THREE-LEVEL BDDC IN TWO DIMENSIONS
XUEMIN TU

Abstract. BDDC methods are nonoverlapping iterative substructuring domain decomposition methods for the solution of large sparse linear algebraic systems arising from discretization of elliptic boundary value problems. They are similar to the balancing Neumann-Neumann algorithm. However, in BDDC methods, a small number of continuity constraints are enforced across the interface, and these constraints form a new coarse, global component. An important advantage of using such constraints is that the Schur complements that arise in the computation will all be strictly positive definite. The matrix of the coarse problem is generated and factored by direct solvers at the beginning of the computation. However, this problem can ultimately become a bottleneck, if the number of subdomains is very large. In this paper, two three-level BDDC methods are introduced for solving the coarse problem approximately in two dimensional space, while still maintaining a good convergence rate. Estimates of the condition numbers are provided for the two three-level BDDC methods and numerical experiments are also discussed.

Key words. BDDC, three-level, domain decomposition, coarse problem, condition number, Chebyshev iteration

AMS subject classifications. 65N30, 65N55

1. Introduction. In this paper, we introduce two three-level BDDC (Balancing Domain Decomposition by Constraints) methods. The BDDC algorithms, so far developed for two levels [2, 7, 8], are similar to the balancing Neumann-Neumann algorithms. However, their coarse problems, in BDDC, are given in terms of a set of primal constraints and they are generated and factored by direct solvers at the beginning of the computation. We note that there are now computer systems with more than 100,000 powerful processors, which allow very large and detailed simulations. Since at least one subdomain will be assigned to each processor, the coarse component of the preconditioners can then ultimately become a bottleneck if the number of processors is very large. We will try to remove this difficulty by using one or several additional levels. We introduce two three-level BDDC methods. Here, we only consider two dimensional problems and vertex constraints. We also provide estimates of the condition numbers of the system with these two new preconditioners. We note that, since this paper was submitted, we have extended our algorithms and results to three dimensional cases; see [11, 12].

The rest of the paper is organized as follows. We first review the two-level BDDC methods briefly in Section 2. We introduce our first three-level BDDC method and the corresponding preconditioner \( \hat{M}^{-1} \) in Section 3. We give some auxiliary results in Section 4. In Section 5, we provide an estimate of the condition number for the system with the preconditioner \( \hat{M}^{-1} \) which is of the form \( C \left( 1 + \log \frac{\hat{H}}{H} \right)^2 \left( 1 + \log \frac{H}{h} \right)^2 \), where \( \hat{H}, H, \) and \( h \) are the diameters of the subregions, subdomains, and elements, respectively. In Section 6, we introduce a second three-level BDDC method which uses Chebyshev iterations. We denote the corresponding preconditioner by \( \tilde{M}^{-1} \). We show that the condition number of the system with the preconditioner \( \tilde{M}^{-1} \) is of the form \( CC(k) \left( 1 + \log \frac{H}{h} \right)^2 \), where \( C(k) \) is a constant depending on the eigenvalues of the preconditioned coarse problem, the two parameters chosen for the Chebyshev

\*Courant Institute of Mathematical Sciences, New York University, 251 Mercer Street, New York, NY 10012, USA. E-mail: xuemin@cims.nyu.edu. This work was supported in part by the US Department of Energy under Contract DE-FG02-92ER25127.
iterations, and \( k \), the number of Chebyshev iterations. \( C(k) \) goes to 1 as \( k \) goes to \( \infty \), i.e., the condition number will decrease to that of the two-level case. Finally, some computational results are presented in Section 7.

2. The two-level BDDC method. We consider a second order scalar elliptic problem in a two dimensional region \( \Omega \): find \( u \in H^{1}_{0}(\Omega) \), such that

\[
\int_{\Omega} \rho \nabla u \cdot \nabla v = \int_{\Omega} fv \quad \forall v \in H^{1}_{0}(\Omega),
\]

where \( \rho(x) > 0 \) for all \( x \in \Omega \). We decompose \( \Omega \) into \( N \) nonoverlapping subdomains \( \Omega_i \) with diameters \( H_i \), \( i = 1, \ldots, N \), and set \( H = \max_i H_i \). We then introduce a triangulation of all the subdomains. Let \( \Gamma \) be the interface between subdomains and the set of interface nodes \( \Gamma_h \) is defined as \( \Gamma_h = (\cup_{i} \partial \Omega_{i,h}) \setminus \partial \Omega_h \), where \( \partial \Omega_{i,h} \) is the set of nodes on \( \partial \Omega_i \) and \( \partial \Omega_h \) is the set of nodes on \( \partial \Omega \).

Let \( W_{i} \) be the standard finite element space of continuous, piecewise linear functions on \( \Omega_i \). We assume that these functions vanish on \( \partial \Omega_i \). Each \( W_{i} \) can be decomposed into a subdomain interior part \( W_{I,i} \) and a subdomain interface part \( W_{I,s} \).

The subdomain interface part \( W_{I,i} \) will be further decomposed into a primal subspace \( W_{I,i} \) and a dual subspace \( W_{I,i} \), i.e., \( W_{i} = W_{I,i} \oplus W_{I,i} = W_{I} \oplus W_{I} \oplus W_{I} \).

We denote the associated product spaces by \( W_{I} := \prod_{i=1}^{N} W_{I,i} \), \( W_{I} := \prod_{i=1}^{N} W_{I,i} \), \( W_{I} := \prod_{i=1}^{N} W_{I,i} \), and \( W_{I} := \prod_{i=1}^{N} W_{I,i} \). Correspondingly, we have

\[
W = W_{I} \oplus W_{I} \oplus W_{I}.
\]

We will consider elements of a product space which are discontinuous across the interface. However, the finite element approximation of the elliptic problem is continuous across \( \Gamma \). We denote the corresponding subspace of \( W \) by \( \overline{W_{I}} \).

We further introduce an interface subspace \( \overline{W_{I}} \subset W_{I} \), for which certain primal constraints are enforced. Here, we only consider vertex constraints at the corners of each subdomain. The continuous primal subspace denoted by \( \overline{W_{I}} \) is then spanned by the continuous finite element basis functions of the vertex nodes. The space \( \overline{W_{I}} \) can be decomposed into \( \overline{W_{I}} = \overline{W_{I}} \oplus \overline{W_{I}} \).

The global problem is: find \( (u_{I}, u_{I}, u_{I}) \in (W_{I}, \overline{W_{I}}, \overline{W_{I}}) \), such that

\[
\begin{pmatrix}
A_{II} & A_{II}^T & A_{II}^T \\
A_{II} & A_{II} & A_{II}^T \\
A_{II} & A_{II} & A_{II}^T
\end{pmatrix}
\begin{pmatrix}
u_{I} \\
u_{I} \\
u_{I}
\end{pmatrix}
= \begin{pmatrix}
f_{I} \\
f_{I} \\
f_{I}
\end{pmatrix}.
\]

This problem is assembled from subdomain problems

\[
\begin{pmatrix}
A_{II} & A_{II}^T & A_{II}^T \\
A_{II} & A_{II} & A_{II}^T \\
A_{II} & A_{II} & A_{II}^T
\end{pmatrix}
\begin{pmatrix}
u_{I} \\
u_{I} \\
u_{I}
\end{pmatrix}
= \begin{pmatrix}
f_{I} \\
f_{I} \\
f_{I}
\end{pmatrix}.
\]

We also denote by \( F_{I}, \overline{F}_{I}, \) and \( \overline{F}_{I} \), the right hand side spaces corresponding to \( W_{I}, \overline{W_{I}}, \) and \( \overline{W_{I}} \), respectively.

In order to describe the BDDC algorithm, we need to introduce several restriction, extension, and scaling operators between different spaces. The restriction operator
$R^{(i)}_\Gamma$ maps a vector of the space $\tilde{W}_\Gamma$ to a vector of the subdomain subspace $W^{(i)}_\Gamma$. Each column of $R^{(i)}_\Gamma$ with a nonzero entry corresponds to an interface node, $x \in \partial \Omega_i \cap \Gamma_h$, shared by the subdomain $\Omega_i$ and its next neighbor subdomains. $R^{(i)}_\Gamma$ is similar to $R^{(i)}_\Pi$, and represents the restriction from $\tilde{W}_\Gamma$ to $W^{(i)}_\Gamma$. $R^{(i)}_\Delta : W_\Delta \to W^{(i)}_\Delta$ is the restriction matrix which extracts the subdomain part, in the space $W^{(i)}_\Delta$, of the functions in the space $W_\Delta$. $R^{(i)}_\Pi$ is the restriction operator from the space $\tilde{W}_\Pi$ to $W^{(i)}_\Pi$. Multiplying each element of $R^{(i)}_\Gamma$, $R^{(i)}_\Pi$, and $R^{(i)}_\Delta$, which corresponds to an $x \in \partial \Omega_i \cap \Gamma_h$, with $\delta_j^i(x)$ gives us $R^{(i)}_{D, \Gamma}$, $R^{(i)}_{D, \Pi}$, and $R^{(i)}_{D, \Delta}$, respectively. Here, we define $\delta_j^i(x)$ as follows: for $\gamma \in [1/2, \infty)$,

$$\delta_j^i(x) = \frac{\rho_j^i(x)}{\sum_{j \in N_x} \rho_j^i(x)}, \quad x \in \partial \Omega_i \cap \Gamma_h,$$

where $N_x$ is the set of indices $j$ of the subdomains such that $x \in \partial \Omega_j$ and $\rho_j(x)$ is the coefficient of (2.1) at $x$ in the subdomain $\Omega_j$. Furthermore, $R^{(i)}_\Delta$ and $R^{(i)}_\Pi$ are the restriction operators from the space $\tilde{W}_\Gamma$ onto its subspace $W_\Delta$ and $W_\Pi$ respectively. $R^{(i)}_\Gamma : \tilde{W}_\Gamma \to W_\Gamma$ and $\overline{R^{(i)}_\Gamma} : \tilde{W}_\Gamma \to \tilde{W}_\Gamma$ are the direct sums of $R^{(i)}_\Gamma$ and $\overline{R^{(i)}_\Gamma}$, respectively. $R^{(i)}_\Gamma : \tilde{W}_\Gamma \to \tilde{W}_\Gamma$ is the direct sum of $R^{(i)}_\Pi$ and the $\overline{R^{(i)}_\Gamma} R^{(i)}_\Gamma$. The scaled operators $R^{(i)}_{D, \Gamma}$ and $R^{(i)}_{D, \Delta}$ are the direct sums of $R^{(i)}_{D, \Gamma}$ and $R^{(i)}_{D, \Pi}$, respectively.

We also use the same restriction, extension, and scaled restriction operators for the right hand side spaces $F_\Gamma$, $\overline{F_\Gamma}$, and $F_\Gamma$.

We define an operator $\overline{S}_\Gamma : \tilde{W}_\Gamma \to \overline{F_\Gamma}$, which is of the form: given $u_\Gamma = u_\Pi \oplus u_\Delta \in \tilde{W}_\Pi \oplus W_\Delta = \tilde{W}_\Gamma$, find $\overline{S}_\Gamma u_\Gamma \in \overline{F_\Gamma}$ by eliminating the interior variables of the system:

$$\begin{equation} (2.2) \quad A \begin{pmatrix} u_\Gamma^{(i)} \\ u_\Delta^{(i)} \\ \vdots \\ u_\Pi^{(i)} \\ u_\Delta \\ \vdots \\ u_\Pi \\ 0 \end{pmatrix} = \begin{pmatrix} 0 \\ R^{(i)}_\Delta R^{(i)}_\Gamma \overline{S}_\Gamma u_\Gamma \\ \vdots \\ 0 \\ R^{(i)}_\Delta R^{(i)}_\Gamma \overline{S}_\Gamma u_\Gamma \\ \vdots \\ R^{(i)}_\Pi \overline{S}_\Gamma u_\Gamma \end{pmatrix}, \end{equation}$$

where $A$ is of the form

$$\begin{pmatrix} A^{(i)}_{\Pi \Pi} & A^{(i)T}_{\Pi \Delta} & \cdots & \cdots & \cdots & A^{(i)T}_{\Pi} & R^{(i)}_\Pi \\ A^{(i)}_{\Pi \Delta} & A^{(i)T}_{\Pi \Pi} & \cdots & \cdots & \cdots & A^{(i)T}_{\Pi \Delta} & R^{(i)}_{\Pi \Pi} \\ \vdots & \vdots & \ddots & \cdots & \cdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & A^{(i)}_{\Pi} & A^{(i)T}_{\Pi} & \cdots & \cdots \\ \cdots & \cdots & \cdots & A^{(i)}_{\Delta \Pi} & A^{(i)T}_{\Delta \Delta} & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots & \cdots & \vdots \\ R^{(i)T}_{\Pi} & A^{(i)}_{\Pi \Pi} & A^{(i)}_{\Pi \Delta} & \cdots & \cdots & A^{(i)T}_{\Pi} & \sum_{i=1}^{N} R^{(i)}_{\Pi} A^{(i)}_{\Pi \Pi} R^{(i)}_{\Pi} \end{pmatrix}.$$
The reduced interface problem can be written as: find $\mathbf{u}_\Gamma \in \mathbf{W}_\Gamma$ such that

$$
\tilde{R}_\Gamma^T \tilde{S}_\Gamma \tilde{R}_\Gamma \mathbf{u}_\Gamma = \mathbf{g}_\Gamma,
$$

where the operator $\tilde{S}_\Gamma$ is defined in (2.2), and

$$
\mathbf{g}_\Gamma = \sum_{i=1}^N R_\Gamma^{(i)T} \left\{ \left( \begin{array}{c} f_\Delta^{(i)} \\ f_\Gamma^{(i)} \end{array} \right) - \left( \begin{array}{cc} A_{\Delta}^{(i)} & A_{\Delta\Gamma}^{(i)} \\ A_{\Gamma\Delta}^{(i)} & A_{\Gamma}^{(i)} \end{array} \right) A_{\Gamma}^{(i)-1} f_\Gamma^{(i)} \right\}.
$$

The two-level BDDC method is of the form

$$
M^{-1} \tilde{R}_D^T \tilde{S}_\Gamma \tilde{R}_\Gamma \mathbf{u}_\Gamma = M^{-1} \mathbf{g}_\Gamma,
$$

where the preconditioner $M^{-1} = \tilde{R}_D^T \tilde{S}_\Gamma \tilde{R}_\Gamma$ has the following form:

$$
(2.3)
$$

$$
\tilde{R}_D^T \tilde{S}_\Gamma \tilde{R}_\Gamma \left\{ \sum_{i=1}^N \left( \begin{array}{c} 0 \\ R_\Delta^{(i)} \end{array} \right) \left( \begin{array}{cc} A_{\Delta}^{(i)} & A_{\Delta\Gamma}^{(i)} \\ A_{\Gamma\Delta}^{(i)} & A_{\Gamma}^{(i)} \end{array} \right)^{-1} \left( \begin{array}{c} 0 \\ R_\Gamma^{(i)} \end{array} \right) \right\} R_D + \Phi S_{\Pi}^{-1} \Phi^T \right\} \tilde{R}_D.
$$

Here $\Phi$ is the matrix given by the coarse level basis functions with minimal energy, and it is defined by

$$
\Phi = R_{\Pi}^T - R_{\Delta}^T \Delta \left\{ \sum_{i=1}^N \left( \begin{array}{c} 0 \\ R_\Delta^{(i)} \end{array} \right) \left( \begin{array}{cc} A_{\Delta}^{(i)} & A_{\Delta\Gamma}^{(i)} \\ A_{\Gamma\Delta}^{(i)} & A_{\Gamma}^{(i)} \end{array} \right)^{-1} \left( \begin{array}{c} 0 \\ A_{\Gamma}^{(i)} \end{array} \right) \right\} R_{\Pi}.
$$

The coarse level problem matrix $S_{\Pi}$ is determined by

$$
(2.4)
$$

$$
S_{\Pi} = \sum_{i=1}^N R_{\Pi}^{(i)T} \left\{ \left( \begin{array}{cc} A_{\Pi}^{(i)} & A_{\Pi\Delta}^{(i)} \\ A_{\Delta\Pi}^{(i)} & A_{\Delta}^{(i)} \end{array} \right) \left( \begin{array}{cc} A_{\Delta}^{(i)} & A_{\Delta\Gamma}^{(i)} \\ A_{\Gamma\Delta}^{(i)} & A_{\Gamma}^{(i)} \end{array} \right)^{-1} \left( \begin{array}{c} A_{\Pi}^{(i)T} \\ A_{\Pi\Gamma}^{(i)} \end{array} \right) \right\} R_{\Pi}^{(i)},
$$

which is obtained by assembling subdomain matrices; for more details, see [2, 7, 6].

We know that, for any $\mathbf{u}_\Gamma \in \mathbf{W}_\Gamma$,

$$
\mathbf{u}_\Gamma^T M \mathbf{u}_\Gamma \leq \mathbf{u}_\Gamma^T \tilde{R}_\Gamma^T \tilde{S}_\Gamma \tilde{R}_\Gamma \mathbf{u}_\Gamma \leq C (1 + \log(H/h))^2 \mathbf{u}_\Gamma^T M \mathbf{u}_\Gamma,
$$

see [9, 8, 6], provided that the coefficient $\rho(x)$ of (2.1) varies moderately in each subdomain. In developing these results, we also assume that each subdomain is a union of shape-regular coarse triangles and that the number of such triangles forming an individual subdomain is uniformly bounded. Moreover, we assume that the triangulation of each subdomain is quasi uniform.

3. A three-level BDDC method. In the three-level algorithms, we will not factor the coarse problem matrix $S_{\Pi}$ defined in (2.4) by a direct solver. Instead, we will try to solve the coarse problem approximately by using a similar idea as for the two-level preconditioners.

We decompose $\Omega$ into $N$ subregions $\Omega^j$ with diameters $H^j$, $j = 1, \ldots, N$. Each subregion $\Omega^j$ has $N_j$ subdomains $\Omega_i^j$ with diameters $H_i^j$. Let $H = \max_j H^j$ and $H = \max_{i,j} H_i^j$, for $j = 1, \ldots, N$ and $i = 1, \ldots, N_j$. We introduce the notation

$$
S_{\Pi}^{(j)} = \sum_{i=1}^{N_j} R_{\Pi}^{(i)T} \left\{ A_{\Pi}^{(i)} - \left( A_{\Pi\Pi}^{(i)} A_{\Pi\Delta}^{(i)} \right) \left( \begin{array}{cc} A_{\Delta}^{(i)} & A_{\Delta\Gamma}^{(i)} \\ A_{\Gamma\Delta}^{(i)} & A_{\Gamma}^{(i)} \end{array} \right)^{-1} \left( \begin{array}{c} A_{\Pi}^{(i)T} \\ A_{\Pi\Gamma}^{(i)} \end{array} \right) \right\} R_{\Pi}^{(i)},
$$

$$
\cdots
$$
and note that the coarse problem matrix $S_{II}$ can be assembled from the subregion matrices $S^{(j)}_{II}$.

Let $\hat{\Gamma}$ be the interface between the subregions. We denote the set of interior crosspoints, which contains the vertices of the subdomains in each subregion, by $I_H$, and the set of interface nodes, which contains the vertices of the subdomains on the boundary of the subregions, by $\hat{\Gamma}_H$. We note that $\hat{\Gamma} \subseteq \Gamma$.

We denote the vector space corresponding to the nodes of the subregion $\Omega^i$, which are the vertices of the subdomains in $\Omega^i$, by $W^{(i)}_c$. Each $W^{(i)}_c$ can be decomposed into a subregion interior part $W \cap_i$ and a subregion interface part $W^{(i)}_{c,\hat{F}}$. The subregion interface part $W^{(i)}_{c,\hat{F}}$ can be further decomposed into a primal subspace $W^{(i)}_{c,\hat{F}}$ and a dual subspace $W^{(i)}_{c,\Delta}$, i.e., $W^{(i)}_c = W^{(i)}_{c,\hat{F}} \oplus W^{(i)}_{c,\hat{F}} = W^{(i)}_{c,\hat{F}} \oplus W^{(i)}_{c,\hat{F}} \oplus W^{(i)}_{c,\Delta}$. We denote the associated product spaces by $W_c := \prod_{i=1}^N W^{(i)}_c$, $W^{(i)}_{c,\hat{F}} := \prod_{i=1}^N W^{(i)}_{c,\hat{F}}$, $W^{(i)}_{c,\Delta} := \prod_{i=1}^N W^{(i)}_{c,\Delta}$, $W^{(i)}_{c,\hat{F}} := \prod_{i=1}^N W^{(i)}_{c,\hat{F}}$, and $W^{(i)}_c := \prod_{i=1}^N W^{(i)}_c$. Correspondingly, we have $W^{(i)}_c = W^{(i)}_{c,\hat{F}} \oplus W^{(i)}_{c,\hat{F}} \oplus W^{(i)}_{c,\Delta}$.

We denote by $\tilde{W}_c$ the subspace of $W_c$ of functions that are continuous across $\hat{\Gamma}$. We introduce an interface subspace $W^{(i)}_{c,\hat{F}} \subset W_{c,\hat{F}}$, for which the primal constraints are enforced. Here, we only consider vertex constraints. The continuous primal subspace is denoted by $\tilde{W}_{c,\hat{F}}$. The space $W^{(i)}_{c,\hat{F}}$ can be decomposed into $W^{(i)}_{c,\hat{F}} = W^{(i)}_{c,\hat{F}} \oplus W^{(i)}_{c,\Delta}$.

We also denote by $F^{(i)}_{c,\hat{F}}, F^{(i)}_{c,\hat{F}}$, and $\tilde{F}^{(i)}_{c,\hat{F}}$, the right hand side spaces corresponding to $W^{(i)}_{c,\Gamma}, \tilde{W}^{(i)}_{c,\Gamma},$ and $\tilde{W}^{(i)}_{c,\Gamma}$, respectively.

In the three-level BDDC algorithm, we need to introduce several restriction, extension, and scaling operators between different subregion spaces. The restriction operator $\hat{R}^{(i)}_{\hat{F}}$ maps a vector of the space $\tilde{W}^{(i)}_{c,\hat{F}}$ to a vector of the subdomain subspace $W^{(i)}_c$. Each column of $\hat{R}^{(i)}_{\hat{F}}$ with a nonzero entry corresponds to an interface node, $x \in \partial \Omega^i \cap \partial \Omega^j$, shared by the subregion $\Omega^i$ and certain neighboring subregions. $\hat{R}^{(i)}_{\hat{F}}$ is similar to $\hat{R}^{(i)}_{\hat{F}}$, and represents the restriction from $\tilde{W}^{(i)}_{c,\hat{F}}$ to $W^{(i)}_{c,\hat{F}}$. $\hat{R}^{(i)}_{\hat{F}}$ is the restriction matrix which extracts the subregion part, in the space $W^{(i)}_{c,\Delta}$, of the functions in the space $W^{(i)}_{c,\hat{F}}$. Multiplying each such element of $\hat{R}^{(i)}_{\hat{F}}$, $\hat{R}^{(i)}_{\hat{F}}$, and $\hat{R}^{(i)}_{\hat{F}}$ with $\gamma(x)$ gives us $\hat{R}^{(i)}_{\hat{F},\hat{F}}$, $\hat{R}^{(i)}_{\hat{F},\hat{F}}$, and $\hat{R}^{(i)}_{\hat{F},\hat{F}}$, respectively. Here, we define $\gamma(x)$ as follows: for $\gamma \in [1/2, \infty)$,

$$\gamma(x) = \sum_{j \in N_x, \rho_j(x)}^{\rho_j(x)}$$

where $N_x$ is the set of indices $j$ of the subdomains such that $x \in \partial \Omega^j \cap \partial \Omega^j$ and $\rho_j(x)$ is the coefficient of (2.1) at $x$ in the subregion $\Omega^j$. (In our theory, we assume the $\rho_i$ are constant in the subregions.) Furthermore, $\hat{R}^{(i)}_{\hat{F}}$ and $\hat{R}^{(i)}_{\hat{F}^{(i)}}$ are the restriction operators from the space $\tilde{W}^{(i)}_{c,\hat{F}}$ onto its subspace $W^{(i)}_{c,\hat{F}}$ and $W^{(i)}_{c,\hat{F}}$ respectively. $\hat{R}^{(i)}_{\hat{F}} : \tilde{W}^{(i)}_{c,\hat{F}} \rightarrow W^{(i)}_{c,\hat{F}}$ and $\hat{R}^{(i)}_{\hat{F}} : \tilde{W}^{(i)}_{c,\hat{F}} \rightarrow W^{(i)}_{c,\hat{F}}$ are the direct sum of $\hat{R}^{(i)}_{\hat{F}}$ and $\hat{R}^{(i)}_{\hat{F}}$, respectively. $\hat{R}^{(i)}_{\hat{F}} : \tilde{W}^{(i)}_{c,\Gamma} \rightarrow W^{(i)}_{c,\hat{F}}$ is the direct sum of $\hat{R}^{(i)}_{\hat{F}}$ and the $\hat{R}^{(i)}_{\hat{F}} \hat{R}^{(i)}_{\hat{F}}$, respectively. $\hat{R}^{(i)}_{\hat{F}}$ and $\hat{R}^{(i)}_{\hat{F}}$ are the direct sums of $\hat{R}^{(i)}_{\hat{F}}$ and $\hat{R}^{(i)}_{\hat{F}}$. $\hat{R}^{(i)}_{\hat{F}}$ is the direct sum of
Let $\tilde{S}_i^{-1}$ be an approximation of $S_{ii}^{-1}$ and defined as follows: given any right hand side $\Psi \in \hat{\mathbf{F}}_{c,\hat{\Gamma}}$, let $y = S_{ii}^{-1}\Psi$ and $\bar{y} = \tilde{S}_i^{-1}\Psi$. Here $\Psi = \left(\Psi_i^{(1)}, \ldots, \Psi_i^{(N)}\right)^T$, $y = \left(y_i^{(1)}, \ldots, y_i^{(N)}\right)^T$, and $\bar{y} = \left(\bar{y}_i^{(1)}, \ldots, \bar{y}_i^{(N)}\right)^T$.

To solve $S_{ii}\mathbf{y} = \Psi$ by block factorization, we write

$$
\begin{pmatrix}
S_{ii}^{(1)} & \cdots & \cdots & \cdots & S_{ii}^{(1)}
\vdots & \ddots & \ddots & \ddots & \vdots
\vdots & \ddots & \ddots & \cdots & \vdots
\end{pmatrix}
\begin{pmatrix}
S_{ii}^{(1)}
\vdots
\vdots
\vdots
\end{pmatrix}
\mathbf{y}_{1}^{(1)}
= 
\begin{pmatrix}
\Psi_{1}^{(1)}
\vdots
\vdots
\end{pmatrix}.
$$

We have

$$(3.1) \quad y_{1}^{(i)} = S_{ii}^{(i)^{-1}} \left( \Psi_{1}^{(i)} - S_{ii}^{(i)^{T}} \tilde{R}_{i}^{(i)} \bar{y}_{i}^{(i)} \right),$$

and

$$(3.2) \quad \left( \sum_{i=1}^{N} \tilde{R}_{i}^{(i)^{T}} \left( S_{ii}^{(i)} - S_{ii}^{(i)^{T}} S_{ii}^{(i)^{-1}} S_{ii}^{(i)^{T}} \right) \tilde{R}_{i}^{(i)} \right) \bar{y}_{i} = \Psi_{i} - \sum_{i=1}^{N} \tilde{R}_{i}^{(i)^{T}} S_{ii}^{(i)^{-1}} S_{ii}^{(i)^{T}} \Psi_{i}.$$

Let $T^{(i)} = S_{ii}^{(i)} - S_{ii}^{(i)^{T}} S_{ii}^{(i)^{-1}} S_{ii}^{(i)^{T}}$ and $T = diag(T^{(1)}, \ldots, T^{(N)})$. We then introduce a partial assembled Schur complement of $S_{ii}$, $\tilde{T} : \hat{\mathbf{W}}_{c,\hat{\Gamma}} \to \hat{\mathbf{F}}_{c,\hat{\Gamma}}$ by

$$(3.2) \quad \tilde{T} = \tilde{R}_{\hat{\Gamma}}^{T} \tilde{T} \tilde{R}_{\hat{\Gamma}},$$

and define $h_{\hat{\Gamma}} \in \hat{\mathbf{F}}_{c,\hat{\Gamma}}$, by

$$(3.3) \quad h_{\hat{\Gamma}} = \Psi_{\hat{\Gamma}} - \sum_{i=1}^{N} \tilde{R}_{\hat{\Gamma}}^{(i)^{T}} S_{ii}^{(i)^{-1}} S_{ii}^{(i)^{T}} \Psi_{\hat{\Gamma}}.$$

The reduced subregion interface problem can be written as: find $y_{\hat{\Gamma}} \in \hat{\mathbf{W}}_{c,\hat{\Gamma}}$, such that

$$(3.4) \quad \tilde{T}^{T} \tilde{R}_{\hat{\Gamma}}^{T} \tilde{R}_{\hat{\Gamma}} y_{\hat{\Gamma}} = h_{\hat{\Gamma}}.$$

When using the three-level preconditioner $\tilde{M}^{-1}$, we do not solve (3.4) exactly. Instead, we replace $y_{\hat{\Gamma}}$ by

$$
(3.5) \quad \tilde{y}_{\hat{\Gamma}} = \tilde{R}_{\hat{\Gamma}} \tilde{R}_{\hat{\Gamma}}^{-1} \tilde{R}_{\hat{\Gamma}}^{T} h_{\hat{\Gamma}}.
$$

Here $\tilde{T}^{-1}$ has the same structure as $\tilde{S}_{i}^{-1}$ defined in (2.3), except that the components defined in subdomains are replaced by the components defined in subregions.
We will maintain the same relation between \( y^{(i)}_I \) and \( y^{(i)}_F \), i.e.,

\[ y^{(i)}_I = S^{(i)}_{H_I}^{-1} \left( \Psi^{(i)}_I - S^{(i)}_{H_I}^{-1} R^{(i)}_F y^{(i)}_F \right) \]

(3.6)

We define our three-level preconditioner \( \widetilde{M}^{-1} \) by

\[ \widetilde{R}^D_{D,I} \left\{ R^D_{I,D} \left( \sum_{i=1}^{N} \binom{0}{A_{i}} A^{(i)}_{H} A^{(i)}_{\Delta} \right)^{-1} \left( \begin{array}{c} 0 \\ R^{(i)}_{\Delta} \end{array} \right) \right\} \]

(3.7)

cf. (2.3). We note that we only replace \( S^{-1}_H \) by \( \tilde{S}^{-1}_H \) in \( M^{-1} \) to obtain \( \widetilde{M}^{-1} \). The condition number estimate with the new preconditioner can be obtained if we can get a good estimate for \( \tilde{S}^{-1}_H \) in comparison with \( S^{-1}_H \).

With (3.4), we have to solve a global subregion interface problem for \( y_F \). However, with (3.5), we only need to solve several subregion local problems and one coarse problem on the subregion level to obtain \( y_F \). The subregion coarse and local problems are much smaller than the global subregion interface problem and can be solved exactly, independently, and in parallel; see also Section 8.

4. Some auxiliary results. In this section, we will collect a number of results which are needed in our theory. In order to avoid a proliferation of constants, we will use the notation \( A \approx B \). This means that there are two constants \( c \) and \( C \), independent of any parameters, such that \( c A \leq B \leq C A \), where \( C < \infty \) and \( c > 0 \). For the definition of discrete harmonic functions, see [10, Section 4.4].

**Lemma 4.1.** Let \( \mathcal{D} \) be a square with vertices \( A = (0,0), B = (H,0), C = (H,H), \) and \( D = (0,H) \), with a quasi-uniform triangulation of mesh size \( h \). Then, there exists a discrete harmonic function \( v \) defined on \( \mathcal{D} \) such that \( \|v\|_L^2(\mathcal{D}) = v(A) \approx 1 + \log \frac{L}{h} \), \( v(B) = v(C) = v(D) = 0 \) and \( |v|^2_2(\mathcal{H}(\mathcal{D})) \approx 1 + \log \frac{L}{h} \).

**Proof:** This lemma follows from a result by Brenner and Sung [1, Lemma 4.2]: let \( N \) be an integer and \( G_N \) be the function defined on \((0,1)\) by

\[ G_N(x) = \sum_{n=1}^{N} \left( \frac{1}{4n-3} \sin((4n-3)\pi x) \right). \]

\( G_N \) is symmetric with respect to the midpoint of \((0,1)\), where it attains its maximum in absolute value. Moreover, we have:

\[ |G_N|_{H^{1/2}_{00}(0,1)} \approx 1 + \log N, \]

and

\[ \|G_N\|_{L^\infty(0,1)} = G_N(1/2) \approx 1 + \log N; \]

see [1, Lemma 3.2].

Let \([-H,0] \) and \([0,H] \) have the mesh inherited from the quasi-uniform mesh on \( DA \) and \( AB \) respectively and let \( g_h(x) \) be the nodal interpolation of \( G_N(\frac{x+H}{2H}) \). Then we have

\[ |g_h|^2_{H^{1/2}_{00}(-H,H)} \approx 1 + \log \frac{H}{h} \text{ and } \|g_h\|_{L^\infty(0,1)} \approx 1 + \log \frac{H}{h}; \]
see [1, Corollary 3.6]. We point out that in [1, Corollary 3.6], a uniform mesh is used. But in the proof of the bound for \(| \cdot |_{H^{1/2}_0(-H,H)}\), we only need an interpolation error estimate theorem and the fact that \(H^{1/2}_0(-H,H)\) is the interpolation space halfway between \(L^2(-H,H)\) and \(H^1_0(-H,H)\). Therefore the result is also valid for a quasi-uniform mesh.

We can define \(v\) as 0 on the line segments \(CD\) and \(CB\) and by
\[
v(x,0) = g_h(x), \quad \text{for} \quad 0 \leq x \leq H,
\]
and
\[
v(0,y) = g_h(-y), \quad \text{for} \quad 0 \leq y \leq H.
\]

Since \(v\) is a discrete harmonic function in \(D\), we have,
\[
|v|^2_{H^1(D)} = |v|^2_{H^{1/2}(\partial D)} \approx |g_h|^2_{H^{1/2}_0(-H,H)} \approx 1 + \log \frac{H}{h}.
\]

**Remark:** In Lemma 4.1, we have constructed the function \(v\) for the square \(D\). By using similar ideas, we can easily construct a function \(v\) for other shape-regular polygons which satisfies the same properties.

**Lemma 4.2.** Let \(V_i^h\) be the standard continuous piecewise linear finite element function space for a subregion \(\Omega^i\) with a quasi-uniform coarse mesh with mesh size \(H\). And let \(V_{i,j}^h\), \(j = 1, \ldots, N_i\), be the space for a subdomain \(\Omega^i_j\) with a quasi-uniform fine mesh with mesh size \(h\). Moreover, each subdomain is a union of coarse triangles with vertices on the boundary of the subdomain. Given \(u \in V_i^h\), let \(\hat{u}\) interpolate \(u\) at each coarse node and be the discrete \(V_{i,j}^h\)-harmonic extension in each subdomain \(\Omega^i_j\), constrained only at the vertices of \(\Omega^i_j\), \(j = 1, \ldots, N_i\). Then, there exist two positive constants \(C_1\) and \(C_2\), which are independent of \(H, H, h, \) and \(h\), such that
\[
C_1(1 + \log \frac{H}{h}) \left( \sum_{j=1}^{N_i} |\hat{u}|^2_{H^1(\Omega^i_j)} \right) \leq |u|^2_{H^1(\Omega^i)} \leq C_2(1 + \log \frac{H}{h}) \left( \sum_{j=1}^{N_i} |\hat{u}|^2_{H^1(\Omega^i_j)} \right).
\]

**Proof:** Without loss of generality, we assume that the subdomains are quadrilaterals. Denote the vertices of the subdomain \(\Omega^i_j\) by \(a_j, b_j, c_j,\) and \(d_j\), and denote the nodal values of \(u\) at these four crosspoints by \(u(a_j), u(b_j), u(c_j),\) and \(u(d_j)\), respectively. Since \(u\) is a piecewise linear function, we have,
\[
|u|^2_{H^1(\Omega^i)} = \sum_{j=1}^{N_i} |u|^2_{H^1(\Omega^i_j)},
\]
and
\[
|u|^2_{H^1(\Omega^i_j)} = |u - u(a_j)|^2_{H^1(\Omega^i_j)} \approx C \left( \sum_{m=b,c,d} ((u(m_j) - u(a_j))^2) \right).
\]

According to Lemma 4.1, we can construct three discrete harmonic functions \(\phi_b, \phi_c,\) and \(\phi_d\) on \(\Omega^i_j\) such that
\[
\phi_b(b_j) = (u(b_j) - u(a_j)) \left( 1 + \log \frac{H}{h} \right), \quad \phi_b(a_j) = \phi_b(c_j) = \phi_b(d_j) = 0,
\]
\[
\phi_c(c_j) = (u(c_j) - u(a_j))(1 + \log \frac{H}{h}), \quad \phi_c(a_j) = \phi_c(b_j) = \phi_c(d_j) = 0,
\]
\[
\phi_d(d_j) = (u(d_j) - u(a_j))(1 + \log \frac{H}{h}), \quad \phi_d(a_j) = \phi_d(b_j) = \phi_d(c_j) = 0,
\]

and with
\[
|\phi_m|_{H^1(\Omega_j')}^2 \approx (u(m_j) - u(a_j))^2 (1 + \log \frac{H}{h}), \quad m = b, c, d.
\]
Let \( v_j = \frac{1}{1 + \log \frac{H}{h}} (\phi_b + \phi_c + \phi_d) + u(a_j) \); we then have \( v_j(m_j) = u(m_j) \), \( m = a, b, c, d \), and
\[
|v_j|_{H^1(\Omega_j')}^2 = \left( \frac{1}{1 + \log \frac{H}{h}} \right)^2 |\phi_b + \phi_c + \phi_d|_{H^1(\Omega_j')}^2
\leq 3 \left( \frac{1}{1 + \log \frac{H}{h}} \right)^2 \sum_{m=b,c,d} |\phi_m|_{H^1(\Omega_j')}^2
\leq 3 \left( \frac{1}{1 + \log \frac{H}{h}} \right)^2 \left( 1 + \log \frac{H}{h} \right) \sum_{m=b,c,d} (u(m_j) - u(a_j))^2
\leq \frac{1}{C_1(1 + \log \frac{H}{h})} |u|_{H^1(\Omega_j')}^2.
\]

Here, we have used (4.1) and (4.2) for the last two inequalities.

By the definition of \( \hat{u} \), we have,
\[
|\hat{u}|_{H^1(\Omega_j')}^2 \leq |v|_{H^1(\Omega_j')}^2 \leq \frac{1}{C_1(1 + \log \frac{H}{h})} |u|_{H^1(\Omega_j')}^2.
\]

Summing over all the subdomains in the subregion \( \Omega^i \), we have,
\[
C_1 \left( 1 + \log \frac{H}{h} \right) \left( \sum_{j=1}^{N_i} |\hat{u}|_{H^1(\Omega_j')}^2 \right) \leq \sum_{j=1}^{N_i} |u|_{H^1(\Omega_j')}^2 = |u|_{H^1(\Omega^i)}^2.
\]

This proves the first inequality.

We prove the second inequality as follows:
\[
|u|_{H^1(\Omega^i)}^2 = \sum_{j=1}^{N_i} |u|_{H^1(\Omega_j')}^2 = \sum_{j=1}^{N_i} (u - u(a_j))^2_{H^1(\Omega_j')}
\leq C_2 \left( \sum_{j=1}^{N_i} \max_{m=b,c,d} (u(m_j) - u(a_j))^2 \right) \leq C_2 \left( \sum_{j=1}^{N_i} ||\hat{u} - u(a_j)||_{L^\infty(\Omega_j')}^2 \right)
\leq C_2 \left( 1 + \log \frac{H}{h} \right) \left( \sum_{j=1}^{N_i} |\hat{u}|_{H^1(\Omega_j')}^2 \right).
\]

Here, we have used a standard finite element Sobolev inequality, see [10, Lemma 4.15].
We next list several results for the two-level BDDC methods. To be fully rigorous, we assume that each subregion is a union of shape-regular coarse triangles and that the number of such triangles forming an individual subregion is uniformly bounded. Thus, there is a quasi-uniform coarse triangulation of each subregion. Similarly, each subdomain is a union of shape-regular coarse triangles with the vertices on the boundary of the subdomain. Moreover the fine triangulation of each subdomain is quasi uniform. We can then get uniform constants $C_1$ and $C_2$ in Lemma 4.2, which hold for all the subregions.

We define the interface averaging operator $\hat{E}_D$ on $\tilde{W}_{c, \Gamma}$ as $\hat{E}_D = \hat{F}_D \hat{R}_D \hat{R}_D^T$, which computes the averages across the subregion interface $\hat{\Gamma}$ and then distributes the averages to the boundary points of the subregions.

The interface averaging operator $\hat{E}_D$ satisfies the following bound:

**Lemma 4.3.** For the two-level BDDC, we have

$$|\hat{E}_D \mathbf{u}_F|^2_{S_F} \leq C \left(1 + \log \frac{\hat{H}}{H}\right)^2 |\mathbf{u}_F|^2_{S_F}, \quad \forall \mathbf{u}_F \in \tilde{W}_F,$$

where $S_F$ and $\tilde{W}_F$, which corresponds to a mesh with size $H$, are analogous to $S_T$ and $\tilde{W}_T$, which corresponds to a mesh with size $h$, respectively.

**Proof:** We can use a result by Mandel and Tezaur for the FETI-DP algorithm in [9] to estimate $(\hat{E}_D - I) \mathbf{u}_F$. We find:

$$|\hat{E}_D \mathbf{u}_F|^2_{S_F} = |(\hat{E}_D - I) \mathbf{u}_F + \mathbf{u}_F|^2_{S_F} \leq C \left(1 + \log \frac{\hat{H}}{H}\right)^2 |\mathbf{u}_F|^2_{S_F}.$$

We also have a result for the condition number of the two-level BDDC, see [7].

**Lemma 4.4.** The condition number for the operator with the two-level preconditioner $M^{-1}$ is bounded by $C (1 + \log \frac{\hat{H}}{H})^2$.

In addition, we have; recall that $\hat{T}$ is defined in (3.2).

**Lemma 4.5.**

$$|\hat{E}_D \mathbf{w}_F|^2_{\hat{T}} \leq C \left(1 + \log \frac{\hat{H}}{H}\right)^2 |\mathbf{w}_F|^2_{\hat{T}},$$

for any $\mathbf{w}_F \in \tilde{W}_{c, \hat{T}}$, where $C$ is a positive constant independent of $\hat{H}$, $H$, and $h$.

**Proof:** Denote by $\mathcal{H}_i$ the discrete harmonic extension in the subregion $\Omega_i$ with respect to $S_i$, given by the values on the boundary of $\Omega_i$, i.e., $\mathcal{H}_i(w) \in W^{(i)}_c$ satisfies:

$$|\mathcal{H}_i(w)|_{S_i} = \min_{v \in W^{(i)}, v = w \text{ on } \partial \Omega_i} \rho_i[v]_{S_i}, \quad w \in W^{(i)}_c.$$

For $w \in W^{(i)}_c$, let $\hat{\mathcal{H}}_i(w) \in W^{(i)}_c$ satisfy:

$$|I_H(\hat{\mathcal{H}}_i(w))|_{H^1(\Omega_i)} = \min_{v \in W^{(i)}_c, v = w \text{ on } \partial \Omega_i} |I_H(v)|_{H^1(\Omega_i)}.$$

The interface averaging operator $\hat{E}_D$ satisfies:

$$|\hat{E}_D \mathbf{w}_F|^2_{\hat{T}} \leq C \left(1 + \log \frac{\hat{H}}{H}\right)^2 |\mathbf{w}_F|^2_{\hat{T}},$$

for any $\mathbf{w}_F \in \tilde{W}_{c, \hat{T}}$, where $C$ is a positive constant independent of $\hat{H}$, $H$, and $h$.
where $I_H(\cdot)$ is the nodal interpolation onto $V_i^H$, a space defined in Lemma 4.2.

Denote by $\hat{\mathcal{H}}_i^j$ the discrete harmonic extension in each subdomain $\Omega_i^j$, with respect to the fine mesh with mesh size $h$, given the crosspoint nodal values, where $i = 1, \cdots, N$, and $j = 1, \cdots, N_i$.

We have

\[
|\hat{E}_Dw_F|^2 \leq \sum_{i=1}^N |\mathcal{H}(\hat{E}_Dw_F)|_{\mathcal{S}_H}^2 \leq \sum_{i=1}^N |\mathcal{H}(\hat{E}_Dw_F)|_{\mathcal{S}_H}^2
\]

\[
= \sum_{i=1}^N \rho_i \left( \sum_{j=1}^{N_i} |\mathcal{H}_j^i| (\mathcal{H}(\hat{E}_Dw_F)) |_{\mathcal{S}_H}^2 \right).
\]

Here we have used the definitions of $\mathcal{H}$, $\hat{\mathcal{H}}$, $\mathcal{H}$, and $S^{(i)}_H$.

By Lemma 4.2,

\[
|\hat{E}_Dw_F|^2 \leq \sum_{i=1}^N \rho_i \left( \sum_{j=1}^{N_i} |\mathcal{H}_j^i| (\mathcal{H}(\hat{E}_Dw_F)) |_{\mathcal{S}_H}^2 \right)
\]

\[
\leq \frac{1}{C_1(1 + \log \frac{H}{h})} \sum_{i=1}^N \rho_i \left( \mathcal{H}(\hat{E}_Dw_F) |_{\mathcal{S}_H}^2 \right)
\]

\[
= \frac{1}{C_1(1 + \log \frac{H}{h})} |\hat{E}_Dw_F|^2_{\mathcal{S}_F}.
\]

Using Lemma 4.3, we obtain

\[
|\hat{E}_Dw_F|^2 \leq \frac{1}{C_1(1 + \log \frac{H}{h})} |\hat{E}_Dw_F|^2_{\mathcal{S}_F}
\]

\[
\leq \frac{C}{C_1(1 + \log \frac{H}{h})} \left( 1 + \log \frac{\hat{H}}{H} \right)^2 |w_F|^2_{\mathcal{S}_F}
\]

\[
= \frac{C}{C_1(1 + \log \frac{H}{h})} \left( 1 + \log \frac{\hat{H}}{H} \right)^2 \left( \sum_{i=1}^N \rho_i |\mathcal{H}(\hat{E}_Dw_F)|_{\mathcal{S}_H}^2 \right)
\]

\[
\leq \frac{C}{C_1(1 + \log \frac{H}{h})} \left( 1 + \log \frac{\hat{H}}{H} \right)^2 \left( \sum_{i=1}^N \rho_i |\mathcal{H}(\hat{E}_Dw_F)|_{\mathcal{S}_H}^2 \right) \cdot
\]

Here we have used the definition of $\mathcal{H}$ and $\hat{\mathcal{H}}$ again.
By Lemma 4.2 and the definition of $\mathcal{H}$, we have

$$
|\hat{E}_D w_F|^2 \leq \frac{C}{C_1(1 + \log \frac{\hat{H}}{H})} \left(1 + \log \frac{\hat{H}}{H}\right)^2 \left(\sum_{i=1}^N \rho_i |I_H \left(\mathcal{H} \left(\mathcal{R}_1^{(i)} w_F \right)\right)|^2_{H(i)}\right)
$$

$$
\leq \frac{C}{C_1(1 + \log \frac{\hat{H}}{H})} \left(1 + \log \frac{\hat{H}}{H}\right)^2 C_2 \left(1 + \log \frac{H}{h}\right) \cdot \\
\left(\sum_{i=1}^N \rho_i \sum_{j=1}^{N_i} \left|\mathcal{H} \left(\mathcal{R}_1^{(i)} w_F \right)\right|^2_{H(i)}\right)
$$

$$
= \frac{CC_2}{C_1} \left(1 + \log \frac{\hat{H}}{H}\right)^2 \left(\sum_{i=1}^N \left|\mathcal{H} \left(\mathcal{R}_1^{(i)} w_F \right)\right|^2_{H(i)}\right)
$$

$$
= \frac{CC_2}{C_1} \left(1 + \log \frac{\hat{H}}{H}\right)^2 \left|w_F\right|^2_{T}.
$$

**Lemma 4.6.** Given any $u_F \in \hat{W}_F$, let $\Psi = \Phi^T \hat{R}_{D,F} u_F$. We have,

$$
\Psi^T S_{\Pi}^{-1} \Psi \leq \Psi^T \hat{S}_{\Pi}^{-1} \Psi \leq C \left(1 + \log \frac{\hat{H}}{H}\right)^2 \Psi^T S_{\Pi}^{-1} \Psi.
$$

**Proof:** Using (3.1), (3.3), and (3.4), we have

(4.4)

$$
\Psi^T S_{\Pi}^{-1} \Psi = \sum_{i=1}^N \Psi_i^{(i)^T} y_i^{(i)} + \Psi_i^{T} y_i^{(i)}
$$

$$
= \sum_{i=1}^N \Psi_i^{(i)^T} \left(S_{\Pi i}^{-1} \Psi_i^{(i)} - S_{\Pi i}^{(i)^T} \hat{R}_i y_i^{(i)}\right) + \left(h_F + \sum_{i=1}^N \hat{R}_i^{(i)^T} S_{\Pi i}^{-1} \Psi_i^{(i)}\right)^T y_i^{(i)}
$$

$$
= \sum_{i=1}^N \Psi_i^{(i)^T} S_{\Pi i}^{-1} \Psi_i^{(i)} + h_F^T y_i^{(i)} = \sum_{i=1}^N \Psi_i^{(i)^T} S_{\Pi i}^{-1} \Psi_i^{(i)} + h_F^T \left(\hat{R}_i y_i^{(i)}\right)^{-1} h_F.
$$

Using (3.6), (3.3), and (3.5), we also have

$$
\Psi^T \hat{S}_{\Pi}^{-1} \Psi = \sum_{i=1}^N \Psi_i^{(i)^T} \hat{y}_i^{(i)} + \Psi_i^{T} \hat{y}_i^{(i)}
$$

$$
= \sum_{i=1}^N \Psi_i^{(i)^T} \left(S_{\Pi i}^{-1} \Psi_i^{(i)} - S_{\Pi i}^{(i)^T} \hat{R}_i \hat{y}_i^{(i)}\right) + \left(h_F + \sum_{i=1}^N \hat{R}_i^{(i)^T} S_{\Pi i}^{-1} \Psi_i^{(i)}\right)^T \hat{y}_i^{(i)}
$$

$$
= \sum_{i=1}^N \Psi_i^{(i)^T} S_{\Pi i}^{-1} \Psi_i^{(i)} + h_F^T \hat{y}_i^{(i)} = \sum_{i=1}^N \Psi_i^{(i)^T} S_{\Pi i}^{-1} \Psi_i^{(i)} + h_F^T \left(\hat{R}_i \hat{y}_i^{(i)}\right)^{-1} h_F.
$$

We only need to compare $h_F^T \left(\hat{R}_i \hat{y}_i^{(i)}\right)^{-1} h_F$ and $h_F^T \left(\hat{R}_{D,F} \hat{y}_i^{(i)}\right) h_F$ for any
We obtain that
\[
\begin{align*}
\mathbf{h}_f & \leq \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right) \\
& \leq C \left( 1 + \log \frac{H}{\bar{H}} \right)^2 \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right).
\end{align*}
\]

Let
\[
\mathbf{w}_f = \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \in \widehat{W}_{c,f}, \quad \mathbf{v}_f = \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \mathbf{h}_f \in \widehat{W}_{c,f}.
\]

Noting the fact that \( \widehat{R}_{D,f} \widehat{R}_{D,f} = \mathbf{I} \) and using (4.5), we have,
\[
\begin{align*}
\mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f &= \mathbf{h}_f^T \mathbf{w}_f = \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{w}_f \\
&= \mathbf{h}_f^T \widehat{R}_{D,f}^{-1} \mathbf{w}_f = \left( \mathbf{v}_f^T \widehat{R}_{D,f} \mathbf{v}_f \right)^{1/2} \left( \frac{1}{2} \right) \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right)^{1/2}.
\end{align*}
\]

We obtain that
\[
\mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \leq \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f.
\]

On the other hand,
\[
\begin{align*}
\mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f &= \mathbf{w}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right) \\
&\leq \mathbf{w}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right)^{1/2} \left( \frac{1}{2} \right) \\
&= \left( \frac{1}{2} \right) \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right)^{1/2} \left( \frac{1}{2} \right) \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right)^{1/2} \\
&\leq C \left( 1 + \log \frac{H}{\bar{H}} \right)^2 \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right)^{1/2} \left( \frac{1}{2} \right) \left( \mathbf{h}_f^T \left( \widehat{R}_{D,f}^{-1} \widehat{R}_{D,f} \right)^{-1} \mathbf{h}_f \right)^{1/2}.
\end{align*}
\]
where we use Lemma 4.5 for the penultimate inequality.

Finally we obtain that

\[
\mathbf{h}_\Gamma^T \left( \frac{\tilde{\mathbf{h}}_\Gamma^T \hat{\mathbf{R}}_{\tilde{\mathbf{D}},\Gamma}^{-1} \tilde{\mathbf{h}}_\Gamma}{\left( 1 + \log \frac{\hat{H}}{H} \right)^2} \right)^{-1} \mathbf{h}_\Gamma \leq C \left( 1 + \log \frac{\hat{H}}{H} \right)^2 \left( \mathbf{h}_\Gamma^T \left( \frac{\tilde{\mathbf{h}}_\Gamma^T \hat{\mathbf{R}}_{\tilde{\mathbf{D}},\Gamma}}{\left( 1 + \log \frac{\hat{H}}{H} \right)^2} \right)^{-1} \mathbf{h}_\Gamma \right).
\]

5. Condition number estimate for the new preconditioner. In order to estimate the condition number for the system with the new preconditioner \( \hat{M}^{-1} \), we compare it to that with the preconditioner \( M^{-1} \).

LEMA 5.1. Given any \( \mathbf{u}_\Gamma \in \tilde{\mathbf{W}}_\Gamma \),

\[
(5.1) \quad \mathbf{u}_\Gamma^T \hat{M}^{-1} \mathbf{u}_\Gamma \leq \mathbf{u}_\Gamma^T \hat{M}^{-1} \mathbf{u}_\Gamma \leq C \left( 1 + \log \frac{\hat{H}}{H} \right)^2 \mathbf{u}_\Gamma^T \hat{M}^{-1} \mathbf{u}_\Gamma.
\]

Proof: We have, for any \( \mathbf{u}_\Gamma \in \tilde{\mathbf{W}}_\Gamma \),

\[
\mathbf{u}_\Gamma^T \hat{M}^{-1} \mathbf{u}_\Gamma = \mathbf{u}_\Gamma^T \tilde{\mathbf{R}}_{D,\Gamma} \left( \tilde{\mathbf{R}}_{\Gamma,\Delta} \sum_{i=1}^N \left( \mathbf{0} \right) \mathbf{R}^{(i)}_{\Delta} \left( \begin{array}{cc} A^{(i)}_{\Delta} & A^{(i)}_{\Delta} \\ A^{(i)}_{\Delta} & A^{(i)}_{\Delta} \end{array} \right)^{-1} \left( \mathbf{0} \right) \mathbf{R}^{(i)}_{\Delta} \right) \tilde{\mathbf{R}}_{D,\Gamma} \mathbf{u}_\Gamma
\]

and

\[
\mathbf{u}_\Gamma^T \hat{M}^{-1} \mathbf{u}_\Gamma = \mathbf{u}_\Gamma^T \tilde{\mathbf{R}}_{D,\Gamma} \left( \tilde{\mathbf{R}}_{\Gamma,\Delta} \sum_{i=1}^N \left( \mathbf{0} \right) \mathbf{R}^{(i)}_{\Delta} \left( \begin{array}{cc} A^{(i)}_{\Delta} & A^{(i)}_{\Delta} \\ A^{(i)}_{\Delta} & A^{(i)}_{\Delta} \end{array} \right)^{-1} \left( \mathbf{0} \right) \mathbf{R}^{(i)}_{\Delta} \right) \tilde{\mathbf{R}}_{D,\Gamma} \mathbf{u}_\Gamma
\]

We obtain our result by using Lemma 4.6.

6. Using Chebyshev iterations. Another approach to the three-level BDDC methods is to use an iterative method with a preconditioner to solve (3.4). Here, we consider a Chebyshev method with a fixed number of iterations and use \( \left( \tilde{\mathbf{h}}_\Gamma^T \hat{\mathbf{R}}_{\tilde{\mathbf{D}},\Gamma}^{-1} \tilde{\mathbf{h}}_\Gamma \right) \) as a preconditioner.

Thus, we do not solve (3.4) exactly. Instead, we replace \( \mathbf{y}_\Gamma \) by \( \mathbf{y}_{\Gamma, k} \), where \( \mathbf{y}_{\Gamma, k} \) is the approximation of \( \mathbf{y}_\Gamma \) given by a \( k \)-step Chebyshev iteration.

We will maintain the same relation between \( \mathbf{y}_{\Gamma, k}^{(i)} \) and \( \mathbf{y}_{\Gamma, k}^{(i)} \) as in (3.1), i.e.,
(6.1) \[ y_{I,k}^{(i)} = S_{II}^{(i)-1} \left( \Psi_{I}^{(i)} - S_{II}^{(i)T} \tilde{R}_{I}^{(i)} y_{I,k} \right). \]

Let \( y_k = (y_{I,k}^{(1)}, \ldots, y_{I,k}^{(N)}, y_{\Gamma,k})^T \), and denote the corresponding new coarse problem matrix by \( S_{II} \). Then,

(6.2) \[ S_{II} y_k = \Psi, \]

and the new preconditioner \( \tilde{M}^{-1} \) is defined by:

\[
\tilde{R}_{I,D,\Gamma}^T \left\{ \begin{array}{c}
R_{I,\Delta}^T \\
R_{I,\Delta}^T
\end{array} \left( \sum_{i=1}^{N} \begin{array}{c}
0 \\
R_{I,\Delta}^T
\end{array} \left( \begin{array}{cc}
A_{II}^{(i)} & A_{I\Delta}^{(i)} \\
A_{I\Delta}^{(i)} & A_{\Delta\Delta}^{(i)}
\end{array} \right)^{-1} \begin{array}{c}
0 \\
R_{\Delta}
\end{array} \right) \right\} = 0. \]

6.1. Algorithm. Denoting the eigenvalues of \( \left( \tilde{R}_{D,\Gamma}^T \tilde{R}_{D,\Gamma}^{-1} \right) \) by \( \lambda_j \), we need two input parameters \( l \) and \( u \) for the Chebyshev iterations, where \( l \) and \( u \) are estimates for the minimum and maximum values of \( \lambda_j \), respectively; see [4]. From our analysis above, we know that \( \min_j \lambda_j = 1 \) and \( \max_j \lambda_j \leq C(1 + \log \frac{u}{l})^2(1 + \log \frac{H}{h})^2 \). We can use the conjugate gradient method to obtain an estimate for the largest eigenvalue at the beginning of the computation to choose a proper \( u \).

Let \( \alpha = \frac{2}{1+u} \) and \( \mu = \frac{u/l}{1+u} \). Let \( c_k \) be the value of the \( k^{th} \) Chebyshev polynomial evaluated at \( \mu \), i.e.,

(6.3) \[ c_{k+1} = 2\mu c_k - c_{k-1}, \quad k = 1, 2, \ldots, \text{ with } c_0 = 1 \text{ and } c_1 = \mu. \]

We use a zero initial guess and the Chebyshev acceleration is defined by, see [4],

(6.4) \[ y_{\Gamma,0} = 0 \text{ and } y_{\Gamma,1} = y_{\Gamma,0} + \alpha z_0, \]

(6.5) \[ y_{\Gamma,k+1} = y_{\Gamma,k} + \omega_{k+1} (\alpha z_k + y_{\Gamma,k} - y_{\Gamma,k-1}), \quad k = 1, 2, \ldots, \]

where

\[ \omega_{k+1} = \frac{2\mu c_k}{c_{k+1}}. \]

6.2. Error analysis. Let \( e_k = y_{\Gamma} - y_{\Gamma,k} \). Using (6.4) and (6.5), we obtain

(6.7) \[ e_{k+1} = \omega_{k+1} Q e_k + (1 - \omega_{k+1}) e_{k-1}, \text{ with } e_0 = y_{\Gamma} \text{ and } e_1 = Q e_0, \]

where \( Q = I - \alpha \left( \tilde{R}_{D,\Gamma}^T \tilde{R}_{D,\Gamma}^{-1} \right) \left( \tilde{R}_{\Gamma}^T \tilde{R}_{\Gamma}^{-1} \right), \)
The symmetrized operator \( \left( R_{\tilde{D},\tilde{T}}^T \tilde{R}_{\tilde{D},\tilde{T}}^{-1} \tilde{R}_{\tilde{D},\tilde{T}} \right)^{\frac{1}{2}} \left( \tilde{R}_{\tilde{T}}^T \tilde{R}_{\tilde{T}} \right) \left( R_{\tilde{D},\tilde{T}}^T \tilde{R}_{\tilde{D},\tilde{T}}^{-1} \tilde{R}_{\tilde{D},\tilde{T}} \right)^{\frac{1}{2}} \) has the following eigenvalue decomposition:

\[
(6.8) \quad \left( \tilde{R}_{\tilde{D},\tilde{T}}^T \tilde{R}_{\tilde{D},\tilde{T}}^{-1} \tilde{R}_{\tilde{D},\tilde{T}} \right)^{\frac{1}{2}} \left( \tilde{R}_{\tilde{T}}^T \tilde{R}_{\tilde{T}} \right) \left( \tilde{R}_{\tilde{D},\tilde{T}}^T \tilde{R}_{\tilde{D},\tilde{T}}^{-1} \tilde{R}_{\tilde{D},\tilde{T}} \right)^{\frac{1}{2}} = P \Lambda P^T,
\]

where \( \Lambda \) is a diagonal matrix and the eigenvalues \( \lambda_j \) its diagonal entries. \( P \) is an orthogonal matrix, and \( P^T \) is its transpose.

Let

\[
(6.9) \quad P_1 = \left( \tilde{R}_{\tilde{D},\tilde{T}}^T \tilde{R}_{\tilde{D},\tilde{T}}^{-1} \tilde{R}_{\tilde{D},\tilde{T}} \right)^{\frac{1}{2}} P.
\]

We note that

\[
(6.10) \quad Q = P_1 \Sigma P_1^{-1},
\]

where \( \Sigma \) is a diagonal matrix with the eigenvalues \( \sigma_j \) of \( Q \) on the diagonal and \( \sigma_j = 1 - \alpha \lambda_j \). Let

\[
(6.11) \quad f_k = c_k P_1^{-1} e_k.
\]

If we substitute (6.11) into (6.7), we then obtain a diagonal system of difference equations by using (6.3), (6.6), and (6.10):

\[
f_{k+1} = 2 \mu \Sigma f_k - f_{k-1}, \quad k = 1, 2, \ldots, \quad \text{with} \quad f_1 = \mu \Sigma f_0 \text{ and } f_0 = P_1^{-1} y_{\tilde{T}}.
\]

Solving this system, see [4], we obtain

\[
f_k = \Theta P_1^{-1} y_{\tilde{T}}, \quad k = 1, 2, \ldots,
\]

where \( \Theta \) is a diagonal matrix with \( \cosh \left( k \cosh^{-1}(\mu \sigma_j) \right) \) on its diagonal. Using (6.11), we obtain:

\[
(6.12) \quad e_k = \left( P_1 \Theta P_1^{-1} \right) \frac{y_{\tilde{T}}}{c_k}, \quad k = 1, 2, \ldots.
\]

Using the definition of \( e_k \), our approximate solution after \( k \) Chebyshev iterations is given by

\[
(6.13) \quad y_{\tilde{T},k} = P_1 J P_1^{-1} y_{\tilde{T}},
\]

where \( J \) is a diagonal matrix with \( 1 - \frac{\cosh(k \cosh^{-1}(\mu \sigma_j))}{c_k} \) on its diagonal. Using (6.3) we obtain \( c_k = \cosh \left( k \cosh^{-1}(\mu) \right) \).

Therefore, we have \( 1 - \frac{\cosh(k \cosh^{-1}(\mu \sigma_j))}{\cosh(k \cosh^{-1}(\mu))} \) as the diagonal entries of the matrix \( J \).

From (6.12), we know that the Chebyshev iteration converges if and only if \( |\sigma_j| < 1 \), i.e., if \( 0 < \lambda_j < l + u \). Since \( \left( \tilde{R}_{\tilde{D},\tilde{T}}^T \tilde{R}_{\tilde{D},\tilde{T}}^{-1} \tilde{R}_{\tilde{D},\tilde{T}} \right) \left( \tilde{R}_{\tilde{T}}^T \tilde{R}_{\tilde{T}} \right) \) has a smallest eigenvalue 1, we can guarantee that \( 0 < \lambda_j \) and choose \( l = 1 \). From our analysis above, we get an upper bound for \( \lambda_j \) and can choose \( u \) to guarantee that \( \lambda_j < l + u \).

Since we choose \( u \) such that \( \lambda_j < l + u \), we find that \( 1 - \frac{\cosh(k \cosh^{-1}(\mu \sigma_j))}{\cosh(k \cosh^{-1}(\mu))} > 0 \), i.e., \( J \) has positive diagonal elements.
6.3. Condition number estimate for the second new preconditioner. We begin with a lemma.

**Lemma 6.1.** Given any \( u_\Gamma \in \hat{W}_\Gamma \), let \( \Psi = \Phi^T \tilde{R}_{D,\Gamma} u_\Gamma \). If we choose \( u \) such that \( \lambda_j < u + l \), then there exist two constants \( C_1(k) \) and \( C_2(k) \) that

\[
C_1(k) \Psi^T \tilde{S}_{\Pi}^{-1} \Psi \leq \Psi^T \tilde{S}_{\Pi}^{-1} \Psi^T \leq C_2(k) \Psi^T \tilde{S}_{\Pi}^{-1} \Psi,
\]

where

\[
(6.14) \quad C_1(k) = \min_j \left( 1 - \frac{\cosh(k \cosh^{-1}(\mu_j))}{\cosh(k \cosh^{-1}(\mu))} \right)
\]

and

\[
(6.15) \quad C_2(k) = \max_j \left( 1 - \frac{\cosh(k \cosh^{-1}(\mu_j))}{\cosh(k \cosh^{-1}(\mu))} \right).
\]

**Proof:** Using (6.1), (3.3), and (3.4), we have,

\[
\Psi^T \tilde{S}_{\Pi}^{-1} \Psi = \sum_{i=1}^{N} \Psi_{\Gamma}^{(i)} y_{\Gamma,k}^T + \Psi_{\Gamma}^T \Psi_{\Gamma,k}^{T}
\]

\[
= \sum_{i=1}^{N} \Psi_{\Gamma}^{(i)} \left( S_{\Pi}^{(i)} \Psi_{\Gamma}^{(i)} - S_{\Pi}^{(i)} \tilde{R}_{\Gamma} y_{\Gamma,k} \right) + \left( h_{\Gamma} + \sum_{i=1}^{N} \tilde{R}_{\Gamma} S_{\Pi}^{(i)} \tilde{S}_{\Pi}^{-1} \Psi_{\Gamma}^{(i)} \right) y_{\Gamma,k}^{T}
\]

\[
= \sum_{i=1}^{N} \Psi_{\Gamma}^{(i)} S_{\Pi}^{(i)} \Psi_{\Gamma}^{(i)} + h_{\Gamma} y_{\Gamma,k} = \sum_{i=1}^{N} \Psi_{\Gamma}^{(i)} S_{\Pi}^{(i)} \Psi_{\Gamma}^{(i)} + y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) y_{\Gamma,k}^{T}.
\]

We only need to compare \( y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) y_{\Gamma,k} \) and \( y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) y_{\Gamma,k} \) by comparing (6.16) with (4.4).

Using (6.13) and (6.9), we obtain

\[
y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) y_{\Gamma,k} = y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) P_{\Gamma} J P_{\Gamma}^{-1} y_{\Gamma}
\]

\[
= y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right)^{\frac{1}{2}} P_{\Gamma} J P_{\Gamma}^{-1} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right)^{\frac{1}{2}} y_{\Gamma}.
\]

Let \( Y_{\Gamma} = P^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right)^{\frac{1}{2}} y_{\Gamma} \). Using (6.8), we have,

\[
y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) y_{\Gamma,k} = y_{\Gamma}^{T} \left( \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} \right) y_{\Gamma,k} = Y_{\Gamma}^{T} P T \tilde{R}_{\Gamma} \tilde{S}_{\Pi} \tilde{R}_{\Gamma} P Y_{\Gamma} = Y_{\Gamma}^{T} A Y_{\Gamma}.
\]

17
and, using (6.17), we have
\[
\begin{align*}
  y^T_{\tilde{\Gamma}} & \left( \tilde{R}_{\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Gamma}} \right) y_{\tilde{\Gamma},k} \\
  &= Y^T_{\tilde{\Gamma}} P^T \left( \tilde{R}_{\tilde{\Delta},\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Delta},\tilde{\Gamma}} \right)^{-\frac{1}{2}} \left( \tilde{R}_{\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Gamma}} \right)^{-\frac{1}{2}} \left( \tilde{R}_{\tilde{\Delta},\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Delta},\tilde{\Gamma}} \right)^{-\frac{1}{2}} \left( \tilde{R}_{\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Gamma}} \right)^{-\frac{1}{2}} P J Y_{\tilde{\Gamma}} \\
  &= Y^T_{\tilde{\Gamma}} P^T P \triangle P^T P J Y_{\tilde{\Gamma}} = Y^T_{\tilde{\Gamma}} \Lambda J Y_{\tilde{\Gamma}}.
\end{align*}
\]

Under our assumption, \( J \) is a diagonal matrix with positive diagonal entries \( \left( 1 - \frac{\cosh(k \cosh^{-1}(\mu \sigma_i)))}{\cosh(k \cosh^{-1}(\mu))} \right) \). Thus, we have,
\[
C_1(k) Y^T_{\tilde{\Gamma}} \left( \tilde{R}_{\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Gamma}} \right) y_{\tilde{\Gamma}} \leq Y^T_{\tilde{\Gamma}} \left( \tilde{R}_{\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Gamma}} \right) y_{\tilde{\Gamma},k} \leq C_2(k) Y^T_{\tilde{\Gamma}} \left( \tilde{R}_{\tilde{\Gamma}} \tilde{T} \tilde{R}_{\tilde{\Gamma}} \right) y_{\tilde{\Gamma}}.
\]

**Lemma 6.2.** Given any \( u_\Gamma \in \hat{W}_\Gamma \),
\[
C_1(k) u^T_\Gamma M^{-1} u_\Gamma \leq u^T_\Gamma \hat{M}^{-1} u_\Gamma \leq C_2(k) u^T_\Gamma M^{-1} u_\Gamma,
\]
where \( C_1(k) \) and \( C_2(k) \) are defined in (6.14) and (6.15), respectively.

**Proof:** We have, for any \( u_\Gamma \in \hat{W}_\Gamma \),
\[
\begin{align*}
  u^T_\Gamma \hat{M}^{-1} u_\Gamma &= u^T_\Gamma \tilde{R}_{\tilde{\Gamma},\Gamma} \left\{ \tilde{R}_{\tilde{\Delta}} \sum_{i=1}^{N} \left( 0 R^{(i)}_{\Delta} \right) \begin{pmatrix} A^{(i)}_{\Delta} & A^{(i)}_{\Delta} \end{pmatrix}^{-1} \begin{pmatrix} 0 \ 0 \end{pmatrix} \tilde{R}_{\Delta,\Gamma} \right\} u_\Gamma \\
  &+ u^T_\Gamma \tilde{R}_{\tilde{\Gamma},\Gamma} \Phi_\tilde{\Gamma}^{-1} \Phi^T \tilde{R}_{\tilde{\Gamma},\Gamma} u_\Gamma.
\end{align*}
\]
Comparing this expression with (5.2), we obtain the result by using Lemma 6.1.

**Theorem 6.3.** The condition number using the three-level preconditioner \( \hat{M}^{-1} \) is bounded by \( C \frac{C_2(k)}{C_1(k)} (1 + \log \frac{H}{n})^2 \), where \( C_1(k) \) and \( C_2(k) \) are defined in (6.14) and (6.15), respectively.

**Proof:** Combining the condition number bound, given in Lemma 4.4, for the two-level BDDC method and Lemma 6.2, we find that the condition number for the system with the three-level preconditioner \( \hat{M}^{-1} \) is bounded by \( C \frac{C_2(k)}{C_1(k)} (1 + \log \frac{H}{n})^2 \).

**7. Numerical experiments.** We have applied our two three-level BDDC algorithms to the model problem (2.1). We decompose the unit square into \( N \times N \) subregions and each subregion into \( N \times N \) subdomains with the side-length \( H = 1/N \) and \( h = H/N \), respectively. Equation (2.1) is discretized, in each subdomain, by conforming piecewise linear elements with a finite element diameter \( h \); see Fig 1. The preconditioned conjugate gradient iteration is stopped when the norm of the residual has been reduced by a factor of \( 10^{-8} \).

We have carried out five different sets of experiments to obtain iteration counts and condition number estimates. In some cases the assumption of our theory are not satisfied. All the experimental results, for the cases where our analysis applies, are fully consistent with our theory.
Table 1
Condition number estimates and iteration counts with the preconditioner $\tilde{M}^{-1}$ for constant coefficients

<table>
<thead>
<tr>
<th>$\frac{H}{H}$</th>
<th>$\frac{H}{h}$</th>
<th>$4 \times 4$ subreg.</th>
<th>$\frac{H}{H}$</th>
<th>$\frac{H}{h}$</th>
<th>$4 \times 4$ subreg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4 \times 4$</td>
<td>4</td>
<td>12</td>
<td>3.0428</td>
<td>4</td>
<td>12</td>
</tr>
<tr>
<td>$8 \times 8$</td>
<td>15</td>
<td>13</td>
<td>4.1662</td>
<td>8</td>
<td>15</td>
</tr>
<tr>
<td>$12 \times 12$</td>
<td>17</td>
<td>13</td>
<td>4.9566</td>
<td>12</td>
<td>16</td>
</tr>
<tr>
<td>$16 \times 16$</td>
<td>17</td>
<td>14</td>
<td>5.5739</td>
<td>16</td>
<td>17</td>
</tr>
<tr>
<td>$20 \times 20$</td>
<td>17</td>
<td>15</td>
<td>6.0840</td>
<td>20</td>
<td>19</td>
</tr>
</tbody>
</table>

Table 2
Condition number estimates and iteration counts with the preconditioner $\tilde{M}^{-1}$ for coefficients in a checkerboard pattern

<table>
<thead>
<tr>
<th>$\frac{H}{H}$</th>
<th>$\frac{H}{h}$</th>
<th>$4 \times 4$ subreg.</th>
<th>$\frac{H}{H}$</th>
<th>$\frac{H}{h}$</th>
<th>$4 \times 4$ subreg.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$4 \times 4$</td>
<td>11</td>
<td>1.8096</td>
<td>4</td>
<td>11</td>
<td>1.8096</td>
</tr>
<tr>
<td>$8 \times 8$</td>
<td>11</td>
<td>1.8145</td>
<td>8</td>
<td>12</td>
<td>1.8536</td>
</tr>
<tr>
<td>$12 \times 12$</td>
<td>12</td>
<td>1.8159</td>
<td>12</td>
<td>12</td>
<td>2.7072</td>
</tr>
<tr>
<td>$16 \times 16$</td>
<td>12</td>
<td>1.8162</td>
<td>16</td>
<td>12</td>
<td>3.3473</td>
</tr>
<tr>
<td>$20 \times 20$</td>
<td>12</td>
<td>1.8164</td>
<td>20</td>
<td>12</td>
<td>3.6546</td>
</tr>
</tbody>
</table>

Table 3
Condition number estimated and iteration counts with the preconditioner $\tilde{M}$ for constant coefficients, $u = 3.2$, $4 \times 4$ subregions, $\frac{H}{H} = 16$ and $\frac{H}{h} = 4$

<table>
<thead>
<tr>
<th>k</th>
<th>Iter.</th>
<th>$C_1(k)$</th>
<th>$\lambda_{min}$</th>
<th>$\lambda_{max}$</th>
<th>Cond. #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>20</td>
<td>0.4762</td>
<td>0.4829</td>
<td>2.7110</td>
<td>5.6141</td>
</tr>
<tr>
<td>2</td>
<td>13</td>
<td>0.8410</td>
<td>0.8540</td>
<td>1.8290</td>
<td>2.0638</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>0.9548</td>
<td>0.9981</td>
<td>1.9061</td>
<td>1.9998</td>
</tr>
<tr>
<td>4</td>
<td>11</td>
<td>0.9872</td>
<td>1.0019</td>
<td>1.8663</td>
<td>1.8629</td>
</tr>
<tr>
<td>5</td>
<td>11</td>
<td>0.9964</td>
<td>1.0006</td>
<td>1.8551</td>
<td>1.8541</td>
</tr>
</tbody>
</table>

Table 4
Condition number estimates and iteration counts with the preconditioner $\tilde{M}$ for constant coefficients, $u = 6$, $4 \times 4$ subregions, $\frac{H}{H} = 16$ and $\frac{H}{h} = 4$

<table>
<thead>
<tr>
<th>k</th>
<th>Iter.</th>
<th>$C_1(k)$</th>
<th>$\lambda_{min}$</th>
<th>$\lambda_{max}$</th>
<th>Cond. #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>0.2857</td>
<td>0.2899</td>
<td>1.8287</td>
<td>6.3086</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>0.6357</td>
<td>0.6670</td>
<td>2.3435</td>
<td>3.5134</td>
</tr>
<tr>
<td>3</td>
<td>12</td>
<td>0.8524</td>
<td>0.9286</td>
<td>1.9628</td>
<td>3.1136</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>0.9377</td>
<td>0.9795</td>
<td>1.9850</td>
<td>2.0266</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>0.9738</td>
<td>0.9983</td>
<td>1.9403</td>
<td>1.9437</td>
</tr>
</tbody>
</table>
The unit square is partitioned into subregions with the side-length $\hat{H}$. Each subregion is partitioned into subdomains with the side-length $H$. Each subdomain is partitioned into elements with mesh size $h$.

![Diagram of a unit square partitioned into subregions and subdomains](image)

**Table 5**

Condition number estimates and iteration counts with the preconditioner $\tilde{M}$ for coefficients in a checkerboard pattern, $u = 2$, $4 \times 4$ subregions, $\frac{H}{h} = 16$ and $\frac{H}{h} = 4$

<table>
<thead>
<tr>
<th>$k$</th>
<th>Iter.</th>
<th>$C_1(k)$</th>
<th>$\lambda_{\min}$</th>
<th>$\lambda_{\max}$</th>
<th>Cond. #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>16</td>
<td>0.6667</td>
<td>0.6773</td>
<td>1.8267</td>
<td>2.6969</td>
</tr>
<tr>
<td>2</td>
<td>12</td>
<td>0.9412</td>
<td>0.9539</td>
<td>1.8258</td>
<td>1.9140</td>
</tr>
<tr>
<td>3</td>
<td>11</td>
<td>0.9899</td>
<td>1.0020</td>
<td>1.8395</td>
<td>1.8359</td>
</tr>
<tr>
<td>4</td>
<td>12</td>
<td>0.9983</td>
<td>1.0021</td>
<td>1.8428</td>
<td>1.8388</td>
</tr>
<tr>
<td>5</td>
<td>12</td>
<td>0.9997</td>
<td>1.0020</td>
<td>1.8394</td>
<td>1.8358</td>
</tr>
</tbody>
</table>

In the first set of experiments, we use the first preconditioner $\tilde{M}^{-1}$ and take the coefficient $\rho \equiv 1$. Table 1 gives the iteration counts and condition number estimates with a change of the number of subregions, number of subdomains, and the size of the subdomain problems.

In the second set of experiments, we use the first preconditioner $\tilde{M}^{-1}$ and take the coefficient $\rho = 1$ in half of the subregions and $\rho = 101$ in the neighboring subregions in a checkerboard pattern. Table 2 gives the iteration counts and condition number estimates with a change of the number of subregions, the number of subdomains, and the size of the subdomain problems.

In the third set of experiments, we use the second preconditioner $\hat{M}^{-1}$ and take the coefficient $\rho \equiv 1$. We use the PCG to estimate the largest eigenvalue of $\tilde{M}$ which is approximately 3.2867. And if we have $64 \times 64$ subdomains and $\frac{H}{h} = 4$ for the two-level BDDC, we have a condition number estimate of 1.8380. We select different values of $u$, the upper bound estimate of the eigenvalues for the preconditioner system, and $k$ to see how the condition number changes. We also evaluate $C_1(k)$ for $k = 1, 2, 3, 4, 5$. From Table 3 and 4, we find that the smallest eigenvalue is bounded from below by $C_1(k)$ and the condition num-
Condition number estimates and iteration counts with the preconditioner $\tilde{M}$ for coefficients in a checkerboard pattern, $u = 6$, $4 \times 4$ subdomains, $H = 16$ and $H = 4$

<table>
<thead>
<tr>
<th>$k$</th>
<th>Iter.</th>
<th>$C_1(k)$</th>
<th>$\lambda_{\min}$</th>
<th>$\lambda_{\max}$</th>
<th>Cond. #</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>25</td>
<td>0.2857</td>
<td>0.2901</td>
<td>1.8270</td>
<td>6.2976</td>
</tr>
<tr>
<td>3</td>
<td>16</td>
<td>0.6575</td>
<td>0.6682</td>
<td>1.8267</td>
<td>2.7338</td>
</tr>
<tr>
<td>3</td>
<td>13</td>
<td>0.8524</td>
<td>0.9281</td>
<td>1.8264</td>
<td>1.9679</td>
</tr>
<tr>
<td>4</td>
<td>13</td>
<td>0.9377</td>
<td>1.0045</td>
<td>1.9100</td>
<td>1.9015</td>
</tr>
<tr>
<td>5</td>
<td>13</td>
<td>0.9738</td>
<td>1.0058</td>
<td>1.9451</td>
<td>1.9340</td>
</tr>
</tbody>
</table>

Condition number estimates and iteration counts with the preconditioner $M^{-1}$ for coefficient in a checkerboard subdomain pattern

<table>
<thead>
<tr>
<th>$\frac{H}{H}$</th>
<th>$\frac{H}{h}$ = 4</th>
<th>$\frac{H}{h}$ = 4</th>
<th>$\frac{H}{h}$ = 4</th>
<th>$\frac{H}{h}$ = 4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subreg.</td>
<td>Iter.</td>
<td>Cond. #</td>
<td>$\frac{H}{H}$</td>
<td>Iter.</td>
</tr>
<tr>
<td>4 $\times$ 4</td>
<td>15</td>
<td>9.5041</td>
<td>4</td>
<td>15</td>
</tr>
<tr>
<td>8 $\times$ 8</td>
<td>29</td>
<td>26.0935</td>
<td>8</td>
<td>17</td>
</tr>
<tr>
<td>12 $\times$ 12</td>
<td>44</td>
<td>39.9770</td>
<td>12</td>
<td>18</td>
</tr>
<tr>
<td>16 $\times$ 16</td>
<td>55</td>
<td>49.8132</td>
<td>16</td>
<td>19</td>
</tr>
<tr>
<td>20 $\times$ 20</td>
<td>65</td>
<td>56.7763</td>
<td>20</td>
<td>20</td>
</tr>
</tbody>
</table>

The condition number estimate becomes closer to 1.8380, the value in the two-level case, as $k$ increases. We also see that if we can get more precise estimate for the largest eigenvalue of $R_D T \tilde{R}^{-1}_D$, we need fewer Chebyshev iterations to get a condition number, similar to that of the two-level case. However, the iteration count is not very sensitive to the choice of $\rho$.

In the fourth set of experiments, we use the second preconditioner $M^{-1}$ and take the coefficient $\rho = 1$ in half of the subdomains and $\rho = 101$ in the neighboring subdomains in a checkerboard pattern. We use the PCG to estimate the largest eigenvalue of $R_D T \tilde{R}_D^{-1} T \tilde{R}_D$. And if we have 64 $\times$ 64 subdomains and $H = 4$ for the two-level BDDC, we have a condition number estimate of 1.8338. We select different values of $u$, the upper bound estimate of the eigenvalues for the preconditioner system, and $k$ to see how the condition number changes. We also evaluate $C_1(k)$ for $k = 1, 2, 3, 4, 5$. From Table 5 and 6, we find that the smallest eigenvalue is bounded from below by $C_1(k)$ and the condition number estimate becomes closer to 1.8338, the value in the two-level case, as $k$ increases.

We also see that if we can get more precise estimate for the largest eigenvalue of $R_D T \tilde{R}_D^{-1} T \tilde{R}_D$, we need fewer Chebyshev iterations to get a condition number, similar to that of the two-level case. However, the iteration count is still not very sensitive to the choice of $u$.

In the last set of experiments, we use the first preconditioner $\tilde{M}^{-1}$ and take the coefficient $\rho = 1$ in half of the subdomains and $\rho = 101$ in the neighboring...
Fig. 2. Estimated condition number from Table 7 (asterisk) and the least-square second order logarithmic polynomial (solid line, \(-11.9198 + 8.5831 \log \frac{H}{h} + 4.8544 \log^2 \frac{H}{h}\), versus \(\frac{H}{h}\). \(\frac{H}{h} = 4\) and \(\frac{H}{h} = 4\).

Fig. 3. Estimated condition number from Table 7 (asterisk) and the least-square second order logarithmic polynomial (solid line, \(3.2814 + 3.8284 \log \frac{H}{h} + 0.4759 \log^2 \frac{H}{h}\), versus \(\frac{H}{h}\). \(\frac{H}{h} = 4\) and \(\frac{H}{h} = 4\).

subdomains in a checkerboard subdomain pattern. The coefficients have big jump inside each subregion, and our analysis does not apply. Table 7 gives the iteration counts and condition number estimates with a change of the number of subregions, the number of subdomains, and the size of the subdomain problems. Fig 2 shows that the condition number grows in proportion to \((1 + \log \frac{1}{H})^2\) with a changing of the number of the subregions. Fig 3 shows that the condition number grows in proportion to \((1 + \log \frac{H}{h})^2\) with a changing of the number of the subdomains. Fig 4 shows that the condition number grows in proportion to \((1 + \log \frac{H}{h})^2\) with a changing of the size of the subdomain problems. The numerical results are still consistent with our theory for changes of the number of subdomains and the size of the size of subdomain problems.
Fig. 4. Estimated condition number from Table 7 (asterisk) and the least-square second order logarithmic polynomial (solid line, \(9.5582 - 0.0974 \log \frac{H}{h} + 0.0420 \log^2 \frac{H}{h}\)), versus \(\frac{H}{h}\). \(\frac{H}{h} = 4\) and \(\frac{H}{h} = 4\).

8. Conclusion. As we have noted before, the coarse components of the two-level BDDC preconditioners can ultimately become a bottleneck if the number of subdomains is very large. In three-level algorithms, we alleviate this difficulty by using one additional level.

Let us look at the cost of the two-level and three-level BDDC methods closely. We assume that the diameter of the computational domain is \(O(1)\). \(\hat{H}, H,\) and \(h\) are the typical diameters of the subregions, subdomains, and elements of subdomains, respectively. The numbers of degrees of freedom of the original system, the subdomain local problems, and the coarse problem are \(O(h^2)\), \(O(H^2 h^2)\), and \(O(H^3)\), respectively. Therefore the optimal flop bounds for the Cholesky factorization of a local matrix and the coarse matrix are \(O(H^3 h^3)\) and \(O(H^3)\), respectively; see [3, Section 8.1]. In each iteration, the cost of the forward and backward substitutions are \(O(\frac{H^2}{h^2} \log(\frac{H^2}{h^2}))\) and \(O(\frac{1}{H^2} \log(\frac{1}{H^2}))\). In a parallel computation, we can assign one subdomain to one processor and the coarse problem to one processor also or assign the coarse problem to each processor. We focus on the first approach here. If we store the coarse basis functions at the beginning of the computation, then the BDDC preconditioner fits an additive framework. This means that we can solve the coarse and local problems simultaneously. We denote the number of the processors by \(N + 1\). Then \(N \sim \frac{1}{H^2}\). When \(N\) is large, then the size of the coarse matrix is large and all the processors for the local problems will wait for the processor for the coarse problem and the coarse problem becomes a bottleneck of the two-level BDDC methods.

For the three-level BDDC methods, we could first reduce the original coarse problem to a subregion interface problem by eliminating independently the subregion interior variables, which are the primal variables on the subdomain interface and interior to the subregions. However, in our three-level BDDC algorithms, we do not solve the subregion interface problem exactly, but replace it by one iteration of the BDDC preconditioner. This means that we only need to solve several subregion local problems and one coarse problem on the subregion level in each iteration. Then the sizes of the subregion local problems and the subregion coarse problem are \(O(\frac{H^2}{h^2})\) and \(O(\frac{1}{H^2})\). Therefore, the optimal flop bounds for the Cholesky factorization for a subdomain
local matrix, a subregion local matrix, and the coarse matrix are $O\left(\frac{H^3}{h^3}\right)$, $O\left(\frac{H^3}{H^3}\right)$, and $O\left(\frac{1}{H}\right)$, respectively. In each iteration, the cost of the forward and backward substitutions are $O\left(\frac{H^2}{h^2}\log\left(\frac{H}{h}\right)\right)$, $O\left(\frac{H^2}{H^2}\log\left(\frac{H}{H}\right)\right)$, and $O\left(\frac{1}{H}\log\left(\frac{1}{H}\right)\right)$. We can assign the subregion local problems to the processors for the subdomain local problems and the subregion coarse problem to the processor for the two-level coarse problem. Usually $H$ is much larger than $h$, therefore the size of the subregion coarse problem is much smaller than that of the original coarse problem.

Acknowledgments The author is grateful to Professor Olof Widlund for suggesting this problem and all the help and time he has devoted to my work. The author also thanks Professor Jing Li for personal communications and useful notes.

REFERENCES