Designing Domain Decomposition Methods Based on Cholesky’s Algorithm

Olof Widlund
Courant Institute of Mathematical Sciences
New York University
http://www.cs.nyu.edu/cs/faculty/widlund/
Based on joint work with
with Jing Li and
Axel Klawonn and Oliver Rheinbach
Block Cholesky Elimination

Consider a block matrix, assumed positive definite, symmetric.

\[
\begin{bmatrix}
A & B^T \\
B & C
\end{bmatrix}.
\]

It can be factored by block Cholesky:

\[
\begin{bmatrix}
A & B^T \\
B & C
\end{bmatrix} = \begin{bmatrix}
I_A & I_C \\
BA^{-1} & I_C
\end{bmatrix} \begin{bmatrix}
A & C - BA^{-1}B^T \\
B & S
\end{bmatrix} \begin{bmatrix}
I_A & A^{-1}B^T \\
B & I_C
\end{bmatrix},
\]

where \(I_A\) and \(I_C\) are appropriate identity matrices. \(S = C - BA^{-1}B^T\) is a Schur complement. Inverting, we find that

\[
\begin{bmatrix}
A & B^T \\
B & C
\end{bmatrix}^{-1} = \begin{bmatrix}
I_A & -A^{-1}B^T \\
B & I_C
\end{bmatrix} \begin{bmatrix}
A^{-1} & S^{-1} \\
B & -BA^{-1}
\end{bmatrix} \begin{bmatrix}
I_A \\
B & I_C
\end{bmatrix}.
\]
By changing basis, we can reduce matrix to a block diagonal form.

Cholesky’s algorithm is used extensively in finite element practice. It is often helpful to apply the block ideas recursively by, e.g., writing the matrix $A$ as a two-by-two block matrix. In fact, $A$ is often block diagonal and we can then use recursion to deal with each block separately.

Generally, the computation of and factoring of $S$ can be very expensive and less amenable to parallelization than other parts. Explore the possibility to decrease the size of the Schur complement. Then the solver will only provide an inexact inverse, a preconditioner. The preconditioner will be applied in each step of an iterative process of conjugate gradient type and it will be crucial that the preconditioned operator is well conditioned, which will translate into rapid convergence. Very desirable to work exclusively with positive definite matrices. (Research on domain decomposition theory is almost entirely focused on establishing bounds for the condition numbers.)
Two Subdomains

Consider a domain $\Omega$ subdivided into two nonoverlapping subdomains $\Omega_1$ and $\Omega_2$. In between the interface $\Gamma$.

Consider a finite element approximation of a selfadjoint elliptic problem on $\Omega$ (scalar elliptic, linear elasticity, or even an incompressible Stokes problem.)

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

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

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

Subassemble:

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

\[A_{\Gamma} = A_{\Gamma}^{(1)} + A_{\Gamma}^{(2)}.\]

Degrees of freedom partitioned into those internal to \(\Omega_1\), and \(\Omega_2\), and those on \(\Gamma\).

Eliminate the interior unknowns. Gives two Schur complements:

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

The given system is reduced to

\[Su_{\Gamma} = (S^{(1)} + S^{(2)})u_{\Gamma} = g_{\Gamma}.\]
We can equally well consider a three dimensional lemon.

If we have Neumann boundary conditions on all of the boundary of a subdomain, it is *floating* and the coefficient matrix and the local Schur complement will be singular. We will return to that case later. Interior subdomains are always floating if there are zero energy modes.
A Neumann-Neumann Method

In this simple case, the basic Neumann-Neumann preconditioner is

\[ M^{-1} = S^{(1)^{-1}} + S^{(2)^{-1}}. \]

Can also use restriction operators, \( R^{(i)} \), and diagonal scaling matrices with \( D^{(1)} + D^{(2)} = I \).

\[ M^{-1} = R^{(1)^T} D^{(1)} S^{(1)^{-1}} D^{(1)} R^{(1)} + R^{(2)^T} D^{(2)} S^{(2)^{-1}} D^{(2)} R^{(2)}. \]  

(1)

Here \( R^{(1)} = R^{(2)} \). Generalizes immediately:

\[ S = \sum R^{(i)^T} S^{(i)} R^{(i)}, \quad M^{-1} = \sum R^{(i)^T} D^{(i)} S^{(i)^{-1}} D^{(i)} R^{(i)}. \]  

(2)

Two problems: some \( S^{(i)} \) singular and no global part of preconditioner.
A FETI Method, Two Subdomains

Consider local mixed Neumann-Dirichlet problems:

\[
\begin{pmatrix}
A_{II}^{(i)} & A_{I\Gamma}^{(i)} \\
A_{\Gamma I}^{(i)} & A_{\Gamma\Gamma}^{(i)}
\end{pmatrix}
\begin{pmatrix}
 u_I^{(i)} \\
u_{\Gamma}^{(i)}
\end{pmatrix}
= \begin{pmatrix}
f_I^{(i)} \\
f_{\Gamma}^{(i)} + \lambda_{\Gamma}^{(i)}
\end{pmatrix}, \quad i = 1, 2.
\]

Here, \( \lambda_{\Gamma} = \lambda_{\Gamma}^{(1)} = -\lambda_{\Gamma}^{(2)} \) is the unknown flux. We obtain,

\[
u_{\Gamma}^{(i)} = S^{(i)-1}(g_{\Gamma}^{(i)} + \lambda_{\Gamma}^{(i)}).
\]

Set \( u_{\Gamma}^{(1)} = u_{\Gamma}^{(2)} \), and obtain \( F\lambda_{\Gamma} = d_{\Gamma} \), with \( F = S^{(1)-1} + S^{(2)-1} \).

Precondition this interface equation with \( S^{(1)} + S^{(2)} \). Weights? Obviously the same spectrum for the two preconditioned operators.
A FETI-DP Method. Two Subdomains.

In 2D FETI–DP algorithms, we maintain continuity at subdomain vertices throughout the whole iteration. In 3D, we need to work with common averages over interface edges or faces. Here we consider a scalar elliptic problem, two subdomains and one average constraint.

Figure 1: Partition into two subdomains, with $\Omega_2$ floating, in the absence of a constraint.
Our lemon, which can be 2D or 3D, now only has Dirichlet boundary conditions on part of one subdomain boundary. In 3D, we could work with a face average or an average over an edge of the face common to the two subdomains. We make a change of variables separating the edge average, the *primal* displacement variable from the *dual* displacement variables. Any dual displacement variable has a zero average. We can also think of the primal variable in terms of a constraint. The primal variables will provide a coarse component for our preconditioners and they live in the lower right corner of the block matrix.

We can again introduce a flux variable $\lambda$ and set up an operator $F$ acting on $\lambda$. A preconditioner based on solving a Dirichlet problem on each subdomain can be designed. The $F\lambda$ requires a solve with Neumann conditions on each subdomain.

One can also design a primal BDDC variant. Restore continuity by averaging the values across the interface. Split the assembled residual too.
Finite Element Model on Subdomains

For all subdomains $\Omega_i$, set up local stiffness matrices $A^{(i)}$ and local load vectors $f^{(i)}$, $i = 1, \ldots, N$,

$$A := \begin{bmatrix} A^{(1)} & \cdots & A^{(N)} \end{bmatrix}, u := \begin{bmatrix} u^{(1)} \\ \vdots \\ u^{(N)} \end{bmatrix}, f := \begin{bmatrix} f^{(1)} \\ \vdots \\ f^{(N)} \end{bmatrix}.$$ 

Define the interface by $\Gamma := \bigcup_{i=1}^{N} \partial \Omega_i \setminus \partial \Omega$.

Discrete problem, to obtain displacements $u = [u^{(1)}, \ldots, u^{(N)}]$, can be formulated as minimization problem with interface continuity constraints

$$B_\Gamma u = 0,$$

where $B_\Gamma = [B_\Gamma^{(1)}, \ldots, B_\Gamma^{(N)}]$ with elements $0, 1, -1$. 


FETI-DP

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

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

\[
\begin{align*}
Au + B_\Gamma^T \lambda &= f \\
B_\Gamma u &= 0
\end{align*}
\]

Eliminate displacement $u$ by block Gauss elimination; solve resulting Schur system by PCG. Block diagonal matrix $A$, in general, only positive semidefinite. Enforce some continuity constraints on primal displacement variables $u_\Pi$ throughout iterations (as in primal methods); other constraints, on $u_\Delta$, enforced by dual Lagrange multipliers $\lambda$ (as in standard FETI). Local problems invertible; provides coarse problem.
History of FETI and FETI-DP


FETI-DP introduced by Farhat/Lesoinne/LeTallec/Pierson/Rixen (2001). Further work by Farhat, Lesoinne, Pierson, Mandel, Tezaur, Brenner, Li, Toselli, Widlund, Dryja, Klawonn, ... h- and hp FEM, BEM, mixed FEM, incompressible Stokes, mortar elements, Maxwell in 3D, ...

Structural mechanics 2nd and 4th order, acoustics, scalar diffusion problems, contact, Stokes, ... Tested for at least 100 million dof on thousands of processors (ASCI Option Red, Sandia National Laboratory, USA). Work by Farhat’s group, also by Oliver Rheinbach, Essen in PETSC.
History of Neumann-Neumann

Early work by Bourgat, Glowinski, De Roeck, LeTallec, and Vidrascu. Introduction of a second level by Mandel and Brezina and by Dryja and Widlund. Used extensively in many large scale applications, in particular, in France by LeTallec and Vidrascu et al.

Extended, in various ways, to convection-diffusion equations by Achdou, LeTallec, Nataf, and Vidrascu, to incompressible Stokes by Pavarino and Widlund and in collaboration with Goldfeld, to almost incompressible elasticity and models of elastic bodies which in part are almost incompressible.

New coarse spaces by Dohrmann, Mandel, and Tezaur. BDDC variants of the N-N algorithms. New theory connecting the spectra FETI-DP and BDDC algorithms. First observed by Fragakis and Papadrakakis for older methods. Also recent work on Stokes and BDDC.
Keep continuity of primal variables at vertices (subassembly); other continuity constraints by Lagrange multipliers.

\[
\begin{bmatrix}
A_{II}^{(1)} & A_{I\Delta}^{(1)} \\
A_{\Delta I}^{(1)} & A_{\Delta\Delta}^{(1)} \\
& & \cdots \\
\tilde{A}_{\Pi I}^{(1)} & \tilde{A}_{\Pi\Delta}^{(1)} & \cdots & \tilde{A}_{\Pi I}^{(N)} & \tilde{A}_{\Pi\Delta}^{(N)} & \tilde{A}_{\Pi\Pi I} \\
\end{bmatrix}
\begin{bmatrix}
\tilde{A}_{\Pi B} \\
\tilde{A}_{\Pi\Pi B} \\
B \\
\tilde{A}_{\Pi I} \\
\tilde{A}_{\Pi\Pi I} \\
O \\
O \\
O \\
\end{bmatrix}
\begin{bmatrix}
u_B \\
\tilde{u}_\Pi \\
\lambda \\
\end{bmatrix}
= 
\begin{bmatrix}
f_B \\
\tilde{f}_\Pi \\
0 \\
\end{bmatrix}
\]
Dual-Primal FETI 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 3D elasticity, minimally six constraints per subdomain.

This approach yields condition number estimate $C(1 + \log(H/h)^2)$ for a family of algorithms. Result independent of jumps in coefficients, if scaling chosen carefully. One of the algorithms can have a quite small coarse problem; Klawonn, Widlund, Dryja (2002). Numerical results for elasticity in 3D using vertex constraints and three face averages on each face (Farhat, Lesoinne, Pierson (2000)). Works numerically; might not be fully scalable. Scalability established for elasticity for somewhat richer primal space. Klawonn and Widlund (2004). PETSc codes by Rheinbach.
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 $\Gamma$. Also one local subspace for each subdomain for which all relevant primal degrees vanish making problems invertible.

Partially subassembled Schur complement of the system is block diagonal after this change of variables. Apply an operator $E_D^T$ to residual. Solve linear systems corresponding to blocks exactly, and compute a weighted average, with operator $E_D$, of results, across the interface. Only one block, with a few variables for each subdomain assembled and factored. Compute residual, remove the interior residuals, and repeat coarse and local solves. Accelerate with conjugate gradient method. Theory focused on estimate of $E_D$. Reexamine 2D case and with the same primal constraints.
Matrix Analysis of FETI-DP and BDDC

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

\[ \tilde{W}_\Gamma \subset \hat{W}_\Gamma \subset W. \]

\( W \): no constraints; \( \hat{W}_\Gamma \): continuity at every point on \( \Gamma \); \( \tilde{W}_\Gamma \): 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 $\tilde{S}_\Gamma$:

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

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

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

The matrix $\tilde{S}_\Gamma$ equally important for FETI-DP.
FETI-DP Matrices

The basic operator is now $B_\Delta \tilde{S}^{-1} B_T^T$. $\tilde{S}$ 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 $\Gamma$.

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 $\Delta$ 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 \( \varphi \) be an eigenvector of \( P_D^T \tilde{\Sigma}_\Gamma P_D \tilde{\Sigma}_\Gamma^{-1} \) with eigenvalue \( \lambda \).

Let \( \psi = E_D \tilde{\Sigma}_\Gamma^{-1} \varphi \).

\[
E_D \tilde{\Sigma}_\Gamma^{-1} E_D^T \tilde{\Sigma}_\Gamma \ast E_D \tilde{\Sigma}_\Gamma^{-1} \varphi = E_D \tilde{\Sigma}_\Gamma^{-1} (I - P_D^T) \tilde{\Sigma}_\Gamma (I - P_D) \tilde{\Sigma}_\Gamma^{-1} \varphi.
\]

Gives three terms

\[
E_D \tilde{\Sigma}_\Gamma^{-1} P_D^T \tilde{\Sigma}_\Gamma P_D \tilde{\Sigma}_\Gamma^{-1} \varphi = \lambda E_D \tilde{\Sigma}_\Gamma^{-1} \varphi
\]

and

\[-E_D P_D \tilde{\Sigma}_\Gamma^{-1} \varphi + E_D \tilde{\Sigma}_\Gamma^{-1} (I - P_D^T) \varphi.
\]

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

We wish to show for FETI–DP that for all $\lambda \in V$

$$\langle M\lambda, \lambda \rangle \leq \langle F\lambda, \lambda \rangle \leq C(1 + \log(H/h))^2 \langle M\lambda, \lambda \rangle.$$  

$F$ is the FETI–DP operator and $M$ is the preconditioner. With $\mathbb{N}_x$ the set of indices $j$ such that $x \in \partial \Omega_j$,

$$(P_D w)^{(i)}(x) := (B_{D\Delta}^T B_{\Delta} w)^{(i)}(x) = \sum_{j \in \mathbb{N}_x} \delta_j^\dagger (w^{(i)}(x) - w^{(j)}(x)), \ x \in \partial \Omega_i \cap \Gamma.$$  

For $x \in \Gamma_h$ and $\gamma \in [1/2, \infty)$,

$$\delta_i(x) = \begin{cases} \frac{\sum_{j \in \mathbb{N}_x} G_j^\gamma(x)}{G_i^\gamma(x)} & x \in \partial \Omega_{i,h}, \\ 0 & x \in (\Gamma_h \cup \partial \Omega_h) \setminus \partial \Omega_{i,h}, \end{cases}$$
and $\delta_j^{\dagger}$ is its pseudo inverse.

The core estimate is

$$|P_D w|_S^2 \leq C(1 + \log(H/h))^2 |w|_S^2 \quad \forall w \in \tilde{W}.$$ 

We could equally well estimate the $E_D$ operator.

We can localize these estimates and consider the contribution to the right hand side from one subdomain $\Omega_i$ at a time; on the left hand side we should sum over the index sets of next neighbors of $\Omega_i$. There is promise to choosing the set of primal constraints adaptly.