Convergence of overlapping domain decomposition methods with PML transmission conditions applied to nontrapping Helmholtz problems

J. Galkowski[*], S. Gong[†], I. G. Graham[‡], D. Lafontaine[§], E. A. Spence[¶]

April 2, 2024

Abstract

We study overlapping Schwarz methods for the Helmholtz equation posed in any dimension with large, real wavenumber and smooth variable wave speed. The radiation condition is approximated by a Cartesian perfectly-matched layer (PML). The domain-decomposition subdomains are overlapping hyperrectangles with Cartesian PMLs at their boundaries. The overlaps of the subdomains and the widths of the PMLs are all taken to be independent of the wavenumber.

For both parallel (i.e., additive) and sequential (i.e., multiplicative) methods, we show that after a specified number of iterations – depending on the behaviour of the geometric-optic rays – the error is smooth and smaller than any negative power of the wavenumber. For the parallel method, the specified number of iterations is less than the maximum number of subdomains, counted with their multiplicity, that a geometric-optic ray can intersect.

These results, which are illustrated by numerical experiments, are the first wavenumber-explicit results about convergence of overlapping Schwarz methods for the Helmholtz equation, and the first wavenumber-explicit results about convergence of *any* domain-decomposition method for the Helmholtz equation with a non-trivial scatterer (here a variable wave speed).

1 Introduction

1.1 Informal description of the problem

We consider the following Helmholtz problem: for arbitrary $d \geq 1$, given $f \in L^2_{\mathrm{comp}}(\mathbb{R}^d)$, (normalised) strictly-positive wavespeed $c \in C^\infty(\mathbb{R}^d)$ with $\mathrm{supp}(1-c)$ compact, and wavenumber $k \gg 1$, find $u \in H^1_{\mathrm{loc}}(\mathbb{R}^d)$ satisfying the Helmholtz equation and Sommerfeld radiation condition:

$$k^{-2}\Delta u + c^{-2}u = f \quad \text{in } \mathbb{R}^d \quad \text{and} \quad (\partial_r - ik)u = o(r^{\frac{1-d}{2}}) \text{ as } r := |x| \to \infty. \tag{1.1}$$

Since $\mathbb{R}^d$ is unbounded, a standard approximation to this Helmholtz problem is to truncate the domain at an artificial boundary chosen so that the truncated domain contains both $\mathrm{supp}(1-c)$ and $\mathrm{supp}f$ (i.e., the scatterer and the data) and approximate the radiation condition by a *perfectly-matched layer* (PML) [3]. In this paper we assume that the artificial boundary is a

---

[*]Department of Mathematics, University College London, 25 Gordon Street, London, WC1H 0AY, UK, `J.Galkowski@ucl.ac.uk`

[†]School of Science and Engineering, The Chinese University of Hong Kong, Shenzhen, Guangdong 518172, China, `gongshihua@cuhk.edu.cn`

[‡]Department of Mathematical Sciences, University of Bath, UK, `i.g.graham@bath.ac.uk`

[§]CNRS and Institut de Mathématiques de Toulouse, UMR5219; Université de Toulouse, CNRS; UPS, F-31062 Toulouse Cedex 9, France; `david.lafontaine@math.univ-toulouse.fr`

[¶]Department of Mathematical Sciences, University of Bath, UK, `e.a.spence@bath.ac.uk`

hyperrectangle, and that the PML is a Cartesian PML (i.e., the scaling is active in Cartesian coordinate directions; see §1.4.2 below for a precise definition).

Helmholtz solutions oscillate on a length scale of k^{-1}, and approximating an arbitrary function oscillating on this scale requires $\sim (kL)^d$ degrees of freedom, where L is the length scale of the domain. For piecewise polynomials of fixed degree, the number of degrees of freedom required is $\gg (kL)^d$ because of the *pollution effect* [1, 21]. The linear systems resulting from finite-element discretisations of (1.1) are therefore very large. Furthermore, since the standard variational formulation of the Helmholtz problem above is not self-adjoint and not coercive, methods that work well for self-adjoint coercive problems, such as Poisson's equation, usually perform very badly for Helmholtz problems (see, e.g., the review [18]).

Domain-decomposition (DD) methods approximate the solution of the Helmholtz problem on the computational domain by solving Helmholtz problems on subdomains (either overlapping or non-overlapping) of the original domain, with each subdomain problem involving fewer degrees of freedom than the original problem. Each iteration of a parallel (a.k.a. additive) method involves solving decoupled problems on each subdomain (i.e., subdomains do not communicate with each other at this stage). In contrast, each iteration of a sequential (a.k.a. multiplicative) method involves communication between subdomains at each solve phase.

1.2 Context: DD methods using PML

The design and analysis of DD methods for solving the Helmholtz equation is a very active area; see, e.g., the reviews [17, 18, 33, 25, 23].

A key question is: what boundary conditions should be imposed on the DD subdomains? It is known that the optimal boundary condition on a particular DD subdomain is the Dirichlet-to-Neumann map for the Helmholtz equation posed in the exterior of that subdomain (as a subset of the whole domain) [42, §2] [16, §2], [12, §2.4]. However, complete knowledge of these maps is equivalent to knowing the solution operator for the original problem. The design of good, practical subdomain boundary conditions is then the goal of *optimised Schwarz methods*; see, e.g., [22].

Using PML as a subdomain boundary condition for Helmholtz DD was first advocated for in [54], and there are now many DD methods for Helmholtz using PML on subdomain boundaries with impressive empirical performance; see, e.g., [15, 52, 10, 58, 53, 35, 47, 5] and the reviews [25, 23].

It is now understood in a k-explicit way how PML approximates the Dirichlet-to-Neumann map associated with the Sommerfeld radiation condition. Indeed, when $c \equiv 1$ and $d = 2$, the error between the true Helmholtz solution and the Cartesian PML approximation decreases exponentially in k, the width of the PML, and the strength of the scaling by [10, Lemma 3.4] (this result also holds for general scatterers in $d \geq 2$ with a radial PML by [20]). For general smooth c and any $d \geq 2$, this error is smaller than any negative power of k by Theorem 4.6 below.

However, there is no rigorous understanding of how well PML approximates the (more complicated) Dirichlet-to-Neumann maps corresponding to the optimal DD boundary conditions for general decompositions and non-trivial scatterers; i.e., there are no rigorous k-explicit results about the convergence of DD methods using PML applied to Helmholtz problems with $k \gg 1$ and non-trivial scatterers.

The only existing k-explicit rigorous convergence results are for the sequential "source transfer" DD methods (which [25] showed can be considered as a particular type of optimised Schwartz method) applied to (1.1) with $c \equiv 1$ (i.e., no scattering). These methods involve subdomains that only overlap via the perfectly-matched layers (i.e., the physically-relevant parts of the subdomains do not overlap). For these methods applied to (1.1) with $c \equiv 1$ (i.e., no scattering), k-explicit convergence of the DD method at the continuous level can be obtained from a k-explicit result about how well (Cartesian) PML approximates the Sommerfeld radiation condition – this

is precisely because the optimal boundary conditions on the subdomains in this case are the Dirichlet-to-Neumann maps associated with the Sommerfeld radiation condition. Therefore, accuracy of Cartesian PML when $c \equiv 1$ [10, Lemma 3.4] is then the heart of the k-explicit convergence proofs of the source-transfer-type methods in [10, 35, 13, 36, 37] for $c \equiv 1$.

1.3 Informal description of the main results

The main results of the present paper, Theorems 1.1-1.4 and 1.6 below, concern both parallel and sequential overlapping Schwarz methods at the PDE level (i.e., before discretisation).

Theorem (Informal summary of Theorems 1.1-1.4 and 1.6). *For both parallel and sequential overlapping Schwarz methods applied to the Cartesian PML approximation of* (1.1), *where the subdomains are hyperrectangles with Cartesian PMLs at their boundaries, the following is true. After a number of iterations depending on the behaviour of the geometric-optic rays, given any $M > 0$ there exists $C > 0$ such that the error is smooth and smaller than Ck^{-M} for all $k > 0$.*

In particular, this implies that the fixed-point iterations converge exponentially quickly in the number of iterations for sufficiently-large k.

These results are the first k-explicit results about convergence of overlapping Schwarz methods for the Helmholtz equation, and the first k-explicit results about convergence of *any* DD method for the Helmholtz equation with a non-trivial scatterer (here a variable wave speed). We highlight the following.

- These results are valid on fixed domains for sufficiently-large k, i.e., the PML widths and DD overlaps are arbitrary, but independent of k. Obtaining results that are also explicit in these geometric parameters of the decomposition will require more technical arguments than those used here.

- We make clear exactly what properties of PML are required to obtain these results, and for what other complex absorption operators these results also hold (see Appendix A).

- These results are the PML analogues of the results in [26, 31] for parallel overlapping Schwarz methods with impedance boundary conditions, although the results of the present paper are much stronger because of the superiority of PML over impedance boundary conditions in this context.

1.4 Definition of the overlapping Schwarz methods considered in this paper

1.4.1 Definition of the subdomains

Let the d-dimensional hyperrectangular domain Ω_{int} be given by the Cartesian product

$$\Omega_{\mathrm{int}} := \prod_{1 \leq \ell \leq d} (0, L_\ell),$$

Let $\{\Omega_{\mathrm{int},j}\}_{j=1}^N$ be an overlapping decomposition of Ω_{int} in hyperrectangular subdomains:

$$\Omega_{\mathrm{int},j} := \prod_{1 \leq \ell \leq d} (a_\ell^j, b_\ell^j) \tag{1.2}$$

We extend Ω_{int} and each $\Omega_{\mathrm{int},j}$, $1 \leq j \leq N$, by adding a PML layer to each, to form the domains Ω and Ω_j. Namely, let $\kappa > 0$ and $\kappa_0 > 0$ denote respectively the PML width on Ω and the interior PML width and let

$$\Omega := \prod_{1 \leq \ell \leq d} (-\kappa, L_\ell + \kappa), \quad \Omega_j := \prod_{1 \leq \ell \leq d} (a_\ell^j - \kappa_\ell^j, b_\ell^j + \kappa_\ell^j),$$

3

where $\kappa_\ell^j = \kappa_0$ if the corresponding edge of $\Omega_{\mathrm{int},j}$ belongs to the interior of Ω, and κ otherwise; i.e., edges of subdomains that touch the boundary have the same PML width as Ω (i.e., κ), and interior subdomain edges have PML width κ_0 which can be different than κ.

1.4.2 The scaled operators

We now define standard Cartesian PMLs of width κ at the boundary of Ω and width κ_0 at the boundary of each Ω_j. For simplicity, we assume that the same PML scaling function is used inside every PML; however, this assumption can be easily removed, with, say, one function used in the PMLs of width κ and another used in the PMLs of width κ_0, at the cost of introducing more notation.

Let $f_{\mathrm{s}} \in C^\infty(\mathbb{R})$ (with subscript s standing for "scaling") be such that

$$\{x : f_{\mathrm{s}}(x) = 0\} = \{x : f_{\mathrm{s}}'(x) = 0\} = \{x : x \leq 0\}, \ f_{\mathrm{s}}'(x) > 0 \text{ for } x > 0, \text{ and } f_{\mathrm{s}}''(x) = 0 \text{ for } x \geq \kappa_{\mathrm{lin}}$$

for some $\kappa_{\mathrm{lin}} < \kappa$ (observe that f_{s} is then linear for $x \geq \kappa_{\mathrm{lin}}$). (This assumption is to avoid technical issues about propagation of singularities – see Remark 2.5 below.)

For any $1 \leq \ell \leq d$, we define the following scaling functions in the ℓ-direction $g_\ell \in C^\infty(\mathbb{R}^d)$ by

$$g_\ell(x_\ell) := \begin{cases} f_{\mathrm{s}}(x_\ell - L_\ell) & \text{if } x_\ell \geq L_\ell, \\ 0 & \text{if } x_\ell \in (0, L_\ell), \\ -f_{\mathrm{s}}(-x_\ell) & \text{if } x_\ell \leq 0, \end{cases}$$

and similarly, the subdomain scaling functions in the ℓ-direction $g_{\ell,j} \in C^\infty(\mathbb{R}^d)$ for $1 \leq j \leq N$ by

$$g_{\ell,j}(x_\ell) := \begin{cases} f_{\mathrm{s}}(x_\ell - b_\ell^j) & \text{if } x_\ell \geq b_\ell^j, \\ 0 & \text{if } x_\ell \in (a_\ell^j, b_\ell^j), \\ -f_{\mathrm{s}}(a_\ell^j - x_\ell) & \text{if } x_\ell \leq a_\ell^j. \end{cases}$$

We now define the scaled operators Δ_{s} and $\Delta_{\mathrm{s},j}$ by

$$\Delta_{\mathrm{s}} := \sum_{\ell=1}^d \left(\frac{1}{1 + ig_\ell'(x_\ell)} \partial_{x_\ell} \right)^2,$$

and

$$\Delta_{\mathrm{s},j} := \sum_{\ell=1}^d \left(\frac{1}{1 + ig_{\ell,j}'(x_\ell)} \partial_{x_\ell} \right)^2 \quad \text{for all } 1 \leq j \leq N. \tag{1.3}$$

Given $c \in C^\infty(\mathbb{R}^d)$ that is strictly positive and such that $\mathrm{supp}(1 - c) \subset \Omega_{\mathrm{int}}$, let

$$P_{\mathrm{s}} := -k^{-2}\Delta_{\mathrm{s}} - c^{-2}, \qquad P_{\mathrm{s}}^j := -k^{-2}\Delta_{\mathrm{s},j} - c^{-2} \ \text{ for all } 1 \leq j \leq N. \tag{1.4}$$

These operators are defined on $H^1(\mathbb{R}^d)$, but we consider P_{s} as a operator $H^1(\Omega) \to H^{-1}(\Omega)(= (H_0^1(\Omega))^*)$, and P_{s}^j as an operator $H^1(\Omega_j) \to H^{-1}(\Omega_j)(= (H_0^1(\Omega_j))^*)$. If Dirichlet boundary data in $H^{1/2}$ is prescribed on the corresponding boundaries, these operators are then invertible.

In the proofs of our main results, a key region is the following subset of Ω_j:

$$\mathrm{supp}(P_{\mathrm{s}}^j - P_{\mathrm{s}}) := \Omega_j \cap \left(\bigcup_{\ell=1}^d \overline{\{x \in \mathbb{R}^d : g_{\ell,j}(x_\ell) \neq g_\ell(x_\ell)\}} \right);$$

$\mathrm{supp}(P_{\mathrm{s}}^j - P_{\mathrm{s}})$ is the part of the PML region of Ω_j where the PML in Ω_j differs from the PML in Ω (i.e., where $P_{\mathrm{s}}^j \neq P_{\mathrm{s}}$). Note that when Ω_j is an interior subdomain, $\mathrm{supp}(P_{\mathrm{s}}^j - P_{\mathrm{s}}) = \Omega_j \setminus \Omega_{\mathrm{int,j}}$.

1.4.3 The Helmholtz problem

Given $f \in H^{-1}(\Omega)$, find $u \in H_0^1(\Omega)$ such that

$$P_{\mathrm{s}} u = f. \tag{1.5}$$

The solution u exists and is unique *either* for fixed k when the PML width κ is sufficiently large [30, Theorem 5.5], [6, Theorem 5.7] *or* for fixed κ and scaling function f_{s} when k is sufficiently large (by Lemma 4.5 below). When $\mathrm{supp}\, f \subset \Omega_{\mathrm{int}}$ and $c \equiv 1$, [10, Lemma 3.4] shows that the error between u and the solution of the true Helmholtz problem in $\mathbb{R}^d$ satisfying the Sommerfeld radiation condition decays exponentially in k and κ. Theorem 4.6 below shows that, for $f \in H^{-1}(\Omega_{\mathrm{int}})$ and $c \not\equiv 1$, for fixed κ, this error is $O(k^{-\infty})$; i.e., smaller than any negative power of k.

1.4.4 The partition of unity

Let $\{\chi_j\}_{j=1}^N$ be a partition of unity subordinate to $\{\Omega_j \setminus \mathrm{supp}(P_{\mathrm{s}}^j - P_{\mathrm{s}})\}_{1 \le j \le N}$, i.e., $\{\chi_j\}_{1 \le j \le N}$ is a family of non-negative elements of $C^\infty(\mathbb{R}^d)$ such that

$$\text{for all } x \in \mathbb{R}^d, \quad \sum_{j=1}^N \chi_j(x) = 1, \qquad \mathrm{supp}\, \chi_j \cap \Omega \subset \left(\Omega_j \setminus \mathrm{supp}(P_{\mathrm{s}}^j - P_{\mathrm{s}}) \right) \text{ for all } j, \tag{1.6}$$

and, additionally, in a neighbourhood of $\partial\Omega$, each χ_j does not vary in the normal direction to the boundary of Ω (this last assumption is for technical reasons, and can easily be achieved in practice).

1.4.5 The parallel overlapping Schwarz method

Given $u_+^n \in H_0^1(\Omega)$, for $n \ge 0$ and any $1 \le j \le N$, let $u_j^{n+1} \in H^1(\Omega_j)$ be the solution to

$$\begin{cases} P_{\mathrm{s}}^j u_j^{n+1} = P_{\mathrm{s}}^j(u_+^n|_{\Omega_j}) - (P_{\mathrm{s}} u_+^n)|_{\Omega_j} + f|_{\Omega_j} \in H^{-1}(\Omega_j) \\ u_j^{n+1} = u_+^n \in H^{1/2}(\partial\Omega_j), \end{cases} \tag{1.7}$$

and then set

$$u_+^{n+1} := \sum_{j=1}^N \chi_j u_j^{n+1} \tag{1.8}$$

(where the subscript $+$ indicates that these are the iterates for the parallel/additive method).

Another way to write this is to introduce the *local corrector* $\mathfrak{c}_j^n := u_j^{n+1} - u_+^n|_{\Omega_j} \in H_0^1(\Omega_j)$, which satisfies

$$P_{\mathrm{s}}^j \mathfrak{c}_j^n = (f - P_{\mathrm{s}} u_+^n)|_{\Omega_j} \in H^{-1}(\Omega_j) \quad \text{and} \quad u_+^{n+1} = u_+^n + \sum_{j=1}^N \chi_j \mathfrak{c}_j^n. \tag{1.9}$$

From this we see that the solution u of (1.5) is a fixed point of (1.7)-(1.8). Indeed, if $u_+^n = u$ then $\mathfrak{c}_j^n = 0$ by uniqueness of the PML problem on Ω_j, and thus $u_+^{n+1} = u_+^n = u$. (Note that the subdomain problems for $\mathfrak{c}_j^n$ are wellposed by the same results that ensure wellposedness of (1.5).)

In §8.1 below, we show that, on the discrete level, this parallel method can be understood as a natural PML-variant of the well-known RAS (restricted additive Schwarz) method.

1.4.6 The sequential overlapping Schwarz methods

The following sequential methods are designed to be used when the subdomains have certain additional structure (see the definitions of "strips" and "checkerboards" in §1.4.7 below), but can be written abstractly for any set of subdomains as follows.

Forward-backward sweeping. Given $u_\times^0 \in H_0^1(\Omega)$, let $u_j^0 := u_\times^0|_{\Omega_j}$. Then, for $n \geq 0$,

1. (Forward sweeping) For $j = 1, \ldots, N$, let $u_j^{2n+1} \in H^1(\Omega_j)$ be the solution to

$$\begin{cases} P_s^j u_j^{2n+1} = P_s^j(u_{j,n}^\rightarrow|_{\Omega_j}) - (P_s u_{j,n}^\rightarrow)|_{\Omega_j} + f|_{\Omega_j} \in H^{-1}(\Omega_j) \\ u_j^{2n+1} = u_{j,n}^\rightarrow \in H^{1/2}(\partial\Omega_j), \end{cases} \tag{1.10}$$

where

$$u_{j,n}^\rightarrow := \sum_{\ell < j} \chi_\ell u_\ell^{2n+1} + \sum_{j \leq \ell} \chi_\ell u_\ell^{2n} \in H_0^1(\Omega).$$

Then set

$$u_\times^{2n+1} := u_{N+1,n}^\rightarrow = \sum_{\ell=1}^N \chi_\ell u_\ell^{2n+1} \in H_0^1(\Omega) \tag{1.11}$$

(where the subscript $\times$ indicates that these are the iterates for a sequential/multiplicative method).

2. (Backward sweeping) For $j = N, \ldots, 1$, let $u_j^{2n+2} \in H^1(\Omega_j)$ be the solution to

$$\begin{cases} P_s^j u_j^{2n+2} = P_s^j(u_{j,n}^\leftarrow|_{\Omega_j}) - (P_s u_{j,n}^\leftarrow)|_{\Omega_j} + f|_{\Omega_j} \in H^{-1}(\Omega_j) \\ u_j^{2n+2} = u_{j,n}^\leftarrow \in H^{1/2}(\partial\Omega_j), \end{cases}$$

where

$$u_{j,n}^\leftarrow := \sum_{\ell \leq j} \chi_\ell u_\ell^{2n+1} + \sum_{j < \ell} \chi_\ell u_\ell^{2n+2} \in H_0^1(\Omega).$$

Then set

$$u_\times^{2n+2} := u_{N+1,n}^\leftarrow = \sum_{\ell=1}^N \chi_\ell u_\ell^{2n+2} \in H_0^1(\Omega). \tag{1.12}$$

Observe that the algorithm only uses the local approximations $\{u_j^m\}_{j=1}^N$; i.e., $u_\times^{m-1}$ is not directly used to define $\{u_j^m\}_{j=1}^N$. However, the approximation to the solution of (1.5) at each step is $u_\times^m$, obtained by combining the local approximations u_j^m via the partition of unity in (1.11)/(1.12).

In forward-backward sweeping, the subdomains are always visited in the same order and its reverse. Some of our results about the sequential method involve changing the order in which the subdomains are visited from iteration to iteration, as described as follows.

General sequential method. Let $\{\sigma_n\}_n$ be a sequence of permutations $\{1, \ldots, N\} \mapsto \{1, \ldots, N\}$. For each n, σ_n corresponds to the order that that subdomains are visited in the nth sweep; e.g., $\sigma_n(1)$ is the first subdomain visited on the nth sweep, $\sigma_n(2)$ is the second subdomain visited on the nth sweep, etc. We therefore call $\{\sigma_n\}_n$ a sequence of orderings.

We highlight immediately that forward-backward sweeping is defined by $\sigma_{2n+1}(j) = j, \sigma_{2n}(j) = N - j + 1$ (i.e., odd-numbered sweeps visit the subdomains in order, and then even-numbered sweeps visit the subdomains in reverse order).

Given $u_\times^0 \in H_0^1(\Omega)$, let $u_j^0 := u_\times^0|_{\Omega_j}$. Then for $n \geq 1$, for $j = 1, \ldots, N$, let $u_{\sigma_n(j)}^{n+1} \in H^1(\Omega_{\sigma_n(j)})$ be the solution to

$$
\begin{cases}
P_s^{\sigma_n(j)} u_{\sigma_n(j)}^{n+1} = P_s^{\sigma_n(j)}(u_{\sigma_n(j),n}^*|_{\Omega_{\sigma_n(j)}}) - (P_s u_{\sigma_n(j),n}^*)|_{\Omega_{\sigma_n(j)}} + f|_{\Omega_{\sigma_n(j)}} \in H^{-1}(\Omega_{\sigma_n(j)}), \\
u_{\sigma_n(j)}^{n+1} = u_{\sigma_n(j),n}^* \in H^{1/2}(\partial\Omega_j),
\end{cases}
\tag{1.13}
$$

where

$$
u_{\sigma_n(j),n}^* := \sum_{\ell < j} \chi_{\sigma_n(\ell)} u_{\sigma_n(\ell)}^{n+1} + \sum_{\ell \geq j} \chi_{\sigma_n(\ell)} u_{\sigma_n(\ell)}^n \in H_0^1(\Omega).
$$

Then set

$$
u_\times^{n+1} := u_{N+1,n}^* = \sum_{\ell=1}^N \chi_\ell u_\ell^{n+1} \in H_0^1(\Omega).
$$

1.4.7 Strips and checkerboards

For $N_\ell \in \mathbb{Z}^+$, $\ell = 1, \ldots, d$, we say that $\{\Omega_j\}_{j=1}^N$ is a *checkerboard* of size $N_1 \times \ldots \times N_\ell$ if the following hold.

(i) For $\ell = 1, \ldots, d$, there exist
$$
0 = y_0^\ell < \ldots < y_{N_\ell}^\ell = L_\ell
$$
such that
$$
\overline{\Omega_{\text{int}}} = \bigcup_{j=1}^N \overline{U_j}, \quad U_j \cap U_i = \emptyset \text{ if } j \neq i,
$$
with each U_j of the form $\prod_{\ell=1}^d (y_{m_\ell-1}^\ell, y_{m_\ell}^\ell)$ for some $m_\ell \in \{1, \ldots, N_\ell\}$, $\ell = 1, \ldots, d$.

(ii) The overlapping decomposition $\{\Omega_{\text{int},j}\}_{j=1}^N$ of Ω_{int} comes from extending each U_j in each coordinate direction (apart from at $\partial\Omega$).

(iii) The extensions in (ii) are such that, for all j, Ω_j overlaps only with $\Omega_{j'}$, where Ω_j and $\Omega_{j'}$ are extensions of adjacent nonoverlapping subdomains.

We say that $\{\Omega_j\}$ is a $\mathfrak{d}$-checkerboard if $\mathfrak{d} := |\{\ell, \ N_\ell \neq 1\}|$ i.e. $\mathfrak{d}$ is the effective dimensionality of the checkerboard.

We say that $\{\Omega_j\}_{j=1}^N$ is a *strip* if it is a 1-checkerboard. To simplify notation, we then assume that the subdomains of a strip are ordered monotonically (i.e., so that Ω_i only overlaps Ω_{i-1} and Ω_{i+1}).

Figure 1.1 shows an example of the grid formed by the points $\{y_m^\ell\}$ that is used in the definition of a particular 2-checkerboard of size 4×5, along with a particular subdomain $\Omega_{\text{int},j}$.

1.5 Statement of the results about strips with $c \equiv 1$

Theorem 1.1 ($c \equiv 1$, **strip, parallel method**). *Assume that $c \equiv 1$ and that $\{\Omega_j\}_{j=1}^N$ is a strip. Then, for any $k_0, M, s > 0$, there exists $C > 0$ such that the following is true for any $f \in H^{-1}(\Omega)$ and $k \geq k_0$. If u is the solution to (1.5), and $\{u_+^n\}_{n\geq 0}$ is a sequence of iterates for the parallel overlapping Schwarz method, then*

$$
\|u - u_+^N\|_{H_k^s(\Omega)} \leq C k^{-M} \|u - u_+^0\|_{H_k^1(\Omega)}.
$$

The norm $\| \cdot \|_{H_k^s(\Omega)}$ appearing in Theorem 1.1 is defined for $s \in \mathbb{N}$ by

$$
\|v\|_{H_k^s(\Omega)}^2 := \sum_{|\alpha| \leq s} \|(k^{-1}\partial)^\alpha v\|_{L^2(\Omega)}^2,
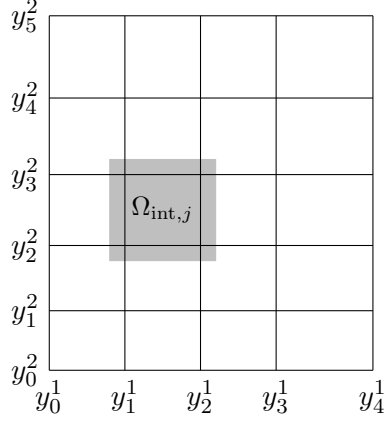\tag{1.14}
$$

7

Figure 1.1: The grid formed by the points $\{y_m^\ell\}$ that is used in the definition of a particular 2-checkerboard of size 4×5, and the subdomain $\Omega_{\text{int},j}$ formed by extending $(y_1^1, y_2^1) \times (y_2^2, y_3^2)$

and for $s \in (0, \infty) \setminus \mathbb{Z}^+$ by interpolation; i.e., $\| \cdot \|_{H_k^s(\Omega)}$ is the standard norm on $H^s(\Omega)$ except that every derivative is weighted by k^{-1} (using this norm is motivated by the fact that Helmholtz solutions oscillate at frequency approximately k).

Theorem 1.2 ($c \equiv 1$, strip, forward-backward sweeping). *Assume that $c \equiv 1$ and that $\{\Omega_j\}_{j=1}^N$ is a strip. Then, for any $k_0, M, s > 0$, there exists $C > 0$ such that the following is true for any $f \in H^{-1}(\Omega)$ and $k \geq k_0$. If u is the solution to (1.5), and $\{u_\times^n\}_{n \geq 0}$ is a sequence of iterates for the forward-backward sweeping overlapping Schwarz method, then*

$$\|u - u_\times^2\|_{H_k^s(\Omega)} \leq C k^{-M} \|u - u_\times^0\|_{H_k^1(\Omega)}$$

(i.e., after one forward and one backward sweep, the error is $O(k^{-\infty})$).

1.6 Statement of the results about checkerboards with $c \equiv 1$

Theorem 1.3 ($c \equiv 1$, checkerboard, parallel method). *Assume that $c \equiv 1$ and that $\{\Omega_j\}_{j=1}^N$ is a $\mathfrak{d}$-checkerboard of size $N_1 \times \cdots \times N_{\mathfrak{d}}$ (with $N_i \geq 2$ for all $1 \leq i \leq \mathfrak{d}$). Then, for any $k_0, M, s > 0$, there exists $C > 0$ such that the following is true for any $f \in H^{-1}(\Omega)$ and $k \geq k_0$. If u is the solution to (1.5), and $\{u_+^n\}_n \geq 0$ is a sequence of iterates for the parallel overlapping Schwarz method, then*

$$\|u - u_+^{N_1 + \cdots + N_{\mathfrak{d}} - (\mathfrak{d}-1)}\|_{H_k^s(\Omega)} \leq C k^{-M} \|u - u_+^0\|_{H_k^1(\Omega)}.$$

Theorem 1.3 contains Theorem 1.1 by recalling that a strip is a 1-checkerboard (with N_1 subdomains), and setting $\mathfrak{d} = 1$ in Theorem 1.3.

To state the result for the sequential method on the checkerboard, we need the following notion. Informally, we say that a sequence of orderings $\{\sigma_n\}$ is *exhaustive* if (i) given any ordering in the sequence, the sequence also contains the reverse ordering (i.e., where the subdomains are visited in reverse order), and (ii) given any directed straight line, there exists n_* such that, in the order of subdomains corresponding to the ordering σ_{n_*}, the subdomains intersected by the straight line are visited in that order (but not necessarily one after the other) in that sweep. For the precise definition of exhaustive, see Definition 7.2 below.

For example, one forward sweep and one backward sweep on a strip corresponds to the ordering $\{\sigma_1, \sigma_2\}$ with $\sigma_1(j) = j$, $\sigma_2(j) = N - j + 1$; this ordering is exhaustive, since (i) the backward sweep is the reverse of the forward sweep (and vice versa), and (ii) any straight line

that intersects more than one subdomain intersects them in the order they appear *either* in the forward sweep *or* in the backward sweep.

Examples 7.3 and 7.4 below describe how to easily construct exhaustive orderings of size $2^{\mathfrak{d}}$ for a $\mathfrak{d}$-checkerboard. In particular, one can take $\{\sigma_n\}_{1 \leq n \leq 2^{\mathfrak{d}}}$ corresponding to lexicographic order with respect to every vertex; see Figure 1.2. Lemma 7.5 shows that an exhaustive sequence of orderings for a 2-checkerboard must contain at least $2^2 = 4$ orderings.

7	8	9		9	8	7		1	2	3		3	2	1
4	5	6		6	5	4		4	5	6		6	5	4
1	2	3		3	2	1		7	8	9		9	8	7

Figure 1.2: A sequence of lexicographic orderings of a 3×3 checkerboard that is exhaustive in the sense of Definition 7.2.

Theorem 1.4 ($c \equiv 1$, checkerboard, sequential method). *Assume that $c \equiv 1$ and that $\{\Omega_j\}_{j=1}^N$ is a $\mathfrak{d}$-checkerboard. Then, for any $k_0, M, s > 0$, there exists $C > 0$ such that the following is true for any $f \in H^{-1}(\Omega)$ and $k \geq k_0$. If u is the solution to (1.5), and $\{u_\times^n\}_{n \geq 0}$ is a sequence of iterates for the general sequential overlapping Schwarz method following an exhaustive sequence of orderings of size S, then*

$$\|u - u_\times^S\|_{H_k^s(\Omega)} \leq Ck^{-M}\|u - u_\times^0\|_{H_k^1(\Omega)}.$$

In particular, we can construct orderings (via Examples 7.3 and 7.4) so that

$$\|u - u_\times^{2^{\mathfrak{d}}}\|_{H_k^s(\Omega)} \leq Ck^{-M}\|u - u_\times^0\|_{H_k^1(\Omega)}.$$

Remark 1.5 (Comparing the number of solves for the parallel and sequential methods on a $N_1 \times \cdots \times N_1$ checkerboard). *For a $N_1^d = N_1 \times \cdots \times N_1$ checkerboard ($N_1 \geq 2$) Theorem 1.3 ensures smooth and $O(k^{-\infty})$ error after $(dN_1 - 1)$ iterations each containing N_1^d parallel subdomain solves, whereas Theorem 1.4 ensures smooth and $O(k^{-\infty})$ error after 2^d iterations each containing N_1^d sequential subdomain solves.*

1.7 Statement of the most general result about the parallel method

We first recall the definition of the geometric-optic rays associated with the Helmholtz equation (1.1). Let $P := -k^{-2}\Delta - c^{-2}$ and let

$$p(x, \xi) := |\xi|^2 - c(x)^{-2}, \quad x, \xi \in \mathbb{R}^d.$$

Given initial data $x_0, \xi_0 \in \mathbb{R}^d$, let $\Phi_t(x_0, \xi_0) := (x(t), \xi(t))$ be the solution of Hamilton's equations

$$\frac{dx}{dt}(t) = \nabla_\xi p\big(x(t), \xi(t)\big), \quad \frac{d\xi}{dt}(t) = -\nabla_x p\big(x(t), \xi(t)\big) \tag{1.15}$$

with initial conditions $x(0) = x_0$ and $\xi(0) = \xi_0$; we call such solutions *trajectories of the flow associated to P*. The geometric-optic rays are then the spatial projections of the trajectories, i.e., their projections in the x-variable. Observe that, when $c \equiv 1$, the solution of (1.15) is just straight line motion $x(t) = x_0 + 2t\xi_0$, $\xi(t) = \xi_0$. Observe also the *time reversibility* property that if $(x(t), \xi(t))$ is a solution to (1.15) then so is $(x(-t), -\xi(-t))$.

The operator P is *non-trapping* if, given $R > 0$ there exists $T > 0$ such that, the x-projection of any trajectory starting in $\{x : |x| < R\}$ has left this set by time $T(R)$; see, e.g., [34, Page 278]. (Strictly this is the definition of being non-trapping forward in time, but, due to time reversibility, this is equivalent to being non-trapping backward in time.)

Theorem 1.6 below depends on a quantity $\mathcal{N}$ that depends on the properties of the trajectories. A precise definition of $\mathcal{N}$ is given in §6.1 below (see also the informal discussion in §2.2), but, informally, we consider trajectories for the flow associated to P travelling from the PML region of one subdomain to the PML region of another subdomain, and so on. $\mathcal{N}$ is then the maximum number of subdomains, counted with their multiplicity, that any such trajectory can travel between. Importantly, $\mathcal{N} < \infty$ when P is nontrapping.

Theorem 1.6 (General c, arbitrary hyperrectangular subdomains, parallel method). *Assume that P is non-trapping. Then, for any $k_0, M, s > 0$, there exists $C > 0$ such that the following is true for any $f \in H^{-1}(\Omega)$ and $k \geq k_0$. If u is the solution to (1.5), and $\{u_+^n\}_{n \geq 0}$ is a sequence of iterates for the parallel overlapping Schwarz method, then*

$$\|u - u_+^{\mathcal{N}}\|_{H_k^s(\Omega)} \leq Ck^{-M}\|u - u_+^0\|_{H_k^1(\Omega)}.$$

When $c \equiv 1$, for a strip, Theorem 1.1 follows from Theorem 1.6 by showing that $\mathcal{N} = N$, and, for a checkerboard, Theorem 1.3 follows from Theorem 1.6 by showing that $\mathcal{N} = N_1 + \ldots + N_{\mathfrak{d}} - (\mathfrak{d} - 1)$; see §7.3 below.

When $c \not\equiv 1$, the quantity $\mathcal{N}$ can, in principle, be calculated from its definition (Definition 2.3) using ray tracing.

Remark 1.7 (The relationship of PML to the optimal subdomain boundary conditions). *Recall from §1.2 that the optimal boundary condition on a particular DD subdomain is the Dirichlet-to-Neumann map for the Helmholtz equation posed in the exterior of that subdomain (as a subset of the whole domain); in particular, the parallel overlapping Schwarz method with a strip decomposition with N subdomains converges in N iterations with these boundary conditions by [42, Proposition 2.4] and [16, Lemma 2.2].*

Theorem 1.1 shows that when $c \equiv 1$ the parallel method with a strip decomposition with PML at the subdomain boundaries has $O(k^{-\infty})$ error after N iterations; i.e., PML approximates well the optimal boundary condition in this case.

When $c \not\equiv 1$, Theorem 1.6 proves that the parallel method with a strip decomposition has $O(k^{-\infty})$ error after $\mathcal{N}$ iterations; we highlight that $\mathcal{N}$ can be much larger than N when $c \not\equiv 1$ and the rays are not straight lines (e.g., if the rays "turn around" at one end of the domain), and PML does not approximate well the optimal boundary condition in this case. The fact that PML at a subdomain boundary condition cannot take into account scattering occurring outside that subdomain was highlighted in [24, §4] and [25, §10].

2 Outline of how the main results are proved

§2.1-§2.4 explain the ideas behind the results for the parallel method (i.e., Theorems 1.1, 1.3, and 1.6). §2.5 and §2.6 then explain the results for the sequential methods (Theorem 1.2 and 1.4).

2.1 The error propagation matrix T for the parallel method

For any $n \geq 1$ and $1 \leq j \leq N$, we define the global and local errors

$$e_+^n := u - u_+^n \in H_0^1(\Omega) \qquad \text{and} \qquad e_{+,j}^n := u|_{\Omega_j} - u_j^n \in H^1(\Omega_j). \tag{2.1}$$

These definitions, the definition of the iterate u_+^n (1.8), and the fact that $\{\chi_j\}_{j=1}^N$ is a partition of unity (1.6) imply that

$$e_+^n = \sum_{j=1}^N \chi_j e_{+,j}^n. \tag{2.2}$$

We introduce the notation

$$\epsilon_{+,j}^n := \chi_j e_{+,j}^n \quad \text{so that} \quad e_+^n = \sum_{j=1}^N \epsilon_{+,j}^n. \tag{2.3}$$

We can interpret $\epsilon_{+,j}^n = \chi_j e_{+,j}^n$, after extension by zero from Ω_j, as an element of $H_0^1(\Omega)$, since $e_{+,j}^n$ is zero on $\partial\Omega_j \cap \partial\Omega$ and $\chi_j = 0$ on $\partial\Omega_j \setminus \partial\Omega$. We call $\epsilon_{+,j}^n$ the *local physical error*. We highlight that the fact that e_+^n is expressed in (2.3) in terms of $\epsilon_{+,j}^n = \chi_j e_{+,j}^n$ and not the whole local error $e_{+,j}^n$ is expected, since $e_{+,j}^n$ contains contributions from the PML layers of Ω_j that are not PML layers of Ω, and hence not part of the original PDE problem (1.5) (and hence "unphysical").

For $1 \le i \le N$ and $v \in H_0^1(\Omega)$ let $\mathcal{T}_i v \in H^1(\Omega_j)$ be the solution to

$$\begin{cases} P_s^i \mathcal{T}_i v = P_s^i(v|_{\Omega_i}) - (P_s v)|_{\Omega_i} & \text{in } \Omega_i, \\ \mathcal{T}_i v = v & \text{on } \partial\Omega_i. \end{cases} \tag{2.4}$$

For technical reasons, we need a set of cut-off functions that have bigger support than the χ_js. Let $\{\chi_j^>\}_{j=1}^N$ be such that, for each j, $\chi_j^> \in C^\infty(\mathbb{R}^d)$, $\mathrm{supp}\chi_j^> \cap \Omega \subset (\Omega_j \setminus \mathrm{supp}(P_s^j - P_s))$ and $\chi_j^> \equiv 1$ on $\mathrm{supp}\chi_j$ (recall that $\Omega_j \setminus \mathrm{supp}(P_s^j - P_s) = \Omega_{\mathrm{int},j}$ for an interior subdomain); such $\{\chi_j^>\}_{j=1}^N$ exist since the distance between $\mathrm{supp}\chi_j$ and $\Omega_j \setminus \mathrm{supp}(P_s^j - P_s)$ is > 0 as a consequence of $\mathrm{supp}\chi_j$ and $\mathrm{supp}(P_s^j - P_s)$ being closed. The functions $\{\chi_j^>\}_{j=1}^N$ and $\{\chi_j\}_{j=1}^N$ are illustrated in the simple case of two subdomains (i.e, $N = 2$) in Figure 2.1.
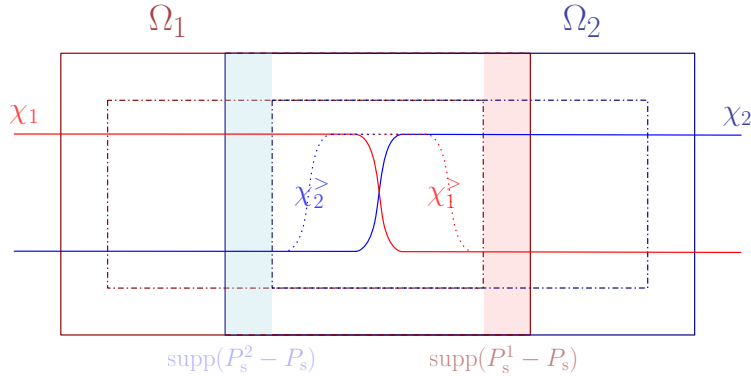

Figure 2.1: The functions $\{\chi_j^>\}$ and $\{\chi_j\}$ for two subdomains

We define the *physical error propagation matrix* $\mathbf{T}$ as

$$\mathbf{T} := (\chi_i \mathcal{T}_i \chi_j^>)_{1 \le i,j \le N} : (H_0^1(\Omega))^N \mapsto (H^1(\Omega))^N \tag{2.5}$$

and, for any $n \ge 0$, the physical errors vector $\boldsymbol{\epsilon}_+^n \in (H_0^1(\Omega))^N$ as

$$(\boldsymbol{\epsilon}_+^n)_j := \epsilon_{+,j}^n, \ 1 \le j \le N.$$

Lemma 2.1 (Physical error propagation for the parallel method). $\mathbf{T} : (H_0^1(\Omega))^N \mapsto (H_0^1(\Omega))^N$ *and*

$$\boldsymbol{\epsilon}_+^{n+1} = \mathbf{T}\boldsymbol{\epsilon}_+^n \quad \text{for all } n \ge 0. \tag{2.6}$$

Lemma 2.1 is proved in §5.1 below. The analogue of $\mathbf{T}$ when impedance boundary conditions are imposed on the subdomains was first studied in [26, §3].

2.2 Relating powers of **T** to trajectories of the flow defined by P

By (2.6), $\epsilon_+^n = \mathbf{T}^n \epsilon_+^0$. By the definition of $\mathbf{T}$ (2.5), powers of $\mathbf{T}$ involve compositions of the maps $\chi_i \mathcal{T}_i \chi_j^>$, and these compositions applied to ϵ_+^0 can be interpreted as the error travelling from subdomain to subdomain through the iterations.

The key ingredient to the proof of Theorem 1.6 is Lemma 2.4 below, which relates compositions of $\chi_i \mathcal{T}_i \chi_j^>$ to properties of the trajectories of the flow travelling between the relevant subdomains.

We now introduce notation to describe trajectories travelling from subdomain to subdomain. Let $\mathcal{W}$ be the set of finite sequences of elements of $\{1, \ldots, N\}$ (i.e., the indices of subdomains) such that, for any $w = \{w_i\}_{1 \le i \le n} \in \mathcal{W}$, $w_i \ne w_{i+1}$ for all $1 \le i \le n$ (i.e., the $(i+1)$th component of w corresponds to a different subdomain than the ith component). We call the elements of $\mathcal{W}$ *words*.

Definition (Informal statement of Definition 6.1 (following a word)). *A trajectory γ for the flow associated to P follows a word $w \in \mathcal{W}$ of length ≥ 2 if it passes through, in order, $\Omega_{w_2}, \ldots, \Omega_{w_n}$ by going through the following sets*

$$\operatorname{supp}\chi_{w_1}^> \cap \operatorname{supp}(P_s^{w_2} - P_s), \quad \ldots \quad, \quad \operatorname{supp}\chi_{w_{n-1}}^> \cap \operatorname{supp}(P_s^{w_n} - P_s), \quad \operatorname{supp}\chi_{w_n}^>. \tag{2.7}$$

We make the following immediate remarks:

- Figure 2.2 illustrates a trajectory following a word involving three subdomains of Ω that do not touch the boundary.

- To see where the sets in (2.7) come from, observe that, by the definition of $\mathcal{T}_i$ (2.4), the functions $\mathcal{T}_i \chi_j^> v$ appearing in the action of $\mathbf{T}$ (2.5) satisfy a PDE on Ω_i with the operator P_s^i and right-hand side supported in $\operatorname{supp}\chi_j^> \cap \operatorname{supp}(P_s^i - P_s)$.

- $\operatorname{supp}\chi_{w_i} \cap \operatorname{supp}(P_s^{w_{i+1}} - P_s)$ is the part of the PML of $\Omega_{w_{i+1}}$ that intersects $\operatorname{supp}\chi_{w_i}$; because χ_j is zero on the support of $P_s^j - P_s$, the sets in (2.7) are disjoint.

- A trajectory γ can follow more than one word because of the overlap of the subdomains.

The flow associated to P only allows passage between certain sequences of subdomains. For example, if the $c \equiv 1$, the trajectories of the flow are straight lines, and thus, since the subdomains are hyperrectangles, in a strip domain there is no trajectory following the word $(1, 2, 1)$. To see this, consider the analogue of Figure 2.2 with the subdomain Ω_1 in the same horizontal strip as Ω_2 and Ω_3; in this modified diagram, a straight-line trajectory cannot leave the blue hatched area, go to the purple hatched area, and then come back to the blue hatched area.

Definition 2.2. *A word $w \in \mathcal{W}$ is* allowed *if there exists a trajectory for P that follows w.*

Definition 2.3.
$$\mathcal{N} := \sup_{w \text{ allowed}} |w|. \tag{2.8}$$

Since the sets in (2.7) are compact and disjoint, the assumption that P is nontrapping implies that the length of allowed words is bounded above, ie $\mathcal{N} < \infty$.

For any $w \in \mathcal{W}$ of length $|w| := n \ge 2$, we define the composite map

$$\mathscr{T}_w := (\mathbf{T})_{w_n, w_{n-1}} (\mathbf{T})_{w_{n-1}, w_{n-2}} \cdots (\mathbf{T})_{w_2, w_1}. \tag{2.9}$$

The presence of the cut-offs in the definition of $(\mathbf{T})_{i,j}$ (2.5) means that $\mathscr{T}_w$ is zero unless Ω_{w_j} and $\Omega_{w_{j+1}}$ overlap for all j.

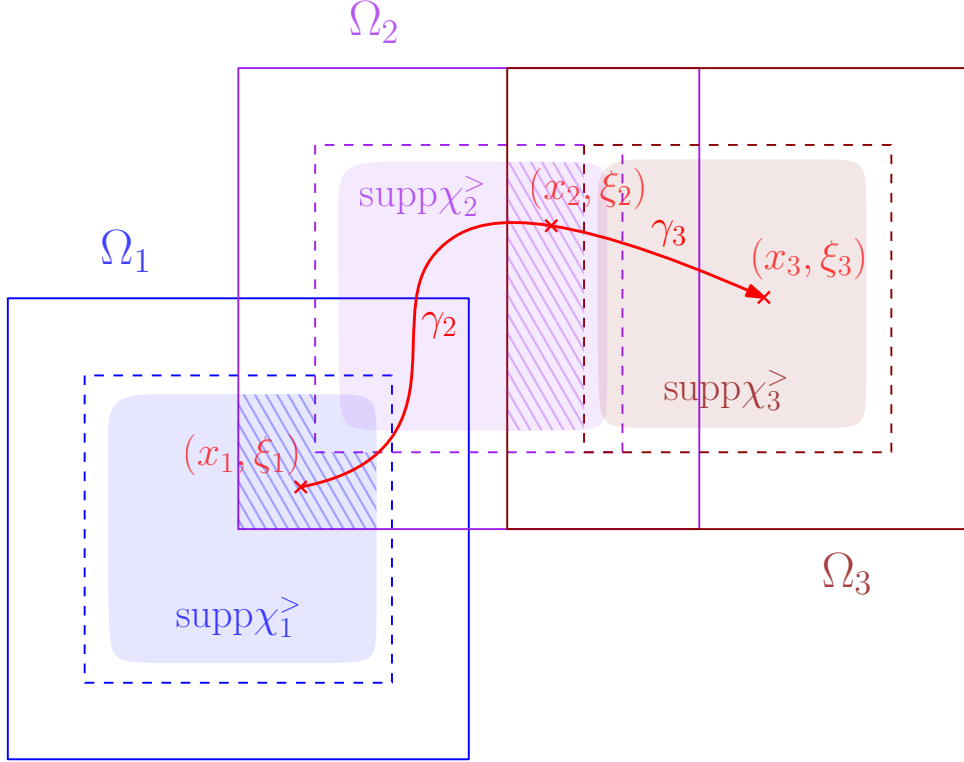The heart of the proof of Theorem 1.6 is the following lemma.

12

Figure 2.2: A trajectory (in red) following the word $(1, 2, 3)$ (here $\Omega_1, \Omega_2, \Omega_3$ are interior subdomains of Ω). The blue hatched shading indicates the domain $\mathrm{supp}\chi_1^> \cap \mathrm{supp}(P_s^2 - P_s)$, and the purple hatched shading indicates the domain $\mathrm{supp}\chi_2^> \cap \mathrm{supp}(P_s^3 - P_s)$. (The points (x_j, ξ_j), $j = 1, 2, 3$, and sub-trajectories γ_2, γ_3 are used in the precise definition of *follow* in Definition 6.1 below.)

Lemma 2.4.

If $w \in \mathcal{W}$ is not allowed then, for any $s \geq 1$, $\quad \|\mathscr{T}_w\|_{H_k^1(\Omega) \to H_k^s(\Omega)} = O(k^{-\infty}).$ (2.10)

Proof of Theorem 1.6 using Lemma 2.4. By the definition of the error propagation matrix $\mathbf{T}$ (2.5) and the composite map $\mathscr{T}_w$ (2.9), for any $m \in \mathbb{Z}^+$, all entries of $\mathbf{T}^m$ are sums of operators $\mathscr{T}_w$ over words w of size $m + 1$. Since words of size $\mathcal{N} + 1$ are not allowed by the definition of $\mathcal{N}$ (2.8), the result that $\|\mathbf{T}^\mathcal{N}\|_{(H_k^1(\Omega))^N \to (H_k^s(\Omega))^N} = O(k^{-\infty})$ follows from Lemma 2.4. $\qquad \square$

We show in §6.2 that Lemma 2.4 reduces to the following result, which describes precisely how, in the $k \to \infty$ limit, the error travels from subdomain to subdomain through the iterations. The lemma is stated in phase space (x, ξ), where x is position and ξ is momentum; note that ξ can be thought of as a Fourier variable measuring the direction of oscillation.

Lemma (Informal statement of Lemma 6.2). *For any $v \in H_0^1(\Omega)$, in the $k \to \infty$ limit, $\mathscr{T}_w v$ is non-negligible (i.e., not $O(k^{-\infty})$) only at points in phase space that are end points of trajectories that follow the word w.*

Lemma 6.2 is in turn proved using the following propagation result.

Lemma (Informal statement of Lemma 4.1). *Let $v \in H(\Omega_j)$ be the solution to*

$$P_s^j v = f \quad \text{in } \Omega_j, \qquad v = g \quad \text{on } \partial\Omega_j.$$

If P is nontrapping, then, in the $k \to \infty$ limit, v is non-negligible only at points in phase space that come from the data f under the flow associated to P.

13

Lemma 4.1 is stated rigorously and proved in §4, and we give the main ideas behind its proof in §2.3 below. Lemma 6.2 is stated in §6.2, and proved in §6.3. The idea is the following.

Sketch proof of Lemma 6.2 using Lemma 4.1. Given $v \in H_0^1(\Omega)$, by the definitions of $(\mathbf{T})_{i,j}$ (2.5) and $\mathcal{T}_j$ (2.4), $(\mathbf{T})_{w_2,w_1}v$ equals χ_{w_2} multiplied by a Helmholtz solution on Ω_{w_2} with data $(P_{\mathrm{s}}^{w_2} - P_{\mathrm{s}})\chi_{w_1}^{>}v$. Therefore, Lemma 4.1 shows that, in the $k \to \infty$ limit, the mass of $(\mathbf{T})_{w_2,w_1}v$ in phase space comes only from $\mathrm{supp}\chi_{w_1} \cap \mathrm{supp}(P_{\mathrm{s}}^{w_2} - P_{\mathrm{s}})$, i.e., the first domain in (2.7). (This concept of "mass in phase space" is made precise by the notion of *wavefront set*; see Definition B.6 below.)

Similarly, $(\mathbf{T})_{w_3,w_2}(\mathbf{T})_{w_2,w_1}v$ satisfies a Helmholtz problem on Ω_{w_w} with data $(P_{\mathrm{s}}^{w_3} - P_{\mathrm{s}})\chi_{w_2}(\mathbf{T})_{w_2,w_1}v$. Therefore, Lemma 4.1 shows that, in the $k \to \infty$ limit, the mass of $(\mathbf{T})_{w_3,w_2}(\mathbf{T})_{w_2,w_1}v$ comes only from the mass in $\mathrm{supp}\chi_{w_2} \cap \mathrm{supp}(P_{\mathrm{s}}^{w_3} - P_{\mathrm{s}})$ that in turn came from $\mathrm{supp}\chi_{w_1} \cap \mathrm{supp}(P_{\mathrm{s}}^{w_2} - P_{\mathrm{s}})$.

In the same way, $\mathcal{T}_w v$ only contains mass that propagates between subdomains following the word w, i.e., the subdomains listed in (2.7). □

2.3 Sketch proof of the propagation result (Lemma 4.1)

The two main concepts from semiclassical analysis used in the proof of Lemma 4.1 are the following.

1. Propagation of singularities, used in the form that the absence of mass (in phase space) of the solution of $Qu = f$ propagates forwards along the flow defined by Q as long as the trajectory does not intersect the data f (see Lemma B.11 below) and the imaginary part of the principal symbol of Q is non-positive.

2. Semiclassical ellipticity: an operator that is elliptic (i.e., has non-zero principal symbol, see §B.4 below) in some region of phase space can be inverted in that region, up to some error that is $O(k^{-\infty})$ (see Lemma B.9 below).

Since P_{s}^j is elliptic in the directions that the scaling of the PML takes place (Lemma 3.5),

3. A trajectory cannot exit Ω_j without passing through a point where P_{s}^j is elliptic (Lemma 4.4).

We now face the issue that the assumptions and conclusions of Lemma 4.1 involve the flow for P (since this is the flow with physical relevance), but to prove the lemma we need to use propagation of singularities in the flow defined by P_{s}^j (since Lemma 4.1 is all about solutions of $P_{\mathrm{s}}^j u = f$), and these two flows are different. The resolution of this apparent difficulty is that

4. The flow defined by P_{s}^j equals the flow defined by P as long as it doesn't reach a point where P_{s}^j is elliptic (Lemma 4.2).

By the assumption that P is nontrapping, any trajectory exits the domain; combining this with Points 3 and 4 we obtain

5. Any point in Ω_j is the endpoint of a trajectory of P_{s}^j coming *either* from an elliptic point without intersecting the data, *or* comes from the data.

Point 5 combined with Point 2 implies that trajectories of P_{s}^j coming from an elliptic point without intersecting the data carry no mass. Therefore Point 5 combined with Point 1 implies the result; i.e., that the solution contains mass only on points in phase space that come from the data under the flow associated to P (this is illustrated graphically in Figure 4.2, next to the proof of Lemma 4.1).

Remark 2.5 (Why we assume that f_{s} is linear near the PML boundary). *The principal symbol of the scaled operator P_{s} has a negative imaginary part; see (3.4) below. For such an operator, the relevant propagation of singularities results in the semiclassical calculus for domains with a boundary have not yet been written down in the literature. Indeed, the relevant propagation of singularities results in the semiclassical calculus has been proved (i) on manifolds without boundary for operators with symbols whose imaginary parts are single-signed [14, Theorem E.47] (see Lemma B.10 below) and (ii) on manifolds with boundary for P [57].*

The assumption that f_{s} is linear near the PML boundary allows us to use a reflection argument (similar to that used for the Cartesian PML analyses in 2-d in [6, Proof of Theorem 5.5], [10, Lemma 3.4]) to extend the solution past $\partial\Omega$ and use the propagation results of (i) above, i.e., bypassing the issue of propagation up to the boundary (exactly where in the proofs of the main results we do this is highlighted in Remark 6.3 below). Once the relevant propagation results are written down in the literature, this assumption can be removed.

Remark 2.6. (Why we do not have a variable coefficient in the highest-order term of the PDE.) *There is also large interest in solving the Helmholtz equation with a variable coefficient in the highest-order term; i.e., the PDE*

$$k^{-2}\nabla \cdot (A\nabla u) + c^{-2}u = -f \tag{2.11}$$

where A is a symmetric-positive-definite-matrix-valued function with $\mathrm{supp}(I - A) \subset \Omega_{\mathrm{int}}$. When $f \in L^2(\Omega_{\mathrm{int}})$ and u satisfies the Sommerfeld radiation condition, the solution of this problem is well-approximated by Cartesian PML. Indeed, Theorem 4.6 below holds verbatim: since $A \equiv I$ when the PML is active, the imaginary part of the principal symbol is unaffected and propagation of singularities (as described informally in Point 1 above) still holds.

However, the analogue of the domain-decomposition methods defined in §1.4 for the PDE (2.11) involve local problems with the operator

$$P_{\mathrm{s}}^j u := k^{-2}\nabla_{\mathrm{s},j} \cdot (A\nabla_{\mathrm{s},j}u) + c^{-2}u, \tag{2.12}$$

where

$$\nabla_{\mathrm{s},j} := \begin{pmatrix} \dfrac{1}{1 + ig'_{1,j}(x_1)}\partial_{x_1} \\ \vdots \\ \dfrac{1}{1 + ig'_{d,j}(x_d)}\partial_{x_d} \end{pmatrix}$$

(compare to (1.4) and (1.3)). For general A, the imaginary part of the principal symbol of P_{s}^j (2.12) no longer has a sign, and propagation of singularities does not hold. This is more than just a technical problem with the proofs: the numerical experiments in §8.5 give an example of a simple A for which the PDE (2.11) is well posed, but the parallel Schwarz method applied to (2.11) with local problems involving (2.12) diverges.

2.4 Summary of the three ingredients used in the proof of Theorem 1.6 (the result about the parallel method)

1. Algebra of the error propagation, showing which words appear in the products $\mathscr{T}_w$ (discussed in §2.1 and §2.2).

2. Semiclassical analysis, showing that $\|\mathscr{T}_w\|_{H_k^1(\Omega)\to H_k^s(\Omega)} = O(k^{-\infty})$ if w is not allowed – Lemma 2.4.

3. Properties of the trajectories of the flow associated with P, dictating which w are allowed.

All the discussion in this section has so far been about the parallel method. However, the three ingredients above show that to prove results about different methods, such as the sequential methods in §1.4.6, we only need to understand Point 1, i.e., the algebra of the error propagation, and then check whether the words appearing in the products $\mathcal{T}_w$ are allowed or not. In the next subsection we illustrate this process for forward-backward sweeping on a strip when $c \equiv 1$ (i.e., the geometric-optic rays are straight lines).

2.5 Error propagation for forward-backward sweeping on a strip (i.e., the idea behind the proof of Theorem 1.2)

We now seek the sweeping analogue of the parallel error propagation result of Lemma 2.1. To do this, we set up notation for the error analogous to that used for the parallel method in §2.1.

Let $\{u_\times^n\}_{n \geq 0}$ be a sequence of iterates for the forward-backward sweeping defined in §1.4.6. Let

$$e_\times^n := u - u_\times^n \in H_0^1(\Omega), \quad \text{and} \quad e_{\times,j}^n := u - u_j^n, \quad 1 \leq j \leq N \tag{2.13}$$

(with the subscript $\times$ indicating that this is the error for the multiplicative/sequential method). We define the localised physical error by

$$\epsilon_{\times,j}^n := \chi_j e_{\times,j}^n \in H_0^1(\Omega), \quad \text{so that} \quad e_\times^n = \sum_{j=1}^N \epsilon_{\times,j}^n \tag{2.14}$$

(in exact analogue with (2.3)), Finally, we define the physical errors vector $\boldsymbol{\epsilon}_\times^n \in (H_0^1(\Omega))^N$ by $(\boldsymbol{\epsilon}_\times^n)_j := \epsilon_{\times,j}^n$.

For any set of subdomains, $(\mathbf{T})_{j,j} = 0$ for all $1 \leq j \leq N$ – this follows from the definitions of $\mathbf{T}$ (2.5) and $\mathcal{T}_j$ (2.4) and the fact that $\chi_j^> = 0$ where $P_s^j - P_s$ is supported. Therefore

$$\mathbf{T} = \mathbf{L} + \mathbf{U}, \tag{2.15}$$

where $\mathbf{L}$ is lower triangular and $\mathbf{U}$ is upper triangular, both with zero on the diagonal.

Using the structure (2.15), we arrive at the analogue of Lemma 2.1 (proved in §5.2 below).

Lemma 2.7 (Error propagation for forward-backward sweeping). *For forward-backward sweeping, for $n \geq 0$,*

$$\boldsymbol{\epsilon}_\times^{2n+1} = \mathbf{L}\boldsymbol{\epsilon}_\times^{2n+1} + \mathbf{U}\boldsymbol{\epsilon}_\times^{2n} \quad \text{and} \quad \boldsymbol{\epsilon}_\times^{2n+2} = \mathbf{L}\boldsymbol{\epsilon}_\times^{2n+1} + \mathbf{U}\boldsymbol{\epsilon}_\times^{2n+2}. \tag{2.16}$$

In the parallel case, Lemma 2.1 gives us immediately that $\boldsymbol{\epsilon}^n = \mathbf{T}^n \boldsymbol{\epsilon}^0$. We now seek the analogue of this for forward-backward sweeping (see (2.18) below), using the additional structure of $\mathbf{L}$ and $\mathbf{U}$ when $\{\Omega_j\}_{j=1}^N$ is a strip (in the sense of §1.4.7).

Lemma 2.8. *If $\{\Omega_j\}_{j=1}^N$ is a strip then*

$$(\mathbf{L})_{i,j} = \begin{cases} (\mathbf{T})_{j+1,j} & \text{if } i = j+1, \\ 0 & \text{otherwise,} \end{cases} \quad \text{and} \quad (\mathbf{U})_{i,j} = \begin{cases} (\mathbf{T})_{i,i+1} & \text{if } j = i+1, \\ 0 & \text{otherwise.} \end{cases} \tag{2.17}$$

Proof. By definition of the partition of unity $\{\chi_j\}_{j=1}^N$ for a strip decomposition, and the "slightly bigger" set of cut-offs $\{\chi_j^>\}_{j=1}^N$, if $|j - i| \neq 1$, then $\chi_i^> = 0$ on Ω_j. Thus $(\mathbf{T})_{i,j} = 0$ if $|j - i| \neq 1$, and the result (2.17) follows. □

Proof of Theorem 1.2 using Lemma 2.4. By the first equation in (2.16) with $n = 0$,

$$(I - \mathbf{L})\boldsymbol{\epsilon}_\times^1 = \mathbf{U}\boldsymbol{\epsilon}_\times^0;$$

hence, since $\mathbf{L}^N = 0$,

$$\epsilon_\times^1 = \sum_{\ell=0}^{N-1} \mathbf{L}^\ell \mathbf{U} \epsilon_\times^0.$$

Similarly, by the second equation in (2.16) and then the above,

$$\epsilon_\times^2 = \sum_{m=0}^{N-1} \mathbf{U}^m \mathbf{L} \epsilon_\times^1 = \sum_{0 \le \ell, m \le N-1} \mathbf{U}^m \mathbf{L}^{\ell+1} \mathbf{U} \epsilon_\times^0. \tag{2.18}$$

By Lemma 2.8, all the entries of $\mathbf{LU}$ consist of $(\mathbf{T})_{j+1,j}(\mathbf{T})_{j,j+1}$ for some j. Therefore, all the entries of any product of matrices appearing in the second sum in (2.18) are of the form $\mathcal{T}_w$, where $w = (\bar{w}, w_0)$ and $w_0 = (j+1, j, j+1)$ (and $\bar{w}$ is empty if $m = 0$ and $\ell = 0$ in the sum).

We have now completed the analogue of Point 1 in §2.4; i.e., we have characterised the words appearing the the error propagation.

Now, since $c \equiv 1$, the trajectories are straight lines and the word $(j+1, j, j+1)$ is not allowed (recall the discussion above Definition 2.2). If w_0 is not allowed, then the word $(\bar{w}, w_0)$ is not allowed (regardless of $\bar{w}$). Therefore, by Lemma 2.4, for any $s \ge 1$,

$$\left\| \sum_{0 \le \ell, m \le N-1} \mathbf{U}^m \mathbf{L}^{\ell+1} \mathbf{U} \right\|_{(H_k^1(\Omega))^N \to (H_k^s(\Omega))^N} = O(k^{-\infty})$$

and the result of Theorem 1.2 follows from (2.18) and (2.14). $\qquad\qquad\square$

2.6 Error propagation for the sequential method on a checkerboard (i.e., the idea behind the proof of Theorem 1.4)

The idea of the proof of the analogous result for the general sequential method (i.e., Theorem 1.4) is the same as for forward-backward sweeping discussed above; i.e., the goal is to show that the words appearing in the products $\mathcal{T}_w$ are not allowed. Theorem 1.4 is proved under the condition that the sequence of subdomain orderings is exhaustive (defined informally above Theorem 1.4 and defined precisely in Definition 7.2 below). This condition ensures that there are enough subdomain orderings so that the words that appear in the error propagation relation correspond to trajectories that come back on themselves, and so are not allowed when the geometric-optic rays are straight lines (i.e., when $c \equiv 1$).

Examples 7.3 and 7.4 construct exhaustive sequences of orderings with $2^\mathfrak{d}$ orderings, where $\mathfrak{d}$ is the effective dimension of the checkerboard, with Lemma 7.5 showing that this number of orderings is optimal when $\mathfrak{d} = 2$. We highlight that the source-transfer-type methods of [36, 37] also require 2^d "sweeps" on a d-checkerboard for the Helmholtz equation with $c \equiv 1$.

For the general sequential method on a checkerboard when $c \not\equiv 1$, the definition of exhaustive has to be modified to depend on the geometric-optic rays (which are no longer straight lines); see Remark 7.8 below.

2.7 Discussion on the importance (or not) of PML conditions on the subdomains

The discussion in §2.3 above shows that the main property of PML used is that P_{s}^j is elliptic in certain directions in the PML region of Ω_j. Appendix A describes how Theorems 1.1-1.6 hold for a much wider class of operators and subdomain boundary conditions, including when the PML is replaced by a *complex absorbing potential*; i.e., the Helmholtz operator in Ω_j is replaced by

$$-k^{-2} \nabla \cdot (A\nabla) - c^{-2} - iV_j$$

where $V_j \in C_{\mathrm{comp}}^\infty(\mathbb{R}^d)$ is supported in what was the PML region of Ω_j and is strictly positive near $\partial\Omega_j$ (see, e.g., [49, 45, 46, 41, 51]); see Example A.4. In particular, Example A.4 shows

that the analogues of Theorems 1.1-1.4 and 1.6 with complex absorption hold when $A \not\equiv I$, in contrast to PML (see Remark 2.6); this will be investigated further elsewhere. Note that complex absorbing potentials can themselves be used to approximate the radiation condition with $O(k^{-\infty})$ error – see Theorem A.2 below – and thus can also be imposed on $\partial\Omega$.

We highlight, however, that the propagation result Lemma 4.1 (stated informally in §2.2) does not hold if the PML is replaced by a local absorbing boundary conditions, such as the impedance boundary condition. Indeed, in the $k \to \infty$ limit, mass is reflected by these boundary conditions [40, 19] and thus the reflections from $\partial\Omega_j$ create mass that comes from the data *not* under the flow associated with the Helmholtz operator on Ω. These additional reflections make analysing DD methods with impedance boundary conditions challenging [2, 31].

3 Preliminary results for the PML operators

It will be convenient to work with the semiclassical small parameter $\hbar := k^{-1}$. We then have

$$P = -\hbar^2\Delta - c^{-2}, \qquad P_{\mathrm{s}} = -\hbar^2\Delta_{\mathrm{s}} - c^{-2}, \quad \text{and} \quad P_{\mathrm{s}}^j = -\hbar^2\Delta_{\mathrm{s},j} - c^{-2}, \quad 1 \le j \le N.$$

3.1 Useful notation in Ω_j

For any $1 \le j \le N$, we define the following, denoting the analogous subsets of Ω by the same notation but omitting the j.

- Let $\mathcal{X}_\ell^j$ for $0 \le \ell \le d-1$ denote the set of ℓ-dimensional generalised edges of Ω_j. For example, if $d = 3$, $\mathcal{X}_0^j$ are the vertices, $\mathcal{X}_1^j$ the edges, and $\mathcal{X}_2^j$ the faces; see Figure 3.1.

- Let $\mathcal{Z}^j := \cup_{0 \le \ell \le d-1} \mathcal{X}_\ell^j$ be the set of generalised edges of Ω_j.

- Let $\mathcal{Z}_{\partial\Omega}^j := \{\mathfrak{e} \in \mathcal{Z}^j, \text{ s.t. } \exists \mathfrak{e}_0 \in \mathcal{Z} \text{ with } \mathfrak{e} \subset \mathfrak{e}_0\}$, the generalised edges of Ω_j that are edges of Ω.

- For $0 \le \ell \le d-1$ and $\mathfrak{e} \in \mathcal{X}_\ell^j$, we write $\xi \parallel \mathfrak{e}$ (informally "ξ is parallel to $\mathfrak{e}$") if $\xi \in \hat{\mathfrak{e}}$, where $\hat{\mathfrak{e}} := \mathrm{Vect}(e_1, \cdots, e_\ell)$, where $(e_1, \cdots, e_\ell)$ are ℓ linearly independent vectors tangent to $\mathfrak{e}$; see Figure 3.2.

- For $\mathfrak{e} \in \mathcal{Z}^j$, let $\kappa(\mathfrak{e}) := \kappa$ if $\mathfrak{e} \in \mathcal{Z}_{\partial\Omega}^j$ and $\kappa(\mathfrak{e}) := \kappa_0$ otherwise.

- For $0 \le \eta < \kappa$, $0 \le \ell \le d-1$, let

$$\mathcal{E}_{\ell,\eta}^j := \bigcup_{\mathfrak{e} \in \mathcal{X}_\ell^j} \left(\mathfrak{e} + B(0, \kappa(\mathfrak{e}) - \eta)\right) \cap \Omega_j$$

(informally, $\mathcal{E}_{\ell,\eta}^j$ is the region a distance η into the ℓ-dimensional edges of the PML layer of Ω_j, taking into account that the PML layer is width κ for edges that are part of $\partial\Omega$ and κ_0 otherwise).

- For $0 \le \eta < \kappa$, $0 \le \ell \le d-1$ and $\mathfrak{e} \in \mathcal{X}_\ell^j$, let

$$\widetilde{\mathfrak{e}\eta} := \left\{x + z + \sum_{1 \le i \le d-\ell} t_i n_i(x), \quad x \in \mathfrak{e}, \ z \in \hat{\mathfrak{e}}, \ t_i > -(\kappa(\mathfrak{e}) - \eta)\right\}, \tag{3.1}$$

where, for any $x \in \mathfrak{e}$, $(n_i(x))_{1 \le i \le d-\ell}$ is a set of $d - \ell$ orthogonal outward-pointing (with respect to Ω_j) normal vectors to $\mathfrak{e}$. Informally, $\widetilde{\mathfrak{e}\eta}$ is the region a distance η into the part of the PML layer of Ω_j corresponding to edge $\mathfrak{e}$ that is then extended out of Ω_j to infinity,

both normally (corresponding to $t_i n_i(x)$ in (3.1)) and tangentially (corresponding to z in (3.1)); see Figure 3.2.

Let
$$\widetilde{\mathcal{E}}^j_{\ell,\eta} := \bigcup_{\mathcal{X} \in \mathcal{X}^j_\ell} \widetilde{\mathfrak{e}^\eta}.$$

- For $0 \le \eta < \kappa$, let $\mathcal{L}^j_\eta := \mathcal{E}^j_{d-1,\eta}$, i.e., the η-strict PML layer of Ω_j, and $\widetilde{\mathcal{L}}^j_\eta := \widetilde{\mathcal{E}}^j_{d-1,\eta}$, i.e., the η-strict extended PML layer of Ω_j.

- For $0 \le \eta < \kappa$, let $\mathcal{C}^j_\eta := \mathcal{E}^j_{0,\eta}$, i.e., the η-strict PML corners of Ω_j, $\widetilde{\mathcal{C}}^j_\eta := \widetilde{\mathcal{E}}^j_{0,\eta}$, i.e., the η-strict extended PML corners of Ω_j.

Additionally, when $\eta = 0$ in the above notation we omit it; e.g., $\widetilde{\mathfrak{e}^0}$ is written as $\widetilde{\mathfrak{e}}$ and $\mathcal{E}^j_{\ell,0}$ is written as $\mathcal{E}^j_\ell$.
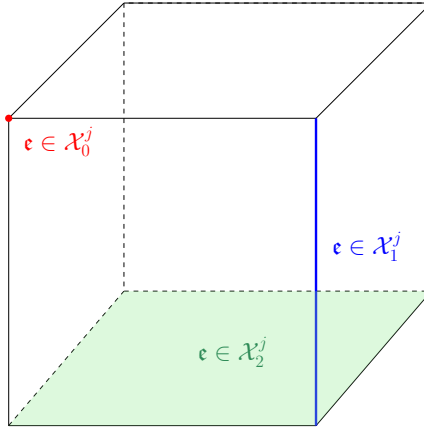

Figure 3.1: Examples of elements of $\mathcal{X}^j_0$, $\mathcal{X}^j_1$ and $\mathcal{X}^j_2$ for a 3-d subdomain.

3.2 Extensions of PML solutions in Ω_j

Odd-symmetric extension on parts of $\partial\Omega_j$ that intersect $\partial\Omega$.

Definition 3.1 (Odd-symmetric extension). *Let $\mathcal{V} \subset \mathbb{R}^d$ be such that there is a unique Euclidien coordinate system (z_1, z_2) so that $\mathcal{V} = \omega \cap \{z_1 \ge 0\}$, where $\omega \subset \mathbb{R}^d$ is open. For any $u : \mathcal{V} \mapsto \mathbb{R}$, we denote $S_\mathcal{V} u : \mathcal{V} \cup S\mathcal{V} \mapsto \mathbb{R}$ its odd-symetric extension to $\mathcal{V} \cup S\mathcal{V}$, $S\mathcal{V} := \{(-z_1, z_2), z \in \mathcal{V}\}$, defined by, for any $z \in \mathcal{V} \cup S\mathcal{V}$*
$$S_\mathcal{V} u(z) := \begin{cases} u(z_1, z_2) & \text{if } z \in \mathcal{V}, \\ -u(-z_1, z_2) & \text{if } z \notin \mathcal{V}. \end{cases}$$

Let $\epsilon_{\text{ext}} > 0$ be a parameter to be fixed later. We define the extended domain
$$\widetilde{\Omega}_j := \Omega_j \cup \bigcup_{\mathfrak{e} \in \mathcal{Z}^j_{\partial\Omega}} \left\{ x + \sum_{1 \le i \le d-\ell} t_i n_i(x), \quad x \in \mathfrak{e}, \, 0 \le t_i < \epsilon_{\text{ext}} \right\},$$

where, for any $x \in \mathfrak{e}$, $(n_i(x))_{1 \le i \le d-\ell}$ is a set of $d - \ell$ orthogonal outward-pointing (with respect to Ω_j) normal vectors to $\mathfrak{e}$; i.e., $\widetilde{\Omega}_j$ is Ω_j extended normally for a distance ϵ_{ext} from the edges of Ω_j that are (subset of) edges of Ω. Note that if $\partial\Omega_j \cap \partial\Omega = \emptyset$, then $\widetilde{\Omega}_j = \Omega_j$.

We now define an extension of functions from Ω_j to $\widetilde{\Omega}_j$ by symmetrising multiple times with respect to the generalised edges of Ω_j (and its partial extensions). The following definition is illustrated in Figure 3.3.
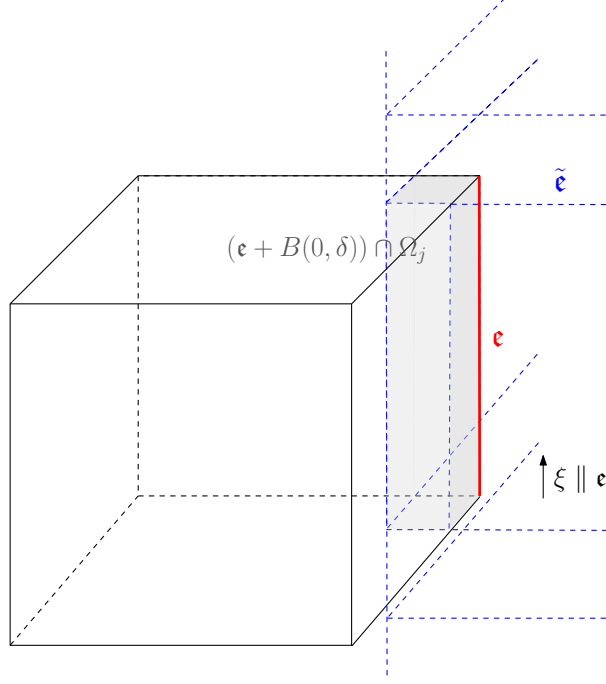
Figure 3.2: Illustration of some elements defined in §3.1 for a 3-d subdomain: an edge $\mathfrak{e} \in \mathcal{X}_1^j$, its associated infinitely extended region $\widetilde{\mathfrak{e}}$, the corresponding region $\big(\mathfrak{e} + B(0, \kappa(\mathfrak{e}) - \eta)\big) \cap \Omega_j$ in the PML of Ω_j, and a vector ξ such that $\xi \parallel \mathfrak{e}$.

Definition 3.2 (Extension from Ω_j to $\widetilde{\Omega}_j$). *For $0 \leq \ell \leq d$, we define the partial extension with respect to edges of dimension $\geq \ell$:*

$$\widetilde{\Omega}_j^\ell := \Omega_j \cup \bigcup_{\substack{\mathfrak{e} \in \mathcal{Z}_{\partial\Omega}^j \cap \mathcal{X}_m^j \\ m \geq \ell}} \left\{ x + \sum_{1 \leq i \leq d - \ell} t_i n_i(x), \quad x \in \mathfrak{e}, \ 0 \leq t_i < \epsilon_{\mathrm{ext}} \right\}.$$

Observe in particular that $\widetilde{\Omega}_j^0 = \widetilde{\Omega}_j$ and $\widetilde{\Omega}_j^d = \Omega_j$. For $1 \leq \ell \leq d$, define the extension S_j^ℓ from $\widetilde{\Omega}_j^\ell$ to $\widetilde{\Omega}_j^{\ell-1}$ in the following way. Let

$$\mathcal{E}_j^\ell := \left(\big(\partial\widetilde{\Omega}_j^\ell \cap \widetilde{\Omega}_j^{\ell-1}\big) + B(0, \epsilon_{\mathrm{ext}}) \right) \cap \Omega_j^\ell, \qquad S_j^\ell := S_{\mathcal{E}_j^\ell},$$

where $S_{\mathcal{E}_j^\ell}$ is defined piecewise by Definition 3.1. When multiples choices of symmetrisation are possible, we chose one arbitrarily. We then set

$$S_j := S_j^1 \circ S_2 \circ \cdots \circ S_j^d.$$

We now fix $\epsilon_{\mathrm{ext}} > 0$ small enough so that $\kappa - \epsilon_{\mathrm{ext}} > \kappa_{\mathrm{lin}}$ and χ_j does not vary normally in an ϵ_{ext}-neighbourhood of $\partial\Omega$ (recall from §1.4 that we assume that in a neighbourhood of $\partial\Omega$, each χ_j does not vary normally). As consequence of the latter,

$$\text{for all } j \in \{1, \cdots, N\}, \quad \big(S_j \chi_j v\big) = \chi_j \big(S_j v\big) \quad \text{for all } v \in H_0^1(\Omega_j). \tag{3.2}$$

Lemma 3.3 (Extension for homogeneous Dirichlet data). *Suppose $u \in H^1(\Omega_j)$, $f \in H^{-1}(\Omega_j)$ and $g \in H^{1/2}(\partial\Omega_j)$ are such that*

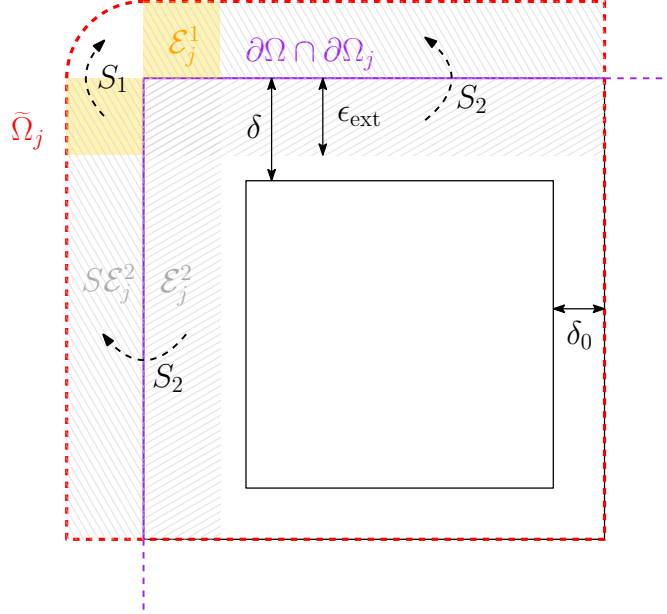$$\begin{cases} P_{\mathrm{s}}^j u = f & \text{in } \Omega_j, \\ u = g & \text{on } \partial\Omega_j. \end{cases}$$

20

Figure 3.3: Extension from Ω_j to $\widetilde{\Omega}_j$ for a subdomain Ω_j belonging to a corner of Ω.

Let $\widetilde{u} := S_j u$ and $\widetilde{f} := S_j f \in H^{-1}(\widetilde{\Omega}_j)$. If $g = 0$ on $\partial\Omega_j \cap \partial\Omega$, then $\widetilde{u} \in H^1(\widetilde{\Omega}_j)$ and

$$P_s^j \widetilde{u} = \widetilde{f} \ in \ \widetilde{\Omega}_j. \tag{3.3}$$

Although we have so far been considering P_s^j as an operator on Ω_j, it is defined on all of $\mathbb{R}^d$ by (1.3) and (1.4), and so its application in (3.3) on $\widetilde{\Omega}_j$ makes sense.

Proof of Lemma 3.3. To prove that $\widetilde{f} \in H^{-1}(\widetilde{\Omega}_j)$, the key calculation in 1-d is that

$$\int_{-\infty}^0 \big(-f(-x)\big)\phi(x)\,dx + \int_0^\infty f(x)\phi(x)\,dx = \int_0^\infty f(x)\Big(\phi(x) - \phi(-x)\Big)\,dx.$$

Since $H^{-1}(\Omega_j) = (H_0^1(\Omega_j))^*$, $\|S_j f\|_{H^{-1}(\mathbb{R})} \leq \|f\|_{(H_0^1(\mathbb{R}^+))^*}$. The result for $d \geq 2$ is then proved in an analogous way.

Since $\kappa - \epsilon_{\text{ext}} > \kappa_{\text{lin}}$, the extension only occurs when the PML scaling functions $g_{\ell,j}$ are either linear or zero. Therefore, by the definitions in §1.4.2, the coefficients of P_s^j are constant near $\partial\Omega_j$. The result (3.3) then follows from the definition of the weak derivative and integrating by parts. Indeed, in 1-d the basic calculation is that, for $\phi \in C_c^\infty(\mathbb{R})$, if $u(0) = 0$ and $u''(x) = f(x)$ for $x > 0$, then $\int_{-\infty}^\infty \widetilde{u}(x)\phi''(x)\,dx = \int_{-\infty}^\infty \widetilde{f}(x)\phi(x)\,dx$; i.e. $\widetilde{u}''(x) = \widetilde{f}(x)$ and so $P(S_j u) = S_j(Pu)$ with $Pu = u''$. For $d \geq 2$, the calculation is more complicated, but the basic idea is the same, crucially using that the coefficients of P_s^j are constant near $\partial\Omega_j$. $\qquad\square$

3.3 Semiclassical ellipticity statements for P_s^j

We begin with a computation quantifying the ellipticity of P_s^j at the symbolic level.

Lemma 3.4. *Let p_s^j denote the semiclassical principal symbol associated with P_s^j. Given a compact set $K \subset \mathbb{R}^d$, there exists $C > 0$ such that for all $x \in K$ and $\xi \in \mathbb{R}^d$ the following is true.*

1. (Bounding the symbol below by the PML scaling function)

$$|p_s^j(x,\xi)| \geq C \sum_{m=1}^d \xi_m^2 g_{x_m,j}'(x_m).$$

2. (Ellipticity at infinity in the ξ variable)

$$if \quad |\xi| \geq C \quad then \quad \langle\xi\rangle^{-2}|p_s^j(x, y, \xi_1, \xi_2)| \geq C^{-1}.$$

Furthermore, the same properties hold for P_s^j replaced with P_s.

Proof. Observe that

$$p_s^j(x, \xi) = \sum_{\ell=1}^{d} \frac{\xi_\ell^2 \left(1 - ig_{\ell,j}'(x_\ell)\right)^2}{\left(1 + g_{\ell,j}'(x_\ell)^2\right)^2} + c^{-2}(x) = \sum_{\ell=1}^{d} \xi_\ell^2 \frac{\left(1 - g_{\ell,j}'(x_\ell)^2 - 2ig_{\ell,j}'(x_\ell)\right)}{\left(1 + g_{\ell,j}'(x_\ell)^2\right)^2} + c^{-2}(x),$$

so that

$$\operatorname{Im} p_s^j(x, \xi) = -2\sum_{\ell=1}^{d} \xi_\ell^2 \frac{g_{\ell,j}'(x_\ell)}{\left(1 + g_{\ell,j}'(x_\ell)^2\right)^2} \quad \text{and} \quad \operatorname{Re} p_s^j(x, \xi) = \sum_{\ell=1}^{d} \xi_\ell^2 \frac{\left(1 - g_{\ell,j}'(x_\ell)^2\right)}{\left(1 + g_{\ell,j}'(x_\ell)^2\right)^2} + c^{-2}(x).$$

$$(3.4)$$

On the one hand, the first equation in (3.4) implies that, for $x \in K$

$$\left| \operatorname{Im} p_s^j(x, \xi) \right| \geq C \sum_{\ell=1}^{d} \xi_\ell^2 g_{\ell,j}'(x_\ell),$$

from which Point 1 follows.

Since $\Im(1 - s^2 - 2is) \leq 0$ for $s \in [0, \infty]$ with equality only when $s = 0$, given $S > 0$ there exists $c_1 > 0$ and $\epsilon > 0$ such that

$$\Re\left(e^{i(\pi/2-\epsilon)}\left(1 - s^2 - 2is\right)\right) \geq c_1 \quad \text{for all } s \in [0, S].$$

Therefore, with $S := \max_{x \in K} \max_j g_{x_\ell,j}'(x_\ell)$,

$$\Re\left(e^{i(\pi/2-\epsilon)}\left(p_s^j(x, \xi) - c^{-2}(x)\right)\right) \geq c_1 \sum_{\ell=1}^{d} \xi_\ell^2 \frac{1}{\left(1 + g_{\ell,j}'(x_\ell)^2\right)^2}$$

from which Point 2 follows. $\qquad\square$

Recall from §B.2 that $T^*\mathbb{R}^d$ can be informally understood as $\{(x, \xi) : x \in \mathbb{R}^d, \xi \in \mathbb{R}^d\}$. Similarly, T^*D can be informally understood as $\{(x, \xi) : x \in D, \xi \in \mathbb{R}^d\}$.

Lemma 3.5 (Directions of ellipticity).

$$\{p_s^j = 0\} \cap T^*\widetilde{\mathcal{L}}_j \subset \left\{(x, \xi) : \xi \parallel \mathfrak{e} \text{ for all } \mathfrak{e} \in \mathcal{Z}^j \text{ s.t. } x \in \widetilde{\mathfrak{e}}\right\}$$

(i.e., at a point x in the extended PML layer of Ω_j, the symbol p_s^j vanishes only when ξ is parallel to every extended generalised edge of Ω_j that x belongs to a κ-neighbourhood of).

Proof. Let $(x, \xi) \in T^*\widetilde{\mathcal{L}}_j$ be such that

$$(x, \xi) \notin \left\{(x', \xi') : \xi' \parallel \mathfrak{e} \text{ for all } \mathfrak{e} \in \mathcal{Z}^j \text{ s.t. } x' \in \widetilde{\mathfrak{e}}\right\}.$$

Then, there exists an edge $\mathfrak{e} \in \mathcal{Z}^j$ so that $x \in \widetilde{\mathfrak{e}}$ and $\xi \nparallel \mathfrak{e}$. Since $x \in \widetilde{\mathfrak{e}}$, if m is such that the Cartesian-coordinates vector e_m is normal to $\mathfrak{e}$ then $g_{x_m,j}'(x_m) > 0$ (i.e., the PML is active in this coordinate direction). On the other hand, since $\xi \nparallel \mathfrak{e}$, we have $\xi_m \neq 0$ for at least one such m. Therefore, by Lemma 3.4 (1), $|p_s^j(x, \xi)| \geq C\xi_m^2 g_{x_m,j}'(x_m) > 0$. $\qquad\square$

4 A propagation result for the PML solutions and its consequences

The main purpose of this section is to prove the following propagation result (stated informally in §2.2, with the main ideas behind the proof described in §2.3).

Lemma 4.1. *Let $1 \leq j \leq N$, $f \in H^{-1}(\Omega_j)$, $g \in H^{1/2}(\partial\Omega_j)$, $\eta > 0$, $\widetilde{\Omega}_j$ the extension defined by Definition 3.2, and $\psi \in C_c^\infty(\widetilde{\Omega}_j)$. We assume that*

1. *any trajectory from $T^*(\operatorname{supp}\psi \cap \Omega_j) \cap \{p_s^j = 0\}$ goes to infinity under the backward flow.*

2. *$g = 0$ on $\partial\Omega_j \cap \partial\Omega$.*

Let $w \in H^1(\Omega_j)$ be the solution to

$$\begin{cases} P_s^j w = f & in \ \Omega_j, \\ w = g & on \ \partial\Omega_j, \end{cases}$$

and assume that f and w are tempered (in the sense of Definition B.5). Then, with $\widetilde{w} := S_j w$, $\widetilde{f} := S_j f$,

$$\mathrm{WF}_\hbar(\psi\widetilde{w}) \subset \bigcup_{\chi \in C_c^\infty(\widetilde{\Omega}_j)} \left\{ \rho \in T^*\big(\operatorname{supp}\psi\big) \ : \ \exists t \leq 0 \ s.t. \ \gamma_{[t,0]}(\rho) \subset T^*\widetilde{\Omega}_j \ and \ \Phi_t(\rho) \in \mathrm{WF}_\hbar(\chi\widetilde{f}) \right\}.$$

The interpretation of Lemma 4.1 is that, in the "measurement location" of $\operatorname{supp}\psi$, the wavefront set of the (extended) solution consists at most of points $\rho = (x, \xi)$ coming from the wavefront set of the (extended) source f under the forward flow; see Figure 4.2. The assumption in Point 1 is a non-trapping assumption.

Lemma 4.1 is proved in §4.3 below, with §4.1 and §4.2 containing intermediate results. §4.4 and §4.5 contain other consequences of the propagation results of this section, namely Lemma 4.5 (a priori bound on the PML solution operator) and Theorem 4.6 (error incurred by Cartesian PML approximation of outgoing Helmholtz solutions), respectively.

4.1 Forward propagation of regularity

Recall that Φ_t is the Hamilton flow for P defined by (1.15). Let Φ_t^θ denote the Hamilton flow for P_s, that is, defined by (1.15) with p replaced by $\operatorname{Re} p_s$. For any $1 \leq j \leq N$, let $\Phi_t^{\theta,j}$ denote the Hamilton flow for P_s^j, that is, defined by (1.15) with p replaced by $\operatorname{Re} p_s^j$. Finally, given an interval I, let

$$\gamma_I(x, \xi) := \left\{ \Phi_t(x, \xi), \ t \in I \right\}. \tag{4.1}$$

Lemma 4.2. *On $\{p_s^j = 0\}$, $\Phi_t^{\theta,j} = \Phi_t$ for any $t \in \mathbb{R}$; and on $\{p_s = 0\}$, $\Phi_t^\theta = \Phi_t$ for any $t \in \mathbb{R}$.*

Proof. We prove the first assertion; the proof of the second one is similar. By the expression for $\operatorname{Re} p_s^j$ in (3.4),

$$\partial_{\xi_\ell} \operatorname{Re} p_s^j = 2\xi_m \frac{\left(1 - g'_{\ell,j}(x_\ell)^2\right)}{\left(1 + g'_{x_\ell,j}(x_\ell)^2\right)^2}, \tag{4.2}$$

and

$$\partial_{x_\ell} \operatorname{Re} p_s^j = -\xi_\ell^2 \frac{2g''_{\ell,j}(x_\ell)g'_{\ell,j}(x_\ell)}{\left(1 + g'_{x_\ell,j}(x_\ell)^2\right)^2} - 4\xi_\ell^2 \frac{g''_{x_\ell,j}(x_\ell)g'_{x_\ell,j}(x_\ell)\left(1 - g'_{\ell,j}(x_\ell)^2\right)}{\left(1 + g'_{x_\ell,j}(x_\ell)^2\right)^3} + \partial_{x_\ell}(c^{-2}). \tag{4.3}$$

23

Let (x, ξ) be such that $p_{\theta,j}(x, \xi) = 0$. Since also $\operatorname{Im} p_{\theta,j}(x, \xi) = 0$, by (3.4),

$$\text{for all } \ell, \quad \text{either } \xi_\ell = 0 \text{ or } g'_{x_\ell,j}(x_\ell) = 0.$$

In both these two cases, (4.2) and (4.3) above simplify to

$$\partial_{\xi_\ell} \operatorname{Re} p_{\mathrm{s}}^j(x, \xi) = 2\xi_m, \quad \partial_{x_\ell} \operatorname{Re} p_{\mathrm{s}}^j(x, \xi) = \partial_{x_\ell}(c^{-2}),$$

and these are the equation defining the flow associated with p (by (1.15)). $\qquad \square$

The following lemma uses the notion of wavefront set defined in Definition B.6 below.

Lemma 4.3 (Forward propagation of regularity from an elliptic point). *Let $v_\hbar$ be a family of $\hbar$-tempered distributions in the sense of Definition B.5. Assume there exists $(x_0, \xi_0) \in T^*\mathbb{R}^d$, $t_0 \leq 0$, and $(s, M) \in [0, +\infty] \times [0, +\infty]$ such that*

$$\Phi_{t_0}(x_0, \xi_0) \in \left\{ (x, \xi) : |p_{\mathrm{s}}^j(x, \xi)| > 0 \right\} \quad \text{and} \quad \gamma_{[t_0,0]}(x_0, \xi_0) \cap \mathrm{WF}_\hbar^{s-1,M}(\hbar^{-1} P_{\mathrm{s}}^j v) = \emptyset. \quad (4.4)$$

Then $(x_0, \xi_0) \notin \mathrm{WF}_\hbar^{s,M}(v)$. The same property holds with P_{s}^j replaced with P_{s}.

Proof. We do the proof for P_{s}^j, the proof for P_{s} is similar. The objects involved in the proof are illustrated in Figure 4.1. We can assume that $p_{\mathrm{s}}^j(x_0, \xi_0) = 0$, otherwise the result follows directly from Lemma B.9.

The idea of the proof is the following: we seek to use forward propagation of regularity given by Lemma B.11. By the second assumption in (4.4), the trajectory under the flow Φ_t (illustrated in red in Figure 4.1) doesn't intersect $\mathrm{WF}_\hbar^{s-1,M}(\hbar^{-1} P_{\mathrm{s}}^j v)$ until at least t_0. However, to use Lemma B.11, we need that the trajectory under the flow $\Phi_t^{\mathrm{s},j}$ doesn't intersect $\mathrm{WF}_\hbar^{s-1,M}(\hbar^{-1} P_{\mathrm{s}}^j v)$, and this flow is not necessarily equal to Φ_t once $p_{\mathrm{s}}^j \neq 0$. A solution to this issue is the following. When travelling backwards from (x_0, ξ_0) under Φ_t (which starts off equal to $\Phi_t^{\mathrm{s},j}$ by Lemma 4.2 and the fact that $p_{\mathrm{s}}^j(x_0, \xi_0) = 0$), by continuity, the two flows must be close to each other in a short time interval after p_{s}^j stops being zero, and $p_{\mathrm{s}}^j > 0$ in this time interval. That is, we know that under the backward $\Phi_t^{\mathrm{s},j}$ flow, (x_0, ξ_0) reaches an elliptic point of p_{s}^j without hitting the data, and the result then follows from Lemma B.11.

In more detail: by Lemma 4.2,

$$\sup \left\{ t \leq 0, \ |p_{\mathrm{s}}^j(\Phi_t(x_0, \xi_0))| > 0 \right\} = \sup \left\{ t \leq 0, \ |p_{\mathrm{s}}^j(\Phi_t^{\mathrm{s},j}(x_0, \xi_0))| > 0 \right\} =: \bar{t}$$

(since the two flows are the same when $p_{\mathrm{s}}^j = 0$). By assumption (i.e., the existence of t_0), $\bar{t} > -\infty$. By Lemma 4.2,

$$\gamma_{[\bar{t},0]}(x_0, \xi_0) = \gamma_{[\bar{t},0]}^{\mathrm{s},j}(x_0, \xi_0), \quad (4.5)$$

(where $\gamma_{[\bar{t},0]}^{\mathrm{s},j}(x_0, \xi_0)$ is defined analogously to (4.1) with Φ_t replaced by $\Phi_t^{\mathrm{s},j}$). By the definition of $\bar{t}$, there exists $T < \bar{t}$ such that

$$|p_{\mathrm{s}}^j(\Phi_t^{\mathrm{s},j}(x_0, \xi_0))| > 0 \quad \text{for all } t \in (T, \bar{t}). \quad (4.6)$$

By (4.5) and the second assumption in (4.4), since $t_0 \leq \bar{t}$ and $\mathrm{WF}_\hbar^{s-1,M}(\hbar^{-1} P_{\mathrm{s}}^j v)$ is a closed set,

$$d\left(\gamma_{[\bar{t},0]}^{\mathrm{s},j}(x_0, \xi_0)), \mathrm{WF}_\hbar^{s-1,M}(\hbar^{-1} P_{\mathrm{s}}^j v) \right) > 0.$$

Hence, since $t \mapsto \Phi_t^{\mathrm{s},j}(x_0, \xi_0)$ is continuous, there exists $T < t_1 < \bar{t}$ close enough to $\bar{t}$ so that

$$\gamma_{[t_1,0]}^{\mathrm{s},j}(x_0, \xi_0) \cap \mathrm{WF}_\hbar^{s-1,M}(\hbar^{-1} P_{\mathrm{s}}^j v) = \emptyset. \quad (4.7)$$

In particular, (4.6) and (4.7) imply by Lemma B.9 that

$$\Phi_{t_1}^{\mathrm{s},j}(x_0, \xi_0) \notin \mathrm{WF}_\hbar^{s,M}(v). \quad (4.8)$$

Forward propagation of regularity given by Lemma B.11 (the sign condition on $\operatorname{Im} p_{\mathrm{s}}^j$ being satisfied thanks to the first equation in (3.4)) allows us to conclude, by (4.8) and (4.7), that $(x_0, \xi_0) \notin \mathrm{WF}_\hbar^{s,M}(v)$. $\qquad \square$
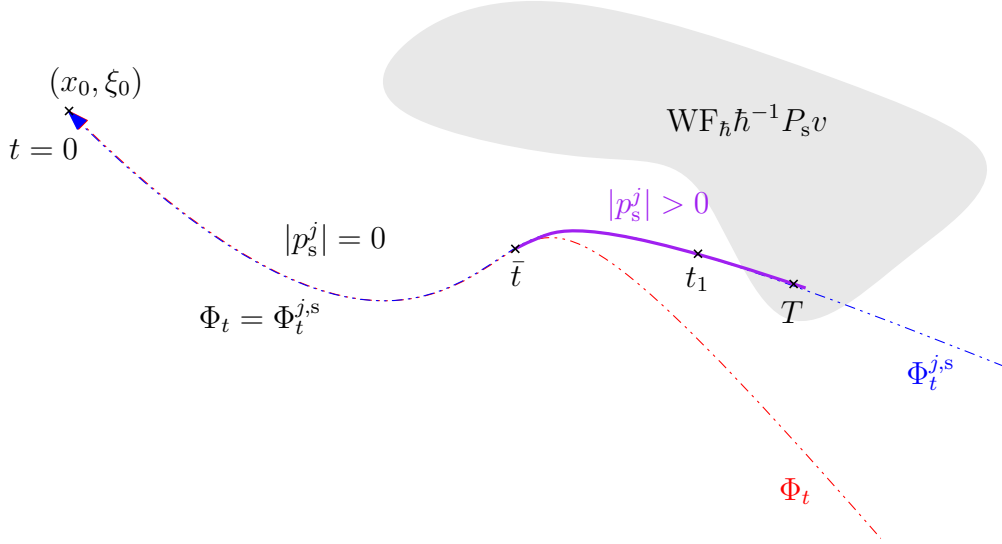
Figure 4.1: The objects involved in the proof of Lemma 4.3.

4.2 Escape to ellipticity

Lemma 4.4 (Escape to ellipticity). *For any $\rho \in T^*\widetilde{\Omega}_j$ such that the backward trajectory from ρ in the flow Φ_t goes to infinity (i.e. ρ is not trapped backward in time), there exists $\tau \leq 0$ such that*

$$\gamma_{[\tau,0]}(\rho) \subset T^*\widetilde{\Omega}_j, \quad and \quad \Phi_\tau(\rho) \in \big\{ (x,\xi) : |p_s^j(x,\xi)| > 0 \big\}.$$

Although Lemma 4.4 is stated (and then used below) with the flow Φ_t associated to P, we note that the result holds for a general continuous flow.

Proof of Lemma 4.4. By Lemma 3.5, it suffices to show the existence of $\tau \leq 0$ such that

$$\gamma_{[\tau,0]}(\rho) \subset T^*\widetilde{\Omega}_j \quad \text{and} \quad \Phi_\tau(\rho) \in T^*\widetilde{\mathcal{L}}_j \backslash \big\{ (x,\xi) : \xi \parallel \mathfrak{e} \text{ for all } \mathfrak{e} \in \mathcal{Z}^j \text{ s.t. } x \in \widetilde{\mathfrak{e}} \big\}.$$

For simplicity we give the proof in dimension 3; we then indicate how to generalise it to any dimension.

By the non-trapping assumption, there exists $t_0 \geq 0$ such that the backward trajectory from ρ enters $T^*\widetilde{\mathcal{L}}^j$ forever after time t_0; i.e., $\gamma_{(-\infty,t_0]}(\rho) \subset T^*\widetilde{\mathcal{L}}^j$. Reducing t_0 if necessary, we have $\gamma_{[t_0,0]}(\rho) \subset T^*\widetilde{\Omega}^j$.

Observe that the trajectory from ρ is in an extended PML-face for all times before time t_0, i.e. $\gamma_{(-\infty,t_0]}(\rho) \in T^*\widetilde{\mathfrak{e}_1}$ for some $\mathfrak{e}_1 \in \mathcal{X}_2^j$. If there exists $\tau \leq t_0$ such that $\pi_\xi \Phi_\tau(\rho) \nparallel \mathfrak{e}_1$ and $\gamma_{[\tau,t_0]}(\rho) \subset T^*\widetilde{\Omega}^j$, then we are done. Otherwise, the backward trajectory from $\Phi_{t_0}(\rho)$ stays parallel to the face $\mathfrak{e}_1$ it entered before leaving $\widetilde{\Omega}_j$, and therefore there exists $t_1 \leq t_0$ such that it enters an extended PML-edge forever; i.e., $\gamma_{(-\infty,t_1]}(\rho) \subset T^*\widetilde{\mathfrak{e}_2}$ for some $\mathfrak{e}_2 \in \mathcal{X}_1^j$. Reducing t_1 if necessary, we have again $\gamma_{[t_1,t_0]}(\rho) \subset T^*\widetilde{\Omega}^j$.

If there exists $\tau \leq t_1$ such that $\pi_\xi \Phi_\tau(\rho) \nparallel \mathfrak{e}_2$ and $\gamma_{[\tau,t_1]}(\rho) \subset T^*\widetilde{\Omega}^j$, then we are done. Otherwise, the backward trajectory from $\Phi_{t_1}(\rho)$ stays parallel to the edge $\mathfrak{e}_2$ it entered before leaving $\widetilde{\Omega}_j$, and therefore there is $\tau \leq t_1$ with $\gamma_{[\tau,t_1]}(\rho) \subset T^*\widetilde{\Omega}_j$ such that it enters a perpendicular PML-edge (because Ω_j is a hyperrectangle). We therefore have $\pi_x \Phi_\tau(\rho) \in \widetilde{\mathfrak{e}_3}$ and $\pi_\xi \Phi_\tau(\rho) \nparallel \mathfrak{e}_3$ for some $\mathfrak{e}_3 \in \mathcal{X}_1^j$ and we are done.

The proof generalises to any dimension $d \geq 2$ by observing that, if $\gamma_{[T,0)}(\rho) \subset T^*\widetilde{\Omega}_j$ and $\gamma_{(-\infty,T]}(\rho) \subset T^*\widetilde{\mathfrak{e}}$ for some $\mathfrak{e} \in \mathcal{X}_\ell^j$ with $2 \leq \ell \leq d-1$, then *either* there exists $\tau \leq T$ such that $\pi_\xi \Phi_\tau(\rho) \nparallel \mathfrak{e}$ and $\gamma_{(\tau,T]} \subset T^*\widetilde{\Omega}^j$, *or* the corresponding trajectory enters a $\ell - 1$ dimensional

25

edge forever: i.e., there exists $T' \leq T$ such that $\gamma_{[T',T]}(\rho) \subset T^*\widetilde{\Omega}_j$ and $\gamma_{(-\infty,T']}(\rho) \subset T^*\widetilde{\mathfrak{e}'}$ with $\mathfrak{e}' \in \mathcal{X}_{\ell-1}^j$. The initialisation of this process (entering $T^*\widetilde{\mathcal{L}}^j$) and the end-case ($\ell = 1$) are the same as in dimension 3. $\qquad\square$
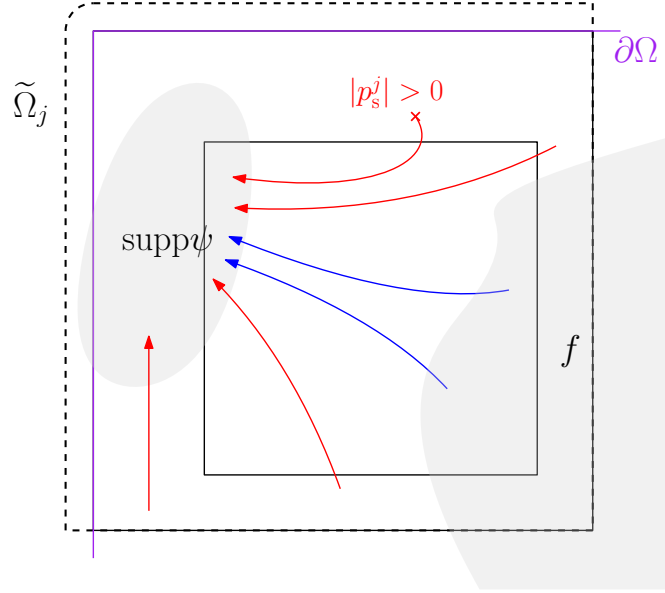
4.3   Proof of Lemma 4.1

Figure 4.2: Idea behind the proof of the propagation Lemma 4.1 illustrated for a subdomain Ω_j belonging to a corner of Ω, extended as $\widetilde{\Omega}_j$ (dashed). Any point of $T^* \operatorname{supp} \psi$ is the end-point of a trajectory coming from an elliptic point in the PML layer. If such a trajectory doesn't intersect $\mathrm{WF}_\hbar f$, it carries no mass by propagation of singularities (red trajectories). Therefore, the solution in the observation region, ψv, has only wavefront corresponding to trajectories coming from $\mathrm{WF}_\hbar f$ (blue trajectories).

By the assumption in Point 2 and Lemma 3.3,

$$P_{\mathrm{s}}^j \widetilde{w} = \widetilde{f} \quad \text{in } \widetilde{\Omega}_j.$$

Let $\rho \in T^* \operatorname{supp} \widetilde{\chi}$ be such that

$$\rho \notin \bigcup_{\chi \in C_c^\infty(\widetilde{\Omega}_j)} \left\{ \exists t < 0, \ \gamma_{[t,0]}(\rho) \subset T^*\widetilde{\Omega}_j \text{ and } \Phi_t(\rho) \in \mathrm{WF}_\hbar(\chi \widetilde{f}) \right\};$$

in other words, the backwards trajectory for the Hamilton flow of P starting at ρ *either* doesn't intersect $\mathrm{WF}_\hbar(\chi \widetilde{f})$ for any $\chi \in C_c^\infty(\widetilde{\Omega}_j)$ *or* leaves $\widetilde{\Omega}_j$ before intersecting such a subset. By Lemma 4.4, there exists $\tau \leq 0$ such that

$$\gamma_{[\tau,0]}(\rho) \subset T^*\widetilde{\Omega}_j, \quad \text{and} \quad \Phi_\tau(\rho) \in \left\{ (x,\xi) \,:\, |p_{\mathrm{s}}^j(x,\xi)| > 0 \right\} \tag{4.9}$$

(where we take $\tau = 0$ if $\rho \in \{p_{\mathrm{s}}^j = 0\}$). From the first part of (4.9), by assumption on ρ, we get that for any $\chi \in C_c^\infty(\widetilde{\Omega}_j)$

$$\gamma_{[\tau,0]}(\rho) \cap \mathrm{WF}_\hbar(\chi \widetilde{f}) = \emptyset. \tag{4.10}$$

We conclude from (4.9) and (4.10), by forward propagation of regularity from an elliptic point, given by Lemma 4.3, that $\rho \notin \mathrm{WF}_\hbar(\psi \widetilde{w})$.

4.4 A priori bound on the local PML solution on the subdomains

Lemma 4.5 (Bound on PML solution operator). *Assume that the flow associated with P is nontrapping. There exists $C > 0$ and $\hbar_0 > 0$ such that the following is true. Given $f \in H^{-1}(\Omega_j)$ and $g \in H^{1/2}(\partial\Omega_j)$, if u satisfies*

$$\begin{cases} P_s^j u = f \ in \ \Omega_j, \\ u = g \ on \ \partial\Omega_j, \end{cases} \tag{4.11}$$

then if $0 < \hbar < \hbar_0$ then

$$\|u\|_{H_\hbar^1(\Omega_j)} \leq C\Big(\hbar^{-1}\|f\|_{H_\hbar^{-1}(\Omega_j)} + \hbar^{-3/2}\|g\|_{H_\hbar^{1/2}(\partial\Omega_j)}\Big). \tag{4.12}$$

The same holds with Ω_j replaced with Ω.

We emphasise that the $\hbar^{-3/2}$ in front of the norm of g on the right-hand side of (4.12) is not sharp. However, the only way in which (4.12) is used in the proofs of Theorems 1.1-1.6 is to show that the solution to (4.11) is tempered (in the sense of Definition B.5) if f and g are tempered, and so this non-sharpness is not important for the results of the paper.

Proof. We first claim that it it sufficient to prove that there exists $C, \hbar_0 > 0$, such that if $v \in H_0^1(\Omega)$ satisfies $P_s^j v = f \in H^{-1}(\Omega)$ then, if $0 < \hbar < \hbar_0$,

$$\|v\|_{H_\hbar^1(\Omega_j)} \leq C\hbar^{-1}\|f\|_{H_\hbar^{-1}(\Omega_j)}. \tag{4.13}$$

Indeed, by [43, Theorem 5.6.4], for any Lipschitz domain D there exists $E_D : H^{1/2}(\partial D) \to H^1(D)$ such that $\gamma E_D \phi = \phi$ for all $\phi \in H^{1/2}(\partial D)$, where $\gamma : H^1(D) \to H^{1/2}(\partial D)$ is the trace operator. Furthermore,

$$\|E_D\|_{H_\hbar^{1/2}(\partial D) \to H_\hbar^1(D)} \leq C\hbar^{-1/2}, \tag{4.14}$$

noting that the weighted $H^{1/2}$ norm on ∂D used in [43] is equivalent to $h^{-1/2}\|\cdot\|_{H_\hbar^{1/2}(\partial D)}$ and the weighted H^1 norm on D used in [43] is equivalent to $h^{-1}\|\cdot\|_{H_\hbar^1(D)}$. Therefore, with u the solution to (4.11), let $v := u - E_D g \in H_0^1(\Omega_j)$. Then, by (4.13),

$$\begin{aligned} \|u\|_{H_\hbar^1(\Omega_j)} &\leq \|v\|_{H_\hbar^1(\Omega_j)} + \|E_D g\|_{H_\hbar^1(\Omega_j)} \\ &\leq C\hbar^{-1}\Big(\|f\|_{H_\hbar^{-1}(\Omega_j)} + \big\|P_s^j E_D g\big\|_{H_\hbar^{-1}(\Omega_j)}\Big) + \|E_D g\|_{H_\hbar^1(\Omega_j)} \\ &\leq C\hbar^{-1}\Big(\|f\|_{H_\hbar^{-1}(\Omega_j)} + C\|E_D g\|_{H_\hbar^1(\Omega_j)}\Big) + \|E_D g\|_{H_\hbar^1(\Omega_j)} \end{aligned}$$

and the bound (4.12) then follows from (4.14).

We now prove the bound (4.13). Seeking a contradiction, we assume that (4.13) does not hold. Then, there exists $\hbar_n \to 0$, $v_n \in H_0^1(\Omega_j)$, $f_n \in L^2(\Omega_j)$, such that $P_s^j(\hbar_n)v_n = f_n$ in Ω_j, and

$$\|v_n\|_{H_{\hbar_n}^1(\Omega_j)} \geq n\hbar_n^{-1}\|f_n\|_{(H_{\hbar_n}^1(\Omega_j))^*}.$$

Renormalising, we can assume that

$$\|v_n\|_{H_{\hbar_n}^1(\Omega_j)} = 1. \tag{4.15}$$

Then

$$\hbar_n^{-1}\|f_n\|_{(H_{\hbar_n}^1(\Omega_j))^*} \to 0. \tag{4.16}$$

We now extend v_n and f_n to $\widetilde{\Omega}_j$ using the extension defined by Lemma 3.2; i.e., we let $\widetilde{v}_n := S_j v_n$ and $\widetilde{f}_n := S_j f_n$. By Lemma 3.3, $P_s^j \widetilde{v}_n = \widetilde{f}_n$. By (4.15) $\widetilde{v}_n$ is tempered, and so propagation of

singularities applies. For any $\chi \in C_c^\infty(\widetilde{\Omega}_j)$, by (3.3) and the fact that $S_j : H^{-1}(\Omega_j) \to H^{-1}(\widetilde{\Omega}_j)$ (by Lemma 3.3),

$$\left\| \chi P_s^j(\hbar_n) \widetilde{v}_n \right\|_{H_\hbar^{-1}(\widetilde{\Omega}_j)} = \left\| \chi \widetilde{f}_n \right\|_{H_\hbar^{-1}(\widetilde{\Omega}_j)} \leq C \left\| f_n \right\|_{H_\hbar^{-1}(\Omega_j)}.$$

Therefore, by (4.16) and the definition of $\mathrm{WF}^{-1,1}$ (Definition B.6),

$$\mathrm{WF}^{-1,1}(\chi P_s^j(\hbar_n) \widetilde{v}_n) = \emptyset, \quad \text{for all } \chi \in C_c^\infty(\widetilde{\Omega}_j). \tag{4.17}$$

By Lemma 4.4 (escape to ellipticity), Lemma 4.3 (forward propagation of regularity from an elliptic point), and (4.17),

$$\mathrm{WF}^{0,1}(\chi \widetilde{v}_n) = \emptyset \quad \text{for all } \chi \in C_c^\infty(\widetilde{\Omega}_j).$$

Therefore, by Lemma B.8, Point 1 together with Point 2 of Lemma 3.4 (ellipticity at infinity) and (4.16),

$$\| \chi \widetilde{v}_n \|_{L^2} \to 0 \quad \text{for all } \chi \in C_c^\infty(\widetilde{\Omega}_j).$$

Lemma B.8, Point 2 together with Point 2 of Lemma 3.4 and (4.16) allow us to strengthen this to

$$\| \chi \widetilde{v}_n \|_{H_{\hbar_n}^1} \to 0 \quad \text{for all } \chi \in C_c^\infty(\widetilde{\Omega}_j)$$

which contradicts (4.15). $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad$ $\square$

4.5 Error incurred by Cartesian PML approximation of outgoing Helmholtz solutions

Theorem 4.6. (The error in Cartesian PML approximation of outgoing Helmholtz solutions $O(k^{-\infty})$.) *Suppose either (i) P is nontrapping, or (ii) with P_s considered as an operator either $H_0^1(\Omega) \to H^{-1}(\Omega)$ or $H^1(\mathbb{R}^d) \to H^{-1}(\mathbb{R}^d)$, its solution operator is bounded polynomially in $\hbar^{-1}$. Then for any $k_0, M, s > 0$, there exists $C > 0$ such that the following holds for all $k \geq k_0$. Let $f \in H^{-1}(\Omega_{\mathrm{int}})$ and let u and v be the solutions to*

$$\begin{cases} P_s u = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases} \qquad\qquad \begin{cases} P v = f \text{ in } \mathbb{R}^d, \\ v \text{ satisfies the Sommerfeld radiation condition.} \end{cases}$$

Then,

$$\| u - v \|_{H_\hbar^s(\Omega_{\mathrm{int}})} \leq C k^{-M} \| f \|_{H_\hbar^{-1}(\Omega_{\mathrm{int}})} \quad \text{for all } k \geq k_0.$$

Theorem 4.6 is not used in the proofs of Theorems 1.1-1.6, but we include it in the paper because of its independent interest. Indeed, as noted in §1.2, the only other k-explicit Cartesian-PML convergence result in the literature is that of [10, Lemma 3.4] for the case $c \equiv 1$ (i.e., no scatterer), with this proof based on the explicit expression for the Helmholtz fundamental solution in this case. k-explicit convergence results for radial PMLs applied to general Helmholtz problems are in [20].

Proof. By dividing u and v by $\| f \|_{H_\hbar^{-1}(\Omega_{\mathrm{int}})}$, we can assume, without loss of generality, that $\| f \|_{H_\hbar^{-1}(\Omega_{\mathrm{int}})} = 1$. It is then sufficient to prove that

$$\| u - v \|_{H_\hbar^s(\Omega_{\mathrm{int}})} \leq C k^{-M} \quad \text{for all } k \geq k_0.$$

Let u_∞ be the solution to

$$P_s u_\infty = f \text{ in } \mathbb{R}^d.$$

By [14, Theorem 4.37],

$$u_\infty |_{\Omega_{\mathrm{int}}} = v; \tag{4.18}$$

28

indeed, the Cartesian PML fits into the framework of [14, §4.5.1] by setting

$$F_\theta(x) = \sum_{\ell=1}^{d} \int^{x_\ell} g_\ell(\widetilde{x}_\ell)\, d\widetilde{x}_\ell,$$

and observing that $(\nabla F(x))_m = g_{x_m}(x_m)$ and $D^2 F \geq 0$ in the sense of quadratic forms (i.e., F_θ is convex) since $g'_{x_m} \geq 0$. The relation (4.18) implies that it is sufficient to prove that

$$\text{for all } s > 0, \quad \|u - u_\infty\|_{H_\hbar^s(\Omega_{\text{int}})} = O(\hbar^\infty). \tag{4.19}$$

We claim that

$$\text{for any } s > 0, \quad \|u_\infty\|_{H_\hbar^s(\partial\Omega)} = O(\hbar^\infty) \tag{4.20}$$

and postpone the proof for now.

We now let $u_\epsilon \in H^1(\Omega)$ be the solution to

$$\begin{cases} P_s u_\epsilon = 0 \text{ in } \Omega, \\ u_\epsilon = u_\infty \text{ on } \partial\Omega. \end{cases}$$

By (4.20) and *either* Lemma 4.5 under Assumption (i), *or* polynomial boundedness of the solution operator on $H^1(\mathbb{R}^d)$ under Assumption (ii),

$$\|u_\epsilon\|_{H_\hbar^1(\Omega)} = O(\hbar^\infty). \tag{4.21}$$

By Lemma B.8, Part 2, the bound (4.21) can be improved to

$$\text{for all } s > 0 \text{ and } \chi \in C_c^\infty(\Omega), \quad \|\chi u_\epsilon\|_{H_\hbar^s(\Omega)} = O(\hbar^\infty). \tag{4.22}$$

The whole point of the definition of u_ϵ is that, by uniqueness of u,

$$u = u_\infty + u_\epsilon.$$

Thus (4.19), and hence also the result, follows from (4.22).

We now prove (4.20). Under Assumption (ii), u_∞ is immediately tempered in the sense of Definition B.5. Under Assumption (i), repeating the proof of Lemma 4.5 (without the extensions), we see that χu_∞ is polynomially bounded in $\hbar^{-1}$ in terms of the data f for any $\chi \in C_c^\infty(\mathbb{R}^d)$, and hence u_∞ is tempered.

Let $\psi \in C_c^\infty(\mathbb{R}^d)$ be such that $\psi \equiv 1$ near $\partial\Omega$ and $\operatorname{supp}\psi \subset \mathbb{R}^d\backslash\Omega_{\text{int}}$. We now seek to apply Lemma 4.1. By Lemma 3.5, any $\rho \in T^*(\operatorname{supp}\psi) \cap \{p_s = 0\}$ is perpendicular to the scaling direction. Therefore, since $c \equiv 1$ in the PML region, any such ρ goes to infinity backwards in time under the flow associated to P. Therefore Point 1 in Lemma 4.1 is satisfied, and the result of this lemma is that

$$\mathrm{WF}_\hbar(\psi u_\infty) \subset \Big\{\rho \in T^* \operatorname{supp}\psi, \quad \exists t < 0, \ \Phi_t(\rho) \in \mathrm{WF}_\hbar(f)\Big\}.$$

By similar arguments, Lemma 3.5 implies that

$$\Big\{\rho \in T^* \operatorname{supp}\psi, \quad \exists t < 0, \ \Phi_t(\rho) \in \mathrm{WF}_\hbar(f)\Big\} \subset \Big\{|p_s| > 0\Big\}.$$

By Lemma B.9,

$$\mathrm{WF}_\hbar(\psi u_\infty) = \emptyset.$$

By Lemma B.8, Part 1, and Lemma 3.4, Part 2, for any $s > 0$,

$$\|\psi u_\infty\|_{H_\hbar^s} = O(\hbar^\infty)$$

(where Lemma B.8 is applied with $\operatorname{supp}\widetilde{\chi} \subset \mathbb{R}^d \setminus \Omega_{\text{int}}$, so that $\widetilde{\chi} P_s u_\infty = \widetilde{\chi} f = 0$). The claim (4.20) then follows from the trace theorem and the fact that $\psi \equiv 1$ near $\partial\Omega$, and the proof is complete. □

5 The propagation of physical error

5.1 The propagation of physical error for the parallel method

Proof of Lemma 2.1. By the definition of $\mathcal{T}_i$ (2.4), $\mathcal{T}_i v = 0$ on $\partial\Omega_i \cap \partial\Omega$ for any $v \in H_0^1(\Omega)$. Thus, since $\mathrm{supp}\chi_i \cap \Omega \subset \Omega_i$ (by (1.6)), $\chi_i \mathcal{T}_i v = 0$ on $\partial\Omega_i \setminus \partial\Omega$ and so $\mathbf{T}$ defined by (2.5) maps $(H_0^1(\Omega))^N \mapsto (H_0^1(\Omega))^N$. Now, in Ω_j, by the definition of e_j^{n+1} (2.1) and the iteration (1.7),

$$
\begin{aligned}
P_{\mathrm{s}}^j e_{+,j}^{n+1} = P_{\mathrm{s}}^j \big(u|_{\Omega_j} - u_j^{n+1}\big) &= P_{\mathrm{s}}^j(u|_{\Omega_j}) - P_{\mathrm{s}}^j(u_+^n|_{\Omega_j}) + (P_{\mathrm{s}}u_+^n)|_{\Omega_j} - f|_{\Omega_j} \\
&= P_{\mathrm{s}}^j\big((u - u_+^n)|_{\Omega_j}\big) - \big(P_{\mathrm{s}}(u - u_+^n)\big)|_{\Omega_j} \\
&= P_{\mathrm{s}}^j\big(e_+^n|_{\Omega_j}\big) - \big(P_{\mathrm{s}}e_+^n\big)|_{\Omega_j}.
\end{aligned}
$$

Therefore, by (2.2),

$$
P_{\mathrm{s}}^j e_{+,j}^{n+1} = \sum_{\ell=1}^N \Big(P_{\mathrm{s}}^j\big((\chi_\ell e_{+,\ell}^n)|_{\Omega_j}\big) - \big(P_{\mathrm{s}}(\chi_\ell e_{+,\ell}^n)\big)|_{\Omega_j} \Big).
$$

Furthermore, by (1.7), $u_j^{n+1} = u_+^n$ on $\partial\Omega_j$, and thus, on $\partial\Omega_j$,

$$
e_{+,j}^{n+1} = u - u_j^{n+1} = u - u_+^n = e_+^n = \sum_{\ell=1}^N \chi_\ell e_{+,\ell}^n.
$$

Therefore, by the definition of $\mathcal{T}_j$ (2.4), and the fact (noted after (2.3)) that $\chi_\ell e_{+,\ell}^n \in H_0^1(\Omega)$,

$$
e_{+,j}^{n+1} = \sum_{\ell=1}^N \mathcal{T}_j \chi_\ell e_{+,\ell}^n.
$$

Therefore, since $\chi_\ell^> \chi_\ell = \chi_\ell$,

$$
\chi_j e_{+,j}^{n+1} = \sum_{\ell=1}^N \Big(\chi_j \mathcal{T}_j \chi_\ell^> \Big) \chi_\ell e_{+,\ell}^n
$$

and the result follows by definitions of $\epsilon_{+,j}^n$ (2.3) and $\mathbf{T}$ (2.5). $\qquad\square$

Remark 5.1 (The rationale for putting $\chi_j^>$ in the definition of T). *The end of the proof of Lemma 2.1 shows that if $\chi_j^>$ is omitted from the definition of $\mathbf{T}$, then the result (2.6) still holds. The advantage of including $\chi_j^>$ is that this builds into the operator the localisation property of the physical error. Indeed, we do not expect the power contractivity proved for $\mathbf{T}$ (which is a result about $\mathbf{T}$ applied to arbitrary functions in $(H_0^1(\Omega))^N$) to hold if $\chi_j^>$ is omitted from the definition of $\mathbf{T}$.*

5.2 The propagation of physical error for the sequential method

Proof of Lemma 2.7. We prove the first equation in (2.16); the proof of the second equation is very similar. To keep the notation concise, in the proof we omit the restrictions onto Ω_j so that, in particular, the PDE defining u_j^{2n+1} in (1.10) becomes

$$
P_{\mathrm{s}}^j u_j^{2n+1} = (P_{\mathrm{s}}^j - P_{\mathrm{s}})(\vec{u}_{j,n}) + f \quad \text{in } \Omega_j.
$$

By the definition of $e_{\times,j}^{2n+1}$ (2.13) and the definition of u_j^{2n+1} (1.10), in Ω_j,

$$P_{\mathrm{s}}^j e_{\times,j}^{2n+1} = P_{\mathrm{s}}^j u - P_{\mathrm{s}}^j u_j^{2n+1} = (P_{\mathrm{s}}^j - P_{\mathrm{s}})u + f - (P_{\mathrm{s}}^j - P_{\mathrm{s}})(u_{j,n}^{\rightarrow}) - f$$

$$= (P_{\mathrm{s}}^j - P_{\mathrm{s}})\Big(u - \sum_{\ell < j}\chi_\ell u_\ell^{2n+1} - \sum_{j \leq \ell}\chi_\ell u_\ell^{2n}\Big)$$

$$= (P_{\mathrm{s}}^j - P_{\mathrm{s}})\Big(\sum_{\ell=1}^{N}\chi_\ell u - \sum_{\ell < j}\chi_\ell u_\ell^{2n+1} - \sum_{j \leq \ell}\chi_\ell u_\ell^{2n}\Big)$$

$$= (P_{\mathrm{s}}^j - P_{\mathrm{s}})\Big(\sum_{\ell < j}\chi_\ell(u - u_\ell^{2n+1}) + \sum_{j \leq \ell}\chi_\ell(u - u_\ell^{2n})\Big)$$

$$= (P_{\mathrm{s}}^j - P_{\mathrm{s}})\Big(\sum_{\ell < j}\chi_\ell e_{\times,\ell}^{2n+1} + \sum_{j \leq \ell}\chi_\ell e_{\times,\ell}^{2n}\Big).$$

Similarly, on $\partial\Omega_j$,

$$e_{\times,j}^{2n+1} = u - u_j^{2n+1} = u - \sum_{\ell < j}\chi_\ell u_\ell^{2n+1} - \sum_{j \leq \ell}\chi_\ell u_\ell^{2n} = \sum_{\ell < j}\chi_\ell e_{\times,\ell}^{2n+1} + \sum_{j \leq \ell}\chi_\ell e_{\times,\ell}^{2n},$$

where, in the last equality, we used the same manipulation as in the previous displayed equation. Therefore, by definition of $\mathcal{T}_j$ (2.4) and the definition of $\chi_j^{>}$ in §2.1,

$$e_{\times,j}^{2n+1} = \sum_{\ell < j}\mathcal{T}_j\chi_\ell e_{\times,\ell}^{2n+1} + \sum_{j \leq \ell}\mathcal{T}_j\chi_\ell e_{\times,\ell}^{2n} = \sum_{\ell < j}\mathcal{T}_j\widetilde{\chi}_\ell\chi_\ell e_{\times,\ell}^{2n+1} + \sum_{j \leq \ell}\mathcal{T}_j\widetilde{\chi}_\ell\chi_\ell e_{\times,\ell}^{2n}.$$

Therefore

$$(\boldsymbol{\epsilon}_{\times}^{2n+1})_j = \chi_j e_{\times,j}^{2n+1} = \sum_{\ell < j}(\chi_j\mathcal{T}_j\chi_\ell^{>})\chi_\ell e_{\times,\ell}^{2n+1} + \sum_{j \leq \ell}(\chi_j\mathcal{T}_j\chi_\ell^{>})\chi_\ell e_{\times,\ell}^{2n}$$

$$= \sum_{\ell < j}(\chi_j\mathcal{T}_j\chi_\ell^{>})\epsilon_{\times,\ell}^{2n+1} + \sum_{j \leq \ell}(\chi_j\mathcal{T}_j\chi_\ell^{>})\epsilon_{\times,\ell}^{2n}.$$

The result (i.e., the first equation in (2.16)) follows by the definition of $\mathbf{T}$ (2.5) and the fact that $\mathbf{L}$ and $\mathbf{U}$ are the lower/upper triangular parts of $\mathbf{T}$. □

The following lemma is proved in an analogous way to Lemma 2.7.

Lemma 5.2. *For the general sequential method with ordering* $\{\sigma_n\}$, *for* $n \geq 0$,

$$\boldsymbol{\epsilon}_{\times}^{n+1} = \mathbf{A}_n\boldsymbol{\epsilon}_{\times}^{n+1} + \mathbf{B}_n\boldsymbol{\epsilon}_{\times}^{n},$$

where

$$(\mathbf{A}_n)_{i,\ell} = \begin{cases}\mathbf{T}_{i,\sigma_n(\ell)} & \text{if } \ell < \sigma_n^{-1}(i), \\ 0 & \text{otherwise}\end{cases}, \quad (\mathbf{B}_n)_{i,\ell} = \begin{cases}\mathbf{T}_{i,\sigma_n(\ell)} & \text{if } \ell > \sigma_n^{-1}(i), \\ 0 & \text{otherwise}\end{cases}.$$

Comparing Lemmas 2.7 and 5.2, we see that, for the forward-backward method, both $\mathbf{A}_n$ and $\mathbf{B}_n$ alternate between equalling $\mathbf{L}$ and $\mathbf{U}$.

By their definitions, we see that $(\mathbf{A}_n)_{i,\ell}$ is non-zero if $\sigma_n(\ell)$ appears before i in the ordering of subdomains $\{\sigma_n(1), \dots, \sigma_n(N)\}$ and $(\mathbf{B}_n)_{i,\ell}$ is non-zero if $\sigma_n(\ell)$ appears after i in this ordering. Compare this to the fact that $(\mathbf{L}_n)_{i,\ell}$ is non-zero if $\ell < i$, i.e., ℓ appears before i in the ordering $\{1, 2, \dots, N\}$ or after i in the ordering $\{N, N-1, \dots, 1\}$, and $(\mathbf{U}_n)_{i,\ell}$ is non-zero if $\ell > i$, i.e., ℓ appears after i in the ordering $\{1, 2, \dots, N\}$ or before i in the ordering $\{N, N-1, \dots, 1\}$.

Proof of Lemma 5.2. Arguing exactly as in the proof of Lemma 2.7, we obtain that

$$(\boldsymbol{\epsilon}_\times^{n+1})_{\sigma(j)} = \sum_{\ell < j} (\chi_{\sigma_n(j)} \mathcal{T}_{\sigma_n(j)} \widetilde{\chi}_{\sigma_n(\ell)}) \epsilon_{\times,\sigma_n(\ell)}^{n+1} + \sum_{j \leq \ell} (\chi_{\sigma_n(j)} \mathcal{T}_{\sigma(j)} \widetilde{\chi}_{\sigma_n(\ell)}) \epsilon_{\times,\sigma_n(\ell)}^n.$$

Therefore, with $i = \sigma_n(j)$,

$$(\boldsymbol{\epsilon}_\times^{n+1})_i = \sum_{\ell < \sigma_n^{-1}(i)} (\chi_i \mathcal{T}_i \widetilde{\chi}_{\sigma_n(\ell)}) \epsilon_{\times,\sigma_n(\ell)}^{n+1} + \sum_{\sigma_n^{-1}(i) \leq \ell} (\chi_i \mathcal{T}_i \widetilde{\chi}_{\sigma_n(\ell)}) \epsilon_{\times,\sigma_n(\ell)}^n$$

and the result follows by the definition of **T** (2.5). $\qquad\qquad\square$

6 Proof of Theorem 1.6

6.1 Precise definition of following a word

Recall from §2.2 the definitions of the set of words $\mathcal{W}$ and the composite map $\mathscr{T}_w$ (2.9). Recall the definition of the trajectory staring at (x, ξ) for a time interval I, $\gamma_I(x, \xi)$ (4.1).

We now define a set of cutoffs that are bigger than the $\chi_j^>$s. Let $\{\chi_j^\gg\}_{j=1}^N$ be such that, for each j, $\chi_j^\gg \in C^\infty(\mathbb{R}^d)$, $\operatorname{supp}\chi_j^\gg \cap \Omega \subset \Omega_j$, $\chi_j^\gg \equiv 1$ on $\operatorname{supp}\chi_j^>$, and $\chi_j^\gg$ is zero on $\operatorname{supp}(P_{\mathrm{s}}^j - P_{\mathrm{s}})$.

Definition 6.1. *A trajectory γ for P follows a word $w \in \mathcal{W}$ of length $n \geq 2$ if*

$$\gamma = \prod_{2 \leq j \leq n} \gamma_j,$$

where the product stands for concatenation, $\gamma_j = \gamma_{(0,T_j]}(x_{j-1}, \xi_{j-1})$ for some $T_j > 0$ and $(x_j, \xi_j) \in T^*\mathbb{R}^d$ with

- *for $1 \leq j \leq n-1$, $x_j \in \operatorname{supp}\chi_{w_j}^\gg \cap \operatorname{supp}\left(P_{\mathrm{s}}^{w_{j+1}} - P_{\mathrm{s}}\right)$,*

- *$\gamma_j \subset T^*\widetilde{\Omega}_{w_j}$ for $2 \leq j \leq n$ (recall Definition 3.2 of the extended domains),*

- *$(x_n, \xi_n) := \Phi_{T_n}(x_{n-1}, \xi_{n-1}) \in T^*(\operatorname{supp}\chi_{w_n}^>).$*

Figure 2.2 illustrates a simple example of a trajectory following a word.

6.2 Reduction to a propagation result for $\mathcal{T}_w$ (Lemma 6.2)

Recall that in §2 we showed how Theorem 1.6 follows from Lemma 2.4; i.e., the result that

$$\text{If } w \in \mathcal{W} \text{ is not allowed then } \|\mathscr{T}_w\|_{H_\hbar^1(\Omega) \to H_\hbar^1(\Omega)} = O(\hbar^\infty).$$

The heart of the proof of Lemma 2.4 is the following result, described informally in §2.2, and proved in the next subsection using the propagation result of Lemma 4.1.

Lemma 6.2. *Suppose that c is nontrapping. Let $w \in \mathcal{W}$ of length $n := |w| \geq 1$. For any $\chi_{\mathrm{meas}} \in C_c^\infty(\widetilde{\Omega}_{w_n})$ and any tempered $\hbar$-family v of elements of $H_0^1(\Omega)$, $\chi_{\mathrm{meas}}\mathscr{T}_w v$ is tempered and*

$$\mathrm{WF}_\hbar(\chi_{\mathrm{meas}}\widetilde{\mathscr{T}_w v}) \subset \left\{\rho \in T^* \operatorname{supp}\chi_{\mathrm{meas}} : \exists t_0 \leq 0 \text{ s.t. } \gamma_{(t_0,0]}(\rho) \text{ follows } w\right\}, \qquad (6.1)$$

where $\widetilde{\mathscr{T}_w v} := S_{w_n}\mathscr{T}_w v$.

Proof of Lemma 2.4 using Lemma 6.2. The fact that $\chi_{\mathrm{meas}}\mathscr{T}_w v$ is tempered is a consequence of the resolvent estimate given by Lemma 4.5. We fix a word $w \in \mathcal{W}$ not allowed. Seeking a contradiction, we assume that (2.10) fails; that is, there exists $\hbar_n \to 0$, $v_n \in H_0^1(\Omega)$ with $\|v_n\|_{H_{\hbar_n}^1} = 1$ and $N_0 \in \mathbb{N}$ such that

$$\text{for all } n \geq 1, \quad \|\mathscr{T}_w(\hbar_n)v_n\|_{H_{\hbar_n}^s(\Omega)} \geq \hbar_n^{N_0}.$$

We show that

$$\|\mathscr{T}_w(\hbar_n)v_n\|_{H_{\hbar_n}^s(\Omega)} = O(\hbar_n^\infty),$$

leading to a contradiction. To lighten the notation, from now on we drop the subscripts n from v_n and $\hbar_n$. By the definition of a word being allowed, Lemma 6.2 implies that, for any $\chi \in C_c^\infty(\widetilde{\Omega}_{w_n})$,

$$\mathrm{WF}_\hbar(\chi \widetilde{\mathscr{T}_w v}) = \emptyset.$$

Therefore, by Lemma B.8, Part 1 together with Lemma 3.4, Part 2,

$$\text{for all } \chi \in C_c^\infty(\widetilde{\Omega}_{w_n}) \quad \|\chi\widetilde{\mathscr{T}_w v}\|_{H_\hbar^s(\widetilde{\Omega}_{w_n})} = O(\hbar_n^\infty). \tag{6.2}$$

Recall that $\mathscr{T}_w v$ has a χ_{w_n} at the front by the definitions of $\mathscr{T}_w$ (2.9) and $\mathbf{T}$ (2.5). We now choose χ in (6.2) to be one on $\mathrm{supp}\chi_{w_n} \cap \Omega_{w_n}$ and supported in $\widetilde{\Omega}_{w_n}$, and obtain that

$$\|\mathscr{T}_w v\|_{H_\hbar^s(\Omega)} = \|\mathscr{T}_w v\|_{H_\hbar^s(\Omega_{w_n})} = O(\hbar_n^\infty).$$

$\square$

Remark 6.3 (Why we don't extend interior subdomains). *Consider the end of the proof of Lemma 2.4 using Lemma 6.2. If Ω_{w_n} is an interior subdomain, i.e., $\partial\Omega_{w_n} \cap \partial\Omega = \emptyset$, then χ_{w_n} is compactly supported in Ω_{w_n} (indeed, its support does not touch the PML layer of Ω_{w_n}), and we can choose χ to be compactly supported in Ω_{w_n}; i.e., we do not need information about $\mathscr{T}_w v$ up to $\partial\Omega_{w_n}$. However, if Ω_{w_n} is not an interior subdomain, then χ_{w_n} can be supported up to the boundary of Ω_{w_n} (by the partition of unity property) and we therefore need information about $\mathscr{T}_w v$ up to $\partial\Omega_{w_n}$. This is why we extend Ω_{w_n} to $\widetilde{\Omega}_{w_n}$, choose χ to have support in $\widetilde{\Omega}_{w_n} \setminus \Omega_{w_n}$, and do the propagation in the interior of $\widetilde{\Omega}_{w_n}$ (to bypass the issue of propagating up to $\partial\Omega_{w_n}$, as discussed in Remark 2.5).*

6.3 Proof of Lemma 6.2 (ending the proof of Theorem 1.6)

It therefore remains to prove Lemma 6.2. The idea of the proof is to iteratively apply Lemma 4.1. We proceed by induction on the length n of w. We first assume that the result of the lemma holds for any word of length n and we show it for any word of length $n+1$. At the end of the proof we show that the result holds when $n = 2$.

If $|w| = n+1$, we decompose w as

$$w = (\bar{w}, w_{n+1}),$$

and write (by the definitions of $\mathscr{T}_w$ (2.9) and $\mathcal{T}_{w_{n+1}}$ (2.4))

$$\mathscr{T}_w v = \chi_{w_{n+1}} \mathcal{T}_{w_{n+1}} \chi_{w_n}^> \mathscr{T}_{\bar{w}} v.$$

We now apply Lemma 4.1. Let

$$W := \mathcal{T}_{w_{n+1}} \chi_{w_n}^> \mathscr{T}_{\bar{w}} v$$

so that $\mathscr{T}_w v = \chi_{w_{n+1}} W$. By (3.2),

$$\chi_{\mathrm{meas}}\chi_{w_{n+1}}\widetilde{W} = \chi_{\mathrm{meas}}(\widetilde{\chi_{w_{n+1}}W}) = \chi_{\mathrm{meas}}\widetilde{\mathscr{T}_w v}. \tag{6.3}$$

By the definition of $\mathcal{T}_{w_{n+1}}$ (2.4),

$$\begin{cases} P_{\mathrm{s}}^{w_{n+1}} W = (P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}})(\chi_{w_n}^> \mathscr{T}_{\bar{w}} v) & \text{in } \Omega_{w_{n+1}}, \\ W = \chi_{w_n}^> \mathscr{T}_{\bar{w}} v & \text{on } \partial\Omega_{w_{n+1}}, \end{cases}$$

where we omit the restriction operators on $\Omega_{w_{n+1}}$ to keep the notation concise (as in the proof of Lemma 2.7 above).

We therefore apply Lemma 4.1 with $j := w_{n+1}$, $f := (P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}})(\chi_{w_n}^> \mathscr{T}_{\bar{w}} v)$, $g := \chi_{w_n}^> \mathscr{T}_{\bar{w}} v$, and $\psi := \chi_{\mathrm{meas}} \chi_{w_{n+1}}$.

We need to check that the assumptions in Points 1-2 are satisfied. The assumption in Point 1 is satisfied because c is nontrapping. The assumption in Point 2 is satisfied since $\mathscr{T}_{\bar{w}} v \in H_0^1(\Omega)$ by Lemma 2.1.

The result of Lemma 4.1 is that

$$\mathrm{WF}_\hbar(\chi_{\mathrm{meas}} \chi_{w_{n+1}} \widetilde{W}) \subset$$
$$\bigcup_{\chi \in C_c^\infty(\widetilde{\Omega}_{w_{n+1}})} \left\{ \rho \in T^* \mathrm{supp}\, \chi_{\mathrm{meas}} : \quad \exists t < 0 \ \text{s.t.} \ \gamma_{[t,0]}(\rho) \subset T^* \widetilde{\Omega}_{w_{n+1}} \ \text{and} \ \Phi_t(\rho) \in \mathrm{WF}_\hbar(\chi \widetilde{f}) \right\},$$

where $\widetilde{f} = S_{w_{n+1}} f$, $\widetilde{W} = S_{w_{n+1}} W$. Recall that in the extended region the PML is linear and the coefficients of both P_{s}^j and P_{s} are constant. The arguments used in the proof of Lemma 3.3 therefore show that the extension operator S_j of Definition 3.2 is such that

$$S_j\Big((P_{\mathrm{s}}^{j+1} - P_{\mathrm{s}})(\chi_j w_j)\Big) = (P_{\mathrm{s}}^{j+1} - P_{\mathrm{s}}) S_j(\chi_j w_j) = (P_{\mathrm{s}}^{j+1} - P_{\mathrm{s}})(\chi_j \widetilde{w}_j). \tag{6.4}$$

where we have used (3.2) in the second step.

Therefore

$$\mathrm{WF}_\hbar(\chi_{\mathrm{meas}} \chi_{w_{n+1}} \widetilde{W}) \subset \bigcup_{\chi \in C_c^\infty(\widetilde{\Omega}_{w_{n+1}})} \left\{ \rho \in T^* \mathrm{supp}\, \chi_{\mathrm{meas}} : \exists t < 0 \ \text{s.t.} \ \gamma_{[t,0]}(\rho) \subset T^* \widetilde{\Omega}_{w_{n+1}} \right.$$
$$\left. \text{and} \ \Phi_t(\rho) \in \mathrm{WF}_\hbar(\chi(P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}})(\chi_{w_n}^> \widetilde{\mathbf{T}_{\bar{w}} v})) \right\}. \tag{6.5}$$

We ultimately use that, by (B.4) and (B.5),

$$\mathrm{WF}_\hbar \chi(P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}})(\chi_{w_n}^> \widetilde{\mathbf{T}_{\bar{w}} v}) \subset \mathrm{WF}_\hbar(\chi^> \chi_{w_n}^> \widetilde{\mathbf{T}_{\bar{w}} v}) \cap T^*(\mathrm{supp}\, \chi_{w_n}^>) \cap T^*(\mathrm{supp}(P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}})), \tag{6.6}$$

where $\chi^> \in C_c^\infty(\widetilde{\Omega}_{w_{n+1}})$ is such that $\chi^> \equiv 1$ on $\mathrm{supp}\chi$. However, we first change the cut-off $\chi \in C_c^\infty(\widetilde{\Omega}_{w_{n+1}})$ into a cut off in $C_c^\infty(\widetilde{\Omega}_{w_n})$ so that we can use the induction hypothesis, which gives information about $\mathrm{WF}_\hbar(\chi \chi_{w_n}^> \mathbf{T}_{\bar{w}} v)$ with $\chi \in C_c^\infty(\widetilde{\Omega}_{w_n})$.

Observe that

for all $\chi \in C_c^\infty(\widetilde{\Omega}_{w_{n+1}})$ there exists $\psi \in C_c^\infty(\widetilde{\Omega}_{w_n})$ s.t. $\chi(P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}}) \chi_{w_n}^> = \psi(P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}}) \chi_{w_n}^>$

(note that this crucially relies on the presence of $\chi_{w_n}^>$ on the right of both sides of the equality).

Using this in (6.5), we have

$$\mathrm{WF}_\hbar(\chi_{\mathrm{meas}} \chi_{w_{n+1}} \widetilde{W}) \subset \bigcup_{\psi \in C_c^\infty(\widetilde{\Omega}_{w_n})} \left\{ \rho \in T^* \mathrm{supp}\, \chi_{\mathrm{meas}} : \exists t < 0 \ \text{s.t.} \ \gamma_{[t,0]}(\rho) \subset T^* \widetilde{\Omega}_{w_{n+1}} \right.$$
$$\left. \text{and} \ \Phi_t(\rho) \in \mathrm{WF}_\hbar(\psi(P_{\mathrm{s}}^{w_{n+1}} - P_{\mathrm{s}})(\chi_{w_n}^> \widetilde{\mathbf{T}_{\bar{w}} v})) \right\}. \tag{6.7}$$

Therefore, by (6.6) with χ replaced by ψ,

$$
\begin{aligned}
\mathrm{WF}_\hbar(&\chi_{\mathrm{meas}}\chi_{w_{n+1}}\widetilde{W}) \\
&\subset \bigcup_{\psi\in C_c^\infty(\widetilde{\Omega}_{w_n})} \Big\{\rho\in T^*\operatorname{supp}\chi_{\mathrm{meas}} \;:\; \exists t<0 \;\text{ s.t. }\; \gamma_{[t,0]}(\rho)\subset T^*\widetilde{\Omega}_{w_{n+1}} \\
&\hspace{3cm} \text{and } \Phi_t(\rho)\in \mathrm{WF}_\hbar(\psi^>\chi_{w_n}^>\,\widetilde{\mathbf{T}_{\overline{w}}v})\cap T^*(\operatorname{supp}\chi_{w_n}^>)\cap T^*\big(\operatorname{supp}(P_{\mathrm{s}}^{w_{n+1}}-P_{\mathrm{s}})\big)\Big\}. \\
&= \bigcup_{\psi\in C_c^\infty(\widetilde{\Omega}_{w_n})} \Big\{\rho\in T^*\operatorname{supp}\chi_{\mathrm{meas}} \;:\; \exists t<0 \;\text{ s.t. }\; \gamma_{[t,0]}(\rho)\subset T^*\widetilde{\Omega}_{w_{n+1}} \\
&\hspace{3cm} \text{and } \Phi_t(\rho)\in \mathrm{WF}_\hbar(\psi\chi_{w_n}^>\,\widetilde{\mathbf{T}_{\overline{w}}v})\cap T^*(\operatorname{supp}\chi_{w_n}^>)\cap T^*\big(\operatorname{supp}(P_{\mathrm{s}}^{w_{n+1}}-P_{\mathrm{s}})\big)\Big\}.
\end{aligned}
$$

Therefore, by the induction hypothesis that (6.1) holds with w replaced by $\overline{w}$ and then the Definition 6.1 of following a word,

$$
\begin{aligned}
\mathrm{WF}_\hbar(\chi_{\mathrm{meas}}\chi_{w_{n+1}}\widetilde{W}) \subset \Big\{&\rho\in T^*\operatorname{supp}\chi_{\mathrm{meas}} \;:\; \exists t<0 \;\text{ s.t. }\; \gamma_{[t,0]}(\rho)\subset T^*\widetilde{\Omega}_{w_{n+1}} \text{ and} \\
&\Phi_t(\rho)\in T^*(\operatorname{supp}\chi_{w_n}^>)\cap T^*\big(\operatorname{supp}(P_{\mathrm{s}}^{w_{n+1}}-P_{\mathrm{s}})\big), \\
&\text{and } \exists t_0\le 0 \;\text{ s.t. }\; \gamma_{(t_0,0]}(\Phi_t(\rho)) \text{ follows } \overline{w}\Big\} \\
\subset \Big\{&\rho\in T^*\operatorname{supp}\chi_{\mathrm{meas}} \;:\; \exists t_1\le 0 \;\text{ s.t. }\; \gamma_{(t_1,0]}(\rho) \text{ follows } w\Big\},
\end{aligned}
$$

and the result for words of length $n+1$, assuming the result for words of length n, follows by (6.3). To complete the induction, we prove the result for words of length 2. If $w=(w_1,w_2)$, then the inclusion (6.7) simplifies to

$$
\begin{aligned}
\mathrm{WF}_\hbar(\chi_{\mathrm{meas}}\chi_{w_2}\widetilde{W}) \subset \bigcup_{\psi\in C_c^\infty(\widetilde{\Omega}_{w_1})} \Big\{&\rho\in T^*\operatorname{supp}\chi_{\mathrm{meas}} \;:\; \exists t<0 \;\text{ s.t. }\; \gamma_{[t,0]}(\rho)\subset T^*\widetilde{\Omega}_{w_2} \\
&\text{and } \Phi_t(\rho)\in \mathrm{WF}_\hbar(\psi(P_{\mathrm{s}}^{w_2}-P_{\mathrm{s}})\chi_{w_1}^>)\Big\}.
\end{aligned}
$$

Arguing as in (6.6) and using the definition of following a word (Definition 6.1), we obtain

$$
\mathrm{WF}_\hbar(\chi_{\mathrm{meas}}\chi_{w_2}\widetilde{W}) \subset \Big\{\rho\in T^*\operatorname{supp}\chi_{\mathrm{meas}} \;:\; \exists t<0 \;\text{ s.t. }\; \gamma_{(t,0]}(\rho) \text{ follows } w\Big\},
$$

and the proof is complete.

7 Proofs of Theorems 1.1, 1.2, 1.3, 1.4

7.1 Checkerboard coordinates

The proofs of Theorems 1.1, 1.2, 1.3, 1.4 use a set of checkerboard coordinates $\{c_\ell\}_{1\le\ell\le\mathfrak{d}}$.

For concreteness we use the natural coordinates coming from the checkerboard construction in §1.4.7. For a d-checkerboard, the coordinates (with origin $(0,\ldots,0)$) of the subdomain Ω_j are $(c_1(j),\ldots,c_d(j))$ where $c_\ell(j)=m_\ell$ if the U_j from which Ω_j is formed equals $\prod_{\ell=1}^d(y_{m_\ell-1}^\ell, y_{m_\ell}^\ell)$.

For a $\mathfrak{d}$-checkerboard, we omit the $(d-\mathfrak{d})$ directions that have $N_\ell=1$ to obtain coordinates $\{c_\ell\}_{1\le\ell\le\mathfrak{d}}$.

Examples 7.3 and 7.4 below use coordinates with respect to other checkerboard vertices; these are defined in an analogous way by first rotating the checkerboard to place that particular vertex at the origin and relabelling all the points y_m^ℓ, $\ell=1,\ldots,d$, $m=0,\ldots,N_\ell$.

7.2 Allowed words in a checkerboard when $c \equiv 1$

Lemma 7.1. *Assume that $\{\Omega_j\}_{j=1}^N$ is a checkerboard and $c = 1$. Then, if $w \in \mathcal{W}$ is allowed, for any $1 \leq \ell \leq \mathfrak{d}$, the map $j \mapsto c_\ell(w_j)$ is monotonic.*

Proof. The underlying idea of the proof is that a straight line through a non-overlapping checkerboard (i.e., the decomposition of Ω_{int} described in Point (i) of §1.4.7) necessarily has monotonic coordinate maps. The subtlety here is that the subdomains $\{\Omega_j\}_{j=1}^N$ overlap; however, the fact that an allowed word must pass through the (disjoint) subdomains in (2.7) essentially puts us back in the non-overlapping case.

Assume that $n := |w| \geq 2$, otherwise the statement is void. By the definition of allowed, there exists a trajectory following w. Since the coefficients are constant, this trajectory is a straight line, and by the definitions of allowed (Definition 2.2) and follow (Definition 6.1), the trajectory intersects

$$\left[\bigcup_{1 \leq j \leq n-1} T^*\widetilde{\Omega}_{w_{j+1}} \cap T^*\left(\operatorname{supp} \chi_{w_j}^> \cap \operatorname{supp}\left(P_\mathrm{s}^{w_{j+1}} - P_\mathrm{s}\right)\right)\right] \cup \left(T^* \operatorname{supp} \chi_{w_n}^> \cap T^*\widetilde{\Omega}_{w_n}\right) =: \bigcup_{1 \leq j \leq n} X_j$$

Since $\chi_\ell^>$ is supported in $T^*\widetilde{\Omega}_\ell$ and away from $\operatorname{supp}\left(P_\mathrm{s}^\ell - P_\mathrm{s}\right)$, this implies that $X_j \cap X_{j+1} = \emptyset$ for $1 \leq j \leq n-1$ and $X_j \subset T^*\widetilde{\Omega}_{w_j}$ for $1 \leq j \leq n$. It follows that there exists $(x_1, \ldots, x_n)$ on a line with $x_j \in \widetilde{\Omega}_{w_j} \backslash \widetilde{\Omega}_{w_{j+1}}$, $j = 1, \ldots, n-1$ and $x_n \in \widetilde{\Omega}_{w_n}$. Because $x_j \in \widetilde{\Omega}_j \backslash \widetilde{\Omega}_{j+1}$ and $\{\Omega_j\}_{j=1}^N$ is a checkerboard, it follows that for any ℓ such that $c_\ell(j) \neq c_\ell(j+1)$, $c_\ell(j) \leq c_\ell(j+1)$ if and only if $(x_j)_\ell \leq (x_{j+1})_\ell$ (where subscript ℓ denotes the ℓth coordinate of x in $\mathbb{R}^\mathfrak{d}$); note that for this to be true it is key that the x_j is not in the overlap of $\widetilde{\Omega}_j$ and $\widetilde{\Omega}_{j+1}$ (which we have ensured above). The result follows. $\square$

7.3 Proofs of Theorems 1.1 and 1.3 (i.e., the results about the parallel method with $c \equiv 1$)

By Lemma 7.1, if $\{\Omega_j\}$ is a $\mathfrak{d}$-checkerboard, then $\mathcal{N} = 1 + (N_1 - 1) + \cdots + (N_\mathfrak{d} - 1) = N_1 + \cdots + N_d - (d-1)$. Indeed, this is the length of the longest component-wise monotonic sequence with values in $\{1, \ldots, N_1\} \times \cdots \times \{1, \ldots, N_d\}$. Thus Theorem 1.3 is a consequence of Theorem 1.6. Theorem 1.1 is then a consequence of Theorem 1.3 with $\mathfrak{d} = 1$.

7.4 Exhaustive sequence of orderings

Given a sequence of orderings $\{\sigma_n\}_{1 \leq n \leq S}$, for every n we define the order relation $\preceq_n$ on $\{1, \cdots, N\}$ by

$$i \preceq_n j \iff \sigma_n^{-1}(i) \leq \sigma_n^{-1}(j);$$

i.e., $i \preceq_n j$ if at step n of the method, Ω_i comes before Ω_j.

Recall that $\{c_\ell\}_{1 \leq \ell \leq \mathfrak{d}}$ is a fixed set of checkerboard coordinates, i.e., $c_\ell : \{1, \ldots, N\} \mapsto \{1, \ldots, N_\ell\}$ is such that $c_\ell(j)$ is the ℓth coordinate of Ω_j on the checkerboard.

Definition 7.2 (Exhaustive ordering). *We say that an ordering $\{\sigma_n\}_{1 \leq n \leq S}$ is exhaustive if (i) for any $1 \leq m \leq S$, there is $1 \leq \overline{m} \leq S$ such that $\sigma_m(i) = \sigma_{\overline{m}}(N - i + 1)$, and (ii) for any sequence $\{j_i\}_{1 \leq i \leq n}$ of elements of $\{1, \cdots, N\}$ such that*

$$\text{for all } 1 \leq \ell \leq \mathfrak{d}, \quad i \mapsto c_\ell(j_i) \text{ is monotonic,} \tag{7.1}$$

there exists an $1 \leq s \leq n$ so that

$$j_1 \preceq_s \cdots \preceq_s j_n. \tag{7.2}$$

6	8	9
4	5	7
1	2	3

4	2	1
6	5	3
9	8	7

9	6	5
8	4	2
7	3	1

1	4	5
2	6	8
3	7	9

Figure 7.1: Example of an exhaustive ordering of a 3×3 checkerboard obtained by following the construction in Example 7.4; the key point is that, moving through the subdomains in order, their coordinates on the checkerboard with respect to the first subdomain are non-decreasing.

Recall the informal definition of exhaustive in §1.6 that (i) given any ordering in the sequence, the sequence also contains the reverse ordering (i.e., where the subdomains are visited in reverse order), and (ii) given any directed straight line (here a sequence of subdomains with monotonic coordinate maps (7.1)) the sequence of sweeps has to contain at least one sweep such that the subdomains in the straight line are visited in order, but not necessarily sequentially, in that sweep; i.e., (7.2).

Observe that, if $m, \overline{m}$ are the indices of reverse orderings, as in Definition 7.2, then $\preceq_m = \succeq_{\overline{m}}$.

In what follows, we assume that the Cartesian directions on the checkerboard are always ordered in the same way, i.e., if $(c_\ell^n)_{1 \le n \le \mathfrak{d}}$ and $(c_\ell^m)_{1 \le n \le \mathfrak{d}}$ are two set of checkerboard coordinates, then c_ℓ^n and c_ℓ^m represent coordinates in the same ℓth direction.

Example 7.3 (Lexicographic orderings). *Let $(\mathfrak{v}_1, \ldots, \mathfrak{v}_{2^\mathfrak{d}})$ be the vertices of the checkerboard. For $1 \le n \le 2^\mathfrak{d}$, let $(c_\ell^n)_{1 \le \ell \le \mathfrak{d}}$ be the set of checkerboard coordinates with $\mathfrak{v}_n$ as origin. Define σ_n, for $1 \le n \le 2^\mathfrak{d}$, to be the lexicographic order with respect to $(c_\ell^n)_{1 \le \ell \le \mathfrak{d}}$, ie the lexicographic order with vertex $\mathfrak{v}_n$ as origin. Then $\{\sigma_n\}_{1 \le n \le 2^\mathfrak{d}}$ is exhaustive. One such ordering is given by defining*

$$\sigma_n(j) := 1 + \sum_{\ell=1}^{\mathfrak{d}} \big(\prod_{m \le \ell-1} N_m \big)(c_\ell^n(j) - 1);$$

this is shown in the case $\mathfrak{d} = d = 2$ and $N_1 = N_2 = 3$ in Figure 1.2.

Example 7.4 (An algorithm for generating exhaustive sequences of orderings).
Let $(\mathfrak{v}_1, \ldots, \mathfrak{v}_{2^\mathfrak{d}})$ be the vertices of the checkerboard. For any $1 \le n \le 2^\mathfrak{d}$, let $(x_\ell^n)_{1 \le \ell \le \mathfrak{d}}$ be the set of checkerboard coordinates with $\mathfrak{v}_n$ as origin. For $n = 1, \ldots, 2^\mathfrak{d}$, if σ_n is not already defined, define σ_n in such a way that all the maps $j \mapsto x_\ell^n(\sigma_n(j))$ are non-decreasing, then, for $\bar{n} \in \{1, \ldots, 2^\mathfrak{d}\}$ such that $\mathfrak{v}_{\bar{n}}$ has for coordinates $(c_1^n(\sigma_n(N)), \ldots, c_\ell^n(\sigma_n(N)))$, define $\sigma_{\bar{n}}(j) := \sigma_n(N - j + 1)$. Then $\{\sigma_n\}_{1 \le n \le 2^\mathfrak{d}}$ is exhaustive.

Figure 7.1 illustrates an exhaustive sequence of orderings for a 3×3 checkerboard following the algorithm in Example 7.4.

Lemma 7.5. *On a 2-checkerboard, if $\{\sigma_n\}_{1 \le n \le S}$ is exhaustive then $S \ge 2^2$.*

Proof. Suppose we have an exhaustive sequence of orderings on a 2×2 2-checkerboard. Label the subdomains

$$\frac{C \mid D}{A \mid B}$$

By the definition of exhaustive, there exist an ordering σ_n that visits (in order) C, A, B and an ordering σ_m that visits (in order) C, D, B, with these subdomain orderings corresponding to diagonal lines from the top left to bottom right passing, respectively, below and above the centre.

If $\sigma_n = \sigma_m$, then σ_n equals either (C, A, D, B) or (C, D, A, B). In either case, neither ordering, nor its reverse, visits A, B, D or A, C, D in order (with these subdomain orderings corresponding to diagonal lines from the bottom left to top right passing, respectively, below

and above the centre). Therefore the sequence of orderings consisting of σ_n and its reverse is not exhaustive, and so $S \geq 4$.

If $\sigma_n \neq \sigma_m$ then the reverse ordering to σ_n (which visits B, A, C in order, and hence B before C) is neither σ_m (which visits C, D, B in order, and hence B after C) or the reverse ordering to σ_m (since $\sigma_n \neq \sigma_m$). Therefore $S \geq 4$.

For an arbitrary 2-checkerboard, we repeat the argument above in the top left 2×2 corner. $\quad\square$

7.5   Proof of Theorem 1.4

Arguing as in the proof of Theorem 1.2 outlined in §2.5, and using Lemma 5.2 instead of Lemma 2.7, we obtain

$$\epsilon_\times^{n+1} = \sum_{l=0}^{N-1} \mathbf{A}_n^l \mathbf{B}_n \epsilon_\times^n,$$

where we used that $\mathbf{A}_n^N = 0$ since $\mathbf{A}_n$ is the rearrangment of an lower-triangular matrix. It follows that

$$\epsilon_\times^S = \sum_{0 \leq \ell_1, \dots, \ell_S \leq N-1} \mathbf{A}_S^{\ell_S} \mathbf{B}_S \dots \mathbf{A}_1^{\ell_1} \mathbf{B}_1 \epsilon_\times^0,$$

and Theorem 1.4 therefore follows immediately from the following lemma.

Lemma 7.6. *Assume that $\{\sigma_n\}_{1 \leq n \leq S}$ is exhaustive (in the sense of Definition 7.2). Then, for any $\ell_1, \dots, \ell_S \in \mathbb{N}$,*

$$\|\mathbf{A}_S^{\ell_S} \mathbf{B}_S \dots \mathbf{A}_1^{\ell_1} \mathbf{B}_1\|_{(H^1(\Omega))^N \mapsto (H^s(\Omega))^N} = O(\hbar^\infty).$$

To prove Lemma 7.6, we identify the words arising in the entries of the products $\mathbf{A}_S^{\ell_S} \mathbf{B}_S \dots \mathbf{A}_1^{\ell_1} \mathbf{B}_1$, show that they are not allowed, and then apply Lemma 2.4.

Lemma 7.7. *Assume that $\{\sigma_n\}_{1 \leq n \leq S}$ is exhaustive. Then, no word $w = (w_1, \dots, w_n)$ satisfying*

$$w_1 \succeq_{j_1} w_2 \succeq_{j_2} w_3 \succeq_{j_3} \cdots \succeq_{j_{n-1}} w_n, \tag{7.3}$$

where $(j_i)_{1 \leq i \leq n-1}$ is a surjective sequence with values in $\{1, \dots, S\}$ is allowed.

Proof of Lemma 7.6 assuming Lemma 7.7. Let

$$\mathbf{M} := \mathbf{A}_S^{\ell_S} \mathbf{B}_S \dots \mathbf{A}_1^{\ell_1} \mathbf{B}_1.$$

Observe that

$$(\mathbf{B}_m \mathbf{X})_{i,j} = \sum_{\ell=1}^N (\mathbf{B}_m)_{i,\ell} (\mathbf{X})_{\ell,j} = \sum_{\ell \geq \sigma_m^{-1}(i)} (\mathbf{T})_{i,\sigma_m(\ell)} (\mathbf{X})_{\ell,j}, \tag{7.4}$$

and

$$(\mathbf{A}_m \mathbf{X})_{i,j} = \sum_{\ell=1}^N (\mathbf{A}_m)_{i,\ell} (\mathbf{X})_{\ell,j} = \sum_{\ell \leq \sigma_m^{-1}(i)} (\mathbf{T})_{i,\sigma_m(\ell)} (\mathbf{X})_{\ell,j}. \tag{7.5}$$

Now, any product of $\mathbf{A}_\ell$ and $\mathbf{B}_{\ell'}$, $1 \leq \ell, \ell' \leq N$, can be understood as the sum of $\mathcal{T}_{w'}$ for suitable words w'. Recall from the definition of $\mathcal{T}_{w'}$ (2.9) that the last letter in w appears in the left-most operator in the composition defining $\mathcal{T}_{w'}$. The equation (7.4) therefore implies that left-multiplying such a product by $\mathbf{B}_m$ results in a linear combination of operators $\mathcal{T}_w$ such that the last letter of w (corresponding to i in (7.4)) is $\preceq_m$ the previous letter (corresponding to $\sigma_m(\ell)$ in (7.4)). Similarly, (7.5) implies that left-multiplying a product of $\mathbf{A}_\ell$ and $\mathbf{B}_{\ell'}$, $1 \leq \ell, \ell' \leq N$, by $\mathbf{A}_m$ results in a linear combination of operators $\mathcal{T}_w$ such that the last letter of w (corresponding to i in (7.5)) is $\succeq_m$ the previous letter (corresponding to $\sigma_m(\ell)$ in (7.5)).

By induction, it follows that the entries of $\mathbf{M}$ are linear combinations of operators of the form $\mathcal{T}_w$, where w is a word of size $|w| := n \geq S$ that can be written as the concatenation

$$w = w^0 w^1 \ldots w^S,$$

where

$$|w^0| = 1,$$

and each word w^m for $1 \leq m \leq S$ is of size $|w^m| = n_m := 1 + \ell_m$ and verifies

$$\begin{cases} w^m = (\sigma_m(i_1^m), \ldots, \sigma_m(i_{n_m}^m)), \\ i_1^m \leq \sigma_m^{-1}(w_{n_{m-1}}^{m-1}), \\ i_k^m \geq \sigma_m^{-1}(w_{k-1}^m) \text{ for all } 2 \leq k \leq n_m. \end{cases}$$

Observe that, by definition of $\succeq_m$, for any $1 \leq m \leq S$

$$w_1^m \preceq_m w_{n_{m-1}}^{m-1} \qquad \text{and} \qquad w_k^m \succeq_m w_{k-1}^m \text{ for all } 2 \leq k \leq n_m.$$

On the other hand, by the definition of exhaustive, for any $1 \leq m \leq S$, there is $1 \leq \overline{m} \leq S$ such that $\preceq_m = \succeq_{\overline{m}}$. We thus have, for any $1 \leq m \leq S$

$$w_{n_{m-1}}^{m-1} \succeq_m w_1^m \succeq_{\overline{m}} w_2^m \succeq_{\overline{m}} \cdots \succeq_{\overline{m}} w_{n_m}^m.$$

It follows that, denoting $w = (w_1, \ldots, w_n)$, there is a *surjective* sequence $(j_i)_{1 \leq i \leq n-1}$ with values in $\{1, \ldots, S\}$ so that

$$w_1 \succeq_{j_1} w_2 \succeq_{j_2} w_3 \succeq_{j_3} \cdots \succeq_{j_{n-1}} w_n.$$

Hence w is not allowed by Lemma 7.7, and Lemma 2.4 gives the result. $\qquad\square$

Proof of Lemma 7.7. Seeking a contradiction, assume that such a word w is allowed. By Lemma 7.1, all the coordinates maps $i \mapsto c_\ell(w_i)$ are monotonic. As a consequence, since $\{\sigma_n\}$ is exhaustive, there exists an $1 \leq s \leq S$ so that

$$w_1 \preceq_s w_2 \preceq_s w_3 \preceq_s \cdots \preceq_s w_n. \tag{7.6}$$

However, there exists $1 \leq i \leq n-1$ such that $j_i = s$. Therefore, by the assumption (7.3) and (7.6)

$$w_i \succeq_{j_i} w_{i+1} \quad \text{and} \quad w_i \preceq_{j_i} w_{i+1},$$

hence $w_i = w_{i+1}$, which contradicts the definition of a word. $\qquad\square$

Remark 7.8 (The analogue of Theorem 1.4 when $c \not\equiv 1$). *In the general case when $c \not\equiv 1$, we define an ordering $\{\sigma_n\}_{1 \leq n \leq S}$ to be* exhaustive *if (i) for any $1 \leq m \leq S$, there is $1 \leq \overline{m} \leq S$ such that $\sigma_m(i) = \sigma_{\overline{m}}(N - i + 1)$, and (ii) for any allowed word w of size $n \geq S + 1$, there exists an $1 \leq s \leq S$ so that*

$$w_1 \preceq_s \cdots \preceq_s w_n.$$

The proofs of Lemma 7.6 and Lemma 7.7 still hold, and this therefore leads immediately to the generalisation of Theorem 1.4 when $c \not\equiv 1$ with the result that, if $\{\sigma_n\}$ is exhaustive (as defined above) and of size S then

$$\|u^S - u\|_{H_k^s(\Omega)} \leq k^{-M} \|u^0 - u\|_{H_k^1(\Omega)}.$$

For such orderings to always exist, one would need in principle to define the sequential method in a more general way by not imposing the σ_n to be bijections, allowing to visit a subdomain multiple times during a single sweep – as would a non-straight trajectory visiting the same subdomain multiple times. Constructing such orderings is, however, difficult, since it depends on the specific dynamic of the Hamilton flow imposed by the non-constant c, and we have therefore not pursued this further in this paper.

8 Numerical experiments

8.1 The overlapping Schwarz methods on the discrete level

Given a finite-dimensional subspace of $H_0^1(\Omega)$, which we denote by $V_h(\Omega)$, the discrete analogues of the parallel and sequential methods described in §1.4.5 and §1.4.6 (respectively) compute approximations to the solution of (1.5) in $V_h(\Omega)$ using Galerkin approximations to the solutions of the local problems (1.7) and (1.13) in appropriate subsets of $V_h(\Omega)$.

We now describe this process in detail for the discretisation of the parallel overlapping method in §1.4.5, showing that it gives a natural PML-variant of the well-known RAS (restricted additive Schwarz) method (introduced in [8]).

The Galerkin approximation of the PDE (1.5). Let $\{\mathcal{T}_h\}$ be a shape-regular conforming sequence of meshes for Ω. We assume that the edges of $\mathcal{T}_h$ resolve the boundaries of Ω, Ω_{int}, $\Omega_{\text{int,j}}$, and Ω_j, for all $j = 1, \ldots, N$; since all these domains are unions of hyperrectangles, this is straightforward to achieve in practice.

Let $V_h(\Omega) \subset H_0^1(\Omega)$ be the standard finite-element space of continuous piecewise-polynomial functions of degree $\leq p$ on $\mathcal{T}_h$. Given $f \in H^{-1}(\Omega)$, the Galerkin approximation to the solution of (1.5) is defined as the solution of the problem: find $u_h \in V_h(\Omega)$ such that

$$a(u_h, v_h) = \langle f, v_h \rangle_{H^{-1}(\Omega) \times H^1(\Omega)} \quad \text{for all } v_h \in V_h(\Omega), \tag{8.1}$$

where the sesquilinear form $a(\cdot, \cdot)$ is obtained by multiplying the PDE (1.5) by a test function and integrating by parts; i.e., when $d = 2$,

$$a(u, v) := \int_{\Omega_j} \left(k^{-2}\big((D\nabla u) \cdot \nabla \overline{v} - (\beta \cdot \nabla u)\overline{v}\big) - c^{-2}u\overline{v} \right) dx \quad \text{for } u, v \in H_0^1(\Omega), \tag{8.2}$$

where

$$D := \begin{pmatrix} \frac{1}{\gamma_1^2(x_1)} & 0 \\ 0 & \frac{1}{\gamma_2^2(x_2)} \end{pmatrix} \quad \text{and} \quad \beta := \left(\frac{\gamma_1'(x_1)}{\gamma_1^3(x_1)}, \frac{\gamma_2'(x_2)}{\gamma_2^3(x_2)} \right)^T$$

with $\gamma_1(x_1) := 1 + ig_1'(x_1)$ and $\gamma_2(x_2) := 1 + ig_{x_2}'(x_2)$.

We abbreviate $V_h(\Omega)$ to V_h, define $\mathcal{A}_h : V_h \to V_h'$ by

$$(\mathcal{A}_h u_h)(v_h) = a(u_h, v_h) \quad \text{for all } u_h, v_h \in V_h \tag{8.3}$$

and denote the functional on the right-hand side of (8.1) by $f_h \in V_h'$, The variational problem (8.1) then becomes $\mathcal{A}_h u_h = f_h$, and is equivalent to a linear system in $\mathbb{C}^m$, where m is the dimension of V_h.

The Galerkin approximations of the local problems. Let $a_j(\cdot, \cdot)$ be the sesquilinear form obtained by multiplying P_{s}^j be a test function and integrating by parts; when $d = 2$ the definition of $a_j(\cdot, \cdot)$ is the same as the definition of $a(\cdot, \cdot)$ (8.2) except that g_ℓ', $\ell = 1, 2$, is replaced by $g_{\ell,j}'$.

Let $V_{h,j} := \{v_h|_{\Omega_j} : v_h \in V_h\} \cap H_0^1(\Omega_j)$ and let $\mathcal{A}_{h,j} : V_{h,j} \to V_{h,j}'$ be the operator associated with $a_j(\cdot, \cdot)$; i.e.,

$$(\mathcal{A}_{h,j} u_{h,j})(v_{h,j}) = a_j(u_{h,j}, v_{h,j}) \quad \text{for all } u_{h,j}, v_{h,j} \in V_{h,j}. \tag{8.4}$$

We assume that that all the operators $\mathcal{A}_h : V_h \to V_h'$ and $\mathcal{A}_{h,j} : V_{h,j} \to V_{h,j}'$, $j = 1, \ldots, N$ are invertible. When the Helmholtz problem on Ω is nontrapping, we expect this to be true when $(hk)^{2p}k$ is sufficiently small; this threshold was famously identified for the 1-d Helmholtz h-FEM in [29], and the latest results for $d \geq 2$ proving existence of the Galerkin solution (along with error bounds) under this threshold are given in [21]. The results of [21], however, are not immediately applicable when Ω is a hyperrectangle because of the low regularity of $\partial\Omega$.

Restriction and prolongation operators. We define the prolongation operator $\mathcal{R}_{h,j}^T :$ $V_{h,j} \to V_h$ by, for $v_h \in V_{h,j}$,

$$\mathcal{R}_{h,j}^T v_{h,j}(x) = \begin{cases} v_{h,j}(x), & \text{when } x \text{ is a node of } \Omega_j, \\ 0, & \text{when } x \text{ is a node of } \Omega \text{ that is not a node of } \Omega_j. \end{cases}$$

Let $\mathcal{R}_{h,j} : V_h' \to V_{h,j}'$ be the adjoint of $\mathcal{R}_{h,j}^\top$, i.e.,

$$\langle \mathcal{R}_{h,j} f, v_{h,j} \rangle_{V_{h,j}' \times V_{h,j}} = \langle f, \mathcal{R}_{h,j}^T v_{h,j} \rangle_{V_h' \times V_h} \quad \text{for all } f \in V_h' \text{ and } v_{h,j} \in V_{h,j}. \tag{8.5}$$

We also define a weighted prolongation operator $\widetilde{\mathcal{R}}_{h,j}^T : V_{h,j} \to V_h$, involving the partition-of-unity function χ_j, by

$$\widetilde{\mathcal{R}}_{h,j}^T v_{h,j} = \mathcal{R}_{h,j}^T \big(\Pi_h(\chi_j v_{h,j})\big) \quad \text{for all } v_{h,j} \in V_{h,j},$$

where Π_h is nodal interpolation onto V_h. Then, for all $w_h \in V_h$ and all finite element nodes $x \in \Omega$,

$$\sum_{j=1}^N \widetilde{\mathcal{R}}_{h,j}^T (w_h|_{\Omega_j})(x) = \sum_{j:x \in \Omega_j} \big(\Pi_h(\chi_j w_h|_{\Omega_j})\big)(x) = \sum_{j=1}^N \chi_j(x) w_h(x) = w_h(x);$$

i.e.,

$$\sum_{j=1}^N \widetilde{\mathcal{R}}_{h,j}^T (w_h|_{\Omega_j}) = w_h \quad \text{for all } w_h \in V_h. \tag{8.6}$$

The discrete parallel overlapping Schwarz method. Let $\mathfrak{c}_{h,j}^n$ be the finite-element approximation to the local corrector $\mathfrak{c}_j^n$ (1.9); i.e., given $u_h^n \in V_h$, $\mathfrak{c}_{h,j}^n \in V_{h,j}$ is the solution of

$$a_j\big(\mathfrak{c}_{h,j}^n, v_{h,j}\big) = \big\langle f, \mathcal{R}_{h,j}^T v_{h,j} \big\rangle - a\big(u_h^n, \mathcal{R}_{h,j}^T v_{h,j}\big) \quad \text{for all } v_{h,j} \in V_{h,j}.$$

Thus, by (8.3), (8.4), and (8.5),

$$\mathcal{A}_{h,j} \mathfrak{c}_{h,j}^n = \mathcal{R}_{h,j}(f - \mathcal{A}_h u_h^n) \in V_{h,j}' \quad \text{and so} \quad \mathfrak{c}_{h,j}^n = \mathcal{A}_{h,j}^{-1} \mathcal{R}_{h,j}(f - \mathcal{A}_h u_h^n) \in V_{h,j}.$$

In analogy with the second equation in (1.9) and recalling (8.6), we then define

$$u_h^{n+1} := u_h^n + \sum_{j=1}^N \widetilde{\mathcal{R}}_j^T \mathcal{A}_{h,j}^{-1} \mathcal{R}_j(f - \mathcal{A}_h u_h^n) \in V_h;$$

i.e., the method is the preconditioned Richardson iteration

$$u_h^{n+1} := u_h^n + \mathcal{B}_h^{-1}(f - \mathcal{A}_h u_h^n) \in V_h \quad \text{with} \quad \mathcal{B}_h^{-1} := \sum_j \widetilde{\mathcal{R}}_{h,j}^\top \mathcal{A}_{h,j}^{-1} \mathcal{R}_{h,j}. \tag{8.7}$$

The preconditioner $\mathcal{B}_h^{-1} : V_h' \to V_h$ is of the form of the RAS preconditioner [8], with the crucial fact that it involves the weighted prolongation operator $\widetilde{\mathcal{R}}_{h,j}^T$ instead of $\mathcal{R}_{h,j}^T$ (compare, e.g., [12, Definitions 1.12 and 1.13]). The classical 'Optimised' RAS (known as ORAS [50], [12, §2.3.2]) for Helmholtz uses the impedance boundary condition on the subdomain boundaries (see, e.g., [26, 27]), but here we have PML at the subdomain boundaries. We note that numerical results on this version of RAS when the PML scaling function is proportional to k^{-1} were given in [4].

GMRES. The main results of the paper concern the overlapping Schwarz methods described in §1.4.5 and §1.4.6 as fixed-point iterations (on the continuous level). We also present numerical results for the generalised minimal residual (GMRES) algorithm [48] acting as an acceleration of the parallel fixed point iteration in (8.7). Since the iteration counts of the sequential method used as a fixed-point iteration are relatively low, we do not consider GMRES acceleration for this iteration here.

We therefore briefly recall the definition of GMRES applied to the abstract linear system $\mathcal{C}\mathbf{x} = \mathbf{d}$ where $\mathcal{C}$ is an nonsingular complex matrix. For GMRES applied to the parallel method above, $\mathcal{C}$ is the matrix corresponding to the operator $\mathcal{B}_h^{-1}\mathcal{A}_h : V_h \to V_h$ and $\mathbf{d}$ is the vector corresponding to $\mathcal{B}_h^{-1}f_h \in V_h$. Given initial guess $\mathbf{x}^0$, let $\mathbf{r}^0 := \mathbf{d} - \mathcal{C}\mathbf{x}^0$ and

$$\mathcal{K}^n(\mathcal{C}, \mathbf{r}^0) := \mathrm{span}\big\{\mathcal{C}^j\mathbf{r}^0 : j = 0, \ldots, n-1\big\}.$$

For $n \geq 1$, define $\mathbf{x}^n$ to be the unique element of $\mathcal{K}^n$ satisfying the minimal residual property:

$$\|\mathbf{r}^n\|_2 := \|\mathbf{d} - \mathcal{C}\mathbf{x}^n\|_2 \;=\; \min_{\mathbf{x}\in\mathcal{K}^n(C,\mathbf{r}^0)} \|\mathbf{d} - \mathcal{C}\mathbf{x}\|_2.$$

8.2 Common set-up for the experiments

The Helmholtz problem. We work in $\mathbb{R}^2$ with $\Omega_{\mathrm{int}} := (0,1)^2$ and PML width $\kappa = 1/40$, so that $\Omega := (-1/40, 1+1/40)^2$. Note that, for $k = 100, 150, 200, 250, 300, 350$, the number of wavelengths in the PML is $(1/40)/(2\pi/k) = 0.40, 0.60, 0.80, 0.99, 1.19, 1.39$, respectively. We take the PML scaling function as $f_s(x) = 5000x^3/3$, and the right-hand side of the PDE (1.5) to be $f(x) = J_0(k|x - x_0|)$ when $x_0 = (0.5, 0.5)$. Note that taking the right-hand side to be a Helmholtz solution, similar to for a scattering problem, ensures that all of the mass of the solution in phase space is oscillating at frequency approximately k, ensuring that the solution propagates.

The domain decomposition. We consider both strip and checkerboard decompositions of Ω. The overlapping decomposition $\{\Omega_{\mathrm{int},j}\}_{j=1}^N$ (1.2) is taken to have overlap $\delta = 1/40$, with then each overlapping subdomain extended by PML layers of width $\kappa = 1/40$.

The smooth partition of unity functions are constructed with respect to a (still overlapping) cover of Ω_{int} consisting of subsets of $\Omega_{\mathrm{int},j}$, so that each POU function χ_j vanishes in a neighbourhood of $\Omega_{\mathrm{int},j}$. In more detail, for strip decompositions with $\Omega_{\mathrm{int},j} = (a^j, b^j) \times (0,1)$, we let

$$\widetilde{\chi}_j(x) = \begin{cases} e^{-\frac{(b^j-a^j)^2}{4(x-a^j-0.3\delta)(b^j-0.3\delta-x)}}, & x \in (a^j + 0.3\delta, b^j - 0.3\delta), \\ 0, & \text{elsewhere,} \end{cases} \quad j = 2, \ldots, N-1,$$

and

$$\widetilde{\chi}_1(x) = \begin{cases} e^{-\frac{b^1-a^1}{2(b^1-0.3\delta-x)}}, & x < b^1 - 0.3\delta, \\ 0, & \text{elsewhere.} \end{cases} \quad \widetilde{\chi}_N(x) = \begin{cases} e^{-\frac{b^1-a^1}{2(x-a^N-0.3\delta)}}, & x > a^N + 0.3\delta, \\ 0, & \text{elsewhere.} \end{cases}$$

The PoU functions χ_j are then defined by $\chi_j := \widetilde{\chi}_j/(\sum_{j=1}^N \widetilde{\chi}_j)$. For checkerboard decompositions, the PoU functions are created by normalising the Cartesian product of $\widetilde{\chi}_j(x)\widetilde{\chi}_\ell(y)$.

We use the abbreviations RAS-PML and RMS-PML to denote the parallel (additive) and sequential (multiplicative) methods, respectively.

The finite-element method. The finite-dimensional subspace V_h is taken to be standard Lagrange finite elements of degree $p = 2$ on uniform meshes of Ω, with h chosen as the largest number $\leq k^{-1.25}$ such that $1/h$ is an integer (recall from the discussion immediately below (8.4) that we expect the relative-error of the finite-element solutions to be bounded uniformly in k under this choice of h). We write the resulting linear system as $\mathsf{A}\mathsf{u} = \mathsf{f}$.

8.3 Experiment I: $c \equiv 1$

Figure 8.1 plots the rate that the residual decreases as a function of k for

- a strip decomposition, with RAS-PML (Plot (a) of the figure) and RMS-PML with forward-backward sweeping (Plot (b)),

- a checkerboard decomposition, with RAS-PML (Plot (c)) and RMS-PML with the exhaustive sequence of four orderings constructed via Example 7.3 (Plot (d)).

In all cases the rate of residual reduction increases as k increases (as expected from Theorems 1.1-1.4 and 1.6).

(a) strip, RAS-PML: $\rho = \frac{\|A(u-u^N)\|_{\ell^2}}{\|A(u-u^0)\|_{\ell^2}}$

(b) strip, RMS-PML: $\rho = \frac{\|A(u-u^2)\|_{\ell^2}}{\|A(u-u^0)\|_{\ell^2}}$

(c) checkerboard, RAS-PML: $\rho = \frac{\|A(u-u^{N_x+N_y-1})\|_{\ell^2}}{\|A(u-u^0)\|_{\ell^2}}$

(d) checkerboard, RMS-PML: $\rho = \frac{\|A(u-u^4)\|_{\ell^2}}{\|A(u-u^0)\|_{\ell^2}}$
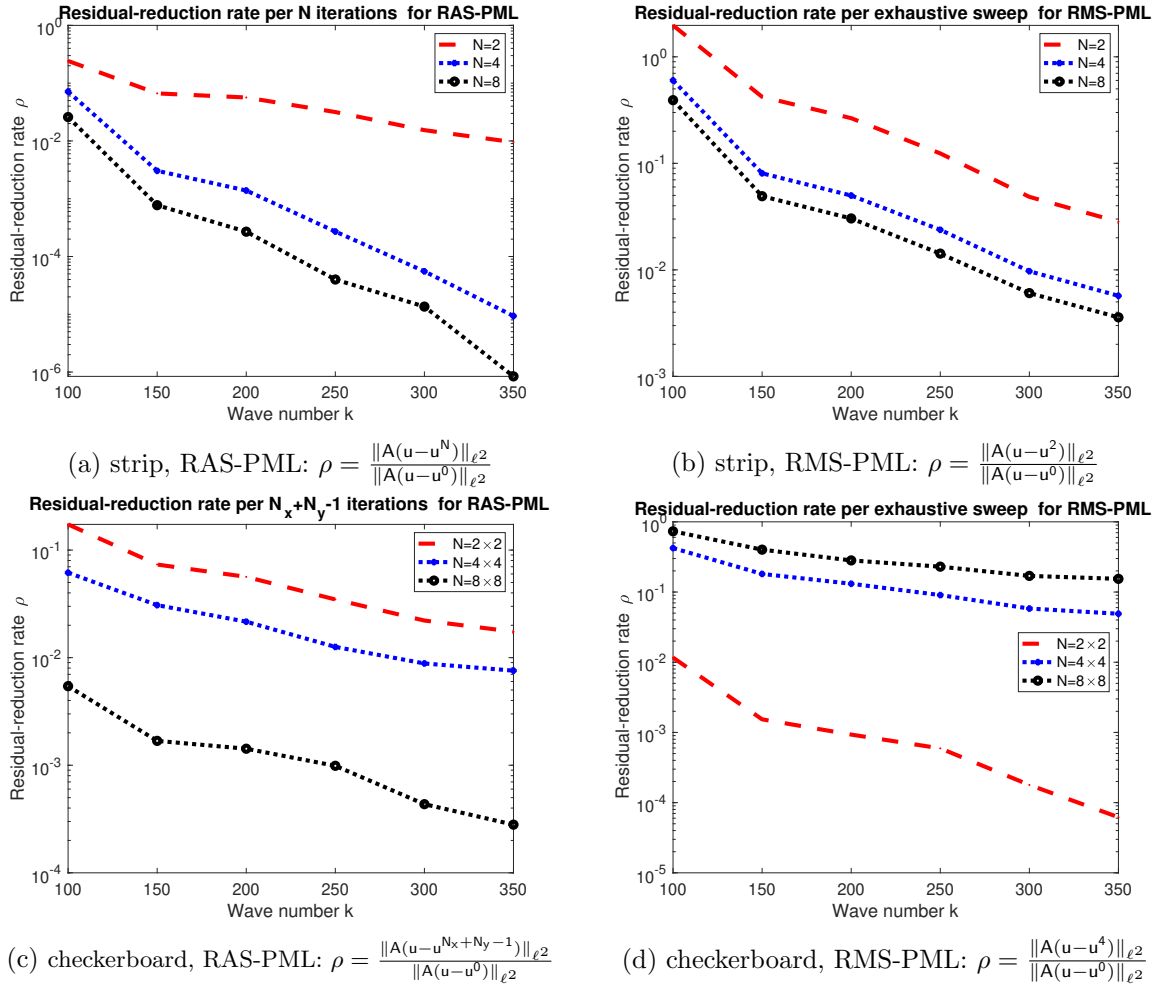
Figure 8.1: The rate of reduction of the residual against k

Figures 8.2, 8.3, 8.4, and 8.5 present the convergence histories for RAS-PML and RMS-PML for different ks and different domain decompositions. We highlight that, for the strip decompositions with $N = 4$ and $N = 8$, Figure 8.2 shows that the residual for RAS-PML on a strip undergoes a significant reduction after every N iterations, as expected from Theorem 1.1.
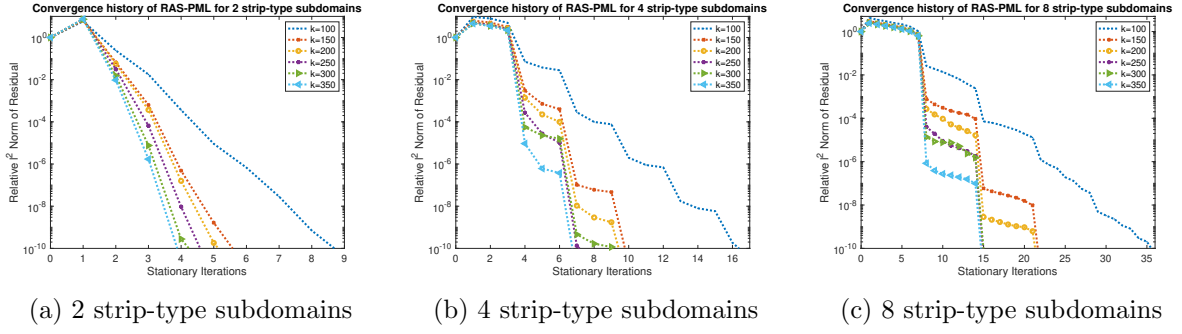
(a) 2 strip-type subdomains (b) 4 strip-type subdomains (c) 8 strip-type subdomains

Figure 8.2: The convergence histories of RAS-PML for strip decompositions

(a) 2 strip-type subdomains (b) 4 strip-type subdomains (c) 8 strip-type subdomains
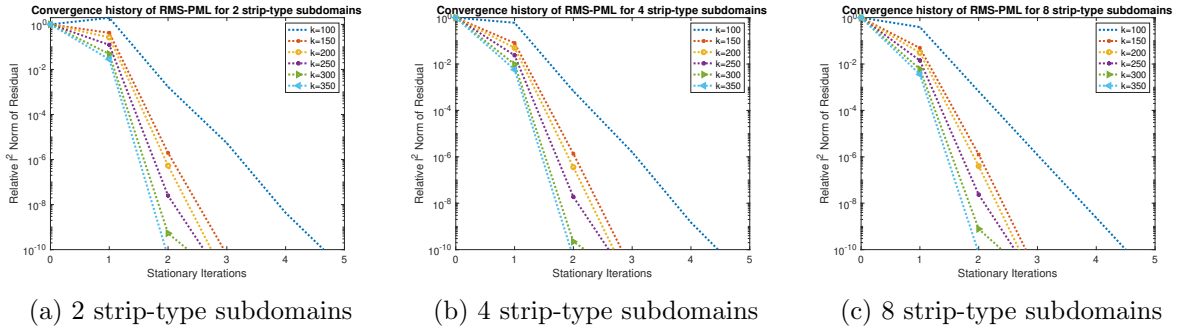
Figure 8.3: The convergence histories of RMS-PML for strip decompositions

(a) 2by2 subdomains (b) 4by4 subdomains (c) 8by8 subdomains
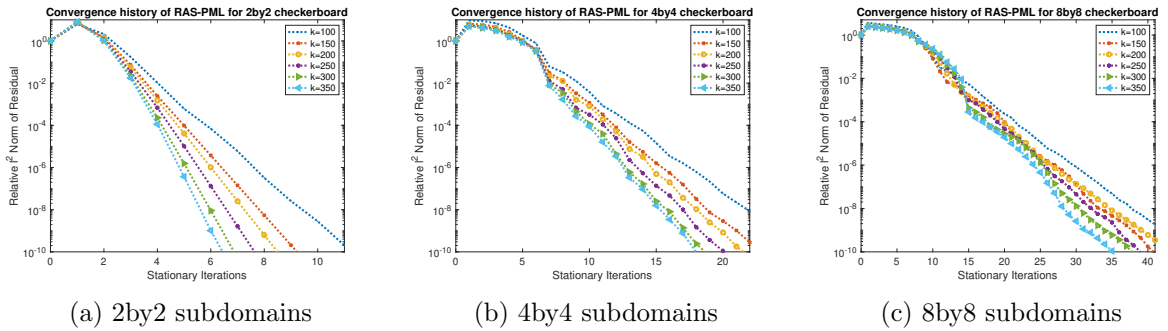
Figure 8.4: The convergence histories of RAS-PML for checkerboard decompositions

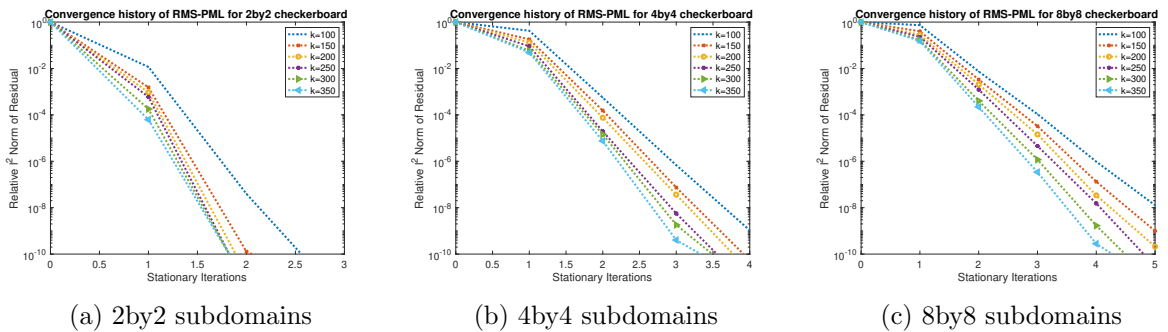(a) 2by2 subdomains (b) 4by4 subdomains (c) 8by8 subdomains

Figure 8.5: The convergence histories of RMS-PML for checkerboard decompositions

8.4 Experiment II: variable c

We consider three cases

Case 1: $c^{-2} = 1$,

Case 2: $c^{-2} \in [2, 1]$ decreases linearly from the center to $|x| = 0.4$ and equals 1 for $|x| > 0.4$,

Case 3: $c^{-2} \in [0.5, 1]$ increases linearly from the center to $|x| = 0.4$ and equals 1 for $|x| > 0.4$.

When c decreases, i.e., c^{-1} increases, as a function of $|x|$, then the problem is nontrapping (see, e.g., [28, Theorem 2.5], [28, §7]). If c increases quickly enough as a function of $|x|$, then the problem is trapping [44], but the growth of the solution operator through an increasing sequence of ks is very sensitive to the particular value of k [32], and depends on the data [11]. Therefore, in these experiments, we do not seek to solve any trapping problems where the solution operator grows faster with k than for nontrapping problems.

Experiment I showed that, in Case 1 ($c \equiv 1$), the rate of residual reduction increases with k for all the methods considered. In this experiment, we compare Cases 1-3 by listing the fixed-point iteration counts (in brackets, GMRES iterations) to obtain the following relative-residual reduction

$$\frac{\|A(u - u^n)\|_{\ell^2}}{\|A(u - u^0)\|_{\ell^2}} < 10^{-6}, \tag{8.8}$$

with Table 1 considering RAS-PML on strips, Table 2 considering RMS-PML on strips, Table 3 considering RAS-PML on checkerboards, and Table 4 considering RMS-PML on checkerboards. In all scenarios the number of iterations are very similar for each of Cases 1, 2, and 3.

Variable c^{-2}	Case 1			Case 2			Case 3		
$k\backslash N$	2	4	8	2	4	8	2	4	8
100	6(6)	11(11)	23(22)	6(6)	13(12)	26(23)	7(6)	11(10)	22(22)
150	4(4)	7(7)	15(15)	5(4)	10(10)	18(18)	5(4)	10(8)	17(15)
200	4(4)	7(7)	15(15)	4(4)	10(10)	18(15)	4(4)	10(8)	15(15)
250	4(4)	7(7)	15(15)	4(4)	10(8)	18(15)	4(4)	8(7)	14(14)
300	4(4)	7(7)	15(15)	4(4)	10(8)	15(15)	4(4)	8(7)	15(15)
350	4(4)	5(5)	8(8)	4(4)	10(7)	13(13)	4(4)	7(7)	13(13)

Table 1: Number of fixed-point iterations (in brackets, GMRES iterations) to obtain the residual reduction (8.8) for RAS-PML with a strip decomposition

Variable c^{-2}	Case 1			Case 2			Case 3		
$k\backslash N$	2	4	8	2	4	8	2	4	8
100	4	4	4	4	3	3	4	4	4
150	3	3	3	3	3	3	3	3	3
200	2	2	2	3	3	2	3	2	3
250	2	2	2	2	2	2	2	2	2
300	2	2	2	2	2	2	2	2	2
350	2	2	2	2	2	2	2	2	2

Table 2: Number of fixed-point iterations to obtain the residual reduction(8.8) for RMS-PML with a strip decomposition, where one iteration equals one forward plus one backward sweep

8.5 Experiment III: variable coefficient in the differential operator

Remark 2.6 above described how the results in Theorems 1.1-1.4 and 1.6 do not hold for the problem

$$k^{-2}\nabla_s \cdot (A\nabla_s u) + c^{-2}u = -f \quad \text{in } \mathbb{R}^d, \qquad u = 0 \quad \text{on } \partial\Omega, \tag{8.9}$$

Variable c^{-2}	Case 1			Case 2			Case 3		
$k \backslash N$	2×2	4×4	8×8	2×2	4×4	8×8	2×2	4×4	8×8
100	8(7)	18(16)	30(27)	8(7)	20(18)	34(30)	8(7)	17(15)	28(27)
150	7(6)	16(15)	28(25)	7(6)	17(16)	30(27)	7(6)	15(14)	27(24)
200	7(6)	15(14)	27(25)	7(6)	16(15)	29(26)	7(6)	14(14)	25(24)
250	6(5)	14(13)	26(24)	6(6)	15(14)	28(26)	7(6)	13(13)	25(23)
300	6(5)	13(13)	26(23)	6(6)	15(14)	26(25)	6(6)	13(12)	25(22)
350	6(5)	13(12)	25(23)	6(6)	15(14)	25(24)	6(6)	12(12)	23(21)

Table 3: Iteration counts of fixed-point iteration (in brackets, GMRES iterations) to obtain the residual reduction (8.8) for RAS-PML with checkerboard decompositions, where one iteration consists of one exhaustive sequence of four orderings.

Variable c^{-2}	Case 1			Case 2			Case 3		
$k \backslash N$	2×2	4×4	8×8	2×2	4×4	8×8	2×2	4×4	8×8
100	2	3	4	2	4	5	2	3	4
150	2	3	4	2	3	4	2	3	4
200	2	3	4	2	3	4	2	3	4
250	2	3	4	2	3	4	2	3	4
300	2	3	4	2	3	3	2	3	3
350	2	3	3	2	3	3	2	3	3

Table 4: Iteration counts of fixed-point iteration to obtain the residual reduction (8.8) for RMS-PML with checkerboard decompositions, where one iteration consists of one exhaustive sequence of four orderings.

where

$$
\nabla_{\mathrm{s}} := \begin{pmatrix} \dfrac{1}{1 + i g_1'(x_1)} \partial_{x_1} \\ \vdots \\ \dfrac{1}{1 + i g_d'(x_d)} \partial_{x_d} \end{pmatrix}
$$

and A is a general symmetric-positive-definite-matrix-valued function with $\operatorname{supp}(I - A) \subset \Omega_{\mathrm{int}}$. This experiment demonstrates that the parallel Schwarz method diverges when applied to (8.9), where the local problems involve the operator (2.12).

In common with the previous two experiments, Ω_{int} is the unit square in 2-d, now centred at the origin to make the expressions for A simpler. We let $\Omega = (-0.5 - \kappa, 0.5 + \kappa) \times (-0.5 - \kappa, 0.5 + \kappa)$ and

$$
A(x) = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} + \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \left(\big(0.4 - |x|\big) + \big|0.4 - |x|\big| \right); \tag{8.10}
$$

i.e., A transitions from being equal to $\begin{pmatrix} 1 & 0.8 \\ 0.8 & 1 \end{pmatrix}$ at the origin to being equal to $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$ near the start of the PML. The variational formulation of (8.9) is (8.1) with the sesquilinear form now

$$
a(u, v) := \int_{\Omega_j} \left(k^{-2} \big(\widetilde{A} \nabla u \big) \cdot \nabla \overline{v} - \big(\widetilde{\beta} \cdot \nabla u \big) \overline{v} \big) - c^{-2} u \overline{v} \right) dx \quad \text{for } u, v \in H_0^1(\Omega),
$$

where

$$\widetilde{A}(x) = \begin{pmatrix} \frac{1}{\gamma_1(x_1)} & 0 \\ 0 & \frac{1}{\gamma_2(x_2)} \end{pmatrix} A(x) \begin{pmatrix} \frac{1}{\gamma_1(x_1)} & 0 \\ 0 & \frac{1}{\gamma_2(x_2)} \end{pmatrix} \quad \text{and} \quad \widetilde{\beta}(x) = \begin{pmatrix} \frac{1}{\gamma_1(x_1)} & 0 \\ 0 & \frac{1}{\gamma_2(x_2)} \end{pmatrix} A(x) \begin{pmatrix} \frac{\gamma_1'(x_1)}{\gamma_1^2(x_1)} \\ \frac{\gamma_2'(x_2)}{\gamma_2^2(x_2)} \end{pmatrix}$$

with $\gamma_1(x_1) := 1 + ig_1'(x_1)$ and $\gamma_2(x_2) := 1 + ig_2'(x_2)$ as before. The local sesquilinear forms $a_j(\cdot, \cdot)$ are defined analogously, with g_ℓ', $\ell = 1, 2$, replaced by $g_{\ell,j}'$.

We consider a strip decomposition with two subdomains and overlap $\delta = 1/40$. We display the results for PML width $\kappa = 1/40, 1/20, 1/10$. The number of wavelengths in the PML at $k = 350$ are then $1.39, 2.78$, and 5.56 respectively.

Table 5 displays the number of fixed-point iterations (in brackets, GMRES iterations) required to achieve the residual reduction (8.8), both for the coefficient A given by (8.10) and (for comparison) $A \equiv I$. Table 5 shows that the parallel Schwarz method applied to (8.9) with A given by (8.10) converges when $\kappa = 1/40$ (with iteration counts almost identical to the case $A \equiv I$), but diverges when $\kappa = 1/20$ and when $\kappa = 1/10$.

When PML is used to approximate the radiation condition, the accuracy increases with the width of the PML (see [30, Theorem 5.5], [6, Theorem 5.7], [20, Theorems 1.2 and 1.5]). The fact the the behaviour of the parallel Schwarz method with $A \not\equiv I$ gets worse as the PML width increases is consistent with the fact, highlighted in Remark 2.6, that PML is not intended to be used when $A \not\equiv I$ (in the sense that the imaginary part of the principal symbol is not single-signed).

$k\backslash\kappa$	A given by (8.10)			$A \equiv I$		
	$\frac{1}{40}$	$\frac{1}{20}$	$\frac{1}{10}$	$\frac{1}{40}$	$\frac{1}{20}$	$\frac{1}{10}$
100	7(7)	diverges(14)	diverges(24)	6(6)	4(4)	7(7)
150	5(5)	diverges(> 100)	diverges(> 100)	4(4)	3(3)	5(5)
200	5(5)	diverges(> 100)	diverges(> 100)	4(4)	3(3)	4(4)
250	4(5)	diverges(> 100)	diverges(> 100)	4(4)	3(3)	3(3)
300	4(5)	diverges(> 100)	diverges(> 100)	4(4)	3(3)	3(3)
350	4(4)	diverges(> 100)	diverges(> 100)	3(4)	3(3)	3(3)

Table 5: Iteration counts (GMRES)(tol=1e-6) for RAS-PML using two subdomains for the problem (8.9) with A given by (8.10).

A   Extending the main results to more general complex absorbing-layer operators

A careful examination of the assumptions used in the proof of Theorems 1.1-1.6 shows that these results hold under the following abstract assumptions, in which $P_\mathcal{A}$ and $P_\mathcal{A}^j$ ($\mathcal{A}$ standing for "absorbing") play the role of (and generalise) P_s and P_s^j, respectively.

Theorem A.1. *Let $\{\Omega_j\}_{j=1}^N$ be an overlapping decomposition of Ω associated with the partition of unity $\{\chi_j\}_{1\le j\le N}$. Let $P, P_\mathcal{A}, P_\mathcal{A}^j \in \Psi_\hbar^2(\mathbb{R}^d)$ be second-order semiclassical pseudodifferential operators (in the sense of §B.2) with $p, p_\mathcal{A}, p_\mathcal{A}^j$ denoting their semiclassical principal symbols and $\Phi_t, \Phi_t^\mathcal{A}, \Phi_t^{\mathcal{A},j}$ denoting the Hamiltonian flows associated to $\operatorname{Re} p$, $\operatorname{Re} p_\mathcal{A}$ and $\operatorname{Re} p_\mathcal{A}^j$ respectively (i.e. defined by (1.15) with p replaced by the appropriate symbol). Assume further that $P, P_\mathcal{A}$, and $P_\mathcal{A}^j$ satisfy the following.*

1. *(Constant coefficients near $\partial\Omega$) $P_\mathcal{A}$ and $P_\mathcal{A}^j$ can be written near $\partial\Omega$ as second-order differential operators with constant coefficients and zero first-order coefficients.*

2. *(Agreement of $P_{\mathcal{A}}$ and $P_{\mathcal{A}}^j$)* $P_{\mathcal{A}} = P_{\mathcal{A}}^j$ *on* $\Omega_{\mathrm{int},j}$.

3. *(Ellipticity)*

    (a) *(Escape to ellipticity (cf. Lemma 4.4))* *For any* $\rho \in T^*\widetilde{\Omega}_j$ *not trapped backward in time, there exists* $\tau \leq 0$ *so that*

    $$\gamma_{[\tau,0]}(\rho) \subset T^*\widetilde{\Omega}_j, \quad and \quad \Phi_\tau(\rho) \in \big\{ |p_{\mathcal{A}}^j(x,\xi)| > 0 \big\}.$$

    Furthermore, the same property holds replacing $\widetilde{\Omega}_j$ *by* $\mathbb{R}^d$.

    (b) *(Ellipticity at infinity (cf. Lemma 3.4))* *For any compact* K, *there exists* $C > 0$ *such that if* $x \in K$ *and* $|\xi| \geq C$ *then* $|p_{\mathcal{A}}^j(x,\xi)| \geq C^{-1}$. *Furthermore the same property holds for* $p_{\mathcal{A}}$.

4. *(Respecting propagation)*

    (a) *On* $\{p_{\mathcal{A}}^j = 0\}$, $\Phi_t^{\mathcal{A},j} = \Phi_t$ *for any* $t \in \mathbb{R}$; *and on* $\{p_{\mathcal{A}} = 0\}$, $\Phi_t^{\mathcal{A}} = \Phi_t$ *for any* $t \in \mathbb{R}$ *(cf. Lemma 4.2).*

    (b) $\operatorname{Im} p_{\mathcal{A}}^j \leq 0$ *near* $\{p_{\mathcal{A}}^j = 0\}$, *and* $\operatorname{Im} p_{\mathcal{A}} \leq 0$ *near* $\{p_{\mathcal{A}} = 0\}$.

Then the results of Theorems 1.1, 1.2, 1.3, 1.4, 1.6 hold (replacing $(P_{\mathrm{s}}, P_{\mathrm{s}}^j)$ *by* $(P_{\mathcal{A}}, P_{\mathcal{A}}^j)$ *in the definition of the iterates).*

Proof of Theorem A.1. As discussed in §2.4, the proofs of Theorems 1.1, 1.2, 1.3, 1.4, 1.6 are the consequences of

(i) Algebra of the error propagation, showing which words appear in the products $\mathcal{T}_w$.

(ii) Semiclassical analysis, showing that $\mathcal{T}_w = O(\hbar^\infty)$ if w is not allowed (Lemma 2.4).

(iii) Properties of the trajectories of the flow associated with P, dictating which w are allowed.

By Point 2, the regions $P_{\mathcal{A}}^j - P_{\mathcal{A}}$ are unchanged, so that, in particular,

$$\chi_j \equiv 0 \quad \text{on} \quad \mathrm{supp}(P_{\mathcal{A}} - P_{\mathcal{A}}^j). \tag{A.1}$$

This implies that (i) above remains unchanged. We define *both* the notion of following a word *and* the quantity $\mathcal{N}$ exactly as before (i.e., by Definition 6.1 and (2.8), respectively).

For (ii), we observe that the proof of Lemma 2.4 uses only the following properties of P, P_{s} and P_{s}^j:

- Point 2 of Lemma 3.4 (ellipticity at infinity): this is Point 3b, and

- Lemma 6.2 (a composite map follows the associated word).

The proof of Lemma 6.2 in turn uses only:

- the fact that in the extended region the PML is linear and the PDE coefficients are constant (this is used to obtain the expression (6.4) involving the extension operator S_j of Definition 3.2): this is a consequence of Point 1,

- the property (A.1), and

- Lemma 4.5 (PML solution-operator bound) and Lemma 4.1 (key propagation lemma).

The proofs of Lemma 4.5 and Lemma 4.1 use only the following: for both Lemma 4.5 and Lemma 4.1:

- Lemma 4.4 (escape to ellipticity): this is Point 3a,

- Lemma 4.3 (forward propagation of regularity from an elliptic point), the proof of which uses only Points 4a and 4b,

- Lemma 3.3 (the homogeneous Dirichlet extension of a solution is a solution on the extended subdomain), the proof of which only uses Point 1.

$\square$

Theorem A.2. (Approximation of outgoing Helmholtz solutions.) *Assume that P and $P_{\mathcal{A}}$ satisfy the assumptions of Theorem A.1, and, in addition, that*

1. $P = -\hbar^2 \nabla \cdot A \nabla - c^{-2}$, *with A smooth and positive definite, c smooth and strictly positive, and $(A, c) = (\mathrm{Id}, 1)$ in a neighbourhood of $\mathbb{R}^d \backslash \Omega_{\mathrm{int}}$.*

2. $P = P_{\mathcal{A}}$ *in Ω_{int}.*

3. *Either (i) P is non-trapping, or (ii) the solution operators for $P_{\mathcal{A}}$ and P, as operators on, respectively, $H_0^1(\Omega)$ and $H^1(\mathbb{R}^d)$, are bounded polynomially in $\hbar^{-1}$.*

Let $f \in H^{-1}(\Omega_{\mathrm{int}})$, and let u and v satisfy

$$\begin{cases} P_{\mathcal{A}} u = f \text{ in } \Omega, \\ u = 0 \text{ on } \partial\Omega, \end{cases} \qquad \begin{cases} P v = f \text{ in } \mathbb{R}^d, \\ v \text{ satisfies the Sommerfeld radiation condition.} \end{cases}$$

Then

$$\text{for all } s > 0 \text{ and } \chi \in C_c^\infty(\Omega_{\mathrm{int}}), \quad \|\chi(u-v)\|_{H_\hbar^s} = O(\hbar^\infty) \|f\|_{H_\hbar^{-1}(\Omega_{\mathrm{int}})}.$$

Proof. Recall from the proof of Theorem A.1 that, under the assumptions of Theorem A.1, Lemma 4.1 and Lemma 4.5 hold with P_s replaced with $P_{\mathcal{A}}$.

By dividing u and v by $\|f\|_{H_\hbar^{-1}(\Omega_{\mathrm{int}})}$, we can assume, without loss of generality, that $\|f\|_{H_\hbar^{-1}(\Omega_{\mathrm{int}})} = 1$. When Point 3.(ii) holds, both u and v are tempered directly by assumption. When Point 3.(i) holds, v is tempered by the non-trapping solution-operator bound for P (from [39] together with either [55, Theorem 3]/ [56, Chapter 10, Theorem 2] or [34], or [7, Theorem 1.3 and §3]) and u is tempered by Lemma 4.5 (with P_s replaced with $P_{\mathcal{A}}$).

Let $\psi \in C_c^\infty(\Omega)$ be such that $\psi \equiv 1$ near Ω_{int}. Let

$$z := u - \psi v,$$

and observe that

$$\begin{cases} P_{\mathcal{A}} z = -[P_{\mathcal{A}}, \psi]v - \psi(P_{\mathcal{A}} - P)v =: F \text{ in } \Omega, \\ z = 0 \text{ on } \partial\Omega. \end{cases}$$

Since v satisfies the Sommerfeld radiation condition,

$$\mathrm{WF}_\hbar v \subset \left\{ \rho, \ \exists t < 0, \Phi_t(\rho) \in T^*(\mathrm{supp}\, f) \right\};$$

indeed, this is a classical consequence of Rellich's uniqueness theorem, the outgoing properties of the free resolvent, and propagation of singularities; see e.g., [19, Proof of Lemma 3.4]. Therefore

$$\mathrm{WF}_\hbar F \subset \left\{ \rho, \ \exists t < 0, \Phi_t(\rho) \in T^*(\mathrm{supp}\, f) \right\} \cap \left(T^*(\mathrm{supp}\, \psi(P_{\mathcal{A}} - P)) \cup T^*(\mathrm{supp}\, \nabla\psi) \right).$$

Let $\chi \in C_c^\infty(\Omega_{\mathrm{int}})$; we aim to show that

$$\mathrm{WF}_\hbar(\chi z) = \emptyset, \tag{A.2}$$

from which Lemma B.8, Part 1, yields the result.

Lemma 4.1 with P_{s} replaced with $P_{\mathcal{A}}$ gives that

$$\mathrm{WF}_\hbar(\chi z) \subset \left\{ \rho \in T^*\big(\operatorname{supp}\chi\big) \quad \text{s.t.} \quad \exists t \leq 0, \ \Phi_t(\rho) \in \mathrm{WF}_\hbar(F) \right\},$$

hence

$$\mathrm{WF}_\hbar(\chi z) \subset \left\{ \rho \in T^*\big(\operatorname{supp}\chi\big) \quad \text{s.t.} \quad \exists t_2 \leq t_1 \leq 0, \right.$$
$$\left. \Phi_{t_1}(\rho) \in T^*\big(\operatorname{supp}\psi(P_{\mathcal{A}} - P)\big) \cup T^*\big(\operatorname{supp}\nabla\psi\big) \ \text{and} \ \Phi_{t_2}(\rho) \in T^*\big(\operatorname{supp}f\big) \right\}.$$

The trajectories for P are straight lines in a neighbourhood of $\mathbb{R}^d \backslash \Omega_{\mathrm{int}}$, but the support properties of χ, $\psi(P_{\mathcal{A}} - P)$, and f, imply that no smooth curve going from $T^*(\operatorname{supp}\chi)$ to $T^*(\operatorname{supp}\psi(P_{\mathcal{A}} - P)) \cup T^*(\operatorname{supp}\nabla\psi)$ then back to $T^*(\operatorname{supp}f)$ can possibly be straight in a neighbourhood of $\mathbb{R}^d \backslash \Omega_{\mathrm{int}}$. We therefore obtain (A.2) and the result follows. $\qquad\square$

Remark A.3 (Comparing Theorems 4.6 and A.2). *Comparing Theorems 4.6 and A.2, we see that the advantage of PML truncation is that it controls the error on the whole of Ω_{int}, whereas the more general approximation considered in Theorem A.2 only controls the error in compact subsets of Ω_{int}. This difference is because of the property of PML that the scaled solution without truncation is precisely the scattering solution (i.e., (4.18) below), thanks to analyticity.*

Example A.4 (Complex absorbing potential). *Let $A \in C^\infty(\mathbb{R}^d; \mathbb{R}^{d\times d}))$ be such that $A(x)$ is symmetric positive definite for every $x \in \mathbb{R}^d$, $c \in C^\infty(\mathbb{R}^d, \mathbb{R})$ be such that $c > 0$, and $(A, c) = (\mathrm{Id}, 1)$ near $\partial\Omega$. Define*

$$P := -k^{-2}\nabla \cdot (A\nabla) - c^{-2},$$
$$P_{\mathcal{A}}^j := -k^{-2}\nabla \cdot (A\nabla) - c^{-2} - iV_j, \quad P_{\mathcal{A}} := -k^{-2}\nabla \cdot (A\nabla) - c^{-2} - iV,$$

where $V_j, V \in C^\infty(\mathbb{R}^d, \mathbb{R})$ are such that $V, V_j \geq 0$, $\{V_j > 0\} = \widetilde{\mathcal{L}}_j$, $\{V > 0\} = \widetilde{\mathcal{L}}$ (in the notation of §3.1), and V and V_j are constant near $\partial\Omega$. Then, the assumptions of Theorem A.1 are satisfied. In particular, the results of Theorems 1.1, 1.2, 1.3, 1.4, 1.6 hold (replacing $(P_{\mathrm{s}}, P_{\mathrm{s}}^j)$ by $(P_{\mathcal{A}}, P_{\mathcal{A}}^j)$ in the definition of the iterates).

Proof. We check the assumptions of Theorem A.1. Point 1 holds because V and V_j are constant near $\partial\Omega$. Point 2 holds by the support properties of V and V_j. Point 3a holds because any trajectory starting from ρ goes to $T^*\widetilde{\mathcal{L}}_j$, but $\{|p_{\mathcal{A}}^j| > 0\} \supset \{|\operatorname{Im}p_{\mathcal{A}}^j| > 0\} = \{V_j > 0\} = T^*\widetilde{\mathcal{L}}_j$. Point 3b holds because $|p_{\mathcal{A}}^j| > |\operatorname{Re}p_{\mathcal{A}}^j|$ and A is symmetric positive definite. Point 4a holds because $\operatorname{Re}p_{\mathcal{A}}^j = \operatorname{Re}p_{\mathcal{A}} = \operatorname{Re}p$, and Point 4b holds because $\operatorname{Im}p_{\mathcal{A}}^j = -V_j \leq 0$ and $\operatorname{Im}p_{\mathcal{A}} = -V \leq 0$. $\qquad\square$

Example A.5 (PML on the main domain, CAP on subsets of the subdomains). *Let $A \in C^\infty(\mathbb{R}^d; \mathbb{R}^{d\times d}))$ be such that $A(x)$ is symmetric positive definite for every $x \in \mathbb{R}^d$, $c \in C^\infty(\mathbb{R}^d, \mathbb{R})$ be such that $c > 0$, and $(A, c) = (\mathrm{Id}, 1)$ near $\partial\Omega$. Define*

$$P := -k^{-2}\nabla \cdot (A\nabla) - c^{-2},$$

$$P_{\mathcal{A}} := -k^{-2}\nabla_{\mathrm{s}} \cdot (A\nabla_{\mathrm{s}}) - c^{-2}, \quad (\nabla_{\mathrm{s}})_\ell := \frac{1}{1 + ig'_\ell(x_\ell)}\partial_{x_\ell},$$

$$P_{\mathcal{A}}^j := -k^{-2}\nabla_{\mathrm{s}} \cdot (A\nabla_{\mathrm{s}}) - c^{-2} - iV_j,$$

where $V_j \in C^\infty(\mathbb{R}^d, \mathbb{R})$ are such that $V_j \geq 0$ and (in the notation of §3.1)

$$\operatorname{supp}V_j = \bigcup_{\mathfrak{e} \in \mathcal{Z}^j \backslash \mathcal{Z}_{\partial\Omega}^j} \widetilde{\mathfrak{e}}$$

50

(in the PML setting of the rest of the paper, the region on the right-hand side of this last equality is $\bigcup_{\ell=1}^{d} \overline{\{x \in \mathbb{R}^d \,:\, g_{\ell,j}(x_\ell) \neq g_\ell(x_\ell)\}}$).

Then, the assumptions of Theorem A.1 are satisfied. In particular, the results of Theorems 1.1, 1.2, 1.3, 1.4, 1.6 hold (replacing (P_s, P_s^j) by $(P_{\mathcal{A}}, P_{\mathcal{A}}^j)$ in the definition of the iterates).

Proof. The assumptions of Theorem A.1 are verified for $P_{\mathcal{A}}$ by remarking that the results of §3 and §4 (in particular, Lemma 3.4, Lemma 4.2 and Lemma 4.3), still hold for $P_{\mathcal{A}}$ instead of P_s using the fact that $A = \mathrm{Id}$ near $\widetilde{\mathcal{L}}$. These results also hold for $P_{\mathcal{A}}^j$ by this same reasoning combined with the arguments in the proof of Example A.4. Observe that the key point is that $A = \mathrm{Id}$ whenever any PML-scaling is active. $\qquad\square$

B   Recap of results from semiclassical analysis

We recap here some standard definitions and results of semiclassical analysis that we use throughout the paper. The two main tools that we use are, as highlighted in §2.3, semiclassical ellipticity and forward propagation of regularity, in the forms stated below as Lemma B.9 and Lemma B.11, respectively.

B.1   Weighted Sobolev spaces

Recall that $\hbar := k^{-1}$. The *semiclassical Fourier transform* is defined by

$$(\mathcal{F}_\hbar u)(\xi) := \int_{\mathbb{R}^d} \exp\big(-ix \cdot \xi/\hbar\big) u(x)\, dx,$$

with inverse

$$(\mathcal{F}_\hbar^{-1} u)(x) := (2\pi\hbar)^{-d} \int_{\mathbb{R}^d} \exp\big(ix \cdot \xi/\hbar\big) u(\xi)\, d\xi;$$

see [59, §3.3]; i.e., the semiclassical Fourier transform is just the usual Fourier transform with the transform variable scaled by $\hbar$. These definitions imply that, with $D := -i\partial$,

$$\mathcal{F}_h\big((\hbar D)^\alpha)u\big) = \xi^\alpha \mathcal{F}_\hbar u \quad \text{and} \quad \|u\|_{L^2(\mathbb{R}^d)} = \frac{1}{(2\pi\hbar)^{d/2}} \|\mathcal{F}_\hbar u\|_{L^2(\mathbb{R}^d)}; \qquad \text{(B.1)}$$

see, e.g., [59, Theorem 3.8]. Let

$$H_\hbar^s(\mathbb{R}^d) := \Big\{ u \in \mathcal{S}'(\mathbb{R}^d) \text{ such that } \langle\xi\rangle^s (\mathcal{F}_\hbar u) \in L^2(\mathbb{R}^d) \Big\},$$

where $\langle\xi\rangle := (1 + |\xi|^2)^{1/2}$, $\mathcal{S}(\mathbb{R}^d)$ is the Schwartz space (see, e.g., [38, Page 72]), and $\mathcal{S}'(\mathbb{R}^d)$ its dual. Define the norm

$$\|u\|_{H_\hbar^m(\mathbb{R}^d)}^2 = \frac{1}{(2\pi\hbar)^d} \int_{\mathbb{R}^d} \langle\xi\rangle^{2m} |\mathcal{F}_\hbar u(\xi)|^2\, d\xi.$$

The properties (B.1) imply that the space $H_\hbar^s(\mathbb{R}^d)$ is the standard Sobolev space $H^s(\mathbb{R}^d)$ with each derivative in the norm weighted by $\hbar := k^{-1}$. The norm on $H_\hbar^s(\Omega)$ for $\Omega \subset \mathbb{R}^d$ is then defined by

$$\||v\||_{H_\hbar^s(\Omega)} := \inf_{V|_\Omega = v, V \in H_\hbar^s(\mathbb{R}^d)} \||V\||_{H_\hbar^s(\mathbb{R}^d)};$$

see, e.g., [38, Page 77]. When working with the wavenumber k (instead of $\hbar$) we write $H_\hbar^s(\mathbb{R}^d)$ as $H_k^s(\mathbb{R}^d)$ and $H_k^s(\Omega)$ for $H_\hbar^s(\Omega)$, and similarly for the norms, to avoid the cumbersome notation $H_{k^{-1}}^s(\mathbb{R}^d)$ and $H_{k^{-1}}^s(\Omega)$.

When Ω is Lipschitz and $s \geq 1$, the norm $\||\cdot\||_{H_k^s(\Omega)}$ is equivalent (with the constants in the norm equivalence independent of k) to the norm $\|\cdot\|_{H_k^s(\Omega)}$ defined by (1.14) for $s \in \mathbb{Z}^+$ and by interpolation for $s > 1$; see, e.g., [9, §4]. For simplicity, from here on we drop the "triple bar" notation and use only $\|\cdot\|_{H_\hbar^s}$ for norms on $H_\hbar^s$.

B.2 Semiclassical pseudodifferential operators

The set of all possible positions x and momenta (i.e. Fourier variables) ξ is denoted by $T^*\mathbb{R}^d$; this is known informally as "phase space". Strictly, $T^*\mathbb{R}^d := \mathbb{R}^d \times (\mathbb{R}^d)^*$, but for our purposes, we can consider $T^*\mathbb{R}^d$ as $\{(x,\xi) : x \in \mathbb{R}^d, \xi \in \mathbb{R}^d\}$.

Definition B.1 (Symbols). *We say that $a \in C^\infty(T^*\mathbb{R}^d)$ is a symbol of order m (and write $a \in S^m(T^*\mathbb{R}^d)$) if, given multiindices α and β, there exists $C_{\alpha,\beta}$ such that, for all $(x,\xi) \in T^*\mathbb{R}^d$,*

$$|\partial_x^\alpha \partial_\xi^\beta a(x,\xi)| \le C_{\alpha,\beta}\langle \xi \rangle^m.$$

Definition B.2 (Semiclassical pseudodifferential operators). *Fix $\chi_0 \in C_c^\infty(\mathbb{R})$ to be identically 1 near 0. We say that an operator $A : C_c^\infty(\mathbb{R}^d) \to \mathcal{D}'(\mathbb{R}^d)$ is a semiclassical pseudodifferential operator of order m, and write $A \in \Psi_\hbar^m(\mathbb{R}^d)$, if A can be written as*

$$Au(x) = \frac{1}{(2\pi\hbar)^d} \int_{\mathbb{R}^d} e^{\frac{i}{\hbar}\langle x-y,\xi\rangle} a(x,\xi)\chi_0(|x-y|)u(y)dyd\xi + O(\hbar^\infty)_{\Psi^{-\infty}}, \tag{B.2}$$

where $a \in S^m(T^*\mathbb{R}^d)$ and $E = O(\hbar^\infty)_{\Psi^{-\infty}}$, if for all $N > 0$ there exists $C_N > 0$ such that

$$\|E\|_{H_\hbar^{-N}(\mathbb{R}^d) \to H_\hbar^N(\mathbb{R}^d)} \le C_N \hbar^N.$$

We use the notation $\mathrm{Op}_\hbar a$ for the operator A in (B.2) with $E = 0$, and define

$$\Psi^{-\infty} := \bigcap_m \Psi^m, \qquad S^{-\infty} := \bigcap_m S^m, \qquad \Psi^\infty := \bigcup_m \Psi^m, \qquad S^\infty := \bigcup_m S^m.$$

Theorem B.3. ([14, Propositions E.17 and E.19].) *If $A \in \Psi_\hbar^{m_1}$ and $B \in \Psi_\hbar^{m_2}$, then*

(i) $AB \in \Psi_\hbar^{m_1+m_2}$,

(ii) For any $s \in \mathbb{R}$, A is bounded uniformly in $\hbar$ as an operator from $H_\hbar^s$ to $H_\hbar^{s-m_1}$.

Principal symbol. There exists a map

$$\sigma_\hbar^m : \Psi^m \to S^m/hS^{m-1}$$

called the *principal symbol map* such that the sequence

$$0 \to hS^{m-1} \xrightarrow{\mathrm{Op}_\hbar} \Psi^m \xrightarrow{\sigma_\hbar^m} S^m/hS^{m-1} \to 0$$

is exact and, for $a \in S^m$,

$$\sigma_\hbar^m\big(\mathrm{Op}_\hbar(a)\big) = a \mod \hbar S^{m-1};$$

see [59, Page 213], [14, Proposition E.14]. When applying the map $\sigma_\hbar^m$ to elements of $\Psi_\hbar^m$, we denote it by $\sigma_\hbar$ (i.e. we omit the m dependence) and we use $\sigma_\hbar(A)$ to denote one of the representatives in S^m (with the results we use then independent of the choice of representative). By [14, Proposition E.17],

$$\sigma_\hbar(AB) = \sigma_\hbar(A)\sigma_\hbar(B). \tag{B.3}$$

Compactification at infinity. To deal with the behaviour of functions on phase space uniformly near $\xi = \infty$ (so-called *fiber infinity*), we consider the *radial compactification* in the ξ variable of $T^*\mathbb{R}^d$. This is defined by

$$\overline{T^*\mathbb{R}^d} := \mathbb{R}^d \times B^d,$$

where B^d denotes the closed unit ball, considered as the closure of the image of $\mathbb{R}^d$ under the radial compactification map

$$\mathsf{RC} : \xi \mapsto \xi/(1 + \langle \xi \rangle);$$

see [14, §E.1.3]. Near the boundary of the ball, $|\xi|^{-1} \circ \mathsf{RC}^{-1}$ is a smooth function, vanishing to first order at the boundary, with $(|\xi|^{-1} \circ \mathsf{RC}^{-1}, \widehat{\xi} \circ \mathsf{RC}^{-1})$ thus giving local coordinates on the ball near its boundary. The boundary of the ball should be considered as a sphere at infinity consisting of all possible *directions* of the momentum variable. When appropriate (e.g., in dealing with finite values of ξ only), we abuse notation by dropping the composition with RC from our notation and simply identifying $\mathbb{R}^d$ with the interior of B^d.

Definition B.4 (Operator wavefront set). *We say that a point $(x_0, \xi_0) \in \overline{T^*\mathbb{R}^d}$ is not in the wavefront set of an operator $A \in \Psi^m$, and write $(x_0, \xi_0) \notin \mathrm{WF}\, A$, if there exists a neighbourhood U of (x_0, ξ_0) such that A can be written as in* (B.2) *with*

$$\sup_{(x,\xi) \in U} |\partial_x^\alpha \partial_\xi^\beta a(x,\xi) \langle \xi \rangle^N| \leq C_{\alpha,\beta N} \hbar^N.$$

By [14, Equation E.2.5],

$$\mathrm{WF}_\hbar(AB) \subset \mathrm{WF}_\hbar(A) \cap \mathrm{WF}_\hbar(B). \tag{B.4}$$

B.3 Tempered distribution

Definition B.5 (Tempered distributions). *An $\hbar$-dependent family of distributions $v_\hbar \in \mathcal{D}'(\mathbb{R}^2)$ is tempered if*

for all $\chi \in C_c^\infty(\mathbb{R}^2)$ there exist $C, K > 0$ such that $\|\chi v_\hbar\|_{H_\hbar^{-K}} \leq C\hbar^{-K}$.

Definition B.6 (Semiclassical wavefront set). *Let v be a $\hbar$-family of tempered distributions on $\mathbb{R}^d$. For any $s, M \in \mathbb{R}$, the strict semiclassical wavefront set of v at regularity (s, M), denoted $\mathrm{WF}_\hbar^{s,M} v$, is defined as follows: $\rho \in T^*\mathbb{R}^d$ is not in $\mathrm{WF}_\hbar^{s,M} v$ if for any $A \in \Psi_\hbar^\infty(\mathbb{R}^d)$ so that $\mathrm{WF}_\hbar A$ is a sufficiently close neighbourhood of ρ,*

$$\hbar^{-M} \|Av\|_{H_\hbar^s} \to 0.$$

In addition

$$\mathrm{WF}_\hbar^{s,\infty} v := \bigcap_{M \geq 0} \mathrm{WF}_\hbar^{s,M} v, \quad \mathrm{WF}_\hbar^{\infty,\infty} v := \bigcap_{s \geq 0} \mathrm{WF}_\hbar^{s,\infty} v,$$

and $\mathrm{WF}_\hbar v := \mathrm{WF}_\hbar^{\infty,\infty} v$.

By [14, Equation E.2.18], if v is tempered (in the sense of Definition B.5) then

$$\mathrm{WF}_\hbar(Av) \subset \mathrm{WF}_\hbar(A) \cap \mathrm{WF}_\hbar(v). \tag{B.5}$$

B.4 Semiclassical ellipticity

We say that $A \in \Psi_\hbar^m(\mathbb{R}^d)$ is *elliptic* at $(x_0, \xi_0) \in \overline{T^*\mathbb{R}^d}$ if there exists a neighbourhood U of (x_0, ξ_0) such that
$$\inf_{(x,\xi) \in U} |\sigma_\hbar(A)\langle\xi\rangle^{-m}| \geq c > 0.$$

Lemma B.7. (Elliptic parametrix [14, Proposition E.32].) *Let $A \in \Psi_\hbar^m$ and $B \in \Psi_\hbar^\ell$ be such that B is elliptic on $\mathrm{WF}_\hbar A$. Then there exist $E, E' \in \Psi_\hbar^{m-\ell}$ such that*
$$A = BE + O(\hbar^\infty)_{\Psi^{-\infty}} = E'B + O(\hbar^\infty)_{\Psi^{-\infty}}.$$

We use the two following applications of Lemma B.7.

Lemma B.8 (Ellipticity at infinity in the ξ variable). *Let $Q \in \Psi_h^m(\mathbb{R}^d)$ be such that, given a compact set $K \subset \mathbb{R}^d$, there exists $C > 0$ such that for all $x \in K$*

$$\text{if} \quad |\xi| \geq C, \quad \text{then} \quad |\sigma_\hbar(Q)(x,\xi)| \geq C^{-1}|\xi|^m. \tag{B.6}$$

Then, for any $(s, M) \in [0, +\infty]^2$, and any $\chi, \widetilde\chi \in C_c^\infty(\mathbb{R}^2)$ such that $\widetilde\chi = 1$ near $\mathrm{supp}\,\chi$, the following holds.

1. *For any $\hbar$-family of tempered distributions $v_\hbar$ such that $\mathrm{WF}_\hbar^{s,M}(\chi v) = \emptyset$ and $\|\widetilde\chi Q v\|_{H_\hbar^{s-m}} = o(\hbar^M)$,*
$$\|\chi v\|_{H_\hbar^s} = o(\hbar^M).$$

2. *Given $s > 0$ there exists $C > 0$ such that, for any $\hbar$-family of tempered distributions $v_\hbar$,*
$$\|\chi v\|_{H_\hbar^s} \leq C\Big(\|\chi v\|_{L^2} + \|\widetilde\chi Q v\|_{H_\hbar^{s-m}} + O(\hbar^\infty)\Big).$$

Proof. We begin with Point 1. We do the proof for $M < +\infty$, the proof for $M = +\infty$ is similar. Let $\psi \in C_c^\infty(\mathbb{R})$ be such that $\psi = 1$ near $B(0, A)$. Then, on the one hand, by definition of the wavefront-set,
$$\hbar^{-M}\|\psi(|\hbar D|)\chi v\|_{H_\hbar^s} \to 0.$$
On the other hand, by the assumption (B.6) and the properties (B.3) and (B.4), $\widetilde\chi Q$ is semi-classicaly elliptic near $\mathrm{WF}_\hbar(1 - \psi(|\hbar D|))\chi$; hence
$$(1 - \psi(|\hbar D|))\chi = E\widetilde\chi Q + O(\hbar^\infty)_{\psi_\hbar^{-\infty}},$$
with $E \in \Psi_\hbar^{-m}(\mathbb{R}^d)$. Applying v to the above gives
$$\hbar^{-M}\|(1 - \psi(|\hbar D|))\chi v\|_{H_\hbar^s} \to 0,$$
and the result follows. The proof of Point 2 is very similar to the proof of Point 1 using that given $s > 0$ there exists $C > 0$ such that $\|\psi(|\hbar D|)\chi v\|_{H_\hbar^s} \leq C\|\chi v\|_{L^2}$; this follows from Part (ii) of Theorem B.3 since $\psi(|\hbar D|) \in \Psi_\hbar^{-\infty}(\mathbb{R}^d)$. $\qquad\square$

Lemma B.9 (WF statement from symbolic ellipticity). *Let $Q \in \Psi_h^m(\mathbb{R}^d)$. Then, for any $\hbar$-tempered family of distributions $v_\hbar$, any $\chi, \widetilde\chi \in C_c^\infty(\mathbb{R}^d)$ such that $\widetilde\chi = 1$ on $\mathrm{supp}\,\chi$ and any $(s, M) \in (-\infty, +\infty]^2$,*
$$\mathrm{WF}_\hbar^{s,M}(\chi v) \subset \mathrm{WF}_\hbar^{s-m,M}(\widetilde\chi Q v) \cup \{\sigma_\hbar(Q) = 0\}.$$

If Q is elliptic, i.e., $\{\sigma_\hbar(Q) = 0\} = \emptyset$, then Lemma B.9 states that an elliptic pseudodifferential operator does not "move around" mass in phase space; i.e., in any compact set in space, the wave front set of v is contained inside the wavefront set of Qv.

Proof of Lemma B.9. Let $\rho \notin \mathrm{WF}_\hbar^{s,M}(\widetilde{\chi}Qv) \cup \{\sigma_\hbar(Q) = 0\}$, and $b \in C_c^\infty(T^*\mathbb{R}^2)$ with $b = 1$ near ρ be such that

$$\mathrm{supp}\, b \cap \mathrm{WF}_\hbar^{s-m,M}(\widetilde{\chi}Qv) = \emptyset, \quad \mathrm{supp}\, b \cap \{\sigma_\hbar(Q) = 0\} = \emptyset.$$

Then, by (B.3), $\widetilde{\chi}Q$ is semiclassically elliptic on $\mathrm{supp}\, b \cap T^* \mathrm{supp}\, \chi$. Therefore, by the elliptic parametrix construction (Lemma B.7),

$$(\mathrm{Op}_\hbar b)\chi = E\widetilde{\chi}Q + O(\hbar^\infty)_{\Psi_\hbar^{-\infty}},$$

with $E \in \Psi_\hbar^{-m}$ and $\mathrm{WF}_\hbar E \subset \mathrm{supp}\, b$. Therefore $\mathrm{WF}_\hbar E \cap \mathrm{WF}_\hbar^{s-m,M}(\widetilde{\chi}Qv) = \emptyset$ and applying v to the last displayed equation we see that $\rho \notin \mathrm{WF}_\hbar^{s,M}(\chi v)$, which implies the result. $\square$

B.5 Propagation of singularities

Lemma B.10 (Propagation of singularities estimate, [14, Theorem E.47]). *Let $Q \in \Psi_\hbar^m(\mathbb{R}^d)$, of semiclassical principal symbol q. Let φ_t be the Hamiltonian flow associated to $\mathrm{Re}\, q$ (i.e. defined by (1.15) with p replaced by $\mathrm{Re}\, q$), and $A, B, B_1 \in \Psi_\hbar^0(\mathbb{R}^d)$ be so that*

$$\mathrm{Im}\, q \leq 0 \ \ on \ \ \mathrm{WF}_\hbar B_1,$$

and

$$there\ exists\ T \geq 0\ \ s.t.\ \ \varphi_{-T}(x,\xi) \in \mathrm{ell}_\hbar B,\ \ and\ for\ all\ t \in [0,T],\ \varphi_{-t}(x,\xi) \in \mathrm{ell}_\hbar B_1.$$

Then, for any $s > 0$ and any family $v_\hbar$ of $\hbar$-tempered distributions (in the sense of Definition B.5)

$$\|Av\|_{H_\hbar^s} \leq C\|Bv\|_{H_\hbar^s} + C\hbar^{-1}\|B_1 Qv\|_{H_\hbar^{s-m+1}} + O(\hbar^\infty).$$

We use the following an application of Lemma B.10.

Lemma B.11 (Forward propagation of regularity). *Let $Q \in \Psi_\hbar^m(\mathbb{R}^d)$ with semiclassical principal symbol q. Let φ_t denote the Hamiltonian flow associated with $\mathrm{Re}\, q$ (i.e. defined by (1.15) with p replaced by $\mathrm{Re}\, q$), and assume that $\mathrm{Im}\, q \leq 0$. Let $v_\hbar$ be a family of $\hbar$-tempered distributions (in the sense of Definition B.5). Let $(x_0, \xi_0) \in T^*\mathbb{R}^d$ and $(s, M) \in [0, +\infty] \times [0, +\infty]$. If $(x_0, \xi_0) \notin \mathrm{WF}_\hbar^{s,M}(v)$ and there exists $t_0 > 0$ so that*

$$\left\{\varphi_t(x_0, \xi_0),\ t \in [0, t_0]\right\} \cap \mathrm{WF}_\hbar^{s-m+1,M}(\hbar^{-1}Qv) = \emptyset,$$

then $\varphi_{t_0}(x_0, \xi_0) \notin \mathrm{WF}_\hbar^{s,M}(v)$.

Lemma B.11 is a rigorous statement of Point 1 in §2.3; i.e., that the absence of mass in phase space of the solution of $Qu = f$ propagates along the flow defined by Q as long as the trajectory does not intersect the data f.

Proof of Lemma B.11. By continuity of the flow and the assumptions in the lemma, there exist $A, B, B_1 \in \Psi_\hbar^0$ such that A is elliptic near $\varphi_{t_0}(x_0, \xi_0)$, $\mathrm{WF}_\hbar B$ is an arbitrarily-small neighbourhood of (x_0, ξ_0), and, for any $(x, \xi) \in \mathrm{WF}_\hbar A$,

$$there\ exists\ T \geq 0\ s.t.\ \ \varphi_{-T}(x,\xi) \in \mathrm{ell}_\hbar B,\ and\ for\ all\ t \in [0,T],\ \varphi_{-t}(x,\xi) \in \mathrm{ell}_\hbar B_1,$$

with

$$\mathrm{WF}_\hbar B_1 \cap \mathrm{WF}_\hbar^{s-m+1,M}(\hbar^{-1}Qv) = \emptyset.$$

Hence, by the propagation of singularities estimate (Lemma B.10), for any $s > 0$,

$$\|Av\|_{H_\hbar^s} \leq C\|Bv\|_{H_\hbar^s} + C\hbar^{-1}\|B_1 Qv\|_{H_\hbar^{s-m+1}} + O(\hbar^\infty),$$

and the result follows taking $\mathrm{WF}_\hbar B$ small enough. $\square$

Remark B.12. *Working in $\mathbb{R}^d$, as opposed to on a general manifold defined by coordinate charts, allows us to remove the proper-support assumption appearing in [14, Proposition E.32] and [14, Theorem E.47].*

Acknowledgements

References

[1] I. M. Babuška and S. A. Sauter. Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wave numbers? *SIAM Review*, pages 451–484, 2000.

[2] T. Beck, Y. Canzani, and J. L. Marzuola. Quantitative bounds on Impedance-to-Impedance operators with applications to fast direct solvers for PDEs. *Pure and Applied Analysis*, 4(2):225–256, 2022.

[3] J.-P. Berenger. A perfectly matched layer for the absorption of electromagnetic waves. *J. Comp. Phys.*, 114:185–200, October 1994.

[4] N. Bootland, S. Borzooei, V. Dolean, and P.-H. Tournier. Numerical assessment of PML transmission conditions in a domain decomposition method for the Helmholtz equation. In *International Conference on Domain Decomposition Methods*, pages 445–453. Springer, 2022.

[5] N. Bouziani, F. Nataf, and P.-H. Tournier. A Unified Framework for Double Sweep Methods for the Helmholtz Equation. *Journal of Computational Physics*, page 112305, 2023.

[6] J. H. Bramble and J. E. Pasciak. Analysis of a Cartesian PML approximation to acoustic scattering problems in $\mathbb{R}^2$ and $\mathbb{R}^3$. *Journal of Computational and Applied Mathematics*, 247:209–230, 2013.

[7] N. Burq. Semi-classical estimates for the resolvent in nontrapping geometries. *International Mathematics Research Notices*, 2002(5):221–241, 2002.

[8] X-C. Cai and M. Sarkis. A restricted additive Schwarz preconditioner for general sparse linear systems. *SIAM J. Sci. Comp.*, 21(2):792–797, 1999.

[9] S. N. Chandler-Wilde, D. P. Hewett, and A. Moiola. Interpolation of Hilbert and Sobolev spaces: quantitative estimates and counterexamples. *Mathematika*, 61:414–443, 2015.

[10] Z. Chen and X. Xiang. A source transfer domain decomposition method for Helmholtz equations in unbounded domain. *SIAM J. Num. Anal.*, 51(4):2331–2356, 2013.

[11] K. Datchev and A. Vasy. Semiclassical resolvent estimates at trapped sets. *Annales de l'Institut Fourier*, 62(6):2379–2384, 2012.

[12] V. Dolean, P. Jolivet, and F. Nataf. *An introduction to domain decomposition methods: algorithms, theory, and parallel implementation*. SIAM, 2015.

[13] Y. Du and H. Wu. A pure source transfer domain decomposition method for Helmholtz equations in unbounded domain. *Journal of Scientific Computing*, DOI: 10.1007/s10915-020-01249-2, 2020.

[14] S. Dyatlov and M. Zworski. *Mathematical theory of scattering resonances*, volume 200. American Mathematical Soc., 2019.

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

[16] B. Engquist and H.-K. Zhao. Absorbing boundary conditions for domain decomposition. *Applied numerical mathematics*, 27(4):341–365, 1998.

[17] Y. A. Erlangga. Advances in iterative methods and preconditioners for the Helmholtz equation. *Archives of Computational Methods in Engineering*, 15(1):37–66, 2008.

[18] O. G. Ernst and M. J. Gander. Why it is difficult to solve Helmholtz problems with classical iterative methods. In I. G. Graham, T. Y. Hou, O. Lakkis, and R. Scheichl, editors, *Numerical Analysis of Multiscale Problems*, volume 83 of *Lecture Notes in Computational Science and Engineering*, pages 325–363. Springer, 2012.

[19] J. Galkowski, D. Lafontaine, and E. A. Spence. Local absorbing boundary conditions on fixed domains give order-one errors for high-frequency waves. *IMA J. Num. Anal.*, 2023.

[20] J. Galkowski, D. Lafontaine, and E. A. Spence. Perfectly-matched-layer truncation is exponentially accurate at high frequency. *SIAM J. Math. Anal.*, 55(4):3344–3394, 2023.

[21] J. Galkowski and E. A. Spence. Sharp preasymptotic error bounds for the Helmholtz h-FEM. *arXiv 2301.03574*, 2023.

[22] M. J. Gander. Optimized Schwarz methods. *SIAM J. Numer. Anal.*, 44:699–731, 2006.

[23] M. J. Gander and H. Zhang. Schwarz methods by domain truncation. *Acta Numerica*, 2022.

[24] M.J. Gander and H. Zhang. Restrictions on the use of sweeping type preconditioners for helmholtz problems. In *Domain Decomposition Methods in Science and Engineering XXIV*, Lecture Notes in Computational Science and Engineering, pages 321–332. Springer-Verlag, 2018.

[25] M.J. Gander and H. Zhang. A class of iterative solvers for the Helmholtz equation: factorizations, sweeping preconditioners, source transfer, single layer potentials, polarized traces, and optimized Schwarz methods. *SIAM Review*, 61(1):3–76, 2019.

[26] S. Gong, M. J. Gander, I. G. Graham, D. Lafontaine, and E. A. Spence. Convergence of parallel overlapping domain decomposition methods for the Helmholtz equation. *Numerische Mathematik*, 152(2):259–306, 2022.

[27] S. Gong, I. G. Graham, and E. A Spence. Convergence of Restricted Additive Schwarz with impedance transmission conditions for discretised Helmholtz problems. *Mathematics of Computation*, 92(339):175–215, 2023.

[28] I. G. Graham, O. R. Pembery, and E. A. Spence. The Helmholtz equation in heterogeneous media: a priori bounds, well-posedness, and resonances. *Journal of Differential Equations*, 266(6):2869–2923, 2019.

[29] F. Ihlenburg and I. Babuska. Finite element solution of the Helmholtz equation with high wave number part II: the *hp* version of the FEM. *SIAM J. Numer. Anal.*, 34(1):315–358, 1997.

[30] S. Kim and J. E. Pasciak. Analysis of a Cartesian PML approximation to acoustic scattering problems in $\mathbb{R}^2$. *Journal of Mathematical Analysis and Applications*, 370(1):168–186, 2010.

[31] D. Lafontaine and E. A. Spence. Sharp bounds on Helmholtz impedance-to-impedance maps and application to overlapping domain decomposition. *Pure and Applied Analysis*, 5(4):927–972, 2023.

[32] D. Lafontaine, E. A. Spence, and J. Wunsch. For most frequencies, strong trapping has a weak effect in frequency-domain scattering. *arXiv preprint arXiv:1903.12172*, 2019.

[33] D. Lahaye, J. Tang, and K. Vuik, editors. *Modern solvers for Helmholtz problems*. Birkhauser series in Geosystems Mathematics, 2017.

[34] P. D. Lax and R. S. Phillips. *Scattering theory*, volume 26 of *Pure and Applied Mathematics*. Academic Press Inc., Boston, MA, second edition, 1989.

[35] W. Leng and L. Ju. An additive overlapping domain decomposition method for the Helmholtz equation. *SIAM Journal on Scientific Computing*, 41(2):A1252–A1277, 2019.

[36] W. Leng and L. Ju. A diagonal sweeping domain decomposition method with source transfer for the Helmholtz equation. *Commun. Comput. Phys.*, 29:357–398, 2021.

[37] W. Leng and L. Ju. Trace transfer-based diagonal sweeping domain decomposition method for the Helmholtz equation: Algorithms and convergence analysis. *Journal of Computational Physics*, 455:110980, 2022.

[38] W. McLean. *Strongly Elliptic Systems and Boundary Integral Equations*. Cambridge University Press, 2000.

[39] R. B. Melrose and J. Sjöstrand. Singularities of boundary value problems. ii. *Communications on Pure and Applied Mathematics*, 35(2):129–168, 1982.

[40] L. Miller. Refraction of high-frequency waves density by sharp interfaces and semiclassical measures at the boundary. *Journal de mathématiques pures et appliquées*, 79(3):227–269, 2000.

[41] J. G. Muga, J. P. Palao, B. Navarro, and I. L. Egusquiza. Complex absorbing potentials. *Physics Reports*, 395(6):357–426, 2004.

[42] F. Nataf, F. Rogier, and E. de Sturler. Optimal interface conditions for domain decomposition methods. *Technical Report 301*, 1994.

[43] J. C. Nédélec. *Acoustic and electromagnetic equations: integral representations for harmonic problems*. Springer Verlag, 2001.

[44] J. V. Ralston. Trapped Rays in Spherically Symmetric Media and Poles of the Scattering Matrix. *Comm. Pure Appl. Math.*, 24(4):571–582, 1971.

[45] U. V. Riss and H.-D. Meyer. Calculation of resonance energies and widths using the complex absorbing potential method. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 26(23):4503, 1993.

[46] U. V. Riss and H.-D. Meyer. Reflection-free complex absorbing potentials. *Journal of Physics B: Atomic, Molecular and Optical Physics*, 28(8):1475, 1995.

[47] A. Royer, C. Geuzaine, E. Béchet, and A. Modave. A non-overlapping domain decomposition method with perfectly matched layer transmission conditions for the Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering*, 395:115006, 2022.

[48] Y. Saad and M. H. Schultz. GMRES: A generalized minimal residual algorithm for solving nonsymmetric linear systems. *SIAM Journal on Scientific and Statistical Computing*, 7(3):856–869, 1986.

[49] T. Seideman and W. H. Miller. Calculation of the cumulative reaction probability via a discrete variable representation with absorbing boundary conditions. *The Journal of chemical physics*, 96(6):4412–4422, 1992.

[50] A. St-Cyr, M. J. Gander, and S. J. Thomas. Optimized multiplicative, additive, and restricted additive Schwarz preconditioning. *SIAM Journal on Scientific Computing*, 29(6):2402–2425, 2007.

[51] P. Stefanov. Approximating resonances with the complex absorbing potential method. *Communications in Partial Differential Equations*, 30(12):1843–1862, 2005.

[52] C. C. Stolk. A rapidly converging domain decomposition method for the Helmholtz equation. *J. Comp. Phys.*, 241:240–252, 2013.

[53] M. Taus, L. Zepeda-Núñez, R. J. Hewett, and L. Demanet. L-Sweeps: A scalable, parallel preconditioner for the high-frequency Helmholtz equation. *J. Comp. Phys.*, 420:109706, 2019.

[54] A. Toselli. Overlapping method with perfectly matched layers for the solution of the Helmholtz equation. In *Proceedings of the 11th International Conference on Domain Decomposition Methods*, pages 551–557, 1998. http://www.ddm.org/DD11/DD11proc.pdf.

[55] B. R. Vainberg. On the short wave asymptotic behaviour of solutions of stationary problems and the asymptotic behaviour as $t \to \infty$ of solutions of non-stationary problems. *Russian Mathematical Surveys*, 30:1, 1975.

[56] B. R. Vainberg. *Asymptotic methods in equations of mathematical physics*. Gordon & Breach Science Publishers, New York, 1989. Translated from the Russian by E. Primrose.

[57] A. Vasy. Propagation of singularities for the wave equation on manifolds with corners. *Séminaire Équations aux dérivées partielles (Polytechnique)*, pages 1–16, 2004.

[58] L. Zepeda-Núñez and L. Demanet. The method of polarized traces for the 2D Helmholtz equation. *J. Comput. Phys.*, 308:347–388, 2016.

[59] M. Zworski. *Semiclassical analysis*. American Mathematical Society Providence, RI, 2012.