Noname manuscript No.
(will be inserted by the editor)

Insights on aliasing driven instabilities for advection equations with application to Gauss-Lobatto discontinuous Galerkin methods

Juan Manzanero · Gonzalo Rubio · Esteban Ferrer · Eusebio Valero · David A. Kopriva

Received: date / Accepted: date

Abstract We analyse instabilities due to aliasing errors when solving one dimensional non-constant advection speed equations and discuss means to alleviate these types of errors when using high order discontinuous Galerkin (DG) schemes. First, we compare analytical bounds for the continuous and discrete version of the PDEs. Whilst traditional L^2 norm energy bounds applied to the discrete PDE do not always predict the physical behaviour of the continuous version of the equation, more strict elliptic norm bounds correctly bound the behaviour of the continuous PDE. Having derived consistent bounds, we analyse the effectiveness of two stabilising techniques: over-integration and split form variations (conservative, non-conservative and skew-symmetric). Whilst the former is shown to not alleviate aliasing in general, the latter ensures an aliasing-free solution if the splitting form of the discrete PDE is consistent with the continuous equation. The success of the split form de-aliasing is restricted to DG schemes with the summation-by-parts simultaneous-approximation-term (SBP-SAT) properties (e.g. DG with Gauss-Lobatto points). Numerical experiments are included to illustrate the theoretical findings.

Keywords discontinuous Galerkin · Gauss-Lobatto spectral element · summation-by-parts · discrete conservation · split formulations.

1 Introduction

High order methods are the preferred discretisation technique when high accuracy is required [37] and have gained popularity to solve incompressible

J. Manzanero (E-mail: juan.manzanero@upm.es) · G. Rubio · E. Ferrer · E. Valero
ETSIAE-UPM - School of Aeronautics, Universidad Politécnica de Madrid. Plaza Cardenal Cisneros 3, E-28040 Madrid, Spain.

David A. Kopriva
Department of Mathematics, The Florida State University, Tallahassee, FL 32306, USA.

[2, 24, 9, 10] and compressible flow problems, e.g. [12, 16, 26, 27]. The low numerical dissipation inherent to high-order methods, however, may not be sufficient to mask instabilities such as aliasing errors arising from the non-linear flux discretisation [31]. In Spectral/hp element methods, the stabilisation of the scheme has been traditionally achieved by means of over-integration techniques (see [34]), by using skew symmetric forms [35], or by adding artificial viscosity [20, 13]. The last can be implemented by upwinding inter-element fluxes (in discontinuous schemes, e.g. DG, see [11]), or by including an artificial dissipation term that maintains high-order accuracy (spectral vanishing viscosity methods [19]).

A different approach, followed by some in the finite difference community, is to develop schemes that are provably stable thanks to the summation-by-parts (SBP) and simultaneous-approximation-term (SAT) properties. The summation-by-parts property can be used to prove stability in each element (see [36]), whilst the simultaneous-approximation-term is fundamental to prove stability at both interior and physical boundaries (see [25]). A review on SBP-SAT finite differences schemes, and their stability properties can be found in [28].

Recent research on Spectral/hp methods that satisfy the summation-by-parts property has enabled the adaption of stabilisation techniques and theorems developed in the finite differences community, while retaining favourable high-order spectral element properties. In [15], for instance, Gassner showed that the discontinuous Galerkin spectral element method with Gauss-Lobatto points satisfies all formal definitions of an SBP-SAT scheme, allowing him to obtain methods that are provably stable. In [15] an energy stable split formulation scheme was developed for the Burgers equation. In [6], schemes that are strongly stable were constructed for the non-constant advection equation, considering multidimensional systems with curved elements. Later, in [7] the methodology was extended to moving geometries, under the ALE (Arbitrary-Lagrangian-Eulerian) formulation. The last step forward made by Gassner et. al. has been the derivation of schemes that are provably stable for the Euler fluid dynamics equations [17], the Magneto-Hydrodynamics [1] and the Navier-Stokes equations [18].

Here, we extend the work of Kopriva and Gassner [6], where schemes that are strongly stable (consistent with the energy bounds of the continuous PDE) were developed for the non-constant advection equation. The traditional energy bounds used there do not always predict the behaviour of the analytical solution so we use more strict bounds (based on elliptic norms [3, 8]) to bound correctly the behaviour of the continuous PDE. Numerical schemes consistent with these new bounds produce numerical solutions with the same behaviour as the analytical solution. We first derive these bounds for the continuous PDE to then extend them, using the SBP-SAT properties of the DGSEM with Gauss-Lobatto nodes, to the discretisation. Then we analyse the effectiveness of two stabilising techniques: over-integration and split form variations (conservative, non-conservative and skew-symmetric).

The paper is organised as follows: we first present a summary of the results in Section 2. We then introduce the PDE to be studied and its traditional energy bounds in Section 3. In Section 4, we derive alternative energy bounds (based on elliptic norms) for the continuous PDE. Then, in Section 5 we derive a discrete version for the DGSEM with Gauss-Lobatto nodes, with a discussion on stabilising techniques. Next, in Section 7, we present numerical experiments to show the validity of the stability bounds. Moreover, the Gauss nodes version of the DGSEM method is recovered to show that the capability to perform stabilisation by means of split formulations does not hold with that set of nodes. Lastly, we give some final remarks in Section 8.

2 Summary of Results: Continuous and Discrete Energy Bounds

Consider the first order one-dimensional initial-boundary-value problem:

$$\begin{aligned} u_t + f_x &= F, \quad 0 < x < L,\ t > 0, \\ u(x,0) &= u_0(x), \\ u(0,t) &= g_L(t), \end{aligned} \tag{1}$$

where $u(x,t)$ is the solution, $f = f(u,x) = a(x)u$ is the flux, and $F = F(u,x)$ is a source term that might depend on the solution itself, and explicitly on the space coordinate. For advection equations, we show, by an alternative bounds technique, how aliasing errors lead to numerical instabilities in the discrete solution.
Our findings are summarised in Table 1. Notice that $a(x)$ is the advection speed, whilst $I^N(a)$ refers to its interpolation to the Gauss-Lobatto points. Any exponential growth experienced by the solution energy is entirely driven by aliasing errors, which are measured with the coefficient γ. This coefficient depends on the advection speed function, $a(x)$, and the polynomial order, N, of the approximation. The parameter α controls the discretisation splitting, and must be chosen properly to cancel the aliasing errors to obtain a stable scheme. For the conservative advection equation, the value $\alpha = 1$ (i.e. conservative DG) suffices to avoid the energy growth, whilst the value $\alpha = 0$ (i.e. non-conservative DG) suffices to remove any aliasing-driven energy growth when approximating the non-conservative equation.

3 Traditional energy bounds

The problem in (1) is said to be *strongly well-posed*, if it is well-posed and the solution energy measured with the L^2 norm, $||u|| = \sqrt{\int_0^L u^2 dx}$, satisfies the bound (see [28]):

$$||u(\cdot,T)|| \leq K(T)\left(||u_0||^2 + \int_0^T \left(|g_L(\tau)|^2 + |g_R(\tau)|^2\right)d\tau\right), \tag{2}$$

Table 1 Stability of continuous advection PDE, and their discontinuous Galerkin discrete version.

PDE:	Conservative advection equation
Equation	$u_t + (au)_x = 0, \quad a(x) > 0 \; ; \quad f = a(x)u$
Split version	$u_t + \alpha(au)_x + (1-\alpha)(a_x u + a u_x) = 0$ • $\alpha = 0$: Non-conservative DG • $\alpha = 1/2$: Skew-symmetric DG • $\alpha = 1$: Conservative DG
Continuous bound	$\min\limits_{x\in[0,L]}\{a(x)\}\|\|u(T)\|\|^2 \leq \max\limits_{x\in[0,L]}\{a(x)\}\|\|u_0\|\|^2 + \int_0^T a(0)^2 g_L^2 dt$
Discrete bound	$\min\limits_{x\in[0,L]}\{I^N(a)\}\|\|U(T)\|\|^2 \leq 3 \max\limits_{x\in[0,L]}\{I^N(a)\}\|\|U_0\|\|^2 e^{2(1-\alpha)\gamma T} + \int_0^T \left(A_0^1 g_L\right)^2 dt$

PDE:	Non-conservative advection equation
Equation	$u_t + a u_x = 0, \quad a(x) > 0 \; ; \quad f_x = a(x)u_x$
Split version	$u_t + \alpha\big((au)_x - a_x u\big) + (1-\alpha)a u_x = 0$ • $\alpha = 0$: Non-conservative DG • $\alpha = 1/2$: Skew-symmetric DG • $\alpha = 1$: Conservative DG
Continuous bound	$\dfrac{\|\|u(T)\|\|^2}{\max\limits_{x\in[0,L]}\{a(x)\}} \leq \dfrac{\|\|u_0\|\|^2}{\min\limits_{x\in[0,L]}\{a(x)\}} + \int_0^T g_L^2 dt$
Discrete bound	$\dfrac{\|\|U(T)\|\|^2}{\max_{x\in[0,L]}\{I^N(a)\}} \leq \dfrac{3\|\|U_0\|\|^2}{\min_{x\in[0,L]}\{I^N(a)\}} e^{2\alpha\gamma T} + \int_0^T g_L^2 dt$

where $K(t)$ is a constant that does not depend on the spatial coordinate, x. Since the source term $F(u,x)$ depends on the solution, its contribution on the integral $\int_0^T ||F||^2 d\tau$ is not included, although this source term will contribute to shape the energy amplitude, $K(t)$. For the one-dimensional advection equation, in which $f = a(x)u$, $F = 0$ for its conservative version, and $F = a_x u$ for its non-conservative version, the precise expression for this bound was shown in [21, 6]:

$$||u(\cdot,T)||^2+\int_0^T \big(\beta u^2(0,t)+\delta u^2(L,t)\big)d\tau \leq e^{2\gamma T}\left(||u_0||^2+\int_0^T \big(\delta|g_L|^2+\beta|g_R|^2\big)d\tau\right), \tag{3}$$

where the coefficients δ and β depends on the behaviour of the left and right boundaries respectively. Hence, $(\delta,\beta)=1$ if they act as inflow, whereas $(\delta,\beta)=0$ when they act as outflow. The value of these two coefficients depend on the sign of the advection speed $a(x)$ at the boundaries. Additionally, note that the value of γ depends on the bounds of the advection speed derivative:

$$\begin{aligned} 2\gamma &= \min_{x\in[0,T]} |a_x|\,, \text{ for the conservative form,} \\ 2\gamma &= \max_{x\in[0,T]} |a_x|\,, \text{ for the non-conservative form.} \end{aligned} \tag{4}$$

In this text we will consider the PDE when $a(x)$ does not change its sign (for the sake of simplicity, we will just consider $a(x)>0$). Nevertheless, the advection speed derivative a_x is allowed to change sign within the computational domain, and γ will be positive, allowing for an exponential growth of the energy. Whether this outcome is consistent or not with the analytical solution must be determined by means of alternative energy bounds.

4 Alternative continuous bounds

In this section, alternative energy bounds that show no exponential growth will actually be seen in the continuous setting. To do so, we define the following alternative inner products and norms:

$$\langle u,v\rangle_{a^\delta} = \int_0^T a(x)^\delta u(x)v(x)dx, \quad ||u||^2_{a^\delta} = \langle u,u\rangle_{a^\delta} = \int_0 a(x)^\delta u(x)^2 dx, \tag{5}$$

where δ selects the $a-$norm ($\delta=1$), or the $1/a-$norm ($\delta=-1$). The $1/a-$norm has been previously used [3, 8] to prove stability of the Fourier collocation method. Since $a(x)$ refers to the advection speed, which has been restricted to be positive, (5) satisfies all formal definitions of a norm. Both the $a-$norm and $1/a-$norm can be related to the L^2 norm by

$$\min_{x\in[0,L]}\{a(x)\}||u||^2 \leq ||u||^2_a \leq \max_{x\in[0,L]}\{a(x)\}||u||^2, \tag{6}$$

and,

$$\frac{||u||^2}{\max\limits_{x\in[0,L]}\{a(x)\}} \leq ||u||^2_{\frac{1}{a}} \leq \frac{||u||^2}{\min\limits_{x\in[0,L]}\{a(x)\}}, \tag{7}$$

respectively. Thus, any bound measured with either the $a-$norm or the $1/a-$norm can be translated later to the L^2 norm.
We will first obtain the continuous bound for the conservative advection equation,

$$u_t + [a(x)u]_x = 0, \tag{8}$$

and that of the non-conservative equation,

$$u_t + a(x)u_x = 0 \tag{9}$$

will be derived afterwards. The domain, initial conditions and boundary conditions are those presented in (1).

4.1 Continuous bound of the conservative advection equation

We will first obtain the energy bound measured with the $a-$norm defined in (5). To do so, instead of multiplying by $u(x,t)$ to obtain the energy estimate (i.e. the traditional form to perform this estimation), (8) is multiplied by the flux $f = a(x)u$:

$$\begin{aligned} a(x)uu_t + ff_x = 0, \\ f(0,t) = a(0)g_L. \end{aligned} \tag{10}$$

To obtain the energy, (10) is integrated over the physical domain,

$$\int_0^L a(x)uu_t dx + \int_0^L ff_x dx = 0. \tag{11}$$

The first term in (11) consists of the solution energy measured with the a-norm,

$$\int_0^L a(x)uu_t dx = \frac{1}{2}\frac{d}{dt}\int_0^L a(x)u^2 dx = \frac{1}{2}\frac{d}{dt}||u||_a^2. \tag{12}$$

The second term represents the contribution of the physical boundaries to the energy,

$$\frac{1}{2}f^2\Big|_0^L = \frac{1}{2}f_R^2 - \frac{1}{2}a(0)^2 g_L^2. \tag{13}$$

The flux at the boundaries has been imposed according to the direction of the travelling physical waves. Thus, the solution energy satisfies

$$\frac{d}{dt}||u||_a^2 + f_R^2 = a(0)^2 g_L^2, \tag{14}$$

and can be bounded by

$$\frac{d}{dt}||u||_a^2 \leq a(0)^2 g_L^2. \tag{15}$$

Furthermore, time integration of (15) yields

$$||u(T)||_a^2 \leq ||u_0||_a^2 + \int_0^T a(0)^2 g_L^2 dt. \tag{16}$$

Using the relationship between the L^2 and the $a-$norm, (6), we obtain the L^2-measured energy bound:

$$\min_{x\in[0,L]}\{a(x)\}||u(T)||^2 \leq \max_{x\in[0,L]}\{a(x)\}||u_0||^2 + \int_0^T a(0)^2 g_L^2 dt. \tag{17}$$

Thus, no exponential growth is expected in the analytical solution. The numerical scheme should mimic a discrete version this behaviour.

4.2 Continuous bound of the non-conservative advection equation

Regarding the non-conservative generic advection problem, its alternative energy estimate is obtained multiplying (9) by u, and then dividing by $a(x)$

$$\begin{aligned} \frac{uu_t}{a(x)} + uu_x &= 0, \\ u(0,t) &= g_L, \end{aligned} \tag{18}$$

which is nonsingular since $a(x) > 0$. Following (10), (18) is integrated over the physical domain

$$\int_0^L \frac{uu_t}{a(x)} dx + \int_0^L uu_x dx = 0. \tag{19}$$

As in the conservative case, the first term is the solution energy, measured with the $1/a$-norm defined in (5) (with $\delta = -1$):

$$\frac{1}{2}\frac{d}{dt}\int_0^L \frac{u^2}{a(x)} dx = \frac{1}{2}\frac{d}{dt}||u||_{\frac{1}{a}}^2, \tag{20}$$

The second term of (19) can be integrated by parts to show the energy entering and leaving the domain through the physical boundaries

$$\frac{1}{2}u^2\Big|_0^L = -\frac{1}{2}g_L^2 + \frac{1}{2}u(L,t)^2. \tag{21}$$

Putting it all together, the energy in terms of the $1/a-$norm varies as

$$\frac{d}{dt}||u||_{\frac{1}{a}}^2 + u(L,t)^2 = g_L^2, \tag{22}$$

which when time integrated leads to the bound

$$||u(T)||_{\frac{1}{a}}^2 \leq ||u_0||_{\frac{1}{a}}^2 + \int_0^T g_L^2 dt. \tag{23}$$

Finally, using the relationship between the L^2 and the $1/a-$norm, (7), we get the energy bound

$$\frac{||u(T)||^2}{\max\limits_{x\in[0,L]}\{a(x)\}} \leq \frac{||u_0||^2}{\min\limits_{x\in[0,L]}\{a(x)\}} + \int_0^T g_L^2 dt, \tag{24}$$

which proves that the energy at any time is bounded, and does not exhibit any exponential growth. As before, the numerical scheme should mimic a discrete version of this behaviour.

4.3 Summary

Whilst traditional L^2 energy bounds show energy growth in general, $\gamma > 0$, the more strict $a-$norm estimate shows no growth when $a(x) > 0$ (or equivalently $a(x) < 0$) in both conservative and non-conservative forms of the equation. This result encourages us to use this approach to review the energy bounds of the discontinuous Galerkin approximation. Deriving the discrete version of the bounds, and bearing in mind that the final goal is to derive DG schemes that follow the continuous bound, gives us insight into the effect of the numerical errors on stability. We note in passing that alternative methods to obtain the energy bounds exist. The von Neumann analysis for non-constant advection speeds developed in [22] is also a powerful tool to understand the behaviour of numerical schemes for periodic problems.

5 Discrete DGSEM-GL version of the bounds

In this work, we consider the discontinuous Galerkin formulation with Gauss-Lobatto points. The scheme satisfies all the formal definitions of a summation-by-parts (SBP) simultaneous-approximation-term (SAT) scheme. This allows the removal of aliasing errors by means of split formulations (see [15, 14, 7, 17]), and therefore, to obtain schemes that are provably stable.

We will first introduce the notation used throughout this work. The physical domain $\Omega = \{x \mid x \in [0, L]\}$ is divided in K non-overlapping elements $\Omega = \{x \mid x \in [x^k, x^{k+1}]\}$, in which the solution is approximated by N degree polynomials (they will be said to belong to the $\mathcal{P}^N$ space). The interpolation is performed at the Gauss-Lobatto points $\{\xi_j\}_{j=0}^N$, which are then mapped individually from the local coordinate $\xi \in [-1, 1]$ frame to each element domain by means of a linear one-dimensional mapping, $X^{el}(\xi)$

$$x_j^{el} = X^{el}(\xi_j) = x^{el} + \frac{1}{2}(x^{el+1} - x^{el})(\xi_j + 1). \tag{25}$$

We adopt the system used in [6] where capital symbols refer to the interpolated version of each variable. For instance, for the solution,

$$U^{el}(\xi) = I^N[u(X^{el}(\xi))] = \sum_{j=0}^{N} U_j^{el} l_j(\xi), \quad U_j^{el} = u(x_j^{el}), \tag{26}$$

and for the fluxes:

$$F^{el}(\xi) = I^N[f(X^{el}(\xi))] = \sum_{j=0}^{N} F_j^{el} l_j(\xi), \quad F_j^{el} = f(x_j^{el}) = a(x_j^{el}) U_j^{el}. \tag{27}$$

We will also obtain the discrete version of the advection speed, $a(x)$, $A^{el}(\xi)$ as

$$A^{el}(\xi) = I^N[a(X^{el})] = \sum_{j=0}^{N} A_j^{el} l_j(\xi), \quad A_j^{el} = a(x_j^{el}). \tag{28}$$

Note that because the Gauss-Lobatto points include the endpoints, the discrete version of the advection speed, $A^{el}(\xi)$, will remain continuous across the inter-element interfaces (assuming that $a(x)$ is continuous).

The polynomial space is spanned by the Lagrange polynomials, $l_j(\xi)$, with nodes at the Gauss-Lobatto points

$$l_j(\xi) = \prod_{\substack{m=0 \\ m \neq j}}^{N} \frac{\xi - \xi_m}{\xi_j - \xi_m}. \tag{29}$$

This basis allows one to compute the required derivatives by means of a derivative matrix, $[\boldsymbol{D}]$, defined as

$$D_{ij} = l_j'(\xi_i). \tag{30}$$

We also adopt the following matrix-vector form: A vector contains the nodal degrees of freedom of a certain variable, for instance $\{\underline{\boldsymbol{A}}^{el}\}$. The notation $[\boldsymbol{A}^{el}] = \text{diag}(\{\underline{\boldsymbol{A}}^{el}\}))$ represents the diagonal matrix whose diagonal entries are those nodal values. This allows us to compactly write the following terms, usually arising from the split formulation, as

$$\begin{aligned} I^N[A^{el} U_\xi^{el}] &= \{\underline{\boldsymbol{l}}^{el}(\xi)\}^T [\boldsymbol{A}^{el}][\boldsymbol{D}^{el}]\{\underline{\boldsymbol{U}}^{el}\}, \\ I^N[A_\xi^{el} U^{el}] &= \{\underline{\boldsymbol{l}}^{el}(\xi)\}^T [\boldsymbol{A}_\xi^{el}]\{\underline{\boldsymbol{U}}^{el}\}, \quad [\boldsymbol{A}_\xi^{el}] = \text{diag}([\boldsymbol{D}^{el}]\{\underline{\boldsymbol{A}}^{el}\}) \\ (I^N[A^{el} U^{el}])_\xi &= \{\underline{\boldsymbol{l}}^{el}(\xi)\}^T [\boldsymbol{D}^{el}][\boldsymbol{A}^{el}]\{\underline{\boldsymbol{U}}^{el}\}. \end{aligned} \tag{31}$$

We will discretise the advection equation, either in conservative or non-conservative form, by means of a general split formulation for the flux, $f = au$. The equation solved depends a the parameter θ, namely

$$u_t + f_x = \theta a_x u. \tag{32}$$

Setting $\theta = 0$ solves the conservative equation, while the case $\theta = 1$ recovers the non-conservative equation. The split form is described in [6],

$$u_t + \alpha f_x + (1 - \alpha)(a_x u + a u_x) = \theta a_x u. \tag{33}$$

Equation (33) is multiplied by a test function, $\Phi \in \mathcal{P}^N$, and integrated in each element to get the weak form

$$\begin{aligned} &\frac{\Delta x^{el}}{2}\left\langle \Phi^{el}, u_t \right\rangle + \alpha \left\langle \Phi^{el}, f_\xi \right\rangle \\ &+(1-\alpha)\big(\left\langle \Phi^{el}, a_\xi u \right\rangle + \left\langle a u_\xi, \Phi^{el} \right\rangle\big) = \theta \left\langle \Phi^{el}, a_\xi u \right\rangle, \end{aligned} \tag{34}$$

where the inner product is $\langle u, v \rangle = \int_{-1}^{1} uv d\xi$.

The second and fourth terms of (34) are integrated by parts, and the interface fluxes that appear are replaced by a numerical flux $f^\star$:

$$\begin{aligned} &\frac{\Delta x^{el}}{2}\left\langle \Phi^{el}, u_t \right\rangle + f^\star \Phi^{el}\Big|_{-1}^{1} - \alpha \left\langle \Phi^{el}_\xi, f \right\rangle \\ &+(1-\alpha)\big(\left\langle \Phi^{el}, a_\xi u \right\rangle - \left\langle u, (a\Phi^{el})_\xi \right\rangle\big) = \theta \left\langle \Phi^{el}, a_\xi u \right\rangle. \end{aligned} \tag{35}$$

Lastly, inner product integrals are computed with numerical quadratures (i.e. Gauss-Lobatto points). Thus, we replace the inner products $\langle \cdot, \cdot \rangle$ by their numerical version $\langle \cdot, \cdot \rangle_N$, and the arguments by their polynomial approximations

$$\begin{aligned} &\frac{\Delta x^{el}}{2}\left\langle \Phi^{el}, U^{el}_t \right\rangle_N + F^\star \Phi^{el}\Big|_{-1}^{1} - \alpha \left\langle \Phi^{el}_\xi, F^{el} \right\rangle_N \\ &+(1-\alpha)\big(\left\langle \Phi^{el}, I^N[A^{el}_\xi U^{el}] \right\rangle_N - \left\langle U^{el}, (I^N[A^{el}\Phi^{el}])_\xi \right\rangle_N\big) = \theta \left\langle \Phi^{el}, I^N[A^{el}_\xi U^{el}] \right\rangle_N. \end{aligned} \tag{36}$$

Using $N+1$ linearly independent test functions (e.g. the Lagrange polynomials), we obtain the differential equations for the solution degrees of freedom, $U^{el}_j(t)$. Henceforth, for the sake of simplicity, we will drop the $el-$ superscript.

The discrete formulation is selected by choosing the split form coefficient, α, where $\alpha = 0$ gives the non-conservative DG, $\alpha = 1/2$ gives the skew-symmetric DG, and $\alpha = 1$ gives the conservative DG. It is also possible to switch between the two forms of the equation through the parameter θ, where with $\theta = 0$ we solve the conservative equation, whilst with $\theta = 1$ we recover the non-conservative equation.

Now we proceed to obtain the discrete version of the continuous bounds in (17) and (24). To do so, we define the discrete version of the $a-$norm, the $a-$inner product, the $1/a$-norm, and the $1/a-$inner product as

$$\begin{aligned} \langle U, V \rangle_{a^\delta, N} &= \sum_{m=0}^{N} w_m A^\delta_m U_m V_m = \{\underline{\boldsymbol{U}}\}^T [\boldsymbol{M}][\boldsymbol{A}^\delta]\{\underline{\boldsymbol{V}}\}, \text{ and} \\ ||U||^2_{a^\delta} = \langle U, U \rangle_{a^\delta, N} &= \sum_{m=0}^{N} w_m A^\delta_m U^2_m = \{\underline{\boldsymbol{U}}\}^T [\boldsymbol{M}][\boldsymbol{A}^\delta]\{\underline{\boldsymbol{U}}\}, \end{aligned} \tag{37}$$

where $\delta = 1$ refers to the $a-$norm, and $\delta = -1$ refers to the $1/a-$norm. In (37), the w_m are the Gauss-Lobatto quadrature weights, and $[\boldsymbol{M}]$ is the (lumped) mass matrix, whose entries are the quadrature weights placed along

the main diagonal (see [33]). The two discrete norms can be also related to the continuous L^2 norm. First, the discrete $a^\delta-$ norm can be related to the Gauss-Lobatto discrete norm by

$$\min\{A_j\}_{j=0}^N ||U||_N^2 \leq ||U||_{a,N}^2 \leq \max\{A_j\}_{j=0}^N ||U||_N^2, \tag{38}$$

and related to the L^2 norm (see Section 5.3 in [4]) as

$$\min\{A_j\}_{j=0}^N ||U||^2 \leq ||U||_{a,N}^2 \leq 3\max\{A_j\}_{j=0}^N ||U||^2. \tag{39}$$

equation (39) is valid as long as $A(\xi) > 0$ in each element. Similarly, for the $1/a-$norm:

$$\frac{||U||^2}{\max\{A_j\}_{j=0}^N} \leq ||U||_{\frac{1}{a},N}^2 \leq \frac{3||U||^2}{\min\{A_j\}_{j=0}^N}. \tag{40}$$

Finally, the a^δ norm and inner products are extended to the whole domain by summing all the elemental contributions

$$\begin{aligned} \langle U, V\rangle_{a^\delta,N} &= \frac{\Delta x}{2}\sum_{el=1}^K \left\langle U^{el}, V^{el}\right\rangle_{a^\delta,N}, \\ ||U||_{a^\delta}^2 &= \frac{\Delta x}{2}\sum_{el=1}^K ||U^{el}||_{a^\delta}^2, \end{aligned} \tag{41}$$

where Δx is the size (length) of the element. Recall that local norms are computed without multiplying by Δx, whilst global norms are constructed weighting local norms with Δx.

5.1 Discrete bound of the conservative advection equation

We set $\theta = 0$ to recover the conservative equation and rearrange (36) as the conservative or standard DG (see [5]), plus a correction term that arises from the split formulation

$$\begin{aligned} &\frac{\Delta x}{2}\langle \Phi, U_t\rangle_N + F^\star \Phi\Big|_{-1}^1 - \langle \Phi_\xi, F\rangle_N \\ &+(1-\alpha)\big(\langle \Phi_\xi, F\rangle_N + \left\langle \Phi, I^N[A_\xi U]\right\rangle_N - \left\langle U, (I^N[A\Phi])_\xi\right\rangle_N\big) = 0. \end{aligned} \tag{42}$$

Hence, the precise form of the correction term is

$$\langle \Phi_\xi, F\rangle_N + \left\langle \Phi, I^N[A_\xi U]\right\rangle_N - \left\langle U, (I^N[A\Phi])_\xi\right\rangle_N = N_1 + N_2 + N_3. \tag{43}$$

To derive an energy bound, similar to its continuous counterpart, the test function is replaced by the polynomial approximation of the fluxes, $\Phi = F$

$$\begin{aligned} &\frac{\Delta x}{2}\langle F, U_t\rangle_N + F^\star F\Big|_{-1}^1 - \langle F_\xi, F\rangle_N \\ &+(1-\alpha)\big(\langle F_\xi, F\rangle_N + \left\langle F, I^N[A_\xi U]\right\rangle_N - \left\langle U, (I^N[AF])_\xi\right\rangle_N\big) = 0. \end{aligned} \tag{44}$$

The first term in (44) reproduces the time derivative of the discrete energy measured with the $a-$norm,

$$\langle F, U_t \rangle_N = \langle AU, U_t \rangle_N = \langle U, U_t \rangle_{a,N} = \frac{1}{2}\frac{d}{dt}||U||^2_{a,N}. \tag{45}$$

Regarding the third term in (44), the summation-by-parts property (see [25]) holds, and thus

$$\langle F_\xi, F \rangle_N = \frac{1}{2}\big(F\big)^2\Big|_{-1}^{1}. \tag{46}$$

Lastly, in the correction term defined in (43), both the first and third term can be rewritten (following the summation-by-parts property) as

$$\begin{aligned} N_1 + N_3 &= \langle F_\xi, F \rangle_N - \big\langle U, (I^N[AF])_\xi \big\rangle_N \\ &= -\langle F_\xi, F \rangle_N + \big\langle U_\xi, I^N[AF] \big\rangle_N , \end{aligned} \tag{47}$$

in which boundary terms arising from the summation-by-parts cancel so only volume integrals contribute to the estimate. All three terms involved in the correction term can be converted to the $a-$inner product. The first is

$$-\langle F_\xi, F \rangle_N = -\langle F_\xi, U \rangle_{a,N} = -\Big\langle \big(I^N[AU]\big)_\xi, U \Big\rangle_{a,N}. \tag{48}$$

The second is

$$\big\langle F, I^N[A_\xi U] \big\rangle_N = \big\langle U, I^N[A_\xi U] \big\rangle_{a,N}, \tag{49}$$

and the third becomes

$$\big\langle U_\xi, I^N[AF] \big\rangle_N = \big\langle I^N[AU_\xi], U] \big\rangle_{a,N}. \tag{50}$$

Therefore, the correction term, $\mathcal{L}(A,U)$, is

$$\mathcal{L}(A,U) = \Big\langle I^N[A_\xi U] + I^N[AU_\xi] - \big(I^N[AU]\big)_\xi, U \Big\rangle_{a,N}, \tag{51}$$

which does not vanish since the product derivative rule does not have a discrete equivalent. This term represents the aliasing errors introduced in the discrete weak formulation of the original equation, which are projected (with the a-norm inner product) onto the solution. We can bound these errors, since, using the matrix form shown in (31) we can write the inner product as

$$\mathcal{L}(A,U) = \{\underline{\boldsymbol{U}}\}^T[\boldsymbol{A}][\boldsymbol{M}]\big([\boldsymbol{A}_\xi] + [\boldsymbol{A}][\boldsymbol{D}] - [\boldsymbol{D}][\boldsymbol{A}]\big)\{\underline{\boldsymbol{U}}\}. \tag{52}$$

Thus, we can use the Cauchy-Schwartz inequality using the $a-$norm to bound the aliasing term as

$$\mathcal{L}(A,U) \leq \gamma ||U||^2_{a,N}, \tag{53}$$

where the energy growth coefficient, γ, is defined as:

$$\gamma = \big\| [\boldsymbol{A}][\boldsymbol{M}]\big([\boldsymbol{A}_\xi] + [\boldsymbol{A}][\boldsymbol{D}] - [\boldsymbol{D}][\boldsymbol{A}]\big) \big\|_{a,N}, \tag{54}$$

and the matrix $a-$norm is induced by the vector norm:

$$\|\boldsymbol{M}\|_{a,N} = \max_{\|\boldsymbol{x}\|_{a,N}=1} \frac{\|\boldsymbol{M}\cdot\boldsymbol{x}\|_{a,N}}{\|\boldsymbol{x}\|_{a,N}}. \tag{55}$$

The coefficient γ depends on the advection speed, $A(\xi)$, and the polynomial degree, N. Gathering all the terms together, we get the elemental contribution to the time derivative of the energy,

$$\frac{\Delta x}{2}\frac{d}{dt}||U^{el}||^2_{a,N} + \big(2F^\star - F^{el}\big)F^{el}\Big|_{-1}^{1} \leq 2(1-\alpha)\gamma^{el}||U^{el}||^2_{a,N}. \tag{56}$$

Note that now we specifically include the $el-$ index. Next, (56) is summed over all elements, to obtain the total energy,

$$\frac{d}{dt}||U||^2_{a,N} + \sum_{el=1}^{K}\big(2F^\star - F^{el}\big)F^{el}\Big|_{-1}^{1} \leq 2(1-\alpha)\gamma||U||^2_{a,N}, \tag{57}$$

where γ is bounded by the largest value over all the elements,

$$\gamma = \max_{el}\frac{\gamma^{el}}{\Delta x^{el}/2}. \tag{58}$$

The interior interface contributions to the total energy vanish as long as central fluxes are considered (see [6]), whereas the physical boundary contributions, computed with upwind fluxes, are

$$\sum_{el=1}^{K}\big(2F^\star - F^{el}\big)F^{el}\Big|_{-1}^{1} = \big(F^K(1)\big)^2 - \big(2A_0^1 g_L - F^1(-1)\big)F^1(-1). \tag{59}$$

The upwind flux stabilises the inflow condition, since

$$\big(2A_0^1 g_L - F^1(-1)\big)F^1(-1) = \big(A_0^1 g_L\big)^2 - \big(A_0^1 g_L - F^1(-1)\big)^2. \tag{60}$$

Thus, we get the energy estimate

$$\frac{d}{dt}||U||^2_{a,N} + \big(A_0^1 g_L - F^1(-1)\big)^2 + \big(F^K(1)\big)^2 \leq 2(1-\alpha)\gamma||U||^2_{a,N} + \big(A_0^1 g_L\big)^2, \tag{61}$$

with the upper bound

$$\frac{d}{dt}||U||^2_{a,N} \leq 2(1-\alpha)\gamma||U||^2_{a,N} + \big(A_0^1 g_L\big)^2. \tag{62}$$

Time integration of (62) yields

$$||U(T)||^2_{a,N} \leq ||U_0||^2_{a,N} e^{2(1-\alpha)\gamma T} + \int_0^T \big(A_0^1 g_L\big)^2 dt, \tag{63}$$

which is related to the continuous L_2 norm using (39) as

$$\min\{A_j\}_{j=0}^{N}||U(T)||^2 \leq 3\max\{A_j\}_{j=0}^{N}||U_0||^2 e^{2(1-\alpha)\gamma T} + \int_0^T \left(A_0^1 g_L\right)^2 dt. \quad (64)$$

As a conclusion, the final outcome is that aliasing errors drive the instability, causing an exponential growth. Choosing the correct value of the parameter α makes it possible to remove those errors from the energy estimate. Precisely, for the conservative equation, the parameter α should be equal to 1, that is, a conservative DG scheme. For this last case ($\alpha = 1$), we have shown that neither the numerical solution nor the physical solution will exhibit exponential growth.

5.2 Discrete bound of the non-conservative advection equation

We now switch θ in (32) to 1 to get the non-conservative equation. As with (42), we rearrange (36) to be regarded as the non-conservative or standard DG (see [5]), plus a correction term that arises from the split formulation,

$$\begin{aligned} &\frac{\Delta x}{2}\left\langle \Phi, U_t\right\rangle_N + F^\star \Phi\Big|_{-1}^{1} - \left\langle \Phi_\xi, F\right\rangle_N = \left\langle \Phi, I^N[A_\xi U]\right\rangle_N \\ &-(1-\alpha)\big(\left\langle \Phi_\xi, F\right\rangle_N + \left\langle \Phi, I^N[A_\xi U]\right\rangle_N - \left\langle U, (I^N[A\Phi])_\xi\right\rangle_N\big). \end{aligned} \quad (65)$$

In (65), the correction term is defined as

$$\left\langle \Phi_\xi, F\right\rangle_N + \left\langle \Phi, I^N[A_\xi U]\right\rangle_N - \left\langle U, (I^N[A\Phi])_\xi\right\rangle_N = N_1 + N_2 + N_3. \quad (66)$$

This time the discrete energy bound is derived by replacing the test function by the interpolation of the quotient between the solution and the advection speed, $\Phi = I^N[U/A]$ to get

$$\begin{aligned} &\frac{\Delta x}{2}\left\langle I^N\left[\frac{U}{A}\right], U_t\right\rangle_N + F^\star\left[\frac{U}{A}\right]\Big|_{-1}^{1} + \left\langle U, (I^N[AI^N\left[\frac{U}{A}\right]])_\xi\right\rangle_N = \\ &+\alpha(\left\langle \left(I^N\left[\frac{U}{A}\right]\right)_\xi, F\right\rangle_N + \left\langle I^N\left[\frac{U}{A}\right], I^N[A_\xi U]\right\rangle_N - \left\langle U, (I^N[AI^N\left[\frac{U}{A}\right]])_\xi\right\rangle_N), \end{aligned} \quad (67)$$

where the correction term has been rearranged. The first term in (67) reproduces the time derivative of the discrete energy measured with the $1/a-$norm:

$$\left\langle I^N\left[\frac{U}{A}\right], U_t\right\rangle_N = \langle U, U_t\rangle_{\frac{1}{a},N} = \frac{1}{2}\frac{d}{dt}||U||^2_{\frac{1}{a},N}. \quad (68)$$

We can use the summation-by-parts property for the third term in (67) to write the volume term in terms of surface quantities

$$\left\langle U, (I^N[AI^N\left[\frac{U}{A}\right]])_\xi\right\rangle_N = \langle U, U_\xi\rangle_N = \frac{1}{2}(U)^2\Big|_{-1}^{1}. \quad (69)$$

The first and third terms of the correction term defined in (66) can be rewritten (following the summation-by-parts property) as

$$\begin{aligned} N_1 + N_3 &= \left\langle \left(I^N\left[\frac{U}{A}\right]\right)_\xi, F \right\rangle_N - \left\langle U, (I^N[AI^N\left[\frac{U}{A}\right]])_\xi \right\rangle_N \\ &= \left\langle I^N\left[\frac{U}{A}\right], AU_\xi \right\rangle_N - \left\langle I^N\left[\frac{U}{A}\right], F_\xi \right\rangle_N . \end{aligned} \tag{70}$$

Again, the three terms in the correction term can be written in terms of the $1/a-$inner product. We write the first as

$$-\left\langle I^N\left[\frac{U}{A}\right], F_\xi \right\rangle_N = -\left\langle U, F_\xi \right\rangle_{\frac{1}{a},N}, \tag{71}$$

the second as

$$\left\langle I^N\left[\frac{U}{A}\right], I^N[A_\xi U] \right\rangle_N = \left\langle U, I^N[A_\xi U] \right\rangle_{\frac{1}{a},N}, \tag{72}$$

and the third as

$$\left\langle I^N\left[\frac{U}{A}\right], AU_\xi \right\rangle_N = \left\langle U, AU_\xi \right\rangle_{\frac{1}{a},N}. \tag{73}$$

Therefore, the correction term, $\mathcal{L}(A,U)$, is

$$\mathcal{L}(A,U) = \left\langle I^N[A_\xi U] + I^N[AU_\xi] - \left(I^N[AU]\right)_\xi, U \right\rangle_{\frac{1}{a},N}, \tag{74}$$

which again does not vanish since the product derivative rule does not have a discrete equivalent. This term represents the aliasing errors incurred in the discrete weak formulation of the original equation. Moreover, following a similar approach to (53), it can be bounded as

$$\mathcal{L}(A,U) \leq \gamma ||U||^2_{\frac{1}{a},N}, \tag{75}$$

where the definition of γ is now induced by the $1/a-$norm

$$\gamma = \left\| [\boldsymbol{A}]^{-1}[\boldsymbol{M}]\left([\boldsymbol{A}_\xi] + [\boldsymbol{A}][\boldsymbol{D}] - [\boldsymbol{D}][\boldsymbol{A}]\right) \right\|_{\frac{1}{a},N}. \tag{76}$$

Putting it all together, the estimate of the time derivative of the energy inside each element (where the $el-$ index is again explicitly written) is

$$\frac{\Delta x}{2}\frac{d}{dt}||U^{el}||^2_{\frac{1}{a},N} + \left(2U^\star - U^{el}\right)U^{el}\Big|_{-1}^{1} \leq 2\alpha\gamma^{el}||U^{el}||^2_{\frac{1}{a},N}, \tag{77}$$

where $U^\star$ is the inter-element flux divided by the inter-element advection speed. Next, (77) is summed across all elements, to obtain the time derivative of the total energy,

$$\frac{d}{dt}||U||^2_{\frac{1}{a},N} + \sum_{el=1}^{K}\left(2U^\star - U^{el}\right)U^{el}\Big|_{-1}^{1} \leq 2\alpha\gamma^{el}||U^{el}||^2_{\frac{1}{a},N}. \tag{78}$$

The contribution to the total energy of the interior interfaces vanishes when central fluxes are used (see [6]). When the physical boundary contributions are computed with upwind fluxes, we get terms like those in (59) and (60). Thus,

$$\frac{d}{dt}||U||^2_{\frac{1}{a},N} + \left(g_L - U^1(-1)\right)^2 + \left(U^K(1)\right)^2 \leq 2\alpha\gamma||U||^2_{\frac{1}{a},N} + \left(g_L\right)^2, \tag{79}$$

which has an upper bound

$$\frac{d}{dt}||U||^2_{\frac{1}{a},N} \leq 2\alpha\gamma||U||^2_{\frac{1}{a},N} + \left(g_L\right)^2. \tag{80}$$

The energy at $t = T$ is obtained integrating (80) in time

$$||U(T)||^2_{\frac{1}{a},N} \leq ||U_0||^2_{\frac{1}{a},N} e^{2\alpha\gamma T} + \int_0^T g_L^2 dt, \tag{81}$$

which can be related to the continuous L_2 norm using (39),

$$\frac{||U(T)||^2}{\max\{A_j\}_{j=0}^N} \leq \frac{3||U_0||^2}{\min\{A_j\}_{j=0}^N} e^{2\alpha\gamma T} + \int_0^T g_L^2 dt. \tag{82}$$

Therefore, we can conclude that, like in the conservative equation approximation, any exponential growth experienced by the numerical solution is due to aliasing errors in the flux interpolation. Notwithstanding, we can remove these errors by choosing the precise split form coefficient. In particular, the parameter α should be set to zero (i.e. a non-conservative discretisation, $\alpha = 0$) to follow the behaviour of the analytical solution.

5.3 Summary of alternative discrete energy bounds

We have obtained the discrete versions of the continuous bounds derived in Section 4 for both conservative and non-conservative DG. These bounds have been derived using a split-form discontinuous Galerkin method with Gauss-Lobatto points. These alternative bounds make it possible to analyse the aliasing error removal by means of the split operator. These aliasing errors, which arise from the fact that the product rule does not hold discretely, drive an exponential energy growth, which can be controlled with the parameter α. Precisely, when solving the conservative equation, it suffices to select $\alpha = 1$ (i.e. conservative DG), whilst selecting $\alpha = 0$ (non-conservative DG) when solving the non-conservative equation is enough to prevent aliasing driven exponential growth. It should be noticed that the alternative bounds derived in Section 4 are only valid for advection speeds $a(x)$ that do not change sign in the domain. In the general case, strongly-stable schemes are recovered for $\alpha = 1/2$, as shown in [6].

6 Effect of over-integration

Over-integration, also known as polynomial de-aliasing or consistent integration, is a technique commonly used to reduce aliasing in non-linear equations [30]. In this work, we apply this methodology to the non-constant speed advection equation and find that, despite being capable of removing the non-linear fluxes' aliasing errors, it introduces additional instabilities through the discrete weak form.

Since we are considering the non-constant speed advection equation, the discrete flux is the product of two $N-$degree polynomials, once the generally non-polynomial advection speed, $a(x)$, has been projected onto the solution space. Therefore, we may change the weak formulation to

$$\frac{\Delta x}{2}\left\langle \Phi, U_t\right\rangle_N + F^\star \Phi\Big|_{-1}^{1} - \left\langle \Phi_\xi, F\right\rangle_M = 0, \tag{83}$$

where F is now a polynomial of degree $2N$ (since $F = AU$), and M is the required number of quadrature points. For the Gauss-Lobatto points, $M > 3N/2$ is required to avoid inexact quadratures. Also, note that any of the split forms we have considered are algebraically identical to the conservative formulation, since all volume integrals are approximated and computed exactly.

To compute the energy estimate, we replace Φ by $I^N[F]$ since the test function should belong to the solution space, $\mathcal{P}^N$,

$$\frac{\Delta x}{2}\left\langle I^N[F], U_t\right\rangle_N + F^\star F\Big|_{-1}^{1} - \left\langle \big(I^N[F]\big)_\xi, F\right\rangle_M = 0. \tag{84}$$

The interpolation operator $I^N[\bullet]$ has been omitted for the surface integral since Gauss-Lobatto points include the endpoints, and therefore both are identical (i.e. interpolation is not necessary at the boundaries). Following (45), we write the first term in (84) as the time derivative of the energy, measured with the $a-$norm, whereas the last term is summed-by-parts

$$\begin{aligned}\left\langle \big(I^N[F]\big)_\xi, F\right\rangle_M &= \left\langle F_\xi, F\right\rangle_M + \left\langle \big(I^N[F]\big)_\xi - F_\xi, F\right\rangle_M \\ &= \frac{1}{2}F\Big|_{-1}^{+1} + \left\langle \big(I^N[F]\big)_\xi - F_\xi, F\right\rangle_M.\end{aligned} \tag{85}$$

Thus, the energy bound reads

$$\frac{\Delta x}{2}\frac{d}{dt}||U||_{a,N}^2 + \big(2F^\star - F\big)F\Big|_{-1}^{1} = \left\langle \big(F - I^N[F]\big)_\xi, F\right\rangle_M. \tag{86}$$

Which summed over all mesh elements gives

$$\frac{d}{dt}||U||_{a,N}^2 \leq (A_0^N g_L)^2 + \left\langle \big(F - I^N[F]\big)_\xi, F\right\rangle_M. \tag{87}$$

The last term on the right hand side represents the aliasing errors incurred when using the interpolant of the flux as the test function. Therefore, computing the integrals exactly may not stabilise the scheme in this norm.

7 Numerical experiments

We will show the accuracy of the bounds in (64), (82), and (87) by examining the eigenvalues of the numerical implementations of the advection equation. This particular example considers the domain $x \in [-1, 1]$ in which the advection speed is

$$a(x) = 1 + (1 - x^2)^5, \tag{88}$$

and with periodic boundary conditions at the two endpoints. This example (extracted from Hesthaven and Warburton's book [23]) is interesting since it contains large spectral content.

We use the split discontinuous Galerkin method shown in (36) to solve both conservative and non-conservative equations. Moreover, we use the central fluxes,

$$F^\star(u_L, u_R) = \frac{a_L u_L + a_R u_R}{2}, \tag{89}$$

for both interior and physical boundaries, since we have enforced periodic boundary conditions. As a consequence of performing the interpolation with Gauss-Lobatto points, the discrete version of $a(x)$ will be continuous, and $a_L = a_R$, equal to the advection speed at the boundary.

Replacing the test function Φ^{el} by the set of Lagrange polynomials $\{l_j\}_{j=0}^N$ in (36) and computing the inner products, one arrives to the following matrix system for each element

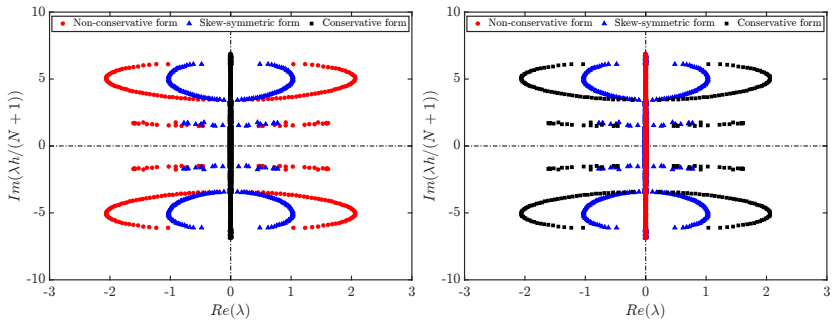
$$\frac{\Delta x^{el}}{2}[\boldsymbol{M}]\frac{d\{\underline{\boldsymbol{U}}^{el}\}}{dt} = [\boldsymbol{L}^{el}]\{\underline{\boldsymbol{U}}^{el-1}\} + [\boldsymbol{C}^{el}]\{\underline{\boldsymbol{U}}^{el}\} + [\boldsymbol{R}^{el}]\{\underline{\boldsymbol{U}}^{el+1}\}, \tag{90}$$

where the matrices $[\boldsymbol{L}^{el}]$, $[\boldsymbol{C}^{el}]$, and $[\boldsymbol{R}^{el}]$ are constant, i.e. they do not depend on $\{\underline{\boldsymbol{U}}\}$, and they just depend on the polynomial order and the discrete advection speed in each element, $A^{el}(\xi)$. Their precise expression is

$$\begin{aligned}
[\boldsymbol{L}^{el}] &= \frac{a^{el-1}}{2}\{\underline{\boldsymbol{l}}(-1)\}\{\underline{\boldsymbol{l}}(1)\}^T, \\
[\boldsymbol{C}^{el}] &= \frac{a^{el-1}}{2}\{\underline{\boldsymbol{l}}(1)\}\{\underline{\boldsymbol{l}}(1)\}^T - \frac{a^{el}}{2}\{\underline{\boldsymbol{l}}(-1)\}\{\underline{\boldsymbol{l}}(-1)\}^T \\
&\quad + \alpha[\boldsymbol{D}]^T[\boldsymbol{M}][\boldsymbol{A}^{el}] - (1-\theta-\alpha)[\boldsymbol{A}^{el}_\xi] + (1-\alpha)[\boldsymbol{A}^{el}][\boldsymbol{D}]^T[\boldsymbol{M}], \\
[\boldsymbol{R}^{el}] &= -\frac{a^{el}}{2}\{\underline{\boldsymbol{l}}(1)\}\{\underline{\boldsymbol{l}}(-1)\}^T,
\end{aligned} \tag{91}$$

where a^{el-1} is the advection speed to the left of the element boundary, and a^{el} is that to the right. Recall that $\theta = 0$ generates the conservative equation, whilst $\theta = 1$ generates the non-conservative equation.

The stability of the scheme is studied using the system of eigenvalues over all elements. We build a mesh with $K = 200$ elements and polynomial order $N = 5$. We have considered the three relevant cases for the split operator



(a) Conservative equation spectra (b) Non-conservative equation spectra

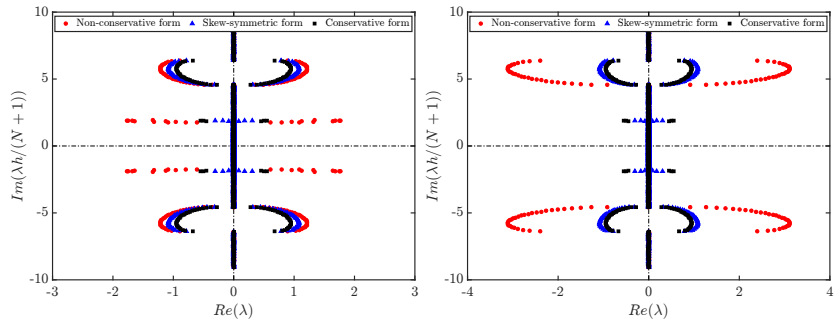
Fig. 1 Eigenvalues of the non-constant speed (88) advection equation using Gauss-Lobatto points. The three relevant versions (i.e. conservative ($\alpha = 1$), skew-symmetric ($\alpha = 1/2$), and non-conservative ($\alpha = 0$) DG) of the split operator coefficient are represented. In this test case, $K = 200$ elements have been used, while the polynomial order is $N = 5$.

coefficient: conservative DG ($\alpha = 1$), skew-symmetric DG ($\alpha = 1/2$), and non-conservative DG ($\alpha = 0$).

The eigenvalues for the conservative equation ($\theta = 0$) are depicted in Figure 1(a). The $x-$axis represents the real part of the spectra (where positive real parts lead to energy growth), whilst the $y-$axes represents their imaginary part. The imaginary part has been scaled with the element sizes and the polynomial order as in [29]. We see that both the non-conservative and skew-symmetric DG are unstable, whereas the conservative DG formulation is stable, consistent with the bound derived in (64). Note that the conservative DG spectra lies on the imaginary axis, and thus, none of the modes will exhibit energy growth or decay. On the other hand, when $\alpha \neq 1$, all modes are arranged by pairs, of which one of the pair shows exponential energy growth, and the other decay.

In Figure 1(b), the non-conservative equation is analysed. We observe that the results obtained are the opposite to the conservative equation ones, as shown by (82). In this case, the non-conservative DG discretisation is stable since its set of eigenvalues lie on the imaginary axes, which is consistent with the original PDE behaviour described in (24).

To show the role played by the Gauss-Lobatto points, the spectra obtained using Gauss points is included in Figure 2 for both the conservative and non-conservative problems. Recall that DG schemes based on Gauss points do not satisfy the summation-by-parts with simultaneous-approximation-term property (SBP-SAT, see [15]), and thus, the energy estimates that we have derived do not hold. An energy estimate for this problem was shown in [23]. The computed eigenvalues show exponential energy growth for all the split operator approximations for both conservative (Figure 2(a)) and non-conservative (2(b)) problems. Hence, it does not seem to effectively remove aliasing errors by means of a split operator when using Gauss points. This energy growth



(a) Conservative equation spectra (b) Non-conservative equation spectra

Fig. 2 Eigenvalues of the non-constant speed (88) advection equation using Gauss points. The three relevant versions (i.e. conservative ($\alpha = 1$), skew-symmetric ($\alpha = 1/2$), and non-conservative ($\alpha = 0$) DG) of the split operator coefficient are represented. In this test case, $K = 200$ elements have been used, while the polynomial order is $N = 5$.

(with Gauss points) must be then dissipated by other stabilisation techniques, such as adding artificial viscosity [32], or by means of interface dissipation with upwind Riemann solvers [11].

Lastly, we show the effect of over-integration in the DGSEM-GL variant. We solve the conservative equation with the conservative DG (recall that when using over-integration, all split formulations are identical) since it was proven to be stable in its reduced (standard) quadrature version. However, when using over-integration, the scheme is not provably stable, according to (87). Figure 3 depicts the numerical eigenvalues showing the aliasing driven instabilities arising as a result of the over-integration. The eigenvalues show the same growth/decay rates as the skew-symmetric form, since both follow the traditional bound, (3). Therefore, over-integration, when considering the discontinuous Galerkin method with Gauss-Lobatto points must be used with care, since this technique does not necessarily stabilise the solution.

8 Conclusions

In this work, we have studied the numerical instabilities that arise in the numerical solution of the non-constant advection speed equation. In particular, we have analysed non-constant advection speeds with constant sign. By means of energy estimates, we have shown that these instabilities are attributed to aliasing errors incurred in the numerical evaluation of the weak formulation integrals. Selecting the appropriate split form coefficient in a DG method with SBP-SAT properties is fundamental to achieve discrete energy conservation, whereas over-integration techniques are not capable of removing aliasing errors. Precisely, the conservative DG form satisfies these requirements when solving the conservative advection equation, whilst the same occurs when solving the non-conservative equation with the non-conservative DG version. Dis-

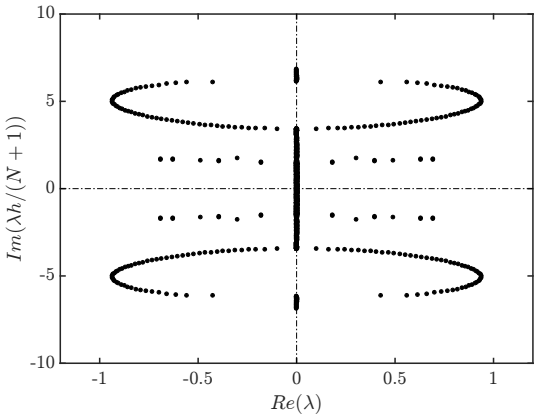


Fig. 3 Eigenvalues of the conservative equation, solved with Gauss-Lobatto points and over-integration of the weak formulation integrals. The polynomial order is $N = 5$, the quadrature order is $M = 10$, and $K = 200$ elements were used. Since the standard quadrature scheme yields a stable scheme, we can conclude that over-integration may lead to aliasing-driven instabilities. Its dissipation equals to that of the skew-symmetric formulation.

crete energy conservation does not occur when using DG versions that do not satisfy the SBP property, e.g. using the Gauss points.

Acknowledgements

This work was supported by a grant from the Simons Foundation (#426393, David Kopriva).

References

1. A.R. Winters and G.J. Gassner: Affordable, entropy conserving and entropy stable flux functions for the ideal MHD equations. Journal of Computational Physics **304**, 72 – 108 (2016)
2. Bassi, F., Crivellini, A., Pietro, D.D., Rebay, S.: An artificial compressibility flux for the discontinuous Galerkin solution of the incompressible Navier–Stokes equations. Journal of Computational Physics **218**(2), 794 – 815 (2006)
3. B. Fornberg: On a Fourier method for the integration of hyperbolic equations. SIAM Journal on Numerical Analysis **12**(4), 509–528 (1975)
4. C. Canuto, M.Y. Hussaini, A. Quarteroni, T.A. Zang: Spectral Methods: Fundamentals in Single Domains (Scientific Computation), 1st ed. 2006. corr. 4th printing 2010 edn. Scientific computation. Springer (2011)
5. D.A. Kopriva: Implementing spectral methods for partial differential equations. Springer Netherlands (2009)

6. D.A. Kopriva and G.J. Gassner: An energy stable discontinuous Galerkin spectral element discretization for variable coefficient advection problems. SIAM Journal on Scientific Computing **36**(4), A2076–A2099 (2014)
7. D.A. Kopriva, A.R. Winters, M. Bohm, and G.J. Gassner: A provably stable discontinuous Galerkin spectral element approximation for moving hexahedral meshes. Computers and Fluids **000**, 1–13 (2016)
8. D. Gottlieb and J.S. Hesthaven: Spectral methods for hyperbolic problems. Journal of Computational and Applied Mathematics **128**(1), 83–131 (2001)
9. E. Ferrer and R.H.J. Willden: A high order discontinuous Galerkin finite element solver for the incompressible Navier–Stokes equations. Computers & Fluids **46**(1), 224–230 (2011)
10. E. Ferrer and R. H.J. Willden: A high order discontinuous Galerkin - Fourier incompressible 3D Navier-Stokes solver with rotating sliding meshes. Journal of Computational Physics **231**(21), 7037–7056 (2012)
11. E. Toro: Riemann solvers and numerical methods for fluid dynamics. Springer (2009)
12. F. Bassi and S. Rebay: A high-order accurate discontinuous finite element method for the numerical solution of the compressible Navier-Stokes equations. Journal of Computational Physics **131**(2), 267 – 279 (1997)
13. Ferrer, E.: An interior penalty stabilised incompressible discontinuous Galerkin-Fourier solver for implicit large eddy simulations. Journal of Computational Physics **348**, 754–775 (2017)
14. G.J. Gassner: A kinetic energy preserving nodal discontinuous Galerkin spectral element method. International Journal for Numerical Methods in Fluids **00**, 1–27 (2013)
15. G.J. Gassner: A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods. SIAM Journal on Scientific Computing **35**(3), 1233–1256 (2013)
16. G.J. Gassner and A.D. Beck pages=221-237, l.: On the accuracy of high-order discretizations for underresolved turbulence simulations. Theoretical and Computational Fluid Dynamics **27**(3-4) (2013). DOI 10.1007/s00162-011-0253-7. URL http://dx.doi.org/10.1007/s00162-011-0253-7
17. G.J. Gassner, A.R. Winters and D.A. Kopriva: Split form nodal discontinuous Galerkin schemes with Summation-By-Parts property for the compressible Euler equations. Journal of Computational Physics **327**, 39–66 (2016)
18. G.J. Gassner, A.R. Winters, F.J. Hindenlang, and D.A. Kopriva: The BR1 scheme is stable for the compressible Navier-Stokes equations. arXiv preprint arXiv:1704.03646 (2017)
19. G.S. Karamanos and G.E. Karniadakis: A spectral vanishing viscosity method for large-eddy simulations. Journal of Computational Physics **163**(1), 22 – 50 (2000)
20. J.L. Guermond and B. Popov: Viscous regularization of the Euler equations and entropy principles. SIAM Journal on Applied Mathematics **74**(2), 284–305 (2014)

21. J. Lorenz and H.O. Kreiss: Initial-Boundary value problems and the Navier-Stokes equations, *Pure and Applied Mathematics*, vol. 136. Academic Press, San Diego, CA, USA (1989)
22. J. Manzanero, G. Rubio, E. Ferrer, and E. Valero: Dispersion-dissipation analysis for advection problems with non-constant coefficients: Application to discontinuous Galerkin formulations. Under review at SIAM Journal of Scientific Computing (2017)
23. J.S. Hesthaven and T. Warburton: Nodal discontinuous Galerkin methods: algorithms, analysis, and applications. Springer Science & Business Media (2008)
24. K. Shahbazi, P.F. Fischer, and C.R. Ethier: A high-order discontinuous Galerkin method for the unsteady incompressible Navier–Stokes equations. Journal of Computational Physics **222**(1), 391 – 407 (2007)
25. M.H. Carpenter, T.C. Fisher, E.J. Nielsen, and S.H. Frankel: Entropy Stable Spectral Collocation Schemes for the Navier-Stokes Equations: Discontinuous Interfaces. SIAM Journal on Scientific Computing **36**(5), B835–B867 (2014)
26. M. Kompenhans, G. Rubio, E. Ferrer, and E. Valero: Adaptation strategies for high order discontinuous galerkin methods based on tau-estimation. Journal of Computational Physics **306**, 216 – 236 (2016)
27. M. Kompenhans, G. Rubio, E. Ferrer, and E. Valero: Comparisons of p–adaptation strategies based on truncation– and discretisation–errors for high order discontinuous Galerkin methods. Computers & Fluids **139**, 36 – 46 (2016). 13th {USNCCM} International Symposium of High-Order Methods for Computational Fluid Dynamics - A special issue dedicated to the 60th birthday of Professor David Kopriva
28. M. Svärd and J. Nordström: Review of summation-by-parts schemes for initial-boundary-value problems. Journal of Computational Physics **268**(1), 1738 (2014)
29. R.C. Moura, S.J. Sherwin, and J. Peiro: Linear dispersion-diffusion analysis and its application to under-resolved turbulence simulations using discontinuous Galerkin spectral/hp methods. Journal of Computational Physics **298**, 695–710 (2015)
30. R.M. Kirby and G. Em Karniadakis: De-aliasing on non-uniform grids: Algorithms and applications. Journal of Computational Physics **191**(1), 249–264 (2003)
31. R.M. Kirby and S.J. Sherwin: Aliasing errors due to quadratic nonlinearities on triangular spectral /hp element discretisations. Journal of Engineering Mathematics **56**(3), 273–288 (2006)
32. R.M. Kirby and S.J. Sherwin: Stabilisation of spectral/hp element methods through spectral vanishing viscosity: application to fluid mechanics modelling. Computer Methods in Applied Mechanics and Engineering **195**(23-24), 3128–3144 (2006)
33. S.A. Teukolsky: Short note on the mass matrix for Gauss-Lobatto grid points. Journal of Computational Physics **283** (2014)

34. S.C. Spiegel, H.T. Huynh, and J.R. DeBonis: De-aliasing through over-integration applied to the flux reconstruction and discontinuous Galerkin methods. 22nd AIAA Computational Fluid Dynamics Conference, AIAA Aviation, (AIAA 2015-2744)
35. T.A. Zang: On the rotation and skew-symmetric forms for incompressible flow simulations. Applied Numerical Mathematics **7**(1), 27–40 (1991)
36. T.C. Fisher and M.H. Carpenter: High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains. Journal of Computational Physics **252**, 518–557 (2013)
37. Z.J. Wang, K. Fidkowski, R. Abgrall, F. Bassi, D. Caraeni, A. Cary, H. Deconinck, R. Hartmann, K. Hillewaert, H.T. Huynh, N. Kroll, G. May, P.O. Persson, B. van Leer, and M. Visbal: High-order CFD methods: current status and perspective. International Journal for Numerical Methods in Fluids **72**(8), 811–845 (2013)