INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2008; **57**:1291–1320
Published online in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1863

Analysis of a consistency recovery method for the 1D convection–diffusion equation using linear finite elements

Prashanth Nadukandi[1,*,†], Eugenio Oñate[1] and Julio Garcia[1,2]

[1]*International Center for Numerical Methods in Engineering (CIMNE), Universidad Politecnica de Catalunya (UPC), Edificio C1, Gran Capitan s/n, 08034 Barcelona, Spain*
[2]*COMPASS Ingenieria y Sistemas, S.A., Tuset 8, 7-2, 08006 Barcelona, Spain*

SUMMARY

For residual-based stabilization methods such as streamline-upwind Petrov–Galerkin (SUPG) and finite calculus (FIC), the higher-order derivatives of the residual that appear in the stabilization term vanish when simplicial elements are used. The sub-grid scale method using orthogonal sub-scales (OSS) attempts to recover the lost consistency by using a fine-scale projected residual in the stabilization term. The FIC method may also be cast into an OSS form with very little manipulation using an auxiliary convective projection equation. This paper discusses the gain/loss by recovering the consistency of the discrete residual in the stabilization terms via the form that includes the convective projection (as in the OSS method). We present the von Neumann analysis of the FIC method with recovered consistency (FIC_RC) for the 1D convection–diffusion problem and we compare it with the standard Bubnov–Galerkin linear finite element method and FIC/SUPG methods. The transient analysis is done by examining the discrete dispersion relation of the stabilization methods. The spectral results for the semi-discrete and fully discrete problem are presented with time integration done by the trapezoidal and second-order backward differencing formula schemes. The effect of lumping the effective mass matrix **T** is considered relative to using a consistent form. The effect of refinement in space and time is also discussed. Finally, an optimal expression for the stabilization parameter for the FIC_RC method on a uniform grid and for the steady state is given and its performance in the transient mode is discussed. Copyright © 2008 John Wiley & Sons, Ltd.

Received 14 November 2007; Revised 14 May 2008; Accepted 19 May 2008

KEY WORDS: finite element; stabilized methods; finite calculus; consistent discrete residual; discrete dispersion relation

*Correspondence to: Prashanth Nadukandi, International Center for Numerical Methods in Engineering (CIMNE), Universidad Politecnica de Catalunya (UPC), Edificio C1, Gran Capitan s/n, 08034 Barcelona, Spain.
†E-mail: npras@cimne.upc.edu

Contract/grant sponsor: Ministry for Science and Education

1. INTRODUCTION

In many transport processes arising in physical problems, convection essentially dominates diffusion. The design of numerical methods for such problems, which reflect their almost hyperbolic nature and guarantee that the discrete solution satisfies the physical conditions, is a subject that has been widely studied. In particular, for the convection–diffusion problem, the standard Galerkin finite element method (FEM) leads to numerical instabilities for the convection-dominated case. Several stabilization methods, for instance, the streamline-upwind Petrov–Galerkin (SUPG), Galerkin least-square, sub-grid scale (SGS), SGS with orthogonal sub-scales (OSS) etc., have been designed to overcome this numerical instability. A thorough comparison of some of these methods from the point of view of their formulation and the motivations that lead to them can be found in [1]. Stabilization procedures based on finite calculus (FIC) have also been developed as a general-purpose tool for improving the stability and accuracy of the convection–diffusion problem [2–5]. A residual correction method based on FIC was presented in [6] and is shown to yield a formulation equivalent to an OSS form [7] with very little manipulation.

For the convection–diffusion problem using the SUPG or FIC methods, the higher-order term (here the diffusion term) that appears in the stabilization term vanishes when simplicial elements are used. In [6] it is shown that for the elasticity problem, the form that includes the projected gradient of pressure into the stabilization terms (motivated by the OSS method) is essential to obtain accurate numerical results that converge in a more monotone manner and are less sensitive to the value of the stabilization parameter. In [8] a method was presented to globally reconstruct a continuous approximation to the diffusive flux for linear FEs using an L_2 projection and it was shown, in some cases (when advection and diffusion are on par), to greatly improve accuracy. It is important to note that again this improved accuracy is demonstrated for other related unknowns of the problem (such as pressure) and not for the transported unknown (say velocity). In addition, when convection dominates diffusion, there is little effect in the inclusion of the recovered diffusive flux. On the other hand, consistency recovery following the OSS philosophy is independent from the diffusive term. These observations are the motivation to investigate in detail the benefits of including similar projections for the convection–diffusion problem following the OSS philosophy. In other words, we try to answer the question—what do we gain/lose by recovering the consistency of the discrete residual in the stabilization terms for the convection-dominated case via the introduction of OSS-type convective projection?

The von Neumann analysis for the Galerkin and SUPG methods is relatively well known [9]. Relevant literature on the type of analysis presented here may be found in [10]. First, we present the von Neumann analysis for the 1D FIC method with recovered consistency (FIC_RC). This is achieved by including the convective projection into the stabilization term (motivated by the OSS method). It is then shown that in 1D the FIC and FIC_RC methods are equivalent to the SUPG and OSS methods, respectively. Consequently, the comparisons made between the former methods may be carried over to the latter methods. The transient analysis is done by examining the discrete dispersion relation (DDR) of the stabilization methods. The explanation for the occurrence of wiggles/oscillations in the transient evolution of the numerical solution was given by examining the dispersion relations of the continuous and discrete problems in [11]. It has been found that beyond a certain wave number ξ_{d}, the continuous and the discrete dispersion relations diverge [11, 12]. This wave number (ξ_{d}) is referred to as the *phase departure wave number* in [12]. If the bandwidth of the amplitude spectra of any given initial function has wave

DOI: 10.1002/fld

numbers greater than ξ_d, the initial function suffers a change of form (with wiggles/oscillations) in its transient evolution. Examining the respective DDRs, we seek to find if the stabilization methods provide any improvements in the solutions. A comparison of the DDR of the FIC_RC/OSS method is done with the DDRs of the Galerkin and FIC/SUPG methods for three standard time integration schemes. In addition, the range of wave numbers to which the DDR agrees with the continuous dispersion relation is shown to extend, should a consistent 'effective' mass matrix be preferred to a lumped one. Next, it is shown that unlike the FIC/SUPG method, the FIC_RC/OSS method introduces a certain rearrangement in the equation stencils at nodes on and adjacent to the domain boundary. Thus, using a uniform expression for the stabilization parameter (α) will lead to enhanced localized oscillations at the boundary. For the 1D steady-state problem, we present a new expression for α, which is optimal for uniform grids and provides negligible damping when used in the transient mode. Unfortunately for non-uniform grids, it leads to weak node-to-node oscillations.

2. TRANSIENT CONVECTION–DIFFUSION EQUATION

2.1. Problem statement

The statement of the multi-dimensional problem is as follows:

$$\frac{\partial \phi}{\partial t} + \mathbf{u} \cdot (\nabla \phi) - \nabla \cdot (k \nabla \phi) - f = 0 \quad \text{on } \Omega \tag{1a}$$

$$\phi(\mathbf{x}, t=0) = \phi_0(\mathbf{x}) \quad \text{on } \Omega \tag{1b}$$

$$\phi = \phi^{\mathrm{p}} \quad \text{on } \Gamma_{\mathrm{D}} \tag{1c}$$

$$k \nabla \phi \cdot \mathbf{n} = q^{\mathrm{p}} \quad \text{on } \Gamma_{\mathrm{N}} \tag{1d}$$

where $\mathbf{u}$ is the convection velocity, k is the diffusion, f is the source, $\phi(\mathbf{x}, t)$ is the transported variable, $\phi_0(\mathbf{x})$ is the initial solution, ϕ^{p} and q^{p} are the prescribed values of ϕ and the diffusive flux at the Dirichlet and Neumann boundaries, respectively. The FIC formulation of this problem (neglecting the time stabilization terms [4]) is as follows:

$$r - \tfrac{1}{2}\mathbf{h} \cdot \nabla(r) = 0 \quad \text{on } \Omega \tag{2a}$$

$$\phi(\mathbf{x}, t=0) = \phi_0(\mathbf{x}) \quad \text{on } \Omega \tag{2b}$$

$$\phi_h = \phi^{\mathrm{p}} \quad \text{on } \Gamma_{\mathrm{D}} \tag{2c}$$

$$k \nabla \phi_h \cdot \mathbf{n} + \tfrac{1}{2}\mathbf{h} \cdot \mathbf{n} r = q^{\mathrm{p}} \quad \text{on } \Gamma_{\mathrm{N}} \tag{2d}$$

$$r := \frac{\partial \phi}{\partial t} + \mathbf{u} \cdot (\nabla \phi) - \nabla \cdot (k \nabla \phi) - f \tag{2e}$$

DOI: 10.1002/fld

Let $(\cdot,\cdot)$ and $(\cdot,\cdot)_{\Gamma_N}$ denote the $L^2(\Omega)$ and $L^2(\Gamma_N)$ inner products, respectively. The variational form of problem (1) can be expressed as follows:

Find $\phi \in H_0^1(\Omega)$ such that $\forall w \in H_0^1(\Omega)$,

$$a(w,\phi) = l(w) \tag{3a}$$

$$a(w,\phi) := \left(w, \frac{\partial \phi}{\partial t}\right) + (w, \mathbf{u}\cdot\nabla\phi) + (\nabla w, k\nabla\phi) \tag{3b}$$

$$l(w) := (w,f) + (w, q^{\mathrm{p}})_{\Gamma_N} \tag{3c}$$

For the FIC equations, the variational form can be expressed as follows:

Find $\phi \in H_0^1(\Omega)$ such that $\forall w \in H_0^1(\Omega)$,

$$a(w,\phi) + \tfrac{1}{2}(\nabla\cdot(\mathbf{h}w), r) = l(w) \tag{4}$$

The calculation of the residual contribution in the stabilization term can be simplified if we introduce the projection of the convective term π via an auxiliary equation defined as

$$\pi = \mathbf{u}\cdot\nabla\phi - r \tag{5}$$

We express the residual, r, that occurs in the stabilization term $(\nabla\cdot(\mathbf{h}w), r)$ as a function of π. Thus, π becomes an addendum to the set of unknowns to be found. The integral equation system is now augmented, forcing the residual r expressed in terms of π using Equation (5) to vanish (in average) over the analysis domain. Problem (4) after the addendum of the convective projection π is expressed as follows:

Find $(\phi,\pi) \in (H_0^1(\Omega), H^1(\Omega))$ such that $\forall (w,z) \in (H_0^1(\Omega), H^1(\Omega))$,

$$a(w,\phi) + \tfrac{1}{2}(\nabla\cdot(\mathbf{h}w), \mathbf{u}\cdot\nabla\phi) - \tfrac{1}{2}(\nabla\cdot(\mathbf{h}w), \pi) = l(w) \tag{6a}$$

$$(z,\pi) = (z, \mathbf{u}\cdot\nabla\phi) \tag{6b}$$

We remark that the projection of the convective term provides consistency to the formulation, i.e. the system of equations (6) have the residual form that vanishes for the exact solution. Henceforth, we refer to the formulation given by Equation (6) as the FIC formulation with recovered consistency (FIC_RC). The convective projection is expected to capture the otherwise lost effect of higher-order terms in the residual when simplicial elements are used. The introduction of the convective projection variable π was deduced from an OSS approach in [7].

2.2. *Dispersion relation in 1D*

Any equation that admits plane wave solutions of the form $\exp[\mathrm{i}(\omega t - \xi x)]$, but with the property that the speed of propagation of these waves is dependent on ξ, is generally referred to as a *dispersive equation*. Here ξ, ω are the angular wave number and the frequency, respectively. The equation that expresses ω as a function of ξ is known as the *dispersion relation*. Generally, the transient convection–diffusion equation is a dispersive equation. In the limit, when diffusion tends to zero and the equation morphs into a pure convection problem, it tends to become non-dispersive [11]. In 1D and for the sourceless case ($f=0$), Equation (1a) can be expressed as follows:

$$\frac{\partial\phi}{\partial t} + u\frac{\partial\phi}{\partial x} - k\frac{\partial^2\phi}{\partial x^2} = 0 \tag{7}$$

DOI: 10.1002/fld

For a given discretization of size ℓ in space and an increment θ in time, we express the *Courant* and *Peclet* numbers as $C=u\theta/\ell$ and $\gamma=u\ell/2k$. We write down the dispersion relation (Equation (8)) for the continuous problem (Equation (7)) by propagating the plane wave solution $\phi=\exp[\mathrm{i}(\omega t-\xi x)]$. From Equation (8) we obtain, taking the limit $k\to 0$ or $\gamma\to\infty$, the dispersion relation for the pure convection problem

$$\omega=u\xi+\mathrm{i}k\xi^2 \tag{8a}$$

$$\omega\theta=C(\xi\ell)+\mathrm{i}\frac{C}{2\gamma}(\xi\ell)^2 \tag{8b}$$

Now let us consider the amplification of the solution from time step n to $n+1$ and at some given spatial point. The *amplification parameter* β as defined in [12] is given by Equation (9). It can be clearly seen that β is stationary in time and uniform in space. The amplification and phase shift per time step are given by the magnitude and argument of Equation (9), respectively. Thus, for the pure convection problem we can see that the amplification is unity, i.e. $|\beta|=1$. The phase and group velocities are given by Equations (10a) and (10b), respectively, [13]

$$\beta=\frac{\phi_j^{n+1}}{\phi_j^n}=\exp[\mathrm{i}\omega\theta]=\exp[-k\theta\xi^2]\exp[\mathrm{i}u\theta\xi]=\exp\left[-\frac{C}{2\gamma}(\xi\ell)^2\right]\exp[\mathrm{i}C(\xi\ell)] \tag{9}$$

$$V_{\mathrm{p}}(\xi)=\frac{\omega(\xi)}{\xi} \tag{10a}$$

$$V_{\mathrm{g}}(\xi)=\frac{\partial\omega(\xi)}{\partial\xi} \tag{10b}$$

3. FE DISCRETIZATION

3.1. Semi-discrete form

The semi-discrete (continuous in time, discrete in space) counterpart of the FIC method (4) can be written as follows:

Find $\phi\in H_0^1(\Omega)$ such that $\forall w\in H_0^1(\Omega)$,

$$a(w_h,\phi_h)+\sum_e\frac{1}{2}(\boldsymbol{\nabla}\cdot(\mathbf{h}w_h),r_h)_{\Omega^e}=l(w_h) \tag{11}$$

The stabilization term in Equation (11) has been expressed as a sum of the element contributions to allow for inter-element discontinuities in the term $\boldsymbol{\nabla}(r_h)$ of Equation (2), where $r_h:=r(\phi_h)$ is the residual of the FE approximation of the infinitesimal governing equation and $(\cdot,\cdot)_{\Omega^e}$ denote the $L^2(\Omega^e)$ inner product. Similarly, the discrete counterpart of the FIC_RC method (6) can be expressed as

Find $(\phi,\pi)\in(H_0^1(\Omega),H^1(\Omega))$ such that $\forall(w,z)\in(H_0^1(\Omega),H^1(\Omega))$,

$$a(w_h,\phi_h)+\sum_e\frac{1}{2}(\boldsymbol{\nabla}\cdot(\mathbf{h}w_h),\mathbf{u}\cdot\boldsymbol{\nabla}\phi_h)_{\Omega^e}-\sum_e\frac{1}{2}(\boldsymbol{\nabla}\cdot(\mathbf{h}w_h),\pi_h)_{\Omega^e}=l(w_h) \tag{12a}$$

$$(z_h,\pi_h)=(z_h,\mathbf{u}\cdot\boldsymbol{\nabla}\phi_h) \tag{12b}$$

DOI: 10.1002/fld

The variables in Equations (11) and (12) interpolated by the FE shape functions N^a can be expressed as follows:

$$\phi_h = N^a \bar{\phi}^a, \quad \pi_h = N^a \bar{\pi}^a, \quad w_h = N^a \bar{w}^a, \quad z_h = N^a \bar{z}^a \tag{13}$$

The discrete problems (11) and (12) can be written in matrix notation using Equations (14) and (15), respectively, as follows:

$$[\mathbf{M}+\mathbf{S}2]\cdot\dot{\bar{\Phi}}+[\mathbf{C}+\mathbf{D}+\mathbf{S}1+\mathbf{S}3]\cdot\bar{\Phi}=\mathbf{fg}+\mathbf{fs} \tag{14}$$

$$\mathbf{M}\cdot\dot{\bar{\Phi}}+[\mathbf{C}+\mathbf{D}+\mathbf{S}1]\cdot\bar{\Phi}-\mathbf{S}2\cdot\bar{\Pi}=\mathbf{fg} \tag{15a}$$

$$\mathbf{M}\cdot\bar{\Pi}-\mathbf{C}\cdot\bar{\Phi}=\mathbf{0} \tag{15b}$$

where the element contributions to the above matrices and vectors are given by

$$\mathbf{C}^e_{ab}=(N^a,\mathbf{u}\cdot\nabla N^b)_{\Omega^e}, \quad \mathbf{D}^e_{ab}=(\nabla N^a, k\nabla N^b)_{\Omega^e} \tag{16a}$$

$$\mathbf{M}^e_{ab}=(N^a,N^b)_{\Omega^e}, \quad \mathbf{S}1^e_{ab}=\tfrac{1}{2}(\nabla\cdot(\mathbf{h}N^a),\mathbf{u}\cdot\nabla N^b)_{\Omega^e} \tag{16b}$$

$$\mathbf{S}2^e_{ab}=\tfrac{1}{2}(\nabla\cdot(\mathbf{h}N^a),N^b)_{\Omega^e}, \quad \mathbf{S}3^e_{ab}=-\tfrac{1}{2}(\nabla\cdot(\mathbf{h}N^a),\nabla\cdot(k\nabla N^b))_{\Omega^e} \tag{16c}$$

$$\mathbf{fg}^e_a=(N^a,f)_{\Omega^e}+(N^a,q^{\mathrm{p}})_{\Gamma_{\mathrm{N}}}, \quad \mathbf{fs}^e_a=\tfrac{1}{2}(\nabla\cdot(\mathbf{h}N^a),f)_{\Omega^e} \tag{16d}$$

Note that Equation (15b) corresponds to the L_2-projection of the term $\mathbf{u}\cdot\nabla(\phi_h)$ onto the space spanned by the shape functions. Whenever the latter term admits discontinuities, the projection is non-monotone [7]. A monotone projection of the convective term can be achieved if the mass matrix $\mathbf{M}$ that appears in Equation (15b) is lumped. In addition, the expression for Π will also have local support only when $\mathbf{M}$ is lumped. This feature allows us to study generic nodal equation stencils for the interior of the domain. Henceforth, we always consider Equation (15b) with $\mathbf{M}$ lumped. Equations (14) and (15) may be expressed in a general form as shown in Equation (17). Table I defines the corresponding matrices for the FIC and FIC_RC methods

$$\mathbf{T}\cdot\dot{\bar{\Phi}}+[\mathbf{C}+\mathbf{D}+\mathbf{S}]\cdot\bar{\Phi}=\mathbf{f} \tag{17}$$

where matrix $\mathbf{T}$ is the 'effective' mass matrix.

Note that for the direction of $\mathbf{h}$ being the same as that of the velocity $\mathbf{u}$, i.e. $\mathbf{h}=2\tau\mathbf{u}$ and assuming τ constant within an element, the form of the stabilization term $(\nabla\cdot(\mathbf{h}w_h),r_h)_{\Omega^e}$ in Equation (11) is identical to that of the standard SUPG method. Thus, with this choice of $\mathbf{h}$ the FIC and FIC_RC methods are identical to the SUPG and OSS methods, respectively. The general direction of $\mathbf{h}$ naturally introduces stabilization along the streamlines and also along the directions

Table I. Matrix definitions for the FIC and FIC_RC methods.

	FIC	FIC_RC
T-lumped	$\mathbf{M}_L+\mathbf{S}2$	$\mathbf{M}_L$
T-consistent	$\mathbf{M}+\mathbf{S}2$	$\mathbf{M}$
S	$\mathbf{S}1+\mathbf{S}3$	$\mathbf{S}1-\mathbf{S}2\cdot\mathbf{M}_L^{-1}\cdot\mathbf{C}$
f	$\mathbf{fg}+\mathbf{fs}$	$\mathbf{fg}$

DOI: 10.1002/fld

of the gradient of the solution transverse to the velocity vector. The FIC formulation therefore incorporates the best features of the SUPG and the shock-capturing methods. Applications of the FIC-FEM formulation to a wide range of convection–diffusion problems with sharp gradients are presented in [14]. We remark that in 1D (assuming h to be constant within an element) the FIC and the FIC_RC methods are identical to the SUPG and OSS methods, respectively. Thus, the comparisons made between the FIC and FIC_RC methods may be carried over to those between the SUPG and OSS methods.

3.2. DDR in 1D

The DDR of the semi-discrete problem (semi-DDR) and when the temporal terms are discretized using two classes of time discretization schemes are investigated in this section. The time discretization schemes considered are the *trapezoidal scheme* and the *second-order backward differencing formula* (*BDF2*). The effects on the numerical dispersion due to the choice of the form of the effective mass matrix **T** (lumped or consistent) in Equation (17) are also studied. The flag *lumped* or *consistent* refers only to matrix **T** as defined in Table I for the FIC and FIC_RC methods. The DDRs are written by inserting a plane wave solution of the form $\phi=\exp[\mathrm{i}(\omega t-\xi x)]$ into the corresponding equation stencils. Taking the limit $\gamma\to\infty$, we recover the DDR for the pure convection problem.

3.2.1. Semi-DDR. We study the equation stencil for an interior node of the semi-discrete problem given by Equation (17) with $f=0$ in 1D. For a compact representation of the stencils, we introduce the following definition:

$$(\star):=\left(\frac{u}{2}\right)(\phi_{j+1}-\phi_{j-1})-\left(\frac{k}{\ell}+\frac{u\alpha}{2}\right)(\phi_{j+1}-2\phi_j+\phi_{j-1}) \tag{18}$$

FIC/SUPG method:

$$\left(\frac{\ell}{6}\right)\left[\left(\frac{-3\alpha}{2}\right)\dot{\phi}_{j+1}+6\dot{\phi}_j+\left(\frac{3\alpha}{2}\right)\dot{\phi}_{j-1}\right]+(\star)=0 \quad \mathbf{T}\text{-lumped} \tag{19a}$$

$$\left(\frac{\ell}{6}\right)\left[\left(1-\frac{3\alpha}{2}\right)\dot{\phi}_{j+1}+4\dot{\phi}_j+\left(1+\frac{3\alpha}{2}\right)\dot{\phi}_{j-1}\right]+(\star)=0 \quad \mathbf{T}\text{-consistent} \tag{19b}$$

FIC_RC/OSS method:

$$\ell\dot{\phi}_j+(\star)+\left(\frac{u\alpha}{8}\right)(\phi_{j+2}-2\phi_j+\phi_{j-2})=0 \quad \mathbf{T}\text{-lumped} \tag{20a}$$

$$\left(\frac{\ell}{6}\right)(\dot{\phi}_{j+1}+4\dot{\phi}_j+\dot{\phi}_{j-1})+(\star)+\left(\frac{u\alpha}{8}\right)(\phi_{j+2}-2\phi_j+\phi_{j-2})=0 \quad \mathbf{T}\text{-consistent} \tag{20b}$$

Making $\alpha=0$ in Equations (19) and (20), we recover the standard Galerkin method. The semi-DDRs for all the methods and for the **T**-lumped, **T**-consistent cases can be expressed in a generic and compact manner as follows:

$$\omega_h=-\mathrm{i}\frac{B}{\theta A} \tag{21}$$

DOI: 10.1002/fld

where

$$A := \begin{cases} \dfrac{1}{C} & \text{Galerkin, FIC\_RC/OSS methods, } \mathbf{T}\text{-lumped} \\ \dfrac{2+\cos(\xi\ell)}{3C} & \text{Galerkin, FIC\_RC/OSS methods, } \mathbf{T}\text{-consistent} \\ \dfrac{1}{C}+\mathrm{i}\dfrac{\alpha}{2C}\sin(\xi\ell) & \text{FIC/SUPG methods, } \mathbf{T}\text{-lumped} \\ \dfrac{2+\cos(\xi\ell)}{3C}+\mathrm{i}\dfrac{\alpha}{2C}\sin(\xi\ell) & \text{FIC/SUPG methods, } \mathbf{T}\text{-consistent} \end{cases} \tag{22a}$$

$$B := \begin{cases} \mathrm{i}\sin(\xi\ell)-2\sin^2\left(\dfrac{\xi\ell}{2}\right)\left(\dfrac{1}{\gamma}\right) & \text{Galerkin method} \\ \mathrm{i}\sin(\xi\ell)-2\sin^2\left(\dfrac{\xi\ell}{2}\right)\left(\dfrac{1}{\gamma}+\alpha\right) & \text{FIC/SUPG methods} \\ \mathrm{i}\sin(\xi\ell)-2\sin^2\left(\dfrac{\xi\ell}{2}\right)\left(\dfrac{1}{\gamma}+\alpha\sin^2\left(\dfrac{\xi\ell}{2}\right)\right) & \text{FIC\_RC/OSS methods} \end{cases} \tag{22b}$$

3.2.2. Trapezoidal scheme. The structure of the equation stencil for an interior node with respect to the spatial indices is the same as in the semi-discrete problem. Henceforth, we express the fully discrete system in the matrix notation only

$$\mathbf{T}\cdot\frac{\Phi^{n+1}-\Phi^n}{\theta}+[\mathbf{C}+\mathbf{D}+\mathbf{S}]\cdot\Phi^{n+\sigma}=\mathbf{O} \tag{23a}$$

$$\Phi^{n+\sigma}:=\sigma\Phi^{n+1}+(1-\sigma)\Phi^n \tag{23b}$$

Making $\sigma=\{0, 0.5, 1\}$ we recover the *forward Euler*, *Crank–Nicholson* and *backward Euler* schemes, respectively. The DDR for the trapezoidal scheme can be expressed in terms of A and B defined in Equation (22) as follows:

$$\exp[\mathrm{i}\omega_h\theta]=\frac{A+(1-\sigma)B}{A-\sigma B} \quad \text{or equivalently} \tag{24a}$$

$$\tan\left(\frac{\omega_h\theta}{2}\right)=-\mathrm{i}\frac{B}{2A+(1-2\sigma)B} \tag{24b}$$

3.2.3. BDF2 scheme. The fully discrete system of equations after time discretization by the BDF2 scheme is given by

$$\mathbf{T}\cdot\frac{3\Phi^{n+1}-4\Phi^n+\Phi^{n-1}}{2\theta}+[\mathbf{C}+\mathbf{D}+\mathbf{S}]\cdot\Phi^{n+1}=\mathbf{O} \tag{25}$$

DOI: 10.1002/fld

The DDR for the BDF2 scheme is a quadratic relation in $\exp[\mathrm{i}\omega_h\theta]$. The solution to the quadratic equation gives two expressions for the DDR, which can be expressed as follows:

$$\exp[\mathrm{i}\omega_h\theta] = \frac{2A+\sqrt{A^2+2AB}}{3A-2B} \tag{26a}$$

$$\exp[\mathrm{i}\omega_h\theta] = \frac{2A-\sqrt{A^2+2AB}}{3A-2B} \tag{26b}$$

We remark that the solution given by Equation (26b) predicts negative values of $\Re(\omega_h)$ for positive wave numbers. Thus, we consider the solution given by Equation (26a) as the only acceptable solution.

3.3. DDR plots

The DDRs presented in the previous section represent the frequency as a function of six independent variables, i.e. $\omega_h := \omega_h(\xi, \ell, \theta, C, \gamma, \alpha)$. For a feasible graphical representation of the DDRs, we freeze some of them and normalize the frequency and wave numbers to retain maximum generality. In the DDR plots, we consider only the pure convection problem ($k=0$). The stabilization parameter $\alpha=1.0$ is chosen. This corresponds to the optimal value for the SUPG/FIC method in 1D and for a uniform mesh. This choice is made for convenience and comparison of the effects of the stabilization term introduced by the considered methods. Note that the DDRs are periodic in ξ and the corresponding fundamental domain is $\xi \in [-\pi/\ell, \pi/\ell]$. The Nyquist frequency in space is $\xi_{\mathrm{nq}} = \pi/\ell$ and in time is $\omega_{\mathrm{nq}} = \pi/\theta$. Thus in the DDR plots, we do not consider wave numbers and frequencies beyond the Nyquist limits. It can be shown with respect to the exact dispersion relation (Equation (8)) that this condition corresponds to choosing $C \leqslant 1$. We normalize the wave number ξ by the Nyquist limit ξ_{nq}, i.e. $\xi^* = \xi/\xi_{\mathrm{nq}}$. The frequency ω_h is normalized as $\omega_h^* = \omega_h/(u\xi_{\mathrm{nq}}) = \omega_h\theta/(C\pi) = \omega_h\ell/(u\pi)$. The DDRs are now expressed with respect to the normalized wave number and frequency, i.e. $\omega_h^* := \omega_h^*(\xi^*, C)$. The plotting domain considered is $(\xi^*, C) = (0,1)\times(0,1)$.

3.3.1. Semi-discrete. For the semi-discrete problem, the DDR no longer depends on θ. The frequency ω_h^* is now only a function of ξ^*, i.e. $\omega_h^* := \omega_h^*(\xi^*)$. The amplification at time $t_n = n\theta$ is given by $\exp[-\Im(\omega_h)t_n] = \exp[-\Im(\omega_h^*)\pi(ut_n/\ell)] = \exp[-\Im(\omega_h^*)\pi Cn]$. This means that should $\Im(\omega_h^*) \neq 0$ the amplification at any given time is independent of the time step θ but dependent on the space discretization ℓ. Thus, we present 1D plots for the following:

- Plot of $\Re(\omega_h^*)$ vs ξ^*: The departure wave number ξ_{d}^* is marked such that $\forall \xi^* \leqslant \xi_{\mathrm{d}}^*$ we have $|\Re(\omega^* - \omega_h^*)| \leqslant 0.001$.
- Amplification plots using $C=0.1$ and at times $\theta, 2\theta, 100\theta, 200\theta$, and 300θ s.

3.3.2. Fully discrete. For the fully discrete case, the time integration schemes considered are the *backward Euler*, *Crank–Nicholson*, and *BDF2*. The frequency is now a function of both ξ^* and C, i.e. $\omega_h^* := \omega_h^*(\xi^*, C)$. The amplification at time $t_n = n\theta$ is given by $\exp[-\Im(\omega_h^*)\pi(ut_n/\ell)]$, the same as for the semi-discrete case except for the fact that ω_h^* is now dependent on C also. Contour plots are presented for the following:

- $\log_{10}(|\Re(\omega^* - \omega_h^*)|)$ vs (ξ^*, C): The contour of values $\{-5, -4, -3, -2, -1\}$ are shown.
- $\log_{10}(\Im(\omega_h^*))$ vs (ξ^*, C): The contour of values $\{-3, -2.5, -2, -1.5, -1, -0.5, 0\}$ are shown.

DOI: 10.1002/fld

Only the contour plots for the normalized group velocity, i.e. $\partial\omega_h^*/\partial\xi^*$ vs (ξ^*, C), for the Crank–Nicholson scheme have been presented. The original group velocity can be recovered as follows: $\partial\omega_h/\partial\xi = u\partial\omega_h^*/\partial\xi^*$ (Figures 1–9).

4. STABILIZATION PARAMETERS

In this section, we propose optimal values for the stabilization parameters for the FIC_RC method in 1D and on a uniform mesh. We consider the steady-state form of the discrete problem (17) for the sourceless case ($f=0$). The equation stencil for an interior node j is as follows:

$$\left(\frac{u}{2}\right)[\phi_{j+1}-\phi_{j-1}]-\left(k+\frac{uh}{2}\right)\left[\frac{\phi_{j+1}-2\phi_j+\phi_{j-1}}{\ell}\right]+\left(\frac{uh}{8}\right)\left[\frac{\phi_{j+2}-2\phi_j+\phi_{j-2}}{\ell}\right]=0 \quad (27)$$

The analytical solution of the steady-state form of problem (17) in the 1D space with only Dirichlet boundary conditions and source $f=0$ is

$$\phi(x)=\phi_l^{\mathrm{p}}+(\phi_r^{\mathrm{p}}-\phi_l^{\mathrm{p}})\left[\frac{\exp[ux/k]-1}{\exp[uL/k]-1}\right] \quad (28)$$

where L is the length of the 1D domain and ϕ_l^{p}, ϕ_r^{p} are the prescribed values of ϕ at the left and right ends of the domain, respectively. We now express the characteristic length in terms of the element size as $h=\alpha\ell$. The optimal value of the stabilization parameter α is found by substituting the analytical solution into the stencil given in Equation (27). This leads to the following expression of α for the interior nodes:

$$\alpha=\frac{-1}{\sinh^2(\gamma)}\left[\coth(\gamma)-\frac{1}{\gamma}\right] \quad (29)$$

where γ is the element *Peclet number* given by $\gamma=u\ell/2k$. The plot of the parameter α vs γ is shown in Figure 10.

It can be seen that $\alpha\approx 0$ for $\gamma\geqslant 3$. Hence, for high Peclet numbers the formulation suggests using very small values of α. That is, for large values of γ the formulation breaks down into the standard Galerkin method without stabilization and one can expect spurious oscillations in the numerical solution. The clues to reason out this behavior can be found by examining the assembly of the linear system.

The 1D problem is discretized by linear elements and the final form of the FE assembly is examined. For simplicity, the nodes are numbered from left to right as shown in Figure 11. For the interior nodes, the equation stencil is as given by Equation (27). For the left penultimate boundary node, here node 2 as per the numbering scheme (Figure 11), we find the following stencil:

$$\left(\frac{u}{2}\right)[\phi_3-\phi_1]-\left(k+\frac{uh}{2}\right)\left[\frac{\phi_3-2\phi_2+\phi_1}{\ell}\right]+\left(\frac{uh}{8}\right)\left[\frac{\phi_4-3\phi_2+2\phi_1}{\ell}\right]=0 \quad (30)$$

DOI: 10.1002/fld

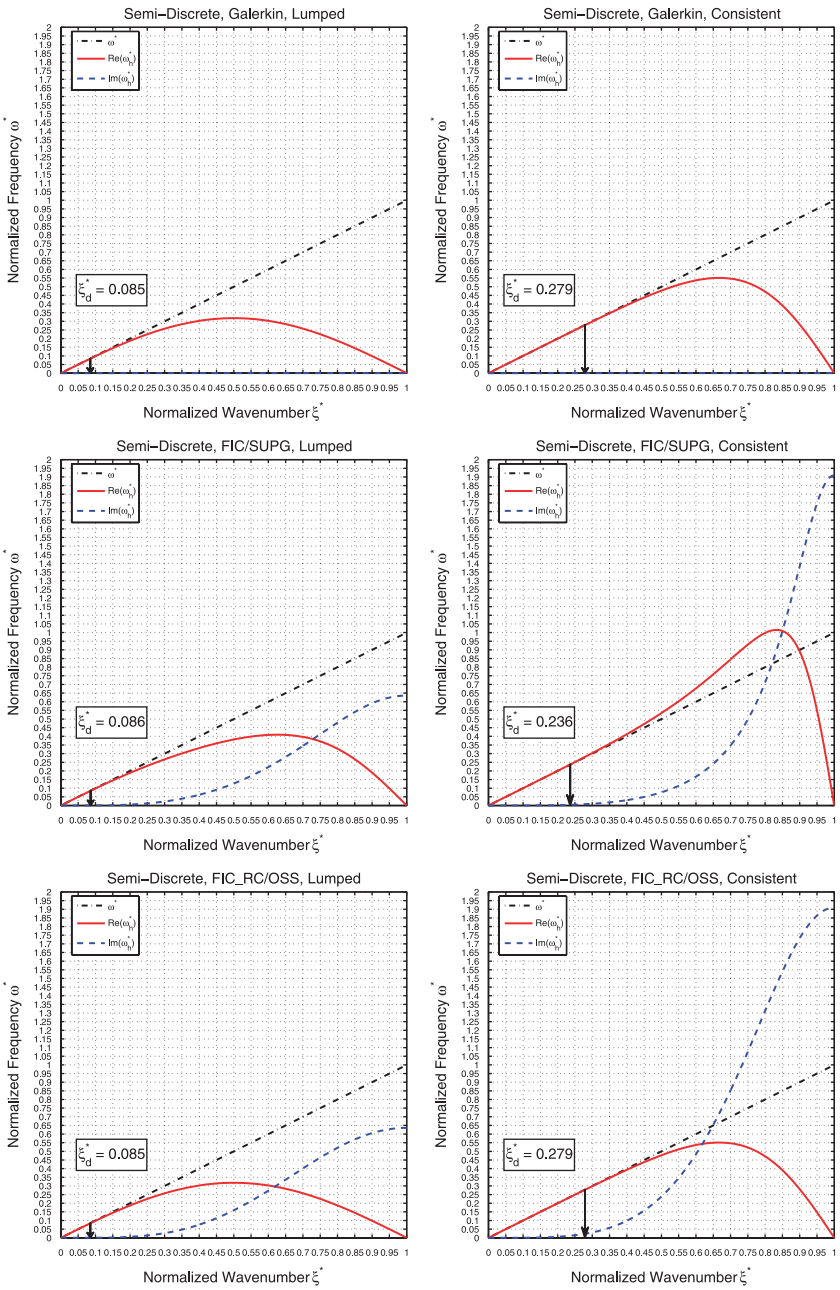

Figure 1. Plot of ω_h^* vs ξ^* for the semi-discrete problem. Frequencies ω^* and ω_h^* correspond to the continuous and discretized problems, respectively. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

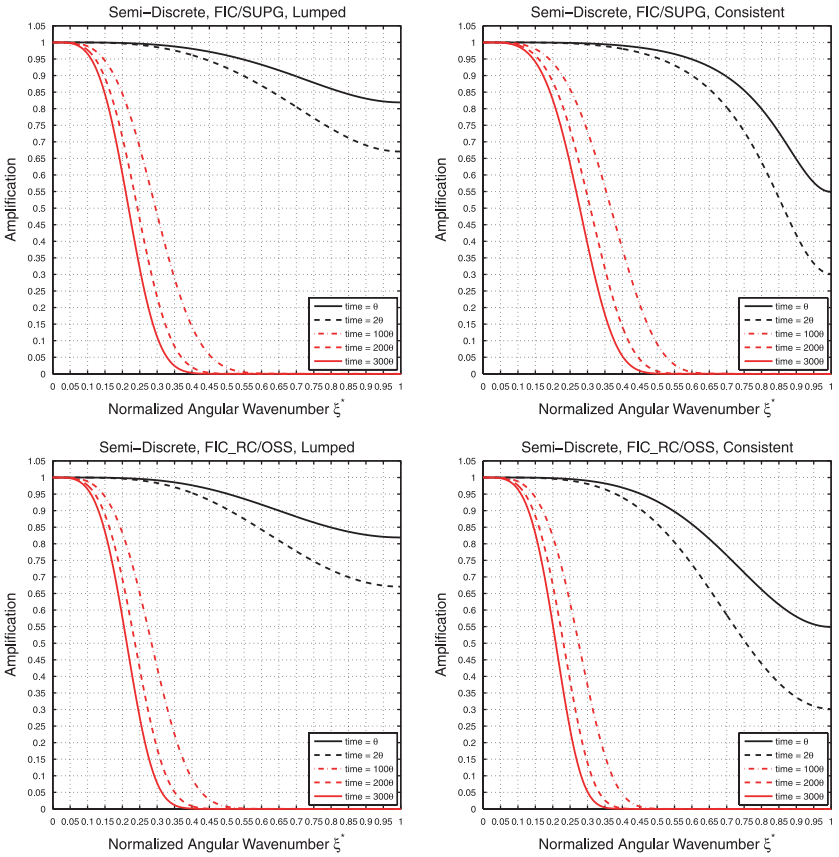

Figure 2. Amplification plots of the semi-discrete problem for the FIC/SUPG and FIC_RC/OSS methods. $\alpha = 1.0$ and $C = 0.1$ are used. The amplification for the Galerkin method is not shown here as it is equal to 1.

For the right penultimate boundary node, i.e. the node n in Figure 11, we find the following stencil:

$$\left(\frac{u}{2}\right)[\phi_{n+1}-\phi_{n-1}]-\left(k+\frac{uh}{2}\right)\left[\frac{\phi_{n+1}-2\phi_n+\phi_{n-1}}{\ell}\right]+\left(\frac{uh}{8}\right)\left[\frac{2\phi_{n+1}-3\phi_n+\phi_{n-2}}{\ell}\right]=0 \quad (31)$$

Thus, we see that there is a rearrangement of the equation stencils for the nodes that lie next to the boundary. The stencils for the boundary nodes also get rearranged, but we do not consider them here as the focus here is to deal with problems with Dirichlet boundary conditions. The deviation of the nodal equations for the penultimate nodes from the interior nodes is responsible for the spurious oscillations. It has been shown using a spectral analysis framework in [15, 16] that the asymmetry in the stencils brings about anti-diffusion even when they are used with central difference schemes and their effect is not localized, thus being responsible for spurious numerical oscillations. A simpler explanation would be that the rearrangement of the stencils near the boundary require different expressions for the stabilization parameter for those nodes.

DOI: 10.1002/fld

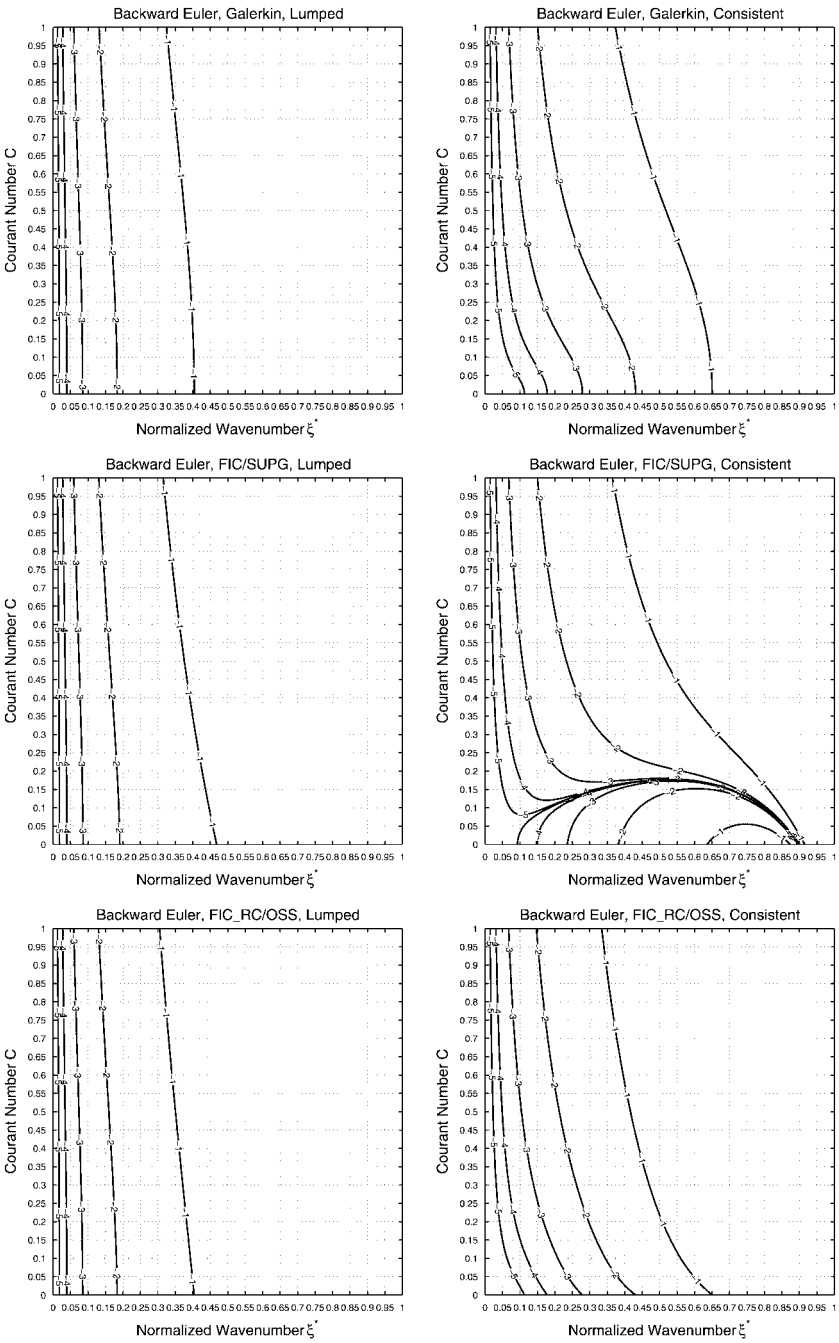

Figure 3. Contour plot of $\log_{10}[\Re(|\omega_h^* - \omega^*|)]$ for the backward Euler scheme. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

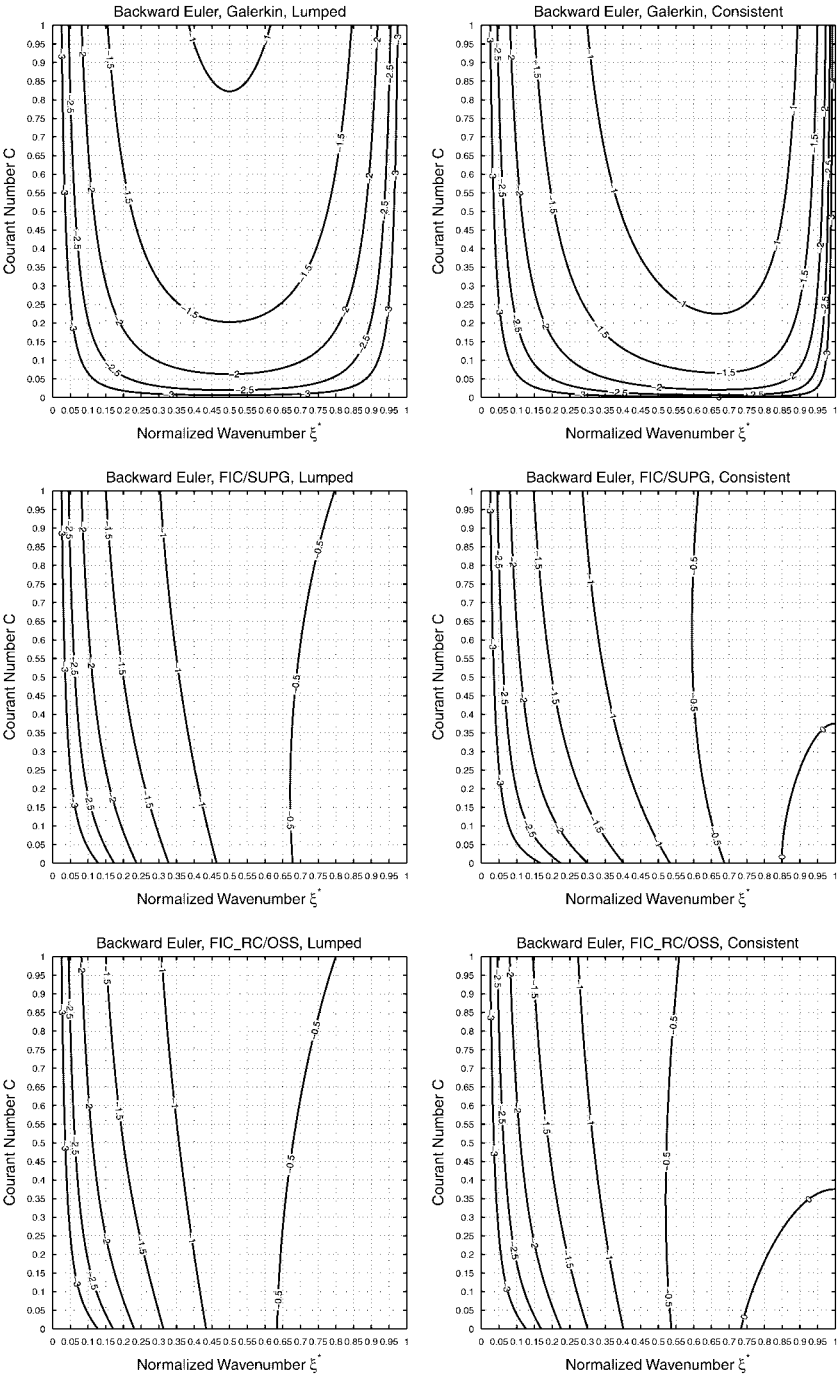

Figure 4. Contour plot of $\log_{10}[\Im(\omega_h^*)]$ for the backward Euler scheme. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

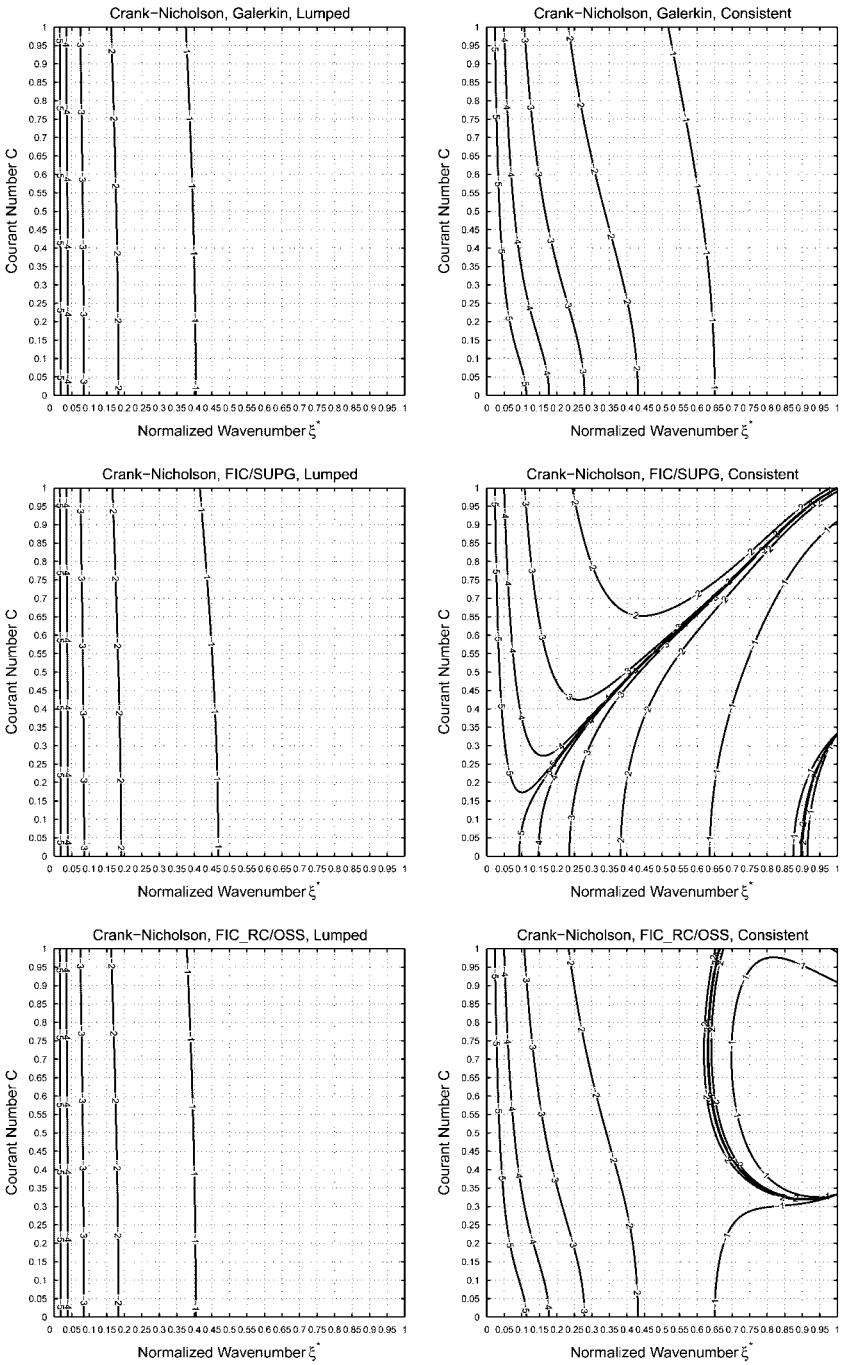

Figure 5. Contour plot of $\log_{10}[\Re(|\omega_h^* - \omega^*|)]$ for the Crank–Nicholson scheme. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

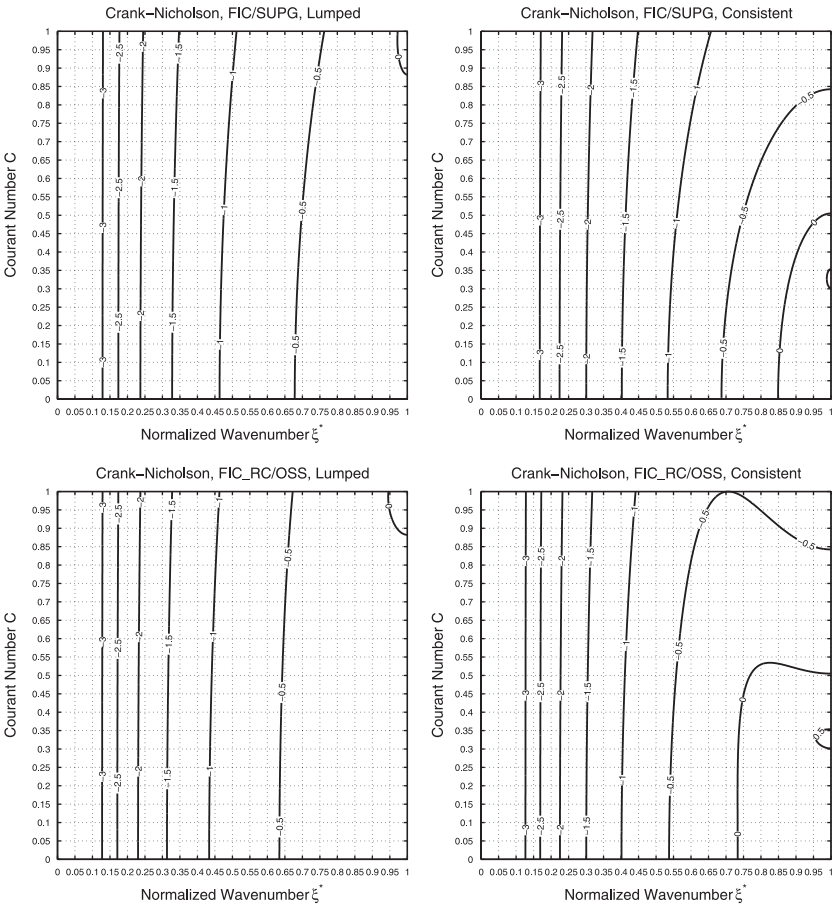

Figure 6. Contour plot of $\log_{10}[\Im(\omega_h^*)]$ for the Crank–Nicholson scheme. $\alpha=1.0$ is used. The plots for the Galerkin method is not shown here as $\Im(\omega_h^*)=0$.

The optimal values of the stabilization parameters for these penultimate nodes on a uniform mesh are as follows. Equations (32a) and (32b) correspond to the optimal values for the nodes 2 and n, respectively. Figure 12 illustrates the variation of α with respect to γ

$$\alpha_2 = \frac{4}{[2-e^{2\gamma}]}\left[\coth(\gamma)-\frac{1}{\gamma}\right] \tag{32a}$$

$$\alpha_n = \frac{4}{[2-e^{-2\gamma}]}\left[\coth(\gamma)-\frac{1}{\gamma}\right] \tag{32b}$$

An alternative to nodal stabilization parameters is to find optimal values of the stabilization parameters for the elements adjacent to the boundary. For elements lying in the domain interior, we use the value given by Equation (29). Denoting the stabilization parameters for elements 1, n and in the domain interior as $\alpha^{(1)}$, $\alpha^{(n)}$ and α, respectively, we find the following equation stencils.

DOI: 10.1002/fld

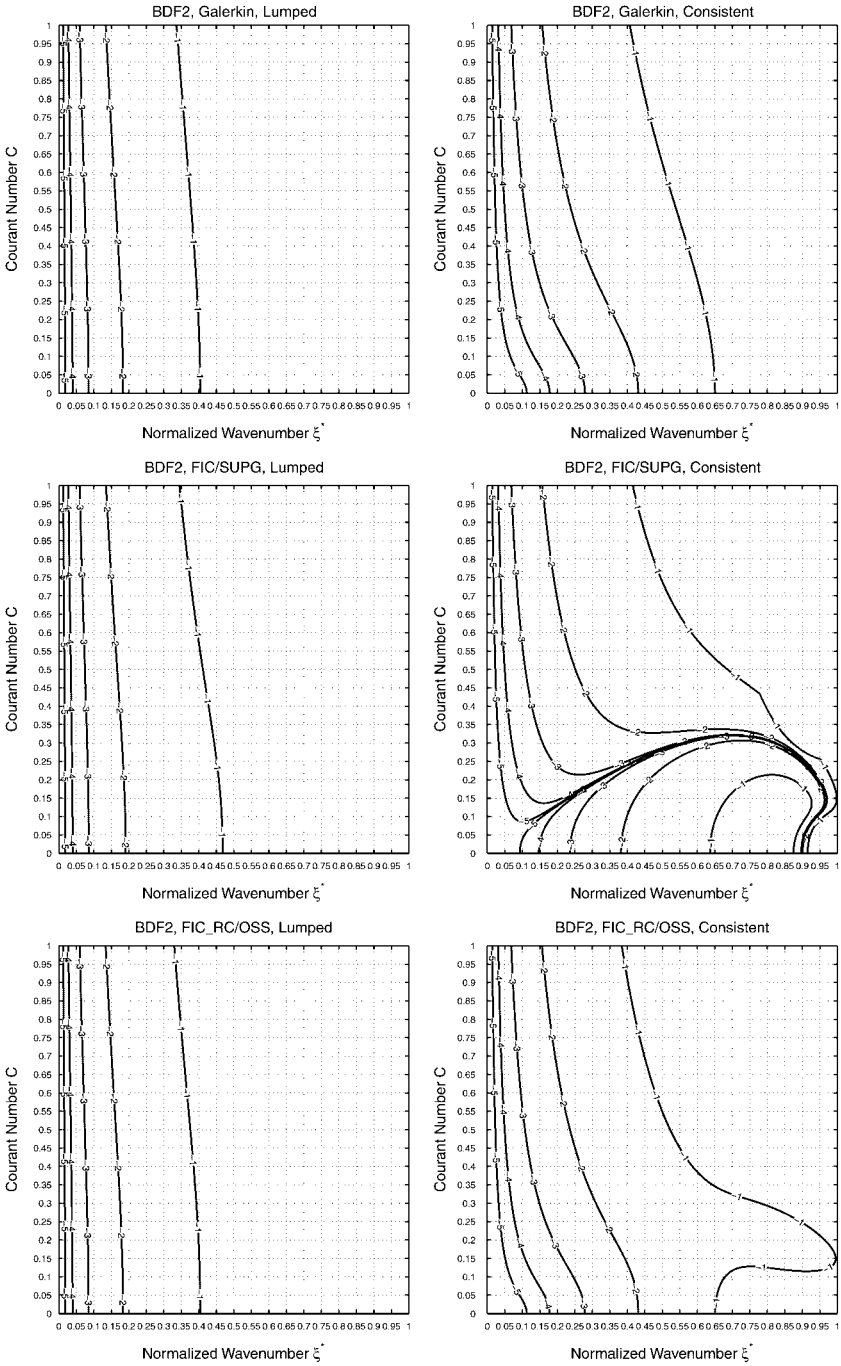

Figure 7. Contour plot of $\log_{10}[\Re(|\omega_h^* - \omega^*|)]$ for the BDF2 scheme. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

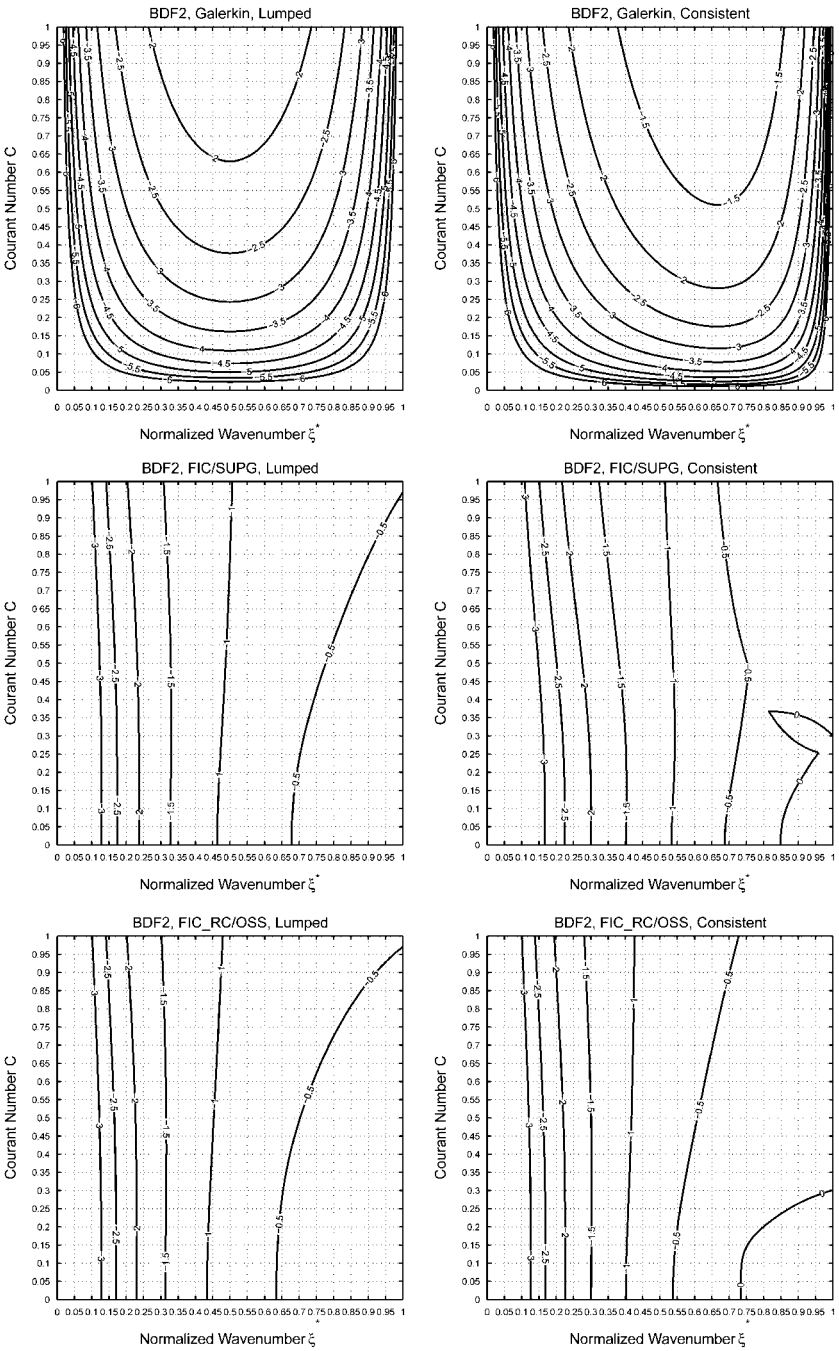

Figure 8. Contour plot of $\log_{10}[\Im(\omega_h^*)]$ for the BDF2 scheme. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

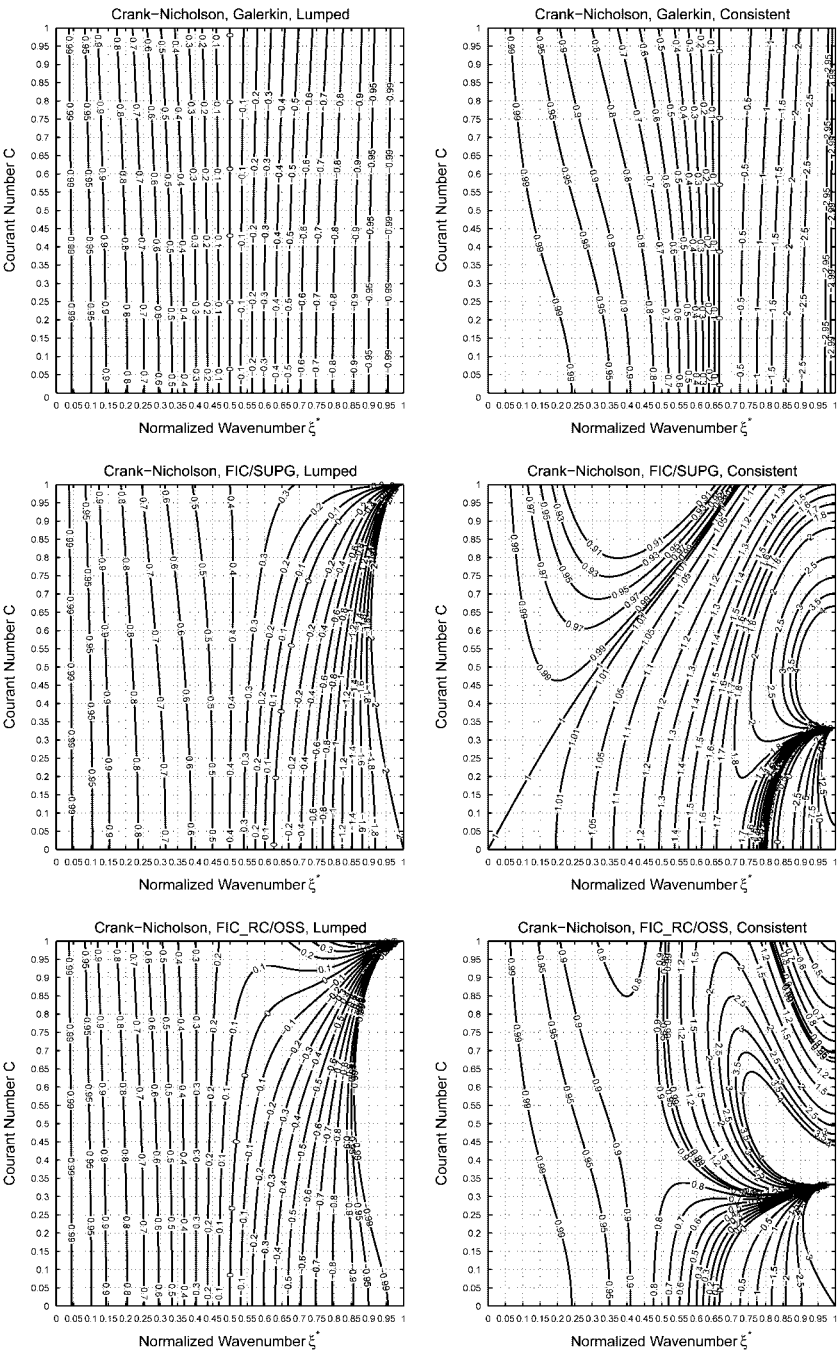

Figure 9. Real part of the normalized group velocity $\partial\omega_h^*/\partial\xi^*$ vs ξ^* for the Crank–Nicholson scheme; $\alpha = 1.0$ is used.

DOI: 10.1002/fld

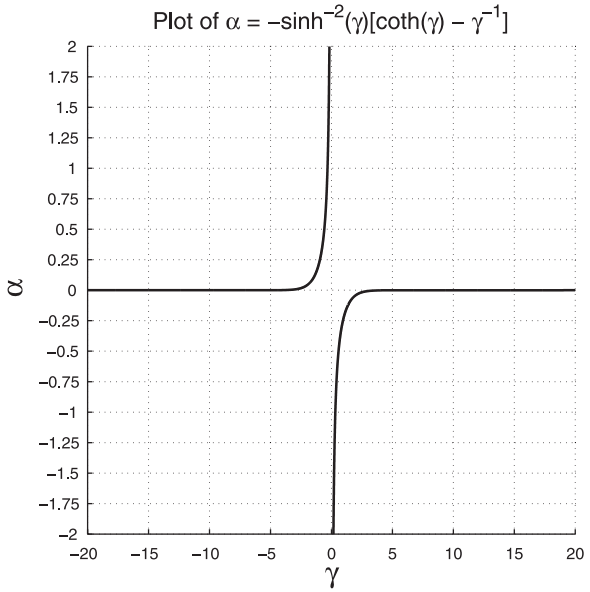

Figure 10. α vs γ for interior nodes.

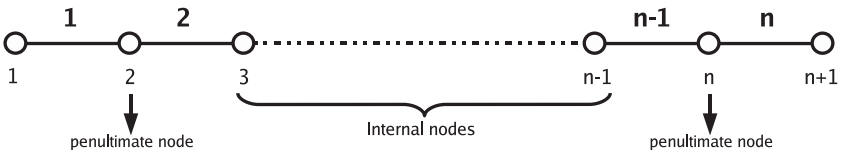

Figure 11. Node stencil for the 1D problem.

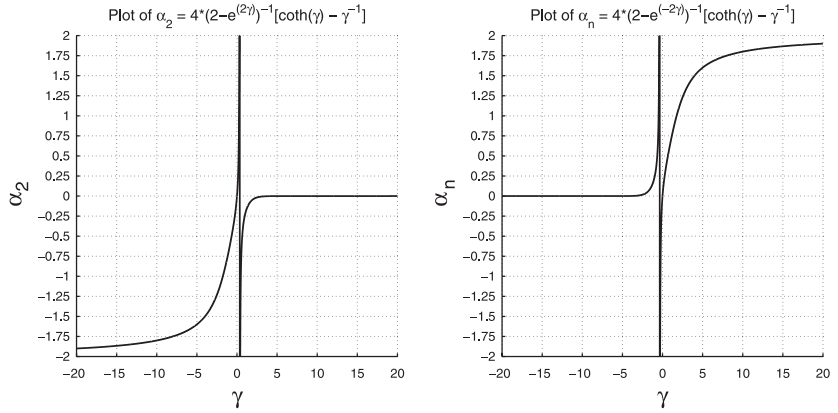

Figure 12. α vs γ for penultimate boundary nodes.

DOI: 10.1002/fld

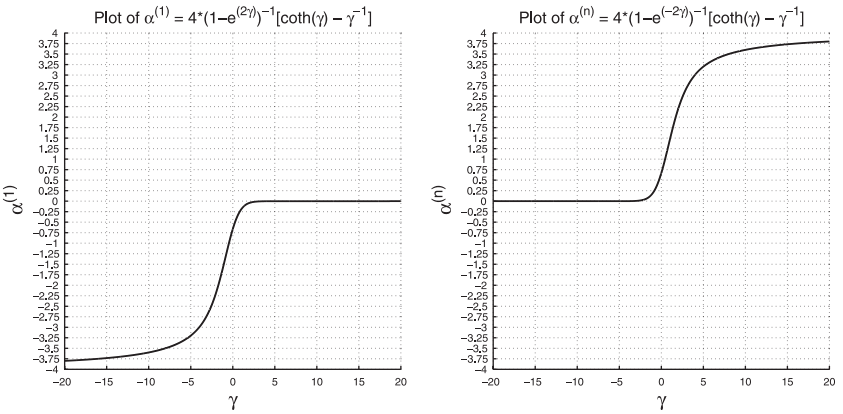

Figure 13. α vs γ for elements adjacent to the boundary.

Node 2:

$$\left(\frac{u}{2}\right)[\phi_3-\phi_1]-\left(\frac{k}{\ell}\right)[\phi_3-2\phi_2+\phi_1]-\left(\frac{u}{2}\right)[\alpha\phi_3-(\alpha^{(1)}+\alpha)\phi_2+\alpha^{(1)}\phi_1]$$
$$+\left(\frac{u}{8}\right)[\alpha\phi_4+(\alpha-\alpha^{(1)})\phi_3-(\alpha+2\alpha^{(1)})\phi_2+(3\alpha^{(1)}-\alpha)\phi_1]=0 \tag{33}$$

Node n:

$$\left(\frac{u}{2}\right)[\phi_{n+1}-\phi_{n-1}]-\left(\frac{k}{\ell}\right)[\phi_{n+1}-2\phi_n+\phi_{n-1}]-\left(\frac{u}{2}\right)[\alpha^{(n)}\phi_{n+1}-(\alpha^{(n)}+\alpha)\phi_n+\alpha\phi_{n-1}]$$
$$+\left(\frac{u}{8}\right)[(3\alpha^{(n)}-\alpha)\phi_{n+1}-(\alpha+2\alpha^{(n)})\phi_n+(\alpha-\alpha^{(n)})\phi_{n-1}+\alpha\phi_{n-2}]=0 \tag{34}$$

The optimal values for $\alpha^{(1)}$ and $\alpha^{(n)}$ are given by Equations (35a) and (35b), respectively. Figure 13 illustrates the variation of α with respect to γ

$$\alpha^{(1)}=\frac{4}{[1-e^{2\gamma}]}\left[\coth(\gamma)-\frac{1}{\gamma}\right] \tag{35a}$$

$$\alpha^{(n)}=\frac{4}{[1-e^{-2\gamma}]}\left[\coth(\gamma)-\frac{1}{\gamma}\right] \tag{35b}$$

It is important to note that α is a signed parameter. The sign of α is the same as that of the Peclet number γ, which in turn is related to the sign of the velocity u. In addition, for the domain interior, α is negative for positive values of γ!

DOI: 10.1002/fld

5. DISCUSSION

It can be seen from the DDR plots that every discrete model and also the semi-discrete model of the continuous problem diverge from the exact dispersion relation beyond a certain wave number, here ξ_{d}^* (Figures 1, 3, 5 and 7). For the semi-discrete case, ξ_{d}^* is marked in the plots and for the fully discrete case ξ_{d}^* is a contour line given by the value -3. ξ_{d}^* is greater when we use a consistent mass matrix for the transient terms in the formulation. Thus, one should expect better phase fidelity over a wider range of wave numbers using a consistent mass matrix. The gain in the value of ξ_{d}^* from the lumped **T** case to the consistent **T** case gradually decreases as the Courant number C increases (except for a certain range of the Courant number C for the FIC/SUPG method).

We now examine the differences in the DDR plots between the Galerkin method and the DDR plots of the FIC/SUPG and FIC_RC/OSS methods. It is interesting that the stabilization terms introduced by the FIC_RC/OSS method do not alter much the location of the phase departure wave number ξ_{d}^*. On the other hand, the stabilization terms introduced by the FIC/SUPG method contribute to the **T** matrix (Table I). Although for higher Courant number C the effect on the location of ξ_{d}^* is negligible, significant alterations in ξ_{d}^* are found for a lower C and a consistent **T** matrix (Figures 3, 5 and 7). When **T** is lumped, all the methods give similar patterns for ξ_{d}^*, indicating that stabilization terms have no role to play in the improvement of the DDRs.

Next, we examine the effect of the variation in C on the $\Re(\omega_h^*)$ vs ξ^* relation for the FIC and SUPG methods with consistent **T** matrix and the Crank–Nicholson scheme. Again, we choose ξ_{d}^* using the criterion $\forall \xi^* \leqslant \xi_{\mathrm{d}}^*$ we have $|\Re(\omega^* - \omega_h^*)| \leqslant 0.001$. In Figure 5, this corresponds to the contour line for the value -3. It can been seen in the semi-discrete case that $\Re(\omega_h^*) > \Re(\omega^*)$ for lower wave numbers (Figure 1). In addition, the error $\Re(\omega_h^* - \omega^*)$ first increases and later decreases. This behavior is exhibited in the fully discrete case too. Thus, there will be multiple contours of the same value in the $\log_{10}(|\omega_h^* - \omega^*|)$ vs (ξ^*, C) contour plots. In this case, ξ_{d}^* is the smallest ξ^* on the contours. For instance, choosing $C = 0.4$ for the FIC and SUPG methods we find $\xi_{\mathrm{d}}^* \approx 0.35$, where as for the Galerkin, FIC_RC and OSS methods we find $\xi_{\mathrm{d}}^* < 0.2$ (Figure 5). Thus, there is a significant improvement in the DDR for the FIC and SUPG methods for this value of C. It is interesting to note that for $C \leqslant 0.1$, the variation of DDR with C is insignificant. A similar conclusion can be made for the backward Euler and BDF2 schemes from their respective DDR plots.

It can be seen from the amplification plots (Figures 2, 4, 6 and 8) that *for the same value of* α (here $\alpha = 1.0$) and using a consistent **T** matrix, the damping associated with the FIC/SUPG method is relatively less than that of the FIC_RC/OSS method, though the gain is not significant for low wave numbers. On the other hand, using a lumped **T** matrix the difference in the amplification associated with FIC/SUPG and FIC_RC/OSS methods is insignificant. It can also be seen that unlike the notable differences in the location of ξ_{d}^*, the differences in the amplification due to lumping the **T** matrix are insignificant.

An important aspect is the effect of group velocity. This is more evident when the transported function is periodic and resembles a sinusoid wave train. For such functions, the Fourier transform is a narrow peak concentrated around the characteristic angular wave number of the function. For such problems, the effect of the group velocity is more significant than that of phase velocity. If the DDR predicts a deviation in the group velocity, then the wave train travels at that deviant velocity (Example 6.3).

Finally, we note that the optimal value of the stabilization parameter α on a uniform grid for the FIC_RC/OSS method is close to zero everywhere within the domain interior (Figure 10). The

DOI: 10.1002/fld

optimal α is negative for positive values of γ. We have also seen that due to a certain rearrangement of the equation stencils at the boundary, the expression for the optimal value of α is different for the boundary-adjacent elements (Figure 13). This is a favorable condition as we do not introduce any artificial (unnatural) damping anywhere within the domain, thus enhancing the accuracy of the numerical solution. The effect of boundary shock layers can also be captured efficiently by the new values of α at the boundary-adjacent elements. Thus, we see that in the transient mode the performance of the FIC_RC/OSS method using the proposed optimal α with boundary correction is similar to the standard Galerkin method. The former method also gives nodally exact solutions (on uniform grids) in the steady-state mode unlike the spurious global oscillations produced by the latter method. Of course, all these features are favorable addenda only if the bandwidth of the amplitude spectra of the transported function lies within the range of the phase departure wave number.

6. NUMERICAL EXAMPLES

6.1. Example 1

In this example, we study the effect of the stabilization introduced by the FIC/SUPG and FIC_RC/OSS methods for transient problems. We study the pure convection problem primarily for two reasons. First, the problem becomes simple as the dispersion effects of natural diffusion are sidelined, thus allowing us to study the effects of the diffusion introduced by the stabilization methods. Next, the convection-dominated problem is the primary concern of stabilization methods. The domain of interest is $x \in [0, 10]$. The problem data are $k = 1 \times 10^{-30}$, $u = 1.0$, time increment $\theta = 0.01$ and the 1D space is discretized by 100 linear elements with uniform mesh size. Thus, $\ell = 0.1$. The Peclet number for this problem is $\gamma = 5 \times 10^{28}$ ($\gamma \approx \infty$). The Courant number is $C = 0.1$. We have also chosen $\alpha = 1.0$, the optimal value for SUPG in this case, throughout the simulations. This choice is made just to study the effect of the artificial diffusion introduced by the FIC_RC/OSS method within the interior of the domain using a non-optimal value for α. Recall that the optimal value of α for the FIC_RC/OSS method in this case is ≈ 0, thus behaving very similar to the standard Galerkin method in the domain interiors. The discretization in time is done by the following schemes: Crank–Nicholson, backward Euler, and BDF2 schemes.

Figure 14 shows the results obtained when a narrow Gaussian pulse centered at $x = 3.0$ is taken as the initial solution. The equation for the initial solution is $\phi_0(x) = \exp[-8(x-3)^2]$. The numerical solution at time 3 s is examined. The solutions obtained using a lumped and consistent **T** matrix in the formulation are also examined. The idea is to validate the conclusions that can be drawn from the DDR plots. The pulse width of the Gaussian function is chosen so as to guarantee that the bandwidth of the amplitude spectra of this function is just less than $\xi_{\mathrm{d}}^{\mathrm{consistent}}$, the phase departure wave number using a consistent **T** matrix. As $\xi_{\mathrm{d}}^{\mathrm{lumped}} \leqslant \xi_{\mathrm{d}}^{\mathrm{consistent}}$, one should expect incorrect superpositions of the wave trains due to their phase differences from the former. This leads to a train of wiggles as seen in the numerical solution when **T** is lumped (Figure 14).

It is interesting that the standard Galerkin method without any stabilization and by using a consistent **T** matrix yields very accurate results. The other methods using a consistent **T** create a slight bump at the foot of the Gaussian bell. This is because of the damping associated with those methods for the significant wave numbers ($\xi^* \leqslant \xi_{\mathrm{d}}^*$). We also notice that using a consistent **T**, the damping associated with the FIC/SUPG method is less than that for the FIC_RC/OSS method. The differences are insignificant for the **T**-lumped case.

DOI: 10.1002/fld

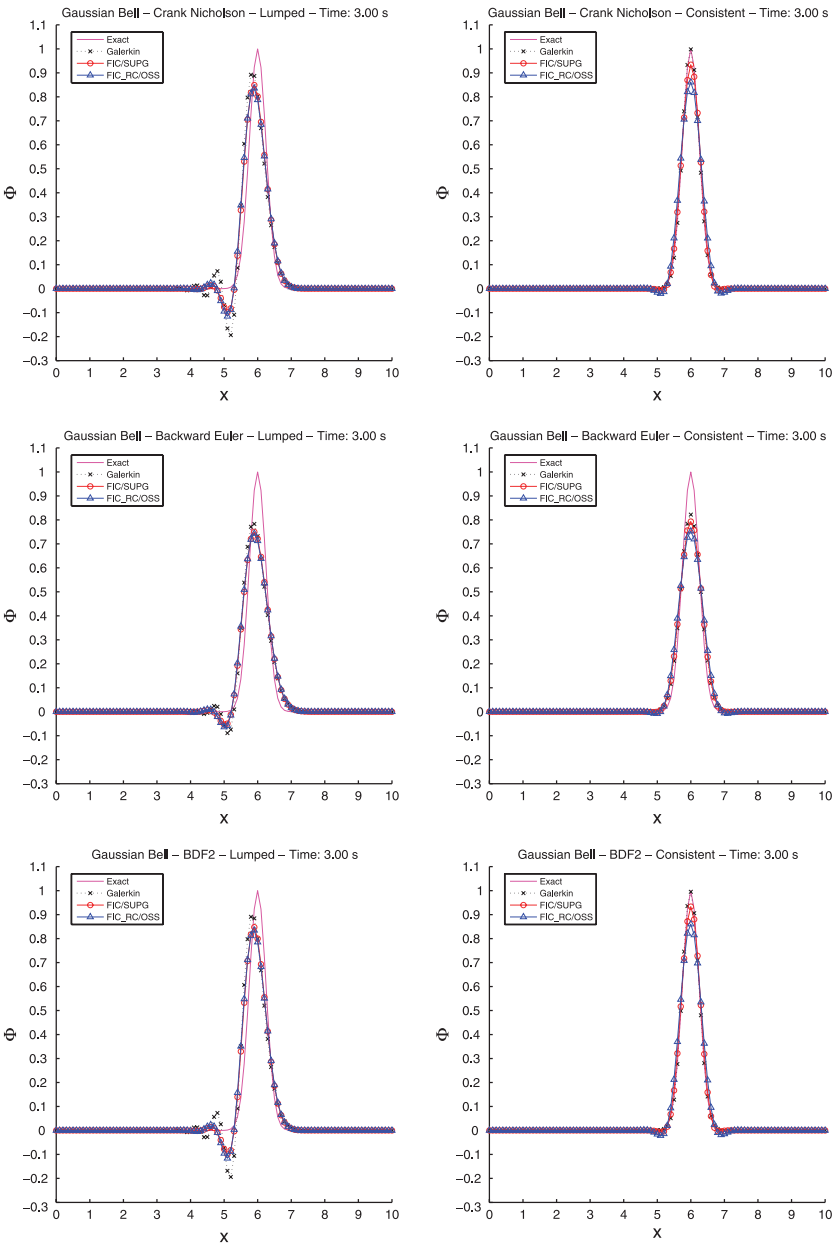

Figure 14. Gaussian pulse: solution at 3 s for the Galerkin, FIC/SUPG and FIC_RC/OSS methods using a lumped (on left) and consistent (on right) **T** matrices; time discretization is performed by Crank–Nicholson, backward Euler and BDF2 schemes. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

6.2. Example 2

In this example, we illustrate the effect of the variation of the Courant number on the DDR. The problem data are the same as in Example 6.1. Time integration is performed by the Crank–Nicholson scheme. Two initial functions are considered: a narrow Gaussian pulse centered at $x=3.0$ as defined in Example 6.1 and a square pulse function defined by $\phi_0(x)=1.0$ if $x\in[2,4]$ else $\phi_0(x)=0.0$. The spectra of the square pulse is broad and the bandwidth extends beyond the ξ_d of all the methods considered here. In other words, in the absence of damping this function will exhibit numerical dispersion. In this example, only the consistent **T** matrix is used.

First, we note that for the higher Courant number (here $C=1.0$), the Gaussian pulse exhibits numerical dispersion even when a consistent **T** matrix is used (Figure 15). As C is reduced to 0.4 and 0.1, we notice that the dispersion errors are minimized. As discussed earlier in Section 5, the FIC/SUPG method with $C=0.4$ should exhibit a better performance over the FIC_RC/OSS method. Unfortunately, the gain in the DDR for the higher wave numbers does not materialize in the simulated results. The solutions for the FIC/SUPG and the FIC_RC/OSS methods are nearly the same for both the initial solutions (Figure 15). This is because all those wave numbers suffer high damping. As $\Im(\omega_h^*)$ does not vary with C (Figure 6), Figure 2 may be referred for the amplification.

6.3. Example 3

In this example we study the transport of a periodic sine wave with angular wave number $\xi_0=7.5$. The problem data are the same as in Example 6.1. The corresponding value of $\xi^*=0.238$. Time integration is performed by the Crank–Nicholson scheme. The idea here is to study the effect of the group velocity in the numerical simulation. The initial function and the boundary condition prescribed at the left boundary are as follows:

$$f(x,0)=\sin(\xi_0 x) \tag{36}$$

$$f(0,t)=\sin(u\xi_0 t) \tag{37}$$

The DDR predicts a group velocity $V_g^{\text{lumped}}\approx 0.75$ and $V_g^{\text{consistent}}\approx 1.0$ for $\xi^*=0.238$ using lumped and consistent **T** matrices, respectively (Figure 9). This property is exhibited in the numerical solution where we find that the wave train moves at a different speed from the one assigned. The results are accurate using a consistent **T** matrix (Figure 16). The damping introduced by the stabilization methods is also in agreement with the one predicted by the DDR of the problem. Using a consistent **T** matrix, the amplifications for the FIC/SUPG and the FIC_RC/OSS methods for $\alpha=1.0$ and the wave number $\xi_0=7.5$ are ≈ 0.7 and 0.3, respectively, after 3 s (Figure 2). The corresponding values for the **T**-lumped case are ≈ 0.4 and 0.32. The numerical results are in agreement with this prediction. We note that the numerical damping associated with the FIC/SUPG method is less than that of the FIC_RC/OSS method for the same value of α. An interesting result is that when expressions (29) and (35) for α are used for the FIC_RC/OSS method, no damping takes place as it behaves similar to the Galerkin method in the interior domain. The numerical solution in this case coincides with that shown for the Galerkin method.

6.4. Example 4

In this example we explore the performance of the stabilization parameter α given by Equations (29) and (35) for the steady-state problem using the FIC_RC/OSS method. The domain of interest

DOI: 10.1002/fld

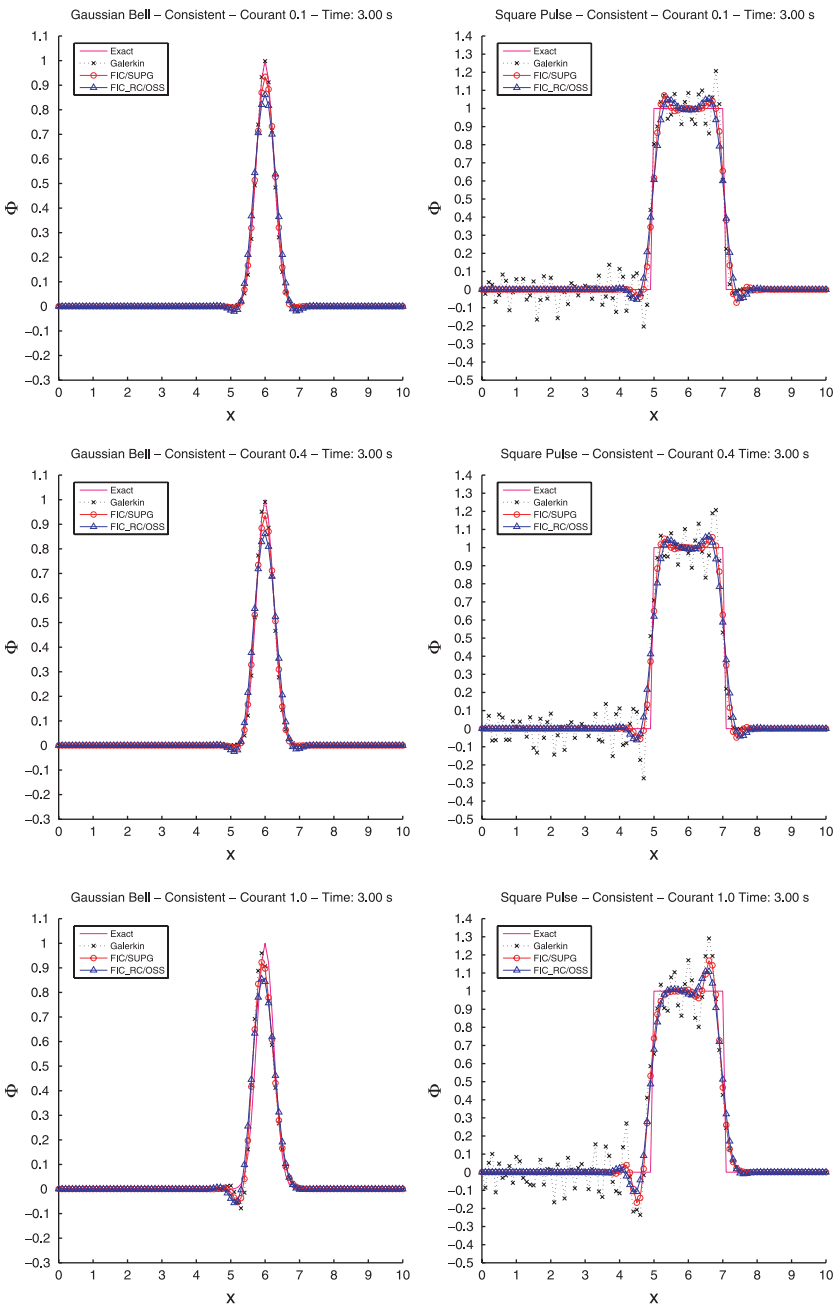

Figure 15. Solution at 3 s for the Galerkin, FIC/SUPG and FIC_RC/OSS methods using a consistent **T** matrix and for Courant numbers 0.1, 0.4 and 1.0; time discretization performed by Crank–Nicholson scheme. $\alpha = 1.0$ is used.

DOI: 10.1002/fld

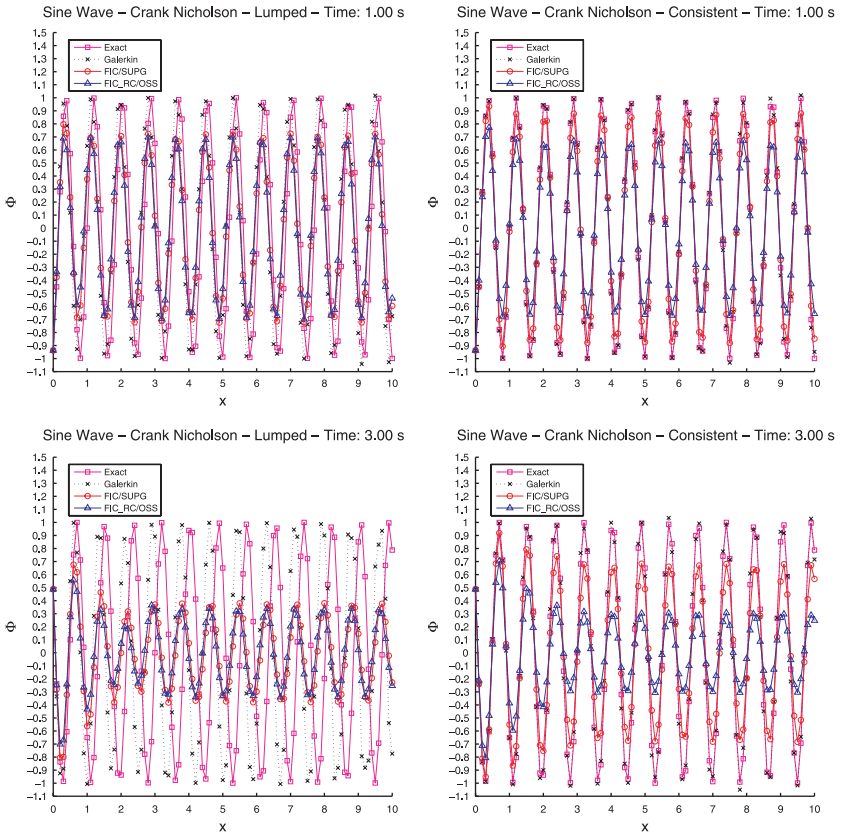

Figure 16. Sine wave: solutions at 1 and 3 s for the Galerkin, FIC/SUPG and FIC_RC/OSS methods using a lumped (on left) and consistent (on right) **T** matrices; time discretization performed by Crank–Nicholson scheme. $\alpha = 1.0$ is used.

is $x \in [0.0, 1.0]$. The problem data are $k = 0.001$, $u = 1.0$ and $f = 1.0$. For the ease of notation and further reference, we define the following:

$$\alpha_a = \left[\coth(\gamma) - \frac{1}{\gamma} \right]$$

$$\alpha_b = \begin{cases} \dfrac{4}{[1 - e^{2\gamma}]} \left[\coth(\gamma) - \dfrac{1}{\gamma} \right] & \text{element 1} \\ \dfrac{4}{[1 - e^{-2\gamma}]} \left[\coth(\gamma) - \dfrac{1}{\gamma} \right] & \text{element } n \\ \dfrac{-1}{\sinh^2(\gamma)} \left[\coth(\gamma) - \dfrac{1}{\gamma} \right] & \text{else} \end{cases}$$

First, we study the solution on a uniform mesh consisting of 20 linear elements ($\ell = 0.05$). We consider the cases when $f = 0$ and 1.0. The numerical solutions of the FIC/SUPG method using the stabilization parameter as α_a and of the FIC_RC/OSS method using α_a and α_b are presented

DOI: 10.1002/fld

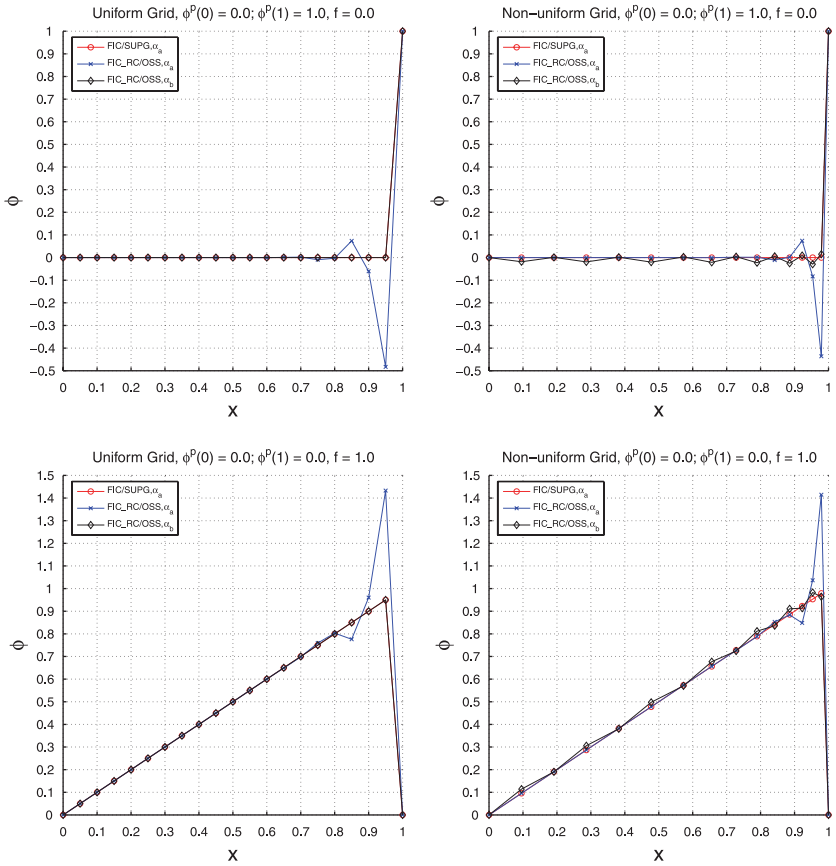

Figure 17. Steady-state solution on uniform and non-uniform grids for the FIC/SUPG and FIC_RC/OSS methods. $\alpha_a := [\coth(\gamma) - \gamma^{-1}]$. α_b is evaluated using Equations (29) and (35).

in Figure 17. We note that the new definition for the stabilization parameter α_b is optimal for the uniform mesh. The boundary correction introduced in Equation (35) also takes effect.

Next, we study the solution on a non-uniform grid consisting of 15 elements. The node coordinates of the discrete 1D space are given by $x = \{0, 0.095, 0.191, 0.2866, 0.382, 0.477, 0.573, 0.656, 0.728, 0.789, 0.841, 0.885, 0.922, 0.953, 0.979, 1\}$. We note that the solution of the FIC/SUPG method is superior to that obtained by the FIC_RC/OSS method. The latter method gives the sharp boundary oscillations using the parameter α_a (also appears on uniform grids). The solution of the FIC_RC/OSS method using α_b on a non-uniform grid is slightly corrupted with weak node-to-node spurious oscillations.

7. CONCLUSION

A detailed transient analysis of a consistency recovery method (FIC_RC/OSS) with respect to the Galerkin and FIC/SUPG methods has been done. The discrete dispersion relations for the

DOI: 10.1002/fld

above methods using the trapezoidal and BDF2 time integration schemes are presented. The phase departure wave number (ξ_d) is greater when a consistent matrix for the transient terms is used. ξ_d proportionally influences the range of wave numbers that have a group velocity close to the convection velocity. The DDR plots predict that the gain in the value of ξ_d from the lumped **T** case to the consistent **T** case gradually decreases with the Courant number (C). An exception to this is on a certain lower range of C for the FIC/SUPG method. Unfortunately, this gain is seldom realized as those higher wave numbers are damped away. The contour plots of $\log[\Re(|\omega_h^* - \omega^*|)]$ are very similar for the Galerkin, FIC/SUPG and the FIC_RC/OSS methods except for the exception mentioned earlier. Neither is the change significant with the choice of the time integration schemes considered. This suggests that the role of the stabilization terms in the improvement of the DDR is insignificant and the enhancement can be achieved only by virtue of the resolution in space and time. It is shown that for the same value of the stabilization parameter (α), the damping associated with the FIC_RC/OSS method is slightly greater than that of the FIC/SUPG method.

It is shown that, unlike the FIC/SUPG method, the FIC_RC/OSS method introduces a certain rearrangement in the equation stencils at nodes on and adjacent to the domain boundary. Thus, using a uniform expression for α will lead to localized oscillations at the boundary. These oscillations are notable in the convection-dominated case. When diffusion is at par with convection, the solution has a smooth profile and these local oscillations are insignificant even though they exist. A proposal for α that gives optimal results for the steady-state 1D convection–diffusion problem on uniform grids is made for the FIC_RC/OSS method. An interesting result is that when the new expressions for α are used for the FIC_RC/OSS method, no damping takes place as $\alpha \approx 0$ in the domain interior. The numerical solution in the transient case coincides with that for the Galerkin method and in the steady state, unlike the Galerkin method, it is stable. Unfortunately, using this new expression for α dispersion errors, whenever present (for instance, the transport of a square pulse), cannot be controlled and also the steady-state solution has weak node-to-node oscillations on a non-uniform grid. On the basis of these results, it appears that with respect to the stabilization of convection, the numerical performance of the FIC/SUPG method is better as one can control to some extent the dispersive errors and at the same time we can assure the stability of the steady-state solution. Ongoing research proceeds in the direction of finding an expression for α for the FIC_RC/OSS method ensuring a non-oscillatory solution for irregular meshes. In future work we also plan to study the effect of using an alternative definition for π involving the diffusive and transient terms.

ACKNOWLEDGEMENTS

The first author thanks Prof. Ramon Codina for many useful discussions. This study was partially supported by the SEDUREC project of the Consolider Ingenio 2010 Programme, Ministry for Science and Education, Spain.

REFERENCES

1. Codina R. Comparison of some finite element methods for solving the diffusion–convection–reaction equation. *Computer Methods in Applied Mechanics and Engineering* 1998; **156**:185–210.
2. Oñate E, Garcia J, Idelsohn S. Computation of the stabilization parameter for the finite element solution of advective–diffusive problems. *International Journal for Numerical Methods in Fluids* 1997; **25**(12):1385–1407.
3. Oñate E. Derivation of the stabilization equations for advective–diffusive fluid transport and fluid flow problems. *Computer Methods in Applied Mechanics and Engineering* 1998; **151**(1–2):233–267.
4. Oñate E, Manzan M. A general procedure for deriving stabilized space–time finite element methods for advective–diffusive problems. *International Journal for Numerical Methods in Engineering* 1999; **31**:203–207.

DOI: 10.1002/fld

5. Oñate E. Possibilities of finite calculus in computational mechanics. *International Journal for Numerical Methods in Engineering* 2004; **60**:255–281.
6. Oñate E, Taylor RL, Zienkiewicz OC, Rojek J. A residual correction method based on finite calculus. *Engineering Computations* 2003; **20**(5/6):629–658.
7. Codina R. Stabilization of incompressibility and convection through orthogonal sub-scales in finite element methods. *Computer Methods in Applied Mechanics and Engineering* 2000; **190**:1579–1599.
8. Jansen KE, Collis SS, Whiting C, Shakib F. A better consistency for low-order stabilized finite element methods. *Computer Methods in Applied Mechanics and Engineering* 1999; **174**:153–170.
9. Gresho PM, Sani RL. The advection–diffusion equation. *Incompressible Flow and Finite Element Method*, *Vol. 1*: *Advection–Diffusion and Isothermal Laminar Flow*. Wiley: Chichester, 2000.
10. Christon MA, Martinez MJ, Voth TE. Generalized Fourier analyses of the advection–diffusion equation—Part I: one-dimensional domains. *International Journal for Numerical Methods in Fluids* 2004; **45**:839–887.
11. Trefethen LN. Group velocities in finite difference schemes. *SIAM Review* 1982; **24**(2):113–136.
12. Davis S. A space–time discretization procedure for wave propagation problems. *Technical Memorandum*, *Report No. NASA TM-102215*, Ames Research Center, NASA, 1989.
13. Grimshaw R. Group velocity. In *Encyclopedia of Nonlinear Science*, Scott AC (ed.). Taylor & Francis: New York, 2004; 385–388.
14. Oñate E, Zarate F, Idelsohn S. Finite element formulation for convective–diffusive problems with sharp gradients using finite calculus. *Computer Methods in Applied Mechanics and Engineering* 2006; **195**:1793–1825.
15. Sengupta TK, Ganeriwal G, De S. Analysis of central and upwind compact schemes. *Journal of Computational Physics* 2003; **192**:677–694.
16. Sengupta TK, Ganerwal G, Dipankar A. High accuracy compact schemes and Gibbs' phenomenon. *Journal of Scientific Computing* 2004; **21**(3):253–268.

DOI: 10.1002/fld