MATHEMATICS OF COMPUTATION
https://doi.org/10.1090/mcom/3460
Article electronically published on July 11, 2019

ANALYSIS OF TWO-GRID METHODS: THE NONNORMAL CASE

YVAN NOTAY

Abstract. Core results about the algebraic analysis of two-grid methods are extended in relations bounding the field of values (or numerical range) of the iteration matrix. On this basis, bounds are obtained on its norm and numerical radius, leading to rigorous convergence estimates. Numerical illustrations show that the theoretical results deliver qualitatively good predictions, allowing one to anticipate success or failure of the two-grid method. They also indicate that the field of values and the associated numerical radius are much more reliable convergence indicators than the eigenvalue distribution and the associated spectral radius. On this basis, some discussion is developed about the role of local Fourier or local mode analysis for nonsymmetric problems.

1. Introduction

Regarding linear systems with nonsymmetric (and, more specifically, nonnormal) system matrices, there is an important gap between the general literature on iterative methods and works focused on multigrid methods. In the former, many studies address nonnormality effects, with the outcome that, except for matrices close to Hermitian, the eigenvalue distribution of the (possibly preconditioned) system matrix may be a completely misleading convergence indicator [20].

For instance, in [19] it is shown any given eigenvalue set can be associated with any prescribed convergence curve when using GMRES. The negative side of this result is that slow or no convergence may take place even when the system matrix has all its eigenvalues well clustered around one. Because GMRES minimizes the residual, this a fortiori holds for stationary iterations.

Many works focus then on alternative convergence indicators. The pseudospectrum is one of them; see [38, Chapter VI] and the references therein. It leads to a rigorous bound on the GMRES residual that is meaningful in some cases, but that can also be too pessimistic [20]. Other works point to the role of the field of values when it can be shown that it does not include the origin of the complex plane; see, e.g., [11] for stationary iterations and [12, 34] for minimal residual-type iterations (see also [20] for a related result).

Nevertheless, the assessment of multigrid methods remains largely dominated by eigenvalue analyses. More precisely, although these methods are also used as a preconditioner in combination with Krylov subspace methods (especially the algebraic multigrid variants), they are mostly analyzed as stationary iterative methods. The

Received by the editor April 3, 2018, and, in revised form, March 5, 2019, and April 12, 2019.
2010 *Mathematics Subject Classification.* Primary 65F08, 65F10, 65F50, 65N22.
Key words and phrases. Iterative methods, convergence analysis, linear systems, multigrid, two-grid, nonnormal matrices, AMG, preconditioning.
The author is Research Director of the Fonds de la Recherche Scientifique – FNRS.

spectral radius of the associated iteration matrix is then used as a main convergence indicator, despite that it rigorously describes only the asymptotic convergence.

Abstract convergence theories (e.g., [21, 31]) provide an alternative. They are based on norms and are thus robust with respect to nonnormality effects. However, their assumptions are difficult to check (especially in the nonsymmetric case), whereas the related bounds involve unknown constants. Hence, it is difficult to use them for the practical assessment of the potentialities of a given multigrid scheme.

To this purpose, one thus most often uses the spectral radius estimation with (rigorous) Fourier analysis and local Fourier analysis (LFA, sometimes referred to as local mode analysis); see, e.g., [2, 4, 5, 37, 39, 42]. This means considering a two-grid variant of the method at hand for model but typical problems, like PDEs with constant coefficients discretized on uniform grids.

For more general problems and methods, like algebraic multigrid methods, algebraic convergence theories have been developed [3, 14, 15, 33], which are more general in that they do not require model problem assumptions. For symmetric and positive definite problems, they allow one to bound the spectral radius of the iteration matrix as a function of more tractable quantities. In [28, 29], this approach is extended to nonsymmetric problems, but the focus is kept on eigenvalues.

Some might therefore be surprised that nonnormality effects with two-grid methods can in fact be observed even in quite simple examples. Consider for instance the $(m-1)^2 \times (m-1)^2$ matrix associated with the stencil

$$\begin{bmatrix} & -1 & \\ -\alpha & 2+\alpha+\beta & -1 \\ & -\beta & \end{bmatrix} \tag{1.1}$$

on a square $(m+1) \times (m+1)$ grid with Dirichlet-type boundary conditions. Restricting to $\alpha, \beta \geq 1$, this corresponds to the upwind finite difference discretization on a uniform mesh of

$$-\nu \Delta u + \overline{v} \left(\overline{\nabla} u \right) = 0 \qquad \text{in } (0, L) \times (0, L) \tag{1.2}$$

with Dirichlet boundary conditions for u on the four sides of the domain and convective field $\overline{v} = \big(h^{-1} \nu (1-\alpha), h^{-1} \nu (1-\beta) \big)$, where $h = m^{-1} L$ is the mesh size.

For this matrix, we consider the basic but standard two-grid method that uses a single pre-smoothing step with damped Jacobi smoothing (no post-smoothing and damping parameter $\omega = \frac{1}{2}$), together with a coarse grid correction based on h–$2h$ coarsening, bilinear interpolation for the prolongation, full weighting for the restriction and the Galerkin coarse grid matrix. The zero vector is used as initial approximation and the right hand side has been normalized, entailing that the initial residual has unit norm.

For $\alpha = 50$, $\beta = 1$, and $m = 30$, in Figure 1, left, we depict the computed spectrum of the iteration matrix, whereas in Figure 1, right, we plot the evolution of the relative residual error $\|\mathbf{r}_k\|$ as a function of the number of stationary iterations, together with its estimates ρ^k based on the computed spectral radius ρ of the iteration matrix. Clearly, these estimates are completely misleading here.

Of course, the spectral radius correctly reflects the *asymptotic* convergence, which takes place after enough iterations have been performed, about 70 in the present example. But such a delay is unacceptable for a multilevel method, and it can in fact prevent any convergence when the two-grid scheme is recursively used in a

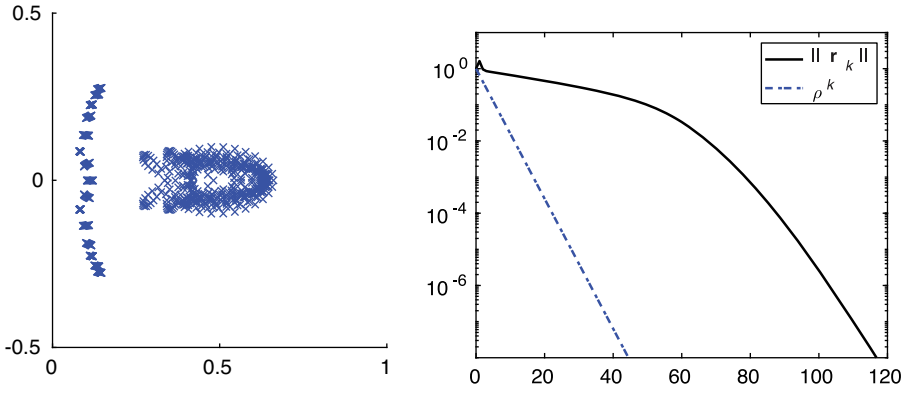

FIGURE 1. Problem (1.1), (1.2) with $\alpha = 50$, $\beta = 1$. Left: spectrum of the iteration matrix T ($\times$). Right: residual norm as a function of the number of iterations and its estimation $(\rho(T))^k$, where k is the iteration index.

multigrid cycle, whose convergence heavily relies on the capacity to have significant error reduction within a few iterations [39, Section 3.2].

On the other hand, any multigrid expert would advise the use of (line) Gauss–Seidel smoothing instead of damped Jacobi in this example, and this would likely solve the observed convergence issue. But one may prefer Jacobi because, e.g., it is straightforwardly parallelizable, and, anyway, independently of this example, some fundamental questions are raised which deserve to be investigated.

In fact, a number of works already mention issues associated with the classical eigenvalue analysis of multigrid methods and develop alternative analysis strategies. Among these, the best known is probably the Brandt half plane analysis for convection dominated (singularly perturbed) problems, which is based on using a semi-infinite grid instead of the fully infinite one used in LFA [5,6]. Other works address more specifically parabolic problems, for which nonnormality effects are most obvious; see, e.g., [16,17,23]. Now, while interesting, these previous studies focus on some particular applications and the proposed approaches remain heuristic, with the notable exception of [10], where rigorous convergence estimates are obtained, which are, however, restricted to a specific parallel-in-time method [13]. One is thus lacking a study that addresses the raised questions from a general viewpoint, while deriving rigorous and widely usable convergence bounds.

In the present work, we bring several contributions to such a study. Our main result is a reworking of the tools proposed in [28,29], which extend to nonsymmetric matrices the algebraic convergence theory of two-grid methods [3,14,15,33]. Among the results in [28, 29], the most easy to deal with concern the case of a single smoothing step with a symmetric and positive definite smoother (as in the above example). They show that the eigenvalues of the iteration matrix are localized in a bounded region of the complex plane described by two easy to assess quantities, which, in fact, are related to a companion symmetric problem (see Section 3 for a precise definition). In view of the above discussion, these results seem to lose most of their relevance. Here we show that the considered region does not only contain the eigenvalues of the iteration matrix, but also the whole field of values (or numerical range) and, as such, gives rise to rigorous convergence estimates.

These results naturally lead to an enlightening discussion of the field of values with respect to the eigenvalues. (Remember that, in the normal case, the field of values is just the convex hull of the eigenvalues, and hence there is no significant difference between both.) Simple numerical experiments suggest that the field of values and the associated numerical radius are not only more rigorous convergence indicators than the eigenvalue distribution and the associated spectral radius; they also offer realistic convergence prediction, and might thus represent the proper generalization of the eigenvalue tools universally used in the symmetric case.

Eventually, our results allow us to enrich the debate about the role of LFA for nonsymmetric problems. Several works discuss the failure of LFA for nonnormal matrices; see [16, 17, 30] and [5, Section 7.5]. Opposite to this, satisfactory convergence predictions for parabolic problems are obtained with LFA in [41] and with a closely related approach in [23, 40]. Here we advocate the idea that, to assess LFA, one should check whether it allows one to correctly predict the extent of the field of values of the iteration matrix, regardless of the actual location of its eigenvalues.

The remainder of this work is organized as follows. In Section 2, we present generic tools—some well known and some new—to characterize the convergence of stationary iterations for nonsymmetric systems, while stressing the connections with the field of values. In Section 3, we develop our analysis of two-grid iterations. The results of numerical experiments are presented in Section 4, where we also discuss LFA. Concluding remarks are given in Section 5.

2. Convergence measures for stationary iterations

We start with a few definitions that clarify notation used throughout this work. In these definitions, $(\mathbf{z},\mathbf{z})$ represents a given scalar product on $\mathbb{C}^n$ and $\|\mathbf{z}\| = (\mathbf{z},\mathbf{z})^{1/2}$ stands for the induced norm.

Definition 2.1. Let C be an $n\times n$ complex matrix, and let $\mathcal{S}$ be a given subspace of $\mathbb{C}^n$. We define the norm, the numerical range (or field of values), and the numerical radius of C with respect to $\mathcal{S}$ by, respectively,

$$\|C\|_{\mathcal{S}} = \max_{\mathbf{z}\in S\setminus\{0\}} \frac{\|C\mathbf{z}\|}{\|\mathbf{z}\|}, \tag{2.3}$$

$$W_{\mathcal{S}}(C) = \left\{ \frac{(\mathbf{z},C\mathbf{z})}{(\mathbf{z},\mathbf{z})} : \mathbf{z}\in S\setminus\{0\} \right\}, \tag{2.4}$$

$$w_{\mathcal{S}}(C) = \max_{\mathbf{z}\in S\setminus\{0\}} \frac{|(\mathbf{z},C\mathbf{z})|}{(\mathbf{z},\mathbf{z})}. \tag{2.5}$$

We further denote with

$$\overline{\mu}_{\mathcal{S}}(C) = \max_{\mathbf{z}\in S\setminus\{0\}} \frac{\Re\mathrm{e}(\mathbf{z},C\mathbf{z})}{(\mathbf{z},\mathbf{z})}, \tag{2.6}$$

$$\underline{\mu}_{\mathcal{S}}(C) = \min_{\mathbf{z}\in S\setminus\{0\}} \frac{\Re\mathrm{e}(\mathbf{z},C\mathbf{z})}{(\mathbf{z},\mathbf{z})}, \tag{2.7}$$

the maximal and minimal real part of all elements in $W_{\mathcal{S}}(C)$.

We shall only use these definitions in contexts where $\mathcal{S}$ is an invariant subspace of C. Note that the standard definitions of matrix norm, numerical range, and

numerical radius are recovered when $\mathcal{S} = \mathbb{C}^n$. Hence, in the following, we write $\|C\|$ for $\|C\|_{\mathbb{C}^n}$ and $W(C)$ for $W_{\mathbb{C}^n}(C)$.

Below, we often use $\underline{\mu}_{\mathcal{S}}(C^{-1})$ for nonsingular C. The following lemma states some useful properties of this quantity. The first of these is borrowed from [34] (we recall the proof for the sake of completeness).

Lemma 2.1. *Let C be a nonsingular $n \times n$ complex matrix, and let $\mathcal{S}$ be a given invariant subspace of C. Then, $\underline{\mu}_{\mathcal{S}}(C^{-1})$ is positive if and only if $\underline{\mu}_{\mathcal{S}}(C)$ is positive, in which case there holds*

$$\underline{\mu}_{\mathcal{S}}(C^{-1}) \geq \frac{\underline{\mu}_{\mathcal{S}}(C)}{\left(\|C\|_{\mathcal{S}}\right)^2} \quad \textit{and} \quad \underline{\mu}_{\mathcal{S}}(C) \geq \frac{\underline{\mu}_{\mathcal{S}}(C^{-1})}{\left(\|C^{-1}\|_{\mathcal{S}}\right)^2}. \tag{2.8}$$

Moreover, for any positive α,

$$\underline{\mu}_{\mathcal{S}}(C^{-1}) \geq \alpha \iff \|2\alpha C - I\|_{\mathcal{S}} \leq 1. \tag{2.9}$$

Proof. First observe that, because $\mathcal{S}$ is an invariant subspace of C, the subspaces $\mathcal{S}$, $C\mathcal{S}$, and $C^{-1}\mathcal{S}$ coincide. Then, for any $\mathbf{z} \in S\backslash\{0\}$ and $\mathbf{w} = C^{-1}\mathbf{z}$ (which is thus also in $\mathcal{S}\backslash\{0\}$),

$$\frac{\Re\mathrm{e}(\mathbf{z}, C^{-1}\mathbf{z})}{(\mathbf{z},\mathbf{z})} = \frac{\Re\mathrm{e}(\mathbf{w}, C\mathbf{w})}{(C\mathbf{w}, C\mathbf{w})} \geq \underline{\mu}_{\mathcal{S}}(C)\frac{(\mathbf{w},\mathbf{w})}{(C\mathbf{w}, C\mathbf{w})};$$

hence the "if" statement and the left inequality (2.8). The "only if" statement and the right inequality (2.8) further follow by permuting the roles of C and C^{-1}.

On the other hand, observe that, for any $\mathbf{z} \in S$, and letting $\mathbf{w} = C\mathbf{z}$,

$$\|(2\alpha C - I)\,\mathbf{z}\|^2 = 4\alpha^2(\mathbf{w},\mathbf{w}) - 4\alpha\Re\mathrm{e}(\mathbf{w}, C^{-1}\mathbf{w}) + (\mathbf{z},\mathbf{z});$$

that is,

$$(\mathbf{z},\mathbf{z}) - \|(2\alpha C - I)\,\mathbf{z}\|^2 = 4\alpha\left(\Re\mathrm{e}(\mathbf{w}, C^{-1}\mathbf{w}) - \alpha(\mathbf{w},\mathbf{w})\right).$$

Hence the left hand side of this equality is always nonnegative if and only if the right hand side is always nonnegative, which is precisely what is stated in (2.9). ☐

Now we are ready to discuss the convergence of iterative methods. Let $A\mathbf{u} = \mathbf{b}$ be a linear system to solve, let $\mathbf{u}_0$ be some initial approximation, and consider the following stationary iteration, where B is a given (nonsingular) preconditioner:

$$\begin{array}{l} \text{For } k = 0, 1, \ldots: \\ \left\{ \begin{array}{rcl} \mathbf{r}_k & = & \mathbf{b} - A\mathbf{u}_k, \\ \mathbf{u}_{k+1} & = & \mathbf{u}_k + B^{-1}\mathbf{r}_k. \end{array} \right. \end{array} \tag{2.10}$$

One has, for any $k \geq 1$,

$$\mathbf{r}_k = (I - AB^{-1})\mathbf{r}_{k-1} = (I - AB^{-1})^k\mathbf{r}_0, \tag{2.11}$$

where $\mathbf{r}_0 = \mathbf{b} - A\mathbf{u}_0$ is the initial residual. As is well known, this implies

$$\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \|(I - AB^{-1})^k\|. \tag{2.12}$$

On the other hand, sometimes $\mathbf{r}_0$ belongs to a given invariant subspace $\mathcal{S}$ of the right preconditioned matrix $C = AB^{-1}$. Then, all subsequent residuals $\mathbf{r}_k$ will be in $\mathcal{S}$ as well, and the above inequality can be improved:

$$\frac{\|\mathbf{r}_k\|}{\|\mathbf{r}_0\|} \leq \|(I - AB^{-1})^k\|_{\mathcal{S}}. \tag{2.13}$$

Before going further, we should discuss whether this latter relation is realistic when working in finite precision arithmetic. In whole generality, it is not, but, in the next section, we apply this result to a case where $I - AB^{-1}$ has a large null space, and where the relevant invariant subspace $\mathcal{S}$ is in fact the range of $I - AB^{-1}$. In such a context, clearly, significant residual components along $\mathcal{S}^\perp$ cannot be regenerated by rounding errors, and the relevance of the above relation is not questionable.

As is well known, one has

$$\|(I - AB^{-1})^k\|_{\mathcal{S}} \leq \left(\|I - AB^{-1}\|_{\mathcal{S}}\right)^k, \tag{2.14}$$

whereas, because $\|T^k\|_{\mathcal{S}} \leq 2w_{\mathcal{S}}(T^k)$ and $w_{\mathcal{S}}(T^k) \leq \left(w_{\mathcal{S}}(T^k)\right)^k$ hold for any square matrix T with invariant subspace $\mathcal{S}$ [1, 18], one has also

$$\|(I - AB^{-1})^k\|_{\mathcal{S}} \leq 2\left(w_{\mathcal{S}}(I - AB^{-1})\right)^k. \tag{2.15}$$

Tools to analyze $\|I - AB^{-1}\|_{\mathcal{S}}$ and $w_{\mathcal{S}}(I - AB^{-1})$ are given in the next theorem, where (2.19) is strongly inspired by closely related results in [34] for minimal residual iterations.

Theorem 2.2. *Let C be a nonsingular $n \times n$ complex matrix, let $\mathcal{S}$ be an invariant subspace of C, and assume that $\underline{\mu}_{\mathcal{S}}(C) > 0$. One has*

$$W_{\mathcal{S}}(I - C) \subset \Big\{ z \in \mathbb{C} : 1 - \overline{\mu}_{\mathcal{S}}(C) \leq \Re\mathrm{e}(z) \leq 1 - \underline{\mu}_{\mathcal{S}}(C) \\ \text{and } |\Im\mathrm{m}(z)| \leq \max\left(\overline{\mu}_{\mathcal{S}}(\imath C), \overline{\mu}_{\mathcal{S}}(-\imath C)\right) \Big\} \tag{2.16}$$

and

$$W_{\mathcal{S}}(I - C) \subset \left\{ z \in \mathbb{C} : |1 - \tfrac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})} - z| \leq \tfrac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})} \right\}, \tag{2.17}$$

while, letting $\overline{\mu} = \min\left(\overline{\mu}_{\mathcal{S}}(C), \frac{1}{\underline{\mu}_{\mathcal{S}}(C^{-1})}\right)$,

$$w_{\mathcal{S}}(I - C) \leq \left(\left(\max\left(\overline{\mu} - 1, 1 - \underline{\mu}_{\mathcal{S}}(C)\right) \right)^2 + \left(\max\left(\overline{\mu}_{\mathcal{S}}(\imath C), \overline{\mu}_{\mathcal{S}}(-\imath C)\right) \right)^2 \right)^{1/2}. \tag{2.18}$$

If, in addition, $\underline{\mu}_{\mathcal{S}}(C^{-1}) > \frac{1}{2}$, one has further

$$\|I - C\|_{\mathcal{S}} \leq \left(1 - \left(2 - \tfrac{1}{\underline{\mu}_{\mathcal{S}}(C^{-1})}\right)\underline{\mu}_{\mathcal{S}}(C)\right)^{1/2}. \tag{2.19}$$

Proof. The relation (2.16) is straightforward. Regarding (2.17), observe that, by (2.9), $\left\|C - \frac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})} I\right\|_{\mathcal{S}} \leq \frac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})}$, and, hence,

$$W_{\mathcal{S}}\left(C - \tfrac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})} I\right) \subset \left\{ z \in \mathbb{C} : |z| \leq \tfrac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})} \right\}.$$

Therefore, since (for any μ)

$$z \in W(I - C) \iff 1 - z \in W(C) \iff 1 - z - (2\mu)^{-1} \in W(C - (2\mu)^{-1} I),$$

one obtains

$$W_{\mathcal{S}}\left(I - C\right) \subset \left\{ z \in \mathbb{C} : \left|1 - z - \tfrac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})}\right| \leq \tfrac{1}{2\underline{\mu}_{\mathcal{S}}(C^{-1})} \right\};$$

i.e., (2.17).

The inequality (2.18) follows from (2.16) together with the fact that, by (2.17), the real part of any element in $W_{\mathcal{S}}(I - C)$ should also not be smaller than

$1 - 1/\underline{\mu}_{\mathcal{S}}(C^{-1})$; hence we can take the best of this latter and of $1 - \overline{\mu}_{\mathcal{S}}(C)$ to bound the minimal (negative) value that can take the real part.

Next, (2.19) holds because, for any $\mathbf{z} \in S$, and letting $\mathbf{w} = C\mathbf{z}$,

$$\begin{aligned}
\|(I-C)\mathbf{z}\|^2 &= \big((1-C)\mathbf{z}, (I-C)\mathbf{z}\big) \\
&= (\mathbf{z},\mathbf{z}) - 2\Re\mathrm{e}(\mathbf{z}, C\mathbf{z}) + (\mathbf{w},\mathbf{w}) \\
&\le (\mathbf{z},\mathbf{z}) - 2\Re\mathrm{e}(\mathbf{z}, C\mathbf{z}) + \frac{1}{\underline{\mu}_{\mathcal{S}}(C^{-1})}\Re\mathrm{e}(\mathbf{w}, C^{-1}\mathbf{w}) \\
&= (\mathbf{z},\mathbf{z}) - \Big(2 - \frac{1}{\underline{\mu}_{\mathcal{S}}(C^{-1})}\Big)\Re\mathrm{e}(\mathbf{z}, C\mathbf{z}) \\
&\le \left(1 - \Big(2 - \frac{1}{\underline{\mu}_{\mathcal{S}}(C^{-1})}\Big)\underline{\mu}_{\mathcal{S}}(C)\right)(\mathbf{z},\mathbf{z}).
\end{aligned}$$

□

Remark. This theorem is apparently lacking a bound on $w_{\mathcal{S}}(I-C)$ based on (2.17) combined with the bound on the real part in (2.16). In fact, developing the calculation yields the same result as that obtained via (2.19) and the standard inequality $w_{\mathcal{S}}(I-C) \le \|I-C\|_{\mathcal{S}}$ (under the same condition $\underline{\mu}_{\mathcal{S}}(C^{-1}) > \frac{1}{2}$).

The strongest result (2.19) requires the condition $\underline{\mu}_{\mathcal{S}}(C^{-1}) > \frac{1}{2}$, which seems restrictive. However, since $C = AB^{-1}$, it just requires a proper scaling of the preconditioner B. Regarding scaling, observe that, when C is multiplied by some factor α, $\underline{\mu}_{\mathcal{S}}(C)$ is multiplied by α while $\underline{\mu}_{\mathcal{S}}(C^{-1})$ is multiplied by α^{-1}. With respect to the bound (2.19), calculations then show that the optimal scaling is when $\underline{\mu}_{\mathcal{S}}(C^{-1}) = 1$. If this holds, the bound (2.19) becomes equal to $\big(1 - \underline{\mu}_{\mathcal{S}}(C)\big)^{1/2}$, which is also equal to $\big(1 - \underline{\mu}_{\mathcal{S}}(C)\underline{\mu}_{\mathcal{S}}(C^{-1})\big)^{1/2}$. Interestingly, this latter expression, which is scaling invariant, coincides with the upper bound on $\frac{\|\mathbf{r}_{k+1}\|}{\|\mathbf{r}_k\|}$ that is proved in [34] for minimal residual iterations.

The following corollary particularizes the results of Theorem 2.2 for the case where $\underline{\mu}_{\mathcal{S}}(C^{-1}) \ge 1$.

Corollary 2.3. *Let the assumptions of Theorem* 2.2 *hold. If, in addition,* $\underline{\mu}_{\mathcal{S}}(C^{-1}) \ge 1$*, there holds*

$$W_{\mathcal{S}}(I-C) \subset \Big\{ z \in \mathbb{C} : \Re\mathrm{e}(z) \le 1 - \underline{\mu}_{\mathcal{S}}(C) \text{ and } |\tfrac{1}{2} - z| \le \tfrac{1}{2} \Big\}, \tag{2.20}$$

$$w_{\mathcal{S}}(I-C) \le \sqrt{\Big(1 - \underline{\mu}_{\mathcal{S}}(C)\Big)^2 + \Big(\max\big(\overline{\mu}_{\mathcal{S}}(\imath C), \overline{\mu}_{\mathcal{S}}(-\imath C)\big)\Big)^2}, \tag{2.21}$$

and

$$\|I-C\|_{\mathcal{S}} \le \sqrt{1 - \underline{\mu}_{\mathcal{S}}(C)}. \tag{2.22}$$

Proof. In (2.17), observe that, the smaller $\underline{\mu}_{\mathcal{S}}(C^{-1})$, the less restrictive is the stated condition for $z \in W$. Therefore, we obtain a wider set when exchanging $\underline{\mu}_{\mathcal{S}}(C^{-1})$ for a lower bound, justifying (2.20). Other relations follow straightforwardly from (2.18) and (2.19) □

3. Convergence of two-grid iterations

Two-grid methods alternate smoothing iterations and coarse grid corrections. Smoothing iterations are stationary iterations with a simple preconditioner, such as Gauss–Seidel or damped Jacobi. Each time a smoothing iteration is applied, the current residual is multiplied by

$$I - AM^{-1},$$

where M is the used preconditioner; e.g., $M = \omega^{-1}\,\mathrm{diag}(A)$ in case of damped Jacobi.

The coarse grid correction is based on solving a coarse representation of the problem with a reduced number of unknowns $n_c < n$. It thus involves: a restriction matrix R, an $n_c \times n$ matrix that provides the coarse grid residual from the fine grid one; a coarse grid matrix A_c, an $n_c \times n_c$ matrix that defines the coarse representation of the fine grid matrix; and a prolongation matrix P, a $n \times n_c$ matrix that extends on the fine grid the correction computed on the coarse grid. In a two-grid scheme (where the coarse problem is solved exactly, as considered in this work), the current residual is multiplied by

$$I - APA_c^{-1}R$$

each time one performs a coarse grid correction.

Our analysis requires some assumptions that restrict the class of methods taken into consideration. First, we assume that the restriction is the transpose or, more generally (leaving the door open to complex valued prolongations [25]), the conjugate transpose of the prolongation: $R = P^H$. This is satisfied with many geometric and algebraic multigrid methods, but it is worth noting several works that advocate to break this rule for nonsymmetric problems [7–9, 25, 35]. Secondly, we assume that one uses the Galerkin coarse grid matrix; i.e., that $A_c = RAP (= P^H AP)$. This choice is nearly universal with algebraic multigrid methods, but, with geometric multigrid ones, rediscretization on the coarse grid is often preferred for practical reasons.

Finally, our analysis requires that M is Hermitian positive definite, which, in practice, for nonsymmetric problems, amounts to restricting the analysis to damped Jacobi smoothing. Moreover, only a single smoothing step is allowed per global two-grid iteration; that is, we assume that one performs exactly one smoothing iteration in between each coarse grid correction, whereas, usually, one performs two or more. Nevertheless, we believe that this provides a right way to assess more complex methods, at least from a qualitative viewpoint. In particular, algebraic multigrid methods are in principle designed to work well with simple smoothers, and one single step with Jacobi may be seen as the simplest possible smoothing scheme.

It is also worth noting that, as commented in [29, Section 4, Remark 5], the classical algebraic multigrid theory for symmetric matrices is, to a large extent, an analysis that resorts to the same assumptions, supplemented with results showing that the convergence does not deteriorate when using more sophisticated smoothing schemes (while not being able to really predict the resulting improvement). From that point of view, the main difference between our analysis below and this classical theory lies in the lack of such complementary results for nonsymmetric matrices, making the use of supposed better smoothing schemes yet more heuristic.

With the assumptions just discussed, if $\mathbf{r}_0$ is the initial residual, the residual after k two-grid iterations is given by

$$\mathbf{r}_k = \left((I - APA_c^{-1}P^H)(I - AM^{-1})\right)^k \mathbf{r}_0 \tag{3.23}$$

if one starts with a smoothing step, and by

$$\mathbf{r}_k = \left((I - AM^{-1})(I - APA_c^{-1}P^H)\right)^k \mathbf{r}_0 \tag{3.24}$$

if one starts with a coarse grid correction. Letting

$$T = (I - APA_c^{-1}P^H)(I - AM^{-1}) \tag{3.25}$$

be the iteration matrix for the two-grid method, one may write, if one starts with a smoothing step,

$$\mathbf{r}_k = T^{k-1}\widehat{\mathbf{r}}_0, \tag{3.26}$$

where $\widehat{\mathbf{r}}_0 = T\mathbf{r}_0$, whereas, if one starts with a coarse grid correction, one has

$$\mathbf{r}_k = (I - AM^{-1})T^{k-1}\widehat{\mathbf{r}}_0, \tag{3.27}$$

where $\widehat{\mathbf{r}}_0 = (I - APA_c^{-1}P^H)\mathbf{r}_0$.

As seen by expanding (3.25), $T = I - C$, where

$$C = A\left(PA_c^{-1}P^H + M^{-1} - PA_c^{-1}P^HAM^{-1}\right) \tag{3.28}$$

has the form of a right preconditioned matrix. Hence, one may apply Theorem 2.2 to analyze $\|T^k\widehat{\mathbf{r}}_0\| = \|(I-C)^k\widehat{\mathbf{r}}_0\|$, which is the main factor governing the convergence. In this view, note that, assuming $A_c = P^HAP$ (i.e., that one uses the Galerkin coarse grid matrix), one has $P^H(I - APA_c^{-1}P^H) = 0$, implying also that $P^HT = 0$. This shows that, in both cases considered above, $\widehat{\mathbf{r}}_0 \in \mathcal{P}^\perp$, where $\mathcal{P}$ stands for the range of P. Moreover, $P^HT = 0$ further entails that $\mathcal{P}^\perp$ is an invariant subspace of T and therefore of C. It thus follows that the assessment of $\|(I - C)^k\|_{\mathcal{P}^\perp}$ is sufficient.

Then, Theorem 2.2 provides convergence estimates if one has proper bounds for $\underline{\mu}_{\mathcal{P}^\perp}(C)$ and $\underline{\mu}_{\mathcal{P}^\perp}(C^{-1})$. Such bounds are given in the following theorem. These results are not based on the standard Euclidean inner product, but consider instead

$$(\mathbf{v}, \mathbf{w})_{M^{-1}} = \mathbf{v}^H M^{-1}\mathbf{w} \tag{3.29}$$

and the induced norm

$$\|\mathbf{v}\|_{M^{-1}} = \left(\mathbf{v}^H M^{-1}\mathbf{v}\right)^{1/2}, \tag{3.30}$$

which is also the energy norm associated with M^{-1}. Note that, for any $\mathbf{r}_0$ and $\mathbf{r}_k$, there holds

$$\left(\|M\|_2\|M^{-1}\|_2\right)^{-1} \frac{\|\mathbf{r}_k\|_{M^{-1}}}{\|\mathbf{r}_0\|_{M^{-1}}} \le \frac{\|\mathbf{r}_k\|_2}{\|\mathbf{r}_0\|_2} \le \left(\|M\|_2\|M^{-1}\|_2\right) \frac{\|\mathbf{r}_k\|_{M^{-1}}}{\|\mathbf{r}_0\|_{M^{-1}}}. \tag{3.31}$$

This kind of equivalence between norms is quite often used in the literature. In the present context, preconditioners M used for smoothing are typically well conditioned matrices. It then follows that the above relations do not only ensure qualitative equivalence, but also show that the related convergence bounds remain quantitatively meaningful even when transposed to the Euclidean norm.

Theorem 3.1. *Let A be a nonsingular $n\times n$ matrix and let M be an $n\times n$ Hermitian positive definite matrix. Let P be an $n \times n_c$ matrix of rank $n_c < n$ and such that $A_c = P^H AP$ is nonsingular, and let C be the matrix defined by (3.28). Let $\mathcal{P}$ denote the range of P.*

With respect to the inner product (3.29), one has

$$W_{\mathcal{P}^\perp}(C) = \left\{ \frac{\mathbf{z}^H A\mathbf{z}}{\mathbf{z}^H M(I - PM_c^{-1}P^H M)\mathbf{z}} : \mathbf{z} \in \mathbb{C}^n\backslash\{0\} \text{ and } P^H A\mathbf{z} = 0 \right\}, \tag{3.32}$$

whereas, letting $A_H = \frac{1}{2}(A + A^H)$ and $M_c = P^H MP$,

$$\underline{\mu}_{\mathcal{P}^\perp}(C) \geq \left(\lambda_{\max}\Big(A_H^{-1} M \left(I - PM_c^{-1}P^H M\right) \Big) \right)^{-1} \tag{3.33}$$

$$= \lambda_{\min}(I - T_{A_H}) \tag{3.34}$$

$$\geq 1 - \rho(T_{A_H}), \tag{3.35}$$

where

$$T_{A_H} = \left(I - A_H P(P^H A_H P)^{-1}P^H\right)\left(I - A_H M^{-1}\right) \tag{3.36}$$

is the iteration matrix for the two-grid method applied to A_H that uses the same prolongation P and smoother M.

On the other hand,

$$\underline{\mu}_{\mathcal{P}^\perp}(C^{-1}) = \underline{\mu}_{\mathcal{P}^\perp}(MA^{-1}) \tag{3.37}$$

$$\geq \underline{\mu}_{\mathbb{C}^n}(MA^{-1}) \tag{3.38}$$

$$= \lambda_{\min}\Big(\tfrac{1}{2}M^{1/2}\big(A^{-1} + A^{-H}\big)M^{1/2}\Big), \tag{3.39}$$

and both $\underline{\mu}_{\mathcal{P}^\perp}(C)$ and $\underline{\mu}_{\mathcal{P}^\perp}(C^{-1})$ are positive if A_H is positive definite.

Proof. First, note that, as seen in [29], C can be written $C = AB^{-1}$ with $B = M - MPM_c^{-1}P^H(M - A)$. Hence, $C^{-1} = BA^{-1}$ satisfies

$$C^{-1}\mathbf{v} = M\left(I - PM_c^{-1}P^H M\right)A^{-1}\mathbf{v} \qquad \forall \mathbf{v} \in \mathcal{P}^\perp. \tag{3.40}$$

Next, observe that $P^H A\mathbf{w} = \mathbf{0}$ implies $\mathbf{w} \notin \mathcal{P}$, because otherwise one would have $P^H AP\mathbf{y} = A_c\mathbf{y} = \mathbf{0}$ for some nontrivial $\mathbf{y}$. Moreover, $\mathcal{P}^\perp$ being an invariant subspace of C, $\mathbf{v} \in \mathcal{P}^\perp \iff \mathbf{w} = C\mathbf{v} \in \mathcal{P}^\perp$. Then, using (3.40), one obtains

$$\begin{aligned}
&\left\{ \tfrac{(\mathbf{v},C\mathbf{v})_{M^{-1}}}{(\mathbf{v},\mathbf{v})_{M^{-1}}} : \mathbf{v} \in \mathcal{P}^\perp\backslash\{0\} \right\} \\
&\quad = \left\{ \tfrac{(C^{-1}\mathbf{w})^H M^{-1}\mathbf{w}}{(C^{-1}\mathbf{w})^H M^{-1}C^{-1}\mathbf{w}} : \mathbf{w} \in \mathcal{P}^\perp\backslash\{0\} \right\} \\
&\quad = \left\{ \tfrac{\left(M\left(I-PM_c^{-1}P^H M\right)A^{-1}\mathbf{w}\right)^H M^{-1}\mathbf{w}}{\left(M\left(I-PM_c^{-1}P^H M\right)A^{-1}\mathbf{w}\right)^H\left(I-PM_c^{-1}P^H M\right)A^{-1}\mathbf{w}} : \mathbf{w} \in \mathcal{P}^\perp\backslash\{0\} \right\} \\
&\quad = \left\{ \tfrac{\left(A^{-1}\mathbf{w}\right)^H\mathbf{w}}{(A^{-1}\mathbf{w})^H M\left(I-PM_c^{-1}P^H M\right)A^{-1}\mathbf{w}} : \mathbf{w} \in \mathcal{P}^\perp\backslash\{0\} \right\} \\
&\quad = \left\{ \tfrac{\mathbf{z}^H A\mathbf{z}}{\mathbf{z}^H M\left(I-PM_c^{-1}P^H M\right)\mathbf{z}} : \mathbf{z} \in \mathbb{C}^n\backslash\{0\} \text{ and } P^H A\mathbf{z} = 0 \right\};
\end{aligned}$$

that is, (3.32).

Using these results, we further obtain

$$\begin{aligned}
\underline{\mu}_{\mathcal{P}^\perp}(C) &= \min_{\substack{\mathbf{z}\in\mathbb{C}^n\setminus\{0\}\\ P^HA\mathbf{z}=0}} \frac{\Re\mathrm{e}\,(\mathbf{z},A\mathbf{z})}{\left(\mathbf{z},M\left(I-PM_c^{-1}P^HM\right)\mathbf{z}\right)}\\
&\geq \min_{\mathbf{z}\notin\mathcal{P}} \frac{\Re\mathrm{e}\,(\mathbf{z},A\mathbf{z})}{\left(\mathbf{z},M\left(I-PM_c^{-1}P^HM\right)\mathbf{z}\right)}\\
&= \left(\lambda_{\max}\Big(\left(\tfrac{1}{2}(A+A^H)\right)^{-1}M\left(I-PM_c^{-1}P^HM\right)\Big)\right)^{-1}.
\end{aligned}$$

Thus, (3.33) is proved whereas (3.34), (3.35) further follow from [28, Theorem 2.1].

On the other hand, using again (3.40),

$$\begin{aligned}
\underline{\mu}_{\mathcal{P}^\perp}(C^{-1}) &= \min_{\mathbf{z}\in\mathcal{P}^\perp\setminus\{0\}} \frac{\Re\mathrm{e}(\mathbf{z},C^{-1}\mathbf{z})_{M^{-1}}}{(\mathbf{z},\mathbf{z})_{M^{-1}}}\\
&= \min_{\mathbf{z}\in\mathcal{P}^\perp\setminus\{0\}} \frac{\Re\mathrm{e}\left(\mathbf{z},\left(I-PM_c^{-1}P^HM\right)A^{-1}\mathbf{z}\right)}{(\mathbf{z},M^{-1}\mathbf{z})}\\
&= \min_{\mathbf{z}\in\mathcal{P}^\perp\setminus\{0\}} \frac{\Re\mathrm{e}(\mathbf{z},A^{-1}\mathbf{z})}{(\mathbf{z},M^{-1}\mathbf{z})}\\
&\geq \min_{\mathbf{z}\in\mathbb{C}^n\setminus\{0\}} \frac{\Re\mathrm{e}(\mathbf{z},A^{-1}\mathbf{z})}{(\mathbf{z},M^{-1}\mathbf{z})}\\
&= \lambda_{\min}\Big(\tfrac{1}{2}M^{1/2}\big(A^{-1}+A^{-H}\big)M^{1/2}\Big),
\end{aligned}$$

proving thus (3.37), (3.38), and (3.39). □

In many cases it should not be too difficult to assess the quantities in the right hand side of (3.33)/(3.34) and (3.38)/(3.39).

Regarding (3.33)/(3.34), observe that T_H is the iteration matrix for the same two-grid method applied to A_H, the Hermitian part of A, the *same* meaning using the same smoother, prolongation and restriction, while defining the coarse grid matrix with the same Galerkin rule. The problem of solving a system with A_H using this method can be seen as a *companion symmetric problem* for the problem at hand. Such symmetric problems are well understood and much easier to analyze. The iteration matrix T_H is not symmetric, but its eigenvalues are real and well characterized [29]. They can be analyzed with (rigorous) Fourier analysis [39] or even LFA, which has always been found reliable for symmetric problems. Other results more adapted to algebraic multigrid methods [26, 33] can be used as well.

Regarding (3.38)/(3.39), we first note that $\underline{\mu}_{\mathcal{P}^\perp}(MA^{-1})$ varies linearly with the scaling of M. Hence, the condition $\underline{\mu}_{\mathcal{P}^\perp}(C^{-1}) \geq 1$ that is useful when applying Theorem 2.2 can be enforced with a proper scaling; that is, with a proper selection of the damping parameter ω in case one uses Jacobi. That said, as pointed out in [29], $\lambda_{\min}\big(\frac{1}{2}M^{1/2}(A^{-1}+A^{-H})M^{1/2}\big)$ is in general smaller than $\lambda_{\min}\big(M^{1/2}(A_H)^{-1}M^{1/2}\big)$, which can be easier to assess. It is then useful to recall the following result from [27], which, combined with Lemma 2.1, leads directly to a bound on $\underline{\mu}_{\mathbb{C}^n}(MA^{-1})$ when $M=\omega^{-1}\operatorname{diag}(A)$.

Theorem 3.2. *Let A be a nonsingular M-matrix with row and column sums both nonnegative, and let $D=\operatorname{diag}(A)$ and $M=\omega^{-1}D$ for some positive parameter ω.*

One has

$$\|I - D^{-1/2}AD^{-1/2}\|_2 = \|I - AD^{-1}\|_{D^{-1}} = \|I - \omega^{-1}AM^{-1}\|_{M^{-1}} \leq 1 \tag{3.41}$$

and, with respect to the inner product (3.29)*,*

$$\underline{\mu}_{\mathbb{C}^n}(MA^{-1}) \geq (2\omega)^{-1}. \tag{3.42}$$

Proof. The inequality $\|I - D^{-1/2}AD^{-1/2}\|_2 \leq 1$ is proved in [27, Lemma 3.1], whereas the equalities in (3.41) are straightforward from the matrix norm definition. The inequality (3.42) is then obtained by applying Lemma 2.1 □

Hence, under the stated assumptions, it suffices to select $\omega = \frac{1}{2}$ to have $\underline{\mu}_{\mathbb{C}^n}(MA^{-1}) \geq 1$ and therefore, by (3.38), to guarantee the scaling condition $\underline{\mu}_{\mathcal{S}}(C^{-1}) \geq 1$. When this latter holds, the combined use of Theorem 2.2 and Corollary 2.3 yields appealing results summarized in the following corollary.

Corollary 3.3. *Let the assumptions of Theorem* 3.1 *hold, and let* $T = I - C$ *be defined by* (3.25)*. If, in addition,* $\underline{\mu}_{\mathcal{S}}(C^{-1}) \geq 1$*, there holds*

$$W_{\mathcal{P}^\perp}(T) \subset \left\{ z \in \mathbb{C} : \Re\mathrm{e}(z) \leq \rho(T_H) \text{ and } |\tfrac{1}{2} - z| \leq \tfrac{1}{2} \right\}, \tag{3.43}$$

$$w_{\mathcal{P}^\perp}(T) \leq \left(\left(1 - \underline{\mu}_{\mathcal{P}^\perp}(C)\right)^2 + \left(\max\left(\overline{\mu}_{\mathcal{P}^\perp}(\imath C), \overline{\mu}_{\mathcal{P}^\perp}(-\imath C)\right)\right)^2 \right)^{1/2} \tag{3.44}$$

$$\leq \left(\left(\rho(T_H)\right)^2 + \left(\max\left(\overline{\mu}_{\mathcal{P}^\perp}(\imath C), \overline{\mu}_{\mathcal{P}^\perp}(-\imath C)\right)\right)^2 \right)^{1/2}, \tag{3.45}$$

and

$$\|T\|_{\mathcal{P}^\perp} \leq \left(1 - \underline{\mu}_{\mathcal{P}^\perp}(C)\right)^{1/2} \tag{3.46}$$

$$\leq \sqrt{\rho(T_H)}. \tag{3.47}$$

Some further insight can be gained by discussing these results in the well-known case of a Hermitian positive definite A, which entails that T coincide with T_H. Then, as stated above, the eigenvalues of T are real and, further, the restriction of T to $\mathcal{P}^\perp$ is self-adjoint.[1] It follows that $W_{\mathcal{P}^\perp}(T)$ is a segment of the real axis (which can also be seen from (3.32)), whereas $\|T\|_{\mathcal{P}^\perp} = w_{\mathcal{P}^\perp}(T) = \rho(T)$. Therefore, since $\overline{\mu}_{\mathcal{P}^\perp}(\pm\imath C) = 0$, (3.44), (3.45) are sharp while (3.46) is not, and relying on this latter leads one to overestimate the number of iterations by a factor of 2.

Going to a nonsymmetric problem, this factor of 2 may be lowered because then the right hand side of (3.44) is larger than the square of the right hand side of (3.46). However, other sources of inaccuracy are introduced, because (3.44) is not necessarily sharp while, in general, $\rho(T_H)$ will be strictly larger than $1 - \underline{\mu}_{\mathcal{P}^\perp}(C)$.

Hence, the bound (3.47) will often lead to an overestimation of the number of iterations. However, in the experiments we made, we always found that $\rho(T_H)$ is a reliable qualitative indicator. When it is significantly below 1, minimal convergence properties are guaranteed via (3.47), hence the right message is delivered even though the actual convergence is faster than predicted by the bound. On the other hand, in all cases where $\rho(T_H) \approx 1$, we always observed unacceptably slow convergence, and how faster it can be compared to what is predicted by the bound (3.47) has no practical importance.

[1] As seen by comparing, with the help of (3.28), $(C\mathbf{v}, \mathbf{w})_{M^{-1}}$ and $(\mathbf{v}, C\mathbf{w})_{M^{-1}}$ for $\mathbf{v}, \mathbf{w} \in \mathcal{P}^\perp$.

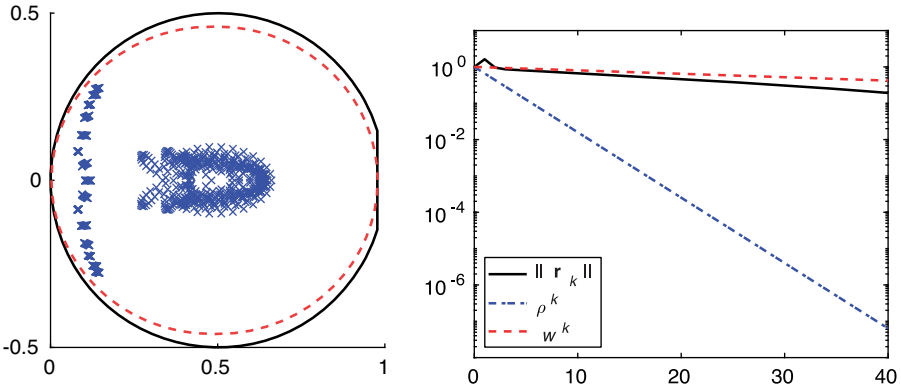

FIGURE 2. Problem (1.1), (1.2) with $\alpha = 50$, $\beta = 1$. Left: spectrum of the iteration matrix T (×), with the boundary of $W_{\mathcal{P}^\perp}(T)$ (- -), and the bound on this latter based on (3.43) (—). Right: residual norm as a function of the number of iterations, and the two estimations $(\rho(T))^k$ and $(w_{\mathcal{P}^\perp}(T))^k$, where k is the iteration index.

Going back to the example presented in Section 1, first note that the system matrix is an M-matrix with both row and column sums nonnegative, while we used damped Jacobi. Hence, (3.42) holds entailing that Corollary 3.3 applies for the selected damping parameter $\omega = \frac{1}{2}$. On the other hand, A_H is associated with the symmetric part of the stencil (1.1):

$$\begin{bmatrix} & -\frac{\beta+1}{2} & \\ -\frac{\alpha+1}{2} & 2+\alpha+\beta & -\frac{\alpha+1}{2} \\ & -\frac{\beta+1}{2} & \end{bmatrix}. \tag{3.48}$$

This corresponds to finite difference for anisotropic diffusion. It is well known that, with Jacobi smoothing, multigrid with geometric-based prolongation and restriction (as considered in the numerical example) performs well only in near isotropic cases; i.e., only when $\alpha \approx \beta$, whereas, in the example of Figure 1, we have used $\alpha = 50$ and $\beta = 1$. Hence, the disappointing performance could have been anticipated just applying ground knowledge to the companion symmetric problem.

In Figure 2, for the same example, in the left picture, we draw the field of values besides the eigenvalues, whereas in the right picture, the residual evolution is further compared with $\big(w_{\mathcal{P}^\perp}(T)\big)^k$. One clearly sees that the field of values spreads significantly away from the convex hull of the eigenvalues, while being nicely estimated by the theoretical results. On the other hand, the numerical radius offers a convergence estimate that is not only rigorous, but also realistic, while, as already noted in Section 1, the spectral radius estimate is completely meaningless.

4. LFA AND NUMERICAL EXAMPLES

LFA or local mode analysis [2,5,39,42] is used to assess geometric two-grid methods applied to model problems with constant coefficients on a regular grid. The

actual grid is extended to an infinite grid so as to get rid of the boundary conditions that prevent the matrix from being diagonal in the Fourier basis. Nearly equivalently, one may stay with a finite grid but modify the true boundary conditions for periodic ones. In both cases, the structure of the prolongation and restriction operators used with classical geometric multigrid methods entails that, in the Fourier basis, the two-grid iteration matrix is block diagonal with, in two dimensional cases, 4×4 diagonal blocks, independently of the grid size. That is why there is nearly no difference between infinite grid and finite grid with periodic boundary conditions. In the former case, there is a continuum of Fourier modes, whereas in the latter one has a discrete sampling of these modes. But even in the infinite grid case, one has, in practice, to discretize the Fourier space to run an actual computation. Note, however, that LFA may also be combined with analytic calculations, in which case one naturally works with the continuum associated with the infinite grid.

We refer to [39, Chapter 4] and [42] for practical details about LFA implementation. To be complete regarding the description of experimental framework, when, below, we refer to LFA computation for the problems (1.1), (1.2) with a given grid size, we mean that the same PDE (1.2) is considered with the same discretization, except that Dirichlet boundary conditions are exchanged for periodic ones ($u(x,0) = u(x,L)$ for $0 \leq x \leq 1$ and $u(0,y) = u(L,y)$ for $0 \leq y \leq 1$). It follows that the system matrix is then $m^2 \times m^2$, with a singular mode which, as usual with LFA, is not taken into consideration when displaying spectra or computing spectral radii. (Because LFA is only used for eigenvalue or field of values computation, the right hand side is irrelevant, as well as the discussion of the iterative method in the presence of a singularity.)

When the matrix is diagonal in the Fourier basis with the actual boundary conditions, LFA often gives the same results as rigorous Fourier analysis (see [32] for a thorough analysis of this point). If not (as is always the case for nonsymmetric problems), one traditionally hopes that the change of boundary conditions will not perturb much the eigenvalues distribution of the two-grid iteration matrix, because effects are supposedly limited to boundaries; see [4, 5, 36] for further discussion, motivation, and validation.

Regarding nonsymmetric problems, however, it is mentioned in several works that the true spectrum may differ significantly from its LFA approximation [16, 17, 30]. In [5, Section 7.5], this is associated with boundary effects that may be visible far in the interior for nonelliptic or singularly perturbed equations.

Here we consider this problem from a different angle. We first observe that a matrix which is diagonal in the Fourier basis is a normal matrix. Hence, with LFA one approximates the actual system matrix by a nearby normal one. Thus, at first sight, LFA is not a proper tool to analyze nonnormality effects, since these are erased from the original problem.

Going to the details, the two-grid iteration matrix associated with an LFA approximation is not strictly speaking a normal matrix, even though the system matrix is normal. However, in the numerical experiments we performed, we never saw a significant difference between its field of values and the convex hull of its eigenvalues. Consider then cases where, as in Figure 2, there is a significant difference between the eigenvalues and the field of values for the corresponding actual problem. Clearly, the LFA approximation can give a good picture of *either* the eigenvalue distribution *or* of the field of values, *but never of both.* In other words, if

it succeeds to predict the eigenvalue distribution and hence the asymptotic convergence, it will be misleading regarding the actual convergence for a realistic number of iterations. In view of our results, we advocate that the relevance of LFA should be assessed on its capacity to approximate the true field of values, independently of the location of the true eigenvalues.

In that context, one may see as good news the observed discrepancy between true and LFA computed spectra as, e.g., in [16, 30]. However, on which basis could we hope that LFA does any better regarding the field of values? A few heuristic arguments go in that direction. First, numerical comparisons between actual convergence and estimates from infinite grids show that these latter can be accurate for parabolic problems [23, 40, 41].

Next, in [40], a connection is made with the theory of Toeplitz matrices and operators, and the related analysis of some (nonmultigrid) iterative methods [24]. As shown in [38, Chapter II], there is a discontinuity between the spectrum of a banded Toeplitz matrix, whatever its size, and the spectrum of the Toeplitz operator obtained at the limit when the size goes up to infinity. However, the pseudo-spectrum is continuous and close to the spectrum of the infinite operator. That is, the eigenvalues of this latter give an accurate picture of the pseudo-spectrum of large finite Toeplitz matrices with the same coefficients; see [38, Chapter II] for details and additional references. Since matrices from discretized PDEs are close to Toeplitz, this suggests that analyzing their infinite grid counterpart (as LFA does) gives a correct picture of the pseudo-spectrum but wrong information about the precise location of the eigenvalues. Because, for nonnormal matrices, the pseudo-spectrum offers a much more reliable convergence indicator than the spectrum [38, Chapter VI], one may then expect that the eigenvalue computation on infinite grids leads to convergence estimates that are at least qualitatively correct.

Similar conclusions can also be reached by considering the field of values. Clearly, the system matrix obtained on a given finite grid is a principal submatrix of the system matrix on any extended grid, implying that the field of values of the former is a subset of the field of values of the latter. Hence, the field of values for the infinite operator should provide a bounding set for the true field of values, and, thus, a worst case convergence estimate. Moreover, the field of values is explicitly known for the special case of a Jordan block [22], and it may be inferred that it smoothly converges towards that of the limiting operator as the size goes up to infinity.

Now, it should be clear that the above reasonings are heuristic, no matter that the supporting theory is rigorous. Indeed, the system matrix is often close to Toeplitz (say, it can be block Toeplitz with Toeplitz blocks), but it is rarely exactly Toeplitz. Moreover, even if it was, the two-grid iteration matrix would in general not be a Toeplitz matrix. Similarly, that the system matrix is a principal submatrix of a bigger matrix does not imply that the associated two-grid iteration matrices satisfy a similar imbrication property.

Hence, numerical experiments are needed to enlighten the discussion. Clearly, what is missing in Figure 2, is: in the left picture, the LFA approximation of the spectrum/field of values, obtained by changing the boundary conditions for periodic ones; in the right picture, the convergence estimate based on the LFA approximation of the spectral/numerical radius. These gaps are filled in Figure 3, where we also consider other values for the main parameter α, β in the stencil (1.1).

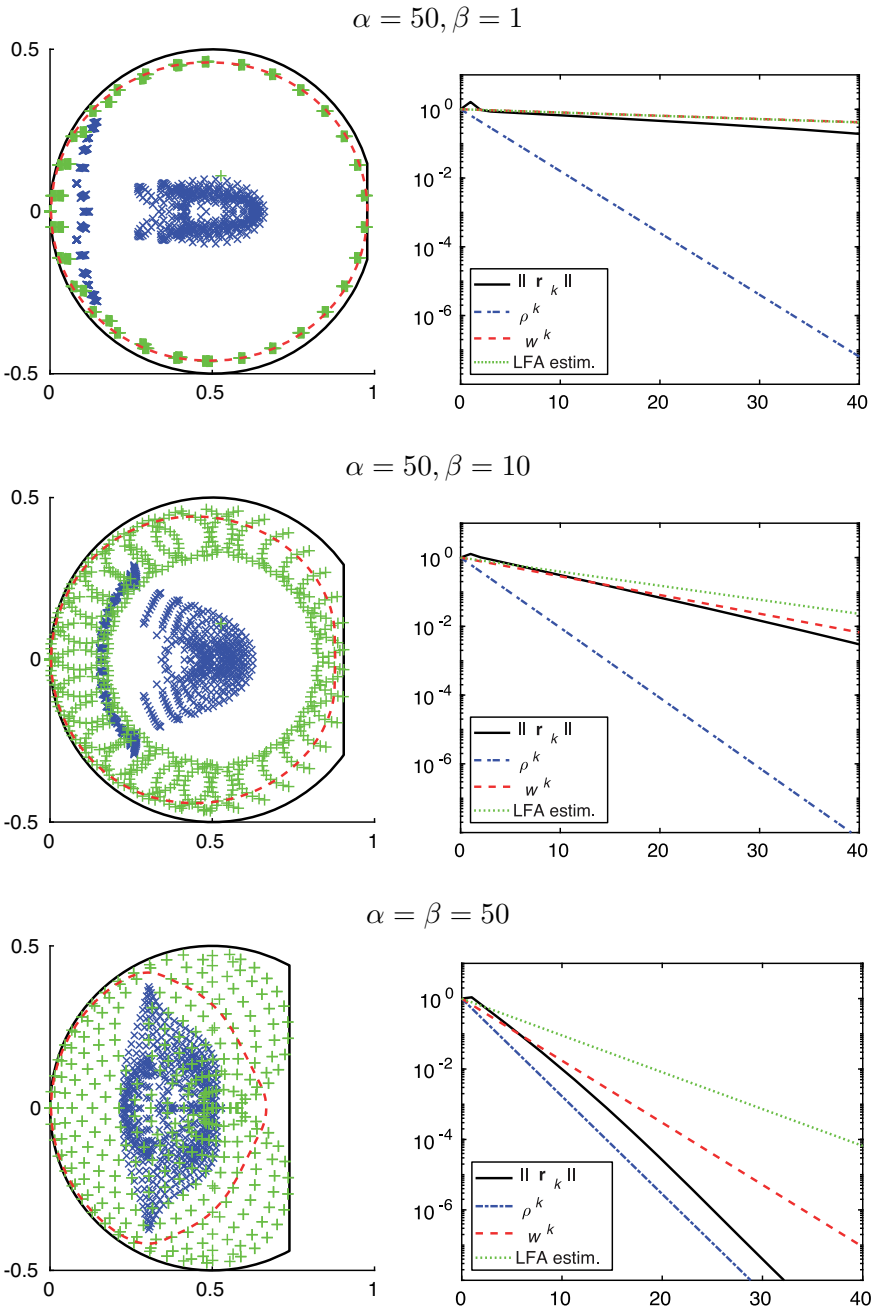

FIGURE 3. Problem (1.1), (1.2). Left: spectrum of the iteration matrix T ($\times$), together with its LFA approximation ($+$), the boundary of $W_{\mathcal{P}^\perp}(T)$ (- -), and the bound on this latter based on (2.20) (—). Right: residual norm as a function of the number of iterations, and the three estimations $(\rho(T))^k$, $(w_{\mathcal{P}^\perp}(T))^k$, and $(\widehat{\rho})^k$, where $\widehat{\rho}$ is the LFA approximation of the numerical/spectral radius, and where k is the iteration index.

Going back to the general discussion of our results, one sees that the numerical radius correctly predicts the actual convergence in all tested configurations. On the other hand, the theoretical bound based on Theorems 2.2, 3.1, and 3.2, as summarized in Corollary 3.3, is qualitatively correct but not always very accurate. Going to the details, one sees that the vertical black line in the left pictures in Figure 3, which represents the upper bound on the real part in (3.43), is always a good approximation of the maximum real part over all elements in $W_{\mathcal{P}^\perp}(T)$. However, the control of the imaginary extension by the second inequality in (3.43) may be more loose. In fact, it remains the same however far we are from a symmetric example with all eigenvalues real. As commented at the end of Section 3, the related loss of accuracy is, however, limited, in the sense that the convergence prediction remains qualitatively correct.

Regarding LFA, it is also in all tested examples qualitatively correct with respect to the objective of predicting the true field of values and the associated numerical radius. In fact, in all considered cases, the convex hull of the LFA spectrum is inside the region defined by the right hand side of (3.43), and is never significantly different. This means that an estimate of the numerical radius based on LFA computation is only slightly more accurate than our theoretical bounds . Hence, in case these latter apply, LFA brings little added value because it remains heuristic.

This heuristic nature is apparently confirmed with the next experiment, reported in Figure 4, where we check, as the grid size is doubled, the evolution of the true spectrum, of the LFA spectrum, and of the field of values boundary. For a full confirmation of the heuristic arguments above, one would need that the boundary of the field of values converges towards the convex hull of the LFA spectrum, which is close to the field of values for the infinite grid. However, that does not seem to be the case, all reported quantities varying only little with the grid size. Because of the heaviness of such computations, experiments are however limited to $m = 24$ and $m = 48$, which is too small to draw definitive conclusions about the asymptotic behavior.

5. Conclusions

We have seen that, for nonnormal matrices, the eigenvalues of the two-grid iteration matrix may be a completely wrong indicator, unable to predict severe convergence failures. Opposite to this, the field of values of this matrix leads to rigorous convergence estimates that have also been found accurate in several numerical examples. When the two-grid method involves only a single smoothing step with Jacobi, we proved that this field of values is located in a region of the complex plane described by two easy to assess quantities, related to a companion symmetric problem. The convergence bounds obtained on this basis, besides being rigorous, have also been found offering qualitatively correct predictions.

The role of LFA has further been discussed in light of these results. We advocate that LFA should be assessed on its ability to predict the extent of the field of values of the iteration matrix, regardless of the actual location of its eigenvalues. In the few numerical examples we tested, LFA delivers results roughly equivalent to those obtained with our theoretical bounds; that is, LFA never leads to overestimate the convergence speed and is always qualitatively correct. This is good news because LFA is more widely applicable than the theoretical bounds. However, one should be cautious, because LFA remains more heuristic than in the symmetric case. This subject thus certainly deserves to be investigated further.

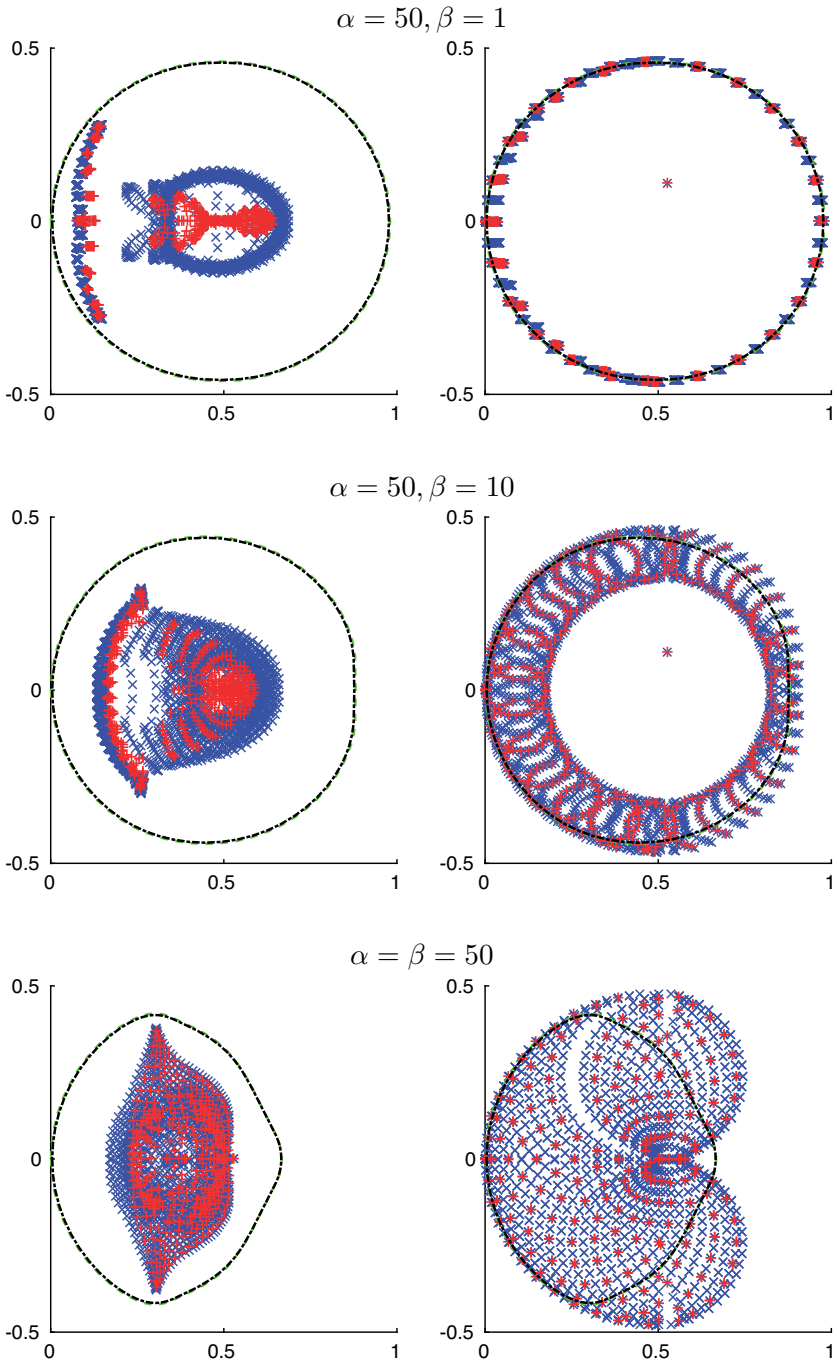

FIGURE 4. Problem (1.1), (1.2). Left and right: boundary of $W_{\mathcal{P}^\perp}(T)$ for $m = 24$ (- -) and $m = 48$ ($- \cdot - \cdot$). Left: spectrum of the iteration matrix T for $m = 24$ (+) and $m = 48$ ($\times$). Right: LFA approximation of the spectrum of the iteration matrix T for $m = 24$ (+) and $m = 48$ ($\times$);

Acknowledgment

The author deeply thanks the anonymous referee whose writing suggestions were very helpful to improve the readability of the paper.

References

[1] C. A. Berger and J. G. Stampfli, *Mapping theorems for the numerical range*, Amer. J. Math. **89** (1967), 1047–1055, DOI 10.2307/2373416. MR0222694

[2] A. Brandt, *Multi-level adaptive solutions to boundary-value problems*, Math. Comp. **31** (1977), no. 138, 333–390, DOI 10.2307/2006422. MR0431719

[3] A. Brandt, *Algebraic multigrid theory: the symmetric case*, Appl. Math. Comput. **19** (1986), no. 1-4, 23–56, DOI 10.1016/0096-3003(86)90095-0. Second Copper Mountain conference on multigrid methods (Copper Mountain, Colo., 1985). MR849831

[4] A. Brandt, *Rigorous quantitative analysis of multigrid. I. Constant coefficients two-level cycle with L_2-norm*, SIAM J. Numer. Anal. **31** (1994), no. 6, 1695–1730, DOI 10.1137/0731087. MR1302681

[5] A. Brandt and O. E. Livne, *Multigrid techniques—1984 guide with applications to fluid dynamics*, Classics in Applied Mathematics, vol. 67, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2011. Revised edition of the 1984 original [MR0772748]. MR3396211

[6] A. Brandt and I. Yavneh, *On multigrid solution of high-Reynolds incompressible entering flows*, J. Comput. Phys. **101** (1992), no. 1, 151–164, DOI 10.1016/0021-9991(92)90049-5. MR1173343

[7] M. Brezina, T. Manteuffel, S. McCormick, J. Ruge, and G. Sanders, *Towards adaptive smoothed aggregation* (αSA) *for nonsymmetric problems*, SIAM J. Sci. Comput. **32** (2010), no. 1, 14–39, DOI 10.1137/080727336. MR2599765

[8] P. M. de Zeeuw, *Matrix-dependent prolongations and restrictions in a blackbox multigrid solver*, J. Comput. Appl. Math. **33** (1990), no. 1, 1–27, DOI 10.1016/0377-0427(90)90252-U. MR1081238

[9] J. E. Dendy Jr., *Black box multigrid for nonsymmetric problems*, Appl. Math. Comput. **13** (1983), no. 3-4, 261–283, DOI 10.1016/0096-3003(83)90016-4. MR726637

[10] V. A. Dobrev, Tz. Kolev, N. A. Petersson, and J. B. Schroder, *Two-level convergence theory for multigrid reduction in time (MGRIT)*, SIAM J. Sci. Comput. **39** (2017), no. 5, S501–S527, DOI 10.1137/16M1074096. MR3716570

[11] M. Eiermann, *Fields of values and iterative methods*, Linear Algebra Appl. **180** (1993), 167–197, DOI 10.1016/0024-3795(93)90530-2. MR1206415

[12] S. C. Eisenstat, H. C. Elman, and M. H. Schultz, *Variational iterative methods for nonsymmetric systems of linear equations*, SIAM J. Numer. Anal. **20** (1983), no. 2, 345–357, DOI 10.1137/0720023. MR694523

[13] R. D. Falgout, S. Friedhoff, Tz. V. Kolev, S. P. MacLachlan, and J. B. Schroder, *Parallel time integration with multigrid*, SIAM J. Sci. Comput. **36** (2014), no. 6, C635–C661, DOI 10.1137/130944230. MR3499068

[14] R. D. Falgout and P. S. Vassilevski, *On generalizing the algebraic multigrid framework*, SIAM J. Numer. Anal. **42** (2004), no. 4, 1669–1693, DOI 10.1137/S0036142903429742. MR2114296

[15] R. D. Falgout, P. S. Vassilevski, and L. T. Zikatanov, *On two-grid convergence estimates*, Numer. Linear Algebra Appl. **12** (2005), no. 5-6, 471–494, DOI 10.1002/nla.437. MR2150164

[16] S. Friedhoff and S. MacLachlan, *A generalized predictive analysis tool for multigrid methods*, Numer. Linear Algebra Appl. **22** (2015), no. 4, 618–647, DOI 10.1002/nla.1977. MR3367826

[17] S. Friedhoff, S. MacLachlan, and C. Börgers, *Local Fourier analysis of space-time relaxation and multigrid schemes*, SIAM J. Sci. Comput. **35** (2013), no. 5, S250–S276, DOI 10.1137/120881361. MR3120772

[18] M. Goldberg and E. Tadmor, *On the numerical radius and its applications*, Linear Algebra Appl. **42** (1982), 263–284, DOI 10.1016/0024-3795(82)90155-0. MR656430

[19] A. Greenbaum, V. Pták, and Z. Strakoš, *Any nonincreasing convergence curve is possible for GMRES*, SIAM J. Matrix Anal. Appl. **17** (1996), no. 3, 465–469, DOI 10.1137/S0895479894275030. MR1397238

[20] A. Greenbaum and Z. Strakoš, *Matrices that generate the same Krylov residual spaces*, Recent advances in iterative methods, IMA Vol. Math. Appl., vol. 60, Springer, New York, 1994, pp. 95–118, DOI 10.1007/978-1-4613-9353-5_7. MR1332745

[21] W. Hackbusch, *Multigrid Methods and Applications*, Springer Series in Computational Mathematics, vol. 4, Springer-Verlag, Berlin, 1985. MR814495

[22] R. A. Horn and C. R. Johnson, *Topics in matrix analysis*, Cambridge University Press, Cambridge, 1991. MR1091716

[23] J. Janssen and S. Vandewalle, *Multigrid waveform relaxation on spatial finite element meshes: the discrete-time case*, SIAM J. Sci. Comput. **17** (1996), no. 1, 133–155, DOI 10.1137/0917011. Special issue on iterative methods in numerical linear algebra (Breckenridge, CO, 1994). MR1375271

[24] A. Lumsdaine and D. Wu, *Spectra and pseudospectra of waveform relaxation operators*, SIAM J. Sci. Comput. **18** (1997), no. 1, 286–304, DOI 10.1137/S106482759528778X. Dedicated to C. William Gear on the occasion of his 60th birthday. MR1433389

[25] S. P. MacLachlan and C. W. Oosterlee, *Algebraic multigrid solvers for complex-valued matrices*, SIAM J. Sci. Comput. **30** (2008), no. 3, 1548–1571, DOI 10.1137/070687232. MR2398878

[26] A. Napov and Y. Notay, *Algebraic analysis of aggregation-based multigrid*, Numer. Linear Algebra Appl. **18** (2011), no. 3, 539–564, DOI 10.1002/nla.741. MR2760067

[27] Y. Notay, *A robust algebraic multilevel preconditioner for non-symmetric M-matrices*, Numer. Linear Algebra Appl. **7** (2000), no. 5, 243–267, DOI 10.1002/1099-1506(200007/08)7:5⟨243::AID-NLA195⟩3.0.CO;2-Y. MR1766914

[28] Y. Notay, *Algebraic analysis of two-grid methods: The nonsymmetric case*, Numer. Linear Algebra Appl. **17** (2010), no. 1, 73–96, DOI 10.1002/nla.649. MR2589584

[29] Y. Notay, *Algebraic theory of two-grid methods*, Numer. Math. Theory Methods Appl. **8** (2015), no. 2, 168–198, DOI 10.4208/nmtma.2015.w04si. MR3395388

[30] C. W. Oosterlee, F. J. Gaspar, T. Washio, and R. Wienands, *Multigrid line smoothers for higher order upwind discretizations of convection-dominated problems*, J. Comput. Phys. **139** (1998), no. 2, 274–307, DOI 10.1006/jcph.1997.5854. MR1614090

[31] P. Oswald, *Multilevel finite element approximation*, Teubner Skripten zur Numerik. [Teubner Scripts on Numerical Mathematics], B. G. Teubner, Stuttgart, 1994. Theory and applications. MR1312165

[32] C. Rodrigo, F. J. Gaspar, and L. T. Zikatanov, *On the validity of the local Fourier analysis*, J. Comput. Math. **37** (2019), no. 3, 340–348, DOI 10.4208/jcm.1803-m2017-0294. MR3866073

[33] S. F. McCormick (ed.), *Multigrid Methods*, Frontiers in Applied Mathematics, vol. 3, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1987. MR972752

[34] Y. Saad, *Further analysis of minimum residual iterations*, Numer. Linear Algebra Appl. **7** (2000), no. 2, 67–93, DOI 10.1002/(SICI)1099-1506(200003)7:2⟨67::AID-NLA186⟩3.3.CO;2-#. MR1755800

[35] M. Sala and R. S. Tuminaro, *A new Petrov-Galerkin smoothed aggregation preconditioner for nonsymmetric linear systems*, SIAM J. Sci. Comput. **31** (2008), no. 1, 143–166, DOI 10.1137/060659545. MR2460774

[36] R. Stevenson, *On the validity of local mode analysis of multi-grid methods*, dissertation, Utrecht University, Utrecht, the Netherlands, 1990.

[37] K. Stüben and U. Trottenberg, *Multigrid Methods: Fundamental Algorithms, Model Problem Analysis and Applications*, Multigrid Methods (Cologne, 1981), Lecture Notes in Math., vol. 960, Springer, Berlin-New York, 1982, pp. 1–176. MR685773

[38] L. N. Trefethen and M. Embree, *Spectra and Pseudospectra*, Princeton University Press, Princeton, NJ, 2005. The Behavior of Nonnormal Matrices and Operators. MR2155029

[39] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, Inc., San Diego, CA, 2001. With contributions by A. Brandt, P. Oswald and K. Stüben. MR1807961

[40] J. Van lent, *Multigrid Methods for Time-Dependent Partial Differential Equations*, dissertation, K.U.Leuven, Leuven, Belgium, 2006.

[41] S. Vandewalle and G. Horton, *Fourier mode analysis of the multigrid waveform relaxation and time-parallel multigrid methods* (English, with English and German summaries), Computing **54** (1995), no. 4, 317–330, DOI 10.1007/BF02238230. MR1334614

[42] R. Wienands and W. Joppich, *Practical Fourier Analysis for Multigrid Methods*, Numerical Insights, vol. 4, Chapman & Hall/CRC, Boca Raton, FL, 2005. With 1 CD-ROM (Windows and UNIX). MR2108045

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