Convergence Behavior of a Two-Level Optimized Schwarz Preconditioner

Olivier Dubois[1] and Martin J. Gander[2]

[1] IMA, University of Minnesota, 207 Church St. SE, Minneapolis, MN 55455
dubois@ima.umn.edu
[2] Section de mathématiques, Université de Genève, 2-4 Rue du Lièvre, CP 64, 1211 Genève martin.gander@math.unige.ch

Summary. Optimized Schwarz methods form a class of domain decomposition algorithms in which the transmission conditions are optimized in order to achieve fast convergence. They are usually derived for a model problem with two subdomains, and give efficient transmission conditions for the local coupling between neighboring subdomains. However, when using a large number of subdomains, a coarse space correction is required to achieve parallel scalability. In this paper we demonstrate with a simple model problem that a two-level optimized Schwarz preconditioner is much more effective than a corresponding two-level Restricted Additive Schwarz preconditioner. The weak dependence on the mesh size is retained from the one-level method, while gaining independence on the number of subdomains. Moreover, the best Robin transmission condition is well approximated by using the analysis from the two subdomain case, under Krylov acceleration.

1 Introduction

In the last ten years, a new class of domain decomposition methods has emerged and has been developed: Optimized Schwarz Methods (OSM). The main idea is to replace the Dirichlet transmission conditions of the classical Schwarz iteration by Robin or higher order conditions, and then optimizing the free parameters in these conditions to obtain the best convergence. In addition to providing fast convergence, the optimized transmission conditions allow us to use very small overlapping regions (as well as no overlap), causing only a weak dependence of the convergence on the mesh size. Optimized Schwarz methods were first introduced in Japhet [1998] for the advection-diffusion equation, and then studied for a variety of problems, for example in Gander et al. [2002], Gander [2006] and Dubois [2007].

In all of these studies, the analysis of the convergence is done only for a model problem with *two* infinite or rectangular subdomains, for which a Fourier transform can be applied, thus making possible the explicit optimization of the transmission conditions. In more practical situations with many subdomains, numerical experiments show that such optimized transmission conditions lead to efficient local coupling between neighboring subdomains, but there are no theoretical estimates on the convergence rate of the Schwarz iteration in that case.

It is well-known that domain decomposition techniques are not scalable with the number of subdomains, unless a global mean of communication between the subdomains is incorporated. This is often achieved by a coarse space or coarse grid correction. For optimized Schwarz methods, it is often claimed that the same coarse grid corrections applied to "classical" Schwarz methods can be employed to remove the dependence on the number of subdomains, but little numerical evidence of this fact are published, and no theoretical results are yet available. In the early paper of Japhet et al. [1998], two coarse space preconditioners were proposed that improve considerably the convergence rate of the Schwarz iteration as we increase the number of subdomains; these preconditioners are specific to non-overlapping decompositions, where the problem is first reformulated as an interface problem.

In this paper, we consider the overlapping Optimized Restricted Additive Schwarz (ORAS) preconditioner from St-Cyr et al. [2007], and apply a standard coarse grid correction to obtain a two-level method (see Toselli and Widlund [2005] and references therein for the two-level Additive Schwarz preconditioner). We verify with experiments that the weak scaling with respect to the mesh size h is preserved, in agreement with the theory on OSM, and that we gain independence on the number of subdomains for generous overlap. Moreover, we investigate whether the formulas for the optimized parameters, derived in the case of two subdomains, still provide good approximations for the best parameters for the two-level ORAS preconditioner when applied to many subdomains.

The paper is organized as follows. In Section 2, we introduce a simple model problem and describe the one-level and two-level preconditioners under consideration. We also discuss some practical implications of an algebraic condition required in the analysis of St-Cyr et al. [2007]. In Section 3, we present several numerical results for the two-level preconditioners, in different scaling scenarios. Finally, in Section 4, we find the best Robin parameter numerically and compare it with the values provided by the formulas, which were derived in the two subdomain case.

2 Domain Decomposition Preconditioners

We consider the simple positive definite elliptic problem $-\Delta u = f$, on the unit square, with homogeneous Dirichlet boundary conditions. We use finite differences to discretize this problem on a uniform grid with $n+2$ points in each dimension ($h := 1/(n+1)$). This leads to a linear system $A\mathbf{u} = \mathbf{b}$. The domain of computation is decomposed into $M_x \times M_y$ rectangular subdomains in the natural way, as illustrated by Figure 1.

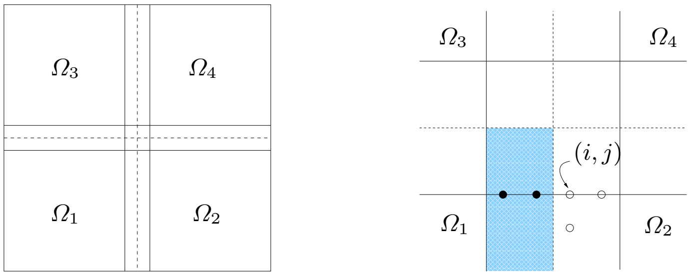

Fig. 1. Example of a uniform domain decomposition into 4 overlapping subdomains. A zoom near the crosspoint is shown on the right.

2.1 One-level preconditioners

Let $\{\tilde{\Omega}_j\}$ denote a non-overlapping partition of the unknowns. By extending these sets with $\frac{C-1}{2}$ layers of unknowns, we get an overlapping decomposition $\{\Omega_j\}$ with a physical overlap of width $L = Ch$. Let $\tilde{R}_j$ and R_j be the restriction operators on the subsets $\{\tilde{\Omega}_j\}$ and $\{\Omega_j\}$ respectively, and $A_j := R_j A R_j^T$ be the induced local matrices.

We consider the following two preconditioners

$$P_{RAS}^{-1} := \sum_{j=1}^{M} \tilde{R}_j^T A_j^{-1} R_j, \qquad P_{ORAS}^{-1} := \sum_{j=1}^{M} \tilde{R}_j^T \tilde{A}_j^{-1} R_j.$$

The first one is the *Restricted Additive Schwarz* (RAS) preconditioner of Cai and Sarkis [1999], while the second denotes the *Optimized Restricted Additive Schwarz* (ORAS) preconditioner, recently introduced in St-Cyr et al. [2007], where the local matrices are modified to implement optimized Robin interface conditions. Note that with this process, the *physical* overlap is reduced by two mesh layers. For example, a physical overlap of $L = 3h$ for RAS will correspond to an overlap of $\tilde{L} = h$ for ORAS. Here, we will use L to denote exclusively the physical overlap corresponding to the RAS preconditioner as a reference. In the present context, we will refer to the case of *minimal overlap* when choosing the width of the overlapping region to be $L = 3h$. In the case of *generous overlap*, we keep the overlap width proportional to the subdomain size, $L = CH$ (where C is chosen in such a way that we always have $L \geq 3h$).

We often think of an optimized Schwarz method as an iteration-by-subdomain of the form

$$\tilde{A}_j \mathbf{u}_j^{n+1} = \mathbf{f}_j + \sum_{j=1}^{M} \tilde{B}_{jk} \mathbf{u}_k^n, \quad j = 1, 2, \ldots, M, \tag{1}$$

whereas in this paper we wish to utilize a stationary iterative method with preconditioner P_{ORAS}^{-1}

$$\mathbf{u}^{n+1} = \mathbf{u}^n + (\sum_{j=1}^{M} \tilde{R}_j^T \tilde{A}_j^{-1} R_j)(\mathbf{f} - A\mathbf{u}^n). \quad (2)$$

It is shown in St-Cyr et al. [2007] that the iterations (1) and (2) are ***equivalent*** under two conditions, one of which is an *algebraic condition* given by

$$\tilde{B}_{jk} R_k \tilde{R}_m^T = 0, \quad \text{for any } j, \text{ and } m \neq k. \quad (3)$$

2.2 Interpretation of the algebraic condition

The algebraic condition (3) relates the discretization of the transmission conditions with the overlapping (Ω_j) and non-overlapping ($\tilde{\Omega}_j$) domain decompositions. For example, in the case of a decomposition into strips and a 5-point stencil discretization of the Laplacian, condition (3) requires that the overlap be at least three mesh layers wide, i.e $L \geq 3h$.

Now, if the domain decomposition has cross-points, how do we interpret the algebraic condition? To gain more insight, let us consider a simple example with 4 subdomains, labeled as in Figure 1. In condition (3), let $m = 1$, $k = 2$ and $j = 4$. Figure 1 also shows a blow-up of the region near the cross-point. After having applied the combination $R_2 \tilde{R}_1^T$ to a vector, the only possible nonzero entries are located in the shaded region. Then, condition (3) imposes that the application of $\tilde{B}_{42}$ should not depend on those nodes. Consider in particular the node (i, j) from Figure 1: it lies on the boundary of Ω_4 and inside $\tilde{\Omega}_2$, hence the operator $\tilde{B}_{42}$ needs to extract a transmission condition there.

If we use a standard finite difference discretization of the normal derivative at (i, j) which is second order accurate ($O(h^2)$), we will need the "illegal node" on the left of (i, j), hence violating the algebraic condition. In practice, we have observed that this causes very slow convergence of iteration (2). To avoid this problem, we have implemented instead a first-order accurate approximation to the normal derivative (using a one-sided finite difference), for which the algebraic condition is satisfied. In that case, modifying the local matrices A_i to $\tilde{A}_i$ also becomes a much simpler task: only diagonal entries for the nodes lying on the boundary of Ω_i need to be modified.

2.3 Two-level preconditioners

To introduce a coarse space correction, we proceed as follows. The (non-overlapping) domain decomposition induces a natural coarse mesh with nodes (iH_x, jH_y), where $H_x = 1/M_x$, $H_y = 1/M_y$. We define P_0 to be the *bilinear interpolation* from the coarse to the fine mesh. This induces a coarse matrix using the relation $A_0 := P_0^T A P_0$. We choose to apply the coarse space correction *sequentially*, after the parallel subdomain solves. In the case of a stationary iterative method, preconditioned by a two-level Restricted Additive Schwarz preconditioner (RAS2), we get the iterates

$$\mathbf{u}^{k+\frac{1}{2}} = \mathbf{u}^k + \sum_{j=1}^{M} \tilde{R}_j^T A_j^{-1} R_j (\mathbf{b} - A\mathbf{u}^k),$$
$$\mathbf{u}^{k+1} = \mathbf{u}^{k+\frac{1}{2}} + P_0 A_0^{-1} P_0^T (\mathbf{b} - A\mathbf{u}^{k+\frac{1}{2}}).$$

The same coarse grid component can be added on top of ORAS to get a two-level Optimized Restricted Additive Schwarz preconditioner (ORAS2). To obtain faster convergence, we can apply a GMRES iteration on the corresponding preconditioned linear systems instead.

In this paper, we will experiment only with the optimized one-sided Robin conditions, using the asymptotic formula valid for small values of h (when $L = Ch$), namely $p^* \approx 2^{-1/3} k_{min}^{2/3} L^{-1/3}$ (see Gander [2006]). In this formula, the minimum frequency in the min-max problem is chosen to be $k_{min} = \pi$ for the one-level preconditioner, and $k_{min} = \pi/H$ for the two-level preconditioner, in which case $p^* = O(H^{-2/3})$. The idea behind this choice is that the coarse grid correction should take care of the frequencies below π/H.

3 Numerical Results

In the following results, we solve the preconditioned linear system in two ways. First, we use a stationary iterative method, with right hand side $f \equiv 0$ and random initial guess $\mathbf{u}^{(0)}$, and check the convergence to 0 in the relative ℓ^∞-norm, with tolerance 10^{-6}. Alternatively, we solve the linear system using a preconditioned GMRES (not restarted), with random right-hand side and zero initial guess, with tolerance 10^{-8} on the preconditioned residual.

Our parallel implementation is based on the PETSc library (Balay et al. [2001]). For the solution of local and coarse problems (i.e. applications of A_j^{-1}), we precompute a full Cholesky factorization.

3.1 Dependence on h

Let us first fix the number of subdomains to $4 \times 4 = 16$ subdomains, use a minimal overlap $L = 3h$, and decrease the fine mesh size. We average the iteration numbers over 25 different random vectors. Figure 2 illustrates that we obtain the theoretical asymptotic convergence; in fact, with GMRES acceleration, we seem to get a slightly better convergence factor than the expected $1 - O(h^{1/6})$.

More importantly, observe that when using the stationary iterative method, the ORAS2 preconditioner (for which we have chosen $k_{min} = \pi/H$) takes *more* iterations than the one-level ORAS preconditioner (for which $k_{min} = \pi$). This indicates that the Robin parameter with $k_{min} = \pi/H$ does not give the appropriate value; we will confirm this in Section 3. On the other hand, this choice of parameter appears to yield good convergence under GMRES acceleration.

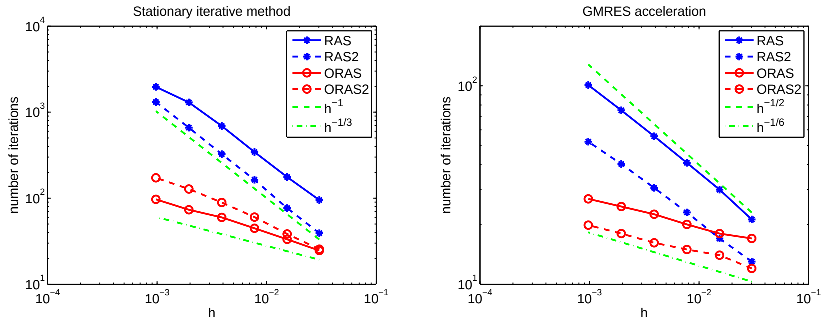

Fig. 2. Example with 16 subdomains, minimal overlap $L = 3h$, and decreasing h.

3.2 Dependence on H, with generous overlap

We now fix the fine mesh size with $n = 512$, and increase the number of subdomains while keeping the overlap proportional to H (as much as possible). The following are results obtained for only one instance of a random vector. Table 1 contains the number of iterations; with the GMRES method, the convergence of ORAS2 appears to be independent of the number of subdomains as expected. Also, we can again observe that the performance of ORAS2 with the choice $k_{min} = \pi/H$ is not acceptable for the stationary iterative method, when compared to the ORAS preconditioner.

$\mathbf{M_x} \times \mathbf{M_y}$ **(L)**	2×2 $(9h)$	4×4 $(5h)$	8×8 $(3h)$	2×2 $(9h)$	4×4 $(5h)$	8×8 $(3h)$
	Stationary iterative method			Preconditioned GMRES		
RAS	306	739	> 2000	33	60	99
ORAS	27	54	136	16	22	32
RAS2	206	395	533	28	32	31
ORAS2	33	76	174	14	16	17

Table 1. Number of iterations with increasing the number of subdomains, while keeping the overlap proportional to H.

3.3 A weak scalability test

Suppose now that each processor handles a problem of fixed size, in this case 192×192, and let's increase the number of processors. In other words, we keep H/h constant, and always use a minimal overlap $L = 3h$. Table 2 clearly shows that the ORAS2 preconditioner provides significant improvement on the convergence over RAS2 (the difference would become even greater if we increase the ratio H/h).

$\mathbf{M_x} \times \mathbf{M_y}$	2×2	4×4	6×6	8×8	9×9
no. of unknowns	$147,456$	$589,824$	$1,327,104$	$2,359,296$	$2,985,984$
	Stationary iterative method				
RAS2	439	1082	1528	1557	1798
ORAS2	325	316	323	332	324
	Preconditioned GMRES				
RAS2	40	47	48	48	48
ORAS2	18	20	21	21	21

Table 2. Number of iterations for a weak scaling experiment.

4 Best Robin Parameter

Does the asymptotic formula for the optimized Robin parameter give a good approximation to the best parameter value for the ORAS2 preconditioner? We provide an answer to this question by minimizing the number of iterations taken by the preconditioned iterative method, to obtain the best Robin parameter numerically. Figure 3 show the results for a fixed problem with mesh size $h = 1/64$ and $4 \times 4 = 16$ subdomains ($H = 1/4$). Figure 4, on the other hand, plots the behavior of the best Robin parameter as the number of subdomains is increased. We can make two interesting remarks, which are in agreement with our previous experiments:

1. For the stationary iterative method, the best p for ORAS and ORAS2 are very close, and the best convergence appears to be the same in both cases. The asymptotic formula for the optimized Robin parameter with $k_{min} = \pi/H$ gives values far from the best possible.
2. On the other hand, in the case of preconditioned GMRES, the optimized Robin parameters with $k_{min} = \pi$ and $k_{min} = \pi/H$ respectively are very close to the best parameter values, and the convergence of the two-level preconditioner offers significant improvement over the one-level version.

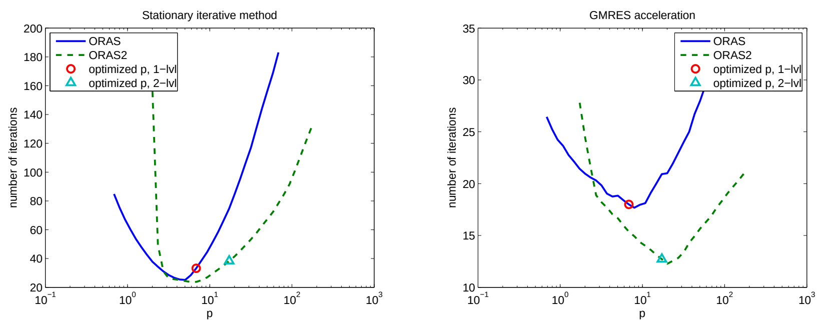

Fig. 3. Convergence for different values of the Robin parameter p, when $h = 1/64$ and $H = 1/4$ (16 subdomains).

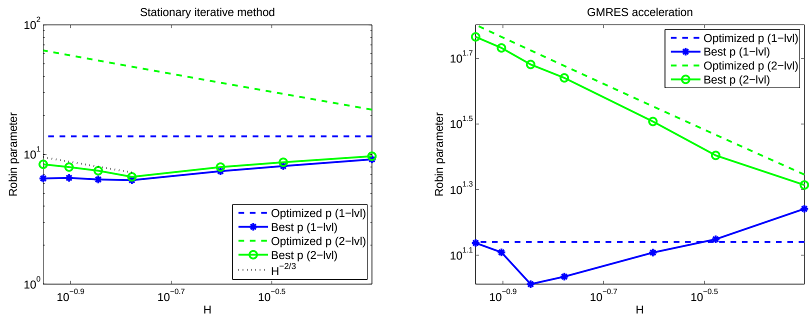

Fig. 4. Comparison of the optimized Robin parameter with the best possible value as we increase the number of subdomains.

References

Satish Balay, Kris Buschelman, William D. Gropp, Dinesh Kaushik, Matthew G. Knepley, Lois Curfman McInnes, Barry F. Smith, and Hong Zhang. PETSc Web page, 2001. http://www.mcs.anl.gov/petsc.

Xiao-Chuan Cai and Marcus Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM Journal on Scientific Computing*, 21:239–247, 1999.

Olivier Dubois. *Optimized Schwarz Methods for the Advection-Diffusion Equation and for Problems with Discontinuous Coefficients*. PhD thesis, McGill University, 2007.

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

Martin J. Gander, Frédéric Magoulès, and Frédéric Nataf. Optimized Schwarz methods without overlap for the Helmholtz equation. *SIAM J. Sci. Comput.*, 24(1):38–60, 2002.

Caroline Japhet. Optimized Krylov-Ventcell method. Application to convection-diffusion problems. In Petter E. Bjørstad, Magne S. Espedal, and David E. Keyes, editors, *Proceedings of the 9th international conference on domain decomposition methods*, pages 382–389. ddm.org, 1998.

Caroline Japhet, Frédéric Nataf, and Francois-Xavier Roux. The Optimized Order 2 Method with a coarse grid preconditioner. application to convection-diffusion problems. In P. Bjorstad, M. Espedal, and D. Keyes, editors, *Ninth International Conference on Domain Decompositon Methods in Science and Engineering*, pages 382–389. John Wiley & Sons, 1998.

Amik St-Cyr, Martin J. Gander, and Stephen J. Thomas. Optimized multiplicative, additive and restricted additive Schwarz preconditioning. *SIAM J. Sci. Comp.*, 29(6):2402–2425, 2007.

Andrea Toselli and Olof Widlund. *Domain Decomposition Methods—Algorithms and Theory*, volume 34 of *Springer Series in Computational Mathematics*. Springer-Verlag, Berlin, 2005.