

Entropy Stable Discontinuous Galerkin Scheme for the Compressible Navier–Stokes Equations

Mohammad Zakerzadeh * and Georg May [†]

Graduate School AICES, RWTH Aachen, 52062 Aachen, Germany

In this paper we present a family of entropy-stable discontinuous Galerkin methods for the compressible Navier-Stokes equations. The discretization presented here is based on the mixed formulation, and is designed to preserve entropy stability of an already existing discretization for the Euler equations. Entropy stable variants of several known schemes are presented, including BR2, SIPG and LDG.

I. Introduction

The class of nonlinear systems of conservation laws contains many important examples. A case in point are the Euler equations and the Navier-Stokes equations, governing inviscid and viscous compressible fluid flow, respectively. In analyzing the stability of numerical schemes for such systems, entropy stability is often the framework of choice.^{3–5,11,12,17–19,24,25,33,34} Furthermore, some convergence proofs have been provided for high order space-time discontinuous Galerkin (DG) schemes and more general hyperbolic systems which heavily exploit the property of entropy stability.^{23,39}

Compared to entropy stability analysis of hyperbolic conservation laws, much less work has been done in extending entropy stability of DG schemes to nonlinear convection-diffusion systems and most of the literature try to mimic the well-known analysis for the linear Poisson problems e.g., in.¹ To the best knowledge of the authors only very few results are available in this direction^{12, 27, 35} (also see¹⁶). More specifically, in Ref. 35 a difference formulation based on entropy variables has been presented and the entropy stability of the method has been proved in the semi-discrete form. In Ref. 12 the symmetric/nonsymmetric interior penalty DG formulation has been presented for the one-dimensional Navier-Stokes equations realized in terms of entropy variables, and entropy stability has been proved in the semi-discrete settings. In Ref. 27 a different formulation of the interior penalty has been presented employing a different penalty scaling, similar to Ref. 22, for the one-dimensional Navier-Stokes equations. Furthermore, a variant of LDG method is presented for this problem, as well as an entropy stability proof in space-time fully-discrete settings.

Another work is the Ref. 15 which presents an entropy-stable space-time formulation for turbulent computation using the Roe flux for the convective part, and the second method of Bassi and Rebay⁷ for the viscous part. Though no proof has been provided for entropy stability, the authors have given illustrative examples by comparison between using conservative and entropy variables. The authors highly emphasized the importance of having an entropy stable method as a base-line scheme, despite issues arising from inexact integration. In fact, the authors showed that while *de-aliasing* process by over-integration²⁶ is sufficiently helpful for an entropy stable scheme in high Reynolds flow computation, it is unable to provide a stable formulation for under-resolved turbulent simulation if one realize the formulation in terms of conservative variables, which does not enjoy the entropy stability.

In this work we consider a more general DG formulation for the Navier-Stokes equations by using a mixed method for discretizing the viscous flow in addition to the entropy stable formulation for the convective flux as we presented in our earlier works^{38,39} (also see e.g.,²³), which is able to represent a family of different known viscous discretization schemes, extended to the nonlinear diffusion of the Navier-Stokes equations in a canonical way. Most of the schemes in the classic paper by Arnold et al.¹ can be included and proved to be entropy stable. We give a variety of examples in this abstract, augmented with numerical examples.

^{*}Graduate student.

[†]Associate professor.

The outline of this paper is as follows: The governing equations and relevant symmetrization concepts are recalled in §II. The DG discretization framework is introduced in §III. This section includes the explicit form of the convective and a variety of viscous discretizations. The stability analysis is presented in §IV, and §V is reserved for numerical examples.

II. Governing Equations

Let us consider the compressible Navier-Stokes equations in \mathbb{R}^d with d = 1, 2, 3 in the divergence form

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}t} + \nabla \cdot \left(\boldsymbol{f}_{c}(\boldsymbol{u}) - \boldsymbol{f}_{v}(\boldsymbol{u}, \nabla \boldsymbol{u})\right) = 0, \qquad \text{in } \Omega,$$
(1)

where $\Omega \subset \mathbb{R}^d$ is open and bounded. Here $u \in \mathbb{R}^m$ is the vector of conservative variables and by f_c and f_v we denote the convective and diffusive fluxes respectively as, e.g. for d = 3 (and m = 5) reads as

$$\boldsymbol{u} = \begin{bmatrix} \rho \\ \rho V_1 \\ \rho V_2 \\ \rho V_3 \\ E \end{bmatrix}, \qquad \boldsymbol{f}_c^i = V_i \boldsymbol{u} + p \begin{bmatrix} 0 \\ \delta_{1,i} \\ \delta_{2,i} \\ \delta_{3i} \\ V_i \end{bmatrix}, \qquad \boldsymbol{f}_v^i = \begin{bmatrix} 0 \\ \tau_{1,i} \\ \tau_{2,i} \\ \tau_{3,i} \\ \tau_{i,j} V_j + q_i \end{bmatrix}, \quad i = 1, \cdots, d$$
(2)

with $V = (V_1, V_2, V_3)^t$ is the velocity vector and the convective/diffusive fluxes $\mathbf{f}_{c/v} = (\mathbf{f}_{c/v}^1, \mathbf{f}_{c/v}^2, \mathbf{f}_{c/v}^3)^t$. Also $\delta_{i,j}$ is the notation for the Kronecker delta function and we used the convention of summing over repeated indices.

Here ρ and E are density and internal energy, respectively. Also p denotes the static pressure defined as

$$p = (\gamma - 1)(E - \frac{1}{2}\rho V^2)$$
(3)

where $\gamma = c_p/c_v$ is the ratio of specific heat capacities at constant pressure (c_p) and constant volume (c_v) . For air we have $\gamma = 1.4$. Moreover, τ is the viscous shear stress tensor and for Newtonian fluid is

$$\tau = \mu \Big(\nabla \boldsymbol{u} + (\nabla \boldsymbol{u})^t - \frac{2}{3} (\nabla \cdot \boldsymbol{u}) I \Big), \tag{4}$$

and the heat flux q is defined by the Fourier's law as

$$q_i = \kappa \frac{\partial T}{\partial x_i}, \qquad i = 1, \cdots, d,$$
(5)

where κ is the heat conductivity and is equal to $\kappa = \frac{\mu c_p}{P_r}$. Here Pr is the Prandtl number, which for air at moderate conditions has a constant value of about Pr = 0.72. Also the temperature T is defined by the ideal gas law as

$$T = \frac{p}{(\gamma - 1)c_v\rho}.$$

Moreover, we need to supplement equation (1) with some boundary conditions on the boundary of Ω denoted by Γ . For simplicity in the formulation we consider here only the Dirichlet boundary conditions in the form of $u_{\Gamma} = g_D$ and refer to the literature, e.g Ref. 21, for changes required for Neumann boundary conditions. In order to have a well-posed problem in the hyperbolic settings we need to consider some special treatment for imposing g_D in different part of the boundary. Let us consider the following decomposition of the boundary Γ as

$$\Gamma = \Gamma_{D, sup} \cup \Gamma_{D, sub-in} \cup \Gamma_{D, sub-out} \cup \Gamma_{W}$$
(6)

to supersonic (inflow/outflow), subsonic inflow, subsonic outflow and the wall boundary conditions, respectively. For brevity we skip the details here and refer again to the literature²² for the treatments of inflow/outflow and wall boundary conditions. We denote the imposed boundary condition by $u_{\Gamma}(u)$ or simply u_{Γ} . One can write (1) in the form²⁴

$$\boldsymbol{u}_t + \boldsymbol{A}_i(\boldsymbol{u}) \nabla_i \boldsymbol{u} - \nabla \cdot (\boldsymbol{K}_{i,j}(\boldsymbol{u}) \nabla_j \boldsymbol{u}) = 0, \qquad i, j = 1, \cdots, d,$$
(7)

where $A_i(u) = \partial_u f^i(u)$ and $f^i_v(u) = K_{i,j}(u) \nabla_j u$.

In general $A_i(u)$ and $K(u) = [K_{i,j}], i, j = 1, \dots, d$ (where each $K_{i,j} \in \mathbb{R}^{m \times d}$) are not symmetric and this makes the energy analysis somewhat cumbersome. An idea to detour this issue is to using entropy variables instead of conservative ones; i.e. using the change of variables $v(u) = U_u(u)$ where U is a convex entropy function of (1). Here we refer for more details to Ref. 24 and just mention that in case of nonzero heat flux the only admissible entropy for symmetrization is an affine function of the specific entropy $s = \log(\frac{p}{q^{\gamma}})$, e.g.

$$U = -\frac{\rho s}{\gamma - 1}, \qquad Q_i = U V_i, \tag{8}$$

where $Q_i, i = 1, \dots, d$, is the corresponding entropy flux. Note that $\partial_u Q = \partial_u U \partial_u f_c$. Then, the symmetric representation of (7) realized in entropy variables can be written as

$$\boldsymbol{u}_{\boldsymbol{v}}\boldsymbol{v}_t + \tilde{\boldsymbol{A}}_i(\boldsymbol{v})\nabla_i\boldsymbol{v} - \nabla\cdot(\tilde{\boldsymbol{K}}_{i,j}(\boldsymbol{v})\nabla_j\boldsymbol{v}) = 0, \qquad i, j = 1, \cdots, d,$$
(9)

such that $\tilde{A}_i(v)$ is symmetric and $\tilde{K}(v)$ is symmetric positive semi-definite. Also note that the matrix u_v is the inversion of the Hessian of the convex entropy function U and is symmetric positive definite. For explicit form of \tilde{A}_i and $\tilde{K}_{i,j}$ we refer to Ref. 24.

Henceforth, we realize the functions in terms of entropy variables v which are the basic unknowns, and the dependent conservative variables are derived via mapping u(v). In our notation, this mapping is sometimes omitted, e.g., f(v) is written rather than f(u(v)). Also for brevity in notation we sometimes might ignore the dependence of $\tilde{K}(v)$ to v and write \tilde{K} . We write the argument whenever it might lead to confusion.

III. Discontinuous Galerkin Formulation

In this section we explain the weak formulation and the DG discretization we apply on (1). This includes the properties of triangulation and the discretization techniques for both the convective and the viscous part.

A. Triangulation

Let us consider the space domain Ω has a polygon boundary Γ . Then we consider a shape-regular triangulation on Ω as $\mathcal{T}_h = \{\kappa\}$ composed of (non-overlapping) triangular or rectangular elements (with possible hanging nodes). Let define h_{κ} as the diameter of each $\kappa \in \mathcal{T}_h$ and $h := \max_{\kappa \in \mathcal{T}_h} h_{\kappa}$. Also we denote ν_{κ} to be the outward normal to $\partial \kappa$. In the following we assume that \mathcal{T}_h is of *bounded variation*, that is, there exists a constant l > 1 such that

$$l^{-1} \le \frac{h_{\kappa}}{h_{\kappa'}} \le l,\tag{10}$$

where $\kappa, \kappa' \in \mathcal{T}_h$ share an edge. This bounded variation property means that there is an upper bound for the number of neighboring elements of each $\kappa \in \mathcal{T}_h$, denoted by N_l . In case that \mathcal{T}_h has no hanging nodes, $N_l = 3$ and $N_l = 4$ for triangular and rectangular elements, respectively. We denote the *skeleton* of the triangulation, i.e. the set of all edges of $\kappa \in \mathcal{T}_h$, by \mathcal{E}_h . Also we denote the set of boundary and interior edges of \mathcal{T}_h by $\mathcal{E}_{h,\partial}$ and $\mathcal{E}_{h,I}$, respectively, and the length of edge e by h_e .

Following standard definitions,¹ let us fix the jump and average of the discontinuous functions on the skeleton \mathcal{E}_h . For any interior edge $e \in \mathcal{E}_{h,I}$, where e is the common edge of $\kappa, \kappa' \in \mathcal{T}_h$, with $\boldsymbol{w}_{\kappa,e} = \boldsymbol{w}_{\kappa}|_e$ we set

$$\{\boldsymbol{w}\} = \frac{1}{2}(\boldsymbol{w}_{\kappa,e} + \boldsymbol{w}_{\kappa',e}), \qquad [\![\boldsymbol{w}]\!] = \boldsymbol{w}_{\kappa,e} \otimes \nu_{\kappa} + \boldsymbol{w}_{\kappa',e} \otimes \nu_{\kappa'}$$
(11)

for all $\boldsymbol{w} \in \prod_{\kappa \in \mathcal{T}_h} [L_2(\partial \kappa)]^m$. Similarly for all $\boldsymbol{\tau} \in \prod_{\kappa \in \mathcal{T}_h} [L_2(\partial \kappa)]^{(m \times d)}$ we set

$$\{\boldsymbol{\tau}\} = \frac{1}{2}(\boldsymbol{\tau}_{\kappa,e} + \boldsymbol{\tau}_{\kappa',e}), \qquad \llbracket \boldsymbol{\tau} \rrbracket = \boldsymbol{\tau}_{\kappa,e} \cdot \boldsymbol{\nu}_{\kappa} + \boldsymbol{\tau}_{\kappa',e} \cdot \boldsymbol{\nu}_{\kappa'}.$$
(12)

Moreover, for any boundary edge $e \in \mathcal{E}_{h,\partial}$ we define

$$\llbracket \boldsymbol{w} \rrbracket = \boldsymbol{w}_{\kappa,e} \otimes \nu_{\kappa}, \qquad \{\boldsymbol{\tau}\} = \boldsymbol{\tau}_{\kappa,e}. \tag{13}$$

Note that we will use two different notations for inner product; $\langle \boldsymbol{w}, \boldsymbol{v} \rangle$ denotes the inner product in the state space between $\boldsymbol{w}, \boldsymbol{v} \in \mathbb{R}^m$, while $a \cdot b$ defines the inner product in the physical space for $a, b \in \mathbb{R}^d$. Moreover we use double inner product notation as $\boldsymbol{\tau} : \boldsymbol{\zeta} = \sum_{i=1}^m \sum_{j=1}^d \boldsymbol{\tau}_{i,j} \boldsymbol{\zeta}_{i,j}$ for $\boldsymbol{\tau}, \boldsymbol{\zeta} \in \mathbb{R}^{m \times d}$. Also for $\boldsymbol{w} \in \mathbb{R}^m$ and $a \in \mathbb{R}^d$ we define the matrix $\boldsymbol{w} \otimes a \in \mathbb{R}^{m \times d}$ as $(\boldsymbol{w} \otimes a)_{i,j} = w_i a_j$.

B. Variational Formulation

The finite dimensional space for the approximate weak solution of (1) is defined as

$$V_{h,q} := \{ \boldsymbol{w}^h \in [L_2(\Omega)]^m : \boldsymbol{w}^h|_{\kappa} \in [\mathcal{P}^q(\kappa)]^m, \quad \forall \kappa \in \mathcal{T}_h \},$$
(14)

where $\mathcal{P}^{q}(\kappa)$ is the space of polynomials of at most degree q on a domain $\kappa \subset \mathbb{R}^{d}$.

The proposed discontinuous Galerkin method has the following quasi-linear (nonlinear in the first argument and linear in the second) variational form in terms of entropy variables: Find $v^h \in V_{h,q}$ such that

$$\langle \boldsymbol{u}_{t}^{h}, \boldsymbol{w}^{h} \rangle + \mathcal{B}(\boldsymbol{v}^{h}, \boldsymbol{w}^{h}) = \langle \boldsymbol{u}_{t}, \boldsymbol{w}^{h} \rangle + \mathcal{B}^{c}(\boldsymbol{v}^{h}, \boldsymbol{w}^{h}) + \mathcal{B}^{v}(\boldsymbol{v}^{h}, \boldsymbol{w}^{h}) = 0, \qquad \forall \boldsymbol{w}^{h} \in V_{h,q}.$$
(15)

Here \mathcal{B}^c and \mathcal{B}^v correspond to convective and viscous discretization, respectively. We are going to present the details of the convective and viscous discretization in sections 1 and 2. Note that in (15) we let the time integration term remain in the semi-discrete form. The reason is that the results we are going to discuss in this paper are quite independent of the time discretization approach. The only important feature of the time treatment process that it should keep the entropy stability property (cf. section IV). There are different types of such time integration methods, e.g. explicit SSP time integration and class of implicit methods like space-time formulation (cf. Ref. 3).

In our previous work for hyperbolic systems,^{23,39} one or two additional stabilization terms, in form of *shock capturing* and *streamline diffusion*, were added to guarantee convergence of the method and controlling the discontinuities. We remark that these terms are not necessary in the entropy analysis of the method, but they do improve the quality of the result and alleviate oscillations, especially for coarser meshes and high Reynolds number settings. In this work, however, we do not add such terms in the analysis and numerical experiments, which focus on simple model equations and smooth solutions, and refer to the literature for more details on further stabilization techniques.^{23,27,28,39}

In the rest of this section we provide details on each of the terms in (15):

1. Convective Discretization

Using the test function $w^h \in \mathcal{V}^q$ to penalize the interior residual of the cell and jump of the fluxes leads to

$$\mathcal{B}^{c}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\langle\nabla\cdot\boldsymbol{f}_{c},\boldsymbol{w}^{h}\rangle\,\mathrm{d}x + \sum_{\kappa\in\mathcal{T}_{h}}\int_{\partial\kappa}\langle\hat{\boldsymbol{f}}_{c}(\boldsymbol{v}^{h}) - \boldsymbol{f}_{c}(\boldsymbol{v}^{h}_{\kappa,e})\cdot\nu_{\kappa},\boldsymbol{w}^{h}_{\kappa,e}\rangle\,\mathrm{d}s.$$
(16)

Here, $\hat{\boldsymbol{f}}_c(\boldsymbol{v}^h) = \hat{\boldsymbol{f}}_c(\boldsymbol{v}^h_{\kappa,e}, \boldsymbol{v}^h_{\kappa',e}; \nu_\kappa)$ denotes the convective numerical flux corresponds to physical convective flux \boldsymbol{f}_c on the interfaces of the elements. This numerical flux is a vector-valued function of two interface states and the interface normal ν_κ on element κ , and is considered to be conservative and consistent with \boldsymbol{f}_c . Also this flux is required to be entropy stable, i.e. it has the following viscosity form as

$$\hat{\boldsymbol{f}}_{c}(\boldsymbol{v}^{h}) = \hat{\boldsymbol{f}}_{c}(\boldsymbol{v}_{\kappa,e}^{h}, \boldsymbol{v}_{\kappa',e}^{h}; \nu_{\kappa}) = \boldsymbol{f}^{\star}(\boldsymbol{v}^{h}) - \frac{1}{2}D(\boldsymbol{v}^{h})[\![\boldsymbol{v}^{h}]\!], \quad \text{on } e \in \mathcal{E}_{h,I}$$
(17)

where $f^{\star}(\boldsymbol{v}^{h}) = f^{\star}(\boldsymbol{v}^{h}_{\kappa,e}, \boldsymbol{v}^{h}_{\kappa',e}; \nu)$ denotes the *entropy conservative flux* and *D* is the numerical diffusion required to obtain entropy stability. This numerical diffusion is set to be a symmetric and uniformly positive definite matrix. The main property of such fluxes is

$$\hat{\boldsymbol{f}}_{c}(\boldsymbol{v}_{\kappa,e}^{h},\boldsymbol{v}_{\kappa',e}^{h};\nu_{\kappa}):[\![\boldsymbol{v}^{h}]\!] \ge C[\![\boldsymbol{v}^{h}]\!]:[\![\boldsymbol{v}^{h}]\!]$$
(18)

where C is some lower bound for the eigenvalues of the uniformly positive definite and symmetric matrix D. Two interesting choices for diffusion matrix D are Roe and Rusanov diffusion. We refer to the literature for more details on the diffusion matrix and properties of the entropy stable fluxes, as well as explicit forms of such fluxes for Euler equations.^{19,25,34} On the boundary faces $\hat{f}_c(\boldsymbol{v}^h)$ is defined as

$$\hat{\boldsymbol{f}}_{c}(\boldsymbol{v}_{\kappa,e}^{h};\boldsymbol{\nu}_{\kappa}) = \boldsymbol{f}_{c}(\boldsymbol{u}_{\Gamma}(\boldsymbol{v}_{\kappa,e}^{h})) \cdot \boldsymbol{\kappa}_{e}, \quad \text{on } e \in \mathcal{E}_{h,\partial}.$$
(19)

Applying integration by parts on (16) leads to the final variational form of the convective part as

$$\mathcal{B}^{c}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = -\sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\boldsymbol{f}_{c}:\nabla\boldsymbol{w}^{h}\,\mathrm{d}x + \sum_{e\in\mathcal{E}_{h,I}}\int_{e}\langle\hat{\boldsymbol{f}}_{c}(\boldsymbol{v}^{h}),\boldsymbol{w}^{h}_{\kappa,e}\rangle\,\mathrm{d}s + \mathcal{B}_{\Gamma}^{c}(\boldsymbol{v}^{h},\boldsymbol{w}^{h})$$
(20)

where the boundary \mathcal{B}_{Γ}^c term is

$$\mathcal{B}_{\Gamma}^{c}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{e \in \mathcal{E}_{h,\partial}} \int_{e} \langle \boldsymbol{f}_{c} \cdot \boldsymbol{\nu}_{\kappa}, \boldsymbol{w}_{\kappa,e}^{h} \rangle \,\mathrm{d}s.$$
(21)

2. Viscous Discretization

For the discretization of the viscous flux $\mathbf{f}_v = \nabla \cdot (\tilde{\mathbf{K}}(v) \nabla v)$ in (9), we follow the formulation presented by Cockburn and Dawson,¹³ later used in other work as well,^{10,30} and consider a first order mixed formulation for three unknown variables; $v, \theta = \nabla v$ and $\boldsymbol{\sigma} = \tilde{\mathbf{K}} \theta$. Then, restricted to local formulations, we present the equivalent primal formulation which is solvable for v^h and can easily fit into the primal formulation (15).

Note that one can rewrite (9) as the following problem

$$\begin{aligned}
-\nabla \cdot \boldsymbol{\sigma} &= R, & x \in \Omega, \\
\boldsymbol{\sigma} &= \tilde{\boldsymbol{K}} \boldsymbol{\theta}, & x \in \Omega, \\
\boldsymbol{\theta} &= \nabla \boldsymbol{v}, & x \in \Omega.
\end{aligned}$$

where $R := -(\boldsymbol{u}_{\boldsymbol{v}}\boldsymbol{v}_t + \tilde{\boldsymbol{A}}_i(\boldsymbol{v})\nabla_i\boldsymbol{v})$, is the sum of time derivative and inviscid convective flux in (1).

Here we approximate the exact solution $(\boldsymbol{\sigma}, \boldsymbol{\theta}, \boldsymbol{u})$ by discrete functions $(\boldsymbol{\sigma}^h, \boldsymbol{\theta}^h, \boldsymbol{u}^h)$ in the finite element space $(\Sigma_{h,p} \times \Sigma_{h,p} \times V_{h,q})$, where $V_{h,q}$ is the same as (14) and $\Sigma_{h,p}$ is defined as

$$\Sigma_{h,p} := \{ \boldsymbol{\theta}^h \in [L_2(\Omega)]^{m \times d} : \boldsymbol{\theta}^h |_{\kappa} \in [\mathcal{P}^p(\kappa)]^{m \times d}, \quad \forall \kappa \in \mathcal{T}_h \},$$

with $q \ge 1$ and p = q or p = q - 1. (To satisfy the property $\nabla V_{h,q} \subset \Sigma_{h,p}$, cf. Ref. 1.) Now we consider the following weak formulation

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \tilde{\boldsymbol{K}} \boldsymbol{\theta}^h : \boldsymbol{\zeta}^h \, \mathrm{d}x = \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \boldsymbol{\sigma}^h : \boldsymbol{\zeta}^h \, \mathrm{d}x, \qquad \qquad \forall \boldsymbol{\zeta}^h \in \Sigma_{h,p}.$$
(22)

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \boldsymbol{\theta}^h : \boldsymbol{\tau}^h \, \mathrm{d}x + \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \langle \boldsymbol{v}^h, \nabla \cdot \boldsymbol{\tau}^h \rangle \, \mathrm{d}x = \sum_{\kappa \in \mathcal{T}_h} \int_{\partial \kappa} \langle \hat{\boldsymbol{v}}, \boldsymbol{\tau}^h \cdot \boldsymbol{\nu}_\kappa \rangle \, \mathrm{d}s, \qquad \forall \boldsymbol{\tau}^h \in \Sigma_{h,p}, \qquad (23)$$

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \boldsymbol{\sigma}^h : \nabla \boldsymbol{w}^h \, \mathrm{d}x - \sum_{\kappa \in \mathcal{T}_h} \int_{\partial \kappa} \langle \hat{\boldsymbol{\sigma}} \cdot \boldsymbol{\nu}_{\kappa}, \boldsymbol{w}^h \rangle \, \mathrm{d}s = (R, \boldsymbol{w}^h), \qquad \forall \boldsymbol{w}^h \in V_{h,q}.$$
(24)

Note that $(R, \boldsymbol{w}^h) = -((\boldsymbol{u}_t, \boldsymbol{w}^h) + \mathcal{B}^c(\boldsymbol{v}^h, \boldsymbol{w}^h)).$

This flux formulation is complete but the definition of the numerical fluxes \hat{v} and $\hat{\sigma}$ which depends on $(\sigma^h, \theta^h, u^h)$, and needs to be designed carefully such that the method has good well-posedness and compatibility properties. We postpone the explicit definition of these numerical fluxes till the end of this section and now we are going to present the equivalent *primal* formulation; i.e., a formulation which has v^h as its only unknowns.

In order to obtain the primal formulation we need to solve the unknowns σ^h and θ^h in terms of v^h . In the first step, using (22) one can solve for σ^h as the Galerkin $[L_2(\Omega)]^{m \times d}$ projection

$$\boldsymbol{\sigma}^{h} = \mathcal{G}_{h} \big(\tilde{\boldsymbol{K}}(\boldsymbol{u}^{h}) \boldsymbol{\theta}^{h} \big), \tag{25}$$

where $\mathcal{G}_h : [L_2(\Omega)]^{m \times d} \to \Sigma_{h,p}$ has the following property; for all $\boldsymbol{\xi} \in [L_2(\Omega)]^{m \times d}$

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \boldsymbol{\xi} : \boldsymbol{\tau} \, \mathrm{d}x = \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \mathcal{G}_h(\boldsymbol{\xi}) : \boldsymbol{\tau} \, \mathrm{d}x, \qquad \forall \boldsymbol{\tau} \in \Sigma_{h,p}.$$
(26)

Moreover, let us remark the following identity; for any $\boldsymbol{v} \in \prod_{\kappa \in \mathcal{T}_h} [L_2(\partial \kappa)]^m$ and $\boldsymbol{\xi} \in \prod_{\kappa \in \mathcal{T}_h} [L_2(\partial \kappa)]^{m \times d}$ the following holds

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\partial \kappa} \langle \boldsymbol{v}, \boldsymbol{\xi} \cdot \boldsymbol{\nu} \rangle \, \mathrm{d}s = \sum_{e \in \mathcal{E}_{h,I}} \int_e \langle \{\boldsymbol{v}\}, \llbracket \boldsymbol{\xi} \rrbracket \rangle \, \mathrm{d}s + \sum_{e \in \mathcal{E}_h} \int_e \llbracket \boldsymbol{v} \rrbracket : \{\boldsymbol{\xi}\} \, \mathrm{d}s.$$
(27)

Applying (27) in (23) and (24), one can write

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \boldsymbol{\theta}^h : \boldsymbol{\tau}^h \, \mathrm{d}x - \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \nabla \boldsymbol{v}^h : \boldsymbol{\tau}^h \, \mathrm{d}x + \sum_{e \in \mathcal{E}_{h,I}} \int_{e} \langle \{\boldsymbol{v}^h - \hat{\boldsymbol{v}}\}, \llbracket \boldsymbol{\tau}^h \rrbracket \rangle \, \mathrm{d}s + \sum_{e \in \mathcal{E}_h} \int_{e} \llbracket \boldsymbol{v}^h - \hat{\boldsymbol{v}} \rrbracket : \{\boldsymbol{\tau}^h\} \, \mathrm{d}s = 0,$$
$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \boldsymbol{\sigma}^h : \nabla \boldsymbol{w}^h \, \mathrm{d}x - \sum_{e \in \mathcal{E}_h} \int_{e} \{\hat{\boldsymbol{\sigma}}\} : \llbracket \boldsymbol{w}^h \rrbracket \, \mathrm{d}s - \sum_{e \in \mathcal{E}_{h,I}} \int_{e} \langle \llbracket \hat{\boldsymbol{\sigma}} \rrbracket, \{\boldsymbol{w}^h\} \rangle \, \mathrm{d}s = (R, \boldsymbol{w}^h).$$

Moreover, let us define two lifting operator $r: [L_2(\mathcal{E}_h)]^{m \times d} \to \Sigma_{h,p}$ and $l: [L_2(\mathcal{E}_{h,I})]^m \to \Sigma_{h,p}$ as,

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} r(\varphi) : \boldsymbol{\tau} \, \mathrm{d}x = -\sum_{e \in \mathcal{E}_h} \int_{e} \varphi : \{\boldsymbol{\tau}\} \, \mathrm{d}s, \qquad \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} l(\varphi) : \boldsymbol{\tau} \, \mathrm{d}x = -\sum_{e \in \mathcal{E}_{h,I}} \int_{e} \langle \varphi, [\![\boldsymbol{\tau}]\!] \rangle \, \mathrm{d}s, \qquad (28)$$

for all $\tau \in \Sigma_{h,p}$. Using the Riesz representation theorem one can prove the existence and uniqueness of the lifting operators introduced in (28) (see e.g., [10, Lemma 3.3]). Also we define edge-wise right lifting operator as $r^e : [L_2(e)]^{m \times d} \to \Sigma_{h,p}$

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} r^e(\varphi) : \boldsymbol{\tau} \, \mathrm{d}x = -\int_e \varphi : \{\boldsymbol{\tau}\} \, \mathrm{d}s, \qquad \forall \boldsymbol{\tau} \in \Sigma_{h,p}$$
(29)

for all edges $e \in \mathcal{E}_h$. Also by noting that $r(\varphi) = \sum_{e \in \mathcal{E}_h} r^e(\varphi)$, and by applying Cauchy-Shwarz inequality and taking the norm over \mathcal{T}_h , one has

$$\|r(\varphi)\|_{L_{2}(\Omega)}^{2} = \|\sum_{e \in \mathcal{E}_{h}} r^{e}(\varphi)\|_{L_{2}(\Omega)}^{2} \le N_{l} \sum_{e \in \mathcal{E}_{h}} \|r^{e}(\varphi)\|_{L_{2}(\Omega)}^{2}.$$
(30)

Also following³⁰ and¹⁰ we define a similar lifting operator on the boundary as

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} r_{\Gamma} : \boldsymbol{\tau}^h \, \mathrm{d}x = -\sum_{e \in \mathcal{E}_{h,\partial}} \int_{e} \langle \boldsymbol{v}_{\Gamma}, \boldsymbol{\tau}^h \cdot \boldsymbol{\nu}_{\kappa} \rangle \, \mathrm{d}s, \qquad \forall \boldsymbol{\tau}^h \in \Sigma_{h,p}$$
(31)

Now let us set additional properties for the numerical flux \hat{v} and $\hat{\sigma}$ and require their conservation property. This property as well as the property of being single valued on the interface $e \in \mathcal{E}_{h,I}$ leads to

$$\llbracket \hat{\boldsymbol{v}} \rrbracket = 0, \qquad \{ \hat{\boldsymbol{v}} \} = \hat{\boldsymbol{v}}, \qquad \llbracket \hat{\boldsymbol{\sigma}} \rrbracket = 0, \qquad \{ \hat{\boldsymbol{\sigma}} \} = \hat{\boldsymbol{\sigma}}. \tag{32}$$

Using (32) and the Dirichlet boundary conditions, one can simplify the weak formulation as

$$\sum_{\kappa \in \mathcal{T}_{h}} \int_{\kappa} \boldsymbol{\theta}^{h} : \boldsymbol{\tau}^{h} \, \mathrm{d}\boldsymbol{x} = \sum_{\kappa \in \mathcal{T}_{h}} \int_{\kappa} \nabla \boldsymbol{v}^{h} : \boldsymbol{\tau}^{h} \, \mathrm{d}\boldsymbol{x} + \sum_{e \in \mathcal{E}_{h,I}} \int_{e} \langle \{\boldsymbol{v}^{h}\} - \hat{\boldsymbol{v}}, [\![\boldsymbol{\tau}^{h}]\!] \rangle \, \mathrm{d}\boldsymbol{s} + \sum_{e \in \mathcal{E}_{h}} \int_{e} [\![\boldsymbol{v}^{h}]\!] : \{\boldsymbol{\tau}^{h}\} \, \mathrm{d}\boldsymbol{s} \qquad (33)$$
$$+ \sum_{e \in \mathcal{E}_{h,\partial}} \int_{e} \langle \hat{\boldsymbol{v}}, \boldsymbol{\tau}^{h} \cdot \boldsymbol{\nu}_{\kappa} \rangle \, \mathrm{d}\boldsymbol{s} = 0,$$
$$\sum_{\kappa \in \mathcal{T}_{h}} \int_{\kappa} \boldsymbol{\sigma}^{h} : \nabla \boldsymbol{w}^{h} \, \mathrm{d}\boldsymbol{x} - \sum_{e \in \mathcal{E}_{h}} \int_{e} \hat{\boldsymbol{\sigma}} : [\![\boldsymbol{w}^{h}]\!] \, \mathrm{d}\boldsymbol{s} = (R, \boldsymbol{w}^{h}). \tag{34}$$

Using the definition of lifting operator (28) and (25), we arrive at

$$\boldsymbol{\theta}^{h} = \nabla \boldsymbol{v}^{h} + r(\llbracket \boldsymbol{v}^{h} \rrbracket) + l(\{\boldsymbol{v}^{h}\} - \hat{\boldsymbol{v}}) + r_{\Gamma}$$
(35)

$$\boldsymbol{\sigma}^{h} = \mathcal{G}_{h} \Big(\tilde{\boldsymbol{K}}(\boldsymbol{v}^{h}) \big(\nabla \boldsymbol{v}^{h} + r(\llbracket \boldsymbol{v}^{h} \rrbracket) + l(\lbrace \boldsymbol{v}^{h} \rbrace - \hat{\boldsymbol{v}}) + r_{\Gamma} \big) \Big).$$
(36)

Now $(\boldsymbol{\theta}^h, \boldsymbol{\sigma}^h)$ can be solved locally in terms of \boldsymbol{v}^h by inserting (35) and (36) in (34). Then, one can obtain the primal formulation as the following; find $\boldsymbol{v}^h \in V_{h,q}$ such that

$$\mathcal{B}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) - (R,\boldsymbol{w}^{h}) = (\boldsymbol{u}_{t}(\boldsymbol{v}^{h}),\boldsymbol{w}^{h}) + \mathcal{B}^{c}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) + \mathcal{B}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = 0, \quad \forall \boldsymbol{w}^{h} \in V_{h,q},$$
(37)

where

$$\mathcal{B}^{\boldsymbol{v}}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}(\nabla\boldsymbol{v}^{h} + r(\llbracket\boldsymbol{v}^{h}\rrbracket) + l(\{\boldsymbol{v}^{h}\} - \hat{\boldsymbol{v}}) + r_{\Gamma}): \nabla\boldsymbol{w}^{h}\,\mathrm{d}x - \sum_{e\in\mathcal{E}_{h}}\int_{e}\hat{\boldsymbol{\sigma}}:\llbracket\boldsymbol{w}^{h}\rrbracket\,\mathrm{d}s.$$
(38)

which still needs to become fully defined by setting \hat{v} and $\hat{\sigma}$. In the following, inspired by some well-known schemes, we examine some choices for these two fluxes:

(i) BR1: This formulation is defined by setting⁶

$$\hat{\boldsymbol{v}} = egin{cases} \{ \boldsymbol{v}^h \} & ext{on } \mathcal{E}_{h,I} \ \boldsymbol{v}_{\Gamma} & ext{on } \mathcal{E}_{h,\partial} \ \end{pmatrix}, \qquad \qquad \hat{\boldsymbol{\sigma}} = \{ \boldsymbol{\sigma}^h \} & ext{on } \mathcal{E}_h \end{cases}$$

which leads to the following primal formulation

$$\mathcal{B}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}}(\nabla\boldsymbol{v}^{h} + r(\llbracket\boldsymbol{v}^{h}\rrbracket)) : (\nabla\boldsymbol{w}^{h} + r(\llbracket\boldsymbol{w}^{h}\rrbracket)) \,\mathrm{d}x + \mathcal{B}_{\Gamma}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}).$$
(39)

with the boundary related terms as

$$\mathcal{B}_{\Gamma}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}}r_{\Gamma} : (\nabla \boldsymbol{w}^{h} + r(\llbracket \boldsymbol{w}^{h} \rrbracket)) \,\mathrm{d}x.$$
(40)

(ii) BR2: Here \hat{v} and $\hat{\sigma}$ are set as⁸

$$\hat{\boldsymbol{v}} = \begin{cases} \{\boldsymbol{v}^h\} & \text{on } \mathcal{E}_{h,I} \\ \boldsymbol{v}_{\Gamma} & \text{on } \mathcal{E}_{h,\partial} \end{cases}, \quad \hat{\boldsymbol{\sigma}} = \begin{cases} \{\mathcal{G}_h(\tilde{\boldsymbol{K}}(\boldsymbol{v}^h)\nabla\boldsymbol{v}^h)\} - \alpha_{r,\tilde{\boldsymbol{K}}}(\llbracket \boldsymbol{v}^h \rrbracket) & \text{on } \mathcal{E}_{h,I} \\ \mathcal{G}_h(\tilde{\boldsymbol{K}}(\boldsymbol{v}_{\Gamma})\nabla\boldsymbol{v}^h) - \alpha_{r,\tilde{\boldsymbol{K}}}^{\partial}((\boldsymbol{v}_{\kappa,e}^h - \boldsymbol{v}_{\Gamma})\otimes\nu_{\kappa}) & \text{on } \mathcal{E}_{h,\partial} \end{cases}$$

with the notation $\alpha_{r, \tilde{K}}$ defined as (noting $\boldsymbol{v}_{\kappa, e}^{h} \otimes \nu_{\kappa} = \llbracket \boldsymbol{v}^{h} \rrbracket$ on $e \in \mathcal{E}_{h, \partial}$)

$$\sum_{e \in \mathcal{E}_h} \int_e \alpha_{r, \tilde{\boldsymbol{K}}}(\llbracket \boldsymbol{v}^h \rrbracket) : \llbracket \boldsymbol{w}^h \rrbracket \, \mathrm{d}s = \sum_{e \in \mathcal{E}_h} \eta_e \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \mathcal{G}_h \left(\tilde{\boldsymbol{K}}(\boldsymbol{v}^h) r^e(\llbracket \boldsymbol{v}^h \rrbracket) \right) : r^e(\llbracket \boldsymbol{w}^h \rrbracket) \, \mathrm{d}x,$$

for any $e \in \mathcal{E}_h$. Also the notation $\alpha_{r, \tilde{K}}^{\Gamma}$ is defined as

$$\sum_{e \in \mathcal{E}_{h,\partial}} \int_{e} \alpha_{r,\tilde{K}}^{\partial} \left((\boldsymbol{v}_{\kappa,e}^{h} - \boldsymbol{v}_{\Gamma}) \otimes \nu_{\kappa} \right) : \llbracket \boldsymbol{w}^{h} \rrbracket \, \mathrm{d}s = \sum_{e \in \mathcal{E}_{h,\partial}} \eta_{e} \sum_{\kappa \in \mathcal{T}_{h}} \int_{\kappa} \mathcal{G}_{h} \left(\tilde{\boldsymbol{K}}(\boldsymbol{v}_{\Gamma}) r^{e} \left((\boldsymbol{v}_{\kappa,e}^{h} - \boldsymbol{v}_{\Gamma}) \otimes \nu_{\kappa} \right) \right) : r^{e} \left(\llbracket \boldsymbol{w}^{h} \rrbracket \right) \, \mathrm{d}x,$$

for any $e \in \mathcal{E}_{h,\partial}$. Using the properties of Galerkin projection \mathcal{G}_h and the definition of the lifting operator (28), the primal formulation of BR2 yields as

$$\mathcal{B}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}\nabla\boldsymbol{v}^{h}:\nabla\boldsymbol{w}^{h}\,\mathrm{d}x + \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\left(\tilde{\boldsymbol{K}}r(\llbracket\boldsymbol{v}^{h}\rrbracket):\nabla\boldsymbol{w}^{h}+\tilde{\boldsymbol{K}}\nabla\boldsymbol{v}^{h}:r(\llbracket\boldsymbol{w}^{h}\rrbracket)\right)\,\mathrm{d}x + \sum_{e\in\mathcal{E}_{h}}\eta_{e}\sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}r^{e}(\llbracket\boldsymbol{v}^{h}\rrbracket):r^{e}(\llbracket\boldsymbol{w}^{h}\rrbracket)\,\mathrm{d}x + \mathcal{B}_{\Gamma}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}).$$
(41)

American Institute of Aeronautics and Astronautics

Here the parameter η_e depends on the properties of the triangulation \mathcal{T}_h and should be chosen such that $\eta_e \geq N_l(\mathcal{T}_h)$. We will discuss the rationale behind this choice more detailed in the stability analysis section IV. Also the additional term for the boundary is

$$\begin{aligned} \mathcal{B}^{v}_{\Gamma}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) &= \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}}(\boldsymbol{v}^{h})r_{\Gamma} : \nabla \boldsymbol{w}^{h} \, \mathrm{d}x + \sum_{e\in\mathcal{E}_{h,\partial}} \int_{e} \mathcal{G}_{h} \Big(\tilde{\boldsymbol{K}}(\boldsymbol{v}^{h}_{\kappa,e}) \nabla \boldsymbol{v}^{h} - \tilde{\boldsymbol{K}}(\boldsymbol{v}_{\Gamma}) \nabla \boldsymbol{v}^{h} \Big) : (\boldsymbol{w}^{h} \otimes \boldsymbol{\nu}_{\kappa}) \, \mathrm{d}s \\ &- \sum_{e\in\mathcal{E}_{h,\partial}} \eta_{e} \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \mathcal{G}_{h} \Big(\tilde{\boldsymbol{K}}(\boldsymbol{v}^{h}) r^{e}(\llbracket \boldsymbol{v}^{h} \rrbracket) - \tilde{\boldsymbol{K}}(\boldsymbol{v}_{\Gamma}) r^{e}(\llbracket \boldsymbol{v}^{h} \rrbracket) \Big) : r^{e}(\llbracket \boldsymbol{w}^{h} \rrbracket) \, \mathrm{d}x \\ &- \sum_{e\in\mathcal{E}_{h,\partial}} \eta_{e} \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \mathcal{G}_{h} \Big(\tilde{\boldsymbol{K}}(\boldsymbol{v}_{\Gamma}) r^{e}(\boldsymbol{v}_{\Gamma} \otimes \boldsymbol{\nu}_{\kappa}) \Big) : r^{e}(\llbracket \boldsymbol{w}^{h} \rrbracket) \, \mathrm{d}x. \end{aligned}$$

(iii) SIPG: In this formulation we set

$$\hat{\boldsymbol{v}} = \begin{cases} \{\boldsymbol{v}^h\} & \text{on } \mathcal{E}_{h,I} \\ \boldsymbol{v}_{\Gamma} & \text{on } \mathcal{E}_{h,\partial} \end{cases}, \quad \hat{\boldsymbol{\sigma}} = \begin{cases} \{\mathcal{G}_h(\tilde{\boldsymbol{K}}(\boldsymbol{v}^h)\nabla\boldsymbol{v}^h)\} - \frac{\mu_e}{h_e}[\![\boldsymbol{v}^h]\!] & \text{on } \mathcal{E}_{h,I}, \\ \mathcal{G}_h\big(\tilde{\boldsymbol{K}}(\boldsymbol{v}_{\Gamma})\nabla\boldsymbol{v}^h)\big) - \frac{\mu_e}{h_e}(\boldsymbol{v}^h_{\kappa,e} - \boldsymbol{v}_{\Gamma}) \otimes \nu_{\kappa} & \text{on } \mathcal{E}_{h,\partial} \end{cases}$$

with some $\mu_e > 0$ as a *penalty* parameter. Similar to³² or,¹² this parameter is dependent on the properties of the triangulation, polynomial order and the diffusion matrix \tilde{K} as

$$\mu_e = C_p(\mathcal{T}_h, \tilde{K})q^2.$$
(42)

We will discuss about the value of C_p later in section IV. Similar to BR2, the primal formulation reads as

$$\mathcal{B}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}\nabla\boldsymbol{v}^{h}:\nabla\boldsymbol{w}^{h}\,\mathrm{d}x + \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\left(\tilde{\boldsymbol{K}}r(\llbracket\boldsymbol{v}^{h}\rrbracket):\nabla\boldsymbol{w}^{h}+\tilde{\boldsymbol{K}}\nabla\boldsymbol{v}^{h}:r(\llbracket\boldsymbol{w}^{h}\rrbracket)\right)\,\mathrm{d}x + \sum_{e\in\mathcal{E}_{h}}\frac{\mu_{e}}{h_{e}}\int_{e}\llbracket\boldsymbol{v}^{h}\rrbracket:\llbracket\boldsymbol{w}^{h}\rrbracket\,\mathrm{d}s + \mathcal{B}_{\Gamma}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}).$$
(43)

where

$$\begin{aligned} \mathcal{B}^{\boldsymbol{v}}_{\Gamma}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) &= \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}(\boldsymbol{v}^{h})r_{\Gamma}:\nabla\boldsymbol{w}^{h}\,\mathrm{d}x + \sum_{e\in\mathcal{E}_{h,\partial}}\int_{e}\mathcal{G}_{h}\Big(\tilde{\boldsymbol{K}}(\boldsymbol{v}^{h}_{\kappa,e})\nabla\boldsymbol{v}^{h} - \tilde{\boldsymbol{K}}(\boldsymbol{v}_{\Gamma})\nabla\boldsymbol{v}^{h}\Big):(\boldsymbol{w}^{h}\otimes\nu_{\kappa})\,\mathrm{d}s\\ &- \sum_{e\in\mathcal{E}_{h,\partial}}\frac{\mu_{e}}{h_{e}}\int_{e}\langle\boldsymbol{v}_{\Gamma},\boldsymbol{w}^{h}\rangle\,\mathrm{d}x\end{aligned}$$

(iv) LDG: In the LDG formulation¹⁴ we set

$$\hat{\boldsymbol{v}} = \begin{cases} \{\boldsymbol{v}^h\} - \beta \cdot \llbracket \boldsymbol{v}^h \rrbracket & \text{on } \mathcal{E}_{h,I} \\ \boldsymbol{v}_{\Gamma} & \text{on } \mathcal{E}_{h,\partial} \end{cases}, \quad \hat{\boldsymbol{\sigma}} = \begin{cases} \{\boldsymbol{\sigma}^h\} + \beta \otimes \llbracket \boldsymbol{\sigma}^h \rrbracket - \alpha \llbracket \boldsymbol{v}^h \rrbracket & \text{on } \mathcal{E}_{h,I} \\ \boldsymbol{\sigma}^h - \alpha (\boldsymbol{v}^h_{\kappa,e} - \boldsymbol{v}_{\Gamma}) \otimes \nu_{\kappa} & \text{on } \mathcal{E}_{h,\partial} \end{cases}$$

where $\beta \in \mathbb{R}^d$ is some mesh-dependent parameter and constant on each edge. Also $\alpha > 0$ is some parameter which is required for obtaining stability. By some lines of straightforward computations the following primal formulation can be obtained (also see³⁰ and¹⁰)

$$\mathcal{B}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}}(\boldsymbol{v}^{h}) \left(\nabla \boldsymbol{v}^{h} + r(\llbracket \boldsymbol{v}^{h} \rrbracket) + l(\beta \cdot \llbracket \boldsymbol{v}^{h} \rrbracket) \right) : \left(\nabla \boldsymbol{w}^{h} + r(\llbracket \boldsymbol{w}^{h} \rrbracket) + l(\beta \cdot \llbracket \boldsymbol{w}^{h} \rrbracket) \right) dx + \sum_{e\in\mathcal{E}_{h}} \int_{e} \alpha \llbracket \boldsymbol{v}^{h} \rrbracket : \llbracket \boldsymbol{w}^{h} \rrbracket ds + \mathcal{B}_{\Gamma}^{v}(\boldsymbol{v}^{h}, \boldsymbol{w}^{h})$$
(44)

where

$$\mathcal{B}_{\Gamma}^{v}(\boldsymbol{v}^{h},\boldsymbol{w}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}r_{\Gamma}:\left(\nabla\boldsymbol{w}^{h} + r(\llbracket\boldsymbol{w}^{h}\rrbracket) + l(\beta\cdot\llbracket\boldsymbol{w}^{h}\rrbracket)\right)\mathrm{d}x - \sum_{e\in\mathcal{E}_{h,\partial}}\alpha\int_{e}\langle\boldsymbol{v}_{\Gamma},\boldsymbol{w}^{h}\rangle\,\mathrm{d}s$$

$$\langle \boldsymbol{u}_t(\boldsymbol{v}^h), \boldsymbol{w}^h \rangle + \mathcal{B}(\boldsymbol{v}^h, \boldsymbol{w}^h) = \langle \boldsymbol{u}_t, \boldsymbol{v}^h \rangle + \mathcal{B}_I^c(\boldsymbol{v}^h, \boldsymbol{w}^h) + \mathcal{B}_I^v(\boldsymbol{v}^h, \boldsymbol{w}^h) + \mathcal{B}_{\Gamma}^c(\boldsymbol{v}^h, \boldsymbol{w}^h) + \mathcal{B}_{\Gamma}^v(\boldsymbol{v}^h, \boldsymbol{w}^h) = 0, \quad (45)$$

where subscript Γ represents the parts of $\mathcal{B}(\boldsymbol{v}^h, \boldsymbol{w}^h)$ (15) which includes boundary conditions and subscript I denotes the rest.

We want to remark that due to the discrete nature of the Galerkin projection (26), the viscous formulations presented here are inconsistent with the exact smooth solution of the problem and consequently adjoint inconsistent as well. However, one might show that in asymptotic limit the consistency and adjoint consistency can be recovered. For more details we refer to²⁹ and the references cited there, as well as our in preparation work.³⁷

IV. Entropy Stability Result

In this section we consider the stability result of the approximate solution of scheme (15). The symmetric formulation we obtained through the change of variables simplifies this analysis, for both the infinitedimensional weak solution and its finite-dimensional numerical approximation counterpart. Taking inner products of the (9) with respect to the entropy variables and using the definition of the entropy flux yields

$$\int_{\Omega} U_t(\boldsymbol{v}) \, \mathrm{d}x + \int_{\Omega} \nabla \cdot Q(\boldsymbol{v}) \, \mathrm{d}x + \int_{\Omega} \tilde{\boldsymbol{K}}(\boldsymbol{v}) \nabla \boldsymbol{v} : \nabla \boldsymbol{v} \, \mathrm{d}x = 0.$$
(46)

From the positive semi-definiteness of matrix \bar{K} one can conclude the non-negativeness of the third term on the left. Moreover if by applying *appropriate* boundary conditions and using the divergence theorem, the second term is non-negative, the following *global entropy inequality* is obtained

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega} U(\boldsymbol{u}) \,\mathrm{d}x \le 0. \tag{47}$$

This property can be viewed as an extension of L_2 stability for systems of conservation laws and is the motivation behind *entropy stable schemes*, which were originally introduced by Tadmor.³³ For more details on these kind of schemes we refer to the seminal review paper.³⁴

From the numerical point of view it is also desirable to retain the entropy stability property for the approximate solution. We are going to present the following theorem on the entropy stability property on the numerical solution of the scheme (15):

Theorem 1. Let us consider v^h as the approximate solution of (1) produced by scheme (15). Also let us assume the following holds for the symmetric diffusion matrix \tilde{K} ; there exists $\Lambda < \infty$ such that for any $w^h \neq 0$

$$0 \le \langle \boldsymbol{w}^h, \tilde{\boldsymbol{K}} \boldsymbol{w}^h \rangle \le \Lambda \langle \boldsymbol{w}^h, \boldsymbol{w}^h \rangle.$$
(48)

Then the following holds

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\Omega} U(\boldsymbol{v}^h) \,\mathrm{d}x \le 0,\tag{49}$$

where U is defined as (8); i.e. the method (15) is entropy stable in semi-discrete form.

For simplicity we are going to neglect the contribution of boundary conditions appearing in form of \mathcal{B}_{Γ}^{c} and \mathcal{B}_{Γ}^{v} . It's plausible that by choosing the right stabilization parameters as well as 'stable' boundary conditions, one is able to retrieve the entropy stability for the full scheme. We shall not address this problem in the present paper and merely concentrate on the main idea of discretization.

Also, before presenting the proof we mention the following lemma which will be used later:

Lemma 1. There exist two positive constants $C_r, C_R > 0$, such that for all $e \in \mathcal{E}_h$ we have

$$C_r h_e^{-1/2} \| \llbracket \boldsymbol{w} \rrbracket \|_{L_2(e)} \le \| r^e(\llbracket \boldsymbol{w} \rrbracket) \|_{L_2(\Omega)} \le C_R h_e^{-1/2} \| \llbracket \boldsymbol{w} \rrbracket \|_{L_2(e)}, \qquad \forall \boldsymbol{w} \in V_{h,q},$$
(50)

where the constants are h independent and only depend on the minimum angle of the triangles and the polynomial degree q.

The proof of this lemma can be found, e.g., in [9, Lemma 2] and we skip it. Also note that the validity of the assumption (48) lies in our symmetric form.²⁴ Now we are ready to present the proof of Theorem 1:

Proof. The proof can be obtained from the convective discretization properties of entropy stable flux (17), and the viscous discretization proposed in section 2. Inserting $w^h = v^h$ in (15) and using the definition of entropy variables yields

$$\frac{\mathrm{d}}{\mathrm{d}t} \Big(\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} U(\boldsymbol{v}^h) \,\mathrm{d}x \Big) + \mathcal{B}^c(\boldsymbol{v}^h, \boldsymbol{v}^h) + \mathcal{B}^v(\boldsymbol{v}^h, \boldsymbol{v}^h) = 0.$$
(51)

Now we deal with second and third terms on the left hand side of (51) as below:

(i) Along the same lines as in related work^{5, 23, 39} which uses the entropy dissipation structure D and the properties of entropy conservative flux as (18), one can show

$$\mathcal{B}_{I}^{c}(\boldsymbol{v}^{h},\boldsymbol{v}^{h}) = -\sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\sum_{k=1}^{d}\langle\boldsymbol{f}^{k}(\boldsymbol{v}^{h}),\boldsymbol{w}_{x_{k}}^{h}\rangle\,\mathrm{d}x + \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\langle\hat{\boldsymbol{f}}_{c}(\boldsymbol{v}^{h}),\boldsymbol{w}_{\kappa,e}^{h}\rangle\,\mathrm{d}s \geq C\sum_{\kappa\in\mathcal{T}_{h}}\int_{\partial\kappa}|[\![\boldsymbol{v}^{h}]\!]|^{2}\,\mathrm{d}s.$$

- (ii) To show the non-negativity of $\mathcal{B}^{v}(\boldsymbol{v}^{h},\boldsymbol{v}^{h})$ for different viscous discretization we consider different cases:
 - (a) BR1: Using the primal formulation (39) and (48) one has

$$\mathcal{B}_{I}^{v}(\boldsymbol{v}^{h},\boldsymbol{v}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}} \left(\nabla \boldsymbol{v}^{h} + r(\llbracket \boldsymbol{v}^{h} \rrbracket) \right) : \left(\nabla \boldsymbol{v}^{h} + r(\llbracket \boldsymbol{v}^{h} \rrbracket) \right) \, \mathrm{d}x \ge 0.$$
(52)

(b) BR2: From (41) one has

$$\mathcal{B}_{I}^{v}(\boldsymbol{v}^{h},\boldsymbol{v}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}} \nabla \boldsymbol{v}^{h} : \nabla \boldsymbol{v}^{h} \, \mathrm{d}x + 2 \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}} r(\llbracket \boldsymbol{v}^{h} \rrbracket) : \nabla \boldsymbol{v}^{h} \, \mathrm{d}x + \sum_{e\in\mathcal{E}_{h}} \eta_{e} \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}} r^{e}(\llbracket \boldsymbol{v}^{h} \rrbracket) : r^{e}(\llbracket \boldsymbol{v}^{h} \rrbracket) \, \mathrm{d}x$$
(53)

Employing Young's inequality for symmetric positive semi-definite matrix (eg. see [27, Lemma 4.4.]), and by setting $0 < \delta < 1$ we can write

$$2\sum_{\kappa\in\mathcal{T}_h}\int_{\kappa}\tilde{\boldsymbol{K}}r(\llbracket\boldsymbol{v}^h\rrbracket):\nabla\boldsymbol{v}^h\,\mathrm{d}x\leq\delta\sum_{\kappa\in\mathcal{T}_h}\int_{\kappa}\tilde{\boldsymbol{K}}\nabla\boldsymbol{v}^h:\nabla\boldsymbol{v}^h\,\mathrm{d}x+\frac{1}{\delta}\sum_{\kappa\in\mathcal{T}_h}\int_{\kappa}\tilde{\boldsymbol{K}}r(\llbracket\boldsymbol{v}^h\rrbracket):r(\llbracket\boldsymbol{v}^h\rrbracket)\,\mathrm{d}x.$$

which gives

$$\mathcal{B}_{I}^{v}(\boldsymbol{v}^{h},\boldsymbol{v}^{h}) \geq (1-\delta) \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}} \nabla \boldsymbol{v}^{h} : \nabla \boldsymbol{v}^{h} \, \mathrm{d}x - \frac{1}{\delta} \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}} r(\llbracket \boldsymbol{v}^{h} \rrbracket) : r(\llbracket \boldsymbol{v}^{h} \rrbracket) \, \mathrm{d}x + \sum_{e\in\mathcal{E}_{h}} \eta_{e} \sum_{\kappa\in\mathcal{T}_{h}} \int_{\kappa} \tilde{\boldsymbol{K}} r^{e}(\llbracket \boldsymbol{v}^{h} \rrbracket) : r^{e}(\llbracket \boldsymbol{v}^{h} \rrbracket) \, \mathrm{d}x$$
(54)

On the other hand, in the same fashion as^1 and (30), one can show

$$\sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \tilde{\boldsymbol{K}} r(\llbracket \boldsymbol{v}^h \rrbracket) : r(\llbracket \boldsymbol{v}^h \rrbracket) \, \mathrm{d}x \le N_l(\mathcal{T}_h) \sum_{e \in \mathcal{E}_h} \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} \tilde{\boldsymbol{K}} r^e(\llbracket \boldsymbol{v}^h \rrbracket) \cdot r^e(\llbracket \boldsymbol{v}^h \rrbracket) \, \mathrm{d}x.$$
(55)

Comparing (54) and (55) one can easily see that $\mathcal{B}^{v}(\boldsymbol{v}^{h}, \boldsymbol{v}^{h}) \geq 0$ is guaranteed if $\eta_{e} > N_{l}$ for all $e \in \mathcal{E}_{h}$.

(c) SIPG: From the SIPG primal formulation (43) we have

$$\mathcal{B}_{I}^{v}(\boldsymbol{v}^{h},\boldsymbol{v}^{h}) = \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}\nabla\boldsymbol{v}^{h}:\nabla\boldsymbol{v}^{h}\,\mathrm{d}x + 2\sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}r(\llbracket\boldsymbol{v}^{h}\rrbracket):\nabla\boldsymbol{v}^{h}\,\mathrm{d}x + \sum_{e\in\mathcal{E}_{h}}\int_{e}\frac{\mu_{e}}{h_{e}}\llbracket\boldsymbol{v}^{h}\rrbracket:\llbracket\boldsymbol{v}^{h}\rrbracket\,\mathrm{d}s.$$

The steps of the proof follow the same lines as for BR2 above. The main difference is the replacement of the relation (55) by using the inequality (50), which gives

$$\sum_{e \in \mathcal{E}_h} \int_e \frac{\mu_e}{h_e} \llbracket \boldsymbol{v}^h \rrbracket : \llbracket \boldsymbol{v}^h \rrbracket \, \mathrm{d}s \ge \sum_{e \in \mathcal{E}_h} \frac{\mu_e}{C_R^2} \sum_{\kappa \in \mathcal{T}_h} \int_{\kappa} r^e(\llbracket \boldsymbol{v}^h \rrbracket) : r^e(\llbracket \boldsymbol{v}^h \rrbracket) \, \mathrm{d}x.$$
(56)

Comparing to the argument for BR2 method and by (48), one can obtain the following condition on the penalty parameter μ_e

$$\mu_e \ge N_l \Lambda(\vec{K}) C_R^2. \tag{57}$$

in order to satisfy $\mathcal{B}^{v}(\boldsymbol{v}^{h}, \boldsymbol{v}^{h}) \geq 0$.

Let us remark that the penalty parameter μ_e in (57) is independent of the condition number of \tilde{K} which would be infinity in case of Navier-Stokes equations. This is in contrast with the analysis e.g. provided in.²⁷ Actually in²⁷ the specific structure of the \tilde{K} has been exploited to avoid this technical problem. This has been only done in one dimensional case and the extension to higher dimensions is not obvious.

(d) LDG: From the primal formulation of LDG (44) we have

$$\begin{split} \mathcal{B}^{\boldsymbol{v}}(\boldsymbol{v}^{h},\boldsymbol{v}^{h}) &= \sum_{\kappa\in\mathcal{T}_{h}}\int_{\kappa}\tilde{\boldsymbol{K}}\big(\nabla\boldsymbol{v}^{h} + r(\llbracket\boldsymbol{v}^{h}\rrbracket) + l(\beta\cdot\llbracket\boldsymbol{v}^{h}\rrbracket)\big) : \big(\nabla\boldsymbol{v}^{h} + r(\llbracket\boldsymbol{v}^{h}\rrbracket) + l(\beta\cdot\llbracket\boldsymbol{v}^{h}\rrbracket)\big) \,\mathrm{d}x \\ &+ \sum_{e\in\mathcal{E}_{h}}\int_{e}\alpha\llbracket\boldsymbol{v}^{h}\rrbracket : \llbracket\boldsymbol{v}^{h}\rrbracket. \end{split}$$

Using (48), $\mathcal{B}^{v}(\boldsymbol{v}^{h}, \boldsymbol{v}^{h})$ for the LDG formulation is obviously non-negative by setting $\alpha > 0$.

Now, the proof of the theorem follows directly of parts (i)-(ii) and (51).

V. Numerical Results

In this section we provide some numerical results to test and validate the code as well to investigate the applicability of the method on more realistic problems like flow around an airfoil. In order to avoid the additional technicalities from the time discretization which we have not considered in the stability analysis in section IV we consider only steady problems in this section. Moreover, in the numerical result section we provide the results only for both BR2 and SIPG formulations and we do not test LDG scheme. First in section A we look into the advection-diffusion problem in the scalar setting to assess the order of accuracy of the method. Afterwards in section B we apply our formulation to Navier-Stokes equations to present its performance.

The Netgen/Ngsolve library³¹ has been used for geometry handling and mesh generation as well as quadrature rules and the evaluation of basis functions. The nonlinear system obtained from the scheme is solved using a damped Newton method utilizing the ILU preconditioned GMRES available through the PETSc library.² Also note that the additional Galekin projection appeared in $\hat{\sigma}$ has not been differentiated exactly, hence we have non-exact Jacobian in the Newton solver. We provide the convergence history of the residual to show that the effect (at least in our test cases) is negligible.

A. Scalar Advection-Diffusion

Here we are going to apply the formulation presented in (15) to the following steady linear advectionnonlinear diffusion problem on $\Omega = [0, 1]^2$,

$$\nabla \cdot \left(\boldsymbol{f}_{c}(u) - \boldsymbol{f}_{v}(u, \nabla u) \right) = s, \qquad \text{in } \Omega \qquad (58a)$$
$$u = 0, \qquad \text{on } \partial\Omega \qquad (58b)$$

where $\mathbf{f}_c(u) = \mathbf{c}u$ and $\mathbf{f}_v(u, \nabla u) = \epsilon(1+u)\nabla u$. Here $\epsilon > 0$ is some constant and $\mathbf{c} = (1, 1)^T$ is the specified velocity field. Moreover we set the source term s on the right hand such that the exact solution of the problem is

$$u(x,y) = \frac{1}{2}\sin(2\pi x)\sin(2\pi y).$$
(59)

The convective flux set for this case is a Lax-Friedrich flux combined with either BR2 or SIPG for discretizing the viscous flux. Moreover the stabilization parameter for SIPG and BR2 have been set $C_P = 10$ and $\eta_e = 4$ respectively. The results provided in Tables 1 and 2 shows that the scheme achieved the optimal q + 1 order of accuracy in the asymptotic limit for both both methods .

Ν	q	$\ e\ _{L_2}$	order	Ν	q	$\ e\ _{L_2}$	order
6	1	2.07e-01		6	2	9.23e-02	
24		4.69e-02	2.139	24		3.22e-02	1.518
96		3.40e-02	0.4638	96		2.95e-03	3.448
384		8.99e-03	1.92	384		3.71e-04	2.995
1536		2.28e-03	1.978	1536		4.63e-05	3
6144		5.72 e- 04	1.996	6144		5.79e-06	3.001
6	3	6.41e-02		6	4	7.48e-03	
24		1.17e-03	5.773	24		8.72e-04	3.096
96		2.71e-04	2.113	96		2.02e-05	5.429
384		1.72e-05	3.975	384		6.35e-07	4.995
1536		1.08e-06	3.998	1536		1.98e-08	5.002
6144		6.74 e- 08	4.001	6144		6.18e-10	5.001

Table 1: Convergence table for advection-diffusion problem with BR2 scheme, $\epsilon = 10$

Table 2: Convergence table for advection-diffusion problem with SIPG scheme, $\epsilon = 10$

Ν	q	$\ e\ _{L_2}$	order	Ν	q	$\ e\ _{L_2}$	order
6	1	1.83e-01		6	2	9.75e-02	
24		4.42e-02	2.053	24		3.50e-02	1.477
96		3.03e-02	0.5464	96		3.24e-03	3.435
384		8.37e-03	1.854	384		4.16e-04	2.96
1536		2.16e-03	1.954	1536		5.25e-05	2.987
6144		5.45e-04	1.985	6144		6.58e-06	2.997
6	3	7.23e-02		6	4	7.80e-03	
24		1.21e-03	5.906	24		1.01e-03	2.954
96		3.00e-04	2.005	96		2.28e-05	5.464
384		1.87e-05	4.001	384		7.27e-07	4.972
1536		1.17e-06	4.008	1536		2.28e-08	4.993
6144		7.26e-08	4.004	6144		7.13e-10	4.999

B. Navier-Stokes Equations

In this section first in order to see the accuracy of the method for the Navier-Stokes equations (1), we consider the following manufactured solution similar to²² and,²⁰ as

$$\rho(x, y) = 4 + \sin(2(x+y))$$
$$\rho V_1(x, y) = \rho V_2(x, y) = 2 + \frac{\sin(2(x+y))}{10}$$
$$E(x, y) = (4 + \sin(2(x+y)))^2$$

on a domain $\Omega = [0, \pi]^2$. Also we set the viscosity μ such that the Reynolds number defined as

$$Re = \frac{\rho_0 u_0 L_0}{\mu} = 500,\tag{60}$$

with the reference values $u_0 = 0.5$, $\rho_0 = 4$ and $L_0 = \pi$. The numerical convective flux is Lax-Friedrich and the method is tested with $\eta_e = 4$ for BR2 and with $\mu_e = 0.1q^2$ for SIPG. On the boundaries we applied inflow boundary conditions on left and bottom, and outflow on top and right. The results which presented in Tables 3 and 4 show the optimal order of convergence q + 1 in the asymptotic limit.

Ν	q	$\ e\ _{L_2}$	order	Ν	q	$\ e\ _{L_2}$	order
34	1	$4.50e{+}00$		34	2	6.63 e- 01	
136		$1.16e{+}00$	1.952	136		1.59e-01	2.061
544		2.97 e- 01	1.969	544		1.95e-02	3.027
2176		7.51e-02	1.985	2176		2.58e-03	2.919
34	3	3.65e-01		34	4	9.29e-02	
136		2.93e-02	3.641	136		5.07 e- 03	4.197
544		1.83e-03	4.005	544		2.30e-04	4.462
2176		1.21e-04	3.912	2176		7.10e-06	5.017

Table 3: Convergence table for the Navier-Stokes problems with BR2 scheme, Re = 500

Table 4: Convergence table for the Navier-Stokes problems with SIPG scheme, Re = 500

Ν	q	$\ e\ _{L_2}$	order	Ν	q	$\ e\ _{L_2}$	order
34	1	$4.26e{+}00$		34	2	5.65e-01	
136		$1.14e{+}00$	1.899	136		1.43e-01	1.983
544		2.89e-01	1.983	544		1.79e-02	2.997
2176		7.18e-02	2.009	2176		2.36e-03	2.922
34	3	3.31e-01		34	4	1.07e-01	
136		2.71e-02	3.608	136		4.93e-03	4.44
544		1.74e-03	3.962	544		2.18e-04	4.499
2176		1.17e-04	3.9	2176		7.25e-06	4.910

Next we are going to compute an actual flow computation around the NACA 0012 airfoil. The geometry of the airfoil is defined by

$$y(x) = \pm 0.6(0.2969\sqrt{x} - 0.1260x - 0.3516x^2 + 0.2843x^3 - 0.1036x^4), \qquad 0 \le x \le 1.$$
(61)

Our settings is laminar viscous flow with a free-stream Mach number of $Ma_{\infty} = 0.5$, angle of attack $\alpha = 2^{\circ}$ and Reynolds number of Re = 500. Also we set far field conditions as $\rho_{\infty} = 1$ and $p_{\infty} = 1$. The mesh is shown in Figure 1 consists of 2155 triangular elements which is created using the Netgen mesh generator.³¹ The far field is a circle, centered at the airfoil mid chord with a radius of 1000 chords.



Figure 1: Computational mesh, 2155 elements, zoomed around airfoil

We compute the solution with DG polynomial of degree q = 3 for both BR2 and SIPG method, with parameters $\eta_e = 4$ for BR2 and with $\mu_e = 0.01q^2$ for SIPG. As a reference solution we computed the same settings with our standard DG code (cf.³⁶). In Figure 2 we compared the distribution of Mach number around the airfoil for BR2 and SIPG method.



Figure 2: Distribution of flow Mach number around the airfoil NACA 0012.

In Figure 3 we compared the pressure coefficient on the airfoil surface, defined as

$$C_{\text{pressure}} = \frac{p - p_{\infty}}{\frac{1}{2}\rho_{\infty}V_{\infty}^2},\tag{62}$$

for both BR2 and SIPG method with our standard BR2 discretization. The results show that the pressure coefficient obtained from both methods are very similar to the results from the reference solution.



Figure 3: Pressure coefficient (C_{pressure}) on the airfoil, computed by BR2 and SIPG (red circles) with q = 3, versus the reference value computed by our standard DG code³⁶ (dashed blue line)

In order to see the effect of non-exact Jacobian on the convergence, in Figure 4 we present the convergence history of both SIPG and BR2 method. The y axis is the residual of the right hand side of the Newton iteration and the x axis is the iteration number. The big increases in the residual shows that steps of ramping in DG polynomial; i.e., we start by q = 0 and use the computed result as the initial condition for q = 1 and the process continues till the highest polynomial degree.



Figure 4: Convergence history of the Newton solver, residual versus the iteration number.

From the result in Figure 4 one can clearly see, that at least in this test case, the effect on non-exactness of the Jacobian is negligible in the convergence behaviour.

VI. Conclusion and Outlook

We presented a family of entropy-stable discontinuous Galerkin methods for the convection-diffusion systems, especially designed for the compressible Navier-Stokes equations. The discretization starts from a mixed formulation and we develop different well-known discretization in a canonical way. We also proved that all of these discretizations are entropy stable in their primal form. In the numerical experiments we showed some numerical result with BR2 and SIPG formulation to investigate their order of accuracy as well as their performance in flow computations in practice. As the results claimed, the order of accuracy of both methods are optimal in L_2 norm and they perform quite similar to the standard BR2 formulation we already developed (cf.³⁶). For the future work one might look at the effect of the boundary conditions on the claimed entropy stability and adjoint consistency.

Acknowledgments

The authors are supported by the Deutsche Forschungsgemeinschaft (German Research Association) through grant GSC 111.

References

¹D. N. ARNOLD, F. BREZZI, B. COCKBURN, AND L. D. MARINI, Unified analysis of discontinuous Galerkin methods for elliptic problems, SIAM journal on numerical analysis, 39 (2002), pp. 1749–1779.

²S. BALAY, S. ABHYANKAR, M. ADAMS, J. BROWN, P. BRUNE, K. BUSCHELMAN, V. EIJKHOUT, W. GROPP, D. KAUSHIK, M. KNEPLEY, ET AL., *PETSc users manual revision 3.5*, tech. rep., Argonne National Laboratory (ANL), 2014.

³T. J. BARTH, Numerical methods for gasdynamic systems on unstructured meshes, in An introduction to recent developments in theory and numerics for conservation laws, Springer, 1999, pp. 195–285.

⁴——, Simplified discontinuous Galerkin methods for systems of conservation laws with convex extension, in Discontinuous Galerkin Methods, Springer, 2000, pp. 63–75.

⁵——, On discontinuous Galerkin approximations of Boltzmann moment systems with Levermore closure, Computer Methods in Applied Mechanics and Engineering, 195 (2006), pp. 3311–3330.

⁶F. BASSI AND S. REBAY, A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier–Stokes equations, Journal of computational physics, 131 (1997), pp. 267–279.

⁷F. BASSI AND S. REBAY, *GMRES discontinuous Galerkin solution of the compressible Navier–Stokes equations*, in Discontinuous Galerkin Methods, Springer, 2000, pp. 197–208.

⁸F. BASSI, S. REBAY, G. MARIOTTI, S. PEDINOTTI, AND M. SAVINI, A high-order accurate discontinuous finite element method for inviscid and viscous turbomachinery flows, in Proceedings of the 2nd European Conference on Turbomachinery Fluid Dynamics and Thermodynamics, Technologisch Instituut, Antwerpen, Belgium, 1997, pp. 99–109.

⁹F. BREZZI, G. MANZINI, D. MARINI, P. PIETRA, AND A. RUSSO, *Discontinuous Galerkin approximations for elliptic problems*.

¹⁰R. BUSTINZA AND G. N. GATICA, A local discontinuous Galerkin method for nonlinear diffusion problems with mixed boundary conditions, SIAM Journal on Scientific Computing, 26 (2004), pp. 152–177.

¹¹M. H. CARPENTER, T. C. FISHER, E. J. NIELSEN, AND S. H. FRANKEL, *Entropy stable spectral collocation schemes for* the Navier–Stokes equations: Discontinuous interfaces, SIAM Journal on Scientific Computing, 36 (2014), pp. B835–B867.

¹²P. CHANDRASHEKAR, Discontinuous Galerkin method for Navier-Stokes equations using kinetic flux vector splitting, Journal of Computational Physics, 233 (2013), pp. 527–551.

¹³B. COCKBURN AND C. DAWSON, Some extensions of the local discontinuous Galerkin method for convection-diffusion equations in multidimensions, in The mathematics of finite elements and applications, X, MAFELAP 1999 (Uxbridge), Elsevier, Oxford, 2000, pp. 225–238.

¹⁴B. COCKBURN AND C.-W. SHU, The local discontinuous Galerkin method for time-dependent convection-diffusion systems, SIAM Journal on Numerical Analysis, 35 (1998), pp. 2440–2463.

¹⁵L. DIOSADY AND S. MURMAN, Higher-order methods for compressible turbulent flows using entropy variables, AIAA Paper, 294 (2015).

¹⁶P. DUTT, Stable boundary conditions and difference schemes for Navier–Stokes equations, SIAM Journal on Numerical Analysis, 25 (1988), pp. 245–267.

¹⁷T. C. FISHER AND M. H. CARPENTER, High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains, Journal of Computational Physics, 252 (2013), pp. 518–557.

¹⁸U. S. FJORDHOLM, *High-order accurate entropy stable numerical schemes for hyperbolic conservation laws*, PhD thesis, ETH Zürich, 2013.

¹⁹U. S. FJORDHOLM, S. MISHRA, AND E. TADMOR, Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws, SIAM Journal on Numerical Analysis, 50 (2012), pp. 544–573.

²⁰G. GASSNER, F. LÖRCHER, AND C.-D. MUNZ, A discontinuous Galerkin scheme based on a space-time expansion ii. viscous flow equations in multi dimensions, Journal of Scientific Computing, 34 (2008), pp. 260–286.

²¹R. HARTMANN AND P. HOUSTON, Symmetric interior penalty DG methods for the compressible Navier–Stokes equations I: Method formulation, Int. J. Num. Anal. Model., 3 (2006), pp. 1–20.

²²R. HARTMANN AND P. HOUSTON, An optimal order interior penalty discontinuous Galerkin discretization of the compressible Navier–Stokes equations, Journal of Computational Physics, 227 (2008), pp. 9670–9685.

²³A. HILTEBRAND AND S. MISHRA, Entropy stable shock capturing space-time discontinuous Galerkin schemes for systems of conservation laws, Numerische Mathematik, 126 (2014), pp. 103–151.

²⁴T. J. HUGHES, L. FRANCA, AND M. MALLET, A new finite element formulation for computational fluid dynamics: I. symmetric forms of the compressible Euler and Navier–Stokes equations and the second law of thermodynamics, Computer Methods in Applied Mechanics and Engineering, 54 (1986), pp. 223–234.

²⁵F. ISMAIL AND P. L. ROE, Affordable, entropy-consistent Euler flux functions ii: Entropy production at shocks, Journal of Computational Physics, 228 (2009), pp. 5410–5436.

²⁶R. M. KIRBY AND G. E. KARNIADAKIS, *De-aliasing on non-uniform grids: algorithms and applications*, Journal of Computational Physics, 191 (2003), pp. 249–264.

²⁷S. MAY, Spacetime discontinuous Galerkin methods for solving convection-diffusion systems, Tech. Rep. 2015-05, Seminar for Applied Mathematics, ETH Zürich, Switzerland, 2015.

²⁸S. MAY, Spacetime discontinuous Galerkin methods for convection-diffusion equations, Bulletin of the Brazilian Mathematical Society, New Series, 47 (2016), pp. 561–573.

²⁹T. A. OLIVER AND D. L. DARMOFAL, Analysis of dual consistency for discontinuous Galerkin discretizations of source terms, SIAM Journal on Numerical Analysis, 47 (2009), pp. 3507–3525.

³⁰I. PERUGIA AND D. SCHÖTZAU, An hp-analysis of the local discontinuous Galerkin method for diffusion problems, Journal of Scientific Computing, 17 (2002), pp. 561–571.

³¹J. SCHÖBERL, NETGEN an advancing front 2D/3D-mesh generator based on abstract rules, Computing and Visualization in Science, 1 (1997), pp. 41–52.

³²K. SHAHBAZI, An explicit expression for the penalty parameter of the interior penalty method, Journal of Computational Physics, 205 (2005), pp. 401–407.

³³E. TADMOR, The numerical viscosity of entropy stable schemes for systems of conservation laws. I, Mathematics of Computation, 49 (1987), pp. 91–103.

³⁴——, Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems, Acta Numerica, 12 (2003), pp. 451–512.

³⁵E. TADMOR AND W. ZHONG, Entropy stable approximations of Navier–Stokes equations with no artificial numerical viscosity, Journal of Hyperbolic Differential Equations, 3 (2006), pp. 529–559.

 36 M. WOOPEN, A. BALAN, G. MAY, AND J. SCHÜTZ, A comparison of hybridized and standard DG methods for target-based hp-adaptive simulation of compressible flow, Computers & Fluids, 98 (2014), pp. 3–16.

³⁷M. ZAKERZADEH AND G. MAY, Analysis of mixed discontinuous Galerkin formulation of quasilinear elliptic problems. In preparation, 2016.

³⁸M. ZAKERZADEH AND G. MAY, High order entropy stable discontinuous Galerkin schemes in a space-time computational framework, in 46th AIAA Fluid Dynamics Conference, 2016, p. 3496.

³⁹—, On the convergence of a shock capturing discontinuous Galerkin method for nonlinear hyperbolic systems of conservation laws, SIAM Journal on Numerical Analysis, 54 (2016), pp. 874–898.