THESIS FOR THE DEGREE OF LICENTIATE OF PHILOSOPHY

Generalized finite element methods for time-dependent multiscale problems

Per Ljung

CHALMERS
UNIVERSITY OF TECHNOLOGY

UNIVERSITY OF GOTHENBURG

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

Generalized finite element methods for time-dependent multiscale problems
Per Ljung
Gothenburg 2021

Division of Applied Mathematics and Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg
SE-412 96 Gothenburg
Sweden
Telephone +46 (0)31 772 1000

Author e-mail: perlj@chalmers.se

Typeset with LaTeX
Printed in Gothenburg, Sweden 2021

Generalized finite element methods for time-dependent multiscale problems

Per Ljung

Division of Applied Mathematics and Statistics
Department of Mathematical Sciences
Chalmers University of Technology and University of Gothenburg

Abstract

In this thesis we develop and analyze generalized finite element methods for time-dependent partial differential equations (PDEs). The focus lies on equations with rapidly varying coefficients, for which the classical finite element method is insufficient, as it requires a mesh fine enough to resolve the data. The framework for the novel methods are based on the localized orthogonal decomposition technique, introduced in [35]. The main idea of this method is to construct a modified finite element space whose basis functions contain information about the variations in the coefficients, hence yielding better approximation properties.

At first, the localized orthogonal decomposition framework is extended to the strongly damped wave equation, where two different highly varying coefficients are present (Paper I). The dependency of the solution on the different coefficients vary with time, which the proposed method accounts for automatically. Then we consider a parabolic equation where the diffusion is rapidly varying in both time and space (Paper II). Here, the framework is extended so that the modified finite element space uses space-time basis functions that contain the information of the diffusion coefficient. In both papers we prove error estimates for the methods, and confirm the theoretical findings with numerical examples.

Keywords: Strongly damped wave equation, multiscale, localized orthogonal decomposition, finite element method, parabolic equations.

List of appended papers

Paper I Per Ljung, Axel Målqvist, and Anna Persson
A generalized finite element method for the strongly damped wave equation with rapidly varying data.
Submitted

Paper II Per Ljung, Roland Maier, and Axel Målqvist
A space-time multiscale method for parabolic problems.
Preprint

Acknowledgements

First of all, I wish to express my gratitude to my supervisor Axel Målqvist for his fantastic guidance and support throughout the work of this thesis. Your expertise and positivity have been invaluable to me these last few years. I would also like to thank my co-authors Anna Persson and Roland Maier. I have learned so much from all of you about the academic world, and about mathematics in general. I am truly lucky to have had the luxury of working with the three of you.

Furthermore, I want to thank my colleagues at the department for providing such a friendly working environment. In particular, thanks to Gustav Lindwall for always letting me knock on your door and helping me with all sorts of work related things, and for our many coffee break strolls.

At last, I want to thank my many close friends for all online and offline activities we have shared the last year. Especially, thanks to Therese for your constant love and support, and for always lifting my spirits. It is thanks to these people I have managed to maintain my sanity through this ongoing pandemic, and I am truly grateful to all of them.

Contents

Introduction

1 Background

The modeling of partial differential equations (PDEs) is a major topic in both science as well as industry. Applications range from simulating the aerodynamics of large aircraft to modeling of atoms on a quantum mechanic level. In this thesis, we focus on time-dependent partial differential equations. In particular, our interest lies in simulations on materials that are strongly heterogeneous, e.g. composite materials. The modeling of such materials result in PDEs with highly oscillatory coefficients. This type of problems, where the data is varying rapidly, is referred to as *multiscale problems*.

One of the most common numerical approaches to solving PDEs is the finite element method (FEM). However, for multiscale problems, FEM approximations are not accurate unless the mesh is fine enough to resolve the variations in the data. This quickly becomes challenging in terms of computational cost and memory. For the purpose of solving multiscale equations, several so called *multiscale methods* have been developed. Some of these methods are derived from analytical homogenization theory, such as the Heterogeneous Multiscale Method (HMM) [11, 1] or the Multiscale Finite Element Method [19]. As these techniques are based on an analytical framework, they require assumptions in terms of periodicity and scale separation. In addition, there have been several so-called numerical multiscale methods developed that circumvent these sorts of requirements. Examples of such methods include generalized (Multiscale) Finite Element Methods [7, 6, 12] and gamblets [37, 38]. In particular, in [35], a generalized finite element method (GFEM) known as the localized orthogonal decomposition method (LOD) is introduced.

The LOD method is based on the Variational Multiscale Method (VMS), first introduced in [20]. The main idea of the method is to construct a modified coarse-scale finite element space enriched by problem-dependent fine-scale

correctors. In turn, it was proven in [35] (and later improved in [18]) that these correctors satisfy an exponential decay, making it possible to compute them solely on local patches, which is one of the advantages of the LOD method. Another strength is that the correctors are completely independent and can be computed in parallel. In [35], convergence of optimal order is proven independent of the variations in the data, with no assumptions made on neither periodicity nor scale separation. The LOD method was first introduced for elliptic equations, but has since then been further developed and analyzed for several types of problems, including time-dependent PDEs such as parabolic ones in [33, 34, 3, 31] and wave-type equations in [2, 32]. For more results on the LOD method, see e.g. [17, 16, 30].

The purpose of this thesis is to extend the LOD framework for problems that require the time-dependency to be taken into account in the construction of the multiscale method. This thesis discusses two such cases. In Paper I, we derive a GFEM with the LOD framework as basis for the so called *strongly damped wave equation*. This case is unique as the PDE consists of two different multiscale coefficients, while multiscale methods such as LOD are in general designed to handle only one coefficient. In Paper II, we consider a parabolic equation, where the diffusion varies rapidly in both time and space. In both papers, convergence of optimal order is proven independent of the high variations in the data, and numerical examples are presented that demonstrate the performance of the methods.

In the upcoming sections, we introduce the LOD method with its prerequisites, and further present the methods and main results derived in Paper I and Paper II. In Section 2 we present the classical finite element method and illustrate why it is not sufficient for problems where the data is highly varying. In Section 3 we look at the LOD method introduced in [35] for elliptic equations. Finally, Section 4 presents the methods and summarizes the main results from Paper I and Paper II.

2 The finite element method

In this section, we discuss the finite element method and why it is insufficient when the data is highly oscillatory. We restrict the discussion to the case of elliptic PDEs with homogeneous Dirichlet boundary condition. That is, we

consider the equation

$$\begin{aligned} -\nabla \cdot (A\nabla u) &= f, \text{ in } \Omega, \\ u &= 0, \text{ on } \partial\Omega, \end{aligned} \tag{2.1}$$

where $\Omega \subset \mathbb{R}^d$, $d = 2, 3$ is a bounded Lipschitz-domain with polygonal boundary, $f \in L_2(\Omega)$ is the source function, and $A := A(x) \in L_\infty(\Omega, \mathbb{R}^{d\times d}_{\text{sym}})$ is the (highly oscillatory) diffusion coefficient that satisfies

$$0 < \alpha := \operatorname*{ess\,inf}_{x\in\Omega} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{A(x)v\cdot v}{v\cdot v} \leq \operatorname*{ess\,sup}_{x\in\Omega} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{A(x)v\cdot v}{v\cdot v} =: \beta < \infty.$$

We begin by introducing the standard Sobolev spaces used in finite element theory. Therefore, let $\alpha = (\alpha_1, \dots \alpha_d)$ be a multi-index, and define

$$\mathrm{D}^\alpha \varphi = \frac{\partial^{|\alpha|}\varphi}{\partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}}.$$

We say that v is the *α'th order weak derivative of u* if

$$\int_\Omega u\mathrm{D}^\alpha \varphi \,\mathrm{d}x = (-1)^{|\alpha|} \int_\Omega v\varphi \,\mathrm{d}x, \quad \forall \varphi \in C_0^{|\alpha|}(\Omega),$$

where $C_0^{|\alpha|}(\Omega)$ is the space of $|\alpha|$ times continuously differentiable functions with compact support in Ω. With the weak derivative defined, we may construct the Sobolev space $H^k(\Omega)$ for $k \geq 0$ as the space of all functions whose weak derivatives of order smaller or equal to k belong to $L_2(\Omega)$, i.e.,

$$H^k(\Omega) := \{v \in L_2(\Omega) \colon \mathrm{D}^\alpha v \in L_2(\Omega) \text{ for } |\alpha| \leq k\},$$

equipped with inner product and corresponding norm

$$\begin{aligned} (v, w)_{H^k} &:= \sum_{|\alpha|\leq k} \int_\Omega \mathrm{D}^\alpha v \mathrm{D}^\alpha w \,\mathrm{d}x, \\ \|v\|_k^2 := (v, v)_{H^k} &= \sum_{|\alpha|\leq k} \int_\Omega (\mathrm{D}^\alpha v)^2 \,\mathrm{d}x. \end{aligned}$$

We moreover define the corresponding seminorm on $H^k(\Omega)$ as

$$|v|_k^2 = \sum_{|\alpha|=k} \int_\Omega (\mathrm{D}^\alpha v)^2 \,\mathrm{d}x.$$

In particular, we introduce $H_0^1(\Omega)$ as the classical Sobolev space with norm

$$\|v\|_{H^1(\Omega)}^2 = \|v\|_{L_2(\Omega)}^2 + \|\nabla v\|_{L_2(\Omega)}^2$$

whose functions vanish on $\partial\Omega$, in the sense of traces. In the following, we abbreviate the $L_2(\Omega)$-norm as $\|\cdot\| := \|\cdot\|_{L_2(\Omega)}$.

We derive the weak formulation corresponding to (2.1), on which we base our finite element method. By standard procedure we multiply the equation by a test function $v \in H_0^1(\Omega)$ and integrate by parts over the domain Ω. The weak formulation becomes: find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v), \quad \forall v \in H_0^1(\Omega), \tag{2.2}$$

where $(\cdot, \cdot)$ is the standard L_2-scalar product and $a(\cdot, \cdot) := (A\nabla\cdot, \nabla\cdot)$. We wish to define a FEM for the problem by constructing a discretized version of the weak formulation (2.2). To this end, let $\{\mathcal{T}_h\}_{h>0}$ denote a family of shape regular elements that form a partition of the domain Ω. For an element $T \in \mathcal{T}_h$ we define the corresponding mesh size as $h_T := \operatorname{diam}(T)$, and denote the largest diameter by $h := \max_{T\in\mathcal{T}_h} h_T$. Now we construct the classical finite element space using continuous piecewise linear polynomials as

$$V_h := \{v \in H_0^1(\Omega) : \ v|_T, \, T \in \mathcal{T}_h, \text{ is a polynomial of partial degree } \leq 1\}.$$

The finite element formulation now follows by considering a Galerkin ansatz based on the discretized space V_h. More precisely: find $u_h \in V_h$ such that

$$a(u_h, v) = (f, v), \quad \forall v \in V_h. \tag{2.3}$$

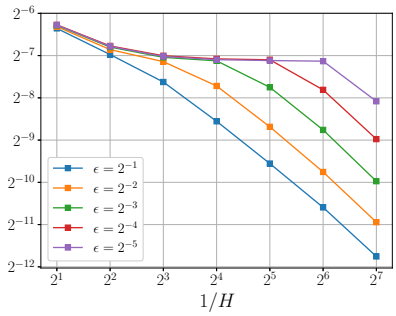
From standard a priori error analysis (see e.g. [26, Theorem 5.1]), the following error bound is derived for the finite element approximations in (2.3)

$$\|u_h - u\|_{H^1} \leq Ch\|u\|_2,$$

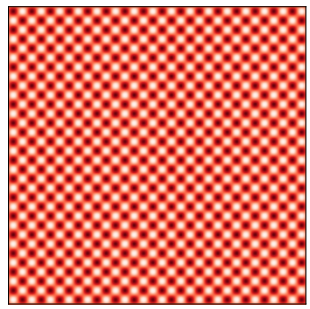
This convergence result is valid for $u \in H^2(\Omega)$, which is satisfied under the assumption $A \in C^1(\Omega)$ for a convex domain Ω. Assume for the moment that A is a scalar valued coefficient. The H^2-seminorm can further be bounded as

$$\begin{aligned} |u|_2 &\leq C\|\Delta u\| \leq C\|A\nabla\cdot(\nabla u)\| = C\|\nabla\cdot(A\nabla u) - \nabla A\nabla u\| \\ &\leq C\|\nabla\cdot(A\nabla u)\| + \|\nabla A\|_{L_\infty}\|\nabla u\| \leq C(1 + \|\nabla A\|_{L_\infty})\|f\| \end{aligned}$$

where the first inequality follows from elliptic regularity and the last inequality from $\|u\|_{H^1} \leq C\|f\|$, which can be derived from the weak formulation. Note here that the error is bounded by the derivative of the diffusion A. Hence, if A



(a) Energy error.



(b) $A(x_1, x_2)$.

Figure 1: The left image shows the energy error $|||u_h - u_{\text{ref}}|||$ as function of mesh size for different choices of scales ε. The right image is the coefficient used in the case $\varepsilon = 2^{-4}$.

is rapidly varying with frequency ε^{-1}, then $\|\nabla A\|_{L_\infty} = \mathcal{O}(\varepsilon^{-1})$. Consequently, the error for the FEM has the upper bound

$$\|u_h - u\|_{H^1} \leq C \min\{h + h\varepsilon^{-1}, 1\}\|f\|.$$

In practice this means that the sought convergence rate may not be achieved unless $h < \varepsilon$. As ε gets smaller, i.e. the diffusion varies more rapidly, this condition quickly becomes tough to satisfy both in terms of computational complexity and in terms of memory.

To demonstrate this phenomenon, we present a numerical example. We set the domain to be the unit square, i.e. $\Omega = [0, 1] \times [0, 1]$, and let the diffusion coefficient be given by

$$A(x_1, x_2) = 102 + 100 \sin\left(\frac{2\pi x_1}{\varepsilon}\right) \sin\left(\frac{2\pi x_2}{\varepsilon}\right), \tag{2.4}$$

where ε denotes the scale at which the diffusion varies. An example of this coefficient when $\varepsilon = 2^{-4}$ is seen in Figure 1b. We compute a reference solution, denoted u_{ref}, on a fine mesh with mesh size $h = 2^{-8}$, and compute the energy error between the reference and finite element solution, u_h, for $h = 2^{-1}, 2^{-2}, \ldots, 2^{-7}$. This is done for different choices of $\varepsilon = 2^{-1}, 2^{-2}, \ldots, 2^{-5}$. The convergence plots are seen in Figure 1a. Here we can see how the error remains on a constant level until the mesh is able to resolve the variations defined by ε, at which point we start to see the linear convergence behavior.

3 Localized orthogonal decomposition (LOD)

For the purpose of solving PDEs with rapidly varying data, several multiscale methods have been developed. In this thesis the focus lies on the Localized Orthogonal Decomposition (LOD) method, first introduced in [35]. Widely speaking, the goal of the LOD method is to decompose the solution space into a coarse and a fine part. Here, the fine part is assumed to be refined enough to be able to resolve the variations in the data, while the coarse part is used for the main (cheap) computations.

We begin by deriving the method in its ideal case. However, there are computational difficulties arising in this method due to the resulting matrix systems not being sparse, but dense. For this purpose, a localized version that is computationally feasible is further presented. We finish this section by revisiting the numerical example from Section 2.

3.1 Ideal method

At first, assume the mesh size h to be fixed and sufficiently small, i.e. $h < \varepsilon$, so that the FE-space V_h can approximate the solution accurately. We define the space V_H similarly to V_h but with a larger mesh size $H > h$. Moreover, let $\mathcal{N}$ denote the set of interior nodes of $\mathcal{T}_H$, and $\{\lambda_x\}_{x\in\mathcal{N}}$ be the set of standard piecewise linear basis function that span V_H. Recall that computing a solution in V_H is cheap, but inaccurate as described in Section 2. The goal of LOD is to incorporate the finescale behavior of the diffusion into V_H to define a new so called multiscale space V_{ms}, which has the property that $\dim(V_{\mathrm{ms}}) = \dim(V_H)$ so that the computations are cheap, but with an error bound independent of ε.

For the construction of the multiscale space we require an interpolant $I_H : V_h \to V_H$ with the projection property $I_H \circ I_H = I_H$ that for all $T \in \mathcal{T}_H$ satisfies

$$H_T^{-1}\|v - I_H v\|_{L_2(T)} + \|\nabla I_H v\|_{L_2(T)} \leq C_I \|\nabla v\|_{L_2(N(T))}, \quad v \in V_h, \tag{3.1}$$

where $N(T) := \{T' \in \mathcal{T}_H : \overline{T'} \cap \overline{T} \neq \emptyset\}$ and $H_T := \operatorname{diam}(T)$. Furthermore, for a shape-regular and quasi-uniform partition, the estimate (3.1) can be summed into the global estimate

$$H^{-1}\|v - I_H v\|_{L_2(\Omega)} + \|\nabla I_H v\|_{L_2(\Omega)} \leq C_\gamma \|\nabla v\|_{L_2(\Omega)},$$

where C_γ depends on the interpolation constant C_I and the shape regularity

parameter defined as

$$\gamma := \max_{T \in \mathcal{T}_H} \gamma_T, \text{ where } \gamma_T = \frac{\text{diam}(B_T)}{H_T}.$$

Here B_T denotes the largest ball inside T. There are several choices of interpolants that can be used for the construction. A commonly used example, used for the numerical examples in both Paper I and Paper II, is $I_H = E_H \circ \Pi_H$, where Π_H is the piecewise L_2-projection onto $P_1(\mathcal{T}_H)$, the space of affine functions on each triangle $T \in \mathcal{T}_H$, and $E_H : P_1(\mathcal{T}_H) \to V_H$ is an averaging operator that, to each free node $x \in \mathcal{N}$, assigns the arithmetic mean of corresponding function values on intersecting elements, i.e.

$$(E_H(v))(x) = \frac{1}{\text{card}\{K \in \mathcal{T}_H : x \in \overline{K}\}} \sum_{K \in \mathcal{T}_H : x \in \overline{K}} v\big|_K(x).$$

For more discussion regarding possible choices of interpolants, see e.g. [13] or [39].

For any function $v \in H_0^1$, $I_H v$ describes the coarse part of the solution in the space V_H. The remainder part, $(1 - I_H)v$, contains fine-scale features of v that are not captured by the coarse space. These fine-scale functions configure the so-called fine-scale space, defined by the kernel of the interpolant, i.e.,

$$V_{\text{f}} := \ker(I_H) = \{v \in V_h : I_H v = 0\}.$$

That is, V_{f} consists of the finescale features of the solution which the FE-space is unable to capture. Consequently, this leads to the solution space being decomposed as

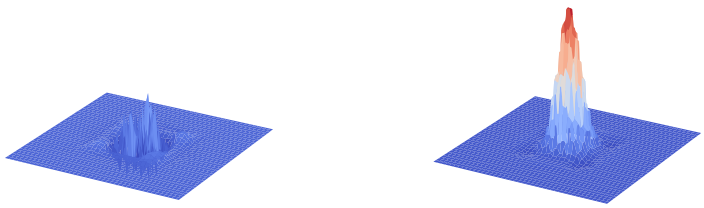
$$V_h = V_H \oplus V_{\text{f}},$$

so that every function $v \in V_h$ can be uniquely written as $v = v_H + v_{\text{f}}$ where $v_H \in V_H$ and $v_{\text{f}} \in V_{\text{f}}$. The LOD method is characterized by correcting coarse functions by appropriate projection into V_{f}. Thus, let $\mathcal{Q}_{\text{f}} : V_h \to V_{\text{f}}$ be the Ritz-projection onto V_{f}, i.e., $\mathcal{Q}_{\text{f}} v \in V_{\text{f}}$ satisfies

$$a(\mathcal{Q}_{\text{f}} v, w) = a(v, w), \quad \forall w \in V_{\text{f}}.$$

Using this Ritz-projection we create our multiscale space as

$$V_{\text{ms}} := V_H - \mathcal{Q}_{\text{f}} V_H.$$



(a) ϕ_x. **(b)** $\lambda_x - \phi_x$.

Figure 2: The basis correction ϕ_x (left) and corresponding modified basis function $\lambda_x - \phi_x$ (right) for a coarse node $x \in \mathcal{N}$ positioned at $(0.375, 0.500)$ in the unit square.

Note that for all $v_{\text{ms}} \in V_{\text{ms}}$ and $v_{\text{f}} \in V_{\text{f}}$ it holds that

$$a(v_{\text{ms}}, v_{\text{f}}) = a(v_H - \mathcal{Q}_{\text{f}} v_H, v_{\text{f}}) = a(v_H, v_{\text{f}}) - a(\mathcal{Q}_{\text{f}} v_H, v_{\text{f}}) = 0.$$

Hence, the Ritz-projection $\mathcal{Q}_{\text{f}}$ yields an orthogonal splitting with respect to the bilinear form $a(\cdot, \cdot)$ as

$$V_h = V_{\text{ms}} \oplus V_{\text{f}}.$$

Since V_{ms} is the orthogonal complement to V_{f} it holds that $\dim(V_{\text{ms}}) = \dim(V_H)$, but unlike V_H it moreover contains fine-scale features of the diffusion due to the Ritz-projection. For the construction of the multiscale space, we want to compute the projection for a fixed set of functions. Hence, we use the Ritz-projection to construct so-called basis correctors, $\phi_x := \mathcal{Q}_{\text{f}} \lambda_x \in V_{\text{f}}$ for each coarse node $x \in \mathcal{N}$ as solutions to the (global) corrector problem

$$a(\phi_x, w) = a(\lambda_x, w), \quad \forall w \in V_{\text{f}}. \tag{3.2}$$

The basis for V_{ms} is then given by $\{\lambda_x - \phi_x\}_{x \in \mathcal{N}}$, which can be viewed as a modified basis that holds information on the fine-scale behavior of the diffusion. An example of a computed basis corrector ϕ_x and its corresponding modified basis function, $\lambda_x - \phi_x$, is illustrated in Figure 2.

Given the multiscale space, the ideal LOD method reads: find $u_{\text{ms}} \in V_{\text{ms}}$ such that

$$a(u_{\text{ms}}, v) = (f, v), \quad \forall v \in V_{\text{ms}}. \tag{3.3}$$

In [35], the following theorem on an a priori error bound is derived for the

method.

Theorem 3.1. *Let u_h be the solution to* (2.3) *and u_{ms} the solution to* (3.3)*. Then the error is bounded by*

$$\|u_{\mathrm{ms}} - u_h\|_{H^1} \leq CH\|f\|,$$

where C is independent of the variations in A, but depends on the upper and lower bound of A.

Proof. Let $e := u_{\mathrm{ms}} - u_h$ and note that $e \in V_{\mathrm{f}}$ due to the orthogonal splitting. Hence it holds that $I_H e = 0$. Moreover, recall the Galerkin orthogonality $a(e, v_{\mathrm{ms}}) = 0$ for $v_{\mathrm{ms}} \in V_{\mathrm{ms}}$. We get

$$a(e,e) = -a(e,u_h) = -(f,e) \leq \|f\|\|e\| = \|f\|\|e - I_H e\| \leq CH\|f\|\|e\|_{H^1}.$$

The desired estimate now follows from the equivalence of norms between the H^1-norm and the norm induced by $a(\cdot,\cdot)$. □

The theorem states that the LOD method achieves convergence of optimal order, but unlike the standard FEM, the constant is independent of the derivatives of A. Although the method seems promising, it is in its current state based on the global projection (3.2) onto the fine-scale space V_{f}. That is, it is as expensive to solve for one basis corrector as solving the finite element problem (2.3) on the fine scale. Moreover, by definition, each basis corrector ϕ_x has a global support which consequently makes the linear system corresponding to (3.3) dense. To circumvent these issues, we wish to localize the computations onto coarse patches in order to obtain a sparse matrix system. This act of localization is justified by the fact that each basis correction ϕ_x decays exponentially fast away from its corresponding node $x \in \mathcal{N}$, as proven in [35].

3.2 Localized method

To localize the corrector problem, we begin by defining the patches to which the support of each basis function is to be restricted. For $\omega \subset \Omega$, let $N(\omega) := \{T \in \mathcal{T}_H : \overline{T} \cap \overline{\omega} \neq \emptyset\}$, and define a patch $N^k(\omega)$ of size k as

$$\begin{aligned} N^1(\omega) &:= N(\omega), \\ N^k(\omega) &:= N(N^{k-1}(\omega)), \quad \text{for } k \geq 2. \end{aligned}$$

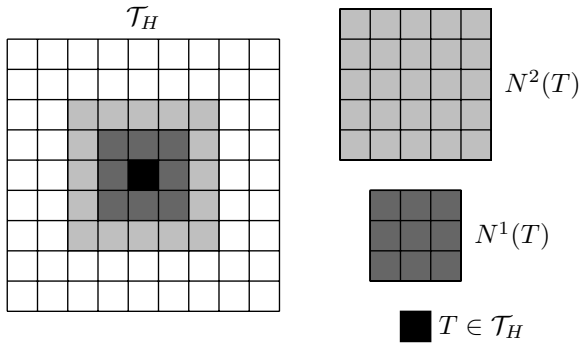


Figure 3: Illustration of patches based on an element $T \in \mathcal{T}_H$.

Given a coarse grid patch, we may restrict the finescale space V_{f} to it by defining

$$V_{\mathrm{f},k}^{\omega} := \{v \in V_{\mathrm{f}} : \mathrm{supp}(v) \subseteq N^k(\omega)\}.$$

In particular, we will commonly use $\omega = T \in \mathcal{T}_H$ and $\omega = x \in \mathcal{N}$ as subdomains. An example of how the patches spread across the grid with increasing k is illustrated in Figure 3.

We aim to localize the computation and support of our basis correctors, ϕ_x, by utilizing the newly defined coarse patches. For this purpose, define the element restricted Ritz-projection $\mathcal{Q}_{\mathrm{f}}^T$ such that $\mathcal{Q}_{\mathrm{f}}^T v \in V_{\mathrm{f}}$ is the solution to the system

$$a(\mathcal{Q}_{\mathrm{f}}^T v, w) = \int_T A\nabla v \cdot \nabla w \,\mathrm{d}x, \quad \forall w \in V_{\mathrm{f}}.$$

Note here that if we sum over all elements $T \in \mathcal{T}_H$ we get

$$a\Big(\sum_T \mathcal{Q}_{\mathrm{f}}^T v, w\Big) = \sum_T a(\mathcal{Q}_{\mathrm{f}}^T v, w) = \sum_T \int_T A\nabla v \cdot \nabla w \,\mathrm{d}x = a(v, w), \quad \forall w \in V_{\mathrm{f}}.$$

That is, the global Ritz-projection is constructed by the summation

$$\mathcal{Q}_{\mathrm{f}} v = \sum_{T \in \mathcal{T}_H} \mathcal{Q}_{\mathrm{f}}^T v.$$

For $k \in \mathbb{N}$, we may restrict the projection to a patch by letting $\mathcal{Q}_{\mathrm{f},k}^T : V_H \to V_{\mathrm{f},k}^T$ be such that $\mathcal{Q}_{\mathrm{f},k}^T v \in V_{\mathrm{f},k}^T$ solves

$$a(\mathcal{Q}_{\mathrm{f},k}^T v, w) = \int_T A\nabla v \cdot \nabla w \,\mathrm{d}x, \quad \forall w \in V_{\mathrm{f},k}^T.$$

By summation this yields the corresponding global version as

$$\mathcal{Q}_{\mathrm{f},k} v = \sum_{T \in \mathcal{T}_H} \mathcal{Q}_{\mathrm{f},k}^T v.$$

Finally, we may construct a localized multiscale space as $V_{\mathrm{ms},k} := V_H - \mathcal{Q}_{\mathrm{f},k} V_H$, spanned by $\{\lambda_x - \mathcal{Q}_{\mathrm{f},k}\lambda_x\}_{x \in \mathcal{N}}$.

We replace the multiscale space V_{ms} by its localized version $V_{\mathrm{ms},k}$ and obtain the localized LOD method that reads: find $u_{\mathrm{ms},k} \in V_{\mathrm{ms},k}$ such that

$$a(u_{\mathrm{ms},k}, v) = (f, v), \quad \forall v \in V_{\mathrm{ms},k}. \tag{3.4}$$

For a given element $T \in \mathcal{T}_H$, the dimension of $V_{\mathrm{f},k}^T$ is significantly smaller than that of V_{f}. Hence, the problem of finding $\mathcal{Q}_{\mathrm{f},k}\lambda_x$ is computationally cheaper than finding $\mathcal{Q}_{\mathrm{f}}\lambda_x$. Moreover, due to the restricted support of the basis correctors, the resulting linear system is also sparse (where the sparsity is determined by the size of corresponding grid patches). Another computational benefit is that the corrector problems are all independent and can be solved in parallel.

In [16], the following theorem is proved for the error bound.

Theorem 3.2. *Let u_h be the solution to* (2.3) *and $u_{\mathrm{ms},k}$ the solution to* (3.4)*. Then there exists $\xi \in (0,1)$ such that*

$$\|u_{\mathrm{ms},k} - u_h\|_{H^1} \leq C(H + k^{d/2}\xi^k)\|f\|,$$

where C is independent of the variations in A, but depends on the upper and lower bound of A.

The convergence is thus dependent on the choice of k. To achieve linear convergence for the method, k should be chosen proportional to $\log(1/H)$.

3.3 Numerical example

To demonstrate the performance of the LOD method we revisit the example from Section 2. Once again the domain is set to the unit square, and the same coefficient as defined in (2.4) is used, where the scale at which the values vary is set to $\varepsilon = 2^{-6}$. The fine mesh is once again set to $h = 2^{-8}$ so that it resolves the fine variations of the coefficient. We compute the localized solution $u_{\mathrm{ms},k}$, where $k = \log_2(1/H)$, for $H = 2^{-2}, \ldots, 2^{-6}$ and plot the energy error $|||u_{\mathrm{ms},k} - u_{\mathrm{ref}}|||$ where the reference solution u_{ref} is obtained using the finite

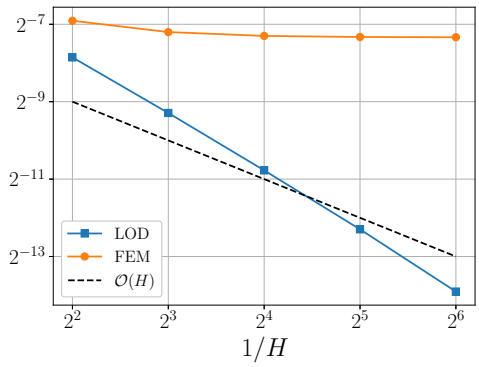


Figure 4: The energy error $|||u_{\mathrm{ms},k} - u_{\mathrm{ref}}|||$ (blue) for different coarse mesh sizes H. For comparison, the FEM-error $|||u_H - u_{\mathrm{ref}}|||$ (orange) is also plotted. The dashed line is an $\mathcal{O}(H)$-reference line.

element method on the fine mesh. For comparison, we also plot the error of the finite element solution based on the same mesh sizes. The error plot can be seen in Figure 4 and shows how the FEM-error remains on a constant level throughout all mesh sizes, while the LOD-error instantly decays according to the established theory.

4 LOD for time-dependent PDEs

The LOD method as stated so far is by now well-established for several types of equations, where the previous section showed its usefulness in the elliptic case. The method becomes even more advantageous with time-dependent equations, such as the parabolic case, since the multiscale space only needs to be computed once and can then be reused in all time steps. In this thesis, we further extend the LOD framework by considering equations that require the time-dependency to be taken into account in the construction of the multiscale method. As a first example, we consider the strongly damped wave equation that consists of two different multiscale coefficients. Here, the way the solution depends on the different coefficients is strongly affected by what phase of the time interval we consider. Hence the multiscale space must be constructed so that it is automatically adjusted for this effect over time. In Paper I, a GFEM that achieves this is derived and analyzed. The second case we consider is a parabolic equation, where we let the diffusion be time-dependent with rapid

variations in both time and space. The LOD method for parabolic equations was rigorously analyzed in [34], but is only applicable for high oscillations in spatial sense. An extension of the LOD method to the space- and time-dependent case is developed in Paper II.

4.1 Strongly damped wave equation

Consider the strongly damped wave equation

$$\ddot{u} - \nabla \cdot (A\nabla\dot{u} + B\nabla u) = f, \quad \text{in } \Omega \times (0,T], \tag{4.1}$$

$$u = 0, \quad \text{on } \partial\Omega \times (0,T], \tag{4.2}$$

$$u(0) = u_0, \quad \text{in } \Omega, \tag{4.3}$$

$$\dot{u}(0) = v_0 \quad \text{in } \Omega, \tag{4.4}$$

where $T > 0$ and Ω is a polygonal (or polyhedral) domain in $\mathbb{R}^d,\ d = 2, 3$. Here $A := A(x)$ represents the damping coefficient, $B := B(x)$ represents the wave propagation speed, $f := f(x,t)$ denotes the source function of the system, and the solution u is a displacement function. This equation is common in the modeling of viscoelastic materials, where the strong damping $-\nabla \cdot A\nabla\dot{u}$ appears when representing the stress as the sum of an elastic part and a viscous part [8, 15]. Viscoelastic materials have several applications in engineering, including noise dampening, vibration isolation, and shock absorption (see [24] for more applications). In multiscale applications, both A and B are rapidly varying. It is noteworthy that the solution is highly dependent on the damping A in the transient phase due to the time derivative, and that in the steady state phase it solely depends on the wave propagation speed B.

The strongly damped wave equation has been thoroughly analyzed in several areas recently. For instance, well-posedness of the equation is discussed in [9, 23, 25], asymptotic behavior in [10, 5, 36], solution blowup in [14, 4], and decay estimates in [22]. In particular, the FEM for the strongly damped wave equation has been analyzed in [28] using the Ritz–Volterra projection, and [27] uses the classical Ritz-projection in the homogeneous case with Rayleigh damping.

We begin by considering the finite element method that corresponds to the system (4.1)-(4.4). For the spatial discretization of the problem, let V_h be defined as in Section 2. The semi-discrete FEM becomes: find $u_h(t) \in V_h$ such that

$$(\ddot{u}_h, v) + a(\dot{u}_h, v) + b(u_h, v) = (f, v), \quad \forall v \in V_h,\ t > 0,$$

with initial values $u_h(0) = u_{h,0}$ and $\dot{u}_h = v_{h,0}$, where $u_{h,0}, v_{h,0} \in V_h$ are appropriate approximations of u_0 and v_0 respectively. Here, the bilinear forms

are defined as $a(\cdot,\cdot) := (A\nabla\cdot, \nabla\cdot)$ and $b(\cdot,\cdot) := (B\nabla\cdot, \nabla\cdot)$.

For the temporal discretization, let $0 =: t_0 < t_1 < ... < t_N := T$ be a partition with uniform time step $\tau := t_n - t_{n-1}$. By applying a backward Euler scheme, the fully discrete system reads: find $u_h^n \in V_h$ such that

$$(\bar{\partial}_t^2 u_h^n, v) + a(\bar{\partial}_t u_h^n, v) + b(u_h^n, v) = (f^n, v), \quad \forall v \in V_h, \tag{4.5}$$

for $n \geq 2$. Here, the discrete derivative is defined as $\bar{\partial}_t u_h^n = (u_h^n - u_h^{n-1})/\tau$. The first initial value is given by $u_h^0 \in V_h$. The second initial value u_h^1 should be an approximation of $u(t_1)$ and could be chosen as $u_h^1 = u_h^0 + \tau v_h^0$. For results on regularity and error estimates, we refer to [27]. However, although convergence of optimal order is proven here, the involved constants are dependent on the variations in the data, and hence not applicable in the multiscale case.

For the development of our GFEM for the strongly damped wave equation, we begin by defining the coarse FE-space V_H and the finescale space $V_{\mathrm{f}} := \ker(I_H)$ in complete analogy to Section 3. For the standard LOD method (as used in the elliptic case [35, 18, 16], parabolic case [34, 33], and for the wave equation [2, 32]), the definition of the Ritz-projection is based on solely the diffusion coefficient. However, since we have two different multiscale coefficients to incorporate in this case, we define the Ritz-projection $R_{\mathrm{f}} := V_H \to V_{\mathrm{f}}$ by

$$a(R_{\mathrm{f}} v, w) + \tau b(R_{\mathrm{f}} v, w) = a(v, w) + \tau b(v, w), \quad \forall w \in V_{\mathrm{f}}.$$

Here, the operator R_{f} takes the role of $\mathcal{Q}_{\mathrm{f}}$ from Section 3, so that the notation stays in line with that of Paper I. The particular choice of scalar product, $a(\cdot,\cdot) + \tau b(\cdot,\cdot)$, comes from the backward Euler scheme. Using this projection, we may define the multiscale space $V_{\mathrm{ms}} := V_H - R_{\mathrm{f}} V_H$ such that

$$V_h = V_{\mathrm{ms}} \oplus V_{\mathrm{f}}, \quad \text{and} \quad a(v_{\mathrm{ms}}, v_{\mathrm{f}}) + \tau b(v_{\mathrm{ms}}, v_{\mathrm{f}}) = 0. \tag{4.6}$$

The basis functions for this multiscale space is, similarly to the standard LOD method, defined by applying the Ritz-projection to the finite element basis function. That is, $R_{\mathrm{f}}\lambda_x \in V_{\mathrm{f}}$ solves the global corrector problem

$$a(R_{\mathrm{f}}\lambda_x, w) + \tau b(R_{\mathrm{f}}\lambda_x, w) = a(\lambda_x, w) + \tau b(\lambda_x, w), \quad \forall w \in V_{\mathrm{f}}. \tag{4.7}$$

We can now construct our basis for V_{ms} as $\{\lambda_x - R_{\mathrm{f}}\lambda_x\}_{x\in\mathcal{N}}$, where $R_{\mathrm{f}}\lambda_x$ contains fine-scale information on both of the multiscale coefficients.

We may now formulate our ideal method. Since the solution space can be decomposed as $V_h = V_{\mathrm{ms}} \oplus V_{\mathrm{f}}$, the idea is to solve a coarse scale problem

in V_{ms}, and then add additional corrections from a problem on the fine scale to account for the time-dependency of the problem. The method reads: find $u_{\mathrm{lod}}^n = v^n + w^n$, with $v^n \in V_{\mathrm{ms}}$ and $w^n \in V_{\mathrm{f}}$ such that

$$\tau(\bar{\partial}_t^2 v^n, z) + a(v^n, z) + \tau b(v^n, z) = \tau(f^n, z) + a(u_{\mathrm{lod}}^{n-1}, z), \quad \forall z \in V_{\mathrm{ms}}, \tag{4.8}$$

$$a(w^n, z) + \tau b(w^n, z) = a(u_{\mathrm{lod}}^{n-1}, z), \quad \forall z \in V_{\mathrm{f}}, \tag{4.9}$$

for $n \geq 2$ with initial data $u_{\mathrm{lod}}^0 = u_h^0 \in V_{\mathrm{ms}}$ and $u_{\mathrm{lod}}^1 = u_h^1 \in V_{\mathrm{ms}}$. The initial data is chosen in V_{ms} to simplify the implementation of the finescale correctors. This choice does not affect the performance of the proposed method, as shown in Paper I.

The ideal method as currently stated is defined on the entire fine grid which is not computationally feasible. For this purpose we construct a localized multiscale space $V_{\mathrm{ms},k} := V_H - R_{\mathrm{f},k} V_H$ in complete analogy to the localization procedure in Section 3. For this localization to be valid it is required that each basis corrector $R_{\mathrm{f}}\lambda_x$ satisfies the same exponential decay as $\mathcal{Q}_{\mathrm{f}}\lambda_x$ from Section 3. However, we quickly note that $R_{\mathrm{f}}\lambda_x$ solves the same type of problem as $\mathcal{Q}_{\mathrm{f}}\lambda_x$ but with diffusion $A + \tau B$, and hence satisfies the required decay property by classical LOD theory.

With the space $V_{\mathrm{ms},k}$ defined, we are able to localize the computations corresponding to the system (4.8) by replacing the multiscale space by its localized counterpart. It remains to localize the computations of the finescale system in (4.9), which equivalently can be written as

$$a(\bar{\partial}_t w^n, z) + b(w^n, z) = \frac{1}{\tau} a(v^{n-1}, z).$$

We replace the right-hand side by its localized version $v_k^{n-1} \in V_{\mathrm{ms},k}$ and note that $v_k^{n-1} = \sum_{x \in \mathcal{N}} \alpha_x^{n-1}(\lambda_x - R_{\mathrm{f},k}\lambda_x)$. Thus, we seek our localized finescale solution as $w_k^n = \sum_{x \in \mathcal{N}} w_{k,x}^n$, where $w_{k,x}^n \in V_{\mathrm{f},k}^x$ solves

$$a(\bar{\partial}_t w_{k,x}^n, z) + b(w_{k,x}^n, z) = \frac{1}{\tau} a(\alpha_x^{n-1}(\lambda_x - R_{\mathrm{f},k}\lambda_x), z), \quad \forall z \in V_{\mathrm{f},k}^x, \tag{4.10}$$

so that the computation of this equation is localized to a patch surrounding the node $x \in \mathcal{N}$. We introduce the functions $\xi_{k,x}^l \in V_{\mathrm{f},k}^x$ as solution to the parabolic equation

$$a(\bar{\partial}_t \xi_{k,x}^l, z) + b(\xi_{k,x}^l, z) = a(\frac{1}{\tau}\chi_1(l)(\lambda_x - R_{\mathrm{f},k}\lambda_x), z), \quad \forall z \in V_{\mathrm{f},k}^x, \tag{4.11}$$

for $l = 1, 2, ..., N$ with initial value $\xi_{k,x}^0 = 0$, and where $\chi_1(l)$ is an indicator function that equals 1 when $l = 1$ and 0 otherwise. It holds that $w_{k,x}^n =$

$\sum_{l=1}^{n} \alpha_x^{n-l} \xi_{k,x}^l$ is the solution to (4.10), which can be shown by simply inserting it and canceling terms.

The localized GFEM then reads: find $u_{\mathrm{lod},k}^n = v_k^n + w_k^n$, where $v_k^n = \sum_{x \in \mathcal{N}} \alpha_x^n (\lambda_x - R_{\mathrm{f},k}\lambda_x) \in V_{\mathrm{ms},k}$ solves

$$\tau(\bar{\partial}_t^2 v_k^n, z) + a(v_k^n, z) + \tau b(v_k^n, z) = \tau(f^n, z) + a(u_{\mathrm{lod},k}^{n-1}, z), \quad \forall z \in V_{\mathrm{ms},k}, \tag{4.12}$$

and $w_k^n = \sum_{x \in \mathcal{N}} \sum_{l=1}^{n} \alpha_x^{n-l} \xi_{k,x}^l$, where $\xi_{k,x}^l \in V_{\mathrm{f},k}^x$ solves (4.11).

To justify the fact that we localize the finescale equation, we furthermore require that the functions $\{\xi_x^l\}_{l=1}^N$ satisfy an exponential decay as well. The following theorem, proven in Paper I, provides this requirement.

Theorem 4.1. *For any node $x \in \mathcal{N}$, let $\xi_x^n \in V_{\mathrm{f}}$ be the solution to*

$$a(\bar{\partial}_t \xi_x^n, z) + b(\xi_x^n, z) = a(\frac{1}{\tau}\chi_1(n)(\lambda_x - R_{\mathrm{f}}\lambda_x), z), \quad \forall z \in V_{\mathrm{f}},$$

with initial value $\xi_x^0 = 0$. Then there exist constants $c > 0$ and $C > 0$ such that for any $k \geq 1$

$$\|\xi_x^n\|_{H^1(\Omega \setminus N^k(x))} \leq Ce^{-ck}\|\lambda_x\|_{H^1},$$

for sufficiently small time step τ.

For the error analysis of this method the solution is first decomposed as $u_{\mathrm{lod}}^n = u_{\mathrm{lod},1}^n + u_{\mathrm{lod},2}^n$, where $u_{\mathrm{lod},1}^n$ has zero initial data and $u_{\mathrm{lod},2}^n$ has zero source data. The error of each decomposed part is then analyzed separately. In each case, the error is split as

$$\|u_{\mathrm{lod},i}^n - u_{h,i}^n\|_{H^1} \leq \|u_{\mathrm{lod},i}^n - X^n\|_{H^1} + \|X^n - u_{h,i}^n\|_{H^1} = \|\theta^n\|_{H^1} + \|\rho^n\|_{H^1},$$

where X^n is the solution to an auxiliary problem. We state the final result here and refer to Paper I where the whole analysis is investigated thoroughly.

Theorem 4.2. *Let u_h^n and u_{lod}^n be the solutions to* (4.5) *and* (4.8)-(4.9)*, respectively. The solutions can be split into $u_h^n = u_{h,1}^n + u_{h,2}^n$ and $u_{\mathrm{lod}}^n = u_{\mathrm{lod},1}^n + u_{\mathrm{lod},2}^n$, where the first part has vanishing initial data, and the second part a vanishing right hand side. The errors are bounded by*

$$\sum_{j=2}^{n} \tau \|u_{h,1}^j - u_{\mathrm{lod},1}^j\|_{H^1}^2 \leq CH^2 \left(\sum_{j=1}^{n} \tau(\|f^j\|^2 + \|\bar{\partial}_t f^j\|^2) + \max_{j=1,\ldots,n} \|f^j\|^2 \right),$$

$$\sum_{j=2}^{n} \tau t_j^2 \|u_{h,2}^j - u_{\mathrm{lod},2}^j\|_{H^1}^2 \leq CH^2(\|\bar{\partial}_t u_h^1\|_{H^1}^2 + \|u_h^1\|_{H^1}^2 + \|u_h^0\|_{H^1}^2),$$

for $n \geq 2$, where the constants are independent of variations in A and B, but depend on the upper and lower bounds of A and B.

4.2 Parabolic equation with time-dependent diffusion

The second case we consider for the extension of the LOD framework is the parabolic equation of the form

$$\dot{u} - \nabla \cdot (A\nabla u) = f, \quad \text{in } \Omega \times (0,T], \tag{4.13}$$

$$u = 0, \quad \text{on } \partial\Omega \times (0,T], \tag{4.14}$$

$$u(0) = 0, \quad \text{in } \Omega, \tag{4.15}$$

where $T > 0$ and Ω is a polygonal (or polyhedral) domain in $\mathbb{R}^d,\ d = 2, 3$. In contrast to the parabolic case dealt with in [34], we have $A := A(t, x)$ with rapid oscillations in both time and space. We remark that the choice of zero initial data is made to simplify the presentation of our proposed method, and that nonzero data can be considered with just a few alterations. The parabolic equation (4.13) appears in several real life applications, where common examples include heat transfer and modeling of pressure in compressible flow [40, 29, 21]. In particular, the time-dependency in the diffusion is highly relevant when considering a heat conductor undertaking radioactive decay [42].

We begin by deriving the weak formulation corresponding to (4.13), and its corresponding finite element formulation. Let $L^2(0,T;\mathcal{B})$ and $H^1(0,T;\mathcal{B})$ be the standard Bochner spaces with norm

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

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

where $\mathcal{B}$ is a Banach space with norm $\|\cdot\|_{\mathcal{B}}$. Throughout this section, we abbreviate the Bochner spaces by omitting the interval and the domain and write, e.g., $L^2(H_0^1) := L^2(0,T;H_0^1(\Omega))$. We consider the following weak space-time formulation: find $u \in V_{\mathrm{tr}} := L^2(H_0^1) \cap H^1(H^{-1})$ such that

$$\int_0^T \langle \dot{u}, v \rangle + a(t;u,v) \,\mathrm{d}t = \int_0^T \langle f, v \rangle \,\mathrm{d}t \tag{4.16}$$

for all $v \in V_{\mathrm{te}} := L^2(H_0^1)$. Here, we have denoted by $\langle \cdot, \cdot \rangle$ the dual pairing of $H^{-1}(\Omega)$ and $H_0^1(\Omega)$. Moreover, the bilinear form $a(t;\cdot,\cdot)\colon H_0^1(\Omega) \times H_0^1(\Omega) \to \mathbb{R}$

is defined by

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

for almost all $t \in (0, T)$. From here on, we omit the t-dependence and abbreviate $a(\cdot, \cdot) := a(t; \cdot, \cdot)$.

We begin by introducing the space-time discretization on the fine scale. Let V_h be defined as in Section 2 for the spatial part. For the temporal discretization, we introduce the fine time step τ and set $t_i = i\tau$, $i = 1, \dots, N_\tau$ as the uniform partition with $t_{N_\tau} = T$. Denote by $\mathcal{I}_\tau$ the decomposition of $[0, T]$ into sub-intervals $[t_{i-1}, t_i]$, $i = 1, \dots, N_\tau$. With respect to this discretization, we introduce two discrete spaces, $\hat{V}_\tau$ and V_τ, as the temporal trial and test space respectively, defined by

$$\hat{V}_\tau := \{v \in H^1(0, T) : \, v|_I, \, I \in \mathcal{I}_\tau, \text{ is a polynomial of degree } \leq 1 \text{ and } v(0) = 0\},$$
$$V_\tau := \{v \in L^2(0, T) : \, v|_I, \, I \in \mathcal{I}_\tau, \text{ is constant}\}.$$

Based on the above definitions of spatial and temporal spaces, we introduce corresponding tensor-product space-time finite element spaces with respect to the full domain $[0, T] \times \Omega$. We set

$$\hat{V}_{h,\tau} := \hat{V}_\tau \times V_h, \qquad V_{h,\tau} := V_\tau \times V_h$$

as trial and test space respectively. The finite element formulation of (4.16) now states: find $u_{h,\tau} \in \hat{V}_{h,\tau}$ such that

$$\int_0^T \langle \dot{u}_{h,\tau}, v_{h,\tau} \rangle + a(u_{h,\tau}, v_{h,\tau}) \, \mathrm{d}t = \int_0^T \langle f, v_{h,\tau} \rangle \, \mathrm{d}t \tag{4.17}$$

for all $v_{h,\tau} \in V_{h,\tau}$. Note that, for the parabolic problem, we base our finite element problem on a Petrov–Galerkin ansatz, i.e., the trial and test spaces do not coincide.

For our space-time multiscale method we will, in similarity to previous LOD based methods, introduce corresponding coarse domains. For the spatial discretization, define V_H and the interpolant $I_H = E_H \circ \Pi_H$ as in Section 3, and let $\{\varphi_x\}_{x \in \mathcal{N}_H}$ denote the standard finite element basis functions for V_H. For the temporal discretization, let $\mathcal{T} > 0$ be a coarse time step, and let $T_i = i\mathcal{T}$, $i = 0, 1, \dots, N$. Denote by $\mathcal{I}_\mathcal{T}$ the decomposition of the time interval $[0, T]$ into sub-intervals $[T_{i-1}, T_i]$, $i = 1, \dots, N$ of uniform size with $T_N = T$. Then, in analogy with the fine temporal spaces, we define $\hat{V}_\mathcal{T}$ and $V_\mathcal{T}$ as coarse temporal trial and test spaces based on the decomposition $\mathcal{I}_\mathcal{T}$. Moreover let $\{\zeta_i\}_{i=1}^N$ denote the piecewise linear basis functions that span $\hat{V}_\mathcal{T}$, and $\{\chi_i\}_{i=1}^N$

the piecewise constant basis functions that span $V_\mathcal{T}$, where $\chi_i = \mathbb{1}_{[T_{i-1},T_i]}$. The coarse space-time spaces are then set as $\hat{V}_{H,\mathcal{T}} := \hat{V}_\mathcal{T} \times V_H$ as trial space, and $V_{H,\mathcal{T}} := V_\mathcal{T} \times V_H$ as test space.

Next, use the spatial interpolant to define the remainder spaces

$$\hat{W}_{h,\tau} := \Big\{ w \in \hat{V}_{h,\tau} : \ I_H w(T_i, \cdot) = 0 \text{ for all } i = 0, \dots, N \Big\},$$
$$W_{h,\tau} := \Big\{ w \in V_{h,\tau} : \ \mathcal{T}^{-1} \int_{T_{i-1}}^{T_i} I_H w \, \mathrm{d}t = 0 \text{ for all } i = 1, \dots, N \Big\},$$

which we will use as our fine trial and test spaces respectively. Note that, by construction, $\hat{V}_{h,\tau} = \hat{V}_{H,\mathcal{T}} \oplus \hat{W}_{h,\tau}$ and $V_{h,\tau} = V_{H,\mathcal{T}} \oplus W_{h,\tau}$.

Our proposed space-time multiscale method is based on the Variational Multiscale Method, first introduced in [20], and aims to extend the LOD framework to space- and time-dependent coefficients. The main idea is to decompose the solution into a coarse part in $\hat{V}_{H,\mathcal{T}}$ and a remainder part in $\hat{W}_{h,\tau}$ and then consider (4.16) for test functions in the coarse test space $V_{H,\mathcal{T}}$ and the fine test space $W_{h,\tau}$ separately. In turn, this yields a coarse-scale and a fine-scale equation respectively. The main purpose of the fine-scale equation is to compute certain *correctors*, which we utilize to enrich the coarse-scale equation with the fine-scale behavior of the diffusion. The proposed method reads: find $\tilde{u}_{H,\mathcal{T}} = u_{H,\mathcal{T}} + \mathcal{Q}u_{H,\mathcal{T}} \in (1 + \mathcal{Q})\hat{V}_{H,\mathcal{T}}$ such that $u_{H,\mathcal{T}} \in \hat{V}_{H,\mathcal{T}}$ solves

$$\int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}(1 + \mathcal{Q})u_{H,\mathcal{T}}, v_{H,\mathcal{T}} \rangle + a((1 + \mathcal{Q})u_{H,\mathcal{T}}, v_{H,\mathcal{T}}) \, \mathrm{d}t = \int_0^T \langle f, v_{H,\mathcal{T}} \rangle \, \mathrm{d}t, \quad (4.18)$$

for all $v_{H,\mathcal{T}} \in V_{H,\mathcal{T}}$ and $\mathcal{Q}u_{H,\mathcal{T}} \in \hat{W}_{h,\tau}$ solves

$$\int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{Q}u_{H,\mathcal{T}}, w_{h,\tau} \rangle + a(\mathcal{Q}u_{H,\mathcal{T}}, w_{h,\tau}) \, \mathrm{d}t = -\int_0^T \langle \dot{u}_{H,\mathcal{T}}, w_{h,\tau} \rangle + a(u_{H,\mathcal{T}}, w_{h,\tau}) \, \mathrm{d}t \quad (4.19)$$

for all $w_{h,\tau} \in W_{h,\tau}$. Due to linearity, we may take (4.19) and further decompose it into corrector problems with local space-time basis functions $\Lambda_x^j := \zeta_j \varphi_x \in \hat{V}_{H,\mathcal{T}}$ as source, where $j \in \{1, \dots, N\}$ denotes the temporal node and $x \in \mathcal{N}_H$ the spatial node that the basis function corresponds to. That is, we define a corrector $\mathcal{Q}\Lambda_x^j \in \hat{W}_{h,\tau}$ as the solution to

$$\int_0^T \langle \tfrac{\mathrm{d}}{\mathrm{d}t}\mathcal{Q}\Lambda_x^j, w_{h,\tau} \rangle + a(\mathcal{Q}\Lambda_x^j, w_{h,\tau}) \, \mathrm{d}t = -\int_0^T \langle \dot{\Lambda}_x^j, w_{h,\tau} \rangle + a(\Lambda_x^j, w_{h,\tau}) \, \mathrm{d}t \quad (4.20)$$

for all $w_{h,\tau} \in W_{h,\tau}$, with initial condition $\mathcal{Q}\Lambda_x^j(\cdot,0) = 0$. Note that, due to $\text{supp}(\Lambda_x^j) = [T_{j-1}, T_{j+1}] \times N(x)$ and the initial condition in (4.20), the integrals are immediately reduced from $[0,T]$ to $[T_{j-1},T]$. Without loss of generality, we will therefore restrict the presentation to the case $j = 1$ and abbreviate $\Lambda = \Lambda_x^1$, $\zeta = \zeta_1$, and $\varphi = \varphi_x$.

We wish to solve the system (4.20) without explicitly computing the fine-scale spaces $\hat{W}_{h,\tau}$ and $W_{h,\tau}$. Hence, we reformulate (4.20) as a constraint problem posed in the full discrete space $\hat{V}_{h,\tau}$ and with test functions in $V_{h,\tau}$. That is, let $\psi \in \hat{V}_{h,\tau}$ be the solution of

$$\int_0^T \langle \dot{\psi}, v\rangle + a(\psi, v)\,\mathrm{d}t + \sum_{i=1}^{N} \int_{T_{i-1}}^{T_i} \langle \lambda_i, I_H v\rangle\,\mathrm{d}t = -\int_0^T \langle \dot{\Lambda}, v\rangle + a(\Lambda, v)\,\mathrm{d}t \tag{4.21a}$$

$$\sum_{j=1}^{N_\mathcal{T}} \langle I_H \psi(T_j), \mu_j\rangle = 0, \tag{4.21b}$$

for all $v \in V_{h,\tau}$, $\mu_j \in V_H$, $j = 1,\dots,N$, where $(\lambda_1,\dots,\lambda_N) \in V_H \times \dots \times V_H$ are the associated Lagrange multipliers. Note that, by construction, (4.20) and (4.21) are equivalent, i.e. $\mathcal{Q}\Lambda = \psi$.

The method presented in (4.18)-(4.19) is referred to as our ideal method. For the corresponding error estimate, we introduce the norms for the trial and test space respectively as

$$\|v\|_{V_{\text{tr}}}^2 := \int_0^T \|\nabla \overline{v}(t,\cdot)\|_{L^2(\Omega)}^2 + \|\dot{v}(t,\cdot)\|_{H^{-1}(\Omega)}^2\,\mathrm{d}t + \|v(T)\|_{L^2(\Omega)},$$
$$\|v\|_{V_{\text{te}}}^2 := \int_0^T \|\nabla v(t,\cdot)\|_{L^2(\Omega)}^2\,\mathrm{d}t,$$

where $\overline{v} := \sum_{i=1}^{N_\tau} \left(\tau^{-1}\int_{t_{i-1}}^{t_i} v(s,\cdot)\,\mathrm{d}s\right)\chi_i$ is the mean with respect to the fine temporal discretization. These norms are essential for the analysis of certain space-time Petrov–Galerkin discretizations as in [41] on which our approach relies. The error for the proposed ideal method is then quantified by following theorem, which is proven in Paper II.

Theorem 4.3 (Error of the ideal method)**.** *Assume that the right-hand side fulfills $f \in L^2(L^2) \cap H^1(H^{-1})$. Then the error between the solutions $u_{h,\tau}$ and $\tilde{u}_{H,\mathcal{T}}$ satisfies*

$$\|\tilde{u}_{H,\mathcal{T}} - u_{h,\tau}\|_{V_{\text{tr}}} \le C\,(H + \mathcal{T})\,\|f\|_{L^2(L^2)\cap H^1(H^{-1})}. \tag{4.22}$$

This theorem states that our novel method converges with optimal order. However, similarly to earlier LOD based methods, the formulation as currently stated is ideal, but impractical. The method is based on auxiliary corrector problems defined on the entire fine space-time grid, which is not computationally feasible, as earlier discussed. To circumvent this issue, one observes that a corrector function ψ decays exponentially fast away from the support of the underlying basis function Λ. Without a great impact on the approximation property, it is therefore possible to restrict the fine-scale computations to local spatial patches as for the standard LOD, and to a limited number of coarse time steps, yielding a localized method in both space and time. For the temporal localization, we begin by demonstrating how the basis corrector ψ is computed by a sequential approach. This turns into an efficient scheme whose computations easily can be restricted in time.

To define the sequential approach, we begin by dividing the integral in (4.21) into local integrals over $[T_{j-1}, T_j]$, $j = 1, \dots, N$ and define for given j the local version of $\hat{V}_{h,\tau}$ by

$$\hat{V}_{h,\tau}^j := \{v|_{[T_{j-1},T_j]\times D} : \, v \in V_{h,\tau}\}.$$

Further, we denote with $\xi_j \in \hat{V}_{h,\tau}^j$, $\xi_j(T_{j-1}) = 0$ the solution of

$$\begin{aligned}\int_{T_{j-1}}^{T_j} \langle \dot{\xi}_j, v\rangle + a(\xi_j, v) + \langle \lambda_j, I_H v\rangle \,\mathrm{d}t = &- \int_{T_{j-1}}^{T_j} \langle \dot{\Lambda}, v\rangle + a(\Lambda, v) \\ &- \langle \tfrac{1}{T}\xi_{j-1}(T_{j-1}), v\rangle \\ &+ a(\tfrac{T_j - t}{T}\xi_{j-1}(T_{j-1}), v)\,\mathrm{d}t, \end{aligned} \tag{4.23a}$$

$$\langle I_H \xi_j(T_j), \mu\rangle = 0, \tag{4.23b}$$

for all $v \in V_{h,\tau}$, $\mu \in V_H$, where $\lambda_j \in V_H$ is the associated Lagrange multiplier. For $j = 1$, we explicitly set $\xi_0(T_0) = 0$ such that the third and the fourth term on the right-hand side of (4.23a) vanish. Note that the functions $\{\xi_j\}_{j=1}^N$ are constructed in a way such that

$$\psi = \sum_{j=1}^{N} \big(\xi_j + \tfrac{T_j - t}{T}\xi_{j-1}(T_{j-1})\big)\mathbb{1}_{[T_{j-1},T_j]}.$$

We emphasize that the basis function Λ only has support on the first two coarse intervals. That is, for $j > 2$ the first two terms in (4.23a) (and also in (4.21a)) disappear and, consequently, ψ will begin to decay due to the parabolic nature of the problem.

Due to the decay property of ψ, there will be an $\ell \in \mathbb{N}$ such that for $j > \ell$, the sequential functions ξ_j will be of negligible size compared to the error of the ideal method. Hence, it suffices to restrict the computations to $\{\xi_j\}_{j=1}^{\ell}$. That is, we define our temporally localized corrector function by $\psi_\ell = \psi \mathbb{1}_{[0,T_\ell]}$, where we refer to ℓ as the *temporal localization parameter*. We remark that simply restricting the computations will make ψ_ℓ discontinuous, and thus it will no longer be a function in $\hat{V}_{h,\tau}$. However, this can easily be circumvented by choosing

$$\psi_\ell = \psi \mathbb{1}_{[0,T_\ell]} + \tfrac{T_{\ell+1}-t}{\mathcal{T}} \xi_\ell(T_\ell) \mathbb{1}_{[T_\ell,T_{\ell+1}]}.$$

It remains to apply the spatial localization procedure, which follows in similarity to Section 3. By combining this temporal localization procedure with the spatial one from Section 3, we yield an efficient space-time multiscale method for which the convergence rate from Theorem 4.3 remains valid. Details on the decay of ψ, and the performance of the localized method are examined numerically in Paper II.

5 Summary of papers

Paper I. In Paper I we propose and analyze the GFEM in (4.8)-(4.9), based on the LOD method, for strongly damped wave equations with rapidly varying data. The method is designed to handle independent variations in both the damping coefficient and wave propagation speed respectively. It does so by correcting for the damping in the transient phase, where it is as most effective, and automatically transitions into correcting for the wave propagation speed in the steady state phase, where the damping has vanished. Convergence of optimal order is proven for the ideal method, as well as the exponential decay of the basis corrector functions to justify the localization. Numerical examples are presented that confirm the theoretical findings.

Paper II. In Paper II we present the GFEM (4.18)-(4.19) extending the LOD framework to parabolic equations where the diffusion is highly oscillating in both time and space. The method computes a coarse-scale representation of the differential operator that contains information on the space-time variations in the diffusion. Once the coarse-scale representation is constructed, it can furthermore be reused to solve the system for multiple right-hand sides. Convergence of optimal order is proven for the ideal method. We illustrate the space-time decay of the basis correctors, which is necessary for the localized scheme. Numerical examples that illustrate the error convergence and the performance of the localized method are presented.

Bibliography

[1] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numer.*, 21:1–87, 2012.

[2] A. Abdulle and P. Henning. Localized orthogonal decomposition method for the wave equation with a continuum of scales. *Math. Comp.*, 86(304):549–587, 2017.

[3] R. Altmann, E. Chung, R. Maier, D. Peterseim, and S.-M. Pun. Computational multiscale methods for linear heterogeneous poroelasticity. *J. Comput. Math.*, 38(1):41–57, 2020.

[4] G. Avalos and I. Lasiecka. Optimal blowup rates for the minimal energy null control of the strongly damped abstract wave equation. *Ann. Sc. Norm. Super. Pisa Cl. Sci. (5)*, 2(3):601–616, 2003.

[5] J. Azevedo, C. Cuevas, and H. Soto. Qualitative theory for strongly damped wave equations. *Math. Method Appl. Sci.*, 40, 08 2017.

[6] I. Babuška, G. Caloz, and J. E. Osborn. Special finite element methods for a class of second order elliptic problems with rough coefficients. *SIAM J. Numer. Anal.*, 31(4):945–981, 1994.

[7] I. Babuška and J. E. Osborn. Generalized finite element methods: their performance and their relation to mixed methods. *SIAM J. Numer. Anal.*, 20(3):510–536, 1983.

[8] E. Bonetti, E. Rocca, R. Scala, and G. Schimperna. On the strongly damped wave equation with constraint. *Commun. Partial. Differ. Equ.*, 42(7):1042–1064, 2017.

[9] A. Carvalho and J. Cholewa. Local well posedness for strongly damped wave equations with critical nonlinearities. *Bull. Aust. Math. Soc.*, 66, 12 2002.

[10] C. Cuevas, C. Lizama, and H. Soto. Asymptotic periodicity for strongly damped wave equations. *Abstr. Appl. Anal.*, 2013, 09 2013.

[11] W. E and B. Engquist. The heterogeneous multiscale methods. *Commun. Math. Sci.*, 1(1):87–132, 2003.

[12] Y. Efendiev, J. Galvis, and T. Y. Hou. Generalized multiscale finite element methods (GMsFEM). *J. Comput. Phys.*, 251:116–135, 2013.

[13] C. Engwer, P. Henning, A. Målqvist, and D. Peterseim. Efficient implementation of the localized orthogonal decomposition method. *Comput. Methods Appl. Mech. Eng.*, 350:123–153, 06 2019.

[14] F. Gazzola and M. Squassina. Global solutions and finite time blow up for damped semilinear wave equations. *Ann. Inst. H. Poincaré Anal. Non Linéaire*, 23(2):185–207, 2006.

[15] P. J. Graber and J. L. Shomberg. Attractors for strongly damped wave equations with nonlinear hyperbolic dynamic boundary conditions. *Nonlinearity*, 29(4):1171, 2016.

[16] P. Henning and A. Målqvist. Localized orthogonal decomposition techniques for boundary value problems. *SIAM J. Sci. Comput.*, 36(4):A1609–A1634, 2014.

[17] P. Henning, A. Målqvist, and D. Peterseim. A localized orthogonal decomposition method for semi-linear elliptic problems. *ESAIM Math. Model. Numer. Anal.*, 48(5):1331–1349, 2014.

[18] P. Henning and D. Peterseim. Oversampling for the multiscale finite element method. *Multiscale Model. Simul.*, 11(4):1149–1175, 2013.

[19] T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.*, 134(1):169–189, 1997.

[20] T. J. Hughes, G. R. Feijóo, L. Mazzei, and J.-B. Quincy. The variational multiscale method—a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Eng.*, 166(1-2):3–24, 1998.

[21] D. Hào, N. Duc, and N. Thang. Backward semi-linear parabolic equations with time-dependent coefficients and local lipschitz source. *Inverse Problems*, 34:055010, 05 2018.

[22] R. Ikehata. Decay estimates of solutions for the wave equations with strong damping terms in unbounded domains. *Math. Methods Appl. Sci.*, 24:659 – 670, 06 2001.

[23] V. Kalantarov and S. Zelik. A note on a strongly damped wave equation with fast growing nonlinearities. *J. Math. Phys.*, 01 2015.

[24] P. Kelly. *Solid Mechanics Part I: An Introduction to Solid Mechanics*. University of Auckland, 2019.

[25] A. Khanmamedov. Strongly damped wave equation with exponential nonlinearities. *J. Math. Anal. Appl.*, 419(2):663 – 687, 2014.

[26] S. Larsson and V. Thomée. *Partial Differential Equations with Numerical Methods*, volume 45. Springer-Verlag Berlin Heidelberg, 2003.

[27] S. Larsson, V. Thomée, and L. B. Wahlbin. Finite-element methods for a strongly damped wave equation. *IMA J. Numer. Anal.*, 11(1):115–142, 1991.

[28] Y. Lin, V. Thomée, and L. Wahlbin. *Ritz-Volterra projections to finite element spaces and applications to integro-differential and related equations*. Technical report (Cornell University. Mathematical Sciences Institute). Mathematical Sciences Institute, Cornell University, 1989.

[29] K. Lipnikov, G. Manzini, J. D. Moulton, and M. Shashkov. The mimetic finite difference method for elliptic and parabolic problems with a staggered discretization of diffusion coefficient. *J. Comput. Phys.*, 305:111–126, 2016.

[30] A. Målqvist and D. Peterseim. Generalized finite element methods for quadratic eigenvalue problems. *ESAIM Math. Model. Numer. Anal.*, 51(1):147–163, 2017.

[31] A. Målqvist and D. Peterseim. *Numerical homogenization by localized orthogonal decomposition*, volume 5 of *SIAM Spotlights*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2020.

[32] R. Maier and D. Peterseim. Explicit computational wave propagation in micro-heterogeneous media. *BIT Numer. Math.*, 59(2):443–462, 2019.

[33] A. Målqvist and A. Persson. A generalized finite element method for linear thermoelasticity. *ESAIM Math. Model. Numer. Anal.*, 51(4):1145–1171, 2017.

[34] A. Målqvist and A. Persson. Multiscale techniques for parabolic equations. *Numer. Math.*, 138(1):191–217, 2018.

[35] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014.

[36] P. Massatt. Limiting behavior of strongly damped nonlinear wave equations. *J. Differ. Equ.*, 48:334–349, 06 1983.

[37] H. Owhadi. Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. *SIAM Rev.*, 59(1):99–149, 2017.

[38] H. Owhadi and C. Scovel. *Operator-adapted wavelets, fast solvers, and numerical homogenization*, volume 35 of *Cambridge Monographs on Applied and Computational Mathematics*. Cambridge University Press, Cambridge, 2019.

[39] D. Peterseim. Variational multiscale stabilization and the exponential decay of fine-scale correctors. In *Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations*, volume 114 of *Lect. Notes Comput. Sci. Eng.*, pages 341–367. Springer, Cham, 2016.

[40] H. Rui and H. Pan. Block-centered finite difference methods for parabolic equation with time-dependent coefficient. *Jpn. J. Ind. Appl. Math.*, 30:681–689, 2013.

[41] K. Urban and A. T. Patera. An improved error bound for reduced basis approximation of linear parabolic problems. *Math. Comp.*, 83(288):1599–1615, 2014.

[42] S. Yousefi, D. Lesnic, and Z. Barikbin. Satisfier function in ritz-galerkin method for the identification of a time-dependent diffusivity. *J. Inverse Ill Posed Probl.*, 20, 12 2012.