AN ADAPTIVE DISCONTINUOUS GALERKIN MULTISCALE METHOD FOR ELLIPTIC PROBLEMS

DANIEL ELFVERSON†, EMMANUIL H. GEORGOULIS‡, AND AXEL MÅLQVIST†

Abstract. An adaptive discontinuous Galerkin multiscale method driven by an energy norm a posteriori error bound is proposed. The method is based on splitting the problem into a coarse and fine scale. Localized fine scale constituent problems are solved on patches of the domain and are used to obtain a modified coarse scale equation. The coarse scale equation has considerably less degrees of freedom than the original problem. The a posteriori error bound is used within an adaptive algorithm to tune the critical parameters, i.e., the refinement level and the size of the different patches on which the fine scale constituent problems are solved. The fine scale computations are completely parallelizable, since no communication between different processors is required for solving the constituent fine scale problems. The convergence of the method, the performance of the adaptive strategy, and the computational effort involved are investigated through a series of numerical experiments.

Key words. multiscale, discontinuous Galerkin, a posteriori error bound

AMS subject classifications. 65N30, 65N15

DOI. 10.1137/120863162

1. Introduction. Problems involving features on several different scales, usually termed multiscale problems, can be found in many branches of the engineering sciences. Examples include the modeling of flow in a porous medium and of composite materials. Multiscale problems involving partial differential equations are often impossible to simulate with an acceptable accuracy using standard (single mesh) numerical methods. A different approach, usually coming under the general term of multiscale methods, consists of considering coarse and fine scale contributions to the solution, with the fine scale contributions approximated on localized patches. The fine scale contributions are then used to upscale the problem in order to obtain an approximation to the global multiscale solution.

1.1. Previous work. Numerous multiscale methods have been developed during the last three decades; see, e.g., [8, 7] for early works, or [16, 29, 15] and the references therein for exposition and recent developments. An important development is the Multiscale Finite Element Method (MsFEM) by Hou and Wu [21], which was further developed in [12], with the introduction of oversampling to reduce resonance effects. Another approach is the so-called Variational Multiscale method (VMS) of Hughes and co-workers [22, 23]. The idea in VMS is to decompose the solution space into coarse and fine scale contributions. A modified coarse scale problem is then solved (using a finite element approach), so that the fine scale contribution is taken into account. To maintain the conformity of the resulting modified finite element space, homogeneous Dirichlet boundary conditions are imposed on each fine-problem

---

*Received by the editors January 23, 2012; accepted for publication (in revised form) May 14, 2013; published electronically August 1, 2013.

†Department of Information Technology, Uppsala University, SE-751 05, Uppsala, Sweden (daniel.elfverson@it.uu.se, axel.malqvist@it.uu.se). The first author was supported by The Göran Gustafsson Foundation and The Swedish Research Council. The third author was supported by The Göran Gustafsson Foundation.

‡Department of Mathematics, University of Leicester, Leicester LE1 7RH, UK (emmanuil.georgoulis@le.ac.uk).
patch boundary. The Adaptive variational multiscale method (AVMS) using the VMS framework, introduced by Larson and Målvist [27], makes use of multiscale-type a posteriori error bound to adapt the coarse and fine scale mesh sizes as well as the fine-problem patch-sizes automatically. A priori error analysis can be found in [30].

An interesting alternative to conforming finite element methods is the class of discontinuous Galerkin (DG) methods, whereby the approximation spaces are elementwise discontinuous; the continuity of the underlying exact solutions is imposed weakly. DG methods appeared in the 1970s and in the early 1980s [32, 28, 9, 5, 24] and have recently received renewed interest; we refer to the volumes [13, 14, 20, 33] and the references therein for a literature review. DG methods admit good conservation properties of the state variable and, due to the lack of interelement continuity requirements, are ideally suited for application to complex and/or irregular meshes. Also, there has been work to better cope with the case of high contrast diffusion; see e.g., [19] where a DG method based on weighted average is proposed and analyzed. Discontinuous Galerkin methods for solving multiscale problems have been discussed using the framework of the MsFEM [1] and of the Heterogeneous Multiscale Method (HMM) [2]; see also [37, 36, 35, 34]. An a priori error analysis for the class of discontinuous Galerkin multiscale method studied in this paper can be found in [17].

1.2. New contributions. In this work, we propose an Adaptive Discontinuous Galerkin MultiScale method (ADG-MS) using the framework of VMS. The underlying DG method is based on weighted averages across the element interfaces. The adaptivity is driven by energy norm a posteriori error bounds. The multiscale method is based on solving localized problems on patches, which are then upscaled to solve a coarse scale equation. The lack of any interelement continuity requirements of the approximate solution allows for very general meshes, which is very common in multiscale applications; i.e., meshes that contain several types of elements and/or hanging nodes. The split between the coarse and fine scale is realized using the elementwise $L^2$-projection onto the coarse mesh. This is more natural in a multiscale setting than, e.g., using the nodal interpolant as in [27]. It is also much easier and efficient to construct an $L^2$ orthogonal split using DG as opposed to conforming multiscale methods. The ADG-MS inherits a local conservation property from DG on the coarse scale, which is crucial in many applications such as porous media flow. The fine scale problems can be solved independently with localized right-hand sides, and it is known that the solutions decay exponentially [17], which allows for small patches. In this case the ADG-MS converges to the reference solution, thereby taking full advantage of cancellation between patches; this is not the case for the original AVMS [27] since hanging nodes are not allowed. In the a posteriori error bound, the error is bounded in terms of the size of the different fine-scale patches and on both the fine-scale and the coarse-scale mesh sizes. An adaptive algorithm to tune all these parameters automatically is proposed. The numerical experiments show good performance of the algorithm for a number of benchmark problems.

1.3. Outline. The rest of this work is structured as follows. Section 2 is devoted to setting up the model problem, the basic DG discretization, and some notation. A general framework for multiscale problems along with the discontinuous Galerkin multiscale method is derived in section 3, and the a posteriori error bound is derived in section 4. The implementation of the method and the adaptive algorithm are discussed in section 5. In section 6, a number of numerical experiments are presented, and finally some conclusions are drawn in section 7.
2. Preliminaries. In this section we define some notations and the underlying DG method is presented.

2.1. Notation. Let \( \omega \subseteq \mathbb{R}^d, d = 2, 3 \), be an open polygonal domain. Denote the \( L^2(\omega) \)-inner product by \( (\cdot, \cdot)_{L^2(\omega)} \), and the corresponding norm by \( \| \cdot \|_{L^2(\omega)} \). Also, let \( H^1(\omega) \) be the Sobolev space with norm \( \| \cdot \|_{H^1(\omega)} := (\| \cdot \|_{L^2(\omega)}^2 + \| \nabla \cdot \|_{L^2(\omega)}^2)^{1/2} \), and let \( H^s(\omega) \) be the standard Hilbertian Sobolev space of index \( s \in \mathbb{R} \). We shall also make use of the space \( L^\infty(\omega) \) consisting of almost everywhere bounded functions, with norm \( \| \cdot \|_{L^\infty(\omega)} := \text{ess sup}_{\omega} |\cdot| \); see, e.g., [3] for details. Finally, the \( d \)-dimensional Lebesgue measure will be denoted by \( \mu_d(\cdot) \).

2.2. The model problem. Let \( \Omega \subset \mathbb{R}^d \) be an open polygonal domain with Lipschitz boundary \( \partial \Omega \), \( d = 2, 3 \), and consider the elliptic boundary value problem find \( u \in \{ v \in H^1(\Omega) : v|_{\partial \Omega} = 0 \} \) fulfilling

\[
\begin{align*}
-\nabla \cdot A \nabla u &= f, & u &\in \Omega, \\
u &= 0, & u &\in \partial \Omega,
\end{align*}
\]

with \( f \in L^2(\Omega) \) and \( A \in L^\infty(\Omega, \mathbb{R}^{d, d}) \) such that \( A \) has uniform spectral bounds, bounded below by \( \alpha > 0 \in \mathbb{R} \) almost everywhere.

2.3. Discretization and subdivision. The domain \( \Omega \) is subdivided into a partition \( \mathcal{K} = \{ K \} \) of shape-regular and closed elements \( K \) with boundaries \( \partial K \); i.e., \( \Omega = \bigcup_{K \in \mathcal{K}} K \). On the partition \( \mathcal{K} \), let \( h : \bigcup_{K \in \mathcal{K}} K \to \mathbb{R} \) be a mesh-function defined elementwise by \( h|_K := \text{diam}(K) \), \( K \in \mathcal{K} \). The partition is allowed to be irregular (i.e., hanging nodes are allowed) and it is locally quasi uniform in the sense that the ratio of the mesh function \( h \) for neighboring elements is uniformly bounded from above and below. Let \( \Gamma^B \) be the set of all boundary edges, and let \( \Gamma^I \) be the set of all interior edges (or faces when \( d = 3 \)) such that \( \Gamma = \Gamma^B \cap \Gamma^I \) is the set of all edges in the partition \( \mathcal{K} \). Associated with the diffusion tensor, we consider the elementwise constant functions \( A^0, A_0 : \bigcup_{K \in \mathcal{K}} K \to \mathbb{R} \) defined by the biggest and smallest eigenvalue of \( A \), respectively, on each element \( K \). For \( K_i, K_j \in \mathcal{K} \), with \( \mu_{d-1}(\partial K_i \cap \partial K_j) > 0 \), let \( K_i, K_j \) be denoted by \( K^+ \) and \( K^- \), where \( K^+ \) is the element with the higher index.

On interior element interfaces \( e \in \Gamma^I \) we shall make use of the shorthand notation \( v^+ := v|_{K^+} \), \( v^- := v|_{K^-} \); on boundary edges we set \( v^+ := v|_K \). We also define the weighted mean value by

\[
\{ v \}_e := w_{K^+(e)} v^+ + w_{K^-(e)} v^-,
\]

where

\[
\begin{align*}
w_{K^+(e)} &:= \frac{A^0|_{K^-}}{A^0|_{K^+} + A^0|_{K^-}}, \\
w_{K^-(e)} &:= \frac{A^0|_{K^+}}{A^0|_{K^+} + A^0|_{K^-}}
\end{align*}
\]

for each \( e \in \Gamma^I \) and

\[
w_{K^+(e)} = 1, \quad w_{K^+(e)} = 0
\]

for \( e \in \Gamma^B \). Further, the jump across element interfaces is defined by

\[
[v] := v^+ - v^- \quad \text{for } e \in \Gamma^I \quad \text{and} \quad [v] := v^+ \quad \text{for } e \in \Gamma^B,
\]

and the harmonic mean value \( \gamma_e \) by

\[
\gamma_e := \frac{2A^0|_{K^+} \cdot A^0|_{K^-}}{A^0|_{K^+} + A^0|_{K^-}}.
\]
Also, \( n \) will denote the outward unit normal to \( \partial K^+ \) when \( \mu_{d-1}(\partial K^+ \cap \partial K^-) > 0 \). When \( \mu_{d-1}(\partial K \cap \partial \Omega) > 0 \), \( n \) will be the outward unit normal to \( \partial \Omega \).

### 2.4. The discontinuous Galerkin method.

For a nonnegative integer \( r \), we denote by \( \mathcal{P}_r(K) \) the set of all polynomials on \( K \) of total degree at most \( r \) if \( K \) is the reference \( d \)-simplex, or of degree at most \( r \) in each variable if \( K \) is the reference \( d \)-hypercube.

Consider the space \( V := V_h + H^{1+\epsilon}(\Omega) \) with \( \epsilon > 0 \) but arbitrary small, and let the discontinuous finite element space be given by

\[
V_h := \{ v \in L^2(\Omega) : v \circ F_K|_K \in \mathcal{P}_r(\hat{K}), \hat{K} \in K \},
\]

where \( F_K : \hat{K} \to K \) is the respective elemental map for \( K \in K \), which is allowed to be nonaffine, provided its Jacobian remains nonsingular and uniformly bounded from above and below with respect to all meshes.

The discontinuous Galerkin method then reads as follows: find \( u_h \in V_h \) such that

\[
a(u_h, v) = \ell(v) \quad \forall v \in V_h,
\]

where the bilinear form \( a(\cdot, \cdot) : V \times V \to \mathbb{R} \) and the linear form \( \ell(\cdot) : V \to \mathbb{R} \) are given by

\[
a(v, z) := \sum_{K \in \mathcal{K}} (A \nabla v, \nabla z)_{L^2(K)} - \sum_{e \in \Gamma} \left( (n \cdot \{ A \Pi_1 \nabla v \}_w, [z])_{L^2(e)} + (n \cdot \{ A \Pi_1 \nabla z \}_w, [v])_{L^2(e)} - \frac{\sigma_e \gamma_e}{h_e} ([v], [z])_{L^2(e)} \right),
\]

\[
\ell(v) := (f, v)_{L^2(\Omega)},
\]

respectively. Here \( \Pi : (L^2(\Omega))^d \to (V_h)^d \) denotes the orthogonal \( L^2 \)-projection operator onto \((V_h)^d \), \( h_e := \text{diam}(e) \), and \( \sigma_e \in \mathbb{R} \) is a positive constant. The bilinear form \( (2.11) \) is coercive with respect to the natural energy norm,

\[
|||v||| = \left( \sum_{K \in \mathcal{K}} \| A^{1/2} \nabla v \|_{L^2(K)}^2 + \sum_{e \in \Gamma} \frac{\sigma_e \gamma_e}{h_e} \| [v] \|_{L^2(e)}^2 \right)^{1/2}
\]

if \( \sigma_e \) is chosen to be large enough. We refer, e.g., to [14, 6] and the references therein for details on the analysis of DG methods for elliptic problems. Discontinuous Galerkin methods with weighted averages were introduced in [10, 19].

**Remark 2.1.** For all \( v \in V_h \), we have \( \Pi \nabla v = \nabla v \); therefore, the bilinear form \( (2.10) \) with \( v, z \in V_h \) is reduced to the more familiar form

\[
a(v, z) := \sum_{K \in \mathcal{K}} (A \nabla v, \nabla z)_{L^2(K)} - \sum_{e \in \Gamma} \left( (n \cdot \{ A \nabla v \}_w, [z])_{L^2(e)} + (n \cdot \{ A \nabla z \}_w, [v])_{L^2(e)} - \frac{\sigma_e \gamma_e}{h_e} ([v], [z])_{L^2(e)} \right).
\]

### 3. The multiscale method.

In the VMS framework, the finite element solution space \( V_h \) is decoupled into coarse and fine scale contributions, viz., \( V_h = V_H \oplus V_f \), with \( V_H \subset V_h \). To this end, let \( \Pi_H : L^2(\Omega) \to V_H \) be the (orthogonal) \( L^2 \)-projection onto the coarse mesh. The split between the coarse and fine scales is then determined by \( V_H := \Pi_H V_h \) and \( V_f := (I - \Pi_H)V_h = \{ v \in V_h : \Pi_H v = 0 \} \), where \( I \) is the identity operator.
The multiscale map $\mathcal{T} : \mathcal{V}_H \to \mathcal{V}_f$ from the coarse to the fine scale is defined as

$$a(\mathcal{T} v_H, v_f) = -a(v_H, v_f) \quad \forall v_H \in \mathcal{V}_H \text{ and } \forall v_f \in \mathcal{V}_f.$$  

The next step is to decompose $u_h$ and $v$ in (2.9) into coarse and fine scale components. In particular, we have

$$u_h = u_H + \mathcal{T} u_H + u_f,$$

and $v = v_H + v_f$, with $u_H, v_H \in \mathcal{V}_H$ and $\mathcal{T} u_H, v_f \in \mathcal{V}_f$ for some $u_f \in \mathcal{V}_f$. Equation (2.9) is equivalent to the following problem: find $u_H \in \mathcal{V}_H$ and $v_f \in \mathcal{V}_f$ such that

$$a(u_H + \mathcal{T} u_H + u_f, v_H + v_f) = \ell(v_H + v_f) \quad \forall v_H \in \mathcal{V}_H \text{ and } \forall v_f \in \mathcal{V}_f.$$  

The fine scale component $u_f$ can be computed by letting $v_H = 0$ in (3.3) and using the multiscale map (3.1). We obtain the fine scale problem driven by the right-hand side data $f$: find $u_f \in \mathcal{V}_f$ such that

$$a(u_f, v_f) = \ell(v_f) \quad \forall v_f \in \mathcal{V}_f.$$  

The coarse scale solution is obtained by letting $v_f = 0$ in (3.3): find $u_H \in \mathcal{V}_H$ such that

$$a(u_H + \mathcal{T} u_H, v_H) = \ell(v_H) - a(u_f, v_H) \quad \forall v_H \in \mathcal{V}_H.$$  

In (3.5), $\mathcal{T} v_H$ and $u_f$ are unknown and obtained by solving (3.1) and (3.4). Note that the linear system (3.5) has $\dim(\mathcal{V}_H)$ unknowns.

### 3.1. Localization and discretization

The bilinear form is characterized by more local behavior in $\mathcal{V}_f$ than in $\mathcal{V}_h$ [30, 17]. This motivates us to solve the fine scale equations on (localized) overlapping patches, instead of the whole domain $\Omega$. The patches are chosen large enough to ensure sufficiently accurate computations of $\mathcal{T} v_H$ and $u_f$. The computations of the fine scale components of the solution can be done in parallel with localized right-hand sides. To define the coarse space $\mathcal{V}_H$, we begin by fixing a coarse mesh $\mathcal{K}_H$. Then, $\mathcal{V}_H$ is defined as

$$\mathcal{V}_H := \{ v \in L^2(\Omega) : v \circ F_K|_K \in \mathcal{P}_r(\hat{K}), \hat{K} \in \mathcal{K}_H \}.$$  

**Definition 3.1.** For all $K \in \mathcal{K}_H$, define element patches of size $L$ patch as

$$\begin{align*}
\omega^1_K &= \text{int}(K), \\
\omega^L_K &= \text{int}(\cup \{ K' \in \mathcal{K}_H \mid K' \cap \omega^1_K \}), \quad L = 2, 3, \ldots.
\end{align*}$$  

The patch $\omega^L_K$ will be referred to as a $L$-layer patch. This is illustrated in Figure 1.

On each $L$-layer patch, we let $\mathcal{K}(\omega^L_K)$ be a restriction of $\mathcal{K}$ to $\omega^L_K$, such that $\cup_{K \in \mathcal{K}(\omega^L_K)} = \omega^L_K$. Also let $\Gamma^I(\omega^L_K)$ and $\Gamma^B(\omega^L_K)$ be the interior, respectively, boundary edges on $\mathcal{K}(\omega^L_K)$. Moreover, we assume that $\mathcal{K}_H|_{\omega^L_K}$ and $\mathcal{K}(\omega^L_K)$ are nested; that is, every coarse element $K_H \in \mathcal{K}_H|_{\omega^L_K}$ coincides with a union of fine elements $K \in \mathcal{K}(\omega^L_K)$. Also, the fine test spaces $\mathcal{V}_f(\omega^L_K)$ are defined by

$$\mathcal{V}_f(\omega^L_K) := \{ v \in \mathcal{V}_f : v|_{\Omega \setminus \omega^L_K} = 0 \}.$$  

Finally, let the indicator function be $\chi_K = 1$ on element $K$ and 0 otherwise, and let $\mathcal{M}_K$ be the index set of all basis functions $\phi_j \in \mathcal{V}_H$ that have support on $K$; i.e., $\chi_K = \sum_{j \in \mathcal{M}_K} \phi_j$. 


3.2. The discontinuous Galerkin multiscale method. For each $K \in \mathcal{K}_H$ the following local problems need to be solved: find $\hat{T} \phi_j \in \mathcal{V}_f(\omega^1_K)$ $\forall j \in \mathcal{M}_K$ and $U_{f,K} \in \mathcal{V}_f(\omega^1_K)$ such that

\begin{align}
(3.9) \quad a(\hat{T} \phi_j, v_f) &= -a(\phi_j, v_f) \quad \forall v_f \in \mathcal{V}_f(\omega^1_K), \\
(3.10) \quad a(U_{f,K}, v_f) &= \ell(\chi_K v_f) \quad \forall v_f \in \mathcal{V}_f(\omega^1_K).
\end{align}

The modified coarse scale problem is formulated as follows: find $U_H \in \mathcal{V}_H$ such that

\begin{equation}
(3.11) \quad a(U_H + \hat{T} U_H, v_H) = \ell(v_H) - a(U_f, v_H), \quad \forall v_H \in \mathcal{V}_H,
\end{equation}

where $U_f := \sum_{K \in \mathcal{K}_H} U_{f,K}$. The approximate solution to the multiscale problem is given by

\begin{equation}
(3.12) \quad U = U_H + \hat{T} U_H + U_f.
\end{equation}

The above procedure will be referred to as the discontinuous Galerkin multiscale method.

We note that the approximation $U$ is not equal to $u_h$, in general, since the domains of the fine scale problems are truncated. However, as discussed above, it is expected that $U$ is a good approximation to $u_h$, due to the decaying nature of the fine scale solutions away from the respective patch. For the approximation $U$ to converge to the exact solution $u$ of (2.1) in the limit, the support of the local problems should be gradually extended to both the whole computational domain and the fine scale meshsize $h$ should converge to 0. The multiscale method proposed here differs from the one proposed in [17] in that a right-hand side correction is present. Using the formulation without the presence of a right-hand side correction, the multiscale solution converges to an $H$-perturbation of the exact solution $u$, uniformly with respect to the diffusion coefficient structure.

Remark 3.2. Note that for a nonuniform mesh $\mathcal{K}$ (and/or $\mathcal{K}_H$), the convergence results presented in [17] still hold if the corrected basis functions are computed on patches of a common reference mesh $\mathcal{K}$. On the other hand, if the adaptive algorithm
is used so that the overlap between different corrected basis functions are computed on different meshes (cf., e.g., [27]), less cancellation of the error will occur and convergence can no longer be guaranteed by the argument in [17].

3.3. Local conservation property. The DG methods are known to have good local conservation properties in that the normal fluxes are conservative. The ADG-MS inherits this property on the coarse scale. To see this, we introduce the normal fluxes on element $K_H \in K_H$ as

\begin{equation}
\sigma(U) := \{ n \cdot A \nabla\chi \}_{e} = \sigma_c e^{-1}[U], \quad e \in \partial K_H,
\end{equation}

where $U = U_H + T U_H + U_f$, $\chi_{K_H} = 1$ on element $K_H$ and $\chi_{K_H} = 0$ otherwise ([$\chi_{K_H}$] is either 1 or -1), and each interface $e$ is a face of a fine scale element $K \in K$, i.e., the number of edges can exceed the number of faces for each element $K_H$. By setting $w \in V_H$ to be $w = \chi_{K_H}$ in (2.10), (2.11), and by using the discrete normal fluxes defined in (3.13), we arrive to the discrete elementwise conservation law

\begin{equation}
(f, 1)_{L^2(K_H)} + (\sigma(U), 1)_{L^2(\partial K_H)} = 0
\end{equation}

for all $K_H \in K_H$.

4. A posteriori error bound in energy norm. Let the constant $0 \leq C < \infty$ be any generic constant neither depending on $H$, $h$, $L$, nor $A$; let $a \lesssim b$ abbreviate the inequality $a \leq Cb$. The following approximation results will be used frequently throughout this section. Let $\pi$ be the orthogonal $L^2$-projection operator onto elementwise constant functions. Then $\pi$ satisfies the following approximation properties: for an element $K$, we have

\begin{equation}
|| v - \pi v ||_{L^2(K)} \lesssim \frac{h_K}{\sqrt{A_0}} || A^{1/2} \nabla v ||_{L^2(K)} \quad \forall v \in H^1(K),
\end{equation}

\begin{equation}
|| v - \pi v ||_{L^2(\partial K)} \lesssim \frac{h_K}{\sqrt{A_0}} || A^{1/2} \nabla v ||_{L^2(K)} \quad \forall v \in H^1(K).
\end{equation}

**Lemma 4.1.** Let $T_h^c : V_h \to V_h \cap H^1(\Omega)$ be a averaging interpolation operator defined pointwise as

\begin{equation}
T_h^c(v_h)(\hat{x}) = \frac{1}{|K_{\hat{x}}|} \sum_{K \in K_{\hat{x}}} v_h(\hat{x})|K,
\end{equation}

where $K_{\hat{x}}$ is the set of elements in $K$ for which $\hat{x}$ belong, with the cardinal $|K_{\hat{x}}|$. Then,

\begin{equation}
||v_h - T_h^c v_h||_{L^2(K)} \lesssim || h_e ||_{L^2(\partial K)}
\end{equation}

\begin{equation}
||A^{1/2} \nabla (v_h - T_h^c v_h)||_{L^2(K)} \lesssim A^0 \left( \frac{1}{\sqrt{h_e}} ||v_h||_{L^2(\partial K)} \right)^2
\end{equation}

holds for all $v_h \in V_h$ and $K \in K$.

The proof, omitted here, follows closely that of [25]. Lemma 4.1 can also be extended to irregular meshes. There a hierarchical refinement of the mesh is performed to eliminate the hanging nodes; we refer to [26] for details. For irregular meshes the constant in the bounds of Lemma 4.1 also depends on the number of hanging nodes on each face.
Remark 4.2. The result in Lemma 4.1 can be sharpened if the diffusion tensor is isotropic and a locally quasi-monotone [31] distribution is assumed to hold. Then $A^0|_K$ can be replaced by the harmonic mean value $\gamma_e$ on face $e$; see [11].

First we derive a posteriori bound for the underlying (one scale) DG method.

Theorem 4.3. Let $u, u_h$ be given by (2.1)–(2.2) and (2.9), respectively. Also let $\mathcal{T}_h u_h \in \mathcal{V}_h \cap H^1(\Omega)$ be given by (4.3). Moreover, let $\mathcal{E} := \mathcal{E}_c + \mathcal{E}_d$, where $\mathcal{E}_c := u - \mathcal{T}_h u_h$ and $\mathcal{E}_d := \mathcal{T}_h u_h - u_h$. Then

$$
|||E||| \leq \left( \sum_{K \in \mathcal{K}} \varrho_K^2 \right)^{1/2} + \left( \sum_{K \in \mathcal{K}} \zeta_K^2 \right)^{1/2},
$$

where

$$
\varrho_K = \frac{h_K}{A_0} \left( ||(1 - \Pi)(f + \nabla \cdot A \nabla u_h)||_{L^2(K)} + \sqrt{\frac{h_K}{A_0}} \left( ||(1 - w_K) \nabla \cdot [A \nabla u_h]||_{L^2(\partial K \setminus \Gamma_d)} + \left\| \frac{\sigma_e}{h_e} [u_h] \right\|_{L^2(\partial K)} \right) \right),
$$

$$
\zeta_K^2 = ||A^{1/2} \nabla (u_h - \mathcal{T}_h u_h)||_{L^2(K)}^2 + \left\| \sqrt{\frac{\sigma_e}{h_e}} [u_h] \right\|_{L^2(\partial K)}^2.
$$

Remark 4.4. Using $\mathcal{T}_h u_h$ as the conforming part of $u_h$, we arrive to an a posteriori bound whereby $\mathcal{T}_h u_h$ can either be evaluated directly or bounded using Lemma 4.1. Another possible choice is a weighted averaging interpolation operator with the weights depending on the diffusion tensor [4].

Remark 4.5. Concerning the lower efficiency bounds, the term (4.7) is robust with respect to the diffusion tensor; see [18]. But to prove that (4.8) is robust with respect to the diffusion tensor, to the best of the authors’ knowledge, the diffusion tensor has to be isotropic and satisfy a locally quasi-monotone property [31, 11].

Proof. Note that

$$
|||E||| \leq |||\mathcal{E}_c||| + |||\mathcal{E}_d|||,
$$

where the first part can be bounded by

$$
|||\mathcal{E}_c|||^2 \lesssim a(\mathcal{E}_c, \mathcal{E}_c) = a(\mathcal{E}, \mathcal{E}_c) - a(\mathcal{E}_d, \mathcal{E}_c) \lesssim a(\mathcal{E}, \mathcal{E}_c) + |||\mathcal{E}_d|||||\mathcal{E}_c|||.
$$

Let $\pi_h$ be the $L^2$-orthogonal projection onto the elementwise constant functions and define $\eta := \mathcal{E}_c - \pi_h \mathcal{E}_c$. We then have

$$
a(\mathcal{E}, \mathcal{E}_c) = a(u, \mathcal{E}_c) - a(u_h, \mathcal{E}_c) = \ell(\mathcal{E}_c) - a(u_h, \mathcal{E}_c) = \ell(\eta) - a(\eta),
$$

which implies

$$
|||\mathcal{E}_c|||^2 = a(\mathcal{E}_c, \mathcal{E}_c) = (\ell(\eta) - a(u_h, \eta)) - a(\mathcal{E}_d, \mathcal{E}_c).
$$

Upon integration by parts and using the identity $vw = \{v\}_w [w] + \{w\}_w [v]$, where $\bar{w}$ is the skew-weighted average given by

$$
\{v\}_w := w_{K^-(e)} v^+ + w_{K^+(e)} v^-,
$$
the first term on the right-hand side of (4.12) yields
\[
\ell(\eta) - \alpha(u_h, \eta)
\]
(4.14) \[
\sum_{K \in \mathcal{K}} (f + \nabla \cdot A \nabla u_h, \eta)_{L^2(K)} + \sum_{e \in \Gamma} (-n \cdot [A \nabla u_h], \{\eta\}_{\partial K})_{L^2(\partial K)} + (n \cdot \{A \Pi \nabla \eta\}_{\partial K}, [u_h]_{\partial K})_{L^2(\partial K)}.
\]

The first term on the right-hand side of (4.14) can be bounded as follows:

\[
\sum_{K \in \mathcal{K}} (f + \nabla \cdot A \nabla u_h, \eta)_{L^2(K)} \lesssim \sum_{K \in \mathcal{K}} \frac{h_K}{A_0} \|(1 - \Pi)(f + \nabla \cdot A \nabla u_h)\|_{L^2(K)} \|A^{1/2} \nabla \mathcal{E}_e\|_{L^2(K)},
\]

using (4.1). The second term on the right-hand side of (4.14) gives

\[
\sum_{e \in \Gamma} (n \cdot [A \nabla u_h], \{\eta\}_{\partial K})_{L^2(\partial K)} \lesssim \sum_{K \in \mathcal{K}} \sqrt{\frac{h_K}{A_0}} \|1 - w_{K(e)}\|_{L^2(\partial K)} \|A^{1/2} \nabla \mathcal{E}_e\|_{L^2(K)}
\]

using (4.2). For the third term on the right-hand side of (4.14), noting that \(\nabla \eta = \nabla \mathcal{E}_e\), we deduce that

\[
\sum_{e \in \Gamma} (n \cdot \{A \Pi \nabla \mathcal{E}_e\}_{\partial K}, [\mathcal{E}_d])_{L^2(\partial K)} \lesssim \sum_{K \in \mathcal{K}} \frac{1}{\sqrt{h_K A_0}} \|\gamma_{\mathcal{E}_e}[\mathcal{E}_d]\|_{L^2(\partial K)} \|A^{1/2} \nabla \mathcal{E}_e\|_{L^2(K)}
\]

using an inverse estimate and the \(L^2\)-stability of \(\Pi\). For the last term on the right-hand side of (4.14), we have

\[
\sum_{e \in \Gamma} \frac{\sigma_{\mathcal{E}_e}}{h_e} ([u_h], [\eta])_{L^2(\partial K)} \lesssim \sum_{K \in \mathcal{K}} \frac{1}{\sqrt{h_K A_0}} \|\gamma_{\mathcal{E}_e}[u_h]\|_{L^2(\partial K)} \|A^{1/2} \nabla \mathcal{E}_e\|_{L^2(K)}
\]

The last term on the right-hand side of (4.12) is bounded using the continuity of the bilinear form. Combining all of the above bounds and using Lemma 4.1 to bound the nonconforming part, the result follows.

A posteriori error estimate for the ADG-MS is given below.

**Theorem 4.6.** Let \(u, U\) be defined in (2.1)--(2.2) and (3.12), respectively, and set \(\mathcal{T}_h^2 U \in H^1(\Omega)\). Set \(\mathcal{E} := \mathcal{E}_c + \mathcal{E}_d\), where \(\mathcal{E}_c := u - \mathcal{T}_h^2 U\) and \(\mathcal{E}_d := \mathcal{T}_h^2 U - U\). Define \(U_{K,H} := \sum_{j \in K_H} U_j(\phi_j + \tilde{\phi}_j) + U_f, K_H\), where \(U_j\) are the nodal values calculated by (3.11) for all \(K_H\). Then, \(\mathcal{E}\) satisfies the estimate

\[
\|\|\mathcal{E}\|\| \lesssim \left( \sum_{K \in \mathcal{K}} \frac{\theta^2_\mathcal{K}}{\bar{\theta}_\mathcal{K}} \right)^{1/2} + \left( \sum_{K \in \mathcal{K}} \zeta^2_\mathcal{K} \right)^{1/2} + \left( \sum_{K_H \in \mathcal{K}_H} \rho^2_{\mathcal{E}, K_H} \right)^{1/2},
\]

where

\[
\rho^2_{\mathcal{E}, K_H} := \sum_{e \in \Gamma^n(w_{K_H})} \left( \frac{H^2_{K_H}}{h_{K,H} A_0} \right) \left( \|n \cdot \{A \nabla U_i\}_w\|_{L^2(c)} + \frac{\sigma_{\mathcal{E}_c}}{h_e} \|[U_i]\|_{L^2(c)} \right)^2.
\]
measures the effect of the truncated patches, \( K^O, K_H^O \) are from outside of \( \omega_{K_H}^c \), and

\[
\varrho_K = \frac{h_K}{\sqrt{A_0}} \left\| (1 - \Pi)(f + \nabla \cdot A \nabla U) \right\|_{L^2(K)} + \sqrt{\frac{h_K}{A_0}} \left( \left\| (1 - w_{K(c)}) n \cdot [A \nabla U] \right\|_{L^2(\partial K)} + \left\| \frac{\sigma_e \gamma_c}{h_e} [U] \right\|_{L^2(\partial K)} \right),
\]

\[
\zeta_K^2 = \left\| \sqrt{A} \nabla (U - T_K U) \right\|_{L^2(K)}^2 + \left\| \frac{\sigma_e \gamma_c}{h_e} [U] \right\|_{L^2(\partial K)}^2.
\]

measuring the refinement level of the fine scale.

Remark 4.7. One possible adaptive strategy would be to refine the coarse mesh as much one can afford, using a standard a posteriori error bound (e.g., using Theorem 4.3), and then further improve the approximation using Theorem 4.6. Note that fine scale problems do not have to be solved everywhere.

Remark 4.8. For the estimator \( \rho_{K_H} \) to retain its optimality with respect to the mesh-sizes, one should assume that \( H_{K_H}^2 \lesssim h_K \). We note that this is not an unreasonable requirement, for, otherwise, each fine scale problem would be more expensive to solve than the coarse scale problem.

Proof. Using the same idea as in Theorem 4.3. We first note that

\[
\| \varphi_c \|^2 = a(\varphi_c, \varphi_c) = a(\varphi, \varphi_c) - a(\varphi_c, \varphi_c).
\]

Then, using (2.9) and the fine scale equations (3.9)–(3.10), we have

\[
a(\varphi, \varphi_c) = \ell(\varphi_c) - a(U, \varphi_c)
\]

\[
= \ell(\varphi_c - v_H) - a(U, \varphi_c - v_H)
\]

\[
= \ell(\varphi_c - v_H - v_f) - a(U, \varphi_c - v_H - v_f) + \ell(v_f) - a(U, v_f)
\]

for any \( v_H \in V_H \) and \( v_f \in V_f \). Note that,

\[
\ell(v_f) - a(U, v_f) = \sum_{K_H \in K_H} \ell(\chi_{K_H} v_f) - a(U_{K_H}, v_f)
\]

\[
= \sum_{K_H \in K_H} \sum_{K_H \in \Omega \cap \Theta(\omega_{K_H}^c)} \left( (n \cdot \{ A \nabla U_i \}_w, [\xi_{K_H}^L v_f])_{L^2(\iota)} + (n \cdot \{ A \nabla ^2 \xi_{K_H} v_f \}_w, [U_i])_{L^2(\iota)} - \frac{\sigma_e \gamma_c}{h_e} ([U_i], [\xi_{K_H}^L v_f])_{L^2(\iota)} \right),
\]

where \( \xi_{K_H}^L = 0 \) on \( \omega_{K_H}^c \) and \( \xi_{K_H}^L = 1 \) otherwise; that is, \( v_f = \xi_{K_H}^L v_f + (1 - \xi_{K_H}^L) v_f \), where \((1 - \xi_{K_H}^L) v_f \in V_f(\omega_{K_H}^c) \). Then, applying (4.25), we deduce

\[
a(\varphi, \varphi_c) = \left( \ell(\varphi_c - v_H - v_f) - a(U, \varphi_c - v_H - v_f) \right)
\]

\[
+ \sum_{K_H \in K_H} \sum_{K_H \in \Omega \cap \Theta(\omega_{K_H}^c)} \left( (n \cdot \{ A \nabla U_i \}_w, [\xi_{K_H}^L v_f])_{L^2(\iota)} + (n \cdot \{ A \nabla ^2 \xi_{K_H} v_f \}_w, [U_i])_{L^2(\iota)} - \frac{\sigma_e \gamma_c}{h_e} ([U_i], [\xi_{K_H}^L v_f])_{L^2(\iota)} \right)
\]

\[
=: I + II.
\]
Term I can be estimated as in the proof of Theorem 4.3, upon selecting \( v_H := \pi_H \phi_c \) and \( v_f = \pi_f (\phi_c - \pi_H \phi_c) = \pi_f \phi_c \), where \( \pi_H \) and \( \pi_f \) are the elementwise constant \( L^2 \)-orthogonal projections onto the coarse space \( V_H \) on the fine space \( V_f \), respectively. We note that, by construction, \( \pi_f \pi_H v = 0 \ \forall v \in V_h \).

Since \( v_f \) is chosen to be piecewise constant the second term in II is equal to zero. For each \( K \in \mathcal{K} \), and for each \( e \in \Gamma^B(\omega_{K_H}^e) \setminus \Gamma^B \), we have

\[
\begin{align*}
\left| \langle n \cdot \{ A \nabla U_i \}_w, [\xi^L_{K_H} v_f] \rangle_{L^2(e)} - \frac{\sigma e \gamma e}{h_e} (\langle [U_i], [\xi^L_{K_H} v_f] \rangle_{L^2(e)} \right| \\
&\lesssim \left( \| n \cdot \{ A \nabla U_i \}_w \|_{L^2(e)} + \frac{\sigma e \gamma e}{h_e} \| [U_i] \|_{L^2(e)} \right) \| \xi^L_{K_H} v_f \|_{L^2(e)}
\end{align*}
\]

using (4.28) and the Cauchy–Schwarz inequality, the proof is concluded.

Combining the above and summing over all patches, using the discrete version of the Cauchy–Schwarz inequality, the proof is concluded.

5. Implementation and adaptivity. The system of equations arising from the discretization of the modified coarse multiscale problem (3.11) is given by

\[
K U = b - d,
\]

where \( K_{i,j} = a(\phi_j + \bar{T} \phi_i, \phi_i) \), \( b_i = \ell(\phi_i) \), and \( d_i = a(U_f, \phi_i) \). To assemble the right- and left-hand sides of (5.1), \( \bar{T} \phi_i \) and \( U_f \) need to be computed for all \( i \in N \). This can be done in parallel since no communication is required between the different fine scale problems. For each fine scale problem it is also possible to assemble \( K_{i,j} = a(\phi_j + \bar{T} \phi_i, \phi_i) \), \( b_i = \ell(\phi_i) \), and \( d_i = \sum_{j \in N} a(U_f, \phi_i) \) for a fixed \( i \) and for all \( j \) such that \( \mu_d(\text{supp}(\phi_j) \cap \bar{\omega}_K) > 0 \). The constraints needed on the fine scale test spaces to solve \( \bar{T} \phi_i \) and \( U_f \) are \( V_f = \{ v \in V_h : \Pi_H v = 0 \} \), which are implemented using Lagrange multipliers. The spaces \( V_f \) and \( V_H \) are orthogonal with respect to the \( L^2 \)-inner product.

Let \( V_H = \text{span}\{ \phi_i \} \) and \( V_f = \text{span}\{ \varphi_i \} \). Then, the system of equations to be solved on the fine scale is given by

\[
\begin{pmatrix}
K & P^T \\
P & 0
\end{pmatrix}
\begin{pmatrix}
\xi
\end{pmatrix}
= 
\begin{pmatrix}
b \\
0
\end{pmatrix},
\]

where

\[
P = 
\begin{pmatrix}
(\phi_1, \varphi_1) & (\phi_1, \varphi_2) & \ldots & (\phi_1, \varphi_N) \\
(\phi_2, \varphi_1) & (\phi_2, \varphi_2) & \ldots & (\phi_2, \varphi_N) \\
\vdots & \vdots & \ddots & \vdots \\
(\phi_M, \varphi_1) & (\phi_M, \varphi_2) & \ldots & (\phi_M, \varphi_N)
\end{pmatrix},
\]
with $K_{k,l} = a_i(\varphi_k, \varphi_l)$ and $b$ either $b_k = l_i(\varphi_k)$ for (3.10) or $b_k = -a_i(\varphi_i, \varphi_k)$ for (3.9).

Using the a posteriori error estimate above, it is possible to design an adaptive algorithm that automatically tunes the fine mesh size and the size of the patches. In the numerical experiments below, we have implemented Algorithm 1, which extends the patches in all directions and uses a uniform mesh refinement of the fine scale on each coarse element. A more elaborate algorithm, which only extends in the direction where the error is large and uses adaptive mesh refinement, would be a possible extension, since the a posteriori indicators above contain local contributions of each individual patch-boundary face and of each fine scale element residual.

Algorithm 1. Adaptive discontinuous Galerkin multiscale method.

1: Initialize the coarse mesh, $\mathcal{K}_H$ with mesh function $H$, and a fine mesh, $\mathcal{K}_h$ with meshfunction $h$, by using to uniform refinements of $\mathcal{K}_H$; i.e., $h = H/4$.

2: For all $K_H$ let the size of the patches be $\omega_{K_H}^3$.

3: Set the mesh refinement level to $X\%$.

4: while $(\sum_{K \in K} \varrho^2_{h,K})^{1/2} + (\sum_{K \in K} \zeta^2_{h,K})^{1/2} + (\sum_{K_H \in K_H} \rho^2_{L_{K_H}})^{1/2} > TOL$ do

5: for $K \in \tilde{K}_H$ do

6: Solve the fine scale problems (3.1) and (3.10).

7: Compute the matrix and vector entries on the coarse scale (5.1).

8: end for

9: Solve the modified coarse scale problem (3.11).

10: Mark the indicator with $X\%$ largest error in $\{\varrho^2_{h,K}, \zeta^2_{h,K}, \rho^2_{L_{K_H}}, \omega_{K_H}\}$.

11: for $K_H \in K_H$ do

12: if $\rho^2_{L_{K_H}}$ is marked then

13: $\omega_{K_H}^L := \omega_{K_H}^{L+1}$

14: end if

15: if $\rho^2_{h,K} + \zeta^2_{h,K}$ is marked then

16: $h|_{K_H} := h|_{K_H}/2$

17: end if

18: end for

19: end while

6. Numerical examples. We present some numerical experiments where the converge of the method as well as the performance of the adaptive algorithm is investigated.

6.1. Convergence. We consider the model problem (2.1)–(2.2) on the $L$-shaped domain constructed by removing the lower right quadrant in the unit square, with forcing function $f = 1$. We consider a coarse quadrilateral mesh of size $H = 2^{-4}$. Furthermore, each coarse element $K \in \mathcal{K}_H$ is further subdivided using two uniform refinements to construct the fine mesh. The error is measured in the relative energy norm, (2.12), where $u_h$ is the DG solution on the fine mesh; i.e., there is only a truncation error (due to the fine scale patch size) between the multiscale solution and the DG solution. The permeabilities One and SPE,\(^1\) illustrated in Figure 2, are used. In One, we have $A = 1$, and in SPE the data is taken from the tenth SPE comparative solution project and is projected into the fine mesh. Exponential decay is observed with respect to the number of layers for the different permeabilities One and SPE.

\(^1\)Data is taken from the tenth SPE comparative solution project http://www.spe.org/web/csp/.
6.2. Adaptivity for a problem with analytic solution. Let us consider the model problem (2.1)–(2.2) on a unit square, using the permeability $A = 1$ and the forcing function $f = 4a^2(1 - ar^2)e^{-ar^2}$ for some $a > 0$. Using $a = 400$, the analytic solution can be approximated sufficiently well by the Gaussian pulse $u = ae^{-ar^2}$, centered in the middle of the domain. We consider a coarse quadrilateral mesh of size $H = 2^{-4}$, and a fine mesh of size $h = 2^{-6}$. The adaptive algorithm (Algorithm 1) with 10% refinement level is used. The starting values for $L$ and $h$ used are $L = 3$ layers, and the fine scale mesh is uniformly refined two times. Figure 4 shows the error and the error indicators decay after each iteration of the adaptive algorithm, while Figure 5 shows the locations where the adaptive algorithm has chosen to concentrate the computational effort, which indeed coincides with the position of the pulse.
6.3. Adaptivity on an $L$-shaped domain. Consider the model problem and the same data as in section 6.1. The solution produced by the adaptive algorithm is compared to a reference solution computed with the standard (one scale) DG method on a uniform quadrilateral mesh with mesh-size $h = 2^{-9}$; see Figure 6. Consider a coarse mesh consisting of a uniform quadrilateral mesh of size $H = 2^{-4}$. The starting values in the adaptive algorithm (Algorithm 1) are $L = 3$ and the fine scale mesh is derived by two uniform refinements of the coarse mesh. In each iteration, a refinement level of 30\% is used. Figure 7 shows the error decays after each iteration of the adaptive algorithm. Also, the adaptive algorithm chooses to increase the patches in the beginning since the error from the truncation is initially larger than the discretization error and after a few iterations it is starting to refine the fine scale mesh more and more. When the patch sizes are increased, the error, due to truncation, decays exponentially independent of the regularity of the solution as shown theoretically in [17]. This is not true for the discretization error. This motivates the use of an adaptive algorithm which tunes the error between the truncation and discretization. Figure 8...
Fig. 6. Reference solution for the different permeabilities computed onto a mesh with size $h = 2^{-9}$ and projected onto a mesh with size $h = 2^{-6}$.

Fig. 7. The relative energy norm error for the multiscale solution using the adaptive algorithm; $\rho_L$ denotes the truncation error indicator and $\varrho_K$ and $\zeta_K$ are the discretization error indicators.
6.4. Adaptivity for a porous media flow problem. We consider the problem (2.1)--(2.2) on the unit square $\Omega = [0,1]^2$, with forcing function $f = -1$ in the lower left corner $\{0 \leq x, y \leq 1/128\}$, $f = 1$ in the upper right corner $\{127/128 \leq x, y \leq 1\}$, and $f = 0$ otherwise. The following permeabilities $SPE11$ and $SPE21$ are used and projected into a mesh with $64 \times 64$ elements; see Figure 9. The computational domain $\Omega$ is split into $32 \times 32$ coarse square elements $K_H \in K_H$. The error is measured in the relative energy norm, with the reference solution $u_h$ being the DG solution computed on a $512 \times 512$-element mesh. The adaptive algorithm (Algorithm 1) with refinement level 30% is used. In Iteration 1 the multiscale problem is solved using two refinements on each coarse element and each fine scale problem is solved with $L = 3$, and so on. Even though complicated permeabilities with $\alpha_{max}/\alpha_{min} \sim 10^5$ are used, the proposed adaptive algorithm is able to reduce relative error considerably; this is shown in Figure 10.

**Fig. 8.** The level of refinement and size of the patches illustrated in the upper, respectively lower plots, for the different permeability One (left) and SPE (right). White is where most refinements, respectively larger patch, are used and black is where least refinements, respectively smallest patches, are used.

**Fig. 9.** Permeabilities projection in log scale.

shows where the adaptive algorithm put the most computational effort.

(a) $SPE11$, $\alpha_{max}/\alpha_{min} = 6.1765e - 5$.
(b) $SPE21$, $\alpha_{max}/\alpha_{min} = 5.0193e - 5$. 

---

D. ELFVERSON, E. H. GEORGOULIS, AND A. MÅLQVIST
7. Concluding remarks. An adaptive multiscale method based on discontinuous Galerkin discretization has been proposed and assessed in practice. There are several different advantages in using the proposed multiscale method. The possibility to allow a global underlying reference grid (using the DG framework including hanging nodes) is crucial. This does not only account for cancellation of the error between different fine scale problems in the a posteriori error bound, it also fits the method into the convergence framework presented in [17]. It admits a local conservation of the state variable, which is crucial in many applications, e.g., porous media flow. The multiscale method and the adaptive algorithm admit naturally parallel implementation, which results in further savings in computational time.

An adaptive algorithm for which the coarse scale, the fine scale, and the size of the different patches are taken into account, based on an energy norm a posteriori bound has been proposed. Using the proposed multiscale method, together with the adaptive algorithm, leads to substantial computational savings, while maintaining a good performance when applied to challenging benchmark problems.

REFERENCES


