NUMERICAL LINEAR ALGEBRA WITH APPLICATIONS
Numer. Linear Algebra Appl. 2004; **11**:327–341 (DOI: 10.1002/nla.351)

A hybrid domain decomposition method based on aggregation

Yu. Vassilevski[1,2,*,†]

[1]*Institute of Numerical Mathematics, Russian Academy of Sciences, 8, Gubkina str., 119991, Moscow, Russia*
[2]*Center for Environmental Simulation, The University of Augsburg, Universitätsstrasse 14, D-86135, Augsburg, Germany*

SUMMARY

A new two-level black-box preconditioner based on the hybrid domain decomposition technique is proposed and studied. The preconditioner is a combination of an additive Schwarz preconditioner and a special smoother. The smoother removes dependence of the condition number on the number of subdomains and variations of the diffusion coefficient and leaves minor sensitivity to the problem size. The algorithm is parallel and pure algebraic which makes it a convenient framework for the construction parallel black-box preconditioners on unstructured meshes. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: domain decomposition; aggregation; parallel computation

1. INTRODUCTION

Domain decomposition methods (DDMs) for partial differential equations have been gradually progressing field in the numerical analysis and applications since the end of 1980s. The approach has collected a vast bibliography [1] including several review books [2–4]. The most common objective of DDMs is as follows: Given a partition of the computational domain into subdomains and solvers or preconditioners in the subdomains, one has to construct an iterative solver for the global problem which is parallel and has a low arithmetic complexity. The convergence rate of iterations should not depend on the mesh size parameter (h), diameter of subdomains (H) and other problem parameters, e.g. the diffusion coefficient (ρ) in a diffusion equation. There exist methods (e.g. References [5, 6]) satisfying all the above requirements but they are complicated in implementation and have certain restrictions. A sensible solution is presumably to sacrifice certain properties. For instance, in the methods described in References [7–11], these properties were the arithmetic complexity (exact solution of subproblems) and a weak (polylogarithmic) dependence on the mesh size. Another example is the method proposed in [12] which allows more severe dependence on the mesh size (H/h) but has a

*Correspondence to: Y. Vassilevski, Institute of Numerical Mathematics, Russian Academy of Sciences, 8, Gubkina str., 119991, Moscow, Russia.
†E-mail: vasilevs@dodo.inm.ras.ru

Contract/grant sponsor: Center for Environmental Simulation of the University of Augsburg

Received 30 April 2003
Revised 25 August 2003

simple parallel implementation. Besides, it has low arithmetic complexity due to the usage of efficient subdomain preconditioners.

The methods mentioned above are based on the non-overlapping DD. It is conventional that the non-overlapping DD exploits smooth interfaces between subdomains fitted to material discontinuity interfaces. This implies an *a priori* partition with smooth interfaces which is not appropriate for black-box parallel solvers on unstructured meshes. Another group of methods is based on the overlapping DD. Considerations of parallel efficiency require the minimal (h) overlap of subdomains. The minimal overlap implies a simple technology for an automatic partitioning of unstructured meshes, minimal inter-processor exchanges and avoids a double arithmetic work. Another advantage of the overlapping DD is that the interface between subdomains may be not matching with jumps in the diffusion coefficient (ρ). The basic DD preconditioner is the additive Schwarz preconditioner [13] which is just a block diagonal matrix in the case of minimal overlap. The diagonal blocks are preconditioners to corresponding diagonal blocks in the stiffness matrix. For problems with self-adjoint second-order elliptic operators, the condition number of thus preconditioned system matrix depends on the subdomain diameter ($\sim 1/H^2$), the width of the overlap ($\sim 1/\delta$) and the diffusion coefficient jump ($\sim \max \rho / \min \rho$). If the additive Schwarz preconditioner is equipped with a coarse subspace, the condition number estimate is reduced to $C(\rho)(1 + (H/\delta))^2$ [14] and in the case of a smooth coefficient ρ it may be improved to $C(1 + (H/\delta))$ [4, 15]. It is pertinent to note that the dependence on the coefficient $\rho(x)$ is not very important if $\rho(x)$ is smooth in large subdomains and the preconditioned CG iterations are applied to the linear solution. According to Reference [16], regardless of the linear dependence of the condition number on $\rho(x)$, the number of PCG iterations is slightly increased by a number depending on the geometric structure of the coefficient jumps. However, this result is not applicable for other Krylov subspace iterative methods and for heterogeneous coefficients.

In this paper, we consider a method based on the additive Schwarz algorithm which provides independence of the condition number on the diffusion coefficient in the region of the overlap and the estimate $C(1 + (H/\delta))$. The analysis uses the smoothness of the coefficient in subdomains out of the overlap region although it is not important in practice. Theoretical estimate independent of the coefficient heterogeneity in the overlap region is the important feature of the method. The core of the method is the particular construction of a coarse subspace based on the aggregation in subdomains. Aggregation is rather a popular technique in the computational community. The closest to the proposed technique are methods [17–19]. The common feature with the method from Reference [18] is the usage of non-smooth-coarse subspace basis functions, in contrast to the technique [17, 19]. However, as opposed to Reference [18], our coarse subspace is more rich since it incorporates fine mesh basis functions which are non-trivial in the overlap region. On the one hand, it makes the coarse subspace (aggregated) problem harder to solve. On the other hand, it allows us to eliminate the dependence on the diffusion coefficient. The approximate solution with the coarse subspace matrix is performed by a few BSOR sweeps. Effectiveness of BSOR iterations is based on the assumption that the dimension of the coarse subspace is far less than that of the original finite element space. This provides a low complexity of one BSOR sweep and high convergence rate of BSOR due to smaller order and less stiffness of the coarse subspace matrix. Moreover, the iterative solution is easily parallelized. A practical feature of the proposed method is a simplicity of its implementation on unstructured meshes, both sequential and parallel [20]. Besides the stiffness matrix, the method needs only an assignment of the degrees of freedom

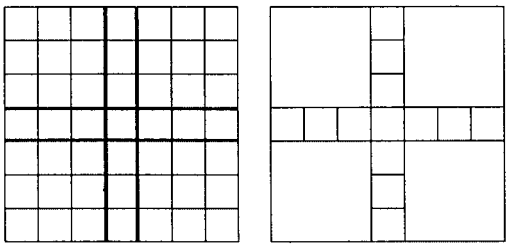

Figure 1. Mesh associated with the additive Schwarz (left) and the correction (right).

to disjoint sets of interior nodes of subdomains provided by a graph partitioning algorithm. Another peculiarity of the preconditioner is its hybrid form [4, 21] where the additive structure on the level of subdomains (for the sake of parallelism) is followed by the coarse subspace multiplicative correction (for the sake of better convergence).

The contents of the paper is as follows. A description of the model problem and assumptions on the domain decomposition in Section 2 are followed by a presentation of the method in Section 3. In Section 4 we analyse the preconditioner on the basis of the additive Schwarz lemma [13]. Numerical experiments illustrate the basic features of the method in Section 5.

2. FINITE ELEMENT PROBLEM

Let a polyhedral domain $\Omega \subset \mathbf{R}^3$, with a boundary $\partial\Omega$ be decomposed into m overlapping regular shaped subdomains Ω_i, with an overlap width δ and diameter H, i.e. $\bar{\Omega} = \bigcup_{i=1}^{m} \bar{\Omega}_i$. Each subdomain is partitioned into the interior part $\Omega_{i,\mathrm{int}}$ and the overlapping part $\Omega_{i,\mathrm{ovr}}$:

$$\Omega_{i,\mathrm{ovr}} = \Omega_i \cap \left(\bigcup_{i \neq j} \bar{\Omega}_j \right), \quad \Omega_{i,\mathrm{int}} \cap \left(\bigcup_{i \neq j} \bar{\Omega}_j \right) = \emptyset$$

We denote by $J(i)$, $i = 1, \ldots, m$, the set of indices of Ω_j such that $\Omega_i \cap \Omega_j \neq \emptyset$.

Let Ω_i^h be a conformal simplicial regular triangulation of Ω_i, $i = 1, \ldots, m$. We assume that in $\Omega_{i,\mathrm{ovr}}$ the mesh Ω_i^h matches with Ω_j^h, $j \in J(i)$. Therefore, $\bigcup_{i=1}^{m} \Omega_i^h$ defines a conformal triangulation Ω^h of Ω. For the sake of simplicity of presentation, we assume that all the nodes of $\Omega_{i,\mathrm{ovr}}^h$ belong to $\partial\Omega_{i,\mathrm{ovr}}$. This implies that we use the domain decomposition with the minimal (one element) overlap (see Figure 1). We denote by V^h the space of P_1 conforming finite elements on triangulation Ω^h with zero trace on $\partial\Omega$.

Let the bilinear form $a(\cdot,\cdot) : H_1(\Omega) \times H_2(\Omega) \to \mathbf{R}$, and the linear functional $l(\cdot) : H_1(\Omega) \to \mathbf{R}$ be given by

$$a(v,w) := \sum_{i=0}^{m} a_i(v,w), \quad a_0(v,w) := \int_{\Omega \setminus \bigcup_{i=1}^{m} \Omega_{i,\mathrm{int}}} \rho \nabla v \cdot \nabla w \, \mathrm{d}x$$

$$a_i(v,w) := \int_{\Omega_{i,\mathrm{int}}} \rho \nabla v \cdot \nabla w \, \mathrm{d}x, \quad l(v) := \int_{\Omega} f v \, \mathrm{d}x, \quad f \in L_2(\Omega)$$

We assume that $\rho(x)$ is a positive piece-wise constant function defined on Ω^h.

The finite element problem reads: find $w \in V^h$ such that

$$a(w, v) = l(v), \quad v \in V^h \tag{1}$$

The algebraic counterpart of (1) is the linear system

$$Au = f \tag{2}$$

with a sparse symmetric positive definite matrix $A \in \mathbf{R}^{n \times n}$. We assume that the solution entries are associated with the mesh nodes.

3. HYBRID DD PRECONDITIONER

The preconditioner is based on the two-level DD-method proposed in References [20, 22]. At the fine grid level, the conventional additive Schwarz preconditioner is applied. We partition the entries of a vector u into m disjoint subsets associated with the meshes $\Omega^h_{i,\mathrm{int}}$, $i = 1, \ldots, m$. The matrix A admits a block representation associated with the above partitioning:

$$A = \begin{pmatrix} A_{11} & \cdots & A_{1m} \\ \vdots & \ddots & \vdots \\ A_{m1} & \cdots & A_{mm} \end{pmatrix}$$

The diagonal blocks $A_{ii} \in \mathbf{R}^{n_i \times n_i}$, $i = 1, \ldots, m$, correspond to boundary value problems in subdomains with the Dirichlet boundary conditions on interior boundaries.

Let $B_{ii} = B_{ii}^{\mathrm{t}} > 0$ be a preconditioner for A_{ii}, $i = 1, \ldots, m$

$$\alpha_1 (B_{ii}^{-1} u, u) \leqslant (A_{ii} u, u) \leqslant \alpha_2 (B_{ii}^{-1} u, u) \tag{3}$$

with positive constants α_1, α_2. Define a block diagonal matrix

$$B_1 = \begin{pmatrix} B_{11} & & \\ & \ddots & \\ & & B_{mm} \end{pmatrix} \tag{4}$$

The matrix B_1 is the additive Schwarz preconditioner for A with the minimal overlap of subdomains. It is a very simple preconditioner for A and can be easily parallelized. However, its efficiency is affected by the number of subdomains m, the width of the overlap δ and the diffusion coefficient ρ. Both numerical results and theoretical considerations [4, 16] show that $\mathrm{cond}(B_1 A) \sim 1/\delta$, $\mathrm{cond}(B_1 A) \sim 1/H^2$, $\mathrm{cond}(B_1 A) \sim \max_x \rho(x) / \min_x \rho(x)$. In order to decrease the negative impact of using a small overlap and eliminate the dependence on the number of subdomains and variations of $\rho(x)$, we apply a correction step associated with a coarse level. The resulting hybrid DD preconditioner B_h is implicitly described by its action on a vector

$u \in \mathbf{R}^n$. Denoting $v = B_h u$, we have

$$\begin{aligned} \bar{v} &= B_2 u \\ w &= B_1(u - A\bar{v}) \\ r &= (u - Aw), \quad v = w + B_2 r \end{aligned} \tag{5}$$

where B_2 is a preconditioner described below. Formally, B_h, may be presented as

$$B_h = (I - B_2 A)B_1(I - AB_2) + B_2. \tag{6}$$

We choose the preconditioner B_2 as follows. Let the number of nontrivial rows in the matrix $A_i = [A_{i1}, \ldots, A_{i,i-1}, A_{i,i+1}, \ldots, A_{i,m}]$ (without the diagonal block) be $\tilde{n}_i$. We define

$$\tilde{n} = \sum_{i=1}^{m} \tilde{n}_i + m$$

and assume that rows of the matrix A are ordered in such a way that in each matrix A_i the non-trivial rows go first. Then the local *aggregation matrix* $T_{ii} \in \mathbf{R}^{n_i \times (\tilde{n}_i + 1)}$ is given by

$$T_{ii} = \begin{pmatrix} I_i & 0 \\ 0 & e_i \end{pmatrix}, \quad e_i = (1, \ldots, 1)^{\mathrm{t}} \in \mathbf{R}^{n_i - \tilde{n}_i}$$

where I_i is the identity matrix. We define the global block diagonal aggregation matrix T by

$$T = \begin{pmatrix} T_{11} & & \\ & \ddots & \\ & & T_{mm} \end{pmatrix}$$

and the coarse subspace stiffness matrix $\tilde{A}$ by

$$\tilde{A} = T^{\mathrm{t}} A T, \quad \tilde{A} \in \mathbf{R}^{\tilde{n} \times \tilde{n}}$$

Let $\tilde{B}$ be the conventional BSOR (BSSOR) smoother [23, 24] for $\tilde{A}$ discussed in the end of the next section. Then, the preconditioner B_2 is defined implicitly by

$$B_2 = T \tilde{B} T^{\mathrm{t}} \tag{7}$$

Note that B_2 is the smoother in a subspace of aggregated vectors $\tilde{V} = \{v \in \mathbf{R}^n : v = T\tilde{v},\ \tilde{v} \in \mathbf{R}^{\tilde{n}}\}$. We motivate its construction as follows. The drawback of the additive Schwarz preconditioner B_1 is that it damps the error locally in subdomains but it does not control the error propagation on the global scale (dependence on H) and does not coordinate the error damping between neighbouring subdomains (dependence on δ). The approximate BSOR inversion of the coarse subspace matrix $\tilde{A}$ coordinates the mean subdomain values, matches the local errors in overlapping strips (see Figure 1) and eliminates the sensitivity of cond$(B_1 A)$ to variations of ρ.

If B_2 has form (7), the evaluation of the three-stage hybrid DD algorithm (6) is equivalent to the evaluation of a two-stage preconditioner [9, 21]

$$B = B_1 + B_2(I - AB_1) \tag{8}$$

on vectors belonging to the subspace $K = \{v \in \mathbf{R}^n : T^t v = 0\}$ (for a general theory of iterative processes in subspaces we refer to Reference [25]). Indeed,

$$B_h v = (I - B_2 A)B_1(I - AB_2)v + B_2 v = B_1 v + B_2(I - AB_1)v = Bv \quad \forall v \in K$$

since $B_2 v = T\tilde{B}T^t v = 0$. Therefore, in spite of unsymmetry, preconditioner (8) may be applied in any right preconditioned Krylov subspace iterative technique (e.g. PCG) provided that it is evaluated on vectors from the subspace K. To this end, we assume that $\tilde{B} = \tilde{A}^{-1}$ and an arbitrary initial vector is corrected as follows:

$$u_0 := u_0 + B_2(f - Au_0) \tag{9}$$

Then all the residuals $r_k := f - Au_k = r_0 - ABP_{k-1}(AB)r_0$ belong to the subspace K. (Here P_{k-1} is a polynomial of degree $k-1$, $P_{k-1}(0) = 1$). Indeed, for the initial residual r_0

$$T^t r_0 = T^t f - T^t A(u_0 + B_2(f - Au_0)) = T^t(I - AB_2)f - T^t(I - AB_2)Au_0 = 0$$

since $T^t(I - AB_2) = T^t - T^t AT\tilde{A}^{-1}T^t = T^t - T^t AT(T^t AT)^{-1}T^t = 0$. For the next residual in a Krylov subspace method

$$T^t r_1 = T^t r_0 - T^t ABP_0(AB)r_0 = 0$$

since $T^t AB = T^t AB_1 - T^t AB_2 AB_1 + T^t AB_2 = T^t(I - AB_2)AB_1 + T^t AT(T^t AT)^{-1}T^t = T^t$. For the residuals of other iterations $T^t r_k = 0$ by induction: $T^t r_k = T^t r_0 - T^t ABP_{k-1}(AB)r_0 \in \text{span}\{T^t r_0, T^t ABr_0, \ldots, T^t ABr_{k-1}\}$.

In practice, however, $\tilde{B}$ is an approximate inverse of $\tilde{A}$ due to using a small number of BSOR sweeps, and the equivalence between (8) and (6) is not valid since the residuals are not strictly in K. Therefore, the theoretical justification of using the preconditioner B in PCG iterations is not complete when $\tilde{B} \neq \tilde{A}^{-1}$. However, according to numerical evidence, the PCG method preserves high convergence rate in this case as well. The two-stage preconditioner (8) is more appealing in applications than its symmetric counterpart (6) since it has lesser arithmetic complexity. On the other hand, the three-stage preconditioner (6) is symmetric and positive definite operator provided $\tilde{A}$ is inverted approximately with BSSOR. The analysis of the preconditioner B_h is given in the next section.

4. ANALYSIS OF THE PRECONDITIONER

For the sake of simplicity, we assume that $\tilde{B} = \tilde{B}^t$, $\alpha_1(\tilde{B}^{-1}v, v) \leqslant (\tilde{A}v, v) \leqslant \alpha_2(\tilde{B}^{-1}v, v)$, $\forall v \in \tilde{V}$.

The abstract form of preconditioners (8) and (6) is the two level hybrid Schwarz preconditioner [4]. In the following, we estimate the condition number $\text{cond}(B_h A)$ for the symmetrized version (6). The first analysis [21] of (6) is based on the observation that the condition number $\text{cond}(B_h A)$ for the hybrid method is smaller than the condition number $\text{cond}(B_a A)$ of the additive method where

$$B_a = B_1 + B_2 \tag{10}$$

Lemma 1 (*Mandel* [*21*])

$$\lambda_{\min}(B_h A) \geqslant \lambda_{\min}(B_a A)$$

$$\lambda_{\max}(B_h A) \leqslant \lambda_{\max}(B_a A)$$

where $\lambda_{\min}(D)$ and $\lambda_{\max}(D)$ stand for the minimum and maximum eigenvalues of a matrix D, respectively.

The analysis of the additive Schwarz method is based on

Theorem 1 (*Matsokin and Nepomnyaschikh* [*13*])
Let the Hilbert space V be split into $m+1$ subspaces V_i and A be a linear self-adjoint continuous positive definite operator. We assume also that for any $v \in V$ there exists a decomposition $V = \sum_{i=0}^{m} v_i$ such that

$$\mu \sum_{i=0}^{m} (Av_i, v_i) \leqslant (Av, v) \tag{11}$$

$$\max_{0 \leqslant j \leqslant m} \sum_{i=0}^{m} \varepsilon_{ij} \leqslant \nu \tag{12}$$

with positive constants μ, ν where $\varepsilon_{ij} = 0$ if $V_i \perp V_j$, $\varepsilon_{ij} = 1$ otherwise.
Then

$$\mu\alpha_1(B^{-1}v, v) \leqslant (Av, v) \leqslant \nu\alpha_2(B^{-1}v, v) \quad \forall v \in V \tag{13}$$

for $B = \sum_{i=0}^{m} B_i$, $B_i = B_i^{\mathrm{t}} : V \to V_i$, $\alpha_1(B_i^{-1}v_i, v_i) \leqslant (Av_i, v_i) \leqslant \alpha_2(B_i^{-1}v_i, v_i)$ $\forall v_i \in V_i$.

The analysis of $\mathrm{cond}(B_a A)$ with B_1, B_2 from (4), (7) is based on the technique [17, 19] adapted to the non-smoothed aggregation [18]. Let ψ_l be the nodal basis functions of the space V^h so that the expansion of a function $v \in V^h$ be $v = \sum_l v_l \psi_l$. Let $L(i)$ and $M(i)$ denote the sets of ψ_l whose support belong to $\Omega_{i,\mathrm{int}}$ and Ω_i, respectively, $i = 1, \ldots, m$, and $L(0)$ stands for the set of ψ_l not contributing to $\bigcup_{i=1}^{m} L(i)$. The coarse mesh basis functions are defined on the basis of ψ_l and $L(i)$

$$\Psi_i = \sum_{l \in L(i)} \psi_l, \quad i = 1, \ldots, m$$

We remark that there exist two partitions of unity

$$\sum_l \psi_l = 1, \quad \sum_{i=1}^{m} \Psi_i + \sum_{l \in L(0)} \psi_l = 1$$

We define the subspaces V_i, $i=0,1,\ldots,m$ as follows:

$$V_0=\left\{v \mid v=\sum_{i=1}^{m}\gamma_i\Psi_i+\sum_{l\in L(0)}\beta_l\psi_l,\ \forall\beta_l,\gamma_i\right\}$$

$$V_i=\left\{v \mid v=\sum_{l\in M(i)}\beta_l\psi_l,\ \forall\beta_l\right\},\quad i>0$$

The projector Q onto V_0 is a combination of L_2-projection and the interpolation operator

$$Qv=\sum_{i=1}^{m}\xi_i\Psi_i+\sum_{l\in L(0)}v_l\psi_l,\quad \xi_i=\int_{\Omega_{i,\mathrm{int}}}v\,\mathrm{d}x\Big/\int_{\Omega_{i,\mathrm{int}}}1\,\mathrm{d}x$$

The projector Q possesses an important property: given a decomposition on $\Omega_{i,\mathrm{int}}$

$$v(x)=\bar{v}(x)+\omega_i,\quad \int_{\Omega_{i,\mathrm{int}}}\bar{v}\,\mathrm{d}x=0,\quad \omega_i\in\mathbf{R}^1$$

Qv may be represented on $\Omega_{i,\mathrm{int}}$ as

$$Qv=Q\bar{v}+Q\omega_i=\bar{\bar{\xi}}_i\Psi_i+\sum_{l\in L(0)\cap M(i)}\bar{v}_l\psi_l+\omega_i\Psi_i+\sum_{l\in L(0)\cap M(i)}\omega_i\psi_l=\sum_{l\in L(0)\cap M(i)}\bar{v}_l\psi_l+\omega_i$$

since $\bar{\bar{\xi}}_i=\int_{\Omega_{i,\mathrm{int}}}\bar{v}\,\mathrm{d}x/\int_{\Omega_{i,\mathrm{int}}}1\,\mathrm{d}x=0$.

Therefore,

$$a(Qv,Qv)=a_0(v,v)+\sum_{i=1}^{m}a_i\left(\sum_{l\in L(0)\cap M(i)}\bar{v}_l\psi_l,\sum_{l\in L(0)\cap M(i)}\bar{v}_l\psi_l\right)\tag{14}$$

Let $\rho_{1,i}\leqslant\rho(x)\leqslant\rho_{2,i}$, $x\in\Omega_{i,\mathrm{int}}$, with positive constants $\rho_{1,i}$, $\rho_{2,i}$, and let $\Gamma_{\delta,i}$ denote the restriction of the support of $\sum_{l\in L(0)\cap M(i)}\psi_l$ onto $\Omega_{i,\mathrm{int}}$. Using the inverse inequality $|\psi_l|^2_{H^1}\leqslant Ch_l^{-2}\|\psi_l\|^2_{L_2}$ and $h_l\sim\delta$ in $\Gamma_{\delta,i}$ (assuming mesh quasi-uniformity in the overlap region), we obtain

$$a_i\left(\sum_{l\in L(0)\cap M(i)}\bar{v}_l\psi_l,\sum_{l\in L(0)\cap M(i)}\bar{v}_l\psi_l\right)\leqslant C\rho_{2,i},\qquad \sum_{l\in L(0)\cap M(i)}\bar{v}_l|\psi_l|^2_{H^1(\Omega_{i,\mathrm{int}})}\leqslant C\,\frac{\rho_{2,i}}{\delta^2}\|\bar{v}\|^2_{L_2(\Gamma_{\delta,i})}\tag{15}$$

Hereinafter, C denotes a generic positive constant. In order to estimate $\|\bar{v}\|^2_{L_2(\Gamma_{\delta,i})}$, we take advantage of the estimate [4]

$$\begin{aligned}\|\bar{v}\|^2_{L_2(\Gamma_{\delta,i})}&\leqslant C\delta^2\left(\left(1+\frac{H}{\delta}\right)|\bar{v}|^2_{H^1(\Omega_{i,\mathrm{int}})}+\frac{1}{H\delta}\|\bar{v}\|^2_{L_2(\Omega_{i,\mathrm{int}})}\right)\\&\leqslant C\delta^2\left(\left(1+\frac{H}{\delta}\right)\frac{a_i(\bar{v},\bar{v})}{\rho_{1,i}}+\frac{1}{H\delta}H^2\,\frac{a_i(\bar{v},\bar{v})}{\rho_{1,i}}\right)\leqslant\frac{C\delta^2}{\rho_{1,i}}\,\frac{H}{\delta}\,a_i(\bar{v},\bar{v})\end{aligned}\tag{16}$$

Here we exploited the Friedrich's inequality

$$\|\bar{v}\|_{L_2(\Omega_{i,\mathrm{int}})} \leqslant CH|\bar{v}|_{H^1(\Omega_{i,\mathrm{int}})}, \quad \int_{\Omega_{i,\mathrm{int}}} \bar{v}\,\mathrm{d}x = 0$$

Combining (14)–(16) we obtain

$$a(Qv, Qv) \leqslant a_0(v,v) + C \sum_{i=1}^{m} \frac{\rho_{2,i}}{\rho_{1,i}} \frac{H}{\delta} a_i(\bar{v},\bar{v}) \leqslant C \max_{1\leqslant i\leqslant m}\left(\frac{\rho_{2,i}}{\rho_{1,i}}\right) \frac{H}{\delta} a(v,v) \tag{17}$$

We apply Theorem 1 for the case of two subspaces V_0 and $\sum_{i=1}^{m} V_i$. The decomposition of a function v

$$v = Qv + (I - Q)v$$

is stable, $\mu^{-1} = 2C \max_{1\leqslant i\leqslant m}(\rho_{2,i}/\rho_{1,i})(H/\delta)$ and $v = 2$. Therefore, we proved

Theorem 2
Let Ω^h be a conformal triangulation of Ω which is partitioned into m overlapping regular shaped subdomains Ω_i with the width of minimal overlap δ and the diameter H and which is quasi-uniform in the overlap region extended by one cell layer. Let the diffusion coefficient $\rho(x)$ be smooth out of the region of the overlap: $\rho_{1,i} \leqslant \rho(x) \leqslant \rho_{2,i},\ x \in \Omega_{i,\mathrm{int}}$ and both $\tilde{B}$ and $B_{ii},\ i = 1,\ldots,m$, be preconditioners for $\tilde{A}$ and A_{ii}, respectively, with constants of equivalence α_1, α_2. Then for methods (6) and (10) the estimate

$$\mathrm{cond}(B_h A) \leqslant \mathrm{cond}(B_a A) \leqslant 2\frac{\alpha_2}{\alpha_1}\left(1 + 2C \max_{1\leqslant i\leqslant m}\left(\frac{\rho_{2,i}}{\rho_{1,i}}\right)\frac{H}{\delta}\right) \tag{18}$$

holds.

Remark
It is pertinent to notice that the diffusion coefficient may be strongly varying and even heterogeneous in the region of the overlap $\bigcup_{i=1}^{m} \Omega_{i,\mathrm{ovr}}$. The numerical evidence shows that the requirement of the smoothness of $\rho(x)$ in $\Omega_{i,\mathrm{int}}$ is not necessary. Theoretically, the mesh Ω^h has to be regular and quasi-uniform in the extended region of overlap $\bigcup_{i=1}^{m}\{\Omega_{i,\mathrm{ovr}} \cup \Gamma_{\delta,i}\}$ only. Numerical experiments show that the method works well for anisotropic and/or locally refined meshes too. We remark that the asymptotic estimate (18) is similar to the estimate for the method [15] in the case of smooth coefficients. However, in the case of jumping coefficients the latter deteriorates to $(H/\delta)^2$ (Remark 3.3 in Reference [12]). The analysis presented here may be extended to the bilinear forms with reaction terms. In this case, additional summands will appear in the right-hand side of identity (14). These summands may be easily estimated and estimate (17) remains valid in this case as well.

Numerical efficiency of the proposed preconditioner essentially depends on the complexity of the evaluation $\tilde{B}$. In contemporary computational practice, the size of the stiffness matrix A ranges from 10^5 to 10^7 while the number of subdomains (and processors) usually does not exceed few tens.[‡] Therefore, from practical point of view, the ratio H/δ does not remain

[‡]This is the typical number of processors in contemporary parallel computers. Although there exist parallel implementations with more than hundred processors, their use is confined to massively parallel computers.

constant in the sequence of refined meshes. Thus, the order of the coarse subspace matrix $\tilde{A}$ becomes essentially smaller than that of A. Typically, the matrix $\tilde{A}$ is much less stiff than A and a few sweeps of BSOR are enough for preconditioning $\tilde{A}$ since the convergence rate of BSOR is relatively high. These combined effects result in a low complexity of the coarse space correction. The techniques for searching the optimal relaxation parameter have been developed since 1970s [24, 26]. The parallel implementations of BSOR are discussed, for example, in References [23, 27]. In the reported numerical experiments, both sequential and parallel, a multi-colouring technique [23] is adopted. We advocate this choice as follows. Matrix $\tilde{A}$ is sparse except P rows which have many entries. Hence, it may be effectively partitioned into blocks by the multi-colouring greedy algorithm. In practice, the number of blocks does not exceed ten. Multi-colouring virtually implies the permutation of $\tilde{A}$ such that a diagonal block corresponding to a colour is the diagonal matrix, and both sequential and parallel implementations of BSOR are very simple. The number of BSOR iterations turns out to be insensitive to variations of the coefficient $\rho(x)$. We remark that the constituents of the method make it very convenient for parallel implementation and for applications on unstructured meshes [20].

5. NUMERICAL EXPERIMENTS

In the first part of this section we compare the additive Schwarz preconditioner and the hybrid Schwarz preconditioner (8). We consider a quasi-uniform unstructured tetrahedrization of the unit cube consisting of $N_{T,1} = 1744$ elements. The hierarchical uniform refinements of the mesh are used for the problem (1) with $\rho(x) = 1$ and $f(x) = 1$. The decomposition into two subdomains with the minimal overlap $\delta = h$ is provided by separating the mesh nodes with respect to the plane $x_1 = 0.5$. In Table I we exhibit the performance of the hybrid Schwarz preconditioner B and the additive Schwarz preconditioner B_1 assuming that $\tilde{B} = \tilde{A}^{-1}$, $B_{ii} = A_{ii}$, $i = 1, 2$. Actually, the coarse subspace matrix $\tilde{A}$ is never inverted, the inversion is simulated by a large number (30) of BSOR iterations. The convergence is measured by the number of PCG iterations required to reduce the initial residual by a factor of 10^6. The number of PCG iterations with the preconditioner B is as much as 2.5 times smaller than that with the preconditioner B_1. However, the linear dependence of the condition numbers on δ^{-1} is clearly observed in both cases.

In the second experiment we show that the presence of the smoother $\tilde{B}$ in (8) damps the sensitivity of the method to the number of subdomains. We set the mesh to be a uniform

Table I. Hybrid DD method versus additive Schwarz, $m = 2$.

Preconditioner		B_1		B	
$N_{T,i}$	h_1/h_i	#iter PCG	cond $B_1^{-1}A$	#iter PCG	cond $B^{-1}A$
1744	1	12	4.6	5	1.2
13 952	2	19	10.2	7	2
111 616	4	29	23.7	11	4
892 928	8	44	52.5	17	8.5

Table II. Hybrid DD method versus additive Schwarz, number of PCG iterations for different decompositions.

Preconditioner / Decomposition\problem	B_1		B	
	Case 1	Case 2	Case 1	Case 2
$m = 4 \times 4 \times 4$	54	90	19	18
$m = 6 \times 6 \times 6$	88	182	17	18
$m = 16 \times 1 \times 1$	76	170	30	53

tetrahedrization of a cubic grid in $\Omega = (0,1)^3$, $h = 2^{-6}$, $N_T = 1\,572\,864$. Again, we consider the ideal case: $B_{ii} = A_{ii}$, $\tilde{B} = \tilde{A}^{-1}$. Along with problem (1) with $\rho(x) = f(x) = 1$ (Case 1), we consider the problem (Case 2)

$$-\Delta u + u = (1 + 3\pi^2)\cos \pi x_1 \cos \pi x_2 \cos \pi x_3 \quad \text{in } \Omega \tag{19}$$

$$\frac{\partial u}{\partial n} = 0, \quad \text{on } \delta\Omega$$

In Table II the number of PCG iterations is exhibited. Two types of the domain partitioning are considered: regular and irregular. Regular decompositions feature isotropic subdomain shapes whereas irregular one is characterized by high anisotropy of subdomain-slices. We remark that for shape regular domain decompositions method (8) is insensitive to the number of subdomains. On the other hand, for irregular decompositions the dependence on the subdomain number exists, although the smoother application reduces the number of iterations.

The next example is the parallel solution of the diffusion equation (1) in the unit cube $\Omega = (0,1)^3$. We study three choices for $\rho(x)$. First, we take the uniform isotropic case with $\rho(x) \equiv 1$. Second, we consider the checkerboard-like jumps in the diffusion coefficient:

$$\rho(x) = \begin{cases} 1, & x \in \Omega_1 \cup \Omega_2 \cup \Omega_3 \cup \Omega_4 \\ 1000 & \text{otherwise} \end{cases}$$

where $\Omega_1 = (0,0.5)^3$, $\Omega_2 = (0.5,1)^2 \times (0,0.5)$, $\Omega_3 = (0.5,1) \times (0,0.5) \times (0.5,1)$ and $\Omega_4 = (0,0.5) \times (0.5,1)^2$. Third, we consider the piecewise constant quasi-random diffusion coefficient given by

$$\rho(x)|_{e_i} = \begin{cases} 1 & \sin(1000x_{i1} + 3000x_{i2} + 5000x_{i3}) > 0 \\ 1000 & \text{otherwise} \end{cases}$$

where $x_i = (x_{i1}, x_{i2}, x_{i3})$ is the barycentre of the tetrahedron e_i. We consider a uniform tetrahedrization of the cubic grid with $h = 2^{-5}$, $N_T = 196608$. In addition, we artificially split the mesh elements between processors either by the plane $x_1 = 0.5$ (for $m = 2$), or by the planes $x_1 = 0.5$ and $x_2 = 0.5$ (for $m = 4$), or by the planes $x_1 = 0.5$, $x_2 = 0.5$ and $x_3 = 0.5$ (for $m = 8$). Therefore, the corresponding subdomains in the Schwarz method overlap along interfaces where $\rho(x)$ has jumps. The subdomain preconditioner is chosen to be the algebraic multigrid

Table III. Number of PCG iterations and root's CPU time.

	Case 1		Case 2		Case 3	
m	#iter PCG	s	#iter PCG	s	#iter PCG	s
2	14	1.22	12	1.02	13	1.13
4	14	0.58	12	0.58	14	0.62
8	14	0.38	13	0.35	13	0.33

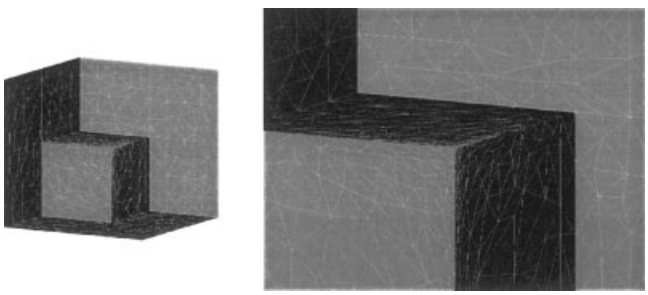

Figure 2. An example of the adaptive mesh in the domain with the reentrant corner.

Table IV. Number of PCG iterations and root's CPU time for $m=4$ on adaptive unstructured meshes.

N_T	9452	19 802	28 754	36 810	52 893	101 344	164 184
#iter PCG	8	11	12	13	14	17	20
s	0.05	0.10	0.12	0.15	0.22	0.53	0.88
$\frac{\text{s}}{\text{\#iter PCG}\, N_T}\, 10^7$	6.6	4.6	3.5	3.1	3.0	3.1	2.7

method (one V(1,1)-cycle of AMG1R5 [28]) and $\tilde{B}$ is the result of 3 BSOR iterations for the aggregated system.

Table III shows that the convergence rate of the PCG method does not depend on the jumps in the diffusion coefficient and on the number of subdomains (processors). Moreover, in all cases, the solver exhibits good parallel properties. The time measurement has been performed on a Compaq TruCluster with processors cadenced to 667 MHz.

Finally, we consider Equation (1) in a domain with a reentrant corner, $\Omega=(0,1)^3\backslash[0,0.5]^3$, and a singular right-hand side $f(x)=1/|x-x_0|$ where $x_0=(0.5,0.5,0.5)$. The solution possesses anisotropic edge singularities and a strong singularity at the reentrant corner point. We consider a sequence of adaptive unstructured meshes with local isotropic (towards the corner point) and anisotropic (towards the reentrant edges) refinements, see Figure 2. For this example, four BSOR sweeps are used for the approximate inversion of $\tilde{A}$ and one V(1,1)-cycle of the AMG method for the evaluation of B_{ii}. In Table IV the convergence rate and the arithmetic

Table V. Number of PCG iterations for $N_T \sim 10^5$.

Mesh type\m	2	4	6	8
1	15	16	17	17
2	17	19	19	19
3	18	20	20	19

complexity of the PCG method is shown for 4-processor runs. We note that the number of iterations grows as the number of elements N_T is increased. However, the dependence on N_T is moderate: 10-fold increase in N_T only doubles the number of PCG iterations (#iter PCG). It is attributable to the two-fold feature of method (8). On the one hand, the smaller order of B_{ii}, the better cond$(B_{ii}A_{ii})$, $i = 1, \ldots, m$, is. On the other hand, for uniform meshes cond(BA) depends on the mesh size h reciprocally (18), i.e., #iter PCG $\sim h^{-1/2}$. In our experiments, the meshes are not uniform but a weak dependence of #iter PCG on N_T is observed as well. The arithmetic complexity per iteration per element is decreased and saturated as N_T grows, resulting in a very good arithmetic scalability on the fine meshes. This is due to the reduction of the relative weight of interprocessor communications for large N_T.

In Table V we compare the convergence rate of the parallel solver for three types of adaptive meshes with $N_T \sim 10^5$: (1) almost quasi-uniform meshes, (2) roughly adapted (towards the corner point mostly), (3) well adapted to problem singularities (Figure 2). The trivial initial guess for PCG is corrected by (9) in all the cases. The most important observation is that the number of PCG iterations is almost insensitive to both the structure of the mesh and the number of processors.

6. CONCLUSION

The iterative solution of the diffusion equation with highly varying diffusion coefficients was addressed. The preconditioner is based on the overlapping hybrid domain decomposition framework. Its coarse subspace is associated with the space of aggregated vectors. It has as many degrees of freedom in the region of the overlap as the original finite element space but only one degree of freedom for the remaining parts of each subdomain. Robustness of the condition number of the preconditioned stiffness matrix is shown with respect to the number of subdomains and variations of the diffusion coefficient in the region of the overlap. The unsymmetric economic counterpart of the proposed preconditioner may be used in the PCG algorithm provided a proper initial guess is chosen. The method is easily parallelized and exploits the sparsity structure of the stiffness matrix only. This gives the framework for a parallel black-box solution algorithm. The numerical results on uniform grids confirm their predicted properties of the method. A parallel implementation shows the efficiency of the approach on both uniform and adaptive unstructured meshes.

ACKNOWLEDGEMENTS

The author is grateful to M. Garbey and D. Tromeur-Dervoux for the assistance in performing numerical experiments, and to K. Lipnikov for providing a mesh generator of adaptive unstructured meshes. He

also thanks Yu. Kuznetsov, K. Lipnikov and S. Nepomnyaschikh for their useful comments. The work has been supported in part by the Center for Environmental Simulation of the University of Augsburg and by the research and project 'Optimization of numerical algorithms for the solution of problems of mathematical physics' within the Russian academic program 'Contemporary problems of theoretical mathematics'.

REFERENCES

1. Domain decomposition page. *http://www.ddm.org*
2. Lebedev V. *Composition Methods*. USSR Academy of Sciences: Moscow, 1986 (in Russian).
3. Quarteroni A, Valli A. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Science Publications: Oxford, 1999.
4. Smith, B, Bjorstad P, Gropp W. *Domain Decomposition*: *Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press: Cambridge, 1996.
5. Nepomnyaschikh S. Application of domain decomposition to elliptic problems on with discontinuous coefficients. In *Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Glowinski R, Kuznetsov Yu, *et al*. (eds). SIAM: Philadelphia, PA, 1991; 242–251.
6. Vaněk P, Brezina M, Tezaur R. Two-grid method for linear elasticity on unstructured meshes. *SIAM Journal of Scientific Computing* 1999; **21**:900–923.
7. Dryja M, Sarkis M, Widlund O. Multilevel Schwarz methods for elliptic problems with discontinuous coefficients in three dimensions. *Numerische Mathematik* 1996; **72**:313–348.
8. Dryja M, Smith B, Widlund O. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM Journal on Numerical Analysis* 1994; **31**:1662–1694.
9. Mandel J. Balancing domain decomposition. *Communications in Numerical Methods in Engineering* 1993; **9**:233–241.
10. Mandel J, Brezina M. Balancing domain decomposition for problems with large jumps in coefficients. *Mathematics of Computation* 1996; **65**(216):1387–1401.
11. Rixen D, Farhat C, Tezaur R, Mandel J. Theoretical comparison of the FETI and algebraically partitioned FETI methods, and performance comparisons with a direct sparse solver. *International Journal for Numerical Methods in Engineering* 1999; **46**:501–534.
12. Bramble J, Pasciak J, Vassilev A. Analysis of non-overlapping domain decomposition algorithms with inexact solves. *Mathematics of Computation* 1998; **67**:1–19.
13. Matsokin A, Nepomnyaschikh S. A Schwarz alternating method in a subspace. *Soviet Mathematics* 1985; **29**(10):78–84.
14. Chan T, Smith B, Zou J. Overlapping Schwarz methods on unstructured meshes using non-matching coarse grids. *Numerische Mathematik* 1996; **73**(2):149–167.
15. Dryja M, Widlund O. Domain decomposition algorithms with small overlap. *SIAM Journal of Scientific Computing* 1994; **15**(3):604–620.
16. Graham I, Hagger M. Unstructured additive Schwarz-CG method for elliptic problems with highly discontinuous coefficients. *SIAM Journal of Scientific Computing* 1999; **20**:2041–2066.
17. Brezina M. Robust iterative solvers on unstructured meshes. *Ph.D. Thesis*, University of Colorado at Denver: Denver, CO, 1997.
18. Jenkins E, Kees C, Kelley C, Miller C. An aggregation-based domain decomposition preconditioner for groundwater flow. *SIAM Journal of Scientific Computing* 2001; **25**:430–441.
19. Vaněk P, Brezina M, Mandel J. Convergence of algebraic multigrid based on smoothed aggregation. *Numerische Mathematik* 2001; **88**:559–579.
20. Lipnikov K, Vassilevski Yu. Parallel adaptive solution of 3D boundary value problems by Hessian recovery. *Computer Methods in Applied Mechanics and Engineering* 2003; **192**(11–12):1495–1513.
21. Mandel J. Hybrid domain decomposition with unstructured subdomains. In *Domain Decomposition Methods in Science and Engineering*: *The Sixth International Conference on Domain Decomposition*, Quarteroni A, Kuznetsov Yu *et al*. (eds.), *Contemporary Mathematics*, vol. 157. American Mathematical Society: Providence, RI, 1994; 103–112.
22. Vassilevski Yu. Iterative solvers for the implicit parallel accurate reservoir simulator (IPARS), II: parallelizatlon issues. *TICAM Report 00-33*, The University of Texas, Austin, 2000, http://www.ticam.utexas.edu/reports/2000/index.html
23. Saad Y. *Iterative Methods for Sparse Linear Systems*. PWS Publishing Co.: Boston, 1996.
24. Young D. *Iterative Solution of Large Linear Systems*. Academic Press: New York, 1971.
25. Kuznetsov Y. Matrix computational processes in subspaces. *Comp. Math. Appl. Sci. Eng.* 1984; North-Holland: Amsterdam, **6**:15–31.

26. Kincaid D, Respess J, Young D, Grimes R. Algorithm 586 ITPACK 2C: a fortran package for solving large sparse linear systems by adaptive accelerated iterative methods. *ACM Transactions on Mathematical Software* 1982; **8**:302–322.
27. Barrett R, Berry M, Chan T, Demmel J, Donato J, Dongarra J, Eijkhout V, Pozo R, Romine C, der Vorst HV. *Templates for the Solution of Linear Systems*: *Building Blocks for Iterative Methods*. SIAM: Philadelphia, PA, 1994.
28. Stüben K. Algebraic multigrid (AMG): experiences and comparisons. *Applied Mathematics and Computation* 1983; **13**:419–452.

