



CHALMERS
UNIVERSITY OF TECHNOLOGY



UNIVERSITY OF GOTHENBURG

Variational crimes in the Localized orthogonal decomposition method

Master's thesis

TIM KEIL

MASTER'S THESIS 2017

**Variational crimes in the Localized
orthogonal decomposition method**

TIM KEIL



CHALMERS
UNIVERSITY OF TECHNOLOGY



UNIVERSITY OF GOTHENBURG

Department of Mathematical Sciences
CHALMERS UNIVERSITY OF TECHNOLOGY
UNIVERSITY OF GOTHENBURG
Gothenburg, Sweden 2017

Variational crimes in the Localized orthogonal decomposition method
TIM KEIL

© TIM KEIL, 2017.

Supervisor: Axel Målqvist, Department of Mathematical Sciences
Examiner: Stig Larsson, Department of Mathematical Sciences

Master's Thesis 2017
Department of Mathematical Sciences
Chalmers University of Technology
SE-412 96 Gothenburg
Telephone +46 31 772 1000

Gothenburg, Sweden 2017

IV Variational crimes in the Localized orthogonal decomposition method

Variational crimes in the Localized orthogonal decomposition method

TIM KEIL

Department of Mathematical Sciences

Chalmers University of Technology

Abstract

This thesis deals with solving multiscale elliptic problems with perturbed coefficients. In this context, the Localized orthogonal decomposition (LOD) method for solving multiscale partial equations is presented and the main results of the error analysis are demonstrated and improved. Moreover, the thesis proposes a method, deduced from the LOD, to solve the variational crimes of perturbations by efficiently taking advantage of the underlying reference configuration. For this method, the numerical analysis and several experiments are presented. Furthermore, numerical experiments for perturbations are discussed and the novel method is assessed. Lastly, the deduced method is applied to weakly random problems as a special instance of perturbations.

Acknowledgements

I would like to thank my supervisor Professor Axel Målqvist for his great support and helpful discussions throughout my whole time in Sweden as well as Fredrik Hellman and Anna Persson for sharing their work and helping whenever it was necessary. Furthermore, I would like to thank Professor Mario Ohlberger for his support from Münster and for leading me into this excellent topic. My thanks also go to my family, which made my year abroad possible, and my friends for inspiration and support, especially to Per Ljung and Lucie Buttkus for their helpful comments on this thesis.

Tim Keil, Gothenburg, June 2017

Contents

1	Introduction	1
2	Problem formulation and mathematical background	5
2.1	The weak formulation	5
2.2	The Finite Element Method	7
2.3	Error bounds for finite element methods	8
2.4	Examples	10
2.4.1	One dimension	10
2.4.2	Two dimensions	12
3	The Localized Orthogonal Decomposition Method	15
3.1	Multiscale splitting	15
3.2	The standard method	16
3.3	Localization	18
3.3.1	Classical nodal patch localization	18
3.3.2	Element patch localization	20
3.4	The Petrov Galerkin formulation	21
3.5	Right hand side correction	22
3.5.1	Application to the LOD	23
3.5.2	Application to the PG-LOD	23
3.6	Interpolation	25
4	Error analysis for the standard LOD	27
4.1	Classical proof	27
4.2	New formulation	32
5	Variational Crimes	42
5.1	The novel Method	45
6	Error analysis for the PG-LOD on perturbed problems	49
6.1	Error indicators	49

6.2	Error analysis	50
6.2.1	The ideal PG-LOD approximation	51
6.2.2	The localized PG-LOD approximation	52
6.2.3	The PG-LOD approximation with the reference LOD space	55
6.2.4	The localized PG-LOD approximation of the reference problem	58
6.2.5	Error bound for reference localized PG-LOD	62
7	Implementation	63
7.1	Discretization	63
7.2	The standard LOD	65
7.3	The Petrov Galerkin LOD	65
7.4	Right hand side Correction	66
7.5	The novel method	67
7.5.1	Error indicators	68
8	Numerical Experiments	69
8.1	One dimensional PG-LOD experiment	69
8.2	Perturbations	71
8.3	The novel method	73
8.4	Numerical examples and discussion	76
8.5	Weakly Random Problems	86
8.6	Monte Carlo simulations	88
9	Conclusion and future work	91
	List of figures	95
	Bibliography	C

1 Introduction

The amount of composite materials used in today's industries and engineering has risen dramatically. They combine various materials to a heterogenous material with the intention to attain better properties than each component on its own. Advantages in robustness or strength, for instance, are achieved, while the weight reduces significantly. Roughly, those materials contain a matrix and a reinforcement, mostly in the form of fibers, depending on the desired properties. A popular example is the utilization in airplane industries. Nowadays, an airplane contains approximately 50% composite materials to make use of lightness and stability properties. Further examples that use composite materials are aerospace engineering, automotive and marine industries or bridge construction. Certainly, each component of the composite material maintains its own chemical, physical and mechanical properties. This might produce rapid changes, high variations and discontinuities in terms of the behavior of the heterogenous material in special situations. In practice, due to the high variations, such materials raise new challenges when it comes to mathematical modeling of processes like flow and diffusion (see [9]). Moreover, those materials produced in industries, potentially have small perturbations, for instance caused by machine failures or unexpected compositions. These circumstances increase the challenge for investigations. Modeling physical processes on such media with or without perturbations usually ends in a partial differential equation that requires vast complexity for computations in order to yield an accurate approximation of the problem. Other examples that produce the same effects are diffusion in porous media, groundwater flow or signal transduction in cell biology.

Partial differential equations that are governed by rapid changes or high variations, as described above, are called multiscale problems. These are problems in which several inherent scales are involved and affect the resulting solutions. Those scales might be non separable, but can be distinguished into the macroscale and microscales. The macroscale indicates the global behavior of the solution, a course average over strong changes, whereas the microscales are responsible for those changes. Consequently, analyzing multiscale processes requires a 'microscopic' accurate investigation, in order to capture every feature correctly. This thesis solely deals with multiscale diffusion problems and methods to resolve the issue of a multiscale setting. In general, the elliptic partial

differential equation for diffusion processes can be formulated as

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

on a bounded Lipschitz domain $\Omega \subset \mathbb{R}^d$. The diffusion coefficient A contains information concerning the substance or the media the problem is subjected to. Whereas f is responsible for outer forces that have influences on the process. For sufficiently strong assumptions on each component of the diffusion, the classical finite element method (FEM) is applicable and achieves good approximations (see [7]). The FEM is a Galerkin method, based on a discretization of the domain into a non-overlapping mesh with, depending on the dimension of the domain, different but consistent shapes of elements. This procedure yields a finite dimensional space, containing finitely many local basis functions. In order to capture every microscopic effect, the mesh size has to be chosen sufficiently small and consequently a high amount of elements is required. This produces a large computational complexity and might reach the limits of today's computer technology. To bypass this issue, a lot of research has been conducted in order to develop and apply novel methods, based on the FEM. Among others, [5], [4], [21], [22] and [14] achieved first promising results and provided the groundwork for further improvements. Most of the methods mainly intend to use the high resolution mesh on only small subdomains and use the received information to incorporate the finescale effects to the method. This adjustment is made by correcting the FE-basis of a Galerkin-method with a coarser mesh. The heterogeneous multiscale method (HMM) in [11] and the multiscale finite element method (MsFEM) in [21] represent two popular examples for methods that follow the exact same strategy as described above. They have been successfully applied and improved in recent research. Both of them as well as several other methods assume periodicity in error analysis or a specific scale separation. Thus, they are applicable only in a limited way. Certainly, further research is also devoted to methods that do not necessitate strong assumptions to be able to deal with non-periodic, discontinuous and rapidly oscillating coefficients, but to still generate a sufficiently accurate solution. Moreover, those methods intend to be most efficient and cheap in terms of computational complexity.

This thesis is devoted to a method that does not make any assumption like the one mentioned above. It is called the localized orthogonal decomposition method (LOD) and was invented by Målqvist and Peterseim (see [28]). This method is based on the variational multiscale method (VMM) in [22] as it decomposes the solution space in a fine and a coarse part. The finescale information that is lost in the coarse FE space can be recaptured by the kernel of an interpolant that maps a function from the infinite

space to the FE space. Thus, it is possible to identify a finescale space on which the subdomain problems for the finescale basis correctors can be computed and subsequently incorporated to the coarse Galerkin method. In the ideal version of the LOD, these utilized correctors have global support. In fact, fine mesh computations get increased instead of avoided which implies that the ideal LOD has no benefit yet. However, Målqvist and Peterseim proved that these correctors decay exponentially outside an area away from the node with which they are associated with. This justifies to compute the correctors solely on a patch around the node which converts the LOD into a feasible method with high practical use. In the recent years, the LOD has been improved and generalized in [16], [18], reviewed in [30] and applied in [26], [15], [29], [17], [27], [1] and [19]. Furthermore, the LOD provided further work in similar approaches such as [23]. In order to achieve a less expensive version, a Petrov Galerkin formulation (PG-LOD) has been proposed in [12]. This method has computational advantages and furthermore, a similar stability and convergence behavior can still be reached. Concerning the implementation of the LOD, a detailed overview has been given in [13]. The first aim of this thesis is to explain the LOD, its improvements as well as the PG-LOD in detail and derive an error estimate for the approximation that proves a sufficient convergence result. This derivation also contains the proof of the exponential decay and hereof, we stress and summarize the improvements which have been built on the classical proof in recent years.

The second part of this thesis deals with variational crimes. Applications of Galerkin methods, such as the finite element method, might violate standard results and the accuracy might get affected. This could be caused by inexact quadrature rules for integrals or inaccurate triangulations of the domain. Gilbert Strang (see [32]) first mentioned these issues in 1973 and developed estimates that work as a replacement for standard inequalities like Céa's lemma. One variational crime is a perturbation of the diffusion coefficient, which basically signifies the same effect like the one caused by the quadrature rule. In case of a high loss of accuracy, the original FEM has to be recomputed entirely for the new coefficient and no utilization of the old coefficient is available. Thus, there is no potential to save the computational cost. We show that this issue can be solved by the PG-LOD. For this purpose, a novel method based on approaches in [15] that contains the original as well as the perturbed coefficient is deduced and investigated analytically.

In the last part of this thesis, we apply this method to weakly random problems. These problems are deterministic diffusion problems, governed by randomly distributed perturbations in the diffusion coefficient. They have recently been introduced and studied by Legoll, Thomines and Le Bris in, among others, [24] and [25]. To achieve stochastic

simulations, they derived a novel method that is based on the MsFEM and saves computational complexity in order to be able to apply Monte Carlo methods. Therefore, their analysis is restricted on periodic assumptions for the deterministic problems as well as the randomness. We clarify that our novel method also results in sufficiently accurate and relatively cheaply computed approximations and covers non periodic instances. For this purpose, we perform various numerical experiments and assess them afterwards.

This thesis is organized as follows: Chapter 2 introduces the mathematical background and presents the original form of the finite element method and the problems that arise with multiscale data. Chapter 3 is devoted to the LOD method, as it stresses the main components, the method itself and improvements such as the PG-LOD. Chapter 4 aims to reveal an overview of the classical numerical analysis, containing error bounds and the exponential decay, while Chapter 5 stresses the issue of variational crimes and proposes the novel PG-LOD method. Chapter 6 presents the error analysis that follows from the proposed method in terms of the required stability and error results for the PG-LOD and Chapter 7 sketches implementation details for the LOD and the PG-LOD method. Chapter 8 contains the application of the novel method and presents the numerical experiments, followed by a conclusion in Chapter 9.

2 Problem formulation and mathematical background

This chapter presents mathematical background information and formulates the problem this thesis is devoted to. The reader is supposed to be familiar with basic linear functional analysis (see [2]) and the finite element method (see [7]). Nevertheless, we give a brief introduction to finite element methods. A short insight in the error analysis of the standard method emphasizes the problems of the FEM regarding multiscale scenarios. Furthermore, we demonstrate two multiscale problems, approximated with the standard FEM.

2.1 The weak formulation

Let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with polygonal boundary and dimension $d = 1, 2, 3$. The underlying problem is a diffusion boundary value problem which can be described as

$$\begin{cases} -\nabla \cdot (A(x)\nabla u(x)) = f(x), & \text{for } x \in \Omega, \\ u(x) = 0, & \text{for } x \in \partial\Omega. \end{cases}$$

Here, f denotes a function in $L^2(\Omega)$ and furthermore, we assume the diffusion coefficient $A \in L^\infty(\Omega, \mathbb{R}^{d \times d})$ to be uniformly elliptic such that

$$0 < \alpha := \operatorname{ess\,inf}_{x \in \Omega} \inf_{v \in \mathbb{R}^d \setminus \{0\}} \frac{(A(x)v) \cdot v}{v \cdot v}, \quad (2.1)$$

$$\infty > \beta := \operatorname{ess\,sup}_{x \in \Omega} \sup_{v \in \mathbb{R}^d \setminus \{0\}} \frac{(A(x)v) \cdot v}{v \cdot v}. \quad (2.2)$$

The function A may be non-periodic as well as subjected to high variations, such as rapidly oscillating scales or discontinuities. Moreover, we assume homogenous Dirichlet boundary conditions. With standard arguments, this problem can be reformulated into a weak formulation. We define $V := H_0^1(\Omega)$, the Hilbert space with homogenous boundary conditions that belongs to the scope of Sobolev spaces (see [2])

$$H_0^1(\Omega) := \left\{ u \in L^2(\Omega) \left| \frac{\partial u}{\partial x_j} \in L^2(\Omega), j = 1, \dots, n, u = 0 \text{ on } \partial\Omega \right. \right\}.$$

For the weak formulation, we use the symmetric, coercive and bounded bilinear form a such that

$$a(u, v) := \int_{\Omega} (A\nabla u) \cdot \nabla v, \quad \forall u, v \in V$$

and the bounded linear functional

$$F(v) := \int_{\Omega} f v, \quad \forall v \in V.$$

The weak formulation is also called a Galerkin method and can be stated as follows.

2.1.1 Definition (Exact solution of the diffusion problem) For V , a and F defined as above, the weak formulation of the diffusion problem is to find $u \in U := V$ such that, for all $v \in W := V$, it holds that

$$a(u, v) = F(v). \quad (2.3)$$

In general, the space U called the trial space and W the test space.

Clearly, the trial space U and the test space W are chosen to be the high resolution space V as this approach results in the exact solution for the underlying problem. The choice of the trial and test space could possibly deviate in order to gain novel Petrov Galerkin methods. However, this requires new conditions on the bilinear form a , since coercivity is not fulfilled anymore. Further details are given in Chapter 3. In the following, we call the solution of (2.3) the exact solution, as we want to approximate it with the methods throughout this thesis. In terms of the existence of a unique solution, we apply the Lax-Milgram Theorem.

2.1.2 Theorem (Lax-Milgram, see [7]) Let X be a real Hilbert space, $B : X \times X \rightarrow \mathbb{R}$ a bounded bilinear form, i.e.,

$$\exists C > 0 : \forall x, y \in X : |B(x, y)| \leq C \|x\|_X \cdot \|y\|_X.$$

Furthermore, B is coercive, which means by definition

$$\exists c_0 > 0 : \forall x \in X : B(x, x) \geq c_0 \|x\|_X^2$$

and $F : X \rightarrow \mathbb{R}$ is a bounded linear functional ($F \in X'$). Then there exists a unique solution $u \in X$ such that, for all $\varphi \in X$,

$$B(u, \varphi) = F(\varphi).$$

Since $V = H_0^1(\Omega)$ is a Hilbert space and a and F satisfy the required assumptions for the theorem, we conclude that the exact solution of (2.3) exists uniquely. This also holds true

for closed subspaces of V , since they are still Hilbert spaces. Now, we want to focus on the approximation of (2.3).

2.2 The Finite Element Method

The finite element method (FEM) and its variations are developed to approximate a solution for problems of the form (2.3). The method is roughly speaking a Galerkin method with specified trial and test space. We define a Finite Element mesh (FE mesh) and introduce the Finite Element space (FE space) V_H as a finite dimensional subspace of V . Let \mathcal{T}_H be a family of coarse, shape regular, conforming triangulations of the domain Ω such that

$$\bigcup_{T \in \mathcal{T}_H} T = \Omega.$$

These triangulations contain, for instance, intervals in one dimension, triangles in two dimensions, and tetrahedrons or hexahedrons in three dimensions. We denote the maximum diameter of an element in \mathcal{T}_H with H and the shape regularity is fulfilled by the existence of a ρ independent on H such that

$$\max_{T \in \mathcal{T}_H} \frac{\text{diam}(T)}{B(T)} \leq \rho \quad \text{and} \quad \max_{T_1, T_2 \in \mathcal{T}_H} \frac{\text{diam}(T_1)}{\text{diam}(T_2)} \leq \rho.$$

$B(T)$ denotes the largest ball that lays in T . With \mathcal{N} , we denote the set of all corresponding interior nodes of the mesh \mathcal{T}_H , which means they are not part of the boundary $\partial\Omega$. A typical finite element space is

$$\mathcal{P}_1(\mathcal{T}_H) := \{v \in C^0(\Omega) \mid v|_T \text{ is a linear polynomial of degree } \leq 1, \text{ for every } T \in \mathcal{T}_H\},$$

containing all \mathcal{T}_H -piecewise linear functions that are continuous on the domain Ω . Other typical finite element spaces like $\mathcal{P}_2(\mathcal{T}_H)$ are also conceivable, but they are not considered further. We set

$$V_H := V \cap \mathcal{P}_1(\mathcal{T}_H).$$

The space V_H is clearly finite-dimensional and thus, we can find a basis $(\lambda_x)_{x \in \mathcal{N}} \subset \mathcal{P}_1(\mathcal{T}_H)$ such that the property

$$\lambda_x(x) = 1 \quad \text{and} \quad \lambda_x(y) = 0, \quad \text{for every other node } x \neq y \in \mathcal{N}$$

is satisfied. Since, for every $x \in \mathcal{N}$, we have $\lambda_x \in \mathcal{P}_1(\mathcal{T}_H)$, the corresponding basis contains 'hat'-functions that form a partition of unity. We call them nodal basis functions. Importantly, these basis functions have a small support as they vanish outside of a vertex patch. The corresponding method is a so-called Galerkin approximation and we call it

the standard finite element method.

2.2.1 Definition (Standard FEM approximation) The Galerkin approximation for the standard FEM is to find $u_H \in V_H \subset V$ such that

$$a(u_H, v) = F(v), \quad \forall v \in V_H. \quad (2.4)$$

This solution is unique, according to Lax-Milgram.

Several choices for the trial space and the test space yield various types of methods, which will be seen later on. For the standard FEM, the trial space equals the test space, which enables the application of the Lax-Milgram uniqueness theorem. The main idea of the FEM is to gain an approximation of the exact solution u with

$$u_H = \sum_{x \in \mathcal{N}} \mathcal{U}_H(x) \lambda_x,$$

where \mathcal{U}_H denotes a vector of size $|\mathcal{N}|$. Inserting this into (2.4) results in a system of linear equations

$$\mathcal{S} \cdot \mathcal{U}_H = \mathcal{L}, \quad (2.5)$$

where each entry of the stiffness matrix \mathcal{S} is determined by

$$\mathcal{S}_{x,y} = a(\lambda_x, \lambda_y), \quad \forall x, y \in \mathcal{N},$$

and the load vector \mathcal{L} is similarly defined by

$$\mathcal{L}_x = F(\lambda_x), \quad \forall x \in \mathcal{N}.$$

As a consequence of the small support of the nodal basis functions, the resulting matrix \mathcal{S} is sparse and, just as the load vector, due to linearity, easy to compute. This is a great advantage of the method, since the linear system of equations with a sparse matrix can be solved efficiently. The size of the stiffness matrix and hence, the size of the load vector increase for smaller H . This implies a bigger amount of elements in \mathcal{T}_H . Thus, the smaller H , the greater the effort to compute the components of (2.5) and to solve the linear system of equations, subsequently.

2.3 Error bounds for finite element methods

The error analysis for the standard FEM emphasizes the space for developing more Galerkin methods that circumvent the drawbacks of the FEM. For this chapter, we omit the proofs, since we state standard results that can be found in [7]. Throughout the whole

thesis, we use the energy norm

$$\|\cdot\| = a(\cdot, \cdot)^{1/2} = \|A^{1/2}\nabla\cdot\|_{L^2(\Omega)},$$

induced by a , and the Cauchy Schwarz inequality for a Hilbert space X

$$|(x, y)_X| \leq \|x\|_X \|y\|_X, \quad \forall x, y \in X.$$

In terms of error bounds for the FEM, Céa's lemma is crucial and is formulated in the following.

2.3.1 Lemma (Céa's lemma, see [7]) For u , the exact solution of (2.3) in Definition 2.1.1, and u_H , the FEM approximation in Definition 2.2.1, the estimate

$$\|u - u_H\|_V \leq \frac{C}{c_0} \inf_{v_H \in V_H} \|u - v_H\|_V$$

holds true. Here, C and c_0 denote the continuity and the coercivity constants. Moreover, we have Galerkin orthogonality, i.e., for all $v_H \in V_H$,

$$a(u - u_H, v_H) = 0.$$

On the basis of Céa's lemma, standard approaches for error bounds of the FEM can be deduced. The following error bound is one of them. It follows from an interpolation estimate.

2.3.2 Theorem (Error bound for the FEM, see [7]) The exact solution $u \in H^2(\Omega)$ of (2.3) and u_H , the standard FEM approximation in (2.4), underlay the error estimate

$$\|u - u_H\| \leq CH \|D^2 u\|_{L^2(\Omega)},$$

where C is a constant independent of H .

This error emphasizes the demand of multiscale methods. The convergence of the method (2.4) is therefore linear, according to the mesh size H , which can be chosen appropriately small and close to 0. Nevertheless, this convergence result works just as long as the solution satisfies sufficiently well smoothness conditions. It might occur that the solution of the diffusion problem has rapid variations on multiple scales, which could, for instance, be induced by high variation in the diffusion coefficient. If we assume that, regardless of the cause, u has such oscillations with at least a frequency of ε^{-1} , for a small $\varepsilon > 0$, then $\|\nabla^2 u\|_{L^2(\Omega)}$ would produce a factor in the estimate which is approximately of the same size. Hence, to oppose this effect, we would have to choose $H \ll \varepsilon$ in order to capture every

effect on each scale. This would imply a vast growth of the computational complexity for the method, if ε is tiny or even goes to zero. Next, we visualize the multiscale setting with examples in one and two dimensions.

2.4 Examples

This section presents two examples for multiscale problems that emphasize the need to develop novel methods.

2.4.1 One dimension

As mentioned above, the considered multiscale problems are subjected to high variations on microscopic scales. For instance, we can describe them with a small ε , in the sense that it represents the length of each periodic oscillation. The following is a standard example and has been introduced in [30]. Consider the partial differential equation

$$\begin{cases} -\frac{\partial}{\partial x} A_\varepsilon(x) \frac{\partial}{\partial x} u_\varepsilon(x) &= 1, & \text{for } x \in (0, 1) \\ u_\varepsilon(0) = u_\varepsilon(1) &= 0, \end{cases}$$

where, for $\varepsilon > 0$,

$$A_\varepsilon(x) := (2 + \cos(2\pi x/\varepsilon))^{-1}.$$

In Figure 2.1, the coefficient is displayed for $\varepsilon = 2^{-5}$ and $\varepsilon = 2^{-6}$. The exact solution can

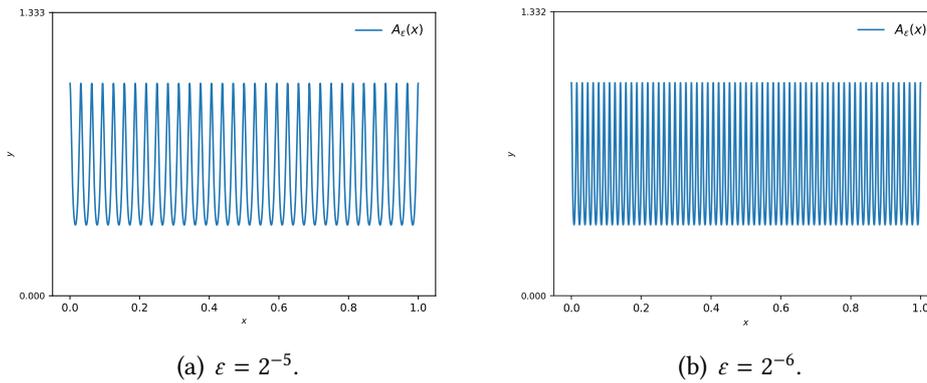


Figure 2.1: Periodic variations of $A_\varepsilon(x)$ on $\Omega = (0, 1)$ for $\varepsilon = 2^{-5}$ and $\varepsilon = 2^{-6}$.

be stated explicitly with

$$u_\varepsilon = 4(x - x^2) - 4\varepsilon \left(\frac{1}{4\pi} \sin(2\pi \frac{x}{\varepsilon}) - \frac{1}{2\pi} x \sin(2\pi \frac{x}{\varepsilon}) - \frac{\varepsilon}{4\pi^2} \cos(2\pi \frac{x}{\varepsilon}) + \frac{\varepsilon}{4\pi^2} \right).$$

In the sequel, we use $\varepsilon = 2^{-5}$. In fact, the oscillations for $A_\varepsilon(x)$ are rather low, but nevertheless, it already emphasizes the problem quite well. Figure 2.2 shows the FEM

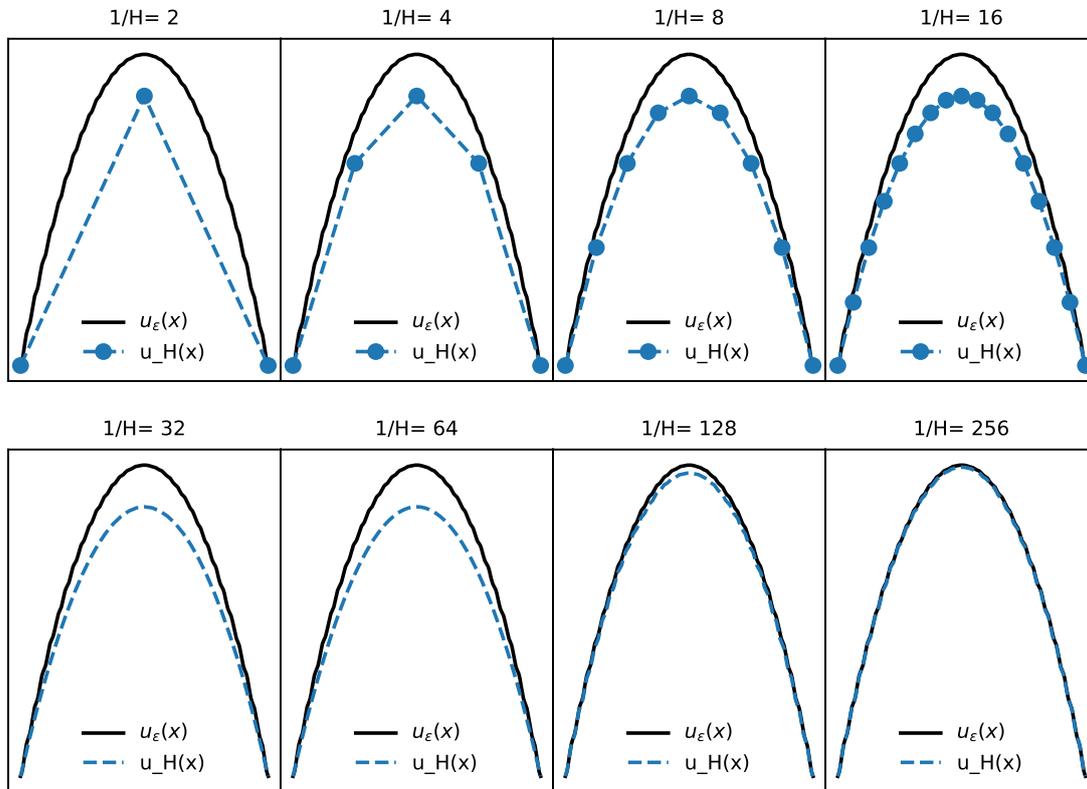


Figure 2.2: FEM approximation for various choices of H .

approximation for several mesh sizes H in comparison to the exact solution u_ε . It turns out that the macroscopic trend of the FEM approximation lays distinctly underneath the macroscopic trend of the exact solution as long as H is chosen too large. Remarkably, the macroscopic approximation failure stays on the same level until $1/H = 64 = 2^6$. This can also be noticed in the error plot in Figure 2.3, which displays the energy error $\|u_\varepsilon - u_H\|$. Clearly, mesh sizes $H \geq \varepsilon$ are not suitable to capture the microscopic effect. Only for $H = 2^7$, the solution slowly begins to adapt its macroscopic behavior. Figure 2.3 emphasizes this approach. The error decreases linearly with respect to H , once the microscopic effect is captured. This error behavior of the solution is not accessible, which was already explained in the previous section. The macroscopic trend is correct for a sufficiently small choice of H and thus, it has to fulfill $H \ll \varepsilon$.

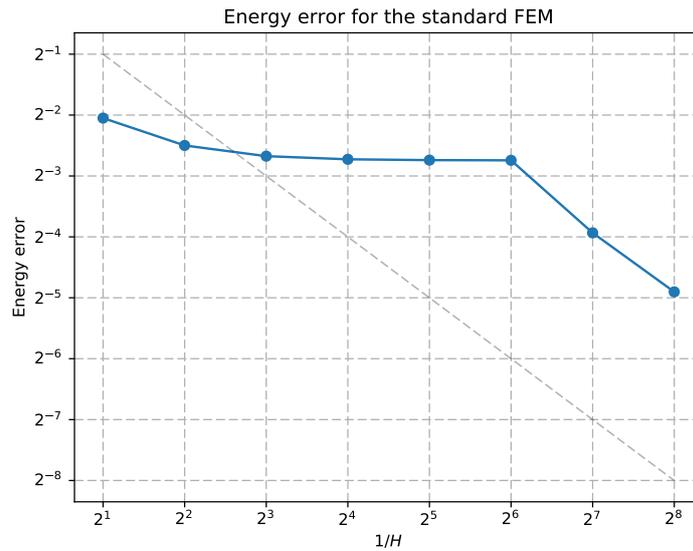


Figure 2.3: Energy error $\|u_\varepsilon - u_H\|$ for various choices of H .

2.4.2 Two dimensions

In two dimensions, the same effects can be observed. Figure 2.4 displays the diffusion coefficient as well as the desired finescale solution of the problem. For this example, we drop the ε periodicity, as it is no required assumption for the approaches and methods in this thesis. However, the coefficient is still subtracted to some periodicity. This is not an essential assumption, but it makes it easier to follow the accuracy of the solutions. The channels in the coefficient are noticeable in the finescale solution. Obviously, the

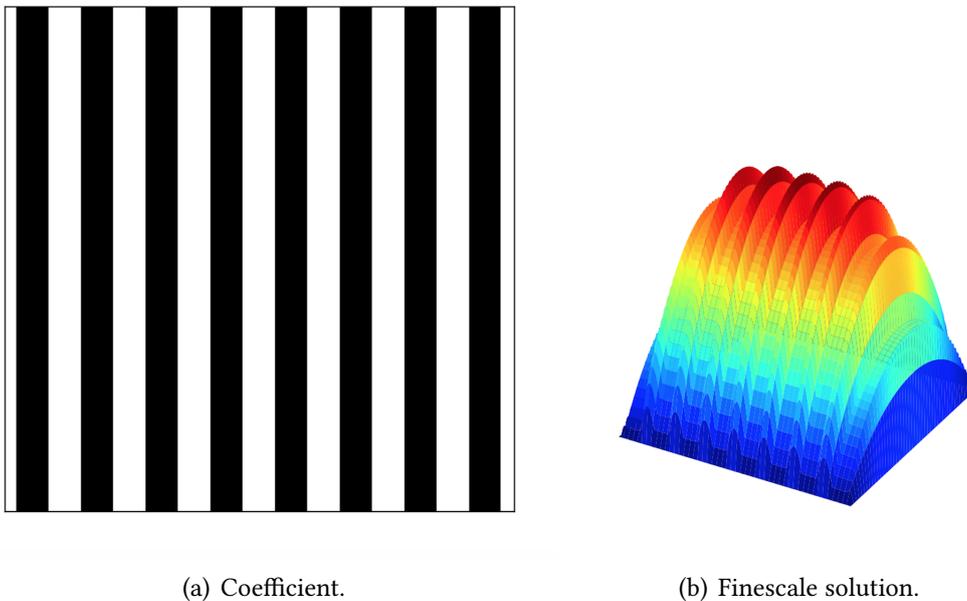


Figure 2.4: $A(x)$ and $u(x)$. Black is 1 and white is 0.01.

diffusion has the largest impact inside the black channels which produce the arcs. Figure 2.5 displays different FEM approximations and Figure 2.6, once again, shows the energy error $\|u - u_H\|$. Although, the coefficient consists of only 16 channels, for $1/H = 16$, the channels are not discernible at all. Furthermore, the macroscopic behavior is rather low. The arcs get recognizable for $1/H = 32$ but they are still not equally high and the solution has only slight accuracy. Figure 2.6 stresses these results as the error first behaves almost constant and, starting from $1/H = 32$, decreases almost linearly.

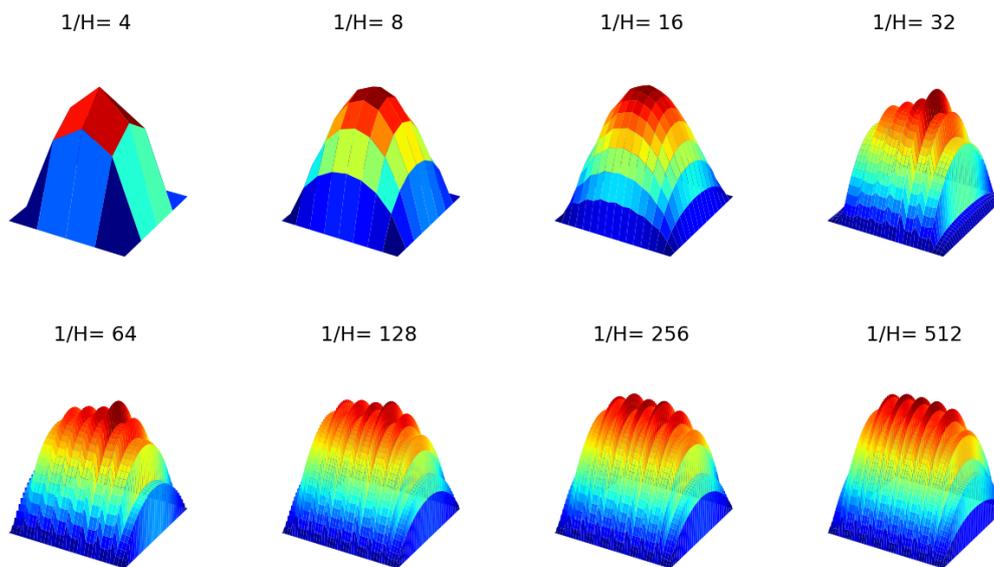


Figure 2.5: FEM approximation for various choices of H .

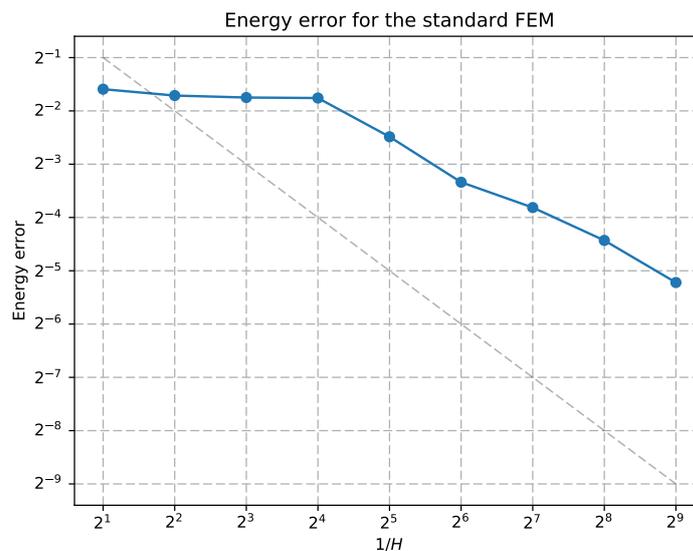


Figure 2.6: Energy error $\|u - u_H\|$ for various choices of H .

Both examples in this section emphasize the need for methods that do not have to cover the microscopic effects with a sufficiently small H , as this might produce too high complexity. The variations for $u_\varepsilon(x)$ and $u(x)$ have been intentionally chosen rather small. The effect clearly gets worse for vast variations. We get to know a lot of examples throughout Chapter 8. Especially for randomly perturbed problems, we have to compute many problems of the same type, which might become impossible for too high variations. This is the motivation for the development of multiscale methods that resolve this issue. Their aim is to derive an approximation u_H^{ms} that satisfies an error bound of the form

$$\|u - u_H^{\text{ms}}\| \leq C_f H,$$

where C_f is only dependent on the right hand side and also on the bounds for A . Furthermore, H does not resolve the variations of the coefficient A . As stated in the introduction, there are plenty of different multiscale methods that have been developed in the past. The tool for this thesis is the Localized orthogonal decomposition method. We present it in the next chapter.

3 The Localized Orthogonal Decomposition Method

This chapter presents the Localized orthogonal decomposition (LOD). It was first introduced by Målqvist and Peterseim in [28]. They show the existence of a (quasi-) optimal basis such that the resulting solution u_H^{LOD} satisfies

$$\|u - u_H^{\text{LOD}}\|_{H^1(\Omega)} \leq C_{f,\alpha,\beta}H,$$

with a constant $C_{f,\alpha,\beta}$, depending on the right hand side and the global bounds of the diffusion A . Furthermore, this constant is independent of the variations of A which allows for the method to solve multiscale problems. In particular, the aim of the LOD is to construct correctors for the basis functions by solving only local problems on coarse element patches. Målqvist and Peterseim justified this by proving that the constructed basis functions decay exponentially in an area outside of their associated node. This chapter explains the multiscale splitting into a fine and coarse part of the solution space and focuses on the interpolation that enables this approach. Subsequently, we formulate the standard LOD method and present old and new approaches for localizations. On top of that, we mention variations of the LOD such as the right hand side correction and the Petrov Galerkin formulation.

3.1 Multiscale splitting

The standard FE space V_H is a finite-dimensional subspace of the space $V = H_0^1(\Omega)$. Finescale features that might occur in the diffusion coefficients and that are hidden in V are no longer captured in the space V_H . Therefore, we would like to characterize the finescale intricacies in V . For this purpose, we introduce a linear surjective (quasi-) interpolation operator

$$I_H : V \rightarrow V_H$$

that maps a function $v \in V$ to a function $v_H \in V_H$ in the coarse FE space. Thus, we regain the finescale space with

$$V^f = \ker(I_H) = \{v \in V \mid I_H(v) = 0\}.$$

This procedure allows us to cover all features of V that are no longer covered in V_H and we yield the multiscale splitting

$$V = V^f \oplus V_H.$$

Figure 3.1 displays a two dimensional function in V , decomposed into the FE- and the finescale space. The sum of Figure 3.1(b) and Figure 3.1(c) yields the high resolution function v in Figure 3.1(a). Importantly, the choice of the interpolant is not unique and

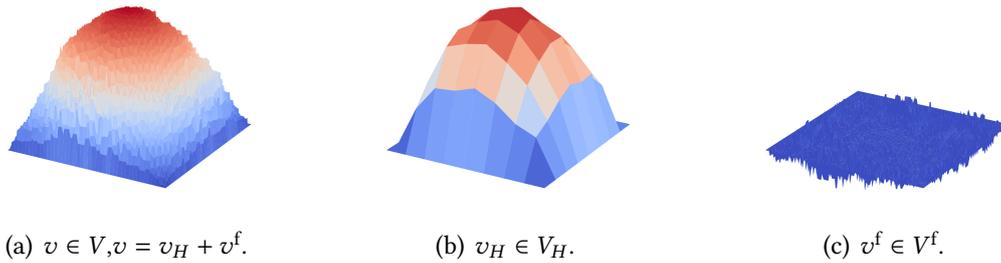


Figure 3.1: Multiscale splitting.

each choice yields another method. However, each interpolation operator needs to satisfy determined properties. These properties and possible different choices are thematized in Section 3.6. First of all, we formulate the method.

3.2 The standard method

We aim to derive an orthogonal decomposition with respect to the bilinear form a , the inner product of the energy norm $\| \cdot \|$. For this purpose, we define a corrector operator Q , for a given $v_H \in V_H$, to be the solution $Qv_H \in V^f$ such that

$$a(Qv_H, w) = a(v_H, w), \quad \forall w \in V^f.$$

Therefore, $Q : V_H \rightarrow V^f$ denotes a finescale projection to V^f . Further, we use this finescale part to define the LOD space

$$V_H^{\text{LOD}} := (V_H - QV_H).$$

Obviously, the dimension does not change, i.e., we still obtain $\dim(V_H^{\text{LOD}}) = \dim(V_H)$. For any $v^f \in V^f$ and $v_H^{\text{LOD}} \in V_H^{\text{LOD}}$, we observe

$$\begin{aligned} a(v_H^{\text{LOD}}, v^f) &= a(v_H - Qv_H, v^f) \\ &= a(v_H, v^f) - a(v_H, v^f) \\ &= 0. \end{aligned}$$

This leads to the orthogonal decomposition with respect to the a -scalar product,

$$V = V_H^{\text{LOD}} \oplus_a V^f.$$

In fact, every $v \in V$ can be decomposed into two parts,

$$v = v_H^{\text{LOD}} + v^f.$$

With this approach, we achieve a space V_H^{LOD} that also contains finescale features of V , as a modification of the standard FE space V_H . We are now prepared to formulate the ideal Galerkin approach of the LOD Method.

3.2.1 Definition (Ideal LOD approximation) The ideal LOD approximation of the exact solution u in (2.3) is to find $u_H^{\text{LOD}} \in V_H^{\text{LOD}}$ that satisfies

$$a(u_H^{\text{LOD}}, v) = F(v), \quad \forall v \in V_H^{\text{LOD}}. \quad (3.1)$$

As mentioned above, various multiscale methods differ in terms of choice of the trial and test space. In the standard LOD-method, we choose both to be equal to V_H^{LOD} , which keeps the coercivity of the bilinear form a and thus, (3.1) has a unique solution, according to Lax-Milgram. In order to form the LOD approximation into a feasible Galerkin method, we need to derive a basis for the new test and trial space V_H^{LOD} . Since V_H^{LOD} and V_H have equal dimensions, it suffices to apply the finescale corrector Q on every single basis function and incorporate it to the underlying basis function. Therefore, for every node $x \in \mathcal{N}$, we define $\phi_x := Q\lambda_x \in V^f$, which means that $\phi_x \in V^f$ solves the following equation

$$a(\phi_x, w) = a(\lambda_x, w), \quad \forall w \in V^f. \quad (3.2)$$

Due to the definition of V_H^{LOD} , we consequently conclude a basis

$$\{\lambda_x - \phi_x \mid x \in \mathcal{N}\}.$$

Clearly, (3.2) requires an additional finite element method to compute the correctors for each node. We furthermore realize that the correctors ϕ_x have a global support. With respect to Section 2.2, the corresponding stiffness matrix \mathcal{S} is no longer a sparse, but a full matrix. These ingredients show that the LOD, in the presented form, does not facilitate the issue of too high computational complexity and hence, it has no practical use. Due to that, Målqvist and Peterseim investigated the behavior of these globally supported correctors and they presented a proof in [28] that shows that the correctors indeed have an exponential decay outside an area of their node. This fact is the justification for cutting of the corrector and compute it on only a patch around the affected node $x \in \mathcal{N}$. Thus,

we achieve a method that is localized on every coarse grid patch. In recent years, it turned out that several strategies of localization imply various results in terms of the error estimates. We now explain both localizations.

3.3 Localization

In this section, we present the localization that turns the LOD into a feasible method. First of all, we define coarse grid patches for arbitrary sets in Ω .

3.3.1 Definition (Patches for sets in Ω) For $k = 1$, we set $U(\omega) := U_1(\omega)$. If $\omega = \{x\}$, for a node $x \in \mathcal{N}$, we call $U_k(x)$ a k -layer nodal patch. For $\omega = T$, where $T \in \mathcal{T}_H$, we call $U_k(T)$ a k -layer element patch. In the sequel, we often use the notation U_k for $U_k(T)$.

The finescale space V^f can also be restricted on these patches with the intuitive definition, for $\omega \subseteq \Omega$ and $k \in \mathbb{N}$,

$$V^f(U_k(\omega)) := \left\{ v \in V^f \mid v|_{\Omega \setminus U_k(\omega)} = 0 \right\}.$$

These local finescale patches enable the 'cut-off' of the corrector in order to avoid the global support for each of them. Functions in $V^f(U_k(\omega))$ vanish outside of the patch $U_k(\omega)$ and therefore, they have local support. Two major strategies for patches are possible, either nodal patches or element patches. In the subsequent, we review both of them.

3.3.1 Classical nodal patch localization

In the classical paper [28], Målqvist and Peterseim proposed a localization for patches around nodes and hereof, they defined

$$\begin{aligned} \omega_{x,1} &:= U(x) = \text{supp } \lambda_x, \\ \omega_{x,k} &:= U_k(x), \quad k = 2, 3, \dots \end{aligned}$$

Figure 3.2 visualizes the nodal patches for several k on a quadrilateral mesh. We denote the locally supported corrector for the node $x \in \mathcal{N}$ by $\phi_{x,k} \in V^f(\omega_{x,k})$ and let it be the solution of

$$a(\phi_{x,k}, w) = a(\lambda_x, w), \quad \forall w \in V^f(\omega_{x,k}).$$

Moreover, we gain a corresponding classical localized LOD space $V_{H,k}^{\text{lod}}$ defined by the basis

$$\{\lambda_x - \phi_{x,k}\}_{x \in \mathcal{N}}.$$

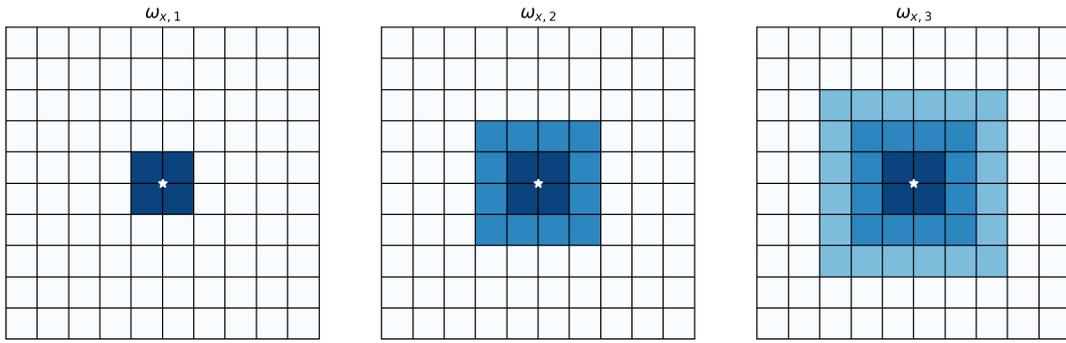


Figure 3.2: Patches around a certain node $x \in \mathcal{N}$ in white.

Note that we intentionally use the small letters to distinguish between the classical nodal patch localization and the element patch localization. Figure 3.3 displays this decomposition. Figure 3.3(b) is a coarse basis function prolonged on the fine mesh in order to gain a better comparison. In the following, we formulate the corresponding localized method.

3.3.2 Definition (nodal patch localized LOD approximation) The classical localized LOD approximation of the exact solution u in (2.3) is to find $u_{H,k}^{\text{lod}} \in V_{H,k}^{\text{lod}}$ that satisfies

$$a(u_{H,k}^{\text{lod}}, v) = F(v), \quad \forall v \in V_{H,k}^{\text{lod}}, \quad (3.3)$$

with $k \in \mathbb{N}$.

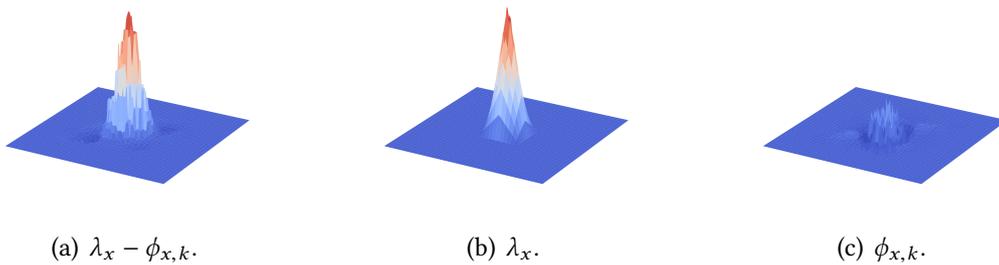


Figure 3.3: Basis function of $V_{H,k}^{\text{lod}}$ and its decomposition for an $x \in \mathcal{N}$ and $k = 4$.

The error analysis that is related to this method is stated in Chapter 4. On top of that, we see that this localization strategy indeed has a disadvantage that also affects the resulting error bound of the method. The alternative localization that reaches a better result is presented now.

3.3.2 Element patch localization

The classical proof is based on patches $\omega_{x,k}$ that are dependent on a node $x \in \mathcal{N}$. Now, we use element patches $U_k(T)$ for $T \in \mathcal{T}_H$ (see Figure 3.4). This is related to a new desired definition for the correctors on elements instead of nodes and it has been proposed in, among others, [26]. For $v \in V$, we can decompose the element corrector operator \mathcal{Q} with

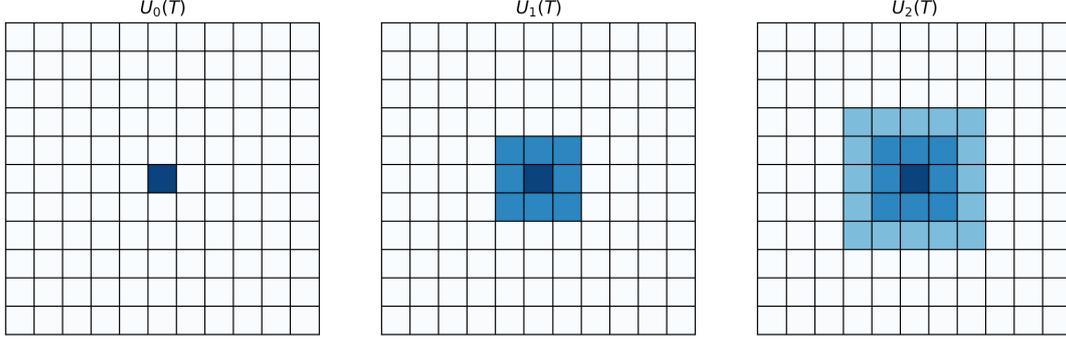


Figure 3.4: Patches for a certain coarse mesh element $T \in \mathcal{T}_H$ in the middle.

$$\mathcal{Q}v = \sum_{T \in \mathcal{T}_H} \mathcal{Q}^T v,$$

where $\mathcal{Q}^T v \in V^f$ are the solutions of

$$a_\Omega(\mathcal{Q}^T v, w) = a_T(v, w) \quad \forall w \in V^f, \quad (3.4)$$

where

$$a_M(u, v) := \int_M A \nabla u \cdot \nabla v,$$

for an arbitrary set $M \subseteq \Omega$. For $k \in \mathbb{N}$, we therefore define the element patch localized corrector operators $\mathcal{Q}_k^T : V \rightarrow V^f(U_k(T))$ by the solution of

$$a_{U_k(T)}(\mathcal{Q}_k^T v, w) = a_T(v, w), \quad \forall w \in V^f(U_k(T)), \quad (3.5)$$

and the corresponding full localized corrector

$$\mathcal{Q}_k v := \sum_{T \in \mathcal{T}_H} \mathcal{Q}_k^T v.$$

The correctors act on an arbitrary $v \in V$. Due to this novel correctors, we get a slightly different definition for the multiscale space $V_{H,k}^{\text{lod}}$ in Section 3.3.1, which we call $V_{H,k}^{\text{LOD}}$, defined by

$$V_{H,k}^{\text{LOD}} = V_H - \mathcal{Q}_k V_H.$$

This space is obviously spanned by $\{\lambda_x - \mathbf{Q}_k \lambda_x\}_{x \in \mathcal{N}}$ and leads to the following Galerkin method.

3.3.3 Definition (Element localized LOD approximation) The localized LOD approximation, based on element patches, is to find $u_{H,k}^{\text{LOD}} \in V_{H,k}^{\text{LOD}}$ such that

$$a(u_{H,k}^{\text{LOD}}, v) = F(v), \quad \forall v \in V_{H,k}^{\text{LOD}}. \quad (3.6)$$

The only difference between this LOD approximation and the classical one is the choice of the space $V_{H,k}^{\text{LOD}}$ instead of $V_{H,k}^{\text{lod}}$. Further, we get to know another method that accrues directly from the LOD approach and is called the Petrov Galerkin formulation.

3.4 The Petrov Galerkin formulation

The Petrov Galerkin formulation of the LOD has been proposed in [12] and is directly related to the standard method, since it uses the same LOD finescale space and therefore, the same correctors. Its ideal formulation reads as follows.

3.4.1 Definition (Ideal PG-LOD approximation) The ideal Petrov Galerkin LOD (PG-LOD) approximation of the exact solution u in (2.3) is to find $u_H^{\text{pg}} \in V_H^{\text{LOD}}$ that satisfies

$$a(u_H^{\text{pg}}, v) = F(v), \quad \forall v \in V_H. \quad (3.7)$$

Once again, we use the small letters to emphasize that we define a different method later on. Due to the same arguments, we also need a localized version. The localization for our PG-LOD is based on element patches.

3.4.2 Definition (Localized PG-LOD approximation) The localized LOD approximation of the exact solution u in (2.3) is to find $u_{H,k}^{\text{pg}} \in V_{H,k}^{\text{LOD}}$ that satisfies

$$a(u_{H,k}^{\text{pg}}, v) = F(v), \quad \forall v \in V_H, \quad (3.8)$$

with $k \in \mathbb{N}$.

Clearly, the standard LOD and the PG-LOD only differ in terms of the test space. We no longer use the LOD space V_H^{LOD} respectively, $V_{H,k}^{\text{LOD}}$, but the FEM space V_H . In practice, this approach has advantages regarding the computational complexity, which is discussed in Chapter 7. Moreover, the PG-LOD still fulfills similar convergence results (see [12]). It is important to notice that the coercivity of a is not valid anymore, since the trial space and the test space are not equal. Thus, the well-posedness of the PG-LOD is not yet

justified. We have to revert to a more general version of coercivity that is the inf-sup stability, defined as follows.

3.4.3 Definition (Inf-sup condition, see [6]) The continuous bilinear form $a : V \times W \rightarrow \mathbb{R}$ satisfies the inf-sup stability, if there exists a $\gamma > 0$ such that

$$\inf_{v \in V} \sup_{w \in W} \frac{|a(v, w)|}{\|v\| \|w\|} \geq \gamma.$$

This inf-sup condition is also related to the Babuška-Brezzi condition, named after Babuška [3] and Brezzi [8]. We also refer to [6] for further details. Once a satisfies this inf-sup condition for the given trial and test space, we conclude well-posedness of the method. This justifies the use of the PG-LOD. The inf-sup stability and the resulting error estimates, as well as the advantages and details, will be focused on later on. In order to gain advantages in the error analysis of especially the PG-LOD, we now present an additional correction tool, which is called 'the right hand side correction'. It has been presented in [26], in order to consider high contrast problems regarding the standard LOD.

3.5 Right hand side correction

To deal with high contrast problems in [26], the right hand side correction is crucial. Although this thesis does not discuss the problem of high contrast, the correction is a key tool in our approaches in Chapter 6. We therefore introduce it and explain the usage in the sequel. We define the right hand side correction, for every $v \in V$, with $\mathcal{R}v \in V^f$ such that, for all $v^f \in V^f$,

$$a(\mathcal{R}v, v^f) = F(v^f).$$

Furthermore, we define the corresponding localized operator

$$\mathcal{R}_k v = \sum_{T \in \mathcal{T}_H} \mathcal{R}_k^T v,$$

where, for every $T \in \mathcal{T}_H$, the operator $\mathcal{R}_k^T : V \rightarrow V^f(U_k(T))$ denotes the solution such that, for all $v^f \in V^f(U_k(T))$,

$$a_{U_k(T)}(\mathcal{R}_k^T v, v^f) = F_T(v^f).$$

This corrector has an application in the classical LOD as well as the PG-LOD.

3.5.1 Application to the LOD

The classical LOD method makes use of the orthogonal decomposition of the exact solution $u = u_H^{\text{LOD}} + u^f$, where u^f denotes a function in the finescale space V^f and $u_H^{\text{LOD}} \in V_H^{\text{LOD}}$. For the error analysis in Chapter 4, it is necessary to derive an estimate for $\|u^f\|$, in order to prove an estimate for $\|u - u_H^{\text{LOD}}\|$ and moreover, for $\|u - u_{H,k}^{\text{LOD}}\|$. The obtained result has the form

$$\|u - u_{H,k}^{\text{LOD}}\| \leq C_1 k^{d/2} \theta^k \|f\|_{L^2(\Omega)} + \|u^f\|.$$

Here, the contrast is hidden in C_1 and θ . Målqvist and Hellman used the right hand side correction to additionally enable a contrast dependency for the second term and thus, to consolidate both terms. Implied by the definition of the right hand side correction, the standard LOD changes to an alternative formulation.

3.5.1 Definition (Standard localized LOD approximation with right hand side correction)

The localized LOD approximation of the exact solution u in (2.3) with right hand side is to find $u_{H,k}^{\text{LOD,rhs}} \in V_{H,k}^{\text{LOD}}$ that satisfies

$$a(u_{H,k}^{\text{LOD,rhs}}, v) = F(v) - a(\mathcal{R}_k f, v), \quad \forall v \in V_{H,k}^{\text{LOD}}, \quad (3.9)$$

with $k \in \mathbb{N}$. The whole solution is then given by $u_{H,k}^{\text{rhs}} = u_{H,k}^{\text{LOD,rhs}} + \mathcal{R}_k f$.

3.5.2 Application to the PG-LOD

The error bounds in Chapter 4 show that already the ideal LOD method generates an error. The right hand side ensures a formulation of the PG-LOD that is only a reformulation and hence, no error occurs in its ideal version. First of all, due to the finescale splitting, we know that we can assume $u = u_H + u^f$ with $u_H \in V_H$ and $u^f \in V^f$. Thus, the weak formulation in Definition 2.1.1

$$a(u_H + u^f, v_H + v^f) = F(v_H + v^f), \quad \forall v_H \in V_H, v^f \in V^f$$

can be rewritten for all $v = v_H + v^f \in V_H \oplus V^f$ such that

$$a(u_H + u^f, v_H) = F(v_H), \quad (3.10)$$

$$a(u^f, v^f) = F(v^f) - a(u_H, v^f). \quad (3.11)$$

We intend to reinsert (3.11) into (3.10) but lose v^f as a test function. For this purpose, we apply the right hand side correction and use linearity in order to replace u^f . According to (3.10) and due to linearity, we can actually write $u^f = -Qu_H + \mathcal{R}f$. Inserting u^f into

(3.10) concludes, for all $v_H \in V_H$,

$$a(u_H - \mathcal{Q}u_H, v_H) = F(v_H) - a(\mathcal{R}f, v_H). \quad (3.12)$$

Finally, the LOD space $V_H^{\text{LOD}} = V_H - \mathcal{Q}V_H$ provides the trial space for the Petrov Galerkin formulation with right hand side correction. As we have already mentioned before, the resulting Petrov Galerkin formulation provides no error compared to the exact solution, since we know

$$\begin{aligned} u &= u_H + u^f \\ &= u_H - \mathcal{Q}u_H + \mathcal{R}f \\ &= u_H^{\text{PG}} + \mathcal{R}f. \end{aligned} \quad (3.13)$$

The resulting Petrov Galerkin method is formulated in the sequel.

3.5.2 Definition (Ideal PG-LOD approximation with right hand side correction) The ideal PG-LOD approximation with right hand side correction of u in (2.3) is to find $u^{\text{PG}} \in V^{\text{LOD}}$ such that, for all $v \in V_H$, it holds that

$$a(u^{\text{PG}}, v) = F(v) - a(\mathcal{R}f, v). \quad (3.14)$$

Due to (3.13), the full solution is given by $u = u^{\text{PG}} + \mathcal{R}f$.

Furthermore, we get the following localized version.

3.5.3 Definition (Localized PG-LOD approximation with right hand side correction) The localized PG-LOD approximation of u in (2.3) is to find $u_k^{\text{PG}} \in V_k^{\text{LOD}}$ such that, for all $v \in V_H$, it holds that

$$a(u_k^{\text{PG}}, v) = F(v) - a(\mathcal{R}_k f, v), \quad (3.15)$$

where $u_k = u_k^{\text{PG}} + \mathcal{R}_k f$ represents the full approximation.

In the following, the PG-LOD with right hand side correction plays an important role, since it provides the main tool for the novel method that is derived in Chapter 5. Moreover, the right hand side is essential for the numerical analysis in Chapter 6. Before we start with the numerical analysis for the standard LOD, we first want to point out details concerning the interpolation, which is the key tool for the multiscale splitting in the LOD.

3.6 Interpolation

In the beginning of this chapter, we used a linear surjective (quasi)-interpolation operator I_H that maps the high resolution space V to the finite element space V_H and enables the definition of the finescale space V^f . Therefore, this interpolant forms the key tool to provide the setting for the LOD. It has the general form

$$I_H v = \sum_{x \in \mathcal{N}} (I_H v)(x) \lambda_x,$$

where $I_H v(x)$ is a function that is zero for $x \notin \mathcal{N}$ and whose value for $x \in \mathcal{N}$ depends on the choice of the interpolant. Throughout this thesis, we use interpolations that satisfy the following assumptions, stated in [31]:

- (AI1) $I_H : V \rightarrow V_H$ is linear and continuous,
- (AI2) the restriction on V_H is an isomorphism,
- (AI3) the stability estimate

$$H_T^{-1} \|v - I_H v\|_{L^2(T)} + \|\nabla I_H v\|_{L^2(T)} \leq C_{I_H} \|\nabla v\|_{L^2(U(T))}, \quad (3.16)$$

holds for every $v \in V$ and $T \in \mathcal{T}_H$, with a generic constant $C_{I_H} > 0$,

- (AI4) there exists a generic constant $C'_{I_H} > 0$, which only depends on ρ , but not on the local mesh size H , such that, for all $v_H \in V_H$, there exists $v \in V$ with the properties

$$\begin{aligned} I_H(v) &= v_H, \\ \|\nabla v\|_{L^2(\Omega)} &\leq C'_{I_H} \|\nabla v_H\|_{L^2(\Omega)}, \text{ and} \\ \text{supp } v &\subset \text{supp } v_H. \end{aligned}$$

These assumptions are, at least in the classical formulation, required for the numerical analysis of the LOD and hence, they need to be satisfied for the chosen interpolant. Assumptions (AI2) and, more importantly, (AI4) are trivially satisfied if I_H is a projection. We will make use of this later on. This thesis uses two different choices of the interpolant. Firstly, we focus on the choice in the classical version. Målqvist and Peterseim chose a Clément-type interpolation. This interpolant has been introduced and analyzed in [10, Chapter 6].

3.6.1 Definition (Clément-type interpolant, see [10, Chapter 6]) For the Clément-type interpolation \mathcal{J}_H , the function $(\mathcal{J}_H v)(x)$ is defined by

$$(\mathcal{J}_H v)(x) := \frac{\left(\int_{\Omega} v \lambda_x \right)}{\left(\int_{\Omega} \lambda_x \right)}.$$

Thus, the Clément-type interpolant consists of weighted averages over the nodal patches $U(x) = \text{supp } \lambda_x$. We refer to [10] and [31] for further details and for a proof for assumptions (AI1)-(AI3). Assumption (AI4) has been proven for example in [28] or [16]. In the classical error analysis, (AI3) and (AI4) is applied a couple of times, whereby the latter is only necessary, because \mathcal{J}_H is not a projection operator. The second choice of the interpolant that is discussed throughout this thesis is the L^2 -projection interpolant.

3.6.2 Definition (L^2 -projection, see [20]) For every $v \in V$, we define the L^2 -projection $\mathcal{I}_H : V \rightarrow V_H$ by

$$\mathcal{I}_H v = \sum_{x \in \mathcal{N}} (\mathcal{I}_H v)(x) \lambda_x,$$

where we define, for every node $x \in \mathcal{N}$,

$$(\mathcal{I}_H v)(x) = (P_x v)(x).$$

The operator P_x defines a projection $P_x v \in V_H|_{U(x)}$ such that

$$\int_{U(x)} P_x v w_H = \int_{U(x)} v w_H,$$

for all $w_H \in V_H|_{U(x)}$.

Due to $\mathcal{I}_H(v_H) = v_H$, \mathcal{I}_H defines a projection on V_H . We conclude that (AI2) and (AI4) are satisfied automatically, since we can choose $v = v_H$ for the latter. Assumption (AI1) holds by definition. The stability estimate in assumption (AI3) has been proven in [31]. The advantage that arises with the help of a projection interpolant is discussed and presented in the following chapter. For other choices of the interpolant, we recommend for instance [26]. Now, we are prepared to dedicate ourselves to the numerical analysis.

4 Error analysis for the standard LOD

This chapter deals with the error bound for the standard LOD method, presented in Chapter 3. To attain this error bound, the proof for the exponential decay of the finescale correctors is crucial and is also the justification for the localization. Basically, we intend to summarize the progress that has been made, starting from the classical up until the most recent knowledge. Recently, various suggestions for the proof have been made. In the first part of this chapter, we state the main approaches of the classical proof, made by Målqvist and Peterseim in [28]. Furthermore, we point out the main details of the proof so that the reader gets an idea which assumptions have been used. Subsequently, we discuss improvements proposed by Målqvist and Peterseim themselves as well as Henning and Hellman (see [16], [30] and [26]). At the end of the chapter, we present a complete proof based on the most recent knowledge.

4.1 Classical proof

This section is based on [28]. The goal is to describe the main results for the proof of Theorem 4.1.5, an error bound of the type

$$\|u - u_{H,k}^{\text{lod}}\| \leq C \|Hf\|_{L^2(\Omega)} + C' H^{-1} \theta^k \|F\|_{H^{-1}(\Omega)},$$

where $u_{H,k}^{\text{lod}} \in V_{H,k}^{\text{lod}}$ is the standard LOD approximation in Definition 3.3.2. To accomplish this, we first of all observe an error bound for the ideal method, formulated in the following lemma. This lemma provides an overview about the techniques that are used in the whole numerical analysis. Moreover, we introduce an overlapping constant C_{ol} , which constitutes an important tool in the classical proof.

4.1.1 Lemma (see Lemma 3.1 in [28]) Let $u \in V$ be an exact solution of (2.3) and let $u_H^{\text{lod}} \in V_H^{\text{lod}}$ be the solution of the approximation of the ideal LOD method (3.1). Then we have

$$\|u - u_H^{\text{lod}}\| \leq C_{\text{ol}}^{1/2} C_{\mathcal{J}_H} \alpha^{-1/2} \|Hf\|_{L^2(\Omega)}, \quad (4.1)$$

where $C_{\mathcal{J}_H}$ is related to the interpolation estimate and α to the boundedness of A . The constant $C_{\text{ol}} > 0$ arises from an overlap.

PROOF: First of all, and representative for the sequel of this thesis, we want to note that the uniformity bounds α and β in (2.1) and (2.2) enable to deduce the estimates

$$\begin{aligned} \|A\nabla v\|_{L^2(\Omega)} &\leq \beta \|\nabla v\|_{L^2(\Omega)}, \\ \|\nabla v\|_{L^2(\Omega)} &\leq \alpha^{-1} \|A\nabla v\|_{L^2(\Omega)}, \end{aligned} \quad (4.2)$$

for every $v \in V$. In order to prove the lemma, we recall the interpolation estimate (3.16) for the interpolation operator \mathcal{J}_H . From the orthogonal multiscale splitting, we know that the exact solution $u \in V$ can be decomposed in

$$u = u_H^{\text{lod}} + u^f.$$

Therefore, we can write $u - u_H^{\text{lod}} = u^f$. Furthermore, we know that $\mathcal{J}_H u^f = 0$, since $u^f \in V^f$. This allows the observation

$$\begin{aligned} \|u - u_H^{\text{lod}}\|^2 &= \|u^f\|^2 \\ &= a(u^f, u^f) \\ &= F(u^f) \\ &= \int_{\Omega} f u^f \\ &= \sum_{T \in \mathcal{T}_H} \int_T f u^f \\ &\stackrel{\text{C.S.}}{\leq} \sum_{T \in \mathcal{T}_H} \|f\|_{L^2(T)} \|u^f - \underbrace{\mathcal{J}_H u^f}_{=0}\|_{L^2(T)} \\ &\stackrel{(3.16)}{\leq} \sum_{T \in \mathcal{T}_H} \|f\|_{L^2(T)} C_{\mathcal{J}_H} H \|\nabla u^f\|_{L^2(U(T))}. \end{aligned} \quad (4.3)$$

Remember Young's inequality

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q},$$

where $1 < p < \infty$ and q such that $\frac{1}{p} + \frac{1}{q} = 1$. This inequality can easily be deduced by using the concavity of the logarithm function. In particular, setting $p = q = 2$, $a' := \frac{a}{\sqrt{\varepsilon}}$ and $b' := \sqrt{\varepsilon}b$, for any $\varepsilon > 0$, reveals

$$ab \leq \frac{1}{2\varepsilon} a^2 + \frac{\varepsilon}{2} b^2. \quad (4.4)$$

Applying this to (4.3) implies

$$\begin{aligned}
 \|\|u - u_H^{\text{lod}}\|\|^2 &\stackrel{(4.3)}{\leq} \sum_{T \in \mathcal{T}_H} \underbrace{\frac{C_{\mathcal{J}_H}}{\alpha^{1/2}} \|Hf\|_{L^2(T)}}_{=a} \underbrace{\|A^{1/2} \nabla u^f\|_{L^2(U(T))}}_{=b} \\
 &\stackrel{(4.4)}{\leq} \frac{C_{\mathcal{J}_H}^2}{2\varepsilon\alpha} \|Hf\|_{L^2(\Omega)}^2 + \frac{\varepsilon}{2} \sum_{T \in \mathcal{T}_H} \|A^{1/2} \nabla u^f\|_{L^2(U(T))}^2,
 \end{aligned} \tag{4.5}$$

for any $\varepsilon > 0$. With respect to the sum in the right hand side, we realize that the proof is done by replacing $U(T)$ by T . This can be achieved by setting ε appropriately. We know that our finite element mesh \mathcal{T}_H is regularly shaped with factor ρ . Thus, we can extract a constant C_{ol} which is able to control the number of elements, covered by $U(T)$, such that it is bounded by C_{ol} . Setting $\varepsilon = C_{\text{ol}}^{-1}$ implies

$$\begin{aligned}
 \|\|u - u_H^{\text{lod}}\|\|^2 &\stackrel{(4.5)}{\leq} \frac{C_{\mathcal{J}_H}^2 C_{\text{ol}}}{2\alpha} \|Hf\|_{L^2(\Omega)}^2 + \frac{1}{2C_{\text{ol}}} \sum_{T \in \mathcal{T}_H} \|A^{1/2} \nabla u^f\|_{L^2(U(T))}^2 \\
 &\leq \frac{C_{\mathcal{J}_H}^2 C_{\text{ol}}}{2\alpha} \|Hf\|_{L^2(\Omega)}^2 + \frac{1}{2} \sum_{T \in \mathcal{T}_H} \|A^{1/2} \nabla u^f\|_{L^2(T)}^2 \\
 &= \frac{C_{\mathcal{J}_H}^2 C_{\text{ol}}}{2\alpha} \|Hf\|_{L^2(\Omega)}^2 + \frac{1}{2} \|A^{1/2} \nabla u^f\|_{L^2(\Omega)}^2.
 \end{aligned}$$

Hence, we obtain the assertion,

$$\|\|u - u_H^{\text{lod}}\|\| \leq C_{\text{ol}}^{1/2} C_{\mathcal{J}_H} \alpha^{-1/2} \|Hf\|_{L^2(\Omega)}.$$

□

This proof makes use of the features of the mesh since the existence of the overlapping constant C_{ol} is necessary. Målqvist and Peterseim used this constant a couple of times in their proofs in order to control the domain of the norms. For further observations, we introduce the so called cut-off functions $\eta_x^{m,M}$, defined in the following.

4.1.2 Definition (Cut-off functions) For $x \in \mathcal{N}$ and $m < M \in \mathbb{N}$, let $\eta_x^{m,M} : \Omega \rightarrow [0, 1]$ be a continuous, weakly differentiable function which satisfies

$$\begin{aligned} (\eta_x^{m,M}) \Big|_{\omega_{x,m}} &= 0, \\ (\eta_x^{m,M}) \Big|_{\Omega \setminus \omega_{x,M}} &= 1, \text{ and} \\ \forall T \in \mathcal{T}_H, \|\nabla \eta_x^{m,M}\|_{L^\infty(T)} &\leq C_{\text{co}}(M - m)^{-1} H_T^{-1}, \end{aligned} \quad (4.6)$$

for some constant C_{co} dependent on only the shape parameter ρ .

We omit a detailed definition for the constants that arise in the next statements. The important aspect is that the constants only depend on $C_{\text{co}}, C_{\text{ol}}, C_{\mathcal{J}_H}, C'_{\mathcal{J}_H}, \alpha, \beta$ and ρ that are related to the cut-off in Definition 4.1.2, the overlap in Lemma 4.1.1, the assumptions for the interpolant (AI3) resp. (AI4), the boundaries for the diffusion and the shape of the elements. Clearly, every constant C_k ($k \in \mathbb{N}$) has the form $C_k = C(C_{\text{co}}, C_{\text{ol}}, C_{\mathcal{J}_H}, C'_{\mathcal{J}_H}, \alpha, \beta, \rho)$. Most importantly, the constants do not depend directly on $|\mathcal{N}|, x, k, l$ or H . The next lemma states the exponential decay, which contains the main work for the classical proof and leads to the desired error bound.

4.1.3 Lemma (Målqvist and Peterseim in [28]) For all $x \in \mathcal{N}$, $k \geq 2 \in \mathbb{N}$ and $l \geq 3 \in \mathbb{N}$, the estimate

$$\|\phi_x - \phi_{x,lk}\| \leq C_2(C_1/l)^{\frac{k-2}{2}} \|\phi_x\|_{\omega_{x,l}} \quad (4.7)$$

holds with constants C_1, C_2 , where $\|\cdot\|_{\omega} := \|A^{1/2} \nabla \cdot\|_{L^2(\omega)}$.

PROOF: See Lemma 3.4 in [28]. □

The lemma is crucial for the localization and for the derivation of error bounds. However, the proof is very technical in its classical version. Appropriated cut-off functions and their properties build the groundwork for the conclusions and the interpolation estimate (3.16) plays an important role. Furthermore, the overlap constant C_{ol} aims to control the domains. The used Clément type interpolant \mathcal{J}_H , defined in Definition 3.6.1, is not a projection, which constitutes a disadvantage. Assumption (AI4) is crucial and utilized multiple times. In particular, some observation require an element of V^f . If $v \notin V^f$, assumption (AI4) provides the existence of a b such that $\mathcal{J}_H(b) = \mathcal{J}_H(v)$. The procedure is to define $v' := b - v \in V^f$ and to take advantage of the support property as well as the given estimate to continue the proof. These observations become quite technical. If the interpolant I_H is a projection, i.e, $I_H(v_H) = v_H$, for all $v_H \in V_H$, we can easily define $v' := v - I_H(v)$ for $v \notin V^f$. This automatically yields $I_H(v') = 0$ and therefore $v' \in V^f$. This approach avoids a lot of technical estimates and, moreover, no workaround like

assumption (AI4) is necessary. The following lemma is also required to prove the main result.

4.1.4 Lemma (Målqvist and Peterseim [28]) There exists a constant C_3 such that

$$\left\| \sum_{x \in \mathcal{N}} v(x)(\phi_x - \phi_{x,lk}) \right\|^2 \leq C_3 (lk)^d \sum_{x \in \mathcal{N}} v^2(x) \|\phi_x - \phi_{x,lk}\|^2. \quad (4.8)$$

PROOF: See Lemma 3.5 in [28]. □

Once again, assumption (AI4), overlapping results and a cut-off function are applied to attain the assertion. The observations are quite similar to the ones in Lemma 4.1.3 and Lemma 4.1.1. Now, we formulate the main result of the classical proof.

4.1.5 Theorem (Målqvist and Peterseim in [28]) Let $u \in V$ be a solution of (2.3) and let $2 \leq k \in \mathbb{N}$ and $3 \leq l \in \mathbb{N}$. Let $u_H^{\text{lod}} \in V_H^{\text{lod}}$ be the solution of the localized LOD approximation (3.3). Then

$$\begin{aligned} \|u - u_{H,lk}^{\text{lod}}\| &\leq C_4 \|H_T^{-1}\|_{L^\infty(\Omega)} (lk)^{d/2} (C_1/l)^{\frac{k-2}{2}} \|F\|_{H^{-1}(\Omega)} \\ &\quad + C_{\text{ol}}^{1/2} C_{\mathcal{J}_H} \alpha^{-1/2} \|Hf\|_{L^2(\Omega)}. \end{aligned}$$

PROOF: See Theorem 3.6 in [28]. □

This proof clearly unites Lemma 4.1.1, Lemma 4.1.3 and Lemma 4.1.4. The achieved estimate contains a factor H^{-1} . This issue can be fixed by choosing the parameter k large enough, for instance, proportional to $\log(1/H)$. We conclude that for l such that $C_1 < l$ and sufficiently big k the method allows for a good approximation. However, we also recognize that there is still room for improvements to yield a better convergence rate. Hereof, the cause of H^{-1} needs to be ascertained. An insight in the proof reveals that especially the application of Lemma 4.1.4 induces problems. The reason why we have to apply this lemma is naturally hidden in the very beginning of the proof, the definition of u_H^{lod} ,

$$u_H^{\text{lod}} := \sum_{x \in \mathcal{N}} u_H^{\text{lod}}(x)(\lambda_x - \phi_x).$$

This representation is due to the nodal patch localization in the standard LOD. In Chapter 3, we have already mentioned that the element patch localization yields a better result in terms of convergence. It is also helpful to avoid the overlapping constant as well as to simplify several observations. In the following section, we present a proof which is based on the most recent improvements of the classical proof.

4.2 New formulation

In the previous section, we briefly demonstrated the room for improvements of the classical LOD localization, interpolation and its numerical analysis. Besides, we have already suggested alternatives in terms of the interpolant and the localization. In the following, we collect all new ideas, to formulate a new theorem for the convergence rate and moreover prove it with novel techniques. The main result of these new approaches is formulated in the next theorem. It can be understood as an improvement of Theorem 4.1.5. This result and the proof are mainly based on [26].

4.2.1 Theorem (Error bound for the LOD approximation with element patch localization) Let $u_{H,k}^{\text{LOD}} \in V_{H,k}^{\text{LOD}}$ be the solution of (3.6) and $u \in V$ the exact solution of (2.3). Then the estimate

$$\|u - u_{H,k}^{\text{LOD}}\| \leq \left(C_1 k^{d/2} \theta^k + HC_2 \right) \|f\|_{L^2(\Omega)}$$

holds, where C_1, C_2 and $0 < \theta < 1$ are positive constants, independent of H and k .

Most of the steps are proven similarly to Theorem 4.1.5. For the sake of completeness, we show every statement in this proof, even though they are similar to the ones in the previous section. We have already seen that the key tool for the method is provided by the interpolant that maps a function from the entire space V to the coarse finite element space V_H . Hence, the interpolant is responsible for the definition of the finescale space $V^f = \ker(I_H)$. For this purpose, the classical LOD uses the Clément-type interpolant \mathcal{J}_H satisfying assumptions (AI3) and (AI4). From now on, we use the L^2 -projection interpolant \mathcal{I}_H , defined in Definition 3.6.2. Clearly, \mathcal{I}_H satisfies the projection property, which means that $\mathcal{I}_H(v_H) = v_H$. Assumption (AI4) is no longer necessary for our approaches, since the projection property enables to drop this assumption, whenever it was necessary in the classical proof. This improvement was already explained above. Certainly, the stability estimate (3.16) in assumption (AI3) is still crucial for our observations. Regarding this assumption, we want to remark that we will use (3.16) not just on a single $T \in \mathcal{T}_H$, but also on arbitrary sets $\omega \subseteq \Omega$ that are a union of arbitrary many $T \in \mathcal{T}_H$ (T -union), or even the whole space Ω . For this purpose, we need to use a sum over all $T \subseteq \omega$. According to (3.16), we get the corresponding patch for every single $T \in \mathcal{T}_H$. Therefore, an overlapping effect occurs. This overlap is clearly dependent on ω and on the triangulation \mathcal{T}_H , but it is just linearly. It gets the worst for $\omega = \Omega$ and it does not exist for $\omega = T$. We will hide this effect by setting $C_{\mathcal{I}_H}$ such that $C_{\mathcal{I}_H}^\omega \leq C_{\mathcal{I}_H}$, for all $\omega \subseteq \Omega$. Note that this constant has not the same usage as the overlapping constant C_{ol} in the classical proof. This constant helped out in order to decrease the patch size, whereas using only $C_{\mathcal{I}_H}$ omits the sake of reducing the patch patch size. In terms of the localization, we choose the element

patch localization, introduced in Section 3.3.2. Clearly, we use the LOD approximation $u_{H,k}^{\text{LOD}} \in V_{H,k}^{\text{LOD}}$, defined in Definition 3.3.3. Motivated by the element patch localization, we define novel cut-off functions.

4.2.2 Definition (Cut-off functions for elements) For $T \in \mathcal{T}_H$ and $0 < k \in \mathbb{N}$, we define $\eta_{T,k}$ such that the following properties are satisfied

$$\begin{aligned} \eta_{T,k}(x) &= 0, & \forall x \in U_{k-1}(T), \\ \eta_{T,k}(x) &= 1, & \forall x \in \Omega \setminus U_k(T), \\ \|\nabla \eta_{T,k}\|_{L^\infty(\Omega)} &\leq C_{co} H^{-1}. \end{aligned} \quad (4.9)$$

These cut-off functions are slightly different to the ones we used in the classical proof. They are defined for every $T \in \mathcal{T}_H$ and no longer for every node $x \in \mathcal{N}$. We separate the proof of Theorem 4.2.1 in several lemmas. The first lemma contains an estimate that can be understood as a Poincaré-type inequality.

4.2.3 Lemma (Poincaré-type inequality for finescale elements) For $v^f \in V^f$ and for every T -union $\omega \subseteq \Omega$, it holds that

$$\|A^{1/2}v^f\|_{L^2(\omega)} \leq C_3 H \|A^{1/2}\nabla v^f\|_{L^2(U(\omega))}, \quad (4.10)$$

where $C_3 = \beta^{1/2}\alpha^{-1/2}C_{\mathcal{I}_H}$.

PROOF: By definition, we have $\mathcal{I}_H(v^f) = 0$, for any $v^f \in V^f$. Using (3.16) yield

$$\begin{aligned} \|A^{1/2}v^f\|_{L^2(\omega)} &\leq \beta^{1/2}\|v^f\|_{L^2(\omega)} \\ &\leq \beta^{1/2}\|v^f - \mathcal{I}_H(v^f)\|_{L^2(\omega)} \\ &\stackrel{(3.16)}{\leq} \beta^{1/2} C_{\mathcal{I}_H} H \|\nabla v^f\|_{L^2(U(\omega))} \\ &\leq \beta^{1/2}\alpha^{-1/2} C_{\mathcal{I}_H} H \|A^{1/2}\nabla v^f\|_{L^2(U(\omega))} \\ &= C_3 H \|A^{1/2}\nabla v^f\|_{L^2(U(\omega))}. \end{aligned}$$

□

The next lemma states two other inequalities that are very helpful for further observations. We state and prove them in the following.

4.2.4 Lemma (Energy of interpolation for v^f after cut-off) Let η be a cut-off function with the same properties like $\eta_{T,k}$, defined in Definition 4.2.2, or even $1 - \eta_{T,k}$. Then, for

any fixed $T \in \mathcal{T}_H$, $0 < k \in \mathbb{N}$, $v^f \in V^f$ and for every T -union $\omega \subseteq \Omega$, it holds that

$$\|A^{1/2} \nabla(\eta v^f)\|_{L^2(\omega)} \leq C'_4 \|A^{1/2} \nabla v^f\|_{L^2(U(\omega))}, \quad (4.11)$$

$$\|A^{1/2} \nabla \mathcal{I}_H(\eta v^f)\|_{L^2(\omega)} \leq C_4 \|A^{1/2} \nabla v^f\|_{L^2(U_2(\omega))}, \quad (4.12)$$

where $C'_4 = \beta^{1/2} \alpha^{-1/2} (C_{\mathcal{I}_H} C_{\text{co}} + 1)$ and $C_4 = C'_4 C_{\mathcal{I}_H}$.

PROOF: First, we want to emphasize that the property (4.9), stated in Definition 4.2.2 and $\|\eta\|_{L^\infty(\Omega)} = 1$ is satisfied for $\eta = \eta_{T,k}$ as well as for $\eta = 1 - \eta_{T,k}$.

For (4.11), we derive, for every $\omega \subseteq \Omega$,

$$\begin{aligned} \|A^{1/2} \nabla(\eta v^f)\|_{L^2(\omega)} &\leq \beta^{1/2} \|\nabla(\eta v^f)\|_{L^2(\omega)} \\ &\leq \beta^{1/2} \left(\|\nabla \eta v^f\|_{L^2(\omega)} + \|\eta \nabla v^f\|_{L^2(\omega)} \right) \\ &\leq \beta^{1/2} \left(\|\nabla \eta\|_{L^\infty(\Omega)} \|v^f\|_{L^2(\omega)} + \|\eta\|_{L^\infty(\Omega)} \|\nabla v^f\|_{L^2(\omega)} \right) \\ &\stackrel{(4.9)}{\leq} \beta^{1/2} \left(C_{\text{co}} H^{-1} \|v^f - \mathcal{I}_H(v^f)\|_{L^2(\omega)} + \|\nabla v^f\|_{L^2(U(\omega))} \right) \\ &\stackrel{(3.16)}{\leq} \beta^{1/2} (C_{\mathcal{I}_H} C_{\text{co}} + 1) \|\nabla v^f\|_{L^2(U(\omega))} \\ &\leq \beta^{1/2} \alpha^{-1/2} (C_{\mathcal{I}_H} C_{\text{co}} + 1) \|A^{1/2} \nabla v^f\|_{L^2(U(\omega))} \\ &= C'_4 \|A^{1/2} \nabla v^f\|_{L^2(U(\omega))}, \end{aligned}$$

Starting again from the second step in the previous estimate concludes

$$\begin{aligned} \|A^{1/2} \nabla \mathcal{I}_H(\eta v^f)\|_{L^2(\omega)} &\leq \beta^{1/2} \|\nabla \mathcal{I}_H(\eta v^f)\|_{L^2(\omega)} \\ &\stackrel{(3.16)}{\leq} \beta^{1/2} C_{\mathcal{I}_H} \|\nabla(\eta v^f)\|_{L^2(U(\omega))} \\ &\stackrel{(4.11)}{\leq} C_4 \|A^{1/2} \nabla v^f\|_{L^2(U_2(\omega))}, \end{aligned}$$

which shows the assertion. \square

For a particular case of ω , such as a ring with the form $U_k \setminus U_{k-1}$, we are able to make use of the properties of the cut-off functions. The resulting improvement is essential for the estimates below and enables a better convergence rate than the one we would reach with using the more rough estimate (4.12). We formulate and proof it in the sequel.

4.2.5 Lemma (Energy of interpolation for v^f after cut-off on rings) Let $\eta := \eta_{T,k}$ be a cut-off function, defined in Definition 4.2.2, for one fixed $T \in \mathcal{T}_H$ and $0 < k \in \mathbb{N}$. Then it

holds for every $v^f \in V^f$ that

$$\|A^{1/2}\nabla\mathcal{I}_H(\eta v^f)\|_{L^2(U_{k+1}\setminus U_{k-2})} \leq C_4 \min\left\{\|A^{1/2}\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-2})}, \|A^{1/2}\nabla v^f\|_{L^2(U_{k+1}\setminus U_{k-3})}\right\}, \quad (4.13)$$

where C_4 is defined as in Lemma 4.2.4.

PROOF: We proceed analogously to Lemma 4.2.4, but this time, we use every feature of η . We obtain

$$\begin{aligned} & \|A^{1/2}\nabla\mathcal{I}_H(\eta v^f)\|_{L^2(U_{k+1}\setminus U_{k-2})} \\ & \stackrel{(3.16)}{\leq} \beta^{1/2} C_{\mathcal{I}_H} \|\nabla(\eta v^f)\|_{L^2(U_{k+2}\setminus U_{k-3})} \\ & \leq \beta^{1/2} C_{\mathcal{I}_H} \left(\|\nabla\eta v^f\|_{L^2(U_{k+2}\setminus U_{k-3})} + \|\eta\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-3})} \right) \\ & \stackrel{(*)}{=} \beta^{1/2} C_{\mathcal{I}_H} \left(\|\nabla\eta v^f\|_{L^2(U_k\setminus U_{k-1})} + \|\eta\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-1})} \right) \\ & \leq \beta^{1/2} C_{\mathcal{I}_H} \left(\|\nabla\eta\|_{L^\infty(\Omega)} \|v^f\|_{L^2(U_k\setminus U_{k-1})} + \|\eta\|_{L^\infty(\Omega)} \|\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-1})} \right) \quad (4.14) \\ & \stackrel{(4.9)}{\leq} \beta^{1/2} C_{\mathcal{I}_H} \left(C_{\text{co}} H^{-1} \|v^f - \mathcal{I}_H(v^f)\|_{L^2(U_k\setminus U_{k-1})} + \|\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-1})} \right) \\ & \stackrel{(3.16)}{\leq} \beta^{1/2} C_{\mathcal{I}_H} \left(C_{\mathcal{I}_H} C_{\text{co}} \|\nabla v^f\|_{L^2(U_{k+1}\setminus U_{k-2})} + \|\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-1})} \right) \\ & \leq \beta^{1/2} \alpha^{-1/2} (C_{\mathcal{I}_H}^2 C_{\text{co}} + C_{\mathcal{I}_H}) \|A^{1/2}\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-2})} \\ & = C_4 \|A^{1/2}\nabla v^f\|_{L^2(U_{k+2}\setminus U_{k-2})}. \end{aligned}$$

In (*), we used the support properties of the cut-off function η . In particular, $\text{supp}(\eta) = \Omega \setminus U_{k-1}$ and $\text{supp}(\nabla\eta) = U_k \setminus U_{k-1}$. Using $\mathcal{I}_H(v^f) = 0$ yields

$$\|A^{1/2}\nabla\mathcal{I}_H(\eta v^f)\|_{L^2(U_{k+1}\setminus U_{k-2})} = \|A^{1/2}\nabla\mathcal{I}_H((1-\eta)v^f)\|_{L^2(U_{k+1}\setminus U_{k-2})}.$$

Proceeding analogously to (4.14), but using the properties of $1-\eta$ instead of η , concludes

$$\|A^{1/2}\nabla\mathcal{I}_H(\eta v^f)\|_{L^2(U_{k+1}\setminus U_{k-2})} \leq C_4 \|A^{1/2}\nabla v^f\|_{L^2(U_{k+1}\setminus U_{k-3})},$$

which attains the assertion. \square

The following lemma is an improvement of Lemma 4.1.3. It is the replacement for the most difficult part in the classical proof.

4.2.6 Lemma (Energy of the operator \mathcal{Q}^T outside of a patch with size k) For any $T \in \mathcal{T}_H$, $v \in V$ and for $k \geq 3$, the correction $q^T := \mathcal{Q}^T v \in V^f$, defined in Section 3.3.2, fulfills

$$\|A^{1/2}\nabla q^T\|_{L^2(\Omega \setminus U_k(T))} \leq \theta^k \|A^{1/2}\nabla q^T\|_{L^2(\Omega)}, \quad (4.15)$$

where $0 < \theta < 1$.

PROOF: Choose $m := k - 1 \geq 2$. To keep clarity, we define $U_m := U_m(T)$ and $\eta := \eta_m^T$. We start with the left hand side and derive

$$\begin{aligned}
 \|A^{1/2}\nabla q^T\|_{L^2(\Omega \setminus U_m)}^2 &= \int_{\Omega \setminus U_m} A\nabla q^T \cdot \nabla q^T \\
 &\stackrel{\text{Def. } \eta}{\leq} \int_{\Omega \setminus U_{m-1}} A\nabla q^T \cdot \eta \nabla q^T \\
 &= \int_{\Omega \setminus U_{m-1}} A\nabla q^T \cdot \eta \nabla q^T + \int_{\Omega \setminus U_{m-1}} A\nabla q^T \cdot q^T \nabla \eta - \int_{\Omega \setminus U_{m-1}} A\nabla q^T \cdot q^T \nabla \eta \\
 &= \underbrace{\int_{\Omega \setminus U_{m-2}} A\nabla q^T \cdot \nabla(\eta q^T)}_{=I} - \underbrace{\int_{\Omega \setminus U_{m-1}} A\nabla q^T \cdot q^T \nabla \eta}_{=II}.
 \end{aligned}$$

For the first term, we take advantage of the circumstance that \mathcal{I}_H is a projection and therefore, $v' := \eta q^T - \mathcal{I}_H(\eta q^T) \in V^f$. Due to the interpolation in v' and $\text{supp}(\eta q^T) = \Omega \setminus U_{m-1}$, we know $\text{supp}(v') = \Omega \setminus U_{m-2}$. Setting $m \geq 2$ implies that v' has no support in $T \subseteq U_{m-2}$. This is also the reason why we changed the domain for the integral in I from $\Omega \setminus U_{m-1}$ to $\Omega \setminus U_{m-2}$. This procedure makes no difference because of the definition for η and $\nabla \eta$. With respect to the definition of the corrector Q^T , we know

$$\begin{aligned}
 I &= \int_{\Omega \setminus U_{m-2}} A\nabla q^T \cdot \nabla(\eta q^T) \\
 &= \int_{\Omega \setminus U_{m-2}} A\nabla q^T \cdot \nabla v' + \int_{\Omega \setminus U_{m-2}} A\nabla q^T \cdot \nabla(\mathcal{I}_H(\eta q^T)) \\
 &= \int_{\Omega} A\nabla q^T \cdot \nabla v' + \int_{\Omega \setminus U_{m-2}} A\nabla q^T \cdot \nabla(\mathcal{I}_H(\eta q^T)) \\
 &\stackrel{(3.4)}{=} \int_T A\nabla v \cdot \nabla v' + \int_{\Omega \setminus U_{m-2}} A\nabla q^T \cdot \nabla(\mathcal{I}_H(\eta q^T)) \\
 &= \int_{\Omega \setminus U_{m-2}} A\nabla q^T \cdot \nabla(\mathcal{I}_H(\eta q^T)).
 \end{aligned}$$

We are also able to restrict the domain of the resulting integral, since $q^T \in V^f$, we get $\text{supp}(\mathcal{I}_H(\eta q^T)) = U_{m+1} \setminus U_{m-2}$. Using Cauchy-Schwarz derives

$$\begin{aligned}
 |I| &= \left| \int_{U_{m+1} \setminus U_{m-2}} A\nabla q^T \cdot \nabla(\mathcal{I}_H(\eta q^T)) \right| \\
 &\stackrel{\text{C.S.}}{\leq} \|A^{1/2}\nabla q^T\|_{L^2(U_{m+1} \setminus U_{m-2})} \|A^{1/2}\nabla \mathcal{I}_H(\eta q^T)\|_{L^2(U_{m+1} \setminus U_{m-2})} \\
 &\stackrel{(4.13)}{\leq} C_4 \|A^{1/2}\nabla q^T\|_{L^2(U_{m+1} \setminus U_{m-2})} \|A^{1/2}\nabla q^T\|_{L^2(U_{m+1} \setminus U_{m-3})}.
 \end{aligned}$$

For the term II , we have to recall a few inequalities from above and that, by definition, $\text{supp}(\nabla\eta) = U_m \setminus U_{m-1}$. We observe

$$\begin{aligned}
 |II| &= \left| \int_{U_m \setminus U_{m-1}} A \nabla q^T \cdot q^T \nabla \eta \right| \\
 &\stackrel{\text{C.S.}}{\leq} \|A^{1/2} \nabla q^T\|_{L^2(U_m \setminus U_{m-1})} \|A^{1/2} q^T \nabla \eta\|_{L^2(U_m \setminus U_{m-1})} \\
 &\leq \|A^{1/2} \nabla q^T\|_{L^2(U_m \setminus U_{m-1})} \|A^{1/2} q^T\|_{L^2(U_m \setminus U_{m-1})} \|\nabla \eta\|_{L^\infty(\Omega)} \\
 &\stackrel{(4.9)}{\leq} C_{\text{co}} H^{-1} \|A^{1/2} \nabla q^T\|_{L^2(U_m \setminus U_{m-1})} \|A^{1/2} q^T\|_{L^2(U_m \setminus U_{m-1})} \\
 &\stackrel{(4.10)}{\leq} C_{\text{co}} C_3 \|A^{1/2} \nabla q^T\|_{L^2(U_m \setminus U_{m-1})} \|A^{1/2} \nabla q^T\|_{L^2(U_{m+1} \setminus U_{m-2})}.
 \end{aligned}$$

With $c := C_4 + C_{\text{co}} C_3$, I and II add up to the estimate

$$\begin{aligned}
 \|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_{m+1})}^2 &\leq \|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_m)}^2 \\
 &\stackrel{I \& II}{\leq} c \|A^{1/2} \nabla q^T\|_{L^2(U_{m+1} \setminus U_{m-3})}^2 \\
 &= c \left(\|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_{m-3})}^2 - \|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_{m+1})}^2 \right),
 \end{aligned}$$

which results in

$$\|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_k)}^2 \leq \frac{c}{1+c} \|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_{k-4})}^2. \quad (4.16)$$

This estimate gives some information about the decay within 4-layers. Starting from $k \geq 4$ and applying (4.16) correspondingly often yield

$$\|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_k)}^2 \leq \left(\frac{c}{1+c} \right)^{\frac{k}{4}} \|A^{1/2} \nabla q^T\|_{L^2(\Omega)}^2. \quad (4.17)$$

The assertion follows with $\theta := \left(\frac{c}{1+c} \right)^{\frac{1}{8}}$. \square

We now use this lemma and see the advantage of the novel corrector operator. The following lemma is more or less the replacement for Lemma 4.1.3 and Lemma 4.1.4.

4.2.7 Lemma (Truncation error for correctors) Let $q := \mathcal{Q}v$ be the corrector operator and $q_k := \mathcal{Q}_k v$ the element patch localized corrector operator, defined in Section 3.3.2. Then for $k \geq 3 \in \mathbb{N}$ and $v \in V$, it holds that

$$\|A^{1/2} \nabla(q - q_k)\|_{L^2(\Omega)} \leq C_Q k^{d/2} \theta^k \|A^{1/2} \nabla q\|_{L^2(\Omega)}, \quad (4.18)$$

where $0 < \theta < 1$ are defined in Lemma 4.2.6. $C_Q = 2(C_4 + C'_4)C_d \theta^{-3}$ contains previously used constants and C_d dependent on the dimension d .

PROOF: Let $q_k^T := \mathbf{Q}_k^T v$ and

$$w := q - q_k = \sum_{T \in \mathcal{T}_H} (q^T - q_k^T).$$

Our purpose is to involve the interpolation and therefore, the definition of the correctors. Set, for all $T \in \mathcal{T}_H$,

$$\bar{w} := \eta_{k+2}^T w - \mathcal{I}_H(\eta_{k+2}^T w).$$

As in the previous lemma, we make use of \mathcal{I}_H being a projection and conclude $\bar{w} \in V^f = \ker(\mathcal{I}_H)$. The index $k + 2$ makes sure that the support of \bar{w} and the support of $q^T - q_k^T$ have an empty cut set. Hence, we use

$$\int_{\Omega} A \nabla \bar{w} \cdot \nabla (q^T - q_k^T) = 0.$$

We observe

$$\begin{aligned} \|A^{1/2} \nabla w\|_{L^2(\Omega)}^2 &= \sum_{T \in \mathcal{T}_H} \int_{\Omega} A \nabla w \cdot \nabla (q^T - q_k^T) \\ &= \sum_{T \in \mathcal{T}_H} \left(\int_{\Omega} A \nabla w \cdot \nabla (q^T - q_k^T) - \int_{\Omega} A \nabla \bar{w} \cdot \nabla (q^T - q_k^T) \right) \\ &= \sum_{T \in \mathcal{T}_H} \int_{\Omega} A \nabla \left((1 - \eta_{k+2}^T) w + \mathcal{I}_H(\eta_{k+2}^T w) \right) \cdot \nabla (q^T - q_k^T) \\ &\stackrel{\text{C.S.}}{\leq} \sum_{T \in \mathcal{T}_H} \left(\underbrace{\|A^{1/2} \nabla \left((1 - \eta_{k+2}^T) w \right)\|_{L^2(\Omega)}}_{=:I} + \underbrace{\|A^{1/2} \nabla \mathcal{I}_H(\eta_{k+2}^T w)\|_{L^2(\Omega)}}_{=:II} \right) \\ &\quad \cdot \underbrace{\|A^{1/2} \nabla (q^T - q_k^T)\|_{L^2(\Omega)}}_{=:III}. \end{aligned}$$

For the first two terms, we only need the estimates that are used and introduced in previous proofs. The third term is important, since we apply the previous lemma. It attains an assertion for the energy of the q^T corrector. We drop the patch index, $U_k := U_k(T)$, and we need to remember previous estimates. Thus, for I , we have

$$\begin{aligned} I &= \|A^{1/2} \nabla \left((1 - \eta_{k+2}^T) w \right)\|_{L^2(\Omega)} \\ &\stackrel{\text{Def.}}{\leq} \eta_{k+2}^T \|A^{1/2} \nabla \left((1 - \eta_{k+2}^T) w \right)\|_{L^2(U_{k+2})} \\ &\stackrel{(4.11)}{\leq} C'_4 \|A^{1/2} \nabla w\|_{L^2(U_{k+3})}. \end{aligned}$$

For the second term, II , the estimate

$$\begin{aligned}
 II &= \|A^{1/2} \nabla \mathcal{I}_H(\eta_{k+2}^T \mathbf{w})\|_{L^2(\Omega)} \\
 &= \|A^{1/2} \nabla \mathcal{I}_H(\eta_{k+2}^T \mathbf{w})\|_{L^2(U_{k+3} \setminus U_k)} \\
 &\stackrel{(4.13)}{\leq} C_4 \|A^{1/2} \nabla \mathbf{w}\|_{L^2(U_{k+3} \setminus U_{k-1})} \\
 &\leq C_4 \|A^{1/2} \nabla \mathbf{w}\|_{L^2(U_{k+3})}
 \end{aligned}$$

holds true. Now, we intend to take advantage of Lemma 4.2.6. By definition, q_k^T can be understood as the best approximation for q^T , in terms of the energy norm on the finescale space $V^f(U_k)$. Therefore, we can pick an arbitrary element of $V^f(U_k)$ to achieve an estimate for III . Set

$$\bar{q}_k^T := (1 - \eta_{k-1}^T) q^T - \mathcal{I}_H((1 - \eta_{k-1}^T) q^T).$$

We easily observe $\bar{q}_k^T \in V^f(U_k)$ and hence, we get

$$\begin{aligned}
 III &= \|A^{1/2} \nabla (q^T - q_k^T)\|_{L^2(\Omega)} \\
 &\leq \|A^{1/2} \nabla (q^T - \bar{q}_k^T)\|_{L^2(\Omega)} \\
 &= \|A^{1/2} \nabla (\eta_{k-1}^T q^T - \mathcal{I}_H(\eta_{k-1}^T q^T) + \mathcal{I}_H(q^T))\|_{L^2(\Omega)} \\
 &\leq \|A^{1/2} \nabla (\eta_{k-1}^T q^T)\|_{L^2(\Omega \setminus U_{k-2})} + \|A^{1/2} \nabla \mathcal{I}_H(\eta_{k-1}^T q^T)\|_{L^2(U_k \setminus U_{k-3})} \\
 &\stackrel{(4.11)}{\leq} C'_4 \|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_{k-3})} + \|A^{1/2} \nabla \mathcal{I}_H(\eta_{k-1}^T q^T)\|_{L^2(U_k \setminus U_{k-3})} \\
 &\stackrel{(4.13)}{\leq} C'_4 \|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_{k-3})} + C_4 \|A^{1/2} \nabla q^T\|_{L^2(U_{k+1} \setminus U_{k-3})} \\
 &\leq C'_Q \|A^{1/2} \nabla q^T\|_{L^2(\Omega \setminus U_{k-3})} \\
 &\stackrel{(4.15)}{\leq} C'_Q \theta^{k-3} \|A^{1/2} \nabla q^T\|_{L^2(\Omega)},
 \end{aligned}$$

where $C'_Q = C_4 + C'_4$ and θ like in Lemma 4.2.6. Combining I , II and III , results finally in

$$\begin{aligned}
 \|A^{1/2} \nabla \mathbf{w}\|_{L^2(\Omega)}^2 &\leq \sum_{T \in \mathcal{T}_H} (I + II) \cdot III \\
 &\leq 2C'_Q \theta^{k-3} \sum_{T \in \mathcal{T}_H} \|A^{1/2} \nabla \mathbf{w}\|_{L^2(U_{k+3})} \|A^{1/2} \nabla q^T\|_{L^2(\Omega)} \\
 &\stackrel{\text{c.s.}}{\leq} 2C'_Q \theta^{k-3} \left(\sum_{T \in \mathcal{T}_H} \|A^{1/2} \nabla \mathbf{w}\|_{L^2(U_{k+3})}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_H} \|A^{1/2} \nabla q^T\|_{L^2(\Omega)}^2 \right)^{1/2} \\
 &\leq 2C'_Q (k+3)^{d/2} \theta^{k-3} \|A^{1/2} \nabla \mathbf{w}\|_{L^2(\Omega)} \|A^{1/2} \nabla q\|_{L^2(\Omega)} \\
 &\leq C_Q k^{d/2} \theta^k \|A^{1/2} \nabla \mathbf{w}\|_{L^2(\Omega)} \|A^{1/2} \nabla q\|_{L^2(\Omega)},
 \end{aligned}$$

where $C_Q = 2C'_Q C_d \theta^{-3}$. The factor $(k+3)^{d/2}$ results from the transformation that

needs to be made to get rid of U_{k+3} and the constant C_d is related to separating $k^{d/2}$ from $(k+3)^d/2$. Dividing with $\|A^{1/2}\nabla w\|_{L^2(\Omega)}$ completes the proof. \square

We are now equipped to prove Theorem 4.2.1.

PROOF OF THEOREM 4.2.1: We start with using Galerkin orthogonality such that, for every $v_{H,k}^{\text{LOD}} \in V_{H,k}^{\text{LOD}}$,

$$\|u - u_{H,k}^{\text{LOD}}\| \leq \|u - v_{H,k}^{\text{LOD}}\|$$

holds true. We recall that $u_H^{\text{LOD}} = u_H - \mathcal{Q}u_H$ and choose $v_{H,k}^{\text{LOD}} = u_H - \mathcal{Q}_k u_H$ for $u_H \in V_H$. Due to the decomposition $V = V_H^{\text{LOD}} + V^f$, we can write

$$u = u_H^{\text{LOD}} + u^f$$

and we observe

$$\begin{aligned} \|u - v_{H,k}^{\text{LOD}}\| &\leq \|u_H^{\text{LOD}} + u^f - v_{H,k}^{\text{LOD}}\| \\ &\leq \|u_H - \mathcal{Q}u_H - u_H + \mathcal{Q}_k u_H\| + \|u^f\| \\ &\leq \|\mathcal{Q}u_H - \mathcal{Q}_k u_H\| + \|u^f\| \\ &\stackrel{(4.18)}{\leq} C_Q k^{d/2} \theta^k \|A^{1/2}\nabla \mathcal{Q}u_H\|_{L^2(\Omega)} + \|u^f\|. \end{aligned}$$

Further, we have

$$\begin{aligned} \|\mathcal{Q}u_H\|^2 &= a(\mathcal{Q}u_H, \mathcal{Q}u_H) \\ &= a(u_H, \mathcal{Q}u_H) \\ &\stackrel{\text{C.S.}}{\leq} \|u_H\| \|\mathcal{Q}u_H\|. \end{aligned} \tag{4.19}$$

Therefore, we obtain because of $\mathcal{I}_H(u_H) = u_H$

$$\begin{aligned} \|\mathcal{Q}u_H\| &\leq \|u_H\| \\ &= \|\mathcal{I}_H(u_H^{\text{LOD}} + u^f)\| \\ &= \|\mathcal{I}_H(u_H^{\text{LOD}})\| \\ &\leq \beta^{1/2} \|\nabla \mathcal{I}_H(u_H^{\text{LOD}})\|_{L^2(\Omega)} \\ &\stackrel{(3.16)}{\leq} \beta^{1/2} C_{\mathcal{I}_H} \|\nabla u_H^{\text{LOD}}\|_{L^2(\Omega)} \\ &\leq \beta^{1/2} \alpha^{-1/2} C_{\mathcal{I}_H} \|u_H^{\text{LOD}}\|. \end{aligned}$$

Combining this with

$$\begin{aligned}
 \|\mathbf{u}_H^{\text{LOD}}\|^2 &= a(\mathbf{u}_H^{\text{LOD}}, \mathbf{u}_H^{\text{LOD}}) \\
 &= F(\mathbf{u}_H^{\text{LOD}}) \\
 &\stackrel{\text{C.S.}}{\leq} \|\mathbf{u}_H^{\text{LOD}}\|_{L^2(\Omega)} \|f\|_{L^2(\Omega)} \\
 &\leq C_P \|\nabla \mathbf{u}_H^{\text{LOD}}\|_{L^2(\Omega)} \|f\|_{L^2(\Omega)} \\
 &\leq C_P \alpha^{-1/2} \|\mathbf{u}_H^{\text{LOD}}\| \|f\|_{L^2(\Omega)}
 \end{aligned}$$

yields with $C'_1 := C_P \alpha^{-1} \beta^{1/2} C_{I_H}$,

$$\|\mathbf{Q}u_H\| \leq C'_1 k^{d/2} \theta^k \|f\|_{L^2(\Omega)}. \quad (4.20)$$

C_P arises from the Poincaré-inequality. The last step is to find an error bound for $\|\mathbf{u}^f\|$, which we already did in Lemma 4.1.1.

$$\begin{aligned}
 \|\mathbf{u}^f\|^2 &= a(\mathbf{u}^f, \mathbf{u}^f) \\
 &= (A^{-1/2} f, A^{1/2} \mathbf{u}^f) \\
 &\stackrel{\text{C.S.}}{\leq} \|A^{-1/2} f\|_{L^2(\Omega)} \|A^{1/2} \mathbf{u}^f\|_{L^2(\Omega)} \\
 &\stackrel{(4.10)}{=} \alpha^{-1/2} C_3 H \|f\|_{L^2(\Omega)} \|A^{1/2} \nabla \mathbf{u}^f\|_{L^2(\Omega)} \\
 &\leq C_2 H \|f\|_{L^2(\Omega)} \|\mathbf{u}^f\|,
 \end{aligned}$$

with $C_2 := \alpha^{1/2} C_3$. Adding all parts together and setting $C_1 := C'_1 C_Q$ leads to the expected result,

$$\|\mathbf{u} - \mathbf{u}_{H,k}^{\text{LOD}}\| \leq \left(C_1 k^{d/2} \theta^k + H C_2 \right) \|f\|_{L^2(\Omega)}.$$

□

5 Variational Crimes

Variational crimes is a popular topic for Galerkin methods, respectively, the FEM. Gilbert Strang already mentioned them in one of the first works about FEMs (see [32]). Strang perceived problems for the accuracy, in case specific approximation properties get touched. One variational crime occurs due to an approximation of the domain $\mathcal{T}_H \not\subset \Omega$. This could be necessary if Ω contains a curved boundary or other shapes that are not exactly approachable. Surely, $\mathcal{T}_H \not\subset \Omega$ implies $V_H \not\subset V$. A second variational crime might occur, when the entries of the stiffness matrix \mathcal{S} or the load vector \mathcal{L} get computed. Clearly, every entry contains integrals. Depending on its functions, those integrals need to be approximated with quadrature rules that might not be exact. In fact, we compute \tilde{a} and \tilde{F} instead of the exact versions a and F . Most importantly for this thesis, we can also derive an error for the entries of the stiffness matrix, if the coefficient is subjected to perturbations or approximations. For this thesis, we only consider the variational crime of an altered a . Regardless of the cause, the ideal Galerkin method,

$$a(u, v) = F(v), \quad \forall v \in V,$$

turns into an approximated form,

$$\tilde{a}(\tilde{u}, v) = F(v), \quad \forall v \in V_H.$$

Either variational crime violates Céa's lemma, as the assumptions are no longer attained. However, Céa's lemma is representative for the error bounds of the FEM. To bypass this issue, we present Strang's lemma, applied to our situation.

5.0.1 Lemma (Strang's lemma, see [7]) Let $u \in V$ be a solution of

$$a(u, v) = F(v), \quad \forall v \in V$$

and $\tilde{u} \in V_H$ a solution of

$$\tilde{a}(\tilde{u}, v) = F(v), \quad \forall v \in V_H.$$

Then for $C_s > 0$, the following error estimate

$$\|u - \tilde{u}\| \leq C_s \left(\inf_{v_H \in \mathbb{V}_H} \left\{ \|u - v_H\| + \sup_{w_H \in \mathbb{V}_H} \frac{|a(v_H, w_H) - \tilde{a}(v_H, w_H)|}{\|w_H\|} \right\} \right) \quad (5.1)$$

holds true.

In many applications in engineering and industries, perturbations of the coefficient can be observed. A composite material that is periodic in theory, often deviates in practice; for example, if the fibers or other material components are unpredictable. However, we can assume that this perturbation is based on an old coefficient \tilde{A} . We call it the reference coefficient since every perturbation in the underlying problem can be ascribed to it. We define the slightly different continuous and bounded bilinear form

$$\tilde{a}(v, w) := \int_{\Omega} \left(\tilde{A}(x) \nabla v(x) \right) \cdot \nabla w(x) \, dx, \quad \forall v \in V, w \in V.$$

The perturbed version of the reference \tilde{a} still reads

$$a(v, w) := \int_{\Omega} (A(x) \nabla v(x)) \cdot \nabla w(x) \, dx, \quad \forall v \in V, w \in V.$$

We choose this reversed notation for convenience with regard to the observations in Chapter 6. Moreover, we keep the linear functional

$$F(w) := \int_{\Omega} f(x) w(x) \, dx, \quad \forall w \in V.$$

Now, the weak formulation for the perturbed diffusion problem reads as follows.

5.0.2 Definition (Exact perturbed problem) For $V = H_0^1(\Omega)$, a , \tilde{a} and F defined above, the weak formulation of a perturbed diffusion problem is to find $u \in V$ such that, for all $v \in V$, it holds that

$$a(u, v) = F(v). \quad (5.2)$$

The reference problem is to find the solution $\tilde{u} \in V$ such that, for all $v \in V$,

$$\tilde{a}(\tilde{u}, v) = F(v), \quad (5.3)$$

where \tilde{a} denotes the non-perturbed version of a , defined above.

We are now able to pursue Strang's lemma, formulated in the following lemma.

5.0.3 Lemma For the solutions u and \tilde{u} of (5.2) and (5.3) in Definition 5.0.2, the following

error bound

$$\|u - \tilde{u}\| \leq C_s \|u - I_H u\| + C \|A - \tilde{A}\|_{L^\infty(\Omega)} \|f\|_{L^2(\Omega)} \quad (5.4)$$

holds.

PROOF: We start with the result of Strang's lemma. Remember that the interpolation $I_H : V \rightarrow V_H$, introduced in Section 3.6, maps V to the FE space V_H .

$$\begin{aligned} \frac{1}{C_s} \|u - \tilde{u}\| &\stackrel{(5.1)}{\leq} \inf_{v_H \in V_H} \left(\|u - v_H\| + \sup_{w_H \in V_H} \frac{|a(v_H, w_H) - \tilde{a}(v_H, w_H)|}{\|w_H\|} \right) \\ &\stackrel{I_H u \in V_H}{\leq} \|u - I_H u\| + \sup_{w_H \in V_H} \frac{|a(I_H u, w_H) - \tilde{a}(I_H u, w_H)|}{\|w_H\|} \\ &\leq \|u - I_H u\| + \sup_{w_H \in V_H} \frac{|((A - \tilde{A})\nabla I_H u, w_H)|}{\|w_H\|} \\ &\stackrel{\text{C.S.}}{\leq} \|u - I_H u\| + \sup_{w_H \in V_H} \frac{\|(A - \tilde{A})\nabla I_H u\|_{L^2(\Omega)} \|w_H\|_{L^2(\Omega)}}{\|w_H\|} \\ &\leq \|u - I_H u\| + \alpha^{-1/2} \|(A - \tilde{A})\nabla I_H u\|_{L^2(\Omega)} \\ &\leq \|u - I_H u\| + \alpha^{-1/2} \|A - \tilde{A}\|_{L^\infty(\Omega)} \|\nabla I_H u\|_{L^2(\Omega)} \\ &\stackrel{(3.16)}{\leq} \|u - I_H u\| + \alpha^{-1/2} \beta^{1/2} C_{I_H} \|A - \tilde{A}\|_{L^\infty(\Omega)} \|u\|. \end{aligned}$$

For $\|u\|$, we conclude with the help of Poincaré's inequality (see [7])

$$\begin{aligned} \|u\|^2 &= a(u, u) \\ &= (f, u) \\ &\leq \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)} \\ &\leq \|f\|_{L^2(\Omega)} \alpha^{-1/2} C_p \|u\|_{L^2(\Omega)}, \end{aligned}$$

which finally attains the assertion. \square

This lemma gives an idea about how the perturbations affect the error bound. Depending on the amount of the perturbation, the maximum norm could potentially have very high values. In the applications, we see instances of dramatically increasing $L^\infty(\Omega)$ norm. In practice, we might want to solve several problems of the same type, based on the same reference coefficient. This is why we start to think about possible ways to decrease computational costs. In terms of the standard FEM, there are two opportunities to deal with perturbations of a reference problem. Either we compute the stiffness matrix with respect to the reference coefficient \tilde{A} , or to the perturbed coefficient A . According to Lemma 5.0.3, approximating a perturbed problem with the reference solution might cause a very high error, whereas a recomputation is much more expensive. We show that the

PG-LOD, presented in Chapter 3, enables to derive a novel method that is able to reduce computational costs significantly and still keeps the accuracy well enough. This method can be applied to normal perturbation problems, but it is even stronger for multiscale problems or in case the FEM is not applicable at all. The method has been proposed in [15] in order to solve time-dependent diffusion problems. Basically, Hellman and Målqvist considered every single time step in the diffusion coefficient as a perturbation of the 'lagging' coefficient. We will show that this method can be generalized to deal with the variational crime of perturbations. In the next section, we explain the novel method on the basis of various methods that we can deduce from the standard PG-LOD approach.

5.1 The novel Method

Obviously, both solutions in (5.2) and (5.3) can be computed with the LOD-technique. For this purpose, we may easily follow the same strategy, that is to use perturbed as well as reference correctors for each coarse element and derive the reference and perturbed versions of the space V_H^{LOD} . We compute the reference corrector functions, which are, analogously to the perturbed correctors, defined by

$$\begin{aligned}\tilde{\mathcal{Q}}_k &= \sum_{T \in \mathcal{T}_H} \tilde{\mathcal{Q}}_k^T, \\ \tilde{\mathcal{R}}_k &= \sum_{T \in \mathcal{T}_H} \tilde{\mathcal{R}}_k^T,\end{aligned}$$

with $\tilde{\mathcal{Q}}_k^T v, \tilde{\mathcal{R}}_k^T f \in V^f(U_k(T))$ such that, for all $v^f \in V^f(U_k(T))$,

$$\begin{aligned}\tilde{a}(\tilde{\mathcal{Q}}_k^T v, v^f) &= \tilde{a}_T(v, v^f), \\ \tilde{a}(\tilde{\mathcal{R}}_k^T f, v^f) &= F_T(v^f).\end{aligned}$$

The resulting methods are formulated in the following. To keep clarity, we omit the notification of H for further observations and definitions. Hence, we use V^{LOD} , based on A , respectively, \tilde{V}^{LOD} , based on \tilde{A} .

As we are interested in the PG-LOD with right hand side correction approximation of (5.2) and (5.3), we formulate the PG-LOD that occurs from (3.12) for the perturbed problem.

5.1.1 Definition (Ideal PG-LOD approximation of the perturbed problem) The ideal PG-LOD approximation with right hand side correction of u in (5.2) is to find $u^{\text{PG}} \in V^{\text{LOD}}$ such that, for all $v \in V_H$, it holds that

$$a(u^{\text{PG}}, v) = F(v) - a(\mathcal{R}f, v). \quad (5.5)$$

Due to (3.13), the full solution is $u = u^{\text{PG}} + \mathcal{R}f$. Besides, the ideal PG-LOD approximation of the reference solution \tilde{u} in (5.3) is to find $\tilde{u}^{\text{PG}} \in \tilde{V}^{\text{LOD}}$ such that, for all $v \in V_H$, it holds that

$$\tilde{a}(\tilde{u}^{\text{PG}}, v) = \tilde{F}(v) - \tilde{a}(\tilde{\mathcal{R}}f, v), \quad (5.6)$$

where the full solution reads $\tilde{u} = \tilde{u}^{\text{PG}} + \tilde{\mathcal{R}}f$.

However, to formulate a feasible and applicable method, we also need the localized versions of the PG-LOD for the reference and the perturbed version of $V_{H,k}^{\text{LOD}}$. This results in the localized PG-LOD approximations $u_k^{\text{PG}} \in V_k^{\text{LOD}}$, based on A , and $\tilde{u}_k^{\text{PG}} \in \tilde{V}_k^{\text{LOD}}$, based on \tilde{A} . The correctors Q_k induce the finescale space V_k^{LOD} and thus, we define the method as follows.

5.1.2 Definition (Localized PG-LOD approximation of the perturbed problem) The localized PG-LOD approximation with right hand side correction of u in (5.2) is to find $u_k^{\text{PG}} \in V_k^{\text{LOD}}$ such that, for all $v \in V_H$, it holds that

$$a(u_k^{\text{PG}}, v) = F(v) - a(\mathcal{R}_k f, v), \quad (5.7)$$

where the full solution is denoted by $u_k = u_k^{\text{PG}} + \mathcal{R}_k f$. Furthermore, the localized PG-LOD approximation of the reference solution \tilde{u} in (5.3) is to find $\tilde{u}_k^{\text{PG}} \in \tilde{V}_k^{\text{LOD}}$ such that, for all $v \in V_H$, it holds that

$$\tilde{a}(\tilde{u}_k^{\text{PG}}, v) = \tilde{F}(v) - \tilde{a}(\tilde{\mathcal{R}}_k f, v) \quad (5.8)$$

and the full solution is given by $\tilde{u}_k = \tilde{u}_k^{\text{PG}} + \tilde{\mathcal{R}}_k f$.

Since A is just a perturbation of the reference problem, several correctors that are based on the perturbed coefficient might result equally or differ just circumstantially to the reference correctors. Thus, depending on the perturbation, a solution of (5.2) might be computed with a huge loss of computational complexity, once we are able to fall back on the already computed reference correctors. As we are interested in the solution of many perturbations, this creates a huge benefit for our purposes. Having this in mind, we may think about two different possibilities to achieve an approximation of (5.2).

1. For sure, the most accurate version is to compute V_k^{LOD} entirely and furthermore, to use the new bilinear form a that is based on A , to compute the stiffness matrix. This approach is already formulated in Definition 5.1.2. Obviously, it does not use the link to the reference problem and requires as high computational complexity as the reference problem itself.
2. The second possible strategy replaces the perturbed bilinear form a with the reference bilinear form \tilde{a} and recomputes no corrector as it uses the reference LOD

space \tilde{V}_k^{LOD} . This approach equals the localized PG-LOD of the reference problem (5.8).

It is obvious that the approach in the first strategy enables the best accuracy, whereas the second minimizes the costs completely. However, it might have a huge error. It turns out that the best way of saving computational complexity is a mix between both strategies, restricted to each coarse element $T \in \mathcal{T}_H$. We intend to develop a novel method that decides for every $T \in \mathcal{T}_H$ which strategy shall be used. This method requires to develop a reasonable error indicator. The error analysis in Chapter 6 reveals an error estimate, which can be understood as the PG-LOD version of Lemma 5.0.3 (Strang's lemma). It is stated in Theorem 6.2.9 and reads

$$\|u - \tilde{u}_k\| \leq ck^{d/2}(\theta^k + \max(e_u, e_f))\|f\|_{L^2(\Omega)}.$$

The error indicators e_u and e_f are a priori computable and defined in Chapter 6. Furthermore, they consist of the maximum of $e_{u,T}$ respectively $e_{f,T}$ over all $T \in \mathcal{T}_H$. If the maximum is comparatively big, the error bound indicates that the reference method might result in an inaccurate approximation. As we have already discussed, the best way of saving computational complexity but not sacrificing too much of the accuracy is a blend between perturbed and reference localized PG-LOD. We therefore compute $e_{u,T}$ and $e_{u,f}$ for every single $T \in \mathcal{T}_H$ and decide separately whether we want to recompute the corrector. For this purpose, we pick a tolerance TOL for the error indicator so that $\max(e_u, e_f) \leq \text{TOL}$. This criteria helps to decide on recomputing and a recomputed element $T \in \mathcal{T}_H$ fulfills $e_{u,T} = e_{f,T} = 0$. Thus, the resulting novel approximation is dependent on the desired percentage of recomputing. In the following, we intend to formulate the discussed procedure. For clarity reasons, we omit the e_f for the correctors $\mathcal{R}^T f$ and formulate the novel method only for the standard correctors Q^T . The involvement of e_f works analogously. As we have already mentioned, we can pick the tolerance TOL arbitrarily. If we aim to achieve a particular percentage of recomputing, we sort the error $e_{u,T}$ for each T and define $\text{TOL} = e_{u,T}$ such that the correct percentage results. After updating the chosen correctors, we get $e_u \leq \text{TOL}$ and a new LOD space with a blend of reference and perturbed correctors. We will call this space $V_{k,p}^{\text{vc}}$, which is obviously dependent on the recompute fraction p . Related to this, we also gain a mixed bilinear form a_{vc} , defined by

$$a_{\text{vc}}(v, w) = \sum_{T \in \mathcal{T}_H \setminus \tilde{\mathcal{T}}_H} a_T(v, w) + \sum_{\tilde{T} \in \tilde{\mathcal{T}}_H} \tilde{a}_{\tilde{T}}(v, w), \quad \forall v, w \in V,$$

where $\tilde{\mathcal{T}}_H$ denotes the set of elements with reference correctors. With these definitions in mind, we formulate the novel method as follows.

5.1.3 Definition (Novel PG-LOD method for perturbations) Fix the patch size $k \in \mathbb{N}$ and a certain recompute fraction $p \in [0, 1]$. Let $\text{TOL}(p)$ be a tolerance such that

$$e_u \leq \text{TOL}(p) \quad \text{and} \quad \frac{\#\mathcal{Q}}{(\#\mathcal{Q} + \#\tilde{\mathcal{Q}})} \approx p,$$

where $\#\mathcal{Q}$ denotes the number of updated correctors and $\#\tilde{\mathcal{Q}}$ the number of old reference correctors. Let $V_{k,p}^{\text{vc}}$ be the resulting LOD space with mixed correctors \mathcal{Q}^T and $\tilde{\mathcal{Q}}^T$. The novel PG-LOD approximation of u in (5.3) is to find $u_{k,p}^{\text{vc}} \in V_{H,k}^{\text{vc}}$ such that, for all $v \in V_H$, it holds that

$$a_{\text{vc}}(u_{k,p}^{\text{vc}}, v) = F(v). \quad (5.9)$$

To simplify the procedure, we present the following pseudo code. Further details regarding

```

Pick  $k$  and  $p$  to fulfill  $e_u < \text{TOL}$ .
Copy precomputed  $\tilde{\mathcal{Q}}_k^T \phi_j$  for all  $T$  and  $j$ 
for all  $T$  do
  Compute  $e_{u,T}$ .
  if  $e_{u,T} \geq \text{TOL}$  then
    | Recompute  $\mathcal{Q}_k^T \phi_j$  for all  $j$ 
  end
end
Apply the PG-LOD to compute  $u_{k,p}^{\text{vc}}$ .

```

the implementation of the PG-LOD are revealed in Chapter 7. We now present the error analysis that justifies this method.

6 Error analysis for the PG-LOD on perturbed problems

The following chapter is devoted to the error analysis for the novel method. It aims to characterize the error indicator that is required for the approach. Moreover, we prove the well-posedness for each approach presented above and gain a PG-LOD error estimate. The approach is based on the observations for time-dependent problems in [15].

6.1 Error indicators

The error indicators that need to be evolved have to be dependent on A and \tilde{A} and, most importantly, they have to be computable with the least possible information and computational complexity. We utilize that the error of the PG-LOD mainly contains factors of the form $\|\mathcal{Q}_k^T - \tilde{\mathcal{Q}}_k^T\|$ and $\|\mathcal{R}_k^T f - \tilde{\mathcal{R}}_k^T f\|$. Note that the energy norm is induced by the perturbed bilinear form, based on A . Deriving an estimate for these factors yield the characterization of the error indicators.

6.1.1 Definition (Error indicators) We define the error indicators, for each $T \in \mathcal{T}_H$, by

$$\begin{aligned} e_{u,T} &= \max_{v|_T \mid v \in V_H, \|v\|_T=1} \|(\tilde{A} - A)A^{-1/2}(\chi_T \nabla v - \nabla \tilde{\mathcal{Q}}_k^T v)\|_{L^2(U_k(T))}, \\ e_{f,T} &= \frac{\|(\tilde{A} - A)^{-1/2} \nabla \tilde{\mathcal{R}}_k^T f\|_{L^2(U_k(T))}}{\|f\|_{L^2(T)}}, \end{aligned} \tag{6.1}$$

where χ_T denotes the indicator function for an element $T \in \mathcal{T}_H$. We have $e_{f,T} = 0$, in case $\|f\|_{L^2(T)} = 0$. Furthermore, we define

$$\begin{aligned} e_u &:= \max_T e_{u,T}, \\ e_f &:= \max_T e_{f,T}. \end{aligned}$$

The following lemma justifies the use of these error indicators.

6.1.2 Lemma For all $v \in V_H$, the error bounds

$$\| \| \mathcal{Q}_k^T v - \tilde{\mathcal{Q}}_k^T v \| \| \leq e_{u,T} \| \| v \| \|_T, \quad (6.2)$$

$$\| \| \mathcal{R}_k^T f - \tilde{\mathcal{R}}_k^T f \| \| \leq e_{f,T} \| \| v \| \|_T \quad (6.3)$$

hold true.

PROOF: For any $v \in V_H$, we define $z := \mathcal{Q}_k^T v - \tilde{\mathcal{Q}}_k^T v \in V^f(U_k(T))$ and we observe

$$\begin{aligned} \| \| z \| \|_{U_k(T)}^2 &= (A \nabla (\mathcal{Q}_k^T v - \tilde{\mathcal{Q}}_k^T v), \nabla z)_{U_k(T)} \\ &\stackrel{z \in V^f}{=} (A \nabla v, \nabla z)_T - (A \nabla \tilde{\mathcal{Q}}_k^T v, \nabla z)_{U_k(T)} + (\tilde{A} \nabla \tilde{\mathcal{Q}}_k^T v, \nabla z)_{U_k(T)} - (\tilde{A} \nabla v, \nabla z)_T \\ &= ((\tilde{A} - A) \nabla \tilde{\mathcal{Q}}_k^T v, \nabla z)_{U_k(T)} - ((\tilde{A} - A) \nabla v, \nabla z)_T \\ &= ((\tilde{A} - A) A^{-1/2} (\nabla \tilde{\mathcal{Q}}_k^T v - \chi_T \nabla v), A^{1/2} \nabla z)_{U_k(T)} \\ &\leq \| \| (\tilde{A} - A) A^{-1/2} (\chi_T \nabla v - \nabla \tilde{\mathcal{Q}}_k^T v) \| \|_{L^2(U_k(T))} \cdot \| \| z \| \|_{U_k(T)}. \end{aligned}$$

Dividing by $\| \| z \| \|_{U_k(T)}$ and taking the maximum on $w \in V_H$, where $\| \| w \| \| = 1$, results in the first assertion. The second estimate follows similarly with $z := \mathcal{R}_k^T f - \tilde{\mathcal{R}}_k^T f$ and with

$$\begin{aligned} \| \| z \| \|_{U_k(T)}^2 &= (A \nabla (\mathcal{R}_k^T f - \tilde{\mathcal{R}}_k^T f), \nabla z)_{U_k(T)} \\ &= (f, z)_T - (A \nabla \tilde{\mathcal{R}}_k^T f, \nabla z)_{U_k(T)} \\ &= ((\tilde{A} - A) \nabla \tilde{\mathcal{R}}_k^T f, \nabla z)_{U_k(T)} \\ &\leq \| \| (\tilde{A} - A) A^{-1/2} \tilde{\mathcal{R}}_k^T f \| \|_{L^2(U_k(T))} \| \| z \| \|_{U_k(T)}. \end{aligned}$$

□

Certainly, we have to mention how to compute the error indicators $e_{u,T}$ and $e_{f,T}$. We refer to Chapter 7, where the implementational details are discussed further. For now, it is only important that they are actually computable a priori with a reasonable complexity.

6.2 Error analysis

This section deals with the problems that have been presented in Chapter 5. The aim is to investigate the well-posedness in terms of the energy norm, which is induced by A . Furthermore, we deduce an error estimate that evaluates the use of the PG-LOD, which leads to the application of the error indicators e_u and e_f . To elaborate the novel method for the perturbed problems, it is useful to obtain an estimate for

$$\| \| u - \tilde{u}_k \| \|,$$

which we have utilized in Section 5.1. Note that we consider the energy norm in terms of A . The error bounds obtained in this section are connected to the stability constants of the inf-sup conditions (see Definition 3.4.3) for each approach. Recall that this inf-sup conditions are responsible for the well-posedness of a Petrov Galerkin formulation, since the trial and test spaces are not equal anymore. Thus, the inf-sup conditions are representative for the lost coercivity of the bilinear form and ensure the uniqueness of the approximations. To summarize PG-LOD methods in Chapter 5, we mention each influent approach and recall each name. Each method aims to gain an approximation for the solution u of (5.2) in the exact perturbed formulation in Definition 5.0.2. We defined the different methods in Definition 5.1.1 and 5.1.2.

Method	Definition
Ideal PG-LOD	$u = u^{\text{PG}} + \mathcal{R}f$
Localized PG-LOD	$u_k = u_k^{\text{PG}} + \mathcal{R}_k f$
Reference localized PG-LOD	$\tilde{u}_k = \tilde{u}_k^{\text{PG}} + \tilde{\mathcal{R}}_k f$

For the following calculations, we need to recall the interpolation estimate (3.16),

$$H_T^{-1} \|v - \mathcal{I}_H v\|_{L^2(T)} + \|\nabla \mathcal{I}_H v\|_{L^2(T)} \leq C_{\mathcal{I}_H} \|\nabla v\|_{L^2(U(T))}.$$

Every overlap effect is captured by $C_{\mathcal{I}_H}$ and furthermore, \mathcal{I}_H is a projection. We also need the uniformly boundedness constants α and β of our coefficient A , defined in (2.1) respectively, (2.2). In the following, we use the constants C and C' several times. They denote constants that we do not specify further, since they are not important for the observations. However, \bar{C} , \check{C} and \tilde{C} denote each constant for the error estimates and will be specified.

6.2.1 The ideal PG-LOD approximation

The ideal PG-LOD solution u in Definition 5.1.1 is the first approach that needs to be investigated. As we have already mentioned, thanks to the right hand side corrector, it is just a reformulation and thus, there is no error. However, we need to derive the well-posedness for the problem to ensure uniqueness of the approximation.

6.2.1.1 Stability

The stability of the ideal PG-LOD method is explained by the following lemma.

6.2.1 Lemma (Stability of the ideal PG-LOD method) There exists a constant $\gamma > 0$ such that the inf-sup condition for a with respect to the trial space V^{LOD} and the test

space V_H

$$\inf_{w^{\text{LOD}} \in V^{\text{LOD}}} \sup_{v \in V_H} \frac{|a(w^{\text{LOD}}, v)|}{\|w^{\text{LOD}}\| \|v\|} \geq \gamma \quad (6.4)$$

is satisfied.

PROOF: We make the observation

$$\begin{aligned} \inf_{w^{\text{LOD}} \in V^{\text{LOD}}} \sup_{v \in V_H} \frac{|a(w^{\text{LOD}}, v)|}{\|w^{\text{LOD}}\| \|v\|} &= \inf_{w \in V_H} \sup_{v \in V_H} \frac{|(A\nabla(w - Qw), \nabla v)|}{\|w - Qw\| \|v\|} \\ &\stackrel{Qv \in V^f}{=} \inf_{w \in V_H} \sup_{v \in V_H} \frac{|(A\nabla(w - Qw), \nabla(v - Qv))|}{\|w - Qw\| \|v\|} \\ &\stackrel{v:=w}{\geq} \inf_{w \in V_H} \frac{\|(w - Qw)\|^2}{\|w - Qw\| \|w\|} \\ &\stackrel{w \in V_H}{=} \inf_{w \in V_H} \frac{\|(w - Qw)\|^2}{\|w - Qw\| \|I_H(w - Qw)\|} \\ &\stackrel{(3.16)}{\geq} C_{I_H}^{-1} \alpha^{1/2} \beta^{-1/2} =: \gamma. \end{aligned}$$

Since $C_{I_H}, \alpha, \beta > 0$, we conclude $\gamma > 0$. □

Thus, the uniqueness of the ideal PG-LOD approximation is justified.

6.2.2 The localized PG-LOD approximation

We are now prepared to investigate the inf-sup condition for the localized PG-LOD method in Definition 5.1.2 and conclude a first error bound for $\|u - u_k\|$. With Lemma 4.2.7, we have already proven the key ingredient for the next inf-sup condition, the estimate for the correctors (4.18). We first formulate and prove the inf-sup condition.

6.2.2.1 Stability

First of all, the stability of the localized PG-LOD is of interest.

6.2.2 Lemma (Stability for the localized PG-LOD method) There exists a constant $\bar{\gamma} > 0$ such that the inf-sup condition for a with respect to the trial space V_k^{LOD} and the test space V_H

$$\inf_{w_k^{\text{LOD}} \in V_k^{\text{LOD}}} \sup_{v \in V_H} \frac{|a(w_k^{\text{LOD}}, v)|}{\|w_k^{\text{LOD}}\| \|v\|} \geq \bar{\gamma}_k, \quad (6.5)$$

for a sufficiently large k , is satisfied.

PROOF: As we have already mentioned, we use (4.18), stated and proved in Lemma

4.2.7, for every v ,

$$\|Qv - Q_k v\| \leq C_Q k^{d/2} \theta^k \|Qv\|,$$

for constants C_Q and $0 < \theta < 1$. We are able to make a further observation if we suppose that $v \in V_H$, since we apply (4.19), the projection property of \mathcal{I}_H and $Qv \in V^f$.

$$\begin{aligned} \|Qv\| &\stackrel{(4.19)}{\leq} \|v\| \\ &= \|\mathcal{I}_H(v - Qv)\| \\ &\stackrel{(3.16)}{\leq} C_{\mathcal{I}_H} \beta^{1/2} \alpha^{-1/2} \|v - Qv\|. \end{aligned} \tag{6.6}$$

Hence, we conclude for $C := C_Q C_{\mathcal{I}_H} \beta^{1/2} \alpha^{-1/2}$

$$\|Qv - Q_k v\| \leq C k^{d/2} \theta^k \|v - Qv\|. \tag{6.7}$$

This estimate can be applied to the inf-sup condition by

$$\begin{aligned} \inf_{w_k^{\text{LOD}} \in V_k^{\text{LOD}}} \sup_{v \in V_H} \frac{|a(w_k^{\text{LOD}}, v)|}{\|w_k^{\text{LOD}}\| \|v\|} &= \inf_{w \in V_H} \sup_{v \in V_H} \frac{|a(w - Q_k w, v)|}{\|w - Q_k w\| \|v\|} \\ &\geq \inf_{w \in V_H} \sup_{v \in V_H} \frac{|a(w - Qw, v)| - |a(Qw - Q_k w, v)|}{(\|w - Qw\| + \|Qw - Q_k w\|) \|v\|} \\ &\stackrel{(6.7)}{\geq} \inf_{w \in V_H} \sup_{v \in V_H} \frac{|a(w - Qw, v)| - C k^{d/2} \theta^k \|w - Qw\| \|v\|}{(1 + C k^{d/2} \theta^k) \|w - Qw\| \|v\|} \\ &\stackrel{(6.4)}{\geq} \frac{\gamma - C k^{d/2} \theta^k}{1 + C k^{d/2} \theta^k} =: \gamma_k. \end{aligned}$$

According to the definition of γ_k , we recognize that for sufficiently large k , there exists a $\bar{\gamma}$ such that $0 < \bar{\gamma} \leq \gamma_k$. \square

6.2.2.2 Error

We now intend to derive an error estimate for $\|u - u_k\|$. First of all, we remark that the results presented in Chapter 4, especially Theorem 4.2.1 are no longer applicable. The way we used the Galerkin orthogonality in the proof is not satisfied, which is a consequence of the use of the Petrov-Galerkin formulation. Therefore, we can not just apply the error analysis in Chapter 4 at this point. We need to derive something similar. This observation emphasizes the use of the right hand side correction.

6.2.3 Lemma (Error for the localized PG-LOD) With the exact solution u of (5.2) and the localized PG-LOD approximation u_k , it holds that

$$\|u - u_k\| \leq \bar{C} k^{d/2} \theta^k \|f\|_{L^2(\Omega)}, \tag{6.8}$$

for a constant \bar{C} and $0 < \theta < 1$.

PROOF: As a workaround for the Galerkin orthogonality, in order to connect u and u_k , we first subtract the equations (5.5) from (5.7) and yield, for $v \in V_H$,

$$a(u_k^{\text{PG}}, v) = a(u^{\text{PG}}, v) + a((\mathcal{R} - \mathcal{R}_k)f, v).$$

Further, we choose an arbitrary $u_I \in V_k^{\text{LOD}}$ to conclude

$$a(u_k^{\text{PG}} - u_I, v) = a(u^{\text{PG}} - u_I, v) + a((\mathcal{R} - \mathcal{R}_k)f, v). \quad (6.9)$$

Thus, the inf-sup condition implies

$$\begin{aligned} \bar{\gamma} \|u_k^{\text{PG}} - u_I\| &\stackrel{(6.5)}{\leq} \sup_{v \in V_H} \frac{|a(u_k^{\text{PG}} - u_I, v)|}{\|v\|} \\ &\stackrel{(6.9)}{\leq} \sup_{v \in V_H} \frac{|a(u^{\text{PG}} - u_I, v)| + |a((\mathcal{R} - \mathcal{R}_k)f, v)|}{\|v\|} \\ &\stackrel{\text{C.S.}}{\leq} \|u^{\text{PG}} - u_I\| + \|(\mathcal{R} - \mathcal{R}_k)f\|, \end{aligned} \quad (6.10)$$

which is a best approximation result to replace the Galerkin orthogonality. We now want to choose u_I reasonably. For this purpose, we first remark that $\mathcal{I}_H u^{\text{PG}} = \mathcal{I}_H u = u_H$ and thus,

$$\begin{aligned} u^{\text{PG}} &= u_H - \mathcal{Q}u_H \\ &= \mathcal{I}_H u^{\text{PG}} - \mathcal{Q}\mathcal{I}_H u^{\text{PG}}. \end{aligned}$$

Setting $u_I := \mathcal{I}_H u^{\text{PG}} - \mathcal{Q}_k \mathcal{I}_H u^{\text{PG}}$ results in

$$u^{\text{PG}} - u_I = -(\mathcal{Q} - \mathcal{Q}_k)\mathcal{I}_H u^{\text{PG}}. \quad (6.11)$$

Plugging this result into the best approximation result concludes

$$\begin{aligned} \|u^{\text{PG}} - u_k^{\text{PG}}\| &\leq \|u^{\text{PG}} - u_I\| + \|u_k^{\text{PG}} - u_I\| \\ &\stackrel{(6.10)}{\leq} (1 + \bar{\gamma}^{-1})\|u^{\text{PG}} - u_I\| + \gamma_k^{-1}\|(\mathcal{R} - \mathcal{R}_k)f\| \\ &\stackrel{(6.11)}{=} (1 + \bar{\gamma}^{-1})\|(\mathcal{Q} - \mathcal{Q}_k)\mathcal{I}_H u^{\text{PG}}\| + \gamma_k^{-1}\|(\mathcal{R} - \mathcal{R}_k)f\|. \end{aligned} \quad (6.12)$$

We already see that the estimate amounts to an estimate for $\|\mathcal{Q} - \mathcal{Q}_k\|$ and $\|\mathcal{R} - \mathcal{R}_k\|$. The former is already given by Lemma 4.2.7. In complete analogy to this lemma and with similar arguments like the ones made for the stability proof in Lemma 6.2.2, we are able to additionally derive the estimate

$$\|\mathcal{R}f - \mathcal{R}_k f\| \leq C_{\mathcal{R}} k^{d/2} \theta^k \|f\|_{L^2}^2, \quad (6.13)$$

for a constant $C_{\mathcal{R}}$ and $0 < \theta < 1$. Furthermore, we note

$$\mathcal{I}_H u^{\text{PG}} = \mathcal{I}_H u \quad (6.14)$$

and the calculation

$$\begin{aligned} \|u\|^2 &= a(u, u) \\ &= (f, u) \\ &\leq \|f\|_{L^2(\Omega)} \|u\|_{L^2(\Omega)}, \\ &\leq \|f\|_{L^2(\Omega)} \alpha^{-1/2} C_p \|u\|_{L^2(\Omega)}, \end{aligned} \quad (6.15)$$

where the last step contains Poincaré's inequality with constant C_p . Recall also

$$\begin{aligned} u &= u^{\text{PG}} + \mathcal{R}f, \\ u_k &= u_k^{\text{PG}} + \mathcal{R}_k f. \end{aligned} \quad (6.16)$$

Finally, we can put everything together to

$$\begin{aligned} \|u - u_k\| &\stackrel{(6.16)}{\leq} \|u^{\text{PG}} - u_k^{\text{PG}}\| + \|(\mathcal{R} - \mathcal{R}_k)f\| \\ &\stackrel{(6.12)}{\leq} (1 + \bar{\gamma}^{-1}) (\|(Q - Q_k)\mathcal{I}_H u^{\text{PG}}\| + \|(\mathcal{R} - \mathcal{R}_k)f\|) \\ &\stackrel{(6.13)}{\leq} C_Q C_{\mathcal{R}} (1 + \bar{\gamma}^{-1}) k^{d/2} \theta^k (\|\mathcal{I}_H u^{\text{PG}}\| + \|f\|_{L^2(\Omega)}) \\ &\stackrel{(4.18)}{\leq} C_Q C_{\mathcal{R}} (1 + \bar{\gamma}^{-1}) k^{d/2} \theta^k (\|\mathcal{I}_H u\| + \|f\|_{L^2(\Omega)}) \\ &\stackrel{(6.14)}{=} C_Q C_{\mathcal{R}} (1 + \bar{\gamma}^{-1}) k^{d/2} \theta^k (\|\mathcal{I}_H u\| + \|f\|_{L^2(\Omega)}) \\ &\stackrel{(3.16)}{\leq} C_Q C_{\mathcal{R}} (1 + \bar{\gamma}^{-1}) k^{d/2} \theta^k (\alpha^{-1/2} \beta^{1/2} C_{\mathcal{I}_H} \|u\| + \|f\|_{L^2(\Omega)}) \\ &\stackrel{(6.15)}{\leq} \bar{C} k^{d/2} \theta^k \|f\|_{L^2(\Omega)}, \end{aligned}$$

where the constant \bar{C} is defined by $\bar{C} := (1 + \bar{\gamma}^{-1})(1 + \beta^{1/2} \alpha^{-1} C_{\mathcal{I}_H} C_p)$. This implies the assertion. \square

6.2.3 The PG-LOD approximation with the reference LOD space

In the next approach we define an auxiliary PG-LOD that makes use of the reference LOD space \tilde{V}^{LOD} . However, we do not use the reference bilinear form \tilde{a} . This surely has a high advantage in terms of computational costs, since no finescale corrector has to be computed. We define the auxiliary PG-LOD in the following.

6.2.4 Definition (Auxiliary localized PG-LOD approximation with reference correctors)

The PG-LOD approximation of u in (5.3) with reference correctors is to find $\check{u}_k^{\text{PG}} \in \tilde{V}_k^{\text{LOD}}$

such that, for all $v \in V_H$, it holds that

$$a(\check{u}_k^{\text{PG}}, v) = F(v) - a(\tilde{\mathcal{R}}_k f, v), \quad (6.17)$$

where the full approximation is defined by $\check{u}_k = \check{u}_k^{\text{PG}} + \tilde{\mathcal{R}}_k f$.

First, we show the stability and subsequently, we derive the error bound for the approximation \check{u}_k .

6.2.3.1 Stability

We derive the related inf-sup condition, stated in the following lemma.

6.2.5 Lemma (Stability for the PG-LOD method with the reference LOD space) There exists a constant $\check{\gamma} > 0$ such that the inf-sup condition for a with respect to the trial space \tilde{V}_k^{LOD} and the test space V_H

$$\inf_{\tilde{w}_k^{\text{LOD}} \in \tilde{V}_k^{\text{LOD}}} \sup_{v \in V_H} \frac{|a(\tilde{w}_k^{\text{LOD}}, v)|}{\|\tilde{w}_k^{\text{LOD}}\| \|v\|} \geq \check{\gamma}_k, \quad (6.18)$$

is satisfied, where $\bar{\gamma}$ is dependent on the error indicator e_u , defined in Definition 6.1.1.

PROOF: We first focus on an estimate we immediately obtain

$$\begin{aligned} \|\mathcal{Q}_k w - \tilde{\mathcal{Q}}_k w\|^2 &= \left\| \sum_{T \in \mathcal{T}_H} (\mathcal{Q}_k^T w - \tilde{\mathcal{Q}}_k^T w) \right\|^2 \\ &\leq C' k^d \sum_{T \in \mathcal{T}_H} \|\mathcal{Q}_k^T w - \tilde{\mathcal{Q}}_k^T w\|_{U_k(T)}^2 \\ &\stackrel{(6.2)}{\leq} C' k^d e_u^2 \|w\|^2, \end{aligned} \quad (6.19)$$

with the help of the error indicator e_u and with a constant C . We use this estimate to derive the desired inf-sup condition

$$\begin{aligned} \inf_{\tilde{w}_k^{\text{LOD}} \in \tilde{V}_k^{\text{LOD}}} \sup_{v \in V_H} \frac{|a(\tilde{w}_k^{\text{LOD}}, v)|}{\|\tilde{w}_k^{\text{LOD}}\| \|v\|} &\geq \inf_{w \in V_H} \sup_{v \in V_H} \frac{|a(w - \mathcal{Q}_k w, v)| - |a(\mathcal{Q}_k w - \tilde{\mathcal{Q}}_k w, v)|}{(\|w - \mathcal{Q}_k w\| + \|\mathcal{Q}_k w - \tilde{\mathcal{Q}}_k w\|) \|v\|} \\ &\stackrel{(6.19)}{\geq} \inf_{w \in V_H} \sup_{v \in V_H} \frac{|a(w - \mathcal{Q}_k w, v)| - Ck^{d/2} e_u \|w - \mathcal{Q}_k w\| \|v\|}{(\|w - \mathcal{Q}_k w\| + Ck^{d/2} e_u \|w\|) \|v\|} \\ &\stackrel{(4.19)}{\geq} \inf_{w \in V_H} \sup_{v \in V_H} \frac{|a(w - \mathcal{Q}_k w, v)| - Ck^{d/2} e_u \|w - \mathcal{Q}_k w\| \|v\|}{(1 + Ck^{d/2} e_u) \|w - \mathcal{Q}_k w\| \|v\|} \\ &\stackrel{(6.5)}{\geq} \frac{\check{\gamma}_k - Ck^{d/2} e_u}{1 + Ck^{d/2} e_u} =: \check{\gamma}_k. \end{aligned}$$

In order to find a lower bound for $\check{\gamma}_k$, we can use e_u to control the effect of k in the constant. The maximum e_u of all $e_{u,T}$ for all $T \in \mathcal{T}_H$. If we set a tolerance $\text{TOL}(k)$ and recompute the correctors for every T' in case that $e_{u,T'} \geq \text{TOL}(k)$, the particular indicator vanishes, i.e. $e_{u,T'} = 0$. Using this strategy for every $T \in \mathcal{T}_H$ yields $e_u < \text{TOL}(k)$. Thus, we conclude that there exists a $\check{\gamma} > 0$ such that $\check{\gamma} \leq \check{\gamma}_k$. \square

6.2.3.2 Error

We now aim to derive an error that relates the localized PG-LOD approximation u_k in Definition 5.1.2 and the solution of the auxiliary PG-LOD approximation \check{u}_k in Definition 6.2.4. This approximation is due to the reference LOD space and the perturbed bilinear form.

6.2.6 Lemma (Error for the localized PG-LOD and the localized PG-LOD with reference finescale space) With the localized PG-LOD approximation u_k in Definition 3.5.3 and the localized PG-LOD approximation with reference finescale space \check{u}_k , it holds that

$$\| \| u_k - \check{u}_k \| \| \leq \check{C} k^{d/2} \max(e_u, e_f) \| f \|_{L^2(\Omega)}, \quad (6.20)$$

for a constant \check{C} and the error indicators e_u and e_f , defined in Definition 6.1.1.

PROOF: We proceed analogously to Lemma 6.2.3, since we are again able to derive a best approximation result. First, we combine (5.7) and (6.17) and add an arbitrary $u_I \in \tilde{V}_k^{\text{LOD}}$ in order to gain

$$a(\check{u}_k^{\text{PG}} - u_I, v) = a(u_k^{\text{PG}} - u_I, v) - a(\mathcal{R}_k - \tilde{\mathcal{R}}_k)f, v). \quad (6.21)$$

Moreover, we use (6.18) to apply the procedure of (6.10), which yields

$$\check{\gamma} \| \| \check{u}_k^{\text{PG}} - u_I \| \| \leq \| \| u_k^{\text{PG}} - u_I \| \| + \| \| (\mathcal{R}_k - \tilde{\mathcal{R}}_k)f \| \|. \quad (6.22)$$

Further,

$$u_k^{\text{LOD}} = \mathcal{I}_H u_k^{\text{PG}} - \mathcal{Q}_k \mathcal{I}_H u_k^{\text{PG}}$$

inspires us to set

$$u_I := \mathcal{I}_H u_k^{\text{PG}} - \tilde{\mathcal{Q}}_k \mathcal{I}_H u_k^{\text{PG}} \in V_k^{\text{LOD}},$$

which consequents with (6.22) to

$$\| \| u_k^{\text{PG}} - \check{u}_k^{\text{PG}} \| \| \leq (1 + \check{\gamma}^{-1}) \| \| (\mathcal{Q}_k - \tilde{\mathcal{Q}}_k) \mathcal{I}_H u_k^{\text{PG}} \| \| + \check{\gamma}^{-1} \| \| (\mathcal{R}_k - \tilde{\mathcal{R}}_k)f \| \|. \quad (6.23)$$

The final estimate is also in analogy to Lemma 6.2.3,

$$\begin{aligned} \|\|u_k - \check{u}_k\|\| &\leq \|\|u_k^{\text{PG}} - \check{u}_k^{\text{PG}}\|\| + \|\|(\mathcal{R}_k - \tilde{\mathcal{R}}_k)f\|\| \\ &\stackrel{(6.23)}{\leq} (1 + \check{\gamma}^{-1})(\|\|(\mathcal{Q}_k - \tilde{\mathcal{Q}}_k)\mathcal{I}_H u_k^{\text{PG}}\|\| + \|\|(\mathcal{R}_k - \tilde{\mathcal{R}}_k)f\|\|). \end{aligned} \quad (6.24)$$

Obviously, we are able to use the estimate for (6.19) for $\|\|(\mathcal{Q}_k - \tilde{\mathcal{Q}}_k)\mathcal{I}_H u_k^{\text{PG}}\|\|$. Furthermore, we can derive a similar approach for the remaining part $\|\|(\mathcal{R}_k - \tilde{\mathcal{R}}_k)f\|\|$. In particular, we use e_f , analogously to (6.19), to gain the estimate

$$\|\|\mathcal{R}_k f - \tilde{\mathcal{R}}_k f\|\| \leq Ck^{d/2} e_f \|f\|_{L^2(\Omega)}. \quad (6.25)$$

We also know that

$$\begin{aligned} \gamma \|\|u_k\|\| &\leq \sup_{v \in V_H} \frac{|a(u_k, v)|}{\|\|v\|\|} \\ &= \sup_{v \in V_H} \frac{(f, v)}{\|\|v\|\|} \\ &\stackrel{\text{C.S.}}{\leq} \alpha^{-1/2} C_p \|f\|_{L^2(\Omega)}. \end{aligned} \quad (6.26)$$

as well as $\mathcal{I}_H u_k^{\text{LOD}} = \mathcal{I}_H u_k$. Adding everything together obtains

$$\begin{aligned} \|\|u_k - \check{u}_k\|\| &\leq (1 + \check{\gamma}^{-1})(\|\|(\mathcal{Q}_k - \tilde{\mathcal{Q}}_k)\mathcal{I}_H u_k^{\text{PG}}\|\| + \|\|(\mathcal{R}_k - \tilde{\mathcal{R}}_k)f\|\|) \\ &\stackrel{(6.19)}{\leq} (1 + \check{\gamma}^{-1})Ck^{d/2}(e_u \|\| \mathcal{I}_H u_k \|\| + e_f \|f\|_{L^2(\Omega)}) \\ &\stackrel{(3.16)}{\leq} (1 + \check{\gamma}^{-1})Ck^{d/2} \max(e_u, e_f)(C_{\mathcal{I}_H} \alpha^{-1/2} \beta^{1/2} \|\|u_k\|\| + \|f\|_{L^2(\Omega)}) \\ &\stackrel{(6.26)}{\leq} \check{C}k^{d/2} \max(e_u, e_f) \|f\|_{L^2(\Omega)}, \end{aligned}$$

where \check{C} is defined by $\check{C} := C(1 + \check{\gamma}^{-1})(\gamma^{-1} C_{\mathcal{I}_H} C_p \beta^{1/2} \alpha^{-1} + 1)$. Thus, the proof is done. \square

6.2.4 The localized PG-LOD approximation of the reference problem

The last method we need to investigate is the reference localized PG-LOD approximation \tilde{u}_k in Definition 5.1.2. It takes the reference LOD space and the reference bilinear form. Obviously, we would be able to deduce the stability in terms of the energy norm induced by the reference coefficient \tilde{A} . This follows analogously to our approach for the perturbed localized PG-LOD u_k . However, in order to gain an error estimate for $\|\|u - \tilde{u}_k\|\|$, we also have to investigate $\|\|\check{u}_k - \tilde{u}_k\|\|$. Therefore, we derive the stability of the reference problem with respect to the perturbed energy norm $\|\|. \|\|$.

6.2.4.1 Stability

The following lemma outlines our desired stability result.

6.2.7 Lemma (Stability for the PG-LOD method with the reference LOD space) There exists a constant $\tilde{\gamma} > 0$ such that the inf-sup condition for \tilde{a} with respect to the trial space \tilde{V}_k^{LOD} and the test space V_H

$$\inf_{\tilde{w}_k^{\text{LOD}} \in \tilde{V}_k^{\text{LOD}}} \sup_{v \in V_H} \frac{|\tilde{a}(\tilde{w}_k^{\text{LOD}}, v)|}{\|\tilde{w}_k^{\text{LOD}}\| \|v\|} \geq \tilde{\gamma}_k \quad (6.27)$$

is satisfied, where $\tilde{\gamma}_k$ is dependent on the error indicator e_u , defined in Definition 6.1.1.

PROOF: The proof of this lemma, once more, makes use of the error indicator e_u . We need to emphasize that, for all $T \in \mathcal{T}_H$, we have $\tilde{Q}_k^T \in V^f(U_k(T))$ and $\tilde{Q}_k = \sum_{T \in \mathcal{T}_H} \tilde{Q}_k^T$. Having this in mind allows for

$$\begin{aligned} & \inf_{\tilde{w}_k^{\text{LOD}} \in \tilde{V}_k^{\text{LOD}}} \sup_{v \in V_H} \frac{|\tilde{a}(\tilde{w}_k^{\text{LOD}}, v)|}{\|\tilde{w}_k^{\text{LOD}}\| \|v\|} \\ & \geq \inf_{\tilde{w}_k^{\text{LOD}} \in \tilde{V}_k^{\text{LOD}}} \sup_{v \in V_H} \frac{|((\tilde{A} - A + A)\nabla \tilde{w}_k^{\text{LOD}}, \nabla v)|}{\|\tilde{w}_k^{\text{LOD}}\| \|v\|} \\ & \geq \inf_{\tilde{w}_k^{\text{LOD}} \in \tilde{V}_k^{\text{LOD}}} \sup_{v \in V_H} \left(\frac{|a(\tilde{w}_k^{\text{LOD}}, v)|}{\|\tilde{w}_k^{\text{LOD}}\| \|v\|} - \frac{|((\tilde{A} - A)\nabla \tilde{w}_k^{\text{LOD}}, \nabla v)|}{\|\tilde{w}_k^{\text{LOD}}\| \|v\|} \right) \\ (6.18) \quad & \geq \tilde{\gamma}_k - \inf_{w \in V_H} \sup_{v \in V_H} \frac{|\sum_T ((\tilde{A} - A)(\chi_T \nabla - \nabla \tilde{Q}_k^T)w, \nabla v)_{U_k(T)}|}{\|w - \tilde{Q}_k w\| \|v\|} \\ \text{C.S.} \quad & \geq \tilde{\gamma}_k - \inf_{w \in V_H} \sup_{v \in V_H} \frac{\sum_T \|(\tilde{A} - A)A^{-1/2}(\chi_T \nabla - \nabla \tilde{Q}_k^T)w\|_{L^2(U_k(T))} \|A^{1/2} \nabla v\|_{L^2(U_k(T))}}{\|w - \tilde{Q}_k w\| \|v\|} \\ (6.1) \quad & \geq \tilde{\gamma}_k - \inf_{w \in V_H} \sup_{v \in V_H} \frac{\sum_T e_{u,T} \|A^{1/2} \nabla w\|_{L^2(T)} \|A^{1/2} \nabla v\|_{L^2(U_k(T))}}{\|w - \tilde{Q}_k w\| \|v\|} \\ & \geq \tilde{\gamma}_k - \inf_{w \in V_H} \sup_{v \in V_H} \frac{e_u \|A^{1/2} \nabla w\|_{L^2(\Omega)} Ck^{d/2} \|v\|}{\|w - \tilde{Q}_k w\| \|v\|} \\ & \geq \tilde{\gamma}_k - \inf_{w \in V_H} \sup_{v \in V_H} \frac{Ck^{d/2} e_u \|A^{1/2} \nabla \mathcal{I}_H(w - \tilde{Q}_k w)\|_{L^2(\Omega)}}{\|w - \tilde{Q}_k w\|} \\ (3.16) \quad & \geq \tilde{\gamma}_k - \inf_{w \in V_H} \sup_{v \in V_H} \frac{C_{\mathcal{I}_H} \alpha^{-1/2} \beta^{1/2} Ck^{d/2} e_u \|w - \tilde{Q}_k w\|}{\|w - \tilde{Q}_k w\|} \\ & \geq \tilde{\gamma}_k - C' k^{d/2} e_u =: \tilde{\gamma}_k. \end{aligned}$$

Analogously to the previous stability approach, the tolerance $\text{TOL}(k)$ can be used to control e_u which justifies the existence of a lower bound $\tilde{\gamma} > 0$ with $\tilde{\gamma} \leq \tilde{\gamma}_k$. \square

Hence, the reference localized PG-LOD is stable with respect to the perturbed energy norm. We use this fact in the remaining error analysis.

6.2.4.2 Error

We derive an error estimate for $\|\check{u}_k - \tilde{u}_k\|$ in order to incorporate it with the estimates, made in the previous approaches. Recall that \check{u}_k , defined in Definition 6.2.4, denotes the approximation with reference LOD space and perturbed bilinear form. The approximation \tilde{u}_k results from Definition 5.1.1. The error bound is formulated in the following.

6.2.8 Lemma (Error for the localized PG-LOD with reference finescale space and reference localized PG-LOD) With the localized PG-LOD approximation \check{u}_k with reference finescale space in Definition 6.2.4 and the localized reference PG-LOD approximation \tilde{u}_k it holds that

$$\|\check{u}_k - \tilde{u}_k\| \leq \tilde{C}k^{d/2} \max(e_u, e_f) \|f\|_{L^2(\Omega)}, \quad (6.28)$$

for a constant \tilde{C} , $0 < \theta < 1$ and the error indicators e_u and e_f , defined in Definition 6.1.1.

PROOF: At first, we realize that \check{u}_k and \tilde{u}_k base on the same LOD space \tilde{V}_k^{LOD} and thus, we know that they only differ in terms of the PG-LOD approach, which means

$$\|\check{u}_k - \tilde{u}_k\| = \|\check{u}_k^{\text{PG}} - \tilde{u}_k^{\text{PG}}\|. \quad (6.29)$$

Another key element marks the use of

$$\begin{aligned} \tilde{w}_k^{\text{LOD}} &= w_H - \tilde{Q}_k w_H \\ &= \tilde{I}_H w_H - \tilde{Q}_k \tilde{I}_H w_H \\ &= (I - \tilde{Q}_k) \tilde{I}_H (w_H - \tilde{Q}_k w_H) \\ &= (I - \tilde{Q}_k) \tilde{I}_H \tilde{w}_k^{\text{LOD}}, \end{aligned} \quad (6.30)$$

for all $\tilde{w}_k^{\text{LOD}} \in \tilde{V}_k^{\text{LOD}}$. We start with combining (5.8) and (6.17) to gain, for every $v \in V_H$,

$$\begin{aligned} a(\check{u}_k^{\text{PG}}, v) - \tilde{a}(\tilde{u}_k^{\text{PG}}, v) &= \tilde{a}(\tilde{\mathcal{R}}_k f, v) - a(\tilde{\mathcal{R}}_k f, v) \\ &= ((\tilde{A} - A) \nabla \tilde{\mathcal{R}}_k f, \nabla v) \\ &= \sum_{T \in \mathcal{T}_H} ((\tilde{A} - A) \nabla \tilde{\mathcal{R}}_k^T f, \nabla v). \end{aligned} \quad (6.31)$$

Since we are going to apply the inf-sup stability (6.18), we need to derive a factor $a(\check{u}_k^{\text{PG}} - \tilde{u}_k^{\text{PG}}, v)$ on the left hand side. We do this by incorporating the factor $\tilde{a}(\tilde{u}_k^{\text{PG}}, v)$ and yield

$$\begin{aligned}
|a(\check{u}_k^{\text{PG}} - \tilde{u}_k^{\text{PG}}, v)| &\stackrel{(6.31)}{=} |\tilde{a}(\tilde{u}_k^{\text{PG}}, v) - a(\tilde{u}_k^{\text{PG}}, v) + \sum_{T \in \mathcal{T}_H} ((\tilde{A} - A) \nabla \tilde{\mathcal{R}}_k^T f, \nabla v)| \\
&= |((\tilde{A} - A) \nabla \tilde{u}_k^{\text{PG}}, \nabla v) + \sum_{T \in \mathcal{T}_H} ((\tilde{A} - A) \nabla \tilde{\mathcal{R}}_k^T f, \nabla v)| \\
&\stackrel{(6.30)}{=} |((\tilde{A} - A) (\nabla - \nabla \tilde{\mathcal{Q}}_k) \mathcal{I}_H \tilde{u}_k^{\text{PG}}, \nabla v) + \sum_{T \in \mathcal{T}_H} ((\tilde{A} - A) \nabla \tilde{\mathcal{R}}_k^T f, \nabla v)| \\
&= \left| \sum_{T \in \mathcal{T}_H} \left(((\tilde{A} - A) (\chi_T \nabla - \nabla \tilde{\mathcal{Q}}_k^T) \mathcal{I}_H \tilde{u}_k^{\text{PG}}, \nabla v) + ((\tilde{A} - A) \nabla \tilde{\mathcal{R}}_k^T f, \nabla v) \right) \right| \\
&\leq \left| \sum_{T \in \mathcal{T}_H} \left(\|(\tilde{A} - A) A^{-1/2} (\chi_T \nabla - \nabla \tilde{\mathcal{Q}}_k^T) \mathcal{I}_H \tilde{u}_k^{\text{PG}}\|_{L^2(U_k(T))} + \right. \right. \\
&\quad \left. \left. \|(\tilde{A} - A) A^{-1/2} \nabla \tilde{\mathcal{R}}_k^T f\|_{L^2(U_k(T))} \right) \|A^{1/2} \nabla v\|_{L^2(U_k(T))} \right| \\
&\stackrel{\text{C.S.}}{\leq} \stackrel{(6.1)}{\left(\sum_{T \in \mathcal{T}_H} e_{u,T}^2 \|\mathcal{I}_H \tilde{u}_k^{\text{PG}}\|^2 + e_{f,T}^2 \|f\|_{L^2(T)}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_H} \|v\|_{U_k(T)}^2 \right)^{1/2}} \\
&\leq Ck^{d/2} \max(e_u, e_f) \left(\|\mathcal{I}_H \tilde{u}_k^{\text{PG}}\| + \|f\|_{L^2(\Omega)} \right) \|v\|.
\end{aligned} \tag{6.32}$$

Now, we proceed in analogy to (6.26) and obtain with the help of (6.27)

$$\check{\gamma} \|\check{u}_k\| \leq \alpha^{-1/2} C_p \|f\|_{L^2(\Omega)}. \tag{6.33}$$

In total, we conclude with the help of $\|\mathcal{I}_H \check{u}_k^{\text{PG}}\| = \|\mathcal{I}_H \tilde{u}_k\|$ and the interpolation inequality that

$$\begin{aligned}
\|\check{u}_k - \tilde{u}_k\| &= \|\check{u}_k^{\text{PG}} - \tilde{u}_k^{\text{PG}}\| \\
&\stackrel{(6.18)}{\leq} \check{\gamma}^{-1} \sup_{v \in V_H} \frac{|a(\check{u}_k^{\text{PG}} - \tilde{u}_k^{\text{PG}}, v)|}{\|v\|} \\
&\stackrel{(6.32)}{\leq} \check{\gamma}^{-1} Ck^{d/2} \max(e_u, e_f) \left(\|\mathcal{I}_H \tilde{u}_k^{\text{PG}}\| + \|f\|_{L^2(\Omega)} \right) \\
&\stackrel{(3.16)}{\leq} \tilde{C}k^{d/2} \max(e_u, e_f) \left(\alpha^{-1/2} \beta^{1/2} C_{\mathcal{I}_H} \|\tilde{u}_k\| + \|f\|_{L^2(\Omega)} \right) \\
&\stackrel{(6.33)}{\leq} \tilde{C}k^{d/2} \max(e_u, e_f) \|f\|_{L^2(\Omega)},
\end{aligned}$$

where $\tilde{C} := C\check{\gamma}^{-1}(\check{\gamma}^{-1}C_pC_{\mathcal{I}_H}\alpha^{-1}\beta^{1/2} + 1)$. □

Finally, all prerequisites to present the main result of this chapter, the justification of the novel method, are covered.

6.2.5 Error bound for reference localized PG-LOD

The previous sections were required to prepare for deriving the desired error bound for $\|u - \tilde{u}_k\|$ that also contains the error indicators e_u and e_f . This error estimate forms the basis for the novel method proposal in Definition 5.1.3.

6.2.9 Theorem (Error bound for perturbed and reference localized PG-LOD) Let u be the exact solution of the perturbed problem (5.2) and $\tilde{u}_k = \tilde{u}_k^{\text{PG}} + \tilde{\mathcal{R}}_k f$ the solution of the localized PG-LOD of the reference problem (5.8), for $k \in \mathbb{N}$. Furthermore, choose TOL small enough such that there exists a lower bound for \tilde{y}_k and \check{y} . Then the error bound

$$\|u - \tilde{u}_k\| \leq ck^{d/2}(\theta^k + \text{TOL})\|f\|_{L^2(\Omega)} \quad (6.34)$$

holds true, where c is independent of H and the patch size k .

PROOF: We combine the previous results from Lemma 6.2.3, Lemma 6.2.6 and Lemma 6.2.8 and observe

$$\begin{aligned} \|u - \tilde{u}_k\| &\leq \|u - u_k\| + \|u_k - \check{u}_k\| + \|\check{u}_k - \tilde{u}_k\| \\ &\leq k^{d/2}(\bar{C}\theta^k + (\check{C} + \tilde{C}) \max(e_u, e_f)) \max(e_u, e_f) \|f\|_{L^2(\Omega)} \\ &\leq ck^{d/2}(\theta^k + \max(e_u, e_f)) \|f\|_{L^2(\Omega)}, \end{aligned}$$

which completes the proof. □

7 Implementation

In the previous chapters, we examined the LOD method by Målqvist and Peterseim and its slight variations from a rather analytic point of view. This chapter is devoted to a presentation of the implementation of the LOD. Besides, we present further explanations for the novel method. The LOD method as well as the PG-LOD method are applicable in practice, which has already been proven by many simulations in recent papers (more details in Chapter 9). We only focus on the main aspects of the implementation. Moreover, we keep the assumption of homogeneous Dirichlet boundary conditions, which simplifies the work. For a detailed algebraic explanation and more special cases such as non-homogeneous and Neumann boundary conditions or Eigenvalue problems, we highly recommend the work of Målqvist, Peterseim, Henning and Engwer in [13].

7.1 Discretization

In order to simulate numerical experiments, we need a discrete setting. The discretization of the space V is achieved by a shape-regular fine mesh \mathcal{T}_h , where h denotes the maximal diameter of the fine elements. In our case this fine mesh is a refinement of the coarse mesh \mathcal{T}_H of the coarse FE space V_H (see Figure 7.1). This means, every coarse mesh element

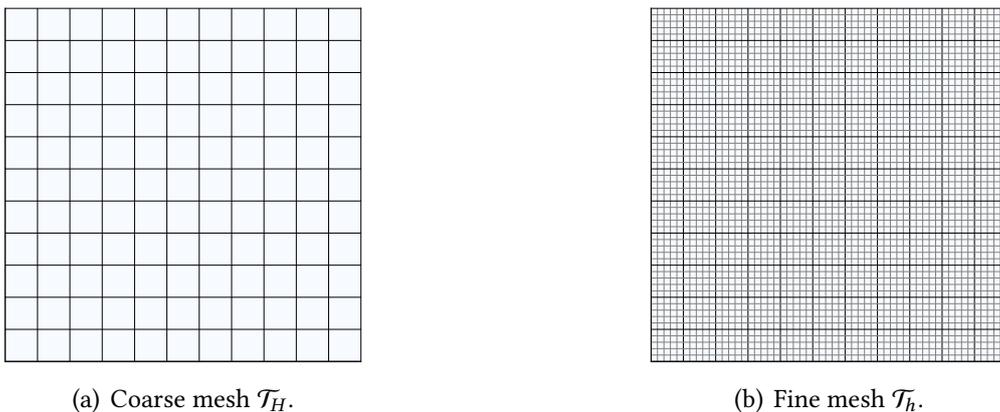


Figure 7.1: Refinement of a quadrilateral coarse mesh.

consists of finitely many fine mesh elements, since $h < H$, and crucially, the finescale elements lay in only one single coarse element. Furthermore, we set N_H and N_h to be the

number of coarse respectively of fine mesh elements. Analogously to Section 2.2 a FE space $\mathcal{P}_1(\mathcal{T}_h)$ can be derived and we yield

$$V_h := V \cap \mathcal{P}_1(\mathcal{T}_h).$$

We assume that this fine discretization is fine enough in order to capture all microscopic features. Thus, to find an $u_h \in V_h$

$$a(u_h, v) = F(v), \quad \forall v \in V_h,$$

is the most accurate FEM. However, this method is not computable, due to high computational complexity or memory issues. Instead, we use this fine mesh to compute the finescale correctors on each coarse mesh element that belongs to \mathcal{T}_H . Since we are computing these correctors on only a small patch, the memory issue is small enough to bypass a high complexity. Surely, the finescale space V^f has also a discretized version, which reads

$$V_h^f(U_k(T)) = V^f(U_k(T)) \cap V_h.$$

For every $v_H \in V_H$, the correctors are computed by a FEM, i.e, they contain solutions $\mathcal{Q}_k^{T,h} v_H \in V^f(U_k(T))$ of

$$a_{U_k(T)}(\mathcal{Q}_k^{T,h} v_H, w_h) = a_T(v_H, w_h) \quad \forall w_h \in V_h^f(U_k(T)), \quad (7.1)$$

and accordingly

$$\mathcal{Q}_k^h v := \sum_{T \in \mathcal{T}_H} \mathcal{Q}_k^{T,h} v.$$

The resulting discretized LOD space is consequently defined by

$$V_{H,k}^{\text{LOD},h} = V_H - \mathcal{Q}_k^h V_H,$$

We end up with the discretized standard LOD method, that is to find $u_{H,k}^{\text{LOD},h} \in V_{H,k}^{\text{LOD},h}$ such that, for all $v \in V_{H,k}^{\text{LOD},h}$, it holds that

$$a(u_{H,k}^{\text{LOD},h}, v) = F(v).$$

Målqvist and Peterseim showed in [28] that this discrete setting still satisfies similar results in terms of the error analysis we presented in Chapter 4. Our next purpose is to sketch the implementational aspects of the standard LOD.

7.2 The standard LOD

The LOD is basically a coarse Galerkin method, based on a finescale space that is computed with the help of an additional Galerkin method on the patch of each coarse element. Thus, the LOD approximation is the result of a system of linear equations with respect to the coarse basis functions. The mesh size is small enough to prevent memory issues and to reduce computational time. The expensive fine mesh is only used to compute the correctors. Once

$$\mathbf{Q}_k^h \mathbf{v} := \sum_{T \in \mathcal{T}_H} \mathbf{Q}_k^{T,h} \mathbf{v}$$

is computed, we can add the correctors to the coarse stiffness matrix $\mathcal{S}_H^{\text{LOD}} \in \mathbb{R}^{N_H \times N_H}$ with entries

$$\mathcal{S}_H^{\text{LOD}}[m][n] := \begin{cases} a(\phi_n + \mathbf{Q}_k^h \phi_n, \phi_m + \mathbf{Q}_k^h \phi_m), & \text{for } z_n, z_m \in \mathcal{N}, \\ 0, & \text{else,} \end{cases}$$

where ϕ_n denotes the basis function of the interior node $z_n \in \mathcal{N}$. Note that \mathcal{N} denotes the set of all interior nodes in V_H . Similarly, the load vector $\mathcal{L}_H \in \mathbb{R}^{N_H}$ can be determined with

$$\mathcal{L}_H^{\text{LOD}}[m] := \begin{cases} F(\phi_m + \mathbf{Q}_k^h \phi_m), & \text{for } z_m \in \mathcal{N}, \\ 0, & \text{else.} \end{cases}$$

Subsequently, we solve the resulting linear system of equations

$$\mathcal{S}_H^{\text{LOD}} \mathbf{u}_H^{\text{LOD}} = \mathcal{L}_H^{\text{LOD}},$$

with the solution $\mathbf{u}_H^{\text{LOD}} \in \mathbb{R}^{N_H}$. We get the final standard LOD approximation by

$$\mathbf{u}^{\text{LOD}} = \sum_{m=0}^{N_H-1} \mathbf{u}_H^{\text{LOD}}[m] (\phi_m + \mathbf{Q}_k^h \phi_m).$$

We realize that this procedure of the standard LOD requires a communication between the correctors. In particular, in order to compute $\mathcal{S}_H^{\text{LOD}}$, all correctors need to be available and stored. This consequents from the fact that the test space is also chosen to be the finescale space $V_{H,k}^{\text{LOD},h}$. Certainly, this circumstance might also become an issue for saving memories. This problem can be resolved by a method, earlier introduced in Chapter 3, the Petrov Galerkin LOD. The advantages are pointed out in the subsequent section.

7.3 The Petrov Galerkin LOD

The Petrov Galerkin LOD was helpful for the analysis of the novel method in Chapter 5 and 6. It also is the method we use for the numerical experiments. As we have already

shown, the difference between the PG-LOD and the standard LOD is the choice of the test space. The test space for the PG-LOD is equal to V_H , whereby the standard LOD uses the finescale space $V_{H,k}^{\text{LOD},h}$. Thus, only the trial space of the PG-LOD equals the finescale space. Obviously, the coarse stiffness matrix of the PG-LOD $\mathcal{S}_H^{\text{PG}} \in \mathbb{R}^{N_H \times N_H}$ contains entries of the form

$$\mathcal{S}_H^{\text{PG}}[m][n] := \begin{cases} a(\phi_n + \mathcal{Q}_k^h \phi_n, \phi_m), & \text{for } z_n, z_m \in \mathcal{N}, \\ 0, & \text{else.} \end{cases}$$

The load vector $\mathcal{L}_H^{\text{PG}} \in \mathbb{R}^{N_H}$ equals the load vector of the standard FEM

$$\mathcal{L}_H^{\text{PG}}[m] := \begin{cases} F(\phi_m), & \text{for } z_m \in \mathcal{N}, \\ 0, & \text{else,} \end{cases}$$

whereas the linear system of equations reads

$$\mathcal{S}_H^{\text{PG}} u_H^{\text{PG}} = \mathcal{L}_H^{\text{PG}}.$$

Thereafter, the solution $u_H^{\text{PG}} \in \mathbb{R}^{N_H}$ can be incorporated to

$$u^{\text{PG}} = \sum_{m=0}^{N_H-1} u_H^{\text{PG}}[m](\phi_m + \mathcal{Q}_k^h \phi_m).$$

The costs to compute $\mathcal{S}_H^{\text{PG}}$ decrease in comparison to $\mathcal{S}_H^{\text{LOD}}$, as there is no communication between correctors required. It is possible to delete the corrector right after the whole support of the basis function has been included.

7.4 Right hand side Correction

In complete analogy to the element corrector \mathcal{Q} , the right hand side corrector could be computed and stored. As a consequence of the approaches presented in Section 3.5, we gain the right hand side correction as the solution $\mathcal{R}_k^{T,h} f \in V^f(U_k(T))$ of

$$a_{U_k(T)}(\mathcal{R}_k^{T,h} f, w_h) = F_T(w_h) \quad \forall w_h \in V_h^f(U_k(T)), \quad (7.2)$$

and in total, we get

$$\mathcal{R}_k^h f := \sum_{T \in \mathcal{T}_H} \mathcal{R}_k^{T,h} f.$$

For the LOD as well as for the PG-LOD, we add the right hand side correction to the load vector by setting

$$\mathcal{L}_H^{\text{LOD}}[m] := \begin{cases} F(\phi_m + \mathbf{Q}_k^h \phi_m) + a(\mathcal{R}_k^h, \phi_m + \mathbf{Q}_k^h \phi_m), & \text{for } z_m \in \mathcal{N}, \\ 0, & \text{else,} \end{cases}$$

and

$$\mathcal{L}_H^{\text{PG}}[m] := \begin{cases} F(\phi_m) + a(\mathbf{Q}_k^h, \phi_m), & \text{for } z_m \in \mathcal{N}, \\ 0, & \text{else.} \end{cases}$$

This procedure is a direct consequence of Section 3.5. For both cases, u^{LOD} and u^{PG} attain the full solution by adding the correction $\mathcal{R}_k^h f$ appropriately. Clearly, the involvement is optional for our approach.

7.5 The novel method

According to the previous approaches and following the explanations in Section 5.1, Algorithm 1 presents the pseudo code for our novel method. As we already know, the involved method is represented by the PG-LOD with right hand side correction. Since we assume that we will not have a memory issue by saving the corrector functions, we call this algorithm the storage method. For this algorithm, we need to store every corrector

```

Pick  $k, p$ 
Copy  $\tilde{\mathbf{Q}}_k^T \phi_j$  and  $\tilde{\mathcal{R}}_k^T f$  for all  $T$  and  $j$ 
Set  $\hat{\mathcal{R}}_k = 0$ 
for all  $T$  do
    Compute  $e_{u,T}$  and  $e_{f,T}$ 
    Pick  $\text{TOL}(p)$  appropriately to  $p$ 
    if  $\max(e_{u,T}, e_{f,T}) \geq \text{TOL}(p)$  then
        Recompute  $\mathbf{Q}_k^T$  and  $\mathcal{R}_k^T f$ 
        Update stiffness matrix  $\mathcal{S}_{ij} += a_T(\phi_j, \phi_i) - a(\mathbf{Q}_k^T \phi_j, \phi_i)$ 
        Update right hand side  $\mathcal{L}_i += F_T(\phi_i)$ 
         $\hat{\mathcal{R}}_k += \mathcal{R}_k^T f$ 
    else
        Update stiffness matrix  $\mathcal{S}_{ij} += \tilde{a}_T(\phi_j, \phi_i) - \tilde{a}(\tilde{\mathbf{Q}}_k^T \phi_j, \phi_i)$ 
        Update right hand side  $\mathcal{L}_i += F_T(\phi_i)$ 
         $\hat{\mathcal{R}}_k += \mathcal{R}_k^T f$ 
    end
end
end
Solve for  $u_{k,p}^{\text{VC}}$ , by computing  $\mathcal{S}^{-1}b$ 
Compute  $u_{k,p}^{\text{vc}} = u_{k,p}^{\text{VC}} + \hat{\mathcal{R}}_k$ 

```

function as well as to have access to the reference correctors. This requirement arises because of the computation of the error indicators $e_{u,T}$ and $e_{f,T}$. This is addressed in the following section.

7.5.1 Error indicators

Recall both error indicators, defined in Definition 6.1.1, for every $T \in \mathcal{T}_H$, by

$$e_{u,T} = \max_{v|_T | v \in \tilde{V}_H, \|v\|_T=1} \|(\tilde{A} - A)A^{-1/2}(\chi_T \nabla v - \nabla \tilde{\mathcal{Q}}_k^T v)\|_{L^2(U_k(T))},$$

$$e_{f,T} = \frac{\|(\tilde{A} - A)^{-1/2} \nabla \tilde{\mathcal{R}}_k^T f\|_{L^2(U_k(T))}}{\|f\|_{L^2(T)}}.$$

We realize that both indicators contain the appropriate corrector functions. Moreover, computing $e_{f,T}$ is a straight forward calculation for every T , since it only consists of the $L^2(\Omega)$ norms. In order to compute $e_{u,T}$, we need to solve the eigenvalue problem

$$\mathcal{B}x_l = \mu_l \mathcal{C}x_l,$$

where

$$\mathcal{B}_{ij} = ((\tilde{A} - A)^2 A^{-1}(\chi_T \nabla \phi_j - \nabla \mathcal{Q}_k^T \phi_j), \chi_T \nabla \phi_i - \nabla \mathcal{Q}_k^T \phi_i)_{U_k(T)},$$

$$\mathcal{C}_{ij} = (A \nabla \phi_j, \nabla \phi_i)_T,$$

for $i, j = 1, \dots, m-1$. In this case, m denotes the number of basis functions. For further information, we refer to [15]. For this eigenvalue problem and for $e_{f,T}$, we need to store the corrector functions as well as \tilde{A} and A . This procedure might result in a memory problem. Due to the PG-LOD, we do not need communication between the correctors and thus, it is useful to derive an error indicator that also enables the delete of the correctors right after it has been involved to the PG-LOD. Hellman and Målqvist presented an error indicator E_u (and similarly E_f) that only requires the storage of the coefficients and satisfies, for every $T \in \mathcal{T}_H$,

$$e_{u,T}^2 \leq E_{u,T}.$$

This indicator still has similar behavior in terms of the decrease. Our method solely requires a measure of which correctors should be updated, if we aim to update, for instance, 20% of all correctors. Using E_u only changes the choice of TOL. For further details on E_u and e_u and memory consumptions, see [15] and [13].

8 Numerical Experiments

The LOD-method as well as the PG-LOD-method have successfully been applied to various problems in several recent research articles, among others, [28], [16], [18], [30] and [13]. This section is devoted to the application of our novel method to the variational crime of random perturbations. For this purpose, Section 8.5 presents Weakly random problems and performs several numeric simulations. Beforehand, we start with a continuation of the multiscale examples in Section 2.4. All simulations have been performed with Python 2.7 and they base on the code that has been used by Hellman and Målqvist in [15]. We always use a quadratic mesh in a two dimensional setting.

8.1 One dimensional PG-LOD experiment

In Section 2.4, we stated two multiscale problems in order to emphasize the effect of high variations in the diffusion coefficient (see Figure 2.1 and Figure 2.4(a)). We realized that the energy error stays on the same logarithmic level as long as the mesh size H does not capture every microscopic effect. For this case, Figure 2.2 and Figure 2.5 showed that the approximation is indeed quite inaccurate. In Figure 8.1, we see the resulting solution of the standard PG-LOD for the one dimensional problem with various coarse mesh size H and fine mesh size $h = 1024$. Clearly, already $H = 1/8$ achieves a reasonable result as the macroscopic behavior is already correct. Figure 8.2 displays the energy error of each PG-LOD and compares it with the FEM. Remarkably, the accuracy of the method is already very good for $k = 2$ and it does not change significantly for a bigger k . Note that the memory consumption of the PG-LOD increases fast for a large k and H . We furthermore want to emphasize that the amount of variations in the example have intentionally chosen rather small. Thus, the FEM is still potentially cheaper. In total, we suspect that for higher dimensions or variations the PG-LOD as well as the LOD outperforms the FEM. Before we start with the numerical simulations, we need to think about examples.

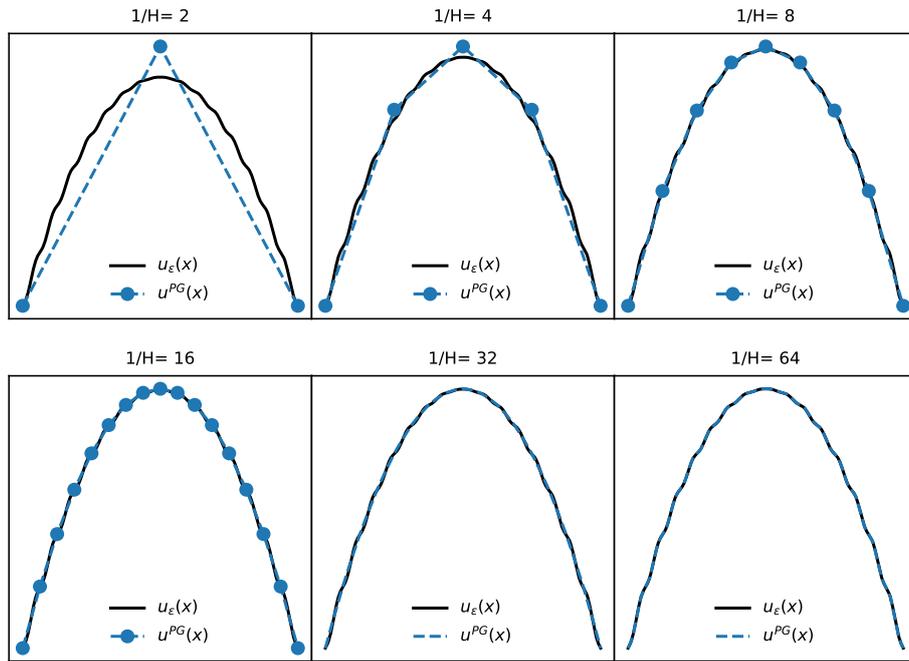


Figure 8.1: PG-LOD approximation of u_ε for various choices of H and $\varepsilon = 2^{-5}$.

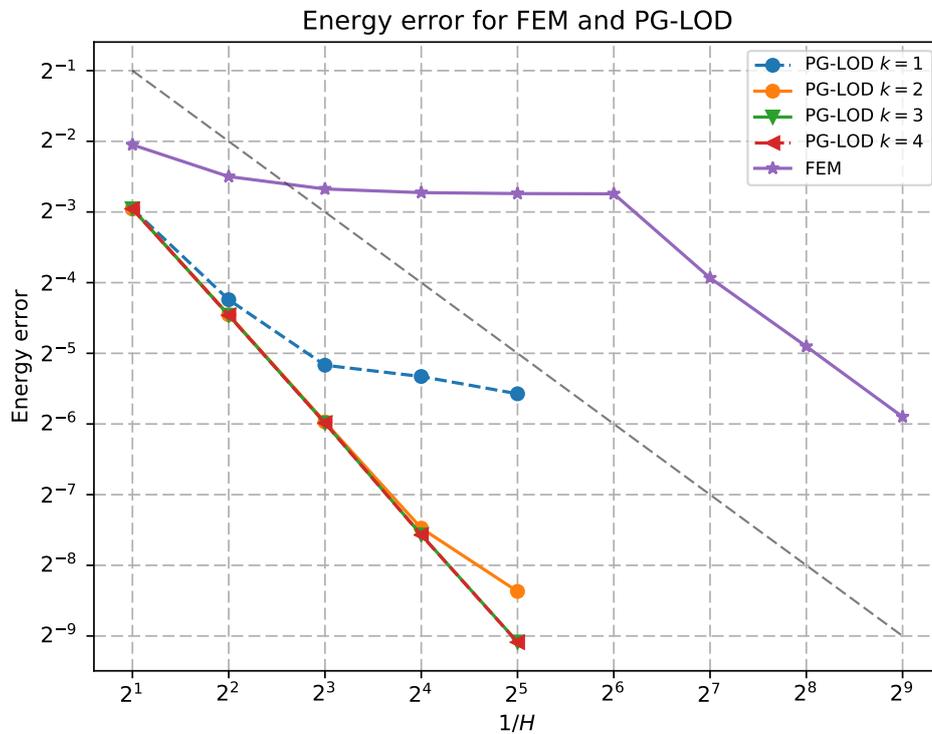


Figure 8.2: Energy error $\|u_\varepsilon - u_H\|$ for $\varepsilon = 2^{-5}$ and for various k .

8.2 Perturbations

In this section, we present various diffusion coefficients and possible perturbations that produce a variational crime in the Galerkin method. In Section 8.5, we get to know the weakly random setting based on [25]. However, in Figure 8.3(c), we already display their standard example for a high variational coefficient. The original diffusion coefficient has two equidistantly distributed values and in the perturbed version, the coefficient is subjected to defects in the form that the entry neutralizes to the other background value and disappears completely. This example is a very particular case for a perturbed problem. We intend to think about some other possibilities of perturbations appearing in

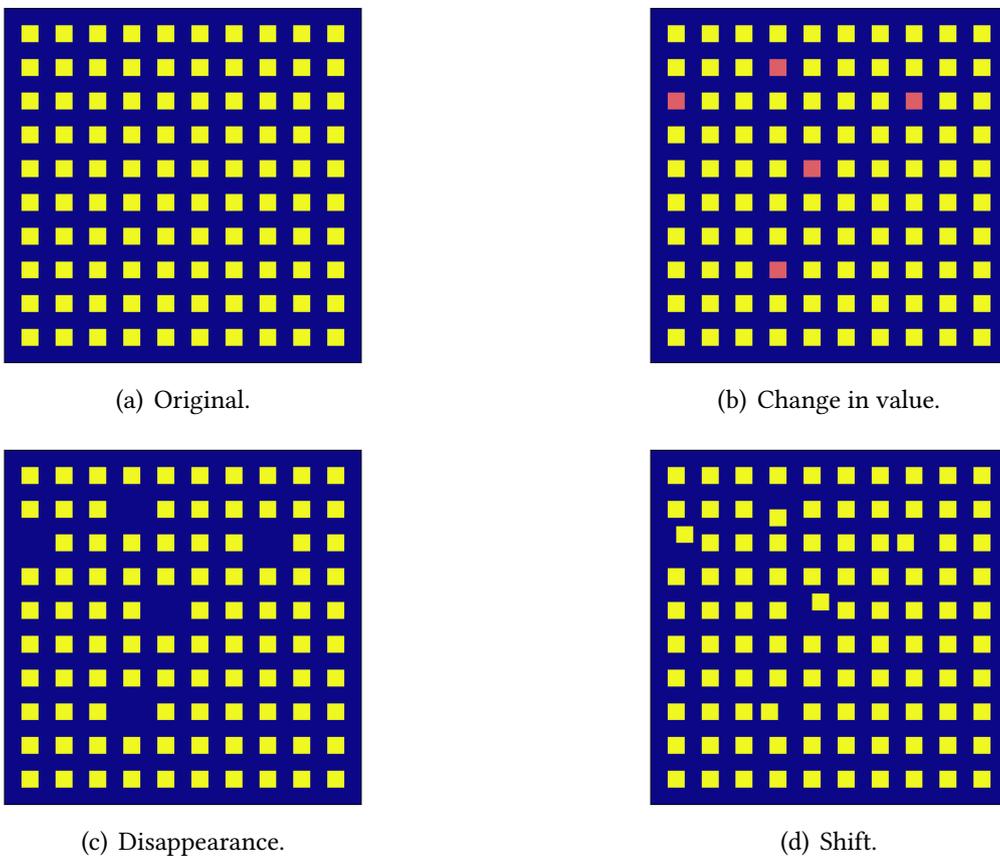
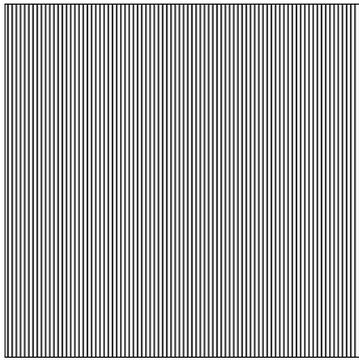


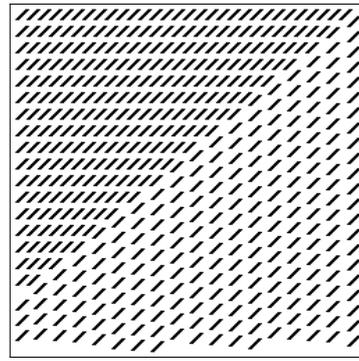
Figure 8.3: 5% defects in a period structure with changing value, shift and disappearance. Blue is 0.05, red is 0.8 and yellow is 1. Motivated by [25].

a certain amount of entries. The entries might change their values or move their position. Figure 8.3 shows those possible perturbations. With a mix of these changes, we are able to achieve many possible perturbations that are based on the reference coefficient. Apart from the standard example in Figure 8.3, there are a lot of instances that belong to the Weakly random problems. In this thesis, we choose additional types, displayed in Figure 8.4. The coefficient with channels in Figure 8.4(a) has already been presented with a

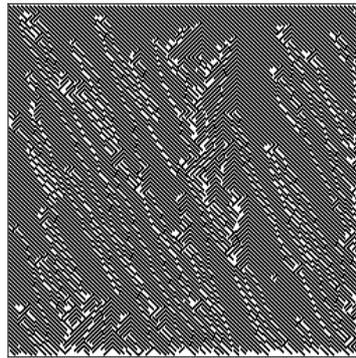
coarser version in Section 2.4. Clearly, every coefficient in Figure 8.4 can be subjected to some perturbation like presented in Figure 8.3. Throughout the entire thesis, the reference coefficient is solely two valued. The black entries have the value 1 and the background (white) is 0.05 which consequents a contrast of 20. Thus, we enable a better comparison and we concentrate on only the essential parts of the simulations.



(a) Coefficient 2, periodic channels.



(b) Coefficient 3, non periodic.



(c) Coefficient 4, non periodic.

Figure 8.4: Instances for diffusion coefficients.

8.3 The novel method

The remarkable feature of the novel method is a mix of old and new corrector functions \mathcal{Q}_k in the space V_k^{LOD} . For each $T \in \mathcal{T}_H$, the error e_u denotes the indicator for recomputation. In the following example, we choose $H = 1/16$, which results in 256 coarse elements, and we take a finer coefficient of the standard example in Figure 8.3(a). Figure 8.5 displays

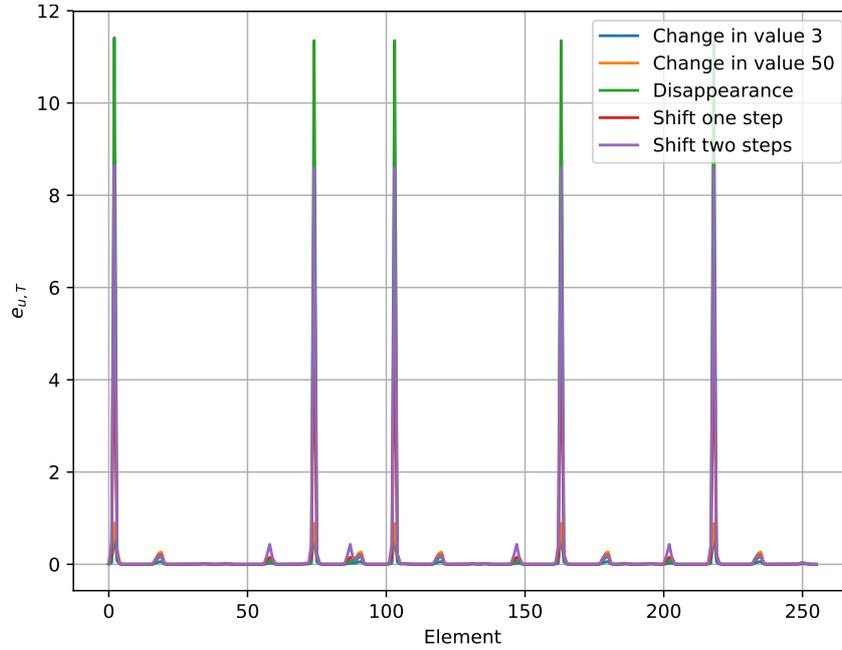


Figure 8.5: Error indicator for the coarse elements.

the error indicator for the coefficient in Figure 8.7, whereas Figure 8.6 only displays the indicator for the elements 70-78. For each element on the x-axis $T \in \mathcal{T}_H$, we get a specific value. In order to enable a better comparison, we also added a two step shift and a much higher change in value. We realize that the impact of the change in value is rather low, even for very large values. The disappearance corresponds to the highest error, whereas the shift is dependent on the particular direction and the step size. We remark that the value of e_u is not sufficient value for the actual impact on the accuracy. Later on, we will see that the shift produces the highest energy error. The indicator $e_{u,T}$ has its maxima in the coarse elements which contain the defects. The effect that is caused by one particular defect has only an impact on the patch with size $k \in \mathbb{N}$. Clearly, the bigger the k , the more coarse element get affected. Figure 8.7 displays the affected correctors for various choices of $k \in \mathbb{N}$. The patch size increases for every k and more elements get affected. Obviously, defects can also appear in multiple elements. In this case, more element correctors get affected. When the error indicator is computed, we choose a certain amount of correctors that we want to recompute. If we take, for example, 20%, then we take the elements with the largest $e_{u,T}$. For this purpose we choose TOL appropriately such that $e_{u,T} \leq \text{TOL}$, for

all $T \in \mathcal{T}_H$. With respect to the error estimate in Chapter 6, we expect a much better approximation of the novel method, whereby we still save 80% recomputation. We want to remark that even 100% recomputing might save some computational effort, since we have $e_{u,T} = 0$ for the correctors that do not get touched at all (see Figure 8.7 for each k).

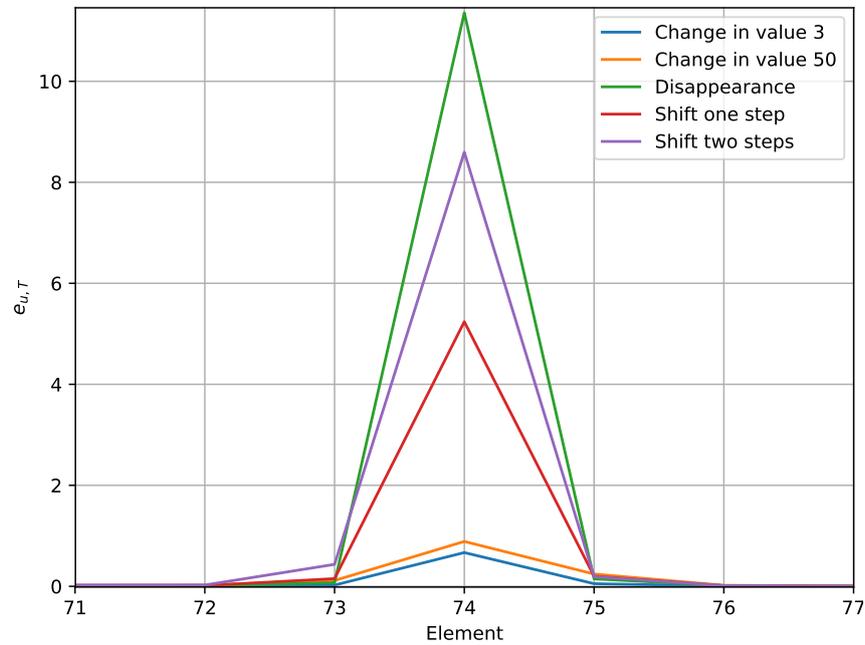


Figure 8.6: Error indicator for particular elements.

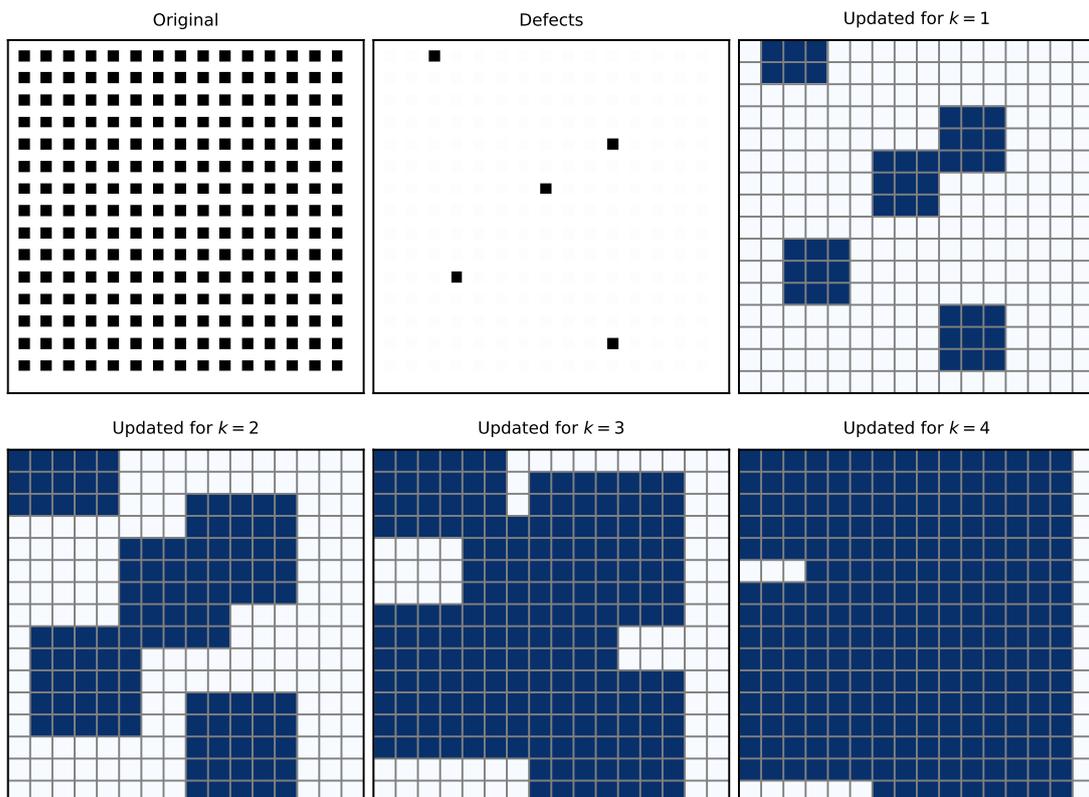


Figure 8.7: Coefficient, defects, and the affected elements for several $k \in \mathbb{N}$. Blue means $e_{u,T} > 0$, white means $e_{u,T} = 0$.

8.4 Numerical examples and discussion

We want to know how much accuracy we actually gain in comparison to the computational consumption we save. For this purpose, we perform an error investigation for each coefficient in Figure 8.4 and the standard coefficient in Figure 8.3(a). Recall that the exact solution we want to approximate u is defined in Definition 5.0.2. With respect to the notation in Section 5.1, the worst PG-LOD error appears for the approximation \tilde{u}_k in Definition 5.1.2, where only the reference correctors are used. It is determined by $\| \|u - \tilde{u}_k\| \|$ and means 0 % recomputing. The best PG-LOD is achieved by updating 100% of the affected correctors. In Definition 5.1.1, we called the result of this strategy u_k and the error is $\| \|u - u_k\| \|$. We expect the approximation error $\| \|u - u_{k,p}^{\text{vc}}\| \|$ that consequents from the novel method approximation $u_{k,p}^{\text{vc}}$, defined in Definition 5.1.3, to be in between the worst and the best case.

$$\| \|u - u_k\| \| \leq \| \|u - u_{k,p}^{\text{vc}}\| \| \leq \| \|u - \tilde{u}_k\| \| \quad (8.1)$$

The goal certainly is to find the best compromise between saving computational effort and accuracy. To compute this errors, we use a FEM on the fine mesh \mathcal{T}_h as a reference solution for u in 5.2 and call it u_h . We set

$$\begin{aligned} e_{\text{ref}} &:= \| \|u_h - \tilde{u}_k\| \|, \\ e_{\text{pert}} &:= \| \|u_h - u_k\| \|, \\ e_{\text{vc}} &:= \| \|u_h - u_{k,p}^{\text{vc}}\| \|. \end{aligned}$$

For each perturbation in Figure 8.9 - 8.12, the subfigures (b),(d) and (f) display the behavior of e_{vc} in comparison to e_{ref} and e_{pert} . Figure 8.13 compares the error for each perturbation for every coefficient. In order to interpret the behavior of the energy error, we furthermore display the error indicator e_u in the subfigures (c),(e) and (g) in Figure 8.9- 8.12, for every coefficient and each perturbation. The comparison between each coefficient is done in Figure 8.14. For every simulation in Figure 8.9 - 8.12, we choose a different two valued coefficient with randomly generated perturbations. The probability is always 1%, except for the channels in Figure 8.10 (2%). In order to enable a comparison between each perturbation, the change in value, the disappearance and the shift always happens at the same place. The change in value is always from the value 1 to 3 which is already rather high. This change in value of plus 200% is realistic and thus, we do not consider higher values (although they gain the same results). The shift is basically one step in terms of the fine mesh to the right. However, we have already determined that the type of the shift does not have a high impact. Figure 8.8 shows the perturbations for each coefficient. We use a coarse mesh size $H = 1/16 = 2^{-4}$ and a fine mesh size $h = 1/256 = 2^{-8}$. Furthermore,

the localization parameter is $k = 4$. With respect to the amount of perturbations, this patch size consequents that indeed every element corrector is affected.

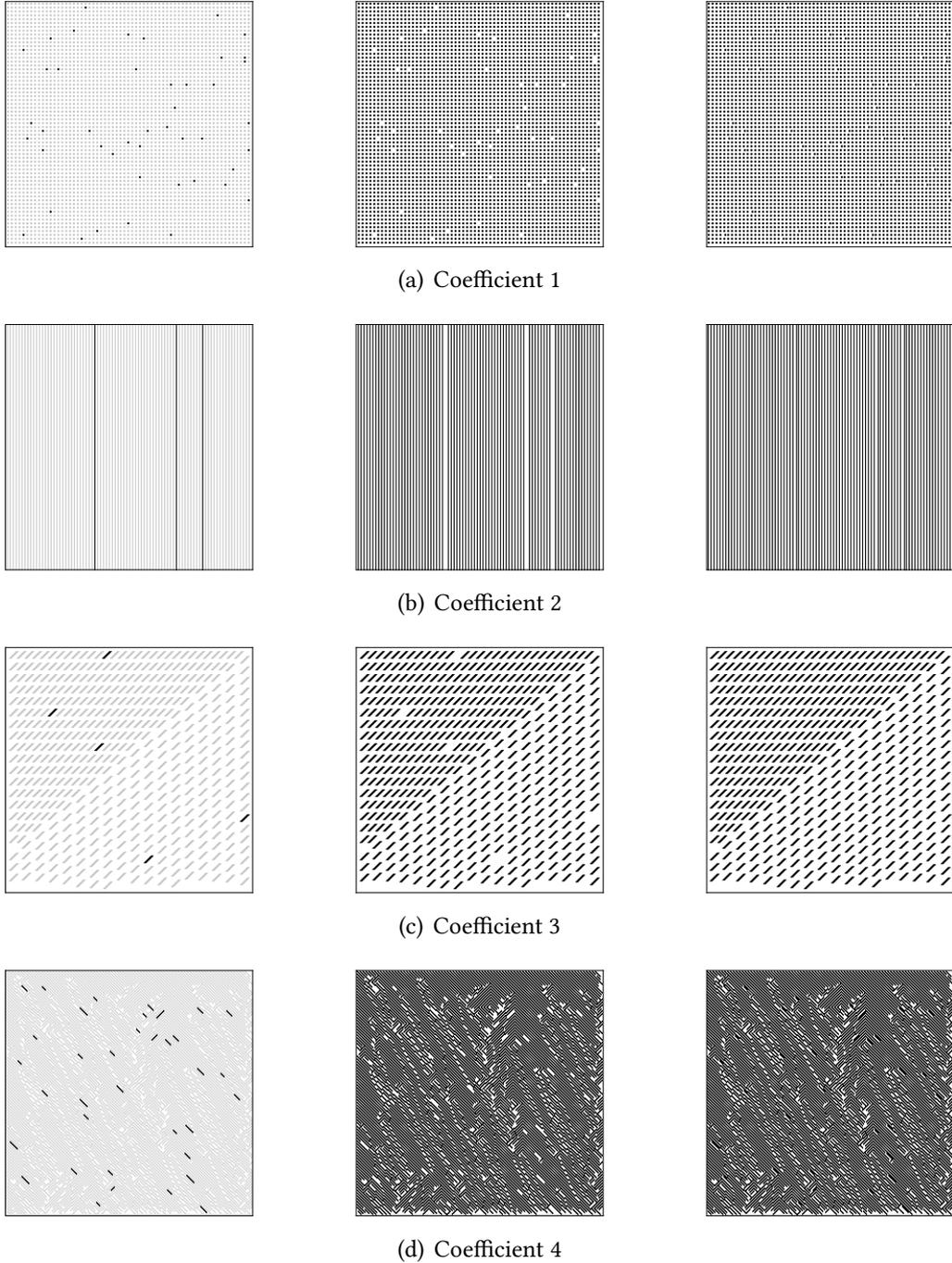
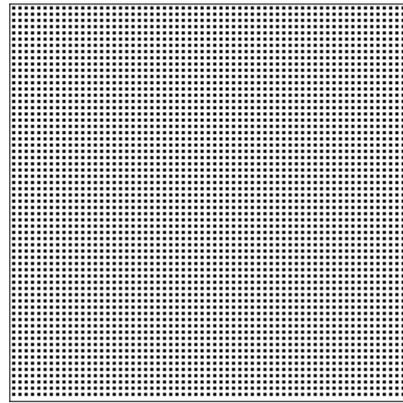
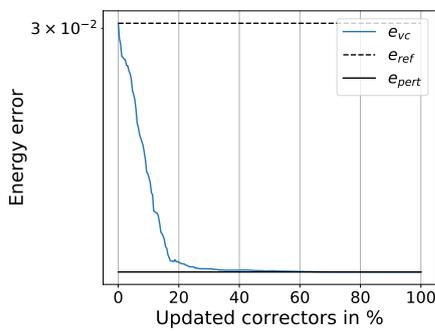


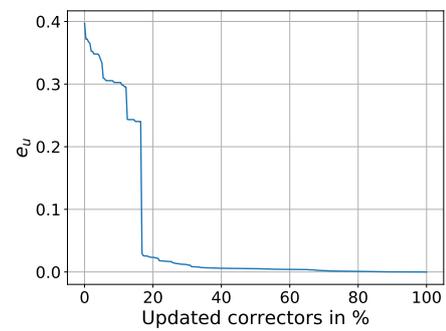
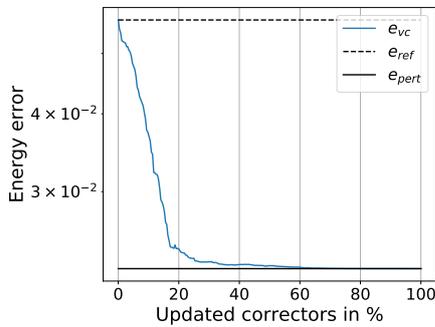
Figure 8.8: Perturbations for each coefficient. Left: change in value, middle: disappearance, right: shift.



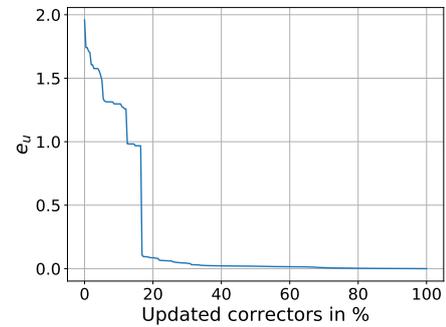
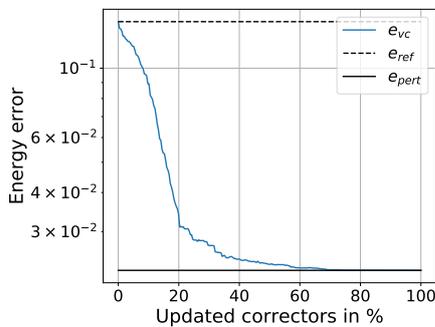
(a) Coefficient



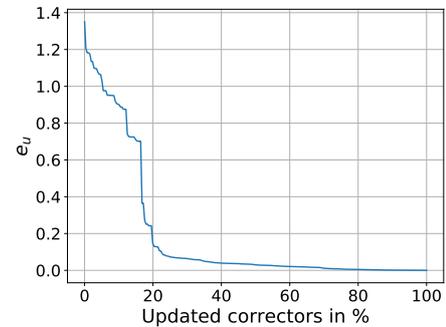
(b) Energy error for change in value.

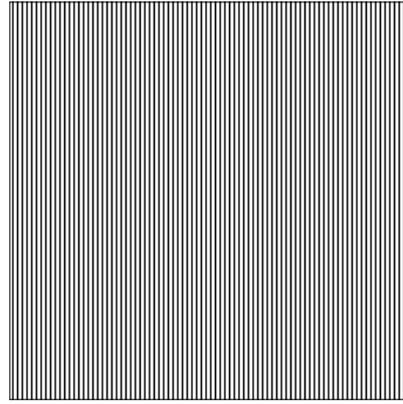
(c) Indicator e_u for change in value.

(d) Energy error for disappearance.

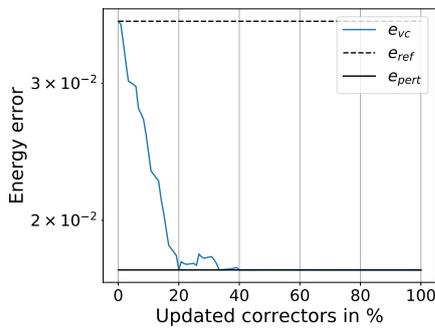
(e) Indicator e_u for disappearance.

(f) Energy error for shift.

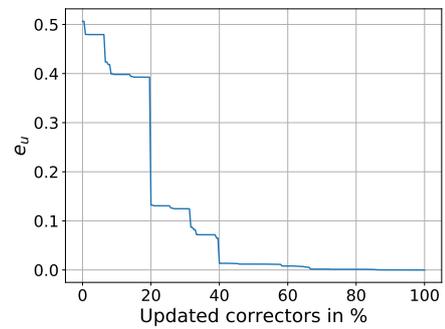
(g) Indicator e_u for shift.Figure 8.9: Energy error comparison and indicator e_u for coefficient 1.



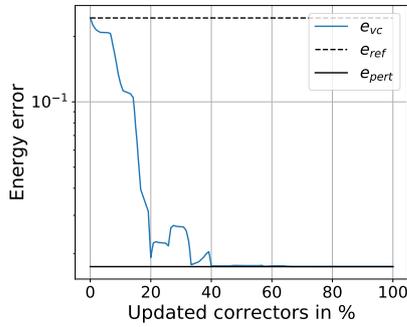
(a) Coefficient



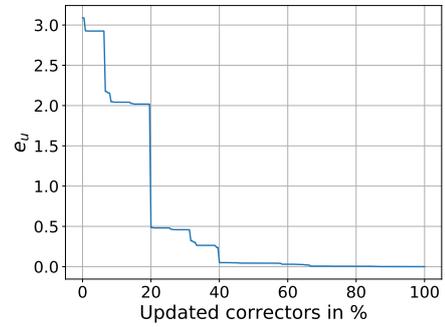
(b) Energy error for change in value.



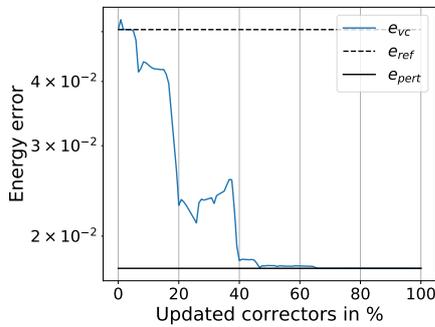
(c) Indicator e_u for change in value.



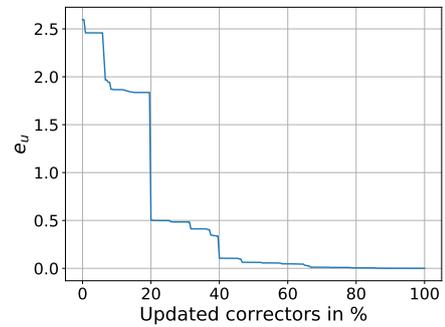
(d) Energy error for disappearance.



(e) Indicator e_u for disappearance.

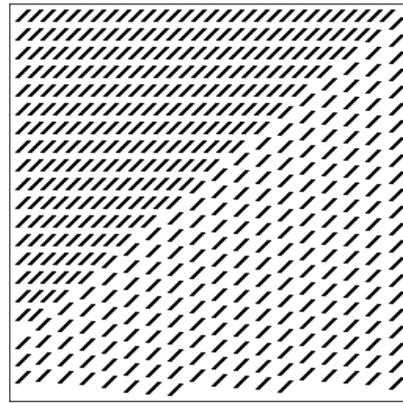


(f) Energy error for shift.

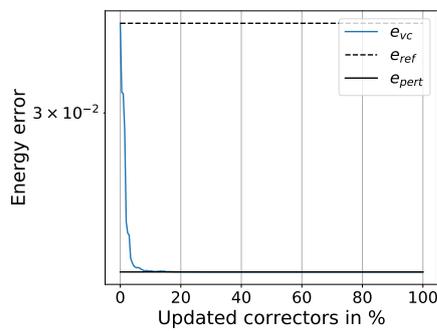


(g) Indicator e_u for shift.

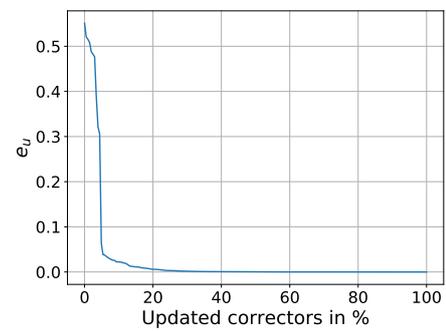
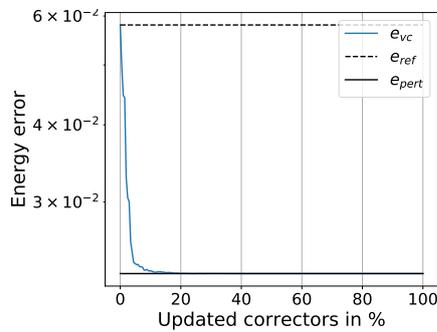
Figure 8.10: Energy error comparison and indicator e_u for coefficient 2.



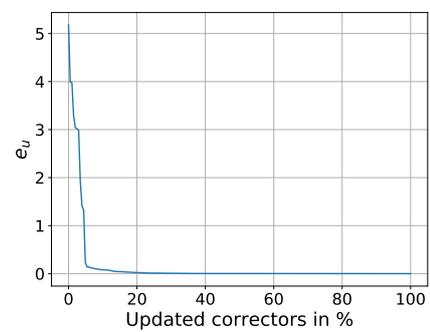
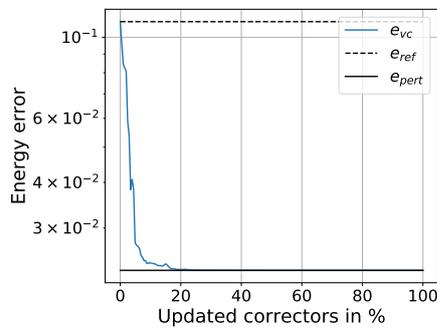
(a) Coefficient



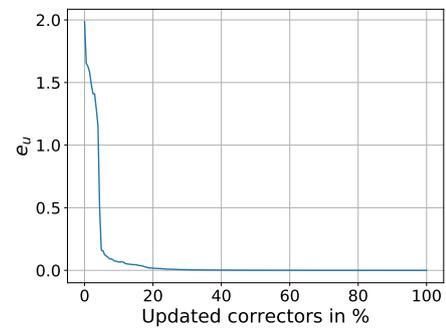
(b) Energy error for change in value.

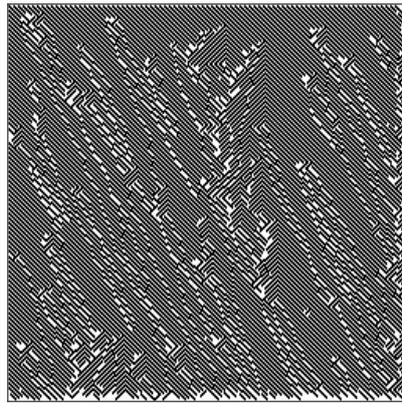
(c) Indicator e_u for change in value.

(d) Energy error for disappearance.

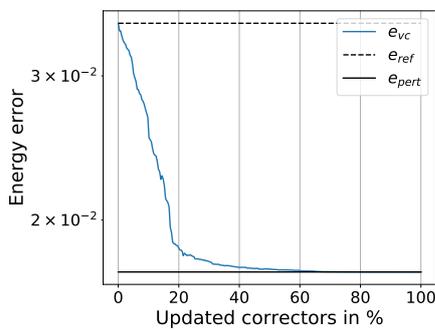
(e) Indicator e_u for disappearance.

(f) Energy error for shift.

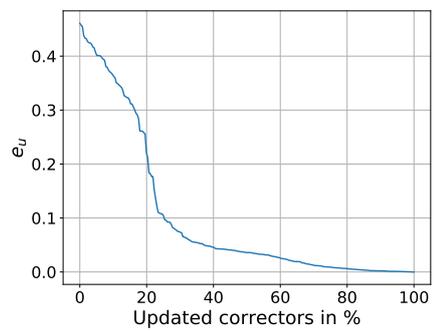
(g) Indicator e_u for shift.Figure 8.11: Energy error comparison and indicator e_u for coefficient 3.



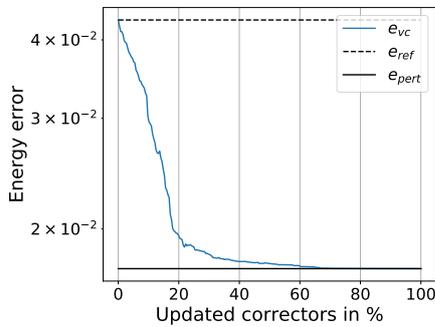
(a) Coefficient



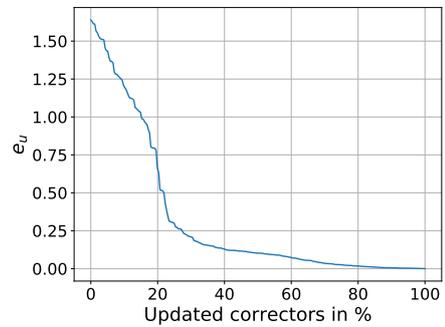
(b) Energy error for change in value.



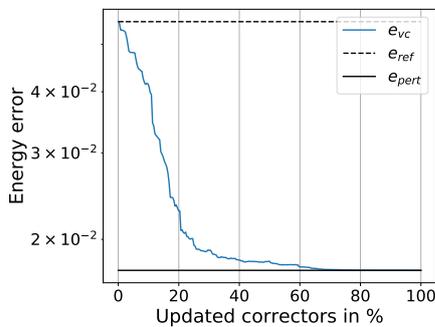
(c) Indicator e_u for change in value.



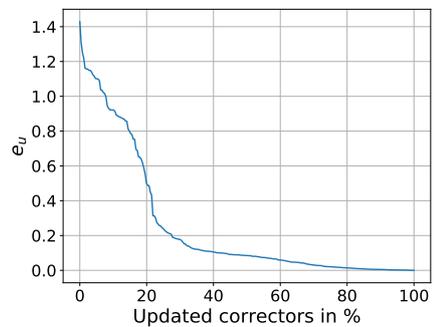
(d) Energy error for disappearance.



(e) Indicator e_u for disappearance.



(f) Energy error for shift.



(g) Indicator e_u for shift.

Figure 8.12: Energy error comparison and indicator e_u for coefficient 4.

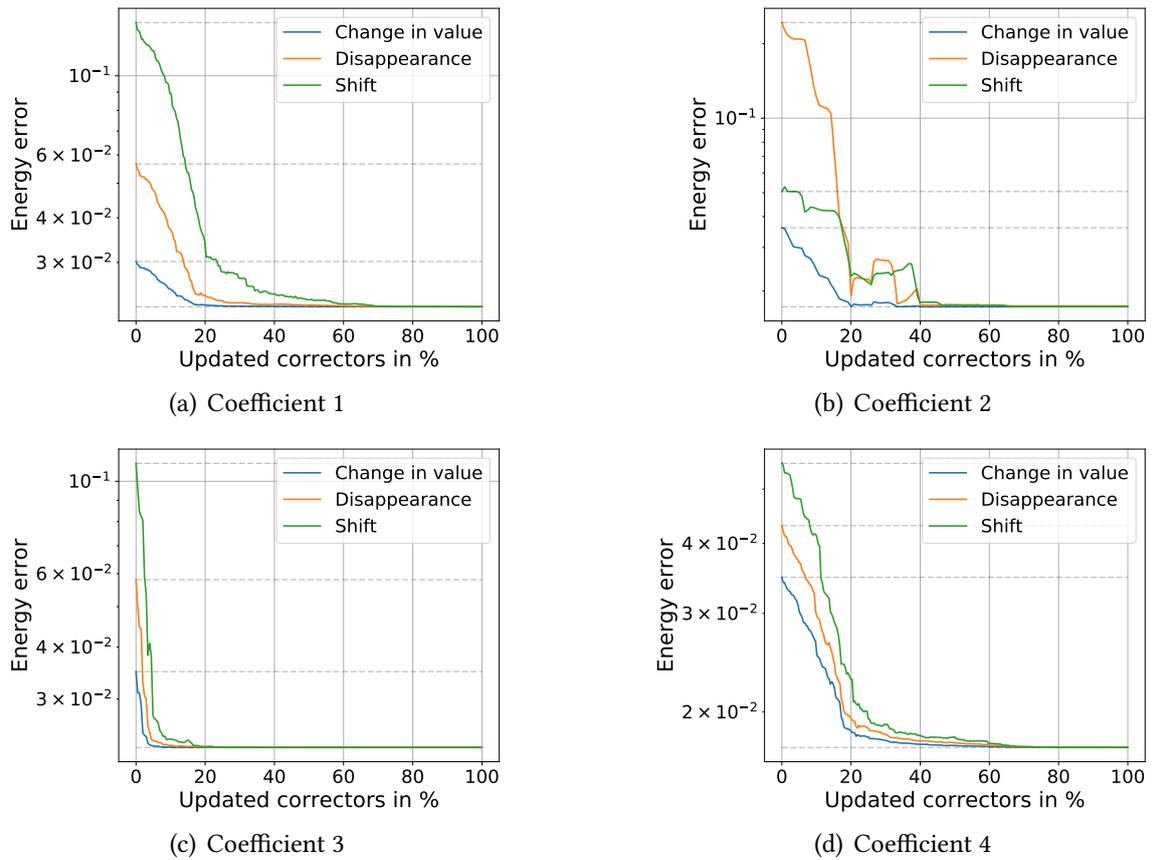


Figure 8.13: Energy errors e_{vc} in comparison.

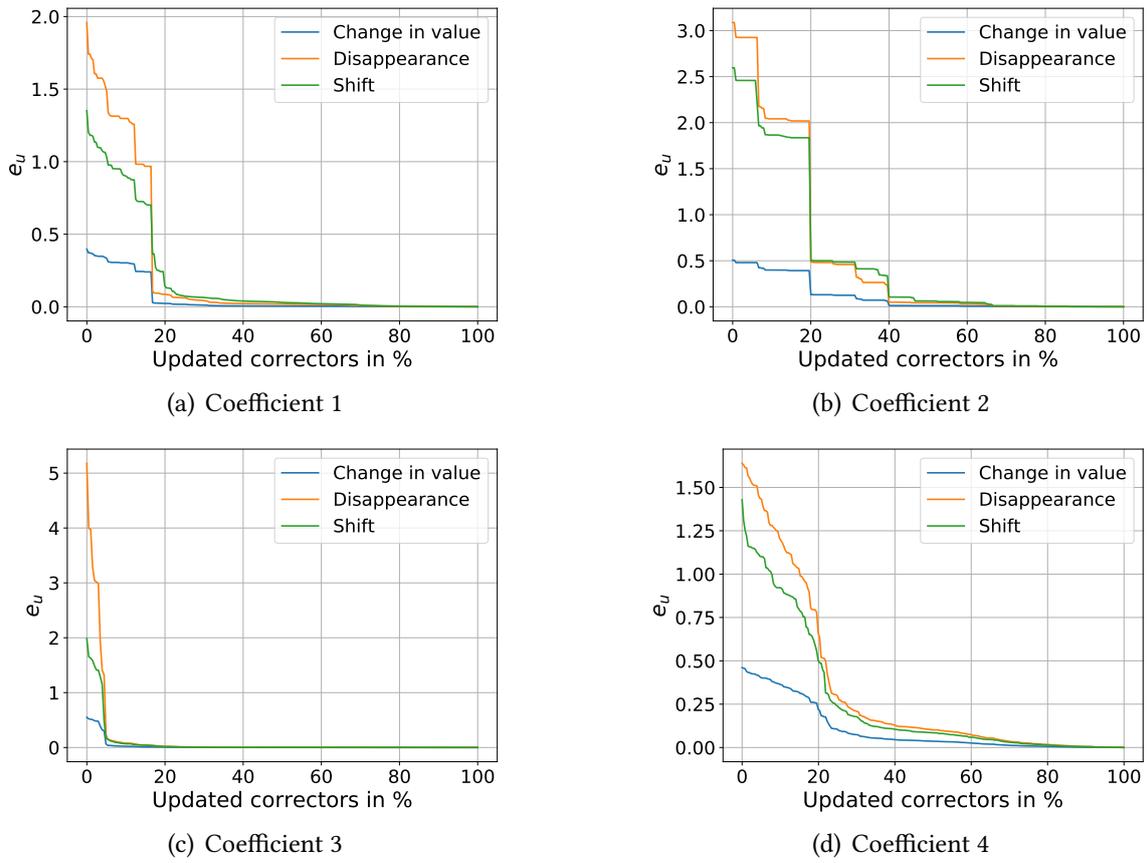


Figure 8.14: Indicator e_u in comparison.

We now discuss the results that are displayed in Figure 8.9-8.14.

- For every coefficient, the behavior of the energy error e_{vc} is quite promising. We yield a graph that decreases fast. In a view of the error indicator $e_{u,T}$ in Figure 8.5, only a few elements get a comparably high value which consequents the peaks. Therefore, those elements with peaks are the first elements for recomputation. With respect to Theorem 6.2.9, we immediately gain a much better energy error for e_{vc} .
- This result can also be determined by the graph of Figure 8.14. The decrease of e_{vc} can also be noticed in e_u . Most of the coefficients possess the best compromise at around 20% of recomputing.
- Regarding Figure 8.13 and Figure 8.14, we conclude that the differences between each perturbation is rather low in terms of the indicator error behavior, since we actually obtain similar graphs. However, the actual absolute impact on the error in Figure 8.13 differs for each coefficient significantly, but it still has the same behavior. Furthermore, we can not distinctly verify which perturbation generally implies the highest impact on the error. Remarkably, the best PG-LOD errors in Figure 8.13 are mostly equal. However, for a very high change in value, the contrast might increase and the error gets naturally bigger.
- The change in value consequents the lowest energy error. Nevertheless, especially the coefficient in Figure 8.10 might produce a high energy error for a drastic change in value.
- Disappearance is more or less a change in value, but while the change in value increases the finescale part, the disappearance eliminates it. This approach is obviously worse, as the disappearance lays always above the change in value. For the channels in Coefficient 2, the disappearance has the highest impact.
- Except for the coefficient in Figure 8.10, the impact of the shift results in the worst error.
- It is remarkable how efficiently the method performs in case of the coefficient in Figure 8.11. The decay of e_{vc} and e_u is very fast. We expect an accurate approach for already 5% recomputing.
- We point out that Figure 8.12 prove that the novel method is applicable for completely non periodic coefficients. We still gain a respectively well performance.

All in all, we conclude that the novel method is applicable for every instance that we simulated in this section. We realize that the decrease of e_{vc} can already be noticed in the behavior of e_u . The different stages of the indicator are also noticeable in the energy error. This can be used in order to detect the best percentage in advance. Thus, a reasonably percentage is already recognizable by a priori considering the error indicator $e_{u,T}$. Now,

we apply our novel PG-LOD method and the knowledge that results from the simulations to the so-called Weakly Random framework.

8.5 Weakly Random Problems

Weakly random problems have been introduced by Le Bris and Legoll in, among others, [25]. The setting is a special case of the variational crime of perturbations, presented in Chapter 5. However, the amount of perturbation underlays a specific probability. In order to apply Monte Carlo methods for stochastic simulations, multiple perturbed multiscale problems based on the same reference problem need to be solved. Le Bris and Legoll develop the weakly stochastic MsFEM as a variation of the MsFEM and investigate it analytically and numerically. Thus, their setting is restricted to periodic instances. Moreover, they assume a periodic-type randomness, which they call weakly randomness. They perform stochastic simulations and apply Monte Carlo methods. The randomly perturbed diffusion problem underlays a probability and reads

$$\begin{cases} -\nabla \cdot (A^{\text{pert}}(x, \omega) \nabla u(x, \omega)) = f(x, \omega), & \text{for } x \in \Omega, \\ u(x, \omega) = 0, & \text{for } x \in \partial\Omega. \end{cases}$$

We do not define a probability space, since we will not make use of it at all. Nevertheless, we want to remark that Le Bris and Legoll define a probability space with special assumptions that they use in the numerical analysis. The coefficient A^{pert} underlies a perturbation that is dependent on ω and still has potentially high variations. Le Bris and Legoll assume this randomness to be restricted by the following property.

8.5.1 Definition (Weakly random property) The coefficient of the weakly random problem A^{pert} can be described by

$$A^{\text{pert}}(x, \omega) := A^{\text{ref}}(x) + \mu A^{\text{rand}}(x, \omega),$$

where $0 < \mu \leq 1$ is a deterministic parameter. The stochastic coefficient A^{rand} contains the same multiscale features like the determined coefficient A^{ref} .

This setting captures a lot of different experiments, as it is a very general formulation. It is important that the resulting stochastic problem is not fully random since the coefficient A^{pert} can be considered as a perturbation of the deterministic coefficient A^{ref} . In Chapter 5 and 6, we showed that every variational crime with a variously high amount of perturbations can be handled by the novel method that we propose in Definition 5.1.3.

Regarding the notation in Chapter 5, we realize

$$\begin{aligned} A &:= A^{\text{pert}}, \\ \tilde{A} &:= A^{\text{ref}}. \end{aligned}$$

We define slightly different continuous and bounded bilinear forms

$$a(v, w, \omega) := \int_{\Omega} (A(x, \omega) \nabla u(x, \omega)) \cdot \nabla v(x) \, dx \quad \forall v \in V, w \in W$$

and

$$\tilde{a}(v, w) := \int_{\Omega} \left(A^{\text{ref}}(x) \nabla v(x) \right) \cdot \nabla w(x) \, dx \quad \forall u \in V, w \in W.$$

The linear functional reads

$$F(w) := \int_{\Omega} f(x) w(x) \, dx \quad \forall w \in W.$$

To be precise, the perturbed problem in Definition 5.0.2, applied to the Weakly random setting, reads as follows.

8.5.2 Definition (Exact stochastic problem) For $V = H_0^1(\Omega)$, a , \tilde{a} and F defined as above, the weak formulation of the stochastic diffusion problem is to find $u(\omega) \in V$ such that, for all $v \in V$, it holds that

$$a(u(\omega), v, \omega) = F(v). \tag{8.2}$$

The reference problem is to find the solution $\tilde{u} \in V$ such that

$$\tilde{a}(\tilde{u}, v) = F(v), \tag{8.3}$$

where \tilde{a} denotes the non-perturbed version of a , defined above.

Furthermore, we assume that the high variations of \tilde{A} and respectively, of A , make it impossible to apply a standard FEM. To enable stochastic results, we have to compute an approximation of the expectation value and the variance. For this purpose, we will use a Monte Carlo method that requires several computations of the same problem, for different values of ω and thus, for a high number of perturbations. At this point, we fall back to the novel method. [25] proposed a lot of different test cases. However, their problems are restricted to some periodic assumptions. We have already mentioned their standard example in Figure 8.3(a). Certainly, our method also works for periodic cases, but it does not make an explicit use of it. Every simulation that we made in the previous section are actually examples of Weakly random problems with several perturbations. It is important to notice that our method covers a lot more scenarios and thus, it has a much

bigger application. However, our method obviously requires a higher computational effort. The fact that our novel method already performs well for 20% recomputing though, justifies the usage.

8.6 Monte Carlo simulations

In order to investigate the accuracy of our approaches, we follow the same strategy like Le Bris and Legoll in [25]. We consider the following error

$$e_{L^2}(u_1, u_2) = \mathbb{E} \left(\frac{\|u_1 - u_2\|_{L^2(\Omega)}}{\|u_2\|_{L^2(\Omega)}} \right)$$

with solutions u_1 and u_2 that we specify later on. To apply the Monte Carlo method in M realizations, we define the sequence of random variables $\{X_m(\omega)\}_{1 \leq m \leq M}$ for every ω with

$$X(\omega) := \frac{\|u_1(\cdot, \omega) - u_2(\cdot, \omega)\|_{L^2(\Omega)}}{\|u_2(\cdot, \omega)\|_{L^2(\Omega)}}.$$

The Monte Carlo method is used to compute the empirical mean μ_M as well as the empirical standard deviation σ_M with

$$\mu_M(X) = \frac{1}{M} \sum_{m=1}^M X_m(\omega), \quad \sigma_M^2(X) = \frac{1}{M-1} \sum_{m=1}^M (X_m(\omega) - \mu_M(X))^2.$$

We can assume that the random variable X_m underlies a normal distribution, which allows for the application of the Central Limit Theorem. Accordingly, we get a confidence interval in the form that

$$|\mathbb{E}(X) - \mu_M(X)| \leq 1.96 \frac{\sigma_M}{\sqrt{M}}.$$

The value 1.96 consequents from the standard level of confidence 95%. In order to compute the confidence interval, we compute μ_M and σ_M with our novel method for variational crimes. In Section 8.4, we verified a first idea for the detection of the best choice of the percentage for recomputation. Obviously, for an arbitrary coefficient, we do not know the perfect choice. However, Section 8.4 showed very similar results for every coefficient. Clearly, the choice is also dependent on the mesh size and the amount of variations. We expect a reasonable result as long as the probability for the random perturbations is appropriately low. Once we know the certain percentage for recomputing, we can apply the novel method to the weakly random setting and perform Monte Carlo simulations. Figure 8.15 displays the result of a Monte Carlo simulation for Coefficient 3 in Section 8.4. This coefficient is displayed in Figure 8.4(b). The coefficient is subjected to 1% of disappearance, displayed in the middle of Figure 8.8(c). For every sample m , we compute

the errors $e_{L^2}(u_h, u_k)$, $e_{L^2}(u_h, u_{k,p}^{\text{vc}})$ and $e_{L^2}(u_k, u_{k,p}^{\text{vc}})$. We set $k = 4$ and we note that u_k corresponds from 100% recomputing and \tilde{u}_k from 0%. The novel method approximation $u_{k,p}^{\text{vc}}$ is due to $p = 0.2$, which results in 20% recomputing. In a view of Figure 8.12(d), we see that 20% corresponds to a reasonable error. Figure 8.16 reveals that the comparison between 100% and 20% is quite promising. Clearly, the confidence interval converges slower compared to the perturbed LOD. With respect to Figure 8.16 and Figure 8.15 we want to remark that a Monte Carlo method might not be the best method for our purposes and thus, the convergence is rather slow. Due to lack of time, we performed no further simulations and left it as a task for the future.

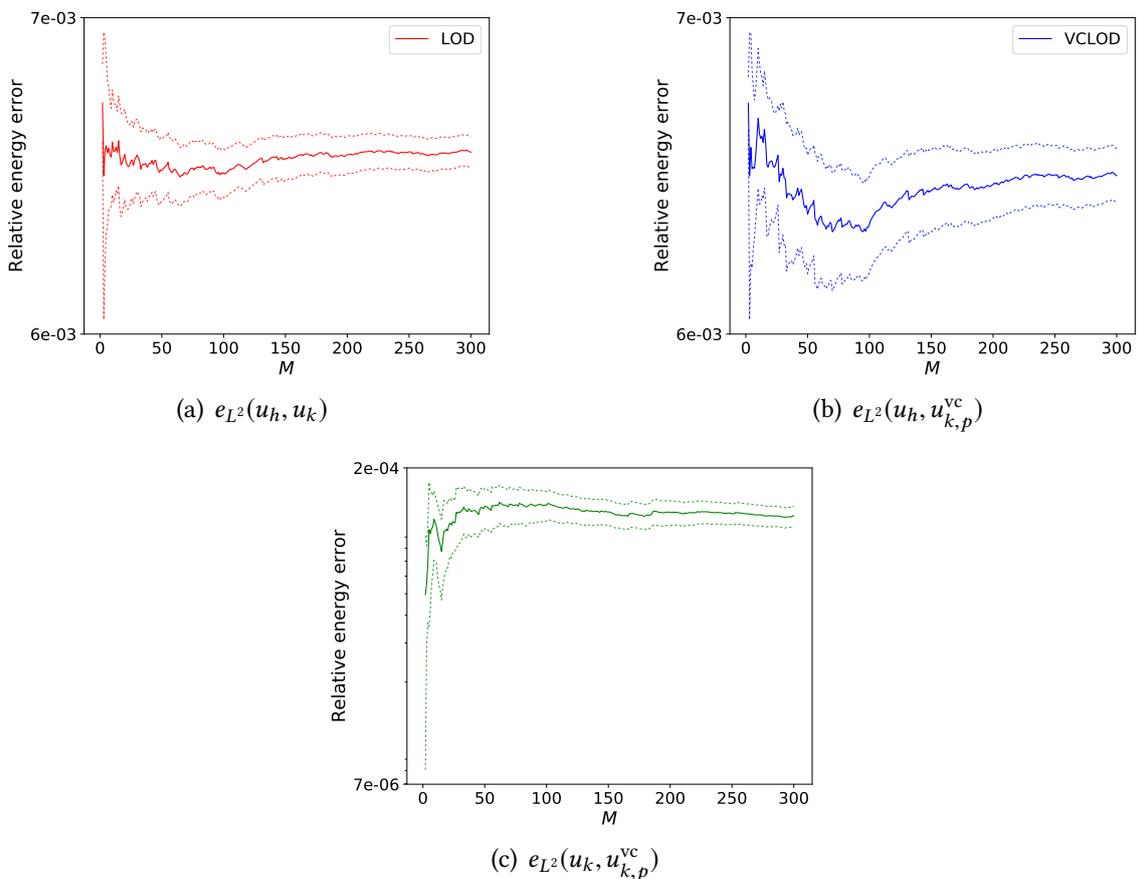


Figure 8.15: Monte Carlo simulation for coefficient 3. The dashed lines represent the confidence interval.

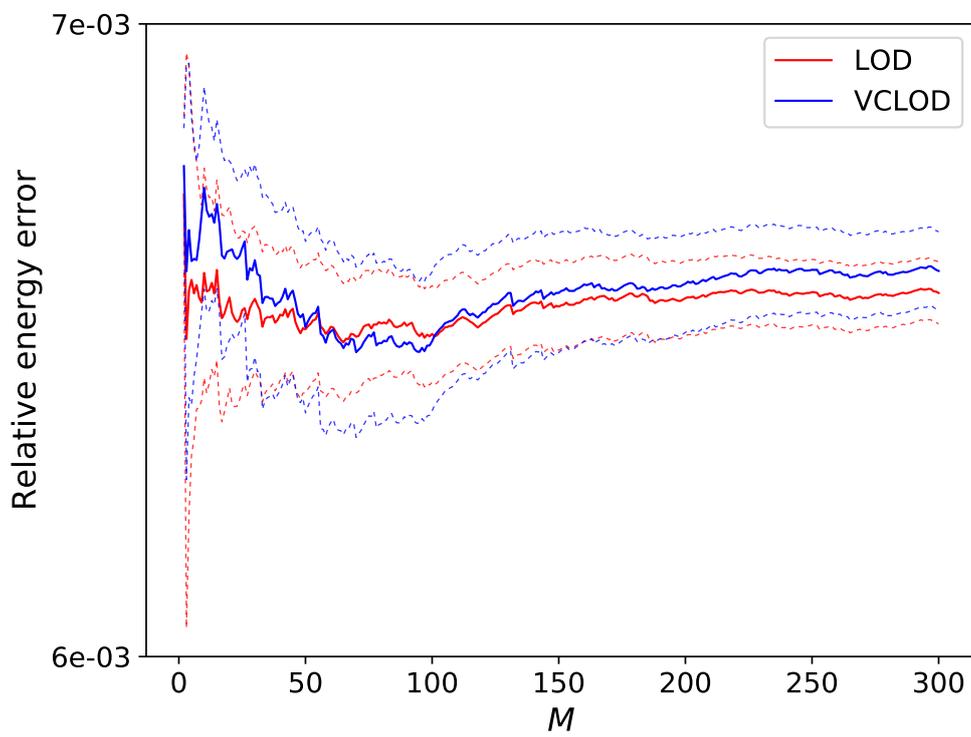


Figure 8.16: Comparison between Monte Carlo simulation for $e_{L^2}(u_h, u_k)$ and $e_{L^2}(u_h, u_{k,p}^{vc})$. The dashed lines represent the confidence interval.

9 Conclusion and future work

This paper was devoted to the Localized orthogonal decomposition method, introduced by Målqvist and Peterseim in [28], and its applications to multiscale diffusion problems. First, we presented the standard FEM and showed that multiscale problems require a mesh size H for the FEM that might reach the limits of today's computer technology. Due to that, we presented the Localized orthogonal decomposition method as an example for a multiscale method. The incorporation of the corrector function into each coarse FE-basis function resulted that the basis functions lose their local support. Thus, a localization was required in order to gain a feasible method. We presented the possibilities for the localization, introduced the right hand side correction and furthermore, we learned about the PG-LOD that aims to reduce the memory consumption of the method. To justify the localization for the LOD, it was crucial to show that the corrector functions decay exponentially outside of an area of their associated node. We presented the classical version and discussed analytical problems and space for improvements. The classical estimate

$$\| \|u - u_{H,lk}^{\text{lod}} \| \| \leq \left(C_4 \|H_T^{-1}\|_{L^\infty(\Omega)} (lk)^{d/2} (C_1/l)^{\frac{k-2}{2}} + HC_2 \right) \|f\|_{L^2(\Omega)},$$

with the exact solution u and its approximation $u_{H,lk}^{\text{lod}}$, showed that the method yields a good accuracy for sufficiently large constants l and k , which are responsible for the patch size of the localization. After presenting this result, we changed the localization as well as the interpolation and we utilized a new proof strategy to proof a similar, but essentially better result

$$\| \|u - u_{H,k}^{\text{LOD}} \| \| \leq \left(Ck^{d/2}\theta^k + HC' \right) \|f\|_{L^2(\Omega)},$$

with a different approximation $u_{H,k}^{\text{LOD}}$. Compared to the former result, the factor $\|H_T^{-1}\|_{L^\infty(\Omega)}$ does not appear anymore and therefore, the method is already applicable for a lower localization constant k . We presented the proof of the latter in detail and emphasized the usage of the changed interpolation and localization. As an application of the LOD, Chapter 5 dealt with variational crimes. We explained that interferences in the stiffness matrix and, more particularly, perturbations in the diffusion coefficient might result in a complete recomputation of the FEM, since the accuracy decreases significantly. Concerning the

FEM, we realized that there is no possibility to make use of the non-perturbed problem in terms of computational savings. Based on the approaches for time-dependent problems by Målqvist and Hellman in [15], we presented a novel method that enables the possibility to actually pay with accuracy in order to gain a less expensive computation. Since this method is based on the PG-LOD with right hand side correction, the error analysis in Chapter 6 required inf-sup conditions for the stability of the PG-LOD. In addition, these results provided the main tools for the error bound

$$\|u - \tilde{u}_k\| \leq ck^{d/2}(\theta^k + \text{TOL})\|f\|_{L^2(\Omega)}$$

for the approximation of the non-perturbed \tilde{u}_k . We pointed out that we recompute only for a certain amount of correctors in the PG-LOD and use old correctors, if they are not, or just slightly, affected. For this purpose, we presented an error indicator e_u that enables the decision on recomputing by setting an upper bound TOL. In Chapter 8, we tested the PG-LOD on multiscale instances and we presented several numerical examples of diffusion coefficients that are potentially subjected to perturbations. The presented perturbations are change in values, shift and disappearance. Related to this, we pointed out the error indicator e_u in order to describe the novel method. In the numerical simulations, we assessed the novel method in terms of the energy error compared to the worst and the best PG-LOD and we compared the results with the indicator e_u . We concluded that the novel method performs well and is actually applicable for every test case we proposed. Lastly, we explained the application of our novel method to the weakly random problems by Le Bris and Legoll in [25]. We proposed to use the novel PG-LOD method for every step in the Monte Carlo method in order to approximate the confidence interval and we concluded well performance.

The instances for diffusion coefficients throughout this thesis were motivated by composite materials in today's industries. We restricted our approaches to two values, which enabled a better comparison. The results can be assigned to more general cases. However, the diffusion properties of composite materials are restricted to the number of their different components. In terms of perturbations, we think about machine failures which are indeed mostly subjected to displacements and defects we addressed in this thesis. With the confidence interval that follows from the Monte Carlo method, we can identify whether we are still able to use the material, although the machine might produce some failure. In total, we actually achieved a strong method throughout this thesis, in case the failure is sufficiently small and the probability appropriately low.

Many different experiments and analytical enhancements have been left for the future.

From an analytic point of view, high contrast problems are of interest regarding to this thesis. Moreover, the proof for the localization in Chapter 4 is still very complex. Recently, there has been another approach made by Kornhuber, Yserentant and Peterseim in [23], in order to gain a similar result that works on a more abstract level, but enables an elegant proof. It is interesting to compare both strategies with the knowledge of Chapter 4. In terms of the error analysis for the novel method in Chapter 6, no estimate for the expectation value in Section 8.5 has been derived, which is also a task for the future. Related to this, more Monte Carlo approaches and stochastic simulations have to be compared and assessed. Moreover, an interesting goal for the novel method is to find an a priori percentage for recomputing.

List of Figures

2.1	Periodic variations of $A_\varepsilon(x)$ on $\Omega = (0, 1)$ for $\varepsilon = 2^{-5}$ and $\varepsilon = 2^{-6}$	10
2.2	FEM approximation for various choices of H	11
2.3	Energy error $\ u_\varepsilon - u_H\ $ for various choices of H	12
2.4	$A(x)$ and $u(x)$. Black is 1 and white is 0.01.	12
2.5	FEM approximation for various choices of H	13
2.6	Energy error $\ u - u_H\ $ for various choices of H	13
3.1	Multiscale splitting.	16
3.2	Patches around a certain node $x \in \mathcal{N}$ in white.	19
3.3	Basis function of $V_{H,k}^{\text{lod}}$ and its decomposition for an $x \in \mathcal{N}$ and $k = 4$	19
3.4	Patches for a certain coarse mesh element $T \in \mathcal{T}_H$ in the middle.	20
7.1	Refinement of a quadrilateral coarse mesh.	63
8.1	PG-LOD approximation of u_ε for various choices of H and $\varepsilon = 2^{-5}$	70
8.2	Energy error $\ u_\varepsilon - u_H\ $ for $\varepsilon = 2^{-5}$ and for various k	70
8.3	5% defects in a period structure with changing value, shift and disappearance. Blue is 0.05, red is 0.8 and yellow is 1. Motivated by [25].	71
8.4	Instances for diffusion coefficients.	72
8.5	Error indicator for the coarse elements.	73
8.6	Error indicator for particular elements.	74
8.7	Coefficient, defects, and the affected elements for several $k \in \mathbb{N}$. Blue means $e_{u,T} > 0$, white means $e_{u,T} = 0$	75
8.8	Perturbations for each coefficient. Left: change in value, middle: disappearance, right: shift.	78
8.9	Energy error comparison and indicator e_u for coefficient 1.	79
8.10	Energy error comparison and indicator e_u for coefficient 2.	80
8.11	Energy error comparison and indicator e_u for coefficient 3.	81
8.12	Energy error comparison and indicator e_u for coefficient 4.	82
8.13	Energy errors e_{vc} in comparison.	83
8.14	Indicator e_u in comparison.	84

8.15	Monte Carlo simulation for coefficient 3. The dashed lines represent the confidence interval.	89
8.16	Comparison between Monte Carlo simulation for $e_{L^2}(u_h, u_k)$ and $e_{L^2}(u_h, u_{k,p}^{vc})$. The dashed lines represent the confidence interval.	90

Bibliography

- [1] Assyr Abdulle and Patrick Henning. “Localized orthogonal decomposition method for the wave equation with a continuum of scales”. In: *Mathematics of Computation* 86.304 (2017), pp. 549–587 (cit. on p. 3).
- [2] Hans Wilhelm Alt. *Lineare Funktionalanalysis. Eine anwendungsorientierte Einführung*. 6th revised ed. Berlin: Springer, 2012. ISBN: 978-3-642-22260-3 (cit. on p. 5).
- [3] Ivo Babuška. “The finite element method with Lagrangian multipliers”. In: *Numerische Mathematik* 20.3 (1973), pp. 179–192 (cit. on p. 22).
- [4] Ivo Babuška, Gabriel Caloz, and John E Osborn. “Special finite element methods for a class of second order elliptic problems with rough coefficients”. In: *SIAM Journal on Numerical Analysis* 31.4 (1994), pp. 945–981 (cit. on p. 2).
- [5] Ivo Babuška and John E Osborn. “Generalized finite element methods: their performance and their relation to mixed methods”. In: *SIAM Journal on Numerical Analysis* 20.3 (1983), pp. 510–536 (cit. on p. 2).
- [6] Klaus-Jürgen Bathe. “The inf–sup condition and its evaluation for mixed finite element methods”. In: *Computers & structures* 79.2 (2001), pp. 243–252 (cit. on p. 22).
- [7] Dietrich Braess. *Finite elements*. Third. Theory, fast solvers, and applications in elasticity theory, Translated from the German by Larry L. Schumaker. Cambridge University Press, Cambridge, 2007. ISBN: 978-0-521-70518-9 (cit. on pp. 2, 5, 6, 8, 9, 42, 44).
- [8] Franco Brezzi. “On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers”. In: *Revue française d’automatique, informatique, recherche opérationnelle. Analyse numérique* 8.2 (1974), pp. 129–151 (cit. on p. 22).
- [9] Flake C Campbell. *Structural composite materials*. ASM international, 2010 (cit. on p. 1).
- [10] C. Carstensen and R. Verfürth. “Edge residuals dominate a posteriori error estimates for low order finite element methods”. In: *SIAM J. Numer. Anal.* 36.5 (1999), pp. 1571–1587 (cit. on pp. 25, 26).
- [11] Weinan E, Bjorn Engquist, and Zhongyi Huang. “Heterogeneous multiscale method: A general methodology for multiscale modeling”. In: *Phys. Rev. B* 67 (9 Mar. 2003), p. 092101. DOI: 10.1103/PhysRevB.67.092101. URL: <https://link.aps.org/doi/10.1103/PhysRevB.67.092101> (cit. on p. 2).
- [12] Daniel Elfverson, Victor Ginting, and Patrick Henning. “On multiscale methods in Petrov–Galerkin formulation”. In: *Numerische Mathematik* 131.4 (2015), pp. 643–682 (cit. on pp. 3, 21).

- [13] C. Engwer et al. “Efficient implementation of the Localized Orthogonal Decomposition method”. In: *ArXiv e-prints* (Feb. 2016). arXiv: 1602.01658 [math.NA] (cit. on pp. 3, 63, 68, 69).
- [14] Victor Ginting. “Analysis of two-scale finite volume element method for elliptic problem”. In: *Journal of Numerical Mathematics jnma* 12.2 (2004), pp. 119–141 (cit. on p. 2).
- [15] Fredrik Hellman and Axel Målqvist. “Numerical homogenization of time-dependent diffusion”. In: *arXiv preprint arXiv:1703.08857* (2017) (cit. on pp. 3, 45, 49, 68, 69, 92).
- [16] P. Henning and A. Målqvist. “Localized orthogonal decomposition techniques for boundary value problems”. In: *SIAM J. Sci. Comput.* 36.4 (2014), A1609–A1634. ISSN: 1064-8275. DOI: 10.1137/130933198 (cit. on pp. 3, 26, 27, 69).
- [17] P. Henning, A. Målqvist, and D. Peterseim. “Two-Level Discretization Techniques for Ground State Computations of Bose-Einstein Condensates”. In: *SIAM J. Numer. Anal.* 52.4 (2014), pp. 1525–1550. ISSN: 0036-1429. DOI: 10.1137/130921520 (cit. on p. 3).
- [18] P. Henning and D. Peterseim. “Oversampling for the Multiscale Finite Element Method.” In: *SIAM Multiscale Mod. Simul.* 11.4 (2013), pp. 1149–1175. DOI: 10.1137/120900332 (cit. on pp. 3, 69).
- [19] Patrick Henning and Anna Persson. “A multiscale method for linear elasticity reducing Poisson locking”. In: *Computer Methods in Applied Mechanics and Engineering* 310 (2016), pp. 156–171 (cit. on p. 3).
- [20] Patrick Henning and Anna Persson. “A multiscale method for linear elasticity reducing Poisson locking”. In: *Computer Methods in Applied Mechanics and Engineering* 310 (2016), pp. 156–171 (cit. on p. 26).
- [21] Thomas Y Hou and Xiao-Hui Wu. “A multiscale finite element method for elliptic problems in composite materials and porous media”. In: *Journal of computational physics* 134.1 (1997), pp. 169–189 (cit. on p. 2).
- [22] Thomas JR Hughes et al. “The variational multiscale method—a paradigm for computational mechanics”. In: *Computer methods in applied mechanics and engineering* 166.1-2 (1998), pp. 3–24 (cit. on p. 2).
- [23] R. Kornhuber, D. Peterseim, and H. Yserentant. “An Analysis of a Class of Variational Multiscale Methods based on Subspace Decomposition”. In: (Nov. 2016). Submitted for publication (cit. on pp. 3, 93).
- [24] Claude Le Bris. “Some numerical approaches for weakly random homogenization”. In: *Numerical mathematics and advanced applications 2009*. Springer, 2010, pp. 29–45 (cit. on p. 3).
- [25] Claude Le Bris, Frédéric Legoll, and Florian Thomines. “Multiscale Finite Element approach for “weakly” random problems and related issues”. In: *ESAIM: Mathematical Modelling and Numerical Analysis* 48.3 (2014), pp. 815–858 (cit. on pp. 3, 71, 86–88, 92).
- [26] Axel Målqvist and Fredrik Hellman. “Contrast independent localization of multiscale problems”. In: *ArXiv e-prints* (Oct. 2016). arXiv: 1610.07398 [math.NA] (cit. on pp. 3, 20, 22, 26, 27, 32).
- [27] Axel Målqvist and Daniel Peterseim. “Computation of eigenvalues by numerical upscaling”. In: *Numerische Mathematik* 130.2 (2015), pp. 337–361 (cit. on p. 3).

-
- [28] Axel Målqvist and Daniel Peterseim. “Localization of elliptic multiscale problems”. In: *Math. Comp.* 83.290 (2014), pp. 2583–2603. ISSN: 0025-5718. DOI: 10.1090/S0025-5718-2014-02868-8. URL: <http://dx.doi.org/10.1090/S0025-5718-2014-02868-8> (cit. on pp. 2, 15, 17, 18, 26, 27, 30, 31, 64, 69, 91).
- [29] M. Oehlberger and B. Verfürth. *Localized Orthogonal Decomposition for two-scale Helmholtz-type problems*. arXiv preprint 1605.03410. arXiv, May 2016. URL: /2016/OV16a (cit. on p. 3).
- [30] Daniel 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*. Ed. by G. R. Barrenechea et al. Vol. 114. Lecture Notes in Computational Science and Engineering. Also available as INS Preprint No. 1509. Springer, May 2016. arXiv: 1505.07611 [math.NA] (cit. on pp. 3, 10, 27, 69).
- [31] Daniel Peterseim and Robert Scheichl. “Robust Numerical Upscaling of Elliptic Multiscale Problems at High Contrast”. In: *Comput. Meth. in Appl. Math.* 16.4 (2016), pp. 579–603. DOI: 10.1515/cmam-2016-0022. URL: <http://dx.doi.org/10.1515/cmam-2016-0022> (cit. on pp. 25, 26).
- [32] Gilbert Strang and George J Fix. *An analysis of the finite element method*. Vol. 212. Prentice-hall Englewood Cliffs, NJ, 1973 (cit. on pp. 3, 42).