ParCFD
32^{nd} International Conference on Parallel Computational Fluid Dynamics
May-17-19 2020, Nice FRANCE

ACCELERATION OF THE CONVERGENCE OF THE ASYNCHRONOUS RESTRICTIVE ADDITIVE SCHWARZ METHOD

Damien TROMEUR-DERVOUT[†]

[†]Université de Lyon, CNRS, Université Lyon 1, Institut Camille Jordan UMR 5208,
43 bd du 11 novembre 1918, F-69622 Villeurbanne-Cedex
e-mail: damien.tromeur-dervout@univ-lyon1.fr

Key words: Asynchronous Schwarz, Aitken acceleration of the convergence

Abstract. This paper focuses on the Aitken's acceleration of the convergence of the asynchronous Restricted Additive Schwarz (RAS) iterates for which transmission conditions on some artificial interfaces generated by the domain decomposition may not have not been updated for some iterates. We show that despite the error operator depends of the iterate, the low-rank approximation of this operator using singular value decomposition of the iterated interface solutions allows to still accelerate. We develop a modeling of the asynchronous RAS allowing to set the level of asynchronism and show how it deteriorates its convergence and its convergence's acceleration. Finaly, some regularization techniques on the iterated interface solutions allow to enhance the acceleration of the convergence .

1 Introduction

Nowadays high performance computers have several thousand of cores and more and more complex hierarchical communication networks. For these architectures, the use of global reduction operation such as the dot products involving in the GMRES acceleration can be the bottleneck for the performance. In these context domain decomposition method as solver, with local communications are becoming particularly interesting. Nevertheless, the probability of temporarly failures/unavailabilty of a set of processors/clusters is non zero leading to the need of fault tolerant algorithms such as asynchronous Schwarz type method. With the asynchronism the transmission conditions at artificial interfaces generated by the domain decomposition may not have been updated for some subdomains and for some iterates. Message passing interface MPI-3 standard provides one side communication protocol where a process can directly write on the local memory of an another process without any synchronizing. The asynchronism in the transmission condition update can also occur in OpenMP implementation where a thread can also write on shared memory whithout any synchronizing. For asynchronous methods is very difficult to know if the update has been perform and most of papers lack to give the level of asynchronism in their implementation results. If the Schwarz DDM converges then the asynchronous Schwarz does the same [5].

From the numerical point of view this asynchronism affects the linear operator of the interface problem. In these context the Aitken's acceleration of convergence should not be applicable as it is based on the pure linear convergence of the DDM [4] [7] [8], i.e. there exists a linear operator P independent of the iteration that links the error with the searched solution at the artificial interfaces of two consecutive iterates.

This paper focuses on the Aitken's acceleration of the convergence of the asynchronous Restricted Additive Schwarz (RAS) iterates. We develop a mathematical model of the Asynchronous RAS allowing to set the percentage of the number of the randomly chosen local artificial interfaces where transmission conditions are not updated. Then we show how this ratio deteriorates the convergence of the Asynchronous RAS and how some regularization techniques on the traces of the iterative solutions at artificial interfaces allow to accelerate the convergence to the searched solution.

The plan of the paper is the following, section 2 gives the notation and the principles of the Aitken-Schwarz using some low-rank approximation of the interface error operator. Section 3 presents the modeling of the asynchronous RAS on a 2D Poisson problem allowing to define the level of asynchronism. Section 4 presents the results of the acceleration with respect to the level of asynchronism and the enhancement of this acceleration with regularisation techniques before to conclude in section 5.

2 Aitken-Schwarz method principles

One Restrictive Additive Schwarz (RAS) iterate[3] to solve $Ax^\infty = b \in \mathbb{R}^n$ writes on subdomain Ω_i: $x_i^{k+1} = A_i^{-1}\left(b_i - E_i x_{i,e}^k\right)$, with $R_i \in \mathbb{R}^{n_i \times n}$ the operator that restricted the global vector to the subdomain Ω_i, including the overlap, $\widetilde{R}_i \in \mathbb{R}^{n_i \times n}$ lthe operator that restricted the global vector to the subdomain Ω_i, with setting to 0 the components of the vector that correspond to the overlap. W_i is the global index set of the unknows belonging to the subdomain Ω_i. A_i is the part of the operator A associated to the subdomain Ω_i: $A_i = R_i A R_i^T$. $x_i = R_i x$ and $b_i = R_i b$ are the restriction to the subdomain Ω_i of the solution and the right hand side respectivelly. $x_{i,e}$ represents the data dependencies external of the subdomain Ω_i : $x_{i,e}$ is composed of the x_j such that $A_{kj} \neq 0$ with $k \in W_i$ and $j \notin W_i$. $R_{i,e}$ is the restriction operator such that $R_{i,e}x = x_{i,e}$. E_i is the part of the matrix A that represents the effect of the unknows external to the subdomain Ω_i on the unknows belonging to the subdomain Ω_i : $E_i = R_i A R_{i,e}^T$. Writting $M_{RAS}^{-1} = \sum_{i=0}^{P-1} \widetilde{R}_i^T A_i^{-1} R_i$, one RAS iterate can be view as a Richardson's process:

$$x^{k+1} = x^k + M_{RAS}^{-1}\left(b - Ax^k\right). \tag{1}$$

The restriction of the Richardson's process (1) to the interface $\Gamma = \cup_{i=0}^{P-1} \partial\Omega_i \backslash \partial\Omega$ writes [9]:

$$\begin{array}{rcl} R_\Gamma x^{k+1} & = & R_\Gamma\left(I - M_{RAS}^{-1}A\right) R_\Gamma^T R_\Gamma x^k + R_\Gamma M_{RAS}^{-1} b. \\ \Leftrightarrow y^k & = & (I-P)y^{k-1} + c \end{array} \tag{2}$$

The pure linear convergence of the RAS at the interface given by : $y^k - y^\infty = (I-P)(y^{k-1} - y^\infty)$ (the operator P does not depend of the iteration k) allows to apply the Aitken acceleration of the convergence technique to obtain the searched solution

y^∞ on the interface Γ: $y^\infty = (I-P)^{-1}(y^k - Py^{k-1})$, and thus after one more local solving the searched solution x^∞. Let us note that we can accelerate the convergence to the solution for a convergent or a divergent iterative method. The only need is that there is no eigenvalue of modulus 1 in P.

Considering $e^k = y^{k+1} - y^k, k = 0, \ldots$, the operator $P \in \mathbb{R}^{n_\Gamma \times n_\Gamma}$ can be computed algebraicaly after n_Γ iterates as $P = [e^{n_\Gamma}, \ldots, e^1][e^{n_\Gamma - 1}, \ldots, e^0]^{-1}$. Nevertheless, for 2D or 3D problems, the value n_Γ may be too large to have an efficient method. So a low-rank approximation of P is computed using the iterated interface solutions and the Aitken's acceleration is performed on the low-rank space of dimension p. As we search, the converged interface solution, we build from the singular value decomposition [6] of the matrix $Y = [y^0, \ldots y^q] = U\Sigma V^T$ a low-rank space with selecting the p singular vectors associated to the most significant singular values.

Algorithm 1 Approximated Aitken's acceleration

Require: y^0 arbitrary initial condition , $\epsilon > 0$ given tolerance

1: **repeat**
2:   **for** $i = 1 \ldots q$ **do**
3:     $y^i \leftarrow Py^{i-1} + c$ // RAS iterate
4:   **end for**
5:   Compute SVD of $[y^0, y^1, \ldots, y^q] = U\Sigma V'$,
    keep the p singular vectors $U_{1:p}$ such that $\sigma_{p+1} < \epsilon$
6:   Compute $[\hat{y}^{q-p-2}, \ldots, \hat{y}^q] = U_{1:p}^T[y^{q-p-2}, \ldots, y^q]$, and $\hat{e}^i = \hat{y}^i - \hat{y}^{i-1}$
7:   Compute $\widehat{P} = [\hat{e}^{q-p} \ldots, \hat{e}^q][\hat{e}^{q-p-1}, \ldots, \hat{e}^{q-1}]^{-1}$
8:   $y^0 \leftarrow U_{1:p}\left(I - \widehat{P}\right)^{-1}\left(U_{1:p}^T\hat{y}^q - \widehat{P}\hat{y}^{q-1}\right)$
9: **until** convergence

This low-rank approximation of the acceleration has been very efficient to solve 3D Darcy flow with highly heterogeneous and randomly generated permeability field [1]. Step 7 of the algorithm may be subject to bad conditioning and matrix inversion can be replaced by pseudo inverse. Other techniques developed in [1] avoid the matrix inversion. For 1D partitioning, we can use the sparsity of P to define a Sparse-Aitken acceleration, numerically more efficient with local SVD for each subdomain [2].

3 Modeling the Asynchronous RAS

Consider the 2D Poisson problem with homogeneous Dirichlet boundary conditions:

$$\left\{ -\left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2}\right)u = f,\ (x,y) \in]0,1[\times]0,1[, \text{ with homogeneous B.C.} \right. \tag{3}$$

Discretizing (3) with second order finite differences on a regular mesh with step sizes h_x and h_y allows us to define an approximate solution $u_{i,j} \simeq u(i\,h_x, j\,h_y)$ at coordinates (i,j) of this mesh of $N_x^g \times N_y^g$ points.

Given a not prime number $P \in \mathbb{N}$, we split the domain $[0,1]^2$ in $P = P_x \times P_y$ subdomains Ω_i with p overlaping points of discretizing in each direction. For sake of simplicity we consider that each subdomain has $N_x^l \times N_y^l$ points of discretizing. We consider the domain partitioning $\cup_{i=0}^{p-1} \Omega_i$ and write Ω_i^O, Ω_i^E, Ω_i^S, Ω_i^N,the neighbour subdomains of Ω_i in the West, East, North and South directions. The asynchronous RAS algorithm does not wait that the update of the transmission conditions are done before starting the next iterate. Especially , when one sided MPI communications are involved one process can directly write to the memory of an another process with any synchronization , and consequently the boundary condition of one subdomain can be totally, partially or not updated. As there is not control on the restraining of the communication network, it is difficult to evaluate the number of update of the local boundaries missing. In order to modelize the Asynchronous RAS, we proposed model where the four transmission conditions of each subdomains are totally update or not following a random draw of four numbers per subdomain (l_W, l_E, l_S, l_N) greater than a fixed limit l (giving the percentage of missing boundary conditions update). The synchronous RAS algorithm is obtained setting $l = 0$ and we note l-RAS the asynchronous RAS with a l level of asynchronism. The l-RAS iterates until the traces of the iterated solution of all the subdomains do not evolve anymore. It writes for the $n+1^{th}$ iterate:

$$\begin{cases} \dfrac{-u_{i+1,j}^{n+1} + 2u_{i,j}^{n+1} - u_{i-1,j}^{n+1}}{h_x^2} + \dfrac{-u_{i,j+1}^{n+1} + 2u_{i,j}^{n+1} - u_{i,j-1}^{n+1}}{h_y^2} = f_{i,j}, \\ \qquad\qquad\qquad 1 \le i \le N_x^l - 2,\ 1 \le j \le N_y^l - 2, \\ u_{0,j}^{n+1} = \begin{cases} u_{N_x^l-p-1,j}^{W,n} & \text{if W exists and } l_W > l \\ 0 & \text{else} \end{cases} \quad , 1 \le j \le N_y^l - 2 \\ u_{N_x^l-1,j}^{n+1} = \begin{cases} u_{p,j}^{E,n} & \text{if E exists and } l_E > l \\ 0 & \text{else} \end{cases} \quad , 1 \le j \le N_y^l - 2 \\ u_{i,0}^{n+1} = \begin{cases} u_{i,N_y^l-p-1}^{S,n} & \text{if S exists and } l_S > l \\ 0 & \text{else} \end{cases} \quad , 1 \le i \le N_x^l - 2 \\ u_{i,N_y^l-1}^{n+1} = \begin{cases} u_{i,p}^{N,n} & \text{if N exists and } l_N > l \\ 0 & \text{else} \end{cases} \quad , 1 \le i \le N_x^l - 2 \end{cases} \tag{4}$$

Figure 1 (left) shows that level of asynchronousm deteriorates the convergence of the RAS. The error between two consecutive iterates oscillates quite strongly with l. This oscillation is smoother for the error with the searched solution.

Table 1 shows that the error with the searched solution of the asynchronous RAS converges to zero for 240 iterates with an increasing variance in the results with respect to the asynchronism level l but with small amplitude between the min,max and mean values. We limited to $p = 40$ first singular values for $l \neq 0$ and to $p = 20$ first for $l = 0$ due to the strong singular values decreasing. The Aitken's acceleration, using the set of 240 iterates, still works even with the asynchronism that also deteriorates the convergence but with more stable variance in results and mean values more close to the max than to the min. Let us notice for this test case $n_\Gamma = 544$ and the approximated acceleration is of size 40.

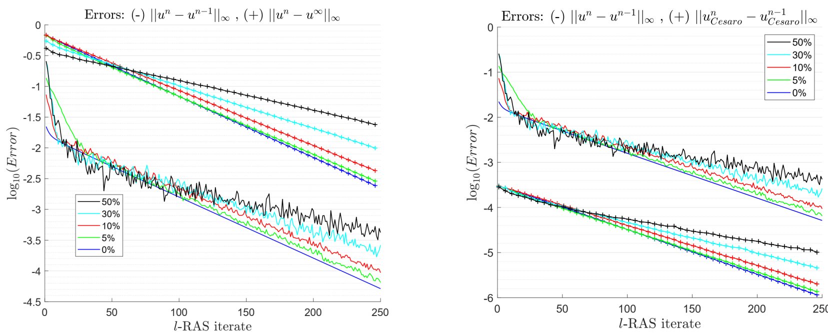

Figure 1: l-RAS convergence with respect to the level of asynchronous l: for two consecutive iterates (continuous line) and (left) with the searched solution (+) , (right) two consecutive iterates after Césaro's summing (+). $(N_x^l = N_y^l = 10, P_x = P_y = 5, p = 40)$

	Aitken l-RAS				l-RAS				Update failures			
l	min	max	mean	σ	min	max	mean	σ	min	max	mean	σ
0.0%	-11.12	-11.12	-11.12	2e-14	-2.543	-2.543	-2.543	3e-15	0	0	0	0
0.5%	-3.666	-5.839	-4.969	4.0e-1	-2.527	-2.556	-2.533	4.8e-3	99	145	120.7	9.9
1.0%	-2.814	-5.440	-4.751	4.7e-1	-2.513	-2.544	-2.524	7.1e-3	202	277	239.48	15.6
5.0%	-2.521	-5.023	-4.284	4.2e-1	-2.415	-2.479	-2.443	1.4e-2	1121	1286	1197.3	34.3
10.%	-1.729	-4.707	-3.956	5.3e-1	-2.303	-2.406	-2.347	2.1e-2	2267	2502	2397.9	43.6
30.%	-1.037	-4.005	-3.280	4.6e-1	-1.868	-2.089	-1.974	4.7e-2	7044	7349	7203.3	66.5
50.%	0.548	-3.613	-2.643	6.1e-1	-1.472	-1.961	-1.678	9.3e-2	11860	12199	12013	66.1

Table 1: Statistics (min,max,mean and variance σ) of $log10(||x^{240} - x^\infty||_\infty)$ for asynchronous RAS and its Aitken's acceleration (with the same data) with respect to the asynchronism level l for 100 runs. $(N_x^l = N_y^l = 10, P_x = P_y = 5, p = 40)$

4 Regularization of the Aitken acceleration of the convergence of the Asynchronous RAS

At first glance, previous results on Aitken's acceleration are surprising as the pure linear convergence of the RAS is destroyed with the asynchronism i.e. the error operator depends of the iterate : $y^{k+1} - y^k = P_k(y^k - y^{k-1})$.

The explanation comes from the low-rank space built from the SVD. Let $Y_l = [y_l^0, \dots, y_l^q]$ the matrix of iterated asynchronous l-RAS interface solutions. As the asynchronous l-RAS converges one can write $Y_l = Y_0 + E_l$ where E_l is a perturbation matrix with smaller and smaller entries with respect to the iterates. Then using the Fan inequality of the svd of a perturbated matrix: $\sigma_{r+s+1}(Y_0 + E_l) \le \sigma_{r+1}(Y_0) + \sigma_{s+1}(E_l)$ with $r, s \ge 0$, $r+s+1 \le q+1$. Considering $s = 0$ we have $|\sigma_{r+1}(Y_0 + E_l) - \sigma_{r+1}(Y_0)| \le \sigma_1(E_l) = ||E_l||_2, \forall r \le q$. Using

	Aitken Césaro l-RAS				Upper	Aitken l-RAS	l-RAS
l	min	max	σ	mean	bound	mean	mean
0.0%	-12.42	-12.42	8.e-15	-12.42	-12.27	-11.111	-2.543
0.5%	-4.059	-6.968	4.5e-1	-6.284	-6.120	-4.969	-2.533
1.0%	-4.667	-6.856	3.9e-1	-6.096	-5.902	-4.751	-2.524
5.0%	-4.184	-6.383	4.9e-1	-5.546	-5.434	-4.284	-2.443
10.%	-3.844	-6.047	4.5e-1	-5.294	-5.106	-3.956	-2.347
30.%	-3.457	-5.261	3.9e-1	-4.500	-4.431	-3.280	-1.974
50.%	-2.505	-4.553	4.7e-1	-3.841	-3.794	-2.643	-1.678

Table 2: Statistics (min,max,mean and variance σ) of $log10(||x^{240} - x^{\infty}||_{\infty})$ Aitken-RAS with Cesaro's summation with respect to the asynchronousm level l for 100 runs. $(N_x^l = N_y^l = 10, P_x = P_y = 5, p = 40, m = 200)$

the Schmidt's theorem on the svd approximation, we can write:

$$\begin{aligned} \min_{X,rankX=k}(||Y_l - X||_2) &= \sigma_{k+1}(Y_l) = \min_{X,rankX=k}(||Y_l - Y_0 + Y_0 - X||_2 \\ &\leq ||Y_l - Y_0||_2 + \min_{X,rankX=k} ||Y_0 - X||_2 \\ &\leq \sigma_1(E_l) + \sigma_{k+1}(Y_0) \end{aligned} \tag{5}$$

This result implies that:

- the low-rank space U_l built from Y_l is an approximation of U_0 with a small perturbation $||E_l||_2$.
- as the perturbation $||E_l||_2$ is dominated by the perturbation of the first iterates better acceleration is obtained with considering the last iterates to build U_l

This last result suggests an enhancement of the Aitken's acceleration with the Césaro's summation of the iterated interface solutions in order to transform the sequence y_l in an another sequence $\tilde{y}_l$ with $\tilde{y}_l^i = \dfrac{1}{m}\sum_{j=0}^{m-1} y_l^{i+j}$ that will smooth the perturbation E_l. Let us notice that this transformation still preserves the pure linear convergence of the synchronous 0%-RAS: $\tilde{y}_0^{k+l} - y^{\infty} = P(\tilde{y}_0^k - y^{\infty})$.

Figure 1 (right) shows that the Césaro summing allows to smooth the convergence oscillations of l-RAS. The difference between two consecutive iterates have a smaller amplitude than l-RAS leading to have a low-rank space U_l more representative of the searched solution. The summation is done with $m = 200$.

Table 2 gives the Aitken's acceleration of the l-RAS with the Césaro summation. It shows that the Aitken's acceleration of the convergence is enhanced, with a smaller value of the variance and smaller amplitude with the min and the max values than for the Aitken procedure applied to the l-RAS. Even the 0%-RAS is better accelerated. Morever it shows a upper bound of the mean acceleration of the l-RAS with Césaro summation of $\dfrac{1}{\sqrt{m}}$ the mean acceleration of the l-RAS.

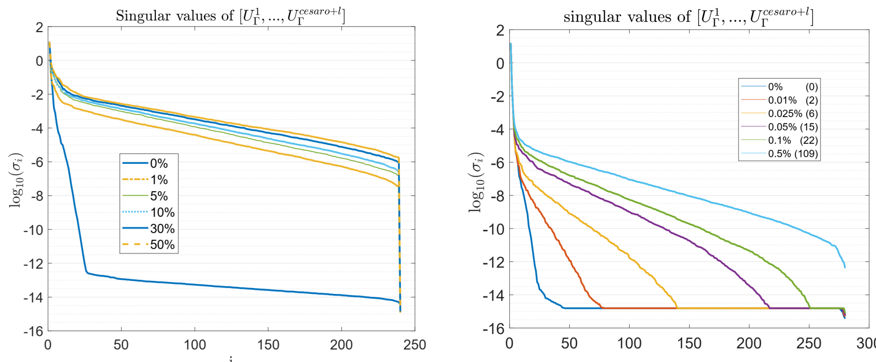

Figure 2: Singular values of one sample of 250 l-RAS iterates for $l = \{0\%, 1\%, 5\%, 10\%, 30\%, 50\%\}$ (left) and for $l = \{0\%, 0.01\%, 0.025\%, 0.05\%, 0.1\%, 0.5\%\}$ and 300 iterates with the number of transmission condition update failures in brackets (right). $(N_x^l = N_y^l = 10, P_x = P_y = 5, p = 40)$

Figure 2 gives the singular values of the SVD of l-RAS interface solution iterates with respect to the level l of asynchronism. It shows that the singular values fast decreasing is lost with the asynchronism. It still exhibits some decreasing that allows the Aitken's acceleration. The right figure shows that even with a very small level l, the decreasing deteriorates even with few transmission condition update failures (total number of update for 300 0%-RAS iterates is $300 \times (4 \times 2 + 12 \times 3 + 9 \times 4) = 24000$).

5 conclusion

In conclusion, we have succeed to accelerate the Asynchronous RAS with the Aitken's acceleration technique based on the low-rank approximation of the error operator with the singular value decomposition of the interface iterated solutions. The SVD allows to smooth the asynchronous effect over the iterations. We proposed a modeling of the asynchronousm that can be used to estimate the asynchronism in real application. Knowing the observed convergence rate of the real application we can extrapolate the level of asynchronism of the implementation. The model proposed here considers a uniform probability for transmission condition update failure (the worst case) but we also can consider that only some part of the domain decomposition can be in fault. Finally, we proposed a regularisation technique based on the Césaro's summation of the l-RAS iterates at the interface that improves the Aitken's acceleration even on the synchronous RAS.

REFERENCES

[1] L. Berenguer, T. Dufaud, and D. Tromeur-Dervout. Aitken's acceleration of the Schwarz process using singular value decomposition for heterogeneous 3D groundwater flow problems. *Computers & Fluids*, 80:320–326, 2013.

[2] L. Berenguer and D. Tromeur-Dervout. Sparse Aitken-Schwarz with application to Darcy flow. https://hal.archives-ouvertes.fr/hal-03090521, 2020.

[3] X.-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM J. Sci. Comput.*, 21(2):792–797 (electronic), 1999.

[4] M. Garbey and D. Tromeur-Dervout. On some Aitken-like acceleration of the Schwarz method. *Internat. J. Numer. Methods Fluids*, 40(12):1493–1513, 2002. LMS Workshop on Domain Decomposition Methods in Fluid Mechanics (London, 2001).

[5] F. Magoules, D. B. Szyld, and C. Venet. Asynchronous optimized Schwarz methods with and without overlap. *Numerische Mathematik*, 137(1):199–227, SEP 2017.

[6] G. W. Stewart. On the early history of the singular value decomposition. *SIAM Rev.*, 35(4):551–566, 1993.

[7] D. Tromeur-Dervout. Meshfree Adaptative Aitken-Schwarz Domain Decomposition with application to Darcy Flow. In Topping, BHV and Ivanyi, P, editor, *Parallel, Distributed and Grid Computing for Engineering*, volume 21 of *Computational Science Engineering and Technology Series*, pages 217–250. Saxe-Coburg Publications, Stirlingshire, UK, 2009.

[8] D. Tromeur-Dervout. Approximating the trace of iterative solutions at the interfaces with nonuniform Fourier transform and singular value decomposition for cost-effectively accelerating the convergence of Schwarz domain decomposition. *ESAIM: PROCEEDINGS*, 42:34–60, 2013.

[9] P Wilders and E Brakkee. Schwarz and Schur: An algebraical note on equivalence properties. *SIAM J. Sci Comput.*, 20(6):2297–2303, JUL 22 1999.