

A NOTE ON MULTIGRID METHODS FOR (MULTILEVEL)
STRUCTURED-PLUS-BANDED UNIFORMLY BOUNDED HERMITIAN
POSITIVE DEFINITE LINEAR SYSTEMS

STEFANO SERRA-CAPIZZANO, CRISTINA TABLINO POSSIO

QUADERNO N. 5/2008 (arxiv:math/0804.3016)



STAMPATO NEL MESE DI APRILE 2008
PRESSO IL DIPARTIMENTO DI MATEMATICA E APPLICAZIONI,
UNIVERSITÀ DEGLI STUDI DI MILANO-BICOCCA, VIA R. COZZI 53, 20125 MILANO, ITALIA.

DISPONIBILE IN FORMATO ELETTRONICO SUL SITO www.matapp.unimib.it.
SEGRETERIA DI REDAZIONE: Ada Osmetti - Giuseppina Cogliandro
tel.: +39 02 6448 5755-5758 fax: +39 02 6448 5705

**Esemplare fuori commercio per il deposito legale agli effetti della Legge 15 aprile 2004
n.106.**

A Note on Multigrid Methods for (Multilevel) Structured-plus-banded Uniformly Bounded Hermitian Positive Definite Linear Systems

Stefano Serra-Capizzano* Cristina Tablino-Possio†

Abstract

In the past few years a lot of attention has been paid in the multigrid solution of multilevel structured (Toeplitz, circulants, Hartley, sine (τ class) and cosine algebras) linear systems, in which the coefficient matrix is banded in a multilevel sense and Hermitian positive definite. In the present paper we provide some theoretical results on the optimality of an existing multigrid procedure, when applied to a properly related algebraic problem. In particular, we propose a modification of previously devised multigrid procedures in order to handle Hermitian positive definite structured-plus-banded uniformly bounded linear systems, arising when an indefinite, and not necessarily structured, banded part is added to the original coefficient matrix. In this context we prove the Two-Grid method optimality.

In such a way, several linear systems arising from the approximation of integro-differential equations with various boundary conditions can be efficiently solved in linear time (with respect to the size of the algebraic problem). Some numerical experiments are presented and discussed, both with respect to Two-Grid and multigrid procedures.

1 Introduction

In the past twenty years, an extensive literature has treated the numerical solution of structured linear systems of large dimensions [11], by means of preconditioned iterative solvers. However, as well known in the multilevel setting, the most popular matrix algebra preconditioners cannot work in general (see [28, 31, 24] and references therein), while the multilevel structures often are the most relevant in practical applications. Therefore, quite recently, more attention has been focused (see [1, 2, 10, 12, 13, 16, 19, 27, 30]) on the multigrid solution of multilevel structured (Toeplitz, circulants, Hartley, sine (τ class) and cosine algebras) linear systems, in which the coefficient matrix is banded in a multilevel sense and Hermitian positive definite. The reason is due to the fact that these techniques are very efficient, the total cost for reaching the solution within a preassigned accuracy being linear as the dimensions of the involved

*Dipartimento di Fisica e Matematica, Università dell'Insubria - sede di Como, via Valleggio, 11, 22100 Como, Italy. E-mail: s.serracapizzano@uninsubria.it

†Dipartimento di Matematica e Applicazioni, Università di Milano Bicocca, Via Cozzi, 53, 20125 Milano, Italy. E-mail: cristina.tablinopossio@unimib.it

linear systems.

In the present note we propose a slight modification of these numerical multigrid procedures in order to handle structured-plus-banded uniformly bounded Hermitian positive definite linear systems, where the banded part which is added to the structured coefficient matrix is indefinite and not necessarily structured. A theoretical analysis of the related Two-Grid Method (TGM) is given in terms of the algebraic multigrid theory considered by Ruge and Stüben [25]. More precisely, we prove that the proposed TGM is optimally convergent with a convergence rate independent of the dimension for a given sequence of linear systems $\{B_n y_n = c_n\}_n$ with uniformly bounded Hermitian positive definite matrix sequence $\{B_n\}_n$, under the assumption that such TGM is optimal for $\{A_n x_n = b_n\}_n$ with a given Hermitian positive definite matrix sequence $\{A_n\}_n$ related to $\{B_n\}_n$ by means of a simple order relation. More precisely, we require $A_n \leq \vartheta B_n$, with $\|B_n\|_2 \leq M$, for some $\vartheta, M > 0$ independent of n and for every n large enough.

As a case study, we may consider the case where $B_n = A_n + \Theta_n$ where A_n is structured, positive definite, ill-conditioned, and for which an optimal multigrid algorithm is already available, and where Θ_n is an indefinite band correction not necessarily structured; moreover, we require that $\{A_n\}_n$ and $\{B_n\}_n$ are uniformly bounded and that $A_n + \Theta_n$ is still positive definite and larger than A_n/ϑ for some $\vartheta > 0$ independent of n .

For instance such a situation is encountered when dealing with standard Finite Difference (FD) discretizations of the problem

$$\mathcal{L}u = -\nabla^2 u(x) + \mu(x)u(x) = h(x), \quad x \in \Omega,$$

where $\mu(x)$ and $h(x)$ are given bounded functions, $\Omega = (0, 1)^d$, $d \geq 1$, and with Dirichlet, periodic or reflective boundary conditions (for a discussion on various boundary conditions see [22, 29]). For specific contexts where structured-plus-diagonal problems arise refer to [14, 18] and [5], when considering also a convection term in the above equation. However, the latter is just an example chosen for the relevance in applications, but the effective range of applicability of our proposal is indeed much wider.

The numerical experimentation suggests that an optimal convergence rate should hold for the MGM as well. Here, for MGM algorithm, we mean the simplest (and less expensive) version of the large family of multigrid methods, i.e., the V-cycle procedure: for a brief description of the TGM and of the V-cycle algorithms we refer to Section 2, while an extensive treatment can be found in [17] and especially in [34]. Indeed, we remark that in all the considered examples the MGM is optimal in the sense that (see [4]):

- a. the observed number of iterations is constant with respect to the size of the algebraic problem;
- b. the cost per iteration (in terms of arithmetic operations) is just linear as the size of the algebraic problem.

Nevertheless, it is worth stressing that the theoretical extension of the optimality result to the MGM is still an open question.

The paper is organized as follows. In Section 2 we report the standard TGM and MGM algorithms and we write explicitly the related iteration matrices. In Section 3 we first recall some classical results related to the algebraic

TGM convergence analysis and then we prove the optimal rate of convergence of the proposed TGM, under some general and weak assumptions; the MGM case is briefly discussed at the end of the section. In Section 4 we analyze in detail the case of the discrete Laplacian-plus-diagonal systems and in Section 5 we report several numerical experiments. Lastly, Section 6 deals with further considerations concerning future work and perspectives.

2 Two-grid and Multigrid Method

Let n_0 be a positive d -index, $d \geq 1$, and let $N(\cdot)$ be an increasing function with respect to n_0 . In devising a TGM and a MGM for the linear system

$$A_{n_0} x_{n_0} = b_{n_0}, \quad (2.1)$$

where $A_{n_0} \in \mathbb{C}^{N(n_0) \times N(n_0)}$ and $x_{n_0}, b_{n_0} \in \mathbb{C}^{N(n_0)}$, the ingredients below must be considered.

Let $n_1 < n_0$ (componentwise) and let $p_{n_0}^{n_1} \in \mathbb{C}^{N(n_0) \times N(n_1)}$ be a given full-rank matrix. In order to simplify the notation, in the following we will refer to any multi-index n_s by means of its subscript s , so that, e.g. $A_s := A_{n_s}$, $b_s := b_{n_s}$, $p_s^{s+1} := p_{n_s}^{n_{s+1}}$, etc. With these notations, a class of stationary iterative methods of the form

$$x_s^{(j+1)} = V_s x_s^{(j)} + \tilde{b}_s$$

is also considered in such a way that $\mathcal{S}mooth(x_s^{(j)}, b_s, V_s, \nu_s)$ denotes the application of this rule ν_s times, with ν_s positive integer number, at the dimension corresponding to the index s .

Thus, the solution of the linear system (2.1) is obtained by applying repeatedly the TGM iteration, where the j^{th} iteration

$$x_0^{(j+1)} = \mathcal{TGM}(x_0^{(j)}, b_0, A_0, V_{0,\text{pre}}, \nu_{0,\text{pre}}, V_{0,\text{post}}, \nu_{0,\text{post}})$$

is defined by the following algorithm [17]:

$$y_0 := \underline{\mathcal{TGM}(x_0, b_0, A_0, V_{0,\text{pre}}, \nu_{0,\text{pre}}, V_{0,\text{post}}, \nu_{0,\text{post}})}$$

1. $\tilde{x}_0 := \mathcal{S}mooth(x_0, b_0, V_{0,\text{pre}}, \nu_{0,\text{pre}})$
2. $r_0 := b_0 - A_0 \tilde{x}_0$
3. $r_1 := (p_0^1)^H r_0$
4. Solve $A_1 y_1 = r_1$, with $A_1 := (p_0^1)^H A_0 p_0^1$
5. $\tilde{y}_0 := \tilde{x}_0 + p_0^1 y_1$
6. $y_0 := \mathcal{S}mooth(\tilde{y}_0, b_0, V_{0,\text{post}}, \nu_{0,\text{post}})$

Steps 1. and 6. concern the application of $\nu_{0,\text{pre}}$ steps of the *pre-smoothing* (or *intermediate*) iteration and of $\nu_{0,\text{post}}$ steps of the *post-smoothing* iteration, respectively. Moreover, steps 2. \rightarrow 5. define the so called *coarse grid correction*, that depends on the projection operator $(p_0^1)^H$. In such a way, the TGM iteration represents a classical stationary iterative method whose iteration matrix is given by

$$\mathcal{TGM}_0 = V_{0,\text{post}}^{\nu_{0,\text{post}}} \mathcal{CGC}_0 V_{0,\text{pre}}^{\nu_{0,\text{pre}}}, \quad (2.2)$$

where

$$\mathcal{CGC}_0 = I_0 - p_0^1 [(p_0^1)^H A_0 p_0^1]^{-1} (p_0^1)^H A_0$$

denotes the coarse grid correction iteration matrix.

The names *intermediate* and *smoothing iteration* used above refer to the multi-iterative terminology [26]: we say that a method is multi-iterative if it is composed by at least two distinct iterations. The idea is that these basic components should have complementary spectral behaviors so that the whole procedure is quickly convergent. In our case the target of the smoothing iteration is to reduce the error in the subspace where A_0 is well-conditioned, but such an iteration will be slowly convergent in the complementary space. The coarse grid correction iteration matrix is a (non-Hermitian) projector (see e.g. [27]) and therefore shows spectral radius equal to 1. As a consequence, the corresponding iterative procedure does not converge at all, but it is very quickly convergent in the subspace where A_0 is ill-conditioned, if p_0^1 is chosen in such a way that its columns span a subspace “close enough” to the ill-conditioned one. Finally, the *intermediate* iteration is strongly convergent in that subspace where the combined effect of the other two iterations resulted to be less effective. Notice that in our setting of Hermitian positive definite and uniformly bounded sequences, the subspace where A_0 is ill-conditioned corresponds to the subspace in which A_0 has small eigenvalues.

Starting from the TGM, we introduce the MGM. Indeed, the main difference with respect to the TGM is as follows: instead of solving directly the linear system with coefficient matrix A_1 , we can apply recursively the projection strategy so obtaining a *multigrid method*.

Let us use the Galerkin formulation and let $n_0 > n_1 > \dots > n_l > 0$, with l being the maximal number of recursive calls and with $N(n_s)$ being the corresponding matrix sizes.

The corresponding multigrid method generates the approximate solution

$$x_0^{(j+1)} = \mathcal{MGM}(0, x_0^{(j)}, b_0, A_0, V_{0,\text{pre}}, \nu_{0,\text{pre}}, V_{0,\text{post}}, \nu_{0,\text{post}})$$

according to the following algorithm:

$$y_s := \mathcal{MGM}(s, x_s, b_s, A_s, V_{s,\text{pre}}, \nu_{s,\text{pre}}, V_{s,\text{post}}, \nu_{s,\text{post}})$$

If $s = l$ then *Solve*($A_s y_s = b_s$)
else

1. $\tilde{x}_s := \text{Smooth}(x_s, b_s, V_{s,\text{pre}}, \nu_{s,\text{pre}})$
2. $r_s := b_s - A_s \tilde{x}_s$
3. $r_{s+1} := (p_s^{s+1})^H r_s$
4. $y_{s+1} := \mathcal{MGM}(s+1, 0_{s+1}, b_{s+1}, A_{s+1}, V_{s+1,\text{pre}}, \nu_{s+1,\text{pre}}, V_{s+1,\text{post}}, \nu_{s+1,\text{post}})$
5. $\tilde{y}_s := \tilde{x}_s + p_s^{s+1} y_{s+1}$
6. $y_s := \text{Smooth}(\tilde{y}_s, b_s, V_{s,\text{post}}, \nu_{s,\text{post}})$

where the matrix $A_{s+1} := (p_s^{s+1})^H A_s p_s^{s+1}$ is more profitably computed in the so called *pre-computing phase*.

Since the multigrid is again a linear fixed-point method, we can express $x_0^{(j+1)}$ as $\mathcal{MGM}_0 x_0^{(j)} + (I_0 - \mathcal{MGM}_0) A_0^{-1} b_0$, where the iteration matrix \mathcal{MGM}_0 is recursively defined according to the following rule (see [34]):

$$\begin{cases} \mathcal{MGM}_l = O, \\ \mathcal{MGM}_s = V_{s,\text{post}}^{\nu_{s,\text{post}}} [I_s - p_s^{s+1} (I_{s+1} - \mathcal{MGM}_{s+1}) A_{s+1}^{-1} (p_s^{s+1})^H A_s] V_{s,\text{pre}}^{\nu_{s,\text{pre}}}, \end{cases} \quad (2.3)$$

$$s = 0, \dots, l-1,$$

and with MGM_s and MGM_{s+1} denoting the iteration matrices of the multigrid procedures at two subsequent levels, $s = 0, \dots, l-1$. At the last recursion level l , the linear system is solved by a direct method and hence it can be interpreted as an iterative method converging in a single step: this motivates the chosen initial condition $MGM_l = O$.

By comparing the TGM and MGM, we observe that the coarse grid correction operator CGC_s is replaced by an approximation, since the matrix A_{s+1}^{-1} is approximated by $(I_{s+1} - MGM_{s+1}) A_{s+1}^{-1}$ as implicitly described in (2.3) for $s = 0, \dots, l-1$. In this way step 4., at the highest level $s = 0$, represents an approximation of the exact solution of step 4. displayed in the TGM algorithm (for the matrix analog compare (2.3) and (2.2)).

Finally, for $l = 1$ the MGM reduces to the TGM if $Solve(A_1 y_1 = b_1)$ is $y_1 = A_1^{-1} b_1$.

3 Discussion and extension of known convergence results

Hereafter, by $\|\cdot\|_2$ we denote the Euclidean norm on \mathbb{C}^m and the associated induced matrix norm over $\mathbb{C}^{m \times m}$. If X is Hermitian positive definite, then its square root obtained via the Schur decomposition is well defined and positive definite. As a consequence we can set $\|\cdot\|_X = \|X^{1/2} \cdot\|_2$ the Euclidean norm weighted by X on \mathbb{C}^m , and the associated induced matrix norm.

In the algebraic multigrid theory some relevant convergence results are due to Ruge and Stüben [25]. In fact, they provide the main theoretical tools to which we refer in order to prove our subsequent convergence results.

More precisely, by referring to the work of Ruge and Stüben [25], we will consider Theorem 5.2 therein in its original form and in the case where both pre-smoothing and post-smoothing iterations are performed. In the following all the constants α , α_{pre} , α_{post} , and β are required to be independent of the actual dimension in order to ensure a TGM convergence rate independent of the size of the algebraic problem.

Theorem 3.1 *Let A_0 be a Hermitian positive definite matrix of size $N(n_0)$, let $p_0^1 \in \mathbb{C}^{N(n_0) \times N(n_1)}$, $n_0 > n_1$, be a given full-rank matrix and let $V_{0,\text{post}}$ be the post-smoothing iteration matrix.*

Suppose that there exists $\alpha_{\text{post}} > 0$, independent of n_0 , such that for all $x \in \mathbb{C}^{N(n_0)}$

$$\|V_{0,\text{post}} x\|_{A_0}^2 \leq \|x\|_{A_0}^2 - \alpha_{\text{post}} \|x\|_{A_0 D_0^{-1} A_0}^2, \quad (3.1)$$

where D_0 is the diagonal matrix formed by the diagonal entries of A_0 .

Assume, also, that there exists $\beta > 0$, independent of n_0 , such that for all $x \in \mathbb{C}^{N(n_0)}$

$$\min_{y \in \mathbb{C}^{N(n_1)}} \|x - p_0^1 y\|_{D_0}^2 \leq \beta \|x\|_{A_0}^2. \quad (3.2)$$

Then, $\beta \geq \alpha_{\text{post}}$ and

$$\|TGM_0\|_{A_0} \leq \sqrt{1 - \alpha_{\text{post}}/\beta} < 1. \quad (3.3)$$

Theorem 3.2 Let A_0 be a Hermitian positive definite matrix of size $N(n_0)$, let $p_0^1 \in \mathbb{C}^{N(n_0) \times N(n_1)}$, $n_0 > n_1$, be a given full-rank matrix and let $V_{0,\text{pre}}$, $V_{0,\text{post}}$ be the pre-smoothing and post-smoothing iteration matrices, respectively. Suppose that there exist $\alpha_{\text{pre}}, \alpha_{\text{post}} > 0$, independent of n_0 , such that for all $x \in \mathbb{C}^{N(n_0)}$

$$\|V_{0,\text{pre}}x\|_{A_0}^2 \leq \|x\|_{A_0}^2 - \alpha_{\text{pre}} \|V_{0,\text{pre}}x\|_{A_0 D_0^{-1} A_0}^2, \quad (3.4)$$

$$\|V_{0,\text{post}}x\|_{A_0}^2 \leq \|x\|_{A_0}^2 - \alpha_{\text{post}} \|x\|_{A_0 D_0^{-1} A_0}^2, \quad (3.5)$$

where D_0 is the diagonal matrix formed by the diagonal entries of A_0 .

Assume, also, that there exists $\beta > 0$, independent of n_0 , such that for all $x \in \mathbb{C}^{N(n_0)}$

$$\|CGC_0x\|_{A_0}^2 \leq \beta \|x\|_{A_0 D_0^{-1} A_0}^2. \quad (3.6)$$

Then, $\beta \geq \alpha_{\text{post}}$ and

$$\|TGM_0\|_{A_0} \leq \sqrt{\frac{1 - \alpha_{\text{post}}/\beta}{1 + \alpha_{\text{pre}}/\beta}} < 1. \quad (3.7)$$

Remark 3.3 Theorems 3.1 and 3.2 still hold if the diagonal matrix D_0 is replaced by any Hermitian positive definite matrix X_0 (see e.g. [2]). More precisely, $X_0 = I$ could be a proper choice for its simplicity, since any contribution due to the use of a different matrix will be subject to a formal simplification in the quotients $\alpha_{\text{pre}}/\beta$ and $\alpha_{\text{post}}/\beta$.

Remark 3.4 For reader convenience, the essential steps of the proof of Theorems 3.1 and 3.2 are reported in Appendix A, where relations (3.1) and (3.4) are called post-smoothing and pre-smoothing property, respectively, and the relation (3.6) is called approximation property. In this respect, we notice that the approximation property deduced by using (3.2) holds only for vectors belonging to the range of CGC_0 , see (A.1); conversely the approximation property described in (3.6) is unconditional, i.e., it is satisfied for all $x \in \mathbb{C}^{N(n_0)}$.

In this paper we are interested in the multigrid solution of special linear systems of the form

$$B_n x = b, \quad B_n \in \mathbb{C}^{N(n) \times N(n)}, \quad x, b \in \mathbb{C}^{N(n)} \quad (3.8)$$

with $\{B_n\}_n$ Hermitian positive definite uniformly bounded matrix sequence, n being a positive d -index, $d \geq 1$ and $N(\cdot)$ an increasing function with respect to it. More precisely, we assume that there exists $\{A_n\}_n$ Hermitian positive definite matrix sequence such that some order relation is linking $\{A_n\}_n$ and $\{B_n\}_n$, for n large enough and we suppose that an optimal algebraic multigrid method is available for the solution of the systems

$$A_n x = b, \quad A_n \in \mathbb{C}^{N(n) \times N(n)}, \quad x, b \in \mathbb{C}^{N(n)}. \quad (3.9)$$

We ask whether the algebraic TGM and MGM considered for the systems (3.9) are effective also for the systems (3.8), i.e., when considering the very same projectors. Since it is well-known that a very crucial role in MGM is played by the choice of projector operator, the quoted choice will give rise to a relevant simplification. The results pertain to the convergence analysis of the TGM and MGM: we provide a positive answer for the TGM case and we only discuss the MGM case, which is substantially more involved.

3.1 TGM convergence and optimality: theoretical results

In this section we give a theoretical analysis of the TGM in terms of the algebraic multigrid theory due to Ruge and Stüben [25] according to Theorem 3.1. Hereafter, the notation $X \leq Y$, with X and Y Hermitian matrices, means that $Y - X$ is nonnegative definite. In addition, $\{X_n\}_n$, with X_n Hermitian positive definite matrices, is a uniformly bounded matrix sequence if there exists $M > 0$ independent of n such that $\|X_n\|_2 \leq M$, for n large enough.

Proposition 3.5 *Let $\{A_n\}_n$ be a matrix sequence with A_n Hermitian positive definite matrices and let $p_0^1 \in \mathbb{C}^{N(n_0) \times N(n_1)}$ be a given full-rank matrix for any $n_0 > 0$ such that there exists $\beta_A > 0$ independent of n_0 so that for all $x \in \mathbb{C}^{N(n_0)}$*

$$\min_{y \in \mathbb{C}^{N(n_1)}} \|x - p_0^1 y\|_2^2 \leq \beta_A \|x\|_{A_0}^2. \quad (3.10)$$

Let $\{B_n\}_n$ be another matrix sequence, with B_n Hermitian positive definite matrices, such that $A_n \leq \vartheta B_n$, for n large enough, with $\vartheta > 0$ absolute constant. Then, for all $x \in \mathbb{C}^{N(n_0)}$ and n_0 large enough, it also holds $\beta_B = \beta_A \vartheta$ and

$$\min_{y \in \mathbb{C}^{N(n_1)}} \|x - p_0^1 y\|_2^2 \leq \beta_B \|x\|_{B_0}^2. \quad (3.11)$$

Proof. From (3.10) and from the assumptions on the order relation, we deduce that for all $x \in \mathbb{C}^{N(n)}$

$$\min_{y \in \mathbb{C}^{N(n_1)}} \|x - p_0^1 y\|_2^2 \leq \beta_A \|x\|_{A_0}^2 \leq \vartheta \beta_A \|x\|_{B_0}^2,$$

i.e., taking into account Remark 3.3, the hypothesis (3.2) of Theorem 3.1 is fulfilled for $\{B_n\}_n$ too, with constant $\beta_B = \beta_A \vartheta$, by considering the very same projector p_0^1 considered for $\{A_n\}_n$. •

Thus, the convergence result in Theorem 3.1 holds true also for the matrix sequence $\{B_n\}_n$, if we are able to guarantee also the validity of condition (3.1). It is worth stressing that in the case of Richardson smoothers such topic is not related to any partial ordering relation connecting the Hermitian matrix sequences $\{A_n\}_n$ and $\{B_n\}_n$. In other words, given a partial ordering between $\{A_n\}_n$ and $\{B_n\}_n$, inequalities (3.1), (3.4), and (3.5) with $\{B_n\}_n$ instead of $\{A_n\}_n$ do not follow from (3.1), (3.4), and (3.5) with $\{A_n\}_n$, but they have to be proved independently. See Proposition 3 in [1] for the analogous claim in the case of $\nu_{\text{pre}}, \nu_{\text{post}} > 0$.

Proposition 3.6 *Let $\{B_n\}_n$ be an uniformly bounded matrix sequence, with B_n Hermitian positive definite matrices. For any $n_0 > 0$, let $V_{n,\text{pre}} = I_n - \omega_{\text{pre}} B_n$, $V_{n,\text{post}} = I_n - \omega_{\text{post}} B_n$ be the pre-smoothing and post-smoothing iteration matrices, respectively considered in the TGM algorithm. Then, there exist $\alpha_{B,\text{pre}}, \alpha_{B,\text{post}} > 0$ independent of n_0 such that for all $x \in \mathbb{C}^{N(n_0)}$*

$$\|V_{0,\text{pre}} x\|_{B_0}^2 \leq \|x\|_{B_0}^2 - \alpha_{B,\text{pre}} \|V_{0,\text{pre}} x\|_{B_0^2}^2, \quad (3.12)$$

$$\|V_{0,\text{post}} x\|_{B_0}^2 \leq \|x\|_{B_0}^2 - \alpha_{B,\text{post}} \|x\|_{B_0^2}^2. \quad (3.13)$$

Proof. Relation (3.13) is equivalent to the existence of an absolute positive constant $\alpha_{B,\text{post}}$ such that

$$(I_0 - \omega_{\text{post}} B_0)^2 B_0 \leq B_0 - \alpha_{B,\text{post}} B_0^2,$$

i.e.,

$$\omega_{\text{post}}^2 B_0 - 2\omega_{\text{post}} I_0 \leq -\alpha_{B,\text{post}} I_0.$$

The latter is equivalent to require that the inequality $\alpha_{B,\text{post}} \leq \omega_{\text{post}}(2 - \omega_{\text{post}}\lambda)$ is satisfied for any eigenvalue λ of the Hermitian matrix B_0 with $\alpha_{B,\text{post}} > 0$ independent of n_0 . Now, let $[m, M]$ be any interval containing the topological closure of the union over all n of all the eigenvalues of B_n . Then it is enough to consider

$$\alpha_{B,\text{post}} \leq \omega_{\text{post}} \min_{\lambda \in [m, M]} (2 - \omega_{\text{post}}\lambda) = \omega_{\text{post}}(2 - \omega_{\text{post}}M),$$

where the condition $\omega_{\text{post}} < 2/M$ ensures $\alpha_{B,\text{post}} > 0$.

By exploiting an analogous technique, in the case of relation (3.12), it is sufficient to consider

$$\begin{aligned} \alpha_{B,\text{pre}} &\leq \omega_{\text{pre}} \min_{\lambda \in [0, M]} \frac{\omega_{\text{pre}}(2 - \omega_{\text{pre}}\lambda)}{(1 - \omega_{\text{pre}}\lambda)^2} \\ &= \begin{cases} 2\omega_{\text{pre}} & \text{if } 0 < \omega_{\text{pre}} \leq 3/(2M), \\ \frac{\omega_{\text{pre}}(2 - \omega_{\text{pre}}M)}{(1 - \omega_{\text{pre}}M)^2} & \text{if } 3/(2M) \leq \omega_{\text{pre}} < 2/M, \end{cases} \end{aligned}$$

where we consider the only interesting case $m = 0$, since $m > 0$ is related to the case of well-conditioned systems. \bullet

In this way, according to the Ruge and Stüben algebraic theory, we have proved the TGM optimality, that is its convergence rate independent of the size $N(n)$ of the involved algebraic problem.

Theorem 3.7 *Let $\{B_n\}_n$ be an uniformly bounded matrix sequence, with B_n Hermitian positive definite matrices. Under the same assumptions of Propositions 3.5 and 3.6 the TGM with only one step of post-smoothing converges to the solution of $B_n x = b$ and its convergence rate is independent of $N(n)$.*

Proof. By referring to Propositions 3.5 and 3.6 the claim follows according to Theorem 3.1. \bullet

Few remarks are useful in order to understand what happens when also a pre-smoothing phase is applied.

A) The first observation is that the convergence analysis can be reduced somehow to the case of only post-smoothing. Indeed, looking at relation (2.2) and recalling that the spectra of AB and BA are the same for any pair (A, B) of square matrices (see [7]), it is evident that

$$TGM_0 = V_{0,\text{post}}^{\nu_0,\text{post}} CGC_0 V_{0,\text{pre}}^{\nu_0,\text{pre}}$$

has the same spectrum, and hence the same spectral radius $\rho(\cdot)$, as

$$V_{0,\text{pre}}^{\nu_0,\text{pre}} V_{0,\text{post}}^{\nu_0,\text{post}} CGC_0,$$

where the latter represents a TGM iteration with only post-smoothing. Therefore

$$\begin{aligned}\rho(TGM_0) &= \rho(V_{0,\text{pre}}^{\nu_{0,\text{pre}}} V_{0,\text{post}}^{\nu_{0,\text{post}}} CGC_0) \\ &\leq \|V_{0,\text{pre}}^{\nu_{0,\text{pre}}} V_{0,\text{post}}^{\nu_{0,\text{post}}} CGC_0\|_{A_0} \\ &\leq \sqrt{1 - \tilde{\alpha}_{\text{post}}/\beta}\end{aligned}$$

where $\tilde{\alpha}_{\text{post}}$ is the post-smoothing constant of the cumulative stationary method described by the iteration matrix $V_{0,\text{pre}}^{\nu_{0,\text{pre}}} V_{0,\text{post}}^{\nu_{0,\text{post}}}$.

- B)** Setting $\nu_{0,\text{pre}} = \nu_{0,\text{post}} = 1$ and with reference to Item A), we easily deduce that $\tilde{\alpha}_{\text{post}} \geq \alpha_{\text{post}}$ where the latter is the post-smoothing constant related to the sole post-smoothing method with iteration matrix $V_{0,\text{post}}$. Furthermore, if the two iteration matrices $V_{0,\text{pre}}$ and $V_{0,\text{post}}$ are chosen carefully, i.e., by taking into account the spectral complementarity principle, then we can expect that $\tilde{\alpha}_{\text{post}}$ is sensibly larger than α_{post} , so that the TGM with both pre-smoothing and post-smoothing is sensibly faster than that with only post-smoothing.
- C)** Items A) and B) show that the TGM iteration with both pre-smoothing and post-smoothing is never worse than the TGM iteration with only post-smoothing. Therefore Theorem 3.7 implies that the TGM with both post-smoothing and pre-smoothing is optimal for systems with matrices B_n under the same assumptions as in Theorem 3.7.
- D)** At this point the natural question arises: is it possible to handle directly assumption (3.6), instead of assumption (3.2)? As observed in Remark 3.4 these two assumptions are tied up and indeed they represent the approximation property on the range of CGC_0 and unconditional, respectively. However, from a technical viewpoint, they are very different and in fact we are unable to state a formal analog of Proposition 3.5 by using (3.6). More precisely, for concluding that

$$\|CGC_0 x\|_{A_0}^2 \leq \beta_A \|x\|_{A_0^2}^2$$

implies

$$\|CGC_0 x\|_{B_0}^2 \leq \beta_B \|x\|_{B_0^2}^2$$

with $X_0 = I$ as in Remark 3.3 and with $\vartheta, \beta_A, \beta_B$ absolute constants, and $A_n \leq \vartheta B_n$, we would need $X \leq Y, X, Y \geq 0$ implies $X^2 \leq \gamma Y^2$ with some γ positive and independent of n . The latter with $\gamma = 1$ is the operator monotonicity of the map $z \rightarrow z^2$ which is known to be false in general [7]. We should acknowledge that there exist important subclasses of matrices for which $X \leq Y, X, Y \geq 0$ implies $X^2 \leq \gamma Y^2$. However, this matrix theoretic analysis of intrinsic interest goes a bit far beyond the scope of the present paper and will be the subject of future investigations.

- E)** Remark 3.4 furnishes an interesting degree of freedom that could be exploited. For instance if we choose $X_0 = A_0$, by assuming suitable order relations between $\{A_n\}_n$ and $\{B_n\}_n$, then proving that

$$\|CGC_0 x\|_{A_0}^2 \leq \beta_A \|x\|_{A_0}^2$$

implies

$$\|CGC_0x\|_{B_0}^2 \leq \beta_B \|x\|_{B_0}^2$$

with ϑ , β_A , β_B absolute constants, becomes easier, but, conversely, the study of the pre-smoothing and post-smoothing properties becomes more involved.

3.2 MGM convergence and optimality: a discussion

In this section we briefly discuss the same question as before, but in connection with the MGM. First of all, we expect that a more severe assumption between $\{A_n\}_n$ and $\{B_n\}_n$ has to be fulfilled in order to infer the MGM optimality for $\{B_n\}_n$ starting from the MGM optimality for $\{A_n\}_n$. The reason is that the TGM is just a special instance of the MGM when setting $l = 1$.

In the TGM setting we have assumed a one side ordering relation: here the most natural step is to consider a two side ordering relation, that is to assume that there exist positive constants ϑ_1, ϑ_2 independent of n such that $\vartheta_1 B_n \leq A_n \leq \vartheta_2 B_n$, for every n large enough. The above relationships simply represent the spectral equivalence condition for sequences of Hermitian positive definite matrices. In the context of the preconditioned conjugate gradient method (see [3]), it is well known that if $\{P_n\}_n$ is a given sequence of optimal (i.e., spectrally equivalent) preconditioners for $\{A_n\}_n$, then $\{P_n\}_n$ is also a sequence of optimal preconditioners for $\{B_n\}_n$ (see e.g. [24]). The latter fact just follows from the observation that the spectral equivalence is an *equivalence* relation and hence is transitive.

In summary, we have enough heuristic motivations in order to conjecture that the spectral equivalence is the correct and needed assumption and, in reality, the numerical experiments reported in Section 5 give a support to the latter statement.

From a theoretical point of view, as done for the TGM, we start from the Ruge-Stüben tools [25] in the slightly modified version contained in Theorem 2.3 in [2], that is taking into account Remark 3.3 and, for the sake of simplicity, we assume no pre-smoothing i.e., $\nu_{\text{pre}} = 0$. The matrix inequalities coming from the assumption (2.9) in [2] are very intricate since they involve simultaneously projector operators and smoothers: whence, it is customary to split it into the *smoothing property* (relation (2.11a) in [2]) and the *approximation property* (relation (2.11b) in [2]). As usual the smoothing property does not pose any problem. However, we encounter a serious technical difficulty in the second inequality, i.e., when dealing with the approximation property. More precisely, we arrive to compare two Hermitian projectors, depending on the same p_s^{s+1} with the first involving A_{n_s} and the second involving B_{n_s} . Unfortunately, they can be compared only in very special and too restricted cases: the needed assumption would not involve ordering, but only the fact that the columns of $A_{n_s}^{1/2} p_s^{s+1}$ and those of $B_{n_s}^{1/2} p_s^{s+1}$ span the same space.

However, as already mentioned, the numerical tests tell us that the latter difficulty is only a technicality and that the right assumption should involve spectral equivalence. Therefore, in future investigations, other directions and proof techniques have to be explored.

4 A case study: discrete Laplacian-plus-diagonal systems

In the present section, we analyze a specific application of the results in Section 3. More precisely, we consider a multigrid strategy for solving Laplacian-plus-diagonal linear systems arising from standard Finite Differences (FD) discretizations of the problem

$$\mathcal{L}u = -\nabla^2 u(x) + \mu(x)u(x) = h(x), \quad x \in \Omega, \quad (4.1)$$

where $\mu(x)$ and $h(x)$ are given bounded functions, $\Omega = (0, 1)^d$, $d \geq 1$, and with Dirichlet, periodic or reflective boundary conditions. Thus, we are facing with a matrix sequence

$$\{B_n\}_n = \{A_n + D_n\}_n, \quad (4.2)$$

where the structure of the matrix sequence $\{A_n\}_n$ is related both to the FD discretization and to the type of the boundary conditions and where $\{D_n\}_n$ is a sequence of uniformly bounded diagonal matrices, due to the hypothesis that $\mu(x)$ is bounded.

Since a fast TGM and MGM working for the Toeplitz (or τ - the τ class is the algebra associated to the most known sine transform [8]) part is well-known [15, 16, 9, 27], we are in position to apply the tools in the preceding section in order to show that the same technique works, and with a cost linear as the dimension, in the context (4.2) too. In the same way, the extension of suitable MGM procedures proposed in the case of the circulant [30], DCT-III cosine [12, 32], or τ [15, 16] algebra, can be considered according to the corresponding boundary conditions. Clearly, this case study is just an example relevant in applications, while the results in Section 3 are of much wider generality.

Once more, we want to remark that, unfortunately, there is a gap in the theory with regard to the MGM, even if the numerical tests reported in Section 5 suggest that the MGM applied to matrices in $\{B_n = A_n + D_n\}_n$ is optimal under the assumptions that the same MGM is optimal for $\{A_n\}_n$, A_n symmetric positive definite matrix, and $\{D_n\}_n$ uniformly bounded matrix sequence, with $A_n \leq \vartheta B_n$ uniformly with respect to n and with some fixed $\vartheta > 0$ independent of n . Clearly if the matrices D_n are also nonnegative definite then the constant ϑ can be set to 1. This result can be plainly extended to the case in which D_n is a (multilevel) banded correction.

4.1 One-Dimensional case

According to the FD approximation of (4.1) with Dirichlet boundary conditions, we obtain the matrix sequence

$$\{B_n\}_n = \{A_n + D_n\}_n$$

where $\{A_n\}_n = \{\text{tridiag}_n[-1, 2, -1]\}_n$ and $\{D_n\}_n$ is a sequence of diagonal matrices whose diagonal entries $d_i^{(n)}$, $i = 1, \dots, n$, are uniformly bounded in modulus by a constant M independent of n . Since

$$\lambda_{\min}(A_n) = 4 \sin^2 \left(\frac{\pi}{2(n+1)} \right) = \frac{\pi^2}{n^2} + O(n^{-3}),$$

we impose the condition

$$n^2 \min_{1 \leq i \leq n} d_i^{(n)} + \pi^2 \geq c$$

for some $c > 0$ independent of n (we consider only the case $\min_{1 \leq i \leq n} d_i^{(n)} < 0$, since the other is trivial). Thus, also $\{B_n\}_n$ is an uniformly bounded positive definite matrix sequence and

$$B_n \geq A_n + \frac{c - \pi^2}{n^2} I \geq \frac{c}{\pi^2} A_n$$

so satisfying the crucial assumption $A_n \leq \vartheta B_n$ in Proposition 3.5 with $\vartheta = \pi^2/c$. Let us consider $B_0 \in \mathbb{R}^{n_0 \times n_0}$, with 1-index $n_0 > 0$. Following [15, 27], we denote by $T_0^1 \in \mathbb{R}^{n_0 \times n_1}$, $n_0 = 2n_1 + 1$, the operator such that

$$(T_0^1)_{i,j} = \begin{cases} 1 & \text{for } i = 2j, \quad j = 1, \dots, n_1, \\ 0 & \text{otherwise,} \end{cases} \quad (4.3)$$

and we define a projector $(p_0^1)^H$, $p_0^1 \in \mathbb{R}^{n_0 \times n_1}$ as

$$p_0^1 = \frac{1}{\sqrt{2}} P_0 T_0^1, \quad P_0 = \text{tridiag}_0 [1, 2, 1]. \quad (4.4)$$

Thus, the basic step in order to prove the TGM optimality result is reported in the proposition below. It is worth stressing that the claim refers to a tridiagonal matrix correction, since, under the quoted assumption, each diagonal correction is projected at the first coarse level into a tridiagonal correction, while the tridiagonal structure is kept unaltered in all the subsequent levels.

Proposition 4.1 *Let $B_0 = \text{tridiag}_0 [-1, 2, -1] + T_0 \in \mathbb{R}^{n_0 \times n_0}$, with $n_0 > 0$ and T_0 being a symmetric uniformly bounded tridiagonal matrix such that $A_0 \leq \vartheta B_0$, with $A_0 = \text{tridiag}_0 [-1, 2, -1]$ and some $\vartheta > 0$. Let $p_0^1 = (1/\sqrt{2})\text{tridiag}_0 [1, 2, 1]T_0^1$, with $n_0 = 2n_1 + 1$. Then,*

$$(p_0^1)^H B_0 p_0^1 = \text{tridiag}_1 [-1, 2, -1] + T_1$$

where $T_1 \in \mathbb{R}^{n_1 \times n_1}$ is a symmetric uniformly bounded tridiagonal matrix with $A_1 \leq \vartheta B_1$, $B_1 = (p_0^1)^H B_0 p_0^1$, $A_1 = \text{tridiag}_1 [-1, 2, -1]$.

Proof. For the Toeplitz part refer to [27]. For the tridiagonal part we just need a simple check. In fact, the product $P_0 T_0^1 P_0$ is a 7-diagonal matrix (P_0 and T_0 are tridiagonal) and the action of T_0^1 , on the left and on the right, selects the even rows and columns so that the resulting matrix is still tridiagonal. Since $A_0 \leq \vartheta B_0$, $A_0 = \text{tridiag}_0 [-1, 2, -1]$, $B_1 = (p_0^1)^H B_0 p_0^1$, and $A_1 = \text{tridiag}_1 [-1, 2, -1] = (p_0^1)^H A_0 p_0^1$ it is evident that $A_1 \leq \vartheta B_1$. Finally, the uniform boundedness is guaranteed by the uniform boundedness of all the involved matrices. •

Corollary 4.2 *Let $B_0 = \text{tridiag}_0 [-1, 2, -1] + T_0 \in \mathbb{R}^{n_0 \times n_0}$, with $n_0 > 0$ and T_0 symmetric tridiagonal matrix such that $A_0 \leq \vartheta B_0$, with $A_0 = \text{tridiag}_0 [-1, 2, -1]$ and some $\vartheta > 0$. Let $p_0^1 = (1/\sqrt{2})\text{tridiag}_0 [1, 2, 1]T_0^1$, with $n_0 = 2n_1 + 1$. Then, there exists $\beta_B > 0$ independent of n_0 so that inequality (3.2) holds true.*

Proof. Let $A_0 = \text{tridiag}_0 [-1, 2, -1] \in \mathbb{R}^{n_0 \times n_0}$. Then relation (3.2) is fulfilled with the operator p_0^1 defined in (4.4) and with a certain β_A independent of n_0 , as proved in [27]. Moreover, from the assumption we have $A_0 \leq \vartheta B_0$ so that Proposition 3.5 implies that (3.2) holds true for B_0 with a constant $\beta_B = \vartheta \beta_A$. •

Corollary 4.3 *Let $\{B_n\}_n$ be the sequence such that $B_n = \text{tridiag}_n [-1, 2, -1] + T_n \in \mathbb{R}^{n \times n}$ with T_n symmetric uniformly bounded tridiagonal matrices. Then, there exist $\alpha_{B,\text{pre}}, \alpha_{B,\text{post}} > 0$ independent of n , so that inequalities (3.12) and (3.13) hold true.*

Proof. It is evident that $\{B_n\}_n$ is a sequence of symmetric positive definite matrices uniformly bounded by $4 + M$, with $\|T_n\|_2 \leq M$ independent of n , so that the thesis follows by the direct application of Proposition 3.6. •

4.2 Two-Dimensional case

Hereafter, we want to consider the TGM and MGM extension to the case $d > 1$. Due to the discretization process, it is natural, and easier, to work with d -indices $n = (n^{(1)}, \dots, n^{(d)})$, with $n^{(r)}$ integer positive number, $r = 1, \dots, d$. In this case the matrix dimension is $N(n_0) = \prod_{r=1}^d n_0^{(r)}$ and when considering the projected matrices of size $N(n_1)$ we have that n_1 is again a d -index and we assume not only $N(n_1) < N(n_0)$, but also $n_1 < n_0$ componentwise. We discuss in detail the two-level case, since the d -level one is a simple generalization. Thus, in the two-level case, we are dealing with the matrix sequence

$$\{B_n\}_n = \{A_n + D_n\}_n$$

where $\{A_n\}_n = \{\text{tridiag}_{n^{(1)}} [-1, 2, -1] \otimes I_{n^{(2)}} + I_{n^{(1)}} \otimes \text{tridiag}_{n^{(2)}} [-1, 2, -1]\}_n$ and $\{D_n\}_n$ is a sequence of diagonal matrices whose diagonal entries $d_i^{(n)}$, $i = 1, \dots, N(n)$, are uniformly bounded in modulus by a constant M independent of n . Since

$$\begin{aligned} \lambda_{\min}(A_n) &= 4 \sin^2 \left(\frac{\pi}{2(n^{(1)} + 1)} \right) + 4 \sin^2 \left(\frac{\pi}{2(n^{(2)} + 1)} \right) \\ &= \frac{\pi^2}{[n^{(1)}]^2} + \frac{\pi^2}{[n^{(2)}]^2} + O(\psi^{-3}), \end{aligned}$$

we impose the condition

$$\psi^2 \min_{1 \leq i \leq N(n)} d_i^{(n)} + \pi^2 \geq c$$

for some $c > 0$ independent of n , with $\psi = \min_j \{n^{(j)}\}$. Thus, also $\{B_n\}_n$ is an uniformly bounded positive definite matrix sequence and

$$B_n \geq A_n + \frac{c - \pi^2}{\psi^2} I \geq \frac{c}{\pi^2} A_n$$

so satisfying the crucial assumption $A_n \leq \vartheta B_n$ in Proposition 3.5 with $\vartheta = \pi^2/c$. The projector definition can be handled in a natural manner by using tensorial

arguments: $(p_0^1)^H$ is constructed in such a way that

$$\begin{aligned} p_0^1 &= P_0 U_0^1 \\ P_0 &= \text{tridiag}_{n_0^{(1)}} [1, 2, 1] \otimes \text{tridiag}_{n_0^{(2)}} [1, 2, 1], \\ U_0^1 &= T_0^1(n_0^{(1)}) \otimes T_0^1(n_0^{(2)}) \end{aligned}$$

with $n_0^{(r)} = 2n_1^{(r)} + 1$ and where $T_0^1(n_0^{(r)}) \in \mathbb{R}^{n_0^{(r)} \times n_1^{(r)}}$ is the unilevel matrix given in (4.3). Notice that this is the most trivial extension of the unilevel projector to the two-level setting and such a choice is also the less expensive from a computational point of view: in fact, $p_0^1 = \tau_0((2 + 2 \cos(t_1))(2 + 2 \cos(t_2)))U_0^1$ equals $[\tau_{n_0^{(1)}}(p(2 + 2 \cos(t_1)))T_0^1(n_0^{(1)})] \otimes [\tau_{n_0^{(2)}}(p(2 + 2 \cos(t_2)))T_0^1(n_0^{(2)})]$. The proposition below refers to a two-level tridiagonal correction for the same reasons as the unilevel case.

Proposition 4.4 *Let*

$$B_0 = \text{tridiag}_{n_0^{(1)}} [-1, 2, -1] \otimes I_{n_0^{(2)}} + I_{n_0^{(1)}} \otimes \text{tridiag}_{n_0^{(2)}} [-1, 2, -1] + T_0 \in \mathbb{R}^{N(n_0) \times N(n_0)},$$

with $n_0 > 0$ and T_0 being a symmetric uniformly bounded tridiagonal block matrix with tridiagonal blocks such that $A_0 \leq \vartheta B_0$, with $A_0 = \text{tridiag}_{n_0^{(1)}} [-1, 2, -1] \otimes I_{n_0^{(2)}} + I_{n_0^{(1)}} \otimes \text{tridiag}_{n_0^{(2)}} [-1, 2, -1]$ and some $\vartheta > 0$. Let

$$p_0^1 = (\text{tridiag}_{n_0^{(1)}} [1, 2, 1] \otimes \text{tridiag}_{n_0^{(2)}} [1, 2, 1])(T_0^1(n_0^{(1)}) \otimes T_0^1(n_0^{(2)})),$$

with $n_0^{(r)} = 2n_1^{(r)} + 1$, $r = 1, 2$. Then, $B_1 := (p_0^1)^H B_0 p_0^1$ coincides with $A_1 + T_1$ where $T_1 \in \mathbb{R}^{N(n_1) \times N(n_1)}$ is a symmetric uniformly bounded tridiagonal block matrix with tridiagonal blocks and where A_1 is a two-level τ tridiagonal block matrix with tridiagonal blocks asymptotic to

$$\text{tridiag}_{n_1^{(1)}} [-1, 2, -1] \otimes I_{n_1^{(2)}} + I_{n_1^{(1)}} \otimes \text{tridiag}_{n_1^{(1)}} [-1, 2, -1],$$

so that $A_1 \leq \vartheta B_1$.

Proof. For the τ part refer to [27]. For the two-level banded part it is a simple check. In fact, the product $P_0 T_0 P_0$ is a 7-diagonal block matrix with 7-diagonal blocks (P_0 is a tridiagonal block matrix with tridiagonal blocks) and the action of U_0^1 , on the left and on the right, selects even rows and columns in even blocks with respect to the rows and columns, so that the resulting matrix has a tridiagonal block pattern with tridiagonal blocks. The order relation follows as a direct consequence of the assumption $A_0 \leq \vartheta B_0$ and the uniform boundedness is implied by the uniform boundedness of all the involved matrices. •

Corollary 4.5 *Let $B_0 = A_0 + T_0 \in \mathbb{R}^{N(n_0) \times N(n_0)}$ with $n_0 > 0$,*

$$A_0 = \text{tridiag}_{n_0^{(1)}} [-1, 2, -1] \otimes I_{n_0^{(2)}} + I_{n_0^{(1)}} \otimes \text{tridiag}_{n_0^{(2)}} [-1, 2, -1],$$

and T_0 symmetric tridiagonal block matrix with tridiagonal blocks such that $A_0 \leq \vartheta B_0$ for some $\vartheta > 0$. Let

$$p_0^1 = (\text{tridiag}_{n_0^{(1)}} [1, 2, 1] \otimes \text{tridiag}_{n_0^{(2)}} [1, 2, 1])(T_0^1(n_0^{(1)}) \otimes T_0^1(n_0^{(2)})),$$

with $n_0^{(r)} = 2n_1^{(r)} + 1$, $r = 1, 2$. Then, there exists $\beta_B > 0$ independent of n_0 so that inequality (3.2) holds true.

Proof. The proof can be done following the same steps as in Corollary 4.2. •

Corollary 4.6 Let $\{B_n\}_n$ be the sequence such that $B_n = A_n + T_n$ with

$$A_n = \text{tridiag}_{n^{(1)}} [-1, 2, -1] \otimes I_{n^{(2)}} + I_{n^{(1)}} \otimes \text{tridiag}_{n^{(2)}} [-1, 2, -1],$$

and with T_n symmetric uniformly bounded tridiagonal block matrix with tridiagonal blocks. Then, there exist $\alpha_{B,\text{pre}}, \alpha_{T,\text{pre}} > 0$ independent of n , so that inequalities (3.12) and (3.13) hold true.

Proof. The proof can be worked out as in Corollary 4.3 since the sequence $\{B_n\}_n$ is uniformly bounded by $8 + M$, with $\|T_n\|_2 \leq M$ independent of n by assumption. •

5 Numerical Examples

We test our TGM and MGM (standard V-cycle according to Section 2) for several examples of matrix corrections $\{D_n\}_n$, $D_n \in \mathbb{C}^{N(n) \times N(n)}$, $N(n) = \prod_{r=1}^d n^{(r)}$, $d = 1, 2$.

We will consider nonnegative definite band corrections and indefinite band corrections. By referring to Section 4, the case of nonnegative definite corrections implies trivially that $A_n \leq B_n$ so that the desired constant is $\vartheta = 1$. However, as observed in real-world applications (see [14]), the most challenging situation is the one of indefinite corrections.

Concerning nonnegative definite corrections, the reference set is defined according to the following notation, in the unilevel and in the two-level setting, respectively:

	$d_s^{(n)}$	d0	d1	d2	d3	d4
1D	$s = 1, \dots, N(n)$	0	$\frac{s}{s+1}$	$ \sin(s) $	$ \sin(s) \frac{s^2-1}{s^2+1}$	$\frac{s}{N(n)}$
2D	$s = (i-1)n_1 + j$ $i = 1, \dots, n_1$ $j = 1, \dots, n_2$	0	$\frac{i}{i+1}$ $+\frac{j}{j+1}$	$ \sin(i) $ $+ \sin(j) $	$ \sin(i) \frac{i^2-1}{i^2+1}$ $+ \sin(j) \frac{j^2-1}{j^2+1}$	$\frac{s}{N(n)}$

The case of indefinite corrections is considered in connection with Laplacian systems with Dirichlet boundary conditions: in that setting the diagonal entries $d_s^{(n)}$ of D_n are generated randomly. Finally higher order differential operators and linear systems arising from integral equations in image restoration are considered at the end of the section.

The aim is to give numerical evidences of the theoretical optimality results of TGM convergence and also to their extension in the case of the MGM application.

The projectors are properly chosen according to the nature of structured part, while we will use, in general, the Richardson smoothing/intermediate iteration step twice in each iteration, before and after the coarse grid correction, with different values of the parameter ω .

According to the definition, when considering the TGM, the exact solution of the system is obtained by using a direct solver in the immediately subsequent coarse grid dimension, while, when considering the MGM, the exact solution of the system is computed by the same direct solver, when the coarse grid dimension equals 16^d (where $d = 1$ for the unilevel case and $d = 2$ for the two-level case).

In all tables we report the numbers of iterations required for the TGM or MGM convergence, assumed to be reached when the Euclidean norm of the relative residual becomes less than 10^{-7} . We point out that the CPU times are consistent with the iteration counts.

Finally, we stress that the matrices A_n at every level (except for the coarsest) are never formed since we need only to store the nonzero Fourier coefficients of the generating function at every level for matrix-vector multiplications. Thus, besides the $O(N(n))$ operations complexity of the proposed MGM both with respect to the structured part and clearly with respect to the non-structured one, the memory requirements of the structured part are also very low since there are only $O(1)$ nonzero Fourier coefficients of the generating function at every level. On the other hand, the projections of the initial diagonal correction are stored at each level according to standard sparse matrix techniques during the pre-computing phase.

5.1 Discrete Laplacian-plus-diagonal systems

The numerical tests below refer to convergence results in the case of matrix sequences arising from the Laplacian discretization, in the unilevel and in the two-level settings, respectively.

5.1.1 Dirichlet boundary conditions

Firstly, we consider the case of Dirichlet boundary conditions so that the obtained matrix sequence is the Toeplitz/ τ matrix sequence $\{\tau_n(f)\}_n$ generated by the function $f(t) = 2 - 2\cos(t)$, $t \in (0, 2\pi]$. The projector is defined as in (4.3) and (4.4), while the parameters ω for the smoothing/intermediate iterations are chosen as

$$\omega_{\text{pre}} = \frac{1}{2(\|f\|_\infty + \|D_n\|_\infty)} \quad \text{and} \quad \omega_{\text{post}} = \frac{1}{\|f\|_\infty + \|D_n\|_\infty},$$

with $\nu_{\text{pre}} = \nu_{\text{post}} = 1$.

The results in Table 1 confirm the optimality of the proposed TGM in the sense that the number of iterations is uniformly bounded by a constant not depending on the size $N(n)$ indicated in the first column. Moreover, it seems that this claim can be extended to the MGM convergence. Notice, also, that the number of iterations is frequently the best possible since it equals the number of TGM iterations.

The case of the diagonal correction $d4$ deserves special attention: as shown in the first column, just one pre-smoothing/intermediate and post-smoothing

$$B_n = \text{tridiag}_n[-1, 2, -1] + \text{Diagonal}$$

TGM						MGM						
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4	
											$\rho=0$	$\rho=1$
31	2	7	7	7	7	31	2	7	8	8	7	7
63	2	7	8	8	7	63	7	7	7	7	7	7
127	2	7	8	8	7	127	8	7	8	8	8	7
255	2	7	8	8	7	255	8	7	8	8	9	7
511	2	6	8	8	7	511	8	7	8	8	16	7

Table 1: Number of iterations required by TGM and MGM - unilevel cases (refer to (5.1) for the definition of the constant ρ).

$$B_n = \text{tridiag}_{n(1)}[-1, 2, -1] \otimes I_{n(2)} + I_{n(1)} \otimes \text{tridiag}_{n(2)}[-1, 2, -1] + \text{Diagonal}$$

TGM						MGM						
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4	
											$\rho=0$	$\rho=1$
31^2	16	10	13	13	16	31^2	16	10	13	13	16	16
63^2	16	10	13	13	16	63^2	16	10	13	13	17	16
127^2	16	10	13	13	16	127^2	16	10	12	12	18	16
255^2	16	10	13	13	16	255^2	16	10	12	12	27	16
511^2	16	10	13	13	16	511^2	16	9	12	12	36	16

Table 2: Number of iterations required by TGM and MGM - two-level cases (refer to (5.1) for the definition of the constant ρ).

iteration at each coarse grid level are not sufficient to ensure the optimality. Moreover, it is enough to consider a trick that keeps unaltered the $O(N(n))$ computational cost as proved in [30] (only the multiplicative constant hidden in the big O can increase): at each projection on a coarser grid the number of smoothing iterations performed at that level is increased by a fixed constant ρ , i.e., according to the MGM notation of Section 2, we set

$$\nu_{s+1} = \nu_s + \rho, \quad s = 0, \dots, l-1, \quad \nu_0 = 1. \quad (5.1)$$

The optimality result in the second column relative to MGM in the $d4$ case is obtained just by considering $\rho = 1$. This phenomenon is probably due to some inefficiency in considering the approximation $\|D_n\|_\infty$ in the tuning of the parameter ω_{pre} and ω_{post} . In fact, it is enough to substitute, for instance, the post-smoother with the Gauss-Seidel method in order to preserve the optimality also for $\rho = 0$.

Other examples of Toeplitz/ τ linear systems plus diagonal correction can be found in [23], corresponding to Sinc-Galerkin discretization of differential problems according to [20].

By using tensor arguments, our results plainly extend to the two-level setting and the comments concerning Table 2 are substantially equivalent as in the unilevel case.

Before dealing with other type of boundary conditions, we want to give a comparison of the performances of the proposed method with respect to those achieved by considering, for instance, the conjugate gradient (cg) method. Table 3 reports, for increasing dimension, the Euclidean matrix condition number $k_2(A_n + D_n)$, together with the number of iterations required by the cg. As

$$B_n = \text{tridiag}_n[-1, 2, -1] + \text{Diagonal}$$

$N(n)$	d0		d1		d2		d3		d4	
	k_2	nit	k_2	nit	k_2	nit	k_2	nit	k_2	nit
31	4.14e+2	31	5.62e+0	18	7.98e+0	22	8.16e+0	22	2.03e+1	27
63	1.65e+3	63	5.67e+0	18	8.02e+0	21	8.19e+0	22	3.30e+1	34
127	6.63e+3	127	5.68e+0	18	8.02e+0	21	8.19e+0	21	5.31e+1	43
255	2.65e+4	255	5.69e+0	17	8.02e+0	21	8.20e+0	21	8.51e+1	54
511	1.06e+5	511	5.69e+0	17	8.05e+0	21	8.20e+0	21	1.35e+2	66

$$B_n = \text{tridiag}_{n^{(1)}}[-1, 2, -1] \otimes I_{n^{(2)}} + I_{n^{(1)}} \otimes \text{tridiag}_{n^{(2)}}[-1, 2, -1] + \text{Diagonal}$$

$N(n)$	d0		d1		d2		d3		d4	
	k_2	nit	k_2	nit	k_2	nit	k_2	nit	k_2	nit
31^2	4.14e+2	82	5.62e+0	18	7.98e+0	22	8.16e+0	22	3.83e+1	46
63^2	1.65e+3	163	5.67e+0	17	8.02e+0	22	8.19e+0	22	6.29e+1	57
127^2	6.63e+3	319	5.68e+0	16	8.02e+0	21	8.19e+0	21	1.01e+2	71
255^2	2.65e+4	623	5.69e+0	16	8.02e+0	21	8.20e+0	21	1.60e+2	83
511^2	1.06e+5	1215	5.69e+0	15	8.05e+0	21	8.20e+0	21	2.54e+2	99

Table 3: Euclidean condition number $k_2(A_n + D_n)$ and number of iterations required by cg - unilevel and two-level cases.

well known in the case $d0$, the cg method requires all the $N(n)$ steps in order to reach the convergence. Moreover, the non-structured part in the cases $d1$, $d2$, $d3$ increases the minimum eigenvalue of the resulting matrix so that the whole condition number becomes moderate. As a consequence the standard cg method is also effective. Notice that this good behavior is no longer observed in the case $d4$, while our MGM technique is still optimal. The same trend is observed in the two-level setting.

5.1.2 Randomly generated indefinite corrections

As a further interesting case, we want to test our proposal in the case of randomly generated matrix corrections. More specifically, we consider diagonal, symmetric tridiagonal, symmetric pentadiagonal matrix corrections with random entries uniformly distributed on the unit interval (cases $d5$, $d7$, and $d9$, respectively) or normally distributed with mean zero and standard deviation one (cases $d6$, $d8$, and $d10$, respectively). Notice that in such a way we are also considering indefinite corrections. Thus, in order to obtain a positive definite matrix B_n and in order to satisfy the crucial relation $A_n \leq \vartheta B_n$ for some positive ϑ independent of n , the arising random matrices corrections are suitable scaled by $1/(\gamma n^2)$ in the unilevel setting, with γ being the number of non-zero diagonals, and by $1/(\gamma(n^{(1)})^2)$ in the two-level setting (assuming $n^{(1)} = n^{(2)}$). Table 4 reports the Euclidean condition number and the mean of the number of iterations required by the MGM in the unilevel and two-level setting by considering, for each case, ten examples of random matrix corrections.

All these results confirm the effectiveness of our proposal. Though the Euclidean condition numbers are fully comparable with those of the $d0$ case, the number of required iterations does not worsen. Conversely, the cg method requires for instance in the $d5$ case $N(n)$ iterations in the unilevel setting, and 83, 163, 318, 621, 1212 in the two-level one.

It is worth stressing that the pentadiagonal corrections are reduced at the first

$$B_n = \text{tridiag}_n[-1, 2, -1] + \text{random correction}$$

$N(n)$	d5		d6		d7		d8		d9		d10	
	k_2	nit	k_2	nit	k_2	nit	k_2	nit	k_2	nit	k_2	nit
31	3.89e+2	3	4.09e+2	3.5	3.83e+2	3	4.16e+2	3	3.84e+2	3	4.11e+2	3
63	1.57e+3	7	1.64e+3	7	1.53e+3	7	1.66e+3	7	1.51e+3	7	1.64e+3	7
12	6.29e+3	8	6.65e+3	8	6.18e+3	8	6.54e+3	8	6.11e+3	8	6.75e+3	8
255	2.52e+4	8	2.65e+4	8	2.46e+4	8	2.63e+4	8	2.44e+4	8	2.65e+4	8
511	1.01e+5	8	1.05e+5	8	9.86e+4	8	1.06e+5	8	9.77e+4	8	1.05e+5	8

$$B_n = \text{tridiag}_{n(1)}[-1, 2, -1] \otimes I_{n(2)} + I_{n(1)} \otimes \text{tridiag}_{n(2)}[-1, 2, -1] + \text{random correction}$$

$N(n)$	d5		d6		d7		d8		d9		d10	
	k_2	nit	k_2	nit	k_2	nit	k_2	nit	k_2	nit	k_2	nit
31^2	4.02e+2	16	4.15e+2	16	3.98e+2	16	4.15e+2	16	3.96e+2	16	4.14e+2	16
63^2	1.61e+3	16	1.66e+3	16	1.59e+3	16	1.66e+3	16	1.59e+3	16	1.65e+3	16
127^2	6.47e+3	16	6.64e+3	16	6.39e+3	16	6.63e+3	16	6.36e+3	16	6.63e+3	16
255^2	2.58e+4	16	2.65e+4	16	2.55e+4	16	2.65e+4	16	2.54e+4	16	2.65e+4	16
511^2	1.03e+5	16	1.06e+5	16	1.02e+5	16	1.06e+5	16	1.01e+5	16	1.06e+5	16

Table 4: Euclidean condition number $k_2(A_n + D_n)$ and mean number of iterations required by MGM - unilevel and two-level cases.

projection to tridiagonal matrices. More in general, bigger patterns are reduced after few steps to a fixed pattern driven by the projector pattern (see [15, 1, 2]).

5.1.3 Periodic and Reflective boundary conditions

In the case of periodic boundary conditions the obtained matrix sequence is the circulant matrix sequence $\{S_n(f)\}_n$ generated by the function $f(t) = 2 - 2 \cos(t), t \in (0, 2\pi]$. Following [30], we consider the operator $T_0^1 \in \mathbb{R}^{n_0 \times n_1}$, $n_0 = 2n_1$, such that

$$(T_0^1)_{i,j} = \begin{cases} 1 & \text{for } i = 2j - 1, \quad j = 1, \dots, n_1, \\ 0 & \text{otherwise,} \end{cases}$$

and we define a projector $(p_0^1)^H, p_0^1 \in \mathbb{R}^{n_0 \times n_1}$, as

$$p_0^1 = P_0 T_0^1, \quad P_0 = S_0(p), \quad p(t) = 2 + 2 \cos(t).$$

It must be outlined that in the $d0$ case the arising matrices are singular, so that we consider the classical Strang correction [33]

$$\tilde{S}_{n_0}(f) = S_{n_0}(f) + f \left(\frac{2\pi}{N(n_0)} \right) \frac{ee^t}{N(n_0)}$$

where e is the vector of all ones. The results in the top part of Table 5 confirm the optimality of the proposed TGM and its extension to MGM (the case $d4$ requires to set $\rho = 4$).

When dealing with reflective boundary conditions, the obtained matrix sequence is the DCT III matrix sequence $C_n(f)_n$ generated by the function $f(t) = 2 - 2 \cos(t), t \in (0, 2\pi]$. Following [12], we consider the operator $T_0^1 \in \mathbb{R}^{n_0 \times n_1}$, $n_0 = 2n_1$, such that

$$(T_0^1)_{i,j} = \begin{cases} 1 & \text{for } i \in \{2j - 1, 2j\}, \quad j = 1, \dots, n_1, \\ 0 & \text{otherwise,} \end{cases}$$

$$B_n = \text{unilevel circulant } S_n(f) + \text{Diagonal}, f(t) = 2 - 2 \cos(t)$$

TGM						MGM						
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4	
											$\rho=0$	$\rho=4$
32	2	6	7	7	7	32	2	6	7	7	7	7
64	2	6	7	7	7	64	7	6	7	7	10	7
128	2	6	7	7	7	128	7	6	7	7	16	7
256	2	6	7	7	7	256	8	6	7	7	22	7
512	2	6	6	7	6	512	8	6	6	7	29	8

$$B_n = \text{unilevel DCT III } C_n(f) + \text{Diagonal}, f(t) = 2 - 2 \cos(t)$$

TGM						MGM						
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4	
											$\rho=0$	$\rho=2$
32	7	6	7	7	6	32	7	6	7	7	6	6
64	7	5	6	6	5	64	7	5	6	6	7	5
128	7	5	7	7	5	128	7	5	7	7	11	5
256	7	4	7	7	4	256	7	5	6	6	17	5
512	7	4	6	6	4	512	7	4	7	7	27	6

Table 5: Number of iterations required by TGM and MGM - unilevel cases (refer to (5.1) for the definition of the constant ρ).

and we define a projector $(p_0^1)^H$, $p_0^1 \in \mathbb{R}^{n_0 \times n_1}$, as

$$p_0^1 = P_0 T_0^1, P_0 = C_0(p), p(t) = 2 + 2 \cos(t).$$

Again, the results in bottom part of Table 5 confirm the optimality of the proposed TGM and its extension to MGM. It is worth stressing that in the $d0$ case we are considering the matrix

$$\tilde{C}_{n_0}(f) = C_{n_0}(f) + f \left(\frac{\pi}{N(n_0)} \right) \frac{ee^t}{N(n_0)}.$$

Furthermore, the case $d4$ requires to set $\rho = 2$ in order to observe optimality. By using tensor arguments, our results plainly extend to the two-level setting and the comments concerning Table 6 are substantially equivalent as in the corresponding unilevel case.

5.2 Other examples

In this section we give numerical evidences of the optimality of TGM and MGM results in a more general setting.

5.2.1 Higher order τ discretizations plus diagonal systems

We consider τ matrix sequences arising from the discretization of higher order differential problems with proper homogeneous boundary conditions on $\partial\Omega$:

$$(-1)^q \sum_{i=1}^d \frac{\partial^{2q}}{\partial x_i^{2q}} u(x) + \mu(x)u(x) = h(x) \text{ on } \Omega = (0, 1)^d, \quad (5.2)$$

i.e., $\{B_n = A_n + D_n\}_n$, where $A_n = \tau_n(f)$ with $f(t) = \sum_{i=1}^d (2 - 2 \cos(t_i))^q$. More specifically, in the unilevel case we define $p(t) = [2 + 2 \cos(t)]^w$ where w is

$$B_n = \text{two-level circulant } S_n(f) + \text{Diagonal, } f(t_1, t_2) = 4 - 2 \cos(t_1) - 2 \cos(t_2)$$

TGM						MGM							
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4	$\rho=0$	$\rho=1$
32^2	15	8	11	11	14	32^2	15	8	11	11	14	14	14
64^2	15	7	11	11	15	64^2	15	7	11	11	15	15	15
128^2	15	7	11	11	15	128^2	15	7	11	11	16	14	14
256^2	15	7	11	11	15	256^2	15	7	11	11	24	14	14
512^2	15	7	11	11	15	512^2	15	7	11	11	34	14	14

$$B_n = \text{two-level DCT III } C_n(f) + \text{Diagonal, } f(t_1, t_2) = 4 - 2 \cos(t_1) - 2 \cos(t_2)$$

TGM						MGM							
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4	$\rho=0$	$\rho=1$
32^2	16	6	10	10	12	32^2	16	6	10	10	12	12	12
64^2	16	6	10	10	11	64^2	16	6	10	10	11	11	11
128^2	16	5	10	10	11	128^2	16	5	10	10	11	10	10
256^2	16	5	9	9	11	256^2	16	5	9	9	17	9	9
512^2	16	5	9	9	11	512^2	16	5	9	9	27	9	9

Table 6: Number of iterations required by TGM and MGM - two-level cases (refer to (5.1) for the definition of the constant ρ).

chosen according to conditions in [15, 9, 27]: in order to have a MGM optimality we must take w at least equal to 1 if $q = 1$ and at least equal to 2 if $q = 2, 3$. Clearly, the lower is the value of w , the greater will be the advantage from a computational viewpoint. Indeed, Table 7 confirms the need of these constraints with respect to the case $q = 2$, this being the only $d0$ case where we observe a growth in the number of iterations with respect to $N(n)$. Nevertheless, it should be noticed that in the same case the contribution of the non-structured part improves the numerical behavior since the minimal eigenvalue is increased. The remaining results in Table 7 confirm the optimality of the corresponding MGM (the $d4$ case requires to set ρ in a proper way as just observed in the Laplacian case).

Notice that the bandwidth of the non-structured diagonal correction is increased by subsequent projections until a maximal value corresponding to $4w - 1$ is reached (for a discussion on the evolution of the bandwidth when a generic (multilevel) band system is encountered see [15, 1, 2]).

With respect to the two-level problem, we consider again the most trivial extension (and less expensive from a computational point of view) of the unilevel projector to the two-level setting, given by $P_n = \tau_n(p)$ with $p(t_1, t_2) = [(2 + 2 \cos(t_1))(2 + 2 \cos(t_2))]^w$, $w = 1, 2, 3$. The comments concerning the two-level setting in Table 7 are of the same type as in the unilevel one.

5.2.2 Higher order circulant discretizations plus diagonal systems

We consider circulant matrix sequences arising from the approximation of higher order differential problems with proper homogeneous/periodic boundary conditions on $\partial\Omega$ as in (5.2), i.e., $\{B_n = A_n + D_n\}_n$, where $A_n = S_n(f)$ with $f(t) = \sum_{i=1}^d (2 - 2 \cos(t_i))^q$. The choice of the generating function for the projector is the same as in the previous section (see [30]). Indeed, Table 8 shows

$$B_n = \text{unilevel } \tau + \text{ Diagonal, } f(t) = (2 - 2 \cos(t))^q$$

q = 2													
w = 1							w = 2						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=4$						$\rho=0$	$\rho=2$
31	20	13	16	16	13	13	31	16	13	14	14	13	13
63	45	13	17	17	17	13	63	16	12	14	14	12	12
127	84	12	15	15	26	13	127	16	12	13	13	13	12
255	149	12	16	16	36	13	255	16	12	13	13	27	12
511	253	12	16	16	42	13	511	16	11	13	13	35	13

q = 3													
w = 2							w = 3						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=1$						$\rho=0$	$\rho=1$
31	35	31	33	33	32	32	31	34	31	33	33	32	32
63	35	31	33	33	31	31	63	34	31	33	33	31	31
127	35	31	32	32	31	32	127	34	30	32	32	31	31
255	35	30	32	32	37	31	255	34	30	32	32	37	30
511	35	30	32	32	39	31	511	34	29	31	32	42	30

$$B_n = \text{two-level } \tau + \text{ Diagonal, } f(t_1, t_2) = (2 - 2 \cos(t_1))^q + (2 - 2 \cos(t_2))^q$$

q = 2													
w = 1							w = 2						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=2$						$\rho=0$	$\rho=2$
31 ²	37	29	32	32	36	36	31 ²	35	28	31	31	34	34
63 ²	44	28	32	32	43	35	63 ²	36	28	31	31	34	34
127 ²	80	28	31	31	73	35	127 ²	36	27	30	30	56	34
255 ²	140	27	30	30	109	35	255 ²	36	27	30	30	89	33
511 ²	235	27	30	30	151	35	511 ²	36	27	30	30	129	33

q = 3													
w = 2							w = 3						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=2$						$\rho=0$	$\rho=2$
31 ²	72	64	67	67	69	69	31 ²	68	61	64	64	66	66
63 ²	73	65	68	68	71	71	63 ²	72	64	67	67	70	70
127 ²	73	64	68	68	137	71	127 ²	72	63	67	67	127	70
255 ²	73	64	67	67	193	70	255 ²	72	63	67	67	182	70
511 ²	73	64	67	67	276	70	511 ²	72	63	67	67	260	70

Table 7: Number of required MGM iterations - unilevel and two-level cases (refer to (5.1) for the definition of the constant ρ).

the importance of these constraints with respect to the case $d0$ with $q = 2$. It is worth mentioning that the optimality of the corresponding MGM is again confirmed (for the case $d4$ the parameter ρ has to be set in a proper way). The comments concerning the two-level setting in Table 8 are of the same type as in the unilevel one.

5.2.3 Reflective BCs discretizations plus diagonal systems

We consider an example of DCT-III matrix sequences arising from the discretization of integral problems with reflective boundary conditions (see [22]), i.e., $\{B_n = A_n + D_n\}_n$, where $A_n = C_n(f)$ with f having nonnegative Fourier coefficients as it is required for the point spread function in the modeling of

$$B_n = \text{unilevel circulant} + \text{Diagonal}, f(t) = (2 - 2 \cos(t))^q$$

q = 2													
w = 1							w = 2						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=4$						$\rho=0$	$\rho=4$
32	19	12	13	14	15	15	32	15	12	13	13	14	14
64	41	12	14	15	29	17	64	15	12	13	13	17	14
128	77	12	14	14	47	19	128	15	12	13	13	34	14
256	137	12	14	14	69	22	256	15	12	12	13	53	14
512	224	12	14	15	94	25	512	15	11	13	13	75	15

q = 3													
w = 2							w = 3						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=4$						$\rho=0$	$\rho=4$
32	32	29	30	31	31	31	32	32	29	30	31	31	31
64	32	29	30	30	34	31	64	31	29	30	30	31	30
128	32	29	30	30	60	31	128	31	28	30	30	54	30
256	32	29	30	30	92	31	256	31	28	29	30	84	30
512	32	28	30	30	133	31	512	31	28	30	30	122	30

$$B_n = \text{two-level circulant} + \text{Diagonal}, f(t_1, t_2) = (2 - 2 \cos(t_1))^q + (2 - 2 \cos(t_2))^q$$

q = 2													
w = 1							w = 2						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=2$						$\rho=0$	$\rho=1$
32 ²	34	23	26	27	32	32	32 ²	34	23	27	27	31	31
64 ²	42	22	25	26	38	33	64 ²	33	22	25	25	31	31
128 ²	76	22	25	25	56	33	128 ²	33	22	25	25	44	31
256 ²	130	21	24	24	75	34	256 ²	33	21	24	24	62	31
512 ²	213	21	25	225	96	34	512 ²	33	20	24	24	85	31

q = 3													
w = 2							w = 3						
N(n)	d0	d1	d2	d3	d4		N(n)	d0	d1	d2	d3	d4	
					$\rho=0$	$\rho=1$						$\rho=0$	$\rho=1$
32 ²	68	56	59	60	65	65	32 ²	68	56	59	60	65	65
64 ²	67	55	58	58	64	64	64 ²	66	54	58	58	64	63
128 ²	67	54	58	58	88	64	128 ²	66	53	57	58	82	63
256 ²	67	53	57	57	118	65	256 ²	66	52	57	57	111	63
512 ²	67	52	57	57	158	65	512 ²	66	51	56	57	149	63

Table 8: Number of required MGM iterations - unilevel and two-level cases (refer to (5.1) for the definition of the constant ρ).

image blurring, see [6]. A simple model is represented by $f(t) = f_d(t) := \sum_{i=1}^d (2 + 2 \cos(t_i))$ where, by the way, the product $f_1(t_1)f_1(t_2)$ is encountered when treating super-resolution or high resolution problems, see e.g. [21]. The choice of the generating function for the projector is the same as in [12].

The results in Table 9 confirm again the optimality of the corresponding MGM (the case $d4$ requires to set ρ in a proper way). The observations regarding the two-level setting are in the same spirit as those of the unilevel one.

$$B_n = \text{unilevel DCT III } C_n(f) + \text{ Diagonal, } f(t) = 2 + 2 \cos(t).$$

TGM						MGM						
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4	
											$\rho=0$	$\rho=3$
32	14	5	8	8	9	32	14	5	8	8	9	9
64	12	4	8	8	9	64	13	5	8	8	11	10
128	11	4	8	8	9	128	13	5	9	9	18	10
256	10	4	8	8	9	256	12	4	9	9	24	11
512	8	3	8	8	9	512	10	4	9	9	30	11

$$B_n = \text{two-level DCT III } C_n(f) + \text{ Diagonal, } f(t_1, t_2) = (2 + 2 \cos(t_1)) + (2 - 2 \cos(t_2)).$$

TGM						MGM					
$N(n)$	d0	d1	d2	d3	d4	$N(n)$	d0	d1	d2	d3	d4
32^2	5	6	7	8	6	32^2	5	6	7	8	6
64^2	4	5	7	8	6	64^2	5	5	9	9	6
128^2	4	5	7	8	6	128^2	4	5	9	9	6
256^2	4	5	7	8	6	256^2	4	5	9	9	6
512^2	4	5	7	8	6	512^2	4	5	9	9	6

Table 9: Number of iterations required by TGM and MGM - unilevel and two level cases (refer to (5.1) for the definition of the constant ρ).

6 Concluding Remarks

The algebraic tools given in Section 3 and Section 4 revealed that, if a suitable TGM for a Hermitian positive definite matrix sequence $\{A_n\}_n$ is available and another Hermitian positive definite uniformly bounded sequence $\{B_n\}_n$ is given such that $A_n \leq \vartheta B_n$, for n large enough, then the same strategy works almost unchanged for $\{B_n\}_n$ too. As an example, this means that if the method is optimal for the first sequence then it is optimal for the second as well. The same results should hold for the MGM procedures, but here only a wide set of numerical evidences has been provided for supporting the claim: the related theory will be a subject of future investigations taking into account the final remarks in Section 3.1 and the discussion in Section 3.2.

We point out that the latter goal is quite important. Indeed, it is not difficult to prove relations of the form $\vartheta_1 A_n \leq B_n \leq \vartheta_2 A_n$ with B_n being discretization of an elliptic variable coefficient problem, A_n being the same discretization in the constant coefficient case, and where ϑ_1, ϑ_2 are positive constants independent of n and mainly depending on the ellipticity parameters of the problem. Therefore, the above mentioned results would represent a link for inferring MGM optimality on a general (possibly high order) variable coefficient elliptic problem, starting from the MGM optimality for the structured part, i.e., the one related to the constant coefficient discretization.

Finally, we point out that the latter idea has been used essentially for the structured plus diagonal systems coming from approximated elliptic partial differential equations with different boundary conditions. However, the same approach is applicable to a wide variety of cases, as sketched for instance in Section 5.2.3.

A Appendix

For reader convenience, we report the essential steps of the proof of Theorems 3.1 and 3.2.

Let us start by proving Theorem 3.1. As demonstrated in Theorem 5.2 in [25], the existence of $\beta > 0$ such that

$$\min_{y \in \mathbb{C}^{N(n_1)}} \|x - p_0^1 y\|_{D_0}^2 \leq \beta \|x\|_{A_0}^2 \quad \forall x \in \mathbb{C}^{N(n_0)}$$

implies the validity of the so called *approximation property* only in the range of CGC_0 , i.e., the existence of $\beta > 0$ such that

$$\|CGC_0 x\|_{A_0}^2 \leq \beta \|CGC_0 x\|_{A_0 D_0^{-1} A_0}^2 \quad \forall x \in \mathbb{C}^{N(n_0)} \quad (\text{A.1})$$

where $CGC_0 = I_0 - p_0^1 A_1^{-1} (p_0^1)^H A_0$.

Thus, by virtue of the *post-smoothing property* (3.1) and of (A.1), for all $x \in \mathbb{C}^{N(n_0)}$ we find

$$\begin{aligned} \|V_{0,\text{post}} CGC_0 x\|_{A_0}^2 &\leq \|CGC_0 x\|_{A_0}^2 - \alpha_{\text{post}} \|CGC_0 x\|_{A_0 D_0^{-1} A_0}^2 \\ &\leq \|CGC_0 x\|_{A_0}^2 - \frac{\alpha_{\text{post}}}{\beta} \|CGC_0 x\|_{A_0}^2 \\ &= \left(1 - \frac{\alpha_{\text{post}}}{\beta}\right) \|CGC_0 x\|_{A_0}^2 \\ &\leq \left(1 - \frac{\alpha_{\text{post}}}{\beta}\right) \|x\|_{A_0}^2, \end{aligned} \quad (\text{A.2})$$

being $\|CGC_0\|_{A_0} = 1$.

Since $TGM_0 = V_{0,\text{post}} CGC_0$ in the case where no pre-smoothing is considered, the latter is the same as $\|TGM_0\|_{A_0} \leq \sqrt{1 - \frac{\alpha_{\text{post}}}{\beta}}$ and hence Theorem 3.1 is proved.

Now let us prove Theorem 3.2. Since the *approximation property* (3.6) implies clearly (A.1), by repeating the very same steps as before and exploiting the *post-smoothing property* (3.5), for all $x \in \mathbb{C}^{N(n_0)}$ we find

$$\begin{aligned} \|V_{0,\text{post}} CGC_0 V_{0,\text{pre}} x\|_{A_0}^2 &\leq \|CGC_0 V_{0,\text{pre}} x\|_{A_0}^2 - \alpha_{\text{post}} \|CGC_0 V_{0,\text{pre}} x\|_{A_0 D_0^{-1} A_0}^2 \\ &\leq \|CGC_0 V_{0,\text{pre}} x\|_{A_0}^2 - \frac{\alpha_{\text{post}}}{\beta} \|CGC_0 V_{0,\text{pre}} x\|_{A_0}^2 \\ &= \left(1 - \frac{\alpha_{\text{post}}}{\beta}\right) \|CGC_0 V_{0,\text{pre}} x\|_{A_0}^2. \end{aligned} \quad (\text{A.3})$$

In addition, by using (3.6) and the *pre-smoothing property* (3.4), respectively, for all $x \in \mathbb{C}^{N(n_0)}$ we obtain

$$\begin{aligned} \|CGC_0 V_{0,\text{pre}} x\|_{A_0}^2 &\leq \beta \|V_{0,\text{pre}} x\|_{A_0 D_0^{-1} A_0}^2 \\ \|V_{0,\text{pre}} x\|_{A_0 D_0^{-1} A_0}^2 &\leq \alpha_{\text{pre}}^{-1} (\|x\|_{A_0}^2 - \|V_{0,\text{pre}} x\|_{A_0}^2). \end{aligned}$$

Hence

$$\begin{aligned} \frac{\alpha_{\text{pre}}}{\beta} \|CGC_0 V_{0,\text{pre}} x\|_{A_0}^2 &\leq \|x\|_{A_0}^2 - \|V_{0,\text{pre}} x\|_{A_0}^2 \\ &\leq \|x\|_{A_0}^2 - \|CGC_0 V_{0,\text{pre}} x\|_{A_0}^2, \end{aligned}$$

since

$$\|CGC_0 V_{0,\text{pre}x}\|_{A_0}^2 \leq \|CGC_0\|_{A_0}^2 \|V_{0,\text{pre}x}\|_{A_0}^2 = \|V_{0,\text{pre}x}\|_{A_0}^2,$$

being $\|CGC_0\|_{A_0} = 1$. Therefore, for all $x \in \mathbb{C}^{N(n_0)}$, it holds

$$\|CGC_0 V_{0,\text{pre}x}\|_{A_0}^2 \leq \left(1 + \frac{\alpha_{\text{pre}}}{\beta}\right)^{-1} \|x\|_{A_0}^2. \quad (\text{A.4})$$

By using inequality (A.4) in (A.3), we have

$$\|V_{0,\text{post}}CGC_0 V_{0,\text{pre}x}\|_{A_0}^2 \leq \frac{1 - \alpha_{\text{post}}/\beta}{1 + \alpha_{\text{pre}}/\beta} \|x\|_{A_0}^2,$$

and the proof of Theorem 3.2 is concluded.

References

- [1] A. Aricò and M. Donatelli, *A V-cycle multigrid for multilevel matrix algebras: proof of optimality*, Numer. Math. 105 (2007), no. 4, pp. 511–547.
- [2] A. Aricò, M. Donatelli, and S. Serra-Capizzano, *V-cycle optimal convergence for certain (multilevel) structured linear system*, SIAM J. Matrix Anal. Appl., 26-1 (2004), pp. 186–214.
- [3] O. Axelsson, *Iterative Solution Methods*, Cambridge University Press, Cambridge, 1994.
- [4] O. Axelsson and M. Neytcheva, *The algebraic multilevel iteration methods – theory and applications*, Proc. 2nd Int. Coll. on Numerical Analysis, D. Bainov Ed., Plovdiv (Bulgaria), August 1993, pp. 13–23.
- [5] D. Bertaccini, G. H. Golub, S. Serra-Capizzano, and C. Tablino-Possio, *Preconditioned HSS methods for the solution of non-Hermitian positive definite linear systems and applications to the discrete convection-diffusion equation*, Numer. Math., 99 (2005), pp. 441–484.
- [6] M. Bertero and P. Boccacci, *Introduction to Inverse Problems in Imaging*, Inst. of Physics Publ. Bristol and Philadelphia, London, 1998.
- [7] R. Bhatia, *Matrix Analysis*, Springer Verlag, New York, 1997.
- [8] D. Bini and M. Capovani, *Spectral and computational properties of band symmetric Toeplitz matrices*, Linear Algebra Appl., 52/53 (1983), pp. 99–125.
- [9] R. H. Chan, Q. Chang, and H. Sun, *Multigrid method for ill-conditioned symmetric Toeplitz systems*, SIAM J. Sci. Comput., 19-2 (1998), pp. 516–529.
- [10] R. H. Chan, M. Donatelli, S. Serra-Capizzano, and C. Tablino-Possio, *Application of multigrid techniques to image restoration problems*, Proc. SPIE - Session: Advanced Signal Processing: Algorithms, Architectures, and Implementations XII, Vol. 4791 (2002), F. Luk Ed., pp. 210–221.

- [11] R. H. Chan and M. K. Ng, *Conjugate gradient methods for Toeplitz systems*, SIAM Rev., 38 (1996), pp. 427–482.
- [12] R. H. Chan, S. Serra-Capizzano, and C. Tablino-Possio, *Two-Grid methods for banded linear systems from DCT III algebra*, Numer. Linear Algebra Appl., 12-2/3 (2005), pp. 241–249.
- [13] Q. Chang, X. Jin, and H. Sun, *Convergence of the multigrid method for ill-conditioned block toeplitz systems*, BIT, 41-1(2001), pp. 179–190.
- [14] U. Elsner, V. Mehrmann, F. Milde, R.A. Römer, and M. Schreiber, *The Anderson Model of Localization: A Challenge for Modern Eigenvalue Methods*, Preprint SFB393/97-31, Sonderforschungsbereich 393, Numerische Simulation auf massiv parallelen Rechnern, Fakultät für Mathematik, TU Chemnitz; SIAM J. Sci. Comput., 20 (1999), pp. 2089–2102.
- [15] G. Fiorentino and S. Serra-Capizzano, *Multigrid methods for Toeplitz matrices*, Calcolo, 28 (1991), pp. 283–305.
- [16] G. Fiorentino and S. Serra-Capizzano, *Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions*, SIAM J. Sci. Comput., 17-4 (1996), pp. 1068–1081.
- [17] W. Hackbusch, *Multigrid Methods and Applications*, Springer Verlag, Berlin, 1985.
- [18] M.K. Ho and M.K. Ng, *Splitting iterations for circulant-plus-diagonal systems*, Numer. Linear Algebra Appl., 12 (2005), pp. 779–792.
- [19] T. Huckle and J. Staudacher, *Multigrid preconditioning and Toeplitz matrices*, Electr. Trans. Numer. Anal., 13 (2002), pp. 81–105.
- [20] J. Lund and K. Bowers, *Sinc Methods for Quadrature and Differential Equations*, SIAM Publications, Philadelphia (PA), 1992.
- [21] M. K. Ng, R. H. Chan, T. Chan, and A. Yip, *Cosine Transform Preconditioners for High Resolution Image Reconstruction*, Linear Algebra Appl., 316 (2000), pp. 89–104.
- [22] M. K. Ng, R. H. Chan, and W. C. Tang, *A fast algorithm for deblurring models with Neumann boundary conditions*, SIAM J. Sci. Comput., 21 (1999), pp. 851–866.
- [23] M. K. Ng, S. Serra-Capizzano, and C. Tablino-Possio, *Numerical behaviour of multigrid methods for symmetric Sinc-Galerkin systems*, Numer. Linear Algebra Appl., 12-2/3 (2005), pp. 261–269.
- [24] D. Noutsos, S. Serra-Capizzano, and P. Vassalos, *Matrix algebra preconditioners for multilevel Toeplitz systems do not insure optimal convergence rate*, Theoret. Computer Science, 315 (2004), pp. 557–579.
- [25] J. Ruge and K. Stüben, *Algebraic Multigrid*, in *Frontiers in Applied Mathematics: Multigrid Methods*, S. McCormick Ed., SIAM, Philadelphia (PA), (1987), pp. 73-130.

- [26] S. Serra-Capizzano, *Multi-iterative methods*, Comput. Math. Appl., 26 (1993), pp. 65–87.
- [27] S. Serra-Capizzano, *Convergence analysis of two grid methods for elliptic Toeplitz and PDEs matrix sequences*, Numer. Math., 92-3 (2002), pp. 433–465.
- [28] S. Serra-Capizzano, *Matrix algebra preconditioners for multilevel Toeplitz matrices are not superlinear*, Linear Algebra Appl., 343-344 (2002), pp. 303–319.
- [29] S. Serra-Capizzano, *A note on anti-reflective boundary conditions and fast deblurring models*, SIAM J. Sci. Comput., 25-3 (2003), pp. 1307–1325.
- [30] S. Serra-Capizzano and C. Tablino-Possio, *Multigrid methods for multilevel circulant matrices*, SIAM J. Sci. Comput., 26-1 (2004), pp. 55–85.
- [31] S. Serra-Capizzano and E. Tyrtysnikov, *How to prove that a preconditioner can not be superlinear*, Math. Comput., 72 (2003), pp. 1305–1316.
- [32] C. Tablino Possio, *V-cycle optimal convergence for DCT-III matrices*, 2007, in press.
- [33] E. Tyrtysnikov, *Circulant preconditioners with unbounded inverses.*, Linear Algebra Appl., 216 (1995), pp. 1–23.
- [34] U. Trottenberg, C. W. Oosterlee, and A. Schüller, *Multigrid*, Academic Press, London, 2001.