J. Numer. Math., Vol. 0, No. 0, pp. 1–29 (2016)
DOI 10.1515/ JNUM.2016.000
 Prepared using jnm.sty *[Version: 20.02.2007 v1.3]*

Duality-based adaptivity in finite element discretization of heterogeneous multiscale problems

Matthias Maier,* Rolf Rannacher*†

Received

Abstract — This paper introduces an framework for adaptivity for a class of heterogeneous multiscale finite element methods for elliptic problems, which is suitable for a posteriori error estimation with separated quantification of the model error as well as the macroscopic and microscopic discretization errors. The method is derived within a general framework for "goal-oriented" adaptivity, the so-called Dual Weighted Residual (DWR) method. This allows for a systematic a posteriori balancing of multiscale modeling and discretization. The developed method is tested numerically at elliptic diffusion problems for different types of heterogeneous oscillatory coefficients.

Keywords: heterogeneous multiscale method, finite element method, mesh adaptation, model adaptation, goal-oriented adaptivity, DWR method

AMS(MOS) subject classification: 35J15 65C20 65N12 65N15 65N30 65N50

1. Introduction

A large class of modeling problems in physics and engineering is of *multiscale character*, meaning that relevant physical processes may act on largely different scales. This usually results in unacceptably high computational cost for a full resolution of all scales. One way to avoid this dilemma are *multiscale techniques*, where, generally speaking, an effective model is solved on a coarse scale with upscaled effective parameters that are determined with the help of localized (possibly coupled) sampling problems on a fine scale.

The usage of upscaling principles in numerical methods can be traced back as early as the 70ies. We exemplarily mention a paper by Hill [33], in which upscaling principles for effective parameters were formulated in the context of elasticity problems. Since then, multiscale methods have become increasingly popular in the engineering community where they are usually referred to as *computational homogenization* schemes (see Geers et al. [30] for an overview).

Different approaches for modeling multiscale phenomena in the context of finite element methods exist and have lead to a number of methods introduced over the last

*Institute of Applied Mathematics, Heidelberg University, Im Neuenheimer Feld 293/294, D-69120 Heidelberg, Germany.

†Corresponding author

years. Most of them either rely on the existence of a periodic or stochastic substructure or on the scale-dependent splitting of variational solution and test spaces. Most notable are the *variational multiscale method* developed by Hughes et al. [35] and Brezzi [19], the *mixed multiscale methods* by Arbogast & Boyd [7] or Chen & Hou [22], the *two-scale* or *generalized finite element method* by Matache & Schwab [37,38], or the *multiscale finite element method* introduced by Hou & Wu [34], Efendiev et al. [29], and variants of these approaches. A mathematically rigorous formulation in the context of finite element theory and homogenization theory was given by E & Engquist [26,27] with the description of the so-called *Heterogeneous Multiscale Method* (HMM). In its original setting the HMM relies on a periodic substructure of the coefficient matrix and can be viewed as a direct discretization of the underlying homogenization process (cf. Babuska [8] and Ohlberger [43]).

Multiscale schemes introduce significant complexity with respect to sources of error, not only are there discretization errors on a coarse and fine scales, but also a model error introduced by modeling assumptions. This makes a priori knowledge and/or suitable a posteriori strategies highly necessary. For the HMM quite a number of theoretical results are available: An *a priori* error estimation for the HMM dealing with the discretization errors was first presented by E & Engquist [26,27], E et al. [28], and later, with improved results, by Abdulle [1] and Abdulle & Vilmart [4] (for a nonlinear case) – without an estimation of the underlying modeling error. *A posteriori* error estimation for the discretization errors was later presented by Ohlberger [43], Henning & Ohlberger [31], Henning et al. [32], and Abdulle [2]; corresponding *goal-oriented* error estimation results (for discretization errors) were formulated by Abdulle & Nonnenmacher [3]. First results for estimating and controlling the *model error* in the context of multiscale schemes were derived by Oden & Vemaganti [39,41] and Braack & Ern [16]. The paper by Romkes & Moody [45] deals with the localization of the dual problems involved.

This paper presents a framework for combined model and discretization adaptivity based on the *Dual Weighted Residual* (DWR) method by Becker & Rannacher [12,13,14] (see also Bangerth & Rannacher [10]) for a class of well-known, rather simple, multiscale schemes. At its heart, independent error indicators for macroscale and microscale discretization as well as the model error are derived and an algorithm is formulated for simultaneous control and balancing of all error sources. Here, the major novelty lies in the simultaneous treatment of discretization and model errors within a unified framework for adaptivity and, in particular, in the introduction of model adaptation in the context of the sampling procedures by tuning the underlying "sampling meshes". This adaptive framework is fairly general in the sense that, in principle, it can be applied to a variety of different multiscale schemes (as long as a complete microscale description is available).

For the sake of simplicity, we restrict the discussion to the usual elliptic model problem with heterogeneous oscillatory coefficient. Because the main subject of the paper is the introduction of an a posteriori framework for combined model and discretization adaptivity, we further restrict the discussion to the class of effective models produced by local averaging strategies using geometric or harmonic mean

values. This is justified by the fact that, for the purpose of this paper, a computationally inexpensive reconstruction strategy is desired—the *efficiency* of the framework will be based on a posteriori error control of the underlying model and discretization parameters. The special case of classical homogenization strategies, such as the HMM, is discussed in remarks. The application of the proposed a posteriori framework to more sophisticated averaging strategies will be the subject of a forthcoming publication.

The outline of this paper is as follows. In Section 2, we introduce on an abstract level our multiscale formulation, which is especially designed to allow for local model adaptation, and prove well-posedness. Then, in Section 3, we give an *a posteriori* error analysis of this scheme, in which the effects of modeling errors are separated from macroscopic and microscopic discretization errors. These error estimates are used for developing a strategy for simultaneously adapting modeling and discretization. Finally, in Section 4 the effectivity of the developed algorithm is demonstrated at prototypical test problems.

2. An abstract multiscale scheme for model adaptation

In this section, we present a multiscale scheme for model adaptation that explicitly decouples all discretization and modeling parameters. It is a reformulation of the classical HMM method by E & Engquist [26,27] and has some similarities with the model adaptation framework introduced by Oden & Vemaganti [39,40,41] and Braack & Ern [16]. The novelty lies in the explicit decoupling of the sampling processes from the macroscopic discretization, as well as the simultaneous treatment of discretization and modeling parameters.

Let us consider the following multi-scale model problem: Find $u^\varepsilon \in H_0^1(\Omega)$ s. t.

$$(A^\varepsilon \nabla u^\varepsilon, \nabla \varphi) = (f, \varphi) \quad \forall\, \varphi \in H_0^1(\Omega), \tag{2.1}$$

on a bounded domain $\Omega \subset \mathbb{R}^d$ $(d = 2,3)$ where the generally tensor-valued function $A^\varepsilon \in L^\infty(\Omega)^{d\times d}$ is of heterogeneous character and highly oscillating on a small length scale indicated by a scaling parameter ε. Here, $H_0^1(\Omega)$ is the usual first-order Sobolev Hilbert space with zero Dirichlet data along the boundary $\partial\Omega$. $(\cdot,\cdot)$ denotes the L^2 scalar product on Ω and $\|\cdot\| = (\cdot,\cdot)^{1/2}$ the corresponding norm. The norms of other function spaces are indicated by subscripts, e. g., $\|\cdot\|_{L^\infty(\Omega)}$ or $\|\cdot\|_K = \|\cdot\|_{L^2(K)}$ for a subset $K \subset \bar{\Omega}$. We assume the coefficient tensor A^ε to be symmetric and positive definite (uniformly in ε),

$$A^\varepsilon_{ij} = A^\varepsilon_{ji}, \quad \text{a. e. on } \Omega, \tag{2.2}$$

$$\alpha|\xi|^2 \leqslant \sum_{i,j=1}^{d} A^\varepsilon_{ij}\xi_i\xi_j \leqslant \beta|\xi|^2, \quad \text{a. e. on } \Omega, \quad \xi \in \mathbb{R}^d, \tag{2.3}$$

with constants $\alpha, \beta \in \mathbb{R}_+$, so that (2.1) admits a unique solution.

We will not make any strong assumptions on A^ε—only the weak conditions given above and Lipschitz-continuity (with ε-dependent constant) as stated in Section 2.4. Especially, we allow A^ε to contain pronounced *heterogeneous* behavior in Ω (e.g. by having micro structures with different, space dependent periodicities, or by exhibiting multiscale features only in a specific part of the domain).

Throughout this paper we assume that we deal with a fixed choice of coefficients A^ε, i. e. we are not concerned with the limit $\varepsilon \to 0$ (for ε being interpreted as a scaling parameter). Nevertheless, the ε is kept as a superscript to indicate oscillating coefficients (and functions).

2.1. The effective problem

In order to be able to adapt *model parameters* separately from the discretization it is necessary to introduce an abstract notion of a model. In context of the mathematical homogenization theory (cf. Allaire [6]), Bensoussan et al. [15], and Cioranescu & Donato [24] for the elliptic problem (2.1) a natural starting point is the following homogenized equation: Find $u^0 \in H_0^1(\Omega)$ s. t.

$$\left(A^0 \nabla u^0, \nabla \varphi\right) = \left(f, \varphi\right) \quad \forall\, \varphi \in H_0^1(\Omega), \tag{2.4}$$

where A^0 is an *effective* tensor whose values are determined by local cell problems (cf. Allaire [6]). In the *heterogeneous* case the matrix A^0 is a function depending on x; $A^0 \in L^\infty(\Omega)^{d\times d}$. This space dependency has to be discretized for a numerical scheme. A possible choice (as employed by the HMM) is to start with a finite element discretization of (2.4) and define a sampling problem for every quadrature point in Ω for which $A^0(x)$ has to be evaluated.

However, this is exactly the type of coupling between (macroscale) discretization and sampling process that we try to avoid. Consequently, in order to decouple the sampling process from the coarse-scale discretization, it is necessary to abstract this choice and incorporate it into an *effective model*.

Definition 2.1 (Effective model). Let Ω be a polygonal/polyhedral domain. Then, an "effective model" is defined to be a pair $\left(\mathbb{T}_\delta(\Omega), \bar{A}^\delta\right)$ consisting of a *decomposition* $\mathbb{T}_\delta(\Omega)$ of Ω into cells, e. g. quadrilaterals/hexahedra, called a "sampling mesh" together with a function

$$\bar{A}^\delta : \Omega \longrightarrow \mathbb{R}^{d\times d} \tag{2.5}$$

with *cell-wise constant* values, more precisely, $\bar{A}^\delta\big|_K$ is constant for every $K \in \mathbb{T}_\delta(\Omega)$ (cf. Fig. 1). We assume that $\bar{A}^\delta$ fulfills (2.2) and (2.3).

The effective model $\left(\mathbb{T}_\delta(\Omega), \bar{A}^\delta\right)$ can be constructed by different means. For the purpose of this paper, where we are interested in a posteriori error control a computationally inexpensive reconstruction strategy is desired. In order to avoid the solution of local sampling problems simple averaging schemes can be used:

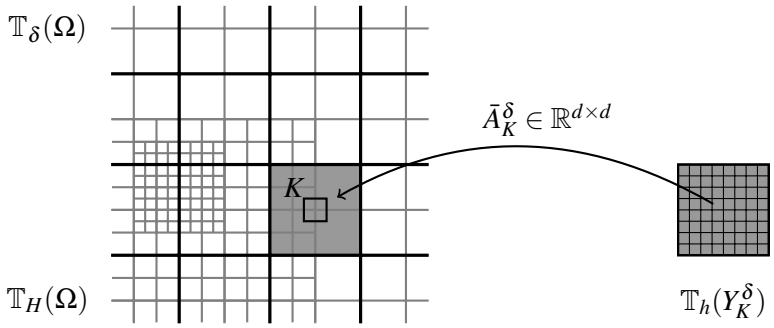

Figure 1. The computational domain Ω together with the *sampling mesh* $\mathbb{T}_\delta(\Omega)$ consisting of sampling regions $K \in \mathbb{T}_\delta$ that are in turn discretized by a fine-scale mesh $\mathbb{T}_h(K)$. The coarse mesh $\mathbb{T}_H(\Omega)$ used for the final finite element discretization is a refinement of the sampling mesh $\mathbb{T}_\delta(\Omega)$.

Definition 2.2 (Averaging sampling strategies). Let $\{Y_K^\delta\}$ be a set of sampling regions, where every Y_K^δ is associated with an individual sampling-mesh cell $K \in \mathbb{T}_\delta(\Omega)$. A sampling region Y_K^δ is further assumed to be a simple translation and rescaling of the unit cell Y (e. g., to the mid-point of K). An averaging process is given by the "geometric mean value"

$$\log \bar{A}_{ij}^\delta(K) := \fint_{Y_K^\delta} \log A_{ij}^\varepsilon(y)\,\mathrm{d}y, \quad \text{for } A_{ij}^\varepsilon(K) \neq 0, \quad \bar{A}_{ij}^\delta = 0, \text{ otherwise}, \tag{2.6}$$

or by the "harmonic mean value"

$$\bar{A}_{ij}^\delta(K)^{-1} := \fint_{Y_K^\delta} \frac{1}{A_{ij}^\varepsilon(y)}\,\mathrm{d}y, \quad \text{for } A_{ij}^\varepsilon(K) \neq 0, \quad \bar{A}_{ij}^\delta = 0, \text{ otherwise}. \tag{2.7}$$

Remark 2.1. It is a well known fact that the arithmetic average

$$\bar{A}_{ij}^\delta(K) := \fint_{Y_K^\delta} A_{ij}^\varepsilon(y)\,\mathrm{d}y, \tag{2.8}$$

is not appropriate in practice: In case of classical homogenization theory it can be shown, e. g., that in 1D the homogenization limit A^0 is given by the harmonic mean value (2.7) and that in higher space dimension the correctors are non-vanishing (cf. Babuska [8] or Cioranescu & Donato [24]).

Remark 2.2. From a physical context it can be shown that, for a large class of porous media, upscaled effective permeabilities lie between the arithmetic and harmonic mean value (cf. Cardwell & Parsons [20]), i. e., transferred to scalar valued, heterogeneous coefficients:

$$\Big(\fint_{Y_K^\delta} \frac{1}{A^\varepsilon(y)}\,\mathrm{d}y\Big)^{-1} \leqslant A^0 \leqslant \fint_{Y_K^\delta} A^\varepsilon(y)\,\mathrm{d}y. \tag{2.9}$$

Furthermore, the geometric average is a reasonable choice for a large class of *log-normally distributed* permeabilities (cf. Warren & Price [49]). For a detailed discussion as well as an overview of newer, more sophisticated averaging methods we refer to Li et al. [36] and references therein.

Remark 2.3. In spirit of classical homogenization theory and the classical HMM formulation, a reconstruction process involving local cell problems on sampling regions can be defined:

$$\bar{A}^{\delta}_{ij}(K) := \fint_{Y^{\delta}_K} A^{\varepsilon}(x)\big(\nabla_x \omega_i(x) + e_i\big) \cdot \big(\nabla_x \omega_j(x) + e_j\big)\,\mathrm{d}x, \tag{2.10}$$

where the $\omega_i \in \tilde{H}^1_{\mathrm{per}}(Y^{\delta}_K)$ are solutions of

$$\int_{Y^{\delta}_K} A^{\varepsilon}(x)\big(\nabla_x \omega_i(x) + e_i\big) \cdot \nabla\varphi(x)\,dx = 0 \quad \forall \varphi \in \tilde{H}^1_{\mathrm{per}}(Y^{\delta}_K). \tag{2.11}$$

In order to avoid a reduction in regularity the patch-wise *constant* (globally discontinuous) coefficients $\bar{A}^{\delta}$ can be further post-processed to yield a globally *continuous* function:

Definition 2.3 (Post-processing). Let $V^{\delta}(\Omega)$ be the space of d-linear finite element functions associated with the sampling mesh $\mathbb{T}_{\delta}(\Omega)$. We define $A^{\delta} \in V^{\delta}(\Omega)^{d\times d}$ with the help of an interpolation of *Clement-type* (cf. Clement [25] and Scott & Zhang [46]): At each nodal point x_i of $\mathbb{T}_{\delta}(\Omega)$, we set

$$A^{\delta}(x_i) := \frac{\sum_{K\in\mathscr{K}_i} \bar{A}^{\delta}(K)\,|K_i|}{\sum_{K\in\mathscr{K}_i} |K_i|}, \tag{2.12}$$

where $\mathscr{K}_i$ is the set of all cells $K \in \mathbb{T}_{\delta}(\Omega)$ with $x_i \in \bar{K}$.

With the help of the effective model $\big(\mathbb{T}_{\delta}(\Omega), A^{\delta}\big)$, we define a corresponding *effective problem* as follows: Find $u^{\delta} \in H^1_0(\Omega)$ s. t.

$$(A^{\delta}\nabla u^{\delta}, \nabla\varphi) = (f,\varphi) \quad \forall\,\varphi \in H^1_0(\Omega). \tag{2.13}$$

Remark 2.4. The post-processed, continuous coefficient function A^{δ} ensures that $u^{\delta} \in H^2(\Omega)$ what will be used in the a priori error estimation to shorten the discussion. From a numerical standpoint it turns out that directly using $\bar{A}^{\delta}$ instead of A^{δ} yields better results, see the discussion in Section 4.

2.2. The semi-discretized problem

A numerical evaluation of the sampling processes (2.8), (2.6), (2.7), or (2.11), introduces quadrature and discretization errors on the fine (sampling) scale that have to be taken into account. Therefore, we denote such a *numerically computed, effective model* with a global discretization parameter h and a sequence $\{h_K\}_{K\in\mathbb{T}_\delta(\Omega)}$ of local discretization parameters for each sampling region Y_K^δ. We define h_K to denote the refinement level of a fine-scale triangulation $\mathbb{T}_h(Y_K^\delta)$ of Y_K^δ, and h to be the maximum of the h_K. With the help of the meshes $\mathbb{T}_h$ a summed quadrature rule $Q_{h,K}$ on Y_K^δ is introduced:

Definition 2.4 (Summed quadrature rule). Given a base quadrature rule $\hat{Q}$ on the reference sampling cell Y with support points $\{\hat{x}_i\}_i$ and weights $\{q_i\}_i$ and denoting by $\mathscr{T}_{\tilde{K}}$ the transformation $\mathscr{T}_{\tilde{K}} : Y \longrightarrow \tilde{K}$ for $\tilde{K} \in \mathbb{T}_h(Y_K^\delta)$, we define a "summed quadrature rule" on K by

$$Q_{h,K}(f) := \sum_{\tilde{K}\in\mathbb{T}_h(Y_K^\delta)} |\tilde{K}|\, Q\big(f\circ\mathscr{T}_{\tilde{K}}\big) = \sum_{\tilde{K}\in\mathbb{T}_h(Y_K^\delta)} |\tilde{K}| \sum_i q_i\,(f\circ\mathscr{T}_{\tilde{K}})(\hat{x}_i). \tag{2.14}$$

With this notation, in correspondence to (2.8), (2.6) and (2.7), we define the *numerically computed, effective model* $\big(\mathbb{T}_\delta(\Omega), \bar{A}^{\delta,h}\big)$ by one of the following options:

$$\bar{A}_{ij}^{\delta,h}(K) := |Y_K^\delta|^{-1} Q_{h,K}\big(A_{ij}^\varepsilon(y)\big), \tag{2.15}$$

$$\big(\bar{A}_{ij}^{\delta,h}(K)\big)^{-1} := |Y_K^\delta|^{-1} Q_{h,K}\Big(\frac{1}{A_{ij}^\varepsilon(y)}\Big), \tag{2.16}$$

$$\log \bar{A}_{ij}^{\delta,h}(K) := |Y_K^\delta|^{-1} Q_{h,K}\big(\log A_{ij}^\varepsilon(y)\big)\,\mathrm{d}y. \tag{2.17}$$

Remark 2.5. In case of the homogenization scheme (2.10), an additional fine-scale discretization has to be taken into account. Choosing a discrete fine-scale space $V^h(Y_K^\delta) \subset \tilde{H}^1_{\mathrm{per}}(Y_K^\delta)$ associated with a mesh $\mathbb{T}_h(Y_K^\delta)$ of Y_K^δ, we set

$$\bar{A}_{ij}^{\delta,h}(x) := |Y_K^\delta|^{-1} Q_{h,K}\big(A^\varepsilon(x)\big(\nabla_x \omega_i^h(x) + e_i\big)\cdot\big(\nabla_x \omega_j^h(x) + e_j\big)\big), \tag{2.18}$$

where

$$Q_{h,K}\big(A^\varepsilon(x)\big(\nabla_x \omega_i(x) + e_i\big)\cdot\nabla\varphi\big) = 0 \quad \forall \varphi \in V^h(Y_K^\delta). \tag{2.19}$$

Finally, introducing a fine-scale discretization for the sampling process leads us to the following *semi-discretized* problem: Find $u^{\delta,h} \in H_0^1(\Omega)$ s. t.

$$(A^{\delta,h}\nabla u^{\delta,h}, \nabla\varphi) = (f,\varphi) \quad \forall\, \varphi \in H_0^1(\Omega). \tag{2.20}$$

The well-posedness and a priori convergence of problem (2.20) will be established in Section 2.4.

2.3. The fully discretized problem

As the last step, we introduce a coarse mesh $\mathbb{T}_H(\Omega)$ of Ω and an associated finite element space $V_H(\Omega) \subset H_0^1(\Omega)$ for the numerical approximation of the macroscale problem (2.20). Then the *fully discrete* problem reads as follows: Find $U \in V_H(\Omega)$ s. t.

$$(A^{\delta,h}\nabla U, \nabla \varphi^H) = (f, \varphi^H) \quad \forall\, \varphi^H \in V_H(\Omega). \tag{2.21}$$

In summary, we have established a framework that explicitly decouples all different error sources, cf. Fig. 2.

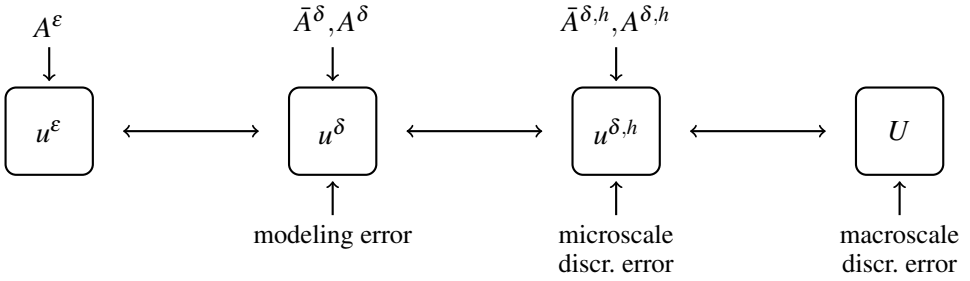

Figure 2. The interplay of the different sources of error ranging from the solution u^ε of the full model problem (2.13), over the solutions u^δ and $u^{\delta,h}$ of the auxiliary problems (2.13) and (2.20), to the solution U of the fully discretized problem (2.21).

Remark 2.6. In Oden & Vemaganti [39,40,41] a similar modeling framework is considered, which can be regarded as a simplification of the above scheme by (i) neglecting fine-scale discretization errors (discussed in Section 2.2) and assuming $u^\delta \equiv u^{\delta,h}$, and (ii) by linking sampling and coarse-grid discretization, i. e., setting $\mathbb{T}_\delta = \mathbb{T}_H$.

2.4. Well-posedness

In this subsection, we briefly discuss well-posedness and a priori convergence of the method. We stress the point that only mild a priori assumptions are made for A^ε and thus the results of this subsection are solely stated for completeness to ensure *effectivity*. The *efficiency* of the method is solely based on the a posteriori error analysis stated in the next section.

As the primary goal of the following short discussion is to ensure well-posedness we prove a priori convergence only for the case of uniformly refined discretization parameters. We thus fix the notation δ_K, H_K, and h_Q to denote the local cell size of a cell $K \in \mathbb{T}_\delta(\Omega)$, $K \in \mathbb{T}_H(\Omega)$, or $Q \in \mathbb{T}_h(K)$, with $K \in \mathbb{T}_\delta(\Omega)$, respectively. The parameters δ, H, and h shall be defined as the corresponding maximum over all local cell diameters.

The following assumptions are made:

(A1) $A^\varepsilon \in L^\infty(\Omega)^{d\times d}$ is Lipschitz continuous with a possibly ε-dependent L-constant:

$$\|A^\varepsilon(y) - A^\varepsilon(x)\| \leqslant C(\varepsilon)\,\|y-x\|. \tag{2.22}$$

(A2) Ω is a polygonal/polyhedral domain and all meshes involved in the multiscale method, $\mathbb{T}_\delta(\Omega)$, $\mathbb{T}_H(\Omega)$, $\{\mathbb{T}_h(Y_K^\delta), K \in \mathbb{T}_\delta(\Omega)\}$, satisfy the following standard regularity conditions (cf. Ciarlet [23] or Brenner & Scott [18]):

- *Structural regularity*: $\bigcup \overline{K} = \overline{\Omega}$, and every intersection $\overline{K}_i \cap \overline{K}_j$, $i \neq j$, is empty, a vertex, an entire edge, or an entire face.
- *Uniform shape regularity*: There exists $C \in \mathbb{R}_+$, such that

$$\|\det \nabla \mathscr{T}_K\| + \|\det \nabla \mathscr{T}_K^{-1}\| \leqslant C, \tag{2.23}$$

  uniformly for all cells K of all families of meshes. Hereby, $\mathscr{T}_K : Y \longrightarrow K$ denotes a d-linear transformation of a unit cell Y onto K. This allows for decompositions of $\bar{\Omega}$ consisting of (closed and convex) triangles/tetrahedra or quadrilaterals/hexahedra.

The ε-dependent coefficient $C(\varepsilon)$ in (A1) can range from no ε-dependence at all, i. e. $C(\varepsilon) = 1$, over to bad exponents, typically $C(\varepsilon) = \varepsilon^{-0.5}$, or $C(\varepsilon) = \varepsilon^{-1}$. Usually strong a priori conditions on A^ε like *periodicity assumptions* (to apply classical homogenization results), or on the type of reconstruction process (e. g. H^1-orthogonality) are done to control this coefficient.

However, for the purpose of this paper that *centers around a posteriori control of discretization and model parameters* we deliberately do not impose any such a priori assumption. In contrast, we assume that we deal with a fixed choice of ε (that might be very small and varying in space) such that a *typical length scale of fluctuations* emerges, that—in the worst case—has to be resolved in full. The aim is now to achieve a *significant* computational saving (of several orders of magnitude) by means of adaptive discretization and model control compared to such a hypothetical full resolution.

Remark 2.7. We note that there is no requirement about "size regularity" of the meshes in order to allow for local mesh adaptation. Further, to ease local mesh refinement the regularity assumption (A2) can be relaxed to allow for "hanging nodes", usually only one per edge or face, where the corresponding degrees of freedom are eliminated by linear interpolation between neighboring "regular" nodes (cf. Carey & Oden [21] or Becker & Braack [11]).

With this prerequisites at hand, the next proposition follows immediately:

Proposition 2.1. *Under above assumptions all intermediate problems, (2.13) for u^{δ}, (2.20) for $u^{\delta,h}$, and (2.21) for U, are well-posed. Due to the fact that $A^{\delta,h}$ is a continuous and patch-wise d-linear interpolation the bound $\|A^{\delta,h}\|_{W^{1,\infty}(\Omega)} \leqslant C(\delta)$ is available, where $C(\delta)$ generally behaves like $1/\delta$. Thus,*

$$\|u^{\delta,h}\|_{H^2(\Omega)} \leqslant C(\delta)\,\|f\|. \tag{2.24}$$

Proposition 2.2. *Under the above assumptions there holds the error estimates*

$$\|u^{\delta} - u^{\varepsilon}\|_{H^1(\Omega)} \leqslant C(\varepsilon)\,\delta, \tag{2.25}$$

$$\|u^{\delta,h} - u^{\delta}\|_{H^1(\Omega)} \leqslant C(\varepsilon)\,h, \tag{2.26}$$

$$\|U - u^{\delta,h}\| + H\,\|\nabla(U - u^{\delta,h})\| \leqslant C(\delta)\,H^2\,\|f\|. \tag{2.27}$$

Proof. The first two statements follow immediately from

$$\|A^{\delta} - A^{\varepsilon}\| \leqslant C(\varepsilon)\,\delta, \quad \|A^{\delta,h} - A^{\delta}\| \leqslant C(\varepsilon)\,h. \tag{2.28}$$

The last statement is known from the standard a priori error analysis for finite element methods (see Ciarlet [23]). The independence of $C(\varepsilon)$ of the other parameters follows from Proposition 2.1. □

3. A posteriori error analysis

In this section, we derive an a posteriori error estimate for our multiscale scheme within the framework of the *Dual Weighted Residual* (DWR) method of Becker & Rannacher [12,13,14]. Hereby, a so called *dual problem* is solved that is defined in terms of the coefficients A^{ε} and a *quantity of interest* represented by a functional value $j(u^{\varepsilon})$. The goal of this discussion is the derivation of separate a posteriori error indicators for all sources of error—the discretization errors on the macro- and microscale, as well as the model error.

A fundamental difficulty that has to be taken care of arises from the fact that in the case of the elliptic model problem the computation of the solution to the dual problem is of the same complexity as that of the primal problem. Thus, a computationally cheap, but still reasonably good approximation is needed.

Remark 3.1. A similar approach has been used by Oden & Vemaganti [39,40,41], Romkes & Moody [45], and Braack & Ern [16] in order to estimate model errors.

3.1. Error identity

Let $j \in H^{-1}(\Omega)$ be a linear, continuous functional on $H_0^1(\Omega)$ and suppose that a *quantity of interest* is given by $\langle j, u^{\varepsilon}\rangle$, where $\langle\cdot,\cdot\rangle$ denotes the duality pairing on

$H^{-1}(\Omega)\times H_0^1(\Omega)$. We define two *dual problems*, a "full" and a "reduced" one as follows: Find $z^\varepsilon \in H_0^1(\Omega)$ and $z^\delta \in H_0^1(\Omega)$, s. t.

$$\left(A^\varepsilon \nabla\varphi, \nabla z^\varepsilon\right) = \langle j,\varphi\rangle \quad \forall\,\varphi\in H_0^1(\Omega), \tag{3.1}$$

$$\left(A^\delta \nabla\varphi, \nabla z^\delta\right) = \langle j,\varphi\rangle \quad \forall\,\varphi\in H_0^1(\Omega). \tag{3.2}$$

Both dual problems are well posed and admit unique solutions. In the following, we require the assumptions of the preceding sections to be satisfied and use the notation introduced above.

First, a straightforward calculation leads to the following basic result.

Proposition 3.1 (Error identity). *There holds the error representation*

$$\begin{aligned}\langle j,u^\varepsilon\rangle - \langle j,U\rangle &= \left[\langle j,u^\delta\rangle - \langle j,U\rangle\right] + \left[\langle j,u^\varepsilon\rangle - \langle j,u^\delta\rangle\right]\\ &= \left[(f,z^\delta) - (A^{\delta,h}\nabla U,\nabla z^\delta)\right] + (\{A^{\delta,h}-A^\delta\}\nabla U,\nabla z^\delta)\\ &\quad + (\{A^\delta - A^\varepsilon\}\nabla u^\delta,\nabla z^\varepsilon),\\ &=: \vartheta^H + \vartheta^h + \vartheta^\delta,\end{aligned} \tag{3.3}$$

where ϑ^H is a residual on the macroscale*, ϑ^h takes the form of a* residual on the microscale *and ϑ^δ has the character of a model error.*

Next, we derive representations of the error estimators ϑ^H, ϑ^h, and ϑ^δ, which can be evaluated numerically.

Proposition 3.2 (Error estimators and indicators). *The macroscale error estimator ϑ^H allows the representation:*

$$\begin{aligned}\vartheta^H &= \sum_{K\in\mathbb{T}_H(\Omega)} \eta_K^H,\\ \eta_K^H &:= (f+\nabla\cdot A^{\delta,h}\nabla U, z^\delta-\varphi^H)_K - (\tfrac{1}{2}[n\cdot A^{\delta,h}\nabla U]_{\partial K}, z^\delta-\varphi^H)_{\partial K},\end{aligned} \tag{3.4}$$

for arbitrary $\varphi^H\in V_H(\Omega)$. Hereby, $[.]_{\partial K}$ denotes the jump across the inter-element boundary ∂K and n is the outward unit normal. The microscale error estimator is given by

$$\begin{aligned}\vartheta^h &= \sum_{K\in\mathbb{T}_\delta(\Omega)} \eta_K^h,\\ \eta_K^h &:= \left(\{A^\delta - A^{\delta,h}\}\nabla U,\nabla z^\delta\right)_K.\end{aligned} \tag{3.5}$$

Proof. Equation (3.4) immediately follows by virtue of Galerkin orthogonality

and by cellwise integration by parts:

$$\begin{aligned}\vartheta^H &= \sum_{K\in\mathbb{T}_H(\Omega)} \big(f, z^\delta - \varphi^H\big)_K - \big(A^{\delta,h}\nabla U, \nabla z^\delta - \nabla\varphi^H\big)_K,\\ &= \sum_{K\in\mathbb{T}_H(\Omega)} \big(f + \nabla\cdot A^{\delta,h}\nabla U, z^\delta - \varphi^H\big)_K - \big(\tfrac{1}{2}[n\cdot A^{\delta,h}\nabla U]_{\partial K}, z^\delta - \varphi^H\big)_{\partial K}.\end{aligned}$$

The equation (3.5) is obvious. □

The residual-type error estimators ϑ^H and ϑ^h are well behaved, i. e., they are uniformly bounded in powers of H and h, respectively. In contrast to the macroscale discretization indicator its microscale pendant only depends on the approximation order of the quadrature rule used.

Proposition 3.3. *The macroscale and microscale error estimators ϑ^H and ϑ^h as well as their local indicators η_K^H and η_K^h admit the asymptotic estimates*

$$|\vartheta^H| \leqslant \sum_{K\in\mathbb{T}_H(\Omega)} |\eta_K^H| \leqslant C(\delta)H, \tag{3.6}$$

$$|\vartheta^h| \leqslant \sum_{K\in\mathbb{T}_\delta(\Omega)} |\eta_K^h| \leqslant C(\varepsilon)\,h. \tag{3.7}$$

Proof. (i) By definition using Galerkin orthogonality, there holds

$$\begin{aligned}|\vartheta^H| &= |(f, z^\delta) - (A^{\delta,h}\nabla U, \nabla z^\delta)|\\ &= |(A^{\delta,h}\nabla u^{\delta,h}, \nabla(z^\delta - \pi_h z^\delta)) - (A^{\delta,h}\nabla U, \nabla(z^\delta - \pi_h z^\delta))|\\ &\leqslant C\|A^{\delta,h}\|_{L^\infty(\Omega)}\|\nabla(u^{\delta,h} - U)\|\,\|\nabla(z^\delta - \pi_h z^\delta)\|.\end{aligned}$$

Here, $\pi_h z^\delta \in V_H(\Omega)$ denotes a generalized nodal interpolation of z^δ, which is stable with respect to the H^1 norm (cf. Scott & Zhang [46]) and satisfies the error estimate

$$\|z^\delta - \pi_h z^\delta\| + H\|\nabla(z^\delta - \pi_H z^\delta)\| \leqslant CH\|z^\delta\|_{H^1(\Omega)}. \tag{3.8}$$

We further have available the bound $\|A^{\delta,h}\|_{L^\infty(\Omega)} \leqslant C(\delta)$ and the a priori estimate (2.2),

$$\|U - u^{\delta,h}\| + H\,\|\nabla(U - u^{\delta,h})\| \leqslant C(\delta)\,H^2. \tag{3.9}$$

Combining the foregoing estimates yields the result

$$|\vartheta^H| \leqslant C(\delta)H. \tag{3.10}$$

The full estimate involving the absolute values $|\eta_K^H|$ can be shown by standard techniques from finite element a posteriori error analysis (cf. Verfürth [47]). We omit the

details.

(ii) The estimate (3.7) is an immediate consequence of Proposition 2.2 and its proof, where the fact was used that $\|A^{\delta,h} - A^{\delta}\|_{L^\infty(\Omega)}$ is bounded with the respective convergence order. □

Remark 3.2. An alternative way of localizing the global residual term ϑ^H, which is cell vertex-oriented and avoids the computation of jumps of derivatives across inter-element boundaries, has been proposed in Braack & Ern [16] (see also Richter& Wick [44]).

Now, based on the definition of ϑ^δ define local *model-error* indicators:

$$\begin{aligned} \vartheta^\delta &= \sum_{K\in\mathbb{T}_\delta(\Omega)} \eta_K^\delta, \\ \eta_K^\delta &:= \big(\{A^\varepsilon - A^\delta\}\nabla u^\delta, \nabla z^\varepsilon\big)_K. \end{aligned} \tag{3.11}$$

Remark 3.3. Without assuming any additional structural properties of the functional j to hold true, the following basic convergence result (in spirit of Proposition 2.2) for the effectivity of the model error estimator can be shown:

$$|\vartheta^\delta| \leqslant \sum_{K\in\mathbb{T}_\delta(\Omega)} |\eta_K^\delta| \leqslant \sum_{K\in\mathbb{T}_\delta(\Omega)} C(\varepsilon)\|\nabla z^\varepsilon\|_K\,\delta \leqslant C(\varepsilon)\,\delta. \tag{3.12}$$

The above estimate is too pessimistic in practice. The *key point* of the whole a posteriori approach formulated in this section is the fact that the model error indicator $\eta_K^\delta = \big(\{A^\varepsilon - A^\delta\}\nabla u^\delta, \nabla z^\varepsilon\big)_K$ contains *the dual solution* as a *weighting factor*. Thus, for a localized functional, where z^ε is of Green's function type, the bad error constant $C(\varepsilon)$ is only present in a small region that needs to be resolved in full (see the numerical results given in Section 4).

3.2. Practical evaluation of the error indicators

The residual-type error indicators η_K^H and η_K^h can be evaluated directly by standard techniques. A straightforward method is to replace $z^\delta - \varphi^H$ and A^δ in (3.4) or (3.5) by a higher order approximation. Given the fact that for the evaluation of η_K^h an approximation of A^δ is necessary, we propose the following strategy.

Definition 3.1 (Approximate discretization-error indicators). We denote by $\bar{A}^{\delta,h/2}$ the numerically computed, effective coefficient with respect to the finer sampling mesh $\{\mathbb{T}_{h/2}(Y_K^\delta) : K \in \mathbb{T}_\delta(\Omega)\}$ and define $\tilde{Z} \in V^H(\Omega)$ to be the solution of

$$(\bar{A}^{\delta,h/2}\nabla\varphi^H, \nabla\tilde{Z}) = \langle j, \varphi^H\rangle \quad \forall\varphi^H \in V^H(\Omega). \tag{3.13}$$

Further, we denote by $\pi_{2H}^{(2)}$ the patchwise interpolation to a higher-order finite element space (d-quadratics) on a once coarsened mesh $\mathbb{T}_{2H}(\Omega)$. Then, the approximate error indicators are defined by

$$\tilde{\eta}_K^H := (f, \pi_{2H}^{(2)}\tilde{Z} - \tilde{Z})_K - (A^{\delta,h}\nabla U, \nabla(\pi_{2H}^{(2)}\tilde{Z} - \tilde{Z}))_K, \tag{3.14}$$

$$\tilde{\eta}_K^h := (\{\bar{A}^{\delta,h/2} - A^{\delta,h}\}\nabla U, \nabla z^\delta)_K. \tag{3.15}$$

Remark 3.4. The replacement of $z^\delta - \pi_H z^\delta$ by $\pi_{2H}^{(2)}\tilde{Z} - \tilde{Z}$ is a well known post-processing technique (cf. Becker & Rannacher [14]). It usually leads to a slight reduction in the accuracy of the estimator (i. e. the quantitative prediction of the estimator). However, the qualitative properties of the resulting error indicators are preserved.

3.2.1. Evaluation of the model error. The model-error indicators η_K^δ are of fundamentally different nature than the residual-type indicators η_K^H and η_K^h, because they are defined in terms of the solution z^ε of the full dual problem (3.1). A *global* approximation of z^ε is (in the case of the elliptic model problem) of the same complexity as the primal problem itself. Consequently, approximating z^ε directly is practically infeasible. Different strategies have been used to circumvent this problem:

– A straightforward idea (e. g. used by Braack & Ern [16]) is to use the homogenized dual solution z^δ directly instead of z^ε in order to estimate model errors. This approach is justified if a *qualitative—not a quantitative*—error estimator is needed, because a critical under- or overestimation of the model error has to be expected. This can be explained by the following heuristic reasoning: The indicator η_K^δ corresponds to a moment of second order,

$$\eta_K^\delta = (\underbrace{(A^\varepsilon - A^\delta)}_{\text{fluct.}}\nabla u^\delta, \underbrace{\nabla z^\varepsilon}_{\text{fluct.}})_K, \tag{3.16}$$

whereas $\tilde{\eta}_K^\delta$ defined in terms of z^δ only corresponds to a moment of first order,

$$\tilde{\eta}_K^\delta = (\underbrace{(A^\varepsilon - A^\delta)}_{\text{fluct.}}\nabla u^\delta, \nabla z^\delta)_K. \tag{3.17}$$

– Oden & Vemaganti [39,40,41] introduced a refined version of the above strategy that uses an additional, cellwise inverse of the coefficients, $I_d - \left(A^\varepsilon\right)^{-1}A^\delta$, in the estimation of the model-error indicator. This results in an estimate of the form:

$$|\langle j, u^\delta\rangle - \langle j, u^\varepsilon\rangle| \leqslant C \sum_{K\in\mathbb{T}_\delta(\Omega)} \left|\left(\nabla u^\delta, [I_d - (A^\varepsilon)^{-1}A^\delta]\nabla z^\delta\right)_K\right| + \mathscr{R}, \tag{3.18}$$

with a (not directly computable) residual-type remainder $\mathscr{R}$.

– An alternative post-processing of z^δ has been proposed by Romkes & Moody [45], in which the global dual problem (3.1) is replaced by entirely local Neumann problems solely defined on the respective sampling cell: Find $\hat{z}^\varepsilon \in H^1(K)$ s. t.

$$(A^\varepsilon \nabla \varphi, \nabla \hat{z}^\varepsilon) = \langle j, \varphi \rangle \quad \forall\, \varphi \in H^1(K). \tag{3.19}$$

This approach is computationally rather expensive and does not provide information about the global error dependencies required within the DWR method.

Remark 3.5. It can be expected, that the approximate model-error indicators $\tilde{\eta}_K^\delta$ defined in terms of the reduced dual solution z^δ still contain enough qualitative information for a (stable) model adaptation strategy that does not need a good quantitative estimate. This is similar to the case of the classical DWR method, where the dual problem is defined on the same (adapted) mesh as the primal problem (cf. Becker & Braack [11]).

In summary, we use the following numerical approximation of the model-error indicators.

Definition 3.2 (Approximate model-error indicators). Let $A^{\delta,h/2}$ be the approximate effective coefficient of Definition 3.1 and let $\tilde{Z}$ be the corresponding solution of the (discrete) dual problem (3.13). Then, the approximate model-error indicators are defined by

$$\tilde{\eta}_K^\delta := (\{A^{\delta,h/2} - A^\varepsilon\}\nabla U, \nabla \tilde{Z})_K. \tag{3.20}$$

Remark 3.6. A local correction $\tilde{\eta}_{K,\mathrm{rec}}^\delta$ of $\tilde{\eta}_K^\delta$, in spirit of the third approximation strategy described above, is given by

$$\tilde{\eta}_{K,\mathrm{rec}}^\delta := (\{A^{\delta,h/2} - A^\varepsilon\}\nabla U, \nabla \tilde{Z}_K)_K, \tag{3.21}$$

with a finite element approximation $\tilde{Z}_K$ solving the local reconstruction problem: Find $\tilde{Z}_k \in V^h(K)$ s. t.

$$(A^\varepsilon \nabla \varphi, \nabla \tilde{Z} + \nabla \tilde{Z}_K)_K = \langle j, \varphi \rangle \quad \forall\, \varphi \in V^h(K). \tag{3.22}$$

3.3. Simultaneous adaptation of model and discretization

The a posteriori error indicators derived in the preceding section can be used in a strategy for simultaneous discretization and model adaptation. For the sake of simplicity, we only consider homogeneous meshes $\mathbb{T}_h(Y_K^\delta)$ of the sampling regions Y_K^δ

and adapt the local finescale discretization solely by the choice of h_K (denoting the uniform refinement parameter of $\mathbb{T}_h(Y_K^\delta)$. In contrast, the macroscale discretization is allowed to be a family of locally refined meshes according to Remark 2.7.

For the model adaptation two fundamentally different possibilities of improving the effective model $\left(\mathbb{T}_\delta(\Omega), A^\delta\right)$ have to be considered:

- *Model switching*: If the local model-error indicator η_K^δ is large, locally switch to a better, but more expensive model. In principle, such an approach could involve several increasingly expensive model formulations of the microscale. In its simplest incarnation the model switching is between a (single) effective cheap model A_K^δ and the full model $A^\varepsilon(x)$ on K.
- Adapt the discretization, i. e., the choice of the sampling mesh $\mathbb{T}_\delta(\Omega)$ while keeping the microscale model.

Remark 3.7. The first strategy is equivalent to the *region of influence* method formulated by Oden & Vemaganti [39,40]. Hereby, the model indicators $\tilde{\eta}_K^\delta$ are used to (iteratively) construct a connected subdomain $I \subset \Omega$ consisting of $K \in \mathbb{T}_\delta(\Omega)$ with large modeling errors. On the so-called *region of influence* I a correction with the full model is computed: Find $U_{\text{corr}} \in V^h(I)$ s. t.

$$(A^\varepsilon(\nabla U + \nabla U_{\text{corr}}), \nabla\varphi) = (f, \varphi) \quad \forall \varphi \in V^h(I), \tag{3.23}$$

$$U_{\text{corr}} = 0 \quad \text{on } \partial I. \tag{3.24}$$

Remark 3.8. In case of the "digital" model switching strategy between reduced and full model both strategies are actually equivalent! This is due to the fact that all averaging strategies in Definition 2.2 can be regarded as an (expensive) quadrature rule. Therefore, a local refinement of $\mathbb{T}_\delta(\Omega)$ is at some point nothing else than a full resolution in the sense of the first strategy. This claim also holds for homogenization strategies with local cell problems. It can be shown that the corrector defined by the local cell problem vanishes asymptotically for $\delta \ll \varepsilon$ and under assumption (A1), i. e., also homogenization processes degrade in this sense to mere quadrature rules.

Remark 3.9. A fundamentally different approach in which the model-error indicators η_K^δ will be used to directly enhance the effective model in a model optimization framework will be the subject of a forthcoming paper.

In order to control the microscale discretization when adapting the sampling region, the local mesh $\mathbb{T}_h(K)$ of a sampling region $K \in \mathbb{T}_\delta(\Omega)$ involved has to be kept at the *same level of resolution*. More precisely, we choose the following approach:

- If η_K^δ is large for some $K \in \mathbb{T}_\delta(\Omega)$, split K into 2^d sampling regions K_i, correspondingly shrink the sampling region Y_K^δ by a factor of 2^{-d} and set

$h_{K_i} = h_K$ – this amounts to a coarsening of one level if expressed in terms of number of grid refinements (c. f. Fig. 3).

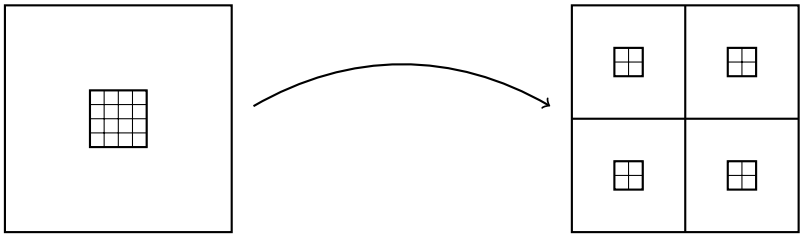

Figure 3. Refinement of a sampling cell K and the associated sampling region Y_K^δ. The degree of microlevel resolution is preserved.

3.3.1. The adaptation algorithm in detail. Let the goal be to reach a prescribed error tolerance TOL in a finite number of adaptation cycles. To achieve this, the adaptation algorithm based on the a posteriori error indicators derived above consists of the following steps.

Step 1. Start with initial meshes $\mathbb{T}_H(\Omega)$, $\mathbb{T}_\delta(\Omega)$, $\left\{\mathbb{T}_h(Y_K^\delta), K \in \mathbb{T}_\delta(\Omega)\right\}$ and choose scaling parameters α_H, α_h, and α_δ.

Step 2. Compute $A^{\delta,h}$ and $\bar{A}^{\delta,h/2}$ according to one of the strategies in Definition 2.4 and compute U and $\tilde{Z}$ with the help of (2.21) and (3.13), respectively.

Step 3. Compute the error estimators and local indicators

$$\tilde{\vartheta}^H = \sum_{K\in\mathbb{T}_H(\Omega)} \tilde{\eta}_K^H, \qquad \tilde{\vartheta}^h = \sum_{K\in\mathbb{T}_\delta(\Omega)} \tilde{\eta}_K^h, \qquad \tilde{\vartheta}^\delta = \sum_{K\in\mathbb{T}_\delta(\Omega)} \tilde{\eta}_K^\delta$$

according to (3.14), (3.15), and (3.20). Optionally, determine a local enhancement $\tilde{\vartheta}_{\mathrm{rec}}^\delta = \sum_{K\in\mathbb{T}_\delta(\Omega)} \tilde{\eta}_{K,\mathrm{rec}}^\delta$ following Remark 3.6.

Step 4. If $|\tilde{\vartheta}^H + \tilde{\vartheta}^h + \tilde{\vartheta}^\delta| \leqslant \mathrm{TOL}$, then stop. Otherwise continue.

Step 5a. Based on the local error indicators $|\eta_K^\nu|$, fo each source of error independently select cells for refinement. This can either be done by selecting a fixed fraction of cells with highest (absolute value) $|\eta_K^\nu|$. Alternatively, a more sophisticated marking strategy based on local optimization is possible (cf. Braack & Richter [17]).

Step 5b. In order to balance the adaptation, not all of the selected cells are used but only the fraction

$$\frac{\alpha_\nu\,|\tilde{\vartheta}^\nu|}{\alpha_H\,|\tilde{\vartheta}^H| + \alpha_\delta\,|\tilde{\vartheta}^\delta| + \alpha_h\,|\tilde{\vartheta}^h|}, \tag{3.25}$$

where α_ν are fixed scaling parameters.

Step 6a. *Microlevel refinement*. Set $h_K \leftarrow h_K/2$ for all selected cells $K \in \mathbb{T}_\delta(\Omega)$.

Step 6b. *Sampling refinement*. In order to maintain $\mathbb{T}_H(\Omega) \supset \mathbb{T}_\delta(\Omega)$, *additionally* mark every macro cell $K \in \mathbb{T}_H(\Omega)$ for refinement that is equal to a selected sampling cell K of $\mathbb{T}_\delta(\Omega)$. Finally execute refinement: Split each selected sampling cell K into a finite number of children cells K_i, associate new sampling regions $Y^\delta_{K_i}$ with half edge length (in 2D) and discretizations $\mathbb{T}_{h_{K_i}}(Y^\delta_{K_i})$ with $h_{K_i} = h_K$, cf. Fig. 3.

Step 6c. *Macrolevel refinement*. Split each selected macrolevel cell $K \in \mathbb{T}_H(\Omega)$ into a finite number of child cells (while ensuring that $\mathbb{T}_H(\Omega) \supset \mathbb{T}_\delta(\Omega)$).

Remark 3.10. In cases where the post-processed coefficient A^δ is used instead of the piecewise constant variant $\bar{A}^\delta$ the marking strategy for the sampling grid in step 6b can be improved by additionally selecting all sampling cells for refinement that are part of the interpolation patch of the originally selected cells (and therefore influence the post-processed coefficients).

4. Numerical results

This section presents results of some computational tests for illustrating the performance of the adaptation algorithm described above. We consider the elliptic boundary value problem (2.1) with two different types of micro structures. The first example comprises an artificial, locally periodic and exponentially scaled coefficient with an additional quadratic scaling towards the corners. In the second example a *log-normally distributed* permeability with Gaussian correlation is considered. In both cases the quantity of interest is given by an approximate point evaluation of the derivative of u^ε in x_2-direction. The adaptation algorithm is compared to both, traditional finite element approximation as well as the HMM variant described in Remark 2.3. All calculations have been performed with the finite element library `deal.II` [9]. The stochastic permeability has been generated with the help of the `QuantIm` library [48].

Further, all computations were done using the averaged and patchwise constant, effective coefficient $\bar{A}^{\delta,h}$ directly instead of its post-processed counterpart. It turns out that the Clement-type post-processing of the sampled, effective coefficient—although necessary for the theoretical analysis (in order to not loose regularity and therefore convergence order on the macroscale)—produces less favorable numerical results. The reason for this lies in the fact that the post-processing step introduces a higher degree of non-locality where the effective value on a given sampling mesh cell is in fact influenced by the sampling process on a patch of (usually) 8 neighboring cells. Therefore, an unfavorable higher degree of refinement of the sampling mesh is necessary to improve such a post-processed effective model. Further, the singularities introduced in patch corners due to patchwise constant coefficients are actually well controlled: In regions, where they matter a high refinement of the macroscale is already necessary, whereas in other regions the small weight (due to the

dual solution) prevents unnecessary refinement.

4.1. Periodic structure

In a first example, we examine the performance of the described sampling process for an artificial periodic structure. This is done in a series of small tests that demonstrate specific aspects of the sampling algorithm and the test case considered. Let Ω be the slit domain as depicted in Fig. 4 and let the quantity of interest be given by a point evaluation of the derivative in x_2-direction:

$$\langle j, \varphi \rangle := \partial_{x_2} \varphi(\hat{x}), \qquad \hat{x} = (0.25, 0.25). \tag{4.1}$$

This specific quantity of interest is chosen because it is equally influenced by macroscale and microscale behavior. Since the corresponding functional (derivative Green function) is not in $H^{-1}(\Omega)$, a regularized approximation depending on the local mesh width $H_{\min}$ is used in our calculations (for a discussion of this approach see Becker & Rannacher [14] or Bangerth & Rannacher [10]).

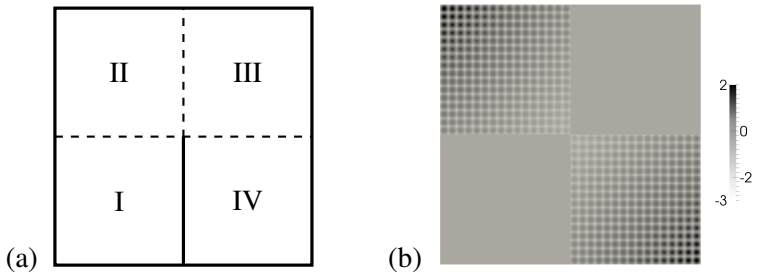

Figure 4. (a) The computational domain Ω. The evaluation point of the quantity of interest is located in the middle of quadrant I. The coefficient A^ε has a strong oscillation in quadrants II and IV and is smooth in III. (b) A logarithmically scaled intensity plot of the trace of the coefficient matrix A^ε of the first numerical example, i. e. $\log_{10}(1/3 \times \mathrm{tr}(A^\varepsilon))$.

We choose A^ε to be a permeability with blockwise different character (cf. Fig. 4):

$$A^\varepsilon(x) = I_d \, \gamma \begin{cases} \exp(3) & \text{in quadrant I and quadrant III,} \\ \exp\big(6\,(1-x_1)\,x_2\,[\cos(\pi\hat{x}_1)+\cos(\pi\hat{x}_2)]\big) & \text{in quadrant II,} \\ \exp\big(6\,(1-x_2)\,x_1\,[\cos(\pi\hat{x}_1)+\cos(\pi\hat{x}_2)]\big) & \text{in quadrant IV.} \end{cases} \tag{4.2}$$

Hereby, $\gamma = 0.001$ and the rescaling $\hat{x}_i$ is defined as $\hat{x}_i := \lfloor x_i/\varepsilon \rfloor - 1/2$. This choice allows us to simultaneously test different aspects of the adaptation algorithm:

– The evaluation of the derivative in quadrant I should enforce a local refinement of the macroscale discretization without an adaptation of microscale and sampling parameters.

– Discretization and sampling parameters for quadrant III should remain coarse because no heterogeneity is present.

– The other two quadrants II and IV have a pronounced heterogeneous fine scale consisting of a periodic substructure and a quadratic scaling of the exponent towards the corners. Quadrant II has a higher influence on the quantity of interest than quadrant IV. Therefore, a higher degree of model adaptation is expected for quadrant II.

For the choice $\varepsilon = 2^{-5} = 0.03125$ a reference solution is computed by a direct finite element approximation on a very fine mesh ($16.8 \cdot 10^6$ degrees of freedom) resulting in the reference values $\partial_{x_2} u^\varepsilon_{\text{ref}}(\hat{x}) \approx 1.699$ and $\|u^\varepsilon_{\text{ref}}\| \approx 0.810$.

4.1.1. Uniform and local refinement. The periodic microstructure exhibits a strong influence on the macroscale behavior. In order to resolve this, a uniformly high resolution is necessary. To exemplify this claim a (direct) finite element approximation with summed quadrature rule (2^d-point Gauss quadrature) is employed for uniform as well as local mesh refinement. The local refinement strategy is based on the macroscale error indicators (3.14) but with (in the absence of an effective model) $A^{\delta,h}$ replaced by A^ε. The results are shown in Tables 1 and 2. We note that in case of uniform refinement (compared to the diffusion problem with smooth parameters) a very high resolution of $6.6 \cdot 10^4$ cells is needed to reach a relative error of 5%, and $1.0 \cdot 10^6$ cells for a relative error of 1%. The local refinement strategy shows almost no improvement compared to uniform refinement. The column next to that for the L^2-error norm contains the logarithmic reduction rate of the error. It shows the typical order reduction which is to be expected for the given macroscopic singularity and strong heterogeneity.

Table 1. Error development in L^2 norm and quantity of interest for uniform refinement with a direct finite element discretization (and high-order quadrature). The second column for the L^2-error norm contains the logarithmic reduction rate.

#macro	L^2 error		$\lvert\langle j, u^\varepsilon - U\rangle\rvert$
1 024	2.60e-1	—	6.31e-1 (37%)
4 096	2.55e-1	0.03	6.21e-1 (36%)
16 384	1.24e-1	1.04	2.81e-1 (17%)
65 536	4.05e-2	1.62	8.62e-2 (5.1%)
262 144	1.20e-2	1.76	2.39e-2 (1.4%)
1 048 576	3.71e-3	1.69	6.64e-3 (0.4%)

Table 2. Error development in L^2 norm and quantity of interest for local refinement with a direct finite element discretization (and high-order quadrature).

	#macro	L^2 error	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$
1	1 024	2.60e-1	6.31e-1 (37%)
2	1 558	2.57e-1	6.24e-1 (37%)
3	2 479	2.52e-1	5.59e-1 (33%)
⋮	⋮	⋮	⋮ ⋮
9	35 377	1.54e-1	1.06e-1 (6.2%)
10	93 331	1.30e-1	6.29e-2 (3.7%)
11	222 064	1.02e-1	3.10e-2 (1.8%)
12	568 201	6.77e-2	1.19e-2 (0.7%)

4.1.2. Model error of averaging schemes and HMM method. Next, in order to determine the typical size of model errors for the test configuration, we examine the different sampling strategies given in Definition 2.2 and the (modified) HMM scheme according to Remark 2.3. Due to the known periodicity of A^ε, we choose $\mathbb{T}_\delta(\Omega)$ to be a uniform mesh with optimal spacing $\delta = \varepsilon = 2^{-5}$ and set the sampling regions to $Y_K^\delta = K$. We choose a high resolution in the macroscale discretization ($2.6 \cdot 10^5$ cells) and microscale discretization ($1.0 \cdot 10^6$ cells) to get an estimate for the model error. The results are shown in Table 3.

The model error for the (modified) HMM is at around 1%, which is expected for a periodic fine scale with a scale separation of around $\varepsilon = 2^{-5}$. The geometric averaging strategy performs comparably well for the periodic coefficient A^ε with a model error of around 6%. In contrast, the averaging strategy has a relative error of 36%. This explains the poor performance of traditional local mesh adaptation (without an effective model); the (wrong) arithmetic average enforces a strong non-local mesh refinement. This is further evidenced by the fact that the error in the L^2 norm for the simple arithmetic averaging scheme is almost one order of magnitude larger than that of the geometric average or the HMM.

Table 3. Model error in L^2 and H^1 norm, and in the quantity of interest for the proposed sampling schemes and the HMM.

strategy	L^2 error	H^1 error	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$
arithmetic averaging	2.55e-1 (31%)	3.21	6.25e-1 (36%)
geometric averaging	5.16e-2 (6.4%)	2.74	1.08e-1 (6.4%)
harmonic averaging	4.41e-1 (54%)	3.98	7.87e-1 (46%)
HMM	1.58e-2 (2.0%)	2.72	1.45e-2 (0.9%)

4.1.3. Adaptive sampling algorithm. As a last test for the periodic permeability, we run the full adaptation algorithm and compare the resulting mesh refinement with uniform and local refinement and the HMM. We choose the geometric averaging

as base sampling strategy for the adaptive sampling algorithm. We start with a very coarse macroscale and microscale discretization of 256 and 1 024 cells, respectively, and 64 distinct sampling regions. The scaling parameters in the balancing step of the adaptive algorithm are set to $\alpha_H = \alpha_\delta = 1$ and $\alpha_h = 10$. This enforces slightly more accuracy on the microscale. The initial microscale resolution is chosen in such a way that the microstructure is coarsely resolved. The results of the adaptation process are shown in Table 4 together with the intermediate values for the different error estimators. A slight initial overestimation, $I_{\text{eff}} > 1$, where

$$I_{\text{eff}} := \left| \frac{\tilde{\vartheta}^H + \tilde{\vartheta}^h + \tilde{\vartheta}^\delta}{\langle j, u^\varepsilon - U \rangle} \right|, \tag{4.3}$$

mainly of the model error changes into a pronounced underestimation in the asymptotic refinement limit. Nevertheless, the qualitative character of the individual error indicators is still well preserved. Fig. 5 shows the adapted meshes for cycle 6 (which corresponds to 5% relative error): The macroscale discretization (Fig. 5a) is locally refined at the point $(0.25, 0.25)$ in quadrant I (due to the choice of quantity of interest) as well as where the sampling discretization enforces a local refinement. The sampling discretization is adapted in quadrant II and IV, with significantly more refinement in quadrant II than in quadrant IV. The same observation holds true for the microscale discretization.

Table 4. Refinement history, error in the quantity of interest and error estimator for the sampling strategy with geometric averaging.

#macro	#sampl.	#micro	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$	$\lvert\vartheta^H\rvert$	$\lvert\vartheta^h\rvert$	$\lvert\vartheta^\delta\rvert$	I_{eff}
256	64	1 024	9.04e-2 (11%)	3.61e-2	1.04e-1	9.99e-1	11.811
262	91	1 456	2.63e-1 (15%)	1.48e-1	1.80e-2	9.00e-1	2.929
385	217	1 696	2.13e-1 (13%)	9.87e-2	8.87e-3	9.05e-1	3.830
886	709	2 032	1.87e-1 (11%)	7.31e-2	5.76e-3	9.06e-1	4.489
2 803	2 620	4 180	1.61e-1 (9.5%)	4.83e-2	9.84e-4	3.36e-1	1.794
7 357	7 102	8 374	6.58e-2 (3.9%)	9.60e-3	2.23e-4	1.17e-1	1.633
17 749	17 407	18 355	3.12e-2 (1.8%)	5.29e-3	6.74e-5	5.25e-2	1.515
39 736	39 262	39 775	2.63e-2 (1.5%)	1.34e-2	1.46e-5	2.15e-2	0.310
89 911	88 675	89 056	1.11e-2 (0.7%)	5.48e-3	3.02e-6	9.06e-3	0.323

The adaptive sampling algorithm leads to a significant improvement in terms of necessary refinement (on macro- and microscale) compared to the direct finite element discretization and HMM (cf. Table 5). The values for the HMM are determined with the help of the adaptive algorithm and a fixed sampling mesh in order to provide a more sensible comparison than uniform refinement. From Table 5 it can be seen that the adaptive sampling is consistently better than uniform and local refinement. This is mainly due to the fact that the geometric average allows for a higher localization in the mesh adaptation than the arithmetic average, which is involuntarily used in the classical finite element discretization. Compared to the

adaptive HMM the adaptive algorithm needs significantly more cells for macroscale and sampling (compared to the fixed 1 024 sampling regions of the HMM). However, the 3.0×10^5 cells on the microscale correspond to 1 024 full sampling problems in case of the HMM, where for the averaging scheme only a significantly less expensive averaging procedure with 9.0×10^4 cells is necessary.

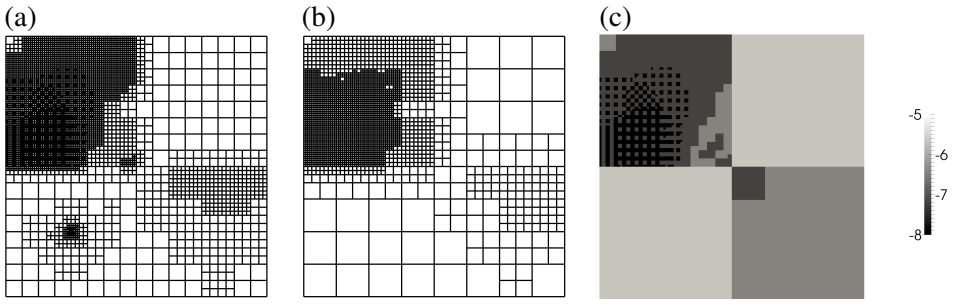

Figure 5. Refinement on cycle 6 (with 7 174 macro cells) on (a) the macroscale $\mathbb{T}_H(\Omega)$, (b) the sampling discretization $\mathbb{T}_\delta(\Omega)$, and (c) on the microscale discretization $\{h_K : K \in \mathbb{T}_\delta(\Omega)\}$ (logarithmically scaled).

Table 5. Refinement levels in number of cells on macro, micro and sampling discretization to achieve at least (a) 5% and (b) 1% error.

(a)	strategy	#macro	#micro	#sampl.	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$
	uniform	65 536	—	—	8.62e-2 (5.1%)
	local	35 377	—	—	1.06e-1 (6.2%)
	ad. sampl.	7 174	6 679	8 149	7.13e-2 (4.1%)

(b)	strategy	#macro	#micro	#sampl.	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$
	uniform	1 048 576	—	—	6.64e-3 (0.4%)
	local	568 201	—	—	1.19e-2 (0.7%)
	HMM	19 330	302 896	1 024	1.55e-2 (0.9%)
	ad. sampl.	93 245	89 017	89 500	1.12e-2 (0.7%)

4.2. Log-normally distributed permeability

In the second numerical example, we keep the computational domain Ω and the quantity of interest, but replace the periodic coefficient A^ε by a *log-normally distributed* permeability. The motivation behind this choice is that a large class of physically relevant permeabilities, such as occurring for example in ground-water flow,

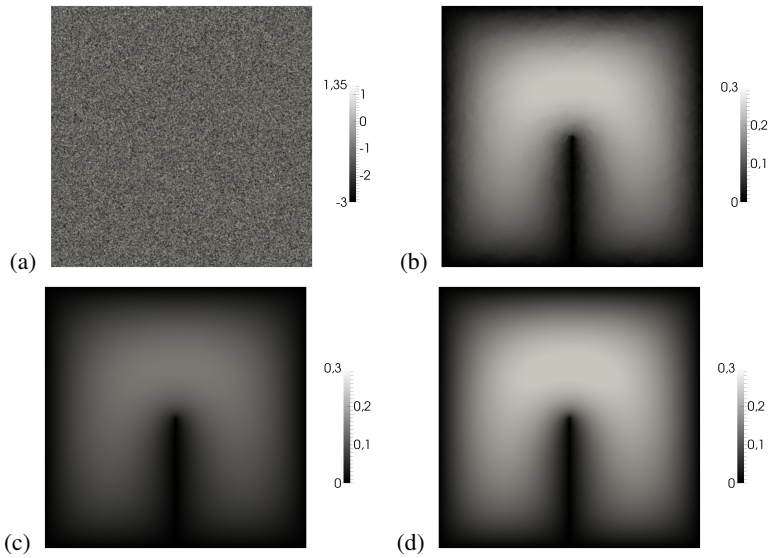

Figure 6. (a) Log-normally distributed permeability with Gaussian correlation shown in log-scale. (b) The corresponding reference solution. (c) Solution for *arithmetic* averaging, and (d) for *geometric* averaging.

exhibits a log-normal distribution (cf. Warren& Price [49]).

In detail, we choose A^ε to be

$$A^\varepsilon(x) := \gamma \exp(10\, g(x)\,/\,255)\, I_d\,, \tag{4.4}$$

where $g(x)$ is an 8 bit grayscale picture (with integral values between 0 and 255) with 1024×1024 pixels resolution (cf. Fig. 6). The grayscale picture is generated using the `QuantIm` library [48]. It is a (discrete) Gaussian random field with an additional Gaussian correlation with a correlation length chosen to be $r = 0.0025$. Due to the small correlation length and exponential scaling, the permeability exhibits a strong influence on the macroscale. A computation with $16\cdot 10^6$ degrees of freedom results in a reference value of $\partial_{x_2} u^\varepsilon_{\mathrm{ref}}(\hat{x}) \approx 1.168$.

4.2.1. Model error of averaging schemes and HMM. Again, we examine the different choices of effective models. Due to the absence of any good a priori knowledge on how to choose δ (except for the correlation length $r = 0.0025$), we set $Y^\delta_K := K$ and test for two different choices of δ, a coarse sampling mesh with $\delta = 2^{-3}$ and a fine version with $\delta = 2^{-5}$. Once again, a high resolution in macroscale discretization ($2.62\cdot 10^5$ cells) and microscale discretization (10^6 cells) is chosen to get an estimate for the model error. Table 6 shows that for the log-normally distributed permeability, none of the effective models reproduces the quantity of interest (which is a localized point value) – every single effective model has a relative error of $50-100\,\%$. However, a clear difference becomes visible when looking at the error in the L^2 norm. Here, both, the geometric averaging and the HMM have

a relative error of about 2 %, whereas the arithmetic averaging is over one order of magnitude worse. This is also clearly observable in the "picture norm", cf. Fig. 6.

Table 6. Model error in L^2 norm and in the quantity of interest for the different sampling schemes and the HMM.

strategy	$\delta = 2^{-3}$		$\delta = 2^{-5}$	
	L^2 error	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$	L^2 error	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$
arithmetic	6.85e-2 (40%)	9.40e-1	6.77e-2 (40%)	9.37e-1
geometric	3.66e-3 (2.2%)	7.93e-1	2.58e-3 (1.5%)	7.60e-1
harmonic	1.21e-1 (71%)	5.55e-1	1.15e-1 (67%)	4.21e-1
HMM	2.87e-3 (1.7%)	7.93e-1	2.23e-3 (1.3%)	7.57e-1

4.2.2. Full adaptive algorithm. As the last numerical test, we perform a full adaptation process with coarse initial macroscale discretization and coarse sampling mesh. In contrast to the first numerical example, a high initial resolution of the microscale with about 10^6 cells is chosen. This is done in order to avoid unnecessary refinement on macroscale and sampling, which turns out to happen if the microscale resolution is not sufficiently good. Given the fact that in case of averaging schemes a high microscale resolution does only introduce a high quadrature, such a choice is still computationally acceptable. A slightly different scaling is chosen: We set $\alpha_\delta = \alpha_h = 1$ and enforce a higher accuracy on the macroscale discretization by setting $\alpha_H = 20$. The results of the adaptation process are given in Table 7.

Table 7. Refinement history: error in the quantity of interest and error estimators for the adaptive sampling strategy with geometric averaging.

	#macro	#sampl.	#micro	$\lvert\langle j, u^\varepsilon - U\rangle\rvert$	$\lvert\vartheta^H\rvert$	$\lvert\vartheta^h\rvert$	$\lvert\vartheta^\delta\rvert$	I_{eff}
1	256	16	$\approx 10^6$	8.09e-1 (70%)	1.07e-2	1.96e-13	2.14e-1	0.25
3	388	37	$\approx 10^6$	7.94e-1 (68%)	7.56e-3	6.25e-15	3.14e-1	0.38
5	652	166	$\approx 10^6$	7.75e-1 (66%)	6.08e-3	2.98e-16	9.84e-2	0.13
7	1 285	760	$\approx 10^6$	5.22e-1 (45%)	1.22e-2	6.29e-15	2.12e-1	0.43
9	3 442	1 981	$\approx 10^6$	1.78e-1 (15%)	2.13e-2	7.47e-14	5.26e-1	2.83
11	11 788	8 581	$\approx 10^6$	1.77e-2 (1.5%)	1.14e-2	5.76e-13	2.59e-1	14.05
13	50 836	26 641	$\approx 10^6$	2.10e-2 (1.8%)	1.37e-2	5.80e-13	1.55e-1	6.73
15	191 368	62 290	$\approx 10^6$	1.61e-2 (1.4%)	9.66e-3	5.84e-13	9.24e-2	5.14
17	658 525	129 910	$\approx 10^6$	1.25e-2 (1.1%)	5.29e-3	6.00e-13	4.68e-2	4.88

Compared to uniform and local refinement with a standard finite element discretization a significant saving can be observed for macroscale and sampling discretization with a difference of almost two orders of magnitude in refinement (cf. Fig. 7

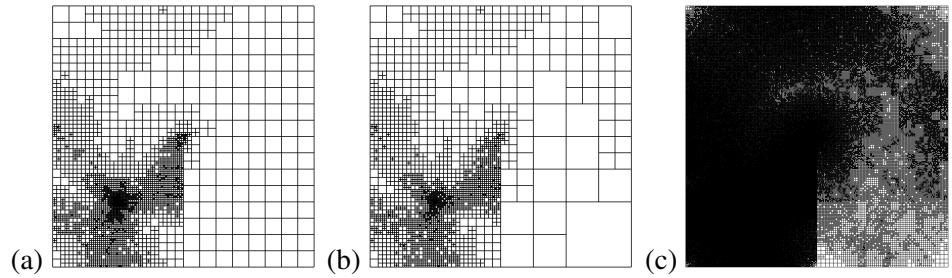

Figure 7. Refinement on cycle 10 (that reaches 2% accuracy) on (a) the macroscale $\mathbb{T}_H(\Omega)$ and (b) the sampling discretization $\mathbb{T}_\delta(\Omega)$. (c) shows a corresponding locally refined mesh to reach 2% accuracy with a direct finite element discretization.

and Fig. 8). The huge saving is due to the fact that the geometric average allows for a very localized refinement process in order to improve the accuracy in the quantity of interest. With this, a relative error of around 2% can be reached with very little coarse scale and sampling refinement. After that convergence stagnates and the adaptive sampling algorithm "degenerates" to local refinement (cf. Fig. 8).

5. Conclusion

In this paper a heterogeneous multiscale finite element method for elliptic problems is derived within the general framework of the DWR method. An a posteriori error analysis is presented for this scheme and independent error estimators as well as localized refinement indicators are derived. This provides the basis of a simultaneous adaptation algorithm for discretization and sampling. The proposed method is tested numerically at two prototypical classes of permeabilities (with mild and strong scale separation). A highly localized quantity of interest is chosen that is influenced equally by (global) macroscale and (local) microscale behavior. The sampling adaptation process shows good results in both cases with a significant reduction in degrees of freedom compared to standard finite element approaches and the classical HMM. Further, it turns out that (in analogy to the standard DWR case) the dual solution can be approximated with a coarse approximation based on the effective parameters. The qualitative character of the refinement indicators and the quantitative character of the error estimators is still sufficient for effective error control.

In principle, the derived error estimators allow for different sampling adaptation strategies; main subject of this paper is a simple averaging scheme based on the geometric mean value. A fundamentally different approach, in which the model-error estimators are used to directly enhance the effective model within a model optimization framework. will be the subject of a forthcoming paper.

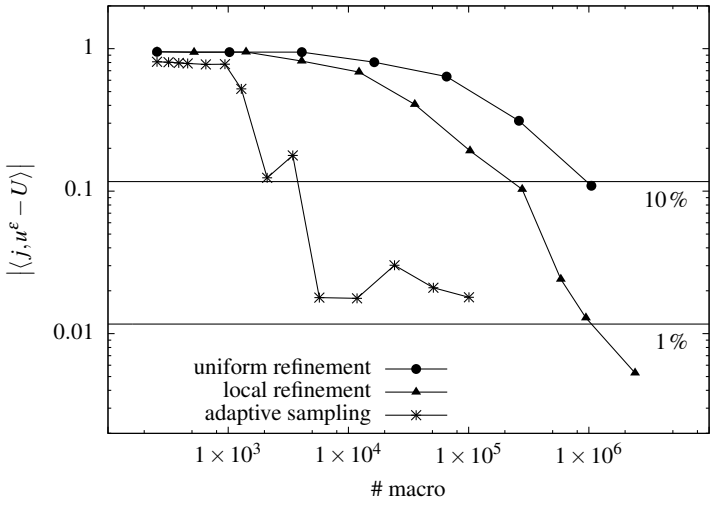

Figure 8. Performance plot showing the error in quantity of interest over macroscale discretization for uniform and local refinement with a standard finite element discretization and the adaptive sampling strategy with geometric averaging

References

1. A. Abdulle. On a priori error analysis of fully discrete heterogeneous multiscale fem. *SIAM Journal of Multiscale Modeling and Simulation*, 4(2):447–459, 2005.
2. A. Abdulle. A priori and a posteriori error analysis for numerical homogenization: A unified framework. In *Multiscale Problems*, volume 16 of *Series in Contemporary Applied Mathematics*, pages 280–305, 2011.
3. A. Abdulle and A. Nonnenmacher. A posteriori error estimate in quantities of interest for the finite element heterogeneous multiscale method. *Numerical Methods for Partial Differential Equations*, 29(5):1629–1656, 2013.
4. A. Abdulle and G. Vilmart. Analysis of the finite element heterogeneous multiscale method for quasilinear elliptic homogenization problems. *Mathematics of Computation*, 83:513–536, 2014.
5. R. A. Adams. *Sobolev Spaces*, *Pure and Applied Mathematics*, Vol. 140, 2003.
6. G. Allaire. Homogenization and two-scale convergence. *SIAM Journal on Mathematical Analysis*, 23(6):1482–1518, 1992.
7. T. Arbogast and K. J. Boyd. Subgrid upscaling and mixed multiscale finite elements. *Journal of Numerical Analysis*, 44(3):1150–1171, 2006.
8. I. Babuska. Homogenization and its application. Mathematical and computational problems. *Numerical solution of Partial Differential Equations III*, Proc. Third Sympos., :363–379, Univ. Maryland, 1975) Academic Press, New York, 1976.
9. W. Bangerth, T. Heister, L. Heltai, G. Kanschat, M. Kronbichler, M. Maier, B. Turcksin, and T. D. Young. The `deal.II` Library, version 8.1. *arXiv preprint,* `http://arxiv.org/abs/1312.2266v4`, 2013.
10. W. Bangerth and R. Rannacher. *Adaptive dinite element methods for differential equations*. Birkhäuser, 2003.

11. R. Becker and M. Braack. Multigrid techniques for finite elements on locally refined meshes. *Numerical Linear Algebra with Applications*, 7:363–379, 2000.

12. R. Becker and R. Rannacher. A feed-back approach to error control in finite element methods: Basic analysis and examples. *East-West Journal of Numerical Mathematics*, 4:237–264, 1996.

13. R. Becker and R. Rannacher. Weighted a posteriori error control in fe methods. In *Proceedings of ENUMATH-97*, pages 621–637, 1998. Lecture at ENUMATH-95, Paris, September 18-22, 1995.

14. R. Becker and R. Rannacher. An optimal control approach to a posteriori error estimation in finite element methods. *Acta Numerica*, 10:1–102, 2001.

15. A. Bensoussan, J.-L. Lions, and G. Papanicolaou. *Asymptotic Analysis for Periodic Structures*. Number 5 in Studies in Mathematics and Its Applications. North-Holland, 1 edition, 1978.

16. M. Braack and A. Ern. A posteriori control of modeling errors and discretization errors. *Multiscale Modeling and Simulation*, 1(2):221–238, 2003.

17. M. Braack and T. Richter. Solutions of 3D Navier-Stokes benchmark problems with adaptive finite element. *Computers and Fluids*, 35:37–392, 2006.

18. S. Brenner and L. R. Scott. *The Mathematical Theory of Finite Element Methods*. Springer, Berlin Heidelberg New York, 1994.

19. F. Brezzi. Interacting with the subgrid world. *Numerical Analysis*, pages 69–82, 1999.

20. W. T. Cardwell Jr. and R. L. Parsons. Average permeability of heterogeneous oil sands. *Transactions of the AIME*, 160(1):34–42, 1945.

21. G. F. Carey and J. T. Oden. *Finite Elements, Computational Aspects*. Vol. III. Prentice-Hall, 1984.

22. Z. Chen and T. Y. Hou. A mixed multiscale finite element method for elliptic problems with oscillating coefficients. *Mathematics of Computation*, 72(242):541–576, 2003.

23. P. G. Ciarlet. *The Finite Element Method for Elliptic Problems*. Classics Appl. Math. 40, SIAM, Philadelphia, 2002

24. D. Cioranescu and P. Donato. *An Introduction to Homogenization*. Number 17 in Oxford Lecture Series in Mathematics and Its Applications. Oxford University Press, 1 edition, 1999.

25. Ph. Clement. Approximation by finite element functions using local regularization. *RAIRO Analyse Numerique* 9:77–84, 1975.

26. W. E and B. Engquist. The heterogeneous multiscale methods. *Communications in Mathematical Sciences*, 1(1):87–132, 2003.

27. W. E and B. Engquist. Multiscale modeling and computation. *Notices of the American Mathematical Society*, 509):1062–1070, 2003b.

28. W. E, P. Ming, and P. Zhang. Analysis of the heterogeneous multiscale method for elliptic homogenization problems. *Journal of the American Mathematical Society*, 18:121–156, 2005.

29. Y. Efendiev, V. Ginting, and T. Y. Hou. Multiscale finite element methods for nonlinear problems and their applications. *Communications in Mathematical Sciences*, 2(4):553–589, 2004.

30. M. G. D. Geers, V. G. Kouznetsova, and W. A. M. Brekelmans. Multi-scale computational homogenization: Trends and challenges. *Journal of Computational and Applied Mathematics*, 234(7):2175–2182, 2010.

31. P. Henning and M. Ohlberger. A note on homogenization of advection-diffusion problems with large expected drift. *Zeitschrift fr Analysis und ihre Anwendungen*, 30(3):319–339, 2011.

32. P. Henning, M. Ohlberger, and B. Schweizer. An adaptive multiscale finite element method.

Technical Report 05/12 - N, FB 10 , Universitt Mnster, 2012.

33. R. Hill. On constitutive macro-variables for heterogeneous solids at finite strain. *Proceedings of the Royal Society*, A326:131–147, 1972.

34. T. Y. Hou and X.-H. Wu. A multiscale finite element method for elliptic problems in composite materials and porous media. *Journal of Computational Physics*, 134:169–189, 1997.

35. T. J. R. Hughes, G. R. Feijeo, L. Mazzei, and J.-B. Quincy. The variational multiscale method—a paradigm for computational mechanics. *Computer Methods in Applied Mechanics and Engineering*, 166(1-2):3–24, 1998.

36. D. Li, B. Beckner, and A. Kumar. Spe 56554 - a new efficient averaging technique for scaleup of multimillion-cell geologic models. *SPE Papers*, pages 495–510, 1999.

37. A.-M. Matache and C. Schwab. Generalized fem for homogenization problems. *Multiscale and Multiresolution Methods*, 20:197–237, 2002a.

38. A.-M. Matache and C. Schwab. Two-scale fem for homogenization problems. *Mathematical Modelling and Numerical Analysis*, 36(4):536–572, 2002b.

39. J. T. Oden and K. S. Vemaganti. Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. part i: Error estimates and adaptive algorithms. *Journal of Computational Physics*, 164:22–47, 2000a.

40. J. T. Oden and K. S. Vemaganti. Adaptive modeling of composite structures: Modeling error estimation. *International Journal for Civil and Structural Engineering*, 1:1–16, 2000b.

41. J. T. Oden and K. S. Vemaganti. Estimation of local modeling error and goal-oriented adaptive modeling of heterogeneous materials. part ii: A computational environment for adaptive modeling of heterogeneous elastic solids. *Computer Methods in Applied Mechanics and Engineering*, 190(46–47):6089–6124, 2001.

42. J. T. Oden, S. Prudhomme, A. Romkes, and P. T. Bauman. Multiscale modeling of physical phenomena: Adaptive control of models. *Siam Journal on Scientific Computing*, 28(6):2359–2389, 2006.

43. M. Ohlberger. A posteriori error estimates for the heterogeneous multiscale finite element method for elliptic homogenization problems. *Multiscale Modeling and Simulation*, 4(1):88–114, 2005.

44. T. Richter and T. Wick. Variational localizations of the dual weighted residual estimator. *Journal of Computational and Applied Mathematics*, to appear: 2014.

45. A. Romkes and T. C. Moody. Local goal-oriented estimation of modeling error for multi-scale modeling of heterogeneous elastic materials. *International Journal for Computational Methods in Engineering Science and Mechanics*, (4), 2007.

46. L. R. Scott and S. Zhang. Finite element interpolation of nonsmooth functions satisfying boundary conditions. *Math. Comp.*, 54:483–493, 1990.

47. R. Verfürth.. *A Review of A Posteriori Error Estimation and Adaptive Mesh-Refine- ment Techniques*. John Wiley/Teubner, 1996.

48. H.-J. Vogel. `QuantIm`, C/C++ library for scientific image processing, version 4.01. *`http://www.ufz.de/index.php?en=16562`*, 2008.

49. J. E. Warren and H. S. Price. Flow in heterogeneous porous media. *Society of Petroleum Engineers Journal*, 1(3):153–169, 1961.