LOW-RANK CORRECTIONS FOR INCREASING THE ARITHMETIC INTENSITY OF THE PRECONDITIONERS

Àdel Alsalti-Baldellou[a,b], Carlo Janna[c,d], Xavier Álvarez-Farré[e] and F. Xavier Trias[a]

[a]Heat and Mass Transfer Technological Center,
Universitat Politècnica de Catalunya – BarcelonaTech
Carrer de Colom 11, 08222 Terrassa (Barcelona), Spain
{adel.alsalti, francesc.xavier.trias}@upc.edu

[b]Termo Fluids S.L., Carrer de Magí Colet 8, 08204 Sabadell (Barcelona), Spain

[c]Department of Civil, Environmental and Architectural Engineering, University of Padova
Via Marzolo 9, 35131 Padova, Italy
carlo.janna@unipd.it

[d]M3E S.r.l., Via Giambattista Morgagni, 44, 35121 Padova, Italy

[e]High-Performance Computing Team, SURF
Science Park 140, 1098 XG Amsterdam, Netherlands
xavier.alvarezfarre@surf.nl

Key words: Spatial symmetries, Poisson equation, FSAI, AMG, SpMM, CFD

Abstract. Spatial reflection symmetries are common in many academic and industrial CFD configurations. Given a mesh with s symmetries, it is possible to transform Poisson's equation into a set of 2^s fully-decoupled subsystems. By doing so, we can increase the arithmetic intensity, reduce the memory footprint, and improve the convergence of the linear solvers. We have recently investigated ways to exploit symmetries for preconditioning Poisson's equation. With this aim, we have recalled the close similarity of the different subsystems to combine low-rank corrections with standard preconditioning techniques– namely, FSAI and AMG. This talk will overview the different strategies considered and their resulting performance.

1 INTRODUCTION

Divergence constraints are present in many disciplines and usually lead to a Poisson equation whose solution is one of the most computationally intensive parts of scientific simulation codes. Despite the relevance of theoretical aspects like the solvers' time complexity or stability, unrelated computational factors may make them impractical for large-scale simulations. The fact that most algorithms have a low arithmetic intensity makes them enjoy a small fraction of the systems' peak performances. Approaches trying to remedy this are generally based on reducing memory traffic, solving multiple right-hand sides (RHSs), using mixed-precision algorithms or adapting more compute-intensive methods.

This work focuses on a strategy exploiting spatial reflection symmetries to block diagonalise Poisson's equation [1]. Then, thanks to the subsystems' close similarity, we add another level of approximation by applying to each of them the same preconditioner, later corrected by means of relatively cheap low-rank corrections. This strategy allows replacing the standard sparse matrix-vector product (`SpMV`) with the more compute-intensive sparse matrix-matrix product (`SpMM`) [2] in the application of factored sparse approximate inverse (FSAI) and Algebraic Multigrid (AMG) preconditioners.

Without loss of generality, we will focus our numerical experiments on incompressible CFD simulations, arising our Poisson equation from an application of the fractional step method and equalling:

$$\nabla \cdot \left(\frac{1}{\rho} \nabla p \right) = \frac{1}{\Delta t} \nabla \cdot v^p, \tag{1}$$

where ρ, p and v^p stand for the density, pressure and predictor velocity fields, respectively. Then, its discrete version will read:

$$Ax = b, \tag{2}$$

where $b \in \text{range}(A)$, and the coefficient matrix, A, stands for the discrete Laplace operator and is assumed symmetric positive semi-definite. Further details about the discretisation can be found in [3].

2 POISSON'S EQUATION BLOCK DIAGONALISATION

For the sake of clarity, let us consider a mesh with a single reflection symmetry. Additionally, let us impose the same grid points' ordering on each of the sides [1]. As a result, the mesh is halved into two subdomains and all the scalar fields satisfy:

$$x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \in \mathbb{R}^n, \tag{3}$$

where $x_1, x_2 \in \mathbb{R}^{n/2}$ correspond to x's restriction to each of the subdomains. Then, spatially symmetric points are in the same position within the subvectors, and the discrete Laplace operator satisfies:

$$A = \begin{pmatrix} A_{\text{inn}} & A_{\text{out}} \\ A_{\text{out}} & A_{\text{inn}} \end{pmatrix} \in \mathbb{R}^{n \times n}, \tag{4}$$

where n stands for the mesh size and $A_{\text{inn}}, A_{\text{out}} \in \mathbb{R}^{n/2 \times n/2}$ for the inner- and outer-subdomain couplings, respectively. Remarkably enough, virtually all the operators satisfy the following structure:

$$H = \begin{pmatrix} H_{\text{inn}} & H_{\text{out}} \\ H_{\text{out}} & H_{\text{inn}} \end{pmatrix} \in \mathbb{R}^{n \times m}, \tag{5}$$

where $H_{\text{inn}}, H_{\text{out}} \in \mathbb{R}^{n/2 \times m/2}$ stand for the inner- and outer-subdomain couplings, being the latter substantially sparser, if not null.

In such a context, we can define the following change-of-basis matrix:

$$P := \frac{1}{\sqrt{2}} \begin{pmatrix} \mathbb{I}_{n/2} & \mathbb{I}_{n/2} \\ \mathbb{I}_{n/2} & -\mathbb{I}_{n/2} \end{pmatrix} \in \mathbb{R}^{n \times n}, \tag{6}$$

which satisfies $P^{-1} = P$. Then, changing the basis of A by means of P leads to:

$$\hat{A} \coloneqq PAP^{-1} = \begin{pmatrix} A_{\text{inn}} + A_{\text{out}} & 0 \\ 0 & A_{\text{inn}} - A_{\text{out}} \end{pmatrix}, \tag{7}$$

decomposing eq. (2) into two fully-decoupled half-sized subsystems, $\hat{A}_1 \coloneqq A_{\text{inn}} + A_{\text{out}}$ and $\hat{A}_2 \coloneqq A_{\text{inn}} - A_{\text{out}}$. Algorithm 1 summarises the resulting strategy.

Algorithm 1 Poisson solver exploiting one reflection symmetry

Require: $A_{\text{inn}}, A_{\text{out}} \in \mathbb{R}^{n/2 \times n/2}$ and $b \in \text{range}(A) \subseteq \mathbb{R}^n$

procedure SOLVE$(b, \hat{A})$
  Transform $\hat{b} = Pb$
  Decoupled solution of $\hat{A}_1 \hat{x}_1 = \hat{b}_1$ and $\hat{A}_2 \hat{x}_2 = \hat{b}_2$
  Inverse transform $x = P^{-1}\hat{x}$
  return x
end procedure

3 LOW-RANK CORRECTIONS FOR FSAI

The idea of applying low-rank corrections arises from the close similarity between $\hat{A}$'s subsystems. Indeed, in eq. (6), all the outer- are substantially sparser than the inner-couplings. In fact, $\text{rank}(A_{\text{out}}) = \mathcal{O}(n^{2/3})$, whereas $\text{rank}(A_{\text{inn}}) = \mathcal{O}(n)$. Hence:

$$\text{rank}\left(A_{\text{out}}\right) \ll \text{rank}\left(A_{\text{inn}}\right), \tag{8}$$

and it makes sense to introduce another level of approximation to FSAI by assuming that:

$$\hat{A}_1 \simeq A_{\text{inn}} \text{ and } \hat{A}_2 \simeq A_{\text{inn}}. \tag{9}$$

In the context of preconditioning linear systems, much work has recently been devoted to low-rank matrix representations [4, 5, 6]. The FSAI of A_{inn} provides an approximation to the inverse of A_{inn}'s lower Cholesky factor, $G_{\text{inn}} \simeq L_{\text{inn}}^{-1}$, which ensures that $G_{\text{inn}}^T G_{\text{inn}} \simeq A_{\text{inn}}^{-1}$. Then, we can define the following auxiliary matrix for each subsystem $\hat{A}_i$:

$$Y \coloneqq \mathbb{I}_{n/2} - G_{\text{inn}} \hat{A}_i G_{\text{inn}}^T \in \mathbb{R}^{n/2 \times n/2}, \tag{10}$$

whose definition yields $Y(\mathbb{I}_{n/2} - Y)^{-1} = (G_{\text{inn}} \hat{A}_i G_{\text{inn}}^T)^{-1} - \mathbb{I}_{n/2}$, finally leading to:

$$\hat{A}_i^{-1} = G_{\text{inn}}^T G_{\text{inn}} + G_{\text{inn}}^T Y (\mathbb{I}_{n/2} - Y)^{-1} G_{\text{inn}}. \tag{11}$$

By virtue of eq. (8), we can expect the "full-rank" correction of eq. (11) to be well represented by a low-rank approximation [4]. Hence, let us truncate Y's eigendecomposition to account for its k most *relevant* eigenpairs:

$$Y \simeq V_k \Sigma_k V_k^T, \text{ such that } V_k \in \mathbb{R}^{n/2 \times k} \text{ and } \Sigma_k \in \mathbb{R}^{k \times k}. \tag{12}$$

Then, by defining $W_k \coloneqq G_{\text{inn}}^T V_k \in \mathbb{R}^{n/2\times k}$ and $\Theta_k \coloneqq \Sigma_k(\mathbb{I}_k - \Sigma_k)^{-1} \in \mathbb{R}^{k\times k}$, eqs. (11) and (12) can be combined to give the following low-rank corrected FSAI:

$$\hat{A}^{-1} \simeq \mathbb{I}_2 \otimes G_{\text{inn}}^T G_{\text{inn}} + \begin{pmatrix} W_{k,1}\Theta_{k,1}W_{k,1}^T & 0 \\ 0 & W_{k,2}\Theta_{k,2}W_{k,2}^T \end{pmatrix}, \tag{13}$$

which allows replacing `SpMV` with `SpMM` in G_{inn}'s application.

In the conference, we will present the strategy with more detail, extend it to AMG, and provide meaningful numerical results illustrating the advantages and disadvantages of our proposal.

ACKNOWLEDGEMENTS

We would like to thank the team of M3E s.r.l., and most notably PhD. Matteo Frigo, for their support with Chronos throughout the implementation of these preconditioners.

This work was financially supported by two competitive R+D projects: RETOtwin (PDC2021-120970-I00), given by MCIN/AEI/ 10.13039/501100011033 and European Union Next GenerationEU, and FusionCAT (001-P-001722), given by *Generalitat de Catalunya* RIS3CAT-FEDER. Àdel Alsalti-Baldellou has also been supported by the predoctoral grants DIN2018-010061 and 2019-DI-90, given by MCIN/AEI/10.13039/501100011033 and the Catalan Agency for Management of University and Research Grants (AGAUR), respectively.

REFERENCES

[1] A. Gorobets, F. Trias, R. Borrell, O. Lehmkuhl, and A. Oliva, "Hybrid MPI+OpenMP parallelization of an FFT-based 3D Poisson solver with one periodic direction," *Comput. Fluids*, vol. 49, no. 1, pp. 101–109, 2011.

[2] H. Anzt, S. Tomov, and J. Dongarra, "On the performance and energy efficiency of sparse linear algebra on GPUs," *Int. J. High Perform. Comput. Appl.*, vol. 31, pp. 375–390, sep 2017.

[3] F. X. Trias, O. Lehmkuhl, A. Oliva, C.-D. Pérez-Segarra, and R. W. C. P. Verstappen, "Symmetry-preserving discretization of Navier-Stokes equations on collocated unstructured grids," *J. Comput. Phys.*, vol. 258, pp. 246–267, feb 2014.

[4] R. Li and Y. Saad, "Divide and conquer low-rank preconditioners for symmetric matrices," *SIAM Journal on Scientific Computing*, vol. 35, pp. A2069–A2095, 2013.

[5] P. Ghysels, X. S. Li, F. H. Rouet, S. Williams, and A. Napov, "An efficient multicore implementation of a novel HSS-structured multifrontal solver using randomized sampling," *SIAM Journal on Scientific Computing*, vol. 38, pp. S358–S384, 2016.

[6] A. Franceschini, V. A. P. Magri, M. Ferronato, and C. Janna, "A robust multilevel approximate inverse preconditioner for symmetric positive definite matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 39, pp. 123–147, 1 2018.