Direct and Iterative Methods for Numerical Homogenization

Ralf Kornhuber[1], Joscha Podlesny[1], and Harry Yserentant[2]

Abstract
Elliptic problems with oscillating coefficients can be approximated up to arbitrary accuracy by using sufficiently fine meshes, i.e., by resolving the fine scale. Well-known multiscale finite elements [5, 9] can be regarded as direct numerical homogenization methods in the sense that they provide approximations of the corresponding (unfeasibly) large linear systems by much smaller systems while preserving the fine-grid discretization accuracy (model reduction). As an alternative, we present iterative numerical homogenization methods that provide approximations up to fine-grid discretization accuracy and discuss differences and commonalities.

Acknowledgements This research has been funded by Deutsche Forschungsgemeinschaft (DFG) through grant CRC 1114.

1 Introduction

Numerical approximation usually aims at modifications of standard finite element approximations of partial differential equations with highly oscillatory coefficients that preserve the accuracy known in the smooth case. Using classical homogenization as a guideline, these modifications are obtained from local auxiliary problems [2, 4, 7]. The error analysis for these kinds of methods is typically restricted to coefficients with separated scales and often requires periodicity [1, 2, 6]. These restrictions were overcome in a recent paper by Målqvist and Peterseim [9] that provides quasioptimal energy and L^2 error estimates without any additional assumptions on periodicity and scale separation [5, 9]. While their approach relies on (approximate) orthog-

[1]FU Berlin: ralf.kornhuber@fu-berlin.de, joscha.podlesny@fu-berlin.de
[2]TU Berlin: yserentant@math.tu-berlin.de

onal subspace decomposition, alternative decompositions into a coarse space and local fine-grid spaces associated with low and high frequencies has been recently considered by Kornhuber and Yserentant [8]. Here, we review these two decomposition techniques providing direct [9] and iterative methods [8] for numerical homogenization in order to better understand conceptual similarities and differences. We also illustrate the performance of the iterative variant by first numerical experiments in $d = 3$ space dimensions.

Both approaches rely on subspace decomposition in function space while practical, discrete variants aim at approximating a sufficiently accurate, computationally unfeasible fine-grid solution up to discretization accuracy. This approximation is either obtained directly from a linear system as derived from local fine-grid problems [9] or iteratively by repeated solution of coarse- and local fine-grid problems [8]. Comparing the computational effort, the direct method requires assembly of the multiscale stiffness matrix and usually leads to larger local fine-grid problems than the iterative approach. In addition, the local fine-grid problems involve a saddle point structure [9, Remark 4.5] rather than positive-definite stiffness matrices [8]. However, in contrast to iterative homogenization the direct approach provides a reduced multiscale basis that incorporates all relevant features and has various advantages, e.g., in case of many different right-hand sides.

2 Elliptic problems with oscillating coefficients

Let $\Omega \subset \mathbb{R}^d$, $d = 2$ or $d = 3$, be a bounded convex domain with polygonal or polyhedral boundary $\partial\Omega$. We consider the variational problem

$$u \in V : \qquad a(u, v) = (f, v) \qquad \forall v \in V, \tag{1}$$

where $V = H_0^1(\Omega)$ is a closed subspace of $H^1(\Omega)$, $(\cdot, \cdot)$ is the canonical scalar product in $L^2(\Omega)$, and $f \in L^2(\Omega)$. The bilinear form $a(\cdot, \cdot)$ takes the form $a(v, w) = \int_\Omega \nabla v(x) \cdot A(x) \nabla w(x)\, dx$, $v, w \in V$, where $A(x) \in \mathbb{R}^{d\times d}$ is a symmetric matrix with sufficiently smooth, but intentionally highly oscillating entries and

$$\delta|\eta|^2 \le \eta \cdot A(x)\eta \le M|\eta|^2 \tag{2}$$

holds for all $\eta \in \mathbb{R}^d$ and almost all $x \in \Omega$ with positive constants δ, M independent of x and η. It is well-known that (1) admits a unique solution and, for ease of presentation, we assume $u \in V \cap H^2(\Omega)$. As a model problem, one might think of two separate scales

$$A(x) = \alpha\left(x, \frac{x}{\varepsilon}\right) I, \qquad x \in \Omega, \tag{3}$$

with the identity matrix I and a fine-scale parameter $\varepsilon > 0$. For periodic coefficients α, the oscillatory problem (1) can be treated by classical homog-

enization via the solution of certain continuous cell problems. However, no scale separation, periodicity, or exact solvability of continuous cell problems will be assumed throughout the rest of the presentation.

Let $\mathcal{T}_H$ denote a regular partition of Ω into simplices with maximal diameter $H > 0$. The corresponding space of piecewise affine finite elements

$$\mathcal{S}_H = \{v \in C(\overline{\Omega}) \mid v|_{\partial\Omega} = 0 \text{ and } v|_t \text{ affine } \forall t \in \mathcal{T}_H\}$$

is spanned by the nodal basis $\lambda_p \in \mathcal{S}_H$, $p \in \mathcal{N}_H$, where $\mathcal{N}_H$ stands for the set of interior vertices of $\mathcal{T}_H$. The usual finite element approximation is given by $u_H = P_{\mathcal{S}_H} u$ with $P_{\mathcal{S}_H} : V \to \mathcal{S}_H$ denoting the Ritz projection defined by

$$P_{\mathcal{S}_H} w \in \mathcal{S}_H : \qquad a(P_{\mathcal{S}_H} w, v) = a(w, v) \qquad \forall v \in \mathcal{S}_H.$$

We have the well-known error estimate $\|u - u_H\| \lesssim H\|u\|_{H^2(\Omega)}$, where $\|\cdot\| = a(\cdot,\cdot)^{1/2}$ signifies the energy norm. Here and throughout this paper, we write $a \lesssim b$, if $a \le cb$ holds with a constant c only depending on the contrast M/δ and on the shape regularity of $\mathcal{T}_H$. Unfortunately, $\|u\|_{H^2(\Omega)}$ depends on the oscillatory behavior of A. For example, we have $\|u\|_{H^2(\Omega)} = \mathcal{O}(\varepsilon^{-1})$ and thus $\|u - u_H\| \lesssim \varepsilon^{-1} H$ in the model case (3). Numerical homogenization is aiming at a modified finite element space $\mathcal{S}_H^{ms}$ with $\dim \mathcal{S}_H^{ms} = \dim \mathcal{S}_H$ such that $u_H^{ms} = P_{\mathcal{S}_H^{ms}}$ satisfies $\|u - u_H^{ms}\| \lesssim H$.

3 Direct homogenization by localized orthogonal decomposition

Let $\Pi : V \to S_H$ denote a quasi-interpolation with the property

$$\|v - \Pi v\|_{0,t} \le C_\Pi H \|\nabla v\|_{0,\omega_t} \quad \forall t \in \mathcal{T}_H, \quad \forall v \in V, \tag{4}$$

with local L^2-norms $\|\cdot\|_{0,t}$, $\|\cdot\|_{0,\omega_t}$ on t, ω_t, respectively, and let ω_t be the union of $t' \in \mathcal{T}_H$ with $t \cap t' \neq \emptyset$. A possible choice is the Clément-type operator [3]

$$\Pi v = \sum_{p \in \mathcal{N}_H} v_p \lambda_p, \qquad v_p = \frac{1}{\omega_p} \int_{\omega_p} v \, dx, \quad \omega_p = \text{int supp } \lambda_p. \tag{5}$$

The main idea taken from Målqvist and Peterseim [9] is to consider the a-orthogonal decomposition

$$V = \mathcal{S}_H^{ms} + V^f \tag{6}$$

into the kernel V^f of Π and its a-orthogonal complement $\mathcal{S}_H^{ms} = (I - P_{V^f})V$.

Proposition 1. *The Ritz projection $u_H^{ms} \in \mathcal{S}_H^{ms}$ of u on $\mathcal{S}_H^{ms}$ satisfies*

$$\|u - u_H^{ms}\| \lesssim H. \tag{7}$$

Proof. Orthogonality of the splitting (6) implies that $w = u - u_H^{ms} \in V^f$ fulfills $\|w\|^2 = (f, w)$. Utilizing the local L^2 scalar product $(\cdot,\cdot)_t$, (4), (2), the local energy norm $\|\cdot\|_t$, the binomial formula, and the L^2 norm $\|\cdot\|_0$, we get

$$\begin{aligned}(f,w) = \sum_{t\in\mathcal{T}_H} (f,w)_t = \sum_{t\in\mathcal{T}_H} (f, w - \Pi w)_t &\lesssim \sum_{t\in\mathcal{T}_H} \|f\|_{0,t} H \|\nabla w\|_{0,\omega_t} \\ \lesssim \sum_{t\in\mathcal{T}_H} s^{-1} H \|f\|_{0,t} s\|w\|_{\omega_t} &\lesssim \tfrac{1}{2} s^{-2} H^2 \|f\|_0^2 + \tfrac{1}{2} c s^2 \|w\|^2\end{aligned}$$

with positive $s \in \mathbb{R}$. The assertion follows by choosing s sufficiently small. □

Note that different choices of Π give rise to different multiscale methods. We refer to [5, 9] for a detailed discussion.

A basis $\lambda_p^{ms} = (I - P_{V^f})\lambda_p$ of $\mathcal{S}_H^{ms}$ is obtained from the local problems

$$\mu_p^{ms} \in V^f : \qquad a(\mu_p^{ms}, v) = a(\lambda_p, v) \qquad \forall v \in V^f \tag{8}$$

for the multiscale corrections $\mu_p^{ms} = P_{V^f}\lambda_p$. Unfortunately, the resulting multiscale basis functions λ_p^{ms} have global support so that sparsity of the corresponding stiffness matrix is lost. As a way out, Målqvist and Peterseim [9] consider the localized orthogonal projection

$$\mu_p^k \in V^f(\omega_{p,k}) : \qquad a(\mu_p^k, v) = a(\lambda_p, v) \qquad \forall v \in V^f(\omega_{p,k}) \tag{9}$$

with local patches $\omega_{p,k}$ of order $k \in \mathbb{N}$ defined by

$$\omega_{p,1} = \omega_p, \qquad \omega_{p,k} = \text{int}\,\{t \in \mathcal{T}_H \mid t \cap \omega_{p,k-1} \neq \emptyset\}, \quad k > 1, \tag{10}$$

and $V^f(\omega_{p,k}) = \{v \in V^f \mid \text{int supp } v \in \omega_{p,k}\}$. The resulting multiscale finite element space now reads $\mathcal{S}_H^k = \text{span}\,\{\lambda_p^k = \lambda_p - \mu_p^k \mid\ p \in \mathcal{N}_H\}$. Exploiting the decay properties of Green's functions Målqvist and Peterseim [9] (see [5] for a later, more elegant proof) were able to show that the desired error estimate (7) is preserved under localization (9).

Theorem 1. *The Ritz projection u_H^k of the solution u of (1) to $\mathcal{S}_H^k$ admits the error estimate $\|u - u_H^k\| \lesssim H$ for sufficiently large $k \gtrsim H^{-1}$.*

The solution of the localized problems (9) is computationally unfeasible, because $\dim V^f = \infty$. As a way out, the continuous solution space V is replaced by a possibly unfeasibly fine finite element space $\mathcal{S}_h$ providing an approximation $u_h = P_{\mathcal{S}_h} u$ with accuracy $\|u - u_h\| \lesssim H$. In the model case (3), we might choose $\mathcal{S}_h$ associated with a uniform partition $\mathcal{T}_h$ with mesh size $h = H\varepsilon^{-1}$. Repeating the above reasoning with V^f replaced by $V_h^f = \ker \Pi|_{\mathcal{S}_h}$, $V^f(\omega_{p,k})$ replaced by $V_h^f(\omega_{p,k}) = V^f(\omega_{p,k}) \cap V_h^f$, etc., we obtain the multiscale finite element space $\mathcal{S}_{H,h}^k = \text{span}\,\{\lambda_{p,h}^k = \lambda_p - \mu_{p,h}^k \mid\ p \in \mathcal{N}_H\}$ with discrete multiscale corrections $\mu_{p,h}^k$ obtained from

$$\mu_{p,h}^k \in V_h^f(\omega_{p,k}): \qquad a(\mu_{p,h}^k, v) = a(\lambda_p, v) \qquad \forall v \in V_h^f(\omega_{p,k}). \tag{11}$$

For quasi-interpolations Π like the one defined in (5), there is no local basis of the linearly constrained subspaces $V_h^f = \ker \Pi|_{\mathcal{S}_h}$. Hence, the constraint $\Pi v = 0$ is usually enforced by a Lagrange multiplier so that the algebraic solution of (11) amounts to solving a saddle point problem. Utilizing essentially the same arguments as before, the error estimates in Proposition 1 and Theorem 1 directly carry over to the discrete case.

Theorem 2. *The Ritz projection $u_{H,h}^k$ of the solution u of (1) to $\mathcal{S}_{H,h}^k$ admits the error estimate $\|u - u_{H,h}^k\| \lesssim H$ for sufficiently large $k \gtrsim H^{-1}$.*

Note that localized orthogonal decomposition can be regarded as a direct method to approximate u_h up to the discretization error by the solution $u_{H,h}^k$ of a much smaller problem. From such a prespective, multiscale finite element methods appears to be a kind of model reduction.

4 Iterative homogenization by subspace correction

The main idea of iterative homogenization is to derive an iterative scheme that allows for solving the given boundary value problem (1) up to a prescribed accuracy with a number of steps that depends only on the contrast M/δ from (2) and on the shape regularity of $\mathcal{T}_H$. To this end, we consider the splitting

$$V = \mathcal{S}_H + \sum_{p \in \overline{\mathcal{N}}_H} V_p, \qquad V_p = H_0^1(\omega_p), \tag{12}$$

with ω_p defined in (5) and $\overline{\mathcal{N}}_H$ consisting of all vertices of $\mathcal{T}_H$. This splitting induces a parallel subspace correction method providing the preconditioner

$$T = P_{\mathcal{S}_H} + \sum_{p \in \overline{\mathcal{N}}_H} P_{V_p}. \tag{13}$$

Utilizing basic results from subspace correction [10, 11], spectral equivalence

$$K_1^{-1}\, a(v,v) \le a(Tv, v) \le K_2\, a(v,v) \qquad \forall v \in V, \tag{14}$$

follows from the stability of the splitting (12). This means that for any $v \in V$ there is a decomposition $v = v_H + \sum_{p \in \overline{\mathcal{N}}_H} v_p$ into $v_H \in \mathcal{S}_H$ and $v_p \in V_p$, $p \in \overline{\mathcal{N}}_H$, such that

$$\|v_H\|^2 + \sum_{p \in \overline{\mathcal{N}}_H} \|v_p\|^2 \le K_1 \|v\|^2 \tag{15}$$

is satisfied with a constant $K_1 > 0$ and such that

$$\|v\|^2 \le K_2(\|v_H\|^2 + \sum_{p\in\overline{\mathcal{N}}_H} \|v_p\|^2) \tag{16}$$

holds with a constant $K_2 > 0$ for any such decomposition. The following proposition taken from [8] is crucial for the rest of this exposition.

Proposition 2. *The splitting* (12) *is stable with positive constants* K_1, K_2 *depending only on the contrast* M/δ *and on the shape regularity of* $\mathcal{T}_H$.

It is not difficult to realize that (16) with $K_2 = d+2$ follows from the Cauchy-Schwarz inequality. Exploiting the quasi-interpolation Π defined in (5) and that the functions λ_p, $p \in \overline{\mathcal{N}}_H$, form a partition of unity, it turns out that (15) holds for the decomposition $v_H = \Pi v$, $v_p = \lambda_p(v - \Pi v)$, $p \in \overline{\mathcal{N}}_H$. We refer to [8] for details.

Note that, in contrast to direct numerical homogenization as explained above, the quasi-interpolation Π now only enters the proof of the condition number estimate, but not the algorithm itself.

Employing spectral equivalence (14), we can use the spectral mapping theorem to obtain usual error bounds for preconditioned cg iterations in function space.

Theorem 3. *The convergence rate* ρ *of the preconditioned cg iteration with preconditioner* T *satisfies* $\rho \le \frac{\sqrt{\kappa}-1}{\sqrt{\kappa}+1}$,$\kappa \le K_1K_2$, *so that the error estimate* $\|u - u^\nu\| \lesssim Tol$ *holds for* $\nu \gtrsim \log(Tol^{-1})$ *and any given tolerance* $Tol > 0$.

Note that, in contrast to direct numerical homogenization, the achievable accuracy is independent of the choice of $\mathcal{S}_H$.

Of course, the preconditioner (13) is computationally unfeasible, because the evaluation of the local Ritz projections P_{V_p}, $p \in \overline{\mathcal{N}}_H$, amounts to the solution of continuous variational problems. As in the previous section, the continuous solution space V is therefore replaced by a, possibly unfeasibly large, finite element space $\mathcal{S}_h \subset V$ that provides an approximation $u_h = P_{\mathcal{S}_H} u$ with accuracy of order H. We then consider the discrete splitting

$$\mathcal{S}_h = \mathcal{S}_H + \sum_{p\in\overline{\mathcal{N}}_H} V_{p,h}, \qquad V_{p,h} = \mathcal{S}_h \cap H_0^1(\omega_p), \tag{17}$$

and the associated preconditioner

$$T_h = P_{\mathcal{S}_H} + \sum_{p\in\overline{\mathcal{N}}_H} P_{V_{p,h}}. \tag{18}$$

Similar arguments as in the continuous case provide the stability of the discrete splitting (17) with constants K_1, K_2 depending only on the contrast M/δ from (2) and on the shape regularity of $\mathcal{T}_H$. Hence, spectral equivalence

$$K_1^{-1}\, a(v,v) \le a(T_h v, v) \le K_2\, a(v,v) \qquad \forall v \in \mathcal{S}_h \tag{19}$$

follows from well-known results, e.g., in [10, 11]. As a consequence, the preconditioned cg iteration in $\mathcal{S}_h$ with preconditioner T_h exhibits mesh-independent convergence rates.

Theorem 4. *The preconditioned cg iteration with preconditioner T_h provides the error estimate $\|u - u_h^\nu\| \lesssim H$ for $\nu \gtrsim \log(H^{-1})$ iteration steps as applied to a fixed initial iterate $u_h^0 \in \mathcal{S}_h$.*

Note that the achievable accuracy is limited only by the selection of the space $\mathcal{S}_h$ but not by the space $\mathcal{S}_H$ as opposed to the direct approach.

Each evaluation of the preconditioner T_h requires the evaluation of the Ritz projections to $\mathcal{S}_H$ and $V_{p,h}$, $p \in \overline{\mathcal{N}}_H$, respectively. As local bases of these subspaces are readily available, this amounts to the solution of symmetric, positive-definite, linear systems associated with the coarse grid $\mathcal{T}_H$ and with the local fine grids $\omega_p \cap \mathcal{T}_h$, $p \in \overline{\mathcal{N}}_H$, and not to saddle point problems (11) as in direct numerical homogenization.

Similar results can be achieved for successive subspace corrections based on the splitting (17). We refer to [8] for further information.

5 Numerical experiments

We consider the unit cube $\Omega = (0,1)^3$ and its uniform partition into cubes of edge length $H = 1/8$ which are further subdivided into cubes of edge length $h = 1/32$ (one more uniform refinement step would lead to computations with more than $2 \cdot 10^6$ unknowns). The simplical partitions $\mathcal{T}_H$ and $\mathcal{T}_h$ are obtained by subdividing each cube into six tetrahedra by the Coxeter-Freudenthal-Kuhn triangulation. We consider (1) with $f \equiv 1$ in the model case (3) with a scalar coefficient $\alpha(x)$ which is piecewise constant on a $32 \times 32 \times 32$ cube grid, with values that are uniformly distributed random numbers in an interval with lower bound $\delta = 1$ and upper bound M.

The reduction factors for the energy error $\|u_h - u_h^\nu\|$ of the preconditioned cg iteration with preconditioner T_h given in (18) and initial iterate $u_h^0 = u_H$ is listed in Table 1 for the ratios $M/\delta = 1$, 10, 10^2, 10^4, and 10^6. The convergence speed does not decrease significantly from $M/\delta = 10^0$, i.e., the simple Laplace equation, to larger and larger contrast, less and less covered by theory. The stopping criterion $\|u_h - u_h^\nu\| \leq \|u_{h/2} - u_h\| \leq \|u - u_h\|$ was reached with at most $\nu = 2$ iteration steps for all considered values of M/δ. Replacing ω_p in (12) by $\omega_{p,k}$, $k > 1$, thus introducing larger overlap, leads to a further improvement of reduction factors. Though error reduction will probably saturate at slightly larger values for mesh sizes $h < 1/32$, we found a similar convergence behavior for $h = 1/512$ in 2D and these computations confirm the potential of iterative methods for numerical homogenization.

step	$M/\delta = 10^0$	$M/\delta = 10^1$	$M/\delta = 10^2$	$M/\delta = 10^4$	$M/\delta = 10^6$
1	0.42289	0.43180	0.43730	0.43673	0.43747
2	0.40494	0.43488	0.44331	0.44399	0.44364
3	0.29253	0.34578	0.34930	0.34953	0.35052
4	0.32946	0.30560	0.30561	0.30714	0.30635
5	0.38972	0.39920	0.40461	0.39976	0.39907
6	0.38917	0.37999	0.38262	0.37489	0.37601
7	0.30847	0.34791	0.35729	0.35498	0.35238
8	0.33201	0.36407	0.38412	0.38667	0.37269
9	0.40475	0.45993	0.47379	0.47412	0.46402
10	0.34971	0.41312	0.41947	0.42260	0.41620

Table 1 Error reduction factors of preconditioned cg iteration with preconditioner T_h.

References

[1] A. Abdulle. A priori and a posteriori error analysis for numerical homogenization: a unified framework. *Series in Contemporary Applied Mathematics*, 16:280–305, 2011.

[2] A. Abdulle, W. E, B. Engquist, and E. Vanden-Eijnden. The hetereogeneous multiscale method. *Acta Numerica*, 21:1–87, 2012.

[3] Ph. Clément. Approximation by finite element functions using local regularization. *Rev. Franc. Automat. Inf. Rech. Operat.*, 9:77–84, 1975.

[4] Y. Efendiev and T. Y. Hou. *Multiscale Finite Element Methods: Theory and Applications*. Springer, New York, 2009.

[5] P. Henning, A. Målqvist, and D. Peterseim. A localized orthogonal decomposition method for semi-linear elliptic problems. *ESAIM: Math. Model. Numer. Anal.*, 48:1331–1349, 2014.

[6] T. Y. Hou, X.-H. Wu, and Z. Cai. Convergence of a multiscale finite element method for elliptic problems with rapidly oscillating coefficients. *Math. Comp.*, 68:913–943, 1999.

[7] T. J. R. Hughes, G. R. Feijó, L. M. Mazzei, and J.-B. Quincy. The variational multiscale method - a paradigm for computational mechanics. *Comput. Methods Appl. Mech. Engrg.*, 166:3–24, 1998.

[8] R. Kornhuber and H. Yserentant. Numerical homogenization of elliptic multiscale problems by subspace decomposition. *Multiscale Model. Simul.*, 2015. submitted.

[9] A. Målqvist and D. Peterseim. Localization of elliptic multiscale problems. *Math. Comp.*, 83:2583–2603, 2014.

[10] J. Xu. Iterative methods by space decomposition and subspace correction. *SIAM Review*, 34:581–613, 1992.

[11] H. Yserentant. Old and new convergence proofs for multigrid methods. *Acta Numerica*, 2:285–326, 1993.