

Computational high frequency scattering from high contrast heterogeneous media^{*}

Daniel Peterseim[†] Barbara Verfürth[†]

Abstract. This article considers the computational (acoustic) wave propagation in strongly heterogeneous structures beyond the assumption of periodicity. A high contrast between the constituents of microstructured multiphase materials can lead to unusual wave scattering and absorption, which are interesting and relevant from a physical viewpoint, for instance, in the case of crystals with defects. We present a computational multiscale method in the spirit of the Localized Orthogonal Decomposition and provide its rigorous a priori error analysis for two-phase diffusion coefficients that vary between 1 and very small values. Special attention is paid to the extreme regimes of high frequency, high contrast, and their previously unexplored coexistence. A series of numerical experiments confirms the theoretical results and demonstrates the ability of the multiscale approach to efficiently capture relevant physical phenomena.

Key words. multiscale method; wave propagation; Helmholtz equation; high contrast; photonic crystal

AMS subject classifications. 35J05, 65N12, 65N15, 65N30

1 Introduction

Wave propagation in heterogeneous high-contrast materials has received a growing interest in recent years because the combination of microstructures and high contrast can lead to unusual material properties such as opening band gaps, artificial magnetism, or surface plasmons, see, for instance, [19, 34, 47, 52]. In periodic structures, homogenization techniques based on two-scale convergence [1, 36] or Bloch-wave techniques [2, 3, 15] are able to reduce the problem's complexity and to provide effective models for the overall macroscopic behavior of the wave. In practical applications, however, the periodic structure may easily be destroyed by local defects due to production errors. Perturbation theory for infinite periodic structures indicates that the combination of certain local defects and wave numbers may (suddenly) lead to localized waves, whereas in the perfectly periodic case the corresponding wave cannot propagate, see [4, 34]. While this can be highly favorable, local defects may also destroy the desired tuned (unusual) properties of the periodic structures. However, the above described analytical examination of the phenomena is quite limited to the periodic case with (mild) perturbations, where each case has to be considered separately. Moreover, it is not clear how these asymptotic results can predict the particular behavior for a given (finite) structure.

Therefore, numerical simulations of wave propagation and scattering in heterogeneous high-contrast materials are crucial to predict and also understand more situations and phenomena.

^{*}The authors would like to acknowledge the kind hospitality of the Erwin Schrödinger International Institute for Mathematics and Physics (ESI), where part of this research was developed under the frame of the Thematic Programme *Numerical Analysis of Complex PDE Models in the Sciences*.

[†]Institut für Mathematik, Universität Augsburg, Universitätsstr. 14, D-86159 Augsburg, {daniel.peterseim, barbara.verfuerth}@math.uni-augsburg.de

Yet, a direct discretization of realistic problems using standard (finite element) methods easily exceeds today’s computational resources because the mesh has to resolve all the material inhomogeneities. Additional challenges and restrictions on the mesh size are encountered in the high-frequency regime by the so-called pollution effect, see [5]. Therefore, multiscale methods need to be employed to reliably simulate wave propagation in high-contrast heterogeneous materials.

By now, the challenges for wave propagation problems on the one side and for high-contrast coefficients on the other side mostly have been studied separately. To reduce the pollution effect for the Helmholtz equation with constant or low-contrast coefficients, a number of approaches has been designed and analyzed, such as (hybridizable) discontinuous Galerkin methods [12, 27], the *hp*-version of the finite element method [21, 40, 41, 42], (plane wave) Trefftz methods [24, 33, 48], or the Localized Orthogonal Decomposition [10, 23, 50]. High-contrast coefficients mainly have been studied for elliptic problems using multiscale finite element methods [13, 18, 45] and the Localized Orthogonal Decomposition [29, 51], just to name a few. Finally, we mention that high-contrast Helmholtz problems have been studied based upon homogenization results in [44]. We shall emphasize that high contrast in the diffusion coefficient can be triggered by the degeneration of the upper or the lower bound of the elliptic operator. In the context of time-harmonic wave propagation, the observed effects are quite different in the two regimes, see, for instance, [34]. Locally, the variation of the diffusion coefficient can be rephrased in terms of a variable wave number. Small diffusion coefficients correspond to large wave numbers which amplifies the aforementioned pollution effect and leads to much more critical conditions on the minimal numerical resolution. That is why we focus in this manuscript on the specific setting where the diffusion coefficient takes a small value ε^2 on one part of the domain and 1 on the complement.

The paper presents a computational multiscale method in the spirit of the Localized Orthogonal Decomposition (LOD) and provides its rigorous analysis for high-contrast Helmholtz problems beyond periodicity and scale separation. We build upon previous ideas for high-frequency wave propagation in [10, 23, 50] and high-contrast elliptic equations in [11, 29, 51]. The combination of these techniques and approaches yields a (quasi-)localized method which does not require structures in the coefficients such as periodicity. Furthermore, a scale separation between the wave length and the size of individual “valleys” (where the coefficient is very small) is not required. The multiscale method can be viewed as a generalized finite element method with special multiscale basis functions. As standard finite element functions on a coarse mesh alone do not yield a faithful approximation space, problem-dependent multiscale functions are added. The latter are defined as solutions of local fine-scale problems. Our numerical analysis shows that the (coarse) mesh, and, hence, the dimension of the space, is coupled linearly to the effective wave number. Moreover, the size of the fine-scale problems to solve scales only logarithmically with this effective wave number. This generalizes the results of [10, 23, 50] for the constant or low-contrast coefficient case. We carefully address the high contrast in our analysis which is crucial to obtain the mildest dependencies on the contrast (encoded in the quantity ε) as possible. We provide a priori error estimates in various norms, which can be motivated from the periodic (high contrast) case and are of the expected order. The error analysis contains the stability constant of the exact solution (see (2.5) and (2.6)). Unfortunately, only little is known in the general setting about bounds for this constant that are explicit in the wave number and the coefficients. Hence, the multiscale method is proved to be contrast- and wave number-robust only under the assumption that the solution depends moderately on these parameters.

The paper is organized as follows. In Section 2 we introduce our model problem and illustrate properties of the solution in the periodic case. We introduce the multiscale method with all necessary notation in Section 3 and analyze its well-posedness and a priori errors in Section 4. The main proofs are detailed in Section 5. Finally, numerical experiments in Section 6 confirm

our predicted theoretical results and show the applicability of our method for physically relevant cases concerning band gap simulations and localized defects.

2 Helmholtz problem with high contrast

In this section, we are concerned with the Helmholtz problem with high contrast in general. First, we formulate our model setting in Section 2.1 and also introduce the necessary notation. In Section 2.2, we then explain in detail the considered set-up by motivating it from homogenization theory and Bloch-wave analysis. The discussion moreover sheds a light on what approximation results we can expect for our multiscale method and compares it to low-contrast Helmholtz problems.

2.1 Model problem

Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral Lipschitz domain. We consider the following model problem: Find $u : \Omega \rightarrow \mathbb{C}$ such that

$$\begin{aligned} -\operatorname{div}(A\nabla u) - k^2 u &= f && \text{in } \Omega, \\ \nabla u \cdot n - iku &= 0 && \text{on } \partial\Omega. \end{aligned}$$

Here, we assume $f \in L^2(\Omega)$ and $k \geq k_0 > 0$. Inhomogeneous (Robin) boundary conditions on $\partial\Omega$ can be treated following the theory of [31], which we omit for simplicity. Other types of boundary conditions [23, 50] and also perfectly matched layers [22] have been studied for homogeneous media and their incorporation in the present work is possible as well. The scalar diffusion coefficient A is piece-wise constant with respect to a quadrilateral background mesh \mathcal{T}_ε with mesh size $O(\varepsilon)$ and $0 < \varepsilon \ll 1$. On each quadrilateral, A takes either the value ε^2 or 1. We introduce the non-overlapping partition $\Omega = \Omega_1 \cup \Omega_\varepsilon$, where Ω_1 and Ω_ε are the parts of the domain where A takes the value 1 or ε^2 , respectively. We assume that $\Omega_\varepsilon \subset\subset \Omega$, so that A does not appear in the Robin boundary condition. This set-up requires some comments. Although we fix the two values of A , the present setting still allows much freedom since the geometry of Ω_ε is relatively flexible. In particular, no periodicity is assumed. We emphasize that the scaling of A is independent of the space dimension. The special choice of 1 and ε^2 induces a high contrast between the two materials, which is important to obtain unusual macroscopic effects in (periodic) homogenization theory, see Section 2.2 below for a detailed discussion. However, we also underline that the results of the present contribution can directly be transferred to other scalings of A , see Remark 4.7. We stick with the above choice of ε^2 in Ω_ε for simplicity and since this seems to be the physically (most) relevant set-up. As mentioned in the introduction, we stress that the present choice of high contrast deals with very small values of A in parts of the domain (in contrast to, for instance, [13]). This scaling in particular implies an almost degeneracy of the equation and is therefore harder than the case where A may vary between 1 and very large values, see Remark 4.7. Finally, note that the assumption of a piece-wise constant A can be relaxed if it is still possible to construct a suitable interpolation operator as described in Section 3.1 for that diffusion coefficient.

Throughout the whole article, we use standard notation on Sobolev spaces. All functions are complex-valued if not mentioned otherwise. The complex conjugate of v is denoted by \bar{v} . We write $a \lesssim b$ in short for $a \leq Cb$ with a constant C independent of k , ε , the mesh size H , and the oversampling parameter m . Similarly, $a \approx b$ stands for $a \leq Cb$ and $b \leq C'a$ with constants C, C' independent of k , ε , H , and m . We then consider the (weak) problem: Find $u \in H^1(\Omega)$

such that

$$\mathcal{B}(u, v) := \int_{\Omega} A \nabla u \cdot \nabla \bar{v} - k^2 u \bar{v} \, dx - ik \int_{\partial\Omega} u \bar{v} \, d\sigma = \int_{\Omega} f \bar{v} \, dx \quad \forall v \in H^1(\Omega). \quad (2.1)$$

Let $\|\cdot\|$ denote the standard $L^2(\Omega)$ -norm, we then define the following weighted (semi) norms:

$$\|u\|_{0,A} := \|A^{1/2}u\|, \quad (2.2)$$

$$\|u\|_{1,A} := \|A^{1/2}\nabla u\|, \quad (2.3)$$

$$\|u\|_{1,A,k} := (\|A^{1/2}\nabla u\|^2 + k^2\|u\|^2)^{1/2}. \quad (2.4)$$

If necessary, we will indicate by an additional subscript the domain where the norm is taken. Note that $\|\cdot\|_{1,A,k}$ represents the energy norm associated with the sesquilinear form \mathcal{B} . If we assume that $\text{dist}(\Omega_\varepsilon, \partial\Omega)$ is bounded from below by an ε -independent constant, we can deduce the following trace inequality (cf. [9])

$$k^{1/2}\|v\|_{\partial\Omega} \lesssim \|v\|_{1,A,k}.$$

Otherwise, if Ω_ε is arbitrarily closed to $\partial\Omega$, we would obtain an additional ε -weight in the trace inequality of the following form

$$\|v\|_{\partial\Omega} \lesssim \varepsilon^{-1}\|v\|_{1,A,k}.$$

Since the trace inequality is needed for the continuity of \mathcal{B} with respect to the energy norm, we assume that $\text{dist}(\Omega_\varepsilon, \partial\Omega)$ is bounded from below by an ε -independent constant. Then, the continuity constant of \mathcal{B} is ε -independent, which simplifies the analysis of the multiscale method below, cf. also Remark 4.6.

The unique continuation principle and Fredholm theory provide existence and uniqueness of a (weak) solution to (2.1) in the two-dimensional case or under certain geometrical assumptions on Ω_ε in general dimensions, see [26, 25] and the discussions therein. Note that if Ω_ε consists of a finite collection of bounded subdomains as we consider it in the numerical experiments, the unique continuation principle holds in general dimensions, see [25]. The solution depends continuously on the right-hand side in the sense of the stability estimate

$$\|u\|_{1,A,k} \lesssim C_{\text{stab}}(k, \varepsilon) \|f\|_{L^2(\Omega)}. \quad (2.5)$$

However, the dependencies of the constant $C_{\text{stab}}(k, \varepsilon)$ on ε and k are not known explicitly in general, even for homogeneous media; see [39] for a positive and [6] for a negative result. Recently, various works have studied stability for the heterogeneous Helmholtz equation, see [10, 25, 43, 53]. We emphasize that the (common) assumptions on the coefficients are not even close to the target setting of this paper, but the recent work [35] indicates that polynomial in k stability results hold for almost all k . The stability estimates readily implies an inf-sup condition of the form

$$\inf_{v \in H^1(\Omega)} \sup_{\psi \in H^1(\Omega)} \frac{\text{Re } \mathcal{B}(v, \psi)}{\|v\|_{1,A,k} \|\psi\|_{1,A,k}} \geq \gamma(\varepsilon, k) \quad (2.6)$$

with $\gamma(\varepsilon, k) \approx (1 + kC_{\text{stab}}(\varepsilon, k))^{-1}$, see, e.g., [39].

2.2 Photonic crystals and homogenization results

To motivate some of our results and connect and differentiate them to the existing literature, we consider the case of a bounded scatterer with periodic micro-structure. To be more precise, we consider $\Omega_\varepsilon = \bigcup_{j \in \mathcal{I} \subset \mathbb{Z}^2} \varepsilon(j + \Sigma)$, where Σ is an open Lipschitz subset of the unit square and \mathcal{I} is an index set such that Ω_ε lies completely in a compact subset of Ω . The last condition is necessary to have a uniformly bounded distance between Ω_ε and $\partial\Omega$.

Two-scale homogenization In the perfectly periodic case and under the assumption that ε is much smaller than the wave length $\lambda \sim k^{-1}$, homogenization using two-scale convergence [1, 36] gives information about the limit problem for fixed but arbitrary k and $\varepsilon \rightarrow 0$. The surprising result is that one does not only obtain an effective diffusion matrix but also an effective wave number of the form $k^2\mu$, where μ is defined by an additional problem on Σ . The sign of μ depends on the choice of k , which highly influences the behavior of the macroscopic solution: For positive μ , standard wave propagation is experienced whereas for negative μ evanescent waves form that decay exponentially inside the scatterer with the periodic microstructure. The above considerations hold for compactly embedded inclusions Σ , see [7], whereas a layered structure results in a degenerate diffusion matrix and induces so-called surface plasmons, see [8].

Moreover, due to the high contrast inclusions there is no weak L^2 -convergence of the exact solution u (for positive ε) to the macroscopic limit solution (named u_0), but instead an additional finescale-type contribution in the inclusion has to be added to u_0 , see [1]. In contrast to problems without high contrast as in [14], we therefore cannot hope to find a good macroscopic L^2 -approximation of u – this is only possible in Ω_1 . However, with an additional correction, we can expect to recover the exact solution in a good L^2 -sense, cf. [44] for corresponding numerical results. This explains parts of the a priori error estimates of our multiscale method in Section 4.2. Roughly speaking, we find a macroscopic approximation which is close to the exact solution in the A -weighted L^2 norm $\|\cdot\|_{0,A}$ (which, in particular, means a good approximation on Ω_1 .) Adding a suitable correction, we obtain an approximation which is good in the $\|\cdot\|_{1,A,k}$ sense. This in particular implies a good approximation in the standard L^2 norm.

Bloch wave homogenization and photonic crystals If the wave length $\lambda \sim k^{-1}$ is of the same order as the periodicity length ε and hence, also the diameter of one high-contrast inclusion $\varepsilon\Sigma$, other types of resonances occur which can no longer be predicted by two-scale homogenization theory, see [16]. Instead, Bloch wave homogenization methods are employed which are based on the Floquet-Bloch theorem, see [3, 15]. To get an idea of the (optical) properties of the finite material with periodic microstructure (also called a photonic crystal), one can consider its infinite counterpart on the whole space \mathbb{R}^d . If the wave number k is an eigenvalue of the elliptic operator $-\nabla \cdot (A\nabla \cdot)$ on the whole space, the corresponding waves can propagate inside the (infinite) crystal. Due to the periodicity, the spectrum can be computed as the union of all spectra on one periodicity cell with varying quasi-periodic boundary conditions, the so-called Bloch spectra, see [15]. Band gaps correspond to “forbidden” wave numbers inducing evanescent waves as described above. A high contrast between the material properties allows to open a considerably large gap in the spectrum. However, when the periodic structure is destroyed by (local) defects, perturbation theory indicates that eigenvalues inside the band gaps may appear. The corresponding eigenfunctions are often localized near the defect, see [34] for a general discussion and [4] for the specific example of a line defect. As mentioned, this considers the infinite crystal with the hope that a considerably large finite photonic crystal will have similar properties. The multiscale method presented in this contribution does not rely on two-scale homogenization results and therefore, can also cope with the regime of Bloch wave homogenization. Hence, we can directly study finite photonic crystals with possible local defects and their properties.

3 Multiscale method

In this section, we introduce the multiscale method based on the ideas of the Localized Orthogonal Decomposition (LOD) [38, 32]. First in Section 3.1, we need some more notation on meshes and in particular a suitable interpolation operator which pays attention to the high contrast. After

these preliminaries, we can define our method in detail in Section 3.2. We close with some remarks concerning implementation in Section 3.3.

3.1 Meshes, finite element spaces, and interpolation operators

We cover Ω with a regular mesh \mathcal{T}_H consisting either of simplices or of parallelograms/parallel-epipeds. The mesh is assumed to be shape regular in the sense that the aspect ratio of the elements of \mathcal{T}_H is bounded uniformly from below. We introduce the mesh size $H = \max_{T \in \mathcal{T}_H} \text{diam } T$. Although we will assume $H = O(\varepsilon)$ for most of this article, see (4.2), note that \mathcal{T}_H does not necessarily resolve the interfaces between Ω_1 and Ω_ε . Associated with an element $T \in \mathcal{T}_H$ we define its neighborhood as

$$N(T) = \bigcup_{K \in \mathcal{T}_H, T \cap K \neq \emptyset} K.$$

Thereby, for any $m \in \mathbb{N}_0$, the m -layer patches are defined inductively via $N^{m+1}(T) = N(N^m(T))$ with $N^0(T) := T$. The shape regularity implies that there is a bound $C_{\text{ol},m}$ (depending only on m) of the number of the elements in the m -layer patch, i.e.,

$$\max_{T \in \mathcal{T}_H} \text{card}\{K \in \mathcal{T}_H : K \subset N^m(T)\} \leq C_{\text{ol},m}. \quad (3.1)$$

Throughout this article, we assume that \mathcal{T}_H is quasi-uniform, which implies that $C_{\text{ol},m}$ grows at most polynomially with m . We discretize the space $H_0^1(\Omega)$ with the lowest order Lagrange elements over \mathcal{T}_H , and denote this space by V_H . This means that $V_H = H^1(\Omega) \cap \mathcal{S}^1(\mathcal{T}_H)$, where $\mathcal{S}^1(\mathcal{T}_H)$ denotes the space of globally continuous functions which are polynomials of partial degree ≤ 1 (for quadrilateral elements) or polynomials of total degree ≤ 1 (for simplicial elements). In the case of quadrilateral meshes, one can easily exploit a possible (periodic) structure or pattern in the coefficient A , see Section 3.3.

Let $I_H : H^1(\Omega) \rightarrow V_H$ denote a bounded local linear projection operator, i.e., $I_H \circ I_H = I_H$, with the following stability and approximation properties for all $v \in H^1(\Omega)$ and all $w \in \ker I_H$

$$|I_H v|_{1,A,T} \lesssim |v|_{1,A,N(T)}, \quad (3.2)$$

$$\|w\|_{0,A,T} \lesssim H|w|_{1,A,N(T)}, \quad (3.3)$$

$$\|I_H v\|_{1,A,k,T} \lesssim \|v\|_{1,A,k,N(T)}, \quad (3.4)$$

where we recall that the constant hidden in \lesssim is independent of H and ε . We emphasize that we assume stability and approximation properties of I_H in A -weighted norms, which is the crucial difference between the low-contrast and the high-contrast case. The stability in the energy norm (3.4) can be deduced from (3.2) and (3.3) at least in the following two cases: The mesh satisfies the resolution condition (4.2) (as assumed later on), which restricts the mesh size significantly and may be too pessimistic. The other possibility is that for each $T \in \mathcal{T}_H$, the quotient $\frac{|T|}{|T \cap \Omega_1|}$ is uniformly bounded from above, which mainly implies that H may not be too fine. Anyhow, the crucial assumptions are (3.2) and (3.3). They have been verified under certain geometric assumptions using special A -weighted interpolants in [29, 51]. A possible choice of [51] that we also use in our numerical implementation builds upon A -weighted local L^2 projections in the following sense. Let \mathcal{N}_{int} denote the interior nodes of \mathcal{T}_H and λ_z the hat function associated with the node z . We then define

$$I_H(v) = \sum_{z \in \mathcal{N}_{\text{int}}} P_{A,\omega_z}(v)(z) \lambda_z \quad (3.5)$$

with the local A -weighted L^2 -projection P_{A,ω_z} on the patch $\omega_z := \{K \in \mathcal{T}_H : z \in K\}$. More precisely, for any $v \in H^1(\omega_z)$, $P_{A,\omega_z}(v) \in V_H|_{\omega_z}$ is given by

$$\int_{\omega_z} AP_{A,\omega_z}(v) \psi_H dx = \int_{\omega_z} Av \psi_H dx \quad \text{for all } \psi_H \in V_H|_{\omega_z}.$$

Assumptions (3.2) and (3.3) hold for this interpolant if the coefficient A is quasi-monotone. For instance, in the two-dimensional case, this condition is satisfied if – roughly speaking – Ω_ε consists of small inclusions of diameter ε that do not touch each other. This setup is considered in our numerical experiments. For more details on quasi-monotonicity we refer to [51] and the original works in the context of domain decomposition [17, 46]. In [29], a Scott-Zhang-type construction is employed, where the geometrical assumptions on Ω_ε are not so easy to describe. We emphasize, however, that the periodic structure underlying our numerical experiments in Section 6 is mentioned as an example in [29] (with circle inclusion instead of square inclusions) and is also analyzed in [11]. For this choice, our numerical experiments indicate that for a high contrast, the above described A -weighted interpolation operator I_H (cf. [51]) is favorable over its unweighted variant

$$I_H^1(v) := \sum_{z \in \mathcal{N}_{\text{int}}} P_{1,\omega_z}(v)(z) \lambda_z, \quad (3.6)$$

which utilizes local projections w.r.t. the standard L^2 -norm. Note that this unweighted construction is used for low-contrast problems, e.g. in [49]. Moreover, we stress that especially the construction in [51] can be extended to coefficients which take more than two distinct values (see the discussion in [51]).

3.2 Definition of the method

We formulate a (generalized) Petrov-Galerkin method to discretize (2.1). The method uses special multiscale trial and test functions, see [50]. The idea is to interpret the kernel of the interpolation operator I_H as the space of functions with finescale heterogeneities and to incorporate parts of these information by special localized finescale problems.

Let $W(N^m(T)) := \{w \in \ker I_H, w = 0 \text{ in } \Omega \setminus N^m(T)\}$. Note that $\ker I_H$ is infinite-dimensional and hence $W(N^m(T))$ is non-trivial. We introduce the element-wise localized corrector approximation $\mathcal{Q}_{T,m} : V_H \rightarrow W(N^m(T))$ via

$$\mathcal{B}_{N^m(T)}(\mathcal{Q}_{T,m} v_H, w) = -\mathcal{B}_T(v_H, w) \quad \forall w \in W(N^m(T)). \quad (3.7)$$

Here, \mathcal{B}_ω denotes the restriction of \mathcal{B} to a subset $\omega \subset \Omega$. Note that (3.7) requires only the local computation of finescale problems if m is small. The parameter m is commonly called oversampling parameter and will be coupled to the mesh size H later on. Using the element-wise correctors $\mathcal{Q}_{T,m}$ we define the localized corrector approximation via

$$\mathcal{Q}_m = \sum_{T \in \mathcal{T}_H} \mathcal{Q}_{T,m}. \quad (3.8)$$

The dual localized corrector is $\mathcal{Q}_m^* v_H := \overline{\mathcal{Q}_m(v_H)}$.

The localized method now uses the pair $((\text{id} + \mathcal{Q}_m)V_H, (\text{id} + \mathcal{Q}_m^*)V_H)$ as the new trial and test spaces in the Petrov-Galerkin formulation: Find $u_{H,m} \in V_H$ as solution to

$$\mathcal{B}((\text{id} + \mathcal{Q}_m)u_{H,m}, (\text{id} + \mathcal{Q}_m^*)v_H) = (f, (\text{id} + \mathcal{Q}_m^*)v_H) \quad \forall v_H \in V_H. \quad (3.9)$$

Note that $u_{H,m}$ lies in the standard finite element space and, hence, is called a macroscopic approximation since it cannot contain finescale information of the exact solution u . This information is only recovered in

$$u_{\text{LOD},m} := (\text{id} + \mathcal{Q}_m)u_{H,m}. \quad (3.10)$$

The well-posedness of the scheme and the error between the exact and the numerical solution will be analyzed in the next sections.

3.3 Practical aspects

The method presented in the previous section is not yet ready to use since the corrector problems (3.7) are still infinite-dimensional. Moreover, the number of these problems to solve increases with smaller mesh size H so that one has to think about an efficient computation. Both issues are briefly addressed in this section. For general aspects of the implementation, we refer to [20].

Fully discrete version of the multiscale method To discretize the corrector problems (3.7), we introduce a finescale shape-regular quadrilateral or simplicial mesh \mathcal{T}_h of Ω , which resolves all features and discontinuities of A . Moreover, we assume that the mesh is fine enough that a direct finite element simulation on \mathcal{T}_h would yield a faithful reference solution u_h to u . We stress, however, that this reference solution is not needed and not computed in our method. Only the local corrector problems are solved on the finescale mesh using the space $W_h(N^m(T)) := \{w \in H^1(\Omega) \cap \mathcal{S}^1(\mathcal{T}_h), w = 0 \text{ in } \Omega \setminus N^m(T) \text{ and } I_H w = 0\}$. The assumption that \mathcal{T}_h is sufficiently fine (as described above) is required for two reasons: (i) It guarantees that the multiscale method (3.9) still possesses a unique solution because for a sufficiently fine mesh, \mathcal{B} fulfills an inf-sup condition over $H^1(\Omega) \cap \mathcal{S}^1(\mathcal{T}_h)$ (which is a central argument in the proof of Theorem 4.4 below). (ii) It implies that the error committed by this additional discretization stays negligible. Due to the latter fact, we only analyze (for simplicity) the semi-discrete method as formulated in Section 3.2 and refer to, e.g., [23] for details on the minor technical changes in the proofs for the fully discrete method.

Efficient computation of the correctors First, we note that all corrector problems are independent and hence, can be easily computed in parallel. What is even more important is that additional structure in the coarse mesh \mathcal{T}_H and the coefficient can be exploited to reduce the number of corrector problems. For instance, in the periodic setting the corrector problems are translation invariant if we use a quadratic mesh with a mesh size which is an integer multiple of the periodicity length. Apart from a few problems at the boundary, only one interior corrector problem needs to be solved, cf. [23] for the homogeneous Helmholtz equation. Similar arguments allow to reduce the number of corrector problems for a (periodic) photonic crystal with a few local defects. If one has already computed the correctors for the periodic photonic crystals and then introduces local defects (as perturbations of the periodic structure), one can employ local error indicators as in [30] to automatically detect locations where the correctors need to be recomputed. Numerical experiments in [28] have shown that thereby a large part of the old correctors can be reused for local defects, which also lowers the computational complexity. Here, however, we do not consider this approach in more detail since we assume the geometry of Ω_ε to be a priori given and not as a deviation from a previous, more structured, set-up.

Efficient computation of the LOD solution The efficient computation of the LOD solution as well as CPU timings have been addressed in full detail in [20] for diffusion and eigenvalue problems. The algorithms turn over to the Helmholtz case in a very natural way. Without any assumptions on the structure of the coefficient, we need to solve $O(H^{-d})$ cell problems of

type (3.8). If these are discretized on a mesh of size h (see above), the number of degrees of freedom for each problem is $O(m^d(H/h)^d)$. Given their coercivity and since H resolves both the wave number and the oscillations of A , they can be solved efficiently. As pointed out in the previous paragraph, the number of cell problems can be drastically reduced for structured coefficients. The linear system of the LOD associated with (3.9) is of small dimension, namely of $\dim(V_H)$. Note that numerical experiments show that $m = 2, 3$ is sufficient and therefore the sparsity pattern of the matrix is only slightly enlarged by a factor $O(m^d)$ in comparison to standard FEM. The conditioning is comparable to standard FEM on the coarse mesh \mathcal{T}_H . The most efficient iterative solution of such systems is an open question in numerical analysis and beyond the scope of this article. The present approach does neither simplify nor overcomplicate this issue of numerical linear algebra. However, the tailored discretization allows the reduction of the degrees of freedom and hence the size of the system to almost optimal in the sense of sampling theory, whereas standard FEM would require much higher costs even in the regime of constant coefficients.

4 Error analysis

In this section, we analyze the multiscale method of Section 3.2 for the high-contrast Helmholtz problem. We first show in Section 4.1 that under a reasonable resolution condition $kH\varepsilon^{-1} \lesssim 1$, the corrector problems are coercive and therefore well-posed. Then, we give the main results of this paper in Section 4.2, namely the inf-sup condition of the method as well as several a priori error estimates. These results are derived under the above resolution condition and the oversampling condition $m \gtrsim |\log(kH\varepsilon^{-1})|$ if the original model problem satisfies a polynomial k - and ε -dependence of the stability constant $C_{\text{stab}}(k, \varepsilon)$. The resolution condition $kH\varepsilon^{-1} \lesssim 1$ should be compared to the usual resolution condition encountered for the quasi-optimality in standard finite element methods, which presumably is $k^2h\varepsilon^{-2} \lesssim 1$ at best in non-smooth scenarios. Note that the condition for existence of a FEM solution can be relaxed to $h(k\varepsilon^{-1})^\alpha \lesssim 1$, but one will always have $\alpha > 1$ for standard FEM.

4.1 The corrector problems

We first consider an idealized variant of (3.9) with “ $m = \infty$ ”, i.e., with $N^m(T) = \Omega$ in the corrector problems (3.7). For notational convenience, we denote by W the kernel of the interpolation operator I_H . Define the ideal correction operator $\mathcal{Q} : H^1(\Omega) \rightarrow W$ via

$$\mathcal{B}(\mathcal{Q}v, w) = -\mathcal{B}(v, w) \quad \forall w \in W. \quad (4.1)$$

Note that $\mathcal{Q} = \sum_{T \in \mathcal{T}_H} \mathcal{Q}_T$, where \mathcal{Q}_T solves (4.1) with right-hand side \mathcal{B}_T . Since the sesquilinear form \mathcal{B} is indefinite, we first have to show the well-posedness of these corrector problems.

Proposition 4.1. *If the mesh size H , the wave number k , and diffusion parameter ε satisfy the condition*

$$kH\varepsilon^{-1} \lesssim 1, \quad (4.2)$$

we have the following equivalence of norms on W

$$|w|_{1,A} \leq \|w\|_{1,A,k} \lesssim |w|_{1,A}.$$

Moreover, the sesquilinear form \mathcal{B} is elliptic over W , i.e., there exists $\alpha > 0$ (independent of H , ε , and k) such that

$$\operatorname{Re} \mathcal{B}(w, w) \geq \alpha |w|_{1,A}^2 \gtrsim \alpha \|w\|_{1,A,k}^2 \quad \forall w \in W.$$

Proof. Similar to [50], the crucial observation is that (3.3) implies

$$k\|w\| \lesssim k\varepsilon^{-1}\|w\|_{0,A} \lesssim kH\varepsilon^{-1}\|A^{1/2}\nabla w\| \quad (4.3)$$

for all $w \in W$. Thus, the energy norm $\|\cdot\|_{1,A,k}$ is equivalent to the weighted H^1 semi norm $|\cdot|_{1,A}$ if (4.2) is satisfied. Moreover, the L^2 part of the sesquilinear form \mathcal{B} can be absorbed in the gradient part if (4.2) is satisfied, leading to the ellipticity. \square

The ellipticity of \mathcal{B} over W implies the unique solvability of (4.1) due to the theory Lax-Milgram-Babuška. The same argument can be applied to obtain the well-posedness of the localized corrector problems (3.8).

Remark 4.2. The resolution condition (4.2) is a natural consequence and generalization of the resolution condition $kH \lesssim 1$ for the homogeneous or low-contrast Helmholtz equation, cf. [23, 50]. In fact, on Ω_ε , $k\varepsilon^{-1}$ is the effective wave number which needs to be resolved. There are certainly geometries of Ω_ε where this condition is necessary and sharp, but it is in general not clear how “likely” such situations are. As discussed, at least conditions like $\varepsilon^{-2}k^2h \lesssim 1$ and $h \ll \varepsilon$ are expected for general non-smooth coefficients for standard finite element methods where possible advantages of high-order methods [41, 42] cannot easily be exploited. (An exact statement or estimate of this condition is not available for this setting as the dependence of the stability constant on ε and k is not known.)

The error between the idealized corrector and its localized (or truncated) approximation decays exponentially in the number of layers m (also called oversampling parameter) in the following way.

Theorem 4.3. *Let \mathcal{Q} be defined by (4.1) and \mathcal{Q}_m be defined by (3.8). Let (4.2) be satisfied. Then there exists a constant $0 < \beta < 1$, independent of H , m , k , and ε , such that for all $v_H \in V_H$ it holds that*

$$|(\mathcal{Q} - \mathcal{Q}_m)v_H|_{1,A} \lesssim C_{\text{ol},m}^{1/2}\beta^m|v_H|_{1,A}. \quad (4.4)$$

The proof is given in Section 5.1. Employing that both correctors map into the kernel space and using (3.3), we deduce furthermore that

$$\|(\mathcal{Q} - \mathcal{Q}_m)v_H\|_{0,A} \lesssim C_{\text{ol},m}^{1/2}H\beta^m\|v_H\|_{1,A}.$$

We emphasize that neither the constant hidden in \lesssim nor β depend on H , m , k , or ε . In particular the independence of β from ε is the so-called contrast-independent localization. It can only be obtained with the A -weighted interpolation operator and is the major difference of this work from the previous ones on the low-contrast Helmholtz equation [10, 23, 50].

4.2 Stability of the method and a priori error estimates

The previous section showed that the corrector problems have a unique (and stable) solution. Before analyzing the error of the multiscale method (3.9), we need to show the stability (and well-posedness) of the method by proving an inf-sup condition. Since the arguments utilize the corrector error estimate (4.4), we first examine what happens if we replace \mathcal{Q}_m and \mathcal{Q}_m^* in (3.9) by their idealized counterparts \mathcal{Q} from (4.1) and its adjoint \mathcal{Q}^* . Hence, $\mathcal{Q}^* : H^1(\Omega) \rightarrow W$ is defined via

$$\mathcal{B}(w, \mathcal{Q}^*v) = -\mathcal{B}(w, v) \quad \forall w \in W$$

and we note that $\mathcal{Q}^*v := \overline{\mathcal{Q}(\bar{v})}$. By the definition of \mathcal{Q}^* and W , it holds that W and $(\text{id} + \mathcal{Q}^*)V_H$ are orthogonal with respect to \mathcal{B} . More precisely, we have $\mathcal{B}(w, (\text{id} + \mathcal{Q}^*)v_H) = 0$ for all $w \in W$

and all $v_H \in V_H$. Note that the same holds true for $(\text{id} + \mathcal{Q})V_H$ and W . Next we observe that the interpolation $I_H u$ of the exact solution satisfies

$$\begin{aligned} \mathcal{B}(I_H u, (\text{id} + \mathcal{Q}^*)v_H) &= \mathcal{B}(u, (\text{id} + \mathcal{Q}^*)v_H) + \underbrace{\mathcal{B}(I_H u - u, (\text{id} + \mathcal{Q}^*)v_H)}_{\in W} \\ &= (f, (\text{id} + \mathcal{Q}^*)v_H). \end{aligned}$$

Using once more the \mathcal{B} -orthogonality of W and $(\text{id} + \mathcal{Q}^*)V_H$, this implies that $u_{\text{LOD}} := (\text{id} + \mathcal{Q})I_H u$ solves

$$\mathcal{B}(u_{\text{LOD}}, (\text{id} + \mathcal{Q}^*)v_H) = (f, (\text{id} + \mathcal{Q}^*)v_H) \quad \forall v_H \in V_H.$$

The well-posedness and the a priori estimates now follow from the properties of the above idealized method and the fact that the ideal and the localized correctors are exponentially close.

Theorem 4.4. *Under the resolution condition (4.2) and the oversampling condition*

$$m \gtrsim |\log(\gamma^{-1}(\varepsilon, k)C_{\text{ol},m}^{1/2})|/|\log(\beta)| \quad (4.5)$$

\mathcal{B} satisfies the following inf-sup condition: there exists $\gamma_{\text{LOD}} \approx \gamma(\varepsilon, k) > 0$ such that

$$\inf_{v_H \in V_H} \sup_{\psi_H \in V_H} \frac{\text{Re} \mathcal{B}((\text{id} + \mathcal{Q}_m)v_H, (\text{id} + \mathcal{Q}_m^*)\psi_H)}{\|v_H\|_{1,A,k} \|\psi_H\|_{1,A,k}} \geq \gamma_{\text{LOD}}. \quad (4.6)$$

The proof is similar to [50] and needs the contrast-independent exponential decay of the corrector error, for details see Section 5.2. The oversampling condition depends on the stability constant for the original problem so that the dependence of m on ε and k is not analytically known. If we assume $\gamma(\varepsilon, k) \approx k^{-q}\varepsilon^p$ for non-negative numbers p and q independent of k and ε , the oversampling condition roughly reads $m \gtrsim |\log(k\varepsilon^{-1})|$ (with a multiplicative factor pq hidden in the \lesssim notation). This logarithmic dependence on the effective wave number $k\varepsilon^{-1}$ in Ω_ε is in agreement with the results for the homogeneous or low-contrast Helmholtz equation in [10, 23, 50]. Numerical experiments moreover indicate that rather small numbers $m = 2, 3$ are sufficient.

Theorem 4.5. *Let u be the solution to (2.1) and let $u_{H,m}$ be the solution to (3.9) and set $u_{\text{LOD},m} := (\text{id} + \mathcal{Q}_m)u_{H,m}$. Assume that (4.2) and the oversampling condition (4.5) are satisfied. Then it holds that*

$$\|u - u_{\text{LOD},m}\|_{1,A,k} \lesssim H\|A^{-1/2}f\| + C_{\text{ol},m}^{1/2}\beta^m\gamma^{-1}(\varepsilon, k)\|f\|, \quad (4.7)$$

$$\|u - u_{\text{LOD},m}\|_{0,A} \lesssim (H + C_{\text{ol},m}^{1/2}\beta^m\varepsilon\gamma^{-1}(\varepsilon, k))\|u - u_{\text{LOD},m}\|_{1,A,k} \quad (4.8)$$

$$\begin{aligned} \|u - u_{H,m}\|_{0,A} &\lesssim H \inf_{v_H \in V_H} \|u - v_H\|_{1,A} \\ &\quad + C_{\text{ol},m}^{1/2}\beta^m\gamma_{\text{LOD}}^{-1}(H\|A^{-1/2}f\| + C_{\text{ol},m}^{1/2}\beta^m\gamma^{-1}(\varepsilon, k)\|f\|). \end{aligned} \quad (4.9)$$

The proof is detailed in Section 5.2.

If m does not only fulfill the oversampling condition (4.5), but is also coupled to H like $m \approx |\log(\gamma^{-1}(\varepsilon, k)C_{\text{ol},m}^{1/2}H)|$, Theorem 4.5 simplifies to the following convergence orders: (i) The error between the exact and the full multiscale solution converges linearly in the energy norm and quadratically in an A -weighted L^2 norm, cf. (4.7) and (4.8). (ii) The error between the exact solution and the FE part of the multiscale solution converges (at least) linearly in the A -weighted L^2 norm, cf. (4.9). For the latter, note that $\inf_{v_H \in V_H} \|u - v_H\|_{1,A} \lesssim \|u\|_{1,A} \lesssim C_{\text{stab}}(\varepsilon, k)\|f\|$ independent of the regularity of u . If u has more than $H^1(\Omega)$ regularity, the rate in (4.9) may

improve. We stress that the remaining terms on the right-hand side of (4.9) are of order H^2 if the coupling $m \approx |\log(\gamma^{-1}(\varepsilon, k)C_{\text{ol},m}^{1/2}H)|$ is satisfied.

These estimates generalize the (expected) approximation results formulated in Section 2.2. We can only find a macroscopic approximation of the exact solution in Ω_1 , which is reflected in (4.9) by the A -weighting in the L^2 norm. For a good approximation in the energy norm, which also includes the standard L^2 norm, we need an additional (finescale) corrector, which is also present in (4.7). In the main error estimate (4.7), we furthermore emphasize that the error from the volume term is weighted by $A^{-1/2}$, which induces a factor of ε^{-1} on the right-hand side of the estimate if the support of f intersects with Ω_ε .

Remark 4.6. Our main results of this paper rely on the ε -independent trace inequality mentioned in Section 2.1, which needs that Ω_ε is bounded uniformly away from $\partial\Omega$. For settings where this is not the case, one can carefully trace the occurrence of extra ε powers because of the continuity of \mathcal{B} and we find that this mainly changes the oversampling condition to $m \gtrsim |\log(\gamma^{-1}(\varepsilon, k)\varepsilon C_{\text{ol},m}^{1/2})|/|\log(\beta)|$. This, however, is not critical since we expect a dependence of m on ε anyway because of $\gamma(\varepsilon, k)$.

Remark 4.7. The proofs of Theorems 4.4 and 4.5 reveal that the ε dependence in the resolution and the oversampling condition come from the ε dependence of the lower bound on A (and not the minimal diameter ε of an individual “valley” of Ω_ε). In fact, one can directly generalize the results to the case where A jumps between the values 1 and $\alpha^2 \ll 1$ by replacing all occurrences of ε by α . In particular, this means that our results remain valid for other scalings of A beyond the (physically interesting) one of, for instance, [7].

In the “complementary” high contrast setting, where A jumps between values 1 and $\alpha \gg 1$, one also needs A -weighted interpolation operators to get a contrast-independent decay of the correctors and thereby contrast-independent error estimates for the LOD. Since in this case it trivially holds $\|v\| \lesssim \|v\|_{0,A}$ for any $v \in H^1(\Omega)$, the resolution condition (4.2) reduces to the usual $kH \lesssim 1$ (independent of α), cf. for instance the proof of Proposition 4.1. For the same reason, one can replace $\|A^{-1/2}f\|$ by $\|f\|$ in (4.7) and (4.9). These improvements of the results reflect and underline the fact that this scaling of A is “easier” as it does not lead to high (effective) wave numbers as discussed in the introduction.

5 Proofs of the main results

5.1 Proof of the corrector error

The goal of this section is to prove (4.4). The main ingredient is the following exponential decay result. We emphasize once more that the following decay is contrast-independent, i.e., the rate $\tilde{\beta}$ and also the multiplicative constant hidden in the notation \lesssim both do *not* depend on ε . This is achieved by using an A -weighted interpolation operator and distinguishes the present analysis from previous works on the Helmholtz equation [10, 23, 50].

Proposition 5.1. *Let $\mathcal{Q} = \sum_{T \in \mathcal{T}_H} \mathcal{Q}_T$ be defined by (4.1) and $v_H \in V_H$. There exists a constant $0 < \tilde{\beta} < 1$, independent of H , m , and ε such that*

$$|\mathcal{Q}_T v_H|_{1,A,\Omega \setminus N^m(T)} \lesssim \tilde{\beta}^m |v_H|_{1,A,T}. \quad (5.1)$$

Proof. We define the cutoff function $\eta \in V_H$ via

$$\eta = 0 \quad \text{in} \quad N^{m-3}(T), \quad \eta = 1 \quad \text{in} \quad \Omega \setminus N^{m-2}(T)$$

and set $\mathcal{R} = \text{supp}(\nabla\eta)$. Let $v_H \in V_H$ and denote $\phi := \mathcal{Q}_T v_H$. Elementary estimates lead to

$$\begin{aligned} |\phi|_{1,A,\Omega \setminus N^m(T)}^2 &\lesssim |\text{Re}(A\nabla\phi, \eta\nabla\phi)_\Omega| \leq |\text{Re}(A\nabla\phi, \nabla(\eta\phi))_\Omega| + |\text{Re}(A\nabla\phi, (\nabla\eta)\phi)_\Omega| \\ &\leq M_1 + M_2 + M_3 \end{aligned}$$

with

$$\begin{aligned} M_1 &:= |\text{Re}(A\nabla\phi, \nabla(\text{id} - I_H)(\eta\phi))_\Omega|, \\ M_2 &:= |\text{Re}(A\nabla\phi, \nabla I_H(\eta\phi))_\Omega|, \\ M_3 &:= |\text{Re}(A\nabla\phi, (\nabla\eta)\phi)_\Omega|. \end{aligned}$$

Since $w := (\text{id} - I_H)(\eta\phi) \in W$, the idealized corrector problem (4.1) and the fact that w has only support outside T imply that $\mathcal{B}(\phi, w) = \mathcal{B}_T(v_H, w) = 0$. Therefore, we obtain

$$M_1 = |\text{Re}(A\nabla\phi, \nabla w)| \leq \left| \text{Re}(\mathcal{B}(\phi, w) + k^2(\phi, w)) \right| = |k^2 \text{Re}(\phi, w)|.$$

Hence, the stability and approximation estimates (3.2) and (3.3) for I_H , the properties of η , and the resolution condition (4.2) give

$$\begin{aligned} M_1 &\lesssim \varepsilon^{-2} k^2 H^2 (|\phi|_{1,A,\Omega \setminus N^m(T)}^2 + \|\nabla\eta\|_{N(\mathcal{R})} \|\phi\|_{0,A,N(\mathcal{R})} |\phi|_{1,A,N^2(\mathcal{R})}) \\ &\lesssim \varepsilon^{-2} k^2 H^2 (|\phi|_{1,A,\Omega \setminus N^m(T)}^2 + |\phi|_{1,A,N^2(\mathcal{R})}^2) \\ &\lesssim \frac{1}{2} (|\phi|_{1,A,\Omega \setminus N^m(T)}^2 + |\phi|_{1,A,N^2(\mathcal{R})}^2), \end{aligned}$$

where the first term can be hidden on the left-hand side.

Because of $\text{supp}(I_H(\eta\phi)) \subset N(\mathcal{R})$, the properties of I_H and the above estimate for $|\eta\phi|_{1,A}$ lead to

$$M_2 \lesssim |\phi|_{1,A,N(\mathcal{R})} |\eta\phi|_{1,A,N(\mathcal{R})} \lesssim |\phi|_{1,A,N^2(\mathcal{R})}^2.$$

Finally, the properties of η and the approximation result (3.3) for I_H show that

$$M_3 \lesssim |\phi|_{1,A,\mathcal{R}} |\phi|_{1,A,N(\mathcal{R})}.$$

All in all, this gives

$$|\phi|_{1,A,\Omega \setminus N^m(T)}^2 \leq C |\phi|_{1,A,N^2(\mathcal{R})}^2$$

with a constant C independent of H , m , k , and ε . We recall that $N^2(\mathcal{R}) = N^m(T) \setminus N^{m-5}(T)$. Because of

$$|\phi|_{1,A,\Omega \setminus N^m(T)}^2 + |\phi|_{1,A,N^m(T) \setminus N^{m-5}(T)}^2 = |\phi|_{1,A,N^{m-5}(T)}^2$$

we obtain

$$|\phi|_{1,A,\Omega \setminus N^m(T)}^2 \leq (1 + 2C^{-1})^{-1} |\phi|_{1,A,\Omega \setminus N^{m-5}(T)}^2.$$

Note that $(1 + 2C^{-1})^{-1} < 1$, so that a repeated application of the above argument and algebraic manipulations finish the proof. Note in particular that $\tilde{\beta}$ is independent of H , m , k , and ε because C is independent of these quantities. \square

The idea for the proof of Theorem 4.3 is now that the ideal corrector is only truncated in $\Omega \setminus N^m(T)$, where we know by the previous Proposition that the contribution is exponentially small.

Proof of Theorem 4.3. We define the cutoff function $\eta \in \mathcal{S}^1(\mathcal{T}_H)$ via

$$\eta = 1 \quad \text{in } N^{m-2}(T), \quad \eta = 0 \quad \text{in } \Omega \setminus N^{m-1}(T).$$

With the ellipticity and continuity of \mathcal{B} over W , we deduce Céa's Lemma

$$|(\mathcal{Q}_{T,m} - \mathcal{Q}_T)v_H|_{1,A} \lesssim \inf_{w_{T,m} \in W(N^m(T))} |\mathcal{Q}_T v_H - w_{T,m}|_{1,A}.$$

Inserting $w_{T,m} = (\text{id} - I_H)(\eta \mathcal{Q}_T v_H)$ and using stability and approximation properties of I_H yields – similar to the proof of Proposition 5.1 –

$$|(\mathcal{Q}_{T,m} - \mathcal{Q}_T)v_H|_{1,A} \lesssim |\mathcal{Q}_T v_H|_{1,A,\Omega \setminus N^m(T)}. \quad (5.2)$$

Applying (5.1) results in an exponential decay result for the error between the element correctors \mathcal{Q}_T and $\mathcal{Q}_{T,m}$.

To obtain the global estimate, we set $z := (\mathcal{Q} - \mathcal{Q}_m)v_H$ and $z_T := (\mathcal{Q}_T - \mathcal{Q}_{T,m})v_H$ and define the cutoff function $\eta \in \mathcal{S}^1(\mathcal{T}_H)$ via

$$\eta = 0 \quad \text{in } N^{m+1}(T), \quad \eta = 1 \quad \text{in } \Omega \setminus N^{m+2}(T).$$

The ellipticity of \mathcal{B} yields

$$|z|_{1,A}^2 \lesssim \sum_{T \in \mathcal{T}_H} |\mathcal{B}(z_T, z)| \leq \sum_{T \in \mathcal{T}_H} |\mathcal{B}(z_T, (1 - \eta)z)| + |\mathcal{B}(z_T, (\text{id} - I_H)(\eta z))| + |\mathcal{B}(z_T, I_H(\eta z))|.$$

The second term vanishes because $(\text{id} - I_H)(\eta z) \in W$ with support outside $N^m(T)$. The function $(1 - \eta)z$ vanishes on $S := \{\eta = 1\}$ and we obtain with the scaling properties of η , (3.3), (3.2), and the resolution condition (4.2) that

$$\begin{aligned} |\mathcal{B}(z_T, (1 - \eta)z)| &\lesssim \|(1 - \eta)z\|_{1,k,A,\Omega \setminus S} |z_T|_{1,A} \\ &\lesssim (k \|z\|_{\Omega \setminus S} + |z|_{1,A,\Omega \setminus S} + H^{-1} \|z\|_{0,A,\Omega \setminus S}) |z_T|_{1,A} \\ &\lesssim |z|_{1,A,\Omega \setminus S} |z_T|_{1,A}. \end{aligned}$$

$I_H(\eta z)$ vanishes on $\Omega \setminus N(\text{supp}(1 - \eta))$. Hence, we infer with the properties of η , (3.3), (3.2), and the resolution condition (4.2) that

$$\begin{aligned} &|\mathcal{B}(z_T, I_H(\eta z))| \\ &\lesssim \|I_H(\eta z)\|_{1,A,k,N(\text{supp}(1 - \eta))} |z_T|_{1,A} \\ &\lesssim \left(|z|_{1,A,N^2(\text{supp}(1 - \eta))} + k \|\eta z - I_H(\eta z)\|_{N(\text{supp}(1 - \eta))} + k \|\eta z\|_{N(\text{supp}(1 - \eta))} \right) |z_T|_{1,A} \\ &\lesssim \left(|z|_{1,A,N^2(\text{supp}(1 - \eta))} + k H \varepsilon^{-1} |\eta z|_{1,A,N^2(\text{supp}(1 - \eta))} + k H \varepsilon^{-1} |z|_{1,A,N(\text{supp}(1 - \eta))} \right) |z_T|_{1,A} \\ &\lesssim |z|_{1,A,N^2(\text{supp}(1 - \eta))} |z_T|_{1,A}. \end{aligned}$$

All in all, summation over all T and Cauchy inequality yield with the finite overlap of patches

$$|z|_{1,A}^2 \lesssim C_{\text{ol},m}^{1/2} |z|_{1,A} \left(\sum_{T \in \mathcal{T}_H} |z_T|_{1,A}^2 \right)^{1/2}.$$

Combination with (5.2) finishes the proof. \square

5.2 Proof of the well-posedness of the method and of the error estimates

We first prove the discrete inf-sup condition which implies that the solution of the multiscale method is indeed well-defined.

Proof of Theorem 4.4. Let $v_H \in V_H$. From the inf-sup condition for the model problem we infer that there exists $\psi \in H^1(\Omega)$ with $\|\psi\|_{1,A,k} = 1$ such that

$$\operatorname{Re} \mathcal{B}((\operatorname{id} + \mathcal{Q})v_H, \psi) \geq \gamma(\varepsilon, k) \|(\operatorname{id} + \mathcal{Q})v_H\|_{1,A,k}.$$

Define now $\psi_H := I_H \psi$. We have

$$\begin{aligned} & \mathcal{B}((\operatorname{id} + \mathcal{Q}_m)v_H, (\operatorname{id} + \mathcal{Q}_m^*)\psi_H) \\ &= \mathcal{B}((\operatorname{id} + \mathcal{Q}_m)v_H, (\operatorname{id} + \mathcal{Q}_m^*)\psi_H - (\operatorname{id} + \mathcal{Q}^*)\psi_H) + \mathcal{B}((\operatorname{id} + \mathcal{Q}_m)v_H, (\operatorname{id} + \mathcal{Q}^*)\psi_H) \\ &= \mathcal{B}((\operatorname{id} + \mathcal{Q}_m)v_H, (\mathcal{Q}_m^* - \mathcal{Q}^*)\psi_H) + \mathcal{B}((\operatorname{id} + \mathcal{Q})v_H, (\operatorname{id} + \mathcal{Q}^*)\psi_H). \end{aligned}$$

Since \mathcal{Q}^* is a projection onto W , we have $(\operatorname{id} + \mathcal{Q}^*)I_H \psi = (\operatorname{id} + \mathcal{Q}^*)\psi$. The solution properties of \mathcal{Q}^* and (4.1) of \mathcal{Q} imply $\mathcal{B}((\operatorname{id} + \mathcal{Q})v_H, (\operatorname{id} + \mathcal{Q}^*)\psi_H) = \mathcal{B}((\operatorname{id} + \mathcal{Q})v_H, \psi)$. Hence, the continuity of \mathcal{B} and (4.4) imply

$$\begin{aligned} & \operatorname{Re} \mathcal{B}((\operatorname{id} + \mathcal{Q}_m)v_H, (\operatorname{id} + \mathcal{Q}_m^*)\psi_H) \\ & \geq \gamma(\varepsilon, k) \|(\operatorname{id} + \mathcal{Q})v_H\|_{1,A,k} - \|(\operatorname{id} + \mathcal{Q}_m)v_H\|_{1,A,k} \|(\mathcal{Q}_m^* - \mathcal{Q}^*)\psi_H\|_{1,A} \\ & \gtrsim \gamma(\varepsilon, k) \|(\operatorname{id} + \mathcal{Q})v_H\|_{1,A,k} - C_{\text{ol},m}^{1/2} \beta^m \|(\operatorname{id} + \mathcal{Q}_m)v_H\|_{1,A,k} \|\psi_H\|_{1,A,k}. \end{aligned}$$

The stability of I_H yields $\|\psi_H\|_{1,A,k} \lesssim \|\psi\|_{1,A,k} = 1$ and

$$\|v_H\|_{1,A,k} = \|I_H((\operatorname{id} + \mathcal{Q})v_H)\|_{1,A,k} \lesssim \|(\operatorname{id} + \mathcal{Q})v_H\|_{1,A,k}.$$

Moreover, we have the following stability for \mathcal{Q}_m :

$$\|\mathcal{Q}_m v_H\|_{1,A,k} \lesssim \|v_H\|_{1,A,k}.$$

This finally gives

$$\operatorname{Re} \mathcal{B}((\operatorname{id} + \mathcal{Q}_m)v_H, (\operatorname{id} + \mathcal{Q}_m^*)\psi_H) \gtrsim (\gamma(\varepsilon, k) - C_{\text{ol},m}^{1/2} \beta^m) \|v_H\|_{1,A,k} \|\psi_H\|_{1,A,k},$$

which together with the oversampling condition (4.5) finishes the proof. \square

We can now prove the a priori error estimates which shed alight onto the approximation properties of the multiscale method.

Proof of Theorem 4.5. Proof of (4.7): Denote $e := u - u_{\text{LOD},m} = u - (\operatorname{id} + \mathcal{Q}_m)u_{H,m}$ and define $e_{H,m} := (\operatorname{id} + \mathcal{Q}_m)I_H e = (\operatorname{id} + \mathcal{Q}_m)(I_H u - u_{H,m})$. The triangle inequality gives

$$\|e\|_{1,A,k} \leq \|e - e_{H,m}\|_{1,A,k} + \|e_{H,m}\|_{1,A,k}.$$

We will show that under the oversampling condition (4.5) the second term can be bounded by the first term.

Let $z_H \in V_H$ be the solution of the adjoint problem

$$\mathcal{B}((\operatorname{id} + \mathcal{Q}_m)v_H, (\operatorname{id} + \mathcal{Q}_m^*)z_H) = ((\operatorname{id} + \mathcal{Q}_m)v_H, e_{H,m})_{1,A,k} \quad \forall v_H \in V_H.$$

Then we obtain with the orthogonality of W and $(\text{id} + \mathcal{Q})V_H$ and the Galerkin orthogonality $\mathcal{B}(e, (\text{id} + \mathcal{Q}_m^*)\psi_H) = 0$ for all $\psi_H \in V_H$ that

$$\begin{aligned} \|e_{H,m}\|_{1,A,k}^2 &= \mathcal{B}(e_{H,m}, (\text{id} + \mathcal{Q}_m^*)z_H) \\ &= \mathcal{B}(e_{H,m}, (\mathcal{Q}_m^* - \mathcal{Q}^*)z_H) + \mathcal{B}(e_{H,m}, (\text{id} + \mathcal{Q}^*)z_H) \\ &= \mathcal{B}(e_{H,m}, (\mathcal{Q}_m^* - \mathcal{Q}^*)z_H) + \mathcal{B}(e, (\text{id} + \mathcal{Q}^*)z_H) \\ &= \mathcal{B}(e_{H,m}, (\mathcal{Q}_m^* - \mathcal{Q}^*)z_H) + \mathcal{B}(e, (\mathcal{Q}^* - \mathcal{Q}_m^*)z_H) \\ &= \mathcal{B}(e - e_{H,m}, (\mathcal{Q}^* - \mathcal{Q}_m^*)z_H) \end{aligned}$$

The solution z_H of the adjoint problem fulfills the following stability

$$\|z_H\|_{1,A,k} \leq \gamma_{\text{LOD}}^{-1} \|e_{H,m}\|_{1,A,k}.$$

Using this stability and the corrector error (4.4) results in

$$\|e_{H,m}\|_{1,A,k}^2 \lesssim C_{\text{ol},m}^{1/2} \gamma_{\text{LOD}}^{-1} \|e - e_{H,m}\|_{1,A,k} \|e_{H,m}\|_{1,A,k}.$$

Dividing by $\|e_{H,m}\|_{1,A,k}$ on both sides and using the oversampling condition (4.5) gives the bound $\|e_{H,m}\|_{1,A,k} \lesssim \|e - e_{H,m}\|_{1,A,k}$. Second, we estimate the error $e - e_{H,m} = u - (\text{id} + \mathcal{Q}_m)I_H u \in W$. With the ellipticity of \mathcal{B} over W , the \mathcal{B} -orthogonality between W and $(\text{id} + \mathcal{Q})V_H$ and Galerkin orthogonality we obtain

$$\begin{aligned} \|e - e_{H,m}\|_{1,A,k}^2 &\lesssim \mathcal{B}(e - e_{H,m}, e - e_{H,m}) \\ &= \mathcal{B}(u - (\text{id} + \mathcal{Q})I_H u, e - e_{H,m}) + \mathcal{B}((\mathcal{Q} - \mathcal{Q}_m)I_H u, e - e_{H,m}) \\ &= \mathcal{B}(u, e - e_{H,m}) + \mathcal{B}((\mathcal{Q} - \mathcal{Q}_m)I_H u, e - e_{H,m}) \\ &= (f, e - e_{H,m}) + \mathcal{B}((\mathcal{Q} - \mathcal{Q}_m)I_H u, e - e_{H,m}). \end{aligned}$$

The last term can be estimated employing (4.4) as

$$|\mathcal{B}((\mathcal{Q} - \mathcal{Q}_m)I_H u, e - e_{H,m})| \lesssim C_{\text{ol},m}^{1/2} \beta^m \|u\|_{1,A,k} \|e - e_{H,m}\|_{1,A,k},$$

where we also used the continuity of \mathcal{B} and the stability of I_H . $\|u\|_{1,A,k}$ can be estimated against the data with the inf-sup constant of the model problem. For the first term, we can insert $I_H(e - e_{H,m})$ and use the approximation properties of I_H to obtain

$$|(f, e - e_{H,m})| \lesssim H \|A^{-1/2} f\| \|e - e_{H,m}\|_{1,A,k}.$$

Proof of (4.8): We again abbreviate $e := u - u_{\text{LOD},m}$ and consider the following two adjoint problems: (i) Find $z \in H^1(\Omega)$ such that

$$\mathcal{B}(\psi, z) = (\psi, e)_{0,A} \quad \forall \psi \in H^1(\Omega);$$

(ii) Find $z_{H,m} \in V_H$ such that

$$\mathcal{B}((\text{id} + \mathcal{Q}_m)\psi_H, (\text{id} + \mathcal{Q}_m^*)z_{H,m}) = ((\text{id} + \mathcal{Q}_m)\psi_H, e)_{0,A} \quad \forall \psi \in H^1(\Omega).$$

Set $z_{\text{LOD},m} := (\text{id} + \mathcal{Q}_m^*)z_{H,m}$. Then we deduce with Galerkin orthogonality

$$\|e\|_{0,A}^2 = |\mathcal{B}(e, z)| = |\mathcal{B}(e, z - z_{\text{LOD},m})| \lesssim \|e\|_{1,A,k} \|z - z_{\text{LOD},m}\|_{1,A,k}.$$

Note that z and $z_{H,m}$ are solutions to adjoint problems with a volume term on the right-hand side of the form Ae . Hence, we deduce similar to the above estimate for $\|e\|_{1,A,k}$ that

$$\|z - z_{\text{LOD},m}\|_{1,A,k} \lesssim H \|A^{1/2} e\| + C_{\text{ol},m}^{1/2} \beta^m \gamma^{-1}(\varepsilon, k) \|Ae\|.$$

Combination of the foregoing estimates finishes the proof.

Proof of (4.9): Let u_{LOD} be the solution to the idealized variant of (3.9), i.e., with $N^m(T) = \Omega$ for all elements T , as introduced at the beginning of Section 4.2. Because of (3.3) and the definition of the norms in (2.2) and (2.4), we obtain

$$\begin{aligned} \|u - u_{H,m}\|_{0,A} &\leq \|u - I_H u\|_{0,A} + \|I_H u - u_{H,m}\|_{0,A} \\ &\lesssim H \|u - I_H u\|_{1,A} + \|I_H u - u_{H,m}\|_{1,A,k}. \end{aligned}$$

The projection property of I_H and its stability (3.2) imply for the first term

$$\|u - I_H u\|_{1,A} \lesssim \inf_{v_H \in V_H} \|u - v_H\|_{1,A}.$$

To estimate the second term, we abbreviate $e_H := I_H u - u_{H,m} \in V_H$ and note that by the definition of \mathcal{Q} and the stability (3.4), we have

$$\|e_H\|_{1,A,k} = \|I_H(\text{id} + \mathcal{Q})e_H\|_{1,A,k} \lesssim \|(\text{id} + \mathcal{Q})e_H\|_{1,A,k}.$$

Because of the discrete inf-sup condition (4.6) (which also holds for the idealized version of the LOD), there exist $z_H \in V_H$ with $\|z_H\|_{1,A,k} = 1$ such that

$$\|(\text{id} + \mathcal{Q})e_H\|_{1,A,k} \lesssim \gamma_{\text{LOD}}^{-1} |\mathcal{B}((\text{id} + \mathcal{Q})e_H, (\text{id} + \mathcal{Q}^*)z_H)|.$$

Proceeding in a similar way as for the estimate of $e_{H,m}$ above, we obtain with the definition of \mathcal{Q} and Galerkin orthogonality that

$$\begin{aligned} \|e_H\|_{1,A,k} &\lesssim \gamma_{\text{LOD}}^{-1} |\mathcal{B}((\text{id} + \mathcal{Q})e_H, (\text{id} + \mathcal{Q}^*)z_H)| \\ &= \gamma_{\text{LOD}}^{-1} |\mathcal{B}((\text{id} + \mathcal{Q})I_H u - (\text{id} + \mathcal{Q})u_{H,m}, (\text{id} + \mathcal{Q}^*)z_H)| \\ &= \gamma_{\text{LOD}}^{-1} |\mathcal{B}(u - (\text{id} + \mathcal{Q}_m)u_{H,m}, (\text{id} + \mathcal{Q}^*)z_H)| \\ &= \gamma_{\text{LOD}}^{-1} |\mathcal{B}(u - u_{\text{LOD},m}, (\mathcal{Q}^* - \mathcal{Q}_m)z_H)| \\ &\lesssim \gamma_{\text{LOD}}^{-1} C_{\text{ol},m}^{1/2} \beta^m \|u - u_{\text{LOD},m}\|_{1,A,k}, \end{aligned}$$

where we used (4.4) in the last estimate. The remaining error $\|u - u_{\text{LOD},m}\|_{1,A,k}$ has been estimated in (4.7). \square

6 Numerical experiments

We investigate the method and the a priori estimates in several experiments. The convergence history plots display the relative error in the indicated norm versus the (coarse) mesh size H . We consider the LOD solution $u_{\text{LOD},m}$ defined in (3.10) as well as its FE part $u_{H,m}$, cf. (3.9). For comparison, we also compute a standard FE solution (denoted as P1FEM in the convergence plots) on V_H as well as the best approximation in V_H with respect to the indicated norm (denoted as P1-best in the convergence plots). The solution plots display the real part with a color map truncated to the interval $[-2, 2]$ to visualize the wave behavior outside of the scatterer, which is our main interest. We use a uniform simplicial mesh on the unit square $\Omega = (0, 1)^2$ in all our experiments and compute all correctors, i.e., we do not exploit the periodic structure as discussed in Section 3.3. As right-hand side we use

$$f(x) = \begin{cases} 10000 \exp\left(-\frac{1}{1 - \left(\frac{|x-x_0|}{0.05}\right)^2}\right) & \text{if } \frac{|x-x_0|}{0.05} < 1, \\ 0 & \text{else,} \end{cases}$$

with varying center x_0 . We compare results for the A -weighted interpolation operator I_H (3.5) as well as its unweighted variant I_H^1 (3.6).

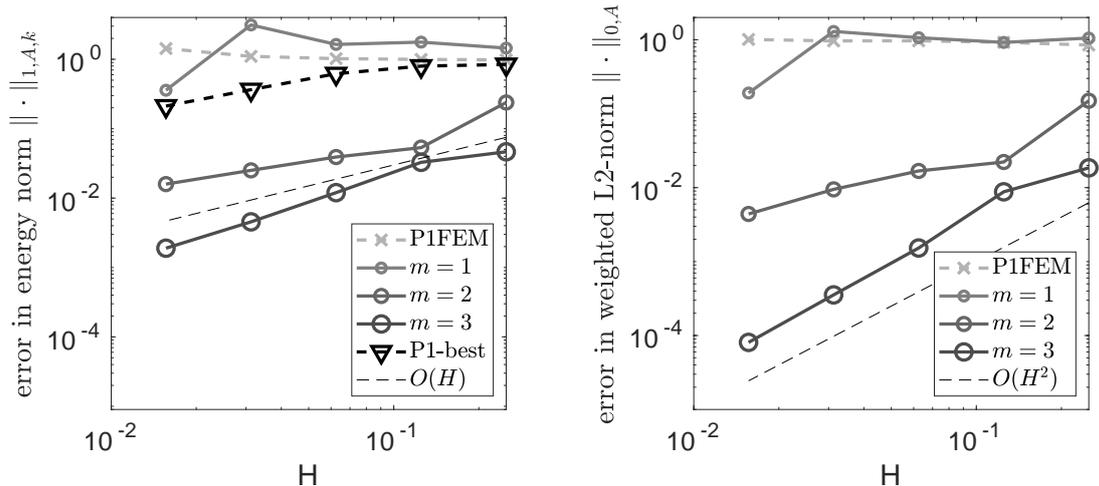


Figure 6.1: Convergence history of the (relative) error $u - u_{\text{LOD},m}$ for $\varepsilon = 2^{-3}$. In the left figure, P1-best denotes the best approximation in V_H w.r.t. $\| \cdot \|_{1,A,k}$.

6.1 Periodic structure with Mie resonances

We consider a periodic scatterer in the domain $(0.25, 0.75)^2$ with

$$\Omega_\varepsilon = (0.25, 0.75)^2 \cap \bigcup_{j \in \mathbb{Z}^2} \varepsilon(j + (0.25, 0.75)^2).$$

We choose the center for f as $x_0 = (0.125, 0.5)$ and the wave number $k = 9$. This choice of k is connected to a negative-valued effective wave number $k^2 \mu$ in homogenization theory and this unusual behavior is induced by (Mie) resonances in the small inclusions, cf. [44]. Therefore, we expect that an approximation of the exact solution might be hard in that case and consider it as a good test for the multiscale method. Note that this setting fulfills our assumptions, in particular Ω_ε is uniformly bounded away from $\partial\Omega$. A reference solution (still denoted u) is computed with standard finite elements with mesh size $h = 2^{-9}$ which is also the fine mesh used for the discretization of the corrector problems as explained in Section 3.3. The convergence history in H for the errors of $u_{\text{LOD},m}$ and $u_{H,m}$ with respect to the reference solution u is studied for patch sizes $m = 1, 2, 3$.

For $\varepsilon = 2^{-3}$, the results obtained with I_H and I_H^1 are very similar so that we restrict the confirmation of the convergence rates to I_H^1 . Figure 6.1 shows the convergence histories for $u - u_{\text{LOD},m}$ and we verify the linear rate in the energy norm (and the standard L^2 norm) as well as the quadratic rate in the weighted L^2 norm. A patch size of $m = 2, 3$ seems to be sufficient. We observe a clear superiority of the multiscale method over a standard finite element method on the coarse mesh, which fails to yield good approximations for these values of H . The standard FEM even deviates significantly from the FE best approximations with respect to the energy norm and weighted L^2 -norm (computed by projecting the reference solution onto V_H). The FE best approximations are outperformed by the multiscale method which takes fine-scale features into account.

In this context it is very interesting to analyze the error $u - u_{H,m}$ of the FE part $u_{H,m}$ of the multiscale method in Figure 6.2. In the standard as well as the weighted L^2 norm, we observe a convergence which closely follows the behavior of the FE best approximations in the space V_H with respect to the corresponding norms. This clearly underlines that there exists a macroscopic

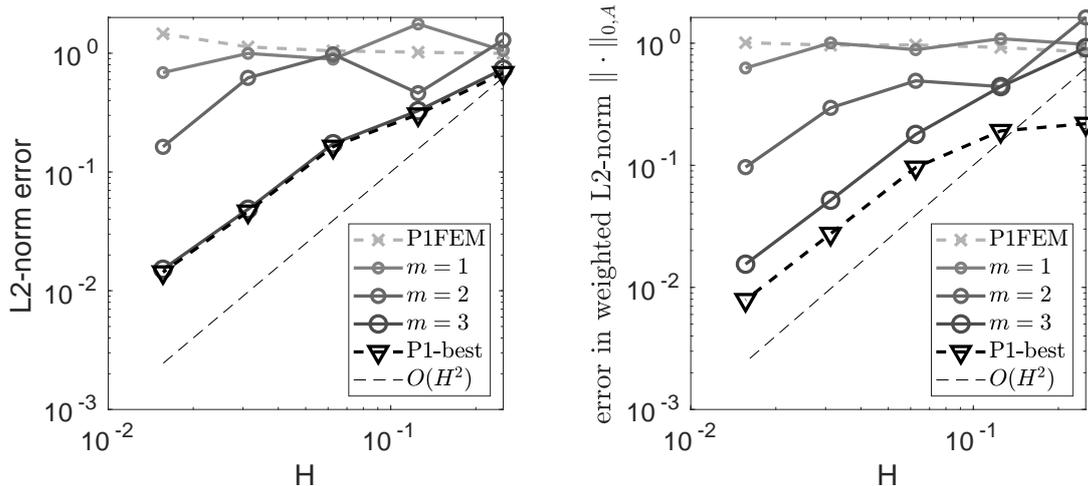


Figure 6.2: Convergence history of the (relative) error $u - u_{H,m}$ for $\varepsilon = 2^{-3}$. Note that P1-best stands for the best approximation in V_H w.r.t. the L^2 -norm (left) and the weighted L^2 -norm $\|\cdot\|_{0,A}$ (right), respectively.

approximation to the exact solution which is good in an L^2 sense. While the multiscale method is able to (almost) find this best approximation, the FEM (using the same approximation space) fails completely. Note that we observe a numerical convergence rate of about 2 for the error $u - u_{H,m}$ in this example. This is explained by higher regularity of the exact solution, which results in an order H for the best-approximation error $\inf_{v_H \in V_H} |u - v_H|_{1,A,k}$ as discussed after Theorem 4.5.

For $\varepsilon = 2^{-4}$, we compare $\|u - u_{\text{LOD},m}\|_{1,A,k}$ and $\|u - u_{H,m}\|_{0,A}$ for the unweighted interpolation operator I_H^1 (3.6) and its A -weighted variant I_H (3.5). As Figure 6.3 shows, only the weighted interpolant succeeds to yield good approximations (for patch size $m = 3$ at least). In the $\|\cdot\|_{0,A}$ -norm, the error even follows the best approximation error for the weighted operator (bottom right), which is the best we can hope for. We moreover observe a larger pre-asymptotic range for $\varepsilon = 2^{-4}$ than before for $\varepsilon = 2^{-3}$ which indicates that the resolution condition between k , ε and H is sharp in this setting.

6.2 Periodic structure with local defects

We now place the periodic structure described in the previous section into a slab-like scatterer, i.e., we set

$$\Omega_\varepsilon = ((0.25, 0.75) \times (0, 1)) \cap \bigcup_{j \in \mathbb{Z}^2} \varepsilon(j + (0.25, 0.75)^2).$$

We fix the periodicity to $\varepsilon = 2^{-3}$, see Figure 6.4 for a representation of A . Since the contrast is rather moderate, the results using I_H and I_H^1 are very similar and we only depict computations with I_H^1 . We choose the same function f as data term, this time at $x_0 = (0.25, 0.5)$, i.e., closer to the periodic structure. Moreover, we change the wave number and study $k = 28$. The reference solution is computed on a fine mesh with $h = 2^{-9}$. We emphasize that this a regime where the wave comes into (geometric) resonance with the periodic structure and where homogenization arguments can no longer be applied, see also the discussion in [16].

In this section, we consider the periodic structure and two perturbations (a point and a line

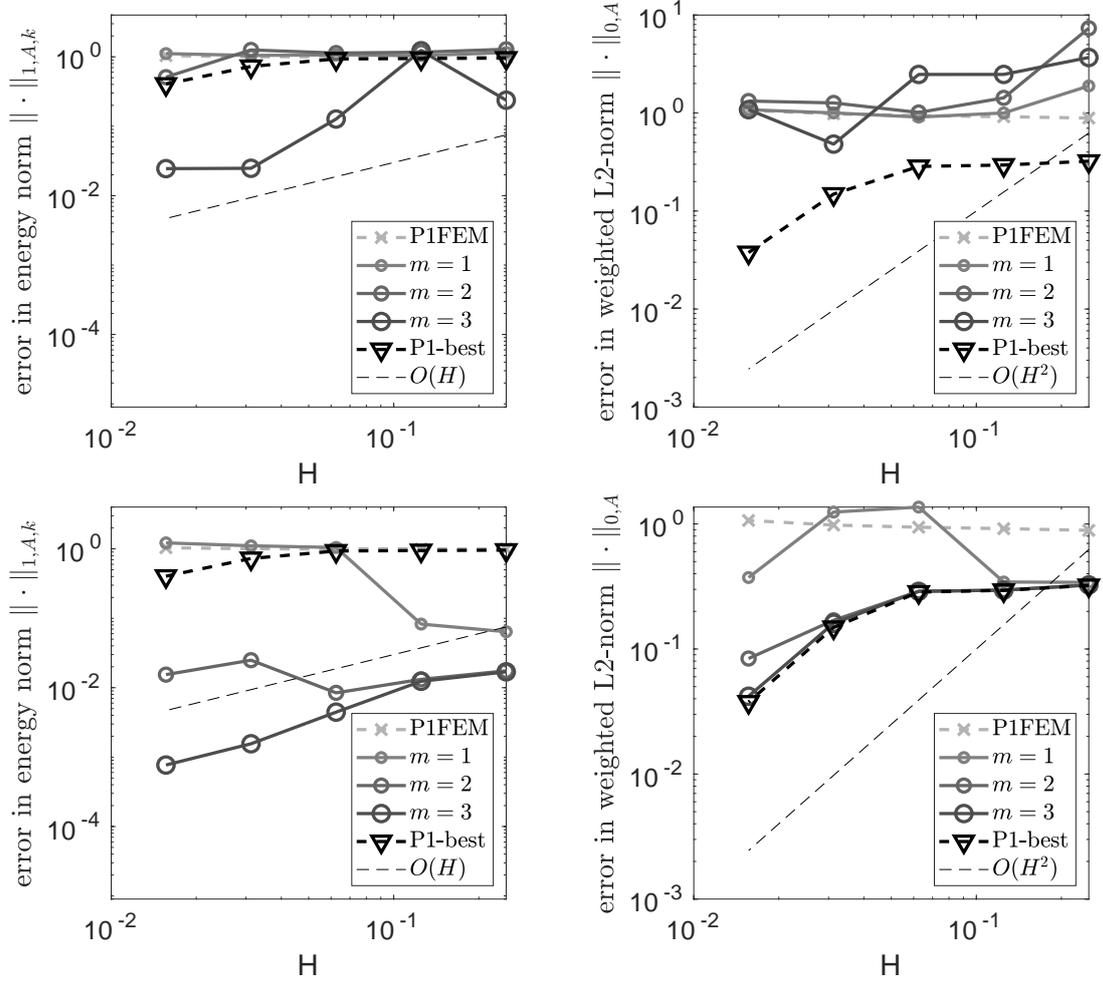


Figure 6.3: Convergence history of the (relative) errors for unweighted (top) and weighted (bottom) interpolation operator, both for $\varepsilon = 2^{-4}$. Note that P1-best stands for the best approximation in V_H w.r.t. $\|\cdot\|_{1,A,k}$ (left) and $\|\cdot\|_{0,A}$ (right), respectively.

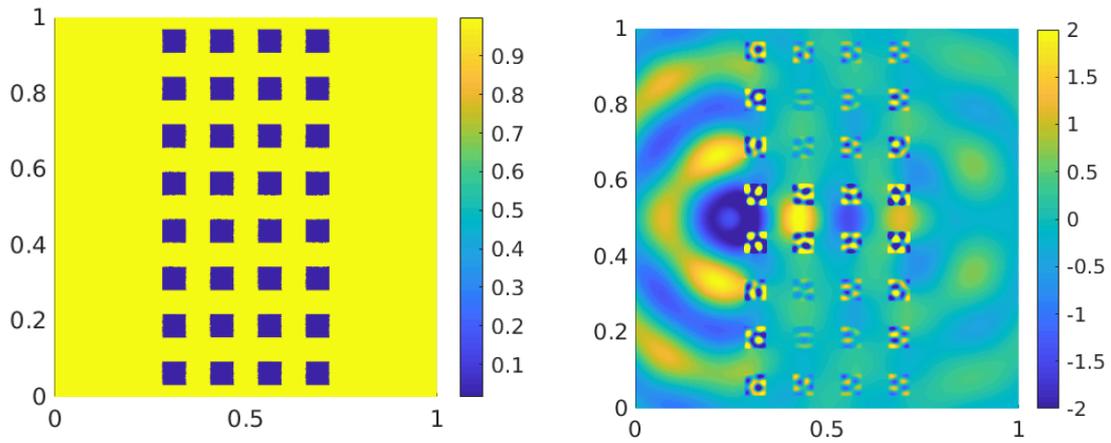


Figure 6.4: Lensing by a periodic structure, left: coefficient, right: (upscaled) LOD approximation $u_{\text{LOD},m}$ for $H = 2^{-6}$, $m = 2$.

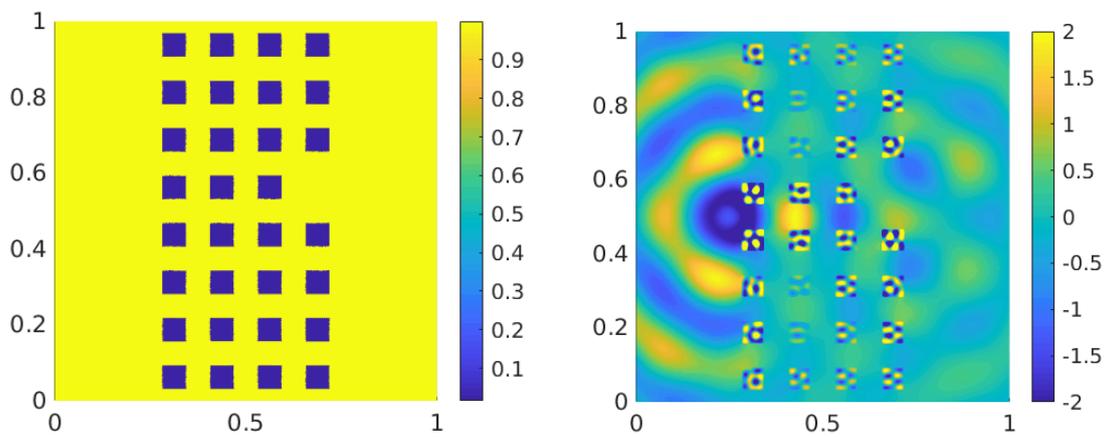


Figure 6.5: Periodic structure with a point defect, left: coefficient, right: (upscaled) LOD approximation $u_{\text{LOD},m}$ for $H = 2^{-6}$, $m = 2$.

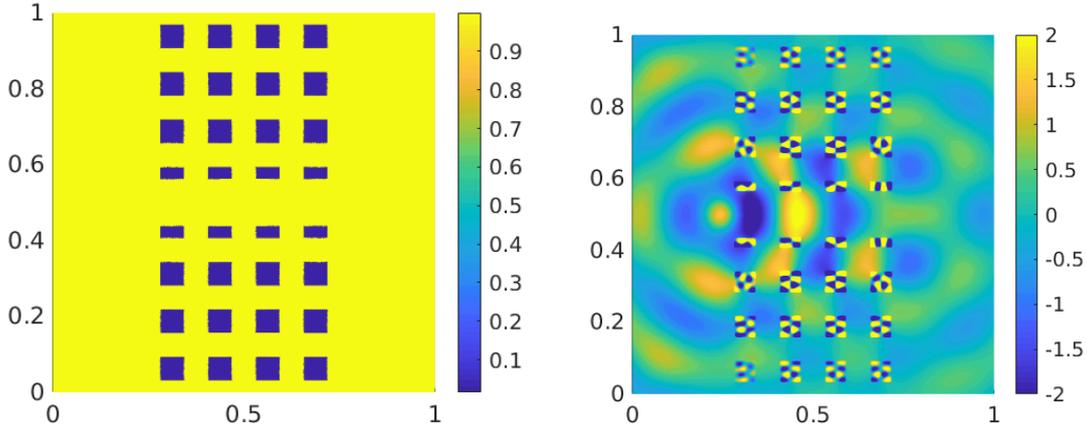


Figure 6.6: Wave guide induced by a line defect, left: coefficient, right: (upscaled) LOD approximation $u_{\text{LOD},m}$ for $H = 2^{-6}$, $m = 2$.

defect, see below), which yield physically interesting effects. As we mainly focus on the phenomena here, we do not study convergence histories, but show how (and for which meshes) the upscaled approximation $u_{\text{LOD},m}$ and its macroscopic part $u_{H,m}$ are able to faithfully represent the (qualitative) behavior of the exact solution. Compared to the wave number $k = 28$, a rather coarse mesh of $H = 2^{-6}$ and a moderate patch size of $m = 2$ is sufficient to produce a qualitatively good approximation with our multiscale method in all experiments. Already in the previous subsection, we observed that the periodic setting with $\varepsilon = 2^{-3}$ is not such a great challenge from the high contrast point of view so that a rough coupling of $kH \lesssim 1$ seems to be sufficient. We also observe that we need $H = 2^{-7}$ and $m = 2$ to have a faithful approximation by the macroscopic part $u_{H,m}$. All in all, the experiments of this section show that the multiscale method is able to capture physically relevant settings and is applicable also in the regime where no scale separation between wave length and fine-scale geometry exists.

For the perfectly periodic setting, we observe a sort of lensing effect: The wave pattern behind the scatterer looks (qualitatively) as if a source were located also behind the scatterer, see Figure 6.4. This effect is closely related to negative refraction of the material, see [16, 37]. Next, we introduce a point defect into the periodic structure: We eliminate one inclusion and leave the setting otherwise unchanged, see Figure 6.5. This local defect, however, although rather far away from the point source has a tremendous effect on the behavior of the wave. The wave pattern behind the scatterer is different from the periodic case (for instance, it is no longer symmetric around the axis $y = 0.5$), see Figure 6.5. Hence, one may conclude that the lensing effect is destroyed. From the point defect we now move to a line defect: We widen the area where $A = 1$ in the middle of the scatterer (around the line $y = 0.5$) from 0.5ε to ε , see Figure 6.6. The new (thin) channel now acts as a wave guide (cf. [4]), i.e., the wave issued from the source is mainly traveling through the new channel, see Figure 6.6.

Conclusion

We analyzed the Localized Orthogonal Decomposition for high-contrast high-frequency Helmholtz problems beyond periodicity and scale separation. It yields faithful approximations of the exact solution if the (coarse) mesh size is of the order of the effective wave length. The method relies on the solution of local finescale problems where the size of the local patches only grows loga-

rhythmically with the effective wave number. By effective wave number or wave length we denote the actual wave number or wave length divided by square root of the lowest value of the diffusion coefficient. We proved optimal convergence rates in the energy, the L^2 , and a weighted L^2 norm and linked the results to expectations from the periodic case. The numerical experiments confirmed the theoretical convergence rates and showed the applicability of the method to the Bloch wave regime and periodic structures perturbed by local (point or line) defects. We thereby simulated interesting physical effects such as lensing and wave guides. Since the method is also applicable for totally unstructured or disordered coefficients, unusual and interesting phenomena in random media can be studied in future work.

Acknowledgments

We thank S. Sauter for his comment on the unique continuation principle for the Helmholtz equation and E. A. Spence for the discussion on stability estimates. We are grateful for the remarks of the anonymous reviewers which helped to improve the paper.

References

- [1] G. Allaire. Homogenization and two-scale convergence. *SIAM J. Math. Anal.*, 23(6):1482–1518, 1992.
- [2] G. Allaire, M. Briane, and M. Vanninathan. A comparison between two-scale asymptotic expansions and Bloch wave expansions for the homogenization of periodic structures. *SeMA J.*, 73(3):237–259, 2016.
- [3] G. Allaire and C. Conca. Bloch wave homogenization and spectral asymptotic analysis. *J. Math. Pures Appl. (9)*, 77(2):153–208, 1998.
- [4] H. Ammari and F. Santosa. Guided waves in a photonic bandgap structure with a line defect. *SIAM J. Appl. Math.*, 64(6):2018–2033, 2004.
- [5] I. Babuška and S. A. Sauter. Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers? *SIAM Rev.*, 42(3):451–484 (electronic), 2000. Reprint of *SIAM J. Numer. Anal.* **34** (1997), no. 6, 2392–2423.
- [6] T. Betcke, S. N. Chandler-Wilde, I. G. Graham, S. Langdon, and M. Lindner. Condition number estimates for combined potential integral operators in acoustics and their boundary element discretisation. *Numer. Methods Partial Differential Equations*, 27(1):31–69, 2011.
- [7] G. Bouchitté and D. Felbacq. Homogenization near resonances and artificial magnetism from dielectrics. *C. R. Math. Acad. Sci. Paris*, 339(5):377–382, 2004.
- [8] G. Bouchitté and B. Schweizer. Plasmonic waves allow perfect transmission through sub-wavelength metallic gratings. *Netw. Heterog. Media*, 8(4):857–878, 2013.
- [9] S. C. Brenner and L. R. Scott. *The mathematical theory of finite element methods*, volume 15 of *Texts in Applied Mathematics*. Springer, New York, 3rd edition, 2008.
- [10] D. L. Brown, D. Gollistl, and D. Peterseim. Multiscale Petrov-Galerkin method for high-frequency heterogeneous Helmholtz equations. In M. Griebel and M. A. Schweitzer, editors, *Meshfree Methods for Partial Differential Equations VIII*, volume 115 of *Lect. Notes Comput. Sci. Eng.*, pages 85–115. Springer, Cham, 2017.

- [11] D. L. Brown and D. Peterseim. A multiscale method for porous microstructures. *Multiscale Model. Simul.*, 14(3):1123–1152, 2016.
- [12] H. Chen, P. Lu, and X. Xu. A hybridizable discontinuous Galerkin method for the Helmholtz equation with high wave number. *SIAM J. Numer. Anal.*, 51(4):2166–2188, 2013.
- [13] C.-C. Chu, I. G. Graham, and T. Y. Hou. A new multiscale finite element method for high-contrast elliptic interface problems. *Math. Comp.*, 79(272):1915–1955, 2010.
- [14] P. Ciarlet, Jr. and C. Stohrer. Finite-element heterogeneous multiscale method for the Helmholtz equation. *C. R. Math. Acad. Sci. Paris*, 352(9):755–760, 2014.
- [15] C. Conca and M. Vanninathan. Homogenization of periodic structures via Bloch decomposition. *SIAM J. Appl. Math.*, 57(6):1639–1659, 1997.
- [16] T. Dohnal and B. Schweizer. A Bloch wave numerical scheme for scattering problems in periodic wave-guides. *SIAM J. Numer. Anal.*, 56(3):1848–1870, 2018.
- [17] M. Dryja, M. V. Sarkis, and O. B. Widlund. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numer. Math.*, 72(3):313–348, 1996.
- [18] Y. Efendiev, J. Galvis, and T. Y. Hou. Generalized multiscale finite element methods (GMsFEM). *J. Comput. Phys.*, 251:116–135, 2013.
- [19] A. Efros and A. Pokrovsky. Dielectric photonic crystal as medium with negative electric permittivity and magnetic permeability. *Solid State Communications*, 129(10):643–647, 2004.
- [20] C. Engwer, P. Henning, A. Målqvist, and D. Peterseim. Efficient implementation of the localized orthogonal decomposition method. *Comput. Methods Appl. Mech. Engrg.*, 350:123–153, 2019.
- [21] S. Esterhazy and J. M. Melenk. On stability of discretizations of the Helmholtz equation. In *Numerical analysis of multiscale problems*, volume 83 of *Lect. Notes Comput. Sci. Eng.*, pages 285–324. Springer, Heidelberg, 2012. extended version on arXiv, arXiv:1105.2112.
- [22] D. Gallistl, T. Chaumont-Frelet, S. Nicaise, and J. Tomezyk. Wavenumber explicit convergence analysis for finite element discretizations of time-harmonic wave propagation problems with perfectly matched layers. *HAL preprint*, 2018.
- [23] D. Gallistl and D. Peterseim. Stable multiscale Petrov-Galerkin finite element method for high frequency acoustic scattering. *Comput. Methods Appl. Mech. Engrg.*, 295:1–17, 2015.
- [24] C. J. Gittelsohn, R. Hiptmair, and I. Perugia. Plane wave discontinuous Galerkin methods: analysis of the h -version. *M2AN Math. Model. Numer. Anal.*, 43(2):297–331, 2009.
- [25] I. G. Graham, O. R. Pembrey, and E. A. Spence. The Helmholtz equation in heterogeneous media: a priori bounds, well-posedness, and resonances. *J. Differential Equations*, 266(6):2869–2923, 2019.
- [26] I. G. Graham and S. A. Sauter. Stability and finite element error analysis for the Helmholtz equation with variable coefficients. *Math. Comp.*, 89(321):105–138, 2020.
- [27] R. Griesmaier and P. Monk. Error analysis for a hybridizable discontinuous Galerkin method for the Helmholtz equation. *J. Sci. Comput.*, 49(3):291–310, 2011.

- [28] F. Hellman, T. Keil, and A. Målqvist. Numerical upscaling of perturbed diffusion problems. arXiv preprint, 2019.
- [29] F. Hellman and A. Målqvist. Contrast independent localization of multiscale problems. *Multiscale Model. Simul.*, 15(4):1325–1355, 2017.
- [30] F. Hellman and A. Målqvist. Numerical homogenization of elliptic PDEs with similar coefficients. *Multiscale Model. Simul.*, 17(2):650–674, 2019.
- [31] P. Henning and A. Målqvist. Localized orthogonal decomposition techniques for boundary value problems. *SIAM J. Sci. Comput.*, 36(4):A1609–A1634, 2014.
- [32] P. Henning and D. Peterseim. Oversampling for the multiscale finite element method. *Multiscale Model. Simul.*, 11(4):1149–1175, 2013.
- [33] R. Hiptmair, A. Moiola, and I. Perugia. Plane wave discontinuous Galerkin methods: exponential convergence of the *hp*-version. *Found. Comput. Math.*, 16(3):637–675, 2016.
- [34] J. D. Joannapolous, S. G. Johnson, J. N. Winn, and R. D. Meade. *Photonic Crystals: Molding the Flow of Light*. Princeton University Press, Princeton, 2nd edition, 2008.
- [35] D. Lafontaine, E. A. Spence, and J. Wunsch. For most frequencies, strong trapping has a weak effect in frequency-domain scattering. arXiv preprint, 2019.
- [36] D. Lukkassen, G. Nguetseng, and P. Wall. Two-scale convergence. *Int. J. Pure Appl. Math.*, 2(1):35–86, 2002.
- [37] C. Luo, S. G. Johnson, J. Joannapolous, and J. Pendry. All-angle negative refraction without negative effective index. *Phys. Rev. B*, 65(2001104), May 2002.
- [38] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83(290):2583–2603, 2014.
- [39] J. M. Melenk. *On generalized finite-element methods*. ProQuest LLC, Ann Arbor, MI, 1995. Thesis (Ph.D.)—University of Maryland, College Park.
- [40] J. M. Melenk, A. Parsania, and S. Sauter. General DG-methods for highly indefinite Helmholtz problems. *J. Sci. Comput.*, 57(3):536–581, 2013.
- [41] J. M. Melenk and S. A. Sauter. Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions. *Math. Comp.*, 79(272):1871–1914, 2010.
- [42] J. M. Melenk and S. A. Sauter. Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation. *SIAM J. Numer. Anal.*, 49(3):1210–1243, 2011.
- [43] A. Moiola and E. A. Spence. Acoustic transmission problems: wavenumber-explicit bounds and resonance-free regions. *Math. Models Methods Appl. Sci.*, 29(2):317–354, 2019.
- [44] M. Ohlberger and B. Verfürth. A new Heterogeneous Multiscale Method for the Helmholtz equation with high contrast. *Multiscale Model. Simul.*, 16(1):385–411, 2018.
- [45] H. Owhadi and L. Zhang. Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast. *Multiscale Model. Simul.*, 9(4):1373–1398, 2011.

- [46] C. Pechstein and R. Scheichl. Weighted Poincaré inequalities and applications in domain decomposition. In Y. Huang, R. Kornhuber, O. Widlund, and J. Xu, editors, *Domain Decomposition Methods in Science and Engineering XIX*, pages 197–204. Springer Berlin Heidelberg, 2011.
- [47] J. B. Pendry, D. Schurig, and D. R. Smith. Controlling electromagnetic fields. *Science*, 312(5781):1780–1782, 2006.
- [48] I. Perugia, P. Pietra, and A. Russo. A plane wave virtual element method for the Helmholtz problem. *ESAIM Math. Model. Numer. Anal.*, 50(3):783–808, 2016.
- [49] D. Peterseim. Variational multiscale stabilization and the exponential decay of fine-scale correctors. In *Building bridges: connections and challenges in modern approaches to numerical partial differential equations*, volume 114 of *Lect. Notes Comput. Sci. Eng.*, pages 341–367. Springer, Cham, 2016.
- [50] D. Peterseim. Eliminating the pollution effect in Helmholtz problems by local subscale correction. *Math. Comp.*, 86(305):1005–1036, 2017.
- [51] D. Peterseim and R. Scheichl. Robust numerical upscaling of elliptic multiscale problems at high contrast. *Comput. Methods Appl. Math.*, 16(4):579–603, 2016.
- [52] A. Pokrovsky and A. Efros. Diffraction theory and focusing of light by a slab of left-handed material. *Physica B: Condensed Matter*, 338(1-4):333–337, 2003. Proceedings of the Sixth International Conference on Electrical Transport and Optical Properties of Inhomogeneous Media.
- [53] S. Sauter and C. Torres. Stability estimate for the Helmholtz equation with rapidly jumping coefficients. *Z. Angew. Math. Phys.*, 69(6):Art. 139, 30, 2018.