APPLIED AND COMPUTATIONAL HARMONIC ANALYSIS **5,** 129–155 (1998)
ARTICLE NO. HA970226

A Multiresolution Strategy for Reduction of Elliptic PDEs and Eigenvalue Problems

Gregory Beylkin* and Nicholas Coult†

Department of Applied Mathematics, University of Colorado at Boulder, Boulder, Colorado 80309-0526

Communicated by Wolfgang Dahmen

Received April 15, 1996; revised May 12, 1997

In many practical problems coefficients of PDEs are changing across many spatial or temporal scales, whereas we might be interested in the behavior of the solution only on some relatively coarse scale. We approach the problem of capturing the influence of fine scales on the behavior of the solution on a coarse scale using the multiresolution strategy. Considering adjacent scales of a multiresolution analysis, we explicitly eliminate variables associated with the finer scale, which leaves us with a coarse-scale equation. We use the term reduction to designate a recursive application of this procedure over a finite number of scales.

We present a multiresolution strategy for reduction of self-adjoint, strictly elliptic operators in one and two dimensions. It is known that the non-standard form for a wide class of operators has fast off-diagonal decay and the rate of decay is controlled by the number of vanishing moments of the wavelet. We prove that the reduction procedure preserves the rate of decay over any finite number of scales and therefore results in sparse matrices for computational purposes. Furthermore, the reduction procedure approximately preserves small eigenvalues of self-adjoint, strictly elliptic operators. We also introduce a modified reduction procedure which preserves the small eigenvalues with greater accuracy than the standard reduction procedure and obtain estimates for the perturbation of those eigenvalues. Finally, we discuss potential extensions of the reduction procedure to parabolic and hyperbolic problems.

1. INTRODUCTION

Coefficients of partial differential equations are often changing across many spatial or temporal scales, whereas we might be interested in the behavior of the solution only on some relatively coarse scale. In such a case one would like to find a set of

* This research was partially supported by DARPA Grant F49620-93-1-0474 and ONR Grant N00014-91-J4037.

† This research was supported by an NSF Graduate Research Fellowship.

1063-5203/98 $25.00

equations on a coarse scale that reproduces the solution on that scale. The difficulty, of course, is that such a solution is influenced by the finer scale behavior of the coefficients. Capturing the influence of fine scales (exactly or approximately) on the behavior of the solution on the coarse scale is the problem generally known as that of homogenization (although particular formulations may be rather different).

Typically such problems were addressed by using asymptotic methods or weak limits; see for example, [2, 7, 15, 22, 27, and references therein]. The basic limitation of these methods is that they require the fine scale behavior to be fairly well separated from the behavior on the coarser scales, so that small parameters may be found in the problem. Recently, a multiresolution strategy for homogenization has been proposed in [4]. Using the notion of multiresolution analysis (MRA), we consider the transition between two adjacent scales explicitly. Namely, one obtains an equation for the projection of the solution on the coarser scale. This procedure (the so-called reduction) may then be repeated over many scales and thus does not require the small parameter assumptions typical for asymptotic methods.

The basic step of the reduction involves computing a Schur complement. (The use of the Schur complement in multilevel methods is not new and plays a role in algebraic multigrid and domain decomposition methods (see, e.g., [8, 25]). Steinberg and McCoy in [26] use the Schur complement for multiresolution effective medium computations. Additionally, Knapek in [19] has used the Schur complement for a multigrid-based homogenization technique.) Two problems have to be addressed in order for the multiresolution strategy for homogenization to be a practical method. First, the transition between the two scales has to be computationally efficient. However, simply truncating the matrices as has been suggested in some of the references mentioned above is not satisfactory since there is no control of the quality of the approximation. Second, the form of equations has to be preserved so that one can use the reduction step in a recursive manner. By the ''form of the equations'' we understand either algebraic form or some alternative algebraic structure. The only requirement is that it may be used recursively. The meaning of this remark will become clear below.

In [4] the multiresolution strategy for reduction and homogenization has been applied to a system of linear ordinary differential equations. It is observed in [4] that the transition between two consecutive scales may be achieved by eliminating variables locally and that a certain algebraic form of the equations is preserved, thus permitting a multiscale reduction.

Gilbert in [11] has applied the approach in [4] to a system of two ordinary differential equations equivalent to the one-dimensional elliptic problem. It turns out that in this case one reproduces with some modifications the classical results for the homogenization problem, thus establishing a connection between the multiresolution strategy and the classical approach to homogenization. Dorobantu in [10] also applies the technique of MRA homogenization to the one-dimensional elliptic problem, but in a different manner than in [11]. Nevertheless, Dorobantu also derives results related to the classical homogenization theory. In both Gilbert's and Dorobantu's consideration, the Haar basis is used.

In this paper we further develop the multiresolution strategy for reduction of elliptic partial differential equations. One of the important points of our approach is the use of high-order wavelets (or MRA associated with such wavelets) rather than the Haar basis. There are two compelling reasons for using high-order wavelets in elliptic

problems. First, the use of high-order wavelets permits us to develop efficient numerical methods for the reduction procedure. Second, we demonstrate that if the MRA is chosen correctly for a given problem, then the small eigenvalues of the reduced operators differ only slightly from those of the original operator. In particular, correctly here means that the basis must have a sufficient number of vanishing moments. As a result, we obtain a method for constructing a low-dimensional approximation to a multiscale elliptic problem such that this approximation accounts for both small eigenvalues and the corresponding eigenvectors. Computing these quantities is desirable in many applications, e.g., computational chemistry, although we do not address such problems here.

A model equation that we consider is of the form

$$-\nabla\cdot(a(\mathbf{x})\nabla) = f(\mathbf{x}) \tag{1.1}$$

with periodic boundary conditions. However, our method is applicable to other boundary conditions as well, in which case the wavelet basis has to satisfy the boundary conditions. For the eigenvalue problem we also assume that the ratio $\max a(\mathbf{x})/\min a(\mathbf{x})$ is moderate in size over the domain.

We are not familiar with prior numerical algorithms of the type presented in this paper. Although similar goals have been sought by multigrid methods (see [6, 19–21]), the approach and the results of this paper appear to be different. Papers on classical homogenization of elliptic eigenvalue problems (see, e.g., [17, 18]) also yield a different type of results.

We start in Section 2 by introducing notation and briefly reviewing related results on homogenization of elliptic equations. In Section 3 we address the problem of the multiresolution reduction for elliptic equations using high-order wavelets. We prove that the rate of the off-diagonal decay of the blocks of the reduced operator is preserved and is the same as that of the blocks of the non-standard form. Thus, the reduced operator is compressible in wavelet bases and, for a given accuracy, the sparsity is controlled by the number of vanishing moments of the wavelet basis. We demonstrate that the spectral bounds for the reduced operator on all scales are the same as those for the original operator. We obtain estimates for small eigenvalues which show that the reduced operator is better at preserving them than the projection of the original operator on the corresponding scale. We also introduce a modified reduction procedure to improve the accuracy of preservation of small eigenvalues. Finally, in Section 4 we present results of our numerical experiments.

2. PRELIMINARY CONSIDERATIONS

2.1. Notation

In this section we set our notation and give a brief description of the concept of multiresolution analysis and wavelets. For the details we refer to, e.g., [9]. As usual, we consider a chain of subspaces

$$\cdots \subset \mathbf{V}_2 \subset \mathbf{V}_1 \subset \mathbf{V}_0 \subset \mathbf{V}_{-1} \subset \mathbf{V}_{-2} \subset \cdots$$

such that

$$\bigcap_j \mathbf{V}_j = \{0\} \qquad \text{and} \qquad \overline{\bigcup_j \mathbf{V}_j} = \mathbf{L}^2(\mathbf{R}^d).$$

If $d = 1$, the subspace $\mathbf{V}_j$ is spanned by an orthonormal basis $\{\phi_k^j(x) = 2^{-j/2}\phi(2^{-j}x - k)\}_{k\in\mathbf{Z}}$. The function ϕ is the so-called scaling function and satisfies the two-scale difference equation $\phi(x/2) = \sum_k h_k\phi(x - k)$.

We denote by $\mathbf{W}_j$ the orthogonal complement of $\mathbf{V}_j$ in $\mathbf{V}_{j-1}$, $\mathbf{V}_{j-1} = \mathbf{V}_j \oplus \mathbf{W}_j$ and use $\mathbf{P}_j$ and $\mathbf{Q}_j$ to denote the projection operators onto $\mathbf{V}_j$ and $\mathbf{W}_j$. If $x \in \mathbf{V}_j$, we write $s_x = \mathbf{P}_{j+1}x$ and $d_x = \mathbf{Q}_{j+1}x$, where $s_x \in \mathbf{V}_{j+1}$ and $d_x \in \mathbf{W}_{j+1}$.

If $d = 1$, then the subspace $\mathbf{W}_j$ is spanned by an orthonormal basis $\{\psi_k^j(x) = 2^{-j/2}\psi(2^{-j}x - k)\}_{k\in\mathbf{Z}}$. The function ψ is the so-called wavelet and may be computed using the scaling function ϕ via $\psi(x/2) = \sum_k g_k\phi(x - k)$. If $d \geq 2$, then the basis in the subspace $\mathbf{W}_j$ may be constructed using products of wavelets and scaling functions. For example, if $d = 2$, then functions $\{\psi_k^j(x)\psi_{k'}^j(y),\ \phi_k^j(x)\psi_{k'}^j(y),\ \psi_k^j(x)\phi_{k'}^j(y),\}_{k,k'\in\mathbf{Z}}$ form an orthonormal basis of $\mathbf{W}_j$.

Given a bounded linear operator $\mathbf{S}$ on $\mathbf{L}^2(\mathbf{R}^d)$, let us consider its projection $\mathbf{S}_j$ on $\mathbf{V}_j$, $\mathbf{S}_j = \mathbf{P}_j\mathbf{S}\mathbf{P}_j$. Since $\mathbf{V}_j$ is a subspace spanned by translations of ϕ^j, we may represent the operator $\mathbf{S}_j$ as a (possibly infinite) matrix in that basis. With a slight abuse of notation, we will use the same symbol $\mathbf{S}_j$ to represent both the operator and its matrix. Since $\mathbf{V}_j = \mathbf{V}_{j+1} \oplus \mathbf{W}_{j+1}$, we may also write $\mathbf{S}_j\colon \mathbf{V}_j \to \mathbf{V}_j$ in a block form

$$\mathbf{S}_j = \begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{C}_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix} : \mathbf{V}_{j+1} \oplus \mathbf{W}_{j+1} \to \mathbf{V}_{j+1} \oplus \mathbf{W}_{j+1}, \tag{2.1}$$

where

$$\begin{aligned} \mathbf{A}_{\mathbf{S}_j} &= \mathbf{Q}_{j+1}\mathbf{S}_j\mathbf{Q}_{j+1}, \\ \mathbf{B}_{\mathbf{S}_j} &= \mathbf{Q}_{j+1}\mathbf{S}_j\mathbf{P}_{j+1}, \\ \mathbf{C}_{\mathbf{S}_j} &= \mathbf{P}_{j+1}\mathbf{S}_j\mathbf{Q}_{j+1}, \\ \mathbf{T}_{\mathbf{S}_j} &= \mathbf{P}_{j+1}\mathbf{S}_j\mathbf{P}_{j+1}. \end{aligned} \tag{2.2}$$

We note that $\mathbf{T}_{\mathbf{S}_j} = \mathbf{S}_{j+1}$. Each of the operators in (2.2) may be considered as a matrix. We note, however, that in the matrix form the transition from $\mathbf{S}_j$ in (2.1) to $\begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{C}_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix}$ requires application of the wavelet transform. We will use the operator notation throughout this paper and comment, if necessary, on the required numerical computations. For example, if $d = 1$ and $\mathbf{S}_j$ is finite and of size N by N, then each operator block in (2.2) is of size $N/2$ by $N/2$.

The operators (and their matrix representations) in (2.2) are referred to as the $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{T}$ *blocks* of $\mathbf{S}_j$. Also, for an operator $\mathbf{Z}$, we use the notation $\mathbf{A}_\mathbf{Z}$, $\mathbf{B}_\mathbf{Z}$, $\mathbf{C}_\mathbf{Z}$, and $\mathbf{T}_\mathbf{Z}$ to indicate its blocks.

2.2. The Reduction Procedure

The multiresolution strategy for the reduction and homogenization of linear problems has been proposed in [4]. Let us briefly review here the reduction procedure

(in its general form). Consider a bounded linear operator $\mathbf{S}_j$: $\mathbf{V}_j \to \mathbf{V}_j$ together with an equation

$$\mathbf{S}_j x = f, \tag{2.3}$$

which we may write as

$$\begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{C}_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix} \begin{pmatrix} d_x \\ s_x \end{pmatrix} = \begin{pmatrix} d_f \\ s_f \end{pmatrix}. \tag{2.4}$$

Formally eliminating d_x from (2.4) by substituting $d_x = \mathbf{A}_{\mathbf{S}_j}^{-1}(d_f - \mathbf{B}_{\mathbf{S}_j} s_x)$ gives us

$$(\mathbf{T}_{\mathbf{S}_j} - \mathbf{C}_{\mathbf{S}_j}\mathbf{A}_{\mathbf{S}_j}^{-1}\mathbf{B}_{\mathbf{S}_j}) s_x = s_f - \mathbf{C}_{\mathbf{S}_j}\mathbf{A}_{\mathbf{S}_j}^{-1} d_f. \tag{2.5}$$

We call (2.5) the *reduced equation* and the operator

$$\mathbf{R}_{\mathbf{S}_j} = \mathbf{T}_{\mathbf{S}_j} - \mathbf{C}_{\mathbf{S}_j}\mathbf{A}_{\mathbf{S}_j}^{-1}\mathbf{B}_{\mathbf{S}_j} \tag{2.6}$$

the *one-step reduction* of the operator $\mathbf{S}_j$. The right-hand side of (2.6) is also known as the Schur complement of the block matrix $\begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{C}_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix}$.

Note that the solution s_x of the reduced equation is exactly $\mathbf{P}_{j+1}x$, the projection of the solution of the original equation in $\mathbf{V}_{j+1}$. The solution of the reduced equation is the same on the subspace $\mathbf{V}_{j+1}$ as the solution of the original equation (2.3). Once we have obtained the reduced equation, it may be reduced again to produce an equation on $\mathbf{V}_{j+2}$, and the solution of this equation is the same on $\mathbf{V}_{j+2}$ as the solution of (2.3). Likewise, we may reduce n times to produce an equation on $\mathbf{V}_{j+n}$ the solution of which is the projection of the solution of (2.3) on $\mathbf{V}_{j+n}$.

We note that in the finite-dimensional case, the reduced equation (2.5) has half as many unknowns as the original equation (2.3). Reduction therefore preserves the coarse-scale behavior of solutions while reducing the number of unknowns.

2.3. Reduction and Homogenization

As described in the previous section, reduction is an algebraic procedure carried out on matrices over a finite number of scales. It relies on the explicit hierarchy of scales provided by the MRA to algebraically eliminate the fine-scale variables, leaving only the coarse-scale variables.

For partial differential equations, the term homogenization refers typically to the process of finding ''effective'' coefficients. For example, given the partial differential equation

$$-\nabla \cdot (a(\mathbf{x}) \nabla u(\mathbf{x})) = f(\mathbf{x}), \tag{2.7}$$

with coefficients $a(\mathbf{x})$ which are highly oscillatory, the goal of homogenization is to find less oscillatory (or even constant) coefficients $a^h(\mathbf{x})$ such that the solution of the

original partial differential equation (2.7) has the same average or coarse-scale behavior as the solution of the partial differential equation with coefficients given by $a^h(\mathbf{x})$.

The approach of classical homogenization is to consider the family of equations

$$-\nabla \cdot (a(\mathbf{x}/\epsilon)\nabla u^\epsilon(\mathbf{x})) = f(\mathbf{x}), \tag{2.8}$$

where the function $a(\mathbf{x})$ is periodic. Clearly, as $\epsilon \to 0$, the coefficients $a(\mathbf{x})$ become more and more oscillatory. This implies that the coefficients change on a scale that is asymptotically fine relative to the fixed coarse scale of the solution. The problem is to find an equation of the form (2.7) which has the weak limit u^0 of u^ϵ as its solution. The coefficients of this equation are taken to be the effective coefficients of the family of equations given by (2.7). Such formulations of homogenization problems are discussed in detail in, e.g., [2] and references therein.

Multiresolution homogenization as defined in [4] is, like classical homogenization, a limit process. It finds the effective coefficients of ODEs by (i) computing recurrence relations of the coefficients over one scale of reduction, (ii) finding the limit of the coefficients over infinitely many scales, and (iii) identifying an equation with smooth or constant coefficients such that reduction over infinitely many scales results in the same equation as the limit from (ii). This procedure does not assume asymptotic separation of fine and coarse scales.

In this paper we use the term homogenization to refer to a limit procedure. In classical homogenization, the fine scale is associated with a small parameter, and the limit is considered as this small parameter goes to zero. Multiresolution homogenization considers a limit over infinitely many scales. We use the term reduction to refer to an explicit transition between neighboring scales and in this paper study it over finitely many scales. We permit the coefficients to vary on intermediate scales.

The reduction procedure when applied to partial differential equations presents several interesting problems. First let us briefly describe some important points about the reduction procedure for ODEs. It is observed in [4] that, for systems of linear ordinary differential equations, using the Haar basis (and also multiwavelets with disjoint supports; see [1]) provides a technical advantage. Since the functions of the Haar basis on a fixed scale do not have overlapping supports, the recurrence relations for the coefficients and forcing terms in the equation may be written as local relations and solved explicitly. Thus for systems of ODEs, an explicit reduction *and* homogenization procedure is possible. Gilbert [11] has demonstrated the reduction and homogenization of [4] applied to the one-dimensional version of (2.7) and has established a connection to classical homogenization results (see, e.g., [2, 15]). Dorobantu [10] has also connected multiresolution homogenization with classical homogenization in the one-dimensional case.

The situation for partial differential equations is more complicated. Indeed, when the reduction procedure is applied to partial differential equations of the form, e.g., (2.7), the recurrence relations for the reduced operators do not appear to be locally solvable. Therefore, unlike the homogenization and reduction procedure outlined in [4] for ODEs, there does not seem to be an explicit local recurrence relation for the coefficients of the partial differential equation. Since such a recurrence does not appear feasible even with the Haar basis, one might as well consider the general scheme

outlined in [4], where high-order wavelets are used. We show in fact that there are compelling reasons for the use of such wavelets.

In particular, although discretizations of partial differential equations typically yield banded matrices, the matrix $\mathbf{A}_{\mathbf{S}_j}^{-1}$ in the reduction procedure appears in general to be dense, and therefore the reduced operator $\mathbf{R}_{\mathbf{S}_j}$ appears to be dense after only one step of reduction. A dense matrix is characteristic of the discretization of an integral equation, and thus (naively) it appears that the reduction procedure results in an integral equation even if the original equation is a partial differential equation. We demonstrate, however, that for a wide class of operators the non-locality introduced by the Schur complement is weak in the sense that the matrix $\mathbf{C}_{\mathbf{S}_j}\mathbf{A}_{\mathbf{S}_j}^{-1}\mathbf{B}_{\mathbf{S}_j}$ is well-approximated by a banded matrix. Our approach is novel in this regard; in, e.g., [19], either the matrix corresponding to $\mathbf{A}_{\mathbf{S}_j}^{-1}$ is approximated by a diagonal matrix or a directional preference is introduced into the discretization so that locality may be achieved. The quality of the approximation of $\mathbf{C}_{\mathbf{S}_j}\mathbf{A}_{\mathbf{S}_j}^{-1}\mathbf{B}_{\mathbf{S}_j}$ by a banded matrix depends on the rate of decay away from the diagonal of the elements of this matrix. We show that this rate of decay can be controlled by the choice of basis, in particular the number of vanishing moments of the wavelet basis. In general, it does not appear that the banded matrices (which result from reduction applied to discretizations of partial differential equations) themselves can be easily identified with a partial differential equation. However, the fact that the matrices are banded (up to finite but arbitrary precision) in the wavelet basis indicates that the class of pseudo-differential operators appears to be the correct class of operators to consider (rather than purely differential operators).

The use of high-order wavelets in the reduction procedure also has important implications for the eigenvalue problem. We show that reduction approximately preserves small eigenvalues of elliptic operators, and the accuracy of this approximation depends on the order of the wavelets.

3. MULTIRESOLUTION REDUCTION OF ELLIPTIC EQUATIONS USING HIGH-ORDER WAVELETS

The use of high-order wavelets to perform multiresolution reduction is desirable for two distinct reasons which we will explore in this section, namely, the sparsity of reduced operators and the preservation of small eigenvalues.

We show that under the reduction procedure the rate of the off-diagonal decay of the $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ blocks of the reduced operators remains the same. Also, the spectral bounds are preserved as well as (approximately) small eigenvalues and the corresponding eigenvectors. We introduce a modified reduction procedure which better approximates the small eigenvalues. The accuracy of the approximation of small eigenvalues as well as the number of eigenvalues which are preserved with a given accuracy strongly depends on the order of wavelets (and some other properties of the basis).

The approximation of small eigenvalues and their corresponding eigenvectors has some important implications for the numerical solution of hyperbolic and parabolic PDEs, and we make some observations on this topic.

We briefly consider computational issues since in (2.5) computing the matrix $\mathbf{A}_{S_j}^{-1}$ may appear to present some computational difficulty. Using an algorithm from [12], we may compute the operator $\mathbf{R}_{S_j}$ without computing $\mathbf{A}_{S_j}^{-1}$ directly.

3.1. Preservation of Ellipticity

An important observation made in [10] is that the reduction procedure preserves the lower bound in the estimate of ellipticity. The proof is very simple and we present a slightly more general result here, using some relations from [24]. This result is well-known in the field of domain-decomposition methods, where the Schur complement plays a prominent role.

THEOREM 3.1 (Preservation of spectral bounds). *Let $\mathbf{S}_j$ be a self-adjoint positive-definite operator on $\mathbf{V}_j$,*

$$m\|x\|^2 \le (\mathbf{S}_j x, x) \le M\|x\|^2, \tag{3.1}$$

for all $x \in \mathbf{V}_j$, where $0 < m \le M$.

Then

$$\mathbf{R}_{\mathbf{S}_j} = \mathbf{R}_{\mathbf{S}_j}^*, \tag{3.2}$$

and

$$m\|x\|^2 \le (\mathbf{R}_{\mathbf{S}_j} x, x) \le M\|x\|^2, \tag{3.3}$$

for all $x \in \mathbf{V}_{j+1}$.

Proof. Note that using (2.2) we can write

$$\mathbf{C}_{\mathbf{S}_j} = \mathbf{P}_{j+1}\mathbf{S}_j\mathbf{Q}_{j+1} = (\mathbf{Q}_{j+1}\mathbf{S}_j\mathbf{P}_{j+1})^* = \mathbf{B}_{\mathbf{S}_j}^*, \tag{3.4}$$

$$\mathbf{T}_{\mathbf{S}_j} = \mathbf{P}_{j+1}\mathbf{S}_j\mathbf{P}_{j+1} = (\mathbf{P}_{j+1}\mathbf{S}_j\mathbf{P}_{j+1})^* = \mathbf{T}_{\mathbf{S}_j}^*, \tag{3.5}$$

and

$$\mathbf{A}_{\mathbf{S}_j} = \mathbf{Q}_{j+1}\mathbf{S}_j\mathbf{Q}_{j+1} = (\mathbf{Q}_{j+1}\mathbf{S}_j\mathbf{Q}_{j+1})^* = \mathbf{A}_{\mathbf{S}_j}^*. \tag{3.6}$$

Therefore, we have

$$\mathbf{R}_{\mathbf{S}_j}^* = \mathbf{T}_{\mathbf{S}_j}^* - (\mathbf{B}_{\mathbf{S}_j}^*\mathbf{A}_{\mathbf{S}_j}^{-1}\mathbf{B}_{\mathbf{S}_j})^* = \mathbf{T}_{\mathbf{S}_j} - \mathbf{B}_{\mathbf{S}_j}^*\mathbf{A}_{\mathbf{S}_j}^{-1}\mathbf{B}_{\mathbf{S}_j} = \mathbf{R}_{\mathbf{S}_j}. \tag{3.7}$$

Since $\mathbf{S}_j$ is positive definite, so is $\begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{B}_{\mathbf{S}_j}^* & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix}$ and thus it follows that $\mathbf{A}_{\mathbf{S}_j}$ is positive definite and $\mathbf{A}_{\mathbf{S}_j}^{-1}$ exists. Let us consider the operator

$$\mathbf{Z} = \begin{pmatrix} \mathbf{I} & -\mathbf{A}_{\mathbf{S}_j}^{-1}\mathbf{B}_{\mathbf{S}_j} \\ 0 & \mathbf{I} \end{pmatrix}.$$

Then we have

$$\mathbf{Z}^* \begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{B}_{\mathbf{S}_j}^* & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix} \mathbf{Z} = \begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & 0 \\ 0 & \mathbf{R}_{\mathbf{S}_j} \end{pmatrix},$$

and

$$(\mathbf{R}_{\mathbf{S}_j} x, x) = \left(\mathbf{Z} * \begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{B}^*_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix} \mathbf{Z} \begin{pmatrix} 0 \\ x \end{pmatrix}, \begin{pmatrix} 0 \\ x \end{pmatrix} \right).$$

For the lower bound we obtain

$$(\mathbf{R}_{\mathbf{S}_j} x, x)$$

$$= \left(\begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{B}^*_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix} \mathbf{Z} \begin{pmatrix} 0 \\ x \end{pmatrix}, \mathbf{Z} \begin{pmatrix} 0 \\ x \end{pmatrix} \right) \geq m(\|\mathbf{A}_{\mathbf{S}_j}^{-1} \mathbf{B}_{\mathbf{S}_j} x\|^2 + \|x\|^2) \geq m\|x\|^2. \quad (3.8)$$

To estimate the upper bound, we use $\mathbf{R}_{\mathbf{S}_j} + \mathbf{B}^*_{\mathbf{S}_j} \mathbf{A}^{-1}_{\mathbf{S}_j} \mathbf{B}_{\mathbf{S}_j} = \mathbf{T}_{\mathbf{S}_j}$ and positive definiteness of $\mathbf{A}^{-1}_{\mathbf{S}_j}$ to obtain

$$(\mathbf{R}_{\mathbf{S}_j} x, x) \leq (\mathbf{T}_{\mathbf{S}_j} x, x).$$

Since $\begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{B}^*_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix}$ satisfies the same spectral bounds as $\mathbf{S}_j$, we have

$$(\mathbf{T}_{\mathbf{S}_j} x, x) \leq M\|x\|^2.$$

This completes the proof. We note that since we have made no assumptions (other than orthogonality) about the multiresolution analysis, the properties (3.3) and (3.2) do not depend on dimension or the choice of wavelet basis.

The ellipticity estimate of (3.3) raises the important question of whether it is possible (and under which conditions) to have exactly or approximately the lower eigenvalues of $\mathbf{S}_j$ as eigenvalues of $\mathbf{R}_{\mathbf{S}_j}$. We will consider these questions in Section 3.4 below.

3.2. Rate of Off-Diagonal Decay and Sparsity of Reduced Operators

In this section we show that the reduction scheme preserves the rate of the off-diagonal decay in the **A**, **B**, and **C** blocks of the reduced operator at all scales. This rate is affected by the number of vanishing moments of the wavelet function.

As was shown in [3], the elliptic operators considered in this paper (and their Green's functions) are compressible in wavelet bases. Let us represent the operators $\mathbf{A}_j$, $\mathbf{B}_j$, $\mathbf{C}_j$, $\mathbf{T}_j$ by the matrices α^j, β^j, γ^j, s^j, where

$$\alpha^j_{k,k'} = \int\int K(x, y)\psi_{j,k}(x)\psi_{j,k'}(y)\,dxdy, \quad (3.9a)$$

$$\beta^j_{k,k'} = \int\int K(x, y)\psi_{j,k}(x)\phi_{j,k'}(y)\,dxdy, \quad (3.9b)$$

$$\gamma^j_{k,k'} = \int\int K(x, y)\phi_{j,k}(x)\psi_{j,k'}(y)\,dxdy, \quad (3.9c)$$

$$s^j_{k,k'} = \int\int K(x, y)\phi_{j,k}(x)\phi_{j,k'}(y)dxdy, \tag{3.9d}$$

and $K(x, y)$ is the kernel of a Calderón–Zygmund or a pseudo-differential operator **T**. We assume that K satisfies the conditions

$$|K(x, y)| \leq \frac{1}{|x - y|}, \tag{3.10}$$

$$|\partial_x^M K(x, y)| + |\partial_y^M K(x, y)| \leq \frac{C_0}{|x - y|^{1+M}}. \tag{3.11}$$

We also assume that the kernel K defines a bounded operator on L^2 or satisfies a substantially weaker condition (the so-called ''weak cancellation condition''),

$$\left|\int_{I\times I} K(x, y)dxdy\right| \leq C|I|, \tag{3.12}$$

for all dyadic intervals I. Under these conditions we have (see [3])

THEOREM 3.2. *If the wavelet basis has M vanishing moments, then for any kernel K satisfying the conditions* (3.10), (3.11), *and* (3.12) *the matrices* α^j, β^j, γ^j *satisfy the estimate*

$$|\alpha^j_{k,l}| + |\beta^j_{k,l}| + |\gamma^j_{k,l}| \leq C^j_M(1 + |k - l|)^{-M-1}, \tag{3.13}$$

for all integers k, l.

This theorem has a straightforward higher-dimensional analogue.

Bi-infinite matrices $\{m_{kl}\}_{k,l\in\mathbf{Z}}$ which satisfy estimates of the form (3.13) fit into the more general class of matrices which decay away from the diagonal according to the estimate

$$|m_{kl}| < C(1 + |k - l|)^{-r}, \tag{3.14}$$

where $r > 1$ is a parameter and C is a constant. The following elegant theorem dealing with the algebra of invertible matrices $\{m_{kl}\}_{k,l\in\mathbf{Z}}$ has been communicated to us by Ph. Tchamitchian. This theorem is an enhancement of the result presented in [28] (following [13]).

THEOREM 3.3. *If the matrix* $\{m_{kl}\}_{k,l\in\mathbf{Z}}$ *is invertible on* l^2, *then*

$$|m^{-1}_{k,l}| < C'(1 + |k - l|)^{-r}. \tag{3.15}$$

In other words, invertible matrices $\{m_{kl}\}_{k,l\in\mathbf{Z}}$ satisfying (3.14) form an algebra.

The proof uses relations between commutators of an unbounded operator X on l^2 defined by $X(y_k) = \{ky_k\}$ and operators $M = \{m_{k,l}\}_{k,l\in\mathbf{Z}}$ and $M^{-1} = \{m^{-1}_{k,l}\}_{k,l\in\mathbf{Z}}$; it is quite elaborate and we refer to [28] for details.

A two-dimensional analogue of Theorem 3.3 has been proved by the authors. The

basic technique of the proof is the same as that of Tchamitchian in [28]. To indicate that the results of this paper are valid in higher dimensions, we state the two-dimensional version of Theorem 3.3 without proof:

THEOREM 3.4. *If a matrix* $\{m_{k,k',l,l'}\}_{k,k',l,l' \in \mathbf{Z}}$ *satisfies*

$$|m_{k,k',l,l'}| < C(1 + |k - k'| + |l - l'|)^{-2-\alpha} \tag{3.16}$$

(*where* $\alpha \in \mathbf{Z}$, $\alpha \geq 2$) *and if the matrix is invertible on* l^2, *then*

$$|m^{-1}_{k,k',l,l'}| < C''(1 + |k - k'| + |l - l'|)^{-2-\alpha}. \tag{3.17}$$

Matrices which satisfy (3.16) also form an algebra under multiplication.

We use Theorems 3.3 and 3.4 to show that at all stages of the reduction procedure in both one and two dimensions the matrices representing the **A**, **B**, and **C** blocks of the reduced operators (2.6) satisfy the same off-diagonal decay estimate (3.13) as the blocks of the non-standard form in Theorem 3.2 and its two-dimensional analogue. In other words, the reduction procedure preserves sparsity for a wide class of elliptic operators. In this sense the form (or structure) is preserved under the reduction procedure, which allows us to apply it over a finite number of scales. The following theorem applies to the one-dimensional case, but analagous results for two dimensions can be proved using Theorem 3.4.

THEOREM 3.5 (Preservation of structure over finitely many scales). *Let us assume that the operator* **S** *and the wavelet basis satisfy conditions of Theorem* 3.2 *and, in addition,* **S** *is a self-adjoint, strictly elliptic operator. Let* $\mathbf{R}_j$ *be the reduced operator on some scale j, where reduction started at some scale n,* $n \leq j$, $n, j \in \mathbf{Z}$, *and let* $\mathbf{A}_{\mathbf{R}_j}$, $\mathbf{B}_{\mathbf{R}_j}$, *and* $\mathbf{C}_{\mathbf{R}_j}$ *be its blocks. Then the bi-infinite matrices* $\alpha^{r,j}$, $\beta^{r,j}$, *and* $\gamma^{r,j}$ *representing these blocks satisfy*

$$|\alpha^{r,j}_{k,l}| + |\beta^{r,j}_{k,l}| + |\gamma^{r,j}_{k,l}| \leq C^{n,j}_M(1 + |k - l|)^{-M-1}, \tag{3.18}$$

for all integers k, l.

Proof. Our starting point is the operator $\mathbf{S}_n$ and its blocks, $\mathbf{A}_{\mathbf{S}_n}$, $\mathbf{B}_{\mathbf{S}_n}$, $\mathbf{C}_{\mathbf{S}_n}$, and $\mathbf{T}_{\mathbf{S}_n} = \mathbf{S}_{n+1}$.

Matrices representing these blocks satisfy the estimate of Theorem 3.2. Since $\mathbf{S}_n$ is positive definite, so is $\mathbf{A}_{\mathbf{S}_n}$ (see Section 3.1), and thus $\mathbf{A}^{-1}_{\mathbf{S}_n}$ exists and, according to Theorem 3.3, satisfies the estimate in (3.13). Since $\mathbf{B}_{\mathbf{S}_n}$ and $\mathbf{C}_{\mathbf{S}_n}$ satisfy the same estimate (3.13), the product $\mathbf{C}_{\mathbf{S}_n}\mathbf{A}^{-1}_{\mathbf{S}_n}\mathbf{B}_{\mathbf{S}_n}$ satisfies it as well. The reduced operator $\mathbf{R}_{n+1}$,

$$\mathbf{R}_{n+1} = \mathbf{R}_{\mathbf{S}_n} = \mathbf{T}_{\mathbf{S}_n} - \mathbf{C}_{\mathbf{S}_n}\mathbf{A}^{-1}_{\mathbf{S}_n}\mathbf{B}_{\mathbf{S}_n}, \tag{3.19}$$

consists of the difference of two terms,

$$\mathbf{R}_{n+1} = \mathbf{S}_{n+1} - \mathbf{F}_{n+1}, \tag{3.20}$$

where

$$\mathbf{F}_{n+1} = \mathbf{C}_{\mathbf{S}_n}\mathbf{A}^{-1}_{\mathbf{S}_n}\mathbf{B}_{\mathbf{S}_n}. \tag{3.21}$$

The operator $\mathbf{S}_{n+1}$ is the projection on the scale $n+1$ of the operator $\mathbf{T}$ and the operator $\mathbf{F}_{n+1}$ has fast decay and satisfies the estimate (3.13). The blocks $\mathbf{A}_{\mathbf{R}_{n+1}}$, $\mathbf{B}_{\mathbf{R}_{n+1}}$, $\mathbf{C}_{\mathbf{R}_{n+1}}$, and $\mathbf{T}_{\mathbf{R}_{n+1}}$ of the operator $\mathbf{R}_{n+1}$ may be written as a difference of the corresponding blocks of these two terms. Theorem 3.2 guarantees that the contribution from $\mathbf{S}_{n+1}$ has the proper decay. On the other hand, the contributions from $\mathbf{F}_{n+1}$ have at least the same rate of decay as $\mathbf{F}_{n+1}$ itself since the blocks are obtained by a wavelet transform.

We prove Theorem 3.5 by induction assuming that on some scale j we have

$$\mathbf{R}_j = \mathbf{S}_j - \mathbf{F}_j, \tag{3.22}$$

where $\mathbf{S}_j$ is the projection on the scale j of the kernel K and $\mathbf{F}_j$ satisfies the estimate in (3.13). The induction step is a repeat of the considerations above with the additional use of Theorem 3.1 (preservation of spectral bounds) in order to assure the invertibility of the $\mathbf{A}_{\mathbf{R}_j}$ block.

Remark 1. We require $\mathbf{S}$ to be a self-adjoint, strictly elliptic operator to assure existence of the inverses of the A-blocks of the non-standard form. Clearly, there are wider classes of operators for which these inverses exist but we do not pursue this question here.

Remark 2. There are (narrower) classes of operators for which the rate of the off-diagonal decay is faster than that in Theorems 3.2 and 3.5. For example, if we consider strictly elliptic pseudo-differential operators of order n with symbols satisfying

$$|\partial_\xi^\alpha \partial_x^\beta \sigma(x, \xi)| \le C(\alpha, \beta)|\xi|^{n-|\alpha|+|\beta|},$$

then the rate of the off-diagonal decay is faster than that in (3.13), namely (using a wavelet basis with all vanishing moments),

$$|\alpha_{k,l}^j| + |\beta_{k,l}^j| + |\gamma_{k,l}^j| \le C_m 2^{nj}(1 + |k-l|)^{-m}, \tag{3.23}$$

for all integer k, l, and m. Since matrices $\{m_{kl}\}_{k,l\in\mathbf{Z}}$ which are invertible on l^2 and which satisfy for all integer m the inequality

$$|m_{k,l}^{-1}| < \tilde{C}_m(1 + |k-l|)^{-m} \tag{3.24}$$

form an algebra (see [28]), we may repeat the above considerations to prove a version of Theorem 3.5 with the decay condition replaced by a decay condition of the form of (3.24).

Remark 3. It is clear that Theorem 3.5 may be viewed as a combination of the results of Tchamitchian [28] and Beylkin *et al.* [3]. We conjecture that (for a narrower class of operators) this theorem can be extended to reduction over an infinite number of scales, thus showing that the constant $C_M^{n,j}$ does not depend on $n-j$. Such an extension requires more precise estimates to replace (3.15) and (3.17). In particular, in both estimates the constants on the right-hand side need to be bounded more precisely; currently they are merely shown to be finite.

3.3. A Fast Method for Computing the Reduced Operator

In practical application of the reduction procedure (2.5) one of the critical issues is the cost of computing the reduced operator (2.6). The sparsity of the operators involved in the reduction is assured by Theorem 3.5. However, one still needs an algorithm for computing the reduced operator.

It turns out that a multiresolution LU decomposition algorithm may be used to obtain the reduced operator [12]. The multiresolution LU decomposition is performed with respect to the product of non-standard forms rather than the ordinary matrix product. It has complexity $O(N)$ for a fixed relative error ϵ and provides a direct solver for linear systems written using the non-standard form.

The algorithm in [12] provides an alternative to the computation of $\mathbf{A}_{\mathbf{s}_j}^{-1}$ by noting that the decomposition of

$$\begin{pmatrix} \mathbf{A}_{\mathbf{s}_j} & \mathbf{B}_{\mathbf{s}_j} \\ \mathbf{C}_{\mathbf{s}_j} & \mathbf{T}_{\mathbf{s}_j} \end{pmatrix} = \begin{pmatrix} \hat{\mathbf{A}}_{\mathbf{s}_j} & 0 \\ \hat{\mathbf{C}}_{\mathbf{s}_j} & \hat{\mathbf{T}}_{\mathbf{s}_j} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{A}}_{\mathbf{s}_j} & \tilde{\mathbf{B}}_{\mathbf{s}_j} \\ 0 & \tilde{\mathbf{T}}_{\mathbf{s}_j} \end{pmatrix} \tag{3.25}$$

implies that

$$\mathbf{R}_{\mathbf{s}_j} = \mathbf{T}_{\mathbf{s}_j} - \hat{\mathbf{C}}_{\mathbf{s}_j} \tilde{\mathbf{B}}_{\mathbf{s}_j}. \tag{3.26}$$

In the one-dimensional case, if $\mathbf{A}_{\mathbf{s}_j}$ is banded with bandwidth m, then its LU factors will also be banded with bandwidth m, and thus they may be computed in $O(Nm^2)$. If $\mathbf{B}_{\mathbf{s}_j}$ is also banded with this same bandwidth, then we may solve for $\tilde{\mathbf{B}}_{\mathbf{s}_j}$ in $O(Nm^2)$; likewise for $\hat{\mathbf{C}}_{\mathbf{s}_j}$. For fixed relative accuracy ϵ (and hence fixed bandwidth m) this leads directly to the $O(N)$ procedure for computing $\mathbf{R}_{\mathbf{s}_j}$ via the sparse incomplete block LU decomposition given by (3.26).

The two-dimensional case is more complicated. Each of the blocks on the left-hand side of (3.26) will in general exhibit a multibanded structure. Thus, one may expect the LU factors of $\mathbf{A}_{\mathbf{s}_j}$ to fill in between the bands. Indeed, this is the case, but the fill-in which occurs is observed in practice to be fill-in with rapid decay, so that truncating to a given accuracy as we compute the LU factors results in a fast method for computing the reduction (as in the one-dimensional case).

There are many details involved in the description of the multiresolution LU decomposition, and we refer to [12] for a full treatment of them. We note finally that due to this algorithm the reduction procedure requires $O(N)$ operations.

3.4. Eigenvalues and Eigenvectors of the Reduced Operators

In this section we further investigate the relations between the spectra of the operators $\mathbf{S}_j$ and $\mathbf{R}_{\mathbf{s}_j}$. In Section 3.1 we established relations between the spectral bounds of these operators and in this section we consider relations between the small eigenvalues and corresponding eigenvectors of the operators $\mathbf{S}_j$ and $\mathbf{R}_{\mathbf{s}_j}$.

We will consider self-adjoint elliptic operators with compact inverses; this class includes variable-coefficient elliptic operators. For such an operator $\mathbf{S}$, the spectrum consists of isolated eigenvalues with finite multiplicity and the only accumulating point is at infinity. The eigenvalues may be ordered,

$$0 < \lambda_0 \leq \lambda_1 \leq \lambda_2 \cdots.$$

The eigenvectors of such operators form an orthonormal basis in the Hilbert space H, and each eigenspace is a finite-dimensional subspace. Heuristically, e.g., in numerical literature, it is always assumed for elliptic operators that the eigenvectors which correspond to small eigenvalues are less oscillatory than those which correspond to large eigenvalues and the number of oscillations increases as $\lambda_n \to \infty$. For example, such statements typically form the basis for the heuristic justification of multigrid methods. There are many other examples of theorems where this property is a subject of consideration; see, e.g., [14]. Let us formulate a simple, general proposition capturing this property for the purposes of this paper.

DEFINITION. Let S be a subspace of the Hilbert space H. We will say that the subspace $\mathbf{V}_n$ of MRA is an ϵ-approximating subspace for S if any function in S may be approximated by functions from $\mathbf{V}_n$ with relative error ϵ.

Let us denote by S_l the span of eigenvectors of $\mathbf{T}$ which correspond to all eigenvalues λ_k, $\lambda_k \leq \lambda_l$. Clearly,

$$S_0 \subset S_1 \subset S_2 \subset \cdots.$$

PROPOSITION. *For any ϵ there exists a monotone sequence $k_l \geq 0$, $k_l \in \mathbf{Z}$, such that the subspaces $\mathbf{V}_{k_l}$ of the MRA,*

$$\mathbf{V}_{k_0} \subset \mathbf{V}_{k_1} \subset \mathbf{V}_{k_2} \subset \cdots,$$

are each ϵ-approximating subspaces for S_l.

The proof of this proposition is straightforward; since each S_l is finite-dimensional, we need only approximate each function in the basis of S_l by a function in some (sufficiently fine) space $\mathbf{V}_n$ to accuracy ϵ. Since there are finitely many basis functions we may choose a finest $\mathbf{V}_n$ to approximate all of them with relative error ϵ.

In [23], a similar but stronger statement is made for a narrower class of operators; in, e.g., [14] this topic is approached in terms of nodal lines of eigenfunctions.

As stated above, the proposition is quite meaningless for practical purposes. By choosing a fine enough scale, we always may use the MRA to approximate any finite-dimensional space to any accuracy ϵ. The only point of the proposition is that the MRA may be used to approximate the eigenspaces in a natural sequence, proceeding from less oscillatory to more oscillatory. For practical purposes, however, we have to construct the MRA very carefully if we want to achieve this property for the first few scales that are involved. For example, it is clear that in order to have good approximating properties, the basis functions of the MRA have to satisfy the boundary conditions. For the same reasons, in choosing an MRA for equations where the coefficients have singularities it makes good sense to incorporate appropriate singularities into the basis.

Let us now illustrate our approach by a simple example. Suppose that $\lambda > 0$ is an eigenvalue and x an eigenvector of the self-adjoint positive definite operator $\mathbf{S}_j$, $x \in \mathbf{V}_j$ and $\mathbf{Q}_{j+1}x = 0$ (in other words, $x \in \mathbf{V}_{j+1}$). Then we have

$$\mathbf{S}_j x = \lambda x, \tag{3.27}$$

and $\mathbf{Q}_{j+1}x = 0$ implies that $\mathbf{P}_{j+1}x = x$, so that

$$\mathbf{T}_{\mathbf{S}_j} x = \mathbf{P}_{j+1}\mathbf{S}_j\mathbf{P}_{j+1}x = \mathbf{P}_{j+1}\mathbf{S}_j x = \mathbf{P}_{j+1}\lambda x = \lambda x \tag{3.28}$$

and

$$\begin{aligned}\mathbf{R}_{\mathbf{S}_j} x &= \mathbf{T}_{\mathbf{S}_j} x - \mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-1}_{\mathbf{S}_j}\mathbf{B}_{\mathbf{S}_j} x \\ &= \lambda x - \mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-1}_{\mathbf{S}_j}\mathbf{Q}_{j+1}\mathbf{S}_j\mathbf{P}_{j+1}x \\ &= \lambda x - \lambda\mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-1}_{\mathbf{S}_j}\mathbf{Q}_{j+1}x \\ &= \lambda x. \end{aligned} \tag{3.29}$$

In other words, eigenvectors of $\mathbf{S}_j$ which are exactly represented on the subspace $\mathbf{V}_{j+1}$ will be preserved (with the same eigenvalue) under the reduction step.

The condition $\mathbf{Q}_{j+1}x = 0$ is certainly too stringent for a general elliptic operator. However, the ϵ-approximating property of the MRA guarantees that we can attain $\|\mathbf{Q}_{j+1}x\| < \epsilon$ if we consider the eigenvalue problem on a fine enough scale. If we accept that eigenvectors x corresponding to small eigenvalues are not very oscillatory, then the ϵ-approximating property may be achieved by a relatively coarse scale in the MRA. More precisely, we will show that if the MRA is chosen so that a set of eigenvectors may be well approximated at some coarse scale, then, up to that scale, the eigenvalues corresponding to these eigenvectors will not be significantly affected by the reduction procedure.

Given the eigenvalue problem

$$\begin{pmatrix} \mathbf{A}_{\mathbf{S}_j} & \mathbf{B}_{\mathbf{S}_j} \\ \mathbf{B}^*_{\mathbf{S}_j} & \mathbf{T}_{\mathbf{S}_j} \end{pmatrix} \begin{pmatrix} d \\ s \end{pmatrix} = \lambda \begin{pmatrix} d \\ s \end{pmatrix}, \tag{3.30}$$

we use the same approach as in deriving (2.5). Solving for d in terms of s and assuming that $(\mathbf{A}_{\mathbf{S}_j} - \lambda\mathbf{I})^{-1}$ exists, we obtain

$$(\mathbf{T}_{\mathbf{S}_j} - \mathbf{B}^*_{\mathbf{S}_j}(\mathbf{A}_{\mathbf{S}_j} - \lambda\mathbf{I})^{-1}\mathbf{B}_{\mathbf{S}_j})s = \lambda s. \tag{3.31}$$

The existence of

$$\mathbf{G}(\lambda) = (\mathbf{A}_{\mathbf{S}_j} - \lambda\mathbf{I})^{-1} \tag{3.32}$$

is assured if we consider (3.31) for λ smaller than the lower bound of $\mathbf{A}_{\mathbf{S}_j}$.

Let us now consider approximations of the left-hand side of (3.31) and the accuracy of solutions based on these approximations. We will use the following simple lemma:

LEMMA 3.1. *For a normal matrix* $\mathbf{M}$, *if*

$$\mathbf{M}x = \lambda x + \xi, \tag{3.33}$$

then there exists an eigenvalue $\lambda_{\mathbf{M}}$ *of* $\mathbf{M}$ *such that*

$$|\lambda - \lambda_{\mathbf{M}}| \leq \frac{\|\xi\|}{\|x\|}. \tag{3.34}$$

The proof of this lemma is straightforward. Let $\mathbf{G} = \mathbf{M} - \lambda\mathbf{I}$; then there is a singular value σ_0 of $\mathbf{G}$ such that

$$\sigma_0 = \inf_{\|y\| \neq 0} \frac{(\mathbf{G}^*\mathbf{G}y, y)^{1/2}}{\|y\|} \leq \frac{\|\mathbf{G}x\|}{\|x\|} = \frac{\|\xi\|}{\|x\|}. \tag{3.35}$$

Since $\mathbf{G}$ is normal, it is diagonalizable by a unitary matrix $\mathbf{Q}$. Therefore, the singular values of $\mathbf{G}$ are given by the absolute values of its eigenvalues. Since at least one singular value of $\mathbf{G}$ satisfies (3.35), the estimate (3.34) follows.

From (3.30) it is clear that

$$d = -\mathbf{G}(\lambda)\mathbf{B}_{\mathbf{S}_j}s. \tag{3.36}$$

Let us rewrite (3.31) as

$$\mathbf{T}_{\mathbf{S}_j}s = \lambda s + \mathbf{B}^*_{\mathbf{S}_j}d. \tag{3.37}$$

Using

$$\mathbf{G}(\lambda) - \mathbf{G}(0) = \lambda\mathbf{G}(\lambda)\mathbf{G}(0), \tag{3.38}$$

where $\mathbf{G}(0) = \mathbf{A}^{-1}_{\mathbf{S}_j}$, we obtain from (3.37)

$$(\mathbf{T}_{\mathbf{S}_j} - \mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-1}_{\mathbf{S}_j}\mathbf{B}_{\mathbf{S}_j})s = \lambda s + \lambda\mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-1}_{\mathbf{S}_j}d, \tag{3.39}$$

or

$$\mathbf{R}_{\mathbf{S}_j}s = \lambda s + \lambda\mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-1}_{\mathbf{S}_j}d. \tag{3.40}$$

Applying (3.38) one more time, we obtain from (3.40)

$$\mathbf{R}_{\mathbf{S}_j}s = \lambda(\mathbf{I} + \mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-2}_{\mathbf{S}_j}\mathbf{B}_{\mathbf{S}_j})s + \lambda^2\mathbf{B}^*_{\mathbf{S}_j}\mathbf{A}^{-2}_{\mathbf{S}_j}d. \tag{3.41}$$

We approximate the eigenvalue problems in (3.37), (3.40), and (3.41) by

$$\mathbf{T}_{\mathbf{S}_j}s = \lambda s, \tag{3.42}$$

$$\mathbf{R}_{\mathbf{S}_j} s = \lambda s, \tag{3.43}$$

and

$$\mathbf{R}_{\mathbf{S}_j} s = \lambda(\mathbf{I} + \mathbf{B}_{\mathbf{S}_j}^* \mathbf{A}_{\mathbf{S}_j}^{-2} \mathbf{B}_{\mathbf{S}_j}) s. \tag{3.44}$$

The last equation gives rise to what we call the modified reduction procedure. As in (3.43), we would like to iterate the modified reduction procedure over many scales for (3.44) so that this form is preserved. To this end, we factor the operator $\mathbf{I} + \mathbf{B}_{\mathbf{S}_j}^* \mathbf{A}_{\mathbf{S}_j}^{-2} \mathbf{B}_{\mathbf{S}_j}$ by using the Cholesky decomposition and obtain

$$\mathbf{I} + \mathbf{B}_{\mathbf{S}_j}^* \mathbf{A}_{\mathbf{S}_j}^{-2} \mathbf{B}_{\mathbf{S}_j} = \mathbf{L}_{\mathbf{S}_j} \mathbf{L}_{\mathbf{S}_j}^*. \tag{3.45}$$

We rewrite (3.44) as

$$\mathbf{L}_{\mathbf{S}_j}^{-1} \mathbf{R}_{\mathbf{S}_j} (\mathbf{L}_{\mathbf{S}_j}^*)^{-1} z = \lambda z, \tag{3.46}$$

where

$$z = \mathbf{L}_{\mathbf{S}_j}^* s, \tag{3.47}$$

and we define

$$\mathbf{Y}_{\mathbf{S}_j} = \mathbf{L}_{\mathbf{S}_j}^{-1} \mathbf{R}_{\mathbf{S}_j} (\mathbf{L}_{\mathbf{S}_j}^*)^{-1}. \tag{3.48}$$

The equations (3.45), (3.47), and (3.48) represent the modified reduction procedure. The operator $\mathbf{Y}_{\mathbf{S}_j}$ is self-adjoint and positive definite; to iterate the modified reduction procedure we compute the $\mathbf{A}$, $\mathbf{B}$, and $\mathbf{C}$ blocks of the operator $\mathbf{Y}_{\mathbf{S}_j}$ and obtain (3.44) on the next scale. Note that in the modified reduction procedure we have to keep track of the projections of the eigenvector since at each step they are modified via (3.47).

Let us now use Lemma 3.1 to estimate the accuracy of the approximations given by (3.42), (3.43), and (3.44). The lemma allows us to use the size of the neglected terms to bound the perturbation of the eigenvalues. For the term neglected in the approximation of (3.37) given by (3.42), we have

$$\|\mathbf{B}_{\mathbf{S}_j}^* d\|_2 \leq \|\mathbf{B}_{\mathbf{S}_j}\|_2 \|d\|_2 \tag{3.49}$$

By introducing the spectral bounds of the operator $\mathbf{A}_{\mathbf{S}_j}$

$$m_{\mathbf{A}}^j \|x\|^2 \leq (\mathbf{A}_{\mathbf{S}_j} x, x) \leq M_{\mathbf{A}}^j \|x\|^2,$$

we obtain for (3.43) the estimate

$$\|\lambda \mathbf{B}_{\mathbf{S}_j}^* \mathbf{A}_{\mathbf{S}_j}^{-1} d\|_2 \leq \frac{\lambda}{m_{\mathbf{A}}^j} \|\mathbf{B}_{\mathbf{S}_j}\|_2 \|d\|_2. \tag{3.50}$$

For the term neglected in (3.44), we follow the above considerations for the modified reduction procedure. After multiplying by $\mathbf{L}_{\mathbf{S}_j}^{-1}$ on the left and substituting (3.47) in (3.41), we have

$$\mathbf{Y}_{\mathbf{S}_j} z = \lambda z + \lambda^2 \mathbf{L}_{\mathbf{S}_j}^{-1} \mathbf{B}_{\mathbf{S}_j}^{*} \mathbf{A}_{\mathbf{S}_j}^{-2} d, \tag{3.51}$$

to which Lemma 3.1 may be applied. Let $\mathbf{Z}_{\mathbf{S}_j} = \mathbf{I} + \mathbf{B}_{\mathbf{S}_j}^{*} \mathbf{A}_{\mathbf{S}_j}^{-2} \mathbf{B}_{\mathbf{S}_j}$. The lower spectral bound of $\mathbf{Z}_{\mathbf{S}_j}$ is clearly bounded below by one (since $\mathbf{B}_{\mathbf{S}_j}^{*} \mathbf{A}_{\mathbf{S}_j}^{-2} \mathbf{B}_{\mathbf{S}_j}$ is positive-definite), and therefore $\|\mathbf{Z}_{\mathbf{S}_j}^{-1/2}\|_2 \leq 1$. Furthermore, there exists a unitary $\mathbf{Q}$ such that $\mathbf{Z}_{\mathbf{S}_j}^{1/2} = \mathbf{Q}^* \mathbf{L}_{\mathbf{S}_j}$, where $\mathbf{L}_{\mathbf{S}_j}$ is the Cholesky factor of $\mathbf{Z}_{\mathbf{S}_j}$; thus $\mathbf{L}_{\mathbf{S}_j}^{-1} = \mathbf{Z}_{\mathbf{S}_j}^{-1/2} \mathbf{Q}^*$ and

$$\|\mathbf{L}_{\mathbf{S}_j}^{-1}\|_2 \leq \|\mathbf{Z}_{\mathbf{S}_j}^{-1/2}\|_2 \leq 1. \tag{3.52}$$

This yields (from (3.51))

$$\|\lambda^2 \mathbf{L}_{\mathbf{S}_j}^{-1} \mathbf{B}_{\mathbf{S}_j}^{*} \mathbf{A}_{\mathbf{S}_j}^{-2} d\|_2 \leq \left(\frac{\lambda}{m_{\mathbf{A}}^j} \right)^2 \|\mathbf{B}_{\mathbf{S}_j}\|_2 \|d\|_2. \tag{3.53}$$

Lemma 3.1 in conjunction with (3.49), (3.50), (3.53) yields the following result:

THEOREM 3.6. *Given an eigenvector x of $\mathbf{S}_j$ such that $\mathbf{S}_j x = \lambda x$, $\|x\|_2 = 1$, $d = \mathbf{Q}_{j+1} x$, and $\|d\|_2^2 \ll \frac{1}{2}$, there exist real $\lambda_{\mathbf{T}}$, $\lambda_{\mathbf{R}}$, and $\lambda_{\mathbf{Y}}$ which solve* (3.42), (3.43), *and* (3.44), *respectively, such that*

$$|\lambda_{\mathbf{T}} - \lambda| \leq C_d \|\mathbf{B}_{\mathbf{S}_j}\|_2 \|d\|_2 \tag{3.54}$$

$$|\lambda_{\mathbf{R}} - \lambda| \leq C_d \|\mathbf{B}_{\mathbf{S}_j}\|_2 \|d\|_2 \left(\frac{\lambda}{m_{\mathbf{A}}^j} \right) \tag{3.55}$$

$$|\lambda_{\mathbf{Y}} - \lambda| \leq C_d \|\mathbf{B}_{\mathbf{S}_j}\|_2 \|d\|_2 \left(\frac{\lambda}{m_{\mathbf{A}}^j} \right)^2, \tag{3.56}$$

where $1 \leq C_d \leq \sqrt{2}$.

We may now identify two factors that affect the estimate, $\|d\|_2$ and the ratio $\lambda / m_{\mathbf{A}}^j$. In order for $\|d\|_2$ to be small, we have to assume that the eigenvector of the problem in (3.30) has a small projection on the subspace $\mathbf{W}_{j+1}$. If the subspace $\mathbf{V}_{j+1}$ is ϵ-approximating for the subspace of eigenvectors, then $\|d\|_2 \leq \epsilon$ and the perturbation of the corresponding eigenvalues is small. The eigenvalue problem (3.42) is merely the projection of the original eigenvalue problem to the next coarsest scale and reflects the current practice in setting up eigenvalue problems. The estimate (3.49) simply shows that it is safe to project the eigenvalue problem to a coarser scale as long as the eigenvectors are represented on that scale to the desired accuracy.

The reduction and modified reduction procedures improve the eigenvalue estimate with the additional factor $\lambda / m_{\mathbf{A}}^j$. This ratio is small (in a generic situation) since the

TABLE 1
Condition Numbers and Lower Bounds for the A-Block of the Operator $-\nabla\cdot(a(x, y)\nabla)$ on the Unit Square with Periodic Boundary Conditions

N	$\kappa(A_j)$	m_{A_j}	$\kappa(S_j)$
256	6.18	1.4577×10^3	1.16×10^2
1024	8.06	5.0363×10^3	6.26×10^2
2304	10.48	1.0043×10^4	1.53×10^3
4096	11.66	1.5948×10^4	2.86×10^3
5184	13.06	1.9077×10^4	3.66×10^3

Note. Here, N is the number of unknowns in the two-dimensional spatial grid. Multiwavelets with two vanishing moments are used, and the coefficients $a(x, y)$ are set to $a(x, y) = 2 + \cos(16\pi x)\cos(16\pi y)$, which provides a moderate amount of oscillation in the coefficients. The condition number depends only weakly on the scale, unlike the condition number of the original matrix (denoted as $\kappa(S_j)$), which for second-order elliptic operators scales as h^{-2} (where h is the step-size of the discretization). Note that m_{A_j} also scales as h^{-2}.

operator $\mathbf{A}_{s_j}$ is typically well-conditioned (see Table 1) and captures the ''high-frequency'' component of the operator $\mathbf{S}_j$. Thus, for the lower-frequency modes with smaller eigenvalues, we expect that $m_{\mathbf{A}}^j \gg \lambda$. We show later in numerical examples that this factor makes a significant difference.

Of great importance is the fact that all of the considerations in this section are *independent* of dimension; the guarantee of the ϵ-approximating property in arbitrary dimensions provides this. However, in higher dimensions the consideration of optimizing the MRA for a given operator becomes the chief practical difficulty.

Remark. In Section 3.3 we outlined an $O(N)$ procedure for computing the reduced operator to relative accuracy ϵ. For small eigenvalues, however, it might be necessary to maintain absolute rather than relative accuracy while performing the reduction. This puts an additional computational burden on the reduction procedure in the case of ill-conditioned operators.

In particular, if we compute $\hat{\mathbf{A}}_{s_j}$ and $\hat{\mathbf{C}}_{s_j} = \tilde{\mathbf{B}}_{s_j}^*$ to some absolute accuracy δ, and from this compute $\mathbf{R}'_{s_j}$, it is clear (from (3.26)) that

$$\|\mathbf{R}'_{s_j} - \mathbf{R}_{s_j}\| < \delta\|\hat{\mathbf{C}}_{s_j}\|. \tag{3.57}$$

In the worst case, the eigenvalues of $\mathbf{R}'_{s_j}$ will approximate the eigenvalues of $\mathbf{R}_{s_j}$ with accuracy no better than $\delta\|\hat{\mathbf{C}}_{s_j}\|$ (see, e.g., [16]). For a typical second-order elliptic operator $\mathbf{S}$, the norms of each of the blocks $\mathbf{A}_{s_j}$ and $\mathbf{C}_{s_j} = \mathbf{B}_{s_j}^*$ behave like $O(h_j^{-2})$ (where h_j is the step size of the discretization). Furthermore, in the Cholesky decomposition, the norm of the lower triangular factor is equal to the square root of the norm of the matrix. Therefore, if we compute the LU factorization defined in Section 3.3 to absolute accuracy δ, then the resulting matrix $\mathbf{R}'_{s_j}$ approximates $\mathbf{R}_{s_j}$ to absolute accuracy δh_j^{-1}, as can easily be seen from (3.57).

In other words, to compute $\mathbf{R}'_{s_j}$ so that its eigenvalues approximate the small eigenvalues of $\mathbf{R}_{s_j}$ with absolute accuracy ϵ, it is necessary to compute the multiresolution

LU decomposition with working precision $\epsilon \cdot h_j$. For a given accuracy δ, the bandwidth m of matrices which satisfy (3.13) (or its two-dimensional analogue) is given by $m \sim \delta^{-1/M}$, where M is the number of vanishing moments of the wavelet basis (see, e.g., [3] for details). Thus, as h_j decreases (and the scale becomes finer) it is necessary to keep a wider band in the LU decomposition. This thickening of the band as the scale becomes finer means that for the purposes of eigenvalue computations with fixed absolute accuracy, the reduction procedure is $O(N^{1+4/M})$ rather than $O(N)$.

This estimate is obtained if we choose the number of vanishing moments M based on the desired accuracy ϵ. A typical choice is $M \sim -\log(\epsilon)$. With this choice we have the bandwidth $m \sim (\epsilon h_j^{-2})^{1/M}$. For matrices with bandwidth m in n dimensions (where $n = 1, 2$) the multiresolution LU decomposition requires $O(Nm^{2n})$ operations. But, $h_j = N^{-1/n}$, so we see that the multiresolution LU decomposition requires $O(NN^{4n/nM}) = O(N^{1+4/M})$ operations to compute the matrix $\mathbf{R}'_{\mathbf{S}_j}$ so that its eigenvalues approximate the eigenvalues of $\mathbf{R}_{\mathbf{S}_j}$ to absolute accuracy ϵ.

3.5. Hyperbolic and Parabolic Partial Differential Equations

This section outlines further work and explains the importance of the fact that the reduction procedure preserves small eigenvalues of elliptic equations. In particular, there are implications for solving hyperbolic and parabolic initial value problems. Let us consider, for example, the differential equation

$$u_{tt}(x, t) + \mathbf{S}u(x, t) = 0, \tag{3.58}$$

where $\mathbf{S}$ is a second-order elliptic operator with variable coefficients, supplemented with some boundary conditions and the initial conditions

$$u(x, 0) = g(x), \qquad u_t(x, t)|_{t=0} = 0. \tag{3.59}$$

This equation describes (for example) wave propagation in a medium with variable velocity. Let us consider a problem where the velocity is changing very rapidly (one may think of a highly stratified rock structure) but the initial condition $g(x)$ has relatively low wavenumbers, i.e., the wavelength of the initial condition is large compared with the typical length over which the velocity changes. A space discretization of this problem would typically require a step size smaller than the smallest length over which the velocity changes, which may be prohibitively expensive in practical applications. The key point of this section is that this difficulty may be overcome by replacing $\mathbf{S}$ by the reduced operator on some scale.

Let us project $\mathbf{S}$ onto $\mathbf{V}_j$ and write, as usual, $\mathbf{S}_j = \mathbf{P}_j\mathbf{S}\mathbf{P}_j$. We assume that the space $\mathbf{V}_j$ has fine enough resolution to capture smallest features of the behavior of the coefficients of $\mathbf{S}$. Consider the eigenvalue problem

$$\mathbf{S}_j v_n^j(x) = \lambda_n^j v_n^j(x), \tag{3.60}$$

where $v_n^j(x)$ satisfies the boundary conditions. The eigenvalues $\{\lambda_n^j\}$ of $\mathbf{S}_j$ are all real and positive, and we enumerate them in ascending order. The eigenvectors $\{v_n^j(x)\}$ of $\mathbf{S}_j$ form an orthonormal basis for $\mathbf{V}_j$.

We may therefore look for solutions of the equation

$$\left(\frac{\partial}{\partial t}\right)^2 u^j(x, t) + \mathbf{S}_j u^j(x, t) = 0 \tag{3.61}$$

in the form

$$u^j(x, t) = \sum_n (a_n \cos(\sqrt{\lambda_n^j} t) + b_n \sin(\sqrt{\lambda_n^j} t)) v_n^j(x). \tag{3.62}$$

Satisfying the initial conditions

$$u^j(x, 0) = g^j(x), \qquad \frac{\partial}{\partial t} u^j(x, t)|_{t=0} = 0, \tag{3.63}$$

we obtain

$$u^j(x, t) = \sum_n a_n \cos((\lambda_n^j)^{1/2} t) v_n^j(x), \tag{3.64}$$

where the coefficients a_n are obtained from $\mathbf{P}_j g(x) = g^j(x) = \sum_n a_n v_n^j(x)$.

We observe that there is no mixing between the eigenfunctions $\{v_n^j(x)\}$ over time. In particular, if $g^j(x) = \sum_{n=0}^{K} a_n v_n^j(x)$, and $\mathbf{V}_{j+k}$ is an ϵ-approximation of the span of $\{v_n^j(x)\}_{n=0}^{n=K}$, then we may approximate solutions of (3.61) projected onto $\mathbf{V}_{j+k}$ by solutions of

$$\left(\frac{\partial}{\partial t}\right)^2 u^{j+k}(x, t) + \mathbf{R}_{j+k} u^{j+k}(x, t) = 0, \tag{3.65}$$

where $\mathbf{R}_{j+k}$ is the k-step reduction of $\mathbf{S}_j$, with the initial conditions

$$u^{j+k}(x, 0) = g^{j+k}(x), \qquad \frac{\partial}{\partial t} u^{j+k}(x, t)|_{t=0} = 0. \tag{3.66}$$

Solving (3.65) on $\mathbf{V}_{j+k}$ is less expensive than solving (3.61) on $\mathbf{V}_j$ since in a compact domain there are 2^{dk}-times as many degrees of freedom in $\mathbf{V}_j$ than in $\mathbf{V}_{j+k}$, where d is the spatial dimension.

The considerations for the hyperbolic case above also apply in the parabolic case

$$u_t(x, t) + \mathbf{S}u(x, t) = 0 \tag{3.67}$$

(with boundary and initial conditions). The situation for (3.67) is even more favorable. Namely, on $\mathbf{V}_j$ we may write the solution of (3.67) in the form

$$u^j(x, t) = \sum_n a_n \mathrm{e}^{-\lambda_n^j t} v_n^j(x), \tag{3.68}$$

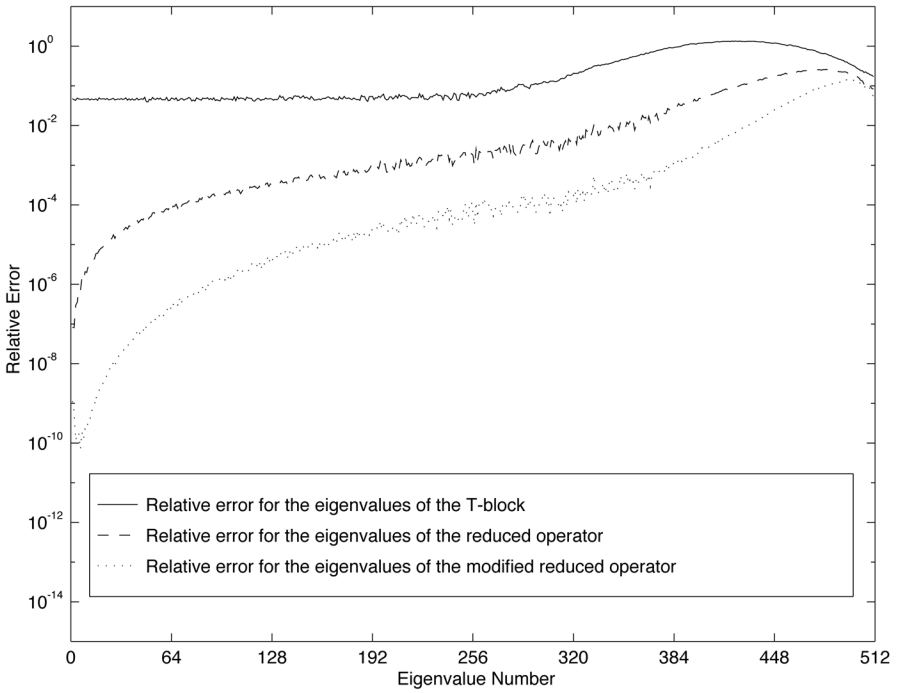

FIG. 1. Relative error of eigenvalues of the coarse scale operators obtained by different methods compared to the eigenvalues of the original operator.

where $\lambda_n^j > 0$. Similar considerations as in the hyperbolic case apply. In addition, if we are interested in the long-time solution (due to the factor $e^{-\lambda_n^j t}$), only those eigenvectors corresponding to small eigenvalues will contribute, and we may replace $\mathbf{S}_j$ by $\mathbf{R}_{j+k}$ *regardless* of which eigenvectors constitute the initial condition $g(x)$.

The efficiency of the reduction procedure in the hyperbolic and parabolic case depends on the quality of the approximations of the eigenvectors of $\mathbf{S}_j$ by functions in subspaces of the MRA. Since the eigenvectors of $\mathbf{S}_j$ satisfy the boundary conditions, it is very important to use an MRA where the scaling functions on coarse scales satisfy the same boundary conditions; further work is required in this direction.

Procedures for reduction of wave propagation models have been extensively studied (see, e.g., [5]), but most results are concerned with situations where there is a preferred direction, thus enabling the use of methods suitable for ODEs. As far as we know, no method has been proposed that addresses the problem where all directions of wave propagation are allowed.

4. NUMERICAL EXPERIMENTS

In this section we present preliminary results of numerical experiments. The goal of these experiments is to study the influence of the number of vanishing moments

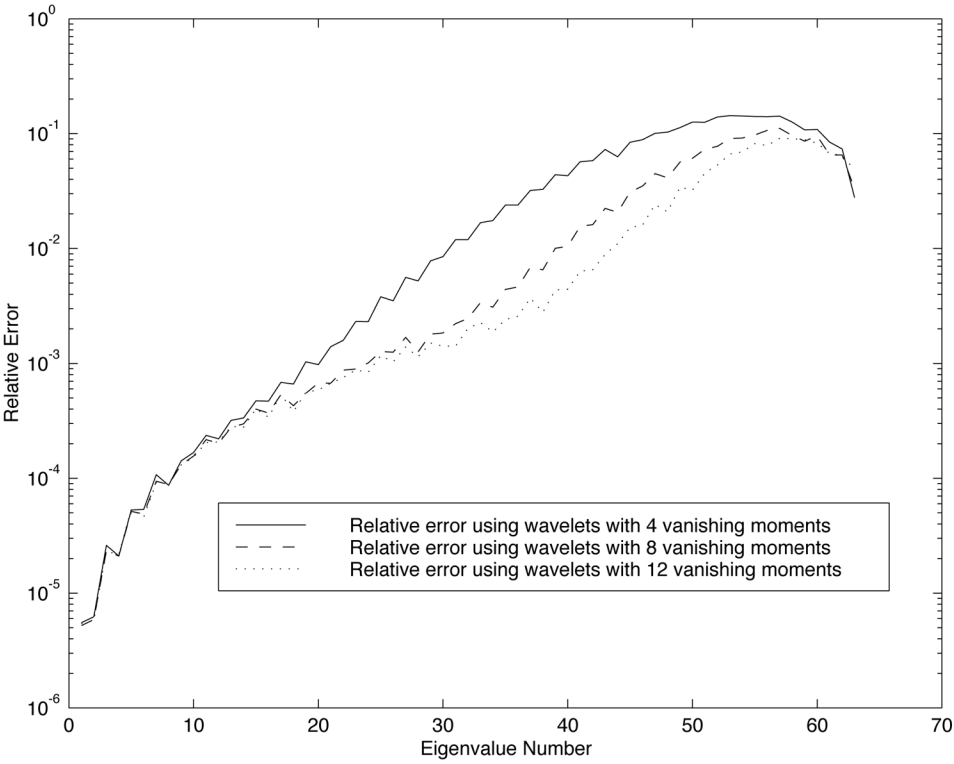

FIG. 2. Relative error of eigenvalues of the one-dimensional example operator reduced over four scales, using wavelets with 4, 8, and 12 vanishing moments.

of the wavelet bases and the effect of using different reduction procedures on the preservation of small eigenvalues.

In our first example we consider the operator

$$\mathbf{S} = \frac{d}{dx} a(x) \frac{d}{dx}$$

on $[0, 1]$ with periodic boundary conditions, and its discretization $\mathbf{S}_0 = \mathbf{D}M\mathbf{D}^T$, where $\mathbf{D}$ is a fifth-order forward-difference approximation to the first derivative, M is a diagonal matrix with uniform samples of $a(x)$ on the diagonal, and the step size $h = \frac{1}{1024}$. Although we could have computed the proper projection of this operator, we prefer to use the finite-difference discretization in our experiments since we have in mind using our method as a linear algebra tool and want to demonstrate the robustness of the method.

We examine eigenvalues of the reduced operators for an $a(x)$ which is pseudo-random (and hence highly oscillatory). The first reduction technique is to simply consider the $\mathbf{T}$ block of $\mathbf{S}_0$ on the coarse scale. The second is to use the reduced operator $\mathbf{R}_{\mathbf{S}_0}$ (2.6). Finally, we consider the modified reduced operator defined by (3.46). Figure 1 compares the performance of these three techniques after one reduction step using compactly supported wavelets with 12 vanishing moments. Experiments clearly show the advantages of using the reduced and modified reduced operators.

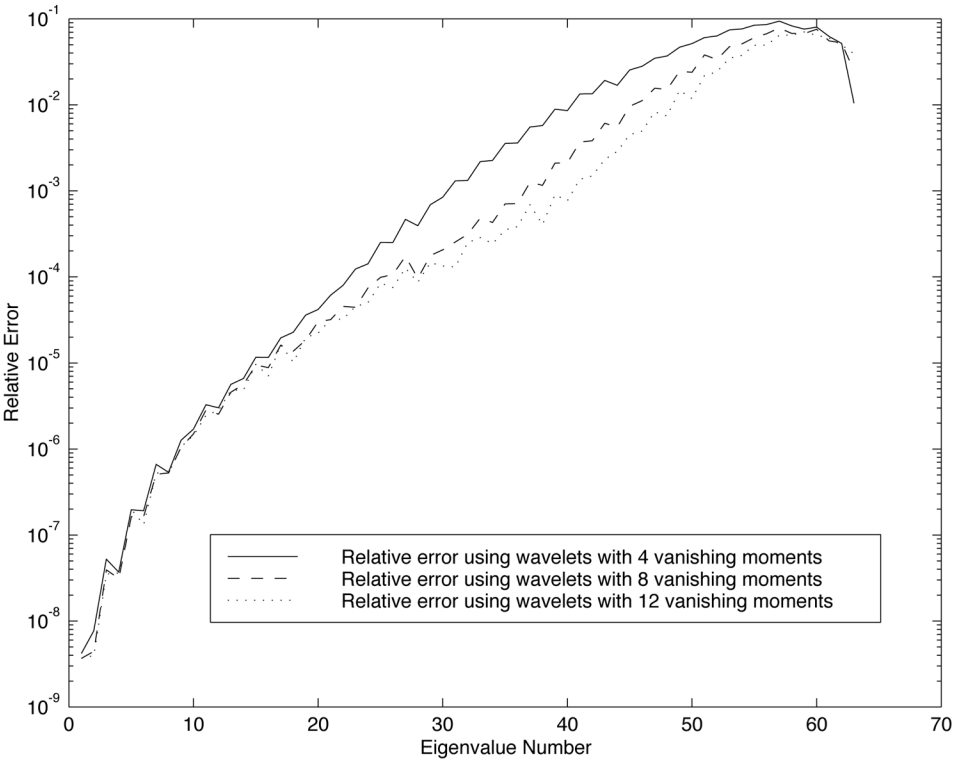

FIG. 3. Relative error of eigenvalues of the one-dimensional example operator reduced via the modified reduction procedure over four scales, using wavelets with 4, 8, and 12 vanishing moments.

In Figs. 2 and 3 we perform reduction over four scales so that the reduced matrix is of size 64×64 (the original matrix is of size 1024×1024) and compare the 64 smallest eigenvalues of the original matrix with eigenvalues of the reduced 64×64 matrix. The three curves correspond to using compactly supported wavelets with different numbers of vanishing moments. Figure 2 was obtained by using the reduced operator $\mathbf{R}_{s_j}$ after four steps of reduction. Figure 3 demonstrates the performance of the modified reduction procedure. For some regimes of the spectrum, we observe that, as expected, increasing the number of vanishing moments increases the accuracy of the approximation.

Our second example illustrates some preliminary two-dimensional results. We consider the operator $\mathbf{S} = -\nabla \cdot (a(x, y)\nabla)$ on the unit square with periodic boundary conditions; we define $a(x, y) = 2 + \cos(32\pi x)$. We discretize this operator in a multiwavelet basis (see [1]) with two vanishing moments, on an interval grid of size 32 by 32. (This results in 4096 unknowns for the fine-scale problem.) Figure 4 shows the relative error for the three techniques after one step of reduction, which reduces the number of unknowns to 1024.

5. CONCLUSIONS AND GENERALIZATIONS

In this paper we discuss only the reduction problem and defer the discussion of the connections between the classical and multiresolution approaches to homogenization

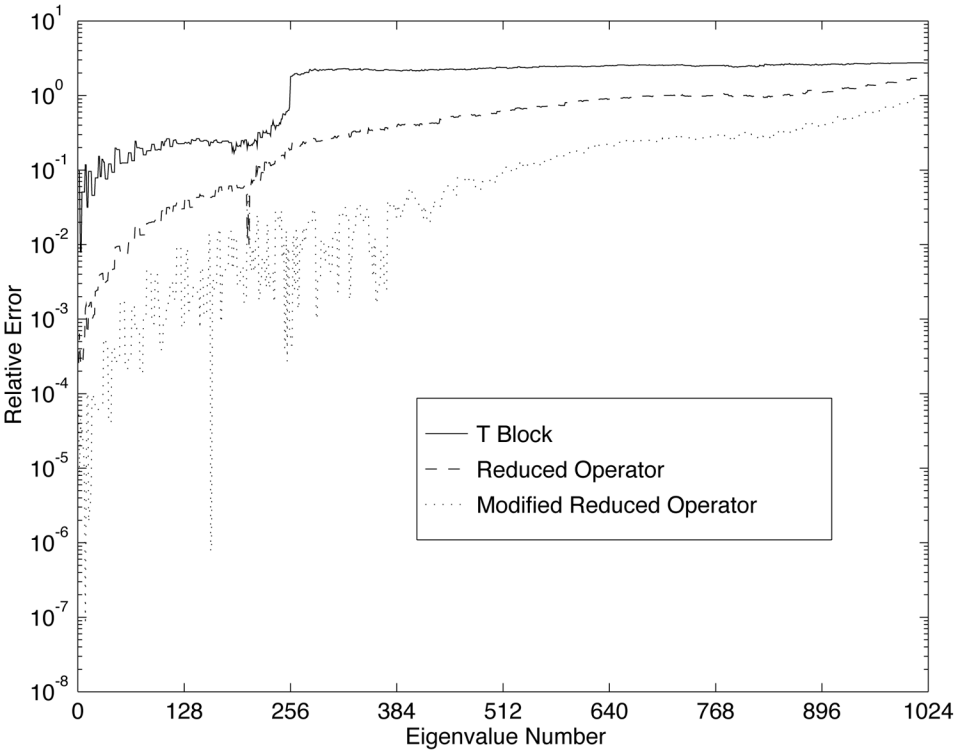

FIG. 4. Relative error of eigenvalues of the operator $-\nabla\cdot(2 + \cos(32\pi x))\nabla$ using the three techniques. Multiwavelets with two vanishing moments are used.

of elliptic PDEs to a future paper. The multiresolution strategy for homogenization introduced in [4] and developed for systems of linear ordinary differential equations is extendable as an effective approach to homogenization of linear partial differential equations and we plan to develop it in detail.

Although we do not present complete results for multiple dimensions, our theoretical results do not use any special properties of the one-dimensional problem. We point out that an effective implementation of the reduction procedure in domains with complicated boundaries does require finding practical constructions of wavelet bases in such domains. A complete treatment of the two-dimensional problem requires a separate paper due to the technical and practical issues involved.

Clearly, it is the reduction procedure that determines the practical applicability of the multiresolution approach (sparsity, fast algorithms, preservation of form), and the homogenization part serves more as an interpretation tool (although we do not want to diminish its importance). The connection established in this paper between the reduction of elliptic operators and approximate preservation of their spectra is of fundamental importance. Our results indicate that it might be possible to develop fast algorithms to find small eigenvalues.

We note that the class of operators on which the reduction procedure can be performed and for which sparsity is preserved is wider than the classes of operators described in this paper but we leave this discussion for the future as well.

Eigenvalue problems for operators of the type (1.1), where $\max a(\mathbf{x})/\min a(\mathbf{x})$ is

large, require special treatment since their eigenvectors have large derivatives in the neighborhood where the above ratio is large. Since such operators represent interesting physical phenomena in elasticity, this also presents an interesting problem for the future.

Finally, the remarks in Section 3.5 appear to open a number of opportunities for reduction and homogenization of hyperbolic and parabolic problems. These problems present a separate subject matter with many practical applications.

ACKNOWLEDGMENTS

We thank one of the anonymous referees for suggesting the simpler proof of Theorem 3.6 which is presented in this paper.

REFERENCES

1. B. Alpert, A class of bases in l^2 for the sparse representation of integral operators, *SIAM J. Math. Anal.* **24**(1) (1993), 246–262.
2. A. Bensoussan, J. L. Lions, and G. Papanicolaou, ''Asymptotic Analysis for Periodic Structures,'' North-Holland, New York, 1978.
3. G. Beylkin, R. R. Coifman, and V. Rokhlin, Fast wavelet transforms and numerical algorithms, I, *Comm. Pure Appl. Math.* **44** (1991), 141–183; Yale University Technical Report YALEU/DCS/RR-696, August 1989.
4. M. E. Brewster and G. Beylkin, A multiresolution strategy for numerical homogenization, *Appl. Comput. Harmon. Anal.* **2** (1995), 327–349; PAM Report 187, 1994.
5. R. Burridge, G. S. Papanicolaou, and B. S. White, One-dimensional wave propagation in a highly discontinuous medium, *Wave Motion* **10**(1) (1988), 19–44.
6. Z. Cai, J. Mandel, and S. McCormick, Multigrid methods for nearly singular linear equations and eigenvalues problems, *SIAM J. Numer. Anal.,* to appear.
7. R. Coifman, P. L. Lions, Y. Meyer, and S. Semmes, Compensated compactness and Hardy spaces, *J. Math. Pures Appl.* **72** (1993), 247–286.
8. W. Dahmen and L. Elsner, Algebraic multigrid methods and the Schur complement, *in* ''Robust Multigrid Methods (Kiel, 1988),'' Notes Numer. Fluid Mech., Vol. 23, pp. 58–68. Vieweg, Braunschweig, 1989.
9. I. Daubechies, ''Ten Lectures on Wavelets,'' CBMS-NSF Series in Applied Mathematics, SIAM, Philadelphia, 1992.
10. M. Dorobantu, ''Wavelet-Based Algorithms for Fast PDE Solvers,'' Dissertation, Royal Institute of Technology, Stockholm University, 1995.
11. A. Gilbert, A comparison of multiresolution and classical homogenization schemes. *Appl. Comput. Harmon Anal.* **5** (1998), 1–35.
12. D. L. Gines, G. Beylkin, and J. Dunn, LU factorization of non-standard forms and direct multiresolution solvers, *Appl. Comput. Harmon. Anal.* **5** (1998) 156–201; PAM Report, 1996.
13. S. Jaffard, Propriétés des matrices bien localisees près de leur diagonale et quelques applications. *Ann. Inst. H. Poincaré Anal. Non. Linéaire* **7**(5) (1990), 461–476.
14. D. Jerison, The diameter of the first nodal line of a convex domain. *Ann. of Math.* **141** (1995), 1–33.
15. V. V. Jikov, S. M. Kozlov, and O. A. Oleinik, ''Homogenization of Differential Operators and Integral Functionals,'' Springer-Verlag, New York, 1994.
16. T. Kato, ''Perturbation Theory for Linear Operators,'' Springer-Verlag, New York, 1996.
17. S. Kesavan, Homogenization of elliptic eigenvalue problems: Part 1, *Appl. Math. Optim.* **5**(2) (1979), 153–167.

18. S. Kesavan, Homogenization of elliptic eigenvalue problems: Part 2, *Appl. Math. Optim.* **5**(3) (1979), 197–216.
19. S. Knapek, Matrix-dependent multigrid-homogenization for diffusion problems, unpublished manuscript.
20. J. Mandel and S. McCormick, A multilevel variational method for $Au = \lambda Bu$ on composite grids, *J. Comput. Phys.* **80** (1989), 442–452.
21. S. F. McCormick, Multigrid adaptive methods for elliptic eigenproblems: A two-level convergence theory, *SINUM* **31**(6) (1994), 1731–1745.
22. F. Murat, Compacité par Compensation, *Ann. Scuola Norm. Sup. Pisa Cl. Sci.* **5**(4) (1978), 485–507.
23. A. R. Nahmod, Generalized uncertainty principles on spaces of homogeneous type, *J. Funct. Anal.* **119** (1994), 171–209.
24. V. V. Prasolov, ''Problems and Theorems in Linear Algebra,'' Translations of Mathematical Monographs, Vol. 134, American Mathematical Society, Providence, RI, 1994.
25. B. F. Smith, P. E. Bjørstad, and W. D. Gropp, ''Domain Decomposition,'' Cambridge Univ. Press, Cambridge, UK, 1996.
26. B. Z. Steinberg and J. J. McCoy, Towards effective parameter theories using multiresolution decomposition, *J. Acoust. Soc. Amer.* **96** (1994), 1130–1143.
27. L. Tartar, ''Compensated Compactness and Applications to Partial Differential Equations,'' Heriot–Watt Symposium, Vol. IV, Pitman, New York, 1979.
28. P. Tchamitchian, *in* ''Wavelets: Theory and Application'' (G. Erlebacher, M. Y. Hussaini, and L. Jameson, Eds.), pp. 83–181, ICASE/LaRC Series in Computational Science and Engineering. Oxford Univ. Press, New York, 1996.