DIRECTIONAL PRECONDITIONER FOR 2D HIGH FREQUENCY OBSTACLE SCATTERING

LEXING YING†

Abstract. The boundary integral method is an efficient approach for solving time-harmonic obstacle scattering problems from bounded scatterers. This paper presents the directional preconditioner for the linear systems of the boundary integral method in two dimensions. This new preconditioner builds a data-sparse approximation of the integral operator, transforms it into a sparse linear system, and computes an approximate inverse with efficient sparse linear algebra algorithms. This preconditioner is efficient and results in small and almost frequency-independent iteration counts for nonresonant scatterers when combined with standard iterative solvers. Numerical results are provided to demonstrate the effectiveness of the new preconditioner.

Key words. boundary integral method, scattering, high-frequency waves, preconditioner, low-rank approximation, sparse linear algebra

AMS subject classifications. 65N38, 65R20, 78A45

DOI. 10.1137/140985135

1. Introduction. This paper is concerned with solving the time-harmonic acoustic obstacle scattering problems in two dimensions. Suppose that \( \Omega \subset \mathbb{R}^2 \) is a bounded scatterer with smooth boundary \( \partial \Omega \). In order to normalize the problem size, we assume that both the diameter and the boundary length of \( \Omega \) are of order 1. Let \( \omega \) be the time frequency and \( e^{i\omega t}u_I(x) \) be the time-harmonic incident wave. In the sound-soft scattering problem, the scattered field \( u(x) \) satisfies the Helmholtz equation with the Dirichlet boundary condition

\[
\Delta u(x) + \omega^2 u(x) = 0, \quad x \in \mathbb{R}^2 \setminus \Omega,
\]

\[
u(x) = -u_I(x), \quad x \in \partial \Omega.
\]

In the sound-hard scattering problem, the scattered field \( u(x) \) satisfies the Helmholtz equation with the Neumann boundary condition

\[
\Delta u(x) + \omega^2 u(x) = 0, \quad x \in \mathbb{R}^2 \setminus \Omega,
\]

\[
\frac{\partial u(x)}{\partial n(x)} = -\frac{\partial u_I(x)}{\partial n(x)}, \quad x \in \partial \Omega.
\]

In both cases, \( u(x) \) satisfies the Sommerfeld radiation condition

\[
\lim_{r \to \infty} r^{1/2} \left( \frac{\partial u}{\partial r} - i\omega u \right) = 0.
\]

An effective way to solve these problems is the boundary integral method, which relies on the free space Green’s function

\[
G(x, y) = \frac{i}{4} H_0^1(\omega |x - y|)
\]
of the Helmholtz equation. For the sound-soft scattering, the combined field integral equation (CFIE) [13, 28] looks for a surface density \( q(x) \) on the boundary such that for each \( x \in \partial \Omega \),

\[
(1) \quad \frac{1}{2} q(x) + \int_{\partial \Omega} \frac{\partial G(x, y)}{\partial n(y)} q(y)dy - i\eta \int_{\partial \Omega} G(x, y) q(y)dy = -u_I(x).
\]

Once \( q(x) \) is computed, the scattered field \( u(x) \) can be evaluated through a boundary integral over \( \partial \Omega \). For the sound-hard scattering, one again looks for \( q(x) \) on the boundary such that for each \( x \in \partial \Omega \),

\[
(2) \quad \frac{1}{2} q(x) - \int_{\partial \Omega} \frac{\partial G(x, y)}{\partial n(x)} q(y)dy + \frac{1}{i\eta} \int_{\partial \Omega} \frac{\partial^2 G(x, y)}{\partial n(x) \partial n(y)} q(y)dy = -\frac{\partial u_I(x)}{\partial n},
\]

with \( \oint \) being the principal value integral. In both cases, \( \eta \) is typically chosen to be of order \( O(\omega) \) (see [23], for example). We refer the reader to [13, 28] for derivations and discussions of these integral equations.

By introducing the operators

\[
(Sq)(x) = \int_{\partial \Omega} G(x, y) q(y)dy,
\]

\[
(Dq)(x) = \int_{\partial \Omega} \frac{\partial G(x, y)}{\partial n(y)} q(y)dy,
\]

\[
(D'q)(x) = \int_{\partial \Omega} \frac{\partial G(x, y)}{\partial n(x)} q(y)dy,
\]

\[
(Nq)(x) = \oint \frac{\partial^2 G(x, y)}{\partial n(x) \partial n(y)} q(y)dy,
\]

we can rewrite (1) and (2) in operator forms:

\[
(3) \quad \left( \frac{1}{2} I + D - i\eta S \right) q = -u_I,
\]

\[
(4) \quad \left( \frac{1}{2} I + D' - \frac{1}{i\eta} N \right) q = -\frac{\partial u_I}{\partial n}.
\]

Standard approaches for discretizing these boundary integral equations include the Nyström method, the Galerkin method, and the collocation method [13, 24]. To simplify the presentation, we assume that the Nyström method is used. The discussion remains similar for the other approaches as long as the basis functions employed are local. A typical discretization of these integral equations requires at least a few quadrature points per wavelength. Since both the diameter and the boundary length of \( \Omega \) are of order 1, this implies that the boundary is discretized with a set \( P \) of \( n = \Theta(\omega) \) points. Here the notation \( f = \Theta(g) \) means that \( cf \leq g \leq c'f \) for some uniform constants \( c \) and \( c' \). For the resulting linear systems, we continue to use \( S, D, D', \) and \( N \) to denote the discrete matrices associated with these operators. Similarly, \( q, u_I, \) and \( \partial u_I/\partial n \) are reused to denote the discrete versions of \( q(x), u_I(x), \) and \( \partial u_I/\partial n(x) \) sampled at the quadrature points. With this slight abuse of notation, the discrete linear systems take the same form as (3) and (4).

There has been a lot of work devoted to the fast solution of these linear systems. Since the system is dense, the standard direct solvers such as the LU factorization
take $O(n^3) = O(\omega^3)$ steps, which is prohibitively expensive when $\omega$ is large. Recently, several linear-complexity approaches based on recursive interpolative decomposition have been proposed by [14, 26, 21] for boundary integral equations with nonoscillatory kernels. However, for high-frequency scattering, the kernel is oscillatory (i.e., $\omega = \Theta(n)$), and the computational complexity of these approaches is still cubic in $n$. The only exception is the algorithm in [27] for quasi-1D domains, where the boundary integral equation essentially reduces to the 1D case and the complexity scales linearly in $\omega$.

For this reason, iterative methods such as GMRES and TFQMR [19, 30, 29] are the main tools for solving these problems. However, though the CFIEs have much improved conditioning properties compared to other integral formulations, the number of iterations can still grow quickly with $\omega$. Therefore, there is a need for preconditioning the CFIEs for high-frequency problems.

Over the past twenty years, there has been a significant amount of research devoted to this task. A few proposed algorithms suggest improving the conditioning property via modifying the standard CFIE formulation. For example, one idea is to replace the $i\eta$ term in (3) with better approximations of the Dirichlet-to-Neumann (DtN) operator, and the $1/(i\eta)$ term in (4) with better approximations of the Neumann-to-Dirichlet (NtD) operator [1, 4, 5, 9]. Typically, these new approximations are derived from leading order terms of the pseudodifferential symbols of the DtN and NtD operators.

A second approach is to precondition the integral equation analytically. Most methods guided by this approach consider the electric field integral equation for electromagnetic scattering and follow the Calderon relationship [31, 11, 12, 2, 3]. The resulting integral equations are of Fredholm second kind with good conditioning properties. However, the number of matrix vector multiplications per iteration is doubled.

There has also been a lot of work on sparsifying the integral operators using special basis functions, such as local cosine bases [6, 8], optimized wavelet packets [17, 18, 20, 22], etc. The resulting sparse representations typically have $O(n^{4/3})$ nonzero entries. Recently in [16], an approach using the wave atom transform [15] generated a sparse representation with $O(n \log n)$ nonzero entries. In [10], Canning constructed a sparse approximation with $O(n)$ nonzero entries via locally mollified exponential functions. However, when good accuracy is required, most of these methods access all entries of the integral operator, thus requiring an $O(n^2)$ precomputation cost for assembling the whole matrix.

In this paper, we propose a new method for preconditioning the CFIEs by incorporating the ideas from sparse representation. This approach builds a data-sparse representation of the boundary integral operator, transforms it into a sparse linear system, and computes an approximate inverse with efficient sparse linear algebra algorithms. This preconditioner is highly efficient to construct and to apply. It results in small and almost frequency-independent iteration counts for nonresonant scatterers when combined with standard iterative solvers. The rest of the paper is organized as follows. Section 2 describes the algorithm. In section 3, we present the results. Future work and open questions are discussed in section 4.

2. Algorithm. For a fixed frequency $\omega$, the wavelength $\lambda$ is equal to $2\pi/\omega$. We assume that the scatterer boundary $\partial \Omega$ is $C^2$ with uniformly bounded curvature and that both the diameter and the boundary length of $\Omega$ are $\Theta(1)$. To simplify the discussion, the length $L$ of $\partial \Omega$ is assumed to be equal to $4^q \lambda$, where $q$ is a positive integer. The actual number $4^q$ is not essential but it makes the presentation simple.
Combining this with $L = \Theta(1)$ implies that $\omega = O(4^q)$.

Suppose that $\rho : \partial \Omega \to [0, L]$ is the arclength parameterization of the boundary and that the boundary is sampled uniformly with $n = 4^q p$ discretization points for some $p = \Theta(1)$, i.e., $p$ points per wavelength $\lambda$.

2.1. Data-sparse approximation. We start by decomposing the boundary into sufficiently planar segments. Let $\kappa$ be the maximum absolute value of the curvature on $\partial \Omega$, and $\ell$ be the largest integer such that $2^{2\ell} \leq \frac{2q}{\sqrt{\kappa}}$. The boundary is partitioned into $2^{2\ell}$ segments, each of which is of length $h := 2^{2\ell} \lambda$ and has $2^{2\ell} p$ points. We denote this set of segments by $G = \{P_1, \ldots, P_m\}$, where $m = 2^{2\ell}$ and the segments $P_i$ are ordered according to their positions on the boundary.

In this paper, we focus on the setting of large $\omega$ and fixed geometry $\Omega$. As $\omega = O(4^q)$ grows, the segment length $h$ scales as $O(2^{2\ell} \lambda) = O(1/\sqrt{\omega})$, and the number of segments $m$ grows like $O(2^q) = O(\sqrt{\omega})$. Notice that since the boundary is assumed to be $C^2$, all segments in $G$ are almost-planar for sufficiently large $\omega$.

The discussion here shall treat the sound-soft case (3) and the sound-hard case (4) in the same way, since their kernels share the same oscillatory pattern. Therefore, it is convenient to use the general form

$$M q = f$$

for both (3) and (4). Based on how the segments are generated, each $P_j$ is of length $2^{2\ell} \lambda$ and contains $2^{2\ell} p$ equally spaced discretization points. After ordering the unknowns according to the ordering of $P_j$, the matrix $M$ can be written in the following block form:

$$M = \begin{bmatrix} M_{11} & \cdots & M_{1m} \\ \vdots & \ddots & \vdots \\ M_{m1} & \cdots & M_{mm} \end{bmatrix},$$

where each $M_{ij}$ is of size $2^{2\ell} p \times 2^{2\ell} p$. The next step is to find data-sparse approximations for the blocks $M_{ij}$.

2.1.1. Diagonal blocks. Let us first consider a diagonal block $M_{jj}$, which represents the interaction between $P_j$ and itself. Since $P_j$ is of length $2^{2\ell} \lambda$, the difference between the geometry of $P_j$ and its linear approximation is bounded by

$$\frac{1}{2} (2^{2\ell} \lambda)^2 \kappa \lesssim \left( \frac{2^{2\ell} \lambda}{\sqrt{\kappa}} \right)^2 \kappa = O(\lambda).$$

Therefore, one can treat $P_j$ approximately as planar. Based on this observation, let us define $B_j$ to be the matrix obtained by restricting the integral operator to a straight segment of length $2^{2\ell} \lambda$ with $2^{2\ell} p$ equally spaced quadrature points. The fact that $P_j$ is almost planar implies that

$$M_{jj} \approx B_j.$$ 

Noticing that $B_j$ depends only on the integer $\ell$ and that there are only a few choices for $\ell$, we can clearly precompute these matrices. Going through all $P_i$ gives the following approximation $B$ to the block-diagonal part of $M$:

$$B = \begin{bmatrix} B_1 \\ \vdots \\ B_m \end{bmatrix}.$$
2.1.2. Off-diagonal blocks. Next we consider the off-diagonal blocks \( M_{ij} \) with \( i \neq j \). We define

- \( c_i \) and \( c_j \) to be the centers of segments \( P_i \) and \( P_j \),
- \( t_i \) and \( t_j \) to be the unit tangent directions of \( \partial \Omega \) at the centers of \( P_i \) and \( P_j \), and
- \( a_{ij} \) to be the unit vector from \( c_j \) to \( c_i \), i.e., \((c_i - c_j)/|c_i - c_j|\) (see Figure 1 for an illustration).

The main challenge here is that the high-frequency Helmholtz kernel \( G(x,y) \) is highly oscillatory. Since the oscillation of \( G(x,y) \) and its derivatives come from the term \( \exp(i\omega|x-y|) \), it is instructive to focus on \( \exp(i\omega|x-y|) \) for the moment. In the following discussion, the symbol \( \sim \) is used to denote an approximation up to a nonoscillatory multiplicative term.

The first observation is that

\[
\exp(i\omega|x-y|) \sim \exp(i\omega a_{ij} \cdot (x-y))
\]

for \( x \in P_i \) and \( y \in P_j \). To see this, we assume without loss of generality that \( a_{ij} = (1,0)^t \) by rotating the coordinates accordingly and placing the origin on \( \partial \Omega \) between \( P_i \) and \( P_j \) (see Figure 1). In the new coordinate system, we have

\[
\exp(i\omega|x-y|) = \exp(i\omega(x_1 - y_1)) \exp(i\omega(|x-y| - (x_1 - y_1)))
\]

\[
= \exp(i\omega(x_1 - y_1)) \exp(i\omega(x_1 - y_1) \left( \sqrt{1 + \frac{|x_2 - y_2|^2}{|x_1 - y_1|^2}} - 1 \right)),
\]

where \( x = (x_1, x_2)^t \) and \( y = (y_1, y_2)^t \). Up to a constant factor, the phase of the last term can be bounded by \( \omega \frac{|x_2 - y_2|^2}{|x_1 - y_1|^2} \), which can be estimated as follows:

- When the distance between the centers of \( P_i \) and \( P_j \) is \( \Theta(1) \), we estimate this by

\[
\omega \frac{|x_2 - y_2|^2}{|x_1 - y_1|^2} \lesssim \omega(2^t \lambda)^2 = \Theta(1)
\]

since \( |x_1 - y_1| = \Theta(1) \) and \( |x_2 - y_2| \) is bounded by the segment length \( 2^t \lambda \).

- When the distance between the centers of \( P_i \) and \( P_j \) is \( o(1) \), we use the second order Taylor approximation based at the origin in the rotated frame (see Figure 1). Let \( P_i \) be the \( k \)th segment to the left of the origin, and \( P_j \) be the \( k \)th segment to the right of the origin. Then both \( x_1 \) and \( y_1 \) are of order
Multiplying this with \( \exp(\kappa x) \) in (8) and (9) as follows. Noticing that \( \omega \) is bounded by a \( \Theta(1) \) constant, and hence (6) is valid up to a nonoscillatory multiplicative term. Next, we rewrite
\[
\exp(i\omega |x-y|) \sim \exp(i\omega a_{ij} \cdot (x-y)) = \exp(i\omega a_{ij} \cdot ((x-c_i) + (c_i - c_j) + (c_j - y))) \tag{7}
\]
To approximate the first term in (7), we perform a Taylor expansion for \( \rho^{-1}(t) \) near \( t = \rho(c_i) \) and evaluate it at \( \rho(x) \):
\[
|\rho^{-1}(\rho(x)) - \rho^{-1}(\rho(c_i)) + t_i(\rho(x) - \rho(c_i))| \leq \frac{1}{2} |\rho(x) - \rho(c_i)|^2 \kappa \leq \frac{1}{2} (2^q \lambda)^2 \kappa \leq \frac{1}{2} \left( \frac{2^q \lambda}{\sqrt{\kappa}} \right)^2 \kappa = O(\lambda),
\]
where \( \kappa \) is the maximum absolute value of the curvature in \( \partial \Omega \). This is equivalent to
\[
(x-c_i) = (\rho(x) - \rho(c_i)) \cdot t_i + O(\lambda).
\]
Multiplying this with \( i\omega a_{ij} \) and taking the exponential gives the approximation
\[
\exp(i\omega a_{ij} \cdot (x-c_i)) \sim \exp(i(\omega a_{ij} \cdot t_i)(\rho(x) - \rho(c_i))). \tag{8}
\]
The same argument works for \( (y-c_j) \) and gives
\[
(y-c_j) = (\rho(y) - \rho(c_j)) \cdot t_j + O(\lambda)
\]
and
\[
\exp(-i\omega a_{ij} \cdot (y-c_j)) \sim \exp(i(-\omega a_{ij} \cdot t_j)(\rho(y) - \rho(c_j))). \tag{9}
\]
We can now further approximate the phase function of the complex exponentials in (8) and (9) as follows. Noticing that \( \omega a_{ij} \cdot t_i \in [-\omega, \omega] \), we partition the interval \([-\omega, \omega]\) into \( 2^{k+1} \) equally spaced subintervals with a set \( K \) of \( 2^{k+1} + 1 \) gridpoints. We define \( [k] \) to be the value of rounding \( k \) to the nearest gridpoint in \( K \). Then for each \( x \in P_i \),
\[
(\omega a_{ij} \cdot t_i - [\omega a_{ij} \cdot t_i])(\rho(x) - \rho(c_i)) \leq \frac{2\omega}{2^{k+1}} \cdot \frac{1}{2} \cdot \frac{2^q \lambda}{2} = \frac{2\pi}{4} = O(1).
\]
Thus, replacing the phase \( \omega a_{ij} \cdot t_i \) with \( [\omega a_{ij} \cdot t_i] \) in (8) introduces an extra nonoscillatory term,
\[
\exp(i\omega a_{ij} \cdot (x-c_i)) \sim \exp(i[\omega a_{ij} \cdot t_i](\rho(x) - \rho(c_i))). \tag{10}
\]
Similarly, we have for \( y \in P_j \) that
\[
-\omega a_{ij} \cdot t_j + [\omega a_{ij} \cdot t_j](\rho(y) - \rho(c_j)) \leq \frac{2\omega}{2\pi + 1} \cdot \frac{1}{2} \cdot \frac{2\lambda}{2} = \frac{2\pi}{4} = O(1).
\]

This change of the phase function also introduces an extra nonoscillatory term,
\[
\exp(-i\omega a_{ij} \cdot (y - c_j)) \sim \exp(-i[\omega a_{ij} \cdot t_j](\rho(y) - \rho(c_j))).
\]

By introducing
\[
\begin{align*}
  k_{ij}^l &= [\omega a_{ij} t_i], \\
  k_{ij}^r &= [\omega a_{ij} t_j], \\
  U_i(x, k) &= \exp(ik(\rho(x) - \rho(c_i))), \\
  U_j(y, k) &= \exp(ik(\rho(y) - \rho(c_j))),
\end{align*}
\]
and putting (10) and (11) into (7), we have the following approximation:
\[
\exp(i\omega|x - y|) \sim U_i(x, k_{ij}^l) \cdot \exp(i\omega(c_i - c_j)a_{ij}) \cdot U_j(y, k_{ij}^r)
\]
for all \( x \in P_i \) and \( y \in P_j \).

Since the kernel \( G(x, y) \) and its derivatives have the same oscillation pattern as \( \exp(i\omega|x - y|) \), \( U_i(x, k_{ij}^l) \) and \( U_j(y, k_{ij}^r) \) also capture the oscillations of \( G(x, y) \) for \( x \in P_i \) and \( y \in P_j \). Therefore, repeating the same argument allows us to write \( M_{ij} \) as
\[
M_{ij}(x, y) = U_i(x, k_{ij}^l) \cdot \tilde{M}_{ij}(x, y) \cdot U_j(y, k_{ij}^r)
\]
for \( x \in P_i \) and \( y \in P_j \), where the term \( \tilde{M}_{ij}(x, y) \) defined through this representation is nonoscillatory. Being nonoscillatory, \( \tilde{M}_{ij}(x, y) \) can be approximated with Chebyshev interpolation. To make this precise, we define
- \( R_i \) to be the Chebyshev grid of a constant size \( m_c \) in \( P_i \),
- \( P_j \) to be the Chebyshev grid of a constant size \( m_c \) in \( P_j \),
- \( I_i \) to be the Chebyshev interpolation operator, with entries given by \( I_i(x, b) \) for \( x \in P_i \) and \( b \in R_i \), and
- \( I_j \) to be the Chebyshev interpolation operator, with entries given by \( I_j(y, b) \) for \( y \in P_j \) and \( b \in R_j \).

This resulting Chebyshev approximation then reads
\[
\tilde{M}_{ij} \approx I_i \tilde{M}_{ij}(R_i, R_j) I_j^T.
\]

Putting this together with (13) and using matrix form gives
\[
(14) \quad M_{ij} \approx \text{diag}(U_i(:, k_{ij}^l)) \, I_i \tilde{M}_{ij}(R_i, R_j) \, I_j \, \text{diag}(U_j(:, k_{ij}^r)).
\]

For the data-sparse representation and the preconditioner, we need an aggressive rank-1 approximation for \( M_{ij} \) of the form
\[
(15) \quad M_{ij} \approx U_i(:, k_{ij}^l) \, e_{ij} \, U_j(:, k_{ij}^r)^*.
\]

To do that, we use the approximation (14) and rewrite the last formula as
\[
\text{diag}(U_i(:, k_{ij}^l)) \, w \, e_{ij} \, w^T \, \text{diag}(U_j(:, k_{ij}^r)),
\]
where \( w \) is defined to be the all-ones vector of length \( 2^\ell p \) and \( e_{ij} \) is a constant to be determined. To determine \( e_{ij} \), we can solve a least squares problem,

\[
e_{ij} = \arg\min_e \| I_i \tilde{M}_{ij}(R_i, R_j) I_j^t - w \circ w^t \|^2.
\]

The solution is

\[
e_{ij} = w^\dagger I_i \tilde{M}_{ij}(R_i, R_j) I_j^t (w^t)^\dagger,
\]

where \((\cdot)^\dagger\) stands for the pseudoinverse. Notice that \( w^\dagger I_i \) and \( I_j^t (w^t)^\dagger \) depend only on \( \ell \). Therefore, they can be precomputed, and the remaining cost of computing \( e_{ij} \) is equal to \( O(m^2) \).

Computing (15) for all pairs \((i, j)\) with \( i \neq j \) yields the following approximation for the off-diagonal part of \( M \):

\[
UEU^*.
\]

Here

\[
U = \begin{bmatrix}
U_1 \\
\vdots \\
U_m
\end{bmatrix},
\]

where each \( U_i \) is a matrix of size \( 2^\ell p \times (2^\ell+1) \) given by (12). The \( E \) matrix also has an \( m \times m \) block form,

\[
E = \begin{bmatrix}
E_{11} & \ldots & E_{1m} \\
\vdots & \ddots & \vdots \\
E_{m1} & \ldots & E_{mm}
\end{bmatrix},
\]

where \( E_{ij} \) is a matrix of size \( (2^\ell+1) \times (2^\ell+1) \) with rows and columns indexed by \( K_i \) and \( K_j \). \( E_{ij} \) is a matrix with value \( e_{ij} \) at entry \((k_{ij}^1, k_{ij}^1)\) and zero everywhere else. Here we emphasize that

- \( U_i \) is a partial Fourier matrix, and
- \( E \) is extremely sparse.

These observations turn out to be essential in the construction of the preconditioner.

Summarizing the discussion for both the diagonal and off-diagonal blocks, we arrive at the data-sparse approximation

\[
M \approx B + UEU^*.
\]

### 2.2. Directional preconditioner.

To precondition (5), we use the approximation (17) and consider the solution \( q \) of

\[
(B + UEU^*)q = f.
\]

First, introducing new vectors \( r = -U^* q \) and \( p = -Er \) gives an equivalent augmented system,

\[
\begin{bmatrix}
B & U & 0 \\
U^* & 0 & I \\
0 & I & E
\end{bmatrix}
\begin{bmatrix}
q \\
p \\
r
\end{bmatrix} =
\begin{bmatrix}
f \\
0 \\
0
\end{bmatrix}.
\]
Factorizing the matrix in (18) gives

\[
\begin{bmatrix}
I & U \ast B \ast^{-1} & I \\
U \ast B \ast^{-1} & I & -T \\
I & -T & I
\end{bmatrix}
\begin{bmatrix}
B & -S \\
- T & W \\
I & -T
\end{bmatrix}
\begin{bmatrix}
I & B \ast^{-1}U \\
I & I \\
B \ast^{-1}U & I
\end{bmatrix},
\]

with

\[
S = U \ast B \ast^{-1}U, \quad T = S \ast^{-1}, \quad W = E \ast T.
\]

Since both \(U\) and \(B\) are block-diagonal, \(S\) and \(T\) are also block-diagonal:

\[
S = \begin{bmatrix} S_1 & \cdots & S_m \end{bmatrix}, \quad T = \begin{bmatrix} T_1 & \cdots & T_m \end{bmatrix},
\]

with \(S_i = U \ast_i B \ast_i^{-1}U\) and \(T_i = S_i \ast^{-1}\). Inverting the factorization (19) gives

\[
\begin{bmatrix}
q \\
p \\
r
\end{bmatrix} =
\begin{bmatrix}
I & -B \ast^{-1}U \\
I & I \\
I & T
\end{bmatrix}
\begin{bmatrix}
I & B \ast^{-1} \\
I & I \\
U \ast B \ast^{-1} & I
\end{bmatrix}
\begin{bmatrix}
f \\
0 \\
0
\end{bmatrix}.
\]

Applying (20) exactly can be quite costly. In order to construct an efficient preconditioner, it is essential to approximate (20) aggressively without sacrificing too much accuracy. In what follows, we shall use the notation \(|A|\) to stand for an approximation to the matrix \(A\), although the actual approximation scheme can be different for different matrices.

- First, the \(2^\ell \times 2^\ell\) matrix \(B_i\) is the restriction of the integral operator to a straight segment of length \(2^\ell \lambda\). Since the geometry is fixed, all \(B_i\) and \(B_i \ast^{-1}\) can be precomputed. Since this is also a 1D problem (i.e., restriction to a straight segment), we can use the hierarchical matrix algebra [7] or the hierarchical semiseparable (HSS) matrices [32] to compress and apply \(B_i \ast^{-1}\) efficiently. We denote by \(|B_i \ast^{-1}|\) the approximation for \(B_i \ast^{-1}\) and accordingly by \(|B_i|\) the one for \(B_i\).

- Second, since each \(U_i\) is a partial Fourier matrix, applying \(U\) and \(U \ast\) reduces to a number of FFTs, which is highly efficient.

- Third, \(S_i\) and \(T_i\) can be precomputed, as they depend only on \(U_i\) and \(B_i\), both of which have already been precomputed. An important observation is that \(T_i\) is numerically sparse (see Figure 2). Therefore, for the sake of efficiency, we approximate \(T_i\) with \(|T_i|\), which is obtained by thresholding the entries in absolute value. In the numerical results, the number of nonzero entries in \(|T_i|\) is kept proportional to the dimension of \(T_i\). This approximation of \(T\) is denoted by \(|T|\).

- The final task is to build an approximate inverse of \(W = E \ast T\). Here, the essential observation is the following:
  - \(T\) concentrates on its diagonal (see Figure 2).
Based on this, we define \(|W|\) to be the sum of \(E\) and the diagonal of \(T\) (i.e., thresholding the rest of the entries of \(T\) to zero). \(|W|\) is extremely sparse, as
the number of nonzeros is about 3/2 times the dimension of the matrix in most cases. To approximate $W^{-1}$, we choose to perform an incomplete LU decomposition,

$$[W] \approx L_{[W]} R_{[W]} ,$$

where $L_{[W]}$ and $R_{[W]}$ are highly sparse upper- and lower triangular. Figure 3 shows the typical sparsity pattern of the matrix $L_{[W]}$ and $R_{[W]}$. Once these two matrices are available, applying $W^{-1}$ approximately reduces to applying backward and forward substitutions with matrices $L_{[W]}$ and $R_{[W]}$. Once these approximations have been formed, we approximate (20) with

$$\begin{bmatrix} q \\ p \\ r \end{bmatrix} \leftarrow \begin{bmatrix} I & -[B^{-1}]U \\ I & I \\ I & [T] \end{bmatrix} \begin{bmatrix} I \\ I \\ [T] \end{bmatrix} \begin{bmatrix} [B^{-1}] \\ [T] \\ -[T] \end{bmatrix} R_{[W]}^{-1} L_{[W]}^{-1} \begin{bmatrix} f \\ 0 \\ 0 \end{bmatrix} . \tag{21}$$

We emphasize again that the following approximations are used for computing (21):

- replacing $B^{-1}$ with $[B^{-1}]$ via hierarchical matrix or HSS approximation for each $B^{-1}_{i}$;
- applying $U$ rapidly via FFT for each $U_{i}$;
- replacing $T$ with $[T]$ via sparse approximation for each $T_{i}$, and
- replacing $W^{-1}$ with $R_{[W]}^{-1} L_{[W]}^{-1}$ via sparse backward and forward substitutions for $L_{[W]}$ and $R_{[W]}$. 

---

**Fig. 2.** The absolute value of the entries of $S_{i}$ (left) and $T_{i}$ (right) in logarithmic scale (base 10) for the sound-soft case (top row) and the sound-hard case (bottom row). For this specific segment, $\ell_i = 5$ and $p = 8$. Hence, there are 256 equally spaced quadrature points in $P_{i}$, and the cardinality of $K$ is 65. Both $S_{i}$ and $T_{i}$ are numerically sparse and diagonally dominant.
Based on (21), our preconditioner is defined as follows. For a given $f$, it

- forms vector $(f^*, 0, 0)^*$,
- carries out the computation of (21), and
- extracts the first component $q$ of the resulting vector.

Since the key step of constructing a data-sparse representation of the operator relies on the directional nature of the kernel $G(x, y)$, we name it the directional preconditioner.

2.3. Complexity analysis. We first consider the setup cost for the preconditioner, i.e., the approximate factorization in (21). Since the matrices $B_i^{-1}$ and $T_i$ for a segment $P_i$ depend only on the integer length parameter $\ell$ of $P_i$, the possible choices for these matrices are fixed and independent of the scatterer. All these possible choices can be precomputed once and stored for future use.

As a result, the setup algorithm consists of only two parts: the evaluation of $E$ and $[W]$, and the incomplete LU factorization $[W] \approx L_{[W]} R_{[W]}$.

- For the first part, since $E$ has only $\omega$ nonzero entries and computing each entry takes $O(1)$ steps (16), the overall cost for this step is $O(\omega) = O(n)$.
- Once $E$ is formed, computing $[W]$ also takes at most $O(n)$ steps.
- In our implementation, the incomplete LU factorization $[W] \approx L_{[W]} R_{[W]}$ is computed with the Crout version of ILU factorization, known as ILUC [25], with the drop tolerance set to be a small value, such as $10^{-6}$. The resulting factors $L_{[W]}$ and $R_{[W]}$ contain on average 5 or 6 nonzero entries per row/column. Therefore, they are extremely sparse. In all our numerical experiments, the computational time for generating the ILUC factorization scales roughly linearly with the dimension of the matrix $[W]$.

Let us now consider the application cost of the preconditioner (21). For the major
steps of applying (21), we have the following estimates:

- The application of \( B^{-1} \) is linear time due to the hierarchical matrix algebra approximation for \( B_i^{-1} \).
- The application of \( U \) is \( O(n \log n) \) since each \( U_i \) is a partial Fourier matrix and the FFT can be used.
- The application of \( T \) is \( O(n) \) since the number of nonzeros in \( T \) is proportional to \( O(n) \) after we threshold each \( T_i \).
- Applying \( R_{[W]}^{-1} L_{[W]}^{-1} \) reduces to the sparse backward and forward substitution algorithm for matrices \( L_{[W]} \) and \( R_{[W]} \). Since the number of nonzero entries in both matrices scales linearly with respect to the matrix dimension, the cost of this step is also \( O(n) \).

Putting these together shows that the application cost of the preconditioner scales like \( O(n \log n) \).

3. Numerical results. The proposed preconditioner is implemented in MATLAB. The numerical results in this section are obtained on a desktop computer with a 3.60 GHz CPU.

In these experiments, the domain boundary is discretized by the Nyström method with \( p = 8 \) points per wavelength. The Chebyshev grid size \( m_c \) used to construct \( e_{ij} \) for the matrix \( E \) is set to be 10. The boundary conditions for both the sound-soft and the sound-hard problems correspond to an incoming plane wave pointing towards the positive \( x_1 \) direction. The fast algorithm described in [33] with accuracy \( 10^{-6} \) is used to accelerate the matrix-vector multiplication of the iterative solve. The coupling constant \( \eta \) is chosen to be \( \frac{\omega^2}{2} \) for the sound-soft case (1) and to be \( \frac{3 \omega^4}{4} \) for the sound-hard case (2).

3.1. Example 1. The first example is an ellipse, shown in Figure 4, with major axis equal to \( r_1 = 1 \) and minor axis equal to \( r_2 = 0.45 \).

![Figure 4. The scatterer of Example 1.](image)

The first test of this example concerns the Dirichlet problem of sound-soft scattering for different values of \( \omega \). For the GMRES solver, the relative tolerance is set to \( 10^{-6} \) to match the accuracy of the matrix-vector multiplication algorithm, and the restart number is equal to 50. The results are reported in Table 1, where

- \( T_s \) is the setup time of the preconditioner in seconds,
- \( T_a \) is the application time of the preconditioner in seconds,
- \( T_m \) is the matrix-vector multiplication time using the algorithm from [33],
- \( n_p \) is the iteration number of the iterative solver when the preconditioner is used, and finally,
- \( n_n \) is the iteration number without any preconditioning.
The ratio $T_a/T_m$ is a good indicator of the computation cost of applying the preconditioner, and $n_p/n_n$ shows the effectiveness of the preconditioner.

**Table 1**

Results of the sound-soft scattering for Example 1.

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$T_s$</th>
<th>$T_a$</th>
<th>$T_m$</th>
<th>$n_p$</th>
<th>$n_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4e+03</td>
<td>8.2e+03</td>
<td>1.2e+00</td>
<td>3.2e-03</td>
<td>3.8e-01</td>
<td>12</td>
<td>31</td>
</tr>
<tr>
<td>5.3e+03</td>
<td>3.3e+04</td>
<td>5.3e+00</td>
<td>2.4e-02</td>
<td>1.5e+00</td>
<td>14</td>
<td>47</td>
</tr>
<tr>
<td>2.1e+04</td>
<td>1.3e+05</td>
<td>2.5e+01</td>
<td>7.4e-02</td>
<td>6.3e+00</td>
<td>16</td>
<td>71</td>
</tr>
</tbody>
</table>

The second test considers the Neumann problem of sound-hard scattering, again for different values of $\omega$. The results are summarized in Table 2.

**Table 2**

Results of the sound-hard scattering for Example 1.

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$T_s$</th>
<th>$T_a$</th>
<th>$T_m$</th>
<th>$n_p$</th>
<th>$n_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.4e+03</td>
<td>8.2e+03</td>
<td>1.7e+00</td>
<td>2.7e-03</td>
<td>3.8e-01</td>
<td>12</td>
<td>24</td>
</tr>
<tr>
<td>5.5e+03</td>
<td>3.3e+04</td>
<td>7.4e+00</td>
<td>9.0e-03</td>
<td>1.6e+00</td>
<td>15</td>
<td>36</td>
</tr>
<tr>
<td>2.2e+04</td>
<td>1.3e+05</td>
<td>3.9e+01</td>
<td>5.1e-02</td>
<td>6.7e+00</td>
<td>17</td>
<td>53</td>
</tr>
</tbody>
</table>

The results of these first two tests show that the setup time of the preconditioner is typically equivalent to a couple of fast matrix-vector multiplications. On the other hand, the application time of the preconditioner is much lower, and, as a result, the cost of applying the preconditioner is almost negligible during the iterative solution. The iteration number of the preconditioned system is significantly lower than that of the unpreconditioned system and in fact scales like $O(\log \omega)$.

The third test examines the dependence of the convergence rate on the curvature of the domain boundary. In this case, $\omega$ is set to be 1.4e+03, and the minor axis $r_2$ is set to be 0.05, 0.15, and 0.45. From the results reported in Table 3, it seems that the convergence rate is quite insensitive to the curvature variation.

**Table 3**

Results of different $r_2$ values for Example 1.

<table>
<thead>
<tr>
<th>$r_2$</th>
<th>$n$</th>
<th>$n_p$</th>
<th>$n_n$</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.05</td>
<td>8.2e+03</td>
<td>13</td>
<td>29</td>
</tr>
<tr>
<td>0.15</td>
<td>8.2e+03</td>
<td>13</td>
<td>24</td>
</tr>
<tr>
<td>0.45</td>
<td>8.2e+03</td>
<td>12</td>
<td>31</td>
</tr>
</tbody>
</table>

The fourth test checks the convergence behavior of different iterative algorithms. For multiple values of $\omega$, the problem is solved with the restarted GMRES, the non-restarted GMRES, and the TFQMR algorithm. For each algorithm, both the preconditioned and unpreconditioned equations are solved. The results, summarized in Table 4, show that the performance for the restarted and the nonrestarted GMRES algorithms is comparable both for preconditioned and unpreconditioned systems. On the other hand, for the TFQMR algorithm, the preconditioned solve exhibits a similar behavior while the unpreconditioned solve takes many more iterations.

The final test studies the spectra of the preconditioned and unpreconditioned systems. Here we let $r_2 = 0.45$, $\omega = 3.4e+02$, and $n = 2.0e+03$. The spectrum plots in Figure 5 show that the eigenvalue distribution of the preconditioned system is much more clustered, and this explains why the preconditioned system converges faster.
Table 4
Results of different iterative algorithms for Example 1. GMRES-R stands for the restarted GMRES; GMRES-NR stands for the nonrestarted GMRES.

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$T_{GMRES}$</th>
<th>$T_{GMRES-NR}$</th>
<th>$T_{TFQMR}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.4e+03$</td>
<td>8.2e+03</td>
<td>12</td>
<td>31</td>
<td>15</td>
</tr>
<tr>
<td>$5.5e+03$</td>
<td>3.3e+04</td>
<td>14</td>
<td>46</td>
<td>16</td>
</tr>
<tr>
<td>$2.2e+04$</td>
<td>1.3e+05</td>
<td>16</td>
<td>74</td>
<td>19</td>
</tr>
</tbody>
</table>

Fig. 5. Spectra of unpreconditioned and preconditioned systems for Example 1.

3.2. Example 2. The second example is a bean-shaped scatter, as shown in Figure 6.

The first test considers the Dirichlet problem of sound-soft scattering for different values of $\omega$. The parameters of the GMRES solve are set to be the same as those for Example 1, i.e., the relative tolerance equal to $10^{-6}$ and the restart number equal to 50. The results are reported in Table 5.

Table 5
Results of the sound-soft scattering for Example 2.

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$T_e$</th>
<th>$T_a$</th>
<th>$T_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$1.3e+03$</td>
<td>8.2e+03</td>
<td>1.2e+00</td>
<td>2.3e-03</td>
<td>4.0e-01</td>
</tr>
<tr>
<td>$5.2e+03$</td>
<td>3.3e+04</td>
<td>5.5e+00</td>
<td>1.3e-02</td>
<td>1.8e-00</td>
</tr>
<tr>
<td>$2.1e+04$</td>
<td>1.3e+05</td>
<td>3.2e+01</td>
<td>2.4e-02</td>
<td>7.5e+00</td>
</tr>
</tbody>
</table>

The second test studies the Neumann problem of sound-hard scattering, again for different values of $\omega$. The results are given in Table 6. The results in these two tests are qualitatively similar to the analogous tests for Example 1.
The third test checks the convergence behavior of different iterative algorithms. For different values $\omega$, the problem is again solved with the restarted GMRES, the nonrestarted GMRES, and the TFQMR. The results, summarized in Table 7, show that the performance for the restarted and the nonrestarted GMRES algorithms is comparable and that the preconditioned system converges much faster for the TFQMR algorithm.

### Table 7

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$T_{\text{GMRES-R}}$</th>
<th>$T_{\text{GMRES-NR}}$</th>
<th>$T_{\text{TFQMR}}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.3e+03</td>
<td>8.2e+03</td>
<td>1.7e+00</td>
<td>1.7e+00</td>
<td>4.1e-01</td>
</tr>
<tr>
<td>5.2e+03</td>
<td>3.3e+04</td>
<td>7.4e+00</td>
<td>8.5e-03</td>
<td>1.8e+00</td>
</tr>
<tr>
<td>2.1e+04</td>
<td>1.3e+05</td>
<td>3.9e+01</td>
<td>4.0e-02</td>
<td>7.4e+00</td>
</tr>
</tbody>
</table>

The final test computes the spectra of the preconditioned and unpreconditioned systems. Here $\omega = 3.4e+02$ and $n = 2.0e+03$. The spectrum plots, given in Figure 7, show that the eigenvalue distribution of the preconditioned system is again much more clustered.

### Figure 7

Spectra of unpreconditioned and preconditioned systems for Example 2.

### 3.3. Example 3

The third example is a resonator, shown in Figure 8, with opening angle equal to $\pi/2$.

The first test considers the Dirichlet problem of sound-soft scattering for different values of $\omega$, again with the GMRES relative tolerance equal to $10^{-6}$ and the restart number equal to 50. The results of this test are reported in Table 8.

The second test considers the Neumann problem of sound-hard scattering for different values of $\omega$. The results are summarized in Table 9.

The third test examines the dependence of the convergence behavior on the opening angle of the resonator. In this case, the $\omega$ is set to be $1.4e+03$, and the opening angle $\theta$ is equal to $3\pi/4$, $\pi/2$, and $\pi/4$. From the results in Table 10, we notice that the iteration count depends highly on the opening angle. When the opening
The scatterer of Example 3.

Table 8

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$T_a$</th>
<th>$T_m$</th>
<th>$n_p$</th>
<th>$n_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1e+02</td>
<td>8.2e+03</td>
<td>1.2e+00</td>
<td>3.6e-03</td>
<td>4.1e-01</td>
<td>23</td>
</tr>
<tr>
<td>3.2e+03</td>
<td>3.3e+04</td>
<td>5.4e+00</td>
<td>1.1e-02</td>
<td>1.8e+00</td>
<td>26</td>
</tr>
<tr>
<td>1.3e+04</td>
<td>1.3e+05</td>
<td>3.2e+01</td>
<td>3.8e-02</td>
<td>7.7e+00</td>
<td>30</td>
</tr>
</tbody>
</table>

Table 9

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$T_a$</th>
<th>$T_m$</th>
<th>$n_p$</th>
<th>$n_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1e+02</td>
<td>8.2e+03</td>
<td>1.7e+00</td>
<td>2.7e-03</td>
<td>4.1e-01</td>
<td>20</td>
</tr>
<tr>
<td>3.2e+03</td>
<td>3.3e+04</td>
<td>7.3e+00</td>
<td>8.1e-03</td>
<td>1.8e+00</td>
<td>24</td>
</tr>
<tr>
<td>1.3e+04</td>
<td>1.3e+05</td>
<td>3.9e+01</td>
<td>3.8e-02</td>
<td>7.7e+00</td>
<td>27</td>
</tr>
</tbody>
</table>

Table 10

<table>
<thead>
<tr>
<th>$\theta$</th>
<th>$n$</th>
<th>$n_p$</th>
<th>$n_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$3\pi/4$</td>
<td>8.2e+03</td>
<td>19</td>
<td>44</td>
</tr>
<tr>
<td>$\pi/2$</td>
<td>8.2e+03</td>
<td>23</td>
<td>53</td>
</tr>
<tr>
<td>$\pi/4$</td>
<td>8.2e+03</td>
<td>76</td>
<td>178</td>
</tr>
</tbody>
</table>

When the opening angle is small, the scatterer is a strong resonator, and both the preconditioned and unpreconditioned iteration counts grow rapidly.

The fourth test again checks the convergence behavior of different iterative algorithms. For different values $\omega$, the problem is still solved with the restarted GMRES, the nonrestarted GMRES, and the TFQMR algorithm. The results in Table 11 again show that the preconditioned system offers a huge speedup for the TFQMR algorithm.

Table 11

<table>
<thead>
<tr>
<th>$\omega$</th>
<th>$n$</th>
<th>$n_p$</th>
<th>$n_m$</th>
</tr>
</thead>
<tbody>
<tr>
<td>8.1e+02</td>
<td>8.2e+03</td>
<td>23</td>
<td>53</td>
</tr>
<tr>
<td>3.2e+03</td>
<td>3.3e+04</td>
<td>26</td>
<td>79</td>
</tr>
<tr>
<td>1.3e+04</td>
<td>1.3e+05</td>
<td>30</td>
<td>122</td>
</tr>
</tbody>
</table>

The final test computes the spectra of the preconditioned and unpreconditioned systems. Here, the opening angle $\theta$ is $\pi/2$, $\omega = 3.4e+02$, and $n = 2.0e+03$. The spectrum plots in Figure 9 show that the eigenvalue distribution of the precondi-
tioned system is still more clustered, though not as significant as the distributions in Examples 1 and 2.

4. Conclusion. This paper presents the directional preconditioner for the combined field integral equations (CFIEs) of high-frequency acoustic obstacle scattering in two dimensions. The main idea is to construct a data-sparse approximation of the linear operator, transform it into an approximate sparse linear system, and form an approximate inverse using efficient sparse and hierarchical linear algebra algorithms.

In this paper, we have assumed that the boundary is discretized with an equally spaced set of discretization points. For nonequally spaced points, the construction goes through as well, except that the FFT has to be replaced with nonuniform FFTs. As a result, some of the scatterer-independent precomputation can become dependent on the discretization pattern.

We would like to extend this approach to 3D scatterers. While the main idea should work, the lack of equally spaced discretization for general surfaces poses a clear challenge for this approach.

The long-term goal is to construct a direct solver for the boundary integral equations of the obstacle scattering problem. It is not clear at this point whether such a direct solver even exists. However, this paper can be viewed a first step toward exploring in this direction.

One important ingredient of our approach is to transform a dense oscillatory matrix into a sparse one. Once it is in a sparse form, we are able to leverage the power of sparse linear algebra algorithms. While traditionally there is relatively little overlap between work in integral equations and that in sparse linear algebra, this work hints at a fruitful exchange of ideas between these two fields.

Acknowledgments. The author thanks Anil Damle and the anonymous referees for comments and suggestions.

REFERENCES


