

# Estimation of the dispersion error in the numerical wave number of standard and stabilized finite element approximations of the Helmholtz equation

Lindaura Maria Steffens, Núria Parés and Pedro Díez\*,<sup>†</sup>

*Laboratori de Càlcul Numèric (LaCàN), Universitat Politècnica de Catalunya, C2 Campus Nord UPC,  
E-08034 Barcelona, Spain*

## SUMMARY

An estimator for the error in the wave number is presented in the context of finite element approximations of the Helmholtz equation. The proposed estimate is an extension of the ideas introduced in Steffens and Díez (*Comput. Methods Appl. Mech. Engng* 2009; **198**:1389–1400). In the previous work, the error assessment technique was developed for standard Galerkin approximations. Here, the methodology is extended to deal also with stabilized approximations of the Helmholtz equation. Thus, the accuracy of the stabilized solutions is analyzed, including also their sensitivity to the stabilization parameters depending on the mesh topology. The procedure builds up an inexpensive approximation of the exact solution, using post-processing techniques standard in error estimation analysis, from which the estimate of the error in the wave number is computed using a simple closed expression. The recovery technique used in Steffens and Díez (*Comput. Methods Appl. Mech. Engng* 2009; **198**:1389–1400) is based in a polynomial least-squares fitting. Here a new recovery strategy is introduced, using exponential (in a complex setup, trigonometric) local approximations respecting the nature of the solution of the wave problem.

KEY WORDS: wave problems; Helmholtz equation; *a posteriori* error estimation; error estimation of wave number; dispersion/pollution error; stabilized methods

## 1. INTRODUCTION

Acoustic wave propagation phenomena are often modeled using the Helmholtz equation, assuming a harmonic character of the solution. Thus, time-dependent acoustic pressure is represented as  $p(\mathbf{x}, t) = u(\mathbf{x})e^{i\omega t}$  for a given angular frequency  $\omega$ , and the unknown  $u(\mathbf{x})$  is the spatial distribution of the pressure. Function  $u(\mathbf{x})$  is the solution of the Helmholtz equation with an associated wave number  $\kappa = \omega/c$ ,  $c$  being the speed of sound [1].

Galerkin approximations of the Helmholtz equation are affected by dispersion (or pollution) errors that may be important especially if the wave number is large with respect to the mesh size. The pollution error, as opposed to the standard interpolation error, is global in nature because the error sources affect (pollute) the solution everywhere in the domain, and not only where the resolution of the mesh is not sufficient to properly approximate the solution. Thus, the pollution error cannot be removed by local refinement, even if the quantity to be assessed is defined locally.

---

\*Correspondence to: Pedro Díez, Laboratori de Càlcul Numèric (LaCàN), Universitat Politècnica de Catalunya, C2 Campus Nord UPC, E-08034 Barcelona, Spain.

<sup>†</sup>E-mail: pedro.diez@upc.edu

The effect of the pollution or dispersion error has been extensively addressed in the literature and, concordantly, *a priori* estimates for the dispersion error have been derived [1–7]. Also, *a posteriori* error estimates assessing the accuracy of the finite element approximations of the Helmholtz equation both in global norms or in some specific quantities of interest have been proposed [8–16]. However, the issue of measuring the dispersion error of the approximations of the Helmholtz equation using *a posteriori* error estimates was first addressed in [17].

The wave number corresponding to the approximate solution is different from the exact one. The corresponding error is directly related to the dispersion error and it is, according to practitioners, a good measure in order to assess the overall quality of the numerical solution. The problem of assessing the error in the wave number is addressed in [17] for standard finite element (Galerkin) approximations. The proposed error estimation strategy is paradoxical in the sense that, in the error to be assessed, the obvious information is the exact value  $\kappa$  and all the efforts are devoted to compute the value of the wave number corresponding to the approximate solution. Note that in the usual error estimation business the situation is the opposite: the approximate value is available and the exact value has to be estimated.

In practice, standard Galerkin methods are not competitive for high wave numbers because controlling the pollution effect requires using extremely fine meshes. Numerous approaches alleviating this deficiency have been proposed based on modifications of the classical Galerkin approximation [4, 18–20]. The Galerkin/least-squares method is one of the most popular techniques. It provides a significant reduction in the dispersion error with an extremely simple implementation using only standard resources available in finite element codes [21].

Stabilized formulations allow eliminating the pollution effect for one-dimensional problems. In two dimensions, the pollution effect is reduced substantially but it cannot be completely eliminated [6]. Thus, also when using stabilized formulations, the end-user of a finite element acoustic computation is concerned with the accuracy of the solution in terms of the dispersion. In this work, an extension of [17] is proposed allowing to assess the dispersion error when the approximate solution is computed using either the standard Galerkin method or the GLS method.

The assessment of the dispersion error aims at obtaining a good estimate of the value of the *numerical wave number*, corresponding to the approximate solution. Here, the definition of the numerical wave number provided in [17], based on the idea of fitting the numerical solution into a modified equation, is adopted. This strategy requires obtaining an inexpensive approximation of the solution of the modified problem using post-processing techniques. Here, a new recovery technique is introduced, using exponential functions rather than polynomials, to take advantage of the nature of the solutions of wave propagation problems.

The remainder of the paper is structured as follows. Section 2 introduces the notation and the description of the problem to be solved along with the standard and stabilized Galerkin formulations. Section 3 describes the main ideas of the paper. First, the basics of the dispersion error assessment are reviewed. Then, the extension to stabilized formulations is described. Finally, the standard polynomial recovery is recalled and the novel exponential post-processing technique is introduced. Section 4 contains four numerical examples demonstrating the efficiency of the proposed technique both in academic and practical examples.

## 2. PROBLEM STATEMENT

### 2.1. Acoustic modeling: the Helmholtz equation

The acoustic pressure  $u(\mathbf{x})$  is a complex function taking values in the spatial domain  $\Omega \subset \mathbb{R}^d$  (being  $d=1, 2$  or  $3$ ). The function  $u$  is determined as the solution of the Helmholtz equation

$$-\Delta u - \kappa^2 u = f \quad \text{in } \Omega, \quad (1)$$

which is stated for a given wave number  $\kappa$  as the Fourier transform of the transient wave equation. Equation (1) has to be complemented with proper boundary conditions on  $\partial\Omega$ . For interior problems, three types of boundary conditions are considered: Dirichlet, Neumann and Robin (or mixed).

Thus, the boundary  $\partial\Omega$  is partitioned into three disjoint sets  $\Gamma_D$ ,  $\Gamma_N$  and  $\Gamma_R$  such that  $\partial\Omega = \overline{\Gamma_D \cup \Gamma_N \cup \Gamma_R}$  and its associated boundary conditions are

$$u = \bar{u} \quad \text{on } \Gamma_D, \quad (2a)$$

$$\nabla u \cdot \mathbf{n} = g \quad \text{on } \Gamma_N, \quad (2b)$$

$$\nabla u \cdot \mathbf{n} = \mathcal{M}u \quad \text{on } \Gamma_R, \quad (2c)$$

where  $\mathbf{n}$  is the outward normal to  $\Omega$  and  $\bar{u}$ ,  $f$ ,  $g$  and  $\mathcal{M}$  are the prescribed data, which are assumed to be sufficiently smooth.

*Remark 1*

For interior acoustic wave propagation problems  $g = -i\rho c\kappa\bar{v}_n$  and  $\mathcal{M}u = -i\rho c\kappa A_n u$ , where  $c$  is the speed of sound in the medium,  $\rho$  is the mass density,  $\bar{v}_n$  corresponds to the normal velocity of a vibrating wall producing the sound that propagates within the medium, and the coefficient  $A_n$  is the admittance and represents the structural damping and  $i$  is the standard imaginary unit. For exterior problems, reduced to fictitious domains,  $\mathcal{M}$  is a linear operator called the Dirichlet-to-Neumann (DtN) map relating Dirichlet data to the outward normal derivative of the solution on the fictitious boundary  $\Gamma_R$ . It is worth noting that in general the data  $g$  and  $\mathcal{M}$  depend on the wave number  $\kappa$ . A notation explicitly stating the dependence of  $\kappa$ , for instance  $g(\kappa)$  and  $\mathcal{M}(\kappa)$ , would be more accurate but for the sake of simplicity this dependence is omitted in the notation.

The boundary value problem defined by Equations (1) and (2) is readily expressed in its weak form introducing the solution and test spaces  $\mathcal{U} := \{u \in \mathcal{H}^1(\Omega), u|_{\Gamma_D} = \bar{u}\}$  and  $\mathcal{V} := \{v \in \mathcal{H}^1(\Omega), v|_{\Gamma_D} = 0\}$ . Here,  $\mathcal{H}^1(\Omega)$  is the standard Sobolev space of complex-valued square integrable functions with square integrable first derivatives. The weak form of the problem then reads: find  $u \in \mathcal{U}$  such that

$$a(\kappa; u, v) = l(\kappa; v) \quad \forall v \in \mathcal{V}, \quad (3)$$

where

$$a(\kappa; u, v) := \int_{\Omega} \nabla u \cdot \nabla \bar{v} \, d\Omega - \int_{\Omega} \kappa^2 u \bar{v} \, d\Omega - \int_{\Gamma_R} \mathcal{M}u \bar{v} \, d\Gamma,$$

$$l(\kappa; v) := \int_{\Omega} f \bar{v} \, d\Omega + \int_{\Gamma_N} g \bar{v} \, d\Gamma$$

the symbol  $\bar{\cdot}$  denotes the complex conjugate,  $a(\kappa; \cdot, \cdot)$  is a sesquilinear form and  $l(\kappa; \cdot)$  is an antilinear functional depending on  $\kappa$  through the Neumann boundary conditions  $g$ . The notation adopted marks the explicit dependence of  $\kappa$  on the forms  $a(\kappa; \cdot, \cdot)$  and  $l(\kappa; \cdot)$ . Although not standard, this is useful in the following to assess the error in the wave number. It is worth noting that the sesquilinear form  $a(\kappa; \cdot, \cdot)$  is not elliptic but satisfies the inf-sup condition and the Gårding inequality. However, for large wave numbers  $\kappa$ , the upper bound for the inf-sup condition is too crude [1]. Moreover, the inf-sup property is not carried over from  $\mathcal{V}$  to a discrete subspace yielding to a loss of stability which produces spurious dispersion in the discrete approximations.

## 2.2. Galerkin finite element approximation

The Galerkin approximation is obtained from a partition  $\mathcal{T}_H$  of the domain  $\Omega$  into nonoverlapping elements and introducing the discrete spaces  $\mathcal{U}_H \subset \mathcal{U}$  and  $\mathcal{V}_H \subset \mathcal{V}$  associated with the parameters of the discretization, namely, the characteristic element size  $H$ , and the degree of the polynomial approximation inside the elements  $p$ . The discrete finite element solution is then  $u_H \in \mathcal{U}_H$  such that

$$a(\kappa; u_H, v) = l(\kappa; v) \quad \forall v \in \mathcal{V}_H. \quad (4)$$

In practice, low-order Galerkin approximations to the Helmholtz equation involving high wave numbers are corrupted by large dispersion or pollution errors due to the loss of stability of  $a(\kappa; \cdot, \cdot)$ .

The wave number  $\kappa$  characterizes the oscillatory behavior of the exact solution: the larger the value of  $\kappa$ , the stronger the oscillations. Hence, the *rule of thumb* is used in computations: each wavelength is resolved by a certain fixed number of elements. For linear elements, the rule of thumb is stated as  $\kappa H = \text{constant} < 1$ . However, it is widely known that this rule is not sufficient to obtain reliable results for large  $\kappa$ . The dispersion error, which is related to the phase lag of the FE-solution, can only be controlled when  $\kappa^2 H/p$  is small. This undermines the practical utility of the Galerkin finite element method since severe mesh refinement is needed for large wave numbers. The performance of finite element computations at high wave numbers can be improved by using stabilization techniques. These techniques, which are extremely simple to implement, alleviate the dispersion effect of the finite element solution without requiring mesh refinement.

### 2.3. Galerkin/least-squares finite element approximation

Stabilized finite element methods were originally developed for fluid problems [22]. The first upwind type stabilized methods [23] subsequently gave rise to consistent stabilization techniques—ensuring that the exact solution  $u$  is also a solution of the weak stabilized problem. Among these techniques, the Galerkin/least-squares method (GLS) has been successfully applied both to fluids and to the Helmholtz equation [24, 25].

The idea behind stabilized finite element methods is to modify the variational form  $a(\kappa; \cdot, \cdot)$  (and, accordingly, the right-hand side) in such a way that the new variational form is *unconditionally stable*. In particular, the weak form of consistent stabilized methods is obtained from (3) by adding extra terms over the element interiors which are a function of the residual of the differential equation to ensure consistency. For instance, the additional stabilization terms of the GLS method are an element-by-element weighted least-squares formulation of the original differential equation.

The weak form of the GLS method associated with the partition  $\mathcal{T}_H$  is: find  $u \in \mathcal{U}$  such that

$$a(\kappa; u, v) + (\mathcal{L}u - f, \tau_H \mathcal{L}\tilde{v})_{\hat{\Omega}} = l(v) \quad \forall v \in \mathcal{V}, \quad (5)$$

where  $\mathcal{L}u = -\Delta u - \kappa^2 u$  is the indefinite Helmholtz operator,  $\hat{\Omega} = \bigcup_{n=1}^{n_{\text{el}}} \Omega_n$  denotes the union of element interiors of  $\mathcal{T}_H$ ,  $n_{\text{el}}$  being the number of elements of  $\mathcal{T}_H$  and  $(\cdot, \cdot)_{\hat{\Omega}}$  is the reduced  $L^2$  inner product, where integration is carried out only on the element interiors (i.e. the singularities at interelement boundaries are suppressed in the reduced inner product). Note that the GLS formulation depends on the stabilization parameter  $\tau_H$  which has to be properly defined to make the form on the l.h.s. unconditionally stable.

#### Remark 2

The exact solution  $u$  verifies Equation (5) for any choice of the stabilization parameter  $\tau_H$  since  $\mathcal{L}u - f = 0$ . That is, the GLS method is consistent for any choice of  $\tau_H$ .

The GLS finite element approximation of  $u$  is  $u_H \in \mathcal{U}_H$  such that

$$a_{\text{GLS}}(\kappa, \tau_H; u_H, v) = l_{\text{GLS}}(\kappa, \tau_H; v) \quad \forall v \in \mathcal{V}_H, \quad (6)$$

where

$$a_{\text{GLS}}(\kappa, \tau; u, v) := a(\kappa; u, v) + (\mathcal{L}u, \tau \mathcal{L}\tilde{v})_{\hat{\Omega}}$$

and

$$l_{\text{GLS}}(\kappa, \tau; v) := l(\kappa; v) + (f, \tau \mathcal{L}\tilde{v})_{\hat{\Omega}}.$$

Note that for the sake of simplicity, the same notation,  $u_H$ , for the Galerkin and GLS finite element approximations has been used. A different notation for the GLS/FE approximation, for instance  $u_H^{\text{GLS}}$ , would be more precise. However, since the error estimation strategy is valid for any approximation  $u_H \in \mathcal{V}_H$  of  $u$ , there is no need to distinguish between  $u_H$  and  $u_H^{\text{GLS}}$  or any other approximation. Moreover, note that  $\tau_H = 0$  results in the Galerkin approximation.

The stabilization parameter  $\tau_H$  is usually determined using discrete dispersion analyses with the aim of eliminating spurious dispersion of plane waves in a user-prescribed direction ( $\theta^{\text{opt}}$ ). That is,

the goal is that the GLS/FE approximation has no phase lag if the exact solution is a plane wave in the direction  $\theta^{\text{opt}}$ . Different definitions for the parameter  $\tau_H$  depending on the underlying size and topology of the mesh may be found in the literature [20, 25].

Unfortunately, it is not possible in general to design a stabilization parameter  $\tau_H$  that confers the ability of fully removing the dispersion error on the GLS method. The reason is twofold. First, a general signal consists of plane waves going in an infinite number of directions. Even if there are directionally prevalent components in this decomposition, they are not necessarily known *a priori*. Moreover, it is not clear whether the GLS method improves the approximations of solutions that are not dominant in the preferred direction. Second, the parameter  $\tau_H$  is derived for particular structured topology meshes. The optimal behavior obtained for some particular structured meshes (which are of limited use in real-life applications) is partially lost when general unstructured meshes are used.

#### 2.4. Matrix form

The Galerkin or GLS finite element approximation  $u_H$  is expressed in terms of the basis-functions  $\{N^j\}_{j=1, \dots, n_{\text{np}}}$  spanning  $\mathcal{U}_H$ , namely

$$u_H = \sum_{j=1}^{n_{\text{np}}} N^j u_H^j = \mathbf{N} \mathbf{u}_H,$$

where  $n_{\text{np}}$  is the number of nodes in the mesh,  $u_H^j$  is the complex nodal value associated with the mesh node  $\mathbf{x}^j$ ,  $\mathbf{N} = [N^1, N^2, \dots, N^{n_{\text{np}}}]$  and  $\mathbf{u}_H^T = [u_H^1, u_H^2, \dots, u_H^{n_{\text{np}}}]$ .

In the case of linear elements ( $p=1$ ),  $\mathcal{L}u_H$  reduces to  $\mathcal{L}u_H = -\kappa^2 u_H$  in  $\hat{\Omega}$ , and the matrix form of (6) reads

$$(\mathbf{K}_H - \mathbf{C}_H - \kappa^2 \mathbf{M}_H^{\tau_H}) \mathbf{u}_H = \mathbf{f}_H^{\tau_H} + \mathbf{f}_H^{\text{N}}, \quad (7)$$

where  $\mathbf{K}_H$ ,  $\mathbf{C}_H$  and  $\mathbf{M}_H^{\tau_H}$  are the so-called stiffness, damping and mass matrices, respectively

$$\mathbf{K}_H := \int_{\Omega} (\nabla \mathbf{N})^T (\nabla \mathbf{N}) d\Omega, \quad \mathbf{C}_H := \int_{\Gamma_{\text{R}}} \mathcal{M} \mathbf{N}^T \mathbf{N} d\Gamma, \quad \mathbf{M}_H^{\tau_H} := \sum_{n=1}^{n_{\text{el}}} \int_{\Omega_n} (1 - \tau_H \kappa^2) \mathbf{N}^T \mathbf{N} d\Omega$$

and the right-hand side vectors accounting for the source term and the Neumann boundary conditions are

$$\mathbf{f}_H^{\tau_H} := \sum_{n=1}^{n_{\text{el}}} \int_{\Omega_n} (1 - \tau_H \kappa^2) \mathbf{N}^T f d\Omega \quad \text{and} \quad \mathbf{f}_H^{\text{N}} := \int_{\Gamma_{\text{N}}} \mathbf{N}^T g d\Gamma.$$

In the particular case where the stabilization parameter  $\tau_H$  is constant in the elements of the mesh,  $\mathbf{M}_H^{\tau_H} = (1 - \tau_H \kappa^2) \mathbf{M}_H$  and  $\mathbf{f}_H^{\tau_H} = (1 - \tau_H \kappa^2) \mathbf{f}_H$ , where

$$\mathbf{M}_H := \int_{\Omega} \mathbf{N}^T \mathbf{N} d\Omega \quad \text{and} \quad \mathbf{f}_H := \int_{\Omega} \mathbf{N}^T f d\Gamma$$

are the standard (non-weighted) mass matrix and vector force. Besides, recall that  $\tau_H = 0$  results in the matrix form of the Galerkin finite element method (4).

### 3. A POSTERIORI ERROR ESTIMATION OF THE WAVE NUMBER

#### 3.1. Basics of error estimation of the wave number for the Galerkin method

It is well known that the error introduced in the numerical solution of wave problems has two different components: *interpolation error* and *pollution error*. The interpolation error is the classical

error arising in elliptic problems and pertains to the ability of the discretization to properly approximate the solution

$$e^{\text{int}} := u - u_H^{\text{int}} = u(\mathbf{x}) - \sum_{j=1}^{n_{\text{np}}} N^j(\mathbf{x})u(\mathbf{x}^j), \quad (8)$$

where  $u_H^{\text{int}}$  is the approximation of  $u$  in  $\mathcal{U}_H$  coinciding with  $u$  at the mesh nodes  $\mathbf{x}^j$ ,  $j = 1, 2, \dots, n_{\text{np}}$ . Thus, the pollution error is defined as:

$$e^{\text{pol}} := u_H^{\text{int}} - u_H = \sum_{j=1}^{n_{\text{np}}} N^j(\mathbf{x})(u(\mathbf{x}^j) - u_H^j).$$

In standard thermal and elasticity problems, the error in the finite element solution is equivalent to the interpolation error, and converges at the same rate. This error is local in nature because it may be reduced in a given zone by reducing the mesh size locally in this zone.

The *pollution* error, however, is especially relevant in the framework of Helmholtz problems due to the blowup of the inf-sup and continuity constants of the weak form when the wave number is large (i.e. the inf-sup constant tends to zero and the continuity constant tends to  $\infty$  as  $\kappa$  tends to  $\infty$ ). In transient wave problems, pollution is associated with the variation of the numerical wave speed with the wavelength. This phenomenon results in the dispersion of the different components of the total wave. In the steady Helmholtz problem, the word dispersion is also used and corresponds to the error in the numerical wave number  $\kappa_H$ , which is therefore identified with the pollution. In other words, the FE error is decomposed into two terms

$$\text{FE error} = u - u_H = e^{\text{int}} + e^{\text{pol}} = \text{Interpolation error} + \text{Dispersion/pollution error},$$

which, in the case of wave problems, behave completely differently (see Figure 1). It has been shown that the pollution term converges at a different rate, lower than the standard interpolation error.

The pollution error  $e^{\text{pol}}$  is related to the phase difference between the exact and FE solutions, that is, the difference between the wave number  $\kappa$  associated with  $u$  and the numerical wave number  $\kappa_H$  associated with  $u_H$ . Usually, the dispersion or pollution error is assessed by obtaining an approximation of the error in the wave number  $\kappa - \kappa_H$  instead of trying to measure the pollution error  $e^{\text{pol}}$  in some predefined norm.

*Remark 3*

For the sake of a simple presentation, the definition adopted here for the interpolation error  $e^{\text{int}} = u - u_H^{\text{int}}$  uses the nodal interpolation as described in (8). Other authors prefer substituting the nodal interpolation by the best approximation in  $\mathcal{U}_H$ , according to some natural norm associated

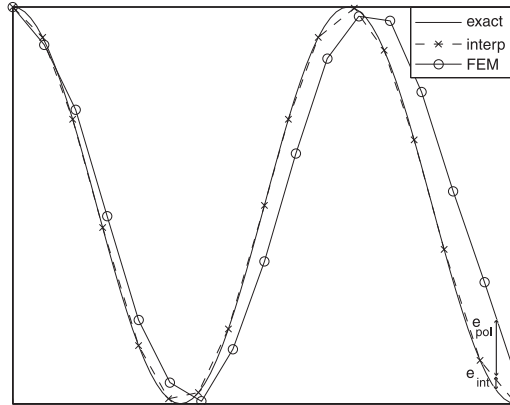


Figure 1. Illustration of the errors arising in the approximation of the Helmholtz equation. The exact solution (solid line, smooth) and interpolant (dashed line) coincide at the nodes, the FEM solution reproduces approximately the shape of the wave with a larger wavelength ( $\kappa_H < \kappa$ ).

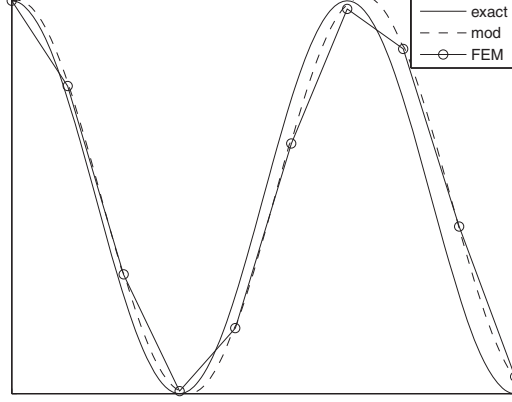


Figure 2. Illustration of the exact solution  $u$ , the approximate solution  $u_H$  and the auxiliary solution  $u_H^m$  coinciding with  $u_H$  at the nodes and sharing its wave number  $\kappa_H$ .

with the problem. For instance, in the present case of the Helmholtz equation, the norm selected can be either the  $H_1$  norm or seminorm. This choice (using a least-squares-like fitting instead of a simple nodal interpolation) provides a definition of the pollution error such that, in the case of the Poisson equation, the pollution vanishes (which is obviously a desirable property). Here, this definition is only used to recall the pollution concept and, consequently, this choice has no practical relevance.

*A priori* error estimates assess the dispersion error by means of providing a closed formula of the numerical wave number  $\kappa_H$ . Recently, a new approach to *a posteriori* estimate the dispersion error, thus using the information given by  $u_H$ , has been developed [17].

The key idea is to define an auxiliary solution  $u_H^m \in \mathcal{U}$  having the same wave number as  $u_H$  and from which to recover the value of  $\kappa_H$ . Intuitively,  $u_H^m \in \mathcal{U}$  is the best solution of the Helmholtz equation (3) associated with a wave number  $\kappa_H$  matching  $u_H$  at the nodes of the mesh, see Figure 2.

To fix the ideas, consider the one-dimensional Helmholtz equation in  $\Omega=(0, 1)$  with boundary conditions  $u(0)=1$  and  $u'(1)=i\kappa u(1)$ . This simple problem admits the analytical solution  $u(x)=e^{i\kappa x}$ . Then, given a uniform finite element mesh and its associated FE approximation  $u_H$ , it turns out that there exists a wave number  $\kappa_H$  such that the solution of Equation (3) associated with  $\kappa_H$ ,  $u_H^m=e^{i\kappa_H x}$ , exactly fulfills the equations of the Galerkin method (7) associated with the interior nodes. This wave number is

$$\kappa_H = \frac{1}{H} \arccos\left(\frac{1-(\kappa H)^2/3}{1+(\kappa H)^2/6}\right) \approx \kappa - \frac{1}{24}\kappa^3 H^2 + \frac{3}{640}\kappa^5 H^4 + \mathcal{O}(\kappa^7 H^6) \quad (9)$$

see [2]. The verification of Equation (7) associated with the interior nodes enforces that the auxiliary solution  $u_H^m$  shares the same wave number as  $u_H$ , although this does not guarantee that  $u_H^m$  matches  $u_H$  exactly at the nodes of the mesh, due to the influence of the Robin boundary conditions. However, the difference between  $u_H^m$  and  $u_H$  at the nodes of the mesh is nearly negligible. Thus, for this particular problem, a very good measure of the dispersion error can be computed as

$$E = \kappa - \kappa_H \approx \frac{1}{24}\kappa^3 H^2 + \mathcal{O}(\kappa^5 H^4). \quad (10)$$

Unfortunately, in general, it is not possible to determine  $u_H^m \in \mathcal{U}$  verifying (3) for a suitable wave number  $\kappa_H \in \mathbb{R}$  and concurrently fulfilling the equations of the Galerkin method associated

with the interior nodes. However, a slight modification of this idea yields a proper definition for  $u_H^m$ . Specifically,  $u_H^m \in \mathcal{U}$  and  $\kappa_H \in \mathbb{R}$  are such that:

- $u_H^m \in \mathcal{U}$  coincides with  $u_H$  at the nodes of the mesh (that is  $u_H^m(\mathbf{x}^j) = u_H(\mathbf{x}^j)$  for  $j = 1, 2, \dots, n_{\text{np}}$ )
- for a given  $\kappa_H$ ,  $u_H^m \in \mathcal{U}$  is such that

$$a(\kappa_H; u_H^m, v) = l(\kappa_H; v) \quad \forall v \in \mathcal{V}_0, \quad (11)$$

where a subspace of  $\mathcal{V}$  is introduced that can be intuitively defined by

$$\mathcal{V}_0 := \{v \in \mathcal{V}, v(\mathbf{x}^j) = 0, j = 1, 2, \dots, n_{\text{np}}\}.$$

A more rigorous definition would require dealing with point values in  $H^1$  similarly as is done when the Dirichlet boundary reduces to a point in a standard boundary value problem. However, in this paper, for all practical purposes, this technicality can be skipped because, as detailed in (13), the methodology is actually referred to a reference space of continuous functions.

- $\kappa_H$  and  $u_H^m$  minimize the norm of the residual functional

$$\|R(\kappa_H, u_H^m; \cdot)\|_* := \max_{v \in \mathcal{H}_0^1 \setminus \{0\}} \frac{R(\kappa_H, u_H^m; v)}{\|v\|}, \quad (12)$$

where  $R(\kappa_H, u_H^m; \cdot) := l(\kappa_H; \cdot) - a(\kappa_H; u_H^m, \cdot)$ ,  $\mathcal{H}_0^1 := \{v \in \mathcal{H}^1(\Omega), v|_{\partial\Omega} = 0\}$  and  $\|v\|$  is some norm defined on  $\mathcal{H}^1(\Omega)$  to be given by the user.

Note that the values of  $u_H^m$  on the boundary of  $\Omega$  do not affect the norm of the residual  $\|\cdot\|_*$ . This definition is used to minimize the influence of the errors due to the boundary conditions (which are considered to be a part of the interpolation error and not of the dispersion error) in the assessment of the dispersion error. Also note that the condition enforcing that  $u_H^m$  and  $u_H$  share the same phase lag, i.e. fulfilling of the equations of the Galerkin method associated with the interior nodes, is replaced by the more simple and equivalent condition of matching  $u_H$  at the nodes of the mesh.

In a compact form,  $\kappa_H$  and  $u_H^m$  are the solutions of the following constrained optimization problem

$$\begin{aligned} (\kappa_H, u_H^m) &= \arg \min_{\substack{\kappa^m \in \mathbb{R} \\ u^m \in \mathcal{U}}} \|R(\kappa^m, u^m; \cdot)\|_* \\ \text{subject to} \quad &a(\kappa^m; u^m, v) = l(\kappa^m; v) \quad \forall v \in \mathcal{V}_0 \\ &u^m(\mathbf{x}^j) = u_H(\mathbf{x}^j), \quad j = 1, 2, \dots, n_{\text{np}}. \end{aligned}$$

The relation between the finite element solution  $u_H$  and the modified solution  $u_H^m$  allows to state that the numerical wave number associated with  $u_H$ , coincides with the wave number associated with the solution  $u_H^m$ . That is, the finite element solutions  $u_H$  and  $u_H^m$  share the same phase lag and therefore the dispersion error associated with  $u_H$  is  $E = \kappa - \kappa_H$ .

It is worth noting that the definition of the numerical wave number through the modified solution  $u_H^m$  is not applicable as a practical error estimation strategy, since  $\kappa_H$  and  $u_H^m$  are even more difficult to compute than the exact solution  $u$ . Nevertheless, this rationale is used as a starting point to obtain a fully computable estimate for the dispersion error, by just introducing two simple modifications.



### 3.2. Practical a posteriori explicit error estimate for the wave number

First, the finite-dimensional reference spaces  $\mathcal{U}_h$  and  $\mathcal{V}_h$  much finer than  $\mathcal{U}_H$  and  $\mathcal{V}_H$  are introduced. These spaces yield to the following approximations of  $\kappa_H$  and  $u_H^m$

$$\begin{aligned} (\kappa_H[h], u_H^m[h]) &= \arg \min_{\substack{\kappa^m \in \mathbb{R} \\ u^m \in \mathcal{U}_h}} \|R(\kappa^m, u^m; \cdot)\|_{*,h} \\ \text{subject to } a(\kappa^m; u^m, v) &= l(\kappa^m; v) \quad \forall v \in \mathcal{V}_h \cap \mathcal{V}_0 \\ u^m(\mathbf{x}^j) &= u_H(\mathbf{x}^j), \quad j = 1, 2, \dots, n_{\text{np}} \end{aligned} \quad (13)$$

and

$$\|R(\kappa_H[h], u_H^m[h]; \cdot)\|_{*,h} := \max_{\substack{v \in \mathcal{V}_h \setminus \{0\} \\ v|_{\partial\Omega} = 0}} \frac{R(\kappa_H[h], u_H^m[h]; v)}{\|v\|}.$$

If the finite element mesh  $\mathcal{V}_h$  is sufficiently fine, one expects that  $u_H^m \approx u_H^m[h]$  and therefore  $\kappa_H[h] \approx \kappa_H$ . If the finite element mesh  $\mathcal{V}_h$  is not fine enough, as mentioned in [17], a correction factor has to be applied to recover a good approximation of  $\kappa_H$  from  $\kappa_H[h]$ , i.e.  $\kappa_H[0] = c_f \kappa_H[h]$ , where  $c_f$  is the correction factor based on a Richardson extrapolation technique.

Second, since the computation of  $\kappa_H[h]$  and  $u_H^m[h]$  is still unaffordable in practical applications another simplification is introduced. An approximation of  $u_H^m[h]$  in  $\mathcal{U}_h$ , denoted by  $u^*$ , is obtained by post-processing  $u_H$ . In general, the approximation  $u^*$  is not obtained solving Equation (11) for some  $\kappa_H$  and thus the computation of  $\kappa_H$  is independent. Indeed  $u^*$  does not verify

$$a(\kappa_H[h]; u^*, v) = l(\kappa_H[h]; v) \quad \forall v \in \mathcal{V}_h \cap \mathcal{V}_0$$

and is therefore no longer linked with the computation of  $\kappa_H[h]$ . Once this approximation  $u^*$  is computed, the wave number  $\kappa_H[h]$  is approximated by  $\kappa^*$  solution of

$$\kappa^* = \arg \min_{\kappa^m \in \mathbb{R}} \|R(\kappa^m, u^*; \cdot)\|_{*,h}.$$

It is worth noting that if the functional norm  $\|\cdot\|$  in (12) is such that its discrete counterpart is the Euclidean norm in  $\mathbb{C}^{n_{\text{np}}}$ , then the norm of the residual  $\|R(\kappa^m, u^*; \cdot)\|_{*,h}$  is a function depending only on the scalar variable  $\kappa^m$  and may be computed as

$$\|R(\kappa^m, u^*; \cdot)\|_{*,h} = \sqrt{\mathbf{r}(\kappa^m, u^*)' \mathbf{r}(\kappa^m, u^*)},$$

where

$$\begin{aligned} \mathbf{r}(\kappa^m, u^*) &:= \mathbf{B}_0((\mathbf{K}_h - \mathbf{C}_h - (\kappa^m)^2 \mathbf{M}_h) \mathbf{u}^* - \mathbf{f}_h - \mathbf{f}_h^N) \\ &= \mathbf{B}_0((\mathbf{K}_h - (\kappa^m)^2 \mathbf{M}_h) \mathbf{u}^* - \mathbf{f}_h) \end{aligned}$$

is the residual associated with the interior nodes of the fine  $h$ -mesh, the approximation  $u^*$  and the wave number  $\kappa^m$ . The symbol  $'$  stands for the conjugated transpose, that is  $v' \equiv \bar{v}^T$ , and  $\mathbf{B}_0$  is a diagonal matrix on the  $h$ -mesh with ones in the positions associated with the interior nodes and zero elsewhere. That is, the matrix  $\mathbf{B}_0$  sets the values of the residual at the boundary (either Dirichlet, Neumann or Robin) to zero.

Thus, for a given value of  $u^* \approx u_H^m[h]$ , the wave number  $\kappa^*$  is the parameter of the modified problem that better accommodates  $u^*$ . In practice,  $\kappa^*$  is determined minimizing the squared norm of the residual, namely

$$\kappa^* := \arg \min_{\kappa^m \in \mathbb{R}} \|R(\kappa^m, u^*; \cdot)\|_{*,h} = \arg \min_{\kappa^m \in \mathbb{R}} \sqrt{\mathbf{r}' \mathbf{r}} = \arg \min_{\kappa^m \in \mathbb{R}} \mathbf{r}' \mathbf{r}. \quad (14)$$

Note that given  $\mathbf{u}^*$ , the squared residual norm  $\mathbf{r}' \mathbf{r}$  is a fourth degree polynomial in  $\kappa^m$  and thus  $\kappa^*$  is computed explicitly, see [17] for the computational details.

In short, the approximation  $\kappa^*$  of the numerical wave number  $\kappa_H$  is assessed by first post-processing the finite element solution  $u_H$  to compute  $u^*$  and then explicitly solving Equation (14). The computable *a posteriori* error estimate for the wave number is then

$$E^* := \kappa - \kappa^*.$$

### 3.3. Assessment of the numerical wave number for stabilized formulations

The dispersion error associated with a stabilized finite element approximation of  $u$  may be assessed using the same methodology detailed for the standard Galerkin approximation. Given the GLS/FE approximation  $u_H$ , a post-processing technique is used to compute an approximation  $u^*$  of the solution  $u_H^m[h]$  of (13). Then, the wave number  $\kappa_H$  is approximated by  $\kappa^*$  solution of (14).

However, the use of stabilized formulations also for the fine mesh solutions in (13) allows to improve the quality of the estimates. Note that the accuracy of the estimate  $\kappa^*$  relies on two facts: first on the quality of the approximation  $u^*$  of  $u_H^m[h]$  and second on the quality of the approximation  $u_H^m[h]$  of  $u_H^m$ . The quality of  $u^*$  depends on the post-processing strategy which will be discussed in the following section. The quality of  $u_H^m[h]$ , on the other hand, depends on the size  $h$  of the reference mesh  $\mathcal{V}_h$ . In fact, it depends on the ratio of  $\kappa$  versus  $h$  since for large values of  $\kappa$  the reference mesh should be finer in order to get good approximations of  $u_H^m$ . Thus, for large wave numbers, the discrete approximation  $u_H^m[h]$  will only be a good approximation of  $u_H^m$  if the reference mesh is taken remarkably fine.

A simple workaround that avoids dealing with fine reference meshes is to stabilize the problem associated with  $u_H^m[h]$ . That is, for a given finite element approximation (either stabilized or not), the stabilized approximation  $u_H^m[h, \tau_h]$  is the solution of

$$\begin{aligned} (\kappa_H[h, \tau_h], u_H^m[h, \tau_h]) &= \arg \min_{\substack{\kappa^m \in \mathbb{R} \\ u^m \in \mathcal{U}_h}} \|R_{\text{GLS}}(\kappa^m, \tau_h, u^m; \cdot)\|_{*,h} \\ \text{subject to} \quad a_{\text{GLS}}(\kappa^m, \tau_h; u^m, v) &= l_{\text{GLS}}(\kappa^m, \tau_h; v) \quad \forall v \in \mathcal{V}_h \cap \mathcal{V}_0 \\ u^m(\mathbf{x}^j) &= u_H(\mathbf{x}^j), \quad j = 1, 2, \dots, n_{\text{np}} \end{aligned} \quad (15)$$

where

$$R_{\text{GLS}}(\kappa^m, \tau_h, u^m; v) := l_{\text{GLS}}(\kappa^m, \tau_h; v) - a_{\text{GLS}}(\kappa^m, \tau_h; u^m, v).$$

This modification yields to the following strategy to assess the error in the numerical wave number:

- (1) compute  $u^*$  approximation of  $u_H^m[h, \tau_h]$  by post-processing  $u_H$ ,
- (2) compute the approximation  $\kappa^*[\tau_h]$  solution of

$$\kappa^*[\tau_h] := \arg \min_{\kappa^m \in \mathbb{R}} \|R_{\text{GLS}}(\kappa^m, \tau_h, u^*; \cdot)\|_{*,h} = \arg \min_{\kappa^m \in \mathbb{R}} \mathbf{r}_{\text{GLS}}(\kappa^m)' \mathbf{r}_{\text{GLS}}(\kappa^m), \quad (16)$$

where

$$\mathbf{r}_{\text{GLS}}(\kappa^m) := \mathbf{B}_0((\mathbf{K}_h - (\kappa^m)^2 \mathbf{M}_h^{\tau_h}) \mathbf{u}^* - \mathbf{f}_h^{\tau_h}).$$

The explicit dependence of the vector  $\mathbf{r}_{\text{GLS}}$  on  $\tau_h$  and  $u^*$ ,  $\mathbf{r}_{\text{GLS}}(\kappa^m, \tau_h, u^*)$ , is omitted for simplicity of presentation. Note that the matrix  $\mathbf{M}_h^{\tau_h}$  and the vector  $\mathbf{f}_h^{\tau_h}$  depend explicitly on the wave number  $\kappa^m$  and also implicitly via the stabilization parameter  $\tau_h$ . Therefore, the dependency of  $\mathbf{r}'_{\text{GLS}} \mathbf{r}_{\text{GLS}}$  with respect to the wave number  $\kappa^m$  is no longer a fourth-order polynomial and the solution of (16) may not be computed explicitly in general.

### 3.4. Computation of the wave number $\kappa^*[\tau_h]$

In order to detail the computation of  $\kappa^*[\tau_h]$  verifying (16) in a simple manner, the stabilization parameter  $\tau_h$  is assumed constant on the elements of the fine mesh. In this case,  $\kappa^*[\tau_h]$  is the solution of (16) where

$$\mathbf{r}_{\text{GLS}}(\kappa^m) := \mathbf{B}_0(\mathbf{K}_h \mathbf{u}^* - (\kappa^m)^2 \mathbf{M}_h \mathbf{u}^* + \tau_h (\kappa^m)^4 \mathbf{M}_h \mathbf{u}^* - \mathbf{f}_h + \tau_h (\kappa^m)^2 \mathbf{f}_h)$$

and  $\tau_h$  depends non-linearly on  $\kappa^m$ . For instance, to minimize the phase lag on the  $x$ -direction for a structured regular quadrilateral mesh, Harari and Magoulès [20] proposed the use of

$$\tau_h(\kappa^m) = \frac{1}{(\kappa^m)^2} - \frac{6}{(\kappa^m)^4 h^2} \frac{1 - \cos(\kappa^m h)}{2 + \cos(\kappa^m h)}.$$

Thus, the computation of  $\kappa^*[\tau_h]$  requires solving a scalar root-finding problem.

Three different options have been considered in this work to approximate  $\kappa^*[\tau_h]$ . The first approach is to compute an approximation of  $\kappa^*[\tau_h]$  using an algorithm to numerically approximate the minimum of  $F(\kappa^m) := \mathbf{r}_{\text{GLS}}(\kappa^m)' \mathbf{r}_{\text{GLS}}(\kappa^m)$ . Namely, a root-finding method on the derivative of  $F(\kappa^m)$  is used taking as initial guess  $\kappa^m = \kappa$ . This approximation is taken to represent the exact value  $\kappa^*[\tau_h]$  since its accuracy can be controlled by the end-user through adjusting the tolerance of the root-finding method.

The second approach assumes that  $\tau_h$  does not vary considerably when varying the parameter  $\kappa^m$ . In this case, the dependency of the parameter  $\tau_h$  with respect to  $\kappa^m$  is removed by setting the value of  $\tau_h(\kappa^m) = \tau_h^\kappa$ , where  $\tau_h^\kappa := \tau_h(\kappa)$ , and the approximation of  $\kappa^*[\tau_h]$  is denoted by  $\kappa^*[\tau_h^\kappa]$ . Note that  $\tau_h^\kappa$  denotes the value of the parameter  $\tau_h$  associated with the wave number  $\kappa$ . Doing this approximation, the residual  $\mathbf{r}_{\text{GLS}}(\kappa^m)$  is approximated by a fourth-order polynomial on  $\kappa^m$

$$\mathbf{r}_{\text{GLS}}(\kappa^m) \approx \mathbf{a}_0 + \mathbf{a}_2(\kappa^m)^2 + \mathbf{a}_4(\kappa^m)^4 \quad (17)$$

for  $\mathbf{a}_0 = \mathbf{B}_0(\mathbf{K}_h \mathbf{u}^* - \mathbf{f}_h)$ ,  $\mathbf{a}_2 = \mathbf{B}_0(-\mathbf{M}_h \mathbf{u}^* + \tau_h^\kappa \mathbf{f}_h)$  and  $\mathbf{a}_4 = \tau_h^\kappa \mathbf{B}_0 \mathbf{M}_h \mathbf{u}^*$ . The minimization of the squared residual  $F(\kappa^m)$  is then reduced to find the critical points of  $F(\kappa^m)$  which is equivalent to find the solutions of

$$\frac{dF}{d\kappa^m} = 2\kappa^m(\mathbf{c}_0 + 2\mathbf{c}_2(\kappa^m)^2 + 3\mathbf{c}_4(\kappa^m)^4 + 4\mathbf{c}_6(\kappa^m)^6) = 0, \quad (18)$$

where  $\mathbf{c}_0 = \mathbf{a}'_0 \mathbf{a}_2 + \mathbf{a}'_2 \mathbf{a}_0$ ,  $\mathbf{c}_2 = \mathbf{a}'_0 \mathbf{a}_4 + \mathbf{a}'_2 \mathbf{a}_2 + \mathbf{a}'_4 \mathbf{a}_0$ ,  $\mathbf{c}_4 = \mathbf{a}'_2 \mathbf{a}_4 + \mathbf{a}'_4 \mathbf{a}_2$ ,  $\mathbf{c}_6 = \mathbf{a}'_4 \mathbf{a}_4$ . Although Equation (18) may have seven real solutions,  $\kappa^*[\tau_h^\kappa]$  is defined to be the solution of (18) closer to  $\kappa$ . Thus, ruling out the trivial solution  $\kappa^m = 0$ ,  $\kappa^*[\tau_h^\kappa]$  is computed by first finding the roots of the bicubic polynomial appearing in Equation (18), which is equivalent to finding the three solutions  $\bar{\kappa}$  of

$$\mathbf{c}_0 + 2\mathbf{c}_2 \bar{\kappa} + 3\mathbf{c}_4 \bar{\kappa}^2 + 4\mathbf{c}_6 \bar{\kappa}^3 = 0$$

and then setting  $\kappa^*[\tau_h^\kappa]$  to be the value of  $\sqrt{\bar{\kappa}}$  nearer to  $\kappa$ , see [17] for the computational details. Thus, the assumption  $\tau_h(\kappa^m) = \tau_h^\kappa$  yields to a simple and explicit algorithm to approximate the exact value of  $\kappa^*[\tau_h^\kappa]$ .

Finally, the third approach directly applies the strategy presented in [17] by considering that the terms added by the GLS method are constant with respect to  $\kappa^m$ , that is, not only the parameter  $\tau_h$  is set to  $\tau_h^\kappa$ , but also the  $(\kappa^m)^2$  associated with the GLS method is set to  $\kappa^2$ . In this way, the residual is approximated by the quadratic function

$$\mathbf{r}_{\text{GLS}}(\kappa^m) \approx \mathbf{B}_0(\mathbf{K}_h \mathbf{u}^* - (\kappa^m)^2 \mathbf{M}_h \mathbf{u}^* + \tau_h^\kappa \kappa^2 (\kappa^m)^2 \mathbf{M}_h \mathbf{u}^* - \mathbf{f}_h + \tau_h^\kappa \kappa^2 \mathbf{f}_h)$$

and the minimization of the fourth-order polynomial  $F(\kappa^m)$  which allows to compute the approximation of  $\kappa^*[\tau_h]$  is done by using the technique detailed in [17].

As will be seen in the numerical examples, the second option yields a fairly good approximation of the exact solution of the one-dimensional non-linear optimization problem (16). The practical and straightforward algorithm to estimate the dispersion error using this second option is summarized in the box shown in Figure 3.

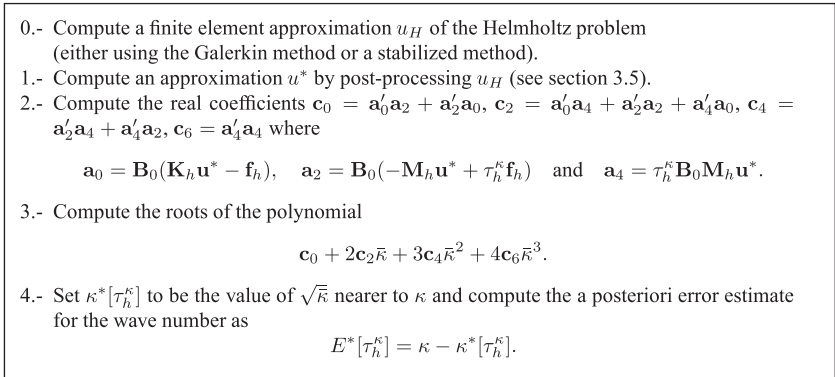


Figure 3. Practical and straightforward algorithm to estimate the dispersion error.

*Remark 4*

Note that the second step of the previous procedure requires to compute the coefficients  $\mathbf{c}_0$ ,  $\mathbf{c}_2$ ,  $\mathbf{c}_4$  and  $\mathbf{c}_6$  associated with the residual  $\mathbf{r}_{\text{GLS}}$ . These coefficients depend on the stabilization parameter  $\tau_h^\kappa$  which in turn depends on a user-prescribed direction  $\theta^{\text{opt}}$  which will be denoted in the following by  $\theta_h^{\text{opt}}$ . In the case that  $u_H$  is computed using the standard Galerkin method, it is not natural to define a direction  $\theta_h^{\text{opt}}$ . However, information about the prevalent wave direction of the exact solution can be used if available. If  $u_H$  is computed using the GLS method with wave direction  $\theta^{\text{opt}}$ , the estimates may be computed using  $\theta_h^{\text{opt}} = \theta^{\text{opt}}$  or again, if information about the exact solution is available, this parameter may be set to adjust the prevalent wave direction of the exact solution. The choice of this parameter will be further discussed in the numerical examples.

3.5. Enhanced solution  $u^*$  by post-processing  $u_H$

The quality of the estimate  $\kappa^*$  depends on the quality of the approximation  $u^*$  of  $u_H^m[h] \in \mathcal{U}_h$  (respectively  $u_H^m[h, \tau_h]$ ). The idea proposed here is to build up an inexpensive approximation using a post-processing technique standard in error estimation analysis [26, 27] and likely having all its features. The post-processing technique starts from the finite element solution  $u_H \in \mathcal{U}_H$  and computes an approximation  $u^*$  of  $u_H^m[h]$  in  $\mathcal{U}_h$ .

Reference [17] presents a procedure to compute  $u^*$  following the work of Calderón and Díez [28]. For each element of the  $H$ -mesh,  $\Omega_n$ , the patch of elements surrounding  $\Omega_n$  is considered and it is denoted by  $\omega_n$ . In this patch, the values of  $\mathbf{u}_H$  at the nodes of the  $H$ -mesh are used as input data and a polynomial is fitted using a constrained least-squares technique. That is, in a two-dimensional setting, for a given polynomial degree  $q$ , a complex-valued polynomial field

$$p(\mathbf{x}) = \sum_{k+l \leq q} p_{kl} x^k y^l$$

is determined from the following constrained least-squares problem

$$\begin{aligned} \min_{p_{kl} \in \mathbb{C}} \quad & \sum_{\mathbf{x}^j \in \omega_n} |u_H^j - p(\mathbf{x}^j)|^2 \\ \text{restricted to} \quad & p(\mathbf{x}^j) = u_H^j \quad \text{for } \mathbf{x}^j \in \Omega_n, \end{aligned}$$

where  $|\cdot|$  denotes the modulus of a complex number. Note that the real and imaginary parts of  $p(\mathbf{x})$  can be computed separately. The real part of  $p(\mathbf{x})$  (and analogously its imaginary part) may

be found solving the real-valued constrained optimization

$$\begin{aligned} \min_{\Re(p_{ki}) \in \mathbb{R}} \quad & \sum_{\mathbf{x}^j \in \omega_n} |\Re(u_H^j) - \Re(p(\mathbf{x}^j))|^2 \\ \text{restricted to} \quad & \Re(p(\mathbf{x}^j)) = \Re(u_H^j) \quad \text{for } \mathbf{x}^j \in \Omega_n. \end{aligned}$$

Once the polynomial is obtained in  $\omega_n$  it is evaluated to find the nodal values of  $u^*$  in the nodes of the  $h$ -mesh lying in element  $\Omega_n$  of the  $H$ -mesh. This approach allows recovering the curvatures of the solution coinciding with  $u_H$  at the nodes where it is computed.

This simple and straightforward strategy provides fairly good results. However, this approach does not use specific information about the differential operator or the exact solution. The use of analytical information about the natural solutions of the differential operator yields an alternative approach to compute  $u^*$ .

The approach to compute  $u^*$  also requires solving a local constrained least-squares problem for each element  $\Omega_n$ . Instead of using a polynomial representation for  $u^*|_{\omega_n}$  an exponential fitting is used. This is a natural choice because the exact solution of the 2D homogeneous Helmholtz equation is an infinite sum of plane waves of the form  $Ae^{i\mathbf{k}\cdot\mathbf{x}}$ , where  $\mathbf{k} = \kappa[\cos(\theta), \sin(\theta)]$ .

Thus, in each patch  $\omega_n$ ,  $u_H$  is approximated by an exponential field of the form

$$A(\mathbf{x})e^{ip(\mathbf{x})},$$

where  $A(\mathbf{x})$  and  $p(\mathbf{x})$  are polynomial fields representing the amplitude and wave direction. The fields  $A(\mathbf{x})$  and  $p(\mathbf{x})$  are determined by a constrained least-squares criterion and hence, they are taken as those minimizing

$$\begin{aligned} \min \quad & \sum_{\mathbf{x}^j \in \omega_n} |u_H^j - A(\mathbf{x}^j)e^{ip(\mathbf{x}^j)}|^2 \\ \text{restricted to} \quad & A(\mathbf{x}^j)e^{ip(\mathbf{x}^j)} = u_H^j \quad \text{for } \mathbf{x}^j \in \Omega_n. \end{aligned}$$

Using a standard technique to linearize the exponential least-squares fitting transforms the previous problem into an equivalent linear constrained least-squares problem

$$\begin{aligned} \min \quad & \sum_{\mathbf{x}^j \in \omega_n} |\ln(u_H^j) - \ln(A(\mathbf{x}^j)e^{ip(\mathbf{x}^j)})|^2 \\ \text{restricted to} \quad & \ln(A(\mathbf{x}^j)e^{ip(\mathbf{x}^j)}) = \ln(u_H^j) \quad \text{for } \mathbf{x}^j \in \Omega_n. \end{aligned}$$

Splitting the real and imaginary part of the previous problem yields a simple strategy to compute  $\ln(A(\mathbf{x}))$  and  $p(\mathbf{x})$  independently using a restricted least-squares fitting, namely:

$$\begin{aligned} \min \quad & \sum_{\mathbf{x}^j \in \omega_n} |\ln(|u_H^j|) - \ln(A(\mathbf{x}^j))|^2 \\ \text{restricted to} \quad & \ln(A(\mathbf{x}^j)) = \ln(|u_H^j|) \quad \text{for } \mathbf{x}^j \in \Omega_n \end{aligned}$$

and

$$\begin{aligned} \min \quad & \sum_{\mathbf{x}^j \in \omega_n} |\arg(u_H^j) - p(\mathbf{x}^j)|^2 \\ \text{restricted to} \quad & p(\mathbf{x}^j) = \arg(u_H^j) \quad \text{for } \mathbf{x}^j \in \Omega_n, \end{aligned}$$

where  $\arg(\cdot)$  denotes the argument of a complex number and a polynomial fitting of  $\ln(A(\mathbf{x}))$  and  $p(\mathbf{x})$  is considered.

The only intricate part of this strategy involves the input data,  $\arg(u_H^j)$ , of the least-squares problem for  $p(\mathbf{x})$ . The non-unique arguments associated with the data  $u_H^j$  have to be carefully selected so that the polynomial fitting yields proper results.

## 4. NUMERICAL EXAMPLES

The strategy to assess the error in the wave number presented in the previous sections is validated in four numerical examples. The performance of the estimates of the dispersion error is shown both for Galerkin and GLS approximations. Moreover, the influence of the post-processing technique yielding  $u^*$  in the resulting effectivity is also discussed.

The finite element approximations are computed using triangular and quadrilateral meshes of linear (resp. bilinear) elements,  $p=1$ . Different definitions of the stabilization parameter  $\tau_H$  are used to compute the GLS approximations depending on the underlying topology of the mesh. In particular, for structured and unstructured quadrilateral meshes the following definition of the parameter, designed to minimize the dispersion error of plane wave in the direction  $\theta^{\text{opt}}$  on cartesian meshes, is used [20, 25]:

$$\tau_H = \frac{1}{\kappa^2} \left( 1 - \frac{6}{(\kappa h)^2} \left( \frac{1 - \cos(\kappa h \cos \theta^{\text{opt}})}{2 + \cos(\kappa h \cos \theta^{\text{opt}})} + \frac{1 - \cos(\kappa h \sin \theta^{\text{opt}})}{2 + \cos(\kappa h \sin \theta^{\text{opt}})} \right) \right).$$

For triangular meshes, the definition derived for hexagonal meshes, namely

$$\tau_H = \frac{1}{\kappa^2} \left( 1 - \frac{8}{(\kappa h)^2} \frac{3 - f(\kappa h, \theta^{\text{opt}})}{3 + f(\kappa h, \theta^{\text{opt}})} \right),$$

where  $f(\kappa h, \theta) = \cos(\kappa h \cos \theta) + 2 \cos(\kappa h \cos \theta / 2) \cos(\sqrt{3} \kappa h \sin \theta / 2)$  is used because it provides good results also for unstructured meshes.

For non-uniform meshes, the stabilization parameter is not constant over the whole mesh. In each element  $\Omega_n$  a different stabilization parameter is used depending on its characteristic element size  $h_n$ . This characteristic element size is taken as the smallest side of the element both for quadrilateral and triangular meshes.

As mentioned in Section 2.3 the parameter  $\tau_H$  depends on a user-prescribed direction  $\theta^{\text{opt}}$ . The influence of the selection of this direction in the reduction of the dispersion error is studied in the following examples.

### 4.1. Example 1: 1D strip

The first example models a plane wave propagating in the  $x$ -direction in a two-dimensional rectangular domain, with length  $L=1$  and width  $V=\sqrt{3}/8$ , see Figure 4. The boundary conditions are specified in order to yield the exact solution  $u(x, y) = e^{i\kappa x}$ : Dirichlet on the left-hand side, Robin on the right-hand side and Neumann homogeneous on the upper and lower sides to maintain the one-dimensional character of the solution. That is, the data entering in Equation (2) are  $\bar{u} = 1$  on  $x=0$ ,  $\mathcal{M}u = i\kappa u$  on  $x=1$  and  $g=0$  on  $y=0$  and  $y=\sqrt{3}/8$ . The performance of the Galerkin and GLS finite element solutions is studied for  $\kappa=8\pi$ . Owing to the 1D character of the problem, the stabilization angle used in all the GLS computations (both for the coarse and fine meshes) is set to 0, that is,  $\theta^{\text{opt}} = \theta_h^{\text{opt}} = 0$ . Note that the solution of the problem is independent of the width of the domain  $V$  and the value  $\sqrt{3}/8$  has been selected in order to accommodate a hexagonal triangular mesh.

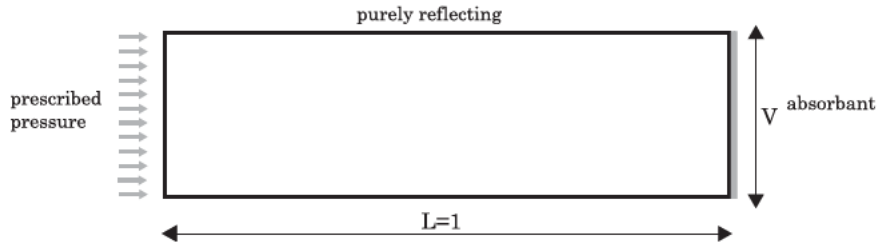


Figure 4. Example 1; 1D strip: problem setup.

Table I. Example 1: assessment of the dispersion error for a uniform coarse quadrilateral mesh ( $24 \times 2$  elements) a successively refined reference meshes for the Galerkin approximations of the solution.

$h$	Galerkin				$E^{\text{pri}} = 1.02211$		
	$E[h]$	$c_f E[h]$	$E[h, \tau_h]$	$c_f E^*$	Option 1 $E^*[\tau_h]$	Option 2 $E^*[\tau_h^k]$	Option 3
$H/2$	0.76790	1.02387	1.02211	1.01428	1.01469	1.01486	1.03682
$H/4$	0.95869	1.02261	1.02211	1.01428	1.01469	1.01486	1.03682
$H/8$	1.00627	1.02224	1.02211	1.01227	1.01232	1.01232	1.01368
$H/16$	1.01815	1.02214	1.02211	1.01214	1.01215	1.01215	1.01249
$H/32$	1.02112	1.02212	1.02211	1.01210	1.01210	1.01210	1.01218
$H/64$	1.02186	1.02211	1.02211	1.01208	1.01208	1.01208	1.01210

The truth error estimates (left) are computed using the fully non-linear solution yielding to  $E[h]$  and  $E[h, \tau_h]$ . The exponential post-processed solution (right)  $u^*$  obtained from  $u_H$  and then different options are used to recover the wave number  $\kappa^*$  associated with  $u^*$  only for the Galerkin approximation.

First, the influence of the selection of the finite reference mesh associated with  $\mathcal{V}_h$  is studied. If the finite element mesh  $\mathcal{V}_h$  is sufficiently fine, one expects that  $u_H^m \approx u_H^m[h, \tau_h] \approx u_H^m[h]$  and therefore  $\kappa_H \approx \kappa_H[h, \tau_h] \approx \kappa_H[h]$ . If the finite element mesh  $\mathcal{V}_h$  is not fine enough, one should apply a correction factor to  $\kappa_H[h]$  to account for the finite size  $h$  of the reference mesh and recover a good approximation of  $\kappa_H$ , see [17]. This correction factor is not necessary for the estimate  $\kappa_H[h, \tau_h]$ . That is when the reference problem is also stabilized.

A uniform coarse mesh of  $24 \times 2$  quadrilateral elements is used for both the Galerkin and the GLS method. The dispersion error associated with the Galerkin approximation can be assessed using the *a priori* estimate of the wave number given by (9)

$$E^{\text{pri}} = \kappa - \kappa^{\text{pri}} = \kappa - \frac{1}{H} \arccos\left(\frac{1 - (\kappa H)^2/3}{1 + (\kappa H)^2/6}\right),$$

which in this case is taken as the actual error in the wave number due to the one-dimensional character of the solution (up to the pollution errors introduced by the Robin boundary conditions). Note that the GLS solution is, for this particular mesh and problem, dispersion free. Thus, the Robin boundary conditions are the unique perturbation producing errors in the approximations of  $\kappa$ .

The different *a posteriori* estimates of the dispersion error are computed using a series of successively nested reference meshes, both triangular and quadrilateral. For the quadrilateral meshes, refinement is performed only in the  $x$ -direction and thus maintaining two rows of elements on all the reference meshes, due to the one-dimensional character of the solution: for  $h = H/2$  each quadrilateral in the coarse mesh is divided into two new ones yielding a mesh of  $48 \times 2$  elements, for  $h = H/4$ , each quadrilateral element is divided into four new ones yielding a mesh of  $96 \times 2$  elements, etc.

The first columns of Table I show the truth estimates of the dispersion error  $E[h] := \kappa - \kappa_H[h]$  and  $E[h, \tau_h] := \kappa - \kappa_H[h, \tau_h]$  where the numerical wave numbers  $\kappa_H[h]$  and  $\kappa_H[h, \tau_h]$  are computed solving the non-linear problems (13) and (15), respectively, and  $c_f = n_r^2 / (n_r^2 - 1)$  stands for the correction factor applied to  $\kappa_H[h]$ , where  $n_r = H/h$ . Note that these truth estimates are computationally unaffordable in real applications, because they involve many resolutions of the problem in the reference mesh. They are computed in academic problems to see the effectivity of the proposed practical estimates. As can be seen, both the estimates  $c_f E[h]$  and  $E[h, \tau_h]$  assessing the dispersion error of the Galerkin approximation are in very good agreement with the *a priori* estimate. It is worth noting that the estimate  $E[h, \tau_h]$  yields very good results even for the case  $h = H/2$  being less sensitive than  $c_f E[h]$  to the choice of the reference mesh size.

The last columns in Table I correspond to the practical estimates obtained from the recovered solution  $u^*$ . In this case  $u^*$  is computed using the exponential fitting. Four different estimates are computed. The first is the estimate proposed by Steffens and Díez [17],  $E^* := \kappa - \kappa^*$ , associated with the assessed wave number obtained from (14) and enhanced by its multiplicative factor.

Table II. Example 1: convergence of the estimates of the dispersion error through a uniform mesh refinement using cartesian quadrilateral meshes (top) and hexagonal triangular meshes (bottom).

$n_{np}$	Galerkin				GLS/FE		
	$E^{pri}$	$E$	$E_{pol}^*$	$E_{exp}^*$	$E$	$E_{pol}^*$	$E_{exp}^*$
75	1.02211	1.02211	1.23174	1.01293	$-4.1 \times 10^{-8}$	0.23026	-0.00626
99	0.60404	0.60404	0.71868	0.59251	$-5.1 \times 10^{-8}$	0.12522	-0.00134
123	0.39584	0.39584	0.46304	0.38942	$5.1 \times 10^{-8}$	0.07167	-0.00035
147	0.27851	0.27851	0.32051	0.27525	$4.1 \times 10^{-8}$	0.04401	-0.00011
172	0.79686	0.79782	0.58502	0.78168	$2.6 \times 10^{-8}$	-0.22714	-0.00542
293	0.47022	0.46319	0.36999	0.45619	$2.9 \times 10^{-8}$	-0.09656	-0.00229
446	0.30565	0.30074	0.25304	0.29794	$5.4 \times 10^{-8}$	-0.04915	-0.00116
631	0.21365	0.21040	0.18306	0.20935	$3.6 \times 10^{-8}$	-0.02829	-0.00066

The other three options correspond to the three approximations of  $\kappa^*[\tau_h]$  detailed in Section 3.3. Recall that Option 1 results from numerically solving the non-linear one-dimensional problem and, since this approximation only depends on an end-user relative tolerance set to  $10^{-12}$ , it is assumed to be exact, that is  $E^*[\tau_h] := \kappa - \kappa^*[\tau_h]$ . Option 2 is associated with  $\kappa^*[\tau_h^K]$  yielding the estimate  $E^*[\tau_h^K] := \kappa - \kappa^*[\tau_h^K]$ , and Option 3 is the most crude approximation of  $\kappa^*[\tau_h]$  since it considers that all the terms in the residual associated with the GLS formulation are constant with respect to the wave number. It is worth noting that all estimates produce similar and sharp approximations to the dispersion error for all the values of the reference mesh size  $h$ .

As expected, the truth estimates provide almost exact values for the dispersion error, fully coinciding with the *a priori* estimate. The effect of correcting the estimate with factor  $c_f$  or considering a stabilized reference problem is equivalent.

Following these results, in the remainder of the numerical examples, the parameter  $h$  is set to  $h = H/4$  (refining only in the  $x$ -direction for this example and uniformly refining the elements in the following examples) and the wave number is approximated using Option 2 that provides really good approximations. Hence, in the following the notation  $E^*$  is used to denote the estimate  $E^*[\tau_h^K]$  (both for the Galerkin and GLS method). A subindex is added to the notation  $E^*$  to specify the type of recovery used to compute  $u^*$ , namely,  $E_{pol}^*$  and  $E_{exp}^*$  for the polynomial and exponential fittings, respectively. Finally, the estimate  $E^*$  is compared with the truth estimate  $E[h, \tau_h]$  which is considered as the one providing the *most accurate-but not computable* approximation of the dispersion error, and it is denoted by  $E$ .

Table II and Figures 5 and 6 present the estimates corresponding to a sequence of uniformly refined meshes. Two series of meshes are used: one of structured quadrilaterals and one of triangular elements following a hexagonal pattern. The two fitting strategies (polynomial and exponential) are compared.

Note that the dispersion error associated with the GLS solution is almost negligible for the truth estimates. The Robin boundary conditions are the unique perturbation producing errors in the approximations of  $\kappa$  for the practical estimates.

Figure 5 shows the convergence of the estimates for the dispersion error of the Galerkin approximation using cartesian quadrilateral meshes. The convergence rate of all the estimates is 2 in the number of points of the mesh, matching the *a priori* expected convergence rate for the dispersion error, since, for a fixed value of  $\kappa$ ,  $E = \kappa - \kappa_H = \mathcal{O}(H^2)$ , see Equation (10). Here, a one-dimensional refinement only along the  $x$ -direction is done, and the characteristic element size  $H$  is inverse proportional to  $n_{np}$ , namely  $H = C/n_{np}$  and therefore  $E = \mathcal{O}(n_{np}^{-2})$ . However, it can be observed that the exponential fitting provides estimates which are in better agreement with the *a priori* or reference estimates.

Finally, Figure 6 shows the convergence of the bounds for both the Galerkin and GLS approximations using either quadrilateral or hexagonal triangular meshes. The reduction in the dispersion error using the stabilized GLS formulation becomes apparent both for quadrilateral and hexagonal



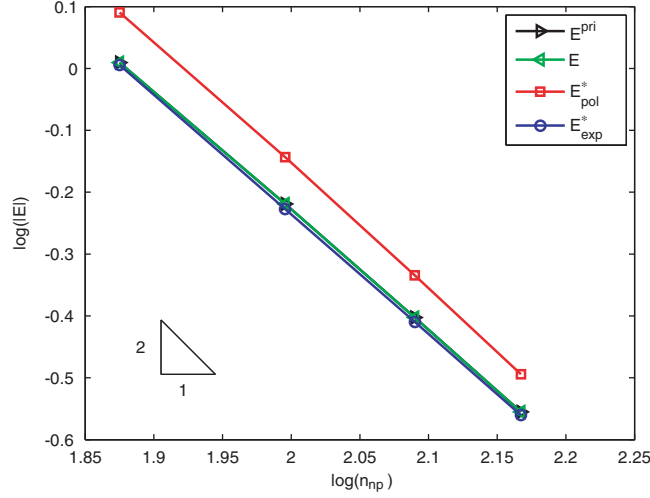


Figure 5. Example 1: convergence of the estimates of the dispersion error of the Galerkin approximations through a uniform mesh refinement using cartesian quadrilateral meshes.

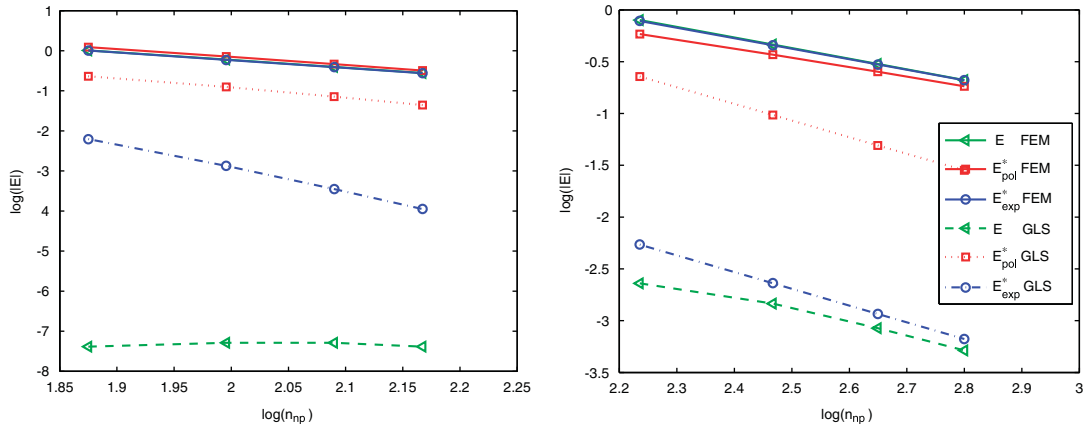


Figure 6. Example 1: convergence of the dispersion error both for the Galerkin and GLS approximations in a series of uniformly refined meshes: quadrilateral meshes (left) and hexagonal triangular meshes (right).

meshes. This important reduction is due to the fact that the stabilization parameters that have been used are particularly designed to eliminate the spurious dispersion of the exact solution  $e^{i\kappa x}$  for the particular quadrilateral and hexagonal meshes at hand. It is also clear that the exponential fitting, in this example, captures more precisely the shape of the solution and thus yields better estimates for the dispersion error.

Although extremely simple, this example demonstrates that the proposed methodology is able to assess the dispersion error in both Galerkin and GLS formulations. The estimate clearly detects that the GLS method reduces the dispersion. As shown in the following examples, the same tools are also useful in more involved situations.

#### 4.2. Example 2: plane wave in square domain

We consider the unit square  $\Omega = ]0, 1[ \times ]0, 1[$  with inhomogeneous Robin boundary conditions specified on all the boundaries of the square so that the exact solution is  $u = e^{i\kappa(\cos \alpha x + \sin \alpha y)}$ . That is, the solution is a plane wave propagating in the direction of angle  $\alpha$ , as illustrated in Figure 7.

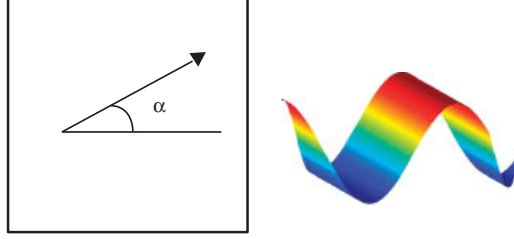


Figure 7. Example 2: problem setup (left) and solution for  $\alpha = \pi/8$  (right).

Table III. Example 2: assessment of the dispersion error of the Galerkin method for uniformly refined structured quadrilateral meshes.

$n_{np}$	Galerkin				
	$E$	$E_{pol}^*$	$E_{exp}^*$	$\eta_{pol}^*$	$\eta_{exp}^*$
81	0.24912	0.41670	0.23725	1.6727	0.9524
289	0.06330	0.09033	0.06328	1.4271	0.9998
1089	0.01563	0.01943	0.01593	1.2434	1.0197

Table IV. Example 2: assessment of the dispersion error of the GLS method for uniformly refined structured quadrilateral meshes. The GLS approximations are computed using different stabilization directions  $\theta^{opt}$ .

$n_{np}$	GLS/FE					
	$\theta^{opt} = 0$		$\theta^{opt} = \pi/8$		$\theta^{opt} = \pi/4$	
	$E$	$E_{exp}^*$	$E$	$E_{exp}^*$	$E$	$E_{exp}^*$
81	$-7.45 \times 10^{-2}$	$-7.17 \times 10^{-2}$	$6.82 \times 10^{-4}$	$3.40 \times 10^{-4}$	$7.71 \times 10^{-2}$	$7.34 \times 10^{-2}$
289	$-1.99 \times 10^{-2}$	$-1.93 \times 10^{-2}$	$-4.43 \times 10^{-4}$	$3.80 \times 10^{-5}$	$1.91 \times 10^{-2}$	$1.95 \times 10^{-2}$
1089	$-5.02 \times 10^{-3}$	$-4.87 \times 10^{-3}$	$-1.84 \times 10^{-4}$	$1.68 \times 10^{-6}$	$4.66 \times 10^{-3}$	$4.88 \times 10^{-3}$

The model parameters are  $\kappa = 8$  and  $\alpha = \pi/8$  and the analytical solution associated with these parameters is depicted in Figure 7.

The performance of the estimates is studied for three different structured uniform quadrilateral meshes ( $8 \times 8$ ,  $16 \times 16$  and  $32 \times 32$  elements). In order to estimate the dispersion error associated with the Galerkin approximation, the stabilization parameters involved in the computation of  $E := E[h, \tau_h]$  and  $E^* := E^*[\tau_h^k]$  in Equations (15) and (17) are computed using the predefined direction  $\theta_h^{opt} = \alpha = \pi/8$ . The results are shown in Table III, where  $\eta_{pol}^* := E_{pol}^*/E$  and  $\eta_{exp}^* := E_{exp}^*/E$  is the effectivity index of the estimates with respect to the reference value  $E$  (truth estimates). Again, the exponential fitting provides better estimates to the dispersion error yielding very good effectivity indices near to one. A linear convergence rate of the estimates with respect to the number of nodes is obtained, matching the *a priori* expected convergence rate for the dispersion error. Indeed, since a uniform refinement is done in both the spatial dimensions, the characteristic element size  $H$  is related to the number of points in the mesh via the relation  $n_{np} \approx C/H^2$  and thus the convergence rate  $E = \mathcal{O}(H^2)$  is equivalent to  $\mathcal{O}(n_{np}^{-1})$ .

The same study is done for the GLS approximations of the problem using the same meshes. Although the exact solution is a plane wave, since the cartesian meshes are not aligned with the wave direction  $\alpha = \pi/8$ , none of the possible choices for the stabilization direction  $\theta^{opt}$  yields a nodally exact solution. Table IV shows the dispersion error of the GLS method for three different stabilization parameters  $\theta^{opt} = 0$ ,  $\theta^{opt} = \pi/8$  and  $\theta^{opt} = \pi/4$ . In all the computations the error estimates are performed using the same value of  $\theta^{opt}$  for the reference  $h$ -mesh, that is  $\theta_h^{opt} = \theta^{opt}$ .

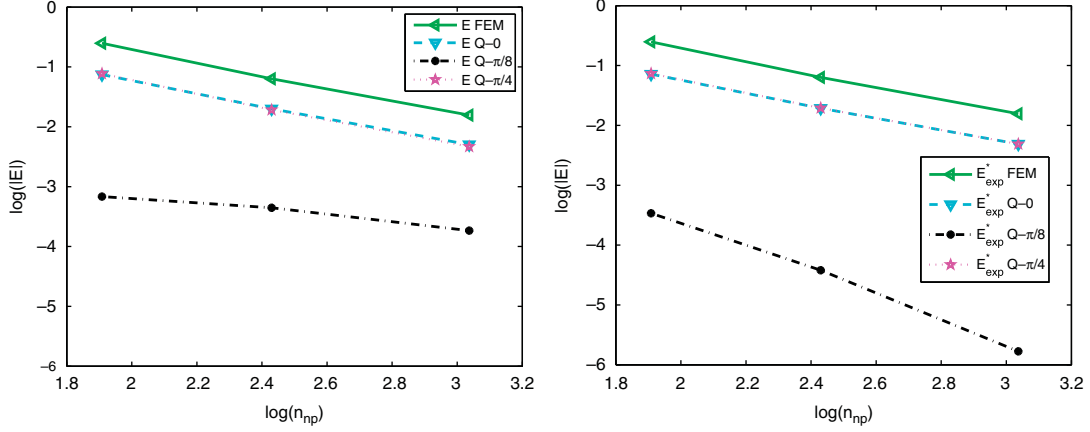


Figure 8. Example 2: performance of the estimates of the dispersion error for the Galerkin method and the GLS method for a plane wave associated with  $\alpha = \pi/8$  using a reference mesh (left) and the exponential fitting (right). The GLS approximations are computed using different stabilization parameters.

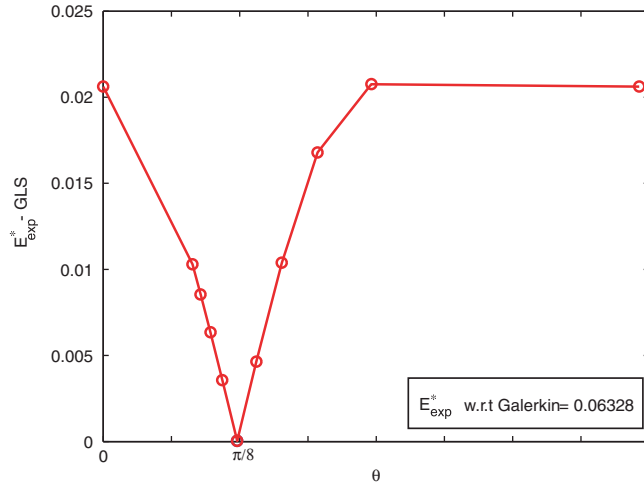


Figure 9. Example 2: influence of the selection of the stabilization angle  $\theta^{\text{opt}}$  in the dispersion error of the GLS approximation for the mesh with 269 nodes. The estimates are computed both using  $\theta^{\text{opt}} = \alpha$ .

The GLS method substantially reduces the dispersion error even for the non-optimal parameters  $\theta^{\text{opt}} = 0$  and  $\theta^{\text{opt}} = \pi/4$ . The error estimate  $E_{\text{exp}}^*$  is properly approximating the truth error  $E$  in all the cases. For  $\theta^{\text{opt}} = \pi/8$  the dispersion error is so small that the resulting effectivity is not as sharp as for the choices producing longer errors.

Figure 8 graphically displays the information shown in the tables in Tables III and IV. As can be seen, the estimates (depicted on the right of the figure) are in very good agreement with the reference mesh computations (depicted on the left of the figure). As mentioned before, the GLS method always performs better than the Galerkin method but there is a qualitative leap of accuracy when the optimal parameter  $\theta^{\text{opt}} = \pi/8$  is used.

Finally, Figure 9 shows the influence of the stabilization direction  $\theta^{\text{opt}}$  used to compute the GLS finite element approximation in the dispersion error. The study is done varying  $\theta^{\text{opt}}$  in the range  $[0, \pi/2]$ . As expected, the optimal performance is reached when the wave direction of the GLS method coincides with the angle of the exact solution,  $\theta^{\text{opt}} = \alpha = \pi/8$ . In any case, if no information about the exact solution is at hand and thus, an arbitrary choice of  $\theta^{\text{opt}}$  is considered, the GLS method provides an important reduction in the dispersion error when compared with the Galerkin

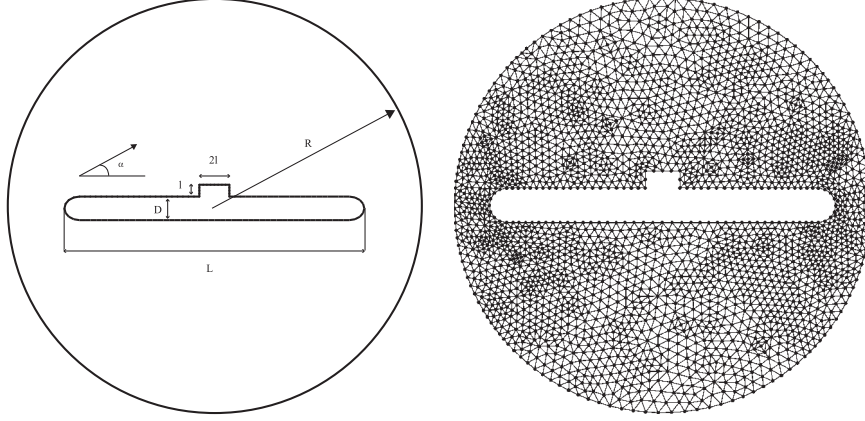


Figure 10. Example 3: geometry of the fictitious bounded domain to study the scattering from a submarine-shaped obstacle problem (top) and a mesh of 2567 nodes (bottom).

approximation: the estimated dispersion error is reduced from  $E_{\text{exp}}^* = 0.06328$  to  $E_{\text{exp}}^* \approx 0.02$  in the worst case.

#### 4.3. Example 3: scattering from submarine-shaped obstacle

The acoustic scattering from an acoustically hard obstacle is studied. The geometry is a submarine-like object parameterized by the distances  $l=3$ ,  $L=60$  and  $D=6$ , see Figure 10. The incident wave is characterized by its wave number  $\kappa=\pi/3$  and an angle of incidence  $\alpha=5\pi/4$ . The original problem is an unbounded Helmholtz problem which is reduced to an interior problem over a bounded computational domain with a circular boundary of radius  $R=36$ . In the fictitious boundary, second-order Bayliss–Gunzberger–Turkel (BGT) [29, 30] absorbing boundary conditions are applied.

The solution of the acoustic scattering problem is decomposed into  $u = u_r + u_i$ , where  $u_r$  and  $u_i$  are the so-called reflected and incident waves, respectively. For a given wave number  $\kappa$  and incident wave direction  $\alpha$ , the incident wave is  $u_i = e^{i\kappa(\cos \alpha x + \sin \alpha y)}$  and the reflected wave  $u_r$  is the solution of the Helmholtz equation (1) with  $f=0$  (because the total wave  $u$  is the solution of the homogeneous Helmholtz problem, Equation (1) with  $f=0$ , and  $-\Delta u_i - \kappa u_i = 0$ ). Neumann boundary conditions are applied on the boundary of the obstacle

$$\nabla u_r \cdot \mathbf{n} = -\nabla u_i \cdot \mathbf{n}$$

( $g = -\nabla u_i \cdot \mathbf{n}$  in (2b) since homogeneous Neumann boundary conditions are applied the total wave  $u$ , that is,  $u$  verifies Equation (2b) with  $g=0$ ), and first-order Bayliss–Gunzberger–Turkel (BGT) non-reflecting boundary conditions are applied to the fictitious boundary

$$\nabla u_r \cdot \mathbf{n} = \mathcal{M} u_r = -i\kappa u_r - \frac{\zeta}{2} u_r$$

in (2c). Here,  $\zeta$  is the curvature of the surface of the scatterer, which for the particular case of a circular boundary of radius  $R$  is  $\zeta = 1/R$ .

In this example, the dispersion error committed in the approximation of the reflected solution  $u_r$  is studied. The total approximated scattered field  $u$  is computed from  $u_r$  adding the known incident field  $u_i$ . Figures 11 and 12 show the approximations obtained using the Galerkin method with a triangular mesh of 10 026 nodes.

The behavior of the estimates of the dispersion error is analyzed for different unstructured triangular meshes both for the Galerkin and SUPG approximations of the reflected solution  $u_r$ . The triangular meshes are obtained from the initial mesh (see Figure 10) using a uniform refinement, that is, a new mesh is obtained from a previous mesh by refining each triangle into four new triangles.

Table V shows the results associated with the Galerkin approximation. In this case, the stabilization parameters involved in the computation of the truth estimate  $E$  and the practical estimate

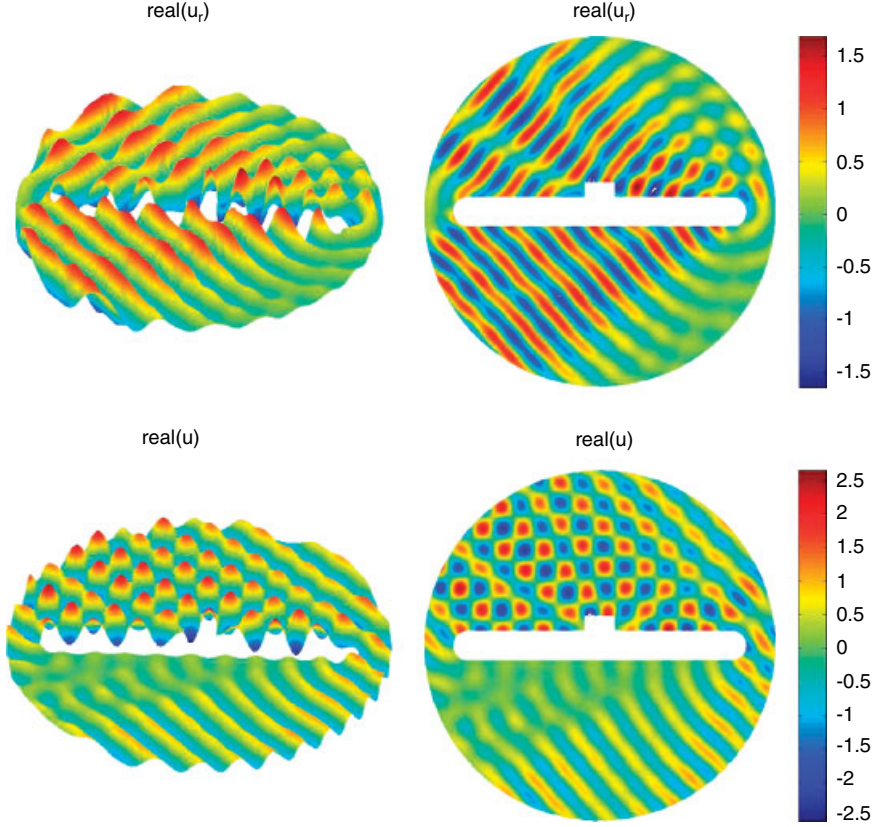


Figure 11. Example 3: real part of the approximated solutions of the scattering problem for  $\alpha=5\pi/4$  and  $\kappa L=62.83$ : scattered wave (top) and total wave (bottom). Approximations computed using the Galerkin method and a mesh of 10 026 nodes.

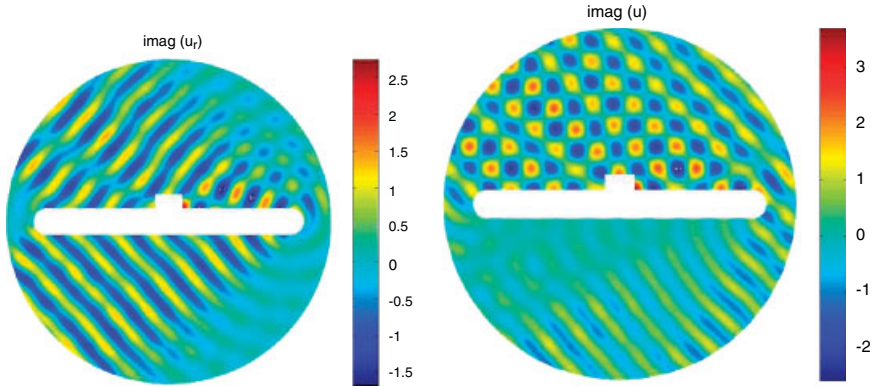


Figure 12. Example 3: imaginary part of the approximated solutions of the scattering problem for  $\alpha=5\pi/4$  and  $\kappa L=62.83$  computed using the Galerkin method and a mesh of 10 026 nodes: scattered wave (left) and total wave (right).

$E^*$  are computed using the predefined direction  $\theta_h^{\text{opt}} = \alpha = 5\pi/4$ . Both the estimates obtained using a polynomial and an exponential fitting provide fairly good approximations to the truth value  $E$ . However, the exponential approach provides better effectivities, closer to one. Moreover, the expected rate of convergence of the estimates of the dispersion error is obtained in all the cases.

Table VI shows the results obtained by the GLS approximations. Three different stabilized approximations are computed associated with the stabilization directions  $\theta^{\text{opt}}=0$ ,  $\theta^{\text{opt}}=\pi/12$  and

Table V. Example 3: assessment of the dispersion error committed by the Galerkin method for uniformly refined unstructured triangular meshes.

$n_{np}$	Galerkin				
	$E$	$E_{pol}^*$	$E_{exp}^*$	$\eta_{pol}^*$	$\eta_{exp}^*$
2567	0.69064	0.55271	0.67186	0.8003	0.9728
10026	0.19509	0.15604	0.20538	0.7998	1.0527
39620	0.04829	0.02959	0.05003	0.6128	1.0360

Table VI. Example 3: assessment of the dispersion error of the GLS method for uniformly refined unstructured triangular meshes.

$n_{np}$	GLS/FE					
	$\theta^{opt}=0$		$\theta^{opt}=\pi/12$		$\theta^{opt}=\pi/6$	
	$E_{pol}^*$	$E_{exp}^*$	$E_{pol}^*$	$E_{exp}^*$	$E_{pol}^*$	$E_{exp}^*$
2567	$1.31 \times 10^{-2}$	$4.357 \times 10^{-3}$	$1.37 \times 10^{-2}$	$4.89 \times 10^{-3}$	$1.42 \times 10^{-2}$	$5.42 \times 10^{-3}$
10026	$4.77 \times 10^{-3}$	$1.771 \times 10^{-3}$	$4.80 \times 10^{-3}$	$1.79 \times 10^{-3}$	$4.83 \times 10^{-3}$	$1.81 \times 10^{-3}$
39620	$2.84 \times 10^{-3}$	$1.866 \times 10^{-4}$	$2.85 \times 10^{-3}$	$1.89 \times 10^{-4}$	$2.85 \times 10^{-3}$	$1.89 \times 10^{-4}$

The GLS approximations are shown for different stabilization directions  $\theta^{opt}$ .

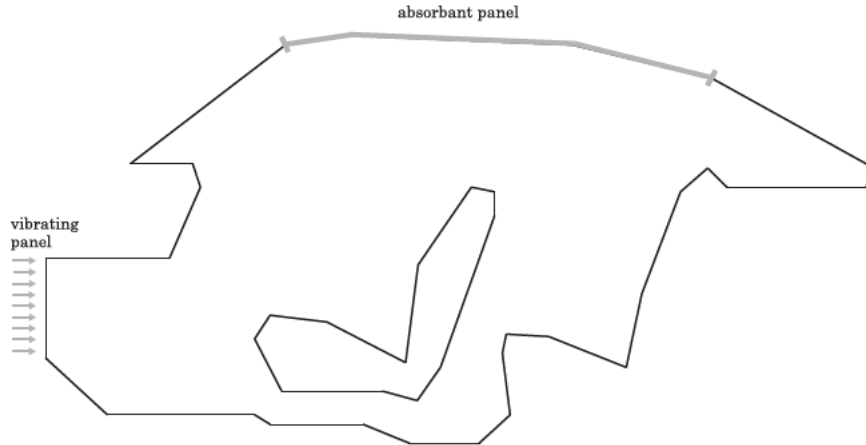


Figure 13. Example 4: description of the two-dimensional section of the cabin of a car and its associated boundary conditions.

$\theta^{opt} = \pi/6$ . The corresponding estimates are computed using the same values of  $\theta^{opt}$  in the reference mesh,  $\theta_h^{opt} = \theta^{opt}$ .

As expected, the use of stabilized formulations reduces considerably the dispersion error. In this case, the three tested stabilization directions provide similar results. Moreover, the exponential fitting which provides really good estimates for the Galerkin solution, yields really low estimates of the dispersion error of the stabilized approximations indicating that, in this example, the dispersion error is nearly negligible when using a GLS approach.

#### 4.4. Example 4: 2D acoustic car cavity

This example studies the noise transmission inside a two-dimensional section of the cabin of a car that is excited by vibrations of the front panel and damped by Robin boundary conditions. This example is frequently used as a benchmark problem in error assessment for interior acoustic problems [20, 31, 32]. The geometry of the cabin is shown in Figure 13. The size of the domain

is characterized by the maximum horizontal and vertical lengths,  $L_x=2.7$  m and  $L_y=1.1$  m, respectively. The source term entering in Equation (1) is  $f=0$ , and as mentioned in Remark 1, for interior acoustic wave propagation problems, the Neumann and Robin boundary conditions entering in Equation (2) are of the form  $g=-i\rho c\kappa\bar{v}_n$  and  $\mathcal{M}u=-i\rho c\kappa A_n u$ , where in this case the material parameters are  $c=340$  m/s standing for the speed of sound of the medium and  $\rho=1.225$  kg/m<sup>3</sup> standing for the mass density. The vibrating front panel is excited with a unit normal velocity  $\bar{v}_n=1$  m/s, whereas the roof is considered to be an absorbent panel with associated admittance  $A_n=1/2000$  m(Pas)<sup>-1</sup>. The rest of the boundary is assumed to be perfectly reflecting and thus  $\bar{v}_n=0$  m/s. Finally, a wave number of  $\kappa\approx 9.7$  has been considered in the computations (equivalent to a frequency of 525 Hz).

In this problem, the exponential fitting presented above yields bad estimates, worse than the standard polynomial fitting. This is due to the fact that the solution is extremely complex without a predominant direction. At many points of the domain, the solution can be expressed as a sum of several plane waves with similar amplitudes. Thus, the exponential fitting fails to properly approximate the local behavior of the modified solution in the vicinity of these points. Actually, the exponential recovery in these zones introduces unrealistic discontinuities resulting in bad estimates. In the following, this phenomenon is described in detail, as well as the proposed remedy.

It is well known that the exact solution of the 2D homogeneous Helmholtz equation can be expressed as an infinite sum of plane waves traveling in different directions. In the previous examples, the solutions were either a single plane wave traveling in a predefined direction (see examples 1 and 2) or had a prevalent plane wave direction, although the prevalent wave direction may vary from different zones of the domain (see the scattered solution of example 3). The sound transmission inside a car cabin is a more complex phenomenon and the solution does not present clear prevalent directions but is a combination of different plane waves with similar amplitudes (see Figure 14).

Even if the exact solution has no prevalent directions, one can consider an exponential representation of the exact solution of the problem

$$u(\mathbf{x})=r(\mathbf{x})e^{i\theta(\mathbf{x})},$$

where  $r(\mathbf{x})$  and  $\theta(\mathbf{x})$  are the real-valued functions providing the modulus and angle of  $u$ , respectively. In the cases where the solution does not have a prevalent direction two phenomena may appear: on the one hand the angle distribution  $\theta(\mathbf{x})$  may present discontinuities coinciding with areas where the modulus vanishes, and, on the other hand, the modulus distribution  $r(\mathbf{x})$  may present a highly non-linear and non-smooth behavior in some regions.

To illustrate these phenomena, the modulus and angle distributions of three simple solutions are shown in Figure 15. First, the solution  $u=2e^{i\kappa x}+e^{i\kappa y}$  is considered. Note that, in this case, the plane wave traveling in the  $x$ -direction,  $e^{i\kappa x}$ , prevails over the wave traveling in the  $y$ -direction,  $e^{i\kappa y}$ . As can be seen in Figure 15, the standard representation of the angle distribution  $\theta(\mathbf{x})$  is a discontinuous function, which can be easily post-processed to recover a continuous angle distribution. Moreover, the modulus does not present large variations over small regions. In this case, the exponential fitting described in Section 3.5 provides accurate approximations of  $u$ . The second example,  $u=e^{i\kappa x}+e^{i\kappa y}$ , shows that if the solution is obtained combining two plane waves of the same amplitude, and thus it does not have any prevalent direction, angle discontinuities appear in some predefined straight lines. As the number of plane waves that comprise the solution  $u$  increases, see for instance the third example  $u=e^{i\kappa x}+e^{i\kappa y}+e^{-i\kappa y}$ , the modulus and angle distributions may present areas with a highly non-linear and non-smooth behavior. Note that, although the angle distribution only presents point or removable discontinuities at nine points of the domain, obtaining a globally smooth angle distribution from the standard angle representation is not a trivial task. Figure 16 shows the behavior of the modulus and angle distribution associated with the acoustic pressure inside the car cabin. As can be seen, it is not easy to clearly identify the regions where the angle distribution is discontinuous.

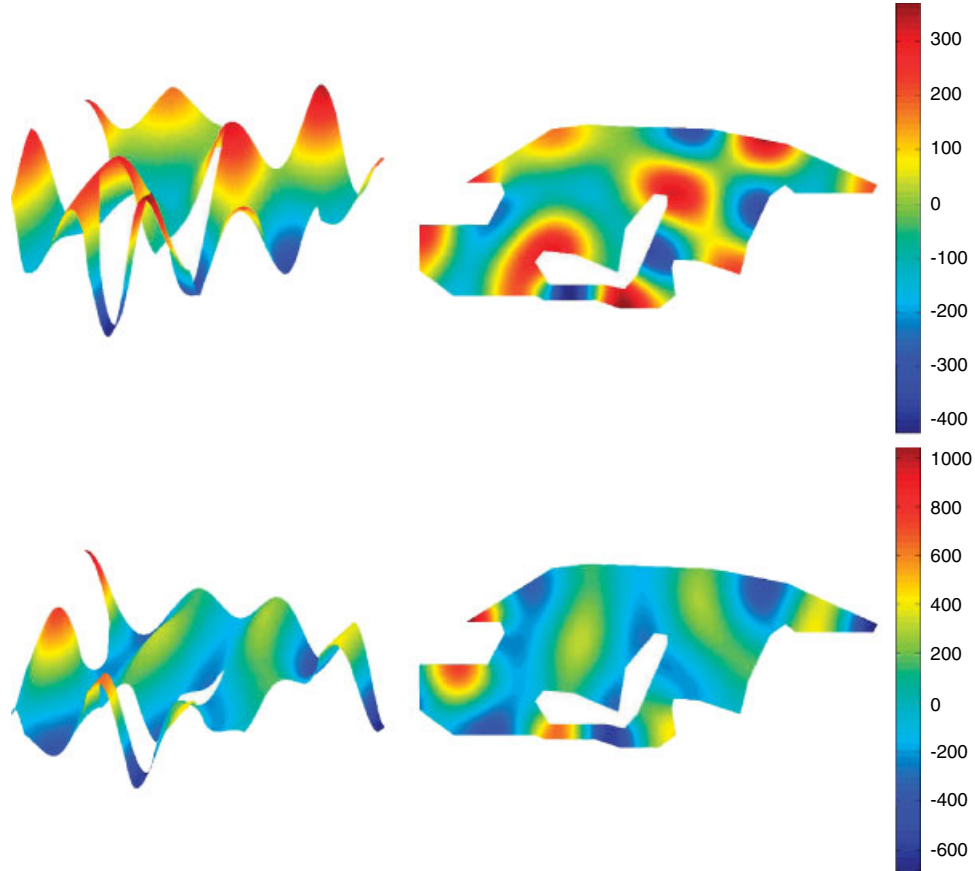


Figure 14. Example 4: solution of the noise transmission problem inside the cabin of a car obtained with an overkill mesh of 20 160 nodes: real part of  $u$  (top) and imaginary part of  $u$  (bottom).

The exponential fitting technique is based on finding a proper local polynomial representation for the modulus and angle distributions. Thus, in regions where either the angle is discontinuous or the modulus presents large oscillations, the exponential representation yields poor results. In this work, a simple workaround is proposed: first, the smoothing technique identifies the elements near the angle discontinuities or near the regions where the modulus has a non-smooth behavior. Then, the exponential fitting is applied only to the non-selected elements while a polynomial fitting is applied to the problematic elements. The estimates obtained with this combined approach are denoted in the following by  $\widehat{E}_{\text{exp}}^*$ .

Estimates of the dispersion error for the Galerkin approximations of the solution are computed for two different triangular meshes of 568 and 2122 nodes, respectively. The results are shown in Table VII. As can be seen, both the polynomial and the combined estimates provide fairly good approximations to the truth value  $E$ . However, using an exponential representation, where possible, allows obtaining effectivities closer to one.

Figure 17 shows the elements that have been selected in the combined approach to apply the polynomial smoothing technique instead of the exponential one. Note that these regions are in good agreement with those highlighted in Figure 16.

The reduction in the dispersion error obtained by using stabilization techniques is shown in Table VIII. This table also shows the influence of the selection of the stabilization parameter. As can be seen the results of the GLS approximations with the three stabilization parameters are nearly identical for the two meshes, and provide significant improvement over the Galerkin method.



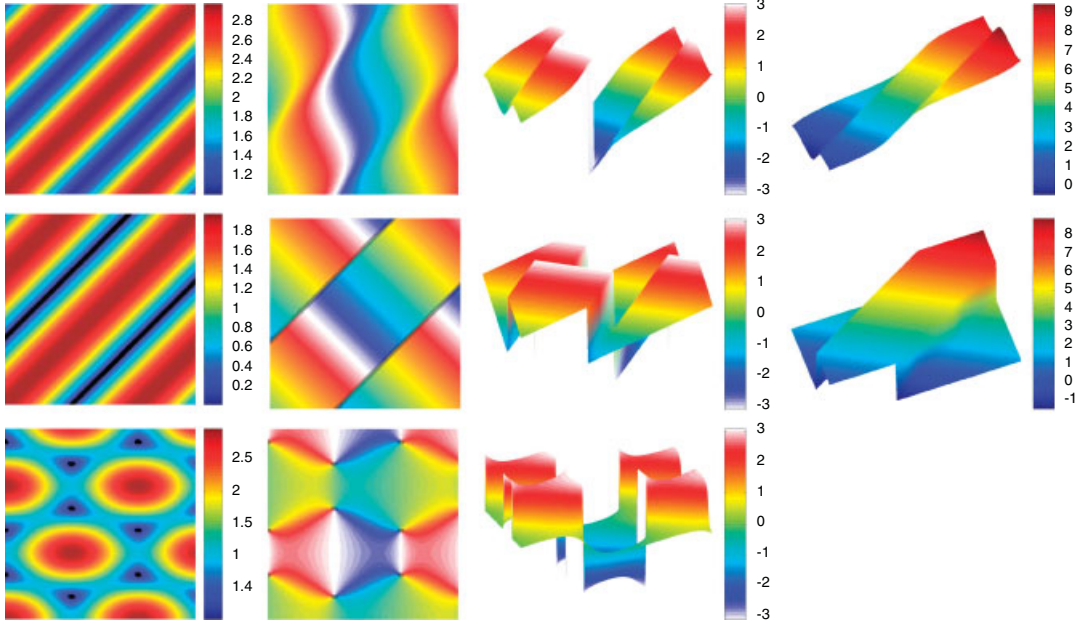


Figure 15. Example 4: behavior of the modulus and angle distributions,  $\theta(\mathbf{x})$  and  $r(\mathbf{x})$ , respectively, for three simple solutions in the unit square. From top to bottom:  $u = 2e^{kix} + e^{kiy}$ ,  $u = e^{kix} + e^{kiy}$  and  $u = e^{kix} + e^{kiy} + e^{-kiy}$  for  $\kappa = 9.7$ . For each solution, the modulus distribution (left) and two views of the angle distributions (middle left, middle right) are shown. When possible, equivalent angle distributions only containing non-removable discontinuities—where the discontinuities associated with a  $2\pi$  angle jump have been smoothed—are shown (right).



Figure 16. Example 4: modulus (left) and angle (middle and right) distribution of the acoustic pressure inside the car cabin. The areas where the modulus is nearly zero are highlighted in the plot in the middle to see the areas where the angle distribution may present discontinuities.

Table VII. Example 4: assessment of the dispersion error of the Galerkin method for unstructured triangular meshes.

$n_{np}$	Galerkin				
	$E$	$E_{pol}^*$	$\hat{E}_{exp}^*$	$\eta_{pol}^*$	$\eta_{exp}^*$
568	0.15001	0.08231	0.12960	0.5486	0.8639
1092	0.07506	0.06694	0.07389	0.8918	0.9845

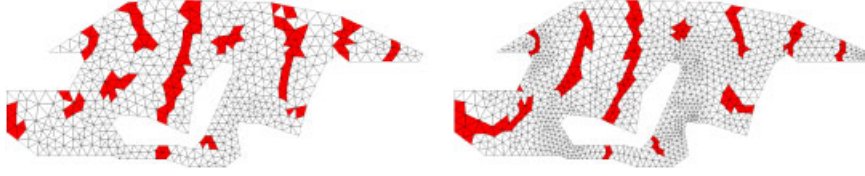


Figure 17. Example 4: two unstructured triangular mesh where the red elements corresponding to solution fitting polynomial.

Table VIII. Example 4: assessment of the dispersion error of the GLS method for an unstructured triangular mesh.

$n_{np}$	GLS/FE		
	$E$	$E_{pol}^*$	$\widehat{E}_{exp}^*$
$\theta^{opt} = 0$			
568	0.03792	0.02267	0.03563
1092	0.00577	0.00653	0.00644
$\theta^{opt} = \pi/12$			
568	0.03808	0.02281	0.03583
1092	0.00583	0.00658	0.00651
$\theta^{opt} = \pi/6$			
568	0.03824	0.02294	0.03601
1092	0.00589	0.00663	0.00656

The GLS approximations are shown for different stabilization directions  $\theta^{opt}$ .

## 5. CONCLUSIONS

This paper introduces an error assessment technique for the numerical wave number  $\kappa_H$  of the Helmholtz problem, both for standard Galerkin and stabilized formulations. The strategy introduced in [17], which determines the numerical wave number  $\kappa_H$  as the one that better accommodates the numerical solution  $u_H$  in a modified problem, has been extended to deal with stabilized formulations. The numerical solution  $u_H$  and the reference modified problem are computed using stabilized methods to obtain both more accurate approximations of the solution and sharper estimates of the dispersion error.

The proposed strategy requires obtaining an inexpensive approximation of the modified problem, using post-processing techniques. Thus, the associated numerical wave number is readily recovered using a closed expression. A new improved recovery technique is developed to take advantage of the nature of the solutions of wave problems. The standard polynomial least-squares techniques is replaced by an exponential fitting yielding much sharper results in most applications. However, both the error estimates computed using a polynomial and exponential fitting provide reasonable approximations of the true errors.

The estimates of the dispersion error reaffirm that using stabilized approximations substantially improves the performance of finite element computations of time-harmonic acoustics at high wave numbers. The sensitivity of the choice of the stabilization parameter for the GLS method has been studied concluding that the change in the orientation of the stabilization parameter has little effect on the results of non-academic problems or when considering non-structured meshes.

## ACKNOWLEDGEMENTS

This work was partially supported by Ministerio de Educación y Ciencia, Grants DPI2007-62395 and BIA2007-66965; Programme Alβan, the European Union Programme of High Level Scholarships for

Latin America, scholarship no. E06D100641BR and Publishing Arts Research Council under grant no. 98-1846389.

#### REFERENCES

1. Ihlenburg F. *Finite Element Analysis of Acoustic Scattering*. Springer: New York, 1998.
2. Ihlenburg F, Babuška I. Dispersion analysis and error estimation of Galerkin finite element methods for the Helmholtz equation. *International Journal for Numerical Methods in Engineering* 1995; **38**:3745-3774.
3. Ihlenburg F, Babuška I. Finite element solution of the Helmholtz equation with high wave number. Part I: the hp-version of the FEM. *Computers and Mathematics with Applications* 1995; **38**:9-37.
4. Babuška I, Ihlenburg F, Paik ET, Sauter SA. A generalized finite element method for solving the Helmholtz equation in two dimensions with minimal pollution. *Computer Methods in Applied Mechanics and Engineering* 1995; **128**:325-359.
5. Gerdes K, Ihlenburg F. On the pollution effect in FE solutions of the 3D-Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering* 1999; **170**(1-2):155-172.
6. Babuška I, Sauter SA. Is the pollution effect of the FEM avoidable for the Helmholtz equation considering high wavenumber? *SIAM Review* 2000; **42**(3):451-484.
7. Ainsworth M. Discrete dispersion relation for hp-version finite element approximation at high wave number. *SIAM Journal on Numerical Analysis* 2004; **42**(2):553-575.
8. Oden JT, Prudhomme S, Demkowicz L. A posteriori error estimation for acoustic wave propagation problems. *Archives of Computational Methods in Engineering* 2005; **12**(4):343-389.
9. Irimie S, Bouillard P. A residual a posteriori error estimator for the finite element solution of the Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering* 2001; **190**(31):4027-4042.
10. Bouillard P. Admissible fields and error estimation for acoustic FEA with low wave numbers. *Computers and Structures* 1999; **73**(1-5):227-237.
11. Bouillard P. Influence of the pollution on the admissible field error estimation for FE solutions of the Helmholtz equation. *International Journal for Numerical Methods in Engineering* 1999; **45**(7):783-800.
12. Peraire J, Patera AT. Asymptotic a posteriori finite element bounds for the outputs of noncoercive problems: the Helmholtz and Burgers equations. *Computer Methods in Applied Mechanics and Engineering* 1999; **171**(1):77-86.
13. Babuška I, Ihlenburg F, Strouboulis T, Gangaraj SK. A posteriori error estimation for finite element solutions of Helmholtz equation. Part I: the quality of local indicators and estimators. *International Journal for Numerical Methods in Engineering* 1997; **40**(18):3443-3462.
14. Babuška I, Ihlenburg F, Strouboulis T, Gangaraj K. A posteriori error estimation for finite element solutions of Helmholtz. Part II: estimation of the pollution error. *International Journal for Numerical Methods in Engineering* 1997; **40**:3883-3900.
15. Stewart JR, Hughes TJR. Explicit residual-based a posteriori error estimation for finite element discretizations of the Helmholtz equation. Computation of the constant and new measures of error estimator quality. *Computer Methods in Applied Mechanics and Engineering* 1996; **131**(3-4):335-363.
16. Stewart JR, Hughes TJR. A posteriori error estimation and adaptive finite element computation of the Helmholtz equation in exterior domains. *Finite Elements in Analysis and Design* 1996; **22**(1):15-24.
17. Steffens LM, Díez P. A simple strategy to assess the error in the numerical wave number of the finite element solution of the Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering* 2009; **198**:1389-1400.
18. Babuška I, Melenk JM. The partition of unity method. *International Journal for Numerical Methods in Engineering* 1997; **40**:727-758.
19. Franca LP, Dutra do Carmo EG. The Galerkin gradient least-squares method. *Computer Methods in Applied Mechanics and Engineering* 1989; **74**(1):41-54.
20. Harari I, Magoulès F. Numerical investigations of stabilized finite element computations for acoustics. *Wave Motion* 2004; **39**:339-349.
21. Harari I, Hughes TJR. Galerkin/least-squares finite element methods for the reduced wave equation with nonreflecting boundary conditions in unbounded domains. *Computer Methods in Applied Mechanics and Engineering* 1992; **98**(3):411-454.
22. Donea J, Huerta A. *Finite Element Methods for Flow Problems*. Wiley: Chichester, 2003.
23. Hughes TJR, Brooks AN. A multi-dimensional upwind scheme with no crosswind diffusion. *Finite Element Methods for Convection Dominated Flows (Collection of Papers Presentation at Winter Annual Meeting of the American Society of Mechanical Engineers (ASME))*, vol. 34, New York, 1979; 19-35.
24. Hughes TJR, Franca LP, Hulbert GM. A new finite element formulation for computational fluid dynamics. *Computer Methods in Applied Mechanics and Engineering* 1989; **73**:173-189.
25. Harari I, Nogueira CL. Reducing dispersion of linear triangular elements for the Helmholtz equation. *Journal of Engineering Mechanics* 2002; **128**(3):351-358.
26. Wiberg N-E, Zeng LF, Li XD. Error estimation and adaptivity in elastodynamics. *Computer Methods in Applied Mechanics and Engineering* 1992; **101**:369-395.

27. Díez P, Calderón G. Goal-oriented error estimation for transient parabolic problems. *Computational Mechanics* 2007; **39**(5):631–646.
28. Calderón G, Díez P. Análisis de diferentes estimadores de error de postproceso para adaptatividad orientada al resultado. *Revista Internacional de Métodos Numéricos Cálculo y Diseño Ingeniería* 2006; **22**(2):193–213.
29. Bayliss A, Gunzburger M, Turkel E. Boundary conditions for the numerical solution of elliptic equations in exterior regions. *SIAM Journal on Applied Mathematics* 1982; **42**(2):430–451.
30. Djellouli R, Farhat C, Macedo A, Tezaur R. Finite element solution of two-dimensional acoustic scattering problems using arbitrarily shaped convex artificial boundaries. *Journal of Computational Acoustics* 2000; **8**(1):81–99.
31. Bouillard P, Ihlenburg F. Error estimation and adaptivity for the finite element method in acoustics: 2D and 3D applications. *Computer Methods in Applied Mechanics and Engineering* 1999; **176**(1):147–163.
32. Suleau S, Deraemaeker A, Bouillard P. Dispersion and pollution of meshless solutions for the Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering* 2000; **190**:639–657.