A Discovery Algorithm for the Algebraic Construction of Optimized Schwarz Preconditioners

Amik St-Cyr[1] and Martin J. Gander[2]

[1] National Center for Atmospheric Research, 1850 Table Mesa Drive, Boulder CO 80305. amik@ucar.edu
[2] University of Geneva, 2-4 rue du Lièvre, CP 64 CH-1211 Genève Martin.Gander@math.unige.ch

Summary. Optimized Schwarz methods have been developed at the continuous level; in order to obtain optimized transmission conditions, the underlying partial differential equation (PDE) needs to be known. Classical Schwarz methods on the other hand can be used in purely algebraic form, which have made them popular. Their performance can however be inferior compared to that of optimized Schwarz methods. We present in this paper a discovery algorithm, which, based purely on algebraic information, allows us to obtain an optimized Schwarz preconditioner for a large class of numerically discretized elliptic PDEs. The algorithm detects the nature of the elliptic PDE, and then modifies a classical algebraic Schwarz preconditioner at the algebraic level, using existing optimization results from the literature on optimized Schwarz methods. Numerical experiments using elliptic problems discretized by Q_1-FEM, P_1-FEM, and FDM demonstrate the algebraic nature and the effectiveness of the discovery algorithm.

1 Introduction

Optimized Schwarz methods are based on transmission conditions between subdomains which are different from the classical Dirichlet conditions. The transmission conditions are adapted to the partial differential equation in order to lead to faster convergence of the method. Optimized transmission conditions are currently available for many types of scalar PDEs: for Poisson problems including a diagonal weight, see [3], for indefinite Helmholtz problems, see [4] and for advection reaction diffusion problems, see [2, 5]. More recently, it was shown that one can easily transform a classical algebraic Schwarz preconditioner such as restricted additive Schwarz (RAS) methods, into an optimized one, by simply changing some matrix entries in the local subdomain matrices, see [6]. However, in order to know what changes to make, one needs to know what the underlying PDE is, and thus the optimized RAS method has so far not become a black box solver, in contrast to classical RAS. We propose in this paper a discovery algorithm which is able to extract all the required information from the given matrix, and thus to make optimized RAS into a black box solver, for discretizations of the elliptic partial differential equation

$$\nu \Delta u + \mathbf{a} \cdot \nabla u - \eta u = f \quad \text{in } \Omega, \tag{1}$$

with suitable boundary conditions. Here, ν, $\mathbf{a}$ and η are all functions of $\mathbf{x}$, and by a suitable choice, we can handle all elliptic PDEs for which currently optimized transmission conditions are known. In this paper, we focus on Robin transmission conditions, which are of the form

$$(\partial_n + p)u_i = (\partial_n + p)u_j \tag{2}$$

at the interface between subdomain i and j. In general, the optimized scalar parameter p depends on the local mesh size h, the overlap width Ch, the interface diameter L, and the coefficients of the underlying PDE, i.e. $p = p(h, Ch, L, \nu, \mathbf{a}, \eta)$.

A discretization of (1) leads to a linear system of equations of the form

$$A\mathbf{u} = (K + S + M)\mathbf{u} = \mathbf{f}, \tag{3}$$

where K is the stiffness matrix, $K\mathbb{I} = 0$ with $\mathbb{I}$ the vector of all ones, S is skew-symmetric from the advection term of the PDE, $S\mathbb{I} = 0$, and M is a mass matrix from the η term in the PDE. For a black box preconditioner, only the matrix A is given, and the decomposition (3) needs to be extracted automatically, in addition to the mesh size and the diameter of the interface, in order to use the existing formulas for the optimized parameter p in a purely algebraic fashion. We also need to extract a normal derivative for (2) algebraically.

In what follows, we assume that we are given the restriction operators R_j and $\tilde{R}_j$ of a restricted additive Schwarz method for the linear system (3), and the associated classical subdomain matrices $A_j := R_j A R_j^T$. The restricted additive Schwarz preconditioner for (3) would then be $\sum_j \tilde{R}_j^T A_j^{-1} R_j$, and the optimized restricted additive Schwarz method is obtained by slightly modifying the local subdomain matrices A_j at interface nodes, in order to obtain $\tilde{A}_j$, which represent discretizations with Robin, instead of Dirichlet boundary conditions, see [6]. In order for this replacement to lead to an optimized Schwarz method, an algebraic condition needs to be satisfied, which requires a minimal overlap and a certain condition at cross-points; for details, see [6].

2 Discovery Algorithm

We now describe how appropriate matrices $\tilde{A}_j$ for an optimized Schwarz preconditioner can be generated algebraically for discretizations of the PDE (1) given in matrix form (3). There are three steps in the algorithm to generate the modified $\tilde{A}_j$:

1. Interface detection.
2. Extraction of physical and discretization parameters.
3. Construction of the optimized transmission condition.

2.1 Interface Detection

For a matrix $A \in \mathbb{R}^{N\times N}$, let $\mathcal{S}(A)$ be its **canonical index** set, i.e. the set of integers going from 1 to N, and let $\mathbf{c} \in \mathbb{N}^N$ be its **multiplicity**, i.e. c_i contains the total number of non-zero entries in the corresponding row of A. For a subdomain decomposition given by restriction matrices R_j, let the matrix $A_j = R_j A R_j^T$ have the canonical index $\mathcal{S}(A_j)$ with multiplicity $\mathbf{c}_j$. Then the set of indices $\mathcal{B}(A_j)$ representing the interfaces of subdomain j corresponds to the non-zero entries of $\mathbf{c}_j - R_j\mathbf{c}$, where $\mathbf{c}$ is the multiplicity of A. The set $\mathcal{B}(A_j)$ indicates which rows of the matrix A_j need to be modified in order to obtain $\tilde{A}_j$ for an optimized preconditioner.

2.2 Extraction of Physical and Discretization Parameters

We start by guessing the decomposition (3) of A by computing

$$S = \frac{1}{2}(A - A^T), \quad M = \mathrm{diag}(A\mathbb{I}), \quad K = \frac{1}{2}(A + A^T) - M. \tag{4}$$

This approach does not necessarily find the same parts one would obtain by knowing the discretization: for example we can only guess a lumped mass matrix and not discover an upwind scheme. The parts we obtain however correspond to a decomposition relevant for the problem.

Definition 1. *Using* (4) *for each interface node i, we define the*

- **local viscosity indicator:** $\nu_i := \sum_j |K_{ij}|/(2(c_i - 1))$,
- **local advection indicator:** $\alpha_i := \max_j |S_{ij}|$,
- **local zeroth order term indicator:** $\eta_i := sign(K_{ii})M_{ii}$.

These three indicators are enough to reveal the PDE-like properties of the matrix at the interface:

1. $\nu_i > 0$, $\eta_i = 0$ and $\alpha_i = 0$: Poisson equation.
2. $\nu_i > 0$, $\eta_i > 0$ and $\alpha_i = 0$: Poisson equation with weight, or implicit heat equation, see [3].
3. $\nu_i > 0$, $\eta_i < 0$ and $\alpha_i = 0$: indefinite Helmholtz equation, see [4].
4. $\nu_i > 0$, $\eta_i = 0$ and $\alpha_i \neq 0$: advection-diffusion equation, see [5].
5. $\nu_i > 0$, $\eta_i > 0$ and $\alpha_i \neq 0$: implicit advection-diffusion equation, see [2].

In the other cases (except if (1) has been multiplied by minus one, which can also be treated similarly), optimized transmission conditions have not yet been analyzed, and we thus simply apply RAS for that particular row. We next have to estimate the local mesh size h_i. The indicators above contain in general this information, for example for a standard five point finite difference discretization of $\eta - \nu\Delta$, we would obtain $\nu_i = \frac{\nu}{h^2}$, but we cannot detect the mesh size h separately without further information. In addition, the algebraic equations could have been scaled by h, or h^2, or any other algebraically useful diagonal scaling. However, in general, the optimized parameter p is also scaled accordingly: the analytical formulas for p all contain the size of the interface L and the mesh spacing h in a certain relation. Since the latter are both interface related quantities, we use the trace of the discovered Laplacian in order to estimate them.

Definition 2. *The relevant local mesh size at point $i \in \mathcal{B}(A_k)$ is*

$$h_i^k \approx \Big(\sum_j |(K_k)_{ij}|/(2c_i - 1)\Big)^{-1/2}, \tag{5}$$

where K_k is the trace of the discovered Laplacian and c_i is its associated multiplicity.

We finally need to estimate the interface diameter L of each interface. To this end, we need to discover the dimension of the problem is. This can be achieved using the ratio of interior nodes versus interface nodes in each subdomain. Solving the equation (# denotes the cardinality of the set)

$$\#\mathcal{B}(A_k) = (\#\mathcal{S}(A_k))^{\frac{d-1}{d}} \tag{6}$$

for d in each subdomain, we obtain an estimate for the dimension denoted by $\bar{d}_k$. We accept a fractional dimension because it is not uncommon for example for three dimensional domains to represent thin shells.

Definition 3. *The diameter of each interface $L = L_{jk}$ between subdomain j and k is estimated for 2d and 3d problems by*

$$L_{jk} := (\#\mathcal{B}_k(A_j))^{-\frac{\bar{d}_k - 2}{\bar{d}_k - 1}} \sum_{i \in \mathcal{B}_k(A_j)} h_i^j, \tag{7}$$

where $\mathcal{B}_k(A_j)$ denotes the interface nodes of subdomain j with subdomain k.

2.3 Construction of the Optimized Transmission Condition

In order to construct an algebraic approximation to the Robin transmission condition (2), we need a normal derivative approximation. Suppose that row i was identified as an interface node. For this row, we can partition the indices denoting the position in the row with non-zero elements into three sets:

1. the diagonal entry denoted by set $\{i\}$,
2. the off-diagonal entries that are not involved in the interface denoted by $\mathcal{I}_i$ for interior,
3. the off-diagonal entries that are on the interface, denoted by $\mathcal{F}_i$.

These indices take values in the set of integers indexing the full matrix A, but in order to simplify what follows, we re-label these indices from 1 to J. Let $\{\mathbf{x}_j\}_{j=1}^J$ be a set of arbitrary spatial points with associated scalar weights $\{w_j\}_{j=1}^J$, and let $\delta\mathbf{x}_{ji} = \mathbf{x}_j - \mathbf{x}_i$. In order to define a normal derivative at the point $\mathbf{x}_i$, we assume that

$$\|\delta \mathbf{x}_{ji}\| \leq h, \quad \text{and} \quad \sum_{j \in \mathfrak{F}_i} w_j \delta \mathbf{x}_{ji} = O(h^2), \tag{8}$$

and we define an approximate unit outward normal vector $\mathbf{n}$ at $\mathbf{x}_i$ by

$$\mathbf{n} = -\sum_{j \in \mathfrak{I}_i} w_j \delta \mathbf{x}_{ji} / \| \sum_{j \in \mathfrak{I}_i} w_j \delta \mathbf{x}_{ji} \|. \tag{9}$$

A situation might arise were the set $\mathfrak{I}_i$ is empty. In this case the connectivity of the matrix must be exploited in order to find a second set of points connected to the points in $\mathfrak{F}_i$. By removing the points lying on any boundary a new set $\mathfrak{I}_i$ can be generated. This procedure can be repeated until the set is non-empty. Let the vectors τ_k, $k = 1, ..., d-1$ be an orthonormal basis spanning the tangent plane implied by $\mathbf{n}$ at $\mathbf{x}_i$, i.e. $\mathbf{n} \cdot \tau_k = 0$.

Proposition 1. *If conditions (8) are satisfied, and in addition $w_i = -\sum_{j \neq i} w_j$, then for a sufficiently differentiable function u around $\mathbf{x}_i$, we have*

$$-\frac{\sum_{j=1}^{J} w_j u(\mathbf{x}_j)}{\| \sum_{j \in \mathfrak{I}_i} w_j \delta \mathbf{x}_{ji} \|} = \mathbf{n} \cdot \nabla u(\mathbf{x}_i) + O(h). \tag{10}$$

Proof. Using a Taylor expansion, and the sum condition on w_j, we obtain

$$\begin{aligned}
\sum_{j=1}^{J} w_j u(\mathbf{x}_j) &= \sum_{j=1}^{J} w_j (u(\mathbf{x}_i) + \delta \mathbf{x}_{ji} \cdot \nabla u(\mathbf{x}_i) + O(h^2)) \\
&= (w_i + \sum_{j \neq i} w_j) u(\mathbf{x}_i) + \sum_{j \neq i} w_j \delta \mathbf{x}_{ji} \cdot \nabla u(\mathbf{x}_i) + O(h^2) \\
&= \nabla u(\mathbf{x}_i) \cdot \sum_{j \neq i} w_j \delta \mathbf{x}_{ji} + O(h^2).
\end{aligned}$$

Now using the second condition in (8), and the decomposition of the gradient into normal and tangential components, $\nabla u(\mathbf{x}_i) = u_0 \mathbf{n} + \sum_{k=1}^{d-1} u_k \tau_k$, we get

$$\begin{aligned}
\sum_{j=1}^{J} w_j u(\mathbf{x}_j) &= \nabla u(\mathbf{x}_i) \cdot \sum_{j \in \mathfrak{I}_i} w_j \delta \mathbf{x}_{ji} + O(h^2), \\
&= u_0 \mathbf{n} \cdot \sum_{j \in \mathfrak{I}_i} w_j \delta \mathbf{x}_{ji} + \sum_{k=1}^{d-1} u_k \tau_k \cdot \sum_{j \in \mathfrak{I}_i} w_j \delta \mathbf{x}_{ji} + O(h^2).
\end{aligned}$$

The double sum vanishes, since the sum on j equals $\mathbf{n}$ up to a multiplicative constant and $\mathbf{n} \cdot \tau_k = 0$. Now using the definition of the approximate normal $\mathbf{n}$, and using that $u_0 = \mathbf{n} \cdot \nabla u(\mathbf{x}_i)$, leads to the desired result.

Note that the formula for the approximation of the normal derivative (10) does not need the explicit computation of a normal or tangential vector at the interface.

Definition 4. *An approximation $A_i^{\mathfrak{I}}$ to the normal derivative is generated from matrix A at a line i having a non-empty set $\mathfrak{I}_i$ by performing (in this order): $a_{ii} = 0$, $a_{ij} = 0$ for $j \in \mathfrak{F}_i$, $a_{ii} = -\sum_{j \neq i} a_{ij}$.*

There are also optimized Schwarz methods with higher order transmission conditions, which use tangential derivatives at the interfaces. Such methods involve for the Poisson equation the Laplace-Beltrami operator at the interface, see [3], or more generally the remaining part of the partial differential operator, see for example [2] or [1]. If we want to use higher order transmission conditions also at the algebraic level, we need to extract the corresponding discretization stencil at the interface as well. This stencil has the same dimensions as $A_i^{\mathfrak{I}}$ and contains all the coefficients lying in $\mathfrak{F}$.

Definition 5. *The complement of $A_i^{\mathfrak{I}}$ is the matrix $A_i^{\mathfrak{F}}$ generated from matrix A at a line i having a non-empty set $\mathfrak{I}_i$ by performing (in this order): $a_{ii} = 0$, $a_{ij} = 0$ for $j \in \mathfrak{I}_i$, $a_{ii} = -\sum_{j \neq i} a_{ij}$.*

The matrices used to detect the nature of the PDE cannot be employed in the construction of the optimized transmission operator, because they might be rank deficient. The detected mass matrix could be employed, if one is present. However, for more generality, we choose the diagonal mass matrix for an interface node i as $D_i = \mathbf{h}_i^2 A_{ii}$: its sign is correct for the elliptic operator for the definite case; for the indefinite case, it needs to be multiplied by -1. From Definitions 4 and 5, and the first of the assumptions (8), we see that the normal and complement matrices are both $O(1)$ (the complement is a difference of 2 normal derivatives at the interface divided by h). However, the entries in the matrix are proportional to $1/h^2$. Thus the normal derivative needs a scaling factor of $1/h$. Consequently, both the mass and complement matrices are divided by h.

The algebraic representation of the transmission condition for domain k in the matrix is then given by

$$T_k \equiv \operatorname{diag}\left(\frac{\mathbf{p}_k}{\mathbf{h}_k}\mathbf{D}_k\right) + A_k^{\mathfrak{I}} + \operatorname{diag}\left(\frac{\mathbf{q}_k}{\mathbf{h}_k}\right) A_k^{\mathfrak{F}}, \tag{11}$$

where the division of a vector by a vector is component wise.

3 Numerical Experiments

We consider three different discretizations, FDM, Q_1-FEM, and P_1-FEM applied to an *a priori unknown* positive definite Helmholtz operator. In all cases the solution is $u(x,y) = \sin(\pi x)\sin(\pi y)$ on the domain $(0,1)\times(0,1)$ with Dirichlet boundary conditions. We present results for the iterative form of the algorithm and its acceleration by GMRES. For all cases a starting vector containing noise in $(0,1)$ was employed.

In the first set of experiments, see Table 1, a square corner $(0,1/2)\times(0,1/2)$ is considered as one of the two subdomains, and the L-shaped rest is the other subdomain. These domains are uniformly discretized for the first experiment by a finite difference method, and for the second experiment by a Q1 finite element method. In each experiment, an overlap of two mesh sizes is added. We can see from Table 1 that the optimized Schwarz methods generated purely algebraically from the global

h		1/8	1/16	1/32	1/64	1/128	1/256
Q_1-FEM:							
iterative:	RAS	6	15	32	67	136	275
iterative:	O0	8	14	23	33	48	65
iterative:	O2	8	13	19	24	30	36
GMRES:	RAS	4	6	10	13	19	26
GMRES:	O0	5	7	9	12	15	19
GMRES:	O2	5	7	9	11	13	16
FDM:							
iterative:	RAS	7	16	32	67	136	275
iterative:	O0	8	16	27	42	63	90
iterative:	O2	8	15	24	33	43	53
GMRES:	RAS	4	8	11	17	24	35
GMRES:	O0	4	6	8	10	14	17
GMRES:	O2	5	6	9	10	12	14

Table 1. Structured corner domain: the same algebraic algorithm was employed

matrix perform significantly better than the classical Schwarz method, both for the iterative and the GMRES accelerated versions.

We next show an example of a triangularly shaped decomposition of the square into two subdomains, as shown in Fig. 1. The discretization is now performed using

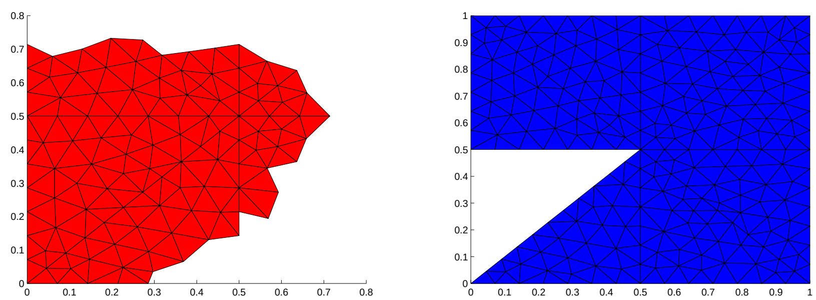

Fig. 1. Left panel: triangularly shaped domain Ω_1 extended by an overlap of size $3h$. Right panel: non-convex domain Ω_2.

an unstructured triangular mesh and a P1 finite element discretization. We show in Table 2 again a comparison of the iteration counts for the classical and various optimized Schwarz methods, obtained purely at the algebraic level with the discovery algorithm. We observe again that substantial gains are possible.

Triangles		534	2080	8278
P_1-FEM:				
iterative:	RAS	16	35	71
iterative:	O0	12	22	32
iterative:	O2	11	18	24
GMRES:	RAS	11	14	20
GMRES:	O0	8	11	15
GMRES:	O2	8	11	14

Table 2. Unstructured corner domain: the same algebraic algorithm was employed

Acknowledgement. The National Center for Atmospheric Research is sponsored by the National Science Foundation. The first author's travel to Geneva, and the second author were partly supported by the Swiss National Science Foundation Grant 200020-1 17577/1.

References

[1] Bennequin, D., Gander, M.J., Halpern, L.: A homographic best approximation problem with application to optimized Schwarz waveform relaxation. *Math. Comp.*, 78:185–223, 2009.

[2] Dubois, O.: *Optimized Schwarz Methods for the Advection-Diffusion Equation and for Problems with Discontinuous Coefficients*. PhD thesis, McGill University, June 2007.

[3] Gander, M.J.: Optimized Schwarz methods. *SIAM J. Numer. Anal.*, 44(2):699–731, 2006.

[4] Gander, M.J., Magoulès, F., Nataf, F.: Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM J. Sci. Comput.*, 24(1):38–60, 2002.

[5] Japhet, C., Nataf, F., Rogier, F.: The optimized order 2 method. Application to convection-diffusion problems. *Future Generation Computer Systems FUTURE*, 18(1):17–30, 2001.

[6] St-Cyr, A., Gander, M.J., Thomas, S.J.: Optimized multiplicative, additive and restricted additive Schwarz preconditioning. *SIAM J. Sci. Comput.*, 29(6):2402–2425, 2007.