Numer. Math. (2011) 118:457–483
DOI 10.1007/s00211-011-0362-7

Numerische
Mathematik

Smoothing factor, order of prolongation and actual multigrid convergence

Artem Napov · Yvan Notay

Received: 30 March 2010 / Published online: 19 February 2011

Abstract We consider the Fourier analysis of multigrid methods (of Galerkin type) for symmetric positive definite and semi-positive definite linear systems arising from the discretization of scalar partial differential equations (PDEs). We relate the so-called smoothing factor to the actual two-grid convergence rate and also to the convergence rate of the V-cycle multigrid. We derive a two-sided bound that defines an interval containing both the two-grid and V-cycle convergence rate. This interval is narrow and away from 1 when both the smoothing factor and an additional parameter are small enough. Besides the smoothing factor, the convergence mainly depends on the angle between the range of the prolongation and the eigenvectors of the system matrix associated with small eigenvalues. Nice V-cycle convergence is guaranteed if the tangent of this angle has an upper bound proportional to the eigenvalue, whereas nice two-grid convergence requires a bound proportional to the square root of the eigenvalue. We also discuss the well-known rule which relates the order of the prolongation to that of the differential operator associated to the problem. We first define a *frequency based* order which in most cases amounts to the so-called *high frequency* order as defined in Hemker (J Comput Appl Math 32:423–429, 1990). We give a firmer basis to the related order rule by showing that, together with the requirement of having the smoothing factor away from one, it provides *necessary* and *sufficient* conditions for having the two-grid convergence rate away from 1. A stronger condition is further shown to be sufficient for optimal convergence with the V-cycle. The presented results apply to rigorous Fourier analysis for regular discrete PDEs, and also to local Fourier analysis

A. Napov was supported by the Belgian FNRS ("Aspirant"). Y. Notay was supported by the Belgian FNRS ("Directeur de recherches").

A. Napov (✉) · Y. Notay
Service de Métrologie Nucléaire, Université Libre de Bruxelles (C.P. 165/84),
50, Av. F.D. Roosevelt, 1050 Brussels, Belgium
e-mail: anapov@ulb.ac.be

via the discussion of semi-positive systems as may arise from the discretization of PDEs with periodic boundary conditions.

Mathematics Subject Classification (2000) 65F10 · 65N55 · 65N12

1 Introduction

We consider Fourier analysis of multigrid methods for symmetric positive definite (SPD) and, more generally, symmetric semi-positive definite $n \times n$ linear systems

$$A\mathbf{x} = \mathbf{b}. \tag{1.1}$$

Multigrid methods are based on the recursive use of a two-grid scheme. A basic two-grid method combines the action of a *smoother*, often a simple iterative method such as Gauss–Seidel, and a *coarse grid correction*, which corresponds to the solution of the residual equation on a coarser grid. A multigrid method is obtained when the residual equation is solved approximately applying few iterations of the two-grid scheme on that level, and so on, until the coarsest level where an exact solve is performed. If the two-grid method is used recursively once on each level, the resulting algorithm is called V-cycle multigrid, whereas more involved cycling strategies (like W- or F-cycle) correspond to more iterations of the two-grid method on given levels (see, e.g., [7,20,22]).

Fourier analysis [19,20,23] is a widely used tool that helps to design efficient multigrid approaches. It exploits the fact that the discretization of a constant-coefficient (elliptic) boundary value problem on simple domains often leads to a system (1.1) of which discrete Fourier modes are eigenvectors. If, in addition, other multigrid components also have a simple block structure in this Fourier basis, the analysis of a multigrid method can be reduced to the analysis of diagonal blocks of small size, which can be done either analytically or numerically. The multigrid components designed for such simple cases are then adapted to more complex problems.

Fourier analysis is in practice limited to a few consecutive grids: generally two, rarely three [24]. Often Fourier analysis is further reduced to the computation of a simpler (one-grid) *smoothing factor*. When assessing this latter, the coarse grid correction is assumed to annihilate the so-called *smooth* (or *low frequency*) error modes, while leaving *rough* (or *high frequency*) modes unchanged. Since this is the limit case of the desired behavior of a coarse grid correction, the smoothing factor is often considered as an ideal two-grid convergence estimate. However, it is so far unclear which condition(s) are to be satisfied by the coarse grid correction for having the actual two-grid convergence close to this ideal. Further, nice two-grid convergence does not necessarily imply optimal convergence of the multigrid method with V-cycle [11], hence the latter likely requires additional conditions.

In this paper we investigate these questions for symmetric multigrid schemes of Galerkin type. We establish a connection between the smoothing factor and the actual two-grid convergence via an additional parameter α. This latter mainly depends on the coefficients of the prolongation in the Fourier basis and therefore allows to quantify the effect of the coarse grid correction. Regarding the V-cycle convergence rate, we use as main tool McCormick's bound [10] (see also [9,15]) which is shown in [13] to be the best convergence estimate among those that can be assessed considering only two consecutive levels at a time. We prove a relation between McCormick's bound and the smoothing factor, using the same parameter α. When this parameter and the smoothing factor are both nicely bounded at each level, our analysis essentially proves that the two-grid and the V-cycle convergence factors are both in a narrow interval, which further goes towards zero as the number of smoothing steps is increased.

Assuming that the smoothing factor is bounded away from 1, we further deduce easy-to-check conditions to be satisfied by the prolongation for optimal two-grid or V-cycle convergence. Doing so, we give in some sense a more precise meaning to statements like "Interpolation must be able to approximate an eigenvector with error bound proportional to the size of the associated eigenvalue" [4, p. 1573], [5, p. 4]. We also highlight that the conditions for guaranteed optimal V-cycle convergence are actually stronger that the conditions for optimal two-grid convergence.

We further discuss the well-known rule on the order of inter-grid transfer operators [2]. When the restriction is the transpose of the prolongation (as considered in this paper), this rule states that the so-called *hight-frequency order* of the prolongation should not be smaller that the half of the order of the differential operator; otherwise, the coarse grid correction operator is unbounded, leading to expect bad two-grid convergence [8]. Here we define a closely related *frequency based order* and, to be rigorous, a *lower* and an *upper* order of the differential operator, noting that both are equal in usual applications. In this latter case, we show in particular that the convergence factor of the two-grid method is away from 1 *if and only if* the smoothing factor is itself bounded away from 1 *and* the frequency based order is not smaller than the half of the order of the differential operator. We further prove that the convergence of the V-cycle is optimal if, in addition, the frequency based order is not smaller than the order of the differential operator.

Note that our results do not only give a firmer basis to the analysis in [8] by providing necessary and sufficient conditions. They also highlights the distinct and complementary role played by the smoothing factor and the order of the prolongation. Our analysis also builds bridges between the works [4,5] quoted above that connect the two-grid convergence to the quality of the prolongation for smooth modes, and the order rule which focus on the behavior for high frequency modes.

In a number of practical cases, when Fourier analysis cannot be applied directly, it is still possible to replace boundary conditions by the periodic ones, to make Fourier analysis work. Such modification has little influence on the convergence rate provided that some extra smoothing is performed on the boundary [3,17]. These approaches are closely related to *local Fourier analysis*, which can often be viewed [23, Remark 5.3], [20, Section 3.4.4] as a (rigorous) Fourier analysis for problems with periodic boundary conditions. Since such boundary conditions often lead to semi-positive definite (singular) systems (1.1), our treatment should be valid for them as well. This is addressed in this work via the extension of McCormick's bound to the semi-positive definite case.

The reminder of this paper is organized as follows. In Sect. 2 we state the general setting of this study for SPD systems and gather the needed assumptions. McCormick's bound is recalled in Sect. 3 and the smoothing factor is related to the two-grid and V-cycle multigrid convergence factors in Sect. 4. The conditions on the quality of the prolongation for smooth modes are presented in Sect. 5 and the rule on the order of inter-grid operators is discussed in Sect. 6. The extension to symmetric semi-positive definite systems is given in Sect. 7.

Notation

Let I denote the identity matrix and O the zero matrix. When the dimensions are not obvious from the context, we write more specifically I_m for the $m \times m$ identity matrix and $O_{m_1 \times m_2}$ for the $m_1 \times m_2$ zero matrix.

For any rectangular matrix B, B^T stands for its transpose and B^H for its complex conjugate transpose. For any $m \times m$ square matrix C, $\rho(C)$ is its spectral radius (that is, its largest eigenvalue in modulus); $\|C\| = \sqrt{\rho(C^H C)}$ is the usual 2-norm and, for a SPD matrix D, $\|C\|_D = \|D^{1/2} C D^{-1/2}\|$ is the D-norm (if $D = A$, it is also called energy norm).

2 General setting

We consider a multigrid method with $J + 1$ levels; $J > 1$ corresponds to a truly *multi*grid method, whereas $J = 1$ leads to a mere *two*-grid scheme. Index J refers to the finest level (on which the system (1.1) is to be solved), and index 0 to the coarsest level. The number of unknowns at level k, $0 \le k \le J$, is noted n_k (with thus $n_J = n$).

Our analysis applies to symmetric multigrid schemes based on the Galerkin principle for the SPD system (1.1); that is, restriction is the transpose of prolongation and the matrix A_k at level k, $k = J - 1, \ldots, 0$, is given by $A_k = P_k^T A_{k+1} P_k$, where P_k is the prolongation operator from level k to level $k + 1$; we also assume that the smoother R_k is SPD and that the number of pre-smoothing steps ν ($\nu > 0$) is equal to the number of post-smoothing steps.

The algorithm for V-cycle multigrid is defined as follows.

Multigrid with V-cycle at level k: $\mathbf{x}_{n+1} \leftarrow \mathrm{MG}(\mathbf{b}, A_k, \mathbf{x}_n, k)$

(1) Relax ν times with smoother R_k:
repeat ν times $\quad \mathbf{x}_n \leftarrow \mathbf{x}_n + R_k^{-1}(\mathbf{b}_k - A_k \mathbf{x}_n)$

(2) Compute residual: $\mathbf{r}_k \leftarrow \mathbf{b} - A_k \mathbf{x}_n$

(3) Restrict residual: $\mathbf{r}_{k-1} \leftarrow P_{k-1}^T \mathbf{r}_k$

(4) Coarse grid correction: **if** $k = 1$, $\mathbf{e}_0 \leftarrow A_0^{-1} \mathbf{r}_0$
else $\mathbf{e}_{k-1} \leftarrow \mathrm{MG}(\mathbf{r}_{k-1}, A_{k-1}, \mathbf{0}, k - 1)$

(5) Prolongate coarse grid correction: $\mathbf{x}_n \leftarrow \mathbf{x}_n + P_{k-1} \mathbf{e}_{k-1}$

(6) Relax ν times with smoother R_k:
repeat ν times $\quad \mathbf{x}_{n+1} \leftarrow \mathbf{x}_n + R_k^{-1}(\mathbf{b}_k - A_k \mathbf{x}_n)$

Observe that for $k = 1$ this algorithm corresponds to a standard two-grid method with exact coarse grid solve. Our analysis makes use of the following general assumptions.

General assumptions

- $n = n_J > n_{J-1} > \cdots > n_0$;
- P_k is an $n_{k+1} \times n_k$ matrix of rank n_k, $k = J-1, \ldots, 0$;
- $A_J = A$ and $A_k = P_k^T A_{k+1} P_k,\ k = J-1, \ldots, 0$;
- R_k is SPD and such that $\rho(I - R_k^{-1} A_k) < 1,\ k = J, \ldots, 1$.

Some of our results are stated with respect to the matrices $N_k^{(\nu)}$ defined from

$$N_k^{(\nu)} = \sum_{j=0}^{\nu-1} (I - R_k^{-1} A_k)^j R_k^{-1}, \tag{2.1}$$

which also satisfy

$$I - N_k^{(\nu)} A_k = (I - R_k^{-1} A_k)^\nu. \tag{2.2}$$

That is, $N_k^{(\nu)}$ is the relaxation operator that provides in 1 step the same effect as ν steps with R_k^{-1}. The results stated with respect to $N_k^{(\nu)}$ may then be seen as results stated for the case of 1 pre- and 1 post-smoothing step, which can be extended to the general case via the relation (2.2).

When applying the V-cycle algorithm, the error satisfies

$$A_k^{-1}\mathbf{b} - \mathbf{x}_{n+1} = E_{MG}^{(k)} (A_k^{-1}\mathbf{b} - \mathbf{x}_n)$$

where the iteration matrix $E_{MG}^{(k)}$ is recursively defined from

$$\begin{aligned} &E_{MG}^{(0)} = O \quad \text{and, for } k = 1, 2, \ldots, J: \\ &E_{MG}^{(k)} = (I - R_k^{-1} A_k)^\nu \left(I - P_{k-1}(I - E_{MG}^{(k-1)}) A_{k-1}^{-1} P_{k-1}^T A_k\right) (I - R_k^{-1} A_k)^\nu \end{aligned} \tag{2.3}$$

(see, e.g., [20, p. 48]). Note that for $J = 1$, (2.3) reduces to the two-grid iteration matrix:

$$E_{TG}^{(J)} = (I - R_J^{-1} A_J)^\nu \left(I - P_{J-1} A_{J-1}^{-1} P_{J-1}^T A_J\right) (I - R_J^{-1} A_J)^\nu. \tag{2.4}$$

The convergence on the finest level is governed by the spectral radius $\rho(E_{MG}^{(J)})$, or, in case of two-grid, $\rho(E_{TG}^{(J)})$. In this paper, we want to discuss assessment of these spectral radii within the framework of Fourier analysis, as may be developed for systems arising from the discretization of scalar partial differential equations (PDEs). It means that the eigenvectors of A_k are explicitly known at each level and form the *Fourier* *basis*. We further assume that the smoother shares the same set of eigenvectors; i.e., is also diagonal when expressed in this Fourier basis.[1] According to (2.1), $N_k^{(\nu)}$ will be diagonal as well for all ν.

Technically, Fourier analysis is then possible if, expressing P_{k-1} in both the coarse (level $k-1$) and fine (level k) Fourier basis, it has the form

$$P_{k-1} = \begin{pmatrix} \mathbf{p}_1^{(k-1)} & & & \\ & \mathbf{p}_2^{(k-1)} & & \\ & & \ddots & \\ & & & \mathbf{p}_{l_k}^{(k-1)} \end{pmatrix}, \tag{2.5}$$

where $\mathbf{p}_j^{(k-1)}$ are non-zero[2] (possibly complex) vectors of size $m_j^{(k)}$, $j = 1, \dots, l_k$. Note that this form induces a block partitioning of A_k, R_k and N_k when these matrices are expressed in Fourier basis. More precisely we write

$$A_k = \begin{pmatrix} \Lambda^{(k,1)} & & & \\ & \Lambda^{(k,2)} & & \\ & & \ddots & \\ & & & \Lambda^{(k,l_k)} \end{pmatrix}, \quad R_k = \begin{pmatrix} \Gamma^{(k,1)} & & & \\ & \Gamma^{(k,2)} & & \\ & & \ddots & \\ & & & \Gamma^{(k,l_k)} \end{pmatrix},$$
$$N_k^{(2\nu)} = \begin{pmatrix} \Sigma^{(k,1)} & & & \\ & \Sigma^{(k,2)} & & \\ & & \ddots & \\ & & & \Sigma^{(k,l_k)} \end{pmatrix}, \tag{2.6}$$

where $\Lambda^{(k,j)} = \mathrm{diag}(\lambda_1^{(k,j)}, \dots, \lambda_{m_j^{(k)}}^{(k,j)})$, $\Gamma^{(k,j)} = \mathrm{diag}(\gamma_1^{(k,j)}, \dots, \gamma_{m_j^{(k)}}^{(k,j)})$, and $\Sigma^{(k,j)} = \mathrm{diag}(\sigma_1^{(k,j)}, \dots, \sigma_{m_j^{(k)}}^{(k,j)})$.

Observe that the partitioning induced by (2.5) associates in a same block the different Fourier modes that, on the coarse grid, are mapped by P_{k-1}^T onto the same coarse Fourier mode. To develop our analysis, we do not need to enter the details about which modes are associated. It is important to note, however, that in usual setting of Fourier analysis (see, e.g., [20,23]), inside each set of associated modes, there is a unique mode classified as "low frequency", all other modes being labelled as "high frequency". Then, the smoothing factor is the worst factor by which high frequency components are reduced per relaxation step; that is,

[1] Here we exclude cases for which the smoother is block diagonal as, e.g., when using red-black Gauss–Seidel relaxation for 5-point discretizations of Poisson equation [20, Section 4.5].

[2] Since P_{k-1} has full rank, see general assumptions.

$$\widetilde{\mu}^{(k)} = \max_{j=1,\ldots,l_k} \max_{\substack{i=1,\ldots,m_j^{(k)} \\ i \text{ is a "high frequency" mode}}} \left| 1 - \gamma_i^{(k,j)\,-1} \lambda_i^{(k,j)} \right|.$$

In our study, we assume that the ordering inside each block is such that

$$\left|1 - \gamma_1^{(k,j)\,-1} \lambda_1^{(k,j)}\right| \geq \left|1 - \gamma_2^{(k,j)\,-1} \lambda_2^{(k,j)}\right| \geq \cdots \geq \left|1 - \gamma_{m_j^{(k)}}^{(k,j)\,-1} \lambda_{m_j^{(k)}}^{(k,j)}\right| \tag{2.7}$$

and report results with respect to

$$\begin{aligned} \mu^{(k)} &= \max_{j=1,\ldots,l_k} \max_{i=2,\ldots,m_j^{(k)}} \left|1 - \gamma_i^{(k,j)\,-1} \lambda_i^{(k,j)}\right| \\ &= \max_{j=1,\ldots,l_k} \left|1 - \gamma_2^{(k,j)\,-1} \lambda_2^{(k,j)}\right|. \end{aligned} \tag{2.8}$$

Clearly, $\mu^{(k)}$ coincides with the classical smoothing factor if, inside each block the low frequency component is also the one less efficiently relaxed by the smoother. This corresponds to usual situations, but may be not true in whole generality. We also note that one has $\mu^{(k)} \leq \widetilde{\mu}^{(k)}$ if there is no more than one low frequency mode per block. In the following, we call $\mu^{(k)}$ the smoothing factor without checking further if $\mu^{(k)} = \widetilde{\mu}^{(k)}$.

Eventually, observe that (2.2) implies $1 - \sigma_i^{(k,j)}\lambda_i^{(k,j)} = (1 - \gamma_i^{(k,j)\,-1} \lambda_i^{(k,j)})^{2\nu}$ and hence (2.7) is equivalent to

$$\sigma_1^{(k,j)}\lambda_1^{(k,j)} \leq \sigma_2^{(k,j)}\lambda_2^{(k,j)} \leq \cdots \leq \sigma_{m_j^{(k)}}^{(k,j)}\lambda_{m_j^{(k)}}^{(k,j)}.$$

Example Let the matrices $A_k, k = 0, \ldots, J$, be obtained from the finite difference discretization of

$$\begin{cases} u_{xx} = f \\ u(0) = u(1) = 0 \end{cases} \tag{2.9}$$

on the uniform grid of mesh size $h_k = (n_k + 1)^{-1}$, where $n_k + 1 = 2(n_{k-1} + 1)$ and $n_0 > 0$. The corresponding stencil is given by

$$\frac{1}{h_k^2}[-1 \quad 2 \quad -1]. \tag{2.10}$$

Let P_{k-1} be defined by the stencil

$$\frac{1}{2\sqrt{2}}[1 \quad 2 \quad 1],$$

which gives (up to a constant factor) the linear interpolation operator and satisfies $A_k = P_k^T A_{k+1} P_k$. The vectors of the Fourier bases (both fine and coarse) are obtained by evaluating the function

$$v_j = \sin(j\pi x) \tag{2.11}$$

at the grid points; the resulting vectors are then normalized. In the Fourier basis, the matrices P_{k-1} and A_k can be checked to have the form (2.5), (2.6), where

$$\begin{aligned}
\mathbf{p}_j^{(k-1)} &= \begin{pmatrix} 1+\cos(\theta_j^{(k)}) \\ -1+\cos(\theta_j^{(k)}) \end{pmatrix}, \\
\Lambda^{(k,j)} &= \frac{1}{h_k^2}\begin{pmatrix} 2-2\cos(\theta_j^{(k)}) & \\ & 2+2\cos(\theta_j^{(k)}) \end{pmatrix},
\end{aligned} \tag{2.12}$$

with $\theta_j^{(k)} = j\pi h_k$, $j = 1, \dots, n_{k-1}$, $k = 1, \dots, J$. Note that A_k has one additional block $\Lambda^{(k,n_{k-1}+1)} = 2h_k^{-2}$, but the corresponding Fourier mode is not represented on the coarse level (and has no associated vector $\mathbf{p}$). Although such situation is not explicitly allowed by (2.5), (2.6), this may be cured by merging this block with any other block of size two. The corresponding vector $\mathbf{p}$ will then have two non-zero and one zero entry.

One may further note that the weighted Jacobi smoother $R_k = \omega_{Jac}^{-1}\mathrm{diag}(A)$ has the same expression in both nodal and Fourier bases, satisfying (2.6) with $\Gamma^{(k,j)} = 2h_k^{-2}\omega_{Jac}^{-1} I$. The weighting factor should satisfy $0 < \omega_{Jac} < 1$ for our general assumptions to hold.

Eventually, observe that a common choice for "low frequency" modes at level k corresponds to (2.11) for $j = 1, \dots, n_{k-1}$; they are represented by the first row/column entry in the blocks (2.12). These modes are indeed the least efficiently reduced by the smoother for all considered values of ω_{Jac}; moreover,

$$\widetilde{\mu}^{(k)} = \mu^{(k)} = \max(1-\omega_{Jac}, 2\omega_{Jac}-1).$$

3 V-cycle analysis and McCormick's bound

We recall here the bound obtained in [10, Lemma 2.3, Theorem 3.4 and Section 5] (see also [9,15] for an alternative proof). The equivalence of definition (3.2) with (3.3) is proved in [13].

Theorem 3.1 *Let $E_{MG}^{(J)}$ and $N_k^{(\nu)}$, $k = 1, \dots, J$, be defined, respectively, by (2.3) and (2.1) with A being SPD and with P_k, $k = 0, \dots, J-1$, A_k, $k = 0, \dots, J$, and R_k, $k = 1, \dots, J$, satisfying the general assumptions stated in Sect. 2.*

Then, letting $\pi_{A_k} = P_{k-1} A_{k-1}^{-1} P_{k-1}^T A_k$, there holds

$$\rho(E_{MG}^{(J)}) \le 1 - \min_{1\le k\le J} \delta_k^{(\nu)}, \tag{3.1}$$

where

$$\delta_k^{(\nu)} = \min_{\mathbf{v}_k \in \mathbb{R}^{n_k}} \frac{\|\mathbf{v}_k\|_{A_k}^2 - \|(I - N_k^{(\nu)} A_k)\mathbf{v}_k\|_{A_k}^2}{\|(I - \pi_{A_k})\mathbf{v}_k\|_{A_k}^2} \tag{3.2}$$

$$= \min_{\mathbf{v}_k \in \mathbb{R}^{n_k}} \frac{\mathbf{v}_k^T N_k^{(2\nu)} \mathbf{v}_k}{\mathbf{v}_k^T (A_k^{-1} - P_{k-1} A_{k-1}^{-1} P_{k-1}^T)\mathbf{v}_k}. \tag{3.3}$$

Moreover,

$$\delta_k^{(\nu)\,-1} \le \frac{1}{\nu}\left(\delta_k^{(1)\,-1} + \nu - 1\right). \tag{3.4}$$

As already mentioned, it is shown in [13] that the McCormick's bound is the best bound for V-cycle multigrid among those characterized by a constant which is a maximum over all levels of an expression involving only two consecutive levels at a time. This latter feature is the key property that allows us, in the next section, to assess the bound in standard Fourier analysis setting, and relate it to the smoothing factor.

4 Smoothing factor and actual multigrid convergence

Let $\widetilde{A}_k = N_k^{(2\nu)\,1/2} A_k N_k^{(2\nu)\,1/2}$ and $\widetilde{P}_{k-1} = N_k^{(2\nu)\,-1/2} P_{k-1}$, with corresponding block structure

$$\widetilde{A}_k = \begin{pmatrix} \widetilde{\Lambda}^{(k,1)} & & & \\ & \widetilde{\Lambda}^{(k,2)} & & \\ & & \ddots & \\ & & & \widetilde{\Lambda}^{(k,l_k)} \end{pmatrix}, \quad \widetilde{P}_{k-1} = \begin{pmatrix} \widetilde{\mathbf{p}}_1^{(k-1)} & & & \\ & \widetilde{\mathbf{p}}_2^{(k-1)} & & \\ & & \ddots & \\ & & & \widetilde{\mathbf{p}}_{l_k-1}^{(k-1)} \end{pmatrix}, \tag{4.1}$$

where $\widetilde{\Lambda}^{(k,j)} = \mathrm{diag}(\widetilde{\lambda}_i^{(k,j)})$ with $\widetilde{\lambda}_i^{(k,j)} = \sigma_i^{(k,j)} \lambda_i^{(k,j)}$. Setting

$$\widetilde{\pi}^{(k,j)} = \widetilde{\mathbf{p}}_j^{(k-1)} (\widetilde{\mathbf{p}}_j^{(k-1)\,H} \widetilde{\Lambda}^{(k,j)} \widetilde{\mathbf{p}}_j^{(k-1)})^{-1} \widetilde{\mathbf{p}}_j^{(k-1)\,H} \widetilde{\Lambda}^{(k,j)},$$

there holds

$$\begin{aligned}
\rho(E_{TG}^{(k)}) &= \rho\left(\left(I - P_{k-1} A_{k-1}^{-1} P_{k-1}^T A_k\right)(I - N_k^{(\nu)} A_k)^2\right) \\
&= \rho\left(\left(I - P_{k-1} A_{k-1}^{-1} P_{k-1}^T A_k\right)(I - N_k^{(2\nu)} A_k)\right) \\
&= \rho\left((I - \widetilde{P}_{k-1} \widetilde{A}_{k-1}^{-1} \widetilde{P}_{k-1}^T \widetilde{A}_k)(I - \widetilde{A}_k)\right) \\
&= \max_{j=1,\dots,l_k} \rho\left((I - \widetilde{\pi}^{(k,j)})(I - \widetilde{\Lambda}^{(k,j)})\right),
\end{aligned}$$

and

$$\begin{aligned}\delta_k^{(\nu)\,-1} &= \max_{\mathbf{v}_k\in\mathbb{R}^{n_k}} \frac{\mathbf{v}_k^T(A_k^{-1}-P_{k-1}A_{k-1}^{-1}P_{k-1}^T)\mathbf{v}_k}{\mathbf{v}_k^T\,N_k^{(2\nu)}\,\mathbf{v}_k}\\ &= \max_{\mathbf{v}_k\in\mathbb{R}^{n_k}} \frac{\mathbf{v}_k^T(I-\widetilde{P}_{k-1}\widetilde{A}_{k-1}^{-1}\widetilde{P}_{k-1}^T\widetilde{A}_k)\widetilde{A}_k^{-1}\mathbf{v}_k}{\mathbf{v}_k^T\mathbf{v}_k}\\ &= \max_{j=1,\dots,l_k}\ \rho\left((I-\widetilde{\pi}^{(k,j)})\ \widetilde{\Lambda}^{(k,j)\,-1}\right).\end{aligned}$$

Now, for each individual block, the quantities one has to take the maximum of may be assessed by applying the following lemma with $\widetilde{\Lambda}=\widetilde{\Lambda}^{(k,j)}$ and $\widetilde{\mathbf{p}}=\widetilde{\mathbf{p}}_j^{(k-1)}$. Observe that the assumption $0<\widetilde{\lambda}_i\le 1$ is then not restrictive since $I-\widetilde{A}_k$ and $(I-N_k^{(\nu)}A_k)^2$ have the same spectra, and hence the eigenvalues of $\widetilde{\Lambda}^{(k,j)}$, being a subset of the eigenvalues of $\widetilde{A}_k$, belong to (0, 1] by virtue of our general assumptions.

Lemma 4.1 *Let $\widetilde{\Lambda}=\mathrm{diag}(\widetilde{\lambda}_i)$ be a $m\times m$ real matrix with $0<\widetilde{\lambda}_1\le\widetilde{\lambda}_2\le\ \cdots\le\widetilde{\lambda}_m\le 1$, and let $\widetilde{\mathbf{p}}=(\widetilde{p}_1\ \ \dots\ \ \widetilde{p}_m)^T$ be a non-zero complex vector. Set*

$$\begin{aligned}\widetilde{\pi}&=\widetilde{\mathbf{p}}\,(\widetilde{\mathbf{p}}^H\widetilde{\Lambda}\widetilde{\mathbf{p}})^{-1}\,\widetilde{\mathbf{p}}^H\widetilde{\Lambda},\\ \rho_{TG}&=\rho\left((I_m-\widetilde{\pi})\ \left(I_m-\widetilde{\Lambda}\right)\right),\end{aligned}$$

and

$$\delta^{-1}=\rho\left((I_m-\widetilde{\pi})\ \widetilde{\Lambda}^{-1}\right).$$

Letting

$$\widetilde{\alpha}=\sum_{i=2}^m\frac{\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{\widetilde{\lambda}_1^2\|\widetilde{\mathbf{p}}\|^2},$$

there holds

$$\frac{\widetilde{\lambda}_2}{1+\widetilde{\alpha}(1-\widetilde{\lambda}_1/\widetilde{\lambda}_2)}\ \le\ \delta\ \le\ \left(\frac{4}{\widetilde{\alpha}}\right)^{1/3}. \tag{4.2}$$

Moreover, if $|\widetilde{p}_1|>0$, letting

$$\widetilde{\beta}=\sum_{i=2}^m\frac{\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{\widetilde{\lambda}_1^2|\widetilde{p}_1|^2},$$

there holds

$$\widetilde{\lambda}_1+\frac{\widetilde{\lambda}_2-\widetilde{\lambda}_1}{1+\widetilde{\beta}}\ \le\ \delta\ \le\ \widetilde{\lambda}_1+\frac{\widetilde{\lambda}_m-\widetilde{\lambda}_1}{1+\widetilde{\beta}} \tag{4.3}$$

and

$$\widetilde{\lambda}_1 + \frac{\widetilde{\lambda}_2 - \widetilde{\lambda}_1}{1 + \widetilde{\lambda}_2^{-1}\widetilde{\lambda}_1\widetilde{\beta}} \;\leq\; 1 - \rho_{TG} \;\leq\; \min\left(\widetilde{\lambda}_1 + \frac{\widetilde{\lambda}_m - \widetilde{\lambda}_1}{1 + \widetilde{\lambda}_m^{-1}\widetilde{\lambda}_1\widetilde{\beta}},\; \widetilde{\lambda}_2\right), \tag{4.4}$$

whereas, if $|\widetilde{p}_1| = 0$,

$$\delta \;=\; 1 - \rho_{TG} \;=\; \widetilde{\lambda}_1. \tag{4.5}$$

Proof. Set $\widetilde{\lambda}_c \;= \widetilde{\mathbf{p}}^H \widetilde{\Lambda}\widetilde{\mathbf{p}} \;= \sum_{i=1}^m \widetilde{\lambda}_i |\widetilde{p}_i|^2$. First, observe that, according to Lemma 2.2 in [11],

$$\begin{aligned} &\widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}\, \widetilde{\lambda}_m^{-1} + (1 - \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}) \widetilde{\lambda}_1^{-1} \\ &\quad \leq \delta^{-1} \;\leq\; \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}\, \widetilde{\lambda}_2^{-1} + (1 - \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}) \widetilde{\lambda}_1^{-1}, \end{aligned} \tag{4.6}$$

and, similarly,

$$\begin{aligned} &\widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}\, \eta_m + (1 - \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}) \eta_1 \\ &\quad \leq \rho_{TG} \leq \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}\, \eta_2 + (1 - \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}) \eta_1, \end{aligned} \tag{4.7}$$

where $\eta_i = 1 - \widetilde{\lambda}_i$. Equality (4.5) then readily follows. Moreover, (4.6) and (4.7) can be further rewritten as

$$\widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1} \left(\widetilde{\lambda}_m^{-1} - \widetilde{\lambda}_1^{-1}\right) + \widetilde{\lambda}_1^{-1} \leq \delta^{-1} \leq \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1} \left(\widetilde{\lambda}_2^{-1} - \widetilde{\lambda}_1^{-1}\right) + \widetilde{\lambda}_1^{-1}, \tag{4.8}$$

$$\widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}\, (\eta_m - \eta_1) + \eta_1 \leq \rho_{TG} \leq \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1}\, (\eta_2 - \eta_1) + \eta_1. \tag{4.9}$$

We now prove the inequalities (4.3) and (4.4). Note that

$$\widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1} = \frac{\widetilde{\lambda}_1 |\widetilde{p}_1|^2}{\sum_{i=1}^m \widetilde{\lambda}_i |\widetilde{p}_i|^2} = \left(1 + \widetilde{\lambda}_1 \frac{\sum_{i=2}^m \widetilde{\lambda}_i |\widetilde{p}_i|^2}{\widetilde{\lambda}_1^2 |\widetilde{p}_1|^2}\right)^{-1}$$

implies

$$(1 + \xi_2)^{-1} \leq \widetilde{\lambda}_1 |\widetilde{p}_1|^2 \widetilde{\lambda}_c^{-1} \leq (1 + \xi_m)^{-1},$$

where $\xi_i = \widetilde{\beta}\widetilde{\lambda}_1 / \widetilde{\lambda}_i$. Using these last inequalities in (4.8) and (4.9) one obtains (since $\widetilde{\lambda}_1^{-1} \geq \cdots \geq \widetilde{\lambda}_m^{-1}$ and $\eta_1 \geq \cdots \geq \eta_m$)

$$\begin{aligned} \frac{1}{\xi_m + 1}\left(\widetilde{\lambda}_m^{-1} - \widetilde{\lambda}_1^{-1}\right) + \widetilde{\lambda}_1^{-1} \leq \delta^{-1} \;&\leq\; \frac{1}{\xi_2 + 1}\left(\widetilde{\lambda}_2^{-1} - \widetilde{\lambda}_1^{-1}\right) + \widetilde{\lambda}_1^{-1}, \\ \frac{1}{\xi_m + 1}(\eta_m - \eta_1) + \eta_1 \leq \rho_{TG} \;&\leq\; \frac{1}{\xi_2 + 1}(\eta_2 - \eta_1) + \eta_1. \end{aligned}$$

Hence, using $\xi_i = \widetilde{\beta}\,\widetilde{\lambda}_1/\widetilde{\lambda}_i$ and $\eta_i = 1 - \widetilde{\lambda}_i$, $i = 2, m$, we have

$$\begin{aligned}\frac{1+\widetilde{\beta}}{\widetilde{\lambda}_m + \widetilde{\lambda}_1\widetilde{\beta}} \le \delta^{-1} &\le \frac{1+\widetilde{\beta}}{\widetilde{\lambda}_2 + \widetilde{\lambda}_1\widetilde{\beta}},\\ 1 - \frac{\widetilde{\lambda}_m^2 + \widetilde{\lambda}_1^2\widetilde{\beta}}{\widetilde{\lambda}_m + \widetilde{\lambda}_1\widetilde{\beta}} \le \rho_{TG} &\le 1 - \frac{\widetilde{\lambda}_2^2 + \widetilde{\lambda}_1^2\widetilde{\beta}}{\widetilde{\lambda}_2 + \widetilde{\lambda}_1\widetilde{\beta}}.\end{aligned}$$

The inequalities (4.3) and (4.4) (except the second term in the minimum) readily follow. To conclude the proof of (4.4), let $\mathbf{e}_k$ be the kth unit vector and set

$$\mathbf{v} = \begin{cases} \mathbf{e}_2 & \text{if } \widetilde{\mathbf{p}}^H\mathbf{e}_2 = 0\\ \mathbf{e}_1 - \mathbf{e}_2\left(\dfrac{\widetilde{\lambda}_1{}^{1/2}\,\widetilde{\mathbf{p}}^H\mathbf{e}_1}{\widetilde{\lambda}_2{}^{1/2}\,\widetilde{\mathbf{p}}^H\mathbf{e}_2}\right) & \text{otherwise.}\end{cases}$$

Note that $\widetilde{\pi}\ \widetilde{\Lambda}^{-1/2}\,\mathbf{v} = \mathbf{0}$ and $\widetilde{\Lambda}\ \widetilde{\pi} = \widetilde{\pi}^H\ \widetilde{\Lambda}$. Hence,

$$\begin{aligned}\rho_{TG} &= \rho\left(\ (I_m - \widetilde{\pi})^2\ \left(I_m - \widetilde{\Lambda}\right)\ \right)\\ &= \rho\left(\ (I_m - \widetilde{\pi})\ \left(I_m - \widetilde{\Lambda}\right)\ (I_m - \widetilde{\pi})\ \right)\\ &= \rho\left(\ \widetilde{\Lambda}^{1/2}\ (I_m - \widetilde{\pi})\ \left(I_m - \widetilde{\Lambda}\right)\ (I_m - \widetilde{\pi})\ \widetilde{\Lambda}^{-1/2}\right)\\ &= \rho\left(\ \widetilde{\Lambda}^{-1/2}\ (I_m - \widetilde{\pi})^H\ \widetilde{\Lambda}^{1/2}\ \left(I_m - \widetilde{\Lambda}\right)\ \widetilde{\Lambda}^{1/2}\ (I_m - \widetilde{\pi})\ \widetilde{\Lambda}^{-1/2}\right)\\ &\ge \frac{\mathbf{v}^H\widetilde{\Lambda}^{-1/2}\ (I_m - \widetilde{\pi})^H\ \widetilde{\Lambda}^{1/2}\ \left(I_m - \widetilde{\Lambda}\right)\ \widetilde{\Lambda}^{1/2}\ (I_m - \widetilde{\pi})\ \widetilde{\Lambda}^{-1/2}\mathbf{v}}{\mathbf{v}^H\mathbf{v}}\\ &= \frac{\mathbf{v}^H\left(I_m - \widetilde{\Lambda}\right)\mathbf{v}}{\mathbf{v}^H\mathbf{v}}\\ &\ge 1 - \widetilde{\lambda}_2.\end{aligned}$$

We next prove the left inequality (4.2). First observe that

$$\widetilde{\lambda}_1|\widetilde{p}_1|^2\widetilde{\lambda}_c^{-1} = \frac{\widetilde{\lambda}_1|\widetilde{p}_1|^2}{\sum_{i=1}^m \widetilde{\lambda}_i|\widetilde{p}_i|^2} = 1 - \frac{\sum_{i=2}^m \widetilde{\lambda}_i|\widetilde{p}_i|^2}{\sum_{i=1}^m \widetilde{\lambda}_i|\widetilde{p}_i|^2} \ge 1 - \frac{\sum_{i=2}^m \widetilde{\lambda}_i|\widetilde{p}_i|^2}{\widetilde{\lambda}_1\sum_{i=1}^m |\widetilde{p}_i|^2} \ge 1 - \frac{\widetilde{\lambda}_1\widetilde{\alpha}}{\widetilde{\lambda}_2},$$

and hence, with (4.8), there holds

$$\delta^{-1} \le \left(1 - \frac{\widetilde{\lambda}_1\widetilde{\alpha}}{\widetilde{\lambda}_2}\right)\left(\widetilde{\lambda}_2^{-1} - \widetilde{\lambda}_1^{-1}\right) + \widetilde{\lambda}_1^{-1} = \widetilde{\lambda}_2^{-1}\left(1 + \widetilde{\alpha}\left(1 - \frac{\widetilde{\lambda}_1}{\widetilde{\lambda}_2}\right)\right).$$

It remains to prove the right inequality (4.2). Note that, according to Theorem 3.2 in [14],

$$\delta \le \min\left(\ 1 - \rho_{TG},\ \ \|\widetilde{\pi}\|^{-2}\ \right).$$

Hence, provided that

$$\widetilde{\alpha} \leq \frac{4}{(1-\rho_{TG})^2}\|\widetilde{\pi}\|^2 \tag{4.10}$$

holds (we prove it below), we have

$$\delta \leq \min\left(1-\rho_{TG},\ \frac{1}{\widetilde{\alpha}}\frac{4}{(1-\rho_{TG})^2}\right) \leq \max_{x>0}\ \min\left(x,\ \frac{1}{\widetilde{\alpha}}\frac{4}{x^2}\right) \leq \left(\frac{4}{\widetilde{\alpha}}\right)^{1/3}.$$

We are thus left with the proof of (4.10), for which we use

$$\|\widetilde{\pi}\|^2 = \rho\left(\widetilde{\mathbf{p}}\,\widetilde{\lambda}_c^{-1}\,\widetilde{\mathbf{p}}^H\widetilde{\Lambda}^2\widetilde{\mathbf{p}}\,\widetilde{\lambda}_c^{-1}\,\widetilde{\mathbf{p}}^H\right) = \frac{\|\widetilde{\mathbf{p}}\|^2\,\widetilde{\mathbf{p}}^H\widetilde{\Lambda}^2\widetilde{\mathbf{p}}}{\widetilde{\lambda}_c^2} = \|\widetilde{\mathbf{p}}\|^2\,\frac{\sum_{i=1}^m\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{\widetilde{\lambda}_c^2}. \tag{4.11}$$

According to (4.9), we have, using $\eta_i = 1-\widetilde{\lambda}_i,\ i=1,m$,

$$\rho_{TG} \geq 1-\widetilde{\lambda}_1-\widetilde{\lambda}_1|\widetilde{p}_1|^2\widetilde{\lambda}_c^{-1}\,(\widetilde{\lambda}_m-\widetilde{\lambda}_1).$$

Hence, considering first the case where $\widetilde{\lambda}_1 \leq \frac{1-\rho_{TG}}{2}$ and $\widetilde{\lambda}_1 \neq \widetilde{\lambda}_m$, there holds

$$\widetilde{\lambda}_1|\widetilde{p}_1|^2\widetilde{\lambda}_c^{-1} \geq \frac{1-\widetilde{\lambda}_1-\rho_{TG}}{\widetilde{\lambda}_m-\widetilde{\lambda}_1} \geq \frac{1-\rho_{TG}}{2(\widetilde{\lambda}_m-\widetilde{\lambda}_1)} \geq \frac{1-\rho_{TG}}{2}, \tag{4.12}$$

the last inequality following from $0 \leq \widetilde{\lambda}_1 < \widetilde{\lambda}_m \leq 1$. Note that (4.12) implies $|\widetilde{p}_1|^2 > 0$. The right inequality (4.10) follows then from

$$\widetilde{\alpha} = \sum_{i=2}^m\frac{\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{\widetilde{\lambda}_1^2\|\widetilde{\mathbf{p}}\|^2} \leq \|\widetilde{\mathbf{p}}\|^2\frac{\sum_{i=1}^m\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{(\widetilde{\lambda}_1|\widetilde{p}_1|^2)^2} \leq \frac{4}{(1-\rho_{TG})^2}\|\widetilde{\mathbf{p}}\|^2\frac{\sum_{i=1}^m\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{\widetilde{\lambda}_c^2},$$

together with (4.11). If $\widetilde{\lambda}_1 = \widetilde{\lambda}_m$, we have

$$\widetilde{\alpha} = \sum_{i=2}^m\frac{|\widetilde{p}_i|^2}{\|\widetilde{\mathbf{p}}\|^2} \leq 1 \leq \frac{4}{(1-\rho_{TG})^2}\|\widetilde{\pi}\|^2,$$

the last inequality coming from the fact that $\widetilde{\pi}$ is a projector, and hence $\|\widetilde{\pi}\| \geq 1$. On the other hand, when $\widetilde{\lambda}_1 \geq \frac{1-\rho_{TG}}{2}$ one has (since $\widetilde{\lambda}_i \leq 1$)

$$\widetilde{\alpha} = \sum_{i=2}^m\frac{\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{\widetilde{\lambda}_1^2\|\widetilde{\mathbf{p}}\|^2} < \frac{4}{(1-\rho_{TG})^2}\sum_{i=2}^m\frac{\widetilde{\lambda}_i^2|\widetilde{p}_i|^2}{\|\widetilde{\mathbf{p}}\|^2} \leq \frac{4}{(1-\rho_{TG})^2} \leq \frac{4}{(1-\rho_{TG})^2}\|\widetilde{\pi}\|^2,$$

the last inequality coming from $\|\widetilde{\pi}\|^2 \geq 1$. □

This lemma can be applied in the context of Fourier analysis, setting $\widetilde{\Lambda} = \widetilde{\Lambda}^{(k,j)}$ and $\widetilde{\mathbf{p}} = \widetilde{\mathbf{p}}_j^{(k-1)}$, where $\widetilde{\Lambda}^{(k,j)}$ and $\widetilde{\mathbf{p}}_j^{(k-1)}$ come from the block representation of $\widetilde{A}_k = N_k^{(2\nu)\,1/2} A_k\, N_k^{(2\nu)\,1/2}$ and $\widetilde{P}_{k-1} = N_k^{(2\nu)\,-1/2}\, P_{k-1}$. Hence, the main constants for block j at level k are

$$\widetilde{\alpha}_\nu^{(k,j)} = \sum_{i=2}^{m_j^{(k)}} \frac{\widetilde{\lambda}_i^{(k,j)\,2}\; |(\widetilde{\mathbf{p}}_j^{(k-1)})_i|^2}{\widetilde{\lambda}_1^{(k,j)\,2}\; \|\widetilde{\mathbf{p}}_j^{(k-1)}\|^2} = \frac{\sum_{i=2}^{m_j^{(k)}} \sigma_i^{(k,j)}\; \lambda_i^{(k,j)\,2}\; |(\mathbf{p}_j^{(k-1)})_i|^2}{\left(\sigma_1^{(k,j)}\lambda_1^{(k,j)}\right)^2 \sum_{i=1}^{m_j^{(k)}} \sigma_i^{(k,j)\,-1}\; |(\mathbf{p}_j^{(k-1)})_i|^2}$$

and

$$\widetilde{\beta}_\nu^{(k,j)} = \sum_{i=2}^{m_j^{(k)}} \frac{\widetilde{\lambda}_i^{(k,j)\,2}\; |(\widetilde{\mathbf{p}}_j^{(k-1)})_i|^2}{\widetilde{\lambda}_1^{(k,j)\,2}\; |(\widetilde{\mathbf{p}}_j^{(k-1)})_1|^2} = \sum_{i=2}^{m_j^{(k)}} \frac{\sigma_i^{(k,j)}\; \lambda_i^{(k,j)\,2}\; |(\mathbf{p}_j^{(k-1)})_i|^2}{\sigma_1^{(k,j)}\; \lambda_1^{(k,j)\,2}\; |(\mathbf{p}_j^{(k-1)})_1|^2},$$

where we use subscript ν to recall that these quantities inherit the dependence of $\sigma_i^{(k,j)}$ on the number of smoothing steps. Taking all blocks into account, we also set

$$\widetilde{\alpha}_\nu^{(k)} = \max_{j=0,\dots,l_k} \widetilde{\alpha}_\nu^{(k,j)}. \tag{4.13}$$

Since our definition (2.8) of the smoothing factor entails

$$\min_{j=1,\dots,l_k} \widetilde{\lambda}_2^{(k,j)} = 1 - \left(\mu^{(k)}\right)^{2\nu}, \tag{4.14}$$

one sees with (4.2) that, when the smoothing factor is away from 1, $\delta_k^{(\nu)}$ remains away from zero if *and only if* $\widetilde{\alpha}_\nu^{(k)}$ is reasonably bounded.

Note that it makes little sense to define in a similar way a "global" $\widetilde{\beta}_\nu^{(k)}$ since, considering (4.3) and (4.4), the magnitude of the parameter $\widetilde{\beta}_\nu^{(k,j)}$ plays an important role only in the blocks for which $\widetilde{\lambda}_1^{(k,j)}$ is close to zero, but may be arbitrary large without dramatic consequences within the other blocks.

Going back to (4.2) and taking all levels into account, we can further state the following theorem.

Theorem 4.1 *Let $E_{MG}^{(J)}$, $E_{TG}^{(J)}$, $\mu^{(k)}$ and $\widetilde{\alpha}_\nu^{(k)}$, $k = 1, \dots, J$, be defined by* (2.3), (2.4), (2.8) *and* (4.13), *respectively, with A_k, $k = 0, \dots, J$, R_k, $k = 1, \dots, J$, and P_k, $k = 0, \dots, J-1$, having the form* (2.5), (2.6) *and satisfying the assumptions stated in Sect.* 2, *and with $N_k^{(\nu)}$, $\widetilde{\alpha}_\nu^{(k,j)}$, $k = 1, \dots, J$, $j = 1, \dots, l_k$, being defined by* (2.1), (4.1), *respectively.*

Then

$$
\begin{aligned}
(\mu^{(J)})^{2\nu} &\le \rho(E_{TG}^{(J)}) \le \rho(E_{MG}^{(J)}) \\
&\le 1 - \min_{1\le k\le J} \delta_k^{(\nu)} \le \max_{1\le k\le J} \frac{\left(\mu^{(k)}\right)^{2\nu} + \widetilde{\alpha}_\nu^{(k)}}{1+\widetilde{\alpha}_\nu^{(k)}}
\end{aligned}
\tag{4.15}
$$

and

$$
\begin{aligned}
(\mu^{(J)})^{2\nu} &\le \rho(E_{TG}^{(J)}) \le \rho(E_{MG}^{(J)}) \le 1 - \min_{1\le k\le J} \delta_k^{(\nu)} \\
&\le \max_{1\le k\le J} \frac{\delta_k^{(1)\,-1} - 1}{\delta_k^{(1)\,-1} - 1 + \nu} \\
&\le \max_{1\le k\le J} \frac{\left(\mu^{(k)}\right)^2 + \widetilde{\alpha}_1^{(k)}}{\left(\mu^{(k)}\right)^2 + \widetilde{\alpha}_1^{(k)} + \nu\left(1-\left(\mu^{(k)}\right)^2\right)}.
\end{aligned}
\tag{4.16}
$$

Proof. The inequality (4.15) is obtained applying successively the right inequality (4.4) (with second term in the minimum) and (4.14), the results in [7, Section 7.2] (for the proof of $\rho(E_{TG}^{(J)}) \le \rho(E_{MG}^{(J)})$), the inequality (3.1), and the left inequality (4.2) (with $1-\widetilde{\lambda}_1/\widetilde{\lambda}_2$ bounded above by 1). The inequality (4.16) is obtained using the same arguments except that (3.1) is further combined with the inequality (3.4). □

This theorem can be considered as complementary to our previous work [14]. There, we show that optimal two-grid convergence implies optimal V-cycle convergence if the norm of the (exact) coarse grid correction operator is bounded at each level. Here, the comparison between the two-grid and the V-cycle convergence rates is further extended to the smoothing factor. Consider in particular the inequalities (4.15) and observe that if $\max_{k=1,\dots,J} \mu^{(k)} \approx \mu^{(J)}$ (which often holds in practice) and if $\widetilde{\alpha}_\nu^{(k)}$ is nicely bounded at each level, these inequalities define a narrow interval containing both the two-grid and V-cycle multigrid convergence factors. On the other hand, if $\widetilde{\alpha}_\nu^{(k)}$ is large at some levels, the right inequality (4.15) becomes ineffective, and the right inequality (4.2) further shows that $1-\min_{1\le k\le J} \delta_k^{(\nu)}$ will be indeed close to 1. As observed in [14], the actual convergence of the V-cycle may then scale poorly with the number of levels.

Example The values of the different terms in the inequalities (4.15) and (4.16) are given in Table 1 for the example introduced at the end of Sect. 2.

5 On the quality of the prolongation

In this and the next section, we derive simple criteria on the prolongation which, in combination with a nicely bounded smoothing factor, guarantee optimal convergence properties. We adopt notation similar to those in [20,23] and use a d-uple

Table 1 The values of the different terms in the inequalities (4.15) and (4.16) for the example with $\omega_{Jac} = 1/2$ and $J + 1 = 15$ grids with $n_0 = 2$ nodes on the coarsest grid (minima/maxima always correspond to $k = J$)

ν	$(\mu^{(J)})^{2\nu}$	$\rho(E_{TG}^{(J)})$	$\rho(E_{MG}^{(J)})$	$1-\delta_J^{(\nu)}$	$\frac{\left(\mu^{(J)}\right)^{2\nu}+\widetilde{\alpha}_\nu^{(J)}}{1+\widetilde{\alpha}_\nu^{(J)}}$	$\frac{\left(\mu^{(J)}\right)^{2}+\widetilde{\alpha}_1^{(J)}}{(\mu^{(J)})^2+\widetilde{\alpha}_1^{(J)}+\nu\left(1-(\mu^{(J)})^2\right)}$
1	0.250	0.250	0.275	0.333	0.537	0.537
2	0.063	0.083	0.126	0.200	0.375	0.367

$\boldsymbol{\theta} = (\theta_1, \theta_2, \dots \theta_d)$ instead of the block index j. Although in the context of (rigorous) Fourier analysis $\theta_i, i = 1, \dots d$, can take only a finite number of values (depending on the assumed mesh size), we follow here the common practice and allow all values inside a fixed domain $(-\pi, \pi)$. We also omit the grid index k since the discussion in this and in the following section does not depends on the choice of a particular grid.

Here our analysis highlights the critical role of the blocks for which $\widetilde{\lambda}_1^{(\boldsymbol{\theta})}$ is small. More precisely, we assume $\widetilde{\lambda}_1^{(\boldsymbol{\theta})}$ as a continuous function of $\boldsymbol{\theta}$ and focus on the values $\boldsymbol{\theta}_0$ of $\boldsymbol{\theta}$ for which $\lim_{\boldsymbol{\theta}\to\boldsymbol{\theta}_0} \widetilde{\lambda}_1^{(\boldsymbol{\theta})} = 0$. We base our discussion on $\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}$, since, considering (4.3) and (4.4), it allows to address at once δ_ν and $\rho(E_{TG})$. Note, for the sake of completeness, that, in the neighborhood of $\boldsymbol{\theta}_0$, $\widetilde{\alpha}_\nu^{(\boldsymbol{\theta})}$ has actually the same qualitative behavior as $\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}$ [12].

Proposition 5.1 *Let the assumption stated in Sect. 2 and in the first paragraph of Sect. 5 hold. Then*

(a) *the two-grid scheme has level-independent convergence properties iff the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from* 1 *and* $|\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}\widetilde{\lambda}_1^{(\boldsymbol{\theta})}|$ *remains bounded above in the neighborhood of any* $\boldsymbol{\theta}_0$ *such that* $\lim_{\boldsymbol{\theta}\to\boldsymbol{\theta}_0} \widetilde{\lambda}_1^{(\boldsymbol{\theta})} = 0$;
(b) *the parameter* $\delta^{(\nu)}$ *in McCormick's bound is bounded away from zero iff the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from* 1 *and* $|\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}|$ *remains bounded above in the neighborhood of any* $\boldsymbol{\theta}_0$ *such that* $\lim_{\boldsymbol{\theta}\to\boldsymbol{\theta}_0} \widetilde{\lambda}_1^{(\boldsymbol{\theta})} = 0$.

Proof. We give the proof of point (a) only; the proof of (b) follows the same lines. First, note that $1 \geq \widetilde{\lambda}_i^{(\boldsymbol{\theta})} \geq 1-\mu$, $i = 2, \dots, m_{\boldsymbol{\theta}}$, and, hence, μ bounded away from 1 implies $\widetilde{\lambda}_i^{(\boldsymbol{\theta})}$, $i = 2, \dots, m_{\boldsymbol{\theta}}$, bounded away from zero. It follows then from left (4.4) that the quantity $1 - \rho_{TG}$ is away from zero if $\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}\widetilde{\lambda}_1^{(\boldsymbol{\theta})}$ remains bounded above in the neighborhood of any $\boldsymbol{\theta}_0$ such that $\lim_{\boldsymbol{\theta}\to\boldsymbol{\theta}_0} \widetilde{\lambda}_1^{(\boldsymbol{\theta})} = 0$. Hence, the sufficient condition. The necessary condition follows from the right inequality (4.4), whose right hand side goes to zero if either $\widetilde{\lambda}_2^{(\boldsymbol{\theta})} \to 0$, or $\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}\widetilde{\lambda}_1^{(\boldsymbol{\theta})} \to \infty$ while with $\widetilde{\lambda}_1^{(\boldsymbol{\theta})} \to 0$. □

Now, the expression of $\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}$ is related to the smoother's Fourier components $\gamma_i^{(\boldsymbol{\theta})}$ only through $\sigma_i^{(\boldsymbol{\theta})}$, which is defined via

$$1 - \sigma_i^{(\boldsymbol{\theta})}\lambda_i^{(\boldsymbol{\theta})} = \left(1 - \gamma_i^{(\boldsymbol{\theta})\,-1}\,\lambda_i^{(\boldsymbol{\theta})}\right)^{2\nu}.$$

This does not make obvious a qualitative discussion. However, as shown in Lemma 5.1 below, $\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}$ has the same qualitative behavior as

$$\beta^{(\boldsymbol{\theta})} = \sum_{i=2}^{m_{\boldsymbol{\theta}}} \frac{\gamma_i^{(\boldsymbol{\theta})\ -1} \lambda_i^{(\boldsymbol{\theta})\ 2} |(\mathbf{p}_{\boldsymbol{\theta}})_i|^2}{\gamma_1^{(\boldsymbol{\theta})\ -1} \lambda_1^{(\boldsymbol{\theta})\ 2} |(\mathbf{p}_{\boldsymbol{\theta}})_1|^2}.$$

In this lemma, we need the technical assumption

$$\max_{1\le i\le m} \gamma_i^{-1}\lambda_i \le 2 - \min_{1\le i\le m} \gamma_i^{-1}\lambda_i. \tag{5.1}$$

In practice, this requires only a proper scaling of the smoother; for instance, if $\gamma_i^{\ -1} = \bar{\omega}\ \bar{\gamma}_i^{\ -1}$, injecting this relation into (5.2) yields the condition

$$\bar{\omega} \ \le\ \frac{2}{\max_{1\le s\le m}\ \bar{\gamma}_i^{\ -1}\,\lambda_i + \min_{1\le s\le m}\ \bar{\gamma}_i^{\ -1}\,\lambda_i}.$$

Lemma 5.1 *Let $\lambda_i > 0$ and $\gamma_i > 0, i = 1, \dots, m$ satisfy $|1 - \lambda_1\gamma_1^{-1}| \ge |1 - \lambda_2\gamma_2^{-1}| \ge \cdots \ge |1 - \lambda_m\gamma_m^{-1}|$ and define $\sigma_i^{(\nu)}$ for some integer $\nu > 0$ as*

$$1 - \sigma_i^{(\nu)}\lambda_i = (1 - \lambda_i\gamma_i^{-1})^{2\nu}, \quad i = 1, \dots, m. \tag{5.2}$$

Let $p_i, i = 1, \dots, m$, be such that $|p_1| > 0$ and set

$$\widetilde{\beta}_\nu = \sum_{i=2}^{m} \frac{\sigma_i^{(\nu)}\lambda_i^2|p_i|^2}{\sigma_1^{(\nu)}\lambda_1^2|p_1|^2} \quad \textit{and} \quad \beta = \sum_{i=2}^{m} \frac{\gamma_i^{-1}\ \lambda_i^2|p_i|^2}{\gamma_1^{-1}\ \lambda_1^2|p_1|^2}.$$

Then, one has

$$\frac{1}{\nu}\widetilde{\beta}_1 \le \widetilde{\beta}_\nu \le \widetilde{\beta}_1. \tag{5.3}$$

Further, if (5.1) holds, one has also

$$\frac{2-\omega}{2\nu}\beta \le \widetilde{\beta}_\nu \le \beta, \tag{5.4}$$

where $\omega = \max_{1\le i\le m} \gamma_i^{-1}\lambda_i$.

Moreover, letting

$$\phi = \arccos\left(\frac{|p_1|}{\sqrt{\sum_{i=1}^m |p_i|^2}}\right), \tag{5.5}$$

one has, for any μ such that $|1-\gamma_i^{-1}\lambda_i| \le \mu$, $i = 2, \dots, m$,

$$(1-\mu)^2 \frac{\min_{2\le i\le m} \gamma_i}{\gamma_1} \left(\frac{\tan\phi}{\gamma_1^{-1}\lambda_1}\right)^2 \le \beta \le \omega^2 \frac{\max_{2\le i\le m} \gamma_i}{\gamma_1} \left(\frac{\tan\phi}{\gamma_1^{-1}\lambda_1}\right)^2. \quad (5.6)$$

Proof. First, using (5.2) we observe that the assumed ordering is equivalent to $\lambda_1\sigma_1^{(\nu)} \le \lambda_2\sigma_2^{(\nu)} \le \cdots \le \lambda_m\sigma_m^{(\nu)}$ for any integer ν.

Then, observe that

$$\frac{\sigma_i^{(\nu)}}{\sigma_i^{(1)}} = \frac{1-(1-\gamma_i^{-1}\lambda_i)^{2\nu}}{1-(1-\gamma_i^{-1}\lambda_i)^2} = \frac{1-(1-s)^\nu}{s} = \sum_{k=0}^{\nu-1}(1-s)^k \quad (5.7)$$

with $s = 1-(1-\gamma_i^{-1}\lambda_i)^2 = \lambda_i\sigma_i^{(1)} \in [0,1]$, is a decreasing function of $\lambda_i\sigma_i^{(1)}$. Hence, since $\lambda_1\sigma_1^{(1)} \le \lambda_2\sigma_2^{(1)} \le \cdots \le \lambda_m\sigma_m^{(1)}$, one has

$$\sigma_i^{(\nu)} \le \frac{\sigma_1^{(\nu)}\sigma_i^{(1)}}{\sigma_1^{(1)}}, \quad i = 1, \dots, m.$$

The right inequality (5.3) straightforwardly follows. On the other hand, from (5.7) we also conclude that

$$1 \le \frac{\sigma_i^{(\nu)}}{\sigma_i^{(1)}} \le \nu;$$

hence the left inequality (5.3).

Moreover, when (5.1) holds, one has also $\min_{1\le i\le m} \gamma_i^{-1}\lambda_i \le 1$ and, hence,

$$\begin{aligned} |1 - \min_{1\le i\le m} \gamma_i^{-1}\lambda_i| &= 1 - \min_{1\le i\le m} \gamma_i^{-1}\lambda_i \\ &= \max\left(1 - \min_{1\le i\le m} \gamma_i^{-1}\lambda_i,\ \max_{1\le i\le m} \gamma_i^{-1}\lambda_i - 1\right) \\ &= \max_{1\le i\le m} |1-\gamma_i^{-1}\lambda_i|, \end{aligned}$$

where the second equality also follows from (5.1). Together with the assumed ordering, this implies $\gamma_1^{-1}\lambda_1 = \min_{1\le i\le m} \gamma_i^{-1}\lambda_i$. Therefore, since

$$\frac{\sigma_i^{(1)}}{\gamma_i^{-1}} = \frac{1-(1-\gamma_i^{-1}\lambda_i)^2}{\gamma_i^{-1}\lambda_i} = 2-\gamma_i^{-1}\lambda_i, \quad (5.8)$$

there holds

$$\sigma_i^{(1)} \leq \frac{\sigma_1^{(1)} \gamma_i^{-1}}{\gamma_1^{-1}}, \quad i = 1, \dots, m.$$

The right inequality (5.4) follows for $\nu = 1$. Further, from (5.8) we also have

$$2 - \omega \leq \frac{\sigma_i^{(1)}}{\gamma_i^{-1}} \leq 2,$$

which in turn implies the left inequality (5.4) for $\nu = 1$. On the other hand, (5.4) with $\nu = 1$ and (5.3) together imply (5.4) for any ν.

Eventually, for the proof of (5.6) we first note that $|1 - \gamma_i^{-1}\lambda_i| \leq \mu$, $i = 2, \dots, m$, and the definition $\omega = \max_{1 \leq i \leq m} \gamma_i^{-1}\lambda_i$ imply $1 - \mu \leq \gamma_i^{-1}\lambda_i \leq \omega$, $i = 2, \dots, m$. Hence since $|p_1| > 0$,

$$\left(\frac{1-\mu}{\gamma_1^{-1}\lambda_1}\right)^2 \sum_{i=2}^{m} \frac{\gamma_i |p_i|^2}{\gamma_1 |p_1|^2} \leq \beta \leq \left(\frac{\omega}{\gamma_1^{-1}\lambda_1}\right)^2 \sum_{i=2}^{m} \frac{\gamma_i |p_i|^2}{\gamma_1 |p_1|^2}.$$

The conclusion follows since

$$\tan^2 \phi = \frac{\sum_{i=2}^{m} |p_i|^2}{|p_1|^2}. \qquad \square$$

The last statement of the lemma is of particular interest when, as often arises, the smoother is well conditioned; that is, when all $\gamma_i^{(\boldsymbol{\theta})}$ are approximately equal (they are all equal for weighted Jacobi smoothing). Then $\beta_\nu^{(\boldsymbol{\theta})}$ and, hence, also $\widetilde{\beta}_\nu^{(\boldsymbol{\theta})}$, behave essentially like $\left(\frac{\tan\phi^{(\boldsymbol{\theta})}}{\gamma^{(\boldsymbol{\theta})}{}_1{}^{-1}\lambda_1^{(\boldsymbol{\theta})}}\right)^2$ where $\phi^{(\boldsymbol{\theta})}$ is the angle between the eigenvector associated to $\lambda_1^{(\boldsymbol{\theta})}$ and the range of the prolongation. This observations, together with points (a) and (b) of Proposition 5.1, lead to the following result.

Proposition 5.2 *Let the assumption stated in Sect. 2 and in the first paragraph of Sect. 5 hold. Assume in addition that the smoother is well conditioned; that is,* $\sup_{\boldsymbol{\theta}} \gamma^{(\boldsymbol{\theta})} / \inf_{\boldsymbol{\theta}} \gamma^{(\boldsymbol{\theta})} < c_\gamma$. *Then*

(a′) *the two-grid scheme has level-independent convergence properties iff the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from 1 and* $\left(\gamma_1^{(\boldsymbol{\theta})\,-1} \lambda_1^{(\boldsymbol{\theta})}\right)^{-\frac{1}{2}} \tan\phi^{(\boldsymbol{\theta})}$ *remains bounded above in the neighborhood of any* $\boldsymbol{\theta}_0$ *such that* $\lim_{\boldsymbol{\theta}\to\boldsymbol{\theta}_0} \widetilde{\lambda}_1^{(\boldsymbol{\theta})} = 0$;

(b′) *the parameter* $\delta^{(\nu)}$ *in McCormick's bound is bounded away from zero iff the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from 1 and* $\left(\gamma_1^{(\boldsymbol{\theta})\,-1} \lambda_1^{(\boldsymbol{\theta})}\right)^{-1} \tan\phi^{(\boldsymbol{\theta})}$ *remains bounded above in the neighborhood of any* $\boldsymbol{\theta}_0$ *such that* $\lim_{\boldsymbol{\theta}\to\boldsymbol{\theta}_0} \widetilde{\lambda}_1^{(\boldsymbol{\theta})} = 0$.

Some heuristics present in the multigrid literature [4, p. 1573] (see also [5, p. 4]) state qualitative rules about the prolongation, such as: “Interpolation must be able to approximate an eigenvector with error bound proportional to the size of the associated eigenvalue”. Our results give a more precise interpretation of such statements. For mere two-grid convergence the tangent of the angle between the eigenvector and the range of the prolongation should have a bound proportional to the square root of the eigenvalue, whereas guaranteed V-cycle convergence requires the bound to be proportional to the eigenvalue.

Example The example presented at the end of Sect.2 can be extended to the continuous case by making θ_j play the role of the continuous parameter $\theta \in (0, \pi/2)$ (possibly extend to $(-\pi, \pi)$). In this context, the mesh size h_k should be interpreted as a parameter independent of θ. Then, $\lambda_1^{(\theta)} = 0$ for $\theta_0 = 0$ (and possibly for $\theta_0 = \pm\pi$); further (for any such θ_0),

$$\lim_{\theta\to\theta_0} \left(\gamma_1^{(\theta)\,-1}\lambda_1^{(\theta)}\right)^{-\frac{1}{2}} \tan\phi^{(\theta)} = 0, \quad \lim_{\theta\to\theta_0} \left(\gamma_1^{(\theta)\,-1}\lambda_1^{(\theta)}\right)^{-1} \tan\phi^{(\theta)} = \frac{1}{2\omega_{Jac}}.$$

6 On the order of the prolongation

Here and in the rest of this section, we assume in addition that $\lambda_1^{(\boldsymbol{\theta})} = \lambda_1^{(\boldsymbol{\theta}+\pi\boldsymbol{\alpha})}$, $\boldsymbol{\alpha} \in \{0, 1\}^d$, and that the only zero of $\lambda_1^{(\boldsymbol{\theta})}$ on $[-\pi/2, \pi/2]^d$ is $\boldsymbol{\theta}_0 = \mathbf{0}$. These assumptions allow us to recover the framework of [8]. (The reasoning below can however be extended without difficulty to more general zero patterns of $\lambda_1^{(\boldsymbol{\theta})}$.) Since in what follows necessary and sufficient conditions are considered, we define the *upper order* $\overline{m}$ (resp. the *lower order* $\underline{m}$) of a differential operator associated to the problem as the smallest (resp. the largest) constant such that

$$c\,|\boldsymbol{\theta}|^{\overline{m}} \le \lambda_1^{(\boldsymbol{\theta})} \le C\,|\boldsymbol{\theta}|^{\underline{m}} \quad \text{for} \quad |\boldsymbol{\theta}| \to 0. \tag{6.1}$$

Note that upper order $\overline{m}$ corresponds to the usual order definition (as, e.g., given in [8]) and that $\overline{m} = \underline{m}$ in usual applications (that is, for PDEs with the same largest order of derivatives along principal directions). Since we assume $\lambda_1^{(\boldsymbol{\theta}_0)} = 0$, only positive $\overline{m}$, $\underline{m}$ can be considered.[3] Further, we define the *frequency based* (FB) *order* $m_{FB}(P)$ of the prolongation P as the largest constant such that

$$\frac{|(\mathbf{p}^{(\boldsymbol{\theta})})_i|}{|(\mathbf{p}^{(\boldsymbol{\theta})})_1|} = \mathcal{O}(|\boldsymbol{\theta}|^{m_{FB}(P)}) \quad \text{for} \quad |\boldsymbol{\theta}| \to 0 \quad \text{and for} \quad i = 2, \ldots, m_{\boldsymbol{\theta}}. \tag{6.2}$$

[3] This is not however a fundamental limitation, since, for $\overline{m} = \underline{m} = 0$, the two-grid scheme has optimal convergence properties regardless the prolongation; the case $\overline{m} > \underline{m} = 0$ is outside the scope of our discussion since $\lambda_1^{(\boldsymbol{\theta})}$ would then be discontinuous, which contradicts the assumptions in Sect. 5.

The following proposition can now be stated.

Proposition 6.1 *Let the assumption stated in Sect. 2 and in the first paragraph of Sects. 5 and 6 hold. Assume in addition that the smoother is well conditioned; that is,* $\sup_{\boldsymbol{\theta}} \gamma^{(\boldsymbol{\theta})} / \inf_{\boldsymbol{\theta}} \gamma^{(\boldsymbol{\theta})} < c_\gamma$*, and that the orders* $\overline{m}$, $\underline{m}$ *are positive. Then*

(a$''$) *two-grid scheme has level-independent convergence properties*
if *both the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from* 1 *and the order rule* $m_{FB}(P) \geq \overline{m}/2$ *holds;*
only if *both the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from* 1 *and the order rule* $m_{FB}(P) \geq \underline{m}/2$ *holds;*

(b$''$) *the parameter* $\delta^{(\nu)}$ *in McCormick's bound is bounded away from zero*
if *both the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from* 1 *and the order rule* $m_{FB}(P) \geq \overline{m}$ *holds;*
only if *both the smoothing factor* $\mu = 1 - \inf_{\boldsymbol{\theta}} \widetilde{\lambda}_2^{(\boldsymbol{\theta})}$ *is bounded away from* 1 *and the order rule* $m_{FB}(P) \geq \underline{m}$ *holds;*

Proof. Equation (5.5) gives

$$\tan^2 \phi^{(\boldsymbol{\theta})} = \frac{1}{\cos^2 \phi^{(\boldsymbol{\theta})}} - 1 = \frac{\|\mathbf{p}^{(\boldsymbol{\theta})}\|^2}{|(\mathbf{p}^{(\boldsymbol{\theta})})_1|^2} - 1 = \frac{\sum_{i=2}^{m_{\boldsymbol{\theta}}} |\left(\mathbf{p}^{(\boldsymbol{\theta})}\right)_i|^2}{|(\mathbf{p}^{(\boldsymbol{\theta})})_1|^2}.$$

Combining this with the definition (6.2) of frequency based order, one can check that m_{FB} is the largest constant such that, for $|\boldsymbol{\theta}| \to 0$,

$$\tan \phi^{(\boldsymbol{\theta})} = \mathcal{O}(|\boldsymbol{\theta}|^{m_{FB}(P)}).$$

Using this latter together with (6.1) and the Proposition 5.2 finishes the proof.

□

In [8], the analysis is not restricted to Galerkin setting for symmetric systems, and the following more general rule is stated

$$m_{HF}(P) + m_{HF}(\bar{P}) \geq \overline{m}, \tag{6.3}$$

where $\bar{P}$ is the transpose of the restriction operator (thus $\bar{P} = P$ in the context of this paper) and where the *high frequency order* $m_{HF}(P)$ is defined as the largest constant such that

$$|\left(\mathbf{p}^{(\boldsymbol{\theta})}\right)_i| = \mathcal{O}(|\boldsymbol{\theta}|^{m_{HF}(P)}) \quad \text{for} \quad |\boldsymbol{\theta}| \to 0 \quad \text{and for} \quad i = 2, \ldots, m_{\boldsymbol{\theta}}. \tag{6.4}$$

This definition (stated here under a form close to that in [25]) coincides with our definition of frequency based order if $|\left(\mathbf{p}^{(\boldsymbol{\theta})}\right)_1| = \mathcal{O}(1)$, in which case the rule (6.3) amounts to that in (a$''$). Our analysis gives then a firmer basis to this often referred rule, since we prove it as a part of a necessary and a sufficient condition for nice two-grid convergence, whereas it appears in [8] only as a necessary condition for the boundness of the coarse grid correction operator. Our analysis also reveals more clearly the complementary role played by the smoothing factor, while giving some further insight on the V-cycle convergence via the assessment of McCormick bound. Note, however, that our approach does not completely cover that in [8] since we assume Galerkin coarse grid matrices, whereas in [8] it is assumed instead that both fine and coarse grid matrices correspond to the same stencil.

The fact that $m_{HF}(P) \geq \overline{m}/2$ can provide a sufficient condition for two-grid convergence, and that $m_{HF}(P) \geq \overline{m}$ can provide a sufficient condition for the convergence of the V-cycle is also shown in [6] via the theory of multigrid methods for Toeplitz matrices (see also [16]). Perhaps because of this latter context, this analysis however seemingly miss the role of the smoothing factor, and that the condition $m_{HF}(P) \geq \overline{m}/2$ is not only sufficient, but (since the case $\overline{m} = \underline{m}$ is considered) also necessary for two-grid convergence. Moreover, the results in [6] are stated with respect to high frequency order as defined in [8], and therefore ignore the subtle but in some cases important difference with our frequency based order (difference, illustrated with the example below).

It is worth mentioning Hackbusch analysis [7], which also proves that an order condition $m(P) \geq \overline{m}/2$ is, in the Galerkin setting for symmetric systems, sufficient for the two-grid convergence. There $m(P) - 1$ is the largest degree of a polynomial that is preserved by the prolongation. Of course, this latter analysis is more general than the one in this paper since it does not require the Fourier analysis framework. However, when this latter applies, it appears that the polynomial-based definition of the order is more restrictive (see [8,25]) (see also [17,18] for more on the relation between different orders).

Example For the considered second order differential equation (2.9) we have $\overline{m} = \underline{m} = 2$, whereas for the linear interpolation, $m_{HF} = m_{FB} = 2$.

Now, it is interesting to observe that, when $\lim_{|\boldsymbol{\theta}|\to\mathbf{0}} |(\mathbf{p}^{(\boldsymbol{\theta})})_1| = 0$, then our Proposition 6.1 still holds whereas the low frequency order can be misleading. For instance, consider instead of linear interpolation the (unexpected, but still possible) prolongation given by the stencil

$$[1 \quad 0 \quad -1].$$

One then has

$$\mathbf{p}^{(\theta)} = \exp(-i\theta) \begin{pmatrix} 1 - \exp(2i\theta) \\ -1 + \exp(2i\theta) \end{pmatrix},$$

and, therefore, there holds $m_{FB} = 0$, whereas $m_{HF} = 1$. Note that the use of the Galerkin coarse grid matrix leads here to the same stencil as on the fine level; hence, both the analysis in [8,25] and the one of Proposition 6.1 apply simultaneously.

7 Semi-positive definite problems and local Fourier analysis

In this section we consider Fourier analysis for symmetric semi-positive definite linear systems. Such extension is motivated by local Fourier analysis (also called local mode analysis) that has a wider scope than (rigorous) Fourier analysis. The main idea is the assessment of the two-grid convergence and/or of the smoothing factor without taking boundary conditions into account. In practice, such approach is often equivalent to the use of periodic boundary conditions and therefore leads to linear systems with non-trivial null space.

Another way to interpret local Fourier analysis is to consider it as a limit case of (rigorous) Fourier analysis for SPD problems on grids of increasing dimension (with, thus, decreasing influence of boundary conditions on estimated parameters). Now, we previously observed for the SPD case that the angle between the range of prolongation and an eigenvector of A should be proportional to the size of the eigenvalue. In the limit case of local Fourier analysis, the modes belonging to the null space $\mathcal{N}(A)$ of A should therefore be interpolated exactly, which also corresponds to a common practice. It then follows that null space components seemingly play no role in the convergence, and hence that Fourier analysis may be carried out ignoring these modes. This, indeed, is the common practice when assessing the two-grid convergence factor (see, for instance [20, p. 109], [23, p. 107] and the references therein).

In Theorem 7.1 below we give theoretical foundation to this approach with respect to V-cycle multigrid, showing that McCormick's bound on the convergence rate can also be computed ignoring singular modes, or, more precisely, restricting the minimum in (3.3) to vectors belonging to the range of A_k. Since (3.3) is at the root of the further analysis developed in Lemmas 4.1 and 5.1, the application of the results in Sects. 4, 5 and 6 to local mode analysis is then straightforward.

Now, to state the theorem, we need to extend our definition of the V-cycle multigrid algorithm in Sect. 2 to A_k possibly singular. The only potential difficulty comes in fact with the bottom level matrix A_0 whose inverse is needed. In Theorem 7.1, we assume that instead one uses any matrix B_0 such that $A_0 B_0 A_0 = A_0$. Such matrices are called $\{1\}$-inverse in [1], and one may check that if $\mathbf{r}_0 \in \mathcal{R}(A_0)$, then $A_0 B_0 \mathbf{r}_0 = \mathbf{r}_0$. On the other hand, to generalize (3.3), we need the inverse of the restriction of A_k to its range. The most convenient way to express it is to use the Moore–Penrose inverse A_k^+ of A_k, since, if $A_k = X \mathrm{diag}(\lambda_i) X^T$, then $A_k^+ = X \mathrm{diag}(\lambda_i^+) X^T$ with

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

The expression of A_k^+ is thus particularly simple when using the Fourier basis which makes A_k diagonal.

Theorem 7.1 *Let $\mathbf{x}_{n+1} = MG(\mathbf{b}, A, \mathbf{x}_n, J)$ be the vector resulting from the application of the multigrid algorithm with V-cycle at level $J > 1$, where A is symmetric semi-positive definite, and, in case A_0 is singular, where A_0^{-1} is exchanged for any matrix $A_0^{(1)}$ such that $A_0 A_0^{(1)} A_0 = A_0$. Assume that P_k, $k = 0, \ldots, J-1$, A_k, $k = 0, \ldots, J$, and R_k, $k = 1, \ldots, J$ satisfy the general assumptions stated in Sect. 2 with $\rho(I - R_k^{-1} A_k) < 1$ being replaced by $\rho(I - R_k^{-1} A_k) \leq 1$, with $(I - R_k^{-1} A_k)\mathbf{z} = \lambda \mathbf{z}$ for $|\lambda| = 1$ if and only if $\mathbf{z} \in \mathcal{N}(A_k)$. Let $\mathcal{P}_{\mathcal{R}(A),\mathcal{N}(A)}$ be the orthogonal projector onto the range of A.*

If $\mathbf{b} \in \mathcal{R}(A)$*, then, for any solution* $\widetilde{\mathbf{x}}$ *to* (1.1)*,*

$$\mathcal{P}_{\mathcal{R}(A),\mathcal{N}(A)}\,(\widetilde{\mathbf{x}} - \mathbf{x}_{n+1}) = E_{MG}^{(J)}\,(\widetilde{\mathbf{x}} - \mathbf{x}_n) = E_{MG}^{(J)}\,\mathcal{P}_{\mathcal{R}(A),\mathcal{N}(A)}\,(\widetilde{\mathbf{x}} - \mathbf{x}_n)$$

for some matrix $E_{MG}^{(J)}$ *satisfying*

$$\rho(E_{MG}^{(J)}) \;\leq\; 1 - \min_{1\leq k\leq J}\,\widetilde{\delta}_k^{(\nu)}\;,$$

with

$$\widetilde{\delta}_k^{(\nu)} \;=\; \min_{\mathbf{v}_k\in\mathcal{R}(A_k)}\,\frac{\mathbf{v}_k^T\,N_k^{(2\nu)}\,\mathbf{v}_k}{\mathbf{v}_k^T\,(A_k^+ - P_{k-1}A_{k-1}^+P_{k-1}^T)\mathbf{v}_k}. \tag{7.1}$$

Proof. Let $q_k = \dim(\mathcal{N}(A_k))$. Observe that, A_k being non-negative definite, $A_{k-1} = P_{k-1}^T A_k P_{k-1}$ is non-negative definite with $q_{k-1} \leq q_k$. Without loss of generality, we can express all the matrices using bases of $\mathbb{R}^{n_k}$, $k = J, \ldots, 0$ such that, when $q_k > 0$, the first q_k canonical vectors span $\mathcal{N}(A_k)$. Hence, A_k admits a block representation

$$A_k = \begin{pmatrix} O_{q_k,q_k} & \\ & A_k^{\mathcal{R}\mathcal{R}} \end{pmatrix}, \tag{7.2}$$

with all but lower right blocks being empty if $q_k = 0$.

Similarly, we partition

$$R_k = \begin{pmatrix} R_k^{\mathcal{N}\mathcal{N}} & R_k^{\mathcal{N}\mathcal{R}} \\ R_k^{\mathcal{R}\mathcal{N}} & R_k^{\mathcal{R}\mathcal{R}} \end{pmatrix}, \quad N_k^{(\nu)} = \begin{pmatrix} N_k^{(\nu)\,\mathcal{N}\mathcal{N}} & N_k^{(\nu)\,\mathcal{N}\mathcal{R}} \\ N_k^{(\nu)\,\mathcal{R}\mathcal{N}} & N_k^{(\nu)\,\mathcal{R}\mathcal{R}} \end{pmatrix},$$
$$P_{k-1} = \begin{pmatrix} P_{k-1}^{\mathcal{N}\mathcal{N}} & P_{k-1}^{\mathcal{N}\mathcal{R}} \\ P_{k-1}^{\mathcal{R}\mathcal{N}} & P_{k-1}^{\mathcal{R}\mathcal{R}} \end{pmatrix},$$

where all but lower right blocks of R_k, $N_k^{(\nu)}$ and P_{k-1} become empty when $q_k = 0$, and where $P_{k-1}^{\mathcal{N}\mathcal{N}}$ and $P_{k-1}^{\mathcal{R}\mathcal{N}}$ are empty when $q_k > 0$ with $q_{k-1} = 0$. If $q_k > 0$, there holds

$$A_{k-1} = P_{k-1}^T A_k P_{k-1} = \begin{pmatrix} {P_{k-1}^{\mathcal{R}\mathcal{N}}}^T A_k^{\mathcal{R}\mathcal{R}} P_{k-1}^{\mathcal{R}\mathcal{N}} & {P_{k-1}^{\mathcal{R}\mathcal{N}}}^T A_k^{\mathcal{R}\mathcal{R}} P_{k-1}^{\mathcal{R}\mathcal{R}} \\ {P_{k-1}^{\mathcal{R}\mathcal{R}}}^T A_k^{\mathcal{R}\mathcal{R}} P_{k-1}^{\mathcal{R}\mathcal{N}} & {P_{k-1}^{\mathcal{R}\mathcal{R}}}^T A_k^{\mathcal{R}\mathcal{R}} P_{k-1}^{\mathcal{R}\mathcal{R}} \end{pmatrix}.$$

Hence, in view of the form (7.2) and the fact that $A_k^{\mathcal{R}\mathcal{R}}$ is SPD, one must have $P_{k-1}^{\mathcal{R}\mathcal{N}} = O$. It then follows that, for any $n_k \times n_k$ matrix

$$B_{k-1} = \begin{pmatrix} * & * \\ * & B_{k-1}^{\mathcal{R}\mathcal{R}} \end{pmatrix},$$

one has

$$P_{k-1}B_{k-1}P_{k-1}^T = \begin{pmatrix} * & * \\ * & P_{k-1}^{\mathcal{RR}} B_k^{\mathcal{RR}} \; P_{k-1}^{\mathcal{RR}\;T} \end{pmatrix}, \tag{7.3}$$

This latter relation also holds for $q_k = 0$, the blocks denoted by a star $*$ being then empty.

Now, $\mathbf{x}_{n+1} = \mathrm{MG}(\mathbf{b}, A_k, \mathbf{x}_n, k)$ may be expressed as

$$\mathbf{x}_{n+1} = \mathbf{x}_n + B_J(\mathbf{b} - A\mathbf{x}_n), \tag{7.4}$$

where the matrix B_J is defined from the recursion

$$\begin{aligned} B_0 &= A_0^{(1)} \\ B_k &= N_k^{(2\nu)} - (I - N_k^{(\nu)} \; A_k)P_{k-1}B_{k-1}P_{k-1}^T(I - A_k \; N_k^{(\nu)}), \quad k = 1, \dots, J \end{aligned}$$

(see, e.g., [21, Section 5.1]; B_k^{-1} in this reference corresponds to B_k here).

Since there holds

$$I - N_k^{(2\nu)} \; A_k = \begin{pmatrix} I_{q_k} & * \\ & I - N_k^{(2\nu)\;\mathcal{RR}} \; A_k^{\mathcal{RR}} \end{pmatrix} \tag{7.5}$$

with all but lower right blocks being empty if $q_k = 0$, letting

$$B_J = \begin{pmatrix} * & * \\ * & B_J^{\mathcal{RR}} \end{pmatrix},$$

it follows from (7.3) that $B_J^{\mathcal{RR}}$ may be computed from the recursion

$$\begin{aligned} B_k^{\mathcal{RR}} = N_k^{(2\nu)\;\mathcal{RR}} &- \left(I - N_k^{(\nu)\;\mathcal{RR}} \; A_k^{\mathcal{RR}}\right) \\ &\times P_{k-1}^{\mathcal{RR}} B_{k-1}^{\mathcal{RR}} \; P_{k-1}^{\mathcal{RR}\;T} \left(I - A_k^{\mathcal{RR}} \; N_k^{(\nu)\;\mathcal{RR}}\right), \\ & k = 1, \dots, J. \end{aligned}$$

On the other hand, when A_0 is singular, $A_0B_0A_0 = A_0$ holds for A_0 of the form (7.2) if and only if $B_0^{\mathcal{RR}} = A_0^{\mathcal{RR}\;-1}$, whereas from (7.5) we deduce

$$I - N_k^{(2\nu)\;\mathcal{RR}} \; A_k^{\mathcal{RR}} = \left(I - N_k^{(\nu)\;\mathcal{RR}} \; A_k^{\mathcal{RR}}\right)^2.$$

Hence $E_k = I - B_k^{\mathcal{RR}} A_k^{\mathcal{RR}}$ obeys the recursion

$$\begin{aligned}
E_0 &= O \\
E_k &= \left(I - N_k^{(\nu)\,\mathcal{RR}} \, A_k^{\mathcal{RR}} \right) \\
&\quad \times \left(I - P_{k-1}^{\mathcal{RR}} (I - E_{k-1}) \, A_{k-1}^{\mathcal{RR}\,-1} \, P_{k-1}^{\mathcal{RR}\,T} \, A_k^{\mathcal{RR}} \right) \left(I - N_k^{(\nu)\,\mathcal{RR}} \, A_k^{\mathcal{RR}} \right), \\
&\quad k = 1, 2, \ldots, J.
\end{aligned}$$

similar to (2.3); that is, corresponding to a multigrid scheme satisfying all assumptions of Theorem 3.1, which therefore implies

$$\rho(E_J) \leq 1 - \max_{1 \leq k \leq J} \delta_k^{(\nu)} \tag{7.6}$$

with

$$\delta_k^{(\nu)} = \min_{\mathbf{v} \in \mathbb{R}^{n_k - q_k}} \frac{\mathbf{v}^T N_k^{\mathcal{RR}} \mathbf{v}}{\mathbf{v}^T (A_k^{\mathcal{RR}\,-1} - P_{k-1}^{\mathcal{RR}} \, A_{k-1}^{\mathcal{RR}\,-1} \, P_{k-1}^{\mathcal{RR}\,T}) \mathbf{v}} .$$

Moreover, $\widetilde{\delta}_k^{(\nu)} = \delta_k^{(\nu)}$ since

$$A_k^+ = \begin{pmatrix} O & \\ & A_k^{\mathcal{RR}\,-1} \end{pmatrix}$$

with all but lower right block being empty when $q_k = 0$.

Finally, using (7.4) and the fact that $\mathbf{b} \in \mathcal{R}(A)$, there holds

$$\widetilde{\mathbf{x}} - \mathbf{x}_{n+1} = (I - B_J A)(\widetilde{\mathbf{x}} - \mathbf{x}_n),$$

and hence

$$\begin{aligned}
\mathcal{P}_{\mathcal{R}(A),\mathcal{N}(A)} \, (\widetilde{\mathbf{x}} - \mathbf{x}_{n+1}) &= \begin{pmatrix} O & O \\ O & I \end{pmatrix} \begin{pmatrix} I & * \\ & I - B_k^{\mathcal{RR}} A_k^{\mathcal{RR}} \end{pmatrix} (\widetilde{\mathbf{x}} - \mathbf{x}_n) \\
&= \begin{pmatrix} O & O \\ O & E_J \end{pmatrix} (\widetilde{\mathbf{x}} - \mathbf{x}_n) \\
&= \begin{pmatrix} O & O \\ O & E_J \end{pmatrix} \mathcal{P}_{\mathcal{R}(A),\mathcal{N}(A)} \, (\widetilde{\mathbf{x}} - \mathbf{x}_n),
\end{aligned}$$

which, together with (7.6), concludes the proof. □

Acknowledgments We thank Cornelis Oosterlee for bringing to our attention the Refs. [8] and [18].

References

1. Ben Israel, A., Greville, T.N.E.: Generalized Inverses: Theory and Applications. Wiley, New York (1974)
2. Brandt, A.: Guide to multigrid development. In: Hackbusch, W., Trottenberg, U. (eds.) Multigrid Methods. Lectures Notes in Mathematics No. 960, pp. 220–312. Springer, Berlin (1982)
3. Brandt, A.: Rigorous quantitative analysis of multigrid, I: constant coefficients two-level cycle with L2-norm. SIAM J. Numer. Anal. **31**, 1695–1730 (1994)
4. Brezina, M., Cleary, A.J., Falgout, R.D., Henson, V.E., Jones, J.E., Manteuffel, T.A., McCormick, S.F., Ruge, J.W.: Algebraic multigrid based on element interpolation (AMGe). SIAM J. Sci. Comput. **22**, 1570–1592 (2000)
5. Chartier, T., Falgout, R.D., Henson, V.E., Jones, J., Manteuffel, T., McCormick, S., Ruge, J., Vassilevski, P.S.: Spectral AMGe (ρAMGe). SIAM J. Sci. Comput. **25**, 1–26 (2004)
6. Donatelli, M.: An algebraic generalization of local fourier analysis for grid transfer operators in multigrid based on Toeplitz matrices. Numer. Linear Algebra Appl. **17**, 179–197 (2010)
7. Hackbusch, W.: Multi-grid Methods and Applications. Springer, Berlin (1985)
8. Hemker, P.: On the order of prolongations and restrictions in multigrid procedures. J. Comput. Appl. Math. **32**, 423–429 (1990)
9. Mandel, J., McCormick, S.F., Ruge, J.W.: An algebraic theory for multigrid methods for variational problems. SIAM J. Numer. Anal. **25**, 91–110 (1988)
10. McCormick, S.F.: Multigrid methods for variational problems: general theory for the V-cycle. SIAM J. Numer. Anal. **22**, 634–643 (1985)
11. Muresan, A.C., Notay, Y.: Analysis of aggregation-based multigrid. SIAM J. Sci. Comput. **30**, 1082–1103 (2008)
12. Napov, A.: Algebraic analysis of V-cycle multigrid and aggregation-based two-grid methods. PhD thesis, Service de Métrologie Nucléaire, Université Libre de Bruxelles, Brussels, Belgium (2010)
13. Napov, A., Notay, Y.: Comparison of bounds for V-cycle multigrid. Appl. Numer. Math. **60**, 176–192 (2010)
14. Napov, A., Notay, Y.: When does two-grid optimality carry over to the V-cycle?. Numer. Linear Algebra Appl. **17**, 273–290 (2010)
15. Ruge, J.W., Stüben, K.: Algebraic multigrid (AMG). In: McCormick, S.F. (ed.) Multigrid Methods. Frontiers in Applied Mathematics, vol. 3, pp. 73–130. SIAM, Philadelphia (1987)
16. Serra-Capizzano, S.: Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences. Numerische Mathematik **92**, 433–465 (2002)
17. Stevenson, R.: On the validity of local mode analysis of multi-grid methods. PhD thesis, Utrecht University (1990)
18. Stevenson, R.: More on the order of prolongations and restrictions. Appl. Numer. Math. **58**, 1875–1880 (2007)
19. Stüben, K., Trottenberg, K.U.: Multigrid methods: fundamental algorithms, model problem analysis and applications. In: Hackbusch, W., Trottenberg, U. (eds.) Multigrid Methods. Lectures Notes in Mathematics No. 960, pp. 1–176. Springer, Berlin (1982)
20. Trottenberg, U., Oosterlee, C.W., Schüller, A.: Multigrid. Academic Press, London (2001)
21. Vassilevski, P.S.: Multilevel Block Factorization Preconditioners. Springer, New York (2008)
22. Wesseling, P.: An Introduction to Multigrid Methods. Wiley, Chichester (1992)
23. Wienands, R., Joppich, W.: Practical Fourier Analysis for Multigrid Methods. Chapman & Hall/CRC Press, Boca Raton (2005)
24. Wienands, R., Oosterlee, C.W.: On three-grid Fourier analysis for multigrid. SIAM J. Sci. Comput. **23**, 651–671 (2001)
25. Yavneh, I.: Coarse-grid correction for nonelliptic and singular perturbation problems. SIAM J. Sci. Comput. **19**, 1682–1699 (1998)