# Non-uniform finite element meshes defined by ray dynamics for high-frequency Helmholtz trapping problems

**Martin Averseng**[1,*], **Jeffrey Galkowski**[2], **Euan Spence**[3]

[1]Laboratoire Angevin de Recherche en Mathématiques, Université d'Angers, Angers, France
[2]Department of Mathematics, University College London, London, UK
[3]Department of Mathematical Sciences, University of Bath, Bath, UK

*Email: martin.averseng@univ-angers.fr

**Abstract**

For the $h$-version of the Finite Element Method applied to the high-frequency Helmholtz equation, the sharp criteria to guarantee a $k$-uniform quasioptimality (QO) and bounded relative error (RE) have been studied since the 90s, and are now well-understood for *uniform* meshes. For instance, the sharp condition for QO is that $(hk)^p\rho(k)$ be sufficiently small, where $h$ is the meshwidth, $k$ the wavenumber, $p$ the polynomial degree, and $\rho(k)$ the norm of the Helmholtz solution operator in suitably normed function spaces, such that $\rho(k) \propto k$ for non-trapping problems. The "pollution effect", seen through the factor $\rho(k)$, is even more pronounced for trapping problems (e.g. when there are cavities) since in this case $\rho(k) \gg k$.

Our main result (of which Theorem 1 below is a particular case) is that for trapping problems, QO and RE can be achieved under weaker conditions than the aforementioned sharp thresholds, provided one uses *non-uniform meshes*, with local mesh refinement dictated by properties of the ray dynamics in the propagation domain. This allows in particular for coarser meshes away from cavities. Numerical experiments show that our results are sharp.

**Keywords:** Helmholtz equation, High-frequency, Trapping, Finite Element Method.

## 1 Helmholtz scattering problem

We consider sound-soft scattering by an obstacle $\Omega_- \subset \mathbb{R}^d$ ($d = 2$ or 3):

$$-k^{-2}\operatorname{div}(A(x)\nabla u) - n(x)u = f \text{ in } \Omega_+, \quad (1)$$

where $\Omega_+ = \mathbb{R}^d \setminus \overline{\Omega_-}$, $A$ is a real smooth s.p.d. matrix, $n \in C^\infty(\overline{\Omega_+};\mathbb{R}_+)$ positive, and $\operatorname{supp}(A-I)\cup\operatorname{supp}(n-1) \Subset \overline{\Omega}_+$. The boundary conditions are $u_{|\partial\Omega_+} = 0$ and the Sommerfeld condition at infinity. This problem is truncated to a bounded domain $\Omega = \Omega_{\mathrm{PML}} \cap \Omega_+$ via the method of perfectly matched layers. Define the norm of the solution operator

$$\rho(k) := \sup\big\{\|u\|_{L^2(\Omega)} : u \text{ solves (1) with } \operatorname{supp} f \subset \Omega,\ \|f\|_{L^2(\Omega)} = 1\big\}. \quad (2)$$

One has $\rho(k) \gtrsim k$ [1, Thm. 1.7]. We restrict our attention to a set of wavenumbers $\mathcal{W}_{\mathrm{poly}} \subset \mathbb{R}_+$ on which $\rho$ is polynomially bounded, i.e. there exists $N > 0$ such that $\rho(k) \lesssim k^N$ for all $k \in \mathcal{W}_{\mathrm{poly}}$ (this can be achieved with $\mathcal{W}_{\mathrm{poly}} = \mathbb{R}_+ \setminus J$ where $J$ has arbitrary small measure, see [2]).

## 2 Regions defined by ray-dynamics

Given a mesh $\mathcal{T}$ of $\Omega$, we let $u_h$ be the standard Galerkin approximation of the problem (1) truncated to $\Omega$, in the space $V^p_{\mathcal{T}}$ of continuous piecewise-polynomial (with respect to $\mathcal{T}$) functions of degree $p$, which vanish on $\partial\Omega$. To describe the relevant local mesh sizes, we now define several regions in terms of the *ray trajectories*, i.e. geodesics for the metric $g^{-1} = A/n$ in $\Omega_+$, continued by reflection at the boundary.[1]

Reg. 1: The **cavity** $K \subset \overline{\Omega_+}$: the set of points lying on a ray which is *both* forward and backward-trapped (meaning that this ray stays in a compact set for all positive and negative times).

Reg. 2: The **visible set** (from the cavity) $\Gamma \subset \overline{\Omega_+}$: the set of points lying on a ray which is *either* forward *or* backward-trapped.

Reg. 3: The **invisible set** $I := \overline{\Omega_+} \setminus \Gamma$.

The problem is *trapping* when $K \neq \emptyset$, in which case it can be proved that $\rho(k) \gg k$ as $k \to \infty$ [5] (at least through some sequence of wavenumbers $k$). In what follows, $\Omega_K$, $\Omega_\Gamma$ and $\Omega_I$ are open neighborhoods of $K$, $\Gamma$ and $I$, and $h_K, h_\Gamma, h_I$ denote upper bounds for the diameter of any element intersecting $\Omega_K, \Omega_\Gamma, \Omega_I$, respectively. Let

[1]For example if $A = n = 1$, these are the straight-line paths reflected by Snell-Descartes laws at the boundary.

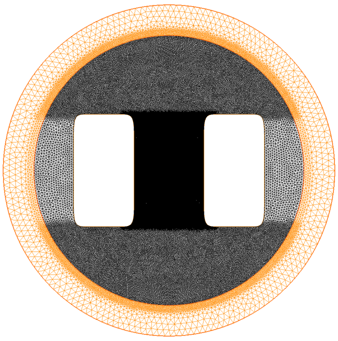

Figure 1: Example of a non-uniform mesh, with the obstacle $\Omega_-$ given by the two rounded rectangles.

$\mathcal{H} := \mathrm{diag}(h_K, h_\Gamma, h_I)$ (the $3 \times 3$ diagonal matrix with entries $h_K, h_\Gamma$ and $h_I$). An example of non-uniform mesh with different thresholds in different regions is shown in Figure 1.

## 3 Main result

To state our main result, we define the matrix

$$\mathcal{C} := \begin{pmatrix} \rho(k) & \sqrt{k\rho(k)} & 0 \\ \sqrt{k\rho(k)} & k & k \\ 0 & k & k \end{pmatrix}.$$

The coefficient $\mathcal{C}_{ij}$ describes how large the solution $u$ can be in region $i$ (up to $O(k^{-\infty})$ remainders), if the data $f$ is supported in region $j$ with unit norm [3, 4]. Given a constant $C_\dagger > 0$, let

$$\mathcal{A}(C_\dagger) := \left( \sum_{n=0}^{+\infty} C_\dagger^n (\mathcal{C}(\mathcal{H}k)^{2p})^n \right) \mathcal{C}(\mathcal{H}k)^p$$

the *accumulation matrix* (which describes the propagation of numerical errors between regions). Finally, for any open set $U$, denote $\|u-u_h\|_U^2 := \|u\|^2_{L^2(U)} + k^{-2}\|\nabla u\|^2_{L^2(U)}$ the “dimension-less" energy norm.

**Theorem 1 (Main result)** *There exists $C_\dagger > 0$ such that for all $0 < c < 1$, there exists $C > 0$ such that for all $k \in \mathcal{W}_{\mathrm{poly}}$ and all meshes $\mathcal{T}$ satisfying*

$$(h_K k)^{2p}\rho(k) + (h_\Gamma k)^{2p} k + (h_I k)^{2p} k < c,$$

*the Galerkin solution, $u_h \in V_\mathcal{T}^p$, to (1) truncated to $\Omega$ by a PML exists, is unique, and satisfies*

$$\begin{pmatrix} \|u-u_h\|_{\Omega_K} \\ \|u-u_h\|_{\Omega_\Gamma} \\ \|u-u_h\|_{\Omega_I} \end{pmatrix} \leq C\Big[\mathrm{Id}+\mathcal{A}(C_\dagger)\Big] \begin{pmatrix} \|u-w_h\|_{\Omega_K} \\ \|u-w_h\|_{\Omega_\Gamma} \\ \|u-w_h\|_{\Omega_I} \end{pmatrix}$$

*for all $w_h \in V_\mathcal{T}^p$, where the inequality is understood component-wise.*

Using this result, one can find mesh conditions to guarantee QO and RE that are weaker than the known ones for uniform meshes:

**Corollary 2 (Condition for QO)** *Under the assumptions of Theorem 1, if*

$$(h_K k)^p \rho(k) + (h_\Gamma k)^p \sqrt{k\rho(k)} + (h_I k)^p k < c,$$

*then the Galerkin solution satisfies*

$$\|u-u_h\|_\Omega \leq C\|u-w_h\|_\Omega$$

*for all $w_h \in V_\mathcal{T}^p$, i.e., it is $k$-uniformly quasi-optimal.*

**Corollary 3 (Condition for RE)** *Under the assumptions of Theorem 1, for any $\varepsilon \leq c$, if*

$$(h_K k)^{2p} \rho(k) + (h_\Gamma k)^{2p} \sqrt{k\rho(k)} + (h_I k)^{2p} k < \varepsilon,$$

*then the Galerkin solution satisfies*

$$\|u-u_h\|_\Omega \leq \sqrt{\varepsilon} C\|u\|_\Omega$$

*for all $w_h \in V_\mathcal{T}^p$, i.e., th relative error is controllably small.*

## References

[1] 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.

[2] D. Lafontaine, E. A. Spence and J. Wunsch, For most frequencies, strong trapping has a weak effect in frequency-domain scattering, *Commun. Pure Appl. Math.* **74** (2021), pp.2025–2063.

[3] N. Burq, Semiclassical estimates for the resolvent in nontrapping geometries, *Int. Math. Res. Not.* **2002**,5 (2002), pp. 221–241.

[4] K. Datchev and A. Vasy, Semiclassical resolvent estimates at trapped sets, *Ann. Inst. Fourier* **62**,6 (2012), pp. 2379–2384.

[5] J.-F. Bony, N. Burq and T. Ramond, Minoration de la résolvante dans le cas captif, *Compte Rendus. Mathématique* **348**,23-24 (2010), pp. 1279–1282