이학박사 학위논문

Nonparametric DSSY Nonconforming Quadrilateral Element and Its Application to Multiscale Methods

DSSY 비순응유한요소와 멀티스케일 방법에 대한 적용

2020 년 2 월

서울대학교 대학원

수리과학부

조 강 훈

Abstract

Nonparametric DSSY Nonconforming Quadrilateral Element and Its Application to Multiscale Methods

Kanghun Cho
Department of Mathematical Sciences
The Graduate School
Seoul National University

We first consider nonparametric DSSY nonconforming quadrilateral element introduced in [26]. The element satisfies the mean value property on each edge and shows optimal convergence for second-order elliptic problems. We estimate the effect of numerical integration on finite element method and construct new quadrature formula for DSSY element. It is shown that only three nodes are enough to get optimal convergence for second-order elliptic problems. Numerical results are presented to compare new quadrature formula with usual Gaussian quadrature rules.

Next we study the nonconforming generalized multiscale finite element method(GMsFEM). The framework of GMsFEM is organized following [38], and every process of constructing nonconforming GMsFE spaces is presented in detail. GMsFE spaces consist of two ingredient. First one is the offline function space, a spectral decomposition of the snapshot space which is used to approximate the solution. Other one is the moment function space, which is used to impose continuity between local offline function spaces. Numerical results are presented based on nonparametric DSSY nonconforming element.

In last chapter, an algebraic multiscale finite element method is investigated. Suppose that the coefficient and the source term of second-order elliptic problems are not available, and we only know the microscale linear system. We try to construct macroscale linear systems only using the algebraic information on the components of microscale systems. One-dimensional case is examined in detail following GMsFEM framework, and two dimensional case is also presented using the DSSY nonconforming finite element space.

Keywords: DSSY nonconforming finite element, nonparametric finite element, generalized multiscale finite element method, algebraic multiscale method, numerical integration, quadrature formula, elliptic problem

Student Number: 2012-20258

Contents

List of Figures

List of Tables

Chapter 1

Nonparametric DSSY Nonconforming Quadrilateral Element

1.1 Introduction

In [19], Rannacher and Turek introduced the two types of rotated Q1 nonconforming elements depending on the choice of their DOFs between the four midpoint values and four integral values over four edges. It is often convenient if the two types of DOFs between the barycenter values and face average values over faces are identical for a finite element (K, P_K, Σ_K). We will coin it as the MVP (Mean Value Property)

$$\frac{1}{|e_j|}\int_{e_j} \phi \, d\sigma = \phi(\mathbf{m}_j), \quad \forall e_j \in \mathcal{F}(K), \quad \forall \phi \in P_K, \tag{1.1}$$

where $\mathcal{F}(K)$ denotes the set of all faces of K. For instance, for simplices, the linear Crouzeix–Raviart element [6] fulfills the MVP (1.1). For rectangular and parallelograms, Douglas *et al.* introduced a quadrilateral nonconforming element, so-called DSSY element [7] Among quadrilateral nonconforming elements which fulfills the MVP (1.1), five DOFs nonconforming element was given in [3], and later a class of nonparametric DSSY element of four DOFs was introduced in [26]. Also the three DOFs nonconforming element on quadrilateral [17] fulfills the MVP (1.1).

In [25], Meng *et al.* introduced an interesting new nonconforming quadrilateral element with a minimum number of quadrature points, although it does not satisfy the MVP (1.1). They adopted a new affine map from new reference element to physical element, which allows to express basis functions explicitly without solving local linear systems.

In this chapter we modify the class of nonparametric quadrilateral element [26] to fulfill the MVP (1.1) with a minimal number of quadrature points, following the approach of Meng *et al.*.

The organization of the rest of chapter is as follows. We review a few quadrilateral nonconforming elements including nonparametric DSSY element in Section 2. Then we introduce a new class of nonparametric quadrilateral element with the MVP (1.1) and compare the proposed element with the nonparametric DSSY element in Section 3. Section 4 is devoted to construction of quadrature formula for newly designed element. We present some numerical results in Section 5.

1.2 Quadrilateral nonconforming elements

In this section we review some quadrilateral nonconforming elements [7, 19, 26].

1.2.1 The Rannacher–Turek element and the DSSY element

Let Ω be a simply connected polygonal domain in $\mathbb{R}^2$ and denote by $(\mathcal{T}_h)_{h>0}$ a family of shape regular convex quadrilateral triangulations of Ω. Also denote by $\mathcal{E}_h$ by the set of all edges of $\mathcal{T}_h$. Here the parameter h is given by

$$h = \max_{K\in\mathcal{T}_h} \operatorname{diam}(K).$$

For a typical quadrilateral $K \in \mathcal{T}_h$, denote its four vertices by $\mathbf{v}_j$ for $j = 1, 2, 3, 4$, and assume $\mathbf{v}_0 := \mathbf{v}_4$ for the convenience. Also denote the edge between $\mathbf{v}_{j-1}$ and $\mathbf{v}_j$ by e_j, and the midpoint of e_j by $\mathbf{m}_j$ for $j = 1, 2, 3, 4$. Denote by $\widehat{K} = [-1, 1]^2$ be the reference element and add the "hat" $(\widehat{\ })$ for the notations for the reference vertices, edges, and midpoints of $\widehat{K}$ such as $\widehat{\mathbf{v}}_j$, $\widehat{e}_j$, and $\widehat{\mathbf{m}}_j$, respectively, for $j = 1, 2, 3, 4$.

Set

$$\mathcal{NC}_{\widehat{K},l} = \operatorname{Span}\{1, \widehat{x}_1, \widehat{x}_2, \widehat{\varphi}_l(\widehat{x}_1) - \widehat{\varphi}_l(\widehat{x}_2)\}, \quad l = 0, 1, 2,$$

where

$$\widehat{\varphi}_l(t) = \begin{cases} t^2, & l = 0, \\ t^2 - \frac{5}{3}t^4, & l = 1, \\ t^2 - \frac{25}{6}t^4 + \frac{7}{2}t^6, & l = 2. \end{cases}$$

The case $l = 0$ defines the Rannacher–Turek elements $RT(\widehat{K})$ with the edge–midpoint value DOFs or the edge–integral average DOFs. The two types of DOFs generate different finite elements. In the meanwhile, the cases $l = 1$ and $l = 2$ define the DSSY elements $\mathcal{NC}^{DSSY}_{\widehat{K},l}$, which fulfill the MVP (1.1), and thus the finite elements generated by both edge–midpoint DOFs and edge–integral average DOFs are identical to each other. Let us focus on the case of $l = 1$. Clearly linear polynomials satisfy the MVP and thus we investigate on the quartic polynomial $\widehat{\varphi}_1(\widehat{x}_1) - \widehat{\varphi}_1(\widehat{x}_2)$. Let us denote $\widehat{\varphi}_1(\widehat{x}_1) - \widehat{\varphi}_1(\widehat{x}_2)$ by

$\widehat{\psi}(\widehat{\mathbf{x}})$ for convenience. On the reference domain $\widehat{K}$, the function $\widehat{\psi}(\widehat{\mathbf{x}})$ can be factorized as

$$\widehat{\psi}(\widehat{\mathbf{x}}) = -\frac{5}{3}(\widehat{x}_1 - \widehat{x}_2)(\widehat{x}_1 + \widehat{x}_2)\Big(\widehat{x}_1^2 + \widehat{x}_2^2 - \frac{3}{5}\Big). \tag{1.2}$$

1.2.2 Nonparametric DSSY quadrilateral element

The nonparametric DSSY quadrilateral element is designed in [26] with retaining the MVP. We decompose the bilinear map $\mathcal{F}_K$ into a composition of an affine map and *simple bilinear map* [11, 17, 18]. A *simple bilinear map* associated with a vector $\widetilde{\mathbf{s}}$ is a bilinear map $S : \mathbb{R}^2 \to \mathbb{R}^2$ satisfying

$$S\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} + x_1 x_2 \, \widetilde{\mathbf{s}}, \quad \forall \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^2.$$

Notice that $\mathcal{F}_K$ can be written as follows:

$$\mathcal{F}_K(\widehat{\mathbf{x}}) = A\widehat{\mathbf{x}} + \widehat{x}_1\widehat{x}_2\,\mathbf{d} + \mathbf{b} = A\left[\widehat{\mathbf{x}} + \widehat{x}_1\widehat{x}_2\,A^{-1}\mathbf{d}\right] + \mathbf{b} = A\left[\widehat{\mathbf{x}} + \widehat{x}_1\widehat{x}_2\,\widetilde{\mathbf{s}}\right] + \mathbf{b}, \tag{1.3}$$

where A is a 2×2 matrix and $\mathbf{b}, \mathbf{d}$, and $\widetilde{\mathbf{s}}$ are two-dimensional vectors given by

$$\begin{aligned} A &= \frac{1}{4}\left(\mathbf{v}_1 - \mathbf{v}_2 - \mathbf{v}_3 + \mathbf{v}_4, \mathbf{v}_1 + \mathbf{v}_2 - \mathbf{v}_3 - \mathbf{v}_4\right), \\ \mathbf{d} &= \frac{\mathbf{v}_1 - \mathbf{v}_2 + \mathbf{v}_3 - \mathbf{v}_4}{4}, \quad \mathbf{b} = \frac{\mathbf{v}_1 + \mathbf{v}_2 + \mathbf{v}_3 + \mathbf{v}_4}{4}, \quad \widetilde{\mathbf{s}} = A^{-1}\mathbf{d}. \end{aligned}$$

Then (1.3) can be understood as the following decomposition of a *simple bilinear map* $\mathcal{S}_K$ associated with $\widetilde{\mathbf{s}}$ followed by an affine map $\mathcal{A}_K$:

$$\mathcal{F}_K = \mathcal{A}_K \circ \mathcal{S}_K,$$

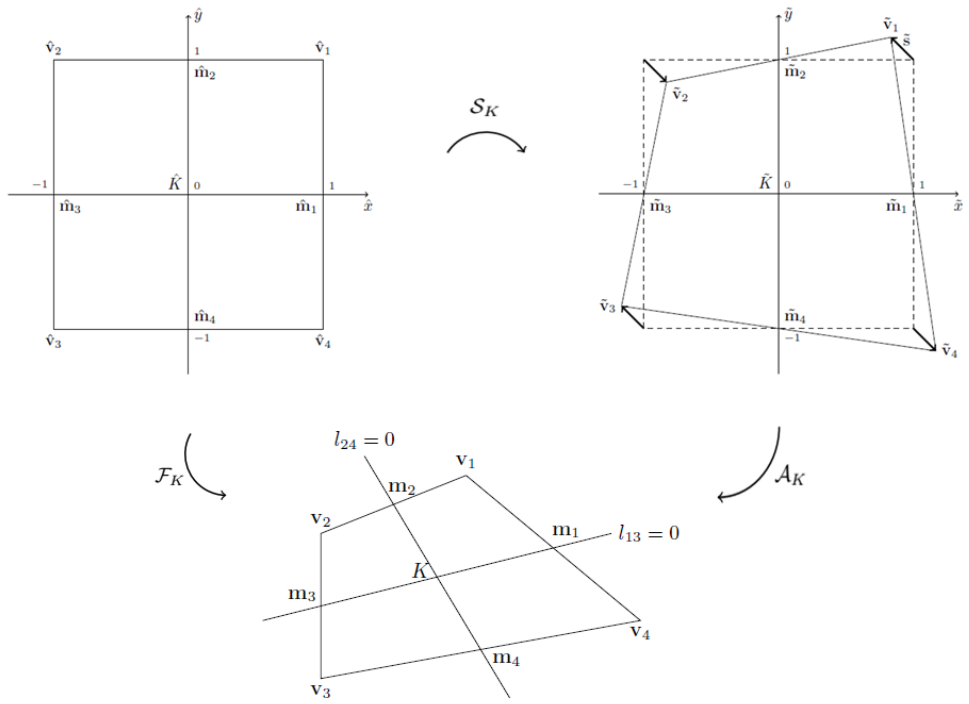


Figure 1.1. A bilinear map $\mathcal{F}_K$ from $\widehat{K}$ to K, a simple bilinear map $\mathcal{S}_K$ from $\widehat{K}$ to $\widetilde{K}$, and an affine map $\mathcal{A}_K$ from $\widetilde{K}$ to K.

where $\mathcal{A}_K : \widetilde{K} \to K$ and $\mathcal{S}_K : \widehat{K} \to \widetilde{K}$ are given by

$$\mathcal{A}_K(\widetilde{\mathbf{x}}) = A\widetilde{\mathbf{x}} + \mathbf{b}, \quad \mathcal{S}_K(\widehat{\mathbf{x}}) = \widehat{\mathbf{x}} + \widehat{x}\,\widehat{y}\,\widetilde{\mathbf{s}}.$$

Here $\widetilde{K} = \mathcal{S}_K(\widehat{K})$ is a quadrilateral with four vertices

$$\widetilde{\mathbf{v}}_1 = \widehat{\mathbf{v}}_1 + \widetilde{\mathbf{s}}, \quad \widetilde{\mathbf{v}}_2 = \widehat{\mathbf{v}}_2 - \widetilde{\mathbf{s}}, \quad \widetilde{\mathbf{v}}_3 = \widehat{\mathbf{v}}_3 + \widetilde{\mathbf{s}}, \quad \widetilde{\mathbf{v}}_4 = \widehat{\mathbf{v}}_4 - \widetilde{\mathbf{s}}.$$

It should be stressed that $\mathcal{S}_K$ is linear on each of four boundaries of $\widehat{K}$, and, in particular, the midpoints of $\widehat{K}$ are invariant under the map $\mathcal{S}_K$ and that $\widetilde{K}$ is a perturbation of $\widehat{K}$ by a single vector $\widetilde{\mathbf{s}}$ such that opposite vertices are moved in the same direction (see Figure 1.1.)

Denote the equations of lines passing through $\widetilde{\mathbf{v}}_1$, $\widetilde{\mathbf{v}}_3$, and $\widetilde{\mathbf{v}}_2$, $\widetilde{\mathbf{v}}_4$ by $\widetilde{\ell}_1(\widetilde{\mathbf{x}}) = 0$ and $\widetilde{\ell}_2(\widetilde{\mathbf{x}}) = 0$, respectively. Then $\widetilde{\ell}_1(\widetilde{\mathbf{x}})$ and $\widetilde{\ell}_2(\widetilde{\mathbf{x}})$ are linear polynomials given (up to multiplicative constants) by

$$\begin{aligned} \widetilde{\ell}_1(\widetilde{\mathbf{x}}) &= \widetilde{x}_1 - \widetilde{x}_2 - \widetilde{s}_1 + \widetilde{s}_2, \\ \widetilde{\ell}_2(\widetilde{\mathbf{x}}) &= \widetilde{x}_1 + \widetilde{x}_2 + \widetilde{s}_1 + \widetilde{s}_2. \end{aligned} \tag{1.4}$$

On the intermediate domain $\widetilde{K}$, the quartic polynomial $\widetilde{\mu}$ of DSSY element can be understood as the multiple of the following factors

$$\widetilde{\mu}(\widetilde{\mathbf{x}}) = -\frac{5}{3}\widetilde{\ell}_1(\widetilde{\mathbf{x}})\widetilde{\ell}_2(\widetilde{\mathbf{x}})\widetilde{\mathscr{Q}}(\widetilde{\mathbf{x}}), \tag{1.5}$$

where $\widetilde{\mathscr{Q}}(\widetilde{\mathbf{x}})$ is a quadratic polynomial. In [26] a class of quadratic polynomials $\widetilde{\mathscr{Q}}(\widetilde{\mathbf{x}};\widetilde{c})$ are chosen such that the quartic polynomial $\widetilde{\mu}(\widetilde{\mathbf{x}};\widetilde{c})$ satisfies the *mean value property* (1.6) in $\widetilde{K}$:

$$\frac{1}{|\widetilde{e}_j|}\int_{\widetilde{e}_j} \widetilde{\mu}\, d\widetilde{\sigma} = \widetilde{\mu}(\widetilde{\mathbf{m}}_j),\ j = 1,2,3,4. \tag{1.6}$$

Indeed, they are given in the following form

$$\begin{aligned} \widetilde{\mathscr{Q}}(\widetilde{\mathbf{x}};\widetilde{c}) = &\left(\widetilde{x}_1 + \frac{2}{5}\widetilde{s}_2\right)^2 + \left(\widetilde{x}_2 + \frac{2}{5}\widetilde{s}_1\right)^2 - \widetilde{r}^2 \\ &+ \widetilde{c}\left[(\widetilde{x}_1 + \frac{2}{5}\widetilde{s}_2)(\widetilde{x}_2 + \frac{2}{5}\widetilde{s}_1) + \frac{6}{25}\widetilde{s}_1\widetilde{s}_2\right], \end{aligned} \tag{1.7}$$

with $\widetilde{r} = \frac{\sqrt{6}}{5}\sqrt{\frac{5}{2} - \widetilde{s}_1^2 - \widetilde{s}_2^2}$ for arbitrary constant $\widetilde{c} \in \mathbb{R}$. Here, we assume that the coefficient of $\widetilde{x}_1$ is normalized. Notice that $\widetilde{r}$ takes a positive real value if $\widetilde{K}$ is convex, which is equivalent to $|\widetilde{s}_1| + |\widetilde{s}_2| \leq 1$. (Remark 2.1 of [26])

Define, for each $\widetilde{c} \in \mathbb{R}$,

$$\widetilde{\mu}(\widetilde{x}_1, \widetilde{x}_2; \widetilde{c}) = -\frac{5}{3}\widetilde{\ell}_1(\widetilde{x}_1, \widetilde{x}_2)\widetilde{\ell}_2(\widetilde{x}_1, \widetilde{x}_2)\widetilde{\mathscr{Q}}(\widetilde{x}_1, \widetilde{x}_2),$$

where $\widetilde{\ell}_1$ and $\widetilde{\ell}_2$ are defined by (1.4) and $\widetilde{\mathscr{Q}}$ by (1.7) depending on $\widetilde{c}$ as well as $\widetilde{\mathbf{s}}$. Then a class of *nonparametric nonconforming elements* are defined on the intermediate quadrilaterals $\widetilde{K}$ with four DOFs as follows:

1. $\widetilde{K} = \mathcal{S}_K(\widehat{K})$;

2. $\widetilde{P}_{\widetilde{K}}(\widetilde{c}) = \mathrm{Span}\{1, \widetilde{x}_1, \widetilde{x}_2, \widetilde{\mu}(\widetilde{x}_1, \widetilde{x}_2; \widetilde{c})\}$;

3. $\widetilde{\Sigma}_{\widetilde{K}} = \{\text{four edge-midpoint values of } \widetilde{K}\} = \{\text{four mean values over edges of } \widetilde{K}\}$.

The above class of intermediate nonparametric elements is unisolvent with $\widetilde{c}$ in most cases.

Theorem 1.2.1. *[26] The intermediate nonparametric element* $\left(\widetilde{K}, \widetilde{P}_{\widetilde{K}}(\widetilde{c}), \widetilde{\Sigma}_{\widetilde{K}}\right)$ *is unisolvent if* $\widetilde{c}$ *satisfies*

$$\widetilde{s}_1^2 + \widetilde{s}_2^2 + \frac{1}{3} + \widetilde{c}\,\widetilde{s}_1\widetilde{s}_2 \neq 0.$$

A class of *nonparametric nonconforming elements on quadrilaterals* K is directly defined by using the affine map $\mathcal{A}_K$ from $\widetilde{K}$ to K. The transformed elements also satisfy the MVP and unisolvency.

1. $K = \mathcal{F}_K(\widehat{K})$;

2. $\mathcal{NC}_K^{np} = P_K(\widetilde{c}) = \mathrm{Span}\{1, x_1, x_2, \mu(x_1, x_2; \widetilde{c})\}$;

3. $\Sigma_K = \{\text{four edge-midpoint values of } K\} = \{\text{four mean values over edges of } K\}$.

Here $\mu(x_1, x_2; \widetilde{c})$ is a quartic polynomial defined by

$$\mu(x_1, x_2; \widetilde{c}) = \widetilde{\mu} \circ \mathcal{A}_K^{-1}(x_1, x_2; \widetilde{c}) = -\frac{5}{3}\ell_1(x_1, x_2)\ell_2(x_1, x_2)q(x_1, x_2; \widetilde{c}),$$

where

$$\ell_1(\mathbf{x}) = \widetilde{\ell}_1 \circ \mathcal{A}_K^{-1}(\mathbf{x}), \quad \ell_2(\mathbf{x}) = \widetilde{\ell}_2 \circ \mathcal{A}_K^{-1}(\mathbf{x}), \quad q(\mathbf{x}; \widetilde{c}) = \widetilde{\mathscr{Q}} \circ \mathcal{A}_K^{-1}(\mathbf{x}).$$

As $\widehat{\psi}(\widehat{\mathbf{x}})$ in (1.2) can be regarded as a product of two lines and a circle, $\mu(\mathbf{x}; \widetilde{c})$ can be interpreted similarly. That is, $\mu(\mathbf{x}; \widetilde{c})$ can be understood as a product of two linear polynomials and one quadratic polynomial such that the straight lines $\ell_1(\mathbf{x}) = 0$ and $\ell_2(\mathbf{x}) = 0$ are passing through $\mathbf{v}_1$, $\mathbf{v}_3$ and $\mathbf{v}_2$, $\mathbf{v}_4$, respectively and $q(\mathbf{x}; \widetilde{c}) = 0$ is an ellipse which is determined to fulfill the MVP for $\widetilde{\mu}(\widetilde{\mathbf{x}})$.

Finally the global nonparametric DSSY element spaces is defined as follows:

$$\begin{aligned}
\mathcal{NC}_h^{np} = \{v_h \in L^2(\Omega) \mid v_h|_K \in \mathcal{NC}_K^{np} \text{ for } K \in \mathcal{T}_h, \\
v_h \text{ is continuous at the midpoint of each } e \in \mathcal{E}_h\}, \\
\mathcal{NC}_{h,0}^{np} = \{v_h \in \mathcal{NC}_h^{np} \mid v_h \text{ is zero at the midpoint of each } e \in \mathcal{E}_h \cap \partial\Omega\}.
\end{aligned}$$

We can simply eliminate the parameter $\widetilde{c}$ everywhere in the above finite element construction by fixing $\widetilde{c} = 0$, which gives

$$\widetilde{\mathscr{Q}}(\widetilde{\mathbf{x}}) = \left(\widetilde{x}_1 + \frac{2}{5}\widetilde{s}_2\right)^2 + \left(\widetilde{x}_2 + \frac{2}{5}\widetilde{s}_1\right)^2 - \frac{6}{25}\left(\frac{5}{2} - \widetilde{s}_1^2 - \widetilde{s}_2^2\right). \tag{1.8}$$

Recalling (1.4), we can rewrite (1.8) in terms of $\widetilde{\ell}_1$ and $\widetilde{\ell}_2$ as follows:

$$\begin{aligned}\widetilde{\mathscr{Q}}(\widetilde{\ell}_1,\widetilde{\ell}_2) &= \left[\frac{1}{2}\left(\widetilde{\ell}_1+\widetilde{\ell}_2-\frac{6}{5}\widetilde{s}_2\right)\right]^2+\left[\frac{1}{2}\left(\widetilde{\ell}_2-\widetilde{\ell}_1-\frac{6}{5}\widetilde{s}_1\right)\right]^2 \\ &\qquad -\frac{6}{25}\left(\frac{5}{2}-\widetilde{s}_1^2-\widetilde{s}_2^2\right) \\ &= \frac{1}{2}\Big[\widetilde{l}_2-\frac{3}{5}(\widetilde{s}_1+\widetilde{s}_2)\Big]^2+\frac{1}{2}\Big[\widetilde{l}_1+\frac{3}{5}(\widetilde{s}_1-\widetilde{s}_2)\Big]^2 \\ &\qquad -\frac{6}{25}\Big(\frac{5}{2}-\widetilde{s}_1^2-\widetilde{s}_2^2\Big).\end{aligned} \tag{1.9}$$

1.3 A new intermediate space $\overline{K}$ for nonparametric DSSY element

In the previous section nonparametric DSSY element is developed on quadrilaterals. Here we modify the element by adopting a new intermediate space $\overline{K}$ with an affine map $\mathcal{C}_K$ from $\overline{K}$ to K. The modified element also have the MVP and its basis functions can be expressed explicitly.

1.3.1 The Meng *et al.* approach

We first review the approach of Meng *et al.* in [25] where a new reference quadrilateral is proposed. The element takes the four integral values over four edges as their DOFs. It does not satisfy the MVP but the basis functions are given explicitly.

We use similar notation in the previous section. Set $l_1(\mathbf{x})$ and $l_2(\mathbf{x})$ to be linear polynomials such that $l_1(\mathbf{x})$ and $l_2(\mathbf{x})$ are line equations passing through $\mathbf{v}_1$, $\mathbf{v}_3$ and $\mathbf{v}_2$, $\mathbf{v}_4$, respectively. Also set

$$l_1(\mathbf{v}_1)=1,\ l_1(\mathbf{v}_3)=h_1 \quad \text{and} \quad l_2(\mathbf{v}_2)=1,\ l_2(\mathbf{v}_4)=h_2.$$

Since K is a convex quadrilateral, h_1 and h_2 satisfy $h_1, h_2 < 0$. Also we may

assume $h_1, h_2 > -1$ by changing the order of the vertices. Now a reference quadrilateral $\overline{K}$ is designed to have four vertices

$$\overline{\mathbf{v}}_1 = (1,0), \overline{\mathbf{v}}_2 = (0,1), \overline{\mathbf{v}}_3 = (h_1,0), \overline{\mathbf{v}}_4 = (0,h_2). \tag{1.10}$$

Obviously there exists a unique affine map $\mathcal{C}_K : \overline{K} \to K$ such that $\mathcal{C}_K(\overline{\mathbf{v}}_j) = \mathbf{v}_j,\ j = 1,2,3,4$. Observe the following property

$$\bar{l}_j(\overline{\mathbf{x}}) = l_j \circ \mathcal{C}_K(\overline{\mathbf{x}}) = \overline{x}_j, \qquad j = 1,2,$$

which indicates that the inverse affine map $\mathcal{C}_K^{-1} : K \to \overline{K}$ can be written as

$$\mathcal{C}_K^{-1}(\mathbf{x}) = (l_1(\mathbf{x}), l_2(\mathbf{x})).$$

The nonconforming element $\left(\overline{K}, \overline{P}_{\overline{K}}, \overline{\Sigma}_{\overline{K}}\right)$ are defined on the reference quadrilaterals $\overline{K}$ as follows:

1. $\overline{K} = \mathcal{C}_K^{-1}(K)$ is the convex quadrilateral defined by (1.10);

2. $\overline{P}_{\overline{K}}^{MCL} = \text{Span}\{1, \overline{x}_1, \overline{x}_2, \bar{l}_1\bar{l}_2 = \overline{x}_1\overline{x}_2\}$,

3. $\overline{\Sigma}_{\overline{K}}^{MCL} = \{\text{four mean values over edges of } \overline{K}\}$.

Theorem 1.3.1. *[25] The nonparametric element* $\left(\overline{K}, \overline{P}_{\overline{K}}^{MCL}, \overline{\Sigma}_{\overline{K}}^{MCL}\right)$ *is unisolvent.*

Proof. Denote the functions $1, \overline{x}_1, \overline{x}_2, \overline{x}_1\overline{x}_2$ by $\overline{\phi}_1, \overline{\phi}_2, \overline{\phi}_3$, and $\overline{\phi}_4$, respectively.

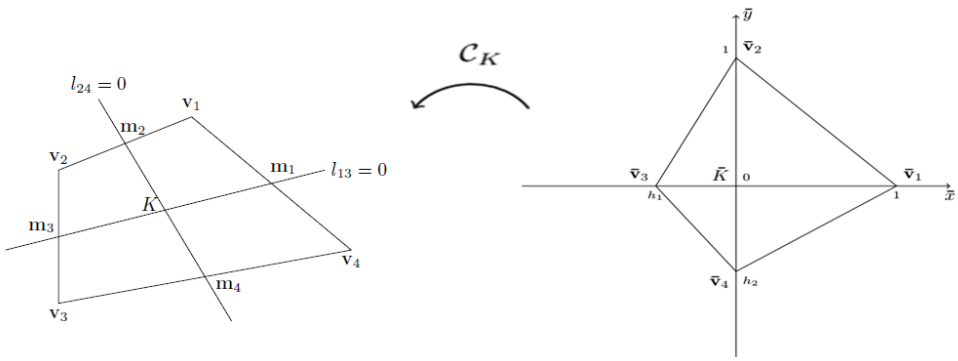


Figure 1.2. An affine map $\mathcal{C}_K$ from $\overline{K}$ to K.

Define $A=(a_{jk})\in M_{4\times 4}(\mathbb{R})$ by $a_{jk}=\frac{1}{|\overline{e}_j|}\int_{\overline{e}_j}\overline{\phi}_j d\overline{\sigma}$. Then

$$A=\begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & \frac{1}{6} \\ 1 & \frac{1}{2}h_1 & \frac{1}{2} & \frac{1}{6}h_1 \\ 1 & \frac{1}{2}h_1 & \frac{1}{2}h_2 & \frac{1}{6}h_1h_2 \\ 1 & \frac{1}{2} & \frac{1}{2}h_2 & \frac{1}{6}h_2 \end{pmatrix} \tag{1.11}$$

with $\det(A)=-\frac{1}{24}(1-h_1)^2(1-h_2)^2$. Since $h_1<0$ and $h_2<0$, A is nonsingular.

□

By computing A^{-1}, we can explicitly present basis functions on $\overline{K}$ as follows:

$$\begin{aligned}
\overline{\phi}_1(\overline{x}_1,\overline{x}_2) &= S_{\overline{K}}\Big(\frac{h_1h_2}{2}-h_2\overline{x}_1-h_1\overline{x}_2+3\,\overline{x}_1\overline{x}_2\Big),\\
\overline{\phi}_2(\overline{x}_1,\overline{x}_2) &= S_{\overline{K}}\Big(-\frac{h_2}{2}+h_2\overline{x}_1+\overline{x}_2-3\,\overline{x}_1\overline{x}_2\Big),\\
\overline{\phi}_3(\overline{x}_1,\overline{x}_2) &= S_{\overline{K}}\Big(\frac{1}{2}-\overline{x}_1-\overline{x}_2+3\,\overline{x}_1\overline{x}_2\Big),\\
\overline{\phi}_4(\overline{x}_1,\overline{x}_2) &= S_{\overline{K}}\Big(-\frac{h_1}{2}+\overline{x}_1+h_1\overline{x}_2-3\,\overline{x}_1\overline{x}_2\Big),
\end{aligned}$$

where

$$S_{\overline{K}} = \frac{2}{(1-h_1)(1-h_2)} = \frac{1}{\text{area}(\overline{K})}.$$

Notice that the mean value property does not hold. For example,

$$\overline{\phi}_1(\overline{\mathbf{m}}_2) = \overline{\phi}_1\Big(\frac{h_1}{2}, \frac{1}{2}\Big) = \frac{h_1}{4} > 0.$$

Here the quadrilateral $\overline{K}$ plays a similar role as the intermediate space $\widetilde{K}$ in nonparametric DSSY element. The nonconforming element (K, P_K, Σ_K) on physical domain K is defined via the affine map $\mathcal{C}_K$:

1. $K = \mathcal{C}_K(\overline{K})$ is a convex quadrilateral;
2. $P_K^{MCL} = \text{Span}\{1, l_1, l_2, l_1 l_2\} = \text{Span}\{1, x_1, x_2, l_1 l_2\}$,
3. $\Sigma_K^{MCL} = \{\text{four mean values over edges of } K\}$.

1.3.2 A class of nonparametric DSSY elements on $\overline{K}$

In this section we define a class of nonparametric DSSY elements on $\overline{K}$ where $\overline{K}$ is proposed in §1.3.1. Our strategy is to find a quartic polynomial on $\overline{K}$ similar to (1.2), where the ansatz is

$$\overline{\mu}(\overline{\mathbf{x}}) = -\frac{5}{3}\overline{\ell}_1(\overline{\mathbf{x}})\overline{\ell}_2(\overline{\mathbf{x}})\overline{\mathscr{Q}}(\overline{\mathbf{x}}) \tag{1.12}$$

with linear polynomials $\overline{\ell}_j(\overline{\mathbf{x}}), j = 1, 2$, and a quadratic polynomial $\overline{\mathscr{Q}}(\overline{\mathbf{x}})$. Here $\overline{\ell}_1(\overline{\mathbf{x}})$ and $\overline{\ell}_2(\overline{\mathbf{x}})$ are linear polynomials similar to (1.4) such that $\overline{\ell}_1(\overline{\mathbf{x}}) = 0$ and $\overline{\ell}_2(\overline{\mathbf{x}}) = 0$ are the line equations passing through $\overline{\mathbf{v}}_1$, $\overline{\mathbf{v}}_3$, and $\overline{\mathbf{v}}_2$, $\overline{\mathbf{v}}_4$, respectively. They are given (up to multiplicative constants) by

$$\begin{aligned} \overline{\ell}_1(\overline{\mathbf{x}}) &= \overline{x}_1, \\ \overline{\ell}_2(\overline{\mathbf{x}}) &= \overline{x}_2. \end{aligned} \tag{1.13}$$

A class of quadratic polynomials $\overline{\mathscr{Q}}(\overline{\mathbf{x}};\overline{c})$ are chosen such that the quartic polynomial $\overline{\mu}(\overline{\mathbf{x}};\overline{c})$ fulfills the mean value property in $\overline{K}$. We use the following Gauss quadrature formula which is exact for quartic polynomials:

$$\int_{-1}^{1} f(t)\ dt \approx \frac{8}{9}f(0) + \frac{5}{9}(f(\xi) + f(-\xi)), \quad \xi = \sqrt{\frac{3}{5}}.$$

Denote by $\overline{\mathbf{d}}_j := \frac{\overline{\mathbf{v}}_j - \overline{\mathbf{v}}_{j-1}}{2}$ for $j = 1, 2, 3, 4$, assuming $\overline{\mathbf{v}}_0 = \overline{\mathbf{v}}_4$. Then the mean value property (1.6) is simplified into the form

$$\overline{\mu}(\overline{\mathbf{g}}_{2j-1}) + \overline{\mu}(\overline{\mathbf{g}}_{2j}) - 2\overline{\mu}(\overline{\mathbf{g}}_j) = 0, \quad j = 1, 2, 3, 4, \tag{1.14}$$

where

$$\begin{aligned}
\overline{\mathbf{g}}_1 &= \overline{\mathbf{m}}_1 - \xi\overline{\mathbf{d}}_1, \quad \overline{\mathbf{g}}_2 = \overline{\mathbf{m}}_1 + \xi\overline{\mathbf{d}}_1,\\
\overline{\mathbf{g}}_3 &= \overline{\mathbf{m}}_2 - \xi\overline{\mathbf{d}}_2, \quad \overline{\mathbf{g}}_4 = \overline{\mathbf{m}}_2 + \xi\overline{\mathbf{d}}_2,\\
\overline{\mathbf{g}}_5 &= \overline{\mathbf{m}}_3 - \xi\overline{\mathbf{d}}_3, \quad \overline{\mathbf{g}}_6 = \overline{\mathbf{m}}_3 + \xi\overline{\mathbf{d}}_3,\\
\overline{\mathbf{g}}_7 &= \overline{\mathbf{m}}_4 - \xi\overline{\mathbf{d}}_4, \quad \overline{\mathbf{g}}_8 = \overline{\mathbf{m}}_4 + \xi\overline{\mathbf{d}}_4.
\end{aligned}$$

Observe that $\overline{\mathbf{g}}_{2j-1}$, $\overline{\mathbf{g}}_{2j}$ and $\overline{\mathbf{m}}_j$ are the Gauss points on $\overline{\mathbf{e}}_j(t)$ for each $j = 1, 2, 3, 4$, since the line equations for edges $\overline{\mathbf{e}}_j$ are written in vector notation as follows:

$$\overline{\mathbf{e}}_j(t) = \overline{\mathbf{m}}_j + t\,\overline{\mathbf{d}}_j,$$

for $t \in [-1, 1]$. Consider the quartic polynomial $\overline{\mu}(\overline{\mathbf{x}})$ in (1.12) restricted to an edge $\overline{\mathbf{e}}_j(t)$. Notice that from (1.13) the following equations hold:

$$\overline{\ell}_1(\overline{\mathbf{g}}_{2j-1})\,\overline{\ell}_2(\overline{\mathbf{g}}_{2j-1}) = \overline{\ell}_1(\overline{\mathbf{g}}_{2j})\,\overline{\ell}_2(\overline{\mathbf{g}}_{2j}) = (1-\xi^2)\overline{\ell}_1(\overline{\mathbf{m}}_j)\,\overline{\ell}_2(\overline{\mathbf{m}}_j). \tag{1.15}$$

By combining (1.14) and (1.15), we can see that (1.1) holds if and only if the quadratic polynomial $\overline{\mathscr{Q}}$ satisfies

$$\overline{\mathscr{Q}}(\overline{\mathbf{g}}_{2j-1}) + \overline{\mathscr{Q}}(\overline{\mathbf{g}}_{2j}) - 5\,\overline{\mathscr{Q}}(\overline{\mathbf{m}}_j) = 0, \quad j = 1, 2, 3, 4. \tag{1.16}$$

The general solution of (1.16) can be found by using symbolic calculation package. The solution is given with arbitrary constant $\overline{c} \in \mathbb{R}$:

$$\overline{\mathscr{Q}}(\overline{\mathbf{x}}; \overline{c}) = \overline{x}_1^2 - \frac{3}{10}(1+h_1)\overline{x}_1 + \frac{3}{20}h_1 + \overline{c}\left[\overline{x}_2^2 - \frac{3}{10}(1+h_2)\overline{x}_2 + \frac{3}{20}h_2\right]. \tag{1.17}$$

Then the quartic polynomial $\overline{\mu}$ is defined for each $\overline{c} \in \mathbb{R}$ as follows:

$$\overline{\mu}(\overline{x}_1, \overline{x}_2; \overline{c}) = -\frac{5}{3}\overline{\ell}_1(\overline{x}_1, \overline{x}_2)\overline{\ell}_2(\overline{x}_1, \overline{x}_2)\overline{\mathscr{Q}}(\overline{x}_1, \overline{x}_2; \overline{c}), \tag{1.18}$$

where $\overline{\ell}_1$ and $\overline{\ell}_2$ are two linear polynomials defined by (1.13) and $\overline{\mathscr{Q}}$ is quadratic polynomial by (1.17).

Now a class of *nonparametric nonconforming elements on the intermediate quadrilaterals* $\overline{K}$ with four DOFs are defined as follows.

1. $\overline{K} = \mathcal{C}_K^{-1}(K)$ is the convex quadrilateral defined by (1.10);
2. $\overline{P}_{\overline{K}} = \mathrm{Span}\{1, \overline{x}_1, \overline{x}_2, \overline{\mu}(\overline{x}_1, \overline{x}_2; \overline{c})\}$,
3. $\overline{\Sigma}_{\overline{K}} = \{\text{four mean values over edges of } \overline{K}\} = \{\text{four mean values over edges of } \overline{K}\}$.

It should be stressed that by the above construction the mean value property holds for any element $\overline{p} \in \overline{P}_{\overline{K}}$:

$$\frac{1}{|\overline{e}_j|}\int_{\overline{e}_j} \overline{p}\; d\overline{\sigma} = \overline{p}(\overline{\mathbf{m}}_j), \qquad j = 1, 2, 3, 4.$$

Also the above class of nonparametric elements is unisolvent for most of $\overline{c}$, such as $\overline{c} > 0$.

Theorem 1.3.2. $(\overline{K}, \overline{P}_{\overline{K}}, \overline{\Sigma}_{\overline{K}})$ *is unisolvent if* $\overline{c}$ *satisfies*

$$h_1^2 + h_1 + 1 + \overline{c}(h_2^2 + h_2 + 1) \neq 0.$$

Proof. Denote the functions $1, \overline{x}_1, \overline{x}_2, \overline{\mu}(\overline{x}_1, \overline{x}_2; \overline{c})$ by $\overline{\phi}_1, \overline{\phi}_2, \overline{\phi}_3$, and $\overline{\phi}_4$, respectively. Define $A = (a_{jk}) \in M_{4\times 4}(\mathbb{R})$ by $a_{jk} = \frac{1}{|\overline{e}_j|}\int_{\overline{e}_j} \overline{\phi}_j d\overline{\sigma}$. By direct calculation, we have

$$\begin{aligned}
&\frac{1}{|\overline{e}_1|}\int_{\overline{e}_1} \overline{x}_1^j \overline{x}_2^k \, d\overline{\sigma} = \frac{j!k!}{(j+k+1)!}, && \frac{1}{|\overline{e}_2|}\int_{\overline{e}_2} \overline{x}_1^j \overline{x}_2^k \, d\overline{\sigma} = \frac{j!k!}{(j+k+1)!} h_1^j, \\
&\frac{1}{|\overline{e}_3|}\int_{\overline{e}_3} \overline{x}_1^j \overline{x}_2^k \, d\overline{\sigma} = \frac{j!k!}{(j+k+1)!} h_1^j \, h_2^k, && \frac{1}{|\overline{e}_4|}\int_{\overline{e}_4} \overline{x}_1^j \overline{x}_2^k \, d\overline{\sigma} = \frac{j!k!}{(j+k+1)!} h_2^k.
\end{aligned}$$

Then

$$A = \begin{pmatrix} 1 & \frac{1}{2} & \frac{1}{2} & -\frac{1}{24}(\overline{c}+1) \\ 1 & \frac{1}{2}h_1 & \frac{1}{2} & -\frac{1}{24}h_1(h_1^2 + \overline{c}) \\ 1 & \frac{1}{2}h_1 & \frac{1}{2}h_2 & -\frac{1}{24}h_1 h_2(h_1^2 + \overline{c}h_2^2) \\ 1 & \frac{1}{2} & \frac{1}{2}h_2 & -\frac{1}{24}h_2(\overline{c}h_2^2 + 1) \end{pmatrix} \tag{1.19}$$

with $\det(A) = \frac{1}{96}(1-h_1)^2(1-h_2)^2\big(h_1^2 + h_1 + 1 + \overline{c}\,(h_2^2 + h_2 + 1)\big)$. Thus A is nonsingular if and only if $h_1^2 + h_1 + 1 + \overline{c}\,(h_2^2 + h_2 + 1) \neq 0$. □

The affine map $\mathcal{C}_K : \overline{K} \to K$ induces a class of *nonparametric nonconforming elements on quadrilaterals* K, denoted by $(K, \mathcal{N}C_K, \Sigma_K)$, which also satisfy the MVP and unisolvency.

1. $K = \mathcal{C}_K(\overline{K})$;

2. $\mathcal{N}C_K = P_K(\overline{c}) = \text{Span}\{1, x_1, x_2, \mu(x_1, x_2; \overline{c})\}$;

3. $\Sigma_K = \{\text{four edge-midpoint values of } K\} = \{\text{four mean values over edges of } K\}$.

Here $\mu(x_1, x_2; \overline{c})$ is a quartic polynomial defined by

$$\mu(x_1, x_2; \overline{c}) = \overline{\mu} \circ \mathcal{C}_K^{-1}(x_1, x_2; \overline{c}) = -\frac{5}{3}\ell_1(x_1, x_2)\ell_2(x_1, x_2)q(x_1, x_2; \overline{c}),$$

where

$$\ell_1(\mathbf{x}) = \overline{\ell}_2 \circ \mathcal{C}_K^{-1}(\mathbf{x}), \quad \ell_2(\mathbf{x}) = \overline{\ell}_1 \circ \mathcal{C}_K^{-1}(\mathbf{x}), \quad q(\mathbf{x}; \overline{c}) = \overline{\mathscr{Q}} \circ \mathcal{C}_K^{-1}(\mathbf{x}).$$

We may choose $\overline{c} = 1$ to have symmetry in (1.17), with which (1.17) reads as follows:

$$\begin{aligned}\overline{\mathscr{Q}}(\overline{\mathbf{x}}) = &\Big(\overline{x}_1 - \frac{3}{20}(1+h_1)\Big)^2 + \Big(\overline{x}_2 - \frac{3}{20}(1+h_2)\Big)^2 \\ &+ \frac{3}{20}h_1 - \frac{9}{400}(1+h_1)^2 + \frac{3}{20}h_2 - \frac{9}{400}(1+h_2)^2.\end{aligned} \tag{1.20}$$

Remark 1.3.3. *We may consider rectangular elements for simple case. Then the basis functions of* $K = [\frac{h_x}{2}, \frac{h_x}{2}] \times [-\frac{h_y}{2}, \frac{h_y}{2}]$ *are given explicitly as follows:*

$$\begin{aligned}\phi_1(\mathbf{x}) &= \frac{1}{4} + \frac{1}{2}\ell_1(\mathbf{x}) - \frac{1}{2}\ell_2(\mathbf{x}) + \frac{6}{\overline{c}+1}\mu(\mathbf{x}), \\ \phi_2(\mathbf{x}) &= \frac{1}{4} + \frac{1}{2}\ell_1(\mathbf{x}) + \frac{1}{2}\ell_2(\mathbf{x}) - \frac{6}{\overline{c}+1}\mu(\mathbf{x}), \\ \phi_3(\mathbf{x}) &= \frac{1}{4} - \frac{1}{2}\ell_1(\mathbf{x}) + \frac{1}{2}\ell_2(\mathbf{x}) + \frac{6}{\overline{c}+1}\mu(\mathbf{x}), \\ \phi_4(\mathbf{x}) &= \frac{1}{4} - \frac{1}{2}\ell_1(\mathbf{x}) - \frac{1}{2}\ell_2(\mathbf{x}) - \frac{6}{\overline{c}+1}\mu(\mathbf{x}),\end{aligned}$$

where $\ell_1(\mathbf{x}) = \frac{1}{h_x}x + \frac{1}{h_y}y, \quad \ell_2(\mathbf{x}) = -\frac{1}{h_x}x + \frac{1}{h_y}y.$

In the end, the global nonconforming element spaces is defined by

$$\begin{aligned}\mathcal{NC}_h = \{v_h \in L^2(\Omega) \mid v_h|_K \in \mathcal{NC}_K \text{ for } K \in \mathcal{T}_h, \\ v_h \text{ is continuous at the midpoint of each } e \in \mathcal{E}_h\}, \\ \mathcal{NC}_{h,0} = \{v_h \in \mathcal{NC}_h \mid v_h \text{ is zero at the midpoint of each } e \in \mathcal{E}_h \cap \partial\Omega\}.\end{aligned}$$

1.4 Construction of quadrature formula

1.4.1 Effect of numerical integration on FEM

Consider the following elliptic boundary problem

$$\begin{cases} -\nabla \cdot (\kappa(\mathbf{x})\nabla\, u) = f \text{ in } \Omega, \\ \qquad\qquad\qquad u = 0 \text{ on } \partial\Omega, \end{cases} \tag{1.21}$$

where Ω is a simply connected polygonal domain in $\mathbb{R}^2$, and $\kappa = \big(\kappa_{ij}(\mathbf{x})\big)$ is a symmetric matrix with smooth functions $\kappa_{ij}(x)$. We assume that κ is uniformly elliptic on Ω so that there is a constant $\lambda > 0$ such that

$$\sum_{i,j=1}^{2} \kappa_{ij}\xi_i\xi_j \geq \lambda|\xi|^2.$$

For any open subset U of $\mathbb{R}^n$, denote the seminorm and norm of the Sobolev space $W^{k,p}(U)$ by $|\cdot|_{k,p,U}$ and $||\cdot||_{k,p,U}$, respectively. Also denote by $H^k(U) = W^{k,2}(U)$ and abbreviate $|\cdot|_{k,p,U}$ and $||\cdot||_{k,p,U}$ as $|\cdot|_{k,U}$ and $||\cdot||_{k,U}$. The variational form of (1.21) is given by finding $u \in H_0^1(\Omega)$ such that

$$a(u, v) = F(v), \quad v \in H_0^1(\Omega) \tag{1.22}$$

where $a(u, v) = \int_\Omega \kappa\nabla\, u \cdot \nabla\, v\, d\mathbf{x}$ and $F(v) = \int_\Omega fv\, d\mathbf{x}$. Consider the nonconforming finite element space $\mathcal{N}C_h$ made up by (K, P_K, Σ_K) in §1.3.2. Then the finite element approximation $u_h \in \mathcal{N}C_{h,0}$ of (1.22) is defined as the solution of discrete problem

$$a_h(u_h, v_h) = F_h(v_h), \quad v_h \in \mathcal{N}C_{h,0}, \tag{1.23}$$

where $a_h(u,v) = \sum_{K\in\mathcal{T}_h} \int_K \kappa\nabla u \cdot \nabla v \, d\mathbf{x}$ and $F_h(v) = \sum_{K\in\mathcal{T}_h} \int_K fv \, d\mathbf{x}$. The energy error estimate for nonconforming method is provided in [7] using the broken energy norm

$$||v||_{1,h} = \sqrt{a_h(v,v)}.$$

Theorem 1.4.1. *[7] Assume that u and u_h are the solutions of* (1.22) *and* (1.23)*, respectively. Then we have the following error estimate*

$$||u - u_h||_{1,h} \leq Ch||u||_{2,\Omega}. \tag{1.24}$$

In actual computation we need to calculate definite integrals in $a_h(u,v)$ and $f(v)$. Gaussian quadrature rules are frequently used, but we want to construct more efficient quadrature formula while the order of convergence is unchanged by numerical integration. There are many papers [27, 28, 29] studying the effect of numerical integration on finite element method. In [27], sufficient conditions for quadrature formula are provided to preserve the order of convergence, where the finite element space consists of $P_k(K)$, piecewise polynomials of degree $\leq k$. If the formula is exact for $P_{2k-2}(K)$, then the optimal order of convergence $\mathcal{O}(h^k)$ is obtained for the energy norm error.

We want to find such conditions based on $\mathcal{N}C_h$. Notice that our nonconforming elements on K are constructed via the affine map $\mathcal{C}_K$ from the reference element $\overline{K}$ onto K. Thus it is natural to construct quadrature formula on $\overline{K}$, which is defined with positive weights $\overline{\omega}_l$ and nodes $\overline{\mathbf{b}}_l$, by

$$\int_{\overline{K}} \overline{\phi}(\overline{\mathbf{x}}) d\overline{\mathbf{x}} \approx \sum_{l=1}^{L} \overline{\omega}_l \, \overline{\phi}(\overline{\mathbf{b}_l}). \tag{1.25}$$

Denote the Jacobian determinant of $\mathcal{C}_K$ by $\det(\mathbf{D}\mathcal{C}_K)$, and observe that

$$\int_K \phi(\mathbf{x})d\mathbf{x} = |\det(\mathbf{D}\mathcal{C}_K)| \int_{\overline{K}} \overline{\phi}(\overline{\mathbf{x}})d\overline{\mathbf{x}}.$$

It induces the quadrature formulae on K from (1.25), which is given by

$$\int_K \phi(\mathbf{x})d\mathbf{x} \approx \sum_{l=1}^{L} \omega_{l,K}\, \phi(\mathbf{b}_{l,K}), \tag{1.26}$$

where $\omega_{l,K} = |\det(\mathbf{D}\mathcal{C}_K)|\,\overline{\omega}_l$ and $\mathbf{b}_{l,K} = \mathcal{C}_K(\overline{\mathbf{b}}_l)$. Suppose that the discrete problem (1.23) is approximated by the above quadrature formulae. Then the numerical solution $\overline{u}_h$ is defined as the solution of approximate problem

$$\overline{a}_h(\overline{u}_h, \overline{v}_h) = \overline{F}_h(\overline{v}_h), \quad \overline{v}_h \in \mathcal{N}C_{h,0}, \tag{1.27}$$

where

$$\begin{aligned} \overline{a}_h(u,v) &= \sum_{K\in\mathcal{T}_h} \omega_{l,K}\Big(\kappa\nabla\, u\cdot\nabla\, v\Big)(\mathbf{b}_{l,K}), \\ \overline{F}_h(v) &= \sum_{K\in\mathcal{T}_h} \omega_{l,K}\Big(fv\Big)(\mathbf{b}_{l,K}). \end{aligned} \tag{1.28}$$

We define the quadrature error functionals to estimate the effect of numerical integration, by

$$\begin{aligned} \overline{E}(\overline{\phi}) &= \int_{\overline{K}} \overline{\phi}(\overline{\mathbf{x}})\, d\overline{\mathbf{x}} - \sum_{l=1}^{L} \overline{\omega}_l\, \phi(\overline{\mathbf{b}}_l), \\ E_K(\phi) &= \int_K \phi(\mathbf{x})\, d\mathbf{x} - \sum_{l=1}^{L} \omega_{l,K}\, \phi(\mathbf{b}_{l,K}). \end{aligned} \tag{1.29}$$

Notice that two error functionals are related by the equation

$$E_K(\phi) = |\det(\mathbf{D}\mathcal{C}_K)|\, \overline{E}(\overline{\phi}).$$

The following lemma is essential for the argument.

Lemma 1.4.2. *[27](Bramble-Hilbert lemma) Let $\Omega \subset \mathbb{R}^n$ be a domain with a Lipschitz continuous boundary. Suppose that L is a continuous linear mapping on $W^{k+1,p}(\Omega)$ for some integer $k \geq 0$ and $p \in \mathbb{R}^+$. If*

$$L(p) = 0 \quad \forall p \in P_k(\Omega), \tag{1.30}$$

then there exists a constant $C(\Omega)$ such that

$$|L(v)| \leq C(\Omega)\big|\big|L\big|\big|\, |v|_{k+1,p,\Omega}. \tag{1.31}$$

Now we are ready to estimate the effect of numerical integration. First we prove uniform ellipticity of the approximate bilinear form $\overline{a}_h$. Denote the space consist of partial derivatives of functions in $\overline{P}_{\overline{K}}$ by $\nabla\, \overline{P}_{\overline{K}}$:

$$\nabla\, \overline{P}_{\overline{K}} := \mathrm{Span}\left\{ \frac{\partial \overline{u}}{\partial \overline{x}_j} \,\middle|\, \overline{u} \in \overline{P}_{\overline{K}},\ j = 1, 2 \right\}.$$

Theorem 1.4.3. *Assume that at least one of following conditions are satisfied:*

1. $\overline{E}(\overline{\phi}) = 0$ *for any* $\overline{\phi} \in \left\{ \nabla\, \overline{u} \cdot \nabla\, \overline{v} \;\middle|\; \overline{u}, \overline{v} \in \overline{P}_{\overline{K}} \right\}$,

2. $\bigcup_{l=1}^{L} \{\overline{\mathbf{b}}_l\}$ *contains a* $\nabla\, \overline{P}_{\overline{K}}$ *unisolvent subset.*

Then there exists a constant $\lambda > 0$ such that

$$\overline{a}_h(v, v) \geq \lambda ||v||^2_{1,\Omega}, \quad \forall v \in \mathcal{N}C_h. \tag{1.32}$$

Proof. Let $v_K := v|_K \in P_K$ and $\overline{v}_{\overline{K}} := v_K \circ \mathcal{C}_K \in \overline{P}_{\overline{K}}$ for arbitrary $v \in \mathcal{N}C_h$. By applying the chain rule to $\mathbf{D}\overline{v}_{\overline{K}}$, we have

$$||\mathbf{D}\overline{v}_{\overline{K}}(\overline{\mathbf{b}}_l)|| \leq ||\mathbf{D}\mathcal{C}_K||\ ||\mathbf{D}v_K(\mathbf{b}_{l,K})||.$$

Since $\mathcal{T}_h$ is a shape regular triangulations of Ω, there is a constant C such that

$$||\mathbf{D}\mathcal{C}_K||\ ||\mathbf{D}\mathcal{C}_K^{-1}|| \leq C.$$

From the scaling argument, we have

$$|v_K|_{1,K} \leq C||\mathbf{D}\mathcal{C}_K^{-1}||^{-1}\big|\det(\mathbf{D}\mathcal{C}_K^{-1})\big|^{1/2}|\overline{v}_{\overline{K}}|_{1,\overline{K}}.$$

First assume that the quadrature formula (1.25) is exact for $|\nabla\,\overline{v}|^2$. It follows that

$$\sum_{l=1}^{L}\overline{\omega}_l|\nabla\,\overline{v}_{\overline{K}}(\overline{\mathbf{b}}_l)|^2 = \int_{\overline{K}}\big|\nabla\,\overline{v}_{\overline{K}}\big|^2 d\overline{\mathbf{x}} = |\overline{v}_{\overline{K}}|^2_{1,\overline{K}} \leq C||\overline{v}_{\overline{K}}||^2_{1,\overline{K}}.$$

Now assume that $\bigcup_{l=1}^{L}\{\overline{\mathbf{b}}_l\}$ contains a $\nabla\,\overline{P}_{\overline{K}}$ unisolvent subset. Then for any $\overline{v}_{\overline{K}} \in \overline{P}_{\overline{K}}$,

$$\begin{aligned}\sum_{l=1}^{L}\overline{\omega}_l|\nabla\,v_K(\overline{\mathbf{b}}_l)|^2 = \sum_{l=1}^{L}\overline{\omega}_l\sum_{j=1}^{2}\frac{\partial\overline{v}_{\overline{K}}}{\partial\overline{x}_j}(\overline{\mathbf{b}}_l) = 0\\ \Rightarrow \frac{\partial\overline{v}_{\overline{K}}}{\partial\overline{x}_j}(\overline{\mathbf{b}}_l) = 0, \quad j = 1,2, \quad 1 \leq l \leq L\\ \Rightarrow \frac{\partial\overline{v}_{\overline{K}}}{\partial\overline{x}_j} = 0.\end{aligned}$$

Therefore the mapping

$$\overline{v}_{\overline{K}} \mapsto \Big(\sum_{l=1}^{L}\overline{\omega}_l|\nabla\,v_K(\overline{\mathbf{b}}_l)|^2\Big)^{1/2}$$

defines a norm over the quotient space $\overline{P}_{\overline{K}}/P_0(\overline{K})$. Since the mapping $\overline{v}_{\overline{K}} \mapsto |\overline{v}_{\overline{K}}|_{1,\overline{K}}$ also defines a norm over $\overline{P}_{\overline{K}}/P_0(\overline{K})$, there exists a constant $C > 0$ such that

$$\sum_{l=1}^{L} \overline{\omega}_l |\nabla v_K(\overline{\mathbf{b}}_l)|^2 \leq C||\overline{v}_{\overline{K}}||^2_{1,\overline{K}}.$$

That is, it leads to the same conclusion when we assume the exactness of quadrature formula. From the uniform ellipticity of κ, we get

$$\begin{aligned}
\sum_{l=1}^{L} \omega_{l,K}\big(\kappa\nabla v_K \cdot \nabla v_K\big)(\mathbf{b}_{l,K}) &\geq C\sum_{l=1}^{L} \omega_{l,K}|\nabla v_K(\mathbf{b}_{l,K})|^2 \\
&\geq C||\mathbf{D}\mathcal{C}_K||^{-2}\sum_{l=1}^{L} \omega_{l,K}|\nabla \overline{v}_{\overline{K}}(\overline{\mathbf{b}}_l)|^2 \\
&= C|\det(\mathbf{D}\mathcal{C}_K)|\,||\mathbf{D}\mathcal{C}_K||^{-2}\sum_{l=1}^{L} \overline{\omega}_l|\nabla \overline{v}_{\overline{K}}(\overline{\mathbf{b}}_l)|^2 \\
&= C|\det(\mathbf{D}\mathcal{C}_K)|\,||\mathbf{D}\mathcal{C}_K||^{-2}||\overline{v}_{\overline{K}}||^2_{1,\overline{K}} \\
&\geq C(||\mathbf{D}\mathcal{C}_K||\,||\mathbf{D}\mathcal{C}_K^{-1}||)^{-2}||v_K||^2_{1,K} \\
&\geq C||v_K||^2_{1,K},
\end{aligned}$$

where C is a generic constant. By combining above results, we conclude that

$$\begin{aligned}
\overline{a}_h(v,v) &= \sum_{K\in\mathcal{T}_h}\sum_{l=1}^{L} \omega_{l,K}\big(\kappa\nabla v_K \cdot \nabla v_K\big)(\mathbf{b}_{l,K}) \\
&\geq \lambda \sum_{K\in\mathcal{T}_h} ||v_K||^2_{1,K} = \lambda||v||^2_{1,\Omega}.
\end{aligned}$$

□

The following theorem estimates the effect of quadrature formulae on the approximate bilinear form $\overline{a}_h$.

Theorem 1.4.4. *Suppose that $\overline{E}(\overline{\phi}) = 0$ for any $\overline{\phi} \in \nabla\,\overline{P}_{\overline{K}}$. Then there exists*

a constant C such that

$$\begin{aligned} &\forall \kappa \in W^{1,\infty}(K), \quad \forall u, v \in P_K, \\ &\left|E_K(\kappa \nabla u \cdot \nabla v)\right| \le C h_K ||\kappa||_{1,\infty,K} |\nabla u|_{0,K} |\nabla v|_{0,K} \\ &\qquad\qquad \le C h_K ||\kappa||_{1,\infty,K} ||u||_{1,K} ||v||_{1,K}, \end{aligned} \tag{1.33}$$

where h_K denotes the diameter of K.

Proof. First we fix $\overline{\phi} \in \nabla \overline{P}_{\overline{K}}$ and estimate $\overline{E}(\overline{\psi}\,\overline{\phi})$ for arbitrary $\psi \in W^{1,\infty}(\overline{K})$. Since $W^{1,\infty}(\overline{K}) \subset C^0(\overline{K})$ and all norms are equivalent in finite dimensional vector spaces, we have

$$\begin{aligned} \left|\overline{E}(\overline{\psi}\,\overline{\phi})\right| &= \Big| \int_{\overline{K}} \overline{\psi}\,\overline{\phi}\, d\overline{\mathbf{x}} - \sum_{l=1}^{L} \overline{w}_l \left(\overline{\psi}\,\overline{\phi}\right)(\overline{\mathbf{b}}_l) \Big| \\ &\le C |\overline{\psi}\,\overline{\phi}|_{0,\infty,\overline{K}} \le C |\overline{\psi}|_{0,\infty,\overline{K}} \, |\overline{\phi}|_{0,\infty,\overline{K}} \\ &\le C ||\overline{\psi}||_{1,\infty,\overline{K}} \, |\overline{\phi}|_{0,\infty,\overline{K}} \le C ||\overline{\psi}||_{1,\infty,\overline{K}} \, |\overline{\phi}|_{0,\overline{K}}. \end{aligned}$$

Thus the linear mapping

$$\overline{\psi} \in W^{1,\infty}(\overline{K}) \mapsto \overline{E}(\overline{\psi}\,\overline{\phi})$$

is continuous with norm less than $C|\overline{\phi}|_{0,\overline{K}}$. Notice that the above mapping vanishes on $P_0(\overline{K})$ by assumption. By Lemma 1.4.2, there is a constant C such that

$$\left|\overline{E}(\overline{\psi}\,\overline{\phi})\right| \le C |\overline{\psi}|_{1,\infty,\overline{K}} \, |\overline{\phi}|_{0,K}, \quad \forall \psi \in W^{1,\infty}(\overline{K}),\ \forall \phi \in \nabla \overline{P}_{\overline{K}}.$$

Next we set $\overline{\psi} = \overline{\kappa}\,\overline{\chi}$ for $\overline{\kappa} \in W^{1,\infty}(\overline{K})$ and $\overline{\chi} \in \nabla \overline{P}(\overline{K})$. Then we have

$$|\overline{\psi}|_{1,\infty,\overline{K}} = |\overline{\kappa}\,\overline{\chi}|_{1,\infty,\overline{K}} \le C |\overline{\kappa}|_{1,\infty,\overline{K}} \, |\overline{\chi}|_{0,\infty,\overline{K}} \le C |\overline{\kappa}|_{1,\infty,\overline{K}} \, |\overline{\chi}|_{0,\overline{K}}.$$

Combining above results, we get

$$\left|\overline{E}(\overline{\kappa}\,\overline{\chi}\,\overline{\phi})\right| \le C|\overline{\kappa}|_{1,\infty,\overline{K}}\,|\overline{\chi}|_{0,\overline{K}}\,|\overline{\phi}|_{0,\overline{K}}.$$

From the inequalities

$$\begin{aligned}
&|\overline{\kappa}|_{1,\infty,\overline{K}} \le Ch_K|\kappa|_{1,\infty,K},\\
&|\overline{\chi}|_{0,\overline{K}} \le C\left|\det(\mathbf{D}\mathcal{C}_K)\right|^{-1/2}|\chi|_{0,K},\\
&|\overline{\phi}|_{0,\overline{K}} \le C\left|\det(\mathbf{D}\mathcal{C}_K)\right|^{-1/2}|\phi|_{0,K},
\end{aligned}$$

we obtain

$$\begin{aligned}
\left|E(\kappa\chi\phi)\right| &= |\det(\mathbf{D}\mathcal{C}_K)|\left|\overline{E}(\overline{\kappa}\,\overline{\chi}\,\overline{\phi})\right|\\
&\le Ch_K|\kappa|_{1,\infty,K}\,|\chi|_{0,K}\,|\phi|_{0,K}.
\end{aligned}$$

Finally set $\chi = \nabla u$ and $\phi = \nabla v$ to get the conclusion. □

Now we estimate the effect of numerical integration on the right hand side linear functional $\overline{F}_h$.

Theorem 1.4.5. *Suppose that $\overline{E}(\overline{\phi}) = 0$ for any $\overline{\phi} \in P_0(\overline{K})$. Then for arbitrary $f \in W^{1,\infty}(\Omega)$ and $\phi \in P_K$, there exists a constant C such that*

$$\left|E_K(f\phi)\right| \le Ch_K\big(area(K)\big)^{1/2}||f||_{1,\infty,K}||\phi||_{1,K}, \tag{1.34}$$

where h_K denotes the diameter of K.

Proof. Since $W^{1,\infty}(\overline{K}) \subset C^0(\overline{K})$, we have for arbitrary $\overline{\phi} \in \overline{P}_{\overline{K}}$ that

$$\left|\overline{E}(\overline{\phi})\right| \le C|\overline{\phi}|_{0,\infty,\overline{K}} \le C||\overline{\phi}||_{1,\infty,\overline{K}}.$$

We apply Lemma 1.4.2 to the linear mapping

$$\overline{\phi} \in W^{1,\infty}(\overline{K}) \mapsto \overline{E}(\overline{\phi}),$$

which is continuous with norm less than C, and vanishes on $P_0(\overline{K})$ by assumption. Then there is a constant C such that

$$\left|\overline{E}(\overline{\phi})\right| \leq C|\overline{\phi}|_{1,\infty,\overline{K}}. \tag{1.35}$$

By generalized Leibniz formula, we have for arbitrary $\overline{f} \in W^{1,\infty}(\overline{K})$ that

$$\left|\overline{f}\,\overline{\phi}\right| \leq C\big(|\overline{f}|_{1,\infty,\overline{K}}\,|\overline{\phi}|_{0,\infty,\overline{K}} + |\overline{f}|_{0,\infty,\overline{K}}\,|\overline{\phi}|_{1,\infty,\overline{K}}\big).$$

Observe the following inequalities with $j = 0, 1$:

$$\begin{aligned}
&|\overline{f}|_{1-j,\infty,\overline{K}} \leq Ch_K^{1-j}|f|_{1-j,\infty,K}, \\
&|\overline{\phi}|_{j,\overline{K}} \leq Ch_K^{j}\left|\det(\mathbf{D}\mathcal{C}_K)\right|^{-1/2}|\phi|_{j,K}.
\end{aligned}$$

Then by (1.35), we conclude that

$$\begin{aligned}
|E(f\phi)| &= |\det(\mathbf{D}\mathcal{C}_K)|\left|\overline{E}(\overline{f}\,\overline{\phi})\right| \\
&\leq Ch_K\left|\det(\mathbf{D}\mathcal{C}_K)\right|^{1/2}\big(|f|_{1,\infty,K}\,|\phi|_{0,\infty,K} + |f|_{0,\infty,K}\,|\phi|_{1,\infty,K}\big) \\
&\leq Ch_K\big(\mathrm{area}(K)\big)^{1/2}||f||_{1,\infty,K}\,||\phi||_{1,K}.
\end{aligned}$$

□

Finally we estimate the effect of numerical integration by combining the above theorems.

Theorem 1.4.6. *Let u and $\overline{u}_h$ are the solutions of* (1.22) *and* (1.27)*, respectively. Assume that $\overline{a}_h$ is uniformly elliptic and $\overline{E}(\overline{\phi}) = 0$ for any $\overline{\phi} \in \nabla \overline{P}_{\overline{K}}$. Then we have the following error estimate*

$$||u - \overline{u}_h||_{1,\Omega} \leq Ch\big(||\kappa||_{1,\infty,\Omega}\,||u||_{2,\Omega} + ||f||_{1,\infty,\Omega}\big). \tag{1.36}$$

Proof. We exploit the uniform ellipticity of $\overline{a}_h$. Let u_h be the solutions of (1.23). Then we have for arbitrary $v_h \in \mathcal{N}C_h$, that

$$\begin{aligned}
\lambda||\overline{u}_h - v_h||^2_{1,\Omega} &\leq \overline{a}_h(\overline{u}_h - v_h, \overline{u}_h - v_h) \\
&= \overline{a}_h(u_h - v_h, \overline{u}_h - v_h) + \overline{a}_h(\overline{u}_h - u_h, \overline{u}_h - v_h) \\
&= \overline{a}_h(u_h - v_h, \overline{u}_h - v_h) + \big(\overline{F}_h(\overline{u}_h - v_h) - \overline{a}_h(u_h, \overline{u}_h - v_h)\big) \\
&\qquad \pm a_h(u_h, \overline{u}_h - v_h) \\
&= \overline{a}_h(u_h - v_h, \overline{u}_h - v_h) + \big(a_h(u_h, \overline{u}_h - v_h) - \overline{a}_h(u_h, \overline{u}_h - v_h)\big) \\
&\qquad + \big(\overline{F}_h(\overline{u}_h - v_h) - F_h(\overline{u}_h - v_h)\big).
\end{aligned}$$

Denote by $w_h := \overline{u}_h - v_h$. It follows that

$$\begin{aligned}
\lambda||\overline{u}_h - v_h||_{1,\Omega} &\leq C||u_h - v_h||_{1,\Omega} + \frac{|a_h(u_h, \overline{u}_h - v_h) - \overline{a}_h(u_h, \overline{u}_h - v_h)|}{||\overline{u}_h - v_h||_{1,\Omega}} \\
&\qquad + \frac{|F_h(\overline{u}_h - v_h) - \overline{F}_h(\overline{u}_h - v_h)|}{||\overline{u}_h - v_h||_{1,\Omega}} \\
&\leq C \inf_{v_h \in \mathcal{N}C^h} ||u_h - v_h||_{1,\Omega} \\
&+ \sup_{w_h \in \mathcal{N}C^h} \Big(\frac{|a_h(u_h, w_h) - \overline{a}_h(u_h, w_h)|}{||w_h||_{1,\Omega}} + \frac{|\overline{F}_h(w_h) - F_h(w_h)|}{||w_h||_{1,\Omega}}\Big).
\end{aligned}$$

If we take $v_h = u_h$, the above inequality is simplified to

$$||\overline{u}_h - u_h||_{1,\Omega} \leq \frac{1}{\lambda} \sup_{w_h \in \mathcal{N}C^h} \Big(\frac{|a_h(u_h, w_h) - \overline{a}_h(u_h, w_h)|}{||w_h||_{1,\Omega}} + \frac{|\overline{F}_h(w_h) - F_h(w_h)|}{||w_h||_{1,\Omega}}\Big).$$

It remains to estimate two consistency error terms. First,

$$\begin{aligned}
|a_h(u_h, w_h) - \overline{a}_h(u_h, w_h)| &\leq \sum_{K \in \mathcal{T}_h} \left| E_K(\kappa \nabla u_h \cdot \nabla w_h) \right| \\
&\leq \sum_{K \in \mathcal{T}_h} h_K ||\kappa||_{1,\infty,K} \, ||u_h||_{1,K} \, ||w_h||_{1,K} \\
&\leq Ch ||\kappa||_{1,\infty,\Omega} \, ||u_h||_{1,\Omega} \, ||w_h||_{1,\Omega} \\
&\leq Ch ||\kappa||_{1,\infty,\Omega} \, ||u||_{2,\Omega} \, ||w_h||_{1,\Omega}.
\end{aligned}$$

In the last inequality, we use

$$\begin{aligned}
||u_h||_{1,\Omega} &\leq ||u||_{1,\Omega} + ||u - u_h||_{1,\Omega} \\
&\leq ||u||_{1,\Omega} + Ch||u||_{2,\Omega} \leq C||u||_{2,\Omega}.
\end{aligned}$$

Second,

$$\begin{aligned}
|\overline{F}_h(w_h) - F_h(w_h)| &\leq \sum_{K \in \mathcal{T}_h} \left| E_K(f w_h) \right| \\
&\leq C \sum_{K \in \mathcal{T}_h} h_K \big(\text{area}(K)\big)^{1/2} ||f||_{1,\infty,K} \, ||w_h||_{1,K} \\
&\leq Ch \big(\text{area}(\Omega)\big)^{1/2} ||f||_{1,\infty,\Omega} \, ||w_h||_{1,\Omega}.
\end{aligned}$$

The theorem follows by combining above two results with triangle inequality. That is,

$$\begin{aligned}
||u - \overline{u}_h||_{1,\Omega} &\leq ||u - u_h||_{1,\Omega} + ||u_h - \overline{u}_h||_{1,\Omega} \\
&\leq Ch||u||_{2,\Omega} + Ch\big(||\kappa||_{1,\infty,\Omega} \, ||u||_{2,\Omega} + \big(\text{area}(\Omega)\big)^{1/2} ||f||_{1,\infty,\Omega}\big) \\
&\leq Ch\big(||\kappa||_{1,\infty,\Omega} \, ||u||_{2,\Omega} + ||f||_{1,\infty,\Omega}\big).
\end{aligned}$$

□

1.4.2 Quadrature formula on $\overline{K}$

Now we turn to develop a quadrature formulae on $\overline{K}$, which is defined as

$$\int_{\overline{K}} \overline{\phi}(\overline{\mathbf{x}})\, d\overline{\mathbf{x}} \approx \sum_{l=1}^{L} \overline{\omega}_l\, \overline{\phi}(\overline{\mathbf{b}_l}), \tag{1.37}$$

where $\overline{\omega}_l$ and $\overline{\mathbf{b}}_l$, $l = 1, \cdots, L$, are positive weights and nodes, respectively. In [25], the basis functions are at most of degree two so that quadrature formula of degree two are found. However our element has high-order degree basis to fulfill the MVP, we require another quadrature formula.

Observe that $\overline{K}$ is a right-angled triangle in each quadrant. It makes us easy to compute integrals of polynomials on $\overline{K}$ exactly. Denote the triangle in j-th quadrant by $\overline{T}_j$ for $j = 1, 2, 3, 4$. Following results are obtained by direct computation with applying integration by parts repeatedly.

$$\begin{aligned}
\int_{\overline{T}_1} \overline{x}_1^i \overline{x}_2^j\, d\overline{\mathbf{x}} &= \frac{i!j!}{(2+i+j)!}, \\
\int_{\overline{T}_2} \overline{x}_1^i \overline{x}_2^j\, d\overline{\mathbf{x}} &= -\frac{i!j!}{(2+i+j)!} h_1^{i+1}, \\
\int_{\overline{T}_3} \overline{x}_1^i \overline{x}_2^j\, d\overline{\mathbf{x}} &= \frac{i!j!}{(2+i+j)!} h_1^{i+1} h_2^{j+1}, \\
\int_{\overline{T}_4} \overline{x}_1^i \overline{x}_2^j\, d\overline{\mathbf{x}} &= -\frac{i!j!}{(2+i+j)!} h_2^{j+1}.
\end{aligned}$$

Then we get

$$\begin{aligned}
\int_{\overline{K}} \overline{x}_1^i \overline{x}_2^j\, d\overline{\mathbf{x}} &= \sum_{k=1}^{4} \int_{\overline{T}_k} \overline{x}_1^i \overline{x}_2^j\, d\overline{\mathbf{x}} \\
&= \frac{i!j!}{(2+i+j)!} (1 - h_1^{i+1})(1 - h_2^{j+1}).
\end{aligned}$$

To preserve the order of convergence, we may find the quadrature formula exact for functions in $\nabla \overline{P}_{\overline{K}}$, where the space $\nabla \overline{P}_{\overline{K}}$ of dimension three is defined as

$$\begin{aligned}\nabla \overline{P}_{\overline{K}} :&= \operatorname{Span}\Big\{ \frac{\partial \overline{u}}{\partial \overline{x}_j} \Big| \ \overline{u} \in \overline{P}_{\overline{K}},\ j = 1, 2 \Big\} \\ &= \operatorname{Span}\Big\{ 1, \frac{\partial \overline{u}}{\partial \overline{x}_1}, \frac{\partial \overline{u}}{\partial \overline{x}_2} \Big\}.\end{aligned}$$

Also we may seek the formula where the set of quadrature nodes $\bigcup_{l=1}^{L} \{\overline{\mathbf{b}}_l\}$ contains a $\nabla \overline{P}_{\overline{K}}$ unisolvent subset. Here we find three-point quadrature formula where the nodes are symmetric with respect to the barycenter $\overline{\mathcal{R}}$ of $\overline{K}$, which is given by

$$\overline{\mathcal{R}} = \Big(\frac{1 + h_1}{3}, \frac{1 + h_2}{3} \Big).$$

We further impose following assumptions to reduce computational burden and get explicit quadrature formula:

1. The nodes are given as

$$\overline{\mathbf{b}}_1 = \overline{\mathcal{R}} + (s, t), \quad \overline{\mathbf{b}}_2 = \overline{\mathcal{R}}, \quad \overline{\mathbf{b}}_3 = \overline{\mathcal{R}} - (s, t).$$

2. The weights are equal to $w = \frac{\operatorname{area}(\overline{K})}{3}$.

Then it reduces to find (s, t) which satisfy

$$\sum_{l=1}^{3} \frac{\partial \overline{u}}{\partial \overline{x}_j} (\overline{\mathbf{b}}_l) = I_j \quad \text{for } j = 1, 2,$$

where

$$I_1 = -\frac{(1 + h_2)(h_1^2 + h_2^2 + h_1 + 2)}{4} \quad \text{and} \quad I_2 = -\frac{(1 + h_1)(h_1^2 + h_2^2 + h_2 + 2)}{4}.$$

A use of symbolic package gives the formula for arbitrary quadrilateral mesh (Figure 1.4) such that t is the root of $a_1 z^4 + a_2 z^2 + a_3 = 0$, where

$$
\begin{aligned}
a_1 &= 17781120\,h_1^4 + 71124480\,h_1^3 + (-34477488\,h_2^2 - 68954976\,h_2 + 72209232)\,h_1^2 \\
&\quad + (-68954976\,h_2^2 - 137909952\,h_2 + 2169504)\,h_1 \\
&\quad + 17781120\,h_2^4 + 71124480\,h_2^3 + 72209232\,h_2^2 + 2169504\,h_2 + 1084752, \\
a_2 &= -1234800\,h_1^6 - 3408048\,h_1^5 + (923112\,h_2^2 + 4513392\,h_2 - 1595880)\,h_1^4 \\
&\quad + (-3103776\,h_2^2 + 4461120\,h_2 - 3795264)\,h_1^3 \\
&\quad + (1767240\,h_2^4 - 841176\,h_2^3 - 13270608\,h_2^2 - 945720\,h_2 - 8805528)\,h_1^2 \\
&\quad + (7344720\,h_2^4 + 13558608\,h_2^3 + 9324000\,h_2^2 + 18019728\,h_2 + 832896)\,h_1 \\
&\quad - 1764000\,h_2^6 - 4868640\,h_2^5 - 1831320\,h_2^4 - 1829016\,h_2^3 \\
&\quad - 7892280\,h_2^2 - 1196424\,h_2 - 308448, \\
a_3 &= (h_1 + 1)^2\,(h_2 + 1)^2\,(229\,h_1^2 - 145\,h_2^2 + 26\,h_1 - 263\,h_2 + 84)^2,
\end{aligned}
$$

and s is given by

$$
s = \frac{(h_1 + 1)(h_2 + 1)(1836\,t^2 - 229\,h_1^2 + 145\,h_2^2 - 26\,h_1 + 263\,h_2 - 84)}{504\,t\,(7\,h_1^2 - 10\,h_2^2 + 14\,h_1 - 20\,h_2 - 3)}.
$$

We report some numerical values of w, s, t for uniform trapezoidal mesh with parameter θ (Figure 1.3) in Table 1.1 for simple use.

θ	w	s	t
0	0.6666666667	α	α
0.2	0.4629629632	0.2888473372	0.2888473372
0.4	0.3401360547	0.2450221177	0.2450221177
0.6	0.2604166667	0.2106058842	0.2106058842
0.8	0.2057613168	0.1823862558	0.1823862558

Table 1.1. Quadrature formula for trapezoidal meshes with parameter θ. We have $s = t$ for these meshes since $h_1 = h_2$. For the rectangular mesh case ($\theta = 0$), any nodes $(\alpha, \alpha) \in \overline{K}$ can be chosen.

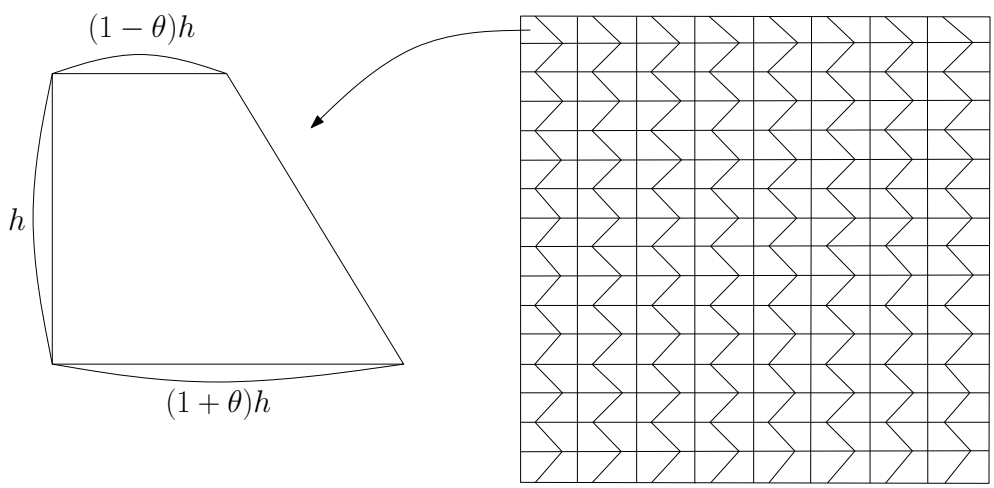


Figure 1.3. An uniform trapezoidal mesh with parameter θ.

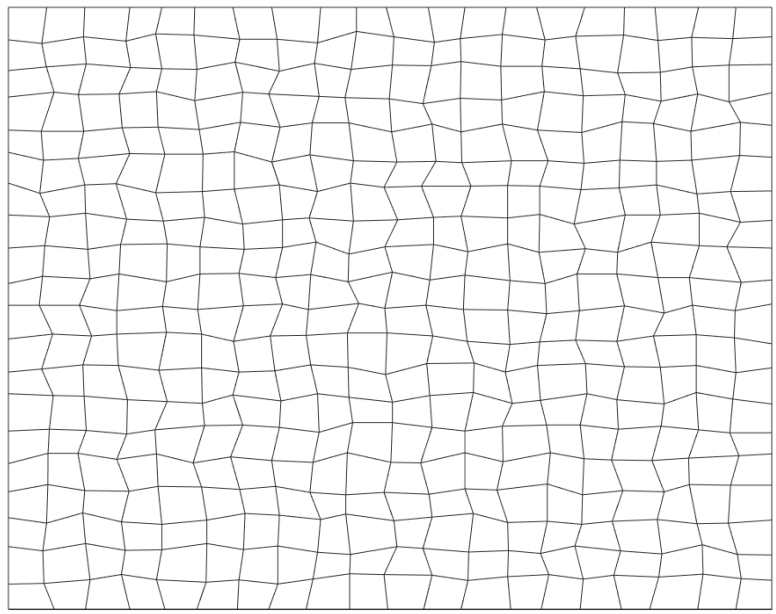


Figure 1.4. A nonuniform quadrilateral mesh.

1.5 Numerical results

Example 1.5.1. *[25] Consider the following elliptic problem:*

$$\begin{cases} -\Delta u = f \ \text{ in } \Omega, \\ \quad u = 0 \ \text{ on } \partial\Omega, \end{cases} \tag{1.38}$$

where $\Omega = (0,1)^2$, *and the source term* f *is generated by the exact solution*

$$u(x_1, x_2) = \sin(2\pi x_1)\sin(2\pi x_2)\big(x_1^3 - x_2^4 + x_1^2 x_2^3\big).$$

We use 4×4 Gauss formula and the proposed formula to compute the

components of corresponding linear system. Error behaviors for uniform trapezoidal mesh with various θ are reported in the below tables, which show the optimal convergence rates for both quadrature formula. Notice that the errors become larger when we use more perturbed meshes with larger θ. We also observe the optimal convergence rate for a nonuniform quadrilateral mesh in Table 1.7.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.380		0.444E-02		0.433		0.455E-02	
32	0.191	0.99	0.111E-02	1.99	0.216	1.00	0.114E-02	2.00
64	0.954E-01	1.00	0.278E-03	2.00	0.108	1.00	0.285E-03	2.00
128	0.477E-01	1.00	0.696E-04	2.00	0.541E-01	1.00	0.713E-04	2.00
256	0.239E-01	1.00	0.174E-04	2.00	0.271E-01	1.00	0.178E-04	2.00

Table 1.2. Results of Example 1.5.1 for uniform trapezoidal mesh of $\theta = 0$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.393		0.496E-02		0.457		0.483E-02	
32	0.198	0.99	0.126E-02	1.98	0.229	1.00	0.122E-02	1.98
64	0.995E-01	1.00	0.317E-03	1.99	0.115	1.00	0.308E-03	1.99
128	0.498E-01	1.00	0.798E-04	1.99	0.572E-01	1.00	0.772E-04	2.00
256	0.249E-01	1.00	0.200E-04	2.00	0.286E-01	1.00	0.193E-04	2.00

Table 1.3. Results of Example 1.5.1 for uniform trapezoidal mesh of $\theta = 0.2$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.426		0.619E-02		0.433		0.637E-02	
32	0.216	0.98	0.160E-02	1.95	0.219	0.99	0.164E-02	1.95
64	0.109	0.99	0.412E-03	1.96	0.110	1.00	0.419E-03	1.97
128	0.549E-01	0.99	0.105E-03	1.98	0.549E-01	1.00	0.106E-03	1.98
256	0.275E-01	1.00	0.264E-04	1.99	0.275E-01	1.00	0.267E-04	1.99

Table 1.4. Results of Example 1.5.1 for uniform trapezoidal mesh of $\theta = 0.4$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.472		0.795E-02		0.476		0.822E-02	
32	0.242	0.96	0.211E-02	1.92	0.243	0.97	0.216E-02	1.93
64	0.123	0.98	0.551E-03	1.93	0.123	0.99	0.559E-03	1.95
128	0.618E-01	0.99	0.142E-03	1.96	0.616E-01	0.99	0.143E-03	1.97
256	0.310E-01	1.00	0.359E-04	1.98	0.309E-01	1.00	0.361E-04	1.98

Table 1.5. Results of Example 1.5.1 for uniform trapezoidal mesh of $\theta = 0.6$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u-\overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u-\overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u-\overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u-\overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.528		0.102E-01		0.532		0.106E-01	
32	0.273	0.95	0.277E-02	1.88	0.274	0.96	0.282E-02	1.91
64	0.139	0.97	0.740E-03	1.90	0.139	0.98	0.744E-03	1.92
128	0.704E-01	0.99	0.192E-03	1.94	0.702E-01	0.99	0.192E-03	1.95
256	0.353E-01	0.99	0.491E-04	1.97	0.352E-01	1.00	0.489E-04	1.98

Table 1.6. Results of Example 1.5.1 for uniform trapezoidal mesh of $\theta = 0.8$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u-\overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u-\overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u-\overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u-\overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.383		0.460E-02		1.42		0.961E-02	
32	0.193	0.99	0.116E-02	1.98	0.677	1.07	0.215E-02	2.16
64	0.966E-01	1.00	0.290E-03	2.00	0.363	0.90	0.597E-03	1.85
128	0.483E-01	1.00	0.726E-04	2.00	0.184	0.99	0.151E-03	1.98
256	0.242E-01	1.00	0.181E-04	2.00	0.921E-01	0.99	0.379E-04	2.00

Table 1.7. Results of Example 1.5.1 for nonuniform quadrilateral mesh.

Example 1.5.2. *Consider the following elliptic problem:*

$$\begin{cases} -\nabla\cdot\big(\kappa(\mathbf{x})\nabla\, u\big) = f \ \textit{in}\ \Omega, \\ \qquad\qquad\qquad u = 0 \ \textit{on}\ \partial\Omega, \end{cases} \tag{1.39}$$

where $\Omega = (0,1)^2$ *and* $\kappa(\mathbf{x}) = 1 + (1+x_1)(1+x_2) + \epsilon\sin(10\pi x_1)\sin(5\pi x_2)$. *The source term* f *is generated by the exact solution*

$$u(x_1,x_2) = \sin(3\pi x_1)x_2(1-x_2) + \epsilon\sin(\pi x_1/\epsilon)\sin(\pi x_2/\epsilon).$$

In this example, the heterogeneous coefficient κ is considered. We only report the result of $\epsilon = 0.2$ case since we have similar error behaviors for other ϵ values. Optimal convergence rates is observed for both quadrature formula.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.899		0.987E-02		0.931		0.122E-01	
32	0.455	0.98	0.247E-02	2.00	0.465	1.00	0.309E-02	1.98
64	0.228	1.00	0.619E-03	2.00	0.233	1.00	0.775E-03	1.99
128	0.114	1.00	0.155E-03	2.00	0.116	1.00	0.194E-03	2.00
256	0.571E-01	1.00	0.387E-04	2.00	0.581E-01	1.00	0.485E-04	2.00

Table 1.8. Results of Example 1.5.2 for uniform trapezoidal mesh of $\theta = 0$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.924		0.112E-01		0.968		0.121E-01	
32	0.470	0.97	0.288E-02	1.95	0.491	0.98	0.307E-02	1.98
64	0.237	0.99	0.732E-03	1.98	0.246	0.99	0.770E-03	1.99
128	0.118	1.00	0.184E-03	1.99	0.123	1.00	0.193E-03	2.00
256	0.593E-01	1.00	0.462E-04	2.00	0.617E-01	1.00	0.482E-04	2.00

Table 1.9. Results of Example 1.5.2 for uniform trapezoidal mesh of $\theta = 0.2$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.978		0.141E-01		0.996		0.154E-01	
32	0.506	0.95	0.381E-02	1.89	0.511	0.96	0.400E-02	1.94
64	0.256	0.98	0.986E-03	1.95	0.257	0.99	0.102E-02	1.97
128	0.129	0.99	0.250E-03	1.98	0.129	1.00	0.256E-03	1.99
256	0.644E-01	1.00	0.630E-04	1.99	0.644E-01	1.00	0.643E-04	2.00

Table 1.10. Results of Example 1.5.2 for uniform trapezoidal mesh of $\theta = 0.4$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	1.05		0.179E-01		1.06		0.200E-01	
32	0.550	0.93	0.498E-02	1.85	0.552	0.94	0.520E-02	1.94
64	0.281	0.97	0.129E-02	1.95	0.281	0.98	0.132E-02	1.97
128	0.141	0.99	0.329E-03	1.97	0.141	0.99	0.334E-03	1.99
256	0.709E-01	1.00	0.830E-04	1.99	0.706E-01	1.00	0.839E-04	1.99

Table 1.11. Results of Example 1.5.2 for uniform trapezoidal mesh of $\theta = 0.6$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	1.13		0.223E-01		1.15		0.256E-01	
32	0.599	0.91	0.633E-02	1.82	0.601	0.93	0.662E-02	1.95
64	0.307	0.97	0.164E-02	1.95	0.306	0.97	0.167E-02	1.99
128	0.155	0.99	0.415E-03	1.98	0.154	0.99	0.420E-03	1.99
256	0.776E-01	1.00	0.105E-03	1.99	0.774E-01	1.00	0.105E-03	1.99

Table 1.12. Results of Example 1.5.2 for uniform trapezoidal mesh of $\theta = 0.8$.

$\frac{1}{h}$	4×4 Gauss formula				Our formula			
	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{1,\Omega}$	ratio	$\|\|u - \overline{u}_h\|\|_{0,\Omega}$	ratio
16	0.906		0.102E-01		1.58		0.223E-01	
32	0.461	0.98	0.259E-02	1.98	0.756	1.06	0.463E-02	2.27
64	0.231	1.00	0.645E-03	2.00	0.377	1.00	0.843E-03	2.46
128	0.115	1.00	0.161E-03	2.00	0.190	0.99	0.209E-03	2.01
256	0.578E-01	1.00	0.404E-04	2.00	0.956E-01	0.99	0.513E-04	2.03

Table 1.13. Results of Example 1.5.2 for nonuniform quadrilateral mesh.

Chapter 2

Nonconforming Generalized Multiscale Finite Element Method

2.1 Introduction

Many real-world problems in science and engineering are modeled with highly heterogeneous coefficients, which are of essentially multiscale nature. Applications include quantum mechanical modeling, groundwater transport, oil reservoir simulation, integrated computational materials engineering, climate modeling, multiscale decision making, and so on. Since fine-resolution discretization is required to capture the high-contrast, it is extremely expensive to solve problems without introducing any model reduction technique. During the last decades many multiscale methods have been actively developed such as multiscale finite element methods [32, 33, 36, 37, 38], heterogeneous multiscale

methods [30, 31], and multiscale hybridizable discontinuous Galerkin methods [34, 35]. All such methods use at least two-scale (micro, macro) grids and build multiscale basis functions by solving local harmonic problems in each macro block. First, one constructs local snapshot spaces which capture the microscale heterogeneity of coefficients. Then dimension reduction techniques are applied to build so-called offline spaces, which are used as the multiscale basis functions. Moment spaces also need to be considered in order to impose continuity between local offline spaces.

This chapter is organized as follows. In section 2, we review a framework of nonconforming generalized multiscale finite element methods. Then the process for constructing multiscale finite element spaces is presented in section 3. We present an energy norm error estimate in section 4. In section 5, some numerical results are provided.

2.2 Framework of nonconforming generalized multiscale finite element methods

In this section we briefly review a framework of generalized multiscale finite element method(GMsFEM) using nonconforming element, following [38]. We only consider two-dimensional elliptic boundary problems here, but the framework can be extended to higher dimensional cases and used for other multiscale problems.

2.2.1 Preliminaries

Let U be any open subset of $\mathbb{R}^2$. We denote the seminorm, norm, and inner product of the Sobolev space $H^k(U)$ by $|\cdot|_{k,U}$, $||\cdot||_{k,U}$, and $(\cdot,\cdot)_{k,U}$ respectively. For the space $H^0(U) = L^2(U)$, we abbreviate $(\cdot,\cdot)_{k,U}$ as $(\cdot,\cdot)_U$. Now for given

$f \in H^{-1}(\Omega)$, consider the following elliptic boundary problem

$$\begin{cases} -\nabla \cdot \big(\kappa(\mathbf{x}) \nabla u\big) = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases} \tag{2.1}$$

where Ω is a simply connected polygonal domain in $\mathbb{R}^2$, and κ is a highly heterogeneous coefficient. The weak formulation of (2.1) is to seek $u \in H_0^1(\Omega)$ such that

$$a(u, v) = F(v), \quad v \in H_0^1(\Omega) \tag{2.2}$$

where $a(u, v) = \int_\Omega \kappa \nabla u \cdot \nabla v \, d\mathbf{x}$ and $F(v) = \int_\Omega f v \, d\mathbf{x}$. Let $\mathcal{T}_h := \bigcup_{j=1}^{N_h} \{T_j\}$ be a family of shape regular triangulations of Ω and V_h be a finite element basis function space based on $\mathcal{T}_h$. The mesh parameter h is given by

$$h = \max_{j=1,\cdots,N_h} \operatorname{diam}(T_j).$$

Let $V_{h,0}$ be the set of all elements in V_h, whose DOFs related to the boundary $\partial\Omega$ vanish. Then the finite element approximation of (2.2) is defined as the solution $u_h \in V_{h,0}$ of the discrete problem

$$a_h(u_h, v_h) = F_h(v_h), \quad v_h \in V_{h,0}, \tag{2.3}$$

where $a_h(u, v) = \sum_{T_j \in \mathcal{T}_h} \int_{T_j} \kappa \nabla u \cdot \nabla v \, d\mathbf{x}$ and $F_h(v) = \sum_{T_j \in \mathcal{T}_h} \int_{T_j} f v \, d\mathbf{x}$. In GMsFEM, we also need to have another shape regular triangulations $\mathcal{T}^H := \bigcup_{J=1}^{N^H} \{T^J\}$ of Ω. We suppose that every $T^J \in \mathcal{T}^H$ consists of a connected union of $T_j \in \mathcal{T}_h$, which makes $\mathcal{T}_h$ be a refinement of $\mathcal{T}^H$. Here, and in what follows, we refer two triangulations $\mathcal{T}_h$ and $\mathcal{T}^H$ to microscale and macroscale

triangulations, respectively. The macro mesh parameter H is given by

$$H = \max_{J=1,\cdots,N^H} \operatorname{diam}(T^J).$$

Let V^H be a finite element basis function space associated with $\mathcal{T}^H$, and $V^{H,0}$ be the set of all elements in V^H, whose DOFs related to $\partial\Omega$ vanish. Then the generalized multiscale finite element approximation of (2.2) is equivalent to find $u^H \in V^{H,0}$ such that

$$a_h(u^H, v^H) = F_h(v^H), \quad v^H \in V^{H,0}. \tag{2.4}$$

2.2.2 Framework of nonconforming GMsFEM

Success of GMsFEM depends on the construction of corresponding finite element space. V^H must contain the essential properties of V_h as well as the coefficient κ, while the dimension of V^H is significantly reduced compared to that of V_h.

The generalized multiscale finite element space V^H is composed of two components. First one is the offline function space which is a spectral decomposition of the snapshot function space, and used to represent the solution in each macro element. Second one is the moment function space which is used to impose continuity between local offline functions. Let microscale basis function space V_h be given. For each macro element $T \in \mathcal{T}^H$, denote the restriction of V_h to T by $V_h(T)$. Also denote the set of all macro edges in $\mathcal{T}^H$ by $\mathcal{E}^H := \bigcup_{J=1}^{N^E}\{E^J\}$, and the set of all interior macro edges by $\mathcal{E}^{H,0}$. Then the process of constructing GMsFE spaces is organized into the following framework:

1. Construct a snapshot function space $V^{\text{snap}} = \bigcup_{T\in\mathcal{T}^H} V^{\text{snap}}(T)$, where

$V^{\text{snap}}(T)$ is a subspace of $V_h(T)$ for each macro element $T \in \mathcal{T}^H$. In general, $V^{\text{snap}}(T)$ is chosen to be the span of $\kappa-$harmonic functions in T.

2. Construct an offline function space $V^{\text{off}} = \bigcup_{T\in\mathcal{T}^H} V^{\text{off}}(T)$, where $V^{\text{off}}(T)$ is obtained by applying a suitable dimension reduction technique to $V^{\text{snap}}(T)$ for each macro element $T \in \mathcal{T}^H$. We may use generalized eigenvalue decomposition, the singular value decomposition, the proper orthogonal decomposition, and so on.

3. Construct a moment function space $\mathcal{M}_H = \bigcup_{E\in\mathcal{E}^H} \mathcal{M}_H(E)$. $\mathcal{M}_H(E)$ may consist of local $\kappa-$harmonic functions in appropriate neighborhood of E. The moment functions are used to glue offline functions through each macro interior edge $E \in \mathcal{E}^{H,0}$.

4. Construct the nonconforming GMsFE spaces V^H and $V^{H,0}$ based on V^{off} and $\mathcal{M}_H$. They are defined as

$$V^H = \Big\{\psi \in V^{\text{off}} \;\Big|\; < [\psi]_E, \zeta >_E = 0,\ \forall \zeta \in \mathcal{M}_H(E),\ \forall E \in \mathcal{E}^{H,0}\Big\},$$
$$V^{H,0} = \Big\{\psi \in V^{\text{off}} \;\Big|\; < [\psi]_E, \zeta >_E = 0,\ \forall \zeta \in \mathcal{M}_H(E),\ \forall E \in \mathcal{E}^{H}\Big\}.$$

Here $[\psi]_E$ stands for the jump of ψ across macro edge E.

2.3 Construction of multiscale finite element spaces

In this section we present the detailed process for constructing GMsFE spaces. We may set microscale space V_h as the DSSY nonconforming finite element space. It is remarkable that V_h can be another finite element space or any space associated with $\mathcal{T}_h$ induced by finite difference method, spectral method, and so on.

2.3.1 Snapshot function space V^{snap}

We first construct local snapshot function space $V^{\text{snap}}(T)$ in each macro element $T \in \mathcal{T}^H$. Since snapshot functions are used to compute multiscale basis functions, we may choose $V^{\text{snap}}(T)$ as all microscale basis functions in T. Or smaller space such as the span of $\kappa-$harmonic functions in T can be considered to reduce the cost of constructing V^H. Let $\widetilde{\phi}_j^T \in V_h(T)$ be the solutions of following local $\kappa-$harmonic problems:

$$\begin{cases} -\nabla \cdot \left(\kappa(\mathbf{x}) \nabla \widetilde{\phi}_j^T\right) = 0 \text{ in } T, \\ \widetilde{\phi}_j^T = \delta_j^T \text{ on } \partial T, \end{cases} \tag{2.5}$$

where $\delta_j^T \in V_h(T)$ is the function which equals to one for the $j-$th microscale mesh DOF on ∂T and zeros for the other DOFs on ∂T. Denote the number of all snapshot functions in T by $\mathcal{N}^{\text{snap}}(T)$ and zero extension of $\widetilde{\phi}_j^T$ outside T by ϕ_j^T. Then the local snapshot function space $V^{\text{snap}}(T)$ is defined as the space spanned by ϕ_j^T:

$$V^{\text{snap}}(T) = \text{Span}\left\{ \phi_j^T \in V_h(T) \,\middle|\, j = 1, \cdots, \mathcal{N}^{\text{snap}}(T) \right\}.$$

Finally the snapshot function space V^{snap} is defined as the union of such local snapshot function spaces:

$$V^{\text{snap}} = \bigcup_{T \in \mathcal{T}^H} V^{\text{snap}}(T).$$

Oversampling technique

We can apply the oversampling technique to reduce the resonance error caused by wrong (local) boundary condition δ_j^T. We solve (2.5) on an extended region

T^+ and restrict the solution $\phi_j^{T^+}$ to the original domain T. We denote the local oversampled snapshot function space by $V^{\mathrm{snap},+}(T^+)$, which is defined as

$$V^{\mathrm{snap},+}(T^+) = \mathrm{Span}\left\{\phi_j^{T^+} \in V_h(T^+) \,\middle|\, j = 1, \cdots, \mathcal{N}^{\mathrm{snap},+}(T^+)\right\}.$$

Then the oversampled snapshot function space $V^{\mathrm{snap},+}$ is given as follows:

$$V^{\mathrm{snap},+} = \bigcup_{T \in \mathcal{T}^H} V^{\mathrm{snap},+}(T^+).$$

2.3.2 Offline function space V^{off}

Offline function space V^{off} is obtained by applying a suitable dimension reduction technique to the snapshot function space V^{snap}. For example, we may use generalized eigenvalue decomposition. For each macro element $T \in \mathcal{T}^H$, consider the following spectral problem to find $(\lambda_j^T, \psi_j^T) \in \mathbb{R} \times V^{\mathrm{snap}}(T)$:

$$a_T(\psi_j^T, \phi^T) = \lambda_j^T (\kappa \psi_j^T, \phi^T)_T, \quad \forall \phi^T \in V^{\mathrm{snap}}(T), \tag{2.6}$$

where $a_T(\psi, \phi) = \sum_{T_j \in T} \int_{T_j} \kappa \nabla \psi \cdot \nabla \phi \, d\mathbf{x}$. We suppose that the eigenvalues are sorted in ascending order as

$$0 \leq \lambda_1^T \leq \lambda_2^T \leq \cdots \leq \lambda_{\mathcal{N}^{\mathrm{snap}}(T)},$$

and the eigenfunctions are normalized by $(\kappa \psi_j^T, \psi_j^T) = 1$. Then the local offline function space $V^{\mathrm{off}}(T)$ is defined as the space spanned by a number of dominant eigenfunctions ψ_j^T, which is related to $j-$th smallest eigenvalue λ_j^T. We may choose $\mathcal{L}(T)$ eigenfunctions, where $\mathcal{L}(T)$ is considerably small number

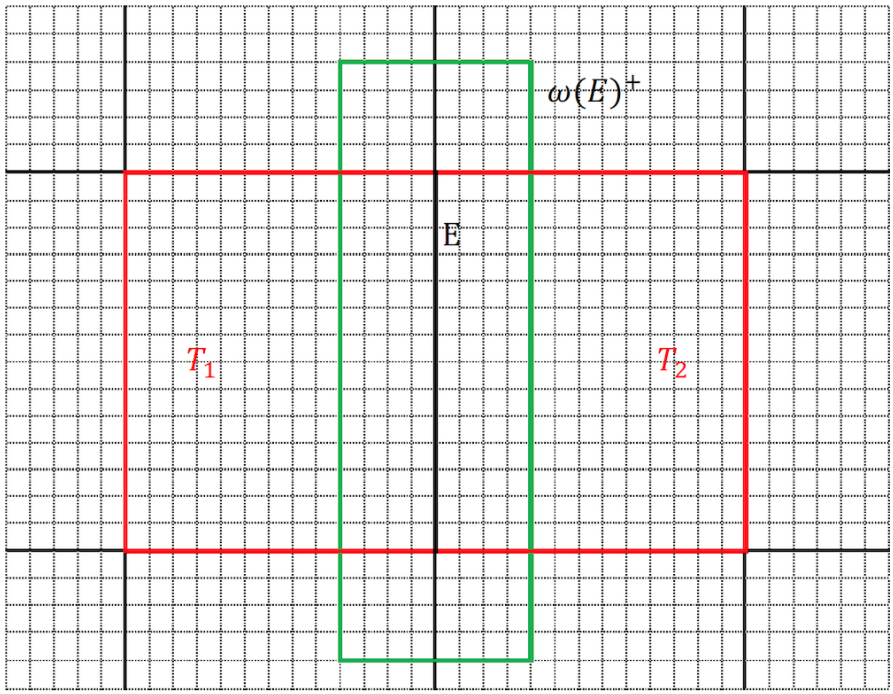


Figure 2.1. Multiscale mesh on Ω. T_1, T_2 are macro elements and $\omega(E)^+$ is an oversampled neighborhood of macro edge E.

compared to $\mathcal{N}^{\text{snap}}(T)$. In short, $V^{\text{off}}(T)$ is given by

$$V^{\text{off}}(T) = \text{Span}\left\{\psi_j^T \in V^{\text{snap}}(T) \,\middle|\, j = 1, \cdots, \mathcal{L}(T)\right\},$$

and the offine function space V^{off} is defined as

$$V^{\text{off}} = \bigcup_{T \in \mathcal{T}^H} V^{\text{off}}(T).$$

2.3.3 Moment function space $\mathcal{M}_H$

Since the offline functions are defined independently in each macro element $T \in \mathcal{T}^H$, we need to glue those functions through each macro interior edge

$E \in \mathcal{E}^{H,0}$. Moment functions play a important role here, as they are used to impose continuity between offline functions in neighboring macro elements. On each macro edge E, let $\omega(E)^+$ be an oversampled neighborhood of E. As we construct local snapshot space, the moment function $\zeta_j^E \in V_h(\omega(E)^+)$ can be obtained by solving local $\kappa-$harmonic problem

$$\begin{cases} -\nabla \cdot \left(\kappa(\mathbf{x}) \nabla \zeta_j^E\right) = 0 \text{ in } \omega(E)^+, \\ \zeta_j^E = \delta_j^E \text{ on } \partial\omega(E)^+, \end{cases} \tag{2.7}$$

where $\delta_j^E \in V_h(\omega(E)^+)$ is the function which equals to one for the $j-$th microscale mesh DOF on $\partial\omega(E)^+$ and zeros for the other DOFs. We collect the traces of ζ_j^E on E and perform a singular value decomposition to them. Denote $m(E)$ linearly independent singular vectors by s_k^E, where s_k^E is arranged in descending order with respect to its norm:

$$||s_k^E||_E^2 = \mu_k^E, \quad \mu_1^E \geq \mu_2^E \geq \cdots \geq \mu_{m(E)}^E > 0. \tag{2.8}$$

Then the local moment function space $\mathcal{M}_H(E)$ on E is given by

$$\mathcal{M}_H(E) = \text{Span}\left\{ s_k^E \,\middle|\, 1 \leq k \leq \mathcal{L}(E) \right\},$$

and the moment function space $\mathcal{M}_H$ is defined as

$$\mathcal{M}_H = \bigcup_{E \in \mathcal{E}^H} \mathcal{M}_H(E).$$

Another method for constructing moment function space

We may consider another method for constructing moment function space in order to reduce the computational cost. That is, the moment function space

can be made up of the traces of the snapshot functions. For each macro edge $E \in \mathcal{E}^H$, denote the collection of such traces by

$$\mathcal{M}_h(E) := \text{Span}\left\{\phi^T|_E \,\middle|\, \phi^T \in V^{\text{snap}}(T), E \subset \partial T\right\}.$$

We perform a singular value decomposition to $\mathcal{M}_h(E)$ and choose the first $\mathcal{L}(E)$ dominant modes of $\mathcal{M}_h(E)$, which span the local moment function space. This method makes us avoid to solve local boundary value problems (2.7).

2.3.4 Nonconforming GMsFE spaces V^H and $V^{H,0}$

The nonconforming GMsFE spaces V^H and $V^{H,0}$ are defined as

$$V^H = \left\{\psi \in V^{\text{off}} \,\middle|\, < [\psi]_E, \zeta >_E = 0,\ \forall \zeta \in \mathcal{M}_H(E),\ \forall E \in \mathcal{E}^{H,0}\right\}, \tag{2.9}$$

$$V^{H,0} = \left\{\psi \in V^{\text{off}} \,\middle|\, < [\psi]_E, \zeta >_E = 0,\ \forall \zeta \in \mathcal{M}_H(E),\ \forall E \in \mathcal{E}^{H}\right\}. \tag{2.10}$$

Since V^{off} and $\mathcal{M}_H$ are defined as the union of local function spaces, it is possible to construct V^H and $V^{H,0}$ locally. Let $E \in \mathcal{E}^{H,0}$ be a common macro edge for two macro elements T_1 and T_2 (see Figure 2.1.) Suppose that the local moment function space $\mathcal{M}_H(E)$ is constructed from $\kappa-$harmonic functions in $\omega(E) := T_1 \cup T_2$. Then the continuity condition for $\psi \in V^{\text{off}}(T_1) \cup V^{\text{off}}(T_2)$ imposed by $\mathcal{M}_H(E)$ is given as follows:

$$\begin{aligned} < [\psi]_E, \zeta >_E &= 0, \quad \forall \zeta \in \mathcal{M}_H(E), \\ < \psi, \zeta >_{E'} &= 0, \quad \forall \zeta \in \mathcal{M}_H(E'), \quad \forall E' \subset \partial\omega(E). \end{aligned} \tag{2.11}$$

Finally we define the local GMsFE space $V^H(\omega(E))$ as

$$V^H(\omega(E)) = \left\{\psi \in V^{\text{off}}(T_1) \cup V^{\text{off}}(T_2) \,\middle|\, \psi \text{ satisfies (2.11)}\right\}.$$

Then the GMsFE space $V^{H,0}$ can be obtained by

$$V^{H,0} = \bigcup_{E \in \mathcal{E}^{H,0}} V^H(\omega(E)).$$

V^H also can be constructed similarly by considering $E \in \mathcal{E}^H$ in the above argument.

Remark 2.3.1. *It is remarkable that there may exist macro bubble functions $\psi \in V^{\text{off}}(T)$ on T, which satisfy*

$$< \psi, \zeta >_{E'} = 0, \quad \forall \zeta \in \mathcal{M}_H(E'), \quad \forall E' \subset \partial T. \tag{2.12}$$

Denote the space of macro bubble functions on T_j by

$$B_H(T) = \Big\{ \psi \in V^{\text{off}}(T) \,\Big|\, \psi \text{ satisfies (2.12)} \Big\}.$$

Then the GMsFE space $V^{H,0}$ is obtained by

$$V^{H,0} = \Big(\bigcup_{T \in \mathcal{T}^H} B_H(T) \Big) \bigcup \Big(\bigcup_{E \in \mathcal{E}^{H,0}} V^H\big(\omega(E)\big) \Big).$$

Remark 2.3.2. *The dimension of GMsFEM space $V^{H,0}$ may depend on the dimension of local moment function space. For each macro element $T \in \mathcal{T}^H$, we practically take the dimension of local offline function space as*

$$\mathcal{L}(T) = \sum_{E' \subset \partial T} \mathcal{L}(E').$$

Then the dimension of $V^H\big(\omega(E)\big)$ is given by

$$\dim \Big(V^H(\omega(E)) \Big) \geq \mathcal{L}(T_1) + \mathcal{L}(T_2) - \mathcal{L}(E) - \sum_{E' \subset \partial\omega(E)} \mathcal{L}(E') = \mathcal{L}(E).$$

If there are no macro bubble functions, it follows that

$$\dim\left(V^{H,0}\right) = \sum_{E\in\mathcal{E}^{H,0}} \mathcal{L}(E).$$

2.4 Error analysis

An error estimate for nonconforming GMsFEM is provided using the broken energy norm

$$||v||_h := \sqrt{a_h(v,v)}, \quad \forall v \in H^1(\Omega) + V_h + V^H.$$

We only state the main theorem here and refer to [38] for the details.

Theorem 2.4.1. *Let u and u^H be the solutions of* (2.1) *and* (2.4)*, respectively. Suppose $\mathcal{M}_H(E)$ is constructed from $\mathcal{M}_h(E)$, which consists of the traces of snapshot functions. Then we have following error estimate*

$$\begin{aligned}||u^H - u||_h \leq CH||f||_\Omega + \Big[\sum_{T\in\mathcal{T}^H}\sum_{j=\mathcal{L}(T)+1}^{\mathcal{N}^{\text{snap}}(T)} \lambda_j^T(\alpha_j^T)^2\Big]^{\frac{1}{2}} + \inf_{q^H\in V^{H,0}}||q^H - \mathcal{P}^H\widetilde{u}_h||_h \\ + \mu_{\max}\Big[\sum_{E\in\mathcal{E}^H}\Big|\Big|\Big|\mathcal{P}_h^E(\kappa\frac{\partial p}{\partial\nu})\Big|\Big|\Big|_E^2\Big]^{\frac{1}{2}} \sup_{w^H\in V^{H,0}}\frac{|||w^H|||}{||w_h||_h},\end{aligned} \tag{2.13}$$

where $\mu_{\max} = \max\Big\{\mu_{\mathcal{L}(E)+1}^E \,\Big|\, E\in\mathcal{E}^H\Big\}$.

On the above theorem, $\widetilde{u}_h \in V^{\text{snap}}$ denotes the snapshot solution of the following problem:

$$\begin{cases} -\nabla\cdot\big(\kappa(\mathbf{x})\nabla\widetilde{u}_h\big) = 0 \text{ in } T, \\ \qquad\qquad \widetilde{u}_h = u_h \text{ on } \partial T, \end{cases} \tag{2.14}$$

where u_h is the microscale solution of (2.3). It is related to the first error

term on the right hand side of (2.13), which measures $||\widetilde{u}_h - u_h||_h$. If we take $V^{\text{snap}}(T) = V_h(T)$, this term will vanish.

For the second error term in (2.13), $\mathcal{P}^H$ denotes the projection operator from V^{snap} to V^{off}. Any function $\phi^T \in V^{\text{snap}}(T)$ can be represented by eigenfunctions ψ_j^T in (2.6):

$$\phi^T = \sum_{j=1}^{\mathcal{N}^{\text{snap}}(T)} \alpha_j^T \psi_j^T.$$

Then the (local) projection $\mathcal{P}_T^H := \mathcal{P}^H|_T : V^{\text{snap}}(T) \to V^{\text{off}}(T)$ is defined by

$$\mathcal{P}_T^H(\phi^T) = \sum_{j=1}^{\mathcal{L}(T)} \alpha_j^T \psi_j^T.$$

It is obvious that the second error term in (2.13) is caused by $\mathcal{P}^H$, while reducing the dimension of V^{snap}.

The third error term measures how well the offline space V^{off} approximates the GMsFE space $V^{H,0}$. To sum up, the first three error terms in (2.13) indicate the approximation error.

The last error term in (2.13) accounts for the consistency error. For each $E \in \mathcal{E}^H$, the projection operator $\mathcal{P}_h^E : \mathcal{M}_h(E) \to \mathcal{M}_H(E)$ is defined by

$$< \mathcal{P}_h^E \phi^E - \phi^E, s_H >_E = 0, \quad \forall s_H \in \mathcal{M}_H(E), \tag{2.15}$$

where $\phi^E \in M_h(E)$. Any function $\phi^E \in M_h(E)$ can be represented by singular vectors s_k^E in (2.8):

$$\phi^E = \sum_{k=1}^{m(E)} \beta_k^E s_k^E,$$

and the norm $|||\cdot|||_E$ on $\mathcal{M}_h(E)$ is defined as

$$|||\phi^E|||_E^2 = \sum_{k=1}^{m(E)} (\beta_k^E)^2.$$

For $w^H \in V^{H,0}$, we define $|||\cdot|||$ as

$$|||w^H||| := \sum_{T\in\mathcal{T}^H} \sum_{E\in\partial T} |||w^H|||_E.$$

2.5 Numerical results

In this section, we provide some numerical results. We use the nonparametric DSSY nonconforming quadrilateral elements to construct microscale space V_h. We take the same number, say k, of moment functions on each interior macro edge $E \in \mathcal{E}^{H,0}$, which determines the dimension of nonconforming GMsFE space $V^{H,0}$ by Remark 2.3.2. We may adopt the oversampling technique for constructing local snapshot function space. For practical use, we may take oversampled region T^+ as the extension of T by $\delta(T)$ layers of micro elements surrounding T. An oversampled neighborhood $\omega(E)^+$ of macro edge E is also taken by extension of $\delta(E)$ microscale layers surrounding E, and used to construct local moment function space.

Example 2.5.1. *Consider the following elliptic problem:*

$$\begin{cases} -\nabla\cdot\big(\kappa(\mathbf{x})\nabla u\big) = f \ \text{ in } \Omega, \\ \qquad\qquad\qquad u = 0 \ \text{ on } \partial\Omega, \end{cases} \tag{2.16}$$

where $\Omega = (0,1)^2$ *and* $\kappa(\mathbf{x}) = 1 + (1+x_1)(1+x_2) + \epsilon\sin(10\pi x_1)\sin(5\pi x_2)$.

The source term f is generated by the exact solution

$$u(x_1, x_2) = \sin(3\pi x_1)x_2(1 - x_2) + \epsilon \sin(\pi x_1/\epsilon) \sin(\pi x_2/\epsilon).$$

We examine error behaviors using uniform trapezoidal meshes with parameter θ. $H \times H$ macro elements and $h \times h$ micro elements are considered respectively. We adopt oversampling technique to construct snapshot and moment function spaces where $\delta(T) = \delta(E) = 1$. Relative energy errors and L^2 errors are reported for various θ and ϵ.

$\frac{1}{H}$	$\frac{1}{h}$	$\dim(V^{H,0})$	$\theta = 0$		$\theta = 0.5$	
			Rel. Energy	Rel. L^2	Rel. Energy	Rel. L^2
5	50	400	0.884	0.388	0.885	0.386
10	100	1800	0.871	0.363	0.871	0.362
20	200	7600	0.346	0.676E-01	0.347	0.674E-01
40	400	31200	0.181	0.181E-01	0.181	0.180E-01

Table 2.1. Convergence for $\epsilon = 0.1$.

$\frac{1}{H}$	$\frac{1}{h}$	$\dim(V^{H,0})$	$\theta = 0$		$\theta = 0.5$	
			Rel. Energy	Rel. L^2	Rel. Energy	Rel. L^2
5	50	400	0.885	0.625	0.886	0.623
10	100	1800	0.355	0.118	0.356	0.118
20	200	7600	0.186	0.316E-01	0.186	0.314E-01
40	400	31200	0.940E-01	0.803E-02	0.942E-01	0.799E-02

Table 2.2. Convergence for $\epsilon = 0.2$.

$\frac{1}{H}$	$\frac{1}{h}$	$\dim(V^{H,0})$	$\theta = 0$		$\theta = 0.5$	
			Rel. Energy	Rel. L^2	Rel. Energy	Rel. L^2
5	50	400	0.335	0.130	0.336	0.130
10	100	1800	0.173	0.342E-01	0.174	0.340E-01
20	200	7600	0.884E-01	0.879E-02	0.888E-01	0.875E-02
40	400	31200	0.444E-01	0.221E-02	0.445E-01	0.220E-02

Table 2.3. Convergence for $\epsilon = 0.5$.

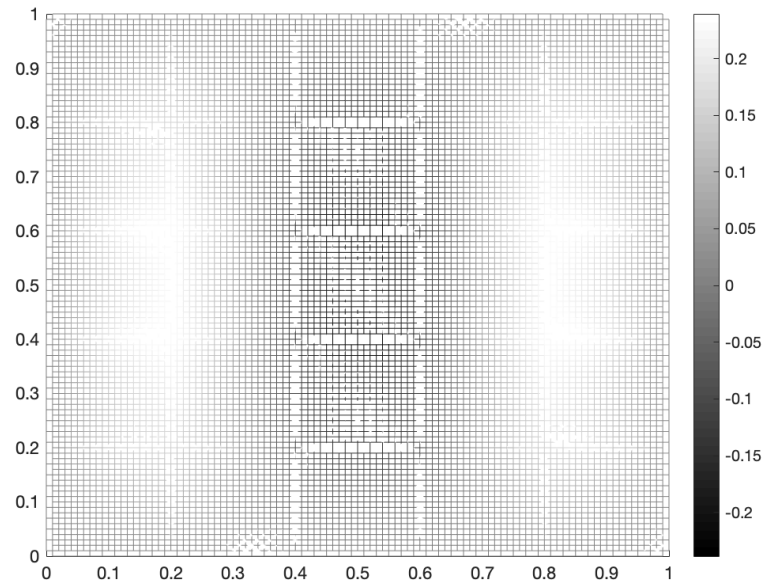


Figure 2.2. Multiscale solution of $\epsilon = 0.2, \theta = 0$ when $1/H = 5, 1/h = 50$.

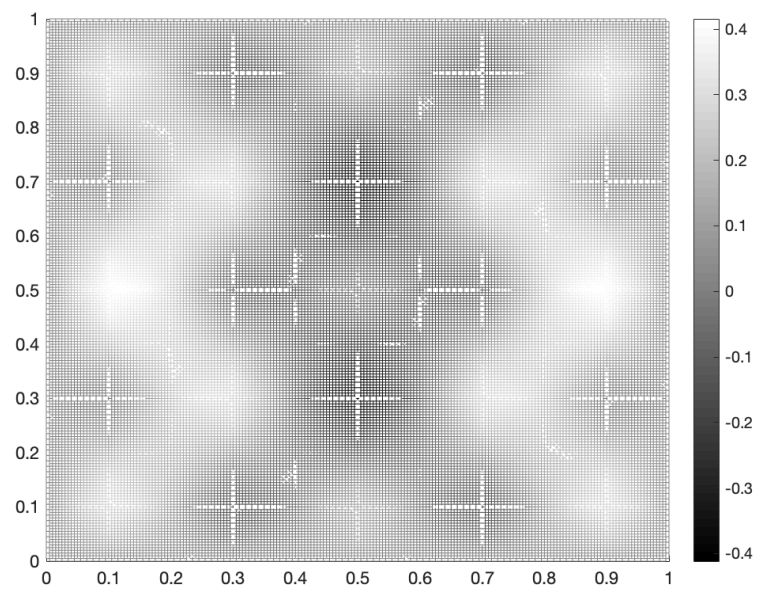


Figure 2.3. Multiscale solution of $\epsilon = 0.2, \theta = 0$ when $1/H = 10, 1/h = 100$.

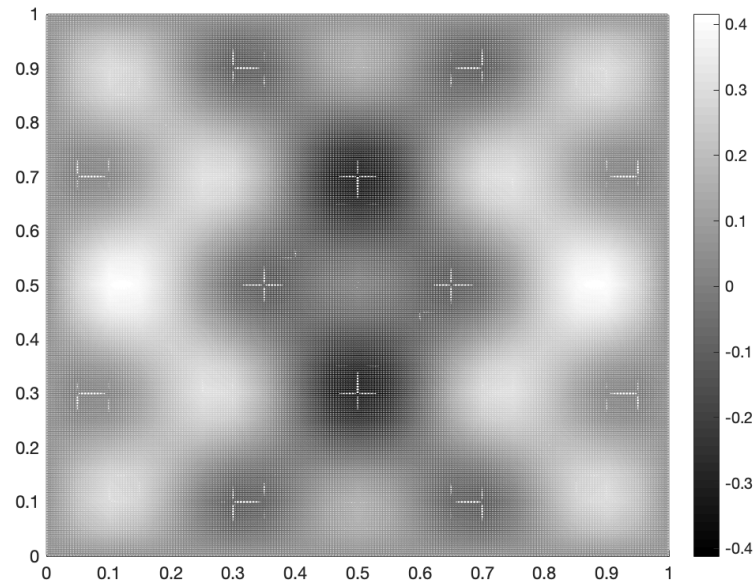


Figure 2.4. Multiscale solution of $\epsilon = 0.2, \theta = 0$ when $1/H = 20, 1/h = 200$.

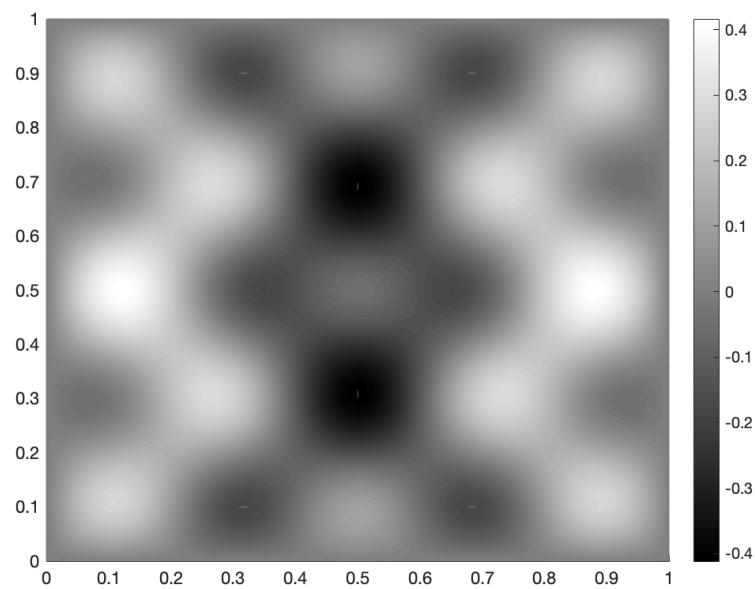


Figure 2.5. Multiscale solution of $\epsilon = 0.2, \theta = 0$ when $1/H = 40, 1/h = 400$.

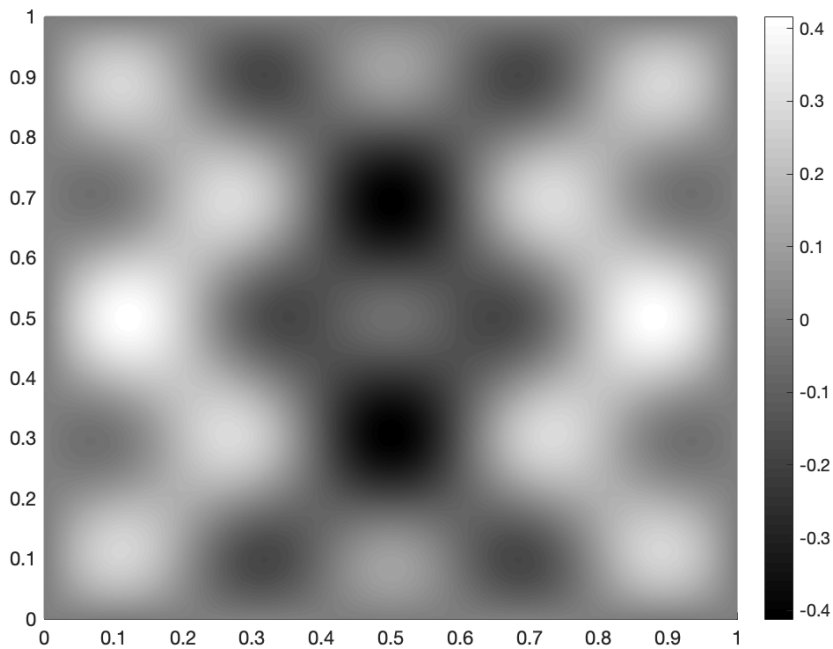


Figure 2.6. Microscale reference solution of $\epsilon = 0.2, \theta = 0$ when $1/h = 400$.

Chapter 3

Algebraic Multiscale Method

3.1 Introduction

In this chapter, we study an AMS (Algebraic MultiScale) finite element method. As a model problem, we consider

$$-\nabla \cdot (\kappa(x)\nabla u) = f \text{ in } \Omega, \tag{3.1}$$

where κ is a heterogeneous coefficient and $f \in H^{-1}(\Omega)$. Assume that a finite element method is used to approximate (3.1) based on a microscale mesh to get a corresponding linear system, say $A_h\eta_h = b_h$. But we assume that the coefficient κ and the source term f are NOT available, although the components of A_h and b_h are available. Then the question is "By using the algebraic information on $A_h\eta_h = b_h$ only, is it possible to provide numerical approximate solutions which contain similar nature and properties of those obtained by the usual multiscale methods?"

Motivated by the AMG (Algebraic MultiGrid) method, we try to build macroscale linear systems, say $A^H\eta^H = b^H$, using "the only algebraic information" obtained from the microscale linear system, *i.e.,* the information on the components of A_h and b_h. The procedure of constructing macroscale linear systems uses the details of building snapshot spaces and offline spaces. We will show that this process can be performed using only the algebraic information on the microscale linear system. We first consider the one-dimensional case in very detail, and present two dimensional case using the DSSY nonconforming finite element space.

We remark that our approach is completely different from that of AMG. The nature of AMG is to solve the original linear system $A_h\eta_h = b_h$ iteratively, but accurately, by using the classical geometric multigrid idea. The spirit of the "AMS (algebraic multiscale method)" is to construct an algebraic system $A^H\eta^H = b^H$ with significant dimension reduction such that the solutions η^H are rough, but reasonable approximation to η_h, which follow the line of thoughts in multiscale methods.

This chapter is organized as follows. In Section 2, we state our model problem and present the multiscale FEM to construct multiscale basis functions. The algebraic multiscale method is then introduced in Section 3. An algebraic formulation of a macroscale linear system $A^H\eta^H = b^H$ is built from the knowledge of components of the microscale linear system $A_h\eta_h = b_h$. Section 4 is devoted to the energy norm error estimate of the proposed method. In Section 5, we present some numerical results. Two-dimensional case is investigated in Section 6.

3.2 Preliminaries

In this section we briefly review the multiscale finite element method in one dimension. For any open interval I, denote by $H^k(I) = W^{k,2}(I)$ the standard Sobolev space equipped with inner product $(\cdot,\cdot)_{k,I}$ and norm $||\cdot||_{k,I}$. For $k = 0$, we abbreviate $(\cdot,\cdot)_{k,I}$ and $||\cdot||_{k,I}$ as $(\cdot,\cdot)_I$ and $||\cdot||_I$, respectively. Consider the following elliptic problem:

$$\begin{cases} -\dfrac{d}{dx}\left(\kappa(x)\dfrac{du}{dx}\right) = f \text{ in } \Omega, \\ \qquad\qquad u = 0 \text{ on } \partial\Omega, \end{cases} \tag{3.2}$$

where $\Omega = (0,1)$, κ is a highly heterogeneous coefficient and $f \in H^{-1}(\Omega)$. Denote by $\mathcal{T}^H = \bigcup_{K=1}^{N^H}\{I^K\}$ and $\mathcal{T}_h = \bigcup_{j=1}^{N_h}\{I_j\}$ two families of macroscale and microscale triangulations of Ω into macroscale and microscale subintervals such that $I^K = [X^{K-1}, X^K]$ and $I_j = [x_{j-1}, x_j]$, where $0 = X^0 < X^1 < \cdots < X^{N^H} = 1$ and $0 = x_0 < x_1 < \cdots < x_{N_h} = 1$. Here, and in what follows, H and h stand for the macroscale and microscale mesh parameters given by

$$H = \max_{K=1,\cdots,N^H}(X^K - X^{K-1}), \quad h = \max_{j=1,\cdots,N_h}(x_j - x_{j-1}).$$

Let V_h be a finite element basis function space associated with $\mathcal{T}_h$. One may solve (3.2) on the microscale mesh $\mathcal{T}_h$ using a finite element method to find $u_h \in V_h$ such that

$$a_h(u_h, v_h) = (f, v_h)_\Omega \quad \forall v_h \in V_h, \tag{3.3}$$

where $a_h(u_h, v_h) = \sum_{j=1}^{N_h}(\kappa u_h', v_h')_{I_j}$. In a multiscale method, we solve κ–harmonic problems in each macroscale interval I^K (see (3.4) and (3.5).) Using these solutions, we construct the multiscale basis function space $V^H =$

$\{\Psi^K\}_{K=1,\ldots,N^H-1}$ and find the multiscale solution in V^H.

To construct V^H, we assume that $\mathcal{T}_h$ is a refinement of $\mathcal{T}^H$ satisfying $0 < h \ll H < 1$. For $K = 1, \cdots, N^H$, denote by H^K the size of K-th macro interval $I^K = (X^{K-1}, X^K)$. Let $\{x_j^K\}_{j=0}^{N_h^K}$ be the set of micro nodes for I^K and designate by I_j^K the j–th subinterval (x_{j-1}^K, x_j^K) with length h_j^K for $j = 1, \cdots, N_h^K$ such that $x_0^K = X^{K-1}$ and $x_{N_h^K}^K = X^K$. For each K, let $V_h(I^K) = \{\phi_j^K\}_{j=0,\cdots,N_h^K}$ be the space of standard basis functions for the C^0–piecewise linear finite element space on I^K.

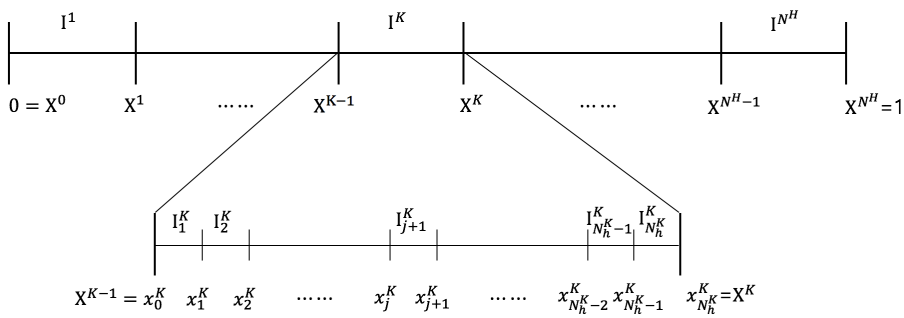


Figure 3.1. Multiscale mesh on Ω.

Denote by $\Psi_\pm^K$ the multiscale basis functions in interval I^K, which can be obtained as the solutions of $\kappa-$harmonic problems

$$\begin{cases} -\dfrac{d}{dx}\big(\kappa\dfrac{d}{dx}\Psi_-^K\big) = 0 \text{ in } I^K, \\ \Psi_-^K(X^{K-1}) = 1,\ \Psi_-^K(X^K) = 0, \end{cases} \tag{3.4}$$

and

$$\begin{cases} -\dfrac{d}{dx}\big(\kappa\dfrac{d}{dx}\Psi_+^K\big) = 0 \text{ in } I^K, \\ \Psi_+^K(X^{K-1}) = 0,\ \Psi_+^K(X^K) = 1. \end{cases} \tag{3.5}$$

Let us seek $\Psi^K_{h,\pm} \in V_h(I^K)$ which approximate $\Psi^K_{\pm}$ in the form

$$\Psi^K_{h,-} = \sum_{j=1}^{N^K_h - 1} \eta^K_{j,-}\phi^K_j + \phi^K_0, \tag{3.6a}$$

$$\Psi^K_{h,+} = \sum_{j=1}^{N^K_h - 1} \eta^K_{j,+}\phi^K_j + \phi^K_{N^K_h}. \tag{3.6b}$$

Since $\Psi^K_{h,\pm}$ are piecewise-linear in I^K, one may set

$$\frac{d}{dx}\Psi^K_{h,\pm} = \gamma^K_{j,\pm} \text{ in } I^K_j \quad \text{for some constant } \gamma^K_{j,\pm},\ j = 1, \cdots, N^K_h. \tag{3.7}$$

Proposition 3.2.1. *Assume that* $\frac{1}{\kappa(x)} \in L^1(a,b)$. *Consider the following differential equation:*

$$\begin{cases} -\dfrac{d}{dx}\big(\kappa \dfrac{dw}{dx}\big) = 0, \quad \text{in } I = (a,b), \\ w(a) = 0,\ w(b) = 1. \end{cases} \tag{3.8}$$

Denote by $\frac{1}{\kappa(x)} = \kappa^{-1}(x)$. *The exact solution is given by*

$$w(x) = \beta \int_a^x \kappa^{-1}(s)\, ds \quad a.e.\, x \in I,$$

where $|I|\beta$ *is the harmonic mean of* $\kappa(x)$ *over the interval* I, *i.e.,*

$$|I|\beta = \left(\frac{1}{|I|}\int_I \kappa^{-1}(s)\, ds\right)^{-1}. \tag{3.9}$$

Proof. We have

$$\kappa(x)\frac{dw}{dx}(x) = c, \quad a.e.\, x \in (a,b)$$

for some constant c. From this, we get

$$w(x) = c \int_0^x \kappa^{-1}(s)\, ds \quad a.e.\, x \in (a, b).$$

With the boundary condition $w(b) = 1$, we see that

$$1 = w(b) = c \int_a^b \kappa^{-1}(s)\, ds.$$

Hence, $c = \beta$ and the proposition follows. □

Thanks to Proposition 3.2.1, it is easy to see that

$$\Psi_-^K(x) = \frac{\int_x^{X^K} \kappa^{-1}(s)\, ds}{\int_{X^{K-1}}^{X^K} \kappa^{-1}(s)\, ds} \text{ for all } x \in I^K, \tag{3.10a}$$

$$\Psi_+^K(x) = \frac{\int_{X^{K-1}}^{x} \kappa^{-1}(s)\, ds}{\int_{X^{K-1}}^{X^K} \kappa^{-1}(s)\, ds} \text{ for all } x \in I^K. \tag{3.10b}$$

Denote an n dimensional vector with parameters K and $\pm$ as follows:

$$\boldsymbol{\alpha}_\pm^K = (\alpha_{1,\pm}^K, \cdots, \alpha_{n,\pm}^K)^t \in \mathbb{R}^n.$$

Recalling (3.6), (3.7), and (3.10), and utilizing the principle of energy norm

minimization of the finite element method, we deduce the following equalities:

$$\begin{aligned}
&\min_{\boldsymbol{\eta}_+^K \in \mathbb{R}^{N_h^K}} \left\{ \int_{I^K} \kappa(x) \left[\frac{d}{dx} \left(\Psi_+^K(x) - \Psi_{h,+}^K(x) \right) \right]^2 \, dx \right\}^{1/2} \\
&= \min_{\boldsymbol{\eta}_+^K \in \mathbb{R}^{N_h^K}} \sum_{j=1}^{N_h^K} \left\{ \int_{I_j^K} \kappa(x) \left[\frac{d}{dx} \left(\Psi_+^K(x) - \Psi_{h,+}^K(x) \right) \right]^2 \, dx \right\}^{1/2} \\
&= \min_{\boldsymbol{\gamma}_+^K} \sum_{j=1}^{N_h^K} \left\{ \int_{I_j^K} \kappa(x) \left(\frac{\beta}{\kappa(x)} - \gamma_{j,+}^K \right)^2 \, dx \right\}^{1/2} \\
&= \min_{\gamma_{j,+}^K} \sum_{j=1}^{N_h^K} \left\{ \int_{I_j^K} \left(\frac{\beta^2}{\kappa(x)} - 2\beta\gamma_{j,+}^K + \kappa(x)(\gamma_{j,+}^K)^2 \right) \, dx \right\}^{1/2}.
\end{aligned}$$

After differentiating the above with respect to $\gamma_{j,+}^K$, we have

$$\gamma_{j,+}^K = \left(\int_{x_{j-1}^K}^{x_j^K} \beta \, dx \right) \left(\int_{x_{j-1}^K}^{x_j^K} \kappa(x) \, dx \right)^{-1} = \beta h_j^K \left(\int_{x_{j-1}^K}^{x_j^K} \kappa(x) \, dx \right)^{-1}.$$

If $\kappa(x) = 1$, $\gamma_{j,\pm}^K = \pm 1$ for $j = 1, \cdots, N_h^K$ to make $\Psi_{h,\pm}^K$ linear. Since $\Psi_{h,+}^K + \Psi_{h,-}^K = 1$ in (3.10), both $\Psi_{h,\pm}^K$ can be computed from $\gamma_{j,+}^K$ values. Now the multiscale basis function $\Psi^K \in V^H$ is constructed from $\Psi_{h,\pm}^K$:

$$\Psi^K = \begin{cases} \Psi_{h,+}^K \text{ in } I^K, \\ \Psi_{h,-}^{K+1} \text{ in } I^{K+1}, \\ 0, \quad \text{otherwise.} \end{cases}$$

Remark 3.2.2. *We may reduce the dimension of multiscale basis function space by applying a suitable dimension reduction technique. For example, we adopt a spectral decomposition method and take a reasonably small number of dominant eigenfunctions as basis functions. We do not perform such process for one-dimensional case.*

3.3 Algebraic Multiscale Method

In this section, we present the procedure of algebraic multiscale method. We assume that we are given all the information on the components of microscale linear system

$$A_h \eta_h = b_h, \tag{3.11}$$

which is obtained from (3.3). However, we assume that κ and f are not known.

We want to find $u^H \in V^H$ satisfying

$$a^H(u^H, v^H) = (f, v^H)_\Omega \quad \forall v^H \in V^H, \tag{3.12}$$

where $a^H(u^H, v^H) = \sum_{K=1}^{N^H-1} (\kappa (u^H)', (v^H)')_{I^K}$. We will construct the corresponding macroscale linear system

$$A^H \eta^H = b^H \tag{3.13}$$

using only the algebraic structure of microscale linear system (3.11).

Throughout the paper, we assume that A_h and b_h are assembled by using the standard C^0 piecewise linear element on $\Omega = (0, 1)$. Also assume that the microscale mesh is sufficiently refined so that the heterogeneous coefficient κ is constant in each micro interval $I_j = (x_{j-1}, x_j)$. Let $\kappa = \kappa_j$ in I_j and denote the size of I_j by h_j.

3.3.1 Algebraic formulation of stiffness matrix

We need to look at the procedure of building the components of stiffness matrices in microscale and macroscale.

Construction of A_h

Let $V_h = \{\phi_j\}_{j=0,\cdots,N_h}$ be the set of basis functions on the microscale mesh. For $1 \le j \le N_h - 1$, we have

$$\frac{d\phi_j}{dx}\Big|_{I_j} = \frac{1}{h_j} \quad \text{and} \quad \frac{d\phi_j}{dx}\Big|_{I_{j+1}} = -\frac{1}{h_{j+1}}. \tag{3.14}$$

The diagonals of A_h are supposed to be expressed as follows:

$$\begin{aligned}
[A_h]_{j,j} &= \int_{x_{j-1}}^{x_{j+1}} \kappa(x) \frac{d\phi_j}{dx} \frac{d\phi_j}{dx}\, dx \\
&= \int_{x_{j-1}}^{x_j} \kappa_j \left(\frac{1}{h_j}\right)^2 dx + \int_{x_j}^{x_{j+1}} \kappa_{j+1} \left(-\frac{1}{h_{j+1}}\right)^2 dx \\
&= \frac{\kappa_j}{h_j} + \frac{\kappa_{j+1}}{h_{j+1}}.
\end{aligned} \tag{3.15}$$

The off-diagonals are given by

$$[A_h]_{j,j-1} = -\frac{\kappa_j}{h_j}, \quad [A_h]_{j,j+1} = -\frac{\kappa_{j+1}}{h_{j+1}}. \tag{3.16}$$

There is an one-to-one correspondence between the off-diagonal element of A_h and the average of coefficient in each micro interval.

Remark 3.3.1. *Instead of assuming κ to be a piecewise constant function, we may adopt a quadrature rule for integration. For instance, we get*

$$[A_h]_{j,j-1} \approx -\frac{1}{h_j}\kappa_{j-1/2}, \quad [A_h]_{j,j+1} \approx -\frac{1}{h_{j+1}}\kappa_{j+1/2}$$

by using the mid-point rule.

Construction of A^H

We present the details of construction of the components of A^H from the knowledge of A_h. This procedure requires the construction of the space of macroscale basis functions $V^H = \{\Psi^K, K = 1, ..., N^H - 1\}$ on $\Omega = (0, 1)$. We state one of the two main results in the following theorem.

Theorem 3.3.2. *All the components of the macroscale stiffness matrix A^H in* (3.13) *are constructed using the components of the microscale stiffness matrix A_h in* (3.11) *only.*

Proof. For each K, by β^K denote the β value defined by (3.9) in the macro interval $I^K = (X^{K-1}, X^K)$. Then from (3.4), (3.5), and Proposition 3.2.1 it follows that

$$\kappa \frac{d}{dx}\Psi^K = \begin{cases} \beta^K & \text{on } I^K, \\ -\beta^{K+1} & \text{on } I^{K+1}. \end{cases}$$

Hence, for $1 \leq K \leq N^H - 1$, the diagonal elements of A^H are given as follows:

$$\begin{aligned} \left[A^H\right]_{K,K} &= \int_{X^{K-1}}^{X^{K+1}} \kappa(x) \frac{d\Psi^K}{dx} \frac{d\Psi^K}{dx}\, dx \\ &= \beta^K \int_{X^{K-1}}^{X^K} \frac{d\Psi^K}{dx}\, dx + \beta^{K+1} \int_{X^K}^{X^{K+1}} \frac{d\Psi^K}{dx}\, dx \\ &= \beta^K + \beta^{K+1}. \end{aligned} \tag{3.17}$$

The off-diagonals can be computed similarly:

$$\left[A^H\right]_{K,K-1} = -\beta^K, \quad \left[A^H\right]_{K,K+1} = -\beta^{K+1}. \tag{3.18}$$

Denote by M_K the total number of micro nodes x_j^L's on $[0, X^K) = \overline{I^1} \cup \overline{I^2} \cup \cdots \cup \overline{I^{K-1}} \cup I^K$. Then $I_j^K = (x_{j-1}^K, x_j^K)$ is same as the $(M_{K-1} + j)-$th micro

interval $I_{M_{K-1}+j} \in \mathcal{T}_h$. Using the relation (3.16), β^K is computed by

$$\begin{aligned}\beta^K &= \left(\int_{I^K} \kappa^{-1}(x)\, dx\right)^{-1} = \left(\sum_{j=1}^{N_h^K} \int_{I_j^K} \kappa^{-1}(x)\, dx\right)^{-1} \\ &= \left(\sum_{j=1}^{N_h^K} \int_{I_{M_{K-1}+j}} \kappa^{-1}(x)\, dx\right)^{-1} = \left(\sum_{j=1}^{N_h^K} \frac{h_{M_{K-1}+j}}{\kappa_{M_{K-1}+j}}\right)^{-1} \\ &= -\left(\sum_{j=1}^{N_h^K} \left(1/\left[A_h\right]_{M_{K-1}+j, M_{K-1}+j-1}\right)\right)^{-1}.\end{aligned} \tag{3.19}$$

Plugging (3.19) in (3.17) and (3.18), we see that all the components of the macroscale stiffness matrix are constructed using the components of the microscale stiffness matrix only. This completes the proof. □

Construction of b^H

We present the details of construction of the components of b^H from the knowledges of A_h and b_h. We use the constructed macroscale basis function space $V^H = \{\Psi^K, K = 1, \cdots, N^H - 1\}$ on $\Omega = (0, 1)$.

Lemma 3.3.3. *The coefficients of $\Psi_{h,-}^K$, $K = 2, \cdots, N^H$ and the coefficients of $\Psi_{h,+}^K$, $K = 1, \cdots, N^H - 1$ in* (3.6) *are computed using the components of the microscale stiffness matrix A_h in* (3.11) *only.*

Proof. The approximate weak problem of (3.5) is to find $\Psi_{h,+}^K \in V_h(I^K)$ such that:

$$\int_{I^K} \kappa(x) \left(\sum_{j'=1}^{N_h^K - 1} \eta_{j',+}^K \frac{d\phi_{j'}^K}{dx}\right) \frac{d\phi_j^K}{dx}\, dx = -\int_{I^K} \kappa(x) \frac{d\phi_{N_h^K}^K}{dx} \frac{d\phi_j^K}{dx}\, dx \tag{3.20}$$

for $j = 1, \cdots, N_h^K - 1$. Since $\phi_j^K = \phi_{M_{K-1}+j}$, from (3.20) we can obtain the

following linear system

$$\begin{bmatrix} d_1^K & u_1^K & & & \\ l_2^K & d_2^K & u_2^K & & \\ & l_3 & \ddots & \ddots & \\ & & \ddots & \ddots & u^K_{N_h^K-2} \\ & & & l^K_{N_h^K-1} & d^K_{N_h^K-1} \end{bmatrix} \begin{bmatrix} \eta^K_{1,+} \\ \eta^K_{2,+} \\ \vdots \\ \vdots \\ \eta^K_{N_h^K-1,+} \end{bmatrix} = \begin{bmatrix} b^K_{1,+} \\ b^K_{2,+} \\ \vdots \\ \vdots \\ b^K_{N_h^K-1,+} \end{bmatrix} \tag{3.21}$$

where

$$\begin{aligned} d_j^K &= [A_h]_{M_{K-1}+j, M_{K-1}+j}\,,\ j = 1, \cdots, N_h^K - 1, \\ u_j^K &= [A_h]_{M_{K-1}+j, M_{K-1}+j+1}\,,\ j = 1, \cdots, N_h^K - 2, \\ l_j^K &= [A_h]_{M_{K-1}+j-1, M_{K-1}+j}\,,\ j = 2, \cdots, N_h^K - 1, \\ b^K_{j,+} &= \begin{cases} 0, \text{ if } j = 1, \cdots, N_h^K - 2, \\ [A_h]_{M_{K-1}+N_h^K-1, M_{K-1}+N_h^K}\,, \text{ if } j = N_h^K - 1. \end{cases} \end{aligned}$$

Note that $b^K_{j,+}$ can be obtained by using the components of the microscale stiffness matrix only, if $K = 1, \cdots, N^H - 1$. Since the values of d_j^K, u_j^K, l_j^K, and $b^K_{j,+}$ come from the microscale stiffness matrix A_h and (3.21) is a tridiagonal system of equations, $\eta^K_{j,+}$ can be computed by using the components of A_h and Thomas' algorithm. $\eta^K_{j,-}$ for $K = 2, \cdots, N^H$ are obtained similarly. We thus see that the coefficients of $\Psi^K_{h,-}$, $K = 2, \cdots, N^H$ and the coefficients of $\Psi^K_{h,+}$, $K = 1, \cdots, N^H - 1$ in (3.6) are computed by using the components of the microscale stiffness matrix only. This completes the proof. □

Now we state the other of the two main results in the following theorem.

Theorem 3.3.4. *All the components of the macroscale right hand side b^H in* (3.13) *are constructed using the components of the microscale stiffness matrix*

A_h in (3.11) *and the components of the microscale right hand side b_h in* (3.11) *only.*

Proof. First we express the multiscale basis function Ψ^K by the microscale basis functions $\{\phi_j\}_{j=1,\dots,N_h}$ using the form of (3.6). For $1 \le K \le N^H - 1$,

$$\begin{aligned}
\Psi^K &= \Psi^K_{h,+} + \Psi^{K+1}_{h,-} \\
&= \left(\sum_{j=1}^{N_h^K - 1} \eta^K_{j,+} \phi^K_j + \phi^K_{N_h^K} \right) + \left(\phi^{K+1}_0 + \sum_{j=1}^{N_h^{K+1}-1} \eta^{K+1}_{j,-} \phi^{K+1}_j \right) \\
&= \sum_{j=1}^{N_h^K - 1} \eta^K_{j,+} \phi_{M_{K-1}+j} + (\phi^K_{N_h^K} + \phi^{K+1}_0) + \sum_{j=1}^{N_h^{K+1}-1} \eta^{K+1}_{j,-} \phi_{M_K+j} \\
&= \sum_{j=1}^{N_h^K - 1} \eta^K_{j,+} \phi_{M_{K-1}+j} + \phi_{M_K} + \sum_{j=1}^{N_h^{K+1}-1} \eta^{K+1}_{j,-} \phi_{M_K+j}.
\end{aligned}$$

Thus the macroscale right hand side vector b^H is computed by

$$\begin{aligned}
[b^H]_K &= \int_{X^{K-1}}^{X^{K+1}} f \Psi^K \, dx \\
&= \int_{X^{K-1}}^{X^{K+1}} f \left(\sum_{j=1}^{N_h^K - 1} \eta^K_{j,+} \phi_{M_{K-1}+j} + \phi_{M_K} + \sum_{j=1}^{N_h^{K+1}-1} \eta^{K+1}_{j,-} \phi_{M_K+j} \right) dx \\
&= \sum_{j=1}^{N_h^K - 1} \eta^K_{j,+} \int_{x_{M_{K-1+j-1}}}^{x_{M_{K-1+j+1}}} f \phi_{M_{K-1}+j} \, dx + \int_{x_{M_{K-1}}}^{x_{M_{K+1}}} f \phi_{M_K} \, dx \\
&\quad + \sum_{j=1}^{N_h^{K+1}-1} \eta^{K+1}_{j,-} \int_{x_{M_{K+j-1}}}^{x_{M_{K+j+1}}} f \phi_{M_K+j} \, dx, \\
&= \sum_{j=1}^{N_h^K - 1} \eta^K_{j,+} [b_h]_{M_{K-1}+j} + [b_h]_{M_K} + \sum_{j=1}^{N_h^{K+1}-1} \eta^{K+1}_{j,-} [b_h]_{M_K+j}, \qquad (3.22)
\end{aligned}$$

where we used the fact that the right hand side vector b_h is given by

$$[b_h]_j = \int_{x_{j-1}}^{x_{j+1}} f\phi_j \, dx, \quad j = 1, ..., N_h - 1.$$

By Lemma 3.3.3 $\eta_{j,+}^K$ and $\eta_{j,-}^{K+1}$ in (3.22) are obtained by using the components of the microscale stiffness matrix A_h. We thus see that all the components of the macroscale right hand side vector are constructed by using the components of the microscale stiffness matrix and the components of the microscale right hand side vector only. This completes the proof. □

3.3.2 Multiscale solution

Now we obtain the macroscale matrix system using the algebraic structure of microscale linear systems. We solve $A^H \eta^H = b^H$ to get the multiscale solution

$$u^H = \sum_{K=1}^{N^H - 1} \eta^K \Psi^K.$$

Since the exact form of Ψ^K is known from (3.10), we may compute the value of u^H at every micro node.

3.4 Error analysis

In this section, we derive an error estimate for the algebraic multiscale method. For every $v \in H^1(\Omega)$, define the energy norm as

$$|||v|||_h := \sqrt{\sum_{I_j \in \mathcal{T}_h} \int_{I_j} \kappa (v')^2 dx} = ||\kappa^{\frac{1}{2}} v'||_h.$$

We will estimate $|||u_{ms}^H - u_h|||_h$ using microscale solution u_h as a reference solution.

Lemma 3.4.1. *[38] Let u_h be the microscale solution of* (3.3)*. Consider $\widetilde{u} \in V_h$ such that for each $I^K \in \mathcal{T}^H$, $\widetilde{u}|_{I^K}$ satisfies*

$$a_h^K(\widetilde{u}, v_h) := \sum_{I_j \in \mathcal{T}_h(I^K)} \int_{I_j} \kappa \widetilde{u}' v_h' = 0 \text{ in } I^K, \quad \forall v_h \in V_h(I^K)$$

$$\widetilde{u} = u^h \text{ on } \partial I^K.$$

Then

$$|||u_h - \widetilde{u}|||_h \leq CH \min\{\kappa^{-\frac{1}{2}}\}||f||_\Omega. \tag{3.23}$$

Proof. Observe that $\widetilde{u} - u^h|_{I^K} \in V_h \subset H_0^1(I^K)$ and $\widetilde{u} \in V^H$ by formulation of $\widetilde{u}$. By Subtracting two equations

$$\begin{aligned} a_h^K(\widetilde{u}, \widetilde{u} - u_h) &= 0, \\ a_h^K(u_h, \widetilde{u} - u_h) &= (f, \widetilde{u} - u_h)_{I^K}, \end{aligned}$$

we obtain

$$\begin{aligned} ||\kappa^{\frac{1}{2}}(\widetilde{u} - u_h)||_h^2 &= a^K(\widetilde{u} - u_h, \widetilde{u} - u_h) \\ &= \sum_{I_j \in \mathcal{T}_h(I^K)} \int_{I_j} \kappa\{(\widetilde{u} - u_h)'\}^2 dx \\ &= -\int_{I^K} f(\widetilde{u} - u_h) dx \\ &= -\int_{I^K} (\kappa^{-\frac{1}{2}} f)(\kappa^{\frac{1}{2}}(\widetilde{u} - u_h)) dx \\ &\leq ||\kappa^{-\frac{1}{2}} f||_{I^K} ||\kappa^{\frac{1}{2}}(\widetilde{u} - u_h)||_{I^K} \\ &\leq CH||\kappa^{-\frac{1}{2}} f||_{I^K} ||\kappa^{\frac{1}{2}}(\widetilde{u} - u_h)')||_{I^K}. \end{aligned}$$

The Poincarè inequality was used in the last inequality. Summation of the above local estimate over all $I^K \in \mathcal{T}^H$ yields (3.23). □

Theorem 3.4.2. *Let u_h and u^H be the microscale and macroscale solution of* (3.3) *and* (3.12) *respectively. Then*

$$|||u^H - u_h|||_h \leq CH \min\{\kappa^{-\frac{1}{2}}\}||f||_\Omega.$$

Proof. By subtracting two equation with $v_h = v^H$,

$$a^H(u^H, v^H) = (f, v^H)_\Omega, \forall v^H \in V^H,$$
$$a_h(u_h, v_h) = (f, v_h)_\Omega, \forall v_h \in V_h,$$

we get $a^H(u^H - u_h, v^H) = 0, \forall v^H \in V^H$. Then for arbitrary $v \in V^H$,

$$\begin{aligned} ||\kappa^{\frac{1}{2}}(u^H - u_h)'||_0^2 &= a^H(u^H - u_h, u^H - u_h) \\ &= a^H(u^H - u_h, (u^H - v) + (v - u_h)) \\ &= a^H(u^H - u_h, v - u_h) \\ &\leq ||\kappa^{\frac{1}{2}}(u^H - u_h)'||_0 \, ||\kappa^{\frac{1}{2}}(v - u_h)'||_0. \end{aligned}$$

The theorem follows by taking $v = \widetilde{u} \in V^H$ in Lemma 3.4.1. □

Remark that the energy norm can be computed by the algebraic information on microscale systems. Since u^H and u_h are piecewise linear functions, for all $1 \leq j \leq N_h$

$$\begin{aligned} \frac{du^H}{dx}\Big|_{(x_{j-1},x_j)} &= \frac{u^H(x_j) - u^H(x_{j-1})}{h_j} = \frac{\Delta u^H(j)}{h_j}, \\ \frac{du_h}{dx}\Big|_{(x_{j-1},x_j)} &= \frac{u_h(x_j) - u_h(x_{j-1})}{h_j} = \frac{\Delta u_h(j)}{h_j}. \end{aligned}$$

Then the energy norm of u^H is computed by

$$\begin{aligned}
|||u^H|||_h &= \left\{ \int_0^1 \kappa(x) \left[\frac{d}{dx} \left(u^H(x) \right) \right]^2 dx \right\}^{1/2} \\
&= \left\{ \sum_{j=1}^{N_h} \int_{x_{j-1}}^{x_j} \kappa(x) \left(\frac{\Delta u_{ms}^H(j)}{h_j} \right)^2 \right\}^{1/2} \\
&= \left\{ \sum_{j=1}^{N_h} \frac{1}{h_j} \kappa_j \left(\Delta u_{ms}^H(j) \right)^2 \right\}^{1/2} \\
&= \left\{ \sum_{j=1}^{N_h} [A_h]_{j,j-1} \left(\Delta u_{ms}^H(j) \right)^2 \right\}^{1/2}.
\end{aligned}$$

The energy norm of u_h is obtained similarly.

3.5 Numerical results

In this section, we investigate some numerical examples to show the optimal convergence of our scheme. In the following examples we take 2^{10} micro element and compute relative energy norm error

$$e_{energy}^H = \frac{|||u_{ms}^H - u^h|||_h}{|||u^h|||_h}.$$

3.5.1 Known Coefficient Case

Example 3.5.1. *Consider the following elliptic problem:*

$$\begin{cases} -\dfrac{d}{dx} \left(\kappa(x) \dfrac{du}{dx} \right) = -1 \ \textit{in} \ \Omega = (0,1), \\ \qquad\qquad\qquad u = 0, \ \textit{if} \ x = 0 \ \textit{or} \ 1, \end{cases}$$

where $\kappa(x) = \frac{2}{3}(1+x)\big(1+\cos(\frac{2\pi x}{\epsilon})^2\big)$.

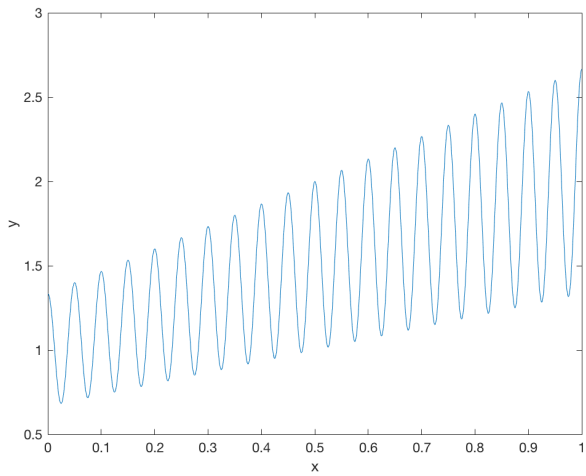


Figure 3.2. Graph of $\kappa(x) = \frac{2}{3}(1+x)(1+\cos(\frac{2\pi x}{\epsilon})^2)$, $\epsilon = \frac{1}{10}$.

The elliptic problem has the homogenized solution

$$u_{\text{hom}}(x) = \frac{3}{2\sqrt{2}}\left(x - \frac{\log(1+x)}{\log 2}\right).$$

In this example, we set $\epsilon = \frac{1}{10}$. Here we recognize the exact coefficient κ but do not know the geometric information on microscale mesh. We may take uniform mesh as an ideal case, or non-uniform mesh is also possible. For example, let $y_0 = 0$ and define

$$y_{j+1} = y_j + 2 \times \frac{rand}{N_h} \text{ for } 1 \le j \le N_h.$$

Then choose $x_j = \frac{y_j}{y_{N_h}}$ as microscale nodes to make $x_{N_h} = 1$. The $rand$ function is used to reflect the ignorance of microscale mesh.

Uniform microscale mesh			Non-uniform microscale mesh		
N^H	e^H_{energy}	Order	N^H	e^H_{energy}	Order
2	5.00E-01		2	5.01E-01	
4	2.50E-01	1.00	4	2.50E-01	1.00
8	1.26E-01	0.99	8	1.26E-01	0.99
16	6.27E-02	1.00	16	6.28E-02	1.00
32	3.08E-02	1.02	32	3.12E-02	1.01
64	1.55E-02	0.99	64	1.60E-02	0.96

Table 3.1. Error of Example 3.5.1.

The relative energy norm error is reported in the Table 3.1, as the number of macroscale nodes is doubled. We observe first-order convergence in both cases regardless of mesh types.

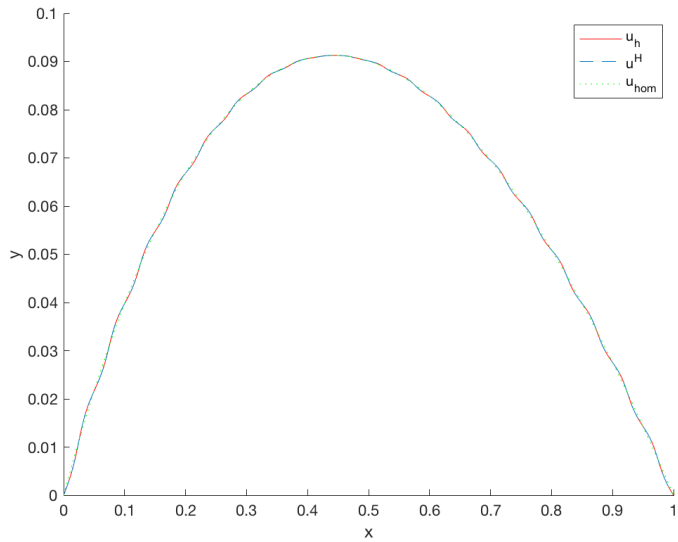
In Figure 3.3, the red line, the blue dashed line and the green dotted line denote the microscale solution u_h, multiscale solution u^H, and the homogenized solution u_{hom}, respectively. All three graphs are almost identical. Since we get similar consequence for other ϵ values, we do not report the result here.

3.5.2 Random Coefficient Case

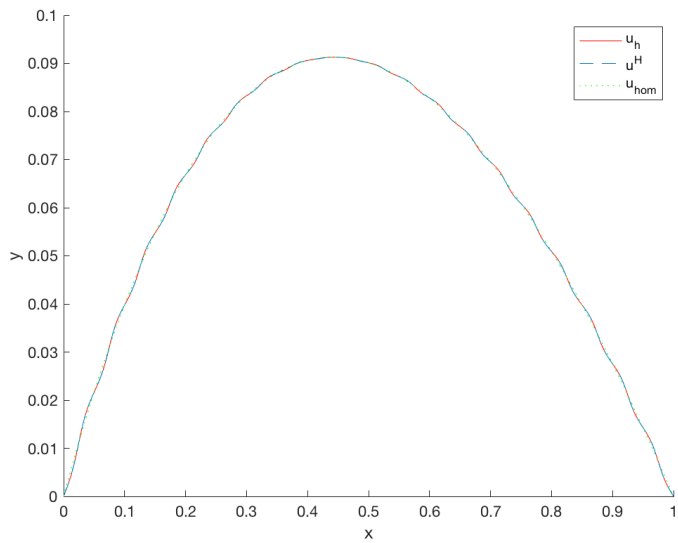
Consider the following elliptic problem:

$$\begin{cases} -\dfrac{d}{dx}\left(\kappa(x)\dfrac{du}{dx}\right) = f \text{ in } \Omega = (0,1), \\ \qquad\qquad\qquad u = 0, \text{ if } x = 0 \text{ or } 1, \end{cases}$$

where κ and f are given randomly. That is, we only have the microscale linear system without knowing the exact form of κ and f. In the previous section



(a) Uniform microscale mesh.



(b) Non-uniform microscale mesh.

Figure 3.3. Solution graph of Example 3.5.1 when $N^H = 64$.

we use the one-to-one correspondence between the off-diagonal element of the microscale stiffness matrix A_h and the average of κ in each micro interval, to formulate macroscale system algebraically. For our simulation, we change the off-diagonal elements of A_h and observe the error. We consider two cases:

1. κ exhibits non-periodic behavior keeping its initial amplitude: Example 3.5.2

2. κ exhibits non-periodic behavior while overall average amplitude grows up: Example 3.5.3

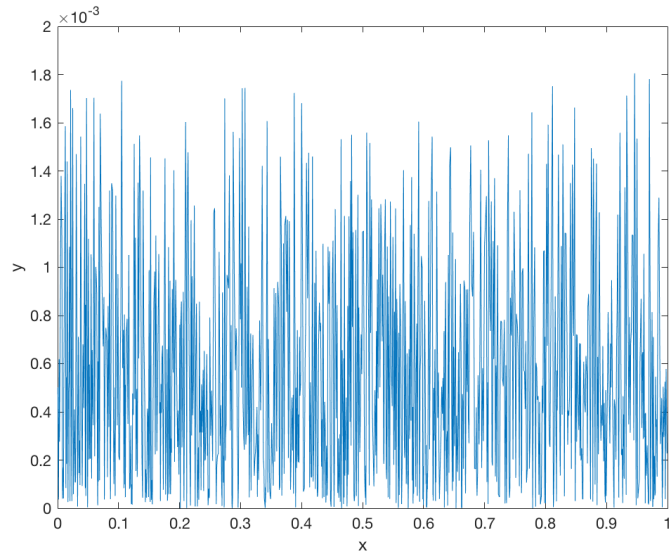
In following examples we use the non-uniform microscale mesh in Example 3.5.1 and the microscale right hand side vector is given by

$$b_j^h = rand \text{ for } 1 \leq j \leq N_h - 1.$$

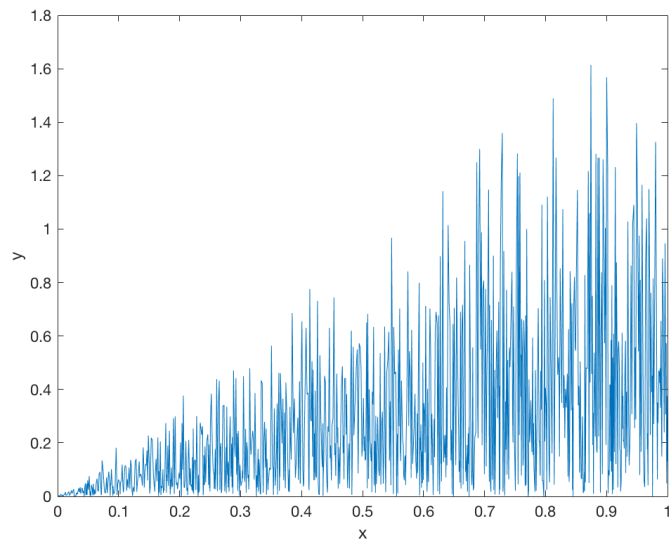
Example 3.5.2. *The off-diagonal elements of A_h are given by*

$$\left[A_h\right]_{j,j-1} = rand \text{ for } 1 \leq j \leq N_h - 1.$$

We observe the relative energy norm error in the Table 3.2. The coefficients are randomly defined by *rand* functions so that two representative simulation results are reported. The convergence order is oscillatory around 1 in both cases. In Figures 3.5 and 3.6, the red solid line and the blue dashed line denote the microscale solution u_h and multiscale solution u^H, respectively. We can see that u^H converges to u_h as the number of macro nodes gradually increases.



(a) Example 3.5.2.



(b) Example 3.5.3.

Figure 3.4. Graph of coefficients κ in simulation 1.

Simulation 1			Simulation 2		
N^H	e^H_{energy}	Order	N^H	e^H_{energy}	Order
2	3.93E-01		2	5.13E-01	
4	2.27E-01	0.79	4	2.66E-01	0.95
8	8.93E-02	1.34	8	1.20E-01	1.15
16	4.40E-02	1.02	16	6.10E-02	0.97
32	2.07E-02	1.09	32	2.86E-02	1.15
64	1.01E-02	1.04	64	1.39E-02	0.99

Table 3.2. Error of Example 3.5.2.

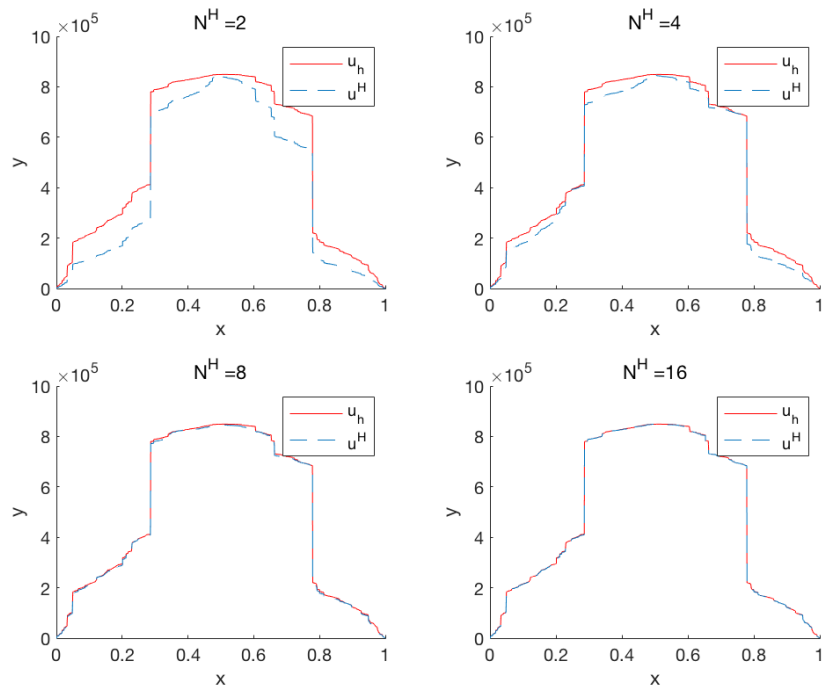


Figure 3.5. Solution graph of Example 3.5.2 for simulation 1.

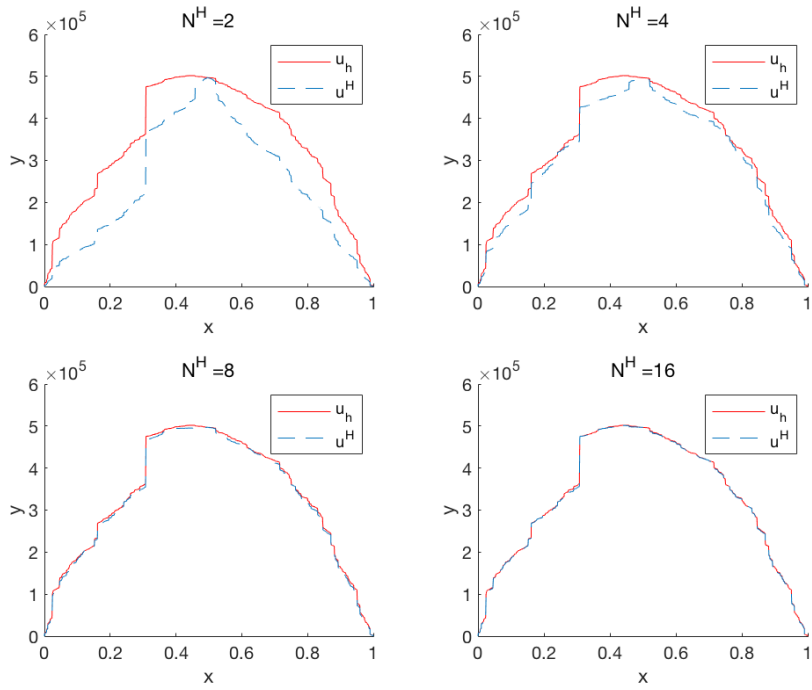


Figure 3.6. Solution graph of Example 3.5.2 for simulation 2.

Example 3.5.3. *The off-diagonal elements of A_h are given by*

$$\left[A_h\right]_{j,j-1} = j * rand \text{ for } 1 \leq j \leq N_h - 1.$$

We see more fluctuating convergence orders in Table 3.3 compared to the previous examples. The range of coefficient in Example 3.5.3 is about 1000 times wider than that of Example 3.5.2 as shown in Figure 3.4. This reflects Theorem 3.4.2 that the error is bounded by the minimum of $\kappa^{-\frac{1}{2}}$. Overall convergence order is almost 1 in both cases. Solution graph is depicted in Figures 3.7 and 3.8 which shows the convergence behavior of u^H to u_h.

Simulation 1			Simulation 2		
N^H	e^H_{energy}	Order	N^H	e^H_{energy}	Order
2	2.85E-01		2	5.86E-01	
4	1.61E-01	0.82	4	4.49E-01	0.38
8	7.60E-01	1.09	8	1.49E-02	1.59
16	4.48E-02	0.76	16	6.80E-02	1.13
32	2.00E-02	1.17	32	3.42E-02	0.99
64	8.70E-03	1.20	64	1.69E-02	1.02

Table 3.3. Error of Example 3.5.3.

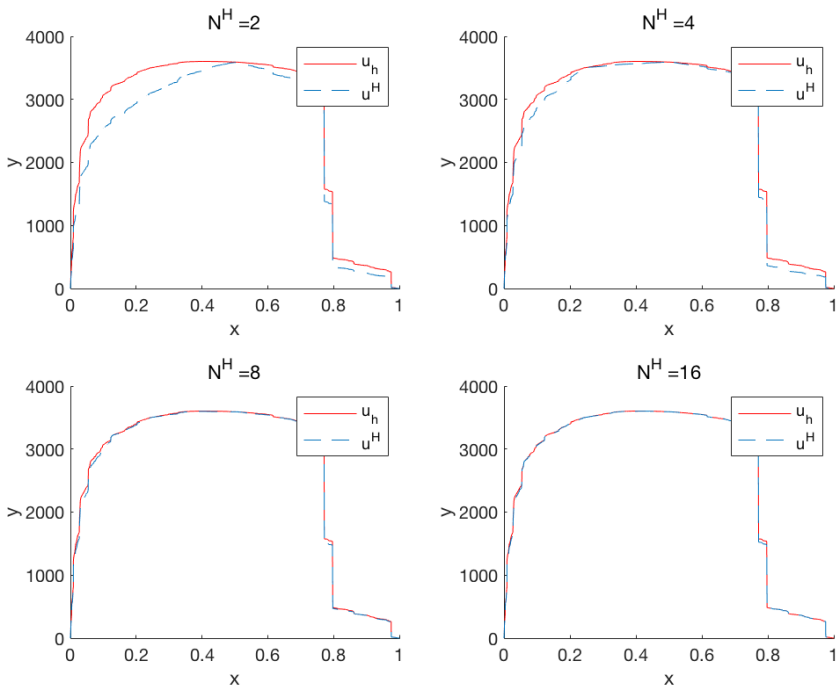


Figure 3.7. Solution graph of Example 3.5.3 for simulation 1.

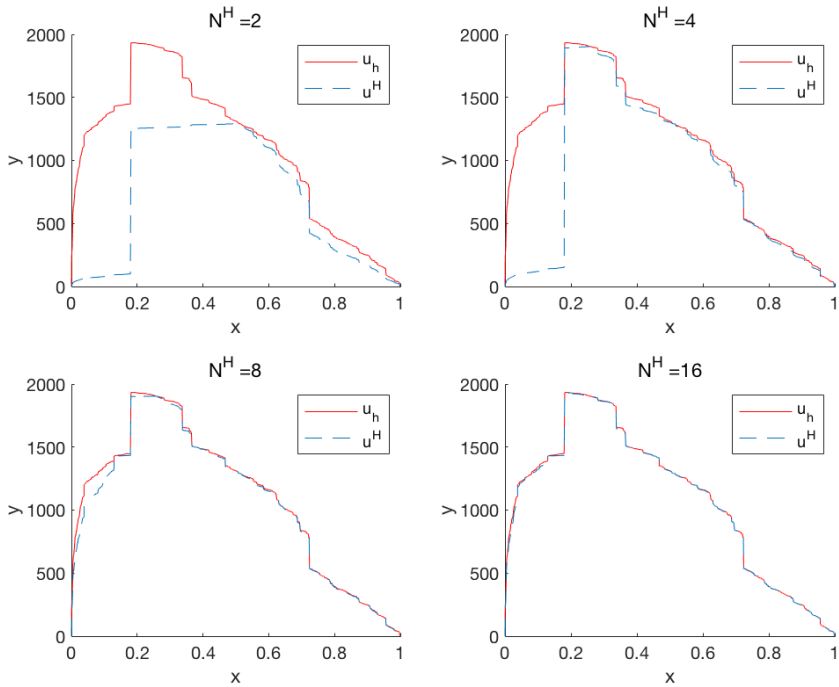


Figure 3.8. Solution graph of Example 3.5.3 for simulation 2.

3.6 2D case

Consider the following elliptic problem:

$$\begin{cases} -\nabla \cdot \big(\kappa(\mathbf{x})\nabla\, u\big) = f \text{ in } \Omega, \\ \qquad\qquad\qquad u = 0 \text{ on } \partial\Omega, \end{cases} \tag{3.24}$$

where $\Omega = (0,1)^2$, κ is a highly heterogeneous coefficient and $f \in H^{-1}(\Omega)$. Denote by $\mathcal{T}^H = \{\Omega^{JK}\}_{J,K}$ and $\mathcal{T}_h = \{\Omega_{jk}\}_{j,k}$ two families of macroscale and microscale triangulations of Ω into macroscale and microscale rectangles such that $\Omega^{JK} = (X^{J-1}, X^J) \times (Y^{K-1}, Y^K)$ and $\Omega_{jk} = (x_{j-1}, x_j) \times (y_{k-1}, y_k)$ where $0 = X^0 < X^1 < \cdots < X^{N_X} = 1$, $0 = Y^0 < Y^1 < \cdots < Y^{N_Y} = 1$ and

$0 = x_0 < x_1 < \cdots < x_{n_x} = 1$, $0 = y_0 < y_1 < \cdots < y_{n_y} = 1$. Also denote by H^{JK} and h_{jk} diameters of Ω^{JK} and Ω_{jk}. H and h represent the macroscale and microscale mesh parameters given by

$$H = \max_{J,K}(H^{JK}), \quad h = \max_{j,k}(h_{jk}). \tag{3.25}$$

3.6.1 Implementation of the DSSY nonconforming element

We use rectangular triangulation to implement the DSSY nonconforming element. Since the DSSY elements are based on the horizontal–type and vertical–type edges, it is more natural to label the edges and basis functions in these two types. For $j = 1, \cdots, N_x$ and $k = 1, \cdots, N_y$, let Ω_{jk} be the (j,k)-rectangle with the four vertices $(x_j, y_k), (x_{j-1}, y_k), (x_{j-1}, y_{k-1}), (x_j, y_{k-1})$, with edges $e_{jk}, f_{jk}, e_{j-1,k}$, and $f_{j,k-1}$, on which the basis functions are respectively given by $\psi_{j,k}(l, \cdot, \cdot) : \Omega_{jk} \to \mathbb{R}, l = 1, 2, 3, 4$. If h_x and h_y denote the lengths of the horizontal and vertical edges, after translating the center of Ω_{jk} to the origin $(0, 0)$ so that Ω_{jk} is translated to $\widehat{K} := (-\frac{h_x}{2}, \frac{h_x}{2}) \times (-\frac{h_y}{2}, \frac{h_y}{2})$, the basis functions are given by

$$\psi(l, x, y) = \begin{cases} \frac{1}{4} + \frac{x}{h_x} + \frac{\theta(\frac{2x}{h_x}) - \theta(\frac{2y}{h_y})}{4\theta(1)}, & l = 1, \\ \frac{1}{4} + \frac{y}{h_y} + \frac{\theta(\frac{2y}{h_y}) - \theta(\frac{2x}{h_x})}{4\theta(1)}, & l = 2, \\ \frac{1}{4} - \frac{x}{h_x} + \frac{\theta(\frac{2x}{h_x}) - \theta(\frac{2y}{h_y})}{4\theta(1)}, & l = 3, \\ \frac{1}{4} - \frac{y}{h_y} + \frac{\theta(\frac{2y}{h_y}) - \theta(\frac{2x}{h_x})}{4\theta(1)}, & l = 4, \end{cases}$$

where $\theta(t) = t^2 - \frac{5}{3}t^4$. The gradients $\nabla\,\psi(l,x,y)$ on $\widehat{K}$ are given by

$$\nabla\,\psi(l,x,y) = \begin{cases} \begin{pmatrix} \frac{1}{h_x} + \frac{\theta'(\frac{2x}{h_x})}{2h_x\theta(1)} \\ -\frac{\theta'(\frac{2y}{h_y})}{2h_y\theta(1)}, \end{pmatrix}, & l = 1, \\ \begin{pmatrix} -\frac{\theta'(\frac{2x}{h_x})}{2h_x\theta(1)} \\ \frac{1}{h_y} + \frac{\theta'(\frac{2y}{h_y})}{2h_y\theta(1)} \end{pmatrix}, & l = 2, \\ \begin{pmatrix} -\frac{1}{h_x} + \frac{\theta'(\frac{2x}{h_x})}{2h_x\theta(1)} \\ -\frac{\theta'(\frac{2y}{h_y})}{2h_y\theta(1)} \end{pmatrix}, & l = 3, \\ \begin{pmatrix} -\frac{\theta'(\frac{2x}{h_x})}{2h_x\theta(1)} \\ -\frac{1}{h_y} + \frac{\theta'(\frac{2y}{h_y})}{2h_y\theta(1)} \end{pmatrix}, & l = 4. \end{cases}$$

One may solve (3.24) on the microscale mesh using finite element method. Let $V_h = \text{Span}(\{\psi_{jk}\}_{j,k} \bigcup \{\phi_{jk}\}_{j,k})$ be the DSSY finite element space associated with $\mathcal{T}_h$, where $\{\psi_{j,k}\}_{j,k}$ is the set of DSSY basis functions associated with DOF at midpoint on each horizontal micro edge, and $\{\phi_{j,k}\}_{j,k}$ is the set of DSSY basis functions associated with DOF at midpoint on each vertical micro edge (see Figure 3.9).

We assume that all the components in A and b in the microscale linear system $Ax = b$ are known, which is constructed by a known finite element method to find $u_h \in V_h$ such that

$$a_h(u_h, v_h) = (f, v_h)_\Omega \quad \forall v_h \in V_h, \tag{3.26}$$

where $a_h(u_h, v_h) = \sum_{j,k}(\kappa\nabla\, u_h, \nabla\, v_h)_{\Omega_{jk}}$. Here, we do not assume that any a priori knowledge is given for the coefficient κ and the exterior source term f. Typically the size of the microscale linear system $Ax = b$ is too huge to solve, and hence we apply certain multiscale finite element method to build a

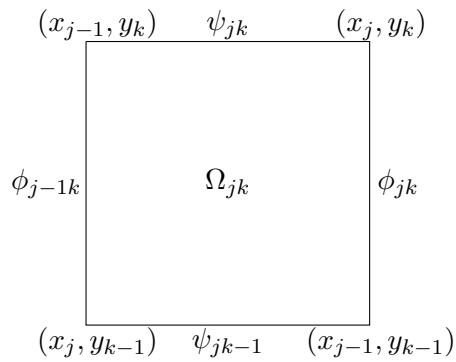


Figure 3.9. Basis functions associated with vertical and horizontal type edges on micro element Ω_{jk}.

reasonable size of macroscale linear system $A^M x^M = b^M$.

In generalized multiscale finite element method, we solve $\kappa-$harmonic problems in each macro element. Using these local solutions, we construct the multiscale basis space V^H and find the solution in V^H. To construct V^H, we assume that $\mathcal{T}_h$ is a refinement of $\mathcal{T}^H$ satisfying $h << H < 1$.

Our approach to the algebraic multiscale method is to construct the macroscale linear system for the macroscale basis functions ψ_{JK}, ϕ_{JK}, which are derived from the microscale linear system $Au = f$ for the microscale basis functions ψ_{jk}, ϕ_{jk}. On our procedure, we use the generalized multiscale finite element method and the following assumptions are imposed:

1. the microscale mesh is rectangular;

2. the linear system is constructed by using the DSSY nonconforming finite element method;

3. the coefficient κ is assumed to be constant on each micro element.

For the sake of convenience, by ϕ_{jk} and ψ_{jk} denote the basis functions

associated with the edges e_{jk} and f_{jk} so that the solution u_h is represented by

$$u_h = \sum_{j'=0}^{n_x}\sum_{k'=1}^{n_y} \alpha_{j'k'}\phi_{j'k'} + \sum_{j'=1}^{n_x}\sum_{k'=0}^{n_y} \beta_{j'k'}\psi_{j'k'}. \tag{3.27}$$

Test (3.26) with u_h represented by (3.27) against $v_h = \phi_{jk}$ and $v_h = \psi_{jk}$ to obtain

$$\begin{aligned}
&\sum_{j'=0}^{n_x}\sum_{k'=1}^{n_y} \alpha_{j'k'} a_h(\phi_{j'k'}, \phi_{jk}) + \sum_{j'=1}^{n_x}\sum_{k'=0}^{n_y} \beta_{j'k'} a_h(\psi_{j'k'}, \phi_{jk}) \\
&\qquad = (f, \phi_{jk}), \quad j = 0, \cdots, n_x,\ k = 1, \cdots, n_y, \qquad (3.28a)\\
&\sum_{j'=0}^{n_x}\sum_{k'=1}^{n_y} \alpha_{j'k'} a_h(\phi_{j'k'}, \psi_{jk}) + \sum_{j'=1}^{n_x}\sum_{k'=0}^{n_y} \beta_{j'k'} a_h(\psi_{j'k'}, \psi_{jk}) \\
&\qquad = (f, \psi_{jk}), \quad j = 1, \cdots, n_x,\ k = 0, \cdots, n_y. \qquad (3.28b)
\end{aligned}$$

Taking into account of the supports of basis functions, we get the following linear system:

$$\begin{aligned}
&A^{\alpha,\alpha}_{jkjk}\alpha_{jk} + A^{\alpha,\alpha}_{j-1kjk}\alpha_{j-1k} + A^{\alpha,\alpha}_{j+1kjk}\alpha_{j+1k} \\
&+A^{\beta,\alpha}_{jkjk}\beta_{jk} + A^{\beta,\alpha}_{jk-1jk}\beta_{jk-1} + A^{\beta,\alpha}_{j+1kjk}\beta_{j+1k} + A^{\beta,\alpha}_{j+1k-1jk}\beta_{j+1k-1} \\
&\qquad = f^{\alpha}_{jk}, \quad j = 0, \cdots, n_x,\ k = 1, \cdots, n_y, \qquad (3.29a)\\
&A^{\beta,\beta}_{jkjk}\beta_{jk} + A^{\beta,\beta}_{jk-1jk}\beta_{jk-1} + A^{\beta,\beta}_{jk+1jk}\beta_{jk+1} \\
&+A^{\alpha,\beta}_{jkjk}\alpha_{jk} + A^{\alpha,\beta}_{j-1kjk}\alpha_{j-1k} + A^{\alpha,\beta}_{jk+1jk}\alpha_{jk+1} + A^{\alpha,\beta}_{j-1k+1jk}\alpha_{j-1k+1} \\
&\qquad = f^{\beta}_{jk}, \quad j = 1, \cdots, n_x,\ k = 0, \cdots, n_y, \qquad (3.29b)
\end{aligned}$$

where $A^{\alpha,\alpha}_{j'k'jk} = a_h(\phi_{j'k'}, \phi_{jk})$, $A^{\alpha,\beta}_{j'k'jk} = a_h(\phi_{j'k'}, \psi_{jk})$, $A^{\beta,\alpha}_{j'k'jk} = a_h(\psi_{j'k'}, \phi_{jk})$, $A^{\beta,\beta}_{j'k'jk} = a_h(\psi_{j'k'}, \psi_{jk})$, $f^{\alpha}_{jk} = (f, \phi_{jk})$ and $f^{\beta}_{jk} = (f, \psi_{jk})$.

A direct computation of the component of the stiffness matrix on $\Omega_{jk} =$

$(x_{j-1}, x_j) \times (y_{k-1}, y_k)$ gives

$$\begin{aligned}
(\nabla\,\phi_{jk}, \nabla\,\phi_{jk})_{\Omega_{jk}} &= \frac{37}{28}\frac{h_{x_j}}{h_{y_k}} + \frac{65}{28}\frac{h_{y_k}}{h_{x_j}}, \\
(\nabla\,\psi_{jk}, \nabla\,\phi_{jk})_{\Omega_{jk}} = (\nabla\,\psi_{jk-1}, \nabla\,\phi_{jk})_{\Omega_{jk}} &= -\frac{37}{28}\frac{h_{x_j}^2 + h_{y_k}^2}{h_{x_j} h_{y_k}}, \\
(\nabla\,\phi_{j-1k}, \nabla\,\phi_{jk})_{\Omega_{jk}} &= \frac{37}{28}\frac{h_{x_j}}{h_{y_k}} + \frac{9}{28}\frac{h_{y_k}}{h_{x_j}}.
\end{aligned} \tag{3.30}$$

Analogous components are obtained by replacing Ω_{jk} by Ω_{jk-1}. Furthermore, we have similar results for ψ_{jk}; just h_{x_j} and h_{y_k} are exchanged in (3.30). Set $\gamma_{jk} = \frac{h_{y_k}}{h_{x_j}}$. By a direct computation, one gets the following expressions:

$$\begin{aligned}
A^{\beta,\beta}_{jkjk} &= \Big(\frac{65}{28}\frac{1}{\gamma_{jk}} + \frac{37}{28}\gamma_{jk}\Big)\kappa_{jk} + \Big(\frac{65}{28}\frac{1}{\gamma_{jk+1}} + \frac{37}{28}\gamma_{jk+1}\Big)\kappa_{jk+1}, \\
A^{\beta,\beta}_{jk-1jk} &= \Big(\frac{9}{28}\frac{1}{\gamma_{jk}} + \frac{37}{28}\gamma_{jk}\Big)\kappa_{jk}, \\
A^{\beta,\beta}_{jk+1jk} &= \Big(\frac{9}{28}\frac{1}{\gamma_{jk+1}} + \frac{37}{28}\gamma_{jk+1}\Big)\kappa_{jk+1}, \\
A^{\alpha,\beta}_{jkjk} &= A^{\alpha,\beta}_{j-1kjk} = -\frac{37}{28}\Big(\frac{1}{\gamma_{jk}} + \gamma_{jk}\Big)\kappa_{jk}, \\
A^{\alpha,\beta}_{jk+1jk} &= A^{\alpha,\beta}_{j-1k+1jk} = -\frac{37}{28}\Big(\frac{1}{\gamma_{jk+1}} + \gamma_{jk+1}\Big)\kappa_{jk+1}.
\end{aligned}$$

First, we need to deduce the coefficient values κ_{jk} and mesh sizes h_{x_j}, h_{y_k} from the microscale linear system (3.29). The result is formulated as the following proposition.

Proposition 3.6.1. *κ_{jk} and h_{x_j}, h_{y_k}, can be determined from the linear system* (3.29).

Proof. At each rectangular elements, except for the 4 corner elements, we can derive at least two information about κ_{jk} from the stiffness matrix. One is $A^{\alpha,\alpha}_{j-1kjk}$ or $A^{\beta,\beta}_{jk-1jk}$ and the other is one of $A^{\alpha,\beta}_{jkjk}$, $A^{\alpha,\beta}_{jkjk-1}$, $A^{\alpha,\beta}_{j-1kjk}$, $A^{\alpha,\beta}_{j-1kjk-1}$ (see Table 3.4.) For example, when we have $A^{\beta,\beta}_{jk-1jk}$ and $A^{\alpha,\beta}_{jkjk}$ for left vertical

element, we can derive the following equalities:

$$A^{\beta,\beta}_{jk-1jk} + A^{\alpha,\beta}_{jkjk} = -\frac{\kappa_{jk}}{\gamma_{jk}} (< 0)$$
$$\frac{A^{\alpha,\beta}_{jkjk}}{A^{\beta,\beta}_{jk-1jk} + A^{\alpha,\beta}_{jkjk}} = \frac{37}{28}(1+\gamma^2_{jk})$$

Hence,

$$\gamma_{jk} = \sqrt{\frac{28}{37}\Big(\frac{A^{\alpha,\beta}_{jkjk}}{A^{\beta,\beta}_{jk-1jk} + A^{\alpha,\beta}_{jkjk}}\Big) - 1}$$
$$\kappa_{jk} = -(A^{\beta,\beta}_{jk-1jk} + A^{\alpha,\beta}_{jkjk})\gamma_{jk}$$

κ_{jk} and γ_{jk} can be derived for other cases in a similar way.

At the corner, we cannot get the value of $A^{\alpha,\alpha}_{j-1kjk}$ or $A^{\beta,\beta}_{jk-1jk}$ from the stiffness matrix. That is, there is only one valid information about κ_{jk} and γ_{jk}. In this case, we need the ratio information from adjacent elements to derive the coefficient and the ratio. First, we can derive γ_{jk} using following relation about ratio, $\gamma_{jk} = \frac{\gamma_{jk+1}\gamma_{j+1k}}{\gamma_{j+1k+1}}$. Since the above formula are valid at every micro element except corners, three of γ_{jk}, γ_{jk+1}, γ_{j+1k}, γ_{j+1k+1} are known and the unknown one would be determined by the ratio information. Then, κ_{jk} can be easily derived from the ratio information and one of $A^{\alpha,\beta}_{jkjk}$, $A^{\alpha,\beta}_{jkjk-1}$, $A^{\beta,\alpha}_{jkjk}$, $A^{\beta,\alpha}_{jkj-1k}$. Now we have every κ_{jk} and γ_{jk} value across all elements. By summation,

$$\sum_{j=1}^{n_x} \frac{1}{\gamma_{jk}} = \frac{\sum_{j=1}^{n_x} h_{x_j}}{h_{y_k}} = \frac{1}{h_{y_k}}$$
$$\sum_{k=1}^{n_y} \gamma_{jk} = \frac{\sum_{k=1}^{n_y} h_{y_j}}{h_{x_j}} = \frac{1}{h_{x_j}}$$

we can determine h_{x_j} and h_{y_k}. □

Position of element	Known information
Left vertical	$A^{\beta,\beta}_{jk-1jk}, A^{\alpha,\beta}_{jkjk}, A^{\alpha,\beta}_{jkjk-1}$
Right vertical	$A^{\beta,\beta}_{jk-1jk}, A^{\alpha,\beta}_{j-1kjk}, A^{\alpha,\beta}_{j-1kjk-1}$
Upper horizontal	$A^{\alpha,\alpha}_{j-1kjk}, A^{\alpha,\beta}_{j-1kjk-1}, A^{\alpha,\beta}_{j-1kjk-1}$
Lower horizontal	$A^{\alpha,\alpha}_{j-1kjk}, A^{\alpha,\beta}_{j-1kjk}, A^{\alpha,\beta}_{jkjk}$
Interior	$A^{\beta,\beta}_{jk-1jk}, A^{\alpha,\alpha}_{j-1kjk}, A^{\alpha,\beta}_{jkjk}, A^{\alpha,\beta}_{jkjk-1}, A^{\alpha,\beta}_{j-1kjk-1}, A^{\alpha,\beta}_{j-1kjk}$

Table 3.4. Known information with respect to the position of element.

3.6.2 Construction of multiscale finite element spaces

We construct multiscale finite element spaces using the approximated values κ_{jk} and h_{x_j}, h_{y_k}. Suppose that $T \in \mathcal{T}^H$ is composed of $n_x^T \times n_y^T$ micro elements. Denote by ϕ_{jk}^T and ψ_{jk}^T the basis functions on T associated with vertical and horizontal edges, respectively.

Snapshot function space

We first consider local snapshot function space $V^{\text{snap}}(T)$ for $T \in \mathcal{T}^H$ using the same notation in §2.3. Recall that the snapshot functions $\widetilde{u}_l^T \in V_h(T)$ are the solutions of following $\kappa-$harmonic problems

$$\begin{cases} -\nabla \cdot \left(\kappa(\mathbf{x}) \nabla \widetilde{u}_l^T\right) = 0 \text{ in } T, \\ \qquad\qquad \widetilde{u}_l^T = \delta_l^T \text{ on } \partial T, \end{cases} \tag{3.31}$$

where δ_l^T is one of $\phi_{0k}^T, \phi_{n_x^T k}^T, \psi_{j0}^T, \psi_{jn_y^T}^T$ for $j = 1, \cdots, n_x^T$ and $k = 1, \cdots, n_y^T$. That is, $\widetilde{u}_l^T$ is the solution of

$$a_T(\widetilde{u}_l^T, v^T) = 0 \quad \forall\, v^T \in V_{h,0}(T) \tag{3.32}$$

satisfying $\widetilde{u}_l^T - \delta_l^T \in V_{h,0}(T)$. Let $\widetilde{u}_l^T \in V_h(T)$ is represented by

$$\widetilde{u}_l^T = \sum_{j'=0}^{n_x^T} \sum_{k'=1}^{n_y^T} \alpha_{j'k'}^T \phi_{j'k'}^T + \sum_{j'=1}^{n_x^T} \sum_{k'=0}^{n_y^T} \beta_{j'k'}^T \psi_{j'k'}^T.$$

Then (3.32) leads to the following equations by setting $v^T = \phi_{jk}^T$ and $v^T = \psi_{jk}^T$:

$$\sum_{j'=0}^{n_x^T} \sum_{k'=1}^{n_y^T} \alpha_{j'k'}^T a_T(\phi_{j'k'}^T, \phi_{jk}^T) + \sum_{j'=1}^{n_x^T} \sum_{k'=0}^{n_y^T} \beta_{j'k'}^T a_T(\psi_{j'k'}^T, \phi_{jk}^T) = 0,$$
$$j = 0, \cdots, n_x^T,\ k = 1, \cdots, n_y^T, \quad (3.33a)$$
$$\sum_{j'=0}^{n_x^T} \sum_{k'=1}^{n_y^T} \alpha_{j'k'}^T a_T(\phi_{j'k'}^T, \psi_{jk}^T) + \sum_{j'=1}^{n_x^T} \sum_{k'=0}^{n_y^T} \beta_{j'k'}^T a_T(\psi_{j'k'}^T, \psi_{jk}^T) = 0,$$
$$j = 1, \cdots, n_x^T,\ k = 0, \cdots, n_y^T. \quad (3.33b)$$

If we take the supports of basis functions into consideration, we have the following linear system for the snapshot function $\widetilde{u}_l^T$:

$$\widetilde{A}_{jkjk}^{\alpha,\alpha}\alpha_{jk} + \widetilde{A}_{j-1kjk}^{\alpha,\alpha}\alpha_{j-1k} + \widetilde{A}_{j+1kjk}^{\alpha,\alpha}\alpha_{j+1k}$$
$$+\widetilde{A}_{jkjk}^{\beta,\alpha}\beta_{jk} + \widetilde{A}_{jk-1jk}^{\beta,\alpha}\beta_{jk-1} + \widetilde{A}_{j+1kjk}^{\beta,\alpha}\beta_{j+1k} + \widetilde{A}_{j+1k-1jk}^{\beta,\alpha}\beta_{j+1k-1}$$
$$= 0, \quad j = 0, \cdots, n_x^T,\ k = 1, \cdots, n_y^T, \quad (3.34a)$$
$$\widetilde{A}_{jkjk}^{\beta,\beta}\beta_{jk} + \widetilde{A}_{jk-1jk}^{\beta,\beta}\beta_{jk-1} + \widetilde{A}_{jk+1jk}^{\beta,\beta}\beta_{jk+1}$$
$$+\widetilde{A}_{jkjk}^{\alpha,\beta}\alpha_{jk} + \widetilde{A}_{j-1kjk}^{\alpha,\beta}\alpha_{j-1k} + \widetilde{A}_{jk+1jk}^{\alpha,\beta}\alpha_{jk+1} + \widetilde{A}_{j-1k+1jk}^{\alpha,\beta}\alpha_{j-1k+1}$$
$$= 0, \quad j = 1, \cdots, n_x^T,\ k = 0, \cdots, n_y^T, \quad (3.34b)$$

where $\widetilde{A}_{j'k'jk}^{\alpha,\alpha} = a_T(\phi_{j'k'}^T, \phi_{jk}^T)$, $\widetilde{A}_{j'k'jk}^{\alpha,\beta} = a_T(\phi_{j'k'}^T, \psi_{jk}^T)$, $\widetilde{A}_{j'k'jk}^{\beta,\alpha} = a_T(\psi_{j'k'}^T, \phi_{jk}^T)$, and $\widetilde{A}_{j'k'jk}^{\beta,\beta} = a_T(\psi_{j'k'}^T, \psi_{jk}^T)$. Since each component of the system (3.34) can be computed from the approximate values of κ_{jk} and h_{x_j}, h_{y_k}, we can construct the snapshot function space $V^{\text{snap}}(T)$. Notice that we can also adopt the over-

sampling technique introduced in §2.3, which only requires us to replace T to a oversampled domain T^+.

Offline function space

For offline functions, we need to solve the spectral problem to find $(\lambda_l^T, u_l^T) \in \mathbb{R} \times V^{\text{snap}}(T)$:

$$a_T(u_l^T, v^T) = \lambda_l^T(\kappa u_l^T, v^T)_T, \quad \forall v^T \in V^{\text{snap}}(T). \tag{3.35}$$

Since any function $v \in V^{\text{snap}}(T)$ is represented by

$$v^T = \sum_{j'=0}^{n_x^T} \sum_{k'=1}^{n_y^T} \alpha_{j'k'}^T \phi_{j'k'}^T + \sum_{j'=1}^{n_x^T} \sum_{k'=0}^{n_y^T} \beta_{j'k'}^T \psi_{j'k'}^T,$$

we can construct the linear system of (3.35) and find offline function u_l^T.

Nonconforming GMsFE space

The moment functions on macro edge E are created by local $\kappa-$harmonic functions in $V_h(\omega(E))$. Thus we can construct the moment function space by the same process to build the snapshot function space. Then the nonconforming GMsFE spaces V^H and $V^{H,0}$ are constructed based on the offline function space and the moment function space. Recall that V^H and $V^{H,0}$ are defined as

$$V^H = \Big\{ \psi \in V^{\text{off}} \Big| \ < [\psi], \zeta >_E = 0, \ \forall \zeta \in \mathcal{M}_H(E), \ \forall E \in \mathcal{E}^{H,0} \Big\},$$
$$V^{H,0} = \Big\{ \psi \in V^{\text{off}} \Big| \ < [\psi], \zeta >_E = 0, \ \forall \zeta \in \mathcal{M}_H(E), \ \forall E \in \mathcal{E}^{H} \Big\}.$$

Construction of b^H

Now we have the GMsFE space $V^{H,0} = \text{Span}(\{\psi^L_{JK}\}_{J,K,L} \bigcup \{\phi^L_{JK}\}_{J,K,L})$, where ψ^L_{JK} and ϕ^L_{JK} denote the L-th multiscale basis function associated with the $JK-$th horizontal macro edge f_{JK} and $JK-$th vertical macro edge e_{JK}, respectively. Suppose that $\Omega_{JK} \cup \Omega_{J+1K}$ is composed of $n_x^{JK} \times n_y^{JK}$ microscale elements. Then ϕ^L_{JK} is represented by

$$\phi^L_{JK} = \sum_{j'=0}^{n_x^{JK}} \sum_{k'=1}^{n_y^{JK}} \alpha^{JK}_{j'k'} \phi^{JK}_{j'k'} + \sum_{j'=1}^{n_x^{JK}} \sum_{k'=0}^{n_y^{JK}} \beta^{JK}_{j'k'} \psi^{JK}_{j'k'},$$

where $\phi^{JK}_{j'k'}$ and $\psi^{JK}_{j'k'}$ is the microscale basis functions of vertical and horizontal type in $\Omega_{JK} \cup \Omega_{J+1K}$, respectively. Therefore it is obvious that the components of b^H can be derived from the summation of that of b_h. We have same argument for ψ^L_{JK}, which completes the construction of b^H.

3.6.3 Numerical results

Example 3.6.2. *Consider the following elliptic problem:*

$$\begin{cases} -\nabla \cdot \big(\kappa(\mathbf{x}) \nabla u\big) = f \ \text{in } \Omega, \\ \qquad\qquad\qquad u = 0 \ \text{on } \partial\Omega, \end{cases}$$

where $\Omega = (0,1)^2$ *and* $\kappa(\mathbf{x}) = 1 + (1+x_1)(1+x_2) + \epsilon \sin(10\pi x_1) \sin(5\pi x_2)$. *The source term* f *is generated by the exact solution*

$$u(x_1, x_2) = \sin(3\pi x_1) x_2 (1 - x_2) + \epsilon \sin(\pi x_1/\epsilon) \sin(\pi x_2/\epsilon).$$

We compare numerical results of GMsFEM and AMS(algebraic multiscale method). We use uniform rectangular mesh and H/h is fixed to 10. Relative

energy and L^2 errors are reported for various ϵ. We observe almost same error behaviors in both methods.

$\frac{1}{H}$	$\frac{1}{h}$	$\dim(V^{H,0})$	GMsFEM		AMS	
			Rel. Energy	Rel. L^2	Rel. Energy	Rel. L^2
5	50	400	0.884	0.388	0.884	0.389
10	100	1800	0.871	0.363	0.871	0.363
20	200	7600	0.346	0.676E-01	0.346	0.678E-01
40	400	31200	0.181	0.181E-01	0.181	0.182E-01

Table 3.5. Convergence for $\epsilon = 0.1$.

$\frac{1}{H}$	$\frac{1}{h}$	$\dim(V^{H,0})$	GMsFEM		AMS	
			Rel. Energy	Rel. L^2	Rel. Energy	Rel. L^2
5	50	400	0.885	0.625	0.885	0.625
10	100	1800	0.355	0.118	0.355	0.119
20	200	7600	0.186	0.316E-01	0.186	0.320E-01
40	400	31200	0.940E-01	0.803E-02	0.939E-01	0.823E-02

Table 3.6. Convergence for $\epsilon = 0.2$.

$\frac{1}{H}$	$\frac{1}{h}$	$\dim(V^{H,0})$	GMsFEM		AMS	
			Rel. Energy	Rel. L^2	Rel. Energy	Rel. L^2
5	50	400	0.335	0.130	0.335	0.132
10	100	1800	0.173	0.342E-01	0.173	0.351E-01
20	200	7600	0.884E-01	0.879E-02	0.885E-02	0.928E-02
40	400	31200	0.444E-01	0.221E-02	0.444E-01	0.246E-02

Table 3.7. Convergence for $\epsilon = 0.5$.

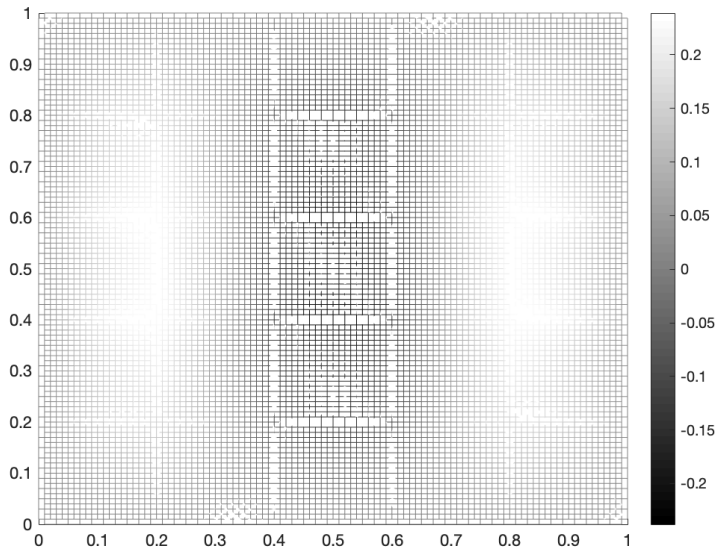


Figure 3.10. Algebraic multiscale solution of $\epsilon = 0.2$ when $1/H = 5, 1/h = 50$.

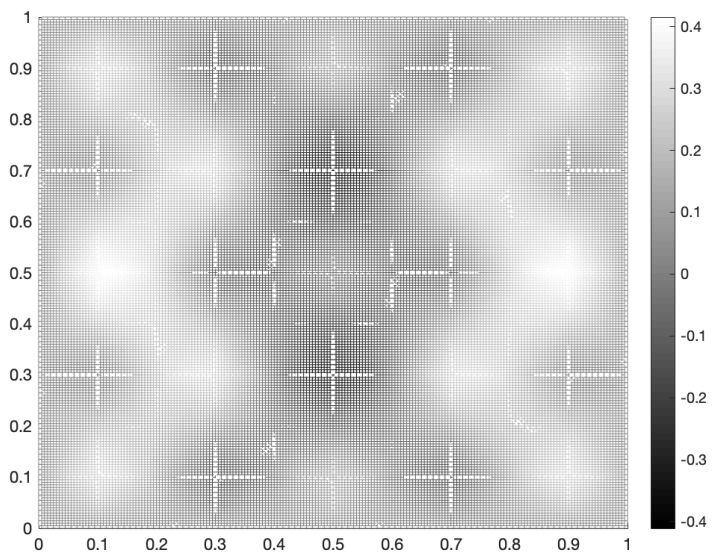


Figure 3.11. Algebraic multiscale solution of $\epsilon = 0.2$ when $1/H = 10, 1/h = 100$.

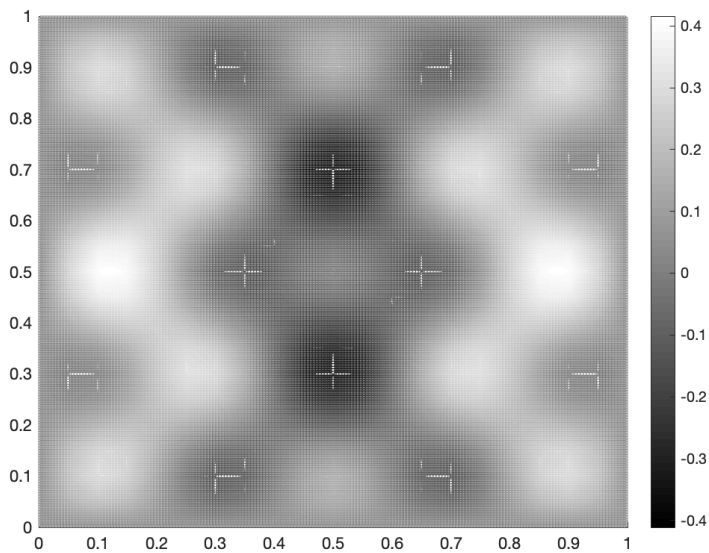


Figure 3.12. Algebraic multiscale solution of $\epsilon = 0.2$ when $1/H = 20, 1/h = 200$.

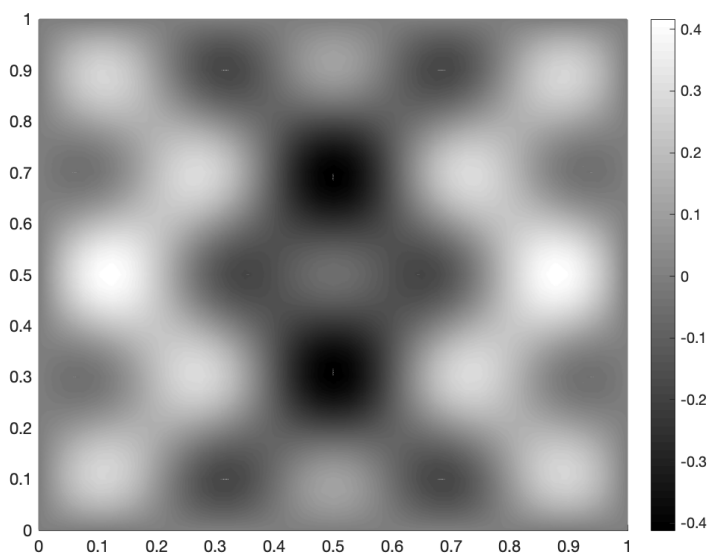


Figure 3.13. Algebraic multiscale solution of $\epsilon = 0.2$ when $1/H = 40, 1/h = 400$.

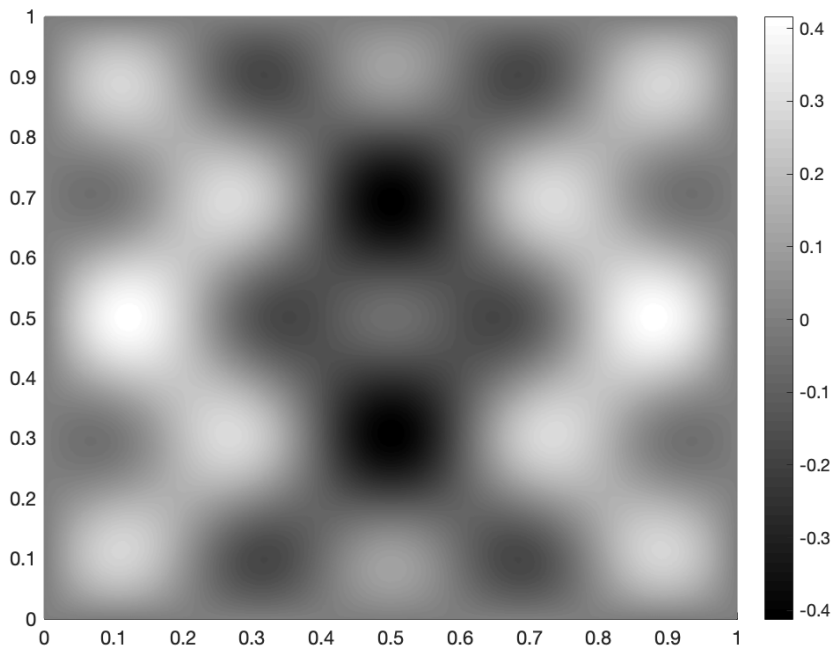


Figure 3.14. Microscale reference solution of $\epsilon = 0.2$ when $1/h = 400$.

Bibliography

[1] D. N. Arnold, D. Boffi, and R. S. Falk. Approximation by quadrilateral finite elements,. *Math. Comp.*, 71:909–922, 2002.

[2] S. C. Brenner and L. Y. Sung. Linear finite element methods for planar elasticity. *Math. Comp.*, 59:321–338, 1992.

[3] Z. Cai, J. Douglas, Jr., J. E. Santos, D. Sheen, and X. Ye. Nonconforming quadrilateral finite elements: A correction. *Calcolo*, 37(4):253–254, 2000.

[4] Z. Cai, J. Douglas, Jr., and X. Ye. A stable nonconforming quadrilateral finite element method for the stationary Stokes and Navier-Stokes equations. *Calcolo*, 36:215–232, 1999.

[5] Z. Chen. Projection finite element methods for semiconductor device equations. *Computers Math. Applic.*, 25:81–88, 1993.

[6] M. Crouzeix and P.-A. Raviart. Conforming and nonconforming finite element methods for solving the stationary Stokes equations. *R.A.I.R.O.–Math. Model. Anal. Numer.*, R-3:33–75, 1973.

[7] J. Douglas, Jr., J. E. Santos, D. Sheen, and X. Ye. Nonconforming Galerkin methods based on quadrilateral elements for second order elliptic problems. *ESAIM–Math. Model. Numer. Anal.*, 33(4):747–770, 1999.

[8] H. Han. Nonconforming elements in the mixed finite element method. *J. Comp. Math.*, 2:223–233, 1984.

[9] J. Hu and Z.-C. Shi. Constrained quadrilateral nonconforming rotated Q_1-element. *J. Comp. Math.*, 23:561–586, 2005.

[10] Y. Jeon, H. Nam, D. Sheen, and K. Shim. A nonparametric DSSY nonconforming quadrilateral element with maximal inf-sup constant. 2013. in preparation.

[11] M. Köster, A. Ouazzi, F. Schieweck, S. Turek, and P. Zajac. New robust nonconforming finite elements of higher order. *Applied Numerical Mathematics*, 62(3):166–184, 2012.

[12] R. Kouhia and R. Stenberg. A linear nonconforming finite element method for nearly incompressible elasticity and Stokes flow. *Comput. Methods Appl. Mech. Engrg.*, 124(3):195–212, 1995.

[13] C.-O. Lee, J. Lee, and D. Sheen. A locking-free nonconforming finite element method for planar elasticity. *Adv. Comput. Math.*, 19(1–3):277–291, 2003.

[14] P. Ming and Z.-C. Shi. Nonconforming rotated Q_1 element for Reissner-Mindlin plate. *Math. Models Methods Appl. Sci.*, 11(8):1311–1342, 2001.

[15] P. Ming and Z.-C. Shi. Two nonconforming quadrilateral elements for the Reissner–Mindlin plate. *Math. Models Methods Appl. Sci.*, 15(10):1503–1517, 2005.

[16] H. Nam, H. J. Choi, C. Park, and D. Sheen. A cheapest nonconforming rectangular finite element for the Stokes problem. *Comput. Methods Appl. Mech. Engrg.*, 257:77–86, 2013.

[17] C. Park and D. Sheen. P_1-nonconforming quadrilateral finite element methods for second-order elliptic problems. *SIAM J. Numer. Anal.*, 41(2):624–640, 2003.

[18] C. Park and D. Sheen. A quadrilateral Morley element for biharmonic equations. *Numer. Math.*, 124(2): 395–413, 2013.

[19] R. Rannacher and S. Turek. Simple nonconforming quadrilateral Stokes element. *Numer. Methods Partial Differential Equations*, 8:97–111, 1992.

[20] Z.-C. Shi. An explicit analysis of Stummel's patch test examples. *Int. J. Numer. Meth. Engrg.*, 20(7):1233–1246, 1984.

[21] Z. C. Shi. The FEM test for convergence of nonconforming finite elements. *Math. Comp.*, 49(180):391–405, 1987.

[22] S. Turek. Efficient solvers for incompressible flow problems, volume 6 of *Lecture Notes in Computational Science and Engineering*. Springer, Berlin, 1999.

[23] M. Wang. On the necessity and sufficiency of the patch test for convergence of nonconforming finite elements. *SIAM J. Numer. Anal.*, 39(2):363–384, 2001.

[24] Z. Zhang. Analysis of some quadrilateral nonconforming elements for incompressible elasticity. *SIAM J. Numer. Anal.*, 34(2):640–663, 1997.

[25] Z. Meng, J. Cui and Z. Luo. A new rotated nonconforming quadrilateral element. J. Sci. Comp., 74(1):324–335, 2018.

[26] Y. Jeon, H. Nam, D. Sheen, and K. Shim. A class of nonparametric DSSY nonconforming quadrilateral elements. *ESAIM–Math. Model. Numer. Anal.*, 47(6):1783–1796, 2013.

[27] P. G. Ciarlet. The finite element method for elliptic problems, volume 4 of *Stiduies in Mathematics and Its Applications*. North-Holland, Amsterdam, 1978.

[28] H. E. Osborn and U. Banerjee Estimation of the effect of numerical integration in finite element eigenvalue approximation. *Numerische Mathematik*, 56:735–762, 1990.

[29] I. Babuška, U. Banerjee, and H. Li. The effect of numerical integration on the finite element approximation of linear functionals. *Numerische Mathematik*, 117:65–88, 2011.

[30] A. Abdulle The finite element heterogeneous multiscale method: a computational strategy for multiscale PDEs. *GAKUTO International Series Mathematical Sciences and Applications*, 31:135–184, 2009.

[31] A. Abdulle, E. Weinan, B. Engquist, and E. Vanden-Eijnden. The heterogeneous multiscale method. *Acta Numerica*, 21:1–87, 2012.

[32] Y. Efendiev, J. Galvis and T. Y. Hou. Generalized multiscale finite element methods (GMsFEM). *Journal of Computational Physics*, 251:116–135, 2013.

[33] Y. Efendiev and T. Y. Hou. Multiscale finite element methods: theory and applications *Springer Science & Business Media*, 2009.

[34] Y. Efendiev, R. Lazarov, M. Moon, and K. Shi. A spectral multiscale hybridizable discontinuous Galerkin method for second order elliptic problems. *Computer Methods in Applied Mechanics and Engineering*, 292:243–256, 2015.

[35] Y. Efendiev, R. Lazarov and K. Shi. A multiscale HDG method for second order elliptic equations. Part I. Polynomial and homogenization-based multiscale spaces. *SIAM Journal on Numerical Analysis*, 53:342–369, 2015.

[36] T. Hou, X.-H. Wu and Z. Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Mathematics of Computation of the American Mathematical Society*, 68(227):913–943, 1999.

[37] Y. Efendiev, T. Hou and X.-H. Wu. Convergence of a nonconforming multiscale finite element method. *SIAM Journal on Numerical Analysis*, 37(3):888–910, 2000.

[38] C. S. Lee and D. Sheen. Nonconforming generalized multiscale finite element methods *Journal of Computational and Applied Mathematics*, 311:215–229, 2017.

국문초록

본 학위논문에서는 일반적인 사각형에서 정의되는 비모수적 DSSY 비순응 유한요소공간을 고려한다. 1장에서는 유한요소법을 이용해 이차 타원형 문제를 해결할 때 수치 적분법이 해의 수렴속도에 작용하는 효과를 분석한다. 최적의 수렴 속도를 변화시키지 않는 수치 적분법의 충분 조건을 구하고, 이를 이용해 DSSY 유한요소에 적합한 새로운 구적법 공식을 고안한다. 단 3개의 점만을 이용해 최적의 수렴 속도를 얻을 수 있음을 보이고 다양한 수치적 결과들을 제시한다.

2장에서는 비모수적 DSSY 비순응유한요소공간을 적용한 일반화된 멀티스케일 비순응유한요소법을 연구한다. 일반화된 멀티스케일 유한요소공간은 두 개의 함수공간으로 구성된다. 첫 번째는 offline 함수공간으로 국소적 조화 문제를 풀어 얻어지는 snapshot 함수공간에 스펙트럼 분해를 적용하여 얻어진다. 두 번째는 moment 함수공간으로 국소적으로 얻어진 offline 함수들 간의 연속성을 부여하는 데 이용된다. 이러한 논의와 함께 1장에서 고안한 구적법 공식을 적용한 수치적 결과들을 제시한다.

3장에서는 대수적 멀티스케일 방법을 소개한다. 이차 타원형 문제의 계수와 소스 항을 모르는 상태에서 단지 미시적 스케일의 선형 시스템만 알고 있을 때, 이 시스템의 구성 성분에 대한 대수적 정보만을 바탕으로 거시적 스케일의 선형 시스템을 건설한다. 먼저 일차원 문제를 구체적으로 분석하고 이차원 문제를 일반화된 멀티스케일 비순응유한요소법을 이용하여 연구한다. 수치적 결과들을 보여준다.

주요어 : DSSY 유한요소, 비모수 유한요소, 일반화된 멀티스케일 유한요소법, 대수적 멀티스케일 방법, 수치 적분, 구적법 공식, 타원형 문제

학번 : 2012-20258