Analysis on Aggregation and Block Smoothers in Multigrid Methods for Block Structured Linear Systems

Matthias Bolten, Marco Donatelli, Paola Ferrari, Isabella Furci

Abstract

In this talk we present a detailed analysis of block smoothers and propose new aggregation-based strategies in multigrid methods for structured linear systems [3].

Multigrid methods are well-known for their efficiency in solving large linear systems, especially when the coefficient matrices exhibit a large (multilevel block) structure [2, 7, 8]. These methods are widely used in various scientific and engineering applications, including the discretization of partial differential equations (PDEs), image processing, and approximation problems.

The key to developing effective multigrid methods lies in the careful selection of the grid transfer operator P and the iterative methods that serve as pre/post smoothers within the multigrid iteration. In previous works [2, 8], a comprehensive convergence analysis for two-grid and V-cycle methods applied to scalar Toeplitz and circulant systems was developed. Specifically, the convergence requirements for P and smoothers were formulated elegantly using conditions on the (scalar-valued) function f associated with the structured coefficient matrix. Extending this *symbol-based* analysis to block systems introduces challenges such as the non-commutativity of matrix-valued functions and the need for suitable grid transfer operators to manage coarse-grid corrections.

Scalar smoothers can be defined by carefully selecting the relaxation parameter based on the matrixvalued function \mathbf{f} , mimicking the approach used for scalar-valued functions. However, we show that, as the block dimension d increases, block smoothers become more appropriate because they align more naturally with the system's block structure.

One of the main goals of this talk is to provide an in-depth analysis of block smoothers, which are more effective in handling block-structured systems. Specifically, we introduce a relaxed block Jacobi method and derive general conditions for the smoothing parameter ω that ensure convergence. This method proves more efficient than scalar smoothers, both in terms of solving time and set-up time. From our comparison of scalar and block Jacobi smoothers, we demonstrate that the block Jacobi method consistently outperforms its scalar counterpart in terms of convergence rate and computational efficiency, particularly for systems with large block dimensions. Moreover, we show that the general conditions on smoothing parameters can be calculated with negligible computational cost in some practical applications.

Regarding the choice of grid transfer operators, a rigorous convergence analysis for the two-grid method (TGM), further extended to the V-cycle, was derived in [6]. The latter demonstrates that certain classical grid transfer operators, such as the geometric projection and standard bisection operators, meet the necessary conditions for convergence in many practical cases. The grid transfer operator in this context is written as

$$P = \mathcal{A}_n(\mathbf{p})(K \otimes I_d),$$

where K is an $n \times k$ matrix that reduces the dimension of the problem by selecting specific rows, and $\mathcal{A}_n(\mathbf{p})$ is a structured matrix analogous to the original problem. A key aspect of such grid transfer operators is that they preserve the block structure at coarser levels, and the proof of the approximation convergence property is based on validating further commutativity conditions on the matrix-valued symbol \mathbf{p} . However, many known grid transfer operators were not covered by the previous theory. To address this, new conditions were obtained in [4], expressing the projector's approximation property in terms of the eigenvector associated with the ill-conditioned subspace, thereby broadening the class of valid operators. Specifically, convergence results in the block-structured setting were derived by exploiting block-symbol analysis. If $\mathbf{f}(\theta)$ is a trigonometric polynomial with exactly one zero eigenvalue at θ_0 and is positive definite in $[0, 2\pi) \setminus \{\theta_0\}$, it can be diagonalized as

$$\mathbf{f}(\theta) = Q(\theta)D(\theta)Q(\theta)^H$$

where $Q(\theta)$ is the matrix of eigenvectors and $D(\theta)$ is the diagonal matrix of eigenvalues. We denote by $q_{\bar{j}}(\theta)$ the normalized eigenvector associated with the zero eigenvalue $\lambda_{\bar{j}}(\mathbf{f}(\theta_0)) = 0$. Under certain assumptions, sufficient conditions for the linear convergence of the TGM involve choosing \mathbf{p} such that the function $\mathbf{p}(\theta)^H \mathbf{p}(\theta) + \mathbf{p}(\theta + \pi)^H \mathbf{p}(\theta + \pi)$ is positive definite for all $\theta \in$ $[0, 2\pi)$ and specific limit conditions as θ approaches θ_0 are satisfied. These requirements can be simplified in specific applications, and the V-cycle convergence conditions are based on these results. An important aspect of the strategy outlined above is that the grid transfer operator maintains the block structure at coarser levels. While this is theoretically convenient for proving V-cycle convergence, it introduces computational challenges, as the block structure remains, and the matrixvalued function can be difficult to analyze.

The second aim of this talk is to present a new symbol-based multigrid method with an aggregationbased approach, which reduces the system to a scalar form at coarser levels. In particular, from the decomposition of the matrix-valued trigonometric polynomial \mathbf{f} , we propose a grid transfer operator of the form

$$P = I_n \otimes q_{\overline{i}}(\theta_0).$$

This approach offers significant computational advantages, especially for large-scale systems [1]. At the coarse level, the coefficient matrix simplifies, resulting in a scalar-valued function

$$\tilde{f}(\theta) = q_{\bar{j}}^{H}(\theta_0)\mathbf{f}(\theta)q_{\bar{j}}(\theta_0),$$

which maintains some properties of the original problem. Indeed, \tilde{f} vanishes at θ_0 (and only at θ_0), and with a zero of order smaller than or equal to that of the original $\lambda_{\bar{j}}(\mathbf{f})$, ensuring that the conditioning of the problem does not worsen.

We derive the convergence and optimality of the TGM by combining the approximation property verified by P and the findings on the block smoothers. We also extend the TGM analysis to the V-cycle, where the properties of the scalar-valued symbol at the coarse level are crucial in determining the convergence rate. In particular, we show that V-cycle convergence can be achieved by combining the TGM results with an analysis of the scalar coarse-level symbol. This allows us to formulate clear conditions for convergence and optimality, even in complex block settings.

From a computational perspective, aggregation-based methods provide substantial savings while maintaining convergence, particularly when combined with over-relaxation strategies. As proposed by Braess [5], over-relaxation of the coarse-grid correction significantly enhances multigrid performance, and we show the effectiveness of this strategy when combined with our symbol-based grid transfer operators. This corresponds to performing an over-relaxation with a parameter $\alpha > 1$ when computing the interpolation of the error, αPy . Consequently, we present a strategy to select the optimal pair (ω, α) to minimize the spectral radius of the TGM iteration matrix, thereby improving results. To validate our theoretical findings, we conduct numerical experiments using the proposed approach both as a standalone method and as a preconditioner for Krylov iterative methods. The tested largescale block circulant, block Toeplitz-like, and Generalized Locally Toeplitz (GLT) linear systems stem from discretizations using \mathbb{Q}_d Lagrangian FEM approximation of second-order differential problems and B-spline discretization with non-maximal regularity. In addition to confirming our theoretical results, these examples demonstrate how the conditions outlined for block smoother convergence simplify in practical scenarios.

References

- C. An, Y. Su. An aggregation-based Two-Grid method for multilevel block Toeplitz linear systems. J. Sci. Comput., 98(3): Paper No. 54, 2024.
- [2] A. Aricò, M. Donatelli. A V-cycle multigrid for multilevel matrix algebras: proof of optimality. Numer. Math., 105 (4): 511–547, 2007.
- [3] M. Bolten, M. Donatelli, P. Ferrari, I. Furci. Analysis on aggregation and block smoothers in multigrid methods for block Toeplitz linear systems. (under revision), arXiv:2403.02139v1
- [4] M. Bolten, M. Donatelli, P. Ferrari, I. Furci. A symbol based analysis for multigrid methods for Block-Circulant and Block-Toeplitz Systems. SIAM J. Matrix Anal. Appl., 43(1): 405–438, 2022.
- [5] D. Braess. Towards algebraic multigrid for elliptic problems of second order. Computing, 55(4): 379–393, 1995.
- [6] M. Donatelli, P. Ferrari, I. Furci, S. Serra–Capizzano, D. Sesana. Multigrid methods for Block-Toeplitz linear systems: Convergence Analysis and Applications. *Numer. Linear Algebra Appl.*, 28(4): e2356, 2021.
- [7] H. Elman, A. Ramage. Fourier Analysis of multigrid for a model two-dimensional convectiondiffusion equation. *BIT Numer Math*, 46, 283-306, 2006
- [8] G. Fiorentino, S. Serra, A. Ramage. Multigrid methods for symmetric positive definite block Toeplitz matrices with nonnegative generating functions. SIAM J. Sci. Comput., 17(5), 1068-1081, 1996.