Stability of Energy Stable Flux Reconstruction for the Diffusion Problem using the Interior Penalty and Bassi and Rebay II Numerical Fluxes for Linear Triangular Elements

Samuel Quaegebeur[a,*], Siva Nadarajah[a]

[a]*Department of Mechanical Engineering, McGill University, Montreal, QC, H3A 0C3, Canada*

Abstract

The flux reconstruction (FR) method has gained popularity within the research community. The approach has been demonstrated to recover high-order methods such as the discontinuous Galerkin (DG) method. Stability analyses have been conducted for a linear advection problem leading to the energy stable flux reconstruction (ESFR) methods also named Vincent-Castonguay-Jameson-Huynh (VCJH) methods. ESFR schemes can be viewed as DG schemes with modally filtered correction fields. Using this class of methods, the linear advection diffusion problem has been shown to be stable using the local discontinuous Galerkin scheme (LDG) to compute the viscous numerical flux. This stability proof has been extended for linear triangular and tetrahedra elements. Although the LDG scheme is commonly used, it requires, on particular meshes, a wide stencil, which raises the computational cost.

As a consequence, many prefer the compact interior penalty (IP) or the Bassi and Rebay II (BR2) numerical fluxes. This article, for the first time, derives, for both schemes, a condition on the penalty term to ensure stability. Moreover the article establishes that for both the IP and BR2 numerical fluxes, the stability of the ESFR scheme is independent of the auxiliary correction field. A von Neumann analysis is conducted to study the maximal time step of various ESFR methods.

Keywords: ESFR correction functions, Diffusion, Stability, Interior Penalty scheme, Bassi and Rebay 2 scheme

---

[*]Corresponding author
  Email address: `samuel.quaegebeur@mail.mcgill.ca` (Samuel Quaegebeur)

1. Introduction

Second-order finite-volume and finite-element methods form the current heart of most, if not all, commercial computational fluid dynamics packages, in addition to in-house codes at most aerospace manufacturers. In the past two decades, discontinuous Galerkin (DG) approaches have been advanced. They offer a combination of the strengths of both the finite-volume and finite-element approaches, where the concept of a numerical flux function, to provide stability, is combined with the finite-element approach of employing high-order shape functions to represent the solution [1]. A numerical flux ensures conservation across control volume faces and a number of these numerical fluxes for the discontinuous Galerkin (DG) approach have been developed for the diffusion equation such as the Bassi-Rebay schemes (BR1 [2] and BR2 [3]), interior penalty (IP) [4], local discontinuous Galerkin (LDG [5]), compact discontinuous Galerkin (CDG) [6]. For each one of these schemes, the DG method has been well-documented in terms of stability. These methods contain a penalty term, which controls the jump of the solution and/or gradient of the solution between the cells or control volumes. A judicious choice of this parameter ensures both stability and the correct order of accuracy of the scheme.

The FR framework developed by Huynh [7], [8] recovers, through the use of correction functions, many high-order methods including the DG and the spectral difference methods [9]. Stability analysis for the linear advection problem have been conducted for one [10], two [11] and three [12] dimensional problems, leading to the stable class of correction functions named Vincent Castonguay Jameson Huynh (VCJH) schemes. The extension from one dimension to higher dimensions has been achieved for linear simplices (triangles and tetrahedra). The mathematical proof is based on the energy of the solution and hence this class has taken the name Energy Stable Flux Reconstruction (ESFR) schemes. These stability proofs were then extended to the linear advection-diffusion problem using the LDG numerical fluxes, [13], [14], [12]. They obtained the stability of the scheme by taking the penalty term to be greater than 0. However the LDG numerical fluxes may require a wide stencil [14] and hence other compact methods are typically preferred. An extension of the stability of the ESFR scheme for one-dimensional problems using the compact IP and BR2 numerical fluxes has been conducted by Quaegebeur et al. [15]. The purpose of this article is to extend the stability proof of the IP and BR2 fluxes to two dimensional problems using linear triangular elements.

This article is composed as follows: Section 3 provides a proof to show that the problem, employing the IP or the BR2 numerical schemes, is independent of the auxiliary correction field;

2

Section 4 contains the theoretical proof of stability for the IP numerical fluxes as well as numerical verifications; Section 5 presents the proof of stability for the BR2 scheme; Section 6 demonstrates a von Neumann analysis showing the maximum time step. We conclude this paper with Section 7 where we present the L_2-errors for a given problem. Therefore this article establishes the necessary bounds for the IP and BR2 penalty terms to ensure stability for ESFR schemes. As the theoretical result is based on the work of Williams et al. [14], we strongly advise the reader to review [14]. The current article will use similar notations in an attempt to be as comprehensible as possible.

2. Preliminaries

In this section, we present the flux reconstruction approach for two dimensional problems using triangular elements, first introduced by Castonguay et al.[11].

Let us consider the diffusion equation

$$\frac{\partial u}{\partial t} = b \Delta u, \ \ (x, y) \in \Omega, \ t \in [0, T], \tag{1}$$

where x and y are the spatial coordinates of the physical domain Ω, b is the diffusion parameter, T is the final time and $\Delta = \left(\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} \right)$ is the Laplacian operator. We write the partial differential equation (1) as a system of two first-order equations. We introduce the operator $\nabla = \begin{pmatrix} \dfrac{\partial}{\partial x} \\ \dfrac{\partial}{\partial y} \end{pmatrix}$,

$$\frac{\partial u}{\partial t} = \nabla \cdot \mathbf{f}(\mathbf{q}), \tag{2a}$$

$$\mathbf{q} = \nabla u, \tag{2b}$$

where $\mathbf{f}$ is the flux, and $\mathbf{q}$ the auxiliary variable of the problem. For a pure diffusion problem, $\mathbf{f}(\mathbf{q}) = b\mathbf{q}$. In this article, a bold lowercase letter, $\mathbf{a}$, denotes a vector and a bold capital letter, $\mathbf{A}$ indicates a matrix.

The physical domain, Ω, can be decomposed into N_K non-overlapping linear triangular elements and thus we apply the tessalation, $\mathcal{T}_h = \sum_{n=1}^{N_K} \Omega_n$, where Ω_n denotes the element n. Before proceeding further on the FR procedure, we first define the computational domain and the reference element Ω_s. We arbitrarily, chose the equilateral element as our reference element. Choosing any other reference element such as the rectangular triangle does not impact our results.

3

Figure 1: Transformation, for triangles, from the physical space to the computational space.

Referring to Figure 1, the affine mapping can be defined as

$$
\mathcal{M}_n: \quad \begin{aligned} \Omega_s &\rightarrow \Omega_n \\ (r,s) &\mapsto \frac{\left(-3r + 2 - \sqrt{3}s\right)}{6}\mathbf{v}_{1,n} + \frac{\left(2 + 3r - \sqrt{3}s\right)}{6}\mathbf{v}_{2,n} + \frac{\left(2 + 2\sqrt{3}s\right)}{6}\mathbf{v}_{3,n}, \end{aligned} \tag{3}
$$

where $\Omega_s = \left\{ (r,s) \mid -1 \le r \le 1\,,\; \frac{-1}{\sqrt{3}} \le s \frac{2}{\sqrt{3}}\,,\; \left|\frac{3}{\sqrt{3}}r + s\right| \le \frac{2}{\sqrt{3}} \right\}$, $\mathbf{v}_{1,n}$, $\mathbf{v}_{2,n}$ and $\mathbf{v}_{3,n}$ are the three vertices of the triangle Ω_n.

With this mapping, we define $\hat{u}$, $\hat{\mathbf{f}}$ and $\hat{\mathbf{q}}$ as the computational quantities of u, $\mathbf{f}$ and $\mathbf{q}$. We then define the Jacobian of the triangle Ω_n and its determinant by $\mathbf{J}_n = \begin{pmatrix} x_r & x_s \\ y_r & y_s \end{pmatrix}$ and $|J_n| = Det\left(\mathbf{J_n}\right)$. We use the notation a_b to indicate the partial derivative of a with respect to b. The elements are linear triangles, therefore $\mathbf{J_n}$ is constant within Ω_n. By convention, the computational quantites are calculated as

$$
\hat{\nabla} = \mathbf{J}_n^T \nabla, \tag{4}
$$

$$
\hat{u}_n = |J_n|\, u_n, \tag{5}
$$

$$
\hat{\mathbf{f}}_n = |J_n|\, \mathbf{J}_n^{-1}\mathbf{f}_n, \tag{6}
$$

$$
\hat{\mathbf{q}}_n = \hat{\nabla}\hat{u} = |J_n|\, \mathbf{J}_n^T \mathbf{q}_n, \tag{7}
$$

$$
\hat{\mathbf{n}}^e = \frac{1}{|J_n|}\,|J^e|\, \mathbf{J}^T \mathbf{n}^e, \tag{8}
$$

where $\mathbf{n}^e$ is the outward normal to the edge e, $|J^e|$ is the length of the edge, and $\hat{\nabla} = \begin{pmatrix} \frac{\partial}{\partial r} \\ \frac{\partial}{\partial s} \end{pmatrix}$ is the divergence operator in the computational domain. We also note $\mathbf{r} = (r,s)$.

4

With these conventional transformations, we obtain

$$\frac{\partial \hat{u}}{\partial t} = \hat{\nabla} \cdot \hat{\mathbf{f}}\left(\hat{\mathbf{q}}\right), \tag{9a}$$

$$\hat{\mathbf{q}} = \hat{\nabla} u, \tag{9b}$$

where, $\hat{\mathbf{f}}_n = \mathbf{J}_n^{-1} \mathbf{J}_n^{-T} \hat{\mathbf{q}}_n$.

The solution u_n is a two-dimensional polynomial of degree p ($u_n \in P_p\left(\Omega_s\right)$). It is calculated through $N_p = \frac{(p+2)(p+1)}{2}$ solution points (SP), represented in Figure 2,

$$u_n\left(\mathbf{r}\right) = \sum_{i=1}^{N_p} \tilde{u}_i l_i\left(\mathbf{r}\right), \tag{10}$$

where l_i is a two-dimensional Lagrange polynomial of degree p, $l_i \in P_p\left(\Omega_s\right)$ (note that the dimension of $P_p\left(\Omega_s\right)$ is N_p). $(\tilde{u}_i)_{i \in [\![1,N_p]\!]}$ are the nodal expansion coefficients associated to the solution points $(r_i)_{i \in [\![1,N_p]\!]}$ (interval of the type $[\![\,\cdot\,,\,\cdot\,]\!]$ denotes an integer interval). Conversely to the 1D problem, there is no analytical expression for the Lagrange polynomial except for certain nodal distributions (e.g. equi-spaced). This issue is resolved by considering the solution in modal form

$$u_n\left(\mathbf{r}\right) = \sum_{i=1}^{N_p} \tilde{u}_i^m L_i\left(\mathbf{r}\right), \tag{11}$$

where $\tilde{u}_i^m$ are the modal coefficients and $(L_i)_{i \in [\![1,N_p]\!]}$ forms an orthonormal basis on the equilateral reference element. Initially, at time t_0, we have the solution on element Ω_n, $u_n^{t_0}$ at its N_p solution points (the N_p nodal coefficients). If we want to calculate the solution elsewhere, we evaluate the solution under its modal form. To do this, we compute the Vandermonde matrix [1], $\mathcal{V}$, and perform the operation $\tilde{u}^m = \mathcal{V}^{-1} \tilde{u}$. Having the modal coefficients $\tilde{u}^m$ and the analytical expression for the orthonormal basis L_i, we can compute u_n at any point with (11). The orthonormal basis (L_i) is of utmost importance as it enables the evaluation of all quantities of the problem.

Figure 2: Reference element for $p = 2$, the green squares represent the Gauss-Legendre Flux Points (FP), the red circles represent the solution points (SP) and the black circles are the vertices.

We now consider an element Ω_n of the domain and apply the FR procedure on equations (9a) and (9b). This results in a correction of the solution, within element Ω_n, on each of its faces. For a given face, this correction is applied on $N_{fp} = p + 1$ flux points, represented in Figure 2.

$$\frac{\partial \hat{u}_n}{\partial t} = \hat{\nabla} \cdot \hat{\mathbf{f}}_n \left(\hat{\mathbf{q}}_n \right) + \hat{\nabla} \cdot \sum_{f=1}^{3} \sum_{j=1}^{Nfp} \left[\left(\hat{\mathbf{f}}_{n,fj}^* - \hat{\mathbf{f}}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \mathbf{h}_{fj} \left(\mathbf{r} \right), \tag{12a}$$

$$\hat{\mathbf{q}}_n = \hat{\nabla} \hat{u}_n + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \left(\hat{u}_{n,fj}^* - \hat{u}_{n,fj} \right) \psi_{fj} \left(\mathbf{r} \right) \hat{\mathbf{n}}_{fj}, \tag{12b}$$

where f denotes the faces of the element, j is the index over the flux points, and lastly u^* and $\mathbf{f}^*$ are the numerical fluxes. $\mathbf{h}_{fj}$ corresponds to the correction function vector of the primary equation associated to the Flux Point (FP) (f, j). In addition ψ_{fj} is the correction field associated to the correction function $\mathbf{g}_{fj}$ of the FP (f, j). We have, for the primary equation, $\hat{\nabla} \cdot \mathbf{h}_{fj} = \phi_{fj}$ and $\hat{\nabla} \cdot \mathbf{g}_{fj} = \psi_{fj}$ for the auxiliary equation. The correction vectors are introduced to create continuous quantities across the edges and hence satisfy

$$\mathbf{h}_{fj} \left(\mathbf{r}_k^l \right) \cdot \hat{\mathbf{n}}_{kl} = \delta_{fk} \delta_{jl}, \tag{13}$$

where δ is the Kronecker delta and $\mathbf{r}_k^l$ is the coordinates of the FP (k, l). Hence, on the boundary of the reference element, the outward component of the correction function $\mathbf{h}_{fj}$ is equal to 1 on the FP (f, j) and 0 on the others.

6

Moreover the correction functions and their divergence must satisfy additional properties to ensure stability of the advection problem. Full details of these properties can be found in the article of Castonguay et al. [11]. For brevity, we present just a few of them. We consider that the vector correction function $\mathbf{h}_{fj}$ associated to the FP (f, j) lies in the Raviart-Thomas (RT) space of order p, $RT_p(\Omega_s) = (P_p(\Omega_s))^2 + \begin{pmatrix} r \\ s \end{pmatrix} P_p(\Omega_s)$. As a result, two properties of $\mathbf{h}_{fj}$ have been demonstrated [11],

$$\mathbf{h}_{fj} \cdot \hat{\mathbf{n}}_{\Gamma_s} \in R_p(\Gamma_s) \; \forall f \in [\![1,3]\!], \forall j \in [\![1, N_{fp}]\!], \tag{14}$$

$$\hat{\nabla} \cdot \mathbf{h}_{fj} \in P_p(\Omega_s) \; \forall f \in [\![1,3]\!], \forall j \in [\![1, N_{fp}]\!], \tag{15}$$

where Γ_s is the boundary of the reference element Ω_s, $\hat{\mathbf{n}}_{\Gamma_s}$ is the normal of one of the faces of Γ_s and $R_p(\Gamma_s)$ is defined as

$$R_p(\Gamma_s) = \left\{ \phi \mid \phi \in L^2(\Gamma_s), \; \phi|_{\Gamma_{s,f}} \in P_p(\Gamma_{s,f}) \, \forall \Gamma_{s,f} \right\}, \tag{16}$$

where $\Gamma_{s,f}$ is one of the edges of Γ_s.

Equation (15) signifies that the divergence of the correction function associated to the FP (f, j) is a two-dimensional polynomial of degree p on the element while equation (14) expresses that the outward component of the correction function associated to FP (f, j) along the edge of any face of the element is a 1D polynomial of degree p.

The following additional property is required to obtain a stable scheme: Castonguay et al. [11], defined a class of correction functions for triangular elements, called ESFR, which is stable for the linear advection problem. This property was proposed such that additional terms, arising from the ESFR norm (defined later in equation (32)), are removed. In the proof [11], using this new norm, lower bounds are found for all the remaining terms and stability is obtained. The correction functions must satisfy,

$$c \sum_{m=1}^{p+1} \binom{p}{m-1} \left(D^{(m,p)} L_i \right) \left(D^{(m,p)} \phi_{fj} \right) = \int_{\Omega_s} \mathbf{h}_{fj} \cdot \hat{\nabla} L_i \, \mathrm{d}\Omega_s, \tag{17}$$

where c is the correction function parameter and $D^{(m,p)} = \dfrac{\partial^p}{\partial r^{p-m+1} \partial s^{m-1}}$ is the derivative operator. We then define

$$\hat{\nabla} \cdot \mathbf{h}_{fj} = \phi_{fj}(\mathbf{r}) = \sum_{i=1}^{N_p} \sigma_{fj,i}^c L_i(\mathbf{r}), \tag{18}$$

7

where $\left(\sigma_{fj,i}^c\right)_{i\in[\![1,N_p]\!]}$ are the coefficients of the correction field ϕ_{fj}. If these coefficients respect the following equation, then ϕ_{fj} is an ESFR correction field and the linear advection problem is stable for triangles using the Lax-Friedrichs numerical flux. Replacing equation (18) in equation (17) yields,

$$c\sum_{k=1}^{N_p}\sigma_{fj,k}^c\sum_{m=1}^{p+1}\binom{p}{m-1}\left(D^{(m,p)}L_i\right)\left(D^{(m,p)}L_k\right)=-\sigma_{fj,i}^c+\int_{\Gamma_s}(\mathbf{h}_{fj}\cdot\hat{\mathbf{n}})\,L_i\,\mathrm{d}\Gamma_s\ \ \forall i\in[\![1,N_p]\!]. \quad (19)$$

Besides ensuring the stability of the advection problem, the above equation enables the correction fields to have mirror and rotational symmetry as explained in [11]. In the following sections, both the primary ($\mathbf{h}_{fj}$) and the auxiliary ($\mathbf{g}_{fj}$) correction functions are taken such that they satisfy the above properties. The auxiliary correction field is parametrized by κ and is defined through the coefficients $\left(\sigma_{fj,i}^{\kappa}\right)_{i\in[\![1,N_p]\!]}$,

$$\kappa\sum_{k=1}^{N_p}\sigma_{fj,k}^{\kappa}\sum_{m=1}^{p+1}\binom{p}{m-1}\left(D^{(m,p)}L_i\right)\left(D^{(m,p)}L_k\right)=-\sigma_{fj,i}^{\kappa}+\int_{\Gamma_s}(\mathbf{g}_{fj}\cdot\hat{\mathbf{n}})\,L_i\,\mathrm{d}\Gamma_s\ \ \forall i\in[\![1,N_p]\!]. \quad (20)$$

Let us now enumerate the different choices of the ESFR method for the diffusion problem:

1. The correction field ψ for the auxiliary equation is parametrized by the parameter κ (Equation (20)). In the following section, we will show that when employing the IP or BR2 numerical fluxes, the method does not depend on κ. To verify this result, we will run numerical simulations for two values, κ_{DG} and κ_+. These values will be further elaborated in the following section.

2. The correction field ϕ for the primary equation. This correction field is parametrized by c. Castonguay et al. [11] studied the influence of c for the advection problem. He defined a particular value of c noted as c_+, which yields the maximal stable time step.

p \ γ	60°	90°
2	4.3e-02	4.3e-02
3	6.4e-04	6.0e-04

Table 1: Numerical values for c_+ for different regular meshes and orders of polynomial p obtained through a von Neumann analysis.

The parameter c_+ given in Table 1 was obtained through a von Neumann analysis for $p = 2$ and $p = 3$ for an advection problem. c_+ also depends on the shape of the elements, and hence on γ, as explained in Section 6.

3. For each element, the solution u is computed at N_p solution points and N_{fp} flux points. The stability proof requires the use of the Gauss-Legendre flux points in order to evaluate exactly the integrals over the edges. However there is no constraint on the solution points. For our numerical simulations, we chose, arbitrarily, the α-optimized solution points proposed in the book of Hesthaven and Warburton [1].

4. Two numerical fluxes will be studied: the IP and BR2 schemes

$$
\begin{cases}
\text{IP:} \\
u^* = \{\{u\}\} \\
\mathbf{q}^* = \{\{\nabla u\}\} - \tau[\![\mathbf{u}]\!]
\end{cases}
\qquad
\begin{cases}
\text{BR2:} \\
u^* = \{\{u\}\} \\
\mathbf{q}^* = \{\{\nabla u\}\} + s\{\{\mathbf{r^e}([\![\mathbf{u}]\!])\}\}
\end{cases}
\tag{21}
$$

where τ is the penalty term for the IP scheme and s is the penalty term for the BR2 scheme; while $\mathbf{r^e}$ is a lifting operator and is defined as

$$
\int_\Omega \mathbf{r^e}([\![\mathbf{u}]\!]) \cdot \mathbf{\Phi} \, d\Omega = -\int_{\Gamma_e} [\![\mathbf{u}]\!] \cdot \{\{\mathbf{\Phi}\}\} \, d\Gamma_e,
\tag{22}
$$

where $\mathbf{\Phi}$ is a vector test function and Γ_e denotes the edge e. The notations $[\![\ ]\!]$ denotes the jump and $\{\{\ \}\}$ the mean value. While the former is the difference of the solution across an edge e, the latter is the average,

$$
[\![\mathbf{u}]\!]_e = u_{e,-}\mathbf{n}_-^e + u_{e,+}\mathbf{n}_+^e,
\tag{23}
$$

$$
\{\{u\}\}_e = \frac{1}{2}(u_{e,+} + u_{e,-}),
\tag{24}
$$

where $u_{e,-}$ and $\mathbf{n}_-^e$ denotes the interior solution and the outward interior normal vector of edge e, while $u_{e,+}$ and $\mathbf{n}_+^e$ signify the exterior solution and the outward exterior edge normal. Figure 4 provides the geometric interpretation of these quantities. Notice, while $[\![\mathbf{u}]\!]$ is a vector, $[\![q]\!]$ is a scalar value. Conversely, $\{\{u\}\}$ is a scalar value but $\{\{\mathbf{q}\}\}$ is a vector.

3. The diffusion equation, independent of κ

The purpose of this section is to prove that using the ESFR schemes with the IP and BR2 numerical fluxes results in an independency of the problem from the parameter κ, associated with the correction function for the auxiliary equation.

From equation (12b), the transformed solution correction at each flux point for both the IP and BR2 schemes as shown in equation (21) yields,

$$
\begin{aligned}
\left(\hat{u}^*_{n,fj} - \hat{u}_{n,fj} \right) &= \{\{\hat{u}\}\}_{fj} - \hat{u}_{fj} \\
&= -\frac{[\![\hat{u}]\!]_{fj}}{2}.
\end{aligned}
\tag{25}
$$

Note that $[\![\hat{u}]\!]_{fj}$ is a scalar value of the jump on face f at the flux point j. However the IP and BR2 schemes differ only for the numerical flux function, $\hat{\mathbf{f}}^*$ for the primary equation as clearly stated in equation (21). The schemes only depend on the mean of the solution gradient and the solution jump but do not depend on the auxiliary solution, $\hat{\mathbf{q}}$ and hence do not depend on the parameter κ.

By including the numerical flux into the transformed solution correction as shown in equation (25) into equations (12a) and (12b), we can represent the ESFR scheme as,

$$
\frac{\partial \hat{u}_n}{\partial t} = \hat{\nabla} \cdot \hat{\mathbf{f}}_n\left(\hat{\mathbf{q}}_n\right) + \hat{\nabla} \cdot \sum_{f=1}^{3}\sum_{j=1}^{N_{fp}} \left[\left(\hat{\mathbf{f}}^*_{n,fj} - \hat{\mathbf{f}}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \mathbf{h}_{fj}\left(\mathbf{r}\right),
\tag{26a}
$$

$$
\hat{\mathbf{q}}_n = \hat{\nabla}\hat{u}_n - \sum_{e=1}^{3}\sum_{i=1}^{N_{fp}} \frac{[\![\hat{u}]\!]_{ei}}{2}\psi_{ei}\left(\mathbf{r}\right)\hat{\mathbf{n}}_{ei}.
\tag{26b}
$$

Note that we have renamed the indices f and j in equation (26b) by e and i.

We now introduce the auxiliary equation (26b) into the primary (26a),

$$
\begin{aligned}
\frac{\partial \hat{u}_n}{\partial t} =\ & \hat{\nabla} \cdot \left(\mathbf{J}_n^{-1}\mathbf{J}_n^{-T}\hat{\nabla}\hat{u}_n \right) + \sum_{f=1}^{3}\sum_{j=1}^{N_{fp}} \left[\left(\hat{\mathbf{f}}^*_{n,fj} - \mathbf{J}_n^{-1}\mathbf{J}_n^{-T}\hat{\nabla}\hat{u}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \phi_{fj}\left(\mathbf{r}\right) \\
& -\hat{\nabla} \cdot \left(\mathbf{J}_n^{-1}\mathbf{J}_n^{-T}\sum_{e=1}^{3}\sum_{i=1}^{N_{fp}} \frac{[\![\hat{u}]\!]_{ei}}{2}\psi_{ei}\left(\mathbf{r}\right)\hat{\mathbf{n}}_{ei} \right) \\
& +\sum_{f=1}^{3}\sum_{j=1}^{N_{fp}} \left[\mathbf{J}_n^{-1}\mathbf{J}_n^{-T}\left(\sum_{e=1}^{3}\sum_{i=1}^{N_{fp}} \frac{[\![\hat{u}]\!]_{ei}}{2}\psi_{ei}\left(\mathbf{r}_f^j\right)\hat{\mathbf{n}}_{ei} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \phi_{fj}\left(\mathbf{r}\right),
\end{aligned}
\tag{27}
$$

where $\psi_{ei}\left(\mathbf{r}_f^j\right)$ is the correction function of the edge e computed at the FP (f,j). On further

simplification, the primary equation can be represented as,

$$
\begin{aligned}
\frac{\partial \hat{u}_n}{\partial t} =\ & \hat{\nabla} \cdot \left(\mathbf{J}_n^{-1} \mathbf{J}_n^{-T} \hat{\nabla} \hat{u}_n \right) + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \left[\left(\hat{\mathbf{f}}_{n,fj}^* - \mathbf{J}_n^{-1} \mathbf{J}_n^{-T} \hat{\nabla} \hat{u}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \phi_{fj}(\mathbf{r}) \\
& + \mathbf{J}_n^{-1} \mathbf{J}_n^{-T} \sum_{e=1}^{3} \sum_{i=1}^{N_{fp}} \left[\frac{[\![\hat{u}]\!]_{ei}}{2} \left(-\hat{\nabla}\psi_{ei}(\mathbf{r}) \cdot \hat{\mathbf{n}}_{ei} + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \psi_{ei}\left(\mathbf{r}_f^j\right) (\hat{\mathbf{n}}_{ei} \cdot \hat{\mathbf{n}}_{fj}) \phi_{fj}(\mathbf{r}) \right) \right].
\end{aligned}
\tag{28}
$$

Postulate 3.1. *Let the solution of the diffusion equation be approximated by a polynomial of degree p on the reference triangle (see Figure 2 for an example where $p = 2$). Let ψ_{ei} be the correction field associated to face e at the flux point i parametrized by κ and ϕ_{fj} the correction field, associated to face f at the flux point j, parametrized by c. Let $\left(\mathbf{r}_f^j\right)_{j \in [\![1,N_{fp}]\!]}$ be the Gauss-Legendre flux points on face f. Then $\forall c \in [0, \infty[$,*

$$
R_{ei}(\mathbf{r}) = \left(-\hat{\nabla}\psi_{ei}(\mathbf{r}) \cdot \hat{\mathbf{n}}_{ei} + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \psi_{ei}\left(\mathbf{r}_f^j\right) (\hat{\mathbf{n}}_{ei} \cdot \hat{\mathbf{n}}_{fj}) \phi_{fj}(\mathbf{r}) \right),
\tag{29}
$$

is independent of the parameter κ.

This two-dimensional postulate was proposed based on the one-dimensional theoretical work [15]. An analytical proof is presented in Appendix A for $p = 1$; however, a general proof, for higher p, currently eludes the authors since specific properties of the orthonormal basis $(L_i)_{i \in [\![1,N_p]\!]}$ are required. This Postulate 3.1 has been, nonetheless, verified numerically for every face e and flux point i, for several values of c and for an order up to $p = 6$. For brevity, we will present only the result for the function,

$$
R_{11} = \left(-\hat{\nabla}\psi_{11}(\mathbf{r}) \cdot \hat{\mathbf{n}}_{11} + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \psi_{11}\left(\mathbf{r}_f^j\right) (\hat{\mathbf{n}}_{11} \cdot \hat{\mathbf{n}}_{fj}) \phi_{fj}(\mathbf{r}) \right).
\tag{30}
$$

Two values of p will be taken, $p = 2$ and $p = 3$. The correction function $\mathbf{h}_{fj}$ will be computed for two values of c: c_{DG} and c_+. In each of the following graphs, $R_{11}(\mathbf{r})$ will be evaluated for all N_p solution points.

11

(a) R_{11} for $c = c_{DG}$ and $p = 2$.

(b) R_{11} for $c = c_{DG}$ $p = 3$.

(c) R_{11} for $c = c_+$ $p = 2$.

(d) R_{11} for $c = c_+$ $p = 3$.

Figure 3: Influence of parameter κ on function R_{11} for various values of parameter c and polynomial order p.

Postulate 3.2. *The diffusion equation is independent of κ when employing the IP or BR2 numerical fluxes.*

Proof. The diffusion equation with the ESFR schemes and employing the IP or BR2 numerical fluxes is written as equation (28). Only the last line of this discretization contains the parameter κ. However from Postulate 3.1, it does not depend on κ. □

4. IP stability condition

4.1. Theoretical result

Having introduced the 2D FR concept, we can now derive the stability condition for the IP scheme using the ESFR class functions. This proof has been derived by Williams et al. [14] for the LDG scheme using linear triangles. For brevity, we will not re-derive all the steps but instead begin from Equation (81) of the aforementioned article. Upon removing the advective term, we obtain:

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|U\|_{p,c}^2 = -b\|\mathbf{Q}\|_{p,\kappa}^2 + b\Theta_{dif}, \tag{31}$$

where,

$$\|U\|_{p,c} = \left\{\sum_{n=1}^{N_K}\int_{\Omega_n}\left[(u_n)^2 + \frac{1}{A_s}\sum_{m=1}^{p+1}c_m\left(D^{(m,p)}u_n\right)^2\right]\mathrm{d}\Omega_n\right\}^{1/2}, \tag{32}$$

$$\|\mathbf{Q}\|_{p,\kappa} = \left\{\sum_{n=1}^{N_K}\int_{\Omega_n}\left[(\mathbf{q}_n)^2 + \frac{1}{A_s}\sum_{m=1}^{p+1}\kappa_m\left(D^{(m,p)}\mathbf{q}_n\right)^2\right]\mathrm{d}\Omega_n\right\}^{1/2}, \tag{33}$$

are broken Sobolev-type norms for the solution and the auxiliary variables, and

$$\Theta_{dif} = \sum_{n=1}^{N_K}\left\{\int_{\Gamma_n}\left[-u_n\left(\mathbf{q}_n\cdot\mathbf{n}\right) + u_n\left(\mathbf{q}_n^*\cdot\mathbf{n}\right) + u_n^*\left(\mathbf{q}_n\cdot\mathbf{n}\right)\right]\mathrm{d}\Gamma_n\right\}, \tag{34}$$

represents contributions of the diffusive fluxes from the boundaries. The coefficients $c_m = c\binom{p}{m-1}$, $\kappa_m = \kappa\binom{p}{m-1}$, A_s denotes the area of the reference equilateral triangle, and Γ_n defines all the edges of an element Ω_n. In the previous equation, the positive scalar value b multiplies every term of Θ_{dif} and is hence factored out. This is a slight difference from [14], where b does not multiply the penalty term. This choice does not impact our result as it is just a nondimensionalization of the penalty term. To ensure energy stability, we must have the right hand side of equation (31) to be less than equal to zero, where $\left(-b\|\mathbf{Q}\|_{p,\kappa}^2\right)$ is ensured to be negative granted that the coefficient b is positive. Hence we must demonstrate that Θ_{dif} is non-positive for appropriate choices of numerical fluxes and interface solutions.

13

Figure 4: Part of an arbitrary physical domain.

For the rest of our proof, we consider the edge e represented in Figure 4. The triangle IJK is the interior element of edge e and triangle IJK_{+1}, the exterior element of edge e.

Lemma 4.1. *Employing the IP scheme and considering periodic boundary conditions, the following equation holds*

$$\Theta_{dif} = \sum_{e=1}^{N_e} \int_{\Gamma_e} \left(-\{\{\mathbf{q} - \nabla u\}\}\right) \cdot [\![\mathbf{u}]\!]_e - \tau [\![\mathbf{u}]\!]_e \cdot [\![\mathbf{u}]\!]_e \, d\Gamma_e. \tag{35}$$

Proof. Instead of considering an element based formulation, we modify Θ_{dif} in (34) to an edge based formulation,

$$\Theta_{dif} = \sum_{e=1}^{N_e} \int_{\Gamma_e} \left[-u_n \left(\mathbf{q}_n \cdot \mathbf{n}\right) + u_{n+1} \left(\mathbf{q}_{n+1} \cdot \mathbf{n}\right) + \mathbf{q}_e^* \cdot \mathbf{n} \left(u_n - u_{n+1}\right) + u_e^* \left(\mathbf{q}_n - \mathbf{q}_{n+1}\right) \cdot \mathbf{n}\right] d\Gamma_e, \tag{36}$$

where $\mathbf{n} = \mathbf{n}_-^e$ as shown in Figure 4.

14

We then use the IP numerical fluxes for both u^* and $\mathbf{q}^*$ as defined in (21) to expand (36)

$$
\begin{aligned}
\Theta_{dif} &= \sum_{e=1}^{N_e} \int_{\Gamma_e} \left(-u_n\mathbf{q}_n + u_{n+1}\mathbf{q}_{n+1}\right) \cdot \mathbf{n} + \left(\{\{\nabla u\}\} - \tau[\![\mathbf{u}]\!]_e\right) \cdot [\![\mathbf{u}]\!]_e + \{\{u\}\}[\![q]\!] \mathrm{d}\Gamma_e, \\
&= \sum_{e=1}^{N_e} \int_{\Gamma_e} \Big[-u_n\mathbf{q}_n \cdot \mathbf{n}\left(1 - \tfrac{1}{2}\right) \\
&\qquad\qquad + u_{n+1}\mathbf{q}_{n+1} \cdot \mathbf{n}\left(1 - \tfrac{1}{2}\right) \\
&\qquad\qquad + \tfrac{1}{2}\left(-u_n\mathbf{q}_{n+1} + u_{n+1}\mathbf{q}_n\right) \cdot \mathbf{n} \\
&\qquad\qquad + \left(\{\{\nabla u\}\} - \tau[\![\mathbf{u}]\!]_e\right) \cdot [\![\mathbf{u}]\!]_e \Big] \mathrm{d}\Gamma_e, \\
&= \sum_{e=1}^{N_e} \int_{\Gamma_e} \tfrac{1}{2}\left[\left(u_{n+1} - u_n\right)\left(\mathbf{q}_n + \mathbf{q}_{n+1}\right) \cdot \mathbf{n}\right] + \left(\{\{\nabla u\}\} - \tau[\![\mathbf{u}]\!]_e\right) \cdot [\![\mathbf{u}]\!]_e \mathrm{d}\Gamma_e \\
&= \sum_{e=1}^{N_e} \int_{\Gamma_e} -\{\{\mathbf{q}\}\} \cdot [\![\mathbf{u}]\!]_e + \left(\{\{\nabla u\}\} - \tau[\![\mathbf{u}]\!]_e\right) \cdot [\![\mathbf{u}]\!]_e \mathrm{d}\Gamma_e.
\end{aligned}
\tag{37}
$$

We finally obtain the desired equation

$$
\Theta_{dif} = \sum_{e=1}^{N_e} \int_{\Gamma_e} \left(-\{\{\mathbf{q} - \nabla u\}\}\right) \cdot [\![\mathbf{u}]\!]_e - \tau[\![\mathbf{u}]\!]_e \cdot [\![\mathbf{u}]\!]_e \, \mathrm{d}\Gamma_e.
\tag{38}
$$

$\square$

Remark 4.1. Notice that if we had taken the LDG scheme then ∇u would have been replaced by $\mathbf{q}$ and we would have had the stability of the scheme for $\tau \geq 0$ as shown by [14].

However, for the IP and BR2 formulations, additional effort is required to demonstrate conditions on stability. We decompose the result of Lemma 4.1 into two components,

$$
\Theta_{dif} = \sum_{e=1}^{N_e} \left[\underbrace{\int_{\Gamma_e} \left(-\{\{\mathbf{q} - \nabla u\}\}\right) \cdot [\![\mathbf{u}]\!]_e \, \mathrm{d}\Gamma_e}_{\Theta_{e,1}} - \underbrace{\int_{\Gamma_e} \tau[\![u]\!]_e^2 \, \mathrm{d}\Gamma_e}_{\Theta_{e,2}} \right],
\tag{39}
$$

where $[\![u]\!]_e = [\![\mathbf{u}]\!]_e \cdot \mathbf{n}$. We now investigate separately $\Theta_{e,1}$ and $\Theta_{e,2}$.

Lemma 4.2. *Using the auxiliary equation, defined in (12b) and the Gauss-Legendre quadratures,*

15

$\Theta_{e,1}$ can be computed exactly,

$$
\begin{aligned}
\Theta_{e,1} &= |J^e| \frac{F^e_{s,-}}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]^2_{ei} \omega_i \psi_{ei,-}(\mathbf{r}_i) + [\![u]\!]_{ei} \omega_i \left(\sum_{\substack{j=1\\j\neq i}}^{N_{fp}} [\![u]\!]_{ej} \psi_{ej,-}(\mathbf{r}_i) \right) \right] \\
&+ |J^e| \frac{F^e_{s,+}}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]^2_{ei} \omega_i \psi_{ei,+}(\mathbf{r}_i) + [\![u]\!]_{ei} \omega_i \left(\sum_{\substack{j=1\\j\neq i}}^{N_{fp}} [\![u]\!]_{ej} \psi_{ej,+}(\mathbf{r}_i) \right) \right] \\
&+ |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei} \omega_i \left(\sum_{\substack{f=1\\f\neq e}}^{N_{fp}} \frac{F^f_{s,-}(\mathbf{n}\cdot\mathbf{n}_{f,-})}{4} \sum_{j=1}^{N_{fp}} [\![u]\!]_{fj,-} \psi_{fj,-}(\mathbf{r}_i) \right) \right] \\
&+ |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei} \omega_i \left(\sum_{\substack{f=1\\f\neq e}}^{N_{fp}} \frac{F^f_{s,+}(\mathbf{n}\cdot\mathbf{n}_{f,+})}{4} \sum_{j=1}^{N_{fp}} [\![u]\!]_{fj,+} \psi_{fj,+}(\mathbf{r}_i) \right) \right],
\end{aligned}
\tag{40}
$$

where $\mathbf{r}_i$ and ω_i are the Gauss-Legendre nodes and weights and $|J^e|$ is the length of edge e. $F^f_{s,-}$ (resp. $F^f_{s,+}$) is the ratio of the length of an edge f over the determinant of the Jacobian of the interior (resp. exterior) element $\left(\frac{|J^f|}{|J_{e,\pm}|} \right)$.

Proof. We use equation (12b) along with the transformation equalities from the physical space to the computational space, equations (4) and (7). Instead of using element indices n, $n+1$, we use the subscript $+$ and $-$ as shown in Figure 4. In the term $\Theta_{e,1}$ in equation (39), we replace $[\![\mathbf{u}]\!]_e$ by $[\![u]\!]_e \mathbf{n}$,

$$
\begin{aligned}
\Theta_{e,1} &= \int_{\Gamma_e} -\frac{[\![u]\!]_e}{2} \left[\frac{1}{|J_{e,-}|} \mathbf{J}^{-T}_{e,-} \left(\hat{\mathbf{q}}_{e,-} - \hat{\nabla}\hat{u}_{e,-} \right) + \frac{1}{|J_{e,+}|} \mathbf{J}^{-T}_{e,+} \left(\hat{\mathbf{q}}_{e,+} - \hat{\nabla}\hat{u}_{e,+} \right) \right] \cdot \mathbf{n}\, \mathrm{d}\Gamma_e \\
&= \int_{\Gamma_e} -\frac{[\![u]\!]_e}{2} \left[\frac{1}{|J_{e,-}|} \mathbf{J}^{-T}_{e,-} \left(\sum_{f=1}^{3}\sum_{j=1}^{N_{fp}} \hat{u}^c_{fj,-} \psi_{fj,-} \hat{\mathbf{n}}_{fj,-} \right) \right] \cdot \mathbf{n}\, \mathrm{d}\Gamma_e \\
&+ \int_{\Gamma_e} -\frac{[\![u]\!]_e}{2} \left[\frac{1}{|J_{e,+}|} \mathbf{J}^{-T}_{e,+} \left(\sum_{f=1}^{3}\sum_{j=1}^{N_{fp}} \hat{u}^c_{fj,+} \psi_{fj,+} \hat{\mathbf{n}}_{fj,+} \right) \right] \cdot \mathbf{n}\, \mathrm{d}\Gamma_e,
\end{aligned}
\tag{41}
$$

16

where $\hat{u}^c_{fj} = (\hat{u}^* - \hat{u})_{fj}$, is the transformed solution correction. Using (5) and (8),

$$
\begin{aligned}
\Theta_{e,1} &= \int_{\Gamma_e} -\frac{[\![u]\!]_e}{2}\left[\frac{1}{|J_{e,-}|}\left(\sum_{f=1}^{3}|J^{f,-}|\sum_{j=1}^{N_{fp}}u^c_{fj,-}\psi_{fj,-}\mathbf{n}_{fj,-}\right)\right]\cdot\mathbf{n}\,\mathrm{d}\Gamma_e \\
&+ \int_{\Gamma_e} -\frac{[\![u]\!]_e}{2}\left[\frac{1}{|J_{e,+}|}\left(\sum_{f=1}^{3}|J^{f,+}|\sum_{j=1}^{N_{fp}}u^c_{fj,+}\psi_{fj,+}\mathbf{n}_{fj,+}\right)\right]\cdot\mathbf{n}\,\mathrm{d}\Gamma_e.
\end{aligned}
\tag{42}
$$

Through a substitution of the numerical flux into the solution correction, we obtain $u^c_{fj} = \{\{u\}\}_{fj} - u_{fj} = -\frac{1}{2}[\![u]\!]_{fj}$. Note that while $[\![u]\!]_e$ is a function of (r,s), $[\![u]\!]_{fj}$ is a constant. We also define the parameter F_s which is the ratio of the length of an edge over the determinant of the Jacobian of an element. Since we employ a straight sided triangle, we can drop the subscript j for the normal of an edge. In the following equation, we expand on the summation over the three faces and separate the jump belonging to edge e, $[\![u]\!]_{ej}$ from the other jumps $[\![u]\!]_{fj}$,

$$
\begin{aligned}
\Theta_{e,1} &= \int_{\Gamma_e}\frac{[\![u]\!]_e}{4}\left[F^e_{s,-}\sum_{j=1}^{N_{fp}}[\![u]\!]_{ej}\psi_{ej,-}\right]\mathrm{d}\Gamma_e \\[2mm]
&+ \int_{\Gamma_e}\frac{[\![u]\!]_e}{4}\left[F^e_{s,+}\sum_{j=1}^{N_{fp}}[\![u]\!]_{ej}\psi_{ej,+}\right]\mathrm{d}\Gamma_e \\[2mm]
&+ \int_{\Gamma_e}\frac{[\![u]\!]_e}{4}\left[\sum_{\substack{f=1\\f\neq e}}F^f_{s,-}\sum_{j=1}^{N_{fp}}[\![u]\!]_{fj}\psi_{fj,-}\mathbf{n}_f\right]\cdot\mathbf{n}\,\mathrm{d}\Gamma_e \\[2mm]
&+ \int_{\Gamma_e}\frac{[\![u]\!]_e}{4}\left[\sum_{\substack{f=1\\f\neq e}}F^f_{s,+}\sum_{j=1}^{N_{fp}}[\![u]\!]_{fj}\psi_{fj,+}\mathbf{n}_f\right]\cdot\mathbf{n}\,\mathrm{d}\Gamma_e,
\end{aligned}
\tag{43}
$$

where $\Theta_{e,1}$ can now be expressed as the sum of four terms. The next step is to transform each integral into the computational space, where each term integrates the product $[\![u]\!]_e\psi_{fj}$ over the computational edge $\Gamma_{s,e}$. Since $u_n \in P_p(\Omega_s)$, we have also $u_n \in R_p(\Gamma_s)$. Similarly, $\nabla\cdot\mathbf{g}_{fj} = \psi_{fj} \in P_p(\Omega_s)$ and hence $\psi_{fj} \in R_p(\Gamma_s)$. Therefore the product $[\![u]\!]_e\psi_{fj}$ is a polynomial of degree less than or equal to $2p$. Using Gauss-Legendre quadratures, we can compute exactly the integrals of $\Theta_{e,1}$.

We introduce $\mathbf{r}_i$ and ω_i as the Gauss-Legendre nodes and weights,

$$
\begin{aligned}
\Theta_{e,1} &= |J^e| \frac{F^e_{s,-}}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]^2_{ei}\omega_i\psi_{ei,-}\left(\mathbf{r}_i\right) + [\![u]\!]_{ei}\omega_i \left(\sum_{\substack{j=1\\j\neq i}}^{N_{fp}} [\![u]\!]_{ej}\psi_{ej,-}\left(\mathbf{r}_i\right) \right) \right] \\
&+ |J^e| \frac{F^e_{s,+}}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]^2_{ei}\omega_i\psi_{ei,+}\left(\mathbf{r}_i\right) + [\![u]\!]_{ei}\omega_i \left(\sum_{\substack{j=1\\j\neq i}}^{N_{fp}} [\![u]\!]_{ej}\psi_{ej,+}\left(\mathbf{r}_i\right) \right) \right] \\
&+ |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}\omega_i \left(\sum_{\substack{f=1\\f\neq e}}^{N_{fp}} \frac{F^f_{s,-}\left(\mathbf{n}\cdot\mathbf{n}_{f,-}\right)}{4} \sum_{j=1}^{N_{fp}} [\![u]\!]_{fj,-}\psi_{fj,-}\left(\mathbf{r}_i\right) \right) \right] \\
&+ |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}\omega_i \left(\sum_{\substack{f=1\\f\neq e}}^{N_{fp}} \frac{F^f_{s,+}\left(\mathbf{n}\cdot\mathbf{n}_{f,+}\right)}{4} \sum_{j=1}^{N_{fp}} [\![u]\!]_{fj,+}\psi_{fj,+}\left(\mathbf{r}_i\right) \right) \right],
\end{aligned}
\tag{44}
$$

where the multiplicative scalar $|J^e|$ in front of each term comes from the transformation from the physical to the computational domain. $\qquad\blacksquare$

Having completed the expansion of the first term, $\Theta_{e,1}$, we now turn our attention to the second term $\Theta_{e,2}$ from (39).

Lemma 4.3. *Using Gauss-Legendre quadratures, we can compute $\Theta_{e,2}$ exactly,*

$$
\Theta_{e,2} = -|J^e| \sum_{i=1}^{N_{fp}} \tau_{ei}\omega_i [\![u]\!]^2_{ei}.
\tag{45}
$$

where $(\omega_i)_{i\in[\![1,N_{fp}]\!]}$ and $(\mathbf{r}_i)_{i\in[\![1,N_{fp}]\!]}$ are the Gauss-Legendre quadrature weights and points.

Proof. Since $[\![u]\!]^2_e$ is a polynomial of degree $2p$ on Γ_s, we can evaluate it exactly with Gauss-Legendre quadratures. Thus the second term $\Theta_{e,2}$ can be written as,

$$
\begin{aligned}
\Theta_{e,2} &= -\int_{\Gamma_e} \tau[\![u]\!]^2 \mathrm{d}\Gamma_e, \\
&= -|J^e| \sum_{i=1}^{N_{fp}} \tau_{ei}\omega_i [\![u]\!]^2_{ei}.
\end{aligned}
\tag{46}
$$

$\qquad\blacksquare$

18

Theorem 4.1. *Employing the IP scheme for the diffusion equation with affine triangular meshes, periodic boundary conditions and using the ESFR methods; for all edges e and for all flux points i, τ_{ei} greater than τ_{ei}^* implies the energy stability of the solution, with*

$$
\tau_{ei}^* = \frac{1}{4} \min_\kappa \sum_k \left[F_{s,k}^e \left(\psi_{ei,k}\left(\mathbf{r}_i^e\right) - \left|\psi_{ei,k}\left(\mathbf{r}_i^e\right)\right| \right) \right.
$$
$$
\left. + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \left(\frac{F_{s,k}^f \left|\mathbf{n} \cdot \mathbf{n}_f\right|}{2} \left(\left|\psi_{fj,k}\left(\mathbf{r}_i^e\right)\right| + \frac{\omega_j}{\omega_i} \left|\psi_{ei,k}\left(\mathbf{r}_j^f\right)\right| \right) \right) \right],
\tag{47}
$$

where $k = \{-, +\}$. $k = -$ signifies triangle IJK, and $k = +$ denotes the adjacent triangle IJK_{+1}.

Proof. First, we employ the triangular inequality $\left(ab \leq \frac{1}{2}\left(a^2 + b^2\right)\right)$ on all products of $[\![u]\!]_{ij}[\![u]\!]_{kl}$ of the result of Lemma 4.2,

$$
\begin{aligned}
\Theta_{e,1} \leq\ & |J^e| \frac{F_{s,-}^e}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \omega_i \psi_{ei,-}\left(\mathbf{r}_i\right) + \frac{1}{2} \left(\sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \omega_i \left|\psi_{ej,-}\left(\mathbf{r}_i\right)\right| \left([\![u]\!]_{ei}^2 + [\![u]\!]_{ej}^2\right) \right) \right] \\
& + |J^e| \frac{F_{s,+}^e}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \omega_i \psi_{ei,+}\left(\mathbf{r}_i\right) + \frac{1}{2} \left(\sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \omega_i \left|\psi_{ej,+}\left(\mathbf{r}_i\right)\right| \left([\![u]\!]_{ei}^2 + [\![u]\!]_{ej}^2\right) \right) \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,-}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,-}\right|}{8} \left|\psi_{fj,-}\left(\mathbf{r}_i\right)\right| \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,+}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,+}\right|}{8} \left|\psi_{fj,+}\left(\mathbf{r}_i\right)\right| \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[\omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,-}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,-}\right|}{8} [\![u]\!]_{fj,-}^2 \left|\psi_{fj,-}\left(\mathbf{r}_i\right)\right| \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[\omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,+}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,+}\right|}{8} [\![u]\!]_{fj,+}^2 \left|\psi_{fj,+}\left(\mathbf{r}_i\right)\right| \right].
\end{aligned}
\tag{48}
$$

The terms $\sum_{i=1}^{N_{fp}} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \omega_i \left|\psi_{ej,-}\left(\mathbf{r}_i\right)\right| [\![u]\!]_{ej}^2$ and $\sum_{i=1}^{N_{fp}} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \omega_i \left|\psi_{ej,+}\left(\mathbf{r}_i\right)\right| [\![u]\!]_{ej}^2$ of the first two lines of the

19

previous equation need further derivations.

$$
\begin{aligned}
A &= \sum_{i=1}^{N_{fp}} \left[\sum_{\substack{j=1 \\ j\neq i}}^{N_{fp}} \omega_i \, |\psi_{ej,-}(\mathbf{r}_i)| \, [\![u]\!]_{ej}^2 \right] \\
&= \sum_{i=1}^{N_{fp}} \left[\sum_{j=1}^{N_{fp}} \omega_i \, |\psi_{ej,-}(\mathbf{r}_i)| \, [\![u]\!]_{ej}^2 \right] - \sum_{i=1}^{N_{fp}} \omega_i \, |\psi_{ei,-}(\mathbf{r}_i)| \, [\![u]\!]_{ei}^2 \\
&= \sum_{j=1}^{N_{fp}} \left[[\![u]\!]_{ej}^2 \sum_{i=1}^{N_{fp}} \omega_i \, |\psi_{ej,-}(\mathbf{r}_i)| \right] - \sum_{i=1}^{N_{fp}} \omega_i \, |\psi_{ei,-}(\mathbf{r}_i)| \, [\![u]\!]_{ei}^2 .
\end{aligned}
\tag{49}
$$

On the first term of A, renaming indices j into i and vice versa, we obtain

$$
\begin{aligned}
A &= \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \sum_{j=1}^{N_{fp}} \omega_j \, |\psi_{ei,-}(\mathbf{r}_j)| \right] - \sum_{i=1}^{N_{fp}} \omega_i \, |\psi_{ej,-}(\mathbf{r}_i)| \, [\![u]\!]_{ej}^2 \\
&= \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \left(\sum_{j=1}^{N_{fp}} (\omega_j \, |\psi_{ei,-}(\mathbf{r}_j)|) - \omega_i \, |\psi_{ej,-}(\mathbf{r}_i)| \right) \right] \\
&= \sum_{i=1}^{N_{fp}} [\![u]\!]_{ei}^2 \left(\sum_{\substack{j=1 \\ j\neq i}}^{N_{fp}} \omega_j \, |\psi_{ei,-}(\mathbf{r}_j)| \right) .
\end{aligned}
\tag{50}
$$

Similar derivations are conducted for $\displaystyle\sum_{i=1}^{N_{fp}}\sum_{\substack{j=1 \\ j\neq i}}^{N_{fp}} \omega_i \, |\psi_{ej,+}(\mathbf{r}_i)| \, [\![u]\!]_{ej}^2$. Including these terms into equa-

tion (48) yields,

$$
\begin{aligned}
\Theta_{e,1} \quad \leq \quad & |J^e| \frac{F_{s,-}^e}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \left(\omega_i \psi_{ei,-}(\mathbf{r}_i) + \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \omega_i \, |\psi_{ej,-}(\mathbf{r}_i)| + \omega_j \, |\psi_{ei,-}(\mathbf{r}_j)| \right) \right] \\
& + |J^e| \frac{F_{s,+}^e}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \left(\omega_i \psi_{ei,+}(\mathbf{r}_i) + \frac{1}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \omega_i \, |\psi_{ej,+}(\mathbf{r}_i)| + \omega_j \, |\psi_{ei,+}(\mathbf{r}_j)| \right) \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,-}^f \, |\mathbf{n} \cdot \mathbf{n}_{f,-}|}{8} \, |\psi_{fj,-}(\mathbf{r}_i)| \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[[\![u]\!]_{ei}^2 \omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,+}^f \, |\mathbf{n} \cdot \mathbf{n}_{f,+}|}{8} \, |\psi_{fj,+}(\mathbf{r}_i)| \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[\omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,-}^f \, |\mathbf{n} \cdot \mathbf{n}_{f,-}|}{8} [\![u]\!]_{fj,-}^2 \, |\psi_{fj,-}(\mathbf{r}_i)| \right] \\
& + |J^e| \sum_{i=1}^{N_{fp}} \left[\omega_i \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \frac{F_{s,+}^f \, |\mathbf{n} \cdot \mathbf{n}_{f,+}|}{8} [\![u]\!]_{fj,+}^2 \, |\psi_{fj,+}(\mathbf{r}_i)| \right] .
\end{aligned}
\tag{51}
$$

Notice that $\Theta_{e,1}$ is only associated to edge e_{IJ} of Θ_{dif}. However it contains the jump squared of the solution on the four faces of the two neighboring triangles that share edge, e_{IJ}.

Figure 5: Representation of the influence of $\Theta_{e,1}$ on the other edges ; the red arrow represents the jump of the solution on the edge e while the blue arrows are the jumps on the neighboring edges; the green squares represent the flux points; in this example $p = 2$.

At this juncture we have only computed the value of $\Theta_{e,1}$ over the edge e_{IJ}. To simplify the final expression further, we require all the contributions to the jump over edge e. From Figure 5 and equation (51), we identify four additional contributions to the jump $[\![u]\!]_e^2$. These contributions are a result of the expansion of equation (51) of other edges that contain a contribution towards the edge e_{IJ}.

- $\Theta_{e_{IK}}$ has a contribution of $|J^e| \sum\limits_{i=1}^{N_{fp}} [\![u]\!]_{ei}^2 \left(\dfrac{F_{s,-}^{IK} |\mathbf{n}_{IJ} \cdot \mathbf{n}_{IK}|}{8} \sum\limits_{j=1}^{N_{fp}} \omega_j \left| \psi_{ei,-} \left(\mathbf{r}_j^{IK} \right) \right| \right).$

- $\Theta_{e_{KJ}}$ has a contribution of $|J^e| \sum\limits_{i=1}^{N_{fp}} [\![u]\!]_{ei}^2 \left(\dfrac{F_{s,-}^{KJ} |\mathbf{n}_{KJ} \cdot \mathbf{n}_{IK}|}{8} \sum\limits_{j=1}^{N_{fp}} \omega_j \left| \psi_{ei,-} \left(\mathbf{r}_j^{KJ} \right) \right| \right).$

- $\Theta_{e_{JK_{+1}}}$ has a contribution of $|J^e| \sum\limits_{i=1}^{N_{fp}} [\![u]\!]_{ei}^2 \left(\dfrac{F_{s,+}^{JK_{+1}} |\mathbf{n}_{JK_{+1}} \cdot \mathbf{n}_{JK_{+1}}|}{8} \sum\limits_{j=1}^{N_{fp}} \omega_j \left| \psi_{ei,+} \left(\mathbf{r}_j^{JK_{+1}} \right) \right| \right).$

- $\Theta_{e_{IK_{+1}}}$ has a contribution of $|J^e| \sum\limits_{i=1}^{N_{fp}} [\![u]\!]_{ei}^2 \left(\dfrac{F_{s,+}^{IK_{+1}} |\mathbf{n}_{IK_{+1}} \cdot \mathbf{n}_{IK_{+1}}|}{8} \sum\limits_{j=1}^{N_{fp}} \omega_j \left| \psi_{ei,+} \left(\mathbf{r}_j^{IK_{+1}} \right) \right| \right).$

Note that while in (51), the correction fields were computed at the flux points of the edge itself, the correction fields of the above terms are calculated at the flux points of the surrounding edges

of e_{IJ}. We indicate the flux point on the corresponding edges through the superscripts, i.e. IK, on the node location, $\mathbf{r}$.

If we sum all the edges, we finally attain

$$
\begin{aligned}
\Theta_{dif,1} &= \sum_{e=1}^{N_e} \Theta_{e,1} \\
&\leq \frac{1}{4} \sum_{e=1}^{N_e} \Bigg[|J^e| \sum_{i=1}^{N_{fp}} \Bigg[\llbracket u \rrbracket_{ei}^2 \Bigg(F_{s,-}^e \omega_i \underbrace{\psi_{ei,-}\left(\mathbf{r}_i^e\right)}_{\text{Term A,-}} + F_{s,+}^e \omega_i \psi_{ei,+}\left(\mathbf{r}_i^e\right) \\
&+ \frac{F_{s,-}^e}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \Bigg(\omega_i \underbrace{\left|\psi_{ej,-}\left(\mathbf{r}_i^e\right)\right|}_{\text{Term B,-}} + \omega_j \underbrace{\left|\psi_{ei,-}\left(\mathbf{r}_j^e\right)\right|}_{\text{Term C,-}} \Bigg) + \frac{F_{s,+}^e}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \left(\omega_i \left|\psi_{ej,+}\left(\mathbf{r}_i^e\right)\right| + \omega_j \left|\psi_{ei,+}\left(\mathbf{r}_j^e\right)\right| \right) \\
&+ \omega_i \Bigg[\sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \Bigg(\frac{F_{s,-}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,-}\right|}{2} \underbrace{\left|\psi_{fj,-}\left(\mathbf{r}_i^e\right)\right|}_{\text{Term D,-}} \Bigg) + \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \left(\frac{F_{s,+}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,+}\right|}{2} \left|\psi_{fj,+}\left(\mathbf{r}_i^e\right)\right| \right) \Bigg] \\
&+ \sum_{\substack{f=1 \\ f \neq e}}^{3} \Bigg[\frac{F_{s,-}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,-}\right|}{2} \sum_{j=1}^{N_{fp}} \Bigg(\omega_j \underbrace{\left|\psi_{ei,-}\left(\mathbf{r}_j^f\right)\right|}_{\text{Term E,-}} \Bigg) \Bigg] + \sum_{\substack{f=1 \\ f \neq e}}^{3} \Bigg[\frac{F_{s,+}^f \left|\mathbf{n} \cdot \mathbf{n}_{f,+}\right|}{2} \sum_{j=1}^{N_{fp}} \left(\omega_j \left|\psi_{ei,+}\left(\mathbf{r}_j^f\right)\right| \right) \Bigg] \Bigg) \Bigg] \Bigg].
\end{aligned}
\tag{52}
$$

For the jump of the flux point i on the edge e, $\llbracket u \rrbracket_{ei}^2$, we have the influence of all the correction fields surrounding e evaluated at the flux point $\mathbf{r}_i^e$ (Terms B and D). In addition, we have the influence of the correction field ψ_{ei} evaluated at all the flux points (Terms A, C and E). Although this last equation may seem complicated, it is quite logical. These individual contributions are best understood through a graphical means as depicted in Figure 6. For simplicity, we just represent the terms arising from triangle IJK, since we have a symmetrical influence from triangle IJK_{+1}.

Figure 6: Influence of the correction fields, for the left side of the edge (Triangle IJK), on the flux point $(f, j) = (1, 2)$ for $p = 2$.

We now return to equation (39) and provide a complete expansion for the diffusion terms toward stability. We combine equations (45) and (52), and introduce a simplification in our notation where the parameter $k = \{-, +\}$; where $k = -$ signifies triangle IJK, while $k = +$ denotes triangle IJK_{+1}. The first term in equation (52) is simplified into,

$$
\begin{aligned}
\Theta_{dif,1} \quad \leq \quad & \sum_{e=1}^{N_e} \Bigg[|J^e| \sum_{i=1}^{N_{fp}} \Bigg[\frac{\llbracket u \rrbracket_{ei}^2}{4} \sum_k \Bigg[F_{s,k}^e \omega_i \psi_{ei,k} \left(\mathbf{r}_i^e \right) + \frac{F_{s,k}^e}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \left(\omega_i \left| \psi_{ej,k} \left(\mathbf{r}_i^e \right) \right| + \omega_j \left| \psi_{ei,k} \left(\mathbf{r}_j^e \right) \right| \right) \\
& + \sum_{\substack{f=1 \\ f \neq e}}^{3} \sum_{j=1}^{N_{fp}} \left(\frac{F_{s,k}^f \left| \mathbf{n} \cdot \mathbf{n}_{f,k} \right|}{2} \left(\omega_i \left| \psi_{fj,k} \left(\mathbf{r}_i^e \right) \right| + \omega_j \left| \psi_{ei,k} \left(\mathbf{r}_j^f \right) \right| \right) \right) \Bigg] \Bigg] \Bigg],
\end{aligned}
$$
(53)

and using Lemma 4.3, we compute Θ_{dif},

$$
\begin{aligned}
\Theta_{dif} &= \sum_{e=1}^{N_e} \left[\Theta_{e,1} + \Theta_{e,2}\right] \\
&\leq \sum_{e=1}^{N_e} |J^e| \sum_{i=1}^{N_{fp}} \left[\llbracket u \rrbracket_{ei}^2 \left[-\tau_{ei}\omega_i + \frac{1}{4}\sum_k \left[F_{s,k}^e \omega_i \psi_{ei,k}\left(\mathbf{r}_i^e\right) \right. \right. \right. \\
&\quad + \frac{F_{s,k}^e}{2} \sum_{\substack{j=1 \\ j\neq i}}^{N_{fp}} \left(\omega_i \left|\psi_{ej,k}\left(\mathbf{r}_i^e\right)\right| + \omega_j \left|\psi_{ei,k}\left(\mathbf{r}_j^e\right)\right| \right) \\
&\quad \left. \left. \left. + \sum_{\substack{f=1 \\ f\neq e}}^{3}\sum_{j=1}^{N_{fp}} \left(\frac{F_{s,k}^f \left|\mathbf{n}\cdot\mathbf{n}_{f,k}\right|}{2} \left(\omega_i \left|\psi_{fj,k}\left(\mathbf{r}_i^e\right)\right| + \omega_j \left|\psi_{ei,k}\left(\mathbf{r}_j^f\right)\right| \right) \right) \right] \right] \right].
\end{aligned}
\tag{54}
$$

To evaluate the limiting value, τ^* to ensure stability, we remove the exclusion of $f \neq e$ and similarly for the second to last term, where we remove the exclusion of $j \neq i$. This results in an additional term $\left|\psi_{ei,k}\left(\mathbf{r}_i^e\right)\right|$ that is subtracted from the first term. We thus ensure energy stability for the diffusion equation in two dimensions for triangles. The IP scheme requires $\tau_{e,i} \geq \tau_{e,i}^*$ $\forall e\, \forall i$ where τ^* is defined as

$$
\begin{aligned}
\tau_{ei}^* &= \frac{1}{4}\sum_k \left[F_{s,k}^e \left(\psi_{ei,k}\left(\mathbf{r}_i^e\right) - \left|\psi_{ei,k}\left(\mathbf{r}_i^e\right)\right| \right) \right. \\
&\quad \left. + \sum_{f=1}^{3}\sum_{j=1}^{N_{fp}} \left(\frac{F_{s,k}^f \left|\mathbf{n}\cdot\mathbf{n}_f\right|}{2} \left(\left|\psi_{fj,k}\left(\mathbf{r}_i^e\right)\right| + \frac{\omega_j}{\omega_i} \left|\psi_{ei,k}\left(\mathbf{r}_j^f\right)\right| \right) \right) \right].
\end{aligned}
\tag{55}
$$

From Postulate 3.2 the ESFR high-order method with either the IP or BR2 schemes for the diffusion problem is independent of κ. As a consequence, the energy stability is also independent of κ. Thus we finally obtain,,

$$
\begin{aligned}
\tau_{ei}^* &= \frac{1}{4}\min_\kappa \sum_k \left[F_{s,k}^e \left(\psi_{ei,k}\left(\mathbf{r}_i^e\right) - \left|\psi_{ei,k}\left(\mathbf{r}_i^e\right)\right| \right) \right. \\
&\quad \left. + \sum_{f=1}^{3}\sum_{j=1}^{N_{fp}} \left(\frac{F_{s,k}^f \left|\mathbf{n}\cdot\mathbf{n}_f\right|}{2} \left(\left|\psi_{fj,k}\left(\mathbf{r}_i^e\right)\right| + \frac{\omega_j}{\omega_i} \left|\psi_{ei,k}\left(\mathbf{r}_j^f\right)\right| \right) \right) \right].
\end{aligned}
\tag{56}
$$

$\square$

Remark 4.2. We could simplify (47) further by noting that $\left|\mathbf{n}\cdot\mathbf{n}_f\right| \leq 1$, and that the quantity $F_{s,max} = \max\left(F_s\right)$. However such simplification would result in a less accurate condition and hence a larger value of τ^*. We refrain from this simplification as increasing τ results in an increase in Δt_{max} [15].

25

Remark 4.3. In the following graphs, we represent the maximum value of the array of τ^*_{theory} evaluated via equation (55). These numerical simulations show that the minimum of τ^*_{ei} is obtained when κ approaches κ_+. In the rest of the article, we will compute τ^*_{ei} in equation (47) with $\kappa = \kappa_+$ for all ESFR schemes.

(a) $\max\left(\tau^*_{theory}\right)$ for $p = 2$.

(b) $\max\left(\tau^*_{theory}\right)$ for $p = 3$.

(c) $\max\left(\tau^*_{theory}\right)$ for $p = 2$ zoom in.

(d) $\max\left(\tau^*_{theory}\right)$ for $p = 3$ zoom in.

Figure 7: Minimization of $\max\left(\tau^*_{theory}\right)$ along κ for different p where the diamond corresponds to $\kappa = \kappa_+$.

26

These numerical simulations are obtained for the mesh represented in Figure 8 with 128 elements.

4.2. Numerical results

The purpose of the following numerical simulation is to find the minimal numerical penalty term $\tau^*_{numerical}$ which guarantees stability for a defined problem. This numerical penalty term will be compared with (47).

The problem solved is of the following: Find $u(x, y, t)$ such that

$$\begin{cases} \dfrac{\partial u}{\partial t} & = & b\Delta u, & \text{for } x \times y \in [-1; 1]^2 \text{ and } t \in [0, 2], \\ u(x, y, 0) & = & \sin(\pi x)\sin(\pi y). \end{cases} \tag{57}$$

We impose periodic boundary conditions and the exact solution for this system is $u_{exact} = e^{-2b\pi^2 t}\sin(\pi x)\sin(\pi y)$.

We take $b = 0.1$ and use the fourth order five stage Runge-Kutta as the time integrator [16]. The flux points are taken as the Gauss-Legendre nodes while the solution points are taken as the α-optimised nodes [1]. The mesh generated is regular as shown in Figure 8.

Figure 8: Representation of a regular mesh $8 \times 8 \times 2$.

The time step is taken as advised by Hesthaven and Warburton [1],

$$\Delta t = CFL \left(\frac{2}{3} \min \Delta r_i^2 \right) \min_\Omega \left(\frac{r_D}{|b|} \right), \tag{58}$$

27

where $CFL = 10^{-2}$ in our cases, Δr_i is a measure of the distance between the solution points and r_D is the ratio between the perimeter of the triangle with its area.

While in our proof, τ^*_{theory} is an array, we consider here, for simplicity, that τ is a constant for all the edges and flux points. We apply the same procedure as [15]: choose a criterion of stability on the upper bound of the solution, $|u(x,t)| \leq u_{max}$. As, for all t, $|u_{exact}(x,t)| \leq 1$, choosing $u_{max} = 2$ is sufficient and allow some instabilities arising from the scheme. Then select a sufficiently low value for τ_0 as a starting point to ensure an unstable solution and increase it by dτ of 0.1 until the solution is stable. Taking the final time equal to 2 ensures that a few thousands iterations are run. The results are presented for four methods: c_{DG}/κ_{DG}, c_{DG}/κ_+, c_+/κ_{DG} and c_+/κ_+ where $c_+ = \kappa_+$ is taken to be equal to the last column of Table 1.

p	2		3	
κ c	κ_{DG}	κ_+	κ_{DG}	κ_+
$\max\left(\tau^*_{theory}\right)$	34.4	34.4	64.4	64.4
c_{DG}	19.3	19.3	37.0	37.0
c_+	15.3	15.3	36.4	36.4

Table 2: $\tau^*_{numerical}$ for the IP scheme for p=2 and p=3 for a $8 \times 8 \times 2$ mesh.

p	2		3	
κ c	κ_{DG}	κ_+	κ_{DG}	κ_+
$\max\left(\tau^*_{theory}\right)$	64.8	64.8	129.9	129.9
c_{DG}	41.7	41.7	76.1	76.1
c_+	39.1	39.2	75.7	75.7

Table 3: $\tau^*_{numerical}$ for the IP scheme for p=2 and p=3 for a $16 \times 16 \times 2$ mesh.

For both tables 2 and 3, the maximum of the criterion found in (47) is greater than $\tau^*_{numerical}$. These results do not validate (47) for every edge and every flux point as $\tau_{numerical}$ is taken to be a constant. However since $\max\left(\tau^*_{theory}\right) \geq \tau^*_{numerical}$, this lead us to believe that the criterion (47) is valid.

5. BR2 stability condition

5.1. Theoretical result

For the BR2 scheme, the steps are identical to that of the IP scheme with the exception of $\mathbf{q}^*$. Therefore Lemma 4.1 is modified into

$$\Theta_{dif} = \sum_{e=1}^{N_e} \left[\underbrace{\int_{\Gamma_e} \left(-\{\{\mathbf{q} - \nabla u\}\} \right) \cdot [\![\mathbf{u}]\!]_e \, d\Gamma_e}_{\Theta_{e,1}} + \underbrace{\int_{\Gamma_e} s \left\{\{\mathbf{r^e}\left([\![\mathbf{u}]\!]\right)\}\right\} \cdot [\![\mathbf{u}]\!]_e \, d\Gamma_e}_{\Theta_{e,2}} \right]. \tag{59}$$

The first term $\Theta_{e,1}$ is the same as the IP scheme and hence $\Theta_{dif,1} = \sum_{e=1}^{N_e} \Theta_{e,1}$ follows the same derivations up to (53); however, the second term requires additional derivations. Similar to the 1D case [15], we expand $\{\{\mathbf{r^e}\left([\![\mathbf{u}]\!]\right)\}\}$ in terms of $[\![\mathbf{u}]\!]_e$. We base our analysis on the article of Huynh [17], where he presented an equivalence between the DG and the FR formulations for the advection equation. This equivalence has been presented for all high dimensions and for curvilinear elements [18]; it is, once again, presented in the following for completeness of the paper. The advection equation is similar to (9a) with $\mathbf{f} = -u\mathbf{a}$, where $\mathbf{a}$ is the velocity. The FR approach, for the advection problem, is similar to (12a),

$$\frac{\partial \hat{u}_n}{\partial t} = \hat{\nabla} \cdot \hat{\mathbf{f}}_n \left(\hat{u}_n \right) + \hat{\nabla} \cdot \sum_{f=1}^{3} \sum_{j=1}^{Nfp} \left[\left(\hat{\mathbf{f}}_{n,fj}^* - \hat{\mathbf{f}}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \mathbf{h}_{fj}\left(\mathbf{r}\right). \tag{60}$$

However the DG formulation, in the strong form, as given in [1], for an element Ω_n is

$$\int_{\Omega_s} \frac{\partial \hat{u}_n}{\partial t} \Phi \, d\Omega_s = \int_{\Omega_s} \hat{\nabla} \cdot \hat{\mathbf{f}}_n \left(u_n \right) \Phi \, d\Omega_s + \int_{\Gamma_s} \left[\left(\hat{\mathbf{f}}_n^* - \hat{\mathbf{f}}_n \right) \cdot \hat{\mathbf{n}} \right] \Phi \, d\Gamma_s, \tag{61}$$

with Φ as the test function. To retrieve the FR formulation from equation (61), we must define, for a face f, the correction field δ_f such that

$$\int_{\Omega_s} \delta_f \Phi \, d\Omega_s = \int_{\Gamma_{s,f}} \left[\left(\hat{\mathbf{f}}_{n,f}^* - \hat{\mathbf{f}}_{n,f} \right) \cdot \hat{\mathbf{n}} \right] \Phi \, d\Gamma_{s,f}, \tag{62}$$

where δ_f is in fact a lifting operator associated to the face f, $\Gamma_{s,f}$, of Ω_s. We evaluate exactly the integral over $\Gamma_{s,f}$ employing Gauss-Legendre quadratures, (nodes $\mathbf{r}_j^f$, weight ω_j),

$$\int_{\Omega_s} \delta_f \Phi \, d\Omega_s = \sum_{j=1}^{N_{fp}} \left[\left(\hat{\mathbf{f}}_f^* - \hat{\mathbf{f}}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{n,fj} \right]\Big|_{\mathbf{r}_j^f} \omega_j \Phi\left(\mathbf{r}_j^f\right). \tag{63}$$

29

Upon removing the corrective flux, we must define δ_{fj}^* such that

$$\int_{\Omega_s} \delta_{fj}^* \Phi \, d\Omega_s = \omega_j \Phi \left(\mathbf{r}_j^f \right). \tag{64}$$

From (61), let us decompose Γ_s into $\sum_{f=1}^{3} \Gamma_{s,f}$,

$$
\begin{aligned}
\int_{\Omega_s} \frac{\partial \hat{u}_n}{\partial t} \Phi \, dr &= \int_{\Omega_s} \hat{\nabla} \cdot \hat{\mathbf{f}}_n \left(\hat{u}_n \right) \Phi \, d\Omega_s + \sum_{f=1}^{3} \int_{\Gamma_{s,f}} \left[\left(\hat{\mathbf{f}}_n^* - \hat{\mathbf{f}}_n \right) \cdot \hat{\mathbf{n}} \right] \Phi \, d\Gamma_{s,f}, \\
&= \int_{\Omega_s} \nabla \cdot \hat{\mathbf{f}}_n \left(\hat{u}_n \right) \Phi \, d\Omega_s + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \left[\left(\hat{\mathbf{f}}_{n,fj}^* - \hat{\mathbf{f}}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \int_{\Omega_s} \delta_{fj}^* \Phi \, d\Omega_s \\
0 &= \int_{\Omega_s} \left[\left(-\frac{\partial \hat{u}_n}{\partial t} + \hat{\nabla} \cdot \hat{\mathbf{f}}_n \left(\hat{u}_n \right) + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \left[\left(\hat{\mathbf{f}}_{n,fj}^* - \hat{\mathbf{f}}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \delta_{fj}^* \right) \Phi \right] d\Omega_s.
\end{aligned}
\tag{65}
$$

Using the lifting operator enables the test function Φ to be factored out and removed from the integral over the different terms of the PDE, resulting into a differential formulation: the FR approach,

$$\frac{\partial \hat{u}_n}{\partial t} = \nabla \cdot \hat{\mathbf{f}}_n \left(\hat{u}_n \right) + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \left[\left(\hat{\mathbf{f}}_{n,fj}^* - \hat{\mathbf{f}}_{n,fj} \right) \cdot \hat{\mathbf{n}}_{fj} \right] \delta_{fj}^*. \tag{66}$$

Lemma 5.1. *Employing the ESFR correction fields in equation* (17) *to define* δ_{fj}^*, *there is a unique value of c such that equation* (64) *is valid and this value is* $c = 0$.

Proof. Replacing L_i by Φ, a test function, in equation (17), we obtain the following,

$$
\begin{aligned}
& \int_{\Omega_s} \mathbf{h}_{fj} \cdot \hat{\nabla} \Phi \, d\Omega_s = c \sum_{m=1}^{p+1} \binom{p}{m-1} \left(D^{(m,p)} \Phi \right) \left(D^{(m,p)} \phi_{fj} \right) \\
\Leftrightarrow \ & \int_{\Omega_s} \hat{\nabla} \left(\mathbf{h}_{fj} \Phi \right) d\Omega_s - \int_{\Omega_s} \Phi \hat{\nabla} \cdot \mathbf{h}_{fj} \, d\Omega_s = c \sum_{m=1}^{p+1} \binom{p}{m-1} \left(D^{(m,p)} \Phi \right) \left(D^{(m,p)} \phi_{fj} \right) \\
\Leftrightarrow \ & \int_{\Omega_s} \phi_{fj} \Phi \, d\Omega_s = \int_{\Gamma_s} \Phi \left(\mathbf{h}_{fj} \cdot \hat{\mathbf{n}} \right) d\Gamma_s - c \sum_{m=1}^{p+1} \binom{p}{m-1} \binom{p}{m-1} \left(D^{(m,p)} \Phi \right) \left(D^{(m,p)} \phi_{fj} \right).
\end{aligned}
\tag{67}
$$

Since $\mathbf{h}_{fj} \cdot \hat{\mathbf{n}} \in R_p \left(\Gamma_s \right)$, we have $\Phi \left(\mathbf{h}_{fj} \cdot \hat{\mathbf{n}} \right)$ a polynomial of degree $2p$ on the edge. We use Gauss-Legendre quadrature (N_{fp} points, same quadrature as (64)) to compute the integral of the right-hand side. Moreover $\mathbf{h}_{fj}$ verifies (13), hence we obtain

$$\int_{\Omega_s} \phi_{fj} \Phi \, d\Omega_s = \omega_j \Phi \left(\mathbf{r}_j^f \right) + c \sum_{m=1}^{p+1} \binom{p}{m-1} \left(D^{(m,p)} \Phi \right) \left(D^{(m,p)} \phi_{fj} \right). \tag{68}$$

30

Therefore,

$$
\int_{\Omega_s} \phi_{fj} \Phi \, \mathrm{d}x = \omega_j \Phi \left(\mathbf{r}_j^f \right)
$$
$$
\Leftrightarrow \quad c \sum_{m=1}^{p+1} \binom{p}{m-1} \left(D^{(m,p)} \Phi \right) \left(D^{(m,p)} \phi_{fj} \right) = 0 \tag{69}
$$

Taking, for instance, $\Phi = \phi_{fj}$, the previous equation is valid if $c = 0$. $\square$

Remark 5.1. With $c = 0$, the FR formulation is equivalent to the DG formulation. Hence the associated correction function is denoted ϕ_{fj}^{DG}.

We now apply the analogy between the correction field, ϕ_{fj}^{DG} and the lifting operator $\mathbf{r^e}$. The support of $\mathbf{r^e}$ is the union of the element of the triangles forming the edge e. On each one of the elements, $\mathbf{r^e}$ is a polynomial of degree p.

We define the space $\Omega_e = \Omega_- \bigcup \Omega_+$. Where Ω_- (resp. Ω_+) is the interior (resp. exterior) element of edge e. Referring to Figure 4, $\Omega_- = \Omega_{IJK} = \Omega_n$. From the affine mapping defined in (3), we define the quantities,

$$
\mathbf{v}_n^a = -\frac{1}{2} \left(\mathbf{v}_{1,n} - \mathbf{v}_{2,n} \right), \tag{70}
$$

$$
\mathbf{v}_n^b = -\frac{\sqrt{3}}{6} \left(\mathbf{v}_{1,n} \mathbf{v}_{2,n} - 2\mathbf{v}_{3,n} \right), \tag{71}
$$

$$
\bar{\mathbf{v}}_n = \frac{1}{3} \left(\mathbf{v}_{1,n} + \mathbf{v}_{2,n} + \mathbf{v}_{3,n} \right), \tag{72}
$$

$$
|\mathbf{V}_n| = Det \begin{pmatrix} v_{n,x}^a & v_{n,x}^b \\ v_{n,y}^a & v_{n,y}^b \end{pmatrix}, \tag{73}
$$

where $v_{n,x}^a$ is the $x-$component of $\mathbf{v}_n^a$. We then define the surjection

$$
\begin{aligned}
\mathcal{M}_e^{-1}: \quad \Omega_e \quad &\to \quad \Omega_s \\
(x,y) \quad &\mapsto \quad \frac{1}{|\mathbf{V}_-|} \begin{pmatrix} (x - \bar{v}_{-,x}) v_{-,y}^b - (y - \bar{v}_{-,y}) v_{-,x}^b \\ -(x - \bar{v}_{-,x}) v_{-,x}^a + (y - \bar{v}_{-,y}) v_{-,x}^a \end{pmatrix} \chi|_- (x,y) \\
&\quad + \frac{1}{|\mathbf{V}_+|} \begin{pmatrix} (x - \bar{v}_{+,x}) v_{+,y}^b - (y - \bar{v}_{+,y}) v_{+,x}^b \\ -(x - \bar{v}_{+,x}) v_{+,x}^a + (y - \bar{v}_{+,y}) v_{+,x}^a \end{pmatrix} \chi|_+ (x,y)
\end{aligned} \tag{74}
$$

where $\chi|_{\Omega_i} (x,y)$ is equal to 1 if $(x,y) \in \Omega_i$, or 0 if $(x,y) \notin \Omega_i$.

Theorem 5.1. *The lifting operator, $\mathbf{r^e}$, employed in the BR2 scheme, is a linear combination of the correction field associated to the DG method and is equivalent to the following formula,*

$$\mathbf{r^e}\left(\llbracket\mathbf{u}\rrbracket\right)(\mathbf{x}) = -\left(\sum_{j=1}^{N_{fp}}\llbracket\mathbf{u}\rrbracket_{ej}\left[\frac{F^e_{s,-}}{2}\phi^{DG}_{ej,-}\left(\mathcal{M}_e^{-1}(\mathbf{x})\right)\chi|_-(\mathbf{x}) + \frac{F^e_{s,+}}{2}\phi^{DG}_{ej,+}\left(\mathcal{M}_e^{-1}(\mathbf{x})\right)\chi|_+(\mathbf{x})\right]\right),$$

(75)

where subscript $-$ (resp. $+$) denotes the interior (resp. exterior) element.

Proof. From the definition of $\mathbf{r^e}$ in (22), we have

$$
\begin{aligned}
\int_\Omega \mathbf{r^e}\left(\llbracket\mathbf{u}\rrbracket\right)\cdot\boldsymbol{\Phi}\,\mathrm{d}\Omega &= \int_{\Omega_e}\mathbf{r^e}\left(\llbracket\mathbf{u}\rrbracket\right)\cdot\boldsymbol{\Phi}\,\mathrm{d}\Omega_e \\
&= -\int_{\Gamma_e}\llbracket\mathbf{u}\rrbracket\cdot\{\{\boldsymbol{\Phi}\}\}\,\mathrm{d}\Gamma_e \\
&= -\frac{1}{2}\left[\int_{\Gamma_e}\llbracket\mathbf{u}\rrbracket\cdot\boldsymbol{\Phi}|_-\,\mathrm{d}\Gamma_e + \int_{\Gamma_e}\llbracket\mathbf{u}\rrbracket\cdot\boldsymbol{\Phi}|_+\,\mathrm{d}\Gamma_e\right] \\
&= -\frac{|J^e|}{2}\left[\int_{\Gamma_{s,e}}\llbracket\mathbf{u}\rrbracket\cdot\boldsymbol{\Phi}|_-\,\mathrm{d}\Gamma_{s,e} + \int_{\Gamma_{s,e}}\llbracket\mathbf{u}\rrbracket\cdot\boldsymbol{\Phi}|_+\,\mathrm{d}\Gamma_{s,e}\right].
\end{aligned}
$$

(76)

Both of these terms are integrals of a polynomial of degree less or equal to $2p$. Using Gauss-Legendre quadratures $((\mathbf{r}_i)_{i\in\llbracket 1,N_{fp}\rrbracket}$ represents the nodes and $(\omega_i)_{i\in\llbracket 1,N_{fp}\rrbracket}$ represents the weights) we have,

$$\int_{\Omega_e}\mathbf{r^e}\left(\llbracket\mathbf{u}\rrbracket\right)\cdot\boldsymbol{\Phi}\,\mathrm{d}\Omega_e = -\frac{1}{2}|J^e|\left[\sum_{j=1}^{N_{fp}}\llbracket\mathbf{u}\rrbracket_{ej}\cdot\left(\boldsymbol{\Phi}|_-\left(\mathbf{r}^e_j\right)\omega_j + \boldsymbol{\Phi}|_+\left(\mathbf{r}^e_j\right)\omega_j\right)\right].$$

(77)

Applying the result of Lemma 5.1, we obtain

$$
\begin{aligned}
\int_{\Omega_e}\mathbf{r^e}\left(\llbracket\mathbf{u}\rrbracket\right)\cdot\boldsymbol{\Phi}\,\mathrm{d}\Omega_e &= -\frac{1}{2}|J^e|\left[\sum_{j=1}^{N_{fp}}\llbracket\mathbf{u}\rrbracket_{ej}\cdot\left(\int_{\Omega_s}\phi^{DG}_{ej,-}\,\boldsymbol{\Phi}|_-\,\mathrm{d}\Omega_s + \int_{\Omega_s}\phi^{DG}_{ej,+}\,\boldsymbol{\Phi}|_+\,\mathrm{d}\Omega_s\right)\right] \\
&= -\sum_{j=1}^{N_{fp}}\llbracket\mathbf{u}\rrbracket_{ej}\cdot\left(\int_{\Omega_-}\left(\frac{F^e_{s,-}}{2}\phi^{DG}_{ej,-}\,\boldsymbol{\Phi}|_-\right)\mathrm{d}\Omega_- + \int_{\Omega_+}\left(\frac{F^e_{s,+}}{2}\phi^{DG}_{ej,+}\,\boldsymbol{\Phi}|_+\right)\mathrm{d}\Omega_+\right) \\
&= -\int_{\Omega_e}\left[\sum_{j=1}^{N_{fp}}\left(\frac{F^e_{s,-}}{2}\phi^{DG}_{ej,-}\llbracket\mathbf{u}\rrbracket_{ej}\cdot\boldsymbol{\Phi}\,\chi|_- + \frac{F^e_{s,+}}{2}\phi^{DG}_{ej,+}\llbracket\mathbf{u}\rrbracket_{ej}\cdot\boldsymbol{\Phi}\,\chi|_+\right)\right]\mathrm{d}\Omega_e \\
&= -\int_{\Omega_e}\left(\sum_{j=1}^{N_{fp}}\llbracket\mathbf{u}\rrbracket_{ej}\left[\frac{F^e_{s,-}}{2}\phi^{DG}_{ej,-}\left(\mathcal{M}_e^{-1}(\mathbf{x})\right)\chi|_-(\mathbf{x}) + \frac{F^e_{s,+}}{2}\phi^{DG}_{ej,+}\left(\mathcal{M}_e^{-1}(\mathbf{x})\right)\chi|_+(\mathbf{x})\right]\right)\cdot\boldsymbol{\Phi}\,\mathrm{d}\Omega_e
\end{aligned}
$$

(78)

Shifting the right-hand side of the previous equation into the left, we can gather the various terms under the integrals and factor out the test function $\boldsymbol{\Phi}$. We finally obtain expression (75). $\square$

Corollary 5.1. *Employing Gauss-Legendre quadratures, the term $\Theta_{e,2}$ defined in equation* (59) *can be computed exactly as,*

$$\Theta_{e,2} = -|J^e| \sum_{i=1}^{N_{fp}} [\![u]\!]_{ei} s\omega_i \left[\sum_{j=1}^{N_{fp}} [\![u]\!]_{ej} \left(\frac{F_{s,-}^e}{4} \psi_{ej,-}^{DG}(\mathbf{r}_i^e) + \frac{F_{s,+}^e}{4} \psi_{ej,+}^{DG}(\mathbf{r}_i^e) \right) \right]. \tag{79}$$

Proof. We apply the result of Theorem 5.1 on equation (59) and then transform the integral from the physical domain to the computational domain,

$$
\begin{aligned}
\Theta_{e,2} &= \int_{\Gamma_e} s\{\{\mathbf{r^e}([\![\mathbf{u}]\!])\}\} \cdot [\![\mathbf{u}]\!] d\Gamma_e \\
&= -|J^e| \left[\int_{\Gamma_{s,e}} \left(s[\![\mathbf{u}]\!] \cdot \left(\sum_{j=1}^{N_{fp}} [\![\mathbf{u}]\!]_{ej} \frac{F_{s,-}^e}{4} \phi_{ej,-}^{DG}(\mathbf{r}) \right) \right) d\Gamma_{s,e} \right. \\
&\quad \left. + \int_{\Gamma_{s,e}} \left(s[\![\mathbf{u}]\!] \cdot \left(\sum_{j=1}^{N_{fp}} [\![\mathbf{u}]\!]_{ej} \frac{F_{s,+}^e}{4} \phi_{ej,+}^{DG}(\mathbf{r}) \right) \right) d\Gamma_{s,e} \right].
\end{aligned}
\tag{80}
$$

Both $[\![\mathbf{u}]\!]$ and ϕ_{ej}^{DG} are polynomials of degree less or equal to p. We compute exactly the integrals using Gauss-Legendre quadratures to obtain,

$$
\begin{aligned}
\Theta_{e,2} &= -|J^e| \left[\sum_{i=1}^{N_{fp}} \left(s[\![u]\!]_{ei}\omega_i \left(\sum_{j=1}^{N_{fp}} [\![u]\!]_{ej} \frac{F_{s,-}^e}{4} \phi_{ej,-}^{DG}(\mathbf{r}_i^e) \right) \right) \right. \\
&\quad \left. + \sum_{i=1}^{N_{fp}} \left(s[\![u]\!]_{ei}\omega_i \left(\sum_{j=1}^{N_{fp}} [\![u]\!]_{ej} \frac{F_{s,+}^e}{4} \phi_{ej,+}^{DG}(\mathbf{r}_i^e) \right) \right) \right] \\
&= -|J^e| \sum_{i=1}^{N_{fp}} [\![u]\!]_{ei} s\omega_i \left[\sum_{j=1}^{N_{fp}} [\![u]\!]_{ej} \left(\frac{F_{s,-}^e}{4} \psi_{ej,-}^{DG}(\mathbf{r}_i^e) + \frac{F_{s,+}^e}{4} \psi_{ej,+}^{DG}(\mathbf{r}_i^e) \right) \right].
\end{aligned}
\tag{81}
$$

□

Theorem 5.2. *Employing the BR2 scheme for the diffusion equation with the ESFR methods, for all edges e and for all flux points i, s_{ei} greater than s_{ei}^* implies the energy stability of the solution, with*

$$s_{ei}^* = \min_\kappa \sum_k \left(\frac{\psi_{ei,k}(\mathbf{r}_i^e) - |\psi_{ei,k}(\mathbf{r}_i^e)| + \sum_{f=1}^{3} \frac{F_{s,k}^f |\mathbf{n} \cdot \mathbf{n}^f|}{2F_{s,k}^e} \left(\sum_{j=1}^{N_{fp}} \left(|\psi_{fj,k}(\mathbf{r}_i^e)| + \frac{\omega_j}{\omega_i} \left| \psi_{ei,k}(\mathbf{r}_j^f) \right| \right) \right)}{\psi_{ei,k}^{DG}(\mathbf{r}_i^e) + \left| \psi_{ei,k}^{DG}(\mathbf{r}_i^e) \right| - \frac{1}{2} \sum_{j=1}^{N_{fp}} \left(\left| \psi_{ej,k}^{DG}(\mathbf{r}_i^e) \right| + \frac{\omega_j}{\omega_i} \left| \psi_{ei,k}^{DG}(\mathbf{r}_j^e) \right| \right)} \right). \tag{82}$$

where $k = \{-, +\}$, $k = -$ signifies interior to triangle IJK, and $k = +$ denotes triangle IJK_{+1}.

33

Proof. The expression of Θ_{e2} in Corollary 5.1 is similar to the second part of the first term in (40). Similarly, we apply the triangular inequality,

$$
\begin{aligned}
\Theta_{e,2} \quad \leq \quad & |J^e| \frac{F^e_{s,-}}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]^2_{ei} \left(-s_{ei}\omega_i \psi^{DG}_{ei,-}\left(\mathbf{r}^e_i\right) + \frac{|s_{ei}|}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \left(\omega_i \left| \psi^{DG}_{ej,-}\left(\mathbf{r}^e_i\right) \right| + \omega_j \left| \psi^{DG}_{ei,-}\left(\mathbf{r}^e_j\right) \right| \right) \right) \right] \\
& + |J^e| \frac{F^e_{s,+}}{4} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]^2_{ei} \left(-s_{ei}\omega_i \psi^{DG}_{ei,+}\left(\mathbf{r}^e_i\right) + \frac{|s_{ei}|}{2} \sum_{\substack{j=1 \\ j \neq i}}^{N_{fp}} \left(\omega_i \left| \psi^{DG}_{ej,+}\left(\mathbf{r}^e_i\right) \right| + \omega_j \left| \psi^{DG}_{ei,+}\left(\mathbf{r}^e_j\right) \right| \right) \right) \right],
\end{aligned}
\tag{83}
$$

where the last term of the two lines are obtained via similar derivations performed in the IP section (from equation (49) up to (50)). We then combine the previous inequality with inequality (53). We can further simplify the expression by removing the exclusion of $j \neq i$ (and retrieve an additional term$- |s_{ei}| \left| \psi^{DG}_{ei,k}\left(\mathbf{r}^e_i\right) \right|$) and introduce the parameter $k = \{-,+\}$ to signify the summation of the terms across the edge to yield,

$$
\begin{aligned}
\Theta_{dif} \quad \leq \quad & \frac{|J^e|}{4} \sum_{e=1}^{N_e} \sum_{i=1}^{N_{fp}} \left[[\![u]\!]^2_{ei} \left(\sum_k \left[F^e_{s,k}\omega_i \left(\psi_{ei,k}\left(\mathbf{r}^e_i\right) - |\psi_{ei,k}\left(\mathbf{r}^e_i\right)| - s_{ei}\psi^{DG}_{ei,k}\left(\mathbf{r}^e_i\right) - |s_{ei}| \left| \psi^{DG}_{ei,k}\left(\mathbf{r}^e_i\right) \right| \right) \right. \right. \right. \\
& \left. + \frac{F^e_{s,k} |s_{ei}|}{2} \sum_{j=1}^{N_{fp}} \left(\omega_i \left| \psi^{DG}_{ej,k}\left(\mathbf{r}^e_i\right) \right| + \omega_j \left| \psi^{DG}_{ei,k}\left(\mathbf{r}^e_j\right) \right| \right) \right. \\
& \left. \left. \left. + \sum_{f=1}^{3} \sum_{j=1}^{N_{fp}} \left(\frac{F^f_{s,k} |\mathbf{n} \cdot \mathbf{n}_f|}{2} \left(\omega_i |\psi_{fj,k}\left(\mathbf{r}^e_i\right)| + \left| \psi_{ei,k}\left(\mathbf{r}^f_j\right) \right| \right) \right) \right] \right) \right].
\end{aligned}
\tag{84}
$$

One minor problem from the previous equation is that we are unable to factor out the parameter s as some of the terms are multiplied by $|s|$. This can be resolved by making the assumption that the parameter s is non-negative, $s \geq 0$. This assumption does not pose a problem for the stability proof as: if s was negative then by taking it to be positive only adds more dissipation and ensures the stability of the scheme.

In order to have $\Theta_{dif} \leq 0$, we require $s_{ei} \geq s^*_{ei} \; \forall e \in [\![1, N_e]\!] \; \forall i \in [\![1, N_{fp}]\!]$, where s^*_{ei} is defined as

$$
s^*_{ei} \quad = \quad \sum_k \left(\frac{\psi_{ei,k}\left(\mathbf{r}^e_i\right) - |\psi_{ei,k}\left(\mathbf{r}^e_i\right)| + \sum_{f=1}^{3} \frac{F^f_{s,k} |\mathbf{n} \cdot \mathbf{n}^f|}{2F^e_{s,k}} \sum_{j=1}^{N_{fp}} \left(|\psi_{fj,k}\left(\mathbf{r}^e_i\right)| + \frac{\omega_j}{\omega_i} \left| \psi_{ei,k}\left(\mathbf{r}^f_j\right) \right| \right)}{\psi^{DG}_{ei,k}\left(\mathbf{r}^e_i\right) + \left| \psi^{DG}_{ei,k}\left(\mathbf{r}^e_i\right) \right| - \frac{1}{2} \sum_{j=1}^{N_{fp}} \left(\left| \psi^{DG}_{ej,k}\left(\mathbf{r}^e_i\right) \right| + \frac{\omega_j}{\omega_i} \left| \psi^{DG}_{ei,k}\left(\mathbf{r}^e_j\right) \right| \right)} \right).
\tag{85}
$$

34

We showed through Postulate 3.2 that the problem is independent of κ. As a consequence, the energy stability is also independent of κ. Minimizing the previous equation results in criterion (82).

$\square$

Remark 5.2. The minimization of s^* with respect to κ yields similar results to Figure 7 and are hence omitted. In the rest of the article, s^*_{theory} is computed with equation (82) with κ_+.

5.2. Numerical results

In order to validate the previous result, we conduct numerical simulations to find the minimum penalty term $s^*_{numerical}$ that ensures stability. We consider the same problem and parameters as in Section 4.2.

p	2		3	
κ ⟍ c	κ_{DG}	κ_+	κ_{DG}	κ_+
$\max\left(\tau^*_{theory}\right)$	1.82	1.82	2.51	2.51
c_{DG}	0.67	0.67	0.76	0.76
c_+	0.53	0.53	0.73	0.73

Table 4: $s^*_{numerical}$ for the BR2 scheme for $p = 2$ and $p = 3$ for a $8 \times 8 \times 2$ mesh.

p	2		3	
κ ⟍ c	κ_{DG}	κ_+	κ_{DG}	κ_+
$\max\left(s^*_{theory}\right)$	1.82	1.82	2.51	2.51
c_{DG}	0.73	0.73	0.78	0.78
c_+	0.68	0.68	0.76	0.76

Table 5: $s^*_{numerical}$ for the BR2 scheme for $p = 2$ and $p = 3$ for a $16 \times 16 \times 2$ mesh.

Both tables 4 and 5 guarantee that $\max\left(s_{theory}\right) \geq s^*_{numerical}$. Similar to the IP results, these tables do not validate the criterion of (82) for every edge and flux point but the results do not contradict our criterion.

6. Von Neumann analysis

This section presents a von Neumann analysis to study the maximal time step, Δt_{max} of the different schemes.

6.1. Maximal time step

ESFR schemes offer a range of methods where the values of both c and κ dictate the amount of filtering or a relaxation of the highest modes of the DG correction fields [18]. The purpose of this

section is to present the methods with the highest time step. Castonguay et al. [11] performed the analysis for the advection equation and studied the influence of c (Table 1). From section 3, we demonstrated that the problem is independent of κ. In this section, we will confirm this theoretical result and study the influence of c and the penalty term (τ and s) of the numerical schemes. Let us consider a 2D periodic pattern, controled by the angle γ, which forms the domain Ω.

Figure 9: Pattern used for the von Neumann analysis.

Referring to Figure 9, the pattern is formed by two triangles, themselves parametrized by two vectors, $\vec{B_1}$ and $\vec{B_2}$. In the study considered, we chose $\vec{B_1} = \begin{pmatrix} \Delta_B \\ 0 \end{pmatrix}$ and $\vec{B_2} = \begin{pmatrix} \Delta_B \cos(\gamma) \\ \Delta_B \sin(\gamma) \end{pmatrix}$. As shown by Castonguay [11], the shape of the pattern has an influence on Δt_{max} hence two values of γ will be studied (60° and 90°). We then create the column vector solution of the pattern element $\bar{u}_{mn} = \begin{pmatrix} \bar{u}_1 \\ \bar{u}_2 \end{pmatrix}$, where $\bar{u}_1$ contains the discrete value on the N_p nodes of the solution on the first triangle and $\bar{u}_2$ contains the discrete values of the solution on the second triangle, represented in Figure 9. Hence $\bar{u}_{mn}$ is a $2N_p \times 1$ vector. The solution points choosen for the analysis are the α-optimized points.

36

We then nondimensionalize the partial differential equation (1) by introducing $\hat{x} = \frac{x}{\Delta_B}$, $\hat{y} = \frac{y}{\Delta_B}$, $\hat{t} = \frac{t\,b}{(\Delta_B)^2}$,

$$\frac{\partial u}{\partial \hat{t}} = \left(\frac{\partial^2 u}{\partial \hat{x}^2} + \frac{\partial^2 u}{\partial \hat{y}^2} \right). \tag{86}$$

Equation (1) is recovered by taking $b = 1$ and a unitary element size $\Delta_B = 1$. We now rename the variable $\hat{x}$ and $\hat{y}$ by x and y. We apply the FR procedure with the IP and BR2 numerical fluxes. Both schemes have compact stencils, where the solution requires information only from its closest neighbours. We write (86) in discrete form as

$$\frac{d\bar{u}_{mn}}{dt} = \left(\mathbf{A}\bar{u}_{m,n} + \mathbf{B}\bar{u}_{m,n+1} + \mathbf{C}\bar{u}_{m,n-1} + \mathbf{D}\bar{u}_{m-1,n} + \mathbf{E}\bar{u}_{m+1,n} \right), \tag{87}$$

where, $\mathbf{A}$ corresponds to a matrix of size $2N_p \times 2N_p$ taking $\bar{u}_{m,n}$ as argument to compute the Laplacian of u_n, and the same can be said of matrices $\mathbf{B}$, $\mathbf{C}$, $\mathbf{D}$ and $\mathbf{E}$ by taking the respective neighbours of the pattern (m, n) as argument. Lowercase letter with a bar, $\bar{a}$ indicates a column vector containing the discrete values of a scalar quantity. Assuming a Bloch-wave solution [19], we have

$$\bar{u}_{m,n} = e^{i|k|(x_{mn} \cos(\theta) + y_{m,n} \sin(\theta)) - \omega t} \bar{v}, \tag{88}$$

where $\vec{k} = |k| \begin{pmatrix} \cos(\theta) \\ \sin(\theta) \end{pmatrix}$ represents the prescribed wave vector, both $|k|$ and θ vary between $[0, 2\pi]$, ω^δ is the discrete frequency and $\bar{v}$ is a vector independent of the elements. The solution is periodic and hence we obtain,

$$\begin{aligned} \bar{u}_{m,n+1} &= e^{i|k|((x_{mn} + \Delta_B \cos(\gamma)) \cos(\theta) + (y_{m,n} + \Delta_B \sin(\gamma)) \sin(\theta)) - \omega t} \bar{v} \\ &= \bar{u}_{m,n} e^{i|k|\Delta_B (\cos(\gamma - \theta))}, \end{aligned} \tag{89}$$

and the quantities $\bar{u}_{m,n-1}$, $\bar{u}_{m-1,n}$ and $\bar{u}_{m+1,n}$ are calculated similarly. We finally obtain

$$\frac{d\bar{u}_{mn}}{dt} = \mathbf{S}\left(|k|, \theta \right) \bar{u}_{m,n}, \tag{90}$$

where $\mathbf{S} = \left(\mathbf{A} + \mathbf{B}e^{i|k|\Delta_B \cos(\gamma - \theta)} + \mathbf{C}e^{-i|k|\Delta_B \cos(\gamma - \theta)} + \mathbf{D}e^{-i|k|\Delta_B \cos(\theta)} + \mathbf{E}e^{i|k|\Delta_B \cos(\theta)} \right)$. The quantities ω^δ and $\bar{v}$ can be computed from the eigenvalues and eigenvectors of $\mathbf{S}$.

We employ the fourth order five stage Runge-Kutta (RK54) method [16]. The column-vector of the solution at time $n + 1$, $\bar{u}_{mn}^{t_{n+1}}$, can be expressed as,

$$\bar{u}_{mn}^{t_{n+1}} = \mathbf{M}\left(|k|, \theta, \Delta t \right) \bar{u}_{m,n}^{t_n}, \tag{91}$$

37

where, $\mathbf{M}$ is defined as

$$\mathbf{M}(k, \Delta t) = 1 + \Delta t \mathbf{S}(k) + \frac{1}{2!}(\Delta t \mathbf{S}(k))^2 + \frac{1}{3!}(\Delta t \mathbf{S}(k))^3 + \frac{1}{4!}(\Delta t \mathbf{S}(k))^4 + \frac{1}{200}(\Delta t \mathbf{S}(k))^5. \quad (92)$$

The matrix $\mathbf{M}$ depends on $|k|$, θ and Δt. Whereas $|k|$ and θ vary between $[0, 2\pi]$, Δt is choosen to ensure the stability of the scheme, i.e the moduli of the spectral radius of $\mathbf{M}$ must be less than 1. The algorithm resumes to: start at an initial Δt_0 sufficiently high to produce an unstable solution, then scan over the range of $|k|$ and θ to compute the highest eigenvalue: $|\lambda|_{max}$, decrease Δt until $|\lambda|_{max} \leq 1$. We apply the analysis for both the IP and BR2 schemes, where the value of c is taken to be equal to c_{DG} and c_+ (Table 1), while two values of the penalty term are taken (either equal or 1.5 times the criterion found: (47) for the IP scheme and (82) for the BR2 scheme).

p		2		3	
c	κ	τ_{theory}	$1.5\tau_{theory}$	τ_{theory}	$1.5\tau_{theory}$
c_{DG}	κ_{DG}	1.85e-02	1.12e-02	6.23e-03	3.85e-03
	κ_+	1.85e-02	1.12e-02	6.23e-03	3.85e-03
c_+	κ_{DG}	2.74e-02	1.76e-02	8.22e-03	5.20e-03
	κ_+	2.74e-02	1.76e-02	8.22e-03	5.20e-03

Table 6: Δt_{max} for triangles for the IP scheme for $\gamma = 60°$.

p		2		3	
c	κ	τ_{theory}	$1.5\tau_{theory}$	τ_{theory}	$1.5\tau_{theory}$
c_{DG}	κ_{DG}	1.77e-02	1.08e-02	5.82e-03	3.59e-03
	κ_+	1.77e-02	1.08e-02	5.82e-03	3.59e-03
c_+	κ_{DG}	2.64e-02	1.69e-02	7.83e-03	4.93e-03
	κ_+	2.64e-02	1.69e-02	7.83e-03	4.93e-03

Table 7: Δt_{max} for triangles for the IP scheme for $\gamma = 90°$.

p		2		3	
c	κ	s_{theory}	$1.5\,s_{theory}$	s_{theory}	$1.5\,s_{theory}$
c_{DG}	κ_{DG}	1.08e-02	6.62e-03	3.31e-03	2.06e-03
	κ_+	1.08e-02	6.62e-03	3.31e-03	2.06e-03
c_+	κ_{DG}	1.56e-02	1.02e-02	4.13e-03	2.66e-03
	κ_+	1.56e-02	1.02e-02	4.13e-03	2.66e-03

Table 8: Δt_{max} for triangles for the BR2 scheme for $\gamma = 60°$.

p		2		3	
c	κ	s_{theory}	$1.5\,s_{theory}$	s_{theory}	$1.5\,s_{theory}$
c_{DG}	κ_{DG}	1.00e-02	6.09e-03	2.81e-03	1.72e-03
	κ_+	1.00e-02	6.09e-03	2.81e-03	1.72e-03
c_+	κ_{DG}	1.39e-02	8.99e-03	3.58e-03	2.27e-03
	κ_+	1.39e-02	8.99e-03	3.58e-03	2.27e-03

Table 9: Δt_{max} for triangles for the BR2 scheme for $\gamma = 90°$.

Through all these tables, for both the IP and BR2 numerical fluxes, we observe, that the lower the value of the penalty term the higher the maximal time step. Conversely, increasing c leads to higher maximal time steps. As expected, κ has no influence on Δt_{max}. Comparing Tables 6 and 7 for the IP scheme, we observe, the maximal time step is always higher for $\gamma = 60°$ than $90°$. Indeed $\gamma = 60°$ results in a domain where each element is an equilateral triangle. Therefore this domain

is more regular than $\gamma = 90°$ and it results in an increase of the maximal time step. Equivalent trends are observed for the BR2 scheme as shown in Tables 8 and 9.

Comparing the BR2 and the IP schemes, we observe that the latter has the higher time step for every case. A likely explanation is that the criterion for the IP scheme (47) is sharper than the one for the BR2 scheme (82).

As expected , Δt_{max} does not depend on κ.

(a) Δt_{max} for $p = 2$.

(b) Δt_{max} for $p = 3$.

Figure 10: Δt_{max} along c for $\gamma = 90°$ for p, the diamond marker represents $c = c_+$; using log scale the DG case couldn't be represented.

Figure 10 lays emphasis to the IP scheme providing a higher Δt_{max} than the BR2 scheme. Moreover, we observe that by taking c closer to c_+ provides for a higher time step. Therefore it would seem that $c = c_+$ for an advection-diffusion problem would lead to the highest Δt_{max} possible.

The insights of the features of the ESFR method given by the von Neumann analysis is only true for regular meshes. However real CFD applications often use irregular grids due to geometric complexity and flow anisotropy. Thus we recommend against any generalization from this section.

7. L_2 errors and order of accuracy

The criterion for both the IP and BR2 schemes have been mostly confirmed with the analysis in Sections 4.2 and 5.2. Our previous von Neumann analysis has enlightened methods which allow for a time step much higher than the classical DG method. In this section, we present the numerical L_2 errors and the order of convergence to verify that we get the expected OOA of $p + 1$ for the different methods. The problem considered is the same as the one before. We consider a regular mesh (Figure 8) of size $N_x \times N_x \times 2$ where N_x has been taken equal to 16, 32 and 64. We use two values for τ for the IP schemes: τ_{theory} given by (47) and $1.5\tau_{theory}$ and two values for the BR2 schemes s_{theory} given by (82) and $1.5s_{theory}$. The L_2-error is computed as

$$
L_2 - \text{error} = \sqrt{\frac{\displaystyle\sum_{n=1}^{2N_x^2}\sum_{i=1}^{N_p}\left(u_{n,i} - u_{exact,n,i}\right)^2}{2N_x^2 N_p}}, \tag{93}
$$

where $u_{n,i}$ is the numerical solution evaluated on element n at the solution point i and $u_{exact,n,i}$ is the exact solution on the same element and solution point. The final time t_{fin} was taken to be equal to 1.

The simulations were conducted for four methods: c_{DG}/κ_{DG}, c_{DG}/κ_+, c_+/κ_{DG} and c_+/κ_+ where $c_+ = \kappa_+$ is taken equal to the second column of Table 1. We provided the maximal time step for $N_x = 16$ elements. Maximal time steps were evaluated through an iterative approach while ensuring the solution remains bounded at $t = 2$.

c	κ	τ_{theory}						$1.5\tau_{theory}$					
		N_x=16	N_x=32	N_x=64	OOA		Δt_{max}	N_x=16	N_x=32	N_x=64	OOA		Δt_{max}
c_{DG}	κ_{DG}	1.16e-04	1.45e-05	1.81e-06	- 3.00	3.00	2.70e-03	8.91e-05	1.12e-05	1.40e-06	- 3.00	3.00	1.70e-03
	κ_+	1.16e-04	1.45e-05	1.81e-06	- 3.00	3.00	2.70e-03	8.91e-05	1.12e-05	1.40e-06	- 3.00	3.00	1.70e-03
c_+	κ_{DG}	1.26e-04	1.48e-05	1.82e-06	- 3.09	3.02	4.10e-03	9.63e-05	1.14e-05	1.40e-06	- 3.08	3.02	2.60e-03
	κ_+	1.26e-04	1.48e-05	1.82e-06	- 3.09	3.02	4.10e-03	9.63e-05	1.14e-05	1.40e-06	- 3.08	3.02	2.60e-03

Table 10: L_2 errors using the IP scheme for $p = 2$ for triangles.

c	κ	s_{theory}							$1.5s_{theory}$						
		$N_x=16$	$N_x=32$	$N_x=64$		OOA		Δt_{max}	$N_x=16$	$N_x=32$	$N_x=64$		OOA		Δt_{max}
c_{DG}	κ_{DG}	9.57e-05	1.20e-05	1.50e-06	-	2.99	3.00	1.50e-03	8.22e-05	1.03e-05	1.28e-06	-	3.00	3.00	9.50e-04
	κ_+	9.57e-05	1.20e-05	1.50e-06	-	2.99	3.00	1.50e-03	8.22e-05	1.03e-05	1.28e-06	-	3.00	3.00	9.50e-04
c_+	κ_{DG}	1.03e-04	1.22e-05	1.51e-06	-	3.07	3.02	2.10e-03	8.77e-05	1.04e-05	1.29e-06	-	3.07	3.02	1.40e-03
	κ_+	1.03e-04	1.22e-05	1.51e-06	-	3.07	3.02	2.10e-03	8.77e-05	1.04e-05	1.29e-06	-	3.07	3.02	1.40e-03

Table 11: L_2 errors using the BR2 scheme for $p = 2$ for triangles.

c	κ	τ_{theory}							$1.5\tau_{theory}$						
		$N_x=16$	$N_x=32$	$N_x=64$		OOA		Δt_{max}	$N_x=16$	$N_x=32$	$N_x=64$		OOA		Δt_{max}
c_{DG}	κ_{DG}	3.94e-06	2.43e-07	1.51e-08	-	4.02	4.01	9.10e-04	3.41e-06	2.14e-07	1.34e-08	-	4.00	4.00	5.60e-04
	κ_+	3.94e-06	2.43e-07	1.51e-08	-	4.02	4.01	9.10e-04	3.41e-06	2.14e-07	1.34e-08	-	4.00	4.00	5.60e-04
c_+	κ_{DG}	3.92e-06	2.42e-07	1.51e-08	-	4.02	4.00	1.20e-03	3.45e-06	2.14e-07	1.34e-08	-	4.01	4.00	7.70e-04
	κ_+	3.92e-06	2.42e-07	1.51e-08	-	4.02	4.00	1.20e-03	3.45e-06	2.14e-07	1.34e-08	-	4.01	4.00	7.70e-04

Table 12: L_2 errors using the IP scheme for $p = 3$ for triangles.

c	κ	s_{theory}							$1.5s_{theory}$						
		$N_x=16$	$N_x=32$	$N_x=64$		OOA		Δt_{max}	$N_x=16$	$N_x=32$	$N_x=64$		OOA		Δt_{max}
c_{DG}	κ_{DG}	3.65e-06	2.29e-07	1.43e-08	-	3.99	4.00	4.40e-04	3.39e-06	2.14e-07	1.34e-08	-	3.98	4.00	2.70e-04
	κ_+	3.65e-06	2.29e-07	1.43e-08	-	3.99	4.00	4.40e-04	3.39e-06	2.14e-07	1.34e-08	-	3.98	4.00	2.70e-04
c_+	κ_{DG}	3.71e-06	2.30e-07	1.44e-08	-	4.01	4.00	5.60e-04	3.47e-06	2.16e-07	1.35e-08	-	4.01	4.00	3.50e-04
	κ_+	3.71e-06	2.30e-07	1.44e-08	-	4.01	4.00	5.60e-04	3.47e-06	2.16e-07	1.35e-08	-	4.01	4.00	3.50e-04

Table 13: L_2 errors using the BR2 scheme for $p = 3$ for triangles.

The maximal time step provided by the above tables can be compared with the von Neumann analysis: $\Delta t_{max} = \dfrac{\Delta \hat{t}_{max}}{b} \left(\dfrac{2}{N_x} \right)^2$. The maximum relative error is at 2.4%, which concurs with the maximal time steps obtained in this section.

For both the IP and BR2 numerical fluxes and for both $p = 2$ and $p = 3$, we obtain the expected order of accuracy: $p+1$. Similarly, we observe that the IP scheme provides a higher time step than the BR2 method but the error from the IP scheme is also higher.

8. Conclusion

This article provides a theoretical proof of energy stability for the diffusion case for triangles using the IP and BR2 schemes. Bounds for the penalty term τ for the IP scheme and s for the BR2

scheme to ensure stability for various ESFR schemes were obtained. These theoretical proofs were validated through numerical simulations and orders of accuracy were provided. It was established that for both the IP and BR2 numerical fluxes, the stability of the ESFR scheme is independent of the auxiliary correction field. A von-Neumann analysis was conducted to present methods which procure a higher maximal time step than the classical DG method. The trade-off is that the L_2-error increases for these methods. While the BR2 scheme procures the least amount of error, the IP scheme has the highest time step. Further analysis will be conducted to extend this proof for tetrahedra elements.

Acknowledgements

Appendix

A. Theoretical proof of the independecy of κ for $p = 1$

This section proposes a theoretical proof of Postulate 3.1 and hence of Postulate 3.2 for $p = 1$.

Theorem A.1. *Let the solution of the diffusion equation be approximated by a polynomial of degree $p = 1$ on the reference triangle. Let ψ_{ei} be the correction field associated to face e at the flux point i parametrized by κ and ϕ_{fj} the correction field, associated to face f at the flux point j, parametrized by c. Let $\left(\mathbf{r}_f^j\right)_{j \in [\![1, N_{fp}]\!]}$ be the Gauss-Legendre flux points on face f. Then*
$\forall c \in [0, \infty[,$

$$R_{ei}\left(\mathbf{r}\right) = \left(-\hat{\nabla}\psi_{ei}\left(\mathbf{r}\right) \cdot \hat{\mathbf{n}}_{ei} + \sum_{f=1}^{3}\sum_{j=1}^{N_{fp}}\psi_{ei}\left(\mathbf{r}_f^j\right)\left(\hat{\mathbf{n}}_{ei} \cdot \hat{\mathbf{n}}_{fj}\right)\phi_{fj}\left(\mathbf{r}\right)\right), \qquad (A.1)$$

is independent of the parameter κ.

Proof. The normal $\hat{\mathbf{n}}_{ei}$ is independent of κ. As a result, proving this Theorem is equivalent to showing that,

$$\mathbf{R}_{ei}\left(\mathbf{r}\right) = \left(-\hat{\nabla}\psi_{ei}\left(\mathbf{r}\right) + \sum_{f=1}^{3}\sum_{j=1}^{N_{fp}}\psi_{ei}\left(\mathbf{r}_f^j\right)\phi_{fj}\left(\mathbf{r}\right)\hat{\mathbf{n}}_{fj}\right), \qquad (A.2)$$

42

is independent of κ.

Figure A.11: Reference element for $p = 1$, the green squares represent the FP of the Gauss-Legendre quadrature

The reference element is equilateral, thus the flux points are symmetric, Figure A.11. Moreover, the weights of the Gauss-Legendre quadrature are equal to 1 ($\omega_1 = \omega_2 = 1$).

(r,s)	$\mathbf{r}_1^1$	$\mathbf{r}_1^2$	$\mathbf{r}_2^1$	$\mathbf{r}_2^2$	$\mathbf{r}_3^1$	$\mathbf{r}_3^2$
r	$-\dfrac{1}{\sqrt{3}}$	$-r_1^1$	$\dfrac{1}{2} + \dfrac{\sqrt{3}}{6}$	$\dfrac{1}{2} - \dfrac{\sqrt{3}}{6}$	$-r_2^2$	$-r_2^1$
s	r_1^1	r_1^1	$-r_2^2$	r_2^1	s_2^2	s_2^1

Table A.14: Numerical values of coordinates of the flux points.

As no mathematical formula, known to the authors, enables to simplify equation (A.2), we expand the correction fields ψ_{ei} and ϕ_{fj} through equation (18),

$$\mathbf{R}_{ei}\left(\mathbf{r}\right) = \left(\sum_{k=1}^{3} \sigma_{ei,k}^{\kappa} \left[-\hat{\nabla} L_k\left(\mathbf{r}\right) + \sum_{f=1}^{3}\sum_{j=1}^{2}\sum_{n=1}^{3} \left(L_k\left(r_f^j\right) \sigma_{fj,n}^{c} L_n\left(\mathbf{r}\right) \hat{\mathbf{n}}_{fj} \right) \right] \right). \qquad (A.3)$$

Before expanding further, we write the analytical formula for the Dubiner basis $(L_i)_{i \in [\![1,3]\!]}$ [11],

$$
\begin{aligned}
L_1(r,s) &= \frac{1}{3^{1/4}} \\
L_2(r,s) &= \frac{\sqrt{6}}{3^{1/4}}s \\
L_3(r,s) &= \frac{\sqrt{6}}{3^{1/4}}r
\end{aligned}
\tag{A.4}
$$

Then, we evaluate the coefficients $\left(\sigma_{ei,k}^{\kappa}\right)_{k \in [\![1,3]\!]}$ via equation (20),

$$
\mathbf{A}\sigma_{ei}^{\kappa} = \mathbf{b}_{ei},
\tag{A.5}
$$

where $\mathbf{A}$ has non-zero values only along its diagonal ($A_{11} = 1$, $A_{22} = 1 + \kappa \left(D^{(2,1)}L_2\right)^2$, $A_{33} = 1 + \kappa \left(D^{(1,2)}L_3\right)^2$) and $b_{ei,j} = L_j\left(\mathbf{r}_e^i\right)$. To simplify the derivations, we denote $\alpha = \frac{1}{3^{1/4}}$. Hence, we obtain,

$$
\begin{bmatrix} \sigma_{ei,1}^{\kappa} \\ \sigma_{ei,2}^{\kappa} \\ \sigma_{ei,3}^{\kappa} \end{bmatrix} = \begin{bmatrix} \alpha \\ \dfrac{\alpha\sqrt{6}s_e^i}{1 + 6\alpha^2\kappa} \\ \dfrac{\alpha\sqrt{6}r_e^i}{1 + 6\alpha^2\kappa} \end{bmatrix}.
\tag{A.6}
$$

Only the last two terms depend on κ; therefore we only need to prove that the projection of $\mathbf{R}_{ei,1}(\mathbf{r})$ on both $\mathbf{e}_r$ and $\mathbf{e}_s$ is independent of κ, where

$$
\mathbf{R}_{ei,1}(\mathbf{r}) = \left(\sum_{k=2}^{3} \sigma_{ei,k}^{\kappa} \left[-\hat{\nabla}L_k(\mathbf{r}) + \sum_{f=1}^{3}\sum_{j=1}^{2}\sum_{n=1}^{3} \left(L_k\left(\mathbf{r}_f^j\right) \sigma_{fj,n}^c L_n(\mathbf{r})\hat{\mathbf{n}}_{fj} \right) \right] \right).
\tag{A.7}
$$

In the following, we will show that these two projections are equal to 0.

Projection on $\mathbf{e}_r$

We can observe from Figure A.11 that the FP are symmetric. As a consequence the terms $\left(L_k\left(\mathbf{r}_f^j\right)\right)_{(f,j) \in [\![1,3]\!] \times [\![1,2]\!]}$ only require to be computed at the FP $(1,1)$, $(2,1)$ and $(2,2)$. Moreover

$\hat{\mathbf{n}}_{1j} \cdot \mathbf{e}_r = 0$. These properties yield,

$$
\begin{aligned}
\mathbf{R}_{ei,1}\left(\mathbf{r}\right) \cdot \mathbf{e}_r &= \frac{\alpha\sqrt{6}s_e^i}{1+6\alpha^2\kappa}\Bigg(&& \frac{\sqrt{3}}{2}L_2\left(s_2^1\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s + r_2^1 r\right)\right) \\
&&+\ & \frac{\sqrt{3}}{2}L_2\left(s_2^2\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s + r_2^2 r\right)\right) \\
&&-\ & \frac{\sqrt{3}}{2}L_2\left(s_2^2\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s - r_2^2 r\right)\right) \\
&&-\ & \frac{\sqrt{3}}{2}L_2\left(s_2^1\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s - r_2^1 r\right)\right)\Bigg) \\
&+\frac{\alpha\sqrt{6}r_e^i}{1+6\alpha^2\kappa}\Bigg(-\sqrt{6}\alpha\ +\ & \frac{\sqrt{3}}{2}L_3\left(r_2^1\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s + r_2^1 r\right)\right) \\
&&+\ & \frac{\sqrt{3}}{2}L_3\left(r_2^2\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s + r_2^2 r\right)\right) \\
&&+\ & \frac{\sqrt{3}}{2}L_3\left(r_2^2\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s - r_2^2 r\right)\right) \\
&&+\ & \frac{\sqrt{3}}{2}L_3\left(r_2^1\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s - r_2^1 r\right)\right)\Bigg) \\
&= \frac{\alpha\sqrt{6}s_e^i}{1+6\alpha^2\kappa}\Bigg(&& \sqrt{3}L_2\left(s_2^1\right) && \left(\frac{6\alpha^2}{1+6\alpha^2 c}r_2^1 r\right) \\
&&+\ & \sqrt{3}L_2\left(s_2^2\right) && \left(\frac{6\alpha^2}{1+6\alpha^2 c}r_2^2 r\right)\Bigg) \\
&+\frac{\alpha\sqrt{6}r_e^i}{1+6\alpha^2\kappa}\Bigg(-\sqrt{6}\alpha\ +\ & \sqrt{3}L_3\left(r_2^1\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}s_2^1 s\right) \\
&&+\ & \sqrt{3}L_3\left(r_2^2\right) && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}s_2^2 s\right)\Bigg).
\end{aligned}
$$
(A.8)

We now replace the values of s_f^j with the values of r_f^j according to Table A.14 and expand the remaining functions L_k,

$$
\begin{aligned}
\mathbf{R}_{ei,1}\left(\mathbf{r}\right) \cdot \mathbf{e}_r &= \frac{\alpha\sqrt{6}s_e^i}{1+6\alpha^2\kappa}\Bigg(&& -3\sqrt{2}\alpha r_2^2 && \left(\frac{6\alpha^2}{1+6\alpha^2 c}r_2^1 r\right) \\
&&+\ & 3\sqrt{2}\alpha r_2^1 && \left(\frac{6\alpha^2}{1+6\alpha^2 c}r_2^2 r\right)\Bigg) \\
&+\frac{\alpha\sqrt{6}r_e^i}{1+6\alpha^2\kappa}\Bigg(-\sqrt{6}\alpha\ +\ & 3\sqrt{2}\alpha r_2^1 && \left(\alpha^2 - \frac{6\alpha^2}{1+6\alpha^2 c}r_2^2 s\right) \\
&&+\ & 3\sqrt{2}\alpha r_2^2 && \left(\alpha^2 + \frac{6\alpha^2}{1+6\alpha^2 c}r_2^1 s\right)\Bigg).
\end{aligned}
$$
(A.9)

We finally obtain

$$\mathbf{R}_{ei,1}\left(\mathbf{r}\right)\cdot\mathbf{e}_r = -\frac{6\alpha^2 r_e^i}{1+6\alpha^2\kappa}\left(-1+\sqrt{3}\alpha^2\left(r_2^1+r_2^2\right)\right), \tag{A.10}$$

since $\sqrt{3}\alpha^2 =$ and $\left(r_2^1+r_2^2\right)$ are equal to 1, we retrieve the expected result.

The projection on $\mathbf{e}_s$ is similar hence only the main steps will be given.

Projection on $\mathbf{e}_s$

As $\hat{\mathbf{n}}_{1j}\cdot\mathbf{e}_s = -1$, additional terms need to be derived.

$$
\begin{aligned}
\mathbf{R}_{ei,1}\left(\mathbf{r}\right)\cdot\mathbf{e}_s =\ & \frac{\alpha\sqrt{6}s_e^i}{1+6\alpha^2\kappa}\Bigg(-\sqrt{6}\alpha \ -\ L_2\left(s_1^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_1^1 s+r_1^1 r\right)\right) \\
& -\ L_2\left(s_1^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_1^1 s-r_1^1 r\right)\right) \\
& +\ \frac{1}{2}L_2\left(s_2^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s+r_2^1 r\right)\right) \\
& +\ \frac{1}{2}L_2\left(s_2^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s+r_2^2 r\right)\right) \\
& +\ \frac{1}{2}L_2\left(s_2^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s-r_2^2 r\right)\right) \\
& +\ \frac{1}{2}L_2\left(s_2^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s-r_2^1 r\right)\right) \Bigg) \\
& +\frac{\alpha\sqrt{6}r_e^i}{1+6\alpha^2\kappa}\Bigg(-\ L_3\left(r_1^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_1^1 s+r_1^1 r\right)\right) \\
& +\ L_3\left(r_1^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_1^1 s-r_1^1 r\right)\right) \\
& +\ \frac{1}{2}L_3\left(r_2^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s+r_2^1 r\right)\right) \\
& +\ \frac{1}{2}L_3\left(r_2^2\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s+r_2^2 r\right)\right) \\
& -\ \frac{1}{2}L_3\left(r_2^2\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^2 s-r_2^2 r\right)\right) \\
& -\ \frac{1}{2}L_3\left(r_2^1\right)\left(\alpha^2+\frac{6\alpha^2}{1+6\alpha^2 c}\left(s_2^1 s-r_2^1 r\right)\right) \Bigg).
\end{aligned}
\tag{A.11}
$$

After further simplifications we obtain,

$$
\begin{aligned}
\mathbf{R}_{ei,1}\left(\mathbf{r}\right)\cdot\mathbf{e}_s =\ & \frac{6\alpha^2 s_e^i}{1+6\alpha^2\kappa}\left(\left(-1+\alpha^2\left(-2s_1^1+s_2^1+s_2^2\right)\right)\right. \\
& +\ \frac{6\alpha^2 s_e^i}{1+6\alpha^2\kappa}\frac{6\alpha^2 s}{1+6\alpha^2 c}\left(-2\left(s_1^1\right)^2+\left(s_2^1\right)^2+\left(s_2^2\right)^2\right) \\
& +\ \frac{6\alpha^2 r_e^i}{1+6\alpha^2\kappa}\frac{6\alpha^2 r}{1+6\alpha^2 c}\left(-2\left(r_1^1\right)^2+\left(r_2^1\right)^2+\left(r_2^2\right)^2\right).
\end{aligned}
\tag{A.12}
$$

46

Replacing r_f^j and s_f^j with their numerical values in Table A.14, the last two lines are equal to 0. Futhermore,

$$
\begin{aligned}
\alpha^2 \left(-2s_1^1 + s_2^1 + s_2^2 \right) &= \frac{1}{\sqrt{3}} \left(\frac{2}{\sqrt{3}} + \left(\frac{\sqrt{3}}{6} - \frac{1}{2} \right) + \left(\frac{1}{2} + \frac{\sqrt{3}}{6} \right) \right) \\
&= 1
\end{aligned}
\tag{A.13}
$$

Hence we have $\mathbf{R}_{ei,1}\left(\mathbf{r}\right) \cdot \mathbf{e}_s = 0$. $\qquad\square$

Corollary A.1. *The diffusion equation is independent of κ when employing the IP or BR2 numerical fluxes with a $p = 1$ interpolation.*

Proof. We let the reader refer to the proof of Postulate 3.2. $\qquad\square$

References

References

[1] T. W. Jan S. Hesthaven, Nodal Discontinuous Galerkin methods Algorithms, Analysis and Applications, Springer, 2008.

[2] F. Bassi, S. Rebay, A High-Order Accurate Discontinuous Finite Element Method for the Numerical Solution of the Compressible NavierStokes Equations, Journal of Computational Physics 131 (2) (1997) 267–279. doi:10.1006/jcph.1996.5572.

[3] F. Bassi, S. Rebay, A High Order Discontinuous Galerkin Method for Compressible Turbulent Flows, in: Discontinuous Galerkin Methods, Springer, Berlin, Heidelberg, 2000, pp. 77–88, dOI: 10.1007/978-3-642-59721-3_4.

[4] D. N. Arnold, An Interior Penalty Finite Element Method with Discontinuous Elements, SIAM Journal on Numerical Analysis 19 (4) (1982) 742–760. doi:10.1137/0719052.

[5] B. Cockburn, C.-W. Shu, The Local Discontinuous Galerkin Method for Time-Dependent Convection-Diffusion Systems, SIAM Journal on Numerical Analysis 35 (6) (1998) 2440–2463. doi:10.1137/S0036142997316712.

[6] J. Peraire, P.-O. Persson, The Compact Discontinuous Galerkin (CDG) Method for Elliptic Problems, SIAM Journal on Scientific Computing 30 (4) (2008) 1806–1824. doi:10.1137/070685518.

[7] H. T. Huynh, A Flux Reconstruction Approach to High-Order Schemes Including Discontinuous Galerkin Methods, American Institute of Aeronautics and Astronautics, 2007. doi:10.2514/6.2007-4079.

[8] H. T. Huynh, A Reconstruction Approach to High-Order Schemes Including Discontinuous Galerkin for Diffusion, American Institute of Aeronautics and Astronautics, 2009. doi:10.2514/6.2009-403.

[9] Y. Liu, M. Vinokur, Z. J. Wang, Spectral difference method for unstructured grids I: Basic formulation, Journal of Computational Physics 216 (2) (2006) 780–801. doi:10.1016/j.jcp.2006.01.024.

[10] P. E. Vincent, P. Castonguay, A. Jameson, A New Class of High-Order Energy Stable Flux Reconstruction Schemes, Journal of Scientific Computing 47 (1) (2011) 50–72. doi:10.1007/s10915-010-9420-z.

[11] P. Castonguay, P. E. Vincent, A. Jameson, A New Class of High-Order Energy Stable Flux Reconstruction Schemes for Triangular Elements, Journal of Scientific Computing 51 (1) (2012) 224–256. doi:10.1007/s10915-011-9505-3.

[12] D. M. Williams, A. Jameson, Energy Stable Flux Reconstruction Schemes for Advection Diffusion Problems on Tetrahedra, Journal of Scientific Computing 59 (3) (2014) 721–759. doi:10.1007/s10915-013-9780-2.

[13] P. Castonguay, D. M. Williams, P. Vincent, A. Jameson, Energy stable flux reconstruction schemes for advection diffusion problems, Computer Methods in Applied Mechanics and Engineering 267 (2013) 400–417. doi:10.1016/j.cma.2013.08.012.

[14] D. M. Williams, P. Castonguay, P. E. Vincent, A. Jameson, Energy stable flux reconstruction schemes for advection diffusion problems on triangles, Journal of Computational Physics 250 (Supplement C) (2013) 53–76. doi:10.1016/j.jcp.2013.05.007.

[15] S. Quaegebeur, S. Nadarajah, F. Navah, P. Zwanenburg, Stability of Energy Stable Flux Reconstruction for the diffusion problem using the Interior Penalty and Bassi Rebay II schemes, submitted to SIAM SISC.

[16] M. H. Carpenter, C. A. Kennedy, Fourth-order 2N-storage Runge-Kutta schemes, Tech. rep. (Jun. 1994).

[17] H. T. Huynh, High-Order Methods Including Discontinuous Galerkin by Reconstructions on Triangular Meshes, American Institute of Aeronautics and Astronautics, 2011. `doi:10.2514/6.2011-44`.

[18] P. Zwanenburg, S. Nadarajah, Equivalence between the Energy Stable Flux Reconstruction and Filtered Discontinuous Galerkin Schemes, Journal of Computational Physics 306 (2016) 343–369. `doi:10.1016/j.jcp.2015.11.036`.

[19] M. Ainsworth, Dispersive and dissipative behaviour of high order discontinuous galerkin finite element methods, Journal of Computational Physics 198 (1) (2004) 106 – 130. `doi:https://doi.org/10.1016/j.jcp.2004.01.004`.