OPTIMIZED SCHWARZ METHODS WITH OVERLAP FOR THE HELMHOLTZ EQUATION

MARTIN J. GANDER* AND HUI ZHANG†

Abstract. Optimized Schwarz methods are based on optimized transmission conditions between subdomains and can have substantially improved convergence behavior compared to classical Schwarz methods. This is especially true when the method is applied to the Helmholtz equation, and better transmission conditions in form of perfectly matched layers have for example led to the new class of sweeping preconditioners. We present here for the first time a complete analysis of optimized Schwarz methods with overlap for the Helmholtz equation. We obtain closed form asymptotically optimized transmission conditions for the case of two subdomains, and study numerically the influence of the number of subdomains on this optimized choice.

Key words. Overlapping Optimized Schwarz Methods, Helmholtz Equation

AMS subject classifications. 65N55, 65N22, 65F10

1. Introduction.

The Helmholtz equation, also known as the time-harmonic wave equation, is used for modeling acoustic and electromagnetic waves arising in many applications, like seismic inversion, antenna design, and tomography in medicine. The numerical solution of the Helmholtz equation is difficult because of the highly oscillatory nature of its solutions. When the number of oscillations becomes large, engineering practice often suggests to use a fixed number of degrees of freedom per oscillation, but this is not enough for accuracy because of the well known pollution effect [1]. The oscillatory kernel (Green's function) of the Helmholtz equation makes any approximation that is essential for iterative methods difficult, see e.g. [2], and this becomes even harder in the presence of strong reflections.

Optimized Schwarz methods, see e.g. [3], [4], [5], [6], [7], [8], [9], use Robin or higher order transmission conditions on interfaces between subdomains, while the parameters contained in these transmission conditions are optimized for rapid convergence. For the Helmholtz equation, a classical first-order absorbing condition was used without overlap in [10] and with overlap in [11], [12]. Without overlap, optimized Robin parameters can be found in [13] with a simplified proposal also for a higher order optimized transmission condition, which then was fully optimized in [14]. Related work for Maxwell's equations can be found in e.g. [15], [16], [17]. One can also use perfectly matched layers or other high-order transmission conditions, see [18, 19, 20, 21]. Recently, it has become popular to use these conditions in a double sweep manner; c.f. [22], [23], [24], [25], [26]. The difference of the double sweep and the parallel optimized Schwarz methods is similar to the difference of the symmetric Gauss-Seidel and the Jacobi iterative methods. The double sweep version requires however a partition into subdomains that are linked chain-wise, i.e. a one dimensional sequence of subdomains, whereas the parallel version does not rely on any special type of partitioning. We will study in this paper mainly the parallel version, but also test the double sweep version in subsection 4.4.

*Section of Mathematics, University of Geneva, 1211 Geneva 4, Switzerland (Martin.Gander@unige.ch).

†Corresponding author. Key Laboratory of Oceanographic Big Data Mining & Application of Zhejiang Province, Zhejiang Ocean University, Zhoushan 316022, China (huiz@zjou.edu.cn); Section of Mathematics, University of Geneva, 1211 Geneva 4, Switzerland. His work was partially supported by Research Start Funding of Zhejiang Ocean University and by the National Natural Science Foundation of China (Grant No. 11371287).

We analyze here for the first time the overlapping case of optimized Schwarz methods for the Helmholtz equation, and solve the corresponding optimization problems asymptotically. Numerically we find a new phenomenon that appears as soon as there are more than two subdomains with this overlapping variant of the algorithm. Our study indicates that in this case a much harder best approximation problem involving many subdomains would need to be solved. Indeed, the visualization of the spectra and the convergence factors based on Fourier analysis for many subdomains show that the number of subdomains is much more influential on the propagating modes than on the evanescent modes (see Figure 4.2 and Figure 4.3). In other words, the new insight is that we should put more weight on the propagating modes than on the evanescent modes for minimizing the convergence factor. Despite the fact that we do not explore the many-subdomain optimization problem in this paper further, the new insight did lead us to a heuristic modification (subsection 4.2) of the asymptotic formulas from section 3, and our numerical experiments show that the heuristic works very well for the open cavity problem. We also compare this approach with a higher-order transmission condition for solving an open cavity problem with spatially varying wave speed, and the results show that the optimized second-order transmission condition is arguably competitive with the higher-order condition.

We organize the paper as follows: we first introduce the Schwarz method with overlap in section 2 and derive the convergence factor. Then, we study the optimization problem in various settings in section 3. One-sided zeroth-order parameters with unequal real and imaginary parts are optimized in subsection 3.1, and we find that the optimal parameters have almost equal real and imaginary parts. A similar result for Maxwell's equations was found in [17]. Based on this observation, we then optimize the two-sided parameters with equal real and imaginary parts in subsection 3.2. The two-sided optimized transmission condition is then transformed to the second-order form and compared with some other conditions in subsection 3.3. We illustrate our results for two subdomains and study also the many subdomain case with numerical experiments in section 4.

2. Schwarz method with overlap. As a model problem, we consider the Helmholtz equation

$$(\omega^2 + \Delta)u = f(x, \mathbf{y}), \qquad u, f : \mathbb{R} \times \mathbb{R}^{d-1} \to \mathbb{C}, \quad \omega \in \mathbb{C},$$

equipped with the Sommerfeld radiation condition

$$\lim_{r\to\infty} r^{\frac{d-1}{2}}(\frac{\partial u}{\partial r} - i\omega) = 0,$$

where $r = \sqrt{x^2 + |\mathbf{y}|^2}$. As it is usually done to study optimized Schwarz methods, we decompose the domain into two overlapping subdomains $\Omega_1 = (-\infty, L) \times \mathbb{R}^{d-1}$ and $\Omega_2 = (0, \infty) \times \mathbb{R}^{d-1}$ with overlap size $L > 0$. We want to compute $u_l := u|_{\Omega_l}$ $(l = 1, 2)$ iteratively by an optimized Schwarz method, which means, with the superscripts being the iteration numbers, we compute

$$\begin{aligned} \omega^2 u_1^{n+1} + \Delta u_1^{n+1} &= f(x, \mathbf{y}), && (x, \mathbf{y}) \in \Omega_1, \\ (\partial_x + \mathcal{S}_1)(u_1^{n+1})(L, \mathbf{y}) &= (\partial_x + \mathcal{S}_1)(u_2^n)(L, \mathbf{y}), && \mathbf{y} \in \mathbb{R}^{d-1}, \end{aligned}$$

and

$$\begin{aligned} \omega^2 u_2^{n+1} + \Delta u_2^{n+1} &= f(x, \mathbf{y}), && (x, \mathbf{y}) \in \Omega_2, \\ (-\partial_x + \mathcal{S}_2)(u_2^{n+1})(0, \mathbf{y}) &= (-\partial_x + \mathcal{S}_2)(u_1^n)(0, \mathbf{y}), && \mathbf{y} \in \mathbb{R}^{d-1}, \end{aligned}$$

where $\mathcal{S}_j$, $j = 1, 2$ are two linear operators in some trace spaces along $\{L\} \times \mathbb{R}^{d-1}$ and $\{0\} \times \mathbb{R}^{d-1}$. For the analysis it suffices to show convergence of the errors $\epsilon_l^n := u_l - u_l^n$ to zero. By the superposition principle for linear equations, the errors satisfy the Helmholtz equation with $f(x, \mathbf{y}) = 0$. We take a Fourier transform in the $\mathbf{y}$ direction to obtain

$$
\begin{aligned}
&(\omega^2 - |\mathbf{k}|^2)\hat{\epsilon}_1^{n+1} + \partial_{xx}^2 \hat{\epsilon}_1^{n+1} = 0, && x \in (-\infty, L),\\
&(\partial_x + s_1)(\hat{\epsilon}_1^{n+1})(L, \mathbf{k}) = (\partial_x + s_1)(\hat{\epsilon}_2^{n})(L, \mathbf{k}),
\end{aligned} \tag{2.1}
$$

and

$$
\begin{aligned}
&(\omega^2 - |\mathbf{k}|^2)\hat{\epsilon}_2^{n+1} + \partial_{xx}^2 \hat{\epsilon}_2^{n+1} = 0, && x \in (0, \infty),\\
&(-\partial_x + s_2)(\hat{\epsilon}_2^{n+1})(0, \mathbf{k}) = (-\partial_x + s_2)(\hat{\epsilon}_1^{n})(0, \mathbf{k}),
\end{aligned} \tag{2.2}
$$

where $\mathbf{k}$ is the Fourier variable of $\mathbf{y}$ and s_j denotes the symbol of $\mathcal{S}_j$. For example, $\mathcal{S}_j = -i\omega$ corresponds to $s_j = -i\omega$, and $\mathcal{S}_j = -i\omega + \frac{1}{2i\omega}\Delta_{\mathbf{y}}$ corresponds to $s_j = -i\omega - \frac{|\mathbf{k}|^2}{2i\omega}$. Since the Sommerfeld radiation condition excludes growing solutions at infinity as well as incoming modes from infinity we obtain the solutions

$$
\begin{aligned}
\hat{\epsilon}_1^{n+1}(x, \mathbf{k}) &= \hat{\epsilon}_1^{n+1}(L, \mathbf{k}) e^{\lambda(\mathbf{k})(x-L)},\\
\hat{\epsilon}_2^{n+1}(x, \mathbf{k}) &= \hat{\epsilon}_2^{n+1}(0, \mathbf{k}) e^{-\lambda(\mathbf{k})x},
\end{aligned}
$$

where $\lambda(\mathbf{k})$ denotes the root of the characteristic equation $\lambda^2 + (\omega^2 - |\mathbf{k}|^2) = 0$ with positive real part or negative imaginary part,

$$
\lambda(\mathbf{k}) := \sqrt{|\mathbf{k}|^2 - \omega^2} \text{ for } |\mathbf{k}| \geq \omega, \qquad \lambda(\mathbf{k}) := -i\sqrt{\omega^2 - |\mathbf{k}|^2} \text{ for } |\mathbf{k}| < \omega.
$$

Substitution of the solutions into the transmission conditions (2.1) and (2.2) yields

$$
\begin{aligned}
\hat{\epsilon}_1^{n+1}(L, \mathbf{k}) &= \frac{s_1(\mathbf{k}) - \lambda(\mathbf{k})}{s_1(\mathbf{k}) + \lambda(\mathbf{k})} e^{-\lambda(\mathbf{k})L} \hat{\epsilon}_2^{n}(0, \mathbf{k}),\\
\hat{\epsilon}_2^{n+1}(0, \mathbf{k}) &= \frac{s_2(\mathbf{k}) - \lambda(\mathbf{k})}{s_2(\mathbf{k}) + \lambda(\mathbf{k})} e^{-\lambda(\mathbf{k})L} \hat{\epsilon}_1^{n}(L, \mathbf{k}).
\end{aligned}
$$

By recursion we have $\hat{\epsilon}_1^{n+1}(L, \mathbf{k}) = \rho(\mathbf{k})\hat{\epsilon}_1^{n-1}(L, \mathbf{k})$ and $\hat{\epsilon}_2^{n+1}(0, \mathbf{k}) = \rho(\mathbf{k})\hat{\epsilon}_2^{n-1}(0, \mathbf{k})$, where the convergence factor ρ for a double iteration is defined by

$$
\rho(\mathbf{k}) = \frac{s_1(\mathbf{k}) - \lambda(\mathbf{k})}{s_1(\mathbf{k}) + \lambda(\mathbf{k})} \cdot \frac{s_2(\mathbf{k}) - \lambda(\mathbf{k})}{s_2(\mathbf{k}) + \lambda(\mathbf{k})} e^{-2\lambda(\mathbf{k})L}. \tag{2.3}
$$

Setting the two complex parameters $s_1 = p_1 - iq_1$ and $s_2 = p_2 - iq_2$, with $p_j, q_j \in \mathbb{R}$, and inserting s_1 and s_2 into the convergence factor (2.3), we find after simplifying

$$
|\rho(p_1, q_1, p_2, q_2, \mathbf{k})|^2 = \begin{cases} \frac{p_1^2 + (q_1 - \sqrt{\omega^2 - |\mathbf{k}|^2})^2}{p_1^2 + (q_1 + \sqrt{\omega^2 - |\mathbf{k}|^2})^2} \frac{p_2^2 + (q_2 - \sqrt{\omega^2 - |\mathbf{k}|^2})^2}{p_2^2 + (q_2 + \sqrt{\omega^2 - |\mathbf{k}|^2})^2}, & |\mathbf{k}|^2 \leq \omega^2, \\ \frac{q_1^2 + (p_1 - \sqrt{|\mathbf{k}|^2 - \omega^2})^2}{q_1^2 + (p_1 + \sqrt{|\mathbf{k}|^2 - \omega^2})^2} \frac{q_2^2 + (p_2 - \sqrt{|\mathbf{k}|^2 - \omega^2})^2}{q_2^2 + (p_2 + \sqrt{|\mathbf{k}|^2 - \omega^2})^2} e^{-4\sqrt{|\mathbf{k}|^2 - \omega^2}L}, & |\mathbf{k}|^2 > \omega^2. \end{cases} \tag{2.4}
$$

As long as $|\mathbf{k}| \neq \omega$ and $p_j, q_j > 0$, we have $|\rho| < 1$. We also see from this convergence factor that the overlap L only helps for $|\mathbf{k}|^2 > \omega^2$ (corresponding to the evanescent modes), while for $|\mathbf{k}|^2 < \omega^2$ (corresponding to the propagating modes) the convergence

is not affected by the overlap in this two subdomain analysis. From (2.3), exact convergence in two steps is achievable if and only if $s_j(\mathbf{k}) = \lambda(\mathbf{k})$ for both $j = 1, 2$. Since this optimal choice leads to non-local operators in real space which are expensive to use, one in general prefers to use approximations of $\lambda(\mathbf{k})$ leading to local operators instead. The quality of the approximation is naturally measured by (2.4). This type of problem occurs also in studying absorbing boundary conditions for domain truncation (see e.g. [27, 28]), where the convergence factor becomes the relative error and L is the distance of the source to the truncation boundary.

REMARK 1. *Our analysis works almost equally for $\Omega = \mathbb{R} \times R$ ($R \subset \mathbb{R}^{d-1}$) as long as the boundary condition on $\mathbb{R} \times \partial R$ permits a diagonalization of $\Delta_{\mathbf{y}}$ into negative real spectra. For instance $R = (0, 1)$ with Dirichlet boundary conditions was analyzed in [14]. A harder situation arises from truncation of free space $\mathbb{R}^d$, which introduces complex eigenvalues of $\Delta_{\mathbf{y}}$ and can not be covered by our present analysis.*

3. Optimized transmission conditions. Our goal is to find good parameters p_j, q_j such that the modulus of the convergence factor in (2.4) is as small as possible over a range of frequencies $|\mathbf{k}| \in [k_{\min}, k_-] \cup [k_+, k_{\max}]$, where $\mathbf{k}_- < \omega < \mathbf{k}_+$. We require $|\mathbf{k}|$ to be away from ω because $|\rho| = 1$ when $|\mathbf{k}| = \omega$, independently of what one chooses for the parameters p_j and q_j. Since in general we do not know how the Fourier coefficients of the initial error are distributed over the frequencies, in optimized Schwarz methods one minimizes $|\rho|$ for the worst case, that is, we solve

$$\underset{(p_1,q_1,p_2,q_2)\in\mathbb{P}}{\operatorname{argmin}} \left(\max_{|\mathbf{k}|\in[k_{\min},k_-]\cup[k_+,k_{\max}]} |\rho(p_1, q_1, p_2, q_2, \mathbf{k})|^2 \right), \tag{3.1}$$

where $\mathbb{P}$ is the search domain of the parameters and $|\rho|^2$ is given in (2.4). For well-posedness of the problems on Ω_1 and Ω_2, we should choose $\mathbb{P} \subset [0, \infty)^2$. The min-max problem (3.1) is too difficult to solve in closed form. Instead, we give asymptotic formulas for the parameters such that the convergence factor is as small as possible in different limiting processes.

In subsection 3.1 and subsection 3.2, we will consider zeroth-order transmission conditions with the symbol $s_j = p_j - iq_j$ *independent of* $\mathbf{k}$ and the corresponding operator $\mathcal{S}_j$ reduces to the scalar s_j. We can then transform the zeroth-order operators to the one-sided second-order operators by the formulas of [3]: $\tilde{\mathcal{S}}_1 = \tilde{\mathcal{S}}_2 = r_0 - r_2 \Delta_{\mathbf{y}}$ with $r_0 = \frac{-\omega^2 + s_1 s_2}{s_1 + s_2}$, $r_2 = \frac{1}{s_1 + s_2}$. From subsection 3.3 to the end of this paper, we will always study the second-order form of the optimized transmission conditions. In this case the convergence factor becomes

$$\tilde{\rho}(\mathbf{k}) = \left(\frac{s_1 - \lambda(\mathbf{k})}{s_1 + \lambda(\mathbf{k})} \cdot \frac{s_2 - \lambda(\mathbf{k})}{s_2 + \lambda(\mathbf{k})} \right)^2 e^{-2\lambda(\mathbf{k})L}.$$

Note that the modulus $|\tilde{\rho}|$ is just (2.4) with L replaced by $L/2$.

3.1. One-sided parameters. In this subsection, we optimize the one-sided zeroth-order operators $\mathcal{S}_1 = \mathcal{S}_2 = s_1 = s_2 = p - \mathbf{i}q$ with p, q independent of $\mathbf{k}$. With this ansatz, we will give without proof the solution of (3.1) for small mesh sizes and for large wavenumbers. We will see in every case that the optimized parameters p^* and q^* are close to being equal (i.e. $p^* \approx q^*$): their leading orders are always the same and their leading terms are approximately equal as long as the excluding interval $[k_-, k_+]$ is not strongly unsymmetric about ω. This motivates us to choose

$q_j = p_j$ for the two-sided parameters in subsection 3.2.

THEOREM 3.1. *Let* $L = C_L h$, $k_{\max} \in [C/h, \infty]$, $C_L, C, k_{\min}, k_-, k_+$ *and* ω *be positive and independent of* h, $k_{\min} < k_- < \omega$, $k_{\max} > k_+ > \omega$ *and* $\mathbb{P} = \{(p,q,p,q)\,|\,p,q \in [0,\infty) \text{ independent of } \mathbf{k}\}$. *Denote by*

$$x_- = \sqrt{\omega^2 - k_-^2}, \quad y_+ = \sqrt{k_+^2 - \omega^2}.$$

For small h, $y_+ > x_-$ *and* $x_- \le 1/(2y_+)$, *any solution of* (3.1) *must satisfy* $p = p^* = q = q^* = x_-^{2/3} L^{-1/3}/2 + o(h^{-1/3})$ *at which* $\max_{\mathbf{k}} |\rho| = 1 - 4(Lx_-)^{1/3} + o(h^{1/3})$; *otherwise, the solution for small* h *must satisfy*

$$\begin{aligned} p = p^* &= \tfrac{x_-^{4/3} y_+^{2/3}}{(x_-^2+y_+^2)^{3/2}} (2L)^{-1/3} + o(h^{-1/3}), \\ q = q^* &= \tfrac{x_-^{1/3} y_+^{5/3}}{(x_-^2+y_+^2)^{3/2}} (2L)^{-1/3} + o(h^{-1/3}), \end{aligned}$$

at which $\max_{\mathbf{k}} |\rho| = 1 - \frac{4x_-^{2/3} y_+^{1/3} (2L)^{1/3}}{(x_-^2+y_+^2)^{1/3}} + o(h^{1/3})$.

THEOREM 3.2. *Let the overlap* $L = C_L h$, $h = C_h/\omega^\gamma$, $\gamma \in [1,2]$, $k_{\max} \in [C/h, \infty]$, $\delta_\omega = \min\{\omega - k_-, k_+ - \omega\}$ *with* C_L, C_h, C, k_- *and* k_+ *positive constants independent of* ω, $k_{\min} < k_- < \omega$, $k_{\max} > k_+ > \omega$ *and* $\mathbb{P} = \{(p,q,p,q)\,|\,p,q \in [0,\infty) \text{ independent of } \mathbf{k}\}$. *Denote by* $\delta_\pm = \omega - k_\pm$. *Suppose* ω *is large, then up to higher-order remainders, for* $1 \le \gamma < 5/4$ *any solution of* (3.1) *must satisfy*

$$p = p^* = \begin{cases} \frac{2^{1/4}\delta_-^{3/4}}{\sqrt{\delta_- + \delta_+}} \omega^{3/4}, & \delta_+ \ge \delta_-, \\ (\delta_+/2)^{1/4} \omega^{3/4}, & \text{otherwise}, \end{cases} \quad q = q^* = \begin{cases} \frac{2^{1/4}\delta_-^{1/4}\delta_+^{1/2}}{\sqrt{\delta_- + \delta_+}} \omega^{3/4}, & \delta_+ \ge \delta_-, \\ p^*, & \text{otherwise}, \end{cases}$$

at which

$$\max_{\mathbf{k}} |\rho| = \begin{cases} 1 - \frac{2^{9/4}\delta_+^{1/2}\delta_-^{1/4}}{\sqrt{\delta_- + \delta_+}} \omega^{-1/4}, & \delta_+ \ge \delta_-, \\ 1 - 4(\delta_+/2)^{1/4} \omega^{-1/4}, & \text{otherwise}; \end{cases}$$

for $5/4 < \gamma \le 2$ *any solution must satisfy*

$$\begin{aligned} p = p^* &= \begin{cases} 2^{-1/6} (\frac{\delta_-}{C_L C_h})^{1/3}, & \delta_+ \ge \delta_-, \\ (\frac{\delta_-}{\delta_+ + \delta_-})^{2/3} (\delta_+ \omega / L)^{1/3}, & \text{otherwise}, \end{cases} \\ q = q^* &= \begin{cases} p^*, & \delta_+ \ge \delta_-, \\ (\frac{\delta_+}{\delta_+ + \delta_-})^{2/3} (\sqrt{\delta_+ \delta_-}\, \omega / L)^{1/3}, & \text{otherwise}, \end{cases} \end{aligned}$$

at which

$$\max_{\mathbf{k}} |\rho| = \begin{cases} 1 - 4(2\delta_-)^{1/6} (C_L C_h)^{1/3} \omega^{1/6 - \gamma/3}, & \delta_+ \ge \delta_-, \\ 1 - 4\sqrt{2} (\frac{\delta_- C_L C_h}{\delta_- + \delta_+})^{1/3} \delta_+^{1/6} \omega^{1/6 - \gamma/3}, & \text{otherwise}. \end{cases}$$

REMARK 2. *The case of* $\gamma = 5/4$ *is too complicated to be shown here.*

3.2. Two-sided parameters. In this subsection, we optimize the two-sided zeroth-order operators $\mathcal{S}_l = s_l = p_l(1-i)$, $l = 1, 2$, with p_l independent of $\mathbf{k}$. Note that we have chosen equal real and imaginary parts, which is motivated by the observation in subsection 3.1. Moreover, if $p_1 = p_1^*$, $p_2 = p_2^*$ solve (3.1), then $p_1 = p_2^*$, $p_2 = p_1^*$

also do. This is a consequence of the symmetric setting of the subdomains Ω_1 and Ω_2. Therefore, between the two possible solutions attaining the same convergence factor we are free to choose either of them. In the following, we simply assume that p_1 used for Ω_1 is smaller than p_2 for Ω_2. We state our main results in Theorem 3.3 and Theorem 3.4. The proofs of the two theorems are lengthy, so we put them into the Appendix.

THEOREM 3.3. *Let $L = C_L h$, $k_{\max} \in [C/h, \infty]$, $C_L, C, k_{\min}, k_-, k_+$ and ω be positive and independent of h, $k_{\min} < k_- < \omega$, $k_{\max} > k_+ > \omega$ and $\mathbb{P} = \{(p_1, p_1, p_2, p_2) \,|\, 0 \le p_1 \le p_2 < \infty$ independent of $\mathbf{k}\}$. Suppose h is small, then any solution of (3.1) must satisfy*

$$\begin{aligned} p_1 &= p_1^* = C_\omega^{2/5}\,(4L)^{-1/5}/2 + o(h^{-1/5}), \\ p_2 &= p_2^* = C_\omega^{1/5}\,(4L)^{-3/5} + o(h^{-3/5}), \end{aligned} \tag{3.2}$$

where $C_\omega = \min\left\{\omega^2 - k_-^2, k_+^2 - \omega^2\right\}$, and $\max_{\mathbf{k}} |\rho|^2 = 1 - 4\,(4L\sqrt{C_\omega})^{1/5} + o(h^{1/5})$ is attained at the given parameters.

REMARK 3. *In practice, we use only the leading terms of the optimized parameters. But it is also possible to derive higher order terms.*

THEOREM 3.4. *Let the overlap $L = C_L h$, $h = C_h/\omega^\gamma$, $\gamma \ge 1$, $k_{\max} \in [C/h, \infty]$, $\delta_\omega = \min\{\omega - k_-, k_+ - \omega\}$ with C_L, C_h, C, k_- and k_+ positive constants independent of ω, $k_{\min} < k_- < \omega$, $k_{\max} > k_+ > \omega$ and $\mathbb{P} = \{(p_1, p_1, p_2, p_2) \,|\, 0 \le p_1 \le p_2 < \infty$ independent of $\mathbf{k}\}$. Suppose ω is large, then, for $1 \le \gamma < 9/8$ any solution of (3.1) must satisfy*

$$\begin{aligned} p_1 &= p_1^* = \delta_\omega^{3/8}\,(\omega/2)^{5/8} + o(\omega^{5/8}), \\ p_2 &= p_2^* = (2\delta_\omega)^{1/8}\,\omega^{7/8} + o(\omega^{7/8}), \end{aligned}$$

at which $\max_{\mathbf{k}} |\rho|^2 = 1 - 4\cdot 2^{1/8}\,\delta_\omega^{1/8}\,\omega^{-1/8} + o(\omega^{-1/8})$; for $\gamma > 9/8$ any solution must satisfy

$$\begin{aligned} p_1 &= p_1^* = (\delta_\omega\,\omega)^{2/5}\,L^{-1/5}/2 + o(\omega^{2/5+\gamma/5}), \\ p_2 &= p_2^* = (\delta_\omega\,\omega)^{1/5}\,L^{-3/5}/2 + o(\omega^{1/5+3\gamma/5}), \end{aligned}$$

at which $\max_{\mathbf{k}} |\rho|^2 = 1 - 4\sqrt{2}\,(C_h C_L)^{1/5}\,\delta_\omega^{1/10}\omega^{1/10-\gamma/5} + o(\omega^{1/10-\gamma/5})$; and for $\gamma = 9/8$ any solution must satisfy

$$\begin{aligned} p_1 &= p_1^* = C_1\,\omega^{5/8} + o(\omega^{5/8}),\ C_1 \in [\sqrt{2\,\delta_\omega}/(8\,C_h C_L), 32\,C_h^3 C_L^3], \\ p_2 &= p_2^* = C_h C_L\,\omega^{7/8} + o(\omega^{7/8}), \end{aligned}$$

at which the following value of $\max_{\mathbf{k}} |\rho|^2$ is attained

$$\begin{cases} 1 - 16\,C_h C_L\,\omega^{-1/8} + o(\omega^{-1/8}), & \text{if } 2^{-15/8}\delta_\omega^{1/8} \le C_h C_L, \\ 1 - 2\sqrt{2}\,\delta_\omega^{1/6} C_h^{-1/3} C_L^{-1/3}\,\omega^{-1/8} + o(\omega^{-1/8}), & \text{if } 2^{-15/8}\delta_\omega^{1/8} \ge C_h C_L. \end{cases}$$

Moreover, for $\gamma = 9/8$, if we take $C_1 = (C_h C_L \delta_\omega)^{1/3}$, then we are simultaneously minimizing the maximum of the other local (but not global) maxima.

REMARK 4. *The scaling $h = C_h/\omega^\gamma$ with $\gamma > 1$ is deemed necessary for accurate low-order discretizations, see e.g. [29]. At first glance, one may think that C_h can be chosen independently of the spatial domain size H because h scales directly with the wavelength $\lambda = \frac{2\pi}{\omega}$. On the other hand, the Helmholtz problem is essentially the same*

if we scale up H and keep ωH constant (i.e. number of wavelength H/λ is fixed); if we use the same C_h for different H, we will see that $h = O(H^\gamma)$ or $H/h = O(H^{1-\gamma})$. This does not make sense as $\gamma > 1$: given that we are solving the same problem, why can we use less number of mesh points for bigger H? So we conclude that C_h must depend also on H and we can take $C_h = H^{1-\gamma}\widehat{C}_h$ with $\widehat{C}_h$ the constant that one would use for the unit domain.

REMARK 5. *It is also possible to solve the min-max problem (3.1) numerically, but this is substantially more expensive than using our asymptotic formulas, and this becomes important for spatially varying wave speed, where in the frozen coefficient approach one generates the optimized parameters at many spatial locations using the local wave speed. Numerical optimization is still useful when no formulas are available, or for a comparison with the formulas. In our study for example, if we solve numerically (3.1) in $\mathbb{P} = \{(p_1, q_1, p_2, q_2) : p_j, q_j \geq 0, j = 1, 2\}$ and compare with our formulas in $\mathbb{P} = \{(p_1, p_1, p_2, p_2) : p_1, p_2 \geq 0\}$, we obtain for $\omega = 20, 200, 2000, 20000$, $L = h = \frac{1}{2\omega}$, $k_{\min} = \pi$, $k_- = \omega - \pi$, $k_+ = \omega + \pi$, $k_{\max} = \pi/h$ using 'fminsearch' of Matlab the convergence factors 0.1056, 0.2723, 0.3623, 0.4640, while our formula for the scaling $\omega h = 1/2$ gives the convergence factors 0.2366, 0.3184, 0.4032, 0.4939.*

3.3. Comparison of convergence factors. We will compare the convergence factors of the following methods, which differ only in the transmission conditions and/or overlap. The overlap size is $L = 2h$. The TO0 method will not be compared here but later in section 4.

- *Classical Schwarz*: using the transmission condition $u_1^{n+1} = u_2^n$ on $\partial\Omega_1 \cap \Omega_2$ and $u_2^{n+1} = u_1^n$ on $\partial\Omega_2 \cap \Omega_1$ for iteration $(n+1)$, which is equivalent to using $\mathcal{S}_1 = \mathcal{S}_2 = \infty$.
- *TO0*: Taylor zeroth-order method using $\mathcal{S}_1 = \mathcal{S}_2 = -i\omega$, which corresponds to the zeroth-order Taylor expansion of the square-root symbol $-i\sqrt{\omega^2 - |\mathbf{k}|^2}$.
- *TO2*: Taylor second-order method using $\mathcal{S}_1 = \mathcal{S}_2 = -i\omega + \frac{1}{2i\omega}\Delta_{\mathbf{y}}$, which corresponds to the second-order Taylor expansion of the square-root symbol $-i\sqrt{\omega^2 - |\mathbf{k}|^2}$.
- *Zolotarev*: using the two-sided transmission operators $\mathcal{S}_1 = -i(r_0 - r_2\Delta_{\mathbf{y}})$ $(r_0, r_2 \in \mathbb{R})$ for Ω_1 and $\mathcal{S}_2 = r_0' - r_2'\Delta_{\mathbf{y}}$ $(r_0', r_2' \in \mathbb{R})$ for Ω_2, where r_0, r_2 are generated by the method of [28] for $\{\mathbf{k} : k_{\min} \leq |\mathbf{k}| \leq k_- < \omega\}$ and r_0', r_2' for $\{\mathbf{k} : \omega < k_+ \leq |\mathbf{k}| \leq k_*\}$, more precisely by the Matlab-like pseudo-code

$$
\begin{aligned}
&a = \sqrt{\omega^2 - k_-^2}; \quad b = \sqrt{\omega^2 - k_{\min}^2}; \quad K = ellipke(1 - a^2/b^2);\\
&[S, C, D] = ellipj(3K/4, 1 - a^2/b^2); \quad p_1 = bD;\\
&[S, C, D] = ellipj(K/4, 1 - a^2/b^2); \quad p_2 = bD;\\
&r_0 = (-p_1p_2 - \omega^2)/(p_1 + p_2); \;\; r_2 = 1/(p_1 + p_2);\\
&a = \sqrt{k_+^2 - \omega^2}; \quad b = \sqrt{k_*^2 - \omega^2}; \quad K = ellipke(1 - a^2/b^2);\\
&[S, C, D] = ellipj(3K/4, 1 - a^2/b^2); \quad p_1 = bD;\\
&[S, C, D] = ellipj(K/4, 1 - a^2/b^2); \quad p_2 = bD;\\
&r_0' = (p_1p_2 - \omega^2)/(p_1 + p_2); \;\; r_2' = 1/(p_1 + p_2);
\end{aligned}
$$

- *OO2*: the optimized Schwarz method using the one-sided second-order operators $\mathcal{S}_1 = \mathcal{S}_2 = r_0 - r_2\Delta_{\mathbf{y}}$ without overlap [14] and with overlap (this paper), where $r_0 = \frac{-\omega^2 - 2ip_1p_2}{(p_1+p_2)(1-i)}$, $r_2 = \frac{1}{(p_1+p_2)(1-i)}$ and p_1, p_2 are from the two-sided

TABLE 3.1

Maxima of the convergence factors for the propagating modes $|\mathbf{k}| \in [\pi, \omega - \pi]$ and the evanescent modes $|\mathbf{k}| \in [\omega + \pi, \pi/h]$, with $\omega h = \pi/30$, overlap size $L = 2h$

$\|\mathbf{k}\| \in [\pi, \omega - \pi]$					
ω	Class.	TO2	OO2 ($L = 0$)	OO2	Zolotarev
20π	1.000000	0.075449	0.290048	0.076742	0.001603
200π	1.000000	0.448578	0.387287	0.129444	0.018937
2000π	1.000000	0.776438	0.488563	0.200557	0.067667
20000π	1.000000	0.923114	0.583468	0.290742	0.147525
$\|\mathbf{k}\| \in [\omega + \pi, \pi/h]$					
ω	Class.	TO2	OO2 ($L = 0$)	OO2	Zolotarev
20π	0.874496	(see left)	0.286252	0.083000	0.080959
200π	0.958927		0.386869	0.128233	0.156277
2000π	0.986840		0.488527	0.197864	0.256676
20000π	0.995819		0.583466	0.289484	0.345952

zeroth-order formulas corresponding to the scaling $h = C_h/\omega$ (Theorem 4.2 in [14] and Theorem 3.4 in this paper).

Note that when computing the Zolotarev parameters of $\mathcal{S}_2$, the method of [28] does not take into account the overlap, because the goal function to be minimized in [28] does not contain the exponential factor as in (2.3). So we decided to modify the value of k_* or the real upper bound of the square-root value $\sqrt{k_*^2 - \omega^2}$ to be put into the method of [28]. The idea behind is that the overlap can take care of the higher spatial frequencies $\{\mathbf{k} : |\mathbf{k}| > k_*\}$ due to the exponential decay. We perform the modification simply by hand until the min-max problem for our actual convergence factor (2.3) over the actual frequencies $\{\mathbf{k} : k_+ \leq |\mathbf{k}| \leq k_{\max}\}$ ($k_{\max} > k_*$) is best solved by the Zolotarev parameters.

The convergence factors are plotted in Figure 3.1. We can see that the OO2 methods have balanced convergence factors for the propagating and the evanescent modes. In particular, the asymptotics of the overlapping OO2 has not set in for the lower mesh density (first column) but has for the higher mesh density (second column). We can also see that Zolotarev's parameters solve the best approximation problem perfectly for the propagating modes; the behavior for the evanescent modes is almost as good, but a closed formula is yet to be found in this case with overlap. For higher-order conditions without overlap, Zolotarev's parameters can be balanced for all the modes, see [28]. Note that we are comparing the *one-sided* second-order conditions OO2 with the *two-sided* second-order condition of Zolotarev's parameters. Nonetheless, in the second column of Figure 3.1 we find the min-max problem is better solved by OO2 with overlap. It would thus be very interesting to optimize the two-sided second-order transmission conditions with a total of four complex parameters. This task, however, would require tremendous further efforts.

The maximal convergence factors are listed in Table 3.1, from which we see again that the OO2 methods equilibrate well for the propagating and evanescent modes and the overlapping OO2 does the best job for minimizing the maxima overall.

4. Numerical experiments. For the implementation, we use the volume formulation called Optimized Restrict Additive Schwarz (ORAS) described in [4]. Let $\tilde{A}_j$ be the matrix of the j-th subproblem equipped with transmission conditions and

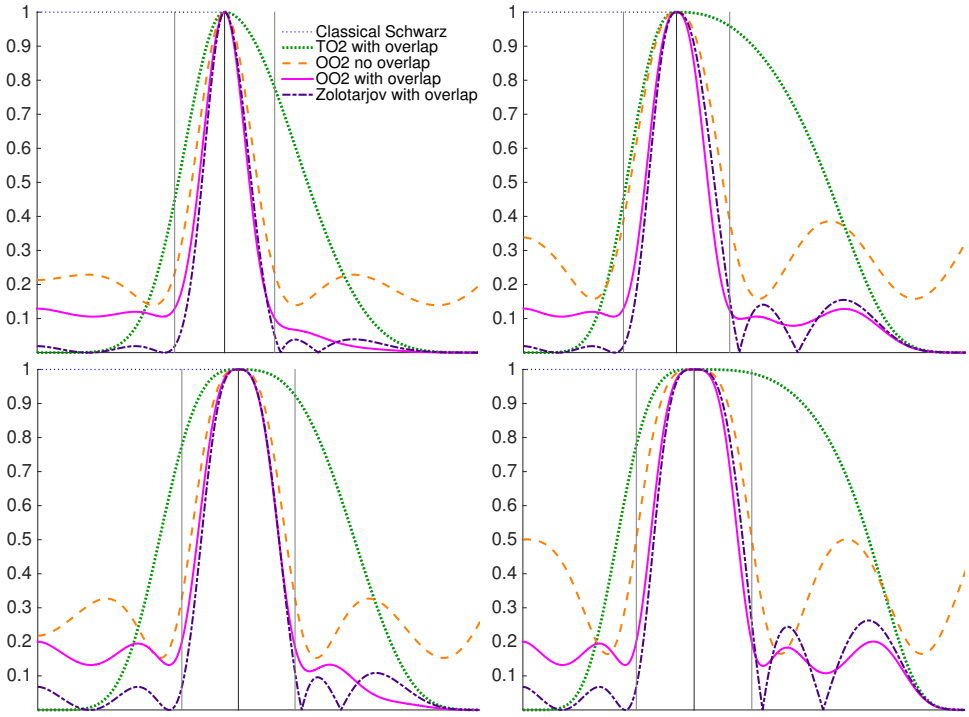

FIG. 3.1. *Convergence factors (vertical axis) against the Fourier parameter* $|\mathbf{k}|$ *(horizontal axis, rescaled from* $|\mathbf{k}| = \omega$ *towards the two ends, to show the details near* $|\mathbf{k}| = \omega$*). The vertical lines indicate the values* ω*,* $\omega - \pi$ *and* $\omega + \pi$ *which are used as excluding interval for the OO2 methods. Top row:* $\omega = 200\pi$*,* $h = 1/10000, 1/30000$*. Bottom row:* $\omega = 2000\pi$ $h = 1/10000, 1/350000$*.*

A be the matrix of the original problem. We assume $\{\Omega_j\}$ is obtained by extending a non-overlapping partition $\{\tilde{\Omega}_j\}$. Let R_j be the restriction from Ω to Ω_j and $\tilde{R}_j$ be the restriction from Ω to $\tilde{\Omega}_j$. Then, the ORAS preconditioner is $\sum_j \tilde{R}_j^T \tilde{A}_j^{-1} R_j$. For high-order transmission conditions, we use $\tilde{A}_j^{-1} = \bar{R}_j \bar{A}_j^{-1} \bar{R}_j^T$ with $\bar{A}_j$ obtained by augmenting Ω_j with auxiliary degrees of freedom introduced by the high-order transmission conditions and $\bar{R}_j$ is the restriction from the augmented subdomain to Ω_j. Alternatively, one can also use the substructured form, see e.g. [30]. The volume form takes less effort for the implementation from our experience and allows inexact subdomain solves, whereas the substructured form iterates with smaller vectors. Yet another form in between is using the volume form as right preconditioning and (when exact subdomain solves are performed) reducing the iterates to their non-zero components (by noting that the iterates are residuals which have compact support near the interfaces), see e.g. [31], [24].

4.1. Two subdomain problems. We solve the homogeneous equation with the zero solution and use a *random* initial guess to have all frequency components present in the error. We use the domain decomposition $\Omega_1 = (0, \frac{1}{2} + h) \times (0, 1)$, $\Omega_2 = (\frac{1}{2} - h, 1) \times (0, 1)$ so that the overlap size $L = 2h$ where h is the mesh size. We iterate until the relative residual of the original system is less than 10^{-8}. We compare the transmission conditions presented in section 3.3 except that the non-overlapping OO2 is replaced (because the overlapping OO2 is better) with the first-order Taylor

TABLE 4.1
Iteration numbers for the open cavity problem, $\omega = 9.5\pi$ (and $\omega = 10\pi$ in parentheses).

	Stationary				GMRES				
$1/h$	TO0	TO2	OO2	Zol	Cl.	TO0	TO2	OO2	Zol
50	34 (67)	35 (70)	14(19)	18(17)	25(26)	16(15)	15(14)	12(14)	11(11)
100	74(227)	84(222)	17(31)	26(43)	30(30)	22(21)	22(22)	13(15)	11(11)
200	166(469)	172(371)	20(44)	37(69)	38(38)	27(28)	32(32)	14(15)	12(12)
400	343(681)	345(455)	20(51)	42(83)	49(51)	33(34)	41(42)	14(15)	13(14)
800	662(864)	717(504)	21(55)	50(97)	67(68)	40(41)	50(52)	16(17)	15(16)

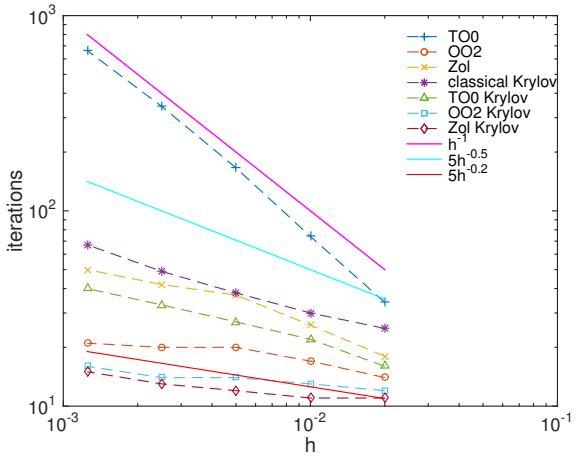

FIG. 4.1. *Asymptotic behavior of the Schwarz methods for the open cavity, $\omega = 9.5\pi$.*

condition ([10], [11], [12]) with overlap denoted by TO0. Since the Schwarz methods can be used as stationary iterative solvers or as a preconditioner for the GMRES iteration, both cases are tested. The classical Schwarz method which uses Dirichlet transmission conditions is not tested as stationary iteration, because in this case it can not converge.

We first consider an open cavity problem with homogeneous Dirichlet boundary conditions on the top and bottom of the unit square and second-order absorbing conditions [32] on the left and right sides.

We fix $\omega = 9.5\pi$ ($\omega = 10\pi$, resp.) which is away from (on, resp.) the eigenvalues of $-\partial_y^2$ on the unit interval with homogeneous Dirichlet conditions. The iteration numbers are listed in Table 4.1. We can see that the minimum distance from ω to the frequencies at the discrete level in the y-direction plays an important role in all the *stationary* iterations, while in the *GMRES* iterations this effect is negligible. Figure 4.1 shows the asymptotic behavior of different Schwarz methods as $h \to 0$, which confirms our Fourier analysis results of Theorem 3.3.

Now we fix $h\omega$ or $h\omega^{3/2}$ constant to see how the Schwarz methods behave for higher and higher wavenumbers, which corresponds to Theorem 3.4. The iteration numbers for GMRES are listed in Table 4.2. We also show in the same Table 4.2 the results for the truncated free space problem with the second-order absorbing conditions at the four sides of the unit square. We can see that the optimized methods

TABLE 4.2
GMRES iteration numbers for $h\omega = \pi/5$ ($h\omega^{3/2} \approx 3.52$ in parentheses).

	open cavity					truncated free space				
$1/h$	Cl.	TO0	TO2	OO2	Zol	Cl	TO0	TO2	OO2	Zol
100	38(34)	20(22)	19(20)	15(13)	13(12)	35(30)	20(21)	18(18)	15(14)	12(13)
200	48(43)	25(27)	22(27)	17(18)	14(15)	43(34)	25(27)	21(24)	16(14)	13(13)
400	69(49)	38(33)	32(32)	18(17)	18(14)	53(41)	28(35)	25(30)	17(15)	14(14)
800	76(70)	42(46)	35(41)	22(15)	18(16)	65(47)	32(43)	30(37)	17(15)	15(14)

TABLE 4.3
GMRES iteration numbers, $h = 1/256$, $\omega = 51.2\pi$, overlap $2h$.

	open cavity					truncated free space				
Sub.	Cl.	TO0	TO2	OO2	Zol	Cl.	TO0	TO2	OO2	Zol
2×1	52	28	24	18 (16, 15)	16 (23)	48	25	22	16	15
4×1	396	68	46	68 (40, 40)	53 (39)	163	29	24	45	40
8×1	–	160	102	162 (87, 87)	127 (84)	–	44	32	108	86
16×1	–	338	212	335(173,174)	257(170)	–	88	65	225	167
2×2	118	66	63	61 (58, 58)	58 (60)	49	27	25	20	19
4×4	2192	184	172	183(167,166)	175(166)	372	38	33	49	41
8×8	–	408	364	379(349,349)	369(360)	–	69	60	107	80
16×16	–	809	732	767(708,706)	759(719)	–	122	112	199	128

OO2 and Zolotarev's parameters converge substantially faster than the others. Note that Zolotarev's parameters are used here as a *two-sided* second-order transmission condition with overlap and no closed-form formula is available to take the overlap into account. Note also that OO2 here is a *one-sided* second-order transmission condition whose asymptotic formulas have been obtained in section 3.

4.2. More than two subdomains. In contrast to the Laplacian case, the local two subdomain analysis does not simply generalize when wave propagation phenomena are present. We thus investigate now numerically the case of more than two subdomains. We monitor the relative residual of the preconditioned system and iterate until it is less than 10^{-8}. The final GMRES iteration numbers are shown in Table 4.3, where we use a bar to represent iteration numbers larger than 3000. We first observe that the classical Schwarz method, which uses Dirichlet transmission conditions, very quickly fails, it is not suitable for Helmholtz type problems. We also see, neglecting the numbers in the parentheses for the moment, that the other methods deteriorate and interestingly the overlapping TO2 method becomes the fastest method for more than two subdomains. In particular, the overlapping optimized methods based on two subdomain analysis including the one using Zolotarev's parameters are not superior any more.

To address this problem, we first note that with increasing number of subdomains the GMRES iteration numbers always increase so it might be worthy to exclude more frequencies from the optimization and let GMRES take care of them. To this end, we choose $\delta_\omega = N\pi$ (N is the number of subdomains in normal direction of the interface concerned) in the min-max problem (3.1) for OO2 (this approach does not improve the performance for Zolotarev's parameters). Second, we note that the propagating modes are the main reason of the non-scalability w. r. t. the number of subdomains

TABLE 4.4
GMRES iteration numbers of the overlapping TO2 method for the truncated free space problem

$\frac{\omega}{2\pi}$	$\frac{1}{h}$	subdomains							
		2×1	4×1	8×1	16×1	4×4	8×8	16×16	32×32
12.8	128(256)	19(27)	22(30)	34(36)	67(67)	32(36)	62(64)	106(118)	166(187)
25.6	256(512)	22(31)	24(33)	32(37)	65(63)	33(39)	60(62)	112(117)	192(214)
51.2	512(1024)	26(35)	28(37)	34(40)	63(60)	34(42)	57(61)	109(112)	198(211)
102.4	1024(2048)	31(40)	33(43)	37(45)	62(61)	38(47)	57(63)	106(109)	197(206)

because the propagating modes can go much further than the evanescent modes. But convergence of the propagating modes is slowed down by large real parts of the transmission parameters. So we can simply diminish the real parts of the OO2 parameters by the factor $1/N$. We let all of Zolotarev's parameters aim only at the propagating modes. For the open cavity problem, we made these changes and the resulting iteration numbers are listed in the parentheses (first column for OO2) in Table 4.3. We then formulated the optimization problems analytically for the many-subdomain model in $\mathbb{R}^2$, like we did it in section 2 for two subdomains. We can solve this optimization problem however only numerically, and we list the iteration numbers as the second column in the parentheses for OO2 in Table 4.3. This works as well as the previous strategy.

For the free space problem, the above adjustments, however, can not make the optimized methods better than TO2 except on less than eight subdomains. So we did more experiments on TO2 and found its remarkable scalability property w. r. t. wavenumber and mesh size in solving the free space problem, see Table 4.4, which we could not expect from Table 4.2. Let us observe how the numbers vary inside each column of Table 4.4 and then how the columns differ from each other. It seems that for a fixed number of subdomains there is a range of ω and h for which the iteration numbers are stable. Intuitively, we think in this regime the global effect of wave propagation determines the convergence speed, which makes the number of subdomains and their adjacent relations more influential than ω and h. When ω and/or $1/h$ become large enough, however, the local effect of the two-subdomain interaction, which we have analyzed in section 2, starts to be significant on many subdomains. This may explain the eventually increasing iteration numbers for the partitions 8×1 and 4×4. When ω is sufficiently small, propagating waves do not exist inside subdomains any more and the solver performance is better on the same partition. This explains for example the increase in the iteration numbers in the 32×32 subdomains case when going from $\frac{\omega}{2\pi}=12.8$ to $\frac{\omega}{2\pi}=25.6$, whereas at all the other wavenumbers the iteration numbers are stable.

Note that for the truncated free space problem, due to the absorbing boundary conditions on the bounds of y, the frequencies in y are actually complex-valued, which we have not taken into account in our analysis. With two-way partitions, we also have non-Dirichlet, non-Neumann boundary conditions on the top and/or bottom of a subdomain, which affects the frequencies in y. Our analysis can also not capture the specific convergence properties of the Krylov method we use, so we next study numerically the best performing parameters and associated spectra of the preconditioned operator.

4.3. Brute force search of best performing parameters. We can find the practically best parameters by running directly the algorithm with a range of param-

TABLE 4.5

Best performing parameters found for the open cavity problem: left (right) panel for one-sided (two-sided) zeroth-order conditions with complex (imaginary) parameters, $\omega h = \pi/5$

$\frac{\omega}{2\pi}$	2×1	4×1	8×1	2×1	4×1	8×1
10	4+ 44**i**	2+ 46**i**	4+ 46**i**	20**i**, 90**i**	35**i**, 65**i**	10**i**, 45**i**
20	10+ 84**i**	6+ 84**i**	2+ 94**i**	85**i**, 339**i**	15**i**, 95**i**	10**i**, 60**i**
40	23+132**i**	0+160**i**	4+182**i**	35**i**, 153**i**	85**i**, 185**i**	10**i**, 120**i**
80	32+230**i**	2+278**i**	-6+345**i**	200**i**, 620**i**	85**i**, 305**i**	36**i**, 230**i**

TABLE 4.6

GMRES iteration numbers using $N\times1$ subdomains for the open cavity problem with $\omega h = \pi/5$; each method has three columns corresponding to $N = 2, 4, 8$

$\frac{\omega}{2\pi}$	TO0			OO0			BO0			TO2			OO2			Zol			BO2		
10	15	39	86	18	43	82	13	33	78	13	29	64	13	28	64	13	28	63	13	27	62
20	19	57	134	20	66	134	16	51	118	15	42	97	15	38	86	15	38	90	15	37	85
40	29	77	168	26	95	199	23	63	144	23	54	112	18	47	101	20	46	101	19	46	99
80	32	112	311	27	139	374	23	87	271	25	74	194	22	53	133	22	58	146	22	54	128

eters. We investigate how these numerically best parameters scale with ω, h, and N. We fix the overlap to be four elements. We solve the open cavity problem on the unit square. Three point sources located at $x = 0.23, 0.53, 0.76$ along $y = 0.9$ and zero initial guess are used. We keep $h\omega = \frac{\pi}{5}$, i.e. ten grid points per wavelength. For $\frac{\omega}{2\pi} \in \{10, 20, 40, 80\}$ on $N\times1$ subdomains ($N \in \{2, 4, 8\}$), we do a number of GMRES iterations and record the residuals in the end. Hence, the best parameter for a given wavenumber and a given partition is the one that delivers the smallest residual. The resulting parameters are listed in Table 4.5. Using these parameters, the GMRES iteration numbers for solving the open cavity problem are compared with the other methods in Table 4.6, where BO0 (BO2, resp.) means the best one-sided zeroth-order (second-order, resp.) parameters, OO0 and OO2 are the modified version (see section 4.2), and Zol is aiming at the propagating modes only. We see that the OO2 parameters work as good as the brute force searched BO2 parameters. By Fourier analysis, the spectra (convergence factors, resp.) of BO2 and OO2 in $\mathbb{R}^2$ are depicted in Figure 4.2 (Figure 4.3, resp.). From the figures, we can see that the modification of OO2 helps clustering the majority of the spectra.

4.4. Numerical example: Marmousi open cavity. We use the velocity model of Marmousi from geophysics [33] which spans a 2-D rectangle of side length $9200, 3000$, see Figure 4.4. The Helmholtz equation to be solved in this subsection is

$$(\frac{\omega^2}{c^2} + \Delta)u = -f.$$

The spatially varying wavespeed c has an average 2500 so that at $\omega = 64\pi$ the average wavelength is $\frac{2\pi c}{\omega} = 78.125$ and the horizontal dimension of the domain contains approximately 120 wavelength. For the open cavity problem, we impose Dirichlet conditions on the top and bottom and complex-coordinate perfectly matched layers (PML, c.f. [34]) on the left and right. The thickness of the PML is $D = 8h$. After a variable transformation, in the PML the Helmholtz equation is modified by replacing ∂_x with $\alpha(x)\partial_x$, where $\alpha(x) = 1/(1 - \frac{i8\pi d(x)^2 c_*}{D^3\omega})$, $c_* = 2500$ and d is the distance to the

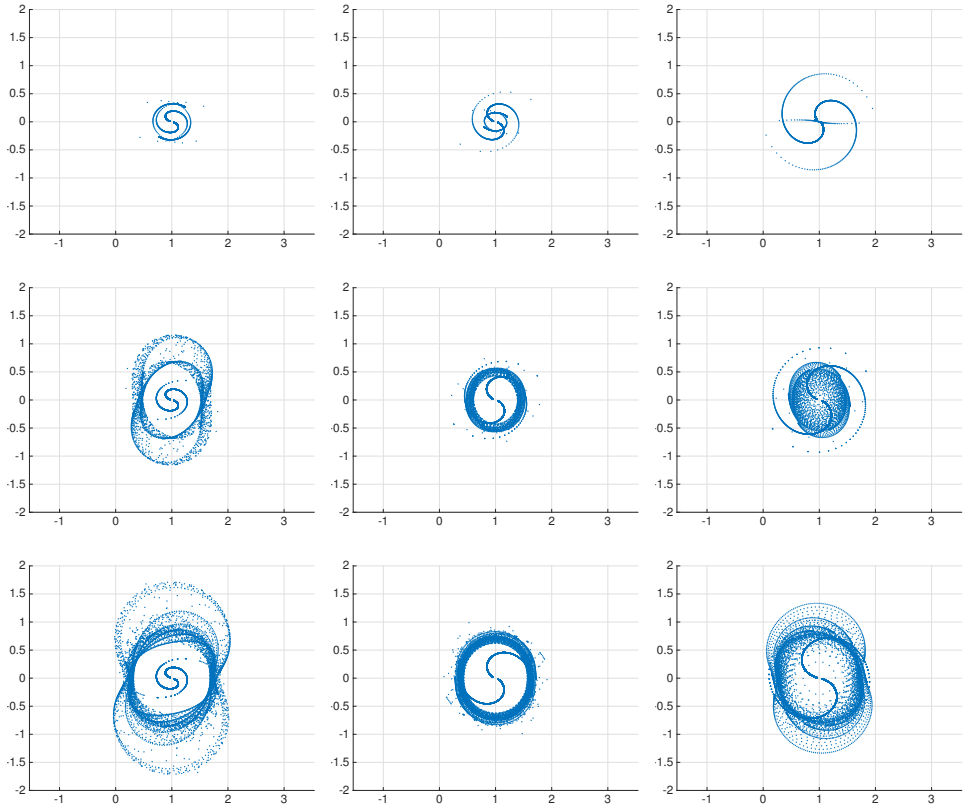

FIG. 4.2. *Complex plane spectral picture of original and modified OO2, BO2 corresponding to the three columns, for $\omega = 160\pi$, and on 2, 4 and 8 subdomains corresponding to the three rows. Real axis is horizontal and imaginary axis is vertical.*

left or right side of the physical domain. We decompose the horizontal dimension into N intervals with overlap twice the mesh size ($L = 2h$). We perform GMRES iterations until the residual of the original system is reduced by a factor of TOL. We set the OO2 parameters with the real part zero and $\delta_\omega = 4\pi/3000$ and let the two-sided Zolotarev's parameters focus on the propagating modes only with $\delta_\omega = 2\pi/3000$. Here, we perform the subdomain solves in double sweeps (left to right and back) and one double-sweep constitutes the Schwarz preconditioner; c.f. [35]. For the spatially varying wavenumber $\frac{\omega}{c}$, we use the frozen coefficient approach, i.e. we generate the OO2 and second-order Zolotarev's parameters locally in space using the local wavenumber. To compare, the PML transmission condition is also tested. Note that PML is a much higher-order transmission condition and gives something like a super optimized Schwarz method, whose implementation is not as simple as OO2. It also increases the subdomain thickness by 16 discrete layers. Interestingly, we found that this super optimized Schwarz method with PML transmission conditions converges much faster when a homogeneous *Dirichlet* rather than a *Neumann* condition is imposed on the left and right sides of the PML-augmented domain and subdomains, while the exact discrete solutions with the two conditions are only slightly different (relative difference of 1.84% in the 2-norm and absolute difference 0.004 in the maximum norm for the

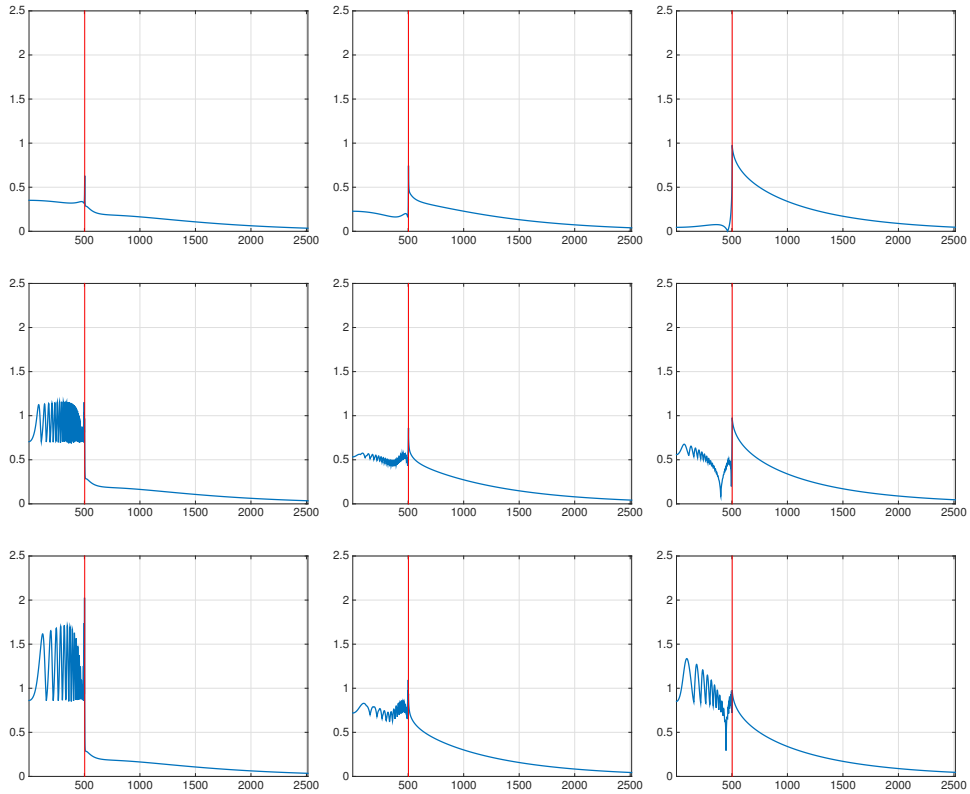

FIG. 4.3. *Convergence factors (vertical axis) against the Fourier parameter* $|\mathbf{k}|$ *(horizontal axis): the three columns correspond to original OO2, modified OO2 and BO2, the three rows correspond to 2, 4 and 8 subdomains, the vertical line indicates* $\omega = 160\pi$.

problem shown in Figure 4.4). The following results are from the better Dirichlet case.

With a point source at the location $(6100, 2200)$, $h \approx 2.93$, $N = 160$ and TOL$= 10^{-6}$, we obtain the solutions shown in Figure 4.4. It takes 88 (98, 26, resp.) GMRES iterations using OO2 (Zolotarev, PML, resp.) transmission conditions and the maximum error to the exact discrete solution is 6.981×10^{-9} (1.037×10^{-8}, 2.262×10^{-8}, resp.). The errors and the residuals are restricted to the physical domain only excluding any PMLs. In principle, without round-off error we should see all the residuals concentrating near the internal interfaces. We observed however that some residuals in the PML for the original domain were generated by GMRES, but not by stationary iterations.

Now we do the ω-scaling tests using zero initial guesses for random exact discrete solutions u_h. The source terms are generated by multiplying the system matrix with u_h. We use a very small tolerance TOL$= 10^{-12}$ to show ample iterations. We fix $\omega h = 187.5\pi$ and set the number of subdomains to $N = 5\frac{\omega}{2\pi}$. In Figure 4.5, we plot the full iterative history of the residuals of the original system and maximum errors to the exact discrete solutions. We can see that all the three methods converge about linearly and the super optimized Schwarz method with PML transmission conditions converges about three times faster than the OO2 and Zolotarev methods.

REMARK 6. *We did also the above experiments for the Marmousi model with*

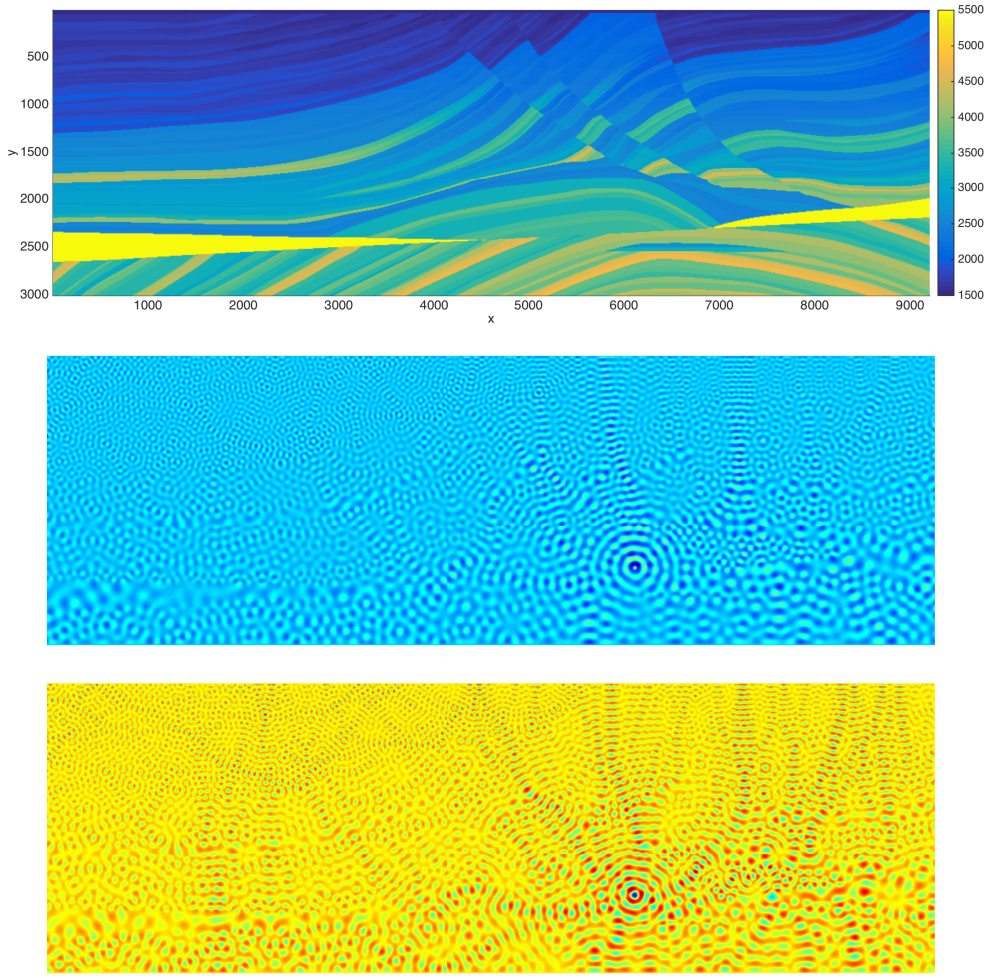

FIG. 4.4. *From top to bottom: Marmousi velocity model, real and imaginary parts of the solution for the Marmousi open cavity problem with a point source at* $(6100, 2200)$ *and* $\omega = 64\pi$.

all sides of the original domain padded with PML. We found in this case the second-order transmission conditions discussed in this paper can not compete with the PML transmission condition. Since a thorough study of the parameter optimization problem in this case is yet to be carried out, we will not report these partial results here.

5. Conclusion. We have studied analytically the best approximation problem for the optimization of zeroth- and second-order transmission (absorbing) conditions with overlap, which is based on a two-subdomain analysis. We found numerically that the optimized transmission conditions for the optimized Schwarz method work perfectly on two subdomains but deteriorate when the number of subdomains increases. We then proposed a simple modification of the parameters to better scale with the number of subdomains, which indeed works very well for the open cavity problem. We found that for the free space problem the Taylor second-order condition

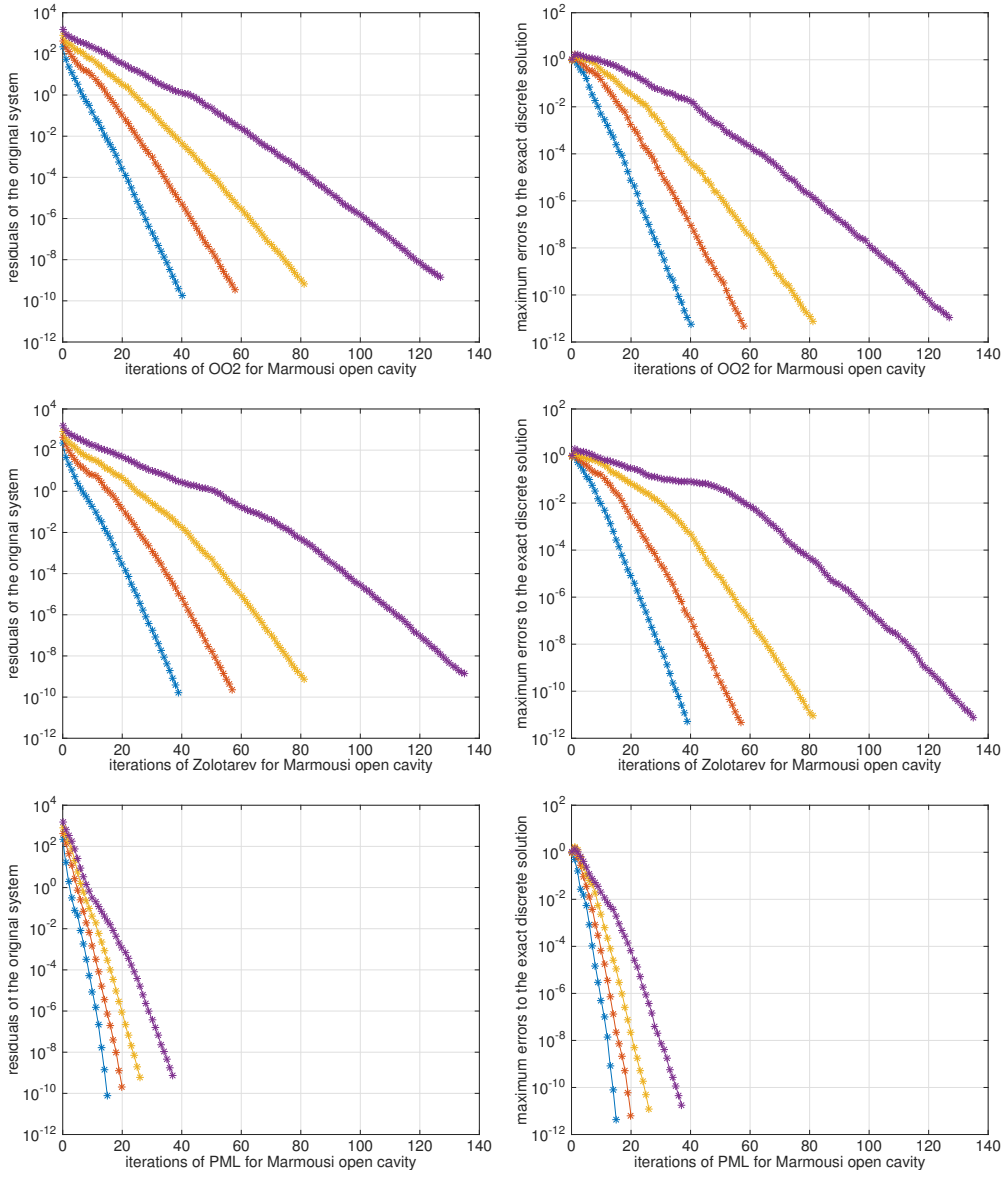

FIG. 4.5. *Iterative history of the residuals (left) and maximum errors (right) for the Marmousi open cavity. In each subplot, from left to right are the curves for* $\frac{\omega}{2\pi} = 4, 8, 16, 32$*, respectively. The mesh size* h *is such that* $\omega h = 187.5\pi$*. Number of subdomains* N *is* $5\frac{\omega}{2\pi}$*. Initial guesses are zero and the exact discrete solutions are random.*

is almost the best among all the one-sided second-order conditions. In our study, we also compared our optimized parameters with Zolotarev's parameters (specialized for a two-sided second-order transmission condition). It turns out that our method is as good as using Zolotarev's parameters and sometimes better. We hope that our techniques presented here can be used for analyzing and optimizing other transmission conditions with overlap, e.g. two-sided second-order transmission conditions with

four complex parameters, and also the super optimized Schwarz method with PML transmission conditions. We should also carry out a detailed analysis of the many-subdomain case to improve our understanding of the numerical observations. Finally, to get a truly parallel scalable solver, we will need to find a scalable coarse problem and/or a scalable subdomain solver. The codes used for this work have been made publicly available at `https://bitbucket.org/MikeHuiZhang/oo2-with-overlap`.

Appendix A. Proof of Theorem 3.3.

We first introduce some notation and do some basic calculations. We denote by

$$g_{low}(x;p_1,p_2) := \frac{p_1^2+(p_1-x)^2}{p_1^2+(p_1+x)^2}\cdot\frac{p_2^2+(p_2-x)^2}{p_2^2+(p_2+x)^2} \tag{A.1}$$

$$= \left(1-\frac{4p_1x}{p_1^2+(p_1+x)^2}\right)\cdot\left(1-\frac{4p_2x}{p_2^2+(p_2+x)^2}\right), \tag{A.2}$$

$$g_{high}(y;p_1,p_2) := g_{low}(y;p_1,p_2)\cdot e^{-4yL}. \tag{A.3}$$

Then, the objective function of (3.1) can be written as

$$|\rho(p_1,p_1,p_2,p_2,\mathbf{k})|^2 = \begin{cases} g_{low}\left(x(|\mathbf{k}|);p_1,p_2\right), & |\mathbf{k}|^2<\omega^2,\\ g_{high}\left(y(|\mathbf{k}|);p_1,p_2\right), & |\mathbf{k}|^2\geq\omega^2,\end{cases}$$

where $x(t)=\sqrt{\omega^2-t^2}\in[x_-,x_{min}]$, $y(t)=\sqrt{t^2-\omega^2}\in[y_+,y_{\max}]$, and $x_-=x(k_-)$, $x_{\min}=x(k_{\min})$, $y_+=y(k_+)$, $y_{\max}=y(k_{\max})$. We thus reformulate (3.1) as

$$\underset{(p_1,p_2)\in\mathbb{P}}{\operatorname{argmin}}\left(\max\left\{\max_{x\in[x_-,x_{\min}]} g_{low}(x;p_1,p_2),\ \max_{y\in[y_+,y_{\max}]} g_{high}(y;p_1,p_2)\right\}\right). \tag{A.4}$$

To compute the maxima of $g_{low}(x)$ and $g_{high}(y)$, we need the derivatives

$$g'_{low}(x) = 4\,\frac{\left(p_2+p_1\right)x^6+\left(-2\,p_1^3-2\,p_2^3\right)x^4+\left(4\,p_1p_2^4+4\,p_1^4p_2\right)x^2-8\,p_1^3p_2^4-8\,p_1^4p_2^3}{\left(x^2+2\,xp_1+2\,p_1^2\right)^2\left(x^2+2\,xp_2+2\,p_2^2\right)^2},$$

$$\begin{aligned} g'_{high}(y) = -4\,\mathrm{e}^{-4\,Ly}\big\{&Ly^8+\left(-p_2-p_1\right)y^6+\left(4\,Lp_2^4+4\,Lp_1^4+2\,p_2^3+2\,p_1^3\right)y^4\\ &+\left(-4\,p_1p_2^4-4\,p_1^4p_2\right)y^2+8\,p_1^3p_2^4+8\,p_1^4p_2^3+16\,Lp_1^4p_2^4\big\}\\ &/\left\{\left(y^2+2\,yp_1+2\,p_1^2\right)^2\left(y^2+2\,yp_2+2\,p_2^2\right)^2\right\}. \end{aligned}$$

We denote by x_{low} a general positive root of $g'_{low}(x)$ and y_{high} a general positive root of $g'_{high}(y)$. Let $z_{low}:=x_{low}^2$ and $z_{high}:=y_{high}^2$. From the above calculations, we know that

$$\left(p_2+p_1\right)z_{low}^3+\left(-2\,p_1^3-2\,p_2^3\right)z_{low}^2+\left(4\,p_1p_2^4+4\,p_1^4p_2\right)z_{low}-8\,p_1^3p_2^4-8\,p_1^4p_2^3=0, \tag{A.5}$$

$$\begin{aligned} &Lz_{high}^4+\left(-p_2-p_1\right)z_{high}^3+\left(4\,Lp_2^4+4\,Lp_1^4+2\,p_2^3+2\,p_1^3\right)z_{high}^2\\ &\qquad+\left(-4\,p_1p_2^4-4\,p_1^4p_2\right)z_{high}+8\,p_1^3p_2^4+8\,p_1^4p_2^3+16\,Lp_1^4p_2^4=0. \end{aligned} \tag{A.6}$$

LEMMA A.1. *Let* $L = C_L h$, $k_{\max} \in [C/h, \infty]$, $C_L, C, k_{\min}, k_-, k_+$ *and* ω *be positive and independent of* h, $k_{\min} < k_- < \omega$, $k_{\max} > k_+ > \omega$ *and*

$$\mathbb{P} = \big\{(p_1, p_2) \,|\, \textit{there exist } C_1, C_2 \in (0,\infty), 0 < a_1 < a_2 < 1, \textit{ such that } \\ p_1 = C_1 h^{-a_1} + o(h^{-a_1}) \textit{ and } p_2 = C_2 h^{-a_2} + o(h^{-a_2}) \textit{ when } h \to 0\big\}.$$

Suppose h *is small, then any solution of (3.1) must be as given in Theorem 3.3.*

Proof. We split the proof into the following steps.

1° Note that $x_-, x_{\min}$ are fixed. From (A.5) and the definition of $\mathbb{P}$, one can easily see that there is no $x_{low} \in [x_-, x_{\min}]$ solving (A.5) when h is sufficiently small.

2° From (A.6), $L = C_L h$ and the definition of $\mathbb{P}$, we have for z_{high}

$$\text{(A.7)}\quad C_L h z_{high}^4 + \left(-C_2 h^{-a_2} + o(h^{-a_2})\right) z_{high}^3 + \left(2\,C_2^3 h^{-3a_2} + o(h^{-3a_2})\right) z_{high}^2 \\ + \left(-4\,C_1 C_2^4 h^{-a_1-4a_2} + o(h^{-a_1-4a_2})\right) z_{high} + 8\,C_1^3 C_2^4 h^{-3a_1-4a_2} + o(h^{-3a_1-4a_2}) = 0.$$

By the transform $z_{high} = C_z h^{-b}$, we find

$$\text{(A.8)}\quad C_L h^{1-4b} C_z^4 + \left(-C_2 h^{-a_2-3b} + o(..)\right) C_z^3 + \left(2\,C_2^3 h^{-3a_2-2b} + o(..)\right) C_z^2 \\ + \left(-4\,C_1 C_2^4 h^{-a_1-4a_2-b} + o(..)\right) C_z + 8\,C_1^3 C_2^4 h^{-3a_1-4a_2} + o(..) = 0,$$

where here and afterwards we use ".." to represent the term immediately before. We then rescale (A.8) with h^{4b-1}, namely

$$C_L C_z^4 + \left(-C_2 h^{-a_2-1+b} + o(..)\right) C_z^3 + \left(2\,C_2^3 h^{-3a_2-1+2b} + o(..)\right) C_z^2 \\ + \left(-4\,C_1 C_2^4 h^{-a_1-4a_2-1+3b} + o(..)\right) C_z + 8\,C_1^3 C_2^4 h^{-3a_1-4a_2-1+4b} + o(..) = 0.$$

Now we take $b = a_2 + 1$ and find

$$\text{(A.9)}\quad F(C_z, h) := C_L C_z^4 + \left(-C_2 + o(1)\right) C_z^3 + \left(2\,C_2^3 h^{-a_2+1} + o(..)\right) C_z^2 \\ + \left(-4\,C_1 C_2^4 h^{-a_1-a_2+2} + o(..)\right) C_z + 8\,C_1^3 C_2^4 h^{-3a_1+3} + o(..) = 0.$$

Letting $h \to 0$ in (A.9) we obtain the limit equation

$$C_L C_z^4 - C_2 C_z^3 = 0,$$

which has the positive root $C_z = C_z^* := C_2/C_L$. We can also verify that the derivative $D_1 F(C_z^*, 0)$ is non-degenerate. Hence the implicit function theorem applies and tells us when $h \to 0$ (A.9) has a root $C_z = C_z^* + o(1) = C_2/C_L + o(1)$. In other words, (A.7) has a root

$$z_{high} = z_{high,1} := (C_2/C_L)\, h^{-a_2-1} + o(..).$$

In the same way, dividing (A.8) by the other powers of h appearing in (A.8), we can find that (A.7) has also the following roots:

$$z_{high} = z_{high,2} := 2\,C_2^2\, h^{-2a_2} + o(..),$$
$$z_{high} = z_{high,3} := 2\,C_1 C_2\, h^{-a_1-a_2} + o(..),$$
$$z_{high} = z_{high,4} := 2\,C_1^2\, h^{-2a_1} + o(..).$$

We are interested in $y_{high} = \sqrt{z_{high}}$. Note that all the roots y_{high} corresponding to the above z_{high} lie in the feasible region $[y_+, y_{\max}]$.

3° Based on 1°, to find $\max_{[x_-, x_{\min}]} g_{low}(x)$ we need only to compare $g_{low}(x_-)$ and $g_{low}(x_{\min})$. By definition of $\mathbb{P}$, (A.1) and the fact that x_-, $x_{\min}$ are fixed independent of h, one can easily find

$$\begin{aligned} g_{low}(x_-) &= 1 - (2\,x_-/C_1)\,h^{a_1} + o(..), \\ g_{low}(x_{\min}) &= 1 - (2\,x_{\min}/C_1)\,h^{a_1} + o(..). \end{aligned} \tag{A.10}$$

Since $x_- < x_{\min}$, we obtain $\max_{[x_-, x_{\min}]} g_{low}(x) = g_{low}(x_-)$.

4° Based on 2°, to find $\max_{[y_+, y_{\max}]} g_{high}(y)$ we need only to compare the values of g_{high} at y_+, $y_{\max} = C/h + o(..)$ and $y_{high,i} := \sqrt{z_{high,i}} (i = 1, \ldots, 4)$. By definition of $\mathbb{P}$, (A.3) and the scalings of these points when $h \to 0$, we find

$$\begin{aligned} g_{high}(y_+) &= 1 - (2\,y_+/C_1)\,h^{a_1} + o(..), \\ g_{high}(y_{\max}) &= 1 - \exp(-4\,C_L C) + o(1), \\ g_{high}(y_{high,1}) &= 1 - 8\,\sqrt{C_L C_2}\,h^{(1-a_2)/2} + o(..), \\ g_{high}(y_{high,2}) &= 1 - \frac{2\sqrt{2}}{2+\sqrt{2}} + o(1), \\ g_{high}(y_{high,3}) &= 1 - 4\,\sqrt{2\,C_1/C_2}\,h^{(a_2-a_1)/2} + o(..), \\ g_{high}(y_{high,4}) &= 1 - \frac{2\sqrt{2}}{2+\sqrt{2}} + o(1). \end{aligned} \tag{A.11}$$

5° In view of (3.1) and (A.4), and using the results from 2° and 3°, we find

$$\begin{aligned} &\max_{|k| \in [k_{\min}, k_-] \cup [k_+, k_{\max}]} |\rho(p_1, p_1, p_2, p_2, k)|^2 \\ &= 1 - \min\left\{(2\,x_-/C_1)\,h^{a_1}, (2\,y_+/C_1)\,h^{a_1}, 8\,\sqrt{C_L C_2}\,h^{(1-a_2)/2}, 4\,\sqrt{2\,C_1/C_2}\,h^{(a_2-a_1)/2}\right\} + o(..). \end{aligned} \tag{A.12}$$

To minimize, we can first optimize the exponents

$$\operatorname*{argmin}_{0<a_1<a_2<1} \max\{a_1, (1-a_2)/2, (a_2-a_1)/2\},$$

and we get the optimal $a_1 = a_1^* = 1/5$ and $a_2 = a_2^* = 3/5$. Then we substitute the optimal a_1, a_2 into (A.12) and optimize the constants $C_1, C_2 > 0$ to obtain (3.2). □

LEMMA A.2. *Let the conditions of Lemma A.1 hold except that we modify the definition of $\mathbb{P}$ to the following cases, respectively, (i) $0 < a_1 < 1 < a_2$, (ii) $0 < a_1 < 1 = a_2$, (iii) $0 < a_1 = a_2 < 1$, (iv) $1 \le a_1 \le a_2$, (v) $a_1 = 0-$, $a_2 \ge 0-$, (vi) $a_1 = 0$, $a_2 \ge 0$, (vii) $a_1 = 0+$, $a_2 \ge 0+$, (viii) $a_1 > 0$, $a_2 = \infty$, (ix) $a_1 = \infty$, $a_2 = \infty$, where*

- *$a_i = 0-$ means $p_i = o(1)$,*
- *$a_i = 0+$ means $p_i^{-1} = o(1)$ and $p_i = o(h^{-a})$ for all $a > 0$,*
- *$a_i = \infty$ means $p_i^{-1} = o(h^a)$ for all $a > 0$,*
- *$a_i \ge 0-$ allows $a_i = 0-, 0, 0+, \infty$ or $a_i > 0$,*
- *$a_i \ge 0$ allows $a_i = 0, 0+, \infty$ or $a_i > 0$,*
- *$a_i \ge 0+$ allows $a_i = 0+, \infty$ or $a_i > 0$.*

Then the solution of (3.1) *can not be better than that of Lemma A.1, i.e. the minimized value is larger than the minimized value of Lemma A.1.*

Proof. We treat the above cases (i)–(ix) as follows.
(i) Note that 1° and 2° of the proof of Lemma A.1 still hold. We find there are two and only two positive roots of (A.6) (the other two roots have non-zero imaginary parts) given by

$$z_{high} = z_{high,1} := C_1/C_L\, h^{-1-a_1} + o(..), \quad z_{high} = z_{high,2} := 2C_1^2\, h^{-2a_1} + o(..).$$

We have all the candidates for the maximum convergence factor: (A.10), (A.11), and

$$g_{high}(y_{high,1}) = 1 - 8\,\sqrt{C_L C_1}\, h^{(1-a_1)/2} + o(..).$$

The optimal choice of $a_1 = 1/3$ gives the minimized value $1 - |\mathcal{O}(h^{1/3})|$ for (3.1), larger than that of Lemma A.1.
(ii), (iii) The equation (A.6) has a positive root $z_{high,1}$ of the same order as in the case (i) and $g_{low}(x_-)$, $g_{high}(y_{high,1})$ are also of the same order as in the case (i). So it can not be better.
(iv) Note that $g_{low}(x_-) = 1 - (2\,x_-/C_1)\, h^{a_1} + o(..)$ and $a_1 \geq 1$.
(v) If $a_2 = 0-, 0$, the order of $1 - g_{high}(h^{-1/2})$ in h is $1/2$. If $a_2 = 0+$, $g_{high}(h^{-1/2}) = 1 - o(h^{1/2-\epsilon})$ for arbitrary $\epsilon > 0$. If $0 < a_2 < 1$, the order of $1 - g_{high}(h^{-a_2/2})$ and $1 - g_{high}(h^{-a_2/2-1/2})$ in h are $a_2/2$ and $(1-a_2)/2$, resp., so the order of $1 - \max g_{high}$ in h is at least $1/4$. If $1 < a_2 \leq \infty$, the order of $g_{high}(h^{-1/2})$ in h is $1/2$.
(vi), (vii) The corresponding claims in (v) stay true.
(viii) If $0 < a_1 < 1$, the order of $1 - g_{low}(x_-)$ and $1 - g_{high}(h^{-(1+a_1)/2})$ in h is a_1 and $(1 - a_1)/2$, resp., so the order of $1 - \max\{\max g_{low}, \max g_{high}\}$ in h is at least $1/4$. The case $a_1 \geq 1$ is the same as (iv).
(ix) Note that $1 - g_{low}(x_-) = o(h^a)$ for all $a > 0$. □

LEMMA A.3. *Let the conditions of Lemma A.1 hold except that we modify the definition of $\mathbb{P}$ in one of the following ways:*

- *(a) we perturb one or both of p_1, p_2 from Lemma A.1 to $p_i = h^{-a_i}\Phi_i(h)$ or $p_i = h^{-a_i}/\Phi_i(h)$, with $(\Phi_i(h))^{-1} = o(1)$, $\Phi_i(h) = o(h^{-b})\ \forall\, b > 0$,*
- *(b) we first take one case from Lemma A.2, if only one of a_1, a_2 is a positive constant, we perturb the corresponding p_i as in (a), else if both of a_1, a_2 are positive constants, we perturb one or both of p_1, p_2 as in (a).*

Then the solution of (3.1) *can not be better than that of Lemma A.1, i.e. the minimized value is larger than the minimized value of Lemma A.1.*

Proof. (a) We consider only the case that p_1 is left unperturbed and $p_2 = h^{-a_2}\Phi_i(h)$. The other cases can be analyzed in a similar way. Note that (A.11) still holds and

$$g_{high}\left(\sqrt{C_L h^{-a_2-1}\Phi_2(h)}\right) = 1 - 8\,\sqrt{C_L h^{1-a_2}\Phi_2(h)} + o(..),$$

$$g_{high}\left(\sqrt{C_1 h^{-a_2-a_1}\Phi_2(h)}\right) = 1 - 6\,\sqrt{C_1 h^{a_2-a_1}/\Phi_2(h)} + o(..),$$

we conclude that the best choice to minimize the maximum of the above three values are $a_1 = 1/5$, $a_2 = 3/5$ but the resulting minimized value is $1 - o(h^{1/5})$.
(b) We can consider the values of the objective function at the same points as in the proof of Lemma A.2. Since the perturbation $\Phi_i(h)$ can not change the exponents of h, the non-optimalities stay true. □

Now we can prove Theorem 3.3.

Proof. Let $p_1^*(h), p_2^*(h)$ be the solution of (3.1). We proceed as follows.
(i) If there exist finite constants $C_1, a_1 > 0$ such that $\overline{\lim}_{h\to 0} p_1^*/h^{-a_1} = C_1$, there will be a subsequence $\lim_{n\to\infty} h_n = 0$, $\lim_{n\to\infty} p_1^*(h_n)/h_n^{-a_1} = C_1$. Now we consider the corresponding subsequence $p_2^*(h_n)$.

(i.i) If there exist finite constants $C_2, a_2 > 0$ such that $\overline{\lim}_{n\to\infty} p_2^*(h_n)/h_n^{-a_2} = C_2$, we can extract a subsequence h_{n_k} such that $\lim_{k\to\infty} p_2^*(h_{n_k})/h_{n_k}^{-a_2} = C_2$. By Lemma A.1 and Lemma A.2, C_1, a_1 must be given as in (3.2).

(i.ii) If $\overline{\lim}_{n\to\infty} p_2^*(h_n)/h_n^{-a_2} = 0$ or ∞ for arbitrary $a_2 > 0$, we will have three possibilities: $\overline{\lim}_{n\to\infty} p_2^*(h_n)/h_n^{-a_2} = 0$ for all $a_2 > 0$, or $\overline{\lim}_{n\to\infty} p_2^*(h_n)/h_n^{-a_2} = \infty$ for all $a_2 > 0$, or there exists $a_2 \in (0,\infty)$ such that $\overline{\lim}_{n\to\infty} p_2^*(h_n)/h_n^{-b} = 0$ for all $b \in (a_2, \infty)$ and $\overline{\lim}_{n\to\infty} p_2^*(h_n)/h_n^{-b} = \infty$ for all $b \in (0, a_2)$. The first two cases can not happen because of Lemma A.2. The third case actually is in Lemma A.3. So the case of (i.ii) can not happen.
(ii) If $\overline{\lim}_{h\to 0} p_1^*/h^{-a_1} = 0$ or ∞ for arbitrary a_1, by extracting subsequences and invoking Lemma A.2 and Lemma A.3, we will get a contradiction as in (i.ii).
(iii) In summary of (i) and (ii), we always have $\overline{\lim}_{h\to 0} p_1^*/h^{-a_1} = C_1$ with C_1, a_1 as in Lemma A.1.
(iv) Similarly to (i)–(iii), we can prove $\underline{\lim}_{h\to 0} p_1^*/h^{-a_1} = C_1$ with C_1, a_1 as in Lemma A.1. Therefore, we have $\lim_{h\to 0} p_1^*/h^{-a_1} = C_1$.
(v) Similarly to (i)–(iv) we have $\lim_{h\to 0} p_2^*/h^{-a_2} = C_2$ with a_2, C_2 as in Lemma A.1.
This completes the proof of Theorem 3.3. □

Appendix B. Proof of Theorem 3.4.

Proof. We first consider the parameter space with the ansatz

$$\mathbb{P} = \{(p_1,p_2)\,|\,\text{there exist } C_1, C_2 \in (0,\infty), 0 < a_1 < a_2 < \gamma, \text{ such that} \\ p_1 = C_1\,\omega^{a_1} + o(\omega^{a_1}) \text{ and } p_2 = C_2\,\omega^{a_2} + o(\omega^{a_2}) \text{ when } \omega \to \infty\}.$$

With this ansatz, we find the three positive roots of $g'_{low}(x)$ as

$$\begin{aligned} x_{low} &= x_{low,1} = \sqrt{2}\,C_2\,\omega^{a_2} + o(..),\\ x_{low} &= x_{low,2} = \sqrt{2\,C_1C_2\,\omega^{a_1+a_2}} + o(..),\\ x_{low} &= x_{low,3} = \sqrt{2}\,C_1\,\omega^{a_1} + o(..), \end{aligned}$$

and the four positive roots of $g'_{high}(y)$ as

$$\begin{aligned} y_{high} &= y_{high,1} = \sqrt{\frac{C_2}{C_LC_h}\,\omega^{\gamma+a_2}} + o(..),\\ y_{high} &= y_{high,2} = \sqrt{2}\,C_2\,\omega^{a_2} + o(..),\\ y_{high} &= y_{high,3} = \sqrt{2\,C_1C_2\,\omega^{a_1+a_2}} + o(..),\\ y_{high} &= y_{high,4} = \sqrt{2}\,C_1\,\omega^{a_1} + o(..). \end{aligned}$$

Furthermore, by (A.1), (A.3) and (A.2) we find

$$\begin{aligned} g_{low}(x_{low,2}) &= 1 - 4\sqrt{\frac{2\,C_1}{C_2}}\,\omega^{(a_1-a_2)/2} + o(..),\\ g_{high}(y_{high,1}) &= 1 - 8\sqrt{C_2C_LC_h}\,\omega^{(a_2-\gamma)/2} + o(..),\\ g_{high}(y_{high,3}) &= 1 - 4\sqrt{\frac{2\,C_1}{C_2}}\,\omega^{(a_1-a_2)/2} + o(..), \end{aligned}$$

and the leading terms of $g_{low}(x_{low,1})$, $g_{low}(x_{low.3})$, $g_{high}(y_{high,2})$ and $g_{high}(y_{high,4})$ are positive constants strictly less than one. We also find the leading orders of $1-g_{low}(x_{\min})$, $1-g_{low}(x_-)$, $1-g_{high}(y_+)$ and $1-g_{high}(y_{\max})$ in ω to be $\max\{-|1-a_1|,-|1-a_2|\}$, $\max\{-|\frac{1}{2}-a_1|,-|\frac{1}{2}-a_2|\}$, $\max\{-|\frac{1}{2}-a_1|,-|\frac{1}{2}-a_2|\}$ and $\exp(-4\,C_L C_h)$. We next optimize a_1, a_2 such that the minimum of the feasible leading orders is maximized, where 'feasible' means the corresponding point lies in $[x_-, x_{\min}]$ or $[y_+, y_{\max}]$. Note that $x_{low,2} \in [x_-, x_{\min}]$ implies $a_1 + a_2 \in [1, 2]$, $y_{high,3} \in [y_+, y_{\max}]$ implies $a_1 + a_2 \geq 1$, and $y_{high,1} \in [y_+, y_{\max}]$ always holds. We split $0 \leq a_1 \leq a_2 \leq \gamma$ into the following seven cases.

(i) $0 \leq a_1 \leq a_2 \leq \frac{1}{2}$. We find the solution, i.e. the maximized minimum of all the feasible leading orders, to be

$$\max_{0\leq a_1\leq a_2\leq \frac{1}{2}} \min\left\{\frac{a_2-\gamma}{2}, a_2-1\right\} = \min\left\{\frac{1}{4}-\frac{\gamma}{2}, -\frac{1}{2}\right\} = \begin{cases} -\frac{1}{2} & \text{if } 1\leq\gamma\leq\frac{3}{2}, \\ \frac{1}{4}-\frac{\gamma}{2} & \text{otherwise.} \end{cases}$$

(ii) $0 \leq a_1 \leq \frac{1}{2} \leq a_2, a_1 + a_2 \leq 1$. We have the solution

$$\max_{\substack{0\leq a_1\leq\frac{1}{2}\leq a_2,\\ a_1+a_2\leq 1}} \min\left\{a_2-1, \frac{1}{2}-a_2, \frac{a_2-\gamma}{2}\right\} = \begin{cases} -\frac{1}{4} & \text{if } 1\leq\gamma\leq\frac{5}{4}, \\ \frac{1}{6}-\frac{\gamma}{3} & \text{if } \frac{5}{4}\leq\gamma\leq 2, \\ \frac{1-\gamma}{2} & \text{otherwise.} \end{cases}$$

(iii) $0 \leq a_1 \leq \frac{1}{2} \leq a_2 \leq 1, a_1 + a_2 \geq 1$. We have the solution (note that $(a_1 - a_2)/2 = a_1 - 1/2$ when $a_1 + a_2 = 1$ so $(a_1 - a_2)/2$ is always feasible)

$$\max_{\substack{0<a_1\leq\frac{1}{2}\leq a_2\leq 1,\\ a_1+a_2\geq 1}} \min\left\{a_2-1, a_1-\frac{1}{2}, \frac{a_2-\gamma}{2}, \frac{a_1-a_2}{2}\right\} = \begin{cases} -\frac{1}{4} & \text{if } 1\leq\gamma\leq\frac{5}{4}, \\ \frac{1}{6}-\frac{\gamma}{3} & \text{if } \frac{5}{4}\leq\gamma\leq 2, \\ \frac{1-\gamma}{2} & \text{otherwise.} \end{cases}$$

(iv) $0 \leq a_1 \leq \frac{1}{2}$, $1 \leq a_2 \leq \gamma$,. We have the solution

$$\max_{\substack{0\leq a_1\leq\frac{1}{2},\\ 1\leq a_2\leq\gamma}} \min\left\{-\min_i |a_i-1|, a_1-\frac{1}{2}, \frac{a_2-\gamma}{2}, \frac{a_1-a_2}{2}\right\} = \begin{cases} -\frac{1}{4} & \text{if } 1\leq\gamma\leq\frac{3}{2}, \\ \frac{1}{8}-\frac{\gamma}{4} & \text{otherwise.} \end{cases}$$

(v) $\frac{1}{2} \leq a_1 \leq a_2 \leq 1$. We have the solution

$$\max_{\frac{1}{2}\leq a_1\leq a_2\leq 1} \min\left\{a_2-1, \frac{1}{2}-a_1, \frac{a_2-\gamma}{2}, \frac{a_1-a_2}{2}\right\} = \begin{cases} -\frac{1}{8} & \text{if } 1\leq\gamma\leq\frac{9}{8}, \\ \frac{1}{10}-\frac{\gamma}{5} & \text{if } \frac{9}{8}\leq\gamma\leq\frac{4}{3}, \\ \frac{1-\gamma}{2} & \text{otherwise.} \end{cases}$$

(vi) $\frac{1}{2} \leq a_1 \leq 1 \leq a_2 \leq \gamma$. We have the solution

$$\max_{\frac{1}{2}\leq a_1\leq 1\leq a_2\leq\gamma} \min\left\{-\min_i |a_i-1|, \frac{1}{2}-a_1, \frac{a_2-\gamma}{2}, \frac{a_1-a_2}{2}\right\} = \begin{cases} -\frac{1}{6} & \text{if } 1\leq\gamma\leq\frac{4}{3}, \\ \frac{1}{10}-\frac{\gamma}{5} & \text{if } \frac{4}{3}\leq\gamma\leq 3, \\ \frac{1}{8}-\frac{\gamma}{4} & \text{otherwise.} \end{cases}$$

(vii) $1 \leq a_1 \leq a_2 \leq \gamma$. We have the solution

$$\max_{1\leq a_1\leq a_2\leq\gamma} \min\left\{1-a_1, \frac{1}{2}-a_1, \frac{a_2-\gamma}{2}, \frac{a_1-a_2}{2}\right\} = \begin{cases} -\frac{1}{2} & \text{if } 1\leq\gamma<3, \\ \frac{1}{10}-\frac{\gamma}{5} & \text{otherwise.} \end{cases}$$

In summary from the seven cases above, the optimal solutions of a_1, a_2 are

$$a_1 = \begin{cases} \frac{5}{8} & \text{if } 1 \le \gamma \le \frac{9}{8}, \\ \frac{2}{5} + \frac{\gamma}{5} & \text{otherwise,} \end{cases} \qquad a_2 = \begin{cases} \frac{7}{8} & \text{if } 1 \le \gamma \le \frac{9}{8}, \\ \frac{1}{5} + \frac{3\gamma}{5} & \text{otherwise,} \end{cases}$$

and they maximize the minimum of all the feasible leading orders as

$$\begin{cases} -\frac{1}{8} & \text{if } 1 \le \gamma \le \frac{9}{8}, \\ \frac{1}{10} - \frac{\gamma}{5} & \text{otherwise.} \end{cases}$$

The rest of the proof is essentially following the same lines as the proof of Theorem 3.3, so we omit the details for brevity. □

Acknowledgments. We thank the editor and the referees who helped us to greatly improve the manuscript.

REFERENCES

[1] I. M. Babuska and S. A. Sauter, *Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers?*, SIAM J. Numer. Anal., 34(6) (1997), pp.2392–2423.

[2] O. Ernst and M. J. Gander, *Why it is difficult to solve Helmholtz problems with classical iterative methods*, in I. Graham, T. Hou, O. Lakkis and R. Scheichl (Eds.), Numerical Analysis of Multiscale Problems, Springer-Verlag, Berlin, 2012, pp. 325–363.

[3] M. J. Gander, *Optimized Schwarz methods*, SIAM J. Numer. Anal., 44 (2006), pp. 699–731.

[4] A. St-Cyr, M. J. Gander and S. J. Thomas, *Optimized multiplicative, additive, and restrict additive Schwarz preconditioning*, SIAM J. Sci. Comput., 29 (2007), pp. 2402–2425.

[5] S. Loisel and D. B. Szyld, *On the geometric convergence of optimized Schwarz methods with applications to elliptic problems*, Numer. Math., 114 (2010), pp. 697-728.

[6] M. J. Gander and F. Kwok, *Best Robin parameters for optimized Schwarz methods at cross points*, SIAM J. Sci. Comput., 34 (2012), pp. A1849–A1879.

[7] S. Loisel, *Condition number estimates for the nonoverlapping optimized Schwarz method and the 2-Lagrange multiplier method for general domains and cross points*, SIAM J. Numer. Anal., 51 (2013), pp. 3062–3083.

[8] M. J. Gander and Y. Xu, *Optimized Schwarz methods for circular domain decompositions with overlap*, SIAM J. Numer. Anal., 52 (2014), pp. 1981–2004.

[9] M. J. Gander and S. Hajian, *Analysis of Schwarz methods for a hybridizable discontinuous Galerkin discretization*, SIAM J. Numer. Anal., 53 (2015), pp. 573–579.

[10] B. Després, *Domain decomposition method and the Helmholtz problem*, in Mathematical and numerical aspects of wave propagation phenomena, G. Cohen, L. Halpern and P. Joly, eds., SIAM, Philadelphia, 1991, pp. 44–52.

[11] X.-C. Cai, M. A. Casarin, F. W. Elliott and O. B. Widlund, *Overlapping Schwarz algorithms for solving Helmholtz's equation*, in Proceedings of the 10th Intl. Conf. on Domain Decomposition Methods, J. Mandel, C. Farhat and X.-C. Cai, eds., AMS, 1998, pp. 437–445.

[12] J.-H. Kimn and M. Sarkis, *Restricted overlapping balancing domain decomposition methods and restricted coarse problems for the Helmholtz problem*, Comput. Methods Appl. Mech. Engrg., 196 (2007), pp. 1507–1514.

[13] M. J. Gander, F. Magoulès and F. Nataf, *Optimized Schwarz methods without overlap for the Helmholtz equation*, SIAM J. Sci. Comput., 24 (2002), pp. 38–60.

[14] M. J. Gander, L. Halpern and F. Magoulès, *An optimized Schwarz method with two-sided Robin transmission conditions for the Helmholtz equation*, Int. J. Numer. Meth. Fluids, 55 (2007), pp. 163–175.

[15] V. Dolean, M. J. Gander and L. Gerardo-Giorda, *Optimized Schwarz methods for Maxwell's equations*, SIAM J. Sci. Comput., 31 (2009), pp. 2193–2213.

[16] Z. Peng and J.-F. Lee, *A scalable nonoverlapping and nonconformal domain decomposition method for solving time-harmonic Maxwell equations in* $\mathbb{R}^3$, SIAM J. Sci. Comput., 34 (2012), pp. A1266–A1295.

[17] M. El Bouajaji, V. Dolean, M. J. Gander and S. Lanteri, *Comparison of a one and two parameter family of transmission conditions for Maxwell's equations with damping*, in Domain

Decomposition Methods in Science and Engineering XX, R. Bank, M. Holst, O. Widulund and J. Xu, eds., Springer, 2013, pp. 271–278.

[18] A. Toselli, *Overlapping methods with perfectly matched layers for the solution of the Helmholtz equation*, in Eleventh International Conference on Domain Decomposition Methods, C.H. Lai, P.E. Bjorstad, M. Cross and O.B. Widlund, eds., DDM.org, 1999, pp. 551–558.

[19] A. Schadle, L. Zschiedrich, *Additive Schwarz method for scattering problems using the PML method at interfaces*, in Domain Decomposition Methods in Science and Engineering XVI, O.B. Widlund, and D.E. Keyes, eds., Springer, 2007, pp. 205–212.

[20] Y. Boubendir, X. Antoine and C. Geuzaine, *A quasi-optimal non-overlapping domain decomposition algorithm for the Helmholtz equation*, J. Comput. Phys., 231 (2012), pp. 262–280.

[21] S. Kim and H. Zhang, *Optimized Schwarz method with complete radiation transmission conditions for the Helmholtz equation*, SIAM J. Numer. Anal., 53 (2015), pp. 1537–1558.

[22] B. Engquist and L. Ying, *Sweeping preconditioner for the Helmholtz equation: Moving perfectly matched layers*, Multiscale Model. Simul., 9 (2011), pp. 686–710.

[23] Z. Chen and X. Xiang, *A source transfer domain decomposition method for Helmholtz equations in unbounded domain*, SIAM J. Numer. Anal., 51 (2013), pp. 2331–2356.

[24] C. C. Stolk, *A rapidly converging domain decomposition method for the Helmholtz equation*, J. Comput. Phys., 241 (2013), pp. 240–252.

[25] A. Vion and C. Geuzaine, *Double sweep preconditioner for optimized Schwarz methods applied to the Helmholtz problem*, J. Comput. Phys. 266 (2014), pp. 171–190.

[26] L. Zepeda-Nunez and L. Demanet, *The method of polarized traces for the 2D Helmholtz equation*, arXiv:1410.5910v2, 2014.

[27] B. Engquist and A. Majda, *Absorbing boundary conditions for the numerical simulation of waves*, Math. Comp. 31 (1977), pp. 629–651.

[28] V. Druskin, S. Guttel and L. Knizhnerman, *Near-optimal perfectly matched layers for indefinite Helmholtz problems*, SIAM Review 58 (2016), pp. 90–116.

[29] L. Zhu and H. Wu, *Preasymptotic error analysis of CIP-FEM and FEM for Helmholtz equation with high wave number. Part II: hp version*, SIAM J. Numer. Anal., 51 (2013), pp. 1828–1852.

[30] M. J. Gander and H. Zhang, *Domain decomposition methods for the Helmholtz equation: a numerical investigation*, in Domain Decomposition Methods in Science and Engineering XX, R. Bank, M. Holst, O. Widlund and J. Xu, eds., Springer, 2013, pp. 215–222.

[31] K. Ito and J. Toivanen, *Preconditioned iterative methods on sparse subspaces*, Appl. Math. Lett., 19 (2006), pp. 1191–1197.

[32] A. Bamberger, P. Joly and J. E. Roberts, *Second-order absorbing boundary conditions for the wave equation: a solution for the corner problem*, SIAM J. Numer. Anal., 27 (1990), pp. 323–352.

[33] R. Versteeg and G. Grau (editors), *The Marmousi Experience, Proceedings of 1990 EAEG Workshop on Practical Aspects of Seismic Data Inversion*, Eur. Assoc. Expl. Geophys., Zeist, 1991.

[34] W. C. Chew, J. M. Jin and E. Michielssen, *Complex coordinate stretching as a generalized absorbing boundary condition*, Microwave and Optical Technology Letters, 15 (1997), pp. 363–369.

[35] F. Nataf, *On the use of open boundary conditions in block Gauss-Seidel methods for convection-diffusion equations*, CMAP (Ecole Polytechnique), 1993.