Calcolo (2024) 61:57
https://doi.org/10.1007/s10092-024-00612-1

Fast algebraic multigrid for block-structured dense systems arising from nonlocal diffusion problems

Minghua Chen[1] · Rongjun Cao[2] · Stefano Serra-Capizzano[3,4]

Received: 30 August 2023 / Revised: 3 May 2024 / Accepted: 26 August 2024 /
Published online: 13 September 2024

Abstract

Algebraic multigrid (AMG) is one of the most efficient iterative methods for solving large structured systems of equations. However, how to build/check restriction and prolongation operators in practical AMG methods for nonsymmetric *structured* systems is still an interesting open question in its full generality. The present paper deals with the block-structured dense and Toeplitz-like-plus-cross systems, including *nonsymmetric* indefinite and symmetric positive definite (SPD) ones, arising from nonlocal diffusion problems. The simple (traditional) restriction operator and prolongation operator are employed in order to handle such block-structured dense and Toeplitz-like-plus-cross systems, which are convenient and efficient when employing a fast AMG. We provide a detailed proof of the two-grid convergence of the method for the considered SPD structures. The numerical experiments are performed in order to verify the convergence with a computational cost of only $\mathscr{O}(N \log N)$ arithmetic operations, by exploiting the fast Fourier transform, where N is the number of the grid points. To the

The work of the first author was supported by the National Natural Science Foundation of China under Grant No. 12471381, the Science Fund for Distinguished Young Scholars of Gansu Province under Grant No. 23JRRA1020 and the Fundamental Research Funds for the Central Universities under grant lzujbky-2023-06.

✉ Stefano Serra-Capizzano
stefano.serrac@uninsubria.it ; stefano.serra@it.uu.se

Minghua Chen
chenmh@lzu.edu.cn

Rongjun Cao
20230126@ynu.edu.cn

1 School of Mathematics and Statistics, Gansu Key Laboratory of Applied Mathematics and Complex Systems, Lanzhou University, Lanzhou 730000, People's Republic of China

2 School of Mathematics and Statistics, Yunnan University, Kunming 650500, People's Republic of China

3 Department of Science and High Technology, University of Insubria, Via Valleggio 11, 22100 Como, Italy

4 Division of Scientific Computing, Department of Information Technology, Uppsala University-ITC, Lägerhyddsv. 2, hus 2, P.O. Box 337, 751 05 Uppsala, Sweden

best of our knowledge, this is the first contribution regarding Toeplitz-like-plus-cross linear systems solved by means of a fast AMG.

Keywords Algebraic multigrid · Nonlocal diffusion problem · Block-structured dense system · Toeplitz-like-plus-cross system · Fast Fourier transform

Mathematics Subject Classification 65N55 · 65N35 · 15B05 · 15A18 · 65T50

1 Introduction

Large, sparse, block-structured linear systems arise in a wide variety of applications throughout computational science and engineering including advection–diffusion flow [48], image processing [37], Markov chains [49], Biot's consolidation model [40], specific saddle point problems arising e.g. in Navier-Stokes equations [8]. In the current paper we study fast algebraic multigrid methods for solving block-structured dense linear systems, stemming from nonlocal problems [2, 6, 15, 21, 29, 30, 47]. The continuous problem is discretized via piecewise quadratic polynomial collocation approximations, whose associated coefficient matrices can be expressed as 2 × 2 block structures

$$\mathscr{A}u = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} b_f \\ b_g \end{bmatrix}, \tag{1}$$

with coefficient matrices $A \in \mathbb{R}^{M\times M}$, $B \in \mathbb{R}^{M\times N}$, $C \in \mathbb{R}^{N\times M}$ and $D \in \mathbb{R}^{N\times N}$ and where the size M may not necessarily be smaller than N.

Algebraic multigrid (AMG) is one of the most efficient iterative methods for solving large-scale system of equations [41, 54]. In the past decades, AMG methods for linear systems having Toeplitz coefficient matrices with scalar entries have been widely studied [19] including elliptic PDEs [41, 44, 53, 54], fractional PDEs [25, 28, 39] and nonlocal PDEs [20, 22]. Few works have investigated the case of block entries, where the entries are small generic matrices of fixed size instead of scalars [10, 11, 23, 27, 31] especially in the case where the spectral information of the related large matrices are encoded in a spectral symbol [33]. We observe that only a limited number of papers have studied block-structured linear systems of the form (1), when dense blocks occur. For example, by defining partial prolongations operators in connection with a Galerkin coarse grid matrix, in [38] the authors design a new AMG approach for the Stokes problem. Furthermore paper [52] contains a fully aggregation-based AMG for nonlinear contact problems of saddle point type, where ad hoc transfer operators are constructed. Using approximate ideal restriction (AIR) operators, in [48] AIR AMG methods for space-time hybridizable discontinuous Galerkin discretization of advection-dominated flows are investigated. By defining the interpolation matrix with the coarse coefficient vector, the authors in [49] propose and study multilevel Markov chain Monte Carlo AMG algorithms. When involving sparse integral transfer operators towards adaptive smoothed aggregation, in [14] specialized AMG for nonsymmetric problems are considered. Finally transfer operator based on fractional approximation properties and two-grid methods convergence in norm of nonsymmetric algebraic multigrid are presented in the work [36]. However, how to build/check restriction and prolongation operators in the practice of AMG methods for nonsymmetric *sparse* systems is still an interesting open question [14, 36]. In particular, how to develop/design fast AMG for block-structured dense linear systems (1) is still a problem to be explored further, since the above special prolongation/transfer operators are not easy to be employed in connection with the fast Fourier transform.

In the current work, the simple (traditional) restriction operator and prolongation operator are used in order to handle such block-structured dense systems (1), including nonsymmetric indefinite systems, symmetric positive definite (SPD) systems, Toeplitz-plus-diagonal systems, which derive from the nonlocal problems discussed in [2, 15, 24, 29]. In general, it is still not at all easy to analyse AMG for dense stiffness matrices [3, 4, 12, 22], unless we can reduce the problem to the Toeplitz setting and we know the symbol, its zeros, and their orders [44]. Instead we focus our attention on answering such a question for a two-grid setting, since it is useful from a theoretical point of view as first step: in fact the study of the AMG convergence usually begins from the convergence analysis of the two-grid method (TGM) [39, 41, 54]. We focus our attention in providing a detailed proof of the convergence of TGM for the considered SPD linear systems. To the best of our knowledge, following previous ideas in [22], this is the first time that a fast AMG is studied for the block-structured dense linear systems as those reported in (1).

The outline of this paper is as follows. In Sect. 2, we introduce block-structured dense systems, including applications in nonlocal diffusion problems by the piecewise quadratic polynomial collocation. In Sect. 3, block-structured V-cycle AMG algorithm using fast Fourier transforms are designed for Toeplitz-like-plus-cross systems. The convergence rate of the two-grid method is analyzed in Sect. 4. To show the effectiveness of the presented schemes, results of numerical experiments are reported in Sect. 5, including a comparison with a Krylov technique. Finally, in Sect. 6 we conclude our study with relevant remarks and open problems.

2 Block-structured dense systems applications

Nonlocal diffusion problems have been used to model very different scientific phenomena occurring in various applied fields, for example in biology, particle systems, image processing, coagulation models, mathematical finance, etc. [2, 29]. Recently, the nonlocal volume-constrained diffusion problems, the so-called nonlocal model for distinguishing the nonlocal diffusion problems, attracted a wide interest of scientists [29], where the linear scalar peridynamic model can be considered as a special case [29, 47]. For example, the nonlocal peridynamic (PD) model has become an attractive emerging tool for the multiscale material simulations of crack nucleation and growth, fracture, and failure of composites [47].

Let us consider the following time-dependent nonlocal volume-constrained diffusion problem [22, 29, 47]

$$\begin{cases} u_t - \mathscr{L}_\delta u = f & \text{on } \Omega,\ t > 0, \\ u(x,0) = u_0 & \text{on } \Omega \cup \Omega_{\mathscr{I}}, \\ u = g & \text{on } \Omega_{\mathscr{I}}, t > 0, \end{cases} \tag{2}$$

where u is a sufficiently smooth function of x and t with $u_t = \frac{\partial u}{\partial t}$. The nonlocal operator $\mathscr{L}_\delta$ is defined by [29]

$$\mathscr{L}_\delta u(x,t) = \int_{B_\delta(x)} \gamma_\delta(|x-y|)\,[u(y,t) - u(x,t)]\,dy, \quad \forall x \in \Omega, \quad t \in [0,T],$$

with $B_\delta(x) = \{y \in \mathbb{R} : |y - x| < \delta\}$ denoting a neighborhood centered at x of radius δ, which is the horizon parameter and represents the size of nonlocality. The symmetric nonlocal kernel is defined as $\gamma_\delta(|x-y|) = 0$ if $y \notin B_\delta(x)$.

Before starting to discuss problem (2), we briefly review few preliminary notions regarding the piecewise quadratic polynomial collocation approximations for the corresponding stationary problem

$$\begin{cases} -\mathscr{L}_\delta u = f & \text{on } \Omega, \\ u = g & \text{on } \Omega_{\mathscr{I}}, \end{cases} \tag{3}$$

where $u = g$ denotes a volumetric constraint imposed on a volume $\Omega_{\mathscr{I}}$ that has a nonzero volume and is made to be disjoint from Ω. In order to keep the expression simple, below we assume the $\Omega = [0, 1]$ with the volumetric constraint domain $\Omega_{\mathscr{I}} = [-\delta, 0] \cup [1, 1+\delta]$, but everything can be shifted to an arbitrary interval $[a, b]$. For convenience, we focus on the special case where the kernel $\gamma_\delta(s)$ is taken to be a constant, i.e., $\gamma_\delta(s) = 3\delta^{-3}$ [22, 29, 50]. More general kernel types [29, 50] can be studied in a similar manner.

Consider the nonlocal model on the interval $\Omega = [a, b]$ with the volumetric constraint domain $\Omega_{\mathscr{I}} = [a - \delta, a] \cup [b, b + \delta]$, $0 \le \delta < b$. Define the mesh grid with uniform spatial stepsize $h = (b - a)/N$ and with mesh points

$$\begin{aligned} x_{-r} < x_{-r+\frac{1}{2}} < \cdots < x_{-\frac{1}{2}} < a = x_0 < x_{\frac{1}{2}} < x_1 < \cdots \\ \cdots < x_{N-\frac{1}{2}} < x_N = b < x_{N+\frac{1}{2}} < \cdots < x_{N+r}, \end{aligned} \tag{4}$$

where

$$r = \begin{cases} \lfloor \delta/h \rfloor, & \delta > h, \\ \lceil \delta/h \rceil, & 0 < \delta \le h, \\ 0, & \delta = 0. \end{cases}$$

Here $\lfloor \delta/h \rfloor$ denotes the greatest integer that is less than or equal to δ/h and $\lceil \delta/h \rceil$ denotes the smallest integer that is greater than or equal to δ/h.

Let the piecewise quadratic base functions $\phi_i(x)$ and $\phi_{i-\frac{1}{2}}(x)$ be defined as in [5, p. 37]. Then the piecewise Lagrange quadratic interpolant of $u(x)$ is given in [15] in the following way

$$u_Q(x) = \sum_{j=-r}^{N+r} u\left(x_j\right)\phi_j(x) + \sum_{j=-r}^{N+r-1} u\left(x_{j+\frac{1}{2}}\right)\phi_{j+\frac{1}{2}}(x), \tag{5}$$

which is used in dealing the nonlocal model in (2), leading to the block-structured dense systems expressed in (1).

Now, we introduce and discuss the discretization scheme of (3). According to the mesh grid (4), we set

$$\mathscr{N} = \left\{-r, -r+\frac{1}{2}, \ldots, -\frac{1}{2}, 0, \frac{1}{2}, \ldots, N-\frac{1}{2}, N, \ldots, N+r-\frac{1}{2}, N+r\right\},$$
$$\mathscr{N}_{in} = \left\{\frac{1}{2}, 1, \frac{3}{2}, \ldots, N-1, N-\frac{1}{2}\right\}, \quad \mathscr{N}_{out} = \mathscr{N} \setminus \mathscr{N}_{in}.$$

Define u_i as the approximate value of $u(x_i)$, $f_i = f(x_i)$ and $g_i = g(x_i)$. For convenience of implementation, we use the matrix form of the grid functions as follows

$$U_h = \left[u_1, u_2, \ldots, u_{N-1}, u_{\frac{1}{2}}, u_{\frac{3}{2}}, \ldots, u_{N-\frac{1}{2}}\right]^T,$$
$$F_h = \left[f_1, f_2, \ldots, f_{N-1}, f_{\frac{1}{2}}, f_{\frac{3}{2}}, \ldots, f_{N-\frac{1}{2}}\right]^T.$$

2.1 Nonsymmetric indefinite block-structured dense systems

By the piecewise quadratic polynomial collocation (5), it is easy to check that the standard collocation method of stationary problem (3) has the following form [16]

$$\begin{cases} -\mathscr{L}_\delta u_i = f_i, & i \in \mathscr{N}_{in}, \\ \quad\; u_i = g_i, & i \in \mathscr{N}_{out}. \end{cases} \tag{6}$$

As a consequence the numerical scheme (6) can be recast as

$$\mathscr{A}_h^N U_h = \eta_h F_h^N \text{ with } \mathscr{A}_h^N = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \quad \eta_h = 2\delta^3/h. \tag{7}$$

Here the discrete source function F_h^S absorbs the boundary conditions, which can be treated as in [22].

Note that the above matrices $A \in \mathbb{R}^{(N-1)\times(N-1)}$, $B \in \mathbb{R}^{(N-1)\times N}$, $C \in \mathbb{R}^{N\times(N-1)}$ and $D \in \mathbb{R}^{N\times N}$ are all Toeplitz matrices [13, 33]. A matrix X is Toeplitz if it is constant along the diagonals that is $X_{s,t} = x_{s-t}$ with x_l complex coefficients and l ranging according to the sizes of X. The latter implies that $X_{s,t} = X_{s+\Delta,t+\Delta}$ for every Δ

compatible with the dimensions of X: from the previous property of invariance under shift, these matrices are named also shift-invariant, especially in signal processing and operator theory. In our case we use the notation toeplitz($[c]$), $c \in \mathbb{C}^m$, for a square complex symmetric Toeplitz matrix of order m whose first column is the vector c, and the notation toeplitz($[c], [d]$), $c \in \mathbb{C}^{m_1}$, $d \in \mathbb{C}^{m_2}$, $c_1 = d_1$, for a generic matrix of sizes $m_1 \times m_2$, with first column c and first row d.

Furthermore if $\phi \in L^1(-\pi, \pi)$ by $T_n(f)$ we denote the Toeplitz matrix generated by ϕ, i.e. $(T_n(\phi))_{s,t} = \hat{\phi}_{s-t}, s, t = 1, \dots, n$, with $\hat{\phi}_k$ being the k-th Fourier coefficient of f, that is

$$\hat{\phi}_k = \frac{1}{2\pi} \int_{-\pi}^{\pi} \phi(\theta)\, e^{-\mathbf{i}k\theta}\, \mathrm{d}\theta, \quad \mathbf{i}^2 = -1, \quad k \in \mathbb{Z}.$$

In that case ϕ is called the generating function of $T_n(\phi)$ [13, 33]. Moreover $T_n(\phi)$ coincides with the matrix toeplitz($[c], [d]$), whenever c contains the Fourier coefficients $\hat{\phi}_k$, $k = 0, 1, \dots, n-1$ and d contains the Fourier coefficients $\hat{\phi}_k$, $k = 0, -1, \dots, 1-n$. Below we report in detail the matrices $A \in \mathbb{R}^{(N-1)\times(N-1)}$, $B \in \mathbb{R}^{(N-1)\times N}$, $C \in \mathbb{R}^{N\times(N-1)}$, $D \in \mathbb{R}^{N\times N}$, which, in accordance to the previous notations, are expressed as

$$\begin{aligned}
A &= \text{toeplitz}\left([a_0, a_1, a_2, \cdots, a_r, \mathbf{0}_{1\times(N-r-2)}]\right),\\
B &= \text{toeplitz}\left(\left[a_{\frac{1}{2}}, a_{\frac{3}{2}}, \cdots, a_{r-\frac{1}{2}}, \mathbf{0}_{1\times(N-1-r)}\right], \left[a_{\frac{1}{2}}, a_{\frac{1}{2}}, a_{\frac{3}{2}}, \cdots, a_{r-\frac{1}{2}}, \mathbf{0}_{1\times(N-r-1)}\right]\right),\\
C &= \text{toeplitz}\left([c_0, c_0, c_1, c_2, \cdots, c_r, \mathbf{0}_{1\times(N-r-2)}], [c_0, c_1, c_2, \cdots, c_r, \mathbf{0}_{1\times(N-r-2)}]\right),\\
D &= \text{toeplitz}\left([d_0, d_1, d_2, \cdots, d_r, \mathbf{0}_{1\times(N-r-1)}]\right),
\end{aligned}$$

with the coefficients [16]

$$\begin{aligned}
&a_0 = 12r - 2, \quad a_m = -2, \quad a_r = -1, \quad 1 \le m \le r-1,\\
&a_{m+\frac{1}{2}} = -4, \quad 0 \le m \le r-1,
\end{aligned}$$

and

$$\begin{aligned}
&c_m = -2, \quad c_{r-1} = -\frac{9}{4}, \quad c_r = \frac{1}{4}, \quad 0 \le m \le r-2,\\
&d_0 = 12r - 4, \quad d_m = -4, \quad d_r = -2, \quad 1 \le m \le r-1.
\end{aligned}$$

Since all the reported coefficients are real, the square Toeplitz matrices A and D are both real and symmetric.

2.2 Symmetric positive definite block-structured dense systems

In Remark 2.2 of [16], it is shown that the discrete maximum principle is not satisfied for the above nonsymmetric indefinite system (7), which might be trickier for the stability analysis of the high-order numerical schemes [26, 35]. As a consequence, the shifted-symmetric piecewise quadratic polynomial collocation method for nonlocal model (3) has been considered in [16], which satisfies the discrete maximum principle and admits symmetric positive definite block-structured dense coefficient matrices. Namely, the shifted-symmetric system of (6) can be recast as [16]

$$\mathscr{A}_h^S U_h = \eta_h F_h^S \quad \text{with} \quad \mathscr{A}_h^S = \begin{bmatrix} A & B \\ B^T & \widehat{A} \end{bmatrix}, \quad \eta_h = 2\delta^3/h. \tag{8}$$

Here the function F_h^S is computed as done in (6), the Toeplitz matrices $A \in \mathbb{R}^{(N-1)\times(N-1)}$, $B \in \mathbb{R}^{(N-1)\times N}$, and $\widehat{A} \in \mathbb{R}^{N\times N}$ are defined by

$$\begin{aligned} A &= \text{toeplitz}\left([a_0, a_1, a_2, \cdots, a_r, \mathbf{0}_{1\times(N-r-2)}]\right), \\ B &= \text{toeplitz}\left(\left[a_{\frac{1}{2}}, a_{\frac{3}{2}}, \cdots, a_{r-\frac{1}{2}}, \mathbf{0}_{1\times(N-1-r)}\right], \left[a_{\frac{1}{2}}, a_{\frac{1}{2}}, a_{\frac{3}{2}}, \cdots, a_{r-\frac{1}{2}}, \mathbf{0}_{1\times(N-r-1)}\right]\right), \\ \widehat{A} &= \text{toeplitz}\left([a_0, a_1, a_2, \cdots, a_r, \mathbf{0}_{1\times(N-r-1)}]\right), \end{aligned}$$

with

$$\begin{aligned} &a_0 = 12r - 2, \quad a_m = -2, \quad a_r = -1, \quad 1 \le m \le r-1, \\ &a_{m+\frac{1}{2}} = -4, \quad 0 \le m \le r-1. \end{aligned} \tag{9}$$

3 Fast AMG for block-structured dense systems

Multigrid methods are among the most efficient iterative methods for solving large scale systems of equations [41, 54]. To the best of our knowledge, there is no fast AMG for block-structured dense linear systems of the type in (1), since the special prolongation/transfer operators are not easy to be employed in connection with the fast Fourier transform. Here, the simple (traditional) transfer operator are employed in order to handle such block-structured dense systems to ensure a fast AMG showing a $\mathscr{O}(N\log N)$ complexity.

3.1 Multigrid methods

Let us first review the basic multigrid technique when applied to a scalar algebraic linear system, having in mind that our target is the efficient solution of the block-structured dense linear systems as those reported in (1). Let the finest mesh points be $x_i = a + ih$, $h = (b-a)/N$, $N = 2^K$, $i = 1 : N$, with $\Omega = (a, b)$. Define the multiple levels of grids $\mathfrak{B}^k$, $k = 1, \ldots, K$, as follows

$$\mathfrak{B}^k = \left\{ x_i^k = a + \frac{i}{2^k}(b-a), i = 1 : N_k \right\} \quad \text{with} \quad N_k = 2^k - 1, \quad k = 1, \ldots, K,$$

where $\mathfrak{B}^k$ represents not only the grid with grid spacing $h_k = 2^{K-k}h$, but also the space of vectors defined on that grid.

Given an algebraic system of the form

$$A_h u^h = b^h, \tag{10}$$

we define a sequence of subsystems on different levels

$$A_k u^k = b^k, \quad u^k \in \mathfrak{B}^k, \quad k = 1, \ldots, K.$$

Here K is the total number of levels, with $k = K$ being the finest level, i.e., $A_K = A_h$.

The traditional restriction operator $I_k^{k-1} : \mathbb{R}^{N_k} \to \mathbb{R}^{N_{k-1}}$ and prolongation operator $I_{k-1}^k : \mathbb{R}^{N_{k-1}} \to \mathbb{R}^{N_k}$ are defined by [42, pp. 438–454]

$$I_k^{k-1} = \frac{1}{4}\begin{bmatrix} 1\ 2\ 1 & & & \\ & 1\ 2\ \ 1 & & \\ & & \cdots\cdots & \\ & & & 1\ \ 2\ 1 \end{bmatrix}_{N_{k-1}\times N_k} \quad \text{and} \quad I_{k-1}^k = 2\left(I_k^{k-1}\right)^T, \tag{11}$$

which should be convenient and efficient for block-structured dense linear systems as those in (1), using AMG and fast Fourier transform. Let

$$\nu^{k-1} = I_k^{k-1}\nu^k \ \text{ with } \ \nu^{k-1} = \frac{1}{4}\left(\nu_{2i-1}^k + 2\nu_{2i}^k + \nu_{2i+1}^k\right), \quad i = 1, \ldots, N_k,$$

and

$$\nu^k = I_{k-1}^k \nu^{k-1}.$$

It may be more useful to define the linear system by using the Galerkin projection in the AMG method, where the coarse grid problem is defined by

$$A_{k-1} = I_k^{k-1} A_k I_{k-1}^k, \tag{12}$$

and the intermediate $(k, k-1)$ coarse grid correction operator is

$$T_k = I_k - I_{k-1}^k A_{k-1}^{-1} I_k^{k-1} A_k.$$

We introduce the damped Jacobi iterative operator

$$S_k = I - \omega D_k^{-1} A_k \tag{13}$$

with a weighting factor ω, as the smoothing operator, and D_k being the diagonal of A_k. Then the V-cycle multigrid algorithm can be designed similarly to Algorithm 1 in [22]: see also [53].

The basic AMG idea for solving the block-structured dense linear systems in (1) is the same as in the scalar case (10). Define a sequence of block-structured subsystems

$$\mathscr{A}_k u^k = b^k, \ \ u^k \in \mathfrak{M}^k, \ \ k = 1, \ldots, K,$$

with the multiple level of grids

$$\mathfrak{M}^k = \left\{ x^k_{i/2} = a + \frac{i/2}{2^k}(b-a), i = 1 : 2N_k + 1 \right\} \text{ with } N_k = 2^k - 1, \ \ k = 1, \ldots, K. \tag{14}$$

We have

$$\mathscr{A}_k = \begin{bmatrix} A^{(k)} & B^{(k)} \\ C^{(k)} & D^{(k)} \end{bmatrix}, \ \ u^k = \begin{bmatrix} v^k \\ w^k \end{bmatrix}, \ \ b^k = \begin{bmatrix} b^k_f \\ b^k_g \end{bmatrix}.$$

Moreover, we define the intermediate $(k, k-1)$ coarse grid correction operator as

$$\mathscr{T}_k = I_k - I^k_{k-1} \mathscr{A}^{-1}_{k-1} I^{k-1}_k \mathscr{A}_k. \tag{15}$$

Following (13), the smoothing operator is chosen to be

$$\mathscr{S}_k = I - \omega \mathscr{D}^{-1}_k \mathscr{A}_k \tag{16}$$

with weighting factor ω, and $\mathscr{D}_k$ being the diagonal matrix of $\mathscr{A}_k$. Hence the block-structured dense V-cycle multigrid method is developed in Algorithm 1, with m_1 pre-smoothing steps, m_2 post-smoothing steps, $m_1 + m_2 \geq 1$, m_1, m_2 nonnegative integers independent of the matrix order.

3.2 Fast Fourier transform for block-structured dense systems

It is well-known that for any N-by-1 vector $\mathbf{x}$, the product of a Toeplitz matrix T_N and a vector $\mathbf{x}$ can be computed by the fast Fourier transform (FFT) with the computational cost of $\mathscr{O}(N \log N)$ arithmetic operations [17, p. 12]. Indeed, given a Toeplitz matrix T_N, it can be embedded in a circulant matrix C_N of size $2N$-by-$2N$ as follows:

$$\begin{bmatrix} T_N & * \\ * & T_N \end{bmatrix} \begin{bmatrix} \mathbf{x} \\ \mathbf{0} \end{bmatrix} = \begin{bmatrix} T_N \mathbf{x} \\ \ddagger \end{bmatrix} \text{ with } C_N = \begin{bmatrix} T_N & * \\ * & T_N \end{bmatrix}. \tag{17}$$

Let us consider the fast Fourier transform algorithm for block-structured dense systems (1) at the finest level, namely,

$$\mathscr{A}_h u = \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = \begin{bmatrix} Av + Bw \\ Cv + Dw \end{bmatrix} \tag{18}$$

with Toeplitz matrices $A \in \mathbb{R}^{M \times M}$, $B \in \mathbb{R}^{M \times N}$, $C \in \mathbb{R}^{N \times M}$ and $D \in \mathbb{R}^{N \times N}$, $M < N$.

Note that a few early works on matrix–vector multiplication with Toeplitz algebraic system were focused on square matrices by fast Fourier transform (FFT) [17, 18]. As reported in [15, 17], the FFT algorithm for the rectangular matrices B and C in (18) also can be realized, which leads to the computational cost of $\mathscr{O}(N \log N)$ arithmetic

Algorithm 1 Block-structured dense V-cycle multigrid method: Define $\mathscr{B}_1 = \mathscr{A}_1^{-1}$. Assume that $\mathscr{B}_{k-1} : \mathfrak{M}^{k-1} \mapsto \mathfrak{M}^{k-1}$ is defined. We now define $\mathscr{B}_k : \mathfrak{M}^k \mapsto \mathfrak{M}^k$ as an approximate iterative solver for the equation associated with $\mathscr{A}_k u^k = b^k$.

1: Pre-smooth: Let $\mathscr{S}_{k,\omega_{pre}}$ be defined by (16) and $\begin{bmatrix} v_0^k \\ w_0^k \end{bmatrix} = \mathbf{0}$; for $l = 1 : m_1$,

$$\begin{bmatrix} v_l^k \\ w_l^k \end{bmatrix} = \begin{bmatrix} v_{l-1}^k \\ w_{l-1}^k \end{bmatrix} + \mathscr{S}_{k,\omega_{pre}} \left(\begin{bmatrix} b_f^k \\ b_g^k \end{bmatrix} - \mathscr{A}_k \begin{bmatrix} v_{l-1}^k \\ w_{l-1}^k \end{bmatrix} \right)$$

2: Coarse grid correction: Denote $\mathbf{e}^{k-1} = \begin{bmatrix} e_v^{k-1} \\ e_w^{k-1} \end{bmatrix} \in \mathfrak{M}_{k-1}$ as the approximate solution of the residual equation $\mathscr{A}_{k-1}\mathbf{e} = I_k^{k-1} \left(\begin{bmatrix} b_f^k \\ b_g^k \end{bmatrix} - \mathscr{A}_k \begin{bmatrix} v_{m_1}^k \\ w_{m_1}^k \end{bmatrix} \right)$ with the iterator $\mathscr{B}_{k-1}$ an approximate inverse of $\mathscr{A}_{k-1}$,

$$\begin{bmatrix} e_u^{k-1} \\ e_v^{k-1} \end{bmatrix} = \mathscr{B}_{k-1} I_k^{k-1} \left(\begin{bmatrix} f_k \\ g_k \end{bmatrix} - \mathscr{A}_k \begin{bmatrix} u_{m_1}^k \\ v_{m_1}^k \end{bmatrix} \right)$$

3: Post-smooth: $\begin{bmatrix} v_{m_1+1}^k \\ w_{m_1+1}^k \end{bmatrix} = \begin{bmatrix} v_{m_1}^k \\ w_{m_1}^k \end{bmatrix} + I_{k-1}^k \begin{bmatrix} e_v^{k-1} \\ e_w^{k-1} \end{bmatrix}$ and $\mathscr{S}_{k,\omega_{post}}$ is defined by (16),

$$\begin{bmatrix} v_l^k \\ w_l^k \end{bmatrix} = \begin{bmatrix} v_{l-1}^k \\ w_{l-1}^k \end{bmatrix} + \mathscr{S}_{k,\omega_{post}} \left(\begin{bmatrix} b_f^k \\ b_g^k \end{bmatrix} - \mathscr{A}_k \begin{bmatrix} v_{l-1}^k \\ w_{l-1}^k \end{bmatrix} \right), \quad l = m_1 + 2 : m_1 + m_2$$

4: Define: $\mathscr{B}_k \begin{bmatrix} b_f^k \\ b_g^k \end{bmatrix} = \begin{bmatrix} v_{m_1+m_2}^k \\ w_{m_1+m_2}^k \end{bmatrix}$.

operations and a required $\mathscr{O}(N)$ storage: indeed, this is an old idea already present in the classical book by Bini and Pan (see [9] [Chapter 3, Problem 5.1]).

In fact, the technique of embedding rectangular matrices B and C in (18) into the square Toeplitz matrices is invalid for the coarser level, since it does not keep block-structured Toeplitz properties, see Example 1 below. Let

$$\mathscr{A}_k \begin{bmatrix} u^k \\ v^k \end{bmatrix} = \begin{bmatrix} A^{(k)}u^k + B^{(k)}v^k \\ C^{(k)}u^k + D^{(k)}v^k \end{bmatrix}.$$

In AMG, the coarse problem at the level $k < K$ is typically defined using the Galerkin approach, i.e., the coefficient matrix on the coarser grid can be computed by

$$\mathscr{A}_{k-1} = I_k^{k-1} \mathscr{A}_k I_{k-1}^k. \tag{19}$$

More concretely,

$$\mathscr{A}_{k-1} = \begin{bmatrix} A^{(k-1)} & B^{(k-1)} \\ C^{(k-1)} & D^{(k-1)} \end{bmatrix}_{(2N_{k-1}+1)\times(2N_{k-1}+1)} = I_k^{k-1} \begin{bmatrix} A^{(k)} & B^{(k)} \\ C^{(k)} & D^{(k)} \end{bmatrix}_{(2N_k+1)\times(2N_k+1)} I_{k-1}^k,$$

where $I_k^{k-1} \in \mathbb{R}^{N_k \times (2N_k+1)}$ is given in (11).

It should be noted that $A^{(k)}, B^{(k)}, C^{(k)}, D^{(k)}$ are Toeplitz matrices if $k = K$, which corresponds to the block-structured dense system (18) at the finest level with the computational count of $\mathscr{O}(N \log N)$ arithmetic operations. However, there exists a substantial difference for $\mathscr{A}_{k-1}$ at the coarser level, namely, it does not preserve block-structured Toeplitz properties, see Example 1. In fact, the resulting cross structure at the coarser level is not only dependent on the cross structure at the fine level but also dependent on the Toeplitz blocks.

Example 1 Choose the identity matrices $A^{(k)} \in \mathbb{R}^{7\times 7}$, $D^{(k)} \in \mathbb{R}^{8\times 8}$ and the rectangular matrices $B^{(k)} \in \mathbb{R}^{7\times 8}$, $C^{(k)} \in \mathbb{R}^{8\times 7}$ with all the entries equal to 1. Using the Galerkin approximation (19), we deduce

$$\left[\begin{array}{c|c} A^{(k-1)} & B^{(k-1)} \\ \hline C^{(k-1)} & D^{(k-1)} \end{array}\right] = I_k^{k-1}\mathscr{A}_k I_{k-1}^k = \frac{1}{8}\left[\begin{array}{ccc|c|ccc} 6 & 1 & 0 & 12 & 16 & 16 & 16 \\ 1 & 6 & 1 & 12 & 16 & 16 & 16 \\ 0 & 1 & 6 & 13 & 16 & 16 & 16 \\ \hline 12 & 12 & 13 & 12 & 5 & 4 & 4 \\ \hline 16 & 16 & 16 & 5 & 6 & 1 & 0 \\ 16 & 16 & 16 & 4 & 1 & 6 & 1 \\ 16 & 16 & 16 & 4 & 0 & 1 & 6 \end{array}\right],$$

and

$$\left[\begin{array}{c|c} A^{(k-2)} & B^{(k-2)} \\ \hline C^{(k-2)} & D^{(k-2)} \end{array}\right] = I_{k-1}^{k-2}\underbrace{I_k^{k-1}\mathscr{A}_k I_{k-1}^k}_{\mathscr{A}_{k-1}} I_{k-2}^{k-1} = \frac{1}{64}\left[\begin{array}{c|c|c} 44 & 170 & 256 \\ \hline 170 & 164 & 106 \\ \hline 256 & 106 & 44 \end{array}\right].$$

We next design fast Fourier transform for block-structured dense systems at the coarser level in AMG. Let

$$\mathscr{A}_k = \begin{bmatrix} A^{(k)} & \mathbf{0} & \bar{B}^{(k)} \\ \mathbf{0}' & 0 & \mathbf{0}' \\ \bar{C}^{(k)} & \mathbf{0} & \bar{D}^{(k)} \end{bmatrix}_{(2N_k+1)\times(2N_k+1)} + \begin{bmatrix} \mathbf{0} & p^{(k)} & \mathbf{0} \\ q^{(k)} & o^{(k)} & \zeta^{(k)} \\ \mathbf{0} & \xi^{(k)} & \mathbf{0} \end{bmatrix}_{(2N_k+1)\times(2N_k+1)}, \quad (20)$$

where $A^{(k)}, \bar{B}^{(k)}, \bar{C}^{(k)}, \bar{D}^{(k)}$ are the square matrices with

$$B^{(k)} = \left[\, p^{(k)} \;\; \bar{B}^{(k)} \right], \quad C^{(k)} = \begin{bmatrix} q^{(k)} \\ \bar{C}^{(k)} \end{bmatrix}, \quad D^{(k)} = \begin{bmatrix} o^{(k)} & \zeta^{(k)} \\ \xi^{(k)} & \bar{D}^{(k)} \end{bmatrix}.$$

The symbol $o^{(k)}$ is a real number, 0 denotes a zero number, and the bold $\mathbf{0}$ denotes a zero matrix/vector with the corresponding size. The coefficients $p^{(k)}, \xi^{(k)}$ denote the column vectors and $q^{(k)}, \zeta^{(k)}$ denote the row vectors. We may call it the *cross-splitting* technique, since we mainly focus on the fast Fourier transform for cross-type matrix in (20).

From (20), we find

$$\mathscr{A}_k \begin{bmatrix} v^k \\ w^k \end{bmatrix} = \left(\begin{bmatrix} A^{(k)} & \mathbf{0} & \bar{B}^{(k)} \\ \mathbf{0}' & 0 & \mathbf{0}' \\ \bar{C}^{(k)} & \mathbf{0} & \bar{D}^{(k)} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & p^{(k)} & \mathbf{0} \\ q^{(k)} & o^{(k)} & \zeta^{(k)} \\ \mathbf{0} & \xi^{(k)} & \mathbf{0} \end{bmatrix} \right) \begin{bmatrix} v^k \\ w_o^k \\ \bar{w}^k \end{bmatrix}$$
$$= \begin{bmatrix} A^{(k)} v^k + \bar{B}^{(k)} \bar{w}^k \\ 0 \\ \bar{C}^{(k)} v^k + \bar{D}^{(k)} \bar{w}^k \end{bmatrix} + \begin{bmatrix} p^{(k)} w_o^k \\ q^{(k)} v^k + o^{(k)} w_o^k + \zeta^{(k)} \bar{w}^k \\ \xi^{(k)} w_o^k \end{bmatrix}.$$

Obviously, since $A^{(k)}$, $\bar{B}^{(k)}$, $\bar{C}^{(k)}$, $\bar{D}^{(k)}$ are Toeplitz matrices, the computation of $A^{(k)} v^k$, $\bar{B}^{(k)} \bar{w}^k$, $\bar{C}^{(k)} v^k$, and $\bar{D}^{(k)} \bar{w}^k$ by FFT needs $\mathscr{O}(N_k \log N_k)$, and with required storage $\mathscr{O}(N_k)$. For the cross matrix, see $p^{(k)}$, $q^{(k)}$, $\zeta^{(k)}$, $\xi^{(k)}$ and $o^{(k)}$, we deduce $\mathscr{O}(N_k)$ complexity and storage operations.

Let the stiffness matrix of the coarser level be

$$\mathscr{A}_{k-1} = \begin{bmatrix} A^{(k-1)} & \mathbf{0} & \bar{B}^{(k-1)} \\ \mathbf{0}' & 0 & \mathbf{0}' \\ \bar{C}^{(k-1)} & \mathbf{0} & \bar{D}^{(k-1)} \end{bmatrix}_{N_k \times N_k} + \begin{bmatrix} \mathbf{0} & p^{(k-1)} & \mathbf{0} \\ q^{(k-1)} & o^{(k-1)} & \zeta^{(k-1)} \\ \mathbf{0} & \xi^{(k-1)} & \mathbf{0} \end{bmatrix}_{N_k \times N_k}. \tag{21}$$

Then using (20), (21), and a Galerkin approach in (19), it is easy to obtain the following results.

Lemma 1 *Let $A^{(k)} = \{a_{i,j}^{(k)}\}_{i,j=1}^{N_k}$ with $a_{i,j}^{(k)} = a_{j-i}^{(k)}$ be a Toeplitz matrix in (20). Then the elements of $A^{(k-1)}$ in (21) can be computed by*

$$8a_0^{(k-1)} = a_{-2}^{(k)} + 4a_{-1}^{(k)} + 6a_0^{(k)} + 4a_1^{(k)} + a_2^{(k)},$$

and

$$8a_i^{(k-1)} = a_{2i-2}^{(k)} + 4a_{2i-1}^{(k)} + 6a_{2i}^{(k)} + 4a_{2i+1}^{(k)} + a_{2i+2}^{(k)},$$
$$8a_{-i}^{(k-1)} = a_{-2i-2}^{(k)} + 4a_{-2i-1}^{(k)} + 6a_{-2i}^{(k)} + 4a_{-2i+1}^{(k)} + a_{-2i+2}^{(k)}, \quad i \geq 1.$$

Moreover, when $\bar{B}^{(k)}$, $\bar{C}^{(k)}$, $\bar{D}^{(k)}$ are similarly defined as $A^{(k)}$, assume $p^{(k)} = \{p_i^{(k)}\}_{i=1}^{N_k}$, and $\xi^{(k)}$, $q^{(k)}$ and $\zeta^{(k)}$ similarly ordered, with a given $o^{(k)}$. Then the cross matrix on the coarser level N_{k-1} can be computed as follows

$$8p_i^{(k-1)} = \left(a_{N_k-2i}^{(k)} + 2a_{N_k-2i-1}^{(k)} + a_{N_k-2i-2}^{(k)}\right) + 2\left(p_{2i-1}^{(k)} + 2p_{2i}^{(k)} + p_{2i+1}^{(k)}\right)$$
$$+ \left(b_{-2i+2}^{(k)} + 2b_{-2i+1}^{(k)} + b_{-2i}^{(k)}\right),$$
$$8q_i^{(k-1)} = \left(a_{-N_k+2i}^{(k)} + 2a_{-N_k+2i+1}^{(k)} + a_{-N_k+2i+2}^{(k)}\right) + 2\left(q_{2i-1}^{(k)} + 2q_{2i}^{(k)} + q_{2i+1}^{(k)}\right)$$
$$+ \left(c_{2i-2}^{(k)} + 2c_{2i-1}^{(k)} + c_{2i}^{(k)}\right),$$
$$8o^{(k-1)} = \left(a_0^{(k)} + 2p_{N_k-1}^{(k)} + b_{-N_k+2}^{(k)}\right) + 2\left(q_{N_k-1}^{(k)} + 2o^{(k)} + \zeta_1^{(k)}\right)$$
$$+ \left(c_{N_k-2}^{(k)} + 2\xi_1^{(k)} + d_0^{(k)}\right),$$

and

$$\begin{aligned} 8\xi_i^{(k-1)} &= \left(c_{N_k-2i}^{(k)} + 2c_{N_k-2i-1}^{(k)} + c_{N_k-2i-2}^{(k)}\right) + 2\left(\xi_{2i-1}^{(k)} + 2\xi_{2i}^{(k)} + \xi_{2i+1}^{(k)}\right) \\ &\quad + \left(d_{-2i+2}^{(k)} + 2d_{-2i+1}^{(k)} + d_{-2i}^{(k)}\right), \\ 8\zeta_i^{(k-1)} &= \left(b_{-N_k+2i}^{(k)} + 2b_{-N_k+2i+1}^{(k)} + b_{-N_k+2i+2}^{(k)}\right) + 2\left(\zeta_{2i-1}^{(k)} + 2\zeta_{2i}^{(k)} + \zeta_{2i+1}^{(k)}\right) \\ &\quad + \left(d_{2i-2}^{(k)} + 2d_{2i-1}^{(k)} + d_{2i}^{(k)}\right). \end{aligned}$$

Proof The formulas above are derived by using the Galerkin projection in (19) in order to obtain a fast computation. □

3.3 The operation count and storage requirement

We now study the computational complexity by the fast Fourier transform and the required storage for the block-structured dense system (18) in AMG, arising from the nonlocal problems in Section 2.

From (18), we know that the matrix $\mathscr{A}_h$ is a block-structured Toeplitz-like system. Then, we only need to store the first (last) column, first (last) row and principal diagonal $\mathscr{A}_h$, which have $\mathscr{O}(N)$ parameters, instead of the full matrix $\mathscr{A}_h$ with N^2 entries. From Example 1 and Lemma 1, we know that $\{A_k\}$ represents a sequence of matrices with Toeplitz-like-plus-cross structure requiring a $2^{k-K}\mathscr{O}(N)$ storage. Adding these terms together, we deduce

$$\text{Storage} = \mathscr{O}(N)\cdot\left(1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^{K-1}}\right) = \left(2-2^{1-K}\right)\mathscr{O}(N) = \mathscr{O}(N).$$

Regarding the computational complexity, the matrix–vector product associated with the matrix $\mathscr{A}_h$ can be computed by discrete convolutions, i.e., by a few FFTs. Indeed, the cost of a cross matrix–vector product is of $O(N)$ arithmetic operations by using its algebraic expression, while the cost of a dense Toeplitz matrix–vector product is of $O(N\log N)$ arithmetic operations using FFTs. Thus, the total per V-cycle AMG operation count is

$$\text{Operation count} = \mathscr{O}(N\log N)\cdot\left(1+\frac{1}{2}+\frac{1}{2^2}+\cdots+\frac{1}{2^{K-1}}\right) = \mathscr{O}(N\log N),$$

with the constant hidden in the big $\mathscr{O}$ being moderate due to the complexity of the FFT.

4 Convergence of TGM for block-structured dense system (8)

The TGM is rarely used in practice since the coarse grid operator may still be too large to be solved exactly. However, from a theoretical point of view, its study is useful as first step for evaluating the AMG convergence speed, whose analysis usually begins from that of the TGM [3, 4, 39, 41, 42, 44]. In the following we consider the convergence of the TGM for symmetric block-structured dense system (8). For any symmetric positive definite matrix $\mathscr{S}$ we can define the associated energy norm with the Euclidean inner product $||u||^2_{\mathscr{S}} = (\mathscr{S}u, u)$. Since the matrix $\mathscr{A} := \mathscr{A}_h := \mathscr{A}_K$ is symmetric positive definite, its diagonal $\mathscr{D}$ is a diagonal positive definite matrix and hence $||u||^2_{\mathscr{D}} = (\mathscr{D}u, u)$ is well defined. If we replace $\mathscr{D}$ with $\mathscr{A}\mathscr{D}^{-1}\mathscr{A}$, we observe

$$||u||^2_{\mathscr{A}\mathscr{D}^{-1}\mathscr{A}} = (\mathscr{A}u, \mathscr{A}v)_{\mathscr{D}^{-1}}.$$

First, we give some Lemmas to be used later.

Lemma 2 [53, p. 7] *Let $\mathscr{A}_K$ be a symmetric positive definite matrix. Then*

(1) the Jacobi method converges if and only if $2\mathscr{D}_K - \mathscr{A}_K$ is symmetric positive definite;

(2) the damped Jacobi method converges if and only if $0 < \omega < 2/\lambda_{\max}(\mathscr{D}_K^{-1}\mathscr{A}_K)$.

Lemma 3 [41, p. 84] *Let $\mathscr{A}_K$ be a symmetric positive definite matrix. Let $\eta \leq \omega(2 - \omega\eta_0)$ with $0 < \omega < 2/\eta_0$, $\eta_0 \geq \lambda_{\max}(\mathscr{D}_K^{-1}\mathscr{A}_K)$. Then the smoothing operator $\mathscr{S}_K$ in (16) satisfies*

$$||\mathscr{S}_K \nu^K||^2_{\mathscr{A}_K} \leq ||\nu^K||^2_{\mathscr{A}_K} - \eta||\mathscr{A}_K \nu^K||^2_{\mathscr{D}_K^{-1}}, \quad \forall \nu^K \in \mathfrak{M}^K \tag{22}$$

with $\mathfrak{M}^K$ in (14).

Lemma 4 [41, p. 89] *Let $\mathscr{A}_K$ be a symmetric positive definite matrix and smoothing operator $\mathscr{S}_K$ in (16) satisfies (22) and*

$$\min_{\nu^{K-1} \in \mathfrak{M}^{K-1}} ||\nu^K - I_{K-1}^K \nu^{K-1}||^2_{\mathscr{D}_K} \leq \mu||\nu^K||^2_{\mathscr{A}_K}, \quad \forall \nu^K \in \mathfrak{M}^K \tag{23}$$

with $\mu > 0$ independent of ν^K. Then, $\mu \geq \eta > 0$ and the convergence factor of TGM satisfies

$$||\mathscr{S}_K \mathscr{T}_K||_{\mathscr{A}_K} \leq \sqrt{1 - \eta/\mu}, \quad \forall \nu^K \in \mathfrak{M}^K,$$

where $\mathscr{T}_K$ is the coarse grid correction operator defined in (15).

Next we need to check the smoothing condition (22) and approximation property (23), respectively. We use the notion of weakly and strictly diagonal dominant matrix. A matrix is weakly diagonal dominant if the modulus of any diagonal element of the considered matrix is at least as large as the sum of the absolute value of off-diagonal elements in the same row or column and at least one diagonal element has modulus strictly larger. Along the same lines, a matrix is strictly diagonal dominant if the modulus of any diagonal element of the considered matrix is larger than the sum of the absolute value of off-diagonal elements in the same row or column.

Lemma 5 [51, p. 23] *Let $\mathscr{A}_K \in \mathbb{R}^{N\times N}$ be a symmetric matrix. If $\mathscr{A}_K$ is a strictly diagonally dominant or irreducibly weakly diagonally dominant matrix with positive real diagonal entries, then $\mathscr{A}_K$ is positive definite.*

Lemma 6 *Let $\mathscr{A}_K := \mathscr{A}_h^S$ be defined by (8) and (9). Then $\mathscr{A}_K$ is a weakly diagonally dominant symmetric matrix with positive entries on the diagonal and nonpositive off-diagonal entries.*

Proof From (9), we have

$$a_0 > 0, \quad a_m < 0, \quad 1 \le m \le r \quad \text{and} \quad a_{m+\frac{1}{2}} < 0, \quad 0 \le m \le r-1,$$

and

$$a_0 + 2\sum_{m=1}^{r} a_m + 2\sum_{m=0}^{r-1} a_{m+\frac{1}{2}} = 0. \tag{24}$$

The proof is completed. □

We now prove the positive definiteness of the matrix $\mathscr{A}_K$.

Lemma 7 *Let $\mathscr{A}_K := \mathscr{A}_h^S$ be defined by (8). Then $\mathscr{A}_K$ is a symmetric positive definite matrix with positive entries on the diagonal and nonpositive off-diagonal entries.*

Proof Let $L_N = \text{tridiag}(-1, 2, -1) \in \mathbb{R}^{N\times N}$ be the discrete Laplacian operator. Let

$$\mathscr{A}_K = \mathscr{A}_{\text{res}} - a_1\mathscr{A}_{\text{main}} \;\text{ with }\; \mathscr{A}_{\text{main}} = \begin{bmatrix} L_{N-1} & O \\ O & L_N \end{bmatrix}. \tag{25}$$

We can check the matrix $-a_1\mathscr{A}_{\text{main}}$ with $a_1 < 0$ is positive definite. We next prove $\mathscr{A}_{\text{res}}$ is a semi-positive definite matrix. From (25), we observe $\mathscr{A}_{\text{res}} = \mathscr{A}_K + a_1\mathscr{A}_{\text{main}}$, which implies that the principal diagonal elements are positive and the off-diagonal are non-positive of the matrix $\mathscr{A}_{\text{res}}$. Using Lemma 6, we know $\mathscr{A}_{\text{res}}$ is a weakly diagonally dominant symmetric matrix. Thus, $\mathscr{A}_{\text{res}}$ is a semi-positive definite matrix by the Geršgorin disc theorem [34, p. 388]. The proof is completed. □

Remark 1 Regarding Lemma 7, an alternative proof of the positive definite character of the matrix $\mathscr{A}_K$ can rely completely on Lemma 5. The positivity of the diagonal entries and $\mathscr{A}_K$ and its weak diagonal dominance are established in Lemma 6.

First we observe that a matrix of the form

$$\begin{bmatrix} X & Y \\ Z & V \end{bmatrix}$$

with square diagonal blocks X, V is not necessarily irreducible, even in presence of irreducible diagonal blocks X, V. Indeed the latter is obvious if one takes either Y or Z equal to the null block. In our setting, i.e. (8), we see that $X = A$ and $V = \widehat{A}$ are irreducible since their tridiagonal parts are irreducible (as in can be plainly checked looking at the coefficients in (9)).

For checking the irreducible character of $\mathscr{A}_K$, we consider its directed graph starting from node 1 in $\mathscr{G}(A)$, $\mathscr{G}(A)$ being the direct graph of A, and visiting all nodes of $\mathscr{G}(A)$ with repetitions and ending in node 2. From node 2 we can jump to node $N+1$ thanks to the fact that all diagonal entries of the rectangular matrix $Y = B$ are equal to -4, as in can be checked at the end of Subsection 2.1. Node $N+1$ refers to the block $V = \widehat{A}$ which is irreducible and hence from it we can visit with repetitions all nodes of $\mathscr{G}(\widehat{A})$, $\mathscr{G}(\widehat{A})$ being the direct graph of $\widehat{A}$, stopping at node N. From node N we can jump to node 1 owing to the relation $Z = B^T$ so that all the diagonal entries of B^T are again equal to -4.

Hence the directed graph associated to the matrix $\mathscr{A}_K$ is strongly connected that is $\mathscr{A}_K$ is irreducible. In conclusion the matrix $\mathscr{A}_K$ is symmetric positive definite thanks to Lemma 5.

Lemma 8 *Let the discrete Laplacian-like operators* $\{L_j^M\}_{j=1}^{M-1} \in \mathbb{R}^{M\times M}$ *and discrete block-structured Laplacian operators* $\mathscr{L}_j$ *be, respectively, defined by*

$$L_j^M = \begin{bmatrix} 2 & \overbrace{\cdots}^{j-1\,zeros} & -1 & & \\ \vdots & \ddots & \ddots & \ddots & \\ -1 & \ddots & \ddots & \ddots & -1 \\ & \ddots & \ddots & \ddots & \vdots \\ & & -1 & \cdots & 2 \end{bmatrix}_{M\times M} \quad \text{and} \quad \mathscr{L}_j = \begin{bmatrix} L_j^M & \mathbf{0} \\ \mathbf{0} & L_j^N \end{bmatrix}.$$

Assume that the integer parameter l *belongs to the interval* $[1, M-1]$*. Then the matrices*

$$\left(\frac{2}{l}\sum_{j=1}^{l} L_j^M\right) - L_1^M, \quad \left(\frac{2}{l}\sum_{j=1}^{l} \mathscr{L}_j\right) - \mathscr{L}_1 \ \text{and}\ 2\mathscr{L}_1 - L_1^{M+N} \tag{26}$$

are all positive definite.

Proof The first results of this lemma can be seen in Lemma 3.10 of [22], which implies that the second one is also satisfied, since

$$\begin{aligned}\left(\frac{2}{l}\sum_{j=1}^{l} \mathscr{L}_j\right) - \mathscr{L}_1 &= \frac{2}{l}\sum_{j=1}^{l} \begin{bmatrix} L_j^M & \mathbf{0} \\ \mathbf{0} & L_j^N \end{bmatrix} - \begin{bmatrix} L_1^M & \mathbf{0} \\ \mathbf{0} & L_1^N \end{bmatrix} \\ &= \begin{bmatrix} \left(\frac{2}{l}\sum_{j=1}^{l} L_j^M\right) - L_1^M & \mathbf{0} \\ \mathbf{0} & \left(\frac{2}{l}\sum_{j=1}^{l} L_j^N\right) - L_1^N \end{bmatrix}.\end{aligned}$$

On the other hand, we can check that $2\mathscr{L}_1 - L_1^{M+N}$ is an irreducible and weakly diagonally dominant symmetric matrix, which means that it is positive definite by Lemma 5. The proof is completed. □

Remark 2 In order to understand the spectral features of the matrices considered in Lemma 8, we can adopt the analysis via the related generating functions, since all the matrices in Lemma 8 are of real symmetric banded Toeplitz type, so they admit real-valued trigonometric polynomials as generating functions (see [44] and references therein): on the other hand the matrices $\mathscr{L}_j$ are block diagonal and hence their spectral analysis reduces to the Toeplitz setting. For instance, according to the notation in [44] concisely recalled in Sect. 2.1, $L_j^M = T_M(2-2\cos(j\theta))$ that is the function $2-2\cos(j\theta)$ is the generating function of L_j^M. From classical results, we know that $T_M(f)$ is positive definite for any matrix-size M if f is essentially bounded and nonnegative, with positive essential supremum. In the present setting, the maximum of $2-2\cos(j\theta)$ is 4 and its minimum is zero and hence $L_j^M = T_M(2-2\cos(j\theta))$ is positive definite. Not only this: if the nonnegative generating function f has a unique zero of order $\alpha > 0$ then the minimal eigenvalue of $T_M(f)$ is positive and, for $M \to \infty$, converges monotonically to zero as c/M^α with c depending on the second derivative of f at the zero if it is positive. Based on these results, we can deduce that L_j^M is positive definite, has minimal eigenvalue positive converging monotonically to zero as c_j/M^2 with positive c_j independent of M, and c_j related to the second derivative of $2-2\cos(j\theta)$ at $\theta = 0$.

Of course, by linearity, the Toeplitz matrix, i.e., the first one in (26), has generating function given by $f_l(\theta) \equiv \left(\frac{2}{l}\sum_{j=1}^{l}[2-2\cos(j\theta)]\right) - (2-2\cos(\theta))$ and hence by studying this generating function, we deduce that this Toeplitz matrix is positive definite, has minimal eigenvalue positive converging monotonically to zero as c/M^2 with positive c independent of M. Hence its condition number grow exactly as M^2 and since the related generating function has a unique zero of order 2 at $\theta = 0$, there is a formal justification in using standard projectors and restriction operators, like those employed in the standard AMG, for the classical discrete Laplacian (see [3, 32, 44] and references therein).

Lemma 9 *Let $\mathscr{A}_K := \mathscr{A}_h^S$ be defined by (8). Then the damped Jacobi iteration convergences with relaxation parameter $0 < \omega \le 1$, and the smoothing operator $\mathscr{S}_K$ in (16) satisfies*

$$||\mathscr{S}_K \nu^K||_{\mathscr{A}_K}^2 \le ||\nu^K||_{\mathscr{A}_K}^2 - \frac{1}{2}||\mathscr{A}_K \nu^K||_{\mathscr{D}_K^{-1}}^2, \quad \forall \nu^K \in \mathfrak{M}^K.$$

Proof According to (8) and (9), using Lemma 6, we know that the matrix $\mathscr{A}_K$ is weakly diagonal dominant, since

$$a_0 + 2\sum_{m=1}^{r} a_m + 2\sum_{m=0}^{r-1} a_{m+\frac{1}{2}} = 0.$$

Ordering the entries in $\mathscr{A}_K$ and taking $\mathscr{A}_K = \left\{a_{i,j}^{(K)}\right\}_{i,j=1}^{2N-1}$, we find

$$r_i^{(K)} := \sum_{j\neq i} |a_{i,j}^{(K)}| \leq a_{i,i}^{(K)} = a_0.$$

Using the Geršgorin disc theorem [34, p. 388], the eigenvalues of $\mathscr{A}_K$ belong to the union of the disks centered at $a_{i,i}^{(K)}$ with radius $r_i^{(K)}$. Namely, the eigenvalues λ of the matrix $\mathscr{A}_K$ satisfy

$$|\lambda - a_{i,i}^{(K)}| \leq r_i^{(K)},$$

which leads to $\lambda_{\max}(\mathscr{A}_K) \leq a_{i,i}^{(K)} + r_i^{(K)} < 2a_{i,i}^{(K)} = 2a_{1,1}^{(K)} = 2a_0$.

From Rayleigh theorem [34, p. 235], we deduce

$$\lambda_{\max}(\mathscr{A}_K) = \max_{x\neq 0} \frac{x^T \mathscr{A}_K x}{x^T x}, \quad \forall x \in \mathbb{R}^n.$$

Take $x = [1, 0, \ldots, 0]^T$. Then

$$\lambda_{\max}(\mathscr{A}_K) \geq \frac{x^T \mathscr{A}_K x}{x^T x} = a_{1,1}^{(K)} = a_0,$$

and

$$\lambda_{\max}\left((\mathscr{D}_K)^{-1}\mathscr{A}_K\right) = \frac{\lambda_{\max}(\mathscr{A}_K)}{a_{1,1}^{(K)}} = \frac{\lambda_{\max}(\mathscr{A}_K)}{a_0} \in [1, 2],$$

where $\mathscr{D}_K$ is the diagonal part of $\mathscr{A}_K$.

From Lemma 2, the damped Jacobi method converges with $0 < \omega < 1$. By following either a similar proof as in Lemma 7 or using the arguments in Remark 1, we can check that $2\mathscr{D}_K - \mathscr{A}_K$ is symmetric positive definite. Then using Lemma 2 again, the Jacobi method converges. Hence, the damped Jacobi iteration with relaxation parameter $0 < \omega \leq 1$ converges. As a consequence, the desired results are obtained by employing Lemma 3. □

Lemma 10 *Let $\mathscr{A}_K = \mathscr{A}_h^S$ be defined by (8). Then*

$$\min_{\nu^{K-1}\in\mathfrak{M}^{K-1}} ||\nu^K - I_{K-1}^K \nu^{K-1}||_{\mathscr{D}_K}^2 \leq 24||\nu^K||_{\mathscr{A}_K}^2, \quad \forall \nu^K \in \mathfrak{M}^K.$$

Proof Let the discrete block-structured Laplacian-like operators be defined by

$$\mathscr{L}_j = \begin{bmatrix} L_j^{N-1} & \mathbf{0} \\ \mathbf{0} & L_j^N \end{bmatrix} = \begin{bmatrix} L_j^{N-1} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix} + \begin{bmatrix} \mathbf{0} & \mathbf{0} \\ \mathbf{0} & L_j^N \end{bmatrix}, \quad j = 1, 2, \ldots, r,$$

where the discrete Laplacian-like operators $\{L_j^{N-1}\}_{j=1}^r \in \mathbb{R}^{(N-1)\times(N-1)}$, $\{L_j^N\}_{j=1}^r \in \mathbb{R}^{N\times N}$ are given in Lemma 8.

From (8), the block-structured dense matrix $\mathscr{A}_K$ can be denoted as the series sum with Laplacian-like operators $\mathscr{L}_j$, i.e.,

$$\mathscr{A}_K = \begin{bmatrix} A & B \\ B^T & \widehat{A} \end{bmatrix} = -\sum_{j=1}^r a_j \mathscr{L}_j + \mathscr{B} \text{ and } \mathscr{B} = \left(a_0 + 2\sum_{j=1}^r a_j\right) I + \begin{bmatrix} \mathbf{0} & B \\ B^T & \mathbf{0} \end{bmatrix} \tag{27}$$

with I an identity matrix.

Using (9), we obtain $a_0 + 2\sum_{m=1}^r a_m + 2\sum_{m=0}^{r-1} a_{m+\frac{1}{2}} = 0$, which implies that $\mathscr{B}$ is a weakly diagonally dominant symmetric matrix with positive entries on the diagonal part and nonpositive off-diagonal entries. As a consequence, by applying the Geršgorin disc theorem, we deduce that the matrix $\mathscr{B}$ is symmetric and semi-positive definite.

From (27), (9) and Lemma 8, we obtain

$$\begin{aligned} \left\|\nu^K\right\|_{\mathscr{A}_K}^2 &= \left(\mathscr{A}_K \nu^K, \nu^K\right) \geq -\frac{a_1}{2}\left(\sum_{j=1}^r \mathscr{L}_j \nu^K, \nu^K\right) \geq -\frac{r a_1}{4}\left(\mathscr{L}_1 \nu^K, \nu^K\right) \\ &\geq -\frac{r a_1}{8}\left(L_1^{2N-1}\nu^K, \nu^K\right) \geq \frac{a_0}{48}\left(L_1^{2N-1}\nu^K, \nu^K\right) \\ &\geq \frac{a_0}{24}\left\|\nu^K - I_{K-1}^K \nu^K\right\|^2 = \frac{1}{24}\left\|\nu^K - I_{K-1}^K \nu^J\right\|_{\mathscr{D}_K}^2 . \end{aligned}$$

The proof is completed. □

Theorem 1 *Let $\mathscr{A}_K = \mathscr{A}_h^S$ be defined by (8). Then the convergence factor of the TGM satisfies*

$$||\mathscr{S}_K \mathscr{T}_K||_{\mathscr{A}_K} \leq \sqrt{47/48} < 1.$$

Proof By combining the results in Lemmas 3, 4, 9 and 10, the desired result follows. □

Remark 3 The problem treated in the present work is one-dimensional. when pure Toeplitz structures appear or even GLT matrix-sequences are considered the d dimensional AMG design is somehow guided by the symbol and and hence d dimensional problems are not a problem with $d \geq 2$; see [10, 11, 27] and references therein. Here the mixture of block and cross structures is more delicate and more work is required for a d-dimensional generalization.

5 Numerical results

We employ the V-cycle block-structured AMG described in Algorithm 1 to solve the time-dependent nonlocal problems in Section 2. The stopping criterion is taken as

$$\frac{||r^{(i)}||}{||r^0||} < 10^{-15},$$

where $r^{(i)}$ is the residual vector after i iterations, the number of pre-smoothing step $m_1 = 1$ and post-smoothing step $m_2 = 1$, and the weighted Jacobi relaxation $(w_{pre}, w_{post}) = (1, 1/2)$. In all tables, N denotes the number of spatial grid points, and the numerical errors ("Error") are measured by the l_∞ (maximum) vector norm, which is computed exactly since we know the analytic solution in our example, "Rate" denotes the convergence order, i.e.,

$$\text{Rate} = \frac{\ln\left(||U_{2h}^{N_t} - U_h^{N_t}||_\infty / ||U_h^{N_t} - U_{h/2}^{N_t}||_\infty\right)}{\ln 2}.$$

"CPU" denotes the total CPU time in seconds (s) for solving the discretized systems, "Iter" denotes the average number of iterations required to solve algebraic systems at each time step. Here, u_i^n denotes the approximated value of $u(x_i, t_n)$ and $f_i^n = f(x_i, t_n)$ with the mesh points $0 = t_0 < t_1 < \cdots < t_{N_t} = T$ and $\tau = T/N_t$.

All numerical experiments are programmed in Matlab, and the computations are carried out on a desktop computer with the configuration: Intel(R) Core(TM) i7-7700 3.60 GHZ and 8 GB RAM and a 64 bit Windows 10 operating system.

First we consider the time-dependent nonlocal models (2) in Section 2. The initial value and the forcing term are chosen such that the exact solution of the considered equations is

$$u(x, t) = e^t\,(1 + x)^6, \quad 0 \leqslant x \leqslant 1, \quad 0 \leqslant t \leqslant 1.$$

We apply the following BDF4 method to such nonlocal models

$$\left(\frac{25}{12} I + \tau \mathscr{A}\right) U^n = 4U^{n-1} - 3U^{n-2} + \frac{4}{3} U^{n-3} - \frac{1}{4} U^{n-4} + \tau F^n, \quad n = 4, 5, \ldots, N_t. \tag{28}$$

Here the operator $\mathscr{A}$ denotes the block-structured systems (7) and (8), respectively. It should be noted that the convergence analysis of time-dependent nonlocal model (2) with $\mathscr{O}\left(\tau^4 + h^{\max\{2,4-2\beta\}}\right)$, $\delta = \mathscr{O}\left(h^\beta\right)$, $\beta \geq 0$ can be found in [1, 16]. The convergence rate of the two-grid method for time-dependent block-structured systems (8) can be directly obtained by Theorem 1 and [25].

Remark 4 Regarding the approximation error we observe that the formula

$$\mathscr{O}\left(\tau^4 + h^{\max\{2,4-2\beta\}}\right), \quad \delta = \mathscr{O}\left(h^\beta\right), \quad \beta \geq 0,$$

implies that the optimal relation between τ and h is of the form $\tau = \sqrt{h}$ for $\beta \geq 1$, while for $\beta \in [0, 1)$ the best choice is $\tau = h^{1-\beta/2}$. Hence when selecting a constant δ we would have $\tau = h$, while for $\delta = \sqrt{h}$ we would have $\tau = h^{3/4}$. In the following numerical experiments, for $\delta = 1/4$ ($\beta = 0$) and $\delta = \sqrt{h}$ ($\beta = 1/2$), we have uniformly chosen $\tau = h$, since it is optimal in the first case and not far from the optimal choice even in the case $\delta = \sqrt{h}$.

Remark 5 Regarding the linear systems in (28), with τ going to zero as the matrix-size tends to infinity, we deduce that the related matrix-sequence has spectrum clustered at $25/12$ and this makes the related linear systems not difficult to solve.

We further extend the V-cycle block-structured algebraic multigrid Algorithm 1 to simulate the nonlocal models with nonsymmetric indefinite block-structured dense systems and symmetric positive definite block-structured dense systems, respectively.

5.1 Nonsymmetric indefinite block-structured dense systems

Table 1 shows that fast algebraic multigrid (FAMG) proposed for the BDF4 scheme (28) stemming from time-dependent nonlocal model (2) with nonsymmetric indefinite block-structured dense systems is efficient and robust. Indeed, the proposed method requires a computational cost of $\mathcal{O}(N\log N)$ arithmetic operations. This can be seen by combining the $\mathcal{O}(N\log N)$ cost per iteration proven in Sect. 3.3 and a number of iteration which is uniformly bounded. In fact the number of iterations decreases mildly and this agree with Remark 5. We also observe a favorable CPU time following a almost linear growth with the matrix-size. We also present for comparison the performances of the fast conjugate gradient least square method (FCGS) for nonsymmetric indefinite systems (7), the result shows that the convergence rate is the same as that simulated by FAMG. However for $\delta = \sqrt{h}$, the CPU time is not competitive- Furthermore, the optimality is lost because the number of iterations grows with the matrix-size. Hence the related overall cost is not longer of $\mathcal{O}(N\log N)$ arithmetic operations, at least for the FCGS method.

Finally in order to check the robustness of our technique and taking into account Remark 5, we consider a variation of the linear systems in (28) where the constant $25/12$ is replaced with $\alpha = 1, 1/10, 1/100$. In fact, the smaller $\alpha > 0$ is, the more the system is ill-conditioned. With this setting of parameters the algorithm FAMG is robust while we observe a certain increase in the iteration count for FCGS (Table 2).

5.2 Symmetric positive definite block-structured dense systems

Table 3 shows that the proposed FAMG for solving the BDF4 scheme (28) with symmetric block-structured dense systems is robust, which implies a $\mathcal{O}(N\log N)$ complexity and a very good CPU time. As comparative tests, Table 3 presents the results of solving symmetric indefinite systems (8) by means of FCGS. Again we observe that the CPU timing deteriorates with the matrix-size, since the method is not optimal and the iteration number grows when increasing the dimension and when $\delta = \sqrt{h}$, at least for the FCGS method.

Table 1 Nonsymmetric indefinite block-structured dense systems (7): FAMG versus FCGS to solve the time-dependent nonlocal model (2) with $\tau = h = 1/N$ and zero initial guess

	N	$\delta = 1/4$				$\delta = \sqrt{h}$			
		Error	Rate	CPU	Iter	Error	Rate	CPU	Iter
$FAMG$	2^5	4.3254e−05		0.1460s	7	9.9042e−05		0.2276s	10
	2^6	2.7166e−06	3.9930	0.3628s	6	9.3791e−06	3.4005	0.4737s	8
	2^7	1.7022e−07	3.9963	0.6684s	5	1.1886e−06	2.9802	1.5029s	8
	2^8	1.0652e−08	3.9981	2.0338s	4	1.3694e−07	3.1177	3.5816s	7
$FCGS$	2^5	4.3254e−05		0.1191s	10	9.9042e−05		0.0741s	14
	2^6	2.7166e−06	3.9930	0.1475s	9	9.3791e−06	3.4005	0.1185s	15
	2^7	1.7022e−07	3.9963	0.3306s	8	1.1886e−06	2.9802	0.5467s	18
	2^8	1.0652e−08	3.9981	0.5336s	7	1.3694e−07	3.1177	1.5552s	22

Table 2 Nonsymmetric indefinite block-structured dense systems performed as (28): here 25/12 is replaced with $\alpha = 1, 1/10, 1/100$, respectively, $\delta = 1/4$ and stopping criterion is taken as 10^{-6}; for FAMG and FCGS we have $\tau = h = 1/N$

	N	$\alpha = 1$		$\alpha = 1/10$		$\alpha = 1/100$	
		CPU	Iter	CPU	Iter	CPU	Iter
$FAMG$	2^5	0.1332s	3	0.1928s	5	0.2028s	5
	2^6	0.3307s	3	0.4293s	4	0.5313s	5
	2^7	0.6399s	2	0.8627s	3	1.4893s	5
$FCGS$	2^5	0.0767s	8	0.1598s	18	0.1612s	16
	2^6	0.2348s	9	0.4054s	17	0.3670s	20
	2^7	0.7737s	15	2.5597s	52	2.5330s	51

6 Conclusions

In this paper, we considered the solutions of block-structured dense and Toeplitz-like-plus-cross systems arising from nonlocal diffusion problem. We designed an AMG for block-structured dense and Toeplitz-like-plus-cross systems, by making also use of fast Fourier transform, and we provided an estimate of the TGM convergence rate for the nonlocal problem with symmetric positive definite block-structured dense linear systems. In this specific context, we answered the question on how to define coarsening and interpolation operators, when the stiffness matrix leads to nonsymmetric systems [14, 36]. The simple (traditional) restriction operator and prolongation operator are employed for such Toeplitz-like-plus-cross systems, so that the entries of the sequence of subsystems are explicitly determined on different levels.

For the future, at least three questions arise and we plan to investigate them. More precisely:

- since the structures arising from the same type of problems but in d dimensions, $d \geq 2$, are definitely more involved due to the simultaneous presence of tensor and cross operations, the related design of efficient AMG solvers is not trivial and it represents a subject to be investigated;
- we would like to consider the study of the TGM convergence analysis for nonsymmetric block-structured dense systems and the analysis of the full AMG for symmetric block-structured dense systems, based on the ideas presented in [20, 22];
- we plan a more complete comparison with preconditioned Krylov solvers designed for block problems [7, 8] and indefinite structured problems [43], taking into account the theoretical barriers in the multilevel setting [45, 46].

Table 3 Symmetric positive definite block-structured dense systems (8): FAMG versus FCGS to solve the time-dependent nonlocal model (2) with $\tau = h = 1/N$ and zero initial guess

	N	$\delta = 1/4$				$\delta = \sqrt{h}$			
		Error	Rate	CPU	Iter	Error	Rate	CPU	Iter
$FAMG$	2^5	1.1628e−05		0.2341s	9	2.5257e−05		0.2987s	12
	2^6	7.3840e−07	3.9771	0.4819s	7	2.3810e−06	3.4070	0.5999s	10
	2^7	4.6514e−08	3.9887	1.0638s	6	2.9930e−07	2.9919	2.0004s	10
	2^8	2.9182e−09	3.9945	2.6782s	5	3.4366e−08	3.1226	4.6659s	9
$FCGS$	2^5	1.1628e−05		0.0552s	11	2.5257e−05		0.0776s	15
	2^6	7.3840e−07	3.9771	0.1206s	10	2.3810e−06	3.4070	0.1983s	16
	2^7	4.6514e−08	3.9887	0.3567s	8	2.9930e−07	2.9919	0.6076s	19
	2^8	2.9182e−09	3.9945	0.5302s	7	3.4366e−08	3.1226	1.6952s	23

Acknowledgements The research of S. Serra-Capizzano was partly supported by the Italian INdAM GNCS agency and by the Theory, Economics, and Systems Laboratory of the Department of Informatics, AUEB (Athens - Greece). Furthermore, the work of Stefano Serra-Capizzano was funded from the European High-Performance Computing Joint Undertaking (JU) under grant agreement No 955701. The JU receives support from the European Union's Horizon 2020 research and innovation programme and Belgium, France, Germany, Switzerland. We warmly thank the Reviewer for pertinent and constructive suggestions and comments.

Funding Open access funding provided by Uppsala University.

References

1. Akrivis, G., Chen, M.H., Yu, F., Zhou, Z.: The energy technique for the six-step BDF method. SIAM J. Numer. Anal. **59**, 2449–2472 (2021)
2. Andreu-Vaillo, F., Mazón, J.M., Rossi, J.D., Toledo-Melero, J.J.: Nonlocal Diffusion Problems. Mathematical Surveys and Monographs, vol. 165. AMS, Providence (2010)
3. Aricò, A., Donatelli, M.: A V-cycle multigrid for multilevel matrix algebras: proof of optimality. Numer. Math. **105**, 511–547 (2007)
4. Aricò, A., Donatelli, M., Serra-Capizzano, S.: V-cycle optimal convergence for certain (multilevel) structured linear systems. SIAM J. Matrix Anal. Appl. **26**, 186–214 (2004)
5. Atkinson, K.E.: The Numerical Solution of Integral Equations of the Second Kind. Cambridge University Press, New York (2009)
6. Bates, P.: On some nonlocal evolution equations arising in materials science. In: Brunner, H., Zhao, X., Zou, X. (eds.) Nonlinear Dynamics and Evolution Equations. Fields Institute Communications, pp. 13–52. Providence, AMS (2006)
7. Benzi, M.: Preconditioning techniques for large linear systems: a survey. J. Comput. Phys. **182–2**, 418–477 (2002)
8. Benzi, M., Golub, G.H., Liesen, J.: Numerical solution of saddle point problems. Acta Numer. **14**, 1–137 (2005)
9. Bini, D.A., Pan, V.Y.: Polynomial and Matrix Computations. Fundamental Algorithms, vol. I. Birkhäuser Boston Inc, Boston (1994)
10. Bolten, M., Donatelli, M., Ferrari, P., Furci, I.: A symbol-based analysis for multigrid methods for block-circulant and block-Toeplitz systems. SIAM J. Matrix Anal. Appl. **43–1**, 405–438 (2022)
11. Bolten, M., Donatelli, M., Ferrari, P., Furci, I.: Symbol based convergence analysis in block multigrid methods with applications for Stokes problems. Appl. Numer. Math. **193**, 109–130 (2023)
12. Bolten, M., Donatelli, M., Huckle, T., Kravvaritis, C.: Generalized grid transfer operators for multigrid methods applied on Toeplitz matrices. BIT **55**, 341–366 (2015)
13. Böttcher, A., Grudsky, S.M.: Spectral Properties of Banded Toeplitz Matrices. SIAM, Philadelphia (2005)
14. Brezina, M., Manteuffel, T., Mccormick, S., Ruge, J., Sanders, G.: Towards adaptive smoothed aggregation (αSA) for nonsymmetric problems. SIAM J. Sci. Comput. **32**, 14–39 (2010)
15. Cao, R.J., Chen, M.H., Ng, M.K., Wu, Y.J.: Fast and high-order accuracy numerical methods for time-dependent nonlocal problems in $\mathbb{R}^2$. J. Sci. Comput. **84**, Paper No. 8 (2020)
16. Cao, R.J., Chen, M.H., Qi, Y.F., Shi, J.K., Yin, X.B.: Analysis of (shifted) piecewise quadratic polynomial collocation for nonlocal diffusion model. Appl. Numer. Math. **185**, 120–140 (2023)
17. Chan, R.H., Jin, X.Q.: An Introduction to Interative Toeplitz Solvers. SIAM, Phildelphia (2007)

18. Chan, R.H., Ng, M.K.: Conjugate gradient methods for Toeplitz systems. SIAM Rev. **38**, 427–482 (1996)
19. Chan, R.H., Chang, Q.S., Sun, H.W.: Multigrid method for ill-conditioned symmetric Toeplitz systems. SIAM J. Sci. Comput. **19**, 516–529 (1998)
20. Chen, M.H., Ekström, S.E., Serra-Capizzano, S.: A Multigrid method for nonlocal problems: non-diagonally dominant or Toeplitz-plus-tridiagonal systems. SIAM J. Matrix Anal. Appl. **41**, 1546–1570 (2020)
21. Chen, M.H., Deng, W.H.: Fourth order accurate scheme for the space fractional diffusion equations. SIAM J. Numer. Anal. **52**, 1418–1438 (2014)
22. Chen, M.H., Deng, W.H.: Convergence analysis of a multigrid method for a nonlocal model. SIAM J. Matrix Anal. Appl. **38**, 869–890 (2017)
23. Chen, M.H., Deng, W.H., Serra-Capizzano, S.: Uniform convergence of V-cycle multigrid algorithms for two-dimensional fractional Feynman-Kac equation. J. Sci. Comput. **74**, 1034–1059 (2018)
24. Chen, M.H., Qi, W.Y., Shi, J.K., Wu, J.M.: A sharp error estimate of piecewise polynomial collocation for nonlocal problems with weakly singular kernels. IMA J. Numer. Anal. **41**, 3145–3174 (2021)
25. Chen, M.H., Wang, Y.T., Cheng, X., Deng, W.H.: Second-order LOD multigrid method for multidimensional Riesz fractional diffusion equation. BIT **54**, 623–647 (2014)
26. D'Elia, M., Du, Q., Glusa, C., Gunzburger, M., Tian, X.C., Zhou, Z.: Numerical methods for nonlocal and fractional models. Acta Numer. **29**, 1–124 (2020)
27. Donatelli, M., Ferrari, P., Furci, I., Serra-Capizzano, S., Sesana, D.: Multigrid methods for block-Toeplitz linear systems: convergence analysis and applications. Numer. Linear Algebra. Appl. **28**, e2356 (2021)
28. Donatelli, M., Mazza, M., Serra-Capizzano, S.: Spectral analysis and multigrid methods for finite volume approximations of space-fractional diffusion equations. SIAM J. Sci. Comput. **40**, A4007–A4039 (2018)
29. Du, Q.: Nonlocal Modeling, Analysis, and Computation. SIAM, Philadelphia (2019)
30. Du, Q., Gunzburger, M., Lehoucq, R., Zhou, K.: Analysis and approximation of nonlocal diffusion problems with volume constraints. SIAM Rev. **56**, 676–696 (2012)
31. Dumbser, M., Fambri, F., Furci, I., Mazza, M., Serra-Capizzano, S., Tavelli, M.: Staggered discontinuous Galerkin methods for the incompressible Navier–Stokes equations: spectral analysis and computational results. Numer. Linear Algebra Appl. **25–5**, e2151 (2018)
32. Fiorentino, G., Serra, S.: Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions. SIAM J. Sci. Comput. **17**, 1068–1081 (1996)
33. Garoni, C., Serra-Capizzano, S.: Generalized Locally Toeplitz Sequences: Theory and Applications, vol. I. Springer, Cham (2017)
34. Horn, R.A., Johnson, C.R.: Matrix Analysis. Cambridge University Press, New York (2013)
35. Leng, Y., Tian, X., Trask, N., Foster, J.T.: Asymptotically compatible reproducing kernel collocation and meshfree integration for nonlocal diffusion. SIAM J. Numer. Anal. **59**, 88–118 (2021)
36. Manteuffel, T., Southworth, S.: Convergence in norm of nonsymmetric algebraic multigrid. SIAM J. Sci. Comput. **41**, S269–S296 (2019)
37. Ng, M.K., Pan, J.Y.: Weighted Toeplit regularized least squares computation for image restoration. SIAM J. Sci. Comput. **36**, B94–B121 (2014)
38. Notay, Y.: A new algebraic multigrid approach for Stokes problems. Numer. Math. **132**, 51–84 (2016)
39. Pang, H., Sun, H.: Multigrid method for fractional diffusion equations. J. Comput. Phys. **231**, 693–703 (2012)
40. Qi, W.Y., Seshaiyer, P., Wang, J.P.: Finite element method with the total stress variable for Biot's consolidation model. Numer. Methods Partial Differ. Equ. **37**, 2409–2428 (2021)
41. Ruge, J., Stüben, K.: Algebraic multigrid. In: McCormick, S. (ed.) Multigrid Methods, pp. 73–130. SIAM, Philadelphia (1987)
42. Saad, Y.: Iterative Methods for Sparse Linear Systems. SIAM, Philadelphia (2003)
43. Serra, S.: Preconditioning strategies for Hermitian Toeplitz systems with nondefinite generating functions. SIAM J. Matrix Anal. Appl. **17**, 1007–1019 (1996)
44. Serra-Capizzano, S.: Convergence analysis of two-grid methods for elliptic Toeplitz and PDEs matrix-sequences. Numer. Math. **92**, 433–465 (2002)
45. Serra-Capizzano: Matrix algebra preconditioners for multilevel Toeplitz matrices are not superlinear. Linear Algebra Appl. **343**, 303–319 (2002)

46. Serra-Capizzano, S., Tyrtyshnikov, E.: How to prove that a preconditioner cannot be superlinear. Math. Comput. **72–243**, 1305–1316 (2003)
47. Silling, S.A.: Reformulation of elasticity theory for discontinuities and long-range forces. J. Mech. Phys. Solids **48**, 175–209 (2000)
48. Sivas, A.A., Southworth, B.S., Rhebergen, S.: Air algebraic multigrid for a space-time hybridizable discontinuous Garlerkin discretization of advection(-diffusion). SIAM J. Sci. Comput. **43**, A3393–A3416 (2021)
49. Stewart, W.J.: Introduction to the Numerical Solution of Markov Chains. Princeton University Press, Princeton (1994)
50. Tian, X.C., Du, Q.: Analysis and comparison of different approximations to nonlocal diffusion and linear peridynamic equations. SIAM J. Numer. Anal. **51**, 3458–3482 (2013)
51. Varga, R.V.: Matrix Iterative Analysis. Springer, Berlin Heidelberg (2000)
52. Wiesner, T.A., Mayr, M., Popp, A., Gee, M.W., Wall, W.A.: Algebraic multigrid methods for saddle point systems arising from mortar contact formulations. Int. J. Numer. Methods Eng. **122**, 3749–3779 (2021)
53. Xu, J.: An introduction to multilevel methods. In: Ainsworth, M., Levesley, J., Light, W.A., Marletta, M. (eds.) Wavelets, Iterative Methods for Sparse Linear Systems, Multilevel Methods and Elliptic PDEs. Leicester, 1996, pp. 213–302. Oxford University Press, New York (1997)
54. Xu, J., Zikatanov, L.: Algebraic multigrid methods. Acta Numer. **26**, 591–721 (2017)