An Introduction to Domain Decomposition Algorithms

Olof Widlund
Courant Institute of Mathematical Sciences
New York University
http://www.cs.nyu.edu/cs/faculty/widlund/

Scope of This Tutorial

I will exclusively look at domain decomposition algorithms for positive definite, symmetric problems arising from a low order finite element approximation of elliptic problems. All subproblems will be solved exactly by Cholesky's algorithm. I will also give few references to the literature.

I will adopt the view that a domain decomposition algorithm provides preconditioners (approximate inverses) M of the large and often very ill-conditioned stiffness matrices A that arise in finite element practice.

They are designed with parallel computing systems in mind and the best of them have proven to scale very well on systems with very many processors.

All this work aims at designing preconditioners M such that $\kappa(M^{-1}A)$, the condition number of the preconditioned operator, is small, while keeping the costs of applying M^{-1} acceptable. A preconditioned Krylov space method is almost always used to accelerate the convergence of the iteration.

The development of theory has greatly assisted in the development of improved algorithms; it can be viewed as a subfield of finite element theory. In particular, some of the good choices of primal constraints and scalings for some of these methods are unlikely to have been found without theoretical work.

Poisson's Equation and a Simple Finite Element Model

By using Green's formula, we can write Poisson's equation as a variational problem: Find $u \in V$ such that $\forall v \in V$,

$$a(u,v) := \int_\Omega \nabla u \cdot \nabla v dx = F(v) := \int_\Omega f v dx + \int_{\partial\Omega_N} g_N v ds.$$

Here f is the load, i.e., the right hand side and g_N the Neumann data given on $\partial\Omega_N \subset \partial\Omega$. All elements of $V \subset H^1(\Omega)$ vanish on the set $\partial\Omega_D := \partial\Omega \setminus \partial\Omega_N$, first assumed to be non empty. This problem is then uniquely solvable.

Introduce a triangulation $\mathcal{T}_h$ of Ω and $V^h \subset V$, the standard piecewise linear finite elements on the triangulation. A linear system $Au = F$ results where u is the vector of nodal values at all interior nodes and those on $\partial\Omega_N$. The stiffness matrix A is sparse, symmetric, and positive definite and can be very large. The resulting finite element solution $u^h(x)$ is well defined and converges to the solution of the differential equation when $h \to 0$.

The smallest eigenvalue $\lambda_1(\Omega)$ of the differential operator, and indirectly that of the stiffness matrix, can be estimated by using Friedrichs' inequality

$$\|u\|^2_{L^2(\Omega)} \leq C_1 a(u,u) + C_2 (\int_{\partial\Omega_D} u ds)^2.$$

For $u \in V$ the second integral vanishes and we get a positive lower bound of the Rayleigh quotient $a(u,u)/\|u\|^2_{L^2(\Omega)}$ and of λ_1.

In a pure Neumann problem, $\partial\Omega_N = \partial\Omega$, the Laplace operator and the stiffness matrix have a common null space of constants and the problem is solvable, modulo a constant, iff $F(1) = 0$. The second eigenvalue $\lambda_2(\Omega)$ of the operator is directly related to Poincaré's inequality:

$$\|u\|^2_{L^2(\Omega)} \leq C_1 a(u,u) + C_2 (\int_\Omega u dx)^2.$$

Note that the second term on the right vanishes if u is orthogonal to the null space.

The largest eigenvalue of the stiffness matrices can be estimated by using Gershgorin's theorem.

It is important to understand what happens to these two inequalities when the diameter of the domain changes under a dilation; a simple change of variables gives the answer. Certain powers of the dilation factor will appear with the constants. Similarly, the full $H^1(\Omega)-$norm should be defined by

$$\|u\|^2_{H^1(\Omega)} := |u|^2_{H^1(\Omega)} + 1/diam(\Omega)^2 \|u\|^2_{L_2(\Omega)} = a(u,u) + 1/diam(\Omega)^2 \|u\|^2_{L_2(\Omega)}.$$

This formula is obtained by using the standard norm for a domain with diameter 1 and dilation.

Using these inequalities and a few additional, elementary arguments, we can show that the condition numbers of the stiffness matrices grow as Ch^{-2} in the case of quasi-uniform meshes. This accounts for the relatively slow convergence of the conjugate gradient method without preconditioning.

We will consider the same type of stiffness matrices for subdomains Ω_i obtained by integrating over the subset $\Omega_i \subset \Omega$. These matrices will be important building blocks for our finite element models and domain decomposition algorithms.

Two Subdomains

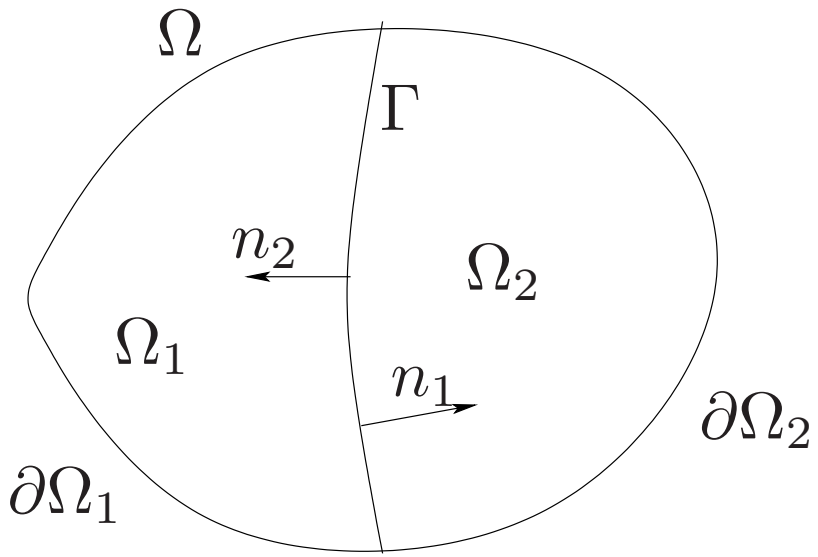


Figure 1: Partition into two non-overlapping subdomains.

Consider a domain Ω subdivided into two non-overlapping subdomains Ω_1 and Ω_2. In between the interface Γ.

Consider a finite element approximation of a Poisson problem on Ω (or scalar elliptic, linear elasticity, or even an incompressible Stokes problem.)

Set up a load vector and a stiffness matrix for each subdomain

$$f^{(i)} = \begin{pmatrix} f_I^{(i)} \\ f_\Gamma^{(i)} \end{pmatrix}, \quad A^{(i)} = \begin{pmatrix} A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\ A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)} \end{pmatrix}, \quad i = 1, 2.$$

Homogeneous Dirichlet condition on $\partial\Omega_i \setminus \Gamma$, Neumann on Γ.

Subassemble:

$$A = \begin{pmatrix} A_{II}^{(1)} & 0 & A_{I\Gamma}^{(1)} \\ 0 & A_{II}^{(2)} & A_{I\Gamma}^{(2)} \\ A_{\Gamma I}^{(1)} & A_{\Gamma I}^{(2)} & A_{\Gamma\Gamma} \end{pmatrix}, \quad u = \begin{pmatrix} u_I^{(1)} \\ u_I^{(2)} \\ u_\Gamma \end{pmatrix}, \quad f = \begin{pmatrix} f_I^{(1)} \\ f_I^{(2)} \\ f_\Gamma \end{pmatrix}.$$

$A_{\Gamma\Gamma} = A_{\Gamma\Gamma}^{(1)} + A_{\Gamma\Gamma}^{(2)}$. Degrees of freedom partitioned into those internal to Ω_1, and internal to Ω_2, and those on Γ.

This is a simple example of how stiffness matrices are assembled from those of the subdomains; we add quadratic forms representing the energy contributed by the subdomains.

Eliminate the interior unknowns. Gives two Schur complements:

$$S^{(i)} := A_{\Gamma\Gamma}^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} A_{I\Gamma}^{(i)}, \quad i = 1, 2.$$

The given system can then be reduced to

$$Su_\Gamma = (S^{(1)} + S^{(2)})u_\Gamma = g_\Gamma.$$

If we use exact solvers for the subdomain problems, we can often reduce our discussion to one about Schur complement problems. We can also take advantage of the reduction in dimension of the Krylov space vectors. Once the interface values are approximated well enough, we can find the values in the interiors by solving a Dirichlet problem for each subdomain. The condition number of a Schur complement of a positive definite symmetric matrix A is always smaller than that of A. In our particular context, the Schur complements will have a condition number on the order of Ch^{-1}. This bound is sharp.

It is easy to see that the product of $S^{(i)}$ times a vector can be obtained at essentially the cost of solving a Dirichlet problem; the elements of the Schur complements need not be computed. This is in contrast to the use of Cholesky's method for the entire problem. It is known that for any symmetric permutation P, factoring $P^T A P$ will require at least quadratic work for a three-dimensional finite element matrix A.

The product of S with a vector, as needed when computing a residual, can then be assembled from matrix-vector products with the two of the subdomain Schur complements.

An important family of domain decomposition methods are the *iterative substructuring* methods; the vocabulary borrowed from structural engineering.

By solving a problem with the matrix $A^{(i)}$ with a right hand side of the form $(0, f_\Gamma)^T$, we obtain a solution with the second component equal to ${S^{(i)}}^{-1} f_\Gamma$; this is an easy exercise on block-Gaussian elimination.

A solution of a system with such a right hand side is $discrete\ harmonic$. It is $A^{(i)}-$orthogonal to any v which vanishes on Γ; it therefore provides the $minimal\ energy\ extension$ for given values on Γ.

Matrix-vector multiplications with $S^{(i)}$ and ${S^{(i)}}^{-1}$ are completely local operations and it does not matter if we have two or many more subdomains; we can use one processor for each subdomain problem and work in parallel.

Coupled system of PDE

Consider Poisson equation on Ω, 2D or 3D. Zero Dirichlet data on $\partial\Omega$, the boundary of Ω. Ω partitioned into two non-overlapping subdomains Ω_i :

$$\overline{\Omega} = \overline{\Omega_1 \cup \Omega_2}, \quad \Omega_1 \cap \Omega_2 = \emptyset, \quad \Gamma = \partial\Omega_1 \cap \partial\Omega_2;$$

$$\text{measure}(\partial\Omega_1 \cap \partial\Omega) \ > 0, \quad \text{measure}(\partial\Omega_2 \cap \partial\Omega) \ > 0.$$

Assume that boundaries of subdomains Lipschitz.

$$\begin{aligned} -\Delta u &= f \qquad \text{in } \Omega, \\ u &= 0 \qquad \text{on } \partial\Omega. \end{aligned}$$

Under suitable assumptions on f (square integrable) and boundaries of the subdomains (Lipschitz) the Poisson problem is equivalent to a coupled problem:

$$
\begin{aligned}
-\Delta u_1 &= f && \text{in } \Omega_1, \\
u_1 &= 0 && \text{on } \partial\Omega_1 \setminus \Gamma, \\
u_1 &= u_2 && \text{on } \Gamma, \\
\frac{\partial u_1}{\partial n_1} &= -\frac{\partial u_2}{\partial n_2} && \text{on } \Gamma, \\
-\Delta u_2 &= f && \text{in } \Omega_2, \\
u_2 &= 0 && \text{on } \partial\Omega_2 \setminus \Gamma.
\end{aligned}
$$

u_i the restriction of u to Ω_i and n_i outward normal to Ω_i. Conditions on interface Γ are *transmission conditions*. Equivalent to the equality of any two independent linear combinations of the traces of the functions and their normal derivatives. By eliminating the interior variables, the transmission conditions gives Poincaré-Steklov operators similar to Schur complements.

Refer to normal derivative as the *flux*.

Approximation of flux: ϕ_j is a nodal basis function for a node on Γ

$$\int_\Gamma \frac{\partial u_i}{\partial n_i}\phi_j \, ds = \int_{\Omega_i} (\Delta u_i \phi_j + \nabla u_i \cdot \nabla \phi_j) \, dx = \int_{\Omega_i} (-f\phi_j + \nabla u_i \cdot \nabla \phi_j) \, dx.$$

In finite element language:

$$\lambda^{(i)} = A_{\Gamma I}^{(i)} u_I^{(i)} + A_{\Gamma\Gamma}^{(i)} u_\Gamma^{(i)} - f_\Gamma^{(i)}.$$

Coincides with residual corresponding to nodes on Γ of a Poisson problem with Neumann condition on Γ.

Setting fluxes equal gives the third equation in the previous block system.

A Dirichlet-Neumann Method

In terms of differential operators, for $n \geq 0$:

$$
\begin{aligned}
(D)\ &\begin{cases} -\Delta u_1^{n+1/2} &= f & \text{in } \Omega_1, \\ u_1^{n+1/2} &= 0 & \text{on } \partial\Omega_1 \setminus \Gamma, \\ u_1^{n+1/2} &= u_\Gamma^n & \text{on } \Gamma, \end{cases} \\
(N)\ &\begin{cases} -\Delta u_2^{n+1} &= f & \text{in } \Omega_2, \\ u_2^{n+1} &= 0 & \text{on } \partial\Omega_2 \setminus \Gamma, \\ \dfrac{\partial u_2^{n+1}}{\partial n_2} &= -\dfrac{\partial u_1^{n+1/2}}{\partial n_1} & \text{on } \Gamma, \end{cases} \\
u_\Gamma^{n+1} &= \theta u_2^{n+1} + (1-\theta) u_\Gamma^n & \text{on } \Gamma,
\end{aligned}
$$

We can also use conjugate gradients.

After working with matrices, we find that the finite element version gives:

$$S^{(2)}(u_\Gamma^{n+1} - u_\Gamma^n) = \theta(g_\Gamma - S u_\Gamma^n),$$

Preconditioned operator: ${S^{(2)}}^{-1} S = I + {S^{(2)}}^{-1} S^{(1)}$.

Theory: Prove that $S^{(1)}$ is spectrally equivalent to $S^{(2)}$, i.e.,

$$c u_\Gamma^T S^{(2)} u_\Gamma \leq u_\Gamma^T S^{(1)} u_\Gamma \leq C u_\Gamma^T S^{(2)} u_\Gamma.$$

Then use the right inequality.

Neumann-Neumann and FETI algorithms can be described using the same frame work. The preconditioner for N-N: ${S^{(1)}}^{-1} + {S^{(2)}}^{-1}$. The preconditioned FETI operator is

$$(S^{(1)} + S^{(2)})({S^{(1)}}^{-1} + {S^{(2)}}^{-1}).$$

The proofs of the optimality of these methods reduces to an extension theorem for finite elements: Given any value on Γ, estimate the energy of contributed by one subdomain in terms of that of the other.

Many subdomains

Now Ω is partitioned into a family of non-overlapping subdomains $\{\Omega_i,\ 1 \leq i \leq N\}$ with

$$\overline{\Omega} = \bigcup_i \overline{\Omega_i}; \quad \Omega_i \cap \Omega_j = \emptyset \quad i \neq j.$$

If $\Gamma_i = \partial\Omega_i \setminus \partial\Omega$, the interface Γ is defined as

$$\Gamma = \bigcup_i \Gamma_i.$$

Linear system written as

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ A_{\Gamma I} & A_{\Gamma\Gamma} \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_I \\ f_\Gamma \end{pmatrix}.$$

We have, interior degrees of freedom, collected in u_I and those on Γ in u_Γ. Block Gauss elimination, in parallel across the subdomains, gives

$$\begin{pmatrix} A_{II} & A_{I\Gamma} \\ 0 & S \end{pmatrix} \begin{pmatrix} u_I \\ u_\Gamma \end{pmatrix} = \begin{pmatrix} f_I \\ g_\Gamma \end{pmatrix}.$$

The Schur complement S and vector g_Γ subassembled from subdomain quantities. The restriction operators R_i, of zeros and ones, map values on Γ onto those on $\Gamma_i := \partial\Omega_i \cap \Gamma$. Then,

$$S = \sum_{i=1}^{N} R_i^T S^{(i)} R_i,$$
$$g_\Gamma = \sum_{i=1}^{N} R_i^T (f_\Gamma^{(i)} - A_{\Gamma I}^{(i)} {A_{II}^{(i)}}^{-1} f_I^{(i)}).$$

Neumann-Neumann and Dirichlet-Neumann

How to precondition S? Try, for N-N,

$$S_{NN}^{-1} S = \sum_{i=1}^{N} R_i^T {S^{(i)}}^{-1} R_i \; S.$$

Not scalable; no mechanism for global transportation of information across the domain in each iteration step. The number of steps required for good progress with conjugate gradients is at least on the order of $1/H$. Also some $S^{(i)}$ singular.

Color subdomains red and black. Use Dirichlet conditions on black and Neumann on red and glue together the red subdomains at the cross points. Gives scalable algorithm in 2D. Condition number bound: $C(1+\log(H/h))^2$.

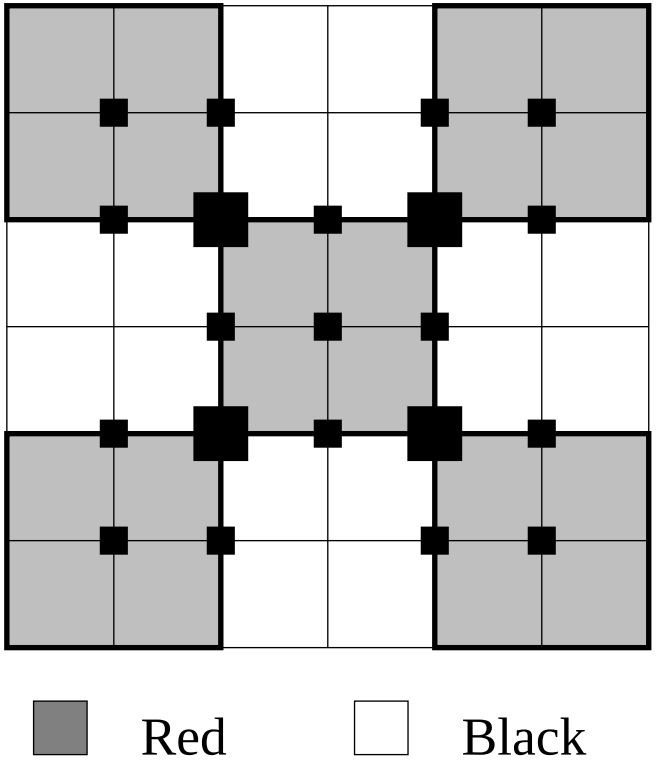
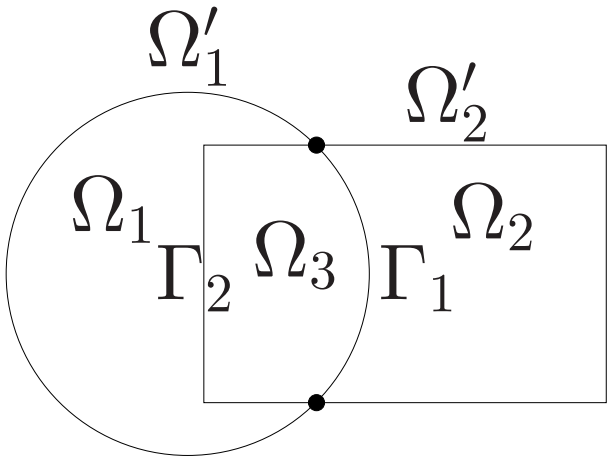


Figure 2: Red-black coloring of the subdomains.

The Schwarz Alternating Method



Dates back to 1870 and H.A. Schwarz.

Given an initial guess u^0, which vanishes on $\partial\Omega$, the iterate u^{n+1} is determined from the previous iterate u^n in two sequential steps.

$$\begin{cases} -\Delta u^{n+1/2} = f & \text{in } \Omega_1', \\ u^{n+1/2} = u^n & \text{on } \partial\Omega_1', \\ u^{n+1/2} = u^n & \text{in } \Omega_2 = \Omega_2' \setminus \overline{\Omega_1'}, \\ -\Delta u^{n+1} = f & \text{in } \Omega_2', \\ u^{n+1} = u^{n+1/2} & \text{on } \partial\Omega_2', \\ u^{n+1} = u^{n+1/2} & \text{in } \Omega_1 = \Omega_1' \setminus \overline{\Omega_2'}\,. \end{cases}$$

We can also write this algorithm in terms of projections onto subspaces:

$$u^{n+1} - u = (I - P_2)(u^{n+1/2} - u) = (I - P_2)(I - P_1)(u^n - u),$$

where $P_i := R_i^T A_i^{-1} R_i A$. This is the basic multiplicative Schwarz method.

Extends immediately to many subdomains by recursion; Schwarz did. We are solving

$$P_{mu}u := (P_1 + P_2 - P_2P_1)u = g.$$

We could simplify and just use the two linear terms. We get the basic additive (parallel) Schwarz method:

$$P_{ad}u := (P_1 + P_2)u = g_{ad}.$$

This is a symmetric operator even for more than two subdomains.

There are other symmetric Schwarz methods such as, for three subdomains,

$$(I - P_1)(I - P_2)(I - P_3)(I - P_2)(I - P_1).$$

There are at least three ways of analyzing the Schwarz methods.

Schwarz used a maximum principle in 1870.

We can use an abstract Schwarz theory to be briefly discussed.

For two subdomains, one can also argue about Schur complements and show,

$$e_{\Gamma_1}^{n+1} = \left(I - (S_{\Gamma_1}^{(2)} + S_{\Gamma_1}^{(3)})^{-1} (S_{\Gamma_1}^{(1)} + S_{\Gamma_1}^{(2)}) \right) e_{\Gamma_1}^{n}.$$

We can view the iteration in terms of an update of the values on Γ_1. The Schur complements correspond to Ω_1', Ω_2, and Ω_3.

Block Jacobi Preconditioners

Precondition A by

$$A_J^{-1} = \begin{pmatrix} A_1^{-1} & 0 \\ 0 & 0 \end{pmatrix} + \begin{pmatrix} 0 & 0 \\ 0 & A_2^{-1} \end{pmatrix} = \begin{pmatrix} A_1 & 0 \\ 0 & A_2 \end{pmatrix}^{-1}.$$

Here $A_i = R_i A R_i^T, \quad i = 1, 2$, and the space is split into two subspaces: $V = R_1^T V_1 \oplus R_2^T V_2$. We can write the preconditioned operator as $P_{ad} = A_J^{-1} A$, i.e., as an additive Schwarz operator. We can also introduce overlap to enhance the convergence. Note that A_J is obtained by a classical splitting, by removing off diagonal block. With overlap, the formulas needs to be modified a little. We can of course also use three or more subregions.

The convergence rate of the block Jacobi method can be estimated by a generalized Rayleigh quotient. We find that

$$\begin{aligned} u^T A(P_{ad}^{-1})u &= u^T(R_1^T A_1^{-1} R_1 + R_2^T A_2^{-1} R_2)^{-1} u \\ &= u^T(R_1^T A_1 R_1 + R_2^T A_2 R_2)u \\ &= u_1^T A_1 u_1 + u_2^T A_2 u_2. \end{aligned}$$

An estimate

$$u_1^T A_1 u_1 + u_2^T A_2 u_2 \leq C_0^2 \, u^T A u,$$

provides the bound

$$\sup_{u \in V} \frac{u^T A(P_{ad}^{-1})u}{u^T A u} \leq C_0^2,$$

and thus a lower bound for the smallest eigenvalue of P_{ad}. There is an upper bound of 2; P_{ad} is a sum of two projections.

The same type of bounds are equally relevant for overlapping domains. There is an abstract, relatively elementary theory: Estimate C_0^2 such that

$$\sum_0^N a(R_i^T u_i, R_i^T u_i) \leq C_0^2 a(u,u), \ \forall u \ \text{ with } u = \sum_0^N R_i^T u_i, \ u_i \in V_i.$$

The smallest $C_0^2 = 1/\lambda_{min}(P_{ad})$.

$$a(u,u) = \sum a(u, R_i^T u_i) = \sum a(P_i u, R_i^T u) \leq$$
$$(\sum (P_i u, P_i u))^{1/2} (\sum a(R_i^T u_i, R_i^T u_i))^{1/2}.$$

Then, $a(u,u) \leq C_0^2 a(P_{ad} u, u)$.

An upper bound for P_{ad} can be obtained by coloring the subspaces V_i.

Overlapping Schwarz methods, for many subdomains Ω_i', should be improved by introducing a coarse component of the preconditioner defined on a coarse triangulation $\mathcal{T}_H$ and with a coarse space $V_0 = V^H$. This can be done even without $\mathcal{T}_h$ being a refinement of $\mathcal{T}_H$, at a cost of more complicated programming. The coarse mesh size should be comparable to the diameters of the subdomains. The basic, sharp result for second order elliptic problems is:

$$\kappa(P_{ad}) \leq C \left(1 + \frac{H}{\delta}\right),$$

where δ measures the overlap between the neighboring subdomains. The proof does not work if the material properties change a lot across the interface Γ.

In the decomposition, the local components can be defined by

$$u_i = R_i(I^h(\theta_i w)) \in V_i, \quad 1 \leq i \leq N,$$

where $w = u - u_0$. Here $u_0 \in V_0$ is a coarse interpolant or quasi-interpolant of u; this operator should reproduce constants. The $\{\theta_i\}$ is a piecewise linear partition of unity associated with the overlapping partition. We have $\sum_1^N \theta_i(x) = 1$, and $|\nabla \theta_i| \leq C/\delta$. In the proof, we use Poincaré's and Friedrichs' inequality. Without a coarse component, we will have large L_2-terms and a poor convergence rates: Consider $\nabla(\theta_i u)$; we obtain a large coefficient in front of one term.

Linear Elasticity

Find displacement $\mathbf{u} \in \mathbf{V}$ of the domain Ω, fixed along $\partial\Omega_D$, with a surface force of density $\mathbf{g}$, along $\partial\Omega_N = \partial\Omega \setminus \partial\Omega_D$, and a body force $\mathbf{f}$:

$$2\int_\Omega \mu\ \epsilon(\mathbf{u}) : \epsilon(\mathbf{v})\ dx + \int_\Omega \lambda\ \text{div}\ \boldsymbol{u}\ \text{div}\ \boldsymbol{v}\ dx\ = < \mathbf{F}, \mathbf{v} > \quad \forall \mathbf{v} \in \mathbf{V}.$$

Here $\lambda(x)$ and $\mu(x)$ are the Lamé parameters and

$$\epsilon_{ij}(\mathbf{u}) = \frac{1}{2}(\frac{\partial u_i}{\partial x_j} + \frac{\partial u_j}{\partial x_i})$$

the linearized strain tensor.

Two inner products defined by

$$\epsilon(\mathbf{u}) : \epsilon(\mathbf{v}) = \sum_{i=1}^{3}\sum_{j=1}^{3} \epsilon_{ij}(\mathbf{u})\epsilon_{ij}(\mathbf{v}),$$

$$< \mathbf{F}, \mathbf{v} > = \int_{\Omega} \sum_{i=1}^{3} f_i v_i \; dx + \int_{\partial\Omega_N} \sum_{i=1}^{3} g_i v_i \; dA.$$

The Lamé parameters expressed in terms of the Poisson ratio ν and Young's modulus E:

$$\lambda = \frac{E\nu}{(1+\nu)(1-2\nu)}, \qquad \mu = \frac{E}{2(1+\nu)}.$$

When $\nu \rightarrow 1/2$, we go to the incompressible limit.

Rigid Body Modes and Korn's Inequality

For $n = 3$, there are six rigid body modes with zero energy, three translations

$$r_1 := \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \quad r_2 := \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \quad r_3 := \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}$$

and three rotations

$$r_4 := \frac{1}{H_i} \begin{bmatrix} 0 \\ -x_3 + \hat{x}_3 \\ x_2 - \hat{x}_2 \end{bmatrix}, r_5 := \frac{1}{H_i} \begin{bmatrix} x_3 - \hat{x}_3 \\ 0 \\ -x_1 + \hat{x}_1 \end{bmatrix}, r_6 := \frac{1}{H_i} \begin{bmatrix} -x_2 + \hat{x}_2 \\ x_1 - \hat{x}_1 \\ 0 \end{bmatrix},$$

where $\hat{x}$ is a shift.

Poincaré's inequality replaced by Korn's second inequality

$$\|\mathbf{v}\|^2_{\boldsymbol{H}^1(\Omega_i)} \leq C \left(a_i(\mathbf{v},\mathbf{v}) + \frac{1}{H_i^2} \|\mathbf{v}\|^2_{\boldsymbol{L}^2(\Omega_i)} \right).$$

We also have, more importantly, with $\mathcal{RB}$, the space of rigid body modes,

$$\inf_{\mathbf{r}\in\mathcal{RB}} \|\mathbf{v}-\mathbf{r}\|^2_{\boldsymbol{H}^1(\Omega_i)} \leq C a_i(\mathbf{v},\mathbf{v}).$$

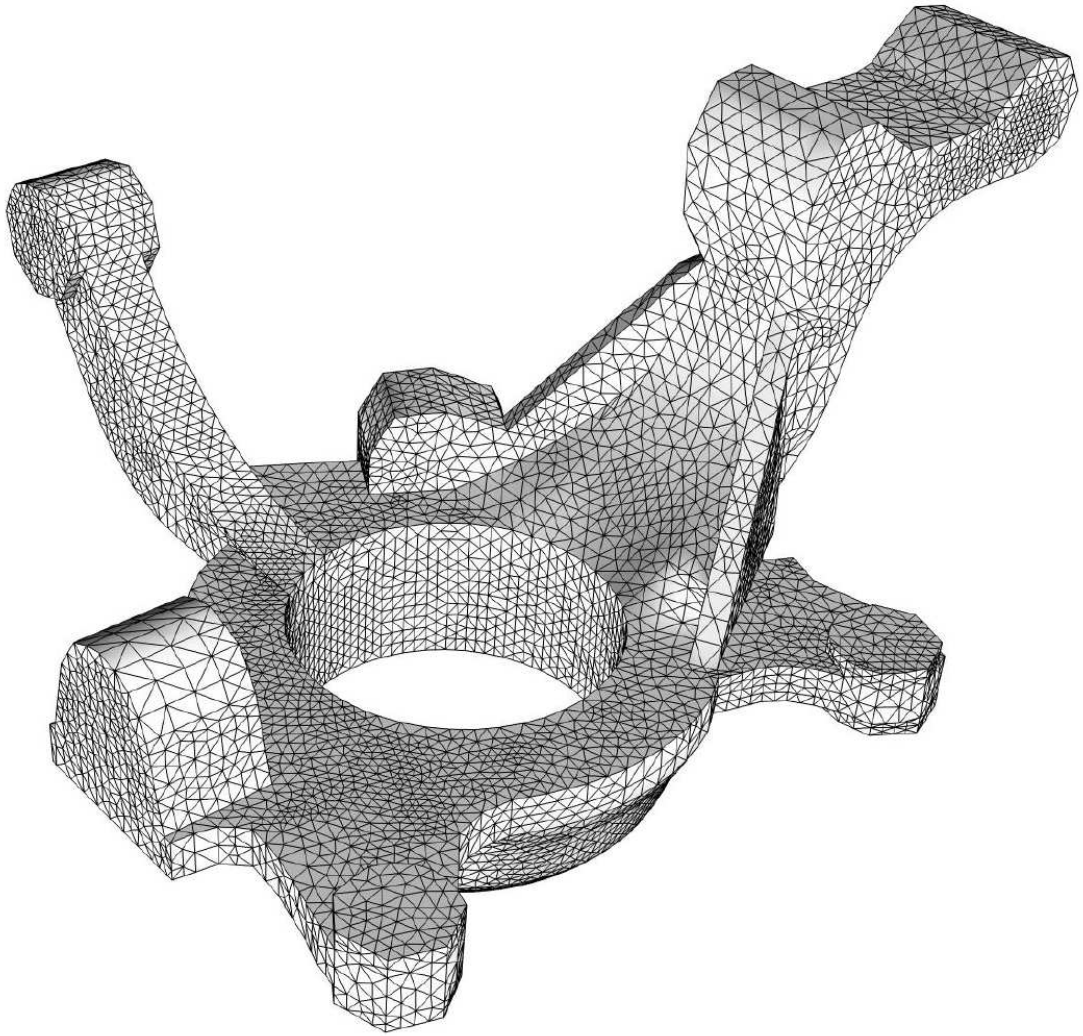


Figure 3: Finite element meshing of a mechanical object.

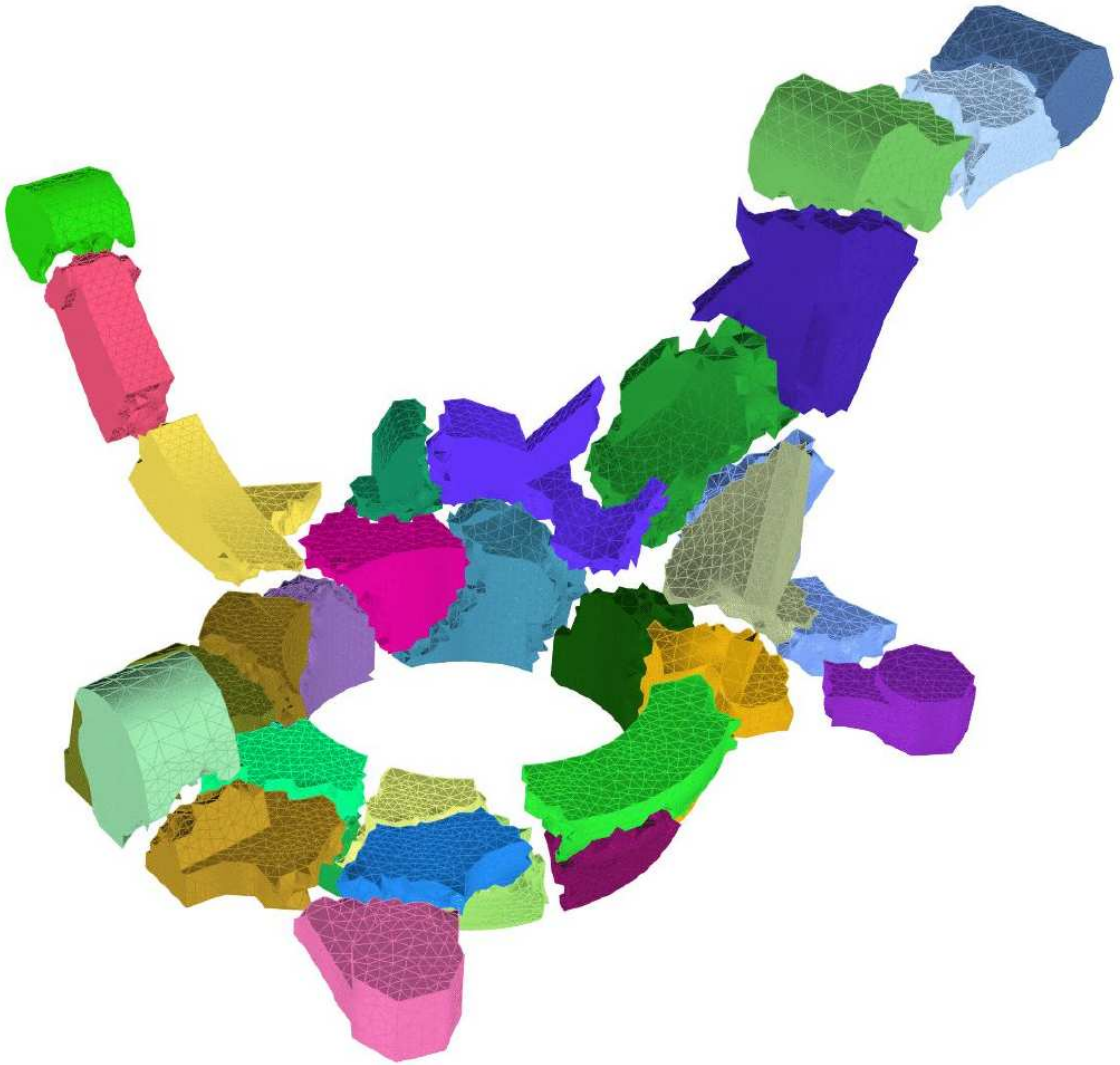


Figure 4: Partition into thirty subdomains. Courtesy Charbel Farhat.

These subdomains effectively provide our coarse mesh.

Faces, edges, and vertices of quite general subdomains can be defined in terms of certain equivalence classes of finite element nodes. These geometric objects are central in an alternative construction of coarse problems and in the theory. Also highly relevant for parallel computing.

We will use face, edge, and vertex functions, providing a partition of unity on the interface. A face function θ_{Fij} equals 1 at all nodes of a face common to two subdomains Ω_i and Ω_j and vanishes at all other interface nodes. They are extended as $discrete\ harmonic$ functions, i.e., with minimal elastic energy; this determines the values at interior nodes. Similarly, we have edge functions and vertex functions. The restriction of the rigid body modes – all linear functions – to faces and edges, are used for problems of elasticity. The coarse space needs to accommodate all rigid body modes.

Alternative Overlapping Schwarz Methods

Consider a scalar elliptic problem, in three dimensions, and a coarse space which is the range of the interpolation operator

$$I_B^h u(x) = \sum_{V^k \in \Gamma} u(V^k)\theta_{V_k}(x) + \sum_{E^\ell \subset \Gamma} \bar{u}_{E^\ell}\theta_{E^\ell}(x) + \sum_{F^{ij} \subset \Gamma} \bar{u}_{F^{ij}}\theta_{F^{ij}}(x).$$

Here $\bar{u}_{E^i}$ and $\bar{u}_{F^k}$ are averages over edges and faces of the subdomains.

$\theta_{V_k}(x)$ is essentially the standard nodal basis functions of a vertex of the subdomains, $\theta_{E^\ell}(x) = 1$ at the nodes of the edge E^i and vanishes at all other interface nodes, and $\theta_{F^{ij}}(x)$ is the similar function already defined for the face F^{ij}. These functions are extended as discrete harmonic functions into the interior of the subdomains. This interpolation operator I_B^h reproduces constants.

For "nice enough" subregions, we have

$$|u - I_B^h u|^2_{H^1(\Omega_i)} \leq C(1 + \log(H_i/h_i))|u|^2_{H^1(\Omega_i)}.$$

We use Cauchy-Schwarz inequality, a trace theorem, bounds on face and edge functions, and finite element Sobolev inequalities. We can now also handle coefficient jumps across the interface.

As noted, good spaces for elasticity are obtained by multiplying the rigid body modes by the face and edge functions. The interpolation operator can then preserve all rigid body modes. The coefficients built from averages and first order moments. Results in a large coarse space.

We can shrink the coarse space by replacing vertex, edge and face contributions by fewer terms. The new coarse basis functions are defined as linear combinations of those of the larger space and in terms of simple least squares problems. The dimension of this coarse space can be made about half of that of the older one.

Submatrices of assembled stiffness matrices can be used to compute the interior values of the basis elements of the coarse space.

The region is also covered by overlapping subregions Ω_i'. δ_i/H_i measures the relative overlap between adjacent subregions, each of which is a union of elements. The local spaces chosen for the Schwarz methods are

$$V_i = V^h \cap H_0^1(\Omega_i'), \quad i > 0.$$

The standard overlapping subdomains Ω_i' are obtained by repeatedly adding layers of elements starting with Ω_i.

Another interesting choice is to work with the Ω_i and $\Omega_{i\delta}$. The latter obtained by adding layers of elements on both sides of $\Gamma_i := \partial\Omega_i \cap \Gamma$. By using a hybrid Schwarz methods, we can make all residuals interior to the Ω_i vanish in each step.

Schwarz Methods

Schwarz 1870, Pierre-Louis Lions 1987: $(I - P_2)(I - P_1)$.

Standard two-level additive 1988: $P_0 + \sum_{i \geq 1} P_i$.

Standard hybrid, e.g., as in balancing N-N: $P_0 + (I - P_0) \sum_{i \geq 1} P_i (I - P_0)$.

New hybrid: $(I - \sum_{i \geq 1} P_i)(P_0 + \sum_{i \geq 1} P_{i\delta})(I - \sum_{i \geq 1} P_i)$.

Note that $(I - \sum_{i \geq 1} P_i)$ is a projection since the subdomains Ω_i do not intersect. Therefore, after the first iteration, we need only apply this operator once per step.

Also note that the residuals vanish in the interior of the subdomains, which allow us to save storage.

Result for Overlapping Schwarz Method

Theorem *The condition number of the preconditioned operator T_{as} satisfies*

$$\kappa(P_{ad}) \leq C(1 + H/\delta)(1 + \log(H/h)).$$

Here C is independent of the mesh size, the number of subdomains, the Lamé parameters, as long as the material compressible. H/δ measures the relative overlap between neighboring overlapping subregions, and H/h the maximum number of elements across any subregion.

What needs to be done in the almost incompressible case?. That is the story for another day.

FETI and FETI–DP

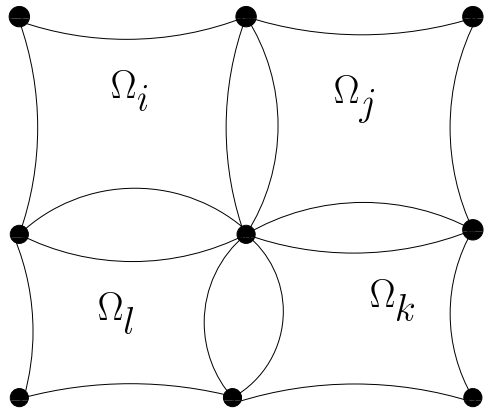
Introduce Lagrange multipliers $\lambda \in U := \mathbf{range}(B_\Gamma)$ and consider the problem:

Find $(u, \lambda) \in W \times U$, such that

$$\left.\begin{array}{rcrcl} Au & + & B_\Gamma^T \lambda & = & f \\ B_\Gamma u & & & = & 0 \end{array}\right\}$$

Eliminate displacement u by block Gauss elimination; solve resulting Schur system by PCG. Block diagonal matrix A, in general, only positive semidefinite. Enforce continuity constraints on $primal$ displacement variables u_Π throughout iterations (as in primal methods); other constraints, on u_Δ, enforced by Lagrange multipliers λ. Local problems invertible; primal variables provide a coarse problem.

Two dimensions. Maintain continuity of primal variables at vertices (subassembly); enforce continuity constraints elsewhere by Lagrange multipliers.



$$\begin{bmatrix} A_{II}^{(1)} & A_{I\Delta}^{(1)} & & & & \tilde{A}_{\Pi I}^{(1)T} \\ A_{\Delta I}^{(1)} & A_{\Delta\Delta}^{(1)} & & & & \tilde{A}_{\Pi\Delta}^{(1)T} \\ & & \ddots & & & \vdots \\ & & & A_{II}^{(N)} & A_{I\Delta}^{(N)} & \tilde{A}_{\Pi I}^{(N)T} \\ & & & A_{\Delta I}^{(N)} & A_{\Delta\Delta}^{(N)} & \tilde{A}_{\Pi\Delta}^{(N)T} \\ \tilde{A}_{\Pi I}^{(1)} & \tilde{A}_{\Pi\Delta}^{(1)} & \cdots & \tilde{A}_{\Pi I}^{(N)} & \tilde{A}_{\Pi\Delta}^{(N)} & \tilde{A}_{\Pi\Pi} \end{bmatrix}$$

$$\begin{bmatrix} A_{BB} & \tilde{A}_{\Pi B}^T & B^T \\ \tilde{A}_{\Pi B} & \tilde{A}_{\Pi\Pi} & O \\ B & O & O \end{bmatrix} \begin{bmatrix} \mathbf{u}_B \\ \tilde{\mathbf{u}}_\Pi \\ \boldsymbol{\lambda} \end{bmatrix} = \begin{bmatrix} \mathbf{f}_B \\ \tilde{\mathbf{f}}_\Pi \\ 0 \end{bmatrix}$$

FETI–DP in 3D

Good numerical results in 2D; not always very good in 3D. Idea: In addition to (or instead of) continuity of primal variables at vertices, constrain certain average values (and moments) of primal variables over individual edges and faces to take common values across the interface.

For scalar second order elliptic equations, this approach yields a condition number estimate $C(1+\log(H/h)^2)$ for certain families of algorithms. Result independent of jumps in coefficients, if scaling is chosen carefully. There are algorithms with quite small coarse problems, i.e., relatively few primal constraints.

Reliable recipes exist to select sets of primal constraints for elasticity in 3D primarily using edge averages and first order moments as primal constraints. High quality PETSc-based codes have been developed and successfully tested on very large parallel computing systems.

N-N Methods of Same Flavor: BDDC

Introduce a coarse basis function for each primal constraint; set one primal variable $= 1$ and all others $= 0$, one at a time. Extend with minimum energy one subdomain at a time. Results in basis functions discontinuous across the interface Γ.

One local subspace for each subdomain. Set all relevant primal degrees of freedom $= 0$. Local problems invertible.

Partially subassembled Schur complement of the system is block diagonal. Apply E_D^T to residual. Solve the linear systems corresponding to these blocks exactly, and compute a weighted average, with operator E_D, of results, across the interface. Only one block assembled. Compute residual, remove the interior residuals, and repeat coarse and local solves. Accelerate with the preconditioned conjugate gradient method. The theory can be focused on estimate of norm of E_D.

Matrix Analysis of FETI–DP and BDDC

Consider three product space of finite element functions/vectors of nodal values.

$$\widehat{W}_\Gamma \subset \widetilde{W}_\Gamma \subset W.$$

W: no constraints; $\widehat{W}_\Gamma$: continuity at every point on Γ; $\widetilde{W}_\Gamma$: common values of primal variables.

Change variables, explicitly introducing primal variables and complementary sets of dual displacement variables. Write Schur complements as

$$S^{(i)} = \begin{pmatrix} S^{(i)}_{\Delta\Delta} & S^{(i)}_{\Delta\Pi} \\ S^{(i)}_{\Pi\Delta} & S^{(i)}_{\Pi\Pi} \end{pmatrix}.$$

Let $\tilde{S}_\Gamma$ denote the partially assembled Schur complement. (In practice, work with interior variables as well when solving linear systems.)

BDDC matrices

For the BDDC method, we use the fully assembled Schur complement $\tilde{R}_\Gamma^T \tilde{S}_\Gamma \tilde{R}_\Gamma$; it is used to compute the residual. Using the preconditioner involves solving a system with the matrix $\tilde{S}_\Gamma$:

$$M_{BDDC}^{-1} = \tilde{R}_{D\Gamma}^T \tilde{S}_\Gamma^{-1} \tilde{R}_{D\Gamma},$$

where $\tilde{R}_{D\Gamma}$ is a scaled variant of $\tilde{R}_\Gamma$ with scale factors computed from the PDE coefficients.

Scaling chosen so that $E_D := \tilde{R}_\Gamma \tilde{R}_{D\Gamma}^T$ is a projection.

FETI–DP Matrices

The basic operator is now $B_\Delta \tilde{S}^{-1} B_\Delta^T$. $\tilde{S}$ is a Schur complement of $\tilde{S}_\Gamma$ obtained after eliminating all primal variables. It is elementary to show that $\tilde{S}^{-1} = R_{\Gamma\Delta} \tilde{S}_\Gamma^{-1} R_{\Gamma\Delta}^T$, where $R_{\Gamma\Delta}$ removes the primal part of a vector defined on Γ.

The preconditioner is now

$$M_{FETI}^{-1} = B_{D\Delta} S_{\Delta\Delta} B_{D\Delta}^T,$$

where $S_{\Delta\Delta} = R_{\Gamma\Delta} \tilde{S}_\Gamma R_{\Gamma\Delta}^T$ is the Δ block of $\tilde{S}_\Gamma$ and $B_{D\Delta}$ is a scaled jump operator. Scale factors depend on PDE coefficients.

Scaling chosen so that $P_D := R_{\Gamma\Delta}^T B_{D\Delta}^T B_\Delta R_{\Gamma\Delta}$ is a projection.
Also, $E_D + P_D = I$ and $E_D P_D = P_D E_D = 0$.

Same Eigenvalues

FETI–DP preconditioned operator:

$$B_{D\Delta}S_{\Delta\Delta}B_{D\Delta}^T * B_\Delta \tilde{S}^{-1} B_\Delta^T = B_{D\Delta}R_{\Gamma\Delta}\tilde{S}_\Gamma R_{\Gamma\Delta}^T B_{D\Delta}^T * B_\Delta R_{\Gamma\Delta}\tilde{S}_\Gamma^{-1} R_{\Gamma\Delta}^T B_\Delta^T.$$

Multiply by $R_{\Gamma\Delta}^T B_\Delta^T$ on left and remove same factor on right:

$$P_D^T \tilde{S}_\Gamma P_D \tilde{S}_\Gamma^{-1}.$$

BDDC preconditioned operator:

$$\tilde{R}_{D\Gamma}^T \tilde{S}_\Gamma^{-1} \tilde{R}_{D\Gamma} * \tilde{R}_\Gamma^T \tilde{S}_\Gamma \tilde{R}_\Gamma.$$

Multiply by $\tilde{R}_\Gamma$ on left and remove same factor on right:

$$E_D \tilde{S}_\Gamma^{-1} E_D^T \tilde{S}_\Gamma.$$

Let φ be an eigenvector of $P_D^T \tilde{S}_\Gamma P_D \tilde{S}_\Gamma^{-1}$ with eigenvalue λ.
Let $\psi = E_D \tilde{S}_\Gamma^{-1} \varphi$.

$$E_D \tilde{S}_\Gamma^{-1} E_D^T \tilde{S}_\Gamma * E_D \tilde{S}_\Gamma^{-1} \varphi = E_D \tilde{S}_\Gamma^{-1} (I - P_D^T) \tilde{S}_\Gamma (I - P_D) \tilde{S}_\Gamma^{-1} \varphi.$$

Gives three terms

$$E_D \tilde{S}_\Gamma^{-1} P_D^T \tilde{S}_\Gamma P_D \tilde{S}_\Gamma^{-1} \varphi = \lambda E_D \tilde{S}_\Gamma^{-1} \varphi$$

and

$$-E_D P_D \tilde{S}_\Gamma^{-1} \varphi + E_D \tilde{S}_\Gamma^{-1} (I - P_D^T) \varphi.$$

$E_D P_D = 0$. $(I - P_D^T)\varphi = 0$ since $\varphi \in \mathbf{range}(P_D^T)$. Similarly, any eigenvalue of the BDDC operator is an eigenvalue of the FETI–DP operator.

What we just did is not quite correct. BDDC always has an eigenvalue equal to 1; FETI–DP does not always.

The analysis of BDDC requires a bound of the $\widetilde{S}-$norm of the average operator E_D. Interestingly enough, a main role, in 2D, is played by the special edge functions θ_E and by the finite element extension theorem. Both were encountered previously. In 3D, the face functions θ_F also comes into play. A sharp $C(1+\log(H/h))^2$ condition number estimate results if the primal constraints and certain scale factors are chosen carefully.