INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2006; **50**:347–376
Published online 7 September 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/fld.1060

Comparison of Galerkin and control volume finite element for advection–diffusion problems‡

M. J. Martinez*,†

Engineering Sciences Center, Sandia National Laboratories, Albuquerque, NM 87185-0735, U.S.A.

SUMMARY

The control volume finite element method (CVFEM) was developed to combine the local numerical conservation property of control volume methods with the unstructured grid and generality of finite element methods (FEMs). Most implementations of CVFEM include mass-lumping and upwinding techniques typical of control volume schemes. In this work we compare, via numerical error analysis, CVFEM and FEM utilizing consistent and lumped mass implementations, and stabilized Petrov–Galerkin streamline upwind schemes in the context of advection–diffusion processes. For this type of problem, we find no apparent advantage to the local numerical conservation aspect of CVFEM as compared to FEM. The stabilized schemes improve accuracy and degree of positivity on coarse grids, and also reduce iteration counts for advection-dominated problems. Published in 2005 by John Wiley & Sons, Ltd.

KEY WORDS: advection–diffusion; control volume finite element; streamline upwind control volume; numerical conservation

1. INTRODUCTION

In this paper we compare control volume finite element methods (CVFEM) with 'traditional' finite element methods (FEM), Galerkin finite element (GFEM) and streamline upwind Petrov–Galerkin (SUPG), in the context of scalar advection–diffusion. The advection–diffusion equation (ADE) is the simplest model of more complicated flow processes, such as viscous flow. The CVFEM method, as it is known today, appears to have been developed in 1980 by Baliga and Patankar [1] for advection–diffusion problems using triangular elements with directional upwind interpolation, and by Ramadhyani and Patankar [2] for the Poisson equation using quadrilateral elements. Ramadhyani and Patankar [3] extended the method

*Correspondence to: M. J. Martinez, Sandia National Laboratories, Engineering Sciences Center, P.O. Box 5800, Albuquerque, NM 87185-0735, U.S.A.
†E-mail: mjmarti@sandia.gov

Contract/grant sponsor: Sandia Corporation; contract/grant number: DE-AC04-94AL85000

Received 16 March 2005
Revised 30 June 2005
Accepted 1 July 2005

Published in 2005 by John Wiley & Sons, Ltd.

to advection–diffusion problems using quadrilateral elements and upwind interpolation. For certain problems, these upwinding schemes were found to suffer from excess cross-wind diffusion and oscillations due to generation of negative coefficients, hence Schneider and Raw [4] developed a skew-upwind scheme on bilinear quadrilateral elements. Subsequent developments include tetrahedral elements for incompressible flow [5, 6], and various upwinding schemes. Gresho and Sani [7] have presented a detailed description of the CVFEM scheme, without upwinding or mass lumping, and compared it to the GFEM, including Fourier spectral analysis of dispersion characteristics in one dimension. Given that the CVFEM seemingly combines the best attributes of GFEM and control volume methods, it is perhaps surprising that it has not attained the popularity of these methods. The present comparison will attempt to clarify this issue.

It is worthwhile to make the distinction with another method in the literature, which often also goes by the name control volume finite element method. An early example of this technique is in Reference [8] for oil reservoir problems; current examples include several papers by Forsyth and co-workers for nonisothermal two-phase flow in porous media; see, for example, References [9–11], and their included references. Neises and Steinbach [12] discuss this scheme for incompressible flow and Lohner [13] devotes Chapter 10 to discussing this technique, which he calls an edge-based FEM. While the approximations used by these researchers to arrive at their discrete equations can vary somewhat, they all start with finite element formulations and arrive at discrete equations that are edge based, as in a traditional control volume scheme. By edge based, we mean that the resulting discrete equations have the pairwise difference form between nodal variables. What distinguishes the edge-based method from the CVFEM, is that in the latter method the control-volumes are explicitly defined in space (see Figure 1); they form a dual-mesh with respect to the finite element mesh. In the edge-based method, the control volumes are not explicitly defined, they essentially overlap within the finite elements. Furthermore, the edge-based form is derived from finite element equations via manipulations which exploit the properties of the basis functions; there are no explicit surface integral calculations as in the CVFEM.

A primary motivation for developing the CVFEM was to combine the local conservation property of control volume methods with the unstructured grid and generality of the FEM. By contrast, it is generally agreed that the GFEM obeys a global conservation property (e.g. Section 2.2.3 of Reference [7]), however, the local conservation property is perhaps not as well-understood. Comini *et al*. [14] argue that the GFEM is locally conservative by deriving balances at the element and node level; their argument is rather subtle. Gresho and Sani [7] discuss the conservation properties of GFEM but are ambivalent on whether GFEM is locally conservative. They discuss both viewpoints, using arguments similar to Comini *et al*. [14] in their Appendix 2 and leave it to the reader to decide. Recently, Hughes *et al*. [15] have demonstrated the local conservation property of the GFEM, albeit with respect to conservative nodal fluxes whose calculation requires a post-processing step. In a similar vein, we show that the global conservation properties of the CVFEM scheme also require the post-processing of certain (computational domain) boundary fluxes to assure global conservation in the presence of Dirichlet boundaries.

In addition to conservation issues, another classic problem for numerical methods is their treatment of advectively dominated flows. The developers of CVFEM referenced above treated advectively dominated problems with *ad hoc* upwind schemes. By *ad hoc*, we mean that the advection schemes were developed apart from the weighted residual principle upon which

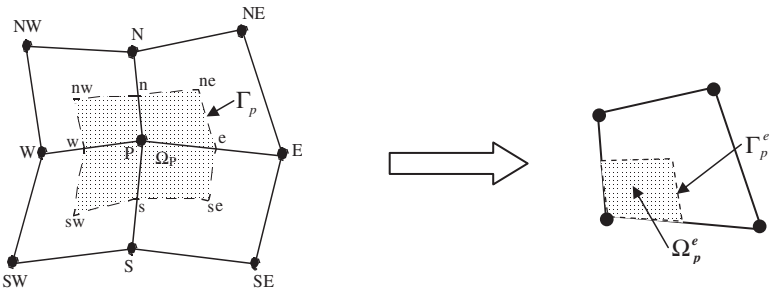

Figure 1. Grid system for the CVFEM discretization. The shaded region is Ω_p and $\Gamma_p = \partial\Omega_p$.

CVFEM is derived. Furthermore, most of these methods depart further from the weighted residual formulation by lumping the mass matrix, which, according to Gresho and Sani ([7], Sections 2.2.6 and 2.5.3), destroys the conservation property of the original scheme. An exception is the work of Swaminathan and Voller [16, 17] who discussed the development of a streamline upwind version of the CVFEM. They coined the method SUCV, for streamline upwind control volume, as a follow-on to the SUPG scheme of Hughes and co-workers [18, 19]. The scheme was formulated in an analogous way to SUPG, in the context of a Petrov–Galerkin method. These methods have recently been analysed and compared via Fourier analysis by Christon *et al.* [20] in 1D and Voth *et al.* [21] in 2D, as part of a larger set of methods for advection–diffusion problems.

In the following, we review the formulation of the CVFEM and SUCV scheme in a slightly different way than Swaminathan and Voller [16]. By recognizing that the CVFEM derives from a weighted residual statement in which the weight function is a distribution, the derivation follows simply by exploiting the operational characteristics of the particular distribution. The numerical implementation of the methods is then discussed. Next, the conservation properties of the CVFEM are analysed. It is shown that consistent fluxes on Dirichlet boundaries are necessary to assure global conservation. Finally, the methods are compared in detail via error analyses on several test problems chosen to elucidate different aspects of the operators. The analysis compares accuracy, conservation and positivity of CVFEM, SUCV, GFEM and SUPG.

2. PROBLEM FORMULATION

The initial-boundary value problem for the ADE is formulated as follows:

$$R(\mathbf{x},t) = \rho C \frac{\partial T}{\partial t} + \nabla \bullet (\boldsymbol{\sigma}(\mathbf{u})) - Q = 0 \quad \text{in } \Omega \tag{1}$$

where the flux vector is defined by

$$\begin{aligned} \boldsymbol{\sigma}^{\mathrm{a}} &= \rho C \mathbf{u} T \quad \text{(advective flux)} \\ \boldsymbol{\sigma}^{\mathrm{d}} &= -k \nabla T \quad \text{(diffusive flux)} \\ \boldsymbol{\sigma}(\mathbf{u}) &= \boldsymbol{\sigma}^{\mathrm{a}} + \boldsymbol{\sigma}^{\mathrm{d}} \quad \text{(total flux)} \end{aligned} \tag{2}$$

The ADE is subject an initial condition,

$$T(\mathbf{x},0)=T_0 \quad \text{in } \Omega \tag{3}$$

and boundary conditions,

$$\begin{aligned} T &= g(\mathbf{x},t) \text{ on } \Gamma_g \\ \boldsymbol{\sigma}(u)\bullet\mathbf{n} &= h(\mathbf{x},t) \text{ on } \Gamma_h \end{aligned} \tag{4}$$

Here Ω is an open bounded region with piecewise smooth boundary $\Gamma=\partial\Omega$; Γ_g, Γ_h are nonoverlapping partitions of Γ. The outward-pointing unit normal to Γ is denoted by $\mathbf{n}=(n_x, n_y, n_z)$. The equation is written as an energy equation, but the dependent variable could represent a number of other (passive) scalars. The specified data include a divergence-free velocity field ($\mathbf{u}$), a source function (Q), volumetric heat capacity (ρC), and diffusion coefficient (k), which may be zero in some cases. Boundary conditions include Dirichlet (essential) boundaries (Γ_g) and flux boundaries (Γ_h). The flux boundary condition needs clarification. On inflow regions of the boundary ($\mathbf{u}\bullet\mathbf{n}<0$) the convective and diffusive fluxes can be specified, while on outflow portions of the boundary, only the diffusive flux may be specified.

In the following we discuss the discrete formulation of the subject problem in the context of the CVFEM. The formulation of GFEM and SUPG methods is well established [18, 22, 23] and will not be discussed here. These latter methods have been implemented in the code MPSalsa [24] developed at Sandia National Laboratories, which was used in this study for comparison with the CVFEM-based schemes.

3. CONTROL VOLUME FINITE ELEMENT METHOD

3.1. Mathematical formulation

A derivation of the control volume equations in the context of a weighted residual scheme was presented previously by Swaminathan and Voller [16, 17] and Swaminathan *et al*. [25]. They derive their final equations by direct integration, allowing the flux and/or weight function to be discontinuous across certain internal (finite element) boundaries. Here we review the derivation and present it in a slightly different way, using the apparatus of generalized functions.

A weak solution is sought to the ADE by formulating a weighted residual statement

$$\int_\Omega wR\,\mathrm{d}\Omega = 0 \tag{5}$$

in a domain Ω with boundary Γ. Here R is the residual equation defined in Equation (1) and w is the weight function. The idea of the weighted residual statement is to find a weak solution such that the residual is orthogonal to the weight function. Substituting from Equation (1) and integrating the flux term by parts we get,

$$\int_\Omega \left\{ w\left(\rho C\frac{\partial T}{\partial t} - Q\right) + \nabla\bullet(w\boldsymbol{\sigma}) - \nabla w\bullet\boldsymbol{\sigma}\right\}\mathrm{d}\Omega = 0 \tag{6}$$

Now, rather than choosing the weight functions from typical FEM theory (C^0 functions with square-integrable derivatives), here we choose the following volume indicator distribution:

$$w = \gamma_p = \begin{cases} 1 & \text{in } \Omega_p \subset \Omega \\ 0 & \text{in the complement} \end{cases} \tag{7}$$

Useful properties of this generalized function include (Gray *et al.* [26], Equation (3.37))

$$\nabla \gamma_p = -\mathbf{n}\delta(\mathbf{x} - \mathbf{x}_{\Gamma_p})$$

and (Gray *et al.*, [26], Equation (6.20b)),

$$\begin{aligned} \int_{\Omega_\infty} \nabla \bullet (\gamma_p \boldsymbol{\sigma})\, \mathrm{d}\Omega = 0 \\ \int_{\Omega_\infty} \nabla \gamma_p \bullet \boldsymbol{\sigma}\, \mathrm{d}\Omega = -\int_{\Gamma_p = \partial\Omega_p} \boldsymbol{\sigma} \bullet \mathbf{n}\, \mathrm{d}\Gamma \end{aligned} \tag{8}$$

where $\boldsymbol{\sigma}$ in this case denotes any smooth vector field, $\mathbf{n}$ is the outward-pointing unit normal to Ω_p, $\delta(\mathbf{x} - \mathbf{x}_{\Gamma_p})$ is the Dirac δ function, which is zero everywhere except on $\Gamma_p (= \partial\Omega_p)$, and Ω_∞ denotes all of free space. To make use of these formulae, we extend the integral in Equation (6) over an infinite volume, and upon substitution for the terms involving the fluxes, we obtain the desired control volume equation for control volume p

$$\int_{\Omega_p} \left(\rho C \frac{\partial T}{\partial t} - Q \right) \mathrm{d}\Omega + \int_{\Gamma_p} \boldsymbol{\sigma} \bullet \mathbf{n}\, \mathrm{d}\Gamma + \int_{\Gamma_p \cap \Gamma_h} h\, \mathrm{d}\Gamma = 0 \tag{9}$$

where the last integral on the LHS applies when part of the subdomain volume coincides with the flux boundary (and in which case the second integral would apply to the control volume boundary modulo the intersection with Γ_h). This is also the control volume formulation of the ADE obtained from continuum mechanics. Note that in the numerical method we do not typically apply this conservation equation to Dirichlet control volumes (DCVs), since the solution is given there by the boundary conditions. However, this has important implications for global conservation properties as we shall see in the following. Upon discretization, this becomes the grid-point control volume equation. This form of the conservation equation explicitly displays the control-volumewise conservation property of the associated numerical scheme.

To derive the CVFEM form of the SUPG scheme, the Petrov–Galerkin method of FEM [22] is followed. The weighting function is now defined by

$$W = w + p \tag{10}$$

where w is the weight function described above for the CVFEM, and the perturbation weighting is defined by,

$$p = \tau \mathbf{u} \bullet \nabla w \tag{11}$$

The time-scale factor τ will be discussed shortly. The Petrov–Galerkin weight function used in the derivation of Swaminathan and Voller [16] includes a sign change, $W = w - p$, in order to obtain the appropriate sign for the artificial diffusion term. For the present derivation, substituting this Petrov–Galerkin weight function into the weighted residual statement,

Equation (5), and choosing the volume indicator distribution as the weight function for the control volume Ω_p, will result in Equation (9) plus a 'stabilizing' term,

$$\int_{\Omega_p} (\tau R\mathbf{u}) \bullet \nabla \gamma_p \, \mathrm{d}\Omega = -\int_{\Gamma_p} \tau R\mathbf{u} \bullet \mathbf{n} \, \mathrm{d}\Gamma \tag{12}$$

where we have applied the properties of the volume indicator distribution. Thus, the streamline upwind form of CVFEM, which we will refer to as SUCV, following Swaminathan and Voller [16, 17], is

$$\int_{\Omega_p} \left(\rho C \frac{\partial T}{\partial t} - Q \right) \mathrm{d}\Omega + \int_{\Gamma_p} \boldsymbol{\sigma} \bullet \mathbf{n} \, \mathrm{d}\Gamma + \int_{\Gamma_p \cap \Gamma_h} h \, \mathrm{d}\Gamma - \int_{\Gamma_p} \tau R \, \mathbf{u} \bullet \mathbf{n} \, \mathrm{d}\Gamma = 0 \tag{13}$$

where we have applied the flux boundary condition, if appropriate to this control volume.[§] This SUCV scheme is the analogue of the SUPG method for finite elements [18, 23]. The present formulation generalizes the derivation of Swaminathan and Voller [16, 17]. The explicit form of the artificial diffusion term in this method can be exposed by substituting the advection term, $\rho C\mathbf{u} \bullet \nabla T$, in place of the full residual in the stabilizing term of Equation (13), resulting in

$$\begin{aligned} &\int_{\Omega_p} \left(\rho C \frac{\partial T}{\partial t} - Q \right) \mathrm{d}\Omega + \int_{\Gamma_p} \rho CT\mathbf{u} \bullet \mathbf{n} \, \mathrm{d}\Gamma + \int_{\Gamma_p \cap \Gamma_h} h \, \mathrm{d}\Gamma \\ &\quad = \int_{\Gamma_p} \mathbf{n} \bullet (k\mathbf{I} + \tau \rho C \, \mathbf{uu}) \bullet \nabla T \, \mathrm{d}\Gamma \end{aligned} \tag{14}$$

The diffusion term on the RHS includes the isotropic physical diffusion plus a (tensor-valued) streamline upwind artificial diffusion term. This provides the streamline upwind stabilization for this scheme.

Note that the conservation law has been modified in the SUCV formulation. From Equation (13) the control volume equation is seen to conserve a modified effective flux,

$$\boldsymbol{\sigma}_{\text{eff}} = (\rho CT - \tau R)\mathbf{u} - k\nabla T \tag{15}$$

In the discrete form, the appended term $(\tau \, \mathbf{u} \, R)$ is never nonzero, hence, the stabilized formulation has an inherent conservation error with respect to the true physical flux. This aspect is also true of the SUPG and related stabilized FEM formulations, and in all methods with so-called artificial diffusion. In the Fourier analysis of Christon *et al.* [20] the spectral nature of artificial diffusion is illustrated for many popular numerical schemes, including those discussed here. The better numerical schemes inject appreciable artificial diffusion only for short wavelength modes, in the same spectral range where dispersion errors also become significant.

[§]To be precise, the limits on the flux integral (second term) should read $\Gamma_p \backslash (\Gamma_p \cap \Gamma_h)$ to indicate that this integral is performed only over portions of the control volume boundary that do not coincide with any applied flux boundary condition. The flux boundary condition is applied in the third integral, if it applies. We will continue to use this simplified notation in the following.

3.1.1. Stabilization parameter. We use the stabilization parameter as defined by Shakib *et al.* [27],

$$\tau = \beta \left[\left(\frac{2\rho C}{\Delta t} \right)^2 + (\rho C)^2 u_i g_{ij} u_j + 9k^2 g_{ij} g_{ij} \right]^{-1/2} \tag{16}$$

where g_{ij} is the metric tensor for the element mapping. This metric provides a consistent method for computing the representative element size for unstructured grids. For large Peclet number, $\tau = O(h/\|\mathbf{u}\|)$, while for vanishing Peclet number, $\tau = O(h^2/\kappa)$, where h denotes a mesh size measure and $\kappa = k/\rho C$. For steady problems the term involving the time step Δt is omitted. The only modification to Shakib's formula is the factor β, which can be used to optimize the dispersive properties of the SUCV scheme under pure advection conditions on Cartesian grids [20]. In general, the time term will dominate as $\Delta t \to 0$, which has the effect of annihilating the stabilization. As we demonstrate later in the examples, for pure advection we find improved performance by omitting the time term entirely and specifying $\beta = 1$ for SUCV and $\beta = 2/\sqrt{15}$ for SUPG.

3.2. Remarks on comparison with GFEM and SUPG

As originally discussed by Brooks and Hughes [18], the SUPG formulation can be written as follows:

$$\int_\Omega \left\{ w \left(\rho C \frac{\partial T}{\partial t} - Q \right) - \nabla w \bullet \boldsymbol{\sigma} \right\} \mathrm{d}\Omega + \sum_e \int_{\Omega^e} pR \, \mathrm{d}\Omega = \int_{\Gamma_h} wh \, \mathrm{d}\Gamma \tag{17}$$

with p as defined in Equations (10) and (11), and residual R as defined in Equation (1); the second integral is the so-called stabilizing term. The GFEM formulation results by omitting the stabilizing term. The τ stabilizing parameter used in the SUPG method implemented here is the same as used for SUCV, given in the previous section. The SUPG is called a consistent method, because the weak formulation is satisfied by the exact solution. However, as typically implemented for low-order elements, and in the following applications, the physical diffusion term in the residual R is usually omitted. Similarly, the CVFEM and SUCV derive from the variational statement in Equation (5), and could also be called consistent in this sense. Here too, typical implementations include various *ad hoc* approximations which would seem to obviate the consistency, including mass lumping, which is almost universally applied. Furthermore, it is clear from the formulations of both SUPG and SUCV that the original advection–diffusion PDE has been modified. Thus, both of these models conserve a fictitious numerical flux as opposed to the true physical flux, $\boldsymbol{\sigma}$, given in Equation (2). This is the price to be paid for these 'stabilized' methods, which allow one to compute reasonable solutions on underresolved grids. The originating methods are not unstable, they work fine on resolved grids, assuming the problem is continuous (no shocks or square waves). Unfortunately, a resolved grid is simply unfeasible for many practical problems with sharp layers and high (with respect to a practical mesh) gradients.

3.3. Discrete (vertex-centred) formulation

3.3.1. Spatial discretization. Spatial discretization is accomplished by using finite element techniques. Figure 1 shows the 2D grid system for quadrilateral elements used for defining

the vertex-centred control volumes [28] as a 'subgrid' of the finite elements. The figure shows a patch of 4 elements and a break-out of one particular finite element showing the subcontrol volume (sometimes called a co-volume) and its boundary inside the element.

Within each element, the unknown is represented in terms of the FEM basis

$$T(\mathbf{x},t) = \sum_J N_J(\mathbf{x}) T_J(t) \tag{18}$$

where N_J are the nodal interpolation (basis) functions and T_J are the nodal values of temperature. Bilinear quadrilaterals were applied in this work. With this representation, the numerical approximation to the ADE, including the stabilizing SUCV term, can be written for the Ith control volume as

$$\int_{\Omega_I} \sum_J (\rho C \dot{T} - Q)_J N_J \, \mathrm{d}\Omega + \int_{\Gamma_I} \sum_J \mathbf{n} \bullet (\rho C \mathbf{u} N_J - k \nabla N_J) T_J \, \mathrm{d}\Gamma - \int_{\Gamma_I} \tau R \mathbf{u} \bullet \mathbf{n} \, \mathrm{d}\Gamma = 0 \tag{19}$$

The control volume boundary within the element is along the bisectors of the (square) element in the mapped domain. Details of the calculation of the stabilizing terms are given in the following. Note that the foregoing discrete problem will lead, in general, to an unsymmetric mass matrix on general unstructured grids. Only for Cartesian grids will the mass matrix be symmetric.

3.3.2. Time integration. The variable-step, predictor–corrector method introduced by Gresho *et al.* [29] is implemented, including both a first-order (in time) and second-order scheme. The first-order scheme employs a forward Euler predictor with a backward Euler corrector. The second-order scheme employs an Adams–Bashforth predictor,

$$T^{p,n+1} = T^n + \frac{\Delta t_n}{2} \left(\left(2 + \frac{\Delta t_n}{\Delta t_{n-1}} \right) \dot{T}^n - \frac{\Delta t_n}{\Delta t_{n-1}} \dot{T}^{n-1} \right)$$

to provide an initial estimate of the temperature at the next time step, with a trapezoid rule corrector

$$\dot{T}^{n+1} = \frac{\eta}{\Delta t} (T^{n+1} - T^n) - (\eta - 1) \dot{T}^n$$

where $\eta = 2$ for trapezoid rule and $\eta = 1$ for backward Euler. This trapezoid rule is second-order accurate in time. The trapezoid method requires a start-up technique; the Euler scheme is used for the first two time steps in a transient simulation.

3.3.3. Finite element calculations. The foregoing discrete approximations follow standard finite element practice [22]. The actual computer processing of the discrete equations can also be performed in the same manner as for FEM. The FEM data structure is retained and the discrete equations are processed in element loops. The processing over each element involves computing element control volume equations for each subdomain (see Figure 1) which can then be assembled into a global matrix. Applying the foregoing approximations, the discrete

equations pertaining to control volume (node point) I take the following form:

$$\int_{\Gamma_I} \boldsymbol{\sigma} \bullet \mathbf{n}\, \mathrm{d}\Gamma = \int_{\Gamma_I} (\rho C \mathbf{u} \bullet \mathbf{n}) \lfloor N_1, \ldots, N_n \rfloor \, \mathrm{d}\Gamma \begin{pmatrix} T_1 \\ \vdots \\ T_n \end{pmatrix} - \int_{\Gamma_I} k \left(\begin{bmatrix} \dfrac{\partial N_1}{\partial x} & \cdots & \dfrac{\partial N_n}{\partial x} \end{bmatrix} n_x + \begin{bmatrix} \dfrac{\partial N_1}{\partial y} & \cdots & \dfrac{\partial N_n}{\partial y} \end{bmatrix} n_y \right) \mathrm{d}\Gamma \begin{pmatrix} T_1 \\ \vdots \\ T_n \end{pmatrix} \tag{20}$$

for the flux terms, and

$$\int_{\Omega_I} (\rho C \dot{T} - Q)\, \mathrm{d}\Omega = \frac{\eta}{\Delta t} \int_{\Omega_I} \rho C \lfloor N_1, \ldots, N_n \rfloor \, \mathrm{d}\Omega \begin{pmatrix} T_1 - T_1^n \\ \vdots \\ T_n - T_n^n \end{pmatrix} - \int_{\Omega_I} \lfloor N_1, \ldots, N_n \rfloor \begin{pmatrix} Q_1 + (\eta - 1)\dot{T}_1^n \\ \vdots \\ Q_n + (\eta - 1)\dot{T}_n^n \end{pmatrix} \mathrm{d}\Omega \tag{21}$$

for the volume terms, and where we have applied the time integration scheme. The superscript refers to the previous (nth) time plane and unsubscripted variable refer to the current time plane. Note that these CVFEMs require both volume integrals and surface integrals, whereas processing of FEM requires mostly volume integrals, except for (usually) a modest number of boundary terms. Calculation of these integrals follows standard FEM practice [7, 22].

3.3.4. Stabilizing term. The SUCV stabilizing term has the form

$$\int_{\Gamma} \tau \mathbf{u} \bullet \mathbf{n} \left(\rho C \frac{\partial T}{\partial t} + \nabla \bullet (\rho C \mathbf{u} T - k \nabla T) - Q \right) \mathrm{d}\Gamma = \int_{\Gamma} \tau \left[\mathbf{u} \bullet \mathbf{n} \left(\rho C \frac{\partial T}{\partial t} - Q \right) \right] + \tau \left[\underbrace{\rho C \mathbf{n} \bullet \mathbf{u}\mathbf{u} \bullet \nabla T}_{\text{streamwise}} \underbrace{- k \mathbf{u} \bullet \mathbf{n} \nabla^2 T}_{\text{diffusion}} \right] \mathrm{d}\Gamma \tag{22}$$

in which an artificial diffusion term is explicitly marked. With the exception of the second-order diffusion term, the operators can be expanded in terms of the bilinear (trilinear in 3D) basis discussed in the previous section. Brooks and Hughes [18] discuss conditions under which the second-order diffusion term can be dropped for low-order elements. This is common practice, although not necessary in terms of implementation. This second-order diffusion term could be approximated either by a projection or by a gradient reconstruction method, for

example, by a discrete application of the divergence theorem. The point of either method is to compute the temperature gradient at the nodes so that it can be represented throughout the element via the basis. Either method of calculating the gradient representation involves a larger stencil than simply the local values restricted to one finite element. Using such a method, the second-order diffusion term becomes

$$\int_{\Gamma_I} \tau k \mathbf{u} \bullet \mathbf{n} \nabla \bullet \mathbf{S} \, \mathrm{d}\Gamma = \int_{\Gamma_I} \tau k \mathbf{u} \bullet \mathbf{n} \left(\begin{bmatrix} \frac{\partial N_1}{\partial x} & \cdots & \frac{\partial N_n}{\partial x} \end{bmatrix} \begin{pmatrix} S_{x1} \\ \vdots \\ S_{xn} \end{pmatrix} + \begin{bmatrix} \frac{\partial N_1}{\partial y} & \cdots & \frac{\partial N_n}{\partial y} \end{bmatrix} \begin{pmatrix} S_{y1} \\ \vdots \\ S_{yn} \end{pmatrix} \right) \mathrm{d}\Gamma \tag{23}$$

where we have written $\mathbf{S}$ for the nodal representation of ∇T. If we use a FEM-based projection scheme, this is obtained by solving (in 2D)

$$\begin{bmatrix} \mathbf{M} & 0 \\ 0 & \mathbf{M} \end{bmatrix} \begin{pmatrix} \mathbf{S}_x \\ \mathbf{S}_y \end{pmatrix} = \int_\Omega \sum_J N_I \begin{pmatrix} \partial N_J / \partial x \\ \partial N_J / \partial y \end{pmatrix} T_J \, \mathrm{d}\Omega \tag{24}$$

where $\mathbf{M}$ is a finite element mass matrix.

3.3.5. Boundary conditions. Dirichlet boundary conditions are applied by replacing the conservation equation pertaining to a Dirichlet boundary node with an identity enforcing the desired boundary value. This is accomplished by specifying the residual equation for this node in the form $R_I = T_I - g(\mathbf{x}_I, t)$.

Flux boundaries with inflow ($\mathbf{u} \bullet \mathbf{n} < 0$) are treated by simply substituting the specified total flux $h(\mathbf{x}, t)$ into the flux integral. On outflow boundaries, only the diffusive component of flux can be specified; the advective component is treated implicitly. A common treatment of an outflow boundary (denoted Γ_h^+) is to specify the diffusive component to be zero. In this case the boundary flux term becomes

$$\int_{\Gamma_I \cap \Gamma_h^+} \boldsymbol{\sigma} \bullet \mathbf{n} \, \mathrm{d}\Gamma = \int_{\Gamma_I \cap \Gamma_h^+} \rho C \mathbf{u} \bullet \mathbf{n} \, \lfloor N_1, \ldots, N_n \rfloor \, \mathrm{d}\Gamma \begin{pmatrix} T_1 \\ \vdots \\ T_n \end{pmatrix} \tag{25}$$

which forms an implicit equation involving the boundary values of temperature on the outflow boundary.

3.3.6. Solution procedures. The FEM code MPSalsa [24], designed to solve reactive flow problems, was used as a platform for computer implementation of the foregoing numerical schemes. This facilitates an implementation that includes advanced features such as unstructured grid and parallel processing. The discrete system of equations is formulated in 'residual'

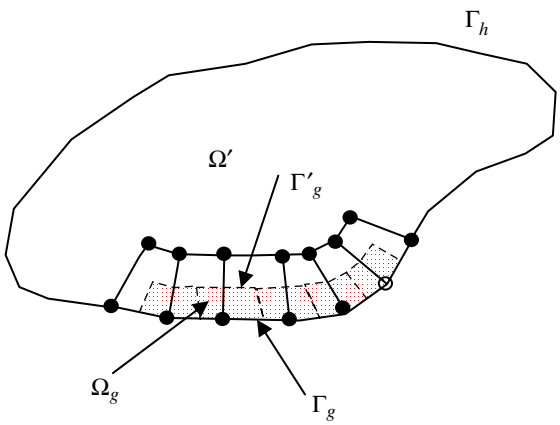

Figure 2. Domain with highlighted Dirichlet control volumes.

form to allow a Newton iteration scheme in the general case of temperature-dependent material properties, or to accommodate other nonlinearities in the specified data. The resulting linear systems were solved using the Aztec [30] parallel processing linear equation package.

In the following, CVFEM/SUCV simulations are compared to simulations using GFEM/SUPG. These latter methods are included in the MPSalsa code, which was used to provide the comparison calculations.

3.4. Conservation properties of the CVFEM and SUCV

A significant motivation for the development of the SUCV by Swaminathan and Voller [16, 17] was to combine the favourable aspects of FEM (especially SUPG) with the local, control-volumewise numerical conservation inherent in control volume methods. Since the CVFEMs are locally (on each control volume) conservative, by their construction, it would seem obvious to conclude that they must be globally conservative as well. In fact, this is almost true, except for consideration of conservation on Dirichlet boundaries. Global numerical conservation for CVFEM (and SUCV) requires that the fluxes on Dirichlet boundaries be computed consistently, using the CVFEM conservation equations, which are abandoned when applying Dirichlet BC values.

Consider Figure 2 which depicts a discretized computational region including both flux and Dirichlet boundaries. The DCVs are highlighted in Figure 2, and the finite elements to which they belong are explicitly outlined. The global conservation equation is determined by summing Equation (13) over control volumes. Note that integrals of the boundary fluxes are continuous across the control volume surfaces and therefore cancel out in this process. In the numerical method, we do not apply the control volume equations to DCVs. Rather, these equations are replaced by an identity that specifies the known Dirichlet boundary values. So, if we sum together the remaining control volume equations, we get the conservation equation for a subdomain $\Omega' = \Omega - \Omega_g$, where

$$\Omega_g = \bigcup_{p \in \mathrm{DCV}} \Omega_p \tag{26}$$

is the set of DCVs, see Figure 2. Performing the summation results in

$$\int_{\Omega'} \left(\rho C \frac{\partial T}{\partial t} - Q \right) \mathrm{d}\Omega + \int_{\Gamma'_g} \boldsymbol{\sigma} \bullet \mathbf{n} \, \mathrm{d}\Gamma + \int_{\Gamma_h} h \, \mathrm{d}\Gamma - \int_{\Gamma_h + \Gamma'_g} \tau R \mathbf{u} \bullet \mathbf{n} \, \mathrm{d}\Gamma = 0 \tag{27}$$

where $\Gamma'_g = \partial\Omega_g - \Gamma_g$, the interface between Ω' and Ω_g. Note that if $\Gamma_g = \varnothing$, and hence $\Gamma'_p = \varnothing$, this becomes a global conservation statement for the ADE in the absence of Dirichlet boundaries.

To this point the numerical scheme has not considered the flux through Dirichlet boundaries and therefore we cannot say whether the method obeys a global conservation principal. Equation (27) applies only to the subdomain which excludes the DCVs. Now, consider the control volume equation for a DCV, i.e.

$$\int_{\Omega_p} \left(\rho C \frac{\partial T}{\partial t} - Q \right) \mathrm{d}\Omega + \int_{\Gamma_p} \boldsymbol{\sigma} \bullet \mathbf{n} \, \mathrm{d}\Gamma + \int_{\Gamma_p \cap \Gamma_g} (\rho C g \mathbf{u} \bullet \mathbf{n} + \sigma_n^{\mathrm{d}}) \, \mathrm{d}\Gamma - \int_{\Gamma_p} \tau R \mathbf{u} \bullet \mathbf{n} \, \mathrm{d}\Gamma = 0 \tag{28}$$

where $\sigma_n^{\mathrm{d}} \equiv -k \nabla T \bullet \mathbf{n}$ is the diffusive flux normal to the Dirichlet boundary and the Dirichlet boundary values ($T = g(\mathbf{x}, t)$) have been applied. Part of the integral on Γ_p is over surfaces in common with other DCVs (Γ'_g) and part is over Γ_g. This equation can be formed for each DCV after the solution for T has been obtained, since all entries in the equation would be known, *except for the diffusive flux on* Γ_g. However, the set of equations for all DCVs can be discretized and solved for σ_n^{d} to compute a consistent, scalar-conserving diffusive normal flux on Γ_g. The procedure would involve representing the normal flux in terms of the same basis as the temperature, $\sigma_n^{\mathrm{d}} = \sum_J N_J \sigma_{n,J}^{\mathrm{d}}$. Note that the stabilizing term contains the gradient of this diffusive flux, which can also be represented by the gradient of this expansion in terms of the basis. Substituting this representation into the conservation equations for all DCVs, along with the known values of T, would result in a (small) system of equations to be solved for the grid-point values of normal diffusive flux, $\sigma_{n,J}^{\mathrm{d}}$.

The diffusive flux thus obtained completes the data required to form a globally conservative method. If we now sum the control volume equations for Ω' and Ω_g, making use of the fact that the outward normal on Γ'_g with respect to Ω_g and Ω' are equal in magnitude but oriented in opposite directions, we get the desired global conservation equation for this method,

$$\int_{\Omega} \left(\rho C \frac{\partial T}{\partial t} - Q \right) \mathrm{d}\Omega + \int_{\Gamma_h} h \, \mathrm{d}\Gamma + \int_{\Gamma_g} (\rho C g \mathbf{u} \bullet \mathbf{n} + \sigma_n^{\mathrm{d}}) \, \mathrm{d}\Gamma - \int_{\Gamma} \tau R \mathbf{u} \bullet \mathbf{n} \, \mathrm{d}\Gamma = 0 \tag{29}$$

Thus, by post-processing the conservation equations on the DCVs to compute control-volume-consistent diffusive fluxes, we get a globally (numerically) conservative formulation, as well as control-volumewise conservation for the CVFEM. However, the stabilized methods (SUCV) have a built-in error associated with the appended stabilizing terms with respect to the original ADE. They are numerically conservative, but with respect to a perturbed (advective plus diffusive) flux vector.

4. ERROR ANALYSIS

In this section we compare the CVFEM and GFEM schemes, including their stabilized counterparts, on several test problems chosen to focus on specific aspects of advection–diffusion

processes. The first set of test problems considers transient advection while the last considers steady advection–diffusion. The methods are compared in terms of error norms measuring accuracy, conservation and degree of positivity.

4.1. Molenkamp problem

There are several versions of the so-called Molenkamp problem, apparently deriving from the problem discussed in Reference [31], which, in general, involves the advection of an initial scalar distribution by the 2D flow field for solid body rotation:

$$u = -\omega y, \quad v = \omega x, \quad \omega = 2\pi \tag{30}$$

Differences in later specifications of this problem involve variations in the form of the initial profile. The current version follows the specification of Vreugdenhil and Koren [32] in which a Gaussian initial profile is used. The exact solution is

$$T(x, y, t) = 0.01^{4r^2}, \quad r = ((x + 1/2\cos(\omega t))^2 + (y + 1/2\sin(\omega t))^2)^{1/2} \tag{31}$$

from which is specified the initial condition ($t = 0$) and boundary condition at inflow. The solution is to be computed for one full revolution ($t = 1$) in the region $-1 \leqslant x, y \leqslant 1$ on a sequence of uniform grids (20×20, 40×40 and 80×80). The extent of this domain is such that the solution values on the boundaries are small but finite. Therefore, inflow and outflow boundary conditions need be specified in general. The inflow boundaries are specified as Dirichlet. Outflow boundaries are natural boundary conditions for the GFEM, and no boundary condition need be specified explicitly. For the CVFEM, the outflow boundary conditions are treated implicitly by integrating the outflow of advected energy using the current value of T, as in Equation (25).

In addition to this 'standard' sequence of grids, a sequence of spatially extended grids was also considered. These spatially extended grids cover $-1.5 \leqslant x, y \leqslant 1.5$, while retaining the same uniform grid spacing of the standard grids, $h = 1/10, 1/20$ and $1/40$, where h denotes the mesh spacing in either coordinate direction. These spatially extended grids were studied to compare the effects of discrete boundary conditions on the error analysis; boundary conditions errors should be negligible on the spatially extended meshes.

4.1.1. Metrics. Overall accuracy of the methods is measured by the error in $|1 - T_{\max}|$ $(= 1 - \max(T_{\text{numerical}}(i,j))|)$, which is a measure of peak resolution, and by the L_1 error norm, here defined by

$$\|\Delta T\|_1 = \frac{1}{N} \sum_{i,j} |T_{\text{exact}}(i,j) - T_{\text{numerical}}(i,j)| \tag{32}$$

Positivity and conservation are measured by

$$|T_{\min}| = |\min(T_{\text{numerical}}(i,j))| \tag{33}$$

and

$$|1-r| = \left| 1 - \frac{\sum_{i,j} T_{\text{numerical}}(i,j)}{\sum_{i,j} T_{\text{exact}}(i,j)} \right| \tag{34}$$

respectively. These metrics were used by Vreugdenhil and Koren [32], and allow comparison with the results they give.

We note here that we may potentially do GFEM methods a disservice by these choices for error norms. The GFEM provides the optimal numerical solution (T_h) in the energy norm, $\left(\int_\Omega \nabla(T-T_h) \bullet \nabla(T-T_h)\,\mathrm{d}\Omega\right)^{1/2}$, for the Poisson equation (see Reference [7], Section 2.2.6).

On the other hand, the energy norm does not have any special significance to the CVFEM. This is a continuing dilemma when comparing different numerical methods via numerical analysis. Gresho and Sani (Section 2.2.6) have verified, for particular solutions of the Poisson equation, that the bilinear GFEM outperforms CVFEM in the energy norm, but the 'GFEM does not win by much'. While we acknowledge these issues with respect to the norm, we will stay with the norms defined above as they at least have no preference to either method. We shall see that, aside from diffusion-dominated advection–diffusion problems discussed in a subsequent test problem, the GFEM schemes outperform CVFEM on the Molenkamp test problems.

4.1.2. Results. Time integration was performed using the second-order trapezoid rule. To ensure that the time integration error is rendered negligible compared to spatial error, multiple simulations were run with decreasing (fixed) time step values until convergence was noted. The spatial discretization utilized 4-node bilinear elements for both FEM and CVFEMs. Sample solutions are shown in Figure 3.

Figure 4 depicts the results of the error analysis on the Molenkamp problem. The legend labels results for the various methods studied with a different symbol. The solid lines correspond to results obtained on the standard grids ($-1 \leqslant x, y \leqslant 1$) and dashed lines are for simulations performed on the spatially extended grids ($-1.5 \leqslant x, y \leqslant 1.5$). On the standard grids the Gaussian distribution just grazes the boundaries, for example, attaining a value of 0.01 on the mid-sides of the mesh. This results in sufficient boundary interaction that the inflow and outflow boundaries must be carefully applied. On the extended meshes the computational boundaries are sufficiently remote that the effect of boundary condition treatment is negligible. The effects of boundary treatment will be evident with respect to the conservation metric.

The results for SUPG and SUCV are obtained by specifying a modified version of the Shakib τ parameter given in Equation (16). The τ parameter used in this particular problem excludes the time term, $(2\rho C_p/\Delta t)^2$. The reason for excluding this term is that it dominates as $\Delta t \to 0$. As noted above, these results were obtained with time steps small enough to eliminate time truncation error compared to spatial error. This is accomplished by computing solutions with decreasing time-step size until convergence is observed. With the unmodified Shakib parameter, the stabilization disappears in this time-step-refinement process. The Shakib parameter imposes a relationship between acceptable values of time-step size and spatial-mesh size. Instead, we chose to use the β factor to optimize the phase error characteristics of these stabilized methods for pure advection. The value $\beta = 2/\sqrt{15}$ is used to 'optimize' the SUPG

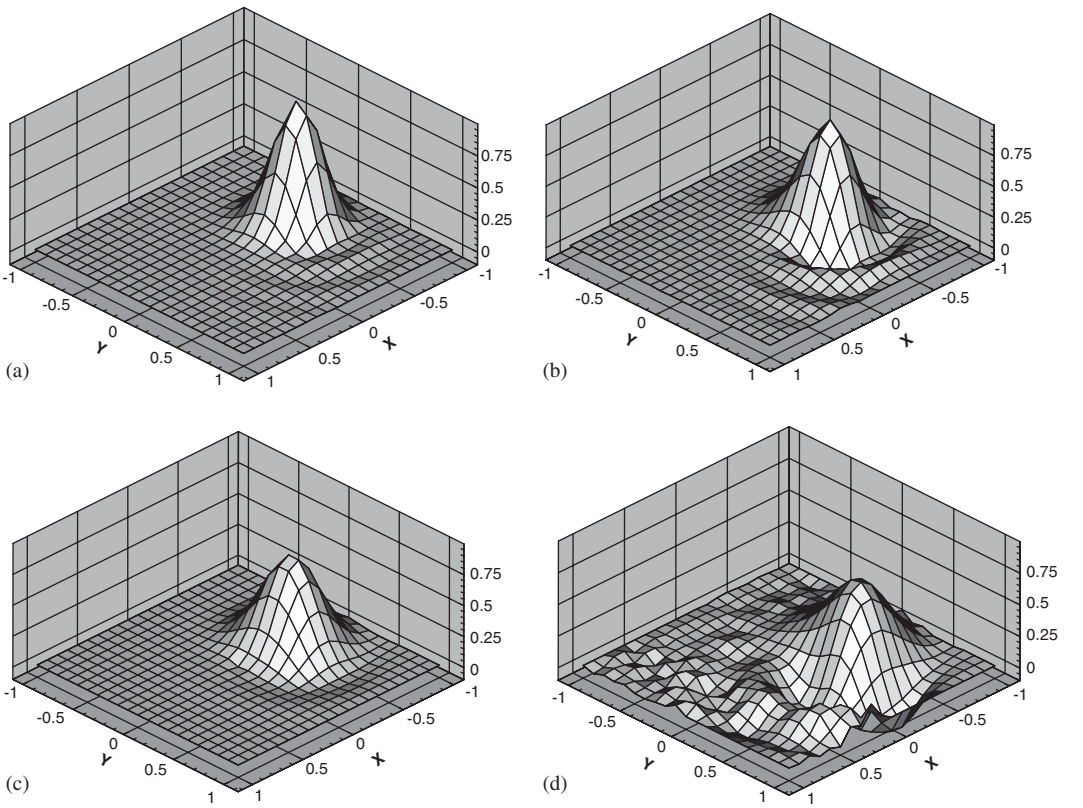

Figure 3. Solutions of Molenkamp problem after one revolution on a 20×20 uniform grid computed using: (a) GFEM; (b) CVFEM; (c) SUCV; and (d) MLGFEM.

for dispersion error; for SUCV $\beta = 1$ is used.¶ See Reference [20], for a discussion of how these values are optimal for reducing dispersion errors.

4.1.3. Accuracy. The L_1 (Figure 4(a)) and $T_{\max}$ error (Figure 4(b)) metrics measure plain accuracy of the methods, the former measuring global error while the latter peak resolution. The best global accuracy, as measured by the L_1 error, is obtained by the GFEM, by a significant margin. Next, in the order of decreasing accuracy are SUPG, CVFEM and SUCV, followed by the lumped mass versions of GFEM (MLGFEM) and CVFEM (MLCVFEM). These lumped mass versions produce nearly identical errors, suggesting that MLGFEM and MLCVFEM are nearly equivalent methods. In 1D, these two schemes are identical (see Reference [7]); in 2D the schemes are not identical, but they appear to be equivalent. These lumped methods display error levels of roughly 5 times CVFEM (and SUCV) and more than 2 orders-of-magnitude

¶The element mapping g_{ij} metric includes an embedded factor of $\frac{1}{2}$ (on these uniform grids), such that the effective β is $1/\sqrt{15}$ for SUPG and $\frac{1}{2}$ for SUCV.

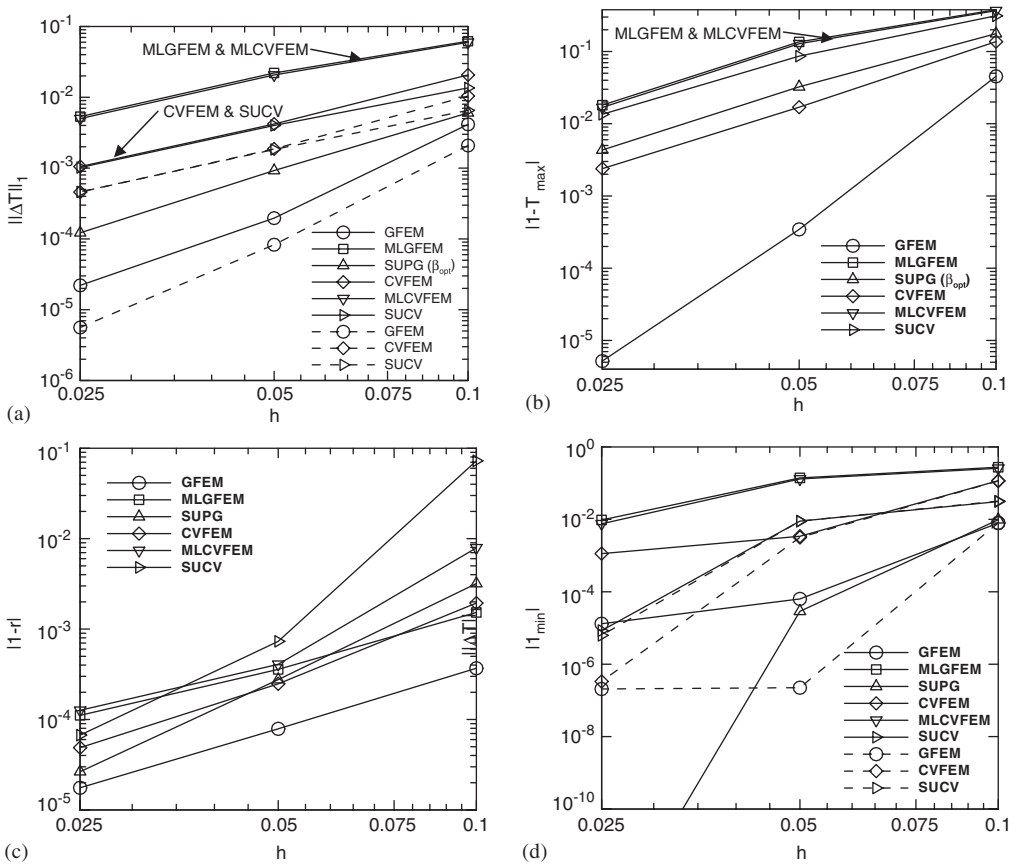

Figure 4. Error analysis of Molenkamp problem in terms of: (a) L_1 norm; (b) $T_{\max}$ norm; (c) conservation norm; and (d) positivity norm. Full and dashed lines denote results from standard and extended meshes, respectively.

larger than GFEM on the finest grids. Also, on the finest (standard) grid, the SUPG L_1 error is about 10 times less than CVFEM and SUCV, which converge to the same error on this mesh. Errors due to boundary effects are notably reduced on the extended meshes when measured with the L_1 norm (Figure 4(a)) for both GFEM and CVFEM.

The deleterious effects of introducing artificial diffusion ($O(h^3)$) in the SUPG and SUCV are manifest in the peak resolution, measured by the $T_{\max}$ norm. GFEM and CVFEM have no artificial diffusion and produce the best peak resolution, although the GFEM clearly outperforms CVFEM. SUPG produces nearly twice the error of CVFEM, while the SUCV error is about 3 times that of SUPG.

The boundary conditions have little effect on the $T_{\max}$ metric, hence, results on the extended grids are not shown for this metric.

4.1.4. Conservation. The conservation error norm, defined in Equation (34), measures the conservation of (thermal) energy with respect to the exact solution (It is not a measure of numerical conservation.). Figure 4(c) displays the conservation error on the standard grids. Conservation error is not shown in Figure 4(c) for the extended meshes because, except on the coarsest mesh, they are zero to better than six significant figures for all methods (not including the mass-lumped schemes). This result confirms the assertion that the mass conservation errors on the standard grid are largely due to the approximate nature of the boundary conditions; inflow does not equal outflow because of the inherent approximations in the discrete boundary conditions. Once again, the GFEM yields the least error in terms of conservation error as shown in Figure 4(c). It is perhaps surprising that the CVFEM produces more error than GFEM, since the CVFEM is based on a conservative discretization. Apparently, the outflow boundary condition for GFEM conserves better than the outflow boundary condition of CVFEM. The conservation norm for SUPG (SUCV) shows convergence to GFEM (CVFEM) on the finest mesh, but the stabilization terms have a deleterious effect at the coarser grids levels. On the finest grid, the mass-lumped methods have the most conservation error, while the error in the stabilized methods approach their corresponding unstabilized version. It should be noted that the stabilized methods are intended to improve performance with respect to nonphysical oscillations on coarse grids. This analysis shows that for pure advection the conservation of these methods is compromised (even more than mass lumping) on coarse grids.

4.1.5. Positivity. Neither of the methods considered is positive as indicated in Figure 4(d) for $|T_{\min}|$ (The actual values are negative.), however, the SUPG comes close on the finest grid where it yields $T_{\min} = -1.69 \times 10^{-14}$, basically machine precision. The 'least' positive methods are the mass-lumped schemes. The GFEM results are much better on the same mesh than the CVFEM values. The extended grid produces much-improved results over the standard grid for GFEM, except for the coarsest mesh. For CVFEM, improvement is only seen for the finest extended grid. SUCV produces the same results on the standard and extended grids.

4.1.6. Apparent convergence rates. Based on results from the two smallest mesh sizes of the standard grids, the rates of spatial convergence are shown in Table I for the various methods, based on the L_1 error norm.

The rates are all second-order, except for GFEM and SUPG which attain third-order. Although the convergence rates may be similar, the absolute error is much greater for the mass-lumped schemes, illustrating once again the deleterious effects of mass lumping on numerical dispersion in numerical schemes for pure advection.

Table I. Apparent spatial convergence rates.

Method	Convergence rate
GFEM	3.16
MLGFEM	2.05
SUPG	2.92
CVFEM	2.00
ML-CVFEM	2.04
SUCV	1.97

4.1.7. Summary remarks on Molenkamp problem. Unless positivity is an overriding requirement, plain GFEM is the best method on this particular problem, and by a substantial margin. It is well known that this will not hold true on problems which are not as smooth and well resolved as the Gaussian distribution studied here. However, on this resolved, nonsingular profile, the GFEM performance is excellent. SUPG also displays very good overall performance; however, it is not able to show its full potential because of the highly resolved nature of this smooth problem.

This study also indicates the potential effectiveness of the SUCV, on a sufficiently spatially refined grid. The overall (L_1) error and mass conservation error of CVFEM is retained, and the negative values are much reduced. However, the latter is only reduced to the level of GFEM, at least on the present problem.

In general, given sufficient mesh resolution, all of these methods are capable of yielding quite good accuracy for pure advection of a *smooth* initial profile if the *consistent* mass matrix is retained. The performance is much degraded, both in terms of dispersion and conservation errors, with mass lumping—this is not a new finding.

4.2. Modified Molenkamp problem

4.2.1. Problem definition. As noted above, the efficacy of the stabilized methods cannot be demonstrated fairly on the Molenkamp problem owing to its smoothness and the good resolution of the grids considered. In this section we make the problem 'harder' by specifying a sharper initial profile, leaving other problem specifications unchanged. The steepness of the initial profile is chosen such that the profile is not well resolved on the coarse (20×20) grid but is reasonably resolved on the fine (80×80) grid. The initial profile is given by

$$T = \begin{cases} 1 & r < R_1 \\ \dfrac{1}{2}\left(1 - \tanh\left(\dfrac{2b}{R_2 - R_1} r - \dfrac{R_1 + R_2}{2}\right)\right) & R_1 < r < R_2 \\ 0 & r > R_2 \end{cases} \tag{35}$$

with r given in Equation (31), and $b = 3$, $R_1 = 0.1$ and $R_2 = 0.4$. This results in a profile with compact support on the computational domain, thus the inflow boundaries are specified as homogeneous Dirichlet boundaries. Outflow boundaries are specified on the remaining domain boundary. The initial profiles on each grid are shown in Figure 5 and illustrate the variation in resolving power of each grid.

Results of the error analysis are shown in Figure 6, using the same metrics as in the Molenkamp problem, except for the $T_{\max}$ norm, which is replaced by the standard infinity norm (L_∞), since the former is more appropriate for the Gaussian profile. The present profile is flat in the vicinity of its centroid.

Results shown by the error analysis can be further understood when compared with the solution profiles after one revolution on the grid. The resulting solution profiles on the 20×20 grid, given in Figure 7, have lost their shape after just one revolution, resembling a Gaussian profile more than the initial profile (which, in pure advection, is the exact solution after one revolution). The SUCV displays the most artificial diffusion and the CVFEM the most dispersion error. These solutions were obtained with the same numerical schemes applied in the standard Molenkamp problem, including using the modified form of stabilizing parameter

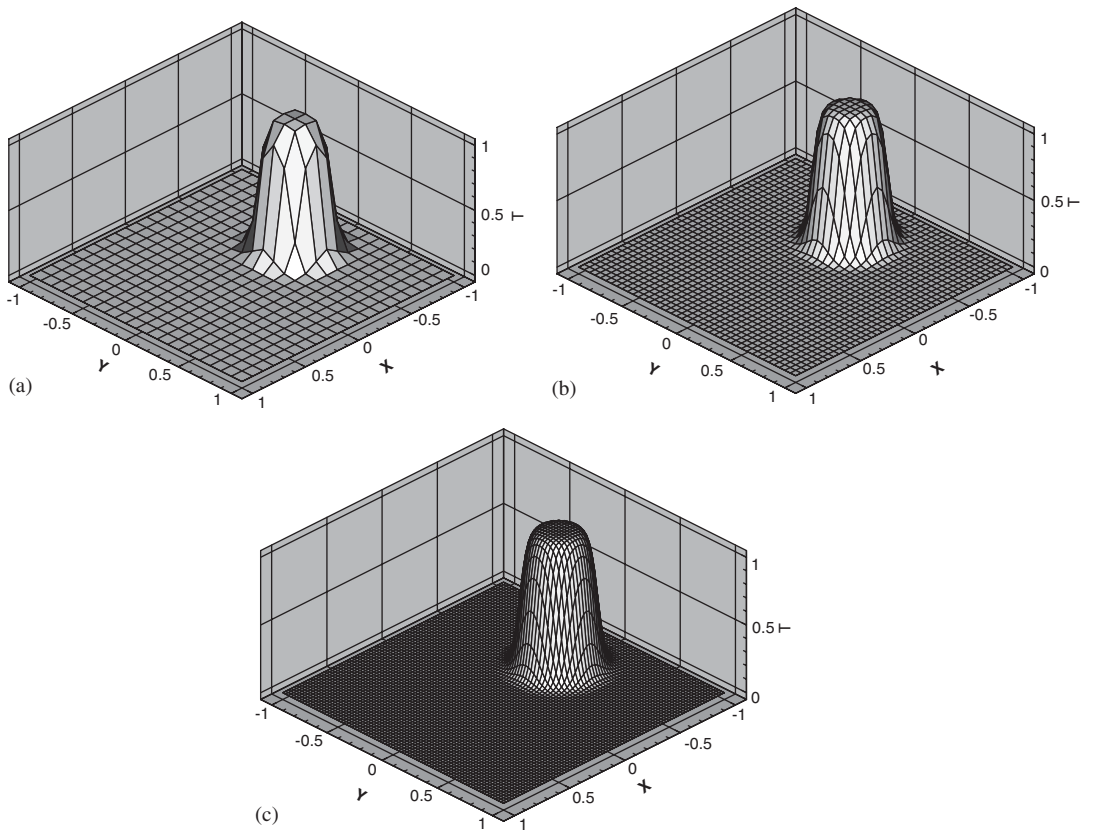

Figure 5. Initial solution representation for modified Molenkamp problem on: (a) 20×20; (b) 40×40; and (c) 80×80 grids.

as discussed in Section 4.1.2. Furthermore, the un-stabilized schemes (CVFEM and GFEM) result in much phase error and thus many 'wiggles' in the resulting solution. The stabilized methods, which contain artificial diffusion, tame the wiggles, but at the expense of diffusing the profiles.

On the 40×40 grid (Figure 8), GFEM and SUPG are beginning to resolve the 'flat spot' at the centroid of the profile, although this mesh is still too coarse for these methods to resolve it well. The CVFEM and GFEM still produce significant phase error as indicated by the wiggles. The $T_{\min}$ norm also shows that these two methods produce the more negative values on this grid. SUCV is not much better in terms of $|T_{\min}|$ (a single point norm), but in terms of the 'eyeball' norm it is seen to have tamed the wiggles. Note that the SUPG produces both leading and lagging phase error, whereas the SUCV displays only lagging phase error.

The solution profiles on the 80×80 grid (Figure 9) for GFEM and especially SUPG are a reasonably good facsimile of the analytical solution (cf. Figure 5). CVFEM and SUCV are not as accurate at this grid spacing.

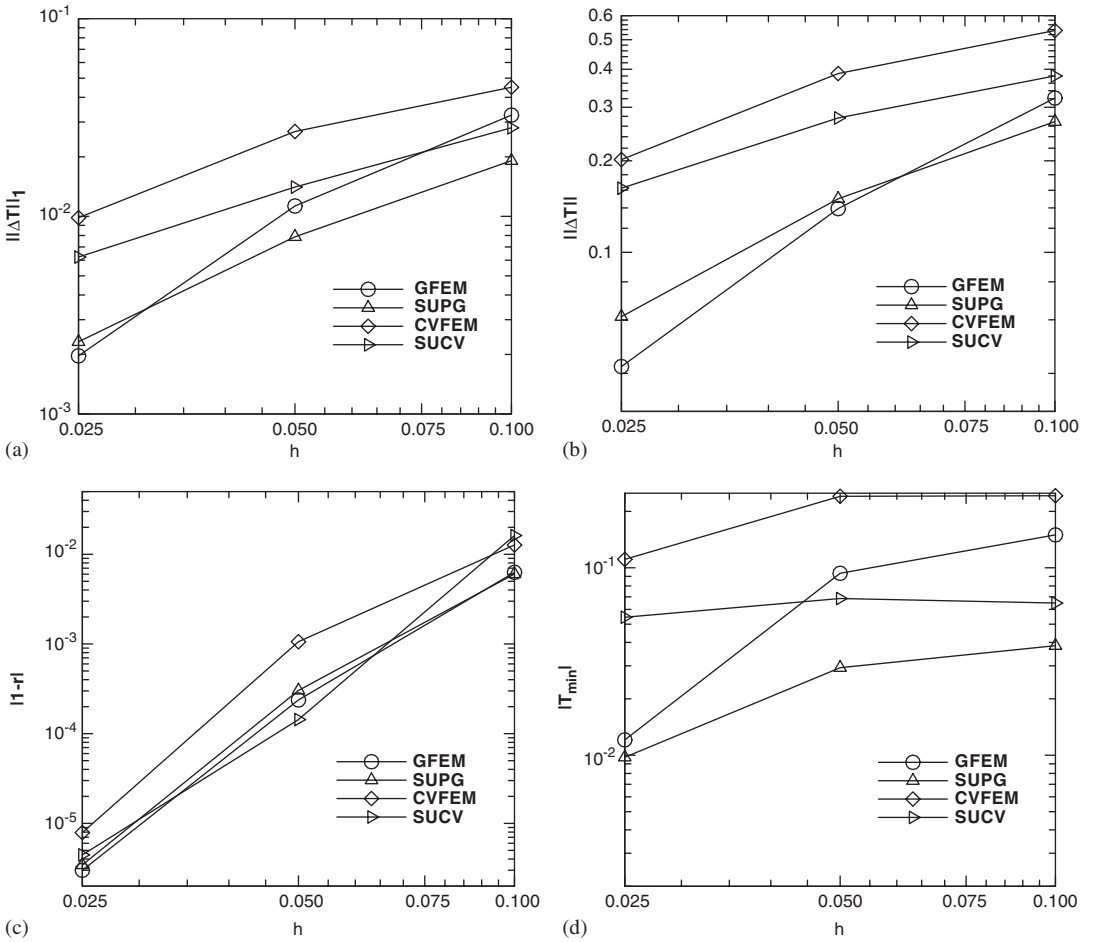

Figure 6. Error analysis of the modified Molenkamp problem in terms of: (a) L_1 norm; (b) L_∞ norm; (c) conservation norm; and (d) positivity norm.

4.2.2. Accuracy. As shown in Figure 6(a), SUCV represents an improvement in accuracy on all grids when compared to CVFEM. Similarly, SUPG outperforms SUCV on all grids. GFEM, however, crosses over in going from the coarse to finest grids, ending up with similar error as SUPG on the 80×80 grid. For GFEM and CVFEM, errors in this problem derive solely from dispersion (phase and group speed errors). Fourier analysis [7, 20, 21] of these methods predicts greater phase error for the latter method; this agrees with present results. Considering Figure 2.6-15 in Reference [7], which shows dispersion error as a function of dimensionless wave number (k) or wave length (λ),

$$\frac{k\Delta x}{\pi} = \frac{2\Delta x}{\lambda}$$

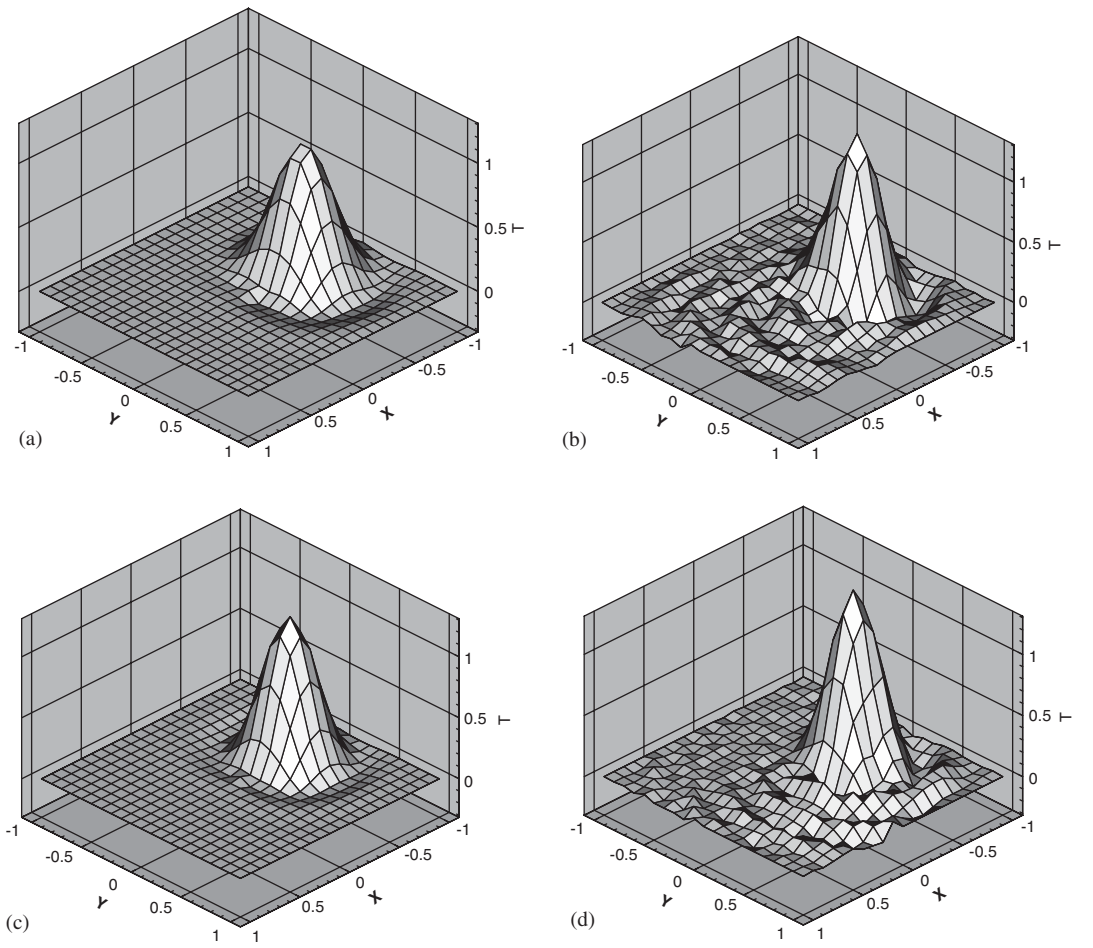

Figure 7. Solutions of modified Molenkamp problem after one revolution on the 20×20 grid from: (a) SUCV; (b) CVFEM; (c) SUPG; and (d) GFEM.

one can estimate the grid resolution necessary to maintain accuracy with respect to the frequency content of the initial profile. From Equation (35), a rough estimate for the characteristic wavelength in the initial profile is to take $2\lambda = R_2 - R_1 = 0.3$. This leads to $2\Delta x/\lambda = 1.33$, 0.67 and 0.33 on the 20×20, 40×40 and 80×80 grids, respectively. With respect to the phase error function from Fourier analysis, this would suggest that the 20×20 grid will produce significant dispersion error (up to and beyond the Nyquist frequency, $2\Delta x/\lambda = 1$), whereas the 80×80 grid may result in less than 10% phase error for all methods. On a qualitative basis, this is roughly what resulted; the 20×20 grid produces much phase error whereas the 80×80 grid produces reasonably good solutions (this is in fact how we chose the 'width' of the mollifier function in the initial profile). It must be noted that the Fourier analysis requires simplifying assumptions (unidirectional propagation of monochromatic waves) compared to

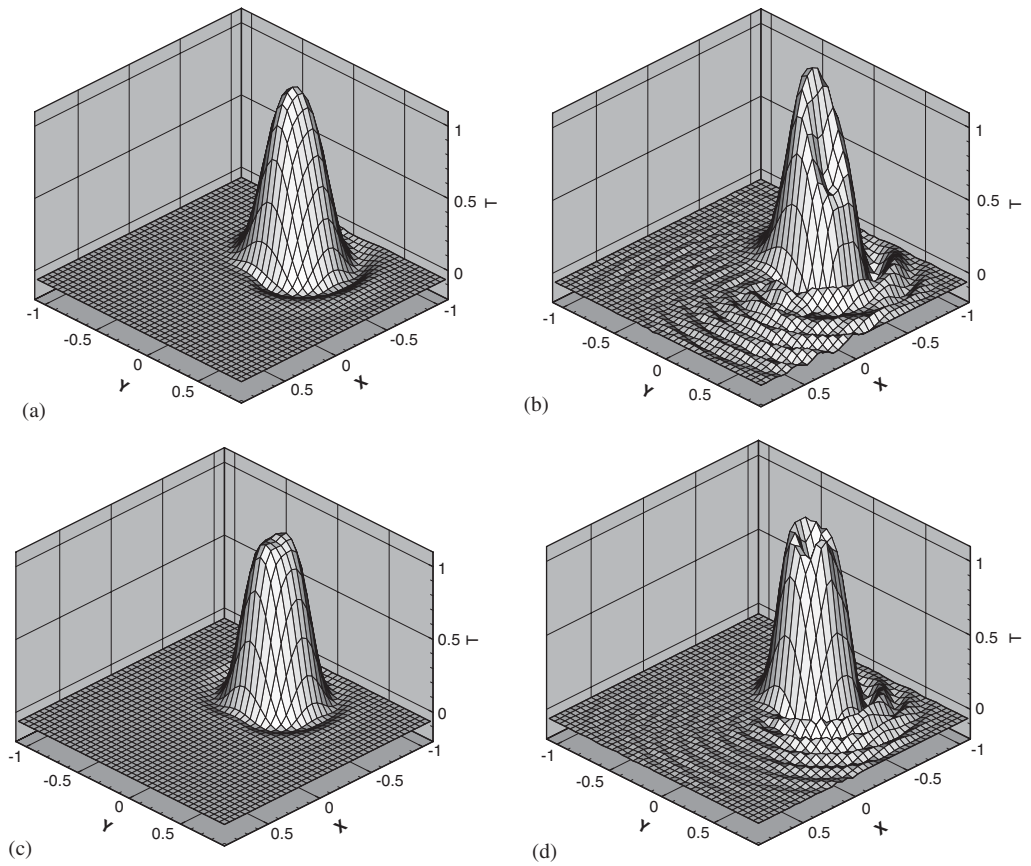

Figure 8. Solutions of modified Molenkamp problem after one revolution on the 40×40 grid from: (a) SUCV; (b) CVFEM; (c) SUPG; and (d) GFEM.

the conditions in the present simulations, and so the results from the former analysis cannot be expected to necessarily match quantitatively.

4.2.3. Conservation. As noted in the discussion of the standard Molenkamp problem, conservation errors for this problem are largely due to error at the boundary. Inflow boundaries were specified as Dirichlet with zero grid-point values for the solution variable. Wiggles generated in the interior which propagate to the boundaries produce error in the outflow regions of the grid. The conservation norm displays a monotone decrease as the grid is refined. Plots of the solutions (Figures 7–9) after one revolution also show the much reduced boundary error with grid refinement.

4.2.4. Positivity. In terms of positivity, it is clear why the stabilized methods, SUPG and SUCV, are preferred on coarse grids. These methods reduce the wiggles, as can be seen in Figures 7–9, and quantitatively in the $T_{\min}$ norm. These methods also improve overall accuracy

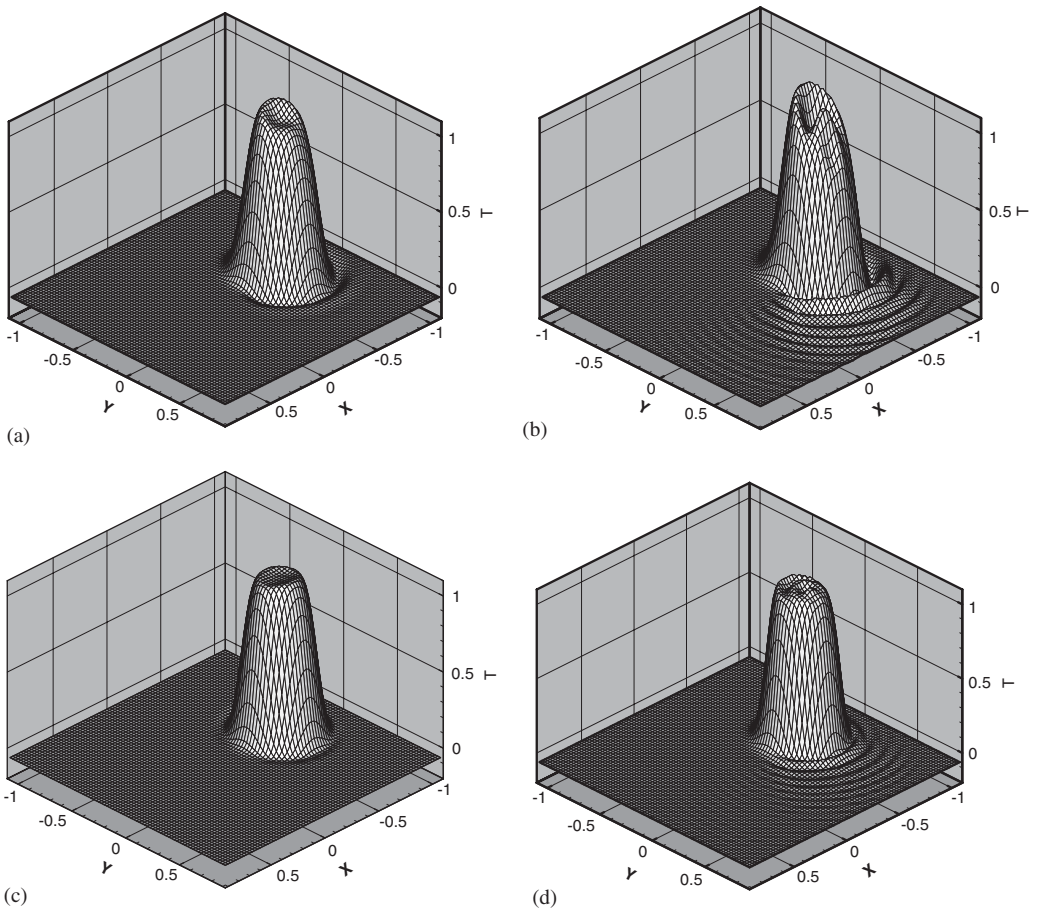

Figure 9. Solutions of modified Molenkamp problem after one revolution on the 80×80 grid from: (a) SUCV; (b) CVFEM; (c) SUPG; and (d) GFEM.

relative to their unstabilized counterparts. The exception to this is GFEM. On the finest grid, GFEM and SUPG produce very similar error levels.

4.2.5. Summary remarks on modified Molenkamp problem. On the same grid, GFEM outperforms CVFEM, at least for the grids considered here. In the asymptotic limit $\Delta x \to 0$, both should be second-order accurate, as shown in the standard Molenkamp problem. These trends holds true for SUPG vs SUCV. On coarse grids, the stabilized methods improve plain accuracy in addition to positivity, although none of these methods is strictly positive.

4.3. Smith and Hutton problem

The Smith and Hutton [33] problem involves the steady advection–diffusion of a scalar in a circular flow field given by,

$$u = 2y(1 - x^2), \quad v = -2x(1 - y^2) \tag{36}$$

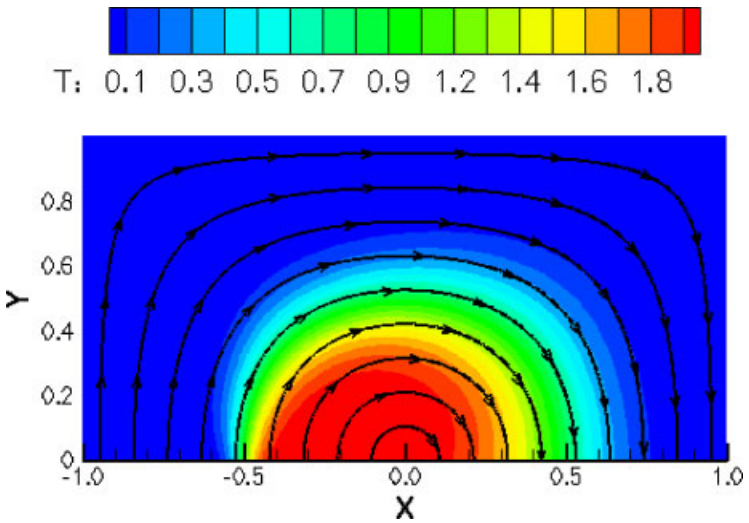

Figure 10. Temperature and streamlines for Smith and Hutton problem.

in a rectangular domain $\Omega = \{(x, y) : -1 < x < 1, 0 < y < 1\}$. A steep inlet temperature profile is specified as

$$T(x, 0) = 1 + \tanh[\alpha(2x + 1)], \quad -1 \leqslant x \leqslant 0 \tag{37}$$

with $\alpha = 10$. The task is to transport (by advection and diffusion) this profile through 180° for various values of the diffusion coefficient. Except for the outflow boundary, the remainder of the computational boundary is specified with the compatible Dirichlet condition, $T = 1 - \tanh(10)$. The outflow boundary is located on $y = 0$ for $0 < x < 1$. We will follow Morton's specification [28], and compute solutions for $\kappa \equiv k/\rho C = 10^{-6}, 0.002, 0.01$ and 0.1, on a uniform 20×10 grid. On this standard grid, these diffusivities result in mesh Peclet numbers, $Pe = \|\mathbf{u}\|h/\kappa = 10^5, 50, 10$ and 1, respectively, where $\|\mathbf{u}\| = 1$ and $h = \Delta x = \Delta y$ denotes the mesh size. Figure 10 depicts the (specified) streamline pattern and temperature distribution for the case $\kappa = 0.01$ computed on a 160×80 grid.

4.3.1. Results. Table II shows estimated L_1 errors for solutions computed on a uniform 20×10 grid ($h = 1/10$). The benchmark solutions used to calculate these error norms were computed on highly refined grids, with $h = 1/80$ for all but the case of $\kappa = 10^{-6}$ which utilized a finer grid, $h = 1/160$. There are several notable results. For each method, the error increases with (mesh) *Pe* number. For a given value of *Pe* number, the error norms are similar (less than a factor of 2 difference) for all methods. At low *Pe* number, the SUCV scheme performs best, while the GFEM scheme does the best at large *Pe* number. However, the differences are modest.

Figure 11 compares the solutions obtained on the 20×10 mesh with the benchmark solutions used for computing L_1 error norms presented in the table above. As noted earlier, there is no remarkable difference in the solutions given by the various methods. Furthermore, the figure illustrates that the stabilized methods do not display a significant improvement in the solutions.

Table II. L_1 error ($\times 100$) for 20×10 grid.

Method	$Pe = 1$	10	50	10^5
CVFEM	0.368	0.586	1.41	2.95
SUCV	0.322	0.548	2.21	4.18
GFEM	0.420	0.686	1.30	2.55
SUPG	0.393	0.613	2.04	3.83

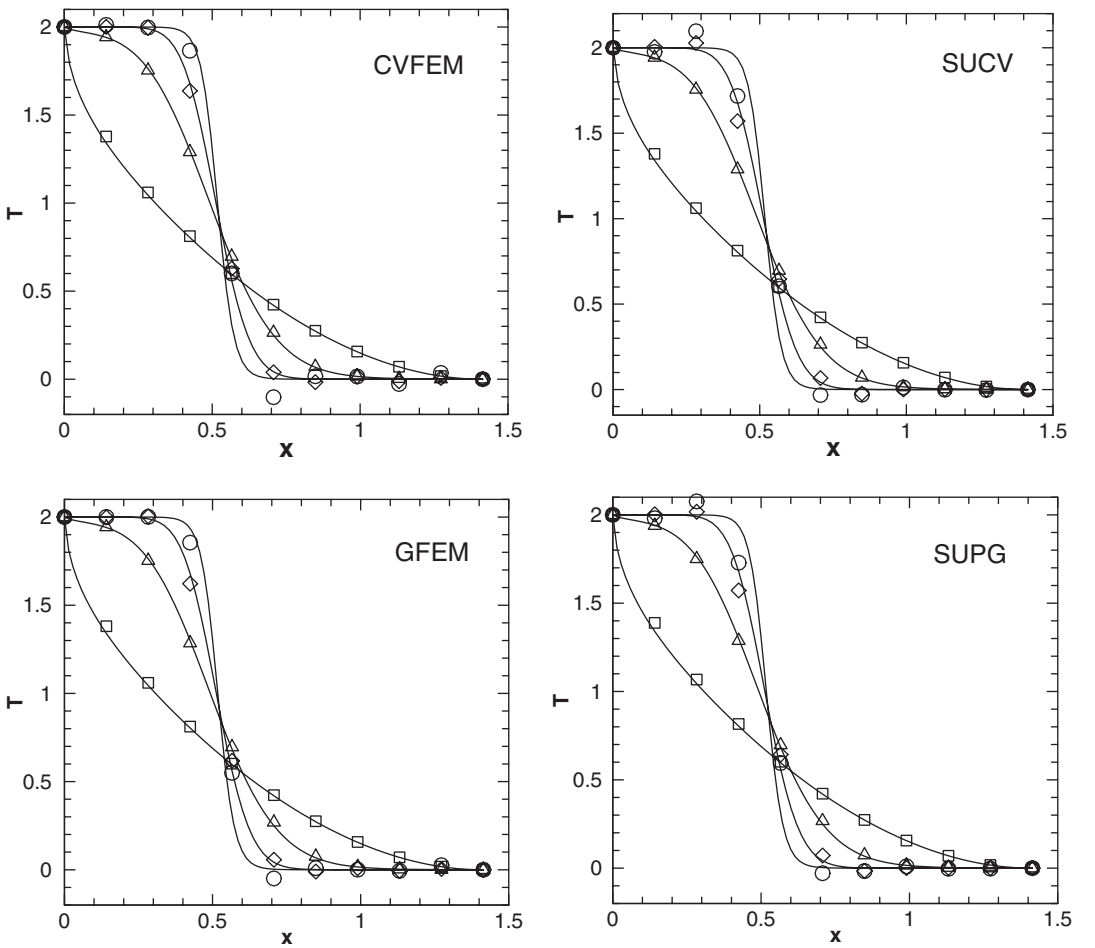

Figure 11. Solution profiles along the diagonal (0,0)–(1,1) of the 20×10 mesh with differing symbols for each value of diffusion coefficient, $\kappa = 10^{-6}$ (circles), 0.002 (diamonds), 0.01 (triangles), and 0.1 (squares). The full lines are the benchmark solutions, computed on highly refined grids.

In fact, it can be argued that the unstabilized methods give better solutions at the higher *Pe* numbers, where the stabilized methods should provide an advantage. It should be noted that the fine grid (benchmark) solutions at $\kappa = 10^{-6}$ were problematic to obtain with the iterative solvers using ILU preconditioning. The reference solutions in this case were obtained with the stabilized schemes on a 320×160 grid ($h = 0.00625$).

4.3.2. Iterative solution characteristics. Although, in terms of solution error, the stabilized methods show little advantage, they do make the linear systems easier to solve. Table III lists the number of iterations required to solve the linear systems using GMRES iteration with ILU(0) preconditioning (except for $Pe = 10^5$, see below) from the Aztec solver package [30]. The convergence criteria was that the initial residual be reduced by a factor of 10^{-8}; all solves were started with a zero initial value everywhere on the grid. Except for the largest Peclet number (smallest diffusivity), the iteration counts are very similar for CVFEM vs GFEM and SUCV vs SUPG.

The utility of the stabilized schemes is manifested in the large *Pe* number cases, especially for $Pe = 10^5$ ($\kappa = 10^{-6}$). In this case, with vanishing diffusion, the diffusion operators are all but annihilated. For the GFEM and CVFEM schemes, this leaves only the advection operators in the stencils, which are skew-symmetric (see, for example, Appendix A of Reference [20], and therefore have zero diagonal. For this reason, this case could not be solved using ILU(0)/GMRES; the ILU(0) preconditioner could not be applied on the skew-symmetric matrices. Instead, for CVFEM and GFEM this case was solved using a third-order least-squares polynomial (LS3) preconditioner with GMRES. The SUPG and SUCV matrices include artificial diffusion, therefore having nonzero diagonal entries, and were solved with ILU(0)/GMRES. The table shows an enormous increase in the work required to solve the linear systems using GFEM and CVFEM (LS3/GMRES) compared to using SUPG and SUCV (ILU/GMRES) at large *Pe* number.

Table IV compares iteration counts for GFEM and CVFEM as a function of grid spacing (and of mesh Peclet number) using ILU(0)/GMRES and requiring a residual reduction tolerance of 10^{-8}. Results from the stabilized versions are not included, as they produce similar results to the unstabilized methods as the grids are refined (grid refinement diminishes artificial diffusion). GFEM and CVFEM display similar iteration counts as a function of grid size and diffusion coefficient. For the diffusion-dominated case ($\kappa = 0.1$), the iteration count (roughly) doubles as the mesh size is halved. This behaviour is similar to stationary iteration methods (where the iteration matrix is independent of the iteration history), whose

Table III. Linear solve iterations on 20×10 grid.

	Method			
Peclet No.	CVFEM	SUCV	GFEM	SUPG
1	10	10	11	11
10	9	7	8	7
50	36	8	35	8
10^5	504*	8	563*	9

*Solved with LS3 preconditioner; all others solved with ILU (0) preconditioner.

Table IV. Iteration count as a function of mesh Peclet number.

	$\kappa = 0.1$			$\kappa = 0.002$		
h	Pe	GFEM	CVFEM	Pe	GFEM	CVFEM
0.1	1.0	11	10	50	35	36
0.05	0.5	20	19	25	18	19
0.025	0.25	38	37	12.5	13	14
0.0125	0.125	77	91	6.25	17	17
0.00625	0.0625	151	193	3.125	37	36

convergence depends on the spectral radius of the iteration matrix (e.g. successive overrelaxation). The spectral radius increases as the grid size decreases. On the other hand, the iteration count distribution for the convection-dominated case ($\kappa = 0.002$) displays a minimum for $h = 0.025$ (80×40 grid). This is because the grid Peclet number decreases with decreasing mesh size, changing the (numerical) problem from advection-dominated towards a diffusion-dominated problem.

4.3.3. A note on convergence rate. In the course of performing the error analysis, it was noted that the rate of convergence depends on the 'smoothness' of the solution, which in turn depends on Peclet number. For example, the profiles in Figure 11 for $\kappa = 0.1$ display an irregularity at the origin, where they appear to have an infinite slope. Profiles for other (smaller) values of diffusivity are smooth. The steep profile is due to the large amount of diffusion away from the specified Dirichlet value imposed on $x < 0, y = 0$.

The order of convergence can be estimated if solutions on three grids are available (e.g. Reference [34]). For the square grids used here, we assume the truncation error behaves as

$$T_h(x, y) = T(x, y) + c(x, y)h^p$$

where T_h represents the numerical solution, T the exact solution and c is a constant independent of h. With a known convergence rate, p, this formula can be used to perform Richardson extrapolation pointwise on the grid. We can also use this truncation error formula to form a bound on the error norm between two numerical solutions computed on different grids, with grid spacings h_i

$$\|T_{h_1} - T_{h_2}\| \leqslant C(h_1^p - h_2^p) \tag{38}$$

If we have solutions on three grids, and if the refinement factor is uniform ($h_1/h_2 = h_2/h_3 = r$), then the order of the method can be estimated from

$$p = \log(\|T_{h_2} - T_{h_3}\|/\|T_{h_1} - T_{h_2}\|)/\log r \tag{39}$$

Table V shows convergence rates computed with Equation (39) using the L_1 error norm. The table shows 3 values for each method in each case. They correspond to the rates estimated from Equation (39) for each triplet of grids (coarse to fine from left to right) from the five grids on which the solution was computed, viz. $h = 0.1, 0.05, 0.025, 0.0125$ and 0.00625. For example, a convergence rate of 1.48 is obtained for GFEM using the solutions on grids 20×10, 40×20 and 80×40 for the case $\kappa = 0.1$.

Table V. Convergence rates.

Method	$\kappa = 0.1$	$\kappa = 0.002$
GFEM	1.48, 1.25, 1.18	3.02, 2.06, 1.97
CVFEM	1.35, 1.11, 1.06	3.00, 2.24, 2.02

The case for $\kappa = 0.002$ asymptotes to second-order, in agreement with the mathematical truncation error formula for these methods [20]. Note that third-order is suggested on the coarser meshes; this indicates these grids are not in the asymptotic convergence region of these methods. In contrast, the diffusion-dominated case ($\kappa = 0.1$) displays less than second-order rate of convergence. Even though the solution is smooth throughout most of the computational domain, the irregular region noted above will eventually dominate the error norm as the mesh is refined. The results for $\kappa = 0.1$ suggest the rate of convergence in this irregular region is tending to first-order for both GFEM and CVFEM.

This analysis demonstrates the diligence required in choosing a problem for which to perform verification of numerical methods. The test problems chosen must be free of singularities or other irregularities that can give a false impression of the convergence rate of a numerical scheme.

5. CONCLUDING REMARKS

In this work we endeavoured to carefully compare the control volume finite element method to traditional Galerkin finite element methods, including so-called stabilized formulations for both. We discussed the formulation of the control volume finite element methods in terms of weighted residual methods with generalized weight function. The resulting methods are locally (control-volumewise) conservative by construction, although the SUCV conserves a perturbed physical flux and therefore includes an inherent (energy) conservation error owing to the stabilization. This is also true of SUPG. It is shown that global numerical conservation in the presence of Dirichlet boundaries requires consistent post-processing of fluxes on the Dirichlet boundaries, in a similar way that Galerkin finite element methods require such post-processing of fluxes.

The error analysis indicates that there is no apparent advantage in using the locally conservative CVFEMs over FEMs for *linear* advection–diffusion problems. There is a large literature on the desirable feature of local conservation in numerical schemes. For linear advection–diffusion processes, there are no general arguments to require discrete conservation. If two methods are both consistent and $O(h^2)$, then with proper mesh resolution, they should both converge to the exact solution. Besides, in either method exact conservation is only approximated to $O(h^2)$.

Unfortunately, these observations cannot be extended to nonlinear problems. For general nonlinear problems, numerical conservation may be more important, certainly degree of positivity often is. For example, problems involving real material equations of state require composition variables (mass fractions) to be positive; solutions such as those depicted in Figure 7, if representing composition, would be catastrophic without special treatment in an equation of state function.

A summary of the more notable findings of the numerical analysis are indicated in the following list.

For linear transient advection:

- Stabilized methods improve accuracy on coarse grids and positivity on coarse and fine grids.
- There is no apparent advantage to the locally conservative CVFEM over the GFEM for *linear* advection–diffusion problems.
- GFEM is hard to beat on *resolved* problems, and stabilized variants do not offer much of an improvement.

For linear steady advection–diffusion:

- Stabilized methods reduce iteration counts significantly for advection-dominated problems.
- No significant difference emerged between CVFEM and GFEM.

ACKNOWLEDGEMENTS

This work was supported under the ASCI Advanced Spatial Discretization project. Sandia is a multi-program laboratory operated by Sandia Corporation, a Lockheed Martin Company for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000.

REFERENCES

1. Baliga BR, Patankar SV. A new finite-element formulation for convection–diffusion problems. *Numerical Heat Transfer* 1980; **3**:393–409.
2. Ramadhyani S, Patankar SV. Solution of the Poisson equation: comparison of the Galerkin and control-volume methods. *International Journal for Numerical Methods in Engineering* 1980; **15**:1395–1418.
3. Ramadhyani S, Patankar SV. Solution of the convection–diffusion equation by a finite-element method using quadrilateral elements. *Numerical Heat Transfer* 1985; **8**:595–612.
4. Schneider GE, Raw MJ. A skewed, positive influence coefficient upwinding procedure for control-volume-based finite-element convection–diffusion computation. *Numerical Heat Transfer* 1985; **9**:1–26.
5. Saabas HJ, Baliga BR. Co-located equal-order control-volume finite-element method for multidimensional incompressible fluid flow—part I: formulation. *Numerical Heat Transfer*, *Part B* 1994; **26**(4):381–407.
6. Saabas HJ, Baliga BR. Co-located equal-order control-volume finite-element method for multidimensional incompressible fluid flow—part II: verification. *Numerical Heat Transfer*, *Part B* 1994; **26**(4):409–424.
7. Gresho PM, Sani RL. *Incompressible Flow and the Finite Element Method*: *Volume One*: *Advection–Diffusion*. Wiley: New York, 2000.
8. Dalen V. Simplified finite-element models for reservoir flow problems. *Society of Petroleum Engineers Journal* 1979; **19**(5):333–343.
9. Forsyth PA. Control volume finite element approach to NAPL groundwater contamination. *SIAM Journal on Scientific and Statistical Computing* 1991; **12**(5):1029–1057.
10. Letniowski FW, Forsyth PA. A control volume finite element method for three-dimensional NAPL groundwater contamination. *International Journal for Numerical Methods in Fluids* 1991; **13**(8):955–970.
11. Forsyth PA, Wu YS, Pruess K. Robust numerical methods for saturated-unsaturated flow with dry initial conditions in heterogeneous media. *Advances in Water Resources* 1995; **18**:25–38.
12. Neises J, Steinbach I. Finite element integration for the control volume method. *Communications in Numerical Methods in Engineering* 1996; **12**:543–555.
13. Lohner R. *Applied Computational Fluid Dynamics Techniques*. Wiley: New York, 2001.
14. Comini G, Del Giudice S, Nonino C. Energy balances in CVFEM and GFEM formulations of convection-type problems. *International Journal for Numerical Methods in Engineering* 1996; **39**:2249–2263.
15. Hughes TJR, Engel G, Mazzei L, Larson MG. The continuous Galerkin method is locally conservative. *Journal of Computational Physics* 2000; **163**:467–488.
16. Swaminathan CR, Voller VR. Streamline upwind scheme for control-volume finite elements, part 1. Formulations. *Numerical Heat Transfer*, *Part B* 1992; **22**:95–107.

17. Swaminathan CR, Voller VR. Streamline upwind scheme for control-volume finite elements, Part 2. Implementation and comparison with the SUPG finite-element scheme. *Numerical Heat Transfer*, *Part B* 1992; **22**:109–124.
18. Brooks AN, Hughes TJR. Streamline upwind/Petrov–Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering* 1980; **32**:199–259.
19. Hughes TJR, Mallet M. A new finite element formulation for computational fluid dynamics: III. The generalized streamline operator for multidimensional advective–diffusive systems. *Computer Methods in Applied Mechanics and Engineering* 1986; **58**:305–328.
20. Christon MA, Martinez MJ, Voth TE. Generalized Fourier analysis of the advection–diffusion equation—Part I: one-dimensional domains. *International Journal for Numerical Methods in Fluids* 2004; **45**:839–887.
21. Voth TE, Martinez MJ, Christon MA. Generalized Fourier analysis of the advection–diffusion equation—Part II: two-dimensional domains. *International Journal for Numerical Methods in Fluids* 2004; **45**:889–920.
22. Hughes TJR. *The Finite Element Method*. Prentice-Hall Inc: Englewood Cliffs, NJ, 1987.
23. Hughes TJR, Franca LP, Hulbert GM. A new finite element formulation for computational fluid dynamics: VIII. The Galerkin least-squares method for advective–diffusive equations. *Computer Methods in Applied Mechanics and Engineering* 1989; **73**:173–189.
24. Shadid JN, Salinger AG, Schmidt RC, Smith TM, Hutchinson SA, Hennigan GL, Devine KD, Moffat HK. MPSalsa version 1.5: a finite element computer program for reacting flow problems. *SAND98-2864*, Sandia National Laboratories, Albuquerque, New Mexico, 1996; 62.
25. Swaminathan CR, Voller VR, Patankar SV. A streamline upwind control volume finite element method for modeling fluid flow and heat transfer problems. *Finite Elements in Analysis and Design* 1993; **13**:169–184.
26. Gray WG, Leijnse A, Kolar RL, Bain CA. *Mathematical Tools for Changing Spatial Scales in the Analysis of Physical Systems*. CRC Press Inc: Boca Raton, FL, 1993.
27. Shakib F, Hughes TJR, Johan Z. A new finite element formulation for computational fluid dynamics: X. The compressible Euler and Navier–Stokes equations. *Computer Methods in Applied Mechanics and Engineering* 1991; **89**:141–219.
28. Morton KW. *Numerical Solution of Convection–Diffusion Problems*. Chapman & Hall: London, 1996.
29. Gresho PM, Lee RL, Sani RL. On the time-dependent solution of the incompressible Navier–Stokes equations in two and three dimensions. *Recent Advances in Numerical Methods in Fluids*, vol. 1. Pineridge Press Ltd.: Swansea, Wales, U.K., 1980; 27–79 (Chapter 2).
30. Tuminaro RS, Heroux M, Hutchinson SA, Shadid JN. Official Aztec user's guide, version 2.1. *SAND99-8801J*, Sandia National Laboratories, Albuquerque, New Mexico, 1999; 63.
31. Molenkamp C. Accuracy of finite-difference methods applied to the advection equation. *Journal of Applied Meteorology* 1968; **7**:160–167.
32. Vreugdenhil CB, Koren B (eds). *Numerical Methods for Advection Diffusion Problems*. Notes on Numerical Fluid Mechanics, vol. 45. Vieweg (Bartelsmann Publishing Group International): Braunschweig, Wiesbaden, 1993, ISBN 3-528-07645-3
33. Smith RM, Hutton AG. The numerical treatment of advection: a performance comparison of current methods. *Numerical Heat Transfer* 1982; **5**:439–461.
34. De Vahl Davis G. Natural convection of air in a square cavity: a bench mark numerical solution. *International Journal for Numerical Methods in Fluids* 1983; **3**:249–264.