

# Analysis for Implicit and Implicit-Explicit ADER and DeC Methods for Ordinary Differential Equations, Advection-Diffusion and Advection-Dispersion Equations: Extra material

Philipp Öffner\*, Louis Petri† Davide Torlo‡

## 1 Scope of the document

In this work, we collect all the stability analysis and numerical results that did not fit the work “Analysis for Implicit and Implicit-Explicit ADER and DeC Methods for Ordinary Differential Equations, Advection-Diffusion and Advection-Dispersion Equations” [16]. We will not introduce the methods that are studied in this document, but we refer to [16] for all the definitions. On the other hand, we keep the discussion of the numerical analysis in the text to give some context to the stability results.

The structure of the document is as follows. In Section 2, we study the stability regions of ADER, DeC and sDeC methods as ODE solver in their explicit, implicit and IMEX versions. Next, in Sections 3 and 4 we extend the stability analysis to the PDE scenario by applying our IMEX methods to advection-diffusion and advection-dispersion equations.

## 2 Numerical Stability Analysis

In this section, we study the stability of the presented method for the linear Dahlquist equation  $u' = -\lambda u$  or  $u' = -\lambda_I u - \lambda_E u$  for IMEX methods. All methods can be rewritten as  $u_{n+1} = R(z)u_n$  or  $u_{n+1} = R(z_I, z_E)u_n$  being  $R$  stability functions and  $z, z_I, z_E \in \mathbb{C}$ . We will study the stability regions  $\{|R| \leq 1\} \subset \mathbb{C}$  for the explicit and implicit ADER/DeC, while for the IMEX schemes we consider different approaches. We will use equispaced (eq) or Gauss-Lobatto (GLB) nodes as quadrature nodes. We will numerically compute the stability regions obtained from the stability functions of ADER/DeC that are defined through their Butcher tableaux, see [4]. In detail, we compute on  $200 \times 200$  grid points with an offset of +0.01 from the origin for both axes to avoid singularities. The plot bounds are dependent on the type of scheme and their stability regions. We decreased the offset in Figures 4 to a fraction of  $10^{-2}$  of the largest real value displayed when zooming on small areas.

To distinguish the different orders, we apply different colors and line styles to the outer and inner bounds according to the legend that are plot next to each stability region plot.

### 2.1 Stability analysis of explicit schemes

In figure 1 (left), we show the results obtained in [5], extended up to order 13. It was pointed out, that the regions of the explicit ADER and DeC coincide and that they are independent of the chosen interpolation

---

\*Institute of Mathematics, Johannes Gutenberg University, Mainz, and Institute of Numerical Analysis, TU Clausthal, Clausthal-Zellerfeld, Germany, mail@philippoeffner.de

†Institute of Mathematics, Johannes Gutenberg University, Mainz, Germany, lpetri01@uni-mainz.de

‡Dipartimento di Matematica Guido Castelnuovo, Università di Roma La Sapienza, Roma, Italy, davide.torlo@uniroma1.it

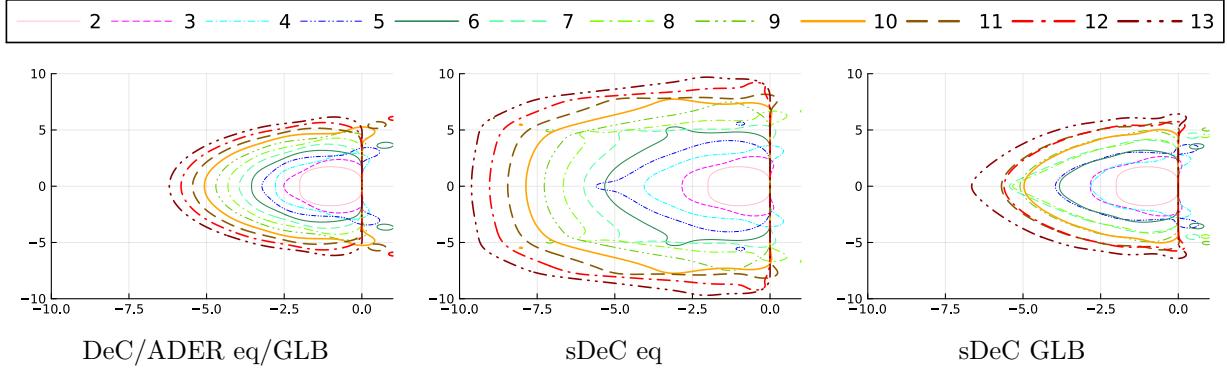


Figure 1: Stability regions for the explicit ADER and DeC methods with GLB or equispaced nodes for orders 2 to 13 (left), sDeC equispaced (center) and sDeC GLB (right).

Table 1: Table with orders of the ImsDeC methods that have a bounded stability region on the negative half-plane

ImsDeC nodes specification	orders of the method
Gauss-Lobatto	9, 10, 11, 12, 13, 14, 15
equispaced	12, 13, 16, 17, 18, 19, 20

nodes. Here, we just display them once. We can highlight the growth of stability by increasing the order of the respective method.

Furthermore, we want to take a look at the explicit sDeC, whose stability regions can be observed in figure 1 (center and right). This method differs not only from the ADER and DeC, but also from sDeC with different nodes. The qualitative shape is still similar to the others, but it is just remarkable that the sDeC methods with equispaced interpolation points have respectively larger stability regions.

## 2.2 Implicit schemes

In the following, we plot the contour lines of the bounds of the stability regions of various implicit methods. We start from ImDeC (left) and ImADER (right) schemes with equispaced nodes (top) and GLB nodes (bottom) in Figure 2 up to order 13. Clearly, all these stability regions are unbounded, but they are not all A-stable, as we will see soon. Moreover, we can observe a great variability changing the scheme or the nodes, in opposition to the explicit case [5]. In most of the cases, ImADER have larger stability regions than ImDeC.

The stability regions for the implicit sDeC (ImsDeC) are also shown in Figure 2 and, surprisingly, do not behave like the other methods. Up to a certain order (i.e. sDeC8 with Gauss-Lobatto and sDeC11 with equispaced nodes), the stability regions are unbounded and *seem* A-stable, but for higher orders, we lose this property, obtaining large, but finite, stability regions. This behavior is not uniform and, at certain orders, the stability region will be unbounded again, as for example shown in Figure 3 for very high order ImsDeC. A detailed list of the bounded methods until order 20 is given in Table 1. For the sDeC, we can conclude that the choice of an implicit version does not guarantee an unbounded stability region. Nevertheless, even these implicit sDeC methods have larger stability regions than their explicit counterparts and therefore may be applicable to mildly stiff problems. We notice again that this odd loss of stability in the left half plane could not be found in the ImDeC and ImADER methods. We checked it numerically up to order 50.

Taking a closer look at the implicit methods, we additionally detect some minor instability regions on the negative half-plane, see Figure 4. It turns out that these instabilities appear for all ImDeC and ImsDeC methods of orders larger than 2 and both types of nodes. We display the sDeC only with GLB nodes, as the equispaced nodes version is very similar to that. For a very small scale, the same can be seen for the ImADER methods with equispaced nodes of orders at least larger than 4, displayed on the top right

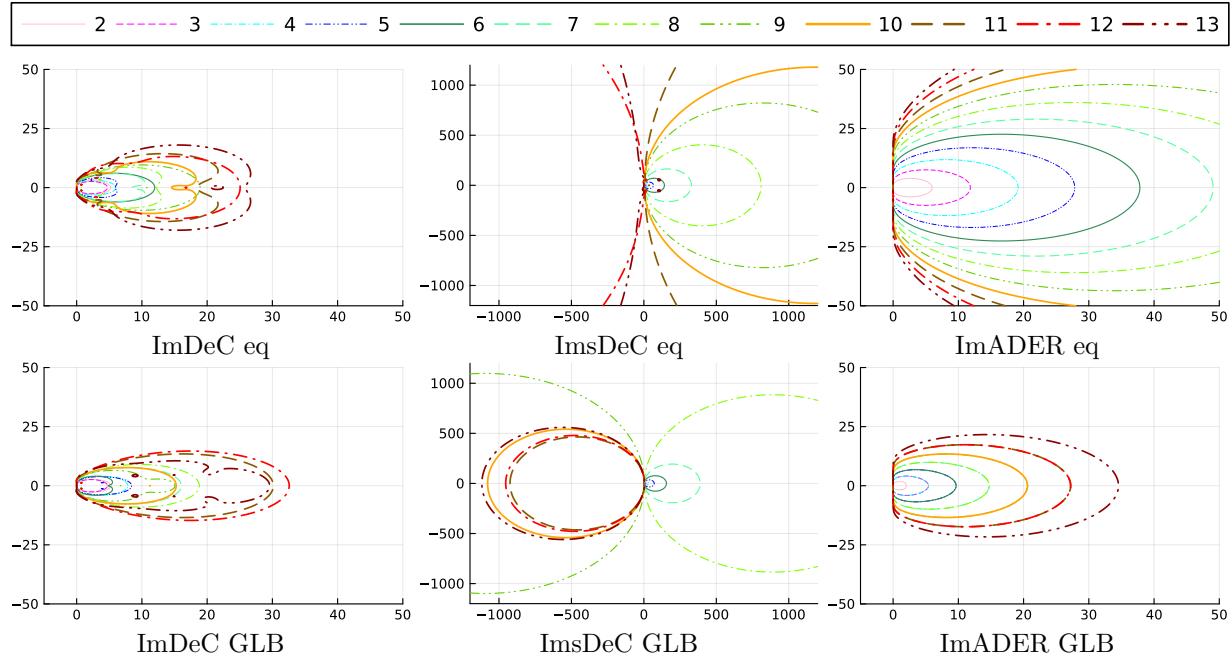


Figure 2: Implicit DeC (left), sDeC (center) and ADER (right) with equispaced (top) and Gauss-Lobatto (bottom) nodes for orders 2 to 13

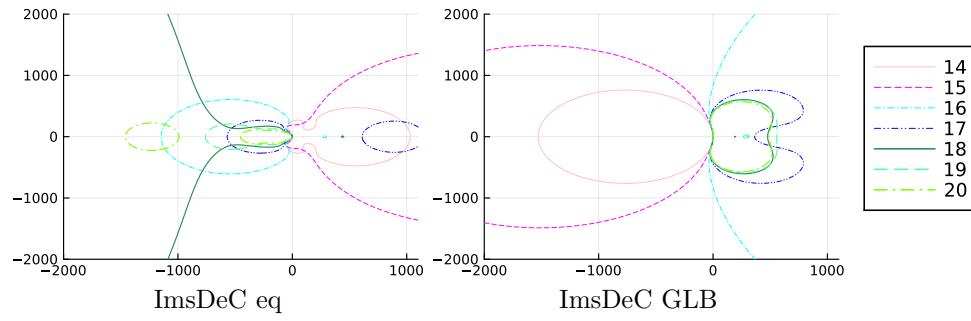


Figure 3: Implicit sDeC for orders 14 to 20

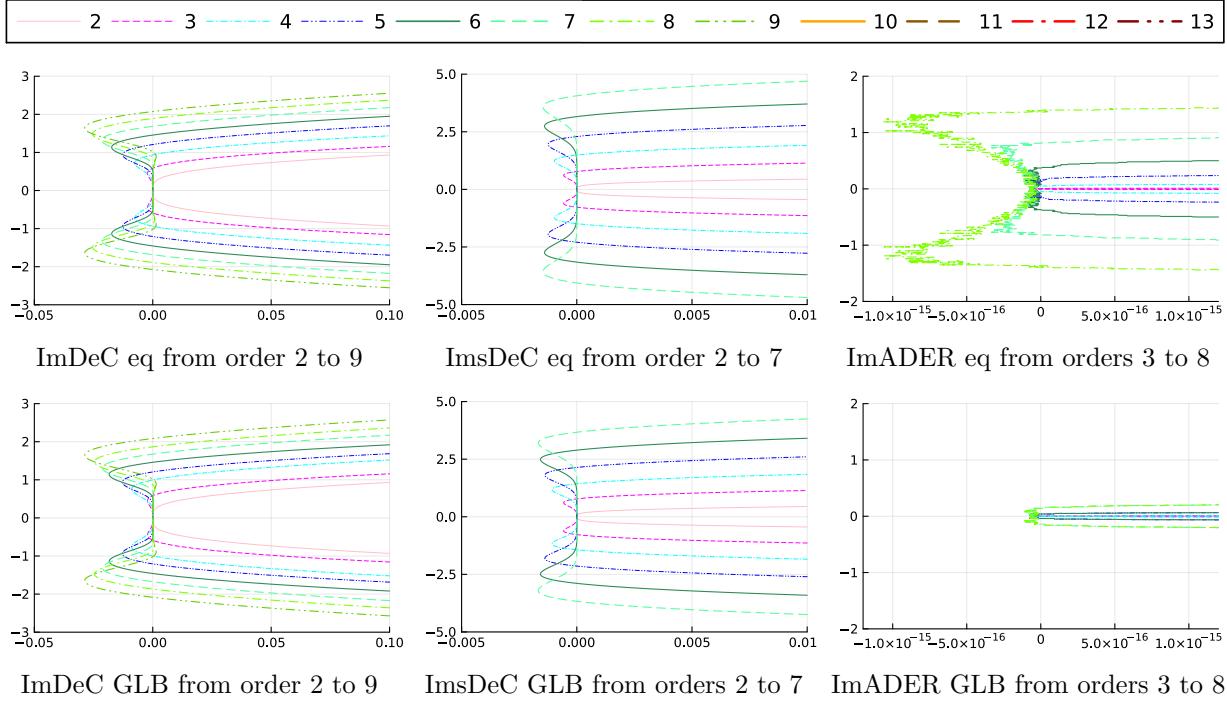


Figure 4: Zoomed stability region of implicit schemes (watch out at the different scales).

of Figure 4. Notice that the sizes of the unstable regions are close to machine precision, which results in non-smooth boundaries and it is unclear if the ImADER with equispaced nodes are not A-stable or the visualization of the unstable area is given by machine precision errors.

As proved in [16], also numerically we observe that the ImADER methods with Gauss-Lobatto nodes are A-stable for all orders. Here, the noise in the stability region (bottom right of Figure 4) is really of the size of machine precision.

Summarizing, we can categorize our methods in 3 different classes:

- The A-stable schemes: all ImADER GLB and all second order implicit methods;
- The *almost A-stable* schemes, when the stability region is unbounded and it *almost* includes the whole left half-plane: high order ImDeC, some ImsDeC and ImADER equispaced;
- The bounded stability schemes: some ImsDeC.

Remark that for *almost A-stable* methods these minor instabilities do not influence the behavior of the scheme on many stiff problems. Nevertheless, when the eigenvalues  $\lambda$  of the system are (almost) purely imaginary (typical for high order advection operators), they might encounter instabilities for some discretizations.

### 2.3 IMEX schemes

To study the stability of IMEX schemes, we will use the RK stability function

$$R(z_I, z_E) = 1 + \left( z_I b^T + z_E \hat{b}^T \right) \underline{u} = 1 + \left( z_I b^T + z_E \hat{b}^T \right) \left( \underline{Id} - z_I \underline{A} - z_E \hat{\underline{A}} \right)^{-1} \underline{1} \quad (1)$$

that uses the matrices defined by the Butcher tableau of an IMEX RK. Remark that the standard approach of A-stability cannot be used anymore. Indeed, the region of absolute stability

$$S = \{(z_I, z_E) \in \mathbb{C}^2 : |R(z_I, z_E)| \leq 1\}$$

lays in a larger space, with respect to classical RK schemes, therefore, its study, computation and visualization are challenging. Hence, we need to rely on some simplifications. In [14], Minion simplifies the Dahlquist equation by imposing

$$\lambda_I \in \mathbb{R}, \quad \lambda_E = i\lambda'_E, \quad \lambda'_E \in \mathbb{R}.$$

This procedure neglects respectively the imaginary or real part of the coefficients in the Dahlquist equation to display a two-dimensional region. This idea is lead by classical PDE discrete operators, where typically the diffusion is symmetric negative definite, while the advection is mainly with imaginary eigenvalues. A second approach where for each  $\lambda_E$  the A-stability is required for the implicit part of the scheme was originally studied in [27, 1] and formalized in [9]. Another approach studies, instead, the stability for each  $\lambda_I$  requiring at least the stability region of the explicit Euler method to the explicit part [6]. We collect these definitions of stability region in the following.

**Definition 2.1** (Stability regions). *Consider the modified test equation with stability function (1). Then, we define multiple approaches for IMEX stability regions by*

- $S := \{(z_I, z_E) \in \mathbb{C}^2 : |R(z_I, z_E)| \leq 1\}$  (Region of absolute stability),
- $\mathcal{D}_M := \{(z_I, z_E) \in \mathbb{R}^2 : |R(z_I, iz_E)| \leq 1\}$  (Minion's stability region) [14],
- $\mathcal{D}_0 := \{z_E \in \mathbb{C} : |R(z_I, z_E)| \leq 1 \text{ for any } z_I \in \mathbb{C}^-\}$  [9],
- $\mathcal{D}_1 := \{z_I \in \mathbb{C} : |R(z_I, z_E)| \leq 1 \text{ for any } z_E \in \mathcal{S}_0\}$  [6],

where  $\mathcal{S}_0 = \{z_E \in \mathbb{C} : |1 + z_E| \leq 1\}$  is the stability region of the explicit Euler method.

$\mathcal{D}_0$  is a very strict condition of IMEX stability, in particular for the considered high order schemes. Theoretically, the terms A-stability and A( $\alpha$ )-stability may be applied for all 3 of these subsets of  $\mathbb{C}$  analogously to the classical cases, so we will make use of this terminology too.

### $\mathcal{D}_M$ stability region

We recall that the plots of the stability regions have very different meaning according to the chosen approach. Starting from Minion's approach [14], we evaluate the IMEX stability function (1) numerically to calculate the respective stability regions.

For the IMEX DeC, we can observe in Figure 5 (left) that the choice of nodes change the regions on some details but the qualitative behavior is the same. We can also conclude on an  $A(\alpha)$ -stability for approximately  $\alpha = 35^\circ$ .

Going on to the IMEX ADER, we can see in Figure 5 (right) a similar behavior, even if the stability regions differ in small details, we observe  $A(\alpha)$ -stability for at least  $\alpha = 35^\circ$ .

Finally, for the IMEX sDeC method in Figure 5 (center) we see a slightly different behavior, still resulting in an  $A(\alpha)$ -stability, but for significantly smaller angles, approximately  $\alpha = 18^\circ$ . Nevertheless, the result on the bottom center in Figure 5 with GLB nodes coincides with the one in [14], as expected. It is also noticeable that the IMEX sDeC stability region of order 2 is  $A(\alpha)$ -stable with larger  $\alpha$  as it coincides with the IMEX DeC2.

With the  $\mathcal{D}_M$  approach, we do not observe strong variations between equispaced and Gauss-Lobatto points.

### $\mathcal{D}_0$ stability region

Now, we want to evaluate  $\mathcal{D}_0$  stability for our IMEX methods. We want to emphasize that the requirements here are stricter than in Minion's approach. Indeed, for  $\mathcal{D}_0$  we require the method to be at least fully A-stable for the implicit part and we look at the stability of the explicit part. The IMEX DeC and IMEX sDeC have  $\mathcal{D}_0 = \emptyset$  and this is probably related to the fact that their implicit counterpart is not A-stable. For the IMEX ADER, only few orders have non-empty  $\mathcal{D}_0$  stability region. In Figure 6, we show the few stability regions, which eventually vanish when increasing the order of accuracy.

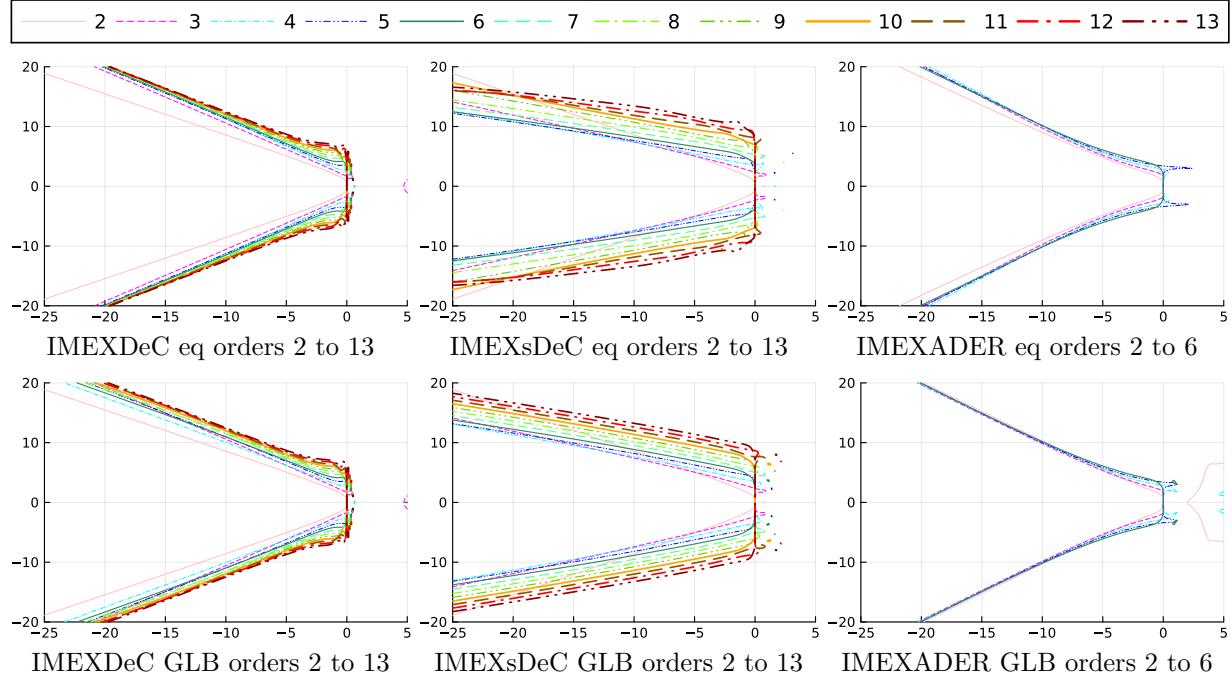


Figure 5: Minion's stability region for IMEX DeC (left), sDeC (center) and ADER (right) with equispaced (top) and GLB (bottom) nodes

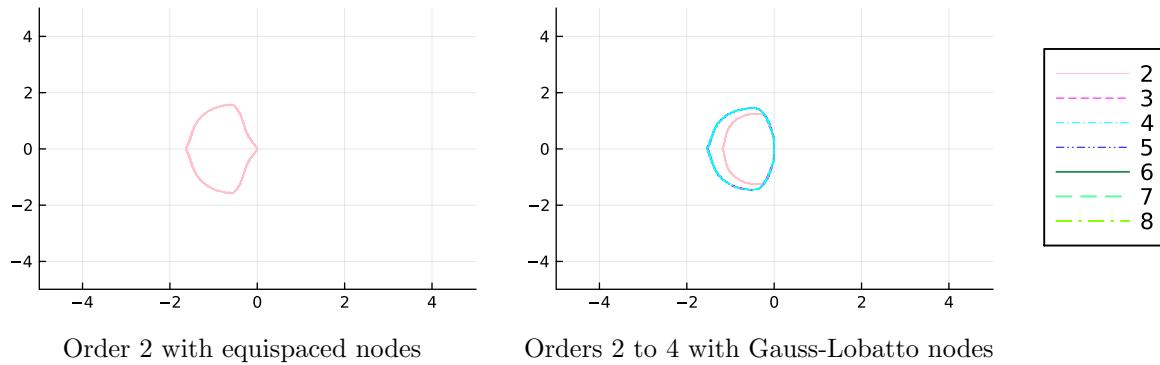


Figure 6:  $\mathcal{D}_0$  Stability regions for IMEX ADER. The smaller stability region displays order 2, while the larger one displays order 3 and 4 (right)

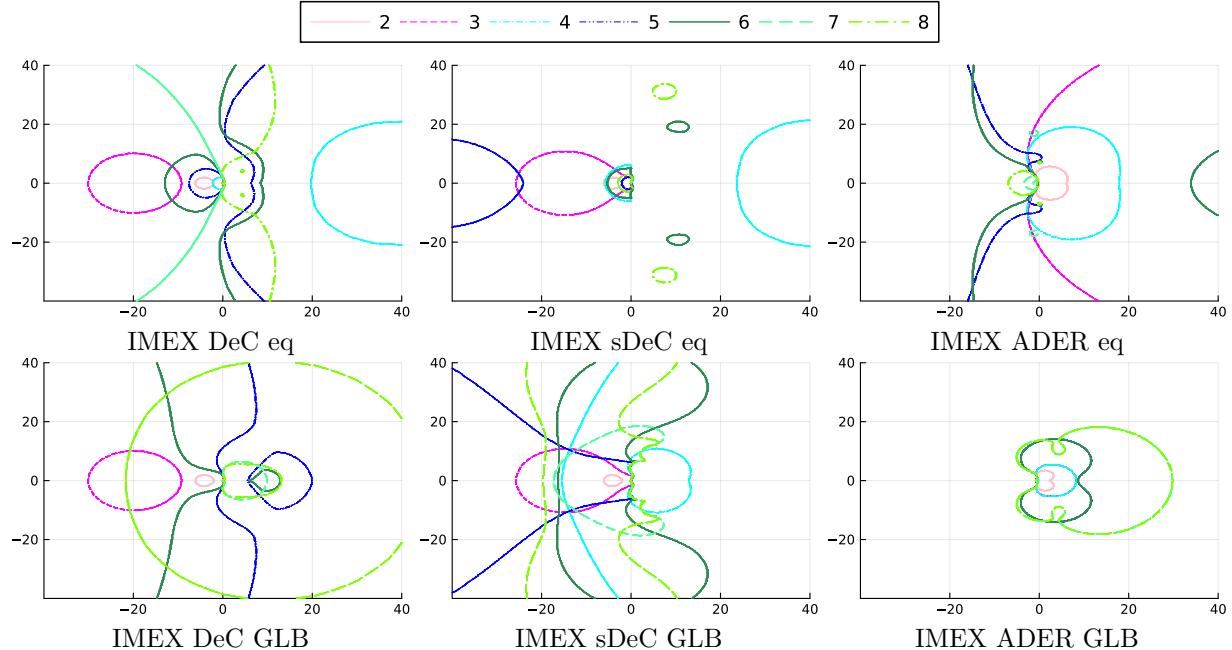


Figure 7:  $\mathcal{D}_1$  Stability Region for IMEX DeC (left), sDeC (center) and ADER (right) with equispaced (top) and GLB (bottom) nodes: orders 2 to 8

### $\mathcal{D}_1$ stability region

We plot the  $\mathcal{D}_1$  stability regions for DeC and sDeC methods in Figures 7, where we require the explicit part to cover the stability region of the explicit Euler method and we look at the stability of the implicit part. Contrary to the  $\mathcal{D}_0$  cases, we observe non-empty, limited regions of stability for every order for the IMEX DeC methods. Moreover, there is no regularity in their shape and their size grow significantly as the order of accuracy increases. Notice that the plots do not show the full stability regions of higher orders, for example for orders 6, 7, and 8 with equispaced nodes, but they are anyway bounded regions.

For the IMEX sDeC, see Figure 7 (center), we observe some remarkable differences. In the case of equispaced nodes, even orders just show the small bounded stability regions in the negative half-plane nearby the origin, odd orders smaller than 6 show large stability regions, while they are unstable starting from order 7. Also in the GLB case, we do not observe much regularity. We notice that the largest stability region is obtained for order 5, while, for higher orders, the stability region almost fit in the plot.

In Figure 7 (right), we show the results for the IMEX ADER methods. We note that most of the methods fulfill nearly A-stability by almost covering the negative half-plane. We highlight that for the equispaced case we do not show the full outreach of the stability regions. While in most of the cases, for the almost A-stable cases, these contour lines represent the inner bounds of unlimited stability regions, in the cases of order 5 and 8 we just have large, limited stability regions, as we also could observe for example in the  $\mathcal{D}_1$  IMEX DeC case for equispaced nodes. Therefore, it seems like we can not guarantee this almost A-stability for the IMEX ADER, but just for some of the orders of accuracy.

Therefore, we can conclude, that some of the IMEX ADER stability regions cover the areas in the complex plane, that we assumed for a stable method in the context of  $\mathcal{D}_0$  and  $\mathcal{D}_1$ , while the DeC and sDeC methods have their limitations with  $\mathcal{D}_0 = \emptyset$  in every case and bounded stability regions for most of the cases in the scope of  $\mathcal{D}_1$ . Nevertheless, we need to keep in mind that the set conditions are very strict, so, the methods might still be applicable to some stiff equations. These results reflect what we have seen for the respective implicit methods.

Table 2: Central finite difference discretizations of  $\partial_{xx}$  applied onto  $w$  centered in  $j$  [3]

order	finite difference for $\partial_{\Delta x}^2(u(x_j))$
2	$\frac{1}{\Delta x^2}(w_{j-1} - 2w_j + w_{j+1})$
4	$\frac{1}{\Delta x^2}(-\frac{1}{12}w_{j-2} + \frac{4}{3}w_{j-1} - \frac{5}{2}w_j + \frac{4}{3}w_{j+1} - \frac{1}{12}w_{j+2})$
6	$\frac{1}{\Delta x^2}(\frac{1}{90}w_{j-3} - \frac{3}{20}w_{j-2} + \frac{3}{2}w_j - \frac{49}{18}w_{j+1} + \frac{3}{2}w_{j+2} - \frac{3}{20}w_{j+3} + \frac{1}{90}w_{j+4})$
8	$\frac{1}{\Delta x^2}(-\frac{1}{56}w_{j-4} + \frac{1}{420}w_{j-3} - \frac{1}{5}w_{j-2} + \frac{8}{5}w_{j-1} - \frac{205}{72}w_j + \frac{8}{5}w_{j+1} - \frac{1}{5}w_{j+2} + \frac{1}{420}w_{j+3} - \frac{1}{56}w_{j+4})$

### 3 PDE: analysis of advection-diffusion

In this section, we want to extend our stability analysis to the one-dimensional advection-diffusion equation

$$u_t(x, t) + au_x(x, t) = du_{xx}(x, t), \quad a \geq 0, \quad d \geq 0, \quad x \in \Omega \subset \mathbb{R}, \quad (2)$$

where  $a$  is the coefficient of the advection term and  $d$  the coefficient of the diffusion term, using the von Neumann stability analysis. After the space discretization, we discretize the advection part with an explicit time-integration scheme and the diffusion with an implicit one. The linear stability of the DeC method in PDE contexts was studied for explicit methods for advection equations with FEM spatial discretizations and various stabilization techniques in [12, 13], while in the IMEX context for FEM methods applied to kinetic models in [24]. For the ADER method, a von Neumann stability analysis was applied to the original formulation [23, 2], but not on the modern version that we are studying. We close this gap with our investigation in the following.

#### 3.1 Finite Difference discretization

We apply spatial discretizations to the spatial derivative operators, namely  $\partial_x$  and  $\partial_{xx}$ . We consider a uniformed grid  $\Omega_{\Delta x} = \{x_j : x_j = x_0 + j\Delta x, j \in \{0, \dots, J\}\}$  with periodic boundary conditions and we denote the approximation of  $u(x_j) = u(x_j, t)$  by  $w_j$ .

To discretize the advection term in (2), i.e., the first spatial derivative  $\partial_x u(x)$ , we make usage of the stable finite difference stencils introduced in [7]. Assume we discretize  $\partial_x u$  at  $x_j$  by an  $[r, s]$ -discretization

$$\partial_{\Delta x}^{[r,s]}(u(x_j)) = \frac{1}{\Delta x} \sum_{k=-r}^s \alpha_k w_{j+k}, \quad (3)$$

with  $r, s$  such that  $\alpha_{j-r}, \alpha_{j+s} \neq 0$ . The maximum order we can achieve with an  $[r, s]$ -discretization is  $q = r + s$  and this discretization actually reaches order  $q$  and is unique by setting the coefficients in (3) as

$$\alpha_0 = \begin{cases} \sum_{k=r+1}^s \frac{1}{k}, & s \geq r+1, \\ 0, & s = r, \\ \sum_{k=s+1}^r \frac{1}{k}, & r \geq s+1, \end{cases} \quad \alpha_k = \frac{(-1)^{k+1}}{k} \cdot \frac{r!s!}{(r+k)!(s-k)!}, \quad -r \leq k \leq s, \quad k \neq 0. \quad (4)$$

It is also proven in [7] that these so-called *optimal-order* schemes of order  $q$  are stable if and only if  $s \leq r \leq s+2$  for  $a > 0$ . We involve these stable *optimal-order* schemes into our analysis. We introduce upwinding in the choice of the stencils, in particular, we will consider  $[r, r+1]$  stencils for odd optimal-order scheme and  $[r, r+2]$  stencils for an even optimal-order scheme for the advection part.

For the diffusion term, we will just use a central finite difference discretization of the second spatial derivative  $\partial_{xx} u(x)$  given in Table 2.

#### 3.2 von Neumann analysis

To analyze the stability of the described methods, we make use of the von Neumann stability analysis for linear partial differential equations [8]. Briefly summarized, we investigate the behavior inside the numerical

scheme of the Fourier modes

$$w_j^n = v^n e^{ikx_j}, \quad (5)$$

where  $w_j^n$  is the discretization of  $u(x_j, t_n)$  and  $k$  is the wavenumber and we focus on the representation coefficient  $v^n$ . Indeed,  $e^{ikx}$  are eigenfunctions of the differential operator  $\partial_x$  and therefore for any linear differential operator. If we use (5) in our discretized system, we obtain a system of the form

$$v^{n+1} = G(k, \Delta x, \Delta t, a, d)v^n. \quad (6)$$

with the amplification factor  $G \in \mathbb{C}$  independent on the mesh point  $x_j$ . Stability means in our context that  $|v^{n+1}| = |G(k, \Delta x, \Delta t, a, d)v^n| \leq |v^n|$  holds. In practice, we check that  $|G(k, \Delta x, \Delta t, a, d)| \leq 1$ . This implies stability for the related method and due to the Lax-Richtmeyer theorem [8] convergence can be ensured. Note that for consistency, we included the parameters  $a$  and  $d$  into the dependency of  $G$  to cover all advection-diffusion equations.

Typically, in order to estimate the stability of the advection-diffusion equation, an analytical study of  $G$  in all the parameters should be performed. This is not feasible when considering high order schemes as ADER and DeC. Hence, we will evaluate the amplification factor numerically, similarly to what we did with the stability functions in the ODE case.

Before running all the simulations, we need to understand what are the free variables of the function  $G$ . First of all, the wavenumbers should be bounded  $k \in \{-n_0 - 1, \dots, n_0 + 1\} \subset \mathbb{Z}$  and the maximum wavenumber  $n_0 + 1$  is strongly related with the discretization scale  $\Delta x$ . Indeed, by Nyquist-Shannon sampling theorem, only functions with frequency less than  $\frac{|x_J - x_0|}{2\Delta x}$  can be represented on our discretization. Hence, we will choose  $n_0 = 10^3$  to take in consideration fine grids.

Then, we have further variables  $a, d, \Delta x, \Delta t$  that are actually coupled together: in the advection term ( $a\Delta t/\Delta x$ ) and in the diffusion term ( $d\Delta t/\Delta x^2$ ). We keep this in mind when studying the behavior of  $G$ , as we can recast few methods to the same coefficients.

### 3.2.1 Displaying stability

It is stated and numerically shown in [20, 25, 26] that several schemes as the local discontinuous Galerkin scheme [25, 26] and other finite difference schemes combined with an IMEX RK method are stable if the time step is upper bounded by some  $\tau_0$ . This  $\tau_0$  is proportional to  $\frac{d}{a^2}$ , i.e., if  $\Delta t \leq \tau_0 = E_0 \cdot \frac{d}{a^2}$  for some  $E_0 > 0$ . Considering the before mentioned parameters  $\Delta x, \Delta t, a, d$ , we introduce two new coefficients

$$C = \frac{a\Delta t}{\Delta x}, \quad D = \frac{d\Delta t}{(\Delta x)^2}. \quad (7)$$

Moreover, using the coefficients  $C$  and  $D$  reveals an equivalent condition to [20], if we assume that the quotient

$$E := \frac{C^2}{D} = \frac{\Delta t^2 a^2}{\Delta x^2} \frac{\Delta x^2}{d\Delta t} = \frac{a^2}{d} \Delta t$$

is bounded by some constant  $E_0$ , indeed,

$$E = \frac{a^2}{d} \Delta t \leq E_0 \iff \Delta t \leq E_0 \cdot \frac{d}{a^2} = \tau_0.$$

Therefore, for a given method solving the advection-diffusion equation, we can rewrite the amplification factor in (6) as

$$g(k, C, E) = G(k, \Delta x, \Delta t, a, d). \quad (8)$$

**Definition 3.1** (Scheme notation). *To shorten the notation, we denote the considered method for the advection-diffusion equation by [TMM, NODES,  $N$ ,  $A_n$ ,  $D_m$ ], where*

- TMM stands for the respective IMEX time-marching method, among DEC, ADER, sDEC,

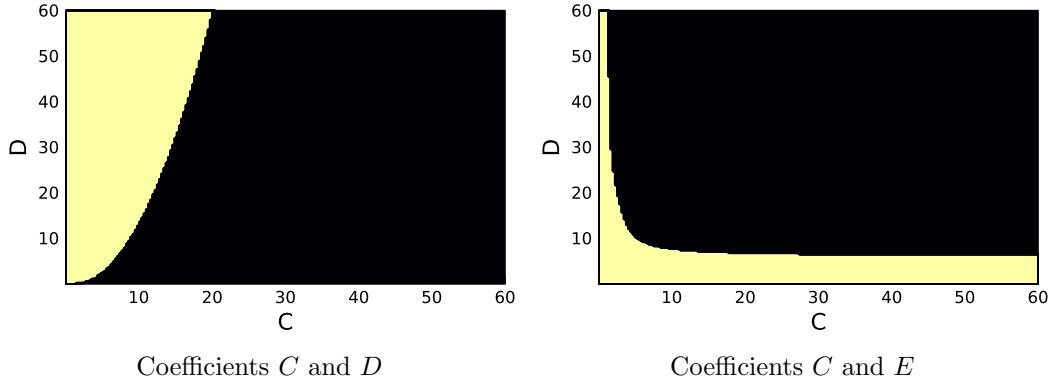


Figure 8: Stability areas (yellow) for the [DEC, EQ, 3,  $A_1$ ,  $D_2$ ].

- NODES stands for the used quadrature nodes for the TMM, among EQ or GLB,
- $N$  stands for the order of the considered time-marching method TMM,
- $A_n$  denotes the optimal first derivative stencil of order  $n$  defined in (4) used for the advection term,
- $D_m$  denotes the central second derivative stencil of order  $m$  in Table 2 used for the diffusion term.

**Example 3.2** (Stability of IMEX DeC3 with  $A_1$  and  $D_2$  operators). To give an example of how to study the stability region, we show in Figure 8 the stability areas  $\{|g(k, C, E)| \leq 1, \forall k \in [-n_0 - 1, n_0 + 1]\}$  for the [DEC, EQ, 3,  $A_1$ ,  $D_2$ ]. On the left, we plot the stability area as a function of  $C$  and  $D$ , while on the right as a function of  $C$  and  $E$ . The black area is associated to the unstable area, while the yellow displays the stable region. We recognize two sufficient conditions to obtain stability:

- the well known CFL-condition, i.e., if  $C$  is lower than some constant  $C_0$  only dependent on the method, then the method is stable.
- the new numerically obtained condition, designated as the  $E_0$ -condition: If  $E$  is lower than some constant  $E_0$  dependent on the method, then the method is stable.

The two parameters  $C$  and  $E$  include all the remaining ones and are enough to characterize the whole scheme.

As we can see in Figure 8, the unstable area of this specific method seems to be bounded by the linear constraints  $C \geq C_0$  and  $E \geq E_0$ . We will observe numerically that these unstable regions are indeed similarly bounded in most of our methods.

**Definition 3.3** (Stability parameters  $C_0$  and  $E_0$ ). Given the amplification factor  $g(k, C, E)$  of a discretization of the advection-diffusion equation, we define the two stability parameters  $C_0$  and  $E_0$  by

- $C_0 := \max_{C \in \mathcal{S}} C$  with  $\mathcal{S} = \{C : |g(k, C, E)| \leq 1, \forall E > 0, \forall k \in [-n_0 - 1, n_0 + 1]\}$ ,
- $E_0 := \max_{E \in \mathcal{R}} E$  with  $\mathcal{R} = \{E : |g(k, C, E)| \leq 1, \forall C > 0, \forall k \in [-n_0 - 1, n_0 + 1]\}$ .

Therefore, the strategy we want to follow is to look at the areas of stability by evaluating the amplification factor  $\max_k |g(k, C, E)|$  like in Figure 8 and numerically calculating the parameters  $C_0$  and  $E_0$  for our methods, when possible.

Note, that the condition  $E < E_0$  does not depend on  $\Delta x$  and avoids CFL restrictions. We should always keep in mind that our numerical evaluations can just cover finite ranges of  $C$  and  $E$ . Hence, we checked that the displayed limits for  $E$  and  $C$  are actually bounds also for larger domains of  $C$  and  $E$ . Moreover, we also observed that the considered results do not vary much for large values of  $n_0$ , hence, we set  $n_0 = 10^3$ . Due to time-efficiency, we will use this value for every evaluation in the von Neumann stability analysis context for the rest of this work. Further, all plots in this section will be evaluated and displayed at  $400 \times 400$  grid points.

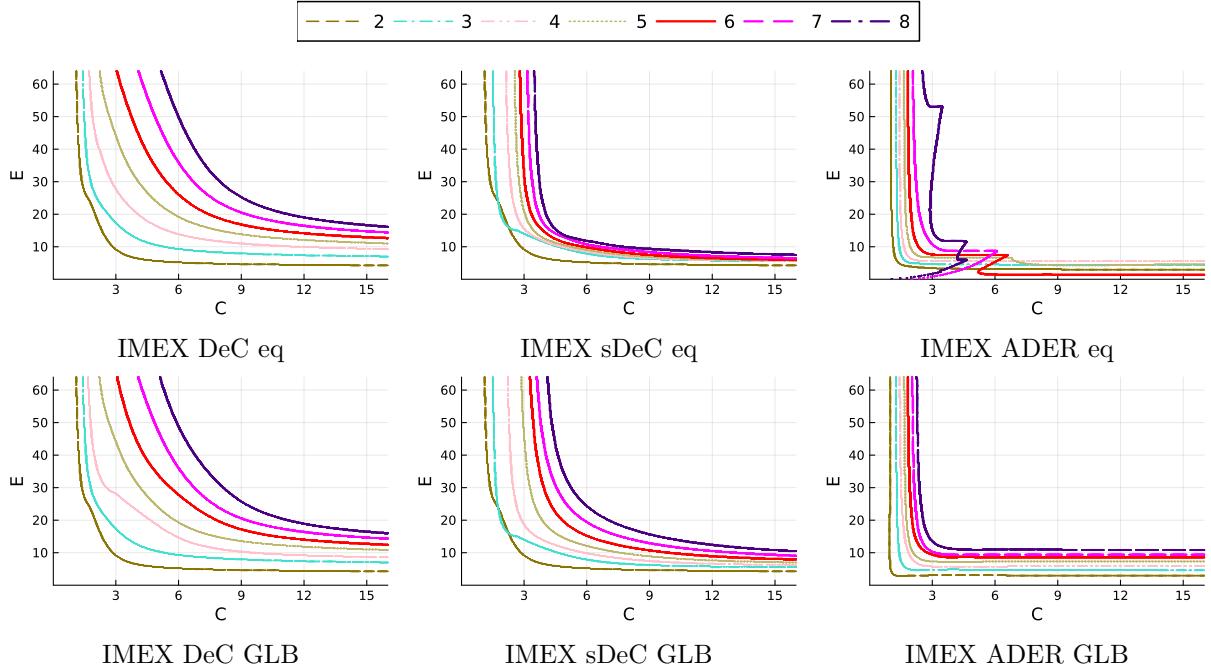


Figure 9: Stability areas for  $[TMM, NODES, k, A_1, D_2]$  for TMM as IMEX DeC (left), IMEX sDeC (center) and IMEX ADER (right) with eq (top) and GLB (bottom) nodes: TMM order  $k$  from 2 to 8

### 3.2.2 Numerical analysis

In the following plots, we will study various configurations of methods and the variation of the stability region changing the order of the methods. In particular, we will focus on changing the order of the time discretization only and varying the order of all operators. In the repository [17], we include other variations of the order of only the advection or the diffusion operator that we discard here for brevity. Moreover, we will focus on Gauss–Lobatto nodes for the time discretization only mentioning the peculiarity of the equispaced ones that can be found, again, in the repository [17].

We start by varying only the time scheme order. In Figure 9, the stability areas of several methods are shown, i.e.,  $[TMM, NODES, k, A_1, D_2]$  varying the time scheme and the nodes. As in Figure 8, the plotted lines separates the stable region in the lower left part from the unstable region in the upper right side of the  $C$ - $E$  plane. We see a similar behavior for mostly all methods. Increasing the order of the time marching method results in a larger stability region and in larger values for  $E_0$  and  $C_0$ .

In the equispaced case we do not observe major differences for the DeC and sDeC methods, while, for the ADER cases, the usage of equispaced nodes results in an irregular reduction to  $E_0 = 0$  for some orders (7 and 8), meaning that we can not ensure stability as we did in the other cases for high order methods, see Figure 13 (left). This is probably due to the numerical cancellation of Newton-Cotes quadratures with more than 8 points that occur for orders larger than 6, used in the considered IMEX ADER method.

In Figure 10, we display the stability areas for the  $[DEC, EQ, 8, A_k, D_2]$ ,  $[SDEC, EQ, 8, A_k, D_8]$  and  $[ADER, GLB, 8, A_k, D_8]$  increasing the order for the finite difference scheme for the advection operator. For DeC (left), this does not lead to higher values for  $E_0$  or  $C_0$ , but seemingly decreases for odd orders and increases for even orders, while starting with relatively big values for order 1 and relatively small values for order 2. It seems that they all guide towards a specific border line. It is also surprising that we do not get a border  $E_0 > 0$  for order 6 in the advection term for equispaced nodes, while we do for GLB ones (not displayed). These phenomena hold for all remaining combinations of time-marching methods, respective orders and nodes also, i.e. if we increase the order of the diffusion term, we can eliminate instabilities that occur due to the advection term, while the remaining variables just play a minor role. Exemplary, we display in Figure 10 on the right the

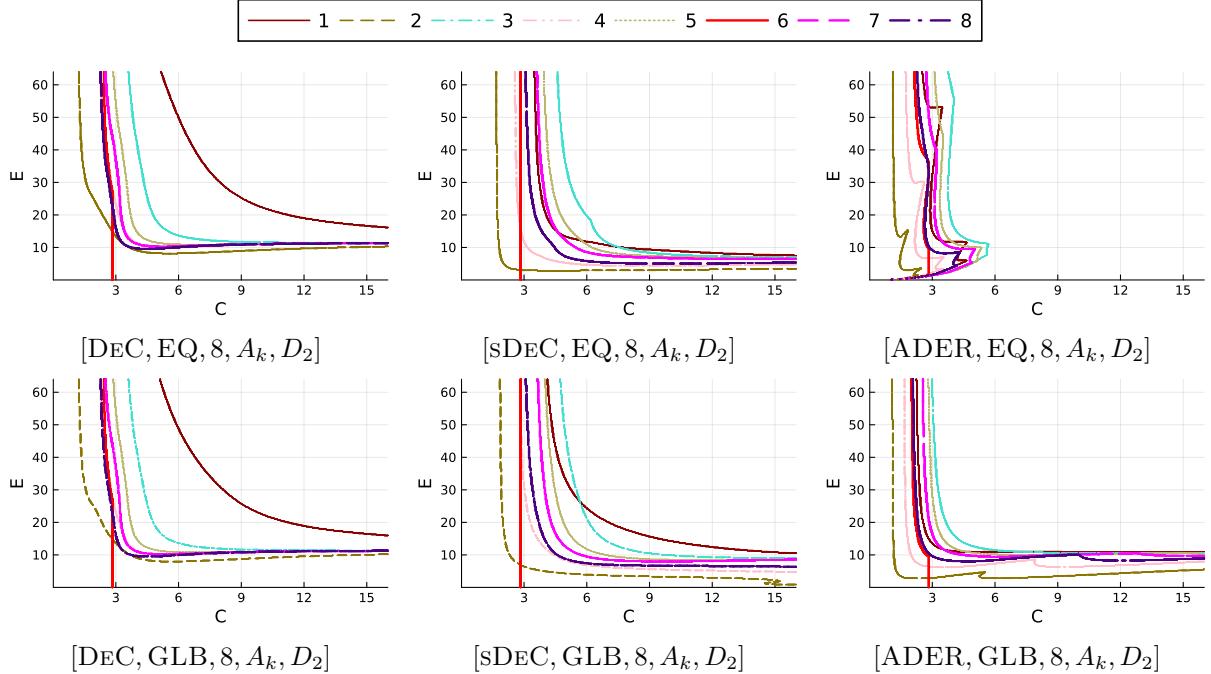


Figure 10: DeC Stability areas varying the order of the advection method

[ADER, GLB, 8,  $A_k, D_2$ ] to underline this statement. This is again not true for order 6 for which all time methods do not show a stability region of type  $E \leq E_0$  and for all IMEX ADER with equispaced nodes. In Figure 11, we show the stability regions for increasing order of the diffusion operator, keeping the order 8 in time and  $A_1$  for the advection operator, i.e., [TMM, eq, 8,  $A_1, D_k$ ], where TMM is every of the 3 considered methods and  $k \in \{2, 4, 6, 8\}$ . We observe that the stability areas of all methods vary slightly, but do not increase or decrease the stability region in a significant way.

In Figure 12, we study the space–time discretizations matching the orders of the time scheme with the order of the spatial terms, i.e. [TMM, GLB,  $k, A_k, D_{2[k/2]}$ ], where TMM is one of the 3 considered methods and  $k \in \{2, \dots, 8\}$ . Here, we can observe for all cases that the higher order terms results in slightly larger stability areas and also in bigger  $C_0$  and  $E_0$ , which leads us to the behavior we have already seen in Figure 9 varying only the order of the time scheme.

We can conclude that increasing the order of the time-marching method results in higher values for both  $C_0$  and  $E_0$ . However, the considered higher order finite differences for the spatial discretizations do not grant significant improvements. A special mention to high order IMEX ADER with equispaced nodes is necessary also here. The border  $E_0$  vanishes, indeed, for order  $> 7$ , see Figure 9 (top right). Moreover, also varying the spatial discretization we obtain the same results.

As an example, we plot the stability region of [sDEC, GLB, 2,  $A_1, D_2$ ] in Figure 13. We observe an asymptotic behavior both for  $E \rightarrow \infty$  against a line  $C = C_0$  and also some sort of border line  $E < E_0$  for large values of  $C$ , which ensures stability for an arbitrary  $C$ , if  $E \leq E_0$ . These are the desired values for  $C_0$  and  $E_0$ . In Table 3, we study the operators with order  $k$  matching for the time and spatial discretization for time schemes defined by GLB nodes. We display in that table the maximal values  $C_0$  and  $E_0$ . We clearly see that they increase as the order increases. The only value that is not uniform among the methods is  $E_0$  for ADER GLB of order 2, for which we have a restrictive bound. We also want to highlight, that  $C_0$  matches in between the DeC and ADER methods for the same orders. This is probably due to their coinciding explicit stability regions for ODEs, as pointed out in [5].

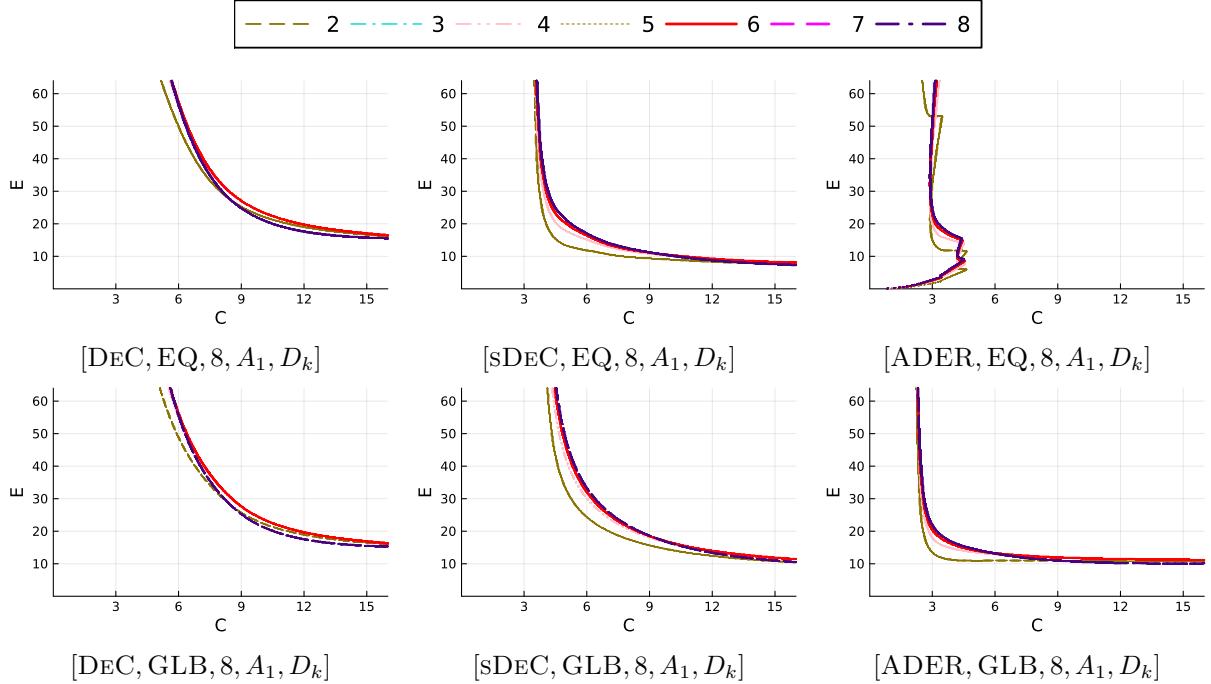


Figure 11: Stability areas varying the order of the diffusion method

## 4 PDE: analysis of advection-dispersion

In this section, we extend the analysis and results to observe the behavior of IMEX ADER and DeC methods onto the advection-dispersion equation

$$u_t(x, t) + au_x(x, t) + \beta u_{xxx}(x, t) = 0, \quad a \geq 0, \quad \beta \geq 0. \quad (9)$$

### 4.1 FD discretization

First, we introduce at this point the considered spatial discretizations for the advection-dispersion equation. Thereby, we consider the same discretization for the advection term, as introduced in Section 3.1. For the dispersion term, we will take two different types of methods into account: Central finite difference and upwind schemes. The underlying theory and assumptions are the same as we saw previously and their

Table 3: Approximated border values  $C_0$  (up to 2 decimals) and  $E_0$  (up to 1 decimal) for Gauss–Lobatto methods with operators with optimal order  $k$

$k$	[DEC, GLB, $k$ , $A_k$ , $D_{2\lceil k/2 \rceil}$ ]		[SDEC, GLB, $k$ , $A_k$ , $D_{2\lceil k/2 \rceil}$ ]		[ADER, GLB, $k$ , $A_k$ , $D_{2\lceil k/2 \rceil}$ ]	
	$C_0$	$E_0$	$C_0$	$E_0$	$C_0$	$E_0$
2	0.50	2.5	0.50	2.5	0.50	0.7
3	1.63	6.1	1.69	5.1	1.63	4.5
4	1.04	6.9	1.43	4.9	1.04	4.2
5	1.74	8.8	2.31	6.6	1.74	7.2
6	1.60	4.1	2.33	4.2	1.60	4.1
7	1.94	9.5	3.12	7.5	1.94	8.5
8	2.00	10.2	2.85	5.9	2.00	9.8

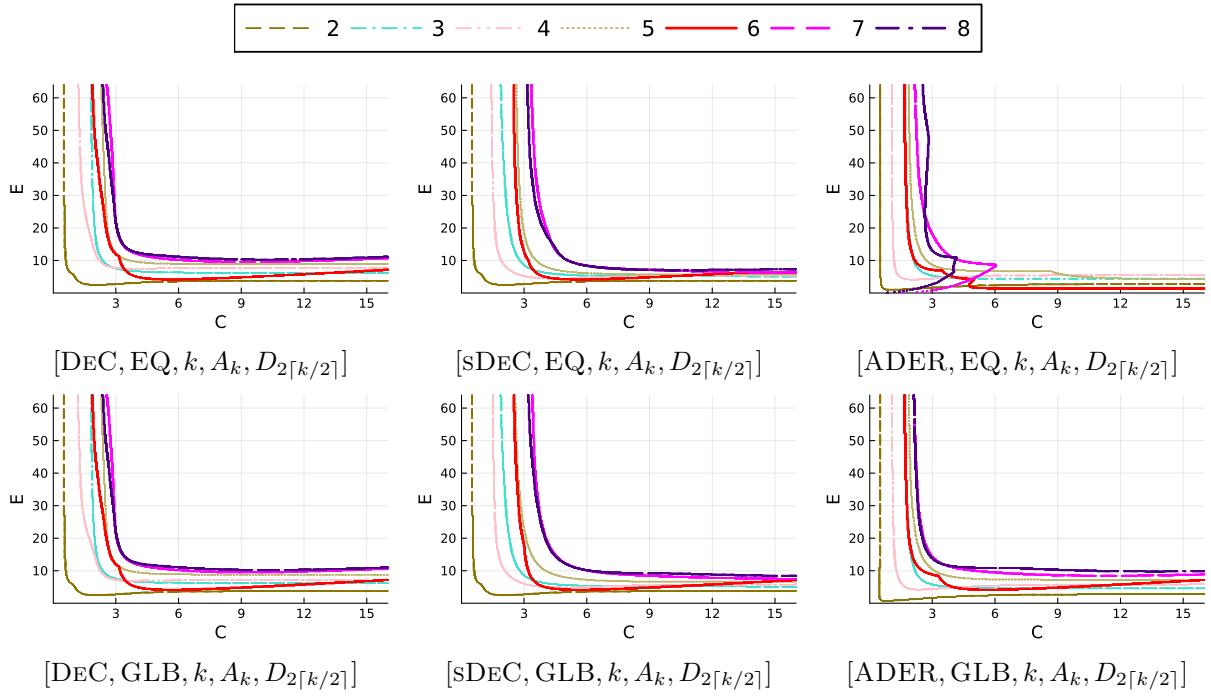


Figure 12: Stability areas varying the order of all partial methods.

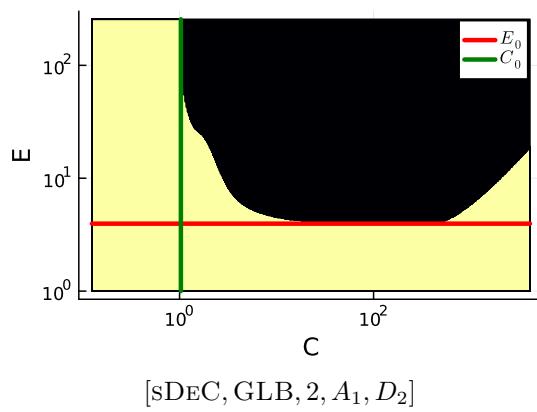


Figure 13: Bounds by  $C_0$  and  $E_0$  of the stability region (in yellow) for an sDeC method.

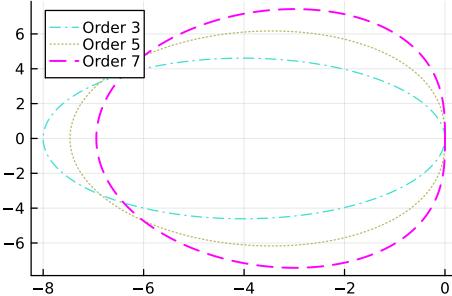


Figure 14: Fourier coefficient of the dispersion operators  $B_{2r+1}$  varying the Fourier mode.

orders can be proven analogously. We consider the following central finite difference scheme of order 2

$$B_2 : \quad \partial_{\Delta x}^3(u(x_j)) = \frac{1}{\Delta x^3} \left( -\frac{1}{2}w_{j-2} + w_{j-1} - w_{j+1} + \frac{1}{2}w_{j+2} \right) \quad (10)$$

and the upwind scheme used in [20] to test stability for the advection-dispersion equation (9). It is of order 3 and given by

$$B_3 : \quad \partial_{\Delta x}^3(u(x_j)) = \frac{1}{4\Delta x^3} (-w_{j-2} - w_{j-1} + 10w_j - 14w_{j+1} + 7w_{j+2} - w_{j+3}). \quad (11)$$

For higher orders, we have used the optimal  $2r + 1$  order formula on stencils of the type  $[-r, r + 1]$  with the tool provided in [22] and we denote them by  $B_{2r+1}$ . In Figure 14, the Fourier value of these dispersion operators are shown. We do not plot  $B_2$  that lies on the imaginary axis.

## 4.2 von Neumann analysis

As previously done for the advection–diffusion problem, we will perform the von Neumann analysis by looking at the coefficients of the finite difference schemes, i.e.,

$$C = a \frac{\Delta t}{\Delta x}, \quad P = \beta \frac{\Delta t}{\Delta x^3}.$$

The procedure is analogous to the advection–diffusion one, with  $C, P$  instead of  $C, E$ .

### 4.2.1 Displaying stability

To denote the considered methods, we use again the notation introduced for the advection-diffusion equation

$$[\text{TMM, NODES, } N, A_n, B_m],$$

where  $B_m$  refers to the upwind  $m$ -th order stencils of type  $[-r, r + 1]$ . We proceed evaluating the amplification factor

$$G(k, \Delta x, \Delta t, a, \beta) = g(k, C, P)$$

to observe the stability region as a function of  $C$  and  $P$ . In opposition to the advection–diffusion case, in [20] only a CFL condition is found, even if, numerically, they observe larger stability regions with a little of dispersion. We want to give a more comprehensive study of this behavior for different schemes and, as before, we look for meaningful coefficients that bounded by some constants give the stability. To find such coefficients, we proceed with an example.

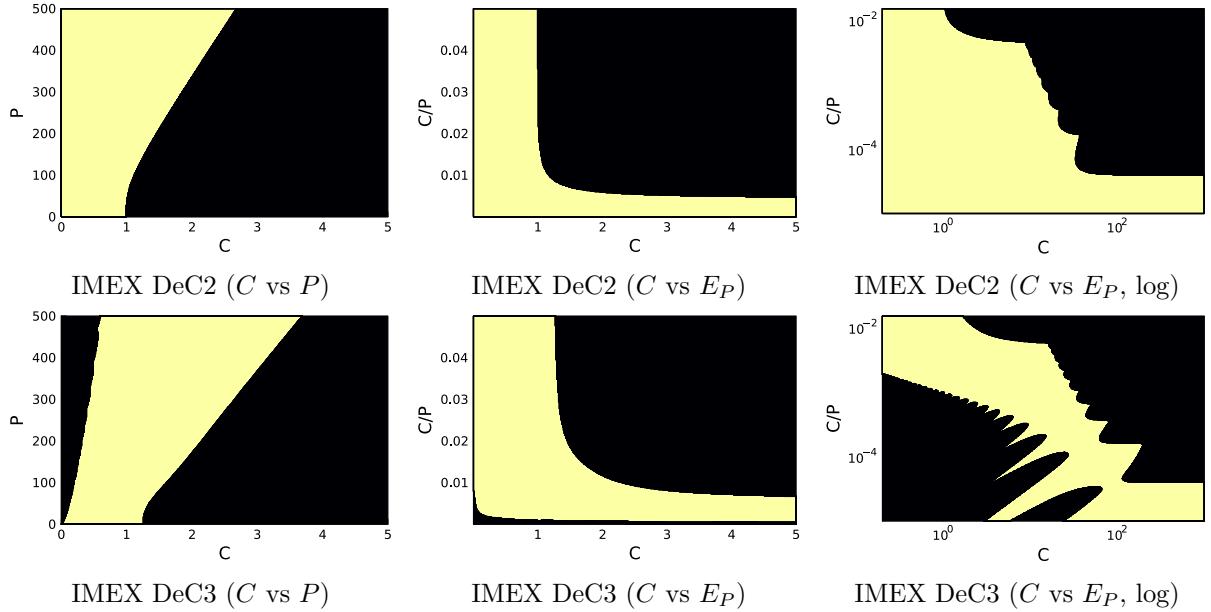


Figure 15: Stability areas for [DEC, EQ, 2,  $A_1, B_2$ ] and [DEC, EQ, 3,  $A_1, B_2$ ] with second order central dispersion operator

**Example 4.1.** In Figure 15 we display the stability areas for the [DEC, EQ, 2,  $A_1, B_2$ ] and [DEC, EQ, 3,  $A_1, B_2$ ] on the  $(C, P)$  plane (left). In the IMEXDeC2 case, we note that for low  $P$  a CFL constraint  $C \leq 1$  guarantees stability, while for large  $P$  we see a linear constraint of the type  $P \gtrsim E_0 C$ . In the IMEXDeC3 case, there is a further unstable region close to the  $C = 0$  axis. This extra unstable region is due to the fact that IMDeC3 is not A-stable and, hence, for low values of  $C$  not enough numerical dissipation is brought to the system.

Anyway, the linear constraint on the large  $P$  motivates the following definition of

$$E_P := \frac{C}{P} = \frac{\Delta t a}{\Delta x} \frac{\Delta x^3}{\beta \Delta t} = \frac{a \Delta x^2}{\beta}.$$

Now, looking at the right plot for IMEXDeC2 in Figure 15, we observe that either  $C \leq 1$  or  $E_P \leq E_0 \approx 0.007$  guarantee stability. This is a peculiar result as  $E_P = \frac{a}{\beta \Delta x^2}$  does not depend on the time discretization. The same does not hold of IMEXDeC3, where this area is stable only for large values of  $C$ , which leads to very inaccurate schemes, as we need  $E_P$  small enough, i.e.,  $\Delta x$  large enough and  $C$  large enough, which implies  $\Delta t$  even larger. Moreover, for the IMEXDeC3 the  $E_0$ -border is much smaller than order 2, so the stable region is even tighter for large  $C$ .

These exemplary stability regions hold for most of the considered cases, i.e. all methods of order 2 do not have the instability areas for small  $C$  and large  $P$ , as well as the IMEX ADER methods with equispaced nodes until order 4 and all IMEX ADER methods with Gauss-Lobatto nodes. Remark that these are exactly the methods which seem to be A-stable in their implicit ODE application as discussed in section 2. All remaining methods possess this unfavorable stability region.

**Example 4.2.** In Figure 16, we display the stability areas for the [DEC, EQ, 2,  $A_1, B_3$ ] and [DEC, EQ, 3,  $A_1, B_3$ ] to compare the results to example 4.1, in which the only difference is the dispersion term:  $B_3$  instead of  $B_2$ . On average, the stability regions increase in this example, in particular  $C_0$  is larger, while the small CFL unstable area remains almost unchanged. This shows that also using upwinded dispersion terms we can not resolve the instability for small  $C$  and large  $P$  for non A-stable schemes. Moreover, among the  $B_r$  schemes

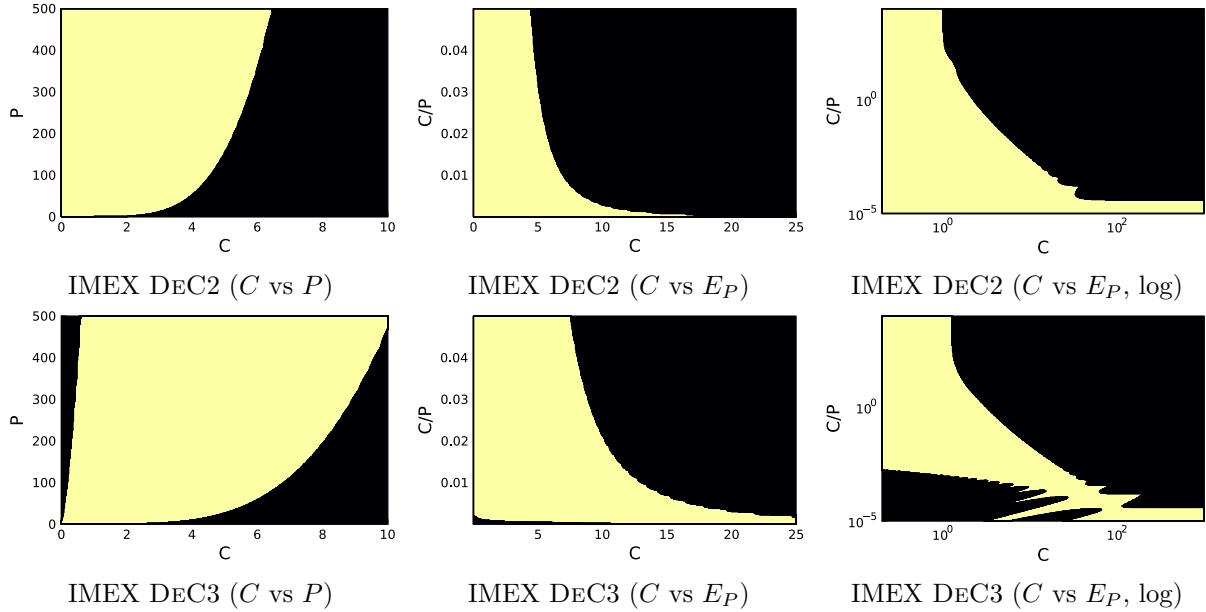


Figure 16: Stability areas for  $[DEC, EQ, 2, A_1, B_3]$  and  $[DEC, EQ, 3, A_1, B_3]$  with third order upwind dispersion operator

*presented,  $B_3$  is the one that leaves the imaginary axis the fastest, so we have less hope for the higher order operators.*

#### 4.2.2 Results for IMEX DeC, sDeC and ADER

In this section, we present the analysis results as displayed in Examples 4.1 and 4.2, varying numerical methods.

We proceed now studying the stability regions increasing the order of the time scheme only, keeping fixed the advection and dispersion operators ( $A_1$  and  $B_3$ ), later on we also increase the accuracy of the advection and dispersion operators. In Figure 17, we can observe the stability regions changing the time scheme order from 2 to 6 for GLB and EQ nodes. For DeC methods of order larger than 2, we cannot not provide bounds that guarantee the stability for small  $C$  and  $E_P$ . Anyway, away from this area, we observe stable regions for both  $C \leq C_0$  with  $C_0$  values similar to the ones of the advection-diffusion section, see Table 3, and for  $E_P \leq E_{P,0}$  with  $E_{P,0} \approx 4 \cdot 10^{-5}$  independently on the DeC method used. In general, sDeC guarantees more stability in the region with  $C \in [1, 10]$  and  $C/P \in [10^{-4}, 10^{-2}]$ . The differences between equispaced and GLB are not so relevant.

On the other hand, IMEXADER with GLB nodes is very stable and there are clear bounds  $C \leq C_0 \leq 3$  and  $E_P \leq E_{P,0} \leq 10^{-4}$  that guarantee two stability regions. Moreover, there is a large stability area for large  $C \leq 10$  and not so small  $E_P$ . On the contrary, IMEXADER with equispaced points for order more than 4 is much more unstable and only the orders 2, 3 and 4 have bounds  $C \leq C_0$  and  $E_P \geq E_{P,0}$  that guarantee stability. This is quite restrictive.

Again, the behavior of all these schemes reflects the A-stability property of the corresponding implicit methods.

In Figure 18, we check the stability regions varying the advection and time order of accuracy for DeC and ADER GLB methods. We find a loss of stability by increasing the order. We already see a slight reduction of our border  $C_0$  for order 2 and 3. Going onto orders 4, 5 and 6, we obtain way larger unstable regions, in particular in the low  $E_P$  region that were stable in the only high order in time scheme presented above. Only the ADER GLB with order different from 6 keep the stability properties, but with much lower  $C_0$  coefficients.

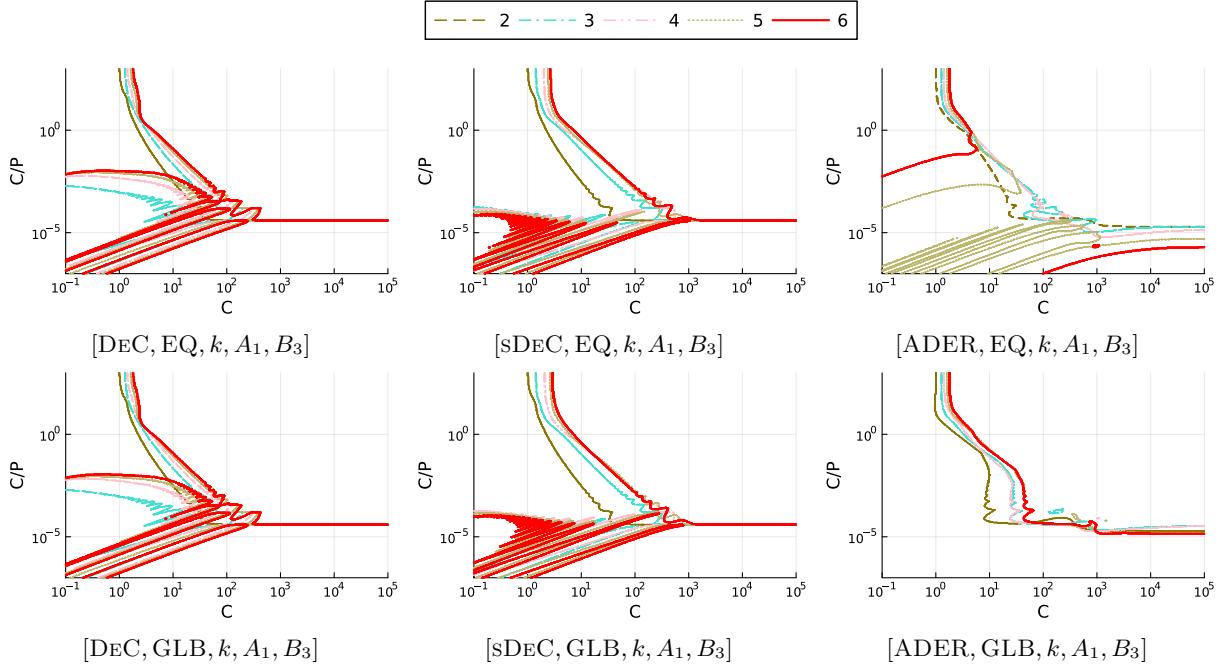


Figure 17: Stability areas for orders 2 to 6 with GLB (top) and equi (bottom) nodes, the upwind scheme of (11) for the dispersion and an first order backward scheme for the advection term

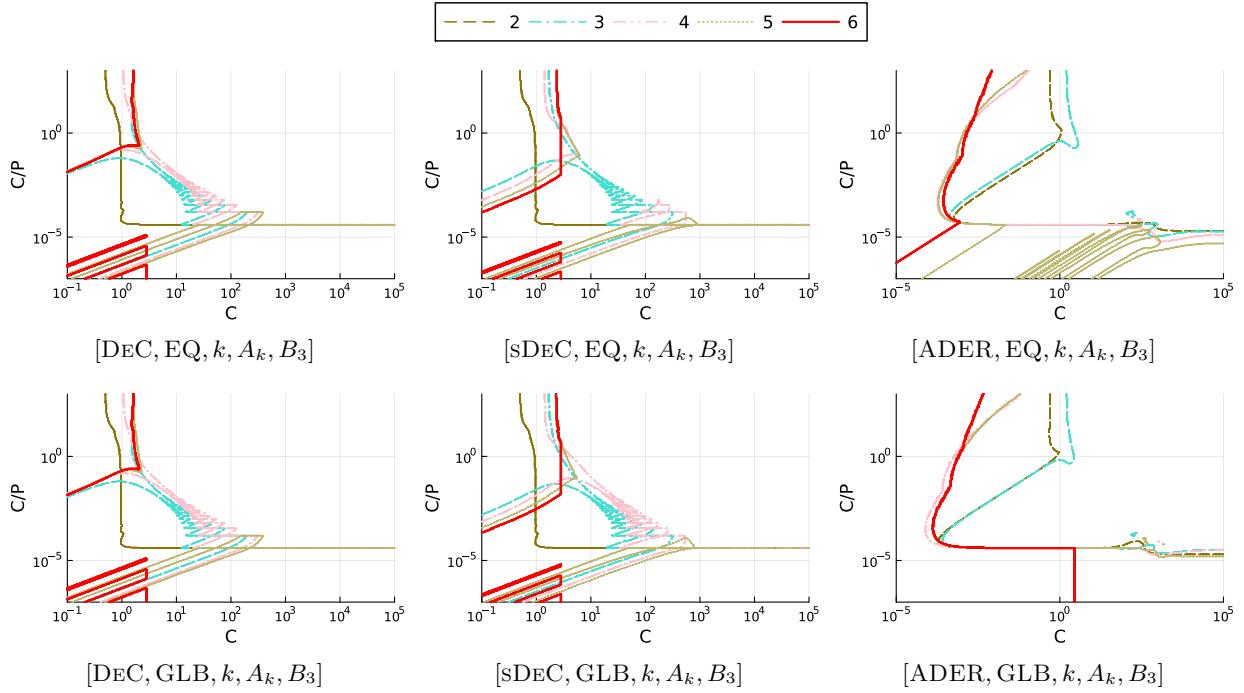


Figure 18: Stability areas varying orders 2 to 6 of the advection scheme and time scheme

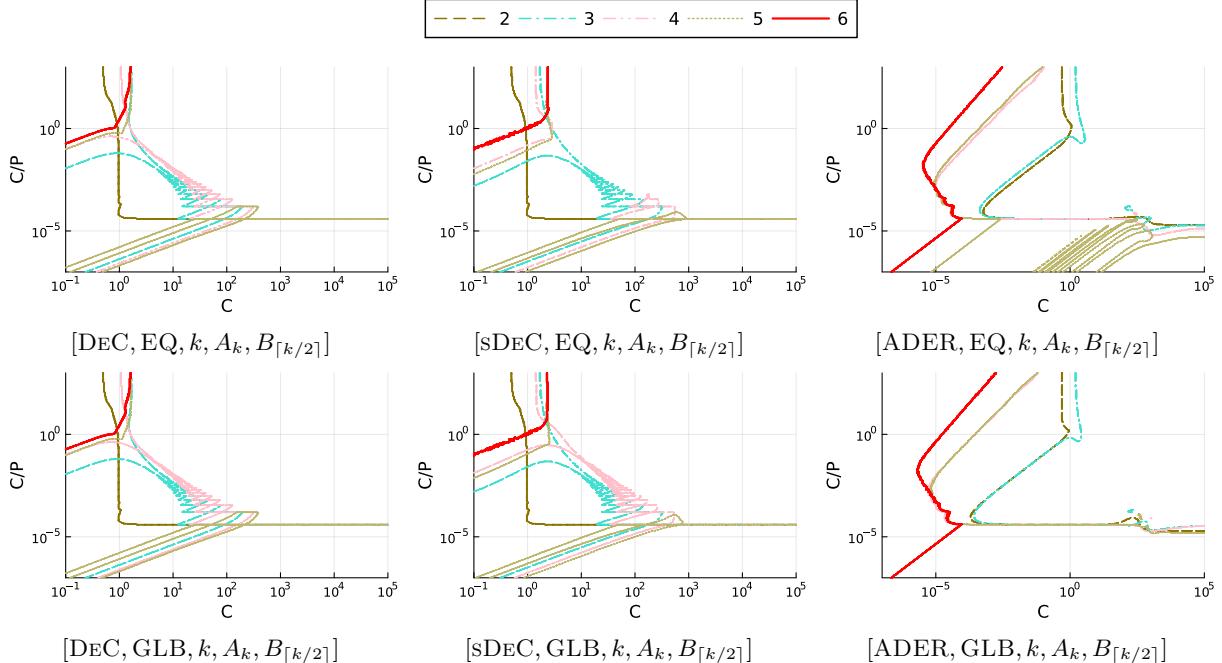


Figure 19: Stability areas for orders 2 to 6 with GLB (top) and equi (bottom) nodes, dispersion with stencil  $[-\lceil(k+2)/2\rceil + 1, \lceil(k+2)/2\rceil]$ , advection of order  $k$  as in (3) and time scheme of order  $k$

In Figure 19, we increase all together the order of all operators. In particular, for a given order  $k$  for the dispersion operator we use the optimal stencil with support  $[-\lceil(k+2)/2\rceil + 1, \lceil(k+2)/2\rceil]$ . The Fourier symbol of the stencils of order 5 and 7 take values very close to the imaginary axis also for quite large imaginary values, see Figure 14. This means that schemes that are not A-stable will poorly perform on such higher orders. On the other hand, the dispersion operator of order 3 is only tangent to the imaginary axis, but it quickly has real values away from zero for large imaginary values. This will influence the stability regions.

We immediately see that the stability regions shrink and for high order DeC (greater than 5), we lose the stability region  $E_P \leq E_{P,0}$ . For ADER methods again the region with  $C \leq C_0$  and moderate  $E_P$  shrinks quite a lot and for order 6 the stability region  $E_P \leq E_{P,0}$  essentially disappears. For the equispaced case, as for the time only case, from order 5 on there is no stable region for low  $E_P$  [17].

We conclude that the observed IMEX methods combined with the finite difference stencils for the spatial discretization do not possess a spatial-independent condition on the time step (as for the diffusion case). Still, in most of the methods a classical stability region for  $C \leq C_0$  and  $E_P \geq E_{P,1}$ , i.e.,  $P \geq C/E_{P,1}$ , is observable, while a time independent stability region for  $E_P \leq E_{P,0}$  is present only in few low order cases and it is really linked to the used spatial discretization.

## Acknowledgements

P.Ö. is supported by the DFG within SPP 2410, project 525866748 (OE 661/5-1) and under the personal grant 520756621 (OE 661/4-1). P.Ö. gratefully acknowledges the support of the Gutenberg Research College, JGU Mainz. D. T. is funded by a SISSA Mathematical Fellowship. L. P. is supported by the DFG within SPP 2410, project number 526031774 and the Daimler und Benz Stiftung (Daimler and Benz foundation), project number 32-10/22.

D. T. and P. Ö. wish to express their gratitude to prof. Giovanni Russo, who brought to their attention

various seminal works on the stability of IMEX schemes. Engaging in discussions with him has deepened our comprehension in this domain.

## References

- [1] R. E. Caflisch, S. Jin, and G. Russo. Uniformly accurate schemes for hyperbolic systems with relaxation. *SIAM Journal on Numerical Analysis*, 34(1):246–281, 1997.
- [2] R. Dematté, V. Titarev, G. Montecinos, and E. Toro. ADER methods for hyperbolic equations with a time-reconstruction solver for the generalized Riemann problem: the scalar case. *Communications on Applied Mathematics and Computation*, 2:369–402, 2020.
- [3] B. Fornberg. Generation of finite difference formulas on arbitrarily spaced grids. *Mathematics of Computation*, 51(184):699–706, 1988.
- [4] E. Hairer, S. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems*. Springer-Verlag, Berlin, 1987.
- [5] M. Han Veiga, P. Öffner, and D. Torlo. DeC and ADER: similarities, differences and a unified framework. *Journal of Scientific Computing*, 87(1), Feb. 2021.
- [6] W. Hundsdorfer and J. Verwer. *Numerical Solution of Time-Dependent Advection-Diffusion-Reaction Equations*. Springer Berlin Heidelberg, 2003.
- [7] A. Iserles. Order stars and a saturation theorem for first-order hyperbolics. *IMA Journal of Numerical Analysis*, 2(1):49–61, 1982.
- [8] R. J. LeVeque. *Finite difference methods for ordinary and partial differential equations: steady-state and time-dependent problems*. SIAM, 2007.
- [9] S. F. Liotta, V. Romano, and G. Russo. Central schemes for balance laws of relaxation type. *SIAM Journal on Numerical Analysis*, 38(4):1337–1356, 2000.
- [10] C. Lozano. Entropy production by explicit Runge-Kutta schemes. *J. Sci. Comput.*, 76(1):521–564, 2018.
- [11] C. Lozano. Entropy production by implicit Runge-Kutta schemes. *J. Sci. Comput.*, 79(3):1832–1853, 2019.
- [12] S. Michel, D. Torlo, M. Ricchiuto, and R. Abgrall. Spectral analysis of continuous FEM for hyperbolic PDEs: influence of approximation, stabilization, and time-stepping. *Journal of Scientific Computing*, 89:1–41, 2021.
- [13] S. Michel, D. Torlo, M. Ricchiuto, and R. Abgrall. Spectral analysis of high order continuous FEM for hyperbolic PDEs on triangular meshes: influence of approximation, stabilization, and time-stepping. *Journal of Scientific Computing*, 94(3):49, 2023.
- [14] M. L. Minion. Semi-implicit spectral deferred correction methods for ordinary differential equations. *Commun. Math. Sci.*, 1(3):471–500, 09 2003.
- [15] P. Öffner, J. Glaubitz, and H. Ranocha. Analysis of artificial dissipation of explicit and implicit time-integration methods. *Int. J. Numer. Anal. Model.*, 17(3):332–349, 2020.
- [16] P. Öffner, L. Petri, and D. Torlo. Analysis for Implicit and Implicit-Explicit ADER and DeC Methods for Ordinary Differential Equations, Advection-Diffusion and Advection-Dispersion Equations. *arXiv*, 2024.

- [17] P. Öffner, L. Petri, and D. Torlo. IMEX\_DeCADER github repository. [https://github.com/accdavlo/IMEX\\_DeCADER.git](https://github.com/accdavlo/IMEX_DeCADER.git), 2024.
- [18] S. Ortlob.  $L^2$ -stability analysis of IMEX- $(\sigma, \mu)$  DG schemes for linear advection-diffusion equations. *Appl. Numer. Math.*, 147:43–65, 2020.
- [19] S. Ortlob. On the stability of IMEX Upwind gSBP schemes for 1D linear advection-diffusion equations. *Communications on Applied Mathematics and Computation*, pages 1–30, 2023.
- [20] M. Tan, J. Cheng, and C.-W. Shu. Stability of high order finite difference schemes with implicit-explicit time-marching for convection-diffusion and convection-dispersion equations. *Int. J. Numer. Anal. Model.*, 18(3):362–383, 2021.
- [21] M. Tan, J. Cheng, and C.-W. Shu. Stability of high order finite difference and local discontinuous Galerkin schemes with explicit-implicit-null time-marching for high order dissipative and dispersive equations. *Journal of Computational Physics*, 464:111314, 2022.
- [22] C. R. Taylor. Finite difference coefficients calculator. <https://web.media.mit.edu/~crtaylor/calculator.html>, 2016.
- [23] V. Titarev and E. F. Toro. Analysis of ADER and ADER-WAF schemes. *IMA journal of numerical analysis*, 27(3):616–630, 2007.
- [24] D. Torlo. *Hyperbolic problems: high order methods and model order reduction*. PhD thesis, Universität Zürich, 2020.
- [25] H. Wang, C.-W. Shu, and Q. Zhang. Stability and error estimates of local discontinuous Galerkin methods with implicit-explicit time-marching for advection-diffusion problems. *SIAM Journal on Numerical Analysis*, 53(1):206–227, Jan. 2015.
- [26] H. Wang, C.-W. Shu, and Q. Zhang. Stability analysis and error estimates of local discontinuous Galerkin methods with implicit-explicit time-marching for nonlinear convection-diffusion problems. *Applied Mathematics and Computation*, 272:237–258, Jan. 2016.
- [27] X. Zhong. Additive semi-implicit Runge–Kutta methods for computing high-speed nonequilibrium reactive flows. *Journal of Computational Physics*, 128(1):19–31, 1996.