INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2010; **00**:1–25
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nme

A Petrov–Galerkin formulation for the alpha-interpolation of FEM and FDM stencils. Applications to the Helmholtz equation

Prashanth Nadukandi*, Eugenio Oñate and Julio Garcia

Centre Internacional de Mètodos Numèrics en Enginyeria (CIMNE), Universitat Politècnica de Catalunya (UPC), Campus Nord, Edifici C1, Gran Capitan s/n, 08034 Barcelona, Spain

SUMMARY

A new Petrov–Galerkin (PG) method involving two parameters viz. α_1, α_2 is presented which yields the following schemes on rectangular meshes: a) a compact stencil obtained by the linear interpolation of the Galerkin FEM and the classical central FDM, should the parameters be equal, i. e. $\alpha_1 = \alpha_2 = \alpha$ and b) the nonstandard compact stencil presented in [doi:10.1002/nme.3043] for the Helmholtz equation if the parameters are distinct, i. e. $\alpha_1 \neq \alpha_2$. The nonstandard compact stencil is obtained by taking the linear interpolation of the diffusive terms (specified by α_1) and the mass terms (specified by α_2) that appear in the stencils obtained by the standard Galerkin FEM and the classical central FDM, respectively. On square meshes, these two schemes were shown to provide solutions to the Helmholtz equation that have a dispersion accuracy of fourth and sixth order respectively [doi:10.1002/nme.3043]. The objective of this paper is to study the performance of this Petrov–Galerkin method for the Helmholtz equation using nonuniform meshes and the treatment of natural boundary conditions. Copyright © 2010 John Wiley & Sons, Ltd.

Received . . .

KEY WORDS: Petrov–Galerkin; alpha interpolation of FEM and FDM; Helmholtz equation; dispersion accuracy; nonconforming test spaces

1. INTRODUCTION

This paper is a continuation of [1] wherein a simple domain-based higher-order compact numerical scheme involving two parameters viz. α_1, α_2 was presented for the Helmholtz equation. The stencil obtained by choosing the parameters as distinct, i. e. $\alpha_1 \neq \alpha_2$ was denoted therein as the 'nonstandard compact stencil'. The nonstandard compact stencil is obtained by taking the linear interpolation of the diffusive terms (specified by α_1) and the mass terms (specified by α_2) that appear in the stencils obtained by the standard Galerkin finite element method (FEM) and the classical central finite difference method (FDM), respectively. Taking $\alpha_1 = \alpha_2 = \alpha$, the nonstandard compact stencil simplifies to the α-interpolation of the Galerkin FEM and the classical central FDM stencils. For the Helmholtz equation, generic expressions for the parameters were given that guarantees a dispersion accuracy of sixth-order should $\alpha_1 \neq \alpha_2$ and fourth-order should $\alpha_1 = \alpha_2$. As the findings reported therein and the corresponding analysis was done for compact stencils, the contribution of the Galerkin FEM to the equation stencil corresponds to the choice of the lowest order rectangular block finite elements. By blocks we mean Cartesian product of intervals and by lowest order we refer to multilinear finite-element (FE) interpolation on these blocks. In this paper we extend this scheme to unstructured meshes. The focus of this paper is twofold: a) to design a Petrov–Galerkin (PG) method that reproduces on structured meshes the aforesaid numerical

*Correspondence to: Centre Internacional de Mètodos Numèrics en Enginyeria (CIMNE), Universitat Politècnica de Catalunya (UPC), Campus Nord, Edifici C1, Gran Capitan s/n, 08034 Barcelona, Spain. E-mail: npras@cimne.upc.edu

scheme and b) to study the performance of this PG method on nonuniform meshes and for problems subjected to natural boundary conditions.

The basic idea is to construct the basis functions of the test space from the standard FE shape functions such that their scalar product results in the lumped mass matrix. These basis functions are designed to have the following features: a) to be piecewise polynomials of the same degree as the FE shape functions, b) to be a partition of unity (only in the sense that they add up to unity) and c) to have a compact support. The last condition allows us to construct test spaces that vanish at the Dirichlet boundaries and thus, advocating its admittance into weak formulations. However, this condition makes these basis functions discontinuous at the element boundaries. In other words, these basis functions belong to the class of *regular*[†] generalized functions and their derivatives must be understood in the sense of a distribution. Hence the test space spanned by these basis functions is nonconforming. As the row lumping technique is a critical step in the design of these basis functions (to fulfill the partition of unity constraint), the current work is restricted only to those FEs where this technique makes sense—simplicial FEs and multi-linear block FEs. We show that using these basis functions with an appropriate single-valued model on the element boundaries, it is possible to recover the classical FDM stencil of the Helmholtz equation on structured meshes. The linear interpolation on the element boundaries (specified by α_1) and the element interiors (specified by α_2) of these basis functions with the standard FE shape functions, will result in a new class of basis functions. These new basis functions involving two parameters α_1, α_2 define the test space of the proposed PG method that yields the nonstandard compact stencil of the Helmholtz equation on structured meshes. The proposed PG method provides the counterpart of the aforesaid scheme on unstructured meshes and allows the treatment of natural boundary conditions (Neumann or Robin) and the source terms in a straight-forward manner.

This paper is organized as follows. In Section 2, we present the statement of the Helmholtz equation in both the strong and the weak forms. In Section 3, we summarize the salient features of the domain-based higher-order compact schemes proposed for the Helmholtz equation in [1]. In Section 4, we present the trial and test spaces involved with the PG method proposed in this paper. The trial space is spanned by the standard conforming FE shape functions. The basis functions that span the test space are defined using the FE shape functions in a piecewise manner in the interiors and the edges of the elements, respectively. The precise definition of the basis functions just in the interior of the elements is given here and only the properties of the basis functions on the element edges are discussed. The weak form associated with the PG method is also presented here. In Section 5, the weak form of the proposed PG method which involves distributional derivatives of the test functions is presented in a form that is easier to compute and implement. Remarks are also made here on the possibility to attain the sparsity pattern of the Galerkin FEM. In Section 6, we present the definition of the basis functions on the element edges for the 1D linear and 2D bilinear FEs. It is also shown here that on structured meshes the proposed PG method is able to recover the higher-order compact schemes summarized earlier in Section 3. In Section 7, using some structured simplicial meshes it is shown that the alpha-interpolation of the FEM and FDM stencils would yield a scheme identical to the alpha-interpolation method (AIM) [4, 5] wherein the mass matrix that appears in the Galerkin FEM is replaced by an alpha-interpolated mass matrix. The dispersion accuracy of the schemes is discussed here and remarks are made on recovering the AIM via the proposed PG method. Some examples are presented in Section 8, using uniform and nonuniform meshes in 2D made up of bilinear FEs. These examples illustrate the pollution effect associated with the proposed PG method through convergence studies in the L^2 norm, the H^1 semi-norm and the l^∞ Euclidean norm. Finally in Section 9, we arrive at some conclusions.

[†]to distinguish with *singular* generalized functions such as the Dirac's 'delta-function'. For further details see [2, 3].

DOI: 10.1002/nme

2. PROBLEM STATEMENT

The statement of the multidimensional Helmholtz equation in the strong form is,

$$\mathcal{L}\phi := -\Delta\phi - \xi_o^2\phi = f(\mathbf{x}) \quad \text{in} \quad \Omega \tag{1a}$$
$$\phi = \phi^p \quad \text{on} \quad \Gamma_{\mathrm{D}} \tag{1b}$$
$$\mathbf{n}\cdot\boldsymbol{\nabla}\phi - \mathcal{M}\phi = q \quad \text{on} \quad \Gamma_{\mathrm{R}} \tag{1c}$$

where ξ_o is the wavenumber, $f(\mathbf{x})$ is the source term, ϕ^p is the prescribed value of ϕ on the Dirichlet boundary Γ_{D}. The operator $\mathcal{M}$ models either the Dirichlet-to-Neumann (DtN) map should the boundary-value problem (BVP) be posed on a domain with an exterior DtN boundary or the Neumann/Robin boundary conditions should the BVP be posed on an interior domain.

For the solution of the BVP (1) we introduce the following set of functions:

$$H^1_{\mathrm{E}} := \{\psi : \psi \in H^1(\Omega) \text{ and } \psi = \phi^p \text{ on } \Gamma_{\mathrm{D}}\} \tag{2a}$$
$$H^1_0 := \{\psi : \psi \in H^1(\Omega) \text{ and } \psi = 0 \text{ on } \Gamma_{\mathrm{D}}\} \tag{2b}$$

where $H^m(\Omega)$ is the usual Sobolev space of functions with m^{th} derivatives square integrable. The variational statement of the BVP (1) can be expressed as follows: Find $\phi \in H^1_{\mathrm{E}}(\Omega)$ such that,

$$B(\psi,\phi) = F(\psi) \quad \forall\psi \in H^1_0(\Omega) \tag{3a}$$
$$B(\psi,\phi) := \int_\Omega \left[\boldsymbol{\nabla}\psi\cdot\boldsymbol{\nabla}\phi - \xi_o^2\psi\phi\right]\,\mathrm{d}\Omega - \int_{\Gamma_{\mathrm{R}}} \psi\mathcal{M}\phi\,\mathrm{d}\Gamma \tag{3b}$$
$$F(\psi) = \int_\Omega \psi f\,\mathrm{d}\Omega + \int_{\Gamma_{\mathrm{R}}} \psi q\,\mathrm{d}\Gamma \tag{3c}$$

Let $U^h \subset H^1(\Omega)$ and $V^h \subset H^1(\Omega)$ be subspaces obtained via any appropriate discretization with h being the discretization size parameter. Then corresponding to Equation (2) we define

$$U^h_{\mathrm{E}} := \{\psi_h : \psi_h \in U^h \text{ and } \psi_h = \phi^p_h \text{ on } \Gamma_{\mathrm{D}}\} \tag{4a}$$
$$V^h_0 := \{\psi_h : \psi_h \in V^h \text{ and } \psi_h = 0 \text{ on } \Gamma_{\mathrm{D}}\} \tag{4b}$$

The statement of the so-called generalized Galerkin method applied to the weak form of the BVP (1) is: Find $\phi_h \in U^h_{\mathrm{E}}$ such that,

$$B(\psi_h,\phi_h) = F(\psi_h) \quad \forall\psi_h \in V^h_0 \tag{5}$$

Taking the discrete test and trial spaces to be distinct, i. e. $U^h \neq V^h$, we get a Petrov–Galerkin method. Otherwise, i. e. taking $U^h = V^h$ we get a Bubnov–Galerkin method. Discretizing both the trial and test spaces by finite elements we obtain the standard Galerkin FEM. This leads to the approximations $\phi_h = N^a\Phi^a$, $\psi_h = N^a\Psi^a$ and Equation (5) reduces into the following system of equations:

$$\left[\mathbf{D} - \xi_o^2\mathbf{M} - \mathbf{R}\right]\boldsymbol{\Phi} = \mathbf{f} \tag{6a}$$
$$\mathrm{D}^{ab} = \int_\Omega \boldsymbol{\nabla}N^a\cdot\boldsymbol{\nabla}N^b\,\mathrm{d}\Omega, \quad \mathrm{M}^{ab} = \int_\Omega N^aN^b\,\mathrm{d}\Omega, \quad \mathrm{R}^{ab} = \int_{\Gamma_{\mathrm{R}}} N^a\mathcal{M}N^b\,\mathrm{d}\Gamma \tag{6b}$$
$$\mathrm{f}^a = \int_\Omega N^a f(\mathbf{x})\,\mathrm{d}\Omega + \int_{\Gamma_{\mathrm{R}}} N^a q\,\mathrm{d}\Gamma \tag{6c}$$

3. ALPHA-INTERPOLATION OF FEM AND FDM STENCILS

Consider the BVP (1) posed on an interior 2D domain subjected to Dirichlet boundary conditions and let $f(\mathbf{x}) = 0$. Further, let the domain be such that it permits a partition of the same using a

DOI: 10.1002/nme

structured mesh consisting of rectangular bilinear FEs. For the considered case we use the following notation to represent a generic compact stencil corresponding to any interior node (i,j) of the structured mesh.

$$\{\circ^{j+1},\circ^{j},\circ^{j-1}\}\mathbf{S}\{\circ^{i-1},\circ^{i},\circ^{i+1}\}^{\mathrm{t}}=0 \tag{7}$$

where $\mathbf{S}$ represents the matrix of the stencil coefficients. For instance, if the standard mass matrix obtained in the Galerkin FEM be assembled for a structured rectangular mesh then we may express the corresponding stencil as follows:

$$\mathbf{S}^{\mathrm{m}}:=\frac{\ell_1\ell_2}{36}\{1,4,1\}^{\mathrm{t}}\{1,4,1\}=\frac{\ell_1\ell_2}{36}\begin{bmatrix}1&4&1\\4&16&4\\1&4&1\end{bmatrix} \tag{8}$$

$$\{\circ^{j+1},\circ^{j},\circ^{j-1}\}\mathbf{S}^{\mathrm{m}}\{\circ^{i-1},\circ^{i},\circ^{i+1}\}^{\mathrm{t}}:=\frac{\ell_1\ell_2}{36}\left\{\begin{array}{l}(\Phi^{i-1,j+1}+4\Phi^{i,j+1}+\Phi^{i+1,j+1})+\\(4\Phi^{i-1,j}+16\Phi^{i,j}+4\Phi^{i+1,j})+\\(\Phi^{i-1,j-1}+4\Phi^{i,j-1}+\Phi^{i+1,j-1})\end{array}\right\} \tag{9}$$

The equation stencil for the Galerkin FEM method corresponding to any interior node (i,j) can be written as Equation (7) with the following definition of the stencil coefficient matrix ($\mathbf{S}$):

$$\mathbf{S}^{\mathrm{fem}}:=\frac{\ell_2}{6\ell_1}\{1,4,1\}^{\mathrm{t}}\{-1,2,-1\}+\frac{\ell_1}{6\ell_2}\{-1,2,-1\}^{\mathrm{t}}\{1,4,1\}-\frac{\xi_o^2\ell_1\ell_2}{36}\{1,4,1\}^{\mathrm{t}}\{1,4,1\} \tag{10}$$

The stencil for the classical FDM method corresponding to any interior node (i,j) can be written as Equation (7) with the following definition of $\mathbf{S}$:

$$\mathbf{S}^{\mathrm{fdm}}:=\frac{\ell_2}{6\ell_1}\{0,6,0\}^{\mathrm{t}}\{-1,2,-1\}+\frac{\ell_1}{6\ell_2}\{-1,2,-1\}^{\mathrm{t}}\{0,6,0\}-\frac{\xi_o^2\ell_1\ell_2}{36}\{0,6,0\}^{\mathrm{t}}\{0,6,0\} \tag{11}$$

The nonstandard compact stencil presented in [1] can be written as Equation (7) with the following definition of $\mathbf{S}$:

$$\begin{aligned}\mathbf{S}^{\alpha_1,\alpha_2}:=\quad&(1-\alpha_1)\frac{\ell_2}{6\ell_1}\{1,4,1\}^{\mathrm{t}}\{-1,2,-1\}+\alpha_1\frac{\ell_2}{6\ell_1}\{0,6,0\}^{\mathrm{t}}\{-1,2,-1\}\\&+(1-\alpha_1)\frac{\ell_1}{6\ell_2}\{-1,2,-1\}^{\mathrm{t}}\{1,4,1\}+\alpha_1\frac{\ell_1}{6\ell_2}\{-1,2,-1\}^{\mathrm{t}}\{0,6,0\}\\&-(1-\alpha_2)\frac{\xi_o^2\ell_1\ell_2}{36}\{1,4,1\}^{\mathrm{t}}\{1,4,1\}-\alpha_2\frac{\xi_o^2\ell_1\ell_2}{36}\{0,6,0\}^{\mathrm{t}}\{0,6,0\}\end{aligned} \tag{12}$$

where α_1,α_2 are two nondimensional parameters. Note that we can obtain $\mathbf{S}^{\alpha_1,\alpha_2}$ by taking the linear interpolation of the diffusive terms (specified by α_1) and the mass terms (specified by α_2) that appear in $\mathbf{S}^{\mathrm{fem}}$ and $\mathbf{S}^{\mathrm{fdm}}$, respectively. Taking $\alpha_1=\alpha_2=\alpha$, we arrive at a stencil that is the α-interpolation of the Galerkin FEM and the classical central FDM stencils. i. e. $\mathbf{S}^{\alpha,\alpha}=(1-\alpha)\mathbf{S}^{\mathrm{fem}}+\alpha\mathbf{S}^{\mathrm{fdm}}$. Choosing $\alpha_1=\alpha_2=0.5$ we get a stencil that is the average of the FEM and FDM stencils in 2D and it can be shown [1] to be equal to the stencil obtained by the generalized fourth-order compact Padé approximation [6, 7] (therein using the parameter $\gamma=2$). Likewise taking $\alpha_1=0$ and $\alpha_2=\alpha$ we get a stencil that results from the Galerkin FEM using an α-interpolated mass matrix $\mathbf{M}^{\alpha}:=(1-\alpha)\mathbf{M}+\alpha\mathbf{M}_{\mathrm{L}}$. Taking $\alpha_1=0$ and $\alpha_2=0.5$, we obtain the higher-order mass matrix scheme introduced in [8, 9], which can also be obtained using special quadrature rules (cf. [10, pp. 446]). Further details on the choice of the parameters to recover various stencils can be found in [1].

Considering square meshes, i. e. $\ell_1=\ell_2=\ell$, the parameters α_1,α_2 that appear in $\mathbf{S}^{\alpha_1,\alpha_2}$ can be chosen such that the numerical solution be sixth-order accurate, i. e. $O\left((\xi_o\ell)^6\right)$ or equivalently $O\left(\omega^3\right)$ where $\omega:=(\xi_o\ell)^2$. Recall that this is the maximum order of dispersion accuracy that can be attained on any compact stencil [11]. All such α_1,α_2 should obey the following series expansion in terms of ω.

$$\alpha_1=\frac{1}{2}-\frac{\omega}{60}+\sum_{m=2}^{\infty}a_m\omega^m\quad;\quad\alpha_2=\frac{1}{2}-\frac{\omega}{40}+\sum_{m=2}^{\infty}b_m\omega^m \tag{13}$$

DOI: 10.1002/nme

where a_m, b_m are coefficients independent of ω. The relative phase error $\mathbb{P}$ and local truncation error $\mathbb{T}$ of these sixth-order schemes can be expressed as follows:

$$\mathbb{P} = r_3\omega^3 + O\left(\omega^4\right), \quad \mathbb{T} = -2r_3\omega^3 + O\left(\omega^4\right) \tag{14}$$

$$r_3 = \left[\frac{5}{110592} - \left(\frac{a_2 - 4b_2}{48}\right) + \left(\frac{1 + 576a_2}{27648}\right)\cos(4\beta) + \frac{\cos(8\beta)}{774144}\right] \tag{15}$$

As a_m, b_m $(m \geq 2)$ can be chosen arbitrarily, infinitely many sixth-order schemes can be designed through $\mathbf{S}^{\alpha_1,\alpha_2}$. Of course some particular choice of a_m, b_m may yield a scheme with better features. For instance, a_m, b_m may be chosen such that the local truncation error $\mathbb{T}$ be zero along some chosen directions. Choosing $\alpha_1 = \alpha_2$, the dispersion accuracy can at most be fourth-order and to attain this all such α_1, α_2 should obey the following series expansion in terms of ω.

$$\alpha_1 = \alpha_2 = \frac{1}{2} + \sum_{m=1}^{\infty} a_m\omega^m \tag{16}$$

The relative phase error $\mathbb{P}$ and local truncation error $\mathbb{T}$ of these fourth-order schemes can be expressed as follows:

$$\mathbb{P} = r_2\omega^2 + O\left(\omega^3\right), \quad \mathbb{T} = -2r_2\omega^2 + O\left(\omega^3\right) \tag{17}$$

$$r_2 = \left[\left(\frac{1 + 36a_1}{576}\right) + \left(\frac{1 + 60a_1}{2880}\right)\cos(4\beta)\right] \tag{18}$$

4. TRIAL AND TEST SPACES

We use standard conforming finite elements to construct the trial space U^h. Thus, if N^a represents the corresponding FE shape function of an arbitrary node a and Φ^a the corresponding nodal unknowns of the FE discretization, then every $\phi_h \in U^h$ can be expressed as $\phi_h = N^a\Phi^a$. Let $\mathbf{M}$ denote the mass matrix resulting from the inner product of the FE shape functions, cf. Equation 6b, and let $\mathbf{M}_\mathrm{L}$ denote the lumped mass matrix. The standard row lumping technique used to obtain $\mathbf{M}_\mathrm{L}$ from the consistent mass matrix $\mathbf{M}$ can be expressed as follows.

$$\mathbf{M}_\mathrm{L}^{ab} := \delta^{ab}\sum_c \mathbf{M}^{ac} = \delta^{ab}\sum_c \int_\Omega N^a N^c \,\mathrm{d}\Omega = \delta^{ab}\int_\Omega N^a \,\mathrm{d}\Omega \tag{19}$$

The fact that the shape functions N^a being a partition of unity is used to arrive at the last part of Equation (19). For every element K, we define a local transformation matrix $\mathbb{W}$ as follows:

$$\mathbb{W} = (1 - \alpha_2)\mathbf{I} + \alpha_2\mathbf{M}_\mathrm{L}\mathbf{M}^{-1} \tag{20}$$

where, $\mathbf{I}$ is the identity matrix and α_2 is a nondimensional parameter. We now construct new basis functions $\widetilde{W}^a$ using the shape functions N^a of the trial space. As is done for the FE shape functions N^a, we first define $\widetilde{W}^a$ locally for each element and later patch them together to obtain their global definitions. Thus within each element K, we define

$$\widetilde{W}^a := \mathbb{W}^{ab}N^b \tag{21}$$

Usually the above definition of $\widetilde{W}^a$ will lead to a loss of C^0 continuity at the element edges. By construction the domain of $\widetilde{W}^a$ is defined as the interior of the elements. Hence the support of the global basis function $\widetilde{W}^a$, denoted as $\mathrm{supp}(\widetilde{W}^a)$, is the interior of the elements of a patch containing the node a.

Prepared using nmeauth.cls
DOI: 10.1002/nme

Table I. Local definition of the basis $\widetilde{W}^a$ corresponding to some finite elements

Shape functions N^a	Basis $\widetilde{W}^a$ (choosing $\alpha_2 = 1$)	Remarks
$\frac{1+\bar{\xi}^a\xi}{2}$	$\frac{1+3\bar{\xi}^a\xi}{2}$	1D linear FE. $\{\bar{\xi}^a\} = \{-1, 1\}$
$\left(\frac{1+\bar{\xi}^a\xi}{2}\right)\left(\frac{1+\bar{\eta}^a\eta}{2}\right)$	$\left(\frac{1+3\bar{\xi}^a\xi}{2}\right)\left(\frac{1+3\bar{\eta}^a\eta}{2}\right)$	2D rectangular bilinear FE. $\{\bar{\xi}^a\} = \{-1, 1, 1, -1\}$, $\{\bar{\eta}^a\} = \{-1, -1, 1, 1\}$.
$\{(1-\xi-\eta), \xi, \eta\}$	$\{(3-4\xi-4\eta), (4\xi-1), (4\eta-1)\}$	2D linear triangle FE.

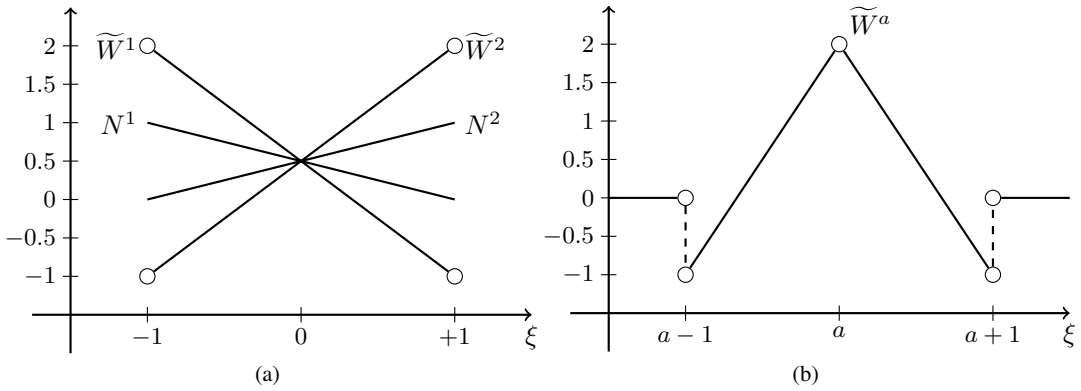


Figure 1. Basis functions $\widetilde{W}^a$ corresponding to the 1D linear FE and choosing $\alpha_2 = 1$. (a) Element-wise comparison of $\widetilde{W}^a$ with the 1D FE shape functions N^a. (b) Illustration of the global basis $\widetilde{W}^a$ corresponding to an arbitrary node a patched element-wise. The open circles in these illustrations signify that the function values taken by $\widetilde{W}^a$ on the element edges are omitted.

If $\mathbf{M}_\mathrm{L}$ be obtained via the row lumping technique, then the basis functions $\widetilde{W}^a$ also form a partition of unity. This statement can be verified within each element as follows:

$$\sum_a \mathrm{M}_\mathrm{L}^{ac}\mathrm{M}^{-cb} = \sum_a \mathrm{M}^{ac}\mathrm{M}^{-cb} = \sum_a \delta^{ab} \tag{22a}$$

$$\Rightarrow \sum_a \mathbb{W}^{ab} = \sum_a (1-\alpha_2)\delta^{ab} + \alpha_2 \mathrm{M}_\mathrm{L}^{ac}\mathrm{M}^{-cb} = \sum_a \delta^{ab} \tag{22b}$$

$$\Rightarrow \sum_a \widetilde{W}^a = \sum_a \mathbb{W}^{ab}N^b = \sum_a \delta^{ab}N^b = \sum_a N^a = 1 \tag{22c}$$

where M^{-cb} denotes the entry corresponding to the indices (c, b) of the matrix $\mathbf{M}^{-1}$. Note that the global basis functions $\widetilde{W}^a$, i. e. the element contributions obtained from Equation (21) patched in a piecewise manner, are no longer a linear combination of the global FE shape functions N^a. Thus, $\widetilde{W}^a$ will span a function space distinct from the trial space U^h but with the same dimensions as U^h. The local expressions (using $\alpha_2 = 1$) of the basis functions $\widetilde{W}^a$ corresponding to three different element types is listed in Table I. Figure 1 illustrates the construction of the global basis $\widetilde{W}^a$ corresponding to the 1D linear FE shape functions. Note the loss of C^0 continuity at the element edges in Figure 1b. The open circles shown in Figure 1b signify that the function values taken by $\widetilde{W}^a$ on the element edges are omitted.

We now construct a new composite basis W^a defined in a piecewise manner as follows:

$$W^a(x) := \begin{cases} \widetilde{W}^a(x) & \text{in the element interiors} \\ \widehat{W}^a(x) & \text{on the element edges} \end{cases} \tag{23}$$

DOI: 10.1002/nme

The composite basis W^a is introduced for two reasons, viz. a) to ensure that W^a be a partition of unity also on the element edges and b) to be able to model $\widehat{W}^a$ such that we recover the sparsity pattern of the Galerkin FEM. The later condition also allows us to construct test spaces with functions that vanish on the Dirichlet boundary. Thus by construction we require that $\widehat{W}^a$ be single-valued functions on the element edges with the following properties:

$$\sum_a \widehat{W}^a = 1, \quad \text{and} \quad \text{supp}(\widehat{W}^a) = \text{supp}(N^a|_{\mathcal{E}_h}) \tag{24}$$

where $\mathcal{E}_h$ represents the collection of all the element edges and $N^a|_{\mathcal{E}_h}$ represents the restriction of N^a on $\mathcal{E}_h$. The precise definition of $\widehat{W}^a$ is delayed until Section 6.

From the properties of $\widetilde{W}^a$ and $\widehat{W}^a$ we have $\text{supp}(W^a) = \text{supp}(N^a)$. Note that the global basis W^a are *regular* generalized functions, i. e. they are ordinary functions with a predefined jump discontinuity at the element edges. Thus using Equations (20) and (21), the following result is straight-forward.

$$\int_\Omega W^a N^b \, \mathrm{d}\Omega = \sum_K \int_K \widetilde{W}^a N^b \, \mathrm{d}\Omega = (1-\alpha_2)\mathbf{M}^{ab} + \alpha_2 \mathbf{M}_{\mathrm{L}}^{ab} \tag{25}$$

We use the composite basis W^a to construct the test space (denoted as V^{h*}) of the PG method. Therefore every $w_h \in V^{h*}$ can be expressed as $w_h = W^a \Psi^a$, where Ψ^a is an arbitrary constant associated with the node a. By construction the test space V^{h*} has the same dimensions as that of the trial space U^h. In the notation V^{h*}, the symbol '$*$' is used to emphasize that generally the test space V^{h*} is nonconforming, i. e. $V^{h*} \not\subset H^1(\Omega)$. The statement of the proposed PG method applied to the BVP (1) is : Find $\phi_h \in U^h_{\mathrm{E}}$ such that,

$$B(w_h, \phi_h) = F(w_h) \quad \forall w_h \in V_0^{h*} \tag{26}$$

In order to compute the integral $\int_\Omega \boldsymbol{\nabla} w_h \cdot \boldsymbol{\nabla} \phi_h \, \mathrm{d}\Omega$ that appears in Equation (26), the derivatives associated with w_h must be understood in the sense of a distribution. Thus the proper setting for the PG method is in the space of generalized functions. Recall that to arrive at the term $\int_\Omega \boldsymbol{\nabla} w_h \cdot \boldsymbol{\nabla} \phi_h \, \mathrm{d}\Omega$ integration by parts needs to be done for an integral form of Equation (1a) containing discontinuous test functions. This is the distinction of the current work from existing stabilized FEM based PG methods that follows the theoretical framework originally proposed for the Streamline–Upwind/Petrov–Galerkin (SUPG) method [12].

The distinction of the current work with Discontinuous–Galerkin (DG) methods is illustrated via a schematic representation of the same in Figure 2. Figure 2a illustrates a generic DG method. Recall that the weights on either side of an element edge in a DG method are not only discontinuous but are also independent. The same applies to the trial solutions (ϕ_h) and in addition to this, models $\widehat{\phi}_h$ for ϕ_h are specified on the element edges. For conservative DG methods, $\widehat{\phi}_h$ which is sometimes named as the *scalar numerical flux*, is single valued on the element edges [13]. On the other hand, Figure 2b illustrates the current PG method. Note that the test functions (w_h) remain discontinuous but they are no longer independent. The restriction of w_h to the element interiors and on the element edges are denoted as $\widetilde{w}_h$ and $\widehat{w}_h$, respectively. The trial solutions for the current PG method are the standard FE solutions which are C^0-continuous and are not independent on either sides of the element edge.

The proposed PG method given by Equation (26) is similar to the generalized difference method (GDM) presented in [14] and the finite volume methods (FVM) analyzed in [15]. The similarity is in the nature of the trial and the test spaces—the trial space is taken as the standard FE space and the test space consist of regular generalized functions. The distinction is in the definition of the test space. In our work both the trial and test spaces are defined on the primary partition of the domain. In the GDM and the FVM the test spaces are defined on the dual partition of a given primary partition.

Finally, we make note that test functions similar to $\widetilde{W}^a$ (with $\alpha_2 = 1$) were introduced earlier in the context of dual mortar methods for non-overlapping domain decomposition techniques [16].

DOI: 10.1002/nme

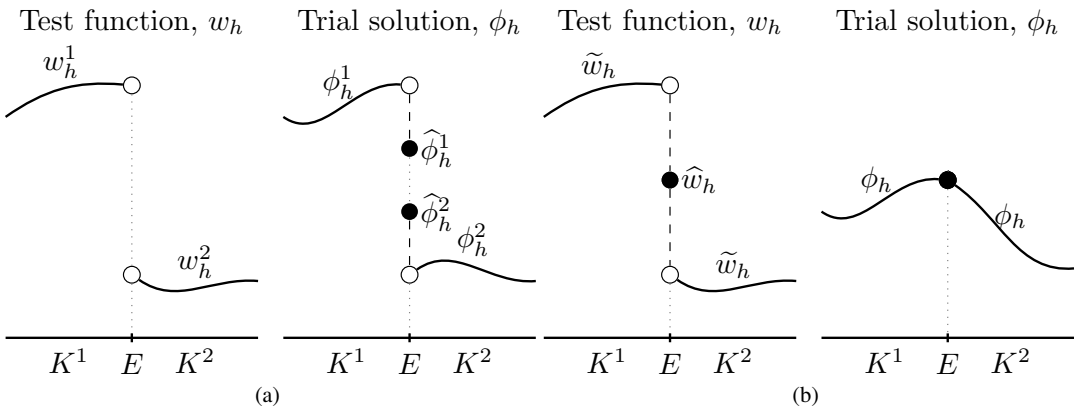


Figure 2. Comparison of the test function w_h and trial solution ϕ_h of a generic Discontinuous–Galerkin (DG) method with those of the current Petrov–Galerkin (PG) method. Schematic representations of w_h and ϕ_h for (a) a DG method and (b) the current PG method. Note that unlike for the DG method, w_h and ϕ_h for the current PG method are not independent on either sides of the edge E. Also note that w_h is a regular generalized function and its derivatives must be understood in the sense of a distribution.

Therein, such test functions were called as *local dual basis functions* and were used to construct the discrete Lagrange multiplier space. In the standard mortar methods, the interface solution on the slave side depends globally on the values on the master side. The motivation for the introduction of the local dual basis functions is to reduce this global dependence to a local one without compromising the a priori error estimates obtained for the standard mortar methods. Thus, the mortar map is represented by a diagonal matrix which allows the matching/coupling condition to be realized explicitly.

5. WEAK FORM WITH GENERALIZED TEST FUNCTIONS

In this section we express Equation (26) in a form that is easier to compute and implement. By choosing a regularization parameter ε, we first construct a sequence of piecewise continuous test functions $w_h^\varepsilon \in V^h \subset H^1(\Omega)$ which converges to $w_h \in V^{h*}$ as $\varepsilon \to 0$. Substituting w_h^ε in Equation (5) and taking the limit $\varepsilon \to 0$ we get the weak form of the PG method using w_h as defined earlier in Equation (26).

Consider an arbitrary element K with boundary ∂K and define two sub-domains within it viz. K_o and K_ε as shown in Fig 3a. The boundary that K_ε shares with K_o is denoted by ∂K_o. The external normals to ∂K and ∂K_o are denoted by $\mathbf{n}$ and $\mathbf{n}^{o+}$ respectively. The normal $\mathbf{n}^{o-} := -\mathbf{n}^{o+}$. The regularization parameter ε characterizes the width of the K_ε sub-domain. Consider a regularized piecewise continuous test function w_h^ε over K whose definition can be split over K_o, K_ε and ∂K as follows:

$$w_h^\varepsilon(x) := \begin{cases} \widetilde{w}_h(x) & \forall x \in K_o \\ \tau_h(x) & \forall x \in K_\varepsilon \\ \widehat{w}_h(x) & \forall x \in \partial K \end{cases} \tag{27}$$

$$\widetilde{w}_h|_{\partial K_o} = \tau_h|_{\partial K_o}, \quad \tau_h|_{\partial K} = \widehat{w}_h \tag{28}$$

Thus, as shown in Figure 3b, taking the limit $\varepsilon \to 0$ the test function w_h^ε develops a sharp layer at the element boundary ∂K and we arrive at a class of generalized test function w_h which was represented schematically in Figure 2b. Likewise, in the limit $\varepsilon \to 0$, the term $\boldsymbol{\nabla} w_h^\varepsilon$ will represent the generalized derivative of the test function w_h. Consider the term $\int_\Omega \boldsymbol{\nabla} w_h^\varepsilon \cdot \boldsymbol{\nabla} \phi_h \, \mathrm{d}\Omega$ which can

DOI: 10.1002/nme

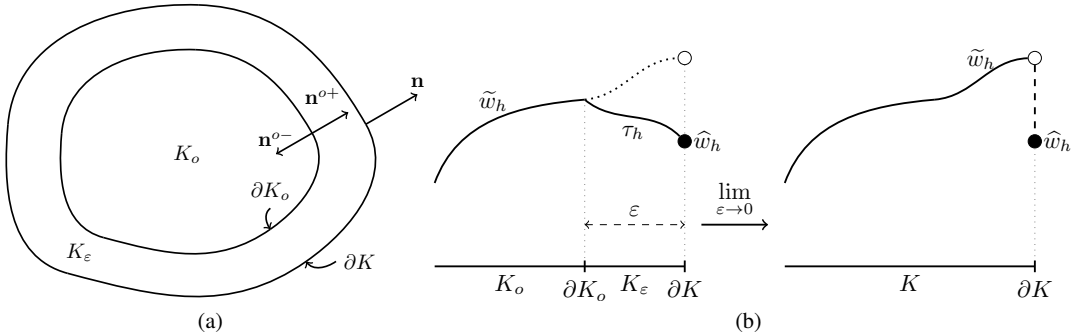


Figure 3. Schematic diagrams of an arbitrary element K and a regularized test function w_h^ε defined over it. (a) The element K is further divided into two sub-domains K_o and K_ε. (b) The test function w_h^ε is defined piecewise as follows: $\widetilde{w}_h$ over K_o, τ_h over K_ε, $\widehat{w}_h$ on ∂K and with $\widetilde{w}_h|_{\partial K_o} = \tau_h|_{\partial K_o}$. The regularization parameter ε characterizes the width of the K_ε domain and taking the limit $\varepsilon \to 0$ we recover the generalized function w_h

be written in the following equivalent forms.

$$\int_\Omega \boldsymbol{\nabla} w_h^\varepsilon \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega = \sum_K \left[\int_{K_o} \boldsymbol{\nabla}\widetilde{w}_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega + \int_{K_\varepsilon} \boldsymbol{\nabla}\tau_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega\right] \tag{29a}$$

$$= \sum_K \left[\int_{K_o} \boldsymbol{\nabla}\widetilde{w}_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega - \int_{K_\varepsilon} \tau_h \Delta\phi_h \,\mathrm{d}\Omega - \int_{\partial K_o} \widetilde{w}_h \mathbf{n}^{o+} \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Gamma + \int_{\partial K} \widehat{w}_h \mathbf{n} \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Gamma\right] \tag{29b}$$

$$= \sum_K \left[-\int_{K_o} \widetilde{w}_h \Delta\phi_h \,\mathrm{d}\Omega - \int_{K_\varepsilon} \tau_h \Delta\phi_h \,\mathrm{d}\Omega + \int_{\partial K} \widehat{w}_h \mathbf{n} \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Gamma\right] \tag{29c}$$

Equations (29b) and (29c) result from integrating by parts the corresponding terms in Equation (29a) and using the relations in Equation (28). Taking the limit $\varepsilon \to 0$ on both sides of Equation (29) we get,

$$\int_\Omega \boldsymbol{\nabla} w_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega = \int_{\mathcal{T}_h} \boldsymbol{\nabla}\widetilde{w}_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega + \lim_{\varepsilon\to 0} \sum_K \int_{K_\varepsilon} \boldsymbol{\nabla}\tau_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega \tag{30a}$$

$$= \int_{\mathcal{T}_h} \boldsymbol{\nabla}\widetilde{w}_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega + \sum_K \int_{\partial K} (\widehat{w}_h - \widetilde{w}_h)\mathbf{n} \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Gamma \tag{30b}$$

$$= -\int_{\mathcal{T}_h} \widetilde{w}_h \Delta\phi_h \,\mathrm{d}\Omega + \sum_K \int_{\partial K} \widehat{w}_h \mathbf{n} \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Gamma \tag{30c}$$

where $\int_{\mathcal{T}_h}$ represents the piecewise integral $\sum_K \int_K$. Note that the integral $\int_{K_\varepsilon}$ that appears in Equation (30a) does not vanish as $\varepsilon \to 0$. On the other hand, using the form expressed in Equation (30b) the extra labor just involves the evaluation of the element boundary integrals. This can be easily incorporated within an 'assemble-by-elements' data structure. Hence for the implementation of the PG method, we use Equation (30b) to compute the bilinear form in Equation (26).

$$\boxed{\begin{aligned} B(w_h, \phi_h) := &\int_{\mathcal{T}_h} \boldsymbol{\nabla}\widetilde{w}_h \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Omega + \sum_K \int_{\partial K} (\widehat{w}_h - \widetilde{w}_h)\mathbf{n} \cdot \boldsymbol{\nabla}\phi_h \,\mathrm{d}\Gamma \\ &- \int_{\mathcal{T}_h} \xi_o^2 \widetilde{w}_h \phi_h \,\mathrm{d}\Omega - \int_{\Gamma_R} \widehat{w}_h \mathcal{M}\phi_h \,\mathrm{d}\Gamma \end{aligned}} \tag{31}$$

We get the discrete system matrix by making the approximations $\phi_h = N^a \Phi^a$, $w_h = W^a \Psi^a$ and substituting it into Equation (31). Note that the approximation $w_h = W^a \Psi^a$ implies $\widetilde{w}_h = \widetilde{W}^a \Psi^a$

DOI: 10.1002/nme

and $\widehat{w}_h = \widehat{W}^a \Psi^a$. However, to remark on the sparsity pattern of the discrete system, it is more appropriate to express the bilinear form in Equation (26) using Equation (30c). Following this line, the discrete system matrix $\mathbf{A}$ can be expressed as follows:

$$\mathrm{A}^{ab} = -\int_{\mathcal{T}_h} \widetilde{W}^a \Delta N^b \,\mathrm{d}\Omega + \sum_K \int_{\partial K} \widehat{W}^a \mathbf{n} \cdot \boldsymbol{\nabla} N^b \,\mathrm{d}\Gamma - \int_{\mathcal{T}_h} \xi_o^2 \widetilde{W}^a N^b \,\mathrm{d}\Omega - \int_{\Gamma_{\mathrm{R}}} \widehat{W}^a \mathcal{M} N^b \,\mathrm{d}\Gamma \tag{32}$$

Recall that by construction, $\mathrm{supp}(\widetilde{W}^a)$ is the interior of the elements of a patch containing the node a. Further, if $\widehat{W}^a$ be designed such that $\mathrm{supp}(\widehat{W}^a) = \mathrm{supp}(N^a|_{\mathcal{E}_h})$, then from Equation (32) we see that the resulting discrete system $\mathbf{A}$ will have a sparsity pattern equivalent to that of the Galerkin FEM. In other words, to attain the Galerkin FEM sparsity pattern, $\widehat{W}^a$ should be zero wherever N^a be zero.

6. BLOCK FINITE ELEMENTS

In this section we complete the definition of the composite basis functions W^a of the test space given by Equation (23) when the trial spaces are spanned by the lowest order block finite elements. In other words we define $\widehat{W}^a$ on the element edges for the 1D linear and the 2D bilinear FEs.

6.1. 1D linear FE

Recall that by construction the basis functions $\widetilde{W}^a$ are defined only in the interior of the elements. Hence in Figure 1b, open circles were used to indicate that the values taken by $\widetilde{W}^a$ on the nodes $a-1$, a and $a+1$ are omitted. Let $\widehat{W}^a|_{a-1}$, $\widehat{W}^a|_a$ and $\widehat{W}^a|_{a+1}$ be the corresponding function values assigned to $\widehat{W}^a$ on these edges. For $\widehat{W}^a$ to be a partition of unity on the element edges the following relation should hold:

$$\widehat{W}^a|_{a-1} + \widehat{W}^a|_a + \widehat{W}^a|_{a+1} = 1 \tag{33}$$

There exists an infinity of solutions for Equation (33) but only the choice $\{\widehat{W}^a|_{a-1}, \widehat{W}^a|_a, \widehat{W}^a|_{a+1}\} = \{0, 1, 0\}$ satisfies the properties given in Equation (24). This choice will result in a discrete system that has the same sparsity structure as that of the Galerkin FEM or the classical FDM. Also, the space spanned by these weights can be restricted to zero on the Dirichlet boundary without being trivially zero inside the domain and thus, justifying their admittance in weak formulations. Thus using the 1D linear FE, Equation (24) is satisfied if and only if $\widehat{W}^a$ is defined as follows.

$$\widehat{W}^a = N^a|_{\mathcal{E}_h} \tag{34}$$

Following Equation (23) and using Equations (21) and (34) the local expression of the composite basis W^a within each element can be expressed as,

$$W^a = \begin{cases} \dfrac{1 + (1+2\alpha_2)\xi\bar{\xi}^a}{2} & -1 < \xi < 1 \\ \dfrac{1 + \xi\bar{\xi}^a}{2} & \xi = \pm 1 \end{cases} \tag{35}$$

Choosing $\alpha_2 = 0$ in Equation (35) the composite basis W^a simplifies to the standard 1D FE shape functions N^a. Likewise, choosing $\alpha_2 = 1$ the composite basis W^a corresponding to the 1D linear FE can be represented as shown in Figure 4.

Consider the BVP (1) subjected to Dirichlet boundary conditions and let $f = 0$. Using W^a given by Equation (35) in the weak form of the PG method given by Equation (26), the following equation

DOI: 10.1002/nme

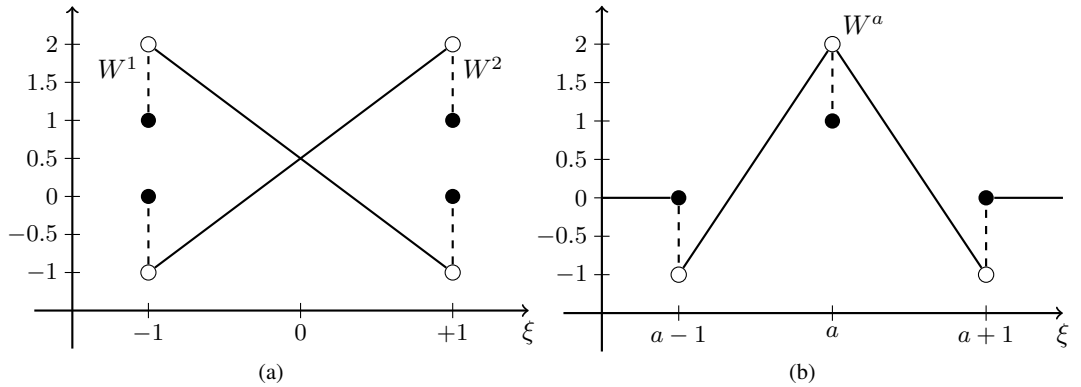


Figure 4. The composite basis functions W^a corresponding to the 1D linear FE and choosing $\alpha_2 = 1$. The filled circles in these illustrations represent the chosen model for $\widehat{W}^a$ on the element edges. (a) Illustration of the basis W^a defined locally within an element. (b) The global basis W^a corresponding to an arbitrary node a patched element-wise.

stencil is obtained:

$$\left(\frac{1}{\ell}\right)\left(-\Phi^{i-1}+2\Phi^{i}-\Phi^{i+1}\right)-(1-\alpha_2)\left(\frac{\xi_o^2\ell}{6}\right)\left(\Phi^{i-1}+4\Phi^{i}+\Phi^{i+1}\right)-\alpha_2\xi_o^2\ell\Phi^{i}=0 \tag{36}$$

$$\begin{aligned}\Rightarrow\quad &(1-\alpha_2)\left[\left(\frac{1}{\ell}\right)\left(-\Phi^{i-1}+2\Phi^{i}-\Phi^{i+1}\right)-\left(\frac{\xi_o^2\ell}{6}\right)\left(\Phi^{i-1}+4\Phi^{i}+\Phi^{i+1}\right)\right]\\ &+\alpha_2\left[\left(\frac{1}{\ell}\right)\left(-\Phi^{i-1}+2\Phi^{i}-\Phi^{i+1}\right)-\xi_o^2\ell\Phi^{i}\right]=0\end{aligned} \tag{37}$$

Equation (36) is precisely the stencil obtained from the Galerkin FEM using an alpha-interpolated mass matrix. For the 1D case using linear FE and as shown in Equation (37), it is equivalent to the alpha-interpolation of the stencils obtained by the Galerkin FEM and the classical FDM methods .

6.2. 2D bilinear FE

Consider the following definition of $\widehat{W}^a$ defined locally on the edges of the 2D bilinear FE,

$$\widehat{W}^a=\begin{cases}\left(\dfrac{1+(1+2\alpha_1)\xi\bar{\xi}^a}{2}\right)\left(\dfrac{1+\eta\bar{\eta}^a}{2}\right) & (\xi,\eta)\in(-1,1)\times\{\pm1\}\\ \left(\dfrac{1+\xi\bar{\xi}^a}{2}\right)\left(\dfrac{1+(1+2\alpha_1)\eta\bar{\eta}^a}{2}\right) & (\xi,\eta)\in\{\pm1\}\times(-1,1)\\ \left(\dfrac{1+\xi\bar{\xi}^a}{2}\right)\left(\dfrac{1+\eta\bar{\eta}^a}{2}\right) & (\xi,\eta)\in\{\pm1\}\times\{\pm1\}\end{cases} \tag{38}$$

One can arrive at the above definition by taking the Cartesian product of the 1D counterparts of W^a defined earlier in Equation (35) and then replacing the parameter α_2 therein by α_1. Clearly the functions $\widehat{W}^a$ defined via Equation (38) are a partition of unity. Likewise, on the edges whenever the expression for α_1 is single-valued, we have simultaneously a single-valued model for $\widehat{W}^a$. Thus, should any length scale appear within the expression for α_1, then it should be proportional to the corresponding edge length. Note that on the element edges wherever $N^a = 0$, we have simultaneously $\widehat{W}^a = 0$. In this way it is possible to retain the sparsity pattern of the Galerkin FEM.

Consider the discrete diffusion term $\int_\Omega \boldsymbol{\nabla} W^a \cdot \boldsymbol{\nabla} N^b \, \mathrm{d}\Omega$ calculated using a structured mesh in 2D made up of rectangular bilinear FEs. As the Laplacian of the shape functions N^b is zero in

DOI: 10.1002/nme

the interior of a rectangular bilinear FE, we use the form given in Equation (30c) to calculate the considered diffusion term. Thus,

$$\int_\Omega \boldsymbol{\nabla} W^a \cdot \boldsymbol{\nabla} N^b \,\mathrm{d}\Omega = \sum_K \int_{\partial K} \widehat{W}^a \mathbf{n} \cdot \boldsymbol{\nabla} N^b \,\mathrm{d}\Gamma \tag{39a}$$

$$\begin{aligned}\int_{\partial K} \widehat{W}^a \mathbf{n} \cdot \boldsymbol{\nabla} N^b \,\mathrm{d}\Gamma = \frac{\ell_2}{6\ell_1} &\begin{bmatrix} (2+\alpha_1) & -(2+\alpha_1) & -(1-\alpha_1) & (1-\alpha_1) \\ -(2+\alpha_1) & (2+\alpha_1) & (1-\alpha_1) & -(1-\alpha_1) \\ -(1-\alpha_1) & (1-\alpha_1) & (2+\alpha_1) & -(2+\alpha_1) \\ (1-\alpha_1) & -(1-\alpha_1) & -(2+\alpha_1) & (2+\alpha_1) \end{bmatrix} \\ + \frac{\ell_1}{6\ell_2} &\begin{bmatrix} (2+\alpha_1) & (1-\alpha_1) & -(1-\alpha_1) & -(2+\alpha_1) \\ (1-\alpha_1) & (2+\alpha_1) & -(2+\alpha_1) & -(1-\alpha_1) \\ -(1-\alpha_1) & -(2+\alpha_1) & (2+\alpha_1) & (1-\alpha_1) \\ -(2+\alpha_1) & -(1-\alpha_1) & (1-\alpha_1) & (2+\alpha_1) \end{bmatrix}\end{aligned} \tag{39b}$$

The stencil coefficient matrix $\mathbf{S}^{\mathrm{d}}$ corresponding to the assembly of the element matrices given by Equation (39b) can be expressed as follows:

$$\begin{aligned}\mathbf{S}^{\mathrm{d}} := &\frac{\ell_2}{6\ell_1}\{(1-\alpha_1),(4+2\alpha_1),(1-\alpha_1)\}^t\{-1,2,-1\} \\ &+ \frac{\ell_1}{6\ell_2}\{-1,2,-1\}^t\{(1-\alpha_1),(4+2\alpha_1),(1-\alpha_1)\}\end{aligned} \tag{40a}$$

$$\begin{aligned}\Rightarrow \mathbf{S}^{\mathrm{d}} := &(1-\alpha_1)\frac{\ell_2}{6\ell_1}\{1,4,1\}^t\{-1,2,-1\} + \alpha_1\frac{\ell_2}{6\ell_1}\{0,6,0\}^t\{-1,2,-1\} \\ &+ (1-\alpha_1)\frac{\ell_1}{6\ell_2}\{-1,2,-1\}^t\{1,4,1\} + \alpha_1\frac{\ell_1}{6\ell_2}\{-1,2,-1\}^t\{0,6,0\}\end{aligned} \tag{40b}$$

Equation (40b) is precisely the linear interpolation (specified by α_1) of the diffusion terms obtained by using the Galerkin FEM and classical FDM stencils in 2D, cf. Equation (12). Unlike in 1D where we had a unique way to model $\widehat{W}^a$ so as to retain the sparsity pattern of the Galerkin-FEM, in 2D many alternatives models exist. However, all acceptable models for $\widehat{W}^a$ have to be a partition of unity for every element and be single-valued on the element edges.

Following Equation (23) and using Equations (21) and (38) the local expression of the composite basis W^a within each element can be expressed as,

$$W^a = \begin{cases} \mathbb{W}^{ab} N^b & (\xi,\eta) \in (-1,1)\times(-1,1) \\ \left(\frac{1+(1+2\alpha_1)\xi\bar{\xi}^a}{2}\right)\left(\frac{1+\eta\bar{\eta}^a}{2}\right) & (\xi,\eta) \in (-1,1)\times\{\pm 1\} \\ \left(\frac{1+\xi\bar{\xi}^a}{2}\right)\left(\frac{1+(1+2\alpha_1)\eta\bar{\eta}^a}{2}\right) & (\xi,\eta) \in \{\pm 1\}\times(-1,1) \\ \left(\frac{1+\xi\bar{\xi}^a}{2}\right)\left(\frac{1+\eta\bar{\eta}^a}{2}\right) & (\xi,\eta) \in \{\pm 1\}\times\{\pm 1\} \end{cases} \tag{41}$$

where $\mathbb{W}$ is the matrix of constant coefficients given by Equation (20). Recall that the matrix $\mathbb{W}$ involves the parameter α_2 in its definition. Note that on structured 2D meshes using the basis W^a given by Equation (41), the PG method defined in Equation (26) will yield the nonstandard compact stencil summarized in Equation (12). This observation follows by using the results given in Equations (25), (39) and (40) in the discrete system matrix given in Equation (32).

6.3. Stabilization parameters on unstructured meshes

As most of the expressions for α_1, α_2 optimized for square meshes need not be optimal for unstructured meshes, in the current work we consider only the simplest expressions that would guarantee fourth-order (cf. Equation (42a)) and sixth-order (cf. Equation (42b)) dispersion accuracy

DOI: 10.1002/nme

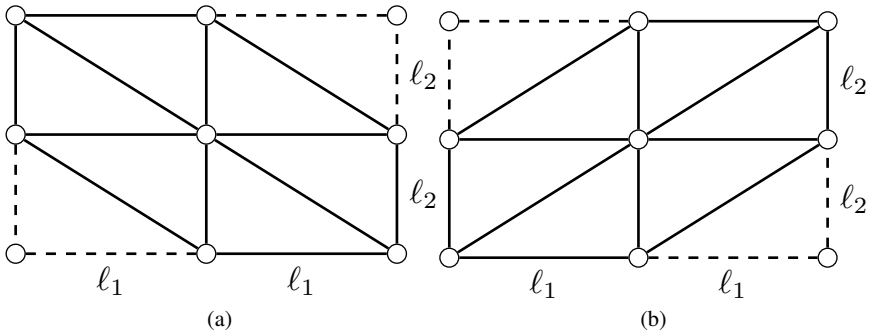


Figure 5. Stencils obtained by using a structured simplicial finite element mesh with the hypotenuse oriented/tilted along (a) left, i. e. $o = l$ and (b) right , i. e. $o = r$. The flag 'o' indicates the stencil tilt.

on square meshes. On unstructured meshes the expressions for α_1, α_2 corresponding to these two choices can be written as follows:

$$\alpha_1 = \alpha_2 = \frac{1}{2} \tag{42a}$$

$$\alpha_1 = \frac{1}{2} - \frac{\widehat{\omega}}{60} \quad ; \quad \alpha_2 = \frac{1}{2} - \frac{\widetilde{\omega}}{40} \tag{42b}$$

where $\widehat{\omega} := (\xi_o \widehat{\ell})^2$ and $\widetilde{\omega} := (\xi_o \widetilde{\ell})^2$. $\widehat{\ell}$ and $\widetilde{\ell}$ represent the models used for the length measures corresponding to the element edges and the interior, respectively. In the current study for each element we have chosen $\widehat{\ell}$ equal to the edge length (will vary from edge to edge) and $\widetilde{\ell}$ equal to the maximum edge length. Note that using this model, α_1 is always single-valued on the edges. On square meshes using Equation (42b) we recover α_1, α_2 as given in Equation (13) up to the first two terms which is sufficient to attain sixth-order dispersion accuracy.

7. SIMPLICIAL FINITE ELEMENTS

Consider a rectangular domain discretized by structured simplicial FEs. Such discretization would typically yield stencils as shown in Figure 5. The stencils with the hypotenuse oriented along left and right are labeled using the markers $o = l$ and $o = r$ respectively. The flag 'o' indicates the stencil tilt.

The equation stencil for the Galerkin FEM corresponding to any interior node (i, j) can be written as Equation (7) with the following definition of stencil coefficient matrix:

$$\mathbf{S}^{fem} = \frac{\ell_2}{\ell_1}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{\ell_1}{\ell_2}\begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix} - \frac{\xi_o^2 \ell_1 \ell_2}{12}\begin{bmatrix} \delta_{ol} & 1 & \delta_{or} \\ 1 & 6 & 1 \\ \delta_{or} & 1 & \delta_{ol} \end{bmatrix} \tag{43}$$

where δ_{ol} and δ_{or} are Kronecker deltas. Note that using simplicial FEs the contribution of the diffusion term in Equation (43) is identical to that obtained in the FDM stencil given by Equation (11). Thus, the stencil obtained via an α-interpolation of the Galerkin FEM and the FDM stencils will lead to the following stencil coefficient matrix:

$$\mathbf{S}^{\alpha} = \frac{\ell_2}{\ell_1}\begin{bmatrix} 0 & 0 & 0 \\ -1 & 2 & -1 \\ 0 & 0 & 0 \end{bmatrix} + \frac{\ell_1}{\ell_2}\begin{bmatrix} 0 & -1 & 0 \\ 0 & 2 & 0 \\ 0 & -1 & 0 \end{bmatrix} - \frac{\xi_o^2 \ell_1 \ell_2}{12}\begin{bmatrix} (1-\alpha)\delta_{ol} & (1-\alpha) & (1-\alpha)\delta_{or} \\ (1-\alpha) & 6(1+\alpha) & (1-\alpha) \\ (1-\alpha)\delta_{or} & (1-\alpha) & (1-\alpha)\delta_{ol} \end{bmatrix} \tag{44}$$

We see that using simplicial FEs in 2D, the α-interpolation of Galerkin FEM and FDM is equivalent to the alpha-interpolation method (AIM) [4, 5]. In the AIM, the consistent mass matrix

DOI: 10.1002/nme

$\mathbf{M}$ that appears in the Galerkin FEM is replaced by the α-interpolated mass matrix $\mathbf{M}^\alpha := (1-\alpha)\mathbf{M} + \alpha\mathbf{M}_\mathrm{L}$. Consider the following definition for the composite basis W^a when using simplicial FEs,

$$W^a = \begin{cases} \widetilde{W}^a := \mathbb{W}^{ab} N^b & \text{in the element interior} \\ \widehat{W}^a := N^a & \text{on the element edges} \end{cases} \tag{45}$$

Consider the BVP (1) posed on an interior 2D domain subjected to Dirichlet boundary conditions and let $f(\mathbf{x}) = 0$. Using the basis W^a given by Equation (45) in the PG method defined in Equation (26), we recover the AIM. In particular for the structured simplicial FE meshes shown in Figure 5, we recover the stencil given in Equation (44). We can guess that a solution to any generic stencil takes the form $\Phi^{i,j} := \phi(x_1^i, x_2^j) = \exp[\mathrm{i}(\xi_1^h x_1^i + \xi_2^h x_2^j)]$. Substituting this solution into the stencil formed by $\mathbf{S}^\alpha$ given in Equation (44) and defining $\lambda_1 := \exp(\mathrm{i}\xi_1^h \ell_1)$ and $\lambda_2 := \exp(\mathrm{i}\xi_2^h \ell_2)$ we get the characteristic equation as follows:

$$\begin{aligned} \frac{[2-\lambda_1-\lambda_1^{-1}]}{\omega_1} + \frac{[2-\lambda_2-\lambda_2^{-1}]}{\omega_2} &= \frac{(1+\alpha)}{2} + \frac{(1-\alpha)}{12}\Big(\lambda_1+\lambda_1^{-1}+\lambda_2+\lambda_2^{-1}\Big) \\ &+ \frac{(1-\alpha)}{12}\Big(\delta_{ol}[\lambda_1\lambda_2^{-1}+\lambda_1^{-1}\lambda_2] + \delta_{or}[\lambda_1\lambda_2+\lambda_1^{-1}\lambda_2^{-1}]\Big) \end{aligned} \tag{46}$$

where, $\omega_1 := (\xi_o \ell_1)^2$ and $\omega_2 := (\xi_o \ell_2)^2$. For the dispersion analysis of $\mathbf{S}^\alpha$ given in Equation (44) we restrict to the case $\ell_1 = \ell_2 = \ell$. In this case, the stencil coefficient matrix $\mathbf{S}^\alpha$ simplifies to,

$$\mathbf{S}^\alpha = \begin{bmatrix} \delta_{ol}S_2 & S_1 & \delta_{or}S_2 \\ S_1 & S_0 & S_1 \\ \delta_{or}S_2 & S_1 & \delta_{ol}S_2 \end{bmatrix} \quad ; \quad \begin{aligned} S_0 &:= 4-(1+\alpha)(\omega/2) \\ S_1 &:= -1-(1-\alpha)(\omega/12) \\ S_2 &:= -(1-\alpha)(\omega/12) \end{aligned} \tag{47}$$

where, $\omega := (\xi_o \ell)^2$. The characteristic equation given in Equation (46) now gets simplified to the following:

$$S_0 + 2S_1[\cos(\xi_1^h \ell) + \cos(\xi_2^h \ell)] + 2S_2 \cos(\xi_1^h \ell \pm \xi_2^h \ell) = 0 \tag{48}$$

The '$\pm$' that appears in the above equation corresponds to the cases $o = r$ and $o = l$ respectively (see Figure 5). The parameter α may be expressed as a generic series expansion in terms of ω as follows:

$$\alpha := \sum_{m=0}^{\infty} a_m \omega^m \approx a_0 + a_1\omega + a_2\omega^2 + a_3\omega^3 + O\left(\omega^4\right) \tag{49}$$

where a_m, b_m are coefficients independent of ω. Following the approach used in [1] which was originally presented in [11], the relative phase error ($\mathbb{P}$) and local truncation error ($\mathbb{T}$) along any direction β can be written as,

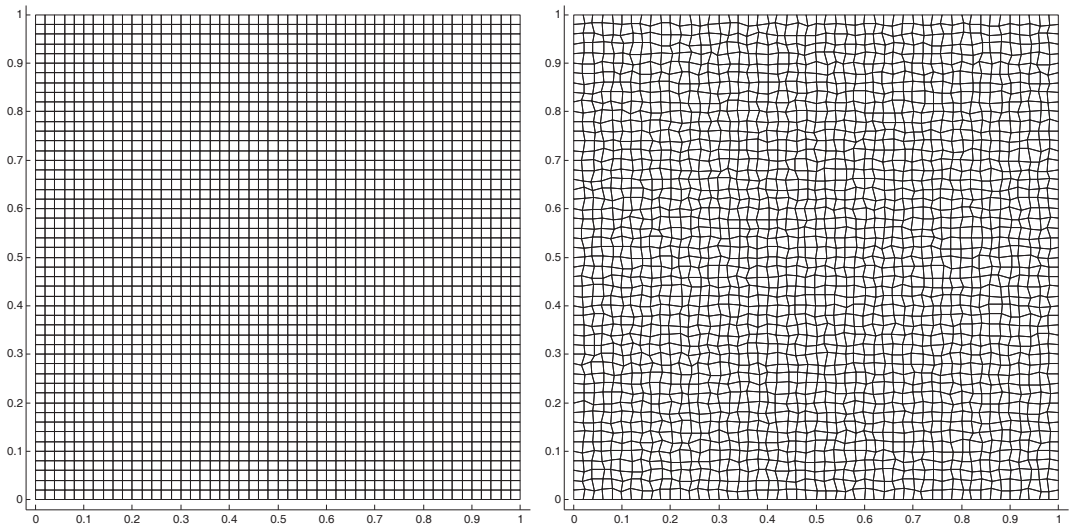
$$\mathbb{P} = r_1\omega + O\left(\omega^2\right) \quad ; \quad \mathbb{T} = -2r_1\omega + O\left(\omega^2\right) \tag{50}$$

$$r_1 := \frac{(a_0-1)}{24}[2 \pm \sin(2\beta)] + \left[\frac{3+\cos(4\beta)}{96}\right] \tag{51}$$

Clearly it is impossible to obtain the condition $r_1 = 0$ by a choice of the coefficient a_0 that is independent of the angle β. Thus, unlike for the structured bilinear block FEs, for the structured simplicial FEs shown in Figure 5, the pollution is essentially of the same order as for those of the Galerkin FEM, the FDM and the GLS-FEM [17, 18]. Nevertheless, just like for the GLS-FEM, the coefficient a_0 can be chosen so as to arrive at a higher-order modification of the interior stencil of the Galerkin FEM. Similar studies for eigenvalue problems using the AIM with simplicial FEs was done in [4, 5, 19, 20].

Remark: Following the approach taken for bilinear block FEs, it is possible to provide different models for the PG weights on the elements edges. This idea will be explored in future works.

DOI: 10.1002/nme



(a) (b)

Figure 6. Meshes made of bilinear block-FEs. (a) Uniform mesh, $\delta = 0$. (b) Nonuniform mesh, $\delta = 0.2$.

8. EXAMPLES

In this section we present some examples in 2D for the problem defined by Equation (1) and considering the following problem data: the wavenumber $\xi_o \in \{50, 100\}$, the source $f = 0$, the direction of wave propagation $\beta = (\pi/9)$ and the domain $\Omega = [0, 1] \times [0, 1]$. The domain Ω is discretized by considering both uniform and nonuniform meshes made up of just the bilinear block FEs. The nonuniform meshes are obtained by randomly perturbing the interior nodes of uniform meshes with coordinates (x_i, y_i) as follows [21, 22]:

$$x_i^{'} = x_i + \ell_1 \delta \,\mathrm{rand}() \quad ; \quad y_i^{'} = y_i + \ell_2 \delta \,\mathrm{rand}() \tag{52}$$

where, $(x_i^{'}, y_i^{'})$ represent the corresponding coordinates of the uniform mesh, δ is a mesh distortion parameter and $\mathrm{rand}()$ is a function that returns random numbers uniformly distributed in the interval $[-1, 1]$. Figure 6 illustrates an instance of an unstructured mesh obtained by this procedure using a 50×50 square mesh and the parameter $\delta = 0.2$.

We consider the following four cases concerned with the choice of the stabilization parameters α_1, α_2:

I: $\alpha_1 = \alpha_2 = 0$. This case corresponds to the Galerkin FEM.

II: $\alpha_1 = \alpha_2 = 1$. This case on rectangular meshes corresponds to the classical FDM. We denote this case as FDM/PG as it is obtained within a Petrov–Galerkin framework. FDM/PG is a straight-forward extension of the FDM to unstructured meshes.

III: $\alpha_1 = \alpha_2 = (1/2)$. This case corresponds to a discrete system that is equivalent to the average of the Galerkin FEM and the FDM/PG. On rectangular meshes we obtain the stencil associated with $(\mathbf{S}^{fem} + \mathbf{S}^{fdm})/2$, which is equivalent to the one obtained using the generalized Padé approximation in 2D [6, 7]. The dispersion accuracy on square meshes is of fourth-order.

IV: $\alpha_1 \neq \alpha_2 \neq 0$ and given by Equation (42b). On rectangular meshes this case yields the nonstandard compact stencil presented in [1] and summarized in Equation (12). Recall that on square meshes these expressions for the parameters α_1, α_2, guarantee sixth-order dispersion accuracy.

DOI: 10.1002/nme

For these considerations we study the convergence of the relative error in the following norms:

$$L^2 \text{ norm} \qquad \frac{\|\phi-\phi_h\|_0}{\|\phi\|_0} := \frac{[\int_\Omega |\phi-\phi_h|^2 \,\mathrm{d}\Omega]^{1/2}}{[\int_\Omega |\phi|^2 \,\mathrm{d}\Omega]^{1/2}} \tag{53a}$$

$$H^1 \text{ semi-norm} \qquad \frac{\|\phi-\phi_h\|_1}{\|\phi\|_1} := \frac{[\int_\Omega |\boldsymbol{\nabla}(\phi-\phi_h)|^2 \,\mathrm{d}\Omega]^{1/2}}{[\int_\Omega |\boldsymbol{\nabla}\phi|^2 \,\mathrm{d}\Omega]^{1/2}} \tag{53b}$$

$$l^\infty \text{ Euclidean norm} \qquad \frac{|\Phi_e-\Phi_h|_\infty}{|\Phi_e|_\infty} := \frac{\max_i |\Phi_e^i-\Phi_h^i|}{\max_i |\Phi_e^i|} \tag{53c}$$

where Φ_e is the exact solution sampled at the nodes of the mesh. In the convergence studies done here, the numerical solutions corresponding to the four cases viz. I-IV, are compared with the following solutions: the nodally exact FE interpolant denoted by $I_h\phi$ and the best approximations with respect to the L^2 norm and the H^1 semi-norm denoted by $P_h^0\phi$ and $P_h^1\phi$ respectively. The solutions $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$ can be found as shown in Equation (54).

$$I_h\phi := N^a\Phi_e^a \tag{54a}$$

$$\int_\Omega w_h(\phi-P_h^0\phi)\,\mathrm{d}\Omega = 0 \;\forall\; w_h \in U_0^h \;\Rightarrow\; \|\,\phi-P_h^0\phi\,\|_0 \leq \|\,\phi-\phi_h\,\|_0 \;\forall\; \phi_h \in U_{\mathrm{E}}^h \tag{54b}$$

$$\int_\Omega \boldsymbol{\nabla}w_h\cdot\boldsymbol{\nabla}(\phi-P_h^1\phi)\,\mathrm{d}\Omega = 0 \;\forall\; w_h \in U_0^h \;\Rightarrow\; \|\,\phi-P_h^1\phi\,\|_1 \leq \|\,\phi-\phi_h\,\|_1 \;\forall\; \phi_h \in U_{\mathrm{E}}^h \tag{54c}$$

As the exact solution ϕ is sinusoidal, we have used a third-order Gauss quadrature rule to evaluate the expressions involving ϕ in Equation (53) and Equation (54).

8.1. Example 1: Dirichlet boundary conditions

In this example, only the Dirichlet boundary conditions are prescribed such that the exact solution is $\phi(\mathbf{x}) = \sin(\boldsymbol{\xi}^\beta\cdot\mathbf{x})$, where $\boldsymbol{\xi}^\beta := \xi_o(\cos(\beta),\sin(\beta))$. Uniform meshes with $n\times n$ square elements are considered with n given by the following expression.

$$n = \mathrm{ceil}(50\times 2^{m/8}) \quad ; \quad m\in\{0,1,2,\ldots 28\} \tag{55}$$

where $\mathrm{ceil}(m)$ is a function that returns the nearest integer greater than or equal to m. Nonuniform meshes are obtained corresponding to each uniform mesh using the procedure described earlier. For these considerations we present the plots of the relative error vs. the mesh-size.

Figure 7 illustrates the convergence of the relative error in the L^2 norm. Clearly the error lines of the considered solutions are bounded from below by the error line of $P_h^0\phi$ (L^2-BA) and show a tendency to become parallel to the error line of $P_h^0\phi$ as $\ell\to 0$. Figures 7a and 7b show the L^2 error considering $\xi_o = 50$ and for uniform ($\delta = 0$) and nonuniform ($\delta = 0.2$) meshes respectively. As expected the error lines corresponding to cases I and II differs substantially from those of $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$. The error lines corresponding to cases III and IV are very close to that of $I_h\phi$. As the solution in case IV has sixth-order dispersion accuracy on square meshes it is almost the same as $I_h\phi$. On nonuniform meshes the quality of the solution in case IV deteriorates and is similar to that of case III. Figures 7c and 7d show the error lines considering $\xi_o = 100$ and for the choices $\delta = 0$ and $\delta = 0.2$ respectively. As expected all the error lines corresponding to cases I-IV deviate further from the error lines of $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$ (the pollution effect). On square meshes the solution of case IV shows the least deviation and is practically identical to $I_h\phi$ (Figure 7c). The pollution associated with the solution of case III is similar to that of cases I and II on coarse meshes but it diminishes rapidly on further mesh refinement. Again, on nonuniform meshes the quality of the solution in case IV deteriorates showing an appreciable deviation from the error lines of $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$ and is similar to that of case III (Figure 7d). A distinctive feature in these plots is the formation of spikes in the error lines. Their presence is more evident for higher wavenumbers and on nonuniform meshes where the dispersion errors are relatively higher. As here we have prescribed only the Dirichlet boundary conditions the numerical solutions might suffer spurious amplitude and/or phase

DOI: 10.1002/nme

modulations to satisfy them [17]. Encounters with zones of degeneracy (wherein the solution might blow up) also contributes to huge errors in the amplitude [1, 17, 23]. Fortunately, these spurious modulations reduce should other choices for the boundary conditions be employed viz. an exterior problem with DtN boundary conditions [17], an interior problem with Robin boundary conditions [11].

Figure 8 illustrates the convergence of the relative error in the H^1 semi-norm. Clearly the error lines of the considered solutions are bounded from below by the error line of $P_h^1\phi$ (H^1-BA). Unlike the errors measured in the L^2 norm, the errors measured in the H^1 semi-norm show a tendency to merge with the error line of $P_h^1\phi$. Figures 8a and 8b ($\xi_o = 50$) show that the error lines of case III and IV are practically the same as of $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$. Figures 8c and 8d ($\xi_o = 100$) show that the deviations of the error lines of cases III and IV from the error line of $P_h^1\phi$ even though they exist, it is smaller than that observed using the L^2 norm.

Figure 9 illustrates the convergence of the relative error in the l^∞ Euclidean norm which is a measure of nodal exactness. Figures 9a and 9c show that on unifrom meshes ($\delta = 0$) the error lines of case III and IV converge at a rate of fourth and sixth order respectively. Figures 9b and 9d show that on nonuniform meshes ($\delta = 0.2$) the higher order accuracy of case IV deteriorates and has a trend similar to that of case III. Also, in an average sense both the cases III and IV have second-order convergence rate similar to $P_h^0\phi$ and $P_h^1\phi$. For the wavenumber $\xi_o = 50$ the errors found for the cases III and IV are similar to that of $P_h^0\phi$ (Figure 9b).

8.2. Example 2: Robin boundary conditions

In this example, only the Robin boundary conditions are prescribed such that the exact solution is $\phi(\mathbf{x}) = \exp(\mathrm{i}\boldsymbol{\xi}^\beta \cdot \mathbf{x})$, where $\boldsymbol{\xi}^\beta := \xi_o(\cos(\beta), \sin(\beta))$. The operator $\mathcal{M}$ that appears in Equation (1c) is chosen as $\mathcal{M} := \mathrm{i}\xi_o$. Thus, $q(\mathbf{x}) := \mathrm{i}(\mathbf{n} \cdot \boldsymbol{\xi}^\beta - \xi_o)\exp(\mathrm{i}\boldsymbol{\xi}^\beta \cdot \mathbf{x})$. Uniform meshes with $n \times n$ square elements are considered with n given by the following expression.

$$n = \mathrm{ceil}(\frac{m\xi_o}{2\pi}) \quad ; \quad m \in \{10, 10.5, 11, 11.5, \ldots 25\} \tag{56}$$

Choosing n by the above expression guarantees the presence of at least m elements per wavelength. Nonuniform meshes are obtained corresponding to each uniform mesh using the procedure described earlier. For these considerations, we present the plots of the relative error vs. ξ^*, where $\xi^* := (\xi_o\ell/\pi)$. The choice of ξ^* as the abscissa in the plots allows us to single out the pollution effect.

Figures 10, 11 and 12 illustrate the convergence of the relative error in the L^2 norm, the H^1 semi-norm and the l^∞ Euclidean norm respectively. Clearly, all the spurious modulations that appeared in the error lines considering only the Dirichlet boundary conditions (Figures 7, 8 and 9) diminish when the Robin boundary conditions are prescribed. Also, in all the Figures (10, 11 and 12) by freezing the value of δ and increasing the value of ξ_o we observe the following trait. The location of the error lines of $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$ is practically unaffected by an increase in ξ_o (no pollution). As expected the error lines of cases I and II not only are located high above the error lines of $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$ but also shift higher with an increase in ξ_o (pollution effect).

On uniform meshes ($\delta = 0$) the error lines of cases III and IV not only are located close to the respective best approximations but also show negligible upward shift with an increase in ξ_o (small pollution). Clearly, on uniform meshes the performance of case IV is relatively better than that of case III (although the difference is small). The pollution effect is more visible for these cases on nonuniform meshes ($\delta = 0.2$). In the L^2 norm the error lines of cases III and IV show an accuracy at par with $I_h\phi$ and $P_h^1\phi$ (Figures 10c and 10d). In the H^1 semi-norm the error lines of cases III and IV are practically the same as those corresponding to $I_h\phi$, $P_h^0\phi$ and $P_h^1\phi$ (Figures 11c and 11d). In the l^∞ Euclidean norm the error lines of cases III and IV are close to the error line of $P_h^0\phi$ (Figures 12c and 12d). Further, in Figure 12 note that in an average sense all the error lines have second-order convergence rate in the l^∞ Euclidean norm. This result is due to the error in the approximation of the Robin boundary condition. Thus, unlike in Figure 9 wherein the error lines of cases III and IV showed fourth-order and sixth-order convergence rates respectively, here it drops to second-order.

DOI: 10.1002/nme

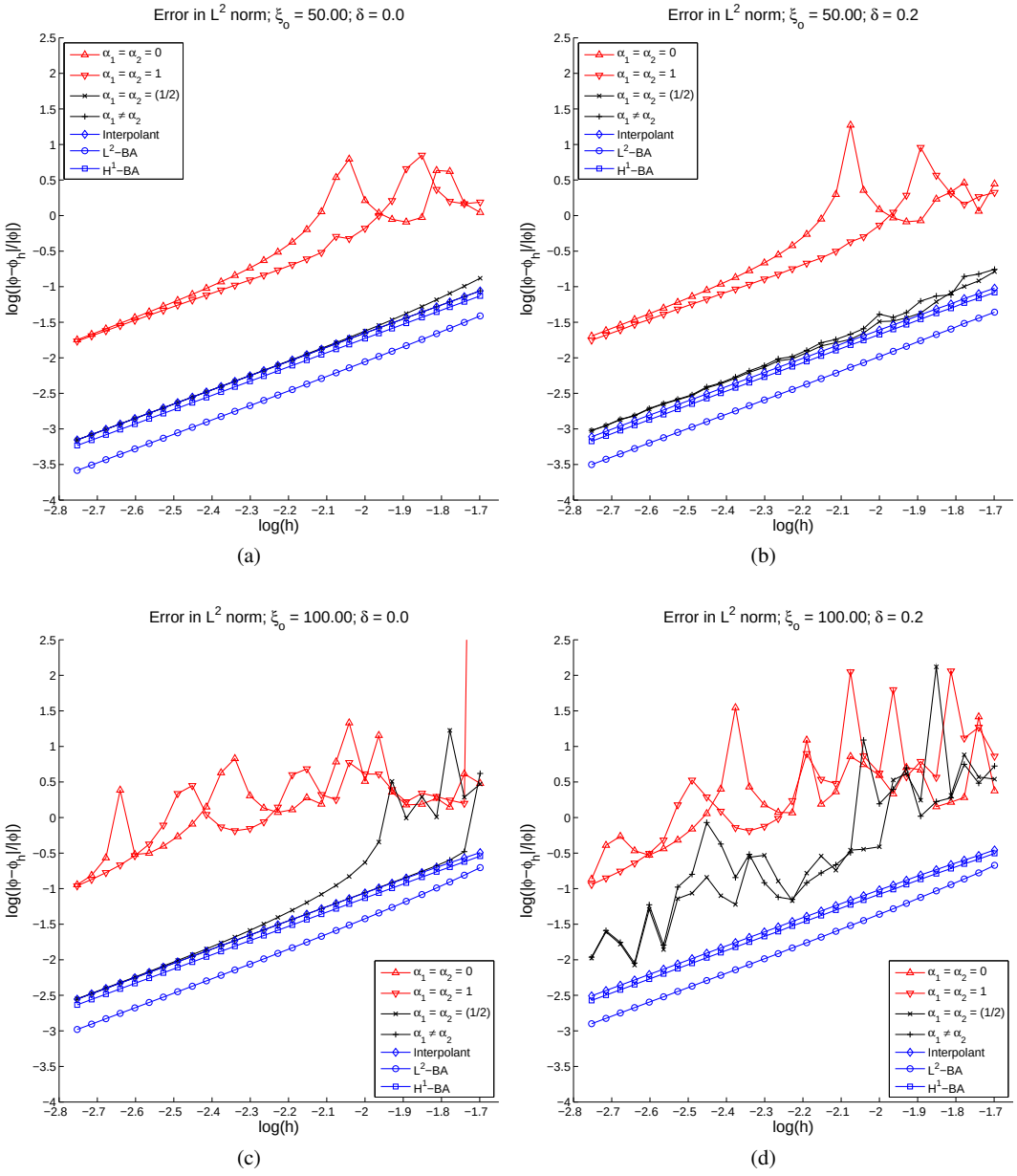


Figure 7. Convergence of the relative error in the L^2 norm using $\beta = (\pi/9)$ and Dirichlet boundary conditions. The wavenumber ξ_o and the mesh distortion parameter used are: (a) $\xi_o = 50$, $\delta = 0$; (b) $\xi_o = 50$, $\delta = 0.2$; (c) $\xi_o = 100$, $\delta = 0$ and (d) $\xi_o = 100$, $\delta = 0.2$.

9. CONCLUSIONS

A new Petrov–Galerkin (PG) method involving two parameters viz. α_1, α_2 is presented which yields the following schemes on rectangular meshes: a) a compact stencil obtained by the α-interpolation of the Galerkin FEM and the classical central FDM, should the two parameters be made equal, i. e. $\alpha_1 = \alpha_2 = \alpha$ and b) the nonstandard compact stencil presented in [1] for the Helmholtz equation if the parameters are distinct, i. e. $\alpha_1 \neq \alpha_2$. On square meshes, these two schemes were shown to provide solutions to the Helmholtz equation that have a dispersion accuracy of fourth and sixth

DOI: 10.1002/nme

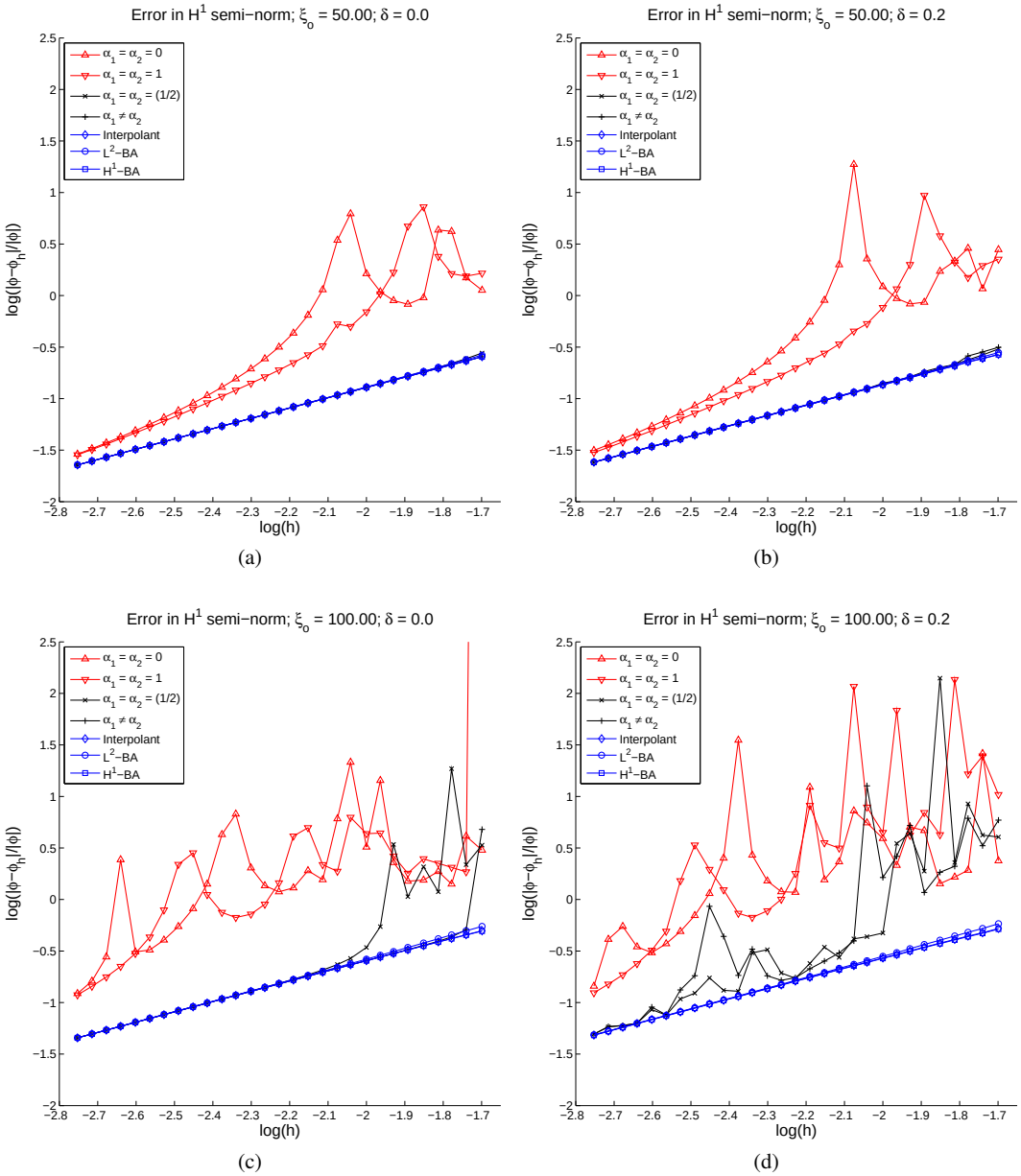


Figure 8. Convergence of the relative error in the H^1 semi-norm using $\beta = (\pi/9)$ and Dirichlet boundary conditions. The wavenumber ξ_o and the mesh distortion parameter used are: (a) $\xi_o = 50$, $\delta = 0$; (b) $\xi_o = 50$, $\delta = 0.2$; (c) $\xi_o = 100$, $\delta = 0$ and (d) $\xi_o = 100$, $\delta = 0.2$.

order respectively [1]. Thus, this Petrov–Galerkin method yields in a straight-forward manner the counterparts of these two schemes on unstructured meshes.

The salient features of this new PG method include the following. The solution space is discretized by standard C^0-continuous finite elements. The test functions/weights are piecewise polynomials of the same degree as the FE shape functions and are generally discontinuous at the inter-element boundaries. Models for the weights on the inter-element boundaries are provided such that the sparsity pattern is the same as that for the Galerkin FEM. The parameters α_1, α_2 determine the shape of the weights on the element edges and the interiors, respectively. The choice $\alpha_1 = \alpha_2 = 0$ yield weights that are identical to the FE shape functions and hence we recover the

DOI: 10.1002/nme

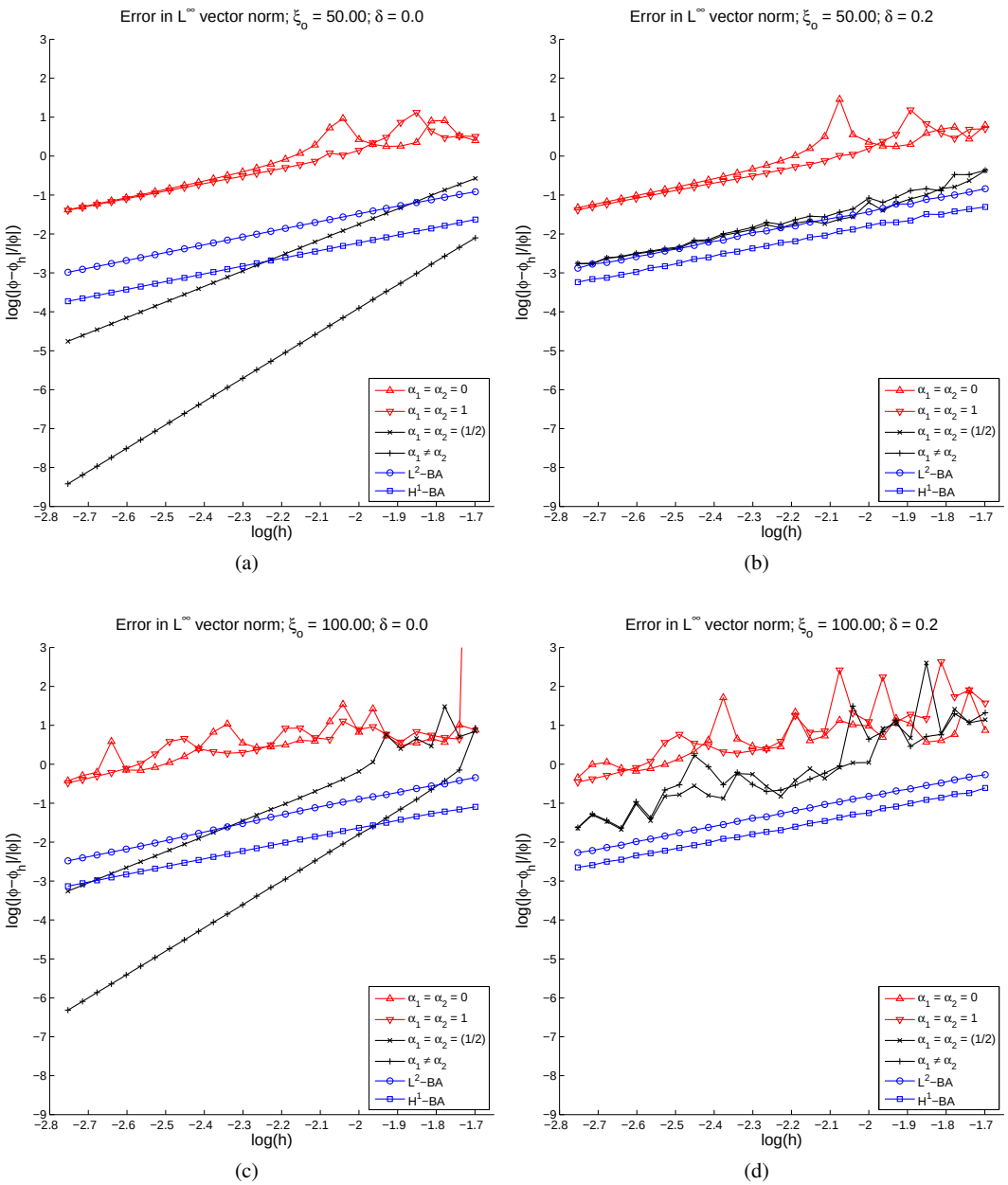


Figure 9. Convergence of the relative error in the l^∞ Euclidean norm using $\beta = (\pi/9)$ and Dirichlet boundary conditions. The wavenumber ξ_o and the mesh distortion parameter used are: (a) $\xi_o = 50$, $\delta = 0$; (b) $\xi_o = 50$, $\delta = 0.2$; (c) $\xi_o = 100$, $\delta = 0$ and (d) $\xi_o = 100$, $\delta = 0.2$.

Galekin FEM. The weights are a partition of unity only in the sense that they add up to unity. As the row lumping technique for the FEM mass matrices is a critical step in the design of these weights (to fulfill the partition of unity constraint), the current PG method is restricted only to those FEs where this technique makes sense, i. e. linear interpolation on simplices and multilinear interpolation on blocks.

The α-interpolation of FEM and FDM on a rectangular domain discretized by structured simplicial FE mesh would yield a scheme identical to the alpha-interpolation method (AIM) [4, 5] wherein the mass matrix that appears in the Galerkin FEM is replaced by an α-interpolated mass

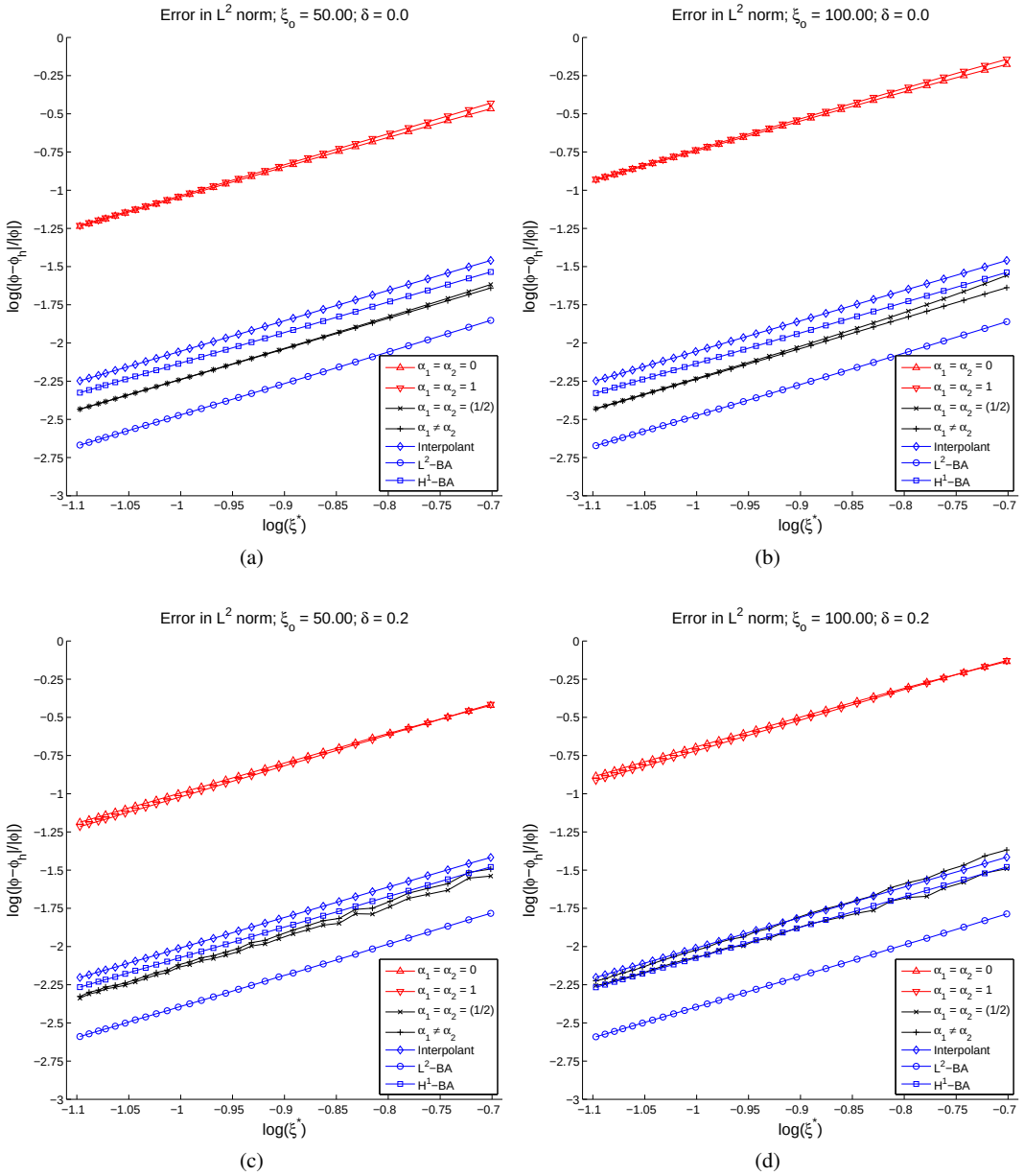


Figure 10. Convergence of the relative error in the L^2 norm using $\beta = (\pi/9)$ and Robin boundary conditions. The wavenumber ξ_o and the mesh distortion parameter used are: (a) $\xi_o = 50$, $\delta = 0$; (b) $\xi_o = 100$, $\delta = 0$; (c) $\xi_o = 50$, $\delta = 0.2$ and (d) $\xi_o = 100$, $\delta = 0.2$.

matrix. In the current PG method we recover the AIM (even on unstructured simplicial meshes) by making the choice $\widehat{W}^a = N^a|_{\mathcal{E}_h}$. Unfortunately in this case the dispersion accuracy drops to second-order.

Recall that on square meshes many existing higher-order compact schemes (including the QSFEM [11]) can be recovered by an appropriate choice of the parameters α_1, α_2 [1]. As most of the expressions for α_1, α_2 optimized for square meshes need not be optimal for unstructured meshes in the presented examples we have considered only the simplest expressions that would guarantee fourth-order (choosing $\alpha_1 = \alpha_2 = (1/2)$) and sixth-order ($\alpha_1, \alpha_2$ given by Equation (42b)) dispersion accuracy on square meshes. Convergence studies of the solution error

DOI: 10.1002/nme

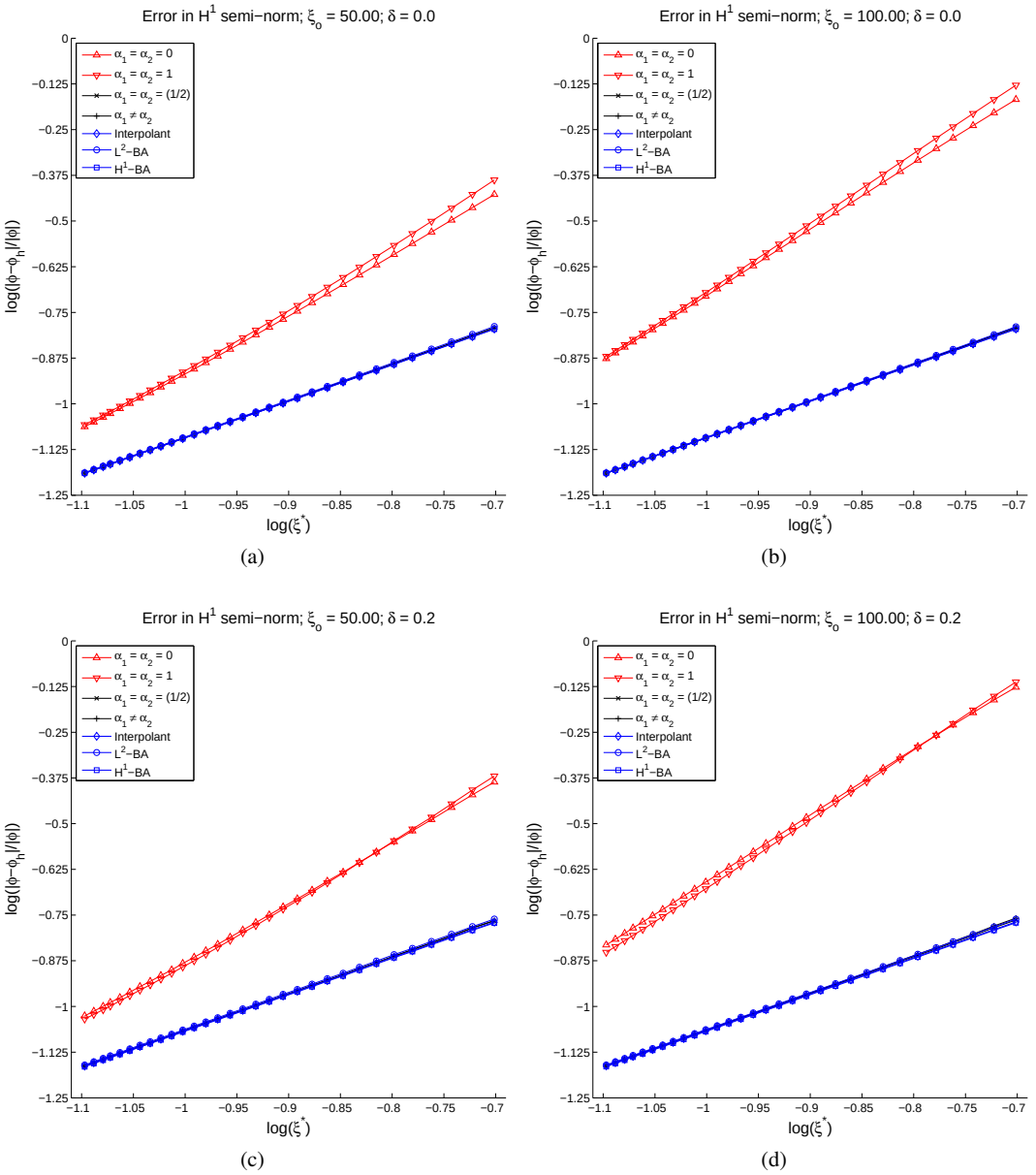


Figure 11. Convergence of the relative error in the H^1 semi-norm using $\beta = (\pi/9)$ and Robin boundary conditions. The wavenumber ξ_o and the mesh distortion parameter used are: (a) $\xi_o = 50,\ \delta = 0$; (b) $\xi_o = 100, \delta = 0$; (c) $\xi_o = 50, \delta = 0.2$ and (d) $\xi_o = 100, \delta = 0.2$.

corresponding to these two choices are done to quantify the pollution effect and comparisons are made with respect to the errors of the Galekin FEM, the nodally exact FE interpolant and the best approximations in the L^2 norm and the H^1 semi-norm respectively. Both the Dirichlet and Robin boundary conditions were considered in the examples. The wavenumbers $\xi_o = 50$ and $\xi_o = 100$ were chosen to represent values in the mid-frequency and high-frequency range respectively.

For the Dirichlet problem, the results on square meshes verify the higher-order dispersion accuracy and the low pollution effect. However on nonuniform meshes the dispersion accuracy of the current PG method drops down to second-order (verified by the errors in the l^∞ Euclidean norm). Also, the performance of both the choices for the parameters α_1, α_2 is similar on nonuniform

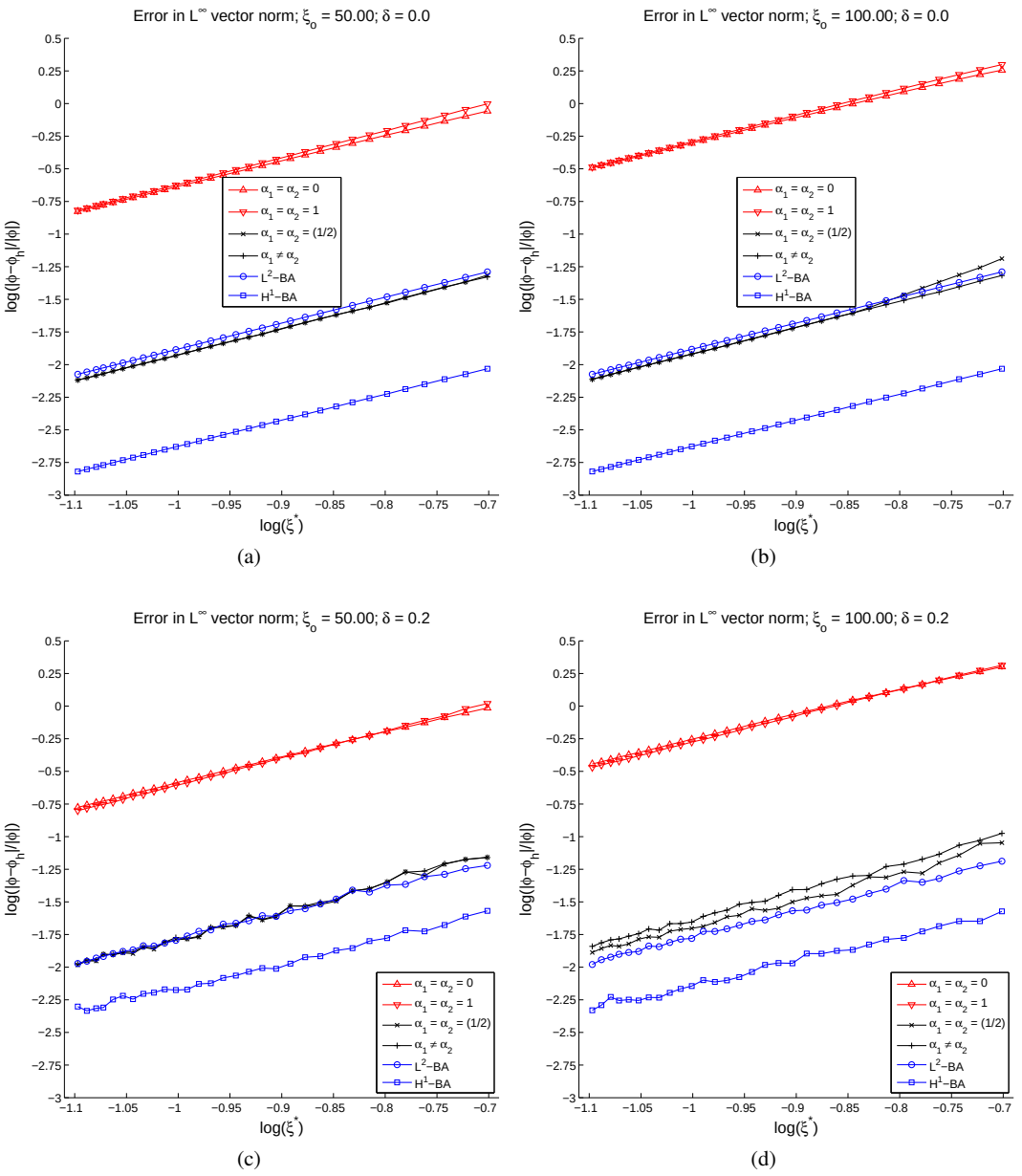


Figure 12. Convergence of the relative error in the l^∞ Euclidean norm using $\beta = (\pi/9)$ and Robin boundary conditions. The wavenumber ξ_o and the mesh distortion parameter used are: (a) $\xi_o = 50,\ \delta = 0$; (b) $\xi_o = 100,\ \delta = 0$; (c) $\xi_o = 50,\ \delta = 0.2$ and (d) $\xi_o = 100,\ \delta = 0.2$.

meshes. For the mid-frequency range, i. e. $\xi_o = 50$ the errors in the l^∞ Euclidean norm for both the parameter choices is close to the error of the best approximation in the L^2 norm. In the high-frequency range, i. e. $\xi_o = 100$, the improvement with respect to the Galerkin FEM is significant. However, the solutions exhibit spurious modulations indicating that there is still room for improvement.

For the Robin problem, these spurious modulations in the solutions cease to exist. The pollution effect on square meshes is greatly reduced and on nonuniform meshes it is small. Also, the location of the error lines of the current PG method is in between the error lines of pollution-free solutions,

DOI: 10.1002/nme

viz. the nodally exact FE interpolant and the best approximations in the L^2 norm and the H^1 semi-norm, thus indicating high accuracy.

The additional cost of implementation of the current PG method is just the evaluation of the element boundary integrals, cf. Equation (31). All the algebraic evaluations are done at the element level unlike the QOPG method [21] where it is done at the patch level. This feature allows the current PG method to be easily incorporated within an 'assemble-by-elements' data structure. The choice of the parameters $\alpha_1 = \alpha_2 = (1/2)$ render the current PG method independent of the problem and mesh data. In this sense and for this choice, the current PG method could be labeled 'parameter-free'.

ACKNOWLEDGEMENTS

The first author acknowledges the economic support received through the FI pre-doctoral grant from the *Department of Universities, Research and Information Society* (Generalitat de Catalunya) and the *European Social Fund*. He also thanks Profs. Ramon Codina, Assad Oberai and Carlos Felippa for many useful discussions.

REFERENCES

1. Nadukandi P, Oñate E, Garcia J. A fourth-order compact scheme for the Helmholtz equation: Alpha-interpolation of FEM and FDM stencils. *International Journal for Numerical Methods in Engineering* 2011; **86**(1):18–46, doi:10.1002/nme.3043. URL http://doi.wiley.com/10.1002/nme.3043.
2. Güttinger W. Generalized Functions and Dispersion Relations in Physics. *Fortschritte der Physik* 1966; **14**(1-12):483–602, doi:10.1002/prop.19660140114. URL http://doi.wiley.com/10.1002/prop.19660140114.
3. Farassat F. Introduction to Generalized Functions with Applications in Aerodynamics and Aeroacoustics. *Technical Report NASA-TP-3428*, NASA Langley Research Center, Hampton, VA 23681-0001, USA May 1994. URL http://hdl.handle.net/2060/19940029887.
4. Ishihara K. Convergence of the Finite Element Method Applied to the Eigenvalue Problem $\Delta u + \lambda u = 0$. *Publications of the Research Institute for Mathematical Sciences* 1977; **13**(1):47–60. URL http://ci.nii.ac.jp/naid/110004714382/en/.
5. Niki H, Sawami H, Ikeuchi M, Okamoto N. The alpha interpolation method for the solution of an eigenvalue problem. *Journal of Computational and Applied Mathematics* Mar 1982; **8**(1):15–19, doi:10.1016/0771-050X(82)90002-X. URL http://linkinghub.elsevier.com/retrieve/pii/0771050X8290002X.
6. Harari I, Turkel E. Accurate Finite Difference Methods for Time-Harmonic Wave Propagation. *Journal of Computational Physics* Jul 1995; **119**(2):252–270, doi:10.1006/jcph.1995.1134. URL http://linkinghub.elsevier.com/retrieve/doi/10.1006/jcph.1995.1134.
7. Singer I, Turkel E. High-order finite difference methods for the Helmholtz equation. *Computer Methods in Applied Mechanics and Engineering* Sep 1998; **163**(1-4):343–358, doi:10.1016/S0045-7825(98)00023-1. URL http://linkinghub.elsevier.com/retrieve/pii/S0045782598000231.
8. Goudreau GL. Evaluation of discrete methods for the linear dynamic response of elastic and viscoelastic solids. PhD Thesis, University of California, Berkeley, CA 1970.
9. Goudreau GL, Taylor RL. Evaluation of numerical integration methods in elastodynamics. *Computer Methods in Applied Mechanics and Engineering* Feb 1973; **2**(1):69–97, doi:10.1016/0045-7825(73)90023-6. URL http://linkinghub.elsevier.com/retrieve/pii/0045782573900236.
10. Hughes TJR. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover Publications: New York, 2000. URL http://store.doverpublications.com/0486411818.html.
11. Babuska IM, Ihlenburg F, Paik E, 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* Dec 1995; **128**(3-4):325–359, doi:10.1016/0045-7825(95)00890-X. URL http://linkinghub.elsevier.com/retrieve/pii/004578259500890X.
12. Hughes TJR, Brooks AN. A theoretical framework for Petrov–Galerkin methods with discontinuous weighting functions: application to the streamline upwind procedure. *Finite Elements in Fluids*, vol. IV, Gallagher R, Norrie D, Oden J, Zienkiewicz OC (eds.). John Wiley and Sons Ltd: Chichester, 1982; 47–65.
13. Arnold DN, Brezzi F, Cockburn B, Marini LD. Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems. *SIAM Journal on Numerical Analysis* 2002; **39**(5):1749–1779, doi:10.1137/S0036142901384162. URL http://link.aip.org/link/SJNAAM/v39/i5/p1749/s1&Agg=doi.
14. Ronghua L. Generalized Difference Methods for a Nonlinear Dirichlet Problem. *SIAM Journal on Numerical Analysis* 1987; **24**(1):77, doi:10.1137/0724007. URL http://link.aip.org/link/SJNAAM/v24/i1/p77/s1&Agg=doi.
15. Chou SH, Ye X. Unified Analysis of Finite Volume Methods for Second Order Elliptic Problems. *SIAM Journal on Numerical Analysis* 2007; **45**(4):1639, doi:10.1137/050643994. URL http://link.aip.org/link/SJNAAM/v45/i4/p1639/s1&Agg=doi.

16. Wohlmuth BI. A Mortar Finite Element Method Using Dual Spaces for the Lagrange Multiplier. *SIAM Journal on Numerical Analysis* 2000; **38**(3):989–1012, doi:10.1137/S0036142999350929. URL http://link.aip.org/link/SJNAAM/v38/i3/p989/s1&Agg=doi.
17. Harari I, Hughes TJR. Finite element methods for the helmholtz equation in an exterior domain: Model problems. *Computer Methods in Applied Mechanics and Engineering* May 1991; **87**(1):59–96, doi:10.1016/0045-7825(91)90146-W. URL http://linkinghub.elsevier.com/retrieve/pii/004578259190146W.
18. Thompson LL, Pinsky PM. A Galerkin least-squares finite element method for the two-dimensional Helmholtz equation. *International Journal for Numerical Methods in Engineering* Feb 1995; **38**(3):371–397, doi:10.1002/nme.1620380303. URL http://doi.wiley.com/10.1002/nme.1620380303.
19. Ikeuchi M, Inoue K, Sawami H, Niki H. Arbitrarily Shaped Hollow Waveguide Analysis by the α-Interpolation Method. *SIAM Journal on Applied Mathematics* 1981; **40**(1):90–98, doi:10.1137/0140007. URL http://link.aip.org/link/SMJMAP/v40/i1/p90/s1&Agg=doi.
20. Nakamura M, Hirasawa M. Eigenvalues of the Schrödinger Equation by the α-Interpolation Method. *SIAM Journal on Applied Mathematics* 1983; **43**(6):1286–1293, doi:10.1137/0143086. URL http://link.aip.org/link/SMJMAP/v43/i6/p1286/s1&Agg=doi.
21. Loula AFD, Fernandes DT. A quasi optimal Petrov–Galerkin method for Helmholtz problem. *International Journal for Numerical Methods in Engineering* Dec 2009; **80**(12):1595–1622, doi:10.1002/nme.2677. URL http://doi.wiley.com/10.1002/nme.2677.
22. Fernandes DT, Loula AFD. Quasi optimal finite difference method for Helmholtz problem on unstructured grids. *International Journal for Numerical Methods in Engineering* 2009; **82**:1244–1281, doi:10.1002/nme.2795. URL http://doi.wiley.com/10.1002/nme.2795.
23. Demkowicz L. Asymptotic convergence in finite and boundary element methods: part 1: theoretical results. *Computers & Mathematics with Applications* Jun 1994; **27**(12):69–84, doi:10.1016/0898-1221(94)90087-6. URL http://linkinghub.elsevier.com/retrieve/pii/0898122194900876.