

ON WAVELET-BASED NUMERICAL HOMOGENIZATION*

ALINA CHERTOCK[†] AND DORON LEVY[‡]

Abstract. Recently, a wavelet-based method was introduced for the systematic derivation of subgrid scale models in the numerical solution of partial differential equations. Starting from a discretization of the multiscale differential operator, the discrete operator is represented in a wavelet space and projected onto a coarser subspace. The coarse (homogenized) operator is then replaced by a sparse approximation to increase the efficiency of the resulting algorithm.

In this work we show how to improve the efficiency of this numerical homogenization method by choosing a different compact representation of the homogenized operator. In two dimensions our approach for obtaining a sparse representation is significantly simpler than the alternative sparse representations. L^∞ error estimates are derived for a sample elliptic problem. An additional improvement we propose is a natural fine-scales correction that can be implemented in the final homogenization step. This modification of the scheme improves the resolution of the approximation without any significant increase in the computational cost. We apply our method to a variety of test problems including one- and two-dimensional elliptic models as well as wave propagation problems in materials with subgrid inhomogeneities.

Key words. numerical homogenization, wavelets, Helmholtz equation

AMS subject classifications. Primary, 65N99; Secondary, 74Q99

DOI. 10.1137/030600783

1. Introduction. We are concerned with approximating solutions of partial differential equations (PDEs) that involve information on different scales. Our main interest is in subgrid models in which the computational grid is too coarse to resolve the fine scales of the solution. The precise form of the PDE is of no interest in this work. However, since it is sometimes convenient to have a specific model, the reader is encouraged to think about multiscale differential operators that include, e.g., variable coefficients that are either rapidly oscillating or random.

Our starting point is a particular discretization of the PDE of interest, which we write as

$$(1.1) \quad Lu = f$$

with $L = P(\Delta, \mathcal{A}, h)$. Here, Δ represents discrete difference operators, \mathcal{A} represents a discretization of the coefficients of the equation, and h is a given mesh size, which is assumed to be sufficiently small to resolve all scales. We will consider equation (1.1) as our starting point even when all the scales are not fully resolved. The ultimate goal is to find an alternative “effective” equation,

$$(1.2) \quad \bar{L}\bar{u} = \bar{f},$$

such that \bar{L} is of the same structure of L , i.e., $\bar{L} = P(\Delta, \mathcal{H}, \bar{h})$. Here, \mathcal{H} is a subgrid model of \mathcal{A} , and $\bar{h} \gg h$. Due to the similarity with classical homogenization [3], such

*Received by the editors August 29, 2003; accepted for publication (in revised form) April 12, 2004; published electronically October 14, 2004.

<http://www.siam.org/journals/mms/3-1/60078.html>

[†]Department of Mathematics, North Carolina State University, Campus Box 8205, Raleigh, NC 27695 (chertock@math.ncsu.edu).

[‡]Department of Mathematics, Stanford University, Stanford, CA 94305-2125 (dlevy@math.stanford.edu). This author’s research was supported in part by the National Science Foundation under Career grant DMS-0133511.

a procedure will be referred to as *numerical homogenization*, or, for simplicity, we will usually refer to it as *homogenization* and to the operator \bar{L} as the *homogenized operator*. Since, in general, the homogenized operator is dense, an additional reduction in the complexity of the computations is sought by finding a sparse representation of \bar{L} .

In the past decade, several authors addressed the problem of finding a suitable homogenized representation (1.2). A multiresolution approach presented by Brewster and Beylkin in [9] reduces the problem of finding the effective equation (1.2) into the problem of finding the Schur complement of a certain projection of the equation on a coarse-scales/fine-scales representation of the solution. This approach was further developed by Gilbert [15], where she established the connection between the multiresolution strategy and methods of classical homogenization. Related results were obtained by Dorobantu [11] and by Dorobantu and Engquist [12]. We note that the wavelet basis used in [11, 12, 15] was the Haar basis.

We would like to note in passing that it is possible to use basis functions other than the Haar basis. For example, high-order wavelets were used for approximating solutions of elliptic problems by Beylkin and Coult [7]. There, increasing the number of vanishing moments of the wavelets basis provided control over the rate of decay of the off-diagonal blocks of the reduced operator, hence enabling one to obtain good approximation properties with reduced operators. For other related works see [6, 20] and the references therein.

Generally, the operator L in (1.1) is sparse and enjoys a certain structure. The operator \bar{L} in (1.2), however, is typically dense. Nevertheless, there are important cases in which \bar{L} is diagonally dominant [1]. An exponential decay for the elements away from the main diagonal can be obtained, e.g., for certain elliptic problems [12]. This property motivated Andersson et al. [1] to propose different ways of obtaining a sparse approximation of \bar{L} (see also [14]). It is important to note that in all the numerical simulations performed in [1, 14] when several homogenization steps were performed sequentially, the homogenized operator was replaced with its sparse approximation only in the last homogenization step. Indeed, such an approximation does reduce the overall computational cost. However, from a practical point of view, it is desirable to find a sparse representation of the homogenized operator that can be invoked at every homogenization step. Unfortunately, a direct application of the compactification algorithm of [1, 14] at every stage yields poor approximation results.

In this paper we propose a couple of ways to improve the method of [1, 12, 14]. First, we propose to generate a sparse representation of the homogenized operator by approximating only the high-frequencies component of the homogenized operator. Such an approximation still allows us to control the size of the homogenized operator, which then enables us to use any standard algorithms for handling sparse matrices at every homogenization step. Our numerical results indicate that such an approach provides a significantly more accurate solution compared with [1] when the sparse approximation is performed at every homogenization step. In two dimensions, this algorithm is noticeably simpler than the alternative approaches of [1, 12, 14]. An additional modification we propose is a fine-scales correction, which allows us to improve the resolution at the final computational stage without any significant computational cost.

The structure of the paper is as follows: we start in section 2 with a brief review of multiresolution analysis and of wavelet-based numerical homogenization algorithms. In section 3 we discuss different compact representations of the homogenized operator, focusing on our approach, which involves the compactification of the high-frequencies

part of the homogenized operator. In section 4 we describe the fine-scales correction to improve the resolution at the final computational stage. We conclude in section 5 with numerical simulations of a variety of elliptic and hyperbolic test cases, all of which involve subgrid inhomogeneities.

For completeness of this presentation, we refer the reader to related modern “state-of-the-art” approaches, such as those in [4, 5, 8, 13, 16, 17, 18, 19, 22, 24] and the references therein.

2. Background.

2.1. Multiresolution analysis. We start with a brief review of multiresolution analysis. For more details we refer the reader to [21, 23].

DEFINITION 2.1. *A multiresolution approximation (MRA) of $L^2(\mathbb{R}^n)$ is an increasing sequence V_j , $j \in \mathbb{Z}$, of closed subspaces of $L^2(\mathbb{R}^n)$ with the following properties:*

1. $\bigcap_{j \in \mathbb{Z}} V_j = \{0\}$.
2. $\bigcup_{j \in \mathbb{Z}} V_j$ is dense in $L^2(\mathbb{R}^n)$.
3. For all $f \in L^2(\mathbb{R}^n)$ and $j \in \mathbb{Z}$, $f(x) \in V_j \Leftrightarrow f(2x) \in V_{j+1}$.
4. For all $f \in L^2(\mathbb{R}^n)$ and $k \in \mathbb{Z}^n$, $f(x) \in V_0 \Leftrightarrow f(x - k) \in V_0$.
5. There exists a function $\theta(x) \in V_0$ such that $\{\theta(x - k)\}_{k \in \mathbb{Z}^n}$ is a Riesz basis of the space V_0 .

We recall that a Riesz basis of a Hilbert space \mathcal{H} is a sequence $e_0, \dots, e_k, \dots \in \mathcal{H}$, such that the vector space of finite sums $\sum \alpha_k e_k$ is dense in \mathcal{H} and there exists constants $0 < a_1 < a_2$, such that for all $\alpha_0, \alpha_1, \dots$, we have $a_1 (\sum |\alpha_k|^2)^{1/2} \leq \|\sum \alpha_k e_k\| \leq a_2 (\sum |\alpha_k|^2)^{1/2}$. It is possible to construct from the Riesz basis $\{\theta(x - k)\}_{k \in \mathbb{Z}^n}$ an orthonormal basis for the MRA as stated by the following theorem.

THEOREM 2.2 (see [23], p. 26). *Let V_j , $j \in \mathbb{Z}$, be a MRA of $L^2(\mathbb{R}^n)$. Then there exist constants $0 < c_1 \leq c_2$ such that for almost all $\xi \in \mathbb{R}^n$, $c_1 \leq (\sum_{k \in \mathbb{Z}^n} |\hat{\theta}(\xi + 2k\pi)|^2)^{1/2} \leq c_2$. Also, if $\phi \in L^2(\mathbb{R}^n)$ is defined by $\hat{\phi}(\xi) = \hat{\theta}(\xi) (\sum_{k \in \mathbb{Z}^n} |\hat{\theta}(\xi + 2k\pi)|^2)^{-1/2}$, then $\{\phi(x - k)\}_{k \in \mathbb{Z}^n}$ is an orthonormal basis of V_0 .*

With a proper normalization of the “scale function,” $\phi(x)$, one can generate an orthonormal basis for V_j : $\{\phi_{j,k}(x) = 2^{j/2} \phi(2^j x - k)\}_{k \in \mathbb{Z}^n}$.

For a given MRA, $\{V_j\}_{j \in \mathbb{Z}}$, the “details” space, W_j , is the orthogonal complement of V_j in V_{j+1} , i.e., $V_{j+1} = V_j \oplus W_j$. As a direct consequence of the definition of the MRA, we have $L^2(\mathbb{R}^n) = V_J \oplus \bigoplus_{j=J}^{\infty} W_j = \bigoplus_{j=-\infty}^{\infty} W_j$.

In one dimension we write an orthonormal basis for the “details” space, W_j , in terms of the “wavelets” $\{2^{j/2} \psi(2^j x - k)\}_{k \in \mathbb{Z}}$. The generating function, $\psi(x)$, the so-called mother wavelet of class m , is assumed to satisfy the following properties:

1. $\psi(x)^{(l)} \in L^\infty(\mathbb{R})$ for $l = 0, \dots, m$ (the l th derivative of ψ).
2. $\psi(x)^{(l)}$ decrease rapidly as $x \rightarrow \pm\infty$ for $l = 0, \dots, m$.
3. $\int_{-\infty}^{\infty} x^k \psi(x) dx = 0$ for $k = 0, \dots, m$.
4. $\{2^{j/2} \psi(2^j x - k)\}_{(j,k) \in \mathbb{Z}^2}$ is an orthonormal basis of $L^2(\mathbb{R})$.

In more than one dimension we consider only the separable case in which the multidimensional MRA can be written as the tensor product of one-dimensional MRAs. For example, a separable two-dimensional MRA can be written as $\mathbf{V}_j = V_j \otimes V_j$, where V_j is a one-dimensional MRA. In this case $L^2(\mathbb{R}^2) = \bigoplus_{j \in \mathbb{Z}} \mathbf{W}_j$, and the two-dimensional details space, \mathbf{W}_j , can be decomposed as $\mathbf{W}_j = (W_j \otimes W_j) \oplus (V_j \otimes W_j) \oplus (W_j \otimes V_j)$. In terms of basis functions, if $\{\phi_{j,k}(x)\}_{k \in \mathbb{Z}}$ is an orthonormal basis of V_j , then $\{\phi_{j,k,l}(x_1, x_2) = \phi_{j,k}(x_1) \otimes \phi_{j,l}(x_2)\}_{(k,l) \in \mathbb{Z}^2}$ is an orthonormal basis of \mathbf{V}_j .

2.2. Wavelet-based numerical homogenization. For simplicity we start with a one-dimensional problem and consider

$$(2.1) \quad Lu = f,$$

where L is assumed to be a differential operator and $u, f \in L^2(\mathbb{R})$. A discretization of (2.1) can be viewed in terms of replacing (2.1) with an approximation in V_{j+1} . The operator L is replaced by $L_{j+1} \in \mathcal{L}(V_{j+1}, V_{j+1})$, where $\mathcal{L}(X, Y)$ represents the bounded linear maps between X and Y . Similarly u, f are approximated by $U, F \in V_{j+1}$. Hence,

$$(2.2) \quad L_{j+1}U = F.$$

Our main goal is to reduce the dimensionality of the system (2.2). This will be done by separating U into fine and coarse components and writing a suitable equation only for the coarse part of the solution. Given the decomposition $V_{j+1} = W_j \oplus V_j$, we introduce the orthogonal projection operator $\mathcal{M}_j : V_{j+1} \rightarrow W_j \times V_j$. In other words, we can separate $U \in V_{j+1}$ into two parts: $\mathcal{M}_j U = \begin{pmatrix} U_f \\ U_c \end{pmatrix}$. Here, $U_f \in W_j$ and $U_c \in V_j$ denote the fine- and coarse-scales components of U , respectively. Projecting (2.1) on $V_j \oplus W_j$ we have $\mathcal{M}_j L_{j+1} \mathcal{M}_j^T U = \mathcal{M}_j F$, or, equivalently,

$$(2.3) \quad \mathcal{M}_j L_{j+1} \mathcal{M}_j^T \begin{pmatrix} U_f \\ U_c \end{pmatrix} = \begin{pmatrix} F_f \\ F_c \end{pmatrix},$$

where $F_f \in W_j$ and $F_c \in V_j$ denote the fine- and coarse-scales components of F , respectively. If we now split $\mathcal{M}_j = \begin{pmatrix} Q_j & P_j \end{pmatrix}^T$, where P_j is the projection on V_j and $Q_j = I - P_j$ is the projection on W_j , we can rewrite the operator on the left-hand side of (2.3) as

$$(2.4) \quad \mathcal{M}_j L_{j+1} \mathcal{M}_j^T = \begin{pmatrix} Q_j L_{j+1} Q_j^T & Q_j L_{j+1} P_j^T \\ P_j L_{j+1} Q_j^T & P_j L_{j+1} P_j^T \end{pmatrix} := \begin{pmatrix} A_j & B_j \\ C_j & L_j \end{pmatrix}.$$

An effective equation for the coarse scales of the solution U_c can now be obtained via a block Gaussian elimination of (2.3), which yields

$$(2.5) \quad \bar{L}_j U_c = \bar{F}_j$$

with

$$(2.6) \quad \bar{L}_j = L_j - C_j A_j^{-1} B_j, \quad \bar{F}_j = F_c - C_j A_j^{-1} F_f.$$

Hence, the coarse grid operator \bar{L}_j is the Schur complement of $\mathcal{M}_j L_{j+1} \mathcal{M}_j^T$.

As an example, in the case of the Haar basis the projection on V_j is simply an approximation with a piecewise constant function on the resolution level set by V_j . In this case, Q_j and P_j are given by

$$Q_j = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & -1 & 0 & \cdots & & \\ 0 & 0 & 1 & -1 & 0 & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix}, \quad P_j = \frac{1}{\sqrt{2}} \begin{pmatrix} 1 & 1 & 0 & \cdots & & \\ 0 & 0 & 1 & 1 & 0 & \cdots \\ \vdots & \vdots & & \ddots & \ddots & \\ 0 & 0 & \cdots & 0 & 1 & 1 \end{pmatrix}.$$

We would like to emphasize that the numerical homogenization technique discussed above is not limited to the Haar wavelets and that other types of wavelets can be used.

An analogous derivation holds in the multidimensional case. Similarly to the one-dimensional case, we start with $L_{j+1}U = F$, where $U, F \in \mathbf{V}_{j+1}$, and $L_{j+1} \in \mathcal{L}(\mathbf{V}_{j+1}, \mathbf{V}_{j+1})$. Hence, (2.3) holds for the d -dimensional orthogonal projection operator, $\mathcal{M}_j^d : \mathbf{V}_{j+1} \rightarrow \mathbf{W}_j \times \mathbf{V}_j$, with $U_f, F_f \in \mathbf{W}_j$ and $U_c, F_c \in \mathbf{V}_j$. The multidimensional homogenization procedure has the same form as in one dimension, i.e.,

$$(2.7) \quad \bar{L}_j = L_j - C_j A_j^{-1} B_j, \quad \bar{F}_j = F_c - C_j A_j^{-1} F_f.$$

For example, in the two-dimensional separable case, the coarse scales U_c, F_c belong to $V_j \otimes V_j$, and the fine scales can be decomposed as

$$(2.8) \quad U_f = \begin{pmatrix} U_{ff} \\ U_{fc} \\ U_{cf} \end{pmatrix}, \quad \begin{array}{l} U_{ff} \in W_j \otimes W_j, \\ U_{fc} \in W_j \otimes V_j, \\ U_{cf} \in V_j \otimes W_j. \end{array}$$

In this case, the two-dimensional projection operator \mathcal{M}_j^2 can be written as the tensor product of two one-dimensional projection operators, i.e.,

$$\mathcal{M}_j^2 = \mathcal{M}_j \otimes \mathcal{M}_j = \begin{pmatrix} Q_j \otimes Q_j \\ Q_j \otimes P_j \\ P_j \otimes Q_j \\ P_j \otimes P_j \end{pmatrix}.$$

Remarks.

1. To make sense of (2.5)–(2.6), A_j^{-1} must exist. Since A_j may not be invertible, \bar{L}_j and \bar{F}_j do not exist for all elements of $L_{j+1} \in \mathcal{L}(V_{j+1}, V_{j+1})$. A direct consequence of the Lax–Milgram theorem is that A_j is invertible if L_{j+1} is bounded and strictly positive definite (see Proposition 1 in [14]). In this case the homogenized operator \bar{L}_j remains symmetric and positive definite, and hence it is possible to apply the homogenization procedure several times: $L_{j+1} \rightarrow \bar{L}_j \rightarrow \bar{L}_{j-1} \rightarrow \dots$.
2. The coarse grid operator, \bar{L}_j , preserves some of the properties of L_{j+1} . For example, if L_{j+1} is symmetric, then so is \bar{L}_j . (In the Haar case, if L_{j+1} is triangular, then so is \bar{L}_j). If, in addition to being symmetric, L_{j+1} is also positive definite, then the condition number does not increase in the homogenization procedure, i.e., $\text{cond}(\bar{L}_j) \leq \text{cond}(L_{j+1})$. In certain cases it is possible to show that some additional structure is preserved. For example, in [12] it was shown that the action of the homogenization procedure on the one-dimensional elliptic operator $-(au_x)_x$ preserves the divergence form of the operator. This property was then used in [1] to generate a compact representation of the homogenized operator.

3. A compact representation of the homogenized operator. Our goal in this section is to increase the efficiency of a successive implementation of the homogenization procedure (2.6). We start by assuming that L_{j+1} is a matrix of dimension $n \times n$, and hence the matrix \bar{L}_j obtained by algorithm (2.6) is of dimensions $(n/2) \times (n/2)$. Since $\bar{L}_j = L_j - C_j A_j^{-1} B_j$, computing \bar{L}_j requires one to invert A_j , which is also of dimensions $(n/2) \times (n/2)$. Hence, to compute U_j , one needs to invert

two matrices of dimension $(n/2) \times (n/2)$. This, together with the additional matrix multiplications, accumulates to a computational work that is equivalent to the one required to invert the original matrix L_{j+1} . There is therefore no computational gain in one application of the homogenization algorithm.

The situation slightly improves if the homogenization algorithm is repeated several times. If this is the case, every step requires the inversion of a matrix that is half the size of the previous one. Only the last step requires inverting the full homogenized operator. In either way, when the homogenization algorithm is used only once, or even when it is being used several times, its efficiency depends on being able to generate a sparse representation of the homogenized operator, \bar{L}_j . In general, the matrix A_j^{-1} is dense, and hence the matrix \bar{L}_j is not sparse. However, at least in some cases of interest such as certain elliptic problems, the elements of \bar{L}_j decay exponentially fast away from the main diagonal [12]. This decay property can be used to generate a compact representation of the homogenized operator \bar{L}_j , which can be obtained as follows (see, e.g., [1, 12, 14]):

1. Approximate \bar{L}_j with a band-diagonal matrix, keeping only ν diagonals that are symmetrically distributed around the main diagonal and setting to zero the rest of the elements. For any matrix \mathcal{R} we denote this operation by

$$(3.1) \quad \text{band}(\mathcal{R}, \nu)_{ij} = \begin{cases} \mathcal{R}_{ij} & \text{if } 2|i-j| \leq \nu-1, \\ 0 & \text{otherwise.} \end{cases}$$

If the matrix \mathcal{R} is upper (or lower) triangular, then (3.1) is replaced by an equivalent expression keeping ν diagonals above (or below) the main diagonal. This approach was proposed in a somewhat related but nevertheless different context [6] and was then used in [12].

2. Another compact representation can be obtained via thresholding, i.e., keeping all the elements that are above a certain given threshold ϵ . We use the following notation for this operation:

$$(3.2) \quad \text{trunc}(\mathcal{R}, \epsilon)_{ij} = \begin{cases} \mathcal{R}_{ij} & \text{if } |\mathcal{R}_{ij}| > \epsilon, \\ 0 & \text{otherwise.} \end{cases}$$

In some cases this approach is less practical than the previous one because one has to search for the small elements of the matrix and to adjust the rest of the algorithms accordingly.

3. A technique that was introduced in [1] and is similar to the probing method of [2, 10] is to approximate a matrix by projecting it onto the subspace of matrices with bandwidth ν such that the projected matrix will give the same result as the original matrix on a given subspace. This subspace is typically chosen such that it represents a certain class of smooth functions. Hence, in the so-called band projection method, we are looking for a matrix, $\text{bandproj}(\mathcal{R}, \nu)$, such that for given linearly independent $\{v_j\}_{j=1}^\nu$,

$$(3.3) \quad \mathcal{R}x = \text{bandproj}(\mathcal{R}, \nu)x$$

for all $x \in \text{span}\{v_1, \dots, v_\nu\}$. The results obtained with this method in [1] were of good quality. However, it is limited to smooth solutions. When the solutions are not sufficiently regular, which is the case, e.g., with most of our examples in section 5, the results are not as good. This approach is similar to the frequency-filtering approach; see, e.g., [25].

It is important to note that in all the numerical simulations presented in [1, 12, 14], when several homogenization steps were performed, \bar{L}_j was replaced by its sparse approximation only in the final homogenization step. All other homogenization steps were computed exactly according to (2.6). Our numerical simulations in section 5 indicate that applying the first approximation method at every homogenization step produces poor results. A method that is applicable only for the final homogenization step provides a somewhat moderate gain as far as the efficiency of the algorithm is concerned, and it remains desirable to find an approximation method that can be applied at every stage.

In certain cases it is possible to show that the homogenized operator preserves the structure of the differential operator. For example, in [12] (see also [14]) it was shown that the one-dimensional elliptic equation

$$(3.4) \quad -(au')' = f,$$

which is discretized as $-\Delta_+ \text{diag}(a) \Delta_- U = h^2 F$, preserves its divergence form during the homogenization process. In other words, the homogenized operator can be written as

$$(3.5) \quad \bar{L}_j = \frac{1}{h^2} \Delta_+ H_j \Delta_-,$$

where H_j is a diagonally dominant matrix that can be interpreted as the effective material coefficients related to a . A sparse representation of the homogenized operator \bar{L}_j can be obtained by approximating the corresponding H_j matrix (e.g., by banding according to (3.1)). This approach assumes, of course, that a representation of the form (3.5) is available. The results of [1, 14] indicate that better results can be obtained by approximating H_j instead of \bar{L}_j , and this is reflected, e.g., by having to keep fewer diagonals in $\text{bandproj}(H_j, \nu)$ than in $\text{bandproj}(\bar{L}_j, \nu)$. We will see below that in multidimensional problems it is less obvious how to generate a sparse approximation of \bar{L}_j by approximating the multidimensional generalization of H_j in (3.5).

There are additional approaches for the compact representation of \bar{L}_j . For example, it is possible to connect the decay rate away from the main diagonal in the matrices A_j , B_j , and C_j with the number of vanishing moments of the wavelets that are being used. This strategy is described in [7].

We would like to propose a different approach for obtaining a compact representation of the operator \bar{L}_j . The idea is to approximate only the high-frequencies component of the homogenized operator. More specifically, we band the matrix A_j^{-1} and replace \bar{L}_j by an approximated matrix \bar{L}_j^b that is defined as

$$(3.6) \quad \bar{L}_j^b = L_j - C_j \text{band}(A_j^{-1}, \nu) B_j.$$

We demonstrate below that this seemingly minor modification of the homogenization algorithm has several advantages over previous ideas. For one, it allows us to invoke the approximation at every homogenization step. Also, in two dimensions this approach significantly simplifies the procedure when compared with banding H_j . A related idea was proposed in [20] where the matrix A_j^{-1} was replaced by a diagonal matrix for certain elliptic problems.

One property of \bar{L}_j that is preserved with \bar{L}_j^b is symmetry as stated in the following proposition.

PROPOSITION 3.1. *If L_{j+1} is self-adjoint and A_j is invertible, then \bar{L}_j^b is also self-adjoint.*

Proof. If $L_{j+1} = L_{j+1}^*$ then $\mathcal{M}_j L_{j+1} \mathcal{M}_j^*$ is also self-adjoint. Hence, $A_j = A_j^*$, $L_j = L_j^*$, and $B_j = C_j^*$. Hence, when A_j is invertible $(A_j^{-1})^* = A_j^{-1}$, and since the operators $\text{band}(\cdot, \nu)$ and $(\cdot)^*$ commute, we have

$$(\bar{L}_j^b)^* = L_j^* - B_j^* (\text{band}(A_j^{-1}, \nu))^* C_j^* = L_j - C_j \text{band}(A_j^{-1}, \nu) B_j = \bar{L}_j^b. \quad \square$$

Clearly, banding A_j^{-1} generates a compact representation of the homogenized operator \bar{L}_j^b . This can be easily demonstrated in the case of Haar wavelets. We assume that the number of nonzero diagonals (the width) of the different matrices is given by the following table:

Matrix	Width
L_{j+1}	α
A_j, B_j, C_j, L_j	β
$\text{band}(A_j^{-1}, \nu)$	ν
\bar{L}_j^b	δ

The width, α , of the matrix that represents the initial discretization is small for compact approximation. The width β in this case is $\frac{\alpha-1}{2} + 2$ if $\frac{\alpha-1}{2}$ is odd and $\frac{\alpha-1}{2} + 1$ otherwise. Given a width, ν , of $\text{band}(A_j^{-1}, \nu)$, the width, δ , of the homogenized matrix, \bar{L}_j^b , is

$$\delta = \nu + 2(\beta - 1) = \begin{cases} \alpha + \nu + 1 & \text{if } \frac{\alpha-1}{2} \text{ odd,} \\ \alpha + \nu - 1 & \text{otherwise,} \end{cases}$$

and this is indeed a sparse representation of \bar{L}_j^b .

In two dimensions one has to take into account the block structure that results from the tensor product between the different matrices. The structure for the banding operator that was proposed in [1] was the following: given a matrix \mathcal{R} in the form of the tensor product of two $N \times N$ matrices, the two-dimensional banding operator is defined as

$$(3.7) \quad \text{band}(\mathcal{R}, \nu)_{ij} = \begin{cases} \mathcal{R}_{ij} & \text{if } 2|i - j - rN| \leq \nu - 1 - |2r|, \\ 0 & \text{otherwise} \end{cases}$$

for all r such that $|2r| + 1 \leq \nu$.

The extension of (3.7) to banding A_j^{-1} is as follows. If the matrix \bar{L}_j is of dimensions $N^2 \times N^2$, the matrix A_j^{-1} is of dimensions $3N^2 \times 3N^2$, and for all r such that $|2r| + 1 \leq \nu$,

$$(3.8) \quad \text{band}(A_j^{-1}, \nu)_{ik} = \begin{cases} (A_j^{-1})_{ik} & \text{if } 2|(i \bmod N^2) - (k \bmod N^2) - rN| \leq \nu - 1 - |2r|, \\ 0 & \text{otherwise.} \end{cases}$$

Generally, in more than one dimension, manipulating A_j^{-1} is much simpler than manipulating H_j . This can be demonstrated, e.g., with the following case study. Consider the two-dimensional elliptic equation

$$(3.9) \quad -\nabla(a(x, y)\nabla u) = f,$$

and assume that its left-hand side is discretized as

$$(3.10) \quad L_{j+1} = -\frac{1}{h^2} \Delta_+^x \mathcal{A} \Delta_-^x - \frac{1}{h^2} \Delta_+^y \mathcal{A} \Delta_-^y,$$

where \mathcal{A} is the appropriately sampled a . It was shown in [14] that for L_{j+1} given by (3.10), the homogenized operator \bar{L}_j contains the mixed derivatives,

$$(3.11) \quad \bar{L}_j = -\frac{1}{h^2} \Delta_+^x H^{xx} \Delta_-^x - \frac{1}{h^2} \Delta_+^y H^{yx} \Delta_-^x - \frac{1}{h^2} \Delta_+^x H^{xy} \Delta_-^y - \frac{1}{h^2} \Delta_+^y H^{yy} \Delta_-^y,$$

and hence is of a different form than L_{j+1} . In general, the divergence form of the Schur complement is difficult to obtain in the multidimensional case. Equation (3.11) is the divergence form that corresponds to the elliptic operator in (3.9). Even in this case, it is not obvious what the resulting form of the different H 's is and how to obtain a compact representation for them.

For future reference it is important to note that even though we band matrices of different dimensions in (3.7) and (3.8), the parameter ν in both equations plays an identical role. If we define the sparsity of a matrix as the ratio between the number of (presumably) nonzero elements and the total number of elements, then it is easy to show that due to the structure of the banding and the dimensions of the matrices, the sparsity of the banded matrices will be exactly the same in both methods (3.7) and (3.8) for any given value of ν .

In order to get a good approximation of the homogenization solution, the sparse representation, $band(A_j^{-1}, \nu)$, should be a good approximation of A_j^{-1} . We cannot expect this to be the case for a general operator L_{j+1} . However, in many cases of interest, A_j^{-1} is diagonally dominant, and hence it is natural to expect good approximation properties from (3.6). As an example we study the quality of the approximation (3.6) when implemented on the elliptic problem

$$(3.12) \quad -\left(\frac{d}{dx} a(x) \frac{d}{dx} u\right) = f, \quad a(x) > 0, \quad x \in \mathbb{R}.$$

We denote the mesh spacing that corresponds to L_{j+1} by h and perform the homogenization in the Haar basis. We assume that $\mathcal{M}_j a = (\tilde{a} \ \bar{a})^T$ and that the coefficient $a(x)$ oscillates on the mesh scale, $a(x_m) = \bar{a} + (-1)^m \tilde{a}$. Here, \bar{a} and \tilde{a} are assumed to be positive constants such that $0 \leq \tilde{a} < \bar{a}$. Following [12], we now discretize (3.12) as

$$(3.13) \quad -\frac{1}{h^2} \Delta_+ \text{diag}(a) \Delta_- U = F.$$

Dropping the $\text{diag}(\cdot)$ notation, a direct computation gives the decomposition of $-(\Delta_+ a \Delta_-)/h^2$ in the Haar basis as

$$(3.14) \quad \mathcal{M}_j \left(-\frac{1}{h^2} \Delta_+ a \Delta_- \right) \mathcal{M}_j^T = \begin{pmatrix} A_j & B_j \\ C_j & L_j \end{pmatrix}$$

with

$$\begin{aligned} A_j &= c [\bar{a}(\Delta_+ \Delta_- - M M^T) + \tilde{a}(\Delta_+ M^T - M \Delta_-)], \\ B_j &= c [(M - \Delta_+)(\bar{a} + \tilde{a}) \Delta_-], \\ C_j &= -c [\Delta_+(\bar{a} + \tilde{a})(\Delta_- + M^T)], \\ L_j &= 2c \Delta_+(\bar{a} + \tilde{a}) \Delta_-. \end{aligned}$$

Here, $c = -1/(4\sqrt{2}h^2)$, $M = -3I - S_1$, S_n is the shift matrix such that $S_{k,l}^{(n)} = \delta_{k+n,l}$, and Δ_{\pm} denotes the usual forward and backward differencing. The expression (3.14) can be easily computed by decomposing

$$\mathcal{M}_j \left(-\frac{1}{h^2} \Delta_+ a \Delta_- \right) \mathcal{M}_j^T = -\frac{1}{h^2} (\mathcal{M}_j \Delta_+ \mathcal{M}_j^T) (\mathcal{M}_j a \mathcal{M}_j^T) (\mathcal{M}_j \Delta_- \mathcal{M}_j^T).$$

Hence, the homogenized operator is given by

$$(3.15) \quad \begin{aligned} \bar{L}_j &= L_j - C_j A_j^{-1} B_j \\ &= L_j + \frac{1}{\sqrt{2}(2h)^2} \Delta_+ (\bar{a} + \tilde{a}) (\Delta_- + M^T) A_j^{-1} (M - \Delta_+) (\bar{a} + \tilde{a}) \Delta_-. \end{aligned}$$

Equation (3.15) can then be rewritten in the divergence form

$$\bar{L}_j = -\frac{1}{(2h)^2} \Delta_+ H \Delta_-$$

with

$$H = \sqrt{2}(\bar{a} + \tilde{a}) \left[I - (\bar{a} + \tilde{a}) (\Delta_- + M^T) \tilde{A}_j^{-1} (M - \Delta_+) \right]$$

and

$$\tilde{A}_j = 2\sqrt{2}h^2 A_j = -\frac{1}{2} \left[\bar{a} (\Delta_+ \Delta_- - M M^T) + \tilde{a} (\Delta_+ M^T - M \Delta_-) \right].$$

With the identities

$$\begin{aligned} \Delta_- + M^T &= -2(I + S_{-1}), & M - \Delta_+ &= -2(I + S_1), \\ \Delta_+ \Delta_- - M M^T &= -2(S_{-1} + 6I + S_1), & \Delta_+ M^T - M \Delta_- &= -2\Delta_+ \Delta_-, \end{aligned}$$

we can rewrite H as

$$H = \sqrt{2}(\bar{a} + \tilde{a}) \left[I - (\bar{a} + \tilde{a}) (I + S_{-1}) \tilde{A}_j^{-1} (I + S_1) \right],$$

where

$$(3.16) \quad \tilde{A}_j = \bar{a}(S_{-1} + 6I + S_1) + \tilde{a}\Delta_+\Delta_-.$$

Since $0 \leq \tilde{a} < \bar{a}$, the sum of the elements off the main diagonal is $2\bar{a} + 2\tilde{a} < 4\bar{a}$, while the diagonal elements are $6\bar{a} - 2\tilde{a} > 4\bar{a}$, and hence the matrix \tilde{A}_j is diagonally dominant. Rearranging the terms in (3.16) we have

$$\begin{aligned} \tilde{A}_j &= \bar{a}(S_{-1} + 6I + S_1) + \tilde{a}(S_1 + S_{-1} - 2I) \\ &= (6\bar{a} - 2\tilde{a})(I + q(S_1 + S_{-1})), \quad q = \frac{\bar{a} + \tilde{a}}{6\bar{a} - 2\tilde{a}}. \end{aligned}$$

Clearly, $|q| < 1/2$, and hence the following expansion converges:

$$(3.17) \quad \tilde{A}_j^{-1} = \frac{1}{6\bar{a} - 2\tilde{a}} \left(I + \sum_k (-1)^k q^k (S_1 + S_{-1})^k \right).$$

Here, we see the rapid decay of the entries of \tilde{A}_j^{-1} away from the main diagonal. According to (3.17) \tilde{A}_j^{-1} can be written as a converging sequence of the form $\tilde{A}_j^{-1} = \sum_{n \in \mathbb{N}} a_n S_n$ with coefficients a_n that are given by (3.17). Hence

$$\tilde{A}_j^{-1} - \text{band}(\tilde{A}_j^{-1}, \nu) = \sum_{|n| > \frac{\nu-1}{2}} a_n S_n.$$

We now denote the approximation of H that is obtained by banding \tilde{A}_j^{-1} as \tilde{H} , i.e.,

$$\tilde{H} = \sqrt{2}(\bar{a} + \tilde{a})[I - (\bar{a} + \tilde{a})(I + S_{-1})\text{band}(\tilde{A}_j^{-1}, \nu)(I + S_1)].$$

Hence,

$$H - \tilde{H} = \sqrt{2}(\bar{a} + \tilde{a})^2 \mathcal{G}$$

with

$$\begin{aligned} \mathcal{G} &= (I + S_{-1})(\tilde{A}_j^{-1} - \text{band}(\tilde{A}_j^{-1}, \nu))(I + S_1) \\ &= 2 \sum_{|n| > \frac{\nu-1}{2}} a_n S_n + \sum_{|n| > \frac{\nu+1}{2}} a_{n-1} S_n + \sum_{|n| > \frac{\nu-3}{2}} a_{n+1} S_n := \sum_{|n| > \frac{\nu-3}{2}} \gamma_n S_n. \end{aligned}$$

It is clear from the expansion in (3.17) that \tilde{A}_j^{-1} is generated from an infinitely long stencil with an exponential decay rate $\rho = 2q < 1$. Hence, $|a_n| < K\rho^{|n|}$ with $K = 1/(6\bar{a} - 2\tilde{a})$. Therefore, a simple bound for the decay of the elements of $H - \tilde{H}$ away from the main diagonal is

$$\sqrt{2}(\bar{a} + \tilde{a})^2 |\gamma_n| \leq 4\sqrt{2}(\bar{a} + \tilde{a})^2 |a_n| < 4\sqrt{2}(\bar{a} + \tilde{a})^2 K\rho^{|n|} := K'\rho^{|n|}$$

with $K' = 4\sqrt{2}(\bar{a} + \tilde{a})^2 K$. Due to the circulant structure of $H - \tilde{H}$ we can now estimate

$$(3.18) \quad \|H - \tilde{H}\|_\infty = 2 \sum_{n=\frac{\nu-1}{2}}^{\infty} |\gamma_n| < 2K' \sum_{n=\frac{\nu-1}{2}}^{\infty} \rho^n = 2K'\rho^{\frac{\nu-1}{2}} \frac{1}{1-\rho}.$$

Similarly to what was done in [12], the estimate (3.18) can be translated into an estimate for $\bar{L}_j - \bar{L}_j^b$ for \bar{L}_j^b that is given by (3.6). Assume that v is a smooth function, and denote it by $v_i = v(x_i)$. Since H (and \tilde{H}) commute with Δ_+ , we have

$$(\bar{L}_j - \bar{L}_j^b)v_i = (H - \tilde{H})\frac{1}{h^2}\Delta_+\Delta_-v_i = (H - \tilde{H})v''(\xi),$$

where $\xi \in \mathbb{R}$. Hence, taking (3.18) into account, we have

$$(3.19) \quad |(\bar{L}_j - \bar{L}_j^b)v_i| \leq 2K'\rho^{\frac{\nu-1}{2}} \frac{1}{1-\rho} |v''(\xi)|.$$

Taking the sup over all ξ and the maximum over i we can obtain the following theorem.

THEOREM 3.2. *Consider (3.12), which is discretized as (3.13) on the line on a uniform grid with mesh spacing h . Assume that the coefficients oscillate on the mesh*

scale, $a(x_m) = \bar{a} + (-1)^m \tilde{a}$. Let \bar{L}_j be given by (2.6), and let \bar{L}_j^b be given by (3.6). Finally, assume that v is smooth. Then

$$(3.20) \quad \|(\bar{L}_j - \bar{L}_j^b)v\|_\infty \leq C\rho^{\frac{\nu}{2}}\|v''\|_\infty, \quad C = \frac{8\sqrt{2}(\bar{a} + \tilde{a})^2}{(6\bar{a} - 2\tilde{a})\rho^{1/2}(1 - \rho)}.$$

Remark. Theorem 3.2 is similar to Theorem 2 in [12], which deals with banding H_j instead of \tilde{A}_j^{-1} . The following bound can be obtained when banding H_j :

$$\|(\bar{L}_j - \bar{L}_j^b)v\|_\infty \leq C\rho^{\frac{\nu+1}{2}}\|v''\|_\infty$$

with the same constant C as in (3.20). Note that this is a correction of the equivalent expression in [12]. Regarding the proofs of Theorems 1 and 2 in [12] we note the following: (a) The operator on page 544 has the wrong sign and hence is not elliptic. This is the reason why the eigenvalue of A_j given in [12, eq. (11)] should be $-8\bar{a}$ and not $8\bar{a}$ as stated (unless the sign of the operator is reversed). (b) The expression on the bottom of [12, p. 548] should be $\Delta_+ M^T - M \Delta_- = -2\Delta_+ \Delta_-$. (c) The estimate in [12, eq. (14)] should be $|\gamma_n| < K'\rho^{|n|}$. (d) The first line of the proof of Theorem 2 in [12, p. 552] reads $\|H - \text{band}(H, \nu)\| < K' \sum_{n=\nu}^\infty \rho^n$, and this should be corrected (compare with (3.18)).

4. A fine-scales correction. The fine-scales component of the solution can be recovered from (2.3) with the notation in (2.4) and is given by

$$(4.1) \quad U_f = -A_j^{-1} B_j U_c + A_j^{-1} F_f.$$

Here $F_f = Qf$ with $f \in V_{j+1}$ and $F_f \in W_j$. Equation (4.1) holds also in more than one dimension.

A straightforward observation is that if the homogenization procedure is repeated several times, then in the last step we already know all the ingredients that are required for computing U_f via (4.1). This is true for the final homogenization step, because that is when we invert \bar{L}_j to compute U_c , and since U_c , A_j^{-1} , as well as the other terms in (4.1), are known at this last step, computing U_f requires no additional excessive computational cost besides multiplying known matrices. Once U_f is known it can be combined with U_c to recover a better resolved solution U .

In two dimensions, the fine-scales correction is also given by (4.1). In this case, the solution should be recovered from four components, as the dimensionality of the fine-scales part of the solution is three times the dimensionality of the coarse-scales part of the solution, as given in (2.8).

In some sense, once we involve fine scales in the computation, it is less natural to address the method as a “subgrid” method. Nevertheless, being able to improve the solution using the fine scales from the final stage with no additional computational cost is a positive consequence of this formulation and should not be ignored.

We would also like to note that in some cases, the additional structure of the problem reduces the advantages in making such a fine-scales correction. For example, if one considers hyperbolic problems which are discretized with a semidiscrete numerical scheme, homogenizing only the spatial part of the operator and evolving in time using an ordinary differential equation (ODE) solver, one can still improve the quality of the homogenized operator by adding the fine scales. However, this will increase the size of the ODE system that then has to be solved, so it might actually be too computationally expensive to implement this improvement.

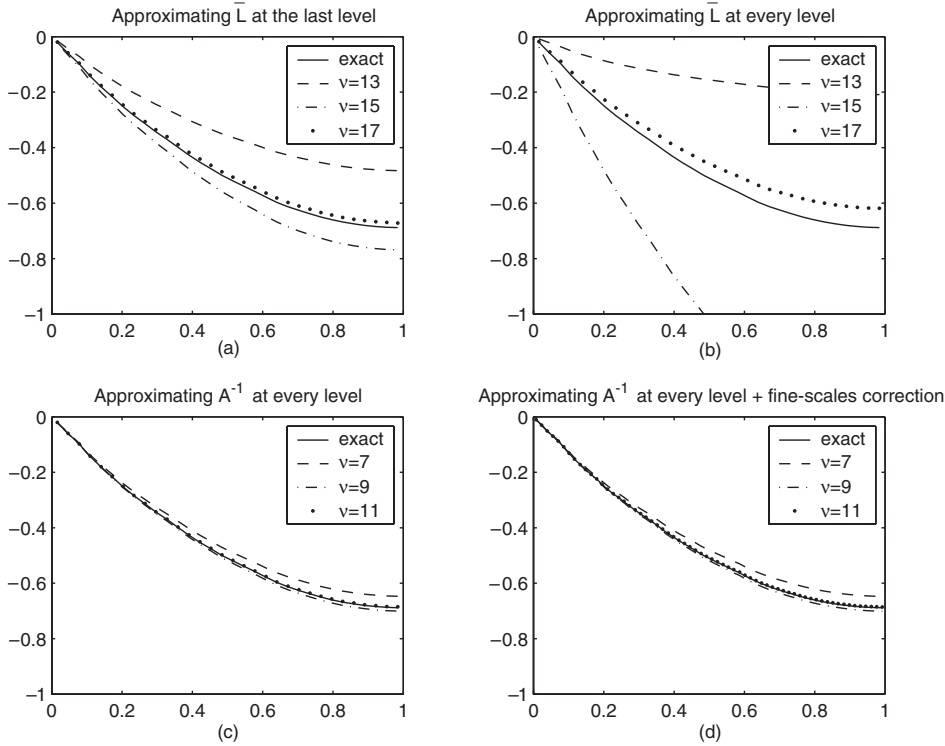


FIG. 1. The one-dimensional elliptic problem (5.1) discretized as (5.2)–(5.3). $a(x)$ is given by the random distribution (5.4). We show three homogenization steps starting from $n = 256$.

5. Numerical examples. In this section we demonstrate the performance of our method on a series of test cases taken from [1, 14]. The examples are a one-dimensional elliptic model problem with discontinuous or random coefficients, a Helmholtz equation in one and two dimensions, and a one-dimensional hyperbolic problem with variable coefficients.

Example 1: One-dimensional elliptic model problem. We consider the elliptic problem

$$(5.1) \quad \begin{cases} -\frac{d}{dx} \left(a(x) \frac{du}{dx} \right) = f(x), \\ u(0) = u'(1) = 0 \end{cases}$$

with $a(x) > 0$, which we discretize as

$$(5.2) \quad -\frac{1}{h^2} \Delta_+ a_i \Delta_- u_i = f_i, \quad i = 1, \dots, n,$$

with $f_i = -1$ for all i . The boundary conditions are approximated by

$$(5.3) \quad u_0 = -u_1, \quad u_{n+1} = u_n.$$

In Figure 1 we show the results obtained when the coefficient $a(x)$ is taken as

$$(5.4) \quad a(x) = U[0.5, 1].$$

Here $U[b, d]$ denotes a uniform random distribution in the interval $[b, d]$. We start with $n = 256$ and reduce the number of variables to $n = 32$ coarse variables using three homogenization steps.

In Figure 1(a), the operator \bar{L}_j is truncated according to (3.1) after the last homogenization step. ν represents the number of diagonals that is kept in \bar{L}_j . In Figure 1(b) we show the results obtained when \bar{L}_j is approximated (again by band truncation) at every homogenization step. The results are much worse than those in Figure 1(a). In Figure 1(c) we approximate A_j^{-1} according to (3.1) at every homogenization level. Here, ν is the number of diagonals we choose to keep in A_j^{-1} . Finally, in Figure 1(d) we show the solution after we apply the fine-scales correction at the last level. This is done after truncating A_j^{-1} at every step, similarly to what was shown in Figure 1(c). In this case, due to the regularity of the solution, the fine-scales correction does not seem to improve the solution. The positive effect of the fine-scales correction will be more apparent in problems that have nonsmooth solutions.

The “exact solution” is the solution obtained on a coarse grid with the exact homogenization procedure, i.e., with no intermediate approximations. In Figure 1(d), the exact solution is the cell averages with $n = 64$; in all other Figures 1(a)–(c) the exact solution is the cell averages with $n = 32$.

In Figure 2 we show the results obtained when the coefficient $a(x)$ is taken as a slit, i.e.,

$$(5.5) \quad a(x) = \begin{cases} 1/6, & 0.45 < x < 0.55, \\ 1 & \text{otherwise.} \end{cases}$$

Once again, we reduce the number of variables from $n = 256$ to $n = 32$ with three homogenization steps. The order of the panels is similar to the one in Figure 1. In this case, the solution is significantly improved with the fine-scales correction. The interfaces between the smooth regions are captured with a much better resolution.

Example 2: One-dimensional Helmholtz equation. We now consider the one-dimensional Helmholtz equation

$$(5.6) \quad \begin{cases} -\frac{d}{dx} \left(a(x) \frac{du}{dx} \right) - \omega^2 u = 0, \\ u(0) = 1, \\ u'(1) = 0. \end{cases}$$

We assume that the mass matrix $a(x)$ given by (5.5), set $\omega = 2\pi$, and discretize (5.6) as

$$(5.7) \quad -\frac{1}{h^2} \Delta_+ a_i \Delta_- u_i - \omega^2 u_i = 0, \quad i = 1, \dots, n.$$

The boundary conditions are approximated by

$$(5.8) \quad u_0 = 2u(0) - u_1, \quad u_{n+1} = u_n.$$

Once again, the number of variables, n , is reduced from 256 to 32 in three homogenization steps. The results are shown in Figure 3 for $\nu = 7, 9, 11$. The arrangement of the panels is similar to those in Figure 2. Clearly, the solution is significantly improved with the fine-scales correction, as the resolution of the interfaces between the smooth regions is significantly better.

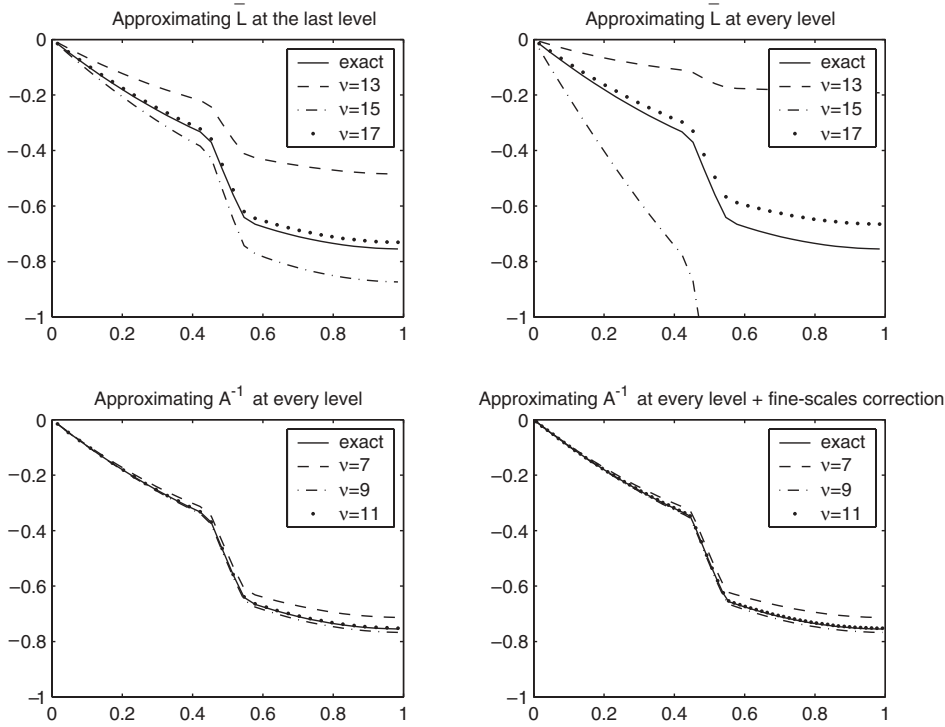


FIG. 2. The one-dimensional elliptic problem (5.1) discretized as (5.2)–(5.3). $a(x)$ is given by the slit (5.5). We show three homogenization steps starting from $n = 256$.

Example 3: Two-dimensional Helmholtz equation. We consider the two-dimensional Helmholtz equation

$$(5.9) \quad \nabla \cdot (a(x, y) \nabla u) + \omega^2 u = 0$$

with periodic boundary conditions in the y -direction and $u(0, y) = 1$, $u_x(1, y) = 0$ on the other two boundaries. This problem corresponds to a plane time-harmonic wave of amplitude one entering the domain at $x = 0$ and leaving at $x = 1$. Following [1] we discretize (5.9) as

$$(5.10) \quad \frac{1}{h^2} \Delta_+^x a_{il} \Delta_-^x u_{il} + \frac{1}{h^2} \Delta_+^y a_{il} \Delta_-^y u_{il} + \omega^2 u_{il} = 0, \quad i, l = 1, \dots, n,$$

$$u_{i,0} = u_{i,n}, \quad u_{i,n+1} = u_{i,1}, \quad u_{n+1,l} = u_{n,l}, \quad u_{0,l} = 2 - u_{1,l}.$$

Here, Δ_\pm^x denotes the usual forward (+) or backward (−) differencing in the x -direction, and a similar notation holds for the y -direction. We solve (5.10) for $\omega = 3\pi$, $n = 48$, and a slit-type $a(x, y)$ that is given by

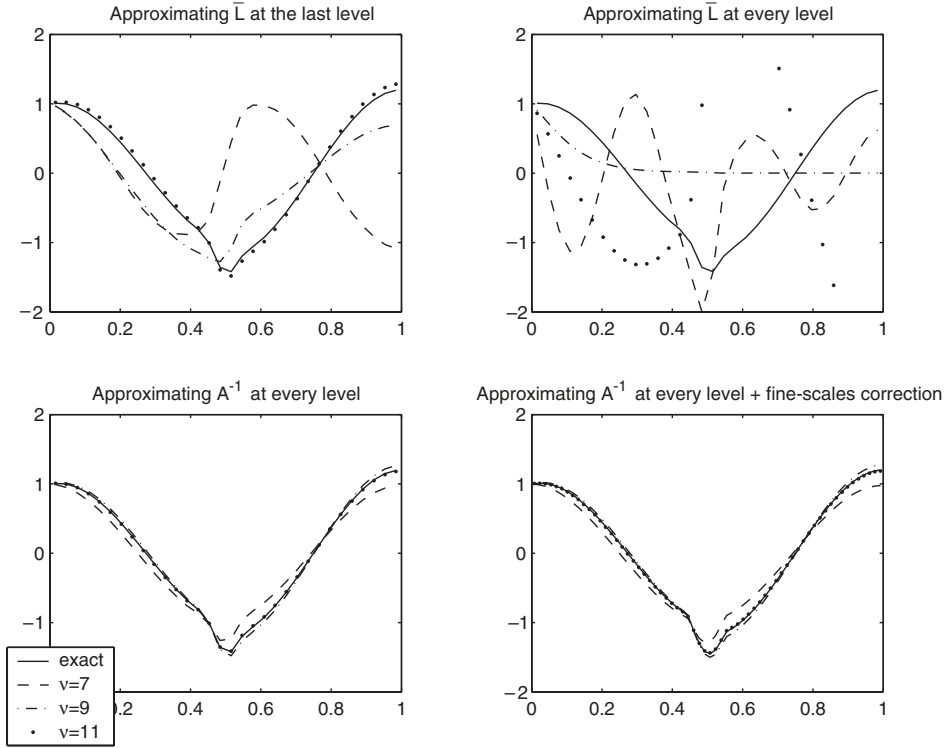


FIG. 3. The one-dimensional Helmholtz equation (5.6) with a slit mass matrix (5.5), discretized as (5.7)–(5.8). We show three homogenization steps starting from $n = 256$.

$$(5.11) \quad a(x, y) = \begin{cases} 10^{-4}, & 0.4 < x < 0.5, \quad |y - 0.5| > 0.05, \\ 1 & \text{otherwise.} \end{cases}$$

The results obtained with one, two, and three homogenization steps according to (2.7) are shown in Figure 4. In Figure 5 we show what happens when the operator \bar{L}_j is truncated after one homogenization step according to (3.7) with $\nu = 5, 7, 9$. Note the large oscillations for $\nu = 5, 7$. In Figure 6 we show the results obtained when truncating A_j^{-1} during the first homogenization step according to (3.8). Here, $\nu = 1, 3, 5$. Clearly, $\nu = 5$ gives results that are nearly identical to the exact homogenized solution. This case of $\nu = 5$ corresponds to a compression rate of approximately 2.2% of the original size of A_j^{-1} . This has to be compared with the case of $\nu = 9$ in banding \bar{L}_j , which is the first instance in which the results are good when truncating \bar{L}_j . The case of $\nu = 9$ corresponds to a compression rate of approximately 6.7% of the original size of \bar{L}_j . In banding A_j^{-1} even the results with $\nu = 1$ are fine in most of the domain, and this is for a compression rate of approximately 0.2%. In Figure 7 we perform two homogenization steps starting from $n = 64$. In this figure we compare banding \bar{L}_j only in the last homogenization step with banding A_j^{-1} in both homogenization steps. In both cases we choose $\nu = 5, 7, 9$. Clearly, the results with banding A_j^{-1} are significantly better and provide a good approximation already for $\nu = 5$. In Figure 8 we demonstrate the fine-scales correction of section 4 with this two-dimensional problem. Here, we implement the fine-scales correction after one homogenization step in which

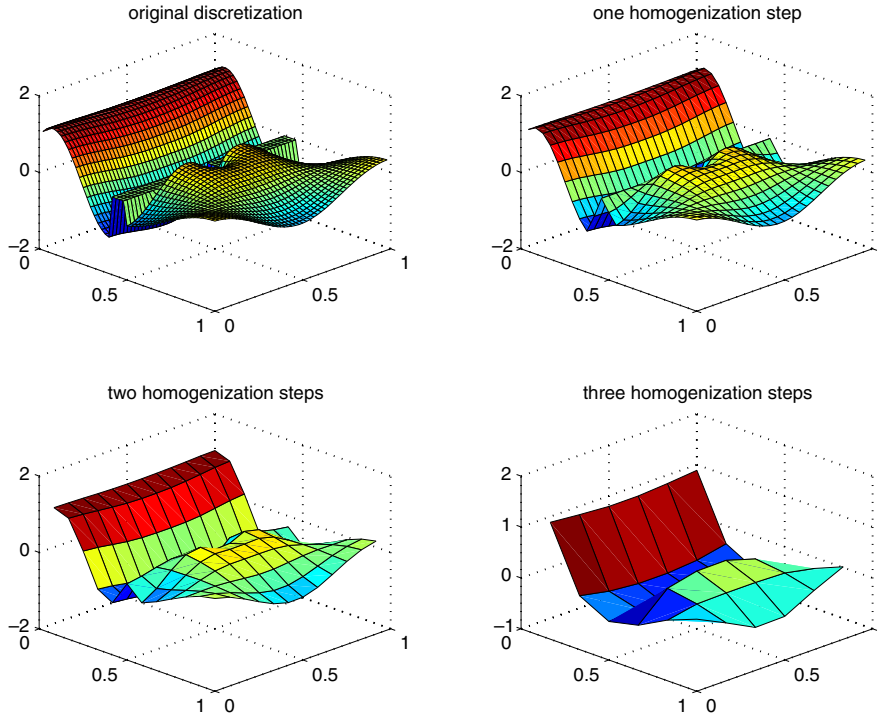


FIG. 4. Results for the two-dimensional Helmholtz example (5.9)–(5.11) showing the solution of the original discretization of dimension 48×48 and the coarse scales solution after one, two, and three homogenization steps.

we wither band \bar{L}_j or A_j^{-1} (in both cases with $\nu = 7$). The figure shows that in both cases, the results improve as a result of the fine-scales correction. This should be compared with the “exact” solution of the original discretization that is given in the upper-left image in Figure 4. Finally, in Figure 9 we show the sparse structure of the operators \bar{L}_j and A_j^{-1} . In each matrix we present elements larger than 0.1% of the max value.

Example 4: Hyperbolic problems. We consider the one-dimensional inhomogeneous advection equation

$$(5.12) \quad \frac{\partial u}{\partial t} + a(x) \frac{\partial u}{\partial x} = f(x, t),$$

which is augmented with the initial data $u(x, 0)$ and subject to the boundary condition $u(0, t) = 0$. Here, we assume that the coefficient $a(x)$ is positive and bounded.

There are different ways of homogenizing (5.12). We follow [1, 14] and consider a semidiscrete approximation of (5.12) in which the time derivative is kept and the spatial operator is discretized. The homogenization is then conducted only on the spatial operator. Alternatively, one can discretize (5.12) both in space and in time and homogenize the resulting two-dimensional operator.

Following [1, 14] we write a semidiscrete approximation of (5.12) as

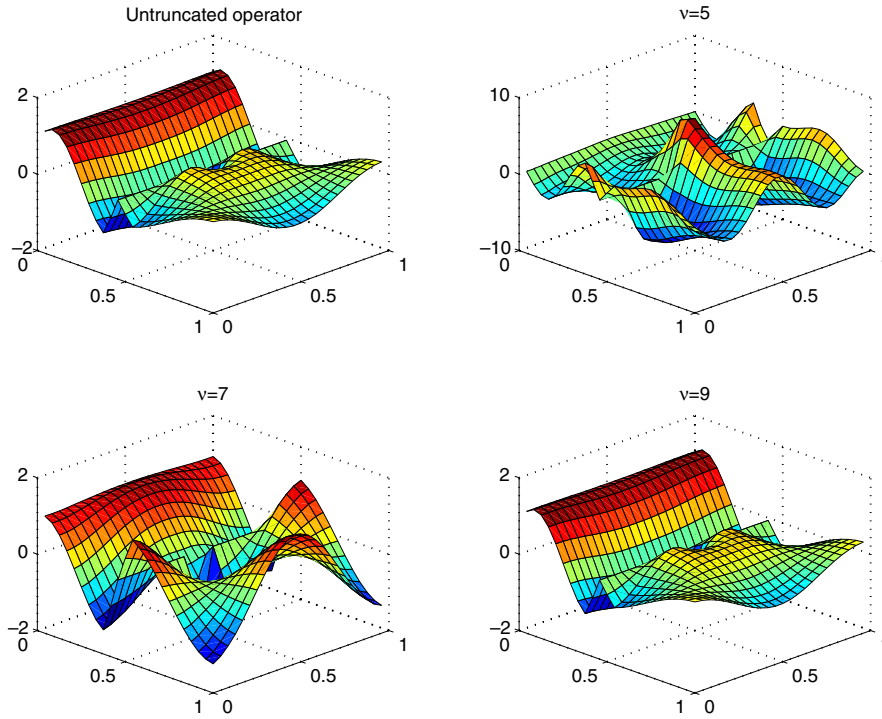


FIG. 5. Results for the two-dimensional Helmholtz example (5.9)–(5.11) with one homogenization step starting from $n = 48$. The upper-left image shows the solution obtained with an untruncated operator. In the other images, \bar{L}_j is truncated with different values of ν according to (3.7). Note the completely different vertical scale in the upper-right image.

$$(5.13) \quad \frac{\partial u_i}{\partial t} + \frac{1}{h} a_i \Delta_- u_i = f_i(t), \quad i = 1, \dots, n, \quad u_0 = -u_1, \quad u_i(0) = b_i,$$

with $a_i = a((i - 1/2)h)$. We assume $f \equiv 0$ and compute the solution at $T = 1$ with initial data that is taken as

$$(5.14) \quad u(x, 0) = \begin{cases} \sin^2(4\pi x), & 0 \leq x \leq 0.25, \\ 0, & 0.25 < x \leq 1. \end{cases}$$

We homogenize the spatial operator only and solve the resulting system of ODEs with a fourth-order Runge–Kutta solver.

In Figure 10 we present the results obtained for $a(x)$ that is given by (5.5). The solution is shown for $n = 128$ (left column) and $n = 256$ (right column). Figures 10(a)–(b) show the exact homogenized solution that is obtained with two homogenization steps. Figures 10(c)–(d) show the result that is obtained when truncating the operator \bar{L}_j at the last stage only. Figures 10(e)–(f) show the results obtained when truncating A_j^{-1} at every homogenization level. Since all the matrices involved in this computation are lower triangular, the parameter ν counts the number of diagonals we keep in the truncation starting from the main diagonal.

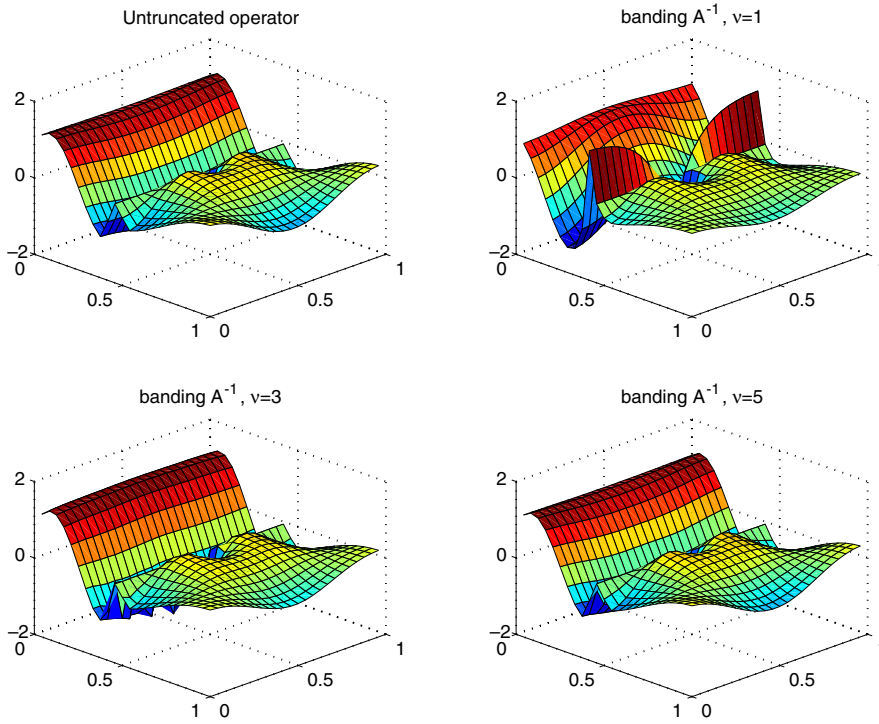


FIG. 6. Results for the two-dimensional Helmholtz example (5.9)–(5.11) with one homogenization step starting from $n = 48$. The upper-left image shows the solution obtained with an untruncated operator. In the other images, A_j^{-1} is truncated with different values of ν according to (3.8).

Figure 11 presents the results obtained for $a(x) = U[0.1, 2]$ (a uniform random distribution). The solution is shown for $n = 256$ (left column) and $n = 512$ (right column). The order of the panels is similar to Figure 10. In both examples not only does the truncation of A_j^{-1} reduce the complexity of the computations, but it also provides a substantial improvement in the quality of the results.

Finally, we would like to note that it is still possible to increase the quality of the approximation with a fine-scales correction. However, since this correction applies to the homogenized operator, it will not help a lot in this case. It is still required to solve a system of ODEs, which due to the fine-scales correction will double its size.

6. Conclusions. In this paper we showed how to improve the efficiency of certain wavelet-based numerical homogenization algorithms with a new compact representation of the homogenized operator. Our method becomes particularly simple in two dimensions when it is difficult to find the divergence form of the Schur complement. Some theoretical aspects of the method were demonstrated in Theorem 3.2 for a one-dimensional problem on an infinite domain. While proving this theorem, we also corrected some of the mistakes in [12].

It is important to note that at present, there is no theoretical understanding why this class of methods is successful with handling multidimensional problems. Such theoretical issues, as well as the implementation of this technique to problems of significant physical interest, remain the subject of future work.

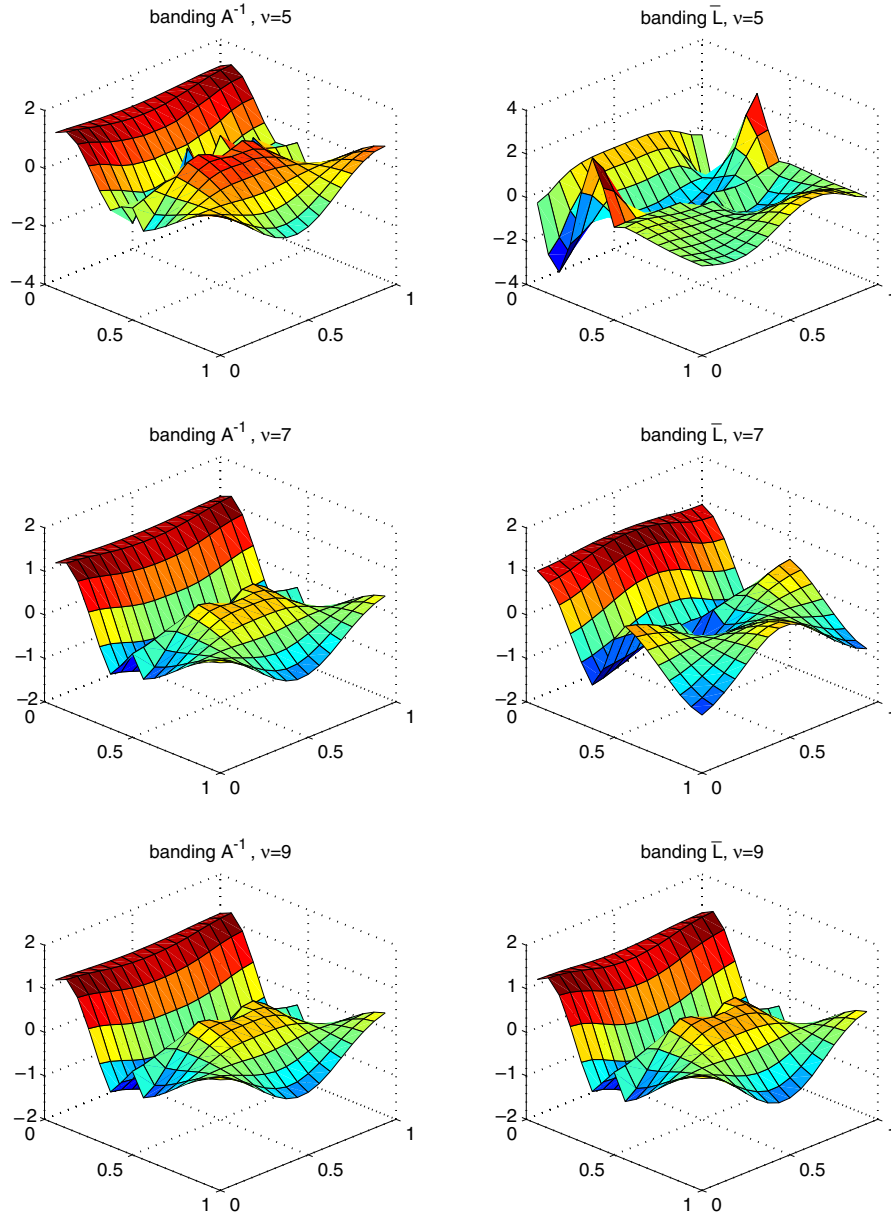


FIG. 7. Results for the two-dimensional Helmholtz example (5.9)–(5.11) using two homogenization steps starting from $n = 64$. Left column: A_j^{-1} is truncated in both homogenization steps according to (3.8). Right column: \bar{L}_j is truncated only in the last homogenization step according to (3.7). For both matrices, $\nu = 5, 7, 9$.

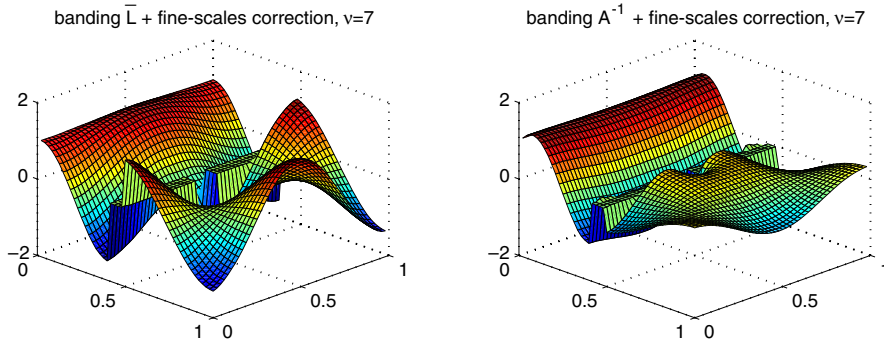


FIG. 8. Results for the two-dimensional Helmholtz example (5.9)–(5.11) using one homogenization step starting from $n = 48$ with a fine-scales correction. Left: \bar{L}_j is truncated according to (3.7). Right: A_j^{-1} is truncated according to (3.8). In both cases, $\nu = 7$.

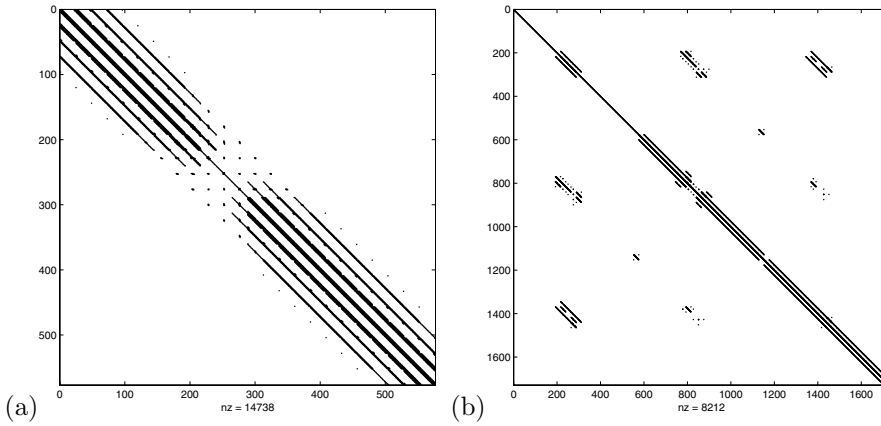


FIG. 9. The structure of the operators (a) \bar{L}_j and (b) A_j^{-1} after one homogenization step for the two-dimensional Helmholtz example (5.9)–(5.11). Elements larger than 0.1% of max value are shown.

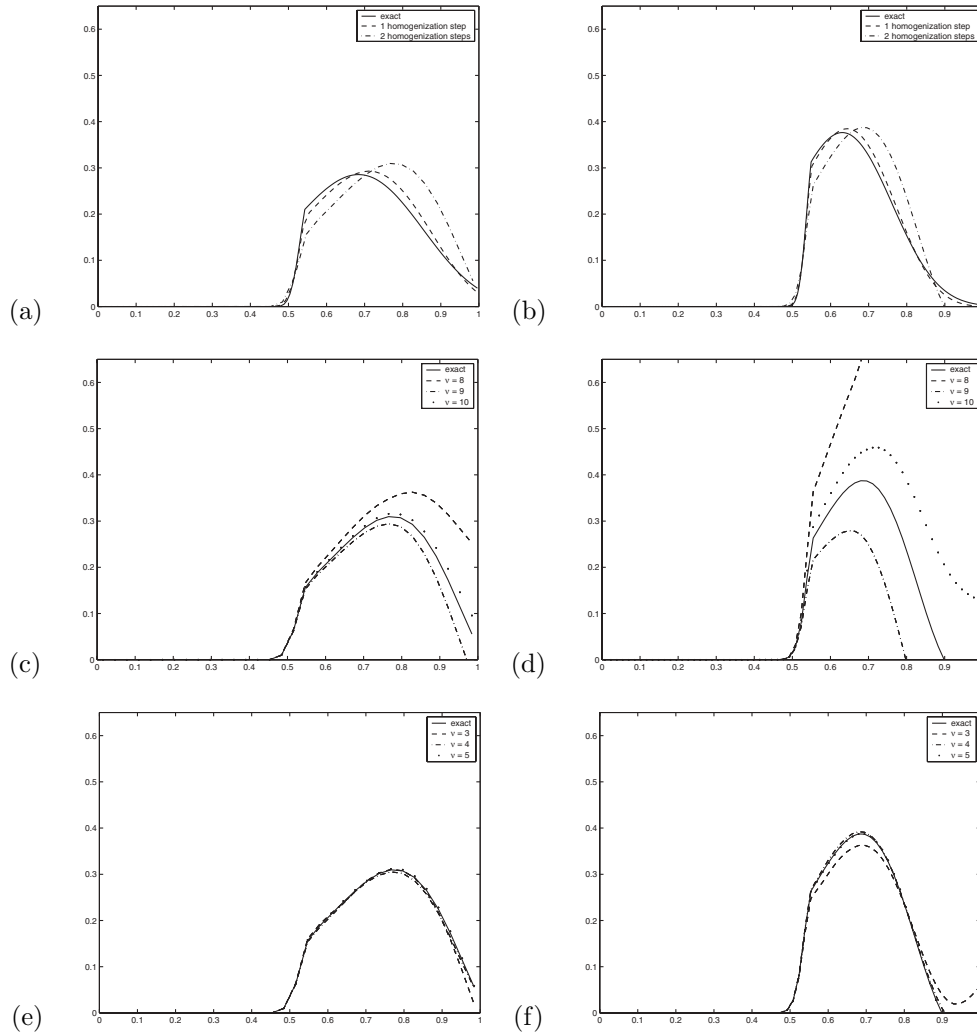


FIG. 10. Hyperbolic equation (5.12) at $T = 1$ with a as in (5.5) and initial conditions (5.14). Left column: $n = 128$. Right column: $n = 256$. (a)–(b): Exact homogenized solution with two homogenization levels. (c)–(d): Approximation of \bar{L}_j at the last stage. (e)–(f): Approximation of A_j^{-1} at every level. ν is the number of diagonals we keep starting from the main diagonal.

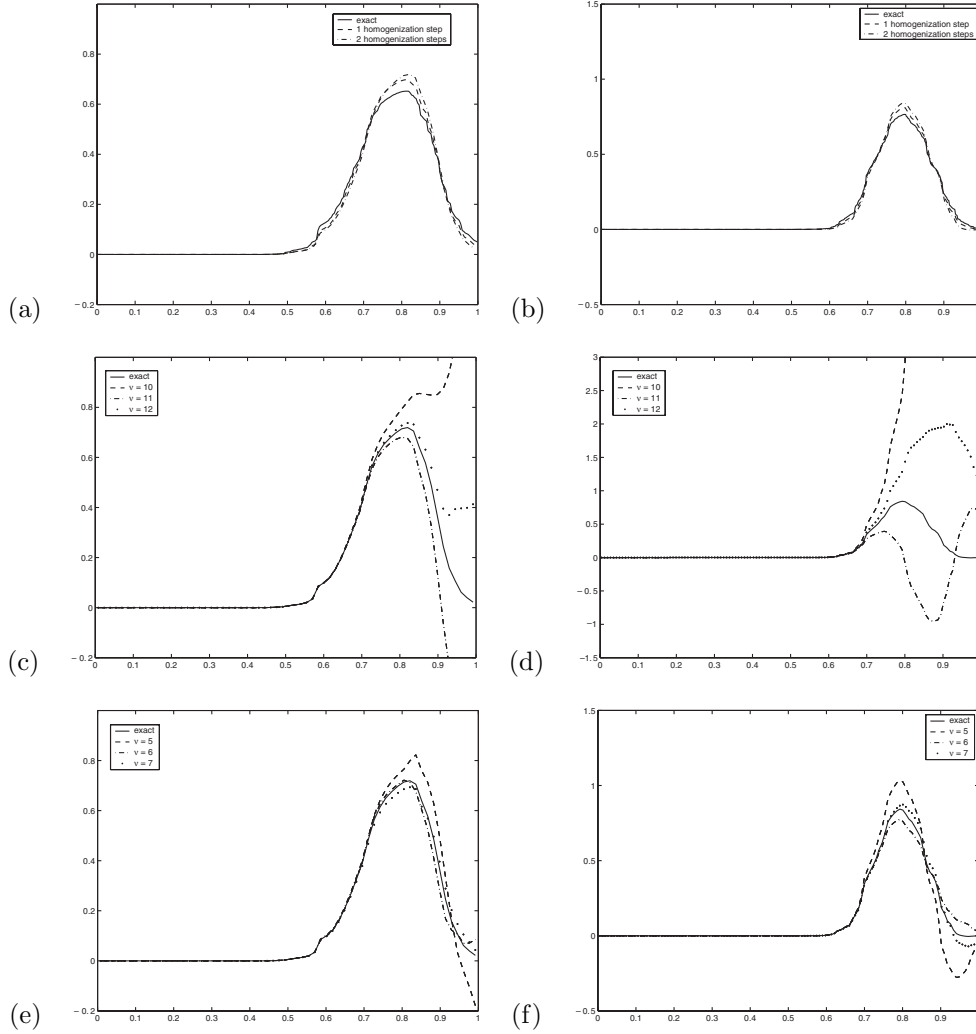


FIG. 11. Hyperbolic equation (5.12) at $T = 1$ with a uniformly distributed in $[0.1, 2]$ and initial conditions (5.14). Left column: $n = 256$. Right column: $n = 512$. (a)–(b): Exact homogenized solution with two homogenization levels. (c)–(d): Approximation of \bar{L}_j at the last stage. (e)–(f): Approximation of A_j^{-1} at every level. ν is the number of diagonals we keep starting from the main diagonal. Note the different scale in (d).

Acknowledgment. We would like to thank Jonathan Mattingly for helpful discussions and for his help in coding the basic homogenization algorithm in the preliminary stages of this work.

REFERENCES

[1] U. ANDERSON, B. ENGQUIST, G. LEDFELT, AND O. RUNBORG, *A contribution to wavelet based subgrid modeling*, Appl. Comput. Harmon. Anal., 7 (1999), pp. 151–164.
 [2] O. AXELSSON, *Iterative Solution Methods*, Cambridge University Press, Cambridge, UK, 1994.
 [3] A. BENSOUSSAN, J.-L. LIONS, AND G. PAPANICOLAOU, *Asymptotic Analysis for Periodic Structures*, North-Holland, Amsterdam, 1978.

- [4] G. BEYLKIN, *On multiresolution methods in numerical analysis*, Doc. Math., 3 (1998), pp. 481–490.
- [5] G. BEYLKIN, *Wavelets and fast numerical algorithms*, in Different Perspectives on Wavelets, Proc. Sympos. Appl. Math. 47, AMS, Providence, RI, 1993, pp. 89–117.
- [6] G. BEYLKIN, R. COIFMAN, AND V. ROKHLIN, *Fast wavelet transforms and numerical algorithms I*, Comm. Pure Appl. Math., 44 (1991), pp. 141–183.
- [7] G. BEYLKIN AND N. COULT, *A multiresolution strategy for the reduction of elliptic PDEs and eigenvalue problems*, Appl. Comput. Harmon. Anal., 4 (1998), pp. 129–155.
- [8] G. BEYLKIN AND J. M. KEISER, *On the adaptive numerical solution of nonlinear partial differential equations in wavelet bases*, J. Comput. Phys., 132 (1997), pp. 233–259.
- [9] M. BREWSTER AND G. BEYLKIN, *A multiresolution strategy for numerical homogenization*, Appl. Comput. Harmon. Anal., 2 (1995), pp. 327–349.
- [10] T. F. C. CHAN AND T. P. MATHEW, *The interface probing technique in domain decomposition*, SIAM J. Matrix Anal. Appl., 13 (1992), pp. 212–238.
- [11] M. DOROBANTU, *Wavelet-Based Algorithms for Fast PDE Solvers*, Dissertation, Royal Institute of Technology, Stockholm University, Stockholm, Sweden, 1995.
- [12] M. DOROBANTU AND B. ENGQUIST, *Wavelet-based numerical homogenization*, SIAM J. Numer. Anal., 35 (1998), pp. 540–559.
- [13] B. ENGQUIST AND T. Y. HOU, *Particle method approximation of oscillatory solutions to hyperbolic differential equations*, SIAM J. Numer. Anal., 26 (1989), pp. 289–319.
- [14] B. ENGQUIST AND O. RUNBORG, *Wavelet-based numerical homogenization with applications*, in Multiscale and Multiresolution Methods, Lect. Notes Comput. Sci. Eng. 20, Springer, Berlin, 2002, pp. 97–148.
- [15] A. GILBERT, *A comparison of multiresolution and classical homogenization schemes*, Appl. Comput. Harmon. Anal., 5 (1998), pp. 1–35.
- [16] W. HACKBUSCH, *The frequency decomposition multi-grid method. I. Application to anisotropic equations*, Numer. Math., 56 (1989), pp. 229–245.
- [17] W. HACKBUSCH, *The frequency decomposition multi-grid method. II. Convergence analysis based on the additive Schwarz method*, Numer. Math., 63 (1992), pp. 433–453.
- [18] T. Y. HOU AND X.-H. WU, *A multiscale finite elements method for elliptic problems in composite materials and porous media*, J. Comput. Phys., 134 (1997), pp. 169–189.
- [19] 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 (1999), pp. 913–943.
- [20] S. KNAPEK, *Matrix-dependent multigrid homogenization for diffusion problems*, SIAM J. Sci. Comput., 20 (1998), pp. 515–533.
- [21] S. MALLAT, *A Wavelet Tour of Signal Processing*, 2nd ed., Academic Press, Cambridge, UK, 1999.
- [22] A.-M. MATAACHE AND C. SCHWAB, *Finite dimensional approximations for elliptic problems with rapidly oscillating coefficients*, in Multiscale Problems in Science and Technology, Springer, Berlin, 2002, pp. 203–242.
- [23] Y. MEYER, *Wavelets and Operators*, Cambridge University Press, Cambridge, UK, 1992.
- [24] C. SCHWAB AND A.-M. MATAACHE, *Generalized FEM for homogenization problems*, in Multiscale and Multiresolution Methods, Lect. Notes Comput. Sci. Eng. 20, Springer, Berlin, 2002, pp. 197–237.
- [25] C. WAGNER AND G. WITTUM, *Filtering decompositions with respect to adaptive test vectors*, in Multigrid Method V, Lect. Notes Comput. Sci. Eng. 3, Springer, Berlin, 1998, pp. 320–334.