CONVERGENCE OF A DISCONTINUOUS GALERKIN MULTISCALE METHOD*

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

Abstract. We present a discontinuous Galerkin multiscale method for second order elliptic problems and prove convergence. We consider a heterogeneous and highly varying diffusion coefficient in $L^\infty(\Omega, \mathbb{R}^{d \times d}_{sym})$ with uniform spectral bounds without any assumption on scale separation or periodicity. The multiscale method uses a corrected basis that is computed on patches/subdomains. The error, due to truncation of the corrected basis, decreases exponentially with the size of the patches. Hence, to achieve an algebraic convergence rate of the multiscale solution on a uniform mesh with mesh size H to a reference solution, it is sufficient to choose the patch sizes $\mathcal{O}(H|\log H|)$. We also discuss a way to further localize the corrected basis to elementwise support. Improved convergence rate can be achieved depending on the piecewise regularity of the forcing function. Linear convergence in energy norm and quadratic convergence in the L^2-norm is obtained independently of the forcing function. A series of numerical experiments confirms the theoretical rates of convergence.

Key words. multiscale method, discontinuous Galerkin, a priori error estimate, convergence

AMS subject classifications. 65N12, 65N30

DOI. 10.1137/120900113

1. Introduction. This work considers the numerical solution of second order elliptic problems with a heterogeneous and highly varying (nonperiodic) diffusion coefficient. The heterogeneities and oscillations of the coefficient may appear on several nonseparated scales. More specifically, let $\Omega \subset \mathbb{R}^d$ be a bounded Lipschitz domain with polygonal boundary Γ. The boundary Γ may be partitioned into some subset Γ_D (the Dirichlet boundary) with positive measure and its complement $\Gamma_N := \Gamma \setminus \Gamma_D$ (the, possibly empty, Neumann boundary). We assume that the diffusion matrix $A \in L^\infty(\Omega, \mathbb{R}^{d \times d}_{\text{sym}})$ has uniform spectral bounds $0 < \alpha, \beta < \infty$, defined by

$$(1.1) \qquad 0 < \alpha := \operatorname*{ess\,inf}_{x \in \Omega} \inf_{v \in \mathbb{R}^d \setminus \{0\}} \frac{(A(x)v) \cdot v}{v \cdot v} \leq \operatorname*{ess\,sup}_{x \in \Omega} \sup_{v \in \mathbb{R}^d \setminus \{0\}} \frac{(A(x)v) \cdot v}{v \cdot v} =: \beta < \infty.$$

Given $f \in L^2(\Omega)$, we seek the weak solution to the boundary-value problem

$$-\nabla \cdot A\nabla u = f \quad \text{in } \Omega,$$
$$u = 0 \quad \text{on } \Gamma_D,$$
$$\nu \cdot A\nabla u = 0 \quad \text{on } \Gamma_N,$$

i.e., we seek $u \in H_D^1(\Omega) := \{v \in H^1(\Omega) \mid v|_{\Gamma_D} = 0\}$ such that

$$(1.2) \qquad a(u,v) := \int_\Omega A\nabla u \cdot \nabla v \, \mathrm{d}x = \int_\Omega fv \, \mathrm{d}x =: F(v) \quad \text{for all } v \in H_D^1(\Omega).$$

Many methods have been developed in recent years to overcome the lack of performance of classical finite element methods when A is rough, meaning that A has discontinuities and/or high variation; we refer to [6, 4, 19, 8, 1, 2], among others. Common to all the aforementioned approaches is the idea to solve problems on small subdomains and to use the results to construct a better basis for some Galerkin method or to modify the coarse scale operator. However, apart from the one-dimensional setting, the performance of those methods correlates strongly with periodicity and scale separation of the diffusion coefficient. There has also been work to design a hierarchical basis such that the multigrid convergence rate does not depended on the variation in the coefficients, e.g., [29], where they assume that the diffusion coefficient fulfills a so-called quasi-monotone property.

Other approaches [7, 28, 5, 11, 12] perform well without any assumptions on periodicity or scale separation in the diffusion coefficient at the price of a high computational cost: in [7, 28] the support of the modified basis functions is large and in [5, 11, 12] the computation of the basis functions involves the solutions of local eigenvalue problems.

In the framework of the variational multiscale method (VMS), introduced in [21, 22], the space for which the solution is sought is split into a coarse and a fine scale contribution. Writing the fine scale contribution in terms of the coarse scale residuals eliminates it from the coarse scale equation. This was first employed in an adaptive setting in the adaptive VMS [24], where the basic idea is to split the fine scale residuals into localized contributions solved on element patches, possibly larger than a single element, with the Dirichlet boundary condition. Using the solution from the fine scale patches a modified nonsymmetric (Petrov–Galerkin) formulation is obtained on the coarse scale. An a posteriori error bound is derived and used within an adaptive algorithm to automatically tune the coarse and fine mesh size as well as the size of the patches.

An abstract framework for constructing multiscale methods for elliptic partial differential equations using the VMS framework is derived in [27]. Both symmetric and nonsymmetric (Petrov–Galerkin) formulations are considered and an a posteriori error bound is derived both for convection-diffusion-reaction problems and for the Poisson's equations on mixed form.

Only recently in [25] the first rigorous a priori error bound for a VMS was derived, which allows for textbook convergence with respect to the mesh size H, $\|u - u_H\|_{H^1(\Omega)} \le C_{f,\beta/\alpha} H$ with a constant $C_{f,\beta/\alpha}$ that depends on f and the global bounds of the diffusion coefficient but not on its variations. This result, which is a natural extension of [24, 27], is achieved using a local orthogonal decomposition technique, where an operator dependent modification of the classical nodal basis is constructed using the solution of local problems on vertex patches of diameter $\mathcal{O}(H|\log H|)$. In the error analysis, the size of the patches depends on the global bound of the diffusion coefficient. This indicates that it may not be suited for high contrast (or degenerate) problems; however, numerical experiments show promising results also for these cases [25]. The methodology has been extended to semilinear elliptic problems [16] and (non)linear eigenvalue problems [26, 15]. In [17] it is shown that the approach may

also be interpreted as a multiscale finite element method, in the sense of [18], with some novel oversampling strategy.

In this work, we present a discontinuous Galerkin (dG) multiscale method with similar performance as [25]. We show that the error between the exact solution and the solution obtained by the dG multiscale method converge as $\tilde{C}_{f,\beta/\alpha}H$ in standard dG energy norm for $f \in L^2(\Omega)$. Higher convergence rates (up to $\tilde{C}_{f,\beta/\alpha}H^3$) can be obtained depending on the elementwise regularity of f. We also give an error bound for a quantity of interest (a linear functional of the solution) and the convergence rate $C'_{f,\beta/\alpha}H^2$ (up to $C'_{f,\beta/\alpha}H^4$) in the L^2-norm follows. Adaptivity for the dG multiscale method is considered in [13] and an extension of the a priori analysis to convection dominated convection-diffusion-reaction problems is considered in [14]. Since the dG method seeks the solution in a nonconforming space, the elementwise L^2-projection as the split between the coarse and fine contribution is now admissible. This is a more natural choice than, e.g., the nodal interpolant used in [24] for multiscale applications and may lead to better performance of the dG-based multiscale method (compared to conforming variants) for eigenvalue computations [26, 15].

The dG finite element method admits good conservation properties of the state variable and also offers the use of very general meshes due to the lack of interelement continuity requirements, e.g., meshes that contain several different types of elements and/or hanging nodes. Both these features are crucial in many multiscale applications.

Although the error analysis presented in this work is restricted to regular simplicial or quadrilateral/hexahedral meshes, we stress that all the results appear to be extendable for the case of irregular meshes (i.e., meshes containing hanging nodes). We refrained from presenting these extensions here for simplicity of the current presentation. Under these assumptions, we provide a complete a priori error analysis of this method including errors caused by the approximation of basis functions.

In this dG multiscale method and in previous related methods [25, 13], the accuracy is ensured by enlarging the support of basis functions appropriately. Hence, supports of basis functions overlap and the communication is no longer restricted to neighboring elements but is present also between elements at a certain distance. This overlap leads to a slight decrease of sparsity of the coarse stiffness matrix. We will show that the overhead is acceptable in the sense that it scales only logarithmically with respect to the coarse mesh size.

In order to retain the dG-typical sparse structure of the stiffness matrix with communication restricted to neighboring elements only, we discuss the possibility of localizing the multiscale basis functions to single elements. Instead of having $\mathcal{O}(1)$ basis functions per element with $\mathcal{O}(H|\log H|)$ support, we would then have $\mathcal{O}(|\log H|)$ basis functions per element with element support. The elementwise application of an eigenvalue decomposition easily prevents ill-conditioning of the element stiffness matrices, while simultaneously offering further compression of the multiscale basis.

The outline of the paper is as follows. In section 2, we recall the dG finite element method. Section 3 defines our multiscale method, which is then analyzed in section 4. Section 5 presents numerical experiments confirming the theoretical developments. Finally, in section 6 we draw some conclusions.

Throughout this paper, standard notation for Lebesgue and Sobolev spaces is employed. Let $0 \leq C < \infty$ be any generic constant that depends on neither the mesh size nor the diffusion matrix A; $a \lesssim b$ abbreviates an inequality $a \leq C b$ and $a \approx b$ abbreviates $a \lesssim b \lesssim a$. Also, let the constant $C_{\beta/\alpha}$ depend on the minimum and maximum bounds (α and β) of the diffusion matrix A in (1.1).

2. Fine scale discretization.

2.1. Finite element meshes and spaces. Let $\mathcal{T}$ denote a subdivision of Ω into (closed) regular simplices or into quadrilaterals (for $d = 2$) or hexahedra (for $d = 3$), i.e., $\bar{\Omega} = \cup_{T \in \mathcal{T}} T$. We assume that $\mathcal{T}$ is conforming in the sense that any two elements are either disjoint or share exactly one edge or vertex.

Let $\mathcal{E}$ denote the set of edges (or faces for $d = 3$) of $\mathcal{T}$; $\mathcal{E}(\Omega)$ denotes the set of interior edges; and $\mathcal{E}(\Gamma)$, $\mathcal{E}(\Gamma_D)$, and $\mathcal{E}(\Gamma_N)$ refer to the set of edges on the boundary of Ω, on the Dirichlet, and on the Neumann boundary, respectively. Let $\hat{T}$, denote the reference simplex or (hyper)cube and let $\mathcal{P}_p(\hat{T})$ and $\mathcal{Q}_p(\hat{T})$ denote the spaces of polynomials of degree less than or equal to p in all and on each variable, respectively. We define the set of piecewise polynomials

$$P_p(\mathcal{T}) := \{v : \Omega \to \mathbb{R} \mid \text{for all } T \in \mathcal{T}, v|_T \circ F_T \in \mathcal{R}_p(\hat{T})\}$$

with $\mathcal{R}_p \in \{\mathcal{P}_p, \mathcal{Q}_p\}$, where $F_T : \hat{T} \to T$, $T \in \mathcal{T}$ is a family of element maps. Let also

$$(2.1) \qquad\qquad \Pi_p(\mathcal{T}) : L^2(\Omega) \to P_p(\mathcal{T})$$

denote the L^2-projection onto $\mathcal{T}$-piecewise polynomial functions of order p. In particular, we have $(\Pi_0(\mathcal{T})f)|_T = |T|^{-1} \int_T f \, dx$, $T \in \mathcal{T}$, for all $f \in L^2(\Omega)$. Note that $v \in P_p(\mathcal{T})$ does not necessarily belong to $H^1(\Omega)$. The $\mathcal{T}$-piecewise gradient $\nabla_\mathcal{T} v$, with $(\nabla_\mathcal{T} v)|_T = \nabla(v|_T)$ for all $T \in \mathcal{T}$, is well-defined and $\nabla_\mathcal{T} v \in (P_{p-1}(\mathcal{T}))^d$.

For any interior edge/face $e \in \mathcal{E}(\Omega)$ there are two adjacent elements T^- and T^+ with $e = \partial T^- \cap \partial T^+$. We define ν to be the normal vector of e that points from T^- to T^+. For boundary edges/faces $e \in \mathcal{E}(\Gamma)$ let ν be the outward unit normal vector of Ω.

Define the jump of $v \in P_k(\mathcal{T})$ across $e \in \mathcal{E}(\Omega)$ by $[v] := v|_{T^-} - v|_{T^+}$ and define $[v] := v|_e$ for $e \in \mathcal{E}(\Gamma)$. The average of $v \in P_p(\mathcal{T})$ across $e \in \mathcal{E}(\Omega)$ is defined by $\{v\} := (v|_{T^-} + v|_{T^+})/2$ and for boundary edges $e \in \mathcal{E}(\Gamma)$ by $\{v\} := v|_e$. Also, we make the shorthand notation $\mathcal{E}(\Omega \cup \Gamma) = \mathcal{E}(\Omega) \cup \mathcal{E}(\Gamma)$.

In the remaining part of this work, we consider two different meshes: a coarse mesh $\mathcal{T}_H$ and a fine mesh $\mathcal{T}_h$, with respective definitions for the edges/faces $\mathcal{E}_H$ and $\mathcal{E}_h$. We denote the $\mathcal{T}_H$-piecewise gradient by $\nabla_H v := \nabla_{\mathcal{T}_H} v$ and, respectively, $\nabla_h v := \nabla_{\mathcal{T}_h} v$ for the $\mathcal{T}_h$-piecewise gradient. We assume that the fine mesh $\mathcal{T}_h$ is the result of one or more refinements of the coarse mesh $\mathcal{T}_H$. The subscripts h, H refer to the corresponding mesh sizes; in particular, we have $H \in P_0(\mathcal{T}_H)$ with $H|_T = \text{diam}(T) =: H_T$ for all $T \in \mathcal{T}_H$, $H_e = \text{diam } e$ for all $e \in \mathcal{E}_H$, $h \in P_0(\mathcal{T}_h)$ with $h|_T = \text{diam}(T) =: h_T$ for all $T \in \mathcal{T}_h$, and $h_e = \text{diam } e$ for all $e \in \mathcal{E}_h$. Obviously, $h \leq H$. For simplicity we assume that the discontinuities in A are aligned with the fine mesh $\mathcal{T}_h$.

2.2. Discretization by the symmetric interior penalty method. We consider the symmetric interior penalty method (SIP) dG method [9, 3, 20]. We seek an approximation in the space $\mathcal{V}_h := P_1(\mathcal{T}_h)$. Given some positive penalty parameter $\sigma > 0$, we define the symmetric bilinear form $a_h : \mathcal{V}_h \times \mathcal{V}_h \to \mathbb{R}$ by

$$(2.2) \qquad a_h(u, v) := (A\nabla_h u, \nabla_h v)_{L^2(\Omega)} - \sum_{e \in \mathcal{E}_h(\Omega \cup \Gamma_D)} \Bigg((\{\nu \cdot A\nabla u\}, [v])_{L^2(e)}$$

$$+ (\{\nu \cdot A\nabla v\}, [u])_{L^2(e)} - \frac{\sigma}{h_e}([u], [v])_{L^2(e)} \Bigg).$$

The jump-seminorm associated with the space $\mathcal{V}_h$ is defined by

$$| \bullet |_h^2 := \sum_{e \in \mathcal{E}_h(\Omega \cup \Gamma_D)} \frac{\sigma}{h_e} \|[\bullet]\|_{L^2(e)}^2, \tag{2.3}$$

while the energy norm in $\mathcal{V}_h$ is then given by

$$||| \bullet |||_h := (\|A^{1/2} \nabla_h \bullet\|_{L^2(\Omega)}^2 + | \bullet |_h^2)^{1/2}. \tag{2.4}$$

If the penalty parameter is chosen sufficiently large, the dG bilinear form (2.2) is coercive and bounded with respect to the energy norm (2.4).

Remark 1. The penalty parameter, σ, depends of the arithmetic mean of diffusion coefficient on edge e. If a SIP with a weighted average would be used [10], the penalty parameter, σ, would instead depend on the harmonic average of the diffusion coefficient. For simplicity of the presentation and since this choice suffices for our purposes, here we consider the standard SIP dG formulation. Hence, there exists a (unique) dG approximation $u_h \in \mathcal{V}_h$, satisfying

$$a_h(u_h, v) = F(v) \quad \text{for all } v \in \mathcal{V}_h. \tag{2.5}$$

We assume that (2.5) is computationally intractable for practical problems, so we shall never seek to solve for u_h directly. Instead, u_h will serve as a reference solution to compare our low dimensional coarse grid multiscale dG approximation with. The underlying assumption is that the mesh $\mathcal{T}_h$ is chosen sufficiently fine so that u_h is sufficiently accurate. The aim of this work is to devise and analyze a multiscale dG discretization with coarse scale H in such a way that the accuracy of u_h is preserved up to an $\mathcal{O}(H)$ perturbation independent of the variation of the coefficient A.

3. Discontinuous Galerkin multiscale method. As mentioned above, the choice of the reference mesh $\mathcal{T}_h$ is not directly related to the desired accuracy but is instead strongly affected by the roughness and variation of the coefficient A. The corresponding coarse mesh $\mathcal{T}_H$, with mesh width function $H \geq h$, is assumed to be completely independent of A. In the spirit of [21, 22] the test space is divided into coarse and fine components, where the fine scale components are computed on the patches (submeshes) of the reference mesh. To encapsulate the fine scale information in the coarse mesh, we shall design coarse generalized finite element spaces based on $\mathcal{T}_H$.

3.1. Multiscale decompositions. We introduce a two-scale splitting for the space $\mathcal{V}_h$. To this end, let $\Pi_H := \Pi_1(\mathcal{T}_H)$, from (2.1), and define $\mathcal{V}_H := \Pi_H \mathcal{V}_h = P_1(\mathcal{T}_H)$ and

$$\mathcal{V}^{\text{f}} := (1 - \Pi_H)\mathcal{V}_h = \{v \in \mathcal{V}_h \mid \Pi_H v = 0\}.$$

LEMMA 2 (L^2-orthogonal multiscale decomposition). *The decomposition*

$$\mathcal{V}_h = \mathcal{V}_H \oplus \mathcal{V}^{\text{f}}$$

is orthogonal in $L^2(\Omega)$.

Proof. The proof is immediate, as any $v \in \mathcal{V}_h$ can be decomposed uniquely into a coarse finite element function $v_H := \Pi_H v \in \mathcal{V}_H$ and a (possibly highly oscillatory) remainder $v^{\text{f}} := (1 - \Pi_H)v \in \mathcal{V}^{\text{f}}$ with $\|v\|_{L^2(\Omega)}^2 = \|v_H\|_{L^2(\Omega)}^2 + \|v^{\text{f}}\|_{L^2(\Omega)}^2$. $\square$

We now orthogonalize the above splitting with respect to the dG scalar product a_h; we keep the space of fine scale oscillations $\mathcal{V}^{\text{f}}$ and simply replace $\mathcal{V}_H$ with the orthogonal complement of $\mathcal{V}^{\text{f}}$ in $\mathcal{V}_h$. We define the fine scale projection $\mathfrak{F} : \mathcal{V}_h \to \mathcal{V}^{\text{f}}$ by

(3.1) $a_h(\mathfrak{F}v, w) = a_h(v, w)$ for all $w \in \mathcal{V}^{\mathrm{f}}$.

Using the fine scale projection, we can define the coarse scale approximation space by

$$\mathcal{V}_H^{ms} := (1 - \mathfrak{F})\mathcal{V}_H.$$

LEMMA 3 (a_h-orthogonal multiscale decomposition). *The decomposition*

$$\mathcal{V}_h = \mathcal{V}_H^{ms} \oplus \mathcal{V}^{\mathrm{f}}$$

is orthogonal with respect to a_h, i.e., any function v in $\mathcal{V}_h$ can be decomposed uniquely into some function $v_H^{ms} \in \mathcal{V}_H^{ms}$ plus $v^{\mathrm{f}} \in \mathcal{V}^{\mathrm{f}}$ with $C^{-1}|||v|||_h^2 \leq |||v_H^{ms}|||_h^2 + |||v^{\mathrm{f}}|||_h^2 \leq C|||v|||_h^2$, where the constant C only depends on the coercivity and continuity constants of the bilinear form. The functions $v_H^{ms} \in \mathcal{V}_H^{ms}$ and $v^{\mathrm{f}} \in \mathcal{V}^{\mathrm{f}}$ are the Galerkin projections of $v \in \mathcal{V}_h$ onto the subspaces $\mathcal{V}_H^{ms}$ and $\mathcal{V}^{\mathrm{f}}$, i.e.,

$$a_h(v_H^{ms}, w) = a_h(v, w) \quad \textit{for all } w \in \mathcal{V}_H^{ms},$$
$$a_h(v^{\mathrm{f}}, w) = a_h(v, w) \quad \textit{for all } w \in \mathcal{V}^{\mathrm{f}}.$$

The unique Galerkin approximation $u_H^{ms} \in \mathcal{V}_H^{ms}$ of $u \in \mathcal{V}$ solves

(3.2) $a_h(u_H^{ms}, v) = F(v)$ for all $v \in \mathcal{V}_H^{ms}$.

We shall see in the error analysis (cf. Theorem 9) that the orthogonality yields error estimates (with respect to a reference solution) for the Galerkin approximation $u_H^{ms} \in \mathcal{V}_H^{ms}$ of (3.2) that are independent of the regularity of the solution and of the variation in the diffusion coefficient A. However, the space $\mathcal{V}_H^{ms}$ is not suitable for practical computations as a local basis for this space is not easily available. Indeed, given a basis of $\mathcal{V}_H$, e.g., the elementwise Lagrange basis functions $\{\lambda_{T,j} \mid T \in \mathcal{T}_H, \ j = 1, \ldots, r\}$, where $r = (1 + d)$ for regular simplices or $r = 2^d$ for quadrilaterals/hexahedra, the space $\mathcal{V}_H^{ms}$ is spanned by the corrected basis functions $(1 - \mathfrak{F})\lambda_{T,j}$, $T \in \mathcal{T}_H$, $j = 1, \ldots, r$. Although $\lambda_{T,j}$ has local support $\mathrm{supp}\,\lambda_{T,j} = T$, its corrected version $(1 - \mathfrak{F})\lambda_{T,j}$ has global support in Ω, as (3.1) is a variational problem on the whole domain Ω. Fortunately, as we shall prove below, the corrector functions $\phi_{T,j}$ decay quickly away from T. (See previous numerical results in [13] and a similar observation for a related conforming method [25].) This decay motivates the local approximation of the corrector functions at the expense of introducing small perturbations in the method's accuracy.

3.2. Localization and computational method. The localized approximations of the corrector functions are supported on element patches in the coarse mesh $\mathcal{T}_H$.

DEFINITION 4. *For all $T \in \mathcal{T}_H$, define element patches with size L as*

$$\omega_T^1 := \mathrm{int}(T),$$
$$\omega_T^L := \mathrm{int}(\cup\{T' \in \mathcal{T}_H \mid T' \cap \bar{\omega}_T^{L-1} \neq \emptyset\}), \quad L = 1, 2, \ldots.$$

We refer to Figure 3.1 *for an illustration.*

We introduce a new discretization parameter $0 < L \in \mathbb{N}$ and define localized corrector functions $\phi_{T,j}^L \in \mathcal{V}^{\mathrm{f}}(\omega_T^L) := \{v \in \mathcal{V}^{\mathrm{f}} \mid v|_{\Omega \setminus \omega_T^L} = 0\}$ by

(3.3) $a_h(\phi_{T,j}^L, w) = a_h(\lambda_{T,j}, w)$ for all $w \in \mathcal{V}^{\mathrm{f}}(\omega_T^L)$.
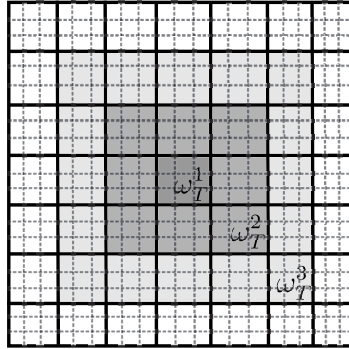
FIG. 3.1. *Example of a one-layer patch ω_T^1, a two-layer patch ω_T^2, and a three-layer patch ω_T^3 on a quadrilateral mesh.*

Further, we define the multiscale approximation space

$$(3.4) \qquad \mathcal{V}_H^{ms,L} = \text{span}\{\lambda_{T,j} - \phi_{T,j}^L \mid T \in \mathcal{T}_H, \ j = 1, \ldots, r\}.$$

The dG multiscale method seeks $u_H^{ms,L} \in \mathcal{V}_H^{ms,L}$ such that

$$(3.5) \qquad a_h(u_H^{ms,L}, v) = F(v) \quad \text{for all } v \in \mathcal{V}_H^{ms,L}.$$

Since $\mathcal{V}_H^{ms,L} \subset \mathcal{V}_h$, this method is a Galerkin method in the Hilbert space $\mathcal{V}_h$ (with scalar product a_h) and hence inherits well-posedness from the reference discretization (2.5).

Moreover, the proposed basis $\{\lambda_{T,j} - \phi_{T,j}^L \mid T \in \mathcal{T}_H, \ j = 1, \ldots, r\}$ is stable with respect to the fine scale parameter h, as we shall see in Lemma 8 below.

3.3. Compressed dG multiscale method. The basis functions in the above multiscale method have enlarged supports (element patches) when compared with standard dG methods (elements). We can decompose the corrector functions into its element contributions

$$\phi_{T,j}^L = \sum_{T' \in \mathcal{T}_H : T' \subset \omega_T^L} \phi_{T,j}^L \chi_{T'},$$

where $\chi_{T'}$ is the indicator function of the element $T' \in \mathcal{T}_H$.

This motivates the coarse approximation space

$$\mathcal{W}_H^{ms,L} = \text{span}\big(\{\lambda_{T,j} \mid T \in \mathcal{T}_H, \ j = 1, \ldots, r\}$$
$$\cup \{\phi_{T,j}^L \chi_{T'} \mid T, T' \in \mathcal{T}_H, \ T' \subset \omega_T^L, \ j = 1, \ldots, r\}\big).$$

This space offers the advantage of a known basis with elementwise support which leads to improved (localized) connectivity in the corresponding stiffness matrix. This is at the expense of a slight increase in the dimension of the space

$$(3.6) \qquad \dim(\mathcal{W}_H^{ms,L}) \approx L^d \dim(\mathcal{V}_H^{ms,L}).$$

The corresponding localized dG multiscale method seeks $w_H^{ms,L} \in \mathcal{W}_H^{ms,L}$ such that

$$(3.7) \qquad a_h(w_H^{ms,L}, v) = F(v) \quad \text{for all } v \in \mathcal{W}_H^{ms,L}.$$

Since $\mathcal{V}_H^{ms,L} \subset \mathcal{W}_H^{ms,L} \subset \mathcal{V}_h$, Galerkin orthogonality yields

$$(3.8) \qquad |||u_h - w_H^{ms,L}|||_h \le |||u_h - u_H^{ms,L}|||_h,$$

i.e., the new localized version (3.7) is never worse than the previous multiscale approximation in terms of accuracy. However, it may lead to very ill-conditioned element stiffness matrices. (See Lemma 11, which shows that $\phi_{T,j}^L \chi_{T'}$ may be very small if the distance between T and T' relative to their sizes is large.)

To circumvent ill-conditioning, one may choose a reduced local approximation space based on an eigendecomposition of the element stiffness matrix. The eigenfunctions which correspond to sufficiently large eigenvalues (principal components) are used as basis functions for the reduced space. Since the dimension of the element stiffness matrix is small (at most proportional to $L^d \times L^d$), the cost of this additional preprocessing step is negligible when compared with the cost of solving the local problems for the corrector functions.

To determine an acceptable level of truncation of the localized basis functions, we can use the a posteriori error estimator contribution of the local problem from [13], which is an estimation of the local fine scale error. Using an adaptive algorithm in [13] to determine the size of the patches may additionally lead to large reduction of the dimension of the local approximation spaces (3.6), since in (3.6) all the patches are assumed to have the same size L.

4. Error analysis. We present an a priori error analysis for the proposed multiscale method (3.5). In view of (3.8), this analysis applies immediately to the modified versions presented in section 3.3. The error analysis will be split into a number of steps. First, in section 4.1, we present some properties of the coarse scale projection operator Π_H. In section 4.2, an error bound for dG multiscale method u_H^{ms} from (3.2) (Theorem 9) is shown, whereby the corrected basis functions are solved globally. Results for the decay of the localized corrected basis function (Lemmas 11 and 12) are shown, along with an error bound for the dG multiscale method $u_H^{ms,L}$ from (3.5) (Theorem 13), where the corrected basis functions are solved locally on element patches. Finally, in section 4.3, we show an error bound given a quantity of interest (Theorem 15), leading to an error bound in the L^2-norm (Corollary 16).

We shall make use of the following (semi)norms. The jump-seminorm and energy norms, associated with the coarse space $\mathcal{V}_H$, are defined by

$$|\bullet|_H^2 := \sum_{e \in \mathcal{E}_H(\Omega \cup \Gamma_D)} \frac{\sigma}{H_e} \|[\bullet]\|_{L^2(e)}^2,$$

$$|||\bullet|||_H := (\|A^{1/2}\nabla_H \bullet\|_{L^2(\Omega)}^2 + |\bullet|_H^2)^{1/2},$$

respectively, along with a localized version of the local jump and energy norms (2.3) and (2.4) on a patch $\omega \subseteq \Omega$, where ω is aligned with the mesh $\mathcal{T}_h$, given by

$$|\bullet|_{h,\omega}^2 := \sum_{\substack{e \in \mathcal{E}_h(\Omega \cup \Gamma_D): \\ e \cap \bar{\omega} \ne 0}} \frac{\sigma}{h_e} \|[\bullet]\|_{L^2(e)}^2,$$

$$|||\bullet|||_{h,\omega} := (\|A^{1/2}\nabla_h \bullet\|_{L^2(\omega)}^2 + |\bullet|_{h,\omega}^2)^{1/2}.$$

The shape-regularity assumptions $h_T \approx h_e$ for all $e \in \partial T : T \in \mathcal{T}_h$ and $H_T \approx H_e$ for all $T \in \partial T : T \in \mathcal{T}_H$ will also be used.

4.1. Properties of the coarse scale projection operator Π_H. The following lemma gives stability and approximation properties of the operator Π_H.

LEMMA 5. *For any $v \in \mathcal{V}_h$, the estimate*

$$H^{-1}\|v - \Pi_H v\|_{L^2(T)} \lesssim \alpha^{-1/2}|||v|||_{h,T}$$

is satisfied for all $T \in \mathcal{T}_H$. Moreover, it holds that

$$\beta^{-1/2}|||\Pi_H v|||_H + \|H^{-1}(v - \Pi_H v)\|_{L^2(\Omega)} \lesssim \alpha^{-1/2}|||v|||_h,$$

where α and β are defined in (1.1).

Proof. Theorem 2.2 in [23] implies that for each $v \in \mathcal{V}_h$, there exists a bounded linear operator $\mathcal{I}_h^c : \mathcal{V}_h \to \mathcal{V}_h \cap H^1(\Omega)$ such that

$$(4.1) \qquad \beta^{-1/2}\|A^{1/2}\nabla_H(v - \mathcal{I}_h^c v)\|_{L^2(T)} + \|h^{-1}(v - \mathcal{I}_h^c v)\|_{L^2(T)} \lesssim \alpha^{-1/2}|v|_{h,T}.$$

We split $v = v^c + v^d \in \mathcal{V}_h$ into a conforming, $v^c = \mathcal{I}_h^c v$, and a nonconforming, $v^d = v - \mathcal{I}_h^c v$, part and obtain

$$(4.2) \qquad \begin{aligned} H^{-1}\|v - \Pi_H v\|_{L^2(T)} &\leq H^{-1}(\|v^c - \Pi_H v^c\|_{L^2(T)} + \|v^d - \Pi_H v^d\|_{L^2(T)}) \\ &\lesssim \|\nabla_h v\| + \|\nabla_h(v - v^c)\|_{L^2(T)} + H^{-1}\|v^d\|_{L^2(T)}) \\ &\lesssim \alpha^{-1/2}|||v|||_{h,T} \end{aligned}$$

using the triangle inequality, stability of the L^2-projection, and (4.1). Furthermore,

$$\begin{aligned} |||\Pi_H v|||_H^2 &= \sum_{T \in \mathcal{T}_H} \|\sqrt{A}\nabla(\Pi_H v - \Pi_0(\mathcal{T}_H)v)\|_{L^2(T)}^2 + \sum_{e \in \mathcal{E}_H(\Omega \cup \Gamma_D)} \frac{\sigma}{H}\|[v_c - \Pi_H v]\|_{L^2(e)}^2 \\ &\lesssim \sum_{T \in \mathcal{T}_H} \beta\left(\frac{1}{H^2}\|v - \Pi_0(\mathcal{T}_H)v\|_{L^2(T)}^2 + \frac{1}{H^2}\|v_c - \Pi_H v\|_{L^2(T)}^2\right) \\ &\lesssim C_{\beta/\alpha}^2 |||v|||_h^2 \end{aligned}$$

using the triangle inequality, (4.1), and (4.2), which concludes the proof. □

The operator Π_H is surjective. The next lemma shows that given some $v_H \in \mathcal{V}_H$ in the image of Π_H there exists a H^1-conforming preimage $v \in \Pi_H^{-1} v_H \subset \mathcal{V}_h$ with comparable support.

LEMMA 6. *For each $v_H \in \mathcal{V}_H$, there exists a $v \in \mathcal{V}_h \cap H^1(\Omega)$ such that $\Pi_H v = v_H$, $|||v|||_h \lesssim C_{\beta/\alpha}|||v_H|||_H$, and $\mathrm{supp}(v) \subseteq \mathrm{supp}(\mathcal{I}_h^c v_H)$. Note that the support of $\mathcal{I}_h^c v_H$ is one layer of coarse element larger than the support of v_H.*

Proof. Using Theorem 2.2 in [23] but on the space $\mathcal{V}_H$ gives for each $v \in \mathcal{V}_H$ that there exists a bounded linear operator $\mathcal{I}_H^c : \mathcal{V}_H \to \mathcal{V}_H \cap H^1(\Omega)$ such that

$$(4.3) \qquad \beta^{-1/2}\|A^{1/2}\nabla_H(v - \mathcal{I}_H^c v)\|_{L^2(T)} + \|H^{-1}(v - \mathcal{I}_H^c v)\|_{L^2(T)} \lesssim \alpha^{-1/2}|v|_{H,T}.$$

We define

$$v := \mathcal{I}_H^c v_H + \sum_{T \in \mathcal{T}_H,\, j=1,\ldots,r} (v_H(x_j) - \mathcal{I}_H^c v_H(x_j))\, \theta_{T,j},$$

where $\theta_{T,j} \in \mathcal{V}_h \cap H_0^1(T)$ are coarse scale bubble functions, supported on each element T, with $\Pi_H \theta_{T,j} = \lambda_{T,j}$ and $|||\theta_{T,j}|||_h^2 \lesssim \beta H^{d-2}$. Observe that $\mathrm{supp}(v) \subseteq \mathrm{supp}(\mathcal{I}_h^c v_H)$. The interpolation property follows from

$$\Pi_H v = \mathcal{I}_H^c v_H + \Pi_H \sum_{T \in \mathcal{T}_H, \, j=1,\ldots,r} (v_H(x_j) - \mathcal{I}_H^c v_H(x_j)) \, \theta_{T,j},$$

$$= \mathcal{I}_H^c v_H + \sum_{T \in \mathcal{T}_H, \, j=1,\ldots,r} (v_H(x_j) - \mathcal{I}_H^c v_H(x_j)) \lambda_{T,j} = v_H.$$

To prove stability, we estimate $|||v|||_h$ as follows:

$$|||v|||_h^2 \le \|A^{1/2} \nabla \mathcal{I}_H^c v_H\|_{L^2(\Omega)}^2 + \sum_{T \in \mathcal{T}_H, \, j=1,\ldots,r} |v_H(x_j) - \mathcal{I}_H^c v_H(x_j)|^2 \, |||\theta_j|||_h^2$$

$$\lesssim \|A^{1/2} \nabla_H \mathcal{I}_H^c v_H\|_{L^2(\Omega)}^2 + \beta \|H^{-1}(v_H - \mathcal{I}_H^c v_H)\|_{L^2(\Omega)}^2$$

$$\lesssim C_{\beta/\alpha}^2 |||v_H|||_H^2,$$

using the inverse estimate $\|v\|_{L^\infty(T)} \le H^{-d/2} \|v\|_{L^2(T)}$ for all $v \in \mathcal{V}_H$ and using the estimate (4.3).   ☐

Remark 7. Note that $\theta_{T,j} \in \mathcal{V}_h \cap H_0^1(T)$ for all $T \in \mathcal{T}_H$ (fulfilling the conditions in Lemma 6) can be constructed using two (or more) refinements of the coarse scale parameter H. We can let $\theta_{T,j} \in \mathcal{V}_{h'} \cap H_0^1(T)$, where $\mathcal{V}_{h'} \subset \mathcal{V}_h$ and $h \le h' \le 2^{-2}H$. This does not put a big restriction on h since the mesh $\mathcal{T}_h$ is assumed to be sufficiently fine to resolve the variation in the coefficient A, while the parameter H does not need to resolve A.

The following lemma says that the corrected basis is stable with respect to the fine scale parameter h in the energy norm (2.4); this is not a trivial result since the basis function $\{\lambda_{T,j} | T \in \mathcal{T}_H, \, j = 1, \ldots, r\}$ is discontinuous.

LEMMA 8 (stability of the corrected basis functions). *For all* $T \in \mathcal{T}_H$, $j = 1, \ldots, r$, *and* $L > 0 \in \mathbb{N}$, *the estimate*

$$|||\lambda_{T,j} - \phi_{T,j}^L|||_h \lesssim C_{\beta/\alpha} |||\lambda_{T,j}|||_H$$

is satisfied, independently of the fine scale parameter h.

Proof. For any $T \in \mathcal{T}_H$, $j = 1, \ldots, r$, by Lemma 6 there exists a b such that $v = \lambda_{T,j} - b \in \mathcal{V}_h^f(\omega_T^L)$, and $|||b|||_h \lesssim C_{\beta/\alpha} |||\lambda_{T,j}|||_H$. We have

$$|||\lambda_{T,j} - \phi_{T,j}^L|||_h^2 \lesssim a_h(\lambda_{T,j} - \phi_{T,j}^L, \lambda_{T,j} - \phi_{T,j}^L) = a_h(\lambda_{T,j} - \phi_{T,j}^L, \lambda_{T,j} - v),$$

$$\lesssim a_h(\lambda_{T,j} - \phi_{T,j}^L, b) \lesssim C_{\beta/\alpha} |||\lambda_{T,j} - \phi_{T,j}^L|||_h |||\lambda_{T,j}|||_H,$$

which concludes the proof.   ☐

4.2. A priori estimates. The following theorem gives an error bound for the idealized dG multiscale method, whereby the correctors for the basis are solved globally.

THEOREM 9. *Let* $u_h \in \mathcal{V}_h$ *solve* (2.5) *and let* $u_H^{ms} \in \mathcal{V}_H^{ms}$ *solve* (3.5); *then the estimate*

$$|||u_h - u_H^{ms}|||_h \le C_1 \alpha^{-1/2} \|H(f - \Pi_H f)\|_{L^2(\Omega)}$$

is satisfied, where C_1 *depends on neither the mesh* (h *or* H) *size nor the diffusion matrix* A.

Proof. Let $e := u_h - u_H^{ms} = u_f \in \mathcal{V}^f$; then

$$|||e|||_h^2 \lesssim a_h(e, e) = (f, e)_{L^2(\Omega)} = (f - \Pi_H f, e - \Pi_H e)_{L^2(\Omega)}$$

$$\le \|H(f - \Pi_H f)\|_{L^2(\Omega)} \|H^{-1}(e - \Pi_H e)\|_{L^2(\Omega)}$$

$$\lesssim \frac{1}{\sqrt{\alpha}} \|H(f - \Pi_H f)\|_{L^2(T)} |||e|||_h$$

using Lemma 3, Lemma 2, the Cauchy–Schwarz inequality, and Lemma 5, respectively. □

DEFINITION 10. *The cutoff functions $\zeta_T^{d,D} \in P_0(\mathcal{T}_h)$ are defined by the conditions*

$$\zeta_T^{d,D}|_{\omega_T^d} = 1,$$

$$\zeta_T^{d,D}|_{\Omega \setminus \omega_T^D} = 0,$$

$$\||[\zeta_T^{d,D}]\||_{L^\infty(\mathcal{E}_h(T))} \lesssim \frac{h_T}{(D-d)H_T} \quad \text{for all } T \in \mathcal{T}_H,$$

and $\zeta_T^{d,D}$ is constant on the boundary $\partial(\omega_T^D \setminus \omega_T^d)$.

The next lemma shows the exponential decay in the corrected basis; this is a key result in the analysis.

LEMMA 11. *For all $T \in \mathcal{T}_H$, $j = 1, \ldots, r$, the estimate*

$$\||(\lambda_{T,j} - \phi_{T,j}) - (\lambda_{T,j} - \phi_{T,j}^L)\||_h = \||\phi_{T,j} - \phi_{T,j}^L\||_h \leq C_3 \gamma^L \||\phi_{T,j} - \lambda_{T,j}\||_h$$

is satisfied with $C_3 = C C_{\beta/\alpha}^3$, $0 < \gamma < 1$ given by $\gamma := (\frac{C_2}{\ell - 1})^{\frac{k-1}{2\ell}}$, $C_2 = C' C_{\beta/\alpha}^2$, and $L = k\ell$, $k, \ell \geq 2 \in \mathbb{N}$, noting that C and C' are positive constants that are independent of the mesh (h or H), of the patch size L, and of the diffusion matrix A.

Proof. Define $e := \phi_{T,j} - \phi_{T,j}^L = \phi_{T,j} - \phi_{T,j}^{\ell k}$. We have

$$(4.4) \qquad \||e\||_h^2 \lesssim a_h(e, \phi_{T,j} - \phi_{T,j}^{\ell k}) = a_h(e, \phi_{T,j} - v) \lesssim \||e\||_h \cdot \||\phi_{T,j} - v\||_h$$

for $v \in \mathcal{V}_h^f(\omega_T^{\ell k})$. Let $\zeta := \zeta_T^{\ell k-1, \ell k}$; then by Lemma 6 there exists a b such that $v = \zeta \phi_{T,j} - b \in \mathcal{V}_h^f(\omega_T^{\ell k})$, $\Pi_H b = \Pi_H \zeta \phi_{T,j}$, $\||b\||_h \lesssim C_{\beta/\alpha} \||\Pi_H \zeta \phi_{T,j}\||_H$, and $\text{supp}(b) \subseteq \text{supp}(\mathcal{I}_h^c \Pi_H \zeta \phi_{T,j})$. We have

$$(4.5) \qquad \||\phi_{T,j} - v\||_h = \||\phi_{T,j} - (\zeta \phi_{T,j} - b)\||_h$$
$$\leq \||\phi_{T,j} - \zeta \phi_{T,j}\||_h + \||b\||_h$$
$$\lesssim \||\phi_{T,j} - \zeta \phi_{T,j}\||_h + C_{\beta/\alpha} \||\Pi_H(\zeta \phi_{T,j} - \phi_{T,j})\||_H$$
$$\lesssim C_{\beta/\alpha}^2 \||\phi_{T,j} - \zeta \phi_{T,j}\||_h.$$

Furthermore, using the properties of ζ we have

$$(4.6) \qquad \|\sqrt{A} \nabla_h (1 - \zeta) \phi_{T,j}\|_{L^2(\Omega)} \leq \|\sqrt{A} \nabla_h \phi_{T,j}\|_{L^2(\Omega \setminus \omega_T^{\ell k-1})}$$

and

(4.7)

$$|(1 - \zeta) \phi_{T,j}|_h^2 = \sum_{e \in \mathcal{E}_h(\Omega \cup \Gamma_D)} \frac{\sigma}{h_e} \|[(1 - \zeta) \phi_{T,j}]\|_{L^2(e)}^2$$

$$\leq \sum_{e \in \mathcal{E}_h(\Omega \cup \Gamma_D)} \frac{\sigma}{h_e} \left(\|\{1 - \zeta\}[\phi_{T,j}]\|_{L^2(e)}^2 + \|\{\phi_{T,j}\}[1 - \zeta]\|_{L^2(e)}^2 \right)$$

$$\leq \sum_{\substack{e \in \mathcal{E}_h(\Omega \cup \Gamma_D): \\ e \cap \Omega \setminus \omega_T^{\ell k-1} \neq 0}} \left(\frac{\sigma}{h_e} \|[\phi_{T,j}]\|_{L^2(e)}^2 + \frac{\sigma h_T^2}{h_e H_T^2} \|\{\phi_{T,j}\}\|_{L^2(e)}^2 \right)$$

$$\leq \sum_{\substack{e \in \mathcal{E}_h(\Omega \cup \Gamma_D): \\ e \cap \Omega \setminus \omega_T^{\ell k-1} \neq 0}} \frac{\sigma}{h_e} \|[\phi_{T,j}]\|_{L^2(e)}^2 + \frac{\sigma}{H_T^2} \|\phi_{T,j} - \Pi_H \phi_{T,j}\|_{L^2(\Omega \setminus \omega_T^{\ell k-1})}^2$$

$$\lesssim C_{\beta/\alpha}^2 \||\phi_{T,j}\||_{h, \Omega \setminus \omega_T^{\ell k-1}}^2$$

using a trace inequality and Lemma 5, respectively. Combining (4.4), (4.5), (4.6), and (4.7) yields

$$(4.8) \qquad |||e|||_h \lesssim C_{\beta/\alpha}^2 |||\phi_{T,j} - \zeta\phi_{T,j}|||_h \lesssim C_{\beta/\alpha}^3 |||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^{\ell k-1}}.$$

To simplify notation, let $m := \ell(k-1) - 1$ and $M := \ell k - 1$. For $\eta_T := 1 - \zeta_T^{m+1,M}$, we obtain

$$(4.9) \qquad |||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^M}^2 \leq |||\eta_T\phi_{T,j}|||_h^2 \lesssim a_h(\eta_T\phi_{T,j}, \eta_T\phi_{T,j}),$$

where

$$(4.10) \quad a_h(\eta_T\phi_{T,j}, \eta_T\phi_{T,j})$$
$$= (A\nabla_h\eta_T\phi_{T,j}, \nabla_h\eta_T\phi_{T,j})_{L^2(\Omega)}$$
$$+ \sum_{e\in\mathcal{E}_h(\Omega\cup\Gamma_D)} \left(-2(\{\nu\cdot A\nabla\eta_T\phi_{T,j}\}, [\eta_T\phi_{T,j}]) + \frac{\sigma}{h_e}([\eta_T\phi_{T,j}], [\eta_T\phi_{T,j}]) \right).$$

For the first term on the right-hand side of (4.10), we have

$$(4.11) \qquad (A\nabla_h\eta_T\phi_{T,j}, \nabla_h\eta_T\phi_{T,j})_{L^2(\Omega)} = (A\nabla_h\phi_{T,j}, \nabla_h\eta_T^2\phi_{T,j})_{L^2(\Omega)}$$

since η_T is constant on each element $T \in \mathcal{T}_h$; for the other terms we use (A.3) and (A.4) (with $v = \eta_T$, $w = \nu\cdot A\nabla\phi_{T,j}$ and $u = \phi_{T,j}$). We can thus arrive at

$$(4.12) \qquad |||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^M}^2 \leq a_h(\eta_T\phi_{T,j}, \eta_T\phi_{T,j}) = a_h(\phi_{T,j}, \eta_T^2\phi_{T,j})$$
$$+ \sum_{e\in\mathcal{E}_h(\Omega)} \big(1/2(\{\nu\cdot A\nabla\phi_{T,j}\}, [\eta_T]^2[\phi_{T,j}])_{L^2(e)}$$
$$- 1/4([\nu\cdot A\nabla\phi_{T,j}], [\eta_T]^2\{\phi_{T,j}\})_{L^2(e)}$$
$$- \frac{\sigma}{4h_e}([\eta_T]^2, [\phi_{T,j}]^2)_{L^2(e)} + \frac{\sigma}{h_e}([\eta_T]^2, \{\phi_{T,j}\}^2)_{L^2(e)}\big)$$

using (4.9), (4.10), and (4.11). Note that

$$(4.13)$$
$$\sum_{e\in\mathcal{E}_h(\Omega)} \big(1/2(\{\nu\cdot A\nabla\phi_{T,j}\}, [\eta_T]^2[\phi_{T,j}])_{L^2(e)} - 1/4([\nu\cdot A\nabla\phi_{T,j}], [\eta_T]^2\{\phi_{T,j}\})_{L^2(e)}$$
$$- \frac{\sigma}{4h_e}([\eta_T]^2, [\phi_{T,j}]^2)_{L^2(e)} + \frac{\sigma}{h_e}([\eta_T]^2, \{\phi_{T,j}\}^2)_{L^2(e)}\big)$$
$$\lesssim \sum_{\substack{e\in\mathcal{E}_h(\Omega): \\ e\cap\omega_T^M\setminus\omega_T^{m+1}\neq 0}} \frac{h_T^2}{\ell^2 H_T^2} \Big(\|\{\nu\cdot A\nabla\phi_{T,j}\}\|_{L^2(e)}\|[\phi_{T,j}]\|_{L^2(e)}$$
$$+ \|[\nu\cdot A\nabla\phi_{T,j}]\|_{L^2(e)}\|\{\phi_{T,j}\}\|_{L^2(e)} + \frac{\sigma}{h_e}\left(\|[\phi_{T,j}]\|_{L^2(e)}^2 + \|\{\phi_{T,j}\}\|_{L^2(e)}^2\right)\Big)$$
$$\lesssim \sum_{\substack{e\in\mathcal{E}_h(\Omega): \\ e\cap\omega_T^M\setminus\omega_T^{m+1}\neq 0}} \left(\frac{h_T}{\ell^2 H_T^2}\|A\nabla\phi_{T,j}\|_{L^2(T^+\cup T^-)}\|\phi_{T,j}\|_{L^2(T^+\cup T^-)} + \frac{\sigma}{\ell^2 H_T^2}\|\phi_{T,j}\|_{L^2(T^+\cup T^-)}^2 \right)$$
$$\lesssim \beta\ell^{-2}\|H_T^{-1}(\phi_{T,j} - \Pi_H\phi_{T,j})\|_{L^2(\omega_T^M\setminus\omega_T^{m+1})}^2 \leq C_{\beta/\alpha}^2 \ell^{-2}|||\phi_{T,j}|||_{h,\omega_T^M\setminus\omega_T^{m+1}}^2.$$

Using that there exist a b such that $\Pi_H b = \Pi_H \eta_T^2 \phi_{T,j}$, $|||b|||_h \lesssim C_{\beta/\alpha} |||\Pi_H \eta_T^2 \phi_{T,j}|||_H$, and $\text{supp}(v) \subseteq \text{supp}(\mathcal{I}_h^c v_H)$ from Lemma 6, we have

$$(4.14) \qquad a_h(\phi_{T,j}, \eta_T^2 \phi_{T,j}) = a_h(\phi_{T,j}, \eta_T^2 \phi_{T,j} - b) + a_h(\phi_{T,j}, b) = a_h(\phi_{T,j}, b)$$
$$\lesssim |||\phi_{T,j}|||_{h,\omega_T^{M+1}\setminus\omega_T^m} |||b|||_{h,\omega_T^{M+1}\setminus\omega_T^m}$$
$$\leq C_{\beta/\alpha} |||\phi_{T,j}|||_{h,\omega_T^{M+1}\setminus\omega_T^m} |||\Pi_H \eta_T^2 \phi_{T,j}|||_{H,\omega_T^M\setminus\omega_T^m}.$$

Furthermore, we have that

$$(4.15)$$
$$|||\Pi_H \eta_T^2 \phi_{T,j}|||_{H,\omega_T^M\setminus\omega_T^m}^2 = |||\Pi_H(\eta_T^2 - \Pi_0(\mathcal{T}_H)\eta_T^2)\phi_{T,j}|||_{H,\omega_T^M\setminus\omega_T^m}^2$$
$$= \|\sqrt{A}\nabla_H \Pi_H(\eta_T^2 - \Pi_0(\mathcal{T}_H)\eta_T^2)\phi_{T,j}\|_{L^2(\omega_T^M\setminus\omega_T^m)}^2$$
$$+ \sum_{\substack{e \in \mathcal{E}_h(\Omega \cup \Gamma_D): \\ e \cap \overline{\omega_T^M \setminus \omega_T^m} \neq 0}} \frac{\sigma}{H_e} \|[\Pi_H(\eta_T^2 - \Pi_0(\mathcal{T}_H)\eta_T^2)\phi_{T,j}]\|_{L^2(e)}^2$$
$$\lesssim \beta \|H_e^{-1}\Pi_H(\eta_T^2 - \Pi_0(\mathcal{T}_H)\eta_T^2)\phi_{T,j}\|_{L^2(\omega_T^M\setminus\omega_T^m)}^2$$
$$\leq \sum_{T \in \mathcal{T}_h(\omega_T^M\setminus\omega_T^m)} \beta \|H_e^{-1}(\eta_T^2 - \Pi_0(\mathcal{T}_H)\eta_T^2)\|_{L^\infty(T)}^2 \|\phi_{T,j}\|_{L^2(T)}^2$$
$$\lesssim \beta \ell^{-2} \|H_e^{-1}(\phi_{T,j} - \Pi_H \phi_{T,j})\|_{L^2(\omega_T^M\setminus\omega_T^m)}^2$$
$$\lesssim C_{\beta/\alpha}^2 \ell^{-2} |||\phi_{T,j}|||_{h,\omega_T^M\setminus\omega_T^m}^2$$

using a trace inequality, an inverse inequality, and Lemma 5, respectively. Combining the inequalities (4.12), (4.13), (4.14), and (4.15) yields

$$|||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^M}^2 \leq \frac{C_2}{\ell-1}|||\phi_{T,j}|||_{h,\omega_T^{M+1}\setminus\omega_T^m}^2 \leq \frac{C_2}{\ell-1}|||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^m}^2,$$

where $C_2 = C'C_{\beta/\alpha}^2$. Substituting back to ℓ and k and using a cutoff function with a slightly different argument yields

$$|||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^{\ell k-1}}^2 \leq \frac{C_2}{\ell-1}|||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^{\ell(k-1)-1}}^2 \leq \left(\frac{C_2}{\ell-1}\right)^2 |||\phi_{T,j}|||_{h,\Omega\setminus\omega_T^{\ell(k-2)-1}}^2$$
$$\leq \cdots \leq \left(\frac{C_2}{\ell-1}\right)^{k-1} |||\phi_{T,j}|||_{h,\omega^\ell\setminus\omega_T^{\ell-1}}^2,$$

which together with (4.8) concludes the proof. $\square$

LEMMA 12. *For all $T \in \mathcal{T}_H$, $j = 1, \ldots, r$, the estimate*

$$\left\|\left| \sum_{T \in \mathcal{T}_H, \, j=1,\ldots,r} v_j(\phi_{T,j} - \phi_{T,j}^L) \right\|\right|_h^2 \leq C_4 L^d \sum_{T \in \mathcal{T}_H, \, j=1,\ldots,r} |v_j|^2 |||\phi_{T,j} - \phi_{T,j}^L|||_h^2$$

is satisfied with $C_4 = CC_{\beta/\alpha}^3$ and C being a positive constant independent of the mesh (h or H), of the patch size L, and of the diffusion matrix A.

Proof. Let $w = \sum_{T \in \mathcal{T}_H, \, j=1,\ldots,r} v_j(\phi_{T,j} - \phi_{T,j}^L)$, and note that

$$(4.16) \qquad \begin{aligned} a_h(\phi_{T,j} - \lambda_{T,j}, w - \zeta_T w + b_T) &= 0, \\ a_h(\phi_{T,j}^L - \lambda_{T,j}, w - \zeta_T w + b_T) &= 0, \end{aligned}$$

where $\zeta_T := \zeta_T^{L+1,L+2}$, using Lemma 6 and the property of the cutoff function. We obtain

(4.17)
$$\left|\left|\left|\sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r} v_j(\phi_{T,j}-\phi_{T,j}^L)\right|\right|\right|_h \lesssim \sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r} v_j a_h(\phi_{T,j}-\phi_{T,j}^L, w)$$
$$= \sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r} v_j a_h(\phi_{T,j}-\phi_{T,j}^L, \zeta_T w - b_T)$$
$$\lesssim \sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r} |v_j|\cdot|||\phi_{T,j}-\phi_{T,j}^L|||_h\,(|||\zeta_T w|||_h + |||b_T|||_h)$$
$$\lesssim \sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r} |v_j|\cdot|||\phi_{T,j}-\phi_{T,j}^L|||_h$$
$$\times\left(|||\zeta_T w|||_h + C_{\beta/\alpha}|||\Pi_H\zeta_T w|||_H\right)$$
$$\lesssim \sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r} |v_j|\cdot|||\phi_{T,j}-\phi_{T,j}^L|||_h C_{\beta/\alpha}^2|||\zeta_T w|||_h.$$

From (4.6) and (4.7), we have

(4.18)
$$|||\zeta_T w|||_h = |||\zeta_T w|||_{h,\omega_T^{L+2}} \lesssim C_{\beta/\alpha}|||w|||_{h,\omega_T^{L+2}}.$$

Then, further estimation of (4.17) can be achieved using (4.18) and the discrete Cauchy–Schwarz inequality:

$$\left|\left|\left|\sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r} v_j(\phi_{T,j}-\phi_{T,j}^L)\right|\right|\right|$$
$$\leq C_{\beta/\alpha}^3\left(\sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r}|v_j|^2|||\phi_{T,j}-\phi_{T,j}^L|||_h^2\right)^{1/2}\left(\sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r}|||w|||_{h,\omega_T^{L+2}}^2\right)^{1/2}$$
$$\leq C_{\beta/\alpha}^3 L^{d/2}\cdot\left(\sum_{T\in\mathcal{T}_H,\,j=1,\ldots,r}|v_j|^2|||\phi_{T,j}-\phi_{T,j}^L|||_h^2\right)^{1/2}\cdot|||w|||_h.$$

Dividing by w on both sides concludes the proof. $\square$

The following theorem gives an error bound for the dG multiscale method.

THEOREM 13. *Let* $u \in H_D^1(\Omega)$ *solve* (1.2) *and let* $u_H^{ms,L} \in \mathcal{V}_H^{ms,L}$ *solve* (3.5). *Then, the estimate*

$$|||u - u_H^{ms,L}|||_h \leq |||u - u_h|||_h + C_1\alpha^{-1/2}||H(f-\Pi_H f)||_{L^2(\Omega)}$$
$$+ C_5\|H^{-1}\|_{L^\infty(\Omega)}L^{d/2}\gamma^L\|f\|_{L^2(\Omega)}$$

is satisfied with $0 < \gamma < 1$, *L from Lemma* 11, C_1 *from Theorem* 9, *and* $C_5 = CC_{\beta/\alpha}^2 C_4^{1/2}C_3$, *where* C_3 *is from Lemma* 11 *and* C_4 *is from Lemma* 12; C *is a positive constant independent of the mesh (h or H), of the patch size L, and of the diffusion matrix A.*

Remark 14. To counteract the factor $\|H^{-1}\|_{L^\infty(\Omega)}$ in the error bound in Theorem 13, we can choose the localization parameter as $L = \lceil C\log(\|H^{-1}\|_{L^\infty(\Omega)})\rceil$. On adaptively refined meshes it is recommended to choose $L = \lceil C\log(H^{-1})\rceil$.

Proof. We define $\tilde{u}_H^{ms,L} := \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} u_{H,T}^{ms}(x_j)\phi_{T,j}^L$. Then, we obtain

$$(4.19) \quad |||u - u_H^{ms,L}|||_h \leq |||u - \tilde{u}_H^{ms,L}|||_h$$

$$\leq |||u - u_h|||_h + |||u_h - u_H^{ms}|||_h + |||u_H^{ms} - \tilde{u}_H^{ms,L}|||_h$$

$$\leq |||u - u_h|||_h + |||u_h - u_H^{ms}|||_h + \left|\left|\left| \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} u_{H,T}^{ms}(x_j)(\phi_{T,j} - \phi_{T,j}^L) \right|\right|\right|_h.$$

Now, estimating the terms in (4.19), we have

$$|||u_h - u_H^{ms}|||_h \leq C_1 \alpha^{-1/2} \|H(f - \Pi_H f)\|_{L^2(\Omega)}$$

using Theorem 9 and

$$(4.20) \quad \left|\left|\left| \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} u_{H,T}^{ms}(x_j)(\phi_{T,j} - \phi_{T,j}^L) \right|\right|\right|_h^2$$

$$\leq C_4 L^d \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} |u_{H,T}^{ms}(x_j)|^2 |||\phi_{T,j} - \phi_{T,j}^L|||_h^2.$$

$$\leq C_4 C_3^2 L^d \gamma^{2L} \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} |u_{H,T}^{ms}(x_j)|^2 |||\phi_{T,j} - \lambda_j|||_h^2$$

using Lemmas 12 and 11, respectively. Further estimation, using Lemma 8, yields

$$(4.21) \quad \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} |u_{H,T}^{ms}(x_j)|^2 |||\phi_{T,j} - \lambda_{T,j}|||_h^2$$

$$\lesssim C_{\beta/\alpha}^2 \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} |u_{H,T}^{ms}(x_j)|^2 |||\lambda_{T,j}|||_H^2$$

$$\lesssim C_{\beta/\alpha}^2 \beta \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} |u_{H,T}^{ms}(x_j)|^2 H_T^{-2} \|\lambda_{T,j}\|_{L^2(T)}^2$$

$$= C_{\beta/\alpha}^2 \beta \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} \|H_T^{-1} u_{H,T}^{ms}(x_j)\lambda_{T,j}\|_{L^2(T)}^2$$

$$\lesssim C_{\beta/\alpha}^2 \beta \left\| \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} H_T^{-1} u_{H,T}^{ms}(x_j)\lambda_{T,j} \right\|_{L^2(\Omega)}^2.$$

Furthermore, using a Poincaré–Friedrichs inequality for piecewise H^1 functions, we deduce

$$(4.22) \quad \left\| \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} H_T^{-1} u_{H,T}^{ms}(x_j)\lambda_{T,j} \right\|_{L^2(\Omega)}^2$$

$$\lesssim \left\| \sum_{T \in \mathcal{T}_H,\, j=1,\dots,r} H_T^{-1} u_{H,T}^{ms}(x_j)\Pi_H(\lambda_{T,j} - \phi_{T,j}) \right\|_{L^2(\Omega)}^2$$

$$\lesssim \alpha^{-1} |||H^{-1} u_H^{ms}|||_h^2$$

$$\lesssim \alpha^{-1} \|H^{-1}\|_{L^\infty(\Omega)}^2 \|f\|_{L^2(\Omega)}^2.$$

Combining (4.20), (4.21), and (4.22), we arrive at

$$|||u_H^{ms} - u_H^{ms,L}|||_h \lesssim C_{\beta/\alpha}^2 C_4^{1/2} C_3 \|H^{-1}\|_{L^\infty(\Omega)} L^{d/2} \gamma^L \|f\|_{L^2(\Omega)}. \qquad \square$$

4.3. Error in a quantity of interest. In engineering applications, we are often interested in a quantity of interest, usually a functional $G(v)$ of the solution. To this end, consider the dual reference solution (2.5): find $\phi_h \in \mathcal{V}_h$ such that

$$(4.23) \qquad a_h(v, \phi_h) = G(v) \quad \text{for all } v \in \mathcal{V}_h;$$

and consider the dual multiscale solution (3.5): find $\phi_H^{ms,L} \in \mathcal{V}_H^{ms,L}$ such that

$$(4.24) \qquad a_h(v, \phi_H^{ms,L}) = G(v) \quad \text{for all } v \in \mathcal{V}_H^{ms,L}.$$

THEOREM 15. *Let $u \in H_D^1(\Omega)$ solve (1.2), let $u_H^{ms,L} \in \mathcal{V}_H^{ms,L}$ solve (3.5), and let $G(v)$ be the quantity of interest. Then, the estimate*

$$|G(u) - G(u_H^{ms,L})| \lesssim |G(u) - G(u_h)| + |||u_h - u_H^{ms,L}|||_h |||\phi_h - \phi_H^{ms,L}|||_h$$

is satisfied.

Proof. From (4.23) and (4.24), we obtain the Galerkin orthogonality

$$(4.25) \qquad a_h(v, \phi_h - \phi_H^{ms,L}) = 0 \quad \text{for all } v \in \mathcal{V}_H^{ms,L}.$$

Using the triangle inequality, we have

$$|G(u) - G(u_H^{ms,L})| \leq |G(u) - G(u_h)| + |G(u_h) - G(u_H^{ms,L})|.$$

Finally, observing that

$$
\begin{aligned}
|G(u_h) - G(u_H^{ms,L})| &= |a_h(u_h - u_H^{ms,L}, \phi_h)| \\
&= |a_h(u_h - u_H^{ms,L}, \phi_h - \phi_H^{ms,L})| \\
&\lesssim |||u_h - u_H^{ms,L}|||_h |||\phi_h - \phi_H^{ms,L}|||_h,
\end{aligned}
$$

using (4.25), concludes the proof. ∎

COROLLARY 16. *For $G(v) = (u_h - u_H^{ms,L}, v)_{L^2(\Omega)}$, the following L^2-norm error estimates hold:*

$$\|u - u_H^{ms,L}\|_{L^2(\Omega)} \lesssim \|u - u_h\|_{L^2(\Omega)} + |||u_h - u_H^{ms,L}|||_h^{1/2} |||\phi_h - \phi_H^{ms,L}|||_h^{1/2}$$

and

$$(4.26) \qquad \|u - u_H^{ms,L}\|_{L^2(\Omega)} \lesssim \|u - u_h\|_{L^2(\Omega)} + H|||u_h - u_H^{ms,L}|||_h$$

for $L = \lceil C \log(H^{-1}) \rceil$ with C a sufficiently large positive constant independent of the mesh parameters.

Remark 17. As expected, if we are interested in a bounded linear functional with additional smoothness, a higher convergence rate is obtained for $|G(u_h) - G(u_H^{ms,L})|$. For example, given the forcing function for the primal problem $f \in H^m(\mathcal{T}_H)$, a quantity of interest $G(v) = (g, v)_{L^2(\Omega)}$, where $g \in H^n(\mathcal{T}_H)$ (with $H^0(\mathcal{T}_H)$ denoting the $L^2(\Omega)$ space), and choosing $L = \lceil C \log(H^{-1}) \rceil$ with large enough C gives

$$
\begin{aligned}
|G(u) - G(u_H^{ms,L})| &\lesssim |G(u) - G(u_h)| \\
&+ H^{2+m+n} \left(\sum_{T \in \mathcal{T}_H} |f|_{H^m(T)}^2 \right)^{1/2} \left(\sum_{T \in \mathcal{T}_H} |g|_{H^n(T)}^2 \right)^{1/2}
\end{aligned}
$$

for $2 \geq m, n \in \mathbb{N}$.

5. Numerical experiments. Let Ω be an L-shaped domain (constructed by removing the lower right quadrant in the unit square) and let the forcing function be $f = 1 + \cos(2\pi x)\cos(2\pi y)$ for $(x,y) \in \Omega$. The boundary Γ is divided into the Neumann boundary $\Gamma_N := \Gamma \cap (\{(x,y) : y = 0\} \cup \{(x,y) : x = 1\})$ and the Dirichlet boundary $\Gamma_D = \Gamma \setminus \Gamma_N$. We shall consider three different permeabilities: constant $A_1 = 1$, $A_2 = A_2(x)$, which is piecewise constant with periodic values of 1 and 0.01 with respect to a Cartesian grid of width 2^{-6} in the x-direction, and $A_3 = A_3(x,y)$, which is piecewise constant with respect to a Cartesian grid of width 2^{-6} in both the x- and y-directions and has a maximum ratio $\beta/\alpha = 4 \cdot 10^6$. The data for A_3 are taken from layer 64 in the SPE benchmark problem (see http://www.spe.org/web/csp). The permeabilities A_2 and A_3 are illustrated in Figure 5.1. For the periodic problem many of the corrected basis functions will be identical. For instance, all the local corrected bases in the interior are solved on identical patches, reducing the computational effort considerably. In the extreme case of a problem with periodic coefficients on a unit hypercube, with period boundary conditions, the correctors $\phi_{T,j}$, $j = 1, \ldots, r$, will be identical for all $T \in \mathcal{T}_H$.

Consider the uniform (coarse) quadrilateral mesh $\mathcal{T}_H$ with size $H = 2^{-i}$, $i = 1, \ldots, 6$. The convergence rate $-p/2$ corresponds to $\mathcal{O}(H^p)$ since the number of degrees of freedom $\approx H^{-2}$. The corrector functions (3.3) are solved on a subgrid of a (fine) quadrilateral mesh $\mathcal{T}_h$ with mesh size 2^{-8}. The mesh $\mathcal{T}_h$ will also act as a reference grid on which we shall compute a reference solution $u_h \in \mathcal{V}_h$ (2.5). Note that the mesh $\mathcal{T}_h$ is chosen so that it resolves the fine scale features of A_i, $i = 1, 2, 3$; hence we assume that the solution u_h is sufficiently accurate.

5.1. Localization parameter. If $f \in H^m(\mathcal{T}_H)$ we have the bound

$$(5.1) \qquad \|H(f - \Pi_H f))\|_{L^2(\Omega)} \lesssim \left(\sum_{T \in \mathcal{T}_H} H^{2k+2} |f|_{H^k(T)}^2 \right)^{1/2},$$

where $k = 0$ for $m = 0$, $k = 1$ for $m = 1$, and $k = 2$ for $m > 1$. Hence, to balance the error between the terms on the right-hand side of the estimate in Theorem 13, a different constant C has to be used for the localization parameter, $L = \lceil C \log(H^{-1}) \rceil$, depending on the elementwise regularity of the forcing function f on $\mathcal{T}_H$. Figure 5.2 shows the relative error in the energy norm $|||u_h - u_H^{ms,L}|||_h / |||u_h|||_h$ and Figure 5.3

(a) $\beta/\alpha = 10^2$ (b) $\beta/\alpha \approx 4 \cdot 10^6$
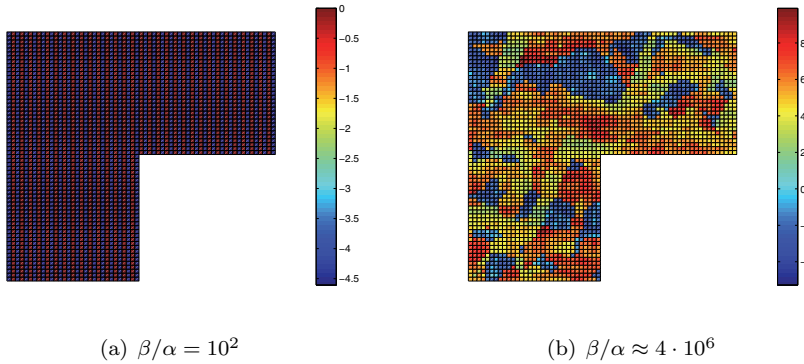
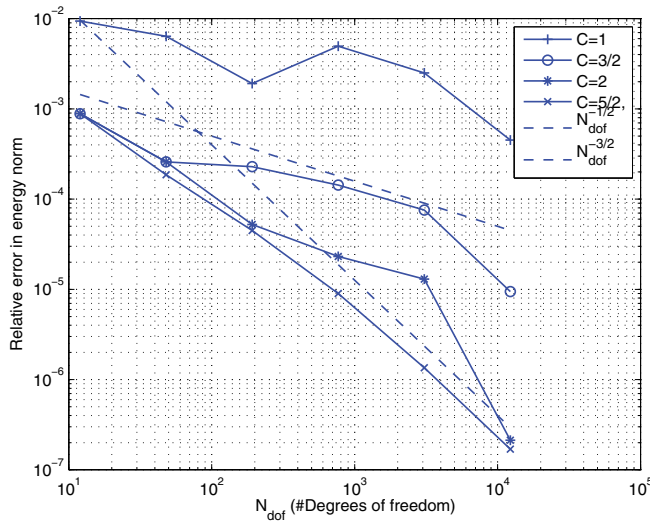FIG. 5.1. *The permeability structure of* (a) A_2 *and* (b) A_3 *in log scale.*

FIG. 5.2. *Diffusion coefficient $A_1 = 1$. Relative energy-norm error against N_{dof} for different values of C for the localization parameter L.*
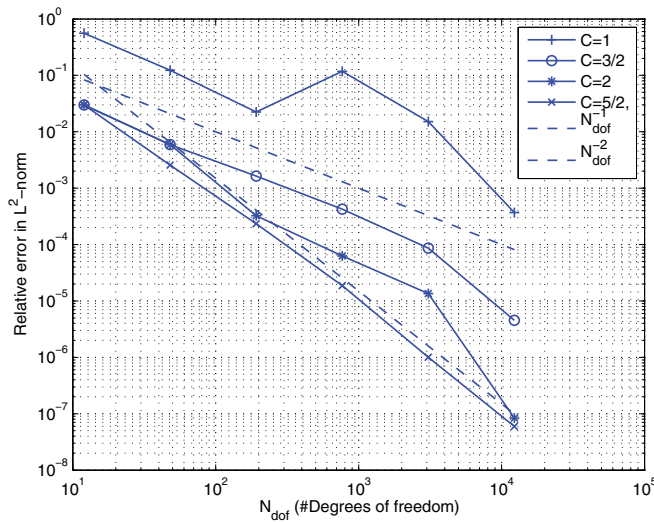

FIG. 5.3. *Diffusion coefficient $A_1 = 1$. Relative L^2-norm error against N_{dof} for different values of C for the localization parameter L.*

the relative error in the L^2-norm $\|u_h - u_H^{ms,L}\|_{L^2(\Omega)}/\|u_h\|_{L^2(\Omega)}$ between u_h and $u_H^{ms,L}$ against the number of degrees of freedom $N_{\text{dof}} \approx O(H^{-2})$, using different constants $C = 1, 3/2, 2, 5/2$. With the choice $C = 5/2$ the errors due to the localization can be neglected compared to the errors from the forcing function for both the energy and the L^2-norm. For $f \notin H^1(\mathcal{T}_h)$, $C = 3/2$ is sufficient since (5.1) gives linear convergence. In the remaining numerical experiments we use $C = 2$, since this value seems to balance the error sufficiently. Note that the numerical overhead increases with C as
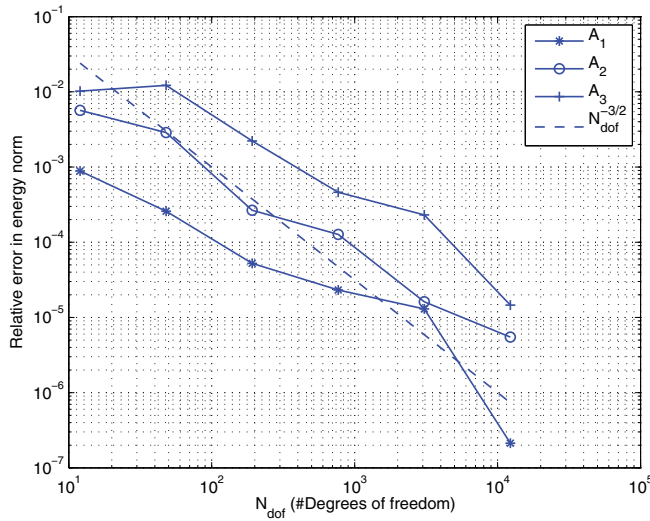
FIG. 5.4. *Relative energy-norm error against N_{dof} for $C = 2$ in the localization parameter L for the the diffusion coefficients A_1, A_2, and A_3.*
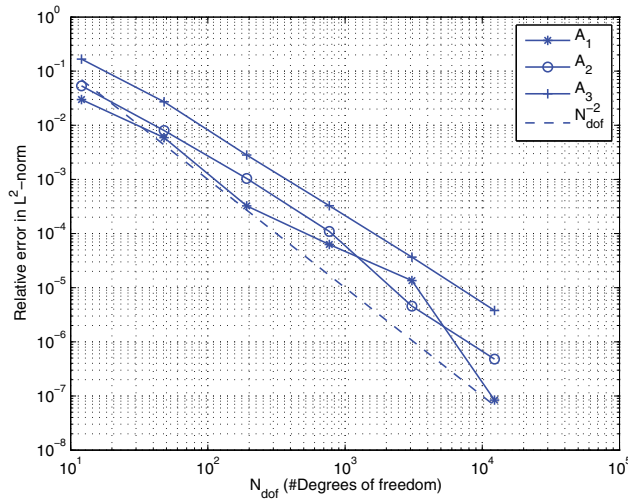

FIG. 5.5. *Relative L^2-norm error against N_{dof} for $C = 2$ in the localization parameter L for the diffusion coefficients A_1, A_2, and A_3.*

the sizes of the patches ω_T^L, $T \in \mathcal{T}_H$, increase with $L = \lceil C \log(H^{-1}) \rceil$. This results in both increased computational cost to compute the corrector functions and reduced sparseness in the coarse scale stiffness matrix.

5.2. Energy-norm convergence. Let the localization parameter be given by $L = \lceil 2 \log(H^{-1}) \rceil$. Figure 5.4 shows the relative error in the energy norm plotted against the number of degrees of freedom. The different permeabilities A_i, $i = 1, 2, 3$, and the singularity arising from the L-shaped domain do not appear to have a
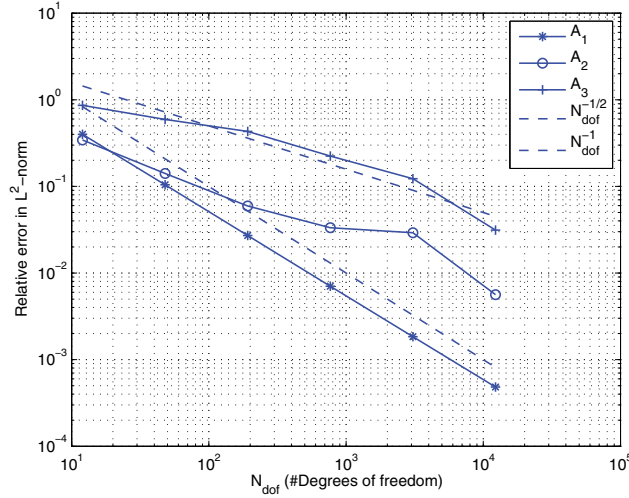
FIG. 5.6. *Relative L^2-norm error against N_{dof} for $C = 2$ in the localization parameter L for the diffusion coefficients A_1, A_2, and A_3.*

substantial impact on the convergence rate, which is about $-3/2$, as expected. We note in passing that using standard dG on the coarse mesh only admits poor convergence behavior for A_2 and for A_3. This is to be expected, since standard dG on the coarse mesh does not resolve the fine scale features.

5.3. L^2-norm convergence. Again, set $L = \lceil 2 \log(H^{-1}) \rceil$. Figures 5.5 and 5.6 show the relative L^2-norm error against the number of degrees of freedom between u_h and $u_H^{ms,L}$ and between u_h and $\Pi_H u_H^{ms,L}$, viz., $\|u_h - \Pi_H u_H^{ms,L}\|_{L^2(\Omega)}/\|u_h\|_{L^2(\Omega)}$, respectively. In Figure 5.5 we see that the L^2-norm error between u_h and $u_H^{ms,L}$ converges at a faster rate than in the energy norm (convergence rate -2 compared to $-3/2$, respectively,) as expected from (4.26). In Figure 5.6 only the coarse part of $u_H^{ms,L}$ is used (i.e., $\Pi_H u_H^{ms,L}$); nevertheless it appears to have a faster convergence rate than $-1/2$, except for the case of the permeability A_3.

6. Concluding remarks. We present a dG multiscale method for second order elliptic problem with heterogeneous and highly varying diffusion coefficients in $L^\infty(\Omega, \mathbb{R}^{d \times d}_{sym})$ with uniform spectral bounds. For $f \in L^2(\Omega)$, the method seeks the solution $u_H^{ms,L}$ (from (3.5)) in a space of corrected basis function (3.4) calculated on patches of size $\mathcal{O}(H \log(H^{-1}))$ (i.e., $L = \lceil C \log(H^{-1}) \rceil$). We have shown the error bounds

$$|||u - u_H^{ms,L}|||_h \leq |||u - u_h|||_h + C_{\alpha/\beta,f} H$$

in the energy norm (Theorem 13) and

$$\|u - u_H^{ms,L}\|_{L^2(\Omega)} \leq \|u - u_h\|_{L^2(\Omega)} + \tilde{C}_{\alpha/\beta,f} H^2$$

in the $L^2(\Omega)$-norm (Corollary 16), where u_h (from (2.5)) is the reference solution on the fine scale and the constants $C_{\alpha/\beta,f}$ and $\tilde{C}_{\alpha/\beta,f}$ depend on the forcing function f and the global bound of the diffusion matrix, but not on its variations. Numerical

experiments show that choosing the localization parameter as $L = \lceil 2\log(H^{-1}) \rceil$ is sufficient to achieve good convergence for the diffusion coefficients (A_1, A_2, and A_3).

Appendix A. Equalities for averages and jump operators. We derive equalities for averages and jump operators across interfaces where the functions v and w have discontinuities. Using $[vw] = \{v\}[w] + [v]\{w\}$ and $\{v\}\{w\} = \{vw\} - 1/4[v][w]$, we have

$$(A.1) \qquad \{vw\}[vu] = \{w\}\{v\}[vu] + 1/4[v][w][vu]$$
$$= \{w\}[v^2u] - \{w\}[v]\{vu\} + 1/4[v][w][vu]$$
$$= \{w\}[v^2u] - [v]\{w\}\{v\}\{u\} - 1/4[v]^2\{w\}[u]$$
$$+ 1/4[v]^2[w]\{u\} + 1/4[v]\{v\}[w][u]$$

and

$$(A.2) \qquad \{vw\}[vu] = \{v\}\{vw\}[u] + \{vw\}[v]\{u\}$$
$$= \{v^2w\}[u] - 1/4[v][vw][u] + \{vw\}[v]\{u\}$$
$$= \{v^2w\}[u] - 1/4[v]^2\{w\}[u] - 1/4[v]\{v\}[w][u]$$
$$+ [v]\{v\}\{w\}\{u\} + 1/4[v]^2[w]\{u\}.$$

Combining (A.1) and (A.2) we obtain

$$(A.3) \qquad 2\{vw\}[vu] = \{w\}[v^2u] + \{v^2w\}[u] + 1/2[v]^2[w]\{u\} - 1/2[v]^2\{w\}[u].$$

Also,

$$(A.4) \qquad [vu][vu] = [u]\{v\}[vu] + [v]\{u\}[vu]$$
$$= [u][v^2u] - [v][u]\{vu\} + [v]\{u\}[vu]$$
$$= [u][v^2u]) - [v][u]\{v\}\{u\} - 1/4[v][u][v][u]$$
$$+ [v]\{u\}[v]\{u\} + [v]\{u\}\{v\}[u]$$
$$= [u][v^2u] - 1/4[v]^2[u]^2 + [v]^2\{u\}^2.$$

REFERENCES

[1] T. Arbogast, G. Pencheva, M. F. Wheeler, and I. Yotov, *A multiscale mortar mixed finite element method*, Multiscale Model. Simul., 6 (2007), pp. 319–346.

[2] T. Arbogast and H. Xiao, *A multiscale mortar mixed space based on homogenization for heterogeneous elliptic problems*, SIAM J. Numer. Anal., 51 (2013), pp. 377–399.

[3] D. N. Arnold, *An interior penalty finite element method with discontinuous elements*, SIAM J. Numer. Anal., 19 (1982), pp. 742–760.

[4] I. Babuška, G. Caloz, and J. E. Osborn, *Special finite element methods for a class of second order elliptic problems with rough coefficients*, SIAM J. Numer. Anal., 31 (1994), pp. 945–981.

[5] I. Babuška and R. Lipton, *Optimal local approximation spaces for generalized finite element methods with application to multiscale problems*, Multiscale Model. Simul., 9 (2011), pp. 373–406.

[6] I. Babuška and J. E. Osborn, *Generalized finite element methods: Their performance and their relation to mixed methods*, SIAM J. Numer. Anal., 20 (1983), pp. 510–536.

[7] L. Berlyand and H. Owhadi, *Flux norm approach to finite dimensional homogenization approximations with non-separated scales and high contrast*, Arch. Ration. Mech. Anal., 198 (2010), pp. 677–721.

[8]   F. Brezzi, L. Franca, T. Hughes, and A. Russo, $b = \int g$, Comput. Methods Appl. Mech.
       Engrg., 145 (1997), pp. 329–339.
[9]   J. Douglas and T. Dupont, *Interior penalty procedures for elliptic and parabolic Galerkin
       methods*, in Computing Methods in Applied Sciences, Lecture Notes in Phys., 58, Springer,
       Berlin, 1976, pp. 207–216.
[10] M. Dryja, *On discontinuous Galerkin methods for elliptic problems with discontinuous coeffi-
       cients*, Comput. Methods Appl. Math., 3 (2003), pp. 76–85.
[11] Y. Efendiev, J. Galvis, and T. Y. Hou, *Generalized Multiscale Finite Element Methods
       (GMsFEM)*, J. Comput. Phys., 251 (2013), pp. 116–135.
[12] Y. Efendiev, J. Galvis, R. Lazarov, M. Moon, and M. Sarkis, *Generalized Multiscale
       Finite Element Method. Symmetric Interior Penalty Coupling*, J. Comput. Phys., 255
       (2013), pp. 1–15.
[13] D. Elfverson, E. Georgoulis, and A. Målqvist, *An adaptive discontinuous Galerkin mul-
       tiscale method for elliptic problems*, Multiscale Model. Simul., 11 (2013), pp. 747–765.
[14] D. Elfverson and A. Målqvist, *Discontinuous Galerkin Multiscale Methods for Convec-
       tion Dominated Problems*, Tech. report 2013-011, Department of Information Technology,
       Uppsala University, Sweden, 2013.
[15] P. Henning, A. Målqvist, and D. Peterseim, *Two-Level Discretization Techniques for
       Ground State Computations of Bose-Einstein Condensates*, arXiv:1305.4080, 2013.
[16] P. Henning, A. Målqvist, and D. Peterseim, *A localized orthogonal decomposition method
       for semi-linear elliptic problems*, arXiv:1211.3551.
[17] P. Henning and D. Peterseim, *Oversampling for the Multiscale Finite Element Method*,
       Multiscale Model. Simul., 11(2013), pp. 1149–1175.
[18] T. Y. Hou, T. Y. Hou, and X. Wu, *A multiscale finite element method for elliptic problems
       in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
[19] T. Y. Hou and X.-H. Wu, *A multiscale finite element method for elliptic problems in composite
       materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
[20] P. Houston, D. Schötzau, and T. P. Wihler, *Energy norm a posteriori error estimation of
       hp-adaptive discontinuous Galerkin methods for elliptic problems*, Math. Models Methods
       Appl. Sci., 17 (2007), pp. 33–62.
[21] T. Hughes, *Multiscale phenomena: Green's functions, the Dirichlet-to-Neumann formulation,
       subgrid scale models, bubbles and the origins of stabilized methods*, Comput. Methods Appl.
       Mech. Engrg., 127 (1995), pp. 387–401.
[22] T. Hughes, G. Feijóo, L. Mazzei, and J.-B. Quincy, *The variational multiscale method: A
       paradigm for computational mechanics*, Comput. Methods Appl. Mech. Engrg., 166 (1998),
       pp. 3–24.
[23] O. Karakashian and F. Pascal, *A posteriori error estimates for a discontinuous Galerkin
       approximation of second-order elliptic problems*, SIAM J. Numer. Anal., 41 (2003), pp.
       2374–2399.
[24] M. G. Larson and A. Målqvist, *Adaptive variational multiscale methods based on a posteriori
       error estimation: Energy norm estimates for elliptic problems*, Comput. Methods Appl.
       Mech. Engrg., 196 (2007), pp. 2313–2324.
[25] A. Målqvist and D. Peterseim, *Localization of elliptic multiscale problems*, Math. Comp.,
       arXiv:1110.0692.
[26] A. Målqvist and D. Peterseim, *Computation of Eigenvalues by Numerical Upscaling*,
       arXiv:1212.0090, 2012.
[27] A. Målqvist, *Multiscale methods for elliptic problems*, Multiscale Model. Simul., 9 (2011),
       pp. 1064–1086.
[28] H. Owhadi and L. Zhang, *Localized bases for finite-dimensional homogenization approxi-
       mations with nonseparated scales and high contrast*, Multiscale Model. Simul., 9 (2011),
       pp. 1373–1398.
[29] R. Scheichl, P. S. Vassilevski, and L. T. Zikatanov, *Multilevel methods for elliptic problems
       with highly varying coefficients on nonaligned coarse grids*, SIAM J. Numer. Anal., 50
       (2012), pp. 1675–1694.