UNIVERSITY OF
GOTHENBURG

Finite Element Methods for PDEs with Algebraic Constraints

Master's thesis in Applied Mathematics

ERIC LINDSTRÖM

Department of Mathematical Sciences
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2021

Master's thesis 2021

Finite Element Methods for PDEs with Algebraic Constraints

ERIC LINDSTRÖM

UNIVERSITY OF
GOTHENBURG

Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Computational Mathematics
University of Gothenburg
Gothenburg, Sweden 2021

Finite Element Methods for PDEs with Algebraic Constraints

ERIC LINDSTRÖM

Supervisor: Axel Målqvist, Department of Mathematical Sciences
Ass. Supervisor: Roland Maier, Department of Mathematical Sciences
Examiner: Klas Modin, Department of Mathematical Sciences

Master's Thesis 2021
Department of Mathematical Sciences
Division of Applied Mathematics and Statistics
Computational Mathematics
University of Gothenburg
SE-412 96 Gothenburg

Finite Element Methods for PDEs with Algebraic Constraints
ERIC LINDSTRÖM
Department of Mathematical Sciences
University of Gothenburg

Abstract

The purpose of this report is to analyze numerical approximations of partial differential equations with algebraic constraints. In particular we consider problems where the algebraic constraint forces the solution to be in the kernel of an interpolation operator. Such constrained PDEs arise for example in numerical solutions of local problems in multiscale methods. We will consider elliptic and parabolic PDEs and present some analytical results as well and numerical simulations. The main points of interest will be to investigate if the said problems are well posed, and to study the decay of the corresponding solutions.

Acknowledgements

First of all I would like to thank Axel Målqvist for being incredibly present and helpful while guiding me through this interesting topic. Also, Roland Maier deserves a lot of gratitude for delivering most of the code for this report and being a great asset when it came to questions surrounding it. Lastly, I would like to thank my friends and family as well for supporting me in this long journey, even long before this thesis.

Eric Lindström, Gothenburg, September 2021

Contents

List of Figures

1
Introduction

The desire to simulate more and more complicated problems calls for sophisticated modeling and efficient numerical methods. A particular case in the context of differential equations are *operator differential algebraic equations* (operator DAEs), or *partial differential algebraic equations* (PDAEs) in the PDE case. In these problems the corresponding solutions are constrained by some operator. These models show up in a variety of different subjects, for example in models of fluid dynamics [23], multibody dynamics [11], electrical circuits [4], reactive transportation [9] and gas transportation networks [1], just to name a few.

However, in this report we will restrict ourselves to one very specific type of PDAE, where we only consider an interpolation operator as the constraining operator. More specifically, we are looking for solutions to PDEs in the kernel of a chosen interpolation operator. This type of PDAEs intuitively brings the focus to scale, and multiscale methods (see [2, 14, 16] for early works and examples of such methods) like the *localized orthogonal decomposition* (LOD) [20, 21]. Multiscale methods targets partial differential equations with highly irregular coefficients, where there is a scale of interest which is smaller than what would be efficient to implement with the classical FEM. Situations where such methods can be beneficial occurs for example in multifunctional materials, where microstuctural properties of the material can result in macroscopic properties like negative optical refraction [10].

Solving general PDEs algebraically is often an insurmountable task, and therefore, we have to use efficient numerical methods to approximate the solution. However, if we want to approximate solutions numerically we have to lay the theoretical groundwork first. This would include proving existence of solutions, as well as investigating approximation errors and stability estimates. The PDAEs we study in this report introduce some complexity into these questions. We will focus on one example of an elliptic PDE, the Poisson equation, and one parabolic example, the heat equation.

In chapter 2 we prepare for the more complicated matters by establishing useful definitions and rehearsing the classical FEM. It also includes a closer look on some interpolation operators.

Thereafter, we approach the questions above both theoretically (chapter 3 and chapter 4) and numerically (chapter 5). In the elliptic case, we acknowledge two different ways of formulating our constrained problem, *the saddle point formulation* and *the constrained variational formula*. We discuss how they have different benefits, and then show that

they are equivalent, as well as well-posed. After that we show that the solution of said problem decays exponentially when the problems involves a localized load vector with an a priori decay estimate. We also investigate the approximation error when one replaces the continuous setting with a discrete setting.

For the parabolic case (chapter 4), we assume a discrete setting immediately, partially for sake of simplicity, but also from realizing the necessity of a discrete settings in implementations. Again we state the two formulations of the problem, similar to the elliptic case, and prove that the problem has an unique solution. The a priori decay estimates of these solutions are open problems, but we prove one temporal a posteriori error estimate, and discuss how temporal and exponential spatial decay probably is inherited from the corresponding unconstrained problem.

Finally, we study numerical solutions of the problems in chapter 5, which are implemented using MATLAB. We do so by presenting solutions with different diffusion coefficients, and to confirm decay we compare them to solutions restricted to smaller domains. The different diffusion coefficients help to bring out certain graphical aspects of the solution which are considered as plausible results of the acting constraints.

2
Background

This first section will consist of definitions as well as some basic theory, that will serve us later in the report and should make the reading more fluent.

2.1 Function Spaces

When dealing with PDEs, one often comes across Sobolev spaces. In this report we will denote them as W_p^k (or $W_p^k(\Omega)$ more precisely) which are defined by

$$W_p^k(\Omega) := \{f \in L^p(\Omega) : f^{(\alpha)} \in L^p(\Omega), \forall |\alpha| \leq k\}. \tag{2.1}$$

Here $\alpha = (\alpha_1, \ldots, \alpha_n)$ is a multi-index, and $f^{(\alpha)}$ is the mixed partial derivative

$$f^{(\alpha)} = \frac{\partial^{|\alpha|} f}{\partial x_1^{\alpha_1} \ldots \partial x_n^{\alpha_n}}. \tag{2.2}$$

Notice that the partial derivatives only have to exist in the weak sense from these definitions. Also, the cases where $p = 2$ are particularly important, since they form Hilbert spaces, and thus have their own notation. We let $H^k(\Omega) := W_2^k(\Omega)$ (or H^k for short).

A frequently used notation in this context is

$$H_0^1(\Omega) := \{v \in H^1(\Omega) \, : \, (\Gamma v)(x) = 0, \quad \forall x \in \partial\Omega\}, \tag{2.3}$$

where Γ is the trace operator onto H^1. In fact, we will let

$$V := H_0^1(\Omega) \tag{2.4}$$

for convenience. We will also often make use of the standard H^1-norm

$$\|\bullet\|_V := \|\bullet\|_{H^1(\Omega)} = \left(\|\bullet\|_{L^2(\Omega)}^2 + \|\nabla \bullet\|_{L^2(\Omega)}^2\right)^{1/2}. \tag{2.5}$$

Sometimes we will leave out the domain in the notation for norms when the chance of misunderstandings is small. The L^2-norm is generated by the standard L^2-inner product which is denoted by $(\bullet, \bullet)$, not to be confused with the standard duality pairing between V and $V' := H^{-1}(\Omega)$, denoted by $\langle \bullet, \bullet \rangle$. Note that the duality pairing reduces to the L^2-inner product if both functions are in L^2.

In sections of the report where we consider parabolic settings we will come across Bochner spaces like $L^2(0,T;\mathcal{B})$ and $H^1(0,T;\mathcal{B})$, where $\mathcal{B}$ is some Banach space. Within these spaces we make use of the standard Bochner-norms

$$\|v\|_{L^2(0,T;\mathcal{B})} = \left(\int_0^T \|v\|_{\mathcal{B}}^2 \, \mathrm{d}t\right)^{1/2}, \tag{2.6}$$

$$\|v\|_{H^1(0,T;\mathcal{B})} = \left(\int_0^T \|v\|_{\mathcal{B}}^2 + \|\dot{v}\|_{\mathcal{B}}^2 \, \mathrm{d}t\right)^{1/2}. \tag{2.7}$$

When referencing Bochner spaces we will also often abbreviate by leaving out the domains. We will also make use of some other norms in our proofs with discrete settings,

$$\|v\|_{\text{tr}}^2 := \int_0^T \|\nabla \bar{v}\|_{L^2}^2 + \|\dot{v}\|_{H^{-1}}^2 \, \mathrm{d}t, \tag{2.8}$$

$$\|v\|_{\text{te}}^2 := \int_0^T \|\nabla v\|_{L^2}^2 \, \mathrm{d}t. \tag{2.9}$$

Here, $\bar{v}$ refers to the temporal mean of v with respect to the fine temporal mesh, see definition in (4.2).

Now we will define the meshes used in this report, which is one key part of FEMs. This is achieved with a collection of closed simplices, $\mathcal{K}_H$. We will define $H_K := \text{diam}(K)$ for all $K \in \mathcal{K}_H$, and let $H := \max_{K\in\mathcal{K}_H} H_K$. Now, the size of elements are restricted by

$$\min_{K\in\mathcal{K}_H} H_K \geq \rho H,$$

where $\rho > 0$ is a constant. The shape of the elements are restricted and characterized by

$$\gamma := \max_{K\in\mathcal{K}_H} \frac{H_K}{\text{diam}(B_K)} > 1.$$

Above, B_K is defined to be the largest ball that fits in $K \in \mathcal{K}_H$. These restrictions are important for error analysis later in the report. We will denote the internal nodes of the mesh with $\mathcal{N}_H$ and the number of internal nodes with N_H. Finally, we will denote a fine mesh with $\mathcal{K}_h$ with the corresponding restrictions but with $h < H$, and the natural notations $\mathcal{N}_h$ and N_h for the nodes and number of internal nodes, respectively.

With our domain Ω discretized we want to simplify the solution space where we search for a solution. This is done by restricting the functions using the mesh. Let us define the restricted function space V_H as

$$V_H := \mathcal{P}^1(\mathcal{K}_H) \cap V,$$

where $\mathcal{P}^1(\mathcal{K}_H)$ is the space of piecewise polynomials of degree ≤ 1 on each $K \in \mathcal{K}_H$, or in other words, functions that are piecewise affine with respect to the mesh. Notice that this space is finite-dimensional, with dimension directly tied to the number of internal

nodes in $\mathcal{N}_H$. To span this space, we will make use of a basis of hat functions $\Lambda_{x_i} \in V_H$, corresponding to internal nodes $x_i \in \mathcal{N}_H$ such that

$$\Lambda_{x_i}(x_j) = \begin{cases} 1, & \text{if } x_i = x_j \\ 0, & \text{otherwise} \end{cases}$$

for all $x_j \in \mathcal{N}_H$. As with the meshes, we will define a fine-scale function space V_h as well, defined in the same manner as V_H but using $\mathcal{K}_h$.

The meshes above suffice to construct our elliptic FEM formulation, and we will now present the definitions corresponding to the parabolic case. The spatial discretization is achieved through a mesh defined as above, $\mathcal{K}_H$. To discretize in the time domain, we simply divide the interval $[0, T]$ into equal parts of some length $\mathcal{T} > 0$. We will let $T_i = i\mathcal{T}$, $i = 0, \ldots, N_\mathcal{T}$, where we assume $N_\mathcal{T}\mathcal{T} = T$. We will also denote the collections of sub-intervals $[T_{i-1}, T_i]$, $i = 1, \ldots, N_\mathcal{T}$, as $\mathcal{J}_\mathcal{T}$. From this discretization we construct two different function spaces,

$$\hat{V}_\mathcal{T} := \{v \in H_0^1([0,T]) \,:\, v \in \mathcal{P}^1(\mathcal{J}_\mathcal{T}) \text{ and } v(0) = 0\}, \tag{2.10a}$$
$$V_\mathcal{T} := \mathcal{P}^0(\mathcal{J}_\mathcal{T}). \tag{2.10b}$$

$\hat{V}_\mathcal{T}$ is spanned by a nodal basis $\{\zeta_i\}_{i=1}^{N_\mathcal{T}}$, where

$$\zeta_i = \left(\frac{t}{\mathcal{T}} - i + 1\right) \mathbb{1}_{[T_{i-1},T_i]} + \left(i + 1 - \frac{t}{\mathcal{T}}\right) \mathbb{1}_{[T_i,T_{i+1}]}, \tag{2.11}$$

with a slight modification to $\zeta_{N_\mathcal{T}}$, where we drop the second term. $V_\mathcal{T}$ is spanned by piecewise constant basis functions $\{\chi_i\}_{i=1}^{N_\mathcal{T}}$ where $\chi_i := \mathbb{1}_{[T_{i-1},T_i]}$. These two function spaces will in turn generate our trial and test spaces given as

$$\hat{V}_{H,\mathcal{T}} := V_H \times \hat{V}_\mathcal{T}, \tag{2.12a}$$
$$V_{H,\mathcal{T}} := V_H \times V_\mathcal{T}. \tag{2.12b}$$

Note that the basis of $\hat{V}_{H,\mathcal{T}}$ and $V_{H,\mathcal{T}}$ are naturally chosen as $\{\zeta_i\Lambda_{x_j}\}_{i=1,\ldots,N_\mathcal{T},x_j\in\mathcal{N}_H}$ and $\{\chi_i\Lambda_{x_j}\}_{i=1,\ldots,N_\mathcal{T},x_j\in\mathcal{N}_H}$ respectively. Note that we may reference variations of the concepts above only varying by index, where H is replaced by $h, 0 < h < H$, and $\mathcal{T}$ by $\tau, 0 < \tau < \mathcal{T}$.

Lastly, we will introduce an important concept to navigate in the mesh, namely element patches. Let $S \in \mathcal{K}_H$ be a collection of elements in $\mathcal{K}_H$, then the element patch $N(S)$ is defined as

$$N(S) := \bigcup \{K \in \mathcal{K}_H : K \cap \overline{S} \neq \emptyset\}.$$

This definition can be extended to any set $S \subset \Omega$, and we are sometimes interested in a patch around a single point z, where we let $N(z) := N(\{z\})$. Also, for $l \geq 2$, we will use the notation $N^l(S) = N(N^{l-1}(S))$, to address the effects of layers in the mesh. Naturally we let $N^1(S) := N(S)$.

2.2 Interpolation

Interpolation is important in the analysis of finite element methods but also in the construction of multiscale methods. In this section we will discuss three common interpolation operators, specifically their benefits, restrictions and error bounds.

2.2.1 Nodal Interpolation

The first operator is also the simplest of the three (the order will be in increasing complexity). Here we will denote the nodal interpolation operator with $\mathcal{I}_{\mathcal{N}}$, and it is defined for all $v \in C^1(\Omega)$ as

$$\mathcal{I}_{\mathcal{N}} v(x) = \sum_{x_i \in \mathcal{N}_H} v(x_i) \Lambda_{x_i}(x).$$

For the nodal interpolation operator we have the error bound

$$\|v - \mathcal{I}_{\mathcal{N}} v\|_{L^2(K)} + h_K \|v - \mathcal{I}_{\mathcal{N}} v\|_{H^1(K)} \leq C_{\mathcal{I}_{\mathcal{N}}} h_K^2 \|v\|_{H^2(K)},$$

where $K \in \mathcal{K}_H$, and $C_{\mathcal{I}_{\mathcal{N}}}$ is a constant.

Note that even if this evaluation is powerful, the restrictions are too invasive for most cases. The functions in H^1 are generally not continuous when we move to higher dimensions, more specifically $H^1(\Omega) \not\subseteq C^0(\Omega)$ if $\Omega \subset \mathbb{R}^d$, $d \geq 2$. This means the operator is not well-defined for all $v \in H^1(\Omega)$ in these cases. Our next candidate will fix this issue.

2.2.2 Clément Interpolation

To define the Clément interpolation operator $\mathcal{I}_{\mathcal{C}} v$, we need the concept of local patches as defined in section 2.1. Using these patches we define $\mathcal{I}_{\mathcal{C}}$ as

$$\mathcal{I}_{\mathcal{C}} v(x) = \sum_{x_i \in \mathcal{N}_H} \frac{\int_{N(x_i)} v(y) dy}{\int_{N(x_i)} dy} \Lambda_{x_i}(x).$$

Note that this operator is defined for all $v \in L^1(\Omega)$. The corresponding error bounds for the operator reads

$$\|v - \mathcal{I}_{\mathcal{C}} v\|_{H^1(K)} \leq C_{\mathcal{I}_{\mathcal{C}}} h_K^s \|v\|_{H^{1+s}(K)}, \forall K \in \mathcal{K}_h, \forall v \in H^{1+s}(\Omega), \tag{2.13}$$

where $s = (0, 1]$.

One disadvantage compared to the nodal interpolation is that the Clément interpolation operator is not a projection. Why this is a disadvantage may not be clear in this moment for the reader but it simplifies future calculations.

Another down-side of the Clément interpolation operator is that it will not conserve homogeneous boundary conditions. However, this can be compensated for by explicitly setting $\mathcal{I}_{\mathcal{C}}$ to zero in the boundary nodes. Even after this modiftication (2.13) still holds [8].

2.2.3 Projective Quasi-Interpolant

In coming applications and code we will use a constructed interpolation operator, $\mathcal{I}_H := E_H \circ \Pi_H^{dg}$. Here, Π_H^{dg} refers to the L^2-projection onto piecewise affine functions, acting on each element of $\mathcal{K}_H$, and E_H is the notation for a weighting operator. It is calculated by

$$E_H v(z) = \sum_{K \in \mathcal{K}_H \,:\, z \in K} \frac{|K|}{|N(z)|} v|_K(z), \quad \forall z \in \mathcal{N}_H.$$

For convenience we interpret $v|_K$ as v on K and 0 elsewhere, to make it defined on the entire set Ω.

This particular operator is constructed to be a projection, and enjoys the following error bound:

$$H_K^{-2} \left\| v - \mathcal{I}_H v \right\|_{L^2(K)}^2 + \left\| \nabla(\mathcal{I}_H v) \right\|_{L^2(K)}^2 \leq \tilde{C}_{\mathcal{I}_H}^2 \left\| \nabla v \right\|_{L^2(N(K))}^2, \quad \forall v \in H^1, \tag{2.14}$$

where $\tilde{C}_{\mathcal{I}_H}$ is a constant depending on the mesh constant γ and the maximum amount of neighbours an element $K \in \mathcal{K}_H$ can have. With a slight change of constants and by summation, one can manipulate (2.14) into a global bound,

$$H^{-2} \left\| v - \mathcal{I}_H v \right\|_{L^2(\Omega)}^2 + \left\| \nabla(\mathcal{I}_H v) \right\|_{L^2(\Omega)}^2 \leq C_{\mathcal{I}_H}^2 \left\| \nabla v \right\|_{L^2(\Omega)}^2, \quad \forall v \in H^1. \tag{2.15}$$

One issue with this interpolation operator is that it does not conserve boundary conditions. However, this is easily fixed for homogeneous boundary conditions, by defining E_H to be 0 on the boundary vertices. This definition also preserves (2.14) for functions $v \in V$ (see [13] for more details).

2.3 A Brief Review of the Finite Element Method

We will use the FEM to solve a PDAE. We start by recapitulating how the FEM is derived for PDEs. This will grant some references to lean back on when we introduce finite element methods for constrained problems, as they for instance arise in connection with multiscale methods.

2.3.1 Elliptic

Now we will set up our elliptic model problem. Our domain $\Omega \subset \mathbb{R}^d$ is assumed to have polyhedral shape with $\operatorname{diam}(\Omega) \approx 1$. We will restrict ourselves to an elliptic equation in $\mathbb{R}^2$,

$$-\operatorname{div}(A \nabla u) = f. \tag{2.16}$$

Here $A : \Omega \to \mathbb{R}^{2\times 2}$ is allowed to be a highly irregular coefficient, but bounded by

$$\begin{aligned} 0 < \alpha := \operatorname*{ess\,inf}_{x\in\Omega} \inf_{v\in\mathbb{R}\setminus\{0\}} \frac{(A(x)v)\cdot v}{v\cdot v} \leq \beta := \operatorname*{ess\,sup}_{x\in\Omega} \sup_{v\in\mathbb{R}\setminus\{0\}} \frac{(A(x)v)\cdot v}{v\cdot v} < \infty, \\ \bar{\beta} := \|A\|_\infty < \infty. \end{aligned} \tag{2.17}$$

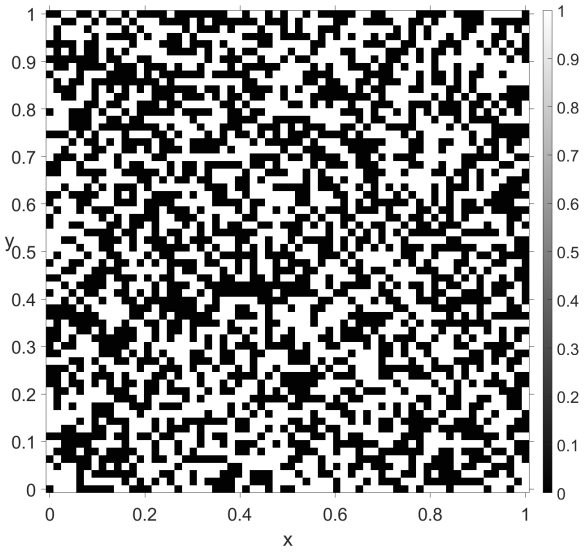


Figure 2.1: *Example of highly irregular diffusion coefficient A visualized, valued 0.01 (black) or 1 (white) in Ω.*

We also assume that A is symmetric in this report, as well as $f \in H^{-1}(\Omega)$.

The first step toward a FEM is to convert (2.16) into its variational formulation

$$a(u, v) := \int_\Omega (A\nabla u) \cdot \nabla v \, \mathrm{d}x = \int_\Omega fv \, \mathrm{d}x, \quad \forall v \in V. \tag{2.18}$$

Since the left-hand side is a bilinear, bounded form, and the right-hand side can be thought of as a bounded linear functional, *the Lax-Milligram Theorem* [17, Theorem A.3 on p. 230] tells us that (2.18) has a unique solution.

Using the solution space defined in section 2.1 and the variational formulation, the classical Galerkin FEM reads: find $u_H \in V_H$ such that

$$\int_\Omega (A\nabla u_H) \cdot \nabla v_H \, \mathrm{d}x = \int_\Omega f v_H \, \mathrm{d}x, \quad \forall v_H \in V_H. \tag{2.19}$$

Since V_H can be spanned by simple hat-functions, we can represent u_H as

$$u_H(x) = \sum_{i=1}^{N_H} U_H^{(i)} \Lambda_i(x),$$

where $U_H^{(i)}$ are scalars. Thus, (2.19) can be reduced to

$$\sum_{i=1}^{N_H} U_H^{(i)} \int_\Omega (A\nabla \Lambda_i) \cdot \Lambda_j \, \mathrm{d}x = \int_\Omega f\Lambda_j \, \mathrm{d}x,$$

for all $j = 1, ..., N^H$. The equations above can be expressed in matrix form as

$$\boldsymbol{AU} = \boldsymbol{f}, \tag{2.20}$$

where $(\boldsymbol{A})_{ij} = a(\Lambda_i, \Lambda_j)$ is called the *stiffness matrix*, $(\boldsymbol{f})_i = \int_\Omega f\Lambda_i \,\mathrm{d}x$ is called *load vector* and $(\boldsymbol{U})_i = U_H^{(i)}$. When using this basis, the stiffness matrix turns out sparse because of the support of the hat functions, which makes this an economic method to implement.

However, one issue with the classical FEM is its poor performance when applied to multiscale problems. Without a fine enough mesh, irregular behaviour on a micro scale can lead to macroscopic error. Occasions like these are examples when our constrained problems arise as part of an alternative numerical approach.

2.3.2 Parabolic

In the parabolic case, we are still considering the same spatial domain Ω, but now with an added time domain, $[0, T]$, with $T > 0$. The problem we are considering is the parabolic equation

$$\dot{u} - \nabla \cdot (A\nabla u) = f \qquad \text{in } \Omega \times (0, T], \tag{2.21a}$$
$$u = 0 \qquad \text{on } \partial\Omega \times (0, T], \tag{2.21b}$$
$$u(0) = 0 \qquad \text{in } \Omega. \tag{2.21c}$$

Once again, $A : [0, T] \times \Omega \to \mathbb{R}^{2\times 2}$ is allowed to be highly irregular, as long as it serves the constraints

$$\begin{aligned} 0 < \alpha :=& \operatorname*{ess\,inf}_{(t,x)\in(0,T)\times\Omega} \inf_{v\in\mathbb{R}^2\setminus\{0\}} \frac{(A(t,x)v)\cdot v}{v\cdot v} \\ \leq& \operatorname*{ess\,sup}_{(t,x)\in(0,T)\times\Omega} \sup_{v\in\mathbb{R}^2\setminus\{0\}} \frac{(A(t,x)v)\cdot v}{v\cdot v} =: \beta < \infty \\ \bar{\beta} :=& \|A\|_\infty < \infty. \end{aligned} \tag{2.22}$$

Additionally, we will also assume that A can be well-approximated by functions piecewise constant on $\mathcal{J}_\tau$. This is to avoid technicalities which this report will not address. This might feel invasive, but note that this assumption is applied on a fine grid, making it less of a restriction. However, it is important to choose τ small enough so that $\mathcal{J}_\tau$ resolves small oscillations in A.

Before we express our variational formulation, we introduce the bilinear form similar to the A-induced one in the elliptic case

$$a(t; v, w) := \int_\Omega (A(t,x)\nabla v)\cdot\nabla w \,\mathrm{d}x. \tag{2.23}$$

Using this notation we can now write the variational formulation for the parabolic problem as finding $u \in V_{\text{tr}} := L^2(V) \cap H^1(H^{-1})$ with $u(0) = 0$ such that

$$\int_0^T \langle \dot{u}, v\rangle + a(t; u, v)\,\mathrm{d}t = \int_0^T \langle f, v\rangle \,\mathrm{d}t \tag{2.24}$$

for all $v \in V_{\text{te}} := L^2(V)$. We assume $f \in L^2(H^{-1})$.

Using (2.24) we will formulate our parabolic FEM problem. We want to find $u_{h,\tau} \in \hat{V}_{h,\tau}$ such that

$$\int_0^T \langle \dot{u}_{h,\tau}, v_{h,\tau} \rangle + a(t; u_{h,\tau}, v_{h,\tau}) \, \mathrm{d}t = \int_0^T \langle f, v_{h,\tau} \rangle \, \mathrm{d}t, \tag{2.25}$$

for all $v_{h,\tau} \in V_{h,\tau}$. Now, to exploit our earlier mentioned discretization, we will express our trial function $u_{h,\tau}$ as

$$u_{h,\tau} := \sum_{i=1}^{N_\tau} U_{h,\tau}^{(i)} \zeta_i, \tag{2.26}$$

where $U_{h,\tau}^{(i)} \in V_h$, $i = 1, \ldots, N_\tau$. We also only consider test functions of the form $v_{h,\tau} = v_h \chi_i$, where $v_h \in V_h$ and $i = 0, \ldots, N_\tau$. By doing this, we can solve (2.25) by each time-step consecutively. Equivalently, we want to find functions $U_{h,\tau}^{(i)} \in V_h$ such that

$$\begin{aligned} &\int_{t_{i-1}}^{t_i} \langle U_{h,\tau}^{(i-1)} \dot{\zeta}_{i-1}, v_h \rangle + \langle U_{h,\tau}^{(i)} \dot{\zeta}_i, v_h \rangle + a(t; U_{h,\tau}^{(i-1)} \zeta_{i-1}, v_h) + a(t; U_{h,\tau}^{(i)} \zeta_i, v_h) \, \mathrm{d}t \\ &= \int_{t_{i-1}}^{t_i} \langle f, v_h \rangle \, \mathrm{d}t \end{aligned} \tag{2.27}$$

for all $v_h \in V_h$ and $i = 1, \ldots, N_\tau$. This is exactly the well-known *Crank-Nicholson scheme* which reads: find $U_{h,\tau}^{(i)} \in V_h$ such that

$$\langle U_{h,\tau}^{(i)}, v_h \rangle + \frac{\tau}{2} a(t_{i-\frac{1}{2}}; U_{h,\tau}^{(i)}, v_h) = \int_{t_{i-1}}^{t_i} \langle f, v_h \rangle \, \mathrm{d}t + \langle U_{h,\tau}^{(i-1)}, v_h \rangle - \frac{\tau}{2} a(t_{i-\frac{1}{2}}; U_{h,\tau}^{(i-1)}, v_h), \tag{2.28}$$

for all $v_h \in V_h$ and $i = 1, \ldots, N_\tau$.

3

Elliptic PDE with Coarse Scale Constraints

This report revolves around the realization that sometimes it might be interesting to solve PDE's in constrained subspaces. Loosely speaking we talk about cases where the solution of the PDE is constrained by some operator, commonly referred to as operator DAEs (or PDAEs in the PDE case). These problems arise naturally in many fields of mathematics and help to increase sophistication of mathematical modeling. Some examples of areas where these problems arise are fluid dynamics [23], multibody dynamics [11] and gas network simulations [1], as previously mentioned. The theory surrounding operator DAEs is vast, and we will only consider a specific case.

In this report we will restrict ourselves to solutions living in the kernel of a chosen interpolation operator from subsection 2.2.3. This chapter will be dedicated to formulate these problems in elliptic cases, and present some results about the properties of the corresponding solutions. Since PDAEs introduce some complexity to otherwise classical PDEs, we will discuss the existence and uniqueness of these solutions, as well as some approximation theory. Problems as these are tightly connected to multiscale methods, and applications can be found in the LOD [20, 21] for example.

3.1 Constrained Elliptic Problem

To begin analyzing our constrained elliptic problems, it seems only natural to begin with the arguably simplest second order PDE, the elliptic problem we formulated in section 2.3. Such a problem is often given as a saddle point problem. In our setting of constraining a PDE on a coarse-scale level, the problem reads: find $\varphi \in V$ and $\lambda_\varphi \in V_H$ such that

$$a(\varphi, w) + (\lambda_\varphi, \mathcal{I}_H w) = (f, w), \qquad \forall w \in V \tag{3.1a}$$

$$(\mathcal{I}_H \varphi, \mu_H) = 0, \qquad \forall \mu_H \in V_H, \tag{3.1b}$$

where $f \in V'$. Note that this is a very specific case of an operator DAE, and if one is interested in reading about a more general setting, one can do so in [12]. To see that there exists a unique solution (3.1), we will use [5, Corollary 4.2.1 on p. 229].

Theorem 1. *Assume that* (2.17) *holds. Then there exists unique solutions* $\varphi \in V$ *and* $\lambda_\varphi \in V_H$ *that solves* (3.1)*. The solution* φ *is bounded by*

$$\|\varphi\|_V \leq \frac{1}{\alpha} \|f\|_{V'} \tag{3.2}$$

Proof. [5, Corollary 4.2.1] applies to our case if we can show that $a(\bullet, \bullet)$ is bounded and coercive, and that the operator $(\mathcal{I}_H \bullet, \bullet)$ fulfills the inf sup-condition

$$\inf_{v_H \in V_H} \sup_{v \in V} \frac{(\mathcal{I}_H v, v_H)}{\|v\|_V \|v_H\|_V} \geq \delta, \tag{3.3}$$

for some $\delta > 0$. We will begin by showing that $a(\bullet, \bullet)$ is bounded, which is true because of (2.17),

$$a(v, v) \leq \beta \|\nabla v\|^2_{L^2(\Omega)} \leq \beta \|v\|^2_V . \tag{3.4}$$

Similarly, $a(\bullet, \bullet)$ is coercive by the following reasoning:

$$a(v, v) \geq \alpha \|\nabla v\|^2_{L^2(\Omega)} \geq \frac{\alpha}{2} \left(\|\nabla v\|^2_{L^2(\Omega)} + C_{\mathcal{P}}^{-2} \|v\|^2_{L^2(\Omega)} \right) \geq \frac{\alpha}{2} \min \left\{ 1, C_{\mathcal{P}}^{-2} \right\} \|v\|^2_V , \tag{3.5}$$

where we used (2.17) in the first inequality, and *the Poincaré inequality* [17, Theorem A.6 on p. 238] in the second.

Lastly, to show that (3.3) holds, one can use $v = v_H$, since $V_H \subset V$

$$\begin{aligned} \sup_{v \in V} \frac{(\mathcal{I}_H v, v_H)}{\|v\|_V \|v_H\|_V} &\geq \frac{(\mathcal{I}_H v_H, v_H)}{\|v_H\|^2_V} = \frac{\|v_H\|^2_{L^2(\Omega)}}{\|v_H\|^2_V} \geq \frac{\|v_H\|^2_{L^2(\Omega)}}{\|v_H\|^2_{L^2(\Omega)} (1 + C^2_{inv} H^{-2})} \\ &= \frac{1}{(1 + C^2_{inv} H^{-2})}, \end{aligned}$$

where we made use of *the inverse inequality* [7, Theorem 4.5.11 on p. 112-113],

$$\|\nabla v_H\|_{L^2(\Omega)} \leq C_{inv} H^{-1} \|v_H\|_{L^2(\Omega)} \quad \forall v_H \in V_H. \tag{3.6}$$

Note that we might make use of and reference the inverse inequality applied to the fine function space V_h as well, where H is replaced by h. Since the reasoning above applies to any $v_H \in V_H$, we have that

$$\inf_{v_H \in V_H} \sup_{v \in V} \frac{(\mathcal{I}_H v, v_H)}{\|v\|_V \|v_H\|_V} \geq \frac{1}{(1 + C^2_{inv} H^{-2})}.$$

The existence and uniqueness of the solutions to (3.1) as well as (3.2) now follows from [5, Corollary 4.2.1].

□

Another way to phrase our constrained problem is to find a solution $\varphi \in W := \ker(\mathcal{I}_H)$ to

$$a(\varphi, w) = \int_\Omega f w \, \mathrm{d}x, \quad \forall w \in W. \tag{3.7}$$

The same bounds are acting on A as in (2.17), and we still assume $f \in V'$, and thus, if one realizes that W is a Hilbert space, the Lax-Milgram Theorem suffices to show that (3.7) has a unique solution. We will state this in a short, informal proof.

Theorem 2. *Assume* (2.17) *holds. Then* (3.7) *has a unique solution.*

Proof. $a(\bullet, \bullet)$ is bounded and coercive by the same reasoning as in (3.4) and (3.5). $(f, \bullet)$ is bounded by applying the *Hölder inequality*,

$$|(f, w)| \leq \|f\|_{H^{-1}} \|w\|_{H^1}, \tag{3.8}$$

for all $w \in W$.

Finally, before applying the Lax-Milgram Theorem, we need W to be a Hilbert space. One comes to the conclusion that W is a Hilbert space by noting that $\ker(\mathcal{I}_H)$ is a kernel of a continuous linear operator, and thus a closed subspace, which makes W a closed subspace of a Hilbert space, making W a Hilbert space aswell. Thereby we can apply the Lax-Milgram Theorem on (3.7) and ensure the existence of an unique solution.

□

Using the existence and uniqueness arguments for solutions to (3.1) and to (3.7) respectively, we will prove in Theorem 3 that the two problems are equivalent. The theorem is a slight modification of a proof in [19, Theorem 2.3.2 p. 18], and mostly varies from it by notation.

Theorem 3. *Assume that A is bounded as in* (2.17)*, and let $\varphi \in W$ be the solution to* (3.7)*. Additionally, let $\lambda_\varphi \in V_H$ be the solution to the problem*

$$(\lambda_\varphi, w_H) = (f, w_H) - a(\varphi, w_H) \tag{3.9}$$

for all $w_H \in V_H$. Then $(\varphi, \lambda_\varphi) \in W \times V_H$ solves (3.1)*, and λ_φ is the associated Lagrange multiplier. The statements are thus equivalent.*

Proof. First of all, the solution to (3.9) exists and is unique by the Lax-Milgram Theorem since the L^2-inner product is bounded and coercive on V_H. We can use Hölder's inequality to show it is bounded, since for all $v_H, w_H \in V_H$ we have

$$|(v_H, w_H)| \leq \|v_H\|_{L^2} \|w_H\|_{L^2} \leq \|v_H\|_{H^1} \|w_H\|_{H^1}.$$

That $(\bullet, \bullet)$ is coercive follows from an application of (3.6),

$$\begin{aligned}(v_H, v_H) = \|v_H\|_{L^2}^2 &\geq \frac{1}{2}(\|v_H\|_{L^2}^2 + C_{inv}^{-2} H^2 \|\nabla v_H\|_{L^2}^2) \\ &\geq \frac{1}{2} \min\{1, C_{inv}^{-2} H^2\} \|v_H\|_{H^1}^2.\end{aligned}$$

Now, let $w \in V$. To arrive at (3.1a), we use (3.9) to realize that

$$\begin{aligned}a(\varphi, w) &= a(\varphi, \mathcal{I}_H w) + a(\varphi, (1 - \mathcal{I}_H)w) + (f, w) - (f, \mathcal{I}_H w) - (f, (1 - \mathcal{I}_H)w) \\ &= -(\lambda_\varphi, \mathcal{I}_H w) + (f, w) + a(\varphi, (1 - \mathcal{I}_H)w) - (f, (1 - \mathcal{I}_H)w).\end{aligned}$$

Our next step is to note that $(1 - \mathcal{I}_H)w \in W$ and φ solves (3.7). Thus

$$a(\varphi, (1 - \mathcal{I}_H)w) - (f, (1 - \mathcal{I}_H)w) = 0.$$

where we used 3.9 and that $\mathcal{I}_H w \in V_H$. To sum it up, we have

$$a(\varphi, w) = -(\lambda_\varphi, \mathcal{I}_H w) + (f, w).$$

The above reasoning holds for all $w \in V$, and thus we have shown (3.1a). To show (3.1b), one just has to keep in mind that $\varphi \in W$ and thus

$$\mathcal{I}_H \varphi = 0$$

making it trivial. By combining the results in this proof and the uniqueness arguments in Theorem 1 and Theorem 2, we have that the two problem formulations (3.1) and (3.7) are equivalent.

□

The saddle point comes with a great advantage. Unlike in (3.7) we do not need an explicit description of the space W, and to solve that problem directly would include finding a local basis of W, which is not trivial. This is not an issue with (3.1).

3.1.1 Exponential Decay

In Theorem 4, we will prove an interesting property of φ, namely that it decays exponentially if the right-hand side of (3.7) has local support. Thus, in the following theorem we will assume that:

$$\begin{cases} f \in H^{-1}, \\ \exists K \in \mathcal{K}_H \,:\, \operatorname{supp}(f) \subseteq K. \end{cases} \tag{3.10}$$

In applications like the LOD, this result can be useful to justify calculating localized solutions to coarse scale constrained problems. Our theorem follows [20, Theorem 4.1 p. 38], but considers a slightly more general right-hand side for coherence sake.

Theorem 4. *Assume that f is restricted by* (3.10) *and let $\varphi \in W$ be the solution of* (3.7). *Then φ is bounded by the following estimate:*

$$\left\| A^{1/2} \nabla \varphi \right\|_{L^2(\Omega \setminus N^l(K))} \lesssim \exp\left(-c_{(3.11)} \frac{\alpha}{\overline{\beta}} l \right) \|f\|_{H^{-1}(\Omega)}, \tag{3.11}$$

where $l \in \mathbb{N}_0$ and $c_{(3.11)} \geq (8C_{\mathcal{I}_H}(1 + C_{\mathcal{I}_H}\gamma + \gamma))^{-1} > 0$ is a constant which depends on the shape regularity constants γ and ρ. The notation $\lesssim$ should be interpreted as $a \lesssim b$ if there exists $k > 0$ independent of a and b such that $a \leq k \cdot b$.

Proof. Let $l \geq 4$. We will introduce a function $\eta \in V_H$ by the following definition:

$$\begin{cases} \eta \equiv 0 \text{ in } N^{l-3}(K), \\ \eta \equiv 1 \text{ in } \Omega \setminus N^{l-2}(K). \end{cases}$$

Because of this definition and the properties of our mesh, we get some useful statement immediately. Namely that

$$\begin{aligned} \operatorname{supp}(\eta) &= \Omega \setminus N^{l-3}(K), \\ \operatorname{supp}(\nabla \eta) &= \Omega \setminus ((\Omega \setminus N^{l-2}(K)) \cup N^{l-3}(K)) \\ &= N^{l-2}(K) \setminus N^{l-3}(K) := R \end{aligned}$$

and $\|\nabla\eta\|_{L^\infty(K)} \leq \frac{\gamma}{H_K}$. One consequence of this (and the positivity of A) is that

$$\begin{aligned}\left\|A^{1/2}\nabla\varphi\right\|_{L^2(\Omega\setminus N^l(K))} &= \int_{L^2(\Omega\setminus N^l(K))} (A\nabla\varphi)\cdot\nabla\varphi\,\mathrm{d}x \\ &\leq \int_{\Omega\setminus N^{l-2}(K)} (A\nabla\varphi)\cdot\nabla\varphi\,\mathrm{d}x \\ &\leq \int_\Omega \eta(A\nabla\varphi)\cdot\nabla\varphi\,\mathrm{d}x. \end{aligned} \tag{3.12}$$

Now this right-hand side can be rewritten using the product rule as

$$\begin{aligned}\int_\Omega \eta(A\nabla\varphi)\cdot\nabla\varphi\,\mathrm{d}x &= \int_\Omega (A\nabla\varphi)\cdot\nabla(\eta\varphi)\,\mathrm{d}x - \int_\Omega \nabla\eta(A\nabla\varphi)\varphi\,\mathrm{d}x \\ &= a(\varphi,\eta\varphi) - \int_R \nabla\eta(A\nabla\varphi)\varphi\,\mathrm{d}x.\end{aligned}$$

This expression is now further expanded with the addition and subtraction of an interpolated term,

$$\int_\Omega \eta(A\nabla\varphi)\cdot\nabla\varphi\,\mathrm{d}x = a(\varphi,(1-\mathcal{I}_H)(\eta\varphi)) + a(\varphi,\mathcal{I}_H(\eta\varphi)) - \int_R \nabla\eta(A\nabla\varphi)\varphi\,\mathrm{d}x. \tag{3.13}$$

The first term in (3.13) can be reduced to 0. To see why, note that $(1-\mathcal{I}_H)(\eta\varphi)\in W$. This is easy to see using the definition of W combined with the fact that our interpolation operator $\mathcal{I}_H$ is a projection. Since φ is the solution to (3.7), we then see that

$$a(\varphi,(1-\mathcal{I}_H)(\eta\varphi)) = \int_\Omega f(1-\mathcal{I}_H)(\eta\varphi)\,\mathrm{d}x = 0,$$

because of the support of $\eta\varphi$ and f.

To find an upper bound on the second term of (3.13), we start by investigating the function $\mathcal{I}_H(\eta\varphi)$. Since $\varphi\in W$ and by the definition of η, one can arrive at $\mathrm{supp}(\mathcal{I}_H(\eta\varphi))\subseteq N(R)$ (note that $\mathcal{I}_H$ might "spread" the support by one layer of elements). Now, using the definition of η, the product rule and the error bound for our interpolation operator (2.14) one can derive

$$\begin{aligned}a(\varphi,\mathcal{I}_H(\eta\varphi)) &= \int_{N(R)} (A\nabla\varphi)\cdot\nabla(\mathcal{I}_H(\eta\varphi))\,\mathrm{d}x \\ &\leq \bar\beta\,\|\nabla\varphi\|_{L^2(N(R))}\,\|\nabla(\mathcal{I}_H(\eta\varphi))\|_{L^2(N(R))} \\ &\leq \bar\beta\,\|\nabla\varphi\|_{L^2(N^{l-1}(K)\setminus N^{l-4}(K))}\,C_{\mathcal{I}_H}\,\|\nabla(\eta\varphi)\|_{L^2(N(N^{l-1}(K)\setminus N^{l-4}(K)))} \\ &\leq C_{\mathcal{I}_H}\bar\beta\,\|\nabla\varphi\|_{L^2(N^{l-1}(K)\setminus N^{l-4}(K))}\,(\|\nabla\varphi\|_{L^2(N^l(K)\setminus N^{l-3}(K))} \\ &+ C_{\mathcal{I}_H}\gamma\,\|\nabla\varphi\|_{L^2(N^{l-1}(K)\setminus N^{l-4}(K))}) \\ &\leq C_{\mathcal{I}_H}\bar\beta(1+C_{\mathcal{I}_H}\gamma)\,\|\nabla\varphi\|^2_{L^2(N^l(K)\setminus N^{l-4}(K))} \\ &\leq C_{\mathcal{I}_H}\frac{\bar\beta}{\alpha}(1+C_{\mathcal{I}_H}\gamma)\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(N^l(K)\setminus N^{l-4}(K))}.\end{aligned}$$

Lastly, we can find an upper bound of the third term in (3.13) by similar strategies,

$$\begin{aligned}
|\int_R \nabla\eta(A\nabla\varphi)\varphi\,\mathrm{d}x| &\leq \|\nabla\eta\|_{L^\infty(R)} \|(A\nabla\varphi)\|_{L^2(R)} \|\varphi\|_{L^2(R)} \\
&\leq \bar{\beta}\gamma \sum_{K\subset R} H_K^{-1} \|\nabla\varphi\|_{L^2(K)} \|\varphi - \mathcal{I}_H\varphi\|_{L^2(K)} \\
&\leq \tilde{C}_{\mathcal{I}_H}\bar{\beta}\gamma \sum_{K\subset R} \|\nabla\varphi\|_{L^2(K)} \|\nabla\varphi\|_{L^2(N(K))} \\
&\leq C_{\mathcal{I}_H}\bar{\beta}\gamma \|\nabla\varphi\|^2_{L^2(N^l(K)\setminus N^{l-4}(K))} \\
&\leq C_{\mathcal{I}_H}\frac{\bar{\beta}\gamma}{\alpha} \left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(N^l(K)\setminus N^{l-4}(K))}.
\end{aligned}$$

We can now use these bounds in (3.12) to get

$$\begin{aligned}
\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(\Omega\setminus N^l(K))} &\leq \tilde{C}\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(N^l(K)\setminus N^{l-4}(K))} \\
&= \tilde{C}(\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(\Omega\setminus N^{l-4}(K))} - \left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(\Omega\setminus N^l(K))}).
\end{aligned}$$

where $\tilde{C} := C_{\mathcal{I}_H}\frac{\bar{\beta}}{\alpha}(1 + C_{\mathcal{I}_H}\gamma + \gamma)$. Therefore, it follows

$$\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(\Omega\setminus N^l(K))} \leq \frac{\tilde{C}}{\tilde{C}+1}\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(\Omega\setminus N^{l-4}(K))}$$

One can also see by the relation of the domains of the norms that $\tilde{C} \geq 1$. The inequality above can be used inductively to arrive at

$$\begin{aligned}
\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(\Omega\setminus N^l(K))} &\leq \delta^{\lfloor l/4\rfloor}\left\|A^{1/2}\nabla\varphi\right\|^2_{L^2(\Omega)} = \delta^{\lfloor l/4\rfloor}a(\varphi,\varphi) = \delta^{\lfloor l/4\rfloor}\int_K f\varphi\,\mathrm{d}x \\
&\leq \delta^{\lfloor l/4\rfloor}\|f\|_{H^{-1}(\Omega)}\|\varphi\|_{H^1(K)} \\
&\leq \delta^{\lfloor l/4\rfloor}\|f\|_{H^{-1}(\Omega)}(1 + C_{\mathcal{P}})\|\nabla\varphi\|_{L^2(K)} \\
&\leq \frac{1+C_{\mathcal{P}}}{\alpha}\delta^{\lfloor l/4\rfloor}\|f\|_{H^{-1}(\Omega)}\left\|A^{1/2}\nabla\varphi\right\|_{L^2(K)} \\
&\leq \frac{1+C_{\mathcal{P}}}{\alpha}\delta^{\lfloor l/4\rfloor}\|f\|_{H^{-1}(\Omega)}\left\|A^{1/2}\nabla\varphi\right\|_{L^2(\Omega)},
\end{aligned}$$

where $\delta := \frac{\tilde{C}}{\tilde{C}+1}$ for simplicity, and where we made use of the Poincaré inequality in the third row. Thus,

$$\left\|A^{1/2}\nabla\varphi\right\|_{L^2(\Omega\setminus N^l(K))} \lesssim \delta^{\lfloor l/4\rfloor}\|f\|_{H^{-1}(\Omega)}. \tag{3.14}$$

Notice that since $\delta < 1$, this estimate holds for $l < 4$ as well (one can also attribute this to the stability estimate (3.2)). Also, we have that

$$\begin{aligned}
\delta^{\lfloor l/4\rfloor} &\leq \delta^{-1}\left(\delta^{\frac{1}{4}}\right)^l \\
&\leq 2\left(\delta^{\frac{1}{4}}\right)^l \\
&= 2\exp\left(-\frac{1}{4}\log\left(\delta^{-1}\right)l\right) \\
&\leq 2\exp\left(-\frac{1}{4}\left(\frac{1}{2\tilde{C}}\right)l\right)
\end{aligned}$$

where we used a rough estimate of $\log\left(\frac{\tilde{C}+1}{\tilde{C}}\right)$ motivated by its McLaurin-expansion in the last step. This new factor applied in (3.14) finishes the proof.

□

Note that this proof is not only useful for localized right-hand sides, but one can easily derive conclusions for the global case by realizing that $\cup_{K\in\mathcal{K}_H} K = \Omega$. In applications, as mentioned earlier, this can allow for localized estimations of φ where one might truncate solutions of the form:

$$a(\varphi_K, w) = \int_\Omega f|_K w \,\mathrm{d}x, \quad \forall w \in W, \tag{3.15}$$

where $f|_K$ is interpreted as being equal to some f (with potentially global support) on K, and 0 elsewhere. Note that $\varphi = \sum_{K\in\mathcal{K}_H} \varphi_K$.

3.1.2 Approximation Error

The last focus of this chapter will be addressing the issue that in numerical applications we are restricted to finite-dimensional spaces, which means we have to approximate continuous spaces, using fine meshes. In our case, that would mean that V_h takes the place of V in calculations, where h is small enough to resolve small oscillations in A. Our next result will show that it is a reasonable approach, and that the approximation error which ensues from this approach is well-behaved. Our theorem will be an application of [5, Theorem 5.2.5 on p. 278].

Theorem 5. *Assume that A is bounded as in* (2.17) *and that $(\varphi, \lambda_\varphi) \in V \times V_H$ solves* (3.1). *Further assume that $(\varphi_h, \lambda_\varphi) \in V_h \times V_H$ solves the following discretized problem:*

$$a(\varphi_h, w_h) + (\lambda_\varphi, \mathcal{I}_H w_h) = (f, w_h), \qquad \forall w_h \in V_h, \tag{3.16a}$$
$$(\mathcal{I}_H \varphi_h, \mu_H) = 0, \qquad \forall \mu_H \in V_H. \tag{3.16b}$$

Then the following bound on the approximation error holds:

$$\|\varphi - \varphi_h\|_V \leq C_{(3.17)} \inf_{w_h\in V_h} \|\varphi - w_h\|_V \,, \tag{3.17}$$

where $C_{(3.17)}$ is a constant independent of h and H.

Proof. The theorem directly follows from [5, Theorem 5.2.5 on p. 278] and the results in (3.5) and (3.3), concerning the coercivity of $a(\bullet,\bullet)$ and the inf-sup condition on $(\mathcal{I}_H\bullet,\bullet)$. Note that $\ker(\mathcal{I}_H|_{V_h}) \subseteq \ker(\mathcal{I}_H|_V)$ obviously holds true in our case, which is why we can enjoy the improved error estimate (3.17).

□

The infimum on the right-hand side in (3.17) might look abstract, but one should keep in mind that this especially holds for interpolation estimates of φ. Thus, one could concretize it further by restricting it with interpolation errors. For example, we could use the error bound (2.13) for the Clemént operator to get the reference error

$$\|\varphi - \varphi_h\|_V \lesssim \inf_{w_h\in V_h} \|\varphi - w_h\|_V \lesssim \|\varphi - \mathcal{I}_C\varphi\|_V \lesssim h^s \|\varphi\|_{H^{1+s}} \,. \tag{3.18}$$

Note that we have to assume $v \in H^{1+s}(\Omega)$, $s \in (0, 1]$, to end up with some factor of h, which is important since it ensures that we can get a better approximation of φ if we use a finer mesh.

4

Parabolic PDE with Coarse Scale Constraints

This chapter will be structurally similar to chapter 3, but in the parabolic setting derived from (2.21). Structurally we will discuss the same results concerning existence and uniqueness of solutions, but we are more limited in proofs concerning the behavior of the solutions, since some problems like a priori error estimates for spatial decay are still open. The chapter follows [18] to a large extent, only distilled to relevant matters and with some minor modifications.

4.1 Constrained Parabolic Problem

Again, we will introduce our coarse-scale constrained parabolic problem as a saddle point problem. However, in this chapter we will consider a fully discrete setting, partially for sake of simplicity but also because as we saw in chapter 3 we eventually have to consider a discrete setting for implementations. Thus the saddle point problem reads: find $\psi_{h,\tau} \in \hat{V}_{h,\tau}$ and the associated Lagrange-multipliers $(\lambda_\psi^{(0)}, \dots, \lambda_\psi^{(N_\mathcal{T})}) \in V_H \times \cdots \times V_H$ such that

$$\int_0^T \langle \dot{\psi}_{h,\tau}, w_{h,\tau} \rangle + a(t;\, \psi_{h,\tau}, w_{h,\tau})\, \mathrm{d}t + \sum_{i=1}^{N_\mathcal{T}} \int_{T_{i-1}}^{T_i} \langle \lambda_\psi^{(i)}, \mathcal{I}_H w_{h,\tau} \rangle\, \mathrm{d}t = \int_0^T \langle f, w_{h,\tau} \rangle\, \mathrm{d}t, \tag{4.1a}$$

$$\sum_{j=0}^{N_\mathcal{T}} \langle \mathcal{I}_H \psi_{h,\tau}(T_j), \mu_H^{(j)} \rangle = 0, \tag{4.1b}$$

for all $w_{h,\tau} \in V_{h,\tau}$, all $\mu_H^{(j)} \in V_H$, $j = 0, \dots, N_\mathcal{T}$ and where $f \in L^2(H^{-1})$.

We will show that (4.1) has a unique solution in a coming theorem, and the reader may recognize the approach from the previous chapter. Before we go into the theorem, we will state a useful notation which will be used numerous times throughout the proof,

$$\bar{v} = \frac{1}{\tau} \sum_{i=1}^{\mathcal{N}_\tau} \int_0^T v(t;\bullet)\chi_i\, \mathrm{d}t. \tag{4.2}$$

Put into words, this function is the temporal mean with respect to our fine temporal mesh.

The coming proof is the combination of [18, Lemma 2.3 on p. 5] and [18, Lemma 3.3 on p. 10] (inspired by [22]), only with a slight modification to the right-hand side bound.

Theorem 6. *Assume that* (2.22) *holds. Then* (4.1) *has a unique solution* $(\psi_{h,\tau}, \mu_H^{(0)}, \ldots, \mu_H^{(N_\mathcal{T})}) \in \hat{V}_{h,\tau} \times V_H \times \cdots \times V_H$*. It is also bounded by the expression*

$$\|\psi_{h,\tau}\|_{tr} \leq c_{\mathfrak{A}}^{-1} \|f\|_{L^2(H^{-1})}, \tag{4.3}$$

where $c_{\mathfrak{A}} = \min\{\alpha, C_{proj}^{-1}\alpha^{1/2}\beta^{-1/2}\}$.

Proof. We will make use of [3, Theorem 2.1 and Corollary 2.1 on p. 1239-1240], but to do so we have to show inf-sup conditions for the following operators:

$$\begin{aligned}
\mathfrak{A} &: \hat{V}_{h,\tau} \times V_{h,\tau} \longrightarrow \mathbb{R}, \\
\mathfrak{A}(v_{h,\tau}, w_{h,\tau}) &:= \int_0^T \langle \dot{v}_{h,\tau}, w_{h,\tau} \rangle + a(t;\, v_{h,\tau}, w_{h,\tau})\, \mathrm{d}t \qquad (4.4) \\
\mathfrak{B}_i &: V_H \times V_{h,\tau} \longrightarrow \mathbb{R}, \\
\mathfrak{B}_i(\lambda_H, w_{h,\tau}) &:= \int_{T_{i-1}}^{T_i} \langle \lambda_H, \mathcal{I}_H w_{h,\tau} \rangle\, \mathrm{d}t, \qquad (4.5) \\
\mathfrak{C}_j &: \hat{V}_{h,\tau} \times V_H \longrightarrow \mathbb{R}, \\
\mathfrak{C}_j(v_{h,\tau}, \mu_H) &:= \langle \mathcal{I}_H v_{h,\tau}(T_j), \mu_H \rangle, \qquad (4.6)
\end{aligned}$$

for $1 \leq i \leq N_\mathcal{T}$ and $0 \leq j \leq N_\mathcal{T}$. This combined with the bounds of the right-hand sides of (4.1) will grant us the results in Theorem 6.

inf-sup condition for $\mathfrak{A}$: To prove our condition we will make use of another operator, $\mathcal{E}_h : V_{h,\tau} \longrightarrow V_{h,\tau}$, This operator is characterized by the equation

$$a(t;\, \mathcal{E}_h v(t), w_h) = \langle v(t), w_h \rangle \tag{4.7}$$

for all $w_h \in V_h$ and $t \in [0, T]$. From this we can derive that if $v_{h,\tau} \in \hat{V}_{h,\tau}$, $v_{h,\tau} \neq 0$, then $\bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau} \in V_{h,\tau}$, where $\bar{v}_{h,\tau}$ is the temporal mean from (4.2). Further we have that

$$\begin{aligned}
\mathfrak{A}(v_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau}) &= \int_0^T \langle \dot{v}_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau} \rangle + a(t;\, v_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau})\, \mathrm{d}t \\
&= \int_0^T a(t;\, \mathcal{E}_h \dot{v}_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau}) + a(t;\, \bar{v}_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau})\, \mathrm{d}t \\
&= \int_0^T a(t;\, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau})\, \mathrm{d}t, \qquad (4.8)
\end{aligned}$$

where we used (4.7) and that $a(t;\, v_{h,\tau}, w_{h,\tau}) = a(t;\, \bar{v}_{h,\tau}, w_{h,\tau})$ for any $w_{h,\tau} \in V_{h,\tau}$. We will

keep expressing (4.8) as

$$
\begin{aligned}
&\mathfrak{A}(v_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau}) \\
&= \int_0^T a(t;\, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h \dot{v}_{h,\tau})\, \mathrm{d}t \\
&= \int_0^T a(t;\, \bar{v}_{h,\tau}, \bar{v}_{h,\tau})\, \mathrm{d}t + 2 \int_0^T a(t;\, \mathcal{E}_h \dot{v}_{h,\tau}, \bar{v}_{h,\tau})\, \mathrm{d}t + \int_0^T a(t;\, \mathcal{E}_h \dot{v}_{h,\tau}, \mathcal{E}_h \dot{v}_{h,\tau})\, \mathrm{d}t \\
&\geq \int_0^T a(t;\, \bar{v}_{h,\tau}, \bar{v}_{h,\tau})\, \mathrm{d}t + 2 \int_0^T \langle \dot{v}_{h,\tau}, \bar{v}_{h,\tau} \rangle\, \mathrm{d}t + \leq C_{\mathit{proj}}^{-2} \beta^{-1} \int_0^T \|\dot{v}_{h,\tau}\|_{H^{-1}}^2\, \mathrm{d}t \\
&= \int_0^T a(t;\, \bar{v}_{h,\tau}, \bar{v}_{h,\tau})\, \mathrm{d}t + \|v_{h,\tau}(T)\|_{L^2}^2\, \mathrm{d}t + \leq C_{\mathit{proj}}^{-2} \beta^{-1} \int_0^T \|\dot{v}_{h,\tau}\|_{H^{-1}}^2\, \mathrm{d}t \\
&\geq \int_0^T a(t;\, \bar{v}_{h,\tau}, \bar{v}_{h,\tau}) + \leq C_{\mathit{proj}}^{-2} \beta^{-1} \|\dot{v}_{h,\tau}\|_{H^{-1}}^2\, \mathrm{d}t,
\end{aligned} \tag{4.9}
$$

where we used (4.7) again and $v_{h,\tau} \in \hat{V}_{h,\tau}$ in the fourth row. In the third row we used that $\|v_{h,\tau}\|_{H^{-1}}^2 \leq C_{\mathit{proj}}^2 \beta a(t; \mathcal{E}_h v_{h,\tau}, \mathcal{E}_h v_{h,\tau})$, which comes from the following reasoning:

$$
\begin{aligned}
\|v_{h,\tau}(t)\|_{H^{-1}} &= \sup_{w \in V} \frac{\langle v_{h,\tau}(t), w \rangle}{\|w\|_{H^1}} \\
&\leq \sup_{w \in V} \frac{\langle v_{h,\tau}(t), \Pi_h w \rangle}{\frac{1}{C_{\mathit{proj}}} \|\Pi_h w\|_{H^1}} \\
&= C_{\mathit{proj}} \sup_{w \in V_h} \frac{\langle v_{h,\tau}(t), w_h \rangle}{\|w_h\|_{H^1}} \\
&= C_{\mathit{proj}} \sup_{w \in V_h} \frac{a(t; \mathcal{E}_h v_{h,\tau}, w_h)}{\|w_h\|_{H^1}} \\
&\leq C_{\mathit{proj}} \sup_{w \in V_h} \frac{\|(A^{1/2} \nabla(\mathcal{E}_h v_{h,\tau}))(t)\|_{L^2} \|(A^{1/2} \nabla w_h)(t)\|_{L^2}}{\|w_h\|_{H^1}} \\
&\leq C_{\mathit{proj}} \sup_{w \in V_h} \frac{\beta^{1/2} \|(A^{1/2} \nabla(\mathcal{E}_h v_{h,\tau}))(t)\|_{L^2} \|\nabla w_h\|_{L^2}}{\|w_h\|_{H^1}} \\
&\leq C_{\mathit{proj}} \beta^{1/2} \|(A^{1/2} \nabla(\mathcal{E}_h v_{h,\tau}))(t)\|_{L^2} \\
&= (C_{\mathit{proj}}^2 \beta a(t; \mathcal{E}_h v_{h,\tau}, \mathcal{E}_h v_{h,\tau}))^{1/2},
\end{aligned} \tag{4.10}
$$

where $\Pi_h : V \longrightarrow V_h$ is the L^2-projection onto V_h. In the reasoning above we made use of the stability estimates of Π_h in H^1 (see [6] for more details) in the second row, (4.7) in the fourth, Hölder's inequality in the fifth and (2.17) in the sixth.

Now we can finally make use of (4.9) to begin showing the inf-sup condition for $\mathfrak{A}$. Let

$v_{h,\tau} \in \hat{V}_{h,\tau}$. Then

$$\begin{aligned}
\sup_{w_{h,\tau}\in V_{h,\tau}} \frac{\mathfrak{A}(v_{h,\tau}, w_{h,\tau})}{\|w_{h,\tau}\|_{\text{te}}} &\geq \frac{\mathfrak{A}(v_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau})}{\|\bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau}\|_{\text{te}}} \\
&= \frac{\int_0^T a(t;\ \bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau})\,\mathrm{d}t}{(\int_0^T (\nabla(\bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau}))^2\,\mathrm{d}t)^{1/2}} \\
&= \frac{\int_0^T \left\|A^{1/2}\nabla(\bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau})\right\|_{L^2}^2\,\mathrm{d}t}{(\int_0^T \|\nabla(\bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau})\|_{L^2}^2\,\mathrm{d}t)^{1/2}} \\
&\geq \frac{\alpha^{1/2}\int_0^T \left\|A^{1/2}\nabla(\bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau})\right\|_{L^2}^2\,\mathrm{d}t}{(\int_0^T \|A^{1/2}\nabla(\bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau})\|_{L^2}^2\,\mathrm{d}t)^{1/2}} \\
&\geq \alpha^{1/2}\left(\int_0^T a(t;\ \bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau}, \bar{v}_{h,\tau} + \mathcal{E}_h\dot{v}_{h,\tau})\,\mathrm{d}t\right)^{1/2} \\
&\geq \alpha^{1/2}\left(\int_0^T a(t;\ \bar{v}_{h,\tau}, \bar{v}_{h,\tau}) + C_{\mathit{proj}}^{-2}\beta^{-1}\|\dot{v}_{h,\tau}\|_{H^{-1}}^2\,\mathrm{d}t\right)^{1/2} \\
&\geq \min\{\alpha, C_{\mathit{proj}}^{-1}\alpha^{1/2}\beta^{-1/2}\}\,\|v_{h,\tau}\|_{\text{tr}}\,. \qquad (4.11)
\end{aligned}$$

Now, if we take the infimum of our left-hand side of (4.11) and re-organize the terms, we get the inf-sup condition for $\mathfrak{A}$

$$\inf_{v_{h,\tau}\in\hat{V}_{h,\tau}} \sup_{w_{h,\tau}\in V_{h,\tau}} \frac{\mathfrak{A}(v_{h,\tau}, w_{h,\tau})}{\|v_{h,\tau}\|_{\text{tr}}\|w_{h,\tau}\|_{\text{te}}} \geq c_{\mathfrak{A}}, \qquad (4.12)$$

where

$$c_{\mathfrak{A}} := \min\{\alpha, C_{\mathit{proj}}^{-1}\alpha^{1/2}\beta^{-1/2}\}. \qquad (4.13)$$

inf-sup condition for $\sum_i \mathfrak{B}_i$: Let $(\lambda_H^{(1)}, \dots, \lambda_H^{(N_\mathcal{T})}) \in V_H \times \cdots \times V_H$ be arbitrary and nonzero, and let $\|(\lambda_H^{(1)}, \dots, \lambda_H^{(N_\mathcal{T})})\|_{V_H\times\ldots\times V_H} := \sum_{i=1}^{N_\mathcal{T}} \|\nabla\lambda_H^{(i)}\|_{L^2}$.

Now, we define $\tilde{w}_{h,\tau} := \sum_{i=1}^{N_\mathcal{T}} \lambda_H^{(i)}\mathbb{1}_{[T_{i-1},T_i]} \in V_{h,\tau}$, which will grant us the inequalities

$$\begin{aligned}
\sup_{w_{h,\tau}\in V_{h,\tau}} \frac{\sum_{i=1}^{N_\mathcal{T}} \mathfrak{B}_i(\lambda_H^{(i)}, w_{h,\tau})}{\|(\lambda_H^{(1)}, \dots, \lambda_H^{(N_\mathcal{T})})\|_{V_H\times\ldots\times V_H}\,\|w_{h,\tau}\|_{\text{te}}} &\geq \frac{\sum_{i=1}^{N_\mathcal{T}} \mathfrak{B}_i(\lambda_H^{(i)}, \tilde{w}_{h,\tau})}{\|(\lambda_H^{(1)}, \dots, \lambda_H^{(N_\mathcal{T})})\|_{V_H\times\ldots\times V_H}\,\|\tilde{w}_{h,\tau}\|_{\text{te}}} \\
&= \frac{\sum_{i=1}^{N_\mathcal{T}} \mathcal{T}\|\lambda_H^{(i)}\|_{L^2}^2}{(\sum_{i=1}^{N_\mathcal{T}} \|\lambda_H^{(i)}\|_{L^2}^2)^{1/2}(\sum_{i=1}^{N_\mathcal{T}} \mathcal{T}\|\nabla\lambda_H^{(i)}\|_{L^2}^2)^{1/2}} \\
&= \mathcal{T}^{1/2}\frac{\sum_{i=1}^{N_\mathcal{T}} \|\lambda_H^{(i)}\|_{L^2}^2}{\sum_{i=1}^{N_\mathcal{T}} \|\nabla\lambda_H^{(i)}\|_{L^2}^2}. \qquad (4.14)
\end{aligned}$$

This expression can be further bounded by an application of the (3.6),

$$\mathcal{T}^{1/2}\frac{\sum_{i=1}^{N_\mathcal{T}} \|\lambda_H^{(i)}\|_{L^2}^2}{\sum_{i=1}^{N_\mathcal{T}} \|\nabla\lambda_H^{(i)}\|_{L^2}^2} \geq \mathcal{T}^{1/2}\frac{\sum_{i=1}^{N_\mathcal{T}} \|\lambda_H^{(i)}\|_{L^2}^2}{C_{\text{inv}}^2 H^{-2}\sum_{i=1}^{N_\mathcal{T}} \|\lambda_H^{(i)}\|_{L^2}^2} = C_{\text{inv}}^{-2}H^2\mathcal{T}^{1/2} > 0. \qquad (4.15)$$

If we combine (4.14) and (4.15) while applying the infimum, we have our inf-sup condition for $\sum_i \mathfrak{B}_i$.

inf-sup condition for $\sum_j \mathfrak{C}_j$: For this condition we will construct $\eta_i \in \hat{V}_\tau$ such that

$$\eta_i := \left(\frac{t}{\mathcal{T}} - i + 1\right) \mathbb{1}_{[T_{i-1},T_i]} + \left(i + 1 - \frac{t}{\mathcal{T}}\right) \mathbb{1}_{[T_i,T_{i+1}]}, \tag{4.16}$$

for $i = 1 \dots N_\mathcal{T}$. Now we choose non-zero $(\mu_H^{(1)}, \dots, \mu_H^{(N_\mathcal{T})}) \in V_H \times \cdots \times V_H$ and let $v_{h,\tau} := \sum_{i=1}^{N_\mathcal{T}} \eta_i \mu_H^{(i)} \in \hat{V}_{h,\tau}$. Examining the norm of $v_{h,\tau}$, we find that

$$\begin{aligned}
\|v_{h,\tau}\|_{\text{tr}} &= \left(\int_0^T \|\nabla \bar{v}_{h,\tau}\|_{L^2}^2 + \|\dot{v}_{h,\tau}\|_{H^{-1}}^2 \, \mathrm{d}t\right)^{1/2} \\
&= \left(\int_0^T \Big\| \sum_{i=1}^{N_\mathcal{T}} \bar{\eta}_i \nabla \mu_H^{(i)} \Big\|_{L^2}^2 + \Big\| \sum_{i=1}^{N_\mathcal{T}} \dot{\eta}_i \mu_H^{(i)} \Big\|_{H^{-1}}^2 \mathrm{d}t\right)^{1/2} \\
&\le \left(2\int_0^T \sum_{i=1}^{N_\mathcal{T}} \bar{\eta}_i^2 \|\nabla \mu_H^{(i)}\|_{L^2}^2 + \sum_{i=1}^{N_\mathcal{T}} \dot{\eta}_i^2 \|\mu_H^{(i)}\|_{L^2}^2 \, \mathrm{d}t\right)^{1/2} \\
&= \left(2\sum_{i=1}^{N_\mathcal{T}} \left(\int_0^T \bar{\eta}_i^2 \|\nabla \mu_H^{(i)}\|_{L^2}^2 + \dot{\eta}_i^2 \|\mu_H^{(i)}\|_{L^2}^2 \, \mathrm{d}t\right)\right)^{1/2} \\
&\le \left(2\mathcal{T} \sum_{i=1}^{N_\mathcal{T}} \|\nabla \mu_H^{(i)}\|_{L^2}^2 + \frac{4}{\mathcal{T}} \sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2\right)^{1/2}.
\end{aligned} \tag{4.17}$$

By using this bound on $v_{h,\tau}$ and the seminorm on $V_H \times \cdots \times V_H$ mentioned earlier in this proof, we get

$$\begin{aligned}
&\sup_{w_{h,\tau} \in \hat{V}_{h,\tau}} \frac{\sum_{i=1}^{N_\mathcal{T}} \mathfrak{C}_i(w_{h,\tau}, \mu_H^{(i)})}{\|w_{h,\tau}\|_{\text{tr}} \|(\mu_H^{(1)}, \dots, \mu_H^{(N_\mathcal{T})})\|_{V_H \times \cdots \times V_H}} \\
&\ge \frac{\sum_{i=1}^{N_\mathcal{T}} \langle \mathcal{I}_H v_{h,\tau}(T_j), \mu_H^{(j)} \rangle}{\|v_{h,\tau}\|_{\text{tr}} \|(\mu_H^{(1)}, \dots, \mu_H^{(N_\mathcal{T})})\|_{V_H \times \cdots \times V_H}} \\
&\ge \frac{\sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2}{\left(2\mathcal{T} \sum_{i=1}^{N_\mathcal{T}} \|\nabla \mu_H^{(i)}\|_{L^2}^2 + \frac{4}{\mathcal{T}} \sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2\right)^{1/2} \left(\sum_{i=1}^{N_\mathcal{T}} \|\nabla \mu_H^{(i)}\|_{L^2}^2\right)^{1/2}} \\
&\ge \frac{\sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2}{\left(2\mathcal{T} C_{\text{inv}}^2 H^{-2} \sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2 + \frac{4}{\mathcal{T}} \sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2\right)^{1/2} \left(C_{\text{inv}}^2 H^{-2} \sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2\right)^{1/2}} \\
&= \frac{\sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2}{\left(2\mathcal{T} C_{\text{inv}}^2 H^{-2} + \frac{4}{\mathcal{T}}\right)^{1/2} (C_{\text{inv}}^2 H^{-2})^{1/2} \sum_{i=1}^{N_\mathcal{T}} \|\mu_H^{(i)}\|_{L^2}^2} \\
&\ge \frac{1}{\frac{2\mathcal{T}+1}{2} C_{\text{inv}}^2 H^{-2} + \frac{2}{\mathcal{T}}} \\
&:= \gamma(H, \mathcal{T}) > 0,
\end{aligned} \tag{4.18}$$

where we made use of (3.6) and Young's inequality. If we apply the infimum to (4.18) we have the inf-sup condition for $\sum_i \mathfrak{C}_i$.

Because of our assumption on f we have that our right-hand side in (4.1a) is bounded. This combined with the proven inf-sup conditions grants us the existence and uniqueness of the solution to (4.1), as well as the bound in (4.3) from [3, Theorem 2.1 and Corollary 2.1 on p. 1239-1240]. □

4.1.1 Alternative Formulation in Constrained Spaces

Similar to the elliptic case, (4.1) can be rephrased as a PDE in a constrained subspace. To do so, we define the following subspaces:

$$\hat{W}_{h,\tau} := \{w_{h,\tau} \in \hat{V}_{h,\tau} \,:\, \mathcal{I}_H w_{h,\tau}(T_i) = 0,\, \forall i = 0, \dots, N_{\mathcal{T}}\}, \tag{4.19}$$

$$W_{h,\tau} := \{w_{h,\tau} \in V_{h,\tau} \,:\, \frac{1}{\mathcal{T}} \int_{T_{i-1}}^{T_i} \mathcal{I}_H w_{h,\tau} \,\mathrm{d}t \,\forall i = 1, \dots N_{\mathcal{T}}\}. \tag{4.20}$$

Using these, our constrained problem simply reads: find $\psi_{h,\tau} \in \hat{W}_{h,\tau}$ such that

$$\mathfrak{A}(v_{h,\tau}, w_{h,\tau}) = \int_0^T \langle f, w_{h,\tau} \rangle \,\mathrm{d}t \quad \forall w_{h,\tau} \in W_{h,\tau}. \tag{4.21}$$

We will not prove that (4.21) is well defined, but instead comment on the equivalence to (4.1) which indeed exists. One should acknowledge the parallels between this chapter and chapter 3, namely that there existing different ways to formulate the same problem. However, we will leave out the formal proof in this report and point the interested reader to [18].

4.1.2 Temporal and Spatial Decay

As for results concerning decay in our parabolic case, we restrict ourselves to a proof of an a posteriori upper bound of the temporal localization, Theorem 7, while the a priori estimate of spatial decay remains an open problem. We make use of parts of [18, Lemma 4.1 on p. 15], with the very slight modification of including a localized f in the calculations.

Theorem 7. *Let $\psi_{h,\tau} \in \hat{V}_{h,\tau}$ be the solution to* (4.1)*. Additionally, assume that f is localized in time, such that $f\mathbb{1}_{[T_2,T]} \equiv 0$. Then we have the a posteriori temporal error for $\psi_{h,\tau}$,*

$$\left\|\psi_{h,\tau}\mathbb{1}_{[T_i,T]}\right\|_{tr} \leq (1+c_{\mathfrak{A}}^{-1}) \left(C_{\mathcal{I}_H} H \mathcal{T}^{-1/2} + \max\{\frac{1}{\sqrt{2}}, \frac{\bar{\beta}}{\sqrt{3}}\} \mathcal{T}^{1/2} \right) \left\|\nabla \psi_{h,\tau}(T_i)\right\|_{L^2}, \tag{4.22}$$

for $i \geq 2$.

Proof. To arrive at (4.22), we define a $\phi \in \hat{V}_{h,\tau}$ such that

$$\psi_{h,\tau} = \psi_{h,\tau}\mathbb{1}_{[0,T_i]} + \frac{T_{i+1} - t}{\mathcal{T}} \psi_{h,\tau}(T_i)\mathbb{1}_{[T_i,T_{i+1}]} + \phi_{h,\tau}. \tag{4.23}$$

Note that this implies $\phi_{h,\tau}\mathbb{1}_{[0,T_i]} \equiv 0$ and $\phi_{h,\tau}\mathbb{1}_{[T_{i+1},T]} = \psi_{h,\tau}\mathbb{1}_{[T_{i+1},T]}$. It also means that

$$\begin{aligned}
&\int_{T_i}^{T} \langle \dot{\phi}_{h,\tau}, w_{h,\tau}\rangle + a(t;\phi_{h,\tau}, w_{h,\tau})\,\mathrm{d}t + \sum_{j=i+1}^{N_{\mathcal{T}}} \int_{T_{i-1}}^{T_i} \langle \lambda_{\psi}^{(i)}, \mathcal{I}_H w_{h,\tau}\rangle\,\mathrm{d}t \\
&= \int_{T_i}^{T} \langle \dot{\psi}_{h,\tau}, w_{h,\tau}\rangle + a(t;\psi_{h,\tau}, w_{h,\tau})\,\mathrm{d}t + \sum_{j=i+1}^{N_{\mathcal{T}}} \int_{T_{i-1}}^{T_i} \langle \lambda_{\psi}^{(i)}, \mathcal{I}_H w_{h,\tau}\rangle\,\mathrm{d}t \\
&+ \int_{T_i}^{T_{i+1}} \frac{1}{\mathcal{T}}\langle \psi_{h,\tau}(T_i), w_{h,\tau}\rangle - a(t;\frac{T_{i+1}-t}{\mathcal{T}}\psi_{h,\tau}(T_i), w_{h,\tau})\,\mathrm{d}t \\
&= \int_{T_i}^{T} \langle f, w_{h,\tau}\rangle\,\mathrm{d}t + \int_{T_i}^{T_{i+1}} \frac{1}{\mathcal{T}}\langle \psi_{h,\tau}(T_i), w_{h,\tau}\rangle - a(t;\frac{T_{i+1}-t}{\mathcal{T}}\psi_{h,\tau}(T_i), w_{h,\tau})\,\mathrm{d}t \\
&= \int_{T_i}^{T_{i+1}} \frac{1}{\mathcal{T}}\langle \psi_{h,\tau}(T_i), w_{h,\tau}\rangle - a(t;\frac{T_{i+1}-t}{\mathcal{T}}\psi_{h,\tau}(T_i), w_{h,\tau})\,\mathrm{d}t
\end{aligned} \tag{4.24}$$

and

$$\langle \mathcal{I}_H \phi_{h,\tau}(T_j), \mu_H^{(j)}\rangle = 0, \tag{4.25}$$

for all $w_{h,\tau} \in V_{h,\tau}$, $\mu_H^{(j)} \in V_H$, $j = i, \dots, N_{\mathcal{T}}$. To arrive at an upper bound for $\phi_{h,\tau}$ we will make use of [3, Corollary 2.1] once again, similar to Theorem 6, but to do so we need to show that the right-hand side of (4.24) is bounded. We have that

$$\begin{aligned}
&\int_{T_i}^{T_{i+1}} \frac{1}{\mathcal{T}}\langle \psi_{h,\tau}(T_i), w_{h,\tau}\rangle - a(t;\frac{T_{i+1}-t}{\mathcal{T}}\psi_{h,\tau}(T_i), w_{h,\tau})\,\mathrm{d}t \\
&\leq \left(\frac{1}{\mathcal{T}} \left(\int_{T_i}^{T_{i+1}} \|\psi_{h,\tau}(T_i)\|_{L^2}^2\,\mathrm{d}t \right)^{1/2} + \bar{\beta} \left(\int_{T_i}^{T_{i+1}} \left\| \frac{T_{i+1}-t}{\mathcal{T}} \nabla\psi_{h,\tau}(T_i) \right\|_{L^2}^2 \mathrm{d}t \right)^{1/2} \right) \|w_{h,\tau}\|_{\mathrm{te}} \\
&= \left(\mathcal{T}^{-1/2} \|\psi_{h,\tau}(T_i)\|_{L^2} + \frac{\bar{\beta}}{\sqrt{3}} \mathcal{T}^{1/2} \|\nabla\psi_{h,\tau}(T_i)\|_{L^2} \right) \|w_{h,\tau}\|_{\mathrm{te}} \\
&= \left(\mathcal{T}^{-1/2} \|\psi_{h,\tau}(T_i) - \mathcal{I}_H(\psi_{h,\tau}(T_i))\|_{L^2} + \frac{\bar{\beta}}{\sqrt{3}} \mathcal{T}^{1/2} \|\nabla\psi_{h,\tau}(T_i)\|_{L^2} \right) \|w_{h,\tau}\|_{\mathrm{te}} \\
&\leq \left(C_{\mathcal{I}_H} H \mathcal{T}^{-1/2} + \frac{\bar{\beta}}{\sqrt{3}} \mathcal{T}^{1/2} \right) \|\nabla\psi_{h,\tau}(T_i)\|_{L^2} \|w_{h,\tau}\|_{\mathrm{te}},
\end{aligned} \tag{4.26}$$

where we used the global application of (2.14) for the last inequality. Thus, from the same arguments as in Theorem 6 we can derive:

$$\|\phi_{h,\tau}\|_{\mathrm{tr}} \leq c_{\mathfrak{A}}^{-1} \left(C_{\mathcal{I}_H} H \mathcal{T}^{-1/2} + \frac{\bar{\beta}}{\sqrt{3}} \mathcal{T}^{1/2} \right) \|\nabla\psi_{h,\tau}(T_i)\|_{L^2}. \tag{4.27}$$

Next, we note that:

$$\begin{aligned}
&\left\| \frac{T_{i+1}-t}{\mathcal{T}} \psi_{h,\tau}(T_i) \mathbb{1}_{[T_i,T_{i+1}]} \right\|_{\text{tr}} \\
&= \frac{1}{\mathcal{T}} \left(\int_{T_i}^{T_{i+1}} \|\psi_{h,\tau}(T_i)\|_{L^2}^2 \, \mathrm{d}t \right)^{1/2} + \left(\int_{T_i}^{T_{i+1}} \left\| \frac{\overline{T_{i+1}-t}}{\mathcal{T}} \nabla \psi_{h,\tau}(T_i) \right\|_{L^2}^2 \, \mathrm{d}t \right)^{1/2} \\
&= \mathcal{T}^{-1/2} \|\psi_{h,\tau}(T_i)\|_{L^2} + \frac{1}{\sqrt{2}} \mathcal{T}^{1/2} \|\nabla \psi_{h,\tau}(T_i)\|_{L^2} \\
&\leq \left(C_{\mathcal{I}_H} H \mathcal{T}^{-1/2} + \frac{1}{\sqrt{2}} \mathcal{T}^{1/2} \right) \|\nabla \psi_{h,\tau}(T_i)\|_{L^2} ,
\end{aligned} \tag{4.28}$$

where we once again applied (2.14) in the last step.

Finally we arrive at (4.22) by the following reasoning:

$$\begin{aligned}
\left\| \psi_{h,\tau} \mathbb{1}_{[T_i,T]} \right\|_{\text{tr}} &= \left\| \frac{T_{i+1}-t}{\mathcal{T}} \psi_{h,\tau}(T_i) \mathbb{1}_{[T_i,T_{i+1}]} + \phi_{h,\tau} \right\|_{\text{tr}} \\
&\leq \left\| \frac{T_{i+1}-t}{\mathcal{T}} \psi_{h,\tau}(T_i) \mathbb{1}_{[T_i,T_{i+1}]} \right\|_{\text{tr}} + \|\phi_{h,\tau}\|_{\text{tr}} \\
&\leq (1 + c_{\mathfrak{A}}^{-1}) \left(C_{\mathcal{I}_H} H \mathcal{T}^{-1/2} + \max\{\frac{1}{\sqrt{2}}, \frac{\bar{\beta}}{\sqrt{3}}\} \mathcal{T}^{1/2} \right) \|\nabla \psi_{h,\tau}(T_i)\|_{L^2} .
\end{aligned} \tag{4.29}$$

□

Theorem 7 is useful since we can acquire a bound for $\psi_{h,\tau}$ by just evaluating it at one fixed point in time. It also grants us a valuable tool to investigate the solution, since it may be sufficient to just evaluate the solution a couple of fixed points in time to confirm the nature of the decay. While more results are not formally in print concerning a priori estimates of decay, it is likely that the spatial decay is similar to what we found true for elliptic problems in chapter 3, which we will see in our results in chapter 5.

5
Numerical Experiments

This chapter will be dedicated to presenting the numerical results from implementations in MATLAB. The implementations will consider the two earlier mentioned constrained problems. This way we can observe some of the discussed properties, like decay for example. We will confirm some of the theory that is shown in earlier chapters, but also discuss some other phenomena and investigate some open problems especially in the parabolic case.

5.1 Prerequisites

We will not go into much details concerning the code used in the report, and we point the interested reader to read more about the elliptic code and parabolic code in [20] and [18] respectively. The only purpose of this section is to establish some parameters and constructions to increase comprehension in coming sections.

First of all, note that we have to approximate the continuous setting with a fine-scale one. This was motivated through Theorem 5 in the elliptic case, and in chapter 4 we assume it immediately. We will discuss this further in chapter 6.

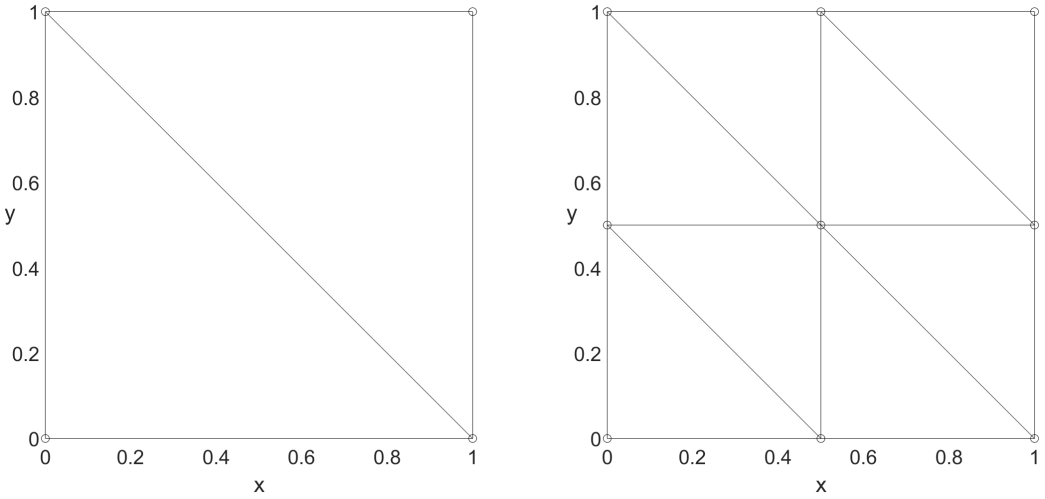


Figure 5.1: *$\mathcal{K}_0$ (to the left) and the resulting mesh after one instance of uniform refinement (to the right).*

The coarse- and fine-scale grids will be constructed in the same way in both the elliptic and the parabolic case. We start of with a simple grid $\mathcal{K}_0$ which is refined uniformly

for some number of instances. This grid and one instance of refinement is illustrated in Figure 5.1. To arrive at our coarse-scale grid $\mathcal{K}_H$ we go through this process four times and for our fine-scale grid seven times, or in other words, $H = 2^{-4}$ and $h = 2^{-7}$. Also, for our construction of f and spatial error analysis we choose a specific element $K \in \mathcal{K}_H$. This element stays the same through our experiment, and has vertices in $(0.6875, 0.5)$, $(0.6875, 0.5625)$ and $(0.75, 0.5)$.

For the temporal discretization, the coarse and fine grid are simply assembled by dividing our interval into subintervals of length 2^{-4} and 2^{-7} respectively.

Finally, we have to choose our diffusion coefficient A. This is achieved through one spatial parameter, N_ε, and one temporal parameter, N_δ. This in practice means that A is divided into N_ε^2 squares over Ω, and these squares stay constant on temporal intervals of length N_δ. Throughout all of the examples we have used $N_\varepsilon = N_\delta := 2^5$. This is motivated by that we want A to be detailed on a scale which is not resolved by $\mathcal{K}_H$, but by $\mathcal{K}_h$.

5.2 Elliptic case

The code we have implemented would be equivalent to solving Equation 3.16, and we will study the behavior of the solution while changing the diffusion coefficient A. We will use a load vector f fulfilling Equation 3.10, simply constructed as $f \equiv 1$ on $K \in \mathcal{K}_H$ with vertices in $(0.6875, 0.5)$, $(0.6875, 0.5625)$ and $(0.75, 0.5)$, and 0 elsewhere. The implemented f is shown in Figure 5.2.

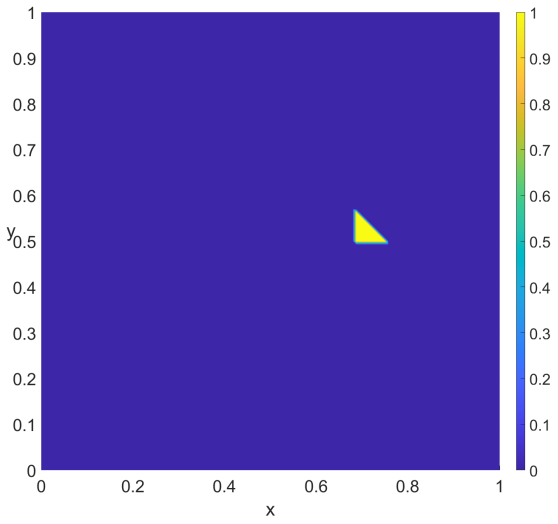


Figure 5.2: *Localized f used in implementations illustrated.*

In our first numerical result we have used a constant diffusion coefficient, $A \equiv 1$. The result is presented in Figure 5.3, and one can see that the decay property we proved in Theorem 4 seems to be confirmed here aswell. One can also see that the underlying mesh vaguely appears, as the edges of the mesh often corresponds to smaller values in the solution. In Figure 5.4 we have constructed an error plot, where we plot the relative error between φ_h and $\varphi_h|_{N^l(K)}$ in L^2-norm and H^1-seminorm. Here we interpret $N^0(K)$ as just being K. The plot makes it easier to interpret the nature of the decay, which appears to be exponential, as would be expected from the theory.

Next, in Figure 5.5 we used a highly irregular A, where the domain Ω is divided into a grid of squares, and every square is randomly assigned a value of 1 or 0.01. The observations concerning exponential decay and the mesh standing out appears once again, although it is less clear. One can also sense a vague resemblance of the diffusion coefficient in the solution. This will be clearer in future examples. We also confirm the exponential decay in Figure 5.9a in both norms, however, the plot looks less smooth and more randomized than in previous example.

To make the mentioned phenomenon where the solution imitates the diffusion coefficient clearer, we have implemented a diffusion coefficient with a thin channel running through it in Figure 5.6, valued 0.01 in the channel and 1 everywhere else. Now one can clearly see the resemblance of the diffusion coefficient. One can also note that the decay of

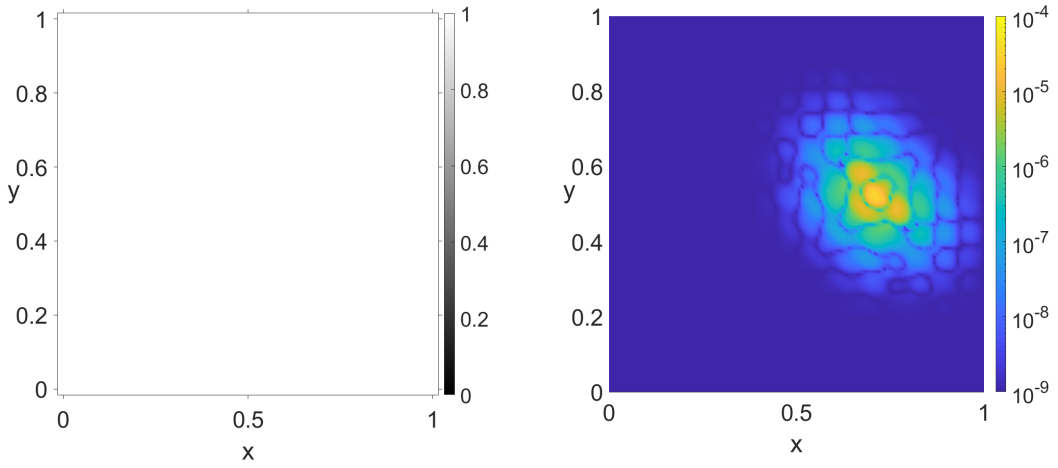


Figure 5.3: *Constant diffusion coefficient to the left, corresponding solution φ_h to the right.*

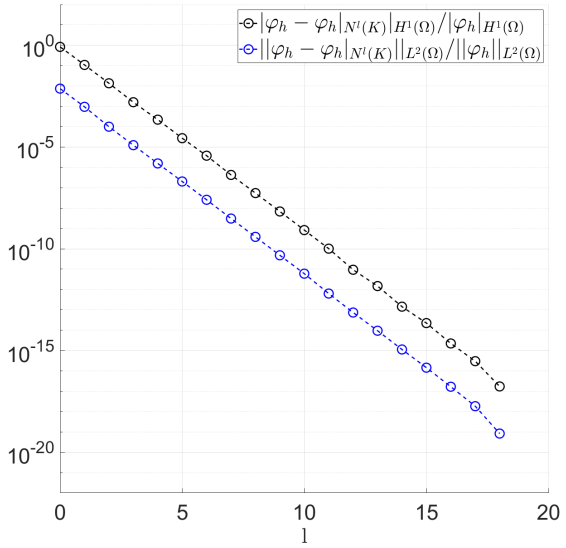


Figure 5.4: *Relative error plot visualizing spatial decay with constant diffusion coefficient.*

the solution is significantly slower within channel. In Figure 5.9b we also observe a quick decay, which also aligns with what we see in Figure 5.6 (see [15]).

Further stressing the point made in the previous paragraph, we implemented an "inverted" version of the thin channel diffusion coefficient. By inverted here we mean that the coefficient is valued 1 within the channel, and 0.01 everywhere else. The result is presented in Figure 5.7, and this time we note a faster decay of the solution within the channel. Note that the error plot in Figure 5.9c is very similar to our previous example, only slightly slower.

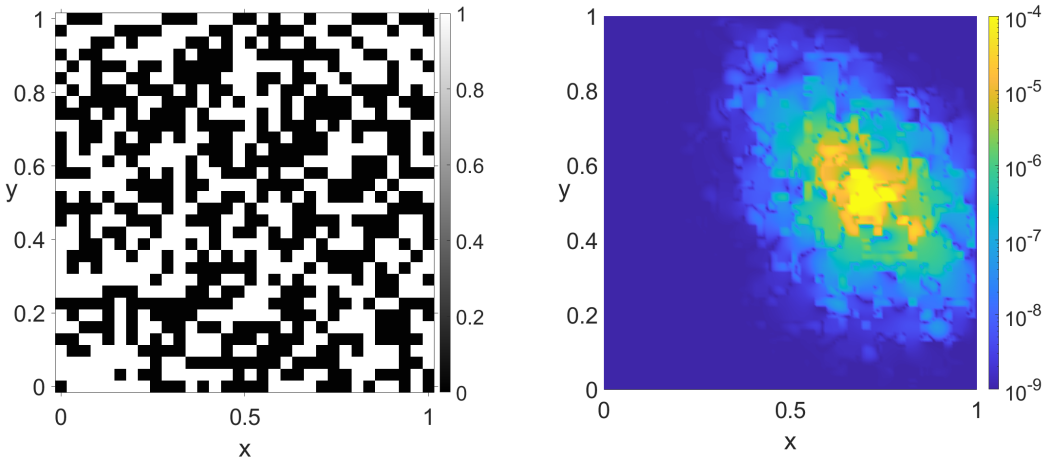


Figure 5.5: *Randomized diffusion coefficient to the left, corresponding solution φ_h to the right.*

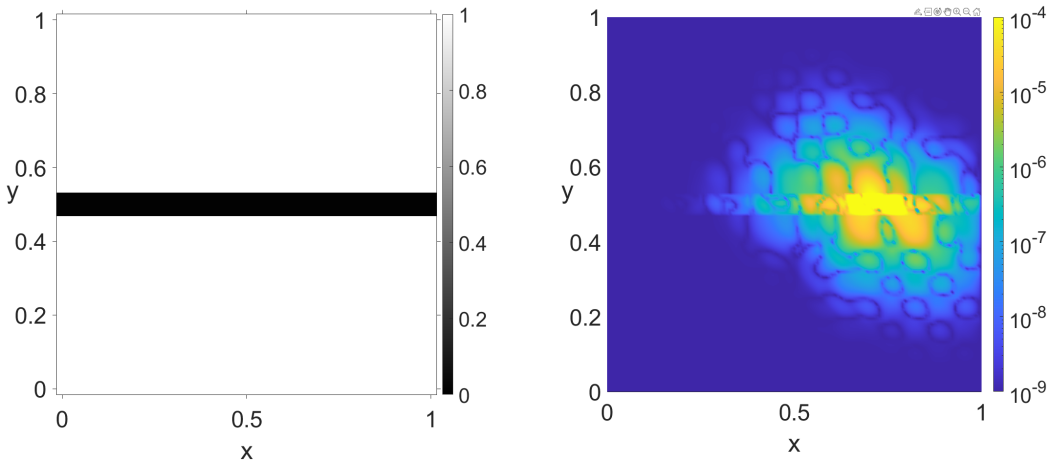


Figure 5.6: *Diffusion coefficient with thin channel valued 0.01 to the left, corresponding solution φ_h to the right.*

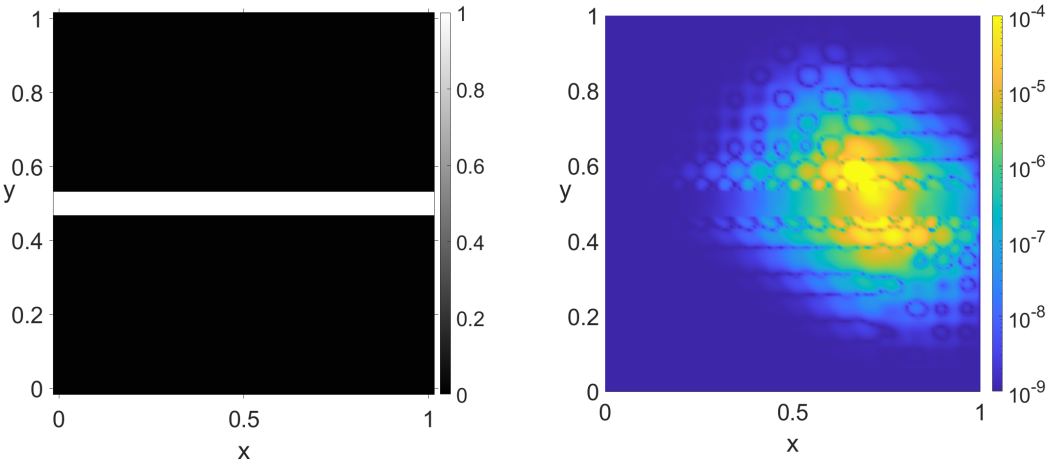


Figure 5.7: *Diffusion coefficient with thin channel valued 1 to the left, corresponding solution φ_h to the right.*

Lastly, in order to further visualize the properties discussed we have implemented an oscillating diffusion coefficient, which oscillates between 0.01 and 1 in the x-direction, and is constant in the y-direction. This is shown in Figure 5.8, and while the mesh is hard to distinguish, one can see the the solution is oscillating slightly. The error plot in Figure 5.9d shows a very quick decay.

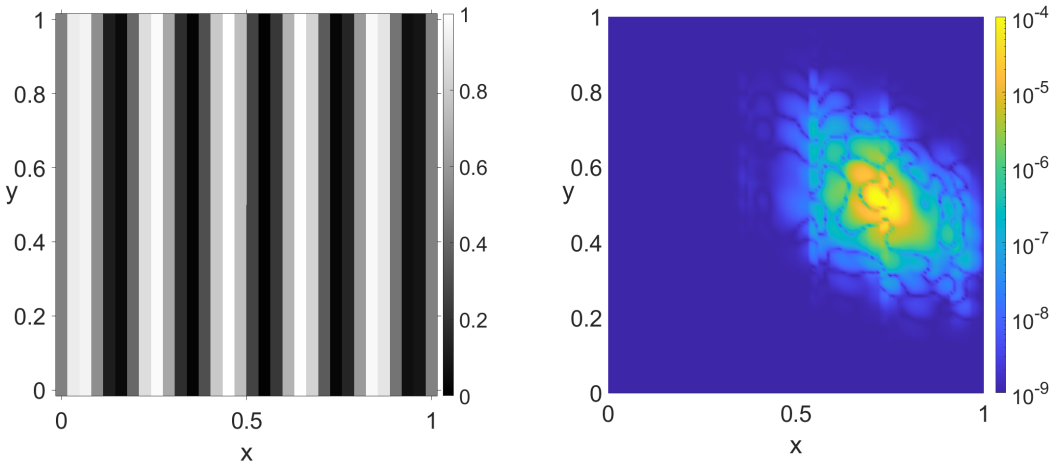
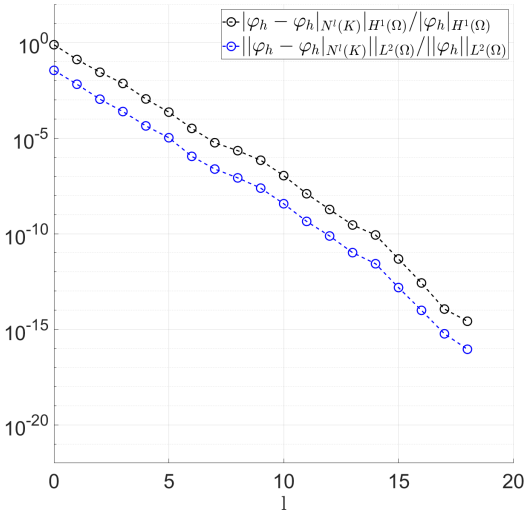
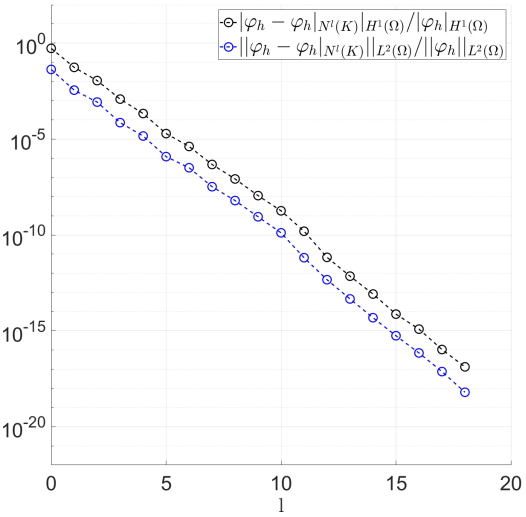


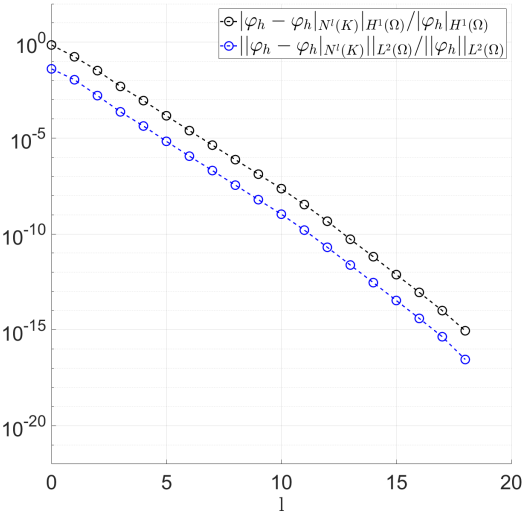
Figure 5.8: *Oscillating coefficient to the left, corresponding solution φ_h to the right.*



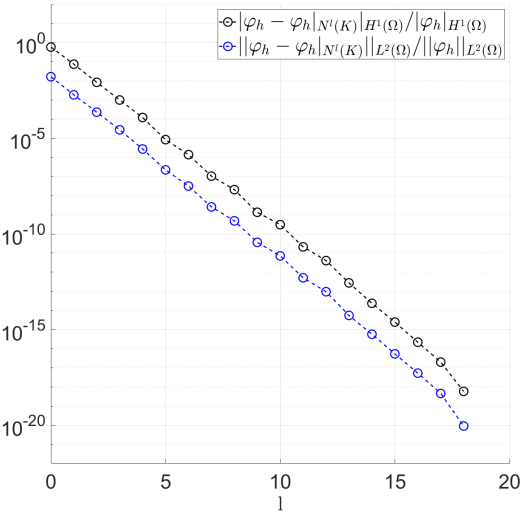
(a) *Relative error plot visualizing spatial decay with randomized diffusion coefficient, corresponding to Figure 5.5*



(b) *Relative error plot visualizing spatial decay with diffusion coefficient with thin channel valued* 0.01 *running through, corresponding to Figure 5.6*



(c) *Relative error plot visualizing spatial decay with diffusion coefficient with thin channel valued* 1 *running through, corresponding to Figure 5.7.*



(d) *Relative error plot visualizing spatial decay with oscillating diffusion coefficient, corresponding to Figure 5.8.*

Figure 5.9: *Error plots corresponding to randomized, thin 0.01 valued channel, thin 1 valued channel and oscillating diffusion coefficients.*

5.3 Parabolic case

We will approach our investigation of the parabolic case in a similar manner as the elliptic case. This time we will use a simple load vector f which is localized in both space and time. It is defined as $f \equiv 1$ on $K \times [0, 0.125]$ for $K \in \mathcal{K}_H$ with vertices in $(0.6875, 0.5)$, $(0.6875, 0.5625)$ and $(0.75, 0.5)$, and as $f \equiv 0$ everywhere else. The implemented function f is visualized in Figure 5.10.

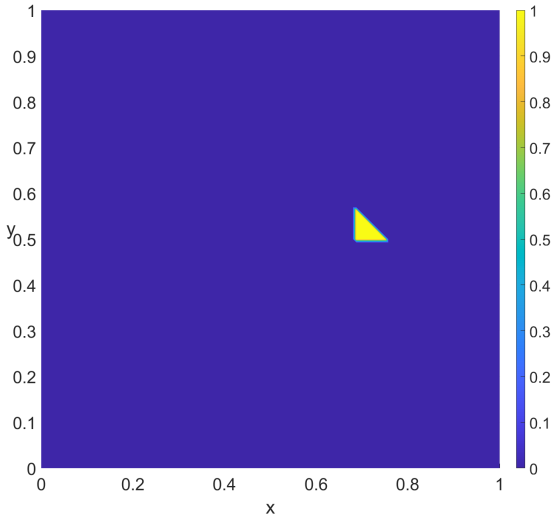


Figure 5.10: *Localized f used in implementations at any $t \in [0, 0.125]$.*

We will apply different diffusion coefficients in (4.1). We have also chosen $T = 5$ as upper time limit while studying the systems. This limit is arbitrary, and one could use a lower or higher one, but this limit is sufficient for making the discussed features of our solutions to emerge. Note also that the time limit essentially works as truncating the solution, and does not affect the solution up until that point in time.

In our first example we will use a constant diffusion coefficient, $A \equiv 1$. The result is shown in Figure 5.11. Firstly, one can confirm that some sort of temporal decay seems to be taking place, since the whole solution seems to fade with time. But one can also suspect that there is an exponential spatial decay present from these images. Note how the solution achieves significantly higher values around K, where f is supported, and then fades out toward the edges. As in the elliptic case, we also see some hints of the spatial grid in the solution, although much less clearly. We have also implemented an error plot for our solution, in this case seen in Figure 5.17. The plots illustrate the relative errors of truncated solutions, spatially and temporally. More specifically, we compare $\psi_{h,\tau}$ to $\psi_{h,\tau}|_{N^l(K)}$ for $l = 1, \ldots, 2^5$ and $\psi_{h,\tau}\mathbb{1}_{[0,T_i]}$ for $i = 0, \ldots, 5 \cdot 2^4$ to confirm spatial and temporal decay respectively. All of the error plots seem to confirm the earlier proposals of temporal decay and exponential spatial decay.

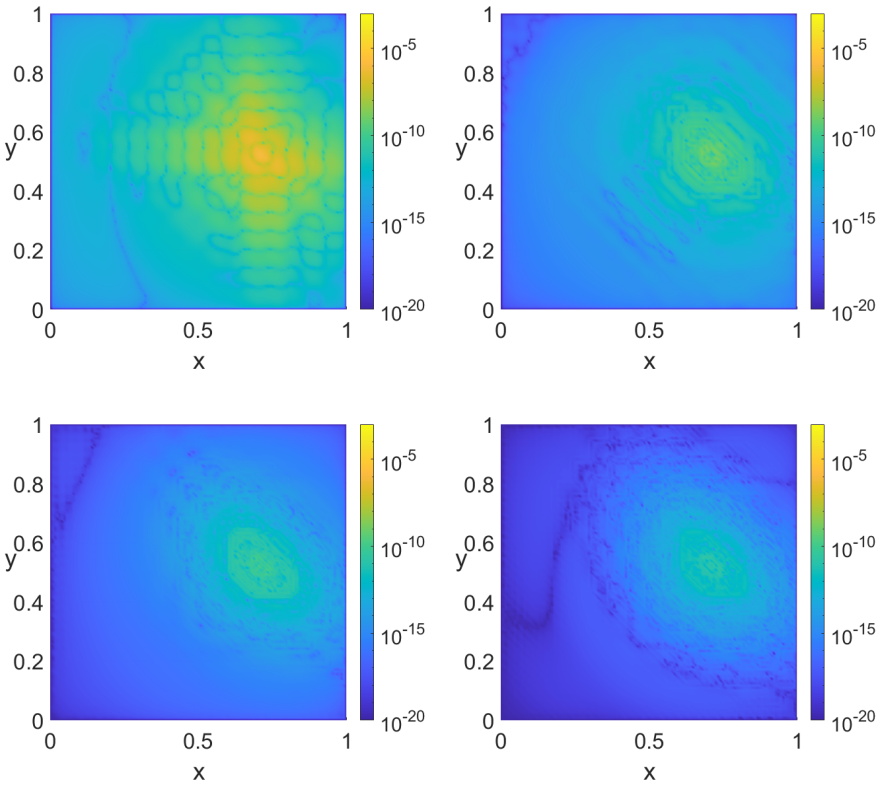


Figure 5.11: *The corresponding solution* $\psi_{h,\tau}$ *to* $A \equiv 1$ *at times* $t \approx 0.001, 1, 2.5, 4.99$.

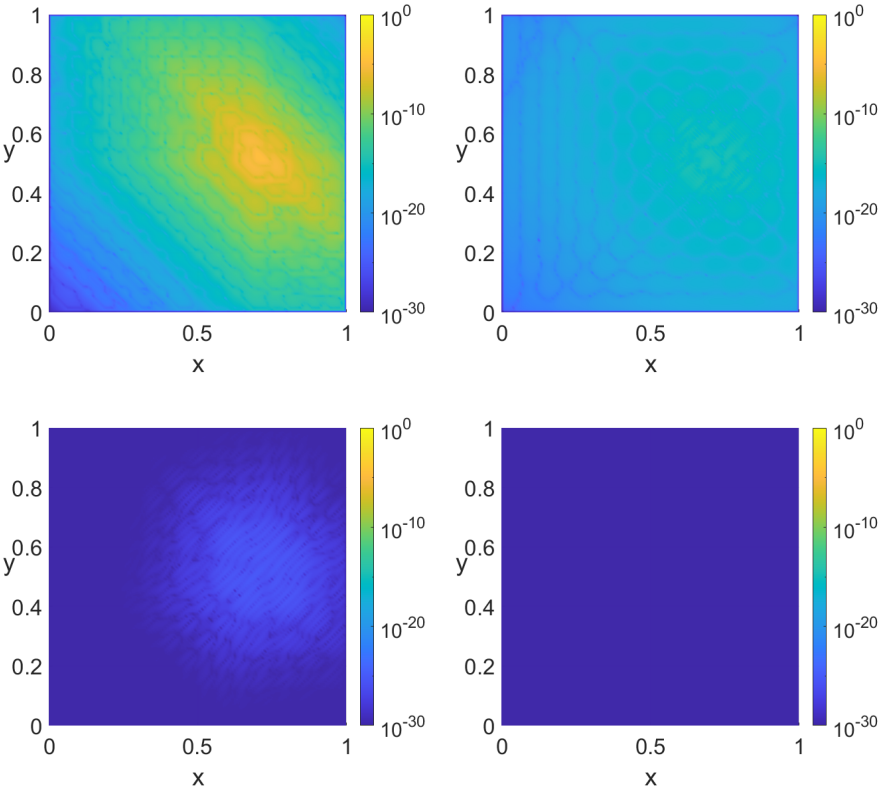


Figure 5.12: *The corresponding solution* $\psi_{h,\tau}$ *to to* $A \equiv 0.01$ *at times* $t \approx 0.001, 1, 2.5, 4.99$. *Note the adjusted scales.*

We have also implemented a second constant diffusion coefficient, $A \equiv 0.1$, where the corresponding plots are shown in Figure 5.12. The same observations goes as for our previous case, but this time the temporal and spatial decay is much faster, which becomes very clear if we compare Figure 5.18 to Figure 5.17, where we adjust the axis for this special case. This tells us that the magnitude of A affects the decay.

In Figure 5.14 we see the results after implementing a highly irregular, randomized coefficient A. Its range is divided into space-time rectangles and randomly assigned values between 0.01 and 1. As in previous cases we see a clear temporal decay. However, if one compares it to the decay in Figure 5.11 and looks at Figure 5.19 and Figure 5.17, it is significantly slower here. This might be explained by that we are comparing a diffusion coefficient which are constant in time to a coefficient which is not.

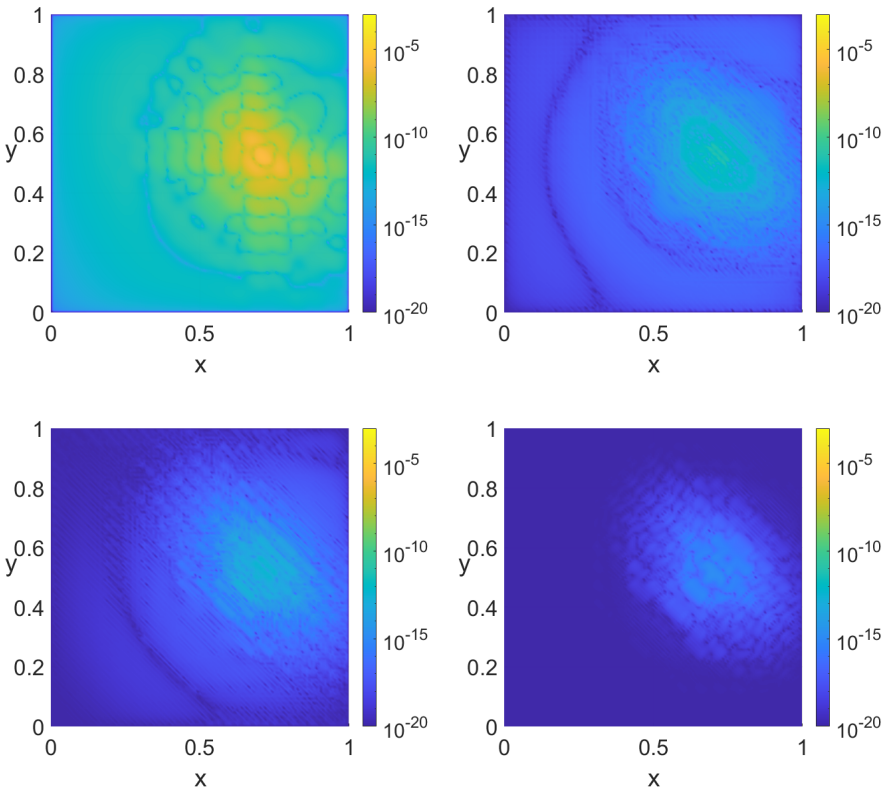


Figure 5.13: *The corresponding solution $\psi_{h,\tau}$ to temporally oscillating diffusion coefficient at times $t \approx 0.001, 1, 2.5, 4.99$. The coefficient is defined as $A(t,x) := \sin(2\pi t)$.*

Another interesting comparison is between Figure 5.13 and Figure 5.14. In Figure 5.13, we have implemented a diffusion coefficient which only oscillates in time. It is a discretized version of a sine wave oscillating between 0.01 and 1, with a period of 1. Even though neither our randomized coefficient or oscillating coefficient is constant in time, we see that regularity pays an important role in the decay. In Figure 5.20 this is confirmed as well, since one can see that the temporal error decreases much faster than in Figure 5.19. Note also how the temporal error graph seems to oscillate in unison with the coefficient.

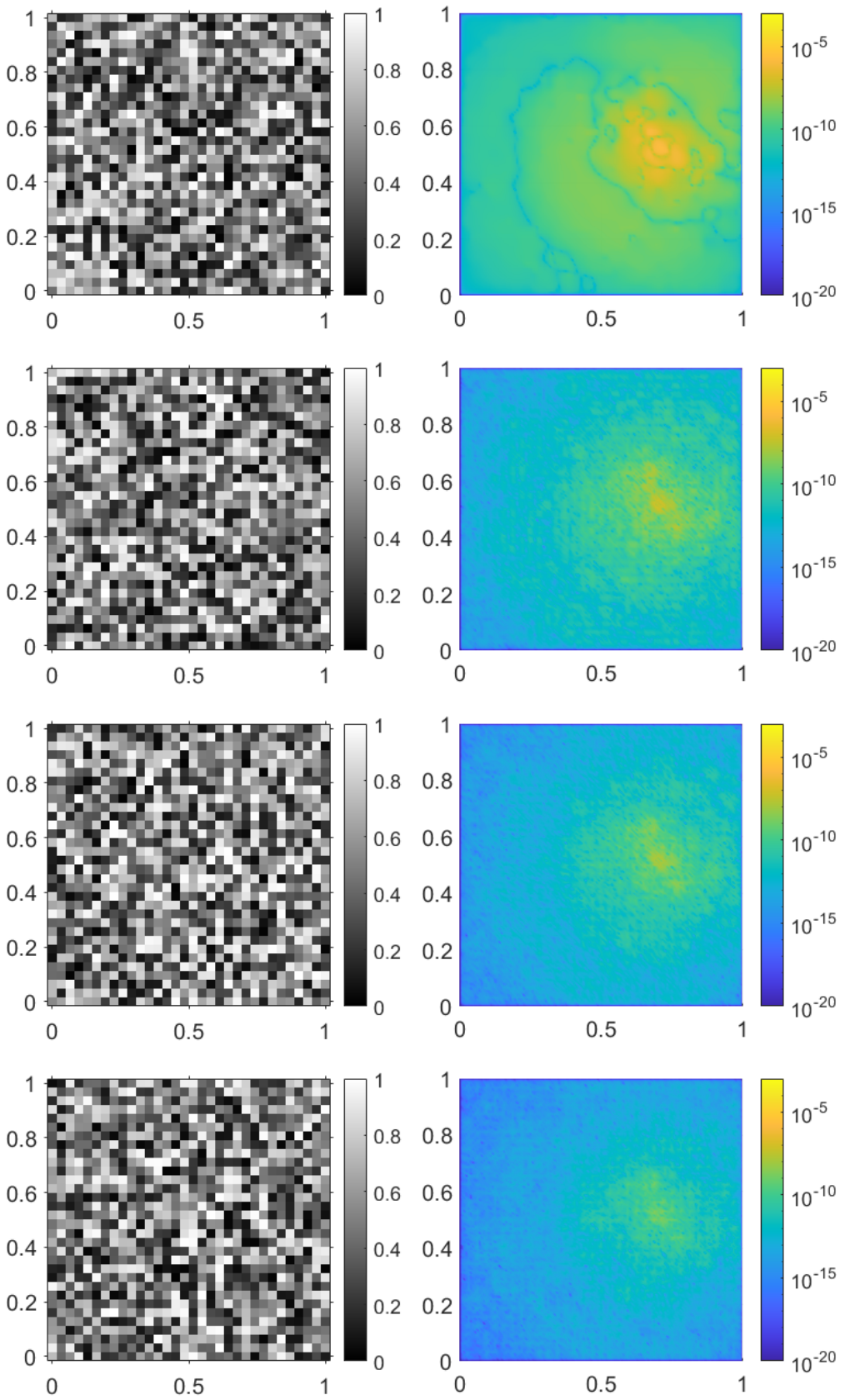


Figure 5.14: *A randomized highly varying diffusion coefficient A (to the left) and the corresponding solution $\psi_{h,\tau}$ (to the right) at times $t \approx 0.001, 1, 2.5, 4.99$.*

One observation we are yet to make, is how the solution imitates the diffusion coefficient, which we saw more clearly in the elliptic case. To make this phenomenon more apparent, we construct two coefficients which are constant in time, but spatially constructed with a thin channel running through the domain. The results are presented in Figure 5.15 and Figure 5.16, where one can clearly see the coefficient reflected in the solution. We also note that in Figure 5.16, the solution decays significantly slower within the value channel, and in Figure 5.16 the opposite seems to happens. The temporal errors in Figure 5.21 and Figure 5.22 do not differ much, but the "high"-valued channel A has a significantly faster spatial decay. Note also the significant bump in spatial decay in Figure 5.22. This bump and increase of speed in the decay coincides with $l = 10$, which happens to be the point where $N^l(K)$ completely covers the channel, which could explain the phenomenon.

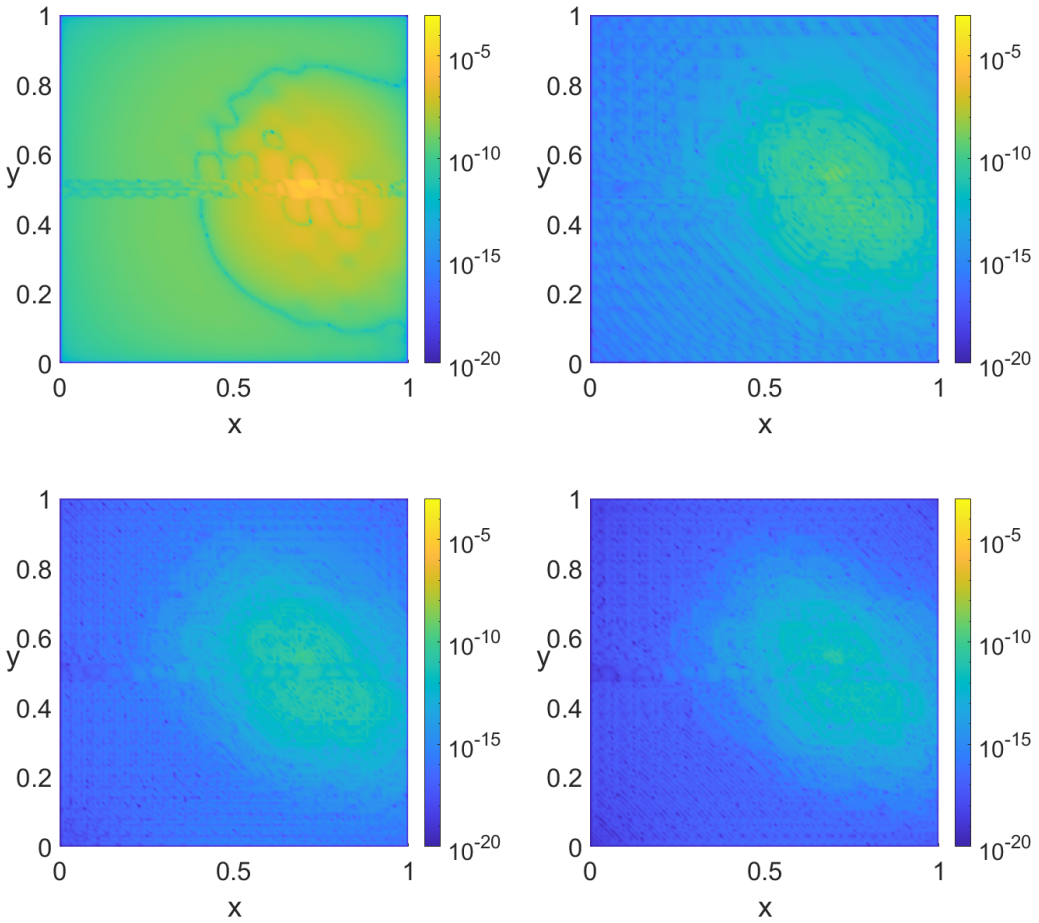


Figure 5.15: *The corresponding solution $\psi_{h,\tau}$ to diffusion coefficient with thin channel valued* 0.01 *(see Figure 5.6) at times* $t \approx 0.001, 1, 2.5, 4.99$.

Lastly, one phenomenon which is difficult to convey in pictures, is that in all cases of different diffusion coefficients, the solution seemed to radiate waves of higher and lower values originating at the element where f is supported. One might suspect that this behaviour stems from the fact that the solution has a temporal coarse scale constraint here, and that is what we see materialized.

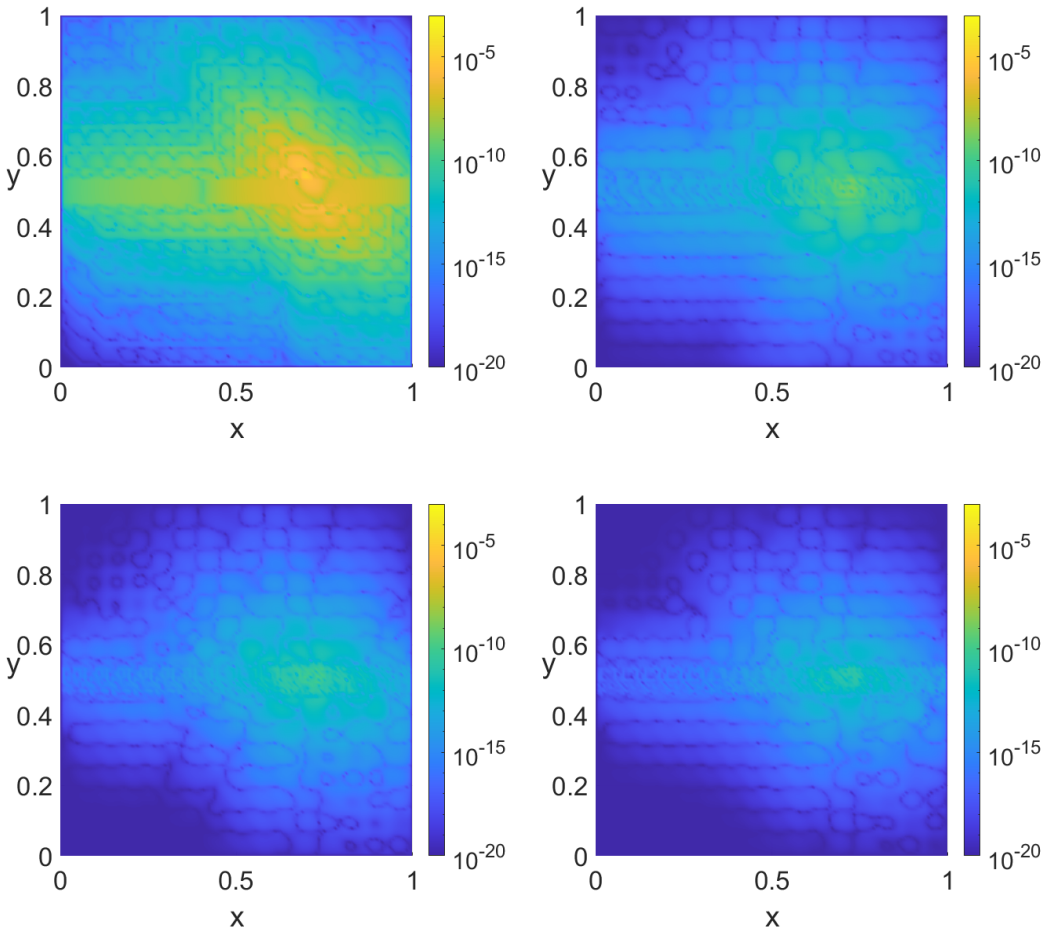


Figure 5.16: *The corresponding solution $\psi_{h,\tau}$ to diffusion coefficient with thin channel valued* 1 *(see Figure 5.7) at times $t \approx 0.001, 1, 2.5, 4.99$.*

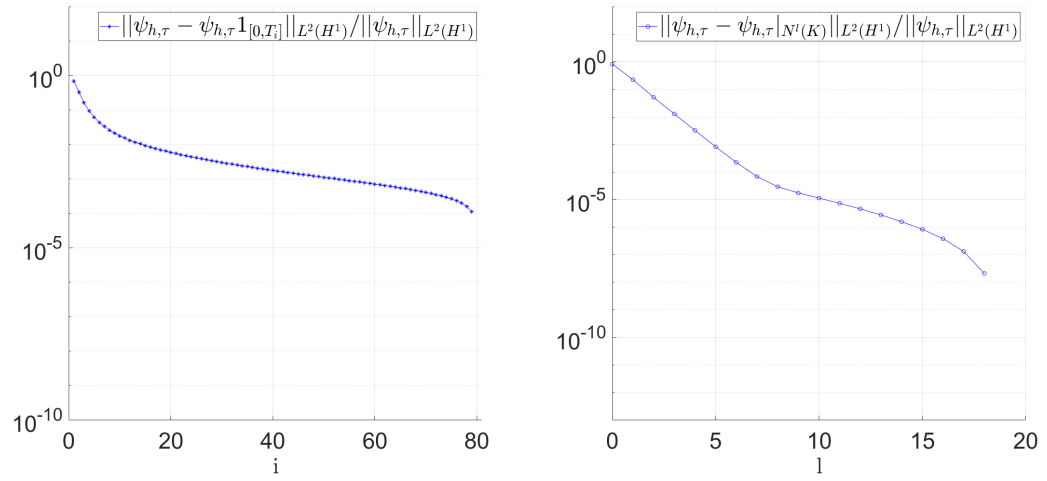


Figure 5.17: *Relative error plots illustrating temporal (to the left) and spatial decay (to the right) for constant diffusion coefficient valued* 1*. Corresponds to Figure 5.11*

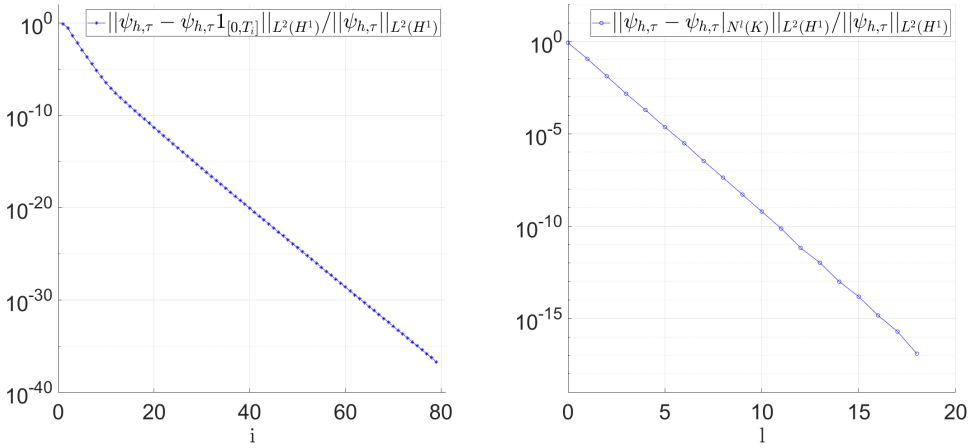


Figure 5.18: *Relative error plots illustrating temporal (to the left) and spatial decay (to the right) for constant diffusion coefficient valued* 0.01*. Corresponds to Figure 5.12.*

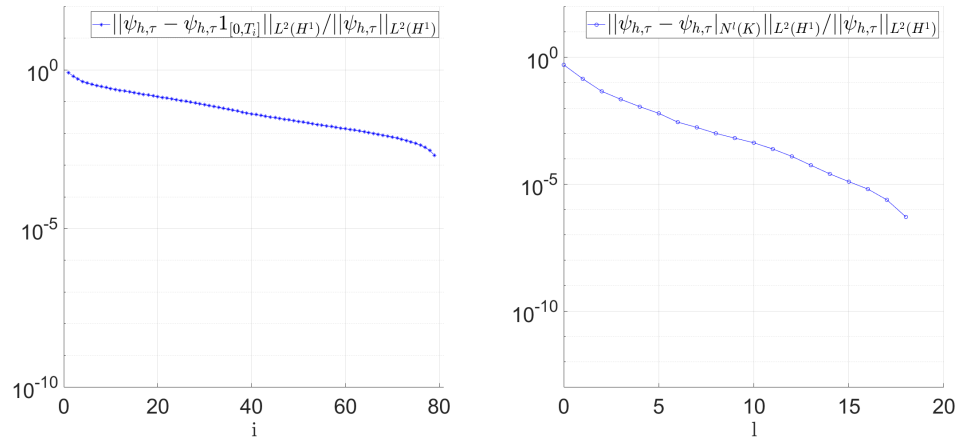


Figure 5.19: *Relative error plots illustrating temporal (to the left) and spatial decay (to the right) for randomized diffusion coefficient. Corresponds to Figure 5.14.*

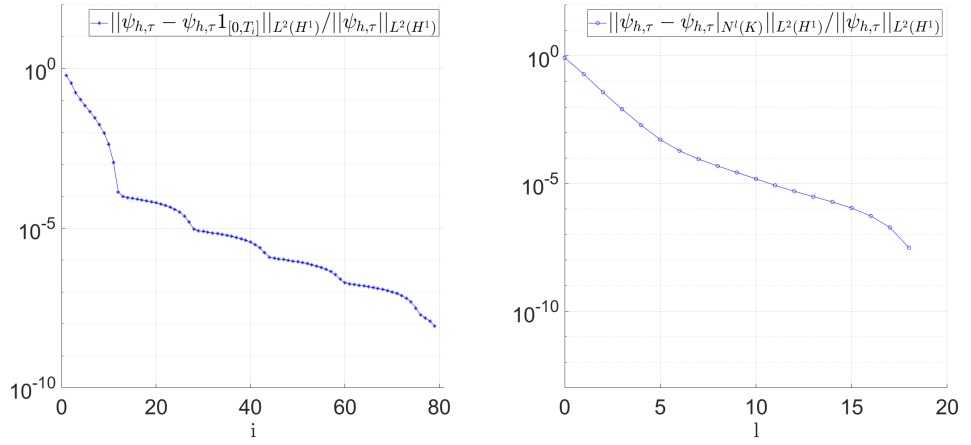


Figure 5.20: *Relative error plots illustrating temporal (to the left) and spatial decay (to the right) for temporally oscillating diffusion coefficient. Corresponds to Figure 5.13.*

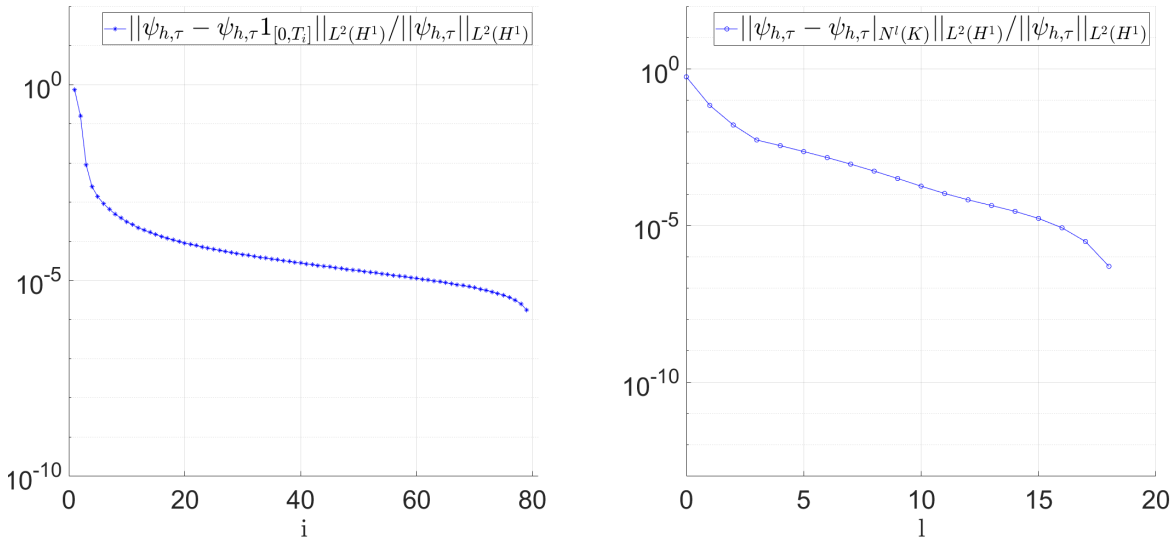


Figure 5.21: *Relative error plots illustrating temporal (to the left) and spatial decay (to the right) for diffusion coefficient with thin channel valued* 0.01*. Corresponds to Figure 5.15.*

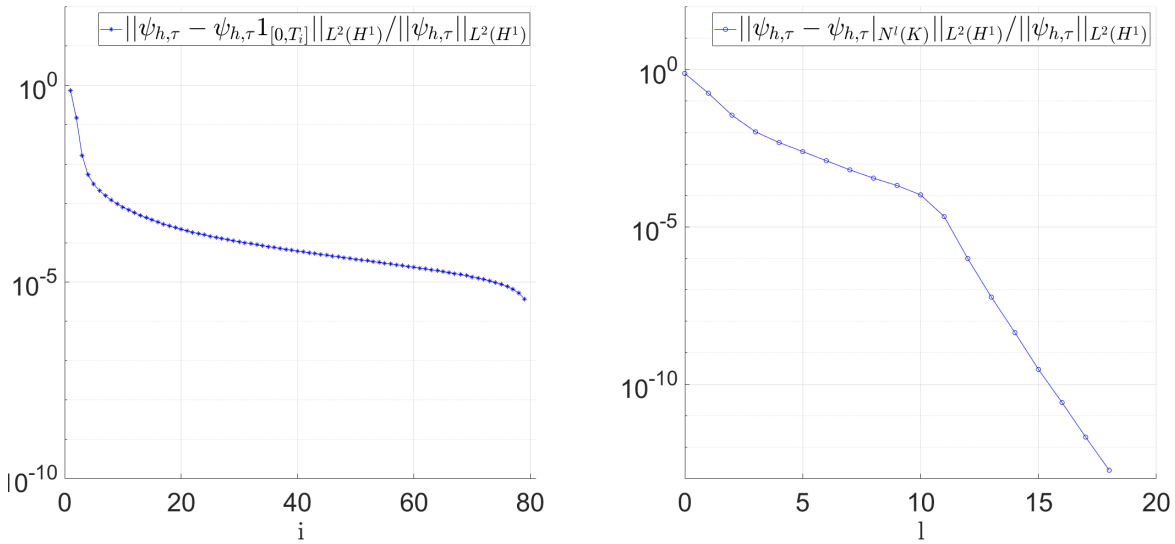


Figure 5.22: *Relative error plots illustrating temporal (to the left) and spatial decay (to the right) for diffusion coefficient with thin channel valued* 1*. Corresponds to Figure 5.16.*

6
Conclusion

The purpose of this report was to investigate the nature of coare-scale constrained problems, a special type of PDAEs. This was done both theoretically, by literature studies, and numerically, using MATLAB. The questions we chose to focus on was existence and uniqueness of solutions, as well as the nature of the decay of said solutions.

From the evidence put forth in this report we can expect these solutions to coarse-scale constrained equations to be well-behaved.

In the elliptic case we showed in chapter 3 that our designated problem has a unique solution, and later that it could be well-approximated by some finite element method. We also showed that the solution decays exponentially when we have a localized load vector. Lastly in section 5.2, we confirmed the claim of exponential decay numerically with MATLAB code, as well as noted some additional features. For example, we saw how the solution reflects different diffusion coefficient visually, and how the mesh structure appears vaguely in the solution, likely because of the coarse-scale constraint.

In the parabolic case we restricted ourselves to a fully discrete case for our theoretical proofs in chapter 4, which was a practical choice since the theory is closer to implementations that way. Again, we proved that the constrained problem in question has a unique solution and provided a stability estimate. As for the questions surrounding decay, we restricted ourselves to a proof of an a posteriori bound for the solution, but pointed out that one could suspect that there is also exponential spatial decay present from our numerical results in chapter 5. The mathematical proof of this is an open problem currently.

As far as future ventures in this subject goes, one might want to look into more general elliptic and parabolic equations, since we restricted this report to one PDE for each case with a second order term. One could introduce various lower order terms, which could prove convenient since the problems then becomes formulated in more general terms. One could also investigate how the choice of right-hand sides affects the solution, since we only made use of one for each case in this report.

Lastly, another natural progression from this report could also be to investigate hyperbolic problems with coarse scale constraints, where many of properties that we discussed here are still open problems.

References

1. Benner, P., Clees, T., Grundel, S., Hornung, N., Jansen, L. & Tischendorf, C. in *Progress in differential-algebraic equations* 183–205 (Springer, Heidelberg, 2014).
2. Berlyand, L., Owhadi, H. & Zhang, L. Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization. *ESAIM: Mathematical Modelling and Numerical Analysis* **48,** 517–552 (Mar. 2014).
3. Bernardi, C., Canuto, C. & Maday, Y. Generalized inf-sup conditions for Chebyshev spectral approximation of the Stokes problem. *SIAM J. Numer. Anal.* **25,** 1237–1271 (1988).
4. Bodestedt, M. & Tischendorf, C. PDAE models of integrated circuits and index analysis. *Mathematical and Computer Modelling of Dynamical Systems* **13,** 1–17 (Feb. 2007).
5. Boffi, D., Brezzi, F. & Fortin, M. *Mixed finite element methods and applications* xiv+685 (Springer, Heidelberg, 2013).
6. Bramble, J. H., Pasciak, J. E. & Steinbach, O. On the stability of the L^2 projection in $H^1(\Omega)$. *Math. Comp.* **71,** 147–156 (2002).
7. Brenner, S. C. & Scott, L. R. *The mathematical theory of finite element methods* Third, xviii+397 (Springer, New York, 2008).
8. Clément, P. Approximation by finite element functions using local regularization. *ESAIM: Mathematical Modelling and Numerical Analysis - Modélisation Mathématique et Analyse Numérique* **9,** 77–84 (1975).
9. Dieuleveult, C., Erhel, J. & Kern, M. A global strategy for solving reactive transport equations. *Journal of Computational Physics* **228,** 6395–6410 (Sept. 2009).
10. Efros, A. & Pokrovsky, A. L. Diffraction theory and focusing of light by a slab of left-handed material. *Physica B-condensed Matter* **338,** 333–337 (2003).
11. Eich-Soellner, E. & Führer, C. *Numerical methods in multibody dynamics* 290 (B. G. Teubner, Stuttgart, 1998).
12. Emmrich, E. & Mehrmann, V. Operator differential-algebraic equations arising in fluid dynamics. *Comput. Methods Appl. Math.* **13,** 443–470 (2013).
13. Ern, A. & Guermond, J.-L. Finite element quasi-interpolation and best approximation. *ESAIM: Mathematical Modelling and Numerical Analysis* **51** (May 2015).
14. Feijóo, G. R., Hughes, T. J., Mazzei, L. & Quincy, J.-B. The variational multiscale method—a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering* **166.** Advances in Stabilized Methods in Computational Mechanics, 3–24 (1998).
15. Hellman, F. & Målqvist, A. Contrast Independent Localization of Multiscale Problems. *Multiscale Modeling & Simulation* **15,** 1325–1355 (2017).

16. Larson, M. G. & Målqvist, A. Goal oriented adaptivity for coupled flow and transport problems with applications in oil reservoir simulation. *Computer Methods in Applied Mechanics and Engineering* **196,** 3546–3561 (2007).
17. Larsson, S. & Thomée, V. *Partial differential equations with numerical methods* x+259 (Springer-Verlag, Berlin, 2003).
18. Ljung, P., Maier, R. & Målqvist, A. *A space-time multiscale method for parabolic problems* 2021. arXiv: 2109.06647 [math.NA].
19. Maier, R. *Computational Multiscale Methods in Unstructured Heterogeneous Media* , Doctoral thesis, (Universität Augsburg, 2020).
20. Målqvist, A. & Peterseim, D. *Numerical homogenization by localized orthogonal decomposition* xii+108 (Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2021).
21. Målqvist, A. & Peterseim, D. Localization of elliptic multiscale problems. *Mathematics of Computation* **83,** 2583–2603 (2014).
22. Patera, A. & Urban, K. An improved error bound for reduced basis approximation of linear parabolic problems. *Mathematics of Computation* **83** (July 2014).
23. Sohr, H. *The Navier-Stokes equations* An elementary functional analytic approach, x+367 (Birkhäuser Verlag, Basel, 2001).