SIAM J. SCI. COMPUT.
Vol. 30, No. 2, pp. 1082–1103

ANALYSIS OF AGGREGATION-BASED MULTIGRID*

ADRIAN C. MURESAN† AND YVAN NOTAY‡

Abstract. Aggregation-based multigrid with standard piecewise constant like prolongation is investigated. Unknowns are aggregated either by pairs or by quadruplets; in the latter case the grouping may be either linewise or boxwise. A Fourier analysis is developed for a model two-dimensional anisotropic problem. Most of the results are stated for an arbitrary smoother (which fits with the Fourier analysis framework). It turns out that the convergence factor of two-grid schemes can be bounded independently of the grid size. With a sensible choice of the (linewise or boxwise) coarsening, the bound is also uniform with respect to the anisotropy ratio, without requiring a specialized smoother. The bound is too large to guarantee optimal convergence properties with the V-cycle or the standard W-cycle, but a W-cycle scheme accelerated by the recursive use of the conjugate gradient method exhibits near grid independent convergence.

Key words. multigrid, aggregation, Fourier analysis, Krylov subspace method, conjugate gradient, preconditioning

AMS subject classifications. 65F10, 65N55, 65F50

DOI. 10.1137/060678397

1. Introduction.

We consider multigrid methods (see, e.g., [14]) for solving large sparse $n \times n$ linear systems

$$A\,\mathbf{u} \;=\; \mathbf{b} \tag{1.1}$$

arising from the discretization of second order elliptic PDEs.

We focus on schemes that use so-called coarsening by aggregation. With such schemes, the fine grid unknowns are grouped into disjoint subsets, and each such subset is associated to a unique coarse level unknown. Prolongation from coarse level to fine level is piecewise constant, that is, a vector defined on the coarse variable set is prolongated by assigning the value at a given coarse variable to all fine grid variables associated to it. Letting n_c be the number of coarse variables, the prolongation P is then a $n \times n_c$ Boolean matrix with exactly 1 nonzero entry in each row. As seen in section 2.1, the coarse grid matrices are then cheap to compute and generally as sparse as the original fine grid matrix.

Such schemes are not new and trace back to [3, 5].[1] They are, however, not very popular because it is difficult to obtain grid independent convergence on this basis [13, pp. 522–524]; see also [18, p. 663], where an accurate three grid analysis is presented for the model Poisson problem. This may be connected to the fact that the piecewise constant prolongation does not correspond to an interpolation which is at least first

*Received by the editors December 22, 2006; accepted for publication (in revised form) November 21, 2007; published electronically March 7, 2008. This work was supported by the Belgian FNRS ("Maître de recherches").

http://www.siam.org/journals/sisc/30-2/67839.html

†Romanian Academy, "T. Popoviciu" Institute of Numerical Analysis, P.O. Box 68-1, 400110 Cluj-Napoca, Romania (a_c_muresan@yahoo.com). This author's work was done while visiting the Université Libre de Bruxelles.

‡Université Libre de Bruxelles, Service de Métrologie Nucléaire (C.P. 165/84), 50, Av. F.D. Roosevelt, B-1050 Brussels, Belgium (ynotay@ulb.ac.be).

[1]The aggregation concept was introduced earlier in other fields; e.g., its use in economics dates back as far as [8].

order accurate, as required by the theory of geometric multigrid [7, sections 3.5 and 6.3.2].

In [15, 16], it is proposed to overcome this difficulty by *smoothing* the interpolation matrix; that is, from a "tentative" aggregation-based prolongation matrix P_0 (Boolean with 1 nonzero entry per row), an effective prolongation is obtained letting $P = M\,P_0$, where M is a matrix that smooths the interpolation (e.g., $M = I - \omega\,A$, where ω is a relaxation parameter). This approach is interesting, but leads to denser coarse grid matrices with the type of aggregation schemes we are interested in (where there are typically four nodes in each aggregate). Here we want to assess the potentialities of "pure" aggregation schemes. In this view, it is worth noting that the above mentioned analyses focus on the conditions to have grid independence convergence with the V-cycle, whereas multigrid methods are more robust with W-cycles, and optimal under mild conditions on the two-grid convergence rate; see [4, pp. 226–228], [14, section 3.2.1], or [11].

In this paper, we first develop the Fourier analysis [14, 17] of several aggregation-based two-grid schemes for a model anisotropic problem. It turns out that the condition number is independent of the grid size and, with a proper choice of the coarsening, also uniformly bounded with respect to the anisotropy ratio.

This condition number is typically between 2 and 5 (or, equivalently, the convergence factor is typically between 0.5 and 0.8). This is too large to hope for optimal order convergence with the V-cycle or even with the standard W-cycle. However, extra robustness can be obtained with K-cycle multigrid schemes introduced in [12]. In these schemes, the coarse-grid system is solved by $\mu \geq 1$ steps of a Krylov subspace iterative method. This approach is followed recursively until the coarsest level where an exact solve is performed. In [12], optimal convergence properties are then proved when, at each level, the condition number of the two-grid method is smaller than the number of inner iterations. In practice, the K-cycle with only 2 inner iterations at each level (that is, the K_2-cycle) is observed to be optimal even for condition numbers up to 5 and close to optimal above that limit. This is confirmed by the numerical results obtained here with aggregation-based multigrid.

The remainder of this paper is organized as follows. In section 2 we analyze the spectral properties of aggregation-based two-grid schemes. Their recursive use is discussed and numerically illustrated in section 3. Concluding remarks are given in section 4.

2. Analysis of aggregation-based two-grid methods.

2.1. General setting. To build an aggregation-based two-grid scheme, one first defines a partitioning of the n fine grid unknowns into *aggregates* (G_i), $i = 1, \ldots, n_c$, where the number n_c of aggregates is also the number of coarse variables. In our theoretical analysis, we consider several model (geometric-based) aggregations schemes. They may be mimicked by "algebraic" algorithms working with matrix entries only, as those proposed in [3, 10].

Once the aggregates are defined, the prolongation P is the $n \times n_c$ matrix given by

$$P_{ij} \;=\; \begin{cases} 1 & \text{if } i \in G_j, \\ 0 & \text{otherwise,} \end{cases} \qquad i = 1, \ldots, n;\; j = 1, \ldots, n_c. \tag{2.1}$$

With such P, it is natural to consider Galerkin coarse grid matrices

$$A_c \;=\; P^T\,A\,P,$$

which implies that A_c is cheap to construct using

$$(A_c)_{ij} \;=\; \sum_{k\in G_i}\sum_{\ell\in G_j} a_{k\ell}. \tag{2.2}$$

With a sensible choice of the aggregates, A_c is then generally as sparse as the original matrix A.

In our analysis, a slight difficulty comes from the fact that the matrix A for our model problem is singular, more precisely symmetric positive semidefinite. The kernel of A is spanned by the constant vector $\mathbf{e} = (1\cdots 1)^T$. Note that $\mathbf{e} = P\mathbf{e}_c$, where $\mathbf{e}_c = (1\cdots 1)^T$ has length n_c. Then, A_c is singular, and its kernel is spanned by $\mathbf{e}_c$. However, assuming the fine grid system compatible, coarse grid systems to solve for a two-grid scheme are of the form

$$A_c\,\mathbf{v}_c \;=\; \mathbf{r}_c,$$

where $\mathbf{r}_c = P^T\mathbf{r}$ with $\mathbf{r}\perp\mathbf{e}$, entailing $\mathbf{e}_c^T\mathbf{r}_c = \mathbf{e}_c^T P^T\mathbf{r} = \mathbf{e}^T\mathbf{r} = 0$. Thus, the coarse grid systems are compatible, and it is sensible to assume that the method computes the solution in the range of A_c; that is, $\mathbf{v}_c = A_c^+\mathbf{r}_c$, where A_c^+ is the Moore–Penrose inverse of A_c [2]. Note that if $A_c = X\,\mathrm{diag}(\lambda_i)\,X^T$ for some orthogonal matrix X, then $A_c^+ = X\,\mathrm{diag}(\lambda_i^+)\,X^T$, where

$$\lambda_i^+ \;=\; \begin{cases}\lambda_i^{-1} & \text{if } \lambda_i\neq 0,\\ 0 & \text{otherwise.}\end{cases}$$

We therefore consider two-grid schemes with the iteration matrix of the form

$$T \;=\; (I-M^{-1}A)^{\nu}\;(I-\alpha\,P\,A_c^+\,P^T\,A)\;(I-M^{-1}A)^{\nu}, \tag{2.3}$$

where M is the smoother, ν is the number of pre- and postsmoothing steps, and α is a parameter that allows us to take into account a possible scaling of the coarse grid matrices as suggested in [3, 13]. We consider symmetric smoothers satisfying $\rho(I-M^{-1}A)\le 1$, where $\rho(\cdot)$ denotes the largest eigenvalue in modulus. Our results are stated for any such smoother which is further compatible with the Fourier analysis framework, that is, which has same eigenvectors as A for the model problem under consideration.

Note that the vectors in the kernel of A are eigenvectors of T with a corresponding eigenvalue equal to 1. These modes (that is, $\mathbf{e}$ for our model problem) do, however, not affect the convergence of iterative methods when solving compatible singular systems. We are therefore interested in eigenvalues associated with other modes. In this respect, observe that if $T\mathbf{v} = \lambda\mathbf{v}$ for some $\mathbf{v}$ not in the kernel of A, then, with $\mathbf{w} = (I-M^{-1}A)^{\nu}\mathbf{v}$ and $\mathbf{z} = A^{1/2}\mathbf{w}$, one has, for $\alpha\ge 1$,

$$\begin{aligned}
\lambda &= \frac{\mathbf{v}^H A\,T\,\mathbf{v}}{\mathbf{v}^H A\,\mathbf{v}}\\
&= \frac{\mathbf{w}^H A\,\left(I-\alpha\,P\,A_c^+\,P^T\,A\right)\mathbf{w}}{\mathbf{w}^H A\,\mathbf{w}}\;\;\frac{\mathbf{w}^H A\,\mathbf{w}}{\mathbf{v}^H A\,\mathbf{v}}\\
&= \left(1-\alpha\,\frac{\mathbf{z}^H A^{1/2}\,P\,A_c^+\,P^T\,A^{1/2}\,\mathbf{z}}{\mathbf{z}^H\,\mathbf{z}}\right)\;\frac{\|A^{1/2}(I-M^{-1}A)^{\nu}\mathbf{v}\|^2}{\|A^{1/2}\mathbf{v}\|^2}\\
&\ge (1-\alpha)\;\left(\rho(I-M^{-1}A)\right)^2 && (2.4)\\
&\ge 1-\alpha, && (2.5)
\end{aligned}$$

where (2.4) holds because $A_c^+(P^T\,A\,P)A_c^+ = A_c^+$; thus $A^{1/2}\,P\,A_c^+\,P^T\,A^{1/2}$ is symmetric and idempotent, i.e., is an orthogonal projector whose norm equals 1.

When $\alpha = 1$ (no scaling of the coarse grid matrix), all eigenvalues of T are nonnegative and the convergence is characterized by

$$\lambda_{\mathbf{e}^\perp}^{(\max)}(T) \;=\; \max_{\lambda:\; T\mathbf{v}=\lambda\mathbf{v},\, A\mathbf{v}\neq 0} \lambda, \tag{2.6}$$

which is the "effective" convergence factor. Note, however, that the methods considered here perform better as a preconditioner than as a stand-alone solver. The K-cycle multigrid scheme considered in section 3 is also based on the recursive use of the two-grid method seen as a preconditioner [12]. In this context, the relevant quantity is the effective condition number

$$\kappa_{\mathbf{e}^\perp} \;=\; \frac{1-\lambda^{(\min)}(T)}{1-\lambda_{\mathbf{e}^\perp}^{(\max)}(T)}, \tag{2.7}$$

where $\lambda^{(\min)}(T) = \min_{\lambda:\; T\mathbf{v}=\lambda\mathbf{v}} \lambda$. With (2.5), one has

$$\kappa_{\mathbf{e}^\perp} \;\leq\; \frac{\alpha}{1-\lambda_{\mathbf{e}^\perp}^{(\max)}(T)}, \tag{2.8}$$

and this upper bound is expected to be accurate since (2.4) is an equality for any $\mathbf{w}$ in the range of P, whereas (2.5) is a near equality if $\mathbf{v}$ is a "smooth" mode (for which $(I-M^{-1}A)\mathbf{v} \approx \mathbf{v}$). Note in this respect that $\mathbf{e}$ is such a smooth mode in the range of P; hence, $T\mathbf{e} \approx (1-\alpha)\mathbf{e}$ for any regular problem. This mode is therefore not crucial for an analysis based on the upper bound (2.8). Hence the analysis below, in which this mode is discarded because it is in the kernel of A, has, despite this particular feature, the same relevance as Fourier analysis in general.

2.2. Model problem and notation. We consider the $N \times N$ periodic grid equipped with stencil

$$\begin{bmatrix} & -1 & \\ -\eta & 2(1+\eta) & -\eta \\ & -1 & \end{bmatrix},$$

where η is a parameter. Letting $\mathcal{A}_N = \text{tridiag}(-1, 2, -1)$, the system matrix A is $\eta\mathcal{I}_N \otimes \mathcal{A}_N + \mathcal{A}_N \otimes \mathcal{I}_N$, where $\otimes$ denotes the tensor product. For future reference we define more generally

$$A_{N,M}^{(\eta)} \;=\; \eta\,\mathcal{I}_N \otimes \mathcal{A}_M \;+\; \mathcal{A}_N \otimes \mathcal{I}_M$$

and thus $A = A_{N,N}^{(\eta)}$. The N normalized eigenvectors $\mathbf{v}_k^N$, $k = 0, \ldots, N-1$, of $\mathcal{A}_N$ are given by

$$\left(\mathbf{v}_k^N\right)_\ell = \frac{1}{\sqrt{N}}\, e^{i\ell\theta_k^N}, \qquad l = 1\ldots, N,$$

where

$$\theta_k^N \;=\; \frac{2k\pi}{N}. \tag{2.9}$$

The eigenvalue corresponding to $\mathbf{v}_k^N$ is

$$\gamma_k^N \;=\; 2\,\left(1-\cos\left(\theta_k^N\right)\right) \;=\; 2\,\left(1-\cos\left(\tfrac{2k\pi}{N}\right)\right). \tag{2.10}$$

Consequently, the NM eigenvectors of $A_{N,M}^{(\eta)}$ are

$$\mathbf{v}_{k,l}^{(N,M)} \;=\; \mathbf{v}_k^N \otimes \mathbf{v}_l^M, \qquad k,l=0,\ldots,N-1,$$

with corresponding eigenvalues

$$\lambda_{k,l}^{(N,M,\eta)} \;=\; \eta\,\gamma_l^M + \gamma_k^N \;=\; 2\left((1+\eta)-\eta\cos(\theta_l^M)-\cos(\theta_k^N)\right).$$

We write

$$\mathcal{V}_N = \begin{pmatrix}\mathbf{v}_0^N & \cdots & \mathbf{v}_{N-1}^N\end{pmatrix},$$

the matrix of eigenvectors of $\mathcal{A}_N$. One then has $\mathcal{A}_N = \mathcal{V}_N\,\Gamma_N\,\mathcal{V}_N^H$, where

$$\Gamma_N \;=\; \begin{pmatrix}\gamma_0^N & & \\ & \ddots & \\ & & \gamma_{N-1}^N\end{pmatrix}.$$

Further, the matrix of eigenvectors of $A_{N,M}^{(\eta)}$ is $\mathcal{V}_N \otimes \mathcal{V}_M$, and one has

$$A_{N,M}^{(\eta)} \;=\; (\mathcal{V}_N \otimes \mathcal{V}_M)\;(\eta\,\mathcal{I}_N \otimes \Gamma_M + \Gamma_N \otimes \mathcal{I}_M)\;(\mathcal{V}_N^H \otimes \mathcal{V}_M^H).$$

With respect to the smoother, we assume that it has the same eigenvectors as A; that is,

$$(I-M^{-1}A)^{2\nu}\,\left(\mathbf{v}_k^N \otimes \mathbf{v}_l^M\right) \;=\; \sigma_{k,l}\,\left(\mathbf{v}_k^N \otimes \mathbf{v}_l^M\right), \qquad k,l=0,\ldots,N-1. \tag{2.11}$$

Hence

$$(I-M^{-1}A)^{2\nu} = (\mathcal{V}_N \otimes \mathcal{V}_N)\;\Sigma\;(\mathcal{V}_N^H \otimes \mathcal{V}_N^H),$$

where Σ is the diagonal matrix with the $\sigma_{k,l}$ on its diagonal. Note that the symmetry of M, $\rho(I-M^{-1}A)\le 1$ and the fact that we consider even power of $I-M^{-1}A$ imply that the $\sigma_{k,l}$ are real and such that

$$0 \;\le\; \sigma_{k,l} \;\le\; 1.$$

Theorems 2.3, 2.4, and 2.5 are stated for any smoother such that the $\sigma_{k,l}$ further satisfies some technical assumptions; see (2.19), (2.20), (2.28), (2.29), (2.33), and (2.34). In fact all of these assumptions are met when the smoother better damps the rougher modes (as it is expected to do); that is,

$$|\theta_l^M-\pi| \;<\; |\theta_{l'}^M-\pi| \qquad \Rightarrow \qquad \sigma_{k,l} \;\le\; \sigma_{k,l'} \tag{2.12}$$

and

$$|\theta_k^N-\pi| \;<\; |\theta_{k'}^N-\pi| \qquad \Rightarrow \qquad \sigma_{k,l} \;\le\; \sigma_{k',l}. \tag{2.13}$$

For damped Jacobi smoothing with a relaxation factor $\omega = \frac{1}{2}$ (i.e., $M = 2\,\mathrm{diag}(A)$), one has

$$\sigma_{k,l} = \left(\frac{\eta\,\left(1+\cos\left(\theta_l^M\right)\right)+\left(1+\cos\left(\theta_k^N\right)\right)}{2\,(1+\eta)}\right)^{2\nu} = \left(\frac{\eta\,\left(1+\cos\left(\frac{2l\pi}{M}\right)\right)+\left(1+\cos\left(\frac{2k\pi}{N}\right)\right)}{2\,(1+\eta)}\right)^{2\nu}, \tag{2.14}$$

and one may check that (2.12), (2.13) indeed hold.

An important tool for the methods analyzed in the following subsections is, assuming N even, the $N \times \frac{N}{2}$ prolongation matrix

$$\mathcal{P}_N = \begin{bmatrix} 1 & 0 & & & \\ 1 & 0 & & & \\ 0 & 1 & & & \\ 0 & 1 & & & \\ & & \ddots & & \\ & & & 1 & 0 \\ & & & 1 & 0 \\ & & & 0 & 1 \\ & & & 0 & 1 \end{bmatrix}.$$

Proposition 2.1. *There holds that*

$$\mathcal{P}_N = \mathcal{V}_N\, C_N\, \mathcal{V}_{\frac{N}{2}}^H \qquad \textit{and} \qquad \mathcal{P}_N^T = \mathcal{V}_{\frac{N}{2}}\, C_N^H\, \mathcal{V}_N^H,$$

where

$$C_N = \begin{pmatrix} c_0^N & & \\ & \ddots & \\ & & c_{\frac{N}{2}-1}^N \\ c_{\frac{N}{2}}^N & & \\ & \ddots & \\ & & c_{N-1}^N \end{pmatrix} \quad \textit{with } c_k^N = \sqrt{2}\,\cos(\tfrac{k\pi}{N})\,e^{-i\frac{k\pi}{N}}, \quad k = 0,\ldots,N-1.$$

Proof. For any eigenvector $\mathbf{v}_k^N$ of $\mathcal{A}_N$ and $\ell = 1,\ldots,N-1$,

$$\begin{aligned}\left(\mathcal{P}_N^T \mathbf{v}_k^N\right)_\ell &= \tfrac{1}{\sqrt{N}}\left(e^{i(2\ell-1)\theta_k^N} + e^{i(2\ell)\theta_k^N}\right) \\ &= \begin{cases} \frac{1}{\sqrt{\frac{N}{2}}}\left(c_k^N\, e^{i\ell\theta_k^{\frac{N}{2}}}\right) = c_k^N\left(\mathbf{v}_k^{\frac{N}{2}}\right)_\ell & \text{if } k < \frac{N}{2}, \\ \frac{1}{\sqrt{\frac{N}{2}}}\left(c_k^N\, e^{i\ell\theta_{k-\frac{N}{2}}^{\frac{N}{2}}}\right) = c_k^N\left(\mathbf{v}_{k-\frac{N}{2}}^{\frac{N}{2}}\right)_\ell & \text{if } k \geq \frac{N}{2}. \end{cases}\end{aligned}$$

The required results then follow from the orthogonality of $\mathcal{V}_N$ and $\mathcal{V}_{\frac{N}{2}}$. ∎

On the other hand, it is worth noting that

$$\mathcal{P}_N^T\,\mathcal{P}_N = 2\,\mathcal{I}_{\frac{N}{2}} \tag{2.15}$$

and

$$\mathcal{P}_N^T \, \mathcal{A}_N \, \mathcal{P}_N \;=\; \mathcal{A}_{\frac{N}{2}}. \tag{2.16}$$

We conclude this subsection with a useful lemma, actually an extension of [6, Theorem 8.6.2].

LEMMA 2.2. *Let* $\mathbf{y}$ *be a vector in* $\mathcal{C}^m$*, and* $\Sigma = \mathrm{diag}(\sigma_i)$*,* $\Delta = \mathrm{diag}(\delta_i)$ *two real nonnegative diagonal matrices. Assume* $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_m$*. For any positive number* β*, the matrix*

$$\left(I - \beta \, \mathbf{y} \, \mathbf{y}^H \, \Delta\right) \, \Sigma$$

has the characteristic equation

$$\prod_{j=1}^{m} (\sigma_j - \lambda) \;-\; \beta \, \sum_{i=1}^{m} \delta_i \, \sigma_i \, |\mathbf{y}_i|^2 \left(\prod_{\substack{j=1 \\ j \neq i}}^{m} (\sigma_j - \lambda) \right) \;=\; 0,$$

and, with a proper ordering, the eigenvalues satisfy $\sigma_1 \geq \lambda_1 \geq \sigma_2 \geq \lambda_2 \geq \cdots \geq \sigma_m \geq \lambda_m$*.*

Moreover, if, in addition, $\beta^{-1} = \mathbf{y}^H \Delta \mathbf{y}$*, then the largest eigenvalue satisfies*

$$\beta \, \delta_1 \, |\mathbf{y}_1|^2 \, \sigma_m + (1 - \beta \, \delta_1 \, |\mathbf{y}_1|^2) \, \sigma_1 \;\leq\; \lambda_1 \;\leq\; \beta \, \delta_1 \, |\mathbf{y}_1|^2 \, \sigma_2 + (1 - \beta \, \delta_1 \, |\mathbf{y}_1|^2) \, \sigma_1. \tag{2.17}$$

Proof. First, consider the case where $\mathbf{y}$ has no zero components, while all σ_i, δ_i are positive and $\sigma_1 > \sigma_2 > \cdots > \sigma_m$.The considered matrix has the same eigenvalues as $\Sigma - \beta \left(\Delta^{1/2}\Sigma^{1/2}\mathbf{y}\right) \left(\Delta^{1/2}\Sigma^{1/2}\mathbf{y}\right)^H$. Then, Theorem 8.6.2 in [6] may be applied to prove the given form of the characteristic equation with $\sigma_1 > \lambda_1 > \sigma_2 > \lambda_2 > \cdots > \sigma_m > \lambda_m$.

Further, the eigenvalues are the same as those of the generalized eigenproblem

$$I - \beta \left(\Delta^{1/2}\mathbf{y}\right) \left(\Delta^{1/2}\mathbf{y}\right)^H \;=\; \lambda \, \Sigma^{-1}.$$

If $\beta^{-1} = \mathbf{y}^H \Delta \mathbf{y}$, then the left-hand side matrix is positive semidefinite, entailing that the largest eigenvalue

$$\lambda_1 \;=\; \max_{\mathbf{z} \neq 0} \; \frac{\mathbf{z}^H \, \left(I - \beta \left(\Delta^{1/2}\mathbf{y}\right) \left(\Delta^{1/2}\mathbf{y}\right)^H\right) \, \mathbf{z}}{\mathbf{z}^H \, \Sigma^{-1} \, \mathbf{z}}$$

is an increasing function of the σ_i. Hence, a lower bound on λ_1 is obtained by exchanging $\sigma_2, \ldots, \sigma_{m-1}$ for σ_m, which leads to the characteristic equation

$$\lambda \, (\sigma_m - \lambda)^{m-2} \, \left((1 - \beta \, \delta_1 \, |\mathbf{y}_1|^2) \, (\sigma_1 - \lambda) + \beta \, \delta_1 \, |\mathbf{y}_1|^2 \, (\sigma_m - \lambda) \right) \;=\; 0.$$

The given lower bound on λ_1 straightforwardly follows. The upper bound is proven in a similar way, exchanging $\sigma_3, \ldots, \sigma_m$ for σ_2.

Finally, consider the general case where some components of $\mathbf{y}$ and/or some σ_i, δ_i may be equal to zero, and/or where some σ_i may be equal to each other. One may define Σ_ε, Δ_ε, and $\mathbf{y}_\varepsilon$ depending continuously on a parameter ε and such that all requirements above are satisfied for any $\varepsilon > 0$, whereas, for $\varepsilon \to 0$, $\Sigma_\varepsilon \to \Sigma$, $\Delta_\varepsilon \to \Delta$ and $\mathbf{y}_\varepsilon \to \mathbf{y}$. The required results then follow because both the characteristic equation and the eigenvalues depend continuously on the matrix entries. (For the second part of the lemma we let $\beta_\varepsilon = (\mathbf{y}_\varepsilon^H \Delta_\varepsilon \mathbf{y}_\varepsilon)^{-1}$.) $\square$

2.3. Pairwise aggregation. This technique aggregates nodes by pairs aligned with x direction. Thus, we assume N is even and the prolongation is the $N^2 \times \frac{N^2}{2}$ matrix

$$P \;=\; \mathcal{I}_N \otimes \mathcal{P}_N. \tag{2.18}$$

With (2.15), (2.16), one sees that the coarse grid matrix is

$$P^T\, A\, P \;=\; \eta\, \mathcal{I}_N \otimes \mathcal{A}_{\frac{N}{2}} \;+\; 2\, \mathcal{A}_N \otimes \mathcal{I}_{\frac{N}{2}} \;=\; 2\, A^{(\frac{\eta}{2})}_{N,\frac{N}{2}}.$$

Hence, T has same eigenvalues as

$$\Big(I \;-\; \alpha\; (\mathcal{I}_N \otimes C_N) \Big(\eta \mathcal{I}_N \otimes \Gamma_{\frac{N}{2}} + 2\Gamma_N \otimes \mathcal{I}_{\frac{N}{2}}\Big)^+ \\ \big(\mathcal{I}_N \otimes C_N^H\big)\, (\eta \mathcal{I}_N \otimes \Gamma_N + \Gamma_N \otimes \mathcal{I}_N)\Big)\; \Sigma.$$

The latter matrix is block diagonal with (after symmetric permutation) 2×2 diagonal blocks

$$T_{kl} \;=\; \left(I \;-\; \alpha \begin{pmatrix} c_l^N \\ c_{l+\frac{N}{2}}^N \end{pmatrix} \Big(2\lambda_{k,l}^{(N,\frac{N}{2},\frac{\eta}{2})}\Big)^+ \begin{pmatrix} \overline{c}_l & \overline{c}_{l+\frac{N}{2}} \end{pmatrix} \begin{pmatrix} \lambda_{k,l}^{(N,N,\eta)} & 0 \\ 0 & \lambda_{k,l+\frac{N}{2}}^{(N,N,\eta)} \end{pmatrix}\right) \\ \begin{pmatrix} \sigma_{k,l} & 0 \\ 0 & \sigma_{k,l+\frac{N}{2}} \end{pmatrix}$$

($k = 0, \ldots, N-1$ and $l = 0, \ldots, \frac{N}{2} - 1$). We can now state the following theorem.

THEOREM 2.3. *Let $A = A^{(\eta)}_{N,N}$ with N even, let P be given by* (2.18), *and let M be a symmetric matrix satisfying* (2.11). *Assume $\alpha \geq 1$. If, for $k = 1, \ldots, N-1$ and $l = 0, \ldots, \frac{N}{2} - 1$,*

$$\begin{cases} \sigma_{k,l} \;\geq\; \sigma_{k,l+\frac{N}{2}} & \text{if } l < \frac{N}{4}, \\ \sigma_{k,l} \;\leq\; \sigma_{k,l+\frac{N}{2}} & \text{otherwise} \end{cases} \tag{2.19}$$

and if, for $l = 1, \ldots, N-1$,

$$\max_{k=0,\ldots,N-1} \sigma_{k,l} \;=\; \sigma_{0,l}, \tag{2.20}$$

then, letting for $l = 0, \ldots, \frac{N}{2} - 1$,

$$s_l = \begin{cases} \frac{\sigma_{0,l} + \sigma_{0,l+\frac{N}{2}}}{2} & \text{if } l > 0, \\ \sigma_{0,\frac{N}{2}} & \text{otherwise}, \end{cases} \tag{2.21}$$

$$g_l = \begin{cases} \sqrt{\sigma_{0,l}\, \sigma_{0,l+\frac{N}{2}}} & \text{if } l > 0, \\ \sigma_{0,\frac{N}{2}} & \text{otherwise}, \end{cases} \tag{2.22}$$

$\lambda^{(\max)}_{\mathbf{e}^\perp}(T)$ defined by (2.3), (2.6) *is given by*

$$\lambda^{(\max)}_{\mathbf{e}^\perp}(T) \;=\; \max_{l=0,\ldots,\frac{N}{2}-1} \; \big(1 - \tfrac{\alpha}{2}\big)\, s_l + \sqrt{\big(1 - \tfrac{\alpha}{2}\big)^2\, s_l^2 + (\alpha - 1)\, g_l^2}. \tag{2.23}$$

Proof. Observe first that the block T_{00} plays a particular role because $(2\lambda_{0,0}^{(N,\frac{N}{2},\frac{\eta}{2})})^+ = 0$. Hence, T_{00} has eigenvalues $\sigma_{0,0}$ and $\sigma_{0,\frac{N}{2}}$, but only the latter has to be taken into account since $\sigma_{0,0}$ is the eigenvalue (in fact equal to 1) corresponding to the singular mode $\mathbf{e} = N\,\mathbf{v}_{0,0}^{(N,N)}$. Now, one sees that $\sigma_{0,\frac{N}{2}} = s_0 = g_0$ is properly taken into account as a possible maximum in (2.23) since the expression in the right-hand side reduces to s_l when $g_l = s_l$. We now consider blocks T_{k0}, $k \geq 1$. Because $c_{\frac{N}{2}}^N = 0$, one has

$$T_{k0} \;=\; \begin{pmatrix} 0 & 0 \\ 0 & \sigma_{k,\frac{N}{2}} \end{pmatrix}, \qquad k = 1, \ldots, N-1.$$

Since by (2.20) $\sigma_{k,\frac{N}{2}} \leq \sigma_{0,\frac{N}{2}}$, and since $\sigma_{0,\frac{N}{2}}$ has already been registered as a possible maximum, none of these blocks can lead to the maximal eigenvalue we are seeking.

We are thus left with the analysis of blocks T_{kl} with $l \geq 1$. Let

$$\zeta_{k,l} = \frac{|c_l^N|^2\,\lambda_{k,l}^{(N,N,\eta)}}{2\,\lambda_{k,l}^{(N,\frac{N}{2},\frac{\eta}{2})}}.$$

One may check that

$$|c_{l+\frac{N}{2}}^N|^2\,\lambda_{k,l+\frac{N}{2}}^{(N,N,\eta)}\,\left(2\,\lambda_{k,l}^{(N,\frac{N}{2},\frac{\eta}{2})}\right)^{-1} = 1 - \zeta_{k,l}$$

either using the explicit formulas or observing that this relation is a consequence of $A_c = P^T\,A\,P$, which entails

$$2\,\lambda_{k,l}^{(N,\frac{N}{2},\frac{\eta}{2})} \;=\; \begin{pmatrix} \bar{c}_l^N & \bar{c}_{l+\frac{N}{2}}^N \end{pmatrix} \begin{pmatrix} \lambda_{k,l}^{(N,N,\eta)} & 0 \\ 0 & \lambda_{k,l+\frac{N}{2}}^{(N,N,\eta)} \end{pmatrix} \begin{pmatrix} c_l^N \\ c_{l+\frac{N}{2}}^N \end{pmatrix}.$$

We may then apply Lemma 2.2, showing that the characteristic equation may be written as

$$(\sigma_{k,l} - \lambda)(\sigma_{k,l+\frac{N}{2}} - \lambda) - \alpha\left(\zeta_{k,l}\,\sigma_{k,l}\,(\sigma_{k,l+\frac{N}{2}} - \lambda) + (1-\zeta_{k,l})\,\sigma_{k,l+\frac{N}{2}}\,(\sigma_{k,l} - \lambda)\right) \;=\; 0,$$

which amounts to

$$\lambda^2 \;-\; \left(\alpha\big((1-\zeta_{k,l})\sigma_{k,l} + \zeta_{k,l}\,\sigma_{k,l+\frac{N}{2}}\big) + 2\,(1-\alpha)\,s_{k,l}\right)\lambda \;-\; (\alpha-1)\,g_{k,l}^2 = 0,$$

where $s_{k,l} = \frac{\sigma_{k,l}+\sigma_{k,l+\frac{N}{2}}}{2}$ and $g_{k,l}^2 = \sigma_{k,l}\,\sigma_{k,l+\frac{N}{2}}$. Using

$$\gamma_l^{\frac{N}{2}} \;=\; 2\,\left(1 - \cos\left(\tfrac{4l\pi}{N}\right)\right) \;=\; 4\,\left(1 - \cos^2\left(\tfrac{2l\pi}{N}\right)\right) \;=\; \gamma_l^N\,\gamma_{l+\frac{N}{2}}^N,$$

one sees that $\lambda_{k,l}^{(N,\frac{N}{2},\frac{\eta}{2})} = \frac{\eta}{2}\gamma_l^N\gamma_{l+\frac{N}{2}}^N + \gamma_k^N$. Because $|c_l^N|^2 = \frac{1}{2}\gamma_{l+\frac{N}{2}}^N$, one then finds

$$\zeta_{k,l} \;=\; \frac{1}{2} + \gamma_k^N\,\frac{\frac{1}{2}\gamma_{l+\frac{N}{2}}^N \;-\; 1}{\eta\gamma_l^N\gamma_{l+\frac{N}{2}}^N + 2\gamma_k^N} \;=\; \frac{1}{2} + \frac{\gamma_k^N\,\cos\left(\frac{2l\pi}{N}\right)}{\eta\gamma_l^N\gamma_{l+\frac{N}{2}}^N + 2\gamma_k^N}.$$

Since $\gamma_k^N \geq 0$, one sees that $\zeta_{k,l} \geq \frac{1}{2}$ if $l < \frac{N}{4}$ and $\zeta_{k,l} \leq \frac{1}{2}$ otherwise (with $\zeta_{0,l} = \frac{1}{2}$ for all l). Hence, with (2.19), we have in all cases

$$(1-\zeta_{k,l})\,\sigma_{k,l} \;+\; \zeta_{k,l}\,\sigma_{k,l+\frac{N}{2}} \;\leq\; s_{k,l}. \tag{2.24}$$

Now, the positive root of $\lambda^2 - b\,\lambda - c = 0$ with $c \geq 0$ is an increasing function of b. Thus, the positive root of

$$\lambda^2 - 2\left(1 - \tfrac{\alpha}{2}\right) s_{k,l}\lambda - (\alpha - 1)\, g_{k,l}^2 \;= 0 \tag{2.25}$$

is an upper bound on the largest eigenvalue of T_{kl} for $l \geq 1$. With the previous discussion of blocks T_{k0} and (2.20), this shows that the right-hand side of (2.23) is an upper bound on $\lambda_{\mathbf{e}^\perp}^{(\max)}(T)$.

On the other hand, a lower bound is obtained by computing the largest eigenvalue of a subset of the blocks. Consider then T_{0l} for $l = 1, \ldots, \frac{N}{2} - 1$. Since $\zeta_{0,l} = \frac{1}{2}$, (2.24) is then an equality, showing that the positive root of (2.25) gives exactly the largest eigenvalue of T_{kl} for $k = 0$. The conclusion readily follows. ∎

By way of illustration, we apply this result to a damped Jacobi smoothing ($\sigma_{k,l}$ given by (2.14)). One has, for all l,

$$s_l \;\leq\; \lim_{N\to\infty} s_1 \;=\; \frac{1 + \left(\frac{1}{1+\eta}\right)^{2\nu}}{2}$$

and

$$g_l \;\leq\; g_{\frac{N}{4}} \;=\; \left(\frac{1 + \frac{\eta}{2}}{1+\eta}\right)^{2\nu}.$$

For $\alpha = 1$, $\lambda_{\mathbf{e}^\perp}^{(\max)}(T) = \max_l s_l$ and (2.8) gives

$$\kappa_{\mathbf{e}^\perp} \;\leq\; \frac{2}{1 - \left(\frac{1}{1+\eta}\right)^{2\nu}};$$

whereas, for $\alpha = 2$, $\lambda_{\mathbf{e}^\perp}^{(\max)}(T) = \max_l g_l$, whence

$$\kappa_{\mathbf{e}^\perp} \;\leq\; \frac{2}{1 - \left(\frac{1+\frac{\eta}{2}}{1+\eta}\right)^{2\nu}}.$$

Numerical computation reveals that the upper bound (2.8) is very tight. Hence, the above expressions also give the exact asymptotic value for $N \to \infty$. Comparing them, one sees that the condition number is always better for $\alpha = 1$, although the difference decreases as ν increases, that is, as the smoother improves. This is illustrated in Figure 1, where we have plotted the condition number as a function of α. One sees that $\alpha = 1$ is optimal for all tested values of η and ν. Hence, scaling the coarse grid matrix does not improve the convergence of the two-grid method.

Our results also show that the condition number decreases as η increases and is in fact very bad for $\eta < 1$. In the latter case, it is be better to coarsen along the y direction.

2.4. Linewise quadruplet aggregation. A faster coarsening can be obtained by repeating the process, defining aggregates by forming pairs of pairs—more precisely, by forming pairs of aggregates from the first pass in x direction. When these pairs of aggregates also follow the x direction, we will then call this *linewise quadruplet aggregation*. *Boxwise quadruplet aggregation* is considered in the next subsection.

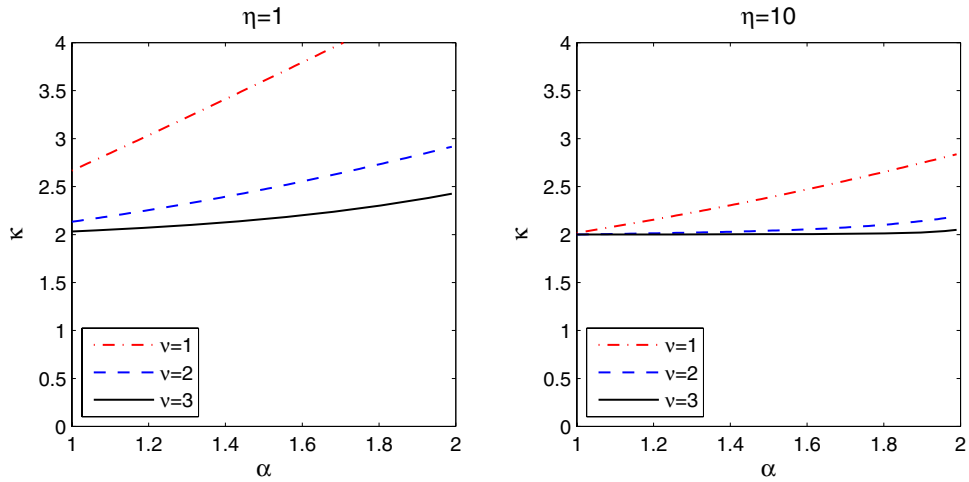

FIG. 1. *Asymptotic ($N \to \infty$) condition number (estimated by* (2.8)*) for pairwise aggregation with damped Jacobi smoothing, as a function of α.*

From Figure 1, one sees that scaling the coarse grid matrix does not improve the condition number in the case of pairwise aggregation. Numerical investigations reveal that this conclusion carries over linewise and boxwise quadruplet aggregation. Hence, here and in the following subsection, we restrict ourselves to the case $\alpha = 1$.

We assume that N is an integer multiple of 4, and the prolongation is the $N^2 \times \frac{N^2}{4}$ matrix

$$P \;=\; \mathcal{I}_N \otimes \mathcal{P}_N \, \mathcal{P}_{\frac{N}{2}}, \tag{2.26}$$

where

$$\mathcal{P}_N \, \mathcal{P}_{\frac{N}{2}} \;=\; \begin{bmatrix} 1 & 0 & & & \\ 1 & 0 & & & \\ 1 & 0 & & & \\ 1 & 0 & & & \\ & & \ddots & & \\ & & & 0 & 1 \\ & & & 0 & 1 \\ & & & 0 & 1 \\ & & & 0 & 1 \end{bmatrix}.$$

From Proposition 2.1, one has

$$\mathcal{P}_N \, \mathcal{P}_{\frac{N}{2}} \;=\; \mathcal{V}_N \; C_N \; C_{\frac{N}{2}} \; \mathcal{V}^H_{\frac{N}{2}} \qquad \text{and} \qquad \mathcal{P}^T_N \;=\; \mathcal{V}_{\frac{N}{2}} \; C^H_{\frac{N}{2}} \; C^H_N \; \mathcal{V}^H_N.$$

On the other hand, with (2.15), (2.16), one sees that the coarse grid matrix is

$$P^T \, A \, P \;=\; \eta \, \mathcal{I}_N \otimes \mathcal{A}_{\frac{N}{4}} \;+\; 4 \, \mathcal{A}_N \otimes \mathcal{I}_{\frac{N}{4}} \;=\; 4 \, A^{(\frac{\eta}{4})}_{N, \frac{N}{4}}. \tag{2.27}$$

Hence, T (with $\alpha = 1$) has the same eigenvalues as

$$\Big(I \;-\; \Big(\mathcal{I}_N \otimes C_N\, C_{\frac{N}{2}}\Big)\Big(\eta\mathcal{I}_N \otimes \Gamma_{\frac{N}{4}} + 4\Gamma_N \otimes \mathcal{I}_{\frac{N}{4}}\Big)^{+} \Big(\mathcal{I}_N \otimes C_{\frac{N}{2}}^H\, C_N^H\Big)\left(\eta\mathcal{I}_N \otimes \Gamma_N + \Gamma_N \otimes \mathcal{I}_N\right)\Big)\,\Sigma.$$

The latter matrix is block diagonal with (after symmetric permutation) 4×4 diagonal blocks

$$T_{kl} \;=\; \left(I - \mathbf{p}_l\,\Big(4\lambda_{k,l}^{(N,\frac{N}{4},\frac{\eta}{4})}\Big)^{+}\,\mathbf{p}_l^H \begin{pmatrix} \lambda_{k,l}^{(N,N,\eta)} & & & \\ & \lambda_{k,l+\frac{N}{4}}^{(N,N,\eta)} & & \\ & & \lambda_{k,l+\frac{N}{2}}^{(N,N,\eta)} & \\ & & & \lambda_{k,l+\frac{3N}{4}}^{(N,N,\eta)} \end{pmatrix}\right) \begin{pmatrix} \sigma_{k,l} & & & \\ & \sigma_{k,l+\frac{N}{4}} & & \\ & & \sigma_{k,l+\frac{N}{2}} & \\ & & & \sigma_{k,l+\frac{3N}{4}} \end{pmatrix},$$

where

$$\mathbf{p}_l \;=\; \begin{pmatrix} p_l^N \\ p_{l+\frac{N}{4}}^N \\ p_{l+\frac{N}{2}}^N \\ p_{l+\frac{3N}{4}}^N \end{pmatrix} \;=\; \begin{pmatrix} c_l^N & c_l^{\frac{N}{2}} \\ c_{l+\frac{N}{4}}^N & c_{l+\frac{N}{4}}^{\frac{N}{2}} \\ c_{l+\frac{N}{2}}^N & c_l^{\frac{N}{2}} \\ c_{l+\frac{3N}{4}}^N & c_{l+\frac{N}{4}}^{\frac{N}{2}} \end{pmatrix}$$

with c_i^N as in Proposition 2.1 ($k = 0, \ldots, N-1$ and $l = 0, \ldots, \frac{N}{4} - 1$).

THEOREM 2.4. *Let $A = A_{N,N}^{(\eta)}$ with an N integer multiple of 4, let P be given by* (2.26)*, and let M be a symmetric matrix satisfying* (2.11)*. If, for $k = 1, \ldots, N-1$ and $l = 0, \ldots, \frac{N}{4} - 1$,*

$$\begin{cases} \sigma_{k,l} \;\geq\; \max\Big(\sigma_{k,l+\frac{N}{4}},\, \sigma_{k,l+\frac{N}{2}},\, \sigma_{k,l+\frac{3N}{4}}\Big) & \text{if } l < \frac{N}{8}, \\ \sigma_{k,l+\frac{3N}{4}} \;\geq\; \max\Big(\sigma_{k,l},\, \sigma_{k,l+\frac{N}{4}},\, \sigma_{k,l+\frac{N}{2}}\Big) & \text{if } l > \frac{N}{8}, \\ \max\Big(\sigma_{k,l},\, \sigma_{k,l+\frac{3N}{4}}\Big) \;\geq\; \max\Big(\sigma_{k,l+\frac{N}{4}},\, \sigma_{k,l+\frac{N}{2}}\Big) & \text{if } l = \frac{N}{8}, \end{cases} \tag{2.28}$$

and if, for $l = 1, \ldots, N-1$,

$$\max_{k=0,\ldots,N-1} \sigma_{k,l} \;=\; \sigma_{0,l}, \tag{2.29}$$

then, for $l = 0, \ldots, \frac{N}{4} - 1$, letting $\tau_l^{(1)}$, $\tau_l^{(2)}$,$\tau_l^{(3)}$, $\tau_l^{(4)}$ be such that

$$\text{for } l > 0: \begin{cases} \Big\{\tau_l^{(1)},\, \tau_l^{(2)},\, \tau_l^{(3)},\, \tau_l^{(4)}\Big\} \;=\; \Big\{\sigma_{0,l},\, \sigma_{0,l+\frac{N}{4}},\, \sigma_{0,l+\frac{N}{2}},\, \sigma_{0,l+\frac{3N}{4}}\Big\}, \\ \tau_l^{(1)} \;\geq\; \tau_l^{(2)} \;\geq\; \tau_l^{(3)} \;\geq\; \tau_l^{(4)}, \end{cases}$$

$$\tau_0^{(1)} = \tau_0^{(2)} = \tau_0^{(3)} = \tau_0^{(4)} = \max\Big(\sigma_{0,\frac{N}{4}},\, \sigma_{0,\frac{N}{2}},\, \sigma_{0,\frac{3N}{4}}\Big),$$

$\lambda_{\mathbf{e}^\perp}^{(\max)}(T)$ *defined by* (2.3) *with* $\alpha = 1$ *and* (2.6) *satisfies*

$$\max_{l=0,\dots,\frac{N}{4}-1} \left(\tfrac{3}{4}\,\tau_l^{(1)} + \tfrac{1}{4}\,\tau_l^{(4)}\right) \;\le\; \lambda_{\mathbf{e}^\perp}^{(\max)}(T) \;\le\; \max_{l=0,\dots,\frac{N}{4}-1} \left(\tfrac{3}{4}\,\tau_l^{(1)} + \tfrac{1}{4}\,\tau_l^{(2)}\right). \tag{2.30}$$

Proof. We first consider the block T_{00}. Since $(4\lambda_{0,0}^{(N,\frac{N}{4},\frac{\eta}{4})})^+ = 0$, T_{00} has eigenvalues $\sigma_{0,0}$, $\sigma_{0,\frac{N}{4}}$, $\sigma_{0,\frac{N}{2}}$, and $\sigma_{0,\frac{3N}{4}}$, but only the last three have to be taken into account, which is properly done with the $\tau_0^{(i)}$. ($\sigma_{0,0}$ is the eigenvalue (in fact equal to 1) corresponding to the singular mode $\mathbf{e} = N\,\mathbf{v}_{0,0}^{(N,N)}$.) We next consider the blocks T_{k0}, $k \ge 1$. Because $p_{\frac{N}{4}}^N = p_{\frac{N}{2}}^N = p_{\frac{3N}{4}}^N = 0$, T_{k0} is diagonal with diagonal entries 0, $\sigma_{k,\frac{N}{4}}$, $\sigma_{k,\frac{N}{2}}$, and $\sigma_{k,\frac{3N}{4}}$. By assumption (2.29), these eigenvalues never exceed those already taken into account previously for the block T_{00}.

We are thus left with the analysis of blocks T_{kl} with $l \ge 1$. Observe that Lemma 2.2 may be applied and that (2.17) holds with $m = 4$ and

$$\beta\,\delta_i\,|\mathbf{y}_i|^2 \;=\; \frac{|p_{l_i}^N|^2\,\lambda_{k,l_i}^{(N,N,\eta)}}{4\,\lambda_{k,l}^{(N,\frac{N}{4},\frac{\eta}{4})}},$$

where $\{l_1, l_2, l_3, l_4\} = \{l,\, l+\frac{N}{4},\, l+\frac{N}{2},\, l+\frac{3N}{4}\}$. Defining γ_k^N by (2.10) (but without assuming $k < N$), one sees that for any such l_i,

$$4\,\lambda_{k,l}^{(N,\frac{N}{4},\frac{\eta}{4})} \;=\; \eta\,\gamma_{l_i}^{\frac{N}{4}} \;+\; 4\,\gamma_k^N.$$

Since $|p_{l_i}^N|^2 = 4\,\cos^2\left(\frac{l_i\pi}{N}\right)\,\cos^2\left(\frac{2l_i\pi}{N}\right) = \frac{1}{4}\,\gamma_{l_i+\frac{N}{2}}^N\,\gamma_{l_i+\frac{N}{4}}^{\frac{N}{2}}$, one then finds

$$\begin{aligned}
\frac{|p_{l_i}^N|^2\,\lambda_{k,l_i}^{(N,N,\eta)}}{4\,\lambda_{k,l}^{(N,\frac{N}{4},\frac{\eta}{4})}} &= \frac{\frac{1}{2}\,\gamma_{l_i+\frac{N}{2}}^N\left(\eta\,\gamma_{l_i}^N + \gamma_k^N\right)}{\eta\,\gamma_{l_i}^{\frac{N}{2}} \;+\; 2\,\gamma_k^N}\;\;\frac{\frac{1}{2}\,\gamma_{l_i+\frac{N}{4}}^{\frac{N}{2}}\left(\eta\,\gamma_{l_i}^{\frac{N}{2}} \;+\; 2\,\gamma_k^N\right)}{\eta\,\gamma_{l_i}^{\frac{N}{4}} \;+\; 4\,\gamma_k^N} \\
&= \frac{\eta\,\gamma_{l_i+\frac{N}{2}}^N\,\gamma_{l_i}^N \;+\; \gamma_{l_i+\frac{N}{2}}^N\,\gamma_k^N}{2\left(\eta\,\gamma_{l_i}^N\,\gamma_{l_i+\frac{N}{2}}^N \;+\; 2\,\gamma_k^N\right)}\;\;\frac{\eta\,\gamma_{l_i+\frac{N}{4}}^{\frac{N}{2}}\,\gamma_{l_i}^{\frac{N}{2}} \;+\; 2\,\gamma_{l_i+\frac{N}{4}}^{\frac{N}{2}}\,\gamma_k^N}{2\left(\eta\,\gamma_{l_i}^{\frac{N}{2}}\,\gamma_{l_i+\frac{N}{4}}^{\frac{N}{2}} \;+\; 4\,\gamma_k^N\right)} \\
&= \left(\frac{1}{2} \;+\; \frac{\gamma_k^N\,\cos\left(\frac{2l_i\pi}{N}\right)}{\eta\,\gamma_{l_i}^N\,\gamma_{l_i+\frac{N}{2}}^N \;+\; 2\,\gamma_k^N}\right)\left(\frac{1}{2} \;+\; \frac{2\,\gamma_k^N\,\cos\left(\frac{4l_i\pi}{N}\right)}{\eta\,\gamma_{l_i}^{\frac{N}{2}}\,\gamma_{l_i+\frac{N}{4}}^{\frac{N}{2}} \;+\; 4\,\gamma_k^N}\right).
\end{aligned}$$

To apply Lemma 2.2, the l_i are to be ordered so that the σ_{k,l_i} are in nonincreasing order. With assumption (2.28), it means that $l_1 = l$ if $l < \frac{N}{8}$, $l_1 = l + \frac{3N}{4}$ if $l > \frac{N}{8}$ and one of these two if $l = \frac{N}{8}$. In all cases, one checks with the above expression that

$$\beta\,\delta_1\,|\mathbf{y}_1|^2 = \frac{|p_{l_1}^N|^2\,\lambda_{k,l_1}^{(N,N,\eta)}}{4\,\lambda_{k,l}^{(N,\frac{N}{4},\frac{\eta}{4})}} \ge \frac{1}{4},$$

which, together with (2.17) and (2.29), yields the required upper bound.

To prove the lower bound, we restrict our attention to the blocks $T_{0,l}$. The block $T_{0,0}$ is already discussed above. On the other hand, for $l \ge 1$, we apply Lemma 2.2 as

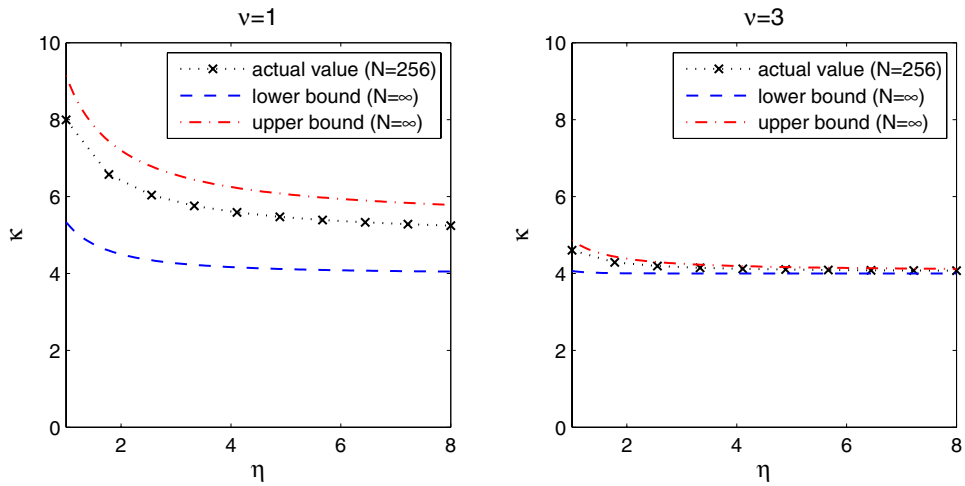

FIG. 2. *Actual condition number and its upper and lower bounds for linewise quadruplet aggregation with damped Jacobi smoothing, as a function of* η *(*$\alpha = 1$*).*

above. But since $\gamma_0^N = 0$, we now have exactly $\beta\,\delta_i\,|\mathbf{y}_i|^2 = \frac{1}{4}$ for all i. The required lower bound straightforwardly follows. $\square$

For damped Jacobi smoothing ($\sigma_{k,l}$ given by (2.14)), the maximum in both the left- and right-hand sides of (2.30) is obtained for $l = 1$. Hence,

$$\frac{4}{1-\left(\frac{1}{1+\eta}\right)^{2\nu}} + \mathcal{O}\left(\frac{1}{N^2}\right) \;\le\; \kappa_{\mathbf{e}^\perp} \;\le\; \frac{4}{1-\left(\frac{1+\frac{\eta}{2}}{1+\eta}\right)^{2\nu}}.$$

This result is illustrated in Figure 2, where we have plotted both bounds (for $N = \infty$) and the actual condition number (for $N = 256$) as a function of η. As for pairwise aggregation, the condition number increases when η decreases, and only results for $\eta \ge 1$ are meaningful because otherwise it is better to coarsen along the y direction.

2.5. Boxwise quadruplet aggregation. Here we consider quadruplets obtained by grouping the nodes boxwise, as obtained with a pairwise aggregation in the x direction followed by a pairwise aggregation in the y direction. We assume N even and the prolongation is the $N^2 \times \frac{N^2}{4}$ matrix

$$P \;=\; \mathcal{P}_N \otimes \mathcal{P}_N. \tag{2.31}$$

With (2.15), (2.16), one then sees that the coarse grid matrix is

$$P^T\,A\,P \;=\; 2\,\eta\,\mathcal{I}_{\frac{N}{2}} \otimes \mathcal{A}_{\frac{N}{2}} \;+\; 2\,\mathcal{A}_{\frac{N}{2}} \otimes \mathcal{I}_{\frac{N}{2}} \;=\; 2\,A^{(\eta)}_{\frac{N}{2},\frac{N}{2}}. \tag{2.32}$$

Hence, T (with $\alpha = 1$) has the same eigenvalues as

$$\Big(I \;-\; \frac{1}{2}\;(C_N \otimes C_N)\left(\eta\mathcal{I}_{\frac{N}{2}} \otimes \Gamma_{\frac{N}{2}} + \Gamma_{\frac{N}{2}} \otimes \mathcal{I}_{\frac{N}{2}}\right)^{+} \\ \left(C_N^H \otimes C_N^H\right)\left(\eta\mathcal{I}_N \otimes \Gamma_N + \Gamma_N \otimes \mathcal{I}_N\right)\Big)\;\Sigma.$$

The latter matrix is block diagonal with (after symmetric permutation) 4×4 diagonal blocks

$$T_{kl} = \left(I - \mathbf{c}_{k,l} \left(2\lambda_{k,l}^{(\frac{N}{2},\frac{N}{2},\eta)}\right)^{+} \mathbf{c}_{k,l}^{H} \begin{pmatrix} \lambda_{k,l}^{(N,N,\eta)} & & & \\ & \lambda_{k,l+\frac{N}{2}}^{(N,N,\eta)} & & \\ & & \lambda_{k+\frac{N}{2},l}^{(N,N,\eta)} & \\ & & & \lambda_{k+\frac{N}{2},l+\frac{N}{2}}^{(N,N,\eta)} \end{pmatrix} \right) \begin{pmatrix} \sigma_{k,l} & & & \\ & \sigma_{k,l+\frac{N}{2}} & & \\ & & \sigma_{k+\frac{N}{2},l} & \\ & & & \sigma_{k+\frac{N}{2},l+\frac{N}{2}} \end{pmatrix},$$

where

$$\mathbf{c}_{k,l} = \begin{pmatrix} c_k^N & c_l^N \\ c_k^N & c_{l+\frac{N}{2}}^N \\ c_{k+\frac{N}{2}}^N & c_l^N \\ c_{k+\frac{N}{2}}^N & c_{l+\frac{N}{2}}^N \end{pmatrix}$$

with c_i^N as in Proposition 2.1 ($k = 0, \ldots, \frac{N}{2} - 1$ and $l = 0, \ldots, \frac{N}{2} - 1$).

THEOREM 2.5. *Let $A = A_{N,N}^{(\eta)}$ with N even, let P be given by* (2.31)*, and let M be a symmetric matrix satisfying* (2.11)*. If, for $k = 0, \ldots, N-1$ and $l = 0, \ldots, \frac{N}{2} - 1$,*

$$\begin{cases} \sigma_{k,l} \;\geq\; \sigma_{k,l+\frac{N}{2}} & \text{and} \quad \sigma_{l,k} \;\geq\; \sigma_{l+\frac{N}{2},k} \quad \text{if } l < \frac{N}{4}, \\ \sigma_{k,l} \;\leq\; \sigma_{k,l+\frac{N}{2}} & \text{and} \quad \sigma_{l,k} \;\leq\; \sigma_{l+\frac{N}{2},k} \quad \text{otherwise} \end{cases} \tag{2.33}$$

and if, for $l = 1, \ldots, N-1$,

$$\max_{k=0,\ldots,N-1} \sigma_{k,l} \;=\; \sigma_{0,l} \quad \text{and} \quad \max_{k=0,\ldots,N-1} \sigma_{l,k} \;=\; \sigma_{l,0}, \tag{2.34}$$

then, for $l = 0, \ldots, \frac{N}{2} - 1$, letting

$$t_l = \begin{cases} \max\left(\dfrac{3\,\sigma_{l,0} \;+\; \sigma_{l+\frac{N}{2},0}}{4}, \; \dfrac{3\,\sigma_{0,l} \;+\; \sigma_{0,l+\frac{N}{2}}}{4} \right) & \text{if } 0 < l < \frac{N}{4}, \\ \max\left(\dfrac{\sigma_{l,0} \;+\; 3\,\sigma_{l+\frac{N}{2},0}}{4}, \; \dfrac{\sigma_{0,l} \;+\; 3\,\sigma_{0,l+\frac{N}{2}}}{4} \right) & \text{if } l \geq \frac{N}{4}, \\ \max\left(\sigma_{\frac{N}{2},0}, \; \sigma_{0,\frac{N}{2}} \right) & \text{if } l = 0, \end{cases} \tag{2.35}$$

$$s_l = \begin{cases} \max\left(\dfrac{\sigma_{l,0} \;+\; \sigma_{l+\frac{N}{2},0}}{2}, \; \dfrac{\sigma_{0,l} \;+\; \sigma_{0,l+\frac{N}{2}}}{2} \right) & \text{if } l > 0, \\ \max\left(\sigma_{\frac{N}{2},0}, \; \sigma_{0,\frac{N}{2}} \right) & \text{if } l = 0, \end{cases} \tag{2.36}$$

$\lambda_{\mathbf{e}^\perp}^{(\max)}(T)$ defined by (2.3) *with $\alpha = 1$ and* (2.6) *satisfies*

$$\max_{l=0,\ldots,\frac{N}{2}-1} s_l \;\leq\; \lambda_{\mathbf{e}^\perp}^{(\max)}(T) \;\leq\; \max_{l=0,\ldots,\frac{N}{2}-1} t_l. \tag{2.37}$$

Proof. We first consider the block T_{00}. Since $(\lambda_{0,0}^{(\frac{N}{2},\frac{N}{2},\eta)})^+ = 0$, T_{00} has eigenvalues $\sigma_{0,0}$, $\sigma_{0,\frac{N}{2}}$, $\sigma_{\frac{N}{2},0}$, and $\sigma_{\frac{N}{2},\frac{N}{2}}$, but only the last three have to be taken into account, which is properly done with the given definitions of t_0 and s_0, noting that by (2.33) (with $k = \frac{N}{2}$ and $l = 0$), $\sigma_{\frac{N}{2},\frac{N}{2}} \le \sigma_{\frac{N}{2},0}$.

For other blocks, Lemma 2.2 may be applied with $m = 4$ and

$$\beta\,\delta_i\,|\mathbf{y}_i|^2 \;=\; \frac{|c_{k_i}^N\,c_{l_i}^N|^2\,\lambda_{k_i,l_i}^{(N,N,\eta)}}{2\,\lambda_{k,l}^{(\frac{N}{2},\frac{N}{2},\eta)}},$$

where $\{(k_1,l_1),\,(k_2,l_2),\,(k_3,l_3),\,(k_4,l_4)\} = \{(k,l),\,(k,l+\frac{N}{2}),\,(k+\frac{N}{2},l),\,(k+\frac{N}{2},l+\frac{N}{2})\}$. Using the same tricks as in the proof of Theorem 2.4, one finds

$$\begin{aligned}\beta\,\delta_i\,|\mathbf{y}_i|^2 &= \frac{\gamma_{k_i+\frac{N}{2}}^N\,\gamma_{l_i+\frac{N}{2}}^N\,\left(\eta\,\gamma_{l_i}^N+\gamma_{k_i}^N\right)}{8\,\left(\eta\,\gamma_{l_i}^N\,\gamma_{l_i+\frac{N}{2}}^N+\gamma_{k_i}^N\,\gamma_{k_i+\frac{N}{2}}^N\right)}\\ &= \frac{1}{4}+\frac{\eta\,\gamma_{l_i}^N\,\gamma_{l_i+\frac{N}{2}}^N\,\cos\left(\frac{2k_i\pi}{N}\right)+\gamma_{k_i}^N\,\gamma_{k_i+\frac{N}{2}}^N\,\cos\left(\frac{2l_i\pi}{N}\right)}{4\,\left(\eta\,\gamma_{l_i}^N\,\gamma_{l_i+\frac{N}{2}}^N+\gamma_{k_i}^N\,\gamma_{k_i+\frac{N}{2}}^N\right)}.\end{aligned} \tag{2.38}$$

We now consider the blocks T_{k0}, $k \ge 1$. One then has $\gamma_{l_i}^N\,\gamma_{l_i+\frac{N}{2}}^N = 0$, and $\beta\,\delta_i\,|\mathbf{y}_i|^2 = \frac{1}{2}$ for the pairs $(k,0)$, $(k+\frac{N}{2},0)$, whereas $\beta\,\delta_i\,|\mathbf{y}_i|^2 = 0$ for the pairs $(k,\frac{N}{2})$, $(k+\frac{N}{2},\frac{N}{2})$. From the characteristic equation given in Lemma 2.2, it then follows that the four eigenvalues are 0, $(\sigma_{k,0}+\sigma_{k+\frac{N}{2},0})/2$, $\sigma_{k,\frac{N}{2}}$, and $\sigma_{k+\frac{N}{2},\frac{N}{2}}$. Note that by (2.34), $\sigma_{k,\frac{N}{2}}$, $\sigma_{k+\frac{N}{2},\frac{N}{2}} \le \sigma_{0,\frac{N}{2}}$. Similarly, the blocks T_{0l} with $l \ge 1$ have the four eigenvalues 0, $(\sigma_{0,l}+\sigma_{0,l+\frac{N}{2}})/2$, $\sigma_{\frac{N}{2},l} \le \sigma_{\frac{N}{2},0}$, and $\sigma_{\frac{N}{2},l+\frac{N}{2}} \le \sigma_{\frac{N}{2},0}$. With the above discussion of T_{00}, it is shown that the lower bound in (2.37) is in fact equal to the maximal eigenvalue of blocks T_{kl} with either $k = 0$ or $l = 0$.

This proves the left inequality (2.37). To prove the right one, and because (2.33) implies $t_l \ge s_l$, we are left with the analysis of blocks T_{kl} with $k,l \ge 1$. With (2.33), (2.38) shows that $\beta\,\delta_i\,|\mathbf{y}_i|^2 \ge \frac{1}{4}$ for $i = 1$ corresponding to the largest σ_{k_i,l_i}. Further, by virtue of (2.33) again, the two largest σ_{k_i,l_i} are either $\sigma_{k_i,l}$, $\sigma_{k_i,l+\frac{N}{2}}$ (with $k_i = k$ or $k_i = k+\frac{N}{2}$), or σ_{k,l_i}, $\sigma_{k+\frac{N}{2},l_i}$ (with $l_i = l$ or $l_i = l+\frac{N}{2}$). In the first case, and if $l < \frac{N}{4}$, Lemma 2.2 yields the upper bound

$$\frac{3\,\sigma_{k_i,l}\;+\;\sigma_{k_i,l+\frac{N}{2}}}{4}\;\le\;\frac{3\,\sigma_{0,l}\;+\;\sigma_{0,l+\frac{N}{2}}}{4}\;\le\;t_l.$$

Other cases can be discussed similarly, showing altogether that the largest eigenvalue of T_{kl} does not exceed $\max(t_k,\,t_l)$. The required upper bound straightforwardly follows. ∎

Note that here the x and y directions play a symmetric role. If the smoother preserves this symmetry, then the condition number will have same value for η and $1/\eta$. Considering damped Jacobi smoothing ($\sigma_{k,l}$ given by (2.14)), we may thus assume $\eta \ge 1$ without loss of generality. The maximum in both the left- and right-hand sides of (2.37) is obtained for $l = 1$ and

$$\frac{2}{1-\left(\frac{\eta}{1+\eta}\right)^{2\nu}}\;+\;\mathcal{O}\left(\frac{1}{N^2}\right)\;\le\;\kappa_{\mathbf{e}^\perp}\;\le\;\frac{4}{1-\left(\frac{\eta}{1+\eta}\right)^{2\nu}}.$$

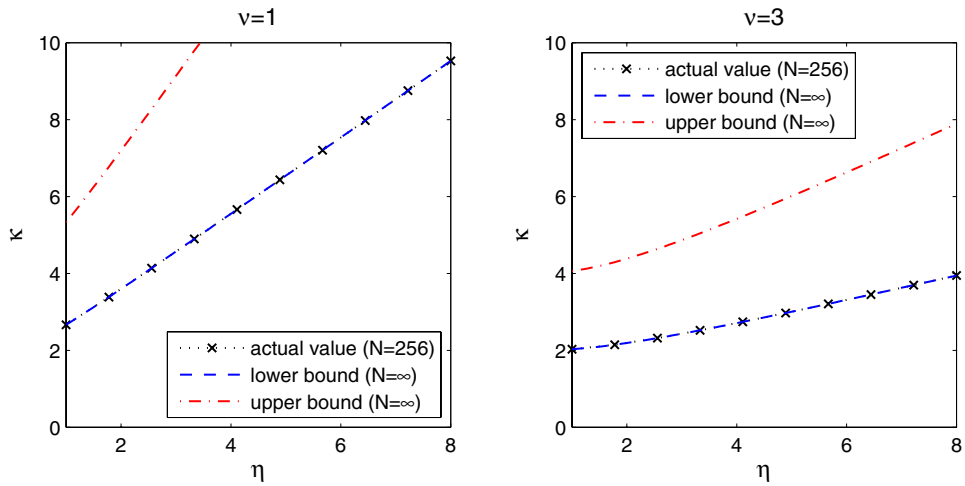

FIG. 3. *Actual condition number and its upper and lower bounds for boxwise quadruplet aggregation with damped Jacobi smoothing, as a function of* η *(*$\alpha = 1$*).*

This result is illustrated in Figure 3, where we have plotted both bounds (for $N = \infty$) and the actual condition number (for $N = 256$) as a function of η. One sees that the actual condition number matches the lower bound and that the upper bound overestimates it by a factor of 2. This stems from the fact that the lower bound is the maximal eigenvalue of the blocks T_{kl} with either $k = 0$ or $l = 0$, while the actual maximum is in one of these blocks (namely T_{10}). In view of (2.34), one may conjecture that these blocks T_{k0} or T_{0l} will indeed most often contain this actual maximum; hence, the gap between the upper and the lower bounds represents a shortcoming in our analysis.

On the other hand, one sees that the condition number here increases as η increases. Our results, which depends only on $\sigma_{k,l}$, show that this may be cured with a smoother that properly handles anisotropy, such as a line Jacobi [14]. One may also follow the philosophy of "algebraic" multigrid (AMG) methods [13], in which the smoother is fixed and the coarsening adapted to the problem. Considering the results obtained in the preceding subsection, one must shift from boxwise to linewise quadruplet aggregation when η is above a given threshold. Note that the optimal threshold depends on the smoother. This approach is illustrated in the next section.

3. Multigrid cycles. As written in the introduction, the results obtained in the preceding section do not allow one to prove optimal order convergence of a multigrid method with V- or W-cycle. They are, however, compatible with the practical requirements stated in [12] for near optimal order convergence with K_2-cycle multigrid. In this scheme, the multigrid method is used as a preconditioner. It is may be seen as an "inexact" two-grid scheme in which the coarse-grid systems are solved by 2 steps of a Krylov subspace iterative method, using the same preconditioning scheme on the coarser level. This procedure is recursively followed until the coarsest level where an exact solve is performed. Note that one has to select for these inner iterations and for the outer iteration a method that accommodates slightly variable preconditioning. For symmetric positive definite systems, the choice method is the "flexible" (or "generalized") conjugate gradient method from [1, 9] (on which the analysis in [12]

is based). In Figure 4, we give the algorithm for the two-grid scheme and both the K_2- and V-cycle multigrid schemes. For the sake of clarity, we simplified somewhat the original formulation of the K_2-cycle, enforcing the use of the same scheme at each level (the algorithm in [12] allows for hybrid schemes; in the numerical experiments reported later we always use the version of Figure 4). On the other hand, we supplement the K_2-cycle with a threshold technique: If, after 1 inner iteration, the relative residual error for the coarse system being solved is below some threshold t, then the second inner iteration is skipped. We indeed observed that, setting $t \approx 0.25$, the convergence remains the same but each iteration is, on average, somewhat cheaper. This version is referred to as $K_{2(t)}$-cycle multigrid.

Input: $\mathbf{r}_k$; Output: $\mathbf{z}_k = \mathrm{MGprec}(\mathbf{r}_k)$.

1. Relax using smoother M_k: $\mathbf{v}_k = M_k^{-1}\,\mathbf{r}_k$.
2. Compute new residual: $\widetilde{\mathbf{r}}_k = \mathbf{r}_k - A_k\,\mathbf{v}_k$.
3. Restrict residual: $\mathbf{r}_{k-1} = P_k^T\,\widetilde{\mathbf{r}}_k$.
4. Compute an (approximate) solution $\widetilde{\mathbf{x}}_{k-1}$ to $A_{k-1}\,\mathbf{x}_{k-1} = \mathbf{r}_{k-1}$:
   if Two-grid **or** $k = 1$ **then** $\widetilde{\mathbf{x}}_{k-1} = A_{k-1}^{-1}\,\mathbf{r}_{k-1}$
   else if V-Cycle **then** $\widetilde{\mathbf{x}}_{k-1} = \mathrm{MGprec}(\mathbf{r}_{k-1})$
   else if $K_{2(t)}$-cycle **then**
   Perform 1 or 2 flexible CG iterations with the multigrid preconditioner:
   $\mathbf{d}_{k-1} = \mathrm{MGprec}(\mathbf{r}_{k-1})$
   $\alpha_{k-1} = \frac{\mathbf{r}_{k-1}^T \mathbf{d}_{k-1}}{\mathbf{d}_{k-1}^T A_{k-1} \mathbf{d}_{k-1}}$
   $\mathbf{x}_{k-1} = \alpha_{k-1}\,\mathbf{d}_{k-1}$
   $\widetilde{\mathbf{r}}_{k-1} = \mathbf{r}_{k-1} - \alpha_{k-1}\,A_{k-1}\mathbf{d}_{k-1}$
   if $\|\widetilde{\mathbf{r}}_{k-1}\| \le t\,\|\mathbf{r}_{k-1}\|$ **then** $\widetilde{\mathbf{x}}_{k-1} = \mathbf{x}_{k-1}$
   else
   $\mathbf{c}_{k-1} = \mathrm{MGprec}(\widetilde{\mathbf{r}}_{k-1})$
   $\widetilde{\mathbf{d}}_{k-1} = \mathbf{d}_{k-1} - \frac{\mathbf{c}_{k-1}^T A_{k-1} \mathbf{d}_{k-1}}{\mathbf{d}_{k-1}^T A_{k-1} \mathbf{d}_{k-1}}\,\mathbf{c}_{k-1}$
   $\widetilde{\mathbf{x}}_{k-1} = \mathbf{x}_{k-1} - \frac{\widetilde{\mathbf{r}}_{k-1}^T \widetilde{\mathbf{d}}_{k-1}}{\widetilde{\mathbf{d}}_{k-1}^T A_{k-1} \widetilde{\mathbf{d}}_{k-1}}\,\widetilde{\mathbf{d}}_{k-1}$
   end if
   end if
5. Prolongate coarse-grid correction: $\mathbf{x}_k = P_k\,\widetilde{\mathbf{x}}_{k-1}$.
6. Compute new residual: $\overline{\mathbf{r}}_k = \widetilde{\mathbf{r}}_k - A_k\,\mathbf{x}_k$.
7. Relax using smoother M_k: $\mathbf{w}_k = M_k^{-1}\,\overline{\mathbf{r}}_k$.
8. $\mathbf{z}_k = \mathbf{v}_k + \mathbf{x}_k + \mathbf{w}_k$.

FIG. 4. *Algorithm defining two-grid, V-cycle multigrid, and $K_{2(t)}$-cycle multigrid preconditioning at level k ($k \ge 1$) for matrix A_k, based on prolongation P_k, on 1 pre- and 1 postsmoothing step with smoother M_k, and on flexible CG as an inner Krylov subspace solver for the $K_{2(t)}$-cycle.*

We first illustrate these multigrid cycles on a model anisotropic problem like in section 2, but with Dirichlet boundary conditions. We select a coarsening method that performs linewise quadruplet aggregation when $\eta > 2$ and boxwise quadruplet aggregation otherwise. We set the threshold to 2 because then the model aggregation investigated here performs roughly as the "black box" aggregation algorithm proposed in [10]. Note that the type of coarsening is not fixed once for all on the finest grid; that is, at each level, the selection of the coarsening scheme is performed independently, based on the matrix as defined at the considered level. Consider, for instance, $\eta = 16$. According to the rule, linewise quadruplet aggregation is selected on the finest

grid. This induces, on the next level, a matrix with the same structure but with the anisotropy ratio $\eta_1 = \eta/4 = 4$ (see (2.27)). Thus, linewise quadruplet aggregation is again selected on this first coarse level. However, the matrix induced at the next level has the anisotropy ratio $\eta_2 = \eta_1/4 = \eta/16 = 1$. Hence, the selected coarsening scheme will be boxwise quadruplet aggregation. Since this leaves the anisotropy ratio unchanged (see (2.32)), boxwise quadruplet aggregation will be further applied at all subsequent levels. Thus, when $\eta = 16$, the aggregation rule automatically selects twice linewise quadruplet aggregation and then shifts to boxwise quadruplet aggregation.

Note that there are critical values of η corresponding to a transition between coarsening schemes. Below, we pay attention to these critical values and always report the results obtained with both η slightly below and η slightly above the limit where the coarsening scheme changes.

In Table 1 we report the CPU time and the number of (outer) conjugate gradient iterations needed to reduce the relative residual error by 10^{-6}, when solving a linear system with the right-hand side equal to $\begin{pmatrix}1 & \cdots & 1\end{pmatrix}^T$ and the zero vector as initial approximation. In all cases, we use the symmetric Gauss–Seidel smoother with 1 pre- and 1 postsmoothing step. The $\mathrm{K}_{2(t_{0.25})}$-cycle refers to the $\mathrm{K}_{2(t)}$-cycle multigrid scheme described previously with a threshold $t = 0.25$. For the two-grid and V-cycle variants, the standard implementation of the conjugate gradient method is used, whereas for the $\mathrm{K}_{2(t_{0.25})}$-cycle the flexible variant is used. The coarsest grid has in most cases 256 nodes. That is, we use 7 levels (6 coarsenings) for the 1024×1024 grid and 5 levels (4 coarsenings) for the 256×256 grid. There are two exceptions, corresponding to the cases where $\eta = 10^4$ and $\eta = 10^5$. For the 1024×1024 grid, 5 linewise quadruplet aggregations result in a coarse grid matrix with 1024 nodes that has a one-dimensional structure, and we keep this matrix as the coarsest one, fixing thereby the number of levels to 6 (5 coarsenings).

One sees that the V-cycle is not optimal, whereas iteration counts for the $\mathrm{K}_{2(t_{0.25})}$ are nearly independent of the grid size, and only slightly larger than those observed for the two-grid method on a grid with moderate size.

We further consider the following more difficult example, with coefficient jumps and a mix of isotropic and anisotropic regions. More precisely, we consider the linear system resulting from the five point finite difference approximation of

$$-\frac{\partial}{\partial x}\left(a_x \frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial y}\left(a_y \frac{\partial u}{\partial y}\right) \;=\; f \qquad \text{in} \quad \Omega = (0,1)\times(0,1)$$

with boundary conditions

$$\begin{cases} u &= 0 \qquad \text{on } \; x=0,\, 0\le y\le 1 \text{ and } y=0,\, 0\le x\le 1\ , \\ \frac{\partial u}{\partial n} &= 0 \qquad \text{elsewhere on } \; \partial\Omega \end{cases}$$

and coefficients given by

$$\begin{cases} a_x = 10^3, & a_y = 1, & f = 0 & \text{in } \; (0,\,1)\times(0,\,0.5), \\ a_x = 1, & a_y = 1, & f = 0 & \text{in } \; (0,\,0.5)\times(0.5,\,1), \\ a_x = 10^{-3}, & a_y = 10^{-3}, & f = 1 & \text{in } \; (0.5,\,1)\times(0.5,\,1). \end{cases}$$

We use a uniform mesh with a constant mesh size h in both directions, $h^{-1} = 256$ or $h^{-1} = 1024$.

For this problem, according our general philosophy, we adapt the coarsening to the coefficients. That is, linewise quadruplet aggregation is applied in the anisotropic

TABLE 1

Number of iterations and CPU time (in seconds) needed to reduce the relative residual error by 10^{-6} for the model anisotropic problem; the second column indicates the successive coarsenings, each ℓ referring to linewise aggregation and each b referring to boxwise aggregation (thus, for instance, ($\ell\ell bbbb$) means 2 linewise quadruplet aggregations followed by 4 boxwise quadruplet aggregations).

		Two-grid	$K_{2(t_{0.25})}$-cycle				V-cycle			
	grid	256×256	256×256		1024×1024		256×256		1024×1024	
		#it.	#it.	time	#it.	time	#it.	time	#it.	time
η	coars.									
1.0	($bbbbbb$)	10	10	0.85	11	17.8	27	1.55	58	57.4
1.9	($bbbbbb$)	10	10	0.82	11	17.4	27	1.54	58	57.2
2.1	($\ell bbbbb$)	15	16	1.33	17	27.2	33	1.90	69	70.7
4.0	($\ell bbbbb$)	15	16	1.37	17	28.0	34	1.98	79	82.0
7.9	($\ell bbbbb$)	15	16	1.34	17	27.2	33	1.90	67	68.8
8.1	($\ell\ell bbbb$)	15	19	1.60	21	33.2	38	2.20	78	78.1
16	($\ell\ell bbbb$)	16	19	1.63	21	33.2	40	2.37	84	84.1
31	($\ell\ell bbbb$)	16	19	1.57	21	32.4	38	2.25	76	76.2
33	($\ell\ell\ell bbb$)	16	19	1.54	22	34.0	42	2.40	88	86.3
64	($\ell\ell\ell bbb$)	16	19	1.54	22	34.4	45	2.56	90	90.0
127	($\ell\ell\ell bbb$)	16	19	1.56	22	33.8	48	2.74	95	93.1
129	($\ell\ell\ell\ell bb$)	16	19	1.53	22	33.7	48	2.76	106	104.
256	($\ell\ell\ell\ell bb$)	16	19	1.55	22	33.8	48	2.75	105	103.
511	($\ell\ell\ell\ell bb$)	16	19	1.55	22	33.9	47	2.71	101	99.1
513	($\ell\ell\ell\ell\ell b$)	16	19	1.55	22	34.0	47	2.71	101	98.9
1.e3	($\ell\ell\ell\ell\ell b$)	16	19	1.55	22	33.8	47	2.71	99	97.2
1.e4	($\ell\ell\ell\ell\ell$)	16	18	1.47	22	33.9	46	2.61	97	95.0
1.e5	($\ell\ell\ell\ell\ell$)	16	18	1.45	20	30.5	42	2.40	93	91.2

region (the first half of the matrix rows when using a rowwise ordering), and boxwise quadruplet aggregation elsewhere. We use 6 coarsenings (7 levels) for $h^{-1} = 256$ and 8 coarsenings (9 levels) for $h^{-1} = 1024$. Hence, the coarsest grid always has 16 nodes.[2] Note that, here again, this model aggregation scheme performs roughly as the "black box" aggregation algorithm in [10].

The results are given in Table 2. For the sake of completeness, we consider scaling the coarse grid matrices according to the parameter α as in (2.3) (that is, all entries in the successive coarse grid matrices are multiplied by α^{-1}). As for the model anisotropic problem, the V-cycle is not optimal whereas the $K_{2(t_{0.25})}$-cycle exhibits near grid independent convergence. As expected from the results in [3, 18], scaling improves the behavior of the V-cycle, but to a limited extent. On the other hand, the convergence of two-grid and $K_{2(t_{0.25})}$-cycle multigrid schemes is nearly independent of the scaling.

4. Conclusion. We have developed the Fourier analysis of multigrid methods with coarsening based on "pure" aggregation schemes. Our results show that the two-grid convergence rates are independent of the grid size and can be made independent of anisotropies by a proper choice of the coarsening direction, as is natural in an AMG-like approach. These two-grid convergence rates are too large for having an optimal order method with standard multigrid cycles. However, the K_2-cycle (where the coarse systems are solved by two steps of a Krylov subspace iterative method)

[2]It means that, at some point in the coarsening process, the matrix graph for the anisotropic region reduces to a vertical line; hence, linewise aggregation along the x direction is no longer possible, and linewise aggregation along that vertical line (y direction) is used instead to ensure that the number of variables is still reduced by a factor of 4.

TABLE 2
Number of iterations and CPU time (in seconds) needed to reduce the relative residual error by 10^{-6} for the problem with coefficients jumps; α is the scaling parameter.

	Two-grid	$K_{2(t_{0.25})}$-cycle				V-cycle			
grid	256×256	256×256		1024×1024		256×256		1024×1024	
	#it.	#it.	time	#it.	time	#it.	time	#it.	time
α									
1.0	19	24	1.92	27	41.5	75	4.22	166	163.
1.1	19	23	1.85	27	41.7	67	3.82	157	156.
1.2	19	22	1.75	28	42.9	65	3.71	148	145.
1.3	19	22	1.75	28	43.0	64	3.64	142	139.
1.4	19	22	1.75	27	40.8	63	3.59	136	135.
1.5	19	22	1.77	27	41.3	62	3.49	133	130.
1.6	19	22	1.75	27	41.4	61	3.46	132	130.
1.7	19	23	1.81	27	41.3	61	3.49	131	130.
1.8	19	23	1.83	27	41.5	61	3.48	131	128.
1.9	19	23	1.83	27	41.2	61	3.45	130	128.
2.0	19	23	1.82	27	41.5	62	3.53	133	130.

exhibits near grid independent convergence. This convergence is too slow to make the approach competitive with geometric multigrid. Nevertheless, combined with a proper automatic aggregation algorithm, the approach is potentially attractive as an AMG-like method.

REFERENCES

[1] O. Axelsson and P. S. Vassilevski, *A black-box generalized conjugate gradient solver with inner iterations and variable-step preconditioning*, SIAM J. Matrix Anal. Appl., 12 (1991), pp. 625–644.

[2] A. Ben Israel and T. N. E. Greville, *Generalized Inverses: Theory and Applications*, John Wiley and Sons, New York, 1974.

[3] D. Braess, *Towards algebraic multigrid for elliptic problems of second order*, Computing, 55 (1995), pp. 379–393.

[4] D. Braess, *Finite Elements. Theory, Fast Solvers, and Applications in Solid Mechanics*, Cambridge University Press, Cambridge, UK, 1997.

[5] V. E. Bulgakov, *Multi-level iterative technique and aggregation concept with semi-analytical preconditioning for solving boundary value problems*, Comm. Numer. Methods in Engrg., 9 (1993), pp. 649–657.

[6] G. H. Golub and C. F. van Loan, *Matrix Computations*, 3rd ed., The Johns Hopkins University Press, Baltimore, MD 1996.

[7] W. Hackbusch, *Multigrid Methods and Applications*, Springer-Verlag, Berlin, 1985.

[8] W. Leontief, *The Structure of the American Economy* 1919–1939, Oxford University Press, New York, 1951.

[9] Y. Notay, *Flexible conjugate gradients*, SIAM J. Sci. Comput., 22 (2000), pp. 1444–1460.

[10] Y. Notay, *Aggregation-based algebraic multilevel preconditioning*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 998–1018.

[11] Y. Notay, *Convergence analysis of perturbed two-grid and multigrid methods*, SIAM J. Numer. Anal., 45 (2007), pp. 1035–1044.

[12] Y. Notay and P. S. Vassilevski, *Recursive Krylov-based multigrid cycles*, Numer. Linear Algebra Appl., to appear. Available online at http://homepages.ulb.ac.be/∼ynotay.

[13] K. Stüben, *An introduction to algebraic multigrid*, in Multigrid, Academic Press, San Diego, 2001, pp. 413–532, Appendix A.

[14] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, San Diego, 2001.

[15] P. Vaněk, M. Brezina, and R. Tezaur, *Two-grid method for linear elasticity on unstructured meshes*, SIAM J. Sci. Comput., 21 (1999), pp. 900–923.

[16] P. Vaněk, J. Mandel, and M. Brezina, *Algebraic multigrid based on smoothed aggregation for second and fourth order elliptic problems*, Computing, 56 (1996), pp. 179–196.
[17] R. Wienands and W. Joppich, *Practical Fourier Analysis for Multigrid Methods*, Chapman & Hall/CRC Press, Boca Raton, FL, 2005.
[18] R. Wienands and C. W. Oosterlee, *On three-grid Fourier analysis for multigrid*, SIAM J. Sci. Comput., 23 (2001), pp. 651–671.