

A Nitsche Hybrid Multiscale Method with Non-matching Grids

Pingbing Ming · Siqu Song

Received: date / Accepted: date

Abstract We propose a Nitsche method for multiscale partial differential equations, which retrieves the macroscopic information and the local microscopic information at one stroke. We prove the convergence of the method for second order elliptic problem with bounded and measurable coefficients. The rate of convergence may be derived for coefficients with further structures such as periodicity and ergodicity. Extensive numerical results confirm the theoretical predictions.

Keywords Multiscale PDE · Hybrid method · Nitsche variational formulation · Non-matching grid

1 Introduction

Consider the elliptic problem with Dirichlet boundary condition

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

where D is a bounded domain in \mathbb{R}^n with $n = 2, 3$, and ε is a small parameter that signifies the multiscale nature of the problem. Problem (1.1) may be viewed as a prototypical model of many multiscale problems arising from a variety of contexts, such as the heat conduction and the electromagnetism in composites, or the transport of the porous media. The main quantities of interest for Problem (1.1) are the macroscopical behavior of the solution and

P.B. Ming · S.Q. Song
LSEC, Institute of Computational Mathematics and Scientific/Engineering Computing,
AMSS, Chinese Academy of Sciences
No. 55, East Road Zhong-Guan-Cun, Beijing 100190, China
and School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing
100049, China
E-mail: mpb@lsec.cc.ac.cn, songsq@lsec.cc.ac.cn

the local microscopical information of the solution [16, 8]. Many numerical methods have been developed in the literature to capture either the macroscopical behaviors or the microscopical information of the solution, such as the heterogeneous multiscale methods (HMM) [17], the multiscale finite element methods [28] and many others.

There are also some methods that aim to retrieve the coarse scale information and the local fine scale information simultaneously. Such methods may be roughly grouped into three classes. The first one is the global-local method, which was originally proposed in [41, 48]. The main idea is to solve the coarse scale problem by a numerical upscaling method firstly, and then solve the local problem around the defects or the places for which the fine scale information is of interest, while the coarse scale information is employed as the constraints. This idea has been incorporated into the HMM framework in [17] and the performance has been thoroughly analyzed in [19] and [7]. The global-local method has been extended to solve an elastodynamical wave equation in [8].

The second method is based on the domain decomposition idea, which has been exploited to solve the multiscale PDEs in [22, 4, 15, 2, 38]. The most relevant is the one in [2]. The authors therein used a discontinuous Galerkin HMM in a region with scale separation, while use a finite element method in a region without scale separation. The unknown boundary condition has been supplied by minimizing the difference between the solutions in the overlapped region. The well-posedness and the convergence of this method have been studied in [2] for the periodic media.

The third method relies on the hybridization idea [29]. One solves the following variational problem: Find $v_h \in X_h$ such that

$$\langle b^\varepsilon \nabla v_h, \nabla w \rangle = \langle f, w \rangle \quad \text{for all } w \in X_h, \quad (1.2)$$

where X_h is any finite element space, and we denote the $L^2(D)$ inner product by $\langle \cdot, \cdot \rangle$. Here $b^\varepsilon(x) := \rho(x)a^\varepsilon(x) + (1 - \rho(x))\mathcal{A}(x)$, where \mathcal{A} is the effective matrix arising from the homogenization problem:

$$\begin{cases} -\operatorname{div}(\mathcal{A}(x)\nabla u_0) = f(x), & x \in D, \\ u_0 = 0, & x \in \partial D. \end{cases} \quad (1.3)$$

The coefficient b^ε is a hybridization of the microscopical coefficient and the macroscopical coefficient with a transition function ρ . Roughly speaking, the transition function takes one in the defected region and zero otherwise. The authors proved the well-posedness and the convergence of (1.2) for the bounded and measurable coefficient a^ε . The rate of convergence was derived for the periodic media and the quasi-periodic media. Numerical results in [29] show that this hybrid method is comparable with the classical global-local method **in terms of** both the accuracy and the efficiency, while it is particularly suitable for the scenario that the microscopic coefficient a^ε is only available in part of the domain, while outside this region, the coarse scale information is available for the coefficient fields.

The present work is a follow-up of [29], and there are two contributions. Firstly we employ the variational formulation of Nitsche [40] to solve (1.2), which allows for non-matching grid across the interface. Such numerical interface is caused by the local support of the transition function. The authors in [29] employed the linear finite element over a body-fitted mesh to solve (1.2). Highly refined mesh has to be used around the defect region to ensure the conformity of the mesh and the resolution of the local defects. From this aspect of view, the non-matching grid is more flexible in implementation. Indeed, as demonstrated in § 5, fewer global degrees of freedom is required to achieve the desired accuracy compared to the original hybrid method. We note that Nitsche's method is a powerful tool to deal with the interface problem in finite element method and the discontinuous Galerkin method; See, e.g., [9, 15, 44, 12]. Another contribution is a general method to construct the transition function, which is an essential ingredient of the hybrid method while seems missing in [29], because only the square defects have been dealt with therein, for which the transition function is a tensor product of a spline function in one dimension. It is nontrivial to find such explicit expression of ρ for defects with irregular shape. **The independent of the smoothness of the transition function in the theoretical analysis inspires a straightforward method to construct it, which help us to handle the defects with irregular shapes. And numerical results show that the method works well for such irregular defects without occurring extra cost.**

To analyze the Nitsche hybrid method, we need a well-defined trace over the element boundary, which demands that $u_0 \in H^{1+s}(D)$ with $s > 1/2$. For smooth solution, the Nitsche hybrid method may be analyzed by combining the technique in [29] and the standard way for analyzing DG method [6]. Unfortunately, such smoothness assumption on u_0 may not be true for a rough coefficient matrix \mathcal{A} , or a point load function f , or a nonconvex domain D . Hence we adapt the **medius** analysis [23] to the present problem. To deal with the non-matching grid that is not covered by the standard **medius** analysis [23, 24, 36], we construct a new enriching operator that measures the difference between the discontinuous finite element space and the Sobolev space H^1 over such triangulation. Such difference may be bounded by the jump of the function across the interface, which is independent of the mesh ratio. The enriching operator stems from [10] and [31], which plays an important role in analyzing DG method [31, 23], nonconforming finite element method [10, 34] and the virtual element method [11], **where we just name a few of them and refer to [11] for an updated review.** The main ingredient of the construction is the mesh ratio dependent weights [45, 27] instead of the standard arithmetic mean [31]. Using this enriching operator, we may prove the error estimate without any regularity assumption on u_0 . Though the error bounds weakly depend on the mesh ratio, we may remove such dependence by adjusting the penalized parameter in Nitsche's variational formulation. Besides the non-matching grids, the bounded measurable coefficient a^ε adds certain new difficulties.

The rest of the paper is organized as follows. We introduce the method in § 2. The well-posedness and the error estimates of the proposed method are

proved in § 3, this is also the main theoretical result of the present work. We prove the main technical lemmas in § 4. Numerical examples for defects with various shapes are reported in § 5. Some technical results are included in the Appendix.

Throughout this paper, we shall use Sobolev spaces $W^{r,p}(D)$ with norm $\|\cdot\|_{r,p,D}$ and semi-norm $|\cdot|_{r,p,D}$, and we shall drop the subscript p when $p = 2$. We refer to [3] for details. We shall use C as a generic constant independent of ε , the mesh size h , H and H/h , which may change from line to line.

2 The Nitsche Hybrid Method

To introduce the method, we fix some notations. Let K_0 be the defected region, and we slightly extend K_0 to K_1 and define $d := \text{dist}(K_0, K_1)$. Denote $K_2 = D \setminus K_1$, and let $|K_i| := \text{mes}K_i$ with $i = 1, 2$ and $\Gamma := \partial K_1 \setminus \partial D$. We construct a transition function ρ satisfying

$$\begin{cases} K_1 = \text{supp } \rho, & 0 \leq \rho \leq 1, \\ \rho(x) \equiv 1 & \text{for } x \in K_0. \end{cases}$$

To this end, we firstly set $\rho \equiv 1$ in K_0 and $\rho \equiv 0$ outside K_1 . A one layer mesh is constructed between K_0 and K_1 . Secondly, we use a linear Lagrange interpolant over this triangulation to generate the transition function ρ in $K_1 \setminus K_0$, which is a continuous function. The triangulation between K_0 and K_1 for three different kinds of defects is plotted in Fig. 1. This construction

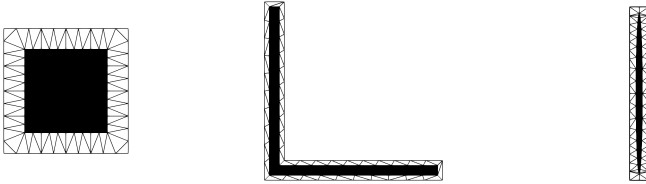


Fig. 1 One layer triangulation of (a) well defect; (b) channel defect; (c) Ellipse defect.

ensures that $\rho(x) = 0$ for $x \in \Gamma$. In what follows, we do not assume any further smoothness on ρ . The effect of the smoothness on ρ will be studied in § 5.2.

We triangulate K_1 and K_2 by \mathcal{T}_h and \mathcal{T}_H with the maximum meshsize h and H , respectively. Hence we have a global triangulation $\mathcal{T}_{h,H} = \mathcal{T}_h \cup \mathcal{T}_H$ over the whole domain though \mathcal{T}_h and \mathcal{T}_H may not be conforming; See Fig. 2 for an illustration of $\mathcal{T}_{h,H}$. We assume that both \mathcal{T}_h and \mathcal{T}_H are shape-regular in the sense of Ciarlet-Raviart [13] with the chunkiness parameter σ .

Denote the set of all edges (faces in $n = 3$) of \mathcal{T}_h and \mathcal{T}_H on Γ by \mathcal{E}_h and \mathcal{E}_H , respectively. Moreover, we denote by \mathcal{E}_Γ the boundary mesh obtained



Fig. 2 Left: the region K_1 and K_2 ; Right: The mesh \mathcal{T}_h in K_1 (blue), \mathcal{T}_H in K_2 (black) and the interface Γ (red).

by intersecting \mathcal{E}_h and \mathcal{E}_H , i.e., $\mathcal{E}_\cap = \{e_\cap = e \cap E : e \in \mathcal{E}_h, E \in \mathcal{E}_H\}$. For convenience, we define

$$\mathcal{T}_h^\Gamma := \{\tau \in \mathcal{T}_h : \bar{\tau} \cap \Gamma \neq \emptyset\} \quad \text{and} \quad \mathcal{T}_H^\Gamma := \{\tau \in \mathcal{T}_H : \bar{\tau} \cap \Gamma \neq \emptyset\},$$

and $\mathcal{T}_{h,H}^\Gamma := \mathcal{T}_h^\Gamma \cup \mathcal{T}_H^\Gamma$. We refer to Fig. 3 for a plot of the mesh around the interface.

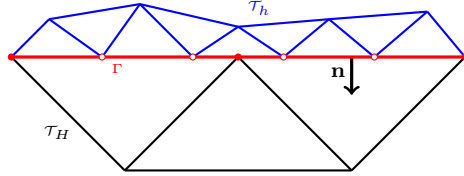


Fig. 3 The collection \mathcal{T}_h^Γ (blue) and \mathcal{T}_H^Γ (black).

Let X_h and X_H be the Lagrange finite element spaces consisting of piecewise polynomials of degree r over \mathcal{T}_h and \mathcal{T}_H , respectively. Over $\mathcal{T}_{h,H}$, we define

$$X_{h,H} := X_h \times X_H \quad \text{and} \quad X_{h,H}^0 := \{v \in X_{h,H} : v|_{\partial D} = 0\}. \quad (2.1)$$

For each $e \in \mathcal{E}_\cap$, there exists $(\tau_1, \tau_2) \in \mathcal{T}_h^\Gamma \times \mathcal{T}_H^\Gamma$ such that $\bar{\tau}_1 \cap \bar{\tau}_2 = e$. Define $h_e := h_{\tau_1}$ and $H_e := h_{\tau_2}$, and the weighted average and jump for $v = (v_1, v_2) \in X_h \times X_H$ on e are defined by

$$\{\{v\}\}_\omega := \omega_1 v_1 + \omega_2 v_2 \quad \text{and} \quad \llbracket v \rrbracket := v_1 - v_2$$

with the weights

$$\omega_1 = \frac{h_e}{h_e + H_e} \quad \text{and} \quad \omega_2 = \frac{H_e}{h_e + H_e}.$$

Similar to the magic DG formula in [6], we have

$$\{\{v\}\}_\omega \llbracket w \rrbracket - \llbracket vw \rrbracket = -\llbracket v \rrbracket \{\{w\}\}_\omega, \quad (2.2)$$

where $\{\{v\}\}^\omega := \omega_2 v_1 + \omega_1 v_2$.

The bilinear form is defined for any $v, w \in X_{h,H}$ as

$$\begin{aligned} B_h^\varepsilon(v, w) := & \sum_{i=1}^2 \int_{K_i} b_h^\varepsilon \nabla v \nabla w \, dx - \sum_{e \in \mathcal{E}_\Gamma} \int_e \{\{b_h^\varepsilon \nabla v \cdot \mathbf{n}\}\}_\omega [w] \, ds \\ & - \sum_{e \in \mathcal{E}_\Gamma} \int_e \{\{(b_h^\varepsilon)^T \nabla w \cdot \mathbf{n}\}\}_\omega [v] \, ds + \sum_{e \in \mathcal{E}_\Gamma} \int_e \frac{\gamma}{H_e + h_e} [v][w] \, ds, \end{aligned}$$

where \mathbf{n} is the unit outward normal vector of Γ from K_1 to K_2 , and γ is the penalized parameter. Here

$$b_h^\varepsilon(x) := \rho(x) a^\varepsilon(x) + (1 - \rho(x)) \mathcal{A}_h(x) \quad (2.3)$$

with \mathcal{A}_h an approximation of \mathcal{A} . To step further, we assume that a^ε belongs to a set $\mathcal{M}(\lambda, A; D)$ defined by

$$\begin{aligned} \mathcal{M}(\lambda, A; D) := & \{a \in [L^\infty(D)]^{n \times n} : \xi \cdot a(x) \xi \geq \lambda |\xi|^2, \xi \cdot a(x) \xi \geq (1/A) |a(x) \xi|^2 \\ & \text{for any } \xi \in \mathbb{R}^n \text{ and a.e. } x \text{ in } D\}, \end{aligned}$$

where $|\cdot|$ denotes the Euclidean norm in \mathbb{R}^n . By the theory of H-convergence [47], we have $\mathcal{A} \in \mathcal{M}(\lambda, A; D)$. For any reasonable approximation \mathcal{A}_h , we may assume that the hybrid coefficient $b_h^\varepsilon \in \mathcal{M}(\lambda', A'; D)$ for certain positive constants λ' and A' . For example, if we use HMM [18, 1, 19, 52] to compute the effective matrix, then $\mathcal{A}_h \in \mathcal{M}(\lambda, A; D)$. Hence $b_h^\varepsilon \in \mathcal{M}(\lambda, A; D)$. If we use some other numerical upscaling methods, e.g., [20, 35, 21, 30], then $b_h^\varepsilon \in \mathcal{M}(\lambda', A'; D)$ with certain constants λ' and A' , which depend on λ and A , but not exactly the same. To quantify the approximation error for the effective matrix, we define $e(\text{HMM}) := \max_{x \in \bar{K}_2} \|(\mathcal{A} - \mathcal{A}_h)(x)\|_F$, where $\|\cdot\|_F$ is the Frobenius norm of a matrix.

The approximation problem is defined as: Find $v_h \in X_{h,H}^0$ such that

$$B_h^\varepsilon(v_h, w) = \langle f, w \rangle \quad \text{for all } w \in X_{h,H}^0. \quad (2.4)$$

3 Error Estimate for the Nitsche Hybrid Method

3.1 Accuracy for retrieving the macroscopic information

In this part, we estimate the error between the hybrid solution and the homogenized solution. For any $v \in X_{h,H}^0$, we define the broken energy norm as

$$\|v\| := \left(\sum_{i=1}^2 |v|_{1,K_i}^2 + \sum_{e \in \mathcal{E}_\Gamma} \frac{\gamma}{H_e + h_e} \| [v] \|_{0,e}^2 \right)^{1/2}. \quad (3.1)$$

The following lemma gives the continuity and coercivity of the bilinear form B_h^ε with respect to the above broken energy norm.

Lemma 1 Assume that b_h^ε is in $\mathcal{M}(\lambda', \lambda'; D)$. Let $\gamma_0 = 8\lambda'^2 C_{inv}^2 / \min(1, \lambda'^2)$, with C_{inv} in (3.4). If $\gamma \geq \gamma_0$, for all $v, w \in X_{h,H}$, then

$$|B_h^\varepsilon(v, w)| \leq 2 \max(1, \lambda') \|v\| \|w\|, \quad (3.2)$$

$$B_h^\varepsilon(v, v) \geq \min(1/2, \lambda'/2) \|v\|^2, \quad (3.3)$$

where the constant C_{inv} depends only on r and σ such that

$$\|v\|_{0,e} \leq C_{inv} |h_\tau|^{-1/2} \|v\|_{0,\tau} \quad \text{for all } v \in P_\tau(\tau), e \subset \partial\tau, \tau \in \mathcal{T}_{h,H}. \quad (3.4)$$

The existence and uniqueness of the solution of Problem (2.4) follow from the Lax-Milgram theorem provided that $\gamma \geq \gamma_0$. The proof is standard and we refer to [6] for details. The explicit form of C_{inv} may be found in [51], which will be used to determine the lower bound γ_0 of the penalized parameter γ .

The following inequality slightly extends [29, Lemma 3.1].

Lemma 2 For any $v \in H^s(D)$ with a positive number s , and for any subset $\Omega \subset D$, there exists C independent of Ω such that

$$\|v\|_{0,\Omega} \leq C |\Omega|^\theta \eta(\Omega) \|v\|_{s,D}, \quad (3.5)$$

where $\theta = \min(s/n, 1/2)$ and

$$\eta(\Omega) := \begin{cases} |\ln|\Omega||^{1/2} & \text{if } n = 2, 3 \text{ and } s = n/2, \\ 1 & \text{otherwise.} \end{cases}$$

Proof If $0 < s < n/2$, then we proceed along the same line that leads to [29, Lemma 3.1] and obtain (3.5) with $\theta = s/n$.

For $s > n/2$, we use the embedding $H^s(D) \hookrightarrow L^\infty(D)$ and obtain

$$\|v\|_{0,\Omega} \leq |\Omega|^{1/2} \|v\|_{L^\infty(\Omega)} \leq |\Omega|^{1/2} \|v\|_{L^\infty(D)} \leq C |\Omega|^{1/2} \|v\|_{s,D},$$

where C depends on D but independent of Ω .

It remains to deal with $s = n/2$. The case $n = 2, s = 1$ has been proved in [29, Lemma 3.1]. The case $n = 3, s = 3/2$ may be proved as follows. Using [33, Theorem 2.1]¹, for any $p > 2$, there exists C depending only on D such that

$$\|v\|_{L^p(D)} \leq C \sqrt{p} \|v\|_{3/2,D}.$$

Therefore,

$$\begin{aligned} \|v\|_{0,\Omega} &\leq |\Omega|^{1/2-1/p} \|v\|_{L^p(\Omega)} \leq |\Omega|^{1/2-1/p} \|v\|_{L^p(D)} \\ &\leq C |\Omega|^{1/2-1/p} \sqrt{p} \|v\|_{3/2,D}. \end{aligned}$$

Taking $p = |\ln|\Omega||$ in the right-hand side of the above inequality, we obtain (3.5) for $n = 3$ and $s = 3/2$.

A combination of the above cases implies (3.5).

¹ The authors therein only considered $D = \mathbb{R}^n$, while the proof may be extended to the domain D by Stein extension as in [3].

The following result concerns the accuracy of the method approximating the homogenized solution u_0 when u_0 is smooth.

Theorem 1 *Let u_0 and v_h be the solutions of Problem (1.3) and Problem (2.4), respectively. If $u_0 \in H^{1+s}(D)$ with $s > 1/2$, and $\psi_g \in H_0^1(D)$ is the unique solution of the adjoint variational problem:*

$$\langle \mathcal{A}\nabla v, \nabla \psi_g \rangle = \langle g, v \rangle \quad \text{for all } v \in H_0^1(D), \quad (3.6)$$

which satisfies the regularity estimate

$$\|\psi_g\|_{2,D} \leq C\|g\|_{0,D}. \quad (3.7)$$

Then there exists C depending on D, λ and Λ such that,

$$\|u_0 - v_h\| \leq C \left(\inf_{\chi \in X_{h,H}^0} \|u_0 - \chi\| + |K_1|^\theta \eta(K_1) + e(\text{HMM}) \right), \quad (3.8)$$

and

$$\begin{aligned} \|u_0 - v_h\|_{0,D} \leq C & \left(\inf_{\chi \in X_{h,H}^0} \|u_0 - \chi\| + |K_1|^\theta \eta(K_1) \right) \left(h + H + |K_1|^{1/n} \eta(K_1) \right) \\ & + Ce(\text{HMM}). \end{aligned} \quad (3.9)$$

Let χ be the Scott-Zhang interpolant [42] in (3.8) and (3.9) and we may obtain the following error estimates.

Corollary 1 *Under the same assumptions as in Theorem 1, there exists C depending on D, λ, Λ and σ such that*

$$\|u_0 - v_h\| \leq C \left(h^{s^*} + H^{s^*} + |K_1|^\theta \eta(K_1) + e(\text{HMM}) \right),$$

and

$$\begin{aligned} \|u_0 - v_h\|_{0,D} \leq C & \left(h^{s^*} + H^{s^*} + |K_1|^\theta \eta(K_1) \right) \left(h + H + |K_1|^{1/n} \eta(K_1) \right) \\ & + Ce(\text{HMM}), \end{aligned}$$

with $s^* = \min(s, r)$.

Proof of Theorem 1 Firstly, we need an auxiliary problem: Find $\tilde{u} \in X_{h,H}^0$ such that

$$A(\tilde{u}, v) = \langle f, v \rangle \quad \text{for all } v \in X_{h,H}^0, \quad (3.10)$$

where the bilinear form A is defined for any $v, w \in X_{h,H}^0$ the same as B_h^ε with b_h^ε replaced by \mathcal{A} . From the Galerkin orthogonality

$$A(u_0 - \tilde{u}, v) = 0 \quad \text{for any } v \in X_{h,H}^0, \quad (3.11)$$

and (3.2), (3.3) we have

$$\|u_0 - \tilde{u}\| \leq \frac{4 \max(\Lambda, 1)}{\min(1, \lambda)} \inf_{\chi \in X_{h,H}^0} \|u_0 - \chi\|, \quad (3.12)$$

and

$$\|u_0 - \tilde{u}\|_{0,D} \leq 2\Lambda \|u_0 - \tilde{u}\| \sup_{g \in L^2(D)} \frac{1}{\|g\|_{0,D}} \inf_{\chi \in X_{h,H}} \|\psi_g - \chi\|. \quad (3.13)$$

Let $w = \tilde{u} - v_h$, we write

$$\begin{aligned} B_h^\varepsilon(w, w) &= B_h^\varepsilon(\tilde{u}, w) - \langle f, w \rangle = B_h^\varepsilon(\tilde{u}, w) - A(\tilde{u}, w) \\ &= \sum_{i=1}^2 \int_{K_i} (b_h^\varepsilon - \mathcal{A}) \nabla \tilde{u} \nabla w \, dx - \sum_{e \in \mathcal{E}_\Gamma} \int_e \{ \{ (b_h^\varepsilon - \mathcal{A}) \nabla \tilde{u} \cdot \mathbf{n} \} \}_\omega [w] \, ds \\ &\quad - \sum_{e \in \mathcal{E}_\Gamma} \int_e \{ \{ (b_h^\varepsilon - \mathcal{A})^T \nabla w \cdot \mathbf{n} \} \}_\omega [\tilde{u}] \, ds. \end{aligned}$$

By $b_h^\varepsilon - \mathcal{A} = \rho(a^\varepsilon - \mathcal{A}) + (1 - \rho)(\mathcal{A}_h - \mathcal{A})$, we have, for any $x \in K_1$,

$$\| (b_h^\varepsilon - \mathcal{A})(x) \|_F \leq \| b_h^\varepsilon(x) \|_F + \| \mathcal{A}(x) \|_F \leq \Lambda' + \Lambda,$$

and for any $x \in K_2$,

$$\| (b_h^\varepsilon - \mathcal{A})(x) \|_F = \| (\mathcal{A}_h - \mathcal{A})(x) \|_F \leq e(\text{HMM}).$$

This implies

$$\left| \sum_{i=1}^2 \int_{K_i} (b_h^\varepsilon - \mathcal{A}) \nabla \tilde{u} \nabla w \, dx \right| \leq (\Lambda + \Lambda') |\tilde{u}|_{1,K_1} |w|_{1,K_1} + e(\text{HMM}) \|\tilde{u}\| \|w\|.$$

Noting that $\rho = 0$ for $e \in \mathcal{E}_\Gamma$, we obtain for any $x \in e$,

$$\| (b_h^\varepsilon - \mathcal{A})(x) \|_F = \| (\mathcal{A}_h - \mathcal{A})(x) \|_F \leq e(\text{HMM}).$$

Hence, we obtain

$$\begin{aligned} & \left| \sum_{e \in \mathcal{E}_\Gamma} \int_e \{ \{ (b_h^\varepsilon - \mathcal{A}) \nabla \tilde{u} \cdot \mathbf{n} \} \}_\omega [w] \, ds \right| \\ & \leq \left(\sum_{e \in \mathcal{E}_\Gamma} \frac{H_e + h_e}{\gamma} \| \{ \{ (b_h^\varepsilon - \mathcal{A}) \nabla \tilde{u} \cdot \mathbf{n} \} \}_\omega \|_{0,e}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_\Gamma} \frac{\gamma}{H_e + h_e} \| [w] \|_{0,e}^2 \right)^{1/2} \\ & \leq \frac{e(\text{HMM})}{\gamma_0} \left(\sum_{e \in \mathcal{E}_\Gamma} (H_e + h_e) \| \{ \{ \nabla \tilde{u} \} \}_\omega \|_{0,e}^2 \right)^{1/2} \|w\|. \end{aligned}$$

Using Cauchy-Schwarz inequality and (3.4), we obtain

$$\begin{aligned} \sum_{e \in \mathcal{E}_\Gamma} (H_e + h_e) \|\{\{\nabla \tilde{u}\}\}_\omega\|_{0,e}^2 &\leq \sum_{e \in \mathcal{E}_\Gamma} (H_e + h_e) (\omega_1 \|\nabla \tilde{u}_1\|_{0,e}^2 + \omega_2 \|\nabla \tilde{u}_2\|_{0,e}^2) \\ &= \sum_{e \in \mathcal{E}_h} h_e \|\nabla \tilde{u}_1\|_{0,e} + \sum_{e \in \mathcal{E}_H} H_e \|\nabla \tilde{u}_2\|_{0,e}^2 \\ &\leq C_{\text{inv}} \left(\sum_{\tau \in \mathcal{T}_h^\Gamma} \|\nabla \tilde{u}_1\|_{0,\tau}^2 + \sum_{\tau \in \mathcal{T}_H^\Gamma} \|\nabla \tilde{u}_2\|_{0,\tau}^2 \right). \end{aligned}$$

A combination of the above two inequalities gives

$$\left| \sum_{e \in \mathcal{E}_\Gamma} \int_e \{(b_h^\varepsilon - \mathcal{A}) \nabla \tilde{u} \cdot \mathbf{n}\}_\omega [w] \, ds \right| \leq \frac{C_{\text{inv}}}{\gamma_0} e(\text{HMM}) \|\tilde{u}\| \|w\|.$$

Exchanging the role of \tilde{u} and w , we obtain

$$\left| \sum_{e \in \mathcal{E}_\Gamma} \int_e \{(b_h^\varepsilon - \mathcal{A})^T \nabla w \cdot \mathbf{n}\}_\omega [\tilde{u}] \, ds \right| \leq \frac{C_{\text{inv}}}{\gamma_0} e(\text{HMM}) \|\tilde{u}\| \|w\|.$$

Combining all the above estimates, we obtain

$$B_h^\varepsilon(w, w) \leq (\Lambda + \Lambda') |\tilde{u}|_{1,K_1} |w|_{1,K_1} + (1 + 2C_{\text{inv}}/\gamma_0) e(\text{HMM}) \|\tilde{u}\| \|w\|. \quad (3.14)$$

Using the triangle inequality and (3.5), we obtain

$$\begin{aligned} \|\tilde{u}\|_{1,K_1} &\leq \|u_0 - \tilde{u}\|_{1,K_1} + \|u_0\|_{1,K_1} \\ &\leq \|u_0 - \tilde{u}\| + C |K_1|^\theta \eta(K_1) \|u_0\|_{1+s,D}. \end{aligned} \quad (3.15)$$

Combining the above three inequalities and the a-priori estimate $\|\tilde{u}\| \leq C \|f\|_{-1,D}$, we obtain

$$\|\tilde{u} - v_h\| \leq C (\|u_0 - \tilde{u}\| + |K_1|^\theta \eta(K_1) \|u_0\|_{1+s,D} + e(\text{HMM}) \|f\|_{-1,D}),$$

which together with (3.12) and the triangle inequality conclude the estimate (3.8).

We exploit *Aubin-Nitsche's dual argument* to prove the L^2 error bound.

For any $g \in L^2(D)$, let $\varphi_g \in H_0^1(D)$ satisfying

$$B_h^\varepsilon(v, \varphi_g) = \langle g, v \rangle \quad \text{for all } v \in X_{h,H}.$$

Substituting $v = \tilde{u} - v_h$ into the above equation, we obtain

$$\begin{aligned} \langle g, \tilde{u} - v_h \rangle &= B_h^\varepsilon(\tilde{u} - v_h, \varphi_g) = B_h^\varepsilon(\tilde{u}, \varphi_g) - \langle f, \varphi_g \rangle \\ &= B_h^\varepsilon(\tilde{u}, \varphi_g) - A(\tilde{u}, \varphi_g). \end{aligned}$$

Proceeding along the same line that leads to (3.14), we obtain

$$|\langle g, \tilde{u} - v_h \rangle| \leq C (|\tilde{u}|_{1,K_1} |\varphi_g|_{1,K_1} + e(\text{HMM}) \|\tilde{u}\| \|\varphi_g\|).$$

Using (3.15) and the a-priori estimates for \tilde{u} and φ_g , we obtain

$$\begin{aligned} |\langle g, \tilde{u} - v_h \rangle| &\leq C|K_1|^{1/n}\eta(K_1)\|g\|_{0,D} (\|u_0 - \tilde{u}\| + |K_1|^\theta\eta(K_1)\|u_0\|_{1+s,D}) \\ &\quad + Ce(\text{HMM})\|f\|_{-1,D}\|g\|_{-1,D}. \end{aligned} \quad (3.16)$$

Then

$$\begin{aligned} \|\tilde{u} - v_h\|_{0,D} &\leq C|K_1|^{1/n}\eta(K_1) (\|u_0 - \tilde{u}\| + |K_1|^\theta\eta(K_1)\|u_0\|_{1+s,D}) \\ &\quad + Ce(\text{HMM})\|f\|_{-1,D}. \end{aligned} \quad (3.17)$$

Combining the above inequality with (3.12), (3.13) and using the triangle inequality, we obtain (3.9). \square

In the following, we shall consider the case that u_0 is not smooth, which covers the case of a rough homogenized coefficient \mathcal{A} , or a point load function f , or a nonconvex domain D . If $u_0 \in H^{1+s}(D)$ with $0 < s \leq 1/2$, the Galerkin orthogonality (3.11) is invalid and we cannot use Céa's lemma to prove (3.12). To overcome this difficulty, we employ the **medius** analysis in [23]. We firstly need the following extra assumptions on $\mathcal{T}_{h,H}$:

Assumption A: \mathcal{T}_h^Γ is quasi-uniform, i.e., there exists a constant ν independent of h_Γ such that for any $\tau \in \mathcal{T}_h^\Gamma$, $h_\tau \geq \nu h_\Gamma$, where $h_\Gamma := \max\{h_\tau : \tau \in \mathcal{T}_h^\Gamma\}$.

Assumption B: \mathcal{E}_h is a subgrid of \mathcal{E}_H , i.e., $\mathcal{E}_\cap = \mathcal{E}_h$.

According to **Assumption A** and **Assumption B**, we may construct a compatible sub-decomposition $\tilde{\mathcal{T}}_H^\Gamma$ out of \mathcal{T}_H^Γ such that $\tilde{\mathcal{T}}_H^\Gamma \cup \mathcal{T}_h^\Gamma$ is a conforming mesh, which is quasi-uniform near the interface Γ , while we never need such refined mesh in the implementation; see Fig. 4. Both assumptions have been used in [31] to prove the a-posteriori error estimate for the DG method.

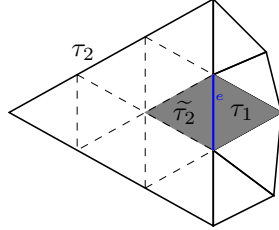


Fig. 4 The split of the element in \mathcal{T}_H^Γ .

We define the oscillation for $f \in L^2(D)$ as

$$\text{osc}(f) := \left[\sum_{\tau \in \mathcal{T}_{h,H}} h_\tau^2 \left(\inf_{\bar{f} \in P_0(\tau)} \|f - \bar{f}\|_{0,\tau}^2 \right) \right]^{1/2},$$

and the oscillation for $\mathcal{A} \in [L^\infty(D)]^{n \times n}$ as

$$\text{osc}(\mathcal{A}) := \max_{\tau \in \mathcal{T}_{h,H}} \left(\inf_{\bar{\mathcal{A}} \in [P_0(\tau)]^{n \times n}} \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau} \right).$$

By Meyers' regularity result [37], there exists $p_0 > 2$ that depends on D , Λ and λ , such that for all $p \leq p_0$,

$$|u_0|_{1,p,D} \leq C \|f\|_{-1,p,D},$$

where C depends on n , D , Λ and λ . Using Hölder's inequality, we obtain

$$\begin{aligned} |u_0|_{1,K_1} &\leq |K_1|^{1/2-1/p} |u_0|_{1,p,K_1} \leq |K_1|^{1/2-1/p} |u_0|_{1,p,D} \\ &\leq C |K_1|^{1/2-1/p} \|f\|_{-1,p,D}. \end{aligned} \quad (3.18)$$

Theorem 2 *Let u_0 and v_h be the solutions of Problem (1.3) and Problem (2.4), respectively. If **Assumption A** and **Assumption B** are valid, then there exists C depending on D , λ , Λ and C_{inv} such that for any $2 < p < p_0$,*

$$\begin{aligned} \|u_0 - v_h\| &\leq C \varrho \left(\inf_{\chi \in X_{h,H}^0} \|u_0 - \chi\| + \text{osc}(\mathcal{A}) + \text{osc}(f) \right) \\ &\quad + C \left(e(\text{HMM}) + |K_1|^{1/2-1/p} \right). \end{aligned} \quad (3.19)$$

Under the same assumptions, for any $2 < p < \tilde{p}_0$ with

$$\tilde{p}_0 = \begin{cases} p_0 & \text{if } n = 2, \\ \min(p_0, 6) & \text{if } n = 3, \end{cases} \quad (3.20)$$

there holds (it holds that?)

$$\begin{aligned} &\|u_0 - v_h\|_{0,D} \\ &\leq C \left(\varrho^2 \left(\inf_{\chi \in X_{h,H}^0} \|u_0 - \chi\| + \text{osc}(f) + \text{osc}(\mathcal{A}) \right) + |K_1|^{1/2-1/p} \right) \\ &\quad \times \left(\sup_{g \in L^2(D)} \frac{1}{\|g\|_{0,D}} \inf_{\tilde{\chi} \in X_{h,H}} \left(\sum_{\tau \in \mathcal{T}_{h,H}} h_\tau^{-1} \|\psi_g - \tilde{\chi}\|_{0,\tau} + \|\psi_g - \tilde{\chi}\| + \text{osc}(g) \right) \right. \\ &\quad \left. + \text{osc}(\mathcal{A}) + |K_1|^{1/2-1/p} \right) + Ce(\text{HMM}), \end{aligned} \quad (3.21)$$

where $\varrho = (H + h)/(\gamma h_\Gamma)$ and $\psi_g \in H_0^1(D)$ is the unique solution of Problem (3.6).

We do not impose the regularity assumption on ψ_g . The proof of this theorem is a combination of the ways that lead to Theorem 1 and Lemma 3 below, provided that we replace (3.5) by (3.18) and replace (3.16) by

$$\begin{aligned} |\langle g, \tilde{u} - v_h \rangle| &\leq C |K_1|^{1/2-1/p} \left(\|u_0 - \tilde{u}\| + |K_1|^{1/2-1/p} \|f\|_{-1,p,D} \right) \|g\|_{-1,p,D} \\ &\quad + Ce(\text{HMM}) \|f\|_{-1,D} \|g\|_{-1,D}, \end{aligned}$$

respectively.

It follows from the Sobolev imbedding $L^2(D) \hookrightarrow W^{-1,p}(D)$ for any $p \geq 1$ if $n = 2$ and for any $1 \leq p \leq 6$ if $n = 3$ [3], the estimate (3.17) is changed to

$$\begin{aligned} \|\tilde{u} - v_h\|_{0,D} &\leq C |K_1|^{1/2-1/p} \left(\|u_0 - \tilde{u}\| + |K_1|^{1/2-1/p} \|f\|_{-1,p,D} \right) \\ &\quad + e(\text{HMM}) \|f\|_{-1,D} \quad \text{for any } 2 < p < \tilde{p}_0. \end{aligned}$$

Lemma 3 *Under the same assumptions of Theorem 2, there exists C depending on D, λ, Λ and C_{inv} such that*

$$\|u_0 - \tilde{u}\| \leq C \varrho \left(\inf_{\chi \in X_{h,H}^0} \|u_0 - \chi\| + \text{OSC}(f) + \text{OSC}(\mathcal{A}) \right), \quad (3.22)$$

and

$$\begin{aligned} \|u_0 - \tilde{u}\|_{0,D} &\leq C (\varrho \|u_0 - \tilde{u}\| + \text{OSC}(f) + \text{OSC}(\mathcal{A})) \\ &\quad \times \left(\sup_{g \in L^2(D)} \frac{1}{\|g\|_{0,D}} \left(\inf_{\tilde{\chi} \in X_{h,H}} \left(\sum_{\tau \in \mathcal{T}_{h,H}} h_\tau^{-1} \|\psi_g - \tilde{\chi}\|_{0,\tau} + \|\psi_g - \tilde{\chi}\| \right) \right. \right. \\ &\quad \left. \left. + \text{OSC}(g) \right) + \text{OSC}(\mathcal{A}) \right), \end{aligned} \quad (3.23)$$

where $\psi_g \in H_0^1(D)$ is the unique solution of Problem (3.6).

Remark 1 By contrast to (3.12) and (3.13), there is no regularity assumption on u_0 in Lemma 3, while it contains extra oscillation terms $\text{OSC}(\mathcal{A})$, $\text{OSC}(f)$ and $\text{OSC}(g)$. Such terms are indispensable by the recent error estimates on Nitsche's methods [36, 25]. We also note that the right hand side of (3.22) and (3.23) depend on ϱ , which seems unpleasant at the first glance because the error blows up for large ϱ , while it may be small or even harmless if we tune the penalized parameter γ . This will be shown in Corollary 2 and § 5. It would be very interesting to know whether such dependance can be removed. We shall leave this for further study.

The proof of the above lemma is quite involved and is of independent interest, we postpone it to §4.

It follows from Theorem 2 that

Corollary 2 *If $\mathcal{A} \in [C^{0,\alpha}(D)]^{n \times n}$ with $\alpha \in (0, 1)$, if we let $\gamma > \max((H + h)/h_\Gamma, \gamma_0)$, and if $u_0 \in H^{1+s}(D)$ with $0 < s \leq 1/2$ and $f \in L^2(D)$, then*

$$\begin{aligned} \|u_0 - v_h\| &\leq C \left(h^\beta + H^\beta + |K_1|^{s/n} \eta(K_1) + e(\text{HMM}) \right), \\ \|u_0 - v_h\|_{0,D} &\leq C \left(h^{2\beta} + H^{2\beta} + |K_1|^{2s/n} \eta^2(K_1) + e(\text{HMM}) \right), \end{aligned} \quad (3.24)$$

where $\beta = \min(\alpha, s)$.

Proof If $\gamma > \max((H + h)/h_\Gamma, \gamma_0)$, then ϱ is a constant which is independent of any other parameters.

If $\mathcal{A} \in [C^{0,\alpha}(D)]^{d \times d}$, then there exists C independent of h and H such that $\text{osc}(\mathcal{A}) \leq C(h^\alpha + H^\alpha)$. If $f \in L^2(D)$, then $\text{osc}(f) \leq (h + H) \|f\|_{0,D}$.

Let χ be the Scott-Zhang interpolant of u_0 in (3.19) and we obtain the first estimate of (3.24).

For any given $g \in L^2(D)$, let χ and $\tilde{\chi}$ be the Scott-Zhang interpolant of u_0 and ϕ_g in (3.21) respectively and we obtain the second estimate of (3.24).

3.2 Accuracy for retrieving the local microscopic information

We assume that $\text{dist}(K_0, \Gamma) = d \geq \kappa h$ for a sufficiently large $\kappa > 0$. For a subset $B \subset D$, we define

$$\begin{aligned} H_{<}^1(B) &:= \{u \in H^1(D) \mid u|_{D \setminus B} = 0\}, \quad \text{and} \\ X_{<}(B) &:= \{u \in X_{h,H} \mid u|_{D \setminus B} = 0\}. \end{aligned}$$

Let G_1 and G be subsets of K_1 with $G_1 \subset G$ and $\text{dist}(G_1, \partial G \setminus \partial D) = \tilde{d} > 0$. In order to prove the localized energy error estimate, we state that some properties of the standard Lagrange finite element space X_h hold [14]:

1. Local interpolant: There exists a local interpolant Iu such that for any $u \in H_{<}^1(G_1)$, $Iu \in X_{<}(G)$.
2. Inverse properties: For each $\chi \in X_{<}(K_1)$, and $\tau \in \mathcal{T}_h$, $1 \leq p \leq q \leq \infty$, and $0 \leq t \leq s \leq r + 1$,

$$\|\chi\|_{s,p,\tau} \leq C h_\tau^{t-s+n/p-n/q} \|\chi\|_{t,p,\tau}. \quad (3.25)$$

3. Superapproximation: Let $\omega \in C^\infty(K_1) \cap H_{<}^1(G_1)$ with $|\omega|_{j,\infty,K_1} \leq C \tilde{d}^{-j}$ for integers $0 \leq j \leq r + 1$ for each $\chi \in X_{<}(G) \cap X_h^0$ and for each $\tau \in \mathcal{T}_h$ satisfying $h_\tau \leq \tilde{d}$,

$$\|\omega^2 \chi - I(\omega^2 \chi)\|_{0,\tau} \leq C \left(\frac{h_\tau}{\tilde{d}} |\omega \chi|_{1,\tau} + \frac{h_\tau}{\tilde{d}^2} \|\chi\|_{0,\tau} \right). \quad (3.26)$$

Theorem 3 Let u^ε and v_h be the solutions of (1.1) and (3.10) respectively. Let $K_0 \subset K_1 \subset D$ be given, and let $\text{dist}(K_0, \Gamma) = d$. Let the above properties hold with $\tilde{d} = d/16$, in addition, let $\max_{\tau \cap \Gamma \neq \emptyset} h_\tau/d \leq 1/16$. Then

$$|u^\varepsilon - v_h|_{1, K_0} \leq C \left(\inf_{\chi \in X_h^0} |u^\varepsilon - \chi|_{1, K_1} + d^{-1} \|u^\varepsilon - v_h\|_{0, K_1} \right), \quad (3.27)$$

where C depends only on λ', A', λ, A , and D .

Proof Let a subset $\tilde{K} \subset K_1$ satisfying $\text{dist}(\tilde{K}, \Gamma) = d_0 > 0$. We have $X_{<}(\tilde{K}) \subset H_{<}^1(K_1)$. Then for any $v, w \in H_{<}^1(K_1)$, the bilinear form $B_h^\varepsilon(v, w)$ degenerates to $\langle b_h^\varepsilon v, w \rangle$. Thus, the proof of above theorem is the same with that in [29, Theorem 3.2], and we omit the details for brevity.

Using the triangle inequality, we have

$$|u^\varepsilon - v_h|_{1, K_0} \leq C \left(\inf_{\chi \in X_{h, H}^0} |u^\varepsilon - \chi|_{1, K_1} + d^{-1} (\|u_0 - v_h\|_{0, D} + \|u_0 - u^\varepsilon\|_{0, D}) \right).$$

The first term in the right-hand side of the above inequality concerns how the local events are resolved. Accurate approximation requires a highly refined mesh, which is allowed by Theorem 3; using Corollary 1 and Corollary 2, we may bound $d^{-1} \|u_0 - v_h\|_{0, D}$; the last term converges to zero as ε tends to zero by H-convergence theory [47] for any bounded and measurable a^ε . Therefore, the method converges for Problem (1.1) with bounded and measurable coefficient. Moreover, if we assume more structural conditions such as periodicity, almost-periodicity or stochastic ergodicity on a^ε , we may expect $\|u^\varepsilon - u_0\|_{0, D} \simeq \mathcal{O}(\varepsilon^\alpha)$ for certain $\alpha > 0$. We refer to [32, 43, 5] for extensive discussions for such L^2 -estimate.

4 Proof of Lemma 3

To prove Lemma 3, we employ the **medius** analysis in [23]. We need an enriching operator $E_h : X_{h, H} \rightarrow X_{h, H} \cap H^1(D)$. The construction of E_h is similar to that in [31] with certain modifications.

Let $\mathcal{N}(\tau)$ be the set of all nodes of $\tau \in \mathcal{T}_{h, H}$. Define $\mathcal{N}_H := \bigcup_{\tau \in \mathcal{T}_H} \mathcal{N}(\tau)$, $\mathcal{N}_h = \bigcup_{\tau \in \mathcal{T}_h} \mathcal{N}(\tau)$, $\mathcal{N}_\Gamma := \{p \in \mathcal{N}_h \cup \mathcal{N}_H \mid p \in \Gamma\}$ and $\mathcal{N}_\Gamma^0 := \mathcal{N}_H \cap \mathcal{N}_\Gamma$.

Lemma 4 *If Assumption A and Assumption B hold for $n = 2, 3$, there exists a linear map $E_h : X_{h, H} \rightarrow X_{h, H} \cap H^1(D)$ satisfying*

$$\sum_{\tau \in \mathcal{T}_{h, H}^\Gamma} h_\tau^{2m-2} |v - E_h v|_{m, \tau}^2 \leq C \sum_{e \in \mathcal{E}_\Gamma} h_e^{-1} \|\llbracket v \rrbracket\|_{0, e}^2 \quad \text{for all } m \leq r, \quad (4.1)$$

where the constant C is independent of h, H and H/h .

Proof See Appendix A.

A direct consequence of the above local enriching estimate is the following

Corollary 3 *If Assumption A and Assumption B are true, then there exists C independent of h, H, γ and the mesh ratio H/h such that*

$$\sum_{\tau \in \mathcal{T}_{h,H}^r} h_\tau^{2m-2} |v - E_h v|_{m,\tau}^2 \leq C \varrho \|v\|^2, \quad 0 \leq m \leq r, \quad (4.2)$$

and

$$\|v - E_h v\|^2 \leq C \varrho \|v\|^2, \quad (4.3)$$

and

$$\sum_{e \in \mathcal{E}_\Omega} h_e^{-1} \| \{ \{ v - E_h v \} \}^\omega \|_{0,e}^2 \leq C \varrho \|v\|^2, \quad (4.4)$$

$$\sum_{e \in \mathcal{E}_\Omega} (H_e + h_e) | \{ \{ v - E_h v \} \} |_{1,e}^2 \leq C \varrho \|v\|^2. \quad (4.5)$$

Proof Using **Assumption A**, we have

$$\sum_{e \in \mathcal{E}_\Omega} h_e^{-1} \| [v] \|_{0,e}^2 \leq \sum_{e \in \mathcal{E}_\Omega} \frac{H_e + h_e}{\gamma h_e} \frac{\gamma}{h_e + H_e} \| [v] \|_{0,e}^2 \leq \varrho \|v\|^2,$$

Therefore, the estimate (4.2) is a direct consequence of (4.1).

Since $v - E_h v \equiv 0$ in $\mathcal{T}_{h,H} \setminus \mathcal{T}_{h,H}^r$, using the above inequality and (4.2) with $m = 1$, we obtain (4.3).

For any $v = (v_1, v_2) \in X_h \times X_H$, by definition, we obtain

$$\sum_{e \in \mathcal{E}_\Omega} h_e^{-1} \| \{ \{ v - E_h v \} \}^\omega \|_{0,e}^2 \leq \sum_{e \in \mathcal{E}_\Omega} (h_e^{-1} \|v_1 - E_h v\|_{0,e}^2 + h_e^{-1} \omega_1^2 \|v_2 - E_h v\|_{0,e}^2).$$

Using **Assumption A**, we have, for any $e \in \mathcal{E}_\Omega$,

$$h_e^{-1} \omega_1^2 \leq h_e \omega_1 \leq H_e^{-1}.$$

Using the trace inequality, there exists C such that

$$\begin{aligned} \sum_{e \in \mathcal{E}_\Omega} h_e^{-1} \| \{ \{ v - E_h v \} \}^\omega \|_{0,e}^2 &\leq \sum_{e \in \mathcal{E}_h} h_e^{-1} |v_1 - E_h v|_{1,e}^2 + \sum_{E \in \mathcal{E}_H} H_E^{-1} |v_2 - E_h v|_{1,E}^2 \\ &\leq C \sum_{\tau \in \mathcal{T}_{h,H}^r} (h_\tau^{-2} \|v - E_h v\|_{0,\tau}^2 + |v - E_h v|_{1,\tau}^2), \end{aligned}$$

which together with (4.2) implies (4.4).

Proceeding along the same line, we obtain

$$\begin{aligned} \sum_{e \in \mathcal{E}_\Omega} (H_e + h_e) | \{ \{ v - E_h v \} \} |_{1,e}^2 &\leq \sum_{e \in \mathcal{E}_h} h_e |v_1 - E_h v|_{1,e}^2 + \sum_{E \in \mathcal{E}_H} H_E |v_2 - E_h v|_{1,E}^2 \\ &\leq C \sum_{\tau \in \mathcal{T}_{h,H}^r} (|v - E_h v|_{1,\tau} + h_\tau^2 |v - E_h v|_{2,\tau}), \end{aligned}$$

which together with (4.2) implies (4.5).

Next lemma concerns the error estimate of \tilde{u} approximating u_0 , which is a type of *Strang's lemma*, the proof follows from the same line of [23, Lemma 2.1 and § 3.2] and we omit the details.

Lemma 5 *There exists a constant C such that*

$$\|u_0 - \tilde{u}\| \leq C \inf_{\chi \in X_{h,H}^0} \left(\|u_0 - \chi\| + \sup_{\psi \in X_{h,H}^0 \setminus \{0\}} \frac{\langle f, \psi - E_h \psi \rangle - A(\chi, \psi - E_h \psi)}{\|\psi\|} \right).$$

To estimate the second term in the right hand side of the above inequality, we shall use the following lemma, which is similar to the discrete local efficiency estimates in the a-posteriori error analysis [49]. The proof is quite standard and we refer to [50].

Lemma 6 *Let $u_0 \in H_0^1(D)$ be the solution of (1.3) and $\chi \in X_{h,H}^0$.*

1. *For each $\tau \in \mathcal{T}_{h,H}$, let $\bar{\mathcal{A}} \in [P_0(\tau)]^{n \times n}$, there exists C such that*

$$\begin{aligned} h_\tau^2 \|f + \nabla \cdot \bar{\mathcal{A}} \nabla \chi\|_{0,\tau}^2 &\leq C \left(|u_0 - \chi|_{1,\tau}^2 + h_\tau^2 \inf_{\bar{f} \in P_0(\tau)} \|f - \bar{f}\|_{0,\tau}^2 \right. \\ &\quad \left. + \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau}^2 |u_0|_{1,\tau}^2 \right). \end{aligned} \quad (4.6)$$

2. *For each $e \in \mathcal{E}_\Gamma$, there exists $(\tau_1, \tilde{\tau}_2) \in \mathcal{T}_h^\Gamma \times \tilde{\mathcal{T}}_H^\Gamma$ such that $\tau_1 \cap \tilde{\tau}_2 = e$ and e is a full edge of both τ_1 and $\tilde{\tau}_2$; See Fig. 4. Let $\bar{\mathcal{A}} \in [P_0(\tau_1) \times P_0(\tilde{\tau}_2)]^{n \times n}$, we have*

$$\begin{aligned} h_e \|\llbracket \bar{\mathcal{A}} \nabla \chi \cdot \mathbf{n} \rrbracket\|_{0,e}^2 &\leq C \sum_{\tau \in \{\tau_1, \tilde{\tau}_2\}} \left(\|u_0 - \chi\|_{0,\tau}^2 + h_\tau^2 \inf_{\bar{f} \in P_0(\tau)} \|f - \bar{f}\|_{0,\tau}^2 \right. \\ &\quad \left. + \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau}^2 |u_0|_{1,\tau}^2 \right). \end{aligned} \quad (4.7)$$

Combining all the above estimates, we are ready to prove Lemma 3.

Proof of Lemma 3 On each element $\tau \in \mathcal{T}_{h,H}$, let $\bar{\mathcal{A}} \in [P_0(\tau)]^{n \times n}$. We define a bilinear form $\bar{A}(\cdot, \cdot)$ the same as $A(\cdot, \cdot)$ with \mathcal{A} replaced by $\bar{\mathcal{A}}$. For any $\psi \in X_{h,H}^0$, using (2.2), we obtain

$$\begin{aligned} &\langle f, \psi - E_h \psi \rangle - A(\chi, \psi - E_h \psi) \\ &= \langle f, \psi - E_h \psi \rangle - \bar{A}(\chi, \psi - E_h \psi) + \bar{A}(\chi, \psi - E_h \psi) - A(\chi, \psi - E_h \psi) \\ &= \sum_{\tau \in \mathcal{T}_{h,H}} \int_\tau (f + \nabla \cdot \bar{\mathcal{A}} \nabla \chi)(\psi - E_h \psi) \, dx - \sum_{e \in \mathcal{E}_\Gamma} \int_e \llbracket \bar{\mathcal{A}} \nabla \chi \cdot \mathbf{n} \rrbracket \{\{\psi - E_h \psi\}\}^\omega \, ds \\ &\quad + \sum_{e \in \mathcal{E}_\Gamma} \int_e \{\{\bar{\mathcal{A}}^T \nabla(\psi - E_h \psi) \cdot \mathbf{n}\}\}^\omega \llbracket u_0 - \chi \rrbracket \, ds \\ &\quad - \sum_{e \in \mathcal{E}_\Gamma} \int_e \frac{\gamma}{H_e + h_e} \llbracket u_0 - \chi \rrbracket \llbracket \psi - E_h \psi \rrbracket \, ds + (\bar{A}(\chi, \psi - E_h \psi) - A(\chi, \psi - E_h \psi)) \\ &= I_1 + \dots + I_5. \end{aligned}$$

Using (4.6) and (4.2) with $m = 0$, we obtain

$$\begin{aligned} |I_1| &\leq \left(\sum_{\tau \in \mathcal{T}_{h,H}} h_\tau^2 \|f + \nabla \cdot \bar{\mathcal{A}} \nabla \chi\|_{0,\tau}^2 \right)^{1/2} \left(\sum_{\tau \in \mathcal{T}_{h,H}^r} h_\tau^{-2} \|\psi - E_h \psi\|_{0,\tau}^2 \right)^{1/2} \\ &\leq C \varrho \left(\inf_{\chi \in X_{h,H}^0} |u_0 - \chi|_{1,D} + \text{OSC}(f) + \text{OSC}(\mathcal{A}) \right) \|\psi\|. \end{aligned}$$

Using (4.7) and (4.4), we obtain

$$\begin{aligned} |I_2| &\leq \left(\sum_{e \in \mathcal{E}_\Gamma} h_e \|\llbracket \bar{\mathcal{A}} \nabla \chi \cdot \mathbf{n} \rrbracket\|_{0,e}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_\Gamma} h_e^{-1} \|\{\{\psi - E_h \psi\}\}^\omega\|_{0,e}^2 \right)^{1/2} \\ &\leq C \varrho \left(\inf_{\chi \in X_{h,H}^0} |u_0 - \chi|_{1,D} + \text{OSC}(f) + \text{OSC}(\mathcal{A}) \right) \|\psi\|. \end{aligned}$$

Using (4.5), we obtain

$$\begin{aligned} |I_3| &\leq C \left(\sum_{e \in \mathcal{E}_\Gamma} \frac{H_e + h_e}{\gamma} |\{\{\psi - E_h \psi\}\}|_{1,e}^2 \right)^{1/2} \left(\sum_{e \in \mathcal{E}_\Gamma} \frac{\gamma}{H_e + h_e} \|\llbracket u_0 - \chi \rrbracket\|_{0,e}^2 \right)^{1/2} \\ &\leq C(\varrho/\gamma_0) \|\llbracket u_0 - \chi \rrbracket\| \|\psi\|. \end{aligned}$$

Using (4.3), we may bound the last two terms as

$$|I_4| \leq C \|\llbracket u_0 - \chi \rrbracket\| \|\psi\|,$$

and

$$|I_5| \leq C \text{OSC}(\mathcal{A}) \|\chi\| \|\psi - E_h \psi\| \leq C \varrho \text{OSC}(\mathcal{A}) (\|u_0 - \chi\| + |u_0|_{1,D}) \|\psi\|.$$

Combining all the above estimates and using Lemma 5, we obtain (3.22).

Proceeding along the same line that leads to [23, Theorem 4.4], we obtain L^2 estimate (3.23), and we postpone the proof to **Appendix B**. \square

5 Numerical Experiments

In this part, we report two examples with different shapes of defects to demonstrate the accuracy and efficiency of the method. The governing equation is (1.1) with domain $D = (0, 1)^2$, $f = 1$ and the homogeneous Dirichlet boundary condition is imposed on ∂D . The finite element solvers are carried on FreeFem++ toolbox [26]².

The first example is taken from [29] with a two-scale coefficient.

² <https://freefem.org/>

Example 1 Let

$$a^\varepsilon(x) = \frac{(R_1 + R_2 \sin(2\pi x_1))(R_1 + R_2 \cos(2\pi x_2))}{(R_1 + R_2 \sin(2\pi x_1/\varepsilon))(R_1 + R_2 \sin(2\pi x_2/\varepsilon))} I, \quad (5.1)$$

where I is a two by two identity matrix. The effective matrix is given by

$$\mathcal{A}(x) = \frac{(R_1 + R_2 \sin(2\pi x_1))(R_1 + R_2 \cos(2\pi x_2))}{R_1 \sqrt{R_1^2 - R_2^2}} I. \quad (5.2)$$

In the simulation we let $R_1 = 2.5$, $R_2 = 1.5$ and $\varepsilon = 0.01$. The reference solutions for u^ε and u_0 are obtained by solving Problem (1.1) and (1.3) with linear element over a uniform mesh with the mesh size around $3.33e - 4$.

The second example is taken from [2] and the coefficient has no clear scale separation inside K_0 .

Example 2 For a subset $B \subset D$, we define δ_B as the characteristic function for the set B . The setup for this example is the same with the first one except that the coefficient is replaced by $a^\varepsilon = \delta_{K_0} \tilde{a} + (1 - \delta_{K_0}) \tilde{a}^\varepsilon$, where

$$\tilde{a}(x) = \left(3 + \frac{1}{7} \sum_{j=0}^4 \sum_{i=1}^j \frac{1}{j+1} \cos \left(\left[8 \left(ix_2 - \frac{x_1}{i+1} \right) \right] + [150ix_1] + [150x_2] \right) \right) I,$$

and

$$\tilde{a}^\varepsilon(x) = (2.1 + \cos(2\pi x_1/\varepsilon) \cos(2\pi x_2/\varepsilon) + \sin(4x_1^2 x_2^2)) I.$$

We let $\varepsilon = 0.0063$ in the simulation.

By [29], the effective matrix $\mathcal{A} = \delta_{K_0} \tilde{a} + (1 - \delta_{K_0}) \tilde{A}$, where \tilde{A} is the effective matrix associated with \tilde{a}^ε . Since there is no analytical formula for \tilde{A} , we use the least-squares reconstruction method in [35, 30] to obtain a higher-order approximation to \tilde{A} so that $e(\text{HMM})$ is negligible. The approximation to \tilde{A} is denoted by \tilde{A}_h . The homogenized solution u_0 is computed by solving Problem (1.3) with $\mathcal{A}_h = \delta_{K_0} \tilde{a} + (1 - \delta_{K_0}) \tilde{A}_h$. The hybrid coefficient is

$$b_h^\varepsilon = \delta_{K_0} \tilde{a} + (1 - \delta_{K_0}) (\rho \tilde{a}^\varepsilon + (1 - \rho) \tilde{A}_h).$$

We use linear element to compute u^ε and u_0 over a refined 3000×3000 mesh as the reference solution.

For both examples, our major interests are the following two quantities:

$$e(u_0): = \frac{|u_0 - v_h|_{1,K_2}}{|u_0|_{1,K_2}} \quad \text{and} \quad e(u^\varepsilon): = \frac{|u^\varepsilon - v_h|_{1,K_0}}{|u^\varepsilon|_{1,K_0}}. \quad (5.3)$$

5.1 The choice of the penalized factor γ

Based on the explicit expressions of the inverse trace inequalities in [51], the lower bound γ_0 in Lemma 1 for the penalized factor may be bounded by

$$\gamma_0 \leq \begin{cases} \frac{16\sqrt{3}}{9} \frac{r(r+1)\sigma A'^2}{\min\{1, \lambda'^2\}} & n = 2, \\ \frac{32\sqrt{3}}{27} \frac{r(r+2)\sigma A'^2}{\min\{1, \lambda'^2\}} & n = 3. \end{cases} \quad (5.4)$$

We test **Example 1** and **Example 2** with the well defect (see § 5.3.1) for different γ and the results are plot in Fig. 5. We observe that the error does

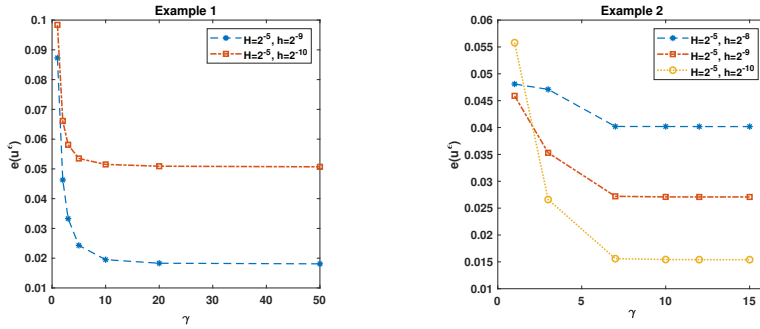


Fig. 5 $|u^\varepsilon - v_h|_{1, K_0} / |u^\varepsilon|_{1, K_0}$ versus γ in the well defect.

not change when γ is bigger than certain threshold value. It is worthwhile to mention that the threshold value is independent of H/h . In what follows, we set $\gamma = 50$ in **Example 1** and $\gamma = 20$ in **Example 2**.

5.2 The effect of the smoothness of the transition function

To study the effect of the smoothness of ρ numerically, we consider a square defect $K_0 := (0.5, 0.5) + (-L, L)^2$ with $L = 0.05$, and define $K_1 := (0.5, 0.5) + (-L - \delta, L + \delta)^2$ with $\delta = 0.05$. The transition function $\rho = \mu(x_1 - 0.5)\mu(x_2 - 0.5)$ with $\mu : \mathbb{R} \rightarrow [0, 1]$ defined by

$$\mu(t) := \begin{cases} \frac{1}{2} \cos(\pi(x+L)/\delta) + \frac{1}{2}, & \text{if } t \in [-L - \delta, -L), \\ 1 & \text{if } t \in (-L, L), \\ \frac{1}{2} \cos(\pi(-x+L)/\delta) + \frac{1}{2} & \text{if } t \in (L, L + \delta], \\ 0 & \text{otherwise.} \end{cases}$$

Such ρ is a C^1 function. We test **Example 1** with this C^1 transition function and the one constructed by a linear interpolant introduced in § 2, which is a continuous function. The results are reported in Table 1. We observe that there is no significant difference between these two choices and we shall use the C^0 transition function in the following tests.

Table 1 The error of Example 1 with different ρ .

$H = 2^{-5}, h = 2^{-8}$	$e(u_0)$	$e(u^\varepsilon)$
C^1 transition func.	4.21e-2	1.12e-1
C^0 transition func.	4.21e-2	1.13e-1
$H = 2^{-6}, h = 2^{-9}$	$e(u_0)$	$e(u^\varepsilon)$
C^1 transition func.	1.62e-2	4.62e-2
C^0 transition func.	1.62e-2	4.65e-2

5.3 The accuracy of the Nitsche hybrid method

We test three defects with different shapes: the well defect, the channel defect and the ellipse defect.

5.3.1 Well defect

The defect $K_0 = (0.5, 0.5) + (-L, L)^2$ with $L = 0.05$ is a square, and we define $K_1 := (0.5, 0.5) + (-L - \delta, L + \delta)^2$ with $\delta = 0.05$. We report the results for **Example 1** in Table 2 by fixing $H = 2^{-5}$ and decreasing h . We observe that locally refined mesh resolves the local events, and $e(u^\varepsilon) \simeq \mathcal{O}(h)$, which is consistent with the following explicit form of the error estimate (3.27):

$$|u^\varepsilon - v_h|_{1, K_0} \leq C \left(h + \frac{1}{\delta} (H^2 + L^2 |\ln L| + \varepsilon^{s-1}) \right).$$

The last term comes from the following error estimate proved in [39]:

$$\|u^\varepsilon - u_0\|_{0, D} \leq C \varepsilon^{s-1} \|u_0\|_{s, D} \quad \text{for } s < 3.$$

The error $e(u_0)$ remains unchanged when h is decreased, which shows that the resolution inside the defect has negligible effect on the accuracy for retrieving the macroscopic information.

To compute $e(u_0)$, we fix the mesh size in K_1 and refine the mesh in K_2 . The result is reported in Table 3. When $H \simeq L$, the first order rate of convergence is observed for the error $e(u_0)$, which is consistent with

$$|u_0 - v_h|_{1, K_2} \leq C(h + H + L |\ln L|^{1/2}).$$

Nevertheless, the error $e(u^\varepsilon)$ remains unchanged when H is decreased.

Table 2 Errors of Example 1 for the well defect with a fixed mesh $H = 2^{-5}$.

h	2^{-7}	2^{-8}	2^{-9}
$e(u^\varepsilon)$	2.32e-1	1.15e-1	4.60e-2
rate		1.01	1.32
$e(u_0)$	4.60e-2	4.67e-2	4.61e-2

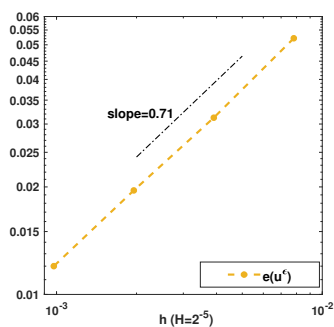
Table 3 The error of Example 1 in the well defect with a fixed $h = 2^{-10}$.

H	2^{-4}	2^{-5}	2^{-6}
$e(u_0)$	1.18e-1	4.61e-2	1.76e-2
rate		1.35	1.39
$e(u^\varepsilon)$	1.96e-2	1.69e-2	1.66e-2

We turn to **Example 2**. It follows from Table 4 that the rate of convergence for $e(u_0)$ is bigger than 1, while we do not know any quantity estimate on $e(\text{HMM})$ in this case. Fig. 6 indicates that the error $e(u^\varepsilon)$ converges at a rate around 0.71, which deteriorates a little bit than **Example 1**. This may be due to the roughness of a^ε inside K_0 .

Table 4 Errors of Example 2 for the well defect with a fixed $h = 2^{-10}$.

H	2^{-4}	2^{-5}	2^{-6}
$e(u_0)$	1.06e-1	4.90e-2	1.56e-2
rate		1.11	1.65
$e(u^\varepsilon)$	1.56e-2	1.02e-2	1.02e-2

**Fig. 6** Errors of **Example 2** for the well defect with a fixed $H = 2^{-5}$.

5.3.2 Channel defect

Let K_0 be a channel with a corner, with width $L = 0.05$ and K_1 is the set within a distance of $\delta = 0.025$ away from the channel; see Fig. 1_b.

We firstly test **Example 1** with different H . The result is shown in Table 5. We observe that the first order rate of convergence for the error $e(u_0)$, which is consistent with the theoretical result.

Table 5 The error of Example 1 in the channel defect with $h = 2^{-9}$.

H	2^{-4}	2^{-5}	2^{-6}
$e(u_0)$	9.74E-1	4.86E-2	2.40E-2
rate		1.19	1.00
$e(u^\varepsilon)$	7.74e-2	7.73e-2	7.73e-2

Next we fix $H = 2^{-5}$ and decrease h , and report the result in Table 6. We observe that the resolution of the defect has more pronounced influence on the error $e(u_0)$.

Table 6 The error of Example 1 in the channel defect with $H = 2^{-5}$.

h	2^{-7}	2^{-8}	2^{-9}
$e(u^\varepsilon)$	3.69e-1	1.99e-1	7.73e-2
rate		0.90	1.36
$e(u_0)$	5.10e-2	4.90e-2	4.86e-2

We report the results for **Example 2** in Table 7 and Fig. 7. The method still works with reasonable accuracy. However, from Fig. 7, we find that the error $e(u^\varepsilon)$ is worse than that in **Example 1**, which may be due to the poor regularity of the solution inside the defect.

Table 7 The error of Example 2 in the channel defect with $h = 2^{-9}$.

H	2^{-4}	2^{-5}	2^{-6}
$e(u_0)$	8.85E-2	4.42E-2	2.17E-2
rate		1.00	1.03
$e(u^\varepsilon)$	5.73e-2	4.80e-2	4.03e-2

5.3.3 Ellipse defects

We choose two slender ellipse defects as K_0 with major axis 0.5 and minor axis 0.02; See Fig. 8. K_1 is a rectangle of size 0.54×0.06 that contains the ellipse; see Fig. 1_c.

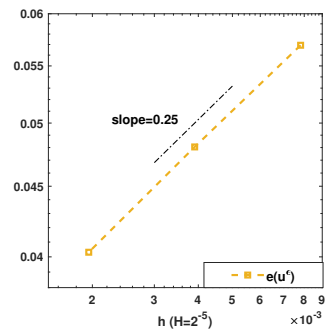


Fig. 7 The error of Example 2 in the channel defect with a fixed $H = 2^{-5}$.

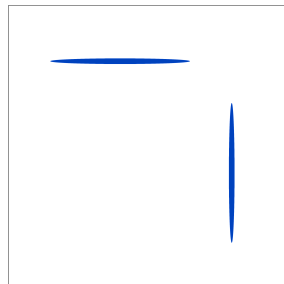


Fig. 8 Ellipse defects

We plot the relative errors in Fig. 9_a and Fig. 9_b with a fixed ratio H/h . The method works for both examples. If we refine the mesh in both subdomains simultaneously, then the energy error is around the first order.

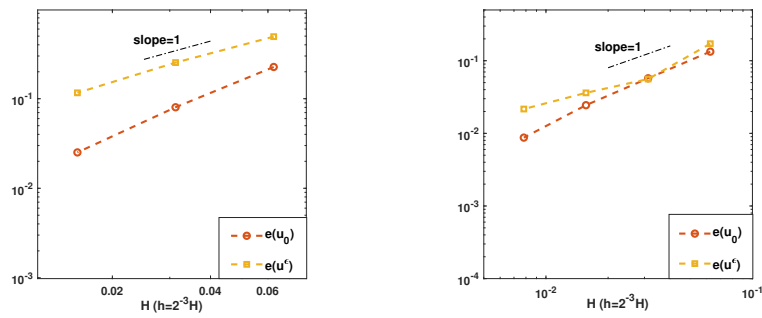


Fig. 9 The relative error for the ellipse defects.

5.4 Comparison with the original hybrid method

In the last test, we compare the present method with the hybrid method with a body fitted mesh [29]. Two kinds of mesh are plot in Fig. 10 for an illustration. Let $K_0 = (0.5, 0.5) + (-L, L)^2$ and $K_1 = (0.5, 0.5) + (-L - \delta, L + \delta)^2$ with $L = 0.05$ and $\delta = 0.01$.

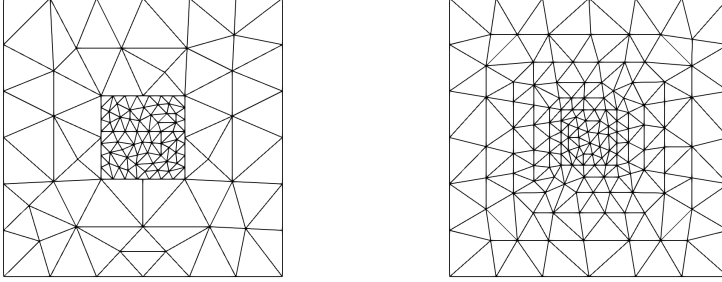


Fig. 10 Left: non-matching grid; Right: matching grid.

We choose a mesh with $N = 28585$ degrees of freedom for the Nitsche hybrid method, and $N = 60564$ degrees of freedom for the original hybrid method. The results for **Example 1** are summarized in Table 8. It seems that the accuracy of both methods are comparable, while the total degrees of freedom of the Nitsche hybrid method is less than one half of the original hybrid method.

Table 8 Comparison between the Nitsche hybrid method and the original hybrid method.

	Nitsche hybrid method	hybrid method
DOF	28585	60564
$e(u^\varepsilon)$	4.60e-2	4.64e-2
$e(u_0)$	9.30e-3	5.66e-3

6 Conclusion

We present a hybrid method that captures the macroscopical and microscopical information simultaneously in the framework of the Nitsche's variational formulation. A general approach for construction the transition function is proposed. This method admits non-matching grids and works for defects with irregular shape. Hence the method is more efficient and flexible than the original hybrid method [29]. We prove that the method converges for problems

with bounded and measurable coefficients. Rate of convergence has been derived for the periodic media and almost-periodic media. A possible extension of the present work is to deal with more realistic problem such as parabolic problems with time varying boundary conditions, which is allowed by the Nitsche's variational formulation [46]. We shall leave this for further pursuit.

Acknowledgements The authors would like to thank Professor Jianfeng Lu and Dr Yufang Huang for the discussion on the topic in the earlier stage of the present work.

A Proof of Lemma 4

We define a extension operator: for any $v \in X_{h,H}$

$$E_h v(p) := \begin{cases} v(p) & \text{if } p \in (\mathcal{N}_H \cup \mathcal{N}_h) \setminus \mathcal{N}_\Gamma, \\ \omega_1 v_1(p) + \omega_2 v_2(p) & \text{if } p \in \mathcal{N}_\Gamma^0, \end{cases}$$

where $v = (v_1, v_2) \in X_h \times X_H$. For each $p \in \mathcal{N}_\Gamma \setminus \mathcal{N}_\Gamma^0$, there exists $\tau \in \mathcal{T}_H$ such that p sits on the boundary of τ by **Assumption B**. Noting that $E_h v$ is well-defined over τ , we may define $E_h v(p) := (E_h v)_\tau(p)$. The definition of E_h and the uniqueness of the Lagrange interpolation over interface mesh \mathcal{E}_\cap ensure that $E_h v$ is continuous across Γ . Therefore, $E_h v \in X_{h,H} \cap H^1(D)$ and $E_h v \equiv v$ in $D \setminus \mathcal{T}_{h,H}^\Gamma$.

Over each $\tau \in \mathcal{T}_H^\Gamma$, we write

$$\begin{aligned} v - E_h v &= \sum_{p \in \mathcal{N}(\tau) \cap \mathcal{N}_\Gamma^0} (v_2 - E_h v)(p) \phi_{\tau,p} \\ &= \sum_{p \in \mathcal{N}(\tau) \cap \mathcal{N}_\Gamma^0} \omega_1 (v_2 - v_1)(p) \phi_{\tau,p}, \end{aligned}$$

where $\phi_{\tau,p}$ is the basis function on τ associated with node p . A scaling argument shows that, there exist c_n and C_n independent of h_τ such that

$$c_n h_\tau^{n-2m} \leq |\phi_{\tau,p}|_{m,\tau}^2 \leq C_n h_\tau^{n-2m}. \quad (\text{A.1})$$

Combining the elements in \mathcal{T}_H^Γ and using the standard inverse inequality [13], we obtain

$$\begin{aligned} \sum_{\tau \in \mathcal{T}_H^\Gamma} h_\tau^{2m-2} |v - E_h v|_{m,\tau}^2 &\leq C \sum_{e \in \mathcal{E}_\cap} \frac{h_e^2}{(h_e + H_e)^2} H_e^{n-2} \|\llbracket v \rrbracket\|_{0,\infty,e}^2 \\ &\leq C \sum_{e \in \mathcal{E}_\cap} \frac{h_e^2}{(h_e + H_e)^2} H_e^{n-2} h_e^{-d+1} \|\llbracket v \rrbracket\|_{0,e}^2 \\ &\leq C \sum_{e \in \mathcal{E}_\cap} h_e^{-1} \|\llbracket v \rrbracket\|_{0,e}^2, \end{aligned} \quad (\text{A.2})$$

where we have used

$$h_e^{-n+3} H_e^{n-2} / (h_e + H_e)^2 \leq (h_e / H_e)^{-n+4} h_e^{-1} \leq h_e^{-1} \quad n = 2, 3.$$

Next, over each $\tau \in \mathcal{T}_h^\Gamma$, we write

$$\begin{aligned} v - E_h v &= \sum_{p \in \mathcal{N}(\tau) \cap \Gamma} (v_1 - E_h v)(p) \phi_{\tau,p} \\ &= \sum_{p \in \mathcal{N}(\tau) \cap \Gamma} (v_1 - v_2)(p) \phi_{\tau,p} + \sum_{p \in \mathcal{N}(\tau) \cap \Gamma} (v_2 - E_h v)(p) \phi_{\tau,p} \end{aligned}$$

By **Assumption B**, there exists $\tilde{\tau} \in \mathcal{T}_H$ such that p sits on the boundary of $\tilde{\tau}$, and

$$\begin{aligned} E_h v(p) - v_2(p) &= \sum_{\tilde{p} \in \mathcal{N}(\tilde{\tau}) \cap \mathcal{N}_\tau^0} (E_h v - v_2)(\tilde{p}) \phi_{\tilde{\tau}, \tilde{p}}(p) \\ &= \sum_{\tilde{p} \in \mathcal{N}(\tilde{\tau}) \cap \mathcal{N}_\tau^0} \omega_1(v_1 - v_2)(\tilde{p}) \phi_{\tilde{\tau}, \tilde{p}}(p). \end{aligned}$$

A scaling argument shows that, there exist \tilde{c}_n and \tilde{C}_n independent of h_τ such that

$$\tilde{c}_n h_\tau^{n-2m} \leq |\phi_{\tau,p}|_{m,\tau}^2 \leq \tilde{C}_n h_\tau^{n-2m}.$$

By **Assumption A**, the number of the hanging nodes on each $e \in \mathcal{E}_H$ may be bounded by $c(H_e/h_\Gamma)^{n-1}$. Thus, proceeding along the same line that leads to (A.2) and using the above estimate and the inverse inequality, we obtain

$$\begin{aligned} \sum_{\tau \in \mathcal{T}_h^\Gamma} h_\tau^{2m-2} |v - E_h v|_{m,\tau}^2 &\leq C \sum_{e \in \mathcal{E}_\Omega} h_e^{n-2} h_e^{-n+1} \|\llbracket v \rrbracket\|_{0,e}^2 \\ &\quad + C \sum_{e \in \mathcal{E}_\Omega} (H_e/h_\Gamma)^{n-1} h_e^{n-2} h_e^{-n+1} \omega_1^2 \|\llbracket v \rrbracket\|_{0,e}^2 \\ &\leq C \sum_{e \in \mathcal{E}_\Omega} h_e^{-1} \|\llbracket v \rrbracket\|_{0,e}^2. \end{aligned} \quad (\text{A.3})$$

where we have used

$$(H_e/h_\Gamma)^{n-1} h_e^{-1} \omega_1^2 \leq (H_e/h_e)^{n-3} h_e^{-1} \leq h_e^{-1} \quad n = 2, 3.$$

Combining (A.2) and (A.3), we obtain (4.1).

B Proof of (3.23)

For any $\chi \in X_{h,H}^0$, using (3.6) and the enriching operator E_h defined in Lemma 4, we obtain

$$\begin{aligned} \langle g, u_0 - \tilde{u} \rangle &= \langle g, u_0 - E_h \tilde{u} \rangle + \langle g, E_h \tilde{u} - \tilde{u} \rangle \\ &= \langle \mathcal{A} \nabla (u_0 - E_h \tilde{u}), \nabla \psi_g \rangle + \langle g, E_h \tilde{u} - \tilde{u} \rangle \\ &= \langle f, \psi_g \rangle - \sum_{i=1}^2 \int_{K_i} \mathcal{A} \nabla \tilde{u} \nabla \psi_g \, dx + \sum_{i=1}^2 \int_{K_i} \mathcal{A} \nabla (\tilde{u} - E_h \tilde{u}) \nabla \psi_g \, dx \\ &\quad + \langle g, E_h \tilde{u} - \tilde{u} \rangle, \end{aligned}$$

which may be further expanded as

$$\begin{aligned} \langle g, u_0 - \tilde{u} \rangle &= \langle f, \chi \rangle - \sum_{i=1}^2 \int_{K_i} \mathcal{A} \nabla \tilde{u} \nabla \chi \, dx + \langle f, \psi_g - \chi \rangle - \sum_{i=1}^2 \int_{K_i} \mathcal{A} \nabla \tilde{u} \nabla (\psi_g - \chi) \, dx \\ &\quad + \langle g, E_h \tilde{u} - \tilde{u} \rangle - \sum_{i=1}^2 \int_{K_i} \mathcal{A} \nabla (E_h \tilde{u} - \tilde{u}) \nabla \chi \, dx \\ &\quad + \sum_{i=1}^2 \int_{K_i} \mathcal{A} \nabla (\tilde{u} - E_h \tilde{u}) \nabla (\psi_g - \chi) \, dx. \end{aligned}$$

Using (3.10), (2.2) and integration by parts, for any piecewise constant matrix $\bar{\mathcal{A}}$ over $\mathcal{T}_{h,H}$, we have

$$\begin{aligned} &\langle g, u_0 - \tilde{u} \rangle \\ &= \left\{ \langle f + \nabla \cdot \bar{\mathcal{A}} \nabla \tilde{u}, \psi_g - \chi \rangle - \sum_{e \in \mathcal{E}_\Omega} \int_e \llbracket \bar{\mathcal{A}} \nabla \tilde{u} \cdot \mathbf{n} \rrbracket \{ \{ \psi_g - \chi \} \}^\omega \, ds \right\} \\ &\quad + \left\{ \langle g + \nabla \cdot \bar{\mathcal{A}}^T \nabla \chi, E_h \tilde{u} - \tilde{u} \rangle - \sum_{e \in \mathcal{E}_\Omega} \int_e \llbracket \bar{\mathcal{A}}^T \nabla \chi \cdot \mathbf{n} \rrbracket \{ \{ E_h \tilde{u} - \tilde{u} \} \}^\omega \, ds \right\} \\ &\quad + \left\{ \sum_{i=1}^2 \int_{K_i} \mathcal{A} \nabla (\tilde{u} - E_h \tilde{u}) \nabla (\psi_g - \chi) \, dx + \sum_{e \in \mathcal{E}_\Omega} \int_e \frac{\gamma}{H_e + h_e} \llbracket \tilde{u} \rrbracket \llbracket \chi \rrbracket \, ds \right\} \\ &\quad + \left\{ - \sum_{i=1}^2 \int_{K_i} (\mathcal{A} - \bar{\mathcal{A}}) \nabla \tilde{u} \nabla (\psi_g - \chi) \, dx - \sum_{e \in \mathcal{E}_\Omega} \int_e \{ \{ (\mathcal{A} - \bar{\mathcal{A}}) \nabla \tilde{u} \} \}_\omega \llbracket \chi \rrbracket \, ds \right\} \\ &\quad + \left\{ - \sum_{i=1}^2 \int_{K_i} (\mathcal{A} - \bar{\mathcal{A}}) \nabla (E_h \tilde{u} - \tilde{u}) \nabla \chi \, dx - \sum_{e \in \mathcal{E}_\Omega} \int_e \{ \{ (\mathcal{A} - \bar{\mathcal{A}}) \nabla \chi \} \}_\omega \llbracket \tilde{u} \rrbracket \, ds \right\} \\ &= I_1 + \cdots + I_5. \end{aligned}$$

Using Lemma 6, we obtain

$$|I_1| \leq C \left(\|u_0 - \tilde{u}\| + \text{Osc}(f) + \max_{\tau \in \mathcal{T}_{h,H}} \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau} \|f\|_{-1,D} \right) \times \sum_{\tau \in \mathcal{T}_{h,H}} (h_\tau^{-1} \|\psi_g - \chi\|_{0,\tau} + |\psi_g - \chi|_{1,\tau}). \quad (\text{B.1})$$

Using Lemma 6 and the enriching estimates (4.2), (4.4), we obtain

$$|I_2| \leq C \varrho \left(\|\psi_g - \chi\| + \text{osc}(g) + \max_{\tau \in \mathcal{T}_{h,H}} \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau} \|g\|_{-1,D} \right) \|u_0 - \tilde{u}\|,$$

where we have used the fact $\llbracket u_0 \rrbracket \equiv 0$ on Γ .

Using (4.2) and the fact $\llbracket \psi_g \rrbracket \equiv \llbracket u_0 \rrbracket \equiv 0$ on Γ , we have

$$|I_3| \leq C \varrho \|\tilde{u} - u_0\| \|\psi_g - \chi\|.$$

Finally,

$$|I_4| \leq C \max_{\tau \in \mathcal{T}_{h,H}} \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau} \left(\sum_{i=1}^2 |\tilde{u}|_{1,K_i}^2 \right)^{1/2} \|\psi_g - \chi\| \\ \leq C \max_{\tau \in \mathcal{T}_{h,H}} \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau} (\|u_0 - \tilde{u}\| + \|f\|_{-1,D}) \|\psi_g - \chi\|,$$

and

$$|I_5| \leq C \varrho \max_{\tau \in \mathcal{T}_{h,H}} \|\mathcal{A} - \bar{\mathcal{A}}\|_{0,\infty,\tau} (\|\psi_g - \chi\| + \|g\|_{-1,D}) \|u_0 - \tilde{u}\|.$$

Combining all the above estimates, we obtain (3.23).

References

1. Abdulle, A., E, W., Engquist, B., Vanden-Eijnden, E.: The heterogeneous multiscale method. *Acta Numer.* **21**, 1–87 (2012)
2. Abdulle, A., Jecker, O.: An optimization-based, heterogeneous to homogeneous coupling method. *Commun. Math. Sci.* **13**, 1639–1648 (2015)
3. Adams, R.A., Fournier, J.J.F.: *Sobolev Spaces, Pure and Applied Mathematics (Amsterdam)*, vol. 140, second edn. Elsevier/Academic Press, Amsterdam (2003)
4. Apoung Kamga, J.B., Pironneau, O.: Numerical zoom for multiscale problems with an application to nuclear waste disposal. *J. Comput. Phys.* **224**, 403–413 (2007)
5. Armstrong, S., Kuusi, T., Mourrat, J.C.: *Quantitative Stochastic Homogenization and Large Scale Regularity*. Springer Nature Switzerland AG (2019)
6. Arnold, D.N., Brezzi, F., Cockburn, B., Marini, L.D.: Unified analysis of discontinuous Galerkin methods for elliptic problems. *SIAM J. Numer. Anal.* **39**, 1749–1779 (2001/02)

7. Babuška, I., Lipton, R.: L^2 -global to local projection: an approach to multiscale analysis. *Math. Models Methods Appl. Sci.* **21**, 2211–2226 (2011)
8. Babuška, I., Motamed, M., Tempone, R.: A stochastic multiscale method for the elastodynamic wave equation arising from fiber composites. *Comput. Methods Appl. Mech. Engrg.* **276**, 190–211 (2014)
9. Becker, R., Hansbo, P., Stenberg, R.: A finite element method for domain decomposition with non-matching grids. *M2AN Math. Model. Numer. Anal.* **37**, 209–225 (2003)
10. Brenner, S.C.: Two-level additive Schwarz preconditioners for nonconforming finite element methods. *Math. Comp.* **65**(215), 897–921 (1996)
11. Brenner, S.C., Sung, L.: Virtual enriching operators. *Calcolo* **56**(4), 44 (2019)
12. Burman, E., Elfverson, D., Hansbo, P., Larson, M.G., Larsson, K.: Hybridized CutFEM for elliptic interface problems. *SIAM J. Sci. Comput.* **41**(5), A3354–A3380 (2019)
13. Ciarlet, P.G.: *The Finite Element Method for Elliptic Problems*. North-Holland Publishing Co., Amsterdam-New York-Oxford (1978). *Studies in Mathematics and its Applications*, Vol. 4
14. Demlow, A., Guzmán, J., Schatz, A.H.: Local energy estimates for the finite element method on sharply varying grids. *Math. Comp.* **80**, 1–9 (2011)
15. Deng, W., Wu, H.: A combined finite element and multiscale finite element method for the multiscale elliptic problems. *Multiscale Model. Simul.* **12**, 1424–1457 (2014)
16. E, W.: *Principles of Multiscale Modeling*. Cambridge University Press, Cambridge (2011)
17. E, W., Engquist, B.: The heterogeneous multiscale methods. *Commun. Math. Sci.* **1**, 87–132 (2003)
18. E, W., Engquist, B., Li, X., Ren, W., Vanden-Eijnden, E.: Heterogeneous multiscale methods: a review. *Commun. Comput. Phys.* **2**, 367–450 (2007)
19. E, W., Ming, P., Zhang, P.: Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *J. Amer. Math. Soc.* **18**, 121–156 (2005)
20. Gloria, A.: Reduction of the resonance error-part 1: approximation of homogenized coefficients. *Math. Models Methods Appl. Sci.* **21**, 1601–1630 (2011)
21. Gloria, A., Habibi, M.: Reduction in the resonance error in numerical homogenization ii: correctors and extrapolation. *Found Comput. Math.* **16**, 217–296 (2016)
22. Glowinski, R., He, J., Rappaz, J., Wagner, J.: Approximation of multiscale elliptic problems using patches of finite elements. *C. R. Math. Acad. Sci. Paris* **337**, 679–684 (2003)
23. Gudi, T.: A new error analysis for discontinuous finite element methods for linear elliptic problems. *Math. Comp.* **79**, 2169–2189 (2010)
24. Gudi, T.: Some nonstandard error analysis of discontinuous Galerkin methods for elliptic problems. *Calcolo* **47**, 239–261 (2010)

25. Gustafsson, T., Stenberg, R., Videman, J.: Error analysis of Nitsche's mortar method. *Numer. Math.* **142**, 973–994 (2019)
26. Hecht, F.: New development in FreeFem++. *J. Numer. Math.* **20**, 251–265 (2012)
27. Heinrich, B., Pietsch, K.: Nitsche type mortaring for some elliptic problem with corner singularities. *Computing* **68**, 217–238 (2002)
28. Hou, T.Y., Wu, X.H.: A multiscale finite element method for elliptic problems in composite materials and porous media. *J. Comput. Phys.* **134**, 169–189 (1997)
29. Huang, Y.F., Lu, J.F., Ming, P.B.: A concurrent global-local numerical method for multiscale PDEs. *J. Sci. Comput.* **76**, 1188–1215 (2018)
30. Huang, Y.F., Ming, P.B., Song, S.Q.: An efficient online-offline method for elliptic homogenization problems. *CSIAM Trans. Appl. Math.* **1**, 556–592 (2020)
31. Karakashian, O.A., Pascal, F.: A posteriori error estimates for a discontinuous Galerkin approximation of second-order elliptic problems. *SIAM J. Numer. Anal.* **41**, 2374–2399 (2003)
32. Kenig, C., Lin, F.H., Shen, Z.W.: Convergence rates in L^2 for elliptic homogenization problems. *Arch. Rational Mech. Anal.* **203**, 1009–1036 (2012)
33. Kozono, H., Wadade, H.: Remarks on Gagliardo-Nirenberg type inequality with critical Sobolev space and BMO. *Math. Z.* pp. 935–950 (2008)
34. Li, H., Ming, P., Wang, H.: H^2 -Korn's inequality and the nonconforming elements for the strain gradient elastic model. *J. Sci. Comput.* **88**(3), 78 (2021)
35. Li, R., Ming, P.B., Tang, F.Y.: An efficient high order heterogeneous multiscale method for elliptic problems. *Multiscale Model. Simul.* **10**, 259–283 (2012)
36. Lüthen, N., Juntunen, M., Stenberg, R.: An improved a priori error analysis of Nitsche's method for Robin boundary conditions. *Numer. Math.* **138**, 1011–1026 (2018)
37. Meyers, N.G.: An L^p -estimate for the gradient of solutions of second order elliptic divergence equations. *Ann. Scuola Norm. Sup. Pisa Cl. Sci. (3)* **17**, 189–206 (1963)
38. Ming, P.B., Xu, X.M.: A multiscale finite element method for oscillating Neumann problem on rough domain. *Multiscale Model. Simu.* **14**, 1276–1300 (2016)
39. Moskow, S., Vogelius, M.: First order corrections to the homogenized eigenvalues of a periodic composite medium. a convergence proof. *Proc. Roy. Soc. Edinburgh* **127**, 1263–1295 (1997)
40. Nitsche, J.A.: Über ein Variationsprinzip zur Lösung von Dirichlet-Problemen bei Verwendung von Teilräumen, die keinen Randbedingungen unterworfen sind. *Abh. Math. Sem. Univ. Hamburg* **36**, 9–15 (1971)
41. Oden, J.T., Vemaganti, K.S.: Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. I. Error estimates and adaptive algorithms. *J. Comput. Phys.* **164**, 22–47 (2000)

42. Scott, L.R., Zhang, S.Y.: Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.* **54**, 483–493 (1990)
43. Shen, Z.W.: Convergence rates and Hölder estimates in almost-periodic homogenization of elliptic systems. *Analysis and PDE* **8**, 1565–1601 (2015)
44. Song, F., Deng, W., Wu, H.: A combined finite element and oversampling multiscale Petrov-Galerkin method for the multiscale elliptic problems with singularities. *J. Comput. Phys.* **305**, 722–743 (2016)
45. Stenberg, R.: Mortaring by a method of J. A. Nitsche. In: *Computational Mechanics (Buenos Aires, 1998)*, pp. CD-ROM file. Centro Internac. Métodos Numér. Ing., Barcelona (1998)
46. Tagliabue, A., Dedé, L., Quarteroni, A.: Nitsche’s method for parabolic partial differential equations with mixed time varying boundary conditions. *ESAIM: M2AN* **50**, 541–563 (2016)
47. Tartar, L.: The General Theory of Homogenization: A Personalized Introduction, *Lecture Notes of the Unione Matematica Italiana*, vol. 7. Springer-Verlag, Berlin; UMI, Bologna (2009)
48. Vemaganti, K.S., Oden, J.T.: Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. II. A computational environment for adaptive modeling of heterogeneous elastic solids. *Comput. Methods Appl. Mech. Engrg.* **190**, 6089–6124 (2001)
49. Verfürth, R.: A posteriori error estimation and adaptive mesh-refinement techniques. In: *Proceedings of the Fifth International Congress on Computational and Applied Mathematics (Leuven, 1992)*, vol. 50, pp. 67–83 (1994)
50. Verfürth, R.: *A Posteriori Error Estimation Techniques for Finite Element Methods*. Numerical Mathematics and Scientific Computation. Oxford University Press, Oxford (2013)
51. Warburton, T., Hesthaven, J.S.: On the constants in hp -finite element trace inverse inequalities. *Comput. Methods Appl. Mech. Engrg.* **192**, 2765–2773 (2003)
52. Yue, X.Y., E, W.: The local microscale problem in the multiscale modeling of strongly heterogeneous media: effects of boundary conditions and cell size. *J. Comput. Phys.* **222**, 556–572 (2007)