Additive Sweeping Preconditioner for the Helmholtz Equation

Fei Liu♯ and Lexing Ying†‡

† Department of Mathematics, Stanford University
‡ Institute for Computational and Mathematical Engineering, Stanford University

April 2015

Abstract

We introduce a new additive sweeping preconditioner for the Helmholtz equation based on the perfect matched layer (PML). This method divides the domain of interest into thin layers and proposes a new transmission condition between the subdomains where the emphasis is on the boundary values of the intermediate waves. This approach can be viewed as an effective approximation of an additive decomposition of the solution operator. When combined with the standard GMRES solver, the iteration number is essentially independent of the frequency. Several numerical examples are tested to show the efficiency of this new approach.

Keyword. Helmholtz equation, perfectly matched layers, preconditioners, high frequency waves.

AMS subject classifications. 65F08, 65N22, 65N80.

1 Introduction

Let the domain of interest be $D = (0, 1)^d$ where $d = 2, 3$. The Helmholtz equation is

$$\Delta u(x) + \frac{\omega^2}{c^2(x)} u(x) = f(x), \quad \forall x \in D,$$

where $u(x)$ is the time-independent wave field generated by the time-independent force $f(x)$, $\omega$ is the angular frequency and $c(x)$ is the velocity field. Commonly used boundary conditions are the approximations of the Sommerfeld radiation condition. By rescaling the system, we assume $c_{\text{min}} \leq c(x) \leq c_{\text{max}}$ where $c_{\text{min}}$ and $c_{\text{max}}$ are of $\Theta(1)$. Then $\omega/(2\pi)$ is the typical wave number and $\lambda = 2\pi/\omega$ is the typical wavelength.

Solving the equation numerically is challenging in high frequency settings for two reasons. First, in most applications, the equation is discretized with at least a constant number of points per wavelength, which makes the number of points in each direction $n = \Omega(\omega)$ and the total degree of freedom $N = n^d = \Omega(\omega^d)$ very large. Second, the system is highly indefinite and has a very oscillatory Green’s function, which makes most of the classical iterative methods no longer effective.

There has been a sequence of papers on developing iterative methods for solving (1). The AILU method by Gander and Nataf [10] is the first to use the incomplete LU factorization to precondition the equation. Engquist and Ying [6, 7] developed a series of sweeping preconditioners based on
approximating the inverse of the Schur complements in the LDU factorization and obtained essentially \( \omega \)-independent iteration numbers. In [15], Stolk proposed a domain decomposition method based on the PML which constructs delicate transmission conditions between the subdomains by considering the “pulses” generated by the intermediate waves. In [19], Vion and Geuzaine proposed a double sweep preconditioner based on the Dirichlet-to-Neumann (DtN) map and several numerical simulations of the DtN map were compared. In [2, 3], Chen and Xiang introduced a source transfer domain decomposition method which emphasizes on transferring the sources between the subdomains. In [20], Zepeda-Núñez and Demanet developed a novel domain decomposition method for the 2D case by pairing up the waves and their normal derivatives at the boundary of the subdomains and splitting the transmission of the waves into two directions. Most recently in [13], Liu and Ying proposed a recursive sweeping preconditioner for 3D Helmholtz problems. Other progresses includes [14, 18, 16, 17] and we refer to [8] by Erlangga and [9] by Ernst and Gander for a complete discussion.

Inspired by [15] and these previous approaches, we propose a new domain decomposition method in this paper which shares some similarities with [7, 15]. The novelty of this new approach is that the transmission conditions are built with the boundary values of the intermediate waves directly. For each wave field on the subdomains, we divide it into three parts – the waves generated by the force to the left of the subdomain, to the right of the subdomain, and within the subdomain itself. This corresponds to an \( L + D + U \) decomposition of the Green’s matrix \( G \) as the sum of its lower triangular part, upper triangular part and diagonal part. This is why we call this new preconditioner the additive sweeping preconditioner.

The rest of this paper is organized as follows. First in Section 2 we use the 1D case to illustrate the idea of the method. Then in Section 3 we introduce the preconditioner in 2D and present the 2D numerical results. Section 4 discusses the 3D case. Conclusions and some future directions are provided in Section 5.

2 1D Illustration

We use the PML [1, 4, 12] to simulate the Sommerfeld condition. The PML introduces the auxiliary functions

\[
\sigma(x) := \begin{cases} 
\frac{C}{\eta} \left( \frac{x - \eta}{\eta} \right)^2, & x \in [0, \eta), \\
0, & x \in [\eta, 1 - \eta], \\
\frac{C}{\eta} \left( \frac{x - 1 + \eta}{\eta} \right)^2, & x \in (1 - \eta, 1],
\end{cases}
\]

\[
s(x) := \left( 1 + \frac{i \sigma(x)}{\omega} \right)^{-1},
\]

where \( C \) is an appropriate positive constant independent of \( \omega \), and \( \eta \) is the PML width which is typically around one wavelength.
The Helmholtz equation with PML in 1D is
\[
\begin{cases}
(s(x) \frac{d}{dx})^2 + \frac{\omega^2}{c^2(x)} u(x) = f(x), \quad \forall x \in (0, 1), \\
u(0) = 0, \\
u(1) = 0.
\end{cases}
\]
We discretize the system with step size \( h = 1/(n+1) \), then \( n \) is the degree of freedom. With the standard central difference numerical scheme the discretized equation is
\[
\frac{s_i}{h} \left( \frac{s_i+1/2}{h} (u_{i+1} - u_i) - \frac{s_i-1/2}{h} (u_i - u_{i-1}) \right) + \frac{\omega^2}{c_i^2} u_i = f_i, \quad \forall 1 \leq i \leq n,
\]
where the subscript \( i \) means that the corresponding function is evaluated at \( x = ih \).

We denote Equation (2) as \( Au = f \), where \( u \) and \( f \) are the discrete array of the wave field and the force
\[
u := [u_1, \ldots, u_n]^T, \quad f := [f_1, \ldots, f_n]^T.
\]
In 1D, \( A \) is tridiagonal and Equation (2) can be solved without any difficulty. However, here we are aiming at an approach which can be generalized to higher dimensions so the rest of this section takes another point of view to solve (2) instead of exploiting the sparsity structure of \( A \) directly.

With the Green’s matrix \( G = A^{-1} \), \( u \) can be written as \( u = Gf \). Now let us divide the domain into \( m \) parts. We assume that \( \eta = \gamma h \) and \( n = 2\gamma + mb - 2 \) where \( \gamma \) and \( b \) are some small constants and \( m \) is comparable to \( n \), and we define
\[
X_1 := \{ih : 1 \leq i \leq \gamma + b - 1\}, \\
X_p := \{ih : \gamma + (p-1)b \leq i \leq \gamma + pb - 1\}, \quad p = 2, \ldots, m - 1, \\
X_m := \{ih : \gamma + (m-1)b \leq i \leq 2\gamma + mb - 2\},
\]
which means, \( X_1 \) is the leftmost part containing the left PML of the original problem and a small piece of grid with \( b \) points, \( X_m \) is the rightmost part containing the right PML and a grid of \( b \) points, and \( X_p, p = 2, \ldots, m - 1 \) are the middle parts each of which contains \( b \) points. \( u_p \) and \( f_p \) are defined as the restrictions of \( u \) and \( f \) on \( X_p \) for \( p = 1, \ldots, m \), respectively,
\[
u_1 := [u_1, \ldots, u_{\gamma+b-1}]^T, \\
u_p := [u_{\gamma+(p-1)b}, \ldots, u_{\gamma+pb-1}]^T, \quad p = 2, \ldots, m - 1, \\
u_m := [u_{\gamma+(m-1)b}, \ldots, u_{2\gamma+mb-2}]^T, \\
f_1 := [f_1, \ldots, f_{\gamma+b-1}]^T, \\
f_p := [f_{\gamma+(p-1)b}, \ldots, f_{\gamma+pb-1}]^T, \quad p = 2, \ldots, m - 1, \\
f_m := [f_{\gamma+(m-1)b}, \ldots, f_{2\gamma+mb-2}]^T.
\]
Then \( u = Gf \) can be written as
\[
\begin{bmatrix}
u_1 \\
u_2 \\
\vdots \\
u_m
\end{bmatrix} = 
\begin{bmatrix}
G_{1,1} & G_{1,2} & \ldots & G_{1,m} \\
G_{2,1} & G_{2,2} & \ldots & G_{2,m} \\
\vdots & \vdots & \ddots & \vdots \\
G_{m,1} & G_{m,2} & \ldots & G_{m,m}
\end{bmatrix}
\begin{bmatrix}
f_1 \\
f_2 \\
\vdots \\
f_m
\end{bmatrix}.
\]
By introducing \( u_{p,q} := G_{p,q} f_q \) for \( 1 \leq p, q \leq m \), one can write \( u_p = \sum_{q=1}^{m} u_{p,q} \). The physical meaning of \( u_{p,q} \) is the contribution of the force \( f_q \) defined on the grid \( X_q \) acting upon the grid \( X_p \). If we know the matrix \( G \), the computation of \( u_{p,q} \) can be carried out directly. However, computing \( G \), or even applying \( G \) to the vector \( f \), is computationally expensive. The additive sweeping method circumvent this difficulty by approximating the blocks of \( G \) sequentially and the idea works in higher dimensions. In what follows, we shall use \( \tilde{u}_{p,q} \) to denote the approximations of \( u_{p,q} \).

### 2.1 Approximating \( u_{p,q} \) with auxiliary PMLs

#### 2.1.1 Wave generated by \( f_1 \)

The components \( u_{p,1} \) for \( p = 1, \ldots, m \) can be regarded as a sequence of right-going waves generated by \( f_1 \). Note that the boundary condition of the system is the approximated Sommerfeld condition. If we assume that the reflection during the transmission of the wave is negligible, then, to approximate \( u_{1,1} \), we can simply put an artificial PML on the right of the grid \( X_1 \) to solve a much smaller problem, since the domain of interest here is only \( X_1 \) (see Figure 1(b)). To be precise, we define

\[
\sigma_1^M(x) := \begin{cases} 
C \left( \frac{x - \eta}{\eta} \right)^2, & x \in [0, \eta), \\
0, & x \in [\eta, \eta + (b - 1)h], \\
C \left( \frac{x - (\eta + (b - 1)h)}{\eta} \right)^2, & x \in (\eta + (b - 1)h, 2\eta + (b - 1)h], 
\end{cases}
\]

\[
s_1^M(x) := \left( 1 + \frac{i \sigma_1^M(x)}{\omega} \right)^{-1}.
\]

We consider a subproblem

\[
\left\{ \begin{aligned}
\left( s_1^M(x) \frac{d}{dx} \right)^2 + \frac{\omega^2}{c^2(x)} v(x) &= g(x), & \forall x \in (0, 2\eta + (b - 1)h), \\
v(0) &= 0, \\
v(2\eta + (b - 1)h) &= 0.
\end{aligned} \right.
\]

With the same discrete numerical scheme and step size \( h \), we have the corresponding discrete system \( H_1^M v = g \) on the extended grid

\( X_1^M := \{ x_i : 1 \leq i \leq 2\gamma + b - 2 \} \).

Then we can define an operator \( \tilde{G}_1^M : y \rightarrow z \), which is an approximation of \( G_{1,1} \), by the following:

1. Introduce a vector \( g \) defined on \( X_1^M \) by setting \( y \) to \( X_1 \) and zero everywhere else.
2. Solve \( H_1^M v = g \) on \( X_1^M \).
3. Set \( z \) as the restriction of \( v \) on \( X_1 \).

Then \( \tilde{u}_{1,1} \) can be set as

\( \tilde{u}_{1,1} := \tilde{G}_1^M f_1 \).
Once we have computed \( \tilde{u}_{1,1} \), we can use the right boundary value of \( \tilde{u}_{1,1} \) to compute \( \tilde{u}_{2,1} \) by introducing an auxiliary PML on the right of \( X_2 \) and solving the boundary value problem with the left boundary value at \( x = (\gamma + b - 1)h \) equal to the right boundary value of \( \tilde{u}_{1,1} \). The same process can be repeated to compute \( \tilde{u}_{p+1,1} \) by exploiting the right boundary value of \( \tilde{u}_{p,1} \) recursively for \( p = 2, \ldots, m - 1 \) (see Figure 1(c)). In the following context of this section, we introduce notations \( g^L, g^R \) for a vector array \( g = [g_1, \ldots, g_s]^T \) by

\[
g^L := g_1, \quad g^R := g_s,
\]

where \( g^L \) and \( g^R \) should be interpreted as the leftmost and the rightmost element of the array \( g \).

To formalize the definition of \( \tilde{u}_{p,1} \) for each \( p = 2, \ldots, m \), we extend \( X_p \) with an auxiliary PML on the right of it to form an extended grid

\[
X^R_p := \{ih : \gamma + (p - 1)b \leq i \leq 2\gamma + pb - 2\},
\]

where the superscript \( R \) means that we are introducing the auxiliary PML on the right. (Note that the grid \( X^R_m \) is \( X_m \) itself since \( X_m \) contains the original right PML region.)

For the PML on \( X^R_p \), we define

\[
\sigma_p^R(x) := \begin{cases} 0, & x \in [\eta + ((p - 1)b - 1)h, \eta + (pb - 1)h], \\ \frac{C}{\eta} \left( \frac{x - (\eta + (pb - 1)h)}{\eta} \right)^2, & x \in (\eta + (pb - 1)h, 2\eta + (pb - 1)h], \\ 1 + i \frac{\sigma_p^R(x)}{\omega}, & x \in (\eta + (pb - 1)h, 2\eta + (pb - 1)h]. \end{cases}
\]

\[
s_p^R := \left( 1 + \frac{\sigma_p^R(x)}{\omega} \right)^{-1},
\]

and consider the problem

\[
\left\{ \begin{align*}
\left( s_p^R(x) \frac{d}{dx} \right)^2 + \frac{\omega^2}{c^2(x)} v(x) &= 0, & \forall x \in (\eta + ((p - 1)b - 1)h, 2\eta + (pb - 1)h], \\
v(\eta + ((p - 1)b - 1)h) &= w, \\
v(2\eta + (pb - 1)h) &= 0.
\end{align*} \right.
\]

We define \( H^R_p v = g \) as the discretization of this problem on \( X^R_p \) where the right-hand side \( g \) is given by \( g := (-1/h^2)[w, 0, \ldots, 0]^T \) as a result of the central discretization. The subproblem \( H^R_p v = g \) for each \( p = 2, \ldots, m \) induces the approximation operator \( \tilde{G}^R_p : y \rightarrow z \) by the following procedure:

1. Set \( g = (-1/h^2)[y, 0, \ldots, 0]^T \).
2. Solve \( H^R_p v = g \) on \( X^R_p \).
3. Set \( z \) as the restriction of \( v \) on \( X_p \).

Then \( \tilde{u}_{p,1} \) can be defined recursively for \( p = 2, \ldots, m \) by

\[
\tilde{u}_{p,1} := \tilde{G}^R_p \tilde{u}_{p-1,1}.
\]

Note that, the operator \( \tilde{G}^R_p \) is not an approximation of the matrix block \( G_{p,1} \), since \( \tilde{G}^R_p \) maps the right boundary value of \( u_{p-1,1} \) to \( u_{p,1} \) while \( G_{p,1} \) maps \( f_1 \) to \( u_{p,1} \).
2.1.2 Wave generated by $f_m$

The components $u_{p,m}$ for $p = 1, \ldots, m$ can be regarded as a sequence of left-going waves generated by $f_m$. The method for approximating them is similar to what was done for $f_1$ (see Figure 1(d)).
More specifically, for $\tilde{u}_{m,m}$ we define

$$X^M_m := \{ih : (m-1)b + 1 \leq i \leq 2\gamma + mb - 2\},$$

$$\sigma^M_m(x) := \begin{cases} C \left( \frac{x - (1 - \eta - (b-1)h)}{\eta} \right)^2, & x \in [1 - 2\eta - (b-1)h, 1 - \eta - (b-1)h), \\ 0, & x \in [1 - \eta - (b-1)h, 1 - \eta], \\ C \left( \frac{x - (1 - \eta)}{\eta} \right)^2, & x \in (1 - \eta, 1], \end{cases}$$

$$s^M_m(x) := \left( 1 + i \frac{\sigma^M_m(x)}{\omega} \right)^{-1}.$$

We consider the continuous problem

$$\begin{cases} \left( (s^M_m(x) \frac{d}{dx})^2 + \frac{\omega^2}{c^2(x)} \right) v(x) = g(x), & \forall x \in (1 - 2\eta - (b-1)h, 1), \\ v(1 - 2\eta - (b-1)h) = 0, & v(1) = 0, \end{cases}$$

and define $H^M_m \mathbf{v} = \mathbf{g}$ as its discretization on $X^M_m$. The operator $\tilde{G}^M_m : \mathbf{y} \mapsto \mathbf{z}$ can be defined as:

1. Introduce a vector $\mathbf{g}$ defined on $X^M_m$ by setting $\mathbf{g}$ to $X_m$ and zero everywhere else.
2. Solve $H^M_m \mathbf{v} = \mathbf{g}$ on $X^M_m$.
3. Set $\mathbf{z}$ as the restriction of $\mathbf{v}$ on $X_m$.

Then

$$\tilde{u}_{m,m} := \tilde{G}^M_m \mathbf{f}_m.$$

For each $\tilde{u}_{p,m}, p = 1, \ldots, m - 1$, we define

$$X^L_p := \{x_i : (p-1)b + 1 \leq i \leq \gamma + pb - 1\},$$

$$\sigma^L_p(x) := \begin{cases} C \left( \frac{x - (\eta + (p-1)bh)}{\eta} \right)^2, & x \in [(p-1)bh, \eta + (p-1)bh), \\ 0, & x \in [\eta + (p-1)bh, \eta + pbh], \end{cases}$$

$$s^L_p(x) := \left( 1 + i \frac{\sigma^L_p(x)}{\omega} \right)^{-1},$$

and consider the continuous problem

$$\begin{cases} \left( (s^L_p(x) \frac{d}{dx})^2 + \frac{\omega^2}{c^2(x)} \right) v(x) = 0, & \forall x \in ((p-1)bh, \eta + pbh), \\ v((p-1)bh) = 0, & v(\eta + pbh) = w, \end{cases}$$

Let $H^L_p \mathbf{v} = \mathbf{g}$ be its discretization on $X^L_p$ with $\mathbf{g} := (-1/h^2)[0, \ldots, 0, w]^T$. We introduce the operator $\tilde{G}^L_p : \mathbf{y} \mapsto \mathbf{z}$ by:
1. Set \( g = (-1/h^2)[0, \ldots, 0, y]^T \).

2. Solve \( H_p^L v = g \) on \( X_p^L \).

3. Set \( z \) as the restriction of \( v \) on \( X_p \).

Then \( \tilde{u}_{p,m} \) can be defined recursively for \( p = m - 1, \ldots, 1 \) by

\[
\tilde{u}_{p,m} := \tilde{G}_p^L u_{p+1,m}^L.
\]

### 2.1.3 Wave generated by \( f_q \) for \( q = 2, \ldots, m - 1 \)

For each \( q \), the components \( u_{p,q} \) for \( p = 1, \ldots, m \) can be regarded as a sequence of left- and right-going waves generated by \( f_q \) (see Figure 1(a)). For \( \tilde{u}_{q,q} \), we introduce

\[
X_q^M := \{ x_i : (q - 1)b + 1 \leq i \leq 2\gamma + qb - 2 \},
\]

\[
\sigma_q^M(x) := \begin{cases} 
C \left( \frac{x - (\eta + (q - 1)bh)}{\eta} \right)^2, & x \in [(q - 1)bh, \eta + (q - 1)bh], \\
0, & x \in [\eta + (q - 1)bh, \eta + (qb - 1)h], \\
C \left( \frac{x - (\eta + (qb - 1)h)}{\eta} \right)^2, & x \in (\eta + (qb - 1)h, 2\eta + (qb - 1)h], 
\end{cases}
\]

\[
s_q^M(x) := \left( 1 + i\frac{\sigma_q^M(x)}{\omega} \right)^{-1},
\]

and define \( H_q^M v = g \) as the discrete problem of the continuous problem

\[
\begin{align*}
\left( (s_q^M(x) \frac{d}{dx})^2 + \frac{\omega^2}{c^2(x)} \right) v(x) &= g(x), \quad \forall x \in [(q - 1)bh, 2\eta + (qb - 1)h), \\
v((q - 1)bh) &= 0, \\
v(2\eta + (qb - 1)h) &= 0,
\end{align*}
\]

We introduce the operator \( \tilde{G}_q^M : y \rightarrow z \) as:

1. Introduce a vector \( g \) defined on \( X_q^M \) by setting \( y \) to \( X_q \) and zero everywhere else.

2. Solve \( H_q^M v = g \) on \( X_q^M \).

3. Set \( z \) as the restriction of \( v \) on \( X_q \).

Then

\[
\tilde{u}_{q,q} := \tilde{G}_q^M f_q.
\]

Following the above discussion, the remaining components \( \tilde{u}_{p,q} \) are defined recursively as

\[
\tilde{u}_{p,q} := \tilde{G}_p^R \tilde{u}_{p-1,q}^R, \quad \text{for } p = q + 1, \ldots, m,
\]

\[
\tilde{u}_{p,q} := \tilde{G}_p^L \tilde{u}_{p+1,q}^L, \quad \text{for } p = q - 1, \ldots, 1.
\]
2.2 Accumulating the boundary values

After all the above are done, an approximation of \( u_p \) is given by (see Figure 2(a))

\[
\tilde{u}_p := \sum_{q=1}^{m} \tilde{u}_{p,q}, \quad p = 1, \ldots, m.
\]

Figure 2: This figure shows how the boundary values are accumulated after each step.

In the algorithm described above, the computation of each component \( \tilde{u}_{p,q} \) requires a separate solution of a problem of form \( H_p^R v = g \) or \( H_p^L v = g \) (see Figure 2(b)). Since there are \( O(m^2) \) such components, the algorithm is computationally expensive. A key observation is that the computation
associated with each \( p \) can be combined in one single shot by exploiting the linearity of the equation. More precisely, we define
\[
\tilde{u}_{p,q_1:q_2} := \sum_{t=q_1}^{q_2} \tilde{u}_{p,t},
\]
which is the total contribution of the waves generated by \( f_{q_1}, \ldots, f_{q_2} \) restricted to the grid \( X_p \). The quantity \( \tilde{u}_{p,1:p-1} \), which is the total right-going wave generated by \( f_1, \ldots, f_{p-1} \) upon \( X_p \), can be computed sequentially for \( p = 2, \ldots, m \) without computing each component and then adding them together as we described above, as long as we accumulate the boundary values after each intermediate step. Specifically, we first compute \( \tilde{u}_{p} = \tilde{G}_p f \) for \( q = 1, \ldots, m \). This step has no difference with what we did above. Then, to compute \( \tilde{u}_{p,p-1} \) we carry out the following steps
\[
\tilde{u}_{p,1:p-1} = \tilde{G}_p \tilde{u}_{p-1,1:p-1}, \quad \tilde{u}_{p,p-1}^R = \tilde{u}_{p,1:p-1}^R + \tilde{u}_{p,p}^R, \quad \text{for } p = 2, \ldots, m.
\]
This means, when computing the total right-going wave on each subdomain, the boundary values of the previous right-going waves are added together before computing the current total wave, so that the the current one can be computed in one shot, eliminating the trouble of solving the subproblems for many times and adding the results together.

For the left going waves \( \tilde{u}_{p,p+1:m} \), a similar process gives rise to the recursive formula
\[
\tilde{u}_{p,p+1:m} = \tilde{G}_p^L \tilde{u}_{p+1,p+1:m}, \quad \tilde{u}_{p,m}^L = \tilde{u}_{p,m}^L + \tilde{u}_{p,p+1:m}, \quad \text{for } p = m - 1, \ldots, 1.
\]
Finally, each \( \tilde{u}_p \) can be computed by summing \( \tilde{u}_{p,1:p-1}, \tilde{u}_{p,p} \) and \( \tilde{u}_{p,p+1:m} \) together (for the leftmost and the rightmost one, \( \tilde{u}_1 \) and \( \tilde{u}_m \), only two terms need to be summed), i.e.,
\[
\begin{align*}
\tilde{u}_1 &= \tilde{u}_{1,1} + \tilde{u}_{1,2:m}, \\
\tilde{u}_p &= \tilde{u}_{p,1:p-1} + \tilde{u}_{p,p} + \tilde{u}_{p,p+1:m}, \quad p = 2, \ldots, m - 1, \\
\tilde{u}_m &= \tilde{u}_{m,1:m-1} + \tilde{u}_{m,m}.
\end{align*}
\]
In this algorithm, the approximation \( \tilde{u}_p \) on each small subdomain is divided into three parts. From a matrix point of view, this is analogous to splitting the block matrix \( G \) into its lower triangular part, diagonal part and upper triangular part, and then approximating each part as an operator and summing the results together (see Equation (3)). This is why we call it the additive sweeping method.

\[
\begin{bmatrix}
\tilde{u}_1 \\
\tilde{u}_2 \\
\vdots \\
\tilde{u}_m
\end{bmatrix} =
\begin{bmatrix}
\tilde{u}_{1,1} + \tilde{u}_{1,2:m} \\
\tilde{u}_{2,1} + \tilde{u}_{2,2} + \tilde{u}_{2,3:m} \\
\vdots \\
\tilde{u}_{m,1:m-1} + \tilde{u}_{m,m}
\end{bmatrix} =
\begin{bmatrix}
G_{1,1} & G_{1,2} & \cdots & G_{1,m} \\
G_{2,1} & G_{2,2} & \cdots & G_{2,m} \\
\vdots & \vdots & \ddots & \vdots \\
G_{m,1} & \cdots & G_{m,m-1} & G_{m,m}
\end{bmatrix}
\begin{bmatrix}
f_1 \\
f_2 \\
\vdots \\
f_m
\end{bmatrix}
\tag{3}
\]

When combined with standard iterative solver, the approximation algorithm serves as a pre-conditioner for Equation (2), and it can be easily generalized to higher dimensions as we shall see in the following sections.
3 Preconditioner in 2D

3.1 Algorithm

The domain of interest is $D = (0,1)^2$. We put PML on the two opposite sides of the boundary, $x_2 = 0$ and $x_2 = 1$, to illustrate the idea. The resulting equation is

$$
\begin{align*}
\left\{ \begin{array}{ll}
\left( \partial_x^2 + (s(x_2) \partial_x^2 + \frac{\omega^2}{C^2(x)}) \right) u(x) = f(x), & \forall x = (x_1, x_2) \in D, \\
u(x) = 0, & \forall x \in \partial D,
\end{array} \right.
\end{align*}
$$

We discretize $D$ with step size $h = 1/(n + 1)$ in each direction, which results the Cartesian grid

$$X := \{ (i_1 h, i_2 h) : 1 \leq i_1, i_2 \leq n \},$$

and the discrete equation

$$
\begin{align*}
&\frac{s_{i_2}^2}{h} \left( \frac{s_{i_2}^2 + 1/2}{h} (u_{i_1, i_2 + 1} - u_{i_1, i_2}) - \frac{s_{i_2}^2 - 1/2}{h} (u_{i_1, i_2} - u_{i_1, i_2 - 1}) \right) \\
&\quad + \frac{u_{i_1+1, i_2 - 2u_{i_1, i_2} + u_{i_1, i_2 - 1}}}{h^2} + \frac{\omega^2}{C_{i_1, i_2}} u_{i_1, i_2} = f_{i_1, i_2}, & \forall 1 \leq i_1, i_2 \leq n,
\end{align*}
$$

where the subscript $(i_1, i_2)$ means that the corresponding function is evaluated at $(i_1 h, i_2 h)$, and since $s(x_2)$ is a function of $x_2$ only, we omit the $i_1$ subscript. $u$ and $f$ are defined to be the column-major ordering of the discrete array $u$ and $f$ on the grid $X$

$$u := [u_{1,1}, \ldots, u_{n,1}, \ldots, u_{n,n}]^T, \quad f := [f_{1,1}, \ldots, f_{n,1}, \ldots, f_{n,n}]^T.$$

Now (4) can be written as $Au = f$.

We divide the domain into $m$ parts along the $x_2$ direction

$$X_1 := \{ (i_1 h, i_2 h) : 1 \leq i_1 \leq n, 1 \leq i_2 \leq \gamma + b - 1 \},$$

$$X_p := \{ (i_1 h, i_2 h) : 1 \leq i_1 \leq n, \gamma + (p-1)b \leq i_2 \leq \gamma + pb - 1 \}, \quad p = 2, \ldots, m - 1,$$

$$X_m := \{ (i_1 h, i_2 h) : 1 \leq i_1 \leq n, \gamma + (m-1)b \leq i_2 \leq 2\gamma + mb - 2 \},$$

and we define $u_p$ and $f_p$ as the column-major ordering restriction of $u$ and $f$ on $X_p$

$$u_1 := [u_{1,1}, \ldots, u_{n,1}, \ldots, u_{n,\gamma+b-1}]^T, \quad u_p := [u_{1,\gamma+(p-1)b}, \ldots, u_{n,\gamma+(p-1)b}, \ldots, u_{n,\gamma+mb-1}]^T, \quad p = 2, \ldots, m - 1,$$

$$u_m := [u_{1,\gamma+(m-1)b}, \ldots, u_{n,\gamma+(m-1)b}, \ldots, u_{n,2\gamma+mb-2}]^T,$$

$$f_1 := [f_{1,1}, \ldots, f_{n,1}, \ldots, f_{n,\gamma+b-1}]^T, \quad f_p := [f_{1,\gamma+(p-1)b}, \ldots, f_{n,\gamma+(p-1)b}, \ldots, f_{n,\gamma+mb-1}]^T, \quad p = 2, \ldots, m - 1,$$

$$f_m := [f_{1,\gamma+(m-1)b}, \ldots, f_{n,\gamma+(m-1)b}, \ldots, f_{n,2\gamma+mb-2}]^T,$$

then $u = Gf$ for $G = A^{-1}$ can be written as

$$
\begin{bmatrix}
u_1 \\
u_2 \\
\vdots \\
u_m
\end{bmatrix} =
\begin{bmatrix}
G_{1,1} & G_{1,2} & \cdots & G_{1,m} \\
G_{2,1} & G_{2,2} & \cdots & G_{2,m} \\
\vdots & \vdots & \ddots & \vdots \\
G_{m,1} & G_{m,2} & \cdots & G_{m,m}
\end{bmatrix}
\begin{bmatrix}
f_1 \\
f_2 \\
\vdots \\
f_m
\end{bmatrix}.
$$
**Auxiliary domains.** Following to the 1D case, the extended subdomains and the corresponding left and right boundaries are defined by

\[
D^M_q = [0, 1] \times [(q - 1)bh, 2\eta + (qb - 1)h], \quad q = 1, \ldots, m,
\]

\[
D^R_p = [0, 1] \times [\eta + ((p - 1)b - 1)h, 2\eta + (pb - 1)h], \quad p = 2, \ldots, m,
\]

\[
D^L_p = [0, 1] \times [(p - 1)bh, \eta + pbh], \quad p = 1, \ldots, m - 1,
\]

\[
\partial^L D^R_p = [0, 1] \times \{\eta + ((p - 1)b - 1)h\}, \quad p = 2, \ldots, m,
\]

\[
\partial^R D^L_p = [0, 1] \times \{\eta + pbh\}, \quad p = 1, \ldots, m - 1.
\]

The extended grid for these domains are

\[
X^M_q := \{(i_1 h, i_2 h) : 1 \leq i_1 \leq n, (q - 1)b + 1 \leq i_2 \leq 2\gamma + qb - 1\}, \quad q = 1, \ldots, m,
\]

\[
X^R_p := \{(i_1 h, i_2 h) : 1 \leq i_1 \leq n, \gamma + (p - 1)b \leq i_2 \leq 2\gamma + pb - 2\}, \quad p = 2, \ldots, m,
\]

\[
X^L_p := \{(i_1 h, i_2 h) : 1 \leq i_1 \leq n, (p - 1)b + 1 \leq i_2 \leq \gamma + pb - 1\}, \quad p = 1, \ldots, m - 1.
\]

**Auxiliary problems.** For \(q = 1, \ldots, m\), we define \(H^M_q v = g\) to be the discretization on \(X^M_q\) of the problem

\[
\begin{cases}
\left(\partial_1^2 + (s^M_q(x_2)\partial_2)^2 + \frac{\omega^2}{c^2(x)}\right)v(x) = g(x), & \forall x \in D^M_q, \\
v(x) = 0, & \forall x \in \partial D^M_q.
\end{cases}
\]

For \(p = 2, \ldots, m\), \(H^R_p v = g\) is the discretization on \(X^R_p\) of the problem

\[
\begin{cases}
\left(\partial_1^2 + (s^R_p(x_2)\partial_2)^2 + \frac{\omega^2}{c^2(x)}\right)v(x) = 0, & \forall x \in D^R_p, \\
v(x) = w(x_1), & \forall x \in \partial D^R_p \setminus \partial^L D^R_p, \\
v(x) = 0, & \forall x \in \partial D^R_p \setminus \partial^L D^R_p.
\end{cases}
\]

where \(g := (-1/h^2)[w_1, \ldots, w_n, 0, \ldots, 0]^T\). Finally, for \(p = 1, \ldots, m - 1\), \(H^L_p v = g\) is the discretization on \(X^L_p\) of the problem

\[
\begin{cases}
\left(\partial_1^2 + (s^L_p(x_2)\partial_2)^2 + \frac{\omega^2}{c^2(x)}\right)v(x) = 0, & \forall x \in D^L_p, \\
v(x) = w(x_1), & \forall x \in \partial D^L_p \setminus \partial^R D^L_p, \\
v(x) = 0, & \forall x \in \partial D^L_p \setminus \partial^R D^L_p.
\end{cases}
\]

where \(g := (-1/h^2)[0, \ldots, 0, w_1, \ldots, w_n]^T\).

**Auxiliary Green’s operators.** For \(q = 1, \ldots, m\), we define \(G^M_q : y \mapsto z\) to be the operator defined by the following operations:

1. Introduce a vector \(g\) defined on \(X^M_q\) by setting \(y\) to \(X_q\) and zero everywhere else.
2. Solve $H_q^M v = g$ on $X_q^M$.
3. Set $z$ as the restriction of $v$ on $X_q$.

For $p = 2, \ldots, m$, the operators $\tilde{G}_p^R : y \mapsto z$ is given by:

1. Set $g = (-1/h^2)[y^T, 0, \ldots, 0]^T$.
2. Solve $H_p^R v = g$ on $X_p^R$.
3. Set $z$ as the restriction of $v$ on $X_p$.

Finally, for $p = 1, \ldots, m - 1$, $\tilde{G}_p^R : y \mapsto z$ is defined as:

1. Set $g = (-1/h^2)[0, \ldots, 0, y^T]^T$.
2. Solve $H_p^L v = g$ on $X_p^L$.
3. Set $z$ as the restriction of $v$ on $X_p$.

**Putting together.** Similar to the previous section, we introduce the left boundary value $g^L$ and the right boundary value $g^R$ for a column-major ordering array $g = [g_{1,1}, \ldots, g_{s_1,1}, \ldots, g_{s_1,s_2}]^T$ induced from some grid with size $s_1 \times s_2$ by

$$g^L := [g_{1,1}, \ldots, g_{s_1,1}]^T, \quad g^R := [g_{1,s_2}, \ldots, g_{s_1,s_2}]^T.$$ 

Then the approximations for $u_{p,p} = 1, \ldots, m$, can be defined step by step as

$$u_{q,q} := \tilde{G}_q^M f_q, \quad q = 1, \ldots, m,$$

$$u_{p,1:p-1} := \tilde{G}_p^R u_{p-1,1:p-1}, \quad u_{p,1:p} := u_{p,1:p-1} + u_{p,p}, \quad \text{for } p = 2, \ldots, m,$$

$$u_{p,p+1:m} := \tilde{G}_p^L u_{p+1,p+1:m}, \quad u_{p,p+1:m} := u_{p,p} + u_{p,p+1:m}, \quad \text{for } p = m - 1, \ldots, 1,$$

$$u_1 := u_{1,1} + u_{1,2:m},$$

$$u_p := u_{p,1:p-1} + u_{p,p} + u_{p,p+1:m}, \quad p = 2, \ldots, m - 1,$$

$$u_m := u_{m,1:m-1} + u_{m,m}.$$ 

To solve the subproblems $H_q^M v = g$, $H_p^R v = g$ and $H_p^L v = g$, we notice that they are indeed quasi-1D problems since $\gamma$ and $b$ are some small constants. Therefore, for each one of them, we can reorder the system by grouping the elements along dimension 2 first and then dimension 1, which results a banded linear system that can be solved by the LU factorization efficiently. These factorization processes induce the factorizations for the operators $\tilde{G}_q^M$, $G_p^R$ and $\tilde{G}_p^L$ symbolically, which leads to our setup algorithm of the preconditioner in 2D as described in Algorithm 1 and the application algorithm as described in Algorithm 2.

To analyze the complexity, we note that, in the setup process, there are $O(n/b)$ subproblems, each of which is a quasi-1D problem with $O(\gamma + b)$ layers along the second dimension. Therefore, the setup cost of each subproblem by the LU factorization is $O(n(\gamma + b)^3)$ and the application cost is $O(n(\gamma + b)^2)$. So the total setup cost is $O(n^2(\gamma + b)^3/b)$. Besides, one needs to solve each subproblem once during the application process so the total application cost is $O(n^2(\gamma + b)^2/b)$.

There are some differences when implementing the method practically:
Algorithm 1 Construction of the 2D additive sweeping preconditioner of the Equation (4). Complexity = $O(n^2(b + \gamma)^3/b) = O(N(b + \gamma)^3/b)$.

for $q = 1, \ldots, m$ do  
  Construct the LU factorization of $H_q^M$, which defines $\tilde{G}_q^M$.  
end for  
for $p = 2, \ldots, m$ do  
  Construct the LU factorization of $H_p^R$, which defines $\tilde{G}_p^R$.  
end for  
for $p = 1, \ldots, m - 1$ do  
  Construct the LU factorization of $H_p^L$, which defines $\tilde{G}_p^L$.  
end for

Algorithm 2 Computation of $\tilde{u} \approx Gf$ using the preconditioner from Algorithm 1. Complexity = $O(n^2(b + \gamma)^2/b) = O(N(b + \gamma)^2/b)$.

for $q = 1, \ldots, m$ do  
  $u_{q,q} = \tilde{G}_q^M f_q$  
end for  
for $p = 2, \ldots, m$ do  
  $u_{p,1:p-1} = \tilde{G}_p^R u_{p-1,1:p-1}$  
  $u_{p,1:p} = u_{p,1:p-1} + u_{p,p}$  
end for  
for $p = m - 1, \ldots, 1$ do  
  $u_{p,p+1:m} = \tilde{G}_p^L u_{p+1,p+1:m}$  
  $u_{p,p:m} = u_{p,p} + u_{p,p+1:m}$  
end for  
$u_1 = u_{1,1} + u_{1,2:m}$  
for $p = 2, \ldots, m - 1$ do  
  $u_p = u_{p,1:p-1} + u_{p,p} + u_{p,p+1:m}$  
end for  
$u_m = u_{m,1:m-1} + u_{m,m}$

1. In the above setting, PMLs are put only on two opposite sides of the unit square for illustration purpose. In reality, PMLs can be put on other sides of the domain if needed. As long as there are two opposite sides with PML boundary condition, the method can be implemented.

2. The thickness of the auxiliary PMLs introduced in the interior part of the domain needs not to be the same with the thickness of the PML at the boundary. In fact, the thickness of the auxiliary PML is typically thinner in order to improve efficiency.

3. The widths of the subdomains are completely arbitrary and they need not to be the same. Practically, the widths can be chosen to be larger for subdomains where the velocity field varies heavily.

4. The symmetric version of the equation can be adopted to save memory and computational cost.
3.2 Numerical results

Here, we present some numerical results in 2D to illustrate the efficiency of the algorithm. The proposed method is implemented in MATLAB and the tests are performed on a 2.0 GHz computer with 256 GB memory. GMRES is used as the iterative solver with relative residual equal to $10^{-3}$ and restart value equal to 40. PMLs are put on all sides of the unit square. The velocity fields tested are given in Figure 3:

(a) A converging lens with a Gaussian profile at the center of the domain.
(b) A vertical waveguide with a Gaussian cross-section.
(c) A random velocity field.

![Figure 3: The three velocity fields tested in 2D.](image)

For each velocity field, two external forces are tested:

(a) A Gaussian point source centered at $(1/2, 1/8)$.
(b) A Gaussian wave packet with wavelength comparable to the typical wavelength of the domain. The packet centers at $(1/8, 1/8)$ and points to the direction $(1/\sqrt{2}, 1/\sqrt{2})$.

In these tests, each typical wavelength is discretized with 8 points. The width of the PML at the boundary and the one of the PMLs introduced in the interior parts of the domain are both $9h$, i.e., $\gamma = 9$. The number of layers in each interior subdomain is $b = 8$, the number of layers in the leftmost subdomain is $b + \gamma - 1 = 16$ and the one in the rightmost is $b + \gamma - 2 = 15$.

We vary the typical wave number $\omega/(2\pi)$ and test the behavior of the algorithm. The test results are presented in Tables 1, 2, and 3. $T_{\text{setup}}$ is the setup time of the algorithm in seconds. $T_{\text{solve}}$ is the total solve time in seconds and $N_{\text{iter}}$ is the iteration number. From these tests we see that the setup time scales like $O(N)$ as well as the solve time per iteration, which is consistent with the algorithm complexity analysis. The iteration number remains constant or grows at most logarithmically, which shows the efficiency of the preconditioner.
4 Preconditioner in 3D

4.1 Algorithm

In this section we briefly state the preconditioner in 3D case. The domain of interest is \( D = (0,1)^3 \). PMLs are put on two opposite faces of the unit cube, \( x_3 = 0 \) and \( x_3 = 1 \), which results the equation

\[
\begin{align*}
\left\{ \begin{array}{l}
\frac{\partial^2 u}{\partial x_1^2} + \frac{\partial^2 u}{\partial x_2^2} + \left(s(x_3)\frac{\partial^3 u}{\partial x_3^3} + \frac{\omega^2}{c^2(x_3)} \right) u(x) = f(x), \\
u(x) = 0,
\end{array} \right.
\forall x = (x_1, x_2, x_3) \in D, \\
\forall x \in \partial D,
\end{align*}
\]

Discretizing \( D \) with step size \( h = 1/(n+1) \) gives the grid

\[
X := \{(i_1h, i_2h, i_3h) : 1 \leq i_1, i_2, i_3 \leq n\},
\]

and the discrete equation

\[
\begin{align*}
\frac{s_{i_3}}{h} \left( \frac{s_{i_3+1/2}}{h} (u_{i_1,i_2,i_3+1} - u_{i_1,i_2,i_3}) - \frac{s_{i_3-1/2}}{h} (u_{i_1,i_2,i_3} - u_{i_1,i_2,i_3-1}) \right) \\
+ \frac{u_{i_1+1,i_2,i_3} - 2u_{i_1,i_2,i_3} + u_{i_1-1,i_2,i_3}}{h^2} \\
+ \frac{\omega^2}{c_{i_1,i_2,i_3}^2} u_{i_1,i_2,i_3} = f_{i_1,i_2,i_3},
\forall 1 \leq i_1, i_2 \leq n.
\end{align*}
\]
force (a)  
force (b)  

<table>
<thead>
<tr>
<th>ω/(2π)</th>
<th>N</th>
<th>T\text{setup}</th>
<th>N\text{iter}</th>
<th>T\text{solve}</th>
<th>N\text{iter}</th>
<th>T\text{solve}</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>127²</td>
<td>7.0834e-01</td>
<td>6</td>
<td>2.9189e-01</td>
<td>4</td>
<td>1.9408e-01</td>
</tr>
<tr>
<td>32</td>
<td>255²</td>
<td>3.2047e+00</td>
<td>8</td>
<td>6.3057e+00</td>
<td>4</td>
<td>7.9303e-01</td>
</tr>
<tr>
<td>64</td>
<td>511²</td>
<td>6.0951e+01</td>
<td>8</td>
<td>3.8908e+00</td>
<td>4</td>
<td>3.8908e+00</td>
</tr>
<tr>
<td>128</td>
<td>1023²</td>
<td>6.0951e+01</td>
<td>8</td>
<td>2.9097e+01</td>
<td>4</td>
<td>1.5287e+01</td>
</tr>
<tr>
<td>256</td>
<td>2047²</td>
<td>2.6025e+02</td>
<td>8</td>
<td>1.1105e+02</td>
<td>5</td>
<td>7.2544e+01</td>
</tr>
</tbody>
</table>

Table 2: Results for velocity field (b) in 2D. Solutions with ω/(2π) = 32 are presented.

\textbf{u} and \textbf{f} are defined as the column-major ordering of \textbf{u} and \textbf{f} on the grid \(X\)

\[
\mathbf{u} := [u_{1,1,1}, \ldots, u_{n,1,1}, \ldots, u_{n,n,n}]^T, \quad \mathbf{f} := [f_{1,1,1}, \ldots, f_{n,1,1}, \ldots, f_{n,n,n}]^T.
\]

\(X\) is divided into \(m\) parts along the \(x_3\) direction

\[
X_1 := \{(i_1 h, i_2 h, i_3 h) : 1 \leq i_1 \leq n, 1 \leq i_2 \leq n, 1 \leq i_3 \leq \gamma + b - 1\},
\]

\[
X_p := \{(i_1 h, i_2 h, i_3 h) : 1 \leq i_1 \leq n, 1 \leq i_2 \leq n, \gamma + (p - 1)b \leq i_3 \leq \gamma + pb - 1\}, \quad p = 2, \ldots, m - 1,
\]

\[
X_m := \{(i_1 h, i_2 h, i_3 h) : 1 \leq i_1 \leq n, 1 \leq i_2 \leq n, \gamma + (m - 1)b \leq i_3 \leq 2\gamma + mb - 2\}.
\]

\(\mathbf{u}_p\) and \(\mathbf{f}_p\) are the column-major ordering restrictions of \(\mathbf{u}\) and \(\mathbf{f}\) on \(X_p\)

\[
\mathbf{u}_1 := [u_{1,1,1}, \ldots, u_{n,1,1}, \ldots, u_{n,n,\gamma+b-1}]^T,
\]

\[
\mathbf{u}_p := [u_{1,1,\gamma+(p-1)b}, \ldots, u_{n,1,\gamma+(p-1)b}, \ldots, u_{n,n,\gamma+(p-1)b}, \ldots, u_{n,n,\gamma+pb-1}]^T, \quad p = 2, \ldots, m - 1,
\]

\[
\mathbf{u}_m := [u_{1,1,\gamma+(m-1)b}, \ldots, u_{n,1,\gamma+(m-1)b}, \ldots, u_{n,n,\gamma+(m-1)b}, \ldots, u_{n,n,2\gamma+mb-2}]^T,
\]

\[
\mathbf{f}_1 := [f_{1,1,1}, \ldots, f_{n,1,1}, \ldots, f_{n,n,\gamma+b-1}]^T,
\]

\[
\mathbf{f}_p := [f_{1,1,\gamma+(p-1)b}, \ldots, f_{n,1,\gamma+(p-1)b}, \ldots, f_{n,n,\gamma+(p-1)b}, \ldots, f_{n,n,\gamma+pb-1}]^T, \quad p = 2, \ldots, m - 1,
\]

\[
\mathbf{f}_m := [f_{1,1,\gamma+(m-1)b}, \ldots, f_{n,1,\gamma+(m-1)b}, \ldots, f_{n,n,\gamma+(m-1)b}, \ldots, f_{n,n,2\gamma+mb-2}]^T.
\]
force (a)  

force (b)  

velocity field (c)  

<table>
<thead>
<tr>
<th>$\omega/(2\pi)$</th>
<th>$N$</th>
<th>$T_{\text{setup}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>16</td>
<td>$127^2$</td>
<td>7.0495e−01</td>
<td>5</td>
<td>2.4056e−01</td>
<td>6</td>
<td>2.8347e−01</td>
</tr>
<tr>
<td>32</td>
<td>$255^2$</td>
<td>3.1760e00</td>
<td>5</td>
<td>1.0506e00</td>
<td>5</td>
<td>9.9551e−01</td>
</tr>
<tr>
<td>64</td>
<td>$511^2$</td>
<td>4.7083e00</td>
<td>6</td>
<td>4.7083e00</td>
<td>7</td>
<td>6.7852e00</td>
</tr>
<tr>
<td>128</td>
<td>$1023^2$</td>
<td>6.1217e01</td>
<td>6</td>
<td>1.8652e01</td>
<td>6</td>
<td>1.9792e01</td>
</tr>
<tr>
<td>256</td>
<td>$2047^2$</td>
<td>2.5762e02</td>
<td>8</td>
<td>1.1214e02</td>
<td>6</td>
<td>8.6936e01</td>
</tr>
</tbody>
</table>

Table 3: Results for velocity field (c) in 2D. Solutions with $\omega/(2\pi) = 32$ are presented.

**Auxiliary domains.** The extended subdomains, the extended grids, and the corresponding left and right boundaries are defined by

$D^M_q = [0, 1] \times [0, 1] \times [(q-1)bh, 2\eta + (qb-1)h]$, $q = 1, \ldots, m$,

$D^R_p = [0, 1] \times [0, 1] \times [\eta + ((p-1)b-1)h, 2\eta + (pb-1)h]$, $p = 2, \ldots, m$,

$D^L_p = [0, 1] \times [0, 1] \times [(p-1)bh, \eta + pbh]$, $p = 1, \ldots, m - 1$,

$\partial^L D^R_p = [0, 1] \times [0, 1] \times \{\eta + ((p-1)b-1)h\}$, $p = 2, \ldots, m$,

$\partial^R D^L_p = [0, 1] \times [0, 1] \times \{\eta + pbh\}$, $p = 1, \ldots, m - 1$.

$X^M_q := \{(i_1h, i_2h, i_3h) : 1 \leq i_1 \leq n, 1 \leq i_2 \leq n, (q-1)b + 1 \leq i_3 \leq 2\gamma + qb - 1\}$, $q = 1, \ldots, m$,

$X^R_p := \{(i_1h, i_2h, i_3h) : 1 \leq i_1 \leq n, 1 \leq i_2 \leq n, \gamma + (p-1)b \leq i_3 \leq 2\gamma + pb - 2\}$, $p = 2, \ldots, m$,

$X^L_p := \{(i_1h, i_2h, i_3h) : 1 \leq i_1 \leq n, 1 \leq i_2 \leq n, (p-1)b + 1 \leq i_3 \leq \gamma + pb - 1\}$, $p = 1, \ldots, m - 1$.

**Auxiliary problems.** For each $q = 1, \ldots, m$, $H_q^M \mathbf{v} = \mathbf{g}$ is defined as the discretization on $X^M_q$ of

$$
\begin{cases}
\n\left(\partial_1^2 + \partial_2^2 + (s_q^M(x_3)\partial_3)^2 + \frac{\omega^2}{c^2(x)}\right)v(x) = g(x), & \forall x \in D^M_q, \\
v(x) = 0, & \forall x \in \partial D^M_q,
\end{cases}
$$
For $p = 2, \ldots, m$, $H^R_p \mathbf{v} = \mathbf{g}$ is defined as the discretization on $X^R_p$ of
\[
\begin{cases}
\left( \partial_1^2 + \partial_2^2 + (s_p^R(x_3) \partial_3)^2 + \frac{\omega^2}{c^2(x)} \right) v(x) = 0, & \forall x \in D^R_p, \\
v(x) = w(x_1, x_2), & \forall x \in \partial^L D^R_p, \\
v(x) = 0, & \forall x \in D^R_p \setminus \partial^L D^R_p,
\end{cases}
\]
where $\mathbf{g} := (-1/h^2)[w_{1,1}, \ldots, w_{n,1}, \ldots, w_{n,n}, 0, \ldots, 0]^T$. Finally, for $p = 1, \ldots, m - 1$, $H^L_p \mathbf{v} = \mathbf{g}$ is the discretization on $X^L_p$ of
\[
\begin{cases}
\left( \partial_1^2 + \partial_2^2 + (s_p^L(x_3) \partial_3)^2 + \frac{\omega^2}{c^2(x)} \right) v(x) = 0, & \forall x \in D^L_p, \\
v(x) = w(x_1, x_2), & \forall x \in \partial^R D^L_p, \\
v(x) = 0, & \forall x \in D^L_p \setminus \partial^R D^L_p,
\end{cases}
\]
where $\mathbf{g} := (-1/h^2)[0, \ldots, 0, w_{1,1}, \ldots, w_{n,1}, \ldots, w_{n,n}]^T$.

**Auxiliary Green’s operators.** For $q = 1, \ldots, m$, $\tilde{G}^M_q : \mathbf{y} \mapsto \mathbf{z}$ is defined using the following operations:

1. Introduce a vector $\mathbf{g}$ defined on $X^M_q$ by setting $\mathbf{y}$ to $X_q$ and zero everywhere else.
2. Solve $H^M_q \mathbf{v} = \mathbf{g}$ on $X^M_q$.
3. Set $\mathbf{z}$ as the restriction of $\mathbf{v}$ on $X_q$.

For $p = 2, \ldots, m$, $\tilde{G}^R_p : \mathbf{y} \mapsto \mathbf{z}$ is given by:

1. Set $\mathbf{g} = (-1/h^2)[\mathbf{y}^T, 0, \ldots, 0]^T$.
2. Solve $H^R_p \mathbf{v} = \mathbf{g}$ on $X^R_p$.
3. Set $\mathbf{z}$ as the restriction of $\mathbf{v}$ on $X_p$.

Finally, for $p = 1, \ldots, m - 1$, the operators $\tilde{G}^L_p : \mathbf{y} \mapsto \mathbf{z}$ is introduced to be:

1. Set $\mathbf{g} = (-1/h^2)[0, \ldots, 0, \mathbf{y}^T]^T$.
2. Solve $H^L_p \mathbf{v} = \mathbf{g}$ on $X^L_p$.
3. Set $\mathbf{z}$ as the restriction of $\mathbf{v}$ on $X_p$.

**Putting together.** In the 3D case, $\mathbf{g}^L$ and $\mathbf{g}^R$ for the column-major ordering array $\mathbf{g} = [g_{1,1,1}, \ldots, g_{s_1,1,1}, \ldots, g_{s_1,s_2,1}, \ldots, g_{s_1,s_2,s_3}]^T$ induced from some 3D grid with size $s_1 \times s_2 \times s_3$ are given by
\[
\mathbf{g}^L := [g_{1,1,1}, \ldots, g_{s_1,1,1}, \ldots, g_{s_1,s_2,1}]^T, \quad \mathbf{g}^R := [g_{1,1,s_3}, \ldots, g_{s_1,1,s_3}, \ldots, g_{s_1,s_2,s_3}]^T.
\]
Algorithm 3 Construction of the 3D additive sweeping preconditioner of the system \((\mathbf{M})\). Complexity = \(O(n^4(b + \gamma)^3/b) = O(N^{4/3}(b + \gamma)^3/b)\).

for \(q = 1, \ldots, m\) do
  Construct the nested dissection factorization of \(H_q^M\), which defines \(\tilde{G}_q^M\).
end for
for \(p = 2, \ldots, m\) do
  Construct the the nested dissection factorization of \(H_p^R\), which defines \(\tilde{G}_p^R\).
end for
for \(p = 1, \ldots, m - 1\) do
  Construct the the nested dissection factorization of \(H_p^L\), which defines \(\tilde{G}_p^L\).
end for

Algorithm 4 Computation of \(\tilde{u} \approx \mathbf{Gf}\) using the preconditioner from Algorithm 3. Complexity = \(O(n^3 \log n (b + \gamma)^2/b) = O(N \log N (b + \gamma)^2/b)\).

for \(q = 1, \ldots, m\) do
  \(\tilde{u}_{q,q} = \tilde{G}_q^M f_q\)
end for
for \(p = 2, \ldots, m\) do
  \(\tilde{u}_{p,1:p-1} = \tilde{G}_p^R \tilde{u}_{p-1,1:p-1}\)
  \(\tilde{u}_{p,1:p} = \tilde{u}_{p,1:p-1} + \tilde{u}_{p,p}\)
end for
for \(p = m - 1, \ldots, 1\) do
  \(\tilde{u}_{p,p+1:m} = \tilde{G}_p^L \tilde{u}_{p+1,p+1:m}\)
  \(\tilde{u}_{p,p:m} = \tilde{u}_{p,p} + \tilde{u}_{p,p+1:m}\)
end for
\(\tilde{u}_1 = \tilde{u}_{1,1} + \tilde{u}_{1,2:m}\)
for \(p = 2, \ldots, m - 1\) do
  \(\tilde{u}_p = \tilde{u}_{p,1:p-1} + \tilde{u}_{p,p} + \tilde{u}_{p,p+1:m}\)
end for
\(\tilde{u}_m = \tilde{u}_{m,1:m-1} + \tilde{u}_{m,m}\)

The subproblems \(H_q^M \mathbf{v} = \mathbf{g}\), \(H_p^R \mathbf{v} = \mathbf{g}\) and \(H_p^L \mathbf{v} = \mathbf{g}\) are quasi-2D. To solve them, we group the elements along dimension 3 first, and then apply the nested dissection method\[11, 5\] to them, as in \[7\]. This gives the setup process of the 3D preconditioner in Algorithm 3 and the application process in Algorithm 4.

For the algorithm analysis, we notice that each quasi-2D subproblem has \(O(b + \gamma)\) layers along the third dimension. Therefore, the setup cost for each subproblem is \(O((\gamma + b)^3 n^3)\) and the application cost is \(O((\gamma + b)^2 n^2 \log n)\). Taking the total number of subproblems into account, the total setup cost for the 3D preconditioner is \(O(n^4(b + \gamma)^3/b)\) and the total application cost is \(O(n^3 \log n (b + \gamma)^2/b)\).
4.2 Numerical results

Here we present the numerical results in 3D. All the settings and notations are kept the same with Section 3.2 unless otherwise stated. The PMLs are put on all sides of the boundary and the symmetric version of the equation is adopted to save memory cost. The PML width is $\eta = 9h$ for the boundary and is $\eta_{\text{aux}} = 5h$ for the interior auxiliary ones. The number of layers in each subdomain is $b = 4$ for the interior ones, $b + \gamma - 1 = 12$ for the leftmost one and $b + \gamma - 2 = 11$ for the rightmost one.

The velocity fields tested are (see Figure 4):

(a) A converging lens with a Gaussian profile at the center of the domain.

(b) A vertical waveguide with a Gaussian cross-section.

(c) A random velocity field.

![Figure 4: The three velocity fields tested in 3D.](image)

The forces tested for each velocity field are:

(a) A Gaussian point source centered at $(1/2, 1/2, 1/4)$.

(b) A Gaussian wave packet with wavelength comparable to the typical wavelength of the domain. The packet centers at $(1/2, 1/4, 1/4)$ and points to the direction $(0, 1/\sqrt{2}, 1/\sqrt{2})$.

The results are given in Tables 4, 5 and 6. From these tests we see that the iteration number grows mildly as the problem size grows. We also notice that the setup cost scales even better than $O(N^{4/3})$, mainly because MATLAB performs dense linear algebra operations in a parallel way, which gives some extra advantages to the nested dissection algorithm as the problem size grows.

5 Conclusion

In this paper, we proposed a new additive sweeping preconditioner for the Helmholtz equation based on the PML. When combined with the standard GMRES solver, the iteration number grows mildly as the problem size grows. The novelty of this approach is that the unknowns are split in an additive way and the boundary values of the intermediate results are utilized directly. The
Table 4: Results for velocity field (a) in 3D. Solutions with $\omega/(2\pi) = 10$ at $x_1 = 0.5$ are presented.

<table>
<thead>
<tr>
<th>$\omega/(2\pi)$</th>
<th>$N$</th>
<th>$T_{\text{setup}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>39</td>
<td>2.3304e+01</td>
<td>3</td>
<td>2.9307e+00</td>
<td>4</td>
<td>3.7770e+00</td>
</tr>
<tr>
<td>10</td>
<td>79</td>
<td>3.2935e+02</td>
<td>3</td>
<td>3.6898e+01</td>
<td>4</td>
<td>4.6176e+01</td>
</tr>
<tr>
<td>20</td>
<td>159</td>
<td>4.2280e+03</td>
<td>4</td>
<td>4.3999e+02</td>
<td>4</td>
<td>4.6941e+02</td>
</tr>
</tbody>
</table>

Table 5: Results for velocity field (b) in 3D. Solutions with $\omega/(2\pi) = 10$ at $x_1 = 0.5$ are presented.

<table>
<thead>
<tr>
<th>$\omega/(2\pi)$</th>
<th>$N$</th>
<th>$T_{\text{setup}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>39</td>
<td>2.1315e+01</td>
<td>3</td>
<td>2.7740e+00</td>
<td>3</td>
<td>2.7718e+00</td>
</tr>
<tr>
<td>10</td>
<td>79</td>
<td>3.4256e+02</td>
<td>4</td>
<td>4.4286e+01</td>
<td>3</td>
<td>3.4500e+01</td>
</tr>
<tr>
<td>20</td>
<td>159</td>
<td>4.3167e+03</td>
<td>5</td>
<td>5.7845e+02</td>
<td>4</td>
<td>4.6462e+02</td>
</tr>
</tbody>
</table>

disadvantage is that, for each subdomains, three subproblems need to be built up, which is time consuming compared to [12] and [15]. However, the costly parts of the algorithm, i.e. the whole
Table 6: Results for velocity field (c) in 3D. Solutions with $\omega/(2\pi) = 10$ at $x_1 = 0.5$ are presented.

<table>
<thead>
<tr>
<th>$\omega/(2\pi)$</th>
<th>$N$</th>
<th>$T_{\text{setup}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
<th>$N_{\text{iter}}$</th>
<th>$T_{\text{solve}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>$39^3$</td>
<td>$2.1063e+01$</td>
<td>4</td>
<td>$3.8074e+00$</td>
<td>4</td>
<td>$3.7975e+00$</td>
</tr>
<tr>
<td>10</td>
<td>$79^3$</td>
<td>$3.4735e+02$</td>
<td>4</td>
<td>$4.4550e+01$</td>
<td>4</td>
<td>$4.5039e+01$</td>
</tr>
<tr>
<td>20</td>
<td>$159^2$</td>
<td>$4.3391e+03$</td>
<td>4</td>
<td>$4.4361e+02$</td>
<td>5</td>
<td>$5.8090e+02$</td>
</tr>
</tbody>
</table>

setup process and the solve processes of the subproblems $H_q^M v = g$, can be done in parallel. The only parts that must be implemented sequentially are the accumulations of the left-going and right-going waves, where only the solve processes of the subproblems $H_q^L v = g$ and $H_q^R v = g$ are involved, which are the cheapest parts of the algorithm. Besides, we think that the whole approximation process is simple and structurally clear from a physics point of view and the idea might be easy to be generalized to other equations.

There are also some other directions to make potential improvements. First, other numerical schemes of the equation and other approximations of the Sommerfeld radiation condition can be used to develop more efficient versions of this additive preconditioner. Second, the parallel version of the nested dissection algorithm can be combined to solve large scale problems. Last, in the 3D case, the quasi-2D subproblems can be solved recursively by sweeping along the $x_2$ direction with the same technique, which reduces the theoretical setup cost to $O(N)$ and the application cost to $O(N)$. However, compared to \cite{7}, the coefficient of the complexity in this new method is larger, so it is not clear whether or not the recursive approach will be more efficient practically. Nevertheless, it is of great theoretical interest to look into it.

References


