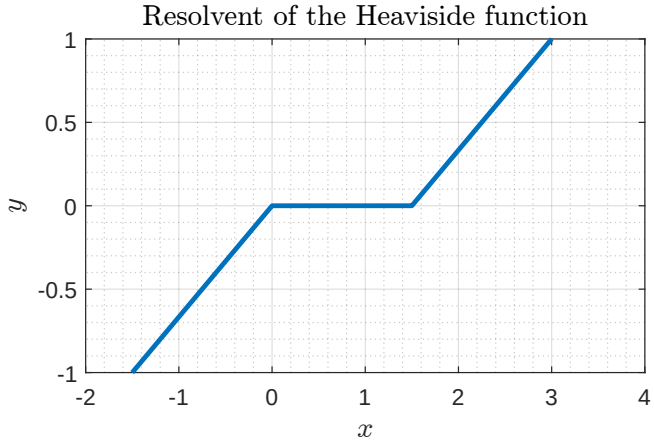

Figure 2-4: Resolvent of the Heaviside function.

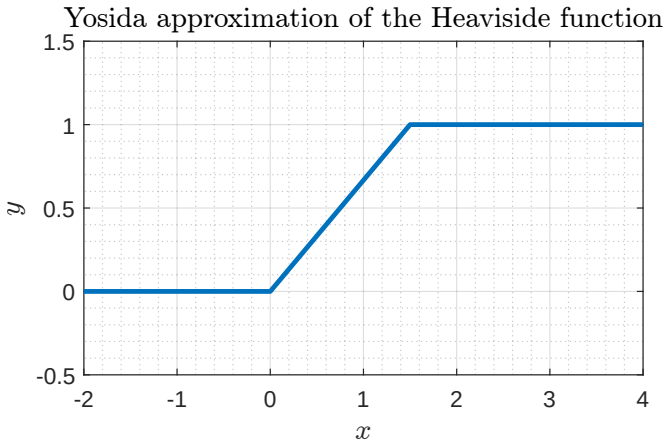

Figure 2-5: Yosida approximation of the Heaviside function.

In general the Yosida approximation for the Heaviside operator for $\lambda\omega = 1/2$ is

$$H^{\omega}_{\frac{1}{2\omega}}(\theta(x)) = \begin{cases} -2\omega\theta(x), & \theta(x) < 0, \\ 2\omega\theta(x), & 0 \leq \theta(x) \leq \frac{1}{2\omega}, \\ 2(1-\omega\theta(x)), & \theta(x) > \frac{1}{2\omega}. \end{cases}$$

2.3.2 Yosida approximation of the subdifferential of the indicatrix function

We consider functions ϕ defined on a set H with values in $(-\infty, \infty]$. The epigraph of φ is the set

$$\text{epi } \varphi = \{(x, \lambda) \in H \times \mathbb{R} : \varphi(x) \leq \lambda\}. \tag{2-67}$$

We recall the following definition.

Definition 11. *A function $\varphi : H \to (-\infty, +\infty]$ is said to be lower semicontinuous (l.s.c.) if for every $\lambda \in \mathbb{R}$ the set*

$$\{x \in H : \varphi(x) \leq \lambda\}, \tag{2-68}$$

is closed.

A definition important for this work is the subdifferential.

Definition 12 (The subgradient and subdifferential)**.** *Let $\varphi : H \to \overline{\mathbb{R}}$ be a functional on the Hilbert space H. Then u in H' is called subgradient of φ at v if and only if $\varphi(v) \neq \pm\infty$ and*

$$\varphi(w) \geq \varphi(v) + (u, w - v) \quad \text{for all } w \in H, \tag{2-69}$$

holds.
The set of all subgradients of φ at v is called subdifferential $\partial\varphi(u)$.

The indicatrix function of a subset V of a Hilbert space H maps elements of the subset V to one and all other elements to infinity. In other words,

$$I_V(v) = \begin{cases} 1 & v \in V, \\ +\infty & v \notin V. \end{cases} \tag{2-70}$$

We have the following assertions (see [23]):

- V is convex if and only if I_V is convex in H.
- V is closed if and only if I_V is lower semicontinuous in H.

An important property of the subdifferential of the indicatrix function of the convex set V is given by the following theorem

Theorem 9 ([53, Theorem A])**.** *If φ is a lower semicontinuous proper convex function on H, then $\partial\varphi$ is a maximal monotone operator.*

An important definition is the concept of orthogonal projection onto a convex set. Let's consider the minimization problem

$$\min_{v \in V} \|v - u\|, \tag{2-71}$$

which seeks points v in V that are at the least distance from u. We denote the unique solution to this minimization problem as $v = P_V u$ (see [23, Proposition 46.5]).
An important property is as follows: for a closed, convex, and non-empty set V in H, $v = P_V u$ if and only if

$$(v - u, w - v) \geq 0, \qquad \text{for all } w \in V. \tag{2-72}$$

If $(I_V)^{\omega}_{\lambda}$ denotes the Yosida approximation of the operator $\partial I_V - \omega I$, we can express this specific Yosida approximation easily in terms of the projection P_V (which represents the projection onto the set V, as defined in [55, 23]). Using equation (2-62), we have

$$(I_V)^{\omega}_{\lambda}(u) = \frac{1}{1-\lambda\omega}(I_V)_{\frac{1}{1-\lambda\omega}}\left(\frac{u}{1-\lambda\omega}\right) - \frac{\omega}{1-\lambda\omega}u. \tag{2-73}$$

Denote

$$\bar{\lambda} = \frac{\lambda}{1-\lambda\omega} \quad \text{and} \quad \bar{u} = \frac{u}{1-\lambda\omega}.$$

We can obtain the following identity

$$(I_V)^{\omega}_{\lambda}(u) = \frac{1}{1-\lambda\omega}(I_V)_{\bar{\lambda}}(\bar{u}) - \omega\bar{u}, \tag{2-74}$$

where

$$(I_V)_{\bar{\lambda}} = \frac{1}{\bar{\lambda}}(\bar{u} - J_{\bar{\lambda}}(\bar{u})) \quad \text{and} \quad J_{\bar{\lambda}} = (I + \bar{\lambda}\partial I_V)^{-1}. \tag{2-75}$$

Next, we compute the resolvent operator $J_{\bar{\lambda}} = (I + \bar{\lambda}\partial I_V)^{-1}$.
Given $u \in H$, let be $v \in J_{\bar{\lambda}}(u) = (I + \bar{\lambda}\partial I_V)^{-1}(u)$, we have the equivalence

$$u \in (I + \bar{\lambda}\partial I_V)(v) \Leftrightarrow \frac{u - v}{\bar{\lambda}} \in \partial I_V(u).$$

If we use the definition of subdifferential, we have

$$I_V(w) - I_V(v) \geq \left(\frac{u - v}{\bar{\lambda}}, w - v\right), \quad \text{for all } w \in V.$$

Therefore, the previous inequality leads to

$$0 \geq \left(\frac{u - v}{\bar{\lambda}}, w - v\right), \quad \text{for all } w \in V.$$

Therefore, we have

$$v \in V, \quad (v - u, w - v) \geq 0, \quad \text{for all } w \in V.$$

and by (2-72), the element v can be expressed as

$$v = P_V(u).$$

Thus, using (2-75) we can express $(\partial I_V)_{\overline{\lambda}}$ of the following form

$$(\partial I_V)_{\overline{\lambda}} = \frac{\overline{u} - P_V(\overline{u})}{\overline{\lambda}}. \tag{2-76}$$

A particular example used in Chapter 4 is the following case. Consider a bounded two-dimensional rectangular domain D, and let $\partial D = \Gamma \cup \Gamma_0 \cup \Gamma_a$ denote its boundary (see Figure **4-1**).
Let $H = H^1(D)$ be the space introduced in Definition 2, and let $V_- = H^1(D) \cap [v|_{\Gamma_0} \leq 0]$, where

$$[v|_{\Gamma_0} \leq 0] = \{\psi \in H(D); \psi|_{\Gamma_a} = 0; \psi|_{\Gamma_0} \leq 0\};$$

see section 4.1, where $(I_V)^{\omega}_{\lambda}$ is used with $\lambda\omega = 1/2$. It can be shown that, in this case, using (2-74) and (2-76) we obtain,

$$\begin{aligned}
(\partial I_{V_-})^{\omega}_{\frac{1}{2\omega}}(u) &= 2\left(\frac{\overline{u} - P_{V_-}(\overline{u})}{\overline{\lambda}}\right) - \omega\overline{u} \\
&= 2\left(\frac{2u - P_{V_-}(2u)}{\frac{1}{\omega}}\right) - 2\omega u \\
&= 2\omega u - 2\omega P_{V_-}(2u) \\
&= 2\omega u - 4\omega P_{V_-}(u).
\end{aligned}$$

2.4 Duality method for solving variational inequalities

This section is based on [10, 12, 23, 55].

2.4.1 Conjugate functionals

When dealing with an optimization problem that lacks differentiability, it becomes crucial to construct an appropriate Lagrangian in order to identify a saddle point. This is where the following definition plays a vital role.

Definition 13 (Conjugate functional)**.** *Let $\varphi : H \to (-\infty, \infty]$ be a functional on the locally convex space V. The conjugate functional $\varphi^* : H \to (-\infty, \infty]$ to φ is defined by*

$$\varphi^*(u) = \sup_{q \in V} \{(u, q) - \varphi(q)\}, \quad \text{for all} \quad u \in H'. \tag{2-77}$$

We have the following relation for $\varphi : H \to \overline{\mathbb{R}}$. Let $\varphi : H \to (-\infty, \infty]$ be a functional on an Hilbert space H.

$$(\varphi^*)^*(u) = \sup_{u^* \in H'} \{(u^*, u) - \varphi^*(u^*)\}, \qquad \text{for all } u \in H. \tag{2-78}$$

Another affirmation necessary for our study is given in the following theorem.

Theorem 10 (Fenchel and Moreau [23, Theorem 51.A])**.** *For $\varphi : H \to \overline{\mathbb{R}}$, we have*

1. *$\varphi = \varphi^{**}$ if and only if φ is convex and lower semicontinuous.*
2. *$u \in \partial\varphi(q)$ if and only if $q \in \partial\varphi(u)$ where φ is lower semicontinuous.*

2.4.2 Variational inequalities

Consider variational inequalities of the form: Find $u \in H$ such that for all $v \in H$,

$$(Au, v - u) + \varphi(v) - \varphi(u) \geq (f, v - u), \tag{2-79}$$

where H is a real Hilbert space, A is a monotone operator, $A \in L(H, H')$ and $\varphi : H \to \overline{\mathbb{R}}$ is a lower semi-continuous convex function over H.

Theorem 11. *If* (2-79) *is satisfied, then the quantity*

$$J(u) = \frac{1}{2}(Au, u) - (f, u) + \varphi(u), \tag{2-80}$$

attains an infimum in H.

Proof. For u, $v \in H$, we have

$$\begin{aligned} J(v) - J(u) &= \frac{1}{2}(Av, v) - (f, v) + \varphi(v) - \frac{1}{2}(Au, u) + (f, u) - \varphi(u), \\ &= \frac{1}{2}((Av, v) - (Au, u)) - (f, v - u) + \varphi(v) - \varphi(u). \end{aligned} \tag{2-81}$$

Take the first expression of equation (2-81) and replace it with the following expression

$$\begin{aligned} (Av, v) &= (A(v - u) + Au, (v - u) + u), \\ &= (A(v - u), v - u) + 2(Au, v - u) + (Au, u). \end{aligned}$$

We obtain

$$\begin{aligned} J(v) - J(u) &= \frac{1}{2}(A(v - u), v - u) + (Au, v - u) - (f, v - u) + \varphi(v) - \varphi(u), \\ &\geq \frac{1}{2}(A(v - u), v - u), \end{aligned}$$

where the ultimate inequality due to (2-79). Given that G is monotone, we obtain

$$J(v) - J(u) \geq 0.$$

Therefore, u is an infimum of J in H. □

Duality methods for the numerical solution of (2-79) try to overcome the difficulty related to the non-differentiability of φ by constructing a convenient Lagrangian (see [10, 12, 23, 55]). We use the conjugate functional of φ and by Theorem 10, we obtain

$$J(u) = \frac{1}{2}(Gu, u) - (f, u) + \sup_{q \in V}\{(q, u) - \varphi^*(q)\} \qquad \text{for all } u \in H. \tag{2-82}$$

This is given that φ is l.s.c. As we find the infimum of J, we obtain

$$\inf_{u \in V} J(u) = \inf_{u \in V} \sup_{q \in V} \mathcal{L}(u, q), \tag{2-83}$$

where $\mathcal{L}$ is the Lagrangian defined by

$$\mathcal{L}(u, q) = \frac{1}{2}(Au, u) - (f, u) + (q, u) - \varphi^*(q). \tag{2-84}$$

Now, for $\mathcal{L}(u, q)$, we find y and p, such that

$$0 \in \frac{\partial \mathcal{L}}{\partial u}(y, p) \text{ and } 0 \in \frac{\partial \mathcal{L}}{\partial q}(y, p). \tag{2-85}$$

From $0 \in \frac{\partial \mathcal{L}}{\partial u}(y, p)$, we obtain

$$Ay + p = f. \tag{2-86}$$

And from $0 \in \frac{\partial \mathcal{L}}{\partial q}(y,p)$, we obtain

$$y \in \partial\varphi^*(p), \tag{2-87}$$

and by Theorem 10 (second item), we have

$$p \in \partial\varphi(y).$$

Therefore, $\{u,p\}$ is a saddle-point of $\mathcal{L}$ and then u is a solution of (2-79) and moreover

$$\begin{aligned} Au + p &= f, \\ p &\in \partial\varphi(u). \end{aligned} \tag{2-88}$$

We introduce a variable α, such that

$$\alpha = p - \omega u. \tag{2-89}$$

Then, if we replace in (2-88) second part, we obtain

$$\alpha \in \partial\varphi(u) - \omega u. \tag{2-90}$$

Then, problem (2-79) is equivalent to the following problem (Lemma 2),

$$\begin{aligned} (A+\omega I)u + \alpha &= f \\ \alpha &= (\partial\varphi)_\lambda^\omega(u + \lambda\alpha) \text{ for every } \lambda > 0, \end{aligned} \tag{2-91}$$

where $(\partial\varphi)_\lambda$ is the Yosida approximation of the maximal monotone operator $\partial\varphi$.
Thus, we can define the following algorithm: Let p_0 be given arbitrarily. Given p_m, we define u_m, p_{m+1} by

$$\begin{aligned} (A+\omega I)u_m + \alpha_m &= f \\ \alpha_{m+1} &= (\partial\varphi)_{\lambda_m}(u_m + \lambda\alpha_m). \end{aligned} \tag{2-92}$$

Therefore,

$$\alpha_{m+1} = (\partial\varphi)_\lambda^\omega\left((A+\omega I)^{-1}(f-\alpha_m) + \lambda_m\alpha_m\right). \tag{2-93}$$

This process is known as the algorithm of Uzawa (see [23]).

3 Generalized multiscale finite elements method (GMsFEM)

In this chapter, the central concepts we study are the MsFEM and the GMsFEM. Two main ingredients of the MsFEM are the overall formulation of the method and the construction of basis functions. We discuss global formulations using various finite element methods. We present simplified computations of basis functions for cases with separation of scale. For the GMsFEM, we begin to vary the number of basis functions taken per neighborhood. As for the multiscale basis functions, the capturing errors of the subgrid are analyzed. We follow [33, 1].

3.1 Multiscale finite element methods (MsFEM)

In this section, we describe in general form, the MsFEM for linear elliptic equations

$$Lp = f \quad \text{in } \Omega, \tag{3-1}$$

where Ω is a domain in $\mathbb{R}^d$ $(d = 2, 3)$ and

$$Lp = -\text{div}(\kappa(x)\nabla p). \tag{3-2}$$

Furthermore, $\kappa(x)$ is a heterogeneous field varying over multiple scales. We note that the tensor $\kappa(x) = (\kappa_{ij}(x))$ is assumed to be symmetric and satisfies $\alpha|\xi|^2 \leq \kappa_{ij}\xi_i\xi_j \leq \beta|\xi|^2$, for all $\xi \in \mathbb{R}^d$ and some $0 < \alpha < \beta$ (see [33]). The MsFEM consists of two important elements, the construction of multiscale basis functions and an appropriate global numerical formulation that combines these multiscale basis functions. The basis functions are designed to capture the multiscale features of the solution. A suitable global formulation combines these basis functions to precisely approximate the solution.

3.1.1 Basis functions

First, we show the basis function construction. Let $\mathcal{T}^H$ be the usual triangulation of Ω into finite elements (triangles, quadrilaterals, and others). We call this partition the coarse grid and assume that the coarse grid can be resolved via a finer resolution which is called the fine grid. Let x_i be the interior nodes of the mesh $\mathcal{T}^H$ and ϕ_i^0 be the nodal basis of the standard finite element space W^H =span(ϕ_i^0). One can assume that W^H consists of piecewise linear

functions if $\mathcal{T}^H$ is a triangular partition. Denote by $\omega_i = \text{supp}(\phi_i^0)$ (the support of ϕ_i^0) and define ϕ_i with support in ω_i as the solution of the problem

$$L\phi_i = 0 \text{ in } K, \quad \phi_i = \phi_i^0 \text{ on } \partial K, \quad \text{for all } K \in \mathcal{T}^H, \quad K \subset \omega_i. \tag{3-3}$$

Multiscale basis functions coincide with standard finite element basis functions on the boundaries of a coarse-grid block K and oscillate in each coarse-grid block's interior. We denote K as a coarse-grid block. Our primary assumption is that the basis functions satisfy the leading-order homogeneous equation when the right-hand side f is a smooth function (L^2 integrable). We denote by P^h the finite element space spanned by ϕ_i,

$$P^h = \text{span}\{\phi_i\}.$$

3.1.2 Global formulation

The representation of the solution at a fine scale through multiscale basis functions allows for reducing the dimension of the calculation. When the approximation of the solution $p^H = \sum_i p_i \phi_i(x)$ (p_i are the approximate values of the solution at the nodal points of the coarse grid) is substituted into the fine-scale equation, the resulting system is projected into coarse-dimensional space to find p_i. This can be done by multiplying the resulting fine-scale equation with coarse-scale test functions.
In the case of Galerkin finite element methods, when the basis functions are conformal ($P^H \subset H_0^1(\Omega)$), the MsFEM is to find $p^H \in P^H$ such that,

$$\sum_K \int_K k \nabla p^H \nabla v^H dx = \int_\Omega f v^H dx, \quad \text{for all } v^H \in P^H. \tag{3-4}$$

We note that in (3-4), the fine-scaled system is multiplied by coarse-scaled trial functions, and the resulting system is coarse-dimensional. The equation (3-4) couples the multiscale basis functions. This gives rise to a linear system of equations to find the solution values at the coarse grid block nodes. Therefore, the resulting linear of system equations determines the coarse grid's solution. Representing the solution in terms of multiscale basis functions, $p^H = \sum_i p_i \phi_i$, it is easy to show that (3-4) is equivalent to the following linear system,

$$A \cdot \hat{p} = b, \tag{3-5}$$

where $A = (a_{ij})$ be a matrix with $a_{ij} = \sum_K \int_K \kappa \nabla \phi_i \nabla \phi_j^0, dx$. The vector $\hat{p} = (p_1, ..., p_i, ...)$ represents the nodal values of the coarse-scale solution, and $b = (b_i)$, where $b_i = \int_\Omega f \phi_i^0, dx$. In this context, we disregard the discretization of the boundary conditions. Similar to standard finite element methods, the stiffness matrix A exhibits a sparse structure and may be poorly conditioned when dealing with highly contrasting media. It is important to note that the computation of the coarse stiffness matrix involves comparing integrals, and the evaluation of a_{ij} requires performing integrals over the fine mesh.

In summary, the MsFEM can be implemented within an existing finite element code:

- Set coarse and fine mesh configuration.
- We construct the basis functions on coarse mesh, then we solve (3-3) for each coarse grid block.
- Assemble stiffness matrix on the coarse mesh.
- Assemble the external force on the coarse mesh.
- Solve the coarse formulation.

For more details see, [33].

As shown in several works, the classical multiscale method is inadequate for high-contrast multiscale problems; see, e.g., [36, 37, 28]. Recall that we refer to high-contrast multiscale problems when the permeability coefficient varies at several scales, including high jumps. The main problem of the MsFEM when the medium has high-contrast and multiscale variations is that it generates systems with very large and ill-conditionate matrices. For this reason, we now study the generalized multiscale finite element method (GMsFEM). We now review the constructions of the GMsFEM basis functions, first in one dimension in Section 3.2 and later in two dimensions in Section 3.3. The construction is similar, but we present both to show all the construction details.

3.2 The GMsFEM in one dimension

Let $D \subset \mathbb{R}$ denote an interval and $\mathcal{T}^H = \{K\}$ be a coarse-grid partition, where H denotes the size of the coarse grid. Then, we consider the elliptic equation with heterogeneous coefficients

$$-\left(\kappa(x)u'\right)' = f,$$

where $\kappa(x)$ is a heterogeneous field with high contrast and multiscale behavior. In particular, we assume that $\kappa(x) \geq C_0 > 0$, while $\kappa(x)$ can have very large values and local variation. The variational formulation of this problem is: find $u \in H_0^1(D)$ such that

$$a(u, v) = F(v) \text{ for all } v \in H_0^1(D),$$

where the bilinear form a and the linear functional F are defined by

$$a(u, v) = \int_D \kappa(x)u'(x)v'(x)dx \text{ for all } u, v \in H_0^1(D),$$

and

$$F(v) = \int_D f(x)v(x)dx.$$

Let $\mathcal{T}^h$ be a fine triangulation, which is a refinement $\mathcal{T}^H$. We denote by $V^h(D)$ the finite element discretization of piecewise linear continuous functions concerning the fine triangulation $\mathcal{T}^h$. Denote also by $V_0^h(D)$ the subset of $V^h(D)$ where the boundary is zero.
The Galerkin formulation of the problem is to find $u \in V_0^h(D)$ such that

$$a(u,v) = F(v), \text{ for all } v \in V_0^h(D),$$

or in matrix form

$$v^t A u = v^t b,$$

where for all $u, v \in V^h(D)$, we have

$$u^t A v = \int_D \kappa(x) u' v' dx \quad \text{and} \quad v^t b = \int_D f v dx.$$

We denote by $\{y_i\}_{i=1}^{N_v}$ the vertices of the coarse mesh $\mathcal{T}^H$ and we define the neighborhood of the node y_i by $\omega_i = \bigcup \{K_j \in \mathcal{T}^H; y_i \in \overline{K}\}$, and the neighborhood of the coarse element K by $\omega_K = \bigcup \{\omega_i \in \mathcal{T}^H; y_i \in \overline{K}\}$, see Figure **3-1**.
Using the coarse mesh $\mathcal{T}^H$, we introduce coarse basis functions $\{\phi_i\}_{i=1}^{L}$, where L is the number of coarse basis functions. The MsFEMs approximates the solution on a coarse grid as $u_0 = \sum_i c_i \phi_i$, where c_i are determined from

$$a(u_0, v) = f(v), \text{ for all } v \in \text{span}\{\phi_i\}_{i=1}^{L},$$

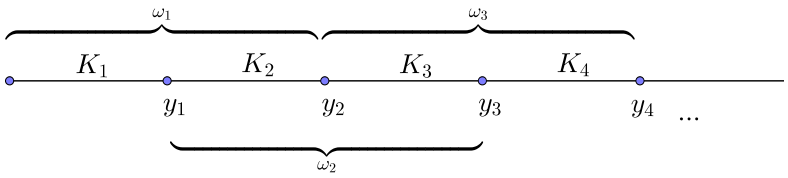

Figure 3-1: Ilustration of a coarse grid.

3.2.1 Coarse space

We consider the following high-contrast eigenvalue problem

$$-(\kappa u')' = \lambda \kappa u \qquad x \in \omega_i, \tag{3-6}$$

where ω_i with homogeneous Neumann boundary condition.
For any $\Omega \subset D$, we define the fine-scale and local matrix A^Ω as follows

$$v^t A^\Omega w = \int_\Omega \kappa v' w' dx, \text{ for all } v, w \in V^h(\Omega), \quad i = 1, ..., N,$$

and the mass matrix of the same dimension M^{Ω} defined by

$$v^t M^{\Omega} w = \int_{\Omega} \kappa v w dx, \text{ for all } v, w \in V^h(\Omega).$$

We consider the finite-dimensional symmetric eigenvalue problem

$$A^{\Omega}\phi = \lambda M^{\Omega}\phi,$$

and denote its eigenvalues and eigenvectors as $\left\{\lambda_{\ell}^{\Omega}\right\}$ and $\left\{\psi_{\ell}^{\Omega}\right\}$ respectively. Note that the eigenvectors form an orthonormal basis of $V^h(\Omega)$ with respect to M^{Ω}. It is important to note that $\lambda_1^{\Omega} = 0$. We order these eigenvalues as

$$\lambda_1^{\Omega} \leq \lambda_2^{\Omega} \leq \lambda_3^{\Omega} \leq .... \tag{3-7}$$

In particular, $\psi_{\ell}^{\omega_i}$ denote the ℓth eigenvector oh the Neumman Matrix associated with the neighborhood of y_i.
We assume that the elements of $\mathcal{T}^H$ contained in a Ω form a triangulation in Ω. For any $v \in V^h(\Omega)$, we have the following properties (see [57, 25]),

$$v = \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa v \psi_{\ell}^{\omega_i} \right) \psi_{\ell}^{\omega_i},$$

also, we have

$$a(v, v) = \int_{\omega_i} \kappa (v')^2 = \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa v \psi_{\ell}^{\omega_i} \right)^2 \lambda_{\ell}^{\omega_i}, \tag{3-8}$$

and

$$m(v, v) = \int_{\omega_i} \kappa (v)^2 = \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa v \psi_{\ell}^{\omega_i} \right)^2. \tag{3-9}$$

For an integer L and $v \in V^h(\omega_i)$, let us define

$$I_L^{\omega_i} v = \sum_{\ell=1}^{L} \left(\int_{\omega_i} \kappa v \psi_{\ell}^{\omega_i} \right) \psi_{\ell}^{\omega_i}. \tag{3-10}$$

Thus, from equations (3-7), (3-9), and (3-8), it is easy to see that

$$\begin{aligned}
\int_{\omega_i} \kappa(v - I_L^{\omega_i} v)^2 &= \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v - I_L^{\omega_i} v)\psi_\ell^{\omega_i} \right)^2 \\
&= \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v - I_L^{\omega_i} v)\psi_\ell^{\omega_i} \right)^2 \frac{\lambda_\ell^{\omega_i}}{\lambda_\ell^{\omega_i}} \\
&\leq \frac{1}{\lambda_{L+1}^{\omega_i}} \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v - I_L^{\omega_i} v)\psi_\ell^{\omega_i} \right)^2 \lambda_\ell^{\omega_i} \\
&\leq \frac{1}{\lambda_{L+1}^{\omega_i}} \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v)\psi_\ell^{\omega_i} \right)^2 \lambda_\ell^{\omega_i} \\
&= \frac{1}{\lambda_{L+1}^{\omega_i}} \int_{\omega_i} \kappa \left(v'\right)^2 .
\end{aligned}$$

And therefore, we get the following inequality

$$\int_{\omega_i} \kappa(v - I_L^{\omega_i} v)^2 \leq \frac{1}{\lambda_{L+1}^{\omega_i}} a\left(v - I_{L_i}^{\omega_i} v, v - I_{L_i}^{\omega_i} v\right) \leq \frac{1}{\lambda_{L+1}^{\omega_i}} a\left(v, v\right). \tag{3-11}$$

3.2.2 Coarse space with a Neumann eigenvalue problem in neighborhoods

Now we construct the coarse space. First note that $\{\omega_i\}_{y_i \in \mathcal{T}^H}$ is a covering of D. Let $\{\chi_i\}_{i=1}^{N_v}$ be a partition of unity subordinated to the covering $\{\omega_i\}$ such that $\chi_i \in V^h(D)$ y $\|\chi_i'\| \leq \frac{1}{H}$, $i = 1, 2, ..., N_v$. Define the set of coarse basis function by

$$\phi_{i,\ell} = \chi_i \psi_\ell^{\omega_i} \text{ for } 1 \leq i \leq N_v \text{ and } 1 \leq \ell \leq L_i,$$

where I^h is the fine-scale nodal value interpolation, and L_i is the number of eigenvalues that we will choose for the node i. In Figure **3-2**, we show the first four elements where

$$\kappa(x) = \begin{cases} 1 & \text{if } x \in K_i,\ i \text{ odd}, \\ \eta & \text{if } x \in K_i,\ i \text{ even}, \end{cases} \tag{3-12}$$

here, $\omega_i = [0, 1]$ and $\eta = 1000$. Denote by V_N, the local spectral space multiscale

$$V_N = \text{span}\left\{\phi_{i,\ell}; 1 \leq i \leq N_v \text{ and } 1 \leq \ell \leq L_i\right\}.$$

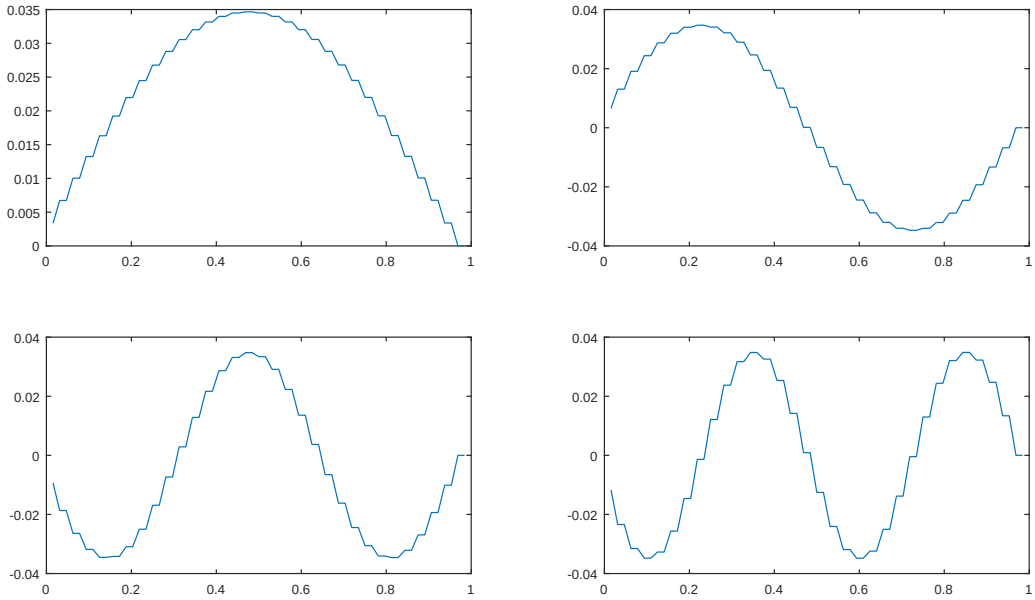

Figure 3-2: First four elements of the spectral basis where $\kappa(x)$ is defined in (3-12) for $\eta = 1000$ in $\omega_i = [0, 1]$.

3.2.3 Approximation and stability using V_N

Let $v \in V^h(D)$, we define

$$I_N v = \sum_{i=1}^{N_v} \sum_{\ell=1}^{L_i} \left(\int_{\omega_i} \kappa v \psi_\ell^{\omega_i} \right) \chi_i \psi_\ell^{\omega_i} = \sum_{i=1}^{N_v} \left(I_{L_i}^{\omega_i} v \right) \chi_i,$$

where $I_{L_i}^{\omega_i}$ is defined in (3-20). Note that we have

$$\begin{aligned} v - I_N v &= v - \sum_{i=1}^{N_v} \chi_i \left(I_{L_i}^{\omega_i} v \right) \\ &= v \sum_{i=1}^{N_v} \chi_i - \sum_{i=1}^{N_v} \chi_i \left(I_{L_i}^{\omega_i} v \right) \\ &= \sum_{i=1}^{N_v} \chi_i \left(v - I_{L_i}^{\omega_i} v \right). \end{aligned} \tag{3-13}$$

For the interpolation, we study the approximation in L^2 and stability in H^1.

First, we show the approximation in L^2. Observe that

$$\begin{aligned}
\int_K \kappa \left(v - I_N v\right)^2 &= \int_K \kappa \left(\sum_{i=1}^{N_v} \chi_i \left(v - I_{L_i}^{\omega_i} v\right)\right)^2 \\
&\leq \int_K \kappa \left(\sum_{i=1}^{N_v} \left(v - I_{L_i}^{\omega_i} v\right)\right)^2 \\
&\leq \int_K \kappa \sum_{i=1}^{N_v} 1^2 \sum_{i=1}^{N_v} \left(v - I_{L_i}^{K} v\right)^2 \\
&\leq N_v \sum_{i=1}^{N_v} \int_K \kappa \left(v - I_{L_i}^{K} v\right)^2 \\
&\leq N_v \sum_{i=1}^{N_v} \frac{1}{\lambda_{L_i}^{\omega_i}} \int_{\omega_i} \kappa \left(v'\right)^2 \\
&\leq \frac{N_v}{\lambda_L^{\omega_i}} \sum_{i=1}^{N_v} \int_{\omega_i} \kappa \left(v'\right)^2.
\end{aligned}$$

The following proof is the stability in H^1. We note that the set ω_i is composed of two coarse blocks, then we have that $\chi_i' + \chi_{i+1}' = 0$. Therefore, we have

$$\begin{aligned}
\left(\chi_i I_{L_i}^{\omega_i} v + \chi_{i+1} I_{L_{i+1}}^{\omega_{i+1}} v\right)' &= \chi_i' \left(I_{L_i}^{\omega_i} v\right) + \chi_i \left(I_{L_i}^{\omega_i} v\right)' + \chi_{i+1}' \left(I_{L_{i+1}}^{\omega_{i+1}} v\right) + \chi_{i+1} \left(I_{L_{i+1}}^{\omega_{i+1}} v\right)' \\
&= \chi_{i+1}' \left(I_{L_{i+1}}^{\omega_{i+1}} v - I_{L_i}^{\omega_i} v\right) + \chi_i \left(I_{L_i}^{\omega_i} v\right)' + \chi_{i+1} \left(I_{L_{i+1}}^{\omega_{i+1}} v\right)'.
\end{aligned}$$

From this, we have the following bound valid in K,

$$\begin{aligned}
\left|\left(\chi_i I_{L_i}^{\omega_i} v + \chi_{i+1} I_{L_{i+1}}^{\omega_{i+1}} v\right)'\right|^2 &= \left(\chi_{i+1}' \left(I_{L_{i+1}}^{\omega_{i+1}} v - I_{L_i}^{\omega_i} v\right) + \chi_i \left(I_{L_i}^{\omega_i} v\right)' + \chi_{i+1} \left(I_{L_{i+1}}^{\omega_{i+1}} v\right)'\right)^2 \\
&\leq 2 \left(\chi_{i+1}' \left(I_{L_{i+1}}^{\omega_{i+1}} v - I_{L_i}^{\omega_i} v\right)\right)^2 + 2 \left(\chi_i \left(I_{L_i}^{\omega_i} v\right)' + \chi_{i+1} \left(I_{L_{i+1}}^{\omega_{i+1}} v\right)'\right)^2 \\
&\leq \frac{2}{H^2} \left(I_{L_{i+1}}^{\omega_{i+1}} v - I_{L_i}^{\omega_i} v\right)^2 + 4 \left(\left(I_{L_i}^{\omega_i} v\right)'\right)^2 + 4 \left(\left(I_{L_{i+1}}^{\omega_{i+1}} v\right)'\right)^2,
\end{aligned}$$

this last step is justified by the fact that $\chi_i \leq 1$ and $|x'| \leq \frac{1}{H}$. Since we want to find an upper bound for $\int_K \kappa \left|(I_N v)'\right|^2$, using the previous inequality, we obtain the following

$$\begin{aligned}\int_K \kappa\left|(I_N v)'\right|^2 &\leq \int_{\omega_i} \kappa\left|(I_N v)'\right|^2 \\ &= \int_{\omega_i} \kappa\left|\left(\chi_i\left(I_{L_i}^{\omega_i}\right)v + \chi_{i+1}\left(I_{L_{i+1}}^{\omega_{i+1}}\right)v\right)'\right|^2 \\ &\leq \frac{2}{H^2}\int_{\omega_i}\kappa\left(I_{L_{i+1}}^{\omega_{i+1}}v - I_{L_i}^{\omega_i}v\right)^2 + \int_{\omega_i}\kappa\left(\left(I_{L_i}^{\omega_i}v\right)'\right)^2 + \int_{\omega_i}\kappa\left(\left(I_{L_{i+1}}^{\omega_{i+1}}v\right)'\right)^2.\end{aligned} \tag{3-14}$$

For the first term on the right-hand side, we use equation (3-11) to obtain that,

$$\begin{aligned}\int_{\omega_i}\kappa\left(I_{L_{i+1}}^{\omega_{i+1}}v - I_{L_i}^{\omega_i}v\right)^2 &= \int_{\omega_i}\kappa\left(I_{L_{i+1}}^{\omega_{i+1}}v - v + v - I_{L_i}^{\omega_i}v\right)^2 \\ &\leq 2\int_{\omega_i}\kappa\left(v - I_{L_{i+1}}^{\omega_{i+1}}v\right)^2 + 2\int_{\omega_i}\kappa\left(v - I_{L_i}^{\omega_i}v\right)^2 \\ &\leq \frac{1}{\lambda_{L+1}^{\omega_K}}\int_{\omega_i}\kappa\left(v'\right)^2.\end{aligned} \tag{3-15}$$

For the second term on the right-hand side of equation (3-14), we use the orthogonality of the eigenvectors, which yields that

$$\int_{\omega_i}\kappa\left(\left(I_{L_i}^{\omega_K}v\right)'\right)^2 \leq \int_{\omega_K}\kappa\left(v'\right)^2. \tag{3-16}$$

Then, using (3-15), (3-16) and (3-14), we get

$$\int_K \kappa\left|(I_N v)'\right|^2 \leq \max\left\{1, \frac{1}{H^2\lambda_K}\right\}\int_K \kappa\left(v'\right)^2. \tag{3-17}$$

For the approximation in H^1, the analysis is the following

$$\begin{aligned}\int_{\omega_i}\kappa\left((v - I_N v)'\right)^2 &\leq \int_{\omega_i}\kappa\left(\left(\sum_{y_i\in K}\chi_i\left(v - I_{L_i}^{\omega_i}v\right)\right)'\right)^2 \\ &\leq 2\int_K \kappa\left(\chi_i'\left(v - I_{L_i}^{\omega_i}v\right) + \chi_i\left(v - I_{L_i}^{\omega_i}v\right)'\right)^2 \\ &\leq 4\int_K \kappa\left(\left(\chi_i'\right)^2\left(v - I_{L_i}^{\omega_i}v\right)^2 + \chi_i^2\left(v - I_{L_i}^{\omega_i}v\right)'^2\right) \\ &\leq 8\frac{1}{H^2}\int_K \kappa\left(v - I_{L_i}^{\omega_i}v\right)^2 + 8\left(\max_{y_i\in K}\lambda_L^{\omega_i}\right)\sum_{\ell=L+1}\left(\int_{\omega_i}\kappa\left(v\psi_\ell^{\omega_i}\right)\right)^2\left(\lambda_\ell^{\omega_i}\right)^2 \\ &\leq 8\frac{1}{H^2}\left(\max_{y_i\in K}\lambda_L^{\omega_i}\right)\int_{\omega_i}\kappa\left(v'\right)^2 + 8\left(\max_{y_i\in K}\lambda_L^{\omega_i}\right)\sum_{\ell=L+1}\left(\int_{\omega_i}\kappa\left(v\psi_\ell^{\omega_i}\right)\right)^2\left(\lambda_\ell^{\omega_i}\right)^2.\end{aligned}$$

Now, from (3-9), we obtain the following inequality:

$$\int_{\omega_i} \kappa \left(|v - I_N v|'\right)^2 \leq 16\frac{H^2}{\widetilde{\lambda}} \sum_{\ell=L+1} \left(\int_{\omega_i} \kappa \left(v\psi_\ell^{\omega_i}\right)\right)^2 (\lambda_\ell^{\omega_i})^2, \tag{3-18}$$

where $\widetilde{\lambda} = \max_{y_i \in K} \lambda_L^{\omega_i}$. It is important to note that we need to restrict the last summation in order to bound the interpolation approximation. This process is provided in the following result.

Theorem 12. *If $(\kappa u')' \in L^2(\omega_i)$, where $\omega_i = [a, b]$ and $u'(a) = u'(b) = 0$ then*

$$\int_{\omega_i} ((\kappa u')')^2 = \int_a^b ((\kappa u')')^2 = \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} v\psi_\ell^{\omega_i}\right)^2 (\lambda_\ell^{\omega_i})^2.$$

Proof. Recall that for $v \in L^2(D)$ we have the expansion $v = \sum_{\ell=1}^{L} \left(\int_{\omega_i} \kappa v \psi_\ell^{\omega_i}\right) \psi_\ell^{\omega_i}$. For any integer n define the truncated approximation of v as:

$$v_n = \sum_{\ell=1}^{n} \left(\int_{\omega_i} \kappa v \psi_\ell^{\omega_i}\right) \psi_\ell^{\omega_i}.$$

With the eigenvalue problem (3-6), we obtain the following equalities

$$\begin{aligned} (\kappa v_n')' &= \sum_{\ell=1}^{n} \kappa \left(\int_{\omega_i} \kappa v \psi_\ell^{\omega_i}\right) \lambda_\ell^{\omega_i} \psi_\ell^{\omega_i} \\ \kappa^{-1} (\kappa v_n')' &= \sum_{\ell=1}^{n} \left(\int_{\omega_i} v \psi_\ell^{\omega_i}\right) \lambda_\ell^{\omega_i} \psi_\ell^{\omega_i}. \end{aligned}$$

Using the orthogonality of the eigenvectors, we can conclude that for every $m > n$

$$\begin{aligned} \left\|\kappa^{-1}(\kappa v_m')' - \kappa^{-1}(\kappa v_n')'\right\|^2 &= \left\| \sum_{\ell=n+1}^{m} \left(\int_{\omega_i} v\psi_\ell^{\omega_i}\right) \lambda_\ell^{\omega_i} \psi_\ell^{\omega_i} \right\|^2 \\ &= \int_{\omega_i} \left(\sum_{\ell=n+1}^{m} \left(\int_{\omega_i} v\psi_\ell^{\omega_i}\right) \lambda_\ell^{\omega_i} \psi_\ell^{\omega_i} \right)^2 \\ &\leq c \sum_{\ell=n+1}^{m} \left(\int_{\omega_i} v\psi_\ell^{\omega_i}\right)^2 (\lambda_\ell^{\omega_i})^2, \end{aligned} \tag{3-19}$$

assuming $\int_{\omega_i} (\psi_\ell^{\omega_i})^2 \leq 1$and given that

$$\|u\|^2 = \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} v\psi_\ell^{\omega_i}\right)^2 (\lambda_\ell^{\omega_i})^2 < \infty,$$

we conclude that there exist an $L^2(\omega_i)$ function, denoted by v, such that $\|U + \kappa^{-1}(\kappa v_n')'\| \to 0$ when $n \to \infty$. We also have that for any $z \in H_0^1(\omega_i)$

$$\int_{\omega_i} \kappa u z = \lim_{n\to\infty} \int_{\omega_i} (\kappa v_n')' z = \int \kappa v' z'.$$

This shows that $\kappa u = (\kappa u')'$. □

3.2.4 Coarse space with Dirichlet eigenvalue problem in blocks

We consider the following high contrast eigenvalue problem. For any $\Omega \subset D$, we define the fine-scale and local Dirichlet matrix A^K by

$$vA^K w = \int_K \kappa v' w' \, dx \quad \text{for all} \quad v, w \in V^h(K), \quad i = 1, ..., N,$$

and the mass matrix of the same dimension M^K is defined by

$$vM^K w = \int_K \kappa v w \, dx, \quad \text{for all} \quad v, w \in V^h(K).$$

We consider the following finite-dimensional symmetric eigenvalue problem:

$$A^K \phi = \mu M^K \phi,$$

where we denote the eigenvalues by $\{\mu_m\}$, the eigenvectors by $\{\varphi_m\}$, and the spectral basis set by

$$V_D = \text{span}\left\{\varphi_m^K : 1 \leq \ell \leq L_i\right\},$$

where L_i is the number of eigenvalues in each neighborhood.
The eigenvalue problem above with Dirichlet conditions is

$$-(\kappa u')' = \lambda \kappa u \qquad x \in K.$$

3.2.5 Approximation and stability using V_D

For any L and $v \in V^h(K)$, define

$$I_D^K v = \sum_{m=1}^{L} \left(\int_K \kappa v \varphi_m^K \right) \varphi_m^K. \tag{3-20}$$

Define also the coarse interpolation $I_D : V^h(D) \to V_0$ by

$$I_D v = \sum_{y_i \in K} I_D^K v.$$

Note that

$$v - I_D v = \sum_{y_i \in K} (v - I_D^K v).$$

Theorem 13. *For all elements in the coarse space K with $v = 0$ on ∂K, we have*

$$\int_0^1 \kappa \left(v - I_D v\right)^2 \leq 8 \frac{H^2}{\widetilde{\lambda}} \int_0^1 \kappa \left(v'\right)^2, \tag{3-21}$$

and

$$\int_K \kappa \left(|v - I_D v|'\right)^2 \leq 2 \left(\max_{y_i \in K} \lambda_M^K \right) \int_K \left((\kappa u')'\right)^2. \tag{3-22}$$

Proof.

- To show (3-21)

$$\begin{aligned}\int_0^1 \kappa\,(v - I_D v)^2 &= \sum_K \int_K \kappa\left(v - I_D^K v\right)^2 \\ &\leq 2\sum_K \max_{y_i \in K}\left(\frac{1}{\lambda_{L+1}^K}\int_K \kappa\,(v')^2\right) \\ &\leq 4\sum_K \left(\max_{y_i \in K}\frac{1}{\lambda_{L+1}^K}\right)\int_{\omega_K} \kappa\,(v')^2 \\ &\leq 8\left(\max_{y_i \in K}\frac{1}{\lambda_{L+1}^K}\right)\int_0^1 \kappa\,(v')^2 \\ &\leq 8\frac{H^2}{\widetilde{\lambda}}\int_0^1 \kappa\,(v')^2.\end{aligned}$$

- To show equation (3-22), let $v = 0$ on ∂K. Then,

$$\begin{aligned}\int_K \kappa\left(|v - I_D v|'\right)^2 &\leq \int_K \kappa\left(\left(\sum_{y_i \in K}(v - I_D^{\omega_i} v)\right)'\right)^2 \\ &\leq 2\int_K \kappa\left(\left(v - I_D^K v\right)'\right)^2 \\ &\leq 2\left(\max_{y_i \in K}\lambda_M^K\right)\sum_{\ell=L+1}\left(\int_K \kappa\left(v\psi_\ell^K\right)\right)^2\left(\lambda_\ell^K\right)^2. \qquad (3\text{-}23)\end{aligned}$$

To complete the proof, we need to bound the last summation. To achieve this, we rely on the following results.

Theorem 14. *If $(\kappa u')' \in L^2(K)$, where $K = [a, b]$ and $u(a) = u(b) = 0$, then*

$$\int_K \left((\kappa u')'\right)^2 = \int_a^b \left((\kappa u')'\right)^2 = \sum_{\ell=1}^{\infty}\left(\int_K v\psi_\ell^K\right)^2\left(\lambda_\ell^K\right)^2.$$

Proof: Recall that for $v \in L^2(K)$ we have the expansion $v = \sum_{\ell=1}^{\infty}\left(\int_K \kappa v\psi_\ell^K\right)\psi_\ell^K$. For any integer n define the truncated approximation of v as

$$v_n = \sum_{\ell=1}^{n}\left(\int_K \kappa v\psi_\ell^K\right)\psi_\ell^K.$$

With the eigenvalue problem (3-6), we obtain the following equality

$$\begin{aligned}(\kappa v_n')' &= \sum_{\ell=1}^{n} \kappa \left(\int_K \kappa v \psi_\ell^K \right) \lambda_\ell^K \psi_\ell^K, \\ \kappa^{-1}(\kappa v_n')' &= \sum_{\ell=1}^{n} \left(\int_K v \psi_\ell^K \right) \lambda_\ell^K \psi_\ell^K.\end{aligned}$$

Using the orthogonality of the eigenvectors, we conclude that for every $m > n$, we have

$$\begin{aligned}\left\|\kappa^{-1}(\kappa v_m')' - \kappa^{-1}(\kappa v_n')'\right\| &= \left\| \sum_{\ell=n+1}^{m} \left(\int_K v \psi_\ell^K \right) \lambda_\ell^K \psi_\ell^K \right\| \\ &= \int_K \left(\sum_{\ell=n+1}^{m} \left(\int_K v \psi_\ell^K \right) \lambda_\ell^K \psi_\ell^K \right)^2 \\ &\leq c \sum_{\ell=n+1}^{m} \left(\int_K v \psi_\ell^K \right)^2 \left(\lambda_\ell^K\right)^2, \end{aligned} \tag{3-24}$$

for $\int_K \left(\psi_\ell^K\right)^2 \leq 1$. On the other hand,

$$\|u\| = \sum_{\ell=1}^{\infty} \left(\int_K v \psi_\ell^K \right)^2 \left(\lambda_\ell^K\right)^2 < \infty,$$

we conclude that there exists an $L^2(K)$ function, denoted by v, such that we have $|U + \kappa^{-1}(\kappa v_n')'| \to 0$ as $n \to \infty$. We also have that for any $z \in H_0^1(K)$

$$\int_K \kappa u z = \lim_{n\to\infty} \int_K (\kappa v_n')' z = \int_K \kappa v' z'.$$

3.2.6 Coarse space

In this section, we present the construction of the coarse-scale finite element spaces within the framework of GMsFEM.
We select basis functions that span the eigenfunctions associated with small eigenvalues. Let $\{\omega_i\}_{y_i \in \mathcal{T}^H}$, where $i = 1, 2, 3, ..., N_v$, denote the set of coarse regions. We define the set of coarse basis functions as follows

$$\Phi_{i,\ell} = \chi_i \Psi_\ell^{\omega_i} \quad \text{for } 1 \leq i \leq N_v \text{ and } 1 \leq \ell \leq L_i, \tag{3-25}$$

where L_i is the number of eigenvalues that will be chosen for the node i. The local spectral multiscale space

$$V_N = \text{span}\{\Phi_{i,\ell} : 1 \leq i \leq N_v \text{ and } 1 \leq \ell \leq L_i\}.$$

Define also

$$V_D = \text{span}\{\psi_\ell : K \in \mathcal{T}^H \text{ and } 1 \leq \ell \leq L_K\}.$$

and

$$V_0 = V_N + V_D.$$

Finally, we define u_H as the Galerkin approximation using the space V_0, that is,

$$a(u_H, v) = f(v), \quad \text{for all } v \in V_0. \tag{3-26}$$

3.2.7 Approximation properties of the coarse space

An important property of the coarse space and its approximation study is that if we have a high contrast medium κ and choose a large enough L_i, then $\lambda_{L_i}^{\omega_i}$ is independent of the contrast (see [32, 29, 56]).

We define the following interpolation,

$$I_N u = \sum_{i=1}^{N_v} \sum_{\ell=1}^{L_i} \left(\int_{\omega_i} \kappa u \varphi_\ell^{\omega_i} \right) \chi_i \varphi_\ell^{\omega_i} = \sum_{i=1}^{N_v} (I_{L_i}^{\omega_i} u_N) \chi_i.$$

The following lemma presents the approximation result in the norm H^1. It is based on results exposed in [1, 57].

Lemma 3. *Assume that $f = (\kappa u')' \in L^2(\omega_K)$ and also assume that for each i, $y_i \in K$ we have $\partial_\eta u = 0$ on $\partial\omega_i$. Then, the following energy approximation holds,*

$$\int_K \kappa |u' - (I_N u)'|^2 \leq \max \left\{ \frac{1}{H^2 \sigma_{K,L+1}^2}, \frac{1}{\sigma_{K,L+1}} \right\} ||f||_{L^2(\omega_K)}^2, \tag{3-27}$$

where $\sigma_{K,L+1} = \min_{y_i \in K} \sigma_{L_i+1}$ and ω^K is defined in (3-46).

Proof. We note that $\sum_{y_i \in K} (\chi_i)' = 0$ in K, and then we can fix $y_j \in K$ and write $(\chi_j)' = -\sum_{y_i \in K \setminus \{y_j\}} (\chi_i)'$.
We obtain,

$$\begin{aligned} \left(\sum_{y_i \in K} (v - I_{L_i}^{\omega_i} v) \chi_i \right)' &= \sum_{y_i \in K} (\chi_i)' (v - I_{L_i}^{\omega_i} v) + \sum_{y_i \in K} \chi_i (v - I_{L_i}^{\omega_i} v)' \\ &= \sum_{y_i \in K \setminus \{y_j\}} (I_{L_i}^{\omega_i} v - I_{L_j}^{\omega_j} v) \chi_i' + \sum_{y_i \in K} \chi_i (v - I_{L_i}^{\omega_i} v)', \end{aligned} \tag{3-28}$$

which gives the following bound valid on K,

$$\left|\left(\sum_{y_i\in K}(v-I_{L_i}^{\omega_i}v)\chi_i\right)'\right|^2 \leq \frac{1}{H^2}\sum_{y_i\in K\setminus\{y_j\}}(I_{L_i}^{\omega_i}v-I_{L_j}^{\omega_j}v)^2+\sum_{y_i\in K}|(v-I_{L_i}^{\omega_i}v)'|^2. \tag{3-29}$$

From (3-29) we get

$$\begin{aligned}\int_K \kappa|(v-I_{L_i}v)'|^2 &\leq \int_K \kappa\left|\left(\sum_{y_i\in K}(v-I_{L_i}^{\omega_i}v)\chi_i\right)'\right|^2\\ &\leq \sum_{y_i\in K}\frac{1}{H^2}\int_K \kappa(I_{L_i}^{\omega_i}v-I_{L_j}^{\omega_j}v)^2+\sum_{y_i\in K}\int_K\kappa|(v-I_{L_i}^{\omega_i}v)'|^2.\end{aligned} \tag{3-30}$$

To bound the first term, use Theorem 12 as follows,

$$\begin{aligned}\int_K \kappa(I_{L_i}^{\omega_i}v-I_{L_j}^{\omega_j}v)^2 &\leq \int_{\omega_i}\kappa(v-I_{L_i}^{\omega_i}v)^2+\int_{\omega_j}\kappa(v-I_{L_i}^{\omega_i}v)^2\\ &\leq \frac{1}{(\sigma_{L+1}^i)^2}\|v-I_{L_i}^{\omega_i}v\|_{2,\omega_i}^2+\frac{1}{(\sigma_{L+1}^j)^2}\|v-I_{L_j}^{\omega_j}v\|_{2,\omega_j}^2\\ &\leq \frac{1}{(\sigma_{K,L+1})^2}\sum_{y_i\in K}\|v-I_{L_i}^{\omega_i}v\|_{2,\omega_i}^2.\end{aligned} \tag{3-31}$$

By Lemma 8, we have

$$\int_{\omega_i}\kappa\,|(v-I_L^{\omega_i})'|^2 \leq \frac{1}{\sigma^{1-s}}\|u\|_{s,\omega_i}^2. \tag{3-32}$$

The second term in (3-30) is estimated using (3-32)

$$\int_K\kappa|(v-I_{L_i}^{\omega_i}v)'|^2 \leq \int_{\omega_i}\kappa|(v-I_{L_i}^{\omega_i}v)'|^2 \leq \frac{1}{\sigma_{L+1}^{\omega_i}}\|v-I_{L_i}^{\omega_i}v\|_{2,\omega_i}^2. \tag{3-33}$$

By combining (3-31), (3-33) and (3-30) we obtain (3-27). □

The case of Dirichlet boundary conditions on K is presented next. First, we define the following interpolation,

$$I_D u=\sum_{K\in\mathcal{T}^H}\sum_{\ell=1}^{L_K}m^K(v,\psi_\ell)\psi_\ell. \tag{3-34}$$

The following lemma a is direct consequence of Lemma 7.

Lemma 4. *Assume that $f=(\kappa u')'\in L^2(\Omega)$ and also assume that for each $K\in\mathcal{T}^H$, we have $u=0$ on ∂K. Then, the following energy approximation holds,*

$$\int_K\kappa|u'-(I_D u)'|^2\leq\frac{1}{\mu_{L+1}^K}\|f\|_{L^2(K)}^2.$$

The idea is to apply I_N to u_N, and the other part, which is u_D, will be approximated by a truncated expansion on V_D. This truncated expansion will approximate u_D. We define the coarse interpolation I_0 by

$$\begin{aligned} I_0 u &= \sum_{i=1}^{N_c}\sum_{\ell=1}^{L_i}\left(\int_{\omega_i}\kappa u_N\varphi_\ell^{\omega_i}\right)\chi_i\varphi_\ell^{\omega_i} + \sum_{K\in\mathcal{T}^H}\sum_{\ell=1}^{L_K}\left(\int_K \kappa u_D\psi_\ell\right)\psi_\ell \\ &= \sum_{i=1}^{N_c}(I_{L_i}^{\omega_i}u_N)\chi_i + \sum_{K\in\mathcal{T}^H}(I^K u_D) = I_N u_N + I_D u_D. \end{aligned} \tag{3-35}$$

Recall that L_i is the number of Newmann eigenfunctions considered in the neighborhood ω_i and L_K is the number of Dirichlet eigenfunctions considered on the element K. We finally present our main approximation result.

Assumption 15. *Let u be the exact solution; that is $-(\kappa u')' = f$. We assume that there exist u_D, u_N and $\epsilon \preceq H$ such that*

1. *We have*
$$\int_\Omega \kappa|(u-u_D-u_N)'|^2 \preceq \epsilon^2\int_\Omega \kappa^{-1}f^2 \preceq H^2\int_\Omega \kappa^{-1}f^2. \tag{3-36}$$
2. *We have the boundary data given by*
$$\partial_\eta u_N = 0 \text{ on } \partial K \text{ for all } K.$$
and
$$u_D = 0 \text{ on } \partial K \text{ for all } K.$$
3. *We have the bounds,*
$$\int_\Omega \kappa^{-1}|(\kappa u_N')'|^2 \preceq \int_\Omega \kappa^{-1}|(\kappa u')'|^2 \text{ and } \int_\Omega \kappa^{-1}|(\kappa u_D')'|^2 \preceq \int_\Omega \kappa^{-1}|(\kappa u')'|^2. \tag{3-37}$$

Theorem 16. *Assume that $|||u|||_{2,\Omega} < \infty$ where u is the solution of (3-43). Then, the following approximation for the energy interpolation error holds,*

$$\int_\Omega \kappa|u'-(I_0u)'|^2 \le \left(\max\left\{\frac{1}{H^2(\sigma_{L+1})^2}, \frac{1}{\sigma_{L+1}}\right\} + \frac{1}{\mu_{L+1}} + \epsilon^2\right)|||\kappa^{-1/2}f|||_0^2, \tag{3-38}$$

where $\sigma_{L+1} = \min_K \sigma_{K,L+1}^{\omega_i}$ and $\mu_{L+1} = \min_K \mu_{L+1}^K$.

Proof. By definition (3-35), our technical assumption, and using the triangular inequality, we have

$$\begin{aligned} \int_\Omega \kappa|u'-(I_0u)'|^2 &= \int_\Omega \kappa|(u+u_N+u_D-(u_N+u_D))' - (I_Nu_N+I_Du_D)'|^2 \\ &\le \int_\Omega \kappa|(u_N-I_Nu_N)'|^2 + \int_\Omega \kappa|(u_D-I_Du_D)'|^2 + \int_\Omega \kappa|(u-u_N-u_D)'|^2. \end{aligned} \tag{3-39}$$

Using Lemma 3 and noting that $\bigcup K \subset \bigcup \omega_K$, we have for the first term in (3-39),

$$\begin{aligned}\int_\Omega \kappa|(u_N - I_N u_N)'|^2 &= \sum_{K\in\mathcal{T}^H}\int_K \kappa|(u_N - I_N u_N)'|^2 \\ &\leq \sum_{\omega_K} \max\left\{\frac{1}{H^2\sigma^2_{K,L+1}}, \frac{1}{\sigma_{K,L+1}}\right\} ||\kappa^{-1}(\kappa u'_N)'||^2_{L^2(\omega_K)} \\ &\leq \max_{K\in\mathcal{T}^H}\left\{\max\left\{\frac{1}{H^2\sigma^2_{K,L+1}}, \frac{1}{\sigma_{K,L+1}}\right\}\right\} \sum_{\omega_K} ||\kappa^{-1}(\kappa u'_N)'||^2_{L^2(\omega_K)} \\ &\leq \max\left\{\frac{1}{H^2\sigma^2_{L+1}}, \frac{1}{\sigma_{L+1}}\right\} ||\kappa^{-1}(\kappa u'_N)'||^2_{L^2(\Omega)}. \end{aligned} \tag{3-40}$$

Now using Theorem 14 for the second term in (3-39), we get

$$\begin{aligned}\int_\Omega \kappa|(u_D - I_D u_D)'|^2 &= \sum_{K\in\mathcal{T}^H}\int_K \kappa|(u_D - I_D u_D)'|^2 \leq \sum_{K\in\mathcal{T}^H} \frac{1}{\mu^K_{L+1}} ||\kappa^{-1}(\kappa u'_D)'||^2_{L^2(K)} \\ &\leq \max_{K\in\mathcal{T}^H}\left\{\frac{1}{\mu_{K,L+1}}\right\} \sum_{K\in\mathcal{T}^H} ||\kappa^{-1}(\kappa u'_D)'||^2_{L^2(K)} \\ &\leq \frac{1}{\mu_{L+1}} ||\kappa^{-1}(\kappa u'_D)'||^2_{L^2(\Omega)}. \end{aligned} \tag{3-41}$$

Given that u is the exact solution, that is $-(\kappa u')' = f$, and with Assumption 20, we obtain the bounds for the third term in equation (3-39)

$$\int_\Omega \kappa|(u - u_N - u_D)'|^2 \leq \epsilon^2 \int_\Omega \kappa^{-1} f^2. \tag{3-42}$$

With (3-42) and (3-41), we obtain (3-38) from (3-39). □

3.3 GMsFEM in two dimensions

Let $\Omega \subset \mathbb{R}^n$ and $\mathcal{T}^H = \{K\}$ a coarse-grid partition, where H denotes the size of the coarse grid. Then, we consider the elliptic equation with heterogeneous coefficients given by

$$-\text{div}\,(\kappa(x)\nabla u) = f, \tag{3-43}$$

where $\kappa(x)$ is a heterogeneous field with high contrast and multiscale behavior. In particular, we assume that $\kappa(x) \geq C_0 > 0$, while $\kappa(x)$ can have very large values and local variation. The variational formulation of this problem is: Find $u \in H^1_0(\Omega)$ such that

$$a(u, v) = f(v) \text{ for all } v \in H^1_0(\Omega), \tag{3-44}$$

where the bilinear form a and the linear functional F is defined by

$$a(u,v) = \int_D \kappa(x)\nabla u(x)\nabla v(x)dx \text{ for all } u, v \in H_0^1(\Omega),$$

and

$$F(v) = \int_D f(x)v(x)dx, \text{ for all } u, v \in H_0^1(\Omega).$$

Let $\mathcal{T}^h$ be a fine triangulation, which is a refinement of coarse mesh $\mathcal{T}^H$. We denote by $V^h(\Omega)$ the finite element discretization of piecewise linear continuous functions concerning the fine triangulation $\mathcal{T}^h$. Denote also by $V_0^h(\Omega)$ the subset of $V^h(\Omega)$ where the boundary is zero.
The Galerkin formulation of the problem is to find $u \in V_0^h(\Omega)$ such that

$$a(u,v) = F(v) \text{ for all } v \in V_0^h(\Omega),$$

or in matrix form

$$v^t A u = v^t b,$$

where for all $u, v \in V^h(D)$ we have

$$u^t A v = \int_D \kappa(x)\nabla u \cdot \nabla v \; dx \quad \text{and} \quad v^t b = \int_D fv \; dx.$$

We denote by $\{y_i\}_{i=1}^{N_v}$ the vertices of the coarse mesh $\mathcal{T}^H$ and define the neighborhood of the node y_i by

$$\omega_i = \bigcup \left\{ K_j \in \mathcal{T}^H; y_i \in \overline{K} \right\}, \tag{3-45}$$

and the neighborhood of the coarse element K by

$$\omega_K = \bigcup \left\{ \omega_i \in \mathcal{T}^H; y_i \in \overline{K} \right\}, \tag{3-46}$$

see Figure **3-3**.

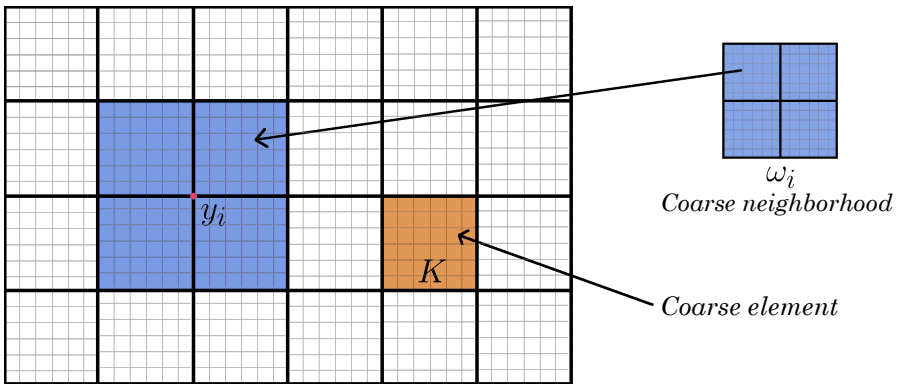

Figure 3-3: Ilustration of a coarse neighborhood.

Using the coarse mesh $\mathcal{T}^H$, we introduce coarse basis functions $\{\phi_i\}_{i=1}^L$, where L is the number of coarse basis functions. The MsFEMs approximate the solution on a coarse grid as $u_0 = \sum_i c_i\phi_i$, where c_i are determined from

$$a(u_0, v) = f(v), \text{ for all } v \in \text{span}\,\{\phi_i\}_{i=1}^{N_c}.$$

3.3.1 Global eigenvalue problem

We consider the following high contrast eigenvalue problem

$$-\text{div}(\kappa\nabla\phi) = \lambda\kappa\phi, \tag{3-47}$$

where Ω is a domain with homogeneous Dirichlet boundary condition on $\partial\Omega$. The weak form of the eigenvalue problem (3-47) is: Find eigenfunctions ϕ and scalar λ such that

$$a(\phi, z) = \lambda m(\phi, z) \quad \text{for all } z \in H_0^1(D). \tag{3-48}$$

Denote its eigenvalues and eigenfuntions by $\{\lambda_\ell\}$ and $\{\phi_\ell\}$, respectively. We order the eigenvalues as

$$\lambda_1 \leq \lambda_2 \leq ... \leq \lambda_\ell \leq ... \tag{3-49}$$

Note that the eigenvectors form an orthonormal basis of $L^2(\Omega)$ with respect to M^Ω and is also orthogonal with respect to the bilinear form a. Given any $v \in H_0^1(\Omega)$, we have

$$v = \sum_{\ell=1}^{\infty}\left(\int_\Omega \kappa v\phi_\ell\right)\phi_\ell,$$

$$a(v, v) = \int_\Omega \kappa(\nabla v)^2 = \sum_{\ell=1}^{\infty}\left(\int_\Omega \kappa v\phi_\ell\right)^2\lambda_\ell, \tag{3-50}$$

and

$$m(v, v) = \int_\Omega \kappa(v)^2 = \sum_{\ell=1}^{\infty}\left(\int_\Omega \kappa v\phi_\ell\right)^2. \tag{3-51}$$

Using (3-44) and (3-48) for each eigenvector function ϕ_ℓ, we obtain

$$\lambda_\ell m(u, \phi_\ell) = a(u, \phi_\ell) = f(\phi_\ell).$$

Then we have

$$u = \sum_{\ell=1}^{\infty}\frac{1}{\lambda_\ell}f(\phi_\ell)\phi_\ell. \tag{3-52}$$

We introduce a norm for studying the convergence of the expansions in terms of eigenvectors. For any $v \in L^2(\Omega)$ and $s > 0$, we introduce the norm $\|\cdot\|_{s,\Omega}$ defined by

$$\|v\|_{s,\Omega}^2 = \sum_{\ell=1}^{\infty}\lambda_\ell^s m(v, \phi_\ell)^2.$$

Note that if $s = 0$ we obtain (3-51) and if $s = 1$, the result is the bilinear form given in (3-50). Mainly we use the norm $\|\cdot\|_{s,\Omega}$ with $s = 2$. We have that

$$\|v\|^2_{2,\Omega} = \sum_{\ell=1}^{\infty} \lambda_\ell^2 m(v,\phi_\ell)^2.$$

We define the operator Υ applied to u by

$$\Upsilon u = \sum_{\ell=1}^{\infty} \lambda_\ell m(u,\phi_\ell).$$

Given that $\|v\|_{2,\Omega} < \infty$, we obtain that

$$\|\Upsilon u\|^2_{0,\Omega} = m(\Upsilon u, \Upsilon u) = \sum_{\ell=0}^{\infty} \lambda_\ell^2 m(u,\phi_\ell)^2 = \|u\|^2_{2,\Omega} < \infty.$$

We also have that if $\|u\|_{2;\Omega} < \infty$ then using integration by parts, we get

$$a(u,v) = \sum_{\ell=1}^{\infty} \lambda_\ell m(u,\phi_\ell) m(v,\phi_\ell) = m(\Upsilon u, v) \quad \text{for all } v \in H^1_0(\Omega). \tag{3-53}$$

Theorem 17 ([1], Theorem 2)**.** *The operator Υ is a locally defined operator. More precisely, if $\|u\|_{2;\Omega} < \infty$, we have that $-\kappa^{-1} div(\kappa\nabla u) \in L^2(\Omega)$ and*

$$\Upsilon u = -\kappa^{-1}\mathrm{div}(\kappa\nabla u). \tag{3-54}$$

Moreover,

$$\|u\|^2_{2;\Omega} = \|\Upsilon u\|^2_0 = \int \kappa^{-1} |\mathrm{div}(\kappa\nabla u)|^2. \tag{3-55}$$

Proof. Recall that for $u \in L^2(\Omega)$, we have the expansion $u = \sum_\ell^\infty m(u,\phi_\ell)\phi_\ell$. For an integer N define, the truncated approximation of u as

$$u^N = \sum_{\ell=1}^{N} m(u,\phi_\ell)\phi_\ell.$$

We construct $-\kappa^{-1}\mathrm{div}(\kappa\nabla u)$ operator as a limit in the m-norm of the sequence of rescaled divergences given by $\left\{\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\right\}_{N=1}^{\infty}$. To this end, we prove that the sequence $\left\{\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\right\}_{N=1}^{\infty}$ is a Cauchy sequence in the m-norm. Indeed, by using the eigenvalue problem (3-48), we have the following identity:

$$\kappa^{-1}\mathrm{div}\left(\kappa\nabla u^N\right) = \sum_{\ell=1}^{N} m(u,\phi_\ell)\lambda_\ell\phi_\ell.$$

Using the orthogonality of the eigenvectors, we conclude that for every $M > N$ we have,

$$\left\|\kappa^{-1}\mathrm{div}(\kappa\nabla u^M) - \kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\right\|^2_0 = \sum_{\ell=N+1}^{m} m(u,\phi_\ell)^2\lambda_\ell^2 < \infty.$$

We conclude that there exists an $L^2(\Omega)$ function, denoted by U, such that $\|U+\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\| \to 0$ when $N \to \infty$. We also have that for any $z \in H_1^0(\Omega)$,

$$\int_\Omega \kappa U z = -\lim_{N\to\infty}\int_\Omega \kappa\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)z = \int_\Omega \kappa\nabla u\nabla z,$$

which proves that $U = -\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)z$.
Finally note that, by using (3-63), for every function $z \in H_0^1(\Omega)$, we have,

$$\int_\Omega \kappa\Upsilon u z = m(\Upsilon u, v) = a(u, z) = \int_\Omega \kappa\nabla u\nabla z.$$

□

Lemma 5 ([1], Lemma 4)**.** *Assume that $\|f\|_s^2 < \infty$ and let u be the solution of* (3-44)*. Consider $t \geq 1$ such that $t - s - 2 \leq 0$. We have that*

$$\|u\|_t^2 \leq \lambda_1^{t-s-2}\|f\|_s^2. \tag{3-56}$$

In particular, if $\|f\|_0 < \infty$, we have that $\|u\|_2 = \|\kappa^{-1}f\|_0 = \int_\Omega \kappa^{-1}f^2$.

Proof. Using the explicit expansion given in (3-52), along with the definition of the norm $|\cdot|_t$, and subsequently ordering the eigenvalues from least to greatest, we obtain

$$\begin{aligned}
\|u\|_t^2 &= \sum_{\ell=1}^{\infty}\lambda_\ell^t m(v,\phi_\ell)^2 \\
&= \sum_{\ell=1}^{\infty}\lambda_\ell^t\frac{1}{\lambda_\ell^2}f(\phi_\ell)^2 \\
&= \sum_{\ell=1}^{\infty}\lambda_\ell^{t-s-2}\lambda_\ell^s f(\phi_\ell)^2 \\
&\leq \lambda_1^{t-s-2}\sum_{\ell=1}^{\infty}\lambda_\ell^s f(\phi_\ell)^2 \\
&= \lambda_1^{t-s-2}\sum_{\ell=1}^{\infty}\lambda_\ell^s\left(\int_\Omega f\phi\right)^2 \\
&= \lambda_1^{t-s-2}\sum_{\ell=1}^{\infty}\lambda_\ell^s\left(\int_\Omega \kappa(\kappa^{-1}f\phi)\right)^2 = \lambda_1^{t-s-2}\|\kappa^{-1}f\|_s^2.
\end{aligned}$$

□

3.3.2 Convergence using global eigenvectors.

In this section, we study the convergence if we use any solution in the space spanned by the first eigenvector.

Given an integer L and $v \in H_0^1(\Omega)$, we define

$$I_L v = \sum_{\ell=1}^{L} m(v, \phi_\ell)\phi_\ell,$$

which is the interpolation of v for the first L eigenvectors. Using equations (3-49), (3-50), and (3-51), we can derive the following expression

$$\begin{aligned}
\int_\Omega \kappa(v - I_L v)^2 &= \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v - I_L v)\phi_\ell \right)^2 \\
&= \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v - I_L v)\phi_\ell \right)^2 \frac{\lambda_\ell}{\lambda_\ell} \\
&\leq \frac{1}{\lambda_{L+1}} \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v - I_L v)\phi_\ell \right)^2 \lambda_\ell \\
&\leq \frac{1}{\lambda_{L+1}} \sum_{\ell=1}^{\infty} \left(\int_{\omega_i} \kappa(v)\phi_\ell \right)^2 \lambda_\ell \\
&= \frac{1}{\lambda_{L+1}} \int_{\omega_i} \kappa \left(\nabla v\right)^2 .
\end{aligned}$$

We obtain the following inequality:

$$\int_\Omega \kappa(v - I_L v)^2 \leq \frac{1}{\lambda_{L+1}} a(v - I_L v, v - I_L v) \leq \frac{1}{\lambda_{L+1}} a(v, v). \tag{3-57}$$

Lemma 6 ([1], Lemma 5)**.** *Let u be the solution of* (3-44)*. If $t - s - 2 < 0$, we have*

$$\|u - I_L u\|_t^2 \leq \lambda_{L+1}^{t-s-2} \|f\|_s^2. \tag{3-58}$$

Proof. Using the explicit expansion in equation (3-52), the definition of the norm $\|\cdot\|_t$, and the ordering of the eigenvalues, we obtain the following

$$\begin{aligned}
\|u - I_L u\|_t^2 = \sum_{\ell=L+1}^{\infty} \lambda_\ell^t \frac{1}{\lambda_\ell^2} f(\phi_\ell)^2 &= \sum_{\ell=L+1}^{\infty} \lambda_\ell^{t-s-2} \lambda^s \left(\int_\Omega f\phi \right)^2 \\
&\leq \lambda_{L+1}^{t-s-2} \sum_{\ell=L+1}^{\infty} \lambda^s \left(\int_\Omega f\phi \right)^2 = \lambda_{L+1}^{t-s-2} \|f\|_s^2.
\end{aligned} \tag{3-59}$$

□

Using Lemma 6 with $t = 1$ and equation (3-57), we obtain the following a priori error estimate

$$a(u - I_L u, u - I_L u) = \|u - I_L u\|_1^2 \leq \lambda_{L+1}^{-(s+2)} \|f\|_s^2.$$

3.3.3 Dirichlet eigenvalue problem in coarse blocks

In this section, we study the eigenvalue problem associated to problem (3-44). For any K, we define the bilinear forms

$$a^K(v,w) = \int_K \kappa \nabla v \nabla w, \quad \text{for all } v, w \in H_0^1(K),$$

and

$$m^K(v,w) = \int_K \kappa v w, \quad \text{for all } v, w \in H_0^1(K).$$

We consider the eigenvalue problem

$$a^K(u,z) = \mu m^K(\psi, z), \quad \text{for all } z \in H_0^1(K),$$

where we denote the eigenvalues and eigenvectors by $\{\mu_\ell\}$ and $\{\psi_\ell\}$ respectively.
We order the eigenvalues as

$$\mu_1 \leq \mu_2 \leq ... \leq \mu_\ell \leq ...$$

The eigenvalue problem above corresponds to the approximation of the eigenvalue problem

$$-\text{div}(\kappa \nabla \psi) = \mu \kappa \psi \quad \text{in } K,$$

with homogeneous Dirichlet boundary condition on ∂K. Given any $v \in L^2(K)$, we can write

$$v = \sum_{\ell=1}^{\infty} m^K(v, \psi_\ell)\psi_\ell,$$

and compute the local energy bilinear form by

$$a^K(v,v) = \sum_{\ell=1}^{\infty} m^K(v, \psi_\ell)^2 \mu_\ell.$$

The local mass bilinear form is

$$m^K(v,v) = \sum_{\ell=1}^{\infty} m^K(v, \psi_\ell)^2.$$

We introduce the local norm

$$\|v\|_{s,K}^2 = \sum_{\ell=1}^{\infty} (\mu_\ell)^s m^K(v, \psi_\ell)^2.$$

We consider the case $s = 2$. If $\|u\|_{2,K} < \infty$ we can define the operator Υ^K by,

$$\Upsilon^K(u) = \sum_{\ell=1}^{\infty} \mu_\ell m^K(u, \psi_\ell)^2 \psi_\ell,$$

which is square integrable since

$$\|\Upsilon^K(u)\|_{0,K}^2 = \sum_{\ell=1}^{\infty}(\mu_\ell)^2 m^K(u,\psi_\ell)^2 = \|u\|_{2,K}^2.$$

We also have that if $\|u\|_{2;\Omega} < \infty$ then we have,

$$a(u,v) = m(\Upsilon u, v) \quad \text{for all } v \in H_0^1(\Omega). \tag{3-60}$$

Theorem 18 ([1], Theorem 6)**.** *The operator Υ^K is a locally defined operator. Specifically, if $|u|_{2;K} < \infty$, we have that $-\kappa^{-1} div(\kappa\nabla u) \in L^2(K)$, and we have:*

$$\Upsilon^K(u,v) = -\int_K \kappa^{-1}\mathrm{div}(\kappa\nabla u)v \quad \textit{for all } v \in H_0^1(K). \tag{3-61}$$

Moreover, we have

$$\|u\|_{2;K}^2 = \|\Upsilon^K u\|_{0,K}^2 = \int_K \kappa^{-1}|\mathrm{div}(\kappa\nabla u)|^2. \tag{3-62}$$

Proof. Recall that for $u \in L^2(\Omega)$, we have the expansion $u = \sum_{\ell=1}^{\infty} m^K(u,\psi_\ell)\psi_\ell$. For an integer N define the truncated approximation of u as,

$$u^N = \sum_{\ell=1}^{N} m^K(u,\psi_\ell)\psi_\ell.$$

We construct the operator $-\kappa^{-1}\mathrm{div}(\kappa\nabla u)$ as a limit in the m-norm of the sequence of rescaled divergences given by $\left\{\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\right\}_{N=1}^{\infty}$. To this end, we prove that the sequence $\left\{\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\right\}_{N=1}^{\infty}$ is a Cauchy sequence in the m-norm. Indeed, we have, by using the eigenvalue problem (3-48), the following identity

$$\kappa^{-1}\mathrm{div}\left(\kappa\nabla u^N\right) = \sum_{\ell=1}^{N} m^K(u,\psi_\ell)\mu_\ell\psi_\ell.$$

Using the orthogonality of the eigenvectors, we conclude that for every $M > N$ we have,

$$\left\|\kappa^{-1}\mathrm{div}(\kappa\nabla u^M) - \kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\right\|_0^2 = \sum_{\ell=N+1}^{m} m^K(u,\phi_\ell)^2\mu_\ell^2 < \infty.$$

We conclude that there exists an $L^2(\Omega)$ function, denoted by U, such that we have $\|U + \kappa^{-1}\mathrm{div}(\kappa\nabla u^N)\| \to 0$, when $N \to \infty$. We also have that for any $z \in H_1^0(\Omega)$

$$\int_K \kappa U z = -\lim_{N\to\infty}\int_K \kappa\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)z = \int_K \kappa\nabla u\nabla z,$$

which proves that $U = -\kappa^{-1}\mathrm{div}(\kappa\nabla u^N)z$.
Finally note that, by using (3-63), for every function $z \in H_0^1(\Omega)$ we have,

$$\int_K \kappa \Upsilon^K uz = m^K(\Upsilon^K u, v) = a^K(u, v) = \int_K \kappa \nabla u \nabla z.$$

□

Given an integer L and $v \in H_0^1(K)$, We define

$$I_L^K v = \sum_{\ell=1}^{L} m^K(v, \psi_\ell)\psi_\ell.$$

Lemma 7 ([1], Lemma 7)**.** *Assume that $u \in H_0^1(K)$ and $\|u\|_{s,K} < \infty$ with $s > 1$. We have for $1 < s \leq t < 2$,*

$$\|u - I_L^K u\|_{t,K}^2 \leq (\mu_{L+1})^{t-s} \|u\|_{s,K}^2.$$

In particular,

$$\|u - I_L^K u\|_{1,K}^2 \leq (\mu_{L+1})^{t-s} \|u\|_{0,K}^2.$$

3.3.4 Neumann eigenvalue problem in coarse neighborhoods

In this section, we study the eigenvalue problem associated to problem (3-44). For any ω_i, we define the following bilinear forms

$$a^{\omega_i}(v, w) = \int_{\omega_i} \kappa \nabla v \nabla w, \quad \text{for all } v, w \in H_0^1(\omega_i),$$

and

$$m^{\omega_i}(v, w) = \int_{\omega_i} \kappa v w, \quad \text{for all } v, w \in H_0^1(\omega_i).$$

We define $V(\omega_i) = \{v \in H^1(\omega_i) : v = 0 \text{ on } \partial\omega_i \cap \partial\Omega\}$ if $\partial\omega_i \cap \partial\Omega \neq \emptyset$ and $V(\omega_i) = \{v \in H^1(\omega_i) : \int_{\omega_i} v = 0\}$. Otherwise, Consider the eigenvalue problem

$$a^{\omega_i}(u, z) = \sigma m^{\omega_i}(\varphi, z) \quad \text{for all } z \in V(\omega_i),$$

where we denote the eigenvalues and eigenvectors by $\{\sigma_\ell\}$ and $\{\varphi_\ell\}$ respectively.
We order the eigenvalues as

$$\sigma_1 \leq \sigma_2 \leq ... \leq \sigma_\ell \leq ...$$

The eigenvalue problem above corresponds to the approximation of the eigenvalue problem

$$-\text{div}(\kappa \nabla \varphi) = \sigma \kappa \varphi \quad \text{in } \omega_i,$$

with homogeneous Neumann boundary condition on $\partial\omega_i \cap \Omega$ and homogeneous Dirichlet boundary condition on $\partial\omega_i \cap \partial\Omega$. Given any $v \in L^2(\omega_i)$, we can write

$$v = \sum_{\ell=1}^{\infty} m^{\omega_i}(v, \varphi_\ell)\varphi_\ell,$$

and compute the local energy bilinear form by

$$a^{\omega_i}(v, v) = \sum_{\ell=1}^{\infty} m^{\omega_i}(v, \varphi_\ell)^2 \sigma_\ell.$$

The local mass bilinear form is

$$m^{\omega_i}(v, v) = \sum_{\ell=1}^{\infty} m^{\omega_i}(v, \varphi_\ell)^2.$$

We introduce the local norm

$$\|v\|^2_{s,\omega_i} = \sum_{\ell=1}^{\infty} (\sigma_\ell)^s m^{\omega_i}(v, \varphi_\ell)^2.$$

We consider the case $s = 2$. If $\|u\|_{2,\omega_i} < \infty$, we can define the operator Υ^{ω_i} by

$$\Upsilon^{\omega_i}(u) = \sum_{\ell=1}^{\infty} \sigma_\ell m^{\omega_i}(u, \varphi_\ell)^2 \varphi_\ell,$$

which is square integrable since

$$\|\Upsilon^{\omega_i}(u)\|^2_{0,\omega_i} = \sum_{\ell=1}^{\infty} (\sigma_\ell)^2 m^{\omega_i}(u, \varphi_\ell)^2 = \|u\|^2_{2,\omega_i}.$$

We also have that if $\|u\|_{2;\Omega} < \infty$ then

$$a^{\omega_i}(u, v) = m^{\omega_i}(\Upsilon^{\omega_i} u, v) \quad \text{for all } v \in H^1_0(\Omega). \tag{3-63}$$

Theorem 19 ([1], Theorem 8)**.** *The operator Υ^{ω_i} is a locally defined operator. More precisely, if $\|u\|_{2;\omega_i} < \infty$, we have that $-\kappa^{-1} div(\kappa \nabla u) \in L^2(\omega_i)$ and we have that*

$$\Upsilon^{\omega_i}(u, v) = -\int_{\omega_i} \kappa^{-1} \mathrm{div}(\kappa \nabla u) v, \quad \textit{for all } v \in H^1_0(\omega_i). \tag{3-64}$$

Moreover, we have

$$\|u\|^2_{2;\omega_i} = \|\Upsilon^{\omega_i} u\|^2_{0,\omega_i} = \int_{\omega_i} \kappa^{-1} |\mathrm{div}(\kappa \nabla u)|^2. \tag{3-65}$$

Proof. Recall that for $u \in L^2(\Omega)$, we have the expansion $u = \sum_{\ell=1}^{\infty} m^{\omega_i}(u, \varphi_\ell)\varphi_\ell$. For an integer N define the truncated approximation of u as,

$$u^N = \sum_{\ell=1}^{N} m^{\omega_i}(u, \varphi_\ell)\varphi_\ell.$$

We construct $-\kappa^{-1}\text{div}(\kappa\nabla u)$ as a limit in the m-norm of the sequence of rescaled divergences given by $\left\{\kappa^{-1}\text{div}(\kappa\nabla u^N)\right\}_{N=1}^{\infty}$. To this end, we prove that the sequence $\left\{\kappa^{-1}\text{div}(\kappa\nabla u^N)\right\}_{N=1}^{\infty}$ is a Cauchy sequence in the m-norm. Indeed, we have, by using the eigenvalue problem (3-48), the following identity

$$\kappa^{-1}\text{div}\left(\kappa\nabla u^N\right) = \sum_{\ell=1}^{N} m^{\omega_i}(u,\varphi_\ell)\sigma_\ell\varphi_\ell.$$

Using the orthogonality of the eigenvectors, we conclude that, for every $M > N$,

$$\left\|\kappa^{-1}\text{div}(\kappa\nabla u^M) - \kappa^{-1}\text{div}(\kappa\nabla u^N)\right\|_0^2 = \sum_{\ell=N+1}^{m} m^{\omega_i}(u,\phi_\ell)^2\sigma_\ell^2 < \infty.$$

We conclude that there exists an $L^2(\Omega)$ function, denoted by U, such that we have $\|U + \kappa^{-1}\text{div}(\kappa\nabla u^N)\| \to 0$ when $N \to \infty$. We also have that for any $z \in H_1^0(\Omega)$

$$\int_{\omega_i} \kappa U z = -\lim_{N\to\infty}\int_{\omega_i} \kappa\kappa^{-1}\text{div}(\kappa\nabla u^N)z = \int_{\omega_i}\kappa\nabla u\nabla z,$$

which proves that $U = -\kappa^{-1}\text{div}(\kappa\nabla u^N)z$.
Finally note that, by using (3-63), for every function $z \in H_0^1(\Omega)$ we have

$$\int_{\omega_i}\kappa\Upsilon^{\omega_i}uz = m^{\omega_i}(\Upsilon^{\omega_i}u, v) = a^{\omega_i}(u,v) = \int_{\omega_i}\kappa\nabla u\nabla z.$$

□

Given an integer L and $v \in H_0^1(\omega_i)$, we define

$$I_L^{\omega_i}v = \sum_{\ell=1}^{L} m^{\omega_i}(v,\varphi_\ell)\varphi_\ell.$$

Lemma 8 ([1], Lemma 10)**.** *Assume that $u \in H_0^1(\omega_i)$ and $\|u\|_{s,\omega_i} < \infty$ with $s > 1$. We have for $1 < s \leq t < 2$,*

$$\|u - I_L^{\omega_i}u\|_{t,\omega_i}^2 \leq (\sigma_{L+1})\|u\|_{s,\omega_i}^2.$$

In particular,

$$\|u - I_L^{\omega_i}u\|_{1,\omega_i}^2 \leq (\sigma_{L+1})^{t-s}\|u\|_{0,\omega_i}^2.$$

3.3.5 The GMsFEM space construction using local eigenvalue problems

In this section, we show the construction of the coarse scale finite element spaces using the GMsFEM framework.

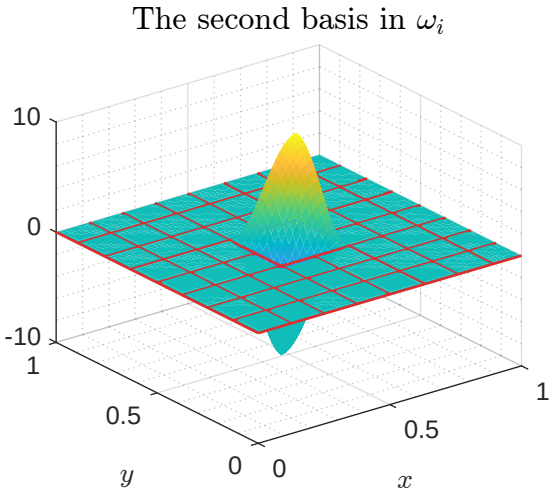

Figure 3-4: Schematic description of basis function construction.

We choose the basis functions that span the eigenfunctions corresponding to small eigenvalues. We consider $\{\omega_i\}_{y_i \in \mathcal{T}^H}$, $i = 1, 2, 3, ..., N_v$. Define the set of coarse basis functions

$$\Phi_{i,\ell} = \chi_i \varphi_\ell^{\omega_i}, \qquad \text{for } 1 \leq i \leq N_v \text{ and } 1 \leq \ell \leq L_i, \tag{3-66}$$

where L_i is the number of eigenvalues that will be chosen for the node i. The local spectral multiscale space is

$$V_N = \text{span}\{\Phi_{i,\ell} : 1 \leq i \leq N_v \text{ and } 1 \leq \ell \leq L_i\},$$

$$V_D = \text{span}\{\psi_\ell : K \in \mathcal{T}^H \text{ and } 1 \leq \ell \leq L_K\},$$

and

$$V_0 = V_N + V_D.$$

Finally, we define u as the Galerkin approximation using the space V_0, that is,

$$a(u_H, v) = f(v), \qquad \text{for all } v \in V_0. \tag{3-67}$$

3.3.6 Approximation properties of the coarse space

An important property of the coarse mesh and its approximation study is that if we have a high contrast medium κ and if we choose a large enough L_i, then $\lambda_{L_i}^{\omega_i}$ is independent of the contrast;which can be see (3-57).

We define

$$I_N u = \sum_{i=1}^{N_v} \sum_{\ell=1}^{L_i} \left(\int_{\omega_i} \kappa u \varphi_\ell^{\omega_i} \right) \chi_i \varphi_\ell^{\omega_i} = \sum_{i=1}^{N_v} (I_{L_i}^{\omega_i} u_N) \chi_i.$$

The following lemma presents the approximation result in the norm H^1. It is done based on results exposed in [1, 57].

Lemma 9 ([1], Lemma 16)**.** *Assume that $f = \operatorname{div}(\kappa \nabla u) \in L^2(\omega_K)$ and also assume that for each i, $y_i \in K$ we have $\partial_\eta u = 0$ on $\partial\omega_i$. Then, the following energy approximation holds,*

$$\int_K \kappa |\nabla u - \nabla I_N u|^2 \leq \max \left\{ \frac{1}{H^2 \sigma_{K,L+1}^2}, \frac{1}{\sigma_{K,L+1}} \right\} ||f||^2_{L^2(\omega_K)}, \tag{3-68}$$

where $\sigma_{K,L+1} = \min_{y_i \in K} \sigma_{L_i+1}$ and ω^K is defined in (3-46)*.*

Proof. We note that $\sum_{y_i \in K} \nabla \chi_i = 0$ in K, and then we can fix $y_j \in K$ and write $\nabla \chi_j = -\sum_{y_i \in K \setminus \{y_j\}} \nabla \chi_i$.
We obtain,

$$\begin{aligned} \nabla \sum_{y_i \in K} (v - I_{L_i}^{\omega_i} v) \chi_i &= \sum_{y_i \in K} \nabla \chi_i (v - I_{L_i}^{\omega_i} v) + \sum_{y_i \in K} \chi_i \nabla (v - I_{L_i}^{\omega_i} v) \\ &= \sum_{y_i \in K \setminus \{y_j\}} (I_{L_i}^{\omega_i} v - I_{L_j}^{\omega_j} v) \nabla \chi_i + \sum_{y_i \in K} \chi_i \nabla (v - I_{L_i}^{\omega_i} v), \end{aligned} \tag{3-69}$$

which gives the following bound valid on K,

$$\left| \nabla \sum_{y_i \in K} (v - I_{L_i}^{\omega_i} v) \chi_i \right|^2 \leq \frac{1}{H^2} \sum_{y_i \in K \setminus \{y_j\}} (I_{L_i}^{\omega_i} v - I_{L_j}^{\omega_j} v)^2 + \sum_{y_i \in K} |\nabla (v - I_{L_i}^{\omega_i} v)|^2. \tag{3-70}$$

From (3-70), we get

$$\begin{aligned} \int_K \kappa |\nabla (v - I_{L_i} v)|^2 &\leq \int_K \kappa |\nabla \sum_{y_i \in K} (v - I_{L_i}^{\omega_i} v) \chi_i|^2 \\ &\leq \sum_{y_i \in K} \frac{1}{H^2} \int_K \kappa (I_{L_i}^{\omega_i} v - I_{L_j}^{\omega_j} v)^2 + \sum_{y_i \in K} \int_K \kappa |\nabla (v - I_{L_i}^{\omega_i} v)|^2. \end{aligned} \tag{3-71}$$

To bound the first term, use Lemma 8 as follows,

$$\begin{aligned} \int_K \kappa (I_{L_i}^{\omega_i} v - I_{L_j}^{\omega_j} v)^2 &\leq \int_{\omega_i} \kappa (v - I_{L_i}^{\omega_i} v)^2 + \int_{\omega_j} \kappa (v - I_{L_i}^{\omega_i} v)^2 \\ &\leq \frac{1}{(\sigma_{L+1}^i)^2} \|v - I_{L_i}^{\omega_i} v\|^2_{2,\omega_i} + \frac{1}{(\sigma_{L+1}^j)^2} \|v - I_{L_j}^{\omega_j} v\|^2_{2,\omega_j} \\ &\leq \frac{1}{(\sigma_{K,L+1})^2} \sum_{y_i \in K} \|v - I_{L_i}^{\omega_i} v\|^2_{2,\omega_i}. \end{aligned} \tag{3-72}$$

Also, according to Lemma 8, if $\partial_n v = 0$ on $\partial\omega_i$, we have

$$\int_{\omega_i} \kappa \, |\nabla(v - I_L^{\omega_i})|^2 \leq \frac{1}{(\sigma_{L+1}^{\omega_i})^{1-s}} \|u\|_{s,\omega_i}^2. \tag{3-73}$$

The second term in (3-71) is estimated using (3-73)

$$\int_K \kappa |\nabla(v - I_{L_i}^{\omega_i} v)|^2 \leq \int_{\omega_i} \kappa |\nabla(v - I_{L_i}^{\omega_i} v)|^2 \leq \frac{1}{\sigma_{L+1}^{\omega_i}} \|v - I_{L_i}^{\omega_i} v\|_{2,\omega_i}^2. \tag{3-74}$$

By combining (3-72), (3-74) and (3-71), we obtain (3-68). □

The case of Dirichlet boundary conditions on K is presented next. First, we define the following interpolation,

$$I_D u = \sum_{K \in \mathcal{T}^H} \sum_{\ell=1}^{L_K} m^K(v, \psi_\ell) \psi_\ell. \tag{3-75}$$

The following lemma is direct consequence of Lemma 7.

Lemma 10 ([1], Lemma 17)**.** *Assume that $f = \operatorname{div}(\kappa \nabla u) \in L^2(\Omega)$ and also assume that for each $K \in \mathcal{T}^H$, we have $u = 0$ on ∂K. Then, the following energy approximation holds,*

$$\int_K \kappa |\nabla u - \nabla I_D u|^2 \leq \frac{1}{\mu_{L+1}^K} \|f\|_{L^2(K)}^2.$$

The idea is to apply I_N to u_N, and the other part, which is u_D, will be approximated by a truncated expansion on V_D. We define the coarse interpolation I_0 as follows

$$\begin{aligned} I_0 u &= \sum_{i=1}^{N_c} \sum_{\ell=1}^{L_i} \left(\int_{\omega_i} \kappa u_N \varphi_\ell^{\omega_i} \right) \chi_i \varphi_\ell^{\omega_i} + \sum_{K \in \mathcal{T}^H} \sum_{\ell=1}^{L_K} \left(\int_K \kappa u_D \psi_\ell \right) \psi_\ell \\ &= \sum_{i=1}^{N_c} (I_{L_i}^{\omega_i} u_N) \chi_i + \sum_{K \in \mathcal{T}^H} (I^K u_D) = I_N u_N + I_D u_D. \end{aligned} \tag{3-76}$$

Recall that L_i is the number of Newmann eigenfunctions considered in the neighborhood ω_i and L_K is the number of Dirichlet eigenfunctions considered on the element K. We finally present our main approximation result.

Assumption 20. *Let u be the exact solution, that is $-\operatorname{div}(\kappa \nabla u) = f$. We assume that there exist u_D, u_N and $\epsilon \preceq H$ such that*

1. *We have*

$$\int_\Omega \kappa |\nabla(u - u_D - u_N)|^2 \preceq \epsilon^2 \int_\Omega \kappa^{-1} f^2 \preceq H^2 \int_\Omega \kappa^{-1} f^2. \tag{3-77}$$

2. We have the boundary data given by

$$\partial_\eta u_N = 0 \text{ on } \partial K \text{ for all } K.$$

and

$$u_D = 0 \text{ on } \partial K \text{ for all } K.$$

3. We have the bounds,

$$\int_\Omega \kappa^{-1}|\mathrm{div}(\kappa\nabla u_N)|^2 \preceq \int_\Omega \kappa^{-1}|\mathrm{div}(\kappa\nabla u)|^2 \text{ and } \int_\Omega \kappa^{-1}|\mathrm{div}(\kappa\nabla u_D)|^2 \preceq \int_\Omega \kappa^{-1}|\mathrm{div}(\kappa\nabla u)|^2. \tag{3-78}$$

Theorem 21. *Assume that $|||u|||_{2,\Omega} < \infty$, where u is the solution of* (3-43)*. Then the following approximation for the energy interpolation error holds,*

$$\int_\Omega \kappa|\nabla u - \nabla I_0 u|^2 \leq \left(\max\left\{\frac{1}{H^2(\sigma_{L+1})^2}, \frac{1}{\sigma_{L+1}}\right\} + \frac{1}{\mu_{L+1}} + \epsilon^2\right) |||\kappa^{-1/2} f|||_0^2, \tag{3-79}$$

where $\sigma_{L+1} = \min_K \sigma^{\omega_i}_{K,L+1}$ and $\mu_{L+1} = \min_K \mu^K_{L+1}$.

Proof. By definition (3-76), our technical assumption, and using the triangular inequality, we have

$$\begin{aligned}\int_\Omega \kappa|\nabla u - \nabla I_0 u|^2 &= \int_\Omega \kappa|\nabla(u + u_N + u_D - (u_N + u_D)) - \nabla(I_N u_N + I_D u_D)|^2 \\ &\leq \int_\Omega \kappa|\nabla(u_N - I_N u_N)|^2 + \int_\Omega \kappa|\nabla(u_D - I_D u_D)|^2 + \int_\Omega \kappa|\nabla(u - u_N - u_D)|^2.\end{aligned} \tag{3-80}$$

Using Lemma 9 and noting that $\bigcup K \subset \bigcup \omega_K$ we have for the first term in (3-80),

$$\begin{aligned}\int_\Omega \kappa|\nabla(u_N - I_N u_N)|^2 &= \sum_{K\in\mathcal{T}^H} \int_K \kappa|\nabla(u_N - I_N u_N)|^2 \\ &\leq \sum_{\omega_K} \max\left\{\frac{1}{H^2\sigma^2_{K,L+1}}, \frac{1}{\sigma_{K,L+1}}\right\} ||\kappa^{-1}\mathrm{div}(\kappa\nabla u_N)||^2_{L^2(\omega_K)} \\ &\leq \max_{K\in\mathcal{T}^H}\left\{\max\left\{\frac{1}{H^2\sigma^2_{K,L+1}}, \frac{1}{\sigma_{K,L+1}}\right\}\right\} \sum_{\omega_K} ||\kappa^{-1}\mathrm{div}(\kappa\nabla u_N)||^2_{L^2(\omega_K)} \\ &\leq \max\left\{\frac{1}{H^2\sigma^2_{L+1}}, \frac{1}{\sigma_{L+1}}\right\} ||\kappa^{-1}\mathrm{div}(\kappa\nabla u_N)||^2_{L^2(\Omega)}.\end{aligned} \tag{3-81}$$

Now, using Lemma 10 to bound the second term in equation (3-80), we obtain the following

$$\begin{aligned}\int_{\Omega}\kappa|\nabla(u_D - I_D u_D)|^2 &= \sum_{K\in\mathcal{T}^H}\int_K \kappa|\nabla(u_D - I_D u_D)|^2 \leq \sum_{K\in\mathcal{T}^H}\frac{1}{\mu_{L+1}^K}||\kappa^{-1}\mathrm{div}(\kappa\nabla u_D)||^2_{L^2(K)}\\ &\leq \max_{K\in\mathcal{T}^H}\left\{\frac{1}{\mu_{K,L+1}}\right\}\sum_{K\in\mathcal{T}^H}||\kappa^{-1}\mathrm{div}(\kappa\nabla u_D)||^2_{L^2(K)}\\ &\leq \frac{1}{\mu_{L+1}}||\kappa^{-1}\mathrm{div}(\kappa\nabla u_D)||^2_{L^2(\Omega)}. \end{aligned} \tag{3-82}$$

Given that u is the exact solution, which satisfies $-\mathrm{div}(\kappa\nabla u) = f$, and with Assumption 20, we can derive bounds for the third term in equation (3-80).

$$\int_{\Omega}\kappa|\nabla(u - u_N - u_D)|^2 \leq \epsilon^2\int_{\Omega}\kappa^{-1}f^2, \tag{3-83}$$

with (3-83) and (3-82), we obtain (3-79) from (3-80). □

4 A duality GMsFEM method applied to high-contrast dam problem

This chapter is based on the manuscript [13]. We address the numerical homogenization approximation of a free-boundary dam problem in a heterogeneous media. More precisely, we propose a GMsFEM for the heterogeneous dam problem. The motivation for using the GMsFEM approach comes from the multiscale nature of the porous media due to its high-contrast permeability. Thus, although we can classically formulate the free-boundary dam problem as in the homogeneous case, a high resolution will be needed by a standard finite element approximation to obtain realistic results that recover the multiscale variations of the solution. First, we introduce a fictitious time variable that motivates a suitable time discretization that can be understood as a fixed point iteration to the steady-state solution. We use a duality method to deal with the involved multivalued nonlinear terms. Next, we compute efficient approximations of the pressure and the saturation using the GMsFEM method, and we can identify the free boundary. More precisely, the GMsFEM method provides numerical results that capture the behavior of the solution due to the variations of the coefficient at the fine resolution by just solving linear systems with the size proportional to the number of coarse blocks of a coarse-grid (that does not need to be adapted to the variations of the coefficient). Finally, we present illustrative numerical results to validate the proposed methodology.

4.1 Introduction

We consider a bounded two-dimensional rectangular domain D and let $\partial D = \Gamma \cup \Gamma_0 \cup \Gamma_a$ denote its boundary, where Γ is an impervious part of the boundary, Γ_0 is the part of the boundary in contact with open air, and Γ_a is the part of the boundary in contact with water; see Figure **4-1** for an illustration. In the proposed dam problem, we aim to compute the pressure p and the saturation θ of water, both defined on D, as well as to identify the free boundary separating the saturated and non-saturated regions of the dam. Moreover, we denote by κ the functional-coefficient that represents the permeability of the porous media and let $\boldsymbol{g} := -g\boldsymbol{e}_2$ denote the gravity. By using Darcy's law for porous media and the relation between pressure and water saturation, we obtain

$$-g\frac{\partial(\theta\kappa)}{\partial x_2} - \operatorname{div}(\kappa\nabla p) = 0, \quad p \geq 0, \quad \theta \in H(p), \tag{4-1}$$

where $H(\cdot)$ denotes the multivalued Heaviside operator, so that for positive pressure ($p > 0$) the porous media is fully saturated ($\theta = 1$) and $\theta \in [0, 1)$ when $p = 0$ in the non saturated region. In order to pose the strong formulation of the free-boundary dam problem, the set of equations (4-1) is completed with the following boundary conditions:

- $p = h_a - x_2$ on Γ_a, with h_a height of the water level in contact with Γ_a,
- $p = 0$ in Γ_0,
- $\big(\theta\kappa \boldsymbol{g} - \kappa\nabla p\big) \cdot \boldsymbol{n} \geq 0$ in Γ_0 where we recall that $\boldsymbol{g} = -g\boldsymbol{e}_2$,
- $\Big(\theta\kappa \boldsymbol{g} - \kappa\nabla p\Big) \cdot \boldsymbol{n} = 0$ in Γ.

In the previous equations, $\boldsymbol{n}$ represents the unitary outwards normal vector to the boundary ∂D. See Figure **4-1** for an illustration of the domain and different boundaries.

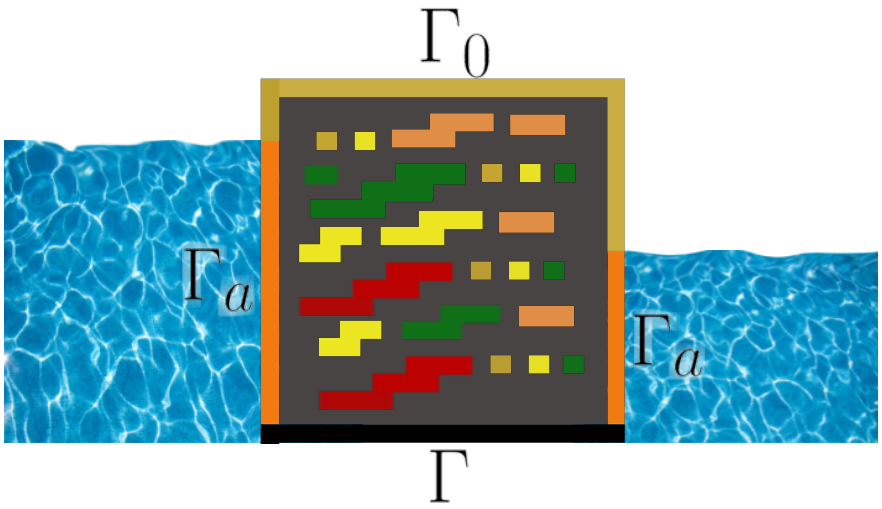

Figure 4-1: Illustration of a free boundary dam problem in multiscale high-contrast porous media.

We remark that the methodology proposed in this paper can be applied to general coefficients (see [30, 28] and related works).
Now, we focus on the case of high-contrast multiscale coefficients. More precisely, we consider piece-wise smooth permeabilities. We assume that the domain D is the union of finitely many sub-domains, that is,

$$D = \bigcup_{i=1}^{N_S} D_i,$$

where $\{D_i\}$ is a non-overlapping decomposition of D. The permeability coefficient can be written as,

$$\kappa(x) = k_i(x) \text{ for } x \in D_i,$$

$i = 1, 2, \ldots, N_S$, where κ_i is a bounded smooth function (that may have oscillations) in D_i. We say that the coefficient $\kappa(x)$ is a multiscale coefficient if it has oscillations at different scales in each subdomain D_i. We also say that κ is a high-contrast coefficient if the ratio $\eta_\Omega = \max_{x,y\in\Omega} \kappa(x)/\kappa(y) >> 1$ for some subsets $\Omega \subset D$ (in this case we say that Ω is a high-contrast sub-region).
In this chapter, we adopt the methodology proposed in [9] and further extended in [49]. To facilitate the analysis, we introduce an artificial time dependence in all the unknowns, denoting them as $p = p(t, x)$ and $\theta = \theta(t, x)$, similar to the notation used in the steady-state case. With this approach, we can express equation (4-1) equivalently as a system of evolution equations, given by the following

$$\frac{\partial}{\partial t}(\theta\kappa) - g\partial_2(\theta\kappa) - \operatorname{div}(\kappa\nabla p) = 0, \quad p \geq 0, \quad \theta \in H(p). \tag{4-2}$$

Note that the first equation in (4-2) is a nonlinear advection-diffusion equation. In order to discretize in time this equation, we are the characteristics method, we first introduce the material or total derivative associated to the vector field driving the convection term which is given by $\boldsymbol{g} = (0, -g)$. Therefore, we can write

$$\frac{Dz}{Dt} = \frac{\partial z}{\partial t} + \mathbf{g} \cdot \nabla z = \frac{\partial z}{\partial t} - g\,\partial_2 z.$$

Thus, in terms of the material derivative we can write (4-2) in the form:

$$\frac{D}{Dt}(\theta\kappa) - \operatorname{div}(\kappa\nabla p) = 0, \quad p \geq 0, \quad \theta \in H(p).$$

For the purpose of the time discretization, we introduce an uniform finite differences time mesh with points $t_0, t_1, \ldots t_M$, and a constant time step δt. Next, we introduce a forward in time approximation of the total derivative by the method of characteristics. More precisely, if we use the notation $f^n(x) = f(t^n, x)$ for $n = 0,\ 1,\ 2, \ldots$, at each time t^{n+1}, we obtain the strong formulation of the problem discretized in time,

$$\frac{\theta^{(n+1)}\kappa - (\theta^{(n)}\kappa) \circ \Phi^n}{\delta t} - \operatorname{div}(\kappa\nabla p^{(n+1)}) = 0, \quad p^{(n+1)} \geq 0, \theta^{(n+1)} \in H(p^{(n+1)}), \tag{4-3}$$

where the index $n+1$ denotes the approximation at the artificial time t^{n+1} of the introduced time dependent functions. Moreover, the function Φ^n is defined at each spatial point of the domain by $\Phi^n(x) = \Phi(t^{n+1}, x; t^n)$, that denotes the position at time t^n of the point placed in x at time t^{n+1} and moved along the integral path (characteristic curve) defined by the velocity field $\mathbf{g}$, so that $\Phi^n(x)$ can be obtained from the solution of the final value ODE problem:

$$\frac{d\Phi}{dt}(t^{n+1}, x; t) = \mathbf{g}\left(t, \Phi(t^{n+1}, x; t)\right), \quad \Phi(t^{n+1}, x; t^{n+1}) = x. \tag{4-4}$$

Therefore, in terms of the previous solution we define $\Phi^n(x) = \Phi(t^{n+1}, x; t^n)$.

Taking into account the particular expression of the velocity field $\mathbf{g}$, for $x = (x_1, x_2)$ we can easily compute

$$\Phi^n(x_1, x_2) = (x_1, x_2 + g\,\Delta t).$$

Remark 22. *As stressed in Chapter 5, the presence of high-contrast and multiscale variation negatively affects the accuracy and stability of time discretizations. As proposed in Chapter 5, one way to work around this problem is to use more sophisticated fine discretization schemes. In this Chapter, we use uniform finite differences in time and assume we can reduce the fine step size to obtain stability. As in the example of Chapter 5, the MsFEM is easily adapted to these procedures.*

Note that as $\mathbf{g}$ does not depend on t and then Φ^n does not depend on n, we will drop the superindex n in Φ^n.
To write a weak form of the problem, we introduce the following functional spaces:

- $V_- = H^1(D) \cap [v|_{\Gamma_0} \leq 0] = \{\psi \in H^1(D); \psi|_{\Gamma_a} = 0; \psi|_{\Gamma_0} \leq 0\}$,
- $V_0 = H^1_0(D, \Gamma_0) = \{v \in H^1(D) : v|_{\Gamma_0} = 0\}$ and note that $V_0 \subset V_-$,
- $V_+ = H^1(D) \cap [v|_{\Gamma_0} \geq 0]$,
- $W_\alpha = H^1(D) \cap [v|_{\Gamma_a} = \alpha]$,
- $W_0 = H^1(D) \cap [v|_{\Gamma_a} = 0]$.

Next, for $\phi \in V_-$ such that $\phi = 0$ on Γ_a, by multiplying the first equation in (4-3) by $\phi - p^{(n+1)}$ and integrating by parts we get the weak formulation of the problem:
Find $p^{(n+1)} \in V_0 \cap W_\alpha \cap V_+ = V_0 \cap W_\alpha$ and $\theta^{n+1} \in L^\infty(D)$ such that

$$\begin{aligned}&\int_D \kappa\theta^{(n+1)}(\phi - p^{n+1}) - \int_D ((\theta^{(n)}\kappa) \circ \Phi)(\phi - p^{n+1}) + \\ &\delta t \int_D \kappa \nabla p^{(n+1)} \nabla(\phi - p^{n+1}) + \delta t \int_{\Gamma_0 \cup \Gamma} \theta^{(n+1)} \kappa \boldsymbol{gn}(\phi - p^{n+1}) \geq 0,\end{aligned} \tag{4-5}$$

jointly with

$$\theta^{(n+1)} \in H(p^{(n+1)}). \tag{4-6}$$

Consider now the indicatrix function of the convex set V_- defined over $H^1(D)$ by

$$I_{V_-}(v) = \begin{cases} 1 & v \in V_-, \\ +\infty & v \notin V_-. \end{cases}$$

(see definition (2-70)). Note that I_{V_-} is a convex semicontinuous function which implies that the subdifferential operator ∂I_{V_-} is a well defined maximal monotone multivalued operator (see Section 2.3.2). Moreover, the multivalued operator ∂I_{V_-} can be characterized as follows:

$$\alpha \in \partial I_{V_-}(u) \Longleftrightarrow I_{V_-}(v) - I_{V_-}(u) \geq \langle \alpha, v - u \rangle, \tag{4-7}$$

for all $v \in H^1(D)$. Here $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^1(D)$ and its dual space. Therefore, if we define

$$\langle L(p^{(n+1)}), \phi - p^{(n+1)} \rangle = \int_D \kappa \theta^{(n+1)}(\phi - p^{n+1}) - \int_D ((\theta^{(n)}\kappa) \circ \Phi)(\phi - p^{n+1}) + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla(\phi - p^{n+1}) + \delta t \int_{\Gamma_0 \cup \Gamma} \theta^{(n+1)} \kappa \boldsymbol{gn}(\phi - p^{n+1}),$$

then from inequality (4-5) and the definition of the indicatrix function, we get

$$\begin{aligned} &\int_D \kappa \theta^{(n+1)}(\phi - p^{n+1}) - \int_D ((\theta^{(n)}\kappa) \circ \Phi)(\phi - p^{n+1}) \\ &+\delta t \int_D \kappa \nabla p^{(n+1)} \nabla(\phi - p^{n+1}) + \delta t \int_{\Gamma_0 \cup \Gamma} \theta^{(n+1)} \kappa \boldsymbol{gn}(\phi - p^{n+1}) \\ &+\frac{1}{\delta t}\left(I_{V-}(\phi) - I_{V-}(p^{(n+1)})\right) \geq 0, \end{aligned}$$

or, in compact notation, (see Section 2.3.2)

$$\langle L(p^{(n+1)}), \phi - p^{(n+1)} \rangle + \frac{1}{\delta t}\left(I_{V-}(\phi) - I_{V-}(p^{(n+1)})\right) \geq 0,$$

for all $\phi \in W_0 = H^1(D) \cap [v|_{\Gamma_a} = 0]$. Moreover, the previous inequality implies that $L(p^{(n+1)}) \in \partial I_{V-}(p^{(n+1)})$. Therefore, we can introduce the new variable

$$q^{(n+1)} = L(p^{(n+1)}) \quad \text{with} \quad -(\delta t) q^{(n+1)} \in \partial I_{V-}(p^{(n+1)}).$$

Note that we can identify $H^1(D)$ with the sum $H^{1/2}(\partial D) \oplus H_0^1(D)$ and therefore we can identify the dual space of $H^1(D)$ with $H^{-1/2}(\partial D) \oplus H^{-1}(D)$. Using this representation, equation (4-7) and the fact that $p^{n+1} \in V_0 \cap W_\alpha$ we can write the time discretized problem at step n as:

Find $p^{(n+1)} \in V_0 \cap W_\alpha$ and $\theta^{n+1} \in L^\infty(D)$, such that

$$\begin{aligned} &\int_D \kappa \theta^{(n+1)} \phi + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla \phi \\ &+\delta t \int_{\Gamma_0 \cup \Gamma} \theta^{(n+1)} \kappa \boldsymbol{gn} \phi + \delta t \int_{\Gamma_0} q^{(n+1)} \phi \;=\; \int_D ((\theta^{(n)}\kappa) \circ \Phi)\phi, \end{aligned} \tag{4-8}$$

for all $\phi \in W_0$, jointly with the following multivalued nonlinear equations

$$\begin{cases} q^{(n+1)} \in \partial I_{V-}(p^{(n+1)}), \\ \theta^{(n+1)} \in H(p^{(n+1)}). \end{cases} \tag{4-9}$$

4.2 A duality method for nonlinear terms

In order to solve (4-8)-(4-9), we follow the methodology used in [9, 49, 55] and Section 4.2 to deal with nonlinear terms associated with multivalued operators in (4-9). These techniques

are based on duality methods for nonlinear maximal monotone operators and are here applied to the multivalued Heaviside and subdifferential operators, (see subsection 2.3.1 and 2.3.2, respectively). In the seminal article [10], these duality methods have been introduced for solving variational inequalities.

For this purpose, we first recall the concept of Yosida approximation (see section 2.3 and [55]). Let G be a maximal monotone operator and let ω and λ be non-negative real numbers such that $\omega\lambda < 1$. The resolvent of G is defined by,

$$J_\lambda^\omega = ((1 - \omega\lambda I) + \lambda G)^{-1}.$$

Next, we introduce the Yosida approximation of $G - \omega I$ of the parameter λ, which is defined by

$$G_\lambda^\omega := \frac{I - J_\lambda^\omega}{\lambda}.$$

As recalled in [49], it can be proved that $u \in G(y) - \omega y$ is equivalent to $u = G_\lambda^\omega(y + \lambda u)$, for further details see also the seminal article [10]. Note that the first expression is written in terms of the multivalued operator while the second one is a nonlinear equation for u in terms of an univalued Yosida operator.

Next, in terms of the non-negative parameters ω_1 and ω_2, we introduce the new variables

$$\alpha^{(n+1)} = q^{(n+1)} - \omega_1 p^{(n+1)} \quad \text{and} \quad \beta^{(n+1)} = \theta^{(n+1)} - \omega_2 p^{(n+1)}. \tag{4-10}$$

Therefore, from (4-9) we have

$$\alpha^{(n+1)} \in \partial I_{V_-}(p^{(n+1)}) - \omega_1 p^{(n+1)}$$

and

$$\beta^{(n+1)} \in H(p^{(n+1)}) - \omega_2 p^{(n+1)}.$$

We can then write the variational formulation in terms of the new variables in the form

$$\begin{aligned}\int_D \kappa(\beta^{(n+1)} + \omega_2 p^{(n+1)})\phi + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla \phi \\ + \delta t \int_{\Gamma_0 \cup \Gamma} ((\beta^{(n+1)} + \omega_2 p^{(n+1)}))\kappa \boldsymbol{gn}\phi \\ + \delta t \int_{\Gamma_0} (\alpha^{(n+1)} + \omega_1 p^{(n+1)})\phi \;\; = \;\; \int_D ((\theta^{(n)}\kappa) \circ \Phi)\phi.\end{aligned}$$

Next, using the previous characterization of the elements of the multivalued operator $G - \omega I$ in terms of the its Yosida approximation for the particular cases $G = \partial I_{V_-}$ and $G = H$, the variational formulation can be equivalently written in the form

$$\begin{aligned}&\int_D \kappa \nabla p^{(n+1)} \nabla \phi + \frac{\omega_2}{\delta t} \int_D \kappa p^{(n+1)} \phi + \omega_1 \int_{\Gamma_0} p^{(n+1)} \phi + \omega_2 \int_{\Gamma_0 \cup \Gamma} p^{(n+1)} \kappa \boldsymbol{gn} \phi, \\ &= \frac{1}{\delta t} \int_D ((\theta^{(n)}\kappa) \circ \Phi)\phi - \frac{1}{\delta t} \int_D \kappa \beta^{(n+1)} \phi - \int_{\Gamma_0 \cup \Gamma} \beta^{(n+1)} \kappa \boldsymbol{gn} \phi - \int_{\Gamma_0} \alpha^{(n+1)} \phi\end{aligned} \tag{4-11}$$

with

$$\alpha^{(n+1)} = (\partial I_{V_-})_{\lambda_1}^{\omega_1}(p^{(n+1)} + \lambda_1\alpha^{n+1}), \tag{4-12}$$

and

$$\beta^{(n+1)} = H_{\lambda_2}^{\omega_2}(p^{(n+1)} + \lambda_2\beta^{n+1}). \tag{4-13}$$

(see Section 2.3.2). Following [49], we propose to solve (4-11), (4-12) and (4-13) numerically using a fixed point iteration as described in the next paragraphs.
Given $\alpha^{(n+1)}$, $\beta^{(n+1)}$ and $\theta^{(n)}$, we solve equation (4-11) for the pressure and denote the solution by $p^{(n+1)} = \mathcal{L}(\alpha^{(n+1)}, \beta^{(n+1)}, \theta^{(n)})$. Schematically, we have the following system of coupled equations

$$\alpha^{(n+1)} = (\partial I_{V_-})_{\lambda_1}^{\omega_1}\left(\mathcal{L}(\alpha^{(n+1)}, \beta^{(n+1)}, \theta^{(n)}) + \lambda_1\alpha^{n+1}\right) \tag{4-14}$$

$$\beta^{(n+1)} = H_{\lambda_2}^{\omega_2}\left(\mathcal{L}(\alpha^{(n+1)}, \beta^{(n+1)}, \theta^{(n)}) + \lambda_2\beta^{n+1}\right). \tag{4-15}$$

Using the results in [49] it can be seen that, given $\theta^{(n)}$, this system can be solved by a fixed point iteration. To start the fixed point iteration, we use previous values of $\alpha^{(n)}$ and $\beta^{(n)}$. The value of θ^n can be updated using (4-10).

For the spatial discretization of the linear problems arising at each step of the fixed point iteration, we consider Finite Elements Methods. For this purpose, let $\mathcal{T}^h$ be a triangular partition of the domain D such that it resolves the variation of the permeability coefficient κ. Consider V the finite element space of piece-wise linear (or bi-linear) finite elements defined on the mesh $\mathcal{T}^h$. At each step of previous iteration, the fully discretized problem can be written in terms of the solution of the following linear system:

$$\begin{aligned}\left(A + \frac{\omega_2}{\delta t}M + \omega_1 M_{\Gamma_0} + \omega_2 M_{\Gamma_0\cup\Gamma}\right)p^{(n+1)} = \\ b^{(n)} - \left(\frac{1}{\delta t}M + M_{\Gamma_0\cup\Gamma}\right)\beta^{(n+1)} - M_{\Gamma_0}\alpha^{(n+1)}.\end{aligned} \tag{4-16}$$

In the linear system (4-16), we introduced the following matrices,

$$A = [a_{ij}] \text{ with } a_{ij} = \int_D \kappa\nabla\phi_i\nabla\phi_j,$$

$$M = [m_{ij}] \text{ with } m_{ij} = \int_D \kappa\phi_i\phi_j,$$

$$M_{\Gamma_0} = [m_{ij;\Gamma_0}] \text{ with } m_{ij;\Gamma_0} = \int_{\Gamma_0} \phi_i\phi_j,$$

and

$$M_{\Gamma\cup\Gamma_0} = [m_{ij;\Gamma\cup\Gamma_0}] \text{ with } m_{ij;\Gamma\cup\Gamma_0} = \int_{\Gamma\cup\Gamma_0} \phi_i\kappa\boldsymbol{gn}\phi_j.$$

Moreover, we have also considered the vector $b^{(n)} = [b_i^{(n)}]$ associated to the method of characteristics, the components of which are given by

$$b_i^{(n)} = \frac{1}{\delta t}\int_D ((\theta^{(n)}\kappa)\circ\Phi)\phi_i.$$

The computation of the integral in $b_i^{(n)}$ requires some interpolation techniques, as the function $\theta^{(n)}\kappa$ needs to be evaluated at points that may not belong to the mesh.
Note that at each step n of the algorithm we recursively solve the linear system (4-16) and update the terms α^{n+1} and β^{n+1} in the second member by using (4-14) and (4-15), respectively. In practice, in the numerical examples in a forthcoming section, we consider $\lambda_1 = \lambda_2 = 1$ and $\omega_1 = \omega_2 = 0.5$, so that we fulfill the condition $\lambda_i\omega_i = 0.5$, as in [49]. Note that this condition proves the convergence of the fixed point iteration in [10] for a variational inequality problem. Also, for an elastohydrodynamic problem in magnetic storage devices, the convergence is theoretically proved under the same condition in [5]. We also mention that the number of fixed point iterations is chosen to be constant independently of the time step iteration. This number of iterations used in the fixed point step was chosen to replace the convergence with the stationary solution (the main target of our computation).
In the next section, we propose using the GMsFEM to solve the fully discretized problem (4-16) so that we replace the fine-scale system with a coarse linear system associated with an appropriate coarse space V_0.

4.3 Generalized multiscale finite element method

This section focuses on high-contrast multiscale problems and summarizes a GMsFEM construction of a coarse space V_0. For a more detailed description of the development of the GMsFEM methodology, see [28, 31], and references therein.
We start by choosing an initial set of basis functions that form a partition of unity. The space generated by this basis function is enriched using a local spectral problem. We use the multiscale basis functions partition of unity with linear boundary conditions (see [33], for example). We have one function per coarse node, and it is defined by

$$\begin{aligned} -\mathrm{div}(\kappa\nabla\chi_i) &= 0 \quad \text{for } K\in\omega_i, \\ \chi_i &= \chi_i^0 \quad \text{on } \partial K, \end{aligned} \tag{4-17}$$

where χ_i^0 is a standard linear partition of unity function.
For each coarse node neighborhood ω_i, consider the eigenvalue problem

$$-\mathrm{div}(\kappa\nabla\psi_\ell^{\omega_i}) = \sigma_\ell^{\omega_i}\widetilde{\kappa}\psi_\ell^{\omega_i}, \tag{4-18}$$

with homogeneous Neumann boundary condition on $\partial\omega_i$. Here $\sigma_\ell^{\omega_i}$ and $\psi_\ell^{\omega_i}$ are eigenvalues and eigenvectors in ω_i and $\widetilde{\kappa}$ defined by

$$\widetilde{\kappa} = \kappa\sum_{j=1}^{N_v} H^2|\nabla\chi_j|^2.$$

We use an ascending ordering on the eigenvectors, $\sigma_1^{\omega_i} \leq \sigma_2^{\omega_i} \leq ....$

Using the partition of unity functions from Eq. (4-17) and eigenfunctions from Eq. (5-11), we then construct a set of enriched multiscale basis functions given by $\chi_i\psi_\ell^{\omega_i}$ for the selected eigenvectors $\psi_\ell^{\omega_i}$. Using L_i to denote the number of basis functions from the coarse region ω_i, we then define the coarse GMsFEM space by

$$V_0 = \text{span}\{\Phi_{i,\ell} = \chi_i\psi_\ell^{\omega_i}, \quad i = 1, \ldots, N_v, \quad \ell = 1, \ldots, L_i\}.$$

For more details, motivation of the construction, approximation properties of the space V_0, and the choice of the initial partition of unity basis functions, we refer the interested reader to [28] and Section 3.3.

In summarizing, in order to solve the problem (4-16), for the pressure, we use the GMsFEM coarse space V_0 constructed in this section. More precisely, let R_0 the matrix whose columns correspond to the coarse basis functions, that is, the column space of R_0 is V_0. Instead of solving the fine-scale linear system (4-16), we solve the coarse-scale linear system

$$S_0 p_0^{(n+1)} = c_0^{(n)}, \tag{4-19}$$

where the matrix S_0 and the second member $c_0^{(n)}$ are given by

$$\begin{aligned} S_0 &= R_0^T \left(A + \frac{\omega_2}{\delta t} M + \omega_1 M_{\Gamma_0} + \omega_2 M_{\Gamma_0 \cup \Gamma} \right) R_0, \\ c_0^{(n)} &= R_0^T \left(b^{(n)} - \left(\frac{1}{\delta t} M + M_{\Gamma_0 \cup \Gamma} \right) \beta^{(n+1)} - M_{\Gamma_0} \alpha^{(n+1)} \right). \end{aligned}$$

We then maintain the duality method explained before but using the approximation $R_0 p_0^{(n+1)} \approx p^{(n+1)}$ that makes the computation more efficient since, instead of solving the full resolution linear system (4-16), we solve the small coarse problem (4-19). We mention that due to the high-contrast multiscale coefficient structure, we need to solve the coarse problem at the right resolution to obtain a good approximation; see [28]. With the GMsFEM methodology, we can adapt the coarse solver's resolution to obtain good results with the duality method. We recall that a linear system has to be solved for each time iteration and each fixed point iteration to compute the current pressure. In the proposed methodology, instead of solving the fine-grid linear system (4-16), we solve the coarse scale linear system (4-19). We stress that the linear system (4-16) is very large and ill-conditioned (with condition number increasing with the contrast in the coefficient). The size of the system (4-19) is of the order of the number of coarse-scale nodes, so it is suitable for factorization methods. This allows us to save computational time. Moreover, we mention that not the basis functions nor the coarse scale operators and matrices change throughout the time and non-linear iteration. Therefore, the set-up cost (constructing a coarse grid, computing local eigenvectors and coarse basis functions, and assembling coarse scale operators) can be considered as a pre-processing

cost; see [28] for more details on the computational implementation of GMsFEM.

A vital aspect of the GMsFEM is that the needed resolution can be considered a priori or a posteriori, depending on the application. In this paper, we show how the resolution of the method, that is, the parameters L_i, the number of eigenvectors in ω_i used in the construction of the coarse space, affects the solution of the free boundary heterogeneous multiscale dam problem. As in other applications leading to iterative corrections using solutions of the diffusion equations, if the approximation of this step is poor, overall poor results are obtained in resolving the problems (see [18]).

In the numerical results presented in the following section, it is evident that the GMsFEM approximation requires a greater number of time steps to achieve convergence to the stationary solution, compared to the fine-scale solution (4-16), while employing the same tolerance for the increment. Specifically, we observe that approximately 10% more time steps are necessary when computing the fine-scale solution. It is important to note that the GMsFEM approximation involves solving the reduced system (4-19), which possesses a significantly smaller dimension compared to the fine-scale system. For a more comprehensive understanding of the results, please refer to Table **4-2**.

4.4 Numerical results

In this section, we provide numerical illustrations to demonstrate the performance of the GMsFEM method compared to the reference solution. The time evolution and duality method remain the same as before, using fine-grid vectors. However, instead of computing the fine-scale solution for the pressure equations, we utilize the GMsFEM solution.
For the boundary partition, we consider the configuration shown in Figure **4-1**. The Dirichlet data is specified as follows: $p = \frac{4}{5} - x_2$ on $\{0\} \times \left[\frac{3}{5}, 1\right]$, $p = \frac{1}{5} - x_2$ on $\{1\} \times \left[\frac{2}{5}, 1\right]$, and $p = 0$ on Γ_0.
In order to numerically study the performance of the GMsFEM method applied to the heterogeneous dam problems, we use the coefficients depicted in Figure **4-2**. We then compare the GMsFEM solution with the reference solutions, that is, we compute the error between the solution of the overall iteration with solutions of the diffusion equation on the fine-grid, with the multiscale solutions, that is, the solutions obtained by using the coarse-scale solution $p_0^{(n+1)}$ (downscaled to the fine-grid as $R_0 p_0^{(n+1)}$) for the approximation of the diffusion equation.

In particular, we use $L_i = 0, 1, \ldots, 10$ for all i. We run the time iteration and the fixed point iteration until the norm of the increment is less than a given tolerance (10^{-4} in our numerical test).

4.4.1 Example with high contrast medium

We consider a structured fine-grid with 100 elements in each direction (yielding a fine-scale linear system matrix of dimension 10000×10000). We also consider a coarse mesh (made of squares) with 10 elements in each direction. see Figure **4-2**.

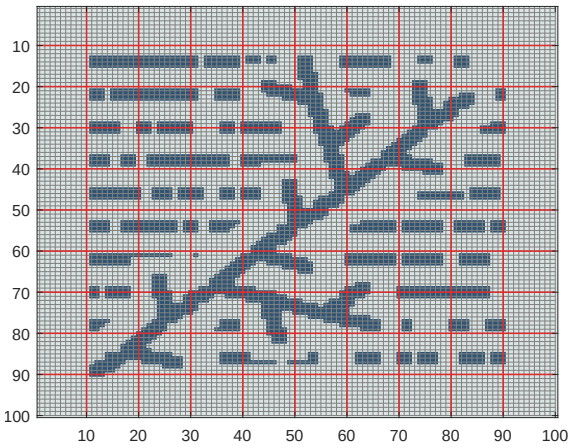

Figure 4-2: High-contrast coefficients used in the numerical experiment. High-conductivity channels in black color. In our numerical experiments we use coefficients of background 1 and high-contrast value 10^2.

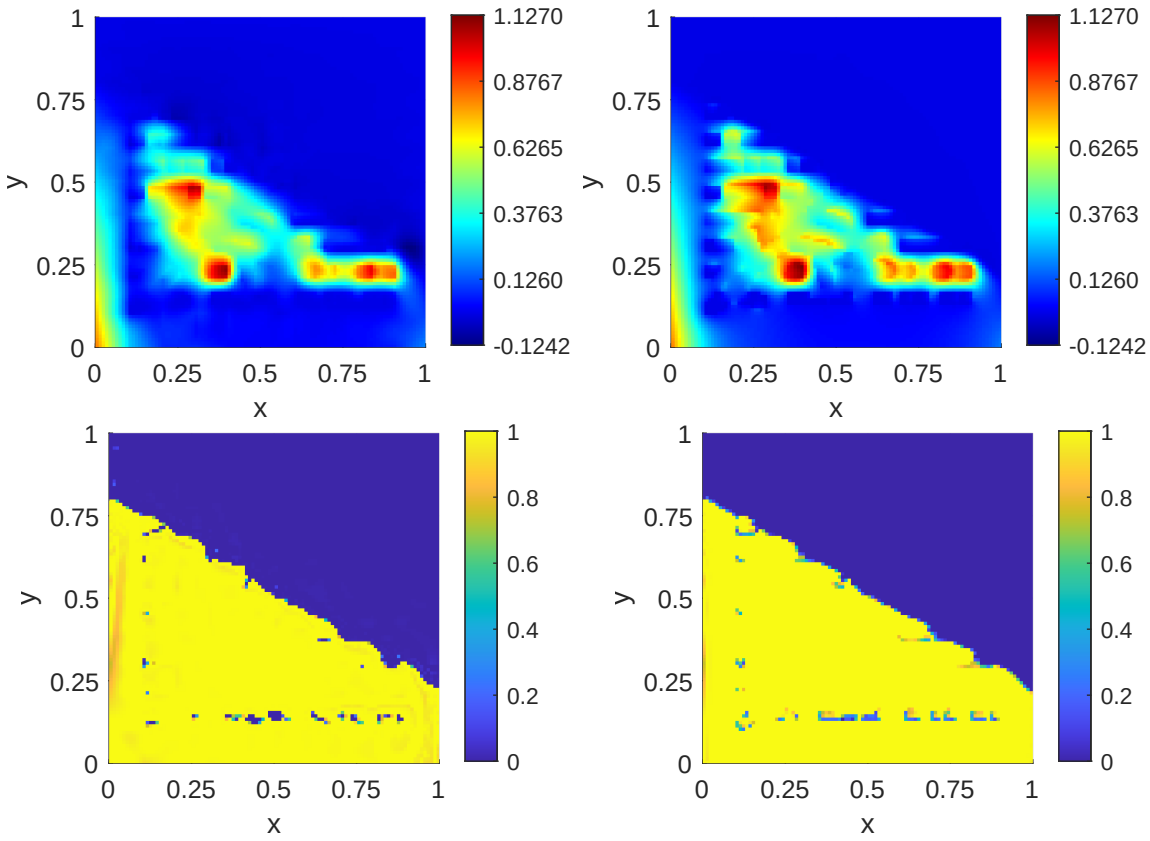

Figure 4-3: Computed pressure and saturation for medium **4-2**. From top to bottom and left to right: Using the fine-grid solution for pressure. Using coarse-grid solution with $L_i = 6$ for pressure. Using the fine-grid solution for saturation. Using coarse-grid solution with $L_i = 6$ for saturation .

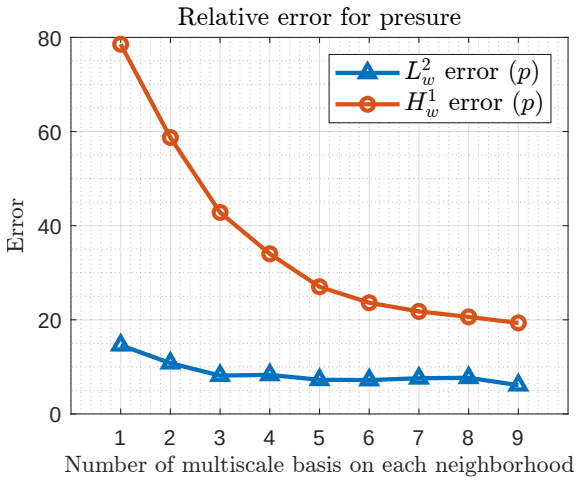

Figure 4-4: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution. The horizontal axis corresponds to the number of basis functions in each neighborhood used in the GMsFEM coarse spaces.

L_i	L_w^2	H_w^1
1	14.60	78.54
2	10.79	58.72
3	8.15	42.81
4	8.30	34.02
5	7.25	27.01
6	7.20	23.62
7	7.58	21.78
8	7.68	20.61
9	6.11	19.34

Table 4-1: The weighted L^2 and H^1 errors between the reference and the coarse-scale for problem (5-39) with respect to pressure.

4.4.2 Example with high contrast medium SPE10

We consider a part of 48^{th} layer of the geological SPE10 porous medium taken from [17, 2] a structured fine grid with 100 elements in each direction (yielding a fine-scale linear system matrix of dimension 10000×10000). We also consider a coarse mesh (made of squares) with ten elements in each direction. see Figure **4-5**.

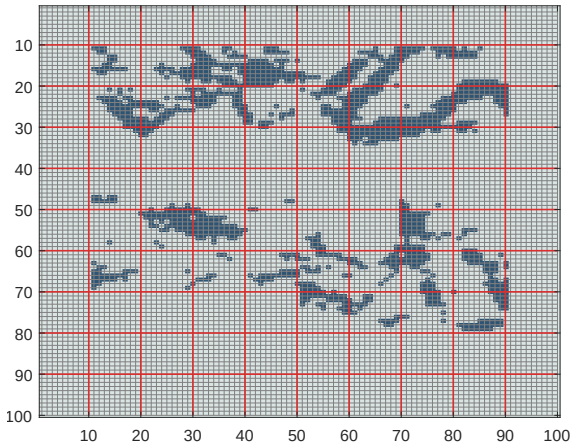

Figure 4-5: High-contrast coefficients used in the numerical experiment. High-conductivity channels in black color. In our numerical experiments we use coefficients of background 1 and high-contrast value 10^2.

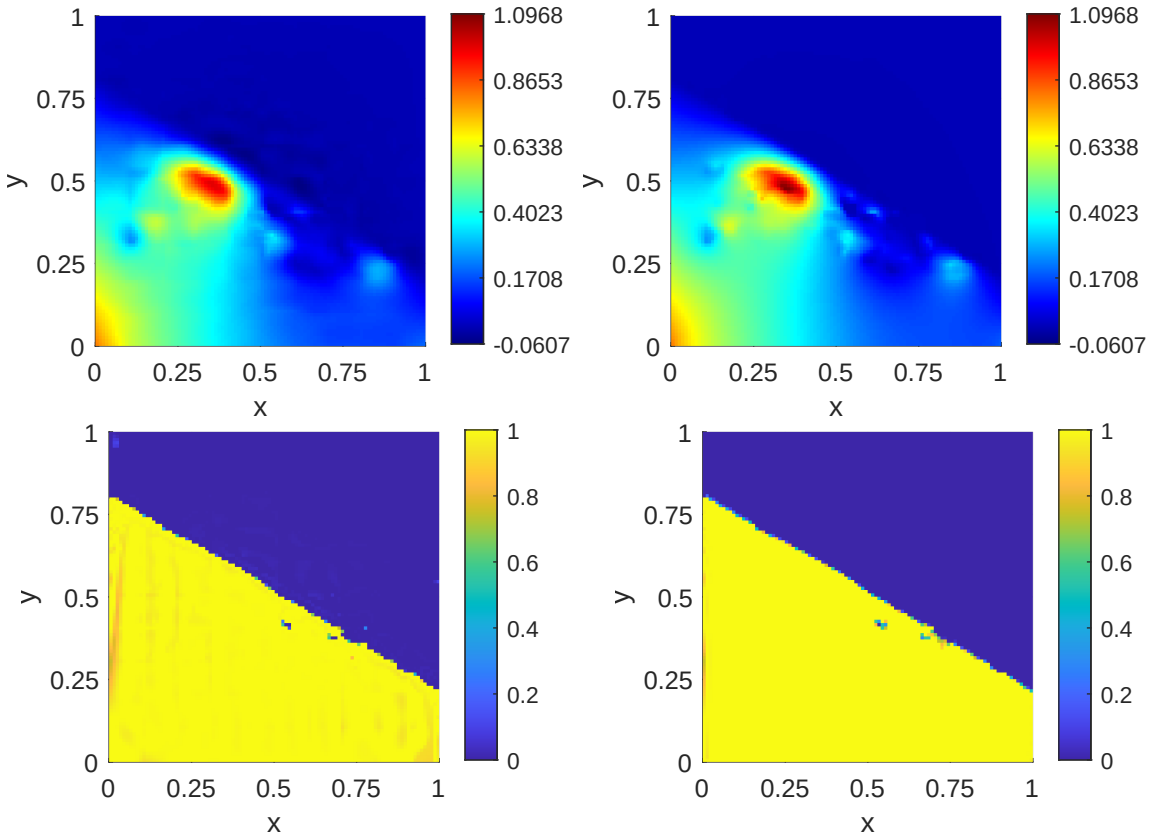

Figure 4-6: Computed pressure and saturation for medium **4-5**. From top to bottom and left to right: Using the fine-grid solution for pressure. Using coarse-grid solution with $L_i = 6$ for pressure. Using the fine-grid solution for saturation. Using coarse-grid solution with $L_i = 6$ for saturation .

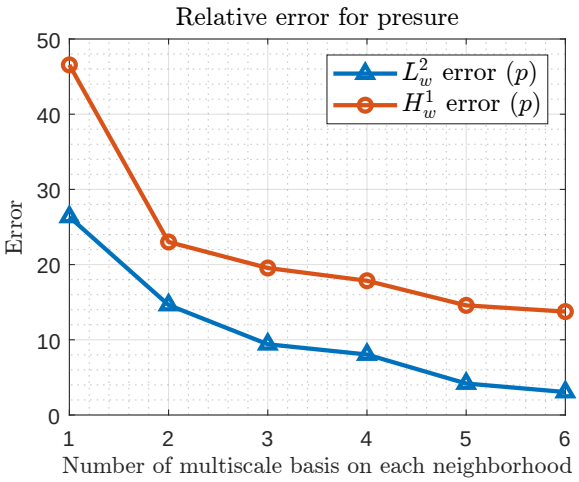

Figure 4-7: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution. The horizontal axis corresponds to the number of basis functions in each neighborhood used in the GMsFEM coarse spaces.

L_i	L_w^2	H_w^1
1	26.33	46.55
2	14.62	23.00
3	9.40	19.54
4	8.04	17.85
5	4.18	14.57
6	3.05	13.74

Table 4-2: The weighted L^2 and H^1 errors between the reference and the coarse-scale for problem (5-39) with respect to pressure.

5 A GMsFEM exponential integrator applied to a high-contrast multiscale parabolic problem

This chapter corresponds to the submitted manuscript [48]. We consider linear and semilinear parabolic problems posed in high-contrast multiscale media in two dimensions. High-contrast multiscale media adversely affects the accuracy, stability, and overall efficiency of numerical approximations such as finite elements in space combined with some time integrator. In many cases, implementing time discretizations such as finite differences or exponential integrators may be impractical because each time iteration needs the computation of matrix operators involving huge and ill-conditioned sparse matrices. Here, we propose an efficient Generalized Multiscale Finite Element Method (GMsFEM) that is robust against the high-contrast diffusion coefficient. We combine the GMsFEM with exponential integration in time to obtain a good approximation of the final time solution. Our approach is efficient and practical because it computes matrix functions of small matrices given by the GMsFEM method. We present representative numerical experiments that show the advantages of combining exponential integration and GMsFEM approximations. The constructions and methods developed here can be easily adapted to three-dimensional domains.

5.1 Introduction

We consider the following semilinear parabolic problem posed in a high-contrast multiscale media,

$$\begin{cases} \partial_t p - \operatorname{div}(\kappa(x)\nabla p) = f(p) & \text{in} \quad \Omega \times I, \\ p = p_D & \text{on} \quad \partial\Omega \times I, \\ p(0,x) = \hat{p}(x), & x \in \Omega. \end{cases} \tag{5-1}$$

Here Ω is a two dimensional convex domain with boundary $\partial\Omega$ and $I = [0,T]$ is the time domain. The field $\kappa(x)$ is a multiscale high-contrast heterogeneous field. Additionally, p is an unknown pressure field satisfying the Dirichlet condition given by p_D and the initial condition given by $\hat{p}$. The constructions and methods developed here can be easily adapted to the three-dimensional domains.

5.2 Variational formulation of the parabolic problem and GMsFEM

To get the variational formulation we proced an in Section 2.2.4. A variational formulation of problem (1-3) is: Find $p(t) \in H^1(\Omega)$ with $(p(t) - p_D) \in H_0^1 = \{w \in H^1(\Omega) : w|_{\partial\Omega} = 0\}$ such that

$$(\partial_t p, v) + a(p, v) = F(p; v) \quad \text{for all } v \in H_0^1(\Omega), \tag{5-2}$$

where $(\cdot, \cdot)$ denotes the usual inner product in $L^2(\Omega)$, the bilinear form a is defined by

$$a(p, v) = \int_\Omega \kappa(x) \nabla p(x) \nabla v(x) dx, \tag{5-3}$$

and the functional F is defined by

$$F(p; v) = \int_\Omega f(p(x)) v(x) dx, \tag{5-4}$$

for $p, v \in H_0^1(\Omega)$. Let $\mathcal{T}^h$ be a triangulation of the domain Ω. As it is usual in multiscale methods, we assume that h is fine enough to completely describe all the variations of the coefficient κ and therefore we refer to $\mathcal{T}^h$ as the fine mesh. We denote by $V^h(\Omega)$ the usual finite elements discretization of piecewise linear continuous functions with respect to $\mathcal{T}^h$. Denote by $V_0^h(\Omega)$ the subset of $V^h(\Omega)$ made of functions that vanish on $\partial\Omega$. The Galerkin formulation of (5-2) is to find $(p(t) - p_D) \in V_0^h(\Omega)$ such that

$$\begin{cases} (\partial_t p, v) + a(p, v) = F(p; v) & \text{for all } v \in V_0^h(\Omega),\ t \in I, \\ (p(0), v) = (\hat{p}, v) & \text{for all } v \in V_0^h(\Omega). \end{cases} \tag{5-5}$$

We consider the following representation for the solution of (5-5),

$$p(x, t) = \sum_{i=1}^{n_v} p_i(t) \phi_i(x), \tag{5-6}$$

where ϕ_i are the usual finite elements basis functions and n_v the number of interior nodes of $\mathcal{T}^h$. Using (5-5) and taking $v = \phi_j$ for $j = 1, ..., n_v$, we have

$$\begin{cases} \sum_{i=1}^{n_v} p_i'(t)(\phi_i, \phi_j) + \sum_{i=1}^{n_v} p_i(t) a(\phi_i, \phi_j) = (f(p), \phi_j) \quad j = 1, ..., n_v,\ t \in I, \\ \sum_{i=1}^{n_v} p_i(0)(\phi_i, \phi_j) = (\hat{p}, \phi_j) \quad j = 1, ..., n_v. \end{cases} \tag{5-7}$$

The equivalent continuous-time matrix form of (5-7) is

$$\begin{aligned} M\partial_t p + Ap &= b(p), \\ Mp(0) &= \hat{p}, \end{aligned} \tag{5-8}$$

where the vector $\hat{p}$ is $\hat{p} = [\int_\Omega p_D \phi_j]$, and the matrices A, M and the vector b are given by

$$u^T A v = \int_\Omega \kappa \nabla u \nabla v, \quad u^T M v = \int_\Omega \kappa u v \quad \text{and} \quad v^T b = \int_\Omega f(p) v, \quad \text{for all } u, v \in V_0^h(\Omega). \tag{5-9}$$

We introduce a coarse-scale mesh $\mathcal{T}^H$, where H indicates the coarse-mesh size, see Section 3.3. In practical applications, the coarse-grid does not resolve all the variations and discontinuities of the coefficient κ. A main goal in multiscale methods is to construct approximation strategies to mimic fine-grid approximation properties but only computing solutions of linear systems at the coarse-scale. The GMsFEM is a multiscale method designed to obtain good approximation of high-contrast multiscale problems. We next review some important aspects in the construction of GMsFEM basis functions. See [28, 1, 13] and references therein for further details.

We denote by $\{y_i\}_{i=1}^{N_v}$ the vertices of the coarse mesh $\mathcal{T}^H$ and define the neighborhood of each node y_i by

$$\omega_i = \bigcup \left\{ K \in \mathcal{T}^H : y_i \in \bar{K} \right\}.$$

See Figure **3-3** for an illustration of coarse elements and coarse neighborhoods.
Note that $\Omega = \bigcup_{y_i \in \tau^H} \{\omega_i\}$. Let $\{\chi_i\}_{i=1}^{N_v}$ be a partition of unity subordinated to the covering $\{\omega_i\}$ and constructed such that $|\nabla \chi_i| \leq \frac{1}{H}$, $i = 1, 2, 3, ..., N_v$, where N_v is the number of nodes in $\mathcal{T}^H$. See [28, 1, 13] for examples of different partition of unity functions that can be used. Now, define the auxiliary coefficient $\widetilde{\kappa}$ by

$$\widetilde{\kappa} = \kappa \sum_{j=1}^{N_v} H^2 |\nabla \chi_j|^2.$$

Coefficient $\widetilde{\kappa}$ can be interpreted as a total pointwise energy for the functions in the partition of unity $\{\chi_i\}_{i=1}^{N_v}$. We define the following local bilinear forms,

$$a^{\omega_i}(p, v) = \int_{\omega_i} \kappa \nabla p \nabla v \quad \text{and} \quad m^{\omega_i}(p, v) = \int_{\omega_i} \widetilde{\kappa} p v \quad \text{for all } p, v \in H^1(\omega_i), \tag{5-10}$$

for every neighborhood ω_i. Also, define $\widetilde{V}(\omega_i) = \{v \in H^1(\omega_i) : v = 0 \text{ on } \partial\omega_i \cap \partial\Omega\}$ if $\partial\omega_i \cap \partial\Omega$ is non-empty and $\widetilde{V}(\omega_i) = \left\{ v \in H^1(\omega_i) : \int_{\omega_i} v = 0 \right\}$ otherwise. We consider the local generalized eigenvalue problem

$$a^{\omega_i}(\psi, z) = \sigma^{\omega_i} m^{\omega_i}(\psi, z) \quad \text{for all } z \in \widetilde{V}(\omega_i), \tag{5-11}$$

with eigenfunction $\psi \in \widetilde{V}(\omega_i)$ and eigenvalue σ. We order eigenvalues as $\sigma_1^{\omega_i} \leq \sigma_2^{\omega_i} \leq ....$ and select the the eigenfunctions corresponding to small eigenvalues. We define the set of GMsFEM coarse basis functions by pointwise multiplication as follows,

$$\Phi_{i,\ell} = \chi_i \psi_\ell^{\omega_i}, \quad \text{for } 1 \leq i \leq N_v \text{ and } 1 \leq \ell \leq L_i, \tag{5-12}$$

where L_i denotes the number of basis functions on the coarse neighborhood ω_i. The support of $\Phi_{i,\ell}$ is ω_i and we remark that there may be multiple basis functions corresponding to this neighborhood since in general $L_i \geq 1$. See [37, 57, 18] for discussion on how to chose L_i. We define the coarse GMsFEM space by

$$V_0 = \text{span}\{\Phi_{i,\ell} = \chi_i \psi_\ell^{\omega_i}, \quad i = 1, \ldots, N_v, \quad \ell = 1, \ldots, L_i\}. \tag{5-13}$$

Let p_D^H be a discrete interpolation of the boundary data p_D. We denote the classical multiscale solution by $p_{ms}(t)$ with $p_{ms}(t) - p_D^H \in V_0$ and such that

$$(\partial_t p_{ms}, v) + a(p_{ms}, v) = F(p_{ms}; v) \quad \text{ for all } v \in V_0. \tag{5-14}$$

We construct the coarse-scale matrix of basis functions,

$$R_0^{\mathrm{T}} = [\Phi_{i,1}, ..., \Phi_{i,L_i}],$$

where $\Phi_{i,\ell}$ was introduced in (5-12). We then define the coarse-scale stiffness matrix and mass matrix by

$$A_0 = R_0 A R_0^{\mathrm{T}} \quad \text{ and } M_0 = R_0 M R_0^{\mathrm{T}}, \tag{5-15}$$

respectively. The coarse-scale load vector is given by $b_{ms} = R_0 b$. With this, we can define the matrix coarse-scale nonlinear system associated to (5-14) as

$$M_0 \partial_t p_{ms} + A_0 p_{ms} = b_{ms}(p_{ms}). \tag{5-16}$$

This matrix problem is the GMsFEM coarse-scale version of the fine-scale matrix system (5-8). Different time stepping methods could be implemented for either (5-8) or (5-16). We could solve system (5-16) and then downscale the final time coarse-scale solution of this matrix problem. However, the approximation may deteriorate as the time advances. We will upscale (project on coarse space) residual vectors and downscale (to the fine-grid) coarse solution at each time step. If h_k denotes the time step, in order to compute the next time solution $p^h(t + h_k)$ from the current time solution $p^h(t)$, we proceed as follows:

1. **Fine-scale residual:** Compute a fine-mesh residual $r^h(t + h_k)$ using the information from the previous time step. Computation of the residual only involves find-grid matrix times find-grid vector products. It does not require solution of fine-grid linear systems neither the computation of functions of fine-grid matrices.

2. **Up-scaling:** Perform up-scaling of the residual vector to obtain a coarse-scale residual: $r_0^H = R_0 r_n^h$.

3. **Coarse-scale solve:** Solve the linear systems and/or function of matrices using coarse-scale matrices A_0, M_0. Here, we obtain a coarse-scale vector representing either the time increment w_0^H or the next time approximation of the solution $p_0^H(t + h_k)$.

4. **Down-scaling:** Compute the next time approximation on the fine grid by downscaling the product of 3. above: here we have, $p^h(t+h_k) = p^h(t) + R_0^T w_0^H$ or $p^h(t+h_k) = R_0^T p_0^H(t+h_k)$.

In the next section, we exemplify this procedure for finite difference and exponential integration time discretizations.

5.3 Time discretizations combined with GMsFEM spatial approximation

We now solve the ODE matrix systems (5-8)-(5-16) in an interval $I = [0, T]$ with the procedure described above. We consider a uniform partition $\{t_0, t_1, \dots, t_M\}$ of I with element size $h_k = T/M$ and approximate the coefficients $p_{ms}(t)$ at the partition points.

5.3.1 GMsFEM finite difference (GMsFEM-FD)

For the FD method, we use the semi-implicit θ-scheme,

$$M\left(\frac{p^{k+1}-p^k}{h_k}\right) = \theta\left(b^{k+1} - Ap^{k+1}\right) + (1-\theta)\left(b^k - Ap^k\right), \tag{5-17}$$

where $\theta \in [0,1]$ and we call $p^k \approx p(t_k)$. We obtain the following solution for p^{k+1}

$$p^{k+1} = (M + h_k\,\theta\,A)^{-1}\left((M - h_k(1-\theta)A)\,p^k + h_k\left((1-\theta)b^k + \theta b^{k+1}\right)\right), \tag{5-18}$$

where $b^k = b(p^k)$. By taking $\theta = \frac{1}{2}$, we obtain the Crank-Nicholson scheme and by taking $\theta = 1$ we obtain the backward Euler method. The term b^{k+1} is unknown at each time step k so we consider a predictor-corrector algorithm as follows. We set $b^{k+1} = b^k$ to predict p^{k+1} by solving (5-18) and obtain a new value of b^{k+1} from it, then we use it to correct p^{k+1} by solving (5-18) again. We can repeat this until we achieve the desired precision.

Remark 23. *From our numerical experiments we found out that, in the presence of high-contrast coefficients, the stability of the θ-scheme is deteriorated with the contrast, specially for $\theta < 1$ due to the presence of the term $(1-\theta)(b^k - Ap^k)$ on the right hand side of the linear system in* (5-18)*. For this reason we consider only the case $\theta = 1$ with small enough time step size.*

As mentioned before, performing this calculation using fine-scale matrices is not practical because this matrix system has a huge dimension and is very ill-conditioned. In particular, for (5-18), we see that the computation of the linear system solution, $(M + h_k\,\theta\,A)^{-1}$ is the real bottleneck for this method in terms of computational time. We propose the use of

GMsFEM for the approximation of the solution of this linear system. Therefore, we arrive to the GMsFEM-FD method

$$p^{k+1} = R_0^T (M_0 + h_k\,\theta\, A_0)^{-1} R_0 \left((M - h_k(1-\theta)A)\, p^k + h_k \left((1-\theta) b^k + \theta b^{k+1} \right) \right). \tag{5-19}$$

Note that, before using the GMsFEM approximation of the solution of the linear system, we upscale the residue of the previous iteration to the coarse mesh. After computing the solution in the coarse space, we downscale it back to the fine mesh. In summary, we perform the following computation in each time step:

1. Fine scale residue: $r_h = (M - h_k(1-\theta)A)\, p^k + h_k \left((1-\theta) b^k + \theta b^{k+1} \right)$,
2. Up-scaling of residue: $r_H = R_0 r_h$,
3. Coarse-scale solve: $p_H^{k+1} = (M_0 + h_k\,\theta\, A_0)^{-1} r_H$,
4. Downscaling of solution: $p_h^{k+1} = R_0^T p_H^{k+1}$.

5.3.2 Exponential Integrator (EI)

We know consider the exponential integrator method for integration in time. From (5-16), we derive the following equivalent system

$$\begin{cases} \partial_t p + Np = F(p) \quad \text{in} \quad I, \\ p(0) = p_0, \end{cases} \tag{5-20}$$

where $N = M^{-1}A$, $F = M^{-1}b(p)$ and $p_0 = M^{-1}\hat{p}$.
We now use exponential integrators to solve (5-20). All this methods are based on the following integral representation of the solution of (5-20) which is called variation-of-constants formula

$$p(t_k) = e^{-h_k N} p(t_{k-1}) + \int_0^{h_k} e^{(\tau - h_k)N} F(p(t_{k-1} + s)) d\tau. \tag{5-21}$$

The main idea here is to find an approximation of the nonlinear term in the variational formula by an algebraic polynomial. In the case of linear problems, the integral in (5-21) is approximated using exponential quadrature rules. Taking s quadrature points $c_i \in [0,1]$, we have

$$p^k = e^{-h_k N} p^{k-1} + h_k \sum_{i=0}^{s} b_i(-h_k N) F_i, \tag{5-22}$$

where $p^k \approx p(t_k)$, $F_i = F(t_{k-1} + c_i h_k)$, and $b_i(z)$, satisfies the following recurrence relations

$$\begin{aligned} z b_0(z) &= e^z - 1 \\ z b_{i+1}(z) + \frac{1}{(i+2)!} &= \left(1 - \frac{1}{(i+2)! z} \right) b_i(z) + ... + (1-z) b_0(z). \end{aligned} \tag{5-23}$$

The coefficient b_i can be rewritten as a linear combination of the φ-functions,

$$\begin{cases} \varphi_0(z) &= e^z \\ \varphi_p(z) &= \displaystyle\int_0^1 e^{(1-\theta)z}\frac{\theta^{p+1}}{(p-1)!}d\theta, \quad p \geq 1, \end{cases} \tag{5-24}$$

which satisfy the following recurrence relation

$$\varphi_{p+1}(z) = \frac{1}{z}\left(\varphi_p(z) - \frac{1}{p!}\right). \tag{5-25}$$

We can then construct the following iterative method. If we select one point $c_1 \in [0,1]$, we have $b_0(z) = \varphi_1(z)$ from (5-25) that $e^z = z\varphi_1(z) + 1$, we obtain the following method

$$p^k = p^{k-1} + h_k\varphi_1(-h_kN)\left(F_1 - Np^{k-1}\right). \tag{5-26}$$

In the case of semilinear problems, the construction of exponential integrators becomes more intricate. This is due to the dependence of F on p, which necessitates the use of internal stages to approximate the solution at various integration points. For a detailed description of the construction of exponential integrators of Runge-Kutta type, we refer to [41]. However, in this article, we focus specifically on the lowest order Exponential Runge-Kutta method. In equation (5-21), we approximate F by employing the value of the solution at the previous time step, denoted as $F^{k-1} = F(p^{k-1})$. By using this approximation, we derive the following time-marching scheme

$$p^k = p^{k-1} + h_k\varphi_1(-h_kN)\left(F^{k-1} - Np^{k-1}\right), \tag{5-27}$$

which is called, the Exponential Euler method.
Observe that the classical eigenvalue problem ([39, 45])

$$-h_kNq = \lambda q,$$

is related to the generalized eigenvalue problem

$$-h_kAq = \lambda Mq.$$

Since M and A are symmetric and positive definite and $N = M^{-1}A$ we factor

$$-h_kN = QDQ^{-1},$$

where the columns of Q are the eigenvectors of $-h_kN$ or the generalized eigenvectors of $-h_kA$ with respect to M. The matrix D is the diagonal matrix of eigenvalues of $-h_kN$ that are also the generalized eigenvalues of $-h_kA$ with respect to M. Given that the eigenvectors q_i are orthonormal with respect to the inner product $(u,v)_M = u^TMv$, we have

$$Q^TMQ = I.$$

Now, using $Q^{-1} = Q^T M$, we rewrite

$$-h_k N = QDQ^T M. \tag{5-28}$$

Using (5-24) and (5-28), we obtain

$$\varphi_p(-h_k N) = Q\varphi_p(D)Q^T M. \tag{5-29}$$

If we use this in (5-27), we obtain

$$\begin{aligned} p^k &= p^{k-1} + h_k Q\varphi_1(D)Q^T M\left(F^{k-1} - Np^{k-1}\right) \\ &= p^{k-1} + h_k Q\varphi_1(D)Q^T\left(MF^{k-1} - Ap^{k-1}\right). \end{aligned} \tag{5-30}$$

As before, the implementation of this iteration using fine scale matrices is inadequate due to the very large computational time needed to numerically approximate the matrix functions. Instead, we use a GMsFEM approximation of the eigenvalue problem to speed up the computations of the φ-functions. As a result, we obtain the GMsFEM-EI iteration. In particular we propose the approximation

$$\varphi_p(-h_k N) \approx R_0^T \varphi_p(-h_k N_0)R_0, \tag{5-31}$$

where $N_0 = M_0^{-1}A_0$ with A_0 and M_0 defined in (5-15). The associated eigenvalue problem is

$$-h_k A_0 q_0 = \lambda M_0 q_0 \quad \text{or} \quad -h_k N_0 = Q_0 D_0 Q_0^T M_0. \tag{5-32}$$

Here, Q_0 is the matrix whose columns are GMsFEM-coarse-scale eigenvectors. Note that we apply the approximation in (5-31) to a fine-scale operator. Given the previous time fine-scale approximation, we first upscale the residual vector to the coarse space, then use the function computed at coarse resolution in the GMsFEM space, and finally, we downscale the result to the fine-grid. For instance, the approximation of the iteration in equation (5-27) using the GMsFEM-EI iteration is given by

$$p^k = p^{k-1} + h_k R_0^T \varphi_1(-h_k N_0)R_0\left(F^{k-1} - Np^{k-1}\right). \tag{5-33}$$

Note that we can use any other procedure to compute $\varphi_1(-h_k N_0)$. In terms of the coarse approximation of eigenvectors and eigenvalues in (5-32) we get

$$p^k = p^{k-1} + h_k R_0^T Q_0 \varphi_1(D_0)Q_0^T R_0\left(MF^{k-1} - Ap^{k-1}\right). \tag{5-34}$$

In summary we compute as follows

1. Fine scale residue: $r_h = MF^{k-1} - Ap^{k-1}$,
2. Up-scaling of residue: $r_H = R_0 r_h$,
3. Coarse-scale function of matrix: $\delta_H^{k+1} = Q_0\varphi_1(D_0)Q_0^T r_H$,

4. Downscaling of solution: $p_h^{k+1} = p_h^k + R_0^T \delta_H^{k+1}$.

For the computation of φ-functions, we use two different methods. The first one is the procedure described in (5-34) that requires the computation of a coarse-scale global generalized eigenvaluen problem. Note that the corresponding fine-scale formula in equation (5-30) will require the computation of a fine-scale global generalized eigenvalue problem involving large sparse and ill conditioned operators. The second method employs the MATLAB package called EXPINT. It is presented in [8] and uses Padé approximations.

Remark 24. *Since we are going to solve the time evolution on the coarse space, as initial condition we use the orthogonal projection of the initial condition*

$$\hat{p}_0 = R_0^T M_0^{-1} R_0 M p_0. \tag{5-35}$$

That is, we solve the approximated initial condition problem

$$\begin{cases} \partial_t p + Np = F(p) \quad in \quad I, \\ p(0) = \hat{p}_0. \end{cases} \tag{5-36}$$

5.4 Numerical examples

We now consider a high-contrast coefficient example. With this example we want to show the effects of the contrast in the stability of the EI approximation. Our reference solution is obtained in the fine-mesh using implicit Euler scheme with a small enough time step size.

5.4.1 Linear problem

We consider the problem,

$$\begin{cases} \partial_t p - \operatorname{div}(\kappa(x)\nabla p) = 0 & \text{in} \quad \Omega = [0,1]^2, \\ p(0, x_1, x_2) = x_1(1-x_1)x_2(1-x_2), & \\ p(t, x_1, x_2) = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{5-37}$$

where the high-contrast coefficient κ is depicted in Figure **5-1**. We consider the case where the value of the contrast is set to 100.
The initial condition is given for the following function,

$$u(0, x_1, x_2) = x_1(1-x_1)x_2(1-x_2), \qquad \text{for all } (x_1, x_2) \in \Omega. \tag{5-38}$$

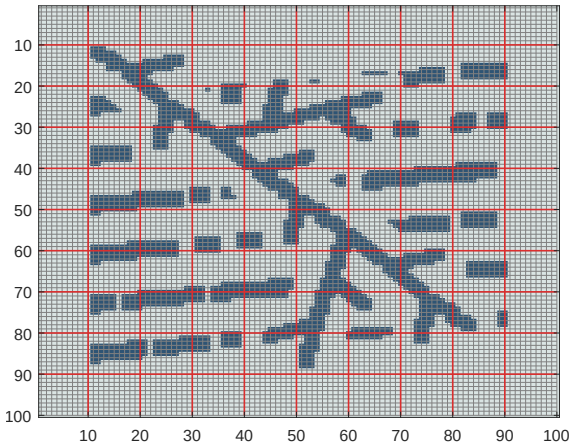

Figure 5-1: High-contrast coefficients used in the numerical experiment. High-conductivity channels in black color. In our numerical experiments we use coefficients of background 1 and high-contrast value 10^2.

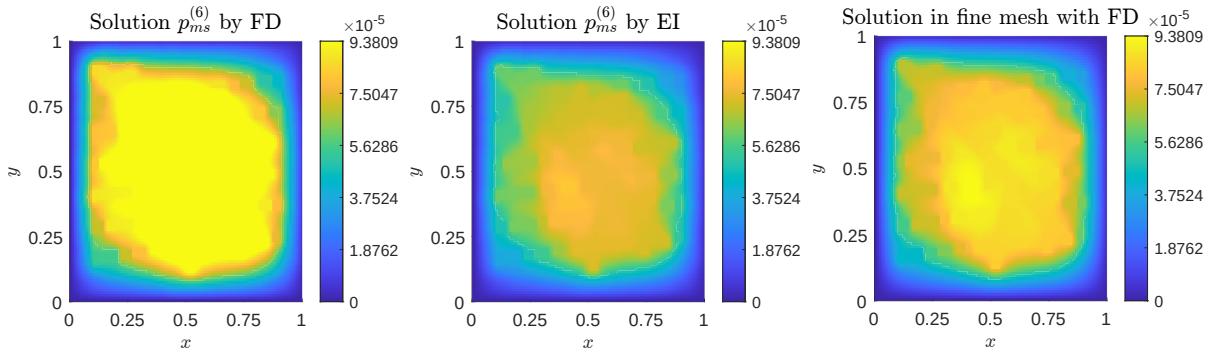

Figure 5-2: Final time ($T = 0.2$) solution for problem (5-37). Computed solution using the MsFEM-FD with 6 basis functions in each neighborhood and 50 times steps (left). Computed solution using the MsFEM-EI with 6 basis functions in each neighborhood and 50 times steps (center). fine mesh solution with 30000 time steps (right).

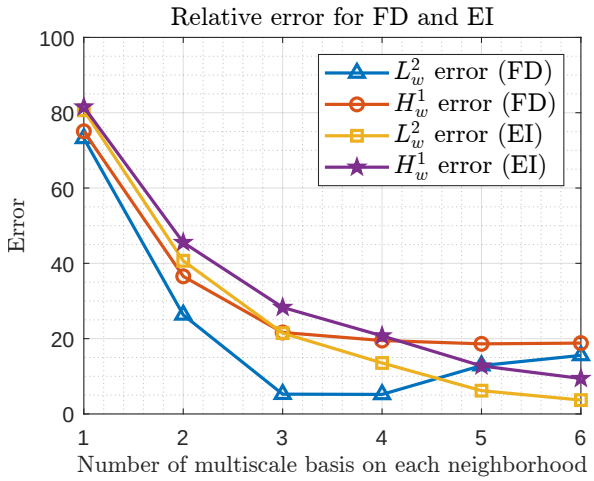

Figure 5-3: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-37). The horizontal axis corresponds to the number of basis functions in each neighborhood used in the GMsFEM coarse spaces.

	FD % Error		EI (eig) % Error		EI (EXPINT) % Error	
L_i	L_w^2	H_w^1	L_w^2	H_w^1	L_w^2	H_w^1
1	73.1	75.0	80.6	81.5	80.6	81.5
2	26.3	36.5	40.6	45.5	40.6	45.5
3	5.2	21.6	21.4	28.2	21.4	28.2
4	5.1	19.4	13.5	20.7	13.5	20.7
5	12.8	18.6	6.1	12.7	6.1	12.7
6	15.5	18.8	3.6	9.4	3.6	9.4

Table 5-1: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-37). In the last column we have added the relative error when the matrix functions are computed using `MatLab expint`.

5.4.2 Semilinear problem with low contrast

We consider the problem,

$$\begin{cases} \partial_t p - \operatorname{div}(\kappa(x)\nabla p) = -p(1-p)(1+p) & \text{in} \quad \Omega = [0,1]^2, \\ p(0, x_1, x_2) = x_1(1-x_1)x_2(1-x_2), & \\ p(t, x_1, x_2) = 0 & \text{on} \quad \partial\Omega, \end{cases} \tag{5-39}$$

where the high-contrast coefficient κ is depicted in Figure **5-1**. We consider the case where the value of the contrast is set to 10.

The initial condition is given for the following function,

$$u(0, x_1, x_2) = x_1(1 - x_1)x_2(1 - x_2), \qquad \text{for all } (x_1, x_2) \in \Omega. \tag{5-40}$$

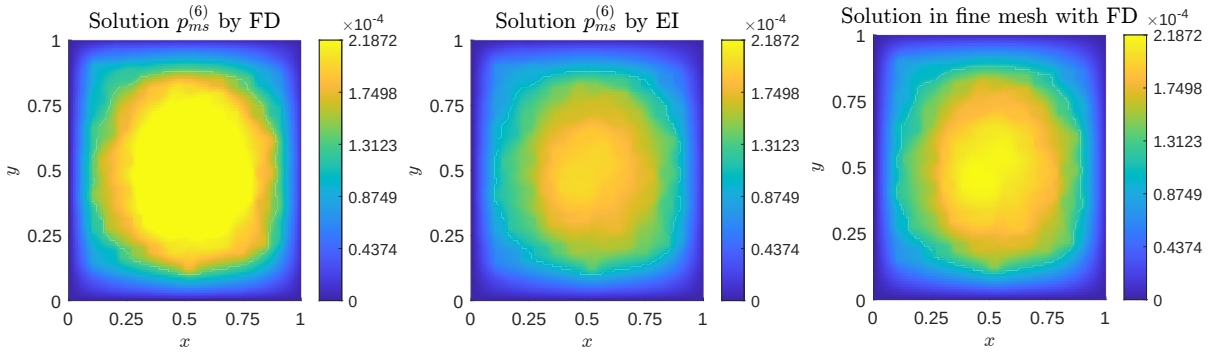

Figure 5-4: Final time ($T = 0.2$) solution for problem (5-39). Computed solution using the MsFEM-FD with 5 basis functions in each neighborhood and 50 times steps (left). Computed solution using the MsFEM-EI with 5 basis functions in each neighborhood and 60 times steps (center). fine mesh solution with 30000 time steps (right) and contrast 10.

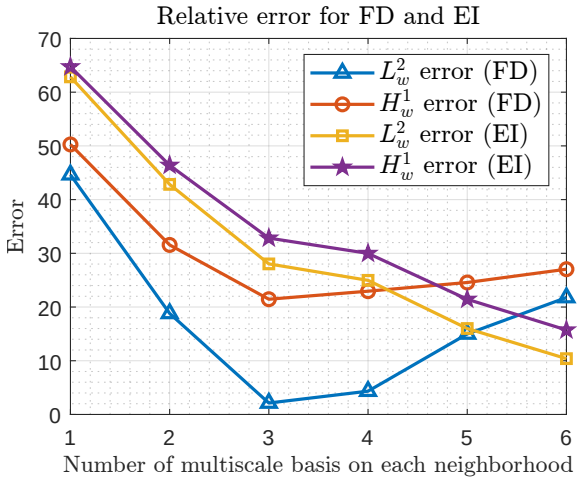

Figure 5-5: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-39). The horizontal axis corresponds to the number of basis functions in each neighborhood used in the GMsFEM coarse spaces and contrast 10.

	FD % Error		EI (eig) % Error		EI (EXPINT) % Error	
L_i	L^2_w	H^1_w	L^2_w	H^1_w	L^2_w	H^1_w
1	44.62	50.26	62.84	64.73	62.84	64.73
2	18.81	31.56	42.83	46.38	42.83	46.38
3	2.13	21.45	28.03	32.83	28.03	32.83
4	4.33	22.94	24.96	29.99	24.96	29.99
5	14.98	24.56	16.00	21.44	16.00	21.44
6	21.76	27.04	10.39	15.72	10.39	15.72

Table 5-2: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-39). In the last column we have added the relative error when the matrix functions are computed using `MatLab expint` and contrast 10.

5.4.3 Semilinear problem with high contrast

We consider the problem (5-39), where the high-contrast coefficient κ is depicted in Figure **5-1**. We consider the case where the contrast value is set to 100.
The initial condition is given for the following function,

$$u(0, x_1, x_2) = x_1(1 - x_1)x_2(1 - x_2), \qquad \text{for all } (x_1, x_2) \in \Omega. \tag{5-41}$$

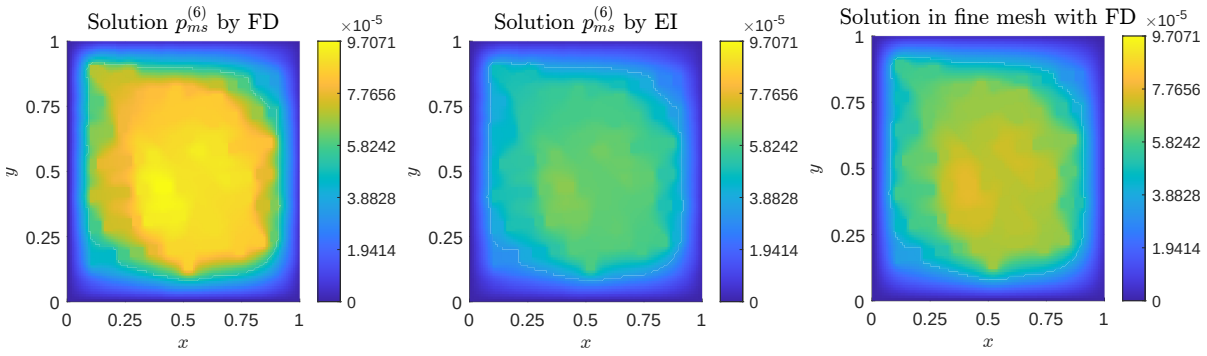

Figure 5-6: Final time ($T = 0.2$) solution for problem (5-39). Computed solution using the MsFEM-FD with 5 basis functions in each neighborhood and 50 times steps (left). Computed solution using the MsFEM-EI with 5 basis functions in each neighborhood and 60 times steps (center). fine mesh solution with 30000 time steps (right) and contrast 100.

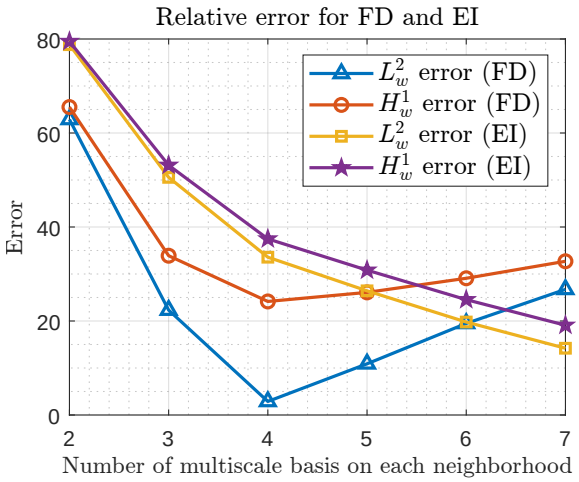

Figure 5-7: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-39). The horizontal axis corresponds to the number of basis functions in each neighborhood used in the GMsFEM coarse spaces and contrast 100.

	FD % Error		EI (eig) % Error		EI (EXPINT) % Error	
L_i	L_w^2	H_w^1	L_w^2	H_w^1	L_w^2	H_w^1
2	62.88	65.52	78.79	79.48	78.79	79.48
3	22.31	33.88	50.56	53.14	50.56	53.14
4	2.88	24.15	33.57	37.46	33.57	37.46
5	10.88	26.08	26.41	30.81	26.41	30.81
6	19.47	29.08	19.78	24.54	19.78	24.54
7	26.72	32.70	14.20	19.09	14.20	19.09

Table 5-3: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-39). In the last column we have added the relative error when the matrix functions are computed using `MatLab` `expint` and contrast 100.

5.4.4 Semilinear problem withmedium SPE10

We consider the problem (5-39), where the high-contrast coefficient κ is depicted in Figure **4-5**. We consider the case where the contrast value is set to 100.
The initial condition is given for the following function,

$$u(0, x_1, x_2) = x_1(1 - x_1)x_2(1 - x_2), \qquad \text{for all } (x_1, x_2) \in \Omega. \tag{5-42}$$

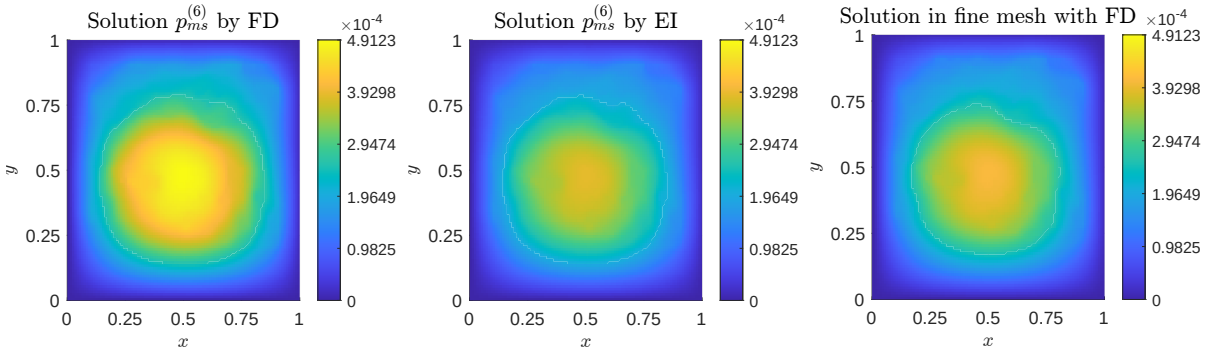

Figure 5-8: Final time ($T = 0.2$) solution for problem (5-39). Computed solution using the MsFEM-FD with 5 basis functions in each neighborhood and 50 times steps (left). Computed solution using the MsFEM-EI with 5 basis functions in each neighborhood and 50 times steps (center). fine mesh solution with 30000 time steps (right) and contrast 10.

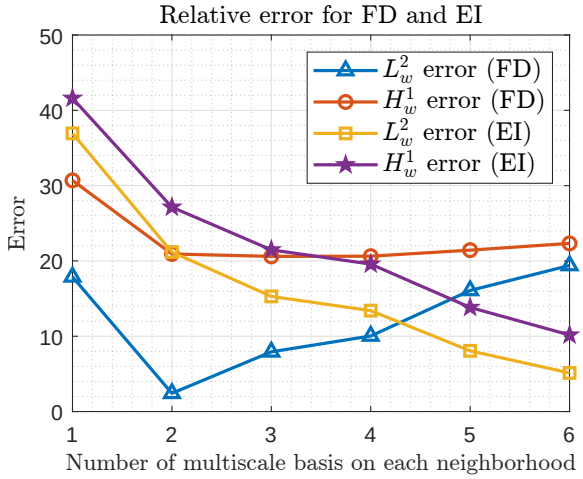

Figure 5-9: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-39). The horizontal axis corresponds to the number of basis functions in each neighborhood used in the GMsFEM coarse spaces and contrast ten.

	FD % Error		EI (eig) % Error		EI (EXPINT) % Error	
L_i	L^2_w	H^1_w	L^2_w	H^1_w	L^2_w	H^1_w
1	17.88	30.69	36.93	41.60	36.93	41.60
2	2.42	20.92	21.17	27.13	21.17	27.13
3	7.92	20.60	15.29	21.47	15.29	21.47
4	10.03	20.63	13.39	19.57	13.39	19.57
5	16.07	21.43	8.06	13.82	8.06	13.82
6	19.39	22.32	5.12	10.16	5.12	10.16

Table 5-4: The weighted L^2 and H^1 errors between the reference and the coarse-scale solution at the final time $T = 0.2$ for problem (5-39). In the last column we have added the relative error when the matrix functions are computed using `MatLab` `expint` and contrast 10.

6 Final comments and future work

In this thesis, the Generalized Multiscale Finite Element Method (GMsFEM) was studied, an error analysis was presented, and two importants applications in which this method is an ideal for efficient approximation are studied. These applications are important because, with these, we show that the GMsFEM is flexible enough to be applied to non-linear problems and variational inequalities as well as exponential integration. When using FEM and the coefficient is of high-contrast, one of the problems is that the system is poorly conditioned and unstable in time due to the discontinuities of the medium. When using the GMsFEM, basis functions are constructed from local eigenvectors, which gather essential information about the high-contrast medium. This enables the reduction of the system while maintaining good accuracy.
In the first chapter, we study some crucial topics to start the investigation: Finite Element Methods (FEM), maximal operators, the method of duality, Yosida's approximation and exponetial integrator. For the second chapter, we obtained error estimates for the GMsFEM approximation of high-contrast multiscale problems. This construction uses local Neumann eigenvectors in each neighborhood and Dirichlet eigenvectors in each cell to construct finite element basis functions. The analysis is based first on the construction of eigenfunctions and the definition of the norms used for the error estimates, where they measure the decay of the expansion of the solution in terms of local eigenfunctions. The norms in the interpolation error estimates can be bounded by the L_2 norm when rescaling the forcing term. For the analysis, we assume that the solution can be approximated by the sum of two functions, one with zero flux through the coarse block boundaries and the other with zero value at the coarse block boundaries. This assumption is easily verified for classical regular problems. The introduction of this assumption allowed us to extend and simplify the convergence analysis.

One of the applications studied is the high-contrast multiscale free boundary dam problem, see Section 4-1; the first thing is that the initial problem is an inequality with multivalued operators; here, it is crucial to apply the duality method that by building a convenient Lagrangian rewrites the initial problem into a simpler one; now, the problem is that the system obtained is multivalued, this is when we can use Yosida's approximation to get a system of single-valued equations which are going to be solved using the fixed-point method since FEM would have to be done in each iteration if the system is large and illconditional the process becomes very slow. By using the GMsFEM, results were achieved more efficiently

and with an acceptable error.

A second application is the solution of a parabolic equation combining the GMsFEM and the method of exponential integrators. The Exponential Integrators (EI) method has many advantages as it does not require many iterations to reach the final simulation time compared to other methods, for example, the finite difference method. One of the problems in exponential integrators have is that their calculation is complicated since it involves calculating the exponential of a vary large and ill conditioned matrix. When the system is large, the calculation becomes quite complicated. The GMsFEM is applied for the residue of the EI formula; that is, the matrix is project to the coarse mesh to calculate the exponential of the matrix then, and then it is sent back to the fine mesh to calculate the iterations of the EI. Finally, the error calculation is compared with finite differences, using tiny time steps to obtain a reasonable estimate of the error.

As we see from these applications studied in this work, an essential contribution of using the GMsFEM is when we have to do post-processing or find solutions in each iteration of a numerical method since it reduces the computational cost of all these processes, in addition to reducing instability and poor conditioning of the system when we are working with high contrast coefficients.

6.1 Future work

In the future, many different applications and adaptations can be made concernig the GMsFEM; one of them that is currently under investigation is the partial differential equation model of resin transfer molding in [51]. We interpret the model as a time-dependent free boundary problem. We apply a time discretization combined with a duality method to deal with nonlinear multivalued terms, [10, 49] . The final solver can be understood as a fixed-point iteration to compute the time step changes to simulate time evolution. This approach was combined with generalized multiscale finite element methods to approximate a heterogeneous dam problem in [7, 13].

6.1.1 A free boundary problem

We consider a bounded two dimensional rectangular domain D and let $\partial D = \Gamma_1 \cup \Gamma_2 \cup \Gamma_3 \cup \Gamma_4$ denote its boundary, where Γ_1 is the inlet, Γ_2 and Γ_3 is the perfectly sealed boundary, and Γ_4 is the outlet. See Figure **6-1**

$$\frac{\partial(\phi\,\theta(p))}{\partial t} - \operatorname{div}(\kappa \nabla p) = 0, \quad p \geq 0, \quad \theta \in H(p), \tag{6-1}$$

where $H(\cdot)$ denotes the multivalued Heaviside operator, so that for positive pressure ($p > 0$) the porous media is fully saturated ($\theta = 1$) and $\theta \in [0, 1)$ when $p = 0$ in the non saturated region. Here, ϕ and κ are the porosity and permeability of the porous media. In order to pose the strong formulation of the free-boundary problem, the set of equations (6-1) is completed with the following boundary conditions:

- $p = p_1$ on Γ_1,
- $p = 0$ in Γ_4,
- $\big(\kappa \nabla p\big) \cdot \boldsymbol{n} \geq 0$ in Γ_4,
- $\Big(- \kappa \nabla p\Big) \cdot \boldsymbol{n} = 0$ in $\Gamma_2 \cup \Gamma_3$.

In previous equations $\boldsymbol{n}$ represents the unitary outwards normal vector to the boundary ∂D.

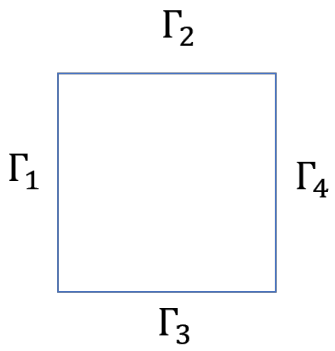

Figure 6-1: Domain configuration illustration. The edge Γ_1 is the inlet boundary. The edges Γ_2 and Γ_3 represent impermeable boundaries and Γ_4 is the outlet boundary.

We assume that the domain D is the union of finitely many sub-domains, that is,

$$D = \bigcup_{i=1}^{N_S} D_i,$$

where $\{D_i\}$ is a non-overlapping decomposition of D. The permeability coefficient can be written as,

$$\kappa(x) = k_i(x) \text{ for } x \in D_i,$$

$i = 1, 2, \ldots, N_S$, where κ_i is a bounded smooth function (that may have oscillations) in D_i.

Time discretization

For the purpose of the time discretization, we introduce an uniform finite differences time mesh with points $t_0, t_1, \ldots t_M$, and a constant time step δt. Next, we introduce a forward in

time approximation of the total derivative by the method of characteristics. More precisely, if we use the notation $f^n(x) = f(t^n, x)$ for $n = 0, 1, 2, \ldots$, at each time t^{n+1} we obtain the strong formulation of the problem discretized in time,

$$\frac{\theta^{(n+1)} - \theta^{(n)}}{\delta t} - \operatorname{div}(\kappa \nabla p^{(n+1)}) = 0, \quad p^{(n+1)} \geq 0, \theta^{(n+1)} \in H(p^{(n+1)}), \tag{6-2}$$

where the index $n+1$ denotes the approximation at the artificial time t^{n+1} of the introduced time dependent functions.
In order to write a weak form of the problem, we introduce the following functional spaces:

- $V_- = H^1(D) \cap [v|_{\Gamma_4} \leq 0] = \{\psi \in H^1(D); \psi|_{\Gamma_1} = 0; \psi|_{\Gamma_4} \leq 0\}$,
- $V_0 = H_0^1(D, \Gamma_4) = \{v \in H^1(D) : v|_{\Gamma_4} = 0\}$ and note that $V_0 \subset V_-$,
- $V_+ = H^1(D) \cap [v|_{\Gamma_4} \geq 0]$,
- $V_- = H^1(D) \cap [v|_{\Gamma_4} \leq 0]$,
- $W_\alpha = H^1(D) \cap [v|_{\Gamma_1} = \alpha]$,
- $W_0 = H^1(D) \cap [v|_{\Gamma_1} = 0]$.

Next, for $\phi \in V_-$ such that $\phi = 0$ on Γ_1, we get the weak formulation of the problem:
Find $p^{(n+1)} \in V_0 \cap W_\alpha \cap V_+ = V_0 \cap W_\alpha$ and $\theta^{n+1} \in L^\infty(D)$ such that

$$\int_D \left(\theta^{(n+1)} - \theta^{(n)}\right)(\phi - p^{n+1}) + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla(\phi - p^{n+1}) \geq 0, \tag{6-3}$$

jointly with

$$\theta^{(n+1)} \in H(p^{(n+1)}).$$

Consider now the indicatrix function of the convex set V_- defined over $H^1(D)$ by

$$I_{V_-}(v) = \begin{cases} 1 & v \in V_-, \\ +\infty & v \notin V_-. \end{cases}$$

Note that I_{V_-} is a convex semicontinuous function which implies that the subdifferential operator ∂I_{V_-} is a well defined maximal monotone multivalued operator. Moreover, the multivalued operator ∂I_{V_-} can be characterized as follows:

$$\alpha \in \partial I_{V_-}(u) \Longleftrightarrow I_{V_-}(v) - I_{V_-}(u) \geq \langle \alpha, v - u \rangle, \tag{6-4}$$

for all $v \in H^1(D)$. Here $\langle \cdot, \cdot \rangle$ denotes the duality pairing between $H^1(D)$ and its dual space. Therefore, if we define

$$\langle -L(p^{(n+1)}), \phi - p^{(n+1)} \rangle = \int_D (\theta^{(n+1)} - \theta^{(n)})(\phi - p^{n+1}) + \\ + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla(\phi - p^{n+1}),$$

then from inequality (6-3) and the definition of the indicatrix function we get

$$\int_D (\theta^{(n+1)} - \theta^{(n)})(\phi - p^{n+1}) + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla(\phi - p^{n+1}) + \\ + I_{V-}(\phi) - I_{V-}(p^{(n+1)}) \geq 0,$$

or, in compact notation,

$$\langle L(p^{(n+1)}), \phi - p^{(n+1)} \rangle + I_{V-}(\phi) - I_{V-}(p^{(n+1)}) \geq 0,$$

for all $\phi \in W_0 = H^1(D) \cap [v|_{\Gamma_1} = 0]$. Moreover, the previous inequality implies that $L(p^{(n+1)}) \in \partial I_{V-}(p^{(n+1)})$. Therefore, we can introduce the new variable

$$q^{(n+1)} = L(p^{(n+1)}) \quad \text{with} \quad q^{(n+1)} \in \partial I_{V-}(p^{(n+1)}).$$

Note that we can identify $H^1(D)$ with the sum $H^{1/2}(\partial D) \oplus H_0^1(D)$ and therefore we can identify the dual space of $H^1(D)$ with $H^{-1/2}(\partial D) \oplus H^{-1}(D)$. Using this representation, equation (6-4) and the fact that $p^{n+1} \in V_0 \cap W_\alpha$ we can write the time discretized problem at step n as:

Find $p^{(n+1)} \in V_0 \cap W_\alpha$ and $\theta^{n+1} \in L^\infty(D)$, such that

$$\int_D \theta^{(n+1)} \phi + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla \phi \\ + \delta t \int_{\Gamma_4} q^{(n+1)} \phi \;=\; \int_D \theta^{(n)} \phi, \tag{6-5}$$

for all $\phi \in W_0$, jointly with the following multivalued nonlinear equations

$$\begin{cases} q^{(n+1)} \in \partial I_{V-}\left(p^{(n+1)}\right), \\ \theta^{(n+1)} \in H\left(p^{(n+1)}\right). \end{cases} \tag{6-6}$$

A duality method for nonlinear terms

In order to solve (6-5)-(6-6), we follow the methodology used in [9, 49] to deal with nonlinear terms associated to multivalued operators in (6-6). These techniques are based on duality methods for nonlinear maximal monotone operators and are here applied to the multivalued Heaviside and subdifferential operators. In the seminal article [10], this duality methods have been introduced for solving variational inequalities.

For this purpose, we first recall the concept of Yosida approximation. Let G be a maximal monotone operator and let ω and λ be non-negative real numbers such that $\omega\lambda < 1$. The resolvent of G is defined by,

$$J_\lambda^\omega = ((1 - \omega\lambda I) + \lambda G)^{-1}.$$

Next, we introduce the Yosida approximation of $G - \omega I$ of parameter λ, which is defined by

$$G_\lambda^\omega := \frac{I - J_\lambda^\omega}{\lambda}.$$

As recalled in [49], it can be proved that $u \in G(y) - \omega y$ is equivalent to $u = G^{\omega}_{\lambda}(y + \lambda u)$, for further details see also the seminal article [10]. Note that the first expression is written in terms of the multivalued operator while the second one is a nonlinear equation for u in terms of an univalued Yosida operator.

Next, in terms of the non-negative parameters ω_1 and ω_2, we introduce the new variables

$$\alpha^{(n+1)} = q^{(n+1)} - \omega_1 p^{(n+1)} \quad \text{and} \quad \beta^{(n+1)} = \theta^{(n+1)} - \omega_2 p^{(n+1)}. \tag{6-7}$$

Therefore, from (6-6) we have

$$\alpha^{(n+1)} \in \partial I_{V_-}(p^{(n+1)}) - \omega_1 p^{(n+1)},$$

and

$$\beta^{(n+1)} \in H(p^{(n+1)}) - \omega_2 p^{(n+1)}.$$

We can then write the variational formulation in terms of the new variables in the form

$$\begin{aligned} \int_D (\beta^{(n+1)} + \omega_2 p^{(n+1)})\phi + \delta t \int_D \kappa \nabla p^{(n+1)} \nabla \phi \\ + \delta t \int_{\Gamma_4} (\alpha^{(n+1)} + \omega_1 p^{(n+1)})\phi \;\; = \;\; \int_D \theta^{(n)} \phi. \end{aligned} \tag{6-8}$$

Next, using the previous characterization of the elements of the multivalued operator $G - \omega I$ in terms of the its Yosida approximation for the particular cases $G = \partial I_{V_-}$ and $G = H$, the variational formulation can be equivalently written in the form

$$\begin{aligned} \int_D \kappa \nabla p^{(n+1)} \nabla \phi + \frac{\omega_2}{\delta t} \int_D p^{(n+1)} \phi + \omega_1 \int_{\Gamma_4} p^{(n+1)} \phi \\ = \frac{1}{\delta t} \int_D \theta^{(n)} \phi - \frac{1}{\delta t} \int_D \beta^{(n+1)} \phi - \int_{\Gamma_4} \alpha^{(n+1)} \phi, \end{aligned} \tag{6-9}$$

with

$$\alpha^{(n+1)} = (\partial I_{V_-})^{\omega_1}_{\lambda_1}(p^{(n+1)} + \lambda_1 \alpha^{n+1}), \tag{6-10}$$

and

$$\beta^{(n+1)} = H^{\omega_2}_{\lambda_2}(p^{(n+1)} + \lambda_2 \beta^{n+1}). \tag{6-11}$$

Following [49], we propose to solve (6-9), (6-10) and (6-11) numerically using a fixed point iteration as described in the next paragraphs.
Given $\alpha^{(n+1)}$, $\beta^{(n+1)}$ and $\theta^{(n)}$ we solve equation (6-9) for the pressure and denote the solution by $p^{(n+1)} = \mathcal{L}(\alpha^{(n+1)}, \beta^{(n+1)}, \theta^{(n)})$. So, schematically, we have the following system of coupled equations

$$\alpha^{(n+1)} = (\partial I_{V_-})^{\omega_1}_{\lambda_1} \left(\mathcal{L}(\alpha^{(n+1)}, \beta^{(n+1)}, \theta^{(n)}) + \lambda_1 \alpha^{n+1} \right) \tag{6-12}$$

$$\beta^{(n+1)} = H^{\omega_2}_{\lambda_2} \left(\mathcal{L}(\alpha^{(n+1)}, \beta^{(n+1)}, \theta^{(n)}) + \lambda_2 \beta^{n+1} \right). \tag{6-13}$$

Using the results in [49] it can be seen that, given $\theta^{(n)}$, this system can be solved by a fixed point iteration. To start the fixed point iteration we use previous values of $\alpha^{(n)}$ and $\beta^{(n)}$. The value of θ^n can be updated using (6-7).

For the spatial discretization of the linear problems arising at each step of the fixed point iteration, we consider Finite Elements Methods. For this purpose, let $\mathcal{T}^h$ be a triangular partition of the domain D such that it resolves the variation of the permeability coefficient κ. Consider V the finite element space of piece-wise linear (or bi-linear) finite elements defined on the mesh $\mathcal{T}^h$. At each step of previous iteration, the fully discretized problem can be written in terms of the solution of the following linear system:

$$\left(A + \frac{\omega_2}{\delta t}M + \omega_1 M_{\Gamma_4}\right) p^{(n+1)} = \frac{1}{\delta t}M\theta^{(n)} - \frac{1}{\delta t}M\beta^{(n+1)} - M_{\Gamma_4}\alpha^{(n+1)}. \tag{6-14}$$

In the linear system (6-14), we introduced the following matrices,

$$A = [a_{ij}] \text{ with } a_{ij} = \int_D \kappa \nabla\phi_i \nabla\phi_j,$$

$$M = [m_{ij}] \text{ with } m_{ij} = \int_D \phi_i\phi_j,$$

and

$$M_{\Gamma_4} = [m_{ij;\Gamma_0}] \text{ with } m_{ij;\Gamma_4} = \int_{\Gamma_4} \phi_i\phi_j,$$

Note that at each step n of the algorithm we recursively solve the linear system (6-14) and update the terms α^{n+1} and β^{n+1} in the second member by using (6-12) and (6-13), respectively. In practice, in the numerical examples in a forthcoming section, we consider $\lambda_1 = \lambda_2 = 1$ and $\omega_1 = \omega_2 = 0.5$, so that we fullfil the condition $\lambda_i\omega_i = 0.5$, as in [49]. Note that this condition allows to prove the convergence of the fixed point iteration in [10] for a variational inequality problem. Also, for an elastohydrodynamic problem in magnetic storage devices the convergence is theoretically proved under the same condition in [5]. We also mention that the number of fixed point iterations is chosen to be a constant number independently of the time step iteration.

The idea is to solve (6-14) using the GMsFEM methodology exactly as in Chapter 4. This is under current research.

Bibliography

[1] E. Abreu, C. Diaz, and J. Galvis. A convergence analysis of generalized multiscale finite element methods. *Journal of Computational Physics*, 396:303–324, 2019.

[2] E. Abreu, C. Diaz, J. Galvis, and J. Perez. On the conservation properties in multiple scale coupling and simulation for darcy flow with hyperbolic-transport in complex flows. *Multiscale Modeling & Simulation*, 18(4):1375–1408, 2020.

[3] Awad H. Al-Mohy and Nicholas J. Higham. Computing the action of the matrix exponential, with an application to exponential integrators. *SIAM Journal on Scientific Computing*, 33(2):488–511, 2011.

[4] Todd Arbogast and Mary F. Wheeler. A nonlinear mixed finite element method for a degenerate parabolic equation arising in flow in porous media. *SIAM Journal on Numerical Analysis*, 33(4):1669–1687, 1996.

[5] I. Arregui, J.J. Cendán, C. Parés, and C. Vázquez. Numerical solution of a 1-d elastohydrodynamic problem in magnetic storage devices. *ESAIM: Math. Model. Num. Anali.*, 42:645–665, 2008.

[6] G. Bayada, S. Martin, and C. Vázquez. Homogenization of a nonlocal elastohydrodynamic lubrication problem: a new free boundary model. *Math. Mod. Meth. Appl. Sci.*, 15(12):1923–1956, 2005.

[7] G. Bayada, S. Martin, and C. Vázquez. Homogéneisation du modéle d'Elrod-Adams hydrodynamique. *J. Asymp. Anali.*, 44(1-2):75–110, 2005.

[8] Håvard Berland, Bård Skaflestad, and Will M. Wright. Expint—a matlab package for exponential integrators. *ACM Trans. Math. Softw.*, 33(1):4–es, mar 2007.

[9] A. Bermúdez and J Durany. Numerical solution of steady-state flow through a porous dam. *Comput. Methods Appl. Mech. Engrg.*, 68(1):55–65, 1988.

[10] A. Bermúdez and C. Moreno. Duality methods for solving variational inequalities. *Comput. Math. Appl.*, 7(1):43–58, 1981.

[11] D. Braess. *FINITE ELEMENTS Theory, Fast Solvers, and Applications in Elasticity Theory*. Cambridge University press, Cambridge, 2007.

[12] H. Brezis. *Functional analysis, Sovolev Space and Partial Differential Equations*. Springer, Rutgers University, 2011.

[13] F. Contreras C. Vazquez, J. Galvis. Numerical upscaling of the free boundary dam problem in multiscale high-contrast media. *Journal of Computational and Applied Mathematics*, 367, 2020.

[14] V. M. Calo, Y. Efendiev, and J. Galvis. Asymptotic expansions for high-contrast elliptic equations. *Math. Models Methods Appl. Sci.*, 24(3):465–494, 2014.

[15] V. M. Calo, Y. Efendiev, J. Galvis, and G. Li. Randomized oversampling for generalized multiscale finite element methods. *Multiscale Model. Simul.*, 14(1):482–501, 2016.

[16] Victor M Calo, Yalchin Efendiev, Juan Galvis, and Guanglian Li. Randomized oversampling for generalized multiscale finite element methods. *Multiscale Modeling & Simulation*, 14(1):482–501, 2016.

[17] M. A. Christie and M. J. Blunt. Tenth SPE Comparative Solution Project: A Comparison of Upscaling Techniques. *SPE Reservoir Evaluation Engineering*, 4(04):308–317, 08 2001.

[18] Eric Chung, Yalchin Efendiev, and Thomas Y. Hou. Adaptive multiscale model reduction with generalized multiscale finite element methods. *Journal of Computational Physics*, 320:69–95, 2016.

[19] Eric Chung, Yalchin Efendiev, Sai-Mang Pun, and Zecheng Zhang. Computational multiscale method for parabolic wave approximations in heterogeneous media. *Applied Mathematics and Computation*, 425:127044, 2022.

[20] Eric T. Chung, Yalchin Efendiev, Wing Tat Leung, and Petr N. Vabishchevich. Contrast-independent partially explicit time discretizations for multiscale flow problems. *Journal of Computational Physics*, 445:110578, 2021.

[21] Zeidler E. *Nonlinear functional analysis and its applications I Variational methods and optimization*. Springer science+business media, Springer verlag New York, 1985.

[22] Zeidler E. *Nonlinear functional analysis and its applications II Variational methods and optimization*. Springer science+business media, Springer verlag New York, 1985.

[23] Zeidler E. *Nonlinear functional analysis and its applications III Variational methods and optimization*. Springer science+business media, Springer verlag New York, 1985.

[24] L. Macul E. Abreu1, P. Ferraz. A multiscale recursive numerical method for semilinear parabolic problems. *CILAMCE, PANACM*, 2021.

[25] Y. Efendiev and J. Galvis. Domain decomposition preconditioner for multiscale high-contrast problems. In *Proceedings of DD19*, 2009.

[26] Y. Efendiev and J. Galvis. A domain decomposition preconditioner for multiscale high-contrast problems. In Y. Huang, R. Kornhuber, O. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering XIX*, volume 78 of *Lect. Notes in Comput. Science and Eng.*, pages 189–196. Springer-Verlag, 2011.

[27] Y. Efendiev and J. Galvis. Domain decomposition preconditioner for multiscale high-contrast problems. In Y. Huang, R. Kornhuber, O. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering XIX*, volume 78 of *Lecture Notes in Computational Science and Engineering*, pages 189–196, Berlin, 2011. Springer-Verlag.

[28] Y. Efendiev, J. Galvis, and T. Hou. Generalized multiscale finite element methods. *Journal of Computational Physics*, 251:116–135, 2013.

[29] Y. Efendiev, J. Galvis, S. Ki Kang, and R.D. Lazarov. Robust multiscale iterative solvers for nonlinear flows in highly heterogeneous media. *Numer. Math. Theory Methods Appl.*, 5(3):359–383, 2012.

[30] Y. Efendiev, J. Galvis, R. Lazarov, and J. Willems. Robust domain decomposition preconditioners for abstract symmetric positive definite bilinear forms. *ESAIM Math. Model. Numer. Anali.*, 46(5):1175–1199, 2012.

[31] Y. Efendiev, J. Galvis, G. Li, and M. Presho. Generalized multiscale finite element methods: Oversampling strategies. *International Journal for Multiscale Computational Engineering*, 12(6), 2014.

[32] Y. Efendiev, J. Galvis, and P.S. Vassilevski. Spectral element agglomerate algebraic multigrid methods for elliptic problems with high-contrast coefficients. In *Domain decomposition methods in science and engineering XIX*, volume 78 of *Lect. Notes Comput. Sci. Eng.*, pages 407–414. Springer, Heidelberg, 2011.

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

[34] Yalchin Efendiev, Sai-Mang Pun, and Petr N. Vabishchevich. Temporal splitting algorithms for non-stationary multiscale problems. *Journal of Computational Physics*, 439:110375, 2021.

[35] L.C. Evans. *Partial Differential Equations*. Graduate studies in mathematics. American Mathematical Society, 2010.

[36] J. Galvis and Y. Efendiev. Domain decomposition preconditioners for multiscale flows in high contrast media. *SIAM J. Multiscale Modeling and Simulation*, 8:1461–1483, 2010.

[37] J. Galvis and Y. Efendiev. Domain decomposition preconditioners for multiscale flows in high contrast media. reduced dimension coarse spaces. *SIAM J. Multiscale Modeling and Simulation*, 8:1621–1644, 2010.

[38] R. Glowinski. *Numerical Methods for Nonlinear Variational Problems*. Computational Physics Series. Springer-Verlag, 1984.

[39] N. Higham. *Functions of matrix theory and computation*. SIAM, University of Manchester, United Kingdom, 2008.

[40] Marlis Hochbruck, Christian Lubich, and Hubert Selhofer. Exponential integrators for large systems of differential equations. *SIAM Journal on Scientific Computing*, 19(5):1552–1574, 1998.

[41] Marlis Hochbruck and Alexander Ostermann. Exponential integrators. *Acta Numerica*, 19:209–286, 2010.

[42] T. Hou and X.H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134:169–189, 1997.

[43] B. Wu J. Huang, L. Ju. A fast compact exponential time differencing method for semilinear parabolic equations with neumann boundary conditions. *Applied Mathematics Letters*, (94):257–265, 2019.

[44] D. Pardo J. Muñoz and L. Demkowicz. Equivalence between the dpg method and the exponential integrator for linear parabolic problems. *Journal of Computational Physics*, 2020.

[45] J. Galvis J. Olmos and F. Martinez. A geometric mean algorithm of symmetric positive definite matrices. *unpublished*.

[46] Lijian Jiang, Yalchin Efendiev, and Victor Ginting. Multiscale methods for parabolic equations with continuum spatial scales. *Discrete and Continuous Dynamical Systems - B*, 8(4):833–859, 2007.

[47] Claes Johnson. Numerical solution of partial differential equations by the finite element method. *Acta Applicandae Mathematica*, 18:184–186, 1988.

[48] E. Abreu C. Diaz J. Muñoz-Matute J. Galvis L. F. Contreras, D. Pardo. An exponential integration generalized multiscale finite element method for parabolic problems. *Submitted*.

[49] S. Martin and C. Vázquez. Homogenization of the layer-structured dam problem with isotropic permeability. *Nonlinear Anali. Real World Appl.*, 14(6):2133–2151, 2013.

[50] Axel Målqvist and Anna Persson. Multiscale techniques for parabolic equations. *Numerische Mathematik*, 138, 01 2018.

[51] M Park and Michael V Tretyakov. Stochastic resin transfer molding process. *SIAM/ASA Journal on Uncertainty Quantification*, 5(1):1110–1135, 2017.

[52] Michael Presho and Juan Galvis. A mass conservative generalized multiscale finite element method applied to two-phase flow in heterogeneous porous media. *Journal of Computational and Applied Mathematics*, 296:376–388, 2016.

[53] R. Tyrrell Rockafellar. On the maximal monotonicity of subdifferential mappings. *Pacific Journal of Mathematics*, 33:209–216, 1970.

[54] Zheng Sun, José A. Carrillo, and Chi-Wang Shu. A discontinuous galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials. *Journal of Computational Physics*, 352:76–104, 2018.

[55] R. Toja. *Contributions of the numerical simulation of coupled models in glaciology.* PhD thesis, Universidade da Coruña, 2010.

[56] H. Thomas Y. Efendiev, J. Galvis. Generalized multiscale finite element methods (gmsfem). *Journal of Computational Physics*, 251:116–135, 2013.

[57] J. Galvis Y. Efendiev and X. Wu. Multiscale finite element methods for high-contrast problems using local spectral basis functions. *Journal of Computational Physics*, 230:937–955, 2011.

[58] Miguel Zambrano, Sintya Serrano, Boyan S Lazarov, and Juan Galvis. Fast multiscale contrast independent preconditioners for linear elastic topology optimization problems. *Journal of Computational and Applied Mathematics*, 389:113366, 2021.