OPEN ACCESS
OPUS AUGSBURG

Universität Augsburg
Universitätsbibliothek

Robust numerical upscaling of elliptic multiscale problems at high contrast

Daniel Peterseim, Robert Scheichl

Angaben zur Veröffentlichung / Publication details:

Peterseim, Daniel, and Robert Scheichl. 2016. "Robust numerical upscaling of elliptic multiscale problems at high contrast." Computational Methods in Applied Mathematics 16 (4): 579–603. https://doi.org/10.1515/cmam-2016-0022.

Nutzungsbedingungen / Terms of use: **licgercopyright**

Dieses Dokument wird unter folgenden Bedingungen zur Verfügung gestellt: / This document is made available under th conditions:
Deutsches Urheberrecht
Weitere Informationen finden Sie unter: / For more information see:
https://www.uni-augsburg.de/de/organisation/bibliothek/publizieren-zitieren-archivieren/publiz/

Research Article

Daniel Peterseim* and Robert Scheichl

Robust Numerical Upscaling of Elliptic Multiscale Problems at High Contrast

DOI: 10.1515/cmam-2016-0022
Received January 24, 2016; revised May 14, 2016; accepted May 15, 2016

Abstract: We present a new approach to the numerical upscaling for elliptic problems with rough diffusion coefficient at high contrast. It is based on the localizable orthogonal decomposition of H^1 into the image and the kernel of some novel stable quasi-interpolation operators with local L^2-approximation properties, independent of the contrast. We identify a set of sufficient assumptions on these quasi-interpolation operators that guarantee in principle optimal convergence without pre-asymptotic effects for high-contrast coefficients. We then give an example of a suitable operator and establish the assumptions for a particular class of high-contrast coefficients. So far this is not possible without any pre-asymptotic effects, but the optimal convergence is independent of the contrast and the asymptotic range is largely improved over other discretization schemes. The new framework is sufficiently flexible to allow also for other choices of quasi-interpolation operators and the potential for fully robust numerical upscaling at high contrast.

Keywords: Finite Element, Multiscale, Upscaling, Computational Homogenization, High Contrast

MSC 2010: 65N30, 65N25, 65N15

1 Introduction

This paper presents and analyzes a novel numerical upscaling technique for the approximate solution of a prototypical partial differential equation with arbitrary positive bounded coefficients. The focus is on coefficients A that are strongly heterogeneous, i.e., A may vary rapidly on several non-separated scales. Moreover, the physical contrast (the ratio between global upper and lower bounds of its spectrum) may be very large.

The precise setting of the paper is as follows. Let $\Omega \subset \mathbb{R}^d$ be a bounded polyhedral domain and let $A \in L^\infty(\Omega, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ be a matrix-valued coefficient with uniform spectral bounds $0 < \alpha \le \beta < \infty$, i.e.,

$$\sigma(A(x)) \subset [\alpha, \beta],$$

for almost all $x \in \Omega$. Given some forcing term $g \in L^2(\Omega)$, we want to approximate the unknown weak solution u of the linear elliptic partial differential equation $-\operatorname{div}(A\nabla u) = g$ with homogeneous Dirichlet boundary condition. The function $u \in V := H^1_0(\Omega)$ is uniquely characterized by the variational problem

$$b(u, v) := \int_\Omega (A\nabla u)\cdot\nabla v \,\mathrm{d}x = \int_\Omega gv \,\mathrm{d}x =: G(v) \quad \text{for all } v \in V. \tag{1.1}$$

The accuracy of standard Galerkin finite element approximations of the unknown function u depends crucially on the regularity of the underlying data, as well as on interior angles of the domain and differentiability properties of A. On the other hand, even if the data is sufficiently regular so that a certain rate of convergence is possible, it may be observed only if the width h of the underlying mesh is sufficiently small. In this context,

*Corresponding author: Daniel Peterseim: Institut für Numerische Simulation der Universität Bonn, Wegelerstr. 6, 53115 Bonn, Germany, e-mail: peterseim@ins.uni-bonn.de
Robert Scheichl: Department of Mathematical Sciences, University of Bath, Claverton Down, Bath BA2 7AY, UK, e-mail: r.scheichl@bath.ac.uk

the notion "sufficiently small" depends on data oscillations and the contrast in a critical way. For a scalar coefficient A that oscillates between α and β at some frequency ε^{-1} with ε small, for example, the asymptotic rate of convergence is not observed unless $h \lesssim \varepsilon$. In addition, even $h \lesssim (\frac{\beta}{\alpha})^{-1}\varepsilon$ is necessary to decrease the energy error below 100%. This is too restrictive in many interesting cases. We emphasize that this condition is sharp for practically relevant right-hand sides g.

We are therefore dealing with pre-asymptotic effects for standard finite element methods and other related schemes such as finite volumes or finite differences. Due to the high variability of the coefficient functions one requires extremely fine computational grids that are able to capture all the fine scale oscillations and discontinuities. Hence, the numerical treatment of such equations is expensive in the sense that standard approaches result in systems of equations of enormous size and, consequently, in a tremendous computational demand that can not be handled in a lot of scenarios.

This paper presents a new approach for numerical upscaling based on localizable orthogonal decompositions (LOD) into a low-dimensional coarse space (where we are looking for our approximation) and a high-dimensional remainder space. Some selectable quasi-interpolation operator serves as the basis of the decompositions. The coarse space is spanned by computable basis functions with local support. The basic methodology was recently introduced in [21] and generalized in [10, 13–15]. For moderate contrast and arbitrary oscillatory coefficients this methodology yields approximations that converge to the true solution at the optimal rate (with respect to the coarse mesh size) without any pre-asymptotic effects. The analysis avoids the strong assumptions usually made in the classical homogenization framework, such as periodicity or scale separation.

The promising numerical results in [10, 12, 21] for high-contrast model coefficients are not yet reflected by the theoretical results for localized bases in those references because the physical contrast β/α appears to be a critical parameter. The dependence on β/α enters the error analysis via norm equivalences

$$\begin{cases} \beta^{-1/2}\|A^{1/2}\nabla\cdot\|_{L^2} \le \|\nabla\cdot\|_{L^2} \le \alpha^{-1/2}\|A^{1/2}\nabla\cdot\|_{L^2}, \\ \beta^{-1/2}\|A^{1/2}\cdot\|_{L^2} \le \|\cdot\|_{L^2} \le \alpha^{-1/2}\|A^{1/2}\cdot\|_{L^2}. \end{cases} \tag{1.2}$$

These equivalences are heavily used to connect variational techniques such as Galerkin orthogonality with approximation properties of standard quasi-interpolation operators in standard coefficient-independent Sobolev spaces. The idea of this paper is to circumvent the critical norm equivalences by using coefficient-dependent quasi-interpolation operators (as, e.g., in [28]) which enjoy optimal approximation properties in A-weighted Sobolev spaces.

Our multiscale method is fully defined by the choice of the quasi-interpolation operator $\mathfrak{I}_H$. We state a sufficient set of conditions on $\mathfrak{I}_H$ that will yield approximations u^{cs} that converge linearly to u in the energy norm with respect to the coarse mesh size H, without any pre-asymptotic effects and independent of the contrast. More precisely, we show that local pre-computations of the coarse basis functions on vertex patches of diameter $\approx H\log(H^{-1}\sqrt{\beta/\alpha})$ suffice to derive the following error bound:

$$\|A^{1/2}\nabla(u - u^{\mathrm{cs}})\|_{L^2(\Omega)} \le CH.$$

Here, C denotes a generic constant that is independent of the computational grid and depends only on the constants in the abstract assumptions that we have made on $\mathfrak{I}_H$. In particular, if $\mathfrak{I}_H$ can be chosen such that all the assumptions hold with constants that are independent of contrast and fine scale heterogeneity, then the convergence is also independent of such pre-asymptotic effects.

Employing (as an example) novel quasi-interpolation techniques related to those analyzed in [28], we are indeed able to satisfy the sufficient conditions with constants that are independent of the contrast. So far this is only possible under some conditions on the geometry of the coefficient relative to the coarse grid. Moreover, the constant C is not independent of H/ε and the method is thus not without pre-asymptotic effects, but it extends the asymptotic regime far beyond that of other methods independently of the contrast. Despite these limitations this result is the first one beyond heuristics to show that numerical upscaling for certain classes of high-contrast problems is possible. It may pave the way towards a comprehensive understanding of general high-contrast coefficients. In fact, our numerical tests do not show any strong pre-asymptotic effects. We

shall emphasize at this point that in the LOD framework the coarse basis functions depend on the particular choice of the quasi-interpolation operator. In that sense, the methods analyzed in this paper differ from those presented in [10, 12, 21]. Our new theoretical results improve the dependence on $\frac{\beta}{\alpha}$ of the convergence rate and of the scaling of the supports of the underlying basis functions in those papers, as well as in the alternative approaches in the literature [1, 2, 19, 22, 23]. The new results apply also to a more general class of coefficients than the analysis in [6] and [25] which is also independent of $\frac{\beta}{\alpha}$.

Our approach belongs to the large class of multiscale methods. These methods typically decouple the necessary fine scale computations into local parts to decrease the computational cost without suffering from a remarkable loss in accuracy. Prominent examples of multiscale methods are the Multiscale Finite Element Method (MsFEM) proposed by Hou and Wu [16] and the Heterogeneous Multiscale Method (HMM) by E and Engquist [8]. In contrast to our approach, MsFEM and HMM are typically not constructed for a direct approximation of the unknown solutions but for homogenized solutions and corresponding correctors instead. Thus, the reliable approximation of the exact solution is up to unknown modeling errors that punish the lack of proper periodicity and scale separation. Our framework is related to another classical multiscale method, the Variational Multiscale Method (VMM) proposed by Hughes et al. [17] (see also [18, 27]). In contrast to MsFEM and HMM, the VMM aims at a direct approximation of the exact solution without suffering from a modeling error remainder arising from homogenization theory. For connections between the methodologies we refer to [15, 27]. An interesting extension of the MsFEM to more general heterogeneous coefficients without assumptions like periodicity and scale separation is the Generalized MsFEM [9].

The remaining part of the paper is structured as follows. Section 2 defines the abstract methodological framework. In particular, abstract axioms on the underlying quasi-interpolation are formulated that guarantee contrast-independent performance of the corresponding method shown in Section 3. In Section 4 we then present particular examples of quasi interpolation operators that satisfy the previous axioms for certain classes of coefficients. Section 5 discusses the results and their limitations in the light of several numerical experiments.

2 An Abstract Multiscale Method

In this section, we propose an abstract multiscale method based on the framework of localizable orthogonal decompositions. The framework is inspired by the Variational Multiscale Method of Hughes et al. [17] but takes a very different point of view and follows the specific constructions proposed in [13, 15, 21]. For a re-interpretation within the stabilization framework of the original VMM, see [27].

The key ingredient is a continuous, surjective and uniformly stable quasi-interpolation operator from some fine scale finite element space to an initial coarse space $\tilde{V}^{\mathrm{cs}}$ that has certain L^2-approximation properties uniformly with respect to the contrast. Following the approach in [21], it is then possible to design a new coarse space that is provably robust even in the high contrast regime. The localization of the basis functions depends only mildly on the contrast.

2.1 Standard Finite Element Discretization

Let $\mathcal{T}_H$ denote a regular triangulation of Ω into closed simplices and let $H : \overline{\Omega} \to \mathbb{R}_{>0}$ denote the $\mathcal{T}_H$-piecewise constant mesh size function with $H|_T = H_T := \operatorname{diam}(T)$ for all $T \in \mathcal{T}_H$. Additionally, let $\mathcal{T}_h$ be a regular triangulation of Ω that is supposed to be a refinement of $\mathcal{T}_H$. We assume that $\mathcal{T}_h$ is sufficiently small so that all fine scale features of the coefficient A are captured. The mesh size h denotes the maximum diameter of an element of $\mathcal{T}_h$. The corresponding classical (conforming) finite element spaces of continuous piecewise polynomials of degree 1 are given by

$$V_H := \{v_H \in H_0^1(\Omega) \mid (v_H)|_T \text{ is affine for all } T \in \mathcal{T}_H\},$$
$$V_h := \{v_h \in H_0^1(\Omega) \mid (v_h)|_K \text{ is affine for all } K \in \mathcal{T}_h\}.$$

By $\mathcal{N}_H$ we denote the set of interior vertices of $\mathcal{T}_H$ (representing the degrees of freedom of the coarse finite element spaces). For every vertex $z \in \mathcal{N}_H$, let $\lambda_z \in V_H$ denote the associated nodal basis function (hat function) characterized by the property $\lambda_y(z) = \delta_{yz}$ for all $y, z \in \mathcal{N}_H$. We will also need the vertex patches

$$\omega_z := \operatorname{supp} \lambda_z = \operatorname{int}\Big(\bigcup \{T \in \mathcal{T}_H \mid x \in T\}\Big).$$

From now on, we denote by $u_h \in V_h$ the classical finite element (FE) approximation of u in the discrete (highly resolved) space V_h, i.e., $u_h \in V_h$ solves

$$b(u_h, v_h) = G(v_h) \quad \text{for all } v_h \in V_h. \tag{2.1}$$

We assume that V_h resolves the micro structure, i.e., that the error $\|u - u_h\|_{H^1(\Omega)}$ becomes sufficiently small by falling below a given tolerance. Moreover, we assume that the contrast relative to the fine mesh $\mathcal{T}_h$ is small in the sense of

$$\operatorname*{ess\,sup}_{x\in\tau} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{(A(x)v)\cdot v}{v\cdot v} \lesssim \operatorname*{ess\,inf}_{x\in\tau} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{(A(x)v)\cdot v}{v\cdot v} \tag{2.2}$$

for all $\tau \in \mathcal{T}_h$.

2.2 Abstract Quasi-Interpolation

As stated above, the key tools in our construction are an initial coarse space $\tilde{V}^{\mathrm{cs}} \subset V_h$ with certain local L^2-approximation properties and a quasi-interpolation operator $\mathcal{I}_H : V_h \to \tilde{V}^{\mathrm{cs}}$ that is linear, continuous and surjective. The kernel of this operator is going to be our fine space (or remainder space) V_h^{fs}.

To simplify the presentation we will for the most part only consider the special case, when the piecewise linear coarse space V_H has the appropriate L^2-approximation properties. As we will see, this allows us to treat a very interesting class of highly varying coefficients, namely those that are locally quasi-monotone (in the sense of [24]). We will comment briefly in Remark 2.2 below on how the framework can be extended also to other initial coarse spaces and to more general highly varying coefficients.

Thus, from now on we set $\tilde{V}^{\mathrm{cs}} := V_H$ and characterize the interpolation operator via some set of assumptions that must be fulfilled in order to derive a contrast-independent convergence result for the constructed multiscale method. Specific constructions are given in Section 4.

Assumption 2.1 (Assumptions on the interpolation). We make the following assumptions on the interpolation operator $\mathcal{I}_H : V_h \to V_H$:

(QI1) $\mathcal{I}_H \in L(V_h, V_H)$ is linear and continuous.

(QI2) The restriction of $\mathcal{I}_H$ to V_H is an isomorphism.

(QI3) There exists a generic constant C_{qip} such that, for all $v_h \in V_h$ and for all $T \in \mathcal{T}_H$,

$$H_T^{-1}\|A^{1/2}(v_h - \mathcal{I}_H v_h)\|_{L^2(T)} + \|A^{1/2}\nabla(v_h - \mathcal{I}_H v_h)\|_{L^2(T)} \le C_{\mathrm{qip}}\|A^{1/2}\nabla v_h\|_{L^2(\omega_T)}$$

with $\omega_T := \operatorname{int}(\bigcup\{K \in \mathcal{T}_H \mid K \cap T \neq \emptyset\})$.

(QI4) There exists a generic constant C'_{qip} such that, for all $v_H \in V_H$, there exists $v_h \in V_h$ with the properties

$$\mathcal{I}_H v_h = v_H, \quad \operatorname{supp} v_h \subset \operatorname{supp} v_H \quad \text{and} \quad \|A^{1/2}\nabla v_h\|_{L^2(\Omega)} \le C'_{\mathrm{qip}}\|A^{1/2}\nabla v_H\|_{L^2(\Omega)}.$$

Some remarks are in order to explain the assumptions (QI1)–(QI4). Linearity and continuity (QI1), as well as invertibility on the finite element space (QI2) are minimal assumptions that are typically satisfied by Clément-type operators. Note that $\mathcal{I}_H$ does not need to be a projection onto the finite element space V_H. The conditions ensure that the concatenation $(\mathcal{I}_H|_{V_H})^{-1}\mathcal{I}_H$ always defines such a projection. Condition (QI3) yields the crucial local approximation (resp. stability) properties in weighted L^2- (resp. energy) norms in $\tilde{V}^{\mathrm{cs}}$. Finally, assumption (QI4) ensures that any coarse finite element function $v_H \in V_H$ is the image of some function $v_h \in V_h$ under $\mathcal{I}_H$ with smaller or equal support. In other words, there exists some bounded left inverse of $\mathcal{I}_H$ that preserves local supports. This property also compensates the possible lack of a projection property. If $\mathcal{I}_H$ was a projection, then (QI4) would be satisfied by choosing $v_h = v_H$.

Remark 2.2. More generally, the initial coarse space $\tilde{V}^{\mathrm{cs}}$ could be any subspace of V_h that admits a local basis

$$\{\tilde{\lambda}_{z,\ell} \in V_h \mid z \in \mathcal{N}_H \text{ and } \ell = 1, \dots, L_z\}$$

with (i) $L_z \geq 1$ basis functions associated with each vertex $z \in \mathcal{N}_H$, (ii) $\operatorname{supp}(\tilde{\lambda}_{z,\ell}) \subset \omega_z$, (iii) $\|\tilde{\lambda}_{z,\ell}\|_{L^\infty(\Omega)} \lesssim 1$, and possibly further conditions such as a partition of unity property; see also [14]. Typical examples in the context of high contrast would be standard or generalized multiscale finite element functions [9, 16] and the associated natural quasi-interpolation operators [28]. The natural L^2-norm in (QI3) will often also be different in those cases.

2.3 Two-Scale Orthogonal Decomposition and Global Coarse Space

In this section, we construct a decomposition of the high resolution finite element space V_h into a low-dimensional space V^{cs} and some high-dimensional remainder space V^{fs}. As subspaces of V_h, both V^{cs} and V^{fs} depend on the fine scale discretization parameter h. Since the choice of h is not the topic of this paper, this dependence will not be reflected by our notation. Note that the subsequent derivation remains valid in the limit $h \to 0$ (cf. [20, 21]).

Let $\mathcal{I}_H : V_h \to V_H$ denote an interpolation operator that satisfies the properties (QI1)–(QI2) from Assumption 2.1. We define V^{fs} as the kernel of $\mathcal{I}_H$ in V_h, i.e.,

$$V^{\mathrm{fs}} := \{v \in V_h \mid \mathcal{I}_H v = 0\}.$$

The space V^{fs} represents the finescale features in V_h not captured by V_H. This definition along with properties (QI1) and (QI2) gives rise to the decomposition $V_h = V_H \oplus V^{\mathrm{fs}}$.

The key step towards the definition of an appropriate coarse space is to orthogonalize this decomposition with respect to the scalar product $b(\cdot,\cdot) = (A\nabla\cdot, \nabla\cdot)_{L^2(\Omega)}$ induced by the problem. For this purpose, we define a corresponding a-orthogonal projection $\mathcal{P}^{\mathrm{fs}} : V_h \to V^{\mathrm{fs}}$ as follows. Given $v \in V_h$, define $\mathcal{P}^{\mathrm{fs}}(v) \in V^{\mathrm{fs}}$ as the unique solution of

$$b(\mathcal{P}^{\mathrm{fs}}(v), w) = b(v, w) \quad \text{for all } w \in V^{\mathrm{fs}}.$$

The coarse scale space is defined by

$$V^{\mathrm{cs}} := (1 - \mathcal{P}^{\mathrm{fs}})V_H$$

and yields the orthogonal splitting

$$V_h = V^{\mathrm{cs}} \oplus V^{\mathrm{fs}} \quad \text{with} \quad b(V^{\mathrm{cs}}, V^{\mathrm{fs}}) = 0.$$

We shall introduce a basis of V^{cs}. The image of the nodal basis function $\lambda_z \in V_H$ under the fine scale projection $\mathcal{P}^{\mathrm{fs}}$ is denoted by $\phi_z = \mathcal{P}^{\mathrm{fs}}(\lambda_z) \in V^{\mathrm{fs}}$, i.e., ϕ_z satisfies the corrector problem

$$b(\phi_z, w) = b(\lambda_z, w) \quad \text{for all } w \in V^{\mathrm{fs}}. \tag{2.3}$$

A basis of V^{cs} is then given by the modified nodal basis

$$\{\psi_z := \lambda_z - \phi_z \mid z \in \mathcal{N}_H\}.$$

Definition 2.3 (Global coarse approximation). The Galerkin approximation $u^{\mathrm{cs}} \in V^{\mathrm{cs}}$ of the exact weak solution u of (1.1) and of the FE reference solution u_h of (2.1) is defined as the solution of

$$b(u^{\mathrm{cs}}, v) = G(v) \quad \text{for all } v \in V^{\mathrm{cs}}. \tag{2.4}$$

In general, the basis functions ψ_z have global support Ω and their pre-computation involves one fine scale computation on the whole domain Ω per coarse degree of freedom. In this sense, the pre-computation of this basis is expensive and the corresponding Galerkin discretization (2.4) yields small but densely populated stiffness and mass matrices. In certain situations, it may still be a reasonable coarsening (see Section 3.1).

A local basis may be achieved by localization of the corrector problems. Since the right-hand side of (2.3) induced by λ_z has small support, the correctors ϕ_z show an exponential decay outside of the support of λ_z. Hence, we are able to localize the correctors and their computation to local subdomains in Section 2.4.

2.4 Localized Coarse Space

We approximate the global coarse space V^{cs} from the previous section by truncating the corrector problems (2.3) for the basis functions to local patches of coarse elements, as suggested in [21].

Let $k \in \mathbb{N}$ be a discretization parameter that reflects the localization of the finescale computations. Define nodal patches $\omega_{z,k}$ of k-th order about $z \in \mathcal{N}_H$ by

$$\begin{aligned}\omega_{z,1} &:= \operatorname{supp} \lambda_z = \cup\{T \in \mathcal{T}_H \mid x \in T\},\\ \omega_{z,k} &:= \bigcup\{T \in \mathcal{T}_H \mid T \cap \omega_{z,k-1} \neq \emptyset\} \quad \text{for } k = 2, 3, 4, \dots.\end{aligned}$$

Define localized finescale spaces

$$V^{\mathrm{fs}}(\omega_{z,k}) := \{v \in V^{\mathrm{fs}} \mid v|_{\Omega\setminus\omega_{z,k}} = 0\}, \quad z \in \mathcal{N}_H,$$

by intersecting V^{fs} with those functions that vanish outside the patch $\omega_{z,k}$.

Definition 2.4 (Local correctors). Local correctors $\phi_{z,k} \in V^{\mathrm{fs}}(\omega_{z,k})$ are unique solutions of

$$b(\phi_{z,k}, w) = b(\lambda_z, w) \quad \text{for all } w \in V^{\mathrm{fs}}(\omega_{z,k}).$$

The local correctors $\phi_{z,k} \in V^{\mathrm{fs}}(\omega_{z,k})$ are approximations of the global correctors $\phi_z \in V^{\mathrm{fs}}$ from (2.3) with local support $\omega_{z,k}$. Note that homogeneous Dirichlet boundary conditions are enforced on $\partial\omega_{z,k}$. We define localized coarse spaces

$$V_k^{\mathrm{cs}} = \operatorname{span}\{\psi_{z,k} := \lambda_z - \phi_{z,k} \mid x \in \mathcal{N}_H\} \subset V.$$

Definition 2.5 (Local coarse approximation). Given some localization parameter $k \in \mathbb{N}$, the Galerkin approximation of (1.1) and (2.1) reads: Find $u_k^{\mathrm{cs}} \in V_k^{\mathrm{cs}}$ such that

$$b(u_k^{\mathrm{cs}}, v) = G(v) \quad \text{for all } v \in V_k^{\mathrm{cs}}.$$

Note that $\dim V_k^{\mathrm{cs}} = |\mathcal{N}_H| = \dim V_H$, i.e., the number of degrees of freedom of the proposed method (2.4) is the same as for the classical finite element method on the coarse mesh $\mathcal{T}_H$, or more generally, the same as for the initial coarse space $\tilde{V}^{\mathrm{cs}}$. The basis functions of the multiscale method have local support. The overlap is proportional to the parameter k. The error analysis of Section 3.2 shows that the choice

$$k \approx 2\log(H^{-1}) + \frac{1}{2}\log\Big(\frac{\beta}{\alpha}\Big)$$

suffices to preserve the desired linear convergence in H.

2.5 Alternative Localization Techniques

A modified technique for localization is presented in [15]. Define element patches of k-th order $\omega_{T,k}$ about $T \in \mathcal{T}_H$ by

$$\begin{aligned}\omega_{T,1} &:= T,\\ \omega_{T,k} &:= \bigcup\{T \in \mathcal{T}_H \mid T \cap \omega_{T,k-1} \neq \emptyset\} \quad \text{for all } k = 2, 3, 4, \dots.\end{aligned}$$

Define localized finescale spaces

$$V^{\mathrm{fs}}(\omega_{T,k}) := \{v \in V^{\mathrm{fs}} \mid v|_{\Omega\setminus\omega_{T,k}} = 0\}, \quad T \in \mathcal{T}_H,$$

by intersecting V^{fs} with those functions that vanish outside the patch $\omega_{T,k}$. The corrections are then computed in a two-step procedure. First, for any element $T \in \mathcal{T}_H$ and for any $y \in \mathcal{N}_H(T) := \mathcal{N}_H \cap T$, compute $\tilde{\psi}_{T,y,k} \in V^{\mathrm{fs}}(\omega_{T,k})$ as the unique solution of

$$b(\tilde{\psi}_{T,y,k}, w) = (A\nabla\lambda_y, \nabla w)_{L^2(T)} \quad \text{for all } w \in V^{\mathrm{fs}}(\omega_{T,k}). \tag{2.5}$$

For any node $z \in \mathcal{N}_H$, the corrector $\phi_{z,k}$ of λ_z is then defined by

$$\tilde{\psi}_{z,k} := \sum_{y \in \mathcal{N}_H \cap \omega_{z,1}} \tilde{\psi}_{T,y,k}. \tag{2.6}$$

The local problems in (2.5) are independent of each other and may be solved in parallel. However, in contrast to the localization of Section 2.4, the computation of the final correctors (2.6) requires communication among neighboring nodes. This two-step technique preserves the partition of unity property of the original basis in V_H and, hence, yields slightly improved error bounds (cf. Remark 3.4) when compared with the localization described above. The improved accuracy has also been observed in numerical experiments (cf. [15]). More general localization techniques with similar properties are discussed in [14]. Nevertheless, with regard to the already very technical error analysis of this paper, we will not include this improved localization strategy in our theory.

3 Abstract A Priori Error Analysis

In this section, we study the error of the coarse scale approximations of Definitions 2.3 and 2.5 under the abstract assumptions (QI1)–(QI4) on the underlying quasi-interpolation operator $\mathfrak{I}_H$.

Here and throughout this paper, the notation $a \lesssim b$ abbreviates $a \leq Cb$ with some multiplicative constant $C > 0$ which only depends on the domain Ω and the shape regularity of underlying finite element meshes. We emphasize that C does *not* depend on discretization parameters or on the coefficient A. Furthermore, $a \approx b$ abbreviates $a \lesssim b \lesssim a$. For parameter-dependent inequalities, $a(\xi) \lesssim b(\xi)$ means that there exists a uniform constant $C > 0$ so that $a(\xi) \leq Cb(\xi)$ holds for all parameters $\xi \in \Xi$, where the parameter set Ξ will always be clear from the context.

3.1 Error Estimates for the Global Basis

The following lemma shows the potential of the coarse space V^{cs} and the corresponding coarse approximation u^{cs}.

Lemma 3.1 (Error of the global method). *Let $u_h \in V$ solve (2.1) and $u^{\mathrm{cs}} \in V^{\mathrm{cs}}$ solve (2.4). Under the conditions (QI1)–(QI3), we have*

$$\|A^{1/2}\nabla(u_h - u^{\mathrm{cs}})\|_{L^2(\Omega)} \lesssim C_{\mathrm{qip}}\alpha^{-1/2}\|Hg\|_{L^2(\Omega)}.$$

The estimate remains valid when u_h is replaced with the weak solution $u \in V$ of (1.1).

Proof. The proof is almost verbatim the same as in [21, Lemma 3]. Let $e := u_h - u^{\mathrm{cs}}$. Galerkin orthogonality and the Cauchy–Schwarz inequality for sums yield

$$\|A^{1/2}\nabla e\|^2_{L^2(\Omega)} = b(u, u - u^{\mathrm{cs}}) = G(e) \leq \sum_{T \in \mathfrak{T}_H} \|g\|_{L^2(T)}\|e\|_{L^2(T)}.$$

Galerkin orthogonality also implies that e is b-orthogonal to the coarse space V^{cs} and, hence, that $e \in V^{\mathrm{fs}}$. Thus, due to (QI3), for any $T \in \mathfrak{T}_H$,

$$\|e\|_{L^2(T)} = \|e - \mathfrak{I}_H e\|_{L^2(T)} \leq \alpha^{-1/2}\|A^{1/2}(e - \mathfrak{I}_H e)\|_{L^2(T)} \leq \alpha^{-1/2}C_{\mathrm{qip}}H\|A^{1/2}\nabla e\|_{L^2(\omega_T)}.$$

Substituting this in the above bound and summing over $T \in \mathfrak{T}_H$ completes the proof. The constant hidden in the $\lesssim$ notation depends on the amount of overlap of the patches ω_T. □

Note that the constant C_{qip} appearing in the error bound of Lemma 3.1 can depend on the contrast if $\mathfrak{I}_H$ is not chosen properly; see also Section 4.1 for a related discussion.

3.2 Decay of Global Correctors

The following lemma is the key result of the paper.

Lemma 3.2 (Decay of global correctors). *Let* (QI1)–(QI4) *be satisfied. For any node* $z \in \mathcal{N}_H$ *and for any* $k \in \mathbb{N}$, *the correctors* ϕ_z *satisfy the estimate*

$$\|A^{1/2}\nabla\phi_z\|_{L^2(\Omega\setminus\omega_{z,k})} \lesssim \exp\Big(-\frac{k}{C_{\mathrm{qip}}^2 C'_{\mathrm{qip}}}\Big)\|A^{1/2}\nabla\phi_z\|_{L^2(\Omega)}$$

with constants C_{qip}, C'_{qip} *from Assumption 2.1.*

Proof. Let $z \in \mathcal{N}_H$ be arbitrary but fixed and, for convenience, define $\phi := \phi_z$ and $\omega_k := \omega_{z,k}$. Define cut-off functions $\zeta_{k,j} : \Omega \to [0,1] \in W^{1,\infty}(\Omega)$ for $j = 1, \dots, k-1$ such that

$$(\zeta_{k,j})|_{\omega_{k-j}} = 0, \tag{3.1}$$

$$(\zeta_{k,j})|_{\Omega\setminus\omega_k} = 1, \tag{3.2}$$

$$\|\nabla\zeta_{k,j}\|_{L^\infty(T)} \lesssim (jH_T)^{-1} \quad \text{for all } T \in \mathcal{T}_H. \tag{3.3}$$

Our particular choice of $\zeta_{k,j}$ is continuous and $\mathcal{T}_H$-piecewise affine with nodal values

$$\zeta_{k,j}(y) := \begin{cases} 0 & \text{for } y \in \mathcal{N}_H \cap \omega_{k-j}, \\ 1 & \text{for } y \in \mathcal{N}_H \cap (\Omega \setminus \omega_k), \\ \frac{m}{j} & \text{for } y \in \mathcal{N}_H \cap \partial\omega_{k-j+m} \text{ and } m = 0, 1, 2, \dots, j. \end{cases}$$

The cut-off function $\zeta_{k,j}$ allows to estimate

$$\|A^{1/2}\nabla\phi\|^2_{L^2(\Omega\setminus\omega_k)} \lesssim (\zeta_{k,j}A\nabla\phi, \nabla\phi)_{L^2(\Omega\setminus\omega_{k-j})} = (A\nabla\phi, \nabla(\zeta_{k,j}\phi))_{L^2(\Omega\setminus\omega_{k-j})} - (\phi A\nabla\phi, \nabla\zeta_{k,j})_{L^2(\Omega\setminus\omega_{k-j})}. \tag{3.4}$$

Let $I_h : V \cap C(\overline{\Omega}) \to V_h$ denote the standard nodal interpolation operator with respect to the fine mesh $\mathcal{T}_h$. Due to (QI4) there exists a $v \in V^{\mathrm{fs}}(\Omega \setminus \omega_{k-j})$ such that $\mathcal{I}_H(v)$ equals the coarse finite element function $\mathcal{I}_H(I_h(\zeta_{k,j}\phi)) \in V_H$. Introducing this into (3.4), expanding and applying the Cauchy–Schwarz inequality yields

$$\begin{aligned}\|A^{1/2}\nabla\phi\|^2_{L^2(\Omega\setminus\omega_k)} &\le \|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_{k-j})}\|A^{1/2}\nabla(\zeta_{k,j}\phi - I_h(\zeta_{k,j}\phi))\|_{L^2(\Omega\setminus\omega_{k-j})} \\ &\quad + \big|(A\nabla\phi, \nabla(I_h(\zeta_{k,j}\phi) - v))_{L^2(\Omega\setminus\omega_{k-j})}\big| + \|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_{k-j})}\|A^{1/2}\nabla v\|_{L^2(\Omega\setminus\omega_{k-j})} \\ &\quad + \|\phi A^{1/2}\nabla\zeta_{k,j}\|_{L^2(\Omega\setminus\omega_{k-j})}\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_{k-j})} \\ &=: M_1 + M_2 + M_3 + M_4. \end{aligned} \tag{3.5}$$

The four terms on the right-hand side of (3.5) are bounded separately as follows.

Bound for M_1. Recall the (local) approximation and stability properties of the nodal interpolation operator I_h (in unweighted norms), i.e.,

$$\|\nabla(v - I_h v)\|_{L^2(t)} \lesssim h_t\|\nabla^2 v\|_{L^2(t)} \quad \text{and} \quad \|\nabla I_h v\|_{L^2(t)} \lesssim \|\nabla v\|_{L^2(t)},$$

for all polynomials v on some element $t \in \mathcal{T}_h$. Since $\zeta_{k,j}\phi$ is $\mathcal{T}_h$-piecewise quadratic polynomial, this leads to

$$\begin{aligned}\|A^{1/2}\nabla(\zeta_{k,j}\phi - I_h(\zeta_{k,j}\phi))\|^2_{L^2(\Omega)} &\le \sum_{t\in\mathcal{T}_h}\|A\|_{L^\infty(t)}\|\nabla(\zeta_{k,j}\phi - I_h(\zeta_{k,j}\phi))\|^2_{L^2(t)} \\ &\lesssim \sum_{t\in\mathcal{T}_h}\|A\|_{L^\infty(t)}h_t^2\|\nabla^2(\zeta_{k,j}\phi)\|^2_{L^2(t)} \\ &\lesssim \sum_{t\in\mathcal{T}_h}\|A\|_{L^\infty(t)}h_t^2\|\nabla\zeta_{k,j}\cdot\nabla\phi\|^2_{L^2(t)} \\ &\lesssim j^{-2}\|A^{1/2}\nabla\phi\|^2_{L^2(\Omega\setminus\omega_{k-j})}. \end{aligned}$$

In the last step above, we used (2.2) and (3.1)–(3.3), as well as the trivial bound $h_t \le H$. In summary, the bound for M_1 reads

$$M_1 \lesssim j^{-1}\|A^{1/2}\nabla\phi\|^2_{L^2(\Omega\setminus\omega_{k-j})}.$$

Bound for M_2. The function v was chosen so that $(I_h(\zeta_{k,j}\phi) - v) \in V^{\mathrm{fs}}$. In conjunction with (2.3), this implies that

$$M_2 = \left|(A\nabla\lambda_z, \nabla(I_h(\zeta_{k,j}\phi) - v))_{L^2(\Omega)}\right| = 0,$$

because the intersection of the supports of λ_z and $I_h(\zeta_{k,j}\phi) - v$ has measure zero.

Bound for M_3. Due to (QI4) v satisfies the estimate

$$\|A^{1/2}\nabla v\|_{L^2(\Omega\setminus\omega_{k-j})} \leq C'_{\mathrm{qip}}\|A^{1/2}\nabla\mathfrak{I}_H(\zeta_{k,j}\phi)\|_{L^2(\Omega\setminus\omega_{k-j})}.$$

Since $\phi \in V^{\mathrm{fs}}$ (i.e., $\mathfrak{I}_H\phi = 0$) and $\phi = I_h\phi$, we have

$$0 = \mathfrak{I}_H\phi = \mathfrak{I}_H I_h\phi = \overline{\zeta}_{k,j}\mathfrak{I}_H I_h\phi = \mathfrak{I}_H I_h(\overline{\zeta}_{k,j}\phi)$$

for all constants $\overline{\zeta}_{k,j} := |T|^{-1}\int_T \zeta_{k,j}\,\mathrm{d}x$ with $T \in \mathcal{T}_H$. Using again (QI3), (2.2) and (3.1)–(3.3), and recalling that $\mathfrak{I}_H\phi = 0$, this implies

$$\begin{aligned}
&\|A^{1/2}\nabla\mathfrak{I}_H(\zeta_{k,j}\phi)\|^2_{L^2(\Omega\setminus\omega_{k-j})}\\
&\quad\leq C'^2_{\mathrm{qip}}\|A^{1/2}\nabla\mathfrak{I}_H I_h((\zeta_{k,j} - \overline{\zeta}_{k,j})\phi)\|^2_{L^2(\Omega\setminus\omega_{k-j})}\\
&\quad\lesssim C'^2_{\mathrm{qip}}C^2_{\mathrm{qip}}\|A^{1/2}\nabla((\zeta_{k,j} - \overline{\zeta}_{k,j})\phi)\|_{L^2(\Omega\setminus\omega_{k-j-1})}\\
&\quad\lesssim C'^2_{\mathrm{qip}}C^2_{\mathrm{qip}}\sum_{T\in\mathcal{T}_H:T\subset\Omega\setminus\omega_{k-j-1}}(\|\zeta_{k,j} - \overline{\zeta}_{k,j}\|^2_{L^\infty(T)}\|A^{1/2}\nabla\phi\|^2_{L^2(T)} + \|\nabla\zeta_{k,j}\|^2_{L^\infty(T)}\|A^{1/2}(\phi - \mathfrak{I}_H\phi)\|^2_{L^2(T)}).
\end{aligned}$$

The bound for M_3 now follows by applying Poincaré's inequality, the approximation property (QI3) of $\mathfrak{I}_H$, and the property (3.3) of $\zeta_{k,j}$:

$$M_3 \lesssim C^2_{\mathrm{qip}}C'_{\mathrm{qip}}j^{-1}\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_{k-j-2})}.$$

Bound for M_4. Similar arguments as before, based on $\mathfrak{I}_H\phi = 0$, the approximation property (QI3) of $\mathfrak{I}_H$, and the property (3.3) of $\zeta_{k,j}$, lead to the following bound:

$$M_4 \lesssim C_{\mathrm{qip}}j^{-1}\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_{k-j-1})}.$$

The combination of (3.5) and the bounds for $M_1, \ldots, M_4$ readily yields

$$\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_k)} \lesssim C^2_{\mathrm{qip}}C'_{\mathrm{qip}}j^{-1}\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_{k-j-2})}.$$

A sufficiently large enough choice of $j \approx C^2_{\mathrm{qip}}C'_{\mathrm{qip}}$ now establishes the following contraction:

$$\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_k)} \leq \exp(-1)\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_{k-j-2})}. \tag{3.6}$$

We emphasize that the choice of j is independent of k, as well as of the mesh sizes H and h. If the constants C_{qip} and C'_{qip} in Assumption 2.1 are independent of A, then so is j.

Repeated application of (3.6) with k replaced by $k - j - 2$ yields

$$\|A^{1/2}\nabla\phi\|_{L^2(\Omega\setminus\omega_k)} \leq \exp\left(\left\lfloor\frac{k}{j-2}\right\rfloor\right)\|A^{1/2}\nabla\phi\|_{L^2(\Omega)}.$$

This is the assertion of the lemma up to rephrasing the decay rate in terms of k, C_{qip} and C'_{qip}, and up to hiding further uncritical constants in the "$\lesssim$" notation. □

3.3 Error Estimates for the Localized Basis

The error estimate for the localized method from Definition 2.5 now follows from the global error bound of Lemma 3.1 and the decay property of the global correctors established in Lemma 3.2 via some algebraic manipulations.

Theorem 3.3 (Energy-error estimate for local coarse approximation). *If* (QI1)–(QI4) *are satisfied with constants* $C_{\mathrm{qip}} \approx C'_{\mathrm{qip}} \approx 1$ *independent of* A, *then there exists* $c \approx 1$ *such that*

$$\|A^{1/2}\nabla(u - u_k^{\mathrm{cs}})\|_{L^2(\Omega)} \lesssim \|A^{1/2}\nabla(u - u_h)\|_{L^2(\Omega)} + \alpha^{-1/2}\|Hg\|_{L^2(\Omega)} + \sqrt{\frac{\beta}{\alpha}}\alpha^{-1/2}H^{-1}e^{-ck}\|g\|_{H^{-1}(\Omega)}.$$

If, moreover, $k \geq \frac{1}{2c}\log(\frac{\beta}{\alpha}) + \frac{2}{c}\log(\frac{1}{H})$, *then*

$$\|A^{1/2}\nabla(u - u_k^{\mathrm{cs}})\|_{L^2(\Omega)} \lesssim \|A^{1/2}\nabla(u - u_h)\|_{L^2(\Omega)} + \alpha^{-1/2}H\|g\|_{L^2(\Omega)}. \tag{3.7}$$

Proof. The proof of [21, Theorem 10] applies almost verbatim to the present setting. We simply replace the contrast-dependent decay of correctors in [21] by our sharper contrast-independent result from Lemma 3.2. Moreover, the proof of [21] involves several applications of the norm equivalences (1.2) followed by L^2-approximation and stability properties. This leads to contrast-dependent constants in [21] but can be avoided here by using the approximation and stability in the A-weighted L^2-norm (QI3) directly. However, the contrast enters our proof via an inverse estimate and leads to the multiplicative constant $\sqrt{\beta/\alpha}$ in front of the exponentially decaying factor e^{-ck}. Finally, this large constant is compensated by choosing k appropriately, as stated. □

Remark 3.4 (Improved estimates with modified localization). The modified localization of Section 2.5 allows to remove the constant H^{-1} in front of the exponentially decaying factor in Theorem 3.3 so that $k \geq \frac{1}{2c}\log(\frac{\beta}{\alpha}) + \frac{1}{c}\log(\frac{1}{H})$ suffices to establish the bound in (3.7).

4 Examples of Quasi-Interpolation Operators

In this section we recall old and introduce new interpolation operators to be used in the framework presented in Section 2.

4.1 *A*-Independent Quasi-Interpolation

Previous papers (cf. [13, 15, 21]) considered a Clément-type (quasi-)interpolation operator $\mathcal{I}_H : V \to V_H$ presented in [5]. Given $v \in V$, define a (weighted) Clément interpolant

$$\mathcal{I}_H v := \sum_{z \in \mathcal{N}_H} (\mathcal{I}_H v)(z)\lambda_z \quad \text{with nodal values } (\mathcal{I}_H v)(z) := \frac{(v, \lambda_z)}{(1, \lambda_z)} \quad \text{for } z \in \mathcal{N}_H.$$

Note that the fine-scale space V^{fs} can then be characterized as L^2-orthogonal complement of V_H in V_h. The operator $\mathcal{I}_H$ does not depend on the coefficient A and satisfies (local) approximation and stability properties only in unweighted norms [5]. In particular, there exists a generic constant C depending only on the shape regularity of the finite element mesh $\mathcal{T}_H$ such that, for all $v \in V$ and for all $T \in \mathcal{T}_H$,

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

As shown in [21], this property suffices to establish an optimal a priori error bound for the global version of the method (cf. Definition 2.3):

$$\|A^{1/2}\nabla(u - u^{\mathrm{cs}})\|_{L^2(\Omega)} \lesssim \|A^{1/2}\nabla(u - u_h)\|_{L^2(\Omega)} + \alpha^{-1/2}\|Hg\|_{L^2(\Omega)}. \tag{4.1}$$

This error estimate does *not* depend on the upper spectral bound β. Hence, the reliability and accuracy of the global version of the method do not suffer from high contrast. Despite its large computational complexity, the approach may be relevant for upscaling to very coarse meshes, where localization has anyway no effect.

A further improvement in terms of accuracy – and, hence, in terms of the complexity to fall below a given error tolerance – can be achieved by substituting $\mathfrak{I}_H$ by the modified partition-of-unity-based Clément interpolation operator $\tilde{\mathfrak{I}}_H$ presented in [4]. Given $v \in V$, we define

$$\tilde{\mathfrak{I}}_H v := \sum_{z\in\mathcal{N}_H} \frac{(\tilde{\lambda}_z, v)}{(\tilde{\lambda}_z, 1)}\lambda_z \quad \text{with} \quad \tilde{\lambda}_z(x) := \frac{\lambda_z(x)}{\sum_{z\in\mathcal{N}_H}\lambda_z(x)}.$$

Since $\{\tilde{\lambda}_z : z \in \mathcal{N}_H\}$ forms a partition of unity up to the boundary, the term $\alpha^{-1/2}\|Hg\|_{L^2(\Omega)}$ in (4.1) can be replaced by data oscillations

$$\Big(\sum_{z\in\mathcal{N}} \|H(g-g_z)\|^2_{L^2(\omega_z)}\Big)^{1/2}$$

with some weighted averages g_z of g on the nodal patch ω_z, $z \in \mathcal{N}_H$; we refer to [4, Section 2] for details. Further smoothness of the right-hand side $g \in H^1(\Omega)$ then leads to quadratic convergence of the global method to the reference solution independent of contrast.

For both operators, localization of the corresponding global basis is possible even for high-contrast coefficients. However, the theory strongly requires (QI3) and (QI4) to be satisfied with constants independent of the contrast and this is not the case in general. Although its performance in the numerical experiments of Section 5 is encouraging, the question whether or not the localized version of the method with these classical quasi-interpolation operators is reliable for high-contrast coefficients remains open.

4.2 A New Quasi-Interpolation Based on A-Weighted L^2 Spaces

This subsection suggests a new quasi-interpolation operator based on A-weighted averages. For this operator we will identify a class of coefficients (with possibly high contrast) that allows us to verify the conditions (QI1)–(QI4) (see Sections 4.3–4.5 below). In particular, this operator allows for contrast-independent constants C_{qip} and C'_{qip} in (QI3) and (QI4), respectively.

The analysis is technical. To get the main ideas across, we will only consider the case of scalar coefficients, i.e. $A = aI_d$. Here, I_d is the $d \times d$ identity matrix and $a \in L^\infty(\Omega)$ with $\alpha \le a(x) \le \beta$, for almost all $x \in \Omega$. We will further assume that the coefficient function $a(x)$ is piecewise constant with respect to $\mathcal{T}_\varepsilon$, for some $h \le \varepsilon \le H$, i.e., we assume that $a(x) = a_\tau$ for all $\tau \in \mathcal{T}_\varepsilon$. Strictly speaking it is not necessary that the grids $\mathcal{T}_\varepsilon$ and $\mathcal{T}_H$ are nested but it simplifies the presentation. We assume that $\mathcal{T}_\varepsilon$ is obtained by uniform refinement from $\mathcal{T}_H$. Similarly, $\mathcal{T}_h$ is obtained by uniform refinement from $\mathcal{T}_\varepsilon$, and thus from $\mathcal{T}_H$. The extension to isotropic or mildly anisotropic tensor coefficients, as well as to coefficients that vary mildly (i.e. with benign contrast but possibly rapidly) within each of the elements $\tau \in \mathcal{T}_\varepsilon$ is also straightforward (see [24] for details).

The quasi-interpolation operator is now a coefficient-weighted generalization of the Clément-type operator presented in Section 4.1 above.

Definition 4.1 (A-weighted quasi-interpolation). Given $v \in V_h$, we define

$$\mathfrak{I}_H v := \sum_{z\in\mathcal{N}_H} \mathfrak{I}_H v(z)\lambda_z \quad \text{with} \quad \mathfrak{I}_H v(z) := \frac{\int_\Omega av\lambda_z\,\mathrm{d}x}{\int_\Omega a\lambda_z\,\mathrm{d}x}. \tag{4.2}$$

Note that the fine-scale space V^{fs} can then be characterized as orthogonal complement of V_H in V_h with respect to the A-weighted L^2-scalar product.

4.3 Characterization of Feasible High-Contrast Coefficients

To satisfy the conditions (QI1)–(QI4) for $\mathfrak{I}_H$ from Definition 4.1 with constants independent of contrast we need to make a further assumption on the type of coefficient distribution. To this end, for each vertex $z \in \mathcal{N}_H$, let $\omega_z = \text{interior}(\text{supp}(\lambda_z))$ and, for all $T \in \mathcal{T}_H$, set $\omega_T := \bigcup_{z\in\mathcal{N}_H\cap T}\omega_z$.

Assumption 4.2. We assume that there exists a generic constant C_{P}, independent of the contrast β/α, such that one of the following two Poincaré-type inequalities holds for all $v \in V_h$ and for all $T \in \mathcal{T}_H$:

$$\inf_{c\in\mathbb{R}} \int_{\omega_T} a(v-c)^2 \,\mathrm{d}x \lesssim C_{\mathrm{P}} H_T^2 \int_{\omega_T} a|\nabla v|^2 \,\mathrm{d}x$$

or

$$\partial\omega_T \cap \partial\Omega \neq \emptyset \quad \text{and} \quad \int_{\omega_T} av^2 \,\mathrm{d}x \lesssim C_{\mathrm{P}} H_T^2 \int_{\omega_T} a|\nabla v|^2 \,\mathrm{d}x.$$

Since any function $v \in V_h$ is zero on the boundary $\partial\Omega$, the existence of a constant is guaranteed for any strictly positive and uniformly bounded coefficient $a(x)$ by applying the standard Poincaré and Friedrichs inequalities on each of the subregions ω_T. Whether C_{P} is independent of the contrast β/α depends on the coefficient distribution.

4.4 On Quasi-Monotonicity

To describe the link between the local coefficient variation and the weighted Poincaré inequalities in Assumption 4.2 in more detail, let us consider a generic coarse element $T \in \mathcal{T}_H$.

We first generalize the notion of quasi-monotonicity coined in [7] by considering the following three directed combinatorial graphs: $\mathcal{G}_T^{(k)} = (\mathcal{N}_T, \mathcal{E}_T^{(k)})$, $k = 0, 1, 2$, with $\mathcal{N}_T = \{\tau \in \mathcal{T}_\varepsilon : \tau \subset \omega_T\}$. The edges are ordered pairs of vertices. To define the edges we now distinguish between three different types of connections.

Definition 4.3. Let $k \in \{1, 2, 3\}$. The ordered pair (τ, τ') is an edge in $\mathcal{E}_T^{(k)}$, if and only if $\gamma^{\tau,\tau'} = \tau \cap \tau'$ is a non-empty manifold of dimension bigger than or equal to k and $a_\tau \lesssim a_{\tau'}$. The edges in $\mathcal{E}_T^{(k)}$ are said to be of *type-k*.

Quasi-monotonicity is related to the connectivity in these graphs. Let $\tau^* = \operatorname{argmax}_{\tau\in\mathcal{N}_T} a_\tau$, i.e. an element in $\mathcal{N}_T$ (not necessarily unique) where the maximum of $a(x)$ is attained on ω_T.

Definition 4.4. The coefficient a is type-k quasi-monotone on ω_T if, for any vertex $\tau \in \mathcal{N}_T$, there exists a path from τ to τ^* in $\mathcal{G}_T^{(k)}$.

Obviously $\mathcal{E}^{(2)} \subset \mathcal{E}^{(1)} \subset \mathcal{E}^{(0)}$, and so type-$k$ quasi-monotonicity implies type-$(k-1)$ quasi-monotonicity. The coefficients in Figure 1 (a)–(c) are examples of quasi-monotone coefficients of type 2, 1 and 0, respectively. The coefficient in Figure 1 (d) is not quasi-monotone.

The following lemma summarizes the results in [24]. It relates the existence of a benign constant C_P in Assumption 4.2 that is independent of β/α directly to quasi-monotonicity. The way in which C_P depends on the ratio H/h depends on the type.

Lemma 4.5. *If a is type-k quasi-monotone on ω_T for all $T \in \mathcal{T}_H$ and for some $0 \le k \le d-1$, then Assumption 4.2 holds with*

$$C_P := \begin{cases} 1 & \text{for } k = d-1, \\ 1 + \log(\frac{H}{h}) & \text{for } k = d-2, \\ \frac{H}{h} & \text{for } k = 0 \text{ when } d = 3. \end{cases}$$

Quasi-monotonicity is a necessary condition. If the coefficient is not quasi-monotone on ω_T, then there exist $\tau, \tau' \in \mathcal{N}_T$ with $a_\tau > a_{\tau'}$ such that $C_P \ge a_\tau/a_{\tau'}$ (cf. [24, Proposition 2.11]). The coefficient C_P will in general depend on the geometry and topology of the coefficient variation. In particular, it depends on the ratio H/ε. Restricting ourselves to type-$(d-1)$ quasi-monotone coefficients, it is shown in [24, Section 4] that

$$C_P \ge \begin{cases} 1 & \text{for } d = 1, \\ 1 + \log(\frac{H}{\varepsilon}) & \text{for } d = 2, \\ \frac{H}{\varepsilon} & \text{for } d = 3. \end{cases}$$

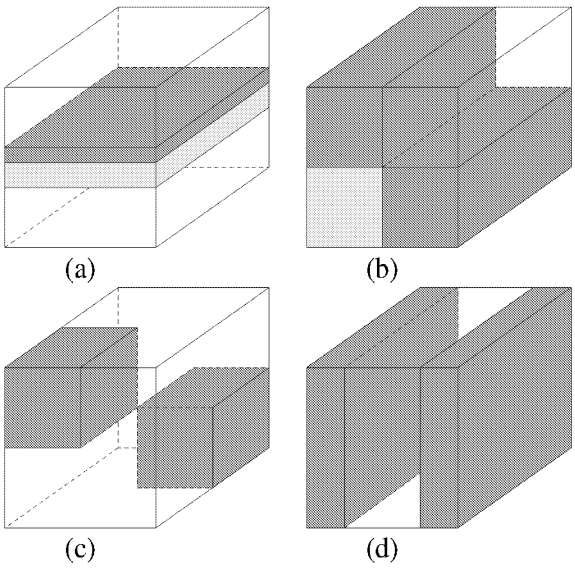


Figure 1. Quasi-monotone coefficient distributions of type 2, 1 and 0 in (a)–(c), respectively.
A darker color indicates a larger coefficient. A typical non-quasi-monotone coefficient is shown in (d).

The bounds are sharp and they are attained when $a_\tau \ll a_{\tau^*}$ for all $\tau \in \mathcal{N}_T$ with $\tau \neq \tau^*$, i.e., when the coefficient is high in only one element $\tau^* \in \mathcal{N}_T$.

4.5 Verification of (QI1)–(QI4) for *A*-Weighted Quasi-Interpolation

To verify conditions (QI1)–(QI4) for the A-weighted quasi-interpolation operator $\mathcal{I}_H$ in Definition 4.2 we need the following two technical lemmas. For the remainder of this section, we assume that $\beta/\alpha \gg H/\varepsilon$, i.e., we consider high-contrast coefficients that do not vary too rapidly relative to the coarse mesh size H.

Lemma 4.6 (Weighted inverse-type estimates). *Let $T \in \mathcal{T}_H$ and $v_H \in V_H$. Then*

$$\|v_H\|_{L^\infty(T)} \leq C_{\mathrm{inv},1}\Big(\int_T a\,\mathrm{d}x\Big)^{-1} \int_T a|v_H|\,\mathrm{d}x, \tag{4.3}$$

$$\|v_H\|_{L^\infty(T)} \leq C_{\mathrm{inv},2}\Big(\int_T a\,\mathrm{d}x\Big)^{-1/2} \|a^{1/2}v_H\|_{L^2(T)} \tag{4.4}$$

with constants $C_{\mathrm{inv},1} \eqsim C_{\mathrm{inv},2} = \mathcal{O}(H/\varepsilon)$ that are independent of the contrast β/α.

Proof. Let $\|v_H\|_{L^\infty(T)} > 0$. Otherwise the results are trivial. Now, set $\widehat{v}_H := \|v_H\|^{-1}_{L^\infty(T)} v_H$. Since $\widehat{v}_H$ is linear on T and equal to 1 at least at one of the vertices of T, it follows from simple geometric arguments that

$$\int_\tau |\widehat{v}_H|\,\mathrm{d}x \gtrsim \frac{\varepsilon}{H}|\tau| \quad \text{and} \quad \int_\tau \widehat{v}_H^2\,\mathrm{d}x \gtrsim \frac{\varepsilon^2}{H^2}|\tau| \quad \text{for all } \tau \in \mathcal{T}_\varepsilon \text{ with } \tau \subset T,$$

as for classical inverse estimates. The implied constants depend only on the dimension d and are independent of the coefficient or of any geometric parameters.

Multiplying each of these inequalities by a_τ and summing over $\tau \in \mathcal{T}_\varepsilon$ with $\tau \subset T$, we get

$$\int_T a|\widehat{v}_H|\,\mathrm{d}x \gtrsim \frac{\varepsilon}{H}\int_T a\,\mathrm{d}x \quad \text{and} \quad \int_T a\widehat{v}_H^2\,\mathrm{d}x \gtrsim \frac{\varepsilon^2}{H^2}\int_T a\,\mathrm{d}x$$

which implies the two inequalities (4.3) and (4.4). □

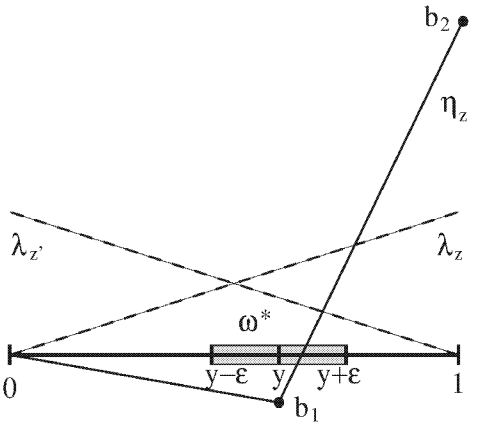


Figure 2. Construction of a suitable function η_z for Lemma 4.7 in one dimension.

Lemma 4.7. *Let Assumption 4.2 hold and let $h < \varepsilon$ be sufficiently small. Then, for every $z \in \mathcal{N}_H$, there exists a function $\eta_z \in V_h$ such that* $\operatorname{supp}(\eta_z) \subset \omega_z$, $\mathcal{I}_H \eta_z = \lambda_z$ *and*

$$\|a^{1/2}\nabla\eta_z\|_{L^2(T)} \le C_{\text{base}}\|a^{1/2}\nabla\lambda_z\|_{L^2(T)} \quad \text{for all } T \in \mathcal{T}_H. \tag{4.5}$$

The constant $C_{\text{base}} = \mathcal{O}(H^2/\varepsilon^2)$ is independent of the contrast β/α.

Proof. We will only give a proof for $d = 1$. The proof in higher dimensions is very technical and not instructive. It suffices to work elementwise and to prove the result by explicitly constructing a suitable piecewise linear function η_z that satisfies the required bound.

To simplify the presentation we focus on the particular case where $a|_{\omega^*} \equiv \beta$ for some interval $\omega^* \subset T$ with diameter $\operatorname{diam}(\omega^*) = 2\varepsilon$, and assume that $a(x) = 1$ otherwise (see Figure 2). This represents in some sense the worst case scenario. Without loss of generality, we work on the reference element $\widehat{T} = [0, 1]$, i.e., $H = 1$. Let y be the center of ω^* which by assumption is a vertex of $\mathcal{T}_h$. Now, let η_z be the piecewise linear function that is 0 at $x = 0$, b_1 at $x = y$ and b_2 at $x = 1$, as depicted in Figure 2.

Imagining a similar construction in the adjoining element containing z, we have $\eta_z \in V_h$ and $\operatorname{supp}(\eta_z) \subset \omega_z$. The values of b_1 and b_2 are chosen such that $\mathcal{I}_H\eta_z = \lambda_z$. Since

$$0 = \int_0^1 a\lambda_{z'}\eta_z\,\mathrm{d}x = \int_0^1 a(1-\lambda_z)\eta_z\,\mathrm{d}x = \int_0^1 a\eta_z\,\mathrm{d}x - \int_0^1 a\lambda_z\eta_z\,\mathrm{d}x,$$

this is equivalent to

$$\int_0^1 a\lambda_z = \int_0^1 a\lambda_z\eta_z\,\mathrm{d}x = \int_0^1 a\eta_z\,\mathrm{d}x. \tag{4.6}$$

An elementary calculation shows that

$$\begin{aligned}
\frac{1}{\beta}\int_0^1 a\lambda_z\,\mathrm{d}x &= \int_{y-\varepsilon}^{y+\varepsilon} x\,\mathrm{d}x + \mathcal{O}(\beta^{-1}) = 2y\varepsilon + \mathcal{O}(\beta^{-1}),\\
\frac{1}{\beta}\int_0^1 a\eta_z\,\mathrm{d}x &= \int_{y-\varepsilon}^{y}\frac{b_1 x}{y}\,\mathrm{d}x + \int_y^{y+\varepsilon}\Big(b_1\frac{1-x}{1-y} + b_2\frac{x-y}{1-y}\Big)\mathrm{d}x + \mathcal{O}(\beta^{-1})\\
&= \frac{\varepsilon}{2}\Big[\frac{4y(1-y)-\varepsilon}{y(1-y)}\,b_1 + \frac{\varepsilon}{1-y}\,b_2\Big] + \mathcal{O}(\beta^{-1}),\\
\frac{1}{\beta}\int_0^1 a\lambda_z\eta_z\,\mathrm{d}x &= \int_{y-\varepsilon}^{y}\frac{b_1 x^2}{y}\,\mathrm{d}x + \int_y^{y+\varepsilon} x\Big(b_1\frac{1-x}{1-y} + b_2\frac{x-y}{1-y}\Big)\mathrm{d}x + \mathcal{O}(\beta^{-1})\\
&= \frac{\varepsilon}{6}\Big[\frac{12y^2(1-y) + (1-2y)2\varepsilon^2 - 3y\varepsilon}{y(1-y)}\,b_1 + \frac{\varepsilon(3y+2\varepsilon)}{1-y}\,b_2\Big] + \mathcal{O}(\beta^{-1}).
\end{aligned}$$

Due to (4.6), all these expressions need to be equal. We equate the first expression with each of the others and neglect terms of $\mathcal{O}(\beta^{-1})$; this is justified since we assumed $\beta^{-1} \ll \varepsilon$:

$$(4y(1-y)-\varepsilon)b_1 + y\varepsilon b_2 = 4y^2(1-y), \tag{4.7}$$

$$(12y^2(1-y) + (1-2y)2\varepsilon^2 - 3y\varepsilon)b_1 + y\varepsilon(3y+2\varepsilon)b_2 = 12y^2(1-y). \tag{4.8}$$

This uniquely defines b_1 and b_2. We see that both values are independent of the contrast β. Subtracting $(3y+2\varepsilon)$ times (4.7) from (4.8) and solving for b_1, we get

$$b_1 = -\frac{y^2(3-3y-2\varepsilon)}{\varepsilon(2y-\varepsilon)}.$$

Substituting this back into (4.7), we can get an expression for b_2.

To finish the proof we need to establish (4.5) and show that $C_{\text{base}} = \mathcal{O}(\varepsilon^{-2})$ independent of the contrast β. Since

$$\frac{1}{\varepsilon\beta}\|a^{1/2}\nabla\lambda_z\|^2_{L^2(\widehat{T})} = \frac{1}{\varepsilon}\int_{y-\varepsilon}^{y+\varepsilon} 1^2\,\mathrm{d}x + \mathcal{O}\Big(\frac{1}{\beta\varepsilon}\Big) = 2 + \mathcal{O}\Big(\frac{1}{\beta\varepsilon}\Big)$$

and

$$\begin{aligned}\frac{1}{\varepsilon\beta}\|a^{1/2}\nabla\eta_z\|^2_{L^2(\widehat{T})} &= \frac{1}{\varepsilon}\int_{y-\varepsilon}^{y}\Big(\frac{b_1}{y}\Big)^2\mathrm{d}x + \frac{1}{\varepsilon}\int_{y}^{y+\varepsilon}\Big(\frac{b_2-b_1}{1-y}\Big)^2\mathrm{d}x + \mathcal{O}\Big(\frac{1}{\beta\varepsilon}\Big)\\ &= \Big(\frac{b_1}{y}\Big)^2 + \Big(\frac{b_2-b_1}{1-y}\Big)^2 + \mathcal{O}\Big(\frac{1}{\beta\varepsilon}\Big),\end{aligned} \tag{4.9}$$

it suffices to prove that the expression in (4.9) is $\mathcal{O}(\varepsilon^{-4})$ independent of β.

First, it is easy to verify that $-b_1/y$ takes its maximum at $y=\varepsilon$ with a value of $(3-5\varepsilon)/\varepsilon$. Second, it follows from (4.7) that

$$\frac{b_2-b_1}{1-y} = \frac{4}{\varepsilon}(y-b_1) + \frac{b_1}{y} \le \frac{4}{\varepsilon}(y-b_1) \le \frac{12y^2(1-y)}{\varepsilon^2(2y-\varepsilon)} \le \frac{3}{\varepsilon^2}$$

which completes the proof. □

Proposition 4.8. *Under Assumption 4.2, the operator $\mathcal{I}_H$ from Definition 4.2 satisfies the conditions* (QI1)–(QI4) *with constants C_{qip} and C'_{qip} that are independent of the contrast β/α, but depend in general on H/ε.*

Proof. Condition (QI1) is satisfied by definition.

To prove (QI2) note that, for any $v_H := \sum_{z\in\mathcal{N}_H} y_z\lambda_z \in V_H$, we have

$$\mathcal{I}_H v_H = \sum_{z\in\mathcal{N}_H} (D^{-1}M\boldsymbol{y})_z\lambda_z.$$

Here, $\boldsymbol{y} := (y_z)_{z\in\mathcal{N}_H}$, M is the mass matrix with entries $M_{z,z'} := \int_\Omega a\lambda_z\lambda_{z'}\,\mathrm{d}x$, and D is a diagonal weighting matrix with strictly positive entries $D_{z,z} := \int_\Omega a\lambda_z\,\mathrm{d}x$. Since M is invertible, the mapping $\boldsymbol{y}\mapsto D^{-1}M\boldsymbol{y}$ is bijective, and so the linear map $\mathcal{I}_H$ is an isomorphism from V_H to V_H.

The proof of (QI3) is analogous to the proof of [29, Lemma 4.1]. Let $v_h \in V_h$ and let $T \in \mathcal{T}_H$. Note first that

$$(\mathcal{I}_H v_h(z))^2 \le \frac{\int_\Omega a v_h^2\lambda_z\,\mathrm{d}x}{\int_\Omega a\lambda_z\,\mathrm{d}x}.$$

Since $\lambda_z \le 1$, this implies

$$\int_T a(\mathcal{I}_H v_h)^2\,\mathrm{d}x \le \sum_{z\in T}\frac{\int_\Omega a v_h^2\lambda_z\,\mathrm{d}x}{\int_\Omega a\lambda_z\,\mathrm{d}x}\int_T a\lambda_z^2\,\mathrm{d}x \le \int_{\omega_T} a v_h^2\,\mathrm{d}x \tag{4.10}$$

and consequently

$$\int_T a(v_h - \mathcal{I}_H v_h)^2 \,\mathrm{d}x \lesssim \int_{\omega_T} a v_h^2 \,\mathrm{d}x. \tag{4.11}$$

Similarly,

$$\int_T a|\nabla \mathcal{I}_H v_h|^2 \,\mathrm{d}x \lesssim \sum_{z\in T} \frac{\int_\Omega a v_h^2 \lambda_z \,\mathrm{d}x}{\int_\Omega a\lambda_z \,\mathrm{d}x} \int_T a|\nabla\lambda_z|^2 \,\mathrm{d}x.$$

Since λ_z is linear on T, we have $|\nabla\lambda_z|^2 = c_{T,z}H_T^{-2}$ for some constant $c_{T,z}$ that only depends on the shape of T. Hence, using the inverse estimate (4.3) with $v_H = \lambda_z$, we have

$$\frac{\int_T a|\nabla\lambda_z|^2 \,\mathrm{d}x}{\int_\Omega a\lambda_z \,\mathrm{d}x} \le c_{T,z}H_T^{-2} \frac{\int_T a \,\mathrm{d}x}{\int_T a\lambda_z \,\mathrm{d}x} \le c_{T,z}H_T^{-2} C_{\mathrm{inv},1}.$$

Since $\lambda_z \le 1$, we can conclude from (4.11) that

$$\int_T a|\nabla \mathcal{I}_H v_h|^2 \,\mathrm{d}x \lesssim C_{\mathrm{inv},1}H_T^{-2} \int_{\omega_T} a v_h^2 \,\mathrm{d}x. \tag{4.12}$$

Now, if T is such that $\partial\omega_T \cap \partial\Omega = \emptyset$, then $\{\lambda_z\}$ forms a partition of unity on ω_T and thus $\mathcal{I}_H$ preserves constants. Consequently, (4.11) and (4.12) remain true if we replace v_h by $\hat{v}_h = v_h - c$ on the right-hand side. If T is such that $\partial\omega_T \cap \partial\Omega \neq \emptyset$, we set $\hat{v}_h = v_h$. In either case, we can apply Assumption 4.2 to bound $\int_{\omega_T} a\hat{v}_h^2 \,\mathrm{d}x$. The bound in (QI3) then follows from (4.11) and (4.12) by summation over all $T \in \mathcal{T}_H$.

The constant in (QI3) satisfies $C_{\mathrm{qip}} \lesssim \sqrt{C_P\, C_{\mathrm{inv},1}}$. In the worst case, for $d = 3$, we have $C_{\mathrm{qip}} = \mathcal{O}(H/\varepsilon)$. In fact, the factor $\sqrt{C_{\mathrm{inv},1}}$ only appears in the bound of the energy error, not in the L^2-part in (QI3). The L^2-part in (QI3) can be bounded with a constant independent of H/ε in one dimension and the constant only grows logarithmically with H/ε in two dimensions.

For (QI4), we proceed as in the proof of [21, Lemma 1]. In Lemma 4.7, we already proved (QI4) for any nodal basis function λ_z, $z \in \mathcal{N}_H$. To prove (QI4) for an arbitrary $v_H := \sum_{z\in\mathcal{N}_H} v_H(z)\lambda_z \in V_H$ we choose

$$v_h = v_H + \sum_{z\in\mathcal{N}_H} (v_H(z) - \mathcal{I}_H v_H(z))\eta_z \in V_h,$$

where $\eta_z \in V_h$ is as defined in Lemma 4.7. With this choice of v_h, it follows immediately that $\mathcal{I}_H v_h = v_H$ and $\operatorname{supp} v_h \subset \operatorname{supp} v_H$.

Recall that $|\nabla\lambda_z|^2 = c_{T,z}H_T^{-2}$ on T and use the inverse estimate (4.4) together with (4.5) to conclude that

$$\begin{aligned}
\|a^{1/2}\nabla v_h\|_{L^2(T)}^2 &\lesssim \|a^{1/2}\nabla v_H\|_{L^2(T)}^2 + \sum_{z\in T} |v_H(z) - \mathcal{I}_H v_H(z)|^2 \,\|a^{1/2}\nabla\eta_z\|_{L^2(T)}^2 \\
&\le \|a^{1/2}\nabla v_H\|_{L^2(T)}^2 + \sum_{z\in T} |v_H(z) - \mathcal{I}_H v_H(z)|^2 \, C_{\mathrm{base}}\Big(\int_T a(x)\Big) c_{T,z}H_T^{-2} \\
&\le \|a^{1/2}\nabla v_H\|_{L^2(T)}^2 + C_{\mathrm{inv},2}^2 C_{\mathrm{base}}\, H_T^{-2} \,\|a^{1/2}(v_H - \mathcal{I}_H v_H)\|_{L^2(T)}^2 \\
&\lesssim C_{\mathrm{inv},2}^2 C_{\mathrm{base}} C_P \,\|a^{1/2}\nabla v_H\|_{L^2(\omega_T)}^2.
\end{aligned}$$

The last step follows as in the proof of (QI3).

Summing over all $T \in \mathcal{T}_H$, we then obtain the stability estimate in (QI4) with a constant

$$C'_{\mathrm{qip}} \lesssim C_{\mathrm{inv},2}\sqrt{C_{\mathrm{base}} C_P}.$$

In the worst case, for $d = 3$, we may have $C'_{\mathrm{qip}} = \mathcal{O}((H/\varepsilon)^{5/2})$. □

4.6 Alternative Quasi-Interpolation Operators

The interpolation operators above are associated with L^2-projections and A-weighted L^2-projections onto classical finite elements. Those projections are global operators. We will now consider local projections. In [3, 11, 26], local projections turned out to be superior over the (weighted) L^2-projections and their corresponding non-projective quasi-interpolations.

Definition 4.9. Given $v \in V_h$ and $z \in \mathcal{N}_H$, let $\mathcal{P}_z v \in V_H|_{\omega_z}$ be defined by

$$\int_{\omega_z} a\, \mathcal{P}_z v\, w_H \,\mathrm{d}x = \int_{\omega_z} a\, v\, w_H \,\mathrm{d}x \quad \text{for all } w_H \in V_H|_{\omega_z}, \tag{4.13}$$

i.e., the weighted L^2-projection of v onto the coarse finite element space restricted to the nodal patch ω_z. The global *A-weighted projective quasi-interpolation operator* is then defined as

$$\mathcal{I}_H^{\mathrm{proj},A} v := \sum_{z\in\mathcal{N}_H} \mathcal{P}_z v(z)\lambda_z.$$

Since $\mathcal{I}_H^{\mathrm{proj},A}$ is a projection, (QI2) and (QI4) are satisfied trivially here with $C'_{\mathrm{qip}} = 1$. Assumption (QI1) is also satisfied by definition. Assumption (QI3) can be verified as in the proof of Proposition 4.8 for $\mathcal{I}_H$ with a constant C_{qip} that is independent of β/α but does depend on H/ε again. The key observation for the proof of (QI3) is that the local mass matrix M_z associated with the patch ω_z, with entries $M_{z;\zeta,\zeta'} := \int_{\omega_z} a\,\lambda_\zeta\,\lambda_{\zeta'}\,\mathrm{d}x$, is spectrally equivalent to $D_z := \mathrm{diag}(M_z)$. Let $N_z := \dim(V_H|_{\omega_z})$. Then this means that

$$\mu_{\min,z}\mathbf{w}^T D_z \mathbf{w} \le \mathbf{w}^T M_z \mathbf{w} \le \mu_{\max,z}\mathbf{w}^T D_z \mathbf{w} \quad \text{for all } \mathbf{w} \in \mathbb{R}^{N_z}, \tag{4.14}$$

which in turn guarantees that

$$(\mathcal{P}_z v(z))^2 \int_{\omega_z} a\,\lambda_z^2 \,\mathrm{d}x \le \mu_{\min,z}^{-1} \int_{\omega_z} a\,(\mathcal{P}_z v)^2 \,\mathrm{d}x \le \mu_{\min,z}^{-1} \int_{\omega_z} a\, v^2 \,\mathrm{d}x$$

and allows to establish a bound akin to (4.10). The remainder follows as in the case of $\mathcal{I}_H$.

Crucially, we require that $\mu_{\min,z}^{-1}$ in (4.14) can be bounded independently of β/α. Note that $\mu_{\min,z}$ is also the smallest eigenvalue of $D_z^{-1}M_z$. As in the proof of Lemma 4.7, we restrict ourselves to a special case in one dimension. Let $d = 1$ and let $\omega^* \subset \omega_z$ with $\mathrm{diam}(\omega^*) = 2\varepsilon$. We consider $a|_{\omega^*} \equiv \beta$ and $a(x) = 1$ otherwise, as depicted in Figure 2. Without loss of generality, we assume that $\omega^* \subset T$ for one of the two elements T making up ω_z and that $H = 1$ again. An elementary calculation shows that

$$M_z = 2\beta\varepsilon \begin{bmatrix} (1-y)^2 + \frac{\varepsilon^2}{3} + \mathcal{O}(\frac{1}{\beta\varepsilon}) & y(1-y) - \frac{\varepsilon^2}{3} + \mathcal{O}(\frac{1}{\beta\varepsilon}) & 0 \\ y(1-y) - \frac{\varepsilon^2}{3} + \mathcal{O}(\frac{1}{\beta\varepsilon}) & y^2 + \frac{\varepsilon^2}{3} + \mathcal{O}(\frac{1}{\beta\varepsilon}) & \frac{1}{12\beta\varepsilon} \\ 0 & \frac{1}{12\beta\varepsilon} & \frac{1}{6\beta\varepsilon} \end{bmatrix}.$$

Considering first $y \gg \varepsilon$ and $1 - y \gg \varepsilon$ and ignoring terms of $\mathcal{O}(\frac{1}{\beta\varepsilon})$ and $\mathcal{O}(\frac{\varepsilon^4}{y^4(1-y)^4})$, we get

$$D_z^{-1}M_z = \begin{bmatrix} 1 & \frac{y}{1-y} - \frac{\varepsilon^2}{3(1-y)^3} & 0 \\ \frac{1-y}{y} - \frac{\varepsilon^2}{3y^3} & 1 & 0 \\ 0 & \frac{1}{2} & 1 \end{bmatrix}.$$

The eigenvalues of $D_z^{-1}M_z$ satisfy

$$0 = \det(\mu I - D_z^{-1}M_z) = (\mu - 1)\Big(\mu^2 - 2\mu + \frac{\varepsilon^2}{3y^2(1-y)^2}\Big)$$

leading to

$$\mu_{\min,z} = \frac{\varepsilon^2}{6y^2(1-y)^2}$$

which is independent of β. For $y = \mathcal{O}(\varepsilon)$ or $1 - y = \mathcal{O}(\varepsilon)$, it is even possible to bound $\mu_{\min,z}^{-1}$ independently of ε.

In the numerical experiments of Section 5, we will also consider the non-weighted variant $\mathcal{I}_H^{\mathrm{proj}}$ that is defined in the same way with classical L^2-inner products (i.e., $a = 1$) in (4.13).

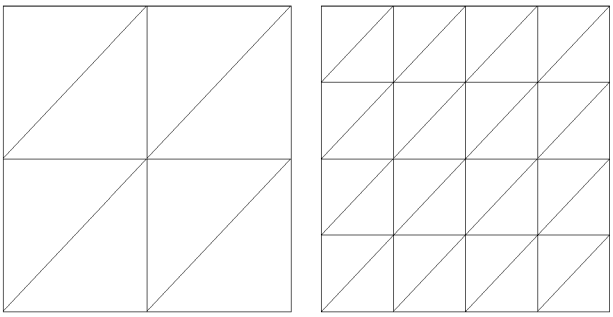


Figure 3. Uniform triangulations of the unit square used as coarse meshes in the numerical experiments of Section 5.

5 Numerical Experiments

Three numerical experiments shall illustrate our theoretical results and illuminate their sharpness and limitations. Numerical experiments with highly oscillatory and high-contrast coefficients have already been documented in [12–15, 21]. While those results were based on the classical coefficient-independent interpolation defined in Section 4.1, this section considers several choices of interpolation operators and investigates the possible benefit of using the A-weighted interpolation operators of Sections 4.2 and 4.6 when high contrast is present.

5.1 High-Contrast Blocks

The first model problem considers a two-phase coefficient with simple topology. The precise data of the first model problem is as follows:

$$\Omega := (0,1)^2,$$

$$g(x) := \begin{cases} 0 & \text{for } x \in [0, \frac{1}{2}) \times [0,1], \\ 1 & \text{for } x \in [\frac{1}{2}, 1] \times [0,1], \end{cases}$$

$$A(x) := \begin{cases} \beta & \text{for } x \in [\frac{11}{32}, \frac{5}{32}] \times [\frac{8}{32}, \frac{11}{32}] \cup [\frac{5}{32}, \frac{11}{32}] \times [\frac{8}{32}, \frac{19}{32}], \\ 1 & \text{elsewhere.} \end{cases}$$

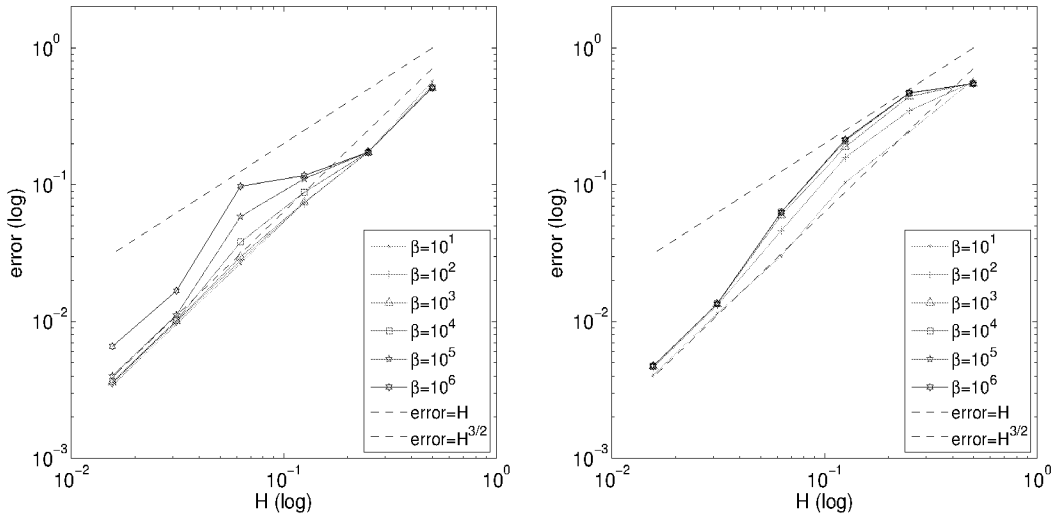
Since A takes only the values 1 and β, parameter $\beta \geq 1$ reflects the contrast. We consider $\beta = 1, 10, \ldots, 10^6$. The numerical experiment aims to study the dependence between these choices of the parameter and the accuracy of the numerical methods.

Consider the uniform triangular coarse meshes on Ω with mesh widths $\sqrt{2}H = 2^{-1}, \ldots, 2^{-6}$, as depicted in Figure 3. The reference mesh $\mathcal{T}_h$ is derived by uniform mesh refinement of the coarse mesh and has maximal mesh width $h = 2^{-8}/\sqrt{2}$. The corresponding $P1$-conforming finite element approximation on the reference mesh $\mathcal{T}_h$ is denoted by V_h. We consider the reference solution $u_h \in V_h$ of (2.1) with the above given data. We compare it with coarse scale approximations $u_k^{\mathrm{cs}} \in V_k^{\mathrm{cs}}$ (cf. Definition 2.5) for various choices of the coarse mesh size H, the localization parameter k and the underlying quasi-interpolation operator $\mathcal{I}_H$. We consider four different quasi-interpolation operators: the A-independent variant $\mathcal{I}_H$ defined in Section 4.1, the A-weighted version $\mathcal{I}_H^A$ from Definition 4.1, the A-independent operator $\mathcal{I}_H^{\mathrm{proj}}$ with projection property defined in Section 4.6 and its A-weighted variant $\mathcal{I}_H^{\mathrm{proj},A}$.

The results are visualized in Figures 4 and 5. Figure 4 shows the relative energy errors

$$\frac{\|A^{1/2}\nabla(u_h - u_{k(H)}^{\mathrm{cs}})\|}{\|A^{1/2}\nabla u_h\|},$$

as a function of the coarse mesh size H, for several choices of the contrast parameter $\beta = 1, 10, \ldots, 10^6$. The localization parameter k is tied to H via the relation $k = k(H) = |\log_2 H| + 1$ (without any dependence on β).



(a) Results for $\mathcal{I}_H$.

(b) Results for $\mathcal{I}_H^A$.

(c) Results for $\mathcal{I}_H^{\mathrm{proj}}$.

(d) Results for $\mathcal{I}_H^{\mathrm{proj},A}$.

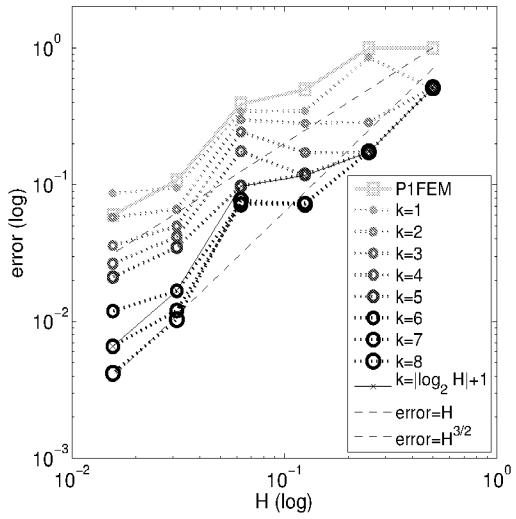
Figure 4. Numerical experiments for Section 5.1: Results for high-contrast blocks with several choices of the contrast parameter β as a function of the coarse mesh size H. The reference mesh size $h = 2^{-8}$ is fixed. The localization parameter is tied to the coarse mesh size via the relation $k = |\log_2 H| + 1$.

For all choices of interpolation operators, only a very mild dependence on β can be observed. In particular, all errors are below the reference curve H. Asymptotically, the experimental convergence rate $H^{3/2}$ is observed. This high rate is related to certain L^2- or $L^2(A)$-orthogonality properties of the interpolation operators as indicated in Section 4.1.

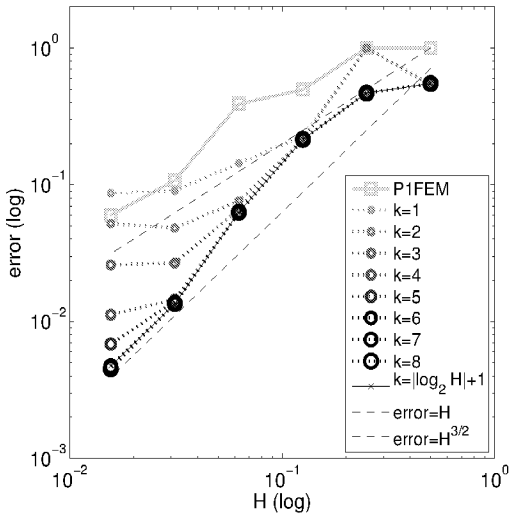
Figure 5 aims to illustrate the role of the localization parameter. It depicts relative energy errors

$$\frac{\|A^{1/2}\nabla(u_h - u_{k(H)}^{\mathrm{cs}})\|}{\|A^{1/2}\nabla u_h\|},$$

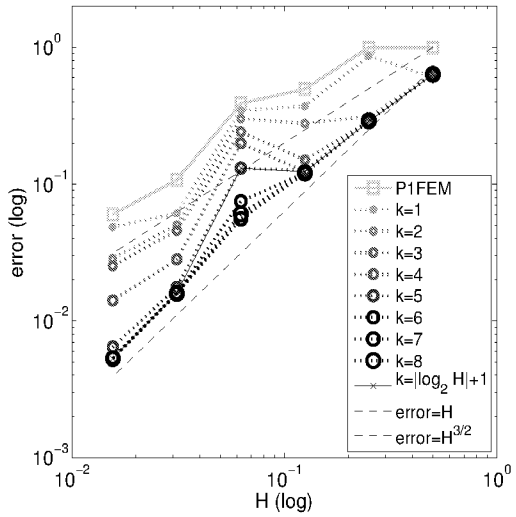
again as a function of the coarse mesh size H, for fixed contrast $\beta = 10^6$ and for several choices of the localization parameter $k = 1, 2, 3, \ldots, 8$. For comparison, we also show relative errors of the standard conforming $P1$-FEM on the coarse meshes. We observe a much faster decay of the error when k is increased for the methods that are based on A-weighted interpolation. For these methods, a fixed choice of $k = 2$ or $k = 3$ already gives very good accuracy for the range of coarse meshes considered. For these small choices of k, the methods based on A-independent interpolation are strongly affected by the high contrast. They are more accurate only for sufficiently large k.



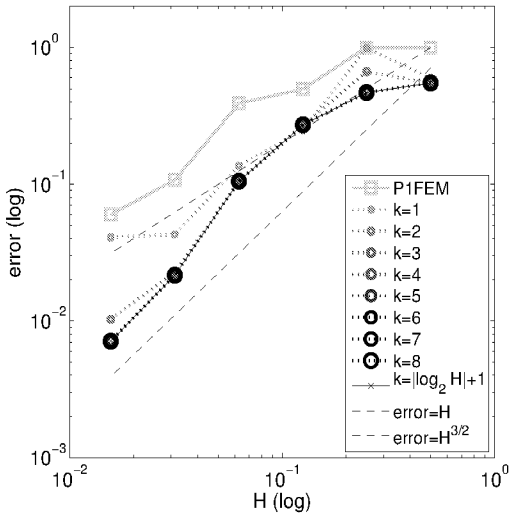
(a) Results for $\mathfrak{I}_H$.



(b) Results for $\mathfrak{I}_H^A$.



(c) Results for $\mathfrak{I}_H^{\mathrm{proj}}$.



(d) Results for $\mathfrak{I}_H^{\mathrm{proj},A}$.

Figure 5. Numerical experiments for Section 5.1: Results for high-contrast blocks with contrast parameter $\beta = 10^6$ as a function of the coarse mesh size H. The reference mesh size $h = 2^{-8}$ is fixed. The localization parameter k is varied between 1 and 8.

5.2 High-Contrast Channels

The second model problem repeats the previous computations for a different two-phase coefficient. The precise data of the second model problem is as follows:

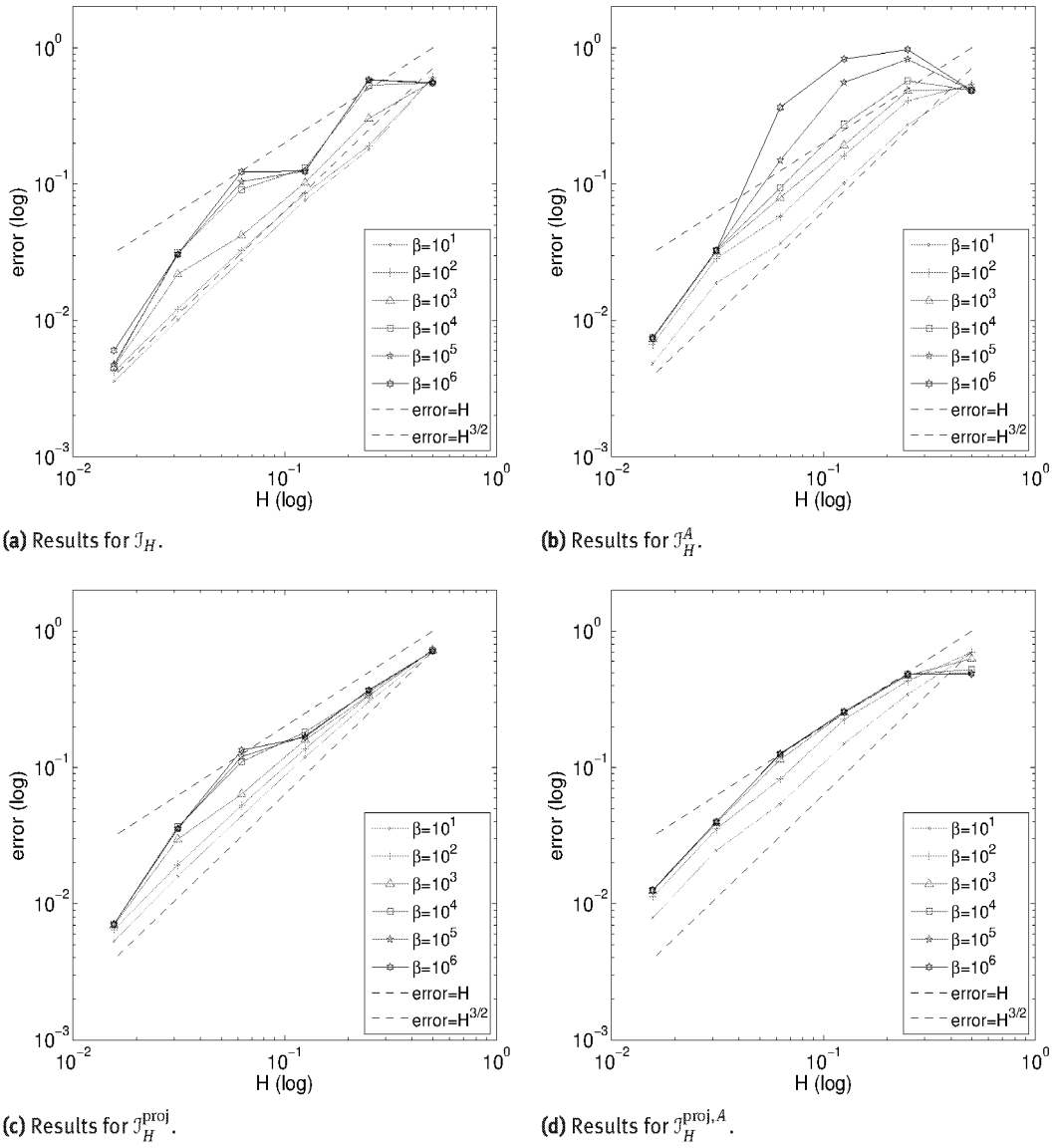
$$\Omega := (0,1)^2,$$

$$g(x) := \begin{cases} 0 & \text{for } x \in [0, \tfrac{1}{2}) \times [0,1], \\ 1 & \text{for } x \in [\tfrac{1}{2}, 1] \times [0,1], \end{cases}$$

$$A(x) = A(x_1, x_2) := A_1(x_1, x_2) + A_1(x_2, x_1),$$

$$A_1(x) := \begin{cases} \beta/2 & \text{for } x \in [\tfrac{8}{32}, \tfrac{9}{32}] \times [\tfrac{1}{32}, \tfrac{31}{32}] \cup [\tfrac{10}{32}, \tfrac{11}{32}] \times [\tfrac{1}{32}, \tfrac{31}{32}], \\ 1 & \text{elsewhere.} \end{cases}$$

Again, the parameter $\beta \geq 1$ reflects the contrast. The numerical experiment aims to study again the accuracy of the numerical methods and its dependence on this parameter.



(a) Results for $\mathfrak{I}_H$.

(b) Results for $\mathfrak{I}_H^A$.

(c) Results for $\mathfrak{I}_H^{\mathrm{proj}}$.

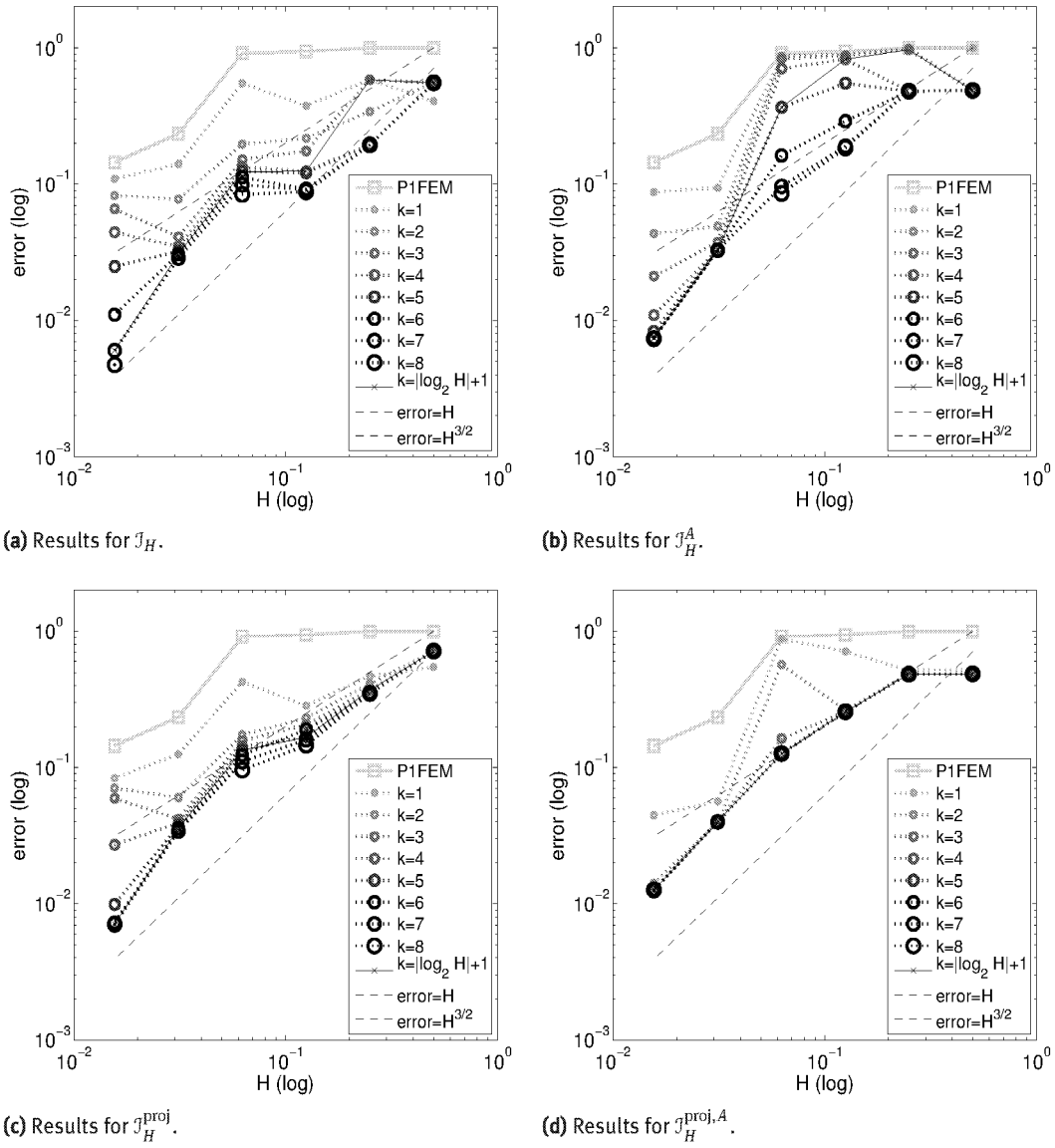
(d) Results for $\mathfrak{I}_H^{\mathrm{proj},A}$.

Figure 6. Numerical experiments for Section 5.2: Results for high-contrast channels with several choices of the contrast parameter β, as a function of the coarse mesh size H. The reference mesh size $h = 2^{-8}$ is fixed. The localization parameter $k = |\log_2 H| + 1$ is tied to the coarse mesh size.

Apart from the coefficient, the experimental setup is exactly the same as in Section 5.1. Figures 6 and 7 show the results. The observations for the operators $\mathfrak{I}_H$, $\mathfrak{I}_H^{\mathrm{proj}}$, and $\mathfrak{I}_H^{\mathrm{proj},A}$ are similar as before. The unweighted choices deliver more accuracy for sufficiently large localization parameter, whereas $\mathfrak{I}_H^{\mathrm{proj},A}$ is significantly more efficient for small k. On the other hand, the operator $\mathfrak{I}_H^A$ performs much worse in this experiment. On the coarse meshes that do not resolve the coefficient, it requires a much larger choice of k than the other operators to be accurate. We emphasize that this effect does neither contradict our theory nor does it explain it. However, it clearly shows that the choice of interpolation operator has a large impact on the actual performance of the methods, motivating further development and analysis of such operators.

5.3 Rough Coefficients with Multiscale Features

Let $\Omega := (0, 1)^2$ be the unit square. In this third problem, the scalar coefficient A is piecewise constant with respect to a uniform Cartesian grid of width 2^{-6} (see Figure 8). Its values are taken from the data of the SPE10 benchmark, see www.spe.org/web/csp/. The coefficient is highly varying and strongly heterogeneous. The



(a) Results for $\mathfrak{I}_H$.

(b) Results for $\mathfrak{I}_H^A$.

(c) Results for $\mathfrak{I}_H^{\mathrm{proj}}$.

(d) Results for $\mathfrak{I}_H^{\mathrm{proj},A}$.

Figure 7. Numerical experiments for Section 5.2: Results for high-contrast channels with contrast parameter β, as a function of the coarse mesh size H. The reference mesh size $h = 2^{-8}$ remains fixed. The localization parameter k is varied between 1 and 8.

contrast in A is large, i.e., $\beta/\alpha \approx 4 \cdot 10^6$. This coefficient is certainly not quasi-monotone with regard to the coarse meshes considered here. The right-hand side term reads

$$g(x) = \begin{cases} 8 & \text{for } x \in [0, \frac{1}{4}] \times [0, \frac{1}{4}] \cup [\frac{3}{4}, 1] \times [\frac{3}{4}, 1], \\ 0 & \text{elsewhere.} \end{cases}$$

Consider uniform coarse meshes for Ω of size $\sqrt{2}H = 2^{-1}, 2^{-2}, \ldots, 2^{-6}$ (cf. Figure 3). Note that these meshes do not have to resolve the rough coefficient A appropriately. Again, the reference mesh $\mathcal{T}_h$ has width $h = 2^{-8}/\sqrt{2}$. We compare the reference solution u_h – the $P1$ conforming finite element approximation on the reference mesh $\mathcal{T}_h$ – with coarse scale approximations. We study their dependence on the coarse mesh size H, on the choice of the interpolation operator and on the localization parameter k. Figure 9 depicts the results.

Here, the methods based on unweighted interpolation perform significantly better than the methods with A-weighted interpolation. This superiority could be related to the approximability properties of the global bases. Note that, for non-quasi-monotone coefficients, the constant in Lemma 3.1 may depend on the contrast

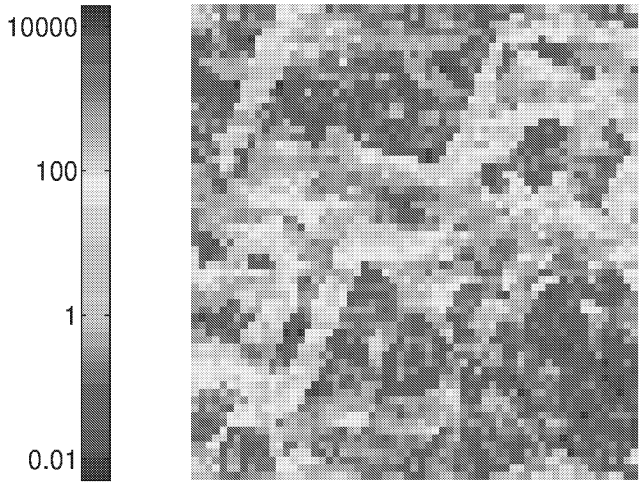
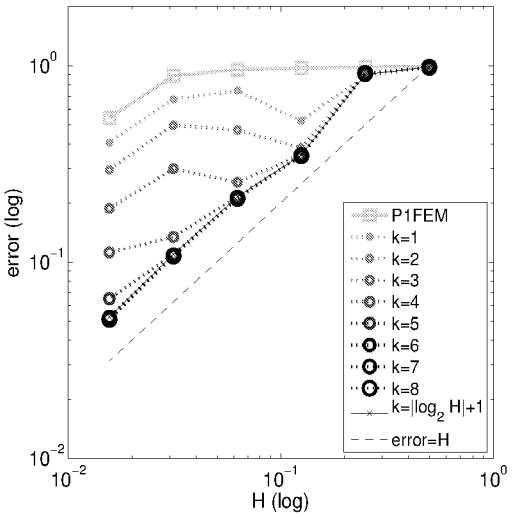
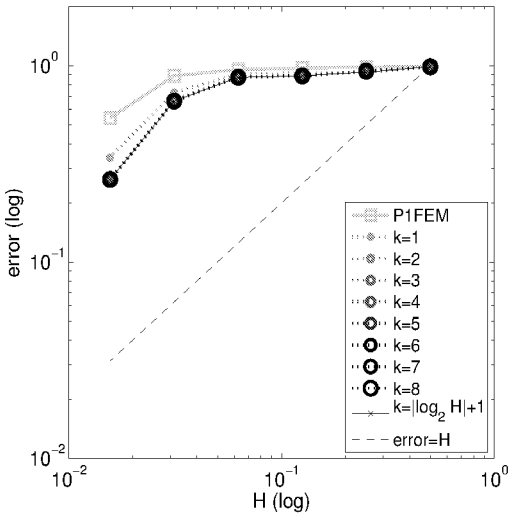


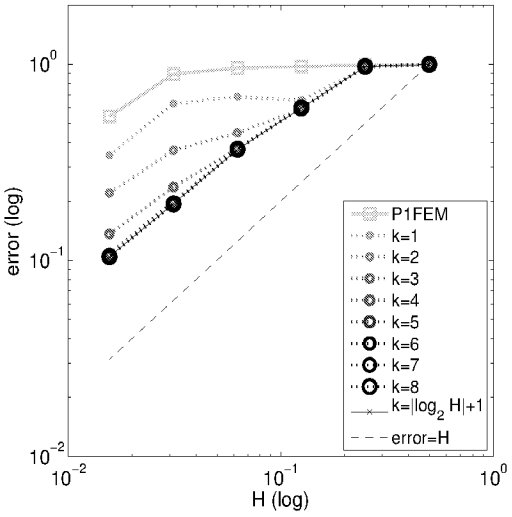
Figure 8. Scalar coefficient A used in the numerical experiment of Section 5.3.



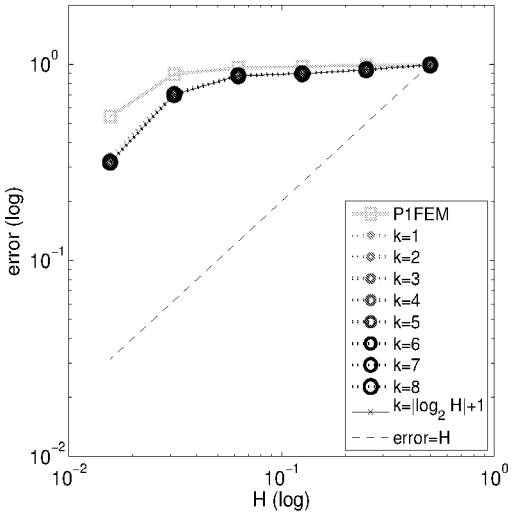
(a) Results for $\mathfrak{I}_H$.



(b) Results for $\mathfrak{I}_H^A$.



(c) Results for $\mathfrak{I}_H^{\mathrm{proj}}$.



(d) Results for $\mathfrak{I}_H^{\mathrm{proj},A}$.

Figure 9. Numerical experiments for Section 5.3: Results for SPE10 data depending on H. The reference mesh size $h = 2^{-8}$ remains fixed. The localization parameter k is varied between 1 and 9.

whereas the accuracy of the global method based on $\mathfrak{I}_H$ is independent of β (cf. equation (4.1)). However, why this nice property of the $\mathfrak{I}_H$-based method is also observed after localization remains completely open.

To sum up, it can be said that the numerical experiments clearly showed the potential of the general methodology for high-contrast problems. They also showed that, in some cases, the decay of the correctors may be accelerated significantly by using A-dependent interpolation operators for the underlying split of coarse and fine scales. This is also supported by our theoretical results. However, the theory remains pessimistic in some cases. It does not yet provide general advice regarding the choice of the interpolation operator along with an optimal choice of the localization parameter.

Acknowledgment: We thank Clemens Pechstein for suggesting the alternative, projective quasi-interpolation operator and providing us with the basic ideas for its analysis.

Funding: Supported by the Sino-German Science Center (grant id 1228) on the occasion of the Chinese-German Workshop on Computational and Applied Mathematics in Augsburg 2015.

References

[1] I. Babuška and R. Lipton, The penetration function and its application to microscale problems, *Multiscale Model. Simul.* **9** (2011), no. 1, 373–406.

[2] L. Berlyand and H. Owhadi, Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast, *Arch. Ration. Mech. Anal.* **198** (2010), 677–721.

[3] D. Brown and D. Peterseim, A multiscale method for porous microstructures, preprint (2014), http://arxiv.org/abs/1411.1944.

[4] C. Carstensen, Quasi-interpolation and a posteriori error analysis in finite element methods, *M2AN Math. Model. Numer. Anal.* **33** (1999), no. 6, 1187–1202.

[5] C. Carstensen and R. Verfürth, Edge residuals dominate a posteriori error estimates for low order finite element methods, *SIAM J. Numer. Anal.* **36** (1999), no. 5, 1571–1587.

[6] 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** (2010), no. 272, 1915–1955.

[7] M. Dryja, M. V. Sarkis and O. B. Widlund, Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions, *Numer. Math.* **72** (1996), no. 3, 313–348.

[8] W. E and B. Engquist, The heterogeneous multiscale methods, *Commun. Math. Sci.* **1** (2003), no. 1, 87–132.

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

[10] D. Elfverson, E. H. Georgoulis, A. Målqvist and D. Peterseim, Convergence of a discontinuous Galerkin multiscale method, *SIAM J. Numer. Anal.* **51** (2013), no. 6, 3351–3372.

[11] D. Gallistl and D. Peterseim, Stable multiscale Petrov–Galerkin finite element method for high frequency acoustic scattering, *Comput. Methods Appl. Mech. Engrg.* **295** (2015), 1–17.

[12] P. Henning and A. Målqvist, Localized orthogonal decomposition techniques for boundary value problems, *SIAM J. Sci. Comput.* **36** (2014), no. 4, A1609–A1634.

[13] P. Henning, A. Målqvist and D. Peterseim, A localized orthogonal decomposition method for semi-linear elliptic problems, *ESAIM Math. Model. Numer. Anal.* **48** (2014), no. 5, 1331–1349.

[14] P. Henning, P. Morgenstern and D. Peterseim, Multiscale partition of unity, in: *Meshfree Methods for Partial Differential Equations VII*, Lect. Notes Comput. Sci. Eng. 100, Springer, Cham (2015), 185–204.

[15] P. Henning and D. Peterseim, Oversampling for the multiscale finite element method, *Multiscale Model. Simul.* **11** (2013), no. 4, 1149–1175.

[16] T. Y. Hou and X.-H. Wu, A multiscale finite element method for elliptic problems in composite materials and porous media, *J. Comput. Phys.* **134** (1997), no. 1, 169–189.

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

[18] T. J. R. Hughes and G. Sangalli, Variational multiscale analysis: The fine-scale Green's function, projection, optimization, localization, and stabilized methods, *SIAM J. Numer. Anal.* **45** (2007), no. 2, 539–557.

[19] R. Kornhuber and H. Yserentant, Numerical homogenization of elliptic multiscale problems by subspace decomposition, preprint (2015).

[20] A. Målqvist and D. Peterseim, Computation of eigenvalues by numerical upscaling, *Numer. Math.* **130** (2014), no. 2, 337–361.

[21] A. Målqvist and D. Peterseim, Localization of elliptic multiscale problems, *Math. Comp.* **83** (2014), no. 290, 2583–2603.
[22] H. Owhadi and L. Zhang, Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast, *Multiscale Model. Simul.* **9** (2011), no. 4, 1373–1398.
[23] H. Owhadi, L. Zhang and L. Berlyand, Polyharmonic homogenization, rough polyharmonic splines and sparse super-localization, *ESAIM Math. Model. Numer. Anal.* **48** (2013), no. 2, 517–552.
[24] C. Pechstein and R. Scheichl, Weighted Poincaré inequalities, *IMA J. Numer. Anal.* **33** (2012), no. 2, 652–686.
[25] D. Peterseim, Composite finite elements for elliptic interface problems, *Math. Comp.* **83** (2014), no. 290, 2657–2674.
[26] D. Peterseim, Eliminating the pollution effect in Helmholtz problems by local subscale correction, preprint (2014), http://arxiv.org/abs/1411.7512.
[27] D. Peterseim, Variational multiscale stabilization and the exponential decay of fine-scale correctors, preprint (2015), http://arxiv.org/abs/1505.07611.
[28] R. Scheichl, P. S. Vassilevski and L. T. Zikatanov, Weak approximation properties of elliptic projections with functional constraints, *Multiscale Model. Simul.* **9** (2011), no. 4, 1677–1699.
[29] R. Scheichl, P. S. Vassilevski and L. T. Zikatanov, Mutilevel methods for elliptic problems with highly varying coefficients on non-aligned coarse grids, *SIAM J. Numer. Anal.* **50** (2012), 1675–1694.