Electronic Transactions on Numerical Analysis.
Volume 54, pp. 234–255, 2021.

ISSN 1068–9613.
DOI: 10.1553/etna_vol54s234

ETNA
Kent State University and
Johann Radon Institute (RICAM)

ADDITIVE SCHWARZ PRECONDITIONERS FOR A LOCALIZED ORTHOGONAL DECOMPOSITION METHOD*

SUSANNE C. BRENNER†, JOSÉ C. GARAY†, AND LI-YENG SUNG†

Abstract. We investigate a variant of the localized orthogonal decomposition method (Henning and Peterseim, [Multiscale Model. Simul., 11 (2013), pp. 1149–1175] and Målqvist and Peterseim, [Math. Comp., 83 (2014), pp. 2583–2603]) for elliptic problems with rough coefficients. The construction of the basis of the multiscale finite element space is based on domain decomposition techniques, which is motivated by the recent work of Kornhuber, Peterseim, and Yserentant [Math. Comp., 87 (2018), pp. 2765–2774]. We also design and analyze additive Schwarz domain decomposition preconditioners for the resulting discrete problems.

Key words. multiscale, localized orthogonal decomposition, domain decomposition, additive Schwarz

AMS subject classifications. 65N12, 65N30, 65N55

1. Introduction.

Let the bounded open set $\Omega \subset \mathbb{R}^d$ ($d = 1, 2, 3$) be an interval ($d = 1$), a polygonal domain ($d = 2$), or a polyhedral domain ($d = 3$). The continuous problem is to find $u \in H_0^1(\Omega)$ such that

$$a(u, v) = (f, v) \qquad \forall\, v \in H_0^1(\Omega), \tag{1.1}$$

where $f \in L_2(\Omega)$,

$$a(u, v) = \int_\Omega (\mathcal{A}\nabla u) \cdot \nabla v \,\mathrm{d}x, \quad \text{and} \quad (f, v) = \int_\Omega f v \,\mathrm{d}x.$$

We assume that (1.1) is an elliptic problem with rough coefficients, i.e., the components of the symmetric diffusion matrix $\mathcal{A}$ only belong to $L_\infty(\Omega)$ and the eigenvalues of $\mathcal{A}$ are bounded below (resp., above) by the positive constant α (resp., β). We will use $\|\cdot\|_a$ to denote the energy norm $\sqrt{a(\cdot,\cdot)}$.

REMARK 1.1. Throughout the paper we will follow the standard notation for differential operators, function spaces, and (semi-)norms that can be found, for example, in [3, 7, 8].

Note that

$$\sqrt{\alpha}\,|v|_{H^1(\Omega)} \le \|v\|_a \le \sqrt{\beta}\,|v|_{H^1(\Omega)} \qquad \forall\, v \in H^1(\Omega), \tag{1.2}$$

and hence, by (1.1) and a standard Poincaré-Friedrichs inequality,

$$\|u\|_a \le (C_{\mathrm{PF}}/\sqrt{\alpha}) D \|f\|_{L_2(\Omega)}, \tag{1.3}$$

where $D = \operatorname{diam}\Omega$ and the positive constant C_{PF} only depends on the shape of Ω.

Let $\mathcal{T}_H$ be a quasi-uniform simplicial (or quadrilateral/hexahedral) triangulation of Ω and $V_H \subset H_0^1(\Omega)$ be the (coarse scale) P_1 (or Q_1) finite element space associated with $\mathcal{T}_H$. The fine scale P_1 (or Q_1) finite element space $V_h \subset H_0^1(\Omega)$ is associated with a uniform refinement $\mathcal{T}_h$ of $\mathcal{T}_H$.

We assume that (i) the approximate solution $u_h \in V_h$ defined by

$$a(u_h, v) = (f, v) \qquad \forall\, v \in V_h \tag{1.4}$$

*Received February 6, 2020. Accepted January 15, 2021. Published online on March 17, 2021. Recommended by O. Widlund.
†Department of Mathematics and Center for Computation & Technology, Louisiana State University, Baton Rouge, LA 70803, USA ({brenner, sung}@math.lsu.edu, jgaray@cct.lsu.edu).

ETNA
Kent State University and
Johann Radon Institute (RICAM)

is a good approximation of u, (ii) the computation of u_h is too expensive, and (iii) the approximate solution $u_H \in V_H$ defined by

$$a(u_H, v) = (f, v) \qquad \forall\, v \in V_H \tag{1.5}$$

is a poor approximation of u. Therefore, we need a generalized finite element space that bridges the two scales.

Elliptic problems with rough coefficients such as (1.1) appear in many multiscale problems. There is a large body of literature on the numerical solutions of these problems by finite element methods, which includes the variational multiscale method (cf. [21, 22, 23] and the references therein), the multiscale finite element method (cf. [13, 19, 20] 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. [17, 18] and the references therein).

Here we will consider the method of Localized Orthogonal Decomposition (LOD) in [16, 25] that does not require periodic structure or scale separation and at the same time can deliver optimal convergence without any elliptic regularity assumption. It is based on the observation that if $\tilde{u}_H$ is the Galerkin approximation defined by (1.5) with V_H replaced by a generalized finite element space $\tilde{V}_H$, then

$$a(u - \tilde{u}_H, u - \tilde{u}_H) = a(u, u - \tilde{u}_H) = (f, u - \tilde{u}_H),$$

and hence

$$\|u - \tilde{u}_H\|_a \leq \|f\|_{L_2(\Omega)} \Big(\frac{\|u - \tilde{u}_H\|_{L_2(\Omega)}}{\|u - \tilde{u}_H\|_a} \Big). \tag{1.6}$$

Therefore, if we can modify the basis functions of V_H so that the resulting $\tilde{V}_H$ is orthogonal (with respect to $a(\cdot,\cdot)$) to a subspace of V_h consisting of highly oscillatory functions, then $u - \tilde{u}_H$, which is orthogonal to $\tilde{V}_H$, is highly oscillatory, and the quotient on the right-hand side of (1.6) would be small. It was established in [25] that such modifications can be obtained by solving local problems and that the magnitude of the quotient on the right-hand side of (1.6) is $O(H)$. This is in great contrast to the fact that the convergence of the standard finite element method based on V_H can be arbitrarily slow for problems with rough coefficients [4].

More precisely, we will consider a variant of LOD motivated by the approach in [24], where the modification of the basis functions is based on domain decomposition techniques. Our key finding is that this variant of LOD has similar properties as the original version in [25] and that standard domain decomposition preconditioners for (1.5) are also effective for this variant of the LOD. The success of this method in bridging the two scales is illustrated in Figure 5.2.

The rest of the paper is organized as follows. We present the variant of LOD in Section 2 and its analysis in Section 3. Preconditioning is discussed in Section 4, and numerical results are presented in Section 5 for a problem whose coefficients exhibit high contrast and in Section 6 for a problem with highly oscillatory coefficients. We end with some concluding remarks in Section 7. For the convenience of the readers, we also include some standard results in Appendix A so that this paper is self-contained with respect to LOD.

All the constants in this paper only depend on the shape regularity of $\mathcal{T}_H$ unless specified otherwise. The dependence on α, β, h, and H are always explicitly stated. To avoid the proliferation of constants, we also use the notation $A \lesssim B$ ($B \gtrsim A$) to represent the statement that $A \leq (\text{constant}) \times B$, where the constant only depends on the shape regularity of $\mathcal{T}_H$. The notation $A \approx B$ is equivalent to the statement that $A \lesssim B$ and $B \lesssim A$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

2. Multiscale methods based on orthogonal decomposition. We will denote the dimension of $\mathcal{T}_H$ (resp., $\mathcal{T}_h$) by m (resp., n).

2.1. The ideal multiscale finite element space $V_H^{\mathrm{ms,h}}$. The construction of $V_H^{\mathrm{ms,h}}$ involves an operator $\Pi : H_0^1(\Omega) \longrightarrow V_H$. Following the suggestion in [24], the function $\Pi v \in V_H$ is defined by taking the nodal average of the local L_2-orthogonal projections of v into the space of linear/bilinear polynomials. More precisely,

$$(\Pi v)(p) = \frac{1}{|\mathcal{T}_p|} \sum_{T \in \mathcal{T}_p} \tilde{v}_T(p), \tag{2.1}$$

where p is any interior vertex of $\mathcal{T}_H$, $\mathcal{T}_p$ is the set of the elements in $\mathcal{T}_H$ that share p as a common vertex, and $\tilde{v}_T \in P_1(T)$ (or $Q_1(T)$) is determined by

$$\int_T \tilde{v}_T w \, \mathrm{d}x = \int_T v w \, \mathrm{d}x \qquad \forall\, w \in P_1(T) \text{ (or } Q_1(T)). \tag{2.2}$$

We have

$$\Pi v = v \qquad \forall\, v \in V_H, \tag{2.3}$$

and, by a direct calculation and the Bramble-Hilbert lemma (cf. Appendix A),

$$H^{-1}\|v - \Pi v\|_{L_2(\Omega)} + |\Pi v|_{H^1(\Omega)} \leq C_\dagger |v|_{H^1(\Omega)} \qquad \forall\, v \in H_0^1(\Omega), \tag{2.4}$$

where $C_\dagger$ only depends on the shape regularity of $\mathcal{T}_H$. In view of (1.2), we also have

$$\|\Pi v\|_a \leq C_\dagger \sqrt{\beta/\alpha}\, \|v\|_a \qquad \forall\, v \in H_0^1(\Omega). \tag{2.5}$$

We will denote by K_h^Π the kernel of the restriction of Π to V_h, i.e.,

$$K_h^\Pi = \{v \in V_h : \Pi v = 0\}.$$

REMARK 2.1. 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. Figure 2.1 for a two-dimensional example with the Q_1 finite element). Then $\{(I - \Pi)\varphi_1, \ldots, (I - \Pi)\varphi_\ell\}$ is a basis of K_h^Π. Indeed, if

$$0 = \sum_{j=1}^{\ell} c_j (I - \Pi)\varphi_j = \sum_{j=1}^{\ell} c_j \varphi_j - \Pi \sum_{j=1}^{\ell} c_j \varphi_j,$$

then the function $\Pi \sum_{j=1}^{\ell} c_j \varphi_j = 0$ because it is a function in V_H that vanishes at the nodes of V_H. It follows that $\sum_{j=1}^{\ell} c_j \varphi_j = 0$ and hence $c_1 = \cdots = c_\ell = 0$.

Let the operator $\mathfrak{C}_h : V_h \longrightarrow K_h^\Pi$ be the orthogonal projection with respect to $a(\cdot, \cdot)$, i.e.,

$$a(\mathfrak{C}_h v, w) = a(v, w) \qquad \forall\, v \in V_h, w \in K_h^\Pi. \tag{2.6}$$

There are two obvious relations between Π and $\mathfrak{C}_h$:

$$\Pi \mathfrak{C}_h = 0 \qquad \text{on } V_h, \tag{2.7}$$

$$(I - \mathfrak{C}_h)(I - \Pi) = 0 \qquad \text{on } V_h. \tag{2.8}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

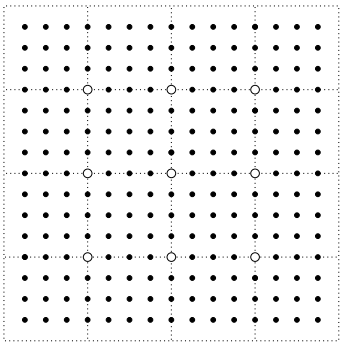

FIG. 2.1. *Nodes for the Q_1 finite element spaces: the nodes for $\phi_1, \ldots, \phi_m \in V_H$ are represented by the circles and the nodes for $\varphi_1, \ldots, \varphi_\ell \in V_h$ are represented by the solid dots.*

The multiscale finite element space $V_H^{\mathrm{ms,h}} \subset V_h$ is the orthogonal complement of K_h^Π with respect to the bilinear form $a(\cdot,\cdot)$, i.e.,

$$V_H^{\mathrm{ms,h}} = \{v \in V_h : a(v,w) = 0 \quad \forall\, w \in K_h^\Pi\}.$$

Note that $V_H^{\mathrm{ms,h}}$ has the same dimension as V_H and

(2.9) $\qquad I - \mathfrak{C}_h : V_h \longrightarrow V_H^{\mathrm{ms,h}}$ is the orthogonal projection with respect to $a(\cdot,\cdot)$.

The restriction of $I - \mathfrak{C}_h$ to V_H is an isomorphism between V_H and $V_H^{\mathrm{ms,h}}$ because the restriction of $\Pi(I - \mathfrak{C}_h)$ to V_H is the identity operator by (2.3) and (2.7). Therefore, if $\{\phi_1, \ldots, \phi_m\}$ is the nodal basis of V_H, then $\{(I - \mathfrak{C}_h)\phi_1, \ldots, (I - \mathfrak{C}_h)\phi_m\}$ is a basis of $V_H^{\mathrm{ms,h}}$.

The approximate solution $u_H^{\mathrm{ms},h} \in V_H^{\mathrm{ms,h}}$ of (1.4) is defined by

$$a(u_H^{\mathrm{ms},h}, v) = (f, v) \qquad \forall\, v \in V_H^{\mathrm{ms,h}}. \tag{2.10}$$

It follows from (1.4), (2.9), and (2.10) that

$$u_H^{\mathrm{ms},h} = (I - \mathfrak{C}_h)u_h, \qquad \text{or equivalently,} \qquad u_h - u_H^{\mathrm{ms},h} = \mathfrak{C}_h u_h, \tag{2.11}$$

from which we can derive the following error estimates (cf. [24, 25] and Appendix A):

$$\|u_h - u_H^{\mathrm{ms},h}\|_a \le (C_\dagger/\sqrt{\alpha})H\|f\|_{L_2(\Omega)}, \tag{2.12}$$

$$\|u_h - u_H^{\mathrm{ms},h}\|_{L_2(\Omega)} \le (C_\dagger^2/\alpha)H^2\|f\|_{L_2(\Omega)}, \tag{2.13}$$

where $C_\dagger$ is the constant that appears in (2.4).

One can treat the ideal multiscale finite element method as a reduced-order method where the functions $\psi_1 = \mathfrak{C}_h\phi_1, \ldots, \psi_m = \mathfrak{C}_h\phi_m$ that correct the original basis functions $\phi_1, \ldots, \phi_m$ of V_H can be computed in parallel and off-line. However, the computation of ψ_i involves solving the equation

$$a(\psi_i, w) = a(\phi_i, w) \qquad \forall\, w \in K_h^\Pi, \tag{2.14}$$

which is expensive. Fortunately, the function ψ_i decays exponentially (cf. [25]), and it is possible to replace ψ_i by a localized version. A two-dimensional example of a Q_1 finite element basis function ϕ_i and the corrected basis function $\phi_i - \psi_i$ is provided in Figure 2.2, where Ω is the unit square $(0,1) \times (0,1)$, $h = 1/160$, and $H = 1/20$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

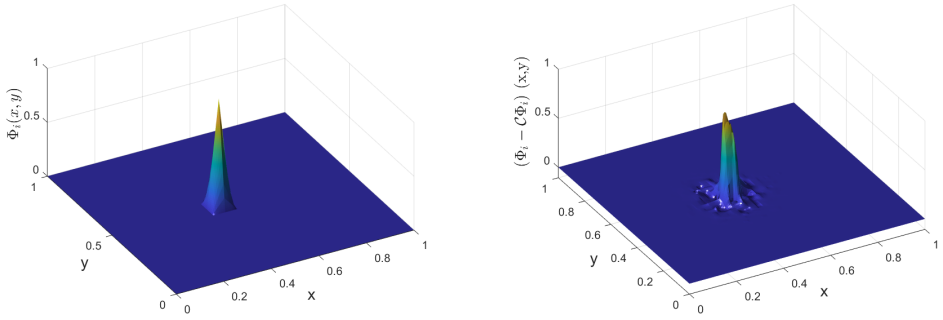

FIG. 2.2. *ϕ_i (left) and $\phi_i - \psi_i$ (right), with $h = 1/160$ and $H = 1/20$.*

2.2. A localized multiscale finite element space $V_{H,k}^{\mathrm{ms,h}}$. We will replace ψ_i by the approximation $\psi_{i,k}$ obtained by k steps of a preconditioned conjugate gradient (PCG) algorithm (cf. [14, 28]) applied to the equation (2.14) with 0 as the initial guess.

We use an additive Schwarz preconditioner (cf. [9]) based on the subspaces K_i $(1 \leq i \leq m)$ of K_h^{Π} defined by

$$K_i = \{(I - \Pi)v : v \in V_h \text{ and } v \text{ vanishes outside } \omega_{x_i}\}. \tag{2.15}$$

Here ω_{x_i} is the union of the elements in $\mathcal{T}_H$ that share the interior node x_i as a common vertex (cf. Figure 2.3 for a two-dimensional example with the Q_1 element). The functions in K_i are supported on the patch $\tilde{\omega}_{x_i}$ generated by adding one layer of elements in $\mathcal{T}_H$ to ω_{x_i} (cf. Figure 2.3).

REMARK 2.2. Let $\varphi_{i,1}, \ldots, \varphi_{i,m_i}$ be the nodal basis functions of V_h that vanish outside ω_{x_i} and at x_i (cf. Figure 2.3). Then, as in Remark 2.1, $\{(I - \Pi)\varphi_{i,1}, \ldots, (I - \Pi)\varphi_{i,m_i}\}$ is a basis of K_i.

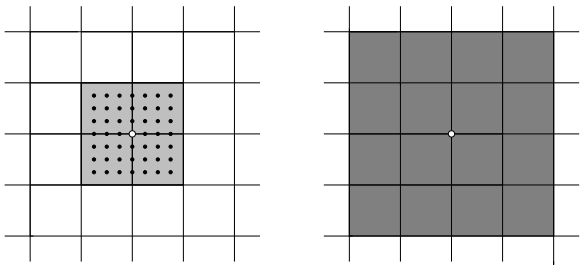

FIG. 2.3. *The patches ω_{x_i} (left) and $\tilde{\omega}_{x_i}$ (right), where the node x_i is represented by the circle and the nodes for $\varphi_{i,1}, \ldots, \varphi_{i,m_i}$ are represented by the solid dots.*

The preconditioned operator $P : K_h^{\Pi} \longrightarrow K_h^{\Pi}$ for (2.14) is given by (cf. [26, 30, 31])

$$P = \sum_{j=1}^{m} P_j, \tag{2.16}$$

where P_j is the orthogonal projection from K_h^{Π} onto K_j with respect to the bilinear form $a(\cdot, \cdot)$. Note that the support of Pv is a subset of the union of all the $\tilde{\omega}_{x_j}$'s whose intersections

ETNA
Kent State University and
Johann Radon Institute (RICAM)

with the support of v have nonempty interiors. Since the function ϕ_i on the right-hand side of (2.14) is supported on ω_{x_i}, at the end of k PCG steps the function $\psi_{i,k}$ is supported in a patch around x_i (with respect to $\mathcal{T}_H$) whose diameter is proportional to k. Thus, $\psi_{i,k}$ is a localized version of ψ_i.

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

$$\mathfrak{C}_{h,k}\phi_i = \psi_{i,k}, \qquad \text{for } 1 \leq i \leq m. \tag{2.17}$$

Then $\mathfrak{C}_{h,k}$ is a localized version of $\mathfrak{C}_h$, and we define $V_{H,k}^{\mathrm{ms,h}}$ to be the subspace of V_h spanned by the functions $\phi_i - \psi_{i,k} = (I - \mathfrak{C}_{h,k})\phi_i$, for $1 \leq i \leq m$.

Again we have an obvious relation

$$\Pi\mathfrak{C}_{h,k} = 0 \qquad \text{on} \quad V_H \tag{2.18}$$

that together with (2.3) implies that $\Pi(I - \mathfrak{C}_{h,k})$ is the identity operator on V_H. Hence, the operator $I - \mathfrak{C}_{h,k} : V_H \longrightarrow V_{H,k}^{\mathrm{ms,h}}$ is an isomorphism, and the dimension of $V_{H,k}^{\mathrm{ms,h}}$ is identical to the dimension of V_H.

The approximate solution $u_{H,k}^{\mathrm{ms},h} \in V_{H,k}^{\mathrm{ms,h}}$ is then defined by

$$a(u_{H,k}^{\mathrm{ms},h}, w) = (f, w) \qquad \forall\, w \in V_{H,k}^{\mathrm{ms,h}}. \tag{2.19}$$

REMARK 2.3. The matrix representing $a(v, w)$ with respect to the basis $\{\phi_i - \psi_{i,k}\}_{i=1}^m$ of $V_{H,k}^{\mathrm{ms,h}}$ is less sparse than the matrix representing $a(v, w)$ with respect to the basis $\{\phi_i\}_{i=1}^m$ of V_H.

3. Analysis of the localized multiscale finite element method. First we observe that (2.4) implies

$$H^{-1}\|v\|_{L_2(\Omega)} = H^{-1}\|v - \Pi v\|_{L_2(\Omega)} \leq C_\dagger |v|_{H^1(\Omega)} \qquad \forall\, v \in K_h^{\Pi}. \tag{3.1}$$

Therefore, the one-level additive Schwarz preconditioner based on $K_1, \ldots, K_m$ in fact behaves like a two-level additive Schwarz preconditioner with generous overlap, as demonstrated in the analysis below.

The condition number of the preconditioned operator P in (2.16) is determined by the formulas (cf. [5, 7, 26, 30, 31])

$$\begin{aligned}
\lambda_{\max}[P] &= \max_{v \in K_h^{\Pi}} \frac{a(v,v)}{\min\limits_{\substack{v = \sum_{j=1}^m v_j \\ v_j \in K_j}} \sum\limits_{j=1}^m a(v_j, v_j)}, \\
\lambda_{\min}[P] &= \min_{v \in K_h^{\Pi}} \frac{a(v,v)}{\min\limits_{\substack{v = \sum_{j=1}^m v_j \\ v_j \in K_j}} \sum\limits_{j=1}^m a(v_j, v_j)}.
\end{aligned} \tag{3.2}$$

Since we are using an overlapping domain decomposition preconditioner, we have a standard estimate (cf. [7, 26, 30, 31])

$$\lambda_{\max}[P] \leq C_\flat, \tag{3.3}$$

where $C_\flat$ is determined by the maximum number of the subspaces $K_1, \ldots, K_m$ that have a nonempty intersection, which in turn is determined by the shape regularity of $\mathcal{T}_H$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

We can construct a partition of unity (cf. [31, Section 3.2]) consisting of nonnegative Lipschitz continuous functions $\rho_1, \ldots, \rho_m$ such that

$$\sum_{j=1}^{m} \rho_j(x) = 1 \qquad \forall\, x \in \bar{\Omega}, \tag{3.4}$$

$$\rho_j \text{ vanishes outside } \omega_{x_j}, \tag{3.5}$$

$$\|\nabla \rho_j\|_{L_\infty(\Omega)} \leq C_\rho / \mathrm{diam}(\omega_{x_j}), \tag{3.6}$$

where C_ρ only depends on the shape regularity of $\mathcal{T}_H$.

Given any $v \in K_h^\Pi$, we define

$$v_j = (I - \Pi) I_h(v \rho_j),$$

where $I_h : C(\bar{\Omega}) \longrightarrow V_h$ is the nodal interpolation operator associated with $\mathcal{T}_h$. Then, $v_j \in K_j$ by (2.15) and (3.5), and

$$\sum_{j=1}^{m} v_j = (I - \Pi) I_h\Big(v \sum_{j=1}^{m} \rho_j\Big) = (I - \Pi) v = v$$

by (3.4). Furthermore, by (1.2), (3.1), (3.6) and direct calculations (cf. [31, Section 3.6], [26, Section 2.5.3], or [7, Section 7.4]), we have,

$$\begin{aligned}
\sum_{j=1}^{m} a(v_j, v_j) &\leq \beta \sum_{j=1}^{m} |(I - \Pi) I_h(v \rho_j)|^2_{H^1(\Omega)} \\
&\lesssim \beta \sum_{j=1}^{m} |I_h(v \rho_j)|^2_{H^1(\Omega)} \\
&\lesssim \beta \sum_{j=1}^{m} \big[|v|^2_{H^1(\omega_{x_j})} + \|\nabla \rho_j\|^2_{L_\infty(\omega_{x_j})} \|v\|^2_{L_2(\omega_{x_j})} \big] \\
&\lesssim \beta \big(|v|^2_{H^1(\Omega)} + H^{-2} \|v\|^2_{L_2(\Omega)} \big) \lesssim \beta |v|^2_{H^1(\Omega)} \lesssim (\beta/\alpha) a(v, v),
\end{aligned} \tag{3.7}$$

which implies through (3.2)

$$\lambda_{\min}[P] \gtrsim (\beta/\alpha)^{-1}. \tag{3.8}$$

Putting (3.3) and (3.8) together, we have

$$\kappa[P] = \frac{\lambda_{\max}[P]}{\lambda_{\min}[P]} \leq C_\sharp (\beta/\alpha), \tag{3.9}$$

where $C_\sharp$ only depends on the shape regularity of $\mathcal{T}_H$. The theory for PCG (cf. [14, 28]) then yields

$$\|\psi_{i,k}\|_a \leq \|\psi_i\|_a = \|\mathfrak{C}_h \phi_i\|_a \leq \|\phi_i\|_a, \tag{3.10}$$

$$\|\psi_i - \psi_{i,k}\|_a \leq \frac{2q^k}{1 + q^{2k}} \|\psi_i\|_a \leq \frac{2q^k}{1 + q^{2k}} \|\phi_i\|_a, \tag{3.11}$$

where

$$q = \frac{\sqrt{\kappa[P]} - 1}{\sqrt{\kappa[P]} + 1} \leq \frac{\sqrt{C_\sharp(\beta/\alpha)} - 1}{\sqrt{C_\sharp(\beta/\alpha)} + 1} \tag{3.12}$$

is bounded away from 1 uniformly with respect to h and H.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

REMARK 3.1. In the case of a large quotient β/α, the number q defined by (3.12) is close to 1, and the performance of the localized multiscale finite element method would be adversely affected by the estimate (3.11). However, there are two mitigating factors. The first factor is that the condition number estimate (3.9) may be too pessimistic. The reason is that the functions v_j that appear in (3.7) belong to the subspaces K_j of K_h^{Π}, and hence the diameters of their supports are of order $O(H)$. Consequently, the effective contrast for such functions can be much smaller than β/α due to averaging. The second factor is that the estimate (3.11) can be improved if the operator P has outlying eigenvalues (cf. [14]).

The following lemma converts the estimates (3.10) and (3.11) into estimates for the operator $\mathfrak{C}_{h,k}$ defined in (2.17), which are useful for the error analysis and the preconditioning of the localized multiscale finite element method. It is based on the observation that

$$|\phi_i|_{H^1(\Omega)} \approx H^{\tau_d}, \quad \text{where} \quad \tau_d = \begin{cases} -\frac{1}{2} & d = 1, \\ 0 & d = 2, \\ \frac{1}{2} & d = 3. \end{cases} \tag{3.13}$$

LEMMA 3.2. *The following estimates hold for the operator* $\mathfrak{C}_{h,k}$:

$$\|\mathfrak{C}_h - \mathfrak{C}_{h,k}\|_a \lesssim \sqrt{\beta/\alpha}\, q^k (H/D)^{-d+\tau_d}, \tag{3.14}$$

$$\|\mathfrak{C}_{h,k}\|_a \lesssim 1 + \sqrt{\beta/\alpha}\, q^k (H/D)^{-d+\tau_d}, \tag{3.15}$$

where $\|\mathfrak{C}_h - \mathfrak{C}_{h,k}\|_a$ *(resp.,* $\|\mathfrak{C}_{h,k}\|_a$*) is the operator norm of* $\mathfrak{C}_h - \mathfrak{C}_{h,k}$ *(resp.,* $\mathfrak{C}_{h,k}$*) induced by the energy norm* $\|\cdot\|_a$ *and* $D = \operatorname{diam}\Omega$.

Proof. We have, by (1.2), (3.11), and (3.13),

$$\begin{aligned} \Big\|(\mathfrak{C}_h - \mathfrak{C}_{h,k}) \sum_{i=1}^m c_i\phi_i\Big\|_a &\le \sum_{i=1}^m |c_i|\, \|\psi_i - \psi_{i,k}\|_a \\ &\lesssim q^k \sum_{i=1}^m |c_i|\, \|\phi_i\|_a \lesssim \sqrt{\beta}\, q^k H^{\tau_d} \sum_{i=1}^m |c_i|, \end{aligned} \tag{3.16}$$

and by a standard discrete estimate together with a standard Poincaré-Friedrichs inequality (and scaling), we also have

$$\begin{aligned} \sum_{i=1}^m |c_i| &\lesssim H^{-d}\Big\|\sum_{i=1}^m c_i\phi_i\Big\|_{L_1(\Omega)} \\ &\lesssim H^{-d} D^{d-\tau_d}\Big|\sum_{i=1}^m c_i\phi_i\Big|_{H^1(\Omega)} \lesssim (1/\sqrt{\alpha}) H^{-d} D^{d-\tau_d}\Big\|\sum_{i=1}^m c_i\phi_i\Big\|_a. \end{aligned} \tag{3.17}$$

The estimate (3.14) follows from (3.16) and (3.17), and the estimate (3.15) follows from (3.14), the triangle inequality, and the fact that $\|\mathfrak{C}_h\|_a = 1$. ∎

Next we derive error estimates for $u_{H,k}^{\mathrm{ms},h}$.

THEOREM 3.3. *The solution* $u_{H,k}^{\mathrm{ms},h}$ *of* (2.19) *satisfies the estimate*

$$\|u_h - u_{H,k}^{\mathrm{ms},h}\|_a \le \|u_h - u_H^{\mathrm{ms},h}\|_a + \big[C_\natural(\beta/\alpha)\|u\|_a\big](H/D)^{-d+\tau_d} q^k, \tag{3.18}$$

where the constant $C_\natural$ *depends only on the shape regularity of* $\mathcal{T}_H$ *and* $D = \operatorname{diam}\Omega$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Proof. Since $(I - \mathfrak{C}_{h,k})\Pi u_h$ belongs to $V_{H,k}^{\mathrm{ms,h}}$, we can use the Galerkin error estimate

$$\|u_h - u_{H,k}^{\mathrm{ms},h}\|_a = \min_{v \in V_{H,k}^{\mathrm{ms,h}}} \|u_h - v\|_a$$

together with (2.8), (2.11), and (3.14) to obtain

$$\begin{aligned}
\|u_h - u_{H,k}^{\mathrm{ms},h}\|_a &\leq \|u_h - (I - \mathfrak{C}_{h,k})\Pi u_h\|_a \\
&= \|[u_h - (I - \mathfrak{C}_h)\Pi u_h] + [(I - \mathfrak{C}_h) - (I - \mathfrak{C}_{h,k})]\Pi u_h\|_a \\
&= \|[u_h - (I - \mathfrak{C}_h)u_h] + (\mathfrak{C}_{h,k} - \mathfrak{C}_h)\Pi u_h\|_a \\
&\leq \|u_h - u_H^{\mathrm{ms},h}\|_a + C_\diamondsuit \sqrt{\beta/\alpha}\, q^k (H/D)^{-d+\tau_d} \|\Pi u_h\|_a.
\end{aligned}$$

The proof is completed by the observation that (1.1), (1.4), and (2.5) imply

$$\|\Pi u_h\|_a \leq C_\dagger \sqrt{\beta/\alpha}\, \|u_h\|_a \leq C_\dagger \sqrt{\beta/\alpha}\, \|u\|_a. \tag{3.19}$$

∎

REMARK 3.4. The dependence on H in the error estimate (3.18) can be eliminated at the expense of increasing the dimension of the localized finite element space (cf. [24]).

It follows from (1.3), (2.12), and (3.18) that

$$\|u_h - u_{H,k}^{\mathrm{ms},h}\|_a \leq (C_\dagger/\sqrt{\alpha}) H \|f\|_{L_2(\Omega)} + C_\natural (C_{\mathrm{PF}}/\sqrt{\alpha})(\beta/\alpha) D \|f\|_{L_2(\Omega)} (H/D)^{-d+\tau_d} q^k.$$

In particular, by choosing k sufficiently large such that

$$q^k \leq (\alpha/\beta)(H/D)^{1+d-\tau_d}, \tag{3.20}$$

we have

$$\|u_h - u_{H,k}^{\mathrm{ms},h}\|_a \leq (1/\sqrt{\alpha})(C_\dagger + C_\natural C_{\mathrm{PF}}) H \|f\|_{L_2(\Omega)}, \tag{3.21}$$

i.e., the performance of the localized multiscale finite element method would be similar to that of the ideal multiscale finite element method.

A standard duality argument then yields

$$\|u_h - u_{H,k}^{\mathrm{ms},h}\|_{L_2(\Omega)} \leq (1/\alpha)(C_\dagger + C_\natural C_{\mathrm{PF}})^2 H^2 \|f\|_{L_2(\Omega)}. \tag{3.22}$$

REMARK 3.5. In the case of a large contrast β/α, the smallest k that satisfies (3.20) would be quite large if both the condition number estimate (3.9) and the error estimate (3.11) are sharp. As mentioned in Remark 3.1, the condition number estimate (3.9) might not be sharp, and the estimate (3.11) can be improved if the operator P has outlying eigenvalues. If the effective value of q in (3.11) is independent of the contrast, then the estimate (3.20) will be satisfied if we choose k to be larger than a multiple of $\log(\beta/\alpha) + \log(D/H)$, a number dominated by $\log(\beta/\alpha)\log(D/H)$ for large β/α and small H/D.

4. Preconditioning the multiscale finite element methods. Both the ideal and the localized multiscale finite element methods are ill-conditioned when H is small, and therefore effective iterative solvers for the discrete problems would require preconditioning.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

4.1. The ideal multiscale finite element method. Let the linear SPD operator $A_H^{\mathrm{ms,h}} : V_H^{\mathrm{ms,h}} \longrightarrow \left[V_H^{\mathrm{ms,h}}\right]'$ be defined by

$$\langle A_H^{\mathrm{ms,h}} v, w\rangle = a(v,w) \qquad \forall\, v, w \in V_H^{\mathrm{ms,h}}, \tag{4.1}$$

where $\langle\cdot,\cdot\rangle$ denotes the canonical bilinear form between a vector space and its dual. Similarly, the linear SPD operator $A_H : V_H \longrightarrow V_H'$ is defined by

$$\langle A_H v, w\rangle = a(v,w) \qquad \forall\, v, w \in V_H. \tag{4.2}$$

Let $Q : V_H \longrightarrow V_H^{\mathrm{ms,h}}$ be the restriction of $I - \mathfrak{C}_h$ to V_H, and $Q^t : [V_H^{\mathrm{ms,h}}]' \longrightarrow V_H'$ be the transpose of Q with respect to the canonical bilinear form, i.e.,

$$\langle Q^t \zeta, v\rangle = \langle \zeta, Qv\rangle \qquad \forall\, \zeta \in [V_H^{\mathrm{ms,h}}]',\ v \in V_H.$$

THEOREM 4.1. *The SPD operator $QA_H^{-1}Q^t : [V_H^{\mathrm{ms,h}}]' \longrightarrow V_H^{\mathrm{ms,h}}$ is an optimal preconditioner of $A_H^{\mathrm{ms,h}}$. More precisely, we have*

$$\kappa\left[(QA_H^{-1}Q^t)A_H^{\mathrm{ms,h}}\right] = \frac{\lambda_{\max}\left[(QA_H^{-1}Q^t)A_H^{\mathrm{ms,h}}\right]}{\lambda_{\min}\left[(QA_H^{-1}Q^t)A_H^{\mathrm{ms,h}}\right]} \le C_\dagger^2(\beta/\alpha), \tag{4.3}$$

where $C_\dagger$ is the constant in (2.4).

Proof. We will use the following Rayleigh quotient formulas (cf. [28]):

$$\lambda_{\max}\left[(QA_H^{-1}Q^t)A_H^{\mathrm{ms,h}}\right] = \lambda_{\max}\left[A_H^{-1}(Q^tA_H^{\mathrm{ms,h}}Q)\right] = \max_{v\in V_H}\frac{\langle A_H^{\mathrm{ms,h}}Qv, Qv\rangle}{\langle A_H v, v\rangle}, \tag{4.4}$$

$$\lambda_{\min}\left[(QA_H^{-1}Q^t)A_H^{\mathrm{ms,h}}\right] = \lambda_{\min}\left[A_H^{-1}(Q^tA_H^{\mathrm{ms,h}}Q)\right] = \min_{v\in V_H}\frac{\langle A_H^{\mathrm{ms,h}}Qv, Qv\rangle}{\langle A_H v, v\rangle}. \tag{4.5}$$

It follows from (2.9), (4.1), and (4.2) that

$$\langle A_H^{\mathrm{ms,h}}Qv, Qv\rangle = a\big((I-\mathfrak{C}_h)v, I-\mathfrak{C}_h)v\big) \le a(v,v) = \langle A_H v, v\rangle \quad \forall\, v \in V_H. \tag{4.6}$$

From (2.3), (2.5), (2.7), (4.1), and (4.2), we also have

$$\begin{aligned}\langle A_H v, v\rangle = a(v,v) &= a\big(\Pi(I-\mathfrak{C}_h)v, \Pi(I-\mathfrak{C}_h)v\big)\\ &\le C_\dagger^2(\beta/\alpha)\, a\big((I-\mathfrak{C}_h)v, (I-\mathfrak{C}_h)v\big)\\ &= C_\dagger^2(\beta/\alpha)\,\langle A_H^{\mathrm{ms,h}}Qv, Qv\rangle \qquad\qquad \forall\, v \in V_H.\end{aligned} \tag{4.7}$$

The estimate (4.3) follows from (4.4)–(4.7). ∎

REMARK 4.2. Note that the estimate (4.7) (and consequently the estimate (4.3)) can be pessimistic. This is due to the fact that the bound (2.5) for functions in $H_0^1(\Omega)$ may be too restrictive for the functions $(I-\mathfrak{C}_h)v \in V_H^{\mathrm{ms,h}}$ that actually appear in (4.7).

COROLLARY 4.3. *We have*

$$\kappa\left[(QBQ^t)A_H^{\mathrm{ms,h}}\right] = \kappa\left[(BA_H)A_H^{-1}(Q^tA_H^{\mathrm{ms,h}}Q)\right] \le C_\dagger^2(\beta/\alpha)\kappa(BA_H),$$

where B is any linear SPD operator that maps V_H' into V_H.

Therefore a good preconditioner for A_H has a similar effect on $A_H^{\mathrm{ms,h}}$ as the preconditioner A_H^{-1}.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

4.2. The localized multiscale finite element method. Here we want to construct a preconditioner for the SPD operator $A_{H,k}^{\mathrm{ms,h}} : V_{H,k}^{\mathrm{ms,h}} \longrightarrow [V_{H,k}^{\mathrm{ms,h}}]'$ defined by

$$\langle A_{H,k}^{\mathrm{ms,h}} v, w\rangle = a(v,w) \qquad \forall\, v, w \in V_{H,k}^{\mathrm{ms,h}}. \tag{4.8}$$

We assume that k is sufficiently large so that (3.20) is satisfied and therefore the error estimates (3.21) and (3.22) are valid.

Let $Q_k = I - \mathfrak{C}_{h,k} : V_H \longrightarrow V_{H,k}^{\mathrm{ms,h}}$ and $Q_k^t : [V_{H,k}^{\mathrm{ms,h}}]' \longrightarrow V_H'$ be the transpose of Q_k with respect to the canonical bilinear forms. We have the following analog of Theorem 4.1.

THEOREM 4.4. *The SPD operator $Q_k A_H^{-1} Q_k^t : [V_{H,k}^{\mathrm{ms,h}}]' \longrightarrow V_{H,k}^{\mathrm{ms,h}}$ is an optimal preconditioner of $A_{H,k}^{\mathrm{ms,h}}$. More precisely, we have*

$$\kappa\left[(Q_k A_H^{-1} Q_k^t) A_{H,k}^{\mathrm{ms,h}}\right] \leq C_\heartsuit(\beta/\alpha), \tag{4.9}$$

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

Proof. Again we use the Rayleigh quotient formulas

$$\lambda_{\max}\left[(Q_k A_H^{-1} Q_k^t) A_{H,k}^{\mathrm{ms,h}}\right] = \max_{v\in V_H} \frac{\langle A_{H,k}^{\mathrm{ms,h}} Q_k v, Q_k v\rangle}{\langle A_H v, v\rangle}, \tag{4.10}$$

$$\lambda_{\min}\left[(Q_k A_H^{-1} Q_k^t) A_{H,k}^{\mathrm{ms,h}}\right] = \min_{v\in V_H} \frac{\langle A_{H,k}^{\mathrm{ms,h}} Q_k v, Q_k v\rangle}{\langle A_H v, v\rangle}. \tag{4.11}$$

According to (3.15), (3.20), (4.2), and (4.8), we have

$$\begin{aligned} \langle A_{H,k}^{\mathrm{ms,h}} Q_k v, Q_k v\rangle &= a\big((I-\mathfrak{C}_{h,k})v, (I-\mathfrak{C}_{h,k})v\big) \\ &\leq 2a(v,v) + 2a(\mathfrak{C}_{h,k}v, \mathfrak{C}_{h,k}v) \lesssim \langle A_H v, v\rangle \qquad \forall\, v \in V_H. \end{aligned} \tag{4.12}$$

In the other direction, we have, by (2.3), (2.5), (2.18), (4.2), and (4.8),

$$\begin{aligned} \langle A_H v, v\rangle = a(v,v) &= a\big(\Pi(I-\mathfrak{C}_{h,k})v, (I-\mathfrak{C}_{h,k})v\big) \\ &\leq C_\dagger^2(\beta/\alpha)\, a\big((I-\mathfrak{C}_{h,k})v, (I-\mathfrak{C}_{h,k})v\big) \\ &= C_\dagger^2(\beta/\alpha)\, \langle A_{H,k}^{\mathrm{ms,h}} Q_k v, Q_k v\rangle \qquad \forall\, v \in V_H. \end{aligned} \tag{4.13}$$

The estimate (4.9) follows from (4.10)–(4.13). ∎

Note that Remark 4.2 also applies to the estimate (4.9).

COROLLARY 4.5. *We have*

$$\begin{aligned} \kappa\left[(Q_k B Q_k^t) A_{H,k}^{\mathrm{ms,h}}\right] &= \kappa\left[(BA_H) A_H^{-1} (Q_k^t A_{H,k}^{\mathrm{ms,h}} Q_k)\right] \\ &\leq \kappa\,[BA_H]\, \kappa\left[(Q_k A_H^{-1} Q_k^t) A_{H,k}^{\mathrm{ms,h}}\right] \leq C_\heartsuit(\beta/\alpha)\kappa\,[BA_H], \end{aligned}$$

where B is any linear SPD operator that maps V_H' into V_H.

Therefore a good preconditioner for A_H has a similar effect on $A_{H,k}^{\mathrm{ms,h}}$ as the preconditioner A_H^{-1}.

5. Numerical results for high contrast coefficients. The domain for the numerical experiments is the unit square $(0,1)\times(0,1)$, and f is the constant function 1. The diffusion matrix $\mathcal{A}$ takes the form of

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

ETNA
Kent State University and
Johann Radon Institute (RICAM)

where $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ are randomly generated piecewise constant functions with respect to a uniform partition of Ω into 40×40 small squares (cf. Figure 5.1). The values of $\mathcal{A}_{11}$ and $\mathcal{A}_{22}$ range between 1 and 1350.

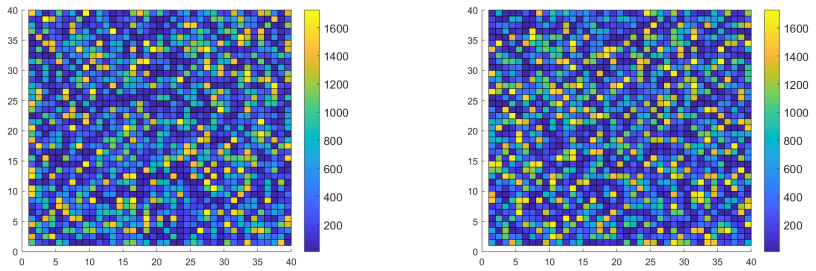

FIG. 5.1. $\mathcal{A}_{11}$ (*left*) *and* $\mathcal{A}_{22}$ (*right*).

5.1. Convergence of the multiscale finite element methods. We take V_h to be the Q_1 finite element space associated with the uniform partition of Ω into 160×160 small squares (i.e., $h = 1/160$). The solution u_h obtained by the standard finite element method (1.4) is treated as the reference solution.

The solutions u_h, $u_H^{\mathrm{ms},h}$, and u_H (with $h = 1/160$ and $H = 1/20$) are depicted in Figure 5.2. It is observed that $u_H^{\mathrm{ms},h}$ provides a good approximation of u_h while u_H suffers from the pre-asymptotic effect and fails to capture u_h.

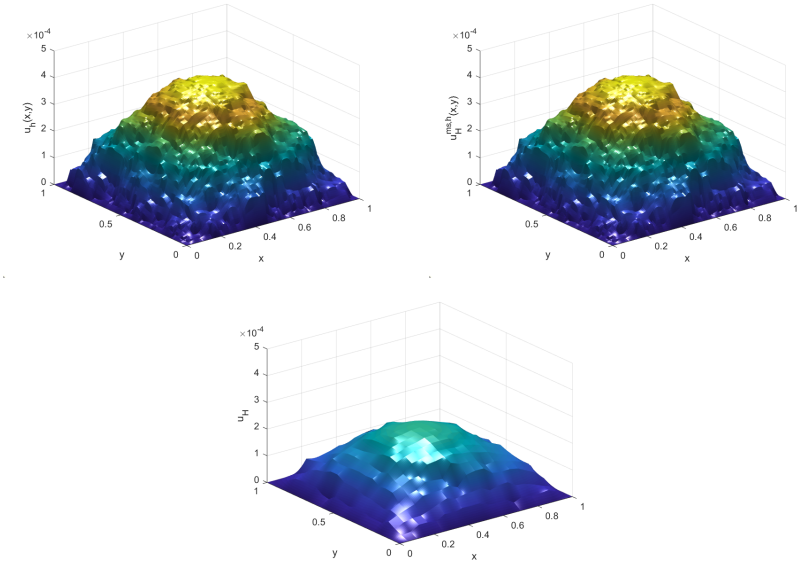

FIG. 5.2. u_h *(top left),* $u_H^{\mathrm{ms},h}$ *(top right) and* u_H *(bottom), with* $h = 1/160$ *and* $H = 1/20$.

The convergence history for the ideal multiscale finite element method and the localized multiscale finite element method are presented in Figure 5.3 for the energy norm and in

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Figure 5.4 for the L_2-norm. In both cases the localized solution $u_{H,k}^{\mathrm{ms},h}$ with $k = \lceil 3\log(1/H)\rceil$

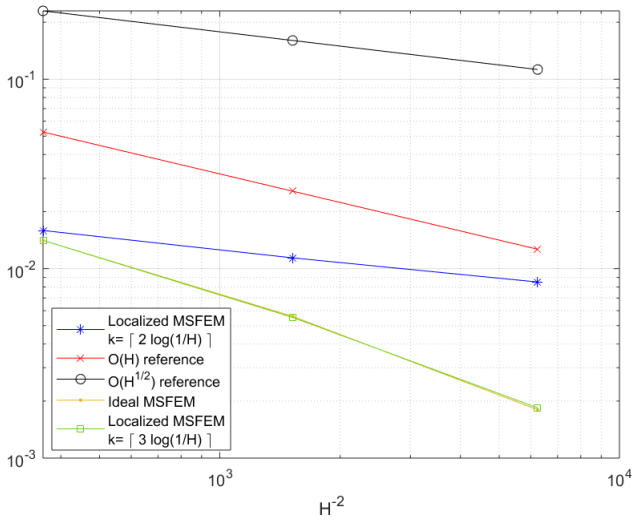

FIG. 5.3. *Relative errors in the energy norm.*

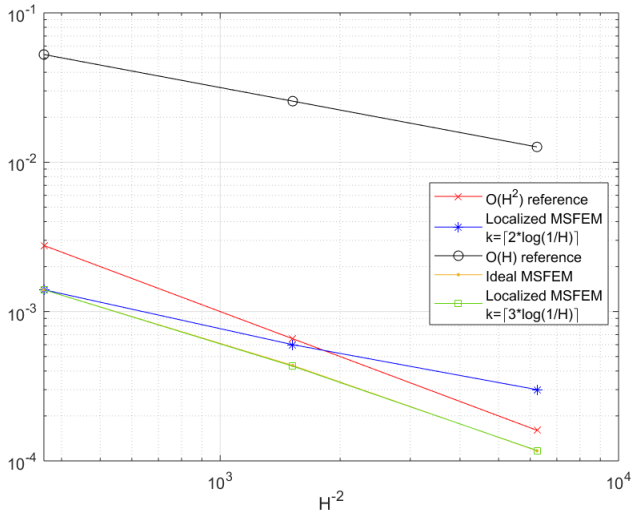

FIG. 5.4. *Relative errors in the L_2-norm.*

and the ideal solution $u_H^{\mathrm{ms},h}$ are almost identical, and their order of convergence agrees with (2.12) and (3.21) in the case of the energy norm and with (2.13) and (3.22) in the case of the L_2-norm. On the other hand, the choice of $k = \lceil 2\log(1/H)\rceil$ results in a reduced order of convergence.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

In order to test the dependence of the performance of the localized multiscale finite element method on the contrast, we have solved (2.19) for three additional randomly generated diffusion matrices, so that the contrast β/α in the four examples increases from the order of 10 to the order of 10^4. The relative errors in the energy norm are displayed in Figure 5.5. The magnitudes are similar for the four examples, and the order of convergence is $O(H)$. The choice of k is $\lceil 3\log(1/H)\rceil$ for the contrasts 10, 1.2×10^2, and 1.3×10, and $\lceil 4\log(1/H)\rceil$ for the contrast 1.4×10^4, which indicates that the effective value of q in (3.11) is likely to be independent of the contrasts (cf. Remark 3.5).

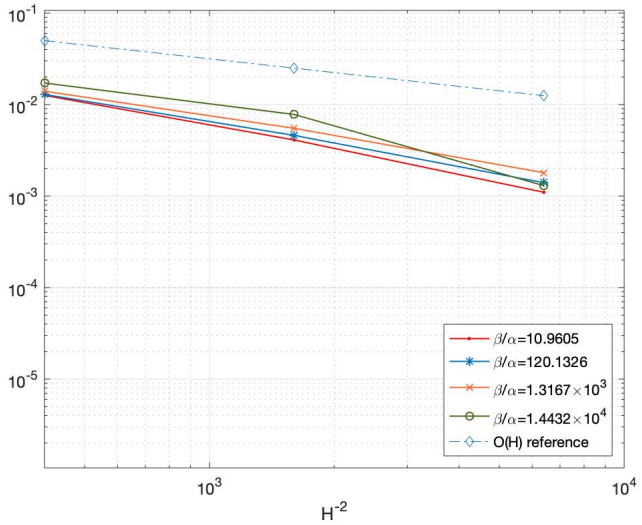

FIG. 5.5. *Relative energy norm errors for the localized multiscale finite element method with respect to four different contrasts.*

5.2. Preconditioning the ideal multiscale finite element method. At first we take $h = 1/160$ and solve (2.10) by the Conjugate Gradient (CG) method (without a preconditioner). The number of iterations required to reduce the residual by a factor of 10^{-6} for various H are reported in Table 5.1.

TABLE 5.1
Iteration counts for CG ($h = 1/160$).

$1/H$	10	20	40	80
Iterations	19	46	122	332

Next we solve the same problem by the Preconditioned Conjugate Gradient (PCG) method with A_H^{-1} as the preconditioner. The number of iterations required to reduce the residual by a factor of 10^{-6} for various H are reported in Table 5.2. The improvement over the CG method and the optimality established in Theorem 4.1 are clearly visible.

We then test the performance of additive Schwarz domain decomposition preconditioners. We divide Ω into J overlapping subdomains with various overlaps measured by the ratio R between the diameter of the subdomains and the amount of overlap. In the case of the two-level

ETNA
Kent State University and
Johann Radon Institute (RICAM)

TABLE 5.2
Iteration counts for PCG with A_H^{-1} as the preconditioner ($h = 1/160$).

$1/H$	10	20	40	80
Iterations	9	13	13	14

TABLE 5.3
Iteration counts for PCG (with one-level and two-level additive Schwarz preconditioners) for (2.10) with $h = 1/320$ and $H = 1/80$.

J \ R	1	2	4	1	2	4
	one-level			two-level		
5^2	33	49	53	31	45	47
10^2	39	48	67	38	36	44
20^2	48	88	119	33	37	42
40^2	103	155	–	33	32	–

preconditioner, the mesh size of the coarse grid is proportional to $1/\sqrt{J}$. We solve (2.10) with $h = 1/320$ and $H = 1/80$ by the PCG method with standard one-level and two-level additive Schwarz preconditioners for A_H as preconditioners for $A_H^{\mathrm{ms,h}}$, and record the number of iterations needed to reduce the residual by a factor of 10^{-6} in Table 5.3.

According to Corollary 4.3, the performance of the PCG for $A_H^{\mathrm{ms,h}}$ is determined by the condition number of BA_H, where B is either the one-level or the two-level additive Schwarz preconditioner for A_H. Therefore, the number of iterations is dictated by the condition number estimates for overlapping additive Schwarz domain preconditioners (cf. [31, Section 3.6], [26, Section 2.5.3], or [7, Section 7.4]). In particular, for a given R, the number of iterations for the one-level method will increase as J increases while the number of iterations for the two-level method will stay bounded. This is observed in Table 5.3.

We also test the weak scalability of the one-level and two-level additive Schwarz methods by fixing a fine mesh with $h = 1/320$ and increasing the number of degrees of freedom for $V_H^{\mathrm{ms,h}}$ and the number of subdomains by a factor of 4 simultaneously, with the initial mesh size $H = 1/40$ and the initial number of subdomains $J = 25$. R is taken to be 8. The results are displayed in Figure 5.6, which again agree with the behavior predicted by Corollary 4.3 and the theory of additive Schwarz preconditioners, i.e, the two-level algorithm is scalable while the one-level algorithm is not.

5.3. Preconditioning the localized multiscale finite element method. We consider the problem (2.19) for the multiscale finite element method with $h = 1/320$, $H = 1/80$, and $k = \lceil 3\log(1/H)\rceil$ and perform the same test for the performance of additive Schwarz domain decomposition preconditioners. The results are identical to the results for the ideal multiscale finite element method reported in Table 5.3.

In order to test the dependence of the performance of the preconditioners on the contrast, we have solved (2.19) by PCG with the two-level preconditioner for the additional randomly generated coefficient matrices in Section 5.1. The mesh size for the localized multiscale finite element method is $H = 1/40$, and the choice of k is $\lceil 3\log(1/H)\rceil$ for the first three contrasts and $\lceil 4\log(1/H)\rceil$ for the largest contrast. The numbers of iterations required to reduce the residual by a factor of 10^{-6} are displayed in Figure 5.7, where the number of subdomains is $J = 5^2, 10^2$, and 20^2, and the ratio between the diameter of the subdomains and the overlap is

ETNA
Kent State University and
Johann Radon Institute (RICAM)

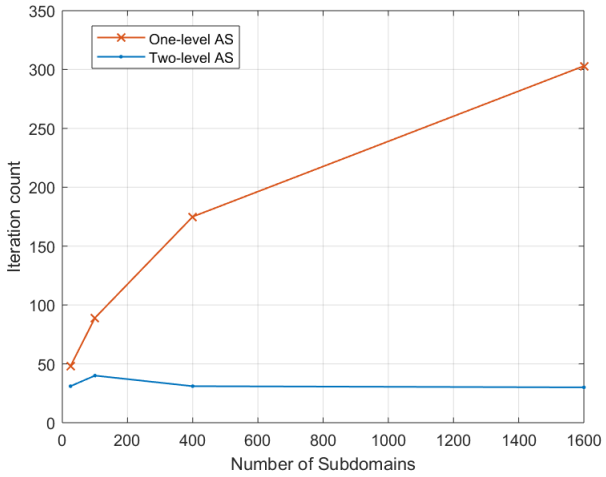

FIG. 5.6. *Iteration counts for the weak scaling test with initial mesh size $H = 1/40$ and initial number of subdomains $J = 25$ ($h = 1/320$, $R = 8$).*

$R = 2$. It can be observed that the number of iterations increases by a factor of at most 2 as the contrast increases by a factor of 10, which indicates that the condition number estimates in Theorem 4.1, Corollary 4.3, Theorem 4.4, and Corollary 4.5 are indeed too pessimistic (cf. Remark 4.2).

6. Numerical results for highly oscillatory coefficients. We consider a problem first treated in [19]. The domain for the numerical experiments is also the unit square $(0,1)\times(0,1)$, and f is the constant function -1. The diffusion matrix $\mathcal{A}$ takes the form of $c(x)I$, where

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

The contrast β/α for this problem is ≈ 16 for all choices of ϵ.

6.1. Convergence of the localized multiscale finite element method. We take V_h to be the Q_1 finite element space associated with the uniform partition of Ω into 256×256 small squares (i.e., $h = 1/256$). The solution u_h obtained by the standard finite element method (1.4) is treated as the reference solution.

The boundary value problem (1.1) is solved by the localized multiscale finite element method (2.19) with $k = \lceil 3\log(1/H)\rceil$ and $H = 1/16$, $1/32$, and $1/64$. The convergence history for the energy (resp., L_2-) norm for $\epsilon = 0.16$, 0.08, and 0.04 are displayed in Figure 6.1 (resp., Figure 6.2). It is observed that the convergence in the energy norm is $O(H)$, which agrees with the estimate (3.21), and the convergence in the L_2-norm is $O(H^2)$, which also agrees with the estimate (3.22). Moreover the performance of the method is independent of ϵ.

6.2. Preconditioning the localized multiscale finite element method. The domain Ω is divided into J subdomains, where the ratio between the diameter of the subdomain and the amount of overlap is denoted by R. In the case of the two-level preconditioner, the mesh size of the coarse grid is proportional to $1/\sqrt{J}$.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

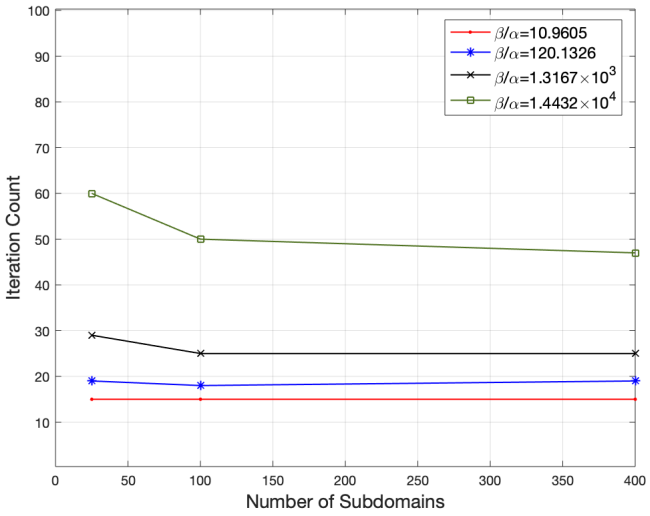

FIG. 5.7. *Iteration counts for the two-level additive Schwarz preconditioner for four different contrasts, where* $H = 1/40$, $R = 2$, *and* $J = 5^2$, 10^2 *and* 20^2.

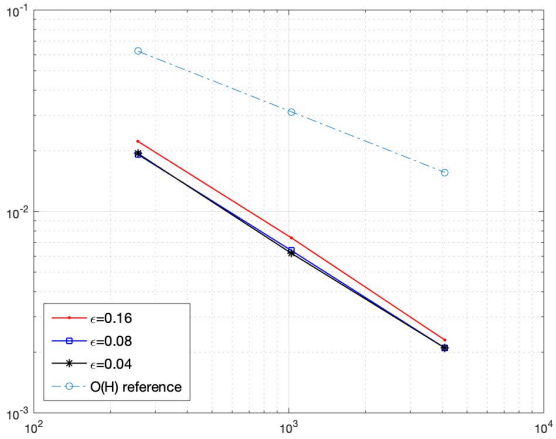

FIG. 6.1. *Relative errors in the energy norm for different* ϵ.

We solve (2.19) by the PCG method with standard one-level and two-level additive Schwarz domain decomposition preconditioners for A_H as preconditioners for $A_{H,k}^{\mathrm{ms,h}}$. For $h = 1/256$ and $H = 1/64$, the number of iterations required to reduce the residual by a factor of 10^{-6} are reported in Table 6.1 for $R = 2$ and Table 6.2 for $R = 4$. The results of solving (2.19) by CG without a preconditioner are also presented in Table 6.3 for comparison.

We observe that the performance of the PCG is more or less independent of ϵ, which is consistent with Corollary 4.5, where the condition number estimate only depends on the contrast. The PCG iteration counts in Table 6.1 and Table 6.2 (especially the ones for the two-level preconditioner) show substantial improvement over the iteration counts in Table 6.3 for the CG algorithm.

We also test the weak scalability of the two-level preconditioner by increasing the number

ETNA
Kent State University and
Johann Radon Institute (RICAM)

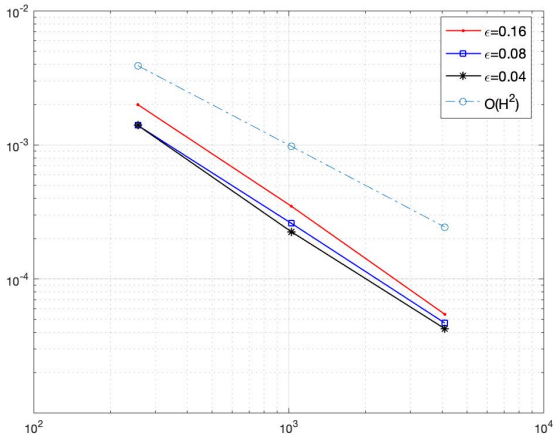

FIG. 6.2. *Relative errors in the L_2-norm for different ϵ.*

TABLE 6.1
Iteration counts for PCG (with one-level and two-level additive Schwarz preconditioners) for (2.19) with $R = 2$, $h = 1/256$, $H = 1/64$ and $\epsilon = 0.16, 0.08,$ and 0.04.

J \ ϵ	0.16	0.08	0.04	0.16	0.08	0.04
	one-level			two-level		
2^2	13	14	15	13	13	15
4^2	15	16	16	15	15	16
8^2	26	24	27	15	15	16
16^2	51	45	45	15	16	15

TABLE 6.2
Iteration counts for PCG (with one-level and two-level additive Schwarz preconditioners) for (2.19) with $R = 4$, $h = 1/256$, $H = 1/64$ and $\epsilon = 0.16, 0.08,$ and 0.04.

J \ ϵ	0.16	0.08	0.04	0.16	0.08	0.04
	one-level			two-level		
2^2	14	15	15	13	14	15
4^2	23	21	24	15	15	16
8^2	42	36	42	16	15	15
16^2	81	70	70	15	16	15

TABLE 6.3
Iteration counts for CG ($h = 1/256$, $H = 1/64$).

ϵ	0.16	0.08	0.04
Iterations	207	149	114

of degrees of freedom in $V_{H,k}^{\mathrm{ms,h}}$ and the number of subdomains by a factor of 4 simultaneously, with the initial mesh size $H = 1/16$ and the initial number of subdomains $J = 4$. The results

ETNA
Kent State University and
Johann Radon Institute (RICAM)

are reported in Figure 6.3. It is clear that the two-level additive Schwarz algorithm is scalable and its performance is independent of ϵ.

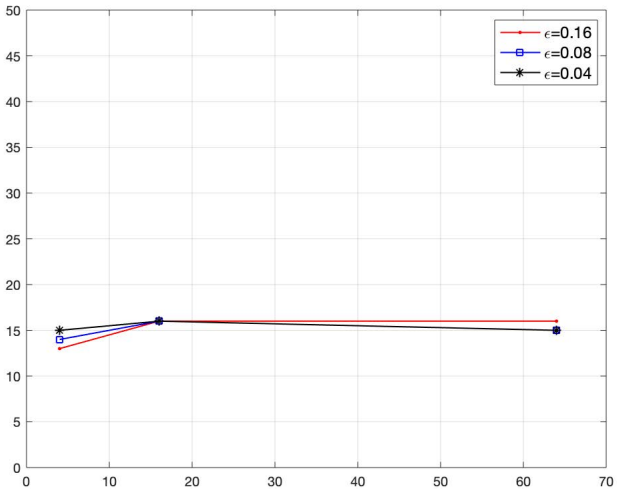

FIG. 6.3. *Iteration counts for the weak scaling test for the two-level preconditioner with initial mesh size* $H = 1/16$ *and initial number of subdomains* $J = 4$ ($h = 1/256$, $R = 2$).

7. Concluding remarks. As a reduced-order method, the variant of the LOD method considered in this paper is effective for applications that involve repeated solves of problems with rough coefficients.

Additive Schwarz domain decomposition preconditioners play a role in both the off-line stage and the on-line stage. They are used in the construction of the localized multiscale finite element space $V_{H,k}^{\mathrm{ms,h}}$, where the subdomains are patches around the nodes of the coarse finite element space V_H. They are also used as preconditioners for the localized multiscale finite element method, where the subdomains are constructed as in overlapping domain decomposition preconditioners for standard second-order problems.

We use a basic interpolation operator Π in this paper. There are more sophisticated coefficient-dependent interpolation operators (cf. [15, 27]) that can lessen the adverse effects of the contrast β/α under additional assumptions on the diffusion matrix $\mathcal{A}$. The approach in this paper can also be applied to LOD methods based on these interpolation operators.

Acknowledgement. This work was supported in part by the National Science Foundation under Grant Nos. DMS-16-20273 and DMS-19-13035. The authors would like to thank two anonymous referees for many helpful comments.

Appendix A. Analysis of the ideal multiscale finite element method. First we provide some details for the estimate (2.4). Let $T \in \mathcal{T}_H$ and ω_T be the interior of the union of all the elements in $\mathcal{T}_H$ that share at least one vertex with T. We will show that

$$H_T^{-1}\|v - \Pi v\|_{L_2(T)} + |\Pi v|_{H^1(T)} \lesssim |v|_{H^1(\omega_T)} \qquad \forall\, v \in H_0^1(\Omega), \tag{A.1}$$

where H_T is the diameter of T, and then (2.4) follows immediately.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Let $v \in H^1(\Omega)$ and $T \in \mathcal{T}_H$ be arbitrary and p be a vertex of T that is interior to Ω. In view of (2.1) and (2.2), we have, by standard inverse estimates (cf. [7, 8]),

$$|(\Pi v)(p)| \lesssim \sum_{T \in \mathcal{T}_p} \|\tilde{v}_T\|_{L_\infty(T)} \lesssim H_T^{-d/2} \sum_{T \in \mathcal{T}_p} \|\tilde{v}_T\|_{L_2(T)} \leq H_T^{-d/2} \sum_{T \in \mathcal{T}_p} \|v\|_{L_2(T)},$$

and hence

$$\|\Pi v\|_{L_2(T)} \lesssim \|v\|_{L_2(\omega_T)}. \tag{A.2}$$

According to the Bramble-Hilbert lemma (cf. [6, 7, 10]), there exists a constant c such that

$$H_T^{-1}\|v - c\|_{L_2(\omega_T)} \lesssim |v|_{H^1(\omega_T)}. \tag{A.3}$$

It follows from (A.2) and (A.3) that

$$\begin{aligned} H_T^{-1}\|v - \Pi v\|_{L_2(T)} &= H_T^{-1}\|(v-c) - \Pi(v-c)\|_{L_2(T)} \\ &\lesssim H_T^{-1}\|v - c\|_{L_2(\omega_T)} \lesssim |v|_{H^1(\omega_T)}. \end{aligned} \tag{A.4}$$

Let $\tilde{\Pi} : H_0^1(\Omega) \longrightarrow V_H$ be a quasi-local operator [7, 29] that satisfies

$$H_T^{-1}\|v - \tilde{\Pi} v\|_{L_2(T)} + |\tilde{\Pi} v|_{H^1(T)} \lesssim |v|_{H^1(\omega_T)} \qquad \forall\, v \in H_0^1(\Omega). \tag{A.5}$$

We can complete the proof of (A.1) by combining (A.4) and (A.5) with a standard inverse estimate:

$$\begin{aligned} |\Pi v|_{H^1(T)} &\leq |\Pi v - \tilde{\Pi} v|_{H^1(T)} + |\tilde{\Pi} v|_{H^1(T)} \\ &\lesssim H_T^{-1}\|\Pi v - \tilde{\Pi} v\|_{L_2(T)} + |v|_{H^1(\omega_T)} \\ &\lesssim H_T^{-1}\|\Pi v - v\|_{L_2(T)} + H_T^{-1}\|v - \tilde{\Pi} v\|_{L_2(T)} + |v|_{H^1(\omega_T)} \lesssim |v|_{H^1(\omega_T)}. \end{aligned}$$

Next we derive the error estimates (2.12) and (2.13). From (1.2), (1.4), (2.4), (2.6), and (2.7), we find

$$\begin{aligned} \|\mathfrak{C}_h u_h\|_a^2 = a(u_h, \mathfrak{C}_h u_h) &= a(u_h, \mathfrak{C}_h u_h - \Pi \mathfrak{C}_h u_h) \\ &= (f, (I - \Pi)\mathfrak{C}_h u_h) \\ &\leq H\|f\|_{L_2(\Omega)} C_\dagger |\mathfrak{C}_h u_h|_{H^1(\Omega)} \leq H\|f\|_{L_2(\Omega)} (C_\dagger/\sqrt{\alpha})\|\mathfrak{C}_h u_h\|_a, \end{aligned}$$

and hence

$$\|\mathfrak{C}_h u_h\|_a \leq (C_\dagger/\sqrt{\alpha}) H\|f\|_{L_2(\Omega)}. \tag{A.6}$$

The estimate (2.12) follows from (2.11) and (A.6).

Finally, we derive the estimate (2.13) by a duality argument. Let $w_h \in V_h$ be defined by

$$a(w_h, v) = (\mathfrak{C}_h u_h, v) \qquad \forall\, v \in V_h. \tag{A.7}$$

Then we have the following analog of (A.6):

$$\|\mathfrak{C}_h w_h\|_a \leq (C_\dagger/\sqrt{\alpha}) H\|\mathfrak{C}_h u_h\|_{L_2(\Omega)}. \tag{A.8}$$

It follows from (2.6) , (A.7), and (A.8) that

$$\begin{aligned} \|\mathfrak{C}_h u_h\|_{L_2(\Omega)}^2 = a(w_h, \mathfrak{C}_h u_h) &= a(\mathfrak{C}_h w_h, \mathfrak{C}_h u_h) \\ &\leq \|\mathfrak{C}_h w_h\|_a \|\mathfrak{C}_h u_h\|_a \leq (C_\dagger/\sqrt{\alpha}) H\|\mathfrak{C}_h u_h\|_{L_2(\Omega)} \|\mathfrak{C}_h u_h\|_a. \end{aligned}$$

ETNA
Kent State University and
Johann Radon Institute (RICAM)

Therefore we have, by (A.6),

$$\|\mathfrak{C}_h u_h\|_{L_2(\Omega)} \leq (C_\dagger/\sqrt{\alpha})H\|\mathfrak{C}_h u_h\|_a \leq (C_\dagger^2/\alpha)H^2\|f\|_{L_2(\Omega)},$$

which is the estimate (2.13) because of (2.11).

REMARK A.1. Since the estimate (A.1) is quasi-local, the results for the multiscale finite element methods can be easily extended to triangulations that are not necessary quasi-uniform (cf. [24, 25]).

REFERENCES

[1] A. ABDULLE, *On a priori error analysis of fully discrete heterogeneous multiscale FEM*, Multiscale Model. Simul., 4 (2005), pp. 447–459.
[2] A. ABDULLE, W. E, B. ENGQUIST, AND E. VANDEN-EIJNDEN, *The heterogeneous multiscale method*, Acta Numer., 21 (2012), pp. 1–87.
[3] R. A. ADAMS AND J. J. F. FOURNIER, *Sobolev Spaces*, 2nd ed., Academic Press, Amsterdam, 2003.
[4] I. BABUŠKA AND J.E. OSBORN, *Can a finite element method perform arbitrarily badly?*, Math. Comp., 69 (2000), pp. 443–462.
[5] P. BJØRSTAD AND J. MANDEL, *On the spectra of sums of orthogonal projections with applications to parallel computing*, BIT, 31 (1991), pp. 76–88.
[6] J.H. BRAMBLE AND S.R. HILBERT, *Estimation of linear functionals on Sobolev spaces with applications to Fourier transforms and spline interpolation*, SIAM J. Numer. Anal., 7 (1970), pp. 113–124.
[7] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, 3rd ed., Springer, New York, 2008.
[8] P. G. CIARLET, *The Finite Element Method for Elliptic Problems,* North-Holland, Amsterdam, 1978.
[9] M. DRYJA AND O. B. WIDLUND, *An additive variant of the Schwarz alternating method in the case of many subregions*, Tech. Report 339, Department of Computer Science, Courant Institute, 1987.
[10] T. DUPONT AND R. SCOTT, *Polynomial approximation of functions in Sobolev spaces*, Math. Comp., 34 (1980), pp. 441–463.
[11] W. E AND B. ENGQUIST, *The heterogeneous multiscale methods*, Commun. Math. Sci., 1 (2003), pp. 87–132.
[12] W. E, P. MING, AND P. ZHANG, *Analysis of the heterogeneous multiscale method for elliptic homogenization problems*, J. Amer. Math. Soc., 18 (2005), pp. 121–156.
[13] Y. EFENDIEV AND T. Y. HOU, *Multiscale Finite Element Methods. Theory and Applications*, Springer, New York, 2009.
[14] A. GREENBAUM, *Iterative Methods for Solving Linear Systems*, SIAM, Philadelphia, 1997.
[15] F. HELLMAN AND A. MÅLQVIST, *Contrast independent localization of multiscale problems*, Multiscale Model. Simul., 15 (2017), pp. 1325–1355.
[16] P. HENNING AND D. PETERSEIM, *Oversampling for the multiscale finite element method*, Multiscale Model. Simul., 11 (2013), pp. 1149–1175.
[17] U. HETMANIUK AND A. KLAWONN, *Error estimates for a two-dimensional special finite element method based on component mode synthesis*, Electron. Trans. Numer. Anal., 41 (2014), pp. 109–132. http://etna.ricam.oeaw.ac.at/vol.41.2014/pp109-132.dir/pp109-132.pdf
[18] U. HETMANIUK AND R. LEHOUCQ, *A special finite element method based on component mode synthesis techniques*, M2AN Math. Model. Numer. Anal., 44 (2010), pp. 401–420.
[19] T. Y. HOU AND X.-H. WU, *A multiscale finite element method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
[20] T. Y. HOU, X.-H. WU, AND Z. CAI, *Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients*, Math. Comp., 68 (1999), pp. 913–943.
[21] T.J.R. HUGHES, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized methods*, Comput. Methods Appl. Mech. Engrg., 127 (1995), pp. 387–401.
[22] 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), pp. 3–24.
[23] 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), pp. 539–557.
[24] R. KORNHUBER, D. PETERSEIM, AND H. YSERENTANT, *An analysis of a class of variational multiscale methods based on subspace decomposition*, Math. Comp., 87 (2018), pp. 2765–2774.
[25] A. MÅLQVIST AND D. PETERSEIM, *Localization of elliptic multiscale problems*, Math. Comp., 83 (2014), pp. 2583–2603.
[26] T. P. A. MATHEW, *Domain Decomposition Methods for the Numerical Solution of Partial Differential Equations*, Springer, Berlin, 2008.

ETNA
Kent State University and
Johann Radon Institute (RICAM)

[27] D. PETERSEIM AND R. SCHEICHL, *Robust numerical upscaling of elliptic multiscale problems at high contrast*, Comput. Methods Appl. Math., 16 (2016), pp. 579–603.
[28] Y. SAAD, *Iterative methods for sparse linear systems*, 2nd ed., SIAM, Philadelphia, 2003.
[29] L. R. SCOTT AND S. ZHANG, *Finite element interpolation of nonsmooth functions satisfying boundary conditions*, Math. Comp., 54 (1990), pp. 483–493.
[30] B. F. SMITH, P. E. BJØRSTAD, AND W. D. GROPP, *Domain Decomposition*, Cambridge University Press, Cambridge, 1996.
[31] A. TOSELLI AND O. B. WIDLUND, *Domain Decomposition Methods—Algorithms and Theory*, Springer, Berlin, 2005.