Available online at www.sciencedirect.com
SCIENCE DIRECT®
Comput. Methods Appl. Mech. Engrg. 193 (2004) 2301–2323
Computer methods in applied mechanics and engineering
www.elsevier.com/locate/cma

Stability of the SUPG finite element method for transient advection–diffusion problems

Pavel B. Bochev [a,*,1,2], Max D. Gunzburger [b,3], John N. Shadid [c,1,2]

[a] Sandia National Laboratories, Computational Mathematics and Algorithms, P.O. Box 5800, MS 1110, Albuquerque, NM 87185-1110, USA
[b] School of Computational Science and Information Technology, Florida State University, Tallahassee, FL 32306-4120, USA
[c] Sandia National Laboratories, Computational Sciences, P.O. Box 5800, MS 1111, Albuquerque, NM 87185-1110, USA

Received 17 October 2003; accepted 7 January 2004

Abstract

Implicit time integration coupled with SUPG discretization in space leads to additional terms that provide consistency and improve the phase accuracy for convection dominated flows. Recently, it has been suggested that for small Courant numbers these terms may dominate the streamline diffusion term, ostensibly causing destabilization of the SUPG method. While consistent with a straightforward finite element stability analysis, this contention is not supported by computational experiments and contradicts earlier Von-Neumann stability analyses of the semidiscrete SUPG equations. This prompts us to re-examine finite element stability of the fully discrete SUPG equations. A careful analysis of the additional terms reveals that, regardless of the time step size, they are always dominated by the consistent mass matrix. Consequently, SUPG cannot be destabilized for small Courant numbers. Numerical results that illustrate our conclusions are reported.

Keywords: Advection–diffusion problems; Stabilized finite element methods; Petrov–Galerkin methods; Generalized trapezoidal rule

1. Introduction

Consider the steady-state scalar advection–diffusion problem

$$-\epsilon\nabla^2\phi + \mathbf{b}\cdot\nabla\phi = f \quad \text{in } \Omega \quad \text{and} \quad \phi = g \quad \text{on } \Gamma, \tag{1}$$

* Corresponding author. Tel.: +1-505-844-1990; fax: +1-505-845-7442.
E-mail addresses: pbboche@sandia.gov (P.B. Bochev), gunzburg@csit.fsu.edu (M.D. Gunzburger), jnshadi@sandia.gov (J.N. Shadid).

[1] Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed-Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC-94AL85000.
[2] This work was partially funded by the Applied Mathematical Sciences program, US Department of Energy, Office of Energy Research.
[3] Supported in part by CSRI, Sandia National Laboratories, under contract 18407.

0045-7825/$ - see front matter © 2004 Elsevier B.V. All rights reserved.
doi:10.1016/j.cma.2004.01.026

where Ω is a bounded open domain in $\mathbf{R}^n$, $n = 1, 2, 3$ with Lipschitz continuous boundary Γ, $\mathbf{b}(\mathbf{x})$ is a given velocity field with $\nabla \cdot \mathbf{b} = 0$, and $\epsilon \geqslant 0$ is a constant diffusion coefficient. When $\epsilon = 0$ boundary conditions are specified only on the inflow part $\Gamma_- = \{\mathbf{x} \in \Gamma | \mathbf{b} \cdot \mathbf{n} < 0\}$ of Γ.

If $\epsilon = 0$, or (1) is advection dominated, Galerkin finite element solutions of this problem develop spurious oscillations unless the exact solution happens to be globally smooth. A popular and efficient remedy is to augment the Galerkin form of (1) by terms that add artificial dissipation but vanish for all sufficiently smooth solutions. Resulting schemes are called *consistently* stabilized methods because the order of the Galerkin approximation is not affected. A consistently stabilized method can be written as

$$G(\phi^h, \psi^h) + \langle R(\phi^h), W(\psi^h) \rangle_h = (f, \psi^h), \tag{2}$$

where $G(\cdot, \cdot)$ is Galerkin form of (1), $R(\phi^h)$ is the residual of (1), $W(\psi^h)$ is weighting operator, and $\langle \cdot, \cdot \rangle_h$ is a broken L^2 inner product defined with respect to a partition $\mathscr{T}_h$ of Ω into finite elements. Of particular interest in this paper is the *streamline upwind* weighting operator

$$W_{\mathrm{SUPG}}(\psi^h) = \mathbf{b} \cdot \nabla \psi^h \tag{3}$$

and the associated SUPG method [13]. Two other possible choices for the weighting function are the Galerkin least-squares operator

$$W_{\mathrm{GLS}}(\psi^h) = -\epsilon \nabla^2 \psi^h + \mathbf{b} \cdot \nabla \psi^h, \tag{4}$$

leading to the GLS method of [15], and the multiscale operator

$$W_{\mathrm{MS}}(\psi^h) = -(-\epsilon \nabla^2 \psi^h - \mathbf{b} \cdot \nabla \psi^h) = +\epsilon \nabla^2 \psi^h + \mathbf{b} \cdot \nabla \psi^h. \tag{5}$$

This operator is obtained from the variational multiscale method [11] and leads to a method originally referred to as the *adjoint* or the *unusual* stabilized Galerkin; see [5,6]. All three stabilization operators are widely used for steady state problems where their properties are well-documented and understood; see e.g. [5–7,12–13,15–16,19].

Consider now the time dependent version of (1)

$$\begin{aligned} &\phi_t - \epsilon \nabla^2 \phi + \mathbf{b} \cdot \nabla \phi = f \quad \text{in } \Omega; \quad \phi = g \quad \text{on } \Gamma, \\ &\phi(0, \mathbf{x}) = \phi_0(\mathbf{x}) \quad \text{in } \Omega, \end{aligned} \tag{6}$$

where ϕ_0 is a given initial data. It is generally agreed that time-space elements are the most natural setting to develop stabilized methods for (6); see e.g. [19,22] or [21]. Already in 1984, Johnson et al. [19] argue that the time derivative and the advective term should be combined into a single "material derivative", so that a natural extension of (3)–(6) is the *time-space* SUPG weighting operator

$$W_{\mathrm{SUPG}}(\psi^h) = \frac{D\psi^h}{Dt} = \dot{\psi}^h + \mathbf{b} \cdot \nabla \psi^h.$$

More recently, Hughes et al. [11,14] demonstrated that stabilized methods for stationary problems can be derived via a variational multiscale framework wherein the solution space is split into resolved and unresolved scales, followed by a defect correction step driven by the residual equation. According to this viewpoint, which has been extended to time-space in [17], stabilization terms originate from approximation of the solution operator (the Green's function) of the defect equation. Therefore, if the problem is time dependent, consistent application of variational multiscale stabilization calls for time-space elements.

Nevertheless, some of the most effective and popular algorithms for treating time-dependent problems can be defined through a process wherein the spatial and temporal discretizations are separated. Such algorithms are especially well adapted to the cylindrical nature of the time-space domain and they reduce (6) to a system of ordinary differential equations (ODE's) that can be solved by many of the available time

integration methods for ODE's. As a result, these algorithms allow reuse of existing spatial finite element frameworks and deploy a time dependent solution method without significant development of new software. Thus, in practice, for several reasons, implicit, fully discrete formulations in which spatial and temporal discretizations are effected separately are in much more common use than are coupled time-space formulations. Additionally, for a large number of computational applications the increased cost in the number of unknowns for coupled time-space formulations is a significant drawback.

As numerical experiences have borne out, separated, fully discrete algorithms are completely adequate for transient calculations carried out for moderate to relatively large time steps. However, in settings that require very fine time resolution, the behavior of such algorithms is not very well understood. Recently, Harari [8,9] demonstrated that for small time steps the implicit time integration of parabolic problems leads to a singularly perturbed elliptic problem with an onset of local spurious oscillations in the vicinity of thin physical layers. Because for small time steps the fully discrete equation can be viewed as discretization of an elliptic boundary value problem with a dominant reaction term, the remedy suggested in [8] is to apply adjoint stabilization to this spatial problem. This is analogous to the approach of [5] but differs from the gradient Galerkin least squares (GGLS) stabilization advocated in [4] and [18].

The main focus of this paper is, however, on another potential source of instability that occurs when implicit time integration is coupled with spatial stabilization. This situation arises whenever, in the development of stabilization methods for (6), one foregoes the time-space setting in favor of the more conventional separated finite difference/finite element approach. After discretization in space one obtains the semidiscrete equation

$$(\phi_t^h, \psi^h) + G(\phi^h, \psi^h) + \left\langle \widetilde{R}(\phi^h), W(\psi^h) \right\rangle_h = (f, \psi^h), \tag{7}$$

where ψ^h varies only spatially and $\widetilde{R}(\phi^h)$ contains the time derivative ϕ_t^h. We can rewrite (7) as

$$(\phi_t^h, \psi^h) + \left\langle \phi_t^h, W(\psi^h) \right\rangle_h + G(\phi^h, \psi^h) + \left\langle R(\phi^h), W(\psi^h) \right\rangle_h = (f, \psi^h), \tag{8}$$

from where it is clear that the fully discrete equation will be a weighted average of a *spatially stabilized Galerkin form* for the steady-state problem (1) and a *modified mass matrix*. The additional "mass" term is contributed by the time derivative in the residual of (6) and is needed for phase consistency. In a recent paper, Bradford and Katopodes [1] argue that in conjunction with Crank–Nicolson implicit time integration this term may have an antidissipative and destabilizing effect for small time steps. Their finite difference analysis leads to a sufficient stability condition that requires Courant numbers greater than one and imposes a lower bound on the admissible time steps. In the next section we introduce the fully discrete equations, review the arguments of [1], repeat some of their numerical experiments, and show that a straightforward finite element stability analysis will lead to essentially the same *sufficient* stability condition for the finite element method. However, our numerical tests fail to excite a true destabilization in the Petrov–Galerkin method, thus raising questions about the sharpness of the stability estimates. Motivated by this discrepancy between analysis and numerical experiments we pursue a more careful stability analysis of this problem.

2. Fully discrete spatially stabilized equations

Let $\mathscr{T}_h$ denote a uniformly regular partition of Ω into finite elements $\mathscr{K}$. We consider affine families of Lagrangian finite element spaces S_d^h where d stands for the polynomial degree. To discretize (6) in space we use the subspace $S_{d,g}^h$ of S_d^h constrained by the essential boundary condition in (6). Approximation of ϕ is sought in the form

$$\phi^h(\mathbf{x},t)=\sum_{i=1}^{N}\alpha_i(t)N_i(\mathbf{x}),$$

where N_i denotes the standard nodal basis of S_d^h.

Let $S_{d,0}^h$ denote the subspace of S_d^h consisting of functions that vanish on Γ (or Γ_- if $\epsilon=0$). The spatially stabilized semidiscrete variational problem is to seek $\phi^h(\mathbf{x},t)\in S_{d,g}^h\times T$ such that

$$M\big(\phi_t^h(\cdot,t),\psi^h\big)+G_S\big(\phi^h(\cdot,t),\psi^h\big)=\big(f(\cdot,t),\psi^h\big)\quad\forall\psi^h\in S_{d,0}^h;\ \ t\in T, \tag{9}$$

where

$$M\big(\phi_t^h(\cdot,t),\psi^h\big)=\big(\phi_t^h(\cdot,t),\psi^h\big)+\sum_{\mathscr{K}\in\mathscr{T}_h}\big(\phi_t^h(\cdot,t),\tau\big(\sigma\epsilon\nabla^2\psi^h+\mathbf{b}\cdot\nabla\psi^h\big)\big)_{0,\mathscr{K}},$$

is an augmented inertial form, and

$$\begin{aligned}G_S\big(\phi^h(\cdot,t),\psi^h\big)&=\big(\epsilon\nabla\phi^h(\cdot,t),\nabla\psi^h\big)+\big(\mathbf{b}\cdot\nabla\phi^h(\cdot,t),\psi^h\big)\\&\quad+\sum_{\mathscr{K}\in\mathscr{T}_h}\big(-\epsilon\triangle\phi^h(\cdot,t)+\mathbf{b}\cdot\nabla\phi^h(\cdot,t),\tau\big(\sigma\epsilon\nabla^2\psi^h+\mathbf{b}\cdot\nabla\psi^h\big)\big)_{0,\mathscr{K}},\end{aligned}$$

is a spatially stabilized Galerkin form. In this formulation τ is the *stability parameter*, and σ takes on the integer values 0, 1 and −1, corresponding to SUPG, MS and GLS, respectively.

In what follows we restrict attention to SUPG spatial stabilization ($\sigma=0$) and use a definition of τ developed in [6]. For the purpose of our study it suffices to consider only advection dominated problems. Therefore, we assume that ϵ, $\mathbf{b}$ and the grid $\mathscr{T}_h$ are such that

$$Pe_{\mathscr{K}}>3, \tag{10}$$

where

$$Pe_{\mathscr{K}}(\mathbf{x})=\frac{m\|\mathbf{b}(\mathbf{x})\|_p h_{\mathscr{K}}}{2\epsilon},$$

is the element *Peclet* number and m is a parameter whose value depends on the inverse constant [4] for $\mathscr{T}_h$. In this case,

$$\tau(\mathbf{x})=\frac{h_{\mathscr{K}}}{2\|\mathbf{b}(\mathbf{x})\|_p} \tag{11}$$

and if $\mathscr{T}_h$ is regular, one can show that

$$\check{\tau}h\leqslant\tau(\mathbf{x})\leqslant\hat{\tau}h\quad\forall\mathscr{K}\in\mathscr{T}_h \tag{12}$$

for some positive constants $\check{\tau}$ and $\hat{\tau}$. In what follows we set $p=2$ in (11).

The semidiscrete equation (9) is a system of ODE's

$$\mathbf{M}\boldsymbol{\alpha}_t(t)+\mathbf{K}\boldsymbol{\alpha}(t)=\mathbf{f}(t)$$

for the unknown coefficient vector $\boldsymbol{\alpha}(t)=(\alpha_1(t),\ldots,\alpha_N(t))$. The matrices $\mathbf{M}$ and $\mathbf{K}$ are generated in the usual manner from the bilinear forms $M(\cdot,\cdot)$ and $G_S(\cdot,\cdot)$, respectively and $\mathbf{f}$ is a vector whose components are L^2 products of the source term and the nodal shape functions N_i. This system may be solved by any of the available ODE solvers. In this paper we use the θ-method, also known as the *generalized trapezoidal rule*. To discretize in time, the interval $(0,T)$ is subdivided into L subintervals $[t_k,t_{k+1}]$, $k=0,\ldots,L$ with

[4] Sharp estimates for the inverse constant and other important constants can be found in [10].

lengths $\Delta^k t$. Throughout, $f_k = f(t_k)$, and ϕ_k^h, $\boldsymbol{\alpha}_k$ denote approximations to $\phi^h(\mathbf{x}, t_k)$ and $\boldsymbol{\alpha}(t_k)$, respectively. Given ϕ_0^h, ϕ_{k+1}^h for $k = 0, 1, \ldots, L-1$ are determined from the equation

$$\frac{1}{\Delta^k t} M\left(\phi_{k+1}^h - \phi_k^h, \psi^h\right) + G_S(\phi_{\theta,k}^h, \psi^h) = \left(f_{\theta,k}, \psi^h\right) \quad \forall \psi^h \in S_{d,0}^h, \tag{13}$$

where $0 \leqslant \theta \leqslant 1$ is a real parameter,

$$\phi_{\theta,k}^h = \theta \phi_{k+1}^h + (1-\theta)\phi_k^h$$

and likewise for $f_{\theta,k}$. The fully discrete problem (13) is a linear system of algebraic equations

$$\left(\mathbf{M} + \theta \Delta^k t \mathbf{K}\right)\boldsymbol{\alpha}_{k+1} = \mathbf{f}_{\theta,k} + \left(\mathbf{M} - (1-\theta)\Delta^k t \mathbf{K}\right)\boldsymbol{\alpha}_k. \tag{14}$$

For $\theta = 0$ the scheme (14) is the explicit Euler method, $\theta = 1/2$ gives the second-order neutrally stable Crank–Nicolson method, and $\theta = 1$ gives the first-order accurate implicit Euler rule. In what follows it will be convenient to introduce the bilinear form

$$B(\phi^h, \psi^h; \rho, \theta) = \rho M(\phi^h, \psi^h) + \theta G_S(\phi^h, \psi^h), \tag{15}$$

that is a weighted average of the inertial form and the spatially stabilized Galerkin form. This form engenders the problem that advances the discrete solution by one time step and will be in the focus of our stability analysis.

The following result holds true; see [5–7] and [19].

Theorem 1. *Assume that* $\nabla \cdot \mathbf{b} = 0$, $g = 0$ *on* Γ *and* $\epsilon \geqslant 0$. *Then, for the weighting operators in* (3)–(5)

$$G_S\left(\psi^h, \psi^h\right) \geqslant \tfrac{1}{2}\left(\epsilon \left\|\nabla \psi^h\right\|_0^2 + \left\|\tau^{1/2}\mathbf{b} \cdot \nabla \psi\right\|_0^2\right) \quad \forall \psi^h \in S_{d,0}^h. \tag{16}$$

2.1. Preliminary analysis

Consider (6) in 1D and assume that $\epsilon = 0$. The discrete equation (14) resulting from the combination of SUPG stabilization in space ($\sigma = 0$) and Crank–Nicolson implicit integration in time ($\theta = 1/2$) is viewed in [1] as a finite difference approximation of the modified equation

$$\phi_t + b\phi_x - \tau(x)\left(b^2\phi_{xx} + b\phi_{xt}\right) = f, \tag{17}$$

where the definition

$$\tau(x) = \frac{\triangle x}{|b|\sqrt{15}} \tag{18}$$

is employed. For 1D pure advection problems, this choice maximizes the phase accuracy in the semidiscrete equation [20].

The "streamline" derivative ϕ_{xx} is contributed by the SUPG stabilization while ϕ_{xt} results from the coupling between ϕ_t and the spatial weight function. Assume now that ϕ^h is a discontinuous pulse and let $\Delta\phi = \phi_R^h - \phi_L^h > 0$ denote its amplitude. The additional terms in (17) are estimated in [1] by

$$\phi_{xx} \approx \mathrm{CFL}\frac{\Delta\phi}{2(\Delta x)^2} \quad \text{and} \quad \phi_{xt} \approx -\mathrm{CFL}\frac{\Delta\phi}{2\Delta x \Delta t},$$

respectively, where

$$\mathrm{CFL} = b\frac{\Delta t}{\Delta x},$$

is the Courant number. The total modification in (17) is then estimated as

$$\tau(x)(b^2\phi_{xx} + b\phi_{xt}) \approx \tau(x)\frac{b^2\Delta\phi}{2(\Delta x)^2}(\mathrm{CFL} - 1)$$

and a conclusion is drawn that for CFL < 1 the term ϕ_{xt} will *dominate* the streamline derivative, causing *destabilization* of the Petrov–Galerkin formulation. To avoid the antidissipative effect of this term, it is suggested that the time step should satisfy the stability condition CFL > 1, or

$$\Delta t > \frac{\Delta x}{|b|}. \tag{19}$$

Next, we obtain a formal finite element stability estimate that leads to the same conclusion. This rather disturbing result ostensibly implies that for stability the CFL number should be greater than 1, however, for accuracy in following transient advection we desire to have CFL < 1.

Theorem 2. *Assume that* ϵ, $\mathbf{b}$ *and* $\mathscr{T}_h$ *are such that* (10) *holds. Then, for* $\sigma = 0$ (*SUPG spatial stabilization*)

$$B(\psi^h, \psi^h; \rho, \theta) \geqslant \frac{\rho}{2}\|\psi^h\|_0^2 + \theta\left(\frac{\epsilon}{2}\|\nabla\psi^h\|_0^2 + \frac{\check{\tau}h}{2}\left(1 - \frac{\rho(\hat{\tau}h)^2}{\theta(\check{\tau}h)}\right)\|\mathbf{b}\cdot\nabla\psi^h\|_0^2\right) \tag{20}$$

for all $\psi^h \in S^h_{d,0}$.

Proof. To prove the theorem we estimate the inertial term $M(\cdot,\cdot)$ and use the available bound (16) for the spatially stabilized component of $B(\cdot,\cdot;\rho,\theta)$. Successive use of Cauchy's inequality and the ε—inequality give

$$\begin{aligned}
M(\psi^h, \psi^h)c &= (\psi^h, \psi^h) + \sum_{\mathscr{K}\in\mathscr{T}_h}(\psi^h, \tau\mathbf{b}\cdot\nabla\psi^h)_{0,\mathscr{K}} \\
&\geqslant \|\psi^h\|_0^2 - \sum_{\mathscr{K}\in\mathscr{T}_h}\hat{\tau}h\|\psi^h\|_{0,\mathscr{K}}\|\mathbf{b}\cdot\nabla\psi^h\|_{0,\mathscr{K}} \\
&\geqslant \|\psi^h\|_0^2 - \frac{1}{2}\left(\sum_{\mathscr{K}\in\mathscr{T}_h}\|\psi^h\|_{0,\mathscr{K}}^2 + (\hat{\tau}h)^2\|\mathbf{b}\cdot\nabla\psi^h\|_{0,\mathscr{K}}^2\right) \\
&\geqslant \frac{1}{2}\left(\|\psi^h\|_0^2 - (\hat{\tau}h)^2\|\mathbf{b}\cdot\nabla\psi^h\|_0^2\right).
\end{aligned}$$

The theorem now easily follows by combining the last bound with (16). □

For a pure advection problem with constant advective velocity $\mathbf{b}$ and a uniform mesh, (11) implies [5] that

$$\tau(\mathbf{x}) = \frac{h}{2\|\mathbf{b}\|_2} = \mathrm{const} \quad \text{and} \quad \check{\tau} = \hat{\tau} = \frac{1}{2\|\mathbf{b}\|_2}.$$

[5] It is possible to extend (18) to multiple dimensions. For example, in [13] the formula $\tau = (\|\mathbf{b}_\xi\|_2 h_\xi + \|\mathbf{b}_\eta\|_2 h_\eta|)/(\|\mathbf{b}\|_2^2\sqrt{15})$ is proposed for quadrilateral elements. Using this formula in lieu of (11) would have changed $\check{\tau}$ and $\hat{\tau}$ to

$$\frac{1}{\|\mathbf{b}\|_2\sqrt{15}} \quad \text{and} \quad \frac{\sqrt{2}}{\|\mathbf{b}\|_2\sqrt{15}},$$

respectively, which is not essential to our discussion.

In this case (20) simplifies to

$$B(\psi^h, \psi^h; \rho, \theta) \geqslant \frac{\rho}{2}\|\psi^h\|_0^2 + \tau\frac{\theta}{2}\left(1 - \frac{\rho\tau}{\theta}\right)\|\mathbf{b}\cdot\nabla\psi^h\|_0^2.$$

For Crank–Nicolson $\theta = 1/2$, and since $\rho = 1/\triangle t$,

$$1 - \frac{\rho\tau}{\theta} = 1 - \frac{h}{\Delta t|\mathbf{b}|} = \frac{1}{\mathrm{CFL}}(\mathrm{CFL} - 1).$$

As a result, the streamline coefficient will be positive if CFL > 1, i.e., we have obtained the same stability condition as in [1]. Let us now check this stability condition against some numerical experiments.

Following [1] we set $\mathbf{b} = 0.001\ \mathrm{m\,s^{-1}}$, $\Delta x = 0.1$, $\Delta t = 1$ s, which makes CFL equal to 0.01. Then we compute solutions of the fully discrete equations with and without SUPG stabilization for different final times using Crank–Nicolson and two different sets of initial and boundary data. The first set

$$\phi_0(\mathbf{x}) = 0 \quad \text{and} \quad g = g(0) = 100, \tag{21}$$

is the same as the one used in [1]. However, for the second set we change the initial condition to a square pulse and set homogeneous data on the inflow:

$$\phi_0(\mathbf{x}) = \begin{cases} 100 & \text{if } 0.25 \leqslant \mathbf{x} \leqslant 0.5 \\ 0 & \text{otherwise} \end{cases} \quad \text{and} \quad g = g(0) = 0. \tag{22}$$

Plots of the Galerkin and SUPG solutions at $t = 100$ s and $t = 500$ s are shown in Figs. 1 and 2.

The left plot in Fig. 1 shows that at early times SUPG solution tends to develop stronger undershoot at the base of the advancing discontinuity. The right plot shows that in later times the undershoots of SUPG and Galerkin solutions are about the same. Nevertheless, in both cases the SUPG solution does not appear to be substantially better than the Galerkin one, which lends some credence to the possibility that the extra mass term is destabilizing.

However, the second set of plots presented in Fig. 2, shows that such a conclusion is unfounded and that each method behaves as advertised: the Galerkin solution quickly develops global spurious oscillations, while SUPG continues to successfully suppress these oscillations, even for very small Courant numbers.

To reconcile the stability criterion (19) with the numerical results shown in Figs. 1 and 2, it is important to recognize that the former is only a sufficient but not a *necessary* condition for stability. As such, (19) does imply stability when satisfied, but it does not imply *instability* when not satisfied. In fact, a sufficient condition may turn out to be too pessimistic. Let us show that this is indeed the case for the examples considered so far. The next Theorem sharpens the stability bound (20) for problems where τ can be set to the same mesh dependent constant throughout the domain Ω.

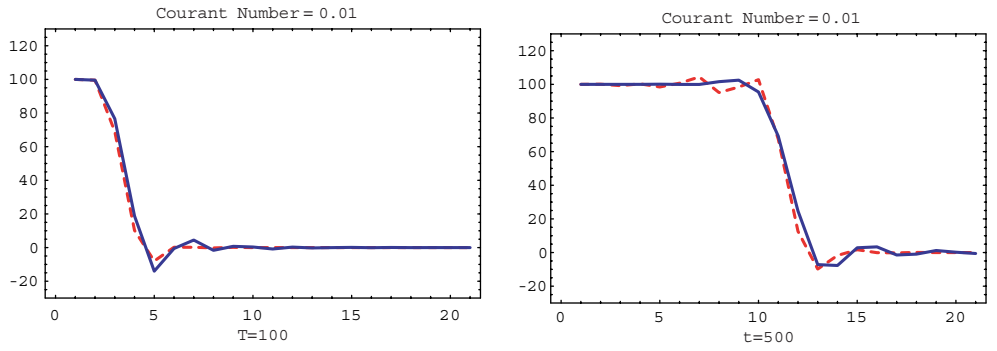

Fig. 1. Galerkin (dashed) vs. SUPG (solid) solutions for (21).

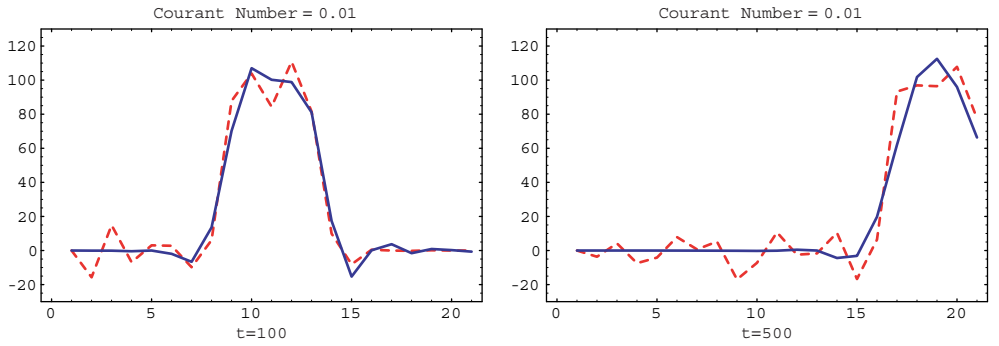

Fig. 2. Galerkin (dashed) vs. SUPG (solid) solutions for (22).

Theorem 3. *Assume that* ϵ, $\mathbf{b}$ *and* $\mathscr{T}_h$ *are such that*

$$\tau(\mathbf{x}) = \delta h \quad \forall \mathbf{x} \in \Omega$$

for some positive constant δ. *Then, for* $\sigma = 0$ (*SUPG spatial stabilization*)

$$B(\psi^h, \psi^h; \rho, \theta) \geqslant \frac{\rho}{2}\|\psi^h\|_0^2 + \theta\left(\frac{\epsilon}{2}\|\nabla\psi^h\|_0^2 + \delta h\|\mathbf{b}\cdot\nabla\psi^h\|_0^2\right) \tag{23}$$

for all $\psi^h \in S^h_{d,0}$.

Proof. As in the proof of Theorem 2 we start by bounding the inertial term $M(\cdot,\cdot)$. The difference is that now τ can be factored out and all element integrals can be collected in a single integral over Ω:

$$M(\psi^h, \psi^h) = (\psi^h, \psi^h) + \sum_{\mathscr{K}\in\mathscr{T}_h}(\psi^h, \tau\mathbf{b}\cdot\nabla\psi^h)_{0,\mathscr{K}} = \|\psi^h\|_0^2 + \tau(\psi^h, \mathbf{b}\cdot\nabla\psi^h)_{0,\Omega}.$$

Consider first the case when $\epsilon > 0$. Because $\mathbf{b}$ is solenoidal and ψ^h vanishes on Γ, integration by parts shows that

$$2\int_\Omega \psi^h(\mathbf{b}\cdot\nabla\psi^h)\,\mathrm{d}\Omega = -\int_\Omega (\psi^h)^2\nabla\cdot\mathbf{b}\,\mathrm{d}\Omega + \int_\Gamma (\psi^h)^2\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}\Gamma = 0.$$

If $\epsilon = 0$ then $\mathbf{n}\cdot\mathbf{b} \geqslant 0$ on Γ_+ and

$$2\int_\Omega \psi^h(\mathbf{b}\cdot\nabla\psi^h)\,\mathrm{d}\Omega = \int_{\Gamma_+} (\psi^h)^2\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}\Gamma \geqslant 0.$$

In either case,

$$M(\psi^h, \psi^h) \geqslant \|\psi^h\|_0^2,$$

which in combination with (16) proves the theorem. □

Theorem 3 leads to a sharper stability bound because it accounts for the fact that for constant τ the additional "mass" term is either skew or gives a non-negative contribution regardless of the time step. Therefore, this term cannot be destabilizing because it will either vanish or *add*, rather than *take away* stability!

Such a conclusion does not contradict the more cautious stability condition (20), because, again, violation of (20) does not imply *instability*. However, (20) is too conservative to be of any predictive value for the model problems used in our numerical experiments. The lack of sharpness in this condition is caused by

the early use of the Cauchy's inequality in the proof of (20). This forces an estimate of the extra mass term by the streamline derivative and leads to the subsequent *subtraction* of beneficial streamline diffusion. Consequently, the proof cannot take advantage of the fact that τ is constant and that element integrals can be combined to form a skew term.

While conclusions of Theorem 3 are valid in a specific setting, they indicate a strong possibility that the sufficient stability condition (19) inferred from Theorem 2 may be unduly restrictive even for a variable τ. In the next section we develop sharp upper bounds for the additional mass term and show that this is indeed the case. Using these bounds we prove stability of SUPG finite elements for arbitrary CFL numbers.

3. Stability analysis of fully discrete equations

The bilinear form $B(\cdot,\cdot;\rho,\theta)$ serves to define the algebraic problem that advances the solution to the next time level. The main goal of this section is to determine whether or not the additional terms engendered by the coupling between the spatial weight function and the time dependent residual can destabilize the time stepping process by destroying the coercivity of $B(\cdot,\cdot;\rho,\theta)$. To avoid unnecessary technical details, in addition to (10) we will assume that $\tau(\mathbf{x})$ is constant on each element, that is, for all $\mathscr{K} \in \mathscr{T}_h$

$$\tau(\mathbf{x})|_{\mathscr{K}} = \tau(\mathscr{K}) \quad \forall \mathbf{x} \in \mathscr{K}.$$

In what follows we will consider general advective–diffusive problems and uniformly regular (but not necessarily uniform) partitions $\mathscr{T}_h$. The key to proving sharp stability conditions will be to obtain tight bounds for the additional "mass" term. For this purpose we begin with a technical lemma that estimates this term for a variable τ.

Lemma 1. *Assume that* ϵ, $\mathbf{b}$ *and* $\mathscr{T}_h$ *are such that* (10) *holds and that* $\nabla\cdot\mathbf{b} = 0$ *and* $\|\mathbf{b}\|_{\infty,\Omega} \leqslant \beta$ *for some positive constant* β. *Then*

$$\sum_{\mathscr{K}\in\mathscr{T}_h} \tau(\mathscr{K})\left(\psi^h, \mathbf{b}\cdot\nabla\psi^h\right)_{0,\mathscr{K}} \leqslant hC\left\|\psi^h\right\|^2_{0,\Omega}, \tag{24}$$

where C *is a positive constant that depends on the diameter of* Ω, *the polynomial degree* d, *the values of* β, $\check{\tau}$ *and* $\hat{\tau}$, *but is independent of* h.

Proof. We give a detailed proof in two space dimensions. The proof in three dimensions follows by minor modifications. Let $\mathscr{K} \in \mathscr{T}_h$ be an arbitrary element. Because $\nabla\cdot\mathbf{b} = 0$ integration by parts gives

$$\int_{\mathscr{K}} \psi^h\left(\mathbf{b}\cdot\nabla\psi^h\right)\mathrm{d}x = \frac{1}{2}\int_{\partial\mathscr{K}} \left(\psi^h\right)^2 \mathbf{n}\cdot\mathbf{b}\,\mathrm{d}S.$$

On each element ψ^h is a polynomial function of degree at most d, and so, its square is a polynomial of degree at most $2d$

$$\left(\psi^h\right)^2 = \left(\sum_{i=0}^{n_{\mathrm{dof}}} \psi_i N_i\right)^2 = a_{00} + a_{10}x + a_{01}y + a_{11}xy + \cdots + a_{dd}x^d y^d,$$

with coefficients

$$a_{ij} = \sum_{k,l=1}^{d} \psi_{k_i}\psi_{l_j},$$

that are linear combinations of the products of the nodal coefficients of ψ^h. Because $\mathbf{b}$ is divergence free,

$$\int_{\partial\mathscr{K}} \mathbf{n}\cdot\mathbf{b}\,\mathrm{d}S = \int_{\mathscr{K}} \nabla\cdot\mathbf{b}\,\mathrm{d}x = 0. \tag{25}$$

As a result, after inserting the polynomial expression for $\left(\psi^h\right)^2$ into the boundary integral, contribution from the constant term a_{00} will vanish so that

$$\int_{\partial\mathscr{K}} \left(\psi^h\right)^2 \mathbf{n}\cdot\mathbf{b}\,\mathrm{d}S = \int_{\partial\mathscr{K}} (a_{10}x + a_{01}y + \cdots)\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}S.$$

Let $\mathbf{x}_P = (x_P, y_P)$ denote one of the vertices of $\mathscr{K}$. Because $\mathscr{T}_h$ is assumed to be uniformly regular, changing variables according to

$$x = \hat{x} + x_P \quad \text{and} \quad y = \hat{y} + y_P,$$

takes $\mathscr{K}$ inside a box $[-Ch, Ch]^2$, where C is a constant that does not depend on the particular element $\mathscr{K}$. Therefore,

$$\begin{aligned}\int_{\partial\mathscr{K}} (a_{10}x + a_{01}y + \cdots)\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}S &= \int_{\partial\hat{\mathscr{K}}} (a_{10}(\hat{x}+x_P) + a_{01}(\hat{y}+y_P) + \cdots)\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}\widehat{S}\\ &= \int_{\partial\hat{\mathscr{K}}} (a_{10}x_P + a_{01}y_P + \cdots)\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}\widehat{S} + \int_{\partial\hat{\mathscr{K}}} (a_{10}\hat{x} + a_{01}\hat{y} + \cdots)\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}\widehat{S}\\ &= \int_{\partial\hat{\mathscr{K}}} (\hat{a}_{10}\hat{x} + \hat{a}_{01}\hat{y} + \cdots)\mathbf{n}\cdot\mathbf{b}\,\mathrm{d}\widehat{S},\end{aligned}$$

where we have used (25) and that $a_{10}x_P + a_{01}y_P + \cdots$ is a constant. The new coefficients

$$\hat{a}_{ij} = \sum_{k,l=1}^{d} \mu_{ij}(\mathbf{x}_P, d)\, a_{k_i l_j}$$

are linear combinations of the old coefficients a_{kl} with factors $\mu_{ij}(\mathbf{x}_P, d)$ that depend only on the polynomial degree and the diameter of Ω. This gives the intermediate bound

$$\int_{\mathscr{K}} \psi^h\left(\mathbf{b}\cdot\nabla\psi^h\right)\mathrm{d}x \leqslant \frac{1}{2}\max|\hat{a}_{ij}|\,\|\mathbf{b}\|_{\infty,\mathscr{K}} \int_{\partial\hat{\mathscr{K}}} (|\hat{x}| + |\hat{y}| + \cdots)\,\mathrm{d}\widehat{S}. \tag{26}$$

To estimate the terms on the right hand side of (26) we first note that

$$\max_{i,j}|\hat{a}_{ij}| \leqslant C(\Omega, d)\max_{k,l}|a_{kl}|$$

and that

$$\max_{k,l}|a_{kl}| \leqslant C(d)\max_{i,j}|\psi_i\psi_j|.$$

Because for nodal finite element bases

$$\max_{0\leqslant i\leqslant N_{\mathrm{dof}}}|\psi_i| \leqslant \left\|\psi^h\right\|_{\infty,\mathscr{K}},$$

it is not hard to see that

$$\max_{i,j}|\hat{a}_{ij}| \leqslant C(\Omega, d)\left\|\psi^h\right\|^2_{\infty,\mathscr{K}},$$

where $C(\Omega, d)$ depends only on the diameter of Ω and the polynomial degree d of the finite element space, but not on mesh parameter h. Because the length of $\partial\widehat{\mathscr{K}}$ is of order $\mathrm{O}(h)$

$$\int_{\partial\hat{\mathscr{K}}} |\hat{x}|\,\mathrm{d}\widehat{S} = \int_{\partial\hat{\mathscr{K}}} |\hat{y}|\,\mathrm{d}\widehat{S} = \mathrm{O}(h^2).$$

Therefore, the line integral on the right hand side in (26) contributes terms of order $\mathrm{O}(h^2)$ and higher. To complete the proof we recall the inverse inequality; see [3,10],

$$\left\|\psi^h\right\|_{\infty,\mathscr{K}} \leqslant C_I h^{-n/2} \left\|\psi^h\right\|_{0,\mathscr{K}}.$$

Combining all estimates together and setting $n = 2$ gives

$$\int_{\mathscr{K}} \psi^h(\mathbf{b}\cdot\nabla\psi^h)\,\mathrm{d}x \leqslant \frac{1}{2}C(\Omega,d)h^2\left\|\psi^h\right\|^2_{\infty,\mathscr{K}}\|\mathbf{b}\|_{\infty,\mathscr{K}} \leqslant \frac{1}{2}C(\Omega,d)\left\|\psi^h\right\|^2_{0,\mathscr{K}}\|\mathbf{b}\|_{\infty,\mathscr{K}}.$$

The Lemma follows by observing that $\tau(\mathscr{K}) = \mathrm{O}(h)$. □

This result shows that for solenoidal advection fields the additional mass term contributed by the coupling between the SUPG operator and the finite difference in time can be completely absorbed in the consistent mass matrix. In particular, it will never dominate the streamline diffusion term and the amount of stabilizing streamline diffusion in $B(\cdot,\cdot,\rho,\theta)$ will not decrease when the time step is reduced. These observations are formalized in the next theorem.

Theorem 4. *Under the same assumptions as in Lemma* 1 *and for* $\sigma = 0$ (*SUPG spatial stabilization*)

$$B\left(\psi^h,\psi^h;\rho,\theta\right) \geqslant \frac{\rho}{2}(1 - C_1 h)\left\|\psi^h\right\|_0^2 + \theta\left(\frac{\epsilon}{2}\left\|\nabla\psi^h\right\|_0^2 + C_2 h\left\|\mathbf{b}\cdot\nabla\psi^h\right\|_0^2\right) \tag{27}$$

for all $\psi^h \in S^h_{d,0}$.

Proof. Follows immediately from Lemma 1 and (16). □

The main conclusion from this theorem is that streamline upwinding in space can be safely coupled with implicit time stepping. Nevertheless, one should be aware of the fact that reduction in the time step will change the balance between the mass and the stiffness matrices in the discrete equation. For very small time steps $B(\cdot,\cdot,\rho,\theta)$ will correspond to a discretization of a singularly perturbed problem and the onset of spurious oscillations in the vicinity of thin layers may be expected; see [9].

4. Numerical results

In this section we test how well the stability theory developed in Theorem 4 matches with computation. Our main focus is on the behavior of the fully discrete equations for small time steps.

According to Theorem 4 application of SUPG stabilization in space leads to a harmless additional term that can be absorbed in the consistent mass matrix for any Courant number. Therefore, this theorem *guarantees* computational stability for small time steps with, perhaps, the exception of small localized oscillations in the vicinity of sharp layers. We remind the reader that these are caused by the singularly perturbed nature of the equations as $\Delta t \mapsto 0$.

To test the conclusion of Theorem 4 we compare Galerkin and SUPG solutions of (6) in the pure advection limit, [6] i.e., for $\epsilon = 0$. Several different time steps are used to provide a representative range of

[6] We note that in this case adjoint and GLS weighting operators reduce to SUPG stabilization.

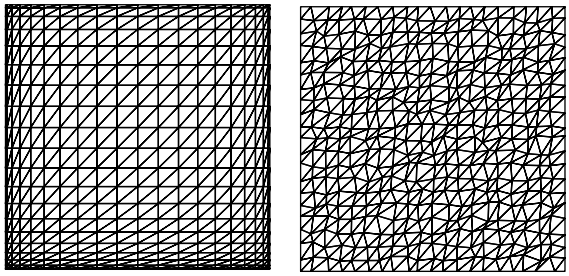

Fig. 3. Non-uniform meshes: mesh (A)—left, mesh (B)—right.

CFL values for each example problem. In all experiments Ω is the unit square, $\mathscr{T}_h$ is a uniform triangulation of Ω into triangles and S_d^h is the standard Lagrangian space consisting of C^0 piecewise polynomial functions whose restrictions to each element $\mathscr{K}$ of $\mathscr{T}_h$ are quadratic polynomials ($d = 2$).

We begin by solving all examples on a uniform mesh obtained by subdividing Ω into 400 squares and then drawing the diagonal in each one of them. This gives a partition containing 800 triangles with a mesh parameter $h = 0.05$, and a finite element space S_2^h with 1681 degrees of freedom. The space $S_{2,0}^h$ is defined by setting all nodal degrees of freedom that belong to Γ_- to zero. All matrices and right hand sides are assembled using a quintic (7 point) quadrature rule [2, p. 343].

Then we repeat the experiments using two different non-uniform meshes shown in Fig. 3 and having the same number of elements. Mesh (A) is a smooth deformation of the original uniform mesh. Mesh (B) is obtained by a random perturbation of the nodes in the uniform mesh.

Results from calculations on uniform grids are plotted by using the nodal values of the finite element solution. For non-uniform grids results are plotted by first generating the values of the finite element solution on a 41×41 uniform interpolation mesh. Thus, in all plots the axes are labeled by the node number, either with respect to the original uniform grid, or with respect to the uniform grid used to interpolate the finite element solution.

Example 1. The first model problem used in the numerical study is (6) with

$$\mathbf{b}_1 = \begin{pmatrix} 1.0 \\ 0.7002075 \end{pmatrix}; \quad g = 0$$

and

$$\phi_0(\mathbf{x}) = \begin{cases} 1 & \text{if } |\mathbf{x} - \mathbf{x}_C| \leqslant 0.2 \\ 0 & \text{otherwise} \end{cases}, \quad \mathbf{x}_C = \begin{pmatrix} 0.25 \\ 0.25 \end{pmatrix}. \tag{28}$$

The choice of $\mathbf{b}_1$ and the initial and boundary data corresponds to an advection of a cylinder of unit height, radius 0.2, and positioned at $\mathbf{x}_C$ in a direction skew to the mesh orientation. The homogeneous boundary data is specified on the inflow portion of the boundary

$$\Gamma_- = \{\mathbf{x} \in \overline{\Omega};\ x = 0\} \cup \{\mathbf{x} \in \overline{\Omega};\ y = 0\}.$$

Example 2. The second model problem is (6) with the variable solenoidal advective field

$$\mathbf{b}_2 = \begin{pmatrix} y \\ x \end{pmatrix} + \begin{pmatrix} 1.0 \\ 0.7002075 \end{pmatrix}$$

and the same initial and boundary data as in (28). The inflow boundary remains the same as in Example 1.

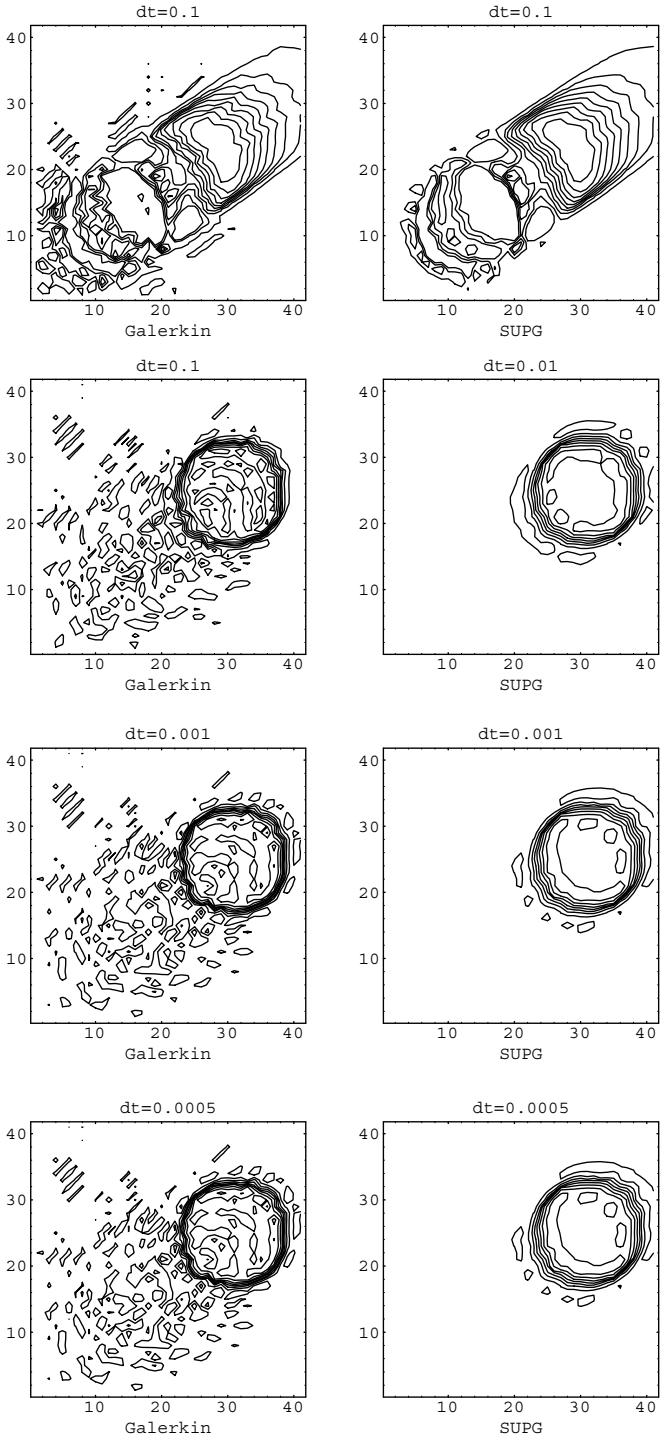

Fig. 4. Example 1. Galerkin (left) and SUPG (right) solutions at $t = 0.5$ computed with $\Delta t = 0.1$, $\Delta t = 0.01$, $\Delta t = 0.001$, and $\Delta t = 0.0005$.

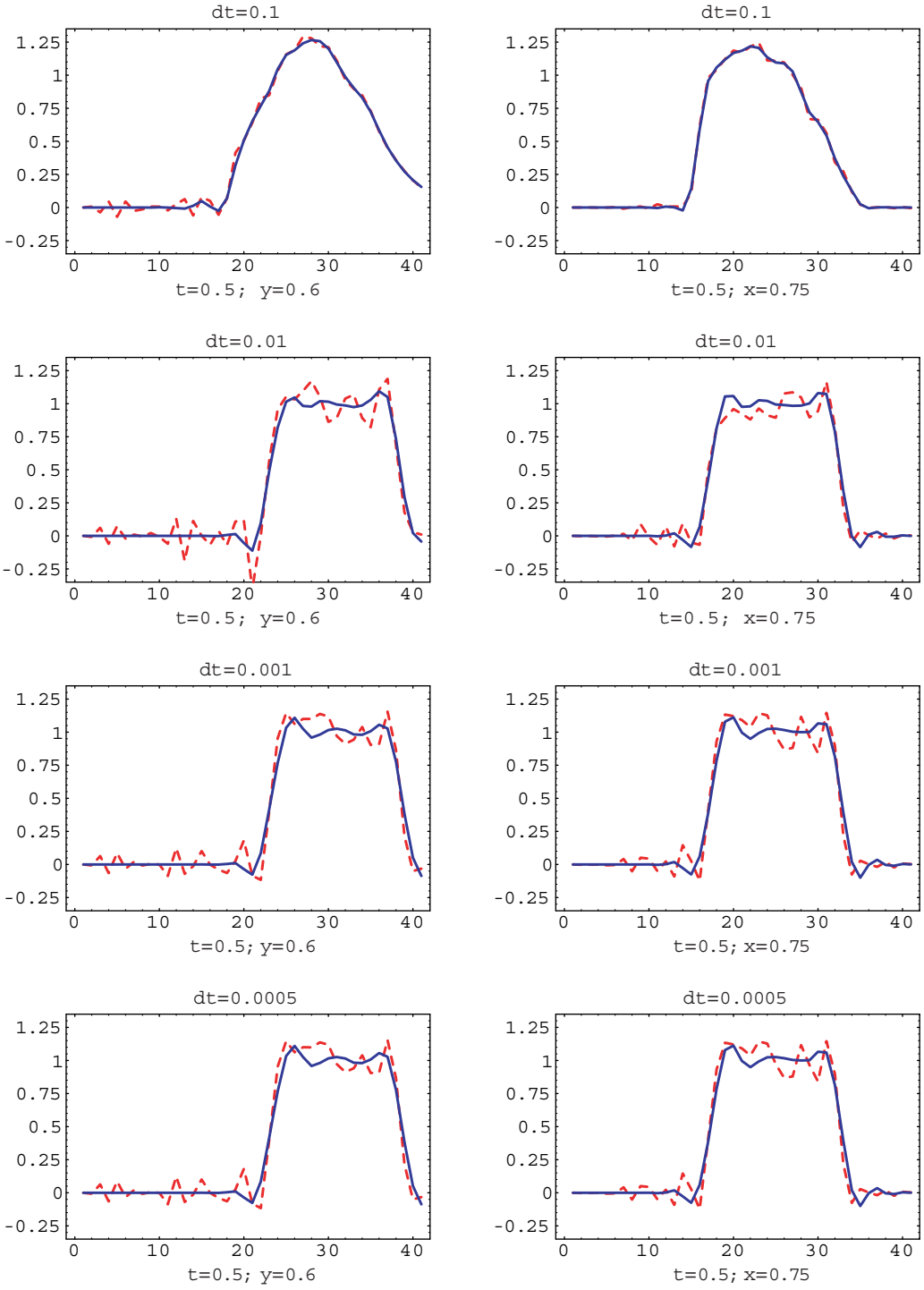

Fig. 5. Example 1. Slices of Galerkin (dashed) and SUPG (solid) solutions at $y = 0.6$ (left), $x = 0.75$ (right) and $t = 0.5$ computed with $\Delta t = 0.1$, $\Delta t = 0.01$, $\Delta t = 0.001$, and $\Delta t = 0.0005$.

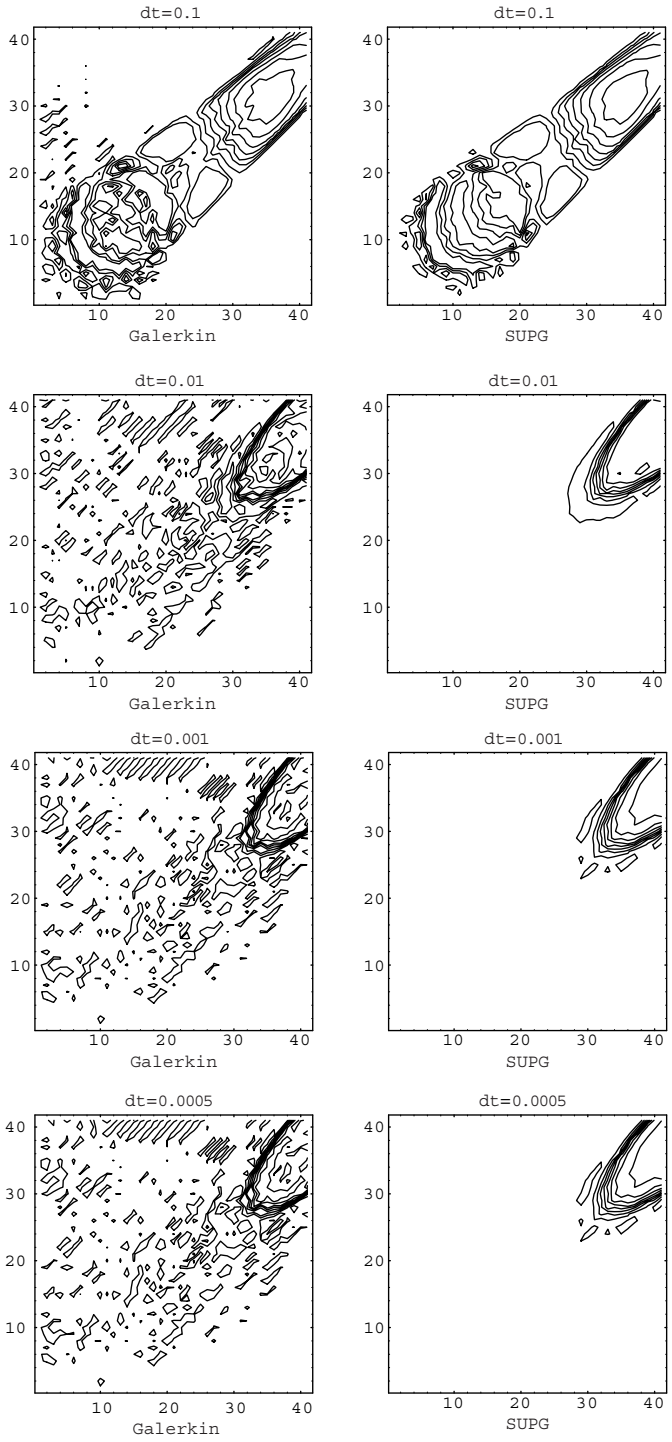

Fig. 6. Example 2. Galerkin (left) and SUPG (right) solutions at $t = 0.5$ computed with $\Delta t = 0.1$, $\Delta t = 0.01$, $\Delta t = 0.001$, and $\Delta t = 0.0005$.

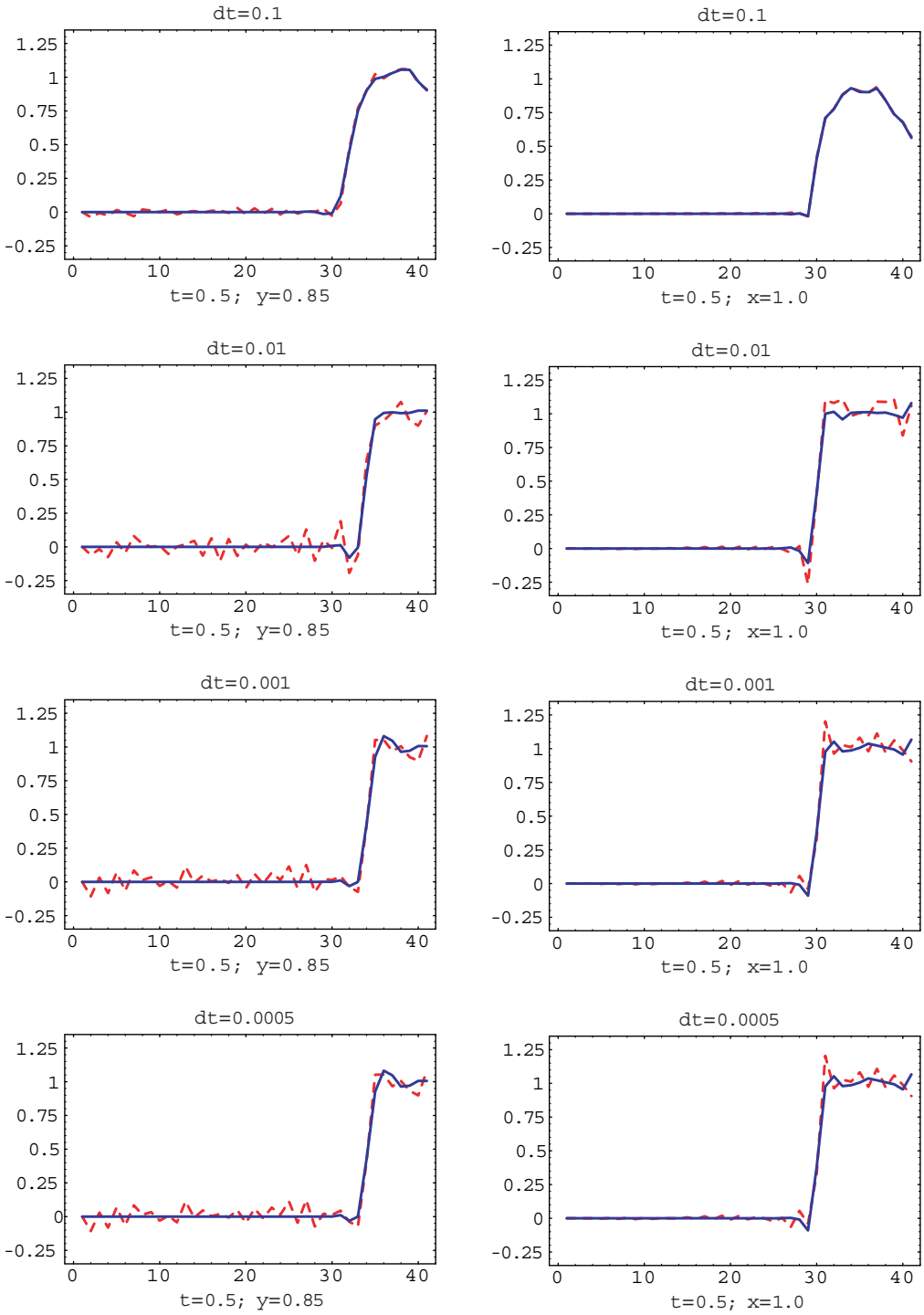

Fig. 7. Example 2. Slices of Galerkin (dashed) and SUPG (solid) solutions at $y = 0.85$ (left), $x = 1.0$ (right) and $t = 0.5$ computed with $\Delta t = 0.1$, $\Delta t = 0.01$, $\Delta t = 0.001$, and $\Delta t = 0.0005$.

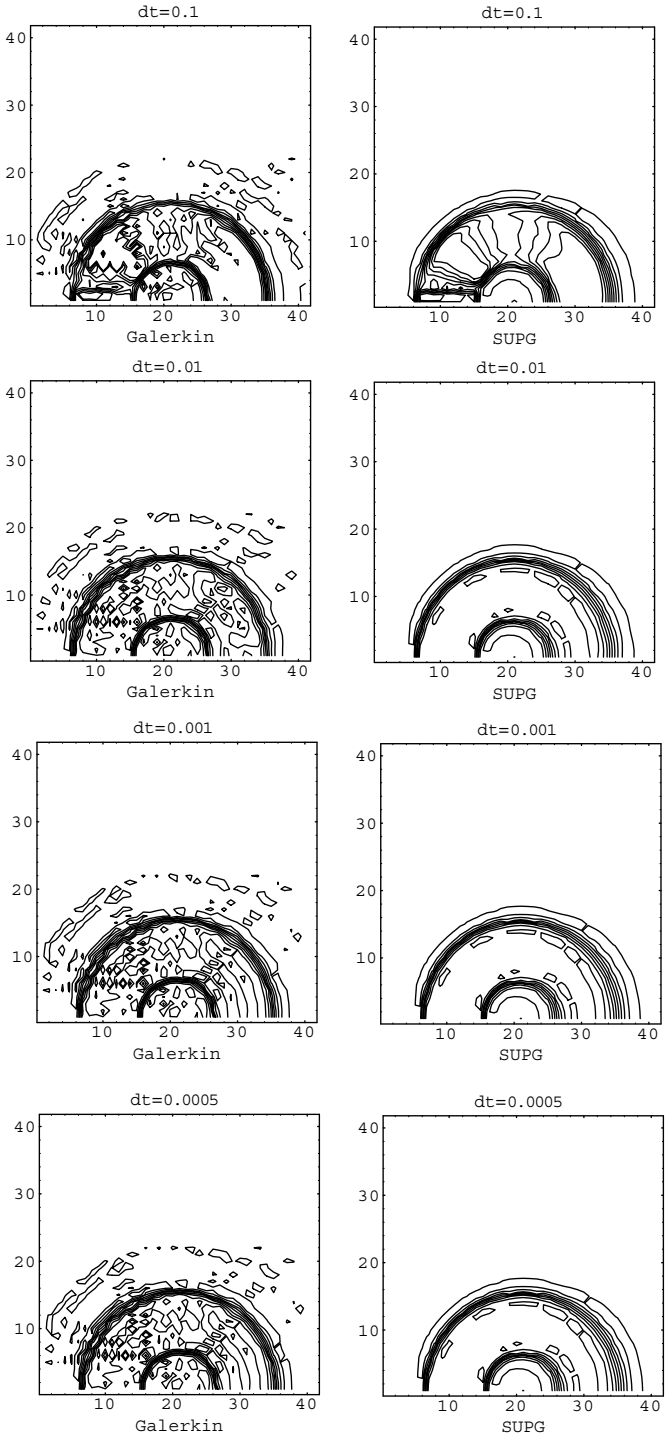

Fig. 8. Example 3. Galerkin (left) and SUPG (right) solutions at $t = 0.5$ computed with $\Delta t = 0.1$, $\Delta t = 0.01$, $\Delta t = 0.001$, and $\Delta t = 0.0005$.

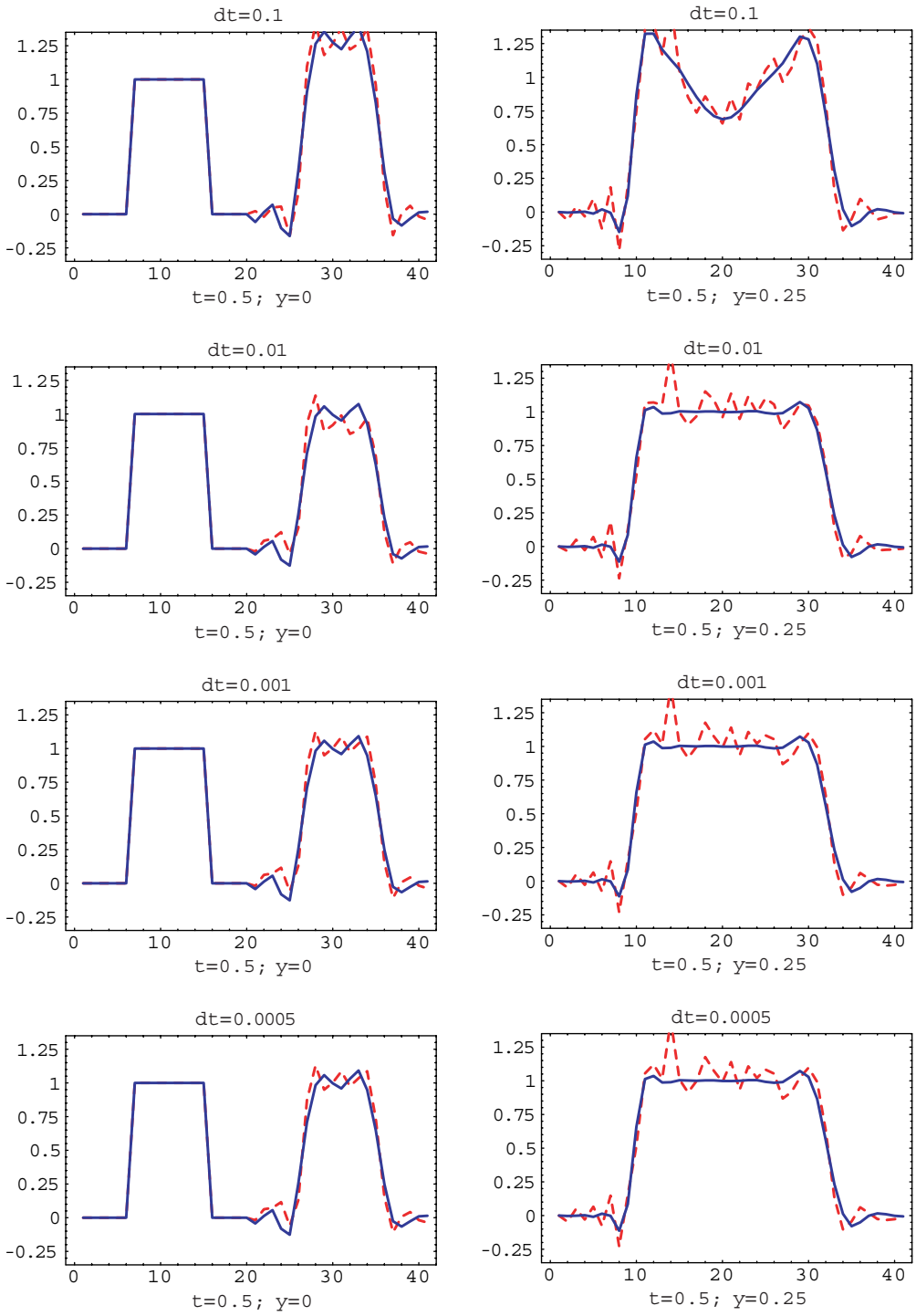

Fig. 9. Example 3. Slices of Galerkin (dashed) and SUPG (solid) solutions at $y = 0$ (left), $y = 0.25$ (right) and $t = 0.5$ computed with $\Delta t = 0.1$, $\Delta t = 0.01$, $\Delta t = 0.001$, and $\Delta t = 0.0005$.

Table 1
Example 1. H^1 seminorm of finite element solutions at $t = 0.5$

Δt	0.1	0.01	0.001	0.0005
CFL	2.442	0.2442	0.02442	0.0122
Method	H^1 seminorm			
Galerkin	0.8357E+01	0.8278E+01	0.8298E+01	0.8300E+01
SUPG	0.6390E+01	0.4715E+01	0.4684E+01	0.4684E+01

Table 2
Example 2. H^1 seminorm of finite element solutions at $t = 0.5$

Δt	0.1	0.01	0.001	0.0005
$\mathrm{CFL}_{\max}$	6.7204	0.67204	0.06720	0.0336
Method	H^1 seminorm			
Galerkin	0.8868E+01	0.8303E+01	0.8073E+01	0.8069E+01
SUPG	0.6943E+01	0.3720E+01	0.3640E+01	0.3639E+01

Table 3
Example 3. H^1 seminorm of finite element solutions at $t = 0.5$

Δt	0.1	0.01	0.001	0.0005
$\mathrm{CFL}_{\max}$	22.361	2.2361	0.22361	0.1118
Method	H^1 seminorm			
Galerkin	0.1030E+02	0.9253E+01	0.9204E+01	0.9205E+01
SUPG	0.7207E+01	0.6290E+01	0.6289E+01	0.6289E+01

Example 3. The last model problem in our numerical study is (6) with

$$\mathbf{b}_3 = 10\begin{pmatrix} y \\ 0.5 - x \end{pmatrix}; \quad \phi_0(\mathbf{x}) = 0$$

and inhomogeneous boundary data

$$g = \begin{cases} 0 & \text{if } y = 0 \text{ and } 0 \leqslant x < 0.125 \text{ or } 0.375 < x \leqslant 0.5, \\ 1 & \text{if } y = 0 \text{ and } 0.125 \leqslant x \leqslant 0.375, \\ 0 & \text{if } x = 0 \text{ or } y = 1 \text{ and } 0.5 \leqslant x \leqslant 1. \end{cases}$$

This example corresponds to a circular advection of an initial square profile. Note that in this case

$$\Gamma_- = \{\mathbf{x} \in \overline{\Omega};\ x = 0\} \cup \{\mathbf{x} \in \overline{\Omega};\ y = 1 \text{ and } 0.5 \leqslant x \leqslant 1\} \cup \{\mathbf{x} \in \overline{\Omega};\ y = 0 \text{ and } 0 \leqslant x \leqslant 0.5\}.$$

The three example problems are discretized in time using the neutrally stable Crank–Nicolson method ($\theta = 0.5$) and a uniform time step Δt. The inhomogeneous initial condition in Example 3 is approximated by its nodal interpolant out of S_2^h The fully discrete equation (13) is solved for different time steps using a direct solver from the LAPACK library. In particular, we choose $\Delta_1 t = 0.1$, $\Delta_2 t = 0.01$, $\Delta_3 t = 0.001$ and $\Delta_4 t = 0.0005$, and integrate in time until $t = 0.5$. The number of time steps required in each case is 5, 50, 500 and 1000, respectively. The choice of time steps ensures that CFL numbers for all three examples include values above and below one.

Figs. 4, 6 and 8 show contours of the three example solutions at $t = 0.5$ computed using the four different time steps. Each refinement of the time step leads to a reduction in the CFL number. From these plots it is clear that stability of the SUPG solution does not suffer when the time step is refined. This conclusion is also confirmed by plots of solution profiles along selected x and y coordinate values, presented in Figs. 5, 7 and 9. In all cases we see that the first two time refinement steps improve the accuracy of SUPG solutions, while the last refinement does not lead to a appreciable change in these solutions, i.e., they have converged in time.

The absence of destabilization in the SUPG solutions, as the time step is being refined, can also be verified by inspecting their H^1 seminorm in Tables 1–3. The top two rows in these tables show the time step and the associated CFL number. (For variable advection fields the maximal CFL number is reported.) In

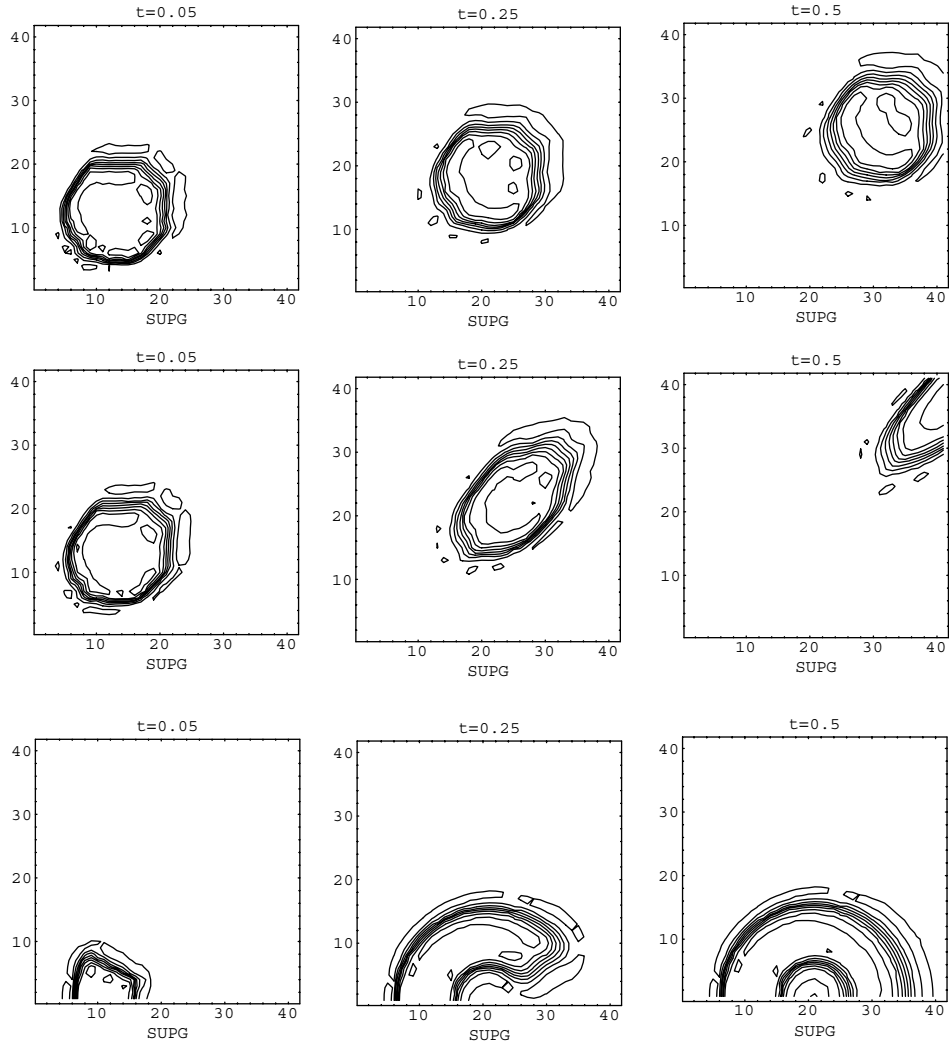

Fig. 10. Snapshots of SUPG solutions for example problems 1 (top), 2 (middle) and 3 (bottom) at $t = 0.05$, $t = 0.25$ and $t = 0.5$ and non-uniform mesh (A).

all three example cases the H^1 seminorm, which measures the amount of oscillation in the solution, does not change as the time step is reduced from $\Delta_3 t$ to the final value $\Delta_4 t$.

When the uniform grid was substituted by either one of the two non-uniform grids shown in Fig. 3, Galerkin and SUPG solutions did not change appreciably in their behavior. For this reason, below we limit ourselves to just a few snapshots of the finite element solutions on the non-uniform meshes computed using the finest time step $\Delta_4 t$. Fig. 10 shows solutions of the three example problems at three different instants in time computed on mesh (A). Fig. 11 shows the same time snapshots but computed using the randomly perturbed mesh (B). In both cases we see that SUPG stabilization performs an exemplary job in suppressing the global spurious oscillations, and that no destabilization is present in the solutions. The absence of destabilization on the two non-uniform grids can also be inferred from the data in Table 4. We see that H^1 seminorms of solutions computed on the non-uniform grids remain bounded by the seminorm values on the uniform grid.

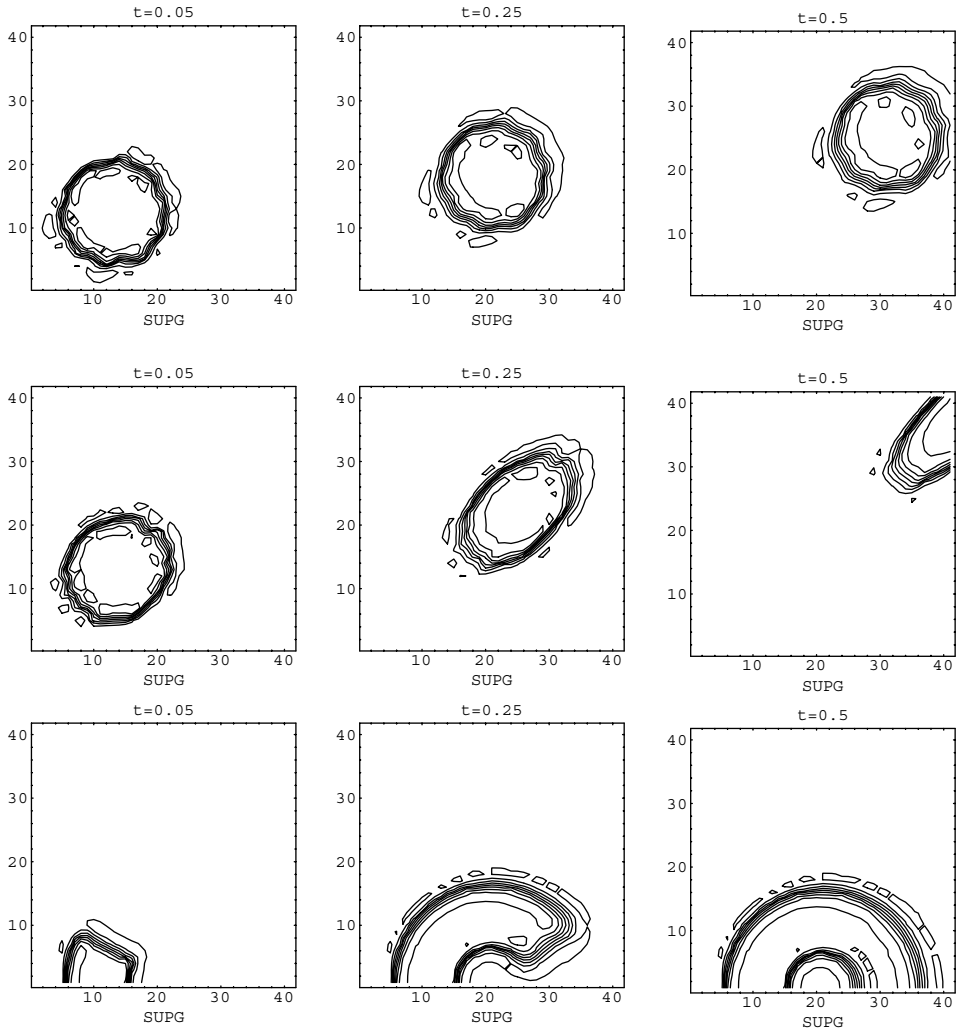

Fig. 11. Snapshots of SUPG solutions for example problems 1 (top), 2 (middle) and 3 (bottom) at $t = 0.05$, $t = 0.25$ and $t = 0.5$ and non-uniform mesh (B).

Table 4
H^1 seminorms of finite element solutions at $t = 0.5$ computed on different meshes and $\Delta t = 0.0005$

Mesh	Example 1	Example 2	Example 3
Uniform	0.4684E+01	0.3639E+01	0.6289E+01
Mesh (A)	0.4157E+01	0.3159E+01	0.5722E+01
Mesh (B)	0.4609E+01	0.3434E+01	0.6026E+01

In summary, our results clearly show the expected pollution by global spurious oscillation in the Galerkin solution and their successful suppression by the SUPG stabilization for all time steps considered in this study. Regarding the small localized oscillations in SUPG solutions we recall that SUPG is not monotonicity preserving, and that such oscillations can be expected in the vicinity of discontinuities and internal layers. Therefore, their presence cannot serve as an indication of a destabilization. Moreover, as the data in Tables 1–3 show, smaller time steps do not lead to an increase in the H^1 seminorm of the solutions, i.e., these oscillations remain bounded for small time steps. An application of a discontinuity capturing operator [16] is recommended for a further suppression of these oscillations.

5. Conclusions

We have considered fully discrete problems obtained by coupling implicit integration in time with spatial advective stabilization. Such formulations serve as an alternative to space-time discretizations and offer many advantages in the algorithmic development.

Our results show that some concerns raised about the possible destabilizing effect of Petrov–Galerkin upwinding in that context, and for small time steps, are unfounded. In fact, application of the streamline upwind stabilization operator in conjunction with implicit time integration can be considered as a safe separated discretization that does not lead to any additional stability restrictions on the Peclet or Courant numbers.

Galerkin least squares and multiscale (adjoint) stabilization cases will be a subject of a forthcoming paper.

In closing, we stress upon the fact that the numerical results presented in this paper are in excellent agreement with the theory and demonstrate that our analytical results are sharp. These results also hold with minor modifications for fully discrete formulation of the advective–diffusive–reactive model. Our conclusions about stability of fully discrete equations are also consistent with an earlier Von-Neumann stability analysis of the semidiscrete SUPG equation carried out in [12] for uniform grids.

Acknowledgements

The authors express their gratitude to professor T.J.R. Hughes for his many helpful suggestions and catching some inaccuracies in an early draft of this paper, and drawing our attention to [12,17,22].

References

[1] S.F. Bradford, N.D. Katopodes, The antidissipative, non-monotone behavior of Petrov–Galerkin upwinding, Int. J. Numer. Meth. Fluids 33 (2000) 583–608.
[2] G. Carey, T. Oden, Finite Elements. Computational Aspects, Prentice-Hall, Englewood Cliffs, NJ, 1984.
[3] P. Ciarlet, The Finite Element Method for Elliptic Problems, SIAM, Philadelphia, 2002.

[4] L.P. Franca, E.G. Dutra do Carmo, The Galerkin gradient least-squares method, Comput. Methods Appl. Mech. Engrg. 74 (1989) 41–54.
[5] L.P. Franca, C. Farhat, Bubble functions prompt unusual stabilized finite element methods, Comput. Methods Appl. Mech. Engrg. 123 (1995) 299–308.
[6] L.P. Franca, S. Frey, T.J.R. Hughes, Stabilized finite element methods: I. Application to the advective–diffusive model, Comput. Methods Appl. Mech. Engrg. 95 (1992) 253–276.
[7] L.P. Franca, F. Valentin, On an improved unusual stabilized finite element method for advective–reactive–diffusive equations, Comput. Methods Appl. Mech. Engrg. 189 (2000) 1785–1800.
[8] I. Harari, Spatial stability of semidiscrete formulations for parabolic problems, in: J. Eberhardsteiner H. Mang, F. Rammerstorfer, (Eds.), Proceedings of the Fifth World Congress on Computational Mechnics, Vienna, Austria, 7–12 July 2002, TU Vienna, Technical University, Vienna.
[9] I. Harari, Stability of semidiscrete formulations for parabolic problems at small time steps, Comput. Methods Appl. Mech. Engrg. 193 (2004) 1491–1516.
[10] I. Harari, T.J.R. Hughes, What are c and h? Inequalities for the analysis and design of finite element methods, Comput. Methods Appl. Mech. Engrg. 97 (1992) 157–192.
[11] T.J.R. Hughes, Multiscale phenomena: Green's function, the Dirichlet-to-Neumann map, subgrid scale models, bubbles and the origins of stabilized methods, Comput. Methods Appl. Mech. Engrg. 127 (1995) 387–401.
[12] T.J.R. Hughes, A. Brooks, Streamline upwind/Petrov–Galerkin formulation for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations, Comput. Methods Appl. Mech. Engrg. 32 (1982) 199–259.
[13] T.J.R. Hughes, A. Brooks, A theoretical framework for Petrov–Galerkin methods with discontinuous weighting functions: application to the streamline-upwind procedure, in: R.H. Gallagher et al. (Eds.), Finite Elements in Fluids, vol. 4, J. Willey & Sons, 1982, pp. 47–65.
[14] T.J.R. Hughes, G.R. Feijoo, L. Mazzei, J.B. Quincy, The variational multiscale method: a paradigm for computational mechanics, Comput. Methods Appl. Mech. Engrg. 166 (1998) 3–24.
[15] T.J.R. Hughes, L.P. Franca, G. Hulbert, A new finite element formulation for computational fluid dynamics: VIII. The Galerkin/least-squares method for advective–diffusive equations, Comput. Methods Appl. Mech. Engrg. 73 (1989) 173–189.
[16] T.J.R. Hughes, M. Mallet, A. Mizukami, A new finite element formulation for computational fluid dynamics: II. Beyond SUPG, Comput. Methods Appl. Mech. Engrg. 54 (1986) 341–355.
[17] T.J.R. Hughes, J.R. Stewart, A space-time formulation for multiscale phenomena, Comput. Methods Appl. Mech. Engrg. 74 (1995) 217–229.
[18] F. Ilinica, J.-F. Hetu, Galerkin gradient least-squares formulations for transient conduction heat transfer, Comput. Methods. Appl. Mech. Engrg. 191 (2002) 3073–3097.
[19] C. Johnson, U. Navert, J. Pitkaranta, Finite element methods for linear hyperbolic problems, Comput. Methods Appl. Mech. Engrg. 45 (1984) 285–312.
[20] W.H. Raymond, A. Gardner, Selective damping in a Galerkin method for solving wave problems with variable grids, Monthly Weather Rev. 104 (1976) 1583–1590.
[21] F. Shakib, T.J.R. Hughes, A new finite element formulation for computational fluid dynamics: IX. Fourier analysis of space-time Galerkin/least-squares algorithms, Comput. Methods Appl. Mech. Engrg. 87 (1991) 35–58.
[22] F. Shakib, T.J.R. Hughes, Z. Johan, A new finite element formulation for computational fluid dynamics: X. The compressible Euler and Navier–Stokes equations, Comput. Methods Appl. Mech. Engrg. 89 (1991) 141–219.