An Efficient Online-Offline Method for Elliptic Homogenization Problems

Yufang Huang[1,2], Pingbing Ming[1,*] and Siqi Song[1]

[1] *LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing, AMSS, Chinese Academy of Sciences, No. 55, Zhong-Guan-Cun East Road, Beijing 100190, China, and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing 100049, China*
[2] *Cornell University, 425 East 61st Street, New York, 10065, USA*

Abstract. We present a new numerical method for solving the elliptic homogenization problem. The main idea is that the missing effective matrix is reconstructed by solving the local least-squares in an offline stage, which shall be served as the input data for the online computation. The accuracy of the proposed method are analyzed with the aid of the refined estimates of the reconstruction operator. Two dimensional and three dimensional numerical tests confirm the efficiency of the proposed method, and illustrate that this online-offline strategy may significantly reduce the cost without loss of accuracy.

1 Introduction

We consider a prototypical elliptic boundary value problem

$$\begin{cases} -\operatorname{div}(a^\varepsilon(x)\nabla u^\varepsilon(x)) = f(x), & x \in D \subset \mathbb{R}^d, \\ u^\varepsilon(x) = 0, & x \in \partial D, \end{cases} \tag{1.1}$$

where ε is a small parameter that signifies explicitly the multiscale nature of the problem. We assume that the coefficient a^ε, which is not necessarily symmetric, belongs to a set $\mathcal{M}(\alpha,\beta,D)$ that is defined by

$$\mathcal{M}(\alpha,\beta,D) := \{\mathcal{B} \in [L^\infty(D)]^{d^2} \,|\, (\mathcal{B}\xi,\xi) \geq \alpha|\xi|^2, |\mathcal{B}(x)\xi| \leq \beta|\xi|,$$
$$\text{for any } \xi \in \mathbb{R}^d \text{ and a.e. } x \in D\}, \tag{1.2}$$

---

*Corresponding author.

where D is a bounded domain in $\mathbb{R}^d$ and $(\cdot,\cdot)$ denotes the inner product on $\mathbb{R}^d$, while $|\cdot|$ is the corresponding norm.

In the sense of H-convergence [53, 54], for every sequence $a^\varepsilon \in \mathcal{M}(\alpha,\beta,D)$ and $f \in H^{-1}(D)$, the sequence u^ε of the solution to (1.1) satisfies

$$\begin{cases} u^\varepsilon \rightharpoonup U_0 & \text{weakly in } H_0^1(D), \\ a^\varepsilon \nabla u^\varepsilon \rightharpoonup \mathcal{A}\nabla u_0 & \text{weakly in } [L^2(D)]^d, \end{cases} \quad \text{as} \quad \varepsilon \to 0, \tag{1.3}$$

where u_0 is the solution of the homogenization problem

$$\begin{cases} -\operatorname{div}(\mathcal{A}(x)\nabla u_0(x)) = f(x), & x \in D, \\ u_0(x) = 0, & x \in \partial D, \end{cases} \tag{1.4}$$

and $\mathcal{A} \in \mathcal{M}(\alpha,\beta,D)$. Here $H_0^1(D), L^2(D)$ and $H^{-1}(D)$ are standard Sobolev space [4].

The quantities of interest for Problem (1.1) and Problem (1.4) are the homogenized solution u_0 over the whole domain and the solution u^ε at certain critical local region. The former stands for the information at the large scale, and the later mimics the information at small scale. There are lots of work devoted to efficiently compute such quantities during the last several decades; see, e.g., [6, 17, 20], among many others. Presently we are interested in the efficient way to compute u_0. A typical way towards this is provided by the heterogeneous multiscale method (HMM) [3, 18], and the FE$^2-$method [40] commonly used in the engineering community is also in the same spirit of HMM. The underlying idea of this approach is to extract $\mathcal{A}$ by solving the cell problems posed on the sampling points of the macroscopic solver. At each point, one needs to solve d cell problems with d the dimensionality. Therefore, the main computational cost comes from solving all these cell problems. The number of the cell problems grows rapidly when higher-order macroscopic solvers are employed. To reduce the cost, certain nonconventional quadrature schemes were proposed in [16] when finite element method is used as the macroscopic solver. The number of the cell problems reduces to one third compared to the standard mid-point quadrature scheme when $\mathbb{P}_2$ Lagrange element is employed as the macroscopic solver. Unfortunately, it does not seem easy to extend such idea to even higher order macroscopic solvers because the quadrature nodes tend to accumulate in the interior of the element [48, 50, 51].

In [35] the authors presented a local least-squares reconstruction of the effective matrix using the solution of the cell problems posed on the vertices of the triangulation, which was dubbed as HMM-LS. The total number of the cell problems equals to the total number of the interior vertices of the triangulation, which is of $\mathcal{O}(h^{-d})$ with h the mesh size of the macroscopic solver. This method achieves higher-order accuracy with almost the same cost of HMM with $\mathbb{P}_1$ Lagrange element as the macroscopic solver [19]. A drawback of this method is that the number of the cell problems is still quite large when mesh refinement is necessary. Moreover, if the adaptive strategy is used in the macroscopic solver, then one has to solve many cell problems around the regions with mesh refinement.

In this work, we propose an offline-online method to compute u_0 efficiently. The main idea is to separate the microscopic solver from the macroscopic solver. In the offline stage, we firstly solve the cell problems posed on a sampling point set and obtain the effective matrix at all these points, then we reconstruct an effective matrix locally by solving the discrete least-squares. The sampling set is constructed from a triangulation of domain, which is usually coarser than the online triangulation. In the online stage, we solve the macroscopic problem with the effective matrix prepared in the offline stage. Such decoupled strategy brings more flexibilities to reduce the number of the cell problems, we may either refine the offline triangulation mesh or increase the reconstruction order, which is guided by the a priori error estimate. The offline computation bears certain similarity with h-p finite element method [7, 47]. Moreover, the offline computation is no longer linked to the macroscopic solvers, which is particularly attractive to higher order macroscopic solver and three dimensional multiscale problems. With the aid of the theoretical results proved in [35, 36], we study the accuracy and the stability of the reconstruction procedure, which is crucial to prove the optimal error estimate of the proposed method. As illustrated by the numerical tests in § 4, the offline-online method converges with optimal order while the cost is smaller than both HMM and the HMM-LS method.

The reduced basis HMM proposed in [1,2] also employed the offline-online idea. The difference between reduced basis HMM and our method lies in the following points: Firstly they employed reduced basis idea in the offline stage while we construct an independent triangulation upon which the cell problems are solved. Secondly they used the empirical interpolation method [8] while we resort to a local least-squares to reconstruct the effective matrix. Finally, a thorough analysis of the least-squares reconstruction is conducted in our work, which concerns the approximation accuracy and the stability of the reconstruction, such rigorous theoretical results put the method on a firm footing. In addition, the analysis of the least-squares reconstruction is of independent interest for other problems such as the construction of the optimal polynomial admissible meshes [13, 45], the discrete norm for polynomials [46], the approximation of the Fekete points [11] and the discontinuous Galerkin method based on patch reconstruction [36,37].

The rest of the paper is organized as follows. In § 2, we introduce the offline-online method that is based on a local discrete least-squares reconstruction. We derive the optimal error estimate in § 3, in particular, we prove the discrete least-squares is stable with respect to small perturbation. In § 4, we report numerical examples in two and three dimensions, the coefficient a^ε may be locally periodic, quasi-periodic and random checker-board. To demonstrate the efficiency of the offline-online method, we compare it with HMM and the HMM-LS method. In addition, we solve a problem posed on L-shape domain with nonsmooth solution. The conclusions are drawn in the last section.

Throughout this paper, we shall use the standard notations for the Sobolev space, norms and semi-norms, cf., [4], e.g.,

$$\|v\|_{H^1(D)} := \|v\|_{L^2(D)} + \|\nabla v\|_{L^2(D)}, \qquad |v|_{W^{m,p}(D)} := \sum_{|\alpha|=m} \|\nabla^\alpha v\|_{L^p(D)}.$$

For any measurable set E, we define the mean of an integrable function g over E as

$$\langle g \rangle_E := \frac{1}{|E|} \int_E g(x) \, dx.$$

We shall also use the discrete ℓ_p norm for any $x \in \mathbb{R}^d$ as

$$\|x\|_{\ell_p} := \begin{cases} \left(\sum_{i=1}^d |x_i|^p \right)^{1/p} & 1 \leq p < \infty, \\ \max_{1 \leq i \leq d} |x_i| & p = \infty. \end{cases}$$

Throughout the paper the generic constant C may be different from line to line, while it is independent of ε and the mesh size parameters h, H.

2   The Offline-online Method

The macroscopic solver is chosen as the standard $\mathbb{P}_l$ Lagrange finite element, which is defined as the set of polynomials with degree less than l for the sum of all variables [14]. The finite element space is denoted by V_h corresponding to the triangulation τ_h with mesh size h that is the maximum of the element size h_τ for all elements $\tau \in \tau_h$, where h_τ is the diameter of τ. We assume that all the elements τ in τ_h satisfy the shape-regular condition in the sense of Ciarlet and Raviart [14], i.e., there exists a constant σ_0 such that $h_\tau / \rho_\tau \leq \sigma_0$, where ρ_τ is the diameter of the smallest ball inscribed into τ, and σ is the so-called chunkiness parameter [12].

The method consists of offline part and online part. In the offline part, we approximate the effective matrix $\mathcal{A}$ as follows.

Offline We firstly construct a sampling triangulation $\mathcal{T}_H$ with mesh size H for domain D. For simplicity, we assume that $\mathcal{T}_H$ consists of simplices, and $\mathcal{T}_H$ is assumed to be shape-regular with the chunkiness parameter σ. On each element $K \in \mathcal{T}_H$, the approximation effective matrix A_H is reconstructed by solving a least-squares: for $i, j = 1, \cdots, d$,

$$(A_H)_{ij} = \arg \min_{p \in \mathbb{P}_m(S(K))} \sum_{x_K \in \mathcal{I}(K)} \left| (A_H(x_K))_{ij} - p(x_K) \right|^2, \tag{2.1}$$

Here $\mathcal{I}(K)$ is the set of all sampling points that belong to $S(K)$, where $S(K)$ is a patch of elements around K, which usually includes K. Its precise definition will be given later on. We refer to Fig. 1 for an example of such $S(K)$.

At each sampling point x_K, the effective matrix $A_H(x_K)$ is defined by averaging the flux arising from the cell problems:
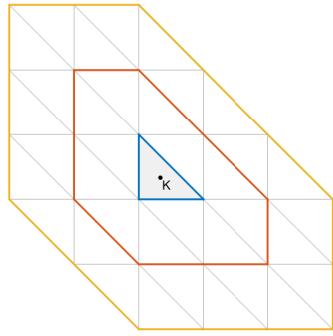
$$A_H(x_K) = \left(\langle a^\varepsilon \nabla v_1^\epsilon \rangle_{I_\delta}, \cdots, \langle a^\epsilon \nabla v_d^\epsilon \rangle_{I_\delta} \right), \tag{2.2}$$

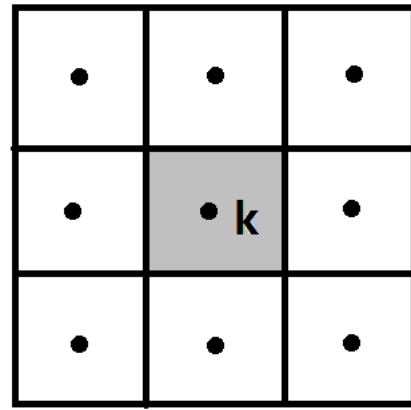where the cell $I_\delta(x_K) := x_K + \delta Y$ with $Y := (-1/2, 1/2)^d$ and δ the cell size. Here for $i = 1, \cdots, d$, v_i^ε satisfies

$$\begin{cases} -\nabla \cdot (a^\varepsilon \nabla v_i^\varepsilon) = 0 & \text{in} \quad I_\delta(x_K), \\ \qquad\qquad v_i^\varepsilon = x_i & \text{on} \quad \partial I_\delta(x_K). \end{cases} \tag{2.3}$$

Online Given A_H, we find $u_h \in V_h$ such that

$$\int_D A_H(x) \nabla u_h \cdot \nabla v \, dx = \int_D f(x) v(x) \, dx \quad \text{for all} \quad v \in V_h. \tag{2.4}$$

(a). Example of $S_2(K)$ constructed by including all the Moore neighbors.

(b). Example of $S(K)$ and the set $\mathcal{I}(K)$ consists of the black dots.

Figure 1: Examples of the element patches and the sample set.

Remark 2.1. The element $K \in \mathcal{T}_H$ may not be a simplex, which may be polygons or polytopes, the corresponding shape-regular condition and other mesh conditions may be found in [36]. Under these mesh conditions, the properties of the reconstruction are still valid.

In what follows, we supplement some details in the algorithm. The first thing is the construction of the element patch $S(K)$ for any element $K \in \mathcal{T}_H$. We start from assigning a threshold value N_{lowest} that is used to control the size of $S(K)$. There are several different ways to find $S(K)$. One way is to define $S(K)$ in an recursive way as in [35]: For any $t \in \mathbb{N}$, we let

$$S_0(K) := K, \quad S_t(K) = \{ K \in \mathcal{T}_H \mid \overline{K} \cap \overline{S_{t-1}(K)} \neq \varnothing \}. \tag{2.5}$$

Once $\#S_t(K) \geq N_{\text{lowest}}$, we stop the construction and let $S(K) = S_t(K)$. This means that we add the Moore neighbors [52] to $S(K)$ in a recursive way. Another way is using the Von

Neumann neighbor [52], i.e., we include the adjacent edge-neighboring elements into the element patch instead of the Moore neighbor. We refer to [36] and [37, Appendix A] for a detailed description for such construction, while $S(K)$ in all the tests in § 4 are defined as in (2.5). We denote by $\mathcal{I}_t(K)$ the set containing all the sampling points that belong to $S_t(K)$. In all the tests in § 4, we use the barycentric of each element K as the sampling point. There are also other choices for construction $\mathcal{I}_t(K)$. However, the vertices are not preferred because the communication cost is quite high for three dimensional problems, though this is a good choice for two dimensional problems; cf., [35]. In what follows, we may drop the subscripts t in $S_t(K)$ and $\mathcal{I}_t(K)$ when there is no confusion may occur.

We remark that there are many other variants for the definitions of the cell problems and the effective matrix in the literatures; see e.g., [24,59]. The periodic cell problem will be discussed in § 4. As illustrated by the examples in § 4, the overhead caused by the local least-squares is small compared to the computational cost of solving all the cell problems, though the number of the local least-squares is also the same with the cell problems.

3   Convergence of the Method

To study the convergence of the method, we define

$$e(\text{MOD}) := \max_{x \in D} \| \mathcal{A}(x) - A_H(x) \|_F,$$

where $\| B \|_F$ is the Frobenius norm of a $d-$by$-d$ matrix B.

Similar to [35, Lemma 3.1, Lemma 3.2], the following lemma gives the error estimates of the proposed method.

Lemma 3.1. *Let u_0 be the homogenized solution with the homogenized matrix in $\mathcal{M}(\alpha,\beta,D)$. If $e(\text{MOD}) < \alpha$, then there exists a unique solution u_h satisfying* (2.4).

If $e(\text{MOD}) < \nu\alpha/(1+\nu)$ for any $\nu > 0$, then

$$\| \nabla(u_0 - u_h) \|_{L^2(D)} \le \frac{\beta}{\alpha} \inf_{v \in V_h} \| \nabla(u_0 - v) \|_{L^2(D)} + \frac{(1+\nu)C_p}{\alpha^2} \| f \|_{H^{-1}(D)} e(\text{MOD}), \qquad (3.1)$$

where C_p appears in the discrete Poincaré inequality that only dependends on D:

$$\| v \|_{H^1(D)} \le C_p \| \nabla v \|_{L^2(D)} \qquad \text{for all} \quad v \in V_h.$$

Moreover, there exists C depending only on α, β, ν, C_p and $\| f \|_{H^{-1}(D)}$ such that

$$\| u_0 - u_h \|_{L^2(D)} \le C \left(\inf_{v \in V_h} \| \nabla(u_0 - v) \|_{L^2(D)} \sup_{\|g\|_{L^2(D)}=1} \inf_{\chi \in V_h} \| \nabla(\phi_g - \chi) \|_{L^2(D)} + e(\text{MOD}) \right),$$

where $\phi_g \in H_0^1(D)$ is the unique solution of the problem:

$$\int_D \mathcal{A}(x) \nabla v \cdot \nabla \phi_g \, dx = \int_D g(x) v(x) \, dx \qquad \text{for all} \quad v \in H_0^1(D). \qquad (3.2)$$

The proof of the above lemma follows the same line of [19, Theorem 1.1] except the explicit constants in (3.1), we omit the proof.

To further elucidate the error structure of the method, we define the reconstruction operator for any piecewise constant function v defined on $\mathcal{T}_H$ as follows. Let $\mathcal{R}_K v$ be the solution of the least-squares

$$\mathcal{R}_K v = \arg \min_{p \in \mathbb{P}_m(S(K))} \sum_{x_K \in \mathcal{I}(K)} |v(x_K) - p(x_K)|^2. \qquad (3.3)$$

We imbed $\mathcal{R}_K$ into a global operator as $\mathcal{R}|_K = \mathcal{R}_K$.

Given the reconstruction operator $\mathcal{R}$, we decompose $e(\text{MOD})$ as

$$\begin{aligned} \mathcal{A}(x) - A_H(x) &= \mathcal{A}(x) - \mathcal{R}\mathcal{A}(x) + \mathcal{R}\mathcal{A}(x) - A_H(x) \\ &= \mathcal{A}(x) - \mathcal{R}\mathcal{A}(x) + \mathcal{R}\left(\mathcal{A}(x) - \overline{A}_H\right), \end{aligned} \qquad (3.4)$$

where $\overline{A}_H(x)$ is a piecewise matrix posed on $\mathcal{T}_H$ that is defined by

$$\overline{A}_H|_K := A_H(x_K)$$

with $A_H(x_K)$ given by (2.2) and (2.3).

The first term in the right-hand side of (3.4) is the reconstruction error while the second one is the so-called estimation error. To quantify these two terms, we need some properties of the reconstruction operator, which is by now well-understood by virtue of [35] and [36]. To state such properties, we make two assumptions on $S(K)$ and $\mathcal{I}(K)$.

Assumption A For every $K \in \mathcal{T}_H$, there exist constants R and r that are independent of K such that $B_r \subset S(K) \subset B_R$ with $R \geq 2r$, and $S(K)$ is star-shaped with respect to B_r, where B_ρ is a disk with radius ρ.

This assumption concerns the geometry of $S(K)$, which is crucial for the uniform boundedness of $\mathcal{R}$. The motivation for this assumption lies in the following Markov inequality [38]:

$$\|\nabla g\|_{L^\infty(S(K))} \leq \frac{4m^2 R}{r^2} \|g\|_{L^\infty(S(K))} \qquad \text{for all} \quad g \in \mathbb{P}_m(S(K)). \qquad (3.5)$$

Here $\|\nabla g\|_{L^\infty(S(K))} := \max_{x \in \overline{S(K)}} \|\nabla g(x)\|_{\ell_2}$. This inequality is proved in [36, Lemma 5], which is a combination of [57, Proposition 11.6] and the fact that $S(K)$ satisfies the uniform interior cone condition [4], which is a direct consequence of **Assumption A**.

In [35], the authors make the following assumption on $S(K)$.

Assumption A′ $S(K)$ is a bounded convex domain and there exists R such that $S(K) \subset B_R$.

By [26, Lemma 1.2.2.2 and Corollary 1.2.2.3], **Assumption A′** implies **Assumption A**. Under **Assumption A′**, WILHELMSEN [58] proved the following Markov inequality:

$$\|\nabla g\|_{L^\infty(S(K))} \leq \frac{4m^2}{w(K)} \|g\|_{L^\infty(S(K))} \qquad \text{for all} \quad g \in \mathbb{P}_m(S(K)), \qquad (3.6)$$

where $w(K)$ is the width of $S(K)$, which is the minimum distance between parallel supporting hyperplanes of $S(K)$.

The above Markov inequalities may be viewed as a type of inverse inequality. By the classical inverse inequality for p-finite element method [47, Theorem 4.76] and a simple scaling argument, we may conclude that there exists C independent of the diameter of $S(K)$ but depends on the shape of $S(K)$ such that for all $g \in \mathbb{P}_m(S(K))$,

$$\|\nabla g\|_{L^\infty(S(K))} \leq \frac{Cm^2}{\operatorname{diam}S(K)}\|g\|_{L^\infty(S(K))}. \tag{3.7}$$

The index 2 is sharp if $S(K)$ is a locally Lipschitz domain. However, for a cuspidal domain, Kroó and Szabodos [34] proved that if $S(K)$ is a Lipγ-domain with $0 < \gamma < 1^\dagger$, then there exists C depends on the shape of $S(K)$ such that

$$\|\nabla g\|_{L^\infty(S(K))} \leq \frac{Cm^{2/\gamma}}{\operatorname{diam}S(K)}\|g\|_{L^\infty(S(K))}, \tag{3.8}$$

where the index $2/\gamma$ is sharp. This would require a larger patch to ensure the reconstruction accuracy, and the reconstruction is less stable than the patch satisfying either **Assumption A** or **Assumption A'**. Moreover, the constant C in (3.7) is only known for $S(K)$ with special shape in the literature, e.g., $S(K)$ is an interval [46], and $S(K)$ is a simplex [33, 56]. All the prefactors are important for us to derive the realistic conditions that ensure the uniform boundedness of $\Lambda(m, \mathcal{I}(K))$; cf., Lemma 3.3.

On the other hand, The authors in [35] derived an explicit expression of C that depends on the recursion depth t and the chunkiness parameter σ under **Assumption A'**. It seems that the convexity of the patch is rather restrictive in implementation, particularly for an L-shape domain. **Assumption A** is less restrictive and is easy to check in practice.

Assumption B For any $K \in \mathcal{T}_H$ and $p \in \mathbb{P}_m(S(K))$,

$$p|_{\mathcal{I}(K)} = 0 \qquad \text{implies} \qquad p|_{S(K)} \equiv 0.$$

This assumption concerns the cardinality of the sampling set $\mathcal{I}(K)$, which gives the uniqueness and hereby the existence of the solution of the discrete least-squares (2.1). **Assumption B** requires that the cardinality of $\mathcal{I}(K)$ is at least $\binom{m+d}{d}$ to ensure the unisolvence of the discrete least-squares. A quantitative version of this assumption is

$$\Lambda(m, \mathcal{I}(K)) < \infty$$

with

$$\Lambda(m, \mathcal{I}(K)) := \max_{p \in \mathbb{P}_m(S(K))} \frac{\|p\|_{L^\infty(S(K))}}{\|p|_{\mathcal{I}(K)}\|_{\ell_\infty}}.$$

---

†A typical Lipγ-domain is an ℓ_γ−ball, i.e.,$\{x \in \mathbb{R}^d \mid |x_1|^\gamma + \cdots + |x_d|^\gamma \leq 1\}$ with $0 < \gamma < 1$.

In practical implementation, the positions of the sampling nodes may be slightly perturbed due to the measure error or certain uncertainties [15]. A natural question arises whether the reconstruction is robust with respect to the uncertainties. We shall prove that the the reconstruction is stable with respect to small perturbation.

Next we prove some properties of the reconstruction operator $\mathcal{R}_K$.

Lemma 3.2. *If **Assumption B** holds, then there exists a unique solution of (3.3) for any $K \in \mathcal{T}_H$. Moreover $\mathcal{R}_K$ satisfies*

$$\mathcal{R}_K g = g \qquad \text{for all} \quad g \in \mathbb{P}_m(S(K)).$$

The stability result is valid for any $K \in \mathcal{T}_H$ and $g \in C^0(S(K))$, and

$$\|\mathcal{R}_K g\|_{L^\infty(K)} \leq \Lambda(m, \mathcal{I}(K)) \sqrt{\#\mathcal{I}(K)} \, \|g|_{\mathcal{I}(K)}\|_{\ell_\infty}. \tag{3.9}$$

The quasi-optimality approximation property is valid in the sense that

$$\|g - \mathcal{R}_K g\|_{L^\infty(K)} \leq (1 + \Lambda(m, \mathcal{I}(K))) \sqrt{\#\mathcal{I}(K)} \inf_{p \in \mathbb{P}_m(S(K))} \|g - p\|_{L^\infty(S(K))}. \tag{3.10}$$

*If **Assumption A** and **Assumption B** are valid, then for any $\delta \in (0,1)$, there exists*

$$\epsilon = \frac{\delta r^2}{4\Lambda(m, \mathcal{I}(K)) m^2 R} \tag{3.11}$$

such that for the perturbed sampling set $\widetilde{I}(K) \subset \mathcal{I}(K) + B_\epsilon(0)$, there exists a unique $\widetilde{R}_K g \in \mathbb{P}_m(S(K))$ satisfying

$$\|\widetilde{R}_K g\|_{L^\infty(K)} \leq \frac{\Lambda(m, \mathcal{I}(K))}{1 - \delta} \sqrt{\#\widetilde{I}(K)} \, \|g|_{\widetilde{I}(K)}\|_{\ell_\infty}, \tag{3.12}$$

where $B_\varepsilon(0)$ is a ball centered at 0 with radius ε.

*If **Assumption A'** and **Assumption B** are valid, then the perturbation result (3.12) remains true with*

$$\varepsilon = \frac{\delta w(K)}{4\Lambda(m, \mathcal{I}(K)) m^2}.$$

The above lemma except the perturbation estimate (3.12) is proved in [35, Theorem 3.3], which is crucial for the accuracy of the reconstruction operator $\mathcal{R}$, more refined estimates on the accuracy of $\mathcal{R}$ may be found in [36, Lemma 4]. The perturbation estimate (3.12) shows that the set of the sampling nodes $\mathcal{I}(K)$ perturbed a little bit remains a norming set with a slightly bigger upper bound, i.e., for all $\delta \in (0,1)$,

$$\Lambda(m, \widetilde{I}(K)) = \frac{\Lambda(m, \mathcal{I}(K))}{1 - \delta}.$$

This perturbation estimate has been encapsulated in an abstract form in [36, Lemma 2], while there is no proof for the perturbed reconstruction operator $\widetilde{R}$.

Proof. For any $p \in \mathbb{P}_m(S(K))$, we let $|p(x^*)| = \|p|_{\mathcal{I}(K)}\|_{\ell_\infty}$, then for any $y \in \widetilde{I}(K) \subset \mathcal{I}(K) + B_\epsilon(0)$ with ϵ given by (3.11), by Taylor's expansion, we obtain

$$|p(x^*)| \leq |p(y)| + \epsilon \|\nabla p\|_{L^\infty(S(K))}.$$

By **Assumption A**, the Markov inequality (3.5) is valid, and we obtain

$$|p(x^*)| \leq |p(y)| + \frac{4m^2 R\epsilon}{r^2} \|p\|_{L^\infty(S(K))}.$$

Using **Assumption B** and the above two inequalities, we obtain

$$
\begin{aligned}
\|p\|_{L^\infty(S(K))} &\leq \Lambda(m, \mathcal{I}(K)) \|p|_{\mathcal{I}(K)}\|_{\ell_\infty} = \Lambda(m, \mathcal{I}(K)) |p(x^*)| \\
&\leq \Lambda(m, \mathcal{I}(K)) |p(y)| + \Lambda(m, \mathcal{I}(K)) \frac{4m^2 R}{r^2} \epsilon \|p\|_{L^\infty(S(K))} \\
&\leq \Lambda(m, \mathcal{I}(K)) \|p|_{\widetilde{I}(K)}\|_{\ell_\infty} + \delta \|p\|_{L^\infty(S(K))},
\end{aligned}
$$

which immediately implies

$$\|p\|_{L^\infty(S(K))} \leq \frac{\Lambda(m, \mathcal{I}(K))}{1 - \delta} \|p|_{\widetilde{I}(K)}\|_{\ell_\infty} \qquad \text{for all} \quad \delta \in (0,1). \tag{3.13}$$

Applying the above perturbation estimate to $\widetilde{R}_K g$, we obtain

$$\|\widetilde{R}_K g\|_{L^\infty(S(K))} \leq \frac{\Lambda(m, \mathcal{I}(K))}{1 - \delta} \|\widetilde{R}_K g|_{\widetilde{I}(K)}\|_{\ell_\infty}.$$

This also gives the existence and uniqueness of $\widetilde{R}_K g$.

By the definition of $\widetilde{R}_K g$, we obtain

$$\|\widetilde{R}_K g|_{\widetilde{I}(K)}\|_{\ell_\infty}^2 \leq \|\widetilde{R}_K g|_{\widetilde{I}(K)}\|_{\ell_2}^2 \leq \|g|_{\widetilde{I}(K)}\|_{\ell_2}^2 \leq \#\widetilde{I}(K) \|g|_{\widetilde{I}(K)}\|_{\ell_\infty}^2.$$

A combination of the above two inequalities and the fact that $\#\widetilde{I}(K) = \#\mathcal{I}(K)$ give (3.12).

If **Assumption A'** and **Assumption B** are valid, then we follow exactly the same line that leads to (3.12) except that we use the Markov inequality (3.6) for a convex patch $S(K)$. □

The following lemma ensures the uniform boundedness of $\Lambda(m, \mathcal{I}(K))$.

Lemma 3.3. *If **Assumption A** holds, then for any $\varepsilon > 0$, if $r > m\sqrt{2RH_K(1 + 1/\varepsilon)}$, then we may take*

$$\Lambda(m, \mathcal{I}(K)) = 1 + \varepsilon. \tag{3.14}$$

Moreover, if $r > 2m\sqrt{RH_K}$, then we may take $\Lambda(m, \mathcal{I}(K)) = 2$.

*If **Assumption A'** holds, then for any $\varepsilon > 0$, if $w(K) > 2m^2 H_K(1 + 1/\varepsilon)$, then we may take*

$$\Lambda(m, \mathcal{I}(K)) = 1 + \varepsilon. \tag{3.15}$$

The first statement is proved in [36, Lemma 5], we only prove the second statement (3.15), which improves [35, Lemma 3.5].

Proof. Let $x^* \in \overline{S(K)}$ such that $|p(x^*)| = \max_{x \in \overline{S(K)}} |p(x)|$, and $x_\ell \in \mathcal{I}(K)$ such that $|x_\ell - x^*| = \min_{y \in \mathcal{I}(K)} |y - x^*|$. Then

$$|x_\ell - x^*| \leq H_K / 2.$$

By Taylor's expansion, we have

$$p(x_\ell) = p(x^*) + (x_\ell - x^*) \cdot \nabla p(\xi_x),$$

where ξ_x lies on the line with endpoints x^* and x_ℓ. This gives

$$|p(x_\ell)| \leq |p(x^*)| + \frac{H_K}{2} \|\nabla p\|_{L^\infty(S(K))},$$

Using the Markov inequality (3.6), we immediately have

$$\|p\|_{L^\infty(S(K))} \leq \|p|_{\mathcal{I}(K)}\|_{\ell_\infty} + \frac{2m^2 H_K}{w(K)} \|p\|_{L^\infty(S(K))}.$$

This implies (3.15). □

If **Assumption A** holds, then we usually have $R \simeq t H_K$, the estimate (3.14) suggests that $r \simeq m\sqrt{t} H_K$ implies the uniform boundedness of $\Lambda(m, \mathcal{I}(K))$. This means that $S(K)$ cannot be too narrow in certain directions. If **Assumption A'** holds, then the estimate (3.15) shows that $w(K) \simeq m^2 H_K$, which immediately implies $t \simeq m^2$. Both conditions show that a relative large patch is required for the reconstruction. Furthermore, if $S(K)$ is a cuspidal domain, then we may use (3.8) to prove that if

$$\frac{\operatorname{diam} S(K)}{H_K} \geq c_0 m^\gamma (1 + 1/\varepsilon)$$

with a constant c_0 depending on the shape of $S(K)$, then the bound (3.14) is also valid. This condition indicates that $S(K)$ has a recursive depth $t \simeq m^\gamma$ with $\gamma > 2$. An even larger patch is required to ensure the stability. This may occur for domain D with complicated boundary or rough boundary.

It remains to find an upper bound for $\#\mathcal{I}(K)$. This is a direct consequence of **Assumption A** or **Assumption A'** and the shape regularity of $\mathcal{T}_H$.

Lemma 3.4. *If $\mathcal{T}_H$ is shape-regular and **Assumption A** or **Assumption A'** is valid, then*

$$\#\mathcal{I}(K) \leq (\sigma R / H_K)^d. \tag{3.16}$$

Proof. For any element $K \in \mathcal{T}_H$, using **Assumption A** or **Assumption A'** and note that there is only one sample point inside each element K, we obtain

$$\# \mathcal{I}(K)|K| \leq \text{Vol } B_d(R),$$

where Vol $B_d(R)$ stands for the volume of a d-dimensional ball with radius R.

Using the shape regularity of $\mathcal{T}_H$, we obtain

$$|K| \geq \text{Vol } B_d(\rho_K) \geq \sigma^{-d} \text{Vol } B_d(h_K).$$

A combination of the above two inequalities yields (3.16).   □

It is clear that the upper bound (3.16) is independent of the way for construction $S(K)$. For the two ways based on either the Moore neighbor or the von Neumann neighbor, we have $\# \mathcal{I}(K) \simeq t^d$ with t the recursion depth, which is consistent with the corresponding upper bound proved in [35, Lemma 3.4] and [36, Lemma 6], in which the two-dimensional problem has been dealt with. Both Lemma 3.3 and Lemma 3.4 require that $\# \mathcal{I}(K)$ should be quite large, equivalently, $S(K)$ is large, so that the uniform boundedness of $\Lambda(m, \mathcal{I}(K))$ is valid. In numerical tests below, we observe that the method still works quite well even when $\# \mathcal{I}(K)$ is far less than the theoretical threshold. We refer to [35] for a list of the size of $S(K)$ and the upper bound of $\max_{K \in \mathcal{T}_H} \Lambda(m, \mathcal{I}(K))$.

Based on the above three lemmas, we are ready to estimate (3.4).

Lemma 3.5. *If **Assumption A** or **Assumption A'** and **Assumption B** are valid and the effective matrix $\mathcal{A}_{ij} \in C^{m+1}(D)$, then there exists C that depends on $\|\mathcal{A}_{ij}\|_{C^{m+1}(D)}, m, R, r, \gamma$ and t but independent of H.*

$$e(MOD) \leq C \left(H^{m+1} + e_1(MOD) \right), \tag{3.17}$$

where $e_1(MOD) := \max_{x \in \mathcal{I}(K), K \in \mathcal{T}_H} \|(\mathcal{A} - A_H)(x)\|_F$.

It is worthwhile to mention that $e_1(MOD)$ is the so-called estimating error in HMM [18]. There are many works devoted to bounding $e_1(MOD)$ and developing new algorithms to improve the estimates; see; e.g., [5,10,19,23,24,43,59] and the references therein.

Proof. We start from the decomposition (3.4). For each $K \in \mathcal{T}_H$, using **Assumption A** and **Assumption B** or **Assumption A'** and **Assumption B**, Lemma 3.3, Lemma 3.4 and (3.10), we obtain, on each element $K \in \mathcal{T}_H$, there exists C depends on t, σ, r, R, d and $\|\mathcal{A}_{ij}\|_{C^{m+1}(D)}$ such that

$$\|\mathcal{A}_{ij} - (\mathcal{R}_K \mathcal{A})_{ij}\|_{L^\infty(K)} \leq C \inf_{p \in \mathbb{P}_m(S(K))} \|\mathcal{A}_{ij}(x) - p\|_{L^\infty(S(K))} \leq CH^{m+1}.$$

Summing up all $i, j = 1, \cdots, d$ and $K \in \mathcal{T}_H$ we obtain

$$\max_{x \in D} \|\mathcal{A}(x) - \mathcal{R}\mathcal{A}(x)\|_F \leq CH^{m+1}. \tag{3.18}$$

Using (3.9), we obtain, for each $K \in \mathcal{T}_H$,

$$\| \mathcal{R}_K \left(\mathcal{A}_{ij} - (\overline{A}_H)_{ij} \right) \|_{L^\infty(K)} \leq \Lambda(m, \mathcal{I}(K)) \sqrt{\# \mathcal{I}(K)} \| \left(\mathcal{A}_{ij} - (A_H)_{ij} \right) |_{\mathcal{I}(K)} \|_{\ell_\infty}.$$

Summing up all $i, j = 1, \cdots, d$, we obtain

$$\max_{x \in K} \| \mathcal{R}_K (\mathcal{A} - A_H)(x) \|_F^2 \leq \Lambda^2(m, \mathcal{I}(K)) \# \mathcal{I}(K) \max_{x \in \mathcal{I}(K)} \| \mathcal{A}(x) - A_H(x) \|_F^2.$$

Using **Assumption A** and **Assumption B** or **Assumption A′** and **Assumption B**, Lemma 3.3 and Lemma 3.4, there exists C depends on R, r, t and σ such that

$$\max_{x \in D} \| \mathcal{R}(\mathcal{A} - A_H) \|_F \leq C e_1(\text{MOD}),$$

which together with (3.18) gives (3.17) and finishes the proof.   □

Substituting the estimate (3.17) into Lemma 3.1, we obtain the main result of this part.

Theorem 3.1. *Let u_0 be the homogenized solution with $\| u_0 \|_{H^{l+1}(D)} < \infty$. If **Assumption A** or **Assumption A′** and **Assumption B** hold, and $e_1(\text{MOD}) < \alpha/4$, then there exists a unique solution u_h of Problem* (2.4) *that satisfying*

$$\| \nabla(u_0 - u_h) \|_{L^2(D)} \leq C \left(h^l + H^{m+1} + e_1(\text{MOD}) \right). \tag{3.19}$$

Moreover, if the solution of the auxiliary problem (3.2) *admits a unique solution ϕ_g satisfying the regularity estimate $\| \phi_g \|_{H^2(D)} \leq C \| g \|_{L^2(D)}$, then*

$$\| u_0 - u_h \|_{L^2(D)} \leq C \left(h^{l+1} + H^{m+1} + e_1(\text{MOD}) \right). \tag{3.20}$$

In view of the above error estimates (3.19) and (3.20), we may have $H \simeq h^{l/(m+1)}$ or $H \simeq h^{(l+1)/(m+1)}$. For a fixed h, one may increase m to decrease the cost in the offline stage. This suggests that higher order reconstruction is preferred to save cost while without loss of accuracy, which is confirmed by the tests in the next section.

The error estimate for the HMM-LS method in [35] reads as

$$\begin{cases} \| \nabla(u_0 - \widetilde{u}_h) \|_{L^2(D)} \leq C \left(h^l + h^{m+1} + e_1(\text{MOD}) \right), \\ \| u_0 - \widetilde{u}_h \|_{L^2(D)} \leq C \left(h^{l+1} + h^{m+1} + e_1(\text{MOD}) \right), \end{cases} \tag{3.21}$$

where $\widetilde{u}_h$ is the solution of HMM-LS method. For a fixed h, the above error estimate indicates that $m \simeq l$ to balance the error provided that $e_1(\text{MOD})$ is sufficiently small, the number of the cell problem in this case is the total number of the vertices of the online triangulation τ_h. We shall compare these two methods in the next part.

4 Numerical Results

In this section, we report a few numerical examples to test the accuracy and efficiency of the proposed offline-online method. The examples include the problems with locally periodic coefficients, with quasi-periodic coefficients and with random coefficients. We also test problems posed on L-shape domain in two dimension and three dimension for which the solutions are usually nonsmooth.

In all the tests, the offline triangulations $\mathcal{T}_H$ are uniform partitions of D, and we list the threshold number (N_{lowest} is defined in § 2) of the sampling points for different orders of reconstruction in Table 1, the number $\#\mathcal{I}(K)$ should be slightly larger than the number list in the table when K abut the boundary of the domain. For $d=3$, we only report $m=3$ because this is the only case we test in this part, the examples for the lower order reconstruction may be found in [28].

Table 1: The number of the threshold number for the sampling point in reconstruction.

	$m=1$	$m=2$	$m=3$
$d=2$	5	7	13
$d=3$	NONE	NONE	27

Noted that the theory covers the case when $S(K)$ consists of polygons or polytopes, we refer to [36] for the demonstration of such patches. It is worth pointing out that the nonuniform $\mathcal{T}_H$ is useful when the shape of D is very complicate. In that case, the sampling points near the boundary has to be dense to ensure the accuracy of the reconstruction. We refer to [28] for the examples of nonuniform patch, which may further reduce the number of the cell problems and makes $e(\text{MOD})$ more evenly distributed over the whole domain, which is useful for adaptivity.

The finite element solvers except Example 4.3 are carried on FreeFem++ toolbox [27][‡], and the test for Example 4.3 is performed in a parallel hierarchical grid platform (PHG) [60][§]. The error of the method is calculated by the relative error measured in H^1 norm and L^2 norm:

$$\frac{\|\nabla(u_0-u_h)\|_{L^2(D)}}{\|\nabla u_0\|_{L^2(D)}} \quad \text{and} \quad \frac{\|u_0-u_h\|_{L^2(D)}}{\|u_0\|_{L^2(D)}}.$$

The *exact* homogenized solution u_0 is generated by discretization (1.4) with $\mathbb{P}_2$ Lagrange element over a very refined 500×500 mesh in most cases unless otherwise stated. In all the tests we set $\varepsilon=10^{-6}$.

4.1 Problems with local periodic coefficient in two dimension

We consider an example with locally periodical coefficient in $d=2$, which is taken from [41].

[‡]https://freefem.org/
[§]http://lsec.cc.ac.cn/phg/

Example 4.1.

$$\begin{cases} -\operatorname{div}(a^{\varepsilon}\nabla u^{\varepsilon}(x)) = f(x), & \text{in} \quad D, \\ \qquad\qquad u^{\varepsilon}(x) = 0, & \text{on} \quad \partial D, \end{cases}$$

where $D = (0,1)^2$, $f(x) = 1$ and

$$a^{\varepsilon}(x) = \frac{(2.5 + 1.5\sin(2\pi x_1))(2.5 + 1.5\cos(2\pi x_2))}{(2.5 + 1.5\sin(2\pi x_1/\varepsilon))(2.5 + 1.5\sin(2\pi x_2/\varepsilon))} 1_{2\times 2}.$$

where $1_{2\times 2}$ denotes the $2-\text{by}-2$ identity matrix. The homogenization problem is

$$\begin{cases} -\operatorname{div}(\mathcal{A}(x)\nabla u_0(x)) = f(x), & \text{in} \quad D, \\ \qquad\qquad u_0(x) = 0, & \text{on} \quad \partial D. \end{cases} \tag{4.1}$$

A direct calculation gives the following analytical expression of $\mathcal{A}$:

$$\mathcal{A}(x_1, x_2) = \frac{1}{5}(2.5 + 1.5\sin(2\pi x_1))(2.5 + 1.5\cos(2\pi x_2))1_{2\times 2}. \tag{4.2}$$

The offline triangulation $\mathcal{T}_H$ consists of a uniform $Q\times Q$ squares. To obtain the online triangulation τ_h, we firstly triangulate D into a uniform $N\times N$ squares and secondly divide each square into two sub-triangles along its diagonal with positive slope. To study the effect of the reconstruction, we use the analytical expression (4.2) of $\mathcal{A}$ for reconstruction so that $e_1(\text{MOD}) = 0$. We denote this numerical solution by u_h^0. Fig. 2 presents the accuracy of the offline computation, which corroborates the estimate (3.17).
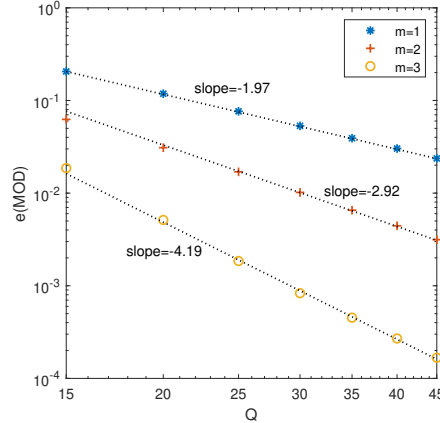

Figure 2: $e(\text{MOD})$ for different reconstruction orders.

To study the convergence rate of the method, it is more convenient to reshape the error estimates in Theorem 3.1 in terms of N and Q as follows.

$$\begin{cases} \|\nabla(u_0 - u_h)\|_{L^2(D)} \le C(N^{-l} + Q^{-(m+1)}), \\ \quad \|u_0 - u_h\|_{L^2(D)} \le C(N^{-(l+1)} + Q^{-(m+1)}). \end{cases} \tag{4.3}$$

Balancing the discretization error and the reconstruction error, we set Q as

$$Q = \begin{cases} \mathcal{O}(N^{\frac{l}{m+1}}), & H^1 \text{ error,} \\ \mathcal{O}(N^{\frac{l+1}{m+1}}), & L^2 \text{ error.} \end{cases}$$

Following this refinement strategy, we plot the relative H^1 error and L^2 error in Fig. 3, which is consistent with the estimates (4.3).
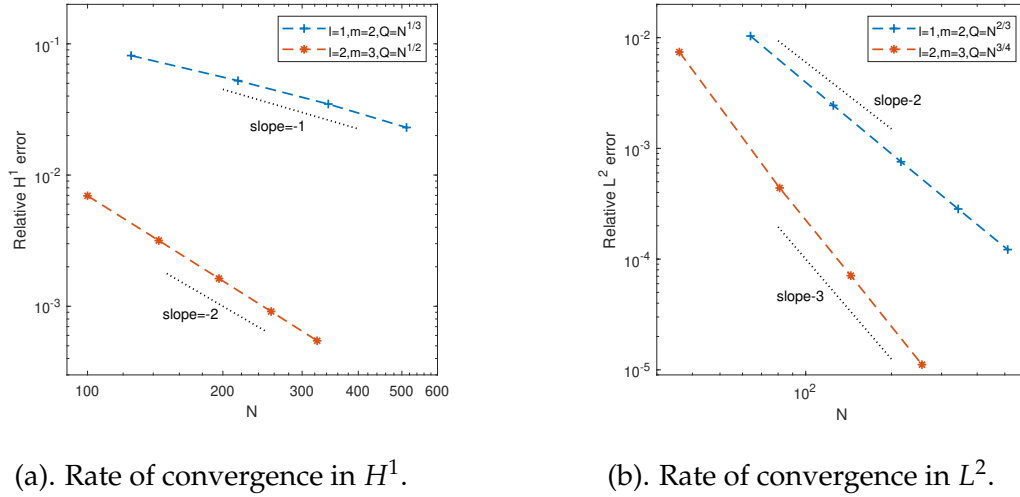

(a). Rate of convergence in H^1.

(b). Rate of convergence in L^2.

Figure 3: Rates of convergence for Example 4.1.

We may also fix the accuracy of the online solver and compare the effect of the reconstruction with different orders in the offline computation. Particularly we choose the online solver as $\mathbb{P}_2$ Lagrange element with the meshsize parameter $N=100$. Fig. 4 clearly shows that the higher-order reconstruction is more accurate with less cost.

Next, we also fix the accuracy of the online solver and compare the number of the cell problems and the running time with reconstructions of different orders. The online solver is still $\mathbb{P}_2$ Lagrange element with the meshsize parameter $N=100$. Instead of using the analytical expression of $\mathcal{A}$ for reconstruction, we solve the periodic cell problems

$$\begin{cases} -\operatorname{div}(a^\varepsilon \nabla v_i^\varepsilon(x)) = 0 & \text{in} \quad I_\varepsilon, \\ v_i^\varepsilon - x_i \quad \text{is periodic} & \text{on} \quad \partial I_\varepsilon, \end{cases} \tag{4.4}$$

with $\mathbb{P}_2$ Lagrange element, and employ (2.2) to obtain A_H. We denote the reconstructed solution as u_h^1, which is very close to u_h^0 in the sense

$$\frac{\|u_h^1 - u_h^0\|_{L^2(D)}}{\|u_h^0\|_{L^2(D)}} \leq 10^{-8}.$$
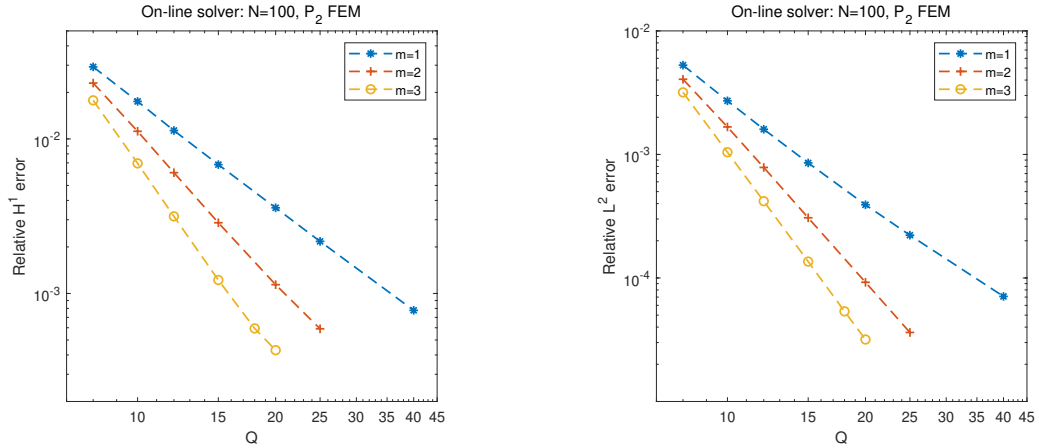
(a). The H^1 error with different reconstruction orders.

(b). The L^2 error with different reconstruction orders.

Figure 4: The effect of Q and the reconstruction orders in Example 4.1; Online solver is $\mathbb{P}_2$ FEM and $N=100$.

This means that the $e_1(\mathrm{MOD})$ is very small, though nonzero. Table 2 shows that the main cost comes from solving the cell problems, and higher-order reconstruction is more efficient in terms of both accuracy and efficiency.

Table 2: Comparison among reconstruction of different orders with the same online solver ($N=100$, $\mathbb{P}_2$ FEM).

	Q	$\sharp$ cell problems	Relative H^1 error	Time (cell problems)	Total Time
$m=1$	45	2025	6.03e-4	348.65s	363.27s
$m=2$	25	625	5.92e-4	106.97s	110.76s
$m=3$	18	324	5.92e-4	54.5s	58.23

In the last test, we compare the offline-online method with the HMM-LS method in [35]. The H^1 error bound in (3.21) may be rewritten as

$$\|\nabla(u_0-\widetilde{u}_h)\|_{L^2(D)} \le C(N^{-l}+N^{-(m+1)}+e_1(\mathrm{MOD})).$$

We follow the setup in the last test to solve the cell problems so that $e_1(\mathrm{MOD})$ is negligible in the above estimate. We take $l=2$ and $m=1$ to balance the error so that

$$\|\nabla(u_0-\widetilde{u}_h)\|_{L^2(D)} \le C N^{-2}.$$

In the offline-online method, we use $\mathbb{P}_2$ Lagrange element as the macroscopic solver and choose the reconstruction order $m=2$ and $m=3$. To balance the error of (4.3), we take $Q=2N^{2/3}$ for $m=2$ and $Q=2.5N^{1/2}$ for $m=3$ and we obtain $\|\nabla(u_0-u_h)\|_{L^2(D)} \le C N^{-2}$.

It seems that the relative H^1 errors for both methods reported in Fig. 5 are comparable, which is consist with the theoretical estimates. The number of the cell problems and the running time plot in Fig. 5 demonstrate that the offline-online method is more efficient. This is easily understood because the number of the cell problems in the offline-online method is of $\mathcal{O}(N^{4/3})$ for $m=2$ and is of $\mathcal{O}(N)$ for $m=3$, while the number of the cell problems in HMM-LS is of $\mathcal{O}(N^2)$.
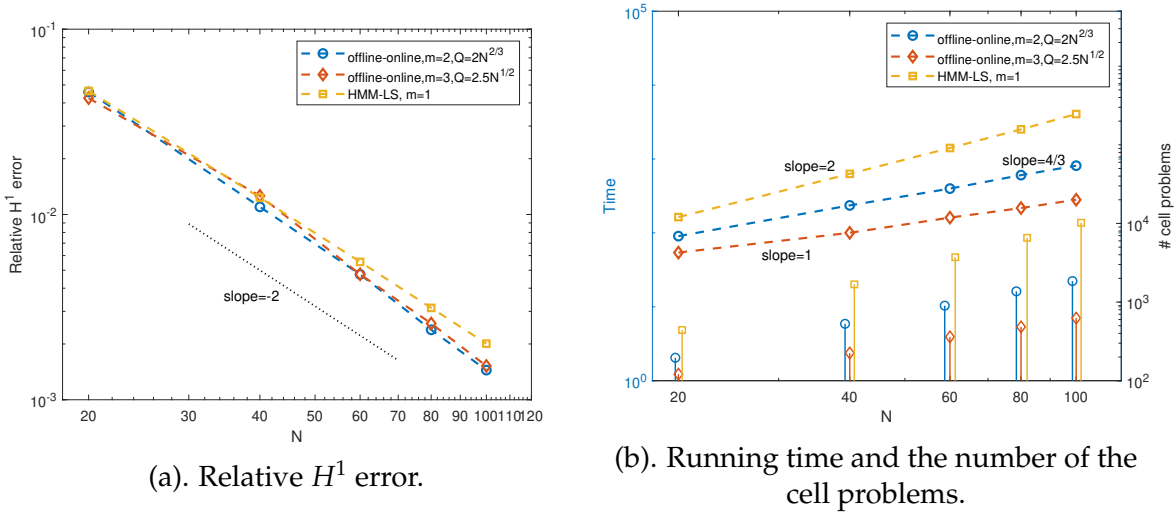

(a). Relative H^1 error.

(b). Running time and the number of the cell problems.

Figure 5: Comparison between the offline-online method and the HMM-LS method in [35].

4.2   Problems posed on L-shape domain in $d=2,3$

In this part, we test the offline-online strategy for a problem posed on L-shape domain. It is well-known that the adaptive strategy has to be used in the macroscopic solver because the solution u_0 is nonsmooth [26].

Example 4.2. The boundary value problem is the same with Example 4.1 except that the domain $D=(-1,1)^2\setminus[0,1]\times[-1,0]$. We let

$$u_0(x)=(x_1^2+x_2^2)^{1/3}.$$

The inhomogeneous boundary condition $g(x)=u_0(x)|_{x\in\partial D}$ and the source term f is computed from the homogenized problem $(4.1)_1$.

The triangulation $\mathcal{T}_H$ is plot in Fig. 6a, we have 972 sampling points in total. The periodic cell problems (4.4) are solved by Fourier spectral method [49]. The third order reconstruction is used to obtain A_H. Under these settings, $e_1(\mathrm{MOD})=2.5E-02$ in the offline stage. $\mathbb{P}_1$ element is used as the macroscopic solver, and we set $\mathrm{TOL}=10^{-2}$ and
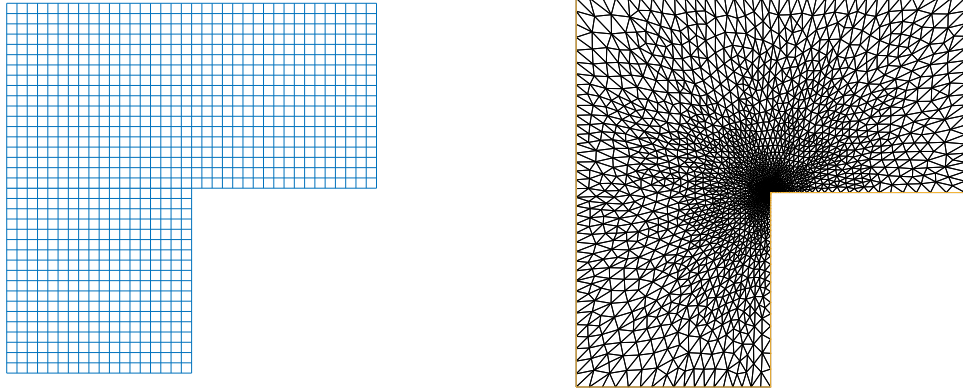
(a). Offline sampling mesh $\mathcal{T}_H$.

(b). Online adaptive mesh τ_h after 8 iterations.

Figure 6: Offline and online triangulations for L-shaped domain.

let $\mathcal{T}_H$ be the initial mesh of the online computation. The mesh is refined by the strategy taken from [27], and we plot the resulting mesh after 8 iterations in Fig. 6b.

The error estimate reads as

$$\|\nabla(u_0-u_h)\|_{L^2(D)} \leq C\left(\text{DOF}^{-1/2}+Q^{-4}+e^{-\beta N_s}\right),\tag{4.5}$$

where DOF is the total degrees of freedom in the online computation, and the factor $e^{-\beta N_s}$ comes from the approximation error caused by solving the cell problem, and β is a universal constant and N_s is the points we used in each cell. Given the error tolerance (TOL), we obtain the offline refinement strategy as

$$Q \lesssim \text{TOL}^{-1/4} \qquad \text{and} \qquad e^{-\beta N_s} \lesssim \text{TOL}.$$

The relative H^1 error is plot in Fig. 7a, and we observe that the error decays with optimal rate of convergence $\mathcal{O}(\text{DOF}^{-1/2})$.

Next we decrease the error tolerance to 10^{-3}, and we also use the third order reconstruction to approximate $\mathcal{A}$ and there are 4332 sampling points in total. The error caused by the reconstruction is $e(\text{MOD}) = 9.48E-04$. The relative H^1 error is plot in Fig. 7b, and we observe that the error decays with optimal rate of convergence $\mathcal{O}(\text{DOF}^{-1/2})$.

Finally we compare the accuracy and efficiency between our method and HMM. The offline computation is the same as the previous test, while in the online stage, we solve the homogenized problem with linear element over the same mesh as HMM. Such mesh is obtained after 8 iterations with 34394 triangles from a uniform mesh dividing each of the uniform squares into two sub-triangles; see Fig. 6b. It is clear that the mesh refinement mostly takes place in the vicinity of the re-entrant corner. It follows from Table 3 that the offline-online method and HMM attain almost the same H^1 error, while the running time

(a). Tolerance$=10^{-2}$.   (b). Tolerance$=10^{-3}$.
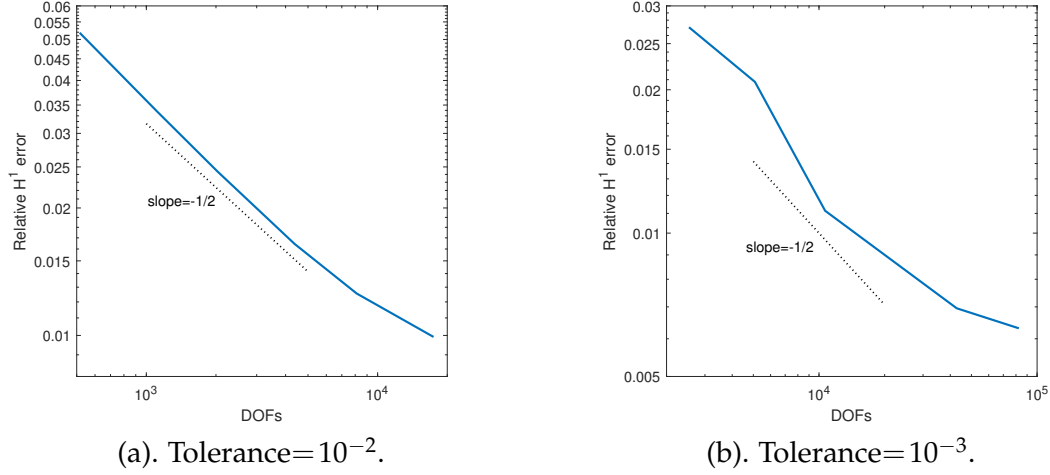
Figure 7: The H^1 error of Example 4.2

of the offline-online method is only around 2.46% of HMM. This is easily understood because the number of the cell problems is only 1.67% of HMM.

Table 3: Comparison between offline-online method and HMM.

	♯ cell problems	Relative H^1 error	Time
Offline-Online	576	3.53e-5	295.95s
HMM	34394	3.62e-5	12020.6s

Next we consider a three-dimensional problem in L-shape domain.

Example 4.3. The example is the same with Example 4.1 except that

$$a^{\varepsilon}(x) = \frac{(2.5+1.5\sin(2\pi x_1))(2.5+1.5\cos(2\pi x_2))(2.5+1.5\cos(2\pi x_3))}{(2.5+1.5\sin(2\pi x_1/\epsilon))(2.5+1.5\sin(2\pi x_2/\epsilon))(2.5+1.5\sin(2\pi x_3/\epsilon))}1_{3\times3\times3},$$

and $D = (0,1)^3 \setminus [0,0.5]^3$. The homogenization problem is the same with (4.1) with the effective matrix as

$$\mathcal{A}(x) = \frac{1}{5}(2.5+1.5\sin(2\pi x_1))(2.5+1.5\cos(2\pi x_2))(2.5+1.5\cos(2\pi x_3))1_{3\times3\times3}. \quad (4.6)$$

We let

$$u_0(x) = \left((x_1-0.5)^2 + (x_2-0.5)^2 + (x_3-0.5)^2\right)^{1/10},$$

and

$$g(x) = u_0(x)|_{\partial D} \qquad f(x) = -\text{div}(\mathcal{A}(x)\nabla u_0).$$

We write

$$\|\nabla(u_0-u_h)\|_{L^2(D)}\leq C\left(\mathrm{DOF}^{-1/3}+Q^{-(m+1)}+e^{-\beta N_s}\right). \tag{4.7}$$

The offline refinement strategy reads as

$$Q\gtrsim \mathrm{TOl}^{-\frac{1}{m+1}} \qquad \text{and} \qquad e^{-\beta N_s}\lesssim \mathrm{TOL}.$$

We set $\mathrm{TOL}=3E-3$ and $m=3$. We firstly divide the domain into 7 subdomain as in the last example, next we triangulate each subdomain by a uniform mesh with $Q=33$. There are 7623 sampling points in total; see Fig. 8a. The H^1 error is plot in Fig. 8b. We observe that the H^1 error decays with an optimal convergence rate $\mathcal{O}(\mathrm{DOF}^{-1/3})$.
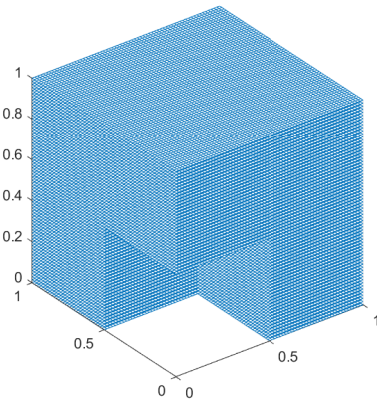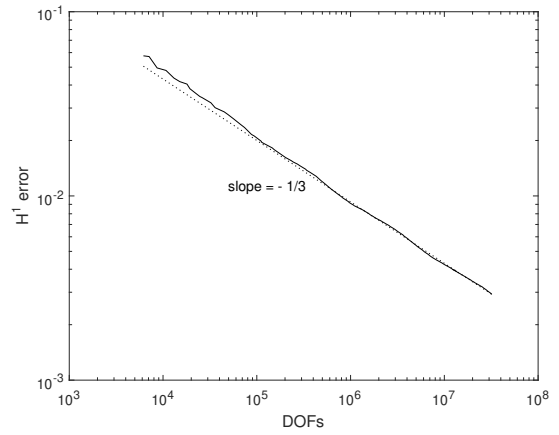
(a). Offline sampling mesh $\mathcal{T}_H$.   (b). H^1 error with tolerance$=3E-3$.

Figure 8: Offline triangulation and H^1 error for Example 4.3.

4.3 Example with quasi-periodic coefficient

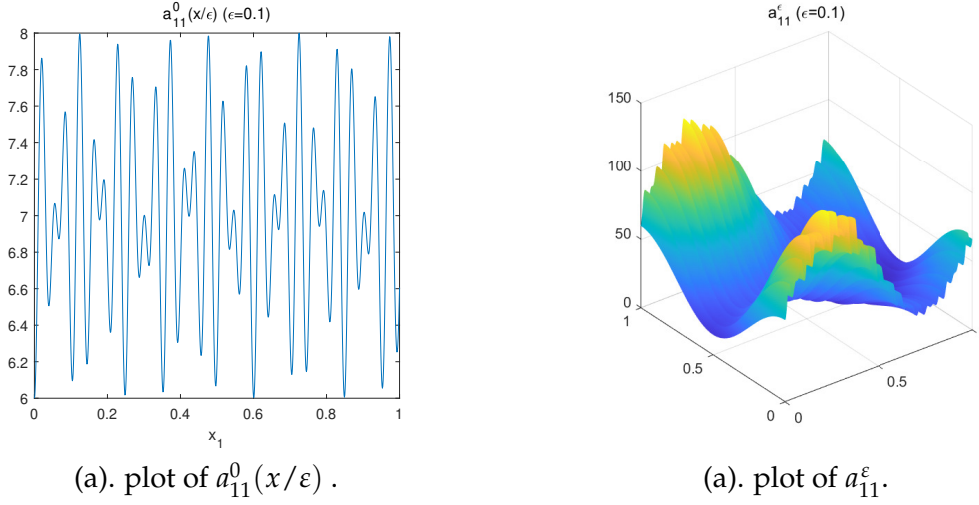**Example 4.4.** The example is the same with Example 4.1 except that

$$a^\varepsilon(x)=a^0(x/\varepsilon)a^1(x)1_{2\times2}, \tag{4.8}$$

where

$$a^0(x)=\begin{pmatrix} 6+\sin(2\pi x_1)^2+\sin\left(2\sqrt{2}\pi x_1\right)^2 & 0 \\ 0 & 6+\sin(2\pi x_2)^2+\sin\left(2\sqrt{2}\pi x_2\right)^2 \end{pmatrix},$$

and

$$a^1(x)=(2.5+1.5\sin(2\pi x_1))(2.5+1.5\cos(2\pi x_2)).$$

(a). plot of $a^0_{11}(x/\varepsilon)$.   (a). plot of a^ε_{11}.

Figure 9: Plots of the coefficients in Example 4.4 ($\varepsilon=0.1$)

Such coefficient belongs to Kozlov class [32] and the example is adapted from [24]. The coefficients $a^0_{11}(x/\varepsilon)$ and a^ε_{11} are visualized in Fig. 9 with $\varepsilon=0.1$. The effective matrix is given by

$$\mathcal{A}_{ij}(x)=\liminf_{R\to+\infty}\left\langle\nabla(\chi^j+x_j)\cdot a^0(x)\nabla(\chi^i+x_i)\right\rangle_{Q_R}*a_1(x),\quad i,j=1,2, \qquad (4.9)$$

where $Q_R=(-R,R)^2$ and $\{\chi^i\}^2_{i=1}$ are the solutions of

$$-\nabla\cdot(a^0(x)\nabla\chi^i)=\nabla\cdot(a^0(x)\nabla x_i)\quad\text{in}\quad\mathbb{R}^2. \qquad (4.10)$$

As oppose to the locally periodic case, there is no explicit expression of $\mathcal{A}$. The naive approach to approximate (4.9) consists in replacing $\{\chi^i\}^2_{i=1}$ by $\{\chi^i_R\}^2_{i=1}$, which are solutions of a truncated cell problem

$$\begin{cases} -\nabla\cdot(a^0(x)\nabla\chi^i_R)=\nabla\cdot(a^0(x)\nabla x_i)\quad\text{in}\quad Q_R, \\ \qquad\qquad\qquad\qquad\quad \chi^i_R=0\quad\text{on}\quad\partial Q_R, \end{cases}$$

and $\mathcal{A}_R$ is defined by

$$(\mathcal{A}_R(x))_{ij}:=\left\langle\nabla(\chi^j_R+x_j)\cdot a^0(x)\nabla(\chi^i_R+x_i)\right\rangle_{Q_R}*a_1(x),\quad i,j=1,2.$$

We take $R=200$ and use $\mathbb{P}_2$ Lagrange element over a 4000×4000 mesh to solve the above truncated cell problem and obtain

$$\mathcal{A}_{200}=\begin{pmatrix} 7.00*a^1(x) & 0 \\ 0 & 7.00*a^1(x) \end{pmatrix}.$$

In the offline stage, we choose $\delta = 10\varepsilon$ and use $\mathbb{P}_2$ Lagrange element over a mesh with size 120×120 to solve the cell problems (2.3). Next we plot the relative H^1 error and the relative L^2 error in Fig. 10, which clearly shows the higher-order reconstruction is more accurate.
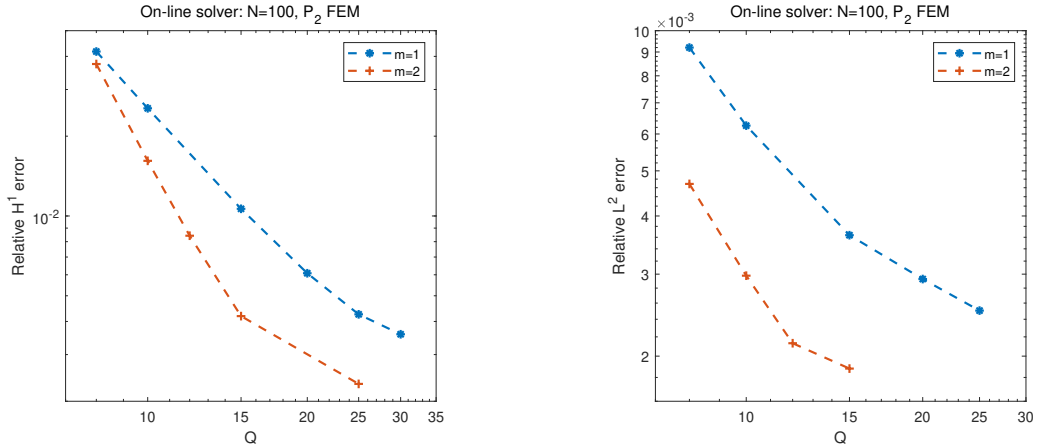
(a). The relative H^1 error with different reconstruction orders.

(b). The relative L^2 error with different reconstruction orders.

Figure 10: The effect of Q and the reconstruction orders in Example 4.4; Online solver is $\mathbb{P}_2$ FEM and $N = 100$.

Finally we compare the number of the cell problems and the running time among reconstructions of different orders in Table 4. The second order reconstruction takes only

Table 4: Comparison of the reconstruction of different orders; online solver is $\mathbb{P}_2$ FEM and $N = 100$.

	Q	$\sharp$ cell problems	Relative H^1 error	Time
$m=1$	20	400	6.08e-3	1083.08s
$m=2$	14	196	4.95e-3	546.24s

about 50% of the time used for the first order reconstruction. It is clear that the higher-order reconstruction is more accurate with less cost.

4.4 Example with random coefficient

Example 4.5. The example is the same with Example 4.1 except that

$$a^\varepsilon(x) = (a^\varepsilon_{\text{rand}} + a_0(x)) 1_{2\times 2}, \tag{4.11}$$

where $a^\varepsilon_{\text{rand}}$ is a random checker-board, and $a_0(x) = (2.5 + 1.5\sin(2\pi x_1))(2.5 + 1.5\cos(2\pi x_2))$. $a^\varepsilon_{\text{rand}}$ is constructed by partitioning $D = (0,1)^2$ into uniform square cells of size ε, each of which is randomly designated as k_1 or k_2 with probability p_1 and $p_2 = 1 - p_1$, respectively.

We visualize one realization of the random coefficients in Fig. 11 with $\varepsilon = 0.02$. Theorem 3.1 remains true with a minor modification of $e_1(\text{MOD})$, we refer to [19] for related result.
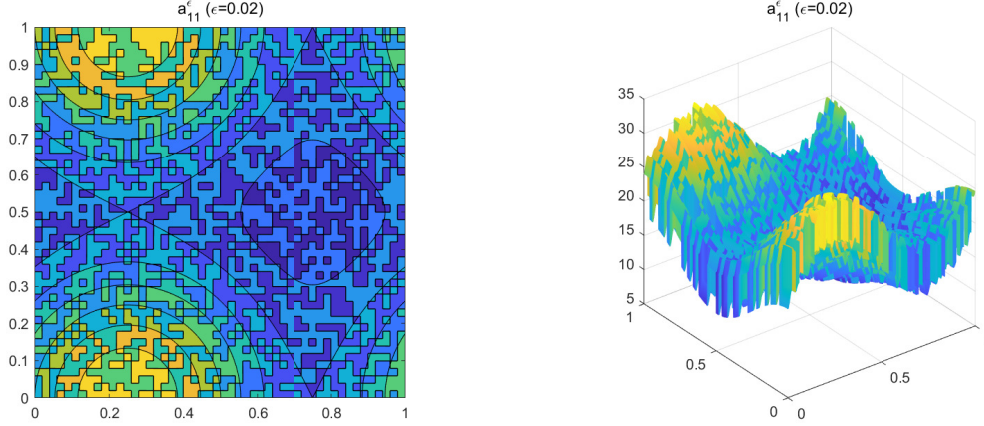

Figure 11: Plot of one realization of a_{11}^{ε} in (4.11). (a) is the contour of a_{11}^{ε}.

As oppose to the standard random checker-board [30], there is no explicit formula for the effective matrix of (4.11). To extract the effective matrix over D, we take $\delta = 16\varepsilon$ as the cell size and use $\mathbb{P}_2$ Lagrange element over a mesh of size 160×160 to solve the cell problem (4.4). We set the offline mesh size as $Q = 20$ and use third order reconstruction to get an approximation effective matrix $\mathcal{A}$, which will be exploited to obtain the reference solution in the following test.

To justify this approach, we choose A_H at three representative points in D: $A = (1/4, 0)$, which is one of the maximum point of a_0 over D; $B = (3/4, 1/2)$, which is one of the minimum point of a_0 over D, and $C = (1, 1/4)$, which is one of the maximum point of ∇a_0 over D. We let $\delta = L\varepsilon$ and use $\mathbb{P}_2$ Lagrange element over a mesh of size $10L \times 10L$ to solve the cell problem (4.4) posed over these three points. We denote by $\mathcal{A}_{16}^n$ the approximation effective matrix for $n-$th realization. In addition, we use the empirical average

$$\mathbb{E}(\mathcal{A}_{16})(x) := \frac{1}{N} \sum_{n=1}^{N} \mathcal{A}_{16}^n(x) \tag{4.12}$$

as the proxy of the expectation $\mathbb{E}$, where N is the total number of the realization. We take $N = 1000$ in the simulation, and the results are reported in Table 5.

Secondly, for $x = A, B, C$, we measure the variance as

$$\sigma_L^{\text{diag}}(x) := \sqrt{\mathbb{E}\left((\mathcal{A}_L(x)_{11} - \mathbb{E}(\mathcal{A}_{16}(x))_{11})^2 + (\mathcal{A}_L(x)_{22} - \mathbb{E}(\mathcal{A}_{16}(x))_{22})^2 \right)},$$

and

$$\sigma_L^{12}(x) := \sqrt{\mathbb{E}(\mathcal{A}_L(x)^2)_{12}}.$$

Table 5: The approximating effective matrix on three points with $L=16$.

	$\mathbb{E}(\mathcal{A})$
$A=(1/4,0)$	$\begin{pmatrix} 20.78 & -2.58e-4 \\ -2.58e-4 & 20.78 \end{pmatrix}$
$B=(3/4,1/2)$	$\begin{pmatrix} 5.20 & -9.47e-4 \\ -9.47e-4 & 5.20 \end{pmatrix}$
$C=(1,1/4)$	$\begin{pmatrix} 10.84 & -4.86e-4 \\ -4.86e-4 & 10.84 \end{pmatrix}$

Fig. 12a and Fig. 12b suggest that $\sigma_L^{\mathrm{diag}}(x)$ and $\sigma_L^{12}(x)$ decay as $\mathcal{O}(L^{-1})$, which is consistent with the theoretical predictions [25]. A systematical numerical tests for the variance may be found in a recent work [31].
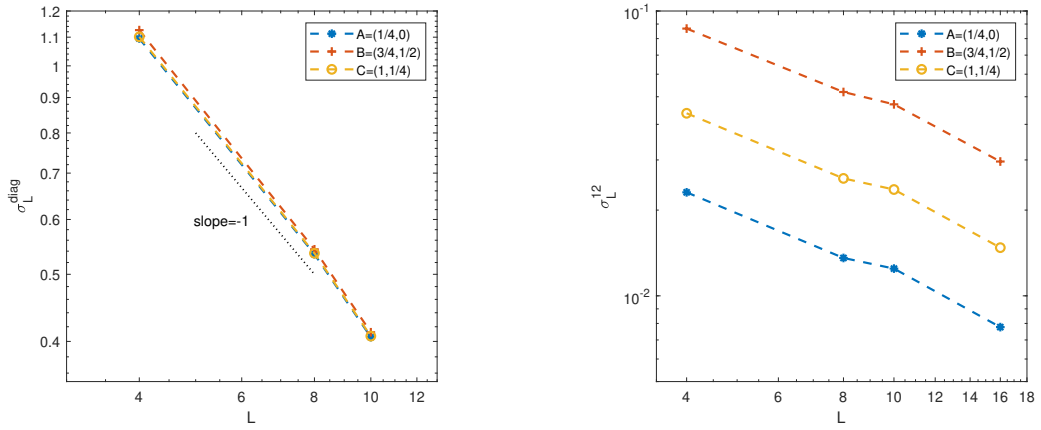
(a). Decays of σ_L^{diag} with respect to the cell size L.

(b). Decays of σ_L^{12} with respect to the cell size L.

Figure 12: The accuracy for the approximating effective matrix.

Finally we compare the effect of the reconstructions with different orders in the offline computation. We fix the online solver as $\mathbb{P}_2$ Lagrange element over a mesh with size 100×100. In the offline stage, we choose $\delta=8\varepsilon$ and use $\mathbb{P}_2$ Lagrange element over a mesh with size 80×80 to solve the cell problem (2.3). We take $Q=8,10,16$ and $m=1,2$ in the tests, and compute the ensemble average of the relative H^1 error and L^2 error

$$\frac{\mathbb{E}\left(\|\nabla(u_0-u_h)\|_{L^2(D)}\right)}{\|\nabla u_0\|_{L^2(D)}} \quad \text{and} \quad \frac{\mathbb{E}\left(\|u_0-u_h\|_{L^2(D)}\right)}{\|u_0\|_{L^2(D)}}.$$

The expectation $\mathbb{E}$ is replaced by the empirical average as that in (4.12) with $N=1000$

realizations. The visualization in Fig. 13a and Fig. 13b clearly shows that the second order reconstruction is more accurate than the first order reconstruction.
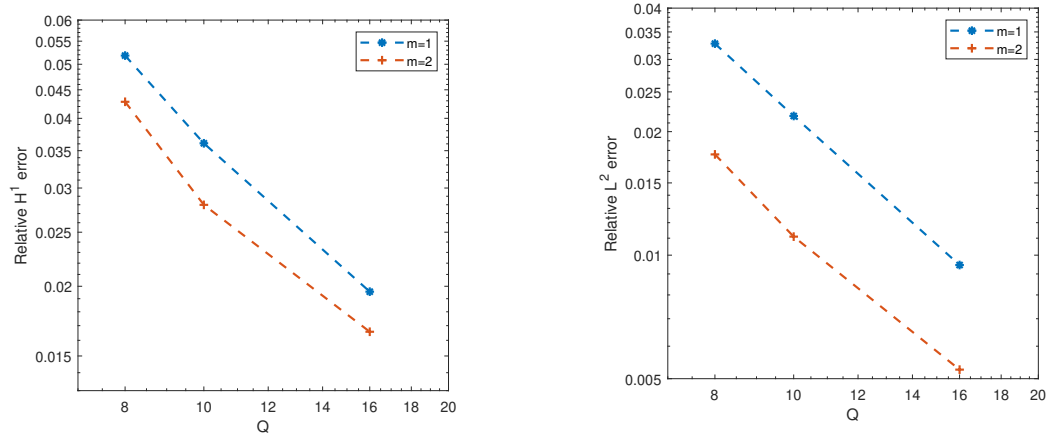
(a). Ensemble average of the H^1 error.

(b). Ensemble average of the L^2 error.

Figure 13: The error between numerical solution and homogenized solution in Example 4.5.

5   Conclusion

We have proposed a new online-offline method to solve the multiscale elliptic problems. Both theoretical and numerical results show that the method significantly reduces the cost while retains the optimal rate of convergence. Moreover, the strategy is problem independent, and it can be extended to time-dependent problems [21, 42]. The implementation of the present method is mainly based on the a priori error estimates, adaptive algorithms should be developed for automatic tuning of the parameters so that the method is more efficient. We shall leave all these issues in the future work.

Acknowledgements

References

[1] A. Abdulle, and Y. Bai, Reduced basis finite element heterogeneous multiscale method for high-order discretizations of elliptic homogenization problems, J. Comput. Phys., 231 (2012), 7014-7036.

[2] A. Abdulle, and Y. Bai, Adaptive reduced basis finite element heterogeneous multiscale method, Comput. Methods Appl. Mech. Engrg., 257 (2013), 203-220.

[3] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden, The heterogeneous multiscale method, Acta. Numeric, 21 (2012), 1-87.

[4] R. A. Adams, and J. J. F. Fournier, Sobolev Spaces, 2nd eds. Elsevier, Amsterdam, 2003.

[5] D. Arjmand, and O. Runborg, A time dependent approach for removing the cell boundary error in elliptic homogenization problems, J. Comput. Phys. 341 (2016), 206-227.

[6] I. Babuška, Homogenization and its applications, mathematical and computational problems. In Numerical Solutions of Partial Differential Equations-III, B. Hubbard, eds., New York, Academic Press, 1976, 89-116.

[7] I. Babuška, B. A. Sabo, and I. Katz, The p-version of the finite element method, SIAM J. Numer. Anal., 18 (1981), 515-545.

[8] M. Barrault, Y. Maday, N. Nguyen, and A. Patera, An empirical interpolation method: application to efficient reduced-basis discretization of partial differential equations, C. R. Acad. Sci. Paris Ser. I., 339 (2004), 667–672.

[9] A. Bensoussan, J.-L. Lions, and G. Papanicolaou. Asymptotic Analysis for Periodic Structures, volume 5 of Studies in Mathematics and its Applications. North-Holland Publishing Co., Amsterdam, 1978.

[10] X. Blanc, and C. Le Bris, Improving on computation of homogenized coefficients in the periodic and quasi-periodic settings, Netw. Heterog. Media, 5 (2010), 1-29.

[11] L. Bos, J.-P. Calvi, N. Levenberg, A. Sommariva, and M. Vianello, Geometric weakly admissible meshes, discrete least squares approximations and approximate Fekete points, Math. Comput., 80(2011), 1623–1638.

[12] S. C. Brenner, and L. R. Scott, The Mathematical Theory of Finite Element Methods, 3rd eds., Springer Science + Business Media, LLC, 2008.

[13] J.P. Calvi, and N. Levenberg, Uniform approximation by discrete least squares polynomials, J. Approx. Theory, 152 (2008), 82-100.

[14] P. G. Ciarlet, The Finite Element Method for the Elliptic Problems, North-Holland, Amsterdam, 1978.

[15] B. DeVolder, J. Glimm, J.W. Grove, Y. Kang, Y. Lee, K. Pao, D.H. Sharp, and K. Ye, Uncertainty quantification for multiscale simulations, J. Fluids Engrg., 124 (2002), 29-41.

[16] R. Du, and P. B. Ming, Heterogeneous multiscale finite element method with novel numerical integration schemes, Comm. Math. Sci., 8 (2010), 863-885.

[17] W. E, Principles of Multiscale Modeling, Cambridge University Press, Cambridge, 2011.

[18] W. E and B. Engquist, The heterogeneous multi-scale methods, Comm. Math. Sci., 1 (2003), 87-132.

[19] W. E, P. B. Ming, and P. W. Zhang, Analysis of the heterogeneous multiscale method for elliptic homogenization problems, J. Amer. Math. Soc., 18(2005), 121-156.

[20] Y. Efendiev, and T. Y. Hou, Multiscale Finite Element Methods, Theory and Applications, Springer, New York, 2009.

[21] B. Engquist, H. Holst, and O. Runborg, Multiscale methods for wave propagation in heterogeneous media over long time, B. Engquist et al. (eds.), Numerical Analysis of Multiscale

Computations, Lecture Notes in Computational Science and Engineering, Springer-Verlag Berlin Heidelberg 2012. 67-186.

[22] B. Engquist, and P. E. Souganidis, Asymptotic and numerical homogenization, Acta Numerica, 17 (2008), 147-190.

[23] A. Gloria. Reduction of the resonance error - Part 1: approximation of homogenized coefficients. Math. Models Methods Appl. Sci., 21 (2011), 1601-1630.

[24] A. Gloria, and M. Habibi, Reduction in the resonance error in numerical homogenization II: correctors and extrapolation, Found Comput. Math., 16 (2016), 217-296.

[25] A. Gloria, S. Neukamm, and F. Otto, Quantification of ergodicity in stochastic homogenization: optimal bounds via spectral gap on Glauber dynamics. Invent. Math., 199 (2015), 455-515.

[26] P. Grisvard, Elliptic Problems in Nonsmooth Domains, Pitman, Boston, 1985.

[27] F. Hecht, New development in FreeFem++, J. Numer. Math., 20(2012), 251-265.

[28] Y. F. Huang, Study on some efficient algorithms for multiscale partial differential equations, PH. D dissertation, Chinese Academy of Sciences, 2018.

[29] P. Kanouté, D. P. Boso, J. L. Chaboche, and B. A. Schrefler, Multiscale methods for composites: a review, Arch. Comput. Meth. Engreg., 16 (2009), 31-75.

[30] J. B. Keller, A theorem on the conductivity of a composite medium, J. Math Phys. 5 (1964) 548-549.

[31] V. Khoromskaia, B. N. Khoromskij, and F. Otto. Numerical study in stochastic homogenization for elliptic partial differential equations: convergence rate in the size of representative volume elements, Numer. Linear Algebra Appl., 27 (2020), e2296.

[32] S. M. Kozlov, Averaging of differential operators with almost periodic rapidly oscillating coefficients, Math. Sb., 107 (1978), 199-217, 317.

[33] A. Kroó, and S. Révész, On Bernstein and Markov-type Inequalities for multivariate polynomials on convex bodies, J. Approx. Theory, 99 (1999), 134-152.

[34] A. Kroó, and J. Szabados, Markov-Bernstein type inequalities for multivariate polynomials on sets with cusps, J. Approx. Theory, 102 (2000), 72-95.

[35] R. Li, P. B. Ming, and F. Y. Tang, An efficient high order heterogeneous multiscale method for elliptic problems, Multiscale Model. & Simul., 10 (2012), 259-283.

[36] R. Li, P. B. Ming, Z. Y. Sun, and Z. J. Yang, An arbitrary-order Discontinuous Galerkin method with one unknown per element, J. Sci. Comput., 80 (2019), 268-288.

[37] R. Li, and F.Y. Yang, A discontinuous Galerkin method by patch reconstruction for elliptic interface problem on unfitted mesh, SIAM J. Sci. Comput. 42 (2020), A1428-A1457.

[38] A. Markov, Sur une question posee par Mendeleieff, Bull. Acad. Sci. St. Petersburg, 62 (1889), 1-24.

[39] K. Matouš, M. G. D. Geers, V.G. Kouznestova, and A. Gillman, A review of predicative nonlinear theories for multiscale modeling of heterogeneous materials, J. Comput. Phys., 330 (2017), 192-220.

[40] J.C. Michel, H. Moulinec, and P. Suquet. Effective properties of composite materials with periodic microstructure: a computational approach, Comput. Meth. in Appl. Mech. Engreg., 172 (1999), 109-143.

[41] P. B. Ming, and X. Y. Yue, Numerical methods for multiscale elliptic problems, J. Comput. Phys., 214 (2006), 421-445.

[42] P.B. Ming, and P.W. Zhang, Analysis of the heterogeneous multiscale method for parabolic homogenization problems, Math. Comput., 76 (2007), 153-177.

[43] J.-C. Mourrat, Efficient methods for the estimation of homogenized coefficients, Found.

Comput. Math., 19 (2019), 435-483.

[44] N. C. Nguyen, A multiscale reduced-basis method for parametrized elliptic partial differential equations with multiple scales, J. Comput. Phys., 227 (2008), 9807-9822.

[45] F. Piazzon, Optimal polynomial admissible meshes on some classes of compact subsets of $\mathbb{R}^d$, J. Approx. Theory, 207 (2016), 241-264.

[46] E. Rakhmanov, Bounds for polynomials with a unit discrete norm, Ann. Math., 165 (2007), 55-88.

[47] C. Schwab, p-and hp-Finite Element Methods, Oxford University Press, 1998.

[48] H. J. Schmid, On cubature formulae with a minimal number of knots, Numer. Math., 31 (1978), 281–297.

[49] J. Shen, T. Tang, and L. L. Wang, Spectral Method: Algorithms, Analysis and Applications, Springer, Berlin, Heidelberg, 2011.

[50] P. Šolín, K. Segeth, and I. Doležel, Higher-order Finite Element Methods, Chapman & Hall/CRC, Boca Raton, FL, 2004.

[51] A. H. Stroud, Approximate Calculation of Multiple Integrals, Prentice Hall, Englewood Cliffs. 1971.

[52] D. O. Sullivan, Exploring spatial process dynamics using irregular cellular automaton models. Geogr. Anal., 33 (2001), 1-18.

[53] L. Tartar, H-convergence, Course Peccot, Collége de France. Partially written by F. Murat. Séminaire d'Analyse Fonctionnelle et Numérique de l'Université d'Alger, 1977–1978, March 1977.

[54] L. Tartar, The General Theory of Homogenization: A Personalized Introduction, Springer-Verlag, 2010.

[55] S. Torquato, Random Heterogeneous Materials: Microstructure and Macroscopic Properties, Springer-Verlag, New York, 2002.

[56] R. Verfürth, On the constants in some inverse inequalities for finite element functions, Technical Report 257, University of Bochum, 1999.

[57] H. Wendland, Scattered Data Approximation, Cambridge University Press, Cambridge, 2005.

[58] D. Wilhelmsen, A Markov inequality in several dimensions, J. Approx. Theory, 11 (1974), 216-220.

[59] X. Y. Yue, and W. E, The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size, J. Comput. Phys., 222 (2007), 556-572.

[60] L. B. Zhang, A Parallel algorithm for adaptive local refinement of tetrahedral meshes using bisection, Numer. Math. Theor. Meth. Appl., 2 (2009), 65-89.

[61] T. Zohdi, and P. Wriggers, Introduction to Computational Micromechanics, Springer-Verlag, Berlin, 2005.