

On the equivalence of semi-discrete Active Flux and Discontinuous Galerkin methods and a comparison of their performance

Wasilij Barsukow¹, Christian Klingenberg², Simon Krotsch²

Abstract

The Active Flux (AF) method employs a globally continuous approximation, like continuous Finite Element methods. This is achieved through the placement of point values at cell interfaces which are shared between adjacent cells. With, on average, $K + 1$ degrees of freedom per cell, Active Flux achieves a polynomial approximation of degree $K + 1$, while the Discontinuous Galerkin (DG) method uses only polynomials of degree K , i.e. one degree less with the same number of degrees of freedom. Despite all the differences, in this paper we show, however, that for linear problems in one and several dimensions as well as—in some sense—for nonlinear ones, semi-discrete AF and DG are the same method. We identify a mapping between their respective degrees of freedom, upon which the updates of these degrees of freedom turn out to agree. On the one hand, AF therefore seems more economical than DG for a given value of the error, and we confirm this in numerical experiments. On the other hand, this is a way to understand superconvergence of DG in a natural way, and we show how Radau polynomials and their zeros appear in the mapping between DG and AF: In the Radau points, AF “shines through” as the background high-order scheme behind DG.

Keywords: Superconvergence, Active Flux, Discontinuous Galerkin, Radau polynomials
Mathematics Subject Classification (2010): 65M08, 65M20, 65M60, 76M10, 76M12

1 Introduction

The numerical error of high-order methods decreases as a high power of the discretization length, at least on smooth solutions. The construction of such methods always amounts to an approximation of the solution by a polynomial of high degree; however, it is possible to distinguish different strategies to construct the polynomial. The most important distinction is associated to whether the stencil is compact. If only one degree of freedom is stored per computational cell, as in Finite Difference (FD) or Finite Volume (FV) methods, then a high-degree polynomial can only be constructed by considering sufficiently many neighbours, increasingly far away. Finite Element (FE) methods by construction consider the numerical solution to be a polynomial in every cell, with varying requirements on what happens at cell interfaces. To this end, they store several degrees of freedom in every cell. The evolution of any degree of freedom in some cell depends at most on degrees of freedom in the cells which are immediate neighbours. Such methods shall be referred to as compact.

Compact methods cannot in general be seen as FD methods on a finer mesh, because first of all, the degrees of freedom might not be associated to point values, and second because even then, and even if the point values are equidistant, their update equations are different. Such a method with, say, N degrees of freedom per cell, could in principle be seen as a Finite Difference method on a finer mesh, if one accepts that there are N different update prescriptions and the value at a FD index $i \in \mathbb{Z}$ is updated according to prescription $i \bmod N$. Such an interpretation, however, amounts to no simplification.

¹CNRS, Institut de Mathématiques de Bordeaux (IMB), UMR 5251, 351 Cours de la Libération, 33405 Talence, France, Email: wasilij.barsukow@math.u-bordeaux.fr. Present address: Imperial College London, IRL2004, Huxley Building, South Kensington Campus, London, United Kingdom

²University of Würzburg, Institute of Mathematics, Emil-Fischer-Straße 40, Würzburg, 97074, Germany

While classical Finite Element methods are usually derived from a variational form of the PDE (Galerkin methods), other ways of evolving a given set of degrees of freedom can be conceived. The Active Flux (AF) method, introduced in [vL77, ER13], specifies directly what the update equations for its degrees of freedom are. These are, classically, a cell average and a shared point value at each cell interface, augmented by further point values or moments to obtain higher than 3rd order accuracy. The shared point values enforce global continuity of the approximation.

The original Active Flux method employs so-called evolution operators (e.g. [ER13, ER15, BHKR19, HKS19, Bar21, CHLM24, Bar25a]) to specify the new value of the degrees of freedom directly, amounting to a single-stage fully discrete method. Such methods have advantages, e.g. a large domain of stability and computational efficiency due to their single-stage nature, but for sufficiently complicated problems general algorithms are unavailable so far. While for linear advection back-tracking of characteristics allows to compute the new point value efficiently, for multi-d problems so far evolution operators are only available for specific equations. Here, we instead focus on the semi-discrete version of Active Flux ([Abg22, AB23]), where only space derivatives are discretized while time is left continuous according to a method-of-lines approach. In the case of the semi-discrete Active Flux method, the update equation of the average is inspired by FV methods, but taking into account the continuity, while the update of the point values is inspired by FD methods. It has, however, also been shown recently ([Bar25b]) that the semi-discrete Active Flux method can actually be derived from a variational form by considering a biorthogonal set of test and basis functions. The latter property implies that the (global) mass matrix is just identity. This exemplifies once more that there are different derivations leading to the same method.

We are interested in hyperbolic PDEs, in particular conservation laws. It is well-known that some form of stabilization is required for the numerical method to be amenable to explicit integration in time. This can be done either by directly introducing artificial viscosity in the PDE and thus effectively discretizing a regularized parabolic problem instead (e.g. SUPG, see [BH82]), or by choosing the discretization to have a directional bias (upwinding), an approach inspired by Riemann solvers for FV methods. One of the most prominent examples of Finite Element methods following the latter approach are Discontinuous Galerkin (DG) methods [LR74, CLS89]. They allow for discontinuities at cell interfaces, and use a Riemann solver to compute the flux. Active Flux includes upwinding only in the update of the point values, while continuity of the approximation allows to retain a central update of averages/moments through an integration by parts.

Here, we intend to compare AF and DG theoretically and experimentally. At first glance it seems that DG, based on a discontinuous approximation and a variational form, must be very different from AF, which approximates the numerical solution continuously with direct updates of the degrees of freedom (mass-matrix-free). As has been mentioned above, AF can actually be derived from a variational form, though. In this work, we show a much more surprising result—AF and DG are in many cases exactly the same method. By this we mean that up to a linear mapping between the degrees of freedom of these methods, a DG method with polynomials of degree K amounts to exactly the same update equations for its degrees of freedom as AF based on polynomials of degree $K + 1$. Both methods, however, have some freedom (the choice of numerical flux for DG, the choice of the point value update in AF, etc.), and by equivalence we thus mean that there exists an AF-type method that is equivalent to a given DG method for some given PDE. In this sense, we show equivalence of DG and AF for linear and nonlinear systems in 1-d and for linear advection on Cartesian meshes in 2-d.

In its simplest form, this result appears already in [Roe17] for linear advection in 1-d, but in our opinion is not accorded sufficient importance, even though it is key to understanding the superconvergence of DG. Our present work can be understood as a generalization of this result. We also show why Radau points/polynomials appear naturally in the mapping between DG and AF.

The superconvergence of DG thus finds a new interpretation: at Radau points, the DG approximation agrees with the AF approximation, and the underlying higher-order accuracy of AF can be measured in the DG setting.

High order of accuracy does not tell everything about the properties of the method. In particular, with increasing order the computational effort might also increase, possibly in a prohibitive manner. In the second part of this work, we thus also perform comparisons of computational complexity and cost. Equivalence of the two methods means that they have the same number of degrees of freedom per cell, but certain degrees of freedom might be computationally easier to update than others. Active Flux tends to make more use of point values, whose updates do not require expensive numerical quadratures. In multi-d it is not clear to what level equivalence between AF and DG can be established for non-tensorial bases, of which some have been recently proposed for AF in [BCKL25]. Some limited, but encouraging results are available concerning superconvergence of DG in these more complicated situations, though, e.g. [KF03, CSZ25]. We thus complement our theoretical study by a set of careful measurements of run time of AF and DG as function of order of accuracy, degrees of freedom and numerical error. We give some practical guidance concerning the tradeoff between cost and error.

In [RMF18], which carries a title very similar to that of the present work, only the Active Flux method based on evolution operators was compared to DG. While [RMF18] finds that this fully discrete version of Active Flux is more efficient than DG, in the present work we find that a similar statement still holds even for the semi-discrete AF, which employs the method of lines and Runge-Kutta time integrators, just as DG. Most importantly, however, the theoretical aspects of equivalence and superconvergence were not addressed in [RMF18], which was entirely devoted to a practical comparison. We recently have also become aware of [AT25], where the relation between AF and DG is studied, but without sufficient attention to the important role played by the upwinding in the point value update/numerical flux in DG, and without making the link to superconvergence of DG.

The paper is structured as follows: Section 2 gives an overview of AF, DG and the all-important Radau polynomials. Section 3 presents the equivalence and superconvergence results. Section 4 is dedicated to an analysis of the expected computational complexity and a comparison to experimental measurements.

A computational cell C_i in 1-d is $[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, with x_i its centroid. We assume $\Delta x := x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$ to be constant. $P^K(I)$ is the space of univariate polynomials of degree at most K on the interval I , which we will omit whenever clear. Certain polynomials carry as a superscript their degree (usually, K or $K + 1$); we believe that no confusion with powers is reasonably possible. Computational cells in 2-d are $C_{ij} := [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}] \times [y_{j-\frac{1}{2}}, y_{j+\frac{1}{2}}]$, with $\Delta y := y_{j+\frac{1}{2}} - y_{j-\frac{1}{2}}$, also considered constant. We restrict ourselves here to tensor-product spaces $P^{K,K} \equiv Q^K := P^K \times P^K$. Indices never denote derivatives.

2 Cast

In one spatial dimension, denote by V^{K+1} and V_{br}^K the following spaces:

- $V^{K+1} := \{v \in C^0, v|_{C_i} \in P^{K+1} \forall i \in \mathbb{Z}\}$ is the space of continuous, piecewise polynomial (of degree $K + 1$) functions (the natural approximation space of AF) and
- $V_{\text{br}}^K := \{v, v|_{C_i} \in P^K \forall i \in \mathbb{Z}\}$ is the space of piecewise polynomial functions of degree K (the natural approximation space of DG).

Both have K degrees of freedom per cell. We use the letter q for the exact solution of the conservation law under consideration. To distinguish approximations and degrees of freedom of DG and AF, we denote the latter in capitals, i.e. the approximations are $q_h \in V_{\text{br}}^K$ and $Q_h \in V^{K+1}$, where the subscript h here, by tradition, merely indicates the fact of being an approximation, since we prefer the notation $\Delta x, \Delta y$ for the grid cell size otherwise. We denote their restrictions onto cell C_i by $q_i \in P^K \left(\left[-\frac{\Delta x}{2}, \frac{\Delta x}{2} \right] \right)$ and $Q_i \in P^{K+1} \left(\left[-\frac{\Delta x}{2}, \frac{\Delta x}{2} \right] \right)$, respectively. In multiple dimensions, we consider tensor-product spaces $V^{K+1} = \{v \in C^0, v|_{C_{ij}} \in Q^{K+1}\}$ and $V_{\text{br}}^K = \{v, v|_{C_{ij}} \in Q^K\}$ instead, with analogous restrictions

$$q_{ij} \in Q^K \left(\left[-\frac{\Delta x}{2}, \frac{\Delta x}{2} \right] \times \left[-\frac{\Delta y}{2}, \frac{\Delta y}{2} \right] \right) \quad Q_{ij} \in Q^{K+1} \left(\left[-\frac{\Delta x}{2}, \frac{\Delta x}{2} \right] \times \left[-\frac{\Delta y}{2}, \frac{\Delta y}{2} \right] \right)$$

The dependence of Q_h or q_h on time will frequently be omitted, and we will not adapt the notation for the spaces in case of PDE systems, i.e. when q actually has several components.

2.1 Active Flux

By Active Flux we mean in the following the semi-discrete arbitrary-order generalization with additional moments (such as [\[AB23\]](#), [\[BCKL25\]](#)), discussed now in one and several space dimensions separately.

2.1.1 One spatial dimension

The arbitrary-order version of AF with additional moments appeared first in [\[Abg22\]](#) (for 1-d). The degrees of freedom are the (shared) point values at cell interfaces $Q_{i+\frac{1}{2}}(t) \simeq q(t, x_{i+\frac{1}{2}})$ and the moments

$$Q_i^{(k)} \simeq \frac{A_k}{\Delta x} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} b_k(x - x_i) q(t, x) dx \quad k = 0, \dots, K-1 \quad (1)$$

where b_k span $P^{K-1} \left(\left[-\frac{\Delta x}{2}, \frac{\Delta x}{2} \right] \right)$, and A_k are normalizations. A possible choice is $b_k = \left(\frac{2x}{\Delta x} \right)^k$ and $A_k = k+1$, which ensures that $Q_i^{(k)} = 1$ for $q \equiv 1$ if k even. Given $Q_{i\pm\frac{1}{2}}$ and $\{Q_i^{(k)}\}_{k=0, \dots, K-1}$, in every cell a polynomial Q_i of degree $K+1$ is uniquely determined. It is called the AF reconstruction or the AF approximation. The interpolation of the shared point values implies global continuity: $Q_i\left(\frac{\Delta x}{2}\right) = Q_{i+\frac{1}{2}} = Q_{i+1}\left(-\frac{\Delta x}{2}\right)$.

The semi-discrete Active Flux method for the $m \times m$ hyperbolic system of conservation laws

$$\partial_t q + \partial_x f(q) = 0 \quad f: \mathbb{R}^m \rightarrow \mathbb{R}^m \quad (2)$$

then amounts to the following moment update

$$\frac{d}{dt} Q^{(k)} - \frac{A_k}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b'_k(x) f(Q_i(x)) dx + A_k \frac{b_k^+ f(Q_{i+\frac{1}{2}}) - b_k^- f(Q_{i-\frac{1}{2}})}{\Delta x} = 0 \quad (3)$$

where $b_k^\pm := b_k(\pm \Delta x/2)$. Observe that no Riemann fluxes are necessary. The integral can be evaluated exactly for linear f or by quadrature. The point value update can, for example, be

$$\frac{d}{dt} Q_{i+\frac{1}{2}} + J^+ \frac{d}{dx} Q_i \Big|_{x=\frac{\Delta x}{2}} + J^- \frac{d}{dx} Q_{i+1} \Big|_{x=-\frac{\Delta x}{2}} = 0 \quad (4)$$

but there exist other versions (see e.g. [\[ABK25\]](#), [\[DBK25\]](#)) for nonlinear problems. Here, $J := \nabla_q f$ is the flux Jacobian evaluated in $Q_{i+\frac{1}{2}}$ and J^\pm are the positive/negative parts of J according to the sign of the eigenvalues.

2.1.2 Two spatial dimensions

The classical (third-order) AF method in two space dimensions ([ER13]) employs point values at nodes and at edge midpoints, as well as a cell average. On Cartesian grids ([BHKR19, HKS19]) this leads to 4 node values, 4 edge values and 1 average accessible to a grid cell, and thus naturally to a locally Q^2 and globally continuous approximation. The point values are shared and there are 4 degrees of freedom per cell: one average, one node value, one value on midpoints of horizontal edges, and one on vertical ones. As will be seen below, the interplay with DG will suggest new choices of degrees of freedom, that will be introduced then.

Several arbitrary-order extensions using additional points along the cell interfaces and additional moments have been suggested in [BCKL25]. Here, a tensorial approach will appear upon the comparison to DG. It is, however, inefficient from a practical point of view, because the number of degrees of freedom is (asymptotically) about twice the number required to reach the desired order of accuracy. Naive ways to minimize the number of moments involved can lead to a loss of unisolvence. In [BCKL25], the most economical choice was that of serendipity-type degrees of freedom; the reader is referred to this work for further details. Our experimental study is based on this more efficient approach.

2.2 Radau polynomials

It is well-known that the zeros of Radau polynomials are important in the context of superconvergence of DG. We briefly review how they appear in the context of numerical quadrature and contrast them with Legendre polynomials. Recall that long division of a polynomial $f \in P^{2K+1}(I)$, $K \geq 0$ allows to write it as $f = p\mathcal{L}^{K+1} + q$, where $p, q \in P^K(I)$ and \mathcal{L}^{K+1} is a Legendre polynomial of degree $K + 1$ chosen such that

$$\int_I b(x)\mathcal{L}^{K+1}(x) dx = 0 \quad \forall b \in P^K(I) \quad (5)$$

over some interval I . Observe that these are $K + 1$ equations for $K + 2$ coefficients of \mathcal{L}^{K+1} , leaving one free for normalization. But then,

$$\int_I f(x) dx = \int_I q(x) dx \quad (6)$$

while the point values of q and f agree in the zeros of \mathcal{L}^{K+1} . This is Gauss-Legendre quadrature. \mathcal{L}^{K+1} does not vanish at the endpoints of I .

Gauss-Radau quadrature arises upon including one of the endpoints of I , say x^* . Thus, instead of (5), the defining relations for the Radau polynomial $\mathcal{R}^{K+1} \in P^{K+1}(I)$ are

$$\mathcal{R}^{K+1}(x^*) = 0 \quad (7a)$$

$$\int_I b(x)\mathcal{R}^{K+1}(x) dx = 0 \quad \forall b \in P^{K-1} \quad (7b)$$

with $K \geq 1$ this time. Observe that these are again $K + 1$ equations for $K + 2$ coefficients of \mathcal{R}^{K+1} , leaving one free for normalization. Observe also that orthogonality now happens with respect to polynomials one degree lower, such that one is only interested in writing $f = p\mathcal{R}^{K+1} + q$ for $p \in P^{K-1}(I)$. Here, q (the remainder of long division of f by \mathcal{R}^{K+1}) is still of degree K , and f therefore is only of degree $2K$, one less than for Gauss-Legendre quadrature. The rest of the arguments is entirely analogous.

Correspondingly, while

$$\int_I b(f - Pf) dx = 0 \quad \forall b \in P^K(I) \quad (8)$$

defines the L^2 projection $Pf \in P^K(I)$ of $f \in L^2$, the Gauss-Radau projection $P_{\text{GR}}f \in P^K(I)$ is, for $K \geq 1$, defined by

$$f(x^*) - P_{\text{GR}}f(x^*) = 0 \quad (9)$$

$$\int_I b(f - P_{\text{GR}}f) dx = 0 \quad \forall b \in P^{K-1}(I) \quad (10)$$

From (7b) one is easily led to the ansatz of a linear combination of \mathcal{L}^{K+1} and \mathcal{L}^K for \mathcal{R}^{K+1} . The exact coefficients in the linear combination depend on the normalizations chosen for the Legendre and the Radau polynomials and on the interval I . With the usual choices, if x^* is the right endpoint of I , \mathcal{R}^{K+1} is a difference of these Legendre polynomials, and if x^* is the left endpoint, it is their sum.

2.3 Discontinuous Galerkin

The DG method for the conservation law (2) is given by the variational form

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dt} q_i dx - \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v' f(q_i) dx + \frac{v^+ \hat{f}_{i+\frac{1}{2}} - v^- \hat{f}_{i-\frac{1}{2}}}{\Delta x} = 0 \quad (11)$$

for all $v \in P^K([-\frac{\Delta x}{2}, \frac{\Delta x}{2}])$, and where $v^\pm := v(\pm \frac{\Delta x}{2})$. The second integral needs to be replaced by a quadrature. With $q_i^\pm := q_i(\pm \frac{\Delta x}{2})$, the numerical flux $\hat{f}_{i+\frac{1}{2}}$ is assumed to be a two-point flux

$$\hat{f}_{i+\frac{1}{2}} = \hat{f}(q_i^+, q_{i+1}^-) \quad (12)$$

The upwind flux, for example, amounts to the choice

$$\hat{f}_{i+\frac{1}{2}} = \frac{f(q_i^+) + f(q_{i+1}^-)}{2} - \frac{1}{2}(J^+ - J^-)(q_{i+1}^- - q_i^+) \quad (13)$$

The superconvergence properties of DG have been known for a long time and are mentioned, studied and used e.g. in [LR74, BDF94, ADFK02], and there is also extensive literature on e.g. parabolic equations. Concerning recent results on hyperbolic PDEs, we refer the reader to [CS08, CS10, YS12, GZQ13, CZZ14, CLYZ17] for the linear 1-d case, to [CSYZ15, XZ22] for the linear scalar tensor-product 2-d case, to [CSZ17] for variable-coefficient linear problems in 1-d and to [CSYZ18] for nonlinear scalar problems in 1-d, among others. High-order solutions can in certain cases be recovered through post-processing, see e.g. [CLSS03, RSA05].

The following orders of accuracy for smooth solutions are known for DG with P^K approximation (possibly under suitable assumptions on the grid and the initial data) at least for linear problems:

- $2K + 1$ for cell averages and numerical fluxes,
- $K + 2$ at Radau points (zeros of the Radau polynomial) or their generalizations,
- $K + 2$ between the DG solution and the Gauss-Radau projection of the exact solution (see e.g. [ZS04]),

The necessity to consider generalizations of Radau points is connected to choices other than that of the upwind numerical flux (see e.g. [CLYZ17]).

3 Equivalence between DG and Active Flux and superconvergence of DG

In this Section, consider $K \geq 1$. Observe that the space V^{K+1} of continuous, piecewise polynomial (of degree $K + 1$) functions has the same dimension as the DG space V_{br}^K of piecewise polynomial functions of degree K : both have $K + 1$ degrees of freedom per cell.

Definition 3.1. *Methods M_1 and M_2 are equivalent, if there exists a bijection B from the degrees of freedom of M_1 to those of M_2 , such that upon the action of B the update equations of the degrees of freedom of M_1 become those of M_2 .*

In the following we show that the DG method, naturally associated to V_{br}^K , is equivalent to Active Flux, which is naturally associated to V^{K+1} . Often methods are equivalent because they consider different bases of the same approximation space (e.g. modal and nodal DG); in the present case the situation is more interesting since the approximation spaces are not the same. Also, a DG method with polynomial degree K is usually associated to a $K + 1$ order of accuracy; Active Flux with polynomial degree $K + 1$ has order of accuracy $K + 2$ ([AB23]). The equivalence between the methods thus is a path towards understanding the superconvergence of DG.

This paper has been greatly influenced by ideas in [Roe17] and [Coc25], one of which, however, lacks a general proof, and the other does not make a link to Active Flux. In this paper we close these gaps and generalize the results significantly.

3.1 One-dimensional linear problems

The semi-discrete Active Flux method (3)–(4) for the linear advection equation $\partial_t q + U \partial_x q = 0$ for $U > 0$ reads as follows:

$$\frac{d}{dt} Q_{i+\frac{1}{2}} + U \frac{d}{dx} Q_i \Big|_{x=\frac{\Delta x}{2}} = 0 \quad (14)$$

$$\frac{d}{dt} Q_i^{(k)} - U \frac{A_k}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b'_k(x) Q_i(x) dx + U A_k \frac{b_k^+ Q_{i+\frac{1}{2}} - b_k^- Q_{i-\frac{1}{2}}}{\Delta x} = 0 \quad (15)$$

For $U < 0$, the point value is upwinded the other way:

$$\frac{d}{dt} Q_{i+\frac{1}{2}} + U \frac{d}{dx} Q_{i+1} \Big|_{x=-\frac{\Delta x}{2}} = 0 \quad (16)$$

In [Bar25b, BKL⁺25] also the average of these updates has been considered (central Active Flux):

$$\frac{d}{dt} Q_{i+\frac{1}{2}} + \frac{1}{2} U \left(\frac{d}{dx} Q_i \Big|_{x=\frac{\Delta x}{2}} + \frac{d}{dx} Q_{i+1} \Big|_{x=-\frac{\Delta x}{2}} \right) = 0 \quad (17)$$

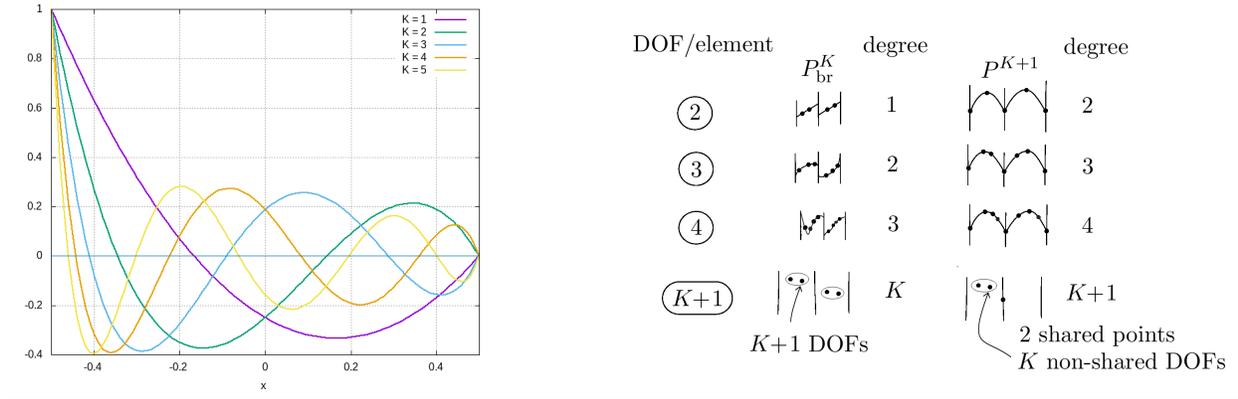
The approximation Q_i in a cell can be written in a basis dual to the degrees of freedom

$$Q_i(x) = Q_{i-\frac{1}{2}} R_{\text{L}}^{K+1}(x) + \sum_{k=0}^{K-1} Q_i^{(k)} S_k^{K+1}(x) + Q_{i+\frac{1}{2}} R_{\text{R}}^{K+1}(x) \quad x \in \left[-\frac{\Delta x}{2}, \frac{\Delta x}{2} \right] \quad (18)$$

with basis functions $R_{\text{L/R}}^{K+1}, S_k^{K+1} \in P^{K+1}$. In fact, $R_{\text{L/R}}^{K+1}$, associated to the point values, are Radau polynomials with the choice $I = [-\frac{\Delta x}{2}, \frac{\Delta x}{2}]$ in (7b) and with the normalization choice is as follows:

$$R_{\text{L}}^{K+1} \left(-\frac{\Delta x}{2} \right) = 1 \quad R_{\text{R}}^{K+1} \left(-\frac{\Delta x}{2} \right) = 0 \quad R_{\text{L}}^{K+1} \left(\frac{\Delta x}{2} \right) = 0 \quad R_{\text{R}}^{K+1} \left(\frac{\Delta x}{2} \right) = 1 \quad (19)$$

Figure 1 Left: Radau polynomials R_L^{K+1} for values $K \in \{1, \dots, 5\}$, with $\Delta x = 1$. Right: Overview of the equivalent DG and AF methods and their degrees of freedom.



The analogue of Equation (7b) for them reads

$$\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b(x) R_L^{K+1}(x) dx = 0 \quad \forall b \in P^{K-1} \quad (20)$$

Radau polynomials are thus a most natural ingredient of the approximation space of Active Flux.

Next, the equivalence between DG and AF for 1-d linear advection is shown. The DG method for the linear advection equation is given by the variational form

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dt} q_i dx - U \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v' q_i dx + U \frac{v^+ \hat{q}_{i+\frac{1}{2}} - v^- \hat{q}_{i-\frac{1}{2}}}{\Delta x} = 0 \quad (21)$$

for all $v \in P^K$ ($[-\frac{\Delta x}{2}, \frac{\Delta x}{2}]$). Here, $\hat{q}_{i+\frac{1}{2}}$ is obtained from the numerical flux $\hat{f}_{i+\frac{1}{2}} \simeq Uq$ as $\hat{q}_{i+\frac{1}{2}} := \hat{f}_{i+\frac{1}{2}}/U$ ($U \neq 0$). The upwind flux amounts to the choices

$$\hat{q}_{i+\frac{1}{2}} = \begin{cases} q_i^+ & U \geq 0 \\ q_{i+1}^- & U < 0 \end{cases} \quad (22)$$

We generalize [Coc25], and adapt the proof to Active Flux. First, upon an integration by parts, (21) can be rewritten as follows:

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dt} q_i dx + U \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v q_i' dx + U \frac{v^+ (\hat{q}_{i+\frac{1}{2}} - q_i^+) - v^- (\hat{q}_{i-\frac{1}{2}} - q_i^-)}{\Delta x} = 0 \quad (23)$$

Observe that now

$$-\frac{v^-}{\Delta x} = \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{d}{dx} (v R_L^{K+1}) dx = \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dx} R_L^{K+1} dx \quad (24)$$

and similarly for v^+ and R_R . The latter equality holds true because $v' \in P^{K-1}$. The DG method thus can be written as

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dt} q_i dx + U \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dx} \left(q_i + (\hat{q}_{i+\frac{1}{2}} - q_i^+) R_R^{K+1} + (\hat{q}_{i-\frac{1}{2}} - q_i^-) R_L^{K+1} \right) dx = 0 \quad (25)$$

Remark 3.2. Observe that the derivative in the second integral is a polynomial of degree K .

Identification 1. Given the DG approximation, the mapping to the degrees of freedom of Active Flux is defined as

$$Q_{i+\frac{1}{2}} := \hat{q}_{i+\frac{1}{2}} \quad Q_i^{(k)} := \frac{A_k}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b_k q_i \, dx \quad \forall k \leq K-1 \quad (26)$$

Remark 3.3. Consider the upwind numerical flux (13) for $U > 0$. Then, $Q_{i+\frac{1}{2}} := q_i^+$, but $q_i^- \neq Q_{i-\frac{1}{2}} = q_{i-1}^+$.

Now, more can be said about the function appearing under the derivative in (25):

Theorem 3.4. The (broken) DG approximation $q_i \in P^K$ augmented by the two Radau polynomials

$$q_i + (\hat{q}_{i+\frac{1}{2}} - q_i^+) R_R^{K+1} + (\hat{q}_{i-\frac{1}{2}} - q_i^-) R_L^{K+1} \quad (27)$$

is the (globally continuous) AF approximation $Q_i \in P^{K+1}$ as it was introduced in Section 2.1.1.

Proof. Using the defining properties (19) of the Radau polynomials and Identification 1 one finds

$$Q_i \left(\frac{\Delta x}{2} \right) = Q_{i+\frac{1}{2}} \quad Q_i \left(-\frac{\Delta x}{2} \right) = Q_{i-\frac{1}{2}} \quad \frac{A_k}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b_k Q_i \, dx = Q_i^{(k)}$$

Thus the DG approximation plus the scaled Radau polynomials is the (continuous) Active Flux approximation (of one polynomial degree higher). \square

The final question is whether the DG method implies the same update equations for the degrees of freedom of Active Flux as the classical Active Flux method. It is answered in the affirmative by the following

Theorem 3.5. For $K \geq 1$, consider the DG method (11) with a linear two-point numerical flux $\hat{f}, \hat{f}: (q_L, q_R) \mapsto \hat{f}(q_L, q_R)$, such that $\hat{f}_{i+\frac{1}{2}} = \hat{f}(q_i^+, q_{i+1}^-)$ is a linear combination $U(\alpha^+ q_i^+ + \alpha^- q_{i+1}^-)$ of the two states. Then, this is the Active Flux method for the degrees of freedom given in the Identification 1, with the average update (4) and the point value update employing the same linear combination of the two adjacent slopes, i.e.

$$\frac{d}{dt} Q_{i+\frac{1}{2}} + U \left(\alpha^+ \frac{d}{dx} Q_i \Big|_{x=\frac{\Delta x}{2}} + \alpha^- \frac{d}{dx} Q_{i+1} \Big|_{x=-\frac{\Delta x}{2}} \right) = 0 \quad (28)$$

Proof. Use the form (25) of the DG method: By Riesz' representation theorem, there exist $v_L, v_R \in P^K$ such that

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v_L q_i \, dx = q_i^- \quad (29)$$

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v_R q_i \, dx = q_i^+ \quad (30)$$

By linearly combining (25) for cells i and $i+1$ with, respectively, v_R and v_L as test function, one obtains the evolution equation for $Q_{i+\frac{1}{2}}$ through Identification 1. Due to Remark 3.2, the second integral in (25) will then be reducing, respectively, to the derivative of Q_i at $x = \frac{\Delta x}{2}$ and that of Q_{i+1} at $x = -\frac{\Delta x}{2}$, giving in total rise to the linear combination given in (28). For the choice of the upwind flux this reduces to Equations (14)/(16) and to (17) for the central flux. The update equation (3) for the moments is immediately obvious by choosing $v = b_k$. \square

The approach immediately generalizes to linear systems in 1-d (see (4)).

Corollary 3.6. *Consider the upwind flux (13) and, for definiteness, $U > 0$. Then, the Active Flux approximation Q_i is*

$$Q_i(x) = q_i(x) + (\hat{q}_{i-\frac{1}{2}} - q_i^-)R_L^{K+1}(x) \quad (31)$$

Active Flux is a method of order of accuracy $K + 2$, while DG is usually associated with an order of accuracy $K + 1$. However, at the zeros of R_L^{K+1} (so-called downwind Radau points, e.g. [YS12]) both approximations agree and the higher order of accuracy becomes visible as superconvergence of DG. We cannot at this point establish the $2K + 1$ superconvergence, valid only at the cell interface and not in the interior Radau points.

Remark 3.7. For the central flux $\hat{q}_{i+\frac{1}{2}} = \frac{q_i^+ + q_{i+1}^-}{2}$, the difference between the DG and the AF approximations remains a weighted sum of both Radau polynomials. Their zeros are distinct such that none of the two sets of points allows to measure superconvergence directly. However, one can study the zeros of the weighted sum of Radau polynomials, see [FR20]. The superconvergence would also become measurable in the moments (which agree), or upon post-processing.

Remark 3.8. The important Remark 3.2 can be rephrased as: The derivative of the AF approximation is in the (broken) DG space. One will see below that the generalization of this statement to multiple dimensions is not straight forward.

3.2 One-dimensional nonlinear problems

While the equivalence is very clear for linear problems, it is less so for nonlinear ones, since there is no unique AF method for them. The point value update requires upwinding, and there are various ways of splitting the Jacobian, and various ways of approximating the flux derivative (compare Jacobian splitting in [ABK25] and Flux-Vector Splitting (FVS) of [DBK25]). We show in the following that in the nonlinear case DG can be rephrased as a new Active Flux method somewhere between the two choices cited.

To this end, consider the $m \times m$ hyperbolic system (2) of conservation laws in this Section. Choosing $v \equiv 1$ in (11) yields the update of the average of the DG approximation

$$\frac{d}{dt} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} q_i dx + \frac{\hat{f}_{i+\frac{1}{2}} - \hat{f}_{i-\frac{1}{2}}}{\Delta x} = 0 \quad (32)$$

The average update (see (3)) for AF reads

$$\frac{d}{dt} Q_i^{(0)} + \frac{f(Q_{i+\frac{1}{2}}) - f(Q_{i-\frac{1}{2}})}{\Delta x} = 0 \quad (33)$$

such that it is natural to identify the two averages and $Q_{i+\frac{1}{2}} := f^{-1}(\hat{f}_{i+\frac{1}{2}})$ wherever f is invertible. (In our opinion this identification really justifies the name “Active Flux”.) For linear f , it reduces to Identification 1 discussed previously. Then, the update equation for the point value must be

$$\frac{d}{dt} Q_{i+\frac{1}{2}} = \left(f'(f^{-1}(\hat{f}_{i+\frac{1}{2}})) \right)^{-1} \left(\frac{\partial \hat{f}}{\partial q_L} \frac{d}{dt} q_i^+ + \frac{\partial \hat{f}}{\partial q_R} \frac{d}{dt} q_{i+1}^- \right) \quad (34)$$

with a two-point numerical flux $\hat{f}(q_L, q_R)$, as defined in (12).

Recall the test functions $v_{L/R}$ defined in (29)–(30), which project out q_i^\pm . Define also $A_{i+\frac{1}{2}} := f'(f^{-1}(\hat{f}_{i+\frac{1}{2}}))$, an (unusual) estimate of the Jacobian. Since $f \circ q$ is not in P^{K+1} (or not even polynomial in general), one cannot proceed as before. Define thus a projection $F_i \in P^{K+1}$ of $f \circ q_i$ into the space of AF approximations such that

$$\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b(F_i - f \circ q_i) dx = 0 \quad \forall b \in P^{K-1} \quad (35a)$$

$$F_i^\pm := F_i \left(\pm \frac{\Delta x}{2} \right) = f(Q_{i\pm\frac{1}{2}}) \equiv \hat{f}_{i\pm\frac{1}{2}} \quad (35b)$$

Observe also that $F \neq f \circ Q_i$ in general. Then, the DG method becomes (no approximation here!), $\forall v \in P^K$

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dt} q_i dx - \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v' F_i dx + \frac{v^+ F_i^+ - v^- F_i^-}{\Delta x} = 0 \quad (36)$$

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dt} q_i dx + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v \frac{d}{dx} F_i dx = 0 \quad (37)$$

which is the equivalent of (25). Since $\frac{d}{dx} F_i \in P^K$, one concludes, as before

$$\frac{d}{dt} Q_{i+\frac{1}{2}} = A_{i+\frac{1}{2}}^{-1} \left(\frac{\partial \hat{f}}{\partial q_L} \frac{d}{dx} F_i \Big|_{x=\frac{\Delta x}{2}} + \frac{\partial \hat{f}}{\partial q_R} \frac{d}{dx} F_{i+1} \Big|_{x=-\frac{\Delta x}{2}} \right) \quad (38)$$

Of course, the derivatives of F_i are different from those of $f \circ q_i$, which is where the projection (35) becomes relevant.

Example 3.9. For $K = 1$ one finds, with $\bar{f}_i := \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} f(q_i) dx$

$$\frac{d}{dt} Q_{i+\frac{1}{2}} = A_{i+\frac{1}{2}}^{-1} \left(\frac{\partial \hat{f}}{\partial q_L} \left(\frac{6\bar{f}_i - 4f(Q_{i+\frac{1}{2}}) - 2f(Q_{i-\frac{1}{2}})}{\Delta x} \right) + \frac{\partial \hat{f}}{\partial q_R} \left(\frac{-6\bar{f}_{i+1} + 2f(Q_{i+\frac{3}{2}}) + 4f(Q_{i+\frac{1}{2}})}{\Delta x} \right) \right)$$

◀

Since the same Finite-Difference-like formulas are used as before, but applied to F instead of Q (as in (4)), it is worth introducing the notation

$$(DQ_i)^+ := \frac{d}{dx} Q_i \Big|_{x=\frac{\Delta x}{2}} \quad (DQ_{i+1})^- := \frac{d}{dx} Q_{i+1} \Big|_{x=-\frac{\Delta x}{2}} \quad (39)$$

Example 3.10. Consider as \hat{f} the global Lax-Friedrichs' flux, such that

$$\hat{f}(q_L, q_R) = \frac{f(q_L) + f(q_R)}{2} - \frac{a}{2}(q_R - q_L) \quad (40)$$

for some $a = \text{const}$. Then

$$\frac{\partial \hat{f}}{\partial q_{L/R}} = \frac{1}{2}(f'(q_{L/R}) \mp a) \quad (41)$$

Thus, $\frac{\partial f}{\partial q_{L/R}}$ are a particular way of choosing the positive/negative part $J_{i+\frac{1}{2}}^\pm$ of the Jacobian f' at $x_{i+\frac{1}{2}}$. With this, the above point value update can be written as

$$\frac{d}{dt}Q_{i+\frac{1}{2}} = - \left(A_{i+\frac{1}{2}}^{-1} J_{i+\frac{1}{2}}^+ (DF_i)^+ + A_{i+\frac{1}{2}}^{-1} J_{i+\frac{1}{2}}^- (DF_{i+1})^- \right) \quad (42)$$

Compare this to the approach

$$\frac{d}{dt}Q_{i+\frac{1}{2}} = - \left(J_{i+\frac{1}{2}}^+ (DQ_i)^+ + J_{i+\frac{1}{2}}^- (DQ_{i+1})^- \right) \quad (43)$$

from [\[ABK25\]](#) (where J^\pm were instead defined via the eigenvalues of J) and the one from [\[DBK25\]](#):

$$\frac{d}{dt}Q_{i+\frac{1}{2}} = - \left((Df_i^+)^+ + (Df_{i+1}^-)^- \right) \quad (44)$$

with the Flux-Vector Splitting $f_i^\pm = \frac{1}{2}(f(Q_i) \pm aQ_i)$ (for example). The DG-inspired AF thus is closely related to existing variants of AF for nonlinear problems. \triangleleft

Most importantly, one thus concludes that DG can be written as an Active Flux method even for nonlinear problems.

3.3 Multi-dimensional linear problems on Cartesian grids

3.3.1 A special choice of degrees of freedom for AF

Above, it has been shown that in 1-d, the classical DG method for linear problems becomes the classical AF method upon a suitable mapping of the degrees of freedom: the two methods are equivalent in the sense of Definition [3.1](#). A natural DG method on 2-d Cartesian grids employs tensor-product basis and test functions.

It turns out (already for third-order accuracy) that it is not equivalent to the classical AF method employing point values at edge midpoints, point values at nodes and cell averages. However, equivalence can be shown for a non-classical point value update. Here, we find it more instructive to instead consider a new, non-classical choice of degrees of freedom for Active Flux, but endowed with a natural update, in a way made precise below. The new third-order accurate, tensorial Active Flux method (see Figure [2](#)) uses as degrees of freedom point values at nodes (as before), a cell average (also unchanged) and one-dimensional averages along the edges (instead of values at their midpoints):

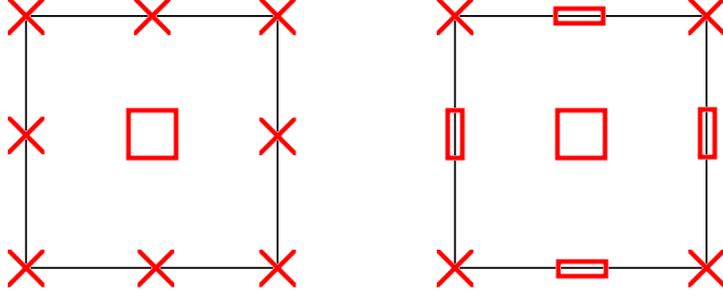
$$\bar{Q}_{i+\frac{1}{2},j} \simeq \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} q(t, x, y) dy \quad (45)$$

$$\bar{Q}_{i,j+\frac{1}{2}} \simeq \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} q(t, x, y) dx \quad (46)$$

These latter will be referred to as edge averages.

These degrees of freedom are those of Virtual Finite Elements (VEM) (see e.g. [\[BdVBC+13\]](#), and in particular [\[ABL25\]](#) for their usage in Active Flux methods on triangular grids). These degrees of freedom might also call for comparison with those of the Whitney finite elements, which associate to each k -cell (in the language of algebraic topology) an average over it. Here, however, orientation is not taken into account and all variables remain scalars. One-dimensional averages were

Figure 2 Left: Classical degrees of freedom of Active Flux on Cartesian grids (third-order accuracy). Right: The modified degrees of freedom proposed here. The edge midpoints are replaced by one-dimensional averages along the edge.



also employed recently in [GHD⁺25], albeit only as an approximation in conjunction with a splitting into one-dimensional problems, and not in an effort to explore alternative multi-dimensional formulations of Active Flux.

The special property associated with the usage of edge averages instead of midpoints is the tensor-product structure of the basis functions. If the one-dimensional AF approximation Q_i is written as

$$Q_i(x) = Q_{i-\frac{1}{2}} R_L^{K+1}(x) + Q_i^{(0)} S_0^{K+1}(x) + Q_{i+\frac{1}{2}} R_R^{K+1}(x) \quad (47)$$

(see Equation (18)), then the two-dimensional tensor-product approximation is

$$\begin{aligned} Q_{ij}(x) &= Q_{i-\frac{1}{2}, j-\frac{1}{2}} R_{L,\Delta x}^{K+1}(x) R_{L,\Delta y}^{K+1}(y) + Q_{i+\frac{1}{2}, j-\frac{1}{2}} R_{R,\Delta x}^{K+1}(x) R_{L,\Delta y}^{K+1}(y) + \dots \\ &+ Q_{ij}^{(0)} S_{0,\Delta x}^{K+1}(x) S_{0,\Delta y}^{K+1}(y) + \bar{Q}_{i-\frac{1}{2}, j} R_{L,\Delta x}^{K+1}(x) S_{0,\Delta y}^{K+1}(y) + \bar{Q}_{i+\frac{1}{2}, j} R_{R,\Delta x}^{K+1}(x) S_{0,\Delta y}^{K+1}(y) \\ &+ \bar{Q}_{i, j-\frac{1}{2}} S_{0,\Delta x}^{K+1}(x) R_{L,\Delta y}^{K+1}(y) + \bar{Q}_{i, j+\frac{1}{2}} S_{0,\Delta x}^{K+1}(x) R_{R,\Delta y}^{K+1}(y) \end{aligned} \quad (48)$$

The usage of classical edge midpoints does not lead to a tensor-product structure of the basis function of AF, as pointed out in [Bar25b]. This property, however, is not used directly to show equivalence between DG and AF below.

Simpson's rule allows to convert between point values and edge averages without losing accuracy:

$$\bar{Q}_{i+\frac{1}{2}, j} = \frac{Q_{i+\frac{1}{2}, j+\frac{1}{2}} + 4Q_{i+\frac{1}{2}, j} + Q_{i+\frac{1}{2}, j-\frac{1}{2}}}{6} \quad (49)$$

Why does it matter then whether a point value or a one-dimensional average is used? The difference lies in what one considers their natural update.

Example 3.11. Consider linear advection with $U^x, U^y > 0$

$$\partial_t q + U^x \partial_x q + U^y \partial_y q = 0 \quad U^x, U^y \in \mathbb{R}^+ \quad (50)$$

The edge midpoint value is naturally updated as

$$\frac{d}{dt} Q_{i+\frac{1}{2}, j} + U^x \frac{\partial}{\partial x} Q_{ij} \Big|_{x=\frac{\Delta x}{2}, y=0} + U^y \frac{\partial}{\partial y} Q_{ij} \Big|_{x=\frac{\Delta x}{2}, y=0} = 0 \quad (51)$$

while the edge-average is naturally updated by integrating (50) over the edge and thus as

$$\frac{d}{dt} \bar{Q}_{i+\frac{1}{2}, j} + U^x \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \frac{\partial}{\partial x} Q_{ij} \Big|_{x=\frac{\Delta x}{2}} dy + U^y \frac{Q_{i+\frac{1}{2}, j+\frac{1}{2}} - Q_{i+\frac{1}{2}, j-\frac{1}{2}}}{\Delta y} = 0 \quad (52)$$

The integral would be replaced by quadrature as:

$$\frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \frac{\partial}{\partial x} Q_{ij} \Big|_{x=\frac{\Delta x}{2}} dy = \frac{1}{6} \left(\frac{\partial}{\partial x} Q_{ij} \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + 4 \frac{\partial}{\partial x} Q_{ij} \Big|_{x=\frac{\Delta x}{2}, y=0} + \frac{\partial}{\partial x} Q_{ij} \Big|_{x=\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} \right) \quad (53)$$

Explicit calculation (omitted) shows that (52) and (51) are not equivalent, in the sense that

$$\frac{d}{dt} \bar{Q}_{i+\frac{1}{2},j} \neq \frac{1}{6} \left(\frac{d}{dt} Q_{i+\frac{1}{2},j+\frac{1}{2}} + 4 \frac{d}{dt} Q_{i+\frac{1}{2},j} + \frac{d}{dt} Q_{i+\frac{1}{2},j-\frac{1}{2}} \right) \quad (54)$$

◁

One can, however, use (54) to define a modified update of $Q_{i+\frac{1}{2},j}$. This way, one will end up with an AF method employing classical degrees of freedom, in particular, point values at edge midpoints, but with a new update prescription for the latter. We prefer the opposite choice here, and use edge averages as non-classical degrees of freedom, but endow them with the natural update (52).

For completeness, the natural update of the nodal point values is (see e.g. [ABK25])

$$\frac{d}{dt} Q_{i+\frac{1}{2},j+\frac{1}{2}} + U^x \frac{\partial}{\partial x} Q_{ij} \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + U^y \frac{\partial}{\partial y} Q_{ij} \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} = 0 \quad (55)$$

and that of the average is

$$\frac{d}{dt} Q_{ij}^{(0)} + U^x \frac{\bar{Q}_{i+\frac{1}{2},j} - \bar{Q}_{i-\frac{1}{2},j}}{\Delta x} + U^y \frac{\bar{Q}_{i,j+\frac{1}{2}} - \bar{Q}_{i,j-\frac{1}{2}}}{\Delta y} = 0 \quad (56)$$

Conversely, it should also be possible to stick to a classical AF method, but a devise a non-classical DG method that is equivalent to it. We will address this as part of future work.

3.3.2 The equivalence of DG and AF

In this Section, we focus on showing equivalence between DG and tensorial AF in the multi-dimensional case for linear advection only, and postpone the nonlinear case and that of systems to future work due to not insignificant technical challenges associated with these.

Consider DG on Cartesian grids with a tensor basis, i.e. test and basis functions $v(x)w(y)$, $v \in P^K \left(\left[-\frac{\Delta x}{2}, \frac{\Delta x}{2} \right] \right)$, $w \in P^K \left(\left[-\frac{\Delta y}{2}, \frac{\Delta y}{2} \right] \right)$, for two-dimensional linear advection (50) with no restrictions on the signs of U^x and U^y :

$$\begin{aligned} & \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} v w \frac{d}{dt} q_{ij} dy dx - \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} (v' w U^x q_{ij} + v w' U^y q_{ij}) dy dx \\ & + \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} w U^x \frac{v^+ \hat{q}_{i+\frac{1}{2},j} - v^- \hat{q}_{i-\frac{1}{2},j}}{\Delta x} dy + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v U^y \frac{w^+ \hat{q}_{i,j+\frac{1}{2}} - w^- \hat{q}_{i,j-\frac{1}{2}}}{\Delta y} dx = 0 \end{aligned} \quad (57)$$

Recall that $\hat{f}_{i+\frac{1}{2},j}^x$ is obtained by means of a numerical flux $\hat{f}^x: \mathbb{R} \times \mathbb{R} \rightarrow \mathbb{R}$ as

$$\hat{f}_{i+\frac{1}{2},j}^x(t, y) := \hat{f}^x \left(q_{ij} \left(t, \frac{\Delta x}{2}, y \right), q_{i+1,j} \left(t, -\frac{\Delta x}{2}, y \right) \right) \quad (58)$$

The time-dependence in q_{ij} will be again omitted now, and analogously for \hat{f}^y .

Here, $\hat{q}_{i+\frac{1}{2},j}, \hat{q}_{i,j+\frac{1}{2}}: \mathbb{R}_0^+ \times \mathbb{R} \rightarrow \mathbb{R}$ are obtained from the numerical fluxes $\hat{f}_{i+\frac{1}{2},j}^x \simeq U^x q, \hat{f}_{i,j+\frac{1}{2}}^y \simeq U^y q$ as

$$\hat{q}_{i+\frac{1}{2},j} := \begin{cases} \hat{f}_{i+\frac{1}{2},j}^x / U^x & U^x \neq 0 \\ 0 & \text{else} \end{cases} \quad \hat{q}_{i,j+\frac{1}{2}} := \begin{cases} \hat{f}_{i,j+\frac{1}{2}}^y / U^y & U^y \neq 0 \\ 0 & \text{else} \end{cases} \quad (59)$$

The upwind flux amounts to choosing

$$\hat{q}_{i+\frac{1}{2},j} = \begin{cases} q_{ij} \left(\frac{\Delta x}{2}, y \right) & U^x > 0 \\ q_{i+1,j} \left(-\frac{\Delta x}{2}, y \right) & \text{else} \end{cases} \quad \hat{q}_{i,j+\frac{1}{2}} = \begin{cases} q_{ij} \left(x, \frac{\Delta y}{2} \right) & U^y > 0 \\ q_{i,j+1} \left(x, -\frac{\Delta y}{2} \right) & \text{else} \end{cases} \quad (60)$$

Upon an integration by parts (57) becomes

$$\begin{aligned} & \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} v w \frac{d}{dt} q_{ij} \, dy dx + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} v w (U^x \partial_x q_{ij} + U^y \partial_y q_{ij}) \, dy dx \\ & + \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} w U^x v^+ \frac{\hat{q}_{i+\frac{1}{2},j} - q_{ij} \left(\frac{\Delta x}{2}, y \right)}{\Delta x} \, dy - \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} w U^x v^- \frac{\hat{q}_{i-\frac{1}{2},j} - q_{ij} \left(-\frac{\Delta x}{2}, y \right)}{\Delta x} \, dy \\ & + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v U^y w^+ \frac{\hat{q}_{i,j+\frac{1}{2}} - q_{ij} \left(x, \frac{\Delta y}{2} \right)}{\Delta y} \, dx - \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v U^y w^- \frac{\hat{q}_{i,j-\frac{1}{2}} - q_{ij} \left(x, -\frac{\Delta y}{2} \right)}{\Delta y} \, dx = 0 \end{aligned} \quad (61)$$

This makes it natural to define

$$r_{ij}^x(x, y) := \left(\hat{q}_{i+\frac{1}{2},j} - q_{ij} \left(\frac{\Delta x}{2}, y \right) \right) R_{R,\Delta x}^{K+1}(x) + \left(\hat{q}_{i-\frac{1}{2},j} - q_{ij} \left(-\frac{\Delta x}{2}, y \right) \right) R_{L,\Delta x}^{K+1}(x) \quad (62)$$

$$r_{ij}^y(x, y) := \left(\hat{q}_{i,j+\frac{1}{2}} - q_{ij} \left(x, \frac{\Delta y}{2} \right) \right) R_{R,\Delta y}^{K+1}(y) + \left(\hat{q}_{i,j-\frac{1}{2}} - q_{ij} \left(x, -\frac{\Delta y}{2} \right) \right) R_{L,\Delta y}^{K+1}(y) \quad (63)$$

Here the length of the interval involved in the definition of the Radau polynomial is made explicit as an additional subscript. Since the test functions factorize, one immediately has

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} v' w r_{ij}^x \, dy dx = 0 \quad \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} v w' r_{ij}^y \, dy dx = 0 \quad (64)$$

since

$$\int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b R_L^{K+1} \, dx = \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} b R_R^{K+1} \, dx = 0 \quad \forall b \in P^{K-1} \quad (65)$$

By using the properties (19) of Radau polynomials one thus finds

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} v w \frac{\partial}{\partial x} r_{ij}^x \, dx dy = \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} w \frac{\partial}{\partial x} (v r_{ij}^x) \, dx dy \quad (66)$$

$$= \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} w(y) \left(v^+ \frac{\hat{q}_{i+\frac{1}{2},j} - q_{ij} \left(\frac{\Delta x}{2}, y \right)}{\Delta x} - v^- \frac{\hat{q}_{i-\frac{1}{2},j} - q_{ij} \left(-\frac{\Delta x}{2}, y \right)}{\Delta x} \right) \, dy \quad (67)$$

and analogously for the other direction. This can be used to replace the corresponding terms in (61) thus obtaining

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} vw \frac{d}{dt} q_{ij} dy dx + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} vw \left(U^x \partial_x (q_{ij} + r_{ij}^x) + U^y \partial_y (q_{ij} + r_{ij}^y) \right) dy dx = 0 \quad (68)$$

Remark 3.12. The situation at this point differs from that of the one-dimensional case (which was Equation (25)): Observe that the derivatives in (68) act on different functions, and neither $q_{ij} + r_{ij}^x$, nor $q_{ij} + r_{ij}^y$ equals the AF approximation (as will be shown next). This is actually very good news: while both $\partial_x (q_{ij} + r_{ij}^x)$ and $\partial_y (q_{ij} + r_{ij}^y)$ are in the DG space Q^K , the same first derivatives of the Q^{K+1} AF approximation would be outside of this space. Having the AF approximation appear there would not allow to generalize the 1-d approach.

Lemma 3.13. *There exist $C_{ij}^{LL}, C_{ij}^{LR}, C_{ij}^{RL}, C_{ij}^{RR} \in \mathbb{R}$ such that*

$$\begin{aligned} Q_{ij} &= q_{ij} + r_{ij}^x + r_{ij}^y \\ &+ C_{ij}^{LL} R_{L,\Delta x}^{K+1}(x) R_{L,\Delta y}^{K+1}(y) + C_{ij}^{LR} R_{L,\Delta x}^{K+1}(x) R_{R,\Delta y}^{K+1}(y) \\ &+ C_{ij}^{RL} R_{R,\Delta x}^{K+1}(x) R_{L,\Delta y}^{K+1}(y) + C_{ij}^{RR} R_{R,\Delta x}^{K+1}(x) R_{R,\Delta y}^{K+1}(y) \end{aligned} \quad (69)$$

with r_{ij}^x, r_{ij}^y given in (62)–(63) is the AF approximation.

The proof can be found in Section B, where it is shown that the statement of this Lemma is independent of the definition of $Q_{i+\frac{1}{2},j+\frac{1}{2}}$.

In 1-d, the point value of Active Flux could be directly linked to the numerical flux of DG (see Identification 1). In multi-d, this has an analogy in the edge averages. The nodal point values, however, involve the x -flux and the y -flux simultaneously. Consider from now on therefore the following linear numerical fluxes:

$$\hat{q}_{i+\frac{1}{2},j}(y) := \alpha^+ q_{ij} \left(\frac{\Delta x}{2}, y \right) + \alpha^- q_{i+1,j} \left(-\frac{\Delta x}{2}, y \right) \quad (70)$$

$$\hat{q}_{i,j+\frac{1}{2}}(x) := \beta^+ q_{ij} \left(x, \frac{\Delta y}{2} \right) + \beta^- q_{i,j+1} \left(x, -\frac{\Delta y}{2} \right) \quad (71)$$

Consistency requires $\alpha^+ + \alpha^- = \beta^+ + \beta^- = 1$. This allows to proceed to the following

Identification 2. *Consider $K = 1$. Starting from the DG approximation, define the degrees of freedom of tensorial AF as*

$$\begin{aligned} Q_{i+\frac{1}{2},j+\frac{1}{2}} &:= \alpha^+ \beta^+ q_{ij} \left(\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) + \alpha^- \beta^+ q_{i+1,j} \left(-\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) \\ &+ \alpha^+ \beta^- q_{i,j+1} \left(\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) + \alpha^- \beta^- q_{i+1,j+1} \left(-\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) \end{aligned} \quad (72)$$

$$\bar{Q}_{i+\frac{1}{2},j} := \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \hat{q}_{i+\frac{1}{2},j}(y) dy \quad \bar{Q}_{i,j+\frac{1}{2}} := \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \hat{q}_{i,j+\frac{1}{2}}(x) dx \quad (73)$$

$$Q_{ij}^{(0)} := \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} q_{ij}(x, y) dy dx \quad (74)$$

This identification is reminiscent of the definition of the generalized Gauss-Radau projections in [XZ22], which, however, do not themselves appear in this discussion. The choice of the point value ensures that

$$\alpha^+ \hat{q}_{i,j+\frac{1}{2}} \left(\frac{\Delta x}{2} \right) + \alpha^- \hat{q}_{i+1,j+\frac{1}{2}} \left(-\frac{\Delta x}{2} \right) = \beta^+ \hat{q}_{i+\frac{1}{2},j} \left(\frac{\Delta y}{2} \right) + \beta^- \hat{q}_{i+\frac{1}{2},j+1} \left(-\frac{\Delta y}{2} \right) = Q_{i+\frac{1}{2},j+\frac{1}{2}} \quad (75)$$

Remark 3.14. Consider $U^x, U^y > 0$ and the upwind flux. Then,

$$Q_{i+\frac{1}{2},j+\frac{1}{2}} := q_{ij} \left(\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) \quad (76)$$

$$\bar{Q}_{i+\frac{1}{2},j} := \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} q_{ij} \left(\frac{\Delta x}{2}, y \right) dy \quad \bar{Q}_{i,j+\frac{1}{2}} := \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} q_{ij} \left(x, \frac{\Delta y}{2} \right) dx \quad (77)$$

The following Lemma establishes the update equations for linear combinations of some special values of the DG approximation:

Lemma 3.15. *For any $\alpha, \beta, \gamma, \delta \in \mathbb{R}$ one has*

$$\begin{aligned} & \frac{d}{dt} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \left(\alpha q_{ij} \left(\frac{\Delta x}{2}, y \right) + \beta q_{i+1,j} \left(-\frac{\Delta x}{2}, y \right) \right) dy \\ & + \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} U^x \left(\alpha \partial_x (q_{ij} + r_{ij}^x) \Big|_{x=\frac{\Delta x}{2}} + \beta \partial_x (q_{i+1,j} + r_{i+1,j}^x) \Big|_{x=-\frac{\Delta x}{2}} \right) dy \\ & + \frac{1}{\Delta y} U^y \left(\left[\alpha \hat{q}_{i,j+\frac{1}{2}} \left(\frac{\Delta x}{2} \right) + \beta \hat{q}_{i+1,j+\frac{1}{2}} \left(-\frac{\Delta x}{2} \right) \right] - \left[\alpha \hat{q}_{i,j-\frac{1}{2}} \left(\frac{\Delta x}{2} \right) + \beta \hat{q}_{i+1,j-\frac{1}{2}} \left(-\frac{\Delta x}{2} \right) \right] \right) = 0 \end{aligned} \quad (78)$$

as well as

$$\begin{aligned} & \frac{d}{dt} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \left(\alpha q_{ij} \left(x, \frac{\Delta y}{2} \right) + \beta q_{i,j+1} \left(x, -\frac{\Delta y}{2} \right) \right) dx \\ & + \frac{1}{\Delta x} U^x \left(\left[\alpha \hat{q}_{i+\frac{1}{2},j} \left(\frac{\Delta y}{2} \right) + \beta \hat{q}_{i+\frac{1}{2},j+1} \left(-\frac{\Delta y}{2} \right) \right] - \left[\alpha \hat{q}_{i-\frac{1}{2},j} \left(\frac{\Delta y}{2} \right) + \beta \hat{q}_{i-\frac{1}{2},j+1} \left(-\frac{\Delta y}{2} \right) \right] \right) \\ & + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} U^y \left(\alpha \partial_y (q_{ij} + r_{ij}^y) \Big|_{y=\frac{\Delta y}{2}} + \beta \partial_y (q_{i,j+1} + r_{i,j+1}^y) \Big|_{y=-\frac{\Delta y}{2}} \right) dx = 0 \end{aligned} \quad (79)$$

and

$$\begin{aligned} & \frac{d}{dt} \left(\alpha q_{ij} \left(\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) + \beta q_{i+1,j} \left(-\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) + \gamma q_{i,j+1} \left(\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) + \delta q_{i+1,j+1} \left(-\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) \right) \\ & + U^x \left(\alpha \partial_x (q_{ij} + r_{ij}^x) \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + \beta \partial_x (q_{i+1,j} + r_{i+1,j}^x) \Big|_{x=-\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} \right. \\ & \quad \left. + \gamma \partial_x (q_{i,j+1} + r_{i,j+1}^x) \Big|_{x=\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} + \delta \partial_x (q_{i+1,j+1} + r_{i+1,j+1}^x) \Big|_{x=-\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} \right) \\ & + U^y \left(\alpha \partial_y (q_{ij} + r_{ij}^y) \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + \beta \partial_y (q_{i+1,j} + r_{i+1,j}^y) \Big|_{x=-\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} \right. \\ & \quad \left. + \gamma \partial_y (q_{i,j+1} + r_{i,j+1}^y) \Big|_{x=\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} + \delta \partial_y (q_{i+1,j+1} + r_{i+1,j+1}^y) \Big|_{x=-\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} \right) = 0 \end{aligned} \quad (80)$$

The proof can be found in Section [A](#).

Even though neither $q_{ij} + r_{ij}^x$, nor $q_{ij} + r_{ij}^y$ agree with the AF approximation Q_{ij} , in certain combinations they do agree along certain edges, as the following Lemma shows:

Lemma 3.16. *One has, for any $x \in [-\frac{\Delta x}{2}, \frac{\Delta x}{2}]$, and for $K = 1$,*

$$\beta^+(q_{ij} + r_{ij}^x) \Big|_{y=\frac{\Delta y}{2}} + \beta^-(q_{i,j+1} + r_{i,j+1}^x) \Big|_{y=-\frac{\Delta y}{2}} = \beta^+ Q_{ij} \Big|_{y=\frac{\Delta y}{2}} + \beta^- Q_{i,j+1} \Big|_{y=\frac{\Delta y}{2}} \quad (81)$$

and analogously, for any $y \in [-\frac{\Delta y}{2}, \frac{\Delta y}{2}]$

$$\alpha^+(q_{ij} + r_{ij}^y) \Big|_{x=\frac{\Delta x}{2}} + \alpha^-(q_{i+1,j} + r_{i+1,j}^y) \Big|_{x=-\frac{\Delta x}{2}} = \alpha^+ Q_{ij} \Big|_{x=\frac{\Delta x}{2}} + \alpha^- Q_{i+1,j} \Big|_{x=-\frac{\Delta x}{2}} \quad (82)$$

The proof can be found in Section [C](#).

Theorem 3.17. *For $K = 1$, the DG method [\(57\)](#) is the Active Flux method [\(52\)](#), [\(55\)](#)–[\(56\)](#) upon identifications [\(72\)](#)–[\(74\)](#).*

Proof. The evolution equation for edge averages is obtained from [\(78\)](#) (Lemma [3.15](#)) with $\alpha = \alpha^+$, $\beta = \alpha^-$:

$$\begin{aligned} \frac{d}{dt} \bar{Q}_{i+\frac{1}{2},j} + \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} U^x \left(\alpha^+ \partial_x (q_{ij} + r_{ij}^x) \Big|_{x=\frac{\Delta x}{2}} + \alpha^- \partial_x (q_{i+1,j} + r_{i+1,j}^x) \Big|_{x=-\frac{\Delta x}{2}} \right) dy \\ + \frac{1}{\Delta y} U^y \left(Q_{i+\frac{1}{2},j+\frac{1}{2}} - Q_{i+\frac{1}{2},j-\frac{1}{2}} \right) = 0 \end{aligned} \quad (83)$$

where [\(75\)](#) has been used. Moreover, due to the integral property of Radau polynomials and the tensor-product structure,

$$\int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \partial_x (q_{ij} + r_{ij}^x) dy = \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \partial_x Q_{ij} dy \quad (84)$$

Then, this equation indeed becomes the natural update [\(52\)](#), with a regularization of the x -derivative by means of the same upwinding as employed in the numerical flux of DG. Analogous arguments apply for the perpendicular edge.

For the point value, consider [\(80\)](#) with $\alpha = \alpha^+ \beta^+$, $\beta = \alpha^- \beta^+$, $\gamma = \alpha^+ \beta^-$, $\delta = \alpha^- \beta^-$:

$$\begin{aligned} \frac{d}{dt} Q_{i+\frac{1}{2},j+\frac{1}{2}} \\ + U^x \left[\alpha^+ \partial_x \left(\beta^+(q_{ij} + r_{ij}^x) \Big|_{y=\frac{\Delta y}{2}} + \beta^-(q_{i,j+1} + r_{i,j+1}^x) \Big|_{y=-\frac{\Delta y}{2}} \right) \Big|_{x=\frac{\Delta x}{2}} \right. \\ \left. + \alpha^- \partial_x \left(\beta^+(q_{i+1,j} + r_{i+1,j}^x) \Big|_{y=\frac{\Delta y}{2}} + \beta^-(q_{i+1,j+1} + r_{i+1,j+1}^x) \Big|_{y=-\frac{\Delta y}{2}} \right) \Big|_{x=-\frac{\Delta x}{2}} \right] \\ + U^y \left[\beta^+ \partial_y \left(\alpha^+(q_{ij} + r_{ij}^y) \Big|_{x=\frac{\Delta x}{2}} + \alpha^-(q_{i+1,j} + r_{i+1,j}^y) \Big|_{x=-\frac{\Delta x}{2}} \right) \Big|_{y=\frac{\Delta y}{2}} \right. \\ \left. + \beta^- \partial_y \left(\alpha^+(q_{i,j+1} + r_{i,j+1}^y) \Big|_{x=\frac{\Delta x}{2}} + \alpha^-(q_{i+1,j+1} + r_{i+1,j+1}^y) \Big|_{x=-\frac{\Delta x}{2}} \right) \Big|_{y=-\frac{\Delta y}{2}} \right] = 0 \end{aligned} \quad (85)$$

With Lemma 3.16 this equation can be rewritten as

$$\begin{aligned}
& \frac{d}{dt} Q_{i+\frac{1}{2},j+\frac{1}{2}} \\
& + U^x \left[\alpha^+ \partial_x \left(\beta^+ Q_{ij} \Big|_{y=\frac{\Delta y}{2}} + \beta^- Q_{i,j+1} \Big|_{y=-\frac{\Delta y}{2}} \right) \Big|_{x=\frac{\Delta x}{2}} \right. \\
& \quad \left. + \alpha^- \partial_x \left(\beta^+ Q_{i+1,j} \Big|_{y=\frac{\Delta y}{2}} + \beta^- Q_{i+1,j+1} \Big|_{y=-\frac{\Delta y}{2}} \right) \Big|_{x=-\frac{\Delta x}{2}} \right] \\
& + U^y \left[\beta^+ \partial_y \left(\alpha^+ Q_{ij} \Big|_{x=\frac{\Delta x}{2}} + \alpha^- Q_{i+1,j} \Big|_{x=-\frac{\Delta x}{2}} \right) \Big|_{y=\frac{\Delta y}{2}} \right. \\
& \quad \left. + \beta^- \partial_y \left(\alpha^+ Q_{i,j+1} \Big|_{x=\frac{\Delta x}{2}} + \alpha^- Q_{i+1,j+1} \Big|_{x=-\frac{\Delta x}{2}} \right) \Big|_{y=-\frac{\Delta y}{2}} \right] = 0
\end{aligned} \tag{86}$$

which indeed is the natural upwind point value update involving a weighted sum of the one-sided derivatives to do multi-dimensional upwinding, and generalizing (55).

The update equation for the cell average follows from (68) with the choice $v = w = 1$:

$$0 = \frac{d}{dt} Q_{ij}^{(0)} + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \left(U^x \partial_x (q_{ij} + r_{ij}^x) + U^y \partial_y (q_{ij} + r_{ij}^y) \right) dy dx \tag{87}$$

$$= \frac{d}{dt} Q_{ij}^{(0)} + U^x \frac{1}{\Delta x} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} (q_{ij} + r_{ij}^x) \Big|_{x=-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} dy + U^y \frac{1}{\Delta x} \frac{1}{\Delta y} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} (q_{ij} + r_{ij}^y) \Big|_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} dx \tag{88}$$

Here, again the integral properties of the Radau polynomials allow to rewrite this as

$$0 = \frac{d}{dt} Q_{ij}^{(0)} + U^x \frac{1}{\Delta x} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} Q_{ij} \Big|_{x=-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} dy + U^y \frac{1}{\Delta x} \frac{1}{\Delta y} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} Q_{ij} \Big|_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} dx \tag{89}$$

which is the usual update (56) of the cell average for AF. □

Remark 3.18. The zeros of $R_{L,\Delta x}^{K+1}$ are straight lines parallel to the y -axis, and those of $R_{L,\Delta y}^{K+1}$ are parallel the x -axis. At the crossing points, i.e. at all points

$$\{(x, y) : R_{L,\Delta x}^{K+1}(x) = 0 \text{ and } R_{L,\Delta y}^{K+1}(y) = 0\} \tag{90}$$

one has $q_{ij} = Q_{ij}$, such that the $K + 2$ order of accuracy of AF becomes immediately visible as superconvergence of DG.

4 Performance Comparison

In this section we investigate the performance of semi-discrete Active Flux methods in two space dimensions of different orders, and compare to that of tensorial modal Discontinuous Galerkin methods. We focus on Active Flux with the minimal possible number of degrees of freedom (point values at the cell interfaces and moments), as proposed in [BCKL25] (serendipity-type, called therein GenAF($M_{\mathcal{T}}$)). This is not the new tensorial version for which equivalence for DG has been shown above, but we prefer to compare with this more established version. Both methods are implemented in a unified way, using identical routines for equivalent computational components wherever applicable. For example, both types of methods use moments as degrees of freedom, and

thus, the computation of them is done in a similar way. Also, we ensured that, for instance, the computational effort to compute the reconstruction is of the same order of magnitude for both types of methods.

In the following, N_{order} denotes the order of the Active Flux or Discontinuous Galerkin methods, N_{dofs} the exclusive degrees of freedom per cell and N_{tdofs} the number of degrees of freedom accessible per cell. This distinction is needed, since the point values of the Active Flux methods are shared between neighbouring cells. The number of moments, edge point values and node point values accessible per cell are named N_{mom} , N_{edge} and N_{node} .

For the modal Discontinuous Galerkin methods we have the following relation:

$$N_{\text{tdofs}} = N_{\text{dofs}} = N_{\text{mom}} = (N_{\text{order}})^2$$

The number of moments and point values per cell of the Active Flux methods is given by

$$N_{\text{node}} = 4, \quad N_{\text{edge}} = 4(N_{\text{order}} - 2) \quad \text{and} \quad N_{\text{mom}} = \max\left(1, \frac{1}{2}(N_{\text{order}} - 4)(N_{\text{order}} - 3)\right)$$

and the number of (exclusive and accessible) degrees of freedom per cell is given by

$$N_{\text{dofs}} = \frac{1}{4}N_{\text{node}} + \frac{1}{2}N_{\text{edge}} + N_{\text{mom}} \quad \text{and} \quad N_{\text{tdofs}} = N_{\text{node}} + N_{\text{edge}} + N_{\text{mom}}.$$

To compute the moments, integrals need to be evaluated numerically, which is done by a Gauss-Legendre quadrature, whose order is chosen as the sum of the order of the method (N_{order}) and the largest exponent of the computed moment. If this value is even, then the next smallest odd number is chosen. As a time discretisation, we use a three-stage, third-order and a five-stage, fourth-order strong stability-preserving Runge-Kutta scheme. In Table [1](#) an overview of the methods considered here is given, where the highest order of integration used by a method of a given order is denoted as N_{int} and the CFL number as C_{CFL} . In our case, the size of the time step is given by $\Delta t = C_{\text{CFL}}\Delta x = N_{\text{order}}^{-\frac{1}{2}}$.

| | Active Flux Method | | | | | Discontinuous Galerkin Method | | | | |
|--------------------|--------------------|-----|------|------|-------|-------------------------------|-----|------|------|------|
| N_{order} | 3 | 4 | 5 | 6 | 7 | 2 | 3 | 4 | 5 | 6 |
| N_{dofs} | 4 | 6 | 8 | 12 | 17 | 4 | 9 | 16 | 25 | 36 |
| N_{tdofs} | 9 | 13 | 17 | 23 | 30 | 4 | 9 | 16 | 25 | 36 |
| N_{mom} | 1 | 1 | 1 | 3 | 6 | 4 | 9 | 16 | 25 | 36 |
| N_{edge} | 4 | 8 | 12 | 16 | 20 | - | - | - | - | - |
| N_{node} | 4 | 4 | 4 | 4 | 4 | - | - | - | - | - |
| N_{int} | 3 | 3 | 5 | 7 | 9 | 3 | 5 | 7 | 9 | 11 |
| C_{CFL} | 0.27 | 0.2 | 0.17 | 0.12 | 0.085 | 0.2 | 0.1 | 0.05 | 0.02 | 0.01 |

Table 1: Overview of the different Active Flux and Discontinuous Galerkin methods.

4.1 Theoretical Expectations

Before we compare the measured results, we will investigate the theoretically expected asymptotic behaviour of the Active Flux and Discontinuous Galerkin methods. First, we consider the runtime, which will be denoted by τ , as a function of the grid and of the method. Using the unified implementation, we expect

$$\tau = \mathcal{O}(N_{\text{timesteps}} \cdot N_{\text{stages}} \cdot N_{\text{cells}} \cdot \max(N_{\text{mom}} \cdot N_{\text{int}} \cdot N_{\text{tdofs}}, N_{\text{mom}} \cdot (N_{\text{int}})^2 \cdot N_{\text{tdofs}})), \quad (91)$$

to be the asymptotic behaviour of the runtime of the Discontinuous Galerkin methods, where N_{cells} is the number of grid cells, $N_{\text{timesteps}}$ the number of time steps and N_{stages} the number of stages of the time discretisation. $N_{\text{timesteps}}$, N_{stages} , N_{tdofs} and N_{int} depend on the specific method and $N_{\text{timesteps}}$ and N_{cells} on the grid. For the Active Flux methods, we expect a slightly different asymptotic behaviour:

$$\tau = \mathcal{O}(N_{\text{timesteps}} \cdot N_{\text{stages}} \cdot N_{\text{cells}} \cdot \max(N_{\text{mom}} \cdot N_{\text{int}} \cdot N_{\text{tdofs}}, N_{\text{mom}} \cdot (N_{\text{int}})^2 \cdot N_{\text{tdofs}}, N_{\text{tdofs}}, N_{\text{edge}} \cdot N_{\text{tdofs}})), \quad (92)$$

where the number of the different degrees of freedom depends on the method. If we fix the used method (varying N_{cells} and thus $N_{\text{timesteps}}$ in (91) and (92) only) and therefore consider τ only as a function of the grid, we expect for the Active Flux and the Discontinuous Galerkin methods:

$$\tau = \mathcal{O}((N_{\text{cells}})^{1.5}),$$

since the number of time steps depends on the chosen spatial step size and C_{CFL} , which is fixed for each method. Another option is to consider the runtime as a function of the numerical error which again depends on the spatial step size. Let $\mathcal{E}_{\text{dofs}}$ be the maximum over all l^2 -errors of the degrees of freedom of a given method. If we consider τ to be a function of $\mathcal{E}_{\text{dofs}}$ and assume that $\mathcal{E}_{\text{dofs}}$ is unaffected by the error introduced by the time discretisation and thereby dominated only by the spatial error, we obtain the following expected asymptotic behaviour of the runtime:

$$\tau = \mathcal{O}\left((\mathcal{E}_{\text{dofs}})^{-\frac{3}{N_{\text{order}}}}\right).$$

Next, we will consider methods of different orders on a fixed grid and with a fixed time discretisation. We define τ_s to be the runtime per time step and obtain for the Discontinuous Galerkin methods from (91) and (92)

$$\tau_s = \mathcal{O}((N_{\text{tdofs}})^3) = \mathcal{O}((N_{\text{dofs}})^3) = \mathcal{O}((N_{\text{order}})^6),$$

if we consider the runtime per time step as a function of the number of degrees of freedom or the order of a specific method. For the Active Flux methods we expect a similar asymptotic behaviour:

$$\tau_s = \mathcal{O}((N_{\text{order}})^6).$$

If we consider the runtime per time step as a function of the number of degrees of freedom, we get

$$\tau_s = \mathcal{O}((N_{\text{tdofs}})^3) \quad \text{and} \quad \tau_s = \mathcal{O}((N_{\text{dofs}})^3).$$

Concerning the total runtime τ , we expect for the Discontinuous Galerkin and the Active Flux methods

$$\tau = \tau_s \cdot \mathcal{O}(N_{\text{timesteps}}),$$

where the number of time steps depends solely on C_{CFL} . Hence, the asymptotic relation between the total runtime and the order or the number of degrees of freedom per cell depends on the asymptotic relation between C_{CFL} and N_{order} , N_{dofs} or N_{tdofs} respectively.

4.2 Numerical Results

For our performance study, we consider the linear advection equation

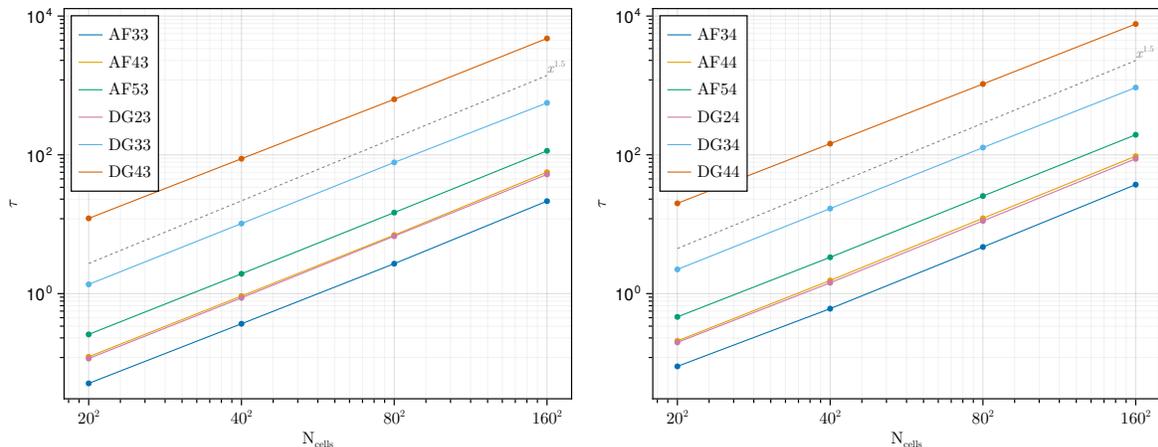
$$\partial_t q + \partial_x q + \partial_y q = 0$$

on $[0, 1]^2$ with Dirichlet boundary conditions and smooth initial data

$$q_0(x, y) = 0.8 + \exp\left(-\left(\frac{x - 0.5}{0.05}\right)^2 - \left(\frac{y - 0.5}{0.05}\right)^2\right).$$

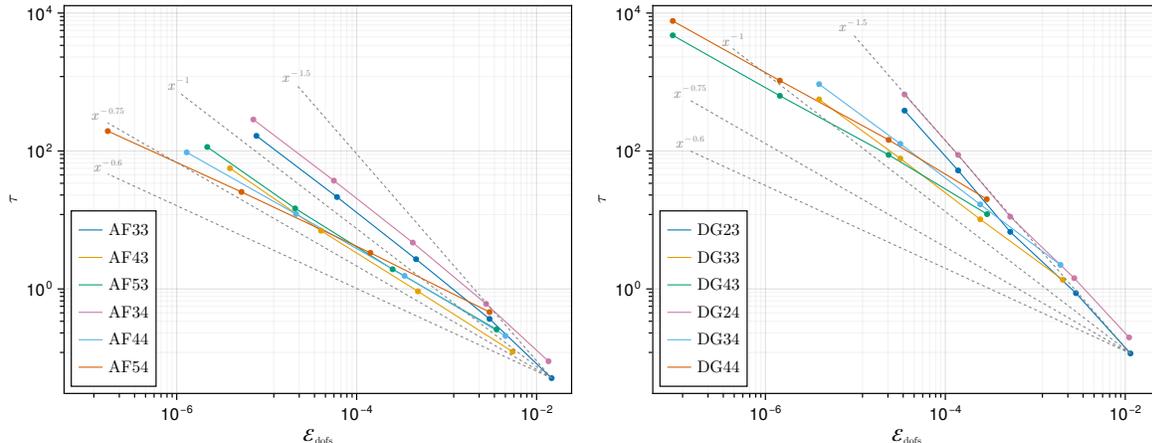
The spatial discretisation uses Cartesian grids with an equal number of cells in x and y directions (i.e., $\Delta x = \Delta y$). To obtain the numerical results, we compute the solution to this problem at time $T = 0.1$ with the different Active Flux and Discontinuous Galerkin methods and measure the total runtime, average runtime per time step and the l^2 -error of the different degrees of freedom. In the following we will refer, for example, to a Discontinuous Galerkin or an Active Flux method of order five with a SSPRK3 time discretisation as DG53 or AF53. For the first numerical test, we consider the runtime τ as a function of N_{cells} . In Figure 3 one can see the measured runtime for the different methods with different values of N_{cells} (i.e., 20^2 , 40^2 , 80^2 and 160^2). As one can see, the Active Flux methods are overall significantly faster than the Discontinuous Galerkin methods with the same order. Note that the curves for AF43 and DG23 and AF44 and DG24 lie on top of each other. In addition, the measured runtime behaves as expected.

Figure 3 The total runtime of the Active Flux and Discontinuous Galerkin methods of different order with respect to the number of cells.



In Figure 4 one can see the runtime with respect to $\mathcal{E}_{\text{dofs}}$. For a given error, the Active Flux methods are capable of computing the numerical results within a shorter runtime. The most runtime-efficient methods with respect to the accuracy are AF33, AF43 and AF54. Furthermore, note that the runtime of AF43 and AF53 appears to have a different behaviour than expected. The reason for this is that the error of methods is presumably dominated by the error of the time discretisation. As one can see in Table 2, the experimental order of convergences of both methods differs from the formal order. By using a fourth-order time discretisation, the fourth and fifth-order Active Flux methods achieve an experimental order of convergence of four and five. This is not needed for the Discontinuous Galerkin method of order four, as the experimental order of convergence for DG43 is already approximately four (see Table 3).

Figure 4 The total runtime of the Active Flux and Discontinuous Galerkin methods of different orders with respect to the error $\mathcal{E}_{\text{dofs}}$.



| Δx | AF33 | AF43 | AF53 | AF34 | AF44 | AF54 |
|------------|--------|--------|--------|--------|--------|--------|
| 0.05 | - | - | - | - | - | - |
| 0.025 | 2.2924 | 3.4956 | 3.8381 | 2.3038 | 3.7278 | 4.3997 |
| 0.0125 | 2.7147 | 3.5786 | 3.5996 | 2.7077 | 4.0076 | 4.761 |
| 0.0625 | 2.9172 | 3.3618 | 3.2449 | 2.9138 | 4.0369 | 4.9365 |
| 0.003125 | 2.9762 | - | - | 2.9743 | - | - |

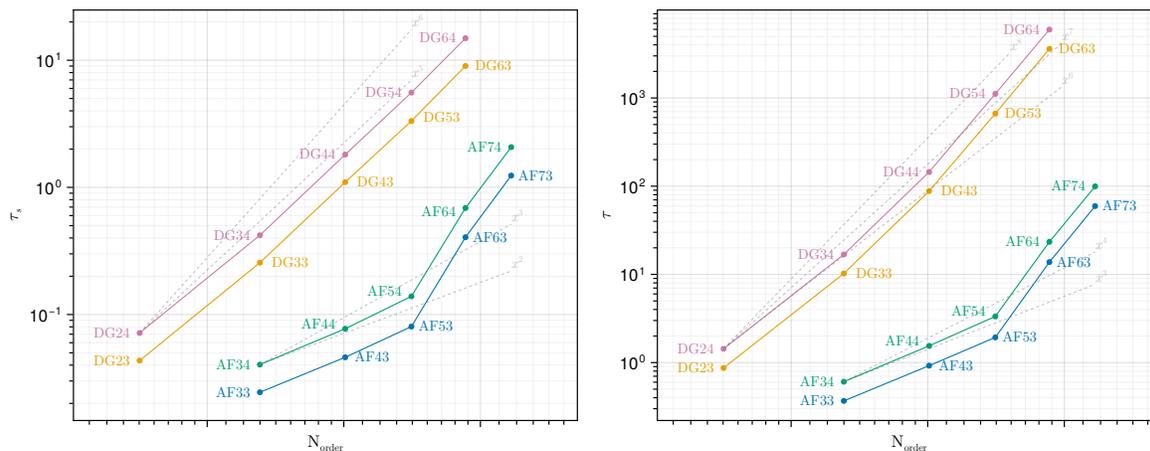
Table 2: Experimental order of convergence of the different Active Flux methods.

| Δx | DG23 | DG33 | DG43 | DG24 | DG34 | DG44 |
|------------|--------|--------|--------|--------|--------|--------|
| 0.05 | - | - | - | - | - | - |
| 0.025 | 2.0138 | 3.0501 | 3.6359 | 2.0268 | 2.9636 | 3.6243 |
| 0.0125 | 2.4265 | 2.9529 | 4.0162 | 2.364 | 2.9496 | 4.0179 |
| 0.0625 | 1.9227 | 2.9984 | 3.9554 | 1.9225 | 2.9983 | 3.9536 |
| 0.003125 | 1.98 | - | - | 1.9801 | - | - |

Table 3: Experimental order of convergence of the different Discontinuous Galerkin methods.

Next, we consider the runtime as a function of the order or the number of exclusive degrees of freedom per cell and therefore all numerical results are computed on a 40×40 grid with $\Delta x = 0.0125$. In Figure 5 one can see the measured average runtime per time step τ_s and the total runtime τ with respect to the orders of the methods. For a given order, the average runtime per time step of the Active Flux methods considered here is shorter than the average runtime per time step of the Discontinuous Galerkin methods. The difference is even greater for the total runtime, because the Active Flux methods allow for a larger CFL number and therefore have fewer time steps. The steep slope between the Active Flux methods of order five and six is due to the introduction of new moments and the associated numerical computation of the integrals, which dominates the runtime. Up to order five, the Active methods only have the zeroth moment (i.e., cell average) as a degree of freedom, whose computation is significantly cheaper than the computation of higher-order moments. Moreover, the measured results for the average runtime per time step correspond approximately to the theoretical expectations.

Figure 5 The average runtime per time step and the total runtime with respect to order of the Active Flux and Discontinuous Galerkin methods.



In Figure 6 one can see the runtime with respect to the number of exclusive degrees of freedom per cell, which corresponds to the memory usage. The average runtime per time step of the Active Flux methods up to order five is faster than the average runtime per time step of the Discontinuous Galerkin methods. For order six and higher, the average runtime per time step of the Active Flux methods is at least as fast as the average runtime per time step of the Discontinuous Galerkin methods with the same time discretisation. The total runtime of the different Active Flux methods is faster than the total runtime of the Discontinuous Galerkin methods. The reason for this is again the larger CFL number of the Active Flux methods. In addition, for a given number of degrees of freedom per cell, the Active Flux methods compute more accurate results, as one can see in Figure 7. Especially the error for the Active Flux methods with a fourth-order time discretisation is significantly lower than the error for the Discontinuous Galerkin methods with a similar number of exclusive degrees of freedom per cell. In contrast, for a given order, the Discontinuous Galerkin methods are more accurate than the Active Flux methods.

Figure 6 The average runtime per time step and the total runtime with respect to number of exclusive degrees of freedom of the Active Flux and Discontinuous Galerkin methods

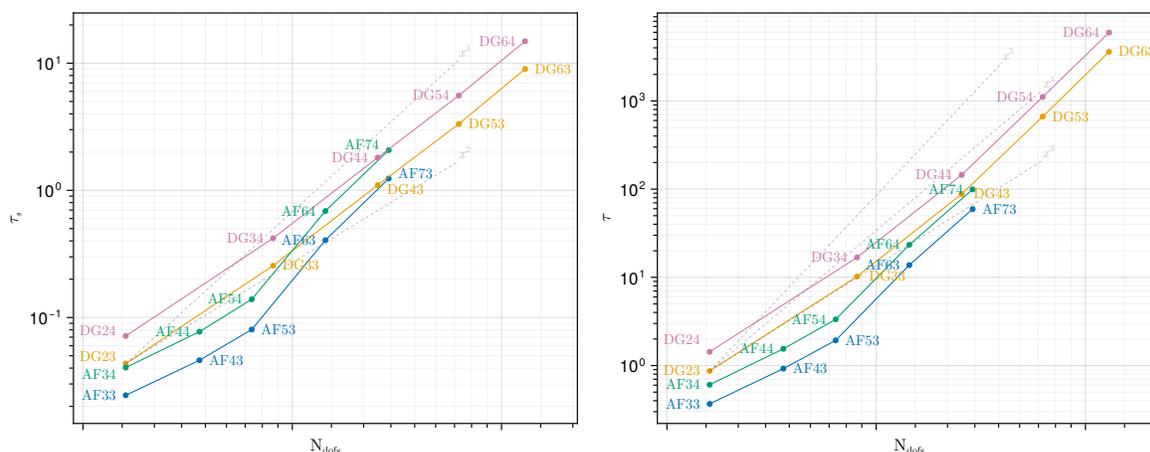
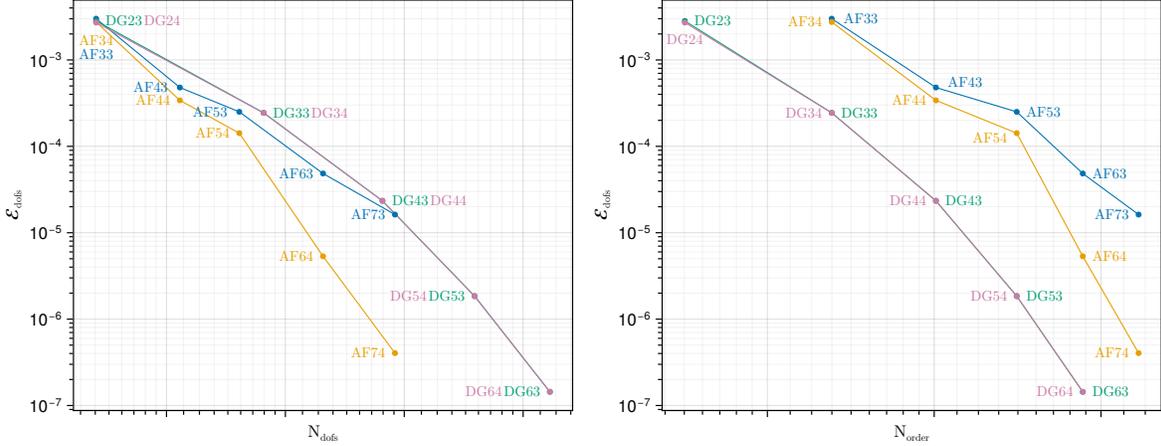


Figure 7 The l^2 -error with respect to the the number of degrees of freedom and the order of the different Active Flux and Discontinuous Galerkin methods.



One way to find the most efficient method with respect to memory usage, accuracy and runtime out of all the methods considered here is to find the one minimising

$$N_{\text{dof}} \cdot \mathcal{E}_{\text{dof}} \cdot \tau.$$

In Table 4 this value for the different methods at different grid sizes is displayed, and the smallest value is highlighted. If one measures efficiency in this way, the Active Flux methods are superior to the Discontinuous Galerkin methods.

An overview of the measured results for the different methods is given in Table 5, where the number of degrees of freedom, total runtime and l^2 -error for different grids are stated with respect to the results for AF54.

5 Conclusions

In this work we have shown that in quite some generality, Active Flux and DG can be considered to be essentially the same method—with the crucial difference, though, that AF is of order $K + 2$, and DG only of order $K + 1$, unless the error is measured on certain points or some post-processing applied. One can say that Active Flux is the high-order method “behind” DG, that becomes visible e.g. in the Radau points. While a choice of bases (modal/nodal/etc.) in DG trivially leads to equivalent methods, the equivalence between AF and DG is more complicated to see since the spaces employed are different.

While the methods might be the same, different choices of degrees of freedom can entail very different computational cost. We have performed experiments in order to measure runtime, accuracy and memory usage in a unified implementation of both AF and DG, ensuring as much as possible their comparability. In our experiments we find Active Flux to be more efficient than DG. Especially the fifth-order Active Flux method with a fourth-order time discretisation appears to be the most accurate and cost-efficient method in the regime considered.

A study of the relationship between AF and DG for multi-dimensional nonlinear systems is going to be subject of future work.

| | 20×20 | 40×40 | 80×80 | 160×160 |
|-------------|----------------|----------------|----------------|------------------|
| DG23 | 0.0053 | 0.0098 | 0.0142 | 0.029 |
| DG24 | 0.0087 | 0.0155 | 0.0235 | 0.0487 |
| AF33 | 0.003 | 0.0044 | 0.0049 | 0.0052 |
| AF34 | 0.0049 | 0.0067 | 0.0079 | 0.0083 |
| DG33 | 0.0248 | 0.0225 | 0.0221 | 0.0199 |
| DG34 | 0.0382 | 0.0369 | 0.0361 | 0.0332 |
| AF43 | 0.004 | 0.0027 | 0.0017 | 0.0013 |
| AF44 | 0.0056 | 0.0032 | 0.0015 | 0.0007 |
| DG43 | 0.0566 | 0.0329 | 0.0146 | 0.0071 |
| DG44 | 0.0924 | 0.0543 | 0.0243 | 0.0115 |
| AF53 | 0.0074 | 0.0039 | 0.0024 | 0.002 |
| AF54 | 0.0111 | 0.0038 | 0.0011 | 0.0003 |
| DG53 | 0.1924 | 0.0307 | - | - |
| DG54 | 0.3185 | 0.0514 | - | - |
| AF63 | 0.023 | 0.008 | - | - |
| AF64 | 0.0281 | 0.0015 | - | - |
| DG63 | 0.0905 | 0.0187 | - | - |
| DG64 | 0.0216 | 0.0164 | - | - |
| AF73 | 0.1284 | 0.0965 | - | - |
| AF74 | 0.0077 | 0.0007 | - | - |

Table 4: The value $N_{\text{dof}} \cdot \mathcal{E}_{\text{dof}} \cdot \tau$. for the different methods and different grid sizes.

Acknowledgement

We acknowledge inspiring and productive discussions with Lisa Lechner and Junming Duan.

References

- [AB23] Remi Abgrall and Wasilij Barsukow. Extensions of Active Flux to arbitrary order of accuracy. ESAIM: Mathematical Modelling and Numerical Analysis, 57(2):991–1027, 2023.
- [Abg22] Rémi Abgrall. A combination of Residual Distribution and the Active Flux formulations or a new class of schemes that can combine several writings of the same hyperbolic problem: application to the 1d Euler equations. Communications on Applied Mathematics and Computation, pages 1–33, 2022.
- [ABK25] Rémi Abgrall, Wasilij Barsukow, and Christian Klingenberg. A semi-discrete Active Flux method for the Euler equations on Cartesian grids. Journal of Scientific Computing, 102(2):36, 2025.
- [ABL25] Rémi Abgrall, Walter Boscheri, and Yongle Liu. Virtual finite element and hyperbolic problems: The pampa algorithm. Journal of Computational Physics, page 114521, 2025.
- [ADFK02] Slimane Adjerid, Karen D Devine, Joseph E Flaherty, and Lilia Krivodonova. A posteriori error estimation for discontinuous galerkin solutions of hyperbolic prob-

| | | 20 × 20 | | 40 × 40 | | 80 × 80 | | 160 × 160 | |
|------|-------------------|----------|-----------------------------|----------|-----------------------------|---------|-----------------------------|-----------|-----------------------------|
| | N_{dofs} | τ | $\mathcal{E}_{\text{dofs}}$ | τ | $\mathcal{E}_{\text{dofs}}$ | τ | $\mathcal{E}_{\text{dofs}}$ | τ | $\mathcal{E}_{\text{dofs}}$ |
| DG23 | 0.5 | 0.251 | 3.8065 | 0.2604 | 19.8958 | 0.2638 | 100.3551 | 0.2683 | 810.5553 |
| DG24 | 0.5 | 0.4288 | 3.6782 | 0.4291 | 19.0528 | 0.4376 | 100.353 | 0.451 | 810.6546 |
| AF33 | 0.5 | 0.1101 | 4.9001 | 0.1101 | 21.1132 | 0.1059 | 87.2066 | 0.1105 | 353.5169 |
| AF34 | 0.5 | 0.1935 | 4.5234 | 0.1818 | 19.3376 | 0.1844 | 80.2651 | 0.1913 | 326.1563 |
| DG33 | 1.125 | 2.9412 | 0.6746 | 3.0652 | 1.7192 | 3.0495 | 6.0202 | 2.881 | 23.0699 |
| DG34 | 1.125 | 4.8306 | 0.6339 | 5.0378 | 1.7152 | 4.9861 | 6.0201 | 4.7938 | 23.0698 |
| AF43 | 0.75 | 0.2663 | 1.8042 | 0.2764 | 3.3763 | 0.2735 | 7.6629 | 0.289 | 22.8259 |
| AF44 | 0.75 | 0.4506 | 1.5042 | 0.4634 | 2.3965 | 0.4778 | 4.0401 | 0.4915 | 7.5368 |
| DG43 | 2.0 | 26.2849 | 0.097 | 26.3351 | 0.1647 | 24.7744 | 0.276 | 24.4461 | 0.5448 |
| DG44 | 2.0 | 43.2942 | 0.0962 | 43.4175 | 0.1647 | 41.1555 | 0.2757 | 39.4917 | 0.5449 |
| AF53 | 1.0 | 0.5596 | 1.1959 | 0.5783 | 1.7651 | 0.5776 | 3.9481 | 0.5874 | 12.7529 |
| AF54 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 | 1.0 |
| DG53 | 3.125 | 207.3767 | 0.0268 | 199.2672 | 0.013 | - | - | - | - |
| DG54 | 3.125 | 343.8791 | 0.0267 | 333.4801 | 0.013 | - | - | - | - |
| AF63 | 1.5 | 3.9886 | 0.3464 | 4.1296 | 0.3412 | - | - | - | - |
| AF64 | 1.5 | 6.59 | 0.2562 | 7.0084 | 0.0376 | - | - | - | - |
| DG63 | 4.5 | 1084.433 | 0.0017 | 1079.589 | 0.001 | - | - | - | - |
| DG64 | 4.5 | 1809.13 | 0.0017 | 1784.803 | 0.001 | - | - | - | - |
| AF73 | 2.125 | 17.4682 | 0.0525 | 17.8019 | 0.1142 | - | - | - | - |
| AF74 | 2.125 | 28.7285 | 0.0114 | 29.7961 | 0.0028 | - | - | - | - |

Table 5: Overview of the results for the different methods with respect to the results for the AF54 method.

lems. *Computer methods in applied mechanics and engineering*, 191(11-12):1097–1112, 2002.

- [AL25] Rémi Abgrall and Yongle Liu. Robust pampa scheme in the dg formulation on unstructured triangular meshes: bound preservation, oscillation elimination, and boundary conditions. [arXiv preprint arXiv:2511.16180](#), 2025.
- [Bar21] Wasilij Barsukow. The active flux scheme for nonlinear problems. *Journal of Scientific Computing*, 86(1):1–34, 2021.
- [Bar25a] Wasilij Barsukow. An active flux method for the euler equations based on the exact acoustic evolution operator. [arXiv preprint arXiv:2506.03291](#), 2025.
- [Bar25b] Wasilij Barsukow. Semi-discrete active flux as a petrov-galerkin method. [arXiv preprint arXiv:2508.15017](#), 2025.
- [BCKL25] Wasilij Barsukow, Praveen Chandrashekar, Christian Klingenberg, and Lisa Lechner. A generalized Active Flux method of arbitrarily high order in two dimensions. [arXiv preprint arXiv:2502.05101](#), 2025.
- [BDF94] Rupak Biswas, Karen D Devine, and Joseph E Flaherty. Parallel, adaptive finite element methods for conservation laws. *Applied Numerical Mathematics*, 14(1-3):255–283, 1994.

- [BdVBC⁺13] L Beirão da Veiga, Franco Brezzi, Andrea Cangiani, Gianmarco Manzini, L Donatella Marini, and Alessandro Russo. Basic principles of virtual element methods. Mathematical Models and Methods in Applied Sciences, 23(01):199–214, 2013.
- [BH82] Alexander N Brooks and Thomas JR Hughes. Streamline upwind/Petrov-Galerkin formulations for convection dominated flows with particular emphasis on the incompressible Navier-Stokes equations. Computer methods in applied mechanics and engineering, 32(1-3):199–259, 1982.
- [BHKR19] Wasilij Barsukow, Jonathan Hohm, Christian Klingenberg, and Philip L Roe. The active flux scheme on Cartesian grids and its low Mach number limit. Journal of Scientific Computing, 81(1):594–622, 2019.
- [BKL⁺25] Wasilij Barsukow, Christian Klingenberg, Lisa Lechner, Jan Nordström, Sigrun Ortleb, and Hendrik Ranocha. Proving stability of the active flux method in the framework of summation-by-parts operators. submitted, 2025.
- [CHLM24] Erik Chudzik, Christiane Helzel, and Mária Lukáčová-Medvid’ová. Active Flux methods for hyperbolic systems using the method of bicharacteristics. Journal of Scientific Computing, 99(1):16, 2024.
- [CLS89] Bernardo Cockburn, San-Yih Lin, and Chi-Wang Shu. TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems. Journal of computational Physics, 84(1):90–113, 1989.
- [CLSS03] Bernardo Cockburn, Mitchell Luskin, Chi-Wang Shu, and Endre Süli. Enhanced accuracy by post-processing for finite element methods for hyperbolic equations. Mathematics of Computation, 72(242):577–606, 2003.
- [CLYZ17] Waixiang Cao, Dongfang Li, Yang Yang, and Zhimin Zhang. Superconvergence of discontinuous galerkin methods based on upwind-biased fluxes for 1d linear hyperbolic equations. ESAIM: Mathematical Modelling and Numerical Analysis, 51(2):467–486, 2017.
- [Coc25] Bernardo Cockburn. The discretizations of the derivative by the continuous galerkin and the discontinuous galerkin methods are exactly the same. arXiv preprint arXiv:2509.22587, 2025.
- [CS08] Yingda Cheng and Chi-Wang Shu. Superconvergence and time evolution of discontinuous galerkin finite element solutions. Journal of Computational Physics, 227(22):9612–9627, 2008.
- [CS10] Yingda Cheng and Chi-Wang Shu. Superconvergence of discontinuous galerkin and local discontinuous galerkin schemes for linear hyperbolic and convection-diffusion equations in one space dimension. SIAM Journal on Numerical Analysis, 47(6):4044–4072, 2010.
- [CSYZ15] Waixiang Cao, Chi-Wang Shu, Yang Yang, and Zhimin Zhang. Superconvergence of discontinuous galerkin methods for two-dimensional hyperbolic equations. SIAM Journal on Numerical Analysis, 53(4):1651–1671, 2015.

- [CSYZ18] Waixiang Cao, Chi-Wang Shu, Yang Yang, and Zhimin Zhang. Superconvergence of discontinuous galerkin method for scalar nonlinear hyperbolic equations. SIAM Journal on Numerical Analysis, 56(2):732–765, 2018.
- [CSZ17] Waixiang Cao, Chi-Wang Shu, and Zhimin Zhang. Superconvergence of discontinuous galerkin methods for 1-d linear hyperbolic equations with degenerate variable coefficients. ESAIM: Mathematical Modelling and Numerical Analysis, 51(6):2213–2235, 2017.
- [CSZ25] Waixiang Cao, Chi-Wang Shu, and Zhimin Zhang. Unified analysis of the semidiscrete discontinuous galerkin methods for 2-d hyperbolic equations on cartesian meshes using Pk elements: optimal error estimates and superconvergence. Mathematics of Computation, 2025.
- [CZZ14] Waixiang Cao, Zhimin Zhang, and Qingsong Zou. Superconvergence of discontinuous galerkin methods for linear hyperbolic equations. SIAM Journal on Numerical Analysis, 52(5):2555–2573, 2014.
- [DBK25] Junming Duan, Wasilij Barsukow, and Christian Klingenberg. Active Flux methods for hyperbolic conservation laws—Flux Vector splitting and bound-preservation. SIAM Journal on Scientific Computing, 47(2):A811–A837, 2025.
- [ER13] Timothy A Eymann and Philip L Roe. Multidimensional active flux schemes. In 21st AIAA computational fluid dynamics conference, 2013.
- [FR15] Doreen Fan and Philip L Roe. Investigations of a new scheme for wave propagation. In 22nd AIAA Computational Fluid Dynamics Conference, page 2449, 2015.
- [FR20] Daniel J Frean and Jennifer K Ryan. Superconvergence and the numerical flux: a study using the upwind-biased flux in discontinuous galerkin methods. Communications on Applied Mathematics and Computation, 2(3):461–486, 2020.
- [GHD⁺25] G Grünwald, L Hensel, M Deisenhofer, S Lautenbach, K Kormann, and R Grauer. Solving the six-dimensional vlasov-maxwell system with active flux and splitting methods. arXiv preprint arXiv:2511.22440, 2025.
- [GZQ13] Wei Guo, Xinghui Zhong, and Jing-Mei Qiu. Superconvergence of discontinuous Galerkin and local discontinuous Galerkin methods: eigen-structure analysis based on Fourier approach. Journal of Computational Physics, 235:458–485, 2013.
- [HKS19] Christiane Helzel, David Kerkmann, and Leonardo Scandurra. A new ADER method inspired by the active flux method. Journal of Scientific Computing, 80(3):1463–1497, 2019.
- [KF03] Lilia Krivodonova and Joseph E Flaherty. Error estimation for discontinuous galerkin solutions of two-dimensional hyperbolic problems. Advances in Computational Mathematics, 19(1):57–71, 2003.
- [LR74] Pierre Lesaint and Pierre-Arnaud Raviart. On a finite element method for solving the neutron transport equation. Publications mathématiques et informatique de Rennes, (S4):1–40, 1974.

- [RMF18] Philip L Roe, Jungyeoul Maeng, and Doreen Fan. Comparing active flux and Discontinuous Galerkin methods for compressible flow. In 2018 AIAA Aerospace Sciences Meeting, page 0836, 2018.
- [Roe17] Philip Roe. A simple explanation of superconvergence for discontinuous galerkin solutions to $ut+ ux= 0$. Communications in Computational Physics, 21(4):905–912, 2017.
- [RSA05] Jennifer Ryan, Chi-Wang Shu, and Harold Atkins. Extension of a post processing technique for the discontinuous galerkin method for hyperbolic equations with application to an aeroacoustic problem. SIAM Journal on Scientific Computing, 26(3):821–843, 2005.
- [vL77] Bram van Leer. Towards the ultimate conservative difference scheme. IV. A new approach to numerical convection. Journal of computational physics, 23(3):276–299, 1977.
- [XZ22] Yuan Xu and Qiang Zhang. Superconvergence analysis of the runge-kutta discontinuous galerkin method with upwind-biased numerical flux for two-dimensional linear hyperbolic equation. Communications on Applied Mathematics and Computation, 4(1):319–352, 2022.
- [YS12] Yang Yang and Chi-Wang Shu. Analysis of optimal superconvergence of discontinuous galerkin method for linear hyperbolic equations. SIAM Journal on Numerical Analysis, 50(6):3110–3133, 2012.
- [ZS04] Qiang Zhang and Chi-Wang Shu. Error estimates to smooth solutions of runge-kutta discontinuous galerkin methods for scalar conservation laws. SIAM Journal on Numerical Analysis, 42(2):641–666, 2004.

A Proof of Lemma [3.15](#)

Proof. The starting point is [\(68\)](#) with $w = 1$. The defining property of v^\pm is that

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} v^\pm(x) q_{ij}(x, y) dx = q_{ij} \left(\pm \frac{\Delta x}{2}, y \right) \quad (93)$$

These test functions exist by Riesz’ representation theorem, and they project out the corresponding values of the derivatives, because these are in the same DG space (see Remark [3.12](#)). Summing up [\(68\)](#) once with αv^+ , and once with βv^- , shifted to $i + 1$, yields

$$\begin{aligned} & \frac{d}{dt} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \left(\alpha q_{ij} \left(\frac{\Delta x}{2}, y \right) + \beta q_{i+1,j} \left(-\frac{\Delta x}{2}, y \right) \right) dy \\ & + \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} U^x \left(\alpha \partial_x (q_{ij} + r_{ij}^x) \Big|_{x=\frac{\Delta x}{2}} + \beta \partial_x (q_{i+1,j} + r_{i+1,j}^x) \Big|_{x=-\frac{\Delta x}{2}} \right) dy \\ & + \frac{1}{\Delta y} U^y \left(\alpha (q_{ij} + r_{ij}^y) \Big|_{x=\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}}^{y=\frac{\Delta y}{2}} + \beta (q_{i+1,j} + r_{i+1,j}^y) \Big|_{x=-\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}}^{y=\frac{\Delta y}{2}} \right) = 0 \end{aligned} \quad (94)$$

Finally, using (63) one finds that

$$\left(q_{ij} + r_{ij}^y \right) \Big|_{x=\pm\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}}^{y=\frac{\Delta y}{2}} = q_{ij} \left(\pm\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) + \hat{q}_{i,j+\frac{1}{2}} \left(\pm\frac{\Delta x}{2} \right) - q_{ij} \left(\pm\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) \quad (95)$$

$$\begin{aligned} & - q_{ij} \left(\pm\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) - \hat{q}_{i,j-\frac{1}{2}} \left(\pm\frac{\Delta x}{2} \right) + q_{ij} \left(\pm\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) \\ & = \hat{q}_{i,j+\frac{1}{2}} \left(\pm\frac{\Delta x}{2} \right) - \hat{q}_{i,j-\frac{1}{2}} \left(\pm\frac{\Delta x}{2} \right) \end{aligned} \quad (96)$$

Similarly, using $v = 1$ and w^\pm one finds the analogous expression

$$\begin{aligned} & \frac{d}{dt} \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \left(\alpha q_{ij} \left(x, \frac{\Delta y}{2} \right) + \beta q_{i,j+1} \left(x, -\frac{\Delta y}{2} \right) \right) dx \\ & + \frac{1}{\Delta x} U^x \left(\alpha (q_{ij} + r_{ij}^x) \Big|_{y=\frac{\Delta y}{2}, x=-\frac{\Delta x}{2}}^{x=\frac{\Delta x}{2}} + \beta (q_{i,j+1} + r_{i,j+1}^x) \Big|_{y=-\frac{\Delta y}{2}, x=-\frac{\Delta x}{2}}^{x=\frac{\Delta x}{2}} \right) \\ & + \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} U^y \left(\alpha \partial_y (q_{ij} + r_{ij}^y) \Big|_{y=\frac{\Delta y}{2}} + \beta \partial_y (q_{i,j+1} + r_{i,j+1}^y) \Big|_{y=-\frac{\Delta y}{2}} \right) dx = 0 \end{aligned} \quad (97)$$

which is simplified by realizing that

$$(q_{ij} + r_{ij}^x) \Big|_{y=\pm\frac{\Delta y}{2}, x=-\frac{\Delta x}{2}}^{x=\frac{\Delta x}{2}} = \hat{q}_{i+\frac{1}{2},j} \left(\pm\frac{\Delta y}{2} \right) - \hat{q}_{i-\frac{1}{2},j} \left(\pm\frac{\Delta y}{2} \right) \quad (98)$$

For the third expression, one combines the two approaches and obtains immediately

$$\begin{aligned} & \frac{d}{dt} \left(\alpha q_{ij} \left(\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) + \beta q_{i+1,j} \left(-\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) + \gamma q_{i,j+1} \left(\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) + \delta q_{i+1,j+1} \left(-\frac{\Delta x}{2}, -\frac{\Delta y}{2} \right) \right) \\ & + U^x \left(\alpha \partial_x (q_{ij} + r_{ij}^x) \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + \beta \partial_x (q_{i+1,j} + r_{i+1,j}^x) \Big|_{x=-\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} \right. \\ & \quad \left. + \gamma \partial_x (q_{i,j+1} + r_{i,j+1}^x) \Big|_{x=\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} + \delta \partial_x (q_{i+1,j+1} + r_{i+1,j+1}^x) \Big|_{x=-\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} \right) \\ & + U^y \left(\alpha \partial_y (q_{ij} + r_{ij}^y) \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + \beta \partial_y (q_{i+1,j} + r_{i+1,j}^y) \Big|_{x=-\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} \right. \\ & \quad \left. + \gamma \partial_y (q_{i,j+1} + r_{i,j+1}^y) \Big|_{x=\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} + \delta \partial_y (q_{i+1,j+1} + r_{i+1,j+1}^y) \Big|_{x=-\frac{\Delta x}{2}, y=-\frac{\Delta y}{2}} \right) = 0 \end{aligned} \quad (99)$$

□

B Proof of Lemma 3.13

Proof. By the fact that Radau polynomials have zero average it is clear that

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} Q_{ij}(x, y) dy dx = \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} q_{ij}(x, y) dy dx \quad (100)$$

and also

$$\frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} Q_{ij} \left(\frac{\Delta x}{2}, y \right) dy = \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} q_{ij} \left(\frac{\Delta x}{2}, y \right) + r_{ij}^x \left(\frac{\Delta x}{2}, y \right) dy \quad (101)$$

$$= \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} q_{ij} \left(\frac{\Delta x}{2}, y \right) + \left(\hat{q}_{i+\frac{1}{2},j} - q_{ij} \left(\frac{\Delta x}{2}, y \right) \right) R_{R,\Delta x}^{K+1} \left(\frac{\Delta x}{2} \right) dy \quad (102)$$

$$= \frac{1}{\Delta y} \int_{-\frac{\Delta y}{2}}^{\frac{\Delta y}{2}} \hat{q}_{i+\frac{1}{2},j} dy \quad (103)$$

The other edge-averages are treated analogously. Finally,

$$Q_{ij} \left(\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) = q_{ij} \left(\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) + r_{ij}^x \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + r_{ij}^y \Big|_{x=\frac{\Delta x}{2}, y=\frac{\Delta y}{2}} + C_{ij}^{\text{RR}} \stackrel{!}{=} Q_{i+\frac{1}{2},j+\frac{1}{2}} \quad (104)$$

which can easily be solved for C_{ij}^{RR} :

$$C_{ij}^{\text{RR}} = Q_{i+\frac{1}{2},j+\frac{1}{2}} - \hat{q}_{i+\frac{1}{2},j} \left(\frac{\Delta y}{2} \right) - \hat{q}_{i,j+\frac{1}{2}} \left(\frac{\Delta x}{2} \right) + q_{ij} \left(\frac{\Delta x}{2}, \frac{\Delta y}{2} \right) \quad (105)$$

$$(106)$$

Analogously, the other three nodal values allow to fix the remaining constants $C_{ij}^{\text{LL}}, C_{ij}^{\text{LR}}, C_{ij}^{\text{RL}}$. \square

C Proof of Lemma [3.16](#)

Proof. The expressions in question are second-degree polynomials. It is therefore sufficient to verify that they agree in average and in value at the endpoints.

Consider first the average of the left-hand side of [\(81\)](#)

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \beta^+ (q_{ij} + r_{ij}^x) \Big|_{y=\frac{\Delta y}{2}} + \beta^- (q_{i,j+1} + r_{i,j+1}^x) \Big|_{y=-\frac{\Delta y}{2}} dx = \frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \hat{q}_{i,j+\frac{1}{2}} dx \quad (107)$$

while the average of the right-hand side is, by continuity of the AF approximation

$$\frac{1}{\Delta x} \int_{-\frac{\Delta x}{2}}^{\frac{\Delta x}{2}} \beta^+ Q_{ij} \Big|_{y=\frac{\Delta y}{2}} + \beta^- Q_{i,j+1} \Big|_{y=-\frac{\Delta y}{2}} dx = (\beta^+ + \beta^-) \bar{Q}_{i,j+\frac{1}{2}} \quad (108)$$

which, with $\beta^+ + \beta^- = 1$ and [\(73\)](#) is indeed the same.

Observe now that the evaluation at $x = \pm \frac{\Delta x}{2}$ allows to simplify the expressions, using [\(62\)](#)–[\(63\)](#) as follows:

$$\left[\beta^+ (q_{ij} + r_{ij}^x) \Big|_{y=\frac{\Delta y}{2}} + \beta^- (q_{i,j+1} + r_{i,j+1}^x) \Big|_{y=-\frac{\Delta y}{2}} \right] \Big|_{x=\pm \frac{\Delta x}{2}} = \beta^+ \hat{q}_{i\pm\frac{1}{2},j} + \beta^- \hat{q}_{i\pm\frac{1}{2},j+1} = Q_{i\pm\frac{1}{2},j+\frac{1}{2}} \quad (109)$$

$$\left[\alpha^+ (q_{ij} + r_{ij}^y) \Big|_{x=\frac{\Delta x}{2}} + \alpha^- (q_{i+1,j} + r_{i+1,j}^y) \Big|_{x=-\frac{\Delta x}{2}} \right] \Big|_{y=\pm \frac{\Delta y}{2}} = \alpha^+ \hat{q}_{i,j\pm\frac{1}{2}} + \alpha^- \hat{q}_{i+1,j\pm\frac{1}{2}} = Q_{i+\frac{1}{2},j\pm\frac{1}{2}} \quad (110)$$

where the last equalities are [\(75\)](#). The rest of the proof follows analogously. \square