Chapter 22
Domain Decomposition Methods and Preconditioning

V. G. Korneev[1] and U. Langer[2]

[1] *St. Petersburg State Polytechnical University, St. Petersburg, Russia, and University of Westminster, London, UK*
[2] *Johannes Kepler University Linz, Linz, Austria*

1 INTRODUCTION

Domain decomposition (DD) methods have been developed for a long time, but most extensively since the first international DD conference held at Paris in 1987. This concerns both the theory and the practical use of DD techniques for creating efficient application software for massively parallel computers. The advances in DD theories and applications are well documented in the proceedings of the annual international DD conferences since 1987 and in numerous papers (see also the DD home page `http://www.ddm.org` for up-to-date information about the annual DD conferences, recent publications, and other DD activities). Two pioneering monographs give an excellent introduction to DD methods from two different points of view: the more algebraic and algorithmic one (Smith, Bjørstad and Gropp, 1996) and the more analytic one (Quarteroni and Vali, 1999). We refer the interested reader also to the survey articles Xu (1992), Le Tallec (1994), Chan and Mathew (1994), and Xu and Zou (1998). We start our chapter with a brief look at the DD history. In this introductory section (Section 2), we provide an exciting journey through the DD history starting in the year 1869 with the classical paper by H.A. Schwarz on the existence of harmonic functions in domains with complicated boundaries, continuing with the variational setting of the alternating Schwarz method by S.L. Sobolev in 1934, looking at the classical finite element (FE) substructuring, or superelement technique intensively used by the engineers in the sixties, and arriving at advanced domain decomposition methods developed mainly during the last 15 years. It is worth mentioning that the classical FE substructuring technique has its roots in calculation methods used in structural mechanics for a long time.

Section 3 gives an introduction to the Schwarz theory (now called Schwarz machinery) that provides a unique framework for constructing and analyzing additive and multiplicative Schwarz methods (preconditioners). Many domain decomposition and multilevel methods (preconditioners) can be put into this framework. Throughout this chapter our model objects are symmetric and positive definite (SPD) systems of algebraic equations typically resulting from finite element discretizations of elliptic problems, such as the heat conduction equation, the potential equation, and the linear elasticity equations. However, the algorithms and some of the results can be extended to more general systems, including systems with indefinite and nonsymmetric system matrices.

In Section 4, overlapping DD methods, which first appeared in Schwarz's original paper in their multiplicative

Encyclopedia of Computational Mechanics, Edited by Erwin Stein, René de Borst and Thomas J.R. Hughes. Volume 1: *Fundamentals*. ISBN: 0-470-84699-2.

version, are considered. We pay main attention to additive versions, since these are more suitable for massively parallel computers. Typically, the main components of advanced (two-level) overlapping DD methods are independent local solvers for Dirichlet problems on overlapping subdomains and a coarse space (grid) solver. The aim of the latter is to avoid strong dependence of the convergence rate (relative condition number) on the number of overlapping subdomains. A proper choice of the overlap is very important. The understanding of these influential factors is crucial for constructing efficient overlapping DD methods (preconditioners). We present the basic algorithms, look at the condition number estimates, and give an overview over some versions, including multilevel overlapping DD methods.

Section 5 is devoted to nonoverlapping DD methods, certainly most interesting for applications. This type of DD methods reflects the classical substructuring finite element technique, where the global sparse finite element system is reduced to a much smaller but denser system (the so-called Schur-complement or interface problem) by condensation of unknowns that are internal for each substructure. Iterative solvers for the interface problems, for example, the conjugate gradient method, are usually most efficient. Together with a good preconditioner they reduce the cost and provide parallelization, in part, by avoiding assembling the Schur complement. Section 5.2 concentrates on various Schur-complement preconditioners. However, the computation of contributions to the Schur complement from the substructures, even without assembling, may, in practice, be more time and memory consuming than direct solving procedures for the original system. For this reason, modern DD algorithms completely avoid the use of Schur complements and use only their preconditioners. Apart from Schur-complement preconditioning, iterative DD methods require efficient solvers for the substructure finite element problems, Dirichlet or others, at each iteration step. They are termed here as *local problems*. Fast direct solvers are hardly available for interesting applications, whereas there exists a variety of well-developed fast iterative solvers, which are easily adapted to local problems in h-versions of the finite element method (FEM). They cover many specific situations, for example, subdomains of complicated and specific shapes, orthotropies, and so on. The implementation of local iterative solvers as inexact solvers may be most efficient, but their use for solving the local Dirichlet FE subproblems arising in the so-called extension (prolongation) and restriction operations (from and to the interface, resp.) is very delicate. However, if these procedures are based on bounded discrete extensions (prolongations) and their transposed operations (restrictions), then stability can be proven. This and other topics related to the inexact iterative substructuring are discussed in Section 5.3. We present the (balanced) Neumann–Neumann method as a special Schur-complement preconditioning technique in Section 5.4, and proceed with the finite element tearing and interconnecting (FETI) and Mortar methods in the next two sections. The FETI method requires a conform triangulation, whereas the FE subspaces are separately given on each substructure including its boundary. The global continuity is then enforced by Lagrange multipliers, resulting in a saddle-point problem

Algorithm 1. Alternating Schwarz Method for solving the boundary value problem (3)–(4).

$u^0 \in \mathbb{C}^2(\Omega) \cap \mathbb{C}(\overline{\Omega})$ given initial guess: $u^0 = 0$ on $\partial\Omega$ {initialization}

{begin iteration loop}

for $n = 0$ **step** 1 **until** Convergence **do**

First step: {update in Ω_1}

Define $\widetilde{u}^{n+1/2} \in \mathbb{C}^2(\Omega_1) \cap \mathbb{C}(\overline{\Omega}_1)$: $-\Delta\widetilde{u}^{n+1/2} = f$ in Ω_1, $\widetilde{u}^{n+1/2} = u^n$ on $\partial\Omega_1$,

$u^{n+1/2}(x) = \widetilde{u}^{n+1/2}(x) \qquad \forall x \in \overline{\Omega}_1,$

$u^{n+1/2}(x) = u^n(x) \qquad \forall x \in \overline{\Omega}_2 \setminus \overline{\Omega}_1.$

Second step: {update in Ω_2}

Define $\widetilde{u}^{n+1} \in \mathbb{C}^2(\Omega_2) \cap \mathbb{C}(\overline{\Omega}_2)$: $-\Delta\widetilde{u}^{n+1} = f$ in Ω_2, $\widetilde{u}^{n+1} = u^{n+1/2}$ on $\partial\Omega_2$,

$u^{n+1}(x) = \widetilde{u}^{n+1}(x) \qquad \forall x \in \overline{\Omega}_2,$

$u^{n+1}(x) = u^{n+1/2}(x) \qquad \forall x \in \overline{\Omega}_1 \setminus \overline{\Omega}_2.$

end for

{end iteration loop}

that can be solved via its dual problem. In Mortar methods, additionally the conformity of the triangulation across the subdomain boundaries is skipped, making the method very flexible.

Let us mention that we are mainly looking for asymptotically (almost) optimal DD preconditioners (solvers) in the sense that the memory requirement and the arithmetical costs should be (almost) proportional to the number of unknowns. In this connection, we sometimes also speak about linear complexity preconditioners (solvers).

Finally, let us mention that this contribution cannot cover all aspects of domain decomposition techniques and their applications. For instance, the *p*- and the *hp*-versions of the FEM have some specific features that are not discussed in this contribution in detail; see **Chapter 5**, **Chapter 6 of this Volume** and **Chapter 3, Volume 3** for more information on these topics. Other discretization techniques like the boundary element methods (BEM) are also not discussed in this paper (see **Chapter 12** and **Chapter 21 of this Volume**). The coupling of FEM and BEM is naturally based on DD techniques (see **Chapter 13, this Volume**). We also refer to the corresponding publications, which have mostly appeared quite recently. The field of the application of DD methods is now very wide. Here, we especially refer to the proceedings of the annual DD conferences mentioned above.

2 DOMAIN DECOMPOSITION HISTORY

Schwarz (1869) investigated the existence of harmonic functions in domains Ω with complicated boundaries $\partial\Omega$: Given some boundary function g, find a function u such that

$$-\Delta u(x) = 0 \quad \forall x \in \Omega \tag{1}$$

$$u(x) = g(x) \quad \forall x \in \partial\Omega \tag{2}$$

The 'complicated' domain $\Omega = \Omega_1 \cup \Omega_2$ is supposed to be the union of two simpler overlapping domains Ω_1 and Ω_2, where the existence of harmonic functions is known (see the left part of Figure 1 for a sketch of Schwarz's original drawing where the simpler domains are a rectangle and a circle). H.A. Schwarz proposed an iteration process for solving (1)–(2) where one has to solve alternately similar problems in Ω_1 and Ω_2. Without loss of generality (homogenization of Dirichlet's boundary conditions), we will explain this Alternating Schwarz Method for the Poisson equation under homogeneous Dirichlet boundary conditions: Given some continuous function $f \in \mathbb{C}(\Omega)$, find a twice continuously differentiable function $u \in \mathbb{C}^2(\Omega) \cap \mathbb{C}(\overline{\Omega})$ such that

$$-\Delta u(x) = f(x) \quad \forall x \in \Omega \tag{3}$$

$$u(x) = 0 \quad \forall x \in \partial\Omega \tag{4}$$

Algorithm 2. Variational Alternating Schwarz Method for solving (3)–(4).

$u^0 \in \mathbb{V}_0 = \mathbb{H}_0^1(\Omega) = \mathbb{V}_1 + \mathbb{V}_2$ given initial guess {initialization}

{begin iteration loop}

for $n = 0$ **step** 1 **until** Convergence **do**

First step: {update in $\mathbb{V}_1$}

$u^{n+1/2} = u^n + w^{n+1/2}$, with $w^{n+1/2} \in \mathbb{V}_1 = \mathbb{H}_0^1(\Omega_1) \subset \mathbb{V}_0$:

$$\int_{\Omega_1} \nabla w^{n+1/2} \cdot \nabla v \, \mathrm{d}x = \int_{\Omega_1} f(x)v(x) \, \mathrm{d}x - \int_{\Omega_1} \nabla u^n \cdot \nabla v \, \mathrm{d}x \quad \forall v \in \mathbb{V}_1 \tag{5}$$

Second step: {update in $\mathbb{V}_2$}

$u^{n+1} = u^{n+1/2} + w^{n+1}$, with $w^{n+1} \in \mathbb{V}_2 = \mathbb{H}_0^1(\Omega_2) \subset \mathbb{V}_0$:

$$\int_{\Omega_2} \nabla w^{n+1} \cdot \nabla v \, \mathrm{d}x = \int_{\Omega_2} f(x)v(x) \, \mathrm{d}x - \int_{\Omega_2} \nabla u^{n+1/2} \cdot \nabla v \, \mathrm{d}x \quad \forall v \in \mathbb{V}_2 \tag{6}$$

end for

{end iteration loop}

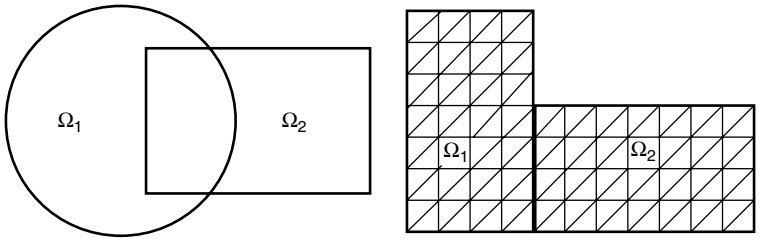

Figure 1. Overlapping and nonoverlapping DD samples.

Algorithm 1 now describes the Alternating Schwarz Method for solving the boundary value problem (3)–(4) as a model problem. The convergence analysis was done by Schwarz (1869) using the maximum principle (see also Nevanlinna, 1939).

Sobolev (1936) gave the variational setting of the Alternating Schwarz Method for solving linear elasticity problems in the form of an alternating minimization procedure in Ω_1 and Ω_2. Algorithm 2 provides the corresponding variational formulation that is nothing else but the weak formulation of the Alternating Schwarz Algorithm 1. Sobolev (1936) proved convergence of the Variational Alternating Schwarz Algorithm 2 in $\mathbb{L}_2$ to the weak solution of the boundary value problem (3)–(4) provided by its variational formulation: Find $u \in \mathbb{V}_0 = \mathbb{H}_0^1(\Omega)$ such that

$$a(u, v) = \langle f, v\rangle \quad \forall v \in \mathbb{V}_0 \tag{7}$$

where the bilinear form $a(\cdot, \cdot)\colon \mathbb{V}_0 \times \mathbb{V}_0 \to \mathbb{R}$ and the linear form $\langle f, \cdot\rangle\colon \mathbb{V}_0 \to \mathbb{R}$ are defined as follows:

$$a(u, v) = \int_\Omega \nabla u \cdot \nabla v \, \mathrm{d}x$$

and

$$\langle f, v\rangle = \int_\Omega f(x)v(x) \, \mathrm{d}x \tag{8}$$

Mikhlin (1951) showed uniform convergence in every closed subdomain of the domain Ω. Since then, the Schwarz Alternating Method was studied by many authors (see, for example, Morgenstern, 1956 and Babuska, 1958).

As just mentioned, the solution of the variational equations (5) and (6) in Algorithm 2 is equivalent to the alternating minimization of the energy functional $E(\cdot) = (1/2)a(\cdot, \cdot) - \langle f, \cdot\rangle$ in $\mathbb{V}_1$ and $\mathbb{V}_2$ respectively, that is,

$$E(u^{n+1/2}) = \min_{w \in \mathbb{V}_1} E(u^n + w)$$

and

$$E(u^{n+1}) = \min_{w \in \mathbb{V}_2} E(u^{n+1/2} + w) \tag{9}$$

Moreover, if we introduce the orthoprojections $P_i\colon \mathbb{V} \to \mathbb{V}_i$ by the identities

$$a(P_i u, v) = a(u, v) \quad \forall v \in \mathbb{V}_i, \quad \forall u \in \mathbb{V} \tag{10}$$

then we immediately observe from (5) and (6) that $w^{n+1/2} = P_1(u - u^n)$ and $w^{n+1} = P_2(u - u^{n+1/2})$ respectively. Thus, the iteration error $z^n = u - u^n$ satisfies the recurrence relation

$$z^{n+1} = (I - P_2)(I - P_1)z^n = (I - P_1 - P_2 + P_2P_1)z^n \tag{11}$$

that is nothing but an alternating orthoprojection of the iteration error to $\mathbb{V}_1^\perp$ and $\mathbb{V}_2^\perp$. This alternating projection procedure can obviously be generalized to a decomposition of $\mathbb{V}$ into many subspaces, to finite element subspaces, to other problems, and so on. Owing to the multiplicative nature of the error transition, these kind of alternating projection procedures are nowadays called multiplicative Schwarz methods (MSM). The main drawback of the MSM is connected with the sequential character of this procedure that makes the parallelization difficult. To overcome this drawback, additive versions of Schwarz algorithms were proposed. These observations led to the modern theory of Schwarz methods that has been developed during the last 15 years (see Section 3).

The substructuring technique developed by mechanical engineers for the finite element analysis of complex structures in the sixties (see e.g. Przemieniecki, 1963) is usually recognized for the other main root of the modern numerical DD algorithms. In the classical finite element substructuring technique, the computational domain Ω is decomposed into J nonoverlapping subdomains (substructures) $\Omega_j (j = 1, 2, \ldots, J)$ such that $\overline{\Omega} = \bigcup_{j=1}^J \overline{\Omega_j}$ and $\Omega_i \cap \Omega_j = \emptyset$ for $i \neq j$, and each subdomain Ω_i is divided into finite elements δ_r such that this discretization process results in a conform triangulation of Ω. In the following, the indices 'C' and 'I' correspond to the nodes belonging to the coupling boundaries (interfaces, skeleton) $\Gamma_C = \bigcup_{j=1}^J \partial\Omega_j \setminus \Gamma_D$ and to the interior $\Omega_I = \bigcup_{j=1}^J \Omega_j$ of the subdomains, respectively, where Γ_D is that part of $\partial\Omega$ where Dirichlet-type boundary conditions are given (see

Algorithm 3. Classical substructuring algorithm for solving (13).

$\widetilde{\mathbf{u}}_I = \mathbf{K}_I^{-1}\mathbf{f}_I$, that is, solve $\mathbf{K}_I\widetilde{\mathbf{u}}_I = \mathbf{f}_I$	{elimination of the internal unknowns in parallel}
$\mathbf{g}_C = \mathbf{f}_C - \mathbf{K}_{CI}\widetilde{\mathbf{u}}_I$	{forming of the right-hand side}
$\mathbf{S}_C = \mathbf{K}_C - \mathbf{K}_{CI}\mathbf{K}_I^{-1}\mathbf{K}_{IC}$	{forming of the Schur complement}
$\mathbf{u}_C = \mathbf{S}_C^{-1}\mathbf{g}_C$	{solving the Schur-complement problem}
$\mathbf{u}_I = \widetilde{\mathbf{u}}_I - \mathbf{K}_I^{-1}\mathbf{K}_{IC}\mathbf{u}_C$	{determination of the internal unknowns in parallel}

also the right part of Figure 1). Boundaries with natural (Neumann, Robin) boundary conditions will be handled as coupling boundaries.

Let us define the usual FE nodal basis

$$\begin{aligned}\Phi &= [\Phi_C, \Phi_I] \\ &= [\phi_1, \ldots, \phi_{N_C}, \phi_{N_C+1}, \ldots, \phi_{N_C+N_{I_1}}, \ldots, \phi_{N=N_C+N_I}]\end{aligned} \tag{12}$$

where the first N_C basis functions belong to Γ_C, the next N_{I_1} to Ω_1, the next N_{I_2} to Ω_2, and so on, with $N_I = N_{I_1} + N_{I_2} + \cdots + N_{I_J}$. The FE space $\mathbb{V} = \mathbb{V}_h$ is obviously a finite-dimensional subspace of the variational function space $\mathbb{V}_0$. Once the FE basis Φ is chosen, the FE scheme leads to a large-scale sparse system $\mathbf{Ku} = \mathbf{f}$ of finite element equations with the SPD stiffness matrix $\mathbf{K}$ provided that the bilinear form has the corresponding properties. Owing to the arrangement of the basis functions made above, the FE system can be rewritten in the block form

$$\begin{pmatrix} \mathbf{K}_C & \mathbf{K}_{CI} \\ \mathbf{K}_{IC} & \mathbf{K}_I \end{pmatrix} \begin{pmatrix} \mathbf{u}_C \\ \mathbf{u}_I \end{pmatrix} = \begin{pmatrix} \mathbf{f}_C \\ \mathbf{f}_I \end{pmatrix} \tag{13}$$

where $\mathbf{K}_I = \mathrm{diag}(\mathbf{K}_{I_i})_{i=1,2,\ldots,J}$ is block diagonal. The block diagonal entries $\mathbf{K}_{I_j}$ are of the dimension $N_{I_j} \times N_{I_j}$ and arise from the FE approximation to the PDE considered in Ω_j under homogenous Dirichlet boundary conditions on $\partial\Omega_j$. Owing to the block diagonal structure of $\mathbf{K}_I$, one can eliminate the internal subdomain unknowns $\mathbf{u}_I$ in parallel by block Gaussian elimination. Solving the resulting Schur-complement problem, we obtain the coupling node (interface) unknowns that allow us to define the internal subdomain unknowns finally. This classical FE substructuring algorithm is described in detail by Algorithm 3. The classical FE substructuring Algorithm 3 is well suited for parallel implementation, but very expensive with respect to the arithmetical operations, especially the forming of the Schur-complement matrix $\mathbf{S}_C$ is very time consuming. If an iterative solver is used for solving the Schur-complement problem, the forming of the Schur-complement matrix can be avoided because only the matrix-by-vector operation $\mathbf{S}_C \star \mathbf{u}_C^n$ is required. This leads us to the iterative substructuring methods, which are the starting point for the modern nonoverlapping domain decomposition methods discussed in Section 5. Iterative nonoverlapping DD algorithms and their mathematical studies appeared in the seventies. The first consistent analysis of the role of Poincaré-Steklov operators (the operator analogue of the Schur-complement matrix $\mathbf{S}_C$) in such algorithms was presented in the book by Lebedev and Agoshkov (1983).

3 FUNDAMENTALS OF SCHWARZ'S METHODS

3.1 Preliminaries

Let us consider a symmetric, elliptic (coercive), and bounded (continuous) abstract variational problem of the following form: Given $f \in \mathbb{V}_0^*$, find $u \in \mathbb{V}_0$ such that the variational equation

$$a(u, v) = \langle f, v\rangle \quad \forall v \in \mathbb{V}_0 \tag{14}$$

holds for all test functions from some Hilbert space $\mathbb{V}_0$ equipped with the scalar product $(\cdot,\cdot)_{\mathbb{V}_0}$ and the corresponding norm $\|\cdot\|_{\mathbb{V}_0}$. The bilinear form $a(\cdot,\cdot) : \mathbb{V}_0 \times \mathbb{V}_0 \to \mathbb{R}$ is supposed to be symmetric, that is,

$$a(u, v) = a(v, u) \quad \forall u, v \in \mathbb{V}_0 \tag{15}$$

$\mathbb{V}_0$-elliptic ($\mathbb{V}_0$-coercive), that is, there exists some positive constant μ_1 such that

$$\mu_1 \|v\|_{\mathbb{V}_0}^2 \le a(v, v) \quad \forall v \in \mathbb{V}_0 \tag{16}$$

and $\mathbb{V}_0$-bounded ($\mathbb{V}_0$-continuous), that is, there exists some positive constant μ_2 such that

$$a(u, v) \le \mu_2 \|u\|_{\mathbb{V}_0} \|v\|_{\mathbb{V}_0} \quad \forall u, v \in \mathbb{V}_0 \tag{17}$$

The value of the bounded (continuous) linear functional f from the dual space $\mathbb{V}_0^*$ at some $v \in \mathbb{V}_0$ is denoted by $\langle f, v\rangle$. Sometimes $\langle \cdot, \cdot\rangle : \mathbb{V}_0^* \times \mathbb{V}_0 \to \mathbb{R}$ is called duality product. Owing to Lax–Milgram's lemma, the $\mathbb{V}_0$-ellipticity (16) and the $\mathbb{V}_0$-boundness (17) ensure the existence and uniqueness of the solution of the abstract variational problem (14), (see e.g. Ciarlet, 1978).

Abstract variational formulations of the form (14) cover a lot of practically very important formally self-adjoint, elliptic boundary value problems. The Dirichlet boundary value problem for the Poisson equation introduced in Section 2 is certainly the most prominent representative of this class. Other representatives are the stationary heat conduction equation (Example 1), the linear elasticity problem (Example 2), linearized mechanical problems, and linear magneto- and electrostatic boundary value problems.

Example 1 Stationary heat conduction problem. Given some heat source intensity function $f \in \mathbb{L}_2(\Omega)$, find the temperature field $u \in \mathbb{V}_0 = \mathbb{H}_0^1(\Omega)$ such that the variational equation (15) holds with the bilinear form $a(\cdot, \cdot)\colon \mathbb{V}_0 \times \mathbb{V}_0 \to \mathbb{R}$ and the linear form $\langle f, \cdot\rangle\colon \mathbb{V}_0 \to \mathbb{R}$ defined by the identities

$$a(u, v) = \int_\Omega \alpha(x)\, \nabla u(x) \cdot \nabla v(x)\, \mathrm{d}x$$

and

$$\langle f, v\rangle = \int_\Omega f(x) v(x)\, \mathrm{d}x \tag{18}$$

respectively. As everywhere else in this chapter, if it is not defined otherwise, the computational domain $\Omega \in \mathbb{R}^d$ $(d = 1, 2, 3)$ is assumed to be bounded and sufficiently smooth (e.g. with a Lipschitz boundary). The given heat conduction coefficient $\alpha(\cdot)$ is supposed to be uniformly positive and bounded. The symmetry of the bilinear form is obvious. The $\mathbb{V}_0$-ellipticity (16) and the $\mathbb{V}_0$-boundness (17) directly follow from the Friedrichs and Cauchy inequalities, respectively (see e.g. Ciarlet, 1978). Here, we consider only homogeneous Dirichlet boundary conditions (vanishing temperature u on the boundary $\partial\Omega$ of Ω). Other boundary conditions (Neumann, Robin, mixed) can be treated in the same way. In many practical cases, the coefficient of conductivity has significant jumps, so that the ratio μ_2/μ_1 is very large. This requires DD algorithms which are robust with respect to μ_2/μ_1 (see Section 5 for robust DD methods). We mention that the heat conduction equation formally describes many other stationary processes, like diffusion or filtration in porous media with variable permeability.

Example 2 The static linear elasticity problem. Given volume forces $f = (f_1, \ldots, f_d)^{\mathrm{T}}$ in Ω and surface tractions $t = (t_1, \ldots, t_d)^{\mathrm{T}}$ on some part $\Gamma_N = \partial\Omega \setminus \Gamma_D$ of the boundary $\partial\Omega$, find the displacement $u = (u_1, \ldots, u_d)^{\mathrm{T}} \in \mathbb{V}_0 = \{v = (v_1, \ldots, v_d)^{\mathrm{T}} : v_i \in \mathbb{H}^1(\Omega),\ v_i = 0 \text{ on } \Gamma_D, i = 1, \ldots, d\}$ of the elastic body $\Omega \subset \mathbb{R}^d$ $(d = 2, 3)$ clamped at Γ_D such that the variational equation (15) holds with the bilinear form $a(\cdot, \cdot)\colon \mathbb{V}_0 \times \mathbb{V}_0 \to \mathbb{R}$ and the linear form $\langle f, \cdot\rangle\colon \mathbb{V}_0 \to \mathbb{R}$ defined by the identities

$$a(u, v) = \int_\Omega \sum_{i,j,k,l=1}^{d} \varepsilon_{ij}(x)\, D_{ijkl}(x)\, \varepsilon_{kl}(x)\, \mathrm{d}x \tag{19}$$

and

$$\langle f, v\rangle = \int_\Omega \sum_{i=1}^{d} f_i(x)\, v_i(x)\, \mathrm{d}x + \int_{\Gamma_N} \sum_{i=1}^{d} t_i(x)\, v_i(x)\, \mathrm{d}s \tag{20}$$

respectively, where the $\varepsilon_{ij} = (1/2)(\partial u_i/\partial x_j + \partial u_j/\partial x_i)$ denote the linearized strains. The given matrix $D(x) = (D_{ijkl}(x))$ of the elastic coefficients is supposed to be symmetric, uniformly positive definite, and uniformly bounded. These assumptions together with Korn's, Friedrichs', and Cauchy's inequalities ensure the symmetry (15), $\mathbb{V}_0$-ellipticity (16) and $\mathbb{V}_0$-boundness (17) of the bilinear form. If the volume forces and surface tractions are chosen in such a way that the corresponding linear functional (20) is continuous on $\mathbb{V}_0$, then the Lax–Milgram lemma again provides existence and uniqueness of the solution of the static linear elasticity problem (see e.g. Ciarlet, 1978 and Korneev and Langer, 1984).

We now approximate the abstract variational equation (14) by some FE Galerkin scheme. Let $\mathbb{V}_h$ be some finite-dimensional FE subspace of the space $\mathbb{V}_0$ spanned by some basis $\Phi_h := [\phi_1, \phi_2, \ldots, \phi_{N_h}]$, that is, $\mathbb{V}_h = \mathrm{span}\Phi_h \subset \mathbb{V}_0$, where h denotes some usual discretization parameter such that the number of unknowns N_h behaves like $O(h^{-d})$ as h tends to 0. Here and in the following, we assume that the FE discretization is based on some quasiuniform triangulation. Note that we use Φ_h as symbol for the set of basis functions $[\phi_i]_{i=1,\ldots,N_h}$ as well as for the FE-Galerkin isomorphism $(u_h \leftrightarrow \mathbf{u}_h)$

$$u_h = \Phi_h \mathbf{u}_h := \sum_{i=1}^{N_h} u_i \phi_i \tag{21}$$

mapping some vector of nodal parameters $\mathbf{u}_h = (u_i)_{i=1,\ldots,N_h} \in \mathbb{R}^{N_h}$ to the corresponding FE function $u_h \in \mathbb{V}_h$. Now the FE-Galerkin solution of the variational equation (14) is nothing but the solution of (14) on the FE subspace $\mathbb{V}_h$:

Given $f \in \mathbb{V}_0^*$, find $u_h \in \mathbb{V}_h$ such that

$$a(u_h, v_h) = \langle f, v_h \rangle \quad \forall v_h \in \mathbb{V}_h \tag{22}$$

Once the basis Φ_h is chosen, the FE scheme (22) is equivalent to the following system of FE equations: Find the nodal parameter vector $\mathbf{u}_h \in \mathbb{R}^{N_h}$ corresponding the FE solution u_h by the Galerkin isomorphism (21) as the solution of the system

$$\mathbf{K}_h \mathbf{u}_h = \mathbf{f}_h \tag{23}$$

where the stiffness matrix $\mathbf{K}_h$ and the load vector $\mathbf{f}_h$ are generated from the identities

$$\begin{aligned}(\mathbf{K}_h \mathbf{u}_h, \mathbf{v}_h) &= a(\Phi_h \mathbf{u}_h, \Phi_h \mathbf{v}_h) \\ &= a(u_h, v_h) \quad \forall u_h, v_h \leftrightarrow \mathbf{u}_h, \mathbf{v}_h \in \mathbb{R}^{N_h}\end{aligned} \tag{24}$$

and

$$(\mathbf{f}_h, \mathbf{v}_h) = \langle f, v_h \rangle \quad \forall v_h \leftrightarrow \mathbf{v}_h \in \mathbb{R}^{N_h} \tag{25}$$

respectively. Here $(\mathbf{f}_h, \mathbf{v}_h) = (\mathbf{f}_h, \mathbf{v}_h)_{R^{N_h}} = \mathbf{f}_h^{\mathrm{T}} \mathbf{v}_h$ denotes the Euclidean scalar product in $\mathbb{R}^{N_h}$.

In order to simplify the notation, we skip the subscript h in the following. Since we are primarily interested in the solution of the FE equations, there will be no confusion of some FE function $u = u_h \in \mathbb{V} = \mathbb{V}_h$ and functions u from the space $\mathbb{V}_0$. Let us now assume that the FE space

$$\mathbb{V} = \mathbb{V}_h = \sum_{j=1}^{J} \mathbb{V}_j \tag{26}$$

can be split into a (not necessarily direct) sum of the J subspaces

$$\mathbb{V}_j = \operatorname{span} \Psi_j = \operatorname{span} \Phi \mathbf{V}_j, \quad j = 1, 2, \ldots, J \tag{27}$$

where the basis $\Psi_j = [\psi_{j1}, \psi_{j2}, \ldots, \psi_{jN_j}] = \Phi \mathbf{V}_j$ of the subspace $\mathbb{V}_j$ is obtained from the original basis Φ by the $N \times N_j$ basis transformation matrix $\mathbf{V}_j$. Therefore, $N_j = \dim \mathbb{V}_j = \operatorname{rank} \mathbf{V}_j$ and $\sum_{j=1}^{J} N_j \geq N$.

The orthoprojection $P_j : \mathbb{V} \to \mathbb{V}_j$ of the space $\mathbb{V}$ onto its subspace $\mathbb{V}_j$ with respect to the energy inner product $a(\cdot, \cdot)$ is uniquely defined by the identity

$$a(P_j u, v_j) = a(u, v_j) \quad \forall v_j \in \mathbb{V}_j, \quad \forall u \in \mathbb{V} \tag{28}$$

Sometimes the orthoprojection P_j is called the Ritz, or energy projection. As orthoprojection, $P_j = P_j^*$ is self-adjoint with respect to the energy inner product, that is,

$$a(P_j u, v) = a(u, P_j v) \quad \forall u, v \in \mathbb{V} \tag{29}$$

and satisfies the projection relation $P_j^2 = P_j$. It is easy to see from (24) and (27) that the orthoprojection $P_j u$ of some $u \leftrightarrow \mathbf{u}$ can be computed by the formula

$$P_j u = \Phi \mathbf{V}_j \mathbf{u}_j = \Phi \mathbf{V}_j (\mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j)^{-1} \mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{u} \tag{30}$$

that is, $\mathbf{u}_j$ is obtained from the solution of a smaller system with the $N_j \times N_j$ system matrix $\mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j$ and the right-hand side $\mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{u}$. Similarly, replacing the original energy inner product $a(\cdot, \cdot)$ on the left-hand side of the identity (28) by some individual symmetric, $\mathbb{V}_j$-elliptic and $\mathbb{V}_j$-bounded bilinear form $a_j(\cdot, \cdot) : \mathbb{V}_j \times \mathbb{V}_j \to \mathbb{R}$, we define some projection-like operator $\widetilde{P}_j : \mathbb{V} \to \mathbb{V}_j$ that is self-adjoint but in general $\widetilde{P}_j^2 \neq \widetilde{P}_j$. Thus, $\widetilde{P}_j$ is not an orthoprojection. Again, $\widetilde{P}_j u$ can be easily calculated by the formula

$$\widetilde{P}_j u = \Phi \mathbf{V}_j \mathbf{u}_j = \Phi \mathbf{V}_j \mathbf{C}_j^{-1} \mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{u} \tag{31}$$

where the $N_j \times N_j$ matrix $\mathbf{C}_j$ is generated from the identity

$$\left(\mathbf{C}_j \mathbf{u}_j, \mathbf{v}_j\right) = a_j(u_j, v_j) \quad \forall u_j, v_j \leftrightarrow \mathbf{u}_j, \mathbf{v}_j \in \mathbb{R}^{N_j} \tag{32}$$

in the same way as $\mathbf{K}$ was generated above from the original bilinear form $a(\cdot, \cdot)$.

3.2 Schwarz methods and preconditioners for elliptic variational problems

In the next three sections, we describe various Schwarz algorithms and the corresponding preconditioners. We mention that the Schwarz algorithms are completely defined by the space splitting (26), the subspace bilinear forms $a_j(\cdot, \cdot)$, and the arrangement of the projection-like operations. Finally, we present some interesting examples.

3.2.1 Additive algorithms

The (inexact) Additive Schwarz Method (ASM) corresponding to the space splitting (26) and to the subspace bilinear forms $a_j(\cdot, \cdot)$ can be written in the form of an iteration process in the FE space $\mathbb{V}$ (function version) and in $\mathbb{R}^N$ (vector/matrix version) as shown in Algorithm 4. Replacing $a_j(\cdot, \cdot)$ by $a(\cdot, \cdot)$, $\mathbf{C}_j$ by $\mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j$, and $\widetilde{P}_j$ by P_j in Algorithm 4, we arrive at the so-called exact ASM. In this sense, we consider the exact ASM as special case of the inexact ASM presented in Algorithm 4.

Algorithm 4. Inexact ASM for solving $\mathbf{Ku} = \mathbf{f}$ (function $\leftrightarrow$ vector/matrix versions).

$u^0 = \Phi \mathbf{u^0} \in \mathbb{V}$ given initial guess and τ given iteration parameter {initialization}

{begin iteration loop}

for $n = 0$ **step** 1 **until** Convergence **do**

{parallel computation of the subspace corrections}

for all $j \in \{1, \ldots, J\}$ **in parallel do**

$w_j^n = \Phi \mathbf{V}_j \mathbf{w}_j^n \in \mathbb{V}_j$: $\quad a_j(w_j^n, v_j) = \langle f, v_j \rangle - a(u^n, v_j) \equiv a(u - u^n, v_j) \quad \forall v_j \in \mathbb{V}_j$

$\updownarrow$

$\mathbf{w}_j^n \in \mathbb{R}^{N_j}$: $\quad \mathbf{C}_j \mathbf{w}_j^n = \mathbf{V}_j^{\mathrm{T}}(\mathbf{f} - \mathbf{K}\mathbf{u}^{\mathbf{n}}) \equiv \mathbf{V}_j^{\mathrm{T}} \mathbf{d}^{\mathbf{n}}$

end for

{updating the old iterate}

$$u^{n+1} = u^n + \tau \sum_{j=1}^{J} w_j^n = u^n + \tau \sum_{j=1}^{J} \widetilde{P}_j (u - u^n)$$

$\updownarrow$

$$\mathbf{u}^{n+1} = \mathbf{u}^n + \tau \sum_{j=1}^{J} \mathbf{V}_j \mathbf{w}_j^n$$

end for

{end iteration loop}

The iteration error $z^n = u - u^n \in \mathbb{V}$ obviously satisfies the error iteration scheme

$$z^{n+1} = (I - \tau \widetilde{P}) z^n \equiv \left(I - \tau \sum_{j=1}^{J} \widetilde{P}_j \right) z^n \tag{33}$$

with the (inexact) ASM operator $\widetilde{P} = \sum_{j=1}^{J} \widetilde{P}_j$. Mention that the error iteration scheme (33) completely defines the ASM iteration scheme given by Algorithm 4. The ASM operator $\widetilde{P} = \widetilde{P}^* > 0$ is self-adjoint and positive definite with respect to the energy inner product $a(\cdot, \cdot)$. Taking the energy norm $\| \cdot \|_a = \sqrt{(a(\cdot, \cdot))}$ of the right-hand side of (33), we arrive at the iteration error estimate

$$\| z^{n+1} \|_a \leq \| I - \tau \widetilde{P} \|_a \| z^n \|_a \tag{34}$$

Therefore, the convergence rate of the ASM iteration is completely defined by the (operator) energy norm $\| I - \tau \widetilde{P} \|_a$ of the ASM iteration operator $I - \tau \widetilde{P}$ (see (38) below for rate estimates).

From the matrix representation of Algorithm 4, we immediately see that the ASM is nothing more than the Richardson iteration

$$\tau^{-1} \mathbf{C}(\mathbf{u}^{n+1} - \mathbf{u}^n) + \mathbf{K}\mathbf{u}^n = \mathbf{f} \tag{35}$$

preconditioned by the inexact ASM preconditioner

$$\mathbf{C}^{-1} = \sum_{j=1}^{J} \mathbf{V}_j \mathbf{C}_j^{-1} \mathbf{V}_j^{\mathrm{T}} \tag{36}$$

where $\mathbf{C}_j^{-1}$ is replaced by $(\mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j)^{-1}$ in the exact version. The error propagation scheme of the Richardson iteration immediately gives the iteration error estimate

$$\| \mathbf{z}^{n+1} \|_{\mathbf{K}} := \sqrt{(\mathbf{K}\mathbf{z}^{n+1}, \mathbf{z}^{n+1})_{\mathbb{R}^N}} \leq \| I - \tau \mathbf{C}^{-1} \mathbf{K} \|_{\mathbf{K}} \| \mathbf{z}^n \|_{\mathbf{K}} \tag{37}$$

where $\mathbf{z}^n = \mathbf{u} - \mathbf{u}^n$ again denotes the iteration error as vector in $\mathbb{R}^N$. The $\mathbf{K}$-energy norm error estimate (37) is the vector counterpart of the iteration error estimate (34). For $\tau \in (0, 2/\lambda_{\max})$, or $\tau \in (0, 2/\overline{\gamma})$, we get the rate estimates

$$\begin{aligned} \| I - \tau \mathbf{C}^{-1} \mathbf{K} \|_{\mathbf{K}} &= \| I - \tau \widetilde{P} \|_a \\ &= \max\{|1 - \tau \lambda_{\min}|, |1 - \tau \lambda_{\max}|\} \\ &\leq q(\tau) := \max\{|1 - \tau \underline{\gamma}|, |1 - \tau \overline{\gamma}|\} < 1 \end{aligned} \tag{38}$$

provided that the minimal eigenvalue $\lambda_{\min} = \lambda_{\min}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\min}(\widetilde{P}) \geq \underline{\gamma}$ and maximal eigenvalue $\lambda_{\max} = \lambda_{\max}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\max}(\widetilde{P}) \leq \overline{\gamma}$ for $\mathbf{C}^{-1}\mathbf{K}$ resp. $\widetilde{P}$, or at least good lower and upper bounds $\underline{\gamma}$ and $\overline{\gamma}$ are known. The lower and upper bounds of the eigenvalues of $\mathbf{C}^{-1}\mathbf{K}$ resp.

$\widetilde{P}$ satisfy the so-called spectral equivalence inequalities

$$\underline{\gamma}(\mathbf{Cv}, \mathbf{v}) \leq (\mathbf{Kv}, \mathbf{v}) \leq \overline{\gamma}(\mathbf{Cv}, \mathbf{v}) \quad \forall \mathbf{v} \in \mathbb{R}^N \tag{39}$$

which are in turn used to determine the bounds $\underline{\gamma}$ and $\overline{\gamma}$. In the following, we frequently use the short notation $\underline{\gamma}\mathbf{C} \leq \mathbf{K} \leq \overline{\gamma}\mathbf{C}$ for the spectral equivalence inequalities (39). Mention that the spectral equivalence inequalities (39) are equivalent to the inequalities

$$\underline{\gamma}\, a(v, v) \leq a(\widetilde{P}v, v) \leq \overline{\gamma}\, a(v, v) \quad \forall v \in \mathbb{V} \tag{40}$$

and

$$\underline{\gamma}\, a(\widetilde{P}^{-1}v, v) \leq a(v, v) \leq \overline{\gamma}\, a(\widetilde{P}^{-1}v, v) \quad \forall v \in \mathbb{V} \tag{41}$$

in the FE space $\mathbb{V}$. The optimal convergence rate

$$q_{\text{opt}} = q(\tau_{\text{opt}}) = \frac{\overline{\gamma} - \underline{\gamma}}{\underline{\gamma} + \overline{\gamma}} < 1 \tag{42}$$

is attained at the optimal iteration parameter $\tau = \tau_{\text{opt}} = 2/(\underline{\gamma} + \overline{\gamma})$. Sometimes it is useful to know that there also holds the iteration error estimate

$$\|\mathbf{z}^{n+1}\|_{\mathbf{KC}^{-1}\mathbf{K}} = \sqrt{(\mathbf{w}^{n+1}, \mathbf{d}^{n+1})_{\mathbb{R}^N}} \leq q(\tau)\|\mathbf{z}^n\|_{\mathbf{KC}^{-1}\mathbf{K}} \tag{43}$$

in the $\mathbf{KC}^{-1}\mathbf{K}$-energy norm with the same $q(\tau)$ as above, where $\mathbf{d}^n = \mathbf{Kz}^n = \mathbf{f} - \mathbf{Ku}^n$ and $\mathbf{w}^n = \mathbf{C}^{-1}\mathbf{d}^n$ denote the defect and the preconditioned defect (correction), respectively. In contrast to the $\mathbf{K}$-energy norm iteration error estimate (37)–(38), the $\mathbf{KC}^{-1}\mathbf{K}$-energy norm estimate (43) is computable and can be used as convergence test in Algorithm 4.

However, in practice, we use the conjugate gradient (CG) acceleration instead of the Richardson iteration, that is, we look at the ASM as a technique for constructing additive, and therefore highly parallelizable preconditioners of the form (36) by space splitting and subspace preconditioning! The preconditioning step $\mathbf{w} = \mathbf{C}^{-1}\mathbf{d}$ in the preconditioned conjugate gradient (PCG) method with the ASM preconditioner (36) is nothing but one iteration step ($n = 1$) of the Algorithm 4 applied to $\mathbf{Ku} = \mathbf{d}$ (i.e. $\mathbf{f} = \mathbf{d}$) with $\tau = 1$ and the zero initial guess $\mathbf{u^0} = \mathbf{0}$ giving $\mathbf{w} = \mathbf{u}^1$.

3.2.2 Multiplicative algorithms

The Multiplicative Schwarz Method (MSM) has its historical roots in the Additive Schwarz Method, as discussed in Section 2. The inexact version of the MSM corresponding to the space splitting (26) and to the subspace bilinear forms $a_j(\cdot, \cdot)$ can be written in the form of an iteration process in the FE space $\mathbb{V}$ (function version) and in $\mathbb{R}^N$ (vector/matrix version), as shown in Algorithm 5. We again consider the exact MSM ($a_j(\cdot, \cdot) := a(\cdot, \cdot)$, $\mathbf{C}_j := \mathbf{V}_j^{\mathrm{T}}\mathbf{K}\mathbf{V}_j$, $\widetilde{P}_j := P_j$) as a special case of the inexact version presented in Algorithm 5.

Algorithm 5. Inexact MSM for solving $\mathbf{Ku} = \mathbf{f}$ (function $\leftrightarrow$ vector/matrix versions).

$u^0 = \Phi\mathbf{u^0} \in \mathbb{V}$ given initial guess {initialization}

{begin iteration loop}

for $n = 0$ **step** 1 **until** Convergence **do**

for $j = 1$ **step** 1 **until** J **do**

{successive computation of the subspace corrections}

$$w_j^{n+(j/J)} = \Phi\mathbf{V}_j\mathbf{w}_j^{n+(j/J)} \in \mathbb{V}_j : a_j\left(w_j^{n+(j/J)}, v_j\right) = \langle f, v_j\rangle - a\left(u^{n+[(j-1)/J]}, v_j\right) \quad \forall v_j \in \mathbb{V}_j$$

$$\updownarrow$$

$$\mathbf{w}_j^{n+(j/J)} \in \mathbb{R}^{N_j} : \qquad \mathbf{C}_j\mathbf{w}_j^{n+(j/J)} = \mathbf{V}_j^{\mathrm{T}}\left(\mathbf{f} - \mathbf{Ku}^{n+[(j-1)/J]}\right) \equiv \mathbf{V}_j^{\mathrm{T}}\mathbf{d}^{n+[(j-1)/J]}$$

{immediate updating of the iterate}

$$u^{n+(j/J)} = u^{n+[(j-1)/J]} + w_j^{n+(j/J)} = u^{n+[(j-1)/J]} + \widetilde{P}_j\left(u - u^{n+[(j-1)/J]}\right)$$

$$\updownarrow$$

$$\mathbf{u}^{n+(j/J)} = \mathbf{u}^{n+[(j-1)/J]} + \mathbf{V}_j\mathbf{w}_j^{n+(j/J)}$$

end for

end for

{end iteration loop}

The iteration error $z^n = u - u^n \in \mathbb{V}$ now satisfies the error iteration scheme

$$z^{n+1} = Ez^n \tag{44}$$

resulting in the energy norm iteration error estimate

$$\|z^{n+1}\|_a \leq \|E\|_a \|z^n\|_a \tag{45}$$

with the MSM error propagation operator (iteration operator)

$$E = (I - \widetilde{P}_J)(I - \widetilde{P}_{J-1}) \cdots (I - \widetilde{P}_1) \tag{46}$$

Therefore, the convergence rate of the MSM iteration is completely defined by the (operator) energy norm $\|E\|_a$ of the MSM error propagation operator E. The same is true for the vector version with respect to the error propagation matrix (iteration matrix) $\mathbf{E} = (\mathbf{I} - \mathbf{V}_J \mathbf{C}_J^{-1} \mathbf{V}_J^{\mathrm{T}} \mathbf{K}) \cdots (\mathbf{I} - \mathbf{V}_1 \mathbf{C}_1^{-1} \mathbf{V}_1^{\mathrm{T}} \mathbf{K})$. Convergence rate estimates of the form

$$\|E\|_a = \|\mathbf{E}\|_{\mathbf{K}} \leq q_{\mathrm{MSM}} < 1 \tag{47}$$

will be presented in Section 3.3. Unfortunately, the MSM preconditioner

$$\mathbf{C} = \mathbf{K}(\mathbf{I} - \mathbf{E})^{-1} \tag{48}$$

is not symmetric and, therefore, cannot be used in the PCG as a preconditioner.

However, repeating the subspace corrections in Algorithm 5 in the reverse direction, we arrive at the so-called symmetric Multiplicative Schwarz Method (sMSM) that is characterized by the error propagation operator

$$E = (I - \widetilde{P}_1) \cdots (I - \widetilde{P}_{J-1})(I - \widetilde{P}_J)(I - \widetilde{P}_J)(I - \widetilde{P}_{J-1}) \cdots (I - \widetilde{P}_1) \tag{49}$$

resp. the error propagation matrix

$$\mathbf{E} = (\mathbf{I} - \mathbf{V}_1 \mathbf{C}_1^{-1} \mathbf{V}_1^{\mathrm{T}} \mathbf{K}) \cdots (\mathbf{I} - \mathbf{V}_J \mathbf{C}_J^{-1} \mathbf{V}_J^{\mathrm{T}} \mathbf{K}) \times (\mathbf{I} - \mathbf{V}_J \mathbf{C}_J^{-1} \mathbf{V}_J^{\mathrm{T}} \mathbf{K}) \cdots (\mathbf{I} - \mathbf{V}_1 \mathbf{C}_1^{-1} \mathbf{V}_1^{\mathrm{T}} \mathbf{K}) \tag{50}$$

resulting in a symmetric preconditioner (48). Mention that, in the exact version of the sMSM, the subspace correction in $\mathbb{V}_J$ has to be carried out only once because $(I - P_J)(I - P_J) = (I - P_J)$.

3.2.3 Hybrid algorithms

There are a lot of useful hybrid Schwarz algorithms corresponding to various possibilities of the arrangement of the subspace correction in an additive and multiplicative manner. The algorithm can be completely defined by its iteration operator E resp. the iteration matrix $\mathbf{E}$.

For instance, the iteration operator

$$E = (I - \widetilde{P}_1)(I - \tau(\widetilde{P}_2 + \cdots + \widetilde{P}_J))(I - \widetilde{P}_1) \tag{51}$$

corresponds to the Hybrid Schwarz Method described by Algorithm 6 (function version only). The error propagation operator (51) resp. Algorithm 6 correspond to a symmetric preconditioner of the form (48). Note that the error propagation operator (51) as well as the corresponding iteration matrix are self-adjoint with respect to the corresponding energy inner products.

3.2.4 Examples

We recall that the Schwarz algorithm (preconditioner) is uniquely defined by the space splitting $\mathbb{V} = \sum_{j=1}^{J} \mathbb{V}_j$, the subspace bilinear forms $a_j(\cdot, \cdot)$ and the arrangement of the projection-like operations.

Example 3 Nodal basis splitting. For $j = \overline{1, J}$ and $J = N$, we define the one-dimensional ($N_j = \dim \mathbb{V}_j = 1$) subspaces

$$\mathbb{V}_j = \mathrm{span}\{\phi_j\} = \mathrm{span}\Phi \mathbf{V}_j \tag{52}$$

of $\mathbb{V}$ giving the so-called nodal basis splitting $\mathbb{V} = \sum_{j=1}^{N} \mathbb{V}_j$, where $\mathbf{V}_j = \mathbf{e}_j$ is the jth unit vector $\mathbf{e}_j = (0, \ldots, 0, 1, 0, \ldots, 0)^{\mathrm{T}}$. Therefore,

$$\mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j = \mathbf{e}_j^{\mathrm{T}} \mathbf{K} \mathbf{e}_j = K_{jj} \tag{53}$$

Taking this relation into account, we observe that the ASM preconditioner (36) in its exact version ($\mathbf{C}_j = \mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j$) is nothing but the well-known Jacobi preconditioner (diagonal scaling)

$$\mathbf{C}^{-1} = \sum_{j=1}^{J} \mathbf{e}_j K_{jj}^{-1} \mathbf{e}_j^{\mathrm{T}} = \mathbf{D}^{-1} = (\mathrm{diag}(\mathbf{K}))^{-1} \tag{54}$$

and the exact ASM coincides with the classical (damped) Jacobi method. Furthermore, the exact MSM corresponding to our nodal basis splitting gives us the classical Gauss–Seidel method. Indeed, from the exact version of the MSM Algorithm 5 and from (53), we see that the jth subspace correction step in the nth iteration step

$$\mathbf{u}^{n+(j/J)} = \mathbf{u}^{n+[(j-1)/J]} + \mathbf{V}_j \mathbf{w}_j^{n+(j/J)} = \mathbf{u}^{n+[(j-1)/J]} + \mathbf{e}_j (\mathbf{e}_j^{\mathrm{T}} \mathbf{K} \mathbf{e}_j)^{-1} \mathbf{e}_j^{\mathrm{T}} (\mathbf{f} - \mathbf{K}\mathbf{u}^{n+[(j-1)/J]}) \tag{55}$$

Algorithm 6. Hybrid Schwarz method corresponding to the error propagation operator (51).

$u^0 = \Phi \mathbf{u^0} \in \mathbb{V}$ given initial guess and τ given iteration parameter {initialization}

{begin iteration loop}

for $n = 0$ **step** 1 **until** Convergence **do**

{first multiplicative subspace correction in $\mathbb{V}_1$}

$w_1^{n,1} \in \mathbb{V}_1 : a_1(w_1^{n,1}, v_1) = \langle f, v_1 \rangle - a(u^n, v_1) \quad \forall v_1 \in \mathbb{V}_1$

$u^{n,1} = u^n + w_1^{n,1} = u^n + \widetilde{P}_1(u - u^n)$

{additive subspace corrections in remaining subspaces}

for all $j \in \{2, \ldots, J\}$ **in parallel do**

$w_j^{n,1} \in \mathbb{V}_j : a_j(w_j^{n,1}, v_j) = \langle f, v_j \rangle - a(u^{n,1}, v_j) \quad \forall v_j \in \mathbb{V}_j$

end for

$u^{n,2} = u^{n,1} + \tau \sum_{j=2}^{J} w_j^{n,1} = u^{n,1} + \tau \sum_{j=2}^{J} \widetilde{P}_j(u - u^{n,1})$

{second multiplicative subspace correction in $\mathbb{V}_1$}

$w_1^{n,2} \in \mathbb{V}_1 : a_1(w_1^{n,2}, v_1) = \langle f, v_1 \rangle - a(u^{n,2}, v_1) \quad \forall v_1 \in \mathbb{V}_1$

$u^{n+1} = u^{n,2} + w_1^{n,2} = u^{n,2} + \widetilde{P}_1(u - u^{n,2})$

end for

{end iteration loop}

updates only the jth component of the iterate

$$u_j^{n+(j/J)} = \frac{1}{K_{jj}} \left(f_j - \sum_{i=1}^{j-1} K_{ji} u_i^{n+[(j-1)/J]} - \sum_{i=j+1}^{J} K_{ji} u_i^{n+[(j-1)/J]} \right) \quad j = 1, 2, \ldots, J \ (J = N) \quad (56)$$

Formulas (56) exactly describe the Gauss–Seidel iteration procedure. In this sense, we look at the ASM and the MSM as the natural generalizations of the Jacobi method and the Gauss–Seidel method, respectively. In an analogous way, the symmetric Gauss–Seidel method corresponds to the exact sMSM. Similar to the SOR and the SSOR methods, we can also introduce overrelaxation parameters into MSM and sMSM aiming at the improvement of convergence (see Griebel and Oswald (1995) for related results).

Therefore, the Jacobi iteration and the Gauss–Seidel method are the classical prototypes of the ASM and the MSM, respectively. Further examples are given in Section 4.3 and correspond to multilevel splittings of the finite element space $\mathbb{V}$. The most prominent one is the so-called BPX preconditioner.

3.3 Spectral equivalence estimates and convergence analysis

In this section, we present some convergence results for the Schwarz methods and some spectral equivalence results for the corresponding Schwarz preconditioners. The main condition for creating good Schwarz methods (preconditioners) consists in a stable splitting of the (FE) space $\mathbb{V}$ into subspaces $\{\mathbb{V}_j\}$. We first consider the simple case of splitting $\mathbb{V}$ into a direct sum of two subspaces $\mathbb{V}_1$ and $\mathbb{V}_2$. This case is very important for the nonoverlapping domain decomposition methods studied in Section 5. Finally, we give some result for the general case of splitting $\mathbb{V}$ into J subspaces.

3.3.1 The simple case: splitting into a direct sum of two subspaces

Let us consider the simple case of splitting

$$\mathbb{V} = \mathbb{V}_1 + \mathbb{V}_2 \quad \text{and} \quad \mathbb{V}_1 \cap \mathbb{V}_2 = \{0\} \quad (57)$$

into the two (nontrivial) subspaces $\mathbb{V}_1 = \text{span}\Phi\mathbf{V}_1$ and $\mathbb{V}_2 = \text{span}\Phi\mathbf{V}_2$, and let us define the cosine of the angle

between $\mathbb{V}_1$ and $\mathbb{V}_2$:

$$\gamma = \cos \measuredangle (\mathbb{V}_1, \mathbb{V}_2) := \sup_{v_1 \in \mathbb{V}_1 \setminus \{0\},\ v_2 \in \mathbb{V}_2 \setminus \{0\}} \frac{a(v_1, v_2)}{\|v_1\|_a \, \|v_2\|_a} < 1 \tag{58}$$

corresponding to the sharp constant γ in the so-called strengthened Cauchy inequality:

$$|a(v_1, v_2)| \le \gamma \|v_1\|_a \, \|v_2\|_a \quad \forall v_1 \in \mathbb{V}_1,\ \forall v_2 \in \mathbb{V}_2 \tag{59}$$

The splitting (57) is called stable if and only if the constant γ stays less than 1 for growing dimensions $N = N_h \to \infty$ ($h \to 0$).

The following lemma gives some useful relations for computing or estimating γ.

Lemma 1. *The following relations are valid:*

$$\cos \measuredangle (\mathbb{V}_1, \mathbb{V}_2) = \cos \measuredangle (\mathbb{V}_1^\perp, \mathbb{V}_2^\perp) \tag{60}$$

$$\sup_{v_1 \in \mathbb{V}_1 \setminus \{0\},\ v_2 \in \mathbb{V}_2 \setminus \{0\}} \frac{a(v_1, v_2)}{\|v_1\|_a \, \|v_2\|_a} = \sup_{v_1 \in \mathbb{V}_1 \setminus \{0\},\ v_2 \in \mathbb{V}_2 \setminus \{0\}} \times \frac{2a(v_1, v_2)}{\|v_1\|_a^2 + \|v_2\|_a^2} \tag{61}$$

$$\sup_{v_1 \in \mathbb{V}_1 \setminus \{0\},\ v_2 \in \mathbb{V}_2 \setminus \{0\}} \frac{a(v_1, v_2)}{\|v_1\|_a \, \|v_2\|_a} = \sup_{\mathbf{v}_1 \in \mathbb{R}^{N_1} \setminus \{0\}, \mathbf{v}_2 \in \mathbb{R}^{N_2} \setminus \{0\}} \times \frac{((\mathbf{V}_1^{\mathrm{T}} \mathbf{K} \mathbf{V}_2)(\mathbf{V}_2^{\mathrm{T}} \mathbf{K} \mathbf{V}_2)^{-1} (\mathbf{V}_2^{\mathrm{T}} \mathbf{K} \mathbf{V}_1) \mathbf{v}_1, \mathbf{v}_2)}{(\mathbf{V}_1^{\mathrm{T}} \mathbf{K} \mathbf{V}_1 \mathbf{v}_1, \mathbf{v}_1)} \tag{62}$$

The proofs of the relations (60), (61), and (62) are elementary and can be found in Bjørstad and Mandel (1991), Axelsson and Vassilevskii (1989), and Haase, Langer and Meyer (1991) respectively. Relation (62) means that γ coincides with the maximal eigenvalue of the generalized eigenvalue problem

$$(\mathbf{V}_1^{\mathrm{T}} \mathbf{K} \mathbf{V}_2)(\mathbf{V}_2^{\mathrm{T}} \mathbf{K} \mathbf{V}_2)^{-1} (\mathbf{V}_2^{\mathrm{T}} \mathbf{K} \mathbf{V}_1) \mathbf{v}_1 = \lambda \, (\mathbf{V}_1^{\mathrm{T}} \mathbf{K} \mathbf{V}_1) \mathbf{v}_1 \tag{63}$$

The error iteration schemes $z^{n+1} = E z^n$ corresponding to the exact ASM ($\tau = 1$) and the exact MSM are illustrated at the left and right sides of Figure 2, respectively. This figure shows that the exact MSM converges twice as fast as the corresponding ASM. More precisely, the following theorem holds.

Theorem 1. *Assume the splitting (57) with $\gamma \in [0, 1)$ defined by (58). Then the exact MSM converges in the energy norm with the rate γ^2, that is,*

$$\|u - u^{n+1}\|_a \le \gamma^2 \|u - u^n\|_a, \ n = 1, 2, \ldots \tag{64}$$

provided that $u^1 \in \mathbb{V}$ is chosen such that $z^1 = (I - P_2) z^0 \in \mathbb{V}_2^\perp$ (initial orthoprojection step onto $\mathbb{V}_2$). The convergence rate of the corresponding exact ASM with $\tau = \tau_{\mathrm{opt}} = 1$ is only γ, that is,

$$\|u - u^{n+1}\|_a \le \gamma \|u - u^n\|_a, \ n = 0, 1, 2, \ldots \tag{65}$$

In this case, the ASM preconditioner C has the form

$$\mathbf{C} = \mathbf{V}^{-T} \begin{pmatrix} \mathbf{V}_1^{\mathrm{T}} \mathbf{K} \mathbf{V}_1 & 0 \\ 0 & \mathbf{V}_2^{\mathrm{T}} \mathbf{K} \mathbf{V}_2 \end{pmatrix} \mathbf{V}^{-1} \tag{66}$$

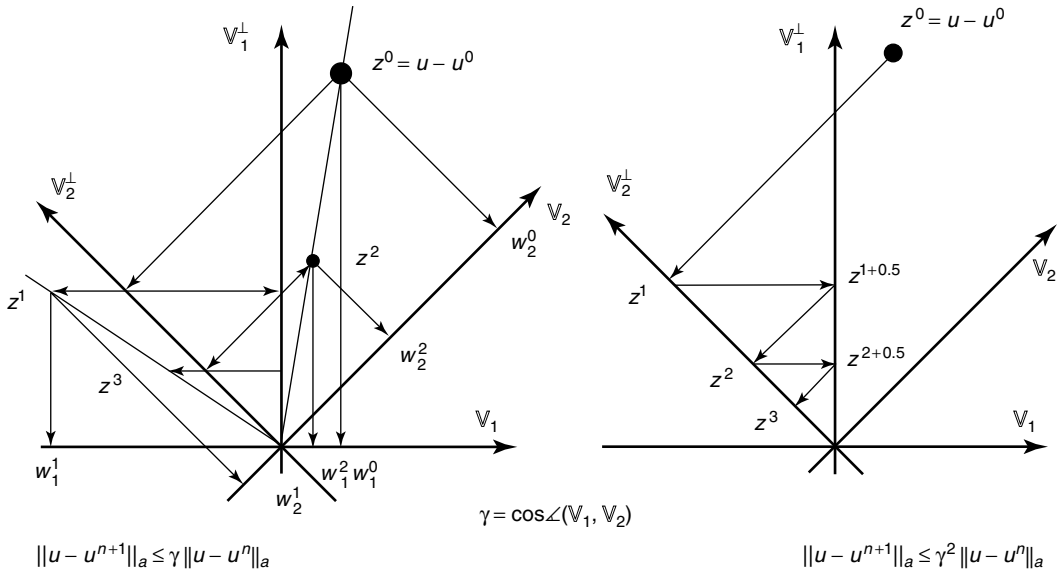

Figure 2. ASM and MSM corresponding to the splitting (57).

with the regular $N \times N$ basis transformation matrix $\mathbf{V} = (\mathbf{V}_1\mathbf{V}_2)$, and satisfies the spectral equivalence inequalities (39) with the sharp spectral equivalence constants $\underline{\gamma} = 1 - \gamma$ and $\overline{\gamma} = 1 + \gamma$.

The following straightforward spectral equivalence estimate for the inexact ASM preconditioner can be very useful in practice.

Corollary 1. *Let us assume that there are SPD subspace preconditioners $\mathbf{C}_1$ and $\mathbf{C}_2$ such that the spectral equivalence inequalities*

$$\underline{\gamma}_j \mathbf{C}_j \le \mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j \le \overline{\gamma}_j \mathbf{C}_j \tag{67}$$

hold with positive $\underline{\gamma}_j$ and $\overline{\gamma}_j$ for $j = 1, 2$. Then the inexact ASM preconditioner

$$\mathbf{C} = \mathbf{V}^{-\mathrm{T}} \begin{pmatrix} \mathbf{C}_1 & 0 \\ 0 & \mathbf{C}_2 \end{pmatrix} \mathbf{V}^{-1} \tag{68}$$

is spectrally equivalent to $\mathbf{K}$ in the sense of the spectral equivalence inequalities (39) with the spectral equivalence constants $\underline{\gamma} = \min\{\underline{\gamma}_1, \underline{\gamma}_2\}(1 - \gamma)$ and $\overline{\gamma} = \min\{\overline{\gamma}_1, \overline{\gamma}_2\}(1 + \gamma)$.

Corollary 1 implies that the inexact ASM converges for $\tau \in (0, 2/\overline{\gamma})$ with the rate given by (38). The inexact sMSM version is discussed in Haase and Langer (1992). Further results for the inexact ASM, MSM, and sMSM versions follow from the general case discussed in the next section. We additionally refer to Aronszjan (1950) and Bjørstad and Mandel (1991) for the special case of the splitting into two subspaces, including the case where $\mathbb{V}_1 \cap \mathbb{V}_2$ is nontrivial like in the classical alternating Schwarz method introduced in Section 3.

3.3.2 The general case: splitting into J subspaces

Let us first consider the inexact ASM and the inexact ASM preconditioner introduced in Section 3.2.1. It was pointed out in this section that the convergence analysis (cf. rate estimate (38)) aims at the spectral estimates (39)–(41) of the ASM preconditioned matrix $\mathbf{C}^{-1}\mathbf{K}$, or equivalently, of the ASM operator $\widetilde{P}$. Defining the so-called splitting norm

$$|||v|||^2 = \inf_{v = \sum_{j=1}^J v_j} \sum_{j=1}^J a_j(v_j, v_j) \tag{69}$$

where the infimum is taken over all possible splittings $v = \sum_{j=1}^J v_j$ of v with $v_j \in \mathbb{V}_j$, we obtain the following exact representation of the minimal and maximal eigenvalues of $\mathbf{C}^{-1}\mathbf{K}$ resp. $\widetilde{P}$.

Theorem 2.

$$\lambda_{\min}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\min}(\widetilde{P}) = \min_{v \in \mathbb{V}} \frac{a(v, v)}{|||v|||^2}$$

and

$$\lambda_{\max}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\max}(\widetilde{P}) = \max_{v \in \mathbb{V}} \frac{a(v, v)}{|||v|||^2} \tag{70}$$

Proof. follows from (39)–(41) and the observation that $|||v||| = a(\widetilde{P}^{-1}v, v)$ (see e.g. Bjørstad and Mandel, 1991; Xu, 1992; and Oswald, 1994). Theorem 2 is closely related to the so-called fictitious space lemma that has been used in the Russian literature since the eighties (Matsokin and Nepomnyaschikh, 1985; Nepomnyaschikh, 1990) (see also Oswald (1994) and Griebel and Oswald (1995) for this relation). □

The following two corollaries immediately follow from Theorem 2 and provide powerful tools for estimating (calculating) $\lambda_{\min}(\widetilde{P})$ and $\lambda_{\max}(\widetilde{P})$.

Corollary 2 (Lions' lemma, 1988). *Assume that there exists a positive constant c_L such that for all $v \in \mathbb{V}$ there exists at least one splitting $v = \sum_{j=1}^J v_j$ such that the inequality*

$$\sum_{j=1}^J a_j(v_j, v_j) \le c_L^2 a(v, v) \tag{71}$$

holds. Then

$$\lambda_{\min}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\min}(\widetilde{P}) \ge 1/c_L^2 \tag{72}$$

Proof. See also the original paper by Lions (1988). □

Corollary 3 (subspace interaction lemma). *Let us define the $J \times J$ subspace interaction matrix $\mathbf{\Gamma} = (\gamma_{ij})_{i,j=1,\ldots,J}$ with the coefficients γ_{ij} stemming from the supremum (cf. also (62))*

$$\gamma_{ij} = \sup_{v_i \in \mathbb{V}_i,\, v_j \in \mathbb{V}_j} \frac{a(v_i, v_j)}{\sqrt{a_i(v_i, v_i)}\sqrt{a_j(v_j, v_j)}} \tag{73}$$

or the generalized strengthened Cauchy inequalities

$$|a(v_i, v_j)| \le \gamma_{ij}\sqrt{a_i(v_i, v_i)}\sqrt{a_j(v_j, v_j)} \quad \forall v_i \in \mathbb{V}_i,\ \forall v_j \in \mathbb{V}_j \tag{74}$$

Then

$$\lambda_{\max}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\max}(\widetilde{P}) \le \rho(\mathbf{\Gamma}) \tag{75}$$

where $\rho(\mathbf{\Gamma})$ *denotes the spectral radius of the subspace interaction matrix* $\mathbf{\Gamma}$.

Proof. See Dryja and Widlund (1995). □

The convergence analysis of inexact multiplicative or hybrid versions of the Schwarz methods is more complicated. First of all, we need the so-called subspace contraction condition stating that the inequalities

$$a_j(\widetilde{P}_j v_j, v_j) \le \omega\, a_j(v_j, v_j) \quad \forall v_j \in \mathbb{V}_j, \quad \forall j = 1, 2, \ldots, J \tag{76}$$

hold for some constant $\omega \in (0, 2)$. Indeed, condition (76) ensures the contraction of the operator $I - \widetilde{P}_j$ on the subspace $\mathbb{V}_j$. We present here only two results concerning the inexact MSM and the inexact sMSM introduced in Section 3.2.2. For more results, we refer the reader to special papers on this topic, for example, Xu (1992), Dryja and Widlund (1995), and Griebel and Oswald (1995).

The sMSM produces a SPD preconditioner $\mathbf{C} = \mathbf{K}(\mathbf{I} - \mathbf{E})^{-1}$ that can be used in the PCG method for solving our FE system (23). The following theorem again provides bounds for the minimal and maximal eigenvalues of the preconditioned matrix $\mathbf{C}^{-1}\mathbf{K}$.

Theorem 3. *Let us assume that the space splitting is stable in the sense of Corollaries 2–3 and that the subspace contraction condition (76) holds for some constant* $\omega \in (1, 2)$. *Then the spectral estimates*

$$\lambda_{\min}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\min}(I - E) \ge \frac{2 - \omega}{\omega \rho^2(\mathbf{\Gamma}) c_L^2}$$

and

$$\lambda_{\max}(\mathbf{C}^{-1}\mathbf{K}) = \lambda_{\max}(I - E) \le 1 \tag{77}$$

hold with c_L *and* $\rho(\mathbf{\Gamma})$ *defined in Corollaries 2 and 3 respectively.*

Proof. See Smith, Bjørstad and Gropp (1996) for a slightly more general case, or the original paper by Dryja and Widlund (1995). □

The following theorem gives an exact representation of the convergence rate $\|E\|_a$ of the MSM in the energy norm, where E again denotes the MSM error propagation (iteration) operator defined by (46):

Theorem 4. *Let us again assume that the subspace contraction condition (76) holds for some constant* $\omega \in (0, 2)$. *Then the energy norm of the MSM error propagation operator E resp. of the MSM iteration matrix* $\mathbf{E}$ *can be represented in the form*

$$\|E\|_a = \|\mathbf{E}\|_K = q := \frac{c}{1 + c} < 1 \tag{78}$$

where

$$c = \sup_{\|v\|_a = 1} \inf_{v = \sum_{j=1}^{J} v_j} \sum_{i=1}^{J} a(\widetilde{P}_i \overline{P}_i^{-1} \widetilde{P}_i w_i, w_i) < \infty \tag{79}$$

with $w_i = \sum_{j=i}^{J} v_j - \widetilde{P}_i^{-1} v_i$ *and* $\overline{P}_i = \widetilde{P}_i^* + \widetilde{P}_i - \widetilde{P}_i^* \widetilde{P}_i = 2\widetilde{P}_i - \widetilde{P}_i^2$ $(\widetilde{P}_i = \widetilde{P}_i^*)$.

Proof. See Xu and Zikatanov (2002) for the more general case of infinite-dimensional Hilbert spaces and of nonsymmetric, but elliptic bilinear forms. □

In the case of the exact MSM, the representation can be simplified because $\widetilde{P}_i = P_i = P_i^* = P_i^2$ is an orthoprojection with respect to the energy inner product $a(\cdot, \cdot)$. More precisely, the constant c in Theorem 4 can directly be rewritten in the form

$$c = \sup_{\|v\|_a = 1} \inf_{v = \sum_{j=1}^{J} v_j} \sum_{i=1}^{J-1} \|P_i \sum_{j=i+1}^{J} v_j\|_a < \infty \tag{80}$$

Moreover, Theorem 4 immediately yields the convergence rate estimate

$$\begin{aligned} \|E_{\mathrm{sMSM}}\|_a &= \|E_{\mathrm{MSM}}^* E_{\mathrm{MSM}}\|_a \le \|E_{\mathrm{MSM}}^*\|_a \, \|E_{\mathrm{MSM}}\|_a \\ &= \|E_{\mathrm{MSM}}\|_a^2 = q^2 < 1 \end{aligned} \tag{81}$$

for sMSM, where E_{MSM} and E_{sMSM} denote the error propagation operators corresponding to the MSM and sMSM, respectively. Estimate (81) implies the spectral equivalence inequalities

$$(1 - q^2)\ \mathbf{C} \le \mathbf{K} \le 1\ \mathbf{C} \tag{82}$$

for the sMSM preconditioner $\mathbf{C} = \mathbf{K}(\mathbf{I} - \mathbf{E}_{\mathrm{sMSM}})^{-1}$. The abstract representations and estimates given above are very essential for obtaining spectral equivalence or convergence rate estimates for specific subspace correction methods in concrete applications.

4 OVERLAPPING DOMAIN DECOMPOSITION METHODS

4.1 Basic construction principles and algorithms with generous overlap

As explained in Sections 2 and 3, the overlapping DD methods have a long history and can be completely treated within the framework of the Schwarz theory. More precisely,

- the splitting of (FE) space $\mathbb{V} = \sum_{j=1}^{J} \mathbb{V}_j$,
- the subspace bilinear forms $a_j(\cdot, \cdot)$, and

- the arrangement of the projection-like operations (additive, multiplicative, hybrid)

completely define the Schwarz method (preconditioner), the complete analysis of which is also covered by the Schwarz theory presented in Section 3.3.

Without loss of generality, we restrict ourselves to the (exact) additive version (preconditioner) that is the most important one in parallel computing. Then the analysis can mainly be reduced to the verification of the conditions formulated in Corollaries 2 and 3, namely,

- the verification of the stability (71) of the space splitting and
- the calculation of the subspace interaction measure $\rho(\mathbf{\Gamma})$.

For definiteness, we consider the heat conduction problem described by Example 1 as a model problem and assume a moderate and smooth behavior of the heat conduction coefficients. In the overlapping DD method, the splitting $\mathbb{V} = \sum_{j=1}^{J} \mathbb{V}_j$ of a (FE) space $\mathbb{V} = \mathbb{V}_h(\Omega) \subset \mathbb{H}_0^1(\Omega)$ corresponds to an overlapping DD $\Omega = \bigcup_{j=1}^{J} \Omega_j$ of the computational domain Ω, where the subspaces usually have the form $\mathbb{V}_j = \mathbb{V} \cap \mathbb{H}_0^1(\Omega_j)$, that is, the subspace problems are local Dirichlet FE problems. There are several methods for constructing overlapping domain decompositions. Let us mention at first two simple techniques that both start from a coarse shape-regular conform triangulation $\mathcal{T}_H$ of $\overline{\Omega} = \bigcup_{j=1}^{\widetilde{J}} \overline{\tau}_{H,j}$ and proceed with its refinement, resulting in a shape-regular conform fine grid discretization $\mathcal{T}_h$ of $\overline{\Omega} = \bigcup_{j=1}^{\widetilde{J}} \bigcup_{r \in \mathcal{T}_{h,j}} \overline{\tau}_{h,j,r} = \bigcup_{r \in \mathcal{T}_h} \overline{\tau}_{h,r}$. For definiteness, we assume that these triangulations are provided by triangles ($d = 2$) or tetrahedra ($d = 3$), and, for simplicity, we also assume that linear elements are used for generating the FE spaces. The parameters H and h stand for the typical sizes of the coarse and the fine quasiuniform triangulations respectively. Now we associate with each coarse grid vertex $x^{(j)}$ ($j = 1, 2, \ldots, J$) some subdomain $\overline{\Omega}_j$ that is built by all coarse simplices containing $x^{(j)}$ as a vertex. This gives our first overlapping domain decomposition (ODD1) of Ω, where the overlap $\delta = O(H)$. Another one is given by associating with each coarse grid simplex $\overline{\tau}_{H,j}$ some subdomain $\overline{\Omega}_j$ that is built by this simplex and all simplices touching this simplex at least in one vertex, where $j = 1, 2, \ldots, J$ with $J = \widetilde{J}$. This again gives us an overlapping domain decomposition (ODD2) of Ω, where the overlap $\delta = O(H)$. Several generalizations of these techniques are feasible. For instance, one can first build a nonoverlapping domain decomposition with subdomains $\widetilde{\Omega}_j$ consisting of one or several coarse grid elements and then extend them by adding some layers of fine grid elements around these subdomains, giving the subdomains Ω_j of an overlapping domain decomposition (ODD(δ)), where δ is the thickness of the layers, that is, the overlap. Thus, ODD2 is a special ODD(H) method.

Let C be the (exact) additive Schwarz preconditioner (cf. (36))

$$\mathbf{C}^{-1} = \sum_{j=1}^{J} \mathbf{V}_j (\mathbf{V}_j^{\mathrm{T}} \mathbf{K} \mathbf{V}_j)^{-1} \mathbf{V}_j^{\mathrm{T}} = \sum_{j=1}^{J} \mathbf{V}_j \mathbf{K}_j^{-1} \mathbf{V}_j^{\mathrm{T}} \tag{83}$$

corresponding to the space splitting

$$\mathbb{V} = \mathbb{V}_h = \sum_{j=1}^{J} \mathbb{V}_j, \quad \mathbb{V}_j = \mathbb{V}_h \cap \mathbb{H}_0^1(\Omega_j) = \mathrm{span}\Phi\mathbf{V}_j \tag{84}$$

that is based on one of the overlapping domain decompositions $\bigcup_{j=1}^{J} \Omega_j$ of Ω with $O(H)$ overlap as described above, where the $N \times N_j$ matrix $\mathbf{V}_j$ picks exactly those basis functions from the fine grid nodal basis Φ, which belong to the inner nodes in Ω_j. The SPD $N_j \times N_j$ matrix $\mathbf{K}_j$ is nothing but the stiffness matrix belonging to the local FE Dirichlet problem in Ω_j. Using the general Schwarz theory, we can prove that

$$\kappa(C^{-1}K) = O(H^{-2}) \tag{85}$$

that is, owing to the generous $O(H)$ overlap, the relative spectral condition number does not depend on the fine grid discretization parameter h in a bad way, but on the domain decomposition parameter H. This is totally unacceptable for the use of this preconditioner in a massively parallel solver environment. The bad dependence on H is due to the absence of some coarse grid solver managing the global information transport that is essential for elliptic problems (see Widlund (1988) and Smith, Bjørstad and Gropp (1996) for a more detailed discussion of this issue in connection with DD methods).

There are a lot of possibilities to include such mechanisms for global information exchange into the Schwarz preconditioner. For the overlapping domain decompositions presented above, the natural way certainly consists in adding the coarse grid FE space $\mathbb{V}_0 = \mathbb{V}_H = \mathrm{span}\Phi\mathbf{V}_0 \subset \mathbb{V} \subset \mathbb{H}_0^1(\Omega)$ to the splitting (84), that is,

$$\mathbb{V} = \mathbb{V}_h = \mathbb{V}_0 + \sum_{j=1}^{J} \mathbb{V}_j = \sum_{j=0}^{J} \mathbb{V}_j \tag{86}$$

Now, the corresponding two-level (coarse level and fine level) overlapping ASM preconditioner

$$\mathbf{C}^{-1} = \sum_{j=0}^{J} \mathbf{V}_j \mathbf{K}_j^{-1} \mathbf{V}_j^{\mathrm{T}} \tag{87}$$

gives an optimal relative spectral condition number estimate.

Theorem 5. *The exact two-level ASM preconditioner (87) based on an overlapping domain decomposition with a uniform overlap width $O(H)$ provides an optimal preconditioner in the sense that $\kappa(C^{-1}K) \leq c$, where the positive constant c does not depend on h, H, and J.*

Proof. was given by Dryja and Widlund (1989) on the basis of the Schwarz theory and some technical lemmas (partition of unity, stability of $\mathbb{L}_2$-projection in $\mathbb{H}^1$) (see also Smith, Bjørstad and Gropp, 1996). □

The theorem remains obviously true for the inexact version where the local stiffness matrices $\mathbf{K}_j$ are replaced by suitable spectrally equivalent preconditioners $\mathbf{C}_j$ for $j = 1, 2, \ldots, J$. The analysis of multiplicative versions follows the same line of the general Schwarz theory (see e.g. Bramble *et al.*, 1991; Xu, 1992; and Xu and Zikatanov, 2002). The results can be extended to coarse grid spaces $\mathbb{V}_H$, which are not subspaces of the FE fine grid space $\mathbb{V} = \mathbb{V}_h$.

However, there are two drawbacks of two-level ASM preconditioners with a generous overlap. The first problem is connected with jumps in the coefficients. In contrast to the nonoverlapping DD methods (cf. Section 5), the influence of the jumps in the coefficients with large jumps is still not completely understood. The second problem is connected with the influence of the overlap width δ on $\kappa(C^{-1}K)$. It is clear that the larger the overlap, the more expensive are the local problems that we have to solve. The computational overhead becomes significant when h becomes small with respect to H. We discuss this problem in the next section.

4.2 Domain decomposition algorithms with a small overlap

In the case of a small overlap δ, Dryja and Widlund (1994) proved the following theorem.

Theorem 6. *The exact two-level ASM preconditioner (87) based on an overlapping domain decomposition with a uniform overlap width $O(\delta)$ gives the estimate*

$$\kappa(C^{-1}K) \leq \overline{c}\left(1 + \frac{H}{\delta}\right) \tag{88}$$

of the relative spectral condition number $\kappa(C^{-1}K)$, where the positive constant $\overline{c}$ does not depend on h, H, J, and δ.

Brenner (2000) showed that this result is sharp in the case of minimal overlap ($\delta = h$), that is, there exists a positive constant $\underline{c}$ that is independent of h, H, and J such that $\kappa(C^{-1}K) \geq \underline{c}(H/h)$. In the same paper, she proved that $\kappa(C^{-1}K) = O((H/h)^3)$ in the case of fourth-order elliptic boundary value problems.

Therefore, a small overlap really affects the preconditioning effect of the two-level ASM preconditioner (87) in a very bad way. On the other hand, the $O(H)$ overlap means that additional $O((H/h)^d)$ unknowns are added to the local problems in contrast to $O((H/h)^{d-1})$ unknowns in the case of an $O(h)$ overlap. Bank *et al.* (2002) have recently proposed a two-level hierarchical overlapping ASM preconditioner that adds only $O((H/h)^{d-1})$ unknowns to the local problems as in the case of small overlap, and that results in a uniformly bounded relative condition number estimate, as in the case of Theorem 6.

4.3 Multilevel versions

The two-level Schwarz methods, described above, use a fine (h) and a coarse (H) mesh capturing the local (high-frequency) and the global (low-frequency) parts in the solution (iteration error) respectively. This approach is satisfactory if efficient local and global solvers (preconditioners) are available. However, if the global problem is large, that is, H is relatively small, then we can again apply a two-level algorithm to the coarse grid problem using again some coarser grid. The recursive application to the coarse grid problems results in a multilevel ASM preconditioner. To be more precise, we assume that the coarse grid (triangulation) $\mathcal{T}_0 = \mathcal{T}_H$ is refined L times giving the finer and finer triangulations $\mathcal{T}_1, \ldots, \mathcal{T}_{L-1}$, and $\mathcal{T}_L = \mathcal{T}_h$. For each level $l = 1, 2, \ldots, L$ of the triangulations, with the exception of the coarsest level $l = 0$, we construct some overlapping domain decomposition $\Omega = \bigcup_{j=1}^{J_l} \Omega_{l,j}$ and connect with this multilevel overlapping domain decomposition the multilevel splitting of the FE space

$$\mathbb{V} = \mathbb{V}_h = \mathbb{V}_0 + \sum_{l=1}^{L} \sum_{j=1}^{J_l} \mathbb{V}_{l,j} \tag{89}$$

in the same way as above, where the subspaces $\mathbb{V}_{l,j} = \text{span}\Phi\mathbf{V}_{l,j}$ can again be generated by using the $N \times N_{l,j}$ basis transformation matrices $\mathbf{V}_{l,j}$. Now the corresponding (exact) multilevel overlapping ASM preconditioner can be written in the form

$$\mathbf{C}^{-1} = \mathbf{V}_0\mathbf{K}_0^{-1}\mathbf{V}_0^{\mathrm{T}} + \sum_{l=1}^{L} \sum_{j=1}^{J_l} \mathbf{V}_{l,j}\mathbf{K}_{l,j}^{-1}\mathbf{V}_{l,j}^{\mathrm{T}} \tag{90}$$

Let us consider one extreme case where the subdomains $\Omega_{l,j}$ are simply the supports of the nodal basis functions

$\phi_{l,j}$ belonging to the node $x_{l,j}$ in the l-level triangulation $\mathcal{T}_l$, that is, now the subspaces $\mathbf{V}_{l,j} = \mathrm{span}(\phi_{l,j}) = \mathrm{span}\Phi\mathbf{V}_{l,j}$ are one-dimensional, where $\Phi = \Phi_h = \Phi_L = [\phi_{L,j}]$ denotes the fine grid basis and the $N \times 1$ matrix $\mathbf{V}_{l,j}$ provides the representation of the basis function $\phi_{l,j}$ in the fine grid basis. We mention that this overlapping DD is closely related to ODD1, but now the elements around the node $x_{l,j}$ are taken from the triangulation $\mathcal{T}_l$ and not $\mathcal{T}_{l-1}$. The exact multilevel overlapping ASM preconditioner corresponding to this overlapping domain decomposition is called *multilevel diagonal scaling* (MDS) preconditioner and was introduced by Zhang (1992). Since for second-order elliptic problems the 1×1 matrices $\mathbf{K}_{l,j}$ obviously behave like h_l^{d-2}, the inexact version of the MDS preconditioner can be written in the form

$$\mathbf{C}^{-1} = \mathbf{V}_0\mathbf{K}_0^{-1}\mathbf{V}_0^{\mathrm{T}} + \sum_{l=1}^{L-1} h_l^{2-d} \sum_{j=1}^{J_l} \mathbf{V}_{l,j}\mathbf{V}_{l,j}^{\mathrm{T}} + h_L^{2-d} I \quad (91)$$

The inexact multilevel overlapping ASM preconditioner (91) was first proposed by Bramble, Pasciak and Xu (1990) and is nowadays known as BPX preconditioner.

Theorem 7. *The MDS and the BPX preconditioners are optimal preconditioners in the sense that there is some positive constant c that does not depend on h and L such that the relative spectral condition number $\kappa(C^{-1}K) \le c$, and the arithmetical cost ops$(C^{-1}\mathbf{d})$ for the preconditioning operation is proportional to the number of unknowns $N_h = O(h^{-d})$ on the finest grid.*

Proof. The original proof by Bramble, Pasciak and Xu (1990) provided weaker (nonoptimal) bounds depending on the number of refinement levels $L = O(\log(1 + (H/h)))$ (see also Zhang (1992) for the MDS preconditioner). The optimality of the BPX preconditioner was first proved by Oswald (1992) using Besov space techniques (see also Oswald, 1994). □

Multiplicative and hybrid versions of these multilevel Schwarz methods are closely related to multigrid methods, which are discussed in **Chapter 20, this Volume**; see also Bramble and Zhang (2000) for this relation.

5 NONOVERLAPPING DOMAIN DECOMPOSITION METHODS

5.1 Iterative substructuring methods

For definiteness, we consider the heat conduction problem described by Example 1 as model problem. In practice, the heat conduction coefficient $\alpha(\cdot)$ typically has jumps due to different materials. As in Section 2, we assume that the computational domain

$$\overline{\Omega} = \bigcup_{j=1}^{J} \overline{\Omega}_j \quad (92)$$

is decomposed into J nonoverlapping subdomains in such a way that the coefficient jumps are along with the boundary of the subdomains. For simplicity, we assume that in each subdomain Ω_j the coefficient $\alpha(\cdot)$ has the constant positive value α_j. Further, we assume that the domain decomposition is quasiregular in the sense that the subdomains are images of some reference domain, or a few reference domains, by a quasiregular mapping with the scaling H. Therefore, H can be viewed as typical subdomain diameter such that $J = O(H^{-d})$. As described in Section 2, we provide every subdomain with a triangulation $\mathcal{T}_j$ such that the triangulation $\mathcal{T}_h$ of the total computational domain $\overline{\Omega} = \bigcup_{j=1}^{J} \bigcup_{r \in \mathcal{T}_j} \overline{\tau}_r$ is conform and quasiregular in the sense that there is a typical element diameter h such that $N = N_h = O(h^{-d})$. Thus, the number of the internal subdomain unknowns N_{I_j} behaves like $O((H/h)^d)$. The FE discretization with the arrangement (12) of the FE basis Φ leads to the block structure (13) of the FE equations. The stiffness matrix $\mathbf{K}$ and the load vector $\mathbf{f}$ can obviously be represented in the form

$$\mathbf{K} = \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}}\mathbf{K}_j\mathbf{A}_j \quad \text{and} \quad \mathbf{f} = \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}}\mathbf{f}_j \quad (93)$$

where the $N_j \times N$ Boolean subdomain connectivity matrices $\mathbf{A}_j$ are mapping some vector $\mathbf{u} \in \mathbb{R}^N$ of all nodal values onto the vector $\mathbf{u}_j = \mathbf{A}_j\mathbf{u} \in \mathbb{R}^{N_j}$ of the subdomain nodal values. The $N_j \times N_j$ subdomain stiffness matrices $\mathbf{K}_j$ and the subdomain load vectors $\mathbf{f}_j = \mathbf{A}_j\mathbf{f} \in \mathbb{R}^{N_j}$ can be structured in the same way as we have structured $\mathbf{K}$ and $\mathbf{f}$ in (13), that is,

$$\mathbf{A}_j = \begin{pmatrix} \mathbf{A}_{C_j} & \mathbf{A}_{C_j I_j} \\ \mathbf{A}_{I_j C_j} & \mathbf{A}_{I_j} \end{pmatrix}$$

$$\mathbf{K}_j = \begin{pmatrix} \mathbf{K}_{C_j} & \mathbf{K}_{C_j I_j} \\ \mathbf{K}_{I_j C_j} & \mathbf{K}_{I_j} \end{pmatrix}, \quad \mathbf{f}_j = \begin{pmatrix} \mathbf{f}_{C_j} \\ \mathbf{f}_{I_j} \end{pmatrix} \quad (94)$$

The matrices $\mathbf{K}_{I_j}$ correspond to the local homogeneous Dirichlet problems, whereas the matrices $\mathbf{K}_j$ arise from the FE discretization of the local Neumann problems, at least, for the subdomains with $\partial\Omega_j \cap \partial\Omega = \emptyset$. For our model problem, these matrices are singular, where the kernel (null space) $\ker(\mathbf{K}_j) = \mathrm{span}(\mathbf{1}_j)$ is spanned by $\mathbf{1}_j = (1, 1, \ldots, 1) \in \mathbb{R}^{N_j}$.

The block Gaussian elimination of the internal subdomain unknowns $\mathbf{u}_I$ reduces the solution of the FE equation (13) to the solution of the Schur-complement problem

$$\mathbf{S}_C \mathbf{u}_C = \mathbf{g}_C \tag{95}$$

that was explicitly formed and directly solved in the classical substructuring Algorithm 3. The Schur complement $\mathbf{S}_C$ and the right-hand side $\mathbf{g}_C$ can be assembled from the local Schur complements $\mathbf{S}_{C_j}$ and the local right-hand sides $\mathbf{g}_{C_j}$ in the same way as $\mathbf{K}$ and $\mathbf{f}$ were assembled in (93) from $\mathbf{K}_j$ and $\mathbf{g}_j$ respectively, that is,

$$\mathbf{S}_C = \sum_{j=1}^{J} \mathbf{A}_{C_j}^{\mathrm{T}} \mathbf{S}_{C_j} \mathbf{A}_{C_j} \quad \text{and} \quad \mathbf{g}_C = \sum_{j=1}^{J} \mathbf{A}_{C_j}^{\mathrm{T}} \mathbf{g}_{C_j} \tag{96}$$

As mentioned in Section 2, the iterative solution of (95) avoids the very expensive forming of the Schur complement $\mathbf{S}_C$. Solving (95) by the PCG methods lead to our first iterative substructuring method called Schur-complement CG. In each iteration step of the Schur complement CG, we need one matrix-by-vector multiplication of the form

$$\begin{aligned} \mathbf{S}_C \mathbf{v}_C^n &= \sum_{j=1}^{J} \mathbf{A}_{C_j}^{\mathrm{T}} \mathbf{S}_{C_j} \mathbf{A}_{C_j} \mathbf{v}_C^n \\ &= \sum_{j=1}^{J} \mathbf{A}_{C_j}^{\mathrm{T}} (\mathbf{K}_{C_j} - \mathbf{K}_{C_j I_j} \mathbf{K}_{I_j}^{-1} \mathbf{K}_{I_j C_j}) \mathbf{A}_{C_j} \mathbf{v}_C^n \end{aligned} \tag{97}$$

requiring the direct solution of J systems (local Dirichlet problems)

$$\mathbf{K}_{I_j} \mathbf{w}_C^n = \mathbf{K}_{I_j C_j} \mathbf{A}_{C_j} \mathbf{v}_C^n, \quad j = 1, \ldots, J \tag{98}$$

which can be done completely in parallel. Moreover, the factorization of the matrices $\mathbf{K}_{I_j}$ in a preprocessing step and the use of sparse direct techniques can make this multiplication operation very efficient (see e.g. the classical monograph by George and Liu (1981) and more recent papers by Demmel *et al.* (1999) and Gupta (2002)). Nevertheless, for real large-scale problems, this operation is a bottleneck of the Schur complement CG. The use of inexact (iterative) solvers for the local Dirichlet problems (98) is dangerous. We discuss the naive use of inexact solvers in Section 5.3.

The forming of the Schur complement is some kind of a preconditioning operation. Indeed, the spectral condition number of the Schur complement $\kappa(\mathbf{S}_C) = \lambda_{\max}(\mathbf{S}_C)/\lambda_{\min}(\mathbf{S}_C) = O(H^{-1}h^{-1})$ is much better than the spectral condition number of the original stiffness matrix $\kappa(\mathbf{K}) = O(h^{-2})$ since $h << H$ (see e.g. Brenner, 1999). However, $\kappa(\mathbf{S}_C)$ still depends on the DD parameter H and the global discretization parameter h as well as on the coefficient jumps in a bad way. Thus, we need a SPD Schur-complement preconditioner $\mathbf{C}_C$ removing the influence of all these parameters in such a way that $\kappa(\mathbf{C}_C^{-1}\mathbf{S}_C)$ does not depend too much on these parameters under the restriction that the preconditioning operation $\mathbf{w}_C^n = \mathbf{C}_C^{-1}\mathbf{d}_C^n$, mapping the defect $\mathbf{d}_C^n$ into preconditioned defect $\mathbf{w}_C^n$, is sufficiently cheap. Since the Schur-complement preconditioner is one of the most important ingredients of many iterative substructuring methods, we discuss this topic in more detail in the next section.

In this chapter, we only consider iterative DD methods for h-versions of the FEM. The hp methods, by which we here mean both finite element and spectral element methods, are gaining growing attention due to the ability to attain exponential convergence even for problems with singularities. Although this noticeable advantage is often damaged by the high cost of the setup procedure caused by the high fill-in of the stiffness matrices and complex algorithms for calculating their entries, the computational cost grows with p algebraically at the worst. Therefore, there is a strong incentive to achieve a better performance by means of smart hp algorithms. The literature on hp methods is very vast. We refer the reader to **Chapter 5**, **Chapter 6 of this Volume** and **Chapter 3, Volume 3** for more information. In spite of the high interest, the toolkit of fast solvers for the systems arising from hp discretizations is much smaller than that for the h-version. In the last decade, the major progress in this area has been achieved on the basis of DD approaches. The formation of the basic features of hp DD algorithms is due to the contributions by Babuska, Craig, Mandel and Pitkäranta (1991), Pavarino (1994), Ivanov and Korneev (1995), Ainsworth (1996), Ivanov and Korneev (1996), Pavarino and Widlund (1996), Widlund (1996), Ainsworth and Senior (1997), Casarin (1997), Oden, Patra and Feng (1997), and Korneev and Jensen (1997). Let us mention that the spectrally equivalent finite-difference-like preconditioners for the local problems were suggested and studied by Orzag (1980), Bernardi, Dauge and Maday (1992), and Casarin (1997) for spectral discretizations, and by Ivanov and Korneev (1996) and Korneev and Jensen (1999) for hierarchical discretizations. Later studies paid more attention to a more elaborate design of all components of DD algorithms, that is, fast solvers for the local Dirichlet problems, efficient prolongations from the edges in 2D and from the faces in 3D cases, respectively, edge and face Schur-complement preconditioners, solvers for the wire-basket problem, and so on. In this relation, we refer to Korneev (2001, 2002a,b), Beuchler (2002), Korneev *et al.* (2002), and Korneev, Langer and Xanthis (2003). These studies resulted in fast

Dirichlet DD preconditioners for second-order elliptic equations. Some useful properties have been added to *hp* DD methods by the use of nonconforming discretizations and, in particular, by the mortar and FETI methods (see e.g. Bernardi, Maday and Sacchi-Landriani, 1989; Bernardi and Maday and Patera, 1993; Ben Belgacem and Maday, 1999; Bernardi, Maday and Sacchi-Landriani, 1989). Note that the components of the fast solvers developed for the Dirichlet DD preconditioners may be applied to these discretizations as well.

5.2 Schur-complement preconditioners

In this section, we look for SPD Schur-complement preconditioners $\mathbf{C}_C$ satisfying the following conditions:

1. Spectral equivalence condition: the spectral equivalence inequalities

$$\underline{\gamma}_C \mathbf{C}_C \le \mathbf{S}_C \le \overline{\gamma}_C \mathbf{C}_C \tag{99}$$

   should hold with positive spectral equivalence constants $\underline{\gamma}_C$ and $\overline{\gamma}_C$ such that $\kappa(\mathbf{C}_C^{-1}\mathbf{S}_C) \le \overline{\gamma}_C/\underline{\gamma}_C$ does not, or only weakly depend on h, H, and the coefficient jumps. The latter property is sometimes also called robustness with respect to coefficient jumps.
2. Efficiency condition: the number of arithmetical operations $\mathrm{ops}(\mathbf{C}_C^{-1}\mathbf{d}_C^n)$ needed for the preconditioning operation should be of the order $O(N_C) \dots O(N)$, or at least should not disturb the overall complexity of the algorithm too much.
3. Parallelizability condition: the preconditioning operation $\mathbf{C}_C^{-1}\mathbf{d}_C^n$ should not disturb the numerical and parallel efficiency of the total algorithm too much. However, we should be aware that in many Schur-complement preconditioners some coarse grid solver managing the global information transport is hidden. Thus, the coarse grid solver requires global communication and is some bottleneck in the parallelization.

In the literature, there are several basic proposals for Schur-complement preconditioners. Many of them are based on the fact that the Schur-complement energy $(\mathbf{S}_C\mathbf{u}_C, \mathbf{u}_C)$ is spectrally equivalent to the broken-weighted $H^{1/2}$-norm

$$\|u_C\|^2_{H^{1/2}(\Gamma_C)} := \sum_{j=1}^{J} \alpha_j |u_{C_j}|^2_{H^{1/2}(\Gamma_j)}$$
$$= \sum_{j=1}^{J} \alpha_j \int_{\Gamma_j}\int_{\Gamma_j} \frac{|u_{C_j}(y) - u_{C_j}(x)|^2}{|y-x|^d}\, \mathrm{d}s_y\, \mathrm{d}s_x \tag{100}$$

with $\Gamma_i = \partial\Omega_i$, that is, there exist positive constants $\underline{\delta}_C$ and $\overline{\delta}_C$, which are independent of h, H, and the coefficient jumps, such that

$$\underline{\delta}_C \|u_C\|^2_{H^{1/2}(\Gamma_C)} \le (\mathbf{S}_C\mathbf{u}_C, \mathbf{u}_C)$$
$$\le \overline{\delta}_C \|u_C\|^2_{H^{1/2}(\Gamma_C)} \quad \forall u_C \leftrightarrow \mathbf{u}_C \in \mathbb{R}^{N_C} \tag{101}$$

Evidently,

$$\|\mathbf{u}_C\|^2_{\mathbf{S}_C} = (\mathbf{S}_C\mathbf{u}_C, \mathbf{u}_C)$$
$$= \sum_{j=1}^{J} \alpha_j (\overline{\mathbf{S}}_{C_j}\mathbf{u}_{C_j}, \mathbf{u}_{C_j}) = \sum_{j=1}^{J} \alpha_j \inf_{v_j|_{\Gamma_j} = u_{C_j}} \|\nabla v_j\|^2_{\mathbb{L}_2(\Omega_j)} \tag{102}$$

where the subdomain Schur complements $\overline{\mathbf{S}}_{C_j}$ arise from the case $\alpha_j = 1$ and the infimum is taken over all FE functions $v_j \in \mathbb{V}_j = V|_{\overline{\Omega}_j}$ living in $\overline{\Omega}_j$ and coinciding with u_{C_j} on Γ_j. Therefore, the equivalence (101) is the same as the equivalence of the infimum in (102) to the $H^{1/2}(\Gamma_j)$ – norm in (100) and requires the trace and lifting (prolongation) theorems for functions from the FE space. In simple cases, the left inequality in (101) is an obvious consequence of the trace theorem for functions from $H^1(\Omega_j)$. The right inequality in (101) requires some special proof, which, for example, may be found in Nepomnyaschikh (1991b).

Let us first describe Dryja's classical Schur-complement preconditioner that is just applicable to the L-shaped domain sketched in Figure 1. Here, the interface Γ_C consists only of one straight piece with N_C (inner) nodal points. For this simple but characteristic example, Dryja (1982) proposed the preconditioner

$$\mathbf{C}_C = \mathbf{B}_C^{1/2}$$
$$= \begin{pmatrix} 2 & -1 & & & \\ -1 & 2 & -1 & & \\ & \ddots & \ddots & \ddots & \\ & & -1 & 2 & -1 \\ & & & -1 & 2 \end{pmatrix}^{1/2} = \mathbf{F}_C^{\mathrm{T}}\, \mathbf{\Lambda}_C^{1/2}\, \mathbf{F}_C \tag{103}$$

that is nothing but the square root of the scaled discretized 1D Laplacian $\mathbf{B}_C$ along the interface under homogeneous Dirichlet boundary conditions at the end points of the interface. Dryja (1982) proved the spectral equivalence inequalities (99) with h-independent spectral equivalence constants $\underline{\gamma}_C$ and $\overline{\gamma}_C$ (see also theorem 1 in Dryja, 1984). It is well known that $\mathbf{B}_C$ has the eigenvalues $\lambda_k = 4\sin^2(k\pi/(2(N_C+1)))$ and the corresponding orthonormal eigenvectors $\mathbf{v}_{C,k} = \left[\sqrt{(2/(N_C+1))}\sin(k\pi l/(N_C+1))\right]_{l=1,\dots,N_C}$, where k is running from 1 to N_C. Therefore, in (103), $\mathbf{\Lambda}_C^{1/2} =$

$\mathrm{diag}(\lambda_k^{1/2})$ and $\mathbf{F}_C = [\mathbf{v}_{C,1}, \ldots, \mathbf{v}_{C,N_C}]$. Since the Fourier matrix $\mathbf{F}_C$ is orthogonal, that is, $\mathbf{F}_C^{-1} = \mathbf{F}_C^{\mathrm{T}}$, the preconditioning equation $\mathbf{C}_C \mathbf{w}_C = \mathbf{d}_C$ can be solved in the following three steps:

$$\begin{aligned} \text{Fourier analysis:} \quad & \mathbf{x}_C = \mathbf{F}_C \mathbf{d}_C \\ \text{Diagonal scaling:} \quad & \mathbf{y}_C = \mathbf{\Lambda}_C^{-1/2} \mathbf{x}_C \\ \text{Fourier synthesis:} \quad & \mathbf{w}_C = \mathbf{F}_C^{\mathrm{T}} \mathbf{y}_C \end{aligned}$$

The complexity of this preconditioning operation is dominated by the Fourier analysis and Fourier synthesis. Using the fast Fourier transformation (FFT), the total complexity of the preconditioning operation is of the order $N_C \ln(N_C)$. Dryja's preconditioner was improved by Golub and Mayers (1984), Bjørstad and Widlund (1986), Chan (1987), and others.

The construction of good Schur-complement preconditioners in the general case of many nonoverlapping subdomains is more involved. In the two-dimensional case, Bramble, Pasciak and Schatz (1986) proposed an ASM preconditioner of the form

$$\mathbf{C}_C^{-1} = \sum_E \mathbf{V}_E \mathbf{C}_E^{-1} \mathbf{V}_E^{\mathrm{T}} + \sum_V \mathbf{V}_V \mathbf{C}_V^{-1} \mathbf{V}_V^{\mathrm{T}} \tag{104}$$

in which the vertices are separated from the edges of the subdomains arising from some coarse grid domain decomposition. Here, $\mathbf{V}_E^{\mathrm{T}}$ takes the edge nodal basis functions (unknowns) belonging to the edge E from the nodal basis functions (unknowns), $\mathbf{V}_V^{\mathrm{T}}$ transforms the nodal basis functions into the coarse grid (hierarchical) vertex basis functions, $\mathbf{C}_E$ is an edge Schur-complement preconditioner as discussed above (Dryja's type), and $\mathbf{C}_V$ denotes a coarse grid preconditioner, for example, $\mathbf{C}_V$ can be the coarse grid stiffness matrix. The latter one manages the global information exchange. Bramble, Pasciak and Schatz (1986) proved that the relative condition number $\kappa(\mathbf{C}_C^{-1}\mathbf{S}_C)$ grows at most like $(1 + \ln(H/h))^2$ as h tends to zero. Moreover, the right averaging of the coefficients of adjacent subdomains makes the BPS preconditioner, how the Schur-complement preconditioner (104) is nowadays called, quite robust with respect to coefficient jumps. In a series of papers, Bramble, Pasciak and Schatz (1986, 1987, 1988, 1989) generalized these ideas in many directions, including Schur-complement preconditioners for the 3D case.

Another type of Schur-complement preconditioners relies on the transformation of the nodal basis to a multilevel basis (generating system). This was very successful in constructing preconditioners for the finite element stiffness matrix $\mathbf{K}$ (the hierarchical preconditioner of Yserantant and the BPX preconditioner). It is easy to see that the same transformations restricted to the coupling boundary (interface) nodes result in Schur-complement preconditioners possessing at least the same quality as the corresponding preconditioners for $\mathbf{K}$. Smith and Widlund (1990) and Haase, Langer and Meyer (1991) proposed hierarchical Schur-complement preconditioners, which asymptotically behave like the BPS preconditioner in the 2D case. The BPX Schur-complement preconditioner was introduced by Tong, Chan and Kuo (1991). It is asymptotically optimal, but is sensitive to coefficient jumps.

There are a lot of other proposals for constructing Schur-complement preconditioners. Let us here mention only the wire-basket-based Schur-complement preconditioners introduced by Dryja, Smith and Widlund (1994), the probing technique proposed by Chan and Mathew (1994) (see also Keyes and Gropp, 1987), and the techniques borrowed from the boundary element method (see e.g. Carstensen, Kuhn and Langer, 1998; Haase *et al.*, 1997; and Steinbach, 2003).

Finally, we refer to the Neumann–Dirichlet and to the Neumann–Neumann preconditioners, which are special Schur-complement preconditioners approved to be very robust in practical applications. The Neumann–Neumann preconditioners are discussed in Section 5.4 in detail.

5.3 Inexact subdomain solvers

5.3.1 Effects of inexact subdomain solvers

Let us consider the discrete harmonic ($a(\cdot,\cdot)$-harmonic) splitting

$$\mathbb{V} = \mathbb{V}_h = \mathbb{V}_C^* \oplus \mathbb{V}_I \tag{105}$$

of the FE space $\mathbb{V}$ into the discrete harmonic space

$$\begin{aligned} \mathbb{V}_C^* &= P^* \mathbb{V}|_{\Gamma_C} = \{u \in \mathbb{V} : a(u, v) = 0 \; \forall v \in \mathbb{V}_I\} \\ &= \mathrm{span}\Psi_C^* = \mathrm{span}\Phi \mathbf{V}_C^* \end{aligned} \tag{106}$$

and the interior subdomain (bubble) space $\mathbb{V}_I = \mathrm{span}\Phi\mathbf{V}_I = \mathbb{V}_{I_1} \oplus \ldots \oplus \mathbb{V}_{I_J}$, with the basis transformation matrices

$$\begin{aligned} \mathbf{V}_C^* &= \begin{pmatrix} \mathbf{I}_C \\ -\mathbf{K}_I^{-1}\mathbf{K}_{IC} \end{pmatrix}_{N \times N_C} \\ &= \begin{pmatrix} \mathbf{I}_C \\ \mathbf{P}_{IC}^* \end{pmatrix}_{N \times N_C} \quad \text{and} \quad \mathbf{V}_I = \begin{pmatrix} \mathbf{0} \\ \mathbf{I}_I \end{pmatrix}_{N \times N_I} \end{aligned} \tag{107}$$

where $\Psi_C^* = \Phi \mathbf{V}_C^*$ is called the discrete harmonic basis and P^* is the discrete harmonic extension (prolongation) operator mapping a FE function u_C living on Γ_C to some FE function $P^* u_C$ that coincides with u_C on Γ_C and is discrete harmonic in all subdomains Ω_j. Owing to the

construction of $\mathbb{V}_C^*$, the splitting is orthogonal with respect to the energy inner product $a(\cdot,\cdot)$. Therefore, the exact ASM preconditioner

$$\mathbf{C} = \begin{pmatrix} \mathbf{I}_C & \mathbf{K}_{CI}\mathbf{K}_I^{-1} \\ \mathbf{0} & \mathbf{I}_I \end{pmatrix} \begin{pmatrix} \mathbf{S}_C & \mathbf{0} \\ \mathbf{0} & \mathbf{K}_I \end{pmatrix} \begin{pmatrix} \mathbf{I}_C & \mathbf{0} \\ \mathbf{K}_I^{-1}\mathbf{K}_{IC} & \mathbf{I}_I \end{pmatrix} = \mathbf{K} \tag{108}$$

must coincide with $\mathbf{K}$. This factorization of $\mathbf{K}$ is used in many applications. Replacing $\mathbf{S}_C$ by a Schur complement preconditioner $\mathbf{C}_C$ (cf. Section 5.3), we arrive at a partially inexact ASM preconditioner that corresponds to the Schur complement iteration methods. Replacing additionally $\mathbf{K}_I$ by some preconditioner $\mathbf{C}_I$ gives us the full inexact ASM preconditioner. Owing to Corollary 1, both inexact ASM preconditioners are covered by the general theory with $\gamma = \cos \measuredangle (\mathbb{V}_C^*, \mathbb{V}_I) = 0$. The naive replacement of $\mathbf{K}_I^{-1}$ in the left and right factors of the factorization (108) is dangerous because it changes the angle between the subspaces. For instance, the 'inversion' of $\mathbf{K}_I$ by one multigrid cycle, that is, replacement of $\mathbf{K}_I^{-1}$ by $(\mathbf{I}_I - \mathbf{E}_I^s)\mathbf{K}_I^{-1}$ $(s = 1)$, will, in general, not result in a stable splitting even if the multigrid convergence rate, that is, the energy norm of the multigrid iteration operator $\|\mathbf{E}_I\|_{\mathbf{K}_I}$, is less than 1 independently of the local discretization parameter H/h (as usual). One needs at least $s = O(\ln(H/h))$ multigrid cycles to get a stable extension (see e.g. Haase, Langer and Meyer, 1991). The next section shows that a stable splitting requires a careful choice of the extension operator $\widetilde{P}$ resp. $\widetilde{\mathbf{P}}_{IC}$ replacing the discrete harmonic extension operator P^* resp. $\mathbf{P}_{IC}^*$.

5.3.2 *The bounded extension splitting*

Let us now consider the bounded extension splitting

$$\mathbb{V} = \mathbb{V}_h = \widetilde{\mathbb{V}}_C + \mathbb{V}_I \tag{109}$$

of the FE space $\mathbb{V}$ in the direct sum of the former subdomain space $\mathbb{V}_I$ that remains unchanged and the bounded extension space

$$\begin{aligned} \widetilde{\mathbb{V}}_C &= \widetilde{P}\mathbb{V}|_{\Gamma_C} = \{\widetilde{P}u \in \mathbb{V} : u_C = u|_{\Gamma_C} \in \mathbb{V}|_{\Gamma_C} \text{ given}\} \\ &= \operatorname{span}\widetilde{\Psi}_C = \operatorname{span}\Phi\widetilde{\mathbf{V}}_C \end{aligned} \tag{110}$$

with the basis transformation matrices

$$\widetilde{\mathbf{V}}_C = \begin{pmatrix} \mathbf{I}_C \\ \widetilde{\mathbf{P}}_{IC} \end{pmatrix}_{N\times N_C} : \mathbb{R}_{N_C} \to \mathbb{R}_N \tag{111}$$

We assume that there exists some constant $c_E^2 \geq 1$ independent of our bad parameters such that

$$\left\| \begin{pmatrix} \mathbf{I}_C \\ \widetilde{\mathbf{P}}_{IC} \end{pmatrix} \mathbf{u}_C \right\|_{\mathbf{K}} \leq c_E \|\mathbf{u}_C\|_{\mathbf{S}_C} \quad \forall \mathbf{u}_C \in \mathbb{R}^{N_C} \tag{112}$$

Together with the minimal energy property (102), inequality (108) gives the following spectral equivalence relations

$$\begin{aligned} (\mathbf{S}_C\mathbf{u}_C, \mathbf{u}_C) \leq (\widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C\mathbf{u}_C, \mathbf{u}_C) \leq c_E^2(\mathbf{S}_C\mathbf{u}_C, \mathbf{u}_C) \\ \forall \mathbf{u}_C \in \mathbb{R}^{N_C} \end{aligned} \tag{113}$$

which are the matrix form of the inequalities

$$\begin{aligned} a(P^*u_C, P^*u_C) &\leq a(\widetilde{P}u_C, \widetilde{P}u_C) \\ &\leq c_E^2 a(P^*u_C, P^*u_C) \quad \forall u_C \in \mathbb{V}|_{\Gamma_C} \end{aligned} \tag{114}$$

Of course, replacing P^* by $\widetilde{P}$, we lose the orthogonality. However, the following lemma shows that the bounded extension gives us a stable splitting.

Lemma 2. *If (108) holds, then*

$$\gamma = \cos \measuredangle (\widetilde{\mathbb{V}}_C, \mathbb{V}_I) \leq \sqrt{1 - c_E^{-2}} < 1 \tag{115}$$

Proof. Follows immediately from relation (62) in Lemma 1 (see e.g. Haase *et al.*, 1994). □

The sharp (minimal) constant c_E in (112)–(113) providing also the sharp constant in (115) is given by the maximal eigenvalue $\lambda_{\max}$ of the generalized eigenvalue problem $\widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C\mathbf{u}_C = \lambda\mathbf{S}_C\mathbf{u}_C$. Now, we can summarize our results in the following theorem.

Theorem 8. *Assume that there is some bounded extension* $\widetilde{\mathbf{P}}_{IC} : \mathbb{R}_{N_C} \to \mathbb{R}_{N_I}$, *satisfying (112) with some constant* $c_E \geq 1$ *and that there are SPD preconditioners* $\mathbf{C}_C$ *and* $\mathbf{C}_I$ *for* $\widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C$ *and* $\mathbf{K}_I$ *respectively, that is, there are positive spectral equivalence constants* $\underline{\gamma}_C$, $\overline{\gamma}_C$, $\underline{\gamma}_I$, *and* $\overline{\gamma}_I$ *such that*

$$\begin{aligned} \underline{\gamma}_C\mathbf{C}_C \leq \widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C \leq \overline{\gamma}_C\mathbf{C}_C \\ \text{and} \\ \underline{\gamma}_I\mathbf{C}_I \leq \mathbf{K}_I \leq \overline{\gamma}_I\mathbf{C}_I \end{aligned} \tag{116}$$

Then, the inexact ASM preconditioner

$$\mathbf{C} = \begin{pmatrix} \mathbf{I}_C & -\widetilde{\mathbf{P}}_{IC}^{\mathrm{T}} \\ 0 & \mathbf{I}_I \end{pmatrix} \begin{pmatrix} \mathbf{C}_C & 0 \\ 0 & \mathbf{C}_I \end{pmatrix} \begin{pmatrix} \mathbf{I}_C & 0 \\ -\widetilde{\mathbf{P}}_{IC} & \mathbf{I}_I \end{pmatrix} \tag{117}$$

based on the bounded extension splitting (105) is spectrally equivalent to $\mathbf{K}$*, that is,*

$$\underline{\gamma}\ \mathbf{C} \leq \mathbf{K} \leq \overline{\gamma}\ \mathbf{C} \tag{118}$$

with the spectral equivalence constants

$$\underline{\gamma} = \min\{\underline{\gamma}_C, \underline{\gamma}_I\}\left(1 - \sqrt{1 - c_E^{-2}}\right)$$

and

$$\overline{\gamma} = \max\{\overline{\gamma}_C, \overline{\gamma}_I\}\left(1 + \sqrt{1 - c_E^{-2}}\right) \tag{119}$$

Proof. Follows easily from Corollary 1 and Lemma 2. □

Theorem 8 provides us with a guide for choosing the ingredients $\mathbf{C}_I$, $\mathbf{C}_C$ and $\widetilde{\mathbf{P}}_{IC}$ of the inexact ASM preconditioner (116) in such a way that the final spectral equivalence constants $\underline{\gamma}$ and $\overline{\gamma}$ in (113) do not depend on h, H and the jumps of coefficient too much. Optimal ingredients will lead to an optimal inexact ASM preconditioner. Let us summarize some concrete proposals for choosing $\mathbf{C}_I$, $\mathbf{C}_C$, and $\widetilde{\mathbf{P}}_{IC}$:

Preconditioners $\boldsymbol{C}_{I_j}$ *for the local Dirichlet problems*
Since $\mathbf{C}_I = \mathrm{diag}(\mathbf{C}_{I_j})_{j=1,\ldots,J}$, we only need good preconditioners $\mathbf{C}_{I_j}$ for the $\mathbf{K}_{I_j}$, arising from the FE discretization of the local Dirichlet problems where the coefficients of the PDE are changing only smoothly. Nowadays, a lot of optimal (linear complexity) preconditioners for such problems are available (see e.g. **Chapter 20, this Volume** or Bramble and Zhang, 2000). For instance, local multigrid preconditioners of the form

$$\mathbf{C}_{I_j} = \mathbf{K}_{I_j}(\mathbf{I}_{I_j} - \mathbf{E}_{I_j}^k)^{-1} \tag{120}$$

will do a good job, where $\mathbf{E}_{I_j}$ denotes the corresponding multigrid iteration operator. They can be generated, for example, by one ($k = 1$) symmetric V-cycle with appropriately chosen multigrid components such that $\mathbf{C}_{I_j}$ is SPD (Jung and Langer, 1991). Since we can assume that the multigrid iteration operators $\mathbf{E}_{I_j}$ are nonnegative with respect to the energy inner product and the multigrid rates $\|\mathbf{E}_{I_j}\|_{\mathbf{K}_{I_j}}$ are bounded by some mesh-independent constant $\eta < 1$, we see that the second spectral inequalities in (116) are fulfilled by $\underline{\gamma}_I = 1 - \eta^k$ and $\overline{\gamma}_I = 1$. The operation count for the local preconditioning operation gives $\mathrm{ops}(\mathbf{C}_{I_j}^{-1}\mathbf{d}_{I_j}) = O(N_{I_j}) = O((H/h)^d)$. Whereas various optimal preconditioners are available in the h-version of the FEM, the situation is quite different for the hp-version. Examples of optimal, or at least almost optimal local hp preconditioners can be found in Korneev and Jensen (1999), Korneev (2001, 2002a), Beuchler (2002), and Korneev, Langer and Xanthis (2003).

Preconditioners $\boldsymbol{C}_C$ *for* $\widetilde{\boldsymbol{V}}_C^{\mathrm{T}}\boldsymbol{K}\widetilde{\boldsymbol{V}}_C$
Owing to the spectral relations (113), every good Schur-complement preconditioner $\mathbf{C}_C$, given in Section 5.2, is also a good preconditioner for $\widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C$ provided that the extension constant c_E is small. More precisely, the spectral equivalence inequalities $\underline{\delta}_C\mathbf{C}_C \leq \mathbf{S}_C \leq \overline{\delta}_C\mathbf{C}_C$ and (109) give us the first spectral inequalities in (116) with $\underline{\gamma}_C = \underline{\delta}_C$ and $\overline{\gamma}_C = c_E^2\overline{\delta}_C$. On the other hand, one can again construct preconditioners of the form

$$\mathbf{C}_C = \widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C(\mathbf{I}_C - \mathbf{E}_C^s)^{-1} \tag{121}$$

applying s iteration steps of some symmetric internal iteration method (e.g. s symmetric V-cycles) with the linear iteration operator $\mathbf{E}_C$ directly to the ASM subspace matrix $\widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C$. Mention that the discrete bounded extension operations discussed below are cheap such that in turn the cost for the matrix-by-vector multiplication $\widetilde{\mathbf{V}}_C^{\mathrm{T}}\mathbf{K}\widetilde{\mathbf{V}}_C\mathbf{d}_C$ is proportional to N.

Discrete bounded extension
Owing to the equivalence

$$\underline{\theta}\,|u|^2_{\mathbb{H}^{1/2}(\Gamma_j)} \leq \inf_{v\in\mathbb{H}^1(\Omega_j):v=u|_{\Gamma_j}} \|\nabla v\|^2_{\mathbb{L}_2(\Omega_j)}$$
$$\leq \overline{\theta}\,|u|^2_{\mathbb{H}^{1/2}(\Gamma_j)} \quad \forall u \in \mathbb{H}^{1/2}(\Gamma_j) \tag{122}$$

and to (102), we can reduce the construction of a bounded extension operator $\widetilde{P}$ to the construction of a local bounded extension operator $\widetilde{P}_j : \mathbb{V}|_{\Gamma_j} \subset \mathbb{H}^{1/2}(\Gamma_j) \to \mathbb{V}|_{\overline{\Omega}_j} \subset \mathbb{H}^1(\Omega_j)$ such that the inequality

$$|\widetilde{P}_j u|_{\mathbb{H}^1(\Omega_j)} = \|\nabla\widetilde{P}_j u\|_{\mathbb{L}_2(\Omega_j)} \leq \widetilde{c}_E\,|u|_{\mathbb{H}^{1/2}(\Gamma_j)} \quad \forall u \in \mathbb{V}|_{\Gamma_j} \tag{123}$$

is valid for some positive constant $\widetilde{c}_E$. If $\widetilde{c}_E$ is independent of H/h, then the extension constant $c_E^2 = \widetilde{c}_E^2/\underline{\theta}$ does not depend on these parameters either.

Let us now review some cheap bounded discrete extension procedures of that kind. The first computable extension procedure was proposed by Matsokin and Nepomnyaschikh (1985) on the basis of some averaging technique that provides a uniform bound (see also Nepomnyaschikh, 1991a). Haase *et al.* (1994) introduced the hierarchical extension that is very cheap. In 2D, it leads to a $\ln(H/h)$ growth of c_E that can be compensated by $O(\ln\ln(H/h))$ multigrid iterations. In 3D, the hierarchical extension is too weak. However, the multilevel extension that was proposed by Haase and Nepomnyaschikh (1997) works fine in 2D as well as in 3D.

5.4 Neumann–Neumann preconditioners

Bourgat *et al.* (1989) introduced the Neumann–Neumann Schur-complement preconditioner (cf. also Section 5.2) $\mathbf{C}_C^{-1} = (1/4)\mathbf{S}_{C_1}^{-1} + (1/4)\mathbf{S}_{C_2}^{-1}$ for the case of two subdomains ($J = 2$) and showed that $\kappa(\mathbf{C}_C^{-1}\mathbf{S}_C) = O(1)$. The operation $\mathbf{w}_{C_j} = \mathbf{S}_{C_j}^{-1}\mathbf{d}_{C_j}$ ($\mathbf{d}_{C_1} = \mathbf{d}_{C_2} = \mathbf{d}_C$) is obviously equivalent to solution of the system

$$\begin{pmatrix} \mathbf{K}_{C_j} & \mathbf{K}_{C_j I_j} \\ \mathbf{K}_{I_j C_j} & \mathbf{K}_{I_j} \end{pmatrix} \begin{pmatrix} \mathbf{w}_{C_j} \\ \mathbf{w}_{I_j} \end{pmatrix} = \begin{pmatrix} \mathbf{d}_{C_j} \\ \mathbf{O}_{I_j} \end{pmatrix} \tag{124}$$

that corresponds to the Neumann boundary condition on $\Gamma_C = \partial\Omega_1 \cap \partial\Omega_2$ for $j = 1, 2$.

De Roeck and Le Tallec (1991) generalized the Neumann–Neumann preconditioner to the general case of J subdomains. To simplify the notation, we skip the subindex C, that is, $\mathbf{S}_j = \mathbf{S}_{C_j}$, $\mathbf{A}_j = \mathbf{A}_{C_j}$, and so on. In order to weight the contributions from the different subdomains Ω_j, we introduce the weight matrices $\mathbf{D}_j$ such that $\sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}}\mathbf{D}_j\mathbf{A}_j = \mathbf{I}_j$. There are different ways to choose these weights. Following Mandel and Brezina (1996), we define $\mathbf{D}_j$ as the diagonal matrix $\mathrm{diag}(d_l^j)_{l=1:N_j}$ with the diagonal entries

$$d_l^j = \alpha_j / \sum_{i:\, x_l \in \partial\Omega_i} \alpha_i \tag{125}$$

This choice avoids the dependence of the condition number on the coefficient jumps in the PDE. In the case that all $\alpha_j = 1$ (Poisson equation), the diagonal entry d_l^j is equal to the reciprocal of the number of subdomains meeting at the nodal point x_l to which the diagonal entry d_l^j belongs. Similar to the case of two subdomains, we can now write the multisubdomain Neumann–Neumann preconditioner $\mathbf{C}$ in the form

$$\mathbf{C}^{-1} = \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}}\mathbf{D}_j^{\mathrm{T}}\mathbf{S}_j^{-1}\mathbf{D}_j\mathbf{A}_j \tag{126}$$

that is again an additive Schwarz preconditioner. The algorithmic form of the preconditioning operation $\mathbf{w} = \mathbf{C}^{-1}\mathbf{d}$ is given in Algorithm 7. In general ($\partial\Omega_j \cap \Gamma_D = \emptyset$), the local Schur complements $\mathbf{S}_j$ as well as the corresponding subdomain stiffness matrices in (124) are singular because they are derived from pure local Neumann problems. Thus, the kernels correspond to the functions that are constant in the subdomain Ω_j (Example 1) and to local rigid body motions in the case of linear elasticity (Example 2). Therefore, to ensure solvability, the right-hand sides of the systems must be orthogonal to the kernel (this is in general not the case!), and if they are orthogonal to the kernel, then the solution is not unique. Another serious drawback of Algorithm 7 consists in the absence of some global information exchange mechanism that causes the H^{-2} dependence of the relative condition number $\kappa(\mathbf{C}^{-1}\mathbf{S})$. The balancing technique introduced by Mandel (1993) removes both drawbacks (see also Dryja and Widlund (1995) for a different approach). Let us introduce $N_j \times M_j$ matrices $\mathbf{Z}_j$ consisting of M_j linear independent column vectors $\mathbf{z}_j^m \in \mathbb{R}^{N_j}$ ($m = 1, \ldots, M_j$) such that

$$\ker\mathbf{S}_j \subset \mathrm{range}\mathbf{Z}_j = \mathrm{span}[\mathbf{z}_j^1, \ldots, \mathbf{z}_j^{M_j}] \tag{127}$$

For our model problem (Example 1), we can simply choose $\ker\mathbf{S}_j = \mathrm{range}\mathbf{Z}_j = \mathrm{span}[(1, \ldots, 1)^{\mathrm{T}}]$ for the singular case and omit the balancing procedures (128) and (130) in Algorithm 8 for those i's which belong to the regular local Schur complements $\mathbf{S}_i$. Then it is clear that some vector $\mathbf{r}$ fulfills the local orthogonality conditions ensuring the solvability of the local Neumann problems if $\mathbf{Z}_j^{\mathrm{T}}\mathbf{D}_j\mathbf{A}_j\mathbf{d} = 0$ for the

Algorithm 7. Neumann–Neumann preconditioning operation $\mathbf{w} = \mathbf{C}^{-1}\mathbf{d}$.

$\mathbf{d} \in \mathbb{R}^{N_C}$ given vector (defect/residual) {initialization}
for all $j = 1, \ldots, J$ **do**
  $\mathbf{d}_j = \mathbf{D}_j\mathbf{A}_j\mathbf{d}$ {distribute $\mathbf{d}$ to the subdomains}
end for
for all $j \in \{1, \ldots, J\}$ **in parallel do**
  $\mathbf{S}_j\mathbf{w}_j = \mathbf{d}_j$ {solve the local Neumann problems (124) in parallel}
end for
$\mathbf{w} = \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}}\mathbf{D}_j^{\mathrm{T}}\mathbf{w}_j$ {average the local solution on the interface}

j's corresponding to the singular cases. Adding this balancing step in a symmetric way to Algorithm 7, we arrive at the so-called balancing Neumann–Neumann preconditioning Algorithm 8.

The following remarks may be useful for the practical implementation of Algorithm 8:

1. On the one hand, the balancing steps (128) and (130) can be omitted for such subdomain Ω_i where the subdomain problems are regular, that is, $M_j = 0$, no λ_j and μ_j. On the other hand, the spaces range$\mathbf{Z}_j$ can always be enriched.
2. The sparse SPD matrix arising from the auxiliary problems (128) and (130) has the dimension $M \times M$ with $M = \sum_{j=1}^{J} M_j$. This matrix can easily be generated in a preprocessing step.
3. Owing to the postbalancing step (130), any solution of the local Neumann problems (129) will be appropriate.
4. If the old residual $\mathbf{d}$ is already balanced, then the prebalancing step (128) can be omitted, that is, if the initial residual of the Schur-complement CG iteration is balanced, then the prebalancing step (128) can always be omitted.

From Algorithm 8, we observe that the balanced Neumann–Neumann preconditioner can be rewritten in the compact form $\mathbf{C}^{-1} = (\mathbf{I} - \mathbf{E})\mathbf{S}^{-1}$ with the iteration matrix

$$\begin{aligned}\mathbf{E} &= (\mathbf{I} - \mathbf{P})(\mathbf{I} - \mathbf{TS})(\mathbf{I} - \mathbf{P}) \\ &= (\mathbf{I} - \mathbf{P})\left[\mathbf{I} - \left(\sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} \mathbf{S}_j^{+} \mathbf{D}_j \mathbf{A}_j\right)\mathbf{S}\right](\mathbf{I} - \mathbf{P}) \quad (131)\end{aligned}$$

where $\mathbf{S}_j^{+}$ denotes the Moore–Penrose pseudo-inverse of $\mathbf{S}_j$, and $\mathbf{P}$ is the $\mathbf{S}$-orthogonal projection onto the space $\mathbb{P} = \left\{\mathbf{v} = \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} \mathbf{z}_j \in \mathbb{R}^N : \mathbf{z}_j \in \text{range}\mathbf{Z}_j\right\}$ that plays the role of some 'coarse grid space'. Thus, the balancing

---

Algorithm 8. Balancing Neumann–Neumann preconditioning operation $\mathbf{w} = \mathbf{C}^{-1}\mathbf{d}$.

$\mathbf{d} \in \mathbb{R}^{N_C}$ given vector (defect/residual) {initialization}
Find $\lambda_l \in \mathbb{R}^{M_l}$, $l = 1, \ldots, J$, such that {balancing the old residual vector}

$$\mathbf{Z}_i^{\mathrm{T}} \mathbf{D}_i \mathbf{A}_i \left(\mathbf{d} - \mathbf{S} \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} \mathbf{Z}_j \lambda_j\right) = 0 \quad \forall i = 1, \ldots, J \tag{128}$$

and set $\mathbf{r} = \mathbf{d} - \mathbf{S} \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} \mathbf{Z}_j \lambda_j$
for all $j = 1, \ldots, J$ **do**
  $\mathbf{r}_j = \mathbf{D}_j \mathbf{A}_j \mathbf{r}$ {distribute $\mathbf{r}$ to the subdomains}
end for
{solve the balanced local Neumann problems (124) in parallel}
for all $j \in \{1, \ldots, J\}$ **in parallel do**

$$\mathbf{S}_j \mathbf{w}_j = \mathbf{r}_j \tag{129}$$

end for
Find $\mu_l \in \mathbb{R}^{M_l}$, $l = 1, \ldots, J$, such that {balancing the new residual vector}

$$\mathbf{Z}_i^{\mathrm{T}} \mathbf{D}_i \mathbf{A}_i \left(\mathbf{d} - \mathbf{S} \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} (\mathbf{w}_j + \mathbf{Z}_j \mu_j)\right) = 0 \quad \forall i = 1, \ldots, J \tag{130}$$

$$\mathbf{w} = \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} (\mathbf{w}_j + \mathbf{Z}_j \mu_j)$$ {average the local solution on the interface}

---

steps are handling not only the singular local Neumann problems but also the global information transport via the space $\mathbb{P}$. This is the reason why some enrichment of space can be meaningful. From (131), we immediately see that the balanced Neumann–Neumann preconditioner is some kind of hybrid Schwarz preconditioner, as described in Section 3.2.3.

Theorem 9. *The balanced Neumann–Neumann preconditioner* $\mathbf{C} = \mathbf{S}(\mathbf{I} - \mathbf{E})^{-1}$ *defined by Algorithm 8 is SPD and*

$$\kappa(\mathbf{C}^{-1}\mathbf{S}) \leq \sup_{\mathbf{u}_j \in \mathbb{X}_j,\ j=1,\ldots,J} \frac{\displaystyle\sum_{i=1}^{J} \|\mathbf{A}_i \sum_{j=1}^{J} \mathbf{A}_j^{\mathrm{T}} \mathbf{D}_j^{\mathrm{T}} \mathbf{u}_j\|_{\mathbf{S}_i}^2}{\displaystyle\sum_{j=1}^{J} \|\mathbf{u}_j\|_{\mathbf{S}_j}^2} \leq c \left(1 + \log \frac{H}{h}\right)^2 \tag{132}$$

where $\mathbb{X}_j = \{\mathbf{u}_j \in \mathbb{R}^{N_j} : (\mathbf{u}_j, \mathbf{v}_j) = 0\ \forall \mathbf{v}_j \in \ker \mathbf{S}_j$ *and* $(\mathbf{S}_j \mathbf{u}_j, \mathbf{v}_j) = 0\ \forall \mathbf{v}_j \in \mathrm{range}\mathbf{Z}_j\}$, *and c denotes a positive constant that is independent of h, H, and the jumps in the coefficients.*

Proof. The first part of estimate (132) was proved by Mandel (1993). This estimate is based on pure linear algebra arguments and is not directly connected to our model problem. This abstract estimate was used by Mandel and Brezina (1996) to produce the bound at the right-hand side of estimate (132) for our model problem. □

The advantages of the balanced Neumann–Neumann Schur-complement preconditioners are certainly their (almost) independence of bad parameters (see also Mandel and Brezina (1996) for impressive numerical results) and the fact that more or less standard software routines can be used. On the other hand, the balanced Neumann–Neumann preconditioned Schur-complement CG that is mostly used in practice is quite expensive with respect to the number of arithmetical operations because one Dirichlet and one Neumann problem must be solved exactly (directly) per subdomain (however, completely in parallel) and per iteration step. Inexact versions are not straightforward (see, however, Sections 3 and 5.3).

5.5 Finite element tearing and interconnecting methods

The FETI methods were introduced by Farhat and Roux (1991) (see also Farhat and Roux (1994) for a more detailed description by the same authors) as a nonoverlapping DD parallel solution method for our system (13) of conform finite element equations that can be reduced to the Schur-complement system (95) after eliminating the internal unknowns, as described in Section 5.1. Tearing the unknowns $\mathbf{u}_C$ at the interface Γ_C first into independent unknowns building the vectors $\mathbf{u}_1, \ldots, \mathbf{u}_J$ (from now on we again omit the index C, as in the preceding section) and then again enforcing the continuity by simple interconnecting constraints $\mathbf{Bu} = \mathbf{0}$, we arrive at the saddle-point problem: Given $\mathbf{f} = (\mathbf{f}_j)_{j=1,\ldots,J} \in \mathbb{R}^N$, find $\mathbf{u} = (\mathbf{u}_j)_{j=1,\ldots,J} \in \mathbb{U} = \mathbb{U}_1 \times \cdots \times \mathbb{U}_J = \mathbb{R}^N$ and $\lambda \in \mathbf{\Lambda} = \mathrm{range}\mathbf{B} \subset \mathbb{R}^M$ such that

$$\begin{pmatrix} \mathbf{S} & \mathbf{B}^{\mathrm{T}} \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{0} \end{pmatrix} \tag{133}$$

where $\mathbf{S} = \mathrm{diag}(\mathbf{S}_j)_{j=1,\ldots,J}$ denotes the $N \times N$ block diagonal matrix with the subdomain Schur complements on the diagonal, $\mathbf{B} = (\mathbf{B}_1, \ldots, \mathbf{B}_J)$ is the $M \times N$ matrix of constraints that measures the jump of a given vector $\mathbf{u}$ across the interface, $N = N_1 + \cdots + N_J$, and M is the number of Lagrange multipliers. Each row of the matrix $\mathbf{B}$ is connected with a pair of matching nodes across the interface. The entries of such a row are 1, −1, and 0 for the indices corresponding to the matching nodes and otherwise. Therefore, $\mathbf{Bu} = \mathbf{0}$ implies that the finite element function u corresponding to $\mathbf{u}$ is continuous across the interface Γ_C. We assume here that the number of constraints at some matching node is equal to the number of matching subdomain minus 1. This method of the minimal number of constraints resp. multipliers is called nonredundant (see e.g. Klawonn and Widlund (2001) for the use of redundant constraints).

Since $\ker\mathbf{S} \cap \ker\mathbf{B} = \{\mathbf{0}\}$, the saddle-point system (133) has a unique solution and is completely equivalent to the Schur-complement problem (95). The subdomain Schur complement $\mathbf{S}_j$ is singular if the corresponding subdomain Ω_j does not touch the Dirichlet boundary Γ_D, that is $\Gamma = \partial\Omega$ for our model problem. Such subdomains are called floating subdomains. Similar to Section 5.4, we assume that $\ker\mathbf{S}$ can be represented by the range of some $N \times L$ matrix $\mathbf{Z}$, that is, now $\ker\mathbf{S} = \mathrm{range}\mathbf{Z}$, with L being the number of floating subdomains. If we assume for the time being that the solvability condition

$$\mathbf{f} - \mathbf{B}^{\mathrm{T}}\lambda \perp \ker\mathbf{S} = \mathrm{range}\mathbf{Z}, \text{ i.e. } \mathbf{Z}^{\mathrm{T}}(\mathbf{f} - \mathbf{B}^{\mathrm{T}}\lambda) = 0 \tag{134}$$

for first equation in (133) is fulfilled, then the solution $\mathbf{u}$ can be represented in the form

$$\mathbf{u} = \mathbf{S}^{+}(\mathbf{f} - \mathbf{B}^{\mathrm{T}}\lambda) + \mathbf{Z}\alpha \tag{135}$$

with some element $\mathbf{Z}\alpha \in \ker\mathbf{S}$ that has to be determined. Substituting now (135) into the second block equation of

(133), we arrive at the dual problem

$$\mathbf{B}\mathbf{S}^{+}\mathbf{B}^{\mathrm{T}}\lambda = \mathbf{B}\mathbf{S}^{+}\mathbf{f} + \mathbf{G}\alpha \tag{136}$$

for defining λ and α with the abbreviation $\mathbf{G} = \mathbf{B}\mathbf{Z}$. Defining now the orthogonal projection $\mathbf{P} = \mathbf{I} - \mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}$ from the space $\mathbf{\Lambda}$ onto the subspace $\mathbf{\Lambda}_0 = \ker\mathbf{G}^{\mathrm{T}} = (\mathrm{range}\mathbf{G})^{\perp}$ with respect to the scalar product $(\cdot,\cdot) = (\cdot,\cdot)_{\mathbf{\Lambda}} = (\cdot,\cdot)_{\mathbb{R}^M}$, we can split the definition of λ from the definition of α. Indeed, applying $\mathbf{P}$ to (136) gives the equation

$$\mathbf{P}\mathbf{B}\mathbf{S}^{+}\mathbf{B}^{\mathrm{T}}\lambda = \mathbf{P}\mathbf{B}\mathbf{S}^{+}\mathbf{f} \tag{137}$$

since $\mathbf{P}\mathbf{G}\alpha = \mathbf{0}$. Together with the solvability condition (134), we get the final dual problem in the following form: Find $\lambda \in \mathbf{\Lambda}$ such that

$$\mathbf{P}\mathbf{F}\lambda = \mathbf{P}\mathbf{d} \text{ subject to } \mathbf{G}^{\mathrm{T}}\lambda = \mathbf{e} \tag{138}$$

with the abbreviations $\mathbf{F} = \mathbf{B}\mathbf{S}^{+}\mathbf{B}^{\mathrm{T}}$, $\mathbf{d} = \mathbf{B}\mathbf{S}^{+}\mathbf{f}$, and $\mathbf{e} = \mathbf{Z}^{\mathrm{T}}\mathbf{f}$. Once λ is defined, from (136) we obtain

$$\alpha = (\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}(\mathbf{F}\lambda - \mathbf{d}) \tag{139}$$

and, finally, we get $\mathbf{u}$ from (135). The solution $\mathbf{u}_C$ of the Schur-complement problem (95) can easily be extracted from $\mathbf{u}$.

The dual problem (138) is now solved by the preconditioned conjugate gradient (PCG) iteration in the subspace $\mathbf{\Lambda}_0$ that is presented in Algorithm 9 as a projected PCG method.

The matrix-by-vector multiplication $\mathbf{F}\mathbf{s}^n = \mathbf{B}\mathbf{S}^{+}\mathbf{B}^{\mathrm{T}}\mathbf{s}^n$ means the concurrent (direct) solution of J local Neumann problems. The orthoprojection $\mathbf{P}$ ensures the solvability of the Neumann problems and the global information exchange. The application of $\mathbf{P} = \mathbf{I} - \mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{G}^{\mathrm{T}}$ to some vector $\mathbf{w} \in \mathbf{\Lambda}$ involves the direct solution of a small system with the $L \times L$ system matrix $\mathbf{G}^{\mathrm{T}}\mathbf{G}$ that plays the role of some kind of a coarse grid problem. The FETI preconditioner $\mathbf{C}$ should be spectrally equivalent to the FETI operator $\mathbf{F}$ on the subspace $\mathbf{\Lambda}_0 = \ker\mathbf{G}^{\mathrm{T}}$, that is,

$$\underline{\gamma}(\mathbf{C}\lambda, \lambda) \le (\mathbf{F}\lambda, \lambda) \le \overline{\gamma}(\mathbf{C}\lambda, \lambda) \quad \forall\lambda \in \mathbf{\Lambda}_0 \tag{140}$$

with positive spectral equivalence constants $\underline{\gamma}$ and $\overline{\gamma}$ such that $\kappa(\mathbf{P}\mathbf{C}^{-1}\mathbf{P}^{\mathrm{T}}\mathbf{P}^{\mathrm{T}}\mathbf{F}\mathbf{P}) \le \underline{\gamma}/\overline{\gamma}$ is as small as possible and the preconditioning operation $\mathbf{C}^{-1}\mathbf{d}$ is as cheap as possible. Farhat and Roux (1991) proposed the FETI preconditioner

$$\mathbf{C}^{-1} = \mathbf{B}\mathbf{S}\mathbf{B}^{\mathrm{T}} \tag{141}$$

Algorithm 9. FETI subspace PCG iteration.

{initialization}
$\lambda^0 = \mathbf{G}(\mathbf{G}^{\mathrm{T}}\mathbf{G})^{-1}\mathbf{e}$ {forcing the constraints $\mathbf{G}^{\mathrm{T}}\lambda^0 = \mathbf{e}$ for the initial guess}
$\mathbf{d}^0 = \mathbf{P}(\mathbf{d} - \mathbf{F}\lambda^0)$ {compute the defect and project to the subspace $\mathbf{\Lambda}_0$}
$\mathbf{w}^0 = \mathbf{C}^{-1}\mathbf{d}^0$ {precondition step}
$\mathbf{s}^0 = \mathbf{z}^0 = \mathbf{P}\mathbf{w}^0$ {project the correction to the subspace $\mathbf{\Lambda}_0$}
$\beta_0 = (\mathbf{w}^0, \mathbf{d}^0) = (\mathbf{z}^0, \mathbf{d}^0)$

{begin iteration loop}
for $n = 0$ **step** 1 **until** $\beta_n \le \varepsilon^2\beta_0$ **do**
  $\mathbf{x}^n = \mathbf{P}\mathbf{F}\mathbf{s}^n$ {matrix-by-vector multiplication + projection}
  $\alpha_n = (\mathbf{x}^n, \mathbf{s}^n)$
  $\alpha = \beta_n/\alpha_n$
  $\lambda^{n+1} = \lambda^n + \alpha\,\mathbf{s}^n$ {update of the iterate in the subspace $\mathbf{\Lambda}_0$}
  $\mathbf{d}^{n+1} = \mathbf{d}^n - \alpha\,\mathbf{x}^n$ {update of the defect in the subspace $\mathbf{\Lambda}_0$}
  $\mathbf{w}^{n+1} = \mathbf{C}^{-1}\mathbf{d}^{n+1}$ {precondition step}
  $\mathbf{z}^{n+1} = \mathbf{P}\mathbf{w}^{n+1}$ {project the correction to the subspace $\mathbf{\Lambda}_0$}
  $\beta_{n+1} = (\mathbf{w}^{n+1}, \mathbf{d}^{n+1}) = (\mathbf{z}^{n+1}, \mathbf{d}^{n+1})$
  $\beta = \beta_{n+1}/\beta_n$
  $\mathbf{s}^{n+1} = \mathbf{z}^n - \beta\,\mathbf{s}^n$ {update of the search direction in the subspace $\mathbf{\Lambda}_0$}
end for

{end iteration loop}

that is now called the Dirichlet preconditioner because the multiplication of $\mathbf{S}$ with some vector requires the concurrent solution of J Dirichlet problems. Mandel and Tezaur (1996) proved that the relative spectral condition number is rigorously bounded by $c(1+\log(H/h))^3$. This polylogarithmic bound could be improved to $c(1+\log(H/h))^2$ for special domain decompositions that do not contain cross points. Numerical studies with the Dirichlet FETI preconditioner (141) can be found in Farhat and Roux (1991) and Stefanica (2001). Similar to the balanced Neumann–Neumann Schur-complement preconditioning technique, Klawonn and Widlund (2001) introduced the preconditioner

$$\mathbf{C}^{-1} = (\mathbf{B}\mathbf{D}^{-1}\mathbf{B}^{\mathrm{T}})^{-1}\mathbf{B}\mathbf{D}^{-1}\mathbf{S}\mathbf{D}^{-1}\mathbf{B}^{\mathrm{T}}(\mathbf{B}^{\mathrm{T}}\mathbf{D}^{-1}\mathbf{B})^{-1} \qquad (142)$$

with the diagonal scaling matrix $D = \mathrm{diag}(\mathbf{D}_j)_{j=1,\ldots,J}$. This scaling and the introduction of an appropriately scaled scalar product in $\mathbf{\Lambda}_0$ (this affects the orthogonal projection $\mathbf{P}$!) lead to the rigorous bound $c(1+\log(H/h))^2$ where the constant c is now independent of the jumps in the coefficients. The numerical and the parallel performances of the classical Dirichlet preconditioner and this new preconditioner are compared in Stefanica (2001). In its exact version, the FETI Algorithm 9 requires the exact (direct) solution of one local Neumann problem (matrix multiplication step) and one local Dirichlet problem (preconditioning step) per subdomain (i.e. in parallel) at each iteration step. In these parts, local direct solvers can be very expensive in practical applications where very complex local problems can appear. The use of inexact solvers (preconditioners) for the local Dirichlet problems is more or less straightforward, whereas the replacement of exact Neumann solvers by inexact ones is not.

Klawonn and Widlund (2000) avoided the reduction of the original problem (13) to the saddle-point problem (133) by eliminating the internal unknowns and related the original problem (13) directly to the saddle-point problem

$$\begin{pmatrix} \mathbf{K} & \mathbf{B}^{\mathrm{T}} \\ \mathbf{B} & \mathbf{0} \end{pmatrix} \begin{pmatrix} \mathbf{u} \\ \lambda \end{pmatrix} = \begin{pmatrix} \mathbf{f} \\ \mathbf{0} \end{pmatrix} \qquad (143)$$

with the torn original stiffness matrix $\mathbf{K} = \mathrm{diag}(\mathbf{K}_j)_{j=1,\ldots,J}$ and the interconnecting matrix $\mathbf{B}$. We refer the reader to Farhat, Lesoinne and Pierson (2000) for some further development of the FETI methodology. In particular, the dual-primal FETI (FETI-DP) methods, introduced by Farhat *et al.* (2001), seem to be very attractive because they avoid the solution of singular problems on the subdomains by fixing some primal unknowns. The first results on the convergence analysis of FETI-DP methods were given by Mandel and Tezaur (2001) and Klawonn and Widlund (2002). Langer and Steinbach (2003) introduced the boundary element tearing and interconnecting (BETI) methods as boundary element counterparts of the FETI methods.

5.6 Mortar methods

In the classical FETI method, we tore (split) the unknowns on the interface, which are conform in our original FE scheme, and interconnected them again by simple $(1,-1)$ equality constraints with the only aim to construct a DD solver that is essentially based on the dual problem for the corresponding Lagrange multipliers. The Mortar technique proposed by Bernardi and Maday and Patera (1993, 1994) goes one step further and allows the triangulation to be nonconforming. Thus, the FE solution cannot globally conform in this general situation and the continuity must be enforced by constraints in an appropriate way. This continuity constraints can be included into the product FE space or can be incorporated by Lagrange multipliers in a saddle-point formulation.

Let us again consider our model problem of Example 1 with piecewise constant coefficients and the nonoverlapping domain decomposition (92), but now we allow the triangulation and the FE functions to be nonconforming across the interfaces $\partial\Omega_j \cap \partial\Omega_i$. Thus, we look for a nonconforming FE solution u in the product FE space $\mathbb{U} = \mathbb{V}_1 \times \ldots \times \mathbb{V}_J$, where the subdomain FE spaces $\mathbb{V}_j = \mathbb{V}_h(\Omega_j) \subset \mathbb{H}^1(\Omega_j) \cap \mathbb{H}^1_0(\Omega)$ are defined on the individual subdomain triangulations $\mathcal{T}_j$ using their individual finite elements. In order to get a proper approximation to our weak solution, we have to enforce weak continuity constraints by Lagrange multipliers in such a way that the approximation and consistency errors are not perturbed. To do this, we first introduce two different, but complementary nonoverlapping decompositions $\overline{\Gamma}_C = \bigcup_{j=1}^{J} \bigcup_{i\in\mathcal{M}(j)} \overline{\Gamma}_{ji}$ and $\overline{\Gamma}_C = \bigcup_{j=1}^{J} \bigcup_{i\in\mathcal{M}(j)} \overline{\Gamma}_{ij}$ of the interface Γ_C into mortar and nonmortar faces $\Gamma_{ji} = \partial\Omega_j \cap \partial\Omega_i \subset \partial\Omega_j$ (edges in 2D). The face Γ_{ji} is considered as a part of $\partial\Omega_j$ and inherits the (surface) triangulation from Ω_j. If some face Γ_{ji} is mortar, that is, $i \in \mathcal{M}(j)$, then its opposite side $\Gamma_{ij} \subset \partial\Omega_i$ is nonmortar, that is, $j \notin \mathcal{M}(i)$. Let us now introduce the discrete Lagrange multiplier space

$$\mathbf{\Lambda} = \prod_{j=1}^{J} \prod_{i\in\mathcal{M}(j)} \mathbf{\Lambda}(\Gamma_{ij}) \subset \prod_{j=1}^{J} \prod_{i\in\mathcal{M}(j)} (\mathbb{H}^{1/2}(\Gamma_{ij}))^{\star} \qquad (144)$$

where the local discrete Lagrange multiplier spaces are all connected with the nonmortar faces. The choice of local discrete Lagrange multiplier spaces $\mathbf{\Lambda}(\Gamma_{ij})$ is crucial not only for the approximation properties but also for efficiency reasons. For instance, in the case of linear triangular elements in 2D, the classical local discrete Lagrange multiplier space

$\mathbf{\Lambda}(\Gamma_{ij})$ is a subspace of codimension two of the trace space $\mathbb{V}_i|_{\Gamma_{ij}}$ on the nonmortar edge Γ_{ij}. More precisely, $\mathbf{\Lambda}(\Gamma_{ij})$ consists of all continuous, piecewise linear functions on Γ_{ij}, that are constant in the first and the last interval of the 1D mesh on Γ_{ij} induced by the mesh of $\overline{\Omega}_i$. We refer to Ben Belgacem and Maday (1999) for the 3D case and to Wohlmuth (2000) for biorthogonal mortar elements.

Now, the mortar scheme can be formulated in the constrained product space $\mathbb{V} = \{v \in \mathbb{U} : b(v, \mu) = 0 \; \forall \mu \in \mathbf{\Lambda}\}$ as nonconforming DD FE scheme: Find $u \in \mathbb{V}$ such that

$$a(u, v) = \langle f, v \rangle \quad \forall v \in \mathbb{V} \tag{145}$$

Alternatively, the mortar scheme can be reformulated as a mixed scheme in the unconstrained product space and the Lagrange multiplier space. Find $u \in \mathbb{U}$ and $\lambda \in \mathbf{\Lambda}$ such that

$$a(u, v) + b(v, \lambda) = \langle f, v \rangle \quad \forall v \in \mathbb{U} \tag{146}$$

$$b(u, \mu) = 0 \quad \forall \mu \in \mathbf{\Lambda} \tag{147}$$

where $a(u, v) = \sum_{j=1}^{L} \alpha_j \int_{\Omega_j} \nabla u(x) \cdot \nabla v(x) \, \mathrm{d}x$, $b(v, \mu) = \sum_{j=1}^{L} \sum_{i \in \mathcal{M}(j)} \int_{\Gamma_{ij}} [v] \, \mu \, \mathrm{d}s$, and $[v] = v|_{\Gamma_{ij}} - v|_{\Gamma_{ji}}$ denotes the jump across face Γ_{ij} that geometrically coincides with Γ_{ji}.

The saddle-point problem can be rewritten in matrix form as the full FETI saddle-point problem (143), or, after eliminating the inner subdomain unknowns, as the reduced FETI saddle-point problem (133). However, the Lagrange multiplier matrix $\mathbf{B}$ is now defined by the mortar conditions (147) across the faces Γ_{ij} instead of the simple hard nodal continuity condition in the FETI method. Now, the FETI solver can be used for solving the mortar saddle-point problem in the same way as in Section 5.6 for solving the original FETI equations (133) (see Stefanica (2001) for more information and numerical experiments). We mention that other nonoverlapping DD algorithms and multilevel methods can successfully be applied to the solution of the mortar equation as well (see Wohlmuth (2001) for more information).

ACKNOWLEDGMENTS

The authors would like to thank the *Mathematisches Forschungsinstitut Oberwolfach* for the hospitality during their RiP stay in summer 2002, the *Austrian Science Fund (FWF)* for supporting the research work of the authors under the grant SFB F013, the Russian Fund of Basic Research, grant 00-01-00772, and the program "Universities of Russia", grant 03.01.017.

REFERENCES

Ainsworth M. A preconditioner based on domain decomposition for h-p finite element approximation on quasi-uniform meshes. *SIAM J. Numer. Anal.* 1996; **33**:1358–1376.

Ainsworth M and Senior B. Aspects of an adaptive hp-finite element method: adaptive strategy, conforming approximation and efficient solvers. *Comput. Methods Appl. Mech. Eng.* 1997; **150**:65–87.

Aronszjan N. Theory of reproducing kernels. *Trans. Am. Math. Soc.* 1950; **68**:337–404.

Axelsson O and Vassilevskii PS. Algebraic multilevel preconditioning methods I. *SIAM J. Numer. Anal.* 1989; **56**:157–177.

Babuska I. On the Schwarz algorithm in the theory of differential equations of mathematical physics. *Czechosl. Math. J.* 1958; **8**(83):328–342 (in Russian).

Babuska I, Craig A, Mandel J and Pitkäranta J. Efficient preconditioning for the p-version finite element method in two dimensions. *SIAM J. Numer. Anal.* 1991; **28**(3):624–661.

Bank RE, Jimack PK, Nadeem SA and Nepomnyaschikh SV. A weakly overlapping domain decomposition preconditioner for the finite element solution of elliptic partial differential equations. *SIAM J. Sci. Comput.* 2002; **23**(6):1818–1242.

Ben Belgacem F and Maday Y. The mortar element method for three dimensional finite elements. *Modell. Math. Anal. Numer.* 1999; **36**(4):1234–1263.

Bernardi Ch, Dauge M and Maday Y. Relévements de traces presérvant les polynômes. *C. R. Acad. Sci. Paris Sér. I Math.* 1992; **315**:333–338 (in French).

Bernardi C, Maday Y and Patera A. Domain decomposition by the mortar element method. In *Asymptotic and Numerical Methods for Partial Differential Equations with Critical Parameters*, Kaper H and Garbey M (eds). Reidel: Dordrecht, 1993; 269–286.

Bernardi C, Maday Y and Patera A. A new nonconforming approach to domain decomposition: the mortar element method. In *Nonlinear Partial Differential Equations and their Applications*, Brezis H and Lions JL (eds). Pitman: New York, 1994; 13–51.

Bernardi Ch, Maday Y and Sacchi-Landriani G. Non conforming matching conditions for coupling spectral and finite element method. *Appl. Numer. Math.* 1989; **54**:64–84.

Beuchler S. Multigrid solver for the inner problem in domain decomposition methods for p-FEM. *SIAM J. Numer. Anal.* 2002; **40**(4):928–944.

Bjørstad PE and Mandel J. On the spectra of sums of orthogonal projections with applications to parallel computing. *BIT* 1991; **31**:76–88.

Bjørstad PE and Widlund OB. Iterative methods for the solution of elliptic problems on regions partitioned into substructures. *SIAM J. Numer. Anal.* 1986; **23**(6):1093–1120.

Bourgat JP, Glowinski R, Le Tallec P and Vidrascu M. Variational formulation and algorithm for trace operator in domain decomposition calculations. In *Second International Symposium on Domain Decomposition Methods for Partial Differential*

Equations, Chan TF, Glowinski R, Périaux J and Widlund OB (eds). SIAM: Philadelphia, 1989; 3–16.

Bramble JH and Zhang X. The analysis of multigrid methods. *Handbook of Numerical Analysis VII*. North Holland: Amsterdam, 2000; 173–415.

Bramble JH, Pasciak JE and Schatz AH. The construction of preconditioners for elliptic problems by substructuring, I. *Math. Comput.* 1986; **47**(175):103–134.

Bramble JH, Pasciak JE and Schatz AH. The construction of preconditioners for elliptic problems by substructuring, II. *Math. Comput.* 1987; **49**(179):1–16.

Bramble JH, Pasciak JE and Schatz AH. The construction of preconditioners for elliptic problems by substructuring, III. *Math. Comput.* 1988; **51**(184):415–430.

Bramble JH, Pasciak JE and Schatz AH. The construction of preconditioners for elliptic problems by substructuring, IV. *Math. Comput.* 1989; **53**(187):1–24.

Bramble JH, Pasciak JE and Xu J. Parallel multilevel preconditioners. *Math. Comput.* 1990; **55**:1–22.

Bramble JH, Pasciak JE, Wang J and Xu J. Convergence estimates for product iterative methods with applications to domain decomposition. *Math. Comput.* 1991; **57**(195):1–21.

Brenner SC. The condition number of the Schur complement in domain decomposition. *Numer. Math.* 1999; **83**:187–203.

Brenner SC. Lower bounds for two-level additive Schwarz preconditioners with small overlap. *SIAM J. Sci. Comput.* 2000; **21**(5):1657–1669.

Casarin MA. Quasi-optimal Schwarz methods for the conforming spectral element discretization. *SIAM J. Numer. Anal.* 1997; **34**(6):2482–2502.

Carstensen C, Kuhn M and Langer U. Fast parallel solvers for symmetric boundary element domain decomposition equations. *Numer. Math.* 1998; **79**(2):321–347.

Chan TF. Iterative methods for the solution of elliptic problems on regions partitioned into substructures. *SIAM J. Numer. Anal.* 1987; **24**(2):382–390.

Chan TF and Mathew TP. Domain decomposition algorithms. *Acta Numer.* 1994; 61–143.

Ciarlet P. *The Finite Element Method for Elliptic Problems*. North Holland Publishing Company: Amsterdam, New York, Oxford, 1978.

Demmel JW, Eisenstat TSC, Gilbert JR, Li XS and Liu JWH. A supernodal approach to sparse partial pivoting. *SIAM J. Matrix Anal. Appl.* 1999; **20**:720–755.

De Roeck YH and Le Tallec P. Analysis and test of local domain decomposition preconditioners. In *Fourth International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Glowinski R, Kuznetsov YA, Meurant G, Périaux J and Widlund OB (eds). SIAM: Philadelphia, 1991; 112–128.

Dryja M. A capacitance matrix method for Dirichlet problems on polygonal regions. *Numer. Math.* 1982; **39**(1):51–64.

Dryja M. A finite element-capacitance matrix method for elliptic problems on regions partitioned into substructures. *Numer. Math.* 1984; **44**(1):153–168.

Dryja M and Widlund OB. Some domain decomposition algorithm for elliptic problems. In *Iterative Methods for Large Linear Systems*, Hayes L and Kincaid D (eds). Academic Press: Orlando, 1989; 273–291.

Dryja M and Widlund OB. Domain decomposition algorithms with small overlap. *SIAM J. Sci. Comput.* 1994; **15**(3):604–620.

Dryja M and Widlund OB. Schwarz methods of Neumann–Neumann type for three-dimensional elliptic finite element problems. *Commun. Pure Appl. Math.* 1995; **48**(2):121–155.

Dryja M, Smith BF and Widlund OB. Schwarz analysis of iterative substructuring algorithms for elliptic problems in three dimensions. *SIAM J. Numer. Anal.* 1994; **31**(6):1662–1694.

Farhat C and Roux FX. A method of finite element tearing and interconnecting and its parallel solution algorithm. *Int. J. Numer. Methods Eng.* 1991; **32**:1205–1227.

Farhat C and Roux FX. Implicite parallel processing in structural mechanics. In *Computational Mechanics Advances*, Oden JT (ed.). North Holland: Amsterdam, 1994; 1–124.

Farhat C, Lesoinne M and Pierson K. A scalable dual-primal domain decomposition method. *Numer. Lin. Alg. Appl.* 2000; **7**:687–714.

Farhat C, Lesoinne M, Le Tellec P, Pierson K and Rixen D. FETI-DP: A dual-primal unified FETI method – part I: A faster alternative to the two-level FETI method. *Int. J. Numer. Methods Eng.* 2001; **50**:1523–1544.

George A and Liu JWH. *Computer Solution of Large Sparse Positive Definite Systems*. Pretice Hall Inc: Englewood Cliffs, 1981.

Golub G and Mayers D. The use of preconditioning over irregular regions. In *Computing Methods in Applied Sciences and Engineering*, Glowinski R and Lions JL(eds). North Holland: Amsterdam, New York, Oxford, 1994; 3–14.

Griebel M and Oswald P. On the abstract theory of additive and multiplicative Schwarz algorithms. *Numer. Math.* 1995; **70**(2):163–180.

Gupta A. Recent advances in direct methods for solving unsymmetric sparse systems of linear equations. *ACM Trans. Math. Softw.* 2002; **28**:301–324.

Haase G and Langer U. The non-overlapping domain decomposition multiplicative Schwarz method. *Int. J. Comput. Math.* 1992; **44**:223–242.

Haase G and Nepomnyaschikh SV. Extension explicite operators on hierarchical grids. *East-West J. Numer. Math.* 1997; **5**(4):231–248.

Haase G, Langer U and Meyer A. The approximate Dirichlet domain decomposition method. Part I and II. *Computing* 1991; **47**:137–167.

Haase G, Langer U, Meyer A and Nepomnyaschikh SV. Hierarchical extension operators and local multigrid methods in domain decomposition preconditioners. *East-West J. Numer. Math.* 1994; **2**(3):173–193.

Haase G, Heise B, Kuhn M and Langer U. Adaptive domain decomposition methods for finite and boundary element equations. In *Boundary Element Topics*, Wendland WL (ed.). Springer-Verlag: Berlin, Heidelberg, New York, 1997; 121–147.

Ivanov SA and Korneev VG. Selection of the high order coordinate functions and preconditioning in the frame of the domain decomposition method. *Izv. Vyshh. Uchebn. Zaved.* 1995; **395**(4):62–81 (in Russian).

Ivanov SA and Korneev VG. Preconditioning in the domain decomposition methods for the *p*-version with the hierarchical bases. *Math. Model.* 1996; **8**(9):63–73.

Jung M and Langer U. Application of multilevel methods to practical problems. *Surveys Math. Ind.* 1991; **1**:217–257.

Keyes DE and Gropp WD. A comparison of domain decomposition techniques for elliptic partial differential equations and their parallel implementation. *SIAM J. Sci. Comput.* 1987; **8**(2):166–202.

Klawonn A and Widlund OB. A domain decomposition method with Lagrange multipliers and inexact solvers for linear elasticity. *SIAM J. Sci. Comput.* 2000; **22**(4):1199–1219.

Klawonn A and Widlund OB. FETI and Neumann-Neumann iterative substructuring methods: connections and new results. *Commun. Pure Appl. Math.* 2001; **54**:57–90.

Klawonn A and Widlund OB. Dual-primal FETI methods for three-dimensional elliptic problems with heterogeneous coefficients. *SIAM J. Numer. Anal.* 2002; **40**(1):159–179.

Korneev VG. Almost optimal solver for Dirichlet problems on subdomains of decomposition. *Differ. Equations* 2001; **37**(7):958–968 (in Russian).

Korneev VG. Local Dirichlet problems on subdomains of decomposition in *hp* discretizations, and optimal algorithms for their solution. *Math. Model.* 2002; **14**(5):51–74.

Korneev VG and Jensen S. Preconditioning of the p-version of the finite element method. *Comput. Methods Appl. Mech. Eng.* 1997; **150**(1–4):215–238.

Korneev VG and Jensen S. Domain decomposition preconditioning in the hierarchical p-version of the finite element method. *Appl. Numer. Math.* 1999; **29**:479–518.

Korneev VG and Langer U. *Approximate Solution of Plastic Flow Theory Problems*, Teubner Texte zur Mathematik. B.G. Teubner: Leipzig, 1984.

Korneev V, Langer U and Xanthis L. On fast domain decomposition solving procedures for *hp*-discretizations of 3-*d* elliptic problems. *Comput. Methods Appl. Math.* 2003; **3**(4):536–559.

Korneev VG, Flaherty J, Oden T and Fish J. Additive Schwarz algorithms for solving hp-version finite element systems on triangular meshes. *Appl. Numer. Math.* 2002; **43**(3):399–421.

Langer U and Steinbach O. Boundary element tearing and interconnecting methods. *Computing* 2003; **71**(3):205–228.

Lebedev VI and Agoshkov VI. *Poincaré-Steklov Operators and their Applications in Analysis*. Academy of Sciences USSR: Moskow, 1983 (in Russian).

Le Tallec P. Domain decomposition methods in computational mechanics. In *Computational Mechanics Advances*, Oden JT (ed.). North Holland: New York, 1994; 121–220.

Lions PL. On the Schwarz alternating method I. In *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Glowinski R, Golub GH, Meurant GA and Périaux J (eds). SIAM: Philadelphia, 1988; 1–42.

Mandel J. Balancing domain decomposition. *Commun. Numer. Methods Eng.* 1993; **9**:233–241.

Mandel J and Brezina M. Balancing domain decomposition for problems with large jumps in coefficients. *Math. Comput.* 1996; **65**:1387–1401.

Mandel J and Tezaur R. Convergence of a substructuring method with Lagrange multipliers. *Numer. Math.* 1996; **73**:473–487.

Mandel J and Tezaur R. On the convergence of a dual-primal substructuring method. *Numer. Math.* 2001; **88**:543–558.

Matsokin AM and Nepomnyaschikh SV. A Schwarz alternating method in a subspace. *Sov. Math.* 1985; **29**(10):78–84.

Mikhlin SG. On the Schwarz algorithm. *Dokl. Akad. Nauk S.S.S.R.* 1951; **77**(4):569–571 (in Russian).

Morgenstern D. Begründung des alternierenden Verfahrens durch Orthogonalprojektion. *ZAMM* 1956; **36**(7/8):255–256 (in German).

Nepomnyaschikh SV. Fictitious components and subdomain alternating methods. *Sov. J. Numer. Anal. Math. Modell.* 1990; **5**(1):53–68.

Nepomnyaschikh SV. Mesh theorems on traces, normalizations of function traces and their inversion. *Sov. J. Numer. Anal. Math. Modell.* 1991; **6**(3):223–242.

Nepomnyaschikh SV. Method of splitting into subspaces for solving elliptic boundary value problems in complex-form domains. *Sov. J. Numer. Anal. Math. Modell.* 1991; **6**(2):151–168.

Nevanlinna R. Über das alternierenden Verfahren von Schwarz. *Crelle's J. Reine Angew. Math.* 1939; **180**:121–128 (in German).

Oden JT, Patra A and Feng YS. Parallel domain decomposition solver for adaptive hp finite element methods. *SIAM J. Numer. Anal.* 1997; **34**:2090–2118.

Orzag A. Spectral methods for problems in complex geometries. *J. Comput. Phys.* 1980; **37**:70–92.

Oswald P. On discrete norm estimates related to multilevel preconditioners in the finite element method. In *Proceedings of the International Conference on Constructive Theory of Functions*, Ivanov KG, Petrusher P and Sendov B (eds). Varna 1991, Bulg. Acad. Sci.:Sofia, 1992; 203–241.

Oswald P. *Multilevel Finite Element Approximation: Theory and Application*. Teubner Skripten zur Numerik, B.G. Teubner: Stuttgart, 1994.

Pavarino LF. Additive Schwarz methods for the *p*-version finite element method. *Numer. Math.* 1994; **66**(4):493–515.

Pavarino LF and Widlund OB. Polylogarithmic bound for an iterative substructuring method for spectral elements in three dimensions. *SIAM J. Numer. Anal.* 1996; **37**(4):1303–1335.

Przemieniecki JS. Matrix structural analysis of substructures. *AIAA J.* 1963; **1**(1):138–147.

Quarteroni A and Vali A. *Domain Decomposition Methods for Partial Differential Equations*. Oxford Sciences Publications, 1999.

Schwarz HA. Über einige Abbildungseigenschaften. *J. Reine Angew. Numer. Math.* 1869; **70**:105–120.

Smith BF and Widlund OB. A domain decomposition algorithm using a hierarchical basis. *SIAM J. Sci. Statist. Comput.* 1990; **11**(4):1212–1220.

Smith B, Bjørstad P and Gropp W. *Domain Decomposition: Parallel Multilevel Methods for Elliptic Partial Differential Equations*. Cambridge University Press, 1996.

Sobolev SL. Schwarz algorithm in the theory of elasticity. *Dokl. Akad. Nauk S.S.S.R.* 1936; **4**:235–238 (in Russian).

Stefanica D. A numerical study of FETI algorithms for mortar finite element methods. *SIAM J. Sci. Statist. Comput.* 2001; **23**(4):1135–1160.

Steinbach O. *Stability Estimates for Hybrid Coupled Domain Decomposition Methods*. Springer-Verlag: Berlin, Heidelberg, New York, 2003.

Tong CH, Chan TF and Kuo CC. A decomposition preconditioner based on a change to a multilevel nodal basis. *SIAM J. Sci. Statist. Comput.* 1991; **12**(6):1486–1495.

Widlund OB. Iterative substructuring methods: algorithms and theory for elliptic problems in the plane. In *First International Symposium on Domain Decomposition Methods for Partial Differential Equations*, Glowinski R, Golub GH, Meurant GA and Périaux J (eds). SIAM: Orlando, Philadelphia, 1988; 113–128.

Widlund OB. Preconditioners for spectral and mortar finite element methods. In *Proceedings of the Eighth International Conference on Domain Decomposition Methods*, Beijing, PRC, 15–19 May, 1995; Wiley-Interscience: Strasbourg, 1996.

Wohlmuth B. A mortar finite element method using dual spaces for Lagrange multipliers. *SIAM J. Numer. Anal.* 2000; **38**:989–1014.

Wohlmuth B. *Discretization Methods and Iterative Solvers Based on Domain Decomposition*. Springer: Heidelberg, 2001.

Xu J. Iterative methods by space decomposition and subspace correction: a unifying approach. *SIAM Rev.* 1992; **34**:581–613.

Xu J and Zikatanov L. The method of alternating projections and the method of subspace corrections in Hilbert space. *J. Am. Math. Soc.* 2002; **15**:573–597.

Xu J and Zou J. Some nonoverlapping domain decomposition methods. *SIAM Rev.* 1998; **40**(4):857–914.

Zhang X. Multilevel Schwarz methods. *Numer. Math.* 1992; **63**(4):521–539.

FURTHER READING

Ben Belgacem F, Seshaiyer P and Suri M. Optimal convergence rates of *hp* mortar finite element methods for second order elliptic problems. *RAIRO Math. Modeling Numer. Anal.* 2000; **34**:591–608.

Chan TF and Resasco DC. *A Survey of Preconditioners for Domain Decomposition*. Technical Report, DCS/RR-414, Yale University 1985.

Domain Decomposition home page. http://www.ddm.org [30 September 2003].