Journal of Scientific Computing (2022) 91:76
https://doi.org/10.1007/s10915-022-01834-7

Multiscale Finite Element Methods for an Elliptic Optimal Control Problem with Rough Coefficients

Susanne C. Brenner[1] · José C. Garay[1] · Li-Yeng Sung[1]

Received: 15 October 2021 / Revised: 6 March 2022 / Accepted: 9 March 2022 /
Published online: 7 May 2022

Abstract
We investigate multiscale finite element methods for an elliptic distributed optimal control problem with rough coefficients. They are based on the (local) orthogonal decomposition methodology of Målqvist and Peterseim.

Keywords Multiscale · Rough coefficients · Elliptic optimal control · Localized orthogonal decomposition · Domain decomposition · P-MINRES

Mathematics Subject Classification 65N30 · 65N15 · 65N55 · 49N10

1 Introduction

Let Ω be a polyhedral domain in $\mathbb{R}^d$ ($d = 1, 2, 3$) and $y_d \in L_2(\Omega)$. We consider the following elliptic distributed optimal control problem:

$$\text{Find} \quad (\bar{y}, \bar{u}) = \underset{(y,u)\in K}{argmin} \frac{1}{2}\big[\|y - y_d\|^2_{L_2(\Omega)} + \gamma\|u\|^2_{L_2(\Omega)}\big], \tag{1.1}$$

where $(y, u) \in H^1_0(\Omega) \times L_2(\Omega)$ belongs to K if and only if

$$a(y, z) = \int_\Omega uz\,dx \qquad \forall\, z \in H^1_0(\Omega), \tag{1.2}$$

This work was supported in part by the National Science Foundation under Grant No. DMS-19-13035.

✉ Susanne C. Brenner
brenner@math.lsu.edu

José C. Garay
jgaray@cct.lsu.edu

Li-Yeng Sung
sung@math.lsu.edu

[1] Department of Mathematics and Center for Computation and Technology, Louisiana State University, Baton Rouge, LA 70803, USA

and the bilinear form $a(\cdot,\cdot)$ is given by

$$a(y,z) = \int_\Omega (\mathcal{A}\nabla y)\cdot\nabla z\, dx. \tag{1.3}$$

Here $\bar{y}$ is the optimal state, $\bar{u}$ is the optimal control and y_d is the desired state.

Remark 1.1 We will follow the standard notation for differential operators, function spaces and norms that can be found for example in [3, 7, 9].

We assume only that the components of the symmetric diffusion matrix $\mathcal{A}$ belong to $L_\infty(\Omega)$ and the eigenvalues of $\mathcal{A}$ are bounded below (resp., above) by the positive constant α (resp., β), which covers many multiscale optimal control problems.

Example 1.2 This example is from [23], where Ω is the unit square $(0,1)\times(0,1)$,

$$\mathcal{A}(x) = \begin{bmatrix} c(x) & 0 \\ 0 & c(x) \end{bmatrix},$$

and

$$c(x) = \frac{2+1.8\sin\left(\frac{2\pi x_1}{\epsilon}\right)}{2+1.8\cos\left(\frac{2\pi x_2}{\epsilon}\right)} + \frac{2+\sin\left(\frac{2\pi x_2}{\epsilon}\right)}{2+1.8\sin\left(\frac{2\pi x_1}{\epsilon}\right)}$$

is highly oscillatory for small ϵ. Note that

$$\alpha = \min_{0\le x\le 1} c(x) \approx 1.248 \quad \text{and} \quad \beta = \max_{0\le x\le 1} c(x) \approx 19.526$$

for any $\epsilon \le 1$.

Example 1.3 This example is from [6], where Ω is the unit square $(0,1)\times(0,1)$,

$$\mathcal{A} = \begin{bmatrix} \mathcal{A}_{11} & 0 \\ 0 & \mathcal{A}_{22} \end{bmatrix},$$

and the components $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ are randomly generated piecewise constant functions with respect to a uniform partition of Ω into 40×40 small squares (cf. Fig. 1). The values of $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ are between $\alpha = 1$ and $\beta = 1331$.

Due to the roughness of the coefficients in (1.3), a standard finite element method can only accurately capture the solution of (1.1) on a very fine mesh (cf. [5]), which can be too expensive, especially when the problem has to be solved repeatedly for different y_d. Our goal is to construct generalized finite element spaces that can produce approximate solutions of (1.1) with $O(H)$ (resp., $O(H^2)$) error in the energy (resp., L_2) norm, where H is the mesh size and the dimensions of the generalized finite element spaces are $O(H^{-d})$. In other words the performance of these generalized finite element methods is similar to standard finite element methods for elliptic problems with smooth coefficients on smooth or convex domains.

Our constructions are based on the Localized Orthogonal Decomposition (LOD) approach in [20, 31] and the ideas in [6, 28] (cf. also [32,Section 4.3]). The basis functions of the generalized finite element spaces are obtained by a correction process that can be carried out offline. The online computation only involves solving a linear system of moderate size. Therefore these generalized finite element methods can also be viewed as reduced order methods that are particularly suitable for repeat solves.

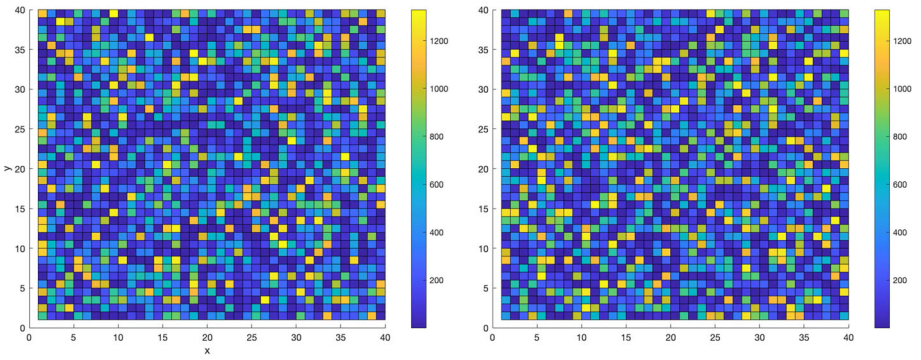

Fig. 1 $\mathcal{A}_{11}$ (left) and $\mathcal{A}_{22}$ (right)

There are many numerical methods for elliptic problems with rough coefficients besides the LOD methods. They include the variational multiscale method (cf. [25–27] and the references therein), the multiscale finite element method (cf. [13, 19, 23, 24] and the references therein), the heterogeneous multiscale method (cf. [1, 2, 11, 12] and the references therein), and the method of approximate component synthesis (cf. [21, 22] and the references therein). We refer the readers to [32, 33] for the discussion of other methods.

On the other hand, as far as we know, there is only one paper [15] that solved the optimal control problem (1.1)–(1.3) (with additional control constraints) by the heterogeneous multiscale method, where scale separation and periodic structure are assumed. In the context of parabolic optimal control problems with rough coefficients, reduced order finite element methods in the same spirit of the current paper are treated in [30, 36]. In particular, the methodology in [30] is also based on the LOD approach. The distinctive feature of our work in this paper is that the construction of the localized multiscale finite element space and its analysis are based entirely on classical techniques from domain decomposition and numerical linear algebra.

The rest of the paper is organized as follows. We recall relevant results for the optimal control problem in Sect. 2. A multiscale finite element method based on orthogonal decomposition is treated in Sect. 3. We introduce a localized multiscale finite element space in Sect. 4 and analyze the corresponding Galerkin method in Sect. 5, where the error estimates in Sect. 3 play a useful role. Numerical results are presented in Sect. 6 and we end with some concluding remarks in Sect. 7.

We will use $\langle \cdot, \cdot \rangle$ to denote the canonical bilinear form on a finite dimensional vector space V and its dual space V'. A linear operator $L : V \longrightarrow V'$ is symmetric if

$$\langle Lv_1, v_2 \rangle = \langle Lv_2, v_1 \rangle \qquad \forall\, v_1, v_2 \in V,$$

and it is symmetric positive definite (SPD) if additionally

$$\langle Lv, v \rangle > 0 \qquad \forall\, v \in V \setminus \{0\}.$$

Given two finite dimensional vector spaces V and W and a linear transform $T : V \longrightarrow W$, the transpose $T^t : W' \longrightarrow V'$ is defined by

$$\langle T^t \mu, v \rangle = \langle \mu, Tv \rangle \qquad \forall\, \mu \in W',\ v \in V.$$

We also assume that all the unspecified positive constants in the paper are greater than or equal to 1.

2 The Continuous Problem

By a standard result [29,Section 2.2], the convex minimization problem defined by (1.1)–(1.3) has a unique solution determined by the following first order optimality conditions:

$$\begin{aligned} a(\bar{y}, z) &= \int_\Omega \bar{u} z \, dx && \forall\, z \in H_0^1(\Omega), \\ a(\bar{p}, q) &= \int_\Omega (y_d - \bar{y}) q \, dx && \forall\, q \in H_0^1(\Omega), \\ \bar{p} &= \gamma \bar{u}, \end{aligned}$$

where the adjoint state $\bar{p}$ belongs to $H_0^1(\Omega)$.

After eliminating $\bar{u}$, we have the following system for $(\bar{p}, \bar{y})$:

$$a(\bar{p}, q) + \int_\Omega \bar{y} q \, dx = \int_\Omega y_d q \, dx \qquad \forall\, q \in H_0^1(\Omega), \tag{2.1}$$

$$\int_\Omega \bar{p} z \, dx - \gamma a(\bar{y}, z) = 0 \qquad \forall\, z \in H_0^1(\Omega). \tag{2.2}$$

Remark 2.1 Note that (2.1)–(2.2) is equivalent to

$$\begin{aligned} \tilde{a}(\bar{p}, q) + \int_\Omega \tilde{y} q \, dx &= \int_\Omega \tilde{y}_d q \, dx && \forall\, q \in H_0^1(\Omega), \\ \int_\Omega \bar{p} z \, dx - \tilde{\gamma} \tilde{a}(\tilde{y}, z) &= 0 && \forall\, z \in H_0^1(\Omega), \end{aligned}$$

where $\tilde{a}(\cdot, \cdot) = \tau a(\cdot, \cdot)$, $\tilde{y} = \tau \bar{y}$, $\tilde{y}_d = \tau y_d$, $\tilde{\gamma} = (\gamma/\tau^2)$ and τ is any positive number. Therefore we can assume that the lower bound α for the eigenvalues of $\mathcal{A}$ (cf. (1.3)) in the definition of the bilinear form $a(\cdot, \cdot)$ in (2.1)–(2.2) is roughly 1, as in Examples 1.2 and 1.3.

Since the dependence on γ is not our main concern here, we will take γ to be 1 in (2.2). We will also simplify the notation by dropping the bars over p and y and consider the problem of finding $(p, y) \in H_0^1(\Omega) \times H_0^1(\Omega)$ such that

$$a(p, q) + \int_\Omega y q \, dx = \int_\Omega y_d q \, dx \qquad \forall\, q \in H_0^1(\Omega), \tag{2.3}$$

$$\int_\Omega p z \, dx - a(y, z) = 0 \qquad \forall\, z \in H_0^1(\Omega). \tag{2.4}$$

We can write (2.3)–(2.4) concisely as

$$\mathcal{B}\big((p, y), (q, z)\big) = \int_\Omega y_d q \, dx \qquad \forall\, (q, z) \in H_0^1(\Omega) \times H_0^1(\Omega), \tag{2.5}$$

where

$$\mathcal{B}\big((p, y), (q, z)\big) = a(p, q) + \int_\Omega y q \, dx + \int_\Omega p z \, dx - a(y, z). \tag{2.6}$$

We will use $\|\cdot\|_a$ to denote the energy norm $\sqrt{a(\cdot, \cdot)}$. Note that

$$\sqrt{\alpha}\, |v|_{H^1(\Omega)} \le \|v\|_a \le \sqrt{\beta}\, |v|_{H^1(\Omega)} \qquad \forall\, v \in H^1(\Omega) \tag{2.7}$$

by our assumption on $\mathcal{A}$, and we have a Poincaré-Friedrichs inequality [3]

$$\|v\|_{L_2(\Omega)}^2 \le C_{\mathrm{PF}} |v|_{H^1(\Omega)}^2 \qquad \forall\, v \in H_0^1(\Omega). \tag{2.8}$$

The following are the salient features of $\mathcal{B}(\cdot,\cdot)$ that follow immediately from (2.6)–(2.8) and the Cauchy-Schwarz inequality:

$$\mathcal{B}\big((q,z),(q,-z)\big) = \|(q,z)\|_{a\times a}^2 \quad \forall\,(q,z) \in H_0^1(\Omega)\times H_0^1(\Omega), \tag{2.9}$$

and

$$\mathcal{B}\big((q,z),(r,s)\big) \le [1 + (C_{\mathrm{PF}}/\alpha)]\|(q,z)\|_{a\times a}\|(r,s)\|_{a\times a} \tag{2.10}$$

for all $(q,z),(r,s) \in H_0^1(\Omega)\times H_0^1(\Omega)$, where the norm $\|\cdot\|_{a\times a}$ is defined by

$$\|(q,z)\|_{a\times a}^2 = \|q\|_a^2 + \|z\|_a^2. \tag{2.11}$$

From here on we will also use boldfaced letters to denote members of the product space $H_0^1(\Omega)\times H_0^1(\Omega)$ in order to improve the readability of the formulas.

Lemma 2.2 *Let V be a subspace of $H_0^1(\Omega)$. We have*

$$\inf_{\boldsymbol{v}\in V\times V} \sup_{\boldsymbol{w}\in V\times V} \frac{\mathcal{B}(\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{v}\|_{a\times a}\|\boldsymbol{w}\|_{a\times a}} \ge 1. \tag{2.12}$$

Proof Let $\boldsymbol{v} = (q,z) \in V\times V$ be arbitrary. According to (2.9), we have

$$\|(q,z)\|_{a\times a}^2 = \mathcal{B}\big((q,z),(q,-z)\big)$$

and consequently

$$\|\boldsymbol{v}\|_{a\times a} = \frac{\mathcal{B}\big((q,z),(q,-z)\big)}{\|(q,z)\|_{a\times a}} = \frac{\mathcal{B}\big((q,z),(q,-z)\big)}{\|(q,-z)\|_{a\times a}} \le \sup_{\boldsymbol{w}\in V\times V} \frac{\mathcal{B}(\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{w}\|_{a\times a}}. \tag{2.13}$$

□

Remark 2.3 Let V be a closed subspace of $H_0^1(\Omega)$ and $\boldsymbol{V} = V\times V$. It follows from (2.10) that we can define a linear transformation $T : \boldsymbol{V} \longrightarrow \boldsymbol{V}'$ by

$$\langle T\boldsymbol{z},\boldsymbol{w}\rangle = \mathcal{B}(\boldsymbol{z},\boldsymbol{w}) \qquad \forall\,\boldsymbol{z},\boldsymbol{w} \in \boldsymbol{V}.$$

Since the bilinear form $\mathcal{B}(\cdot,\cdot)$ is symmetric, the inf-sup condition (2.12) implies that T is an isomorphism and the operator norms of T and T^{-1} (with respect to $\|\cdot\|_{a\times a}$) are bounded by 1 (cf. [4, 8]).

Remark 2.4 In view of Remark 2.3, one can solve (2.5) by a standard finite element method. Let $V_h \subset H_0^1(\Omega)$ (resp., $V_H \subset H_0^1(\Omega)$) be the P_1 or Q_1 finite element space associated with the triangulation $\mathcal{T}_h$ (resp., $\mathcal{T}_H$) of Ω, where $\mathcal{T}_h$ is a refinement of $\mathcal{T}_H$ and hence V_H is a subspace of V_h.

We assume that $h \ll 1$ so that $(p_h, y_h) \in V_h \times V_h$ determined by

$$\mathcal{B}\big((p_h,y_h),(q,z)\big) = \int_\Omega y_d q\,dx \qquad \forall\,(q,z) \in V_h\times V_h \tag{2.14}$$

provides a good approximation of the solution (p,y) of (2.5), but the dimension of V_h is so large that the computational cost is prohibitive, especially if we have to solve (2.14) repeatedly for different y_d.

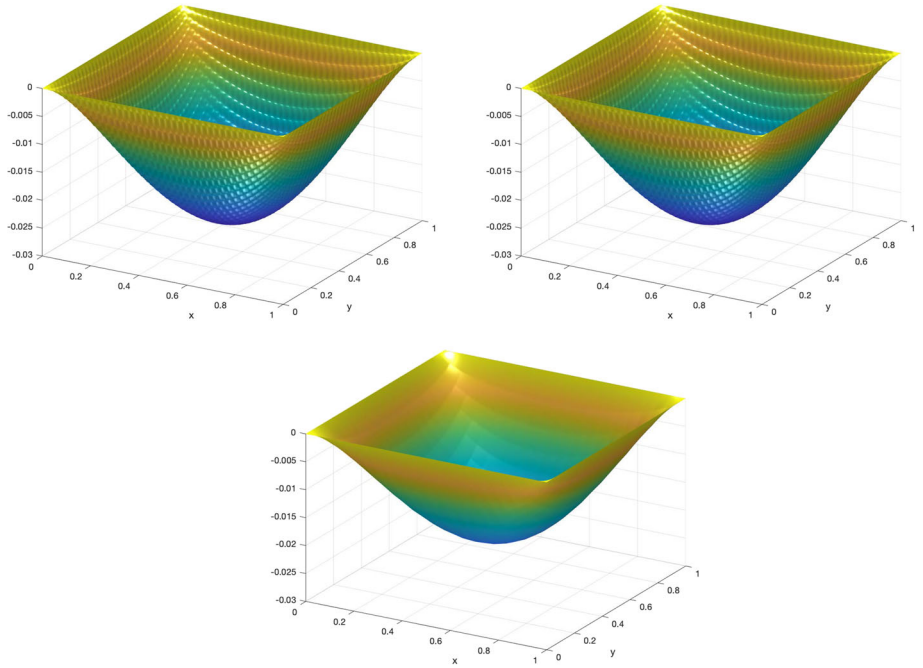

Fig. 2 Finite element solutions of the optimal control in Example 1.2 with $\epsilon = 0.025$: solution on a fine mesh with $h = 1/320$ (top left), solution on a coarse mesh with $H = 1/20$ (bottom), LOD solution with $H = 1/20$ and $h = 1/320$ (top right)

On the other hand, for $H \gg h$, the solution $(p_H, y_H) \in V_H \times V_H$ defined by

$$\mathcal{B}\big((p_H, y_H), (q, z)\big) = \int_\Omega y_d q \, dx \qquad \forall\, (q, z) \in V_H \times V_H \tag{2.15}$$

is computationally feasible but not sufficiently accurate. Therefore we need generalized finite element spaces to bridge the two scales.

Finite element solutions for the optimal controls in Examples 1.2 and 1.3 are displayed in Figs. 2 and 3. It can be observed for both examples that the LOD solutions from Sect. 5 capture the fine scale solutions while the coarse scale solutions fail to do so.

Remark 2.5 It follows from (2.7), (2.8), (2.12) and (2.14) that

$$\|(p_h, y_h)\|_{a\times a} \leq \|y_d\|_{L_2(\Omega)} \sup_{(q,z)\in V_h\times V_h} \frac{\|q\|_{L_2(\Omega)}}{\|(q,z)\|_{a\times a}} \leq \sqrt{C_{\mathrm{PF}}/\alpha}\, \|y_d\|_{L_2(\Omega)}.$$

3 The Ideal Multiscale Finite Element Method

In this section we construct and analyze the ideal multiscale finite element method following the ideas in [28, 31], which begins with the construction of a projection operator. We will denote by n (resp., m) the dimension of the finite element space V_h (resp., V_H) in Remark 2.4.

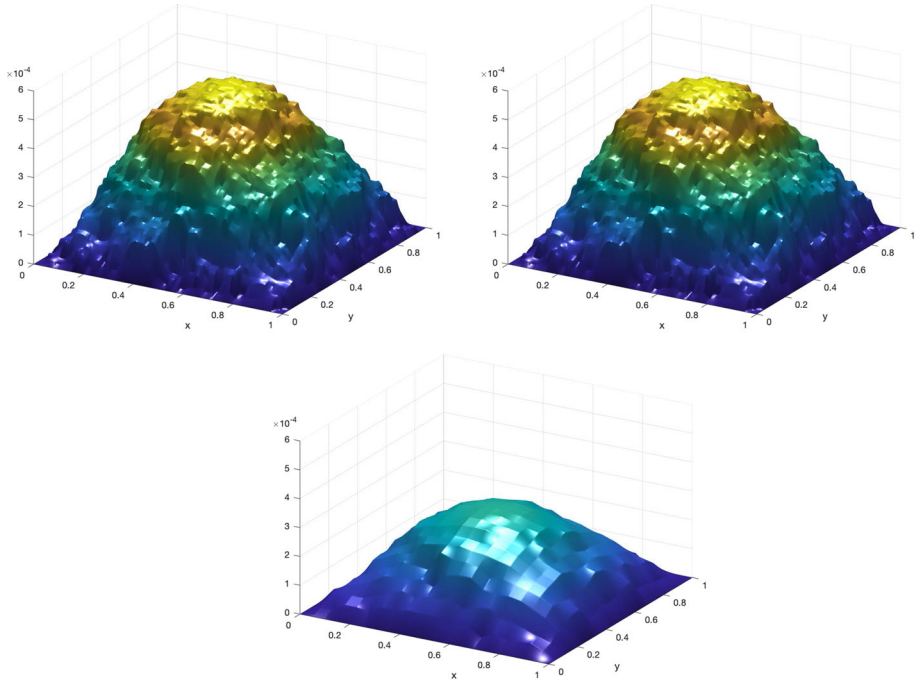

Fig. 3 Finite element solutions of the optimal control in Example 1.3: solution on a fine mesh with $h = 1/320$ (top left), solution on a coarse mesh with $H = 1/20$ (bottom), LOD solution with $H = 1/20$ and $h = 1/320$ (top right)

3.1 The Projection Operator Π_H

The operator $\Pi_H : H_0^1(\Omega) \longrightarrow V_H$ is defined by taking the nodal average of the local L_2 orthogonal projections of $\zeta \in H_0^1(\Omega)$ onto P_1 or Q_1 polynomials. More precisely, we define $\Pi_H \zeta$ by

$$(\Pi_H \zeta)(p) = \frac{1}{|\mathcal{T}_p|} \sum_{T \in \mathcal{T}_p} (Q_T \zeta_T)(p) \qquad \forall\, p \in \mathcal{V}_H, \tag{3.1}$$

where $\mathcal{V}_H$ is the set of all the (interior) vertices of $\mathcal{T}_H$, $\mathcal{T}_p$ is the set of the elements in $\mathcal{T}_H$ that share p as a common vertex, $|\mathcal{T}_p|$ is the number of elements in $\mathcal{T}_p$, ζ_T is the restriction of ζ to T, and Q_T is the orthogonal projection from $L_2(T)$ onto $P_1(T)$ or $Q_1(T)$.

We have an obvious relation

$$\Pi_H v = v \qquad \forall\, v \in V_H \tag{3.2}$$

and also an interpolation error estimate [6,Appendix A]

$$H^{-1}\|v - \Pi_H v\|_{L_2(\Omega)} + |\Pi_H v|_{H^1(\Omega)} \le C_\dagger |v|_{H^1(\Omega)} \qquad \forall\, v \in H_0^1(\Omega), \tag{3.3}$$

where the positive constant $C_\dagger$ depends only on the shape regularity of $\mathcal{T}_H$.

Remark 3.1 We can use (2.7) to translate the estimate for $|\Pi_H v|_{H^1(\Omega)}$ into

$$\|\Pi_H v\|_a \le C_\dagger \sqrt{\beta/\alpha}\, \|v\|_a \qquad \forall\, v \in H_0^1(\Omega).$$

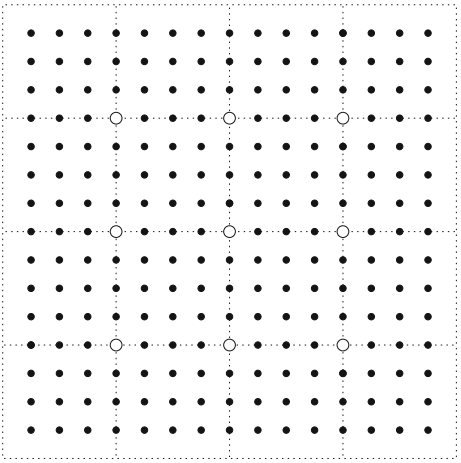

Fig. 4 The nodes for V_H are represented by the circles and the nodes for $\varphi_1, \ldots, \varphi_\ell \in V_h$ are represented by the solid dots

We will denote the kernel of the restriction of Π_H to V_h by $K_h^{\Pi_H}$, i.e.,

$$K_h^{\Pi_H} = \{v \in V_h : \Pi_H v = 0\}. \tag{3.4}$$

It follows from (3.2) and (3.4) that

$$\dim K_h^{\Pi_H} = \dim V_h - \dim V_H = n - m. \tag{3.5}$$

A basis for $K_h^{\Pi_H}$ is given in the following lemma.

Lemma 3.2 *Let $\ell = n - m$ and $\varphi_1, \ldots, \varphi_\ell$ be the nodal basis functions in V_h that vanish at the nodes of V_H (cf. Fig. 4 for a two dimensional example with the Q_1 finite element). Then $(I - \Pi_H)\varphi_1, \ldots, (I - \Pi_H)\varphi_\ell$ form a basis of $K_h^{\Pi_H}$, where I is the identity operator on V_h.*

Proof It follows from (3.2) that $(I - \Pi_H)\varphi_i \in K_h^{\Pi_H}$ for $1 \le i \le \ell$. In view of (3.5), it only remains to show that the functions $(I - \Pi_H)\varphi_1, \ldots, (I - \Pi_H)\varphi_\ell$ are linearly independent.

Suppose $\sum_{i=1}^{\ell} c_i (I - \Pi_H)\varphi_i = 0$. Then the function $\sum_{i=1}^{\ell} c_i \varphi_i = \sum_{i=1}^{\ell} c_i \Pi_H \varphi$ belongs to V_H and at the same time vanishes at the nodes of V_H. It follows that $\sum_{i=1}^{\ell} c_i \varphi_i = 0$ and hence $c_i = 0$ for $1 \le i \le \ell$ because the functions $\varphi_1, \ldots, \varphi_\ell$ are linearly independent. □

3.2 The Projection Operator $\mathcal{C}_h^{\Pi_H}$

According to Remark 2.3, we can define a linear transformation

$$\mathcal{C}_h^{\Pi_H} : V_h \times V_h \longrightarrow K_h^{\Pi_H} \times K_h^{\Pi_H}$$

by

$$\mathcal{B}\big(\mathcal{C}_h^{\Pi_H} \boldsymbol{v}, \boldsymbol{w}\big) = \mathcal{B}\big(\boldsymbol{v}, \boldsymbol{w}\big) \qquad \forall\, \boldsymbol{v} \in V_h \times V_h,\ \boldsymbol{w} \in K_h^{\Pi_H} \times K_h^{\Pi_H}. \tag{3.6}$$

The elementary algebraic properties of $\mathcal{C}_h^{\Pi_H}$ that follow directly from (3.6) are collected in the following lemma.

Lemma 3.3 *We have*

$$\mathcal{B}\big(\mathcal{C}_h^{\Pi_H}\boldsymbol{v},\boldsymbol{w}\big)=\mathcal{B}\big(\boldsymbol{v},\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\big) \qquad \forall\,\boldsymbol{v},\boldsymbol{w}\in V_h\times V_h, \tag{3.7}$$

$$(\Pi_H\times\Pi_H)\mathcal{C}_h^{\Pi_H}\boldsymbol{v}=0 \qquad \forall\,\boldsymbol{v}\in V_h\times V_h, \tag{3.8}$$

$$\mathcal{C}_h^{\Pi_H}\boldsymbol{v}=\boldsymbol{v} \qquad \forall\,\boldsymbol{v}\in K_h^{\Pi_H}\times K_h^{\Pi_H}. \tag{3.9}$$

Remark 3.4 It follows from (3.9) that $\mathcal{C}_h^{\Pi_H}$ is a projection from $V_h\times V_h$ onto $K_h^{\Pi_H}\times K_h^{\Pi_H}$, and that

$$(\boldsymbol{I}-\mathcal{C}_h^{\Pi_H})(\boldsymbol{I}-\Pi_H\times\Pi_H)\boldsymbol{v}=0 \qquad \forall\,\boldsymbol{v}\in V_h\times V_h, \tag{3.10}$$

where $\boldsymbol{I}$ is the identity operator on $V_h\times V_h$.

Lemma 3.5 *We have*

$$\|\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a}\le[1+(\mathrm{C_{PF}}/\alpha)]\|\boldsymbol{v}\|_{a\times a} \qquad \forall\,\boldsymbol{v}\in V_h\times V_h.$$

Proof Let $(q,z)=\mathcal{C}_h^{\Pi_H}\boldsymbol{v}$. Then (q,z) (and hence $(q,-z)$) belongs to $K_h^{\Pi_H}\times K_h^{\Pi_H}$. It follows from (2.9), (2.10) and (3.6) that

$$\begin{aligned}\|\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a}^2=\|(q,z)\|_{a\times a}^2&=\mathcal{B}((q,z),(q,-z))\\&=\mathcal{B}(\boldsymbol{v},(q,-z))\le[1+(\mathrm{C_{PF}}/\alpha)]\|\boldsymbol{v}\|_{a\times a}\|\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a}.\end{aligned}$$

□

Corollary 3.6 *The following relations are valid:*

$$\|\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a}\le[2+(\mathrm{C_{PF}}/\alpha)]\|\boldsymbol{v}\|_{a\times a} \qquad \forall\,\boldsymbol{v}\in V_h\times V_h, \tag{3.11}$$

$$\|\boldsymbol{v}\|_{a\times a}\le C_\dagger\sqrt{\beta/\alpha}\|\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a} \qquad \forall\,\boldsymbol{v}\in V_H\times V_H. \tag{3.12}$$

Proof The inequality (3.11) follows from Lemma 3.5 and the triangle inequality, and the inequality (3.12) follows from (3.2), Remark 3.1 and (3.8) :

$$\|\boldsymbol{v}\|_{a\times a}=\|(\Pi_H\times\Pi_H)(\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v})\|_{a\times a}\le C_\dagger\sqrt{\beta/\alpha}\,\|\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\,\|_{a\times a}\,.$$

□

3.3 The Finite Element Space $V_H^{\mathrm{ms,h}}$

The ideal multiscale finite element space

$$V_H^{\mathrm{ms,h}}\subset V_h\times V_h$$

is defined by

$$V_H^{\mathrm{ms,h}}=\{\boldsymbol{v}\in V_h\times V_h:\ \mathcal{B}\big(\boldsymbol{v},\boldsymbol{w}\big)=0\ \ \forall\,\boldsymbol{w}\in K_h^{\Pi_H}\times K_h^{\Pi_H}\}. \tag{3.13}$$

Let $\boldsymbol{v}\in V_h\times V_h$ be arbitrary. It follows from Lemma 2.2 (with $V=K_h^{\Pi_H}$), (3.6) and (3.13) that

$$\boldsymbol{v}\in V_H^{\mathrm{ms,h}}\quad\Leftrightarrow\quad\mathcal{B}(\mathcal{C}_h^{\Pi_H}\boldsymbol{v},\boldsymbol{w})=0\ \ \forall\,\boldsymbol{w}\in K_h^{\Pi_H}\times K_h^{\Pi_H}\quad\Leftrightarrow\quad\mathcal{C}_h^{\Pi_H}\boldsymbol{v}=0. \tag{3.14}$$

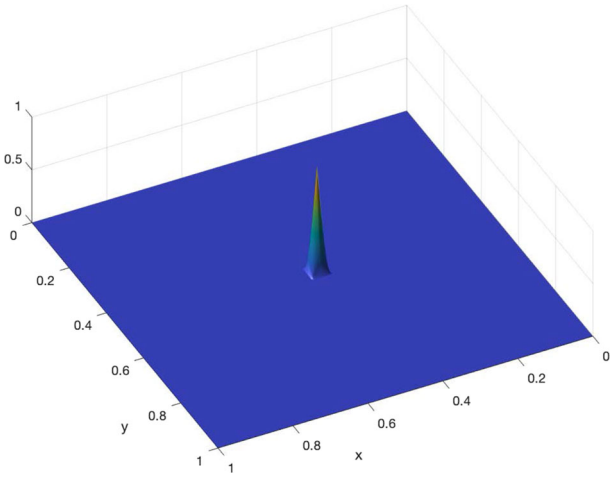

Fig. 5 A basis function ϕ_i of V_H

Therefore we have $V_H^{\mathrm{ms,h}} = (\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(V_h \times V_h)$ and

$$\dim V_H^{\mathrm{ms,h}} = 2n - 2\ell = 2m. \tag{3.15}$$

A basis for $V_H^{\mathrm{ms,h}}$ is given in the following lemma.

Lemma 3.7 *Let* $\{\phi_1, \ldots, \phi_m\}$ *be the nodal basis of* V_H, $\boldsymbol{\psi}_i = \mathcal{C}_h^{\Pi_H}(\phi_i, 0)$ *and* $\boldsymbol{\xi}_i = \mathcal{C}_h^{\Pi_H}(0, \phi_i)$. *Then*

$$\{(\phi_1, 0) - \boldsymbol{\psi}_1, \ldots, (\phi_m, 0) - \boldsymbol{\psi}_m, (0, \phi_1) - \boldsymbol{\xi}_1, \ldots, (0, \phi_m) - \boldsymbol{\xi}_m\}$$

is a basis for $V_H^{\mathrm{ms,h}}$.

Proof In view of (3.15), it suffices to show that the $2m$ functions $(\phi_i, 0) - \boldsymbol{\psi}_i = (\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(\phi_i, 0)$ and $(0, \phi_i) - \boldsymbol{\xi}_i = (\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(0, \phi_i)$ $(1 \le i \le m)$ are linearly independent.

Suppose $\sum_{i=1}^m \left[c_i(\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(\phi_i, 0) + d_i(\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(0, \phi_i)\right] = \boldsymbol{0}$. It then follows from (3.2) and (3.8) that $\sum_{i=1}^m \left[c_i(\phi_i, 0) + d_i(0, \phi_i)\right] = \boldsymbol{0}$ and hence $c_i = d_i = 0$ for $1 \le i \le m$. □

Remark 3.8 It follows from Lemma 3.7 that we also have $V_H^{\mathrm{ms,h}} = (\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(V_H \times V_H)$.

Remark 3.9 Let $(\psi_{i,1}, \psi_{i,2}) = \boldsymbol{\psi}_i = \mathcal{C}_h^{\Pi_H}(\phi_i, 0)$ and $(\xi_{i,1}, \xi_{i,2}) = \boldsymbol{\xi}_i = \mathcal{C}_h^{\Pi_H}(0, \phi_i)$. It follows from (3.6) and the relation (cf. (2.6))

$$\mathcal{B}\big((y, -p), (q, z)\big) = \mathcal{B}\big((p, y), (z, -q)\big) \qquad \forall\, (p, y), (q, z) \in H_0^1(\Omega) \times H_0^1(\Omega)$$

that $\psi_{i,1} = \xi_{i,2}$ and $\psi_{i,2} = -\xi_{i,1}$.

The figure of a typical basis function ϕ_i of V_H is given in Fig. 5, and the figure of the corresponding basis function $(\phi_i, 0) - \mathcal{C}_h^{\Pi_H}(\phi_i, 0)$ is displayed in Fig. 6 for Example 1.2, and in Fig. 7 for Example 1.3.

Remark 3.10 The exponential decay of $\boldsymbol{\psi}_i = \mathcal{C}_h^{\Pi_H}(\phi_i, 0)$ (and hence $\boldsymbol{\xi}_i = \mathcal{C}_h^{\Pi_H}(0, \phi_i)$ in view of Remark 3.9) are clearly observed in Figs. 6 and 7.

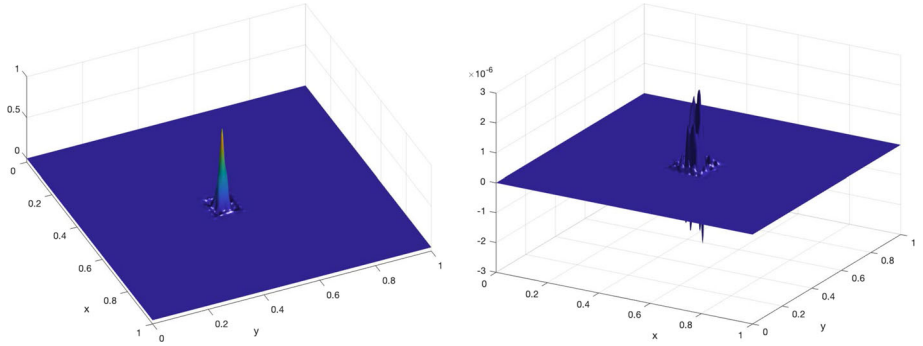

Fig. 6 The basis function $(\phi_i, 0) - \mathcal{C}_h^{\Pi_H}(\phi_i, 0)$ of $V_H^{\mathrm{ms,h}}$ for Example 1.2 with $H = 1/40$, $h = 1/320$ and $\epsilon = 0.0025$: first component (left) and second component (right)

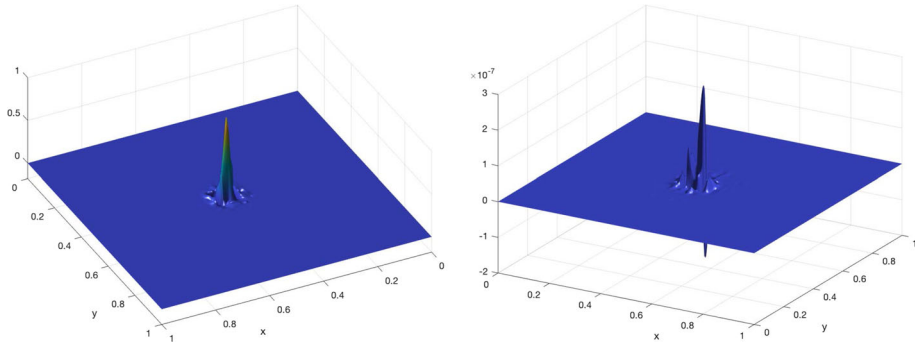

Fig. 7 The basis function $(\phi_i, 0) - \mathcal{C}_h^{\Pi_H}(\phi_i, 0)$ of $V_H^{\mathrm{ms,h}}$ for Example 1.3 with $H = 1/40$ and $h = 1/320$: first component (left) and second component (right)

3.4 The Discrete Problem

The approximate solution $(p_H^{\mathrm{ms},h}, y_H^{\mathrm{ms},h}) \in V_H^{\mathrm{ms,h}}$ is defined by

$$\mathcal{B}\big((p_H^{\mathrm{ms},h}, y_H^{\mathrm{ms},h}), (q, z)\big) = \int_\Omega y_d q \, dx \qquad \forall\, (q, z) \in V_H^{\mathrm{ms,h}}. \tag{3.16}$$

The well-posedness of (3.16) is guaranteed by the following lemma.

Lemma 3.11 *We have*

$$\inf_{\boldsymbol{v} \in V_H^{\mathrm{ms,h}}} \sup_{\boldsymbol{w} \in V_H^{\mathrm{ms,h}}} \frac{\mathcal{B}(\boldsymbol{v}, \boldsymbol{w})}{\|\boldsymbol{v}\|_{a\times a} \|\boldsymbol{w}\|_{a\times a}} \geq [2 + (\mathrm{C_{PF}}/\alpha)]^{-1}. \tag{3.17}$$

Proof Let $\boldsymbol{v} = (q, z) \in V_H^{\mathrm{ms,h}}$ be arbitrary. Then $(q, -z) - \mathcal{C}_h^{\Pi_H}(q, -z) \in V_H^{\mathrm{ms,h}}$ and it follows from (2.9), (3.11) and (3.13) that

$$\begin{aligned} \|\boldsymbol{v}\|_{a\times a} &= \frac{\mathcal{B}\big((q, z), (q, -z)\big)}{\|(q, -z)\|_{a\times a}} \\ &= \frac{\mathcal{B}\big((q, z), (q, -z) - \mathcal{C}_h^{\Pi_H}(q, -z)\big)}{\|(q, -z)\|_{a\times a}} \end{aligned}$$

$$\leq [2 + (\mathrm{C_{PF}}/\alpha)]\left(\frac{\mathcal{B}\big((q,z),(q,-z)-\mathcal{C}_h^{\Pi_H}(q,-z)\big)}{\|(q,-z)-\mathcal{C}_h^{\Pi_H}(q,-z)\|_{a\times a}}\right)$$
$$\leq [2 + (\mathrm{C_{PF}}/\alpha)] \sup_{\boldsymbol{w}\in V_H^{\mathrm{ms,h}}} \frac{\mathcal{B}(\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{w}\|_{a\times a}}$$

□

3.5 Energy Error

It follows from (2.14) and (3.16) that

$$\mathcal{B}\big((p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h}),(q,z)\big)=0 \qquad \forall\,(q,z)\in V_H^{\mathrm{ms,h}}. \tag{3.18}$$

We will use the Galerkin relation (3.18) to derive an error estimate for the ideal multiscale finite element method defined by (3.16).

Theorem 3.12 *We have*

$$\|(p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h})\|_{a\times a}\leq (C_\dagger/\sqrt{\alpha})H\|y_d\|_{L_2(\Omega)}. \tag{3.19}$$

Proof In view of Remark 2.3, (3.13) and (3.18), we have $(p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h})\in K_h^{\Pi_H}\times K_h^{\Pi_H}$ and consequently

$$p_h-p_H^{\mathrm{ms},h} \quad\text{and}\quad y_h-y_H^{\mathrm{ms},h} \quad\text{belong to}\quad K_h^{\Pi_H}. \tag{3.20}$$

Putting (2.7), (2.9), (2.14), (3.3), (3.13) and (3.20) together, we have

$$\begin{aligned}\|(p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h})\|_{a\times a}^2&=\mathcal{B}\big((p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h}),(p_h,-y_h)-(p_H^{\mathrm{ms},h},-y_H^{\mathrm{ms},h})\big)\\&=\mathcal{B}\big((p_h,y_h),(p_h,-y_h)-(p_H^{\mathrm{ms},h},-y_H^{\mathrm{ms},h})\big)\\&=\int_\Omega y_d(p_h-p_H^{\mathrm{ms},h})dx\\&=\int_\Omega y_d[(p_h-p_H^{\mathrm{ms},h})-\Pi_H(p_h-p_H^{\mathrm{ms},h})]dx\\&\leq C_\dagger H\|y_d\|_{L_2(\Omega)}|p_h-p_H^{\mathrm{ms},h}|_{H^1(\Omega)}\\&\leq (C_\dagger/\sqrt{\alpha})H\|y_d\|_{L_2(\Omega)}\|p_h-p_H^{\mathrm{ms},h}\|_a,\end{aligned}$$

and (3.19) follows immediately. □

Remark 3.13 In view of (3.9), (3.14) and (3.20), we can express the error of the ideal multiscale finite element method as

$$(p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h})=\mathcal{C}_h^{\Pi_H}((p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h}))=\mathcal{C}_h^{\Pi_H}(p_h,y_h).$$

3.6 L_2 Error

We will obtain an estimate for the L_2 error by a duality argument.

Theorem 3.14 *We have*

$$\|(p_h,y_h)-(p_H^{\mathrm{ms},h},y_H^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)}\leq [1+(\mathrm{C_{PF}}/\alpha)](C_\dagger^2/\alpha)H^2\|y_d\|_{L_2(\Omega)}. \tag{3.21}$$

Proof In view of Remark 2.3, we can define $(q, z) \in V_h \times V_h$ by

$$\mathcal{B}\big((q,z),(r,s)\big) = \int_\Omega (p_h - p_H^{\mathrm{ms},h}) r \, dx + \int_\Omega (y_h - y_H^{\mathrm{ms},h}) s \, dx \qquad \forall\, (r,s) \in V_h \times V_h. \tag{3.22}$$

Let $\mathcal{C}_h^{\Pi_H}(q,z) \in K_h^{\Pi_H} \times K_h^{\Pi_H}$ be denoted by $(\tilde{q}, \tilde{z})$. It follows from (2.7), (2.9), (3.3), (3.6), (3.7), (3.9), (3.22) and the Cauchy-Schwarz inequality that

$$\begin{aligned}
\|(\tilde{q},\tilde{z})\|_{a\times a}^2 &= \mathcal{B}\big((\tilde{q},\tilde{z}),(\tilde{q},-\tilde{z})\big) \\
&= \mathcal{B}\big((q,z),(\tilde{q},-\tilde{z})\big) \\
&= \int_\Omega (p_h - p_H^{\mathrm{ms},h})\tilde{q}\, dx - \int_\Omega (y_h - y_H^{\mathrm{ms},h})\tilde{z}\, dx \\
&= \int_\Omega (p_h - p_H^{\mathrm{ms},h})(\tilde{q} - \Pi_H \tilde{q}) dx - \int_\Omega (y_h - y_H^{\mathrm{ms},h})(\tilde{z} - \Pi_H \tilde{z}) dx \\
&\le C_\dagger H \big(\|p_h - p_H^{\mathrm{ms},h}\|_{L_2(\Omega)} |\tilde{q}|_{H^1(\Omega)} + \|y_h - y_H^{\mathrm{ms},h}\|_{L_2(\Omega)} |\tilde{z}|_{H^1(\Omega)} \big) \\
&\le C_\dagger H \|(p_h - p_H^{\mathrm{ms},h}, y_h - y_H^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)} (1/\sqrt{\alpha}) \|(\tilde{q},\tilde{z})\|_{a\times a},
\end{aligned}$$

and hence

$$\|\mathcal{C}_h^{\Pi_H}(q,z)\|_{a\times a} \le (C_\dagger/\sqrt{\alpha}) H \|(p_h, y_h) - (p_H^{\mathrm{ms},h}, y_H^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)}. \tag{3.23}$$

On the other hand, we have

$$\begin{aligned}
\|(p_h - p_H^{\mathrm{ms},h}, y_h - y_H^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)}^2 &= \mathcal{B}\big((q,z),(p_h - p_H^{\mathrm{ms},h}, y_h - y_H^{\mathrm{ms},h})\big) \\
&= \mathcal{B}\big(\mathcal{C}_h^{\Pi_H}(q,z),(p_h - p_H^{\mathrm{ms},h}, y_h - y_H^{\mathrm{ms},h})\big) \\
&\le [1 + (\mathrm{C_{PF}}/\alpha)] \|\mathcal{C}_h^{\Pi_H}(q,z)\|_{a\times a} \|(p_h - p_H^{\mathrm{ms},h}, y_h - y_H^{\mathrm{ms},h})\|_{a\times a}
\end{aligned} \tag{3.24}$$

by (2.10), (3.6), (3.20) and (3.22).

Putting (3.19), (3.23) and (3.24) together, we arrive at

$$\begin{aligned}
&\|(p_h, y_h) - (p_H^{\mathrm{ms},h}, y_H^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)} \\
&\qquad \le [1 + (\mathrm{C_{PF}}/\alpha)](C_\dagger/\sqrt{\alpha}) H \|(p_h - p_H^{\mathrm{ms},h}, y_h - y_H^{\mathrm{ms},h})\|_{a\times a} \\
&\qquad \le [1 + (\mathrm{C_{PF}}/\alpha)](C_\dagger^2/\alpha) H^2 \|y_d\|_{L_2(\Omega)}.
\end{aligned}$$

□

4 A Localized Multiscale Finite Element Space

The constructions of $\boldsymbol{\psi}_i = \mathcal{C}_h^{\Pi_H}(\phi_i, 0) \in K_h^{\Pi_H} \times K_h^{\Pi_H}$ and $\boldsymbol{\xi}_i = \mathcal{C}_h^{\Pi_H}(0, \phi_i) \in K_h^{\Pi_H} \times K_h^{\Pi_H}$ require solving the equations

$$\mathcal{B}\big(\boldsymbol{\psi}_i, (q,z)\big) = \mathcal{B}\big((\phi_i, 0), (q,z)\big) \qquad \forall\, (q,z) \in K_h^{\Pi_H} \times K_h^{\Pi_H}, \tag{4.1}$$

$$\mathcal{B}\big(\boldsymbol{\xi}_i, (q,z)\big) = \mathcal{B}\big((0, \phi_i), (q,z)\big) \qquad \forall\, (q,z) \in K_h^{\Pi_H} \times K_h^{\Pi_H}, \tag{4.2}$$

which are expensive. However the exponential decays of $\boldsymbol{\psi}_i$ and $\boldsymbol{\xi}_i$ observed in Figs. 7, 6 and Remark 3.10 indicate that it is possible to capture $\boldsymbol{\psi}_i$ and $\boldsymbol{\xi}_i$ by local approximations. (Note that in practice we only need to solve one of these equations because of the observation in Remark 3.9.)

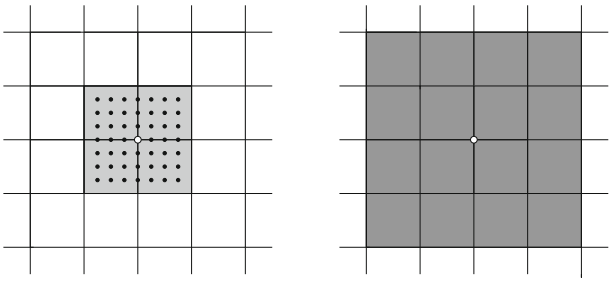

Fig. 8 The patches ω_{x_i} (left) and $\tilde{\omega}_{x_i}$ (right), the node x_i is represented by the circle and the nodes for $\varphi_{i,1}, \ldots, \varphi_{i,m_i}$ are represented by the solid dots

Remark 4.1 For certain problems with high contrast β/α, the decays of $\boldsymbol{\psi}_i$ and $\boldsymbol{\xi}_i$ are sensitive to H if we use the data independent projection operator Π_H in the construction of $V_H^{\mathrm{ms,h}}$. Under additional assumptions on the diffusion matrix $\mathcal{A}$, better decays can be achieved by replacing Π_H with a data adapted interpolation operator (cf. [17, 34]).

We will construct a localized multiscale finite element space by replacing $\boldsymbol{\psi}_i$ (resp., $\boldsymbol{\xi}_i$) with an approximate solution of (4.1) (resp., (4.2)) obtained by a preconditioned minimum residual (P-MINRES) algorithm (cf. [16,Chapter 8] and [14,Section 4.1]). Our construction extends those in [6, 28] to symmetric indefinite problems.

4.1 An Additive Schwarz Preconditioner

Let $A_h^{\Pi_H} : K_h^{\Pi_H} \longrightarrow (K_h^{\Pi_H})'$ be the linear operator defined by

$$\langle A_h^{\Pi_H} v, w\rangle = a(v, w) \qquad \forall\, v, w \in K_h^{\Pi_H}, \tag{4.3}$$

where $a(\cdot,\cdot)$ is given in (1.3). We begin by constructing an additive Schwarz preconditioner (cf. [7, 10, 35]) for $A_h^{\Pi_H}$.

Let $x_1, \ldots, x_m$ be the (interior) nodes for V_H. We define the subspaces $K_{h,i}^{\Pi_H}$ $(1 \le i \le m)$ of $K_h^{\Pi_H}$ by

$$K_{h,i}^{\Pi_H} = \{(I - \Pi_H)v : \ v \in V_h \text{ and } v \text{ vanishes outside } \omega_{x_i}\}, \tag{4.4}$$

where ω_{x_i} is the union of the elements in $\mathcal{T}_H$ that share x_i as a common vertex (cf. Fig. 8 for a two dimensional example with the Q_1 element). The functions in $K_{h,i}^{\Pi_H}$ are supported on the patch $\tilde{\omega}_{x_i}$ obtained from ω_{x_i} by adding one layer of elements in $\mathcal{T}_H$ (cf. Fig. 8). Let $\varphi_{i,1}, \ldots, \varphi_{i,m_i}$ be the nodal basis functions of V_h that vanish at x_i and outside ω_{x_i} (cf. Fig. 8). Then, as in Lemma 3.2, $\{(I - \Pi_H)\varphi_{i,1}, \ldots, (I - \Pi_H)\varphi_{i,m_i}\}$ is a basis of $K_{h,i}^{\Pi_H}$.

Let $I_i : K_{h,i}^{\Pi_H} \longrightarrow K_h^{\Pi_H}$ be the natural injection. The SPD additive Schwarz preconditioner $S_h^{\Pi_H} : (K_h^{\Pi_H})' \longrightarrow K_h^{\Pi_H}$ for $A_h^{\Pi_H}$ is given by

$$S_h^{\Pi_H} = \sum_{i=1}^{m} I_i (A_{h,i}^{\Pi_H})^{-1} I_i^t, \tag{4.5}$$

where $A_{h,i}^{\Pi_H} : K_{h,i}^{\Pi_H} \longrightarrow (K_{h,i}^{\Pi_H})'$ is defined by

$$\langle A_{h,i}^{\Pi_H} v, w\rangle = a(v, w) \qquad \forall\, v, w \in K_{h,i}^{\Pi_H}. \tag{4.6}$$

According to the Raleigh quotient formulas, we have

$$a(v, v) = \langle A_h^{\Pi_H} v, v\rangle \le \lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})\langle (S_h^{\Pi_H})^{-1} v, v\rangle \qquad \forall\, v \in K_h^{\Pi_H}, \tag{4.7}$$

$$a(v, v) = \langle A_h^{\Pi_H} v, v\rangle \ge \lambda_{\min}(S_h^{\Pi_H} A_h^{\Pi_H})\langle (S_h^{\Pi_H})^{-1} v, v\rangle \qquad \forall\, v \in K_h^{\Pi_H}, \tag{4.8}$$

and the following spectral estimates can be found in [6,Section 3]:

$$\lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H}) \le C_{\text{upper}} \quad \text{and} \quad \lambda_{\min}(S_h^{\Pi_H} A_h^{\Pi_H}) \ge C_{\text{lower}}(\alpha/\beta), \tag{4.9}$$

where the positive constants C_{upper} and C_{lower} only depend on the shape regularity of $\mathcal{T}_H$.

4.2 The Generalized Finite Element Space $V_{H,k}^{\text{ms,h}}$

Let

$$\mathbb{B}_h^{\Pi_H} : K_h^{\Pi_H} \times K_h^{\Pi_H} \longrightarrow (K_h^{\Pi_H} \times K_h^{\Pi_H})'$$

be the linear operator defined by

$$\langle \mathbb{B}_h^{\Pi_H} \boldsymbol{v}, \boldsymbol{w}\rangle = \mathcal{B}(\boldsymbol{v}, \boldsymbol{w}) \qquad \forall\, \boldsymbol{v}, \boldsymbol{w} \in K_h^{\Pi_H} \times K_h^{\Pi_H}. \tag{4.10}$$

We can then rewrite (4.1) and (4.2) as

$$\mathbb{B}_h^{\Pi_H} \boldsymbol{\psi}_i = \boldsymbol{f}_i, \tag{4.11}$$

$$\mathbb{B}_h^{\Pi_H} \boldsymbol{\xi}_i = \boldsymbol{g}_i, \tag{4.12}$$

where $\boldsymbol{f}_i, \boldsymbol{g}_i \in (K_h^{\Pi_H} \times K_h^{\Pi_H})'$ are defined by

$$\langle \boldsymbol{f}_i, \boldsymbol{w}\rangle = \mathcal{B}\big((\phi_i, 0), \boldsymbol{w}\big) \quad \text{and} \quad \langle \boldsymbol{g}_i, \boldsymbol{w}\rangle = \mathcal{B}\big((0, \phi_i), \boldsymbol{w}\big) \qquad \forall\, \boldsymbol{w} \in K_h^{\Pi_H} \times K_h^{\Pi_H}.$$

Let $\boldsymbol{\psi}_{i,k} \in K_h^{\Pi_H} \times K_h^{\Pi_H}$ (resp., $\boldsymbol{\xi}_{i,k} \in K_h^{\Pi_H} \times K_h^{\Pi_H}$) be the output of k steps of the P-MINRES algorithm applied to (4.11) (resp., (4.12)) with initial guess 0, where the SPD preconditioner

$$\mathbb{S}_h^{\Pi_H} : (K_h^{\Pi_H} \times K_h^{\Pi_H})' \longrightarrow K_h^{\Pi_H} \times K_h^{\Pi_H}$$

is given by

$$\mathbb{S}_h^{\Pi_H}(\mu, \rho) = (S_h^{\Pi_H} \mu, S_h^{\Pi_H} \rho). \tag{4.13}$$

Then the $2m$ functions

$$(\phi_1, 0) - \boldsymbol{\psi}_{1,k}, \ldots, (\phi_m, 0) - \boldsymbol{\psi}_{m,k}, (0, \phi_1) - \boldsymbol{\xi}_{1,k}, \ldots, (0, \phi_m) - \boldsymbol{\xi}_{m,k} \tag{4.14}$$

are linearly independent because the intersection of $V_H \times V_H$ and $K_h^{\Pi_H} \times K_h^{\Pi_H}$ is trivial, and we define the generalized finite element space $V_{H,k}^{\text{ms,h}}$ by

$$V_{H,k}^{\text{ms,h}} = \text{span}\big\{(\phi_1, 0) - \boldsymbol{\psi}_{1,k}, \ldots, (\phi_m, 0) - \boldsymbol{\psi}_{m,k}, (0, \phi_1) - \boldsymbol{\xi}_{1,k}, \ldots, (0, \phi_m) - \boldsymbol{\xi}_{m,k}\big\}. \tag{4.15}$$

Remark 4.2 It follows from (2.6), (4.5) and (4.10) that the support of $\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}(q,z)$ is a subset of the union of all the $\tilde{\omega}_{x_i}$ whose intersection with the support of (q,z) have nonempty interiors. As the output of a preconditioned Krylov subspace method with initial guess 0, the function $\boldsymbol{\psi}_{i,k}$ belongs to

$$\operatorname{span}\{\mathbb{S}_h^{\Pi_H}\boldsymbol{f}_i, (\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H})\mathbb{S}_h^{\Pi_H}\boldsymbol{f}_i, \ldots, (\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H})^{k-1}\mathbb{S}_h^{\Pi_H}\boldsymbol{f}_i\},$$

and hence is supported in a patch around x_i (with respect to $\mathcal{T}_H$) whose diameter is proportional to kH. This is also true for the function $\boldsymbol{\xi}_{i,k}$. The functions in (4.14) are therefore locally corrected basis functions and $V_{H,k}^{\mathrm{ms,h}}$ defined by (4.15) is a localized multiscale finite element space.

Remark 4.3 In view of (2.6), we can express $\mathbb{B}_h^{\Pi_H}$ in the matrix form

$$\mathbb{B}_h^{\Pi_H} = \begin{bmatrix} A_h^{\Pi_H} & M_h^{\Pi_H} \\ M_h^{\Pi_H} & -A_h^{\Pi_H} \end{bmatrix}, \tag{4.16}$$

where $M_h^{\Pi_H} : K_h^{\Pi_H} \longrightarrow (K_h^{\Pi_H})'$ is the (symmetric) linear operator defined by

$$\langle M_h^{\Pi_H} r, z\rangle = \int_\Omega rz\,dx \qquad \forall\, r, z \in K_h^{\Pi_H}.$$

We can also write $\mathbb{S}_h^{\Pi_H}$ as the diagonal matrix

$$\mathbb{S}_h^{\Pi_H} = \begin{bmatrix} S_h^{\Pi_H} & 0 \\ 0 & S_h^{\Pi_H} \end{bmatrix}. \tag{4.17}$$

4.3 Spectral Analysis of $\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H} : K_h^{\Pi_H} \times K_h^{\Pi_H} \longrightarrow K_h^{\Pi_H} \times K_h^{\Pi_H}$

A spectral analysis of the operator $\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}$ is provided in the following lemma, which is the key to the analysis of the localized multiscale finite element method in Sect. 5.

Lemma 4.4 *The spectrum $\sigma(\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H})$ of $\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}$ satisfies*

$$\sigma(\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}) \subseteq [-d_*, -c_*] \cup [c_*, d_*], \tag{4.18}$$

where

$$c_* = \lambda_{\min}(S_h^{\Pi_H} A_h^{\Pi_H}) \text{ and } d_* = \lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})[1 + (\mathrm{C_{PF}}/\alpha)]. \tag{4.19}$$

Proof In view of (4.16) and (4.17), the eigenvalues of $\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}$ are real numbers. Let λ be one of the eigenvalues and $(r,s) \in K_h^{\Pi_H} \times K_h^{\Pi_H}$ be a corresponding eigenvector. Given any $(q,z) \in K_h^{\Pi_H} \times K_h^{\Pi_H}$, we have, by (4.10),

$$\mathcal{B}((r,s),(q,z)) = \langle (\mathbb{S}_h^{\Pi_H})^{-1}\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}(r,s),(q,z)\rangle = \lambda\langle (\mathbb{S}_h^{\Pi_H})^{-1}(r,s),(q,z)\rangle. \tag{4.20}$$

It follows from (2.9), (4.3), (4.8), (4.13) and (4.20) that

$$\begin{aligned} \|(r,s)\|_{a\times a}^2 = \mathcal{B}((r,s),(r,-s)) &= \lambda\langle (\mathbb{S}_h^{\Pi_H})^{-1}(r,s),(r,-s)\rangle \\ &= \lambda\big(\langle (S_h^{\Pi_H})^{-1}r, r\rangle - \langle (S_h^{\Pi_H})^{-1}s, s\rangle\big) \end{aligned}$$

$$\leq |\lambda|\big(\langle (S_h^{\Pi_H})^{-1} r, r\rangle + \langle (S_h^{\Pi_H})^{-1} s, s\rangle\big)$$
$$\leq |\lambda|[1/\lambda_{\min}(S_h^{\Pi_H} A_h^{\Pi_H})]\|(r,s)\|_{a\times a}^2$$

and hence

$$|\lambda| \geq \lambda_{\min}(S_h^{\Pi_H} A_h^{\Pi_H}).$$

On the other hand we can deduce from (2.10), (4.7) and (4.20) that

$$\begin{aligned}|\lambda|\|(r,s)\|_{a\times a}^2 &\leq \lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})\big|\lambda\big(\langle (S_h^{\Pi_H})^{-1} r, r\rangle + \langle (S_h^{\Pi_H})^{-1} s, s\rangle\big)\big| \\ &= \lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})\big|\lambda\langle (\mathbb{S}_h^{\Pi_H})^{-1}(r,s),(r,s)\rangle\big| \\ &= \lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})|\mathcal{B}((r,s),(r,s))| \\ &\leq \lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})[1 + (\mathrm{C_{PF}}/\alpha)]\|(r,s)\|_{a\times a}^2\end{aligned}$$

and hence

$$|\lambda| \leq \lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})[1 + (\mathrm{C_{PF}}/\alpha)].$$

□

Corollary 4.5 *The following relations hold for any* $\boldsymbol{v} \in K_h^{\Pi_H} \times K_h^{\Pi_H}$:

$$\begin{aligned}\left(\frac{c_*}{\lambda_{\max}(S_h^{\Pi_H} A_h^{\Pi_H})^{\frac{1}{2}}}\right)\|\boldsymbol{v}\|_{a\times a} &\leq \langle \mathbb{B}_h^{\Pi_H}\boldsymbol{v}, \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}\boldsymbol{v}\rangle^{\frac{1}{2}} \\ &\leq \left(\frac{d_*}{\lambda_{\min}(S_h^{\Pi_H} A_h^{\Pi_H})^{\frac{1}{2}}}\right)\|\boldsymbol{v}\|_{a\times a}. \qquad (4.21)\end{aligned}$$

Proof Since the operator $\mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}$ is symmetric with respect to the inner product defined by $\langle (\mathbb{S}_h^{\Pi_H})^{-1}\cdot,\cdot\rangle$, we have, by Lemma 4.4 and the spectral theorem,

$$c_*\langle (\mathbb{S}_h^{\Pi_H})^{-1}\boldsymbol{v},\boldsymbol{v}\rangle^{\frac{1}{2}} \leq \langle \mathbb{B}_h^{\Pi_H}\boldsymbol{v}, \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}\boldsymbol{v}\rangle^{\frac{1}{2}} \leq d_*\langle (\mathbb{S}_h^{\Pi_H})^{-1}\boldsymbol{v},\boldsymbol{v}\rangle^{\frac{1}{2}},$$

which together with (4.7) and (4.8) implies (4.21). □

5 The Localized Multiscale Finite Element Method

The localized multiscale finite element method is to find $(p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h}) \in V_{H,k}^{\mathrm{ms,h}}$ such that

$$\mathcal{B}\big((p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h}),(q,z)\big) = \int_\Omega y_d q\, dx \qquad \forall\,(q,z) \in V_{H,k}^{\mathrm{ms,h}}. \qquad (5.1)$$

We will keep track of all the constants in the error analysis so that the constants that appear in the energy error estimate (cf. Theorem 5.11) and the $L_2(\Omega)\times L_2(\Omega)$ error estimate (cf. Theorem 5.13) are independent of the mesh sizes (h and H) and the contrast β/α.

We begin the error analysis by comparing $\boldsymbol{\psi}_i$ and $\boldsymbol{\psi}_{i,k}$ (resp., $\boldsymbol{\xi}_i$ and $\boldsymbol{\xi}_{i,k}$).

5.1 The Relation Between $(\boldsymbol{\psi}_i, \boldsymbol{\xi}_i)$ and $(\boldsymbol{\psi}_{i,k}, \boldsymbol{\xi}_{i,k})$

It follows from Lemma 4.4 and the theory of the P-MINRES algorithm (cf. [14,Theorem 4.14]) that

$$\langle \mathbb{B}_h^{\Pi_H}(\boldsymbol{\psi}_i - \boldsymbol{\psi}_{i,k}), \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}(\boldsymbol{\psi}_i - \boldsymbol{\psi}_{i,k})\rangle^{\frac{1}{2}} \leq 2q^{\lfloor k/2 \rfloor}\langle \mathbb{B}_h^{\Pi_H}\boldsymbol{\psi}_i, \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}\boldsymbol{\psi}_i\rangle^{\frac{1}{2}}, \tag{5.2}$$

$$\langle \mathbb{B}_h^{\Pi_H}(\boldsymbol{\xi}_i - \boldsymbol{\xi}_{i,k}), \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}(\boldsymbol{\xi}_i - \boldsymbol{\xi}_{i,k})\rangle^{\frac{1}{2}} \leq 2q^{\lfloor k/2 \rfloor}\langle \mathbb{B}_h^{\Pi_H}\boldsymbol{\xi}_i, \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}\boldsymbol{\xi}_i\rangle^{\frac{1}{2}}, \tag{5.3}$$

where (cf. (4.19))

$$\begin{aligned} q &= \frac{d_* - c_*}{d_* + c_*} \\ &= \frac{\kappa(S_h^{\Pi_H}A_h^{\Pi_H})[1 + (\mathrm{C_{PF}}/\alpha)] - 1}{\kappa(S_h^{\Pi_H}A_h^{\Pi_H})[1 + (\mathrm{C_{PF}}/\alpha)] + 1}. \end{aligned} \tag{5.4}$$

Remark 5.1 It follows from (4.9) that the condition number $\kappa(S_h^{\Pi_H}A_h^{\Pi_H})$ has the following (pessimistic) upper bound that is independent of the mesh sizes:

$$\kappa(S_h^{\Pi_H}A_h^{\Pi_H}) = \frac{\lambda_{\max}(S_h^{\Pi_H}A_h^{\Pi_H})}{\lambda_{\min}(S_h^{\Pi_H}A_h^{\Pi_H})} \leq \Big(\frac{C_{\mathrm{upper}}}{C_{\mathrm{lower}}}\Big)\Big(\frac{\beta}{\alpha}\Big). \tag{5.5}$$

Consequently we also have the bound

$$q \leq \frac{\Big(\frac{C_{\mathrm{upper}}}{C_{\mathrm{lower}}}\Big)\Big(\frac{\beta}{\alpha}\Big)[1 + (\mathrm{C_{PF}}/\alpha)] - 1}{\Big(\frac{C_{\mathrm{upper}}}{C_{\mathrm{lower}}}\Big)\Big(\frac{\beta}{\alpha}\Big)[1 + (\mathrm{C_{PF}}/\alpha)] + 1} \tag{5.6}$$

that is independent of the mesh sizes, but which may be too pessimistic.

Lemma 5.2 *We have*

$$\|\boldsymbol{\psi}_i - \boldsymbol{\psi}_{i,k}\|_{a\times a} \leq C_\star q^{\lfloor k/2 \rfloor}\|\boldsymbol{\psi}_i\|_{a\times a} \qquad \text{for} \quad 1 \leq i \leq m \tag{5.7}$$

and

$$\|\boldsymbol{\xi}_i - \boldsymbol{\xi}_{i,k}\|_{a\times a} \leq C_\star q^{\lfloor k/2 \rfloor}\|\boldsymbol{\xi}_i\|_{a\times a} \qquad \text{for} \quad 1 \leq i \leq m, \tag{5.8}$$

where the positive constant $C_\star$ is independent of the mesh sizes.

Proof It follows from (4.21) and (5.2) that

$$\begin{aligned} \|\boldsymbol{\psi}_i - \boldsymbol{\psi}_{i,k}\|_{a\times a} &\leq \frac{\lambda_{\max}(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{1}{2}}}{c_*}\langle \mathbb{B}_h^{\Pi_H}(\boldsymbol{\psi}_i - \boldsymbol{\psi}_{i,k}), \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}(\boldsymbol{\psi}_i - \boldsymbol{\psi}_{i,k})\rangle^{\frac{1}{2}} \\ &\leq 2q^{\lfloor k/2 \rfloor}\frac{\lambda_{\max}(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{1}{2}}}{c_*}\langle \mathbb{B}_h^{\Pi_H}\boldsymbol{\psi}_i, \mathbb{S}_h^{\Pi_H}\mathbb{B}_h^{\Pi_H}\boldsymbol{\psi}_i\rangle^{\frac{1}{2}} \\ &\leq 2q^{\lfloor k/2 \rfloor}\frac{\lambda_{\max}(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{1}{2}}}{c_*}\frac{d_*}{\lambda_{\min}(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{1}{2}}}\|\boldsymbol{\psi}_i\|_{a\times a}, \end{aligned}$$

which implies (5.7) with (cf. (4.19))

$$C_\star = 2\kappa(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{3}{2}}[1 + (\mathrm{C_{PF}}/\alpha)]. \tag{5.9}$$

Similarly we can derive (5.8) from (4.21) and (5.3). □

Remark 5.3 It follows from Lemma 5.2 that the basis corrections $\mathcal{C}_h^{\Pi_H}(\phi_i, 0) = \boldsymbol{\psi}_i$ and $\mathcal{C}_h^{\Pi_H}(0, \phi_i) = \boldsymbol{\xi}_i$ $(1 \le i \le m)$ used in the construction of $V_H^{\mathrm{ms,h}}$ decay exponentially, as observed in Figs. 7, 6 and Remark 3.10.

Let the linear operator $\mathcal{C}_{h,k}^{\Pi_H} : V_H \times V_H \longrightarrow K_h^{\Pi_H} \times K_h^{\Pi_H}$ be defined by

$$\mathcal{C}_{h,k}^{\Pi_H}(\phi_i, 0) = \boldsymbol{\psi}_{i,k} \quad \text{and} \quad \mathcal{C}_{h,k}^{\Pi_H}(0, \phi_i) = \boldsymbol{\xi}_{i,k} \quad \text{for} \quad 1 \le i \le m, \tag{5.10}$$

where $\phi_1, \ldots, \phi_m$ are the nodal basis functions of V_H.

Our next goal is to understand the relation between the operators $\mathcal{C}_h^{\Pi_H}$ and $\mathcal{C}_{h,k}^{\Pi_H}$.

5.2 The Relation Between $\mathcal{C}_h^{\Pi_H}$ and $\mathcal{C}_{h,k}^{\Pi_H}$

We begin with the following lemma.

Lemma 5.4 *There exist positive constants $C_\diamond$ and $C_\heartsuit$ depending only on the shape regularity of $\mathcal{T}_H$ such that*

$$|\phi_i|_{H^1(\Omega)} \le C_\diamond H^{\tau_d} \quad \text{with} \quad \tau_d = \begin{cases} -\frac{1}{2} & d = 1 \\ 0 & d = 2 \\ \frac{1}{2} & d = 3 \end{cases}, \tag{5.11}$$

where d is the dimension of Ω, and

$$\sum_{i=1}^{m}(|c_i| + |d_i|) \le C_\heartsuit(1/\sqrt{\alpha})H^{-d}\Big\|\sum_{i=1}^{m}\big[c_i(\phi_i, 0) + d_i(0, \phi_i)\big]\Big\|_{a\times a} \tag{5.12}$$

for any real numbers $c_1, \ldots, c_m, d_1, \ldots, d_m$.

Proof The estimate (5.11) follows from a scaling argument. For the estimate (5.12), we begin with an inverse estimate

$$\sum_{i=1}^{m}(|c_i| + |d_i|) \le C_\natural H^{-d}\Big\|\sum_{i=1}^{m}\big[c_i(\phi_i, 0) + d_i(0, \phi_i)\big]\Big\|_{L_1(\Omega)\times L_1(\Omega)},$$

where the positive constant $C_\natural$ depends only on the shape regularity of $\mathcal{T}_H$. The proof is then completed by the Poincaré-Friedrichs inequality

$$\|\zeta\|_{L_1(\Omega)} \le C_\Omega|\zeta|_{H^1(\Omega)} \qquad \forall\, \zeta \in H_0^1(\Omega)$$

and the estimate (2.7). □

Lemma 5.5 *The following estimate is valid for any $\boldsymbol{v} \in V_H \times V_H$:*

$$\|(\mathcal{C}_h^{\Pi_H} - \mathcal{C}_{h,k}^{\Pi_H})\boldsymbol{v}\|_{a\times a} \le C_\spadesuit\kappa(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{3}{2}}\Big(\frac{\beta}{\alpha}\Big)^{\frac{1}{2}}[1 + (\mathrm{C_{PF}}/\alpha)]^2 H^{-d+\tau_d}q^{\lfloor k/2\rfloor}\|\boldsymbol{v}\|_{a\times a}, \tag{5.13}$$

where the positive constant $C_\spadesuit$ depends only on the shape regularity of $\mathcal{T}_H$.

Proof Let $\sum_{i=1}^{m}[c_i(\phi_i,0)+d_i(0,\phi_i)]$ be an arbitrary element of $V_H\times V_H$. It follows from (2.7), Lemmas 3.5, 5.2 and 5.4 that

$$\begin{aligned}
&\Big\|(\mathcal{C}_h^{\Pi_H}-\mathcal{C}_{h,k}^{\Pi_H})\sum_{i=1}^{m}[c_i(\phi_i,0)+d_i(0,\phi_i)]\Big\|_{a\times a}\\
&\quad\le\sum_{i=1}^{m}\big[|c_i|\|\boldsymbol{\psi}_i-\boldsymbol{\psi}_{i,k}\|_{a\times a}+|d_i|\|\boldsymbol{\xi}_i-\boldsymbol{\xi}_{i,k}\|_{a\times a}\big]\\
&\quad\le C_\star q^{\lfloor k/2\rfloor}\sum_{i=1}^{m}\big[|c_i|\|\boldsymbol{\psi}_i\|_{a\times a}+|d_i|\|\boldsymbol{\xi}_i\|_{a\times a}\big]\\
&\quad\le C_\star q^{\lfloor k/2\rfloor}[1+(\mathrm{C_{PF}}/\alpha)]\sum_{i=1}^{m}\big[|c_i|\|\phi_i\|_a+|d_i|\|\phi_i\|_a\big]\\
&\quad\le C_\star q^{\lfloor k/2\rfloor}[1+(\mathrm{C_{PF}}/\alpha)]\sqrt{\beta}\sum_{i=1}^{m}(|c_i|+|d_i|)|\phi_i|_{H^1(\Omega)}\\
&\quad\le C_\star q^{\lfloor k/2\rfloor}[1+(\mathrm{C_{PF}}/\alpha)]\sqrt{\beta}\,C_\diamond H^{\tau_d}\sum_{i=1}^{m}(|c_i|+|d_i|)\\
&\quad\le C_\star q^{\lfloor k/2\rfloor}[1+(\mathrm{C_{PF}}/\alpha)]\sqrt{\beta}\,C_\diamond H^{\tau_d}C_\heartsuit(1/\sqrt{\alpha})H^{-d}\\
&\qquad\times\Big\|\sum_{i=1}^{m}\big[c_i(\phi_i,0)+d_i(0,\phi_i)\big]\Big\|_{a\times a}\\
&\quad= C_\star C_\diamond C_\heartsuit\sqrt{\beta/\alpha}[1+(\mathrm{C_{PF}}/\alpha)]H^{-d+\tau_d}q^{\lfloor k/2\rfloor}\Big\|\sum_{i=1}^{m}\big[c_i(\phi_i,0)+d_i(0,\phi_i)\big]\Big\|_{a\times a},
\end{aligned}$$

which together with (5.9) implies (5.13) for $C_\spadesuit=2C_\diamond C_\heartsuit$. □

We have an analog of Corollary 3.6.

Corollary 5.6 *The following relations are valid for any* $\boldsymbol{v}\in V_H\times V_H$:

$$\begin{aligned}
\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a}\le\Big(&C_\spadesuit\kappa(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{3}{2}}\Big(\frac{\beta}{\alpha}\Big)^{\frac{1}{2}}[1+(\mathrm{C_{PF}}/\alpha)]^2H^{-d+\tau_d}q^{\lfloor k/2\rfloor}\\
&+[2+(\mathrm{C_{PF}}/\alpha)]\Big)\|\boldsymbol{v}\|_{a\times a},
\end{aligned}\tag{5.14}$$

$$\|\boldsymbol{v}\|_{a\times a}\le C_\dagger\sqrt{\beta/\alpha}\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a}.\tag{5.15}$$

Proof The estimate (5.14) follows from (3.11), (5.13) and the triangle inequality. The proof of (5.15) is identical to the proof of (3.12). □

5.3 The Well-Posedness of (5.1)

We will use Corollary 3.6, Lemma 5.5 and Corollary 5.6 to derive an analog of Lemma 3.11 under some assumptions on k.

Assumption 1 The number of P-MINRES steps k is sufficiently large so that, by (5.14),

$$\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a}\le[3+(\mathrm{C_{PF}}/\alpha)]\|\boldsymbol{v}\|_{a\times a}\qquad\forall\,\boldsymbol{v}\in V_H\times V_H.\tag{5.16}$$

Remark 5.7 The following condition on k guarantees (5.16):

$$C_{\spadesuit}\kappa(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{3}{2}}\Big(\frac{\beta}{\alpha}\Big)^{\frac{1}{2}}[1+(\mathrm{C_{PF}}/\alpha)]^2 H^{-d+\tau_d}q^{\lfloor k/2\rfloor}\le 1,$$

or equivalently

$$\begin{aligned}&\ln C_{\spadesuit}+\frac{3}{2}\ln\kappa(S_h^{\Pi_H}A_h^{\Pi_H})+\frac{1}{2}\ln(\beta/\alpha)+2\ln[1+(\mathrm{C_{PF}}/\alpha)]\\&+(-d+\tau_d)\ln H+\lfloor k/2\rfloor\ln q\le 0\end{aligned}\tag{5.17}$$

It follows from Corollaries 3.6, 5.6 and (5.16) that

$$\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a}\le C_{\clubsuit}\|\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a}\qquad\forall\,\boldsymbol{v}\in V_H\times V_H,\tag{5.18}$$

where

$$C_{\clubsuit}=[3+(\mathrm{C_{PF}}/\alpha)]C_{\dagger}\sqrt{\beta/\alpha}.\tag{5.19}$$

Assumption 2 The number of P-MINRES steps k is sufficiently large so that, by (3.12), Lemma 5.5 and (5.15),

$$\begin{aligned}&C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)]^2\|(\mathcal{C}_h^{\Pi_H}-\mathcal{C}_{h,k}^{\Pi_H})\boldsymbol{v}\|_{a\times a}\\&\le\frac{1}{3}\min(\|\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a},\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a})\end{aligned}\tag{5.20}$$

for all $\boldsymbol{v}\in V_H\times V_H$.

Remark 5.8 The following condition on k guarantees (5.20):

$$3C_{\spadesuit}C_{\dagger}^2[3+(\mathrm{C_{PF}}/\alpha)]^5\kappa(S_h^{\Pi_H}A_h^{\Pi_H})^{\frac{3}{2}}\Big(\frac{\beta}{\alpha}\Big)^{\frac{3}{2}}H^{-d+\tau_d}q^{\lfloor k/2\rfloor}\le 1,$$

or equivalently

$$\begin{aligned}&\ln(3C_{\spadesuit}C_{\dagger}^2)+5\ln[3+(\mathrm{C_{PF}}/\alpha)]+\frac{3}{2}\ln\kappa(S_h^{\Pi_H}A_h^{\Pi_H})\\&+\frac{3}{2}\ln(\beta/\alpha)+(-d+\tau_d)\ln H+\lfloor k/2\rfloor\ln q\le 0.\end{aligned}\tag{5.21}$$

The well-posedness of (5.1) for a sufficiently large k is addressed by the following lemma.

Lemma 5.9 *The inf-sup condition*

$$\inf_{\boldsymbol{v}\in V_H\times V_H}\sup_{\boldsymbol{w}\in V_H\times V_H}\frac{\mathcal{B}\big(\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v},\boldsymbol{w}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{w}\big)}{\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a}\|\boldsymbol{w}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{w}\|_{a\times a}}\ge\frac{1}{3C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)]}\tag{5.22}$$

holds under Assumption 1 *and Assumption* 2.

Proof Let $\boldsymbol{v}\in V_H\times V_H$ be arbitrary. We have, by (2.10), Remark 3.8, Lemma 3.11, (5.18) and (5.20),

$$\begin{aligned}\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a}&\le C_{\clubsuit}\|\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v}\|_{a\times a}\\&\le C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)]\sup_{\boldsymbol{w}\in V_H\times V_H}\frac{\mathcal{B}\big(\boldsymbol{v}-\mathcal{C}_h^{\Pi_H}\boldsymbol{v},\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\big)}{\|\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\|_{a\times a}}\end{aligned}$$

$$
\begin{aligned}
&\le C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)] \sup_{\boldsymbol{w}\in V_H\times V_H} \frac{\mathcal{B}\big((\mathcal{C}_{h,k}^{\Pi_H}-\mathcal{C}_h^{\Pi_H})\boldsymbol{v},\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\big)}{\|\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\|_{a\times a}} \\
&\quad + C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)] \sup_{\boldsymbol{w}\in V_H\times V_H} \frac{\mathcal{B}\big(\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v},\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\big)}{\|\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\|_{a\times a}} \\
&\le C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)]^2\|(\mathcal{C}_{h,k}^{\Pi_H}-\mathcal{C}_h^{\Pi_H})\boldsymbol{v}\|_{a\times a} \\
&\quad + C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)] \sup_{\boldsymbol{w}\in V_H\times V_H} \frac{\mathcal{B}\big(\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v},(\mathcal{C}_{h,k}^{\Pi_H}-\mathcal{C}_h^{\Pi_H})\boldsymbol{w}\big)}{\|\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\|_{a\times a}} \\
&\quad + C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)] \sup_{\boldsymbol{w}\in V_H\times V_H} \frac{\mathcal{B}\big(\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v},\boldsymbol{w}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{w}\big)}{\|\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\|_{a\times a}} \\
&\le \frac{1}{3}\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a} \\
&\quad + C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)]^2\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a} \sup_{\boldsymbol{w}\in V_H\times V_H} \frac{\|(\mathcal{C}_h^{\Pi_H}-\mathcal{C}_{h,k}^{\Pi_H})\boldsymbol{w}\|_{a\times a}}{\|\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\|_{a\times a}} \\
&\quad + C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)] \sup_{\boldsymbol{w}\in V_H\times V_H} \frac{\mathcal{B}\big(\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v},\boldsymbol{w}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{w}\big)}{\|\boldsymbol{w}-\mathcal{C}_h^{\Pi_H}\boldsymbol{w}\|_{a\times a}} \\
&\le \frac{2}{3}\|\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v}\|_{a\times a} + C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)] \\
&\quad \times \sup_{\boldsymbol{w}\in V_H\times V_H} \frac{\mathcal{B}\big(\boldsymbol{v}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{v},\boldsymbol{w}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{w}\big)}{\|\boldsymbol{w}-\mathcal{C}_{h,k}^{\Pi_H}\boldsymbol{w}\|_{a\times a}},
\end{aligned}
$$

which implies (5.22). □

Remark 5.10 In view of (4.15) and (5.10), we can rewrite the inf-sup condition (5.22) as

$$
\inf_{\boldsymbol{v}\in V_{H,k}^{\mathrm{ms,h}}} \sup_{\boldsymbol{w}\in V_{H,k}^{\mathrm{ms,h}}} \frac{\mathcal{B}(\boldsymbol{v},\boldsymbol{w})}{\|\boldsymbol{v}\|_{a\times a}\|\boldsymbol{w}\|_{a\times a}} \ge \frac{1}{3C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)]}. \tag{5.23}
$$

5.4 Energy Error

Under the inf-sup condition (5.23) we have a standard quasi-optimal error estimate (cf. [8,Theorem 2.1]) for the solution $(p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})$ of (5.1):

$$
\|(p_h,y_h)-(p_{H,k}^{\mathrm{ms},h},y_{H,k}^{\mathrm{ms},h})\|_{a\times a} \le C_{\maltese} \inf_{(q,z)\in V_{H,k}^{\mathrm{ms,h}}} \|(p_h,y_h)-(q,z)\|_{a\times a}, \tag{5.24}
$$

where

$$
C_{\maltese} = 1+3C_{\clubsuit}[2+(\mathrm{C_{PF}}/\alpha)]. \tag{5.25}
$$

It then follows from Remark 3.1, (3.10), Remark 3.13, (4.15), Lemma 5.5 and (5.24) that

$$
\begin{aligned}
&\|(p_h,y_h)-(p_{H,k}^{\mathrm{ms},h},y_{H,k}^{\mathrm{ms},h})\|_{a\times a} \\
&\quad \le C_{\maltese}\big\|(p_h,y_h)-(\boldsymbol{I}-\mathcal{C}_{h,k}^{\Pi_H})(\Pi_H\times\Pi_H)(p_h,y_h)\big\|_{a\times a} \\
&\quad = C_{\maltese}\big\|\big[(p_h,y_h)-(\boldsymbol{I}-\mathcal{C}_h^{\Pi_H})(\Pi_H\times\Pi_H)(p_h,y_h)\big] \\
&\qquad\qquad + (\mathcal{C}_{h,k}^{\Pi_H}-\mathcal{C}_h^{\Pi_H})(\Pi_H\times\Pi_H)(p_h,y_h)\big\|_{a\times a}
\end{aligned}
$$

$$
\begin{aligned}
&= C_{\maltese}\left\|\left[(p_h, y_h) - (\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(p_h, y_h)\right] + (\mathcal{C}_{h,k}^{\Pi_H} - \mathcal{C}_h^{\Pi_H})(\Pi_H \times \Pi_H)(p_h, y_h)\right\|_{a\times a}\\
&\le C_{\maltese}\Big(\|(p_h, y_h) - (p_H^{\mathrm{ms},h}, y_H^{\mathrm{ms},h})\|_{a\times a} + \|(\mathcal{C}_{h,k}^{\Pi_H} - \mathcal{C}_h^{\Pi_H})(\Pi_H \times \Pi_H)(p_h, y_h)\|_{a\times a}\Big)\\
&\le C_{\maltese}\Big(\|(p_h, y_h) - (p_H^{\mathrm{ms},h}, y_H^{\mathrm{ms},h})\|_{a\times a} + C_{\spadesuit} C_{\dagger} \kappa(S_h^{\Pi_H} A_h^{\Pi_H})^{\frac{3}{2}} (\beta/\alpha)\\
&\qquad \times [1 + (\mathrm{C_{PF}}/\alpha)]^2 H^{-d+\tau_d} q^{\lfloor k/2\rfloor} \|(p_h, y_h)\|_{a\times a}\Big),
\end{aligned}
\tag{5.26}
$$

i.e., up to a term that decreases exponentially as k increases, the performance of the localized multiscale finite element method defined by (5.1) is similar to the performance of the ideal multiscale finite element method defined by (3.16).

Assumption 3 The number of P-MINRES steps k is sufficiently large so that

$$
q^{\lfloor k/2\rfloor} \le \kappa(S_h^{\Pi_H} A_h^{\Pi_H})^{-\frac{3}{2}} (\alpha/\beta) H^{1+d-\tau_d}. \tag{5.27}
$$

Theorem 5.11 *Under Assumptions* 1–3*, we have*

$$
\|(p_h, y_h) - (p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})\|_{a\times a} \le C_{\sharp} H \|y_d\|_{L_2(\Omega)}, \tag{5.28}
$$

where

$$
C_{\sharp} = C_{\maltese}(C_{\dagger}/\sqrt{\alpha})\Big(1 + C_{\spadesuit}[1 + (\mathrm{C_{PF}}/\alpha)]^2 \sqrt{\mathrm{C_{PF}}}\Big) \tag{5.29}
$$

is independent of h, H and β/α.

Proof The estimate (5.26) is valid under Assumptions 1 and 2. It then follows from Remark 2.5, Theorem 3.12 and Assumption 3 that

$$
\begin{aligned}
&\|(p_h, y_h) - (p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})\|_{a\times a}\\
&\qquad \le C_{\maltese}\Big((C_{\dagger}/\sqrt{\alpha}) H \|y_d\|_{L_2(\Omega)} + C_{\spadesuit} C_{\dagger} [1 + (\mathrm{C_{PF}}/\alpha)]^2 H \|(p_h, y_h)\|_{a\times a}\Big)\\
&\qquad \le C_{\maltese}(C_{\dagger}/\sqrt{\alpha})\Big(1 + C_{\spadesuit}[1 + (\mathrm{C_{PF}}/\alpha)]^2 \sqrt{\mathrm{C_{PF}}}\Big) H \|y_d\|_{L_2(\Omega)}.
\end{aligned}
$$

□

Remark 5.12 Note that (5.27) is equivalent to

$$
\lfloor k/2\rfloor \ln q + \frac{3}{2} \ln \kappa(S_h^{\Pi_H} A_h^{\Pi_H}) + \ln(\beta/\alpha) - (1 + d - \tau_d) \ln H \le 0. \tag{5.30}
$$

By examining (5.17), (5.21) and (5.30), we see that the impacts of the mesh-independent quantities, including the condition number $\kappa(S_h^{\Pi_H} A_h^{\Pi_H})$ and the contrast β/α, are mitigated by the natural log function, and the dominating condition on k is roughly (cf. (5.30))

$$
k \ge 2(1 + d - \tau_d) \frac{\ln(1/H)}{\ln(1/q)}
$$

From (5.4) and the relation $|\ln(1 + x)| \approx |x|$ (for $|x|$ small), we can also see that

$$
\ln(1/q) \approx \frac{1}{\kappa(S_h^{\Pi_H} A_h^{\Pi_H})}.
$$

Therefore, provided $\kappa(S_h^{\Pi_H} A_h^{\Pi_H})$ is moderate, we can choose $k = \lceil j \ln(1/H)\rceil$ for a moderate positive integer j. In the case of Examples 1.2 and 1.3, this is true for any H. For other high contrast problems the magnitude of $\kappa(S_h^{\Pi_H} A_h^{\Pi_H})$ can be sensitive to H (cf. [34,Section 5]) and we can use it as an indicator of the effectiveness of our method.

5.5 L_2 Error

We can use a duality argument to obtain an L_2 error estimate.

Observe that Assumption 3 and Lemma 5.5 imply

$$\|(\mathcal{C}_h^{\Pi_H} - \mathcal{C}_{h,k}^{\Pi_H})\boldsymbol{v}\|_{a\times a} \le C_\ddagger H \|\boldsymbol{v}\|_{a\times a} \qquad \forall\, \boldsymbol{v} \in V_H \times V_H, \tag{5.31}$$

where

$$C_\ddagger = C_\spadesuit [1 + (\mathrm{C_{PF}}/\alpha)]^2. \tag{5.32}$$

Theorem 5.13 *Under Assumptions* 1–3*, we have*

$$\|(p_h, y_h) - (p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)} \le C_\flat \sqrt{\beta/\alpha}\, H^2 \|y_d\|_{L_2(\Omega)}, \tag{5.33}$$

where

$$C_\flat = 2C_\dagger C_\sharp C_\ddagger [1 + (\mathrm{C_{PF}}/\alpha)]\sqrt{\mathrm{C_{PF}}/\alpha} \tag{5.34}$$

is independent of h, H and β/α.

Proof Let $(q, z) \in V_h \times V_h$ be defined by

$$\begin{aligned} \mathcal{B}\big((q, z), (r, s)\big) \;&= \int_\Omega (p_h - p_{H,k}^{\mathrm{ms},h}) r \, dx \\ &\quad + \int_\Omega (y_h - y_{H,k}^{\mathrm{ms},h}) s \, dx \qquad \forall\, (r, s) \in V_h \times V_h. \end{aligned} \tag{5.35}$$

It follows from (2.7), (2.8), (2.12) and (5.35) that

$$\begin{aligned} \|(q, z)\|_{a\times a} &\le \|(p_h, y_h) - (p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)} \\ &\quad \sup_{(r,s)\in H_0^1(\Omega)\times H_0^1(\Omega)} \frac{\|(r, s)\|_{L_2(\Omega)\times L_2(\Omega)}}{\|(r, s)\|_{a\times a}} \\ &\le \sqrt{\mathrm{C_{PF}}/\alpha}\, \|(p_h, y_h) - (p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)}. \end{aligned} \tag{5.36}$$

By repeating the arguments in the proof of Theorem 3.14 that led to (3.23), we also have

$$\|\mathcal{C}_h^{\Pi_H}(q, z)\|_{a\times a} \le (C_\dagger/\sqrt{\alpha}) H \|(p_h, y_h) - (p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)}. \tag{5.37}$$

Let $(r, s) = (p_h, y_h) - (p_{H,k}^{\mathrm{ms},h}, y_{H,k}^{\mathrm{ms},h})$. Then the Galerkin relation

$$\mathcal{B}\big(\boldsymbol{v}, (r, s)\big) = 0 \qquad \forall\, \boldsymbol{v} \in V_{H,k}^{\mathrm{ms,h}} \tag{5.38}$$

follows from (2.14) and (5.1).

Combining (3.10), (4.15), (5.35) and (5.38), we obtain

$$\begin{aligned} \|(r, s)\|_{L_2(\Omega)\times L_2(\Omega)}^2 &= \mathcal{B}\big((q, z), (r, s)\big) \\ &= \mathcal{B}\big((q, z) - (\boldsymbol{I} - \mathcal{C}_{h,k}^{\Pi_H})(\Pi_H \times \Pi_H)(q, z), (r, s)\big) \\ &= \mathcal{B}\big((\mathcal{C}_{h,k}^{\Pi_H} - \mathcal{C}_h^{\Pi_H})(\Pi_H \times \Pi_H)(q, z), (r, s)\big) \\ &\qquad + \mathcal{B}\big((q, z) - (\boldsymbol{I} - \mathcal{C}_h^{\Pi_H})(q, z), (r, s)\big) \\ &= \mathcal{B}\big((\mathcal{C}_{h,k}^{\Pi_H} - \mathcal{C}_h^{\Pi_H})(\Pi_H \times \Pi_H)(q, z), (r, s)\big) + \mathcal{B}\big(\mathcal{C}_h^{\Pi_H}(q, z), (r, s)\big). \end{aligned} \tag{5.39}$$

Table 1 Choices for h, H, and j for different values of ϵ

ϵ	0.08				0.04			0.025
$1/h$	256				320			
$1/H$	8	16	32	64	10	20	40	80
j	2	2	3	3	2	2	3	3

Using (2.10), Remark 3.1, (5.31), (5.36), we can bound the first term on the right-hand side of (5.39) by

$$\begin{aligned}
&\mathcal{B}\big((\mathcal{C}_{h,k}^{\Pi_H}-\mathcal{C}_h^{\Pi_H})(\Pi_H\times\Pi_H)(q,z),(r,s)\big)\\
&\quad\le[1+(\mathrm{C_{PF}}/\alpha)]\big\|(\mathcal{C}_{h,k}^{\Pi_H}-\mathcal{C}_h^{\Pi_H})(\Pi_H\times\Pi_H)(q,z)\big\|_{a\times a}\|(r,s)\|_{a\times a}\\
&\quad\le[1+(\mathrm{C_{PF}}/\alpha)]C_\ddagger H\|(\Pi_H\times\Pi_H)(q,z)\big\|_{a\times a}\|(r,s)\|_{a\times a}\\
&\quad\le[1+(\mathrm{C_{PF}}/\alpha)]C_\ddagger C_\dagger\sqrt{\beta/\alpha}H\|(q,z)\|_{a\times a}\|(r,s)\|_{a\times a}\\
&\quad\le[1+(\mathrm{C_{PF}}/\alpha)]C_\ddagger C_\dagger\sqrt{\beta/\alpha}\sqrt{\mathrm{C_{PF}}/\alpha}H\|(r,s)\|_{L_2(\Omega)\times L_2(\Omega)}\|(r,s)\|_{a\times a}.
\end{aligned}\tag{5.40}$$

For the second term on the right-hand side of (5.39), we have the bound

$$\begin{aligned}
\mathcal{B}\big(\mathcal{C}_h^{\Pi_H}(q,z),(r,s)\big)&\le[1+(\mathrm{C_{PF}}/\alpha)]\|\mathcal{C}_h^{\Pi_H}(q,z)\|_{a\times a}\|(r,s)\|_{a\times a}\\
&\le[1+(\mathrm{C_{PF}}/\alpha)](C_\dagger/\sqrt{\alpha})H\|(r,s)\|_{L_2(\Omega)\times L_2(\Omega)}\|(r,s)\|_{a\times a}
\end{aligned}\tag{5.41}$$

by (2.10) and (5.37).

Putting (5.28) and (5.39)–(5.41) together, we arrive at the estimate

$$\begin{aligned}
&\|(p_h,y_h)-(p_{H,k}^{\mathrm{ms},h},y_{H,k}^{\mathrm{ms},h})\|_{L_2(\Omega)\times L_2(\Omega)}\\
&\quad\le[1+(\mathrm{C_{PF}}/\alpha)](C_\dagger/\sqrt{\alpha})\big[C_\ddagger\sqrt{\mathrm{C_{PF}}(\beta/\alpha)}+1\big]H\|(p_h,y_h)-(p_{H,k}^{\mathrm{ms},h},y_{H,k}^{\mathrm{ms},h})\|_{a\times a}\\
&\quad\le[1+(\mathrm{C_{PF}}/\alpha)](C_\dagger/\sqrt{\alpha})\big[C_\ddagger\sqrt{\mathrm{C_{PF}}(\beta/\alpha)}+1\big]C_\sharp H^2\|y_d\|_{L_2(\Omega)},
\end{aligned}$$

which implies (5.33) with $C_\flat$ given by (5.34). □

6 Numerical Results

The numerical results for Examples 1.2 and 1.3 are presented in Sects. 6.1 and 6.2. We also describe briefly some computational aspects in Sect. 6.3.

6.1 Highly Oscillatory Problem

We solve the optimal control problem (1.1) on the unit square $(0,1)\times(0,1)$, where $\gamma=1$, $y_d=-1$, and $\mathcal{A}$ is the matrix in Example 1.2 with $\epsilon=0.08$, 0.04 and 0.025. We use the localized multiscale finite element method from Sect. 5.

For this problem $\alpha\approx 1$ and $\beta\approx 20$, the magnitude of $\kappa(S_h^{\Pi_H}A_h^{\Pi_H})$ is moderate by (5.5) and, according to Remark 5.12, we can choose the number of P-MINRES steps k to be $\lceil j\,\ln(1/H)\rceil$ for a moderate positive integer j. The choices for h, H and j for different values of ϵ are described in Table 1.

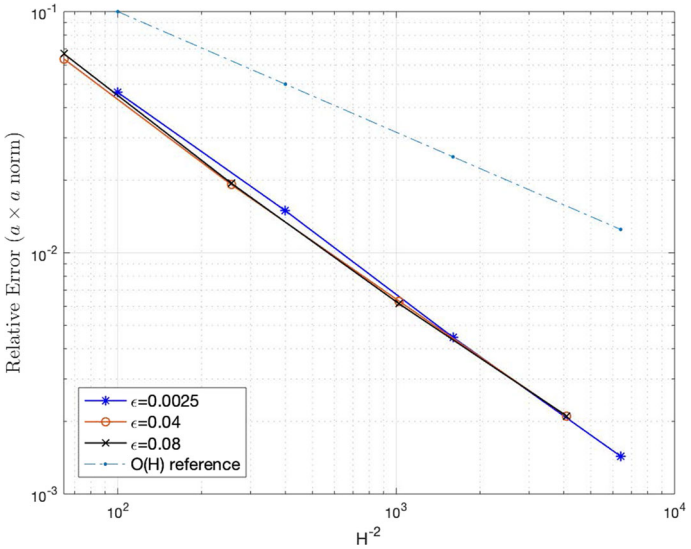

Fig. 9 Relative $\|\cdot\|_{a\times a}$ errors of the localized multiscale approximations of the highly oscillatory problem for different values of ϵ

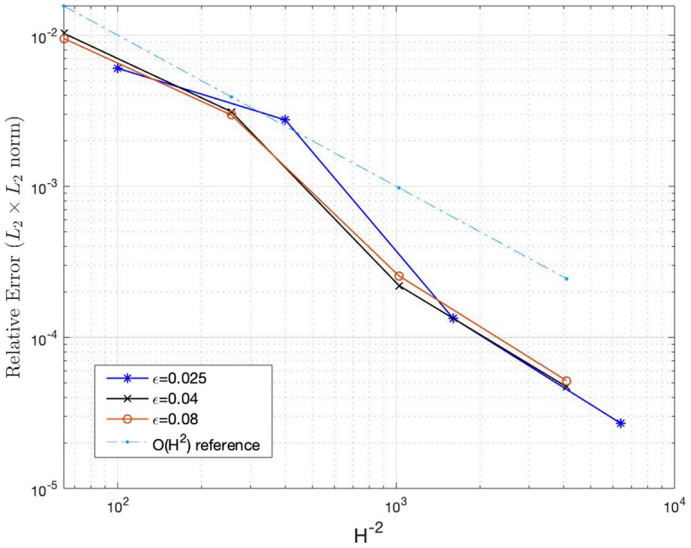

Fig. 10 Relative $\|\cdot\|_{L_2(\Omega)\times L_2(\Omega)}$ errors of the localized multiscale approximations of the highly oscillatory problem for different values of ϵ

The relative errors in the $\|\cdot\|_{a\times a}$ norm and the $\|\cdot\|_{L_2(\Omega)\times L_2(\Omega)}$ norm are presented in Figs. 9 and 10. We observe $O(H)$ convergence in the $\|\cdot\|_{a\times a}$ norm and $O(H^2)$ convergence in the $\|\cdot\|_{L_2(\Omega)\times L_2(\Omega)}$ norm, which agree with Theorems 5.11 and 5.13.

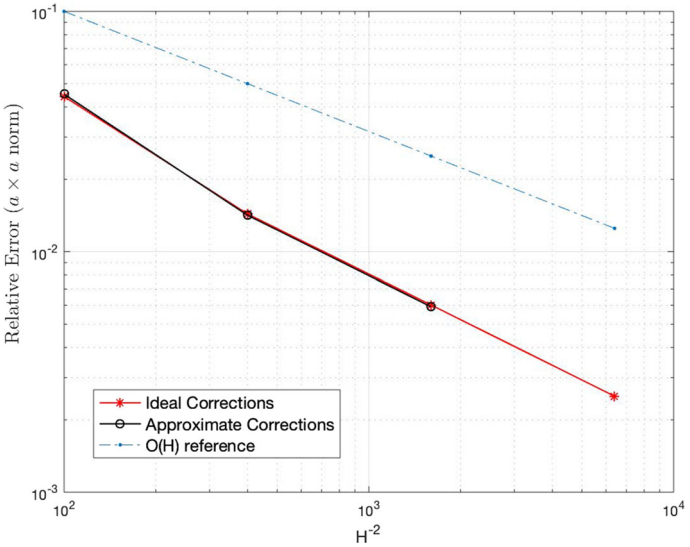

Fig. 11 Relative $\|\cdot\|_{a\times a}$ errors of the multiscale approximations of the highly heterogeneous problem, where $h = 1/320$, $k = \lceil j \ln(1/H)\rceil$, with $j = 2$ for $H = 1/10$, $j = 3$ for $H = 1/20$, and $j = 4$ for $H = 1/40$

6.2 Highly Heterogeneous Problem

We solve the optimal control problem (1.1) on the unit square, where $\gamma = 1$, $y_d = 1$, and $\mathcal{A}$ is the matrix from Example 1.3. We use the ideal multiscale finite element method from Sect. 3 and the localized multiscale finite element method from Sect. 5. The reference mesh size is $h = 1/320$.

For this problem we have $\alpha = 1$ and $\beta = 1331$. The value of the condition number $\kappa(S_h^{\Pi_H} A_h^{\Pi_H})$ is found computationally to be less than 10, which is much better than the pessimistic bound in (5.5). Therefore we can, according to Remark 5.12, choose the number of P-MINRES steps k to be $\lceil j \ln(1/H)\rceil$ for a moderate positive integer j. Here we take $j = 2$ for $H = 1/10$, $j = 3$ for $H = 1/20$, and $j = 4$ for $H = 1/40$.

The relative errors in the $\|\cdot\|_{a\times a}$ norm and the $\|\cdot\|_{L_2(\Omega)\times L_2(\Omega)}$ norm are displayed in Figs. 11 and 12. We observe that the errors for the ideal multiscale finite element method and the localized multiscale finite element method are indistinguishable. The order of convergence in the $\|\cdot\|_{a\times a}$ norm is 1, which agrees with Theorems 3.12 and 5.11. The convergence history for the $\|\cdot\|_{L_2(\Omega)\times L_2(\Omega)}$ norm is similar to the early stage of the history for the highly oscillatory but well-conditioned problem in Fig. 10. Therefore it is reasonable to expect that the order of convergence in the $\|\cdot\|_{L_2(\Omega)\times L_2(\Omega)}$ norm for the ill-conditioned highly heterogenous problem will also approach 2 at a later stage.

6.3 Some Computational Aspects

We will focus on the highly heterogeneous problem in Sect. 6.2, where the reference solution is obtained by a standard finite element method with $h = 1/320$. Below are some observations on the case where the coarse mesh size is $H = 1/20$ and the solution obtained by the localized multiscale method in Sect. 5 is quite reasonable (cf. Fig. 3).

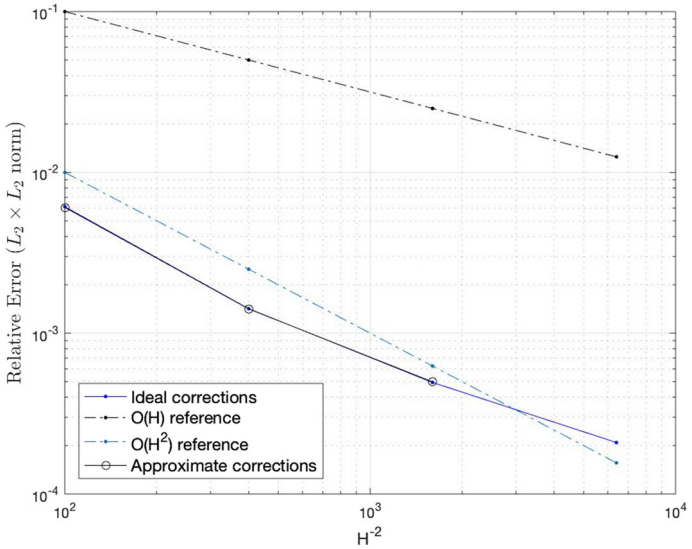

Fig. 12 Relative $\|\cdot\|_{L_2(\Omega)\times L_2(\Omega)}$ errors of multiscale approximations of the highly heterogeneous problem, where $h = 1/320$, $k = \lceil j \ln(1/H)\rceil$, with $j = 2$ for $H = 1/10$, $j = 3$ for $H = 1/20$, and $j = 4$ for $H = 1/40$

The parallel computing was carried out on a cluster with 440 compute nodes running the Red Hat Enterprise Linux 6 operating system and has a 146 TFlops peak performance. Each compute node is equipped with two 8-core Sandy Bridge Xeon 64-bit processors operating at a core frequency of 2.6 GHz, 32GB 1666MHz RAM, 500GB HD, 40 Gigabit/sec Infiniband network interface and a 1 Gigabit Ethernet network interface.

We computed the reference solution with the PETSc library using 128 processors and an ILU(0) preconditioner in PGMRES. The solution time is 0.22 seconds. For comparison, we solved the smaller system (5.1) by Gaussian elimination in MATLAB on a MacBook Pro (2.8 Ghz Quad-Core Intel Core i7 processor and a 16GB 2133 Mhz LPDDR RAM). The solution time is 0.02 seconds.

The total (set-up and solution) time for computing the reference solution with PETSc using 128 processors is 0.47 seconds. For comparison, the total time for solving (5.1) with 128 different right-hand sides simultaneously using PETSc and Gaussian elimination is 1.36 seconds.

Using 1024 processors, it took 2104 seconds in the offline stage to construct the basis functions of the ideal multiscale finite element space $V_H^{\mathrm{ms,h}}$ in Sect. 3.3, and 535 seconds to construct the basis functions of the localized multiscale finite element space $V_{H,k}^{\mathrm{ms,h}}$ in Sect. 4.2.

7 Concluding Remarks

In this paper we have developed multiscale finite element methods for a linear-quadratic elliptic optimal control problem with rough coefficients, where scale separation and periodic structures are not assumed. These methods can be viewed as reduced order methods.

In particular, we have constructed a generalized finite element method with localized basis functions whose performance is similar to standard finite element methods for smooth problems. Both the construction of the generalized finite element space and the analysis of the resulting Galerkin method are based on basic finite element technology and two (by now) classical numerical linear algebra ingredients, namely the additive Schwarz preconditioner and the preconditioned minimum residual algorithm. Our work further illustrates the idea put forth in [28] that multiscale problems can be solved by combining finite element methods, domain decomposition algorithms and iterative Krylov subspace solvers.

The techniques developed in this paper and [6] can be extended to elliptic variational inequalities with rough coefficients, such as the obstacle problem and the optimal control problem with control constraints. They can also be applied to elliptic boundary control problems where we can adopt the LOD techniques for nonhomogeneous boundary conditions in [18].

Acknowledgements Portions of this research were conducted with high performance computing resources provided by Louisiana State University (http://www.hpc.lsu.edu).

Funding Funding is provided by US National Science Foundation (Grant No. DMS-19-13035)

Data Availability The datasets generated during and/or analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Conflict of interest The authors have not disclosed any conflict of interests.

References

1. Abdulle, A.: On a priori error analysis of fully discrete heterogeneous multiscale FEM. Multiscale Model. Simul. **4**, 447–459 (2005)
2. Abdulle, A., E, W., Engquist, B., Vanden-Eijnden, E.: The heterogeneous multiscale method. Acta Numer. **21**, 1–87 (2012)
3. Adams, R.A., Fournier, J.J.F.: Sobolev Spaces, 2nd edn. Academic Press, Amsterdam (2003)
4. Babuška, I.: The finite element method with Lagrange multipliers. Numer. Math. **20**, 179–192 (1973)
5. Babuška, I., Osborn, J.E.: Can a finite element method perform arbitrarily badly? Math. Comput. **69**, 443–462 (2000)
6. Brenner, S.C., Garay, J.C., Sung, L.-Y.: Additive Schwarz preconditioners for a localized orthogonal decomposition method. Electron. Trans. Numer. Anal. **54**, 234–255 (2021)
7. Brenner, S.C., Scott, L.R.: The Mathematical Theory of Finite Element Methods, 3rd edn. Springer, New York (2008)
8. Brezzi, F.: On the existence, uniqueness and approximation of saddle-point problems arising from Lagrangian multipliers. RAIRO Anal. Numer. **8**, 129–151 (1974)
9. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, Amsterdam (1978)
10. Dryja, M., Widlund, O.B.: An additive variant of the Schwarz alternating method in the case of many subregions. Technical Report 339. Department of Computer Science, Courant Institute (1987)
11. Engquist, B., Li, X., Ren, W., Vanden-Eijnden, E.: The heterogeneous multiscale methods. Commun. Math. Sci. **1**, 87–132 (2003)
12. Ela, W., Ming, P., Zhang, P.: Analysis of the heterogeneous multiscale method for elliptic homogenization problems. J. Am. Math. Soc. **18**, 121–156 (2005)
13. Efendiev, Y., Hou, T.Y.: Multiscale Finite Element Methods. Springer, New York (2009)
14. Elman, H.C., Silvester, D.J., Wathen, A.J.: Finite Elements and Fast Iterative Solvers: with Applications in Incompressible Fluid Dynamics, 2nd edn. Oxford University Press, Oxford (2014)
15. Ge, L., Yan, N., Wang, L., Liu, W., Yang, D.: Heterogeneous multiscale method for optimal control problem governed by elliptic equations with highly oscillatory coefficients. J. Comput. Math. **36**, 5 (2018)

16. Greenbaum, A.: Iterative Methods for Solving Linear Systems. SIAM, Philadelphia (1997)
17. Hellman, F., Målqvist, A.: Contrast independent localization of multiscale problems. Multiscale Model. Simul. **15**, 1325–1355 (2017)
18. Henning, P., Målqvist, A.: Localized orthogonal decomposition techniques for boundary value problems. SIAM J. Sci. Comput. **36**, A1609–A1634 (2014)
19. Henning, P., Ohlberger, M., Schweizer, B.: An adaptive multiscale finite element method. Multiscale Model. Simul. **12**, 1078–1107 (2014)
20. Henning, P., Peterseim, D.: Oversampling for the multiscale finite element method. Multiscale Model. Simul. **11**, 1149–1175 (2013)
21. Hetmaniuk, U., Klawonn, A.: Error estimates for a two-dimensional special finite element method based on component mode synthesis. Electron. Trans. Numer. Anal. **41**, 109–132 (2014)
22. Hetmaniuk, U.L., Lehoucq, R.B.: A special finite element method based on component mode synthesis. M2AN Math. Model. Numer. Anal. **44**, 401–420 (2010)
23. Hou, T.Y., Wu, X.-H.: A multiscale finite element method for elliptic problems in composite materials and porous media. J. Comput. Phys. **134**, 169–189 (1997)
24. Hou, T.Y., Wu, X.-H., Cai, Z.: Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. Math. Comput. **68**, 913–943 (1999)
25. Hughes, T.J.R.: Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods. Comput. Methods Appl. Mech. Eng. **127**, 387–401 (1995)
26. Hughes, T.J.R., Feijóo, G.R., Mazzei, L., Quincy, J.-B.: The variational multiscale method—a paradigm for computational mechanics. Comput. Methods Appl. Mech. Eng. **166**, 3–24 (1998)
27. Hughes, T.J.R., Sangalli, G.: Variational multiscale analysis: the fine-scale Green's function, projection, optimization, localization, and stabilized methods. SIAM J. Numer. Anal. **45**, 539–557 (2007)
28. Kornhuber, R., Peterseim, D., Yserentant, H.: An analysis of a class of variational multiscale methods based on subspace decomposition. Math. Comput. **87**, 2765–2774 (2018)
29. Lions, J.-L.: Optimal Control of Systems Governed by Partial Differential Equations. Springer, New York (1971)
30. Målqvist, A., Persson, A., Stillfjord, T.: Multiscale differential Riccati equations for linear quadratic regulator problems. SIAM J. Sci. Comput. **40**, A2406–A2426 (2018)
31. Målqvist, A., Peterseim, D.: Localization of elliptic multiscale problems. Math. Comput. **83**, 2583–2603 (2014)
32. Målqvist, A., Peterseim, D.: Numerical Homogenization by Localized Orthogonal Decomposition. SIAM, Philadelphia (2021)
33. Owhadi, H., Scovel, C.: Operator-Adapted Wavelets, Fast Solvers, and Numerical Homogenization. Cambridge University Press, Cambridge (2019)
34. Peterseim, D., Scheichl, R.: Robust numerical upscaling of elliptic multiscale problems at high contrast. Comput. Methods Appl. Math. **16**, 579–603 (2016)
35. Toselli, A., Widlund, O.B.: Domain Decomposition Methods—Algorithms and Theory. Springer, New York (2005)
36. Zeng, J., Chen, Y., Liu, G.: Rough polyharmonic splines method for optimal control problem governed by parabolic systems with rough coefficient. Comput. Math. Appl. **80**, 121–139 (2020)