Massively Parallel Schwarz Methods for the High Frequency Helmholtz Equation

Yan Xie[0009-0004-2861-3383],
Shihua Gong[0000-0003-3650-2283],
Ivan G. Graham[0000-0002-5730-676X],
Euan A. Spence[0000-0003-1236-4592],
Chen-Song Zhang[0000-0002-2213-0899]

1 The Helmholtz problem

We consider the classical Helmholtz equation given by

$$\Delta u + \frac{k^2}{c^2} u = -f \quad \text{in} \quad \mathbb{R}^d, \tag{1}$$

subject to the Sommerfeld radiation condition:

$$\frac{\partial u}{\partial r} - \mathrm{i}ku = o\left(\frac{1}{r^{(d-1)/2}}\right), \quad r = |x| \to \infty. \tag{2}$$

Here k is the angular frequency, $c \in C^\infty(\mathbb{R}^d)$ represents the (possibly variable) wave speed, and $f \in L^2(\mathbb{R}^d)$ is the source term. We assume that both f and $1 - c$ are compactly supported within a hyper-rectangle $\Omega_{\text{int}} := \prod_{i=1}^d (a^i, b^i)$. To keep the presentation simple, we restrict here to constant wave speed case $c = 1$.

To formulate the problem in a bounded computational domain, we restrict (1) to the extended domain $\Omega := \prod_{i=1}^d (a^i - \kappa_{\mathrm{g}}, b^i + \kappa_{\mathrm{g}})$, where κ_{g} denotes the thickness of the Cartesian PML surrounding Ω_{int}. We introduce a smooth scaling function $g \in C^\infty(\mathbb{R})$, satisfying

Yan Xie · Chen-Song Zhang
SKLMS, Academy of Mathematics and Systems Science, Chinese Academy of Sciences & School of Mathematical Sciences, University of Chinese Academy of Sciences, Beijing, 100049, China, e-mail: xieyan2021,zhangcs@lsec.cc.ac.cn

Shihua Gong
School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China, e-mail: gongshihua@cuhk.edu.cn

Ivan G. Graham · Euan A. Spence
Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK, e-mail: masigg,eas25@bath.ac.uk

$$\begin{cases} g(x) = g'(x) = 0, & x \leq 0 \\ g'(x) > 0, & x > 0 \\ g''(x) = 0, & x \in (\kappa_\infty, +\infty) \end{cases}$$

The scaling function $g_i(x)$ for each direction i is then given by:

$$g_i(x^i) = \begin{cases} g(x^i - b^i), & x^i \geq b^i \\ 0, & x^i \in (a^i, b^i) \\ -g(a^i - x^i), & x^i \leq a^i \end{cases}$$

Using these, we define the PML-modified Laplacian as: $\Delta_{pml} := \sum_{i=1}^d \left(\gamma_i(x^i)^{-1}\partial_{x^i}\right)^2$ where $\gamma_i := 1 + \mathrm{i}g_i'$. This formulation allows for the analytic continuation of the solution into complex coordinates, effectively absorbing outgoing waves. In weak formulation, the truncated Helmholtz problem with PML reads: find $u \in H_0^1(\Omega)$ such that

$$a(u,v) := \int_\Omega (D\nabla u)\cdot\nabla v - (\beta\cdot\nabla u)v - k^2 uv = \int_\Omega fv, \quad \forall v \in H_0^1(\Omega). \tag{3}$$

where D is the diagonal matrix with entries $D_{ii} = \gamma_i(x^i)^{-2}$, and β is the vector field with components $\beta_i = \gamma_i'(x^i)/\gamma_i(x^i)^3$.

2 Restricted additive Schwarz method with PML transmission conditions

The work [3] analyzed Schwarz methods (both additive and multiplicative) with perfectly matched layer transmission conditions, and provided theoretical convergence results for these when applied to the high-frequency Helmholtz equation. In this paper, we implement a practical variant of the additive method and present numerical results to demonstrate that it is scalable to $O(k^d)$ pocessors. Earlier work with lower levels of parallel scaling can be found in [7, 6, 2]. To keep the presentation self-contained, we summarize here the key components of the method.

Cartesian covering. We cover the computational domain Ω with N overlapping subdomains $\Omega_{j,\mathrm{int}} := \prod_{i=1}^d (a_j^i, b_j^i)$, $j = 1, \cdots, N$ obtained by extending a non-overlapping Cartesian partition in each coordinate direction. The overlap width is δ. We then extend each interior boundary $\partial\Omega_{j,\mathrm{int}} \not\subset \partial\Omega$ by a PML layer to obtain $\Omega_j = \prod_{i=1}^d (a_j^i - \kappa_{j,l}^i, b_j^i + \kappa_{j,u}^i)$, where

$$\kappa_{j,l}^i = \begin{cases} 0, & a_j^i = a^i - \kappa_{\mathrm{g}} \\ \kappa, & a_j^i > a^i - \kappa_{\mathrm{g}} \end{cases}, \quad \kappa_{j,u}^i = \begin{cases} 0, & b_j^i = b^i + \kappa_{\mathrm{g}} \\ \kappa, & b_j^i < b^i + \kappa_{\mathrm{g}} \end{cases}$$

and κ is the thickness of the Cartesian PML on subdomains. For each subdomain Ω_j, we define a local PML problem, using the local PML scaling function:

$$g_j^i(x^i) = \begin{cases} g(x^i - b_j^i), & x^i \geq b_j^i \\ g_i(x^i), & x^i \in (a_j^i, b_j^i) \\ -g(a_j^i - x^i), & x^i \leq a_j^i \end{cases}$$

For each $j = 1, ..., N$, we let a_j denote the restriction of the sesquilinear form a from (3) to $H_0^1(\Omega_j)$. To combine local solutions into a global approximation, we introduce a set of non-negative partition of unity (PoU) functions $\{\chi_j\}$ on Ω based on the cover $\{\Omega_{j,\text{int}}\}$, such that $\sum_{j=1}^N \chi_j \equiv 1$, and $\text{supp}(\chi_j) \subset \Omega_{j,\text{int}}$ (i.e., χ_j vanishes on the extra PML of Ω_j) and $\chi_j \equiv 1$ on the non-overlapped part of Ω_j, namely $\Omega_{j,\text{novlp}} := \Omega_{j,\text{int}} \backslash (\bigcup_{l \neq j} \Omega_{l,\text{int}})$.

Restricted additive Schwarz method (RAS). With the above definitions, the RAS-PML method, defined (before discretization) in [3], is given in Algorithm 1.

Algorithm 1: `RAS_IterSolve`

Input: Initial guess $u^{(0)}$, source term f, number of subdomains N, domain Ω
Output: Approximate solution u

Partition Ω into Cartesian subdomains $\{\Omega_j\}_{j=1}^N$;
for $n = 0, 1, 2, \ldots$ ▷ Outer loop over iterations **do**
  for $j = 1, \ldots, N$ ▷ Inner loop (parallel) over subdomains **do**
    Find $\mathfrak{c}_j^{(n+1)} \in H_0^1(\Omega_j)$ such that:

$$a_j(\mathfrak{c}_j^{(n+1)}, v_j) = (f, v_j) - a(u^{(n)}, v_j), \quad \forall v_j \in H_0^1(\Omega_j); \tag{4}$$

  Compute the global update: $u^{(n+1)} = u^{(n)} + \sum_j \chi_j \mathfrak{c}_j^{(n+1)}$;
return $u^{(n+1)}$;

Now let u_h denote the Galerkin solution of (3) in a conforming finite element space $V_h \subset H_0^1(\Omega)$. Then, with $V_{h,j} := \{v_h|_{\Omega_j} : v_h \in V_h\} \cap H_0^1(\Omega_j)$, the discrete version of Algorithm 1 for computing u_h is as follows. Given the current iterate $u_h^{(n)} \in V_h$, we compute local corrections $\mathfrak{c}_{h,j}^{(n+1)} \in V_{h,j}$ by solving the discrete counterpart of (4):

$$a_j(\mathfrak{c}_{h,j}^{(n+1)}, v_{h,j}) = (f, \mathcal{R}_{h,j}^T v_{h,j}) - a(u_h^{(n)}, \mathcal{R}_{h,j}^T v_{h,j}), \quad \forall v_{h,j} \in V_{h,j}, \tag{5}$$

where $\mathcal{R}_{h,j}^\top$ denotes the extension by zero from $V_{h,j}$ to V_h. The new iterate $u_h^{(n+1)}$ is updated as:

$$u_h^{(n+1)} = u_h^{(n)} + \sum_j \tilde{\mathcal{R}}_{h,j}^T \mathfrak{c}_{h,j}^{(n+1)}, \tag{6}$$

where $\tilde{\mathcal{R}}_{h,j}^{\top}$ denotes the weighted extension by χ_j from $V_{h,j}$ to V_h. The corresponding preconditioner is given by

$$\mathcal{B}^{-1} := \sum_j \widetilde{\mathcal{R}}_j^{\top} \mathcal{A}_j^{-1} \mathcal{R}_j, \tag{7}$$

where $\mathcal{A}_j$ is the local operator corresponding to a_j.

As shown in [3, §8], this is a restricted additive Schwarz (RAS) method, where each local subdomain problem is equipped with a PML and a Dirichlet boundary condition.

Theoretical results. Suppose the PoU $\{\chi_j\}$ is C^∞ and the Helmholtz problem is non-trapping, which means that all rays of geometric optics escape the domain in a finite time, avoiding closed cycles or infinite reflections (see [3, §1.7]). Then results from [3, Theorems 1.1-1.4 and 1.6] establish conditions, for which, given any $M > 0$ and integer $s \geq 1$, there exist constants $\mathcal{N} \in \mathbb{N}$ and $k_0 > 0$ (both independent of f) such that

$$\|u - u^{(\mathcal{N})}\|_{H_k^s(\Omega)} \leq k^{-M} \|u - u^{(0)}\|_{H_k^1(\Omega)}, \text{ for } k > k_0. \tag{8}$$

Here the weighted Sobolev norm is defined as $||v||^2_{H_k^s(\Omega)} := \sum_{|\alpha| \leq s} ||(k^{-1}\partial)^\alpha v||^2_{L^2(\Omega)}$. In particular, (8) implies that the fixed-point iterations converge super-algebraically fast in the number of iterations for sufficiently-large k, and that the rate of convergence improves as k increases. These theoretical results apply on arbitrary overlap δ and PML width κ, but these have to remain fixed as k increases.

Mesh refinement and number of subdomains. To resolve the oscillatory solutions of (1), the mesh size must decrease at least as fast as $h = O(k^{-1})$ as $k \to \infty$, leading to a finite element system with at least $O(k^d)$ degrees of freedom (DoFs). To achieve efficient parallelization, we partition the domain first into $O(k)$ non-overlapping subdomains along each coordinate direction, so that the number of DoFs within each remains approximately constant as k grows. Then we add the overlap and PML layers of thickness δ, κ as described above. The case of meshes refined to avoid the pollution effect will also be discussed in future work.

Although the theory in [3] assumes that δ, κ should be fixed with respect to k, the results here show that these parameters can be chosen to decrease quickly as k increases, leading to a scalable algorithm with no loss of convergence rate. (Preliminary experiments were given in [4].)

3 Practical improvements

We propose several practical improvements to Algorithm 1 which will be illustrated by the numerical experiments below. These include:

(i) **Reducing communication by exploiting the sparsity of the residual term.** It can be shown that the right-hand side of (5) is nonzero only in overlapping regions, and thus the communication of the local residuals can be restricted to these regions.

(ii) **Combining PML and impedance boundary conditions for improved robustness.** The fact that the local corrections (5) are computed in $V_{h,j} \subset H_0^1(\Omega_j)$ implies that a Dirichlet condition is applied at the boundary $\partial\Omega_j$. (This is a common set-up when PML is used.) However it is simple to apply instead the impedance boundary condition

$$\partial_{\nu_j} u - \mathrm{i}ku = 0 \quad \text{on} \quad \partial\Omega_j, \tag{9}$$

(as a natural boundary condition on subdomains), where ∂_{ν_j} denotes is the outward normal derivative on $\partial\Omega_j$, i.e. we use a hybrid of the PML with impedance boundary conditions. The local sesquilinear forms a_j then incorporate (9) as a natural boundary condition. This improves robustness when the PML is thin (see Table 1). This hybrid treatment, also used for Maxwell's equations [1], yields smaller errors than PML with Dirichlet boundary condition.

(iii) **Scaling of width of overlap and PML layers as k increases.** To balance communication and convergence, we have found that it is advantageous to let the number of grid points in the overlap and PML layers grow logarithmically with k. Since the mesh diameter in these experiments is of the order of a wavelength $\lambda = 2\pi/k$, we choose overlap δ and PML width κ according to the formulae

$$\delta = C_\delta\, \ell(k)\, \frac{2\pi}{k} \quad \text{and} \quad \kappa = C_\kappa\, \ell(k)\, \frac{2\pi}{k}, \tag{10}$$

where $\ell(k) := \frac{k}{k-k_0}\log_2(k/k_0)$, for some constants k_0, C_δ and C_κ to be chosen.

4 Numerical experiments

In our numerical tests, Ω is the unit square, and the source is the smoothed delta function: $f(x) = \frac{16k^2}{\pi^3}\exp\left(-\frac{16k^2(x-x_c)^2}{\pi^2}\right)$, with the source point $x_c = (0.5, 0.5)$, the global PML width is $\kappa_g = 3\frac{2\pi}{k}$ (i.e., 3 wavelengths) and the PML coordinate scaling function is taken to be $g(x) = 10kx^3$.

For this experiment the problem (1) is discretized on a uniform square mesh with 12 grid points per wavelength using bilinear elements. The domain decomposition consists of N uniform overlapping square subdomains and the PoU functions are taken to be the tensor products of the 1D PoU functions which vary linearly across the overlap regions in each direction. In the tables, `ovlp` and `pml` denote the number of grid points in the overlapping and PML regions respectively. For the formula (10) of δ and κ, we use $\ell(k)$ with $k_0 = 150$. In all cases, the number of processors is equal to the total number of subdomains $N = \mathcal{O}(k^2)$, chosen so that each non-overlapping subdomain has a bounded number of degrees of freedom as $k \to \infty$. Iterations are terminated when the relative residual is below rtol=1E-10 or the number of iterations exceeds 500.

4.1 Different types of boundary conditions

We first compare the performance of three subdomain boundary condition strategies: RAS-PML-Drch means that a PML is combined with a subdomain Dirichlet boundary condition, as implied by (5). RAS-PML-Imp means that a PML is combined with a subdomain impedance boundary condition, as described in §3(ii), while RAS-Imp means that the PML is discarded and the impedance boundary condition (9) is imposed directly on subdomain boundaries. A comparison of these three strategies is given in Table 1. RAS-PML-Imp exhibits superior performance over RAS-PML-Drch, particularly at high frequencies, while RAS-Imp fails to converge at all for high k.

To further enhance convergence, we use the Algorithm 1 as a preconditioner for the Krylov method GMRES. The results in Table 2 show a significant improvement over Table 1, with RAS-PML-Imp remaining the most effective. However, since GMRES entails higher communication costs and offers little advantage when convergence is already satisfactory, we only employ the simple Richardson iteration (5), (6) in the following experiments. From now on, we refer to RAS-PML-Imp simply as RAS-PML for brevity.

Table 1 The performance of RAS-PML-Imp, RAS-PML-Drch and RAS-Imp.

κ, δ in (10) with $C_\kappa = 3/8$, $C_\delta = 1/6$					**RAS-PML-Imp**		**RAS-PML-Drch**		**RAS-Imp**	
k	grid	N	pml	ovlp	iter	relres	iter	relres	iter	relres
300	600^2	4	8	4	12	1.88E-11	13	2.29E-11	30	6.44E-11
600	1200^2	16	11	5	17	9.15E-11	18	7.22E-11	43	7.13E-11
1200	2400^2	64	15	6	34	7.26E-11	35	8.83E-11	500	0.016
2400	4800^2	256	18	8	77	8.98E-11	94	8.91E-11	×	-diverged-
4800	9600^2	1024	22	10	196	9.12E-11	500	1.17E-10	×	-diverged-
9600	19200^2	4096	26	12	500	4.43E-08	×	-diverged-	×	-diverged-

Table 2 Use GMRES to improve the performance of RAS-PML-Imp, RAS-PML-Drch and RAS-Imp with slightly thin PML thickness.

κ, δ in (10) with $C_\kappa = 3/8$, $C_\delta = 1/6$					**RAS-PML-Imp**		**RAS-PML-Drch**		**RAS-Imp**	
k	grid	N	pml	ovlp	iter	relres	iter	relres	iter	relres
300	600^2	4	8	4	11	7.86E-11	12	8.84E-11	29	6.71E-11
600	1200^2	16	11	5	19	6.90E-11	19	6.28E-11	41	8.34E-11
1200	2400^2	64	15	6	48	9.30E-11	49	9.31E-11	95	8.20E-11
2400	4800^2	256	18	8	88	8.50E-11	91	5.30E-11	310	9.24E-11
4800	9600^2	1024	22	10	154	6.93E-11	173	9.62E-11	500	4.49E-10
9600	19200^2	4096	26	12	294	8.90E-11	389	9.92E-11	500	9.24E-06

4.2 Choice of PML and overlapping width

We next compare the convergence behavior of RAS-PML under different choices of PML and overlap widths. Table 3 shows that with κ chosen as a fixed multiple of wavelength, RAS-PML eventually diverges, independently of whether δ contains a fixed or logarithmically growing number of grid-points. In contrast, if κ is chosen as in (10) then convergence is obtained both for δ chosen as a multiple of wavelength or growing more quickly as in (10), with the latter producing the best iteration counts, in fact with close to linear growth as frequency increases. Then Table 4 shows that this strategy results in a total runtime that increases linearly with respect to k.

Table 3 The performance of RAS-PML with different PML and overlapping width.

			$\kappa = 2\frac{2\pi}{k}$						κ in (10) with $C_\kappa = 1/2$							
k	grid	N	$\delta = \frac{2}{3}\frac{2\pi}{k}$			δ in (10) with $C_\delta = 1/6$			$\delta = \frac{1}{3}\frac{2\pi}{k}$				δ in (10) with $C_\delta = 1/6$			
			pml	ovlp	iter	pml	ovlp	iter	pml	ovlp	iter	ratio	pml	ovlp	iter	ratio
300	600^2	4	25	8	6	25	4	7	12	4	7	-	12	4	7	-
600	1200^2	16	25	8	13	25	5	14	16	4	14	2.00	16	5	14	2.00
1200	2400^2	64	25	8	26	25	6	31	21	4	31	2.21	21	6	31	2.21
2400	4800^2	256	25	8	61	25	8	61	26	4	72	2.32	26	8	60	1.94
4800	9600^2	1024	25	8	140	25	10	134	32	4	149	2.07	32	10	119	1.98
9600	19200^2	4096	25	8	$\times$	25	12	> 500	38	4	298	2.00	38	12	224	1.88

Table 4 Time table where the total runtime increase with $O(k)$.

κ, δ in (10) with $C_\kappa = 1/2,\ C_\delta = 1/6$					rtol=1E-10				
k	grid	N	pml	ovlp	iter	setup(s)	solve(s)	total time(s)	total/k
300	600^2	4	12	4	7	0.95	0.27	1.22	0.0041
600	1200^2	16	16	5	14	1.12	0.68	1.80	0.0030
1200	2400^2	64	21	6	31	1.40	1.87	3.27	0.0027
2400	4800^2	256	26	8	60	1.67	4.39	6.06	0.0025
4800	9600^2	1024	32	10	119	1.94	11.20	13.15	0.0027
9600	19200^2	4096	38	12	224	1.90	23.22	25.12	0.0026

5 Conclusion

We describe several improvements of our previous work [5, 3, 4] to develop a practical parallel RAS-PML method for solving high-frequency Helmholtz equations. We apply both PMLs and impedance boundary conditions for subdomains to make the method more robust. We show by experiment that allowing the PML width to contain a logarithmically growing number of grid points can yield good convergence rates without excessive computation and communication. Under a Cartesian covering with $O(k^2)$ subdomains for 2D problems with $O(k^2)$ DoFs, numerical experiments demonstrate that both iteration counts and total runtime grow nearly linearly for the increasing frequency k. Full details, analysis and extensions to variable wavespeed and 3D are given in future work.

Acknowledgements SG was supported by the National Natural Science Foundation of China (No. 12201535, 12411530066) and Guangdong Provincial Program. ES was supported by the ERC synergy grant "PSINumScat" 101167139.

References

1. Collino, F., Monk, P.: The perfectly matched layer in curvilinear coordinates. SIAM J. Sci. Comput. **19**(6), 2061–2090 (1998)
2. Dai, R., Modave, A., Remacle, J.F., Geuzaine, C.: Multidirectional sweeping preconditioners with non-overlapping checkerboard domain decomposition for helmholtz problems. J. Comput. Phys. **453**, 110887 (2022)
3. Galkowski, J., Gong, S., Graham, I.G., Lafontaine, D., Spence, E.A.: Convergence of overlapping domain decomposition methods with PML transmission conditions applied to nontrapping Helmholtz problems. arXiv preprint arXiv:2404.02156 (2024)
4. Galkowski, J., Gong, S., Graham, I.G., Lafontaine, D., Spence, E.A.: Schwarz methods with PMLs for Helmholtz problems: fast convergence at high frequency. arXiv preprint arXiv:2408.16580, to appear in proceedings of DD28 (2025)
5. Gong, S., Gander, M.J., Graham, I.G., Lafontaine, D., Spence, E.A.: Convergence of parallel overlapping domain decomposition methods for the Helmholtz equation. Numer. Math. **152**(2), 259–306 (2022)
6. Leng, W., Ju, L.: Trace transfer-based diagonal sweeping domain decomposition method for the Helmholtz equation: Algorithms and convergence analysis. J. Comput. Phys. **455**, 110980 (2022)
7. Taus, M., Zepeda-Núñez, L., Hewett, R.J., Demanet, L.: L-sweeps: A scalable, parallel preconditioner for the high-frequency Helmholtz equation. J. Comput. Phys. **420**, 109706 (2020)