A high-resolution Petrov–Galerkin method for the 1D convection–diffusion–reaction problem

Prashanth Nadukandi, Eugenio Oñate and Julio Garcia
International Center for Numerical Methods in Engineering (CIMNE),
Universitat Politecnica de Catalunya (UPC), Campus Nord,
Edifici C1, Gran Capitan s/n, 08034 Barcelona, Spain

Abstract

We present the design of a high-resolution Petrov–Galerkin(HRPG) method using linear finite elements for the problem defined by the residual

$$R(\phi) := \frac{\partial \phi}{\partial t} + u\frac{\partial \phi}{\partial x} - k\frac{\partial^2 \phi}{\partial x^2} + s\phi - f$$

where $k, s \geq 0$. The structure of the method in 1D is identical to the *consistent approximate upwind* Petrov–Galerkin (CAU/PG) method [19] except for the definitions of the stabilization parameters. Such a structure may also be attained via the *Finite Calculus* (FIC) procedure [13, 41] by an appropriate definition of the characteristic length. The prefix 'high-resolution' is used here in the sense popularized by Harten, i.e. second order accuracy for smooth/regular regimes and good shock-capturing in nonregular regimes. The design procedure embarks on the problem of circumventing the Gibbs phenomenon observed in L2-projections. Next we study the conditions on the stabilization parameters to circumvent the global oscillations due to the convective term. A conjuncture of the two results is made to deal with the problem at hand that is usually plagued by *Gibbs*, *global* and *dispersive* oscillations in the numerical solution. It is shown that the method indeed reproduces stabilized high-resolution numerical solutions for a wide range of values of u, k, s and f. Finally, some remarks are made on the extension of the HRPG method to multidimensions.

Keywords: convection–diffusion–reaction; finite element; Petrov–Galerkin; stabilized high-resolution methods;

1 Introduction

A singularly perturbed convection–diffusion–reaction problem is an initial-boundary value problem where the diffusion coefficient may take arbitrarily small values. The solution of this problem may exhibit transient and/or exponential boundary layers. In higher dimensions the solution may also exhibit characteristic boundary and/or interior layers. It is well known that the numerical solution of this problem by the Bubnov–Galerkin finite element method (FEM) is prone to exhibit *global, Gibbs* and *dispersive* oscillations. The solution of the stationary problem by the above method exhibits spurious global oscillations for the convection-dominated cases. The local Gibbs oscillations are exhibited along the characteristic layers for the convection-dominated cases. For the reaction-dominated cases Gibbs oscillations may be found near the Dirichlet boundaries and in the regions where the distributed source term is nonregular. The solution of the transient problem may exhibit dispersive oscillations should the initial solution and/or the distributed source term are nonregular.

In the context of variational formulations and weighted residual methods, control over the global instability has been achieved via the streamline-upwind Petrov–Galerkin (SUPG) [1, 2], Taylor–Galerkin [4], characteristic Galerkin [3, 5], Galerkin least squares (GLS) [6], bubble functions [7–9], variational multiscale (VMS) [10, 11], characteristic-based split (CBS) [12] and finite calculus (FIC) based methods [13]. A thorough comparison of some of these methods can be found in [16]. Close connections between the VMS method and stabilization via bubble functions was pointed out in [15]. It was shown that some of the above stabilized methods can be recovered using the FIC equations via an appropriate definition of the stabilization parameters [13, 14]. Nevertheless nonregular solutions continue to exhibit the Gibbs and dispersive oscillations.

Several shock-capturing nonlinear Petrov–Galerkin methods were proposed to control the Gibbs oscillations observed across characteristic internal/boundary layers for the convection-diffusion problem [17–28]. A thorough review, comparison and state of the art of these and several other shock-capturing methods for the convection-diffusion equations, therein named as *spurious oscillations at layers diminishing* methods, was done in [29]. Reactive terms were not considered in the design of these methods and hence they fail to control the localized oscillations in the presence of these terms. Exceptions to this are the *consistent approximate upwind* (CAU) method [19], the methods presented in [24] and those that take the CAU method as the starting point [21, 25, 26]. Nevertheless the expressions for the stabilization parameters therein were never optimized for reactive instability and often the solutions are over-diffusive in these cases.

In the quest to gain reactive stability several methods were built upon the existing frameworks of methods that control Global oscillations. Following the framework of the SUPG method linear Petrov–Galerkin methods were proposed for the convection–diffusion–reaction problem, viz. the DRD [30] and (SU+C)PG [31] methods. Based on the GLS method linear stabilized methods were proposed, viz. the GGLS method [32] for the diffusion–reaction problem and GLSGLS method [33] for the convection-diffusion-production problem. Within the framework of stabilization via bubbles the USFEM method [34] for the diffusion–reaction problem, the improved USFEM method [35] and the *link cutting bubbles* [36] for the convection–diffusion–reaction problem were proposed. Based on the VMS method linear stabilized methods were proposed for the convection–diffusion–reaction problem, viz. the ASGS method [37], the methods presented in [38, 39] and the SGS-GSGS method [40]. Using the FIC equations a nonlinear method based on a single stabilization parameter was proposed for the convection–diffusion–reaction problem [41, 42]. This nonlinear method, though initially formulated within the Petrov–Galerkin framework, subsequent modeling of a simplified form for the numerical nonlinear diffusion, deviated the method from being residual-based (consistency property is violated). Nodally exact Ritz discretizations of the 1D diffusion-absorption/production equations by variational FIC and modified equation methods using a single stabilization parameter were presented in [43]. Generally the homogeneous steady convection–diffusion–reaction problem in 1D has two fundamental solutions. Likewise, the characteristic equation associated with linear stabilized methods which result in compact stencils are quadratic and hence have two solutions. Thus in principle using two stabilization parameters (independent of the boundary conditions) linear stabilized methods which result in compact stencils can be designed to be nodally exact in 1D. Following this line several 'two-parameter methods' viz. (SU+C)PG, GLSGLS and SGS-GSGS methods were designed to be nodally exact for the stationary problem in 1D.

Control over the dispersive oscillations for the transient convection-diffusion problem via linear Petrov–Galerkin methods were discussed in [44] and using space-time finite elements in [45]. As for the linear methods, optimizing the expressions of the stabilization parameters to attain monotonicity will lead to solutions that are at most first-order accurate.

Out of the context of variational formulations and weighted residual methods, a vast literature exists on the design of high-resolution methods. These methods are characterized as *Algebraic Flux Correction/Limiting* methods and are usually developed with in the framework of finite-difference (FDM) or finite-volume (FVM) models. We refer to the books [46–49] for a review of these methods and to the seminal papers in this field [50–54]. The book [55] describes the state of the art in the development of high-resolution schemes based on the Flux-Corrected Transport (FCT) paradigm for unstructured meshes and their generalization to the FEM. Nevertheless the use of these schemes were reported to be rather uncommon in spite of their enormous potential. We refer to the introduction in [56] that discusses the popularity of methods based on variational formulations and weighted residuals. Thus, as encouraged therein, the quest for the design of high-resolution methods based on variational/weighted-residual formulations is active to date.

In this paper we present the design of a FIC-based nonlinear high-resolution Petrov–Galerkin (HRPG) method for the 1D convection–diffusion–reaction problem. The prefix 'high-resolution' is used here in the sense popularized by Harten, i.e. second order accuracy for smooth/regular regimes and good shock-capturing in nonregular regimes. The goal is to design a numerical method within the context of the FIC variational formulation and weighted residuals which is capable of reproducing high-resolution numerical solutions for both the stationary (efficient control of global and Gibbs oscillations as seen in methods [31, 33, 36, 40–43]) and transient regimes (efficient control of dispersive oscillations as seen in *Algebraic Flux Correction/Limiting* methods). In §2 we present the statement of the problem and the HRPG method in higher-dimensions. The statement in higher-dimensions is made only to distinguish the current method with the existing ones. The structure of the method in 1D is identical to the CAU method except for the definitions of the stabilization parameters. The method can be derived via the FIC approach [13] with an adequate (nonlinear) definition of the characteristic length. Thus the results presented here may be extended to these methods. In §4 we focus on the Gibbs phenomenon that is observed in L_2-projections. The design procedure embarks by defining a model L_2-projection problem and establishing the expression for the stabilization parameter to circumvent the Gibbs phenomenon. The target solution for the model problem is chosen to be the one obtained via the mass-lumping procedure. We remark that this solution is used to evaluate the stabilization terms introduced by the HRPG method and an expression for the stabilization parameter is defined that depends only on the problem data. In §5 we extend the methodology to the transient convection–diffusion–reaction problem. We split the design into four model problems and derive the stabilization parameters accordingly. Finally we arrive at an expression for the stabilization parameters depending only on the problem data and representing asymptotically the prior expressions derived for the model problems. We summarize the HRPG design in §5.6. In §5.7 several examples are presented that support the design objectives i.e. stabilization with high-resolution. In §6 some remarks are made on the extension of the HRPG method to multidimensions. Finally we arrive at some conclusions in §7.

2 High-resolution Petrov–Galerkin method

The statement of the multidimensional convection–diffusion–reaction problem is as follows:

$$R(\phi) := \frac{\partial \phi}{\partial t} + \mathbf{u} \cdot (\nabla \phi) - \nabla \cdot (k \nabla \phi) + s\phi - f(\mathbf{x}) = 0 \qquad \text{in} \quad \Omega \tag{1a}$$

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

$$\phi = \phi^p \qquad \text{on} \quad \Gamma_D \tag{1c}$$

$$(k \nabla \phi) \cdot \mathbf{n} + g^p = 0 \qquad \text{on} \quad \Gamma_N \tag{1d}$$

where $\mathbf{u}$ is the convection velocity, k,s are the diffusion and reaction coefficient respectively, $f(\mathbf{x})$ is the source, $\phi_0(\mathbf{x})$ is the initial solution, ϕ^p and g^p are the prescribed values of ϕ and the diffusive flux at the Dirichlet and Neumann boundaries respectively and $\mathbf{n}$ is the normal to the boundary.

The variational statement of the problem (1) can be expressed as follows: Find $\phi : [0, T] \mapsto V$ such that $\forall w \in V_0$ we have,

$$\Big(w, R(\phi)\Big)_\Omega + \Big(w, (k\nabla\phi) \cdot \mathbf{n} + g^p\Big)_{\Gamma_N} = 0 \tag{2}$$

where, if H is the associated Hilbert space then $V := \{w : w \in H \text{ and } w = \phi^p \text{ on } \Gamma_D\}$, $V_0 := \{w : w \in H \text{ and } w = 0 \text{ on } \Gamma_D\}$, $(\cdot, \cdot)_\Omega$ and $(\cdot, \cdot)_{\Gamma_N}$ denote the $\mathrm{L}_2(\Omega)$ and $\mathrm{L}_2(\Gamma_N)$ inner products respectively. The problem (1) may also be expressed in the weak form as follows: Find $\phi : [0, T] \mapsto V$ such that $\forall w \in V_0$ we have,

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

$$a(w, \phi) := \Big(w, \frac{\partial \phi}{\partial t} + u\nabla(\phi) + s\phi\Big)_\Omega + \Big(\nabla(w), k\nabla(\phi)\Big)_\Omega \tag{3b}$$

$$l(w) := \Big(w, f(x)\Big)_\Omega - \Big(w, g^p\Big)_{\Gamma_N} \tag{3c}$$

The statement of the Galerkin method applied to the weak form of the problem (3) is: Find $\phi_h : [0, T] \mapsto V^h$ such that $\forall w_h \in V_0^h$ we have,

$$a(w_h, \phi_h) = l(w_h) \tag{4}$$

We follow [1] to describe a certain class of Petrov–Galerkin methods which account for weights that are discontinuous across element boundaries. The perturbed weighting function is written as $\tilde{w}_h = w_h + p_h$, where p_h is the perturbation that account for the discontinuities. The statement of these class of Petrov–Galerkin methods is as follows: Find $\phi_h : [0, T] \mapsto V^h$ such that $\forall w_h \in V_0^h$ we have,

$$a(w_h, \phi_h) + \sum_e \Big(p_h, R(\phi_h)\Big)_{\Omega_h^e} = l(w_h) \tag{5}$$

The HRPG method whose design in 1D is presented in the subsequent sections, may be defined as Eq.(5) along with the following definitions:

$$p_h := [\mathbf{h} + \mathbf{H} \cdot \hat{\mathbf{u}}^r] \cdot \nabla w_h \tag{6a}$$

$$\mathbf{u}^r := \frac{R(\phi_h)}{|\nabla \phi_h|^2} \nabla \phi_h; \quad \Rightarrow \hat{\mathbf{u}}^r := \frac{\mathbf{u}^r}{|\mathbf{u}^r|} = \frac{\mathrm{sgn}[R(\phi_h)]}{|\nabla \phi_h|} \nabla \phi_h \tag{6b}$$

Method	Perturbation(p_h)	Remarks
SUPG[1]	$\tau\mathbf{u}\cdot\nabla w_h$	
MH[17]	C_i^e	$C_i^e \in \{-\frac{1}{3},\frac{2}{3}\}, i=1,2,3$ $\sum C_i^e = 0$
DC[18]	$\tau_1\mathbf{u}\cdot\nabla w_h + \tau_2\mathbf{u}^{\parallel}\cdot\nabla w_h$	$\mathbf{u}^{\parallel} := \frac{\mathbf{u}\cdot\nabla\phi_h}{\lvert\nabla\phi_h\rvert^2}\nabla\phi_h$
CAU[19], CCAU[21]	$\tau_1\mathbf{u}\cdot\nabla w_h + \tau_2\mathbf{u}^{r}\cdot\nabla w_h$	$\mathbf{u}^{r} := \frac{R(\phi_h)}{\lvert\nabla\phi_h\rvert^2}\nabla\phi_h$
CD[22]	$\tau_1\mathbf{u}\cdot\nabla w_h + \alpha_2\ell\nabla w_h\cdot[\mathbf{I}-\hat{\mathbf{u}}\otimes\hat{\mathbf{u}}]\cdot\hat{\mathbf{u}}^r$	$\hat{\mathbf{u}} := \frac{\mathbf{u}}{\lvert\mathbf{u}\rvert}$ $\hat{\mathbf{u}}^r := \frac{\mathbf{u}^r}{\lvert\mathbf{u}^r\rvert} = \frac{\mathrm{sgn}[R(\phi_h)]}{\lvert\nabla\phi_h\rvert}\nabla\phi_h$
SAUPG[25], Mod.CAU[22]	$\tau[\lambda\mathbf{u}+(1-\lambda)\mathbf{u}^r]\cdot\nabla w_h$	λ is a smoothness measure.
FIC[13]	$\mathbf{h}^{fic}\cdot\nabla w_h$	here $\mathbf{h}^{fic}$ is a characteristic length vector which may be defined in a linear or nonlinear fashion.
HRPG	$[\mathbf{h}+\mathbf{H}\cdot\hat{\mathbf{u}}^r]\cdot\nabla w_h$	$\mathbf{h}$,$\mathbf{H}$ are frame-independent linear characteristic length tensors based on the element geometry (see §6).

Table 1: Perturbations associated with Petrov–Galerkin methods

where, $\mathbf{h}$ and $\mathbf{H}$ are frame-independent 'linear' *characteristic length* tensors that are defined based on the element geometry (see §6). We refer to the Table(1) for a comparison of the HRPG method with the SUPG, FIC and some of the existing shock-capturing methods. From Eqs.(6),(7) and Table(1) the HRPG method could be understood as the combination of upwinding plus a nonlinear discontinuity-capturing operator. The distinction is that in general the upwinding provided by $\mathbf{h}$ is not streamline and the discontinuity-capturing provided by $\mathbf{H}\cdot\hat{\mathbf{u}}^r$ is neither isotropic nor purely crosswind. Of course defining $\mathbf{h} := \tau\mathbf{u}$ and $\mathbf{H} := (\beta\ell)\mathbf{I}$ or $\mathbf{H} := (\beta\ell)[\mathbf{I}-\hat{\mathbf{u}}\otimes\hat{\mathbf{u}}]$ one would recover (except for the definitions of the stabilization parameters) the CAU and the CD methods respectively. We remark that one may arrive at the HRPG method via the finite-calculus(FIC) equations wherein the characteristic length is defined as $\mathbf{h}^{fic} := \mathbf{h}+\mathbf{H}\cdot\hat{\mathbf{u}}^r$. From this point of view the HRPG method can be presented as 'FIC-based'. More details are given in the next section.

Note that in 1D $\mathbf{u}^{\parallel} = \mathbf{u}$ and hence the performance of the DC method [18] is similar to that of the SUPG method. Also note that as the notion of crosswind directions does not exist in 1D, the CD method [22] is identical to the SUPG method. On the other hand the nonlinear shock-capturing terms introduced by the CAU method still exists in 1D and thus in principle are able to control the Gibbs and dispersive oscillations. This feature does carry over to all the methods that have the shock-capturing term similar to that in the CAU method viz. the methods presented in [21],[25],[26],[24]. Unfortunately as pointed out in [29] and in §5.7.1 of this paper, these methods are often over diffusive. The structure of the HRPG method in

1D is identical to the CAU method except for the definitions of the stabilization parameters. In the subsequent sections we design the stabilization parameters of the HRPG method to overcome the shortcomings of the earlier methods.

From Eqs.(5) and (6) the statement of the HRPG method in 1D can be expressed as follows: Find $\phi_h : [0,T] \mapsto V^h$ such that $\forall w_h \in V_0^h$ we have,

$$a(w_h,\phi_h) + \sum_e \left[\Big(\frac{\alpha\ell}{2}\frac{\mathrm{d}w_h}{\mathrm{d}x}, R(\phi_h) \Big)_{\Omega_h^e} + \Big(\frac{\beta\ell}{2}\frac{|R(\phi_h)|}{|\nabla\phi_h|}\frac{\mathrm{d}w_h}{\mathrm{d}x}, \frac{\mathrm{d}\phi_h}{\mathrm{d}x} \Big)_{\Omega_h^e} \right] = l(w_h) \tag{7}$$

where α , β are stabilization parameters to be defined later.

3 Derivation of the HRPG expression via the FIC procedure

The governing equations in the finite calculus (FIC) approach are derived by expressing the balance equations in a domain of finite size and retaining higher order terms. For the 1D convection–diffusion–reaction problem the FIC governing equations are written as [13]

$$R(\phi) - \frac{h}{2}\frac{\partial R(\phi)}{\partial x} = 0 \qquad \text{in} \quad \Omega \tag{8a}$$

$$\phi(x,t=0) = \phi_0(x) \qquad \text{in} \quad \Omega \tag{8b}$$

$$\phi - \phi^p = 0 \qquad \text{on} \quad \Gamma_D \tag{8c}$$

$$k\frac{\partial\phi}{\partial x} + g^p + \frac{h}{2}R(\phi) = 0 \qquad \text{on} \quad \Gamma_N \tag{8d}$$

where the characteristic length h is the dimension of the domain where balance of fluxes is enforced and $R(\phi)$ is defined in Eq.(1a).

The variational statement of Eqs.(8) can be written as: Find $\phi : [0,T] \mapsto V$ such that $\forall w \in V_0$ we have,

$$\Big(w, R(\phi) - \frac{h}{2}\frac{\partial R(\phi)}{\partial x} \Big)_\Omega + \Big(w, k\frac{\partial\phi}{\partial x} + g^p + \frac{h}{2}R(\phi) \Big)_{\Gamma_N} = 0 \tag{9}$$

The corresponding weak form is: Find $\phi : [0,T] \mapsto V$ such that $\forall w \in V_0$ we have,

$$\Big(w + \frac{h}{2}\frac{\mathrm{d}w}{\mathrm{d}x}, R(\phi) \Big)_\Omega + \Big(w, k\frac{\partial\phi}{\partial x} + g^p \Big)_{\Gamma_N} = 0 \tag{10}$$

In the derivation of Eq.(10) we have neglected the change of h within the elements. Clearly Eq.(10) can be seen as a Petrov–Galerkin form with the weighting function defined as $\tilde{w} := w + \frac{h}{2}\frac{\mathrm{d}w}{\mathrm{d}x}$. The term depending on h in Eq.(10) is usually computed in the element interiors only to avoid the discontinuities of the second derivatives terms in $R(\phi)$ along the element boundaries. The discretized form of Eq.(10) is therefore written as: Find $\phi_h : [0,T] \mapsto V^h$ such that $\forall w_h \in V_0^h$ we have,

$$a(w_h,\phi_h) + \sum_e \Big(\frac{h}{2}\frac{\mathrm{d}w_h}{\mathrm{d}x}, R(\phi_h) \Big)_{\Omega_h^e} = l(w_h) \tag{11}$$

The characteristic length can be defined in a number of ways so as to provide an 'optimal'(stabilized) solution. In this work the following nonlinear expression is chosen for h:

$$h := \alpha\ell + \beta\ell\frac{\mathrm{sgn}[R(\phi_h)]}{|\nabla\phi_h|}\nabla\phi_h \tag{12}$$

where ℓ is the element length and α, β are stabilization parameters. A discussion of the alternatives for the definition of h in the FIC context can be found in [13, 28, 41, 42]. Substituting the expression of h into Eq.(11) gives the HRPG form of Eq.(7).

4 Gibbs phenomenon in L_2-Projections

4.1 Introduction

Gibbs phenomenon is a spurious oscillation that occurs when using a truncated Fourier series or other eigen function series at a simple discontinuity. It is characterized by an initial overshoot and then a pattern of undershoot-overshoot oscillations that decrease in amplitude further from the discontinuity. In fact for any given function f and using the metric as the standard L_2-norm, the partial sum of order N of the Fourier series of f denoted as $S_N f$ is the *best approximation* of f in a subspace spanned by trigonometric polynomials of order N. Thus $S_n f$ is the L_2-Projection of f in the considered subspace. This phenomenon is manifested due to the lack of completeness of the approximation space. Similar oscillations appear in the problem of finding the best approximation of a given discontinuous function in any subspace using the L_2-norm as the metric. On every discrete grid/mesh the maximum wavenumber that can be represented is limited by the Nyquist limit. The Nyquist frequency on a uniform grid with grid spacing ℓ is given by π/ℓ. Thus the span of the finite element basis functions associated with this mesh might be viewed as a truncated function series which might be expanded by refining the mesh. Hence the projection of a discontinuous function onto this finite element space exhibits the Gibbs phenomenon. As the amplitude spectrum of a discontinuous function decays only as fast as the *harmonic series*, which is not absolutely convergent, it is impossible to circumvent these oscillations by mere mesh refinement.

The variational statement of the L_2-projection problem is as follows: Find $\phi_h \in V^h$ such that $\forall w_h \in V_0^h$ we have,

$$\Big(w_h, \phi_h - f\Big)_{\Omega_h} = 0 \tag{13}$$

Where f is the given function which might admit discontinuities. If we denote the solution of Eq.(13) as $\phi_h = \mathrm{P}_h(f)$, we have,

$$\| \mathrm{P}_h(f) - f \|_{\mathrm{L}_2(\Omega_h)} \leq \| w_h - f \|_{\mathrm{L}_2(\Omega_h)} \tag{14}$$

In the following sections we consider the scaled L_2-projection problem defined by the residual $R(\phi) = s\phi - f$. In the context of the convection–diffusion–reaction problem, the problem data s physically represents the reaction coefficient. The variational statement of the scaled L_2-projection problem is as follows: Find $\phi_h \in V^h$ such that $\forall w_h \in V_0^h$ we have,

$$\Big(w_h, s\phi_h - f\Big)_{\Omega_h} = 0 \tag{15}$$

Taking $s = 1$ we recover the L_2-projection problem given by Eq.(13).

4.2 Galerkin Method

4.2.1 FE discretization

Discretization of the space by linear finite elements will lead to the approximation $\phi_h = N^a \Phi^a$ and the Eq.(15) reduces into the following system of equations.

$$s\mathbf{M} \cdot \mathbf{\Phi} \;=\; \mathbf{f} \tag{16}$$

$$\mathbf{M}^{ab} = \Big(N^a, N^b\Big)_{\Omega_h} \quad ; \quad \mathbf{f}^a = \Big(N^a, f\Big)_{\Omega_h} \tag{17}$$

It is well known that the Gibbs oscillations can be circumvented in the numerical solution if the standard row-lumping technique is performed on the mass matrix $\mathbf{M}$. Unfortunately, this operation though effective for this specific problem, it cannot be extended to other problems in general. The answer for not advocating this technique can be found in the Godunov's theorem: 'All linear monotone schemes are at most first order accurate'. The only way to circumvent this problem is to design a nonlinear method that would reproduce the same numerical solution as obtained by mass-lumping.

4.2.2 Model Problem 1

The 1D domain is chosen to be of unit length and discretized by 4N linear elements(N>5). The function whose L_2-projection is sought is defined as follows:

$$f(x) = \begin{cases} 0 & \forall \quad x \in [0, 0.25 + \eta_1 \ell] \cup [0.75 - \eta_2 \ell, 1] \\ q & \text{else} \end{cases} \tag{18}$$

Where $\ell = 1/(4N)$ is the element length and $\eta_1, \eta_2 \in [0, 1]$ are parameters that determine the location of the simple discontinuity in the function f. The solution of Eq.(16) using a lumped mass matrix can be expressed as follows:

$$\mathbf{\Phi} = (\frac{q}{s})\{0, \cdots, 0, \frac{(1-\eta_1)^2}{2}, \frac{(2-\eta_1^2)}{2}, 1, \cdots, 1, \frac{(2-\eta_2^2)}{2}, \frac{(1-\eta_2)^2}{2}, 0, \cdots, 0\} \tag{19}$$

Figure (1a) illustrates the function $f(x)$ using $\eta_1 = 0.5$, $\eta_2 = 0.3$ and $q = s = 1$ alongside the numerical solution of the Eq.(16) using both the consistent and lumped mass matrix. Figure (1b) illustrates the profile characteristics of the monotone solution obtained via mass-lumping (Eq.19) with respect to the location of the discontinuity.

4.2.3 Discrete Upwinding

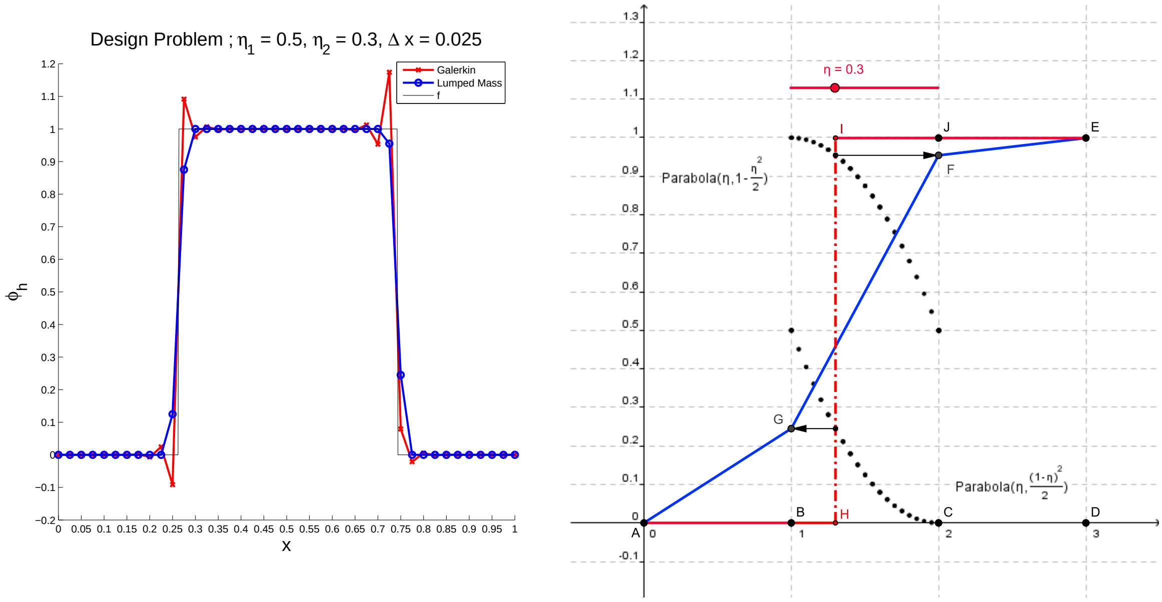
Let the discrete system of equations be represented in the matrix form as follows:

$$\mathbf{A} \cdot \mathbf{x} = \mathbf{b} \tag{20}$$

Discrete upwinding is an algebraic operation to convert the system matrix $\mathbf{A}$ into an M-matrix [55]. Discrete upwinding is the least diffusive linear operation to produce an M-matrix. We denote the discrete upwinding operation on any given matrix $\mathbf{A}$ by $\mathrm{DU}(\mathbf{A})$. The discrete upwinding operation is performed by adding to the matrix $\mathbf{A}$ a discrete diffusion matrix $\tilde{\mathbf{D}}$ as follows:

$$\tilde{d}_{ii} = -\sum_{j \neq i} \tilde{d}_{ij}; \qquad \tilde{d}_{ij} = \tilde{d}_{ji} = -\max\{0, a_{ij}, a_{ji}\} \tag{21}$$

$$\mathrm{DU}(\mathbf{A}) = \;\; \tilde{\mathbf{A}} \;\; = \mathbf{A} + \tilde{\mathbf{D}} \tag{22}$$



(a) (b)

Figure 1: (a) The design problem ; (b) Characteristics of the monotone solution(AGFE). ABHIJE illustrates the discontinuous regime of $f(x)$

It is interesting to note that the discrete upwinding operation on the mass matrix $\mathbf{M}$ will result in the mass-lumping operation.

$$\mathbf{M}^e = \frac{\ell}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \quad ; \quad \tilde{\mathbf{D}}^e = \frac{\ell}{6}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{23}$$

$$\mathrm{DU}(\mathbf{M}^e) = \mathbf{M}^e + \tilde{\mathbf{D}}^e = \frac{\ell}{6}\begin{bmatrix} 3 & 0 \\ 0 & 3 \end{bmatrix} = \mathbf{M}_L^e \tag{24}$$

4.2.4 Total variation

The total variation of a function, say $\phi(x)$, in 1D is given by the following equation:

$$TV(\phi) = \int_x |\nabla\phi| \mathrm{d}x \tag{25}$$

Thus it can be seen that the total variation, as the name suggests, measures the total hike or drop in the function profile as we traverse the 1D domain. It can also be noticed that any spurious oscillation in the numerical approximation of ϕ would cause the total variation to increase. Harten proved that a monotone scheme is total variation non-increasing (TVD) and a TVD scheme is monotonicity preserving [53]. To date various high-resolution schemes have been designed based on the TVD concept often using flux/slope limiters. If linear finite elements are used to approximate the numerical solution ϕ_h the total variation may be calculated as follows:

$$TV(\phi_h) = \sum_i |\Phi_{i+1} - \Phi_i| \tag{26}$$

For the problem under consideration, the sufficient conditions given by Harten in [53] for a numerical scheme to be TVD drops down to the condition for the system matrix to be an M-matrix. Thus in the design of the high-resolution Petrov–Galerkin method we use the total variation of the numerical solution as *a posteriori* verification condition. The total variation of the given discontinuous function $f(x)$ in the design problem is $TV(f) = 2$.

4.3 HRPG design

In this section we design the stabilization parameters of the HRPG method given by Eq.(7) and choosing $\alpha = 0$. For the model problem described in §4.2.2 the statement of the method is as follows: Find $\phi_h \in V^h$ such that $\forall w_h \in V_0^h$ we have,

$$\Big(w_h, R(\phi_h)\Big)_{\Omega_h} + \sum_e \Big(\frac{\beta\ell}{2}\frac{|R(\phi_h)|}{|\nabla\phi_h|}\frac{\mathrm{d}w_h}{\mathrm{d}x}, \frac{\mathrm{d}\phi_h}{\mathrm{d}x}\Big)_{\Omega_h^e} = 0 \tag{27}$$

Where, $R(\phi_h) := s\phi_h - f$ is the residual and β is a stabilization parameter to be defined later. If the domain is discretized by linear finite elements the Eq.(27) can be expressed in the matrix form for each element as follows:

$$\Big[s\mathbf{M}^e + \mathbf{S}^e\Big] \cdot \mathbf{\Phi}^e = \mathbf{f}_g^e \tag{28}$$

where the corresponding matrices are defined as,

$$\mathbf{M}^e = \Big(N^a, N^b\Big)_{\Omega_h^e} = \frac{s\ell}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \tag{29}$$

$$\mathbf{S}^e = (\frac{\beta\ell}{2})\Big(\frac{|R(\phi_h)|}{|\nabla\phi_h|}\frac{\mathrm{d}N^a}{\mathrm{d}x}, \frac{\mathrm{d}N^b}{\mathrm{d}x}\Big)_{\Omega_h^e} = \frac{k^*(\phi_h)}{\ell}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{30}$$

$$\mathbf{f}_g^e = \Big(N^a, f\Big)_{\Omega_h^e} \tag{31}$$

$$k^*(\phi_h) = \frac{\beta}{2}\Big(\frac{|R(\phi_h)|}{|\nabla\phi_h|}, 1\Big)_{\Omega_h^e} \tag{32}$$

$$\mathbf{f}_g = q\ell\ \{0, \cdots, 0, \frac{(1-\eta_1)^2}{2}, \frac{(2-\eta_1^2)}{2}, 1, \cdots, 1, \frac{(2-\eta_2^2)}{2}, \frac{(1-\eta_2)^2}{2}, 0, \cdots, 0\} \tag{33}$$

In order to design the parameter β we assume that the method converges to the solution given by Eq.(19). This is a fair assumption as it can be seen in Eq.(27) that the nonlinear Petrov–Galerkin term is symmetric subjected to the linearization as shown in Eq.(30) and hence there exists a β such that the effect of this term is equivalent to the discrete diffusion introduced by the discrete upwinding operation. If $\eta \in [0, 1]$ be a generic parameter to define the location of the simple discontinuity within the element, we have then for the element containing the discontinuity,

$$\begin{aligned}\Big(\frac{|R(\phi_h)|}{|\nabla\phi_h|}, 1\Big)_{\Omega_h^e} &= (\frac{s\ell^2}{2})\Big[\frac{1+2\eta-6\eta^2+8\eta^3-4\eta^4}{1+2\eta-2\eta^2}\Big] \\ &= (\frac{s\ell^2}{2})\Big[\frac{1+2\eta(1-\eta)[1-2\eta(1-\eta)]}{[1+2\eta(1-\eta)]}\Big]\end{aligned} \tag{34}$$

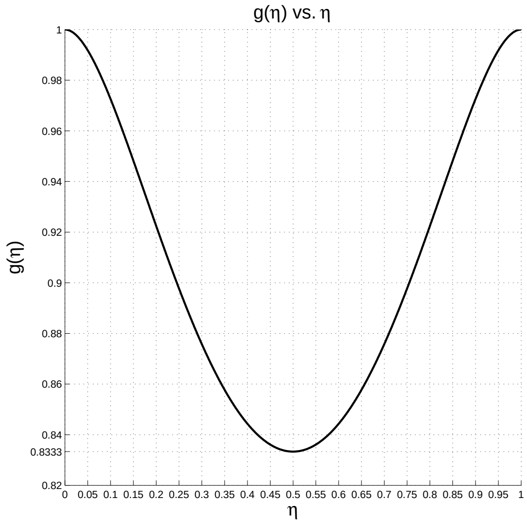


Figure 2: The plot of $g(\eta)$ for $\eta \in [0, 1]$

For the element adjacent to the element containing the discontinuity we have,

$$\Big(\frac{|R(\phi_h)|}{|\nabla\phi_h|}, 1\Big)_{\Omega_h^e} = (\frac{s\ell^2}{2}) \tag{35}$$

Thus, the nonlinear term in Eq.(27) and for the converged solution given by Eq.(19) can be expressed as follows:

$$\Big(\frac{|R(\phi_h)|}{|\nabla\phi_h|}, 1\Big)_{\Omega_h^e} = \begin{cases} (s\ell^2/2)\ g(\eta)\ , & \text{for elements with shock} \\ (s\ell^2/2)\ , & \text{for elements adjacent to the shock} \\ 0\ , & \text{else} \end{cases} \tag{36}$$

where, the function $g(\eta)$ is defined as,

$$g(\eta) \quad := \quad \Big[\frac{1 + 2\eta(1-\eta)[1 - 2\eta(1-\eta)]}{[1 + 2\eta(1-\eta)]}\Big] \tag{37}$$

$$\forall\ \eta \in [0, 1] \quad , \quad g(\eta) \in [(5/6), 1] \tag{38}$$

Figure (2) illustrates the plot of the function $g(\eta)$ vs η. To define the parameter β we require that for the elements in the vicinity of the discontinuity the nonlinear Petrov–Galerkin method reproduces the effect of discrete upwinding. The system matrix for the element containing the discontinuity is as follows,

$$\Big[s\mathbf{M}^e + \mathbf{S}^e\Big] = (\frac{s\ell}{6}) \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} + (\frac{\beta s\ell g(\eta)}{4}) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{39}$$

To reproduce the effect of discrete upwinding the following relation should hold,

$$s\mathbf{M}^e + \mathbf{S}^e \quad = \quad \mathrm{DU}(s\mathbf{M}^e) = s\mathbf{M}^e + \tilde{\mathbf{D}}^e \tag{40}$$

$$\Rightarrow (\frac{\beta s\ell g(\eta)}{4}) \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \quad = \quad \frac{s\ell}{6} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{41}$$

The expression for the parameter β may be expressed as follows:

$$\boxed{\beta g(\eta) \geq \frac{2}{3}} \tag{42}$$

Remarks:

- The definition of β satisfying the equation $\beta g(\eta) = \frac{2}{3}$ would exactly reproduce the solution of Eq.(16) using the lumped mass matrix.
- From the design point-of-view the definition of β involving the function $g(\eta)$ would imply *a priori* knowledge of the solution. Hence we define β using the extremum values of the function $g(\eta)$.

Thus from Eq.(38) and Eq.(42) we have,

$$\text{Type-I}: \quad \min\{g(\eta)\} = \frac{5}{6} \quad \Rightarrow \quad \boxed{\beta \geq \frac{4}{5}} \tag{43}$$

$$\text{Type-II}: \quad \max\{g(\eta)\} = 1 \quad \Rightarrow \quad \boxed{\beta \geq \frac{2}{3}} \tag{44}$$

With these definitions the design of the high-resolution Petrov–Galerkin (HRPG) method for the scaled L_2-Projection problem is complete.

4.4 Examples

4.4.1 Example 1

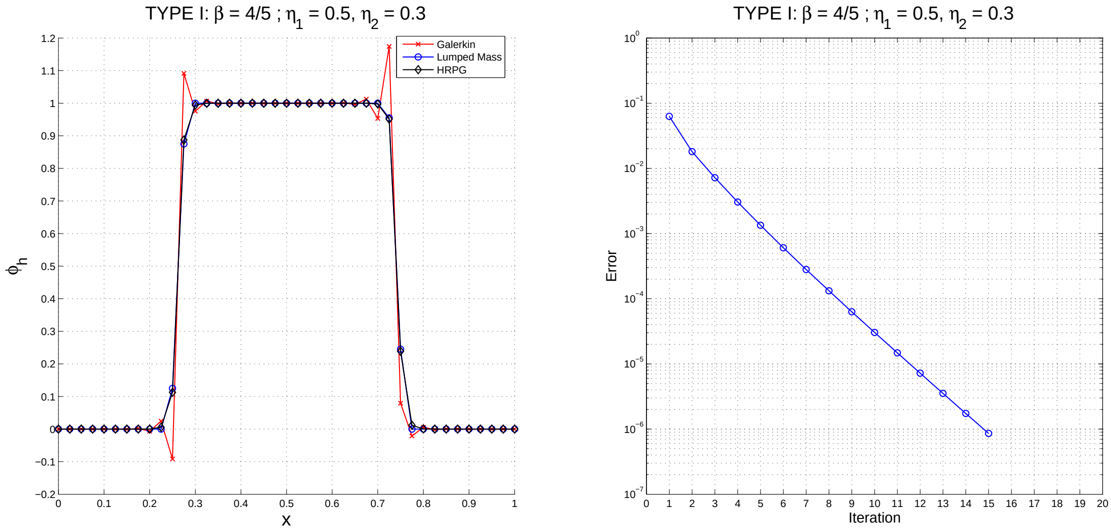
We solve the problem described in §4.2.2 and using $q = s = 1$. The 1D domain is discretized into 40 linear elements. The numerical results of the HRPG method (Type I and II) are compared with the solutions obtained by the Galerkin method using both consistent and lumped mass matrix. Figs.(3,4) illustrate the results of HRPG Type I and Type II respectively for $\eta_1 = 0.5, \eta_2 = 0.3$. Figs.(5,6) illustrate the same for $\eta_1 = 1, \eta_2 = 0$. Both Type I and Type II effectively circumvent the Gibbs phenomenon. HRPG Type I method is clearly more diffusive, nevertheless monotonicity is guaranteed. HRPG Type II is monotone *to-the-eye*. A quantitative analysis based on the measured total variation is studied in §4.4.3.

4.4.2 Example 2

The analysis domain is the same as the problem described in §4.2.2 and using $s = 1$. The 1D domain is discretized into 100 linear elements. The function whose L_2-projection is sought is now defined as follows:

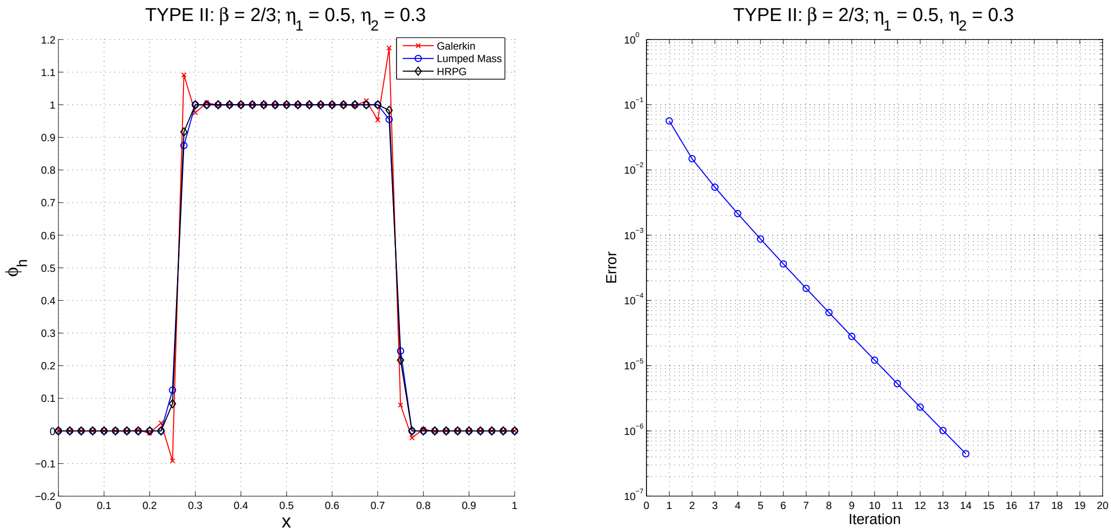
$$f(x) = \begin{cases} \cos(4\pi x - 2\pi) & \forall \quad x \in [0.25, 0.75] \\ 0 & \text{else} \end{cases} \tag{45}$$

The above function has both smooth and shock regimes. Simple discontinuities are present at $x = 0.25$ and $x = 0.75$ and in the rest of the domain the function is smooth. This example studies the efficiency of the HRPG method for mixed regimes. Figures (7,8) illustrate that the accuracy of the solution in the smooth regime is not compromised while effectively circumventing the Gibbs phenomenon around the shocks.



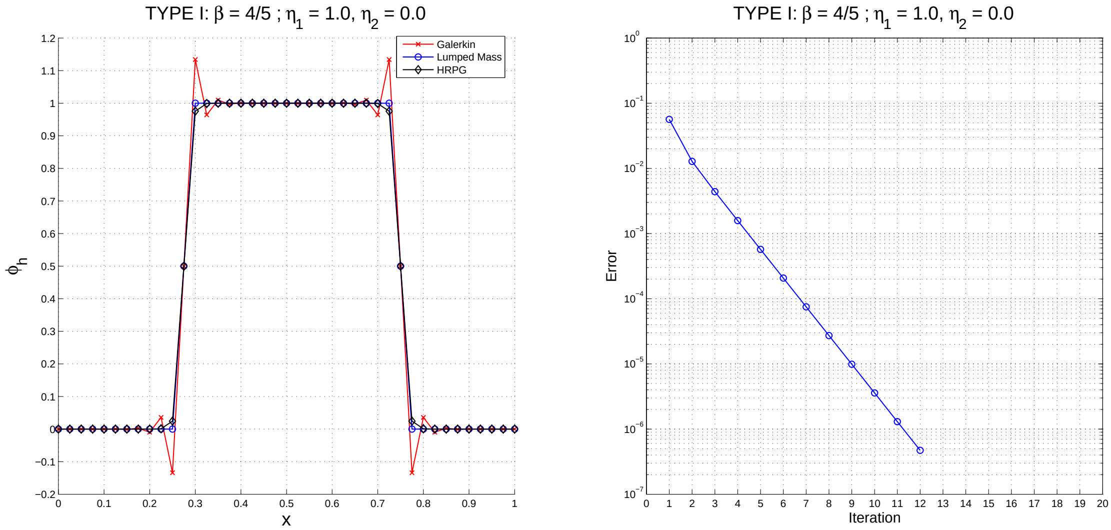
(a) (b)

Figure 3: Example 1: HRPG Type I, (a) numerical solution for $\eta_1 = 0.5, \eta_2 = 0.3$; (b) corresponding nonlinear convergence plot



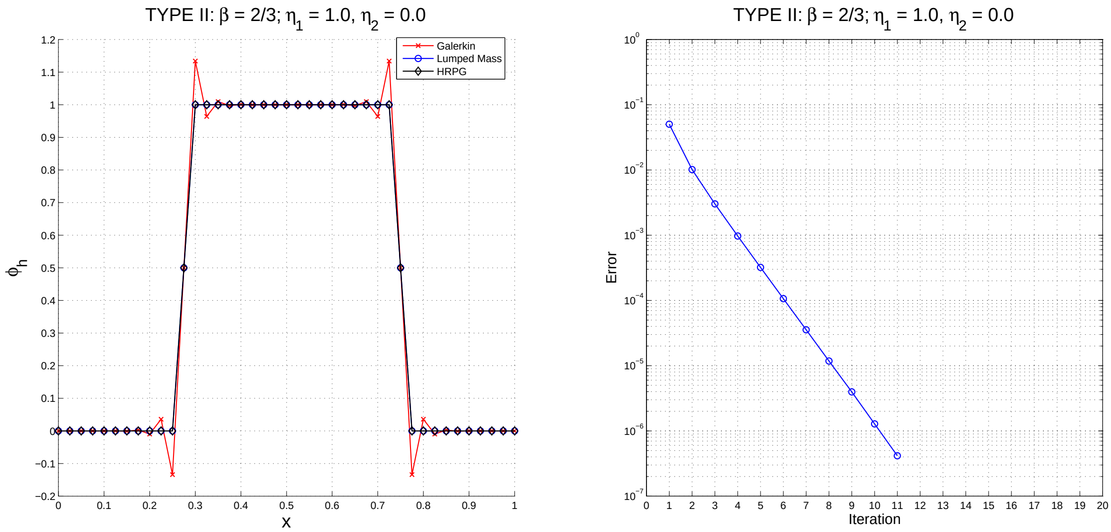
(a) (b)

Figure 4: Example 1: HRPG Type II, (a) numerical solution for $\eta_1 = 0.5, \eta_2 = 0.3$; (b) corresponding nonlinear convergence plot



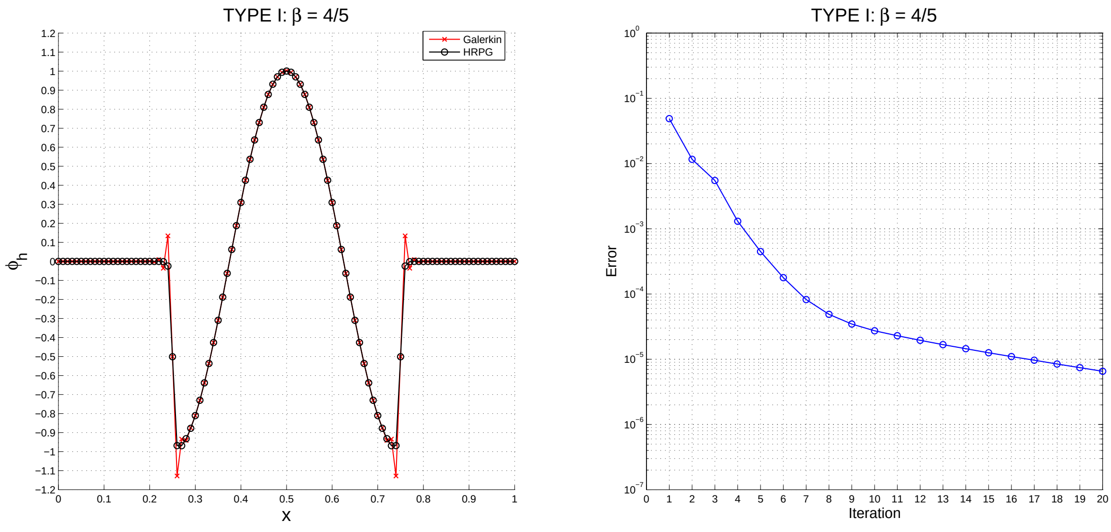
(a) (b)

Figure 5: Example 1: HRPG Type I, (a) numerical solution for $\eta_1 = 1.0, \eta_2 = 0.0$; (b) corresponding nonlinear convergence plot



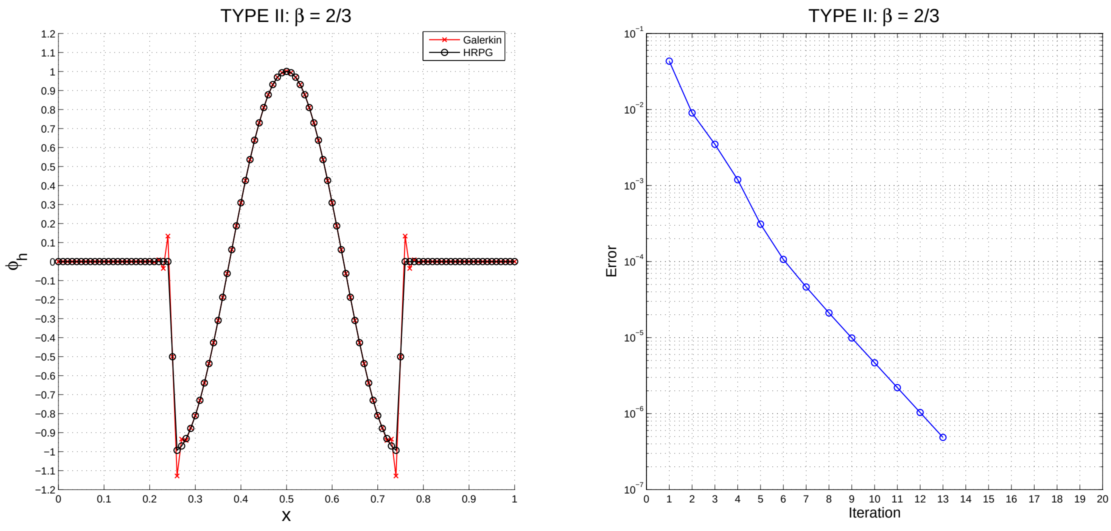
(a) (b)

Figure 6: Example 1: HRPG Type II, (a) numerical solution for $\eta_1 = 1.0, \eta_2 = 0.0$; (b) corresponding nonlinear convergence plot



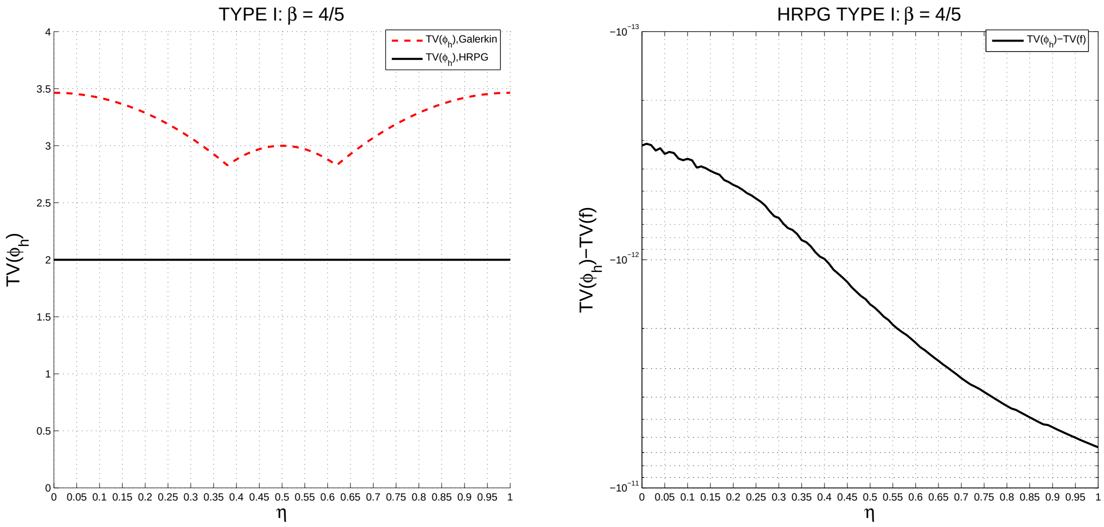
(a) (b)

Figure 7: Example 2: HRPG Type I, (a) numerical solution with both smooth and shock regimes ; (b) corresponding nonlinear convergence plot



(a) (b)

Figure 8: Example 2: HRPG Type II, (a) numerical solution with both smooth and shock regimes ; (b) corresponding nonlinear convergence plot



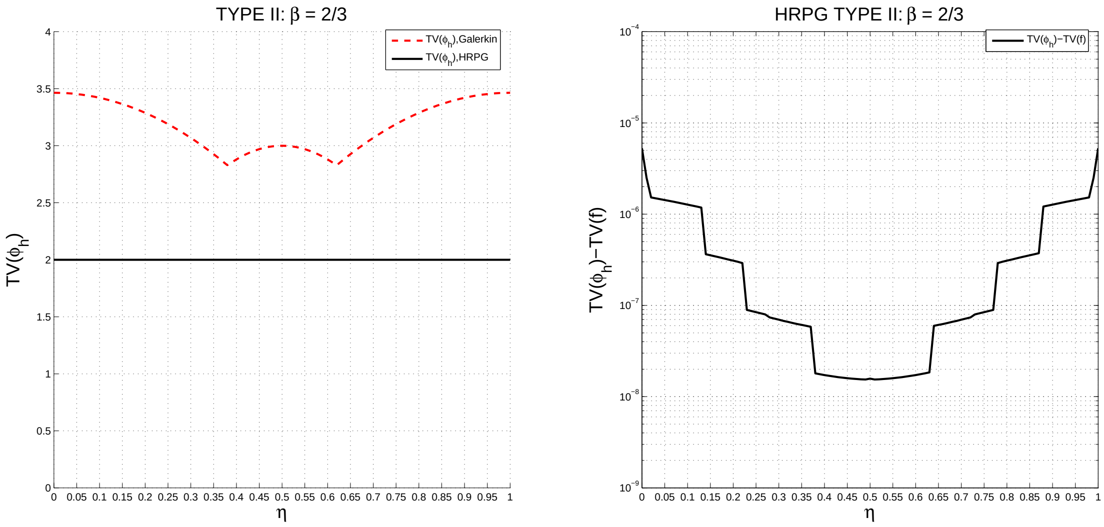
(a) (b)

Figure 9: Example 3: HRPG Type I, (a) TV(ϕ_h) plot ; (b) TV(ϕ_h)-TV(f) plot

4.4.3 Example 3

The problem considered in this example is the same as in §4.4.1. To reduce the variability of the problem data, we have chosen $\eta_1 = \eta_2$. As it is mentioned earlier in §4.2.4, the total variation(TV) of the numerical solution(ϕ_h) is a direct measure (though *a posteriori*) of the presence of spurious oscillations. HRPG Type I method guarantees that the system matrix for the current problem is an M-matrix for all values of η_1 and η_2. This is a sufficient condition to obtain a monotonicity-preserving solution. The system matrix using the HRPG Type II method is an M-matrix only when $\eta_1, \eta_2 = \{0, 1\}$. Thus we study TV($\phi_h$) to have a quantitative measure of performance for the Type I and Type II methods.

Figures (9,10) illustrate with respect to the Galerkin method the $TV(\phi_h)$ vs η plots for the HRPG Type I and Type II methods respectively. It is remarkable that both the methods measure $TV(\phi_h) = 2$ which is the same as $TV(f)$. A study of the error $TV(\phi_h) - TV(f)$ suggests(as expected) that for the Type I method $TV(\phi_h) < TV(f)$ and $TV(\phi_h) - TV(f) =$ O(`1e-11`). For the Type II method, $TV(\phi_h) > TV(f)$ and $TV(\phi_h) - TV(f) =$ O(`1e-5`) which is an acceptable tolerance. In the light of these results the method we currently prefer is HRPG Type II.



(a) (b)

Figure 10: Example 3: HRPG Type II, (a) TV(ϕ_h) plot ; (b) TV(ϕ_h)-TV(f) plot

4.5 Summary

$$
\boxed{
\begin{array}{rcl}
\text{Residual} & : & R(\phi_h) := s\phi_h - f \\
\text{HRPG method} & : & \Big(w_h, R(\phi_h)\Big)_{\Omega_h} + \sum_e \left(\frac{\beta \ell}{2} \frac{|R(\phi_h)|}{|\nabla \phi_h|} \frac{\mathrm{d}w_h}{\mathrm{d}x}, \frac{\mathrm{d}\phi_h}{\mathrm{d}x} \right)_{\Omega_h^e} = 0 \\
\text{Type I} & : & \boxed{\beta \geq \frac{4}{5}} \rightarrow \text{M-matrix guaranteed} \\
\text{Type II} & : & \boxed{\beta \geq \frac{2}{3}} \rightarrow \text{Total variation limit}
\end{array}
}
$$

The Gibbs phenomenon that arises in L_2-projections is studied for the Galerkin method in 1D using linear finite elements. A nonlinear Petrov–Galerkin method (HRPG) is formulated and the stabilization parameter is designed (Type I and Type II) so as to circumvent the Gibbs phenomenon and thus leading to a high-resolution method. The HRPG method is shown to perform well in the presence of both smooth and shock regimes in the solution. HRPG Type II method is shown to be in the total variation limit with acceptable tolerance of O(`1e-5`) and thus is essentially non-oscillatory(monotone *to-the-eye*). Hence it is currently the preferred choice for the extension of the HRPG design to the model problems in the subsequent sections.

5 convection–diffusion–reaction problem

5.1 Galerkin method and discrete upwinding

Consider the convection–diffusion–reaction problem given by Eq.(1) in 1D and subjected only to the Dirichlet boundary conditions. Discretization of the space by linear finite elements will lead to the approximation $\phi_h = N^a \Phi^a$. For the Galerkin method we arrive at the following system of equations.

$$\mathbf{M}\dot{\Phi} + [u\mathbf{C} + k\mathbf{D} + s\mathbf{M}]\Phi = \mathbf{f} \tag{46}$$

where the element contributions to the above matrices and vector are given by,

$$\mathbf{M}^e_{ab} = \Big(N^a, N^b\Big)_{\Omega_h} = \frac{\ell}{6}\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad \mathbf{C}^e_{ab} = \Big(N^a, \nabla(N^b)\Big)_{\Omega_h} = \frac{1}{2}\begin{bmatrix} -1 & 1 \\ -1 & 1 \end{bmatrix} \tag{47}$$

$$\mathbf{f}^e_a = \Big(N^a, f\Big)_{\Omega_h} = \frac{\ell}{2}\begin{Bmatrix} 1 \\ 1 \end{Bmatrix} \qquad \mathbf{D}^e_{ab} = \Big(\nabla(N^a), \nabla(N^b)\Big)_{\Omega_h} = \frac{1}{l}\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \tag{48}$$

The discrete upwinding process applied to the steady state of Eq.(46) will introduce a numerical diffusion k^{du} as follows:

$$\mathrm{DU}(u\mathbf{C} + k\mathbf{D} + s\mathbf{R}) = [u\mathbf{C} + k\mathbf{D} + s\mathbf{R}] + k^{du}\mathbf{D} \tag{49}$$

$$k^{du} = \max\left\{\left[\frac{|u|\ell}{2} + \frac{s\ell^2}{6} - k\right], 0\right\} = k\ \max\left\{\left[|\gamma| + \frac{\omega}{6} - 1\right], 0\right\} \tag{50}$$

The form of this numerical diffusion (Eq.(50)) is identical to that found in [41] using a FIC-based approach. The stabilization method presented in [41] introduces within each element an additional nonlinear diffusion as follows:

$$k^{fic} = k\ \max\left\{\left[\left(\frac{\mathrm{sgn}[\nabla\phi_h]}{\mathrm{sgn}[\Delta\phi_h]}\right)\gamma + \left(\frac{\mathrm{sgn}[\phi_h]}{\mathrm{sgn}[\Delta\phi_h]}\right)\frac{\omega}{6} - 1\right], 0\right\} \tag{51}$$

Clearly the form of Eq.(50) is an upper bound of the value of k^{fic} as defined in Eq.(51).

5.2 Model problem 2

Consider the steady diffusion–reaction problem with a distributed source term given by Eq.(18) and homogeneous Dirichlet boundary conditions:

$$R(\phi) := -k\Delta(\phi) + s\phi - f(x) \tag{52}$$

For the current problem we design the HRPG method with $\alpha = 0$. In the limit as $k \to 0$ the problem reduces to the scaled L_2-projection problem considered in §4. The discrete upwinding operation on the Galerkin method, will introduce an artificial diffusion equivalent to the following,

$$k^{du} = \max\left\{\left[\frac{s\ell^2}{6} - k\right], 0\right\} \tag{53}$$

Note that $\forall k \leq (s\ell^2/6)$ the critical non-oscillatory solution obtained via the discrete upwinding process is identical to the solution obtained with $k = 0$. Thus in order to design the parameter β we can use the solution given by Eq.(19) to estimate the amount of nonlinear diffusion that would be introduced by the HRPG method. Thus,

$$k^*(\phi_h) := \frac{\beta}{2}\Big(\frac{|R(\phi_h)|}{|\nabla\phi_h|}, 1\Big)_{\Omega_h^e} = \beta\frac{s\ell^2}{4}g(\eta) \geq \max\left\{\left[\frac{s\ell^2}{6} - k\right], 0\right\} \tag{54}$$

We define a dimensionless element number $\omega := (s\ell^2/k)$ and consider $g(\eta) = 1$ (Type-II). Thus,

$$\beta \geq \max\left\{\frac{2}{3}\left[1 - \frac{6}{\omega}\right], 0\right\} \tag{55}$$

We remark that β depends only on the problem data and for the current model problem the nonlinear (residual-based) diffusion $k^*(\phi_h)$ implemented in the HRPG method is nonzero and equals k^{du} only for the elements in the vicinity of the discontinuity. In this way it differs from the form of Eq.(51) which for the current model problem introduces a nonlinear diffusion k^{fic} (Eq.(51) using $\gamma = 0$) for all elements.

5.3 Model problem 3

Consider the steady convection-diffusion problem,

$$R(\phi) := u\nabla(\phi) - k\Delta(\phi) \tag{56}$$

For the current problem we design the HRPG method with $\alpha = 0$. The discrete upwinding operation on the Galerkin method, will introduce an artificial diffusion equivalent to the following,

$$k^{du} = \max\left\{\left[\frac{|u|\ell}{2} - k\right], 0\right\} \tag{57}$$

The parameter β may be designed as follows:

$$\frac{|R(\phi_h)|}{|\nabla(\phi_h)|} = \frac{|u\nabla(\phi_h)|}{|\nabla(\phi_h)|} = |u| \tag{58}$$

$$\Rightarrow k^*(\phi_h) := \frac{\beta}{2}\Big(\frac{|R(\phi_h)|}{|\nabla(\phi_h)|}, 1\Big)_{\Omega_h^e} = \frac{\beta}{2}|u|\ell \geq \max\left\{\left[\frac{|u|\ell}{2} - k\right], 0\right\} \tag{59}$$

$$\Rightarrow \beta \geq \max\left\{\left[1 - \frac{1}{|\gamma|}\right], 0\right\} \tag{60}$$

Note that in contrast to the problem considered in §5.2 here we do not need the solution to estimate the nonlinear diffusion introduced by the HRPG method. The expression for β (Eq.60) is identical to the standard critical stabilization parameter obtained for upwind techniques (see [57]).

5.4 Model problem 4

Consider the steady convection–diffusion–reaction problem:

$$R(\phi) := u\nabla(\phi) - k\Delta(\phi) + s\phi - f(x) \tag{61}$$

For the current problem we design the HRPG method again with $\alpha = 0$. If linear finite elements were used, the residual obeys the following relation:

$$
\begin{aligned}
|R(\phi_h)| = |u\nabla(\phi_h) + s\phi_h - f| \quad &\leq \quad |u\nabla(\phi_h)| + |s\phi_h - f| && (62)\\
\Rightarrow \frac{|R(\phi_h)|}{|\nabla(\phi_h)|} = \left| u + \frac{s\phi_h - f}{\nabla(\phi_h)} \right| \quad &\leq \quad |u| + \frac{|s\phi_h - f|}{|\nabla(\phi_h)|} && (63)
\end{aligned}
$$

As it can be seen in Eq.(63), to estimate the nonlinear diffusion introduced by the HRPG method we require the solution of the problem a priori. The simplest idea would be to use the nodally exact solution. Unlike the solution used in §5.2, the analytical solution of the current problem has a complex structure [41]. In order to retain simplicity in the design of β we make a conjuncture of the results obtained in §5.2 and §5.3. The conjuncture is made such that the designed expression for β would approach asymptotically the expressions obtained in §5.2 and §5.3 as $u \to 0$ and $s \to 0$ respectively.

Assume that $u \ll s$ and $f(x)$ be defined as in Eq.(18). Thus we may approximate the solution of the current problem to the one considered in §5.2. This assumption allows us to use the solution defined by Eq.(19) to approximately estimate the following expression:

$$
\Big(\frac{|s\phi_h - f|}{|\nabla(\phi_h)|}, 1\Big)_{\Omega_h^e} \approx \frac{s\ell^2}{2} g(\eta) \tag{64}
$$

As $u \ll s$ we make another approximation using Eq.(63) as follows:

$$
\frac{|R(\phi_h)|}{|\nabla(\phi_h)|} \approx |u| + \frac{|s\phi_h - f|}{|\nabla(\phi_h)|} \tag{65}
$$

Using the two approximations (Eq.64, Eq.65) the parameter β may be designed as follows:

$$
k^*(\phi_h) := \frac{\beta}{2}\Big(\frac{|R(\phi_h)|}{|\nabla(\phi_h)|}, 1\Big)_{\Omega_h^e} \approx \frac{\beta}{2}\left[|u|\ell + \frac{s\ell^2}{2} g(\eta)\right] \geq \max\left\{\left[\frac{|u|\ell}{2} + \frac{s\ell^2}{6} - k\right], 0\right\} \tag{66}
$$

$$
\beta := \max\left\{\left[\frac{2}{3}\left(\frac{|\sigma| + 3}{|\sigma| + 2}\right) - \left(\frac{4}{\omega + 4|\gamma|}\right)\right], 0\right\} \Rightarrow \left\{ \begin{array}{l} \lim\limits_{u \to 0} \; \beta = \max\left\{\dfrac{2}{3}\left[1 - \dfrac{6}{\omega}\right], 0\right\} \\ \lim\limits_{s \to 0} \; \beta = \max\left\{\left[1 - \dfrac{1}{|\gamma|}\right], 0\right\} \end{array} \right. \tag{67}
$$

where, $\gamma := (u\ell/2k)$, $\omega := (s\ell^2/k)$ and $\sigma := (\omega/2\gamma) = (s\ell/u)$ are the element Peclect number, a velocity independent dimensionless number and the Damköler number respectively.
Remark : Eq.(66) does not mean that $k^*(\phi_h) = k^{du}$ for all elements and problem data. We remind that under the assumption $u \ll s$ and $f(x)$ defined as in Eq.(18), the solution to the current model problem is approximated as the one given by Eq.(19). This suggests that, similar to the model problems in §4.2.2 and §5.2, the nonlinear (residual-based) diffusion $k^*(\phi_h)$ equals k^{du} only for the elements in the vicinity of the layers. In general, as β is independent of ϕ_h, the only information known a priori is that $k^*(\phi_h)$ is proportional to the residual $R(\phi_h)$. The argument that this expression for β i.e. Eq.(67), would perform well $\forall u, s$ is a *conjecture* based on the fact that we recover asymptotically the expressions for β i.e. Eq.(55) and Eq.(60) as $u \to 0$ and $s \to 0$ respectively. The efficiency of this expression for β is shown via various numerical examples (see §5.7).

5.5 Model problem 5

Consider the transient convection–diffusion–reaction problem:

$$R(\phi) := \dot{\phi} + u\nabla(\phi) - k\Delta(\phi) + s\phi - f(x) \tag{68}$$

We now design the parameter β associated with the nonlinear perturbation term when the linear perturbation terms exists (i.e. $\alpha \neq 0$). The HRPG method after discretization by linear finite elements will lead to the following system of equations.

$$\begin{array}{rcccccc} \text{(Galerkin terms)} \rightarrow & [\mathbf{M}]\dot{\Phi} & + & [u\mathbf{C} + k\mathbf{D} + s\mathbf{M}]\Phi & - & \mathbf{f_g} & \\ \text{(linear PG terms)} \rightarrow & +\dfrac{\alpha\ell}{2}[\mathbf{C^t}]\dot{\Phi} & + & \dfrac{\alpha\ell}{2}[u\mathbf{D} + s\mathbf{C^t}]\Phi & - & \dfrac{\alpha\ell}{2}\mathbf{f_s} & \\ \text{(nonlinear PG terms)} \rightarrow & & + & \dfrac{\beta}{2}\Big(\dfrac{|R(\phi_h)|}{|\nabla(\phi)|}, 1\Big)_{\Omega^e_h}[\mathbf{D}]\Phi & & & = 0 \end{array} \tag{69}$$

The above equation may be rearranged as follows

$$\mathbf{M}\{\dot{\Phi} + s\Phi\} + \mathbf{C}\{u\Phi\} + \mathbf{C^t}\{\frac{\alpha\ell}{2}\dot{\Phi} + \frac{\alpha\ell s}{2}\Phi\} + (k + \frac{\alpha\ell u}{2} + k^*(\phi_h))\mathbf{D}\Phi = \mathbf{f_g} + \frac{\alpha\ell}{2}\mathbf{f_s} \tag{70}$$

where the expression for $k^*(\phi_h)$ is given by Eq.(32). We define for each element a measure δ with the dimensions of the reaction coefficient.

$$\delta := \dot{\Phi} \oslash \Phi \quad \Rightarrow \dot{\Phi} = \delta \odot \Phi \tag{71}$$

where the vector operators $\oslash$ and $\odot$ are understood to operate point-to-point division and multiplication respectively. Thus in the design of the parameter β, we may model δ as a non-linear reactive coefficient. For the fully discrete problem (after time discretization)δ may be approximated to an element-wise positive constant for simplification. This idea had been pointed out earlier in [44]. Also note that the convection matrix $\mathbf{C}$ is skew-symmetric. Hence the transposed matrix $\mathbf{C^t}$ introduces a negative convection effect. Thus for each element we define the effective convection, diffusion and reaction coefficients as follows:

$$\tilde{u} := u - \frac{\alpha\ell s}{2} - \frac{\alpha\ell\delta}{2} \quad ; \quad \tilde{k} := k + \frac{\alpha\ell u}{2} \quad ; \quad \tilde{s} := s + \delta \tag{72}$$

The effective coefficients $\tilde{u}, \tilde{k}$ and $\tilde{s}$ will be used to design the parameter β. The following effective element dimensionless numbers may be defined:

$$\tilde{\gamma} := \frac{\tilde{u}\ell}{2\tilde{k}} \quad ; \quad \tilde{\omega} := \frac{\tilde{s}\ell^2}{\tilde{k}} \quad ; \quad \tilde{\sigma} := \frac{\tilde{s}\ell}{\tilde{u}} = \frac{\tilde{\omega}}{2\tilde{\gamma}} \tag{73}$$

The parameter β is now defined as in Eq.(67) using these effective element numbers as follows:

$$\beta := \max\left\{ \left[\frac{2}{3}\left(\frac{|\tilde{\sigma}| + 3}{|\tilde{\sigma}| + 2}\right) - \left(\frac{4}{\tilde{\omega} + 4|\tilde{\gamma}|}\right)\right], 0\right\} \tag{74}$$

If the discretization in time is done using the implicit trapezoidal rule, we have

$$\dot{\Phi} = \frac{\tilde{\Phi} - \Phi^n}{\theta\Delta t} \quad ; \quad \Phi^{n+1} = \frac{1}{\theta}\tilde{\Phi} + \frac{\theta - 1}{\theta}\Phi^n \tag{75}$$

where, Δt is the time increment, $n, n+1$ denote the previous and current time steps and $\theta \in [0,1]$ is a parameter that defines the scheme. $\theta = \{0, 0.5, 1\}$ define the forward Euler, implicit midpoint and the backward Euler methods. For the fully discrete system and within each element we evaluate the parameter δ and the residual as follows:

$$
\begin{aligned}
\delta &\approx \frac{1}{\theta \Delta t} \frac{\| \tilde{\phi}_h - \phi_h^n \|_\infty^e}{\| \tilde{\phi}_h \|_\infty^e} && (76) \\
R(\tilde{\phi}_h) &\approx \frac{\tilde{\phi}_h - \phi_h^n}{\theta \Delta t} + u \nabla(\tilde{\phi}_h) - k \Delta(\tilde{\phi}_h) + s\tilde{\phi}_h - f && (77)
\end{aligned}
$$

where $\| \cdot \|_\infty^e$ is the L_∞norm. Note that as steady state is reached $\delta \to 0$. Thus for the steady state problem and using $\alpha = 0$ we recover the definition of the parameter β as given by Eq.(67).

It remains to define the parameter α that controls the fraction of linear perturbation term in the HRPG method. For the 1D convection–diffusion–reaction problem, the HRPG method using $\alpha = 0$ does solve a plethora of examples to give high-resolution stabilized results. Nevertheless for the transient problem the presence of the linear perturbation terms improves the convergence of the nonlinear iterations. Numerical experiments suggest $\alpha \in [0, 1/3]$ which means that the approximations/conjecture used in the design strategy does not hold for larger fractions of the linear perturbation term. The following expression for α was used in the examples to come.

$$
\alpha := \lambda \operatorname{sgn}(u) \max \left\{ \left[1 - \frac{1}{|\gamma|} \right], 0 \right\} \quad ; \quad \lambda := \frac{1}{3(1 + \sqrt{|\sigma|})} \tag{78}
$$

5.6 Summary

Residual $R(\phi_h) := \dfrac{\partial \phi_h}{\partial t} + u\nabla(\phi_h) - k\Delta(\phi_h) + s\phi_h - f(x)$

The HRPG method

Find $\phi_h : [0,T] \mapsto V^h$ such that $\forall w_h \in V_0^h$ we have,

$$a(w_h,\phi_h) + \sum_e \Big(\frac{\alpha\ell}{2}\frac{\mathrm{d}w_h}{\mathrm{d}x}, R(\phi_h)\Big)_{\Omega_h^e} + \Big(\frac{\beta\ell}{2}\frac{|R(\phi_h)|}{|\nabla\phi_h|}\frac{\mathrm{d}w_h}{\mathrm{d}x}, \frac{\mathrm{d}\phi_h}{\mathrm{d}x}\Big)_{\Omega_h^e} = l(w_h)$$

Petrov–Galerkin weight $\rightarrow w_h + \left[\dfrac{\alpha\ell}{2} + \dfrac{\beta\ell}{2}\mathrm{sgn}[R(\phi_h)]\mathrm{sgn}[\nabla(\phi_h)]\right]\dfrac{\mathrm{d}w_h}{\mathrm{d}x}$

Definitions

$$\gamma := \frac{u\ell}{2k} \quad ; \quad \omega := \frac{s\ell^2}{k} \quad ; \quad \sigma := \frac{s\ell}{u}$$

$$R(\tilde{\phi}_h) \approx \frac{\tilde{\phi}_h - \phi_h^n}{\theta\Delta t} + u\nabla(\tilde{\phi}_h) - k\Delta(\tilde{\phi}_h) + s\tilde{\phi}_h - f$$

$$\phi_n^{n+1} = \Big(\frac{1}{\theta}\Big)\tilde{\phi}_h + \Big(\frac{\theta-1}{\theta}\Big)\phi_h^n \quad ; \quad \Delta t = t^{n+1} - t^n \quad ; \quad \theta \in (0,1)$$

$$\lambda := \frac{1}{3(1+\sqrt{|\sigma|})} \quad ; \quad \delta \approx \frac{1}{\theta\Delta t}\frac{\|\tilde{\phi}_h - \phi_h^n\|_\infty}{\|\tilde{\phi}_h\|_\infty}$$

$$\boxed{\alpha := \lambda\ \mathrm{sgn}(u)\max\left\{\left[1 - \frac{1}{|\gamma|}\right], 0\right\}}$$

$$\tilde{u} := u - \frac{\alpha\ell s}{2} - \frac{\alpha\ell\delta}{2} \quad ; \quad \tilde{k} := k + \frac{\alpha\ell u}{2} \quad ; \quad \tilde{s} := s + |\delta|$$

$$\tilde{\gamma} := \frac{\tilde{u}\ell}{2\tilde{k}} \quad ; \quad \tilde{\omega} := \frac{\tilde{s}\ell^2}{\tilde{k}} \quad ; \quad \tilde{\sigma} := \frac{\tilde{s}\ell}{\tilde{u}}$$

$$\boxed{\beta := \max\left\{\left[\frac{2}{3}\left(\frac{|\tilde{\sigma}|+3}{|\tilde{\sigma}|+2}\right) - \left(\frac{4}{\tilde{\omega}+4|\tilde{\gamma}|}\right)\right], 0\right\}}$$

5.7 Examples

5.7.1 Example 1

We consider the convection–diffusion–reaction problem given by Eq.(1) in 1D. We study the steady-state case with the following data : $k = 1$ and $u, s \neq 0$. The 1D domain is taken as $x \in [0,1]$ and it is discretized with eight two-node linear elements. The values of u and s are determined appropriately for different values of γ and ω. The results of the HRPG method (using both $\lambda = 0$ and $\lambda \neq 0$) are compared with that of the Galerkin, Galerkin with discrete upwinding (DU), SUPG, CAU, modified CAU and the FIC based stabilization method presented in [41]. The error in the nonlinear iterations was measured by the following norm:

$$\frac{\|\Phi^{i+1} - \Phi^i\|_{\mathrm{e}}}{\|\Phi^{i+1}\|_{\mathrm{e}}} \tag{79}$$

where, $\|\cdot\|_{\mathrm{e}}$ is the standard Eucleadian vector norm. A tolerance of `1e-5` was chosen as the termination criteria. A maximum of 30 iterations were allowed. Note that the number of iterations required by the nonlinear methods for convergence is displayed next to the corresponding legends. The nonlinear iterations were initialized by the solution obtained by the DU method.

Figures 11 - 16 illustrate the solution obtained for the sourceless case ($f = 0$) and for $(\gamma, \omega) = \{(1, 5), (1, 20), (1, 120), (2, 2), (10, 4), (10, 20)\}$ respectively. The Dirichlet boundary conditions $\phi^p_L := \phi(x = 0) = 8$ and $\phi^p_R := \phi(x = 1) = 3$ were employed. The DU method is robust and provides stable solutions. Unfortunately the accuracy achieved is at most first-order and hence the solutions are generally over-diffusive. The FIC method presented in [41] provides more accurate solutions and remarkably the nonlinear iterations converge with just two iterations. Slight node-to-node oscillations around the exact solutions are observed for the case $\gamma = 10$, $\omega = 4$ viz. Figure 15a and is duly discussed in [41]. As expected the SUPG method provides good solutions to all except the reaction-dominated cases viz. Figures 12b,13b. The CAU and modified CAU methods succeed in circumventing the instabilities for the reaction-dominated cases but for these cases provide solutions that are more diffusive than that of the DU method. The HRPG method provides good solutions for all the cases considered. Note that for the reaction-dominated cases the solutions are less diffusive than the DU, CAU and modified CAU methods. Also note that the solutions obtained by taking $\lambda = 0$ is indistinguishable to that obtained by taking $\lambda \neq 0$. Nevertheless the nonlinear iterations converge faster for the latter.

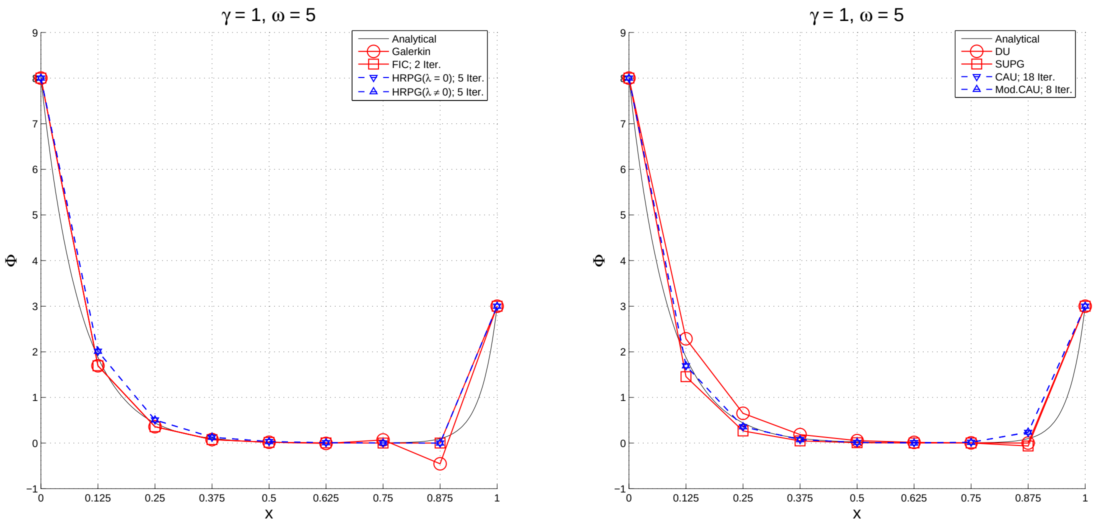
Figures 17,18 illustrate the solution obtained for the sourceless case ($f = 0$) with $(\gamma, \omega) = (10, 200)$ and for Dirichlet boundary conditions $(\phi^p_L, \phi^p_R) = \{(0, 1), (1, 0)\}$ respectively. The FIC method of [41] provides nodally exact *to-the-eye* solutions and the nonlinear iterations converge within 2 iterations. The solutions obtained by the SUPG and CAU methods are indistinguishable and exhibit instabilities for the latter boundary conditions viz. Figure 18b. The modified CAU method circumvents these instabilities but instead provides solutions that are more diffusive than the DU method. The HRPG method (both $\lambda = 0$ and otherwise) succeed to provide stable solutions and are less diffusive than that obtained by the DU method. Figure 17 shows that the HRPG method with $\lambda = 0$ converges in just one iteration while using $\lambda \neq 0$ seven iterations were needed. This is a rare coincidence where the initial solution provided by the DU method and the solution of the HRPG method with $\lambda = 0$ are closer than the specified tolerance.

Figures 19,20 illustrate the solution obtained with $(\gamma, \omega) = (2, 0)$ and $(f, \phi^p_L, \phi^p_R) = \{(0, 0, 1), (u, 0, 0)\}$ respectively. As expected the SUPG method provides nodally exact solutions for these cases. The DU, CAU and HRPG methods provide stable solutions and are indistinguishable from each other. The modified CAU solution is very similar to the solutions of the former methods. For the sourceless case ($f = 0$) the solutions of the DU and FIC methods are indistinguishable. Unfortunately when $f \neq 0$ the nonlinear iterations associated with the latter fail to converge. We believe that this behavior is due to the increased nonlinearity associated with the definition of the stabilization parameters (See [41]).

5.7.2 Example 2

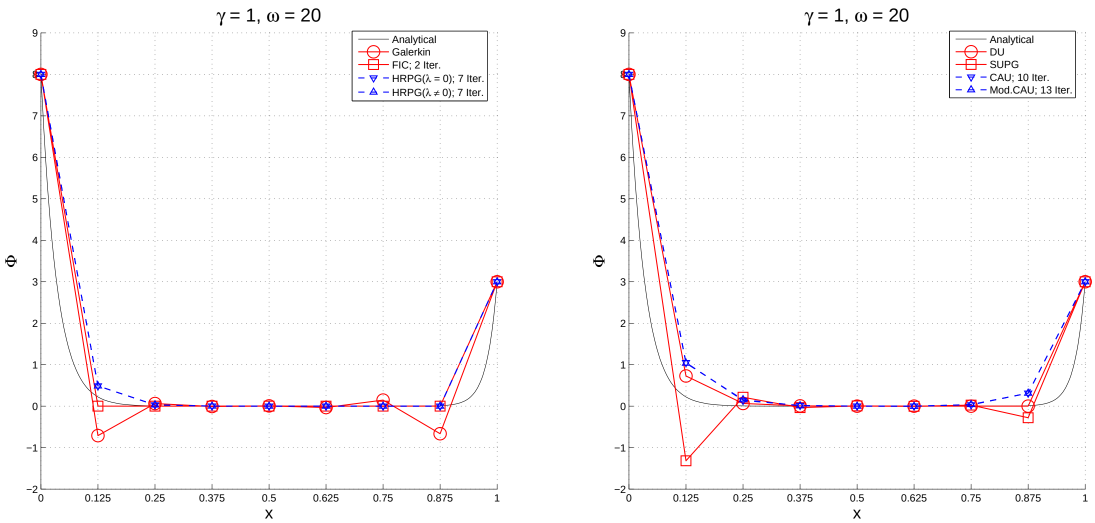
We consider again the convection–diffusion–reaction problem given by Eq.(1) in 1D. Now we study the transient pure-convection problem, i.e. $k, s, f = 0$. The Dirichlet boundary condition $\phi(x = 0) = 0$ is employed. The following initial condition was used:

$$\phi(x, t = 0) = \begin{cases} 1 & \forall \quad x \in [0.1, 0.2] \cup [0.3, 0.4] \\ 0 & \text{else} \end{cases} \tag{80}$$



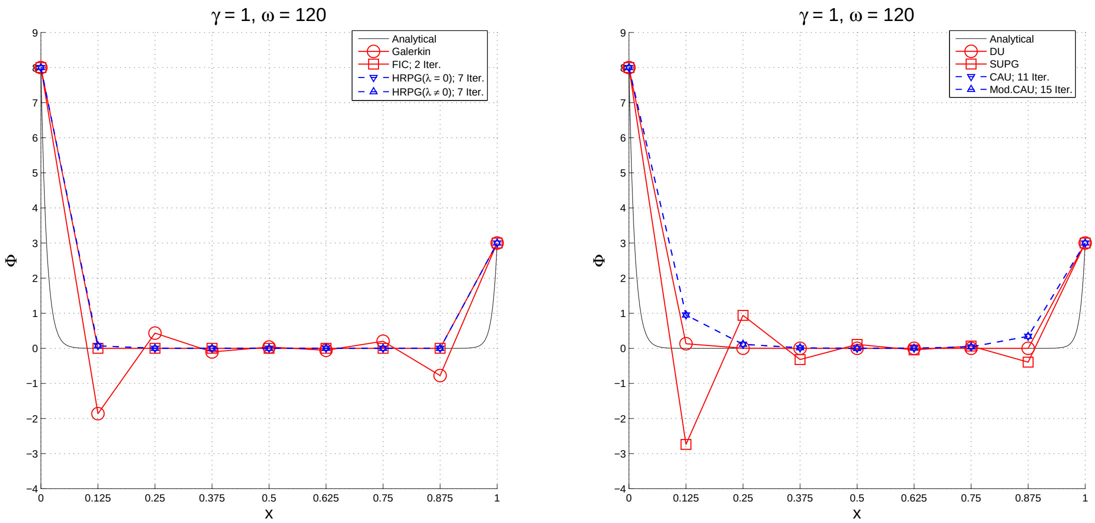
(a) (b)

Figure 11: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (1, 5, 0, 8, 3)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



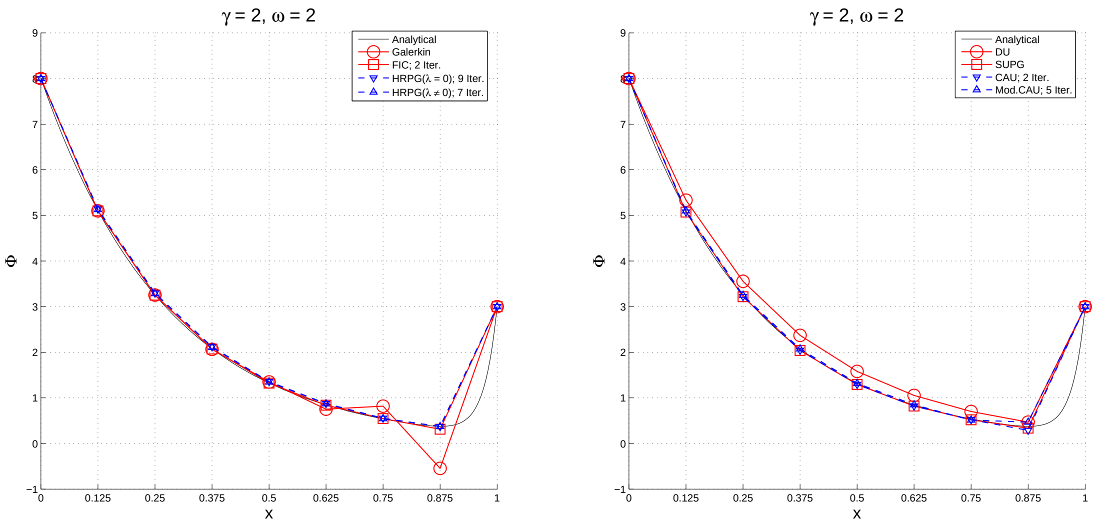
(a) (b)

Figure 12: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (1, 20, 0, 8, 3)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



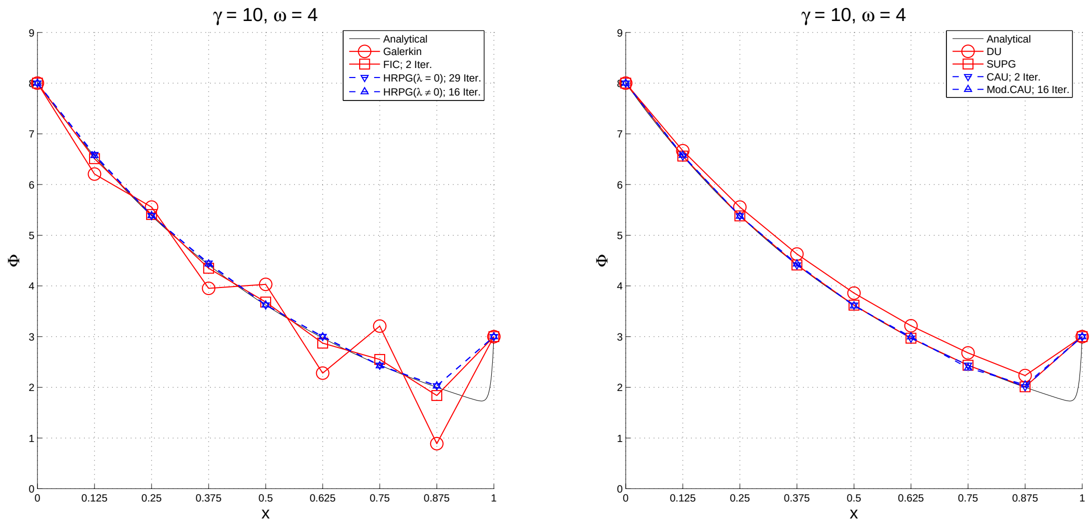
(a) (b)

Figure 13: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (1, 120, 0, 8, 3)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



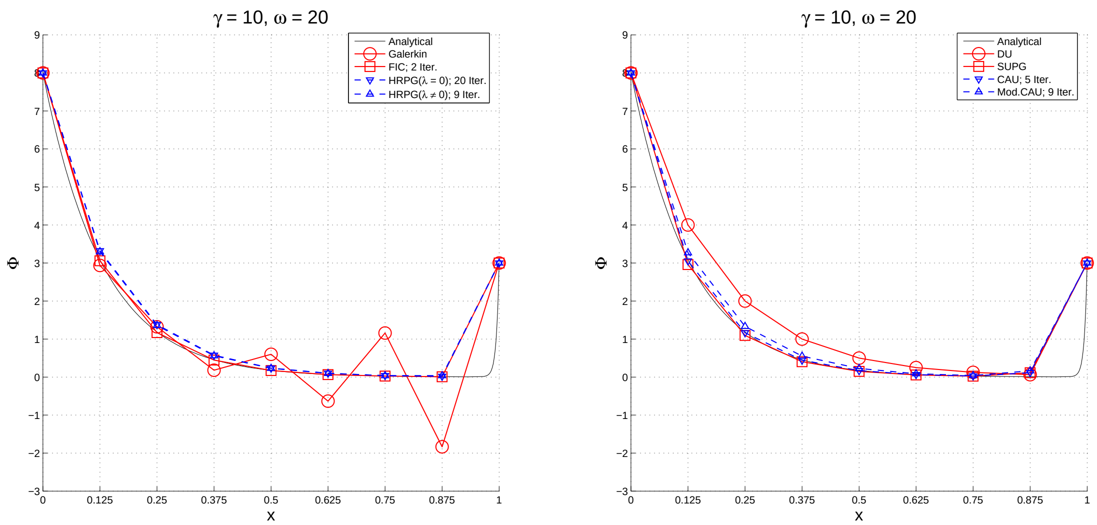
(a) (b)

Figure 14: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (2, 2, 0, 8, 3)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



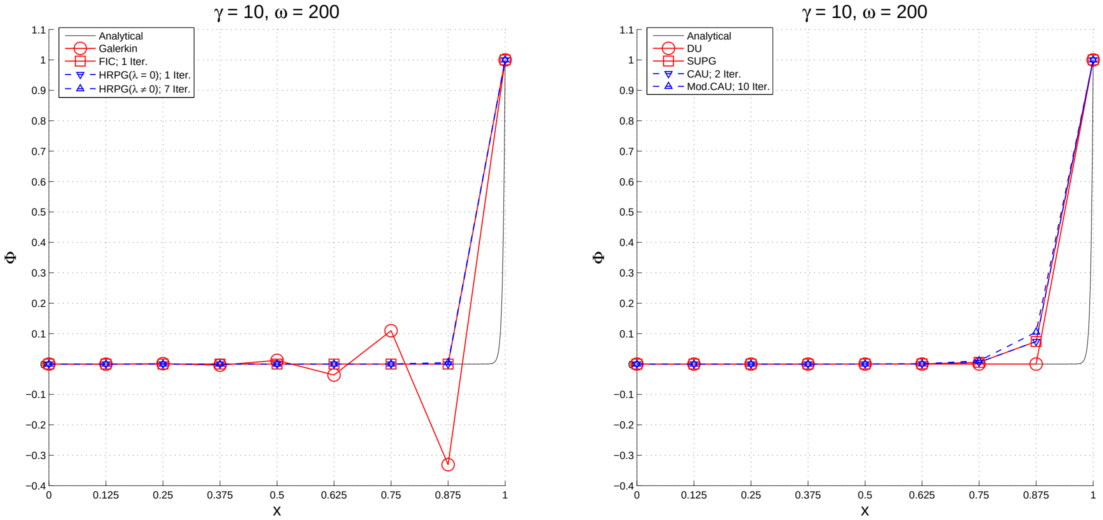
(a) (b)

Figure 15: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (10, 4, 0, 8, 3)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



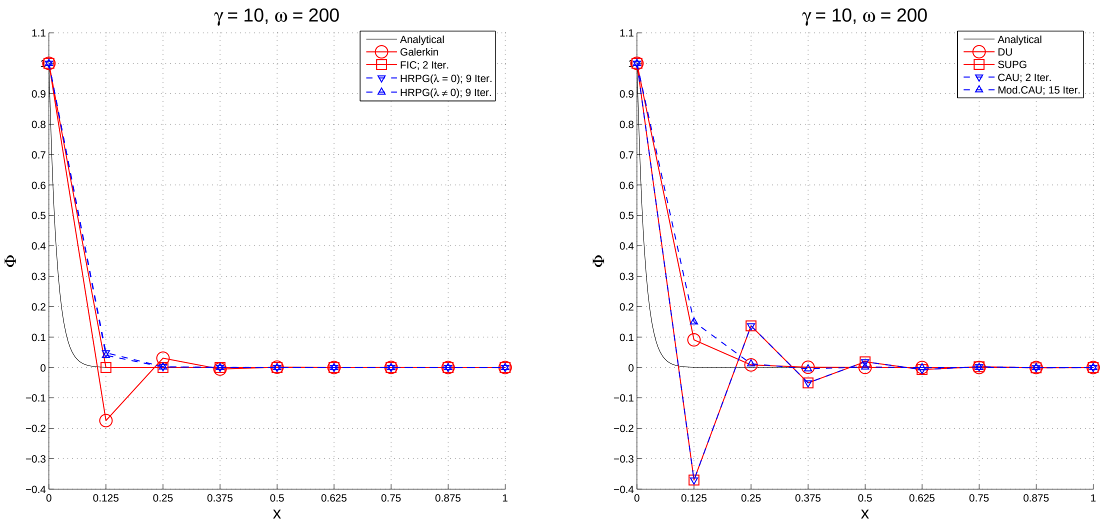
(a) (b)

Figure 16: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (10, 20, 0, 8, 3)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



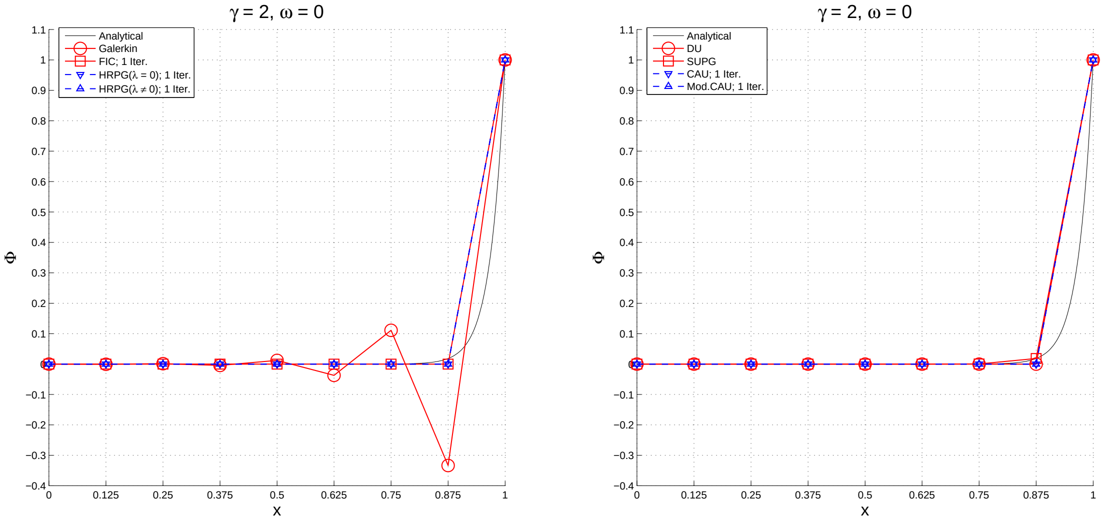
(a) (b)

Figure 17: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (10, 200, 0, 0, 1)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



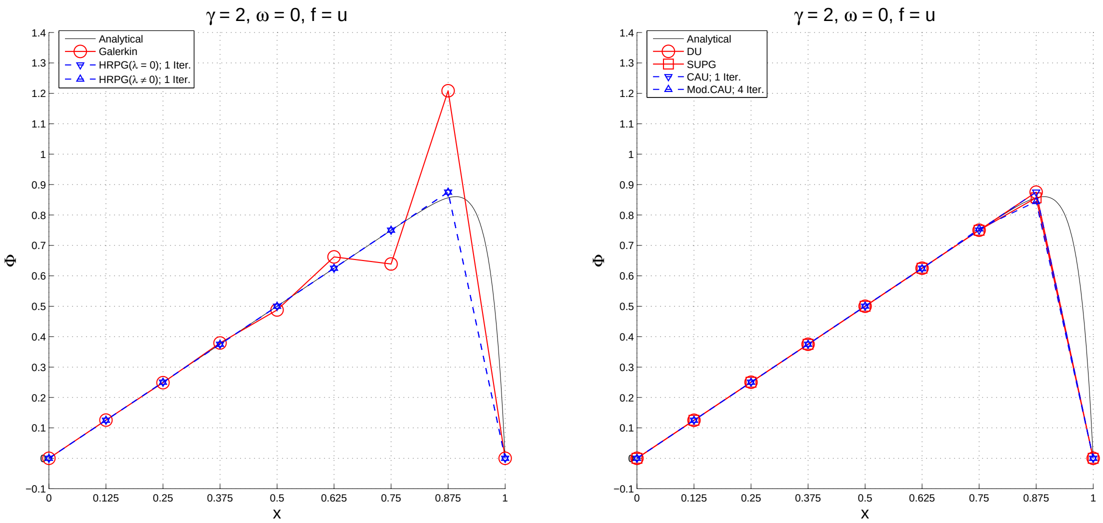
(a) (b)

Figure 18: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (10, 200, 0, 1, 0)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



(a) (b)

Figure 19: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (2, 0, 0, 0, 1)$. (a) Exact, Galerkin, FIC [41], HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions



(a) (b)

Figure 20: Steady state; $(\gamma, \omega, f, \phi_L^p, \phi_R^p) = (2, 0, u, 0, 0)$. (a) Exact, Galerkin, HRPG($\lambda = 0$) and HRPG($\lambda \neq 0$) solutions; (b) Exact, DU, SUPG, CAU and Mod.CAU solutions

The above initial condition models a double rectangular pulse with simple discontinuities. The amplitude spectrum of this function decays only as fast as the *harmonic series* and hence is rich in high wave numbers. It is a challenging problem for the validation of any method for the control of dispersive oscillations and accuracy. The 1D domain is taken as $x \in [0, 1]$ and it is discretized with 200 two-node linear elements. The time step was chosen as $\Delta t = 0.001$s. This corresponds to a Courant number $C = 0.2$. The error was measured using Eq.(79) and a tolerance of `1e-4` was used. For the HRPG method with $\lambda = 0$ a tolerance of `1e-3` was used. The nonlinear iterations at every time step were initialized by the solution obtained by the SUPG method.

Figures 21,22 illustrate the solution obtained with the SUPG, CAU, modified CAU and HRPG methods. As expected the SUPG solution exhibits dispersive oscillations viz. Figure 21a. Appreciable control over the dispersive oscillations is obtained in the CAU method viz. Figure 21b. However slight crests and troughs do appear in the solution that gradually die out in time. These crests and troughs are reduced in the solution obtained by the modified CAU method at the cost of accuracy viz. Figure 21c. Best results were obtained with the HRPG method with $\lambda \neq 0$ which exhibits better control over the dispersive oscillations and maintains the symmetry of the initial solution viz. Figure 22b-c. On the other hand the HRPG method with $\lambda = 0$ is more diffusive than with $\lambda \neq 0$ and needs more iterations per time step for convergence viz. Figure 22a.

6 Extension to multidimensions

It is well known that the solution to the stationary convection–diffusion–reaction problem may develop two types of layers: *exponential* and *parabolic* layers. The exponential layers are usually found in the convection-dominant cases and near the boundary or close to the regions where the source term is nonregular. Parabolic layers, which are of larger width than exponential layers, are found in the reaction-dominant cases near the boundary or close to the regions where the source term is nonregular and in the convection-dominated cases along the characteristics of the solution. The later characteristic internal/boundary layers are usually found only in higher dimensions and hence have no instances in 1D. For this reason a direct extension of the definition of the stabilization parameters α, β derived for 1D will not be efficient to resolve these layers. In the following we outline the approach to extend the HRPG method to multidimensions but delay the details to a later work:

1. To design a nondimensional element number that quantifies the characteristic internal/boundary layers. By quantification we mean that it should serve a similar purpose as the element Peclet number γ for the exponential layers in convection-dominant cases and the dimensionless number $\omega := 2\gamma\sigma$ for the parabolic layers in the reaction-dominant cases.

2. Update the definition of the stabilization parameter β to include this new dimensionless number. Of course this update takes effect only in higher dimensions.

3. Using this updated definition of the stabilization parameters the characteristic lengths are calculated as: $\mathbf{h} := \alpha^i \mathbf{l}^i$, $\mathbf{H} := (\beta^i/|\mathbf{l}^i|)[\mathbf{l}^i \otimes \mathbf{l}^i]$. Where $\mathbf{l}^i$ are frame-independent element length vectors and α^i, β^i are calculated along these $\mathbf{l}^i$. Thus for the 2D bilinear

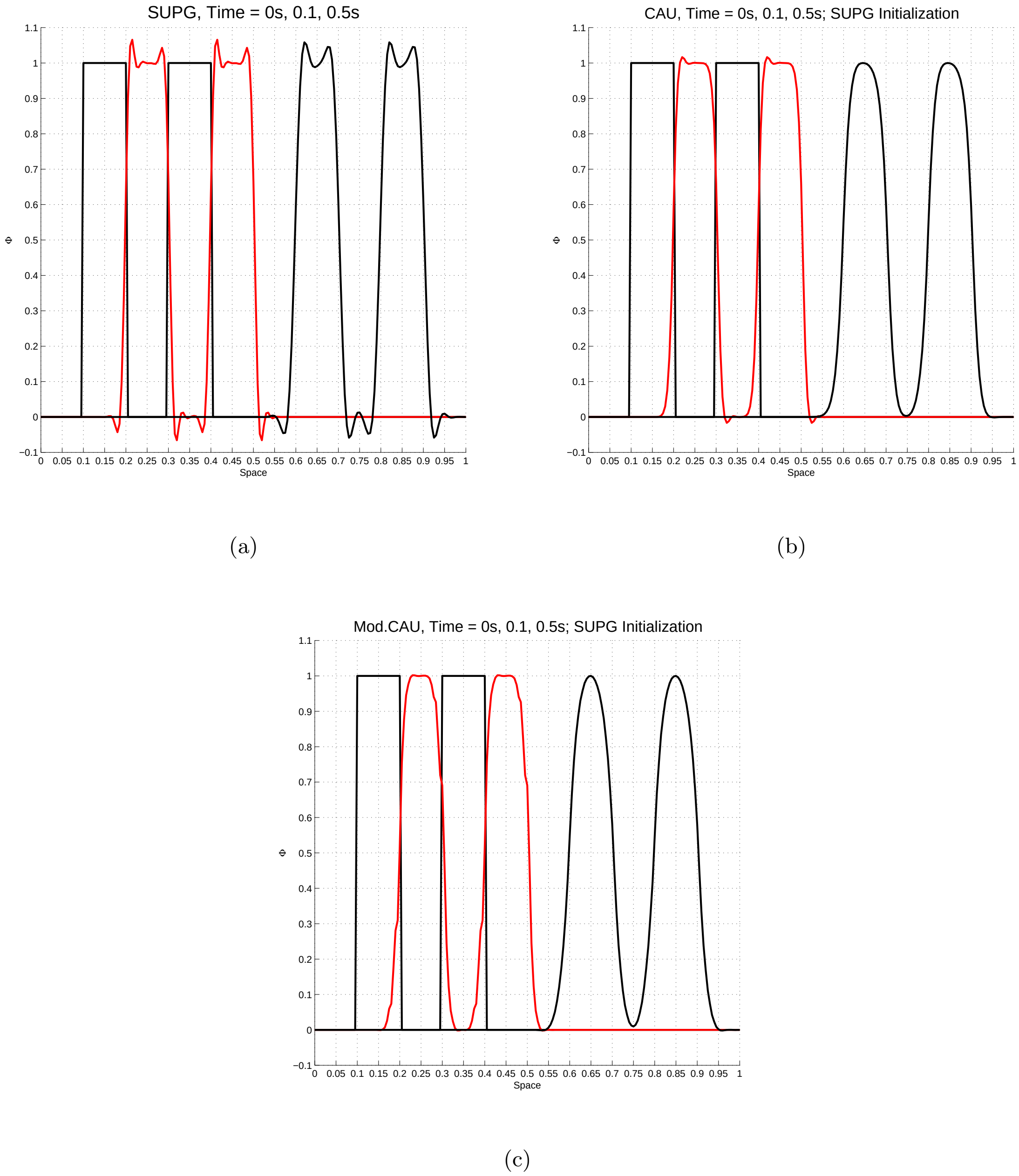
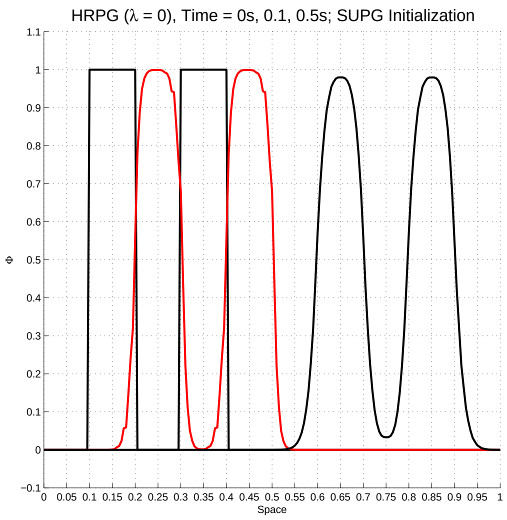
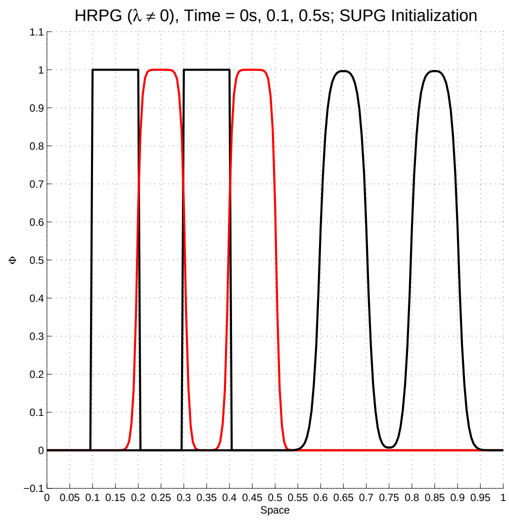


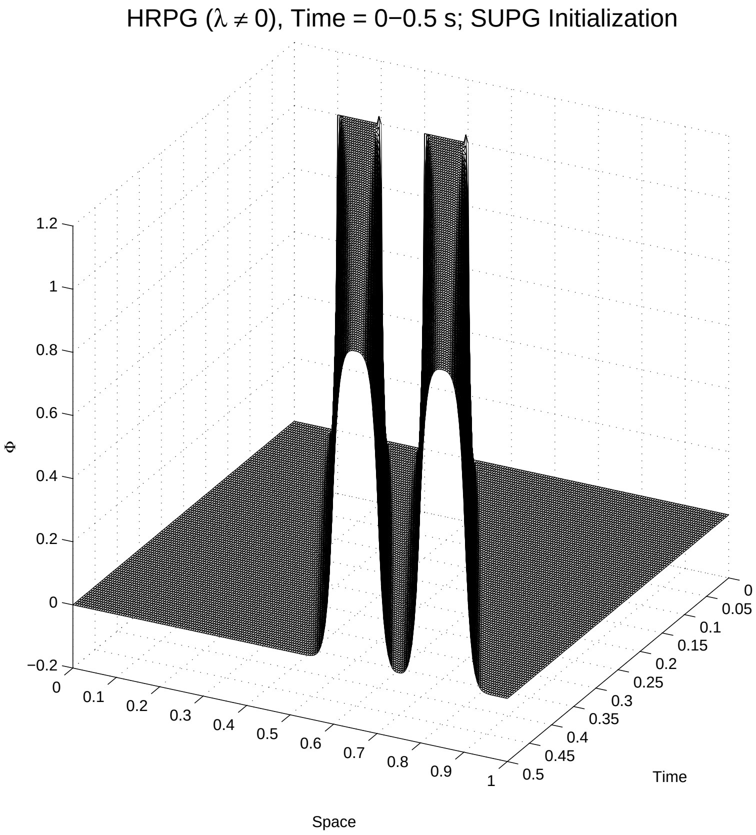
Figure 21: Transient pure convection; $u = 1$m/s, $\ell = 0.005$m, $\Delta t = 0.001$s. (a) SUPG solution; (b) CAU solution; (c) Mod.CAU solution



(a)



(b)



(c)

Figure 22: Transient pure convection; $u = 1$m/s, $\ell = 0.005$m, $\Delta t = 0.001$s. (a) HRPG($\lambda = 0$) solution; (b) HRPG($\lambda \neq 0$) solution; (c) HRPG($\lambda \neq 0$) solution (evolution plot)

quadrilateral elements the characteristic lengths would be:

$$\mathbf{h} := \alpha^1 \mathbf{l}^1 + \alpha^2 \mathbf{l}^2 \quad ; \quad \mathbf{H} := \frac{\beta^1}{|\mathbf{l}^1|}[\mathbf{l}^1 \otimes \mathbf{l}^1] + \frac{\beta^2}{|\mathbf{l}^2|}[\mathbf{l}^2 \otimes \mathbf{l}^2] \tag{81}$$

4. Using $\mathbf{h}, \mathbf{H}$ as defined above we calculate the perturbation p_h associated with the HRPG method via the Eq.(6). This completes the definition of the HRPG method in multidimensions.

Remark : As noted in §2 the multidimensional HRPG method could be arrived at from the FIC equations with an appropriate definition of the characteristic length as $\mathbf{h}^{fic} := \mathbf{h} + \mathbf{H} \cdot \hat{\mathbf{u}}^r$.

7 Conclusions

A high-resolution Petrov–Galerkin method is presented for the 1D convection–diffusion–reaction problem. The prefix 'high-resolution' is used here in the sense popularized by Harten, i.e. second order accuracy for smooth/regular regimes and good shock-capturing in nonregular regimes. The HRPG method could be understood as the combination of upwinding plus a non-linear discontinuity-capturing operator. The distinction is that in general(multidimensions) the upwinding provided by $\mathbf{h}$ is not streamline and the discontinuity-capturing provided by $\mathbf{H} \cdot \hat{\mathbf{u}}^r$ is neither isotropic nor purely crosswind. The HRPG form can be considered as a particular class of the stabilized governing equations obtained via a finite calculus (FIC) procedure. For the 1D problem the HRPG method is similar to the CAU method with new definitions of the stabilization parameters. The 1D examples presented demonstrate that the method provides stabilized and essentially non-oscillatory i.e. monotone *to-the-eye* solutions for a wide range of the physical parameters and boundary conditions. It is interesting to note that the HRPG method without the linear upwinding term, i.e. using $\alpha = 0$ does solve all the steady-state examples to give high-resolution stabilized results. Nevertheless the presence of the linear perturbation terms improves the convergence of the nonlinear iterations especially for the transient problem.

8 Acknowledgments

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

References

[1] Hughes TJR, Brooks AN. "A theoretical framework for Petrov–Galerkin methods, with discontinuous weighting functions: application to the streamline upwind procedure," in Gallagher RH, Norrie DM, Oden JT, Zienkiewicz OC (eds.), *Finite Elements in Fluids* 1982, Vol. IV,Wiley, Chichester.

[2] Brooks AN, Hughes TJR. "Streamline Upwind/Petrov–Galerkin formulations for the convective dominated flows with particular emphasis on the incompressible Navier-Stokes equations," *Computer Methods in Applied Mechanics and Engineering* 1982, Vol. 32, pp. 199-259.

[3] Douglas J, Russell TF. "Numerical methods for convection-dominated diffusion problems based on combining the method of characteristics with the finite element or finite difference procedures," *SIAM Journal of Numerical Analysis* 1982, Vol. 19, pp. 871-885

[4] Donea J. "A Taylor-Galerkin method for convective transport problems," *International Journal for Numerical Methods in Engineering* 1984, Vol. 20, pp. 101-119.

[5] Löhner R, Morgan K, Zienkiewicz OC. "The solution of non-linear hyperbolic equation systems by the finite element method," *International Journal for Numerical Methods in Fluids* 1984, Vol. 4, pp. 1043-1063.

[6] 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, Vol. 73, pp. 173-189.

[7] Brezzi F, Bristeau MO, Franca LP, Mallet M, Rogé G. "A relationship between stabilized finite element methods and the Galerkin method with bubble functions," *Computer Methods in Applied Mechanics and Engineering* 1992, Vol. 96, pp. 117-129.

[8] Baiocchi C, Brezzi F, Franca LP. "Virtual bubbles and the GaLS," *Computer Methods in Applied Mechanics and Engineering* 1993, Vol. 105, pp. 125-141.

[9] Brezzi F, Russo A. "Choosing bubbles for advection-diffusion problems," *Mathematical Models and Methods in Applied Sciences* 1994, Vol. 4, pp. 571-587.

[10] Hughes TJR. "Multiscale phenomena: Green's function, the Dirichlet-to-Neumann formulation, subgrid scale models, bubbles and the origins of stabilized formulations" *Computer Methods in Applied Mechanics and Engineering* 1995, Vol. 127, pp. 387-401.

[11] Hughes TJR, Feijoo GR, Mazzei L, Quincy JB. "The variational multiscale method: a paradigm for computational mechanics," *Computer Methods in Applied Mechanics and Engineering* 1998, Vol. 166, pp. 3-24.

[12] Zienkiewicz OC, Codina R. "A general algorithm for compressible and incompressible flows. Part I: the split, characteristic based scheme," *International Journal for Numerical Methods in Fluids* 1995, Vol. 20, pp. 869-885.

[13] Oñate E. "Derivation of stabilized equations for numerical solution of advective-diffusive transport and fluid flow problems," *Computer Methods in Applied Mechanics and Engineering* 1998, Vol. 151, pp. 233-265.

[14] Oñate E, Manzan M. "Stabilization techniques for finite element analysis of convection-diffusion problems," in Sundén B, Comini G. (eds.), *Computational Analysis of Convection Heat Transfer* 2000, pp. 71-117, WIT Press, Southampton (UK), 2000.

[15] Brezzi F, Franca LP, Hughes TJR, Russo A. "$b = \int g$," *Computer Methods in Applied Mechanics and Engineering* 1997, Vol. 142, pp. 353-360.

[16] Codina R. "Comparison of some finite element methods for solving the diffusion-convection-reaction equation," *Computer Methods in Applied Mechanics and Engineering* 1998, Vol. 156, pp. 185-210.

[17] Mizukami A, Hughes TJR. "A Petrov–Galerkin finite element method for convection-dominated flows: An accurate upwinding technique for satisfying the maximum principle," *Computer Methods in Applied Mechanics and Engineering* 1985, Vol. 50, pp. 181-193.

[18] Hughes TJR, Mallet M, Mizukami A. "A new finite element formulation for computational fluid dynamics: II. Beyond SUPG," *Computer Methods in Applied Mechanics and Engineering* 1986, Vol. 54, pp. 341-355.

[19] Galeão AC, Dutra do Carmo EG. "A consistent approximate upwind Petrov–Galerkin method for the convection-dominated problems," *Computer Methods in Applied Mechanics and Engineering* 1988, Vol. 68, pp. 83-95.

[20] Johnson C, Szepessy A, Hansbo P. "On the convergence of shock-capturing streamline diffusion finite element methods for hyperbolic conservation laws," *Mathematics of Computation* 1990, Vol. 54, pp. 107-129.

[21] Dutra do Carmo EG, Galeão AC. "Feedback Petrov–Galerkin methods for convection-dominated problems," *Computer Methods in Applied Mechanics and Engineering* 1991, Vol. 88, pp. 1-16.

[22] Codina R. "A discontinuity-capturing crosswind-dissipation for the finite element solution of the convection-diffusion equation," *Computer Methods in Applied Mechanics and Engineering* 1993, Vol. 110, pp. 325-342.

[23] de Sampaio PAB, Coutinho ALGA. "A natural derivation of discontinuity capturing operator for convection-diffusion problems," *Computer Methods in Applied Mechanics and Engineering* 2001, Vol. 190, pp. 6291-6308.

[24] Burman E, Ern A. "Nonlinear diffusion and discrete maximum principle for stabilized Galerkin approximations of the convection–diffusion–reaction equation," *Computer Methods in Applied Mechanics and Engineering* 2002, Vol. 191, pp. 3833-3855.

[25] Dutra do Carmo EG, Alvarez GB. "A new stabilized finite element formulation for the scalar convection-diffusion problems: the streamline and approximate upwind/Petrov–Galerkin method," *Computer Methods in Applied Mechanics and Engineering* 2003, Vol. 192, pp. 3379-3396.

[26] Dutra do Carmo EG, Alvarez GB. "A new upwind function in stabilized finite element formulations, using linear and quadratic elements for the scalar convection-diffusion problems," *Computer Methods in Applied Mechanics and Engineering* 2004, Vol. 193, pp. 2383-2402.

[27] Knobloch P. "Improvements of the Mizukami-Hughes method for convection-diffusion equations," *Computer Methods in Applied Mechanics and Engineering* 2006, Vol. 196, pp. 579-594.

[28] Oñate E, Zarate F, Idelsohn SR. "Finite element formulation for the convective-diffusive problems with sharp gradients using finite calculus," *Computer Methods in Applied Mechanics and Engineering* 2006, Vol. 195, pp. 1793-1825.

[29] John V, Knobloch P. "On spurious oscillations at layers diminishing (SOLD) methods for convection-diffusion equations: Part I - A review," *Computer Methods in Applied Mechanics and Engineering* 2007, Vol. 196, pp. 2197-2215.

[30] Tezduyar TE, Park YJ. “Discontinuity-capturing finite element formulations for nonlinear convection–diffusion–reaction equations,” *Computer Methods in Applied Mechanics and Engineering* 1986, Vol. 59, pp. 307-325.

[31] Idelsohn SR, Nigro N, Storti M, Buscaglia G. “A Petrov–Galerkin formulation for advective-reaction-diffusion problems,” *Computer Methods in Applied Mechanics and Engineering* 1996, Vol. 136, pp. 27-46.

[32] Franca LP, Dutra do Carmo EG. “The Galerkin gradient least-squares method,” *Computer Methods in Applied Mechanics and Engineering* 1989, Vol. 74, pp. 41-54.

[33] Harari I, Hughes TJR. “Stabilized finite element methods for steady advection-diffusion with production,” *Computer Methods in Applied Mechanics and Engineering* 1994, Vol. 115, pp. 165-191.

[34] Franca LP, Farhat C. “Bubble functions prompt unusual stabilized finite element methods,” *Computer Methods in Applied Mechanics and Engineering* 1995, Vol. 123, pp. 299-308.

[35] Franca LP, Valentin F. “On an improved unusual stabilized finite element method for the advective-reactive-diffusive equation,” *Computer Methods in Applied Mechanics and Engineering* 2001, Vol. 190, pp. 1785-1800.

[36] Brezzi F, Hauke G, Marini LD, Sangalli G. “Link-cutting bubbles for the stabilization of convection–diffusion–reaction problems,” *Mathematical Models and Methods in Applied Sciences* 2003, Vol. 13, No. 3, pp. 445-461.

[37] Codina R. “On stabilized finite element methods for linear systems of convection–diffusion–reaction equations,” *Computer Methods in Applied Mechanics and Engineering* 2000, Vol. 190, pp. 2681-2706.

[38] Hauke G, García-Olivares A. “Variational subgrid scale formulations for the advection-diffusion–reaction equation,” *Computer Methods in Applied Mechanics and Engineering* 2001, Vol. 190, pp. 6847-6865.

[39] Hauke G. “A simple subgrid scale stabilized method for the advection-diffusion–reaction equation,” *Computer Methods in Applied Mechanics and Engineering* 2002, Vol. 191, pp. 2925-2947.

[40] Hauke G, Sangalli G, Doweidar MH. “Combining adjoint stabilized methods for the advection-diffusion–reaction problem,” *Mathematical Models and Methods in Applied Sciences* 2007, Vol. 17, No. 2, pp. 305-326.

[41] Oñate E, Miquel J, Hauke G. “Stabilized formulation for the advection-diffusion-absorption equation using finite calculus and linear finite elements,” *Computer Methods in Applied Mechanics and Engineering* 2006, Vol. 195, pp. 3926-3946.

[42] Oñate E, Miquel J, Zarate F. “Stabilized solution of the multidimensional advection-diffusion-absorption equation using linear finite elements,” *Computers and Fluids* 2007, Vol. 36, pp. 92-112.

[43] Felippa C, Oñate E. “Nodally exact Ritz discretizations of the 1D diffusion-absorption and Helmholtz equations by variational FIC and modified equation methods,” *Computational Mechanics* 2007, Vol. 39, No.2, pp. 91-111.

[44] Idelsohn SR, Heinrich JC, Oñate E. “Petrov–Galerkin methods for the transient advective-diffusive equation with sharp gradients,” *International Journal for Numerical Methods in Engineering* 1996, Vol. 39, pp. 1455-1473.

[45] Yu CC, Heinrich JC. "Petrov–Galerkin methods for the time-dependent convective transport equations," *International Journal for Numerical Methods in Engineering* 1986, Vol. 23, pp. 883-901.

[46] Hirsch C. *Numerical computation of internal and external flows,* Wiley, 1990, Vols. 1 and 2.

[47] LeVeque RJ. *Numerical methods for conservation laws,* Birkhauser-Verlag: Basel, 1990.

[48] Laney CB. *Computational gas dynamics,* Cambridge University Press, 1998.

[49] Toro EF. *Riemann solvers and numerical methods for fluid dynamics,* Springer-Verlag, 1999.

[50] Boris JP, Book DL. "Flux-corrected transport. I. SHASTA, A fluid transport algorithm that works," *Journal of Computational Physics* 1973, Vol. 11, pp. 38-69.

[51] Zalesak ST. "Fully multidimensional flux-corrected transport algorithms for fluids," *Journal of Computational Physics* 1979, Vol. 31, pp. 335-362.

[52] van Leer B. "Towards the ultimate conservative difference scheme, V. A second order sequel to Godunov's method," *Journal of Computational Physics* 1979, Vol. 32, pp. 101-136.

[53] Harten A. "High resolution schemes for hyperbolic conservation laws," *Journal of Computational Physics* 1983, Vol. 49, pp. 357-393.

[54] Harten A, Engquist B, Osher S, Chakravarthy S. "Uniformly higher order essentially non-oscillatory schemes, III," *Journal of Computational Physics* 1987, Vol. 71, pp. 231-303.

[55] Kuzmin D, Löhner R, Turek S. *Flux-corrected transport: principles, algorithms and applications,* Springer, 2005.

[56] Oden JT. "Historical comments on finite elements" *Proceedings of the ACM conference on history of scientific and numeric computation* Princeton, New Jersey USA, 1987, pp. 125-130.

[57] Zienkiewicz OC, Taylor RL, Nithiarasu P. *The finite element method for fluid dynamics,* Elsevier Butterworth-Heinemann, 2005.