Multiscale Model Reduction for a Class of Optimal Control Problems with Highly Oscillating Coefficients

Tak Shing Au Yeung and Eric Chung

1 Introduction

The paper is concerned with the discretization of a class of elliptic optimal control problems with highly heterogeneous coefficient:

$$\inf J(u) = F(y,u) = \frac{1}{2}\|y - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u\|^2_{L^2(\Omega)} \tag{1}$$

subject to the state equations

$$-\mathrm{div}(\kappa(x)\nabla y) = u, \qquad \text{in } \Omega, \tag{2}$$

$$y = 0, \qquad \text{on } \Gamma, \tag{3}$$

and to the control constraints

$$a \le u(x) \le b \qquad \text{for a.e. } x \in \Omega, \tag{4}$$

where $\Omega \subset \mathbb{R}^2$ is a bounded polygonal Lipschitz domain and Γ is the boundary of Ω; $\kappa(x)$ is a high-contrast heterogeneous permeability field with $0 < \kappa_0 \le \kappa(x) \le \kappa_1$ and a, b are real numbers. In (1), we assume $y_d \in L^2(\Omega)$. Moreover, $\nu > 0$ is a fixed positive number. We denote the set of admissible controls by U_{ad}:

$$U_{ad} = \{u \in L^2(\Omega) : a \le u \le b, \text{ a.e. in } \Omega\}.$$

Tak Shing Au Yeung
Department of Mathematics, The Chinese University of Hong Kong, Hong Kong SAR e-mail: iauyeung@math.cuhk.edu.hk

Eric Chung
Department of Mathematics, The Chinese University of Hong Kong, Hong Kong SAR e-mail: tschung@math.cuhk.edu.hk

In many practical situations, one may encounter heterogeneous media such as fractured media or porous media with high contrast channels. The coefficients for these problems usually contain scale disparity and high contrast regions. Solutions to the problems in these scenarios can contain multiple scales, and very fine computational meshes are typically needed in order to capture these scales. Because of these reasons, some type of model reduction is crucial for these problems. These reduced models are usually constructed based on a coarse grid, whose size does not necessarily resolve any of the scales. In addition, the resulting solutions are required to be robust with respect to the scales and the contrasts of the media, which is the main challenge.

There are existing multiscale approaches, such as [1, 2, 6, 7, 4]. The method in this paper is based on the Constraint Energy Minimizing Generalized Multiscale Finite Element Method (CEM-GMsFEM) [3]. In general, the method has two computational stages, called the offline and the online stages. In the offline stage, some computations are performed once and the reduced model is obtained. In the online stage, the problem formulated using the reduced model is solved when the input arguments and source terms are provided. The key to the success of the method is that the reduced model is only computed once in the offline stage, and the model can be used repeatedly in the online stage for various choices of input parameters and sources. In the offline stage, we will construct some local multiscale basis functions. The construction begins with a local auxiliary space, which is defined for each coarse element. The local auxiliary space is determined using a local spectral problem, which is able to identify high contrast channelized networks and fractures, as well as identify some important modes of the solution. We will use the first few eigenfunctions corresponding to small eigenvalues as the local auxiliary functions. Next, for each auxiliary function on a target coarse element, we will define a corresponding target multiscale basis function. The multiscale basis function is obtained by minimizing an energy over an oversampling region, obtained by extending the target coarse element by a few coarse grid layers, subject to some orthogonality conditions. These orthogonality conditions require that the target multiscale basis function is orthogonal to all auxiliary functions except the one being selected. The resulting multiscale basis functions have several important properties. One of them is that these basis functions are localized, providing the foundation of computing numerically on local oversampling regions. Another property is that the resulting coarse model based on the Galerkin formulation is first order convergent with respect to the coarse mesh size in the natural energy norm. The error bound is independent of the heterogeneities and contrast of the medium parameter κ. Hence the reduced model is very robust.

2 Method description

This section will give the detail of our multiscale method and state the main convergence results. First of all, we introduce the adjoint equation

$$-\mathrm{div}(\kappa(x)\nabla p) = y - y_d, \qquad \text{in } \Omega, \tag{5}$$
$$p = 0, \qquad \text{on } \Gamma. \tag{6}$$

We call the solution y of (2)-(3) for a control u an associated state to u and denote it as $y(u)$. In the same way, we call the solution p of (7)-(6) corresponding to $y(u)$ an associated adjoint state to u and denote it as $p(u)$. We introduce the projection

$$\Pi_{[a,b]}(f(x)) = \max(a, \min(b, f(x))).$$

Then we can formulate the necessary and sufficient first order optimality condition for (1)-(4).

Lemma 1 *A necessary and sufficient condition for the optimality of a control $\bar{u}$ with corresponding state $\bar{y} = y(\bar{u})$ and adjoint state $\bar{p} = p(\bar{u})$, respectively, is that the equation*

$$\bar{u} = \Pi_{[a,b]}\left(-\frac{1}{\nu}\bar{p}\right) \tag{7}$$

holds, where the state and adjoint equations for control $\bar{u}$ is given by:

$$-div(\kappa(x)\nabla\bar{y}) = \bar{u}, \qquad in\ \Omega,$$
$$\bar{y} = 0, \qquad on\ \Gamma$$

and

$$-div(\kappa(x)\nabla\bar{p}) = \bar{y} - y_d, \qquad in\ \Omega,$$
$$\bar{p} = 0, \qquad on\ \Gamma.$$

Moreover, due to (7), we obtain $\bar{u} \in H^1(\Omega)$. See Theorem 2.28 in [8] for details.

We are now in a position to introduce the discretized problem. We apply a multiscale finite element based approximation of the optimal control problem (1)-(4). First, the notions of fine and coarse grids are introduced. Let T_H be a conforming partition of Ω into finite elements. Here, H is the coarse-mesh size and this partition is called coarse grid. We let N_c be the number of vertices and N be the number of elements in the coarse mesh. We assume that each coarse element is partitioned into a connected union of fine-grid cells and this partition is called T_h. Note that T_h is a refinement of the coarse grid T_H with the mesh size h. It is assumed that the fine grid is sufficiently fine to resolve the solution.

Moreover, we set

$$U_H = \{u \in L^\infty(\Omega) : u|_{T'} \text{ is constant on all } T' \in T_H\},$$
$$U_H^{ad} = U_H \cap U_{ad},$$
$$V = H_0^1(\Omega).$$

For each $u_H \in U_H$, the solution $y(u_H)$ of (2)-(3) satisfies

$$a(y(u_H), v) = \int_\Omega u_H v \, dx \qquad \forall v \in V, \tag{8}$$

where $a : V \times V \to \mathbb{R}$ is the bilinear form defined by $a(y, v) = \int_\Omega \kappa \nabla y \cdot \nabla v \, dx$. We define the energy norm $\|y\|_a = a(y, y)^{\frac{1}{2}}$. Notice that our goal is to construct a numerical scheme that gives the cell average of the control on the coarse grid.

2.1 Multiscale basis functions

We will construct V_{ms}, which is the space spanned by all multiscale basis functions. Then the multiscale solution y_{ms} is defined as the solution of the following problem: find $y_{ms} \in V_{ms}$ such that

$$a(y_{ms}(u_H), v) = \int_\Omega u_H v \, dx \qquad \forall v \in V_{ms}. \tag{9}$$

We will first construct our auxiliary multiscale basis functions, which will be constructed for each coarse cell K in the coarse grid. Let K_i be the i-th coarse cell and let $V(K_i)$ be the restriction of V on K_i, which is $H^1(K_i)$. Following the construction from [3], we need a local spectral problem, which is to find a real number $\lambda_j^{(i)}$ and a function $\phi_j^{(i)} \in V(K_i)$ such that

$$a_i(\phi_j^{(i)}, w) = \lambda_j^{(i)} s_i(\phi_j^{(i)}, w), \qquad \forall w \in V(K_i), \tag{10}$$

where a_i is a symmetric non-negative definite bilinear operator and s_i is a symmetric positive definite bilinear operators defined on $V(K_i) \times V(K_i)$. We assume the normalization $s_i(\phi_j^{(i)}, \phi_j^{(i)}) = 1$. Notice that $\lambda_j^{(i)}$ depends on H. In the numerical implementation, we need a fine grid in order to compute $\phi_j^{(i)}$. Based on the analysis in [3], we can choose

$$a_i(v, w) = \int_{K_i} \kappa \nabla v \cdot \nabla w \, dx, \quad s_i(v, w) = \int_{K_i} \tilde{\kappa} v w \, dx$$

where $\tilde{\kappa} = \sum_{j=1}^{N_c} \kappa |\nabla \chi_j^{ms}|^2$ and $\{\chi_j^{ms}\}_{j=1}^{N_c}$ are the standard multiscale finite element (MsFEM) basis functions or piecewise bilinear basis, which satisfy the partition of unity property. We note that $\tilde{\kappa}$ is positive and it is important in the estimate of localization of basis functions, see Lemma 3 of [3]. We let $\lambda_j^{(i)}$ be the eigenvalues of (10) arranged in ascending order. We will use the first l_i eigenfunctions to construct our local auxiliary multiscale space $V_{aux}^{(i)}$, where $V_{aux}^{(i)} = \text{span}\{\phi_j^{(i)} | j \leq l_i\}$. The precise choice of l_i is based on a given tolerance. In particular, we let $\Lambda = \min_{1 \leq i \leq N} \lambda_{l_i+1}^{(i)}$. Then we can choose l_i so that Λ is less than a given tolerance, which can be chosen as $O(1)$. Such tolerances will be introduced in Theorem 1. The global auxiliary multiscale space V_{aux} is the sum of these local auxiliary multiscale spaces, namely

$V_{aux} = \oplus_{i=1}^{N} V_{aux}^{(i)}$. This space is used to construct the target multiscale basis functions that are ϕ-orthogonal to the auxiliary space V_{aux}. The notion of ϕ-orthogonality will be defined next.

For the local auxiliary multiscale space $V_{aux}^{(i)}$, the bilinear form s_i in (10) defines an inner product with norm $\|v\|_{s(K_i)} = s_i(v,v)^{\frac{1}{2}}$. These local inner products and norms provide natural definitions of inner product and norm for the global auxiliary multiscale space V_{aux}, which are defined by

$$s(v,w) = \sum_{i=1}^{N} s_i(v,w), \quad \|v\|_s = s(v,v)^{\frac{1}{2}}, \quad \forall v \in V_{aux}.$$

We note that $s(v,w)$ and $\|v\|_s$ are also an inner product and norm for the space V. Using the above inner product, we can define the notion of ϕ-orthogonality in the space V. Given a function $\phi_j^{(i)} \in V_{aux}$, we say that a function $\psi \in V$ is $\phi_j^{(i)}$-orthogonal if

$$s(\psi, \phi_j^{(i)}) = 1, \quad s(\psi, \phi_{j'}^{(i')}) = 0, \ \text{ if} j' \neq j \text{ or } i' \neq i.$$

We remark that the function $\phi_j^{(i)}$ has support K_i, and we assume that $\phi_j^{(i)}$ is zero outside K_i. Now, we let $\pi_i : L^2(K_i) \to V_{aux}^{(i)}$ be the projection with respect to the inner product $s_i(v,w)$. So, the operator π_i is given by

$$\pi_i(u) = \sum_{j=1}^{l_i} s_i(u, \phi_j^{(i)})\phi_j^{(i)}, \quad \forall u \in V.$$

In addition, we let $\pi : L^2(\Omega) \to V_{aux}$ be the projection with respect to the inner product $s(v,w)$. So, the operator π is given by

$$\pi(u) = \sum_{i=1}^{N} \sum_{j=1}^{l_i} s_i(u, \phi_j^{(i)})\phi_j^{(i)}, \quad \forall u \in V.$$

Note that $\pi = \sum_{i=1}^{N} \pi_i$.

We next present the construction of our multiscale basis functions. For each coarse element K_i, we define an oversampled domain $K_{i,m} \subset \Omega$ by enlarging K_i by m coarse grid layers, where $m \geq 1$ is an integer. An illustration of the fine grid, coarse grid, and oversampling domain are shown in Fig. 1. We emphasize that the basis functions $\psi_{j,ms}^{(i)}$ are supported in the oversampling region $K_{i,m}$ with m being the number of oversampling layers. We will state in Theorem 1 the requirement on this integer m.

We next define the multiscale basis function $\psi_{j,ms}^{(i)} \in V_0(K_{i,m})$ by

$$\psi_{j,ms}^{(i)} = \text{argmin} \left\{ a(\psi,\psi) | \psi \in V_0(K_{i,m}), \quad \psi \text{ is } \phi_j^{(i)}\text{-orthogonal} \right\} \tag{11}$$

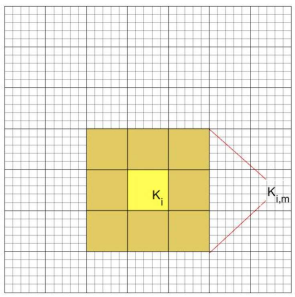

Fig. 1: Illustration of the coarse grid, fine grid and oversampling domain.

where $V(K_{i,m})$ is the restriction of V in $K_{i,m}$ which is $H^1(K_{i,m})$, and $V_0(K_{i,m})$ is the subspace of $V(K_{i,m})$ with zero trace on $\partial K_{i,m}$, i.e. $V_0(K_{i,m}) = H_0^1(K_{i,m})$. Equivalently, we find $\psi_{j,ms}^{(i)} \in V_0(K_{i,m})$ and $\mu \in V_{aux}^{(i,m)}$ that satisfy the following

$$\begin{aligned} a(\psi_{j,ms}^{(i)}, v) + s(v, \mu) &= 0, \quad \forall v \in V_0(K_{i,m}), \\ s(\psi_{j,ms}^{(i)}, \nu) &= s(\phi_j^{(i)}, \nu), \quad \forall \nu \in V_{aux}^{(i,m)}. \end{aligned} \tag{12}$$

In the above, we define $V_{aux}^{(i,m)} = \oplus V_{aux}^{(j)}$ where the sum is over all $K_j \subset K_{i,m}$. Our multiscale finite element space V_{ms} is defined by

$$V_{ms} = \text{span}\left\{\psi_{j,ms}^{(i)} |\ 1 \le j \le l_i,\ 1 \le i \le N\right\}.$$

Finally, we set

$$V_H = V_{ms} \subset H_0^1(\Omega).$$

This is the coarse space for the systems (2)-(3) and (7)-(6).

2.2 The proposed method

We will use the space U_H^{ad} for the approximation of the control u. For the state variables y and p, we will use the space V_H. For each $u_H \in U_H$, the approximate solution $y_H(u_H) \in V_H$ of (2)-(3) satisfies

$$a(y_H(u_H), v_H) = \int_\Omega u_H v_H \, dx, \qquad \forall v_H \in V_H. \tag{13}$$

In other words, $y_H(u_H)$ is the approximated state associated with u_H. The finite dimensional approximation of the optimal control problem is defined as: find $u_H \in U_H^{ad}$ such that is minimizes the following functional

$$J(u_H) = \frac{1}{2}\|y_H(u_H) - y_d\|^2_{L^2(\Omega)} + \frac{\nu}{2}\|u_H\|^2_{L^2(\Omega)}. \tag{14}$$

The adjoint equation is discretized in the same way: find $p_H(u_H) \in V_H$ such that

$$a(p_H(u_H), v_H) = \int_\Omega (y_H(u_H) - y_d)v_H \, dx \qquad \forall v_H \in V_H. \tag{15}$$

From now on, we denote the optimal control of the discrete optimization problem as $\bar{u}_H$ and corresponding state and adjoint solutions as $\bar{y}_H$ and $\bar{p}_H$ respectively. That is $\bar{y}_H = y_H(\bar{u}_H)$ and $\bar{p}_H = p_H(\bar{u}_H)$. We remark that the associated adjoint state $\bar{p}$ belongs to the space $H^1(\Omega)$. The optimal control $\bar{u}$ is obtained by the projection formula (7).

Let $\bar{u}$ be the solution of (1)-(4). We define a piecewise constant function by using the local mean value of $\bar{u}$:

$$w_H(x) = \frac{\int_{T_i} \bar{u}(x)dx}{\int_{T_i} 1dx} \qquad \text{if } x \in T_i \text{ where } T_i \in T_H. \tag{16}$$

It is clear that $w_H \in U_H^{ad}$.

Now we are able to formulate our convergence result.

Theorem 1 *Let $\bar{u}_H$ be the solution of (14). Moreover, if the number of oversampling layers $m = O(\log(H^{-1}\kappa_0^{-1}\kappa_1))$ and χ_i are bilinear partition of unity, then we have the following error bound*

$$\|\bar{u}_H - \bar{u}\|_{L^2(\Omega)} + \|\bar{y}_H - \bar{y}\|_a + \|\bar{p}_H - \bar{p}\|_a \leq CH\Lambda^{-\frac{1}{2}}\nu^{-1}. \tag{17}$$

where Λ is the minimal eigenvalue that the corresponding eigenvector is not included in the auxiliary space, that is, $\Lambda = \min_{1\leq i\leq N} \lambda^{(i)}_{l_i+1}$. Moreover, the constant C is independent of the mesh size and the coefficient κ.

Note that the precise equation for m can be found at the end of Section 5 in [3].

2.3 Outline of error analysis

We will briefly outline the error analysis and a proof of Theorem 1. Using the results in [3], we obtain Lemmas 2-4.

Lemma 2 *Let $u \in L^2(\Omega)$. Moreover, if the number of oversampling layers $m = O(\log(H^{-1}\kappa_0^{-1}\kappa_1))$ and $\{\chi_i\}$ are bilinear partition of unity, then we have*

$$\|y(u) - y_H(u)\|_a \leq CH\Lambda^{-\frac{1}{2}}\|\kappa^{-\frac{1}{2}}u\|_{L^2(\Omega)}, \tag{18}$$

$$\begin{aligned}\|p(u) - p_H(u)\|_a &\leq CH\Lambda^{-\frac{1}{2}}(\|\kappa^{-\frac{1}{2}}y\|_{L^2(\Omega)} + \|\kappa^{-\frac{1}{2}}y_d\|_{L^2(\Omega)}),\\ &\leq CH\Lambda^{-\frac{1}{2}}\kappa_0^{-\frac{1}{2}}(\|u\|_{L^2(\Omega)} + \|y_d\|_{L^2(\Omega)}).\end{aligned} \tag{19}$$

Lemma 3 *Let w_H be the function defined by (16). In addition, suppose that the assumptions of Lemma 2 are fulfilled. Then we have*

$$\|y_H(\bar{u}) - y_H(w_H)\|_a \leq CH\kappa_0^{-\frac{1}{2}}\|\bar{u}\|_{H^1(\Omega)}, \tag{20}$$

$$\|p_H(\bar{u}) - p_H(w_H)\|_a \leq CH\kappa_0^{-\frac{1}{2}}\|\bar{u}\|_{H^1(\Omega)}, \tag{21}$$

$$\|\bar{p} - p_H(w_H)\|_a \leq CH\Lambda^{-\frac{1}{2}}\kappa_0^{-\frac{1}{2}}(\|\bar{u}\|_{H^1(\Omega)} + \|y_d\|_{L^2(\Omega)}). \tag{22}$$

Lemma 4 *The following variational inequalities are necessary and sufficient for the optimality of the unique solutions of (1)-(4) and (14):*

$$(\bar{p} + \nu\bar{u}, u - \bar{u})_{L^2(\Omega)} \geq 0 \qquad \forall u \in U_{ad}, \tag{23}$$

$$(p_H(\bar{u}_H) + \nu\bar{u}_H, \zeta_H - \bar{u}_H)_{L^2(\Omega)} \geq 0 \qquad \forall \zeta_H \in U_H^{ad}. \tag{24}$$

Now, we derive a variational inequality for the function w_H. We define a new function $\hat{p}$ by

$$\hat{p}(x) = \frac{\int_{T_i} \bar{p}(x)\, dx}{\int_{T_i} 1\, dx}, \qquad \text{where } x \in T_i \in T_H.$$

Then, using (23), we obtain

$$(\hat{p} + \nu w_H, \bar{u}_H - w_H)_{L^2(\Omega)} \geq 0. \tag{25}$$

Moreover, we can test inequality (24) with the function w_H and get

$$(p_H(\bar{u}_H) + \nu\bar{u}_H, w_H - \bar{u}_H)_{L^2(\Omega)} \geq 0. \tag{26}$$

Combining the results, we have

$$\nu\|w_H - \bar{u}_H\|^2_{L^2(\Omega)} \leq (\hat{p} - p_H(\bar{u}_H), \bar{u}_H - w_H)_{L^2(\Omega)}. \tag{27}$$

The right-hand side of (27) can be written as

$$\begin{aligned}(\hat{p} - p_H(\bar{u}_H), \bar{u}_H - w_H)_{L^2(\Omega)} &= (p_H(w_H) - p_H(\bar{u}_H), \bar{u}_H - w_H)_{L^2(\Omega)} \\ &\quad + (\bar{p} - p_H(w_H), \bar{u}_H - w_H)_{L^2(\Omega)} \\ &\quad + (\hat{p} - \bar{p}, \bar{u}_H - w_H)_{L^2(\Omega)}.\end{aligned} \tag{28}$$

Next we estimate these three terms. The first term on the right hand side of (28) can be estimated as

$$\begin{aligned}&(p_H(w_H) - p_H(\bar{u}_H), \bar{u}_H - w_H)_{L^2(\Omega)} \\ &= (y_H(w_H) - y_H(\bar{u}_H), y_H(\bar{u}_H) - y_H(w_H))_{L^2(\Omega)} \\ &\leq 0.\end{aligned} \tag{29}$$

The second term on the right hand side of (28) can be estimated using (22):

$$
\begin{aligned}
&(\bar{p} - p_H(w_H), u_H - w_H)_{L^2(\Omega)} \\
&\leq CH\Lambda^{-\frac{1}{2}}\kappa_0^{-\frac{1}{2}}(\|\bar{u}\|_{H^1(\Omega)} + \|y_d\|_{L^2(\Omega)}) \cdot \|w_H - u_H\|_{L^2(\Omega)}.
\end{aligned} \tag{30}
$$

The third term represents a formula for the numerical integration. Using that u_H and w_H are constant on each triangle T_i,

$$
\begin{aligned}
(\hat{p} - \bar{p}, \bar{u}_H - w_H)_{L^2(\Omega)} &= \sum_i \int_{T_i} ((\hat{p}(x) - \bar{p}(x))(\bar{u}_H(x) - w_H(x))\, dx \\
&= \sum_i (\bar{u}_H(x) - w_H(x)) \int_{T_i} (\hat{p}(x) - \bar{p}(x))\, dx \\
&= \sum_i (\bar{u}_H(x) - w_H(x))(\int_{T_i} \bar{p}(x)\, dx - \int_{T_i} \bar{p}(x)\, dx) \\
&= 0.
\end{aligned} \tag{31}
$$

Using (29)-(31) in (28), we get

$$
\begin{aligned}
&(\hat{p} - p_H(\bar{u}_H), \bar{u}_H - w_H)_{L^2(\Omega)} \\
&\leq CH\Lambda^{-\frac{1}{2}}\kappa_0^{-\frac{1}{2}}(\|\bar{u}\|_{H^1(\Omega)} + \|y_d\|_{L^2(\Omega)}) \cdot \|w_H - \bar{u}_H\|_{L^2(\Omega)}.
\end{aligned}
$$

Note that, by the standard finite element interpolation theory, we have

$$
\begin{aligned}
\|\bar{u}_H - \bar{u}\|_{L^2(\Omega)} &\leq \|\bar{u}_H - w_H\|_{L^2(\Omega)} + \|\bar{u} - w_H\|_{L^2(\Omega)} \\
&\leq CH\Lambda^{-\frac{1}{2}}\kappa_0^{-\frac{1}{2}}\nu^{-1}(\|\bar{u}\|_{H^1(\Omega)} + \|y_d\|_{L^2(\Omega)}).
\end{aligned}
$$

By Lemma 2 and Lemma 3, we have

$$
\begin{aligned}
&\|\bar{y}_H - \bar{y}\|_a \\
&\leq \|\bar{y}_H - y_H(w_H)\|_a + \|y_H(w_H) - y(w_H)\|_a + \|y(w_H) - \bar{y}\|_a \\
&\leq CH\Lambda^{-\frac{1}{2}}\kappa_0^{-\frac{1}{2}}\|\bar{u}\|_{H^1(\Omega)} + CH\Lambda^{-\frac{1}{2}}\|\kappa^{-\frac{1}{2}}w_H\|_{L^2(\Omega)} + CH\Lambda^{-\frac{1}{2}}\kappa_0^{-\frac{1}{2}}\|\bar{u}_H\|_{H^1(\Omega)} \\
&\leq CH\Lambda^{-\frac{1}{2}}.
\end{aligned}
$$

Similarly, we have $\|\bar{p}_H - \bar{p}\|_a \leq CH\Lambda^{-\frac{1}{2}}$. This proves Theorem 1.

3 Numerical results

In this section, we will present some numerical tests to validate the convergence of the method. The optimization problems are solved numerically by a primal-dual active set strategy; see, for instance, [5]. The primal-dual active set strategy will be presented here. For this purpose we introduce the active and inactive sets for the solution and define

$$A_+^* = \{x \in \Omega : u^*(x) = b\}, \qquad A_-^* = \{x \in \Omega : u^*(x) = a\},$$

$$\text{and } \ I^* = \{x \in \Omega : a < u^*(x) < b\}.$$

Here and below, the set theoretic definitions are understood in the almost everywhere sense. Given (u_{n-1}, λ_{n-1}), the active sets for the new iterate are chosen according to

$$A_n^+ = \left\{x \in \Omega : u_{n-1}(x) + \frac{\lambda_{n-1}(x)}{c} > b\right\}, \tag{32}$$

$$A_n^- = \left\{x \in \Omega : u_{n-1}(x) + \frac{\lambda_{n-1}(x)}{c} < a\right\}, \tag{33}$$

where $c > 0$. The update strategies for A_n^+ and A_n^- are the key ingredients of the proposed algorithm. The complete algorithm is specified in Algorithm 1.

Algorithm 1 Primal-dual Active Set Strategy.

1: Initialization: Choose u_0 and λ_0, and set $n = 1$.
2: Determine the active sets according to (32)-(33), and set $I_n = \Omega \setminus (A_n^+ \cup A_n^-)$.
3: If $n \geq 2$, $A_n^+ = A_{n-1}^+$, $A_n^- = A_{n-1}^-$, and $I_n = I_{n-1}$, then STOP.
4: Else, find $(y_n, p_n) \in V_H \times V_H$ such that

$$\int_\Omega \kappa \nabla y_n \cdot \nabla v_H \, dx = \int_\Omega u_n v_H \, dx \qquad \forall v_H \in V_H$$

$$\int_\Omega \kappa \nabla p_n \cdot \nabla v_H \, dx = \int_\Omega (y_n - y_d) v_H \, dx \qquad \forall v_H \in V_H$$

where

$$u_n(x) = \begin{cases} b & \text{if } x \in A_n^+, \\ a & \text{if } x \in A_n^-, \\ -\frac{\int_T p_n \, dx}{\nu \int_T 1 \, dx} & \text{if } x \in I_n \cap T, \text{ where } T \in \mathcal{T}_H. \end{cases}$$

5: Set $\lambda_n = -p_n - \nu u_n$, update $n := n + 1$, and goto 2.

In our simulations, we take the medium parameter κ shown in Fig. 2, and the contrast is 10^4. Note that, the state equation is given by

$$\begin{aligned} -\text{div}(\kappa \nabla y) &= u \qquad \text{in } \Omega, \\ y &= 0 \qquad \text{on } \Gamma \end{aligned} \tag{34}$$

Define $u_f(x_1, x_2) = 2\pi x_1 (1 - x_1)^2 \sin(\pi x_2)$. We construct the exact optimal control $\bar{u}$

$$\bar{u}(x_1, x_2) = \begin{cases} a & \text{if } u_f(x_1, x_2) < a, \\ u_f(x_1, x_2) & \text{if } u_f(x_1, x_2) \in [a, b], \\ b & \text{if } u_f(x_1, x_2) > b \end{cases}$$

We also denote the optimal state $\bar{y}$ by solving (34). For the optimal adjoint state $\bar{p}$, we find

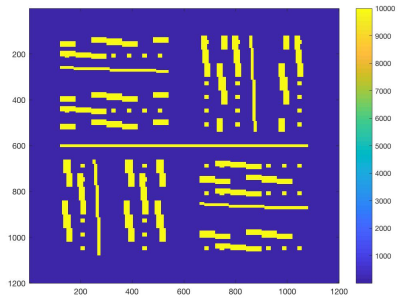

Fig. 2: The high contrast medium κ.

$$\bar{p}(x_1, x_2) = -2\pi\nu x_1(1 - x_1)^2 \sin(\pi x_2)$$

The desired state is given by

$$y_d(x_1, x_2) = \bar{y} + \operatorname{div}(\kappa\nabla\bar{p}).$$

It is easy to see that these functions fulfill the necessary and sufficient first order optimality conditions. Also, we take $c = 2$ and $\nu = 1$.

The solution $\bar{y}$ is calculated by the reference solution using a 200×200 fine mesh. We need 6 iterations to stop the primal-dual active set strategy. Also, if we solve the problem on the fine mesh, we need to solve problems with $2 \times 201 \times 201$ unknowns in one iteration but in our approach, we only need 9600 unknowns even for the finest case $H = \frac{1}{40}$. Fig. 3 and Fig. 4 show the numerical solutions $\bar{u}_H$ for $H = 0.05$ and $H = 0.025$ respectively. Table 1 shows the relative L^2-norm error for $\bar{u} - \bar{u}_H$. The order of the L^2-error is about 1. Table 2 shows the same result with one more number of basis per coarse element.

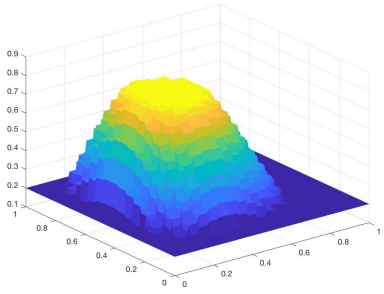

Fig. 3: $\bar{u}_H$ using $H = 0.05$.

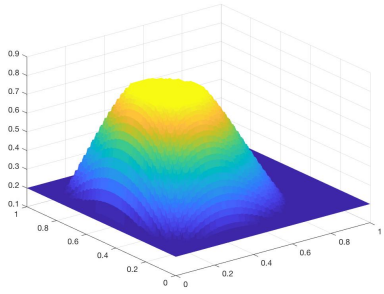

Fig. 4: $\bar{u}_H$ using $H = 0.025$.

Table 1: L^2 error with 3 basis functions per coarse element

number of basis	H	# oversample layers	$\frac{\|\vec{u}_H-\vec{u}\|_{L^2(\Omega)}}{\|\vec{u}\|_{L^2(\Omega)}}$
3	1/5	3	23.8716%
3	1/10	3	9.6973%
3	1/20	4	4.3582%
3	1/40	5	1.6717%

Table 2: L^2 error with 4 basis functions per coarse element

number of basis	H	# oversample layers	$\frac{\|\vec{u}_H-\vec{u}\|_{L^2(\Omega)}}{\|\vec{u}\|_{L^2(\Omega)}}$
4	1/5	3	21.9864%
4	1/10	3	9.6987%
4	1/20	4	4.3463%
4	1/40	5	1.6782%

Acknowledgements The research of Eric Chung is partially supported by the Hong Kong RGC General Research Fund (Project numbers 14304719 and 14302018) and the CUHK Faculty of Science Direct Grant 2020-21.

References

1. Abdulle, A., Weinan, E., Engquist, B. and Vanden-Eijnden, E.: The heterogeneous multiscale method. Acta Numerica **21**, 1–87 (2012)
2. Chung, E., Efendiev, Y. and Hou, T.Y.: Adaptive multiscale model reduction with generalized multiscale finite element methods. Journal of Computational Physics **320**, 69–95 (2016)
3. Chung, E.T., Efendiev, Y. and Leung, W.T.: Constraint energy minimizing generalized multiscale finite element method. Computer Methods in Applied Mechanics and Engineering **339**, 298–319 (2018)
4. Hughes, T.J., Feijoo, G.R., Mazzei, L. and Quincy, J.B.: The variational multiscale method - a paradigm for computational mechanics. Computer methods in applied mechanics and engineering **166**, 3–24 (1998)
5. Kunisch, K. and Rösch, A.: Primal-dual active set strategy for a general class of constrained optimal control problems. SIAM Journal on Optimization, **13**, 321–334 (2002)
6. Malqvist, A. and Peterseim, D.: Numerical homogenization by localized orthogonal decomposition. Society for Industrial and Applied Mathematics (2020)
7. Owhadi, H.: Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games. SIAM Review **59**, 99–149 (2017)
8. Tröltzsch, F.: Optimal control of partial differential equations: theory, methods, and applications. American Mathematical Soc. (2010)