SHARP PREASYMPTOTIC ERROR BOUNDS FOR THE HELMHOLTZ h-FEM

J. GALKOWSKI* AND E. A. SPENCE†

Abstract. In the analysis of the h-version of the finite-element method (FEM), with fixed polynomial degree p, applied to the Helmholtz equation with wavenumber $k \gg 1$, the *asymptotic regime* is when $(hk)^p C_{\rm sol}$ is sufficiently small and the sequence of Galerkin solutions are quasioptimal; here $C_{\rm sol}$ is the $L^2 \to L^2$ norm of the Helmholtz solution operator, with $C_{\rm sol} \sim k$ for nontrapping problems. In the *preasymptotic regime*, one expects that if $(hk)^{2p} C_{\rm sol}$ is sufficiently small, then (for physical data) the relative error of the Galerkin solution is controllably small.

In this paper, we prove the natural error bounds in the preasymptotic regime for the variable-coefficient Helmholtz equation in the exterior of a Dirichlet, or Neumann, or penetrable obstacle (or combinations of these) and with the radiation condition *either* realised exactly using the Dirichlet-to-Neumann map on the boundary of a ball *or* approximated either by a radial perfectly-matched layer (PML) or an impedance boundary condition. Previously, such bounds for $p > 1$ were only available for Dirichlet obstacles with the radiation condition approximated by an impedance boundary condition. Our result is obtained via a novel generalisation of the "elliptic-projection" argument (the argument used to obtain the result for $p = 1$) which can be applied to a wide variety of abstract Helmholtz-type problems.

AMS subject classifications. 35J05, 65N15, 65N30, 78A45

Key words. Helmholtz, FEM, high order, pollution effect, preasymptotic, perfectly-matched layer, elliptic projection.

1. Introduction.

1.1. Informal statement of the main result.

We consider the h-version of the finite-element method (h-FEM), where accuracy is increased by decreasing the meshwidth h while keeping the polynomial degree p constant, applied to the Helmholtz equation.

THEOREM 1.1 (Informal statement of the main result). *Let u be the solution to the variable-coefficient Helmholtz equation, with wavenumber $k > 0$, in the exterior of a Dirichlet, or Neumann, or penetrable obstacle (or combinations of these) and with the radiation condition* either *realised exactly using the Dirichlet-to-Neumann map on the boundary of a ball* or *approximated either by a radial perfectly-matched layer (PML) or an impedance boundary condition. Let $C_{\rm sol}$ be the $L^2 \to L^2$ norm of the solution operator, with $C_{\rm sol} \sim k$ for nontrapping problems.*

Under the natural regularity assumptions on the domain and coefficients for using degree p polynomials, if

$$(hk)^{2p} C_{\rm sol} \text{ is sufficiently small} \tag{1.1}$$

then the Galerkin solution u_h exists, is unique, and satisfies

$$\|u - u_h\|_{H^1_k(\Omega)} \le C\Big(1 + (hk)^p C_{\rm sol}\Big) \min_{v_h \in \mathcal{H}_h} \|u - v_h\|_{H^1_k(\Omega)}, \tag{1.2}$$

$$\|u - u_h\|_{L^2(\Omega)} \le C\Big(hk + (hk)^p C_{\rm sol}\Big) \min_{v_h \in \mathcal{H}_h} \|u - v_h\|_{H^1_k(\Omega)}. \tag{1.3}$$

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

†Department of Mathematical Sciences, University of Bath, Bath, BA2 7AY, UK, E.A.Spence@bath.ac.uk

Furthermore, if the data is k-oscillatory and sufficiently regular (in a sense made precise below), then

$$\frac{\|u-u_h\|_{H^1_k(\Omega)}}{\|u\|_{H^1_k(\Omega)}} \leq C\Big(1+(hk)^p C_{\rm sol}\Big)(hk)^p; \tag{1.4}$$

i.e., the relative H^1_k error can be made controllably small by making $(hk)^{2p}C_{\rm sol}$ sufficiently small.

The norm $\|\cdot\|_{H^1_k(\Omega)}$ in the bounds above is defined by

$$\|v\|^2_{H^1_k(\Omega)} := k^{-2}\,\|\nabla v\|^2_{L^2(\Omega)} + \|v\|^2_{L^2(\Omega)}\,. \tag{1.5}$$

1.2. The context and novelty of the main result. The fact that, for oscillatory data, the relative H^1_k error for the Helmholtz h-FEM is controllably small if $(hk)^{2p}C_{\rm sol}$ is sufficiently small was famously identified for 1-d nontrapping problems by the work of Ihlenburg and Babuška [25, 26] (see [26, Page 350, penultimate displayed equation], [24, Equation 4.7.41]). The bounds (1.2) and (1.3) have previously been obtained for

1. $p=1$, for Helmholtz problems with either an impedance boundary condition [37, Theorem 6.1], [1, Theorem 2], or truncation via the exact Dirichlet-to-Neumann map [27, Theorem 4.1], or truncation via a radial, k-independent PML [31, Theorem 4.4], or truncation via a radial, k-dependent PML [6, Theorem 7.2],
2. $p \in \mathbb{Z}^+$, the constant-coefficient Helmholtz equation with no obstacle and an impedance boundary condition approximating the radiation condition [13, Theorem 5.1],
3. $p \in \mathbb{Z}^+$, the variable-coefficient Helmholtz equation in the exterior of a Dirichlet obstacle with an impedance boundary condition approximating the radiation condition [36, Theorem 2.39].

The bounds in Point 1 for $p=1$ come from the so-called *elliptic projection* argument, which proves error bounds under the condition "$(hk)^{p+1}C_{\rm sol}$ is sufficiently small"; i.e., the sharp condition when $p=1$, but not when $p>1$. The initial ideas behind this argument were introduced in the Helmholtz context in [16, 17] for interior-penalty discontinuous Galerkin methods, and then further developed for the standard FEM and continuous interior-penalty methods in [37, 39].

The present paper proves the bounds (1.2), (1.3), and (1.4) for the h-FEM assuming only that the sesquilinear form is continuous, satisfies a Gårding inequality, and satisfies certain standard elliptic-regularity assumptions, therefore covering a variety of scatterers and methods for truncating the exterior domain. Regarding the latter: in this paper we consider truncating with the exact Dirichlet-to-Neumann map on the boundary of a ball, with a radial PML, or with an impedance boundary condition.

Since the preprint of this paper appeared, its ideas have been used in [38, 9], with [38] analysing high-order continuous interior-penalty methods for the Helmholtz PML problem (generalising the $p=1$ results in [31]) and [9] analysing the geometric error for the Helmholtz h-FEM.

1.3. Statement of the main result in abstract form. Let $\mathcal{H}\subset\mathcal{H}_0\subset\mathcal{H}^*$ be Hilbert spaces, with $\mathcal{H}^*$ the space of anti-linear functionals on $\mathcal{H}$, $\mathcal{H}_0$ identified with its dual, and $\mathcal{H}\subset\mathcal{H}_0$ compact. Let $a:\mathcal{H}\times\mathcal{H}\to\mathbb{C}$ be a sesquilinear form; i.e., a is linear in its first argument, anti-linear in its second argument. We assume that a is continuous, i.e.,

$$|a(u,v)| \le C_{\rm cont} \|u\|_{\mathcal{H}} \|v\|_{\mathcal{H}} \quad \text{ for all } u,v \in \mathcal{H}, \tag{1.6}$$

and satisfies the Gårding inequality

$$\Re a(v,v) \ge C_{\rm G1} \|v\|^2_{\mathcal{H}} - C_{\rm G2} \|v\|^2_{\mathcal{H}_0} \quad \text{ for all } v \in \mathcal{H}, \tag{1.7}$$

for some $C_{\rm cont}, C_{\rm G1}, C_{\rm G2} > 0$.

ASSUMPTION 1.2 (Abstract elliptic regularity bounds for a). *Let $\mathcal{Z}_0 = \mathcal{H}_0$, $\mathcal{Z}_1 = \mathcal{H}$, and $\mathcal{Z}_j \subset \mathcal{Z}_{j-1}$ for $j = 2, \ldots, \ell+1$ be such that $\mathcal{Z}_j$ is dense in $\mathcal{Z}_{j-1}$. There exists a $C_{\rm ell,1} > 0$ such that, for all $u \in \mathcal{H}$,*

$$\|u\|_{\mathcal{Z}_j} \le C_{\rm ell,1} \Big(\|u\|_{\mathcal{H}_0} + \sup_{v \in \mathcal{H},\, \|v\|_{(\mathcal{Z}_{j-2})^*}=1} \big| a(u,v) \big| \Big), \quad j = 2, \ldots, \ell+1, \tag{1.8}$$

with $u \in \mathcal{Z}_j$ if the right-hand side of (1.8) is finite. In addition, for all $u \in \mathcal{H}$,

$$\|u\|_{\mathcal{Z}_j} \le C_{\rm ell,1} \Big(\|u\|_{\mathcal{H}_0} + \sup_{v \in \mathcal{H},\, \|v\|_{(\mathcal{Z}_{j-2})^*}=1} \big| (\Re a)(u,v) \big| \Big), \quad j = 2, \ldots, \ell+1, \tag{1.9}$$

with $u \in \mathcal{Z}_j$ if the right-hand side of (1.9) is finite, where the sesquilinear form $\Re a$ is defined by

$$(\Re a)(u,v) := \tfrac{1}{2} \big(a(u,v) + \overline{a(v,u)} \big). \tag{1.10}$$

Recalling the one-to-one correspondence between sesquilinear forms $a : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ and operators $\mathcal{A} : \mathcal{H} \to \mathcal{H}^*$ given by $a(u,v) = \langle \mathcal{A}u, v\rangle_{\mathcal{H}^* \times \mathcal{H}}$ (see, e.g., [32, Page 42]), we see that $\Re a$ (1.10) is the sesquilinear form corresponding to the operator $\Re \mathcal{A} := (\mathcal{A} + \mathcal{A}^*)/2$.

REMARK 1.3. *Note that $\Re a$ in (1.7) and (1.10) could be replaced by $\Re({\rm e}^{{\rm i}\omega} a)$, so long as one uses the same value of ω in both conditions. Remark 4.4 below describes a situation where this is useful.*

EXAMPLE 1.4. *For the Helmholtz equation outside a Dirichlet obstacle with radial PML truncation and Ω the truncated exterior domain, $\mathcal{H}_0 = L^2(\Omega)$, $\mathcal{H} = H^1_0(\Omega)$, and $\mathcal{Z}_j = H^j(\Omega) \cap H^1_0(\Omega)$. Assumption 1.2 is then elliptic regularity for the Helmholtz PML operator and its real part, which both hold if the coefficients of the Helmholtz equation are in $C^{\ell-1,1}$, the PML scaling function is $C^{\ell,1}$, $\partial\Omega$ is $C^{\ell,1}$ and one works with the Sobolev norms where each derivative is scaled by k^{-1} (see Lemma 4.7 below).*

Given $g \in \mathcal{H}^*$, suppose that $u \in \mathcal{H}$ satisfies

$$a(u,v) = \langle g, v \rangle \quad \text{ for all } v \in \mathcal{H}, \tag{1.11}$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing between $\mathcal{H}^*$ and $\mathcal{H}$. Since a is continuous (1.6) and satisfies the Gårding inequality (1.7) with $\mathcal{H} \subset \mathcal{H}_0$ compact, uniqueness of the solution to (1.11) is equivalent to existence; see, e.g., [32, Theorem 2.32].

Given a sequence of finite dimensional subspace $\{\mathcal{H}_h\}_{h>0}$ with $\mathcal{H}_h \subset \mathcal{H}$, the Galerkin method seeks approximations of u, $\{u_h\}_{h>0}$ with $u_h \in \mathcal{H}_h$, such that

$$a(u_h, v_h) = \langle g, v_h \rangle \quad \text{for all } \; v_h \in \mathcal{H}_h. \tag{1.12}$$

THEOREM 1.5 (Abstract generalisation of the elliptic-projection argument).

Let $a : \mathcal{H}\times\mathcal{H} \to \mathbb{C}$ satisfy (1.6), (1.7), and Assumption 1.2 for some $\ell \in \mathbb{Z}^+$, and suppose that (1.11) has a unique solution. Define $\mathcal{R}^* : \mathcal{H}^* \to \mathcal{H}$ by

$$a(w, \mathcal{R}^* v) = \langle w, v\rangle \qquad \text{for all } w \in \mathcal{H},\ v \in \mathcal{H}^*, \tag{1.13}$$

and

$$\eta(\mathcal{H}_h) := \|(I-\Pi_h)\mathcal{R}^*\|_{\mathcal{H}_0\to\mathcal{H}}, \tag{1.14}$$

where $\Pi_h : \mathcal{H} \to \mathcal{H}_h$ is the orthogonal projection. Then there exist $C_1, C_2, C_3 > 0$ such that if h satisfies

$$\eta(\mathcal{H}_h)\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}} \le C_1, \tag{1.15}$$

then the solution u_h to (1.12) exists, is unique, and satisfies

$$\|u-u_h\|_{\mathcal{H}} \le C_2\big(1+\eta(\mathcal{H}_h)\big)\,\|(I-\Pi_h)u\|_{\mathcal{H}}\,, \tag{1.16}$$

$$\|u-u_h\|_{\mathcal{H}_0} \le C_3\,\eta(\mathcal{H}_h)\,\|(I-\Pi_h)u\|_{\mathcal{H}}\,, \tag{1.17}$$

In addition, for all $C_{\rm osc} > 0$ there exists $C_4 > 0$ such that if

$$\|g\|_{\mathcal{Z}_{\ell-1}} \le C_{\rm osc}\,\|g\|_{\mathcal{H}^*} \tag{1.18}$$

and h satisfies (1.15) then

$$\frac{\|u-u_h\|_{\mathcal{H}}}{\|u\|_{\mathcal{H}}} \le C_4\big(1+\eta(\mathcal{H}_h)\big)\,\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\,; \tag{1.19}$$

i.e., the relative error in $\mathcal{H}$ can be made controllably small by making $\eta(\mathcal{H}_h)\,\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}$ sufficiently small.

By the order of quantifiers in Theorem 1.5, C_1, C_2, and C_3 depend only on $C_{\rm cont}, C_{\rm G1}, C_{\rm G2}, C_{\rm ell,1}$, and ℓ, and C_4 depends only on $C_{\rm cont}, C_{\rm G1}, C_{\rm G2}, C_{\rm ell,1}$, ℓ, and $C_{\rm osc}$.

The bound (1.16) implies the result that the sequence of Galerkin solutions are quasioptimal if $\eta(\mathcal{H}_h)$ is sufficiently small – with this the so-called *asymptotic regime*. Theorem 1.5, however, holds under the less-restrictive condition that $\eta(\mathcal{H}_h)\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}$ be small (i.e., the condition (1.15)), and so is valid in part of the preasymptotic regime.

The bounds (1.16), (1.17), and (1.19) and the meshthreshold (1.15) all involve the quantity $\eta(\mathcal{H}_h)$, which measures how well solutions of the adjoint problem are approximated in the space $\mathcal{H}_h$. Bounds on $\eta(\mathcal{H}_h)$ are given in [33, 34, 14, 7, 28, 19, 20, 2]. The following bound on $\eta(\mathcal{H}_h)$ is essentially the one in [7], although our proof is different. We include this bound here both for completeness, and because, after establishing the key intermediate result for proving Theorem 1.5 (Lemma 2.1 below) our proof of the bound on $\eta(\mathcal{H}_h)$ is very short (see §2.3 below).

ASSUMPTION 1.6 (Elliptic regularity assumption for the adjoint sesquilinear form). *With $\mathcal{Z}_j$, $j = 0,\ldots,\ell+1$, as in Assumption 1.2, there exists a $C_{\rm ell,2} > 0$ such*

$$\|v\|_{\mathcal{Z}_j} \le C_{\rm ell,2}\Big(\|v\|_{\mathcal{H}_0} + \sup_{u\in\mathcal{H},\,\|u\|_{(\mathcal{Z}_{j-2})^*}=1} |a(u,v)|\Big), \quad j = 2,\ldots,\ell+1, \tag{1.20}$$

with $v \in \mathcal{Z}_j$ if the right-hand side of (1.20) is finite.

THEOREM 1.7 (Bound on $\eta(\mathcal{H}_h)$). *Suppose that the assumptions of Theorem 1.5 hold and, additionally, Assumption 1.6 holds. With $\mathcal{R}^*$ and $\eta(\mathcal{H}_h)$ defined by (1.13) and (1.14) respectively, there exists $C > 0$ such that*

$$\eta(\mathcal{H}_h) \leq C\Big(\|(I - \Pi_h)\|_{\mathcal{Z}_2 \to \mathcal{H}} + \|(I - \Pi_h)\|_{\mathcal{Z}_{\ell+1} \to \mathcal{H}} \|\mathcal{R}^*\|_{\mathcal{H}_0 \to \mathcal{H}_0} \Big). \tag{1.21}$$

EXAMPLE 1.8. *In §4 and §5 below we show how Helmholtz problems with the radiation condition* either *realised by the exact Dirichlet-to-Neumann map on the boundary of a ball* or *approximated by either a radial PML or an impedance boundary condition, respectively, fit into the abstract framework of Theorems 1.5 and 1.7. In both these cases, the norm of the adjoint solution operator, i.e., $\|\mathcal{R}^*\|_{\mathcal{H}_0 \to \mathcal{H}_0}$, is the same as the norm of the solution operator of the original (non-adjoint) problem, which we denote by $C_{\rm sol}$. Furthermore, with $\{\mathcal{H}_h\}_{h>0}$ corresponding to the standard finite-element spaces of piecewise degree-p polynomials on shape-regular simplicial triangulations, indexed by the meshwidth h,*

$$\|(I - \Pi_h)\|_{\mathcal{Z}_{m+1} \to \mathcal{H}} \leq C(hk)^m \quad \text{for } 0 \leq m \leq p. \tag{1.22}$$

The meshthreshold (1.15) then becomes that $(hk)^{2\ell} C_{\rm sol}$ is sufficiently small when $\ell \leq p$. Recall that ℓ is a parameter in the elliptic-regularity assumptions (Assumptions 1.2 and 1.6). If the polynomial degree p is taken to be ℓ then (1.15) becomes (1.1); the bounds (1.16) and (1.17) then become (1.2) and (1.3), respectively.

2. Proofs of the main results (Theorems 1.5 and 1.7).

2.1. Construction of a regularizing operator that produces coercivity when added to a.

LEMMA 2.1. *Suppose that $a : \mathcal{H} \times \mathcal{H} \to \mathbb{C}$ satisfies (1.6), (1.7), and Assumption 1.2 for some $\ell \in \mathbb{Z}^+$. Then there exists $S : \mathcal{H}_0 \to \mathcal{H}_0$ self adjoint and $c, C > 0$ such that, with*

$$\widetilde{a}(u, v) := a(u, v) + \langle Su, v\rangle_{\mathcal{H}_0}, \tag{2.1}$$

$$\Re \widetilde{a}(v, v) \geq c\, \|v\|^2_{\mathcal{H}} \quad \text{for all } v \in \mathcal{H}, \tag{2.2}$$

$$\|S\|_{\mathcal{H}_0 \to \mathcal{Z}_j} \leq C, \qquad j = 0, \ldots, \ell - 1, \tag{2.3}$$

and $\widetilde{\mathcal{R}} : \mathcal{H}^* \to \mathcal{H}$ defined by

$$\widetilde{a}(\widetilde{\mathcal{R}} f, v) = \langle f, v\rangle \quad \text{for all } v \in \mathcal{H},\ f \in \mathcal{H}^*, \tag{2.4}$$

is well defined with

$$\|\widetilde{\mathcal{R}}\|_{\mathcal{Z}_{j-2} \to \mathcal{Z}_j} \leq C, \qquad 2 \leq j \leq \ell + 1. \tag{2.5}$$

REMARK 2.2 (Relation with the original elliptic-projection argument). *The original elliptic-projection argument [16, 17] uses coercivity of (2.1) with S a sufficiently large multiple of the identity (see [16, Lemma 5.1], [17, Lemma 4.1]). This particular S satisfies Lemma 2.1 with $\ell = 1$; the threshold (1.15) in Theorem 1.5 is then "$\eta(\mathcal{H}_h)\|I - \Pi_h\|_{\mathcal{Z}_2 \to \mathcal{H}}$ sufficiently small", which (by (1.21) and (1.22)) becomes the condition "$(hk)^{\ell+1} C_{\rm sol}$ sufficiently small" discussed in §1.2.*

The proof of Lemma 2.1 uses the spectral theorem for bounded self-adjoint operators, $B : \mathcal{H} \to \mathcal{H}^*$, which we recap here. With $\mathcal{H}_0$ and $\mathcal{H}$ as in §1.3, let b be a sesquilinear form on $\mathcal{H}$ satisfying $b(u,v) = \overline{b(v,u)}$, with associated operator B; i.e., $b(u,v) = \langle Bu, v\rangle$ for all $u,v \in \mathcal{H}$. If b satisfies the Gårding inequality (1.7) (with a replaced by b) then there exist an orthonormal basis (in $\mathcal{H}_0$) of eigenfunctions of B, $\{\phi_j\}_{j=1}^\infty$, with associated eigenvalues satisfying $\lambda_1 \leq \lambda_2 \leq \ldots$ with $\lambda_j \to \infty$ as $j \to \infty$. Furthermore, for all $u \in \mathcal{H}$,

$$Bu = \sum_{j=1}^{\infty} \lambda_j \langle \phi_j, u\rangle \phi_j \tag{2.6}$$

(where the sum converges in $\mathcal{H}^*$); see, e.g., [32, Theorem 2.37]. Given a bounded function f, we define $f(B) : \mathcal{H}_0 \to \mathcal{H}_0$ by

$$f(B)u := \sum_{j=1}^{\infty} f(\lambda_j)\langle \phi_j, u\rangle \phi_j, \quad \text{so that} \quad \|f(B)\|_{\mathcal{H}_0\to\mathcal{H}_0} \leq \sup_{\lambda\in[\lambda_1,\infty)} |f(\lambda)|. \tag{2.7}$$

Proof of Lemma 2.1. Let $\mathcal{P} : \mathcal{H} \to \mathcal{H}^*$ be the operator associated with the sesquilinear form $\Re a$ defined by (1.10), i.e., $(\Re a)(u,v) = \langle \mathcal{P}u, v\rangle$ for all $u,v \in \mathcal{H}$; observe that $\mathcal{P}$ is self-adjoint. Since $(\Re a)$ also satisfies the Gårding equality satisfied by a (1.7), the spectral theorem recapped above applies. Let $\{\lambda_j\}_{j=1}^\infty$ be the eigenvalues of $\mathcal{P}$ with $\lambda_j \leq \lambda_{j+1}$, $j = 1, \ldots,$ let $\psi \in C^\infty_{\rm comp}(\mathbb{R}; [0,\infty))$ be such that

$$x + \psi(x) \geq 1 \quad \text{for } x \geq \lambda_1, \tag{2.8}$$

and let $S := \psi(\mathcal{P})$, in the sense of (2.7).

We now use (1.9) to prove that $S : \mathcal{H}_0 \to \mathcal{Z}_j$ satisfies (2.3). Since ψ has compact support, the function $t \mapsto t^m\psi(t)$ is bounded for any $m \geq 0$. Thus (2.7) implies that, for any $m \geq 0$,

$$\|\mathcal{P}^m \psi(\mathcal{P})\|_{\mathcal{H}_0\to\mathcal{H}_0} \leq C_m. \tag{2.9}$$

By (1.9),

$$\|\psi(\mathcal{P})\|_{\mathcal{H}_0\to\mathcal{Z}_j} \leq C_{\rm ell,1}\Big(\|\psi(\mathcal{P})\|_{\mathcal{H}_0\to\mathcal{H}_0} + \|\mathcal{P}\psi(\mathcal{P})\|_{\mathcal{H}_0\to\mathcal{Z}_{j-2}}\Big), \quad j = 2, \ldots, \ell+1,$$

so that, by induction and (2.9),

$$\|S\|_{\mathcal{H}_0\to\mathcal{Z}_{\ell-1}} = \|\psi(\mathcal{P})\|_{\mathcal{H}_0\to\mathcal{Z}_{\ell-1}} \leq C_\ell \sum_{j=0}^{\lceil(\ell-1)/2\rceil} \big\|\mathcal{P}^j\psi(\mathcal{P})\big\|_{\mathcal{H}_0\to\mathcal{H}_0} \leq C_\ell.$$

We now show that $\widetilde{a}$ satisfies (2.2). By the definitions of $\mathcal{P}$ and S, the eigenfunction expansions (2.6) and (2.7), and the inequality (2.8), for all $v \in \mathcal{H}$,

$$\Re\widetilde{a}(v,v) = \Re a(v,v) + \langle \psi(\mathcal{P})v, v\rangle = \langle (\mathcal{P} + \psi(\mathcal{P}))v, v\rangle \geq \|v\|^2_{\mathcal{H}_0}.$$

Since $\psi \geq 0$, S is positive, and thus $\Re\widetilde{a}(v,v) \geq \Re a(v,v)$ for all $v \in \mathcal{H}$, for any $\epsilon > 0$ and for all $v \in \mathcal{H}$,

$$\Re\widetilde{a}(v,v) \geq \epsilon\Re a(v,v) + (1-\epsilon)\Re\widetilde{a}(v,v) \geq \epsilon C_{\rm G1}\|v\|^2_{\mathcal{H}} - C_{\rm G2}\epsilon\|v\|^2_{\mathcal{H}_0} + (1-\epsilon)\|v\|^2_{\mathcal{H}_0}.$$

Therefore, if $\epsilon = (1 + C_{\rm G2})^{-1}$, then

$$\Re\widetilde{a}(v,v) \geq C_{\rm G1}(1 + C_{\rm G2})^{-1}\|v\|^2_{\mathcal{H}};$$

i.e., $\widetilde{a}$ is coercive. The existence of $\widetilde{\mathcal{R}} : \mathcal{H}^* \to \mathcal{H}$ satisfying (2.4) and $\|\widetilde{\mathcal{R}}\|_{\mathcal{H}^*\to\mathcal{H}} \le C$ then follows from the Lax–Milgram lemma (see, e.g., [32, Lemma 2.32]). Finally, to see that

$$\|\widetilde{\mathcal{R}}\|_{\mathcal{Z}_{j-2}\to\mathcal{Z}_j} \le C, \qquad 2 \le j \le \ell+1,$$

observe that, since S is self-adjoint and satisfies (2.3), for $v \in (\mathcal{Z}_{j-2})^*$,

$$\begin{aligned}|a(\widetilde{\mathcal{R}}g, v)| = |\widetilde{a}(\widetilde{\mathcal{R}}g, v) - \langle S\widetilde{\mathcal{R}}g, v\rangle| &\le |\widetilde{a}(\widetilde{\mathcal{R}}g, v)| + |\langle S\widetilde{\mathcal{R}}g, v\rangle| \\ &\le |\langle v, g\rangle| + \|v\|_{(\mathcal{Z}_{j-2})^*}\|S\|_{\mathcal{H}\to\mathcal{Z}_{j-2}}\|(\widetilde{\mathcal{R}})^*\|_{\mathcal{H}^*\to\mathcal{H}}\|g\|_{\mathcal{H}^*} \\ &\le \|v\|_{(\mathcal{Z}_{j-2})^*}(\|g\|_{\mathcal{Z}_{j-2}} + C\|g\|_{\mathcal{H}^*}),\end{aligned}$$

and the claim follows from (1.8). ∎

2.2. Proof Theorem 1.5 using Lemma 2.1. By, e.g., [32, Theorem 2.34], the operator associated to the sesquilinear form a is Fredholm of index zero. The fact that the solution to (1.11) is unique therefore implies that the solution exists, and also implies that $\mathcal{R}^*$ is well-defined (by, e.g., [32, Theorem 2.27]).

We claim it is sufficient to prove the bounds (1.16) and (1.17) under the assumption of existence. Indeed, by uniqueness of the variational problem (1.11), either of the bounds (1.16) or (1.17) under the assumption of existence implies uniqueness of u_h, and uniqueness implies existence for the finite-dimensional Galerkin linear system.

We next show that the bound (1.16) follows from (1.17). In the rest of the proof, C denotes a constant (whose value may change from line to line) that only depends on $C_{\rm cont}, C_{\rm G1}, C_{\rm G2}, C_{\rm ell,1}$, and ℓ. By the Gårding inequality (1.7), Galerkin orthogonality

$$a(u-u_h, v_h) = 0 \quad \text{for all } v_h \in \mathcal{H}_h, \tag{2.10}$$

and (1.17), for any $v_h \in \mathcal{H}_h$,

$$\|u-u_h\|^2_{\mathcal{H}} \le C\Big[\big|a(u-u_h, u-v_h)\big| + \|u-u_h\|^2_{\mathcal{H}_0}\Big] \tag{2.11}$$

$$\le C\Big[\,\|u-u_h\|_{\mathcal{H}}\,\|u-v_h\|_{\mathcal{H}} + \Big(\eta(\mathcal{H}_h)\,\|(I-\Pi_h)u\|_{\mathcal{H}}\Big)^2\Big]. \tag{2.12}$$

The bound (1.16) on the error in $\mathcal{H}$ then follows by using the inequality $2ab \le \epsilon a^2 + b^2/\epsilon$ for all $a, b, \epsilon > 0$ in the first term on the right-hand side of (2.12), and then using the inequality $a^2 + b^2 \le (a+b)^2$ for $a, b > 0$.

We now prove (1.17). By the definition of $\mathcal{R}^*$, Galerkin orthogonality (2.10), and the definition of $\widetilde{a}$ (2.1)

$$\begin{aligned}\|u-u_h\|^2_{\mathcal{H}_0} &= a\big(u-u_h, \mathcal{R}^*(u-u_h)\big) = a\big(u-u_h, \mathcal{R}^*(u-u_h) - v_h\big) \\ &= \widetilde{a}\big(u-u_h, \mathcal{R}^*(u-u_h) - v_h\big) - \big\langle S(u-u_h), \mathcal{R}^*(u-u_h) - v_h\big\rangle_{\mathcal{H}_0}.\end{aligned} \tag{2.13}$$

Let $\widetilde{\Pi}_h : \mathcal{H} \to \mathcal{H}_h$ be the solution of the variational problem

$$\widetilde{a}(w_h, \widetilde{\Pi}_h v) = \widetilde{a}(w_h, v) \quad \text{for all } w_h \in \mathcal{H}_h.$$

Since $\widetilde{a}$ is continuous (by (1.6) and (2.3)) and coercive (by (2.2)), by the Lax–Milgram lemma and Céa's lemma (see, e.g., [4, Theorem 2.8.1]), $\widetilde{\Pi}_h$ is well-defined with

$$\big\|(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}} \le C\,\|(I-\Pi_h)v\|_{\mathcal{H}}\,. \tag{2.14}$$

The definition of $\widetilde{\Pi}_h$ implies the Galerkin orthogonality

$$\widetilde{a}\big(w_h, (I-\widetilde{\Pi}_h)u\big) = 0 \quad \text{for all } w_h \in \mathcal{H}_h. \tag{2.15}$$

We now choose $v_h = \widetilde{\Pi}_h \mathcal{R}^*(u-u_h)$ in (2.13) so that, by (2.15),

$$\begin{aligned}
&\|u-u_h\|^2_{\mathcal{H}_0} \\
&= \widetilde{a}\big((I-\Pi_h)u, (I-\widetilde{\Pi}_h)\mathcal{R}^*(u-u_h)\big) - \big\langle u-u_h, S^*(I-\widetilde{\Pi}_h)\mathcal{R}^*(u-u_h)\big\rangle_{\mathcal{H}_0} \\
&\le C\,\|(I-\Pi_h)u\|_{\mathcal{H}}\,\big\|(I-\widetilde{\Pi}_h)\mathcal{R}^*(u-u_h)\big\|_{\mathcal{H}} + \|u-u_h\|_{\mathcal{H}_0}\,\big\|S^*(I-\widetilde{\Pi}_h)\mathcal{R}^*(u-u_h)\big\|_{\mathcal{H}_0}.
\end{aligned} \tag{2.16}$$

By (2.14) and the definition of $\eta(\mathcal{H}_h)$ (1.14),

$$\big\|(I-\widetilde{\Pi}_h)\mathcal{R}^*(u-u_h)\big\|_{\mathcal{H}} \le C\,\|(I-\Pi_h)\mathcal{R}^*(u-u_h)\|_{\mathcal{H}} \le C\eta(\mathcal{H}_h)\,\|u-u_h\|_{\mathcal{H}_0}. \tag{2.17}$$

We now claim that the bound (1.17) follows if we can prove that, for all $v \in \mathcal{H}$,

$$\big\|S^*(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}_0} \le C\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\big\|(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}}. \tag{2.18}$$

Indeed, we use (2.18), with $v = \mathcal{R}^*(u-u_h)$, in the second term on the right-hand side of (2.16) to obtain

$$\begin{aligned}
\|u-u_h\|^2_{\mathcal{H}_0} &\le C\,\|(I-\Pi_h)u\|_{\mathcal{H}}\,\big\|(I-\widetilde{\Pi}_h)\mathcal{R}^*(u-u_h)\big\|_{\mathcal{H}} \\
&\quad + C\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\,\|u-u_h\|_{\mathcal{H}_0}\,\big\|(I-\widetilde{\Pi}_h)\mathcal{R}^*(u-u_h)\big\|_{\mathcal{H}}.
\end{aligned} \tag{2.19}$$

We then use (2.17) in both terms on the right-hand side of (2.19) to obtain

$$\|u-u_h\|^2_{\mathcal{H}_0} \le C\eta(\mathcal{H}_h)\,\|(I-\Pi_h)u\|_{\mathcal{H}}\,\|u-u_h\|_{\mathcal{H}_0} + C\eta(\mathcal{H}_h)\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\,\|u-u_h\|^2_{\mathcal{H}_0},$$

from which (1.17), under the condition (1.15), follows.

We now prove (2.18) by using a standard duality argument. By the definition of $\widetilde{\mathcal{R}}$ (2.4) and Galerkin orthogonality (2.15),

$$\begin{aligned}
\big\|S^*(I-\widetilde{\Pi}_h)v\big\|^2_{\mathcal{H}_0} = \big\langle SS^*(I-\widetilde{\Pi}_h)v, (I-\widetilde{\Pi}_h)v\big\rangle_{\mathcal{H}_0} &= \widetilde{a}\big(\widetilde{\mathcal{R}}SS^*(I-\widetilde{\Pi}_h)v - w_h, (I-\widetilde{\Pi}_h)v\big) \\
&= \widetilde{a}\big((I-\Pi_h)\widetilde{\mathcal{R}}SS^*(I-\widetilde{\Pi}_h)v, (I-\widetilde{\Pi}_h)v\big).
\end{aligned}$$

Then, by continuity of $\widetilde{a}$ and the bounds (2.5) and (2.3),

$$\begin{aligned}
\big\|S^*(I-\widetilde{\Pi}_h)v\big\|^2_{\mathcal{H}_0} &\le C\big\|(I-\Pi_h)\widetilde{\mathcal{R}}SS^*(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}}\big\|(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}} \\
&\le \|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\big\|\widetilde{\mathcal{R}}SS^*(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{Z}_{\ell+1}}\big\|(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}}, \\
&\le C\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\big\|SS^*(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{Z}_{\ell-1}}\big\|(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}}, \\
&\le C\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\big\|S^*(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}_0}\big\|(I-\widetilde{\Pi}_h)v\big\|_{\mathcal{H}}
\end{aligned}$$

which implies the bound (2.18), and hence (1.17).

Finally, we prove (1.19). By (1.11), (1.18), and the abstract elliptic-regularity assumption (1.8), $u \in \mathcal{Z}_{\ell+1}$ with

$$\|u\|_{\mathcal{Z}_{\ell+1}} \le C\big(\,\|u\|_{\mathcal{H}_0} + \|g\|_{\mathcal{Z}_{\ell-1}}\big) \le C\big(\,\|u\|_{\mathcal{H}_0} + \|g\|_{\mathcal{H}^*}\big).$$

The variational problem (1.11) implies that

$$\|g\|_{\mathcal{H}^*} = \sup_{v\in\mathcal{H}, v\neq 0} \frac{|a(u,v)|}{\|v\|_{\mathcal{H}}} \leq C \|u\|_{\mathcal{H}},$$

and thus $\|u\|_{\mathcal{Z}_{\ell+1}} \leq C\|u\|_{\mathcal{H}}$. The bound (1.16) then implies that

$$\|u-u_h\|_{\mathcal{H}} \leq C_2\big(1+\eta(\mathcal{H}_h)\big)\,\|I-\Pi_h\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\,\|u\|_{\mathcal{Z}_{\ell+1}}$$

and (1.19) follows.

2.3. Proof of Theorem 1.7. Let $\widetilde{\mathcal{R}}^* : \mathcal{H}^* \to \mathcal{H}$ be defined by

$$\widetilde{a}(u, \widetilde{\mathcal{R}}^* f) = \langle u, f\rangle \quad \text{for all } u\in\mathcal{H},\ f\in\mathcal{H}^* \tag{2.20}$$

(compare to (2.4)); i.e., $\widetilde{\mathcal{R}}^*$ is the adjoint solution operator for the sesquilinear form $\widetilde{a}$. Repeating the proof of (2.5) with the elliptic-regularity assumption (1.8) replaced by (1.20), we see that

$$\|\widetilde{\mathcal{R}}^*\|_{\mathcal{Z}_{j-2}\to\mathcal{Z}_j} \leq C, \qquad 2\leq j\leq \ell+1. \tag{2.21}$$

We now claim that

$$\mathcal{R}^* = \widetilde{\mathcal{R}}^*(I+S\mathcal{R}^*). \tag{2.22}$$

Indeed, by the definitions of $\widetilde{a}$ (2.1) and $\mathcal{R}^*$ (1.13) the fact that $S=S^*$, and the definition of $\widetilde{\mathcal{R}}^*$ (2.20)

$$\begin{aligned}\widetilde{a}(u,\mathcal{R}^* f) = a(u,\mathcal{R}^* f) + \langle Su, \mathcal{R}^* f\rangle = \langle u, f\rangle + \langle u, S\mathcal{R}^* f\rangle &= \langle u, (I+S\mathcal{R}^*)f\rangle \\ &= \widetilde{a}\big(u, \widetilde{\mathcal{R}}^*(I+S\mathcal{R}^*)f\big)\end{aligned}$$

for all $u\in\mathcal{H}$ and $f\in\mathcal{H}^*$. The expression (2.22) then follows from coercivity of $\widetilde{a}$ (2.2). Then, by (2.22) and the mapping properties (2.21) and (2.3) of $\widetilde{\mathcal{R}}^*$ and S, respectively,

$$\begin{aligned}\|(I-\Pi_h)\mathcal{R}^* g\|_{\mathcal{H}} &\leq C\Big(\|(I-\Pi_h)\widetilde{\mathcal{R}}^* g\|_{\mathcal{H}} + \|(I-\Pi_h)\widetilde{\mathcal{R}}^* S\mathcal{R}^* g\|_{\mathcal{H}}\Big)\\ &\leq C\Big(\|(I-\Pi_h)\|_{\mathcal{Z}_2\to\mathcal{H}}\|\widetilde{\mathcal{R}}^*\|_{\mathcal{H}_0\to\mathcal{Z}_2}\|g\|_{\mathcal{H}_0}\\ &\qquad +\|(I-\Pi_h)\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\|\widetilde{\mathcal{R}}^*\|_{\mathcal{Z}_{\ell-1}\to\mathcal{Z}_{\ell+1}}\|S\|_{\mathcal{H}_0\to\mathcal{Z}_{\ell-1}}\|\mathcal{R}^*\|_{\mathcal{H}_0\to\mathcal{H}_0}\|g\|_{\mathcal{H}_0}\Big)\\ &\leq C\Big(\|(I-\Pi_h)\|_{\mathcal{Z}_2\to\mathcal{H}}\|g\|_{\mathcal{H}_0} + \|(I-\Pi_h)\|_{\mathcal{Z}_{\ell+1}\to\mathcal{H}}\|\mathcal{R}^*\|_{\mathcal{H}_0\to\mathcal{H}_0}\|g\|_{\mathcal{H}_0}\Big);\end{aligned}$$

the result (1.21) then follows from the definition of $\eta(\mathcal{H}_h)$ (1.14).

REMARK 2.3 (The splitting (2.22)). *Recalling the mapping properties of $\widetilde{\mathcal{R}}^*$ (2.21) and S (2.3), we see that (2.22) splits the Helmholtz adjoint solution into a part with finite regularity but norm bounded independently of $\|\mathcal{R}^*\|_{\mathcal{H}_0\to\mathcal{H}_0}$ – namely $\widetilde{\mathcal{R}}^*$ – and a part with the highest possible regularity allowed by the coefficients and the domain (via Assumptions 1.2 and 1.6) and norm bounded by $\|\mathcal{R}^*\|_{\mathcal{H}_0\to\mathcal{H}_0}$ – namely $\widetilde{\mathcal{R}}^* S\mathcal{R}^*$.*

We highlight that such a splitting was previously achieved in the following papers.

- *[33, 34, 14, 28, 19, 20, 2], which split the solution into "high-" and "low-" frequency components; in these papers the scatterer (either an impenetrable obstacle or variable coefficients) is assumed to be analytic, and thus the "low"-frequency part of the solution is analytic (at least near the scatterer).*
- *[7], which worked under elliptic-regularity assumptions analogous to Assumptions 1.2 and 1.6, and expanded the Helmholtz solution in a series whose terms increase with regularity, with the remainder have the highest possible regularity and being norm-bounded by $\|\mathcal{R}^*\|_{\mathcal{H}_0\to\mathcal{H}_0}$.*

3. Elliptic-regularity results. This section collects the elliptic-regularity results that are used to verify that Assumption 1.2 holds for Helmholtz problems with truncation of the exterior domain either by a radial PML (in §4) or an impedance boundary condition/exact Dirichlet-to-Neumann map (in §5). Let

$$\mathcal{L}u = -k^{-2}\nabla\cdot(A\nabla u) - c^{-2}u, \tag{3.1}$$

with associated sesquilinear form

$$a(u,v) = \int_\Omega \Big(k^{-2}(A\nabla u)\cdot\overline{\nabla v} - c^{-2}u\,\overline{v}\Big),$$

where Ω be a bounded Lipschitz domain with outward-pointing unit normal vector n. The conormal derivative $\partial_{n,A}u$ is defined for $u\in H^2(\Omega)$ by $\partial_{n,A}u := n\cdot(A\nabla u)$; recall that $\partial_{n,A}u$ can be defined for $u\in H^1(\Omega)$ with $\mathcal{L}u\in L^2(\Omega)$ by Green's identity; see, e.g., [32, Lemma 4.3].

Let $\|\cdot\|_{H^1_k}$ be defined by (1.5) and define higher-order weighted Sobolev norms by

$$\|v\|^2_{H^m_k(\Omega)} := \sum_{0\le|\alpha|\le m} k^{-2|\alpha|}\,\|\partial^\alpha v\|^2_{L^2(\Omega)}. \tag{3.2}$$

The rationale for using these norms is that if a function v oscillates with frequency k, then $|(k^{-1}\partial)^\alpha v|\sim|v|$ for all α; this is true, e.g., if $v(x)=\exp(ikx\cdot a)$. We highlight that many papers on the FEM applied to the Helmholtz equation use the weighted H^1 norm $|||v|||^2 := \|\nabla v\|^2_{L^2(\Omega)} + k^2\|v\|^2_{L^2(\Omega)}$; we work with (1.5)/(3.2) instead, because weighting the jth derivative with k^{-j} is easier to keep track of than weighting the jth derivative with k^{-j+1}.

ASSUMPTION 3.1. *For all $x\in\Omega$, $A_{j\ell}(x)=A_{\ell j}(x)$ and*

$$\Re\big(A(x)\xi,\xi\big)_2 = \Re\sum_{j=1}^d\sum_{\ell=1}^d A_{j\ell}(x)\xi_\ell\overline{\xi_j} \ge c|\xi|^2 \quad \text{for all } \xi\in\mathbb{C}^d.$$

THEOREM 3.2 (Local elliptic regularity near a Dirichlet or Neumann boundary). *Let Ω be a Lipschitz domain and let G_1, G_2 be open subsets of $\mathbb{R}^d$ with $G_1\Subset G_2$ and $G_1\cap\partial\Omega\neq\emptyset$. Let*

$$\Omega_j := G_j\cap\Omega,\ j=1,2, \qquad \text{and} \qquad \Gamma_2 := G_2\cap\partial\Omega. \tag{3.3}$$

Suppose that A satisfies Assumption 3.1, $A,c\in C^{m,1}(\overline{\Omega_2})$, and $\Gamma_2\in C^{m+1,1}$ for some $m\in\mathbb{N}$. Given $k_0>0$, there exists $C>0$ such that if $k\ge k_0$, $u\in H^1(\Omega_2)$, $\mathcal{L}u\in H^m(\Omega_2)$, and either $u=0$ or $\partial_{n,A}u=0$ on Γ_2, then

$$\|u\|_{H^{m+2}_k(\Omega_1)} \le C\big(\|u\|_{H^1_k(\Omega_2)} + \|\mathcal{L}u\|_{H^m_k(\Omega_2)}\big). \tag{3.4}$$

Proof. In unweighted norms, this follows from, e.g., [32, Theorems 4.7 and 4.16]; the proof in the weighted norms (3.2) is very similar. ∎

THEOREM 3.3 (Local elliptic regularity for the transmission problem). *Let $\Omega_{\rm in}$ be a Lipschitz domain, and let $\Omega_{\rm out} := \mathbb{R}^d \setminus \overline{\Omega_{\rm in}}$. Let G_1, G_2 be open subsets of $\mathbb{R}^d$ with $G_1 \Subset G_2$ and $G_1 \cap \partial\Omega_{\rm in} \neq \emptyset$. Let*

$$\Omega_{\rm in/out,j} := G_j \cap \Omega_{\rm in/out}, \quad j = 1, 2, \qquad \textit{and } \Gamma_2 := G_2 \cap \partial\Omega_{\rm in}.$$

Suppose that A satisfies Assumption 3.1, $A|_{\Omega_{\rm in/out,2}}, c|_{\Omega_{\rm in/out,2}} \in C^{m,1}(\overline{\Omega_{\rm in/out,2}})$, and $\Gamma_2 \in C^{m+1,1}$ for some $m \in \mathbb{N}$. Given $k_0 > 0$, there exists $C > 0$ such that if $k \geq k_0$, $u_{\rm in/out} \in H^1(\Omega_{\rm in/out})$, $\mathcal{L}u \in H^m(\Omega_{\rm in/out,2})$, and $u_{\rm in} = u_{\rm out}$ and $\partial_{n,A} u_{\rm in} = \zeta \partial_{n,A} u_{\rm out}$ on Γ_2 for some $\zeta > 0$, then

$$\begin{aligned}&\|u_{\rm in}\|_{H_k^{m+2}(\Omega_{\rm in,1})} + \|u_{\rm out}\|_{H_k^{m+2}(\Omega_{\rm out,1})}\\ &\quad\leq C\Big(\|u_{\rm in}\|_{H_k^1(\Omega_{\rm in,2})} + \|u_{\rm out}\|_{H_k^1(\Omega_{\rm out,2})} + \|\mathcal{L}u_{\rm in}\|_{H_k^m(\Omega_{\rm in,2})} + \|\mathcal{L}u_{\rm out}\|_{H_k^m(\Omega_{\rm out,2})} \Big).\end{aligned} \tag{3.5}$$

Proof. In unweighted norms, this is, e.g., [12, Theorem 5.2.1(i)] (and [32, Theorems 4.7 and 4.20] when $\zeta = 1$); the proof in the weighted norms (3.2) is very similar. ∎

THEOREM 3.4 (Local elliptic regularity for the impedance problem). *Let Ω be a Lipschitz domain and let G_1, G_2 be open subsets of $\mathbb{R}^d$ with $G_1 \Subset G_2$ and $G_1 \cap \partial\Omega \neq \emptyset$. Let Ω_j and Γ_2 be defined by (3.3). Suppose that, for some $m \in \mathbb{N}$, $\Gamma_2 \in C^{m+1,1}$. Given $k_0 > 0$, there exists $C > 0$ such that if $k \geq k_0$, $u \in H^1(\Omega_2)$, $\Delta u \in H^m(\Omega_2)$, and $(k^{-1}\partial_n - {\rm i})u = 0$ on Γ_2, then*

$$\|u\|_{H_k^{m+2}(\Omega_1)} \leq C\big(\|u\|_{H_k^1(\Omega_2)} + \big\|k^{-2}\Delta u\big\|_{H_k^m(\Omega_2)} \big). \tag{3.6}$$

Proof. When $m = 0$, the result can be obtained from [8, Corollary 4.2/Theorem 4.3] by multiplying by k^{-2} to switch to weighted norms, and using that the trace operator has norm bounded by $Ck^{1/2}$ from H_k^1 to L^2 (which can be obtained from, e.g., [35, Theorem 5.6.4] since the weighted norms there are, up to a constant, the weighted norms (1.5)).

The proof that (3.6) follows for $m > 0$ is then standard and can be found e.g. in [15, §6.3.2, Theorem 5]. We repeat it here in the context of impedance boundary conditions for completeness.

We now prove that if the bound holds for $m = q$, then it holds for $m = q + 1$ (assuming the appropriate regularity of the coefficients and the domain). Without loss of generality, we can change coordinates and work with $U := B(0, s) \cap \{x_d > 0\}$ and $V := B(0, t) \cap \{x_d > 0\}$ for some $0 < t < s$. In these coordinates

$$\widetilde{\mathcal{L}}u := (-k^{-2}a^{ij}\partial_{x_i}\partial_{x_j} - k^{-2}(b^i\partial_{x_i} - c))u = f, \qquad (-k^{-1}\partial_{x_d} - {\rm i})u = 0 \text{ on } \{x_d = 0\} \cap \overline{U}.$$

Suppose that for some $q \geq 0$, for any $0 < t < s$,

$$\|u\|_{H_k^{q+2}(V)} \leq C_t\big(\|u\|_{L^2(U)} + \|f\|_{H_k^q(U)} \big). \tag{3.7}$$

Now suppose that $f \in H_k^{q+1}(U)$ and $a, b, c \in C^{q+1,1}(\overline{U})$, and let $W := B(0, r) \cap \{x_d > 0\}$ with $t < r < s$. By (3.7),

$$\|u\|_{H_k^{q+2}(W)} \leq C\big(\|u\|_{L^2(U)} + \|f\|_{H_k^q(U)} \big), \tag{3.8}$$

and, by interior elliptic regularity, $u \in H^{q+3}_{\rm loc}(U)$.

The next step is to bound tangential derivatives of u: let α be a multiindex with $|\alpha| = q+1$ and $\alpha_d = 0$ (so that ∂_x^α is a tangential derivative). Let

$$\widetilde{f} := \widetilde{\mathcal{L}}\big(k^{-|\alpha|}\partial_x^\alpha u\big) \quad \text{ so that } \quad \widetilde{f} = [\widetilde{\mathcal{L}}, k^{-|\alpha|}\partial_x^\alpha]u + k^{-|\alpha|}\partial_x^\alpha f$$

(where $[A,B] := AB - BA$). Therefore,

$$\|\widetilde{f}\|_{L^2(W)} \leq C\big(\, \|u\|_{H_k^{q+2}(W)} + \|f\|_{H_k^{q+1}(W)}\,\big) \leq C\big(\, \|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\,\big). \tag{3.9}$$

where to obtain the last inequality we have used (3.8) and the fact that the coefficients of $\widetilde{\mathcal{L}}$ are $C^{q+1,1}(\overline{U})$. Furthermore $k^{-|\alpha|}\partial_x^\alpha u$ satisfies the impedance boundary condition, since

$$(-k^{-1}\partial_{x_d} - {\rm i})k^{-|\alpha|}\partial_x^\alpha u|_{x_d=0} = k^{-|\alpha|}\partial_x^\alpha\big[(-k^{-1}\partial_{x_d} - {\rm i})u|_{x_d=0}\big] = 0. \tag{3.10}$$

Therefore, by the analogue of (3.7) with $q=0$ and U replaced by W, (3.8), and (3.9),

$$\big\|k^{-|\alpha|}\partial_x^\alpha u\big\|_{H_k^2(V)} \leq C\big(\big\|k^{-|\alpha|}\partial_x^\alpha u\big\|_{L^2(W)} + \big\|\widetilde{f}\big\|_{L^2(W)}\big) \leq C\big(\, \|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\,\big).$$

In summary, recalling the definition of α, we have proved that

$$\big\|k^{-|\beta|}\partial_x^\beta u\big\|_{L^2(V)} \leq C\big(\, \|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\,\big) \text{ for all } |\beta| = q+3 \text{ with } \beta_d \in \{0,1,2\}.$$

To prove that the bound (3.7) holds with q replaced by $q+1$, i.e.,

$$\|u\|_{H_k^{q+3}(V)} \leq C\big(\, \|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\big),$$

it is sufficient to prove that

$$\big\|k^{-|\beta|}\partial_x^\beta u\big\|_{L^2(V)} \leq C\big(\, \|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\,\big) \quad \text{for all } |\beta| = q+3 \text{ with } \beta_d \in \{0,\ldots,q+3\}.$$

We therefore now prove by induction that if

$$\big\|k^{-|\beta|}\partial_x^\beta u\big\|_{L^2(V)} \leq C\big(\, \|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\,\big) \tag{3.11}$$

for any $|\beta| = q+3$ with $\beta_d \in \{0,\ldots,j\}$ for some $j \in \{2,\ldots,q+2\}$, then (3.11) holds for $|\beta| = q+3$ with $\beta_d = j+1$. Combined with (4), this completes the proof.

We therefore assume that $|\beta| = q+3$ with $\beta_d = j+1$. Then, putting $\beta = \gamma + \delta$ with $\delta = (0,\ldots,0,2)$ and $|\gamma| = q+1$ (so that $\gamma_d = j-1$), and using that $u \in H^{q+3}_{\rm loc}(U)$, we have

$$k^{-|\gamma|}\partial^\gamma \widetilde{\mathcal{L}}u = a^{dd}k^{-|\beta|}\partial^\beta u + Bu \quad \text{ in } V, \tag{3.12}$$

where

$$Bu = \sum_{|\alpha|\leq q+3,\, \alpha_d \leq j} b_\alpha k^{-|\alpha|}\partial_x^\alpha u$$

for appropriate b_α. By the induction hypothesis (3.11),

$$\|Bu\|_{L^2(V)} \leq C\big(\, \|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\,\big).$$

Dividing (3.12) by a^{dd}, taking the $L^2(V)$ norm, and using that $1/a^{dd}$ is bounded, we have

$$\|k^{-|\beta|}\partial^\beta u\|_{L^2(V)} \leq C\big(\|u\|_{L^2(U)} + \|f\|_{H_k^{q+1}(U)}\big);$$

i.e., we have proved that (3.11) holds for $|\beta| = q+3$ with $\beta_d = j+1$, and the proof is complete. ∎

Let $\mathrm{DtN}_k : H^{1/2}(\partial B_R) \to H^{-1/2}(\partial B_R)$ be the Dirichlet-to-Neumann map, $u \mapsto k^{-1}\partial_r u$, for the Helmholtz equation $(k^{-2}\Delta + 1)u = 0$ posed in the exterior of B_R and satisfying the Sommerfeld radiation condition

$$k^{-1}\frac{\partial u}{\partial r}(x) - \mathrm{i}u(x) = o\Big(\frac{1}{r^{(d-1)/2}}\Big) \quad \text{as } r := |x| \to \infty, \text{ uniformly in } \widehat{x} := x/r. \tag{3.13}$$

For explicit expressions for DtN_k in terms of spherical harmonics and Hankel and Bessel functions, see, e.g., [33, Equations 3.7 and 3.10].

Theorem 3.5 (Elliptic regularity for the Laplacian on a ball with boundary condition involving DtN_k). *Given $R, k_0 > 0$, there exists $C > 0$ such that the following is true for all $k \geq k_0$. Let $B_R := \{x : |x| < R\}$ and suppose that $u \in H^1(B_R)$, $\Delta u \in H^m(B_R)$, and $(k^{-1}\partial_n - T)u = 0$ on ∂B_R, where T is one of*

$$\mathrm{DtN}_k, \quad \mathrm{DtN}_k^*, \quad \textit{and} \quad (\mathrm{DtN}_k + \mathrm{DtN}_k^*)/2. \tag{3.14}$$

Then

$$\|u\|_{H_k^{m+2}(B_R)} \leq C\big(\|u\|_{L^2(B_R)} + \big\|k^{-2}\Delta u\big\|_{H_k^m(B_R)}\big). \tag{3.15}$$

Proof. When $m = 0$ and $T = \mathrm{DtN}_k$, the bound (3.15) is contained in [27, Lemma 6.4]. The only specific property of DtN_k used in the proof is that

$$-\Re\langle \mathrm{DtN}_k\phi, \phi\rangle \geq 0 \quad \text{for all } \phi \in H^{1/2}(\partial B_R) \tag{3.16}$$

(see [35, Theorem 2.6.4], [5, Lemma 2.1]); thus (3.15) holds for $m = 0$ and the other two choices of T in (3.14). Furthermore, applying (3.15) with $m = 0$ to φu with $\varphi \in C^\infty(\mathbb{R}^d)$, $\varphi \equiv 0$ on B_{R_2}, and $\varphi \equiv 1$ on $(B_{R_3})^c$, with $R_1 < R_2 < R_3 < R$, we find that

$$\|u\|_{H_k^2(B_R\setminus B_{R_3})} \leq C\Big(\|u\|_{L^2(B_R\setminus B_{R_2})} + \|k^{-2}\Delta u\|_{L^2(B_R\setminus B_{R_2})} + k^{-1}\|u\|_{H_k^1(B_R\setminus B_{R_2})}\Big). \tag{3.17}$$

Let $\widetilde{\varphi} \in C^\infty(\mathbb{R}^d)$ be such that $\widetilde{\varphi} \equiv 1$ on $(B_{R_2})^c$, $\widetilde{\varphi} \equiv 0$ on B_{R_1}, and $|\nabla\widetilde{\varphi}|^2/|\widetilde{\varphi}|$ is bounded (this last condition can be achieved by making $\widetilde{\varphi}$ vanish quadratically at ∂B_{R_1}). Applying Green's identity to u and $\widetilde{\varphi}u$, and using (3.16), we find that

$$\int_{B_R} \widetilde{\varphi}|k^{-1}\nabla u|^2 \leq -\Re\bigg(\int_{B_R} \overline{u}(k^{-1}\nabla\widetilde{\varphi}) \cdot (k^{-1}\nabla u) + \widetilde{\varphi}\overline{u}(k^{-2}\Delta u)\bigg).$$

Using the Cauchy–Schwarz inequality and the inequality $2ab \leq \epsilon a^2 + \epsilon^{-1}b^2$ for all $a, b, \epsilon > 0$, we obtain that

$$\|k^{-1}\nabla u\|_{L^2(B_R\setminus B_{R_2})} \leq C\big(\|u\|_{L^2(B_R\setminus B_{R_1})} + \|k^{-2}\Delta u\|_{L^2(B_R\setminus B_{R_1})}\big). \tag{3.18}$$

Combining (3.17) and (3.18), we obtain that

$$\|u\|_{H_k^2(B_R\setminus B_{R_3})} \leq C\big(\|u\|_{L^2(B_R\setminus B_{R_1})} + \|k^{-2}\Delta u\|_{L^2(B_R\setminus B_{R_1})}\big). \tag{3.19}$$

We now repeat the argument for increasing m used in Theorem 3.4 except that we work in an annulus around ∂B_R in polar coordinates. Without loss of generality $R > 2$, so that, for $u \in C^\infty_{\rm comp}(B(0, R+1) \setminus \overline{B(0, R-1)}$,

$$-k^{-2}\Delta u = \Big(-k^{-2}\partial_r^2 - \frac{d-1}{kr}k^{-1}\partial_r - r^{-2}k^{-2}\Delta_\omega\Big)u,$$

where $[R-1, R+1] \times S^{d-1} \ni (r,\omega) \mapsto r\omega \in \mathbb{R}^d$ and Δ_ω denotes the Laplacian on S^{d-1}. Now, DtN_k and $\Delta_\omega|_{\partial B_R}$ commute: this can be seen, e.g., from the expression for DtN_k in terms of spherical harmonics on ∂B_R.

Let $\chi \in C^\infty_{\rm comp}(B(0,R+1) \setminus \overline{B(0,R-1)})$ with $\chi \equiv 1$ near ∂B_R and put $v = \chi u$ so that

$$-\Delta v = [-\Delta, \chi]u - \chi\Delta u$$

Then,

$$(-\Delta_\omega)^j(-\Delta v) = (-\Delta_\omega)^j\big([-\Delta,\chi]u - \chi\Delta u\big), \qquad (k^{-1}\partial_r - T)(-\Delta_\omega)^j v|_{\partial B_R} = 0.$$

Since $-\Delta_\omega$ commutes with $-\Delta$,

$$-\Delta(-\Delta_\omega)^j v = (-\Delta_\omega)^j(-\Delta v) = (-\Delta_\omega)^j\big([-\Delta,\chi]u - \chi\Delta u\big).$$

In particular, letting $R_1 < R_3 < R$, such that $B_R \setminus B_{R_1} \subset \{\chi \equiv 1\}$, (3.19) implies

$$\begin{aligned}\|k^{-2j}(-\Delta_\omega)^j v\|_{H^2_k(B_R\setminus B_{R_3})} &\le C\|k^{-2j}(-\Delta_\omega)^j u\|_{L^2(B_R\setminus B_{R_1})} \\ &\qquad + \|k^{-2j}(-\Delta_\omega)^j k^{-2}\Delta u\|_{L^2(B_R\setminus B_{R_1})} \\ &\le C\|k^{-2j}(-\Delta_\omega)^j u\|_{L^2(B_R\setminus B_{R_1})} + C\|k^{-2}\Delta u\|_{H^{2j}_k(B_R)}.\end{aligned}$$

Using that

$$C\|k^{-2j}(-\Delta_\omega)^j u\|_{L^2(B_R\setminus B_{R_1})} \le C\|k^{-2(j-1)}(-\Delta_\omega)^{j-1}u\|_{H^2_k(B_R\setminus B_{R_1})}$$

together with the $m = 0$ case and induction on j, we obtain for any $R > R_4 > R_3$ that

$$\|k^{-2j}(-\Delta_\omega)^j v\|_{H^2_k(B_R\setminus B_{R_4})} \le C\|u\|_{L^2(B_R)} + C\|k^{-2}\Delta u\|_{H^{2j}_k(B_R)}.$$

Applying elliptic regularity of $(-\Delta_\omega)$ for each fixed $R_4 \le r \le R$, we obtain for any $|\alpha| + \ell \le 2j+2$ and $\ell \in \{0,1,2\}$ that

$$\|k^{-|\alpha|-\ell}\partial_\omega^\alpha\partial_r^\ell v\|_{L^2(B_R\setminus B_{R_4})} \le C\|u\|_{L^2(B_R)} + C\|k^{-2}\Delta u\|_{H^{2j}_k(B_R)}.$$

We then proceed as in the proof of Theorem 3.4 starting from (3.11) to complete the proof of (3.15). ∎

4. Theorem 1.5 applied to the PML problem.

4.1. Definition of the PML problem.

Obstacles and coefficients for Dirichlet/Neumann/penetrable obstacle problem. Let $\Omega_{\rm p}, \Omega_- \subset B_{R_0} := \{x : |x| < R_0\} \subset \mathbb{R}^d$, $d = 2, 3$, be bounded open sets with Lipschitz boundaries, $\Gamma_{\rm p}$ and Γ_-, respectively, such that $\Gamma_{\rm p} \cap \Gamma_- = \emptyset$, and $\mathbb{R}^d\setminus\overline{\Omega_-}$ is connected. Let $\Omega_{\rm out} := \mathbb{R}^d\setminus\overline{\Omega_- \cup \Omega_{\rm p}}$ and $\Omega_{\rm in} := (\mathbb{R}^d\setminus\overline{\Omega_-}) \cap \Omega_{\rm p}$.

Let $A_{\rm out} \in C^{0,1}(\Omega_{\rm out}, \mathbb{R}^{d\times d})$ and $A_{\rm in} \in C^{0,1}(\Omega_{\rm in}, \mathbb{R}^{d\times d})$ be symmetric positive definite, let $c_{\rm out} \in L^\infty(\Omega_{\rm out};\mathbb{R})$, $c_{\rm in} \in L^\infty(\Omega_{\rm in};\mathbb{R})$ be strictly positive, and let $A_{\rm out}$ and $c_{\rm out}$ be such that there exists $R_{\rm scat} > R_0 > 0$ such that

$$\overline{\Omega_-} \cup \operatorname{supp}(I - A_{\rm out}) \cup \operatorname{supp}(1 - c_{\rm out}) \Subset B_{R_{\rm scat}}.$$

The obstacle Ω_- is the impenetrable obstacle, on which we impose either a zero Dirichlet or a zero Neumann condition, and the obstacle $\Omega_{\rm in}$ is the penetrable obstacle, across whose boundary we impose transmission conditions.

For simplicity, we do not cover the case when Ω_- is disconnected, with Dirichlet boundary conditions on some connected components and Neumann boundary conditions on others, but the main results hold for this problem too (at the cost of introducing more notation).

Definition of the radial PML. Let $R_{\rm tr} > R_{\rm PML,-} > R_{\rm scat}$ and let $\Omega_{\rm tr} \subset \mathbb{R}^d$ be a bounded Lipschitz open set with $B_{R_{\rm tr}} \subset \Omega_{\rm tr} \subset B_{CR_{\rm tr}}$ for some $C > 0$ (i.e., $\Omega_{\rm tr}$ has characteristic length scale $R_{\rm tr}$). Let $\Omega := \Omega_{\rm tr} \cap (\Omega_{\rm in} \cup \Omega_{\rm out})$ and $\Gamma_{\rm tr} := \partial\Omega_{\rm tr}$. For $0 \leq \theta < \pi/2$, let the PML scaling function $f_\theta \in C^3([0,\infty);\mathbb{R})$ be defined by $f_\theta(r) := f(r)\tan\theta$ for some f satisfying

(4.1)
$$\big\{f(r) = 0\big\} = \big\{f'(r) = 0\big\} = \big\{r \leq R_{\rm PML,-}\big\}, \quad f'(r) \geq 0, \quad f(r) \equiv r \text{ on } r \geq R_{\rm PML,+};$$

i.e., the scaling "turns on" at $r = R_{\rm PML,-}$, and is linear when $r \geq R_{\rm PML,+}$. We note that $R_{\rm tr}$ can be $< R_{\rm PML,+}$, i.e., we allow truncation before linear scaling is reached. Given $f_\theta(r)$, let

$$\alpha(r) := 1 + {\rm i} f_\theta'(r) \quad \text{ and } \quad \beta(r) := 1 + {\rm i} f_\theta(r)/r. \tag{4.2}$$

and let

(4.3)
$$A := \begin{cases} A_{\rm in} & \text{in } \Omega_{\rm in}, \\ A_{\rm out} & \text{in } \Omega_{\rm out} \cap B_{R_{\rm PML,-}}, \\ HDH^T & \text{in } (B_{R_{\rm PML,-}})^c \end{cases} \text{ and } \frac{1}{c^2} := \begin{cases} c_{\rm in}^{-2} & \text{in } \Omega_{\rm in}, \\ c_{\rm out}^{-2} & \text{in } \Omega_{\rm out} \cap B_{R_{\rm PML,-}}, \\ \alpha(r)\beta(r)^{d-1} & \text{in } (B_{R_{\rm PML,-}})^c, \end{cases}$$

where, in polar coordinates (r,φ),

(4.4)
$$D = \begin{pmatrix} \beta(r)\alpha(r)^{-1} & 0 \\ 0 & \alpha(r)\beta(r)^{-1} \end{pmatrix} \quad \text{and} \quad H = \begin{pmatrix} \cos\varphi & -\sin\varphi \\ \sin\varphi & \cos\varphi \end{pmatrix} \text{ for } d = 2,$$

and, in spherical polar coordinates (r,φ,ϕ),

(4.5)
$$D = \begin{pmatrix} \beta(r)^2\alpha(r)^{-1} & 0 & 0 \\ 0 & \alpha(r) & 0 \\ 0 & 0 & \alpha(r) \end{pmatrix} \text{ and } H = \begin{pmatrix} \sin\varphi\cos\phi & \cos\varphi\cos\phi & -\sin\phi \\ \sin\varphi\sin\phi & \cos\varphi\sin\phi & \cos\phi \\ \cos\varphi & -\sin\varphi & 0 \end{pmatrix}$$

for $d = 3$ (observe that then $A_{\rm out} = I$ and $c_{\rm out}^{-2} = 1$ when $r = R_{\rm PML,-}$ and thus A and c^{-2} are continuous at $r = R_{\rm PML,-}$).

We highlight that, in other papers on PMLs, the scaled variable, which in our case is $r + {\rm i} f_\theta(r)$, is often written as $r(1 + {\rm i}\widetilde{\sigma}(r))$ with $\widetilde{\sigma}(r) = \sigma_0$ for r sufficiently large; see, e.g., [23, §4], [3, §2]. Therefore, to convert from our notation, set $\widetilde{\sigma}(r) = f_\theta(r)/r$ and $\sigma_0 = \tan\theta$.

Let

$$\mathcal{H} := H_0^1(\Omega) \quad \text{or} \quad \{v \in H^1(\Omega) \,:\, v = 0 \text{ on } \Gamma_{\rm tr}\}, \tag{4.6}$$

with the former corresponding to zero Dirichlet boundary conditions on Ω_- and the latter corresponding to zero Neumann boundary conditions on Ω_-.

DEFINITION 4.1 (A variational formulation of the PML problem). *Given $G \in \mathcal{H}^*$ and $\zeta > 0$,*

$$\text{find } u \in \mathcal{H} \text{ such that } a(u,v) = G(v) \text{ for all } v \in \mathcal{H}, \tag{4.7}$$

where

$$a(u,v) := \left(\int_{\Omega\cap\Omega_{\rm out}} + \frac{1}{\zeta}\int_{\Omega_{\rm in}}\right)\left(k^{-2}(A\nabla u)\cdot\overline{\nabla v} - c^{-2}u\overline{v}\right). \tag{4.8}$$

When

$$G(v) := \left(\int_{B_{R_{\rm PML,-}}\cap\Omega_{\rm out}} + \frac{1}{\zeta}\int_{\Omega_{\rm in}}\right) c^{-2} g\overline{v} \tag{4.9}$$

for $g \in L^2(\Omega)$ with $\operatorname{supp} g \subset B_{R_{\rm PML,-}}$, the variational problem (4.7) is a weak form of the problem

$$\begin{aligned} k^{-2}c_{\rm out}^2\nabla\cdot(A_{\rm out}\nabla u_{\rm out}) + u_{\rm out} &= -g \quad \text{in } \Omega_{\rm out},\\ k^{-2}c_{\rm in}^2\nabla\cdot(A_{\rm in}\nabla u_{\rm in}) + u_{\rm in} &= -g \quad \text{in } \Omega_{\rm in},\\ u_{\rm in} = u_{\rm out} \quad \text{and} \quad \partial_{n,A_{\rm in}}u_{\rm in} &= \zeta\partial_{n,A_{\rm out}}u_{\rm out} \quad \text{on } \Gamma_{\rm p},\\ \text{either} \quad u_{\rm in} = 0 \quad \text{or} \quad \partial_{n,A_{\rm in}}u_{\rm in} &= 0 \quad \text{on } \Gamma_-, \end{aligned} \tag{4.10}$$

and with the Sommerfeld radiation condition approximated by a radial PML ((4.7) is obtained by multiplying the PDEs above by $c_{\rm in/out}^{-2}\alpha\beta^{d-1}$ and integrating by parts).

Using the fact that the solution of the true scattering problem exists and is unique with $A_{\rm out}, A_{\rm in}, c_{\rm out}, c_{\rm in}, \Omega_-$, and $\Omega_{\rm in}$ described above (see, e.g., the discussion and references in [21, §1]), the solution of (4.7) exists and is unique (i) for fixed k and sufficiently large $R_{\rm tr} - R_1$ by [29, Theorem 2.1], [30, Theorem A], [23, Theorem 5.8] and (ii) for fixed $R_{\rm tr} > R_1$ and sufficiently large k by [18, Theorem 1.5].

For the particular data G (4.9), it is well-known that, for fixed k, the error $\|u-v\|_{H_k^1(B_{R_{\rm PML,-}}\setminus\Omega)}$ decays exponentially in $R_{\rm tr}-R_{\rm PML,-}$ and $\tan\theta$; see [29, Theorem 2.1], [30, Theorem A], [23, Theorem 5.8]. It was recently proved in [18, Theorems 1.2 and 1.5] that the error $\|u-v\|_{H_k^1(B_{R_{\rm PML,-}}\setminus\Omega)}$ also decreases exponentially in k.

4.2. Showing that the PML problem fits in the abstract framework used in Theorem 1.5. Recall that $\mathcal{H}$ is defined by (4.6) and let $\mathcal{H}_0 = L^2(\Omega)$. We work with the norm $\|\cdot\|_{H_k^1(\Omega)}$ (1.5) on $\mathcal{H}$.

We first check that the sesquilinear form a (4.8) is continuous and satisfies a Gårding inequality, with constants uniform for $\epsilon \le \theta \le \pi/2-\epsilon$.

LEMMA 4.2 (Bounds on the coefficients A and c). *Given A and c as in (4.3), a scaling function $f(r)$ satisfying (4.1), and $\epsilon > 0$ there exist A_+ and c_- such that, for all $\epsilon \le \theta \le \pi/2-\epsilon$, $x \in \Omega$, and $\xi, \zeta \in \mathbb{C}^d$,*

$$|(A(x)\xi,\zeta)_2| \le A_+\|\xi\|_2\|\zeta\|_2 \quad \text{and} \quad \frac{1}{|c(x)|^2} \ge \frac{1}{c_-^2}.$$

Proof. This follows from the definitions of A and c in (4.3), the definitions of α and β in (4.2), and the fact that $f_\theta(r) := f(r)\tan\theta$. ∎

Continuity of a (1.6) with $C_{\rm cont} := \max\{A_+, c_-^{-2}\}$ then follows from the Cauchy-Schwarz inequality and the definition of $\|\cdot\|_{H^1_k(\Omega)}$ (1.5).

ASSUMPTION 4.3. *When $d = 3$, $f_\theta(r)/r$ is nondecreasing.*

Assumption 4.3 is standard in the literature; e.g., in the alternative notation described above it is that $\widetilde{\sigma}$ is non-decreasing – see [3, §2].

REMARK 4.4. *As noted above, the variational problem (4.7) is obtained by multiplying the PDEs in (4.10) by $c_{\rm in/out}^{-2}\alpha\beta^{d-1}$ and integrating by parts (as in [11, §3]). If one integrates by parts the PDEs directly (as in, e.g., [23, Lemma 4.2 and Equation 4.8]), the resulting sesquilinear form satisfies Assumption 1.2 after multiplication by $e^{i\omega}$, for some suitable ω (see Remark 1.3), without the need for Assumption 4.3.*

LEMMA 4.5. *Suppose that f_θ satisfies Assumption 4.3. With A defined by (4.3), given $\epsilon > 0$ there exists $A_- > 0$ such that, for all $\epsilon \le \theta \le \pi/2 - \epsilon$,*

$$\Re\big(A(x)\xi,\xi\big)_2 \ge A_-\|\xi\|_2^2 \quad \text{for all } \xi \in \mathbb{C}^d \text{ and } x \in \Omega.$$

Reference for the proof. See, e.g., [20, Lemma 2.3]. ∎

COROLLARY 4.6. *If f_θ satisfies Assumption 4.3 then*

$$\Re a(w,w) \ge A_-\|w\|^2_{H^1_k(\Omega)} - \big(A_- + c_{\rm min}^{-2}\big)\|w\|^2_{L^2(\Omega)} \quad \text{for all } w \in \mathcal{H}.$$

Let $\mathcal{R} : L^2(\Omega) \to \mathcal{H}$ be defined by $a(\mathcal{R}g, v) = (g, v)_{L^2(\Omega)}$ for all $v \in \mathcal{H}$; i.e., $\mathcal{R}$ is the solution operator of the PML problem. The definition of a and the facts that (with the matrices H and D defined by (4.4), (4.5)) H is real and the matrix D is diagonal (and hence symmetric) imply that $a(\overline{u}, v) = a(\overline{v}, u)$ for all $u, v \in \mathcal{H}$, and thus $\mathcal{R}g = \overline{\mathcal{R}^*\overline{g}}$. We therefore let

$$C_{\rm sol} := \|\mathcal{R}\|_{L^2(\Omega)\to L^2(\Omega)} = \|\mathcal{R}^*\|_{L^2(\Omega)\to L^2(\Omega)}\,. \tag{4.11}$$

We highlight that (i) $C_{\rm sol}$ is bounded by the norm of the solution operator of the true scattering problem (i.e., with the Sommerfeld radiation condition) by [18, Theorem 1.6], (ii) $C_{\rm sol} \sim k$ when the problem is nontrapping (with this the slowest-possible growth in k), and (iii) an advantage of working with the weighted norms (3.2) is that $C_{\rm sol}$ in fact describes the k-dependence of the Helmholtz solution operator between H^m_k and H^{m+2}_k for any m.

LEMMA 4.7 (The PML problem satisfies Assumptions 1.2 and 1.6). *Suppose that f_θ satisfies Assumption 4.3 and, for some $\ell \in \mathbb{Z}^+$, $A_{\rm out}, A_{\rm in}, c_{\rm out}, c_{\rm in}$ are all $C^{\ell-1,1}$, f_θ is $C^{\ell,1}$, and $\Gamma_{\rm p}, \Gamma_-$, and $\Gamma_{\rm tr}$ are all $C^{\ell,1}$. Let*

$$\mathcal{Z}_j = \big\{v : v_{\rm out} \in H^j(\Omega\cap\Omega_{\rm out}),\, v_{\rm in} \in H^j(\Omega_{\rm in})\big\} \cap \mathcal{H} \tag{4.12}$$

with norm

$$\|v\|^2_{\mathcal{Z}_j} := \|v_{\rm out}\|^2_{H^j_k(\Omega_{\rm out}\cap\Omega)} + \|v_{\rm in}\|^2_{H^j_k(\Omega_{\rm in})}\,. \tag{4.13}$$

where the "out" and "in" subscripts denote restriction to $\Omega_{\rm out}\cap\Omega$ and $\Omega_{\rm in}$, respectively.

Then a defined by (4.8) satisfies Assumptions 1.2 and 1.6 and given $\epsilon > 0$ and $k_0 > 0$ there exists $C > 0$ such the bounds (1.8), (1.9), and (1.20) hold for all $k \ge k_0$ and $\epsilon \le \theta \le \pi/2 - \epsilon$.

Proof. With $\mathcal{L}$ defined by (3.1),

$$\sup_{v\in\mathcal{H},\,\|v\|_{(\mathcal{Z}_{j-2})^*}=1} |a(u,v)| = \|\mathcal{L}u\|_{\mathcal{Z}_{j-2}} \quad \text{and} \quad \sup_{u\in\mathcal{H},\,\|u\|_{(\mathcal{Z}_{j-2})^*}=1} |a(u,v)| = \|\mathcal{L}^*v\|_{\mathcal{Z}_{j-2}}.$$

Assumption 3.1 is then satisfied for both $\mathcal{L}$ and $\mathcal{L}^*$ by the definition (4.3) of A, Lemma 4.5, and the fact that A is symmetric.

The bounds (1.8) and (1.20) then hold by combining Theorem 3.2 (used near Γ_- and $\Gamma_{\rm tr}$) and Theorem 3.3 (used near $\Gamma_{\rm p}$) and using the fact that, by Green's identity, for $u \in H^1_0(\Omega)$ with $\mathcal{L}u \in L^2(\Omega)$ and $\partial_{n,A_{\rm in}} u_{\rm in} = \zeta \partial_{n,A_{\rm out}} u_{\rm out}$ on $\partial\Omega_{\rm in}$,

$$\begin{aligned}&\|u_{\rm in}\|_{H^1_k(\Omega_{\rm in})} + \|u_{\rm out}\|_{H^1_k(\Omega_{\rm out})} \\ &\qquad \le C\Big(\|u_{\rm in}\|_{L^2(\Omega_{\rm in})} + \|u_{\rm out}\|_{L^2(\Omega_{\rm out})} + \|\mathcal{L}u_{\rm in}\|_{L^2(\Omega_{\rm in})} + \|\mathcal{L}u_{\rm out}\|_{L^2(\Omega_{\rm out})} \Big)\end{aligned}$$

(so that the H^1_k norms on the right-hand sides of (3.4) and (3.5) can be replaced by L^2 norms). Since the operator associated with the sesquilinear form $\Re a$ is

$$\left(\frac{\mathcal{L}+\mathcal{L}^*}{2}\right) u = -k^{-2}\nabla\cdot\left(\frac{A+\overline{A}}{2}\nabla u\right) - \left(\frac{c^{-2}+\overline{c}^{-2}}{2}\right) u$$

and the matrix A is symmetric, this operator also satisfies Assumption 3.1. The bound (1.9) then holds by a very similar argument. ∎

4.3. Theorem 1.5 applied to the PML problem.

ASSUMPTION 4.8. *Given $p \in \mathbb{Z}^+$, $(\mathcal{H}_h)_{h>0}$ are such that the following holds. There exists $C > 0$ such that, for all $h > 0$, $0 \le j \le m+1 \le p+1$, and $v \in \mathcal{Z}_{m+1}$ defined by (4.12), there exists $\mathcal{I}_{h,p}v \in \mathcal{H}_h$ such that*

$$\begin{aligned}&\big|v_{\rm out} - (\mathcal{I}_{h,p}v)_{\rm out}\big|_{H^j(\Omega_{\rm out}\cap\Omega)} + \big|v_{\rm in} - (\mathcal{I}_{h,p}v)_{\rm in}\big|_{H^j(\Omega_{\rm in})} \\ &\qquad \le Ch^{m+1-j}\big(\|v_{\rm out}\|_{H^{m+1}(\Omega_{\rm out}\cap\Omega)} + \|v_{\rm in}\|_{H^{m+1}(\Omega_{\rm in})}\big).\end{aligned} \tag{4.14}$$

where the "out" and "in" subscripts denote restriction to $\Omega_{\rm out}\cap\Omega$ and $\Omega_{\rm in}$, respectively.

Assumption 4.8 holds when $(\mathcal{H}_h)_{h>0}$ consists of piecewise degree-p polynomials on shape-regular simplicial triangulations, indexed by the meshwidth; see, e.g., [10, Theorem 17.1], [4, Proposition 4.4.20].

THEOREM 4.9 (Existence, uniqueness, and error bound in the preasymptotic regime for the PML problem). *Suppose that $\zeta > 0$, f_θ satisfies Assumption 4.3, and, for some $\ell \in \mathbb{Z}^+$, $A_{\rm out}, A_{\rm in}, c_{\rm out}, c_{\rm in}$ are all $C^{\ell-1,1}$, f_θ is $C^{\ell,1}$, and $\Gamma_{\rm p}, \Gamma_-$, and $\Gamma_{\rm tr}$ are all $C^{\ell,1}$. Let $C_{\rm sol}$ be defined by (4.11), and assume that $\{\mathcal{H}_h\}_{h>0}$ satisfy Assumption 4.8. Given $\epsilon > 0$ and $p \in \mathbb{Z}^+$ with $p \ge \ell$, there exists $k_0 > 0$ and $C_j, j = 1,2,3$, such that the following is true for all $k \ge k_0$ and $\epsilon \le \theta \le \pi/2 - \epsilon$.*

The solution u of the PML problem (4.7) exists and is unique, and if

$$(hk)^{2\ell} C_{\rm sol} \le C_1 \tag{4.15}$$

then the Galerkin solution u_h, exists, is unique, and satisfies

$$\|u-u_h\|_{H^1_k(\Omega)} \le C_2\Big(1 + (hk)^\ell C_{\rm sol}\Big) \min_{v_h\in\mathcal{H}_h} \|u-v_h\|_{H^1_k(\Omega)}, \tag{4.16}$$

$$\|u-u_h\|_{L^2(\Omega)} \le C_3\Big(hk + (hk)^\ell C_{\rm sol}\Big) \min_{v_h\in\mathcal{H}_h} \|u-v_h\|_{H^1_k(\Omega)}. \tag{4.17}$$

If, in addition, $g \in H^{\ell-1}(\Omega) \cap \mathcal{H}$ *(with* $\mathcal{H}$ *defined by (4.6)) with*

$$\|g\|_{H_k^{\ell-1}(\Omega)} \leq C \, \|g\|_{\mathcal{H}^*} \tag{4.18}$$

for some $C > 0$*, then there exists* $C_4 > 0$ *such that if* h *satisfies (4.15) then*

$$\frac{\|u - u_h\|_{H_k^1(\Omega)}}{\|u\|_{H_k^1(\Omega)}} \leq C_4 \Big(1 + (hk)^\ell C_{\rm sol} \Big) (hk)^\ell. \tag{4.19}$$

When $p = \ell$, i.e., the polynomial degree matches the regularity of the domain and coefficients, (4.15) becomes the condition (1.1), and the bounds (4.16), (4.17), and (4.19) become (1.2), (1.3), and (1.4), respectively.

Proof of Theorem 4.9. By the results in §4.2, a defined by (4.8) satisfies the assumptions of Theorems 1.5 and 1.7 with $C_{\rm cont}, C_{\rm G1}, C_{\rm G2}, C_{{\rm ell},1}$, and $C_{{\rm ell},2}$ all independent of k. By (4.14), the definition of $\| \cdot \|_{\mathcal{Z}_j}$ (4.13), and the definition (3.2) of the weighted norms, $\|I - \Pi_h\|_{\mathcal{Z}_{m+1} \to \mathcal{H}} \leq C(hk)^m$. This bound along with Theorem 1.7 and (4.11) imply the bound $\eta(\mathcal{H}_h) \leq C(hk + (hk)^\ell C_{\rm sol})$, and the result then follows from Theorem 1.5, using that $hk \leq C$ when (4.15) holds. ∎

5. Theorem 1.5 applied to the exact DtN/impedance problems.

5.1. Definition of the exact DtN/impedance problems.

Let $A_{\rm out}, A_{\rm in}, c_{\rm out}, c_{\rm in}, \Omega_-, \Omega_{\rm in}$, and $\Omega_{\rm tr}$ be as in §4.1. Let

$$A := \begin{cases} A_{\rm in} & \text{in } \Omega_{\rm in}, \\ A_{\rm out} & \text{in } \Omega_{\rm out} \cap \Omega, \end{cases} \qquad \text{and} \qquad \frac{1}{c^2} := \begin{cases} c_{\rm in}^{-2} & \text{in } \Omega_{\rm in}, \\ c_{\rm out}^{-2} & \text{in } \Omega_{\rm out} \cap \Omega. \end{cases}$$

Let

$$\mathcal{H} := \{ v \in H^1(\Omega) \, : \, v = 0 \text{ on } \partial\Omega_- \} \quad \text{or} \quad H^1(\Omega), \tag{5.1}$$

with the former corresponding to zero Dirichlet boundary conditions on Ω_- and the latter corresponding to zero Neumann boundary conditions on Ω_-.

DEFINITION 5.1 (Variational formulation of the impedance/exact DtN problems). *Let* either $T = {\rm i}I$ *(with no further constraint on* $\Omega_{\rm tr}$*)* or $T = {\rm DtN}_k$ *with* $\Omega_{\rm tr} = B(0, R_{\rm tr})$*. Given* $G \in \mathcal{H}^*$ *and* $\zeta > 0$*,*

$$\text{find } u \in \mathcal{H} \text{ such that } a(u, v) = G(v) \text{ for all } v \in \mathcal{H}, \tag{5.2}$$

where

$$a(u, v) := \left(\int_{\Omega \cap \Omega_{\rm out}} + \frac{1}{\zeta} \int_{\Omega \cap \Omega_{\rm in}} \right) \Big(k^{-2} (A \nabla u) \cdot \overline{\nabla v} - c^{-2} u \overline{v} \Big) - k^{-1} \langle Tu, v \rangle_{\Gamma_{\rm tr}}. \tag{5.3}$$

The solution of this variational problem exists and is unique; see, e.g., [22, Theorem 2.4] and the discussion and references in [21, §1].

5.2. Showing that the exact DtN/impedance problems fit in the abstract framework used in Theorem 1.5.

The proofs that the sesquilinear form a is continuous and satisfies a Gårding inequality are very similar to those for the PML problem in §4.2 (in fact, they are simpler because there is no PML scaling parameter in which the bounds need to be uniform). When $T = {\rm DtN}_k$, the proof of the Gårding inequality uses (3.16) and the proof of continuity uses that $|\langle {\rm DtN}_k u, v \rangle_{\partial B_{R_{\rm tr}}}| \leq Ck \|u\|_{H_k^1(\Omega)} \|v\|_{H_k^1(\Omega)}$ [33, Equation 3.4a].

LEMMA 5.2 (The exact DtN/impedance problems satisfy Assumptions 1.2 and 1.6). *Suppose that, for some* $\ell \in \mathbb{Z}^+$, $A_{\rm out}, A_{\rm in}, c_{\rm out}, c_{\rm in}$ *are all* $C^{\ell-1,1}$ *and* $\Gamma_{\rm p}$, Γ_-, *and* $\Gamma_{\rm tr}$ *are all* $C^{\ell,1}$. *With* $\mathcal{Z}_j$ *and its norm defined by* (4.12) *and* (4.13), a *defined by* (5.3) *satisfies Assumptions 1.2 and 1.6 and given* $k_0 > 0$ *there exists* $C > 0$ *such the bounds* (1.8), (1.9), *and* (1.20) *hold for all* $k \geq k_0$.

Proof. This is very similar to the proof of Lemma 4.7. For the impedance problem, the regularity assumptions (1.8) and (1.20) follow by combining Theorem 3.2 used near $\partial\Omega_-$, Theorem 3.3 used near $\partial\Omega_{\rm in}$, and Theorem 3.4 used near $\Gamma_{\rm tr}$. The regularity assumption (1.9) follows by combining Theorem 3.2 used near $\partial\Omega_-$, Theorem 3.3 used near $\partial\Omega_{\rm in}$, and now Theorem 3.2 (with Neumann boundary condition) used near $\Gamma_{\rm tr}$. Indeed, near $\Gamma_{\rm tr}$, the operator associated with $(\Re a)$ is $-k^{-2}\Delta - 1$ with Neumann boundary conditions (coming from $A_{\rm out} = I$ and $c_{\rm out} = 1$ near $\Gamma_{\rm tr}$ and the fact that no boundary condition is imposed on $\Gamma_{\rm tr}$ in $\mathcal{H}$ (5.1)).

For the exact DtN problem, let $\chi \in C^\infty(\mathbb{R}^d)$ be such that $\chi \equiv 0$ on $B_{R_{\rm scat}}$ and $\chi \equiv 1$ on $(B_{(R_{\rm scat}+R_{\rm tr})/2})^c$. Then apply Theorem 3.5 to χu and Theorems 3.2 and 3.3 to $(1-\chi)u$. Note that (i) in these applications, the terms arising from the commutator of $\mathcal{L}$ and χ are lower-order in both k and Sobolev index and (ii) the different choices of T in (3.14) correspond to (1.8), (1.20), and (1.9), respectively. ∎

5.3. Theorem 1.5 applied to the exact DtN/impedance problems.

THEOREM 5.3 (Existence, uniqueness, and error bound in the preasymptotic regime for the exact DtN/impedance problems). *Suppose that* $\zeta > 0$ *and, for some* $\ell \in \mathbb{Z}^+$, $A_{\rm out}, A_{\rm in}, c_{\rm out}, c_{\rm in}$ *are all* $C^{\ell-1,1}$ *and* $\Gamma_{\rm p}, \Gamma_-$, *and* $\Gamma_{\rm tr}$ *are all* $C^{\ell,1}$. *Let* $C_{\rm sol}$ *be defined by* (4.11), *and assume that* $\{\mathcal{H}_h\}_{h>0}$ *satisfy Assumption 4.8. Given* $p \in \mathbb{Z}^+$ *with* $p \geq \ell$, *there exists* $k_0 > 0$ *and* $C_j, j = 1, 2, 3$, *such that the following is true for all* $k \geq k_0$.

The solution u *of the exact DtN/impedance problem* (5.2) *exists and is unique, and if* (4.15) *holds then the Galerkin solution* u_h, *exists, is unique, and satisfies the bounds* (4.16) *and* (4.17). *If, in addition,* $g \in H^{\ell-1}(\Omega) \cap \mathcal{H}$ *(with* $\mathcal{H}$ *defined by* (5.1)*) with* (4.18) *for some* $C > 0$, *then there exists* $C_4 > 0$ *such that if* h *satisfies* (4.15) *then the bound* (4.19) *holds.*

Given Lemma 5.2, the proof of Theorem 5.3 is very similar to the proof of Theorem 4.9, and so we omit it for brevity.

Acknowledgements. We thank the referees for their constructive comments. EAS was supported by EPSRC grant EP/R005591/1 and JG was supported by EPSRC grants EP/V001760/1 and EP/V051636/1.

REFERENCES

[1] H. BARUCQ, T. CHAUMONT-FRELET, AND C. GOUT, *Stability analysis of heterogeneous Helmholtz problems and finite element solution based on propagation media approximation*, Math. Comp., 86 (2017), pp. 2129–2157.

[2] M. BERNKOPF, T. CHAUMONT-FRELET, AND J. M. MELENK, *Stability and convergence of Galerkin discretizations of the Helmholtz equation in piecewise smooth media*, Math. Comp., to appear (2024).

[3] J. H. BRAMBLE AND J. PASCIAK, *Analysis of a finite PML approximation for the three dimensional time-harmonic Maxwell and acoustic scattering problems*, Mathematics of Computation, 76 (2007), pp. 597–614.

[4] S. C. BRENNER AND L. R. SCOTT, *The Mathematical Theory of Finite Element Methods*, vol. 15 of Texts in Applied Mathematics, Springer, 3rd ed., 2008.

[5] S. N. Chandler-Wilde and P. Monk, *Wave-number-explicit bounds in time-harmonic scattering*, SIAM J. Math. Anal., 39 (2008), pp. 1428–1455.

[6] T. Chaumont-Frelet, S. Gallistl, D.and Nicaise, and J. Tomezyk, *Wavenumber explicit convergence analysis for finite element discretizations of time-harmonic wave propagation problems with perfectly matched layers*, Comm. Math. Sci., 20 (2022), pp. 1–52.

[7] T. Chaumont-Frelet and S. Nicaise, *Wavenumber explicit convergence analysis for finite element discretizations of general wave propagation problem*, IMA J. Numer. Anal., 40 (2020), pp. 1503–1543.

[8] T. Chaumont-Frelet, S. Nicaise, and J. Tomezyk, *Uniform a priori estimates for elliptic problems with impedance boundary conditions*, Communications on Pure & Applied Analysis, 19 (2020), p. 2445.

[9] T. Chaumont-Frelet and E. A. Spence, *The geometric error is less than the pollution error when solving the high-frequency Helmholtz equation with high-order FEM on curved domains*, arXiv preprint arXiv:2401.16413, (2024).

[10] P. G. Ciarlet, *Basic error estimates for elliptic problems*, in Handbook of numerical analysis, Vol. II, North-Holland, Amsterdam, 1991, pp. 17–351.

[11] F. Collino and P. Monk, *The perfectly matched layer in curvilinear coordinates*, SIAM Journal on Scientific Computing, 19 (1998), pp. 2061–2090.

[12] M. Costabel, M. Dauge, and S. Nicaise, *Corner Singularities and Analytic Regularity for Linear Elliptic Systems. Part I: Smooth domains.*, (2010). https://hal.archives-ouvertes.fr/file/index/docid/453934/filename/CoDaNi_Analytic_Part_I.pdf.

[13] Y. Du and H. Wu, *Preasymptotic error analysis of higher order FEM and CIP-FEM for Helmholtz equation with high wave number*, SIAM J. Numer. Anal., 53 (2015), pp. 782–804.

[14] S. Esterhazy and J. M. Melenk, *On stability of discretizations of the Helmholtz equation*, in Numerical Analysis of Multiscale Problems, I. G. Graham, T. Y. Hou, O. Lakkis, and R. Scheichl, eds., Springer, 2012, pp. 285–324.

[15] L. C. Evans, *Partial differential equations*, AMS, 1998.

[16] X. Feng and H. Wu, *Discontinuous Galerkin methods for the Helmholtz equation with large wave number*, SIAM J. Numer. Anal., 47 (2009), pp. 2872–2896.

[17] X. Feng and H. Wu, *hp-Discontinuous Galerkin methods for the Helmholtz equation with large wave number*, Math. Comp., 80 (2011), pp. 1997–2024.

[18] J. Galkowski, D. Lafontaine, and E. A. Spence, *Perfectly-matched-layer truncation is exponentially accurate at high frequency*, SIAM J. Math. Anal., 55 (2023), pp. 3344–3394.

[19] J. Galkowski, D. Lafontaine, E. A. Spence, and J. Wunsch, *Decompositions of high-frequency Helmholtz solutions via functional calculus, and application to the finite element method*, SIAM J. Math. Anal., 55 (2023), pp. 3903–3958.

[20] J. Galkowski, D. Lafontaine, E. A. Spence, and J. Wunsch, *The hp-FEM applied to the Helmholtz equation with PML truncation does not suffer from the pollution effect*, Comm. Math. Sci., to appear (2024).

[21] 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 (2019), pp. 2869–2923.

[22] I. G. Graham and S. A. Sauter, *Stability and finite element error analysis for the Helmholtz equation with variable coefficients*, Math. Comp., 89 (2020), pp. 105–138.

[23] T. Hohage, F. Schmidt, and L. Zschiedrich, *Solving time-harmonic scattering problems based on the pole condition II: convergence of the PML method*, SIAM Journal on Mathematical Analysis, 35 (2003), pp. 547–560.

[24] F. Ihlenburg, *Finite element analysis of acoustic scattering*, Springer Verlag, 1998.

[25] F. Ihlenburg and I. Babuška, *Finite element solution of the Helmholtz equation with high wave number Part I: The h-version of the FEM*, Comput. Math. Appl., 30 (1995), pp. 9–37.

[26] 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 (1997), pp. 315–358.

[27] D. Lafontaine, E. A. Spence, and J. Wunsch, *A sharp relative-error bound for the Helmholtz h-FEM at high frequency*, Numerische Mathematik, 150 (2022), pp. 137–178.

[28] D. Lafontaine, E. A. Spence, and J. Wunsch, *Wavenumber-explicit convergence of the hp-FEM for the full-space heterogeneous Helmholtz equation with smooth coefficients*, Comp. Math. Appl., 113 (2022), pp. 59–69.

[29] M. Lassas and E. Somersalo, *On the existence and convergence of the solution of PML equations*, Computing, 60 (1998), pp. 229–241.

[30] M. Lassas and E. Somersalo, *Analysis of the PML equations in general convex geome-*

try, Proceedings of the Royal Society of Edinburgh Section A: Mathematics, 131 (2001), pp. 1183–1207.

[31] Y. Li and H. Wu, *FEM and CIP-FEM for Helmholtz Equation with High Wave Number and Perfectly Matched Layer Truncation*, SIAM J. Numer. Anal., 57 (2019), pp. 96–126.

[32] W. McLean, *Strongly elliptic systems and boundary integral equations*, CUP, 2000.

[33] J. M. Melenk and S. Sauter, *Convergence analysis for finite element discretizations of the Helmholtz equation with Dirichlet-to-Neumann boundary conditions*, Math. Comp, 79 (2010), pp. 1871–1914.

[34] J. M. Melenk and S. Sauter, *Wavenumber explicit convergence analysis for Galerkin discretizations of the Helmholtz equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1210–1243.

[35] J. C. Nédélec, *Acoustic and electromagnetic equations: integral representations for harmonic problems*, Springer Verlag, 2001.

[36] O. R. Pembery, *The Helmholtz Equation in Heterogeneous and Random Media: Analysis and Numerics*, PhD thesis, University of Bath, 2020. https://researchportal.bath.ac.uk/en/studentTheses/the-helmholtz-equation-in-heterogeneous-and-random-media-analysis.

[37] H. Wu, *Pre-asymptotic error analysis of CIP-FEM and FEM for the Helmholtz equation with high wave number. Part I: linear version*, IMA J. Numer. Anal., 34 (2014), pp. 1266–1288.

[38] H. W. Y. Li, *Higher-order FEM and CIP-FEM for Helmholtz equation with high wave number and perfectly matched layer truncation*, arXiv preprint arXiv:2312.02476, (2023).

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