A block multigrid strategy for two-dimensional coupled PDEs

Marco Donatelli[1], Ali Dorostkar[2], Mariarosa Mazza[1], Maya Neytcheva[2], and Stefano Serra-Capizzano[1,2]

[1]Department of Science and High Technology, University of Insubria, via Valleggio 11, 22100, Como, Italy
[2]Department of Information Technology, division of Scientific Computing, Uppsala University, Box 337 SE-751 05, Uppsala, Sweden

Abstract

We consider the solution of linear systems of equations, arising from the finite element approximation of coupled differential boundary value problems. Letting the fineness parameter tend to zero gives rise to a sequence of large scale structured two-by-two block matrices. We are interested in the efficient iterative solution of the so arising linear systems, aiming at constructing optimal preconditioning methods that are robust with respect to the relevant parameters of the problem. We consider the case when the originating systems are solved by a preconditioned Krylov method, as inner solver, and propose an efficient preconditioner for that, based on the Generalized Locally Toeplitz framework.

In this paper, we exploit the almost two-level block Toeplitz structure of the arising block matrix. We provide a spectral analysis of the underlying matrices and then, by exploiting the spectral information, we design a multigrid method with an ad hoc grid transfer operator. As shown in the included examples, choosing the damped Jacobi or Gauss-Seidel methods as smoothers and using the resulting solver as preconditioner leads to a competitive strategy that outperforms some aggregation-based algebraic multigrids, widely employed in the relevant literature.

Keywords: multilevel Toeplitz matrices; multigrid methods; coupled PDEs; saddle point problem

1 Problem setting

We are interested in solving large linear systems arising from the finite element approximation of a coupled system of partial differential equations (PDEs). As an example we consider the linear elasticity problem in saddle point form. Such a problem can be viewed as a subproblem of a more general coupled system of PDEs arising from the so-called glacial isostatic adjustment (GIA) model, used in Geophysics to describe the response of the Earth to redistribution of mass due to alternating glaciation and deglaciation periods, cf. e.g., [33, 34, 20]. When the Earth is modeled as a flat homogeneous incompressible material body and only its elastic response is considered, we obtain a two-dimensional coupled system of PDEs, which in its simplest form reads as follows

$$\begin{aligned} -2\mu\,\Delta u_1 + \mu\,\frac{\partial}{\partial x_2}\left(\frac{\partial u_2}{\partial x_1} - \frac{\partial u_1}{\partial x_2}\right) - c_1\frac{\partial u_1}{\partial x_1} + \mu\frac{\partial p}{\partial x_1} &= f_1, \\ -2\mu\,\Delta u_2 + \mu\,\frac{\partial}{\partial x_1}\left(\frac{\partial u_1}{\partial x_2} - \frac{\partial u_2}{\partial x_1}\right) - c_2\frac{\partial u_2}{\partial x_2} + \mu\frac{\partial p}{\partial x_2} &= f_2 \\ \mu\nabla\cdot\mathbf{u} - \rho\,p &= 0, \end{aligned} \tag{1}$$

Here $u_1(x_1, x_2)$ and $u_2(x_1, x_2)$ are the displacements in x_1 and x_2 directions, respectively, $(x_1, x_2) \in \Omega \subset \mathbb{R}^2$, λ and μ are the so-called Lamé coefficients, for simplicity assumed not to vary in space, $\mathbf{c} = [c_1, c_2]^T$ is an advection vector, $\rho = \mu^2/\lambda$ and $\rho = 0$ in the case of compressible materials. The two unknowns are the displacements $\mathbf{u} = [u_1, u_2]^T$ and the pressure p. Discretizing (1) by the stable finite element method (FEM) pair Q1isoQ1 (cf. [5]), we obtain a linear system with a two-by-two block matrix of saddle point form,

Email addresses: marco.donatelli@uninsubria.it (Marco Donatelli), ali.dorostkar@it.uu.se (Ali Dorostkar), mariarosa.mazza@uninsubria.it (Mariarosa Mazza), maya.neytcheva@it.uu.se (Maya Neytcheva), stefano.serrac@uninsubria.it (Stefano Serra-Capizzano)

$$\mathcal{A} = \begin{bmatrix} K & B^T \\ B & -\rho M \end{bmatrix}. \tag{2}$$

Here, M is the mass matrix, ρ is a positive integer, and B and B^T correspond to discrete divergent and gradient operators, respectively. The pivot block K itself is a two-by-two block matrix where the structure is due to imposing the so-called separate displacement ordering on the components of the vector $\mathbf{u}$ (see [13] for details). Furthermore, discretizing the problem of interest for a sequence of discretization parameters we obtain a sequence of matrices of size that grows to infinity as the approximation error tends to zero. In other words, the more accurate the approximation is, the larger the related system size becomes. This rules out the direct methods as demanding too much computer resources, and other methods as preconditioned Krylov or multigrid methods have to be applied. Most of the preconditioning strategies for two-by-two block systems as in (2) are based on the following factorization of $\mathcal{A}$

$$\mathcal{A} = \begin{bmatrix} K & B^T \\ B & -\rho M \end{bmatrix} = \begin{bmatrix} K & 0 \\ B & -S \end{bmatrix} \begin{bmatrix} I_1 & K^{-1}B^T \\ 0 & I_2 \end{bmatrix},$$

where S is the negative *Schur complement* of $\mathcal{A}$ defined as $S = \rho M + BK^{-1}B^T$ and I_1 and I_2 are identity matrices of proper order. Two examples of preconditioners are

$$\mathcal{B}_1 = \begin{bmatrix} K & 0 \\ B & -\widehat{S} \end{bmatrix}, \qquad \mathcal{B}_2 = \begin{bmatrix} K & 0 \\ 0 & -\widehat{S} \end{bmatrix}, \tag{3}$$

where $\widehat{S}$ is an approximation of the exact Schur complement S.

It is well-known that a necessary condition for the above preconditioners to be efficient is that $\widehat{S}$ is an high quality approximation of S. Some approaches can be found in [12, 21, 3]. Another necessary condition is to solve the linear system with K accurately enough, thus, we need efficient inner solvers. Since the considered problem is elliptic, the Algebraic Multigrid (AMG) method is a suitable choice. As observed in [12], however, solving systems with K is the most time consuming part when applying the preconditioner. In addition, in three dimensions the memory demands become rather prohibitive.

Multigrid methods for Toeplitz matrices were firstly investigated in [14, 7, 18] and extended to multilevel Toeplitz matrices in [15, 30, 27]. The main contributions of these works were in the definition of proper grid transfer operators and in the convergence analysis of the corresponding two-grid method. The analysis of the V-cycle has been provided later in [2, 1]. More recently some specific applications [10] and aggregation grid transfer operators have been considered [11, 4]. A multigrid method for Toeplitz matrices with block symbol has been proposed in [19], even though, up to our knowledge, when the block symbol is not diagonal it lacks of a convergence analysis yet. On the other hand, the block symbol is becoming a popular theoretical tool [17] and relevant applications, such as those considered in this paper, are related to block (multilevel) Toeplitz matrices, e.g. [9].

In this paper we exploit the fact that, up to low-rank perturbations, the matrix block K is 2-level block Toeplitz matrix. We provide a spectral analysis of it and its 2×2 matrix-valued symbol to design a 2D block multigrid with an ad hoc grid transfer operator and formulate our multigrid for a more general k-level Toeplitz matrix, associated to a $s \times s$ matrix-valued symbol. Choosing damped Jacobi or Gauss-Seidel methods as smoothers, the resulting method reveals to be more efficient than some of the AMG methods in use.

The paper is organized as follows. In Section 2 we introduce the multilevel Toeplitz matrices and write explicitly K and its symbol. In Section 3 we describe the general notion of multigrid strategies and recall the algebraic multigrid for Toeplitz matrices with scalar-valued symbol. In Section 4 we present the newly developed multigrid method for multilevel Toeplitz matrices with matrix-valued symbol and apply it to the matrix K. In Section 5 we illustrate the performance of the method with numerical tests.

2 Toeplitz matrix-sequences associated to matrix-valued symbols

Our aim is to efficiently solve a linear system

$$K\mathbf{u} = \mathbf{b}, \quad \mathbf{u}, \mathbf{b} \in \mathbb{C}^N, \tag{4}$$

where $K \in \mathbb{C}^{N \times N}$ is, for instance, the head-block of $\mathcal{A}$ defined in (2), taking advantage of its structure and especially of its spectral features. To do that, we need the notion of a multilevel block Toeplitz matrix associated with a matrix-valued symbol.

Throughout the paper, we use the following notation. Let $\mathcal{M}_s$ be the linear space of the complex $s \times s$ matrices and let $f : G \to \mathcal{M}_s$, with $G \subseteq \mathbb{R}^l$, $l \geqslant 1$, measurable set. We say that f belongs to $L^p(G)$ (resp. is measurable) if all its components $f_{ij} : G \to \mathbb{C}$, $i, j = 1, \ldots, s$, belong to $L^p(G)$ (resp. are measurable) for $1 \leqslant p \leqslant \infty$. Moreover, we denote by $\mathcal{I}_k$ the k-dimensional cube $(-\pi, \pi)^k$ and define $L^p(k, s)$ as the linear space of k-variate functions $f : \mathcal{I}_k \to \mathcal{M}_s$, $f \in L^p(\mathcal{I}_k)$. Let $\mathbf{n} := (n_1, \ldots, n_k)$ be a multi-index in $\mathbb{N}^k$ and set $\hat{n} := \prod_{i=1}^k n_i$.

Definition 1. Let the Fourier coefficients of a given function f, defined as $f \in L^1(k, s)$, be

$$\hat{f}_{\mathbf{j}} := \frac{1}{(2\pi)^k} \int_{\mathcal{I}_k} f(\theta) \mathrm{e}^{-\iota \langle \mathbf{j}, \theta \rangle} \, d\theta \in \mathcal{M}_s, \qquad \mathbf{j} = (j_1, \ldots, j_k) \in \mathbb{Z}^k, \tag{5}$$

where $\langle \mathbf{j}, \theta \rangle = \sum_{t=1}^k j_t \theta_t$ and the integrals in (5) are computed componentwise. Then, the $\mathbf{n}$th Toeplitz matrix associated with f is the matrix of order $s\hat{n}$ given by

$$T_{\mathbf{n}}(f) = \sum_{|j_1| < n_1} \cdots \sum_{|j_k| < n_k} \left[J_{n_1}^{(j_1)} \otimes \cdots \otimes J_{n_k}^{(j_k)} \right] \otimes \hat{f}_{\mathbf{j}},$$

where $\otimes$ denotes the (Kronecker) tensor product of matrices. The term $J_t^{(l)}$ is the matrix of order t whose (i, j) entry equals 1 if $i - j = l$ and zero otherwise. The set $\{T_{\mathbf{n}}(f)\}_{\mathbf{n} \in \mathbb{N}^k}$ is called the *family of k-level Toeplitz matrices generated by f*, that in turn is referred to as the *generating function or the symbol of* $\{T_{\mathbf{n}}(f)\}_{\mathbf{n} \in \mathbb{N}^k}$.

As we see in the next section, a multigrid strategy for Toeplitz matrices requires information about the symbol of the coefficient matrix to define both the projector and the matrix at the coarse level. Such spectral information compactly contained in the symbol is crucial, not only in the Toeplitz setting, but even when the coefficient matrix is Toeplitz up to low-rank perturbations. In order to deal with the low-rank perturbations and to show that they do not affect the symbol, we need first to introduce the definition of spectral distribution in the sense of the eigenvalues and of the singular values for a generic matrix-sequence $\{A_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{N}^k}$ and then – the generalized locally Toeplitz (GLT) algebra. In short, the latter is an algebra containing sequences of matrices including the Toeplitz sequences with Lebesgue integrable symbols and virtually any sequence of matrices coming from 'reasonable' approximations by local discretization methods (finite differences, finite elements, isogeometric analysis, etc) of partial differential equations.

Definition 2. Let $f : G \to \mathcal{M}_s$ be a measurable function, defined on a measurable set $G \subset \mathbb{R}^l$ with $l \geqslant 1$, $0 < m_l(G) < \infty$. Let $\mathcal{C}_0(\mathbb{K})$ be the set of continuous functions with compact support over $\mathbb{K} \in \{\mathbb{C}, \mathbb{R}_0^+\}$ and let $\{A_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{N}^k}$ be a sequence of matrices with eigenvalues $\lambda_j(A_{\mathbf{n}})$, $j = 1, \ldots, M$ and singular values $\sigma_j(A_{\mathbf{n}})$, $j = 1, \ldots, M$, where $M \equiv M(\mathbf{n})$ is the size of $A_{\mathbf{n}}$.

- $\{A_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{N}^k}$ is *distributed as the pair (f, G) in the sense of the eigenvalues,* in symbols

  $$\{A_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{N}^k} \sim_\lambda (f, G),$$

  if the following limit relation holds for all $F \in \mathcal{C}_0(\mathbb{C})$:

  $$\lim_{\mathbf{n} \to \infty} \frac{1}{M} \sum_{j=1}^M F(\lambda_j(A_{\mathbf{n}})) = \frac{1}{m_k(G)} \int_G \frac{\mathrm{tr}(F(f(t)))}{s} dt. \tag{6}$$

- $\{A_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{N}^k}$ is *distributed as the pair (f, G) in the sense of the singular values,* in symbols

  $$\{A_{\mathbf{n}}\}_{\mathbf{n} \in \mathbb{N}^k} \sim_\sigma (f, G),$$

  if the following limit relation holds for all $F \in \mathcal{C}_0(\mathbb{R}_0^+)$:

  $$\lim_{\mathbf{n} \to \infty} \frac{1}{M} \sum_{j=1}^M F(\sigma_j(A_{\mathbf{n}})) = \frac{1}{m_k(G)} \int_G \frac{\mathrm{tr}(F(|f(t)|))}{s} dt. \tag{7}$$

Remark 1. Denote by $\lambda_1(f),\dots,\lambda_s(f)$ and by $\sigma_1(f),\dots,\sigma_s(f)$ the eigenvalues and the singular values of a $s\times s$ matrix-valued function f, respectively. If f is smooth enough, an informal interpretation of the limit relation (6) (resp. (7)) is that when the matrix-size of $A_{\mathbf{n}}$ is sufficiently large, then M/s eigenvalues (resp. singular values) of $A_{\mathbf{n}}$ can be approximated by a sampling of $\lambda_1(f)$ (resp. $\sigma_1(f)$) on a uniform equispaced grid of the domain G, and so on until the last M/s eigenvalues can be approximated by an equispaced sampling of $\lambda_s(f)$ (resp. $\sigma_s(f)$) in the domain.

If f is a real-valued function, the following theorem due to Szegö holds:

Theorem 1 ([16])**.** *Let $f\in L^1(1,1)$ be a real-valued function. Then, $\{T_{\mathbf{n}}(f)\}_{\mathbf{n}\in\mathbb{N}^k}\sim_\lambda(f,\mathcal{I}_1)$.*

In the case where f is a Hermitian matrix-valued function, previous theorem can be extended as follows:

Theorem 2 ([32])**.** *Let $f\in L^1(k,s)$ be a Hermitian matrix-valued function. Then, $\{T_{\mathbf{n}}(f)\}_{\mathbf{n}\in\mathbb{N}^k}\sim_\lambda(f,\mathcal{I}_k)$.*

Without going into details of the GLT algebra (see the pioneering work [31] by Tilli for describing the spectrum of one-dimensional differential operators and the generalization contained in [28, 29] for multi-variate differential operators), here we list some properties of the GLT sequences, used when proving that a sequence of Toeplitz matrices, up to low-rank corrections, is a GLT sequence and that its symbol is not affected by the low-rank perturbation.

GLT1 Each GLT sequence has a singular value symbol $f(x,\theta)$ for $(x,\theta)\in[0,1]^k\times[-\pi,\pi]^k$ with $l=2k\geqslant 2$ according to the second item in Definition 2. If the sequence is Hermitian, then the distribution also holds in the eigenvalue sense.

GLT2 The set of GLT sequences form a $*$-algebra, i.e., it is closed under linear combinations, products, inversion (whenever the symbol vanishes, at most, in a set of zero Lebesgue measure), conjugation. Hence, the sequence obtained via algebraic operations on a finite set of given GLT sequences is still a GLT sequence and its symbol is obtained by performing the same algebraic manipulations as on the corresponding symbols of the input GLT sequences.

GLT3 Every Toeplitz sequence generated by an $L^1(k,s)$ function $f=f(\theta)$ is a GLT sequences and its symbol is f, with the specifications reported in item **GLT1**. We note that the function f does not depend on the spacial variables $x\in[0,1]^k$.

GLT4 Every sequence which is distributed as the constant zero in the singular value sense is a GLT sequence with symbol 0.

Using Definition 1, we can now explicitly express the symbol of the matrix K. Denote $\mathbf{n}=(n_1,n_2)$ and $\widehat{n}=n_1n_2$. From (1) we have that K is a two-by-two block matrix of size $N=2\widehat{n}$, that is

$$K=\begin{bmatrix}K_{11}&K_{12}\\K_{21}&K_{22}\end{bmatrix},\qquad K_{ij}\in\mathbb{C}^{\widehat{n}\times\widehat{n}},\qquad i,j=1,2. \tag{8}$$

As is seen from (1), K can be symmetric positive definite ($\mathbf{c}=\mathbf{0}$) or nonsymmetric. Consider the symmetric case. If for simplicity, we discretize the original problem by a finite difference method (FDM), we obtain that $K_{ij}=T_{\mathbf{n}}(f_{ij})$, $i,j=1,2$, where $T_{\mathbf{n}}(f_{ij})$ is a 2-level Toeplitz matrix generated by $f_{ij}:\mathcal{I}_2\to\mathbb{C}$, $i,j=1,2$, with (see [13])

$$\begin{aligned}f_{11}(\theta_1,\theta_2)&=4-2\cos\theta_1(1+\cos\theta_2),\\ f_{12}(\theta_1,\theta_2)&=f_{21}(\theta_1,\theta_2)=\sin\theta_1\sin\theta_2,\\ f_{22}(\theta_1,\theta_2)&=4-2\cos\theta_2(1+\cos\theta_1).\end{aligned} \tag{9}$$

From Theorem 1, it holds that $\{T_{\mathbf{n}}(f_{ij})\}_{\mathbf{n}\in\mathbb{N}^2}\sim_\lambda(f_{ij},\mathcal{I}_1)$, $i,j=1,2$.

It is easy to see that through a proper permutation matrix Π of size $2\widehat{n}$ we can write

$$\Pi K\Pi^T=\Pi\begin{bmatrix}T_{\mathbf{n}}(f_{11})&T_{\mathbf{n}}(f_{12})\\T_{\mathbf{n}}(f_{21})&T_{\mathbf{n}}(f_{22})\end{bmatrix}\Pi^T=T_{\mathbf{n}}(f), \tag{10}$$

where $f:\mathcal{I}_2\to\mathcal{M}_2$ is defined as follows

$$f(\theta_1,\theta_2)=\begin{bmatrix}f_{11}(\theta_1,\theta_2)&f_{12}(\theta_1,\theta_2)\\f_{21}(\theta_1,\theta_2)&f_{22}(\theta_1,\theta_2)\end{bmatrix}.$$

Therefore, K is a 2-level block Toeplitz matrix associated to the $\mathcal{M}_2$-valued function f. Note that this permutation of the matrix K imposes the structure that arises if, when discretizing (1), the displacements are ordered per mesh point, i.e., the separate displacement ordering is not imposed. Since f is symmetric, from Theorem 2 it holds that $\{T_{\mathbf{n}}(f)\}_{\mathbf{n}\in\mathbb{N}^2} \sim_\lambda (f, \mathcal{I}_k)$. To write explicitly the permutation matrix Π, let us define by e_j, $j = 1, \ldots, 2\widehat{n}$ the jth column of the identity matrix of size $2\widehat{n}$ and by π_j, $j = 1, \ldots, 2\widehat{n}$ the jth column of Π. Then,

$$\pi_j = \begin{cases} e_{2j-1} & j = 1, \ldots, \widehat{n} \\ e_{2(j-\widehat{n})} & j = \widehat{n}+1, \ldots, 2\widehat{n} \end{cases}. \tag{11}$$

In other words, Π is the $2\widehat{n} \times 2\widehat{n}$ matrix whose first $\widehat{n}$ columns are the odd columns of $I_{2\widehat{n}}$, while the remaining ones are the even columns of the same matrix.

If instead of finite differences we use the Q1isoQ1 FEM scheme for discretizing the original problem, then we obtain $K_{ij} = T_{\mathbf{n}}(f_{ij}) + E_{\mathbf{n}}^{(ij)}$, $i, j = 1, 2$, where $E_{\mathbf{n}}^{(ij)}$ is a low-rank perturbation whose rank grows at most proportionally to $\sqrt{\widehat{n}}$. This means that, $\{E_{\mathbf{n}}^{(ij)}\}_{\mathbf{n}\in\mathbb{N}^2} \sim_\sigma 0$ and so, by the **GLT4**, the sequence $\{E_{\mathbf{n}}^{(ij)}\}_{\mathbf{n}\in\mathbb{N}^2}$ is a GLT sequence with symbol identically zero. Using **GLT3**, also $\{T_{\mathbf{n}}(f_{ij})\}_{\mathbf{n}\in\mathbb{N}^2}$ is a GLT sequence with symbol f_{ij} and then, by **GLT2**, the sequence $\{T_{\mathbf{n}}(f_{ij}) + E_{\mathbf{n}}^{(ij)}\}_{\mathbf{n}\in\mathbb{N}^2}$ is a GLT sequence with the same symbol. As a consequence, it is clear that the symbol does not depends on the scheme (FDM or Q1isoQ1 FEM) used to discretize the problem. In the light of this, in the next sections, we do not consider the low-rank perturbation and discuss only the construction of multigrid methods for Toeplitz matrices.

Remark 2. Note that discretizing the original problem using the Q1isoQ1 finite elements means that the matrix K corresponds to a Q1 (bilinear basis functions on a quadrilateral mesh) discretization of the part of the first two equations in (1), that contains only derivatives of the displacements.

3 Multigrid methods for Toeplitz matrices

This section is divided in two parts. First, we recall the basic idea of a multigrid method, then we focus on algebraic multigrid methods for multilevel Toeplitz matrices with scalar-valued symbol.

3.1 Multigrid methods

When a classical stationary iterative method is used to solve a linear system, the error components corresponding to large eigenvalues are damped efficiently, while the error components corresponding to the small eigenvalues are reduced slowly. Since in the discretized PDE the former correspond to rough error modes, while the latter to smooth error modes, methods like Jacobi are known as *smoothers*. The main aim of a multigrid (MG) method is to combine a smoother with some strategy able to damp the error components corresponding to the small eigenvalues, using some geometrically or algebraically constructed hierarchy of linear systems.

Let $A\mathbf{u} = \mathbf{b}$ be the linear system we want to solve, with $\mathbf{u}, \mathbf{b} \in \mathbb{C}^N$ and $A \in \mathbb{C}^{N\times N}$ Hermitian positive definite matrix. Fix $m+1$ integers $N = N_0 > N_1 > \cdots > N_m > 0$, where $0 < m < N$ denotes the maximum number of levels we decided to use. To define a multigrid method the following ingredients are needed for every level $i = 0, \ldots, m-1$:

1. appropriate smoothers $\mathcal{S}_i$, $\tilde{\mathcal{S}}_i$, and the corresponding smoothing steps ν_i, $\tilde{\nu}_i$;
2. restriction operators $R_i : \mathbb{C}^{N_i} \to \mathbb{C}^{N_{i+1}}$ and prolongation operators $P_i : \mathbb{C}^{N_{i+1}} \to \mathbb{C}^{N_i}$ to transfer a quantity between levels i and $i+1$;
3. the matrix at the coarser level $A_{i+1} \in \mathbb{C}^{N_{i+1}\times N_{i+1}}$ ($A_0 = A$, $\mathbf{b}_0 = \mathbf{b}$).

One iteration of MG in the *V-cycle version* consists of the following steps:

- ν_i pre-smoothing steps are performed using $\mathcal{S}_i$;
- the current iteration is corrected using the coarser level, process which is known as *coarse grid correction*. More precisely, the residual $r_i \in \mathbb{C}^{N_i}$ is computed and restricted to the coarse grid obtaining r_{i+1}, which is used to solve the error equation on the coarse grid

$$A_{i+1}e_{i+1} = r_{i+1},$$

by a recursive application of MG. The error e_{i+1} is interpolated back to obtain the finer level error e_i which is used to update the current iteration. The iteration matrix of the coarse grid correction is

$$CGC_i = I_{N_i} - P_i A_{i+1}^{-1} R_i A_i;$$

- the iterate is improved by $\tilde{\nu}_i$ steps performed using $\tilde{\mathcal{S}}_i$.

The following algorithm summarize one iteration of a V-cycle MG.

Algorithm 1 $\mathcal{MG}(i, \mathbf{u}_i, \mathbf{b}_i)$

if $i = m$ **then**
  $\mathbf{u}_m \leftarrow A_m^{-1}\mathbf{r}_m$
else
  $\mathbf{u}_i \leftarrow \mathcal{S}_i^{\nu_i}(\mathbf{u}_i, \mathbf{b}_i)$
  $\mathbf{r}_i \leftarrow b_i - A_i\mathbf{u}_i$
  $\mathbf{r}_{i+1} \leftarrow R_i\mathbf{r}_i$
  $\mathbf{e}_{i+1} \leftarrow \mathcal{MG}(i+1, 0_{N_{i+1}}, \mathbf{r}_{i+i})$
  $\mathbf{e}_i \leftarrow P_i\mathbf{e}_{i+1}$
  $\mathbf{u}_i \leftarrow \mathbf{u}_i + \mathbf{e}_i$
  $\mathbf{u}_i \leftarrow \tilde{\mathcal{S}}_i^{\tilde{\nu}_i}(\mathbf{u}_i, \mathbf{b}_i)$
end

For a given initial guess $\mathbf{u}^{(0)}$, one MG iteration can be described as $\mathbf{u}^{(l+1)} = \mathcal{MG}(0, \mathbf{u}^{(l)}, \mathbf{b}_0)$, for $l = 0, 1, \ldots$ A simple choice for both pre- and post-smoothing is relaxed Richardson

$$\begin{aligned} \mathcal{S}_i(\mathbf{u}_i, \mathbf{b}_i) &= S_i\mathbf{u}_i + \omega_i\mathbf{b}_i, \quad S_i = (I_{N_i} - \omega_i A_i), \quad \omega > 0, \\ \tilde{\mathcal{S}}_i(\mathbf{u}_i, \mathbf{b}_i) &= \tilde{S}_i\mathbf{u}_i + \tilde{\omega}_i\mathbf{b}_i, \quad \tilde{S}_i = (I_{N_i} - \tilde{\omega}_i A_i), \quad \tilde{\omega} > 0. \end{aligned} \tag{12}$$

When deriving convergence estimates for MG, usually, R_i is chosen to be the adjoint of P_i and the coarse grid matrix A_{i+1} is chosen as $P_i^H A_i P_i$. These conditions are known in the related literature as *Galerkin conditions* and the resulting method is the so-called *algebraic multigrid* (AMG). Note that, if the projectors have full rank, the matrix at the next coarse level is nonsingular and still Hermitian positive definite.

Let us define $\|\cdot\|_X = \|X^{1/2}\cdot\|_2$, where $\|\cdot\|_2$ is the usual Euclidean norm on $\mathbb{C}^N$ and X is an Hermitian positive definite matrix. The following theorem presents a convergence result for AMG ([1, 23]).

Theorem 3. *Let m, N be integers satisfying $0 < m < N$ and suppose that $A \in \mathbb{C}^{N\times N}$ is a Hermitian positive definite matrix; given a sequence of $m+1$ positive integers $N = N_0 > N_1 > \cdots > N_m > 0$, let $P_i : \mathbb{C}^{N_{i+1}} \to \mathbb{C}^{N_i}$, $i = 0, \ldots, m-1$ be a full-rank matrix and $R_i = P_i^H$. Define $A_0 = A$, $\mathbf{b}_0 = \mathbf{b}$, $A_{i+1} = P_i^H A_i P_i$, $i = 0, \ldots, m-1$ and choose two classes of iterative methods $\mathcal{S}_i$, $\tilde{\mathcal{S}}_i$ whose iteration matrices are S_i, $\tilde{S}_i$, respectively. If there exists three real positive numbers α_i, β_i, γ_i such that*

$$\|S_i^{\nu_i}\mathbf{u}\|_{A_i}^2 \leqslant \|\mathbf{u}\|_{A_i}^2 - \alpha_i\|S_i^{\nu_i}\mathbf{u}\|_{A_i^2}^2 \quad \forall \mathbf{u} \in \mathbb{C}^{N_i} \qquad \text{(pre-smoothing property)} \tag{13}$$

$$\|\tilde{S}_i^{\tilde{\nu}_i}\mathbf{u}\|_{A_i}^2 \leqslant \|\mathbf{u}\|_{A_i}^2 - \beta_i\|\mathbf{u}\|_{A_i^2}^2 \quad \forall \mathbf{u} \in \mathbb{C}^{N_i} \qquad \text{(post-smoothing property)} \tag{14}$$

$$\|CGC_i\mathbf{u}\|_{A_i}^2 \leqslant \gamma_i\|\mathbf{u}\|_{A_i^2}^2 \quad \forall \mathbf{u} \in \mathbb{C}^{N_i} \qquad \text{(approximation property)} \tag{15}$$

for every $i = 0, \ldots, m-1$, defined

$$\delta_{\text{pre}} := \min_{0\leqslant i<m} \frac{\alpha_i}{\gamma_i}, \quad \delta_{\text{post}} := \min_{0\leqslant i<m} \frac{\beta_i}{\gamma_i},$$

it holds that $\delta_{\text{post}} \leqslant 1$ and

$$\|AMG_0\|_A \leqslant \sqrt{\frac{1-\delta_{\text{post}}}{1+\delta_{\text{pre}}}} < 1,$$

where AMG_0 is the V-cycle iteration matrix.

Remark 3. From Theorem 3, the sequence $\{\mathbf{u}^{(l)}\}_{l\in\mathbb{N}}$ generated by AMG converges to the solution of $A\mathbf{u}=\mathbf{b}$. Moreover, when at least one of the following conditions holds

$$\inf_t \min_{0\leqslant i<m(t)} \frac{\alpha_i}{\gamma_i} > 0, \quad \inf_t \min_{0\leqslant i<m(t)} \frac{\beta_i}{\gamma_i} > 0,$$

the convergence is optimal, that is, the method converges with a constant error reduction not depending on N and m.

Remark 4. Observe that the smoothing properties (13)-(14) are related to the smoothers, but not to the projectors, while the approximation property (15) depends only on the choice of the projectors.

3.2 AMG for Toeplitz matrices with scalar-valued symbol

AMG for multilevel Toeplitz matrices with scalar-valued symbols has been investigated in detail in [15, 30, 27, 2, 1]. After a brief presentation of the convergence results for this method, we introduce a AMG for multilevel Toeplitz matrices with matrix-valued symbols. To guarantee the optimal convergence of the AMG method, certain smoothing and approximation properties have to be fulfilled. As observed in Remark 4, they can be treated separately. We start with the conditions that ensure the validity of smoothing properties. First of all, note that since the diagonal of the coefficient matrix is a multiple of the identity, for Toeplitz matrices the relaxed Richardson in (12) are equivalent to the weighted Jacobi iterations. As the following proposition shows, using appropriate weights, these smoothers fulfill the smoothing properties.

Proposition 1 ([1]). *Let $A=T_{\mathbf{n}}(f)$ with $f:\mathcal{I}_k\to\mathbb{C}$ nonnegative and not identically zero. Defined $S=I-\omega A$ and $\tilde{S}=I-\tilde{\omega}A$, if*

$$0\leqslant\omega,\tilde{\omega}\leqslant\frac{2}{\|f\|_\infty},$$

then there exist $\alpha,\beta>0$ such that the two smoothing properties

$$\|S^\nu\mathbf{u}\|_A^2 \leqslant \|\mathbf{u}\|_A^2-\alpha\|S^\nu\mathbf{u}\|_{A^2}^2 \quad \forall\mathbf{u}\in\mathbb{C}^{\hat{n}}, \tag{16}$$

$$\|\tilde{S}^\theta\mathbf{u}\|_A^2 \leqslant \|\mathbf{u}\|_A^2-\beta\|\mathbf{u}\|_{A^2}^2 \quad \forall\mathbf{u}\in\mathbb{C}^{\hat{n}}, \tag{17}$$

hold with $\nu,\theta\in\mathbb{N}$.

Apart from the smoother, to define an AMG method, a further ingredient is necessary, namely the projector.

Assume that f has only one zero of order at most 2. Under this hypothesis, in the 1-level case, the different strategies proposed in [14, 7, 18, 27, 2] are equivalent. For a fixed $N_i=2^{t-i}-1$, with t an integer number, chose as projectors the product between $T_{N_i}(p_i)$, with a nonnegative trigonometric polynomial p_i and the transpose of the following cutting matrix $K_{N_i}\in\mathbb{R}^{N_{i+1}\times N_i}$

$$K_{N_i}=\begin{bmatrix} 0 & 1 & 0 & & & & & \\ & & 0 & 1 & 0 & & & \\ & & & \ddots & \ddots & \ddots & & \\ & & & & & 0 & 1 & 0 \end{bmatrix}. \tag{18}$$

Then the matrix at the so-defined coarse level is still a Toeplitz matrix $A_{i+1}=P_i^HA_iP_i=T_{N_{i+1}}(f_{i+1})$, where

$$f_{i+1}(\theta)=\frac{1}{2}\left[p_i^2f_i\left(\frac{\theta}{2}\right)+p_i^2f_i\left(\frac{\theta}{2}+\pi\right)\right].$$

The k-level case has been discussed in [15, 30, 27, 1]. There, the authors consider k-level Toeplitz matrices $T_{\mathbf{n}}(f)$ of order $\mathbf{n}=(2^t-1)\mathbf{e}\in\mathbb{N}^k$, where $\mathbf{e}=(1,\ldots,1)\in\mathbb{N}^k$ and t is a positive integer. Starting with $\mathbf{n}_0=\mathbf{n}$, the order on the coarse levels is defined as $\mathbf{n}_i=(2^{t-i}-1)\mathbf{e}$. The projector P_i is defined as the product between a matrix $T_{\mathbf{n}_i}(p_i)$ and the transpose of a k-level cutting matrix $K_{\mathbf{n}_i}$, where p_i is a nonnegative k-variate trigonometric polynomial. More precisely, the cutting matrix $K_{\mathbf{n}_i}$ is obtained as the Kronecker product of k 1-level cutting matrices $K_{(n_i)_1},\ldots,K_{(n_i)_k}$. If f has only one zero of order at most 2 in every direction, all these choices preserve the k-level Toeplitz structure on the coarse levels.

Multigrid methods are widely used for solving PDEs, independently of the boundary conditions. In that context, the convergence analysis is performed by Local Fourier analysis (LFA) which does not consider the

boundary effects, i.e. it assumes periodic boundary conditions or an infinite domain [6]. Analogously, since Toeplitz matrices are difficult to manipulate, multigrid convergence results are usually investigated using matrix algebra approximations such as τ or circulant matrices having the same symbol, i.e. spectral distribution, as the original Toeplitz matrix (see [8] for more details).

In the following, we recall the conditions that ensure the validity of the approximation property in the circulant algebra. For a fixed $\theta \in \mathbb{R}^k$, define the set of all *corners points* $\Omega(\theta)$ and the set of all *mirror points* $\mathcal{M}(\theta)$ as

$$\Omega(\theta) = \{\eta | \eta_j \in \{\theta_j, \theta_j + \pi\}\} \text{ and } \mathcal{M}(\theta) = \Omega(\theta) \backslash \{\theta\}.$$

Let $\{f_i\}_{i=0}^m$ be the sequence of symbols on the coarse levels, where

$$f_{i+1}(\theta) = \frac{1}{2^k} \sum_{\eta \in \Omega(\theta)} p_i^2 f_i(\eta). \tag{19}$$

Proposition 2 ([1])**.** *Let $A_i = C_{\mathbf{n}_i}(f_i)$ be the circulant matrix generated by f_i a k-variate nonnegative trigonometric polynomial. Let θ_i^0 be the unique zero of f_i in $[0,\pi]^k$ and let $CGC_i = I_{\hat{n}_i} - P_i(P_i^H A_i P_i)^{-1} P_i^H A_i$ with $P_i = C_{\mathbf{n}_i}(p_i) K_{\mathbf{n}_i}^T$ and p_i a nonnegative k-variate trigonometric polynomial. Then the approximation property* (15) *holds if for all $\theta \in [0,\pi]^k$ p_i is such that*

$$\limsup_{\theta \to \theta_i^0} \left| \frac{p_i(\eta)}{f_i(\theta)} \right| < +\infty, \quad \eta \in \mathcal{M}(\theta), \tag{20}$$

$$\sum_{\eta \in \Omega(\theta)} p_i^2(\eta) > 0. \tag{21}$$

Remark 5. If θ_i^0 is a zero of order q for f_i, by condition (20), $\eta \in \mathcal{M}(\theta_i^0)$ is such that p_i has a zero at η at least of the same order. From conditions (20)–(21) holds that $p_i(\theta_i^0) \neq 0$, which means that the projectors are full-rank and that the ill-conditioned subspace of A_i is in the image of the projector P_i.

Proposition 3. *Let f_{i+1} be defined as in* (19)*, and suppose that p_i satisfies* (20)-(21)*. Then, if θ_i^0 is a zero of order q for f_i, $\theta_{i+1}^0 = 2\theta_i^0$ is a zero of order q for f_{i+1}.*

If θ_i^0 has order (at most) $2q$, a natural choice for p_i is

$$p_i(\theta) = c \cdot \prod_{j=1}^{k} [1 + \cos(\theta_j - (\theta_i^0)_j)]^q. \tag{22}$$

with c constant. Indeed, the polynomial p_i has a zero of order $2q$ at $\eta \in \mathcal{M}(\theta_i^0)$ and does not vanish at θ_i^0.

4 AMG for Toeplitz matrices with matrix-valued symbol

In this section we describe AMG for multilevel Toeplitz matrices with matrix-valued symbol. The 1-level case has been treated in [19]. As observed in that paper, when the generating function is an $s \times s$ diagonal matrix-valued function, a multigrid method on the whole matrix can be seen as s independent multigrid methods for 1-level Toeplitz matrices with scalar-valued symbols. This approach can be applied also to the multilevel case. Let $f : \mathcal{I}_k \to \mathcal{M}_s$ defined as

$$f(\theta) = \begin{bmatrix} f_{11}(\theta) & 0 & \cdots & 0 \\ 0 & f_{22}(\theta) & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & f_{ss}(\theta) \end{bmatrix}.$$

If we assume that $f_{jj} : \mathcal{I}_k \to \mathbb{C}$, $j = 1, \ldots, s$ has only a single isolated zero in $\mathcal{I}_k$ of order (at most) $2q$, we can define s AMG methods as discussed in Subsection 3.2 (one for each f_{jj}) choosing polynomials like in (22) as symbol for the projectors.

In [19], the authors consider the more general case when the generating matrix is not diagonal, but Hermitian and positive definite with a constant basis of eigenvectors. In brief, the main idea is to diagonalize the generating function and to choose the projector in view of the location of the zeroes of its eigenvalues. We formally extend

this idea also in the multidimensional setting and to non-constant basis of eigenvectors. Let $f : \mathcal{I}_k \to \mathcal{M}_s$ be Hermitian positive definite and let us diagonalize $f(\theta)$ as follows

$$f(\theta) = Q(\theta)\Lambda(\theta)Q(\theta)^H, \tag{23}$$

where

$$\Lambda(\theta) = \begin{bmatrix} \lambda_1(\theta) & 0 & \cdots & 0 \\ 0 & \lambda_2(\theta) & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \lambda_s(\theta) \end{bmatrix},$$

with $\lambda_j : \mathcal{I}_k \to \mathbb{C}$, $j = 1, \ldots, s$ is a nonnegative function.

As Proposition 1 can be straightforwardly extended to any positive definite matrix, the smoothing properties are ensured.

Proposition 4. *Let $A = T_{\mathbf{n}}(f)$ with $f : \mathcal{I}_k \to \mathcal{M}_s$ Hermitian positive definite and defined according to equation (23). Let $S = I - \omega A$ and $\tilde{S} = I - \tilde{\omega} A$. If*

$$0 \leqslant \omega, \tilde{\omega} \leqslant \frac{2}{\max_{j=1,\ldots,s} \|\lambda_j\|_\infty},$$

then there exist $\alpha, \beta > 0$ such that the smoothing properties (16) and (17) hold with $\nu, \theta \in \mathbb{N}$.

Proof. Since f is Hermitian positive definite then A is positive definite and so $A = QDQ^H$, where D is the diagonal matrix with the eigenvalues of A on the diagonal. Note that the notation $A \leqslant B$ for any given two Hermitian matrices, means that the matrix $B - A$ is positive semidefinite. Note that (16) is equivalent to require that

$$S^\nu A S^\nu \leqslant A - \alpha S^\nu A^2 S^\nu,$$

which reduces to

$$D(I - \omega D)^{2\nu} \leqslant D - \alpha D^2 (I - \omega D)^{2\nu} \tag{24}$$

thanks to the expression of S. The matrix inequality (24) is implied by the function inequalities

$$\lambda_j (I - \omega \lambda_j)^{2\nu} \leqslant \lambda_j - \alpha \lambda_j^2 (I - \omega \lambda_j)^{2\nu}, \qquad j = 1, \ldots, s.$$

Performing the same function study presented in [1, Proposition 3], we deduce that the smoothing property (16) follows whenever the parameter ω satisfies the inequalities $0 \leqslant \omega \leqslant 2/\max_{j=1,\ldots,s} \|\lambda_j\|_\infty$.

Similarly, we can prove the smoothing property (17) when $0 \leqslant \tilde{\omega} \leqslant 2/\max_{j=1,\ldots,s} \|\lambda_j\|_\infty$. □

To define the projector at level i, $i = 0, \ldots, m-1$ for a fixed $\mathbf{n}_i = (2^{t-i} - 1)\mathbf{e}$, we use the following cutting matrix

$$K_{\mathbf{n}_i}^{[s]} = K_{(n_i)_1} \otimes \cdots \otimes K_{(n_i)_k} \otimes I_s,$$

where I_s is the $s \times s$ identity matrix and $K_{(n_i)_\ell} \in \mathbb{R}^{(n_{i+1})_\ell \times (n_i)_\ell}$, $\ell = 1, \ldots, k$ is either defined as in (18) for $\mathbf{n}_i = (2^{t-i} - 1)\mathbf{e}$ or in the case of $\mathbf{n}_i = (2^{t-i} + 1)\mathbf{e}$ it is chosen as

$$K_{(n_i)_\ell} = \begin{bmatrix} 1 & 0 & & & & \\ & 0 & 1 & 0 & & \\ & & & & \ddots & \\ & & & & 0 & 1 \end{bmatrix}.$$

Assuming that for the matrix A_i the associated symbol f_i has all eigenvalues functions $\lambda_j^{(i)}$ with only a single isolated zero at the same point $\theta_i^0 \in \mathcal{I}_k$ of order (at most) $2q$ for every $j = 1, \ldots, s$, we define the projector as

$$P_{\mathbf{n}_i}^{[s]} = T_{\mathbf{n}_i}(p_i)(K_{\mathbf{n}_i}^{[s]})^T, \qquad p_i(\theta) = c \cdot \prod_{j=1}^{k} [1 + \cos(\theta_j - (\theta_i^0)_j)]^q \cdot I_s. \tag{25}$$

The matrix at the coarse grid is obtained by the Galerkin approach, that is as $A_{i+1} = (P_{\mathbf{n}_i}^{[s]})^T A_i P_{\mathbf{n}_i}^{[s]}$.

4.1 AMG for the block matrix K discretized by Q1 FEM

Let us recall that in our test problem (10) we have $k = 2$, $s = 2$, and $f : \mathcal{I}_2 \to \mathcal{M}_2$ is the following symmetric positive defined matrix function

$$f(\theta_1, \theta_2) = \begin{bmatrix} f_{11}(\theta_1, \theta_2) & f_{12}(\theta_1, \theta_2) \\ f_{12}(\theta_1, \theta_2) & f_{22}(\theta_1, \theta_2) \end{bmatrix} = \begin{bmatrix} 4 - 2\cos\theta_1(1 + \cos\theta_2) & \sin\theta_1 \sin\theta_2 \\ \sin\theta_1 \sin\theta_2 & 4 - 2\cos\theta_2(1 + \cos\theta_1) \end{bmatrix} \tag{26}$$

according to (9). Note that functions f_{11} and f_{22} have a zero in $(0, 0)$ of order 2.

By direct computation of the zeros of the characteristic polynomial of $f(\theta_1, \theta_2)$, we obtain the following two eigenvalue functions

$$\lambda_1(\theta_1, \theta_2) = \frac{f_{11}(\theta_1, \theta_2) + f_{22}(\theta_1, \theta_2) + \sqrt{(f_{11}(\theta_1, \theta_2) - f_{22}(\theta_1, \theta_2))^2 + 4f_{12}^2(\theta_1, \theta_2)}}{2},$$
$$\lambda_2(\theta_1, \theta_2) = \frac{f_{11}(\theta_1, \theta_2) + f_{22}(\theta_1, \theta_2) - \sqrt{(f_{11}(\theta_1, \theta_2) - f_{22}(\theta_1, \theta_2))^2 + 4f_{12}^2(\theta_1, \theta_2)}}{2}.$$

By using the definition of f in (26) and the identity $2 - 2\cos\theta = 4\sin^2\frac{\theta}{2}$ we can write λ_1, λ_2 explicitly as follows

$$\lambda_1(\theta_1, \theta_2) = 4 - (\cos\theta_1 + \cos\theta_2) - 2\cos\theta_1\cos\theta_2 + \sqrt{(\cos\theta_1 - \cos\theta_2)^2 + \frac{1}{4}(1 - \cos 2\theta_1)(1 - \cos 2\theta_2)},$$
$$\lambda_2(\theta_1, \theta_2) = 4 - (\cos\theta_1 + \cos\theta_2) - 2\cos\theta_1\cos\theta_2 - \sqrt{(\cos\theta_1 - \cos\theta_2)^2 + \frac{1}{4}(1 - \cos 2\theta_1)(1 - \cos 2\theta_2)}.$$

Since f is symmetric, there exists a unitary matrix $Q \in \mathcal{M}_2$ such that $f = Q\Lambda Q^H$, where

$$\Lambda = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix} \qquad \text{and} \qquad Q = \begin{pmatrix} q_1 & -q_2 \\ q_2 & q_1 \end{pmatrix}.$$

Computing the eigenvectors corresponding to λ_1, λ_2, we have

$$f_{11}q_1 + f_{12}q_2 = \lambda_1 q_1 \iff q_2 = \frac{\lambda_1 - f_{11}}{f_{12}} q_1.$$

Using the expressions for λ_1 and f_{11} we obtain

$$\lambda_1 - f_{11} = \cos\theta_1 - \cos\theta_2 + \sqrt{(\cos\theta_1 - \cos\theta_2)^2 + \sin^2\theta_1 \sin^2\theta_2}.$$

Note that the factor $(\lambda_1 - f_{11})/f_{12} \to 0$ when θ_1 and θ_2 approach zero. Therefore, for the ill-conditioned subspace associated to $\theta_1, \theta_2 \to 0$ the eigenvector $(q_1(\theta_1, \theta_2), q_2(\theta_1, \theta_2))^T \to (1, 0)^T$ and hence $Q \to I_2$. It follows that when $\theta_1, \theta_2 \to 0$ the symbol (26) is almost diagonal and the construction of the projector (25) requires only the computation of the zeros, with their order, of the eigenvalue functions λ_1 and λ_2.

Both eigenvalues λ_1 and λ_2 have a zero of order 2 in $(0, 0)$ (see Figures 1(a) and 1(b)). In fact, if we fix $\theta_2 = 0$, then

$$\lim_{\theta_1 \to 0} \frac{\lambda_1(\theta_1, 0)}{\theta_1^2} = \lim_{\theta_1 \to 0} \frac{4 - (\cos\theta_1 + 1) - 2\cos\theta_1 + \sqrt{(\cos\theta_1 - 1)^2}}{\theta_1^2} = \lim_{\theta_1 \to 0} \frac{4 - 4\cos\theta_1}{\theta_1^2} = 2,$$
$$\lim_{\theta_1 \to 0} \frac{\lambda_2(\theta_1, 0)}{\theta_1^2} = \lim_{\theta_1 \to 0} \frac{4 - (\cos\theta_1 + 1) - 2\cos\theta_1 - \sqrt{(\cos\theta_1 - 1)^2}}{\theta_1^2} = \lim_{\theta_1 \to 0} \frac{2 - 2\cos\theta_1}{\theta_1^2} = 1. \tag{27}$$

Similarly, for a fixed $\theta_1 = 0$,

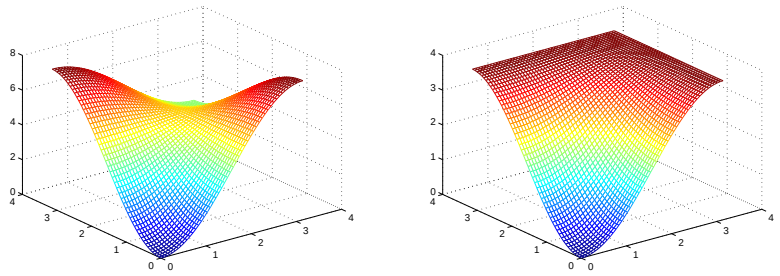
$$\lim_{\theta_2 \to 0} \frac{\lambda_1(0, \theta_2)}{\theta_2^2} = \lim_{\theta_2 \to 0} \frac{4 - (\cos\theta_2 + 1) - 2\cos\theta_2 + \sqrt{(\cos\theta_2 - 1)^2}}{\theta_2^2} = \lim_{\theta_2 \to 0} \frac{4 - 4\cos\theta_2}{\theta_2^2} = 2,$$
$$\lim_{\theta_2 \to 0} \frac{\lambda_2(0, \theta_2)}{\theta_2^2} = \lim_{\theta_2 \to 0} \frac{4 - (\cos\theta_2 + 1) - 2\cos\theta_2 - \sqrt{(\cos\theta_2 - 1)^2}}{\theta_2^2} = \lim_{\theta_2 \to 0} \frac{2 - 2\cos\theta_2}{\theta_2^2} = 1. \tag{28}$$

Due to (27) and (28), we expect the eigenvalues of the generating function at the coarse level to have a zero of order 2 at the origin and therefore we define the projectors as

$$P_{\mathbf{n}_i}^{[2]} = T_{\mathbf{n}_i}(p)(K_{\mathbf{n}_i}^{[2]})^T, \quad K_{\mathbf{n}_i}^{[2]} = K_{(n_i)_1} \otimes K_{(n_i)_2} \otimes I_2,$$

where

$$p(\theta_1,\theta_2) = 4\prod_{j=1}^{2}[1+\cos(\theta_j)]\cdot I_2 = \begin{bmatrix} (2+2\cos\theta_1)(2+2\cos\theta_2) & 0 \\ 0 & (2+2\cos\theta_1)(2+2\cos\theta_2) \end{bmatrix}.$$



(a) $\lambda_1(\theta_1,\theta_2)$, $(\theta_1,\theta_2)\in[0,\pi)^2$ (b) $\lambda_2(\theta_1,\theta_2)$, $(\theta_1,\theta_2)\in[0,\pi)^2$

Note that the restriction of a vector from a fine to a coarser grid is obtained by product with the matrix $(P_{\mathbf{n}_i}^{[2]})^T = K_{\mathbf{n}_i}^{[2]} T_{\mathbf{n}_i}(p)$, where $T_{\mathbf{n}_i}(p)$ has at most nine nonzero entries in every row and $K_{\mathbf{n}_i}^{[2]}$ simply performs a proper down sampling of two entries every four, due to the tensor with I_2. In detail, let z be defined on the fine grid and let y be its projection into the coarser grid obtained by applying $(P_{\mathbf{n}_i}^{[2]})^T$, then the entries of y arranged as a 2D array can be computed as

$$y_{i,j} = z_{2i-2,2j+2} + z_{2i-2,2j-2} + 2z_{2i,2j-2} + 2z_{2i-2,2j} + 4z_{2i,2j} + 2z_{2i+2,2j} + 2z_{2i,2j+2} + z_{2i+2,2j+2} + z_{2i+2,2j-2}.$$

Similarly, the prolongation of a vector from a coarse to the finer grid is obtained by product with the matrix $P_{\mathbf{n}_i}^{[2]} = T_{\mathbf{n}_i}(p)(K_{\mathbf{n}_i}^{[2]})^T$, where $(K_{\mathbf{n}_i}^{[2]})^T$ adds the zeros corresponding to the new grid points and $T_{\mathbf{n}_i}(p)$ performs the average of nine "near" points: again there is a jump of size 2 because of the tensor by I_2 in the formula of $K_{\mathbf{n}_i}^{[2]}$. From a geometrical point of view, the grid transfer operator is the standard bilinear interpolation associated to the linear tensor B-spline [8], again tensored with I_2. The coarser matrices are computed by the Galerkin approach as $A_{i+1} = (P_{\mathbf{n}_i}^{[2]})^T A_i P_{\mathbf{n}_i}^{[2]}$, for $i = 0,\dots,m-1$ in a setup phase. According to the results in [2, 1], the matrix A_{i+1} inherits the same 2-level block Toeplitz structure with 2×2 blocks of A_i. In particular, thanks to (19), the generating function of A_{i+1} is again $f(\theta_1,\theta_2)$ defined in (9) up to a factor 16. Hence A_{i+1} could be simply computed by rediscretization with double step size scaling properly the projector.

Regarding the smoother, because of its simplicity, just as in the scalar-case, we opt for the weighted Jacobi method. Note that since the constant coefficient of f_{11} and f_{22} is the same and equal to 4, then weighted Jacobi iteration is simply the weighted Richardson iteration considered in Proposition 4. Hence, according to Proposition 4 and scaling the weight by a factor 4, we have that $0 \leqslant \omega, \tilde{\omega} \leqslant 8/\max_{j=1,2}\|\lambda_j\|_\infty$. Computing

$$\max_{j=1,2}\|\lambda_j\|_\infty = \|\lambda_1\|_\infty = \lambda_1(0,\pi) = 8,$$

we have that $0 \leqslant \omega, \tilde{\omega} \leqslant 1$. Therefore, following the multi-iterative idea (see [26]), we use both pre-smoother and post-smoother with parameters $\omega = 1$ and $\tilde{\omega} = 2/3$, respectively. Moreover, as shown in the next section, we test also the performances of our method choosing Gauss-Seidel method as smoother.

5 Numerical results

In this section we use the block multigrid defined in Section 4 (denoted as GLT-MG) as a preconditioner for Krylov methods in order to solve the linear system $K\mathbf{u} = \mathbf{b}$, with K as in (8) and $n_1 = n_2 = 2^t + 1$. To be

precise, we solve the $2(2^t+1)^2 \times 2(2^t+1)^2$ linear system associated to the matrix $\Pi K \Pi^T$, with Π defined as in (11). In the following we refer to this permuted matrix again as K.

Although our analysis has been focused on symmetric positive definite matrices, in the following we test GLT-MG also on the nonsymmetric matrices coming from the discretization of the original problem without neglecting the advection term. We choose the preconditioned conjugate gradient (PCG, [24]) method for symmetric matrices and the generalized minimal residual (GMRES [25]) method for the case where the matrices are nonsymmetric. Both methods are applied using the built-in Matlab functions `pcg` and `gmres`. Moreover, for all involved methods (GLT-MG, PCG and GMRES), we choose as initial guess the zero vector and use a relative stopping criterion $tol \in \{10^{-3}, 10^{-6}\}$.

Recall that the GLT-MG uses weighted Jacobi (in this case we label our method as 'GLT-MG$_{\text{J}}$') or Gauss-Seidel (in this case we label our method as 'GLT-MG$_{\text{GS}}$') as pre-smoothers and post-smoothers. The parameters of Jacobi method are $\omega = 1$ and $\tilde{\omega} = 2/3$. In our test we make only one pre- and post-smoothing steps and perform only one V-cycle.

Assuming to know the true solution $\tilde{\mathbf{u}}$, we compute the relative error of the approximated solution $\mathbf{u}$ as the measure of accuracy.

In [12] the authors solve $K\mathbf{u} = \mathbf{b}$ with an aggregation-based algebraic multigrid (AGMG), see [22]. AGMG performs one forward and one backward Gauss-Seidel sweep for pre- and post-smoothing, respectively, and performs also a K-cycle, i.e., two Krylov accelerated iterations at each intermediate level. The main iterative solver in AGMG is the Generalized Conjugate Residual method.

We check the validity of the strategy proposed in this paper and compare it with AGMG both in terms of iterations and of approximation error. We choose $\tilde{\mathbf{u}}$ as an equispaced sampling of the function

$$\varphi(x_1, x_2) = \sin(3x_1) + \sin(3x_2), \qquad (x_1, x_2) \in \Omega$$

and as right-hand side $\mathbf{b} = K\tilde{\mathbf{u}}$. The numerical tests with GLT-MG are performed in Matlab and AGMG is used via its Matlab interface.

Table 1 shows a comparison between of GLT-MG$_{\text{J}}$ and GLT-MG$_{\text{GS}}$ (used as preconditioners for PCG) and AGMG in terms of iterations and approximation error, in the case when K is symmetric. We observe that for both stopping criteria the GLT-MG method requires less iterations than the AGMG method, especially when the Gauss-Seidel smoother is used. Moreover, the accuracy of the computed iterative solution, achieved when using GLT-MG is better than that provided by AGMG. Note that for both methods the number of iterations does not increase when increasing the size of the problem, which means that both GLT-MG and AGMG are of optimal order.

tol	$n_1 = n_2$	GLT-MG$_{\text{J}}$		GLT-MG$_{\text{GS}}$		AGMG	
		Iter	Error	Iter	Error	Iter	Error
10^{-3}	2^5+1	4	4.27e-004	3	2.09e-004	6	1.66e-003
	2^6+1	4	4.24e-004	3	2.17e-004	6	3.19e-003
	2^7+1	4	4.56e-004	3	2.25e-004	6	2.51e-003
10^{-6}	2^5+1	8	8.09e-007	5	5.95e-007	11	8.87e-007
	2^6+1	8	4.47e-007	5	5.73e-007	12	1.88e-006
	2^7+1	8	4.60e-007	5	5.77e-007	12	3.73e-006

Table 1: Symmetric case - Comparison between GLT-MG$_{\text{J,GS}}$ (used as preconditioner for PCG) and AGMG method both in terms of iterations and approximation error fixed $tol = 10^{-3}, 10^{-6}$.

In Table 2 we show the number of iterations and accuracy achieved by GLT-MG$_{\text{J}}$, GLT-MG$_{\text{GS}}$ (used as preconditioners for GMRES) and AGMG when K is a nonsymmetric matrix. Even in this case GLT-MG converges with a fewer number of iterations compared to AGMG and the resulting iteratively computed solution for the considered test problem is more accurate.

Conclusions and future works

In this paper, based on the GLT framework, we derive a very efficient multigrid preconditioner for matrices originating from FEM discretization of a coupled system of PDEs. We illustrate the technique on a 2D linear elasticity problem, discretized using Q1 FEM elements. We take advantage of the 2-level block Toeplitz

tol	$n_1 = n_2$	GLT-MG$_{\rm J}$		GLT-MG$_{\rm GS}$		AGMG	
		Iter	Error	Iter	Error	Iter	Error
10^{-3}	2^5+1	4	4.41e-004	3	9.76e-005	7	4.15e-003
	2^6+1	4	3.23e-004	3	8.55e-005	7	3.74e-003
	2^7+1	4	2.92e-004	3	7.77e-005	6	2.78e-003
	2^8+1	4	2.87e-004	3	7.23e-005	6	2.51e-003
10^{-6}	2^5+1	8	2.45e-007	5	3.28e-007	16	3.71e-006
	2^6+1	8	2.08e-007	5	2.74e-007	17	3.84e-006
	2^7+1	8	1.60e-007	5	2.61e-007	16	7.04e-006
	2^8+1	8	1.53e-007	5	2.47e-007	15	7.75e-006

Table 2: Nonsymmetric case - Comparison between GLT-MG$_{\rm J,GS}$ (used as preconditioner for GMRES) and AGMG method both in terms of iterations and approximation error fixed $tol = 10^{-3}, 10^{-6}$.

structure of the so-arising matrix to design a multigrid method that reveals to be numerically competitive with the AGMG method, used successfully as a preconditioner for the considered class of problems. The preconditioning technique can be applied in a more general context, for instance, for solving inner systems for block-preconditioner as in (3).

A future work will concern a detailed analysis of the proposed multigrid in terms of convergence and optimality. Furthermore, we plan to extend this strategy to linear systems arising from other FEM discretizations.

Acknowledgements

The work of the first and third authors is partly supported by the Italian grant MIUR - PRIN 2012 N. 2012MTE38N and by GNCS-INDAM (Italy), while the work of the fifth author is partly supported by Donation KAW 2013.0341 from the Knut & Alice Wallenberg Foundation in collaboration with the Royal Swedish Academy of Sciences, supporting Swedish research in mathematics.

References

[1] A. Aricò, M. Donatelli, AV-cycle multigrid for multilevel matrix algebras: proof of optimality, Numer. Math. 105 (4) (2007) 511–547.

[2] A. Aricò, M. Donatelli, S. Serra.Capizzano, V-cycle optimal convergence for certain (multilevel) structured linear systems, SIAM J. Matrix Anal. Appl. 26 (1) (2004) 186–214.

[3] E. Bängtsson, M. Neytcheva, Numerical simulations of glacial rebound using preconditioned iterative solution methods, Appl. Math. 50 (3) (2005) 183–201.

[4] M. Bolten, M. Donatelli, T. Huckle, Analysis of smoothed aggregation multigrid methods based on Toeplitz matrices, Electron. Trans. Numer. Anal. 44 (2015) 25–52.

[5] D. Braess, Finite elements: Theory, Fast Solvers, and applications in Solid Mechanics, Second Edition, Cambridge University Press, Cambridge, 2001.

[6] A. Brandt, Rigorous quantitative analysis of multigrid, I: Constant coefficients two-level cycle with l_2-norm, SIAM J. Numer. Anal. 31 (6) (1994) 1695–1730.

[7] R. H. Chan, Q.-S. Chang, H.-W. Sun, Multigrid method for ill-conditioned symmetric Toeplitz systems, SIAM J. Sci. Comp. 19 (2) (1998) 516–529.

[8] M. Donatelli, An algebraic generalization of local Fourier analysis for grid transfer operators in multigrid based on Toeplitz matrices, Numer. Linear Algebra Appl. 17 (2-3) (2010) 179–197.

[9] M. Donatelli, M. Molteni, V. Pennati, S. Serra-Capizzano, Multigrid methods for cubic spline solution of two point (and 2D) boundary value problems, App. Numer. Math. (2014) DOI: 10.1016/j.apnum.2014.04.004.

[10] M. Donatelli, S. Serra-Capizzano, On the regularizing power of multigrid-type algorithms, SIAM J. Sci. Comp. 27 (6) (2006) 2053–2076.

[11] M. Donatelli, S. Serra-Capizzano, D. Sesana, Multigrid methods for Toeplitz linear systems with different size reduction, BIT 52 (2) (2012) 305–327.

[12] A. Dorostkar, M. Neytcheva, B. Lund, Numerical and computational aspects of some block-preconditioners for saddle point systems, Parallel Comput. 49 (2015) 164–178.

[13] A. Dorostkar, M. Neytcheva, S. Serra-Capizzano, Spectral analysis of coupled PDEs and of their Schur complements via the notion of Generalized Locally Toeplitz sequences, Tech. rep., Technical Report 2015-008, Department of Information Technology, Uppsala University (2015).

[14] G. Fiorentino, S. Serra-Capizzano, Multigrid methods for Toeplitz matrices, Calcolo 28 (3-4) (1991) 283–305.

[15] G. Fiorentino, S. Serra-Capizzano, Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions, SIAM J. Sci. Comp. 17 (5) (1996) 1068–1081.

[16] U. Grenander, G. Szegö, Toeplitz Forms and Their Applications, vol. 321, Second Edition, Chelsea, New York, 1984.

[17] T. Huckle, Compact Fourier Analysis for designing multigrid methods, SIAM J. Sci. Comp. 31 (1) (2008) 644–666.

[18] T. Huckle, J. Staudacher, Multigrid preconditioning and Toeplitz matrices, Electr. Trans. Numer. Anal. 13 (2002) 81–105.

[19] T. Huckle, J. Staudacher, Multigrid methods for block Toeplitz matrices with small size blocks, BIT 46 (1) (2006) 61–83.

[20] B. Lund, J. O. Näslund, Glacial isostatic adjustment: implications for glacially induced faulting and nuclear waste repositories, in: C. B. Connor, N. A. Chapman, L. J. Connor (eds.), Volcanic and Tectonic Hazard Assessment for Nuclear Facilities, Cambridge University Press, 2009, pp. 142–155.

[21] M. Neytcheva, On element-by-element Schur complement approximations, Linear Algebra Appl., special Issue: Devoted to the 2nd NASC 08 Conference in Nanjing (NSC) 434 (11) (2011) 2308–2324.

[22] Y. Notay, AGMG, `http://homepages.ulb.ac.be/~ynotay/AGMG/`.

[23] J. Ruge, K. Stüben, Algebraic multigrid, in Frontiers in Applied Mathematics: Multigrid Methods, S. McCormick Ed. SIAM, Philadelphia, 1987.

[24] Y. Saad, Iterative Methods for Sparse Linear Systems, PWS Publishing, Boston, MA, 2003.

[25] Y. Saad, M. H. Schultz, Gmres: A generalized minimal residual algorithm for solving nonsymmetric linear systems, SIAM J. Sci. Stat. Comp. 7 (3) (1986) 856–869.

[26] S. Serra, Multi-iterative methods, Comput. Math. Appl. 26 (4) (1993) 65–87.

[27] S. Serra-Capizzano, Convergence analysis of Two-Grid methods for elliptic Toeplitz and PDEs matrix-sequences, Numer. Math. 92 (3) (2002) 433–465.

[28] S. Serra-Capizzano, Generalized locally Toeplitz sequences: spectral analysis and applications to discretized partial differential equations, Linear Algebra Appl. 366 (2003) 371–402.

[29] S. Serra-Capizzano, The GLT class as a generalized Fourier Analysis and applications, Linear Algebra Appl. 419 (1) (2006) 180–233.

[30] H.-W. Sun, X.-Q. Jin, Q.-S. Chang, Convergence of the multigrid method of ill-conditioned block Toeplitz systems, BIT 41 (1) (2001) 179–190.

[31] P. Tilli, Locally Toeplitz sequences: spectral properties and applications, Linear Algebra Appl. 278 (1) (1998) 91–120.

[32] P. Tilli, A note on the spectral distribution of Toeplitz matrices, Linear Multilinear A. 45 (2-3) (1998) 147–159.

[33] P. Wu, Deformation of an incompressible viscoelastic flat earth with Power Law Creep: a Finite Element approach, Geophys. J. Int. 108 (1) (1992) 35–51.

[34] P. Wu, Using commercial finite element packages for the study of earth deformations, sea levels and the state of stress, Geophys. J. Int. 158 (2) (2004) 401–408.