Wavelet-Based Subgrid Modeling: 1. Principles and Scalar Equations

Ulf Andersson* Björn Engquist† Gunnar Ledfelt*
Olof Runborg*

Abstract

A systematic technique for the derivation of subgrid scale models in the numerical solution of partial differential equations, is described. The technique is based on wavelet projections of the discrete operator followed by a sparse approximation. The resulting numerical method will accurately represent subgrid scale phenomena on a coarse grid. Applications to numerical homogenization and wave propagation in materials with subgrid inhomogeneities are presented.

1 Introduction

In the numerical simulation of partial differential equations, the existence of subgrid scale phenomena poses considerable difficulties. With subgrid scale phenomena, we mean processes which should influence the solution on the computational grid but which have length scales shorter than the grid size. Fine scales in the initial values may for example interact with fine scales in the material properties and produce coarse scale contributions to the solution.

There are many traditional ways to deal with this problem. Local mesh refinement is useful in several applications. A drawback with this method is that it might lead to reflections at mesh size discontinuities, increased

*Numerical Analysis and Computing Science, KTH

†Numerical Analysis and Computing Science, KTH and Department of Mathematics UCLA. Research supported by the grants ARPA/ONR ,N00014-92-J-1890, ONR, N00014-91-J-1034, NSF, DMS-91-03104.

complexity of the algorithm and reduced time steps for explicit methods due to stability constraints.

There also exist a number of methods based on physical considerations for a specific application such as turbulence models in computational fluid dynamics, [12], and analytically derived local subcell models in computational electromagnetics, [11]. Another example is the immersed interface method, [10] used on Cartesian grids. In this method the subgrid variations of the shape of a material interface is taken into consideration by changing the updating stencils near the interfaces. Shock tracking or shock fitting can also be seen as subgrid models, [1].

The most general analytic technique is that of homogenization. Consider a family of differential operators L_ϵ, indexed by the small parameter ϵ, and study

$$L_\epsilon u_\epsilon = f_\epsilon, \tag{1}$$

where f_ϵ is given. Typically L_ϵ has coefficients which are periodic or stochastic on the spatial ϵ-scale. The classical homogenization problem is to show convergence, $u_\epsilon \to \bar{u}$ as $\epsilon \to 0$, and to find an operator $\bar{L}$ and a right hand side $\bar{f}$ such that

$$\bar{L}\bar{u} = \bar{f}. \tag{2}$$

See Benoussan et al., [2], for a thorough description of homogenization. In practice one often wants to solve (1) for a small but fixed ϵ. Since u_ϵ is close to $\bar{u}$ we may take $\bar{u}$ as an approximation. The homogenized equation does not contain the ϵ-scale and is usually much easier to solve. In some cases $\bar{L}$ has a closed form analytic expression, [2]. If that is not the case, but the general structure of $\bar{L}$ is known, a numerical approximation is possible, [8].

In this paper we present a general procedure for including subgrid scale effects in the coarse grid discretization. It is based on multiresolution analysis with wavelet projections and approximation of the discrete operator. This work is a continuation of the work by Dorobantu and Engquist, [7], and based on ideas from Brewster and Beylkin, [4]. In this paper we extend the method in [7] to hyperbolic problems and introduce a new truncation algorithm. The coarse scale contribution from the time variations in the fine scale requires a new level of approximation. The new truncation technique increases the sparsity in the final coarse grid operator. This truncation can also be used for compression of the wavelet based representation of Calderon-Zygmund

operators in [4]. The great advantage of this procedure is its generality. It can be used on any system of differential equations and does not require separation into the distinct $\mathcal{O}(\epsilon)$ and $\mathcal{O}(1)$ scales. It can also be used to test if it is physically reasonable to represent fine scales effects on a coarse scale grid with a local operator.

This paper is organized as follows. In Section 2 we give the mathematical background and in Sections 3 and 4 we study elliptic and hyperbolic model problems respectively and present a number of numerical examples. The examples cover random and localized inhomogeneities. In one example with the 2D Helmholtz equation, we show that the effect of a small geometric detail can be captured on a coarse grid.

2 Wavelet Projection of Difference Operators

2.1 Multiresolution analysis

We will develop our subgrid scale models within the framework of wavelet based *multiresolution analysis* (MRA). A wavelet representation lends itself naturally to analyzing the fine and coarse scales as well as the localization properties of a function. For a detailed description of MRA we defer the reader e.g. to the book by Daubechies, [6].

An MRA is made up of a ladder of closed function spaces,

$$\ldots \subset V_j \subset V_{j+1} \subset \ldots , \tag{3}$$

satisfying

$$\overline{\bigcup_{j\in\mathbb{Z}} V_j} = L_2(\mathbb{R}), \qquad \bigcap_{j\in\mathbb{Z}} V_j = \{0\}. \tag{4}$$

Moreover, the spaces V_j represent different scales in the sense that

$$f(x) \in V_j \;\Leftrightarrow\; f(2^{-j}x) \in V_0. \tag{5}$$

The zero space should also be invariant under integer translations,

$$f(x) \in V_0 \;\Rightarrow\; f(x-k) \in V_0, \qquad \forall k \in \mathbb{Z}, \tag{6}$$

and there should exist a *shape function*, $\varphi \in V_0$, such that

$$\{\varphi(x-k) :\; k \in \mathbb{Z}\}, \tag{7}$$

is an L_2-orthonormal basis for V_0.

From the definition above follows a number of important facts. Dilates and translates of the shape function span the V_j spaces,

$$V_j = \text{span}\{\varphi_{j,k}(x) = 2^{j/2}\varphi(2^j x - k)\}, \qquad j,k \in \mathbb{Z}. \tag{8}$$

Denote the orthogonal complement of V_j in V_{j+1} by W_j,

$$V_{j+1} = V_j \oplus W_j, \qquad V_j \perp W_j. \tag{9}$$

Then, $L_2(\mathbb{R}) = \overline{\bigoplus_{j\in\mathbb{Z}} W_j}$ and there exists a *mother wavelet*, $\psi \in W_0$, with properties corresponding to that of φ,

$$W_j = \text{span}\{\psi_{j,k}(x) = 2^{j/2}\psi(2^j x - k)\}, \qquad j,k \in \mathbb{Z}. \tag{10}$$

This mother wavelet can be constructed explicitly from a given MRA. The functions $\varphi_{j,k}$ form an L_2-orthonormal basis in V_j and so does the functions $\psi_{j,k}$ in W_j.

The simplest MRA is the one built on the Haar basis, where the shape function and the mother wavelet are given by

$$\varphi(x) = \begin{cases} 1, & \text{if } 0 \le x \le 1, \\ 0, & \text{otherwise,} \end{cases} \qquad \psi(x) = \begin{cases} 1, & \text{if } 0 \le x \le 1/2, \\ -1, & \text{if } 1/2 \le x \le 1, \\ 0, & \text{otherwise.} \end{cases} \tag{11}$$

This is the basis that we will use in the remaining part of the paper.

2.2 Wavelet projections

Given the full discrete solution operator on a fine grid, the underlying idea of wavelet homogenization is to find an operator of lower dimension that extracts only the coarse scales of the solution. This is the part of the solution that lies in a space representing a coarse scale. In the Haar basis, for a solution in V_{j+1}, the coarse scale is represented by V_j, and we are thus interested in the operator that yields the solution's projection onto V_j.

Consider a transformation,

$$\mathcal{W}_j : V_{j+1} \to W_j \oplus V_j, \tag{12}$$

that maps the basis $\{\varphi_{j+1,k}\}$ into $\{\psi_{j,k}, \varphi_{j,k}\}$. This is an orthogonal operator and we denote its inverse with $\mathcal{W}^\top$. We further denote by P_j and Q_j the L_2-projections onto V_j and W_j. A linear operator L_{j+1} acting on the space V_{j+1} can be decomposed into four operators $L_{j+1} = A_j + B_j + C_j + L_j$ acting on the subspaces W_j and V_j, where

$$\begin{array}{ll} A_j = Q_j L_{j+1} Q_j : W_j \to W_j, & B_j = Q_j L_{j+1} P_j : V_j \to W_j, \\ C_j = P_j L_{j+1} Q_j : W_j \to V_j, & L_j = P_j L_{j+1} P_j : V_j \to V_j. \end{array} \tag{13}$$

Applying the transformation $\mathcal{W}_j$ on L_{j+1}, we have for $U \in V_{j+1}$,

$$\mathcal{W}_j L_{j+1} \mathcal{W}_j^\top (\mathcal{W}_j U) = \begin{bmatrix} A_j & B_j \\ C_j & L_j \end{bmatrix} \begin{bmatrix} Q_j U \\ P_j U \end{bmatrix}, \tag{14}$$

or simply

$$\mathcal{W}_j L_{j+1} \mathcal{W}_j^\top = \begin{bmatrix} A_j & B_j \\ C_j & L_j \end{bmatrix}. \tag{15}$$

In the Haar basis we get the following matrix representation of $\mathcal{W}$,

$$\mathcal{W} = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 & \cdots & & \\ 0 & 0 & 1 & -1 & 0 & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & 1 & -1 \\ 1 & 1 & 0 & \cdots & & \\ 0 & 0 & 1 & 1 & 0 & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & 1 & 1 \end{pmatrix}. \tag{16}$$

Let us now consider

$$L_{j+1} U = F, \qquad U, F \in V_{j+1}. \tag{17}$$

This equation may originate from a finite difference, finite element or finite volume discretization of a given equation on a regular grid. The continuous function $u(x)$ is identified as piecewise constant over the grid. After the same wavelet transformation as in (14), we have

$$\begin{bmatrix} A_j & B_j \\ C_j & L_j \end{bmatrix} \begin{bmatrix} U_h \\ U_l \end{bmatrix} = \begin{bmatrix} F_h \\ F_l \end{bmatrix}, \qquad U_h, F_h \in W_j, \qquad U_l, F_l \in V_j, \tag{18}$$

where $U_h = Q_j U$ and $U_l = P_j U$ and similarly for F. For the Haar basis this means that U_h is essentially the high frequency part and U_l is the low frequency part of U. Eliminating U_h yields the equation for U_l,

$$(L_j - C_j A_j^{-1} B_j) U_l = F_l - C_j A_j^{-1} F_h. \tag{19}$$

Our new "coarse grid operator" is the Schur complement,

$$\bar{L}_j = L_j - C_j A_j^{-1} B_j, \tag{20}$$

which include subgrid phenomena via $C_j A_j^{-1} B_j$. We also get the homogenized right hand side,

$$\bar{F}_j = F_l - C_j A_j^{-1} F_h. \tag{21}$$

Note that this is in fact a block Gaussian elimination procedure. Further note that the above procedure can be repeated on $\bar{L}_j$ to get $\bar{L}_{j-1}$ and so on. To make this efficient in real applications it is necessary to be able to approximate $\bar{L}_j$ with a sparse matrix. In Section 2.5 we will look closer at such techniques.

2.3 Two-dimensional extension

In more than one dimension the MRA is extended by considering tensor product spaces. Each dimension can be treated separately, using the one-dimensional analysis. The resulting multi-dimensional MRA is a tensor product of one-dimensional MRAs.

In two dimensions we define the tensor product spaces

$$\boldsymbol{V}_j = V_j \otimes V_j, \qquad \ldots \subset \boldsymbol{V}_j \subset \boldsymbol{V}_{j+1} \subset \ldots, \qquad j \in \mathbb{Z}, \tag{22}$$

with V_j being the one-dimensional spaces introduced in Section 2.1. The spaces $\boldsymbol{V}_j$ will satisfy the same provisions, (3)-(7), as the V_j spaces, with $\mathbb{R}, \mathbb{Z}$ replaced by $\mathbb{R}^2, \mathbb{Z}^2$. The shape function of this two-dimensional MRA can be written in terms of the one-dimensional shape function,

$$\boldsymbol{\varphi}(\boldsymbol{x}) = \varphi(x_1) \otimes \varphi(x_2) = \varphi(x_1)\varphi(x_2). \tag{23}$$

It follows, like in the one-dimensional case, that

$$\{\boldsymbol{\varphi}_{j,k,l}(\boldsymbol{x}) = \varphi_{j,k}(x_1) \otimes \varphi_{j,l}(x_2)\}, \qquad k, l \in \mathbb{Z}, \tag{24}$$

is an orthonormal basis for $\boldsymbol{V}_j$. Furthermore, denoting the orthogonal complement of $\boldsymbol{V}_j$ in $\boldsymbol{V}_{j+1}$ by $\boldsymbol{W}_j$, we have that

$$\boldsymbol{V}_{j+1} = \boldsymbol{V}_j \oplus \boldsymbol{W}_j, \qquad \boldsymbol{V}_j \perp \boldsymbol{W}_j, \qquad L_2(\mathbb{R}^2) = \overline{\bigoplus_{j \in \mathbb{Z}} \boldsymbol{W}_j}. \tag{25}$$

The wavelet space $\boldsymbol{W}_j$ is composed of three parts,

$$\boldsymbol{W}_j = (W_j \otimes W_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes V_j). \tag{26}$$

Here, the W_j spaces are the same as in Section 2.1.

The two-dimensional wavelet transform

$$\boldsymbol{\mathcal{W}}_j : \boldsymbol{V}_{j+1} \to \boldsymbol{W}_j \oplus \boldsymbol{V}_j \tag{27}$$

can be written as a tensor product of one-dimensional transforms,

$$\boldsymbol{\mathcal{W}}_j = \mathcal{W}_j \otimes \mathcal{W}_j. \tag{28}$$

The transform maps the standard basis of $\boldsymbol{V}_{j+1}$ into the union of the bases of the three parts of $\boldsymbol{W}_j$ and the standard basis of $\boldsymbol{V}_j$.

A linear operator L_{j+1} that acts on the space $\boldsymbol{V}_{j+1}$ can be decomposed in a way similar to the one-dimensional case. To get a convenient matrix representation we use $\widetilde{\boldsymbol{\mathcal{W}}} = P\boldsymbol{\mathcal{W}}$, instead of $\boldsymbol{\mathcal{W}}$. The matrix P is a suitable permutation. The equation

$$L_{j+1}U = F, \qquad U, F \in \boldsymbol{V}_{j+1}. \tag{29}$$

can then be transformed to

$$\widetilde{\boldsymbol{\mathcal{W}}}_j L_{j+1} \widetilde{\boldsymbol{\mathcal{W}}}_j^{\top} (\widetilde{\boldsymbol{\mathcal{W}}}_j U) = \widetilde{\boldsymbol{\mathcal{W}}}_j F, \tag{30}$$

or

$$\begin{bmatrix} A_j & B_j \\ C_j & L_j \end{bmatrix} \begin{bmatrix} U_h \\ U_l \end{bmatrix} = \begin{bmatrix} F_h \\ F_l \end{bmatrix}, \qquad U_h, F_h \in \boldsymbol{W}_j, \qquad U_l, F_l \in \boldsymbol{V}_j, \tag{31}$$

and the coarse grid operator is again the Schur complement,

$$\bar{L}_j = L_j - C_j A_j^{-1} B_j. \tag{32}$$

Note that the high frequency part of U can be decomposed as

$$U_h = \begin{bmatrix} U_{hh} \\ U_{lh} \\ U_{hl} \end{bmatrix}, \qquad U_{hh} \in W_j \otimes W_j, \qquad U_{lh} \in V_j \otimes W_j, \qquad U_{hl} \in W_j \otimes V_j. \tag{33}$$

Similar tensor product extensions can be made also for higher dimensions.

2.4 Comparison with homogenization

The classical way to approach the problem with small scales analytically, is that of homogenization. The problem is often formulated as follows. Consider a set of operators L_ϵ, indexed by the small parameter ϵ, and a right hand side f. Find the *homogenized operator* $\bar{L}$ defined by

$$L_\epsilon u_\epsilon = f, \qquad \lim_{\epsilon \to 0} u_\epsilon = \bar{u}, \qquad \bar{L}\bar{u} = f. \tag{34}$$

In certain cases the convergence above and existence of the homogenized operator can be proved, [2].

For simple model problems, with coefficients periodic on the fine scale, exact closed form solutions can be obtained. A one-dimensional case is

$$L_\epsilon = -\frac{\partial}{\partial x}\left(a(x/\epsilon)\frac{\partial}{\partial x}\right), \tag{35}$$

where the variable coefficient $a(y)$ is positive, one-periodic and bounded away from zero. In this case the homogenized operator is

$$\bar{L} = -\bar{a}\frac{\partial^2}{\partial x^2}, \qquad \bar{a} = \left(\int_0^1 \frac{dy}{a(y)}\right)^{-1}. \tag{36}$$

We see that the variable coefficient operator is replaced by a constant coefficient operator, the constant being the harmonic average of $a(x)$.

Similarly, in the case of the time-dependent one-dimensional hyperbolic model problem,

$$L_\epsilon = \frac{\partial}{\partial t} + a(x/\epsilon)\frac{\partial}{\partial x}, \tag{37}$$

the solution is

$$\bar{L} = \frac{\partial}{\partial t} + \bar{a}\frac{\partial}{\partial x}. \tag{38}$$

Hence, in the limit only the spatial part of the original operator is affected by the homogenization, and that part is homogenized in exactly the same way as in the elliptic case, (35). The model examples (35), (36) and (37), (38) are used in Sections 3 and 4 respectively.

In higher dimensions, the solution to (34) is a little more complicated, although for model problems the structure of the homogenized operator can

still be written down, as long as the coefficients are periodic. In the elliptic case, we have

$$L_\epsilon = -\sum_{ij} \frac{\partial}{\partial x_i}\left(a_{ij}(x/\epsilon)\frac{\partial}{\partial x_j}\right) \tag{39}$$

with a_{ij} one-periodic in y_i, $i = 1, \ldots, d$, $y = x/\epsilon$. It can be shown, [2], that $\bar{L}$ is given by

$$\bar{L} = -\sum_{ij} A_{ij}\frac{\partial^2}{\partial x_i \partial x_j} \tag{40}$$

where

$$A_{ij} = \sum_k \int (a_{ij} - a_{ik}\frac{\partial \chi_j}{\partial y_k})dy. \tag{41}$$

The function χ_k is given by the solution of the so called cell problem,

$$-\sum_{ij} \frac{\partial}{\partial y_i}\left(a_{ij}(y)\frac{\partial}{\partial y_j}\chi_k(y)\right) = -\sum_i \frac{\partial}{\partial y_i}a_{ik}(y), \tag{42}$$

with periodic boundary conditions for χ.

There is a close relationship between the homogenized operator and the the Schur complement $\bar{L}_j$ in (20), repeated here for convenience,

$$\bar{L}_j = L_j - C_j A_j^{-1} B_j. \tag{43}$$

In the multi-dimensional elliptic case, the first terms in (41) and (43) represent the averaged operator, L_j in a discrete sense and

$$-\sum_{ij} \int a_{ij} dy \frac{\partial^2}{\partial x_i \partial x_j} \tag{44}$$

in a integral sense. The operator C_j is related to the term a_{ik}. The function χ is the solution of an elliptic equation, (42), while A_j^{-1} is a discrete solution operator. The right hand side B_j corresponds to $a_{ik}(y)$ in (42).

Knapek, [9], also uses a Schur complement for the approximation of the homogenized operator. The decomposition in [9] is in terms of grid values and not scales. The coarse grid will thus also contain the high frequency components.

2.5 Compact representation of projected operators

When the operator L_{j+1} is derived from a finite difference, finite element or finite volume discretization, it is sparse and of a certain structure. In one dimension it might for instance be tridiagonal. However, the matrix $\bar{L}_j$ is not sparse since A^{-1} is usually dense. Keeping all components of $\bar{L}_j$ would lead to an inefficient calculation. Fortunately, $\bar{L}_j$ will be diagonal dominant in many important cases. We should then be able to find a sparse matrix that is a close approximation of $\bar{L}_j$.

One simple way to approximate $\bar{L}_j$ is to set all components outside a prescribed bandwidth ν equal to zero. Let us define

$$\text{trunc}(M,\nu)_{ij} = \begin{cases} M_{ij}, & \text{if } 2\,|i-j| \leq \nu - 1 \\ 0, & \text{otherwise.} \end{cases} \tag{45}$$

For $\nu = 1$ we will get a diagonal matrix, for $\nu = 3$ we will get a tridiagonal matrix and so on. This procedure was introduced in [3] and used in [7]. We will refer to it as *truncation*. For lower triangular matrices we define truncation to lower triangular form as

$$\text{trunc}_{\text{L}}(M,\nu)_{ij} = \begin{cases} M_{ij}, & \text{if } 0 \leq i-j \leq \nu - 1 \\ 0, & \text{otherwise.} \end{cases} \tag{46}$$

The basic truncation method used in [3],

$$\text{trunc}_{\text{BCR}}(M,\epsilon)_{ij} = \begin{cases} M_{ij}, & \text{if } |M_{ij}| > \epsilon \\ 0, & \text{otherwise,} \end{cases} \tag{47}$$

is not practical here, since the location of the non-zero elements cannot be controlled.

We propose that $\bar{L}_j$ be projected onto banded form in a more elaborate manner. The aim is that the projected matrix should give the same result as the original matrix on a given subspace, e.g. when applied to vectors representing smooth functions. We define

$$\boldsymbol{v}_j = \{1^j, 2^j, \ldots, N^j\}^T. \tag{48}$$

Let $\mathcal{B}$ be the space of $N \times N$ matrices and $\mathcal{T}_\nu$ the subspace of $\mathcal{B}$ with matrices

essentially[1] of bandwidth ν. Moreover, let

$$\mathcal{L}_\nu = \{M \in \mathcal{B} : \ \mathrm{span}\{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{\nu-1}\} \subset \mathsf{N}(M)\}, \tag{49}$$

where $\mathsf{N}(M)$ represents the null space of A. Then

$$\mathcal{B} = \mathcal{T}_\nu \oplus \mathcal{L}_\nu \tag{50}$$

and we define the *band projection* of a matrix $M \in \mathcal{B}$ as the projection of M onto $\mathcal{T}_\nu$ along $\mathcal{L}_\nu$, with the notation

$$\mathrm{band}(M, \nu) = \mathrm{Proj}_{\mathcal{T}_\nu} M. \tag{51}$$

As a consequence,

$$M\boldsymbol{x} = \mathrm{band}(M, \nu)\boldsymbol{x}, \qquad \forall \boldsymbol{x} \in \mathrm{span}\{\boldsymbol{v}_0, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_{\nu-1}\}. \tag{52}$$

For the case $\nu = 1$ we should remark that we get the standard "masslumping" of a matrix, often used in the context of finite element methods. We conclude by demonstrating an example:

$$M = \begin{pmatrix} 7 & 9 & 0 & 4 & 1 \\ 2 & 9 & 7 & 6 & 7 \\ 3 & 7 & 8 & 4 & -3 \\ -1 & 2 & 6 & 2 & 2 \\ 2 & -4 & 6 & 7 & 10 \end{pmatrix}, \qquad \mathrm{band}(M, 3) = \begin{pmatrix} 14 & -11 & 18 & 0 & 0 \\ 29 & -65 & 67 & 0 & 0 \\ 0 & 13 & 8 & -2 & 0 \\ 0 & 0 & 6 & 4 & 1 \\ 0 & 0 & 6 & 3 & 12 \end{pmatrix} \tag{53}$$

This technique is similar to the probing technique used by Chan et al., [5]. In that case the vectors $\boldsymbol{v}_j$ are sums of unit vectors.

In some cases the band projection technique only gives improvements for small values of ν, see Figure 1. The probing technique of Chan et al., [5], could be used for larger ν in our examples. For small values of ν, the band projection technique is more efficient than the probing technique.

In [7] it was shown that the one-dimensional elliptic model equation $-(au')' = f$ discretized with $-\Delta_+ \mathrm{diag}(a)\Delta_- U = h^2 F$ will preserve its divergence form during homogenization. That is, we will get

$$\bar{L}_j = \frac{1}{h^2}\Delta_+ H_j \Delta_-, \tag{54}$$

[1] We must require that each row of the matrices in $\mathcal{T}_\nu$ has the same number of elements. Therefore, the first and last $\nu - 1$ rows will have additional elements located immediately to the right and left of the band, respectively.

where H_j is a strongly diagonal dominant matrix and the forward and backward undivided differences are defined as

$$\Delta_+ u_i = u_{i+1} - u_i, \qquad \Delta_- u_i = u_i - u_{i-1}. \tag{55}$$

Analogously, for the first order differential operator $a(x)\frac{\partial}{\partial x}$ the discretized form $\mathrm{diag}(a)\Delta_-/h$ is preserved during homogenization, so that

$$\bar{L}_j = \frac{1}{h} H_j \Delta_-. \tag{56}$$

An obvious conclusion from this is that it would be more efficient to do the truncation on H_j rather than on $\bar{L}_j$. The H_j matrix can be seen as the effective material coefficient.

3 Elliptic Problems

In this section we consider first the one-dimensional elliptic model equation,

$$\begin{aligned} -\frac{\partial}{\partial x}\left(a(x)\frac{\partial u}{\partial x}\right) &= f(x), \\ u(0) &= 0, \\ u'(1) &= 0, \end{aligned} \tag{57}$$

with $a(x) > 0$ and then the Helmholtz equation,

$$-\sum_{i=1}^{d} \frac{\partial}{\partial x_i}(a(x)\frac{\partial u}{\partial x_i}) - k^2 u(x) = 0, \tag{58}$$

in one and two dimensions ($d = 1, 2$). The Helmholtz equation can be derived from the wave equation $v_{tt} = \nabla \cdot a(x)\nabla v$ under the assumption that $v(x,t) = u(x)e^{ikt}$. Hence $u(x)$ represents the amplitude and k is the wave number.

We will use a uniform finite volume discretization in the numerical experiments. The computational domain will be the unit interval (square in 2D). The grid size will be denoted n, and the cell size $h = 1/n$. For a given function, $u(x)$, the corresponding grid function will be written u_i (or u_{ij}). The grid function approximates cell averages,

$$u_i \sim \frac{1}{h}\int_{(i-1)h}^{ih} u(x)dx, \qquad i = 1, \ldots, n, \tag{59}$$

in one dimension and

$$u_{ij} \sim \frac{1}{h^2} \int_{(i-1)h}^{ih} \int_{(j-1)h}^{jh} u(x_1, x_2) dx_1 dx_2, \qquad i, j = 1, \ldots, n, \tag{60}$$

in two dimensions. For the whole vector $(u_1, \ldots, u_n)^T$ we will write U. In two dimensions, $U = (u_{11}, \ldots, u_{1n}, u_{21}, \ldots, u_{2n}, \ldots, u_{nn})^T$.

3.1 The 1D model equation

We approximate (57) with the discretization

$$\begin{aligned} -\frac{1}{h^2} \Delta_+ a_i \Delta_- u_i &= f_i, \qquad i = 1, \ldots, n, \\ u_0 &= -u_1, \\ u_{n+1} &= u_n. \end{aligned} \tag{61}$$

First we let $f = -1$ and the coefficients of $a(x)$ be randomly distributed in the interval $[0.1, 2]$. We take $n = 256$ and use three homogenizations, i.e. the coarsest level contains 32 grid points. This is the same case as in Section 3.5.3 of [7]. The random coefficients, however, are not the same.

In Figure 1 we compare different truncation strategies. In the top two subplots we use the truncation, as in [7]. In the bottom two subplots we use the band projection described in Section 2.5. The truncation is performed on H, see (54), and on $\bar{L}$ after all three homogenizations. We see that band projection gives a better approximation. The exact reference solution here is the untruncated solution at the coarsest level.

Next we change the coefficient in the differential equation to

$$a(x) = \begin{cases} 1/6, & \text{if } 0.45 < x < 0.55 \\ 1, & \text{otherwise.} \end{cases} \tag{62}$$

but we keep all other characteristics. The result is given in Figure 2. Again we see that band projection is more efficient than truncation. We also see that it is much more efficient to truncate H than to truncate $\bar{L}$.

The structure of the untruncated $\bar{L}$ and H matrices is shown in Figure 3. It should be noted that the local inhomogeneity of the full operator has spread out over a larger area, but it is still essentially local.

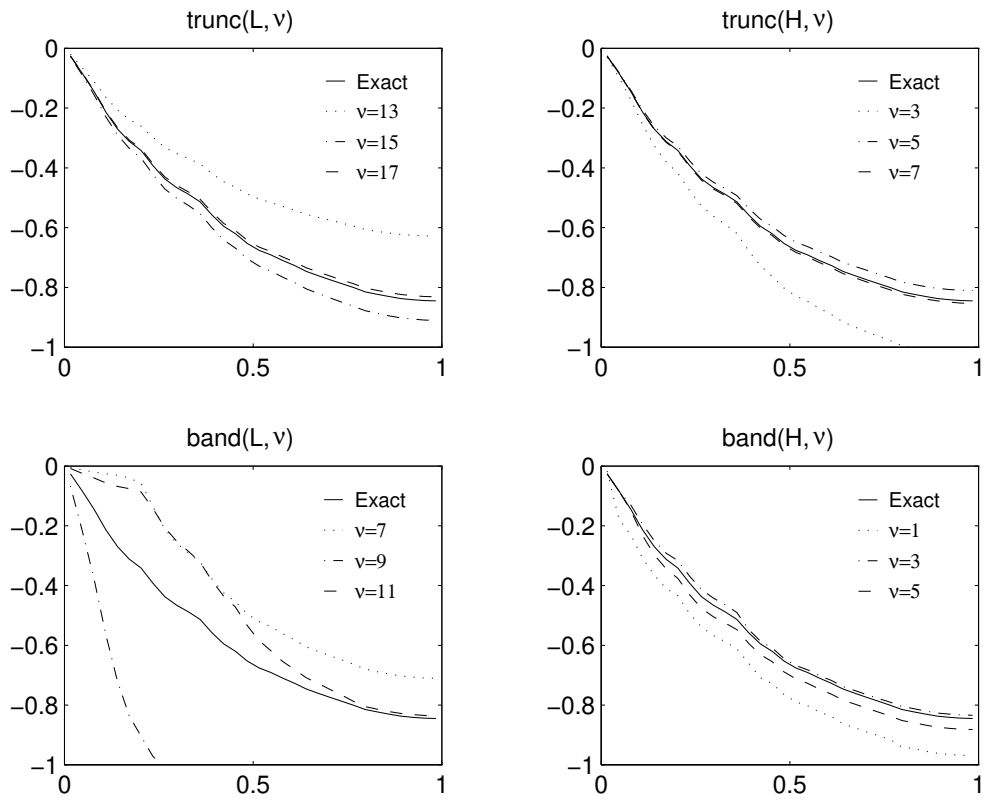


Figure 1: Result for the elliptic model problem, $a(x)$ random, when the homogenized operator is approximated in different ways. The "exact" solution refers to the solution with the full 32×32 homogenized operator.

3.2 The 1D Helmholtz equation

We discretize (58) in one dimension, with the boundary conditions $u(0) = 1$ and $u'(1) = 0$,

$$\begin{aligned} -\frac{1}{h^2}\Delta_+ a_i \Delta_- u_i - k^2 u_i = 0, \quad i = 1, \ldots, n, \\ u_0 = 2u(0) - u_1, \\ u_{n+1} = u_n. \end{aligned} \tag{63}$$

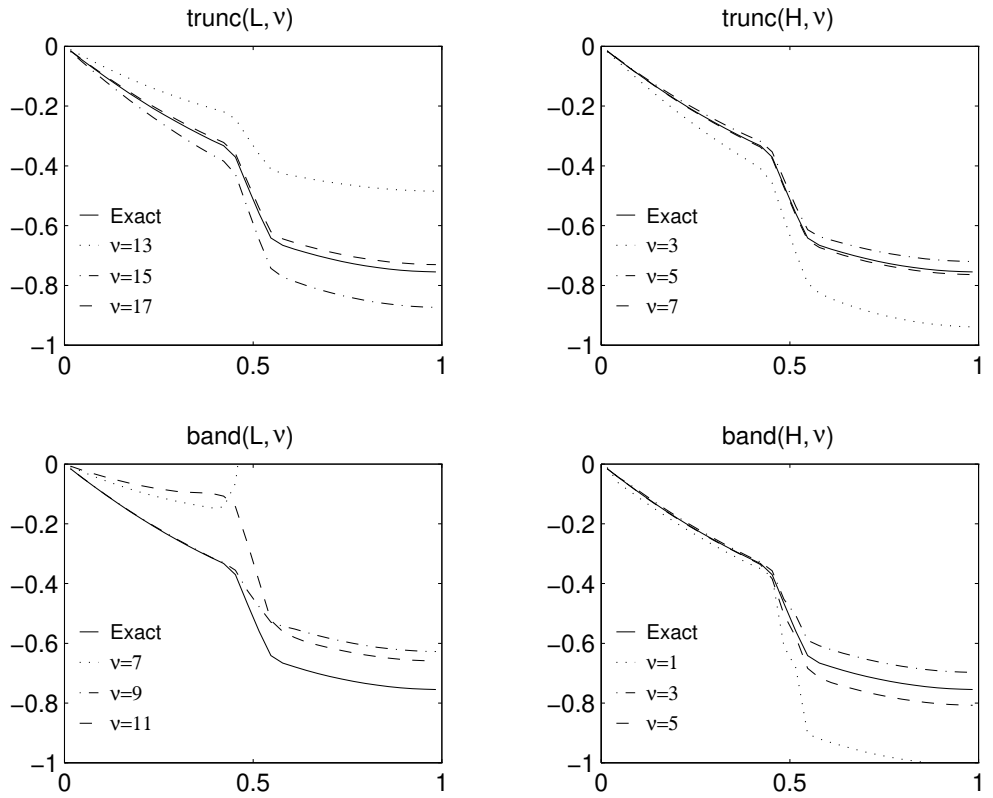


Figure 2: Result for the elliptic model problem, $a(x)$ a slit, when the homogenized operator is approximated in different ways. The "exact" solution refers to the solution with the full 32×32 homogenized operator.

We use $k = 2\pi$ and the same $a(x)$ as in (62) and again we take $n = 256$ and use three homogenizations. We get

$$\bar{L}_j u = (\tilde{L}_j - k^2 I)u = 0. \tag{64}$$

Truncation is performed on $\tilde{L}_j$ (or $\tilde{H}_j$) and not on $\bar{L}_j$. The result is in Figure 4. We see that Helmholtz equation gives results similar to those of the model equation. Again band projection is more efficient than truncation and working on H_j is more efficient than working on $\bar{L}_j$.

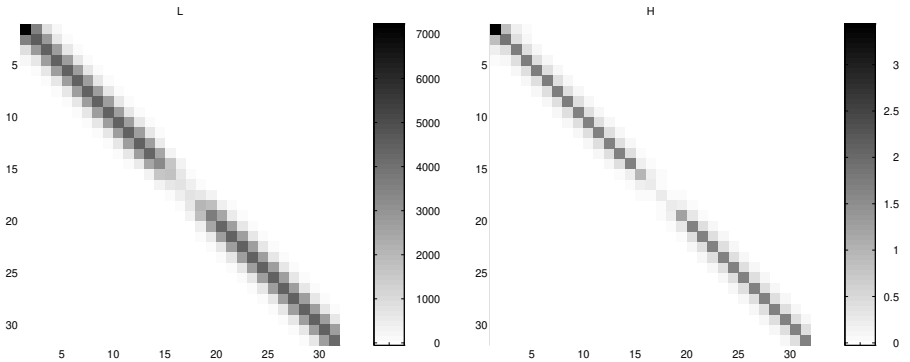


Figure 3: Structure of the untruncated homogenized operators $\bar{L}$ (left) and H (right) for the elliptic model problem, $a(x)$ a slit.

3.3 The 2D Helmholtz equation

We consider the two-dimensional version of (58), with periodic boundary conditions in the y-direction, and $u(0,y) = 1$, $u_x(1,y) = 0$ at the left and right boundaries respectively. This corresponds to a plane time-harmonic wave of amplitude one entering the computational domain from the left and flowing out at the right. The discretization that we use is

$$\begin{aligned} -\frac{1}{h^2}\Delta_+^x a_{i\ell}\Delta_-^x u_{i\ell} - \frac{1}{h^2}\Delta_+^y a_{i\ell}\Delta_-^y u_{i\ell} - k^2 u_{i\ell} = 0, \qquad & i,\ell = 1,\ldots,n, \\ u_{i,0} &= u_{i,n}, \\ u_{i,n+1} &= u_{i,1}, \\ u_{n+1,\ell} &= u_{n,\ell} \\ u_{0,\ell} &= 2 - u_{1,\ell}. \end{aligned} \tag{65}$$

This leads to the matrix equation

$$L_{j+1}U = F, \qquad U, F \in \boldsymbol{V}_{j+1}, \qquad n = 2^{j+1}, \tag{66}$$

where L_{j+1} is homogenized following the theory in Section 2.3.

As an example we choose the $a(x,y)$ shown in Figure 5, which models a wall with a small slit where the incoming wave can pass through. With $k = 3\pi$ and $n = 48$, we obtained the results presented in Figure 6.

The operators in the example above were not truncated in any way. However the structure of them suggests approximations by sparse matrices are feasible, see Figure 7.

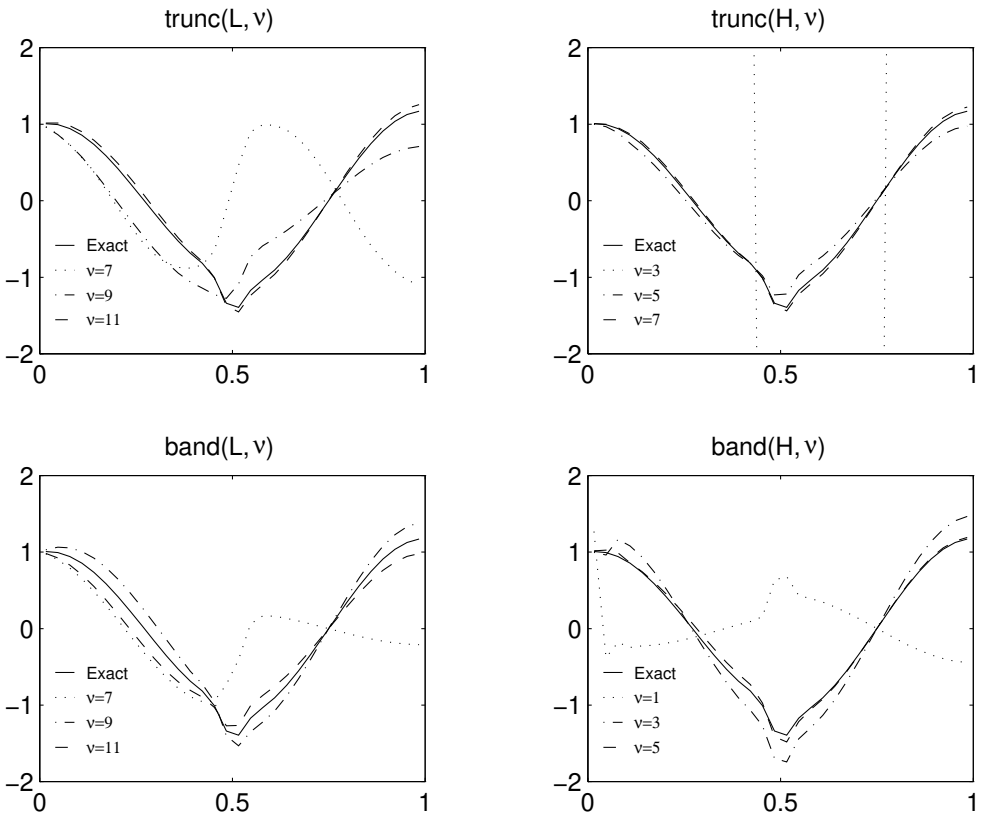


Figure 4: Result for the Helmholtz equation, $a(x)$ a slit, when the homogenized operator is approximated in different ways. The "exact" solution refers to the solution with the full 32×32 homogenized operator.

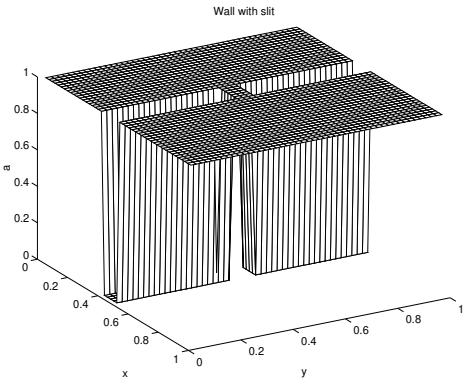
4 Hyperbolic Problems

In this section we consider the time-dependent hyperbolic model equation in one space dimension,

$$\begin{aligned} \frac{\partial u}{\partial t} + a_\epsilon(x)\frac{\partial u}{\partial x} &= f(x), \\ u(0,t) &= 0, \\ u(x,0) &= g(x). \end{aligned} \tag{67}$$

The variable coefficient $a_\epsilon(x)$ is positive and bounded.

We will use three different methods for the numerical homogenization



$$a(x,y) = \begin{cases} 10^{-4}, & \text{if } 0.4 < x < 0.5 \text{ and} \\ & |y - 0.5| > 0.05, \\ 1, & \text{otherwise.} \end{cases}$$

Figure 5: The variable coefficient $a(x,y)$ used in the 2D Helmholtz example.

of (67). First, we consider a semi-discrete form of the problem, and only homogenize the spatial part of the operator. Second, we make a correction for the extra errors introduced when neglecting the time-derivative. Third, we homogenize the full two-dimensional operator.

The most straightforward way to homogenize (67) is by a semi-discrete approximation,

$$\begin{aligned} \frac{\partial u_i}{\partial t} + \frac{1}{h} a_i \Delta_- u_i &= f_i, \qquad i = 1, \ldots, n, \\ u_0 &= -u_1, \\ u_i(0) &= g_i, \end{aligned} \tag{68}$$

with $a_i = a_\epsilon((i-1/2)h)$. In matrix form we will have

$$U_t + L_{j+1} U = F, \qquad 2^{j+1} = n, \tag{69}$$

which can be expressed in a wavelet basis as

$$\frac{\partial}{\partial t}\begin{pmatrix} U_h \\ U_l \end{pmatrix} + \begin{pmatrix} A_j & B_j \\ C_j & L_j \end{pmatrix}\begin{pmatrix} U_h \\ U_l \end{pmatrix} = \begin{pmatrix} F_h \\ F_l \end{pmatrix}, \qquad U_h \in W_j, \qquad U_l \in V_j. \tag{70}$$

We now claim that we can homogenize L_{j+1} in the same way as in the elliptic case, using the Schur complement. The motivation for this is (38), which shows that when the scales of $a_\epsilon(x)$ become very small, $\epsilon \to 0$, the effect of the time-derivative vanishes, $\frac{\partial}{\partial t} U_h \to 0$. As a limiting process in classical

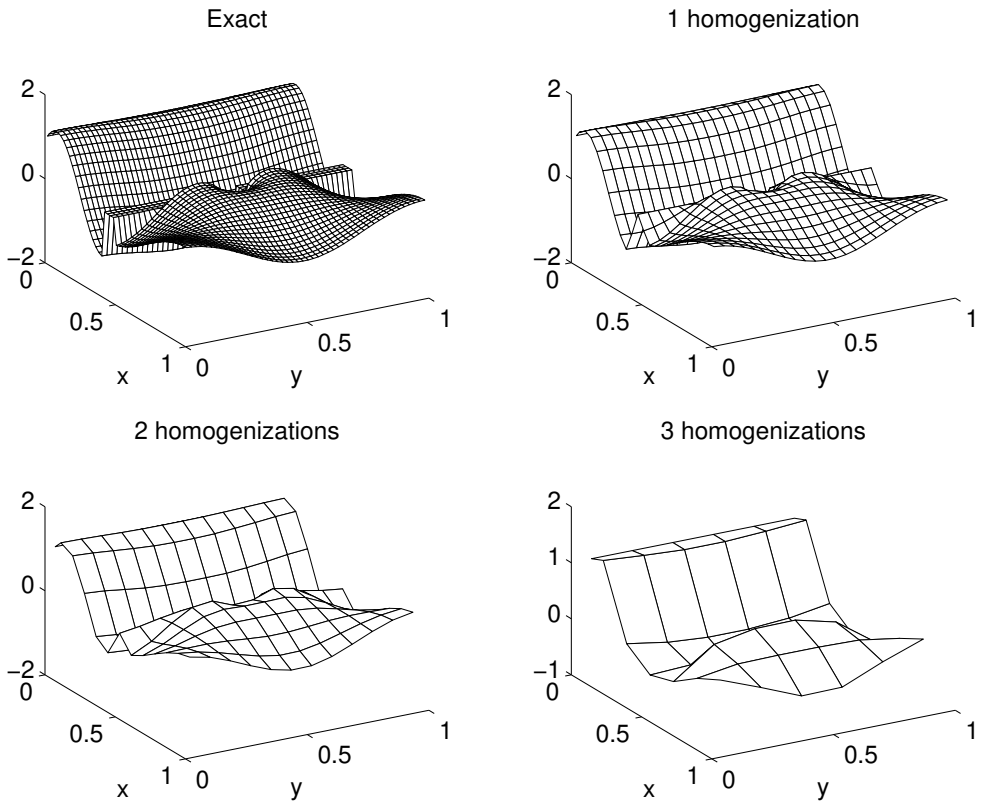


Figure 6: Result for the 2D Helmholtz example. Solution shown for $0, \ldots, 3$ homogenization steps.

homogenization, this is fully justified, [2]. With $\bar{L}_j$ and $\bar{F}_j$ defined by (20) and (21), we get

$$\frac{\partial}{\partial t} U_l + \bar{L}_j U_l = \bar{F}_j. \tag{71}$$

Our second approach is to make a first order correction for the finiteness of ϵ to the computed $\bar{L}$, by assuming that U_h varies approximately linearly with time. This correction is often necessary since in practice ϵ is not small enough. Higher order corrections are possible, but may create stability problems.

Differentiating the first equation in (70) with respect to t, under the assumption that $\frac{\partial^2}{\partial t^2} U_h = 0$, gives us a third equation,

$$A_j \frac{\partial}{\partial t} U_h + B_j \frac{\partial}{\partial t} U_l = 0. \tag{72}$$

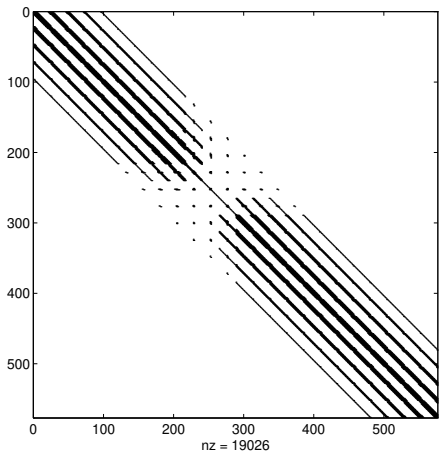


Figure 7: Structure of the homogenized operator $\bar{L}$, after one homogenization step, for the 2D Helmholtz example. Elements larger than 0.1% of max value shown.

We can now solve the system for U_h and $\frac{\partial}{\partial t}U_h$ and write it only in terms of U_l,

$$K_j \frac{\partial U_l}{\partial t} + (L_j - C_j A_j^{-1} B_j) U_l = F_l - C_j A_j^{-1} F_h, \tag{73}$$

with the correction matrix

$$K_j = I + C_j A_j^{-2} B_j. \tag{74}$$

Hence, the first order corrected homogenized operator $\tilde{L}_j$ is

$$\tilde{L}_j = K_j^{-1}(L_j - C_j A_j^{-1} B_j) = K_j^{-1} \bar{L}, \qquad \tilde{F}_j = K_j^{-1}(F_l - C_j A_j^{-1} F_h) = K_j^{-1} \bar{F} \tag{75}$$

and we will solve (71), with $\tilde{L}_j$, $\tilde{F}_j$ substituted for $\bar{L}_j$, $\bar{F}_j$ respectively.

Let $n = 128$, $f \equiv 0$ and

$$g(x) = \begin{cases} \sin^2(4\pi x), & 0 \le x \le 0.25 \\ 0, & 0.25 < x \le 1. \end{cases} \tag{76}$$

The result at time $t = 1$, is given in Figure 8 for the case of $a_\epsilon(x)$ uniformly random distributed in $[0.1, 2]$ and $a_\epsilon(x)$ as in (62). The effect of the first order

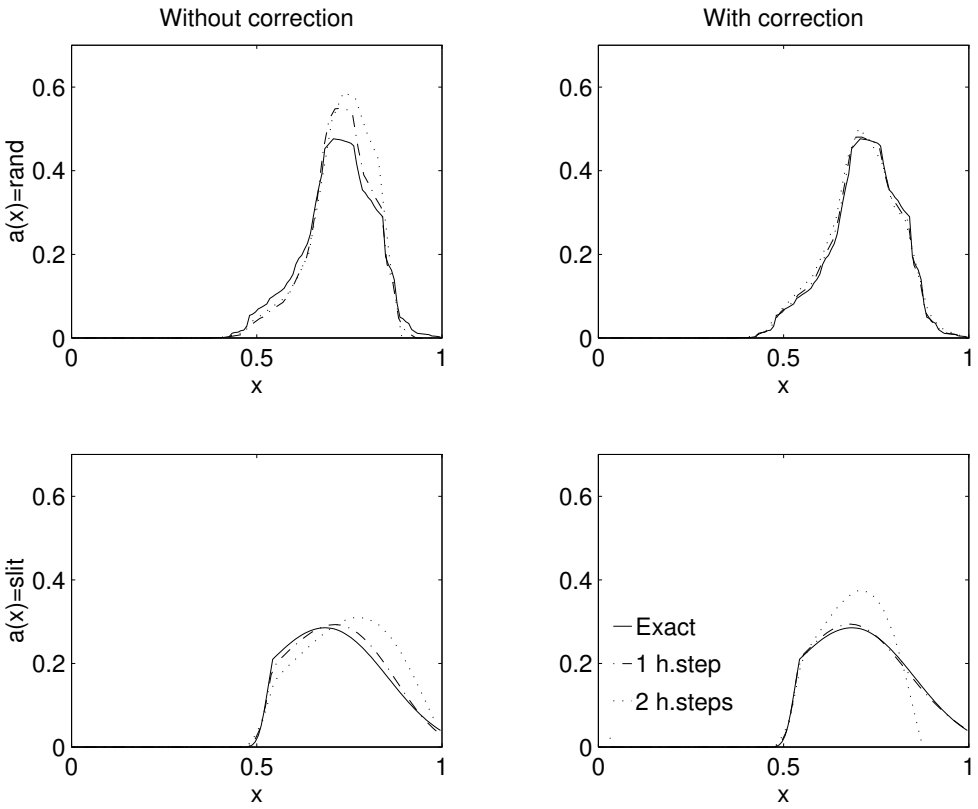


Figure 8: The solution $u(x)$ at $t = 1$ for the hyperbolic case, using different $a_\epsilon(x)$, with and without first order correction. Solutions shown are the exact solution, computed with the full operator, and solutions computed with the operator homogenized one and two steps.

correction is shown to the right. In these calculations, the time-integration was replaced by a fourth-order Runge-Kutta method.

Because of causality the full operator must be lower triangular in the hyperbolic case. The homogenized operator will keep this property. After two homogenization steps the operator, with $a_\epsilon(x)$ a slit, has the structure shown in Figure 9. In view of this we approximate $\bar{L}$ and H, see (56), by truncation to lower triangular form. In Figure 10 results using this truncation is displayed. Like in the elliptic case, truncating the H matrix is more efficient than truncating $\bar{L}$. When the correction method is used it is harder to

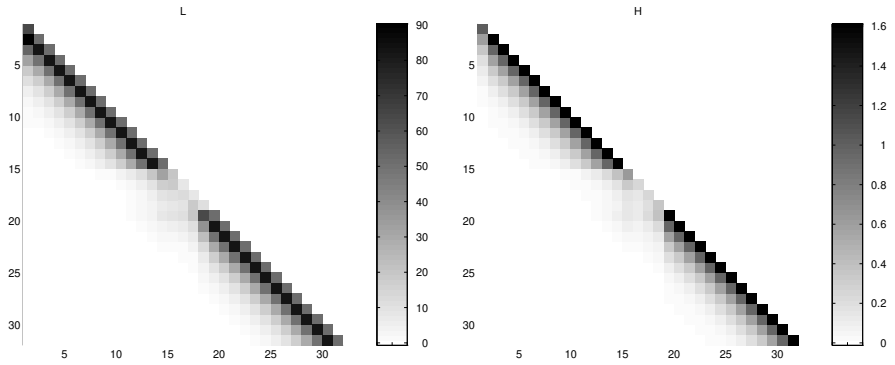


Figure 9: Structure of the untruncated homogenized operators $\bar{L}$ (left) and H (right) for the hyperbolic problem, $a_\epsilon(x)$ a slit.

approximate the resulting homogenized operator, $\tilde{L}$, with a sparse matrix. It will not be as diagonal dominant as in the non-corrected case.

In our third approach we leave the semi-discrete form and instead make a full two-dimensional discretization of (67) in (x,t) space, using the implicit Euler approximation in time,

$$\begin{aligned} \frac{1}{h}\Delta_-^t u_{i\ell} + \frac{1}{h} a_i \Delta_-^x u_{i\ell} &= f_i, \qquad i,\ell = 1,\dots,n, \\ u_{0,\ell} &= -u_{1,\ell}, \\ u_{i,0} &= 2g_i - u_{i,1}. \end{aligned} \tag{77}$$

Like in the two-dimensional Helmholtz case, Section 3.3, we arrive at a linear system of equations,

$$L_{j+1}U = F, \qquad U, F \in \boldsymbol{V}_{j+1}, \qquad n = 2^{j+1}, \tag{78}$$

which is homogenized in the standard way, with the Schur complement.

Using the same data as in the semi-discrete case, but with $n = 48$, we get for the case of a_ϵ as in (62) the results in Figure 11. We used the untruncated homogenized operator, but, like in the 2D Helmholtz case, its structure indicates that good sparse approximations can be found.

5 Conclusions

The new technique of numerical homogenization based on wavelet projections, [7], has been substantially improved and extended.

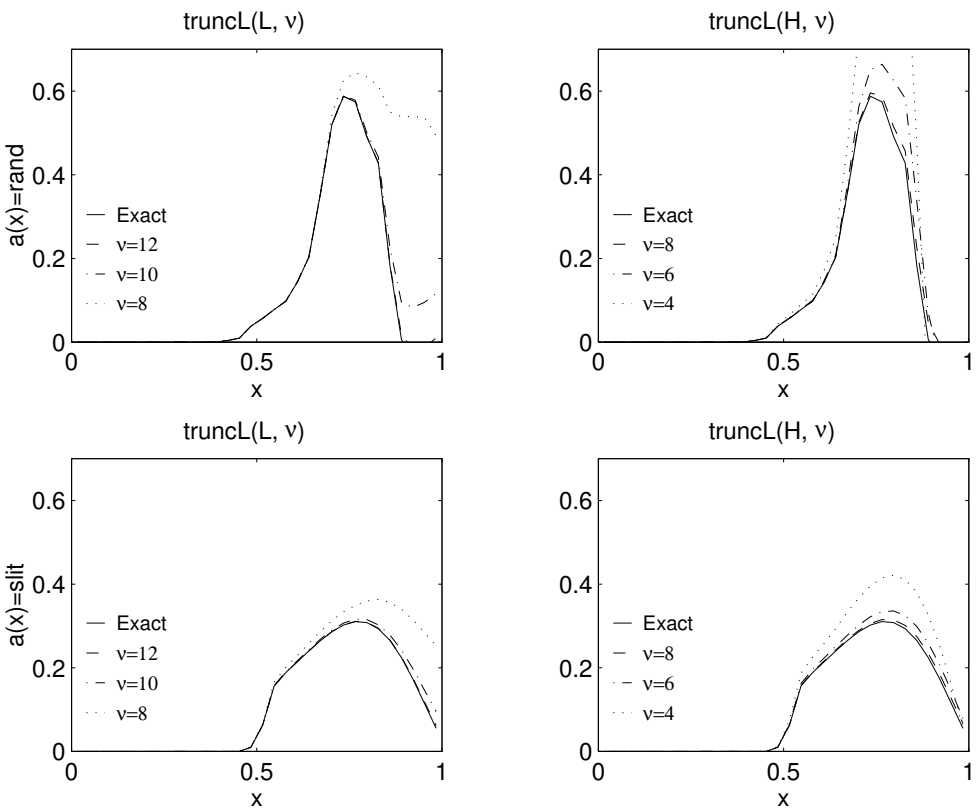


Figure 10: Solution $u(x)$ for the hyperbolic case, using different $a_\epsilon(x)$ and different approximations of the homogenized operator. The "exact" solution refers to the solution with the full 32×32 homogenized operator.

The projection technique has been successfully applied to approximations of differential equations with local variations in the coefficients. This procedure can be seen as a systematic technique for subgrid scale modeling of general boundary value problems. After projection, a coarse grid operator is produced from a local fine grid representation. The coarse grid operator will represent the subgrid scale phenomena but otherwise have a structure similar to a coarse grid approximation of a homogeneous problem. This is useful for example when solving electromagnetic problems where fine geometry details might have large influence on the solution but can not be resolved on a coarse grid. This approach has been tested in an example with the 2D Helmholtz equation.

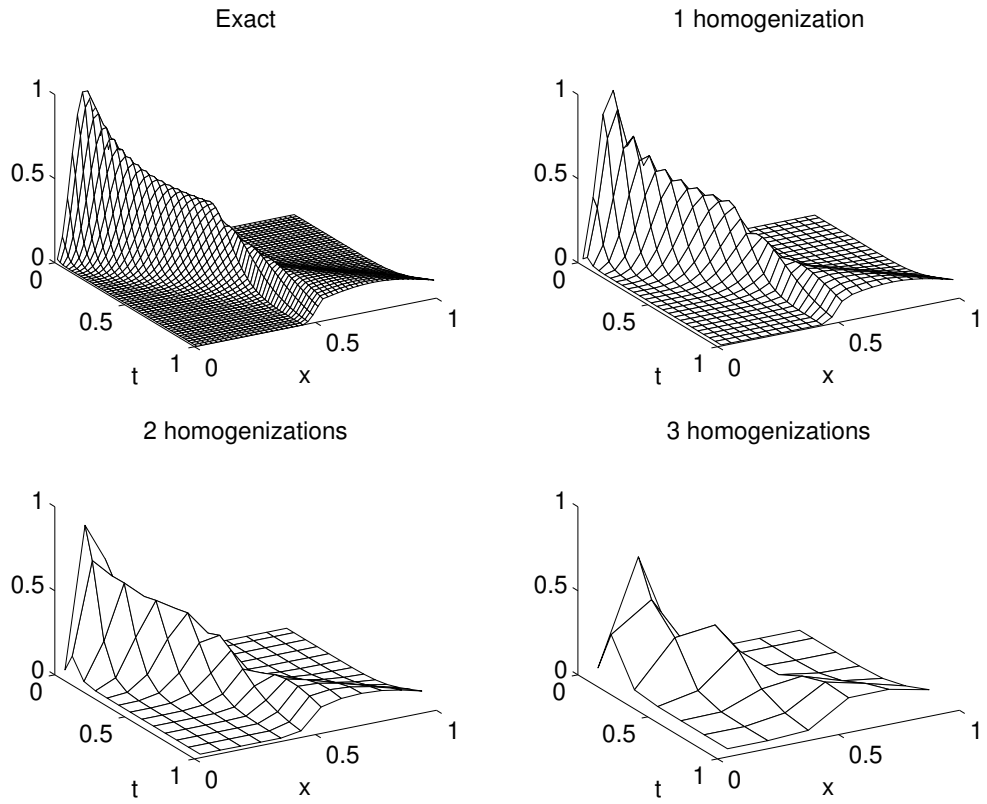


Figure 11: Result for the 2D hyperbolic case, with $a_\epsilon(x)$ a slit.

A new truncation technique has been developed which increases the sparsity over the one described in [3]. The computational stencils of the new coarse grid operators are thus more compact.

The numerical homogenization technique does not apply directly to initial and initial boundary value problems. A new methodology has been developed which handles time dependent problems.

References

[1] J. Anderson. *Computational Fluid Dynamics, The Basics with Applications*. McGraw-Hill, 1995.

[2] A. Benoussan, J.-L. Lions, and G. Papanicolau. *Asymptotic Analysis*

for Periodic Structures. North-Holland Publ. Comp., The Netherlands, 1987.

[3] G. Beylkin, R. Coifman, and V. Rokhlin. Fast wavelet transforms and numerical algorithms I. *Comm. Pure Appl. Math.*, 44:141–183, 1991.

[4] M. Brewster and G. Beylkin. A multiresolution strategy for numerical homogenization. *Applied and Computational Harmonic Analysis*, 2:327–349, 1995.

[5] T. Chan and T. Mathew. The interface probing technique in domain decomposition. *SIAM J. Matrix Anal. Appl.*, 13(1):212–238, January 1992.

[6] I. Daubechies. *Ten Lectures on Wavelets*. SIAM, 1991.

[7] M. Dorobantu and B. Engquist. Wavelet-based numerical homogenization. Technical Report TRITA-NA-9601, C^2M^2, Nada, KTH, 1996. Submitted to SIAM J. on Num. Anal.

[8] L. Durlofsky. Numerical calculation of equivalent grid block permeability tensors for heterogenuous porous media. *Water Resour. Res.*, 27:699–708, 1991.

[9] S. Knapek. Matrix-dependent multigrid-homogenization for diffusion problems. In *Copper Mountain Conference on Iterative Methods*, March 1996.

[10] Z. Li. *The Immersed Interface Method—A Numerical Approach for Partial Differential Equations with Interfaces*. PhD thesis, Department of Applied Mathematics, Univ. of Washington, 1994.

[11] A. Taflove. *Computational Electromagnetics, The Finite-Difference Time-Domain Method*, chapter 10. Artech House, 1995.

[12] D. Wilcox. *Turbulence Modeling for CFD*. DCW Industries, Inc., 1993.