Is Additive Schwarz with Harmonic Extension Just Lions' Method in Disguise?

Felix Kwok

Section de mathématiques, Université de Genève, Geneva, Switzerland,
Felix.Kwok@unige.ch

Summary. The Additive Schwarz Method with Harmonic Extension (ASH) was introduced by Cai and Sarkis (1999) as an efficient variant of the additive Schwarz method that converges faster and requires less communication. We show how ASH, which is defined at the matrix level, can be reformulated as an iteration that bears a close resemblance to the parallel Schwarz method at the continuous level, provided that the decomposition of subdomains contains no cross points. In fact, the iterates of ASH are identical to the iterates of the discretized parallel Schwarz method outside the overlap, whereas inside the overlap they are linear combinations of previous Schwarz iterates. Thus, the two methods converge with the same asymptotic rate, unlike additive Schwarz, which fails to converge inside the overlap (Efstathiou & Gander 2007).

1 The Methods of Lions, AS, RAS and ASH

Let $\Omega \subset \mathbb{R}^n$ be a bounded open set. Suppose we want to solve the elliptic PDE

$$\mathcal{L}u = f \quad \text{on } \Omega, \qquad u = g \quad \text{on } \partial\Omega. \tag{1}$$

Based on the theoretical work of [8, 9] introduced the first domain decomposition methods for solving (1). In the two-subdomain case, let $\Omega_1, \Omega_2 \subset \Omega$ such that $\Omega_1 \cup \Omega_2 = \Omega$ and $\Omega_1 \cap \Omega_2 \neq \emptyset$. We also define $\Gamma_i = \partial\Omega \cap \bar{\Omega}_i$ and $\Gamma_{ij} = \partial\Omega_i \cap \bar{\Omega}_j$ for $i, j = 1, 2$. Then Lions' *parallel Schwarz method* calculates the subdomain iterates $u_i^k : \Omega_i \to \mathbb{R}$, $i = 1, 2$ via

$$\mathcal{L}u_i^{k+1} = f \quad \text{on } \Omega_i, \qquad u_i^{k+1} = g \quad \text{on } \Gamma_i, \qquad u_i^{k+1} = u_{3-i}^k \quad \text{on } \Gamma_{i,3-i}. \tag{2}$$

If we discretize the parallel Schwarz method (2), we obtain for $k = 0, 1, \ldots,$

$$A_1\mathbf{u}_1^{k+1} = f_1 - A_{12}\mathbf{u}_2^k, \qquad A_2\mathbf{u}_2^{k+1} = f_2 - A_{21}\mathbf{u}_1^k, \tag{3}$$

where $A_i = R_i A R_i^T$, $A_{ij} = (R_i A - A_i R_i)R_j^T$, and R_i restricts the set $V = \{1, \ldots, n\}$ of all nodes onto the subset V_i of nodes that lie in Ω_i. The above method trivially generalizes to the case of many subdomains if there are no cross points, i.e., $\Omega_i \cap \Omega_j \cap \Omega_l = \emptyset$ for distinct i, j and l:

$$A_i \mathbf{u}_i^{k+1} = f_i - \sum_{j \neq i} A_{ij} \mathbf{u}_j^k, \qquad \text{for all } i. \tag{4}$$

Note that (4) does not define a global approximate solution $\mathbf{U}^k$ that is valid over the entire domain Ω. In fact, if the subdomains overlap, there is no unique way of defining $\mathbf{U}^k$ in terms of the $\mathbf{u}_j^k$ until the method has converged. Thus, one cannot directly consider parallel Schwarz as a preconditioner for the global system and use it in combination with Krylov subspace methods.

In order to turn parallel Schwarz into a preconditioner, [3] introduced the *additive Schwarz* (AS) method, which is equivalent to a block Jacobi iteration when the subsets V_j are disjoint. However, when the subdomains overlap, the method no longer converges inside the overlapping regions ([4, 6]). To obtain a convergent method, [2] introduced the methods of *Restricted Additive Schwarz* (RAS) and *Additive Schwarz with Harmonic Extension* (ASH), which are defined as follows: let $\tilde{\Omega}_j$ be a partition of Ω such that $\tilde{\Omega}_j \subset \Omega_j$. Let $\tilde{V}_j$ be the nodes that lie in $\tilde{\Omega}_j$, and $\tilde{R}_l$ be a matrix of the same size as R_l, such that

$$[\tilde{R}_l]_{ij} = \begin{cases} 1 & \text{if } [R_l]_{ij} = 1 \text{ and } j \in \tilde{V}_l, \\ 0 & \text{otherwise.} \end{cases}$$

Then, starting from an initial guess of the *global solution* $\mathbf{U}^0$, RAS calculates

$$\mathbf{U}^{k+1} = \mathbf{U}^k + \sum_j \tilde{R}_j^T A_j^{-1} R_j (f - A\mathbf{U}^k), \tag{5}$$

whereas ASH computes

$$\mathbf{U}^{k+1} = \mathbf{U}^k + \sum_j R_j^T A_j^{-1} \tilde{R}_j (f - A\mathbf{U}^k). \tag{6}$$

By restricting either the residual or the update to $\tilde{V}_j$, RAS and ASH avoids the redundant updates that occur within the overlap when Additive Schwarz is used. There exist other methods capable of eliminating the non-converging modes in AS, such as the method of Restricted Additive Schwarz with Harmonic Overlap (RASHO), which was proposed by [1].

It is clear that the RAS and ASH preconditioners are transposes of each other when A is symmetric; one thus expects the two methods to converge at a similar rate. In the case where A is an M-matrix, [5] proved that RAS and ASH both converge as an iterative method. For the RAS method, [6] showed that the iterates produced are equivalent to those of the discretized parallel Schwarz method, regardless of the number of subdomains and whether cross points are present. On the other hand, to our knowledge no such interpretation exists for the ASH method. Our goal is to offer such an interpretation in the case where cross points are absent.

2 Assumptions and the Main Result

Before stating the main result, we make some assumptions that are algebraic manifestations of the fact that there are no cross points. The first one is self-evident based on the definition of the restriction operators R_k.

Assumption 1 (No cross points) For distinct i, j and l, we have

$$R_i R_j^T R_j R_l^T = 0. \tag{7}$$

The next pair of assumptions ensures that $\partial\Omega_j \setminus \partial\Omega$ are partitioned into r connected components, each of which must be a subset of only one $\tilde{\Omega}_i$ for some i (see Fig. 1).

Assumption 2 (Partition of internal boundaries) For all $i \neq j$, we must have

$$(R_i - \tilde{R}_i)(AR_j^T - R_j^T A_j) = 0, \tag{8a}$$

$$(R_i A - A_i R_i)(R_j^T - \tilde{R}_j^T) = 0. \tag{8b}$$

The two conditions are simply transposes of each other; hence, they will be satisfied simultaneously if A has a symmetric nonzero pattern. Also note that when $i = j$, the two relations are trivially satisfied: since $\tilde{R}_i = \tilde{R}_i R_i^T R_i$, we have

$$(R_i - \tilde{R}_i)(AR_i^T - R_i^T A_i) = \underbrace{R_i A R_i^T}_{A_i} - \underbrace{R_i R_i^T}_{I} A_i - \tilde{R}_i R_i^T \underbrace{R_i A R_i^T}_{A_i} + \tilde{R}_i R_i^T A_i = 0. \tag{9}$$

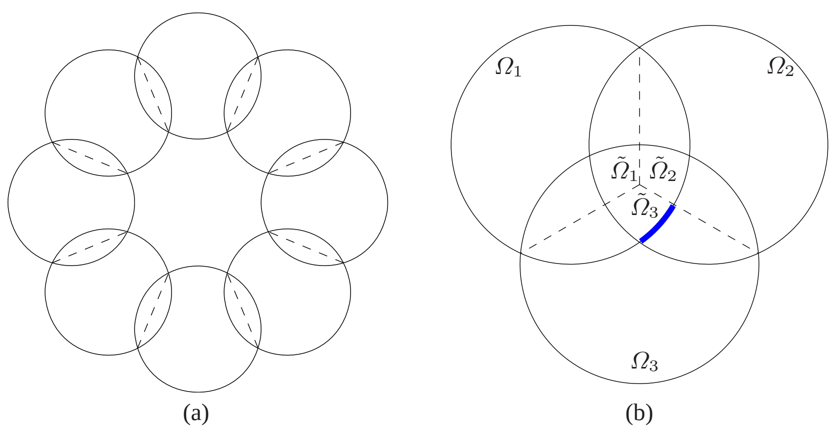


Fig. 1. Some examples of decompositions into subdomains, with solid lines delimiting Ω_i and dashed lines delimiting the $\tilde{\Omega}_i$. In (**a**), Assumptions 1–2 are satisfied, whereas in (**b**) they are not, because the stencil of $AR_1^T - R_1^T A_1$ along the thick portion of $\partial\Omega_1$ would extend into $\Omega_2 \setminus \tilde{\Omega}_2$, violating (8a).

We can interpret (8a) as follows. For any vector w over V_j, the vectors $AR_j^T w$ and $R_j^T A_j w$ agree inside V_j, but $AR_j^T w$ may have nonzero entries outside V_j (which $R_j^T A_j w$ cannot have). For a PDE, these entries are located along the boundary $\partial\Omega_j$. The assumption then says that these entries must correspond to nodes that are either inside $\tilde{\Omega}_i$ or completely outside Ω_i, as in Fig. 1(a). In Fig. 1(b), the thick portion of $\partial\Omega_1$ is inside $\Omega_2 \setminus \tilde{\Omega}_2$, violating (8a).

We are now ready to state our main result.

Theorem 1. *Suppose $\mathbf{U}^0 = 0$ and Assumptions 1 and 2 are satisfied. Then the iterates $\mathbf{U}^k$ of the ASH method are related to the iterates $\mathbf{v}_i^k$ of the discretized parallel Schwarz method with $\mathbf{v}_i^0 = 0$ and*

$$A_i \mathbf{v}_i^1 = \tilde{R}_i f, \tag{10}$$

$$A_i \mathbf{v}_i^k = R_i f - \sum_{j \neq i} A_{ij} \mathbf{v}_j^{k-1} \qquad (k \geq 2) \tag{11}$$

via the relation

$$\sum_{j=1}^{N} R_j^T \mathbf{v}_j^k = \mathbf{U}^k + \Big(\sum_j R_j^T R_j - I \Big) \mathbf{U}^{k-1}. \tag{12}$$

Remark Since $\sum_j R_j R_j^T - I$ is zero outside the overlap, it is clear that the iterates of ASH and Parallel Schwarz are identical outside the overlap, whereas inside they are linear combinations of the current and previous iterates.

3 Proof of the Main Result

We assume throughout this section that Assumptions 1 and 2 hold. We let

$$\mathbf{r}^k := f - A\mathbf{U}^k, \qquad \delta\mathbf{u}_j^k := A_j^{-1} \tilde{R}_j (f - A\mathbf{U}^k), \qquad \mathbf{u}_j^k := \sum_{l=0}^{k-1} \delta\mathbf{u}_j^l.$$

From (6) it is clear that $\mathbf{U}^k = \sum_j R_j^T \mathbf{u}_j^k$. We also define $\mathbf{v}_j^k$ such that $\mathbf{v}_j^0 = \mathbf{u}_j^0 = 0$, $\mathbf{v}_j^1 = \mathbf{u}_j^1$, and

$$\mathbf{v}_j^k = \mathbf{u}_j^k + \sum_{i \neq j} R_j R_i^T \mathbf{u}_i^{k-1} \qquad (k \geq 2). \tag{13}$$

The following properties are elementary and will be used repeatedly:

(i) $R_i R_i^T = I$ for all i,
(ii) $\tilde{R}_i^T R_i = R_i^T \tilde{R}_i = \tilde{R}_i^T \tilde{R}_i$ for all i,
(iii) $\tilde{R}_i R_i^T = R_i \tilde{R}_i^T = \tilde{R}_i \tilde{R}_i^T$ for all i,
(iv) $\sum_j \tilde{R}_j^T \tilde{R}_j = I$.

Lemma 1. *For all $k \geq 1$ and for all i, we have $(R_i - \tilde{R}_i)\mathbf{r}^k = 0$.*

Proof. Fix i and let $k \geq 0$. We calculate

$$(R_i - \tilde{R}_i)\mathbf{r}^k = R_i R_i^T (R_i - \tilde{R}_i)\mathbf{r}^k = R_i \Big[\sum_j \tilde{R}_j^T \tilde{R}_j R_i^T (R_i - \tilde{R}_i)\mathbf{r}^k\Big].$$

Since $R_j^T R_j$ and $R_i^T (R_i - \tilde{R}_i)$ are diagonal matrices, they commute, and hence

$$(R_i - \tilde{R}_i)\mathbf{r}^k = R_i \Big[\sum_j R_i^T (R_i - \tilde{R}_i)\tilde{R}_j^T \tilde{R}_j \mathbf{r}^k\Big] = R_i R_i^T (R_i - \tilde{R}_i) \sum_j \tilde{R}_j^T \tilde{R}_j \mathbf{r}^k.$$

Noting that $R_i R_i^T = I$ and $\tilde{R}_j^T \tilde{R}_j = R_j^T \tilde{R}_j$, we can rewrite $\tilde{R}_j \mathbf{r}^k$ as $A_j \delta \mathbf{u}_j^k$. Then (8a) and (9) together give

$$\begin{aligned}(R_i - \tilde{R}_i)\mathbf{r}^k &= (R_i - \tilde{R}_i) \sum_j A R_j^T \delta \mathbf{u}_j^k \\ &= (R_i - \tilde{R}_i) A(\mathbf{U}^{k+1} - \mathbf{U}^k) = (R_i - \tilde{R}_i)(\mathbf{r}^k - \mathbf{r}^{k+1}).\end{aligned}$$

Cancelling $(R_i - \tilde{R}_i)\mathbf{r}^k$ from both sides gives the required result.

Proof of Theorem 1 We first prove the relation (12). We have

$$\begin{aligned}\sum_j R_j^T \mathbf{v}_j^k &= \sum_j R_j^T \mathbf{u}_j^k + \sum_j R_j^T \sum_{l \neq j} R_j R_l^T \mathbf{u}_l^{k-1} \\ &= \mathbf{U}^k + \sum_j R_j^T R_j \Big(\sum_l R_l^T \mathbf{u}_l^{k-1} - R_j^T \mathbf{u}_j^{k-1}\Big) \\ &= \mathbf{U}^k + \Big(\sum_j R_j^T R_j\Big)\mathbf{U}^{k-1} - \sum_j R_j^T \mathbf{u}_j^{k-1} \\ &= \mathbf{U}^k + \Big(\sum_j R_j^T R_j - I\Big)\mathbf{U}^{k-1},\end{aligned}$$

as required. Now let $A_{i\Gamma} := R_i A - A_i R_i$ be the boundary operator. Multiplying both sides of (12) by $A_{i\Gamma}$ on the left gives

$$\begin{aligned}A_{i\Gamma} \sum_j R_j^T \mathbf{v}_j^k &= A_{i\Gamma}\mathbf{U}^k + A_{i\Gamma}\Big(\sum_j R_j^T R_j - I\Big)\mathbf{U}^{k-1} \\ &= A_{i\Gamma}\mathbf{U}^k + A_{i\Gamma}\Big(\sum_j (R_j^T R_j - \tilde{R}_j^T R_j)\Big)\mathbf{U}^{k-1} \\ &= A_{i\Gamma}\mathbf{U}^k + A_{i\Gamma}\Big(\sum_j (R_j - \tilde{R}_j)^T R_j\Big)\mathbf{U}^{k-1} = A_{i\Gamma}\mathbf{U}^k,\end{aligned}$$

since $A_{i\Gamma}(R_j - \tilde{R}_j)^T = 0$ for all i and j by (8b) and (9). When $i \neq j$, $A_{i\Gamma} R_j^T = A_{ij}$ by definition, and when $i = j$, we have

$$A_{i\Gamma} R_i^T = (R_i A - A_i R_i) R_i^T = R_i A R_i^T - A_i R_i R_i^T = A_i - A_i \cdot I = 0.$$

So in fact we have

$$A_{i\Gamma}\mathbf{U}^k = A_{i\Gamma}\sum_j R_j^T \mathbf{v}_j^k = \sum_{j\neq i} A_{ij}\mathbf{v}_j^k. \tag{14}$$

We now prove the main statement of the theorem. For $k \geq 0$, we have

$$\begin{aligned}
A_i\mathbf{v}_i^{k+1} &= A_i\mathbf{u}_i^{k+1} + A_i\sum_{j\neq i} R_iR_j^T\mathbf{u}_j^k \\
&= A_i\delta\mathbf{u}_i^k + A_i\mathbf{u}_i^k + A_i\sum_{j\neq i} R_iR_j^T\mathbf{u}_j^k \\
&= \tilde{R}_i(f - A\mathbf{U}^k) + A_i\sum_j R_iR_j^T\mathbf{u}_j^k \\
&= \tilde{R}_i f - \tilde{R}_iA\mathbf{U}^k + A_iR_i\mathbf{U}^k.
\end{aligned}$$

If $k = 0$, then all terms other than $\tilde{R}_i f$ vanish because $\mathbf{U}^0 = 0$; we have thus proved (10). We continue by assuming $k \geq 1$:

$$\begin{aligned}
A_i\mathbf{v}_i^{k+1} &= \tilde{R}_i f - \tilde{R}_iA\mathbf{U}^k + (R_iA - A_{i\Gamma})\mathbf{U}^k \\
&= \tilde{R}_i f + (R_i - \tilde{R}_i)A\mathbf{U}^k - A_{i\Gamma}\mathbf{U}^k \\
&= \tilde{R}_i f + (R_i - \tilde{R}_i)(f - \mathbf{r}^k) - A_{i\Gamma}\mathbf{U}^k \\
&= R_i f - \underbrace{(R_i - \tilde{R}_i)\mathbf{r}^k}_{=0} - \sum_{j\neq i} A_{ij}\mathbf{v}_j^k.
\end{aligned}$$

by Lemma 1 and (14), and (11) follows.

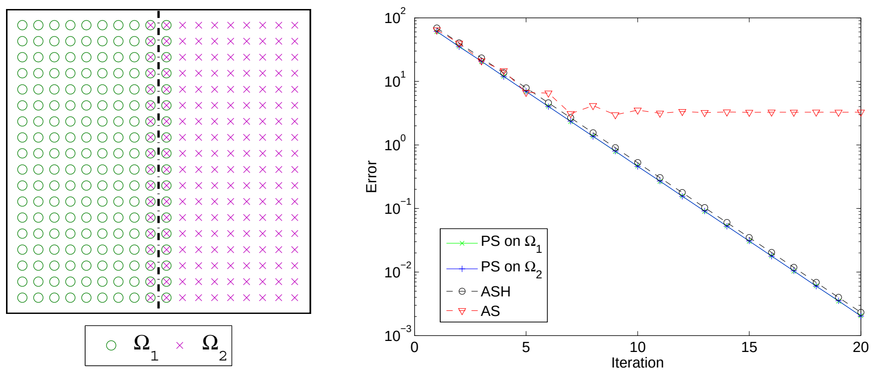
4 Convergence Rate

Given the close relationship between ASH and Parallel Schwarz, one would expect that the two methods converge at the same speed. This is true if the overlap subproblem is well posed.

Theorem 2 (cf. [7]). *Let R_o be the restriction operator onto the union of all overlaps, i.e., R_o is a full row-rank matrix such that $R_o^TR_o = \sum_j R_j^TR_j - I$. If $R_oAR_o^T$ is non-singular, then the ASH method (6) converges if and only if the parallel Schwarz method (10), (11) converges. In addition, when both methods converge, they do so at the same asymptotic rate.*

To illustrate this theorem, we solve Poisson's equation on the unit square with homogeneous Dirichlet boundary condition. We use a 20×20 grid, which is divided into two subdomains with a two-row overlap (Fig. 2(a)). We then solve this problem using (i) the discrete Parallel Schwarz method with Dirichlet boundary conditions, as defined in Eq. (3), (ii) the ASH method, and (iii) overlapping Additive Schwarz.

The convergence history for all three methods is shown in Fig. 2(b). We see that ASH converges linearly, unlike AS, which does not converge because of the overlap. In addition, the curves for Parallel Schwarz and ASH are very close to one another, and their slopes are asymptotically equal. We also see from Table 1 that the ratio of successive errors for parallel Schwarz alternates between 0.5737 and 0.5895 (which is typical for a two-subdomain problem), whereas ASH converges at a rate of $0.5815 = \sqrt{0.5737 \cdot 0.5895}$, the geometric mean. Thus, as iterative methods, ASH and parallel Schwarz converge at the same rate, as stated in Theorem 2.



(a) (b)

Fig. 2. (**a**) A two-subdomain decomposition. (**b**) Convergence behaviour for the parallel Schwarz, additive Schwarz and ASH methods.

Table 1. Error norms for Parallel Schwarz and ASH iterates.

	Parallel Schwarz				ASH	
	Error on Ω_1		Error on Ω_2		Error on Ω	
Iters	L_2-Norm	Ratio	L_2-Norm	Ratio	L_2-Norm	Ratio
1	60.4103		61.2159		69.0920	
2	35.5724	0.5888	35.0937	0.5733	40.1592	0.5812
3	20.4046	0.5736	20.6840	0.5894	23.3519	0.5815
4	12.0281	0.5895	11.8656	0.5737	13.5796	0.5815
5	6.9001	0.5737	6.9947	0.5895	7.8969	0.5815
6	4.0676	0.5895	4.0126	0.5737	4.5923	0.5815
7	2.3335	0.5737	2.3654	0.5895	2.6706	0.5815
8	1.3756	0.5895	1.3570	0.5737	1.5530	0.5815
9	0.7891	0.5737	0.7999	0.5895	0.9031	0.5815
10	0.4652	0.5895	0.4589	0.5737	0.5252	0.5815

5 Conclusions

We have shown that when the domain decomposition contains no cross points, the ASH method and parallel Schwarz have identical iterates outside the overlap. When both methods converge, they do so at the same asymptotic rate, provided the overlap subproblem is well posed. Thus, ASH is simply Lions' parallel method disguised as a preconditioner. The same conclusions hold when optimized transmission conditions are used; the proof is given in a separate article [7]. Such an insight can be used to estimate convergence rates of the optimized ASH method in cases where known results (e.g. [5]) do not apply. It would be interesting to see whether similar ideas can be used to relate RASHO to the parallel Schwarz method. Finally, a crucial assumption throughout this paper is that there must be no cross points. Since no such assumption is required in the interpretation of RAS [6], it would be instructive to study a problem with cross points to see whether similar results hold for arbitrary domain decompositions.

References

1. X.-C. Cai, M. Dryja, and M. Sarkis. Restricted additive Schwarz preconditioners with harmonic overlap for symmetric positive definite linear systems. *SIAM J. Numer. Anal.*, 41 (4):1209–1231 (electronic), 2003.
2. 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.
3. M. Dryja and O.B. Widlund. Some domain decomposition algorithms for elliptic problems. In *Iterative Methods for Large Linear Systems (Austin, TX, 1988)*, pp. 273–291. Academic Press, Boston, MA, 1990.
4. E. Efstathiou and M.J. Gander. Why restricted additive Schwarz converges faster than additive Schwarz. *BIT*, 43(suppl):945–959, 2003.
5. A. Frommer and D.B. Szyld. An algebraic convergence theory for restricted additive Schwarz methods using weighted max norms. *SIAM J. Numer. Anal.*, 39(2):463–479 (electronic), 2001.
6. M.J. Gander. Schwarz methods in the course of time. *ETNA*, 31:228–255, 2008.
7. F. Kwok. Optimized additive Schwarz with harmonic extension as a discretization of continous parallel Schwarz methods, SINUM, 2010.
8. P.-L. Lions. On the Schwarz alternating method. I. In *First International Symposium on Domain Decomposition Methods for Partial Differential Equations (Paris, 1987)*, pp. 1–42. SIAM, Philadelphia, PA, 1988.
9. H.A. Schwarz. Über einen Grenzübergang durch alternierendes Verfahren. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, 15:272–286, 1870.