FINITE ELEMENTS FOR ELLIPTIC PROBLEMS WITH HIGHLY VARYING, NONPERIODIC DIFFUSION MATRIX*

D. PETERSEIM† AND S. SAUTER‡

Abstract. This paper considers the numerical solution of elliptic boundary value problems with a complicated (nonperiodic) diffusion matrix which is smooth but highly oscillating on very different scales. We study the influence of the scales and amplitudes of these oscillations to the regularity of the solution. We introduce weighted Sobolev norms of integer order, where the $(p+1)$st seminorm is weighted by properly scaled pth derivative of the diffusion coefficient. The constants in the regularity estimates then turn out to depend only on global lower and upper bounds of the diffusion matrix but not on its derivatives; in particular, the constants are independent of the scales of the oscillations. These regularity results give rise to error estimates for hp-finite element discretizations with scale-adapted distribution of the mesh cells. The adaptation of the mesh is explicit with respect to the local variations of the diffusion coefficient. Numerical results illustrate the efficiency of these oscillation adapted finite elements, in particular, in the preasymptotic regime.

Key words. weighted regularity, elliptic problem, oscillatory diffusion, hp finite elements

AMS subject classifications. 65N30, 35B65, 35J57

DOI. 10.1137/10081839X

1. Introduction. The numerical solution of elliptic boundary value problems on bounded domains $\Omega \subset \mathbb{R}^d$ by the Galerkin finite element method consists of the construction of an "appropriate" finite element mesh and the choice of the (local) polynomial degrees of approximation. An optimal construction should be adapted to the local behavior of the exact solution and, hence, should take into account the following:

(a) local singularities of the solution (e.g., singularities at reentrant corners or at nonsmooth interfaces),

(b) effects of possibly singular perturbations in the equation (e.g., indefiniteness, boundary layers, etc.), and

(c) scales and amplitudes of oscillatory coefficients.

In this paper, we are interested in case (c), i.e., coefficients in the elliptic PDE which are smooth but oscillatory on possibly very small scales. Emphasis is on the case where the oscillatory behavior of the coefficient varies in space in a nonuniform (especially nonperiodic) way. We generalize the regularity theory for elliptic PDEs with a smooth, periodic diffusion matrix of the form $A^{\mathrm{per}}\left(\frac{x}{\varepsilon}\right)$ which oscillates on a small scale ε (see, e.g., [27], [23], where the theory is based on Fourier transform with special kernel functions) to the case of nonseparated, nonperiodic scales.

In [23], the smooth, periodic case was considered and the error estimate for an hp-finite element discretization

$$\|u - u_h\|_{H^1(\Omega)} \leq C_f \min\left\{1, \left(\frac{h}{\varepsilon}\right)^p\right\} \tag{1.1}$$

*Received by the editors December 16, 2010; accepted for publication (in revised form) March 19, 2012; published electronically July 5, 2012.

http://www.siam.org/journals/mms/10-3/81839.html

†Institut für Mathematik, Humboldt-Universitaet zu Berlin, D-10099 Berlin, Germany (peterseim@math.hu-berlin.de). The work of this author was supported by the DFG Research Center Matheon Berlin through project C33.

‡Institut für Mathematik, Universität Zürich, CH-8057 Zürich, Switzerland (stas@math.uzh.ch).

was presented in [23, p. 539]. If we consider the mesh width h of a finite element mesh as the *discretization scale*, the error estimate (1.1) shows that the coefficient scale ε and the discretization scale h are coupled. In particular, the choice $h \approx \varepsilon$ yields an energy error of order $O(1)$ only. If, e.g., $P1$ finite elements are employed, i.e., $p = 1$, then the conventional finite element method requires $O\left(\varepsilon^{-2d}\right)$ degrees of freedom in order to get an accuracy of order $O(\varepsilon)$ with respect to the energy norm.

This paper presents a new regularity theory for our nonperiodic setting. This theory motivates an algorithm for adapting the local mesh width to the scales of the diffusion coefficient in such a way that only $O\left(\varepsilon^{-d}\log^d \frac{1}{\varepsilon}\right)$ degrees of freedom are sufficient to obtain an accuracy in the energy norm of order $O(\varepsilon)$. The theory is based on the local regularity results derived in [25, Chap. 5]—the main difference is that we use this local regularity to derive a weight function for the definition of weighted Sobolev norms of integer order so that the constants in the regularity estimates become independent of the local variations of the coefficients. More precisely, the seminorms of order $p+1$ are weighted by the pth derivatives of the diffusion matrix in a properly scaled way.

We emphasize that our focus in this paper is *not* on the regularity theory of problems with discontinuous coefficients and/or polygonal/polyhedral domains. In those cases it is well known that the solution exhibits singularities, and the optimal a priori and a posteriori mesh grading is well understood. Early publications on local singularities for elliptic problems are [19], [13]. We refer the reader to, e.g., [5] and the references therein for an overview of the regularity for polygonal/polyhedral domains in d dimensions and smooth interfaces. The regularity in the presence of nonsmooth interfaces is studied in [18], [21], [30], whereas optimally graded finite element meshes are described, e.g., in [2] for singular boundary points and in [16], [14] for interfaces. However, if the coefficient does not jump across a *sharp* interface but, although being smooth, rapidly changes its values within some interface "zone," the regularity of the solution is polluted in a neighborhood of this zone and the weighted regularity estimates for sharp interfaces are not applicable, whereas the estimates of this paper apply.

Nowadays, a posteriori error estimation is commonly used for the control of adaptive mesh refinement or, in general, of the adaptive enrichment of finite element spaces. However, for many singularly perturbed or parameter-dependent problems such as, e.g., convection dominated problems, highly indefinite scattering problems, high-frequency eigenvalue problems, etc., the condition "the mesh width has to be *sufficiently* small" typically arises (see, e.g., [1], [10], [6], [32], [24], [26]). For singularly perturbed problems or high-frequency scattering problems, this condition is often so restrictive that the initial mesh must be chosen very fine and further refinement exceeds computer capacity. Thus, the generation of optimal initial meshes is of utmost importance and our goal is to present a new concept for this purpose.

In [12], diffusion problems with even more general L^∞ coefficients are considered. It is proved that in such cases there also exists a (local) *generalized* finite element basis with the following property: For any shape regular finite element mesh of width h there exist $O\left(\left(\log \frac{1}{h}\right)^{d+1}\right)$ *local* basis functions per nodal point such that the corresponding Galerkin solution u_h satisfies the error estimate

$$\|u - u_h\|_{H^1(\Omega)} \leq C_f h,$$

where C_f depends on the right-hand side f and the global bounds of the diffusion coefficients (cf. (2.1)) but not on its variations. Related approaches in the literature

are [31], [3], [22]. On one hand, this result is more general than ours because no smoothness assumption is imposed on the diffusion matrix and no restriction on the discretization parameter h appears. On the other hand, the definition of the basis functions in [12] and the other references [31, 3, 22] is semidiscrete and/or requires a preprocessing step—in some cases the numerical solution of the PDE with very small mesh width—whereas our construction, which utilizes standard hp-finite element spaces (see section 5), is fully explicit. Moreover, our method may serve as an efficient solver for the fine scale computations which are a necessary component of any reasonable numerical upscaling procedure.

The paper is organized as follows. In section 2 we formulate the model problem and the conditions on the coefficients. In section 3 we introduce an oscillation adapted partition of Ω which motivates the definition of weighted Sobolev norm via scaled derivatives of the diffusion matrix on these subdomains. Section 4 proves the new regularity estimates. The tools which are developed, although being technical, are based on the classical regularity estimates in [28] and at a similar technical level. Some facts on the growth of derivatives of the composition of smooth functions are needed; their proofs are postponed to Appendix A. In section 5 we use the regularity results to prove convergence estimates for hp finite elements where the mesh is adapted to the scales of the diffusion matrix via the algorithm developed in section 3. In section 6 we present the results of numerical experiments which show that in the preasymptotic regime many fewer degrees of freedom are necessary to achieve a prescribed accuracy by the oscillation adapted finite elements compared to standard finite elements. Finally, section 7 contains some conclusions.

2. Setting. Let $\Omega \subset \mathbb{R}^d$ be a Lipschitz domain, and let the diffusion matrix $A \in L^\infty\left(\Omega, \mathbb{R}^{d\times d}_{\text{sym}}\right)$ be uniformly elliptic:

$$
\begin{aligned}
0 < \alpha(A,\Omega) &:= \operatorname*{ess\,inf}_{x\in\Omega} \inf_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\langle A(x)v, v\rangle}{\langle v, v\rangle}, \\
\infty > \beta(A,\Omega) &:= \operatorname*{ess\,sup}_{x\in\Omega} \sup_{v\in\mathbb{R}^d\setminus\{0\}} \frac{\langle A(x)v, v\rangle}{\langle v, v\rangle}.
\end{aligned}
\tag{2.1}
$$

For $m \in \mathbb{N}_0 := \mathbb{N} \cup \{0\}$, let $H^m(\Omega)$ denote the usual Sobolev spaces with norm $\|\cdot\|_{H^m(\Omega)}$, and let $H^m_0(\Omega)$ be the closure of $C^\infty_0(\Omega)$ with respect to the norm $\|\cdot\|_{H^m(\Omega)}$. The dual space of $H^m_0(\Omega)$ is denoted by $H^{-m}(\Omega)$.

Given $f \in L^2(\Omega)$, we are seeking $u \in H^1_0(\Omega)$ such that

$$
a(u,v) := \int_\Omega \langle A\nabla u, \nabla v\rangle \,\mathrm{d}x = \int_\Omega fv \,\mathrm{d}x =: F(v) \quad \forall v \in H^1_0(\Omega). \tag{2.2}
$$

The elliptic regularity theory (see, e.g., [11]) shows that for smooth data (domain Ω, diffusion coefficient A) the condition $f \in H^m(\Omega)$ implies $u \in H^{m+2}(\Omega)$ and there is some constant C (depending on the data and m) such that

$$
\|u\|_{H^{m+2}(\Omega)} \le C \|f\|_{H^m(\Omega)}.
$$

If the domain Ω is polygonal and/or A is only piecewise smooth and discontinuous along polygonal interfaces, high order regularity can be preserved in weighted Sobolev spaces [19], [13], [30], [16], [5].

In this paper, we study the effect of a smooth but oscillatory diffusion coefficient and introduce new types of weighted Sobolev spaces where the regularity constants

are independent of such oscillations. These regularity estimates are then interlinked with the Galerkin discretization of (2.2) by *hp* finite elements because it allows one to balance the local estimates of the interpolation error on the single finite element simplices. The weight function used in the definition of the oscillation adapted Sobolev norms encodes the strength of the oscillations of the diffusion coefficients on the scale of the finite element mesh.

3. Oscillation adapted Sobolev norms. We assume that A, besides (2.1), satisfies $A \in C^p(\overline{\Omega}, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ for some smoothness parameter $p \in \mathbb{N}$. In the subsequent definition we quantify the smoothness of the coefficient relative to subdomains of Ω.

DEFINITION 3.1 (oscillation condition). *Let $A \in C^p(\overline{\Omega}, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ for some $p \in \mathbb{N}$. A subset $\omega \subset \Omega$ fulfills the* oscillation condition of order p *if*

$$\operatorname{osc}(A,\omega,p) := \max_{1\le q\le p} \left\{ \frac{1}{q!} (\operatorname{diam}\omega)^q \, \|\nabla^q A\|_{L^\infty(\omega)} \right\} \le 1. \tag{3.1}$$

Definition 3.1 can be extended to the case $p = \infty$ by replacing $\max_{1\le q\le p}$ with $\sup_{q\in\mathbb{N}}$ in (3.1).

Note that the oscillation condition is fulfilled if and only if

$$\operatorname{diam}\omega \max_{1\le q\le p} \left\{ \left(\frac{1}{q!} \|\nabla^q A\|_{L^\infty(\omega)} \right)^{1/q} \right\} \le 1. \tag{3.2}$$

Furthermore, (3.2) implies that ω resolves the scales of ∇A in the sense that $\operatorname{diam}\omega \le \|\nabla A\|^{-1}_{L^\infty(\omega)}$ holds. Correspondingly, we define a function $H_{p,A} : \Omega \to (0,\infty)$ which turns out to measure the "variation" of the regularity for problem (2.2) from a standard Poisson problem. This function depends on the smoothness parameter p. The construction is as follows. We subdivide some bounding box $Q_0 \supset \bar{\Omega}$ into hypercubes that satisfy the oscillation condition. In the following, a cube $Q := \{x \in \mathbb{R}^d : \|x - c_Q\|_\infty \le R_Q\}$ shall be represented by its center c_Q and its radius R_Q (its halved width). For any parameter $\rho > 0$,

$$B_\rho(Q) := \{x \in \mathbb{R}^d : \ \|x - c_Q\|_\infty \le \rho R_Q\}$$

defines a ρ-scaled version of the cube Q. Clearly, $B_1(Q) = Q$.

ALGORITHM 3.2 (oscillation adapted covering). *Let $Q_0 \supset \bar{\Omega}$ be some closed bounding box of Ω. For $p \in \mathbb{N}$, a subdivision $\mathcal{Q} = \mathcal{Q}_p(A)$ of Q_0 into closed cubes is defined by the following:*

$\mathcal{Q} = \{Q_0\}$, $\mathcal{Q}^* := \emptyset$
while $\mathcal{Q}^* \neq \mathcal{Q}$ ***do***
  $\mathcal{Q}^* := \mathcal{Q}$
  for $Q \in \mathcal{Q}^*$ ***do***
    if $\operatorname{osc}(A, B_2(Q) \cap \Omega, p) > 1$ ***then***
      Q is subdivided into 2^d disjoint, congruent cubes $q_1, \dots, q_{2^d}$ and
      $\mathcal{Q} = (\mathcal{Q} \setminus Q) \cup \{q_1, \dots, q_{2^d}\}$
    end if
  end for
end while

REMARK 3.3.

(a) *Since $A \in C^p(\overline{\Omega}, \mathbb{R}^{d\times d}_{\mathrm{sym}})$, Algorithm 3.2 terminates. More precisely, $A \in C^p(\overline{\Omega}, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ implies $\operatorname{osc}(A,\Omega,p) \le C_{A,\Omega}$. If $Q \in \mathcal{Q}_p(A)$ is the result of ℓ refine-*

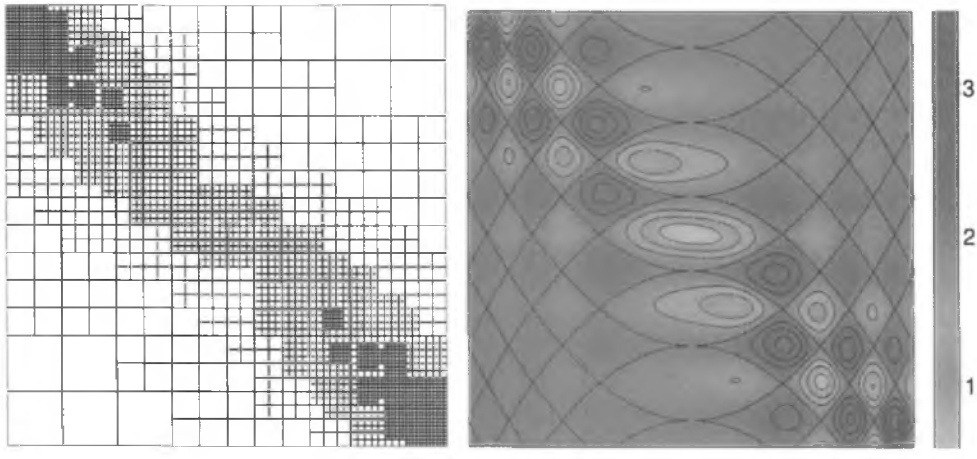

FIG. 3.1. *The oscillation adapted cover $\mathcal{Q}_{1,A}$ (left) for the scalar coefficient A (right) given in* (3.3); *$\mathcal{Q}_{1,A}$ was computed by Algorithm* 3.2 *with $Q_0 = [-1,1]^2$ and $p = 1$.*

ment steps in Algorithm 3.2, *then*

$$\operatorname{osc}(A, Q, p) \le 2^{-\ell} \left(\frac{\operatorname{diam} Q_0}{\operatorname{diam} \Omega} \right) C_{A,\Omega}.$$

Hence, the maximal number of refinement steps is bounded from above by

$$\left\lceil \log_2 (1 + C_{A,\Omega} \tfrac{\operatorname{diam} Q_0}{\operatorname{diam} \Omega}) \right\rceil,$$

where $\lceil \cdot \rceil$ denotes the ceiling function.

(b) *In the if-statement of Algorithm* 3.2, *the oscillation condition is checked in a certain neighborhood of the elements $Q \in \mathcal{Q}^*$, more precisely in $B_2(Q) \cap \Omega$. The overlap of the neighborhoods $B_2(Q) \cap \Omega$ accommodates the nonlocality of the differential operator; even locally, the solution depends on the global data. The choice of the size of the neighborhood affects the grading of the subdivision towards areas of high oscillation. The subsequent analysis shows that the choice* 2 *is appropriate.*

If the coefficient A oscillates periodically at frequency ε^{-1} for some small $\varepsilon > 0$, the output of Algorithm 3.2 is a subdivision of some bounding box into equal cubes with diameter proportional to ε. If the oscillations of A vary in space, for instance

$$A(x) = \left(2 - \frac{\cos(16\pi x_1^2) + \cos(16\pi x_2)}{11\left(\frac{1}{20} + (x+y)^2\right)} \right) I_{2\times 2} \quad \text{for } x \in (-1,1) \times (-1,1), \tag{3.3}$$

where $I_{2\times 2}$ is the 2×2 identity matrix, the corresponding adapted covering $\mathcal{Q}_{1,A}$ is highly nonuniform. Figure 3.1 shows that $\mathcal{Q}_{1,A}$ is coarse in regions of low variations and adaptively refined towards regions of high variation.

We shall now make a few observations concerning the local quasi-uniformity of the subdivisions $\mathcal{Q}_p(A)$ generated by Algorithm 3.2.

NOTATION 3.4. *We call two different cubes $Q_1, Q_2 \in \mathcal{Q}$* neighbored *if their boundaries have a common point. If a cube Q is refined during Algorithm* 3.2, *the resulting new cubes $q_1, \ldots, q_{2^d}$ are called* sons *of Q; correspondingly, Q is denoted the* father *of $q_1, \ldots, q_{2^d}$.*

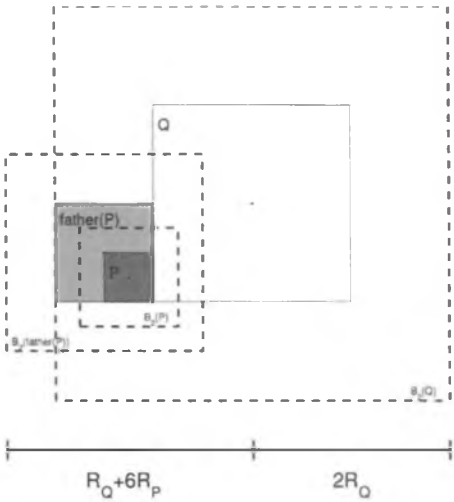

FIG. 3.2. *Illustration of the proof of Proposition* 3.5.

PROPOSITION 3.5. *If $P, Q \in \mathcal{Q}_p(A)$ are neighbored, then $\frac{1}{4}R_Q \le R_P \le 4R_Q$.*

Proof. $P \in \mathcal{Q} := \mathcal{Q}_p(A)$ implies that the father $\tilde{P}$ of P in the hierarchical construction of $\mathcal{Q}$ does not fulfill the oscillation condition. Necessarily, $B_2(\tilde{P}) \not\subset B_2(Q)$. Since $\tilde{P} \cap Q \neq \emptyset$, we get $\|c_Q - c_{\tilde{P}}\|_\infty \le R_Q + R_{\tilde{P}}$. The condition $B_2(\tilde{P}) \not\subset B_2(Q)$ can be rewritten as (see Figure 3.2 for an illustration)

$$R_Q + 3R_{\tilde{P}} = R_Q + 6R_P > R_{B_2(Q)} = 2R_Q,$$

which yields $R_{\tilde{P}} > \frac{1}{3}R_Q$ and $R_P = \frac{1}{2}R_{\tilde{P}} > \frac{1}{6}R_Q$. Since radii are successively halved in Algorithm 3.2, we finally get $R_P \ge \frac{1}{4}R_Q$. □

PROPOSITION 3.6. *There exists $C_{\mathrm{ol}} \in \mathbb{N}$ depending only on d such that for all $Q \in \mathcal{Q}_p(A)$ and for all $\eta \in [0,1)$ it holds that*

$$\#\{P \in \mathcal{Q}_p(A) : |P \cap B_{1+\eta}(Q)| > 0\} \le C_{\mathrm{ol}} M_d(\eta),$$

where $M_1(\eta) = \log(1-\eta)$ and $M_d(\eta) = (1-\eta)^{1-d}$ if $d \ge 2$.

Proof. Let $\mathcal{Q} := \mathcal{Q}_p(A)$ and $\mathcal{P}_Q := \{P \in \mathcal{Q} : |P \cap B_{1+\eta}(Q)| > 0\}$. The number of elements of $\mathcal{P}_Q$ is maximized in the scenario where Q is surrounded by layers each of which consists of congruent cubes of minimal size. Those layers are defined recursively: The first layer contains all neighbors of Q; for $k \ge 2$ the kth layer contains all neighbors of elements of the $(k-1)$st layer that are not contained in the layers $1, 2, \ldots, k-1$. By Proposition 3.5 all neighbors of Q have at least radius $\frac{1}{4}R_Q$. We will show later (see the end of this proof) that elements of the kth layer have at least radius $2^{-(k+1)}R_Q$. Assuming the latter statement is true, we compute that the thickness of the first K layers is $(2(1-(1/2)^{K+1})-1)R_Q$. Thus there are at most $K = \lceil -\log(1-\eta)/\log(2) \rceil$ layers within the η-neighborhood of Q and the proof is finished for $d = 1$. If $d \ge 2$, then the number of elements in $\mathcal{P}_Q$ can be bounded by $\sum_{k=1}^{K} 2^{(d-1)k} = \frac{2^{d-1}}{2^{d-1}-1}\frac{1-(1-\eta)^{d-1}}{(1-\eta)^{d-1}}$. This bound depends only on η and d but not on Q. It can be written in terms of $M_d(\eta)$, and the proof is finished.

The missing estimate on the minimal layer thickness is proved recursively. Assume that the radii of the kth layer elements are bounded from below by $2^{-(k+1)}R_Q$ for all

$k = 1, \ldots, L$ and that P is an element of layer $L+1$. Then

$$\|c_Q - c_P\|_\infty \geq R_Q + \sum_{k=1}^{L} 2(2^{-(k+1)}R_Q) + R_P = \left(\sum_{k=0}^{L} 2^{-k}\right) R_Q + R_P.$$

If $R_P \geq 2^{-(L+1)}R_Q$, nothing has to be shown. If, otherwise, $R_P \leq 2^{-(L+2)}R_Q$, the intersection of $\tilde{P}$ and elements of the layers $1, \ldots, L$ is of measure zero, $\tilde{P}$ being the father of P in the hierarchical construction of $\mathcal{Q}$. This yields

$$\|c_Q - c_{\tilde{P}}\|_\infty \geq \left(\sum_{k=0}^{L} 2^{-k}\right) R_Q + 2R_P,$$

$$\max_{y \in B_2(\tilde{P})} \|y - c_Q\|_\infty \geq \left(\sum_{k=0}^{L} 2^{-k}\right) R_Q + 6R_P.$$

Since $B_2(\tilde{P})$ does not satisfy the resolution condition whereas $B_2(Q)$ does, i.e., $B_2(\tilde{P}) \not\subset B_2(Q)$, we have

$$\left(\sum_{k=0}^{L} 2^{-k}\right) R_Q + 6R_P > 2R_Q.$$

Hence, $R_P > \frac{1}{6}2^{-L}R_Q = \frac{2}{3}2^{-(L+2)}R_Q$. Since radii are successively halved in Algorithm 3.2, we get the desired result $R_P \geq 2^{-(L+2)}R_Q$. □

A density function is now given by the local element size in $\mathcal{Q}_p(A)$.

DEFINITION 3.7 (oscillation adapted density). *Let $\mathcal{Q}_p(A)$, $p \in \mathbb{N}$, be some covering of Ω generated by Algorithm 3.2. Then $\mathcal{Q}_p(A)$-piecewise constant functions $H_{p,A} : \cup\mathcal{Q}_p(A) \to (0, \infty)$ are defined by*

$$H_{p,A}(x) := \min\{\operatorname{diam} Q : Q \in \mathcal{Q}_p(A) \text{ with } x \in Q\} \quad \text{for } x \in \cup\mathcal{Q}_p(A).$$

Figure 3.3 shows the reciprocal of the oscillation adapted density that corresponds to the coefficient given in (3.3). One observes the strong relation between the density $H_{1,A}^{-1}$ and the modulus of the gradient of A.

The analysis in section 4 shows that function $H_{p,A}$ contains important information about the diffusion coefficient A for higher order regularity estimates. However, the construction of $H_{p,A}$ via subdivisions into (overlapping) cubes is not well suited for the representation of the geometry of Ω and for finite element discretizations thereon. Since smooth domains, curvilinear polygons, and curved polyhedra are the relevant geometries for our theory, we construct a regular finite element mesh (cf. [7]) consisting of (possibly curved) simplices. The distribution of the simplices in this mesh is controlled by the oscillation adapted function $H_{p,A}$. In a first step, we introduce an initial coarse mesh that resolves the geometry. In a second step, based on the function $H_{p,A}$, the initial mesh is refined according to the oscillations of the coefficient.

DEFINITION 3.8 (macrotriangulation, refinement, parametrization).

(a) *We assume that there exists a polyhedral (polygonal in two dimensions) domain $\tilde{\Omega}$ along with a bi-Lipschitz mapping $\chi : \tilde{\Omega} \to \Omega$. Let $\tilde{\mathcal{T}}^{\text{macro}} = \{\tilde{K}_i^{\text{macro}} : 1 \leq i \leq q\}$ denote some conforming finite element mesh for $\tilde{\Omega}$ consisting of simplices which are regular in the sense of [7]. $\tilde{\mathcal{T}}^{\text{macro}}$ is considered as a coarse partition of $\tilde{\Omega}$; i.e., the*

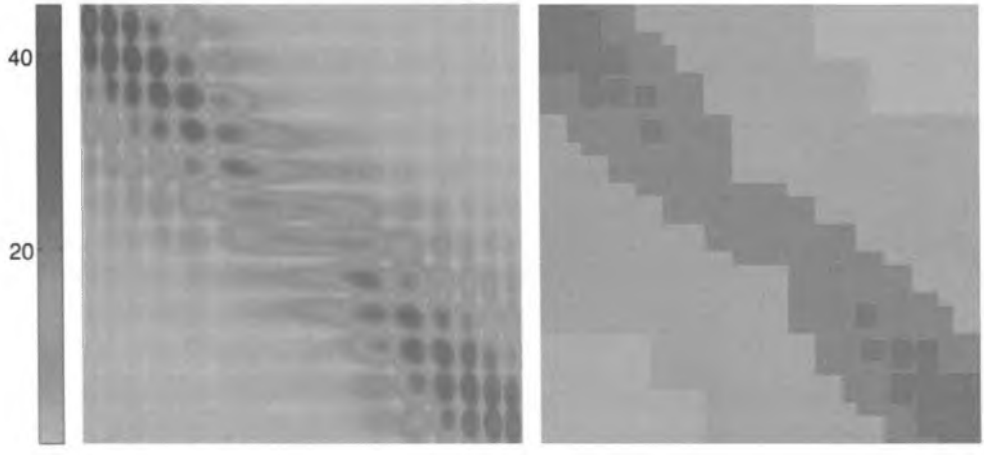

FIG. 3.3. *The modulus of the gradient* $|\nabla A|$ *(left) of the coefficient* A *given in* (3.3) *and the reciprocal of the corresponding oscillation adapted density* $H_{1,A}$ *(right) derived from the oscillation adapted covering* $\mathcal{Q}_{1,A}$ *that is shown in Figure* 3.1.

diameters of the elements in $\widetilde{\mathcal{T}}^{\text{macro}}$ *are of order* 1. *We assume that the restrictions* $\chi_i := \chi|_{\widetilde{K}_i^{\text{macro}}}$ *are analytic for all* $1 \le i \le q$. *The macromesh for* Ω *is then given by*

$$\mathcal{T}^{\text{macro}} := \left\{ K = \chi(\widetilde{K}^{\text{macro}}) : \widetilde{K}^{\text{macro}} \in \widetilde{\mathcal{T}} \right\}.$$

(b) *Using the macromesh as the initial mesh we introduce a recursive refinement procedure* REFINE. *The input of* REFINE *is a finite element mesh* $\mathcal{T}$, *where some elements are marked for refinement, and the output is a new conforming finite element mesh* $\mathcal{T}^{\text{refine}}$ *in the sense of* [7]. *The output is derived by refining the corresponding simplicial mesh* $\widetilde{\mathcal{T}}$ *in a standard way (e.g., in two dimensions, by first connecting the midpoints of the marked triangle edges and then eliminating hanging nodes by some suitable closure algorithm). The resulting mesh is denoted by* $\widetilde{\mathcal{T}}^{\text{refine}} = \{\widetilde{K}_i : 1 \le i \le N\}$. *The corresponding finite element mesh for* Ω *is denoted by* $\mathcal{T}^{\text{refine}} = \left\{K = \chi(\widetilde{K}) : \widetilde{K} \in \widetilde{\mathcal{T}}^{\text{refine}}\right\}$. *As a simplifying assumption on the refinement strategy we assume that the elimination of hanging nodes causes refinement of nonmarked triangles only in the first layer around marked triangles. In certain cases this strategy generates meshes with some "flat" triangles; i.e., the constant measuring the shape regularity of the mesh increases.*[1]

(c) *There exists an affine bijection* $J_K : \widehat{K} \to \widetilde{K}$ *which maps the reference element* $\widehat{K} := \{x \in [0,\infty))^d : \sum_{i=1}^d x_i \le 1\}$ *to the simplex* $\widetilde{K}$ *for any* $K = \chi(\widetilde{K}) \in \mathcal{T}$, *where* $\mathcal{T}$ *is derived from* $\mathcal{T}^{\text{macro}}$ *by repeated application of* REFINE. *A parametrization* $F_K : \widehat{K} \to K$ *can be written as* $F_K = R_K \circ J_K$, *where* J_K *is an affine map and the*

[1]One could avoid this degeneration by allowing the closure algorithm to spread out by more than one layer about the red-refined triangles. The generalization of our theory to this version of the closure algorithm, however, would require further technicalities. To avoid this for the sake of readability we impose our simplifying assumption on the closure algorithm.

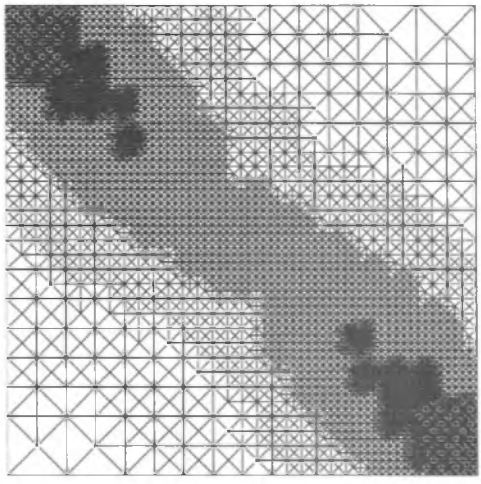

FIG. 3.4. *The oscillation adapted finite element mesh $\mathcal{T}_{1,A}$ derived from the adapted cover $\mathcal{Q}_{1,A}$ shown in Figure 3.1 by Algorithm 3.9; the underlying coefficient A is given in* (3.3).

maps R_K and J_K satisfy for constants C_{affine}, C_{metric}, $\gamma > 0$

$$
\begin{aligned}
\|J_K'\|_{L^\infty(\widehat{K})} &\le C_{\text{affine}} \operatorname{diam}(K),\\
\|(J_K')^{-1}\|_{L^\infty(\widehat{K})} &\le C_{\text{affine}} \operatorname{diam}(K)^{-1},\\
\|(R_K')^{-1}\|_{L^\infty(\widehat{K})} &\le C_{\text{metric}},\\
\|\nabla^n R_K\|_{L^\infty(\widehat{K})} &\le C_{\text{metric}} \gamma^n n! \quad \text{for } n \in \mathbb{N}_0.
\end{aligned}
\tag{3.4}
$$

Driven by the density function $H_{p,A}$, the actual oscillation adapted meshes are derived by successively refining the macromesh as follows.

ALGORITHM 3.9 (oscillation adapted finite element mesh). *Let $\mathcal{T}^{\text{macro}}$ be a subdivision of $\bar{\Omega}$ in the sense of Definition 3.8, and let $p \in \mathbb{N}$. A subdivision $\mathcal{T}_p(A)$ of Ω that (as we will prove) reflects the regularity of the coefficient is defined by the following:*

$\mathcal{T} := \mathcal{T}^{\text{macro}}$
for $q = 1, 2, \ldots, p$ ***do***
  $\mathcal{M} := \mathcal{T}$
  while $\mathcal{M} \neq \emptyset$ ***do***
    $\mathcal{M} := \{K \in \mathcal{T} : \operatorname{diam}(K) > \min_{x \in K} H_{q,A}(x)\}$
    $\mathcal{T} = \text{REFINE}(\mathcal{T}, \mathcal{M})$
  end while
end for

Figure 3.4 depicts the oscillation adapted finite element mesh $\mathcal{T}_{1,A}$ according to the coefficient A from (3.3) and its adapted covering $\mathcal{Q}_{1,A}$ shown in Figure 3.1. Further remarks are in order.

REMARK 3.10.

(a) *The mesh $\mathcal{T}_p(A)$ serves as a starting mesh for further regular refinements. Then the final mesh $\mathcal{T}_h$ is again a finite element mesh for Ω and satisfies the following: For all $K \in \mathcal{T}_p(A)$ there exists a set of sons* $\operatorname{sons}(K) \subset \mathcal{T}_h$ *such that* $K = \bigcup\{K' : K' \in \operatorname{sons}(t)\}$. *The diameter of $K \in \mathcal{T}_h$ is denoted by h_K, and we are using the maximal mesh width $h := \max\{h_K : K \in \mathcal{T}_h\}$ as the index in $\mathcal{T}_h$.*

(b) *Given some initial mesh* $\mathcal{T}^{\text{macro}}$ *and* $\ell, \ell' \in \mathbb{N}$, $\ell < \ell'$, *let* $\mathcal{T}_\ell(A)$ *and* $\mathcal{T}_{\ell'}(A)$ *denote the meshes generated by Algorithm* 3.9. *By construction,* $\mathcal{T}_{\ell'}$ *is a refinement of* $\mathcal{T}_\ell$.

Our goal is to discretize (2.2) by the Galerkin finite element method. It turns out that the ratio $\max\{h_{K'} : K' \in \operatorname{sons} K\}/h_K$, $K \in \mathcal{T}_p(A)$, plays the essential role for the error estimates.

We shall prove that the mesh $\mathcal{T}_p(A)$ has analog properties as the mesh $\mathcal{Q}_p(A)$—it satisfies Propositions 3.5 and 3.6. In addition, it is a simplicial finite element mesh. For $K \in \mathcal{T}$ and $\rho \geq 1$, some scaled neighborhood of K is defined by

$$K_\rho := \{x \in \mathbb{R}^d : \exists y \in K : \|y - x\|_2 \leq \tfrac{\rho}{2} \operatorname{diam}(K)\}. \tag{3.5}$$

Definition 3.8 implies that $\rho_{\mathcal{T}} := \max\{\operatorname{diam}(T)^d/|T| : T \in \mathcal{T}\}$ is bounded from above by a constant which depends only on C_{affine} and C_{metric} from (3.4).

LEMMA 3.11. *Let* $\mathcal{Q} = \mathcal{Q}_p(A)$ *(resp.,* $\mathcal{T} = \mathcal{T}_p(A)$*),* $p \in \mathbb{N}$, *be the subdivisions generated by Algorithm* 3.2 *(resp.,* 3.9*) which "resolve" the coefficient* A. *Then the following hold:*

(a) *There exist* $C_1(d, C_{\text{affine}}, C_{\text{metric}}), C_2(d) \in \mathbb{N}$ *such that for all* $Q \in \mathcal{Q}$ *and all* $K \in \mathcal{T}_p$, $p \in \mathbb{N}$,

$$\#\{T \in \mathcal{T} : T \cap Q \neq \emptyset\} \leq C_1 \quad \textit{and} \quad \#\{P \in \mathcal{Q} : P \cap K \neq \emptyset\} \leq C_2.$$

(b) *For all* $\eta \in [0, 1)$, *there exists* $C'_{\text{ol}}(C_1, C_{\text{ol}}, C_2) > 0$ *such that for all* $K \in \mathcal{T}$,

$$\#\{T \in \mathcal{T}_p : |T \cap K_{1+\eta}| > 0\} \leq C'_{\text{ol}} \begin{cases} \log(1-\eta) & \textit{if } d = 1, \\ (1-\eta)^{1-d} & \textit{if } d \geq 2. \end{cases}$$

Proof. Let $K \in \mathcal{T}$ and $Q \in \mathcal{Q}$ be given such that $K \cap Q \neq \emptyset$. Then, depending on the actual realization of the procedure REFINE, there exists $\theta > 0$ such that

$$\theta \operatorname{diam}(Q) \leq \operatorname{diam}(K) \leq \operatorname{diam}(Q). \tag{3.6}$$

Let $\mathcal{N}(Q) := \bigcup\{P \in \mathcal{Q} : P \text{ and } Q \text{ are neighbored}\}$ denote some neighborhood of Q in $\mathcal{Q}$. Then C_1 can be estimated by

$$\begin{aligned} |\mathcal{N}(Q)| &= \sum_{T \in \mathcal{T} : T \cap \mathcal{N}(Q) \neq \emptyset} |T \cap \mathcal{N}(Q)| \geq \sum_{T \in \mathcal{T} : T \cap Q \neq \emptyset} |T \cap \mathcal{N}(Q)| \\ &\geq \sum_{T \in \mathcal{T} : T \cap Q \neq \emptyset} \tfrac{\theta}{4\rho_{\mathcal{T}}} \operatorname{diam}(Q)^d \geq C_1 \tfrac{\theta}{4\rho_{\mathcal{T}}} |Q|. \end{aligned}$$

This implies that C_1 is bounded in terms of θ, $\rho_{\mathcal{T}}$, and d. An analog argument proves that C_2 is finite and therefore proves part (a).

Part (b) follows from part (a): There are at most C_1 cubes which intersect K. Proposition 3.6 shows that in an η-neighborhood of every such cube there are at most $C_{\text{ol}} M_d(\eta)$ elements of $\mathcal{Q}$. Hence, $K_{1+\eta}$ is covered by at most $C_1 C_{\text{ol}} M_d(\eta)$ cubes. Due to part (a), each of the latter cubes is again intersected by at most C_2 simplices. □

By construction, the elements of the oscillation adapted finite element mesh satisfy the oscillation condition

$$\operatorname{osc}(A, K_2, p) < 1 \quad \forall K \in \mathcal{T}_p(A). \tag{3.7a}$$

LEMMA 3.12. *For all $K \in \mathcal{T}_p(A)$, the lower bound*

$$h_K \geq c \min\left\{\tau, \left(\max_{1\leq q\leq p}\left\{\left(\frac{\|\nabla^q A\|_{L^\infty(K^*)}}{q!}\right)^{1/q}\right\}\right)^{-1}\right\} \tag{3.7b}$$

holds with a constant τ that represents the minimal mesh size in the initial macromesh $\mathcal{T}^{\text{macro}}$ (cf. Definition 3.8), and with a constant $c > 0$ that depends only on the shape parameters in $\mathcal{T}^{\text{macro}}$ and, through (3.6), the procedure REFINE*; $K^* := K_C$ denotes the C-scaled version of K (cf. (3.5)), where the constant C depends only on the shape parameters in $\mathcal{T}^{\text{macro}}$ and the procedure* REFINE.

Proof. If K is an element of the initial macrotriangulation $\mathcal{T}^{\text{macro}}$, then, by choosing τ appropriately, the assertion can always be satisfied. If $K \in \mathcal{T}_p(A)$ originates from some father simplex $\tilde{K}$ through refinement, $\tilde{K}$ or one of its neighbors is marked in Algorithm 3.9. The marking of $\tilde{K}$ implies the existence of some $Q \in \mathcal{Q}_p(A)$ so that $\operatorname{diam}(\tilde{K}) > \operatorname{diam}(Q)$. If $Q \neq Q_0$, then the scaled version of its father $B_2(\tilde{Q})$ violates the oscillation condition (see (3.2)), i.e.,

$$\operatorname{diam}(B_2(\tilde{Q})) \max_{1\leq q\leq p}\left\{\left(\|\nabla^q A\|_{L^\infty(B_2(\tilde{Q}))}/q!\right)^{1/q}\right\} > 1.$$

This yields

$$\begin{aligned}\operatorname{diam}(K) &\geq \theta \operatorname{diam}(\tilde{K}) > (\theta/4)\operatorname{diam}(B_2(\tilde{Q})) \\ &\geq \frac{\theta}{4}\left(\max_{1\leq q\leq p}\left\{\left(\|\nabla^q A\|_{L^\infty(\tilde{Q}_2)}/q!\right)^{1/q}\right\}\right)^{-1}.\end{aligned} \tag{3.8}$$

Based on mesh regularity, a similar estimate can be derived in the case where the refinement of $\tilde{K}$ is due to the preservation of conformity. Therefore (3.7b) is proved. □

Finally, we introduce weighted (mesh-dependent) Sobolev norms.

DEFINITION 3.13 (oscillation adapted Sobolev norms). *Let $\mathcal{T}_p(A)$, $p \in \mathbb{N}$, be the subdivision of Ω generated by Algorithm 3.9. A weighted seminorm $|\cdot|_{p+1,A}$ in $H^{p+1}(\Omega)$ is defined by*

$$|u|_{p+1,A} := \frac{1}{p!}\sqrt{\sum_{K\in\mathcal{T}_p(A)} \operatorname{diam}(K)^{2p}\|\nabla^{p+1}u\|^2_{L^2(K)}},$$

while corresponding full norms are given by

$$\|u\|_{p+1,A} := \sqrt{\|u\|^2_{H^1(\Omega)} + \sum_{\ell=2}^{p+1}|u|^2_{\ell,A}}.$$

4. Oscillation adapted regularity. Subsection 4.1 presents the main result concerning the regularity estimates in weighted Sobolev norms. Its proof, which is based on local interior regularity estimates, is devoted to sections 4.2 and 4.3. Our theory can be regarded as a generalization of the regularity theory of [25] which, in turn, is based on techniques established in the classical book [28]. In contrast to [25] we do not assume that the diffusion matrix is analytic but of finite smoothness. Hence, we *cannot* characterize the growth of the derivatives in terms of only two constants stemming, e.g., from Cauchy's integral formula for analytic functions as in

[25]. Instead, we employ the density function and its derivatives and the oscillation condition as weights in higher order Sobolev norms and prove regularity estimates for these norms, where the constants are independent of the derivatives of A. This requires some nontrivial generalization of the results in [25]. Sections 4.2 and 4.3 contain only the parts of the proofs which are directly related to the regularity estimates. Further technical results concerning the growth of derivatives of composite functions with finite smoothness are contained in Appendix A.

4.1. Main regularity result.

THEOREM 4.1. *Let $A \in C^p(\Omega, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ satisfy (2.1) for some $p \in \mathbb{N}$, and assume $f \in H^{p-1}(\Omega)$. The corresponding solution of (2.2) is denoted by u. Further assume that the mesh $\mathcal{T}_p(A)$ is generated by Algorithm 3.9. Let the boundary $\partial\Omega$ be of class C^p. Then the solution satisfies $u \in H^{p+1}(\Omega)$ and the estimate*

$$\|u\|_{p+1,A} \le C_{11} C_{12}^p \|f\|_{H^{p-1}(\Omega)}, \tag{4.1}$$

where $\|\cdot\|_{p+1,A}$ is the oscillation adapted Sobolev norm as in Definition 3.13. The constants C_{11} and C_{12} are independent of p and the variation of A but depend on α, β as in (2.1), on C_{ol} (cf. Proposition 3.6), on the constants in Definition 3.8(c), on the spatial dimension d, and on the geometry of the domain Ω through its diameter and the constants describing the regularity of the boundary $\partial\Omega$.

Proof. For $K \in \mathcal{T}_p(A)$, let $H_K := \sup\{H_{p,A}(x) \mid x \in K\}$, and let $H_{\max} := \|H_{p,A}\|_{L^\infty(\Omega)}$. Then by choosing $0 < \eta < 1$ as in Lemma 3.11(b) and using Lemmata 4.5, 4.6, and 4.8 we obtain

$$\begin{aligned}
&\sum_{K\in\mathcal{T}_p(A)} \frac{H_K^{2p} \|\nabla^{p+1} u\|^2_{L^2(K)}}{(p!)^2} \\
&\quad \le C_8^2 C_9^{2p} p \sum_{K\in\mathcal{T}_p(A)} \left(\|\nabla u\|^2_{L^2(K_{1+\eta})} + \sum_{i=0}^{p-1} \frac{H_K^{2+2i}}{(i+1)!^2} \|\nabla^i f\|^2_{L^2(K_{1+\eta})} \right) \\
&\quad \overset{(3.7a)}{\le} C_{10}^2 C_9^{2p} p \left(\|\nabla u\|^2_{L^2(\Omega)} + \sum_{i=0}^{p-1} \frac{H_{\max}^{2+2i}}{(i+1)!^2} \|\nabla^i f\|^2_{L^2(\Omega)} \right).
\end{aligned}$$

Since $0 < \alpha := \alpha(A, \Omega)$ is bounded from below and $H_{\max} \le \operatorname{diam}\Omega$, we obtain

$$\sum_{K\in\mathcal{T}_p(A)} \frac{H_K^{2p} \|\nabla^{p+1} u\|^2_{L^2(K)}}{(p)!^2} \le C_{11}^2 C_9^{2p} p \|f\|^2_{H^{p-1}(\Omega)} \le C_{11}^2 (C_9')^{2p} \|f\|^2_{H^{p-1}(\Omega)}, \tag{4.2}$$

where C_{11} also depends on $H_{\max}$. For the estimate of the full norm, we get

$$\begin{aligned}
\|u\|^2_{p+1,A} &= \|u\|^2_{H^1(\Omega)} + \sum_{\ell=1}^{p} |u|^2_{\ell+1,A} \\
&\le C \|f\|^2_{L^2(\Omega)} + \sum_{\ell=1}^{p} C_{11}^2 (C_9')^{2\ell} \|f\|^2_{H^{\ell-1}(\Omega)} \\
&\le C_{11}^2 C_{12}^{2p} \|f\|^2_{H^{p-1}(\Omega)}. \qquad \square
\end{aligned}$$

4.2. Interior regularity. We employ the framework of [28], [25] to derive local high order interior regularity estimates. For technical reasons, which are related to the construction of the oscillation adaptive covering and the finite element method, we replace the Euclidean balls used in [28], [25] by simplices. For $R > 0$, let the d-dimensional scaled unit simplex (with barycenter at the origin) be denoted by

$$\widehat{T}_R := \left\{ x - \frac{1}{d+1}(R, R, \ldots, R)^\intercal \mid x \in [0,\infty)^d \text{ and } \|x\|_{\ell^1} \le R \right\}.$$

LEMMA 4.2 (H^2-regularity). *Let $f \in L^2(\widehat{T}_R)$ and $A \in C^1(\widehat{T}_R, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ such that $0 < \alpha := \alpha(A, \widehat{T}_R)$ and $\beta(A, \widehat{T}_R) =: \beta < \infty$ (cf. (2.1)). Assume that $\operatorname{osc}(A, \widehat{T}_R, 1) \le 1$. Then there exists a constant C_{I} depending only on α and β such that the weak solution u of*

$$-\operatorname{div}(A\nabla u) = f \quad \text{in } \widehat{T}_R, \quad u = 0 \quad \text{on } \partial\widehat{T}_R$$

is in $H^2(\widehat{T}_R)$ and satisfies

$$\left\|\nabla^2 u\right\|_{L^2(\widehat{T}_R)} \le C_{\mathrm{I}} \|f\|_{L^2(\widehat{T}_R)}. \tag{4.3}$$

The proof follows by scaling the problem to the unit simplex $\widehat{T}_1$ (as explained in [25, Lem. 5.5.5]) and by using standard regularity estimates (see, e.g., [9], [11]).

LEMMA 4.3 (interior regularity). *Let the assumptions of Lemma 4.2 be satisfied. Then there exists a constant $C_{\mathrm{I}}' > 0$ depending only on $\alpha(A, \widehat{T}_R)$ and $\beta(A, \widehat{T}_R)$ such that any solution of*

$$-\operatorname{div}(A\nabla u) = f \quad \text{in } \widehat{T}_R \tag{4.4}$$

satisfies

$$\left\|\nabla^2 u\right\|_{L^2(\widehat{T}_r)} \le C_{\mathrm{I}}' \left(\|f\|_{L^2(\widehat{T}_{r+\delta})} + \delta^{-1} \|\nabla u\|_{L^2(\widehat{T}_{r+\delta})} + \delta^{-2} \|u\|_{L^2(\widehat{T}_{r+\delta})} \right)$$

for all $r, \delta > 0$ with $r + \delta < R$.

Proof (see [28, Lem. 5.7.1], [25, Lem. 5.5.11] for ball-shaped domains). We employ a cutoff function χ being identically one on $\widehat{T}_{R-\delta}$, vanishing on $\widehat{T}_R \backslash \widehat{T}_{R-\delta/2}$, and satisfying $\left\|\nabla^j \chi\right\|_{L^\infty(\widehat{T}_R)} \le C\delta^{-j}$, $j = 0, 1, 2$. Then $U = \chi u$ satisfies

$$-\operatorname{div}(A\nabla U) = \chi f - 2\langle \nabla u, A\nabla\chi\rangle - u \operatorname{div}(A\nabla\chi) \quad \text{in } \widehat{T}_R \quad \text{and} \quad U = 0 \quad \text{on } \partial\widehat{T}_R.$$

By using (4.3) and triangle inequalities in combination with Hölder inequalities, we get

$$\begin{aligned}
\left\|\nabla^2 u\right\|_{L^2(\widehat{T}_{R-\delta})} &\le \left\|\nabla^2 U\right\|_{L^2(\widehat{T}_R)} \\
&\le C_{\mathrm{I}} \Big(\|f\|_{L^2(\widehat{T}_R)} \|\chi\|_{L^\infty(\widehat{T}_R)} + 2 \|\nabla u\|_{L^2(\widehat{T}_R)} \|A\|_{L^\infty(\widehat{T}_R)} \|\nabla\chi\|_{L^\infty(\widehat{T}_R)} \\
&\qquad + \|u\|_{L^2(\widehat{T}_R)} \|\operatorname{div}(A\nabla\chi)\|_{L^\infty(\widehat{T}_R)} \Big).
\end{aligned} \tag{4.5}$$

The assumptions on the cutoff function and A imply that $\|\chi\|_{L^\infty(\widehat{T}_R)} \le C$ and

$$\|A\|_{L^\infty(\widehat{T}_R)} \|\nabla\chi\|_{L^\infty(\widehat{T}_R)} \le C\beta\left(A, \widehat{T}_R\right)\delta^{-1}, \tag{4.6a}$$

$$\begin{aligned}
\|\operatorname{div}(A\nabla\chi)\|_{L^\infty(\widehat{T}_R)} &\le C\left(R\|A'\|_{L^\infty(\widehat{T}_R)} (R\delta)^{-1} + \beta\left(A, \widehat{T}_R\right)\delta^{-2} \right) \\
&\le C\delta^{-2}\left(1 + \beta\left(A, \widehat{T}_R\right)\right).
\end{aligned} \tag{4.6b}$$

The combination of (4.5) and (4.6) leads to the assertion. □

For the estimate of the higher order derivatives we need some further notation. Let

$$N_{R,\ell}(v) := \frac{1}{[\ell]!} \sup_{R/2 \le r < R} (R-r)^{2+\ell} \left\|\nabla^{\ell+2} v\right\|_{L^2(\widehat{T}_r)}, \quad \ell \in \mathbb{N}_0 \cup \{-2,-1\}, \tag{4.7a}$$

$$M_{R,\ell}(v) := \frac{1}{\ell!} \sup_{R/2 \le r < R} (R-r)^{2+\ell} \left\|\nabla^{\ell} v\right\|_{L^2(\widehat{T}_r)}, \quad \ell \in \mathbb{N}_0, \tag{4.7b}$$

where $[\ell] := \max\{1,\ell\}$. Note that for any $1/2 \le \eta < 1$

$$\left\|\nabla^{\ell+2} v\right\|_{L^2(\widehat{T}_{\eta R})} \le \frac{[\ell]!}{((1-\eta)R)^{2+\ell}} N_{R,\ell}(v). \tag{4.8}$$

LEMMA 4.4 (interior higher order regularity). *For $\ell \in \mathbb{N}_0$, we assume (cf. (2.1)) that*

$$A \in C^{\ell+1}(\widehat{T}_R, \mathbb{R}^{d\times d}_{\mathrm{sym}}), \quad \operatorname{osc}\left(A, \widehat{T}_R, \ell+1\right) \le \kappa \quad \textit{for some } \kappa > 0,$$
$$0 < \alpha := \alpha\left(A, \widehat{T}_R\right), \quad \beta\left(A, \widehat{T}_R\right) =: \beta < \infty.$$

Then there exists $C_{\mathrm{I}}' > 0$ depending only on α, β, and d such that

$$N_{R,\ell}(u) \le C_{\mathrm{I}}' \left(M_{R,\ell}(f) + (1+\kappa) \sum_{q=1}^{\ell+1} \frac{\ell+1}{2^q [\ell+1-q]} N_{R,\ell-q}(u) \right) \tag{4.9}$$

for any $f \in H^\ell(\widehat{T}_R)$ and any solution u of

$$-\operatorname{div}(A\nabla u) = f \quad \textit{on } \widehat{T}_R.$$

Proof. One easily checks that the proof of [25, Lem. 5.5.12] carries over to the scaled unit simplex (instead of balls) so that

$$N_{R,\ell}(u) \le C_{\mathrm{I}}' \Bigg(M_{R,\ell}(f) + \sum_{q=1}^{\ell+1} \binom{\ell+1}{q} \left(\frac{R}{2}\right)^q \left\|\nabla^q A\right\|_{L^\infty(\widehat{T}_R)} \frac{[\ell-q]!}{\ell!} N_{R,\ell-q}(u) + N_{R,\ell-1}(u) + N_{R,\ell-2}(u) \Bigg). \tag{4.10}$$

From $\operatorname{osc}\left(A, \widehat{T}_R, \ell+1\right) \le \kappa$ we conclude that

$$N_{R,\ell}(u) \le C_{\mathrm{I}}' \left(M_{R,\ell}(f) + \kappa \sum_{q=1}^{\ell+1} \frac{\ell+1}{2^q [\ell+1-q]} N_{R,\ell-q}(u) + N_{R,\ell-1}(u) + N_{R,\ell-2}(u) \right). \tag{4.11}$$

Since the factors in front of $N_{R,\ell-1}(u)$ and $N_{R,\ell-2}(u)$ above are 1, the assertion follows. □

LEMMA 4.5. *For $p \in \mathbb{N}$, we assume that*

$$A \in C^p(\widehat{T}_R, \mathbb{R}^{d\times d}_{\mathrm{sym}}), \quad \operatorname{osc}\left(A, \widehat{T}_R, p\right) \le \kappa \quad \textit{for some } \kappa > 0,$$
$$0 < \alpha := \alpha\left(A, \widehat{T}_R\right), \quad \beta\left(A, \widehat{T}_R\right) =: \beta < \infty.$$

Then for any $f \in H^{p-1}\left(\widehat{T}_R\right)$ *and any solution* u *of*

$$-\operatorname{div}(A\nabla u) = f \quad \text{on } \widehat{T}_R$$

it holds that

$$\frac{R^p}{p!}\left\|\nabla^{p+1}u\right\|_{L^2(\widehat{T}_{\eta R})} \le C_1 C_2^p \left(\|\nabla u\|_{L^2(\widehat{T}_R)} + \sum_{i=0}^{p-1}\frac{R^{1+i}}{(i+1)!}\left\|\nabla^i f\right\|_{L^2(\widehat{T}_R)}\right)$$

for all $1/2 \le \eta < 1$ *with*

$$C_1 := \frac{\lambda + 1 + C_1'}{(1-\eta)(1+\lambda)}, \quad C_2 := \frac{\lambda+1}{2(1-\eta)}, \quad \lambda := 2C_1'(1+\kappa). \tag{4.12}$$

Proof. The estimate

$$N_{R,-1}(u) \le \frac{R}{2}\|\nabla u\|_{L^2(\widehat{T}_R)} \tag{4.13}$$

follows directly from (4.7a). Definition (4.7a) implies that

$$N_{R,0}(u) = \sup_{R/2 \le r < R}(R-r)^2\left\|\nabla^2 u\right\|_{L^2(\widehat{T}_r)}.$$

Next, we estimate the recursion (4.9) and define
(4.14)

$$N_{-1} := N_{R,-1}(u) \quad \text{and, for } \ell = 0,1,2\ldots, \quad N_\ell := C_\ell + \lambda\sum_{q=1}^{\ell+1}\frac{\ell+2}{2^q(\ell+2-q)}N_{\ell-q},$$

where $C_\ell := C_1' M_{R,\ell}(f)$ and λ as in (4.12). It follows directly by comparing (4.9) with (4.14) that $N_{R,\ell}(u) \le N_\ell$. We set $\tilde{N}_\ell := 2^\ell N_\ell/(\ell+2)$ and $\tilde{C}_\ell := C_\ell 2^\ell/(\ell+2)$ to obtain

$$\tilde{N}_{-1} = N_{-1}/2 \quad \text{and, for } \ell = 0,1,2\ldots, \quad \tilde{N}_\ell = \tilde{C}_\ell + \lambda\sum_{q=1}^{\ell+1}\tilde{N}_{\ell-q}.$$

This recursion can be resolved, and we get, for all $\ell \ge 0$,

$$\begin{aligned}\tilde{N}_\ell &\le \tilde{C}_\ell + \lambda(\lambda+1)^\ell\tilde{N}_{-1} + \lambda\sum_{i=0}^{\ell-1}(\lambda+1)^{\ell-1-i}\tilde{C}_i \\ &\le (\lambda+1)^{\ell+1}\tilde{N}_{-1} + \sum_{i=0}^{\ell}(\lambda+1)^{\ell-i}\tilde{C}_i.\end{aligned}$$

By substituting back the original quantities we derive

$$\frac{N_\ell}{\ell+2} \le \left(\frac{\lambda+1}{2}\right)^{\ell+1}N_{-1} + C_1'\sum_{i=0}^{\ell}\left(\frac{\lambda+1}{2}\right)^{\ell-i}\frac{M_{R,i}(f)}{(i+2)}. \tag{4.15}$$

The combination of (4.8), (4.13), (4.15), and $M_{R,i}(f) \le \frac{1}{i!}\left(\frac{R}{2}\right)^{2+i}\left\|\nabla^i f\right\|_{L^2(\widehat{T}_R)}$ with some elementary estimates leads to the assertion. □

4.3. Regularity at the boundary. For $R > 0$, let the d-dimensional scaled unit simplex be denoted by

$$\widehat{T}_R^+ := \left\{x \mid x \in [0,\infty)^d \wedge \|x\|_{\ell^1} \le R\right\},$$

and let $\Gamma_R^+ := \left\{x \in \overline{\widehat{T}_R^+} \mid x_d = 0\right\}$ be its horizontal facet.

For the estimate of the solution in $\widehat{T}_r^+$ by quantities in a certain neighborhood, we proceed along the lines of [25, sect. 5.5.3] and derive estimates for the normal and tangential derivatives at the boundary separately.

4.3.1. Control of tangential derivatives. Let $x = (x_1, x_2 \ldots, x_{d-1})$ denote the tangential variables with respect to Γ_R^+. The derivatives with respect to x are denoted by ∇_x. We need the following notation:

$$N_{R,\ell}^+(v) := \begin{cases} \dfrac{1}{\ell!} \displaystyle\sup_{R/2 \le r < R} (R-r)^{\ell+2} \left\|\nabla^2 \nabla_x^\ell v\right\|_{L^2(\widehat{T}_r^+)} & \text{if } \ell \ge 0, \\ \displaystyle\sup_{R/2 \le r < R} (R-r)^{\ell+2} \left\|\nabla^{2+\ell} v\right\|_{L^2(\widehat{T}_r^+)} & \text{if } \ell = -2, -1, \end{cases} \tag{4.16a}$$

$$M_{R,\ell}^+(v) := \frac{1}{\ell!} \sup_{R/2 \le r < R} (R-r)^{\ell+2} \left\|\nabla_x^\ell v\right\|_{L^2(\widehat{T}_r^+)}. \tag{4.16b}$$

LEMMA 4.6. *For $p \in \mathbb{N}$, we assume (cf. (2.1)) that*

$$A \in C^p(\widehat{T}_R^+, \mathbb{R}^{d\times d}_{\mathrm{sym}}), \qquad \operatorname{osc}\left(A, \widehat{T}_R^+, p\right) \le \kappa \quad \text{for some } \kappa > 0,$$
$$0 < \alpha := \alpha\left(A, \widehat{T}_R^+\right), \quad \beta\left(A, \widehat{T}_R^+\right) =: \beta < \infty.$$

Then there exists $C_{\mathrm{B}}' > 0$ depending only on α, β, and d such that for all $f \in H^{p-1}\left(\widehat{T}_R^+\right)$ and any solution u of

$$-\operatorname{div}(A\nabla u) = f \quad \text{in } \widehat{T}_R^+, \quad u = 0 \quad \text{on } \Gamma_R^+ \tag{4.17}$$

we have

$$\frac{R^p}{p!} \left\|\nabla_x^{p-1} \nabla^2 u\right\|_{L^2(\widehat{T}_{\eta R}^+)} \le C_1 C_2^p \left(\|\nabla u\|_{L^2(\widehat{T}_R^+)} + \sum_{i=0}^{p-1} \frac{R^{1+i}}{(i+1)!} \left\|\nabla^i f\right\|_{L^2(\widehat{T}_R^+)} \right)$$

for all $1/2 \le \eta < 1$ with

$$C_1 := \frac{\lambda_{\mathrm{B}} + 1 + C_{\mathrm{B}}'}{(1-\eta)(1+\lambda_B)}, \quad C_2 := \frac{\lambda_{\mathrm{B}} + 1}{2(1-\eta)}, \quad \lambda_{\mathrm{B}} := 2C_{\mathrm{B}}'(1+\kappa).$$

Proof. Once again, one checks that the proof for [25, Lem. 5.5.15] carries over to the scaled unit simplex so that

$$N_{R,\ell}^+(u) \le C_{\mathrm{B}} \left(M_{R,\ell}^+(f) + \sum_{q=1}^{\ell+1} \binom{\ell+1}{q} \left(\frac{R}{2}\right)^q \|\nabla^q A\|_{L^\infty(\widehat{T}_R^+)} \frac{[\ell-q]!}{\ell!} N_{R,\ell-q}^+(u) \right.$$
$$\left. + N_{R,\ell-1}^+(u) + N_{R,\ell-2}^+(u) \right).$$

Since this estimate has the same form as (4.10), we may conclude that

$$\frac{N_{R,\ell}^+(u)}{\ell+2} \le \left(\frac{\lambda_{\mathrm{B}}+1}{2}\right)^{\ell+1} N_{R,-1}^+(u) + C_{\mathrm{B}}' \sum_{i=0}^{\ell} \left(\frac{\lambda_{\mathrm{B}}+1}{2}\right)^{\ell-i} \frac{M_{R,i}^+(f)}{(i+2)}, \tag{4.18}$$

and, finally, the assertion follows in the same way as in the proof of Lemma 4.5. ∎

4.3.2. Control of normal derivatives. For the control of the normal derivatives, we introduce the quantity

$$N^+_{R,\ell,q}(v) := \frac{1}{[\ell+q]!} \sup_{R/2\le r<R} (R-r)^{\ell+q+2} \left\|\nabla_x^\ell \partial_y^{q+2} v\right\|_{L^2(\widehat{T}_r^+)}, \tag{4.19}$$

where, again, ∇_x denotes the gradient with respect to the tangential variables x_i with respect to Γ_R^+, $1 \le i \le d-1$, and $\partial_y = \partial_{x_n}$ denotes the derivative with respect to the normal direction.

LEMMA 4.7. *For $t \in \mathbb{N}_0$, we assume that*

$$\begin{aligned} &A \in C^{t+1}(\widehat{T}_R^+, \mathbb{R}^{d\times d}_{\mathrm{sym}}), && \operatorname{osc}(A, \widehat{T}_R^+, t+1) \le \kappa \quad \textit{for some} \quad \kappa > 0, \\ &0 < \alpha := \alpha(A, \widehat{T}_R^+), && \frac{R^{\ell+m}}{\ell! m!} \left|\nabla_x^\ell \partial_y^m A\right| \le \kappa && \forall 1 \le \ell+m \le t+1, \\ &\infty > \beta := \beta(A, \widehat{T}_R^+). \end{aligned}$$

Then for all $f \in H^t(\widehat{T}_R^+)$ and corresponding solutions u of (4.17) we have

$$\left|N^+_{R,\ell,q}(u)\right| \le C_5 K_1^\ell K_2^q \left(N^+_{R,-1}(u) + \sum_{s=0}^{\ell+q} M_{R,s}(f)\right)$$

for all $\ell \in \mathbb{N}_0$ and $q \in \mathbb{Z}_{\ge -2}$ with $\ell + q \le t$. The constants C_5, K_1, and K_2 depend only on α, β, d, λ_{B}, C'_{B}, and κ.

Proof. We assume that $f \in C^t(\overline{\widehat{T}_R^+})$ and obtain the result for general $f \in H^t(\widehat{T}_R^+)$ by a standard density argument.

In the following, ℓ, q always denote integers which satisfy $\ell \in \mathbb{N}_0$, $q = -2, -1, 0, \ldots$, and $\ell + q \le t$.

For $q = -2, -1$, the estimate $N^+_{R,\ell,q}(v) \le N^+_{R,\ell+q}(v)$ follows directly from the definitions (4.16) and (4.19). This serves as the start of an induction. We assume that the assertion is proved for all

$$(\ell, q) \in \mathcal{I}_t(q') := \{(r,s) \mid 0 \le r \le t, -2 \le s \le \min\{q', t-r-1\}\}$$

for some $-1 \le q' \le t-1$. In the induction step, we prove the result for all $(\ell, q) \in \mathcal{I}_t(q'+1)$. Taking into account the start of the induction, we may assume from now on that $\ell, q \ge 0$ and $\ell + q \le t$.

Let $\hat{A}$ denote the $d \times d$ matrix with $\hat{A}_{i,j} := A_{i,j}$ for all $1 \le i, j \le d$ with $(i,j) \ne (d,d)$ and $\hat{A}_{d,d} := 0$. Then[2]

$$-A_{d,d}\partial_y^2 u = f + \langle \operatorname{div} A, \nabla u\rangle + \hat{A} : \nabla^2 u$$

and

$$-\partial_y^2 u = \tilde{f} + \langle b, \nabla u\rangle + B : \nabla^2 u \quad \text{with} \quad \tilde{f} = f/A_{d,d},\ b = \frac{\operatorname{div} A}{A_{d,d}},\ B := A_{d,d}^{-1}\hat{A}.$$

With start with the contribution related to $\tilde{f}$. From Lemma A.2 we obtain

$$\frac{(R-r)^{\ell+q+2}}{(\ell+q)!} \left\|\nabla^{\ell+q}\tilde{f}\right\|_{L^2(\widehat{T}_r^+)} \le \frac{2}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-1}{2}} \gamma^{\ell+q} \sum_{s=0}^{\ell+q} M_{R,s}(f),$$

[2]For $d \times d$ matrices $A, B \in \mathbb{R}^{d\times d}$, we set $A : B := \sum_{i,j=1}^d A_{i,j}B_{i,j}$.

where $\gamma := \max\left\{2, \frac{8\kappa}{\alpha}\right\}$ and $\widehat{T}_r$ in the definition of $M_{R,s}(f)$ (cf. (4.7b)) has to be replaced by $\widehat{T}_r^+$.

Next, we bound the term

$$M_{\ell,q}(b,u) := \frac{1}{(\ell+q)!} \sup_{R/2\le r<R} (R-r)^{\ell+q+2} \left\|\nabla_x^\ell \partial_y^q \langle b, \nabla u\rangle\right\|_{L^2(\widehat{T}_r^+)}.$$

From [25, Lem. 5.5.18] we get

$$\begin{aligned} M_{\ell,q}(b,u) \le{}& \frac{\ell!q!}{(\ell+q)!} \sum_{r=0}^{\ell}\sum_{s=0}^{q} \frac{\left\|\partial_y^s \nabla_x^r b\right\|_{L^\infty(\widehat{T}_R^+)}}{r!s!} \left(\frac{R}{2}\right)^{r+s+1} \\ &\times \frac{[\ell-r+q-s-1]!}{(\ell-r)!\,(q-s)!} \left(N^+_{R,\ell-r,q-1-s}(u) + N^+_{R,\ell+1-r,q-s-2}(u)\right). \end{aligned}$$

The bound

$$\frac{1}{r!s!} \left\|\partial_y^s \nabla_x^r b\right\|_{L^\infty(\widehat{T}_R^+)} \left(\frac{R}{2}\right)^{r+s+1} \le \frac{C}{2}\left(\frac{\gamma}{2}\right)^{r+s} \tag{4.20}$$

is proved in Lemma A.3, where C depends on d, α, and κ. Thus

$$\begin{aligned} M_{\ell,q}(b,u) \le{}& \frac{C\ell!q!}{2(\ell+q)!} \sum_{r=0}^{\ell}\sum_{s=0}^{q} \left(\frac{\gamma}{2}\right)^{r+s} \frac{[\ell-r+q-s-1]!}{(\ell-r)!\,(q-s)!} \\ &\times \left(N^+_{R,\ell-r,q-1-s}(u) + N^+_{R,\ell+1-r,q-s-2}(u)\right). \end{aligned}$$

Finally, we consider the term $B : \nabla^2 u$. From [25, Lem. 5.5.17] we derive the estimate

$$\begin{aligned} &\frac{1}{[\ell+q]!} \sup_{R/2\le r<R} (R-r)^{\ell+q+2} \left\|\nabla_x^\ell \partial_y^q (B : \nabla^2 u)\right\|_{L^2(\widehat{T}_r^+)} \\ &\le \sum_{r=0}^{\ell}\sum_{s=0}^{q} \binom{\ell}{r}\binom{q}{s} \left\|\partial_y^s \nabla_x^r B\right\|_{L^\infty(\widehat{T}_R^+)} \left(\frac{R}{2}\right)^{r+s} \\ &\times \frac{[\ell-r+q-s]!}{[\ell+q]!} \left(N^+_{R,\ell+1-r,q-1-s}(u) + N^+_{R,\ell+2-r,q-s-2}(u)\right). \end{aligned}$$

Similarly as for the estimate (4.20) one shows that

$$\frac{1}{r!s!} \left\|\partial_y^s \nabla_x^r B\right\|_{L^\infty(\widehat{T}_R^+)} \left(\frac{R}{2}\right)^{r+s} \le C\left(\frac{\gamma}{2}\right)^{r+s} \quad \text{with} \quad C := \frac{2\kappa}{\alpha(\gamma-1)^2}\left(\frac{8}{3}\right)^{\frac{d-2}{2}} + \beta.$$

Thus

$$\begin{aligned} &\frac{1}{[\ell+q]!} \sup_{R/2\le r<R} (R-r)^{\ell+q+2} \left\|\nabla_x^\ell \partial_y^q (B : \nabla^2 u)\right\|_{L^2(\widehat{T}_R^+)} \\ &\le C\frac{\ell!q!}{[\ell+q]!} \sum_{r=0}^{\ell}\sum_{s=0}^{q} \left(\frac{\gamma}{2}\right)^{r+s} \frac{[\ell-r+q-s]!}{(\ell-r)!\,(q-s)!} \left(N^+_{R,\ell+1-r,q-1-s}(u) + N^+_{R,\ell+2-r,q-s-2}(u)\right). \end{aligned}$$

In this way, we have proved

$$
\begin{aligned}
N^{+}_{R,\ell,q}(u) \le{}& \frac{2}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-1}{2}}\left(\frac{\gamma}{2}\right)^{\ell+q}\sum_{s=0}^{\ell+q} M_{R,s}(f) \\
&+ \frac{\ell!q!}{[\ell+q]!}\frac{C}{2}\sum_{r=0}^{\ell}\sum_{s=0}^{q}\left(\frac{\gamma}{2}\right)^{r+s}\frac{[\ell-r+q-s-1]!}{(\ell-r)!\,(q-s)!} \\
&\times\left(N^{+}_{R,\ell-r,q-1-s}(u)+N^{+}_{R,\ell+1-r,q-s-2}(u)\right) \\
&+ C\frac{\ell!q!}{[\ell+q]!}\sum_{r=0}^{\ell}\sum_{s=0}^{q}\left(\frac{\gamma}{2}\right)^{r+s}\frac{[\ell-r+q-s]!}{(\ell-r)!\,(q-s)!} \\
&\times\left(N^{+}_{R,\ell+1-r,q-1-s}(u)+N^{+}_{R,\ell+2-r,q-s-2}(u)\right).
\end{aligned}
$$

To understand this recursion, we start by introducing

$$
\begin{aligned}
&C_1 = \tfrac{2}{\alpha}\left(\tfrac{8}{3}\right)^{\frac{d-1}{2}}, \qquad C_2 = \gamma/2, \\
&\tilde{N}^{+}_{\ell,q} := N^{+}_{R,\ell,q}(u), \quad \hat{C} := \max\{C_1, C\}
\end{aligned}
$$

to obtain

$$
\begin{aligned}
\tilde{N}^{+}_{\ell,q} \le{}& \hat{C}C_2^{\ell+q}\Bigg(\sum_{s=0}^{\ell+q} M_{R,s}(f) \\
&+ \frac{\ell!q!}{(\ell+q)!}\sum_{r=0}^{\ell}\sum_{s=0}^{q} C_2^{-r-s}\frac{(r+s)!}{r!s!}\left(\frac{\tilde{N}^{+}_{r,s-1}+\tilde{N}^{+}_{r+1,s-2}}{2\,[r+s]}+\tilde{N}^{+}_{r+1,s-1}+\tilde{N}^{+}_{r+2,s-2}\right)\Bigg).
\end{aligned}
$$

By using Stirling's formula and $\ell!/r! \le \ell^{\ell-r}$ we get with $c := \mathrm{e}^{1/12}$

$$
\frac{\ell!q!}{(\ell+q)!}\frac{(r+s)!}{r!s!} \le c\left(\frac{\mathrm{e}}{\ell+q}\right)^{\ell+q-r-s}\ell^{\ell-r}q^{q-s} \le c\left(\frac{\mathrm{e}\,\ell}{\ell+q}\right)^{\ell-r}\left(\frac{\mathrm{e}\,q}{\ell+q}\right)^{q-s} \le c\,\mathrm{e}^{\ell-r+q-s}.
$$

Thus,

$$
\begin{aligned}
\tilde{N}^{+}_{\ell,q} \le{}& \hat{C}C_2^{\ell+q}\sum_{s=0}^{\ell+q} M_{R,s}(f) \\
&+ c\hat{C}\sum_{r=0}^{\ell}\sum_{s=0}^{q}(C_2\,\mathrm{e})^{\ell-r+q-s}\left(\frac{\tilde{N}^{+}_{r,s-1}+\tilde{N}^{+}_{r+1,s-2}}{2}+\tilde{N}^{+}_{r+1,s-1}+\tilde{N}^{+}_{r+2,s-2}\right).
\end{aligned}
$$

By defining the quantities $N^{\mathrm{I}}_{\ell,q}$ via the recursion $N^{\mathrm{I}}_{\ell,q} := N^{+}_{R,\ell,q}(u)$ for $q = -2, -1$ and by

(4.21)

$$
N^{\mathrm{I}}_{\ell,q} = \hat{C}C_2^{\ell+q}\sum_{s=0}^{\ell+q} M_{R,s}(f) + \frac{3c\hat{C}}{2}\sum_{r=0}^{\ell}\sum_{s=0}^{q}(C_2\,\mathrm{e})^{\ell-r+q-s}\left(\tilde{N}^{\mathrm{I}}_{r+1,s-1}+\tilde{N}^{\mathrm{I}}_{r+2,s-2}\right)
$$

for $q \ge 0$, we conclude from obvious monotonicity considerations that $\tilde{N}^{+}_{\ell,q} \le N^{\mathrm{I}}_{\ell,q}$.

Note that (4.18) implies after some simple estimates that

$$
N^{+}_{R,t}(u) \le C_4C_3^{t}\left(N^{+}_{R,-1}(u)+\sum_{i=0}^{t} M^{+}_{R,i}(f)\right) \tag{4.22}
$$

with $C_3 = \mathrm{e}^{\frac{\lambda_B+1}{2}}$ and $C_4 = C_3 + C_B^{\prime}$.

We prove by induction that

$$N^{\mathrm{I}}_{\ell,q} \le C_5 K_1^{\ell} K_2^{q} \left(N^{+}_{R,-1}(u) + \sum_{s=0}^{\ell+q} M_{R,s}(f) \right),$$

where

$$C_5 \ge \max\left\{C_4, 2\hat{C}\right\}, \quad K_1 \ge \max\left\{C_3, 2C_2\,\mathrm{e}\right\}, \quad K_2 \ge K_1 \max\left\{1, 24c\hat{C}\right\}.$$

For $q = -2, -1$, we get from (4.22)

$$N^{\mathrm{I}}_{\ell,q} = N^{+}_{R,\ell,q}(u) \le N^{+}_{R,\ell+q}(u) \le C_4 C_3^{\ell+q} \left(N^{+}_{R,-1}(u) + \sum_{i=0}^{\ell+q} M^{+}_{R,i}(f) \right),$$

and the assumptions on C_5, K_1, K_2 imply the assertion.

For $q \ge 0$, we estimate the right-hand side (r.h.s.) in the recursion (4.21) by

$$\begin{aligned}
\text{r.h.s.} &\le \hat{C} C_2^{\ell+q} \sum_{s=0}^{\ell+q} M_{R,s}(f) \\
&\quad + \frac{3c\hat{C}}{2} K_1^{\ell+1} K_2^{q-1} C_5 \left(1 + \frac{K_1}{K_2}\right) \sum_{r=0}^{\infty}\sum_{s=0}^{\infty} \left(\frac{C_2\,\mathrm{e}}{K_1}\right)^{r} \left(\frac{C_2\,\mathrm{e}}{K_2}\right)^{s} \\
&\quad \times \left(N^{+}_{R,-1}(u) + \sum_{m=0}^{\ell+q} M_{R,m}(f) \right) \\
&\le C_5 K_1^{\ell} K_2^{q} \left(\left\{ 6c\hat{C} \frac{K_1}{K_2} \left(1 + \frac{K_1}{K_2}\right) \right\} N^{+}_{R,-1}(u) \right. \\
&\qquad \left. + \left\{ \frac{\hat{C}}{C_5} \left(\frac{C_2}{K_1}\right)^{\ell} \left(\frac{C_2}{K_2}\right)^{q} + 6c\hat{C} \frac{K_1}{K_2} \left(1 + \frac{K_1}{K_2}\right) \right\} \sum_{m=0}^{\ell+q} M_{R,m}(f) \right).
\end{aligned}$$

The assumptions on K_1, K_2 ensure that the expressions in the curly brackets are bounded by 1 so that the assertion follows. □

LEMMA 4.8. *For $p \in \mathbb{N}$, we assume that*

$$\begin{aligned}
&A \in C^{p}(\widehat{T}^{+}_{R}, \mathbb{R}^{d\times d}_{\mathrm{sym}}), && \operatorname{osc}\left(A, \widehat{T}^{+}_{R}, p\right) \le \kappa && \textit{for some} \quad \kappa > 0, \\
&0 < \alpha := \alpha\left(A, \widehat{T}^{+}_{R}\right), && \frac{R^{\ell+m}}{\ell! m!} \left|\nabla_x^{\ell} \partial_y^{m} A\right| \le \kappa && \forall 1 \le \ell + m \le p, \\
&\infty > \beta := \beta\left(A, \widehat{T}^{+}_{R}\right).
\end{aligned}$$

Then for all $f \in H^{p-1}\left(\widehat{T}^{+}_{R}\right)$ and corresponding solutions u of (4.17) we have

$$\frac{R^{\ell+q+1}}{[\ell+q]!} \left\| \nabla_x^{\ell} \partial_y^{q+2} u \right\|_{L^2(\widehat{T}^{+}_{\eta R})} \le C_6 C_7^{\ell+q} \left(\|\nabla u\|_{L^2(\widehat{T}^{+}_{R})} + \sum_{i=0}^{\ell+q} \frac{1}{i!} \left(\frac{R}{2}\right)^{1+i} \left\|\nabla^{i} f\right\|_{L^2(\widehat{T}_{R})} \right)$$

for all $\ell \in \mathbb{N}_0$ and $q \in \mathbb{Z}_{\ge -2}$ with $\ell + q \le p - 1$ and for all $1/2 \le \eta < 1$, where $C_6 := \frac{C_5}{(1-\eta)^2}$ and $C_7 := \frac{\max\{K_1, K_2\}}{1-\eta}$.

4.3.3. Curved boundaries. We lift the regularity estimates on the scaled unit simplex to possibly curved simplices of the finite element mesh. We explain the arguments only for the case of a simplex $K \in \mathcal{G}$ with one and only one edge, say E, on Γ. We denote the pullback to the (scaled) reference element by $F_K := R_K \circ J_K : \widehat{T}^+_{h_K} \to K$, which is chosen such that $F_K : \Gamma^+_{h_K} \to E$. The scaling of the reference simplex is chosen such that F_K and its derivatives are bounded independently of h.

From the invariance (up to multiplicative constants) of Sobolev norms under analytic coordinate transforms (see [25, Cor. 4.2.21]) we conclude that the estimates in sections 4.3.1 and 4.3.2 remain valid (with the substitutions $R \leftarrow h_K$ and $\widehat{T}^+_R \leftarrow K$)— now with multiplicative constants which depend in addition on bounds of derivatives of the pullback F_K.

5. Oscillation adapted finite elements. As an application of the new regularity estimates we derive error estimates for Galerkin *hp*-finite element discretizations of (2.2) We refer the reader to [2], [8], [14], [15], [33] for further details concerning *hp* methods.

Let $\mathcal{T}_p(A)$ be generated by Algorithm 3.9. We assume that the mesh $\mathcal{T}_h$ is a refinement of $\mathcal{T}_p(A)$ according to Definition 3.8 and satisfies (3.4) with moderate constants. Recall the definition of the subsets $\operatorname{sons}(K) \subset \mathcal{T}_h$ for $K \in \mathcal{T}_p(A)$ as in Remark 3.10(a). The *hp*-finite element space for the mesh $\mathcal{T}_h$ with polynomial degree p is given by

$$S_h^p := \left\{ u \in H_0^1(\Omega) \mid \forall K \in \mathcal{T}_h : u|_K \circ F_K \in \mathbb{P}_p \right\}, \tag{5.1a}$$

where the pullback is as in Definition 3.8(c). The Galerkin discretization of (2.2) reads as follows:

$$\text{Find } u_h^p \in S_h^p \quad \text{such that} \quad a(u_h^p, v) = F(v) \quad \forall v \in S_h^p. \tag{5.1b}$$

It is well known that the Galerkin solution exists, is unique, and satisfies the quasi-optimal error estimate in the form of Céa's lemma

$$\|u - u_h^p\|_{H^1(\Omega)} \le \frac{1}{\alpha} \inf_{v \in S_h^p} \|u - v\|_{H^1(\Omega)}. \tag{5.2}$$

To obtain explicit convergence estimates in terms of h and p one has to construct an *hp*-interpolation operator and to use regularity estimates for the solution u in combination with approximation properties.

THEOREM 5.1. *There exists an interpolation operator* $\Pi_{h,p} : H^k(\Omega) \to S_h^p$ *such that*

$$\|u - \Pi_{h,p} u\|_{H^1(K)} \le C_{\text{apx}} \left(\frac{h_K}{p}\right)^p \|u\|_{H^{p+1}(K)}$$

holds for all $K \in \mathcal{T}_h$. *The constant* C_{apx} *depends only on the constants in* (3.4) *and is independent of* p, u, K, *and the diameter* $h_K := \operatorname{diam} K$.

A construction for the interpolation operator $\Pi_{h,p}$ and the proof of the theorem can be found, e.g., in [4, Lem. 4.5], [29, Lem. 17]. The combination of the local interpolation estimates as in Theorem 5.1 with the new regularity estimates (cf. Theorem 4.1) and Céa's lemma (5.2) gives us the error estimate for the Galerkin solution.

THEOREM 5.2. *Let the assumption of Theorem* 4.1 *be satisfied. Let the hp-finite element discretization be as in* (5.1). *Then the Galerkin solution* u_h *exists, is unique,*

and satisfies the error estimate

$$\|u - u_h\|_{H^1(\Omega)} \le \frac{C_{11} C_{\text{apx}}}{c\alpha} (C_{13} h_{\text{eff}})^p \|f\|_{H^{p-1}(\Omega)}, \tag{5.3a}$$

where

$$h_{\text{eff}} := \max_{K \in \mathcal{T}_p(A)} \left\{ \left(1 + \max_{1 \le q \le p} \left(\frac{\|\nabla^q A\|_{L^\infty(K^\star)}}{q!} \right)^{1/q} \right) \max_{K' \in \text{sons}(K)} h_{K'} \right\} \tag{5.3b}$$

with $K^\star$ and c as in Lemma 3.12.

Proof. We obtain

$$\begin{aligned}
\|u - u_h\|^2_{H^1(\Omega)} &\le \frac{1}{\alpha^2} \sum_{K \in \mathcal{T}_h} \|u - \Pi_{h,p} u\|^2_{H^1(K)} \le \frac{C^2_{\text{apx}}}{\alpha^2} \sum_{K \in \mathcal{T}_h} \left(\frac{h_K}{p} \right)^{2p} \|u\|^2_{H^{p+1}(K)} \\
&= \frac{C^2_{\text{apx}}}{\alpha^2} \sum_{K \in \mathcal{T}_p(A)} \sum_{K' \in \text{sons}(K)} \left(\frac{h_{K'}}{p} \right)^{2p} \|u\|^2_{H^{p+1}(K')} \\
&\le \frac{C^2_{\text{apx}}}{\alpha^2} \sum_{K \in \mathcal{T}_p(A)} \left(\frac{\max_{K' \in \text{sons}(K)} h_{K'}}{h_K} \right)^{2p} \left(\frac{h_K}{p} \right)^{2p} \|u\|^2_{H^{p+1}(K)}.
\end{aligned}$$

Remark 3.10(b) implies that

$$\sum_{K \in \mathcal{T}_p(A)} \left(\frac{h_K}{p} \right)^{2p} \|u\|^2_{H^{p+1}(K)} \le \|u\|^2_{p+1,A}.$$

From (3.1) and (3.7b) we conclude that

$$\|u - u_h\|_{H^1(\Omega)} \le \frac{C_{\text{apx}}}{c\alpha} (C h_{\text{eff}})^p \|u\|_{p+1,A}. \qquad \square$$

COROLLARY 5.3. *Let the assumption of Theorem* 5.2 *be satisfied. Assume that the coefficient A satisfies*

$$\frac{1}{q!} \|\nabla^q A\|_{L^\infty(\Omega)} \le C \varepsilon^{-q} \tag{5.4}$$

for some (small) $\varepsilon > 0$ and for all $1 \le q \le p$. Let p and h be chosen such that

$$p = \left\lceil \frac{\log h}{\log(C_9' h/\varepsilon)} \right\rceil \quad \text{and} \quad C_9' h < \varepsilon$$

holds. Then the Galerkin discretization with the corresponding hp-finite element space S_h^p has a unique solution u_h which converges linearly:

$$\|u - u_h\|_{H^1(\Omega)} \le C h \|f\|_{H^{p-1}(\Omega)},$$

where C is independent of ε, h, and f.

Proof. The combination of (5.3) and (5.4) yields

$$\|u - u_h\|_{H^1(\Omega)} \le C \left(\frac{C_9' h}{\varepsilon} \right)^p \|f\|_{H^{p-1}(\Omega)}$$

with some constant C. By using $p = \lceil \frac{\log h}{\log(C_2^{-1} h/\varepsilon)} \rceil$ and the condition on h we obtain, after some straightforward manipulations, the assertion. □

REMARK 5.4. *The above corollary, which also covers the case of periodic coefficients, shows that for problems with oscillations of characteristic length scale ε in the coefficient, standard hp-finite element basis functions on a uniform mesh of width $h \lesssim \varepsilon$ and polynomial degree $p \gtrsim \log 1/\varepsilon$ suffice to establish an error bound proportional to ε. Note that in this case the dimension of the finite element space is of order $\varepsilon^{-d} \log^d(1/\varepsilon)$. In contrast, a conventional $P1$-finite element method requires for the same error tolerance a mesh of width $h \lesssim \varepsilon^2$ so that the dimension of the $P1$-finite element space is much larger, more precisely of order $(\varepsilon)^{-2d}$.*

6. Numerical experiments. This paper focuses on the derivation of a priori error bounds for elliptic problems with an oscillating diffusion coefficient. The bounds are explicit with respect to the scales in the problem which originate from local oscillations of the diffusion coefficient. Two numerical experiments shall illuminate these results with the help of simplest (one-dimensional) model problems. A comprehensive numerical study of the algorithmic ideas proposed in this paper is the topic of current and future research.

6.1. Illustration of Corollary 5.3. Consider the one-dimensional model of heat diffusion in a periodic laminate. Let $\Omega := (0,1)$, $f = 1$, and let the diffusion coefficient A_ε be given by

$$A_\varepsilon(x) := \left(2 - \sin\left(2\pi \frac{x}{\varepsilon}\right)\right)^{-1}$$

for some small parameter $\varepsilon > 0$. We are seeking $u_\varepsilon \in H_0^1(\Omega)$ such that

$$\int_0^1 A_\varepsilon u_\varepsilon' v' \,\mathrm{d}x = \int_0^1 f v \,\mathrm{d}x \quad \forall v \in H_0^1(\Omega). \tag{6.1}$$

Since the oscillations of the coefficient are periodic with two separated characteristic length scales, 1 and ε, advanced techniques from homogenization and numerical homogenization can be applied, while the focus of the oscillation adaptive finite elements is on problems with nonuniformly distributed scales. Nevertheless, this example nicely illustrates the sharpness of the convergence estimates in Corollary 5.3.

Given $h \in \{2^{-k} \mid k \in \mathbb{N}_0\}$, let $\mathcal{T}_h$ denote the uniform subdivision into intervals of equal length h, i.e., $\mathcal{T}_h := \{[(j-1)h, jh] \mid j = 1, 2, \ldots, h^{-1}\}$. The Galerkin approximation of u_ε based on $\mathcal{T}_h$ and some uniform polynomial degree $p \in \mathbb{N}$ is denoted by $u_{\varepsilon,h,p} = u_{\varepsilon,h(\varepsilon),p(\varepsilon)} \in S_h^p$.

Consider test cases characterized by the parameters

$$\varepsilon \in \mathcal{I} := \left\{1, \tfrac{1}{2}, \tfrac{1}{4}, \tfrac{1}{8}, \tfrac{1}{16}, \ldots\right\}.$$

The corresponding solutions u_ε, $\varepsilon \in \mathcal{I}$, oscillate at the same frequency ε^{-1} as the coefficient A_ε; the amplitude of the oscillation is of order ε. In order to capture this behavior by the standard h-version of the finite element method, a mesh of size ε or smaller is required. Algorithm 3.2 with $Q_0 = [0,1]$ produces a uniform subdivision $\mathcal{Q}_{p,A_\varepsilon}$ into intervals of equal length ε, i.e., $\mathcal{Q}_{p,A_\varepsilon} = \mathcal{T}_\varepsilon$. Moreover, since $\|\nabla^\ell A_\varepsilon\|_{L^\infty(0,1)} \le \varepsilon^{-\ell}/\ell!$, the elements of $\mathcal{T}_\varepsilon$ satisfy the oscillation condition for all $p \in \mathbb{N}$.

Given $\varepsilon \in \mathcal{I}$, we compare three different choices of h and p:

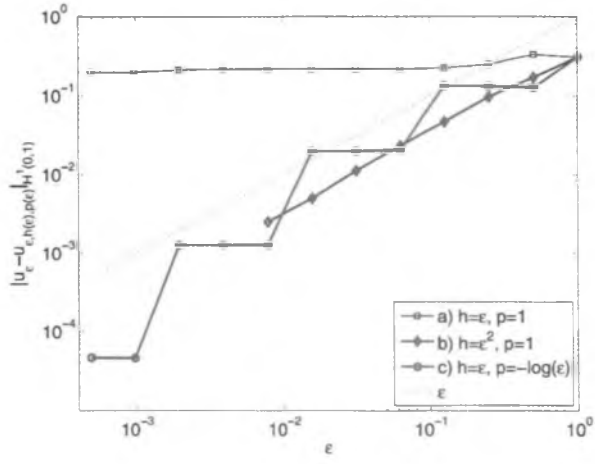

FIG. 6.1. *Example from section 6.1: H^1-error of $h(\varepsilon)p(\varepsilon)$-finite element method vs. ε.*

(a) $h = \varepsilon$ and $p = 1$,
(b) $h = \varepsilon^2$ and $p = 1$, and
(c) $h = \varepsilon$ and $p = \lceil |\log \varepsilon| \rceil$.

Figure 6.1 shows the error $\|u - u_{\varepsilon,h,p}\|_{H^1(\Omega)}$ as a function of ε for the three approaches. Choice (a) does not provide any accuracy; the error equals some constant of order 1 independent of ε as predicted by (1.1). However, the computational cost grows like ε^{-1} as ε tends to zero. The error of choice (b) behaves like ε, but the computational cost to achieve this accuracy scales like ε^{-2} (the numerical results for very small values of ε are missing because of this high cost). Figure 6.1 clearly indicates that choice (c) is superior to the other two. With a computational cost that is proportional to $\varepsilon^{-1}\log(\varepsilon^{-1})$, i.e., with approximately $\log(\varepsilon^{-1})$ degrees of freedom per vertex in the finite element mesh, the error decreases linearly in ε. Hence, this choice recovers textbook convergence up to some logarithmic factor. This observation is in agreement with Corollary 5.3.

6.2. Nonperiodic diffusion. After distortion the material from the previous experiment may no longer be periodic; it may even contain a whole continuum of scales. In this context, a possible diffusion coefficient A after distortion reads as

$$A(x) := \big(2 - \sin(2\pi \tan(\pi(1 - 2^{-4})x/2))\big)^{-1}. \tag{6.2}$$

Again, we are seeking $u \in H_0^1(\Omega)$ such that

$$\int_0^1 Au'v' \,\mathrm{d}x = \int_0^1 fv \,\mathrm{d}x \quad \forall v \in H_0^1(\Omega). \tag{6.3}$$

Figure 6.2 depicts both the coefficient A and the corresponding solution u. In contrast to the periodic case, the oscillations vary strongly in space.

The above example allows one to investigate the finite element method with regard to its adaptation of variations in the diffusion coefficient in space (h-adaptivity). We compare the performance of the finite element method of (uniform, fixed) orders $p = 1, 2, 3$ for two different choices of the initial mesh:

(a) the naive coarse uniform subdivision $\mathcal{T}_{\frac{1}{2}}$, and

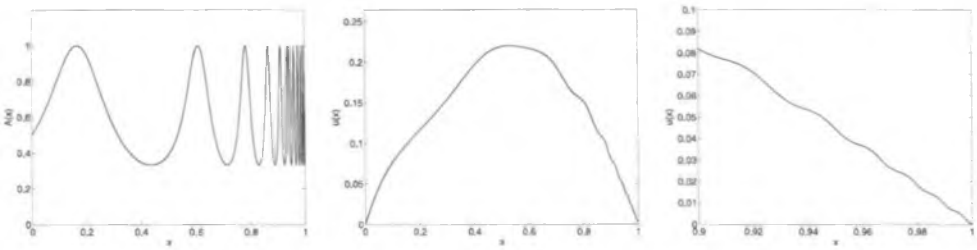

FIG. 6.2. *Diffusion coefficient from* (6.2) *(left) and corresponding solution of* (6.3) *(middle) with zoom into the right end of the interval (right).*

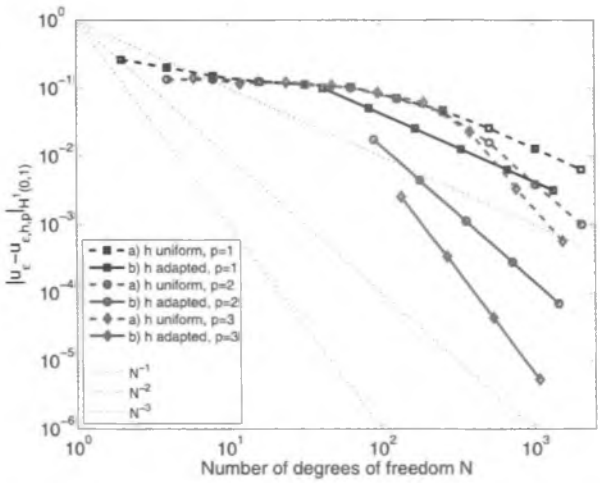

FIG. 6.3. *Example from section* 6.2*: H^1-error of $h(\varepsilon)p(\varepsilon)$-finite element method vs. number of degrees of freedom N.*

(b) the oscillation adapted mesh produced by Algorithm 3.2.

Starting from one of these initial meshes, Figure 6.3 shows the evolution of the $H_0^1(0,1)$-errors under uniform mesh refinement (bisection of all intervals). Asymptotically, both choices show the expected optimal rate of convergence with respect to the number of unknowns N; i.e., the error is proportional to N^{-p} in this one-dimensional setting. However, the asymptotic regime is not reached unless all scales are resolved by the underlying mesh. In this respect, the oscillation adapted mesh (b) is superior to the naive approach (a) by orders of magnitude.

7. Conclusions. In this paper, we have derived new regularity results for elliptic problems with smooth diffusion matrices having possibly highly oscillatory and nonuniform behavior on different scales. This generalizes the regularity theory for periodic problems with a highly oscillatory diffusion matrix on the periodic cells in the domain.

Based on this theory we have introduced *hp*-finite element spaces, where the distribution of the mesh cells and the local polynomial degrees are adapted to the local growth of derivatives of the diffusion matrix. We have introduced an algorithm for generating such types of meshes where the refinement criterion is based on the oscillations in the diffusion matrix. Numerical results show that this approach leads to a significantly improved accuracy in the preasymptotic regime compared to standard

finite elements.

In the context of a posteriori error estimation we emphasize that for many singularly perturbed or parameter-dependent problems such as, e.g., convection dominated problems, highly indefinite scattering problems, high-frequency eigenvalue problems, etc., the condition "the mesh width has to be *sufficiently* small" typically arises (see, e.g., [1], [10], [6], [32], [24], [26]). In this light, the generation of optimal initial meshes is also of utmost importance for adaptive finite element methods for these kinds of problems. Future research will be directed to the generation of a priori adapted meshes for singularly perturbed problems with coefficients which vary over a large range of different scales.

In many applications the diffusion matrix is given by measurement data in a discrete way. Therefore our algorithm for the generation of an oscillation adapted finite element mesh is not directly applicable. However, if it is known that the measured data is expected to approximate a (piecewise) smooth, possibly oscillatory coefficient, one could extract information on derivatives of A in a preprocessing step by filtering techniques or interpolation.

Appendix A. Derivatives of composite functions and of products of functions.

LEMMA A.1. *Let $\omega \subset \mathbb{R}^d$ be a domain, and let $a \in C^{t+1}(\omega)$, which satisfies* $\operatorname{osc}(a, \omega, t+1) \leq \kappa$ *and*

$$\forall x \in \omega \quad 0 < \alpha \leq a(x) \leq \beta < \infty.$$

Then $a^{-1} \in C^{t+1}(\omega)$ and satisfies, for $R := \operatorname{diam} \omega$,

$$\frac{R^{t+1}}{(t+1)!} \left|\nabla^{t+1} a^{-1}(x)\right| \leq \frac{2}{\alpha} \left(\frac{8}{3}\right)^{\frac{d-1}{2}} \gamma^{t+1} \tag{A.1}$$

with $\gamma := \max\left\{2, \frac{8\kappa}{\alpha}\right\}$.

Proof. The existence of a^{-1} and $a \in C^{t+1}(\omega)$ follows readily from the generalized di Bruno formula (cf. [20, Thm. 4.2]).

For $x \in \omega$ fixed and all $y \in \mathbb{C}^d$, let

$$\tilde{a}(y) := \sum_{\ell=0}^{t+1} \frac{1}{\ell!} \langle y - x, \nabla \rangle^{\ell} a(x) \quad \text{with} \quad \frac{1}{\ell!} \langle y - x, \nabla \rangle^{\ell} a := \sum_{\substack{\mu \in \mathbb{N}_0^d \\ |\mu| = \ell}} \frac{(y-x)^{\mu}}{\mu!} \partial^{\mu} a(x).$$

Also from [20, Thm. 4.2], it follows that $\partial^{\mu} a^{-1}(x)$ depends only on $\partial^{\nu} a(x)$ for $\nu_i \leq \mu_i$, $1 \leq i \leq d$. Hence, by choosing $y = x$, we obtain

$$\left(\tilde{a}^{-1}\right)^{(\mu)}(x) = \left(a^{-1}\right)^{(\mu)}(x) \quad \forall \mu \in \mathbb{N}_0^d, \quad |\mu| = t+1.$$

Since $\tilde{a}$ is analytic, we may apply Cauchy's integral formula to estimate the derivatives of $\tilde{a}$:

$$\frac{1}{\mu!} \left(\tilde{a}^{-1}\right)^{(\mu)}(x) = \frac{1}{(2\pi i)^d} \oint_{C_r(x_1)} \oint_{C_r(x_2)} \cdots \oint_{C_r(x_d)} \frac{\tilde{a}^{-1}(v)}{(v-x)^{\mu+\mathbf{1}}} dv \tag{A.2}$$

with $\mathbf{1} = (1, 1, \ldots, 1)^{\mathsf{T}}$ and $C_r(x_i)$ being the circle in $\mathbb{C}$ about x_i with radius $r > 0$. The denominator satisfies $\left|(v-x)^{\mu+\mathbf{1}}\right| = r^{|\mu|+d}$ so that

$$\left|\frac{1}{\mu!} \left(\tilde{a}^{-1}\right)^{(\mu)}(x)\right| \leq r^{-|\mu|} \sup\left\{\left|\tilde{a}^{-1}(v)\right| : v \in \mathbb{C}^d \mid \forall 1 \leq i \leq d \quad v_i \in C_r(x_i)\right\}.$$

The assumptions on a imply that

$$\left|\tilde{a}^{-1}(v)\right| = \frac{1}{a(x) + \langle v - x, \nabla\tilde{a}(\xi)\rangle}$$

for some $\xi \in \overline{vx}$. We set $e := \frac{v-x}{\|v-x\|}$ and obtain

$$\begin{aligned}|\langle v - x, \nabla\tilde{a}(\xi)\rangle| &= \left|\sum_{\ell=0}^{t} \frac{1}{\ell!} \langle \xi - x, \nabla\rangle^{\ell} \langle v - x, \nabla\rangle a(x)\right| \\ &\leq \sum_{\ell=0}^{t} \frac{r^{\ell+1}}{\ell!} \left|\langle e, \nabla\rangle^{\ell+1} a(x)\right|.\end{aligned}$$

Some tedious calculus leads to

$$\begin{aligned}\frac{1}{\ell!}\left|\langle e, \nabla\rangle^{\ell+1} a(x)\right| &= (\ell+1) \sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|=\ell+1}} \frac{e^{\mu}}{\mu!}\partial^{\mu} a(x) \\ &= \frac{1}{\ell!}\sqrt{\sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|=\ell+1}} \frac{(\ell+1)!}{\mu!} e^{2\mu}} \sqrt{\sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|=\ell+1}} \frac{(\ell+1)!}{\mu!} \left|\partial^{\mu} a(x)\right|^2} \\ &= \frac{1}{\ell!}\sqrt{\left(\sum_{i=1}^{d} e_i^2\right)^{\ell+1}} \left|\nabla^{\ell+1} a(x)\right| = \frac{1}{\ell!}\left|\nabla^{\ell+1} a(x)\right|.\end{aligned}$$

Thus, with $R = \operatorname{diam}\omega$, by choosing $r = cR$ in (A.2) for some $0 < c < 1$ and by using the oscillation condition we obtain

$$|\langle v - x, \nabla\tilde{a}(\xi)\rangle| \leq \kappa \sum_{\ell=0}^{t} (\ell+1) c^{\ell+1} \leq \kappa c \sum_{\ell=0}^{\infty} (\ell+1) c^{\ell} = \frac{c\kappa}{(1-c)^2}.$$

By setting $c = \gamma^{-1}$ (cf. (A.1)) we get $|\langle v - x, \nabla\tilde{a}(\xi)\rangle| \leq \frac{\alpha}{2}$ so that $\left|\tilde{a}^{-1}(v)\right| \leq \frac{2}{\alpha}$. Hence,

$$\frac{R^{|\mu|}}{\mu!}\left|\left(\tilde{a}^{-1}\right)^{(\mu)}(x)\right| \leq \frac{2}{\alpha c^{|\mu|}}. \tag{A.3}$$

A summation over all $\mu \in \mathbb{N}_0^d$ with $|\mu| = t+1$ leads to

$$\begin{aligned}\frac{R^{t+1}}{(t+1)!}\left|\nabla^{t+1} a^{-1}(x)\right| &= \frac{1}{(t+1)!}\sqrt{\sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|=t+1}} \frac{(t+1)!}{\mu!}\left|R^{|\mu|}\partial^{\mu}\tilde{a}^{(\mu)}(x)\right|^2} \\ &\leq \frac{2}{\alpha}\frac{1}{\sqrt{(t+1)!}c^{t+1}}\sqrt{\sum_{\substack{\mu\in\mathbb{N}_0^d \\ |\mu|=t+1}} \mu!} \overset{\text{Lemma A.4}}{\leq} \frac{2}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-1}{2}} c^{-t-1}. \qquad \square\end{aligned}$$

LEMMA A.2. *Let $\omega \subset \mathbb{R}^d$ be a domain, and let $a \in C^{t+1}(\omega)$ satisfy the assumptions of Lemma A.1. Then, for $f \in C^{t+1}(\omega)$, it holds that $\tilde{f} := f/a \in C^{t+1}(\omega)$ and $\tilde{f}$*

satisfies, for $R := \operatorname{diam}\omega$ *and* $1 \le \ell \le t+1$,

$$\frac{R^\ell}{\ell!}\left|\nabla^\ell \tilde{f}(x)\right| \le \frac{2}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-1}{2}} \gamma^\ell \sum_{q=0}^{\ell} \frac{R^q\left|\nabla^q f(x)\right|}{q!}$$

with γ *as in Lemma* A.1.

Proof. From [25, Lem. A.1.3] we conclude that

$$\frac{R^\ell}{\ell!}\left|\nabla^\ell \tilde{f}(x)\right| \le \sum_{q=0}^{\ell} \frac{R^q\left|\nabla^q f(x)\right|}{q!}\frac{R^{\ell-q}\left|\nabla^{\ell-q} a^{-1}(x)\right|}{(\ell-q)!}.$$

By using Lemma A.1 we get

$$\frac{R^\ell}{\ell!}\left|\nabla^\ell \tilde{f}(x)\right| \le \frac{2}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-1}{2}} \gamma^\ell \sum_{q=0}^{\ell} \frac{R^q\left|\nabla^q f(x)\right|}{q!}. \qquad \square$$

LEMMA A.3. *Let* $\omega \subset \mathbb{R}^d$ *be a domain, and let* $A \in C^{t+1}(\omega, \mathbb{R}^{d\times d}_{\mathrm{sym}})$ *be such that* $0 < \alpha\left(A, \widehat{T}_R^+\right) =: \alpha$ *and* $\beta := \beta\left(A, \widehat{T}_R^+\right) < \infty$. *For the oscillations we assume that* $\operatorname{osc}(a, \omega, t+1) \le \kappa$ *and*

$$\frac{R^{\ell+m}}{\ell!m!}\left|\nabla_x^\ell \partial_y^m A\right| \le \kappa \quad \forall 1 \le \ell+m \le t+1.$$

Then, for $b = \frac{\operatorname{div} A}{A_{d,d}}$, *it holds that* $b \in C^t(\omega, \mathbb{R}^d)$ *and* b *satisfies, for* $R := \operatorname{diam}\omega$ *and* $1 \le \ell+m \le t$,

$$\frac{R^{\ell+m+1}}{\ell!m!}\left|\nabla_x^\ell \partial_y^m b(x)\right| \le C\gamma^{\ell+m} \quad \textit{with} \quad C := 4\frac{\sqrt{d}\kappa}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-2}{2}}\left(\frac{\gamma}{\gamma-1}\right)^3.$$

Proof. For $1 \le \ell+m \le t$, it holds that

(A.4)
$$\frac{R^{\ell+m+1}}{\ell!m!}\left|\nabla_x^\ell \partial_y^m \frac{\operatorname{div} A}{A_{d,d}}\right| \le \sum_{r=0}^{\ell}\sum_{s=0}^{m} \frac{R^{r+s+1}\left|\nabla_x^r \partial_y^s \operatorname{div} A\right|}{r!s!}\frac{R^{\ell-r+m-s}\left|\nabla_x^{\ell-r}\partial_y^{m-s} A_{d,d}^{-1}\right|}{(\ell-r)!(m-s)!}.$$

Next, observe that

$$\begin{aligned}\frac{R^{r+s+1}}{r!s!}\left|\nabla_x^r \partial_y^s \operatorname{div} A\right| &\le \sqrt{d}\left((r+1)\frac{R^{r+s+1}}{(r+1)!s!}\left|\nabla_x^{r+1}\partial_y^s A\right|\right.\\ &\qquad \left. + (s+1)\frac{R^{r+s+1}}{r!(s+1)!}\left|\nabla_x^r \partial_y^{s+1} A\right|\right)\\ &\le \sqrt{d}(r+s+2)\kappa \qquad \text{(A.5a)}\end{aligned}$$

and

$$\begin{aligned}\frac{R^{r+s}}{r!s!}\left|\nabla_x^r \partial_y^s A_{d,d}^{-1}\right| &= \frac{1}{\sqrt{r!s!}}\sqrt{\sum_{\substack{\mu\in\mathbb{N}_0^{d-1}\\|\mu|=r}} \frac{1}{\mu!s!}\left|R^{r+s}\partial_x^\mu \partial_y^s A_{d,d}^{-1}\right|^2} \overset{\text{(A.3)}}{\le} \frac{2\gamma^{r+s}}{\alpha}\frac{1}{\sqrt{r!}}\sqrt{\sum_{\substack{\mu\in\mathbb{N}_0^{d-1}\\|\mu|=r}} \mu!}\\ &\le \left(\frac{8}{3}\right)^{\frac{d-2}{2}}\frac{2\gamma^{r+s}}{\alpha}. \qquad \text{(A.5b)}\end{aligned}$$

Inserting (A.5) into (A.4) leads to

$$\begin{aligned}\frac{R^{\ell+m+1}}{\ell!m!}\left|\nabla_x^\ell \partial_y^m \frac{\operatorname{div} A}{A_{dd}}\right| &\le 2\frac{\sqrt{d}\kappa}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-2}{2}} \sum_{r=0}^{\ell}\sum_{s=0}^{m}(r+s+2)\,\gamma^{\ell-r+m-s}\\ &\le 2\frac{\sqrt{d}\kappa}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-2}{2}} \gamma^{\ell+m}\sum_{r=0}^{\infty}\sum_{s=0}^{\infty}(r+s+2)\,\gamma^{-r-s}\\ &\le 4\frac{\sqrt{d}\kappa}{\alpha}\left(\frac{8}{3}\right)^{\frac{d-2}{2}} \frac{\gamma^{\ell+m+3}}{(\gamma-1)^3}. \qquad \square\end{aligned}$$

LEMMA A.4. *It holds that*

$$\sum_{\substack{\mu\in\mathbb{N}_0^d\\ |\mu|=\ell}} \mu! \le \ell!\left(\frac{8}{3}\right)^{d-1}. \tag{A.6}$$

Proof. Let

$$\sigma_d(\ell) := \begin{cases} \ell!, & d=1,\\ \sum_{i=0}^{\ell} i!\sigma_{d-1}(\ell-i), & d>1,\end{cases}$$

and observe that the left-hand side in (A.6) equals $\sigma_d(\ell)$. We prove the result by induction over d.

The case $d=1$ is trivial. For $d=2$, we employ the notation as in (cf. [17, eq. (2.5)])

$$(-\ell)_i := (-1)^i \frac{\Gamma(1+\ell)}{\Gamma(1+\ell-i)} \quad \forall \ell\in\mathbb{N} \text{ and } \forall 0\le i\le \ell$$

to obtain

$$\begin{aligned}\frac{1}{\ell!}\sum_{i=0}^{\ell} i!(\ell-i)! &= \sum_{i=0}^{\ell}(-1)^i\frac{i!}{(-\ell)_i} \overset{[17,\,(7.2.4)]}{=} \frac{\ell+1}{2^{\ell+1}}\sum_{k=1}^{\ell+1}\frac{2^k}{k}\\ &= \frac{\ell+1}{2^{\ell+1}}\left(-\mathrm{i}\,\pi + B_2(2+\ell,0)\right) =: \psi(\ell),\end{aligned}$$

where $B_z(a,b)$ is the incomplete beta function. The function $\psi(x)$ is continuous for all $x\ge 0$ and satisfies

$$\psi(0)=1, \quad \psi(\infty)=1.$$

Hence, there exists C such that, for all $\ell\ge 0$, it holds that $\psi(\ell)\le C$. Numerical tests show that the maximum of ψ is attained for $\psi(4)=8/3$ so that $\sigma_2(\ell)\le \frac{8}{3}\ell!$.

Assume that the assertion holds for $d-1$. Then, the recursion formula gives

$$\sigma_d(\ell) \le \left(\frac{8}{3}\right)^{d-2}\sum_{i=0}^{\ell} i!(\ell-i)! \le \left(\frac{8}{3}\right)^{d-1}\ell!. \qquad \square$$

REFERENCES

[1] M. Ainsworth and W. Dörfler, *Fundamental systems of numerical schemes for linear convection-diffusion equations and their relationship to accuracy*, Computing, 66 (2001), pp. 199–229.

[2] I. Babuška and B. Q. Guo, *Approximation properties of the h-p version of the finite element method*, Comput. Methods Appl. Mech. Engrg., 133 (1996), pp. 319–346.

[3] 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.

[4] I. Babuška and M. Suri, *The h-p version of the finite element method with quasi-uniform meshes*, RAIRO Modél. Math. Anal. Numér., 21 (1987), pp. 199–238.

[5] C. Băcuţă, A. L. Mazzucato, V. Nistor, and L. Zikatanov, *Interface and mixed boundary value problems on n-dimensional polyhedral domains*, Doc. Math., 15 (2010), pp. 687–745.

[6] C. Carstensen and J. Gedicke, *An oscillation-free adaptive FEM for symmetric eigenvalue problems*, Numer. Math., 118 (2011), pp. 401–427.

[7] P. Ciarlet, *The Finite Element Method for Elliptic Problems*, North–Holland. Amsterdam, 1987.

[8] L. Demkowicz, *Polynomial exact sequences and projection-based interpolation with applications to Maxwells equations*, in Mixed Finite Elements, Compatibility Conditions, and Applications, D. Boffi, F. Brezzi, L. Demkowicz, L. Durán, R. Falk, and M. Fortin, eds., Lecture Notes in Math. 1939, Springer-Verlag, Berlin, 2008.

[9] L. C. Evans, *Partial Differential Equations*, 2nd ed., American Mathematical Society. Providence, RI, 2010.

[10] S. Giani and I. G. Graham, *A convergent adaptive method for elliptic eigenvalue problems*, SIAM J. Numer. Anal., 47 (2009), pp. 1067–1091.

[11] D. Gilbarg and N. Trudinger, *Elliptic Partial Differential Equations of Second Order*, Springer-Verlag, Berlin, 1983.

[12] L. Grasedyck, I. Greff, and S. Sauter, *The AL basis for the solution of elliptic problems in heterogeneous media*, Multiscale Model. Simul., 10 (2012), pp. 245–258.

[13] P. Grisvard, *Elliptic Problems in Nonsmooth Domains*, Pitman, Boston, MA, 1985.

[14] B. Guo and E. P. Stephan, *The h-p version of the coupling of finite element and boundary element methods for transmission problems in polyhedral domains*. Numer. Math., 80 (1998), pp. 87–107.

[15] B. Q. Guo and I. Babuška, *Countable normed spaces and the h-p version of the finite element method*, in Numerical and Applied Mathematics, Part II (Paris, 1988), IMACS Ann. Comput. Appl. Math. 1.2. Baltzer, Basel, 1989, pp. 525–530.

[16] B. Q. Guo and H. S. Oh, *The h-p version of the finite element method for problems with interfaces*, Internat. J. Numer. Methods Engrg., 37 (1994), pp. 1741–1762.

[17] E. Hansen, *A Table of Series and Products*, Prentice–Hall, Englewood Cliffs, NJ, 1975.

[18] R. B. Kellogg, *Singularities in interface problems*, in Numerical Solution of Partial Differential Equations, II (SYNSPADE 1970) (Proc. Sympos., Univ. of Maryland, College Park, MD, 1970), Academic Press, New York, 1971, pp. 351–400.

[19] V. A. Kondratev, *Boundary value problems for elliptic equations in domains with conical or angular points*, Trudy Moskov. Mat. Obšč., 16 (1967), pp. 209–292.

[20] R. B. Leipnik and C. E. M. Pearce, *The multivariate Faà di Bruno formula and multivariate Taylor expansions with explicit integral remainder term*, ANZIAM J., 48 (2007), pp. 327–341.

[21] K. Lemrabet, *Régularité de la solution d'un problème de transmission*, J. Math. Pures Appl. (9), 56 (1977), pp. 1–38.

[22] A. Malqvist and D. Peterseim, *Localization of elliptic multiscale problems*, ArXiv e-prints, 2011.

[23] A. Matache and C. Schwab, *Two-scale FEM for homogenization problems*, M2AN Math. Model. Numer. Anal., 36 (2002), pp. 537–572.

[24] J. Melenk and S. Sauter, *Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions*, Math. Comp., 79 (2010), pp. 1871–1914.

[25] J. M. Melenk, *hp-Finite Element Methods for Singular Perturbations*, Lecture Notes in Math. 1796, Springer-Verlag, Berlin, 2002.

[26] J. M. Melenk and S. Sauter, *Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1210–1243.

[27] R. C. Morgan and I. Babuška, *An approach for constructing families of homogenized equa-*

tions for periodic media. I: *An integral representation and its consequences*, SIAM J. Math. Anal., 22 (1991), pp. 1–15.

[28] C. B. Morrey, Jr., *Multiple Integrals in the Calculus of Variations*, Die Grundlehren der mathematischen Wissenschaften, Band 130, Springer-Verlag, New York, 1966.

[29] R. Muñoz-Sola, *Polynomial liftings on a tetrahedron and applications to the h-p version of the finite element method in three dimensions*, SIAM J. Numer. Anal., 34 (1997), pp. 282–314.

[30] S. Nicaise, *Singularities in interface problems*, in Problems and Methods in Mathematical Physics (Chemnitz, 1993), Teubner-Texte Math. 134, Teubner, Stuttgart, 1994, pp. 130–137.

[31] H. Owhadi and L. Zhang, *Localized bases for finite-dimensional homogenization approximations with nonseparated scales and high contrast*, Multiscale Model. Simul., 9 (2011), pp. 1373–1398.

[32] S. Sauter and W. Dörfler, *A Posteriori Error Estimation for Highly Indefinite Helmholtz Problems*, Preprint 13-2011, Institut für Mathematik, Universität Zürich. Zürich, Switzerland, 2011.

[33] C. Schwab, *p- and hp-Finite Element Methods. Theory and Applications in Solid and Fluid Mechanics*, Numer. Math. Sci. Comput., The Clarendon Press, Oxford University Press, New York, 1998.