A priori error analysis of a numerical stochastic homogenization method

Julian Fischer, Dietmar Gallistl, Daniel Peterseim

Angaben zur Veröffentlichung / Publication details:

Fischer, Julian, Dietmar Gallistl, and Daniel Peterseim. 2021. "A priori error analysis of a numerical stochastic homogenization method." *SIAM Journal on Numerical Analysis* 59 (2): 660–74. https://doi.org/10.1137/19m1308992.

SIAM J. Numer. Anal.
Vol. 59, No. 2, pp. 660–674

A PRIORI ERROR ANALYSIS OF A NUMERICAL STOCHASTIC HOMOGENIZATION METHOD*

JULIAN FISCHER†, DIETMAR GALLISTL‡, AND DANIEL PETERSEIM§

Abstract. This paper provides an a priori error analysis of a localized orthogonal decomposition method for the numerical stochastic homogenization of a model random diffusion problem. If the uniformly elliptic and bounded random coefficient field of the model problem is stationary and satisfies a quantitative decorrelation assumption in the form of the spectral gap inequality, then the expected L^2 error of the method can be estimated, up to logarithmic factors, by $H + (\varepsilon/H)^{d/2}$, ε being the small correlation length of the random coefficient and H the width of the coarse finite element mesh that determines the spatial resolution. The proof bridges recent results of numerical homogenization and quantitative stochastic homogenization.

Key words. numerical homogenization, stochastic homogenization, quantitative theory, a priori error estimates, uncertainty, model reduction

AMS subject classifications. 35R60, 65N12, 65N15, 65N30, 73B27, 74Q05

DOI. 10.1137/19M1308992

1. Introduction.

We study a prototypical random diffusion problem

$$-\operatorname{div} \mathbf{A}\nabla\mathbf{u} = f$$

with homogeneous Dirichlet boundary conditions on a bounded Lipschitz polytope. The diffusion tensor $\mathbf{A}$ is a random coefficient field with short correlation length $\varepsilon > 0$. We are interested in the approximation of this random PDE by a deterministic finite element model and corresponding estimates of the expected L^2 error. The approximation is based on the localized orthogonal decomposition (LOD) approach to numerical homogenization beyond scale separation and periodicity [35, 31, 40]. This method is well established for deterministic applications ranging from nonlinear, time-dependent, multiphysics problems [30, 1, 36, 44] to the stabilization of numerical wave scattering [41, 21, 42]. Apart from possible reinterpretations of the approach in the frameworks of domain decomposition [33, 43] and Bayesian inference [38, 39], the method can be rephrased as a discrete nonlocal integral operator with a piecewise constant and exponentially decaying integral kernel, thereby connecting the approach to the mathematical theory of homogenization [22]. This particular perspective extends to the present stochastic homogenization problem in a natural way. In [23] it is shown that the expectation of the discrete nonlocal integral representations of the realizations of the random operator provides an approximation of the stochastically homogenized operator. The error bounds of this LOD approach to stochastic homogenization contains the typical a priori terms for the spatial discretization, quantified by the mesh size (or observation scale) H, and an a posteriori estimator that represents

*Received by the editors December 27, 2019; accepted for publication (in revised form) November 30, 2020; published electronically March 9, 2021.
https://doi.org/10.1137/19M1308992
Funding: The work of the third author was supported by the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program grant 865751.
†Institute of Science and Technology Austria (IST Austria), Am Campus 1, 3400 Klosterneuburg, Austria (julian.fischer@mis.mpg.de).
‡Institut für Mathematik, Universität Jena, 07737 Jena, Germany (dietmar.gallistl@uni-jena.de).
§Institut für Mathematik, Universität Augsburg, 86159 Augsburg, Germany (daniel.peterseim@math.uni-augsburg.de).

© 2021 J. Fischer, D. Gallistl, and D. Peterseim

local fluctuations of the deterministic upscaled model. Without any assumptions on the statistical structure of the coefficient $\mathbf{A}$, the a priori quantification of the statistical error estimator seems hardly possible. However, the numerical experiments of [23] revealed that small values of the estimator are achieved given a certain scale separation in the stochastic variable, in particular for random coefficients at finite correlation length ε. This paper makes this plausible observation rigorous in an a priori error analysis that is explicit in H and ε.

The key tools for our present paper are adopted from the recent quantitative theory of homogenization, in particular the framework of functional inequalities from [26, 27, 25, 24] and the regularity theory from [10, 9, 24, 5, 20, 16]. We recall that the quantitative theory of stochastic homogenization has recently led to optimal-order convergence rates for linear elliptic PDEs with random coefficient fields [26, 27, 8, 28], as well as to a corresponding result for monotone operators [19]. For (nonoptimal) convergence rates for further nonlinear problems, we refer the reader to [3, 4, 6]. An overview of computational methods in stochastic homogenization can be found in the review article [2]; see also [7, 29]. Numerical approaches to the computation of effective coefficients in stochastic homogenization have been devised, e.g., in [13, 14, 32, 37]; of particular interest in this context are variance reduction schemes; see [34, 12] for several methods capable of substantially reducing the computational cost and [18] for a theoretical analysis.

Altogether, by merging the theories of LOD and quantitative stochastic homogenization, we achieve rigorous a priori error bounds for a numerical stochastic homogenization in the spirit of LOD. If the uniformly elliptic and bounded random coefficient field $\mathbf{A}$ is stationary and satisfies a quantitative decorrelation assumption in the form of the spectral gap inequality, then the numerical deterministic approximation u_{H} of the random solution field $\mathbf{u}$ fulfills the relative error bound

$$\frac{\mathbb{E}\big[\|\mathbf{u}-u_{\mathsf{H}}\|^2_{\mathsf{L}^2(\mathsf{D})}\big]^{1/2}}{\|f\|_{\mathsf{L}^2(\mathsf{D})}} \lesssim |\log H|^{4+3d/2}\left(H+\left(\frac{\varepsilon}{H}\right)^{d/2}\right). \tag{1.1}$$

Recall that $H > 0$ refers to the mesh size of some possibly coarse simplicial finite element mesh T_{H} that underlies the LOD construction and that ε is the correlation length of $\mathbf{A}$. Estimate (1.1) appears to be optimal in the sense of spatial approximability and CLT scaling (up to the logarithmic factor which is most probably pessimistic). This bound is in agreement with the numerical experiments of [23] for a relevant class of random coefficients in the regime of short-range correlation. We shall emphasize that the method of [23] itself is applicable without the structural assumptions of stationarity and quantitative decorrelation. However, the accuracy of a deterministic approximation of the random solution field is very limited beyond such assumptions.

Apart from the mathematical justification of LOD for stochastic homogenization problems, the numerical analysis of this paper may have an impact on the practical realization of more general multiscale representations of homogenized operators in stochastic homogenization [17]. Moreover, given the aforementioned generalizations of both LOD and the analytical techniques, the present work may be the starting point for the numerical analysis of more involved stochastic homogenization problems beyond the prototypical linear elliptic model problem.

The remaining parts of this paper are structured as follows. Section 2 specifies the model problem, and section 3 reviews the numerical stochastic homogenization method of [23]. Section 4 characterizes the admissible class of random diffusion coefficients and presents and proves the main results.

Standard notation on Lebesgue and Sobolev spaces applies throughout this paper. The notation $a \lesssim b$ abbreviates $a \leq Cb$ for some constant C that is independent of the mesh size and variations of the coefficient $\mathbf{A}$ but may depend on the shape of mesh elements and the contrast (i.e., the ratio of the uniform upper and lower bound) of $\mathbf{A}$ (cf. Assumption (A1) of subsection 4.1). The notation $a \approx b$ abbreviates $a \lesssim b \lesssim a$. For a matrix $\mathbf{A}$, we use the notation $\mathbf{A}^*$ to denote its transpose. The duality product of H^{-1} and H_0^1 is denoted by $\langle \cdot, \cdot \rangle$.

2. Model problem. Let $(\Omega, \mathsf{F}, \mathbb{P})$ be a probability space with set of events Ω, σ-algebra $\mathsf{F} \subseteq 2^\Omega$, and probability measure $\mathbb{P}$. The expectation operator is denoted by $\mathbb{E}$. Let $D \subseteq \mathbb{R}^\mathsf{d}$ for $d \in \{1, 2, 3\}$ be a bounded Lipschitz polytope with a diameter of order 1. For technical reasons in our proofs we assume that D is a cuboid. Let $\mathbf{A}$ be a uniformly elliptic and bounded random coefficient field, and let, for the sake of readability, $\mathbf{A}$ be pointwise symmetric. The proofs can, however, be extended to the unsymmetric case. For a deterministic right-hand side $f \in L^2(D)$, the model problem reads

$$\left\{ \begin{aligned} -\operatorname{div}(\mathbf{A}(\omega)(x)\nabla \mathbf{u}(\omega)(x)) &= f(x), \qquad x \in D \\ \mathbf{u}(\omega)(x) &= 0, \qquad x \in \partial D \end{aligned} \right\} \quad \text{for almost all } \omega \in \Omega. \tag{2.1}$$

The weak formulation of (2.1) is based on the Sobolev space $V := H_0^1(D)$ and seeks a random field $\mathbf{u}$ in the Hilbert space $L^2(\Omega; V)$ such that

$$\int_\Omega \int_\mathsf{D} (\mathbf{A}(\omega)\nabla \mathbf{u}(\omega)(x)) \cdot \nabla \mathbf{v}(\omega)(x)\, dx\, d\mathbb{P}(\omega) = \int_\Omega \int_\mathsf{D} f(x)\mathbf{v}(\omega)(x)\, dx\, d\mathbb{P}(\omega) \tag{2.2}$$

holds for all $\mathbf{v} \in L^2(\Omega; V)$. Well-posedness of this problem follows from coercivity of the bilinear form on the left-hand side.

The numerical stochastic homogenization method introduced below can be applied to the model problem without further statistical assumptions on the diffusion coefficient $\mathbf{A}$. However, its a priori error analysis based on the quantitative theory of stochastic homogenization will require the restriction to the class of stationary random coefficient fields $\mathbf{A}$ satisfying a spectral gap inequality. These structural assumptions will be made specific in subsection 4.1.

3. Numerical stochastic homogenization method. This section reviews the numerical stochastic homogenization method of [23]. This requires the introduction of some basic notation on finite element spaces.

3.1. Finite element notation. Let $\mathsf{T_H}$ denote a quasi-uniform and regular simplicial triangulation of the domain D. Introduce the global mesh size $H := \max\{\operatorname{diam}(T) : T \in \mathsf{T_H}\}$ of the quasi-uniform mesh $\mathsf{T_H}$. The corresponding P_1 finite element space of piecewise affine and globally continuous functions that satisfy the homogeneous Dirichlet boundary condition is denoted by $V_\mathsf{H} \subseteq V$. This space will be used for the approximation of the solution. That is why we refer also to the corresponding discretization scale H as the observation scale. For the approximation of integral kernels that define the numerical homogenization method, we will use the space of piecewise constant functions (resp., $d \times d$ matrix fields), which is denoted by $P_0(\mathsf{T_H})$ (resp., $P_0(\mathsf{T_H}; \mathbb{R}^{\mathsf{d}\times\mathsf{d}})$).

The definition of localized numerical correctors requires the concept of patches. The neighborhood (or first-order patch) of a given subdomain $S \subseteq \overline{D}$ is defined as

$$\mathsf{N}(S) := \operatorname{int}\left(\bigcup \{T \in \mathsf{T_H} : T \cap \overline{S} = \emptyset\} \right).$$

Furthermore, we introduce for any $\ell \geq 2$ the patch extensions

$$\mathsf{N}^1(S) := \mathsf{N}(S) \qquad \text{and} \qquad \mathsf{N}^\ell(S) := \mathsf{N}(\mathsf{N}^{\ell-1}(S)).$$

Note that the number of elements in the ℓth-order patch in a quasi-uniform mesh scales like ℓ^d.

Let $I_H : V \to V_H$ be a surjective quasi-interpolation operator that acts as an H^1-stable and L^2-stable quasi-local projection in the sense that $I_H \circ I_H = I_H$ and that for any $T \in \mathsf{T}_H$ and all $v \in V$ there holds

$$H^{-1}\|v - I_H v\|_{L^2(T)} + \|\nabla I_H v\|_{L^2(T)} \leq C_{I_H} \|\nabla v\|_{L^2(\mathsf{N}(T))} \tag{3.1}$$

$$\|I_H v\|_{L^2(T)} \leq C_{I_H} \|v\|_{L^2(\mathsf{N}(T))}. \tag{3.2}$$

For the discussion in this paper, we choose I_H to be the concatenation of the L^2 projection to (possibly discontinuous) piecewise affine functions over T_H and the averaging operator that maps a piecewise affine function p_H to V_H by assigning to each interior vertex z the average of all values $p_H|_T(z)$ such that $z \in T \in \mathsf{T}_H$ [23]. We note that various other choices are possible.

Remark 3.1. Let $\tilde{\mathsf{T}}_H$ be the triangulation generated from T_H by two barycentric refinements, and let $\tilde{V}_H \subseteq V$ denote the first-order finite element space with respect to $\tilde{\mathsf{T}}_H$. For the above choice of I_H, it is easy to see that for any $v \in V$ there exists $\tilde{v} \in \tilde{V}_H$ such that $I_H v = I_H \tilde{v}$ and $\|\tilde{v}\|_{L^2(\Omega)} \lesssim \|v\|_{L^2(\Omega)}$. This remains true if I_H is replaced by the operator on $H^1(\Omega)$ that averages at boundary vertices as well (i.e., not enforcing Dirichlet conditions). The claim follows from the fact that any piecewise affine p_H can be generated by the L^2 projection of a suitable linear combination of bubble functions from $\tilde{V}_H$.

3.2. Numerical stochastic homogenization method. The LOD approach of [23] to stochastic homogenization computes a quasi-local effective coefficient as a discrete integral operator on finite element spaces. Its construction is described in the following steps.

Following [35], coarse and fine scales are characterized through the quasi-interpolation operator I_H introduced above. The space W of fine-scale functions is defined by $W := \ker I_H \subseteq V$. Given a nonnegative integer *oversampling parameter* ℓ, which throughout this paper is assumed to satisfy $\ell \approx |\log H|$, consider the ℓth-order extended patch $D_T := \mathsf{N}^\ell(T)$ of an element $T \in \mathsf{T}_H$. The space of fine-scale functions that vanish outside D_T is denoted by $W_{D_T} \subseteq W$. Note that this choice encodes a homogeneous Dirichlet boundary condition at the boundary of D_T.

Given the jth Cartesian unit vector e_j $(j = 1, \ldots, d)$, the localized element corrector $\mathbf{q}_{T;j} \in L^2(\Omega; W_{D_T})$ related to the element $T \in \mathsf{T}_H$ is defined as the solution to the following localized problem (cell problem):

$$\int_{D_T} \nabla w \cdot (\mathbf{A}\nabla \mathbf{q}_{T;j})\, dx = \int_T \nabla w \cdot (\mathbf{A} e_j)\, dx \quad \text{for all } w \in W_{D_T}. \tag{3.3}$$

Given $v_H \in V_H$, we define the correction operator $\mathfrak{C} v_H \in L^2(\Omega; W)$ by

$$\mathfrak{C} v_H = \sum_{T \in \mathsf{T}_H} \sum_{j=1}^{d} (\partial_j v_H|_T) \mathbf{q}_{T;j}. \tag{3.4}$$

Note that the element correctors and the correction operator implicitly depend on the parameter ℓ. In the deterministic case it was shown in [22] how the use of corrected test functions leads to a sparse discrete integral operator. In the stochastic setting, a similar representation with a stochastic integral kernel is possible [23], namely, with the piecewise-in-space constant matrix field $\boldsymbol{\mathcal{A}}_H \in L^2(\Omega; P_0(\mathsf{T}_H \times \mathsf{T}_H; \mathbb{R}^{d\times d}))$ over $\mathsf{T}_H \times \mathsf{T}_H$, which, for $T, K \in \mathsf{T}_H$, is defined by

$$(3.5)\qquad (\boldsymbol{\mathcal{A}}_H|_{T;K})_{jk} := \frac{1}{|T|\,|K|}\left(\delta_{T;K}\int_T \mathbf{A}_{jk}\,dx - e_j \cdot \int_K \mathbf{A}\nabla \mathbf{q}_{T;k}\,dx\right)$$

$(j, k = 1, \dots, d)$, where δ is the Kronecker symbol. Note that the operator $\boldsymbol{\mathcal{A}}_H$ is sparse in the sense that $\boldsymbol{\mathcal{A}}_H|_{T;K}$ equals zero for $T, K \in \mathsf{T}_H$ whenever $K \notin \mathsf{N}^{\ell}(T)$, i.e., $\operatorname{dist}(T, K) \gtrsim \ell H$.

The kernel $\boldsymbol{\mathcal{A}}_H$ induces the discrete bilinear form $\mathfrak{a} : V_H \times V_H \to L^2(\Omega; \mathbb{R})$ given by

$$\mathfrak{a}(v_H, z_H) := \int_D \int_D \nabla v_H(x) \cdot (\boldsymbol{\mathcal{A}}_H(x, y)\nabla z_H(y))\,dy\,dx$$

for any $v_H, z_H \in V_H$. V-coercivity and continuity of the form $\mathfrak{a}$ for any $\omega \in \Omega$ under the condition $\ell \approx \mathsf{O}(|\log H|)$ were shown in [22].

As pointed out in [22], there holds for all finite element functions $v_H, z_H \in V_H$ that

$$(3.6)\qquad \int_D \nabla v_H \cdot (\mathbf{A}\nabla(1 - \mathfrak{C})z_H)\,dx = \mathsf{a}(v_H, z_H).$$

This shows how the form $\mathfrak{a}$ is connected to a Petrov–Galerkin variant of the method of [35, 22].

The final approximation by a deterministic model is based on the averaged integral kernel $\bar{\mathsf{A}}_H := \mathbb{E}[\boldsymbol{\mathcal{A}}_H]$, i.e.,

$$(3.7)\qquad (\bar{\mathsf{A}}_H|_{T;K})_{jk} = \frac{1}{|T|\,|K|}\left(\delta_{T;K}\int_T \mathbb{E}[\mathbf{A}_{jk}]\,dx - e_j \cdot \int_K \mathbb{E}[\mathbf{A}\nabla \mathbf{q}_{T;k}]\,dx\right)$$

for any two simplices $T, K \in \mathsf{T}_H$. The corresponding deterministic bilinear form $\bar{\mathsf{a}}(\cdot,\cdot)$ is given by

$$\bar{\mathsf{a}}(v_H, z_H) := \int_D \int_D \nabla v_H(x) \cdot (\bar{\mathsf{A}}_H(x, y)\nabla z_H(y))\,dy\,dx \quad \text{for any } v_H, z_H \in V_H.$$

Given this discrete deterministic approximation of the random partial differential operator, an approximation $u_H \in V_H$ of the solution $\mathbf{u}$ in the coarse finite element space V_H solves

$$(3.8)\qquad \bar{\mathsf{a}}(u_H, v_H) = (f, v_H)_{L^2(D)} \quad \text{for all } v_H \in V_H.$$

This can be phrased as a sparse linear system using the canonical nodal basis of the finite element space. Compared to the direct finite element discretization of (2.2), this system is slightly denser because the degrees of freedom associated with the interior vertices of T_H are directly coupled over distances of order ℓH. In this regard the system is similar to a B-spline discretization of order $\ell + 1$.

Altogether, the numerical stochastic homogenization method consists of two steps. The first step is the assembling of the system matrix associated with (3.8), which in turn requires the solution of the $d \times \operatorname{card} \mathsf{T}_H$ cell problems (3.3). This task is often referred to as the offline phase, which is independent of the right-hand side. In analogy to periodic deterministic coefficients [21], stationarity plus an appropriately chosen structured mesh T_H allow one to reduce the number of cell problems to $\mathsf{O}(\ell^\mathsf{d})$ (namely, $\mathsf{O}(1)$ representative interior problems plus all representative intersections of patches with the domain boundaries). We shall emphasize in this connection that, in contrast to analytical approaches to homogenization, the numerical method depends on the domain D through the mesh and the boundary condition encoded in the cell problems. This dependence can be eliminated by replacing $\bar{\mathsf{A}}_\mathsf{H}$ with a Toeplitz matrix resulting from solving a representative interior cell problem. In this spirit, one may as well approximate $\bar{\mathsf{A}}_\mathsf{H}$ by a diagonal matrix resulting from row averaging to recover a classical finite element system that can be interpreted as the discretization of the homogenized PDE in certain cases [22]. However, simplification steps are beyond rigorous a priori error control and will not be discussed further here.

4. Error analysis. This section presents the novel a priori error analysis that combines arguments from the theories of LOD and quantitative stochastic homogenization. This requires some structural assumptions on the underlying random diffusion field.

4.1. Key assumptions. We will impose three structural assumptions on the random coefficient field $\mathbf{A}$: uniform ellipticity and boundedness, stationarity, and quantitative decorrelation. These conditions are classical in stochastic homogenization (see, for instance, [24]):

(A1) The random coefficient field $\mathbf{A}$ is uniformly elliptic and bounded; i.e., there exist constants $0 < \lambda \leq \Lambda < \infty$ such that almost surely we have $\mathbf{A}(x)v \cdot v \geq \lambda|v|^2$ and $|\mathbf{A}(x)v| \leq \Lambda|v|$ for every $v \in \mathbb{R}^\mathsf{d}$ and almost every $x \in \mathbb{R}^\mathsf{d}$.

(A2) The random coefficient field $\mathbf{A}$ is *stationary*; i.e., the law of shifted coefficient field $\mathbf{A}(\omega)(\cdot + x)$ coincides with the law of $\mathbf{A}$ for all $x \in \mathbb{R}^\mathsf{d}$.

(A3) The random coefficient field $\mathbf{A}$ is subject to a quantitative decorrelation assumption on scales larger than ε in the form of the spectral gap inequality with correlation length $\varepsilon > 0$; i.e., there exists a constant $\rho > 0$ such that for any Fréchet differentiable random variable $F = F(\mathbf{A})$ the estimate

$$\mathbb{E}\left[\left|F - \mathbb{E}[F]\right|^2\right] \leq \frac{\varepsilon^\mathsf{d}}{\rho}\mathbb{E}\left[\int_{\mathsf{R}^d} \quad {\scriptstyle \mathsf{B}_{\varepsilon}(\mathsf{x})} \; \frac{\partial F}{\partial \mathbf{A}}(\tilde{x}) \; d\tilde{x} \;^2 \; dx\right] \tag{4.1}$$

holds.

One example of a coefficient field satisfying the assumptions (A1)–(A3) are coefficient fields arising by applying a nonlinear function to a stationary Gaussian random field with integrable correlations. To be more explicit, let $k \in \mathbb{N}$, and let $Y : \Omega \times \mathbb{R}^\mathsf{d} \to \mathbb{R}^\mathsf{k}$ be a stationary Gaussian random field with integrable correlations in the sense that

$$\int_{\mathsf{R}^d} \sup_{|\tilde{\mathsf{x}}|=|\mathsf{x}|} \quad \mathrm{Cov}[Y(\tilde{x}), Y(0)]\,\big|\, dx \lesssim \varepsilon^\mathsf{d}.$$

Furthermore, let $\xi : \mathbb{R}^\mathsf{k} \to \mathbb{R}^{\mathsf{d}\times\mathsf{d}}$ be a 1-Lipschitz function taking values in the space of matrices subject to the uniform ellipticity and boundedness conditions in (A1). Then the random field

$$\mathbf{A}(\omega)(x) := \xi(Y(\omega, x))$$

satisfies the conditions (A1)–(A3) for some constant $\rho \gtrsim 1$.

Remark 4.1. For simplicity, we assume uniform ellipticity and boundedness of the coefficient field (condition (A1)). Since the ellipticity ratio of log-Gaussian random fields satisfies strong stochastic moment bounds, we believe that similar results may be deduced in the case of log-Gaussian random fields or, more generally, random fields with moment bounds on $\mathbb{E}[|A|^p] + \mathbb{E}[|A^{-1}|^p]$ for $p \gg 1$ by using an adaptation of our strategy.

4.2. Review of a posteriori error bounds. We briefly review the L^2 error estimates from [23] for the numerical method of the previous section, which mark the starting point for the novel a priori error analysis. These estimates require solely Assumption (A1) and no statistical assumptions. This generality results in an error estimate that contains an a posteriori term reflecting statistical errors that cannot be quantified a priori without further assumptions.

The error measure of interest is the $L^2(\Omega; L^2(D))$ norm. Besides the usual explicit convergence rates in terms of the mesh size H, the error bound contains an a priori quantity called worst-case best-approximation error, defined by

$$\mathbf{wcba}(\mathbf{A}(\omega), \mathsf{T}_H) := \sup_{g \in L^2(D)\setminus\{0\}} \inf_{v_H \in V_H} \frac{\|u(g, \mathbf{A}(\omega)) - v_H\|_{L^2(D)}}{\|g\|_{L^2(D)}} \lesssim H, \tag{4.2}$$

where for $g \in L^2(D)$, $u(g, \mathbf{A}(\omega)) \in V$ solves the deterministic model problem with diffusion coefficient $\mathbf{A}(\omega)$ and right-hand side g. This quantity is always controlled from above by H, but it can behave better (up to H^2) in certain regimes [22].

The a posteriori part in the error bound is referred to as model error estimator.

Definition 4.2 (model error estimator). *For any $T \in \mathsf{T}_H$, denote*

$$X(T) := \max_{\substack{K \in \mathsf{T}_H \\ K \cap \mathsf{N}^\ell(T) = \emptyset}} |T| \, |\mathcal{A}_H|_{T;K} - \bar{\mathsf{A}}_H|_{T;K}|.$$

The model error estimator γ is defined by

$$\gamma := \max_{T \in \mathsf{T}_H} \sqrt{\mathbb{E}[X(T)^2]}.$$

The model error estimator γ coincides with the one introduced in [23] up to some scaling factor that was used to improve the efficiency of the estimator in computations. Since the rescaling has no effect on the a priori error analysis, it is not considered here. The following error estimate was shown in [23, Proposition 9].

Proposition 4.3 (error estimate for the quasilocal method). *Let $\ell \approx |\log H|$. Let $\mathbf{u}$ solve* (2.2)*, and let u_H solve* (3.8) *with right-hand side $f \in L^2(D)$. Then the estimate*

$$\begin{aligned} \sqrt{\mathbb{E}[\|\mathbf{u} - u_H\|^2_{L^2(D)}]} &\lesssim (H^2 + \mathbb{E}[\mathbf{wcba}(\mathbf{A}, \mathsf{T}_H)] + \ell^d \gamma)\|f\|_{L^2(D)} \\ &\lesssim (H + \ell^d \gamma)\|f\|_{L^2(D)} \end{aligned} \tag{4.3}$$

holds with the model error estimator γ from Definition 4.2.

We end this paragraph by noting a technical perturbation result that will later be used in the proof of Theorem 4.7.

Lemma 4.4. *Let $T, K \in \mathsf{T}_H$ and $j, k \in \{1, \ldots, d\}$. Then there exist a box $Q \subseteq D_T$ and some $m \lesssim \ell$ such that the patches satisfy the inclusion $\mathsf{N}^\ell(T) \subseteq \overline{Q} \subseteq \overline{\mathsf{N}^m(T)}$. Let*

$\mathbf{q}^Q_{T;j} \in W_Q$ *solve* (3.3) *with* D_T *replaced by* Q, *and let* $(\mathcal{A}^Q_H|_{T;K})_{jk}$ *be defined by* (3.5) *with* $\mathbf{q}_{T;j}$ *replaced by* $\mathbf{q}^Q_{T;j}$. *Then the following perturbation result holds almost surely:*

$$|(\mathcal{A}_H|_{T;K})_{jk} - (\mathcal{A}^Q_H|_{T;K})_{jk}| \lesssim \frac{H}{|T|}.$$

Proof. Denote $\widehat{D}_T := \mathsf{N}^m(T)$ and as before $D_T = \mathsf{N}^\ell(T)$. The claimed inclusion relation follows from the quasi-uniformity of $\mathcal{T}_H$ and the assumption that the domain D is rectangular. From the definition (3.5) applied to D_T and Q and the Hölder inequality, it follows that

$$|(\mathcal{A}_H|_{T;K})_{jk} - (\mathcal{A}^Q_H|_{T;K})_{jk}| \lesssim |T|^{-1}|K|^{-1/2}\|\nabla(\mathbf{q}_{T;j} - \mathbf{q}^Q_{T;j})\|_{L^2(D_T)}.$$

By extending $\mathbf{q}_{T;j}$ and $\mathbf{q}^Q_{T;j}$ by zero to functions from $W_{\widehat{D}_T}$ and noting that both functions are Galerkin projections of the corrector $\widehat{\mathbf{q}}_{T;j} \in W_{\widehat{D}_T}$ defined through (3.3) with respect to $\widehat{D}_T$, we deduce

$$\|\nabla(\mathbf{q}_{T;j} - \mathbf{q}^Q_{T;j})\|_{L^2(D_T)} \lesssim \|\nabla(\mathbf{q}_{T;j} - \widehat{\mathbf{q}}_{T;j})\|_{L^2(\widehat{D}_T)}.$$

By an application of the exponential decay argument from [35, 31], it can be shown that the right-hand side is controlled by $H|T|^{1/2}$. The combination of the foregoing estimates and the shape regularity yield the assertion. ∎

4.3. Main result. The central result of this paper is the a priori quantification of the a posteriori model estimator under the assumptions of stationarity and quantitative decorrelation.

Theorem 4.5 (a priori error estimate for γ). *Let the diffusion tensor* $\mathbf{A}$ *satisfy assumptions* (A1)–(A3) *from subsection* 4.1, *and let* $\ell \approx |\log H|$. *Then* γ *from Definition* 4.2 *satisfies*

$$\gamma \lesssim |\log H|^{4+d/2}\left(H + \left(\frac{\varepsilon}{H}\right)^{d/2}\right).$$

Section 4.4 below is devoted to the proof of this theorem. The combination of Proposition 4.3 and Theorem 4.5 readily yields the desired a priori error bound of the numerical stochastic homogenization method of section 3.

Corollary 4.6 (a priori error estimate for the numerical method). *Let the diffusion tensor* $\mathbf{A}$ *satisfy assumptions* (A1)–(A3) *from subsection* 4.1, *let* $\ell \approx |\log H|$, *let* $\mathbf{u}$ *solve* (2.2), *and let* u_H *solve* (3.8) *with right-hand side* $f \in L^2(D)$. *Then*

$$\sqrt{\mathbb{E}[\|\mathbf{u} - u_H\|^2_{L^2(D)}]} \lesssim |\log H|^{4+3d/2}\left(H + \left(\frac{\varepsilon}{H}\right)^{d/2}\right)\|f\|_{L^2(D)}.$$

Note that the logarithmic factor $|\log H|^{4+3d/2}$ in the preceding estimate is likely nonoptimal; for instance, deriving and using sharper bounds on $|\nabla \mathbf{q}_{T;k}|$ that reflect the exponential decay of $\mathbf{q}_{T;k}$ outside of T in the proof below would give rise to an improved estimate.

4.4. Proof of the main result. This section is devoted to the proof of Theorem 4.5. Subsections 4.4.1 and 4.4.2 provide the necessary variance bounds for the entries of $\mathcal{A}_H$. The final subsection 4.4.3 concludes the proof of Theorem 4.5.

4.4.1. Variance bounds for the entries of $\mathcal{A}_H$.

Theorem 4.7. *Let* $\mathbf{A}$ *be a random coefficient field subject to the assumptions* (A1)–(A3). *Then the entries of the upscaled operator* $\mathcal{A}_H$ *defined in* (3.5) *satisfy the variance estimate*

$$\mathbb{E}\Big[\Big(\mathcal{A}_H|_{T;K} - \bar{A}_H|_{T;K}\Big)^2\Big] \lesssim \frac{H^2}{|T|^2} + \frac{\ell^8}{|T|^2}\Big(\frac{\varepsilon}{H}\Big)^d.$$

Proof. We apply Lemma 4.4 and assume that D_T is a box and note that the error from this replacement is controlled by H. Our goal is to estimate the variance of $\mathcal{A}_H|_{T;K}$ by means of the spectral gap inequality (4.1). To do so, we need to bound the Fréchet derivative of $\mathcal{A}_H|_{T;K}$. By (3.5) we have

$$\frac{\partial(\mathcal{A}_H|_{T;K})_{jk}}{\partial\mathbf{A}}(\delta\mathbf{A}) = \frac{1}{|T|\,|K|}\Big(\delta_{T;K}\int_T (\delta\mathbf{A})_{jk}\,dx - e_j\cdot\int_K \delta\mathbf{A}\nabla\mathbf{q}_{T;k}\,dx - e_j\cdot\int_K \mathbf{A}\nabla\frac{\partial\mathbf{q}_{T;k}}{\partial\mathbf{A}}(\delta\mathbf{A})\,dx\Big).$$

We define the auxiliary functions ${}_{T;K;j} \in W_{D_T}$ as the unique solution to the equation

$$\int_{D_T}\nabla w\cdot(\mathbf{A}^*\nabla\; {}_{T;K})\,dx = \int_K \nabla w\cdot(\mathbf{A}^*e_j)\,dx \quad \text{for all } w\in W_{D_T}. \tag{4.4}$$

Choosing $w = \frac{\partial\mathbf{q}_{T;k}}{\partial\mathbf{A}}(\delta\mathbf{A})$ as a test function, we may rewrite the Fréchet derivative of $\mathcal{A}_H|_{T;K}$ as

$$\frac{\partial(\mathcal{A}_H|_{T;K})_{jk}}{\partial\mathbf{A}}(\delta\mathbf{A}) = \frac{1}{|T|\,|K|}\Big(\delta_{T;K}\int_T (\delta\mathbf{A})_{jk}\,dx - e_j\cdot\int_K \delta\mathbf{A}\nabla\mathbf{q}_{T;k}\,dx - \int_{D_T}\nabla\; {}_{T;K;j}\cdot\mathbf{A}\nabla\frac{\partial\mathbf{q}_{T;k}}{\partial\mathbf{A}}(\delta\mathbf{A})\,dx\Big). \tag{4.5}$$

The differentiation of (3.3) shows that

$$\int_{D_T}\nabla w\cdot\Big(\mathbf{A}\nabla\frac{\partial\mathbf{q}_{T;k}}{\partial\mathbf{A}}(\delta\mathbf{A})\Big)\,dx = \int_T\nabla w\cdot(\delta\mathbf{A}e_k)\,dx - \int_{D_T}\nabla w\cdot(\delta\mathbf{A}\nabla\mathbf{q}_{T;k})\,dx$$

for any $w\in W_{D_T}$. The particular choice $w = {}_{T;K;j}$ allows one to rewrite (4.5) in the form

$$\frac{\partial(\mathcal{A}_H|_{T;K})_{jk}}{\partial\mathbf{A}}(\delta\mathbf{A}) = \frac{1}{|T|\,|K|}\Big(\delta_{T;K}\int_T (\delta\mathbf{A})_{jk}\,dx - e_j\cdot\int_K \delta\mathbf{A}\nabla\mathbf{q}_{T;k}\,dx - \int_T\nabla\; {}_{T;K;j}\cdot(\delta\mathbf{A}e_k)\,dx + \int_{D_T}\nabla\; {}_{T;K;j}\cdot(\delta\mathbf{A}\nabla\mathbf{q}_{T;k})\,dx\Big).$$

This expression characterizes the (L^2 representation of the) Fréchet derivative of the entries of $\mathcal{A}_H|_{T;K}$ with respect to the coefficient field $\mathbf{A}$ as

$$\frac{\partial(\mathcal{A}_H|_{T;K})_{jk}}{\partial\mathbf{A}} = \frac{1}{|T|\,|K|}\Big(\delta_{T;K}\,e_j\otimes e_k\chi_T - e_j\otimes\nabla\mathbf{q}_{T;k}\chi_K - \nabla\; {}_{T;K;j}\otimes e_k\chi_T + \nabla\; {}_{T;K;j}\otimes\nabla\mathbf{q}_{T;k}\chi_{D_T}\Big).$$

Using estimate (4.1) of Assumption (A3), this readily yields

$$\mathbb{E}\Big[\Big(\mathcal{A}_{\mathsf{H}}|_{\mathsf{T;K}} - \bar{\mathsf{A}}_{\mathsf{H}}|_{\mathsf{T;K}}\Big)^2\Big]$$
$$\lesssim \frac{\varepsilon^{\mathsf{d}}}{|T|^2|K|^2}\mathbb{E}\Bigg[\int_{\mathsf{R}^d} \quad {}_{\mathsf{B}"(\mathsf{x})} \chi_{\mathsf{T}}(\delta_{\mathsf{T;K}} + |\nabla \quad {}_{\mathsf{T;K}}|)\, d\tilde{x} \quad ^2 \, dx\Bigg]$$
$$+ \frac{\varepsilon^{\mathsf{d}}}{|T|^2|K|^2}\mathbb{E}\Bigg[\int_{\mathsf{R}^d} \quad {}_{\mathsf{B}"(\mathsf{x})} (\chi_{\mathsf{K}} + \chi_{\mathsf{D}_T}|\nabla \quad {}_{\mathsf{T;K}}|)|\nabla \mathbf{q}_{\mathsf{T}}|\, d\tilde{x} \quad ^2 \, dx\Bigg].$$

Jensen's inequality and Hölder's inequality then imply

$$\mathbb{E}\Big[\Big(\mathcal{A}_{\mathsf{H}}|_{\mathsf{T;K}} - \bar{\mathsf{A}}_{\mathsf{H}}|_{\mathsf{T;K}}\Big)^2\Big]$$
$$\lesssim \frac{\varepsilon^{\mathsf{d}}}{|T|^2|K|^2} \quad |T|\delta_{\mathsf{T;K}} + \mathbb{E}\Bigg[\int_{\mathsf{T}} |\nabla \quad {}_{\mathsf{T;K}}|^2\, dx\Bigg] + \mathbb{E}\Bigg[\int_{\mathsf{K}} |\nabla \mathbf{q}_{\mathsf{T}}|^2\, dx\Bigg]\Bigg)$$
$$+ \frac{\varepsilon^{\mathsf{d}}}{|T|^2|K|^2}\int_{\mathsf{D}_T} \mathbb{E}\Bigg[\quad {}_{\mathsf{B}"(\mathsf{x})} \chi_{\mathsf{D}_T}|\nabla \quad {}_{\mathsf{T;K}}|^2\, d\tilde{x}\Bigg)^2\Bigg]^{1=2}$$
$$\times\, \mathbb{E}\Bigg[\Bigg(\quad {}_{\mathsf{B}"(\mathsf{x})} \chi_{\mathsf{D}_T}|\nabla \mathbf{q}_{\mathsf{T}}|^2\, d\tilde{x}\Bigg)^2\Bigg]^{1=2} dx.$$

At this point we make use of the assumption that D_{T} is a box. We first note that Lemma 4.9 is still valid for the box D_{T} (which need not necessarily match with the triangulation T_{H}) because, after considering a larger patch $\widehat{D}_{\mathsf{T}}$ containing the box domain D_{T} and applying Lemma 4.9 there, the restriction of the resulting right-hand side $\hat{b}_{\mathsf{T;j}}$ satisfies the properties from Lemma 4.9 for the box domain. Using Lemma 4.9 (which applies to ${}_{\mathsf{T;K}}$, as its defining (4.4) is of the same structure as (3.3)) and Lemma 4.8 (which applies to the box D_{T}) below, we deduce the estimates

$$\int_{\mathsf{D}_T} \mathbb{E}\Bigg[\Bigg(\quad {}_{\mathsf{B}"(\mathsf{x})} |\nabla \mathbf{q}_{\mathsf{T}}|^2\, d\tilde{x} \quad ^2\Bigg] dx \le C(\lambda, \Lambda, \rho)\ell^8|T|,$$
$$\int_{\mathsf{D}_T} \mathbb{E}\Bigg[\Bigg(\quad {}_{\mathsf{B}"(\mathsf{x})} |\nabla \quad {}_{\mathsf{T;K}}|^2\, d\tilde{x}\Bigg)^2\Bigg] dx \le C(\lambda, \Lambda, \rho)\ell^8|K|.$$

Furthermore, a simple energy estimate yields

$$\mathbb{E}\Bigg[\int_{\mathsf{D}_T} |\nabla \mathbf{q}_{\mathsf{T}}|^2\, dx\Bigg] \le C(\lambda, \Lambda)|T|,$$
$$\mathbb{E}\Bigg[\int_{\mathsf{D}_T} |\nabla \quad {}_{\mathsf{T;K}}|^2\, dx\Bigg] \le C(\lambda, \Lambda)|K|.$$

Inserting these bounds as well as the relations $H^{\mathsf{d}} \lesssim |T| \lesssim |K| \lesssim H^{\mathsf{d}}$ into the previous estimate, we obtain the desired bound

$$\mathbb{E}\Big[\Big(\mathcal{A}_{\mathsf{H}}|_{\mathsf{T;K}} - \bar{\mathsf{A}}_{\mathsf{H}}|_{\mathsf{T;K}}\Big)^2\Big] \lesssim \frac{\varepsilon^{\mathsf{d}}}{H^{\mathsf{d}}|T|^2}\ell^8. \qquad \square$$

The following result is derived in the case of an equation on the full space $\mathbb{R}^d$ in [16], extending earlier results from [5, 15]. Its proof in the case of the Dirichlet problem on a box is analogous but requires a boundary regularity theory as derived in [20] as well as a regularity theory at edges and corners as an input; we refer the reader to the forthcoming work [11].

Lemma 4.8 (annealed large-scale L^p theory for random elliptic operators on cubes). *Let $d \in \{2,3\}$, and let $\mathbf{A}$ be a random coefficient field subject to the assumptions* (A1)–(A3). *Let $Q \subset \mathbb{R}^d$ be a box, let $b \in L^2(Q)$, and let $u \in L^2(\Omega; H_0^1(Q))$ be a solution to the linear elliptic PDE*

$$
\begin{aligned}
-\nabla \cdot (\mathbf{A}\nabla u) &= \nabla \cdot b && \text{on } Q,\\
u &\equiv 0 && \text{on } \partial Q.
\end{aligned}
$$

Then for any $2 \leq p < \infty$ and any $p < q < \infty$, a regularity estimate of the form

$$
\int_Q \mathbb{E}\left[\left(\fint_{B_\varepsilon(x)} \chi_Q |\nabla u|^2 \, d\tilde{x}\right)^{p/2}\right] dx \leq C(\lambda, \Lambda, \rho, p, q)\left(\int_Q |b|^q \, dx\right)^{p/q}
$$

holds true.

4.4.2. Schur complement representation of the element correctors.

Lemma 4.9. *The element correctors $\mathbf{q}_{T;j}$ satisfy a PDE of the form*

$$
\nabla \cdot (\mathbf{A}\nabla \mathbf{q}_{T;j}) = \nabla \cdot (\mathbf{A} e_j \chi_T + b_{T;j}) \qquad \text{on } D_T
$$

for some $b_{T;j}$ with

$$
\int_{D_T} |b_{T;j}|^{9/2} \, dx \lesssim \ell^8 |T|.
$$

Proof. The proof proceeds by rewriting the defining equation of the element correctors (3.3) as a Schur complement problem.

Let $f_{T;j} := -\nabla \cdot (\mathbf{A} e_j \chi_T) \in H^{-1}(D_T)$. Let $\mathsf{L} : H_0^1(D_T) \to H^{-1}(D_T)$ be defined as $\mathsf{L}u := -\nabla \cdot (\mathbf{A}\nabla u)$ (where the operator L is not to be confused with the upscaled operator $\mathcal{A}_H$), and let $\mathsf{I}_{H;D_T} : H_0^1(D_T) \to \hat{V}_H(D_T)$ be the concatenation of extension by zero, quasi-interpolation I_H, and restriction to the patch D_T. Here, by $\hat{V}_H(D_T)$ we denote the range $\mathsf{I}_{H;D_T}(H_0^1(D_T))$ of the operator $\mathsf{I}_{H;D;T}$. Note that $\hat{V}_H(D_T)$ is a subspace of the space of P_1 finite element functions on the patch D_T with *arbitrary boundary values* but with zero boundary values on $\partial D \cap D_T$. Denote by $p_{T;j} \in \hat{V}_H'(D_T)$ the Lagrange multiplier associated with the constraint $\mathsf{I}_{H;D_T} \mathbf{q}_{T;j} = 0$ (note that this constraint is equivalent to $I_H \mathbf{q}_{T;j} = 0$).

The element correctors $\mathbf{q}_{T;j} \in H_0^1(D_T)$ are then determined by the Schur complement problem

$$
\begin{pmatrix} \mathsf{L} & \mathsf{I}_{H;D_T}^t \\ \mathsf{I}_{H;D_T} & 0 \end{pmatrix}
\begin{pmatrix} \mathbf{q}_{T;j} \\ p_{T;j} \end{pmatrix} = \begin{pmatrix} f_{T;j} \\ 0 \end{pmatrix}. \tag{4.6}
$$

By the standard theory for Schur complement problems, we have

$$
p_{T;j} = (\mathsf{I}_{H;D_T} \mathsf{L}^{-1} \mathsf{I}_{H;D_T}^t)^{-1} \mathsf{I}_{H;D_T} \mathsf{L}^{-1} f_{T;j}. \tag{4.7}
$$

By the Lax–Milgram theorem and the uniform ellipticity and boundedness of $\mathbf{A}$, the operator $\mathsf{L} : H_0^1(D_\mathsf{T}) \to H^{-1}(D_\mathsf{T})$ is invertible, and the operator norm of its inverse is bounded by a constant. Moreover, we have $\|\mathsf{L}^{-1}v\|_{\mathsf{H}_0^1(\mathsf{D}_T)} \approx \|v\|_{\mathsf{H}^{-1}(\mathsf{D}_T)}$. The latter is seen directly by $\mathsf{L}\mathsf{L}^{-1}v = v$ for all $v \in H^{-1}(D_\mathsf{T})$ and $\|\mathsf{L}w\|_{\mathsf{H}^{-1}(\mathsf{D}_T)} \lesssim \|w\|_{\mathsf{H}_0^1(\mathsf{D}_T)}$ for all $w \in H_0^1(D_\mathsf{T})$.

Set $\hat{f} := \mathsf{L}^{-1} f_{\mathsf{T};\mathsf{j}}$. We then have by the Poincaré inequality on the patch D_T, the bound on the operator norm of L^{-1}, and the definition of $f_{\mathsf{T};\mathsf{j}}$

$$\begin{aligned}\|\hat{f}\|_{\mathsf{L}^2(\mathsf{D}_T)} &\lesssim \ell H \|\hat{f}\|_{\mathsf{H}_0^1(\mathsf{D}_T)} \lesssim \ell H \|f_{\mathsf{T};\mathsf{j}}\|_{\mathsf{H}^{-1}(\mathsf{D}_T)} \leq \ell H \|\mathbf{A} e_\mathsf{j} \chi_\mathsf{T}\|_{\mathsf{L}^2(\mathsf{D}_T)} \\ &\lesssim \ell H |T|^{1=2}. \end{aligned} \tag{4.8}$$

Reformulating (4.7), $p_{\mathsf{T};\mathsf{j}}$ is given by the solution to the equation

$$\langle \mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}, \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} w \rangle = \langle \hat{f}, \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} w \rangle \qquad \text{for all } w \in \hat{V}'_\mathsf{H}(D_\mathsf{T}). \tag{4.9}$$

The quadratic form associated with the operator $\mathsf{I}_{\mathsf{H};\mathsf{D}_T} \mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} : \hat{V}'_\mathsf{H}(D_\mathsf{T}) \to \hat{V}_\mathsf{H}(D_\mathsf{T})$ is coercive. This follows from the energy estimate

$$\begin{aligned}\langle \mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v, \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v \rangle &= \langle \mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v, \mathsf{L}\mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v \rangle \\ &= \int_{\mathsf{D}_T} \mathbf{A} \nabla (\mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v) \cdot \nabla (\mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v) \, dx \\ &\gtrsim \int_{\mathsf{D}_T} |\nabla (\mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v)|^2 \, dx \\ &= \|\mathsf{L}^{-1} \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v\|^2_{\mathsf{H}_0^1(\mathsf{D}_T)} \gtrsim \|\mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} v\|^2_{\mathsf{H}^{-1}(\mathsf{D}_T)}, \end{aligned} \tag{4.10}$$

where in the first estimate the lower bound from (A1) has been used. Thus, the Lax–Milgram theorem yields the existence of a unique solution $p_{\mathsf{T};\mathsf{j}}$ to problem (4.9). We furthermore note that Remark 3.1 (with $\tilde{V}_\mathsf{H}$ defined with respect to two barycentric refinements) implies

$$\begin{aligned}\|\mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}\|_{\mathsf{L}^2(\mathsf{D}_T)} &= \sup_{0 = \mathsf{v} \in \mathsf{H}_0^1(\mathsf{D}_T)} \frac{\langle \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}, v \rangle}{\|v\|_{\mathsf{L}^2(\mathsf{D}_T)}} = \sup_{0 = \mathsf{v} \in \mathsf{H}_0^1(\mathsf{D}_T)} \frac{\langle p_{\mathsf{T};\mathsf{j}}, \mathsf{I}_{\mathsf{H};\mathsf{D}_T} v \rangle}{\|v\|_{\mathsf{L}^2(\mathsf{D}_T)}} \\ &\lesssim \sup_{0 = \tilde{\mathsf{v}} \in \tilde{\mathsf{V}}_H(\mathsf{D}_T)} \frac{\langle \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}, \tilde{v} \rangle}{\|\tilde{v}\|_{\mathsf{L}^2(\mathsf{D}_T)}}. \end{aligned}$$

A standard inverse estimate for finite element functions on the submesh therefore yields

$$\begin{aligned}\|\mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}\|_{\mathsf{L}^2(\mathsf{D}_T)} &\lesssim \sup_{0 = \mathsf{v}_H \in \tilde{\mathsf{V}}(\mathsf{D}_T)} \frac{\langle \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}, \tilde{v} \rangle}{H \|\tilde{v}\|_{\mathsf{H}_0^1(\mathsf{D}_T)}} \\ &\lesssim \sup_{0 = \mathsf{v} \in \mathsf{H}_0^1(\mathsf{D}_T)} \frac{\langle \mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}, v \rangle}{H \|v\|_{\mathsf{H}_0^1(\mathsf{D}_T)}} = H^{-1} \|\mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}\|_{\mathsf{H}^{-1}(\mathsf{D}_T)}. \end{aligned}$$

In combination with (4.10), (4.9), and (4.8), this implies

$$\|\mathsf{I}^\mathsf{t}_{\mathsf{H};\mathsf{D}_T} p_{\mathsf{T};\mathsf{j}}\|_{\mathsf{L}^2(\mathsf{D}_T)} \lesssim \ell H^{-1} |T|^{1=2}. \tag{4.11}$$

In total, from (4.6) we see that the element correctors $\mathbf{q}_{T;j}$ solve an equation of the form

$$-\nabla \cdot (\mathbf{A}\nabla \mathbf{q}_{T;j}) = -\nabla \cdot (\mathbf{A}e_j \chi_T) + I^t_{H;D_T} p_{T;j}.$$

We now claim that this may be rewritten as

$$-\nabla \cdot (\mathbf{A}\nabla \mathbf{q}_{T;j}) = -\nabla \cdot (\mathbf{A}e_j \chi_T + b_{T;j})$$

for some $b_{T;j}$ with

$$\int_{D_T} |b_{T;j}|^{9=2}\, dx \lesssim \ell^8 |T|.$$

To see this, one may, for example, choose $b_{T;j} := \nabla v$ for v solving $-\Delta v = I^t_{H;D_T} p_{T;j}$ with homogeneous Dirichlet boundary conditions on a ball $B_{C \cdot H}(y)$, which contains D_T (where we extend $I^t_{H;D_T} p_{T;j}$ to $B_{C \cdot H}(y)$ by zero outside of Q). Using elliptic regularity theory and (4.11), one then has

$$\int_{D_T} |D^2 v|^2\, dx \lesssim \| I^t_{H;D_T} p_{T;j} \|^2_{L^2(D_T)} \lesssim \ell^2 H^{-2} |T|$$

as well as $\int_{D_T} |\nabla v|^2\, dx \lesssim \ell^2 H^2 \| I^t_{H;D_T} p_{T;j} \|^2_{L^2(D_T)} \lesssim \ell^4 |T|$. Finally, the Sobolev embedding and a scaling argument imply $\int_{D_T} |b_{T;j}|^{9=2}\, dx \lesssim \ell^8 |T|$. □

4.4.3. The proof of Theorem 4.5.

Proof of Theorem 4.5. Denote for a given $T \in T_H$ the index set $J := \{K \in T_H : K \cap N^{`}(T) = \emptyset\}$ and, for $K \in J$, $v_K := |T|(\mathcal{A}_H|_{T;K} - \bar{A}_H|_{T;K})$. Then

$$X\ (T) = \max_{K \in J} |v_K|,$$

and elementary arguments show that

$$\mathbb{E}[X\ (T)^2] \le \mathbb{E}\left[\sum_{K \in J} |v_K|^2 \right] \le \mathrm{card}(J) \max_{K \in J} \mathbb{E}[|v_K|^2].$$

Theorem 4.7 shows that

$$\mathbb{E}[|v_K|^2] \lesssim H^2 + \ell^8 \quad \frac{\varepsilon}{H}\Big)^{d}.$$

This and $\mathrm{card}(J) \lesssim \ell^d$ by quasi-uniformity of the mesh imply

$$\max_{T \in T_H} \quad \sqrt{\mathbb{E}[X\ (T)^2]}\Big) \lesssim \sqrt{\mathrm{card}(J)} \left(H + \ell^4 \left(\frac{\varepsilon}{H}\right)^{d=2} \right) \lesssim \ell^{4+d=2} \left(H + \quad \frac{\varepsilon}{H}\right)^{d=2} \quad . \ \square$$

Acknowledgments. This work was initiated while the authors enjoyed the kind hospitality of the Hausdorff Institute for Mathematics in Bonn during the trimester program *Multiscale Problems: Algorithms, Numerical Analysis, and Computation*. D. Peterseim would like to acknowledge the kind hospitality of the Erwin Schrödinger International Institute for Mathematics and Physics (ESI), where parts of this research were developed under the frame of the thematic program *Numerical Analysis of Complex PDE Models in the Sciences*.

REFERENCES

[1] R. Altmann, E. Chung, R. Maier, D. Peterseim, and S.-M. Pun, *Computational multiscale methods for linear heterogeneous poroelasticity*, J. Comput. Math., 38 (2020), pp. 41–57.

[2] A. Anantharaman, R. Costaouec, C. Le Bris, F. Legoll, and F. Thomines, *Introduction to numerical stochastic homogenization and the related computational challenges: Some recent developments*, in Multiscale Modeling and Analysis for Materials Simulation, Lect. Notes Ser. Inst. Math. Sci. Natl. Univ. Singap. 22, World Sci. Publ., Hackensack, NJ, 2012, pp. 197–272, https://doi.org/10.1142/9789814360906_0004.

[3] S. Armstrong and P. Cardaliaguet, *Stochastic homogenization of quasilinear Hamilton-Jacobi equations and geometric motions*, J. Eur. Math. Soc. (JEMS), 20 (2018), pp. 797–864.

[4] S. Armstrong, P. Cardaliaguet, and P. Souganidis, *Error estimates and convergence rates for the stochastic homogenization of Hamilton-Jacobi equations*, J. Amer. Math. Soc., 27 (2014), pp. 479–540.

[5] S. Armstrong and J.-P. Daniel, *Calderon-Zygmund estimates for stochastic homogenization*, J. Funct. Anal., 270 (2016), pp. 312–329.

[6] S. Armstrong, S. Ferguson, and T. Kuusi, *Higher-Order Linearization and Regularity in Nonlinear Homogenization*, preprint, arXiv:1910.03987, 2019.

[7] S. Armstrong, A. Hannukainen, T. Kuusi, and J.-C. Mourrat, *An Iterative Method for Elliptic Problems with Rapidly Oscillating Coefficients*, preprint, arXiv:1803.03551, 2018.

[8] S. Armstrong, T. Kuusi, and J.-C. Mourrat, *The additive structure of elliptic homogenization*, Invent. Math., 208 (2017), pp. 999–1154.

[9] S. Armstrong and J.-C. Mourrat, *Lipschitz regularity for elliptic equations with random coefficients*, Arch. Ration. Mech. Anal., 219 (2016), pp. 255–348.

[10] S. Armstrong and C. Smart, *Quantitative stochastic homogenization of convex integral functionals*, Ann. Sci. Éc. Norm. Supér. (4), 49 (2016), pp. 423–481.

[11] P. Bella, J. Fischer, M. Josien, and C. Raithel, *tba*, in preparation, (2020).

[12] X. Blanc, C. Le Bris, and F. Legoll, *Some variance reduction methods for numerical stochastic homogenization*, Philos. Trans. A, 374 (2016), 20150168.

[13] A. Bourgeat and A. Piatnitski, *Approximations of effective coefficients in stochastic homogenization*, Ann. Inst. Henri Poincaré Probab. Stat., 40 (2004), pp. 153–165, https://doi.org/10.1016/S0246-0203(03)00065-7.

[14] E. Cances, V. Ehrlacher, F. Legoll, and B. Stamm, *An embedded corrector problem to approximate the homogenized coefficients of an elliptic equation*, C. R. Math. Acad. Sci. Paris, 353 (2015), pp. 801–806.

[15] M. Duerinckx, A. Gloria, and F. Otto, *The Structure of Fluctuations in Stochastic Homogenization*, preprint, arXiv:1602.01717, 2016.

[16] M. Duerinckx and F. Otto, *Higher-Order Pathwise Theory of Fluctuations in Stochastic Homogenization*, preprint, arXiv:1903.02329, 2019.

[17] M. Feischl and D. Peterseim, *Sparse Compression of Expected Solution Operators*, preprint, arxiv:1807.01741, 2018.

[18] J. Fischer, *The choice of representative volumes in the approximation of effective properties of random materials*, Arch. Ration. Mech. Anal., 234 (2019), pp. 635–726.

[19] J. Fischer and S. Neukamm, *Optimal Homogenization Rates in Stochastic Homogenization of Nonlinear Uniformly Elliptic Equations and Systems*, preprint, arXiv:1908.02273, 2019.

[20] J. Fischer and C. Raithel, *Liouville principles and a large-scale regularity theory for random elliptic operators on the half-space*, SIAM J. Math. Anal., 49 (2017), pp. 82–114.

[21] D. Gallistl and D. Peterseim, *Stable multiscale Petrov-Galerkin finite element method for high frequency acoustic scattering*, Comput. Methods Appl. Mech. Engrg., 295 (2015), pp. 1–17.

[22] D. Gallistl and D. Peterseim, *Computation of quasilocal effective diffusion tensors and connections to the mathematical theory of homogenization*, Multiscale Model. Simul., 15 (2017), pp. 1530–1552, https://doi.org/10.1137/16M1088533.

[23] D. Gallistl and D. Peterseim, *Numerical stochastic homogenization by quasilocal effective diffusion tensors*, Commun. Math. Sci., 17 (2019), pp. 637–651.

[24] A. Gloria, S. Neukamm, and F. Otto, *A Regularity Theory for Random Elliptic Operators*, preprint, arXiv:1409.2678, 2014.

[25] A. Gloria, S. Neukamm, and F. Otto, *Quantification of ergodicity in stochastic homogenization: Optimal bounds via spectral gap on Glauber dynamics*, Invent. Math., 199 (2015), pp. 455–515, https://doi.org/10.1007/s00222-014-0518-z.

[26] A. Gloria and F. Otto, An optimal variance estimate in stochastic homogenization of discrete elliptic equations, Ann. Probab., 39 (2011), pp. 779–856, https://doi.org/10.1214/10-AOP571.
[27] A. Gloria and F. Otto, An optimal error estimate in stochastic homogenization of discrete elliptic equations, Ann. Appl. Probab., 22 (2012), pp. 1–28, https://doi.org/10.1214/10-AAP745.
[28] A. Gloria and F. Otto, The Corrector in Stochastic Homogenization: Optimal Rates, Stochastic Integrability, and Fluctuations, preprint, arXiv:1510.08290, 2015.
[29] C. Gu, An Efficient Algorithm for Solving Elliptic Problems on Percolation Clusters, preprint, arXiv:1907.13571, 2019.
[30] P. Henning, A. Malqvist, and D. Peterseim, A localized orthogonal decomposition method for semi-linear elliptic problems, ESAIM Math. Model. Numer. Anal., 48 (2014), pp. 1331–1349, https://doi.org/10.1051/m2an/2013141.
[31] P. Henning and D. Peterseim, Oversampling for the multiscale finite element method, Multiscale Model. Simul., 11 (2013), pp. 1149–1175, https://doi.org/10.1137/120900332.
[32] V. Khoromskaia, B. Khoromskij, and F. Otto, Numerical Study in Stochastic Homogenization for Elliptic PDEs: Convergence Rate in the Size of Representative Volume Elements, preprint, arXiv:1903.12227, 2019.
[33] R. Kornhuber, D. Peterseim, and H. Yserentant, An analysis of a class of variational multiscale methods based on subspace decomposition, Math. Comp., 87 (2018), pp. 2765–2774.
[34] C. Le Bris, F. Legoll, and W. Minvielle, Special quasirandom structures: A selection approach for stochastic homogenization, Monte Carlo Methods Appl., 22 (2016), pp. 25–54.
[35] A. Malqvist and D. Peterseim, Localization of elliptic multiscale problems, Math. Comp., 83 (2014), pp. 2583–2603, https://doi.org/10.1090/S0025-5718-2014-02868-8.
[36] A. Malqvist and D. Peterseim, Generalized finite element methods for quadratic eigenvalue problems, ESAIM Math. Model. Numer. Anal., 51 (2017), pp. 147–163, https://doi.org/10.1051/m2an/2016019.
[37] J.-C. Mourrat, Efficient methods for the estimation of homogenized coefficients, Found. Comput. Math., 19 (2019), pp. 435–483.
[38] H. Owhadi, Bayesian numerical homogenization, Multiscale Model. Simul., 13 (2015), pp. 812–828, https://doi.org/10.1137/140974596.
[39] H. Owhadi, Multigrid with rough coefficients and multiresolution operator decomposition from hierarchical information games, SIAM Rev., 59 (2017), pp. 99–149, https://arxiv.org/abs/1503.03467.
[40] D. Peterseim, Variational multiscale stabilization and the exponential decay of fine-scale correctors, in Building Bridges: Connections and Challenges in Modern Approaches to Numerical Partial Differential Equations, G. R. Barrenechea, F. Brezzi, A. Cangiani, and E. H. Georgoulis, eds., Lect. Notes Comput. Sci. Eng. 114, Springer, Berlin, 2016, pp. 341–367.
[41] D. Peterseim, Eliminating the pollution effect in Helmholtz problems by local subscale correction, Math. Comp., 86 (2017), pp. 1005–1036, https://doi.org/10.1090/mcom/3156.
[42] D. Peterseim and M. Schedensack, Relaxing the CFL condition for the wave equation on adaptive meshes, J. Sci. Comput., 72 (2017), pp. 1196–1213, https://doi.org/10.1007/s10915-017-0394-y.
[43] D. Peterseim, D. Varga, and B. Verfürth, From Domain Decomposition to Homogenization Theory, preprint, arXiv:1811.06319, 2018.
[44] B. Verfürth, Heterogeneous multiscale method for the Maxwell equations with high contrast, ESAIM Math. Model. Numer. Anal., 53 (2019), pp. 35–61, https://doi.org/10.1051/m2an/2018064.