A unifying algebraic framework for discontinuous Galerkin and flux reconstruction methods based on the summation-by-parts property

Tristan Montoya · David W. Zingg

Abstract We propose a unifying framework for the matrix-based formulation and analysis of discontinuous Galerkin (DG) and flux reconstruction (FR) methods for conservation laws on general unstructured grids. Within such an algebraic framework, the multidimensional summation-by-parts (SBP) property is used to establish the discrete equivalence of strong and weak formulations, as well as the conservation and energy stability properties of a broad class of DG and FR schemes. Specifically, the analysis enables the extension of the equivalence between the strong and weak forms of the discontinuous Galerkin collocation spectral-element method demonstrated by Kopriva and Gassner (J Sci Comput 44:136–155, 2010) to more general nodal and modal DG formulations, as well as to the Vincent-Castonguay-Jameson-Huynh (VCJH) family of FR methods. Moreover, new algebraic proofs of conservation and energy stability for DG and VCJH schemes are presented, in which the SBP property serves as a unifying mechanism for establishing such results. Numerical experiments are provided for the two-dimensional linear advection and Euler equations, highlighting the design choices afforded for methods within the proposed framework and corroborating the theoretical analysis.

Keywords high-order methods · discontinuous Galerkin · flux reconstruction · summation-by-parts · conservation laws · energy stability

Mathematics Subject Classification (2010) 65M12 · 65M60 · 65M70

1 Introduction

Hyperbolic and convection-dominated systems of conservation laws are of considerable importance in the mathematical modelling of physical phenomena, with numerous

T. Montoya
University of Toronto Institute for Aerospace Studies, Toronto, Canada
E-mail: tristan.montoya@mail.utoronto.ca
ORCiD: 0000-0002-4259-1449

D. W. Zingg
University of Toronto Institute for Aerospace Studies, Toronto, Canada
E-mail: dwz@utias.utoronto.ca

scientific and engineering disciplines relying heavily on efficient and robust numerical methods for the solution of such partial differential equations (PDEs). High-order methods (i.e. those of third order or higher) have shown significant promise for problems involving the propagation of waves over long distances (an early example being the work of Kreiss and Oliger [1]) and the fine-scale resolution of turbulent flow structures (as surveyed by Wang *et al.* [2]), both of which incur prohibitive computational expense when conventional second-order spatial discretizations are used. The most popular high-order methods in current use by practitioners for such problems are arguably the well-established discontinuous Galerkin (DG) methods and the more recent flux reconstruction (FR) schemes. DG methods were originally introduced for the steady neutron transport equation by Reed and Hill [3], and have since been applied successfully to a wide range of problems following extensive development by Cockburn, Shu, and collaborators (see, for example, [4–7]). The FR approach was first proposed by Huynh [8] for one-dimensional conservation laws, extended to triangular elements by Wang and Gao [9] through the so-called lifting collocation penalty (LCP) formulation, and further developed by several other authors, including Vincent *et al.* [10], Castonguay *et al.* [11], and Williams and Jameson [12], who identified energy-stable families of FR schemes on one-dimensional, triangular, and tetrahedral elements, respectively.

Although DG methods are derived from a weak (i.e. variational) formulation whereas FR methods are derived from a strong (i.e. differential) formulation, the two approaches are closely related, as discussed in Huynh's seminal paper [8] as well as in further investigations by Allaneau and Jameson [13], De Grazia *et al.* [14], Mengaldo *et al.* [15], and Zwanenburg and Nadarajah [16]. Moreover, certain DG and FR schemes have been recast in terms of summation-by-parts (SBP) operators, which are discrete differential operators equipped with compatible inner products such that integration by parts is mimicked algebraically (see, for example, the review papers of Del Rey Fernández *et al.* [17] and Svärd and Nordström [18]). Beginning with the work of Kreiss and Scherer [19], this mimetic property, referred to as the SBP property, has been instrumental in the development and analysis of high-order finite-difference methods with discrete stability and conservation properties, and there have been significant advances in recent years towards extending the SBP methodology to encompass a wider range of new and existing methods, facilitated by the one-dimensional and multidimensional generalizations of Del Rey Fernández *et al.* [20] and Hicken *et al.* [21], respectively.

Due to its reliance on discretization-agnostic matrix properties, the SBP approach allows for existing theoretical results established for particular numerical methods to be reinterpreted within a more general and arguably simpler setting, and facilitates the cross-pollination of techniques between different schemes sharing common algebraic properties. This cross-pollination has proved fruitful over the past decade, with Gassner [22] and Ranocha *et al.* [23] employing the SBP property in order to develop nonlinearly stable DG and FR methods, respectively, for Burgers' equation in one space dimension using skew-symmetric split forms originally introduced for finite-difference discretizations. Motivated by these contributions, as well as subsequent extensions to entropy-stable and kinetic-energy-preserving DG-type schemes exploiting the SBP property (see, for example, Gassner *et al.* [24] and Chan [25]), the goal of this work is to develop a comprehensive framework enabling the unification of existing algebraic techniques for the analysis of semi-discrete DG and FR methods applied to time-dependent conservation laws and the extension of the existing theory based on the

SBP property to more general DG and FR formulations. To this end, the analysis has resulted in the following novel contributions.

1. The discrete equivalence of strong and weak formulations, demonstrated by Kopriva and Gassner [26] in the context of DG methods employing collocated tensor-product Legendre-Gauss (LG) and Legendre-Gauss-Lobatto (LGL) quadrature rules (i.e. the LG and LGL variants discontinuous Galerkin collocation spectral-element method, which we refer to as the DGSEM-LG and DGSEM-LGL, respectively), is extended to more general quadrature-based and collocation-based DG formulations by way of the SBP property.
2. The Vincent-Castonguay-Jameson-Huynh (VCJH) family of FR schemes (as described in [10–12]), is recast in a generalized matrix form applicable to general element types, allowing for the use of modal bases and over-integration[1] of the volume and facet terms. Through the SBP property, these formulations are shown to be equivalent to filtered DG schemes in strong or weak form, in the latter case leading to methods which are algebraically equivalent to conventional strong-form FR methods while offering the potential for improved computational efficiency.
3. Proofs of discrete conservation with respect to suitable quadrature rules are presented for DG and VCJH methods in strong and weak form, differing from the existing conservation proofs for the FR approach due to their reliance on the discrete divergence theorem – which holds as a consequence of the SBP property – rather than the continuity of the corrected flux, which is not explicitly reconstructed for multidimensional FR schemes on non-tensor-product elements.
4. Energy stability (with respect to suitable discrete norms) is established as a consequence of the SBP property for DG and FR schemes applied to constant-coefficient linear advection problems on affine polytopal meshes with periodic or inflow-outflow boundary conditions, recovering the existing stability results for VCJH schemes (i.e. those in [10–12]) as special cases of a more general theory.

We now outline the content of the remainder of this paper. In §2, we describe the systems of conservation laws to which we seek numerical solutions, as well as the meshes and approximation spaces employed for the discretizations considered in this work; we then provide a brief overview of the DG and FR approaches. In §3, we develop a unifying matrix formulation for a broad class of DG and FR methods, which we employ in §4 for the analysis of such schemes based on the SBP property. In §5, we present numerical experiments for the two-dimensional linear advection and Euler equations in support of the theory. Concluding remarks are provided in §6.

2 Preliminaries

2.1 Notation

Throughout this paper, any symbol bearing a single underline denotes a vector (treated as a column matrix), whereas any symbol bearing a double underline denotes a matrix. The symbols $\mathbb{R}$, $\mathbb{R}^+$ $\mathbb{R}_0^+$, $\mathbb{N}$ and $\mathbb{N}_0$ denote the real numbers, the positive

[1] In this paper, the term *over-integration* (also referred to as polynomial de-aliasing, and discussed further, for example, by Kirby and Karniadakis [27] and Mengaldo *et al.* [28]) refers to the approximation of a volume or facet integral on a nodal set of cardinality greater than the dimension of the space to which the numerical solution on such an element or facet belongs.

real numbers, the non-negative real numbers, the natural numbers (excluding zero), and the natural numbers including zero, while $\underline{0}$, $\underline{1}$, $\underline{\underline{0}}$, and $\underline{\underline{I}}$ are reserved for vectors of zeros, vectors of ones, matrices of zeros, and the identity matrix, respectively, with their dimensions inferred from the context. Furthermore, the operator $\mathrm{diag}(\cdot)$ is used to denote the assembly of a diagonal matrix from a list of scalars or the assembly of a block-diagonal matrix from a list of matrices, and δ_{ij} denotes the Kronecker delta.

Given a bounded domain $\Omega \subset \mathbb{R}^d$ with $d \in \mathbb{N}$, we use the symbol $\partial\Omega$ to denote its boundary, $\bar{\Omega} := \Omega \cup \partial\Omega$ to denote its closure, and $|\Omega|$ to denote its Lebesgue measure (i.e. d-dimensional volume); the interior of a closed domain Ω is given by $\mathring{\Omega} := \Omega \setminus \partial\Omega$. The unit $(d-1)$-sphere is given by $\mathbb{S}^{d-1} := \{\underline{x} \in \mathbb{R}^d : \|\underline{x}\| = 1\}$, where the Euclidean dot product and norm are defined such that $\underline{x} \cdot \underline{y} := x_1 y_1 + \cdots + x_d y_d$ and $\|\underline{x}\|^2 := \underline{x} \cdot \underline{x}$, respectively. We denote the order of a multi-index $\underline{\alpha} \in \mathbb{N}_0^d$ as $|\underline{\alpha}| := \alpha_1 + \cdots + \alpha_d$ and define the corresponding differential operator such that $\partial^{(\underline{\alpha})} V(\underline{x}) := \partial^{|\underline{\alpha}|} V(\underline{x}) / \partial x_1^{\alpha_1} \cdots \partial x_d^{\alpha_d}$. The gradient of a scalar-valued function is given by $\underline{\nabla} V(\underline{x}) := [\partial V(\underline{x})/\partial x_1, \ldots, \partial V(\underline{x})/\partial x_d]^{\mathrm{T}}$, while the divergence and Jacobian of a vector-valued function are given by $\underline{\nabla} \cdot \underline{V}(\underline{x}) := \partial V_1(\underline{x})/\partial x_1 + \cdots + \partial V_d(\underline{x})/\partial x_d$ and $\underline{\nabla}\underline{V}(\underline{x}) := [\underline{\nabla} V_1(\underline{x}), \ldots, \underline{\nabla} V_d(\underline{x})]^{\mathrm{T}}$, respectively. All other nontrivial notational conventions are introduced as they appear.

2.2 Problem formulation

We consider first-order systems of time-dependent conservation laws governing the evolution of the $N_{\mathrm{eq}} \in \mathbb{N}$ solution variables $\underline{U}(\underline{x}, t) \in \Upsilon \subseteq \mathbb{R}^{N_{\mathrm{eq}}}$, where Υ denotes the set of admissible solution states and $\underline{x}$ and t denote the spatial and temporal coordinates, respectively. Such systems of PDEs take the general form

$$\begin{aligned} \frac{\partial \underline{U}(\underline{x}, t)}{\partial t} + \sum_{m=1}^{d} \frac{\partial \underline{F}^{(m)}(\underline{U}(\underline{x}, t))}{\partial x_m} &= \underline{0}, && \forall\, (\underline{x}, t) \in \Omega \times (0, T), \\ \underline{U}(\underline{x}, 0) &= \underline{U}^0(\underline{x}), && \forall\, \underline{x} \in \Omega, \end{aligned} \tag{2.1}$$

subject to appropriate boundary conditions, if necessary, where $\Omega \subset \mathbb{R}^d$ is a fixed domain of arbitrary dimension $d \in \mathbb{N}$ with a piecewise smooth boundary, $T \in \mathbb{R}^+$ is the final time, $\underline{F}^{(m)} : \Upsilon \to \mathbb{R}^{N_{\mathrm{eq}}}$ defines the flux component corresponding to the coordinate index $m \in \{1, \ldots, d\}$, and $\underline{U}^0 : \Omega \to \Upsilon$ defines the initial data at $t = 0$.

Remark 2.1 To simplify the exposition, we assume that the flux components have no direct dependence on space or time, and that no source terms are present; however, the extension of the present framework to such problems is straightforward.

While a very broad class of problems may be formulated as in (2.1), we highlight two representative examples, which serve as test cases for the numerical experiments in §5. We first consider the linear advection equation,

$$\frac{\partial U(\underline{x}, t)}{\partial t} + \sum_{m=1}^{d} \frac{\partial (a_m U(\underline{x}, t))}{\partial x_m} = 0, \tag{2.2}$$

which governs the transport of a scalar quantity $U(\underline{x}, t) \in \mathbb{R} =: \Upsilon$ at a velocity $\underline{a} \in \mathbb{R}^d$, assumed here to be constant. We are also interested in the Euler equations, which

constitute a nonlinear system governing the conservation of mass, momentum, and energy for a compressible, inviscid, adiabatic fluid, and are given by

$$\frac{\partial}{\partial t}\underbrace{\begin{bmatrix}\rho(\underline{x},t)\\ \rho(\underline{x},t)V_1(\underline{x},t)\\ \vdots\\ \rho(\underline{x},t)V_d(\underline{x},t)\\ E(\underline{x},t)\end{bmatrix}}_{=:\underline{U}(\underline{x},t)}+\sum_{m=1}^{d}\frac{\partial}{\partial x_m}\underbrace{\begin{bmatrix}\rho(\underline{x},t)V_m(\underline{x},t)\\ \rho(\underline{x},t)V_1(\underline{x},t)V_m(\underline{x},t)+P(\underline{x},t)\delta_{1m}\\ \vdots\\ \rho(\underline{x},t)V_d(\underline{x},t)V_m(\underline{x},t)+P(\underline{x},t)\delta_{dm}\\ V_m(\underline{x},t)(E(\underline{x},t)+P(\underline{x},t))\end{bmatrix}}_{=:\underline{F}^{(m)}(\underline{U}(\underline{x},t))}=\underline{0}. \quad (2.3)$$

In the above, $\rho(\underline{x},t)\in\mathbb{R}$ denotes the fluid density, $\underline{V}(\underline{x},t)\in\mathbb{R}^d$ denotes the flow velocity, $E(\underline{x},t)\in\mathbb{R}$ denotes the total energy per unit volume, and $P(\underline{x},t)\in\mathbb{R}$ denotes the pressure, which is related to the other variables through the equation of state given by $P(\underline{x},t)=(\gamma-1)(E(\underline{x},t)-\frac{1}{2}\rho(\underline{x},t)\|\underline{V}(\underline{x},t)\|^2)$ for an ideal gas with constant specific heat and specific heat ratio $\gamma>1$. The set of admissible solution states for (2.3) is therefore given by $\Upsilon:=\{\underline{U}(\underline{x},t)\in\mathbb{R}^{N_{\mathrm{eq}}+2}:P(\underline{x},t),\rho(\underline{x},t)>0\}$.

Remark 2.2 Throughout this paper, we tacitly assume that the values of numerical solutions to (2.1) remain within Υ. This is difficult to ensure *a priori* for nonlinear problems such as (2.3), and often requires the use of bespoke limiting procedures (see, for example, Zhang and Shu [29]), which are not considered in this work.

2.3 Mesh and coordinate transformation

We begin our description of the spatial discretization by introducing a mesh $\mathcal{T}_h:=\{\Omega^{(\kappa)}\}_{\kappa=1}^{N_{\mathrm{el}}}$ consisting of $N_{\mathrm{el}}\in\mathbb{N}$ compact, connected, elements of characteristic size $h\in\mathbb{R}^+$ with nonempty interiors, satisfying the following (standard) assumption.

Assumption 2.1 *The elements in $\mathcal{T}_h$ satisfy $\bigcup_{\kappa=1}^{N_{\mathrm{el}}}\Omega^{(\kappa)}=\bar{\Omega}$ and $\mathring{\Omega}^{(\kappa)}\cap\mathring{\Omega}^{(\nu)}=\emptyset$ for any $\kappa\neq\nu$. The boundary of each element is the union of $N_{\mathrm{fac}}\geq d+1$ smooth facets,[2] which are denoted by $\Gamma^{(\kappa,\zeta)}\subset\partial\Omega^{(\kappa)}$ and indexed as $\zeta\in\{1,\ldots,N_{\mathrm{fac}}\}$, each with an outward unit normal vector given by $\underline{n}^{(\kappa,\zeta)}:\Gamma^{(\kappa,\zeta)}\to\mathbb{S}^{d-1}$. Furthermore, each element is the image of a polytopal reference element $\hat{\Omega}\subset\mathbb{R}^d$ with flat facets $\hat{\Gamma}^{(\zeta)}\subset\partial\hat{\Omega}$ and constant outward unit normal vectors $\underline{\hat{n}}^{(\zeta)}\in\mathbb{S}^{d-1}$ under a smooth, time-invariant mapping $\underline{X}^{(\kappa)}:\hat{\Omega}\to\Omega^{(\kappa)}$ satisfying $J^{(\kappa)}(\underline{\hat{x}}):=\det(\underline{\nabla}\underline{X}^{(\kappa)}(\underline{\hat{x}}))>0$.*

We now recall from Vinokur [30] that systems in the form of (2.1) retain a conservative formulation in reference coordinates, which is given in the present context by

$$\begin{aligned}&\frac{\partial J^{(\kappa)}(\underline{\hat{x}})\underline{U}(\underline{X}^{(\kappa)}(\underline{\hat{x}}),t)}{\partial t}\\ &+\sum_{m=1}^{d}\frac{\partial}{\partial\hat{x}_m}\left(\sum_{n=1}^{d}J^{(\kappa)}(\underline{\hat{x}})\left[(\underline{\nabla}\underline{X}^{(\kappa)}(\underline{\hat{x}}))^{-1}\right]_{mn}\underline{F}^{(n)}(\underline{U}(\underline{X}^{(\kappa)}(\underline{\hat{x}}),t))\right)=\underline{0}.\end{aligned} \quad (2.4)$$

Defining $J^{(\kappa,\zeta)}(\underline{\hat{x}}):=\|J^{(\kappa)}(\underline{\hat{x}})(\underline{\nabla}\underline{X}^{(\kappa)}(\underline{\hat{x}}))^{-\mathrm{T}}\underline{\hat{n}}^{(\zeta)}\|$, we note that the outward unit normal vector to each facet transforms under the change of coordinates according to

[2] For notational convenience, it is assumed that all elements are of the same type and therefore have an equal number of facets, although this is not a limitation of the analysis.

Nanson's formula (see, for example, Gurtin *et al.* [31, §8.1] for a derivation in the context of continuum mechanics),

$$J^{(\kappa,\zeta)}(\underline{\hat{x}})\,\underline{n}^{(\kappa,\zeta)}(\underline{X}^{(\kappa)}(\underline{\hat{x}})) = J^{(\kappa)}(\underline{\hat{x}})(\underline{\nabla}\underline{X}^{(\kappa)}(\underline{\hat{x}}))^{-\mathrm{T}}\underline{\hat{n}}^{(\zeta)}. \tag{2.5}$$

Furthermore, since periodic boundary conditions may be treated identically to interior interfaces through mesh connectivity, we hereinafter use the symbol $\partial\Omega$ to refer only to non-periodic portions of the domain boundary. We may then define the set of element-facet index pairs $\mathcal{F} := \{1,\dots,N_{\mathrm{el}}\}\times\{1,\dots,N_{\mathrm{fac}}\}$, which may be partitioned into the disjoint subsets $\mathcal{F}_{\partial\Omega} := \{(\kappa,\zeta)\in\mathcal{F} : \Gamma^{(\kappa,\zeta)}\subset\partial\Omega\}$ and $\mathcal{F}_{\Omega} := \mathcal{F}\setminus\mathcal{F}_{\partial\Omega}$, corresponding to facets lying on the non-periodic portion of the boundary and those forming interior or periodic interfaces, respectively.

2.4 Polynomial approximation space

Using the symbol $\circ$ to denote function composition and considering the system in (2.4), we seek a semi-discrete approximation $\underline{U}^{(h,\kappa)}(\cdot,t)$ of $\underline{U}(\cdot,t)\circ\underline{X}^{(\kappa)}$ with components belonging to a finite-dimensional polynomial space with real coefficients, as given by

$$\mathbb{P}_{\mathcal{N}}(\hat{\Omega}) := \mathrm{span}\Big\{\hat{\Omega}\ni\underline{\hat{x}}\mapsto\hat{x}_1^{\alpha_1}\cdots\hat{x}_d^{\alpha_d} : \underline{\alpha}\in\mathcal{N}\Big\}, \tag{2.6}$$

where $\mathcal{N}\subset\mathbb{N}_0^d$ denotes a finite set of multi-indices with cardinality $N^*\in\mathbb{N}$. The space obtained by restricting functions in $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ to the facet $\hat{\Gamma}^{(\zeta)}\subset\partial\hat{\Omega}$ is denoted by $\mathbb{P}_{\mathcal{N}}(\hat{\Gamma}^{(\zeta)})$, which is of dimension $N_\zeta^*\in\mathbb{N}$. Such spaces afford the present framework with substantial generality, requiring only the following assumption on choice of $\mathcal{N}$.

Assumption 2.2 *The multi-index set $\mathcal{N}$ defining the space in (2.6) is a lower set (see, for example, Dyn and Floater [32, §3]) in the sense that for any $\underline{\beta}\in\mathcal{N}$, every multi-index $\underline{\alpha}$ satisfying $\alpha_m\le\beta_m$ for all $m\in\{1,\dots,d\}$ also belongs to $\mathcal{N}$.*

The globally discontinuous approximation $\underline{U}^h(\cdot,t)$ of the solution $\underline{U}(\cdot,t)$ to the system in (2.1) is then defined piecewise (up to a set of measure zero) as $\underline{U}^h(\underline{x},t) := \underline{U}^{(h,\kappa)}((\underline{X}^{(\kappa)})^{-1}(\underline{x}),t)$ for $\underline{x}\in\Omega^{(\kappa)}$, with each component thus belonging to the space

$$\mathbb{P}_{\mathcal{N}}(\mathcal{T}_h) := \Big\{V\in L^2(\Omega) : V\big|_{\Omega^{(\kappa)}}\circ\underline{X}^{(\kappa)}\in\mathbb{P}_{\mathcal{N}}(\hat{\Omega}),\ \forall\,\Omega^{(\kappa)}\in\mathcal{T}_h\Big\}, \tag{2.7}$$

where $L^2(\Omega)$ denotes the space of square-integrable measurable functions $V:\Omega\to\mathbb{R}$.

Remark 2.3 The choice of approximation space generally depends on the geometry of the reference element. For example, we recover the standard total-degree and tensor-product polynomial spaces $\mathbb{P}_p(\hat{\Omega})$ and $\mathbb{Q}_p(\hat{\Omega})$ for $p\in\mathbb{N}_0$ as special cases of (2.6) by taking $\mathcal{N} := \{\underline{\alpha}\in\mathbb{N}_0^d : |\underline{\alpha}|\le p\}$ for the former and $\mathcal{N} := \{\underline{\alpha}\in\mathbb{N}_0^d : \max_{m=1}^d\alpha_m\le p\}$ for the latter. These are the natural choices for constructing discontinuous approximation spaces in the form of (2.7) on meshes consisting of triangular/tetrahedral and quadrilateral/hexahedral elements, respectively.

Remark 2.4 Rather than constructing approximations of $\underline{U}(\underline{X}^{(\kappa)}(\underline{\hat{x}}),t)$ such that $\underline{U}^{(h,\kappa)}(\cdot,t)\in\mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_{\mathrm{eq}}}$, it is possible to instead approximate $J^{(\kappa)}(\underline{\hat{x}})\underline{U}(\underline{X}^{(\kappa)}(\underline{\hat{x}}),t)$ such that $J^{(\kappa)}\underline{U}^{(h,\kappa)}(\cdot,t)\in\mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_{\mathrm{eq}}}$. This modification of the approximation space

in (2.7) has no effect for affine elements (in which case $J^{(\kappa)}$ is constant for a given element), but simplifies the treatment of curvilinear coordinates for time-dependent problems. While the analysis in §4 extends in a straightforward manner to such modified formulations, we note that, as discussed by Yu *et al.* [33, §3], such schemes have been observed to suffer from degraded accuracy on curvilinear elements.

2.5 Discontinuous Galerkin and flux reconstruction methods

In this section, we briefly review the formulation of DG and FR methods for systems of conservation laws in curvilinear coordinates and introduce some relevant notation employed throughout our analysis of such schemes.

Discontinuous Galerkin method

Integrating (2.4) by parts over the reference element against a suitable test function and simplifying the resulting facet integrals using (2.5) leads to a local weak formulation of the system, which is given for all $\kappa \in \{1, \dots, N_{\text{el}}\}$ and $t \in (0, T)$ by

$$\begin{aligned}
&\int_{\hat{\Omega}} \Bigg(V(\underline{\hat{x}}) \frac{\partial J^{(\kappa)}(\underline{\hat{x}}) \underline{U}(\underline{X}^{(\kappa)}(\underline{\hat{x}}), t)}{\partial t} \\
&\qquad - \sum_{m=1}^{d} \frac{\partial V(\underline{\hat{x}})}{\partial \hat{x}_m} \sum_{n=1}^{d} J^{(\kappa)}(\underline{\hat{x}}) \left[(\nabla \underline{X}^{(\kappa)}(\underline{\hat{x}}))^{-1} \right]_{mn} \underline{F}^{(n)}(\underline{U}(\underline{X}^{(\kappa)}(\underline{\hat{x}}), t)) \Bigg) \, d\underline{\hat{x}} \\
&+ \sum_{\zeta=1}^{N_{\text{fac}}} \int_{\hat{\Gamma}^{(\zeta)}} V(\underline{\hat{x}}) J^{(\kappa,\gamma)}(\underline{\hat{x}}) \left(\sum_{m=1}^{d} n_m^{(\kappa,\zeta)}(\underline{X}^{(\kappa)}(\underline{\hat{x}})) \underline{F}^{(n)}(\underline{U}(\underline{X}^{(\kappa)}(\underline{\hat{x}}), t)) \right) d\hat{s} = \underline{0}.
\end{aligned} \tag{2.8}$$

Choosing test functions belonging to the same space as each component of the transformed numerical solution and resolving the discontinuity in the global approximation on the element boundary using a directional numerical flux function $\underline{F}^* : \Upsilon \times \Upsilon \times \mathbb{S}^{d-1} \to \mathbb{R}^{N_{\text{eq}}}$, typically corresponding to an approximate Riemann solver developed in the context of finite-volume methods (see, for example, Toro [34]), we obtain a weak-form DG discretization of the system in (2.4), in which, for all $\kappa \in \{1, \dots, N_{\text{el}}\}$ and $t \in (0, T)$, we seek a function $\underline{U}^{(h,\kappa)}(\cdot, t) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_{\text{eq}}}$ satisfying

$$\begin{aligned}
&\int_{\hat{\Omega}} \left(V(\underline{\hat{x}}) \frac{\partial J^{(\kappa)}(\underline{\hat{x}}) \underline{U}^{(h,\kappa)}(\underline{\hat{x}}, t)}{\partial t} - \sum_{m=1}^{d} \frac{\partial V(\underline{\hat{x}})}{\partial \hat{x}_m} \underline{F}^{(h,\kappa,m)}(\underline{\hat{x}}, t) \right) d\underline{\hat{x}} \\
&\qquad + \sum_{\zeta=1}^{N_{\text{fac}}} \int_{\hat{\Gamma}^{(\zeta)}} V(\underline{\hat{x}}) J^{(\kappa,\zeta)}(\underline{\hat{x}}) \underline{F}^{(*,\kappa,\zeta)}(\underline{\hat{x}}, t) \, d\hat{s} = \underline{0}, \quad \forall V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}).
\end{aligned} \tag{2.9}$$

In the above, the volume-weighted contravariant flux components are evaluated as

$$\underline{F}^{(h,\kappa,m)}(\underline{\hat{x}}, t) := \sum_{n=1}^{d} J^{(\kappa)}(\underline{\hat{x}}) \left[(\nabla \underline{X}^{(\kappa)}(\underline{\hat{x}}))^{-1} \right]_{mn} \underline{F}^{(n)}(\underline{U}^{(h,\kappa)}(\underline{\hat{x}}, t)), \tag{2.10}$$

and the numerical flux function is evaluated on each facet of the reference element as

$$\underline{F}^{(*,\kappa,\zeta)}(\underline{\hat{x}}, t) := \underline{F}^*(\underline{U}^{(h,\kappa)}(\underline{\hat{x}}, t), \underline{U}^{(+,\kappa,\zeta)}(\underline{\hat{x}}, t), \underline{n}^{(\kappa,\zeta)}(\underline{X}^{(\kappa)}(\underline{\hat{x}}))), \tag{2.11}$$

where $\underline{U}^{(+,\kappa,\zeta)}(\underline{\hat{x}},t) \in \Upsilon$ denotes the external data corresponding to a weakly imposed boundary or interface condition. The initial condition for a DG scheme is typically imposed through a Galerkin projection with respect to the standard L^2 inner product on the physical element; expressing such a projection problem in reference coordinates, we therefore obtain $\underline{U}^{(h,\kappa)}(\cdot,0) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_{\text{eq}}}$ for $\kappa \in \{1,\dots,N_{\text{el}}\}$ by solving

$$\int_{\hat{\Omega}} V(\underline{\hat{x}}) J^{(\kappa)}(\underline{\hat{x}}) \big(\underline{U}^{(h,\kappa)}(\underline{\hat{x}},0) - \underline{U}^{0}(\underline{X}^{(\kappa)}(\underline{\hat{x}}))\big)\, d\underline{\hat{x}} = \underline{0}, \quad \forall\, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}). \tag{2.12}$$

Remark 2.5 As (2.9) and (2.12) generally contain integrals of non-polynomial functions which are impractical or impossible to compute analytically, DG methods typically employ some form of numerical integration, which we discuss in §3.

Flux reconstruction method

While DG and FR methods both employ discontinuous polynomial approximation spaces as described in §2.4, the FR approach involves the discretization of the strong form of (2.4) rather than the weak form, and traditionally employs a collocation-based approximation rather than a Galerkin approach. To construct an FR scheme, we first require a nodal set $\mathcal{S} := \{\underline{\hat{x}}^{(1)},\dots,\underline{\hat{x}}^{(N)}\} \subset \hat{\Omega}$ of cardinality $N = N^*$ which is unisolvent for $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ in the sense that for a given function[3] $U : \hat{\Omega} \to \mathbb{R}$, there exists a unique interpolating polynomial (i.e. collocation projection) $\mathit{\Pi} U \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ satisfying

$$(\mathit{\Pi} U)(\underline{\hat{x}}^{(i)}) = U(\underline{\hat{x}}^{(i)}), \quad \forall\, i \in \{1,\dots,N\}. \tag{2.13}$$

We also require *correction functions* $\underline{G}^{(\zeta,j)} : \hat{\Omega} \to \mathbb{R}^d$ associated with each node in the set $\mathcal{S}^{(\zeta)} := \{\underline{\hat{x}}^{(\zeta,1)},\dots,\underline{\hat{x}}^{(\zeta,N_\zeta)}\} \subseteq \hat{\Gamma}^{(\zeta)}$ for $\zeta \in \{1,\dots,N_{\text{fac}}\}$, which satisfy

$$\underline{\nabla} \cdot \underline{G}^{(\zeta,j)} \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}) \quad \text{and} \quad \underline{G}^{(\zeta,j)}(\underline{\hat{x}}^{(\eta,k)}) \cdot \underline{\hat{n}}^{(\eta)} = \delta_{\zeta\eta}\delta_{jk}, \tag{2.14}$$

where it is typically assumed that $\mathcal{S}^{(\zeta)}$ is of cardinality $N_\zeta = N_\zeta^*$ and unisolvent for the space $\mathbb{P}_{\mathcal{N}}(\hat{\Gamma}^{(\zeta)})$. Weighting each correction function by the normal flux difference at the corresponding node in $\mathcal{S}^{(\zeta)}$ and adding the resulting flux correction to the projected flux inside the reference divergence operator, an FR discretization of the system in (2.4) is given concisely in terms of $\underline{U}^{(h,\kappa)}(\cdot,t) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_{\text{eq}}}$ as

$$\begin{aligned} \frac{\partial \mathit{\Pi}(J^{(\kappa)}\underline{U}^{(h,\kappa)}(\cdot,t))}{\partial t} + \sum_{m=1}^{d} \frac{\partial}{\partial \hat{x}_m} \Bigg(& \mathit{\Pi}\underline{F}^{(h,\kappa,m)}(\cdot,t) \\ & + \sum_{\zeta=1}^{N_{\text{fac}}} \sum_{j=1}^{N_\zeta} \Big(J^{(\kappa,\zeta)}(\underline{\hat{x}}^{(\zeta,j)}) \underline{F}^{(*,\kappa,\zeta)}(\underline{\hat{x}}^{(\zeta,j)},t) \\ & - \sum_{n=1}^{d} \hat{n}_n^{(\zeta)} \big(\mathit{\Pi}\underline{F}^{(h,\kappa,n)}(\cdot,t)\big)(\underline{\hat{x}}^{(\zeta,j)})\Big) G_m^{(\zeta,j)} \Bigg) = \underline{0} \end{aligned} \tag{2.15}$$

for all $\kappa \in \{1,\dots,N_{\text{el}}\}$, and $t \in (0,T)$, where the operator $\mathit{\Pi}$ is taken to act componentwise on vector-valued functions, and is typically also used to impose the initial condition. The properties of the FR scheme in (2.15) are then determined in part by the choice of correction functions, which we discuss in §3.4.

[3] For the purposes of the present analysis, the regularity of such functions is is not of concern.

Remark 2.6 As with the DG scheme in (2.9), the FR formulation in (2.15) is independent of the basis used to represent the numerical solution and may be used with any approximation space satisfying Assumption 2.2. In a conventional FR implementation, the solution is represented using a Lagrange basis (as described, for example, in Appendix A.2) with nodes collocated with those used for the projection in (2.13); however, we note that other choices (e.g. orthogonal modal bases, as described in Appendix A.1 and used alongside nodal bases for the computations in §5) are possible.

3 Algebraic formulation

3.1 Discrete operators and the summation-by-parts property

We now introduce the matrices required for the algebraic formulation of the schemes described in §2.5 and present several assumptions and lemmas which will be employed throughout the analysis in §4. Considering nodal sets $\mathcal{S}$ and $\mathcal{S}^{(\zeta)}$ defined as in §2.5, we relax the assumptions of unisolvency in order to allow for the possibility of over-integration (i.e. allowing for $N \geq N^*$ and $N_\zeta \geq N_\zeta^*$) and construct approximations of the L^2 inner products on the reference element and each of its facets as

$$\int_{\hat{\Omega}} U(\underline{\hat{x}})V(\underline{\hat{x}})\,d\underline{\hat{x}} \approx \left\langle U, V \right\rangle_{\underline{\underline{W}}} := \sum_{i=1}^{N}\sum_{j=1}^{N} U(\underline{\hat{x}}^{(i)})W_{ij}V(\underline{\hat{x}}^{(j)}) \tag{3.1}$$

and

$$\int_{\hat{\Gamma}^{(\zeta)}} U(\underline{\hat{x}})V(\underline{\hat{x}})\,d\hat{s} \approx \left\langle U, V \right\rangle_{\underline{\underline{B}}^{(\zeta)}} := \sum_{i=1}^{N_\zeta}\sum_{j=1}^{N_\zeta} U(\underline{\hat{x}}^{(\zeta,i)})B_{ij}^{(\zeta)}V(\underline{\hat{x}}^{(\zeta,j)}), \tag{3.2}$$

in terms of the matrices $\underline{\underline{W}} \in \mathbb{R}^{N\times N}$ and $\underline{\underline{B}}^{(\zeta)} \in \mathbb{R}^{N_\zeta \times N_\zeta}$, respectively. Following the literature on spectral and spectral-element methods (see, for example, Canuto *et al.* [35, §2.2.3]),[4] we refer to the bilinear forms $\langle\cdot,\cdot\rangle_{\underline{\underline{W}}}$ and $\langle\cdot,\cdot\rangle_{\underline{\underline{B}}^{(\zeta)}}$ as *discrete inner products* and note that such forms allow for the approximation of each term of the integration-by-parts relation on the reference element for $m \in \{1,\dots,d\}$ as

$$\underbrace{\int_{\hat{\Omega}} U(\underline{\hat{x}})\frac{\partial V(\underline{\hat{x}})}{\partial \hat{x}_m}\,d\underline{\hat{x}}}_{\approx \langle U, \partial V/\partial \hat{x}_m\rangle_{\underline{\underline{W}}}} + \underbrace{\int_{\hat{\Omega}} \frac{\partial U(\underline{\hat{x}})}{\partial \hat{x}_m}V(\underline{\hat{x}})\,d\underline{\hat{x}}}_{\approx \langle \partial U/\partial \hat{x}_m, V\rangle_{\underline{\underline{W}}}} = \sum_{\zeta=1}^{N_{\text{fac}}} \underbrace{\int_{\hat{\Gamma}^{(\zeta)}} U(\underline{\hat{x}})V(\underline{\hat{x}})\hat{n}_m^{(\zeta)}\,d\hat{s}}_{\approx \hat{n}_m^{(\zeta)}\langle U,V\rangle_{\underline{\underline{B}}^{(\zeta)}}}, \tag{3.3}$$

where we make the following assumption regarding the above approximations.

Assumption 3.1 *For any $m \in \{1,\dots,d\}$, the integration-by-parts relation in (3.3) is satisfied under the discrete inner products for all $U, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, as given by*

$$\left\langle U, \frac{\partial V}{\partial \hat{x}_m}\right\rangle_{\underline{\underline{W}}} + \left\langle \frac{\partial U}{\partial \hat{x}_m}, V\right\rangle_{\underline{\underline{W}}} = \sum_{\zeta=1}^{N_{\text{fac}}} \hat{n}_m^{(\zeta)}\left\langle U, V\right\rangle_{\underline{\underline{B}}^{(\zeta)}}. \tag{3.4}$$

[4] In such contexts, the weight matrices are generally assumed to be diagonal; Boland and Duris [36] as well as Hunkins [37] provide more general analyses of such approximations.

Remark 3.1 Recalling the identities $\int_{\Omega^{(\kappa)}} V(\underline{x})\,d\underline{x} = \int_{\hat{\Omega}} V(\underline{X}^{(\kappa)}(\underline{\hat{x}}))J^{(\kappa)}(\underline{\hat{x}})\,d\underline{\hat{x}}$ and $\int_{\Gamma^{(\kappa,\zeta)}} V(\underline{x})\,ds = \int_{\hat{\Gamma}^{(\zeta)}} V(\underline{X}^{(\kappa)}(\underline{\hat{x}}))J^{(\kappa,\zeta)}(\underline{\hat{x}})\,d\hat{s}$, the L^2 inner products on each physical element and facet may be approximated analogously to (3.1) and (3.2) by replacing $\underline{\underline{W}}$ and $\underline{\underline{B}}^{(\zeta)}$ with $\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}$ and $\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}$, respectively, where we define

$$\underline{\underline{J}}^{(\kappa)} := \operatorname{diag}\big(J^{(\kappa)}(\underline{\hat{x}}^{(1)}), \ldots, J^{(\kappa)}(\underline{\hat{x}}^{(N)})\big) \tag{3.5}$$

and

$$\underline{\underline{J}}^{(\kappa,\zeta)} := \operatorname{diag}\big(J^{(\kappa,\zeta)}(\underline{\hat{x}}^{(\zeta,1)}), \ldots, J^{(\kappa)}(\underline{\hat{x}}^{(\zeta,N_\zeta)})\big). \tag{3.6}$$

Introducing an arbitrary (ordered) basis $\mathcal{B} := \{\phi^{(1)}, \ldots, \phi^{(N^*)}\}$ for $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, we may define the matrices $\underline{\underline{V}} \in \mathbb{R}^{N\times N^*}$ and $\underline{\underline{R}}^{(\zeta)} \in \mathbb{R}^{N_\zeta \times N^*}$ with entries given by

$$V_{ij} := \phi^{(j)}(\underline{\hat{x}}^{(i)}) \quad \text{and} \quad R^{(\zeta)}_{ij} := \phi^{(j)}(\underline{\hat{x}}^{(\zeta,i)}), \tag{3.7}$$

respectively. Referring to Appendix A as well as to Hesthaven and Warburton [38, §3.1, §6.1, §10.1] and Karniadakis and Sherwin [39, Ch. 3] for details regarding the construction of such polynomial bases, we make the following additional assumption.

Assumption 3.2 *The generalized Vandermonde matrix $\underline{\underline{V}}$ defined as in (3.7) is of rank N^*. Additionally, $\underline{\underline{W}}$ is symmetric positive definite (SPD), and for all $\zeta \in \{1, \ldots, N_{\text{fac}}\}$, the corresponding $\underline{\underline{B}}^{(\zeta)}$ is symmetric positive semidefinite (SPSD).*

Remark 3.2 The condition on the rank of $\underline{\underline{V}}$ generalizes the unisolvency property of $\mathcal{S}$ assumed in §2.5 to the case of $N \geq N^*$. Details regarding the construction of discrete inner products satisfying Assumptions 3.1 and 3.2 for quadrature-based and collocation-based schemes are provided in Appendices B.1 and B.2, respectively.

The choice of a basis defines an isomorphism (i.e. a bijective linear mapping between vector spaces) associating any polynomial $U \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ with a unique vector $\underline{\tilde{u}} \in \mathbb{R}^{N^*}$ containing its expansion coefficients in terms of the basis $\mathcal{B}$, satisfying

$$U(\underline{\hat{x}}) = \sum_{i=1}^{N^*} \tilde{u}_i \phi^{(i)}(\underline{\hat{x}}). \tag{3.8}$$

Such an expansion may be evaluated the nodes in $\mathcal{S}$ and $\mathcal{S}^{(\zeta)}$ through the matrix-vector products

$$\underline{\underline{V}}\,\underline{\tilde{u}} = \begin{bmatrix} U(\underline{\hat{x}}^{(1)}) \\ \vdots \\ U(\underline{\hat{x}}^{(N)}) \end{bmatrix} \quad \text{and} \quad \underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}} = \begin{bmatrix} U(\underline{\hat{x}}^{(\zeta,1)}) \\ \vdots \\ U(\underline{\hat{x}}^{(\zeta,N_\zeta)}) \end{bmatrix}. \tag{3.9}$$

The positive-definiteness of the mass matrix with entries consisting of the discrete inner products $\langle \phi^{(i)}, \phi^{(j)} \rangle_{\underline{\underline{W}}}$ of all pairs of basis functions $\phi^{(i)}, \phi^{(j)} \in \mathcal{B}$ is then demonstrated with the following lemma.

Lemma 3.1 *Under Assumption 3.2, the reference mass matrix $\underline{\underline{M}} := \underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{\underline{V}}$ is SPD, regardless of the accuracy of the approximation in (3.1).*

Proof The symmetry and positivity of the corresponding quadratic form are clear from the structure of $\underline{\underline{M}}$ and the fact that $\underline{\underline{W}}$ is SPD, while the definiteness property follows from the fact that the nullspace of $\underline{\underline{V}}$ contains only the zero vector. □

Associated with the discrete inner product in (3.1) is a projection operator approximating the L^2 projection of a function onto the space $\mathcal{P}_{\mathcal{N}}(\hat{\Omega})$, which we introduce and relate to the collocation projection in (2.13) with the following lemma.

Lemma 3.2 *Under Assumption 3.2, the Galerkin projection $\Pi U \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ of a given function $U : \hat{\Omega} \to \mathbb{R}$ with respect to the discrete inner product, satisfying*

$$\left\langle V, \Pi U - U \right\rangle_{\underline{\underline{W}}} = 0, \quad \forall V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}), \tag{3.10}$$

may be obtained in terms of any basis $\mathcal{B}$ for the space $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ as

$$(\Pi U)(\underline{\hat{x}}) = \sum_{i=1}^{N^*} \left(\sum_{j=1}^{N} P_{ij} U(\underline{\hat{x}}^{(j)}) \right) \phi^{(i)}(\underline{\hat{x}}), \tag{3.11}$$

where we define $\underline{\underline{P}} := \underline{\underline{M}}^{-1} \underline{\underline{V}}^{\mathrm{T}} \underline{\underline{W}}$. Such a projection operator then satisfies $\Pi U = U$ for any $U \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ and recovers the collocation projection in (2.13) when $N = N^*$.

Proof The proofs of such properties are standard (see, for example, [39, §4.1.5.3]). For our purposes, it is important to highlight that the the invertibility of $\underline{\underline{M}}$ follows from its positive-definiteness, which was established in Lemma 3.1. Moreover, we note that the polynomial exactness of the projection results from substituting (3.8) into (3.11) and using $\underline{\underline{P}}\,\underline{\underline{V}} = \underline{\underline{M}}^{-1} \underline{\underline{V}}^{\mathrm{T}} \underline{\underline{W}}\,\underline{\underline{V}} = \underline{\underline{I}}$, and that we recover (2.13) by evaluating (3.11) at each node in $\mathcal{S}$ and using the fact that a square matrix is invertible if and only if it is of full rank to obtain $\underline{\underline{V}}\,\underline{\underline{P}} = \underline{\underline{V}}\,\underline{\underline{V}}^{-1} = \underline{\underline{I}}$ when $N = N^*$. □

Since Assumption 2.2 implies that the reference approximation space is closed under partial differentiation, the operator $\partial / \partial \hat{x}_m$ from $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ onto itself may be represented algebraically in terms of a given basis $\mathcal{B}$ through the matrix $\underline{\underline{D}}^{(m)} \in \mathbb{R}^{N^* \times N^*}$ defined implicitly (see, for example, [25, Eq. (16)]) for each $m \in \{1, \dots, d\}$ such that for any function given as in (3.8), the corresponding partial derivative may be expanded as

$$\frac{\partial U(\underline{\hat{x}})}{\partial \hat{x}_m} = \sum_{i=1}^{N^*} \left(\sum_{j=1}^{N^*} D_{ij}^{(m)} \tilde{u}_j \right) \phi^{(i)}(\underline{\hat{x}}). \tag{3.12}$$

We now present the following lemma, establishing that, under certain conditions, such a matrix satisfies the SBP property, which serves as a discrete analogue of the integration-by-parts relation in (3.3) and is crucial to the analysis in §4.

Lemma 3.3 *Under Assumptions 2.2 and 3.2, the differentiation matrix satisfying (3.12) for a given basis $\mathcal{B}$ and coordinate index $m \in \{1, \dots, d\}$ may be expressed as*

$$\underline{\underline{D}}^{(m)} = \underline{\underline{M}}^{-1} \underline{\underline{S}}^{(m)}, \tag{3.13}$$

where $\underline{\underline{M}} \in \mathbb{R}^{N^* \times N^*}$ is the SPD mass matrix defined in Lemma 3.1 for a discrete inner product in the form of (3.1) and $\underline{\underline{S}}^{(m)} \in \mathbb{R}^{N^* \times N^*}$ is the corresponding stiffness matrix with entries given in terms of each pair of basis functions $\phi^{(i)}, \phi^{(j)} \in \mathcal{B}$ as

$$S_{ij}^{(m)} := \left\langle \phi^{(i)}, \frac{\partial \phi^{(j)}}{\partial \hat{x}_m} \right\rangle_{\underline{\underline{W}}}. \tag{3.14}$$

The SBP property is then satisfied under Assumption 3.1, as given by

$$\underline{\underline{S}}^{(m)} + \big(\underline{\underline{S}}^{(m)}\big)^{\mathrm{T}} = \sum_{\zeta=1}^{N_{\mathrm{fac}}} \hat{n}_m^{(\zeta)} \big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{R}}^{(\zeta)}. \tag{3.15}$$

Proof Taking $U(\underline{\hat{x}}) = \phi^{(j)}(\underline{\hat{x}})$ in (3.12) and substituting the resulting expression for $\partial\phi^{(j)}/\partial\hat{x}_m$ into (3.14), we see that $\underline{\underline{S}}^{(m)} = \underline{\underline{M}}\,\underline{\underline{D}}^{(m)}$. Since $\underline{\underline{M}}$ is SPD and therefore invertible under Assumption 3.2 as a result of Lemma 3.1, such a decomposition is equivalent to (3.13). As noted in [25, Eq. (26)], the SBP property in (3.15) then follows directly from Assumptions 2.2 and 3.1, which can be seen by expanding the functions U, V, $\partial U/\partial\hat{x}_m$, and $\partial V/\partial\hat{x}_m$ in (3.4) with respect to the basis $\mathcal{B}$ and expressing the discrete inner products in matrix form. □

Remark 3.3 As the operators in (3.15) act on expansion coefficients with respect to a given basis, which do not necessarily correspond to nodal values, such an algebraic relation is sometimes referred to as a "modal" SBP property (see, for example, Chen and Shu [40, §3.1]), to distinuish it from the nodal multidimensional SBP property introduced in [21, Definition 2.1], which serves an analogous purpose for the construction and analysis of nodal discretizations for which the approximation is not necessarily defined in terms of a polynomial basis.

3.2 Algebraic formulation of the discontinuous Galerkin method

Given any basis $\mathcal{B}$ for the space $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, each component of the transformed numerical solution for a DG or FR method may be expressed in the form of (3.8) in terms of the expansion coefficients $\underline{\tilde{u}}^{(h,\kappa,e)}(t) \in \mathbb{R}^{N^*}$ as $U_e^{(h,\kappa)}(\underline{\hat{x}},t) = \sum_{i=1}^{N^*} \tilde{u}_i^{(h,\kappa,e)}(t)\phi^{(i)}(\underline{\hat{x}})$. In the particular case of a DG method, employing such an expansion and approximating the integrals in (2.9) using the corresponding discrete inner products results in

$$\begin{aligned}\sum_{j=1}^{N^*} \Big\langle \phi^{(i)}, \phi^{(j)} \Big\rangle_{\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}} \frac{d\tilde{u}_j^{(h,\kappa,e)}(t)}{dt} - \sum_{m=1}^{d} \left\langle \frac{\partial\phi^{(i)}}{\partial\hat{x}_m}, F_e^{(h,\kappa,m)}(\cdot,t) \right\rangle_{\underline{\underline{W}}} \\ + \sum_{\zeta=1}^{N_{\mathrm{fac}}} \Big\langle \phi^{(i)}, F_e^{(*,\kappa,\zeta,e)}(\cdot,t) \Big\rangle_{\underline{\underline{B}}^{(\zeta)}\,\underline{\underline{J}}^{(\kappa,\zeta)}} = 0\end{aligned} \tag{3.16}$$

for all $\kappa \in \{1,\ldots,N_{\mathrm{el}}\}$, $e \in \{1,\ldots,N_{\mathrm{eq}}\}$, $i \in \{1,\ldots,N^*\}$, and $t \in (0,T)$, where it is sufficient to test with each basis function $\phi^{(i)} \in \mathcal{B}$ due to the linearity of all forms with respect to the test function. Defining the local mass matrix $\underline{\underline{M}}^{(\kappa)} := \underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}$ with entries given by $\langle\phi^{(i)},\phi^{(j)}\rangle_{\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}}$ for all pairs of basis functions $\phi^{(i)},\phi^{(j)} \in \mathcal{B}$, we may express (3.16) algebraically using the matrix operators introduced in §3.1 as

$$\begin{aligned}\underline{\underline{M}}^{(\kappa)} \frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt} - \sum_{m=1}^{d} \big(\underline{\underline{D}}^{(m)}\big)^{\mathrm{T}} \underline{\underline{V}}^{\mathrm{T}} \underline{\underline{W}}\,\underline{f}^{(h,\kappa,m,e)}(t) \\ + \sum_{\zeta=1}^{N_{\mathrm{fac}}} \big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{J}}^{(\kappa,\zeta)}\,\underline{f}^{(*,\kappa,\zeta,e)}(t) = \underline{0}\end{aligned} \tag{3.17}$$

for all $\kappa \in \{1,\dots,N_{\mathrm{el}}\}$, $e \in \{1,\dots,N_{\mathrm{eq}}\}$, and $t \in (0,T)$, where we define

$$\underline{f}^{(h,\kappa,m,e)}(t) := \left[F_e^{(h,\kappa,m)}(\underline{\hat{x}}^{(1)},t),\dots,F_e^{(h,\kappa,m)}(\underline{\hat{x}}^{(N)},t)\right]^{\mathrm{T}} \tag{3.18}$$

and

$$\underline{f}^{(*,\kappa,\zeta,e)}(t) := \left[F_e^{(*,\kappa,\zeta)}(\underline{\hat{x}}^{(\zeta,1)},t),\dots,F_e^{(*,\kappa,\zeta)}(\underline{\hat{x}}^{(\zeta,N_\zeta)},t)\right]^{\mathrm{T}}. \tag{3.19}$$

Remark 3.4 In order to obtain the vectors $\underline{f}^{(h,\kappa,m,e)}(t)$ and $\underline{f}^{(*,\kappa,\zeta,e)}(t)$, the fluxes in (2.10) and (2.11) may be evaluated pointwise following the pre-multiplication of $\underline{\tilde{u}}^{(h,\kappa,e)}(t)$ by $\underline{\underline{V}}$ and $\underline{\underline{R}}^{(\zeta)}$ as in (3.9) in order to obtain the corresponding nodal values of the numerical solution. If $\mathcal{B}$ is taken to be a nodal basis defined on the set $\tilde{\mathcal{S}} \subset \hat{\Omega}$ as in Appendix A.2, the pre-multiplication by $\underline{\underline{V}}$ may be avoided if collocation is exploited with $\mathcal{S} = \tilde{\mathcal{S}}$, while the pre-multiplication by $\underline{\underline{R}}^{(\zeta)}$ may be avoided if $\mathcal{S}^{(\zeta)} \subset \tilde{\mathcal{S}}$, reducing the number of floating-point operations per residual evaluation.

The semi-discrete residual may therefore be obtained from (3.17) through inversion or factorization of the local mass matrix, hence requiring the following assumption.

Assumption 3.3 *The local mass matrix $\underline{\underline{M}}^{(\kappa)}$ is invertible for all $\kappa \in \{1,\dots,N_{\mathrm{el}}\}$.*

For the DG and FR schemes described within the present framework, we impose the initial condition by approximating the projection in (2.12) analogously to (3.10) as

$$\left\langle V, U_e^{(h,\kappa)}(\cdot,0) - U_e^0 \circ \underline{X}^{(\kappa)} \right\rangle_{\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}} = 0, \quad \forall\, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}), \tag{3.20}$$

for all $\kappa \in \{1,\dots,N_{\mathrm{el}}\}$ and $e \in \{1,\dots,N_{\mathrm{eq}}\}$. A unique solution $\underline{U}^{(h,\kappa)}(\cdot,0) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_{\mathrm{eq}}}$ to (3.20) then exists under Assumption 3.3, with coefficients given by

$$\underline{\tilde{u}}^{(h,\kappa,e)}(0) = \left(\underline{\underline{M}}^{(\kappa)}\right)^{-1} \underline{\underline{V}}^{\mathrm{T}}\, \underline{\underline{W}}\, \underline{\underline{J}}^{(\kappa)} \left[U_e^0(\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(1)})),\dots,U_e^0(\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(N)}))\right]^{\mathrm{T}}. \tag{3.21}$$

3.3 Algebraic formulation of the flux reconstruction method

Separating the divergence of the projected flux from that of the correction term, defining the scalar-valued *correction field* associated with each node in $\mathcal{S}^{(\zeta)}$ as $H^{(\zeta,j)}(\underline{\hat{x}}) := \underline{\nabla} \cdot \underline{G}^{(\zeta,j)}(\underline{\hat{x}})$, which belongs to $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ as a result of (2.14), and using the more general definition of the projection operator Π in (3.10) as opposed to (2.13) in order to allow for over-integration, the FR method in (2.15) may be expressed as

$$\begin{aligned} &\frac{\partial \Pi(J^{(\kappa)}\underline{U}^{(h,\kappa)}(\cdot,t))}{\partial t} + \sum_{m=1}^{d} \frac{\partial \Pi \underline{F}^{(h,\kappa,m)}(\cdot,t)}{\partial \hat{x}_m} \\ &\quad + \sum_{\zeta=1}^{N_{\mathrm{fac}}} \sum_{j=1}^{N_\zeta} \Big(J^{(\kappa,\zeta)}(\underline{\hat{x}}^{(\zeta,j)}) \underline{F}^{(*,\kappa,\zeta)}(\underline{\hat{x}}^{(\zeta,j)},t) \\ &\quad - \sum_{m=1}^{d} \hat{n}_m^{(\zeta)} \left(\Pi \underline{F}^{(h,\kappa,m)}(\cdot,t)\right)(\underline{\hat{x}}^{(\zeta,j)}) \Big) H^{(\zeta,j)} = \underline{0}. \end{aligned} \tag{3.22}$$

An algebraic formulation is then obtained by expanding all functions belonging to $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ appearing in (3.22) in terms of $\mathcal{B}$ and using the operators in §3.1 to obtain

$$\begin{aligned}&\underline{\underline{\tilde{J}}}^{(\kappa)}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt}+\sum_{m=1}^{d}\underline{\underline{D}}^{(m)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\\&\quad+\sum_{\zeta=1}^{N_{\text{fac}}}\underline{\underline{L}}^{(\zeta)}\left(\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t)-\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\right)=\underline{0}\end{aligned}\tag{3.23}$$

for all $\kappa\in\{1,\dots,N_{\text{el}}\}$, $e\in\{1,\dots,N_{\text{eq}}\}$, and $t\in(0,T)$, where $\underline{\underline{\tilde{J}}}^{(\kappa)}:=\underline{\underline{P}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}$ and $\underline{\tilde{f}}^{(h,\kappa,m,e)}(t):=\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,e)}(t)$ reduce to simply $\underline{\underline{J}}^{(\kappa)}$ and $\underline{f}^{(h,\kappa,m,e)}(t)$, respectively, when $\mathcal{B}$ is taken to be a collocated nodal basis with $\tilde{\mathcal{S}}=\mathcal{S}$ as in Remark 3.4. The correction fields are encoded within the lifting matrices $\underline{\underline{L}}^{(\zeta)}\in\mathbb{R}^{N^*\times N_\zeta}$ satisfying

$$H^{(\zeta,j)}(\underline{\hat{x}})=\sum_{i=1}^{N^*}L_{ij}^{(\zeta)}\phi^{(i)}(\underline{\hat{x}}),\quad\forall\, j\in\{1,\dots,N_\zeta\},\tag{3.24}$$

and may be viewed as generalizing [25, Eq. (22)] to allow for a broader range of correction procedures (i.e. recovering FR schemes that are not equivalent to DG methods) or as extending the formulation in [23, §3.1] to multiple dimensions.

Remark 3.5 Noting that $\underline{\underline{M}}\,\underline{\underline{\tilde{J}}}^{(\kappa)}=\underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}=\underline{\underline{M}}^{(\kappa)}$, we see that for general (i.e. including non-collocated) choices of basis, the matrix pre-multiplying the time derivative in (3.23) has an inverse given by $(\underline{\underline{\tilde{J}}}^{(\kappa)})^{-1}=(\underline{\underline{M}}^{(\kappa)})^{-1}\underline{\underline{M}}$ under Assumption 3.3, ensuring that such a formulation results in a well-defined semi-discrete residual.

3.4 Vincent-Castonguay-Jameson-Huynh correction functions

To complete our algebraic formulation of the FR method in (3.23), the entries of the lifting matrices $\underline{\underline{L}}^{(\zeta)}$, or, equivalently, the expansion coefficients of each correction field (i.e. the divergence of each correction function) in terms of the basis $\mathcal{B}$, must be specified. The following theorem demonstrates how such lifting matrices may be obtained, beginning with the fundamental assumptions associated with the VCJH family of correction functions introduced in [10–12].

Theorem 3.1 *Given any space satisfying Assumption 2.2, suppose that there exist VCJH correction functions $\underline{G}^{(\zeta,j)}$ and correction fields $H^{(\zeta,j)}$ parametrized by $\mathcal{C}:=\{c^{(\underline{\alpha})}\}_{\underline{\alpha}\in\mathcal{M}}$ with $\mathcal{M}\subset\mathcal{N}$ and $c^{(\underline{\alpha})}\in\mathbb{R}$ for all $\underline{\alpha}\in\mathcal{M}$, satisfying (2.14) as well as*

$$\int_{\hat{\Omega}}\left(\underline{G}^{(\zeta,j)}(\underline{\hat{x}})\cdot\underline{\nabla}V(\underline{\hat{x}})-\sum_{\underline{\alpha}\in\mathcal{M}}\frac{c^{(\underline{\alpha})}}{|\hat{\Omega}|}\partial^{(\underline{\alpha})}V(\underline{\hat{x}})\partial^{(\underline{\alpha})}H^{(\zeta,j)}(\underline{\hat{x}})\right)d\underline{\hat{x}}=0\tag{3.25}$$

for all $V\in\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, $\zeta\in\{1,\dots,N_{\text{fac}}\}$, and $j\in\{1,\dots,N_\zeta\}$. Then, given any basis $\mathcal{B}$ for such a space and discrete inner products defined as in (3.1) and (3.2), satisfying

$$\langle U,V\rangle_{\underline{\underline{W}}}=\int_{\hat{\Omega}}U(\underline{\hat{x}})V(\underline{\hat{x}})\,d\underline{\hat{x}},\quad\forall\, U,V\in\mathbb{P}_{\mathcal{N}}(\hat{\Omega}),\tag{3.26}$$

and

$$\left\langle U, V \right\rangle_{\underline{\underline{B}}^{(\zeta)}} = \int_{\hat{\Gamma}^{(\zeta)}} U(\underline{\hat{x}}) V(\underline{\hat{x}}) \, d\hat{s}, \quad \forall U, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Gamma}^{(\zeta)}), \tag{3.27}$$

such correction fields may be expanded as in (3.24) for any $\zeta \in \{1, \ldots, N_{\text{fac}}\}$ *with*

$$\underline{\underline{L}}^{(\zeta)} := \left(\underline{\underline{M}} + \underline{\underline{K}}\right)^{-1} \left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \quad \textit{and} \quad \underline{\underline{K}} := \sum_{\underline{\alpha} \in \mathcal{M}} \frac{c^{(\underline{\alpha})}}{|\hat{\Omega}|} \left(\underline{\underline{D}}^{(\underline{\alpha})}\right)^{\mathrm{T}} \underline{\underline{M}} \underline{\underline{D}}^{(\underline{\alpha})}, \tag{3.28}$$

where we define $\underline{\underline{D}}^{(\underline{\alpha})} := \left(\underline{\underline{D}}^{(1)}\right)^{\alpha_1} \cdots \left(\underline{\underline{D}}^{(d)}\right)^{\alpha_d}$ *and assume that* $\underline{\underline{M}} + \underline{\underline{K}}$ *is SPD.*

Proof Applying integration by parts to the first term of the integrand in (3.25) and noting that the resulting integrals are L^2 inner products of functions in $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ and $\mathbb{P}_{\mathcal{N}}(\hat{\Gamma}^{(\zeta)})$, we may use the exactness conditions in (3.26) and (3.27) as well as the fact that $H^{(\zeta,j)}$ is defined as the divergence of $\underline{G}^{(\zeta,j)}$ to obtain

$$\begin{aligned} &\left\langle V, H^{(\zeta,j)} \right\rangle_{\underline{\underline{W}}} + \sum_{\underline{\alpha} \in \mathcal{M}} \frac{c^{(\underline{\alpha})}}{|\hat{\Omega}|} \left\langle \partial^{(\underline{\alpha})} V, \partial^{(\underline{\alpha})} H^{(\zeta,j)} \right\rangle_{\underline{\underline{W}}} \\ &\quad = \sum_{\eta=1}^{N_{\text{fac}}} \left\langle V, \underline{G}^{(\zeta,j)} \cdot \underline{\hat{n}}^{(\eta)} \right\rangle_{\underline{\underline{B}}^{(\eta)}}, \quad \forall V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}). \end{aligned} \tag{3.29}$$

Since all terms in (3.29) are linear with respect to V, we can test with $\phi^{(k)} \in \mathcal{B}$ and apply the expansion in (3.24) to $H^{(\zeta,j)} \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ in order to obtain

$$\begin{aligned} &\sum_{i=1}^{N^*} \left(\left\langle \phi^{(k)}, \phi^{(i)} \right\rangle_{\underline{\underline{W}}} + \sum_{\underline{\alpha} \in \mathcal{M}} \frac{c^{(\underline{\alpha})}}{|\hat{\Omega}|} \left\langle \partial^{(\underline{\alpha})} \phi^{(k)}, \partial^{(\underline{\alpha})} \phi^{(i)} \right\rangle_{\underline{\underline{W}}} \right) L_{ij}^{(\zeta)} \\ &\quad = \sum_{\eta=1}^{N_{\text{fac}}} \left\langle \phi^{(k)}, \underline{G}^{(\zeta,j)} \cdot \underline{\hat{n}}^{(\eta)} \right\rangle_{\underline{\underline{B}}^{(\eta)}}, \quad \forall k \in \{1, \ldots, N^*\}. \end{aligned} \tag{3.30}$$

Recognizing the entries of $\underline{\underline{M}}$ and $\underline{\underline{K}}$ on the left-hand side, expanding the discrete inner product on the right-hand side, and using the second condition in (2.14), we see that the j^{th} column of the matrix $\underline{\underline{L}}^{(\zeta)}$, containing the expansion coefficients for the corresponding correction field $H^{(\zeta,j)} \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ in terms of the basis $\mathcal{B}$, may be obtained by solving the linear system of equations given by

$$\sum_{i=1}^{N^*} \left(M_{ki} + K_{ki}\right) L_{ij}^{(\zeta)} = \sum_{i=1}^{N_\zeta} R_{ik}^{(\zeta)} B_{ij}^{(\zeta)}, \quad \forall k \in \{1, \ldots, N^*\}, \tag{3.31}$$

for which a unique solution in the form of (3.28) exists if and only if $\underline{\underline{M}} + \underline{\underline{K}}$ is invertible, which follows from the assumption of positive-definiteness. As such a system is obtained as a consequence of the conditions in (2.14) and (3.25) for all $j \in \{1, \ldots, N_\zeta\}$ and $\zeta \in \{1, \ldots, N_{\text{fac}}\}$, the correction fields are then fully specified under the present assumptions by defining the lifting matrices as in (3.28). □

Remark 3.6 While the conditions on the correction functions in Theorem 3.1 imply the existence of correction fields given by (3.24) in terms of a particular form of $\underline{\underline{L}}^{(\zeta)}$, the converse does not hold in general, as we cannot necessarily associate a given correction field $H^{(\zeta,j)}$ obtained from $\underline{\underline{L}}^{(\zeta)}$ as in (3.24) with a unique correction function $\underline{G}^{(\zeta,j)}$ satisfying the conditions of Theorem 3.1. In fact, correction functions satisfying such conditions have not, to the authors' knowledge, been explicitly constructed for simplicial elements in two or more space dimensions.[5] The analysis in §4 based on properties of the matrix operators appearing in (3.23) is therefore more general than that explicitly relying on properties of the correction functions, and moreover ensures that all mathematical objects involved in the proofs are well defined.

Remark 3.7 The accuracy conditions (3.26) and (3.27) are only necessary to demonstrate the recovery of the existing VCJH correction fields within the present framework. The theory in §4 instead requires Assumptions 3.1 and 3.2, the former being satisfied under slightly weaker accuracy constraints (for example, with $\mathbb{P}_{\mathcal{N}}(\hat{\Omega}) = \mathbb{P}_p(\hat{\Omega})$, a volume quadrature rule of total degree $2p-1$ does not satisfy (3.26), but is sufficient for Assumption 3.1 to hold). Discrete inner products satisfying Assumptions 3.1 and 3.2 may therefore be used within the present framework to construct lifting matrices in the form of (3.28) even when the conditions of Theorem 3.1 are not satisfied, although (as discussed in [23, §3.6] for collocated LGL quadrature) the corresponding correction fields differ in such cases from those of the standard VCJH schemes.

The choices of $\mathcal{M}$ and $\mathcal{C}$ are then constrained by the following assumption.

Assumption 3.4 *The correction fields are given in terms of (3.24) with the matrix $\underline{\underline{L}}^{(\zeta)}$ defined as in (3.28), where the multi-index set $\mathcal{M}$ is chosen such that $\underline{\underline{K}}\,\underline{\underline{D}}^{(m)} = \underline{\underline{0}}$ for all $m \in \{1,\dots,d\}$ and the coefficients $\mathcal{C}$ are chosen such that $\underline{\underline{M}} + \underline{\underline{K}}$ is SPD.*

Conditions for Assumption 3.4 to be satisfied depend on the choice of approximation space $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, and are discussed further in Appendix C. If we begin with a set of correction functions satisfying (2.14) and (3.25), the first condition in Assumption 3.4 amounts to choosing $\mathcal{M}$ such that the second term of the integrand in (3.25) is constant. In such a case, we recover a generalization of the fundamental assumption for energy stability employed in [10, Eqs. (3.31), (3.32)] and [11, Eq. (5.37)],

$$\int_{\hat{\Omega}} \underline{G}^{(\zeta,j)}(\underline{\hat{x}}) \cdot \underline{\nabla} V(\underline{\hat{x}})\, d\underline{\hat{x}} = \sum_{\underline{\alpha} \in \mathcal{M}} c^{(\underline{\alpha})} \left(\partial^{(\underline{\alpha})} V\right) \left(\partial^{(\underline{\alpha})} H^{(\zeta,j)}\right), \quad \forall\, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}), \tag{3.32}$$

where we slightly abuse notation to treat $\partial^{(\underline{\alpha})} V, \partial^{(\underline{\alpha})} H^{(\zeta,j)} \in \mathbb{P}_0(\hat{\Omega})$ as constants.

3.5 Connection between FR methods and strong-form filtered DG schemes

Considering an FR method for which the correction fields are given as in Theorem 3.1 with $\underline{\underline{K}} = \underline{\underline{0}}$, as is obtained when (3.25) is satisfied with $c^{(\underline{\alpha})} = 0$ for all $\underline{\alpha} \in \mathcal{M}$, we may pre-multiply the resulting scheme in the form of (3.23) by the reference mass

[5] Recalling the formulation in (3.22), as well as those appearing elsewhere in the literature (e.g. [11, Eq. (4.35)] and [12, Eq. (29)]), it is the correction fields $H^{(\zeta,j)}$, and not the correction functions $\underline{G}^{(\zeta,j)}$, that appear explicitly in the implementation of a multidimensional FR scheme, and thus such a construction is not necessary.

matrix $\underline{\underline{M}}$ in order to recover a DG method in strong conservation form (see, for example, [10, §3.5.1], [11, §5.4], [12, §5], and [16, §2.3]), which is given by

$$\begin{aligned}&\underline{\underline{M}}^{(\kappa)}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt}+\sum_{m=1}^{d}\underline{\underline{S}}^{(m)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\\&\quad+\sum_{\zeta=1}^{N_{\mathrm{fac}}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\left(\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t)-\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\right)=\underline{0}\end{aligned}\tag{3.33}$$

for all $\kappa\in\{1,\dots,N_{\mathrm{el}}\}$, $e\in\{1,\dots,N_{\mathrm{eq}}\}$, and $t\in(0,T)$, where we note that the above requires the projection in (3.10) to be applied prior to differentiating each transformed flux component as well as prior to evaluating the normal component of the flux on each facet.

Remark 3.8 It is also possible to compute the volume terms of a strong-form DG or FR method by differentiating the flux components in (2.10) exactly using the chain rule and integrating or projecting afterwards, and to compute the facet terms by evaluating the normal component of the flux on each facet directly in terms of the numerical solution on $\mathcal{S}^{(\zeta)}$ without first projecting.[6] However, the theoretical analysis in §4 based on the SBP property does not apply to such alternative formulations, which, as discussed in [26, §4] and [9, §3.3], generally differ from those in (3.23) and (3.33) when applied to nonlinear or variable-coefficient problems and may therefore result in non-conservative strong-form discretizations which are not equivalent to their weak-form counterparts.

For more general choices of $\underline{\underline{K}}\neq\underline{0}$ satisfying Assumption 3.4, it was recognized in [13] (and extended to multidimensional formulations in [12, §5] and [16, §2.4]) that the resulting FR formulation recovers a linearly filtered DG method in the sense that the FR residual may be recovered by pre-multiplying the strong-form DG residual, which we obtain through solving (3.33) for the time derivative, by a constant filter matrix. Employing a transformation to a modal basis, it can be shown that, at least in the case of an affine mapping between reference and physical coordinates, such a filter acts only on polynomial modes of the highest degree contained within the space $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ (see, for example, [13, §3.2] and [12, Appendix B]). The following lemma illustrates such an equivalence within the context of the present framework.

Lemma 3.4 *Defining the VCJH filter matrices on the reference element and physical element, respectively, as*

$$\underline{\underline{F}}:=\left(\underline{\underline{I}}+\underline{\underline{M}}^{-1}\underline{\underline{K}}\right)^{-1}\quad\text{and}\quad\underline{\underline{F}}^{(\kappa)}:=\left(\underline{\tilde{J}}^{(\kappa)}\right)^{-1}\underline{\underline{F}}\,\underline{\tilde{J}}^{(\kappa)},\tag{3.34}$$

the FR method in (3.23) is equivalent, under Assumptions 2.2, 3.2, 3.3, and 3.4, to the strong-form filtered DG scheme given by

$$\begin{aligned}&\underline{\underline{M}}^{(\kappa)}\left(\underline{\underline{F}}^{(\kappa)}\right)^{-1}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt}+\sum_{m=1}^{d}\underline{\underline{S}}^{(m)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\\&\quad+\sum_{\zeta=1}^{N_{\mathrm{fac}}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\left(\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t)-\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\right)=\underline{0}\end{aligned}\tag{3.35}$$

[6] As noted in [26, §4] in the context of the DGSEM-LGL, such a modification to the facet terms has no effect on the discretization when $N=N^*$ and $\mathcal{S}^{(\zeta)}\subset\mathcal{S}$ for all $\zeta\in\{1,\dots,N_{\mathrm{fac}}\}$.

for all $\kappa \in \{1,\dots,N_{\text{el}}\}$, $e \in \{1,\dots,N_{\text{eq}}\}$, *and* $t \in (0,T)$.

Proof Pre-multiplying (3.23) by $\underline{\underline{M}} + \underline{\underline{K}}$ and simplifying the first term by using $\underline{\underline{M}} + \underline{\underline{K}} = \underline{\underline{M}}\underline{\underline{F}}^{-1}$ and $\underline{\underline{M}}^{(\kappa)} = \underline{\underline{M}}\underline{\underline{\tilde{J}}}^{(\kappa)}$ to obtain

$$\left(\underline{\underline{M}} + \underline{\underline{K}}\right)\underline{\tilde{J}}^{(\kappa)} = \underline{\underline{M}}^{(\kappa)}\left(\underline{\underline{F}}^{(\kappa)}\right)^{-1}, \tag{3.36}$$

where all necessary inverses exist under the present assumptions, the FR formulation may be expressed equivalently as

$$\begin{aligned}&\underline{\underline{M}}^{(\kappa)}\left(\underline{\underline{F}}^{(\kappa)}\right)^{-1}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt} + \sum_{m=1}^{d}\underline{\underline{S}}^{(m)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t) + \sum_{m=1}^{d}\underline{\underline{K}}\underline{\underline{D}}^{(m)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\\ &\quad + \sum_{\zeta=1}^{N_{\text{fac}}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\left(\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t) - \sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\right) = \underline{0}.\end{aligned} \tag{3.37}$$

Noting that the third term of (3.37) vanishes when $\underline{\underline{K}}\underline{\underline{D}}^{(m)} = \underline{\underline{0}}$ for all $m \in \{1,\dots,d\}$, as is the case for methods satisfying Assumption 3.4, we therefore recover (3.35). □

Remark 3.9 As shown in [16, §2.4], the filtered DG scheme recovered via FR in (3.35) be expressed equivalently with the filter matrix only pre-multiplying the facet contributions to the residual. In this work, we instead retain the filter on all terms (i.e. as a modification to the mass matrix) in order to facilitate the analysis in §4.

4 Theoretical analysis

4.1 Equivalence between strong and weak forms

Just as integration by parts allows for the weak form of a partial differential equation to be obtained from the strong form, the SBP property in (3.15) may be used analogously to transform a strong-form discretization into a weak-form discretization, and vice versa (as discussed in [17, §8.1]). Although not relying explicitly on the matrix form of the SBP property in (3.15), Kopriva and Gassner proved in [26, §3] that for the DGSEM-LG and DGSEM-LGL on curvilinear hexahedral elements, the strong form in (3.33) is equivalent to the weak form in (3.17). Such an equivalence is extended to more general DG formulations with the following theorem.

Theorem 4.1 *Under Assumptions 2.2, 3.1, 3.2, and 3.3, the strong-form DG method in (3.33) is equivalent to the weak-form DG method in (3.17).*

Proof Beginning with (3.33) and applying the SBP property in (3.15) to the matrix $\underline{\underline{S}}^{(m)}$ for all $m \in \{1,\dots,d\}$, we obtain

$$\begin{aligned}&\underline{\underline{M}}^{(\kappa)}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt} + \sum_{m=1}^{d}\left(\sum_{\zeta=1}^{N_{\text{fac}}}\hat{n}_m^{(\zeta)}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)} - \left(\underline{\underline{S}}^{(m)}\right)^{\mathrm{T}}\right)\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\\ &\quad + \sum_{\zeta=1}^{N_{\text{fac}}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\left(\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t) - \sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)\right) = \underline{0}.\end{aligned} \tag{4.1}$$

Recognizing that $\sum_{\zeta=1}^{N_{\text{fac}}}\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)$ cancels from the second and third terms on the left-hand side of (4.1) and noting that $(\underline{\underline{S}}^{(m)})^{\mathrm{T}}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t) = (\underline{\underline{D}}^{(m)})^{\mathrm{T}}\underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{f}^{(h,\kappa,m,e)}(t)$, we then recover the weak-form DG method in (3.17). □

Exploiting the connection between FR methods and filtered DG schemes established in Lemma 3.4, Theorem 4.1 also implies that an equivalent weak form may be recovered from the FR method in (3.23), as is demonstrated with the following theorem.

Theorem 4.2 *Under Assumptions 2.2, 3.1, 3.2, 3.3, and 3.4, the FR method in (3.23) is equivalent to the weak-form filtered DG scheme given by*

$$\begin{aligned}\underline{\underline{M}}^{(\kappa)}\big(\underline{\underline{F}}^{(\kappa)}\big)^{-1}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt} - \sum_{m=1}^{d}\big(\underline{\underline{D}}^{(m)}\big)^{\mathrm{T}}\underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{f}^{(h,\kappa,m,e)}(t)\\ + \sum_{\zeta=1}^{N_{\text{fac}}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t) = \underline{0}\end{aligned} \tag{4.2}$$

for all $\kappa \in \{1,\dots,N_{\text{el}}\}$, $e \in \{1,\dots,N_{\text{eq}}\}$, *and* $t \in (0,T)$.

Proof As shown in Lemma 3.4, the FR scheme in (2.15) is equivalent under the present assumptions to the strong-form filtered DG scheme in (3.35), which differs from (3.33) only by the mass matrix. We may therefore proceed analogously to the proof of Theorem 4.1 by applying the SBP property in (3.15) to the matrix $\underline{\underline{S}}^{(m)}$ in (3.35) for all $m \in \{1,\dots,d\}$ in order to recover the weak formulation in (4.2). □

Remark 4.1 Although the weak-form filtered and unfiltered DG methods in (4.2) and (3.17) are mathematically equivalent (under the stated assumptions) to the strong-form FR and DG schemes in (3.23) and (3.33), respectively, the weak-form discretizations have the potential for reduced computational expense relative to their corresponding strong forms due to the elimination of contributions of the form $\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t)$ within the facet terms of the semi-discrete residual.

Remark 4.2 The weak form of the VCJH family of FR methods proposed in [16, Eq. (3.7)] may be expressed in the present notation as

$$\begin{aligned}\underline{\underline{M}}^{(\kappa)}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt} - \sum_{m=1}^{d}\big(\underline{\underline{S}}^{(m)}\big)^{\mathrm{T}}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t) + \sum_{\zeta=1}^{N_{\text{fac}}}\underline{\underline{F}}^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t)\\ + \sum_{\zeta=1}^{N_{\text{fac}}}\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\big(\underline{\underline{I}} - \underline{\underline{F}}^{\mathrm{T}}\big)\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{f}}^{(h,\kappa,m,e)}(t) = \underline{0},\end{aligned} \tag{4.3}$$

which is algebraically equivalent to the strong formulation in (3.23) under the assumptions of Theorem 4.2, but has a higher computational cost than the weak-form DG method in (3.17) when $\underline{\underline{K}} \neq \underline{\underline{0}}$. This does not, however, imply that VCJH schemes with $\underline{\underline{K}} \neq \underline{\underline{0}}$ are inherently more computationally expensive than DG methods, since by instead using the weak formulation in (4.2), we are, in fact, able to attain equivalent computational expense to (3.17) with general choices of $\underline{\underline{K}}$ satisfying Assumption 3.4.

4.2 Conservation

The proofs of conservation in this section and the proof of energy stability in §4.3 make use of the following assumption regarding the conformity of the mesh.

Assumption 4.1 *For each pair of element indices $\kappa, \nu \in \{1, \dots, N_{\text{el}}\}$ with $\kappa \neq \nu$ such that $\partial\Omega^{(\kappa)} \cap \partial\Omega^{(\nu)} \neq \emptyset$, there exist facet indices $\zeta, \eta \in \{1, \dots, N_{\text{fac}}\}$ such that either $\Gamma^{(\kappa,\zeta)} = \Gamma^{(\nu,\eta)}$, or the facets are connected via periodic boundary conditions. For such corresponding facets, the nodes align in physical space up to some permutation $\tau^{(\kappa,\zeta)}$ of the nodal indices, satisfying*

$$\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(\zeta,i)}) = \underline{X}^{(\nu)}(\underline{\hat{x}}^{(\eta,\tau^{(\kappa,\zeta)}(i))}), \quad \forall\, i \in \{1, \dots, N_\zeta\}, \tag{4.4}$$

with an additional translation in the periodic case. Moreover, $\underline{\underline{B}}^{(\eta)}\underline{\underline{J}}^{(\nu,\eta)}$ and $\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}$ are congruent via the permutation matrix $\underline{\underline{T}}^{(\kappa,\zeta)} \in \mathbb{R}^{N_\zeta \times N_\zeta}$, as given by

$$\underline{\underline{B}}^{(\eta)}\underline{\underline{J}}^{(\nu,\eta)} = \big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}, \quad \text{where} \quad T_{ij}^{(\kappa,\zeta)} := \delta_{\tau^{(\kappa,\zeta)}(i),j}. \tag{4.5}$$

We may now demonstrate algebraically that the weak-form unfiltered and filtered DG methods are discretely conservative, as given by the following theorem.

Theorem 4.3 *Under Assumptions 2.2, 3.2, 3.3, the weak-form unfiltered and filtered DG methods in (3.17) and (4.2), respectively, the latter with a filter matrix associated with correction fields satisfying Assumption 3.4, are locally conservative, satisfying*

$$\frac{d}{dt}\Big\langle 1, U_e^{(h,\kappa)}(\cdot,t)\Big\rangle_{\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}} + \sum_{\zeta=1}^{N_{\text{fac}}}\Big\langle 1, F_e^{(*,\kappa,\zeta)}(\cdot,t)\Big\rangle_{\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}} = 0 \tag{4.6}$$

for all $\kappa \in \{1, \dots, N_{\text{el}}\}$, $e \in \{1, \dots, N_{\text{eq}}\}$, and $t \in (0, T)$, where, abusing notation, we use $1 \in \mathbb{P}_0(\hat{\Omega})$ to denote the constant function $\hat{\Omega} \mapsto 1$. Moreover, such schemes are globally conservative with the addition of Assumption 4.1, satisfying

$$\frac{d}{dt}\sum_{\kappa=1}^{N_{\text{el}}}\Big\langle 1, U_e^{(h,\kappa)}(\cdot,t)\Big\rangle_{\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}} + \sum_{(\kappa,\zeta)\in\mathcal{F}_{\partial\Omega}}\Big\langle 1, F_e^{(*,\kappa,\zeta)}(\cdot,t)\Big\rangle_{\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}} = 0 \tag{4.7}$$

for all $e \in \{1, \dots, N_{\text{eq}}\}$ and $t \in (0, T)$, provided that $\underline{F}^*$ is conservative in the sense that $\underline{F}^*(\underline{U}^-, \underline{U}^+, \underline{n}) = -\underline{F}^*(\underline{U}^+, \underline{U}^-, -\underline{n})$ for all $\underline{U}^-, \underline{U}^+ \in \Upsilon$ and $\underline{n} \in \mathbb{S}^{d-1}$.

Proof Considering the weak-form filtered DG scheme in (4.2), which recovers the standard DG method in (3.17) when $\underline{\underline{K}} = \underline{\underline{0}}$, we first consider a single element and pre-multiply by the transpose of $\tilde{\underline{1}} := \underline{\underline{P}}\underline{1}$, which, by Assumption 2.2 and Lemma 3.2, contains the expansion coefficients for the constant function $1 \in \mathbb{P}_0(\hat{\Omega})$ in terms of the basis $\mathcal{B}$. Using the relation in (3.36), we therefore obtain

$$\begin{aligned}\tilde{\underline{1}}^{\mathrm{T}}\big(\underline{\underline{M}} + \underline{\underline{K}}\big)\underline{\tilde{J}}^{(\kappa)}\frac{d\underline{\tilde{u}}^{(h,\kappa,e)}(t)}{dt} - \sum_{m=1}^{d}\tilde{\underline{1}}^{\mathrm{T}}\big(\underline{\underline{D}}^{(m)}\big)^{\mathrm{T}}\underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{f}^{(h,\kappa,m,e)}(t) \\ + \sum_{\zeta=1}^{N_{\text{fac}}}\tilde{\underline{1}}^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t) = 0,\end{aligned} \tag{4.8}$$

where the second term on the left-hand side vanishes due to the fact that $\underline{\underline{D}}^{(m)}\underline{\tilde{1}} = \underline{0}$ for all $m \in \{1, \dots, d\}$, as any partial derivative of a constant function is zero. Considering the form of $\underline{\underline{K}}$ given in (3.28), such a property implies that $\underline{\underline{K}}\underline{\tilde{1}} = \underline{\underline{K}}^{\mathrm{T}}\underline{\tilde{1}} = \underline{0}$, which may be used along with $\underline{\underline{V}}\underline{\tilde{1}} = \underline{1}$ and $\underline{\underline{R}}^{(\zeta)}\underline{\tilde{1}} = \underline{1}$ for all $\zeta \in \{1, \dots, N_{\mathrm{fac}}\}$ to obtain

$$\frac{d}{dt}\left(\underline{1}^{\mathrm{T}}\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}\,\underline{\tilde{u}}^{(h,\kappa,e)}(t)\right) + \sum_{\zeta=1}^{N_{\mathrm{fac}}} \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t) = 0, \tag{4.9}$$

recovering the statement of local conservation in (4.6). To establish global conservation, we sum (4.6) over all $\kappa \in \{1, \dots, N_{\mathrm{el}}\}$ and split the net flux contribution $\Phi_e^{(\kappa,\zeta)}(t) := \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t) + \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\eta)}\underline{\underline{J}}^{(\nu,\eta)}\underline{f}^{(*,\nu,\eta,e)}(t)$ from each shared interface equally between the index pairs $(\kappa,\zeta), (\nu,\eta) \in \mathcal{F}_\Omega$ corresponding to facets $\Gamma^{(\kappa,\zeta)} \subset \partial\Omega^{(\kappa)}$ and $\Gamma^{(\nu,\eta)} \subset \partial\Omega^{(\nu)}$ which are either coincident or joined through periodicity, resulting in

$$\begin{aligned}\frac{d}{dt}\sum_{\kappa=1}^{N_{\mathrm{el}}}\left\langle 1, U_e^{(h,\kappa)}(\cdot,t)\right\rangle_{\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}} + \sum_{(\kappa,\zeta)\in\mathcal{F}_\Omega}\frac{1}{2}\Phi_e^{(\kappa,\zeta)}(t) \\ + \sum_{(\kappa,\zeta)\in\mathcal{F}_{\partial\Omega}}\left\langle 1, F_e^{(*,\kappa,\zeta)}(\cdot,t)\right\rangle_{\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}} = 0.\end{aligned} \tag{4.10}$$

Since (4.5) holds under Assumption 4.1 and any permutation matrix $\underline{\underline{T}}^{(\kappa,\zeta)}$ satisfies $\underline{\underline{T}}^{(\kappa,\zeta)}\underline{1} = \underline{1}$, the contribution from each interface may be simplified as

$$\begin{aligned}\Phi_e^{(\kappa,\zeta)}(t) &= \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,e)}(t) + \underline{1}^{\mathrm{T}}\left(\underline{\underline{T}}^{(\kappa,\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{f}^{(*,\nu,\eta,e)}(t) \\ &= \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\left(\underline{f}^{(*,\kappa,\zeta,e)}(t) + \underline{\underline{T}}^{(\kappa,\zeta)}\underline{f}^{(*,\nu,\eta,e)}(t)\right).\end{aligned} \tag{4.11}$$

Using (4.4) and the fact that the outward unit normal vectors for an interior or periodic interface corresponding to $(\kappa,\zeta), (\nu,\eta) \in \mathcal{F}_\Omega$ satisfy

$$\underline{n}^{(\kappa,\zeta)}\left(\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(\zeta,i)})\right) = -\underline{n}^{(\nu,\eta)}\left(\underline{X}^{(\nu)}(\underline{\hat{x}}^{(\eta,\tau^{(\kappa,\zeta)}(i))})\right), \quad \forall\, i \in \{1, \dots, N_\zeta\}, \tag{4.12}$$

the terms in parentheses on the last line of (4.11) are given by

$$\underline{f}^{(*,\kappa,\zeta,e)}(t) = \begin{bmatrix} F_e^*\left(\underline{U}^{(h,\kappa)}(\underline{\hat{x}}^{(\zeta,1)},t), \underline{U}^{(h,\nu)}(\underline{\hat{x}}^{(\eta,\tau^{(\kappa,\zeta)}(1))},t), \right. \\ \left. \qquad\qquad \underline{n}^{(\kappa,\zeta)}(\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(\zeta,1)}))\right) \\ \vdots \\ F_e^*\left(\underline{U}^{(h,\kappa)}(\underline{\hat{x}}^{(\zeta,N_\zeta)},t), \underline{U}^{(h,\nu)}(\underline{\hat{x}}^{(\eta,\tau^{(\kappa,\zeta)}(N_\zeta))},t), \right. \\ \left. \qquad\qquad \underline{n}^{(\kappa,\zeta)}(\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(\zeta,N_\zeta)}))\right) \end{bmatrix} \tag{4.13}$$

and

$$\underline{\underline{T}}^{(\kappa,\zeta)}\underline{f}^{(*,\nu,\eta,e)}(t) = \begin{bmatrix} F_e^*\left(\underline{U}^{(h,\nu)}(\underline{\hat{x}}^{(\eta,\tau^{(\kappa,\zeta)}(1))},t), \underline{U}^{(h,\kappa)}(\underline{\hat{x}}^{(\zeta,1)},t), \right. \\ \left. \qquad\qquad -\underline{n}^{(\kappa,\zeta)}(\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(\zeta,1)}))\right) \\ \vdots \\ F_e^*\left(\underline{U}^{(h,\nu)}(\underline{\hat{x}}^{(\eta,\tau^{(\kappa,\zeta)}(N_\zeta))},t), \underline{U}^{(h,\kappa)}(\underline{\hat{x}}^{(\zeta,N_\zeta)},t), \right. \\ \left. \qquad\qquad -\underline{n}^{(\kappa,\zeta)}(\underline{X}^{(\kappa)}(\underline{\hat{x}}^{(\zeta,N_\zeta)}))\right) \end{bmatrix}, \tag{4.14}$$

which sum to the zero vector due to the conservation property of the numerical flux. Each term of the second sum in (4.10) then vanishes, and hence we obtain (4.7). □

Remark 4.3 The relation in (4.6) may be expressed in terms of quadrature as

$$\begin{aligned}&\frac{d}{dt}\underbrace{\sum_{i=1}^{N}\underline{U}^h(\underline{X}^{(\kappa)}(\hat{\underline{x}}^{(i)}),t)J^{(\kappa)}(\hat{\underline{x}}^{(i)})\,\omega^{(i)}}_{\approx\int_{\Omega^{(\kappa)}}\underline{U}^h(\underline{x},t)\,d\underline{x}}\\&+\sum_{\zeta=1}^{N_{\text{fac}}}\underbrace{\sum_{i=1}^{N_\zeta}\underline{F}^{(*,\kappa,\zeta)}(\hat{\underline{x}}^{(i)},t)J^{(\kappa,\zeta)}(\hat{\underline{x}}^{(\zeta,i)})\,\omega^{(\zeta,i)}}_{\approx\int_{\Gamma^{(\kappa,\zeta)}}\underline{F}^{(*,\kappa,\zeta)}((\underline{X}^{(\kappa)})^{-1}(\underline{x}),t)\,ds}=\underline{0}\end{aligned}\tag{4.15}$$

for all $\kappa\in\{1,\dots,N_{\text{el}}\}$ and $t\in(0,T)$, where the reference volume and facet quadrature weights are given by $\omega^{(i)}:=\sum_{j=1}^{N}W_{ij}$ and $\omega^{(\zeta,i)}:=\sum_{j=1}^{N_\zeta}B_{ij}^{(\zeta)}$, respectively. The accuracy (e.g. in terms of polynomial exactness) of such quadrature rules depends on the methodology by which the discrete inner products are constructed (which we discuss in Appendix B) and not on the matrix $\underline{\underline{K}}$, and hence Theorem 4.3 implies that the conservation properties of the original DG method in (3.17) are preserved under the application of any VCJH filter given as in (3.34).

To establish discrete conservation for strong-form DG and FR schemes, the divergence theorem must be satisfied under the discrete inner products for vector fields with components belonging to the approximation space, as given by

$$\left\langle 1,\underline{\nabla}\cdot\underline{V}\right\rangle_{\underline{\underline{W}}}=\sum_{\zeta=1}^{N_{\text{fac}}}\left\langle 1,\underline{V}\cdot\hat{\underline{n}}^{(\zeta)}\right\rangle_{\underline{\underline{B}}^{(\zeta)}},\quad\forall\,\underline{V}\in\mathbb{P}_{\mathcal{N}}(\hat{\Omega})^d,\tag{4.16}$$

which is implied by Assumptions 2.2 and 3.1, where the latter is a sufficient but not necessary condition.[7] Such a result is demonstrated with the following theorem.

Theorem 4.4 *Theorem 4.3 is valid for the FR scheme in (3.23) and for the strong-form DG scheme in (3.33) with the additional requirement that the discrete divergence theorem in (4.16) be satisfied, as is the case for methods satisfying Assumption 3.1.*

Proof Pre-multiplying (3.23) by $\tilde{\underline{1}}^{\mathrm{T}}(\underline{\underline{M}}+\underline{\underline{K}})$, which for $\underline{\underline{K}}=\underline{\underline{0}}$ is equivalent to pre-multiplying (3.33) by $\tilde{\underline{1}}^{\mathrm{T}}$, we recover (4.8) by using the fact that $(\underline{\underline{M}}+\underline{\underline{K}})\underline{\underline{D}}^{(m)}=\underline{\underline{S}}^{(m)}$ holds for all $m\in\{1,\dots,d\}$ under Assumption 3.4 and noting that the relation

$$\sum_{m=1}^{d}\tilde{\underline{1}}^{\mathrm{T}}\underline{\underline{S}}^{(m)}=\sum_{\zeta=1}^{N_{\text{fac}}}\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\tilde{\underline{1}}^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\tag{4.17}$$

may obtained by expanding (4.16) in terms of $\mathcal{B}$, or, if Assumption 3.1 is satisfied, by pre-multiplying the SBP property in (3.15) by $\tilde{\underline{1}}^{\mathrm{T}}$ and summing over all $m\in\{1,\dots,d\}$. The remainder of the proof is therefore identical to that of Theorem 4.3. □

[7] Considering, for example, the quadrature-based discrete inner products in Appendix B.1 and the total-degree polynomial space $\mathbb{P}_p(\hat{\Omega})$, volume quadrature rules of total degree $p-1$ or greater and facet quadrature rules of total degree p or greater are generally insufficient for Assumption 3.1 to hold, but nonetheless satisfy (4.16).

Remark 4.4 Defining the quadrature-based functional approximating the integral of $V : \Omega \to \mathbb{R}$ over its entire domain as $I^h[V] := \sum_{\kappa=1}^{N_{\mathrm{el}}} \langle 1, V \circ \underline{X}^{(\kappa)} \rangle_{\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}}$, we see that

$$\frac{dI^h[U_e^h(\cdot,t)]}{dt} = 0, \quad \forall\, t \in (0,T), \tag{4.18}$$

holds for all $e \in \{1,\dots,N_{\mathrm{eq}}\}$ as a consequence of Theorem 4.3 (for the weak form) or Theorem 4.4 (for the strong form) in the case of periodic boundary conditions (i.e. $\mathcal{F}_{\partial\Omega} = \emptyset$), under the appropriate assumptions. Such a property then provides a useful numerical test for global conservation, which will be used in §5.

Remark 4.5 Following the discussion in Remark 3.6, we note that the proofs of conservation commonly presented in the FR literature (for example, those in [8, §3] and [11, §4.2]) rely explicitly on properties of the correction functions $\underline{G}^{(\zeta,j)}$, in particular, the second condition in (2.14), in order to show that the normal trace of the corrected flux in physical space is continuous on each shared facet. Our proofs, in contrast, rely only on well-defined algebraic properties of the discretization in (3.23), wherein the influence of the correction fields is contained within the lifting matrices.

4.3 Energy stability

While the theory in §4.1 and §4.2 applies to DG and FR methods for systems in the form of (2.1) on general curvilinear elements, the analysis in this section is restricted to the particular case of the constant-coefficient linear advection equation in (2.2) on meshes consisting solely of affinely mapped polytopes. We therefore make the following assumption.

Assumption 4.2 *The mapping $\underline{X}^{(\kappa)}$ is affine for all $\kappa \in \{1,\dots,N_{\mathrm{el}}\}$, and hence we may abuse notation to treat $\underline{\nabla}\underline{X}^{(\kappa)}$, $J^{(\kappa)}$, $J^{(\kappa,\zeta)}$, and $\underline{n}^{(\kappa,\zeta)}$ as constants.*

The discrete energy functional employed for the stability analysis in this section is introduced with the following lemma, where we note that the equation index is suppressed as we are considering a scalar conservation law.

Lemma 4.1 *Under Assumptions 2.1, 2.2, 3.4, and 4.2, the energy functional*

$$\begin{aligned} E^h[V] := \frac{1}{2}\sum_{\kappa=1}^{N_{\mathrm{el}}} \Big\langle V \circ \underline{X}^{(\kappa)}, V \circ \underline{X}^{(\kappa)} \Big\rangle_{\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}} \\ + \frac{1}{2}\sum_{\kappa=1}^{N_{\mathrm{el}}} \sum_{\underline{\alpha}\in\mathcal{M}} \frac{c^{(\underline{\alpha})}}{|\hat{\Omega}|} \Big\langle \partial^{(\underline{\alpha})}\big(V \circ \underline{X}^{(\kappa)}\big), \partial^{(\underline{\alpha})}\big(V \circ \underline{X}^{(\kappa)}\big) \Big\rangle_{\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}} \end{aligned} \tag{4.19}$$

may be expressed algebraically as a positive-definite quadratic form in terms of the global degrees of freedom $\underline{\tilde{u}}^h(t) := [(\underline{\tilde{u}}^{(h,1)}(t))^{\mathrm{T}}, \dots, (\underline{\tilde{u}}^{(h,N_{\mathrm{el}})}(t))^{\mathrm{T}}]^{\mathrm{T}}$ for the semi-discrete numerical solution $U^h(\cdot,t) \in \mathbb{P}_{\mathcal{N}}(\mathcal{T}_h)$ to a scalar PDE, as given by

$$E^h[U^h(\cdot,t)] = \frac{1}{2}\big(\underline{\tilde{u}}^h(t)\big)^{\mathrm{T}} \mathrm{diag}\Big(\underline{\underline{M}}^{(1)}\big(\underline{\underline{F}}^{(1)}\big)^{-1}, \dots, \underline{\underline{M}}^{(N_{\mathrm{el}})}\big(\underline{\underline{F}}^{(N_{\mathrm{el}})}\big)^{-1}\Big)\underline{\tilde{u}}^h(t). \tag{4.20}$$

Proof Noting that $J^{(\kappa)}$ is constant under Assumption 4.2, we may substitute $V(\underline{x}) = U^h(\underline{x}, t)$ into (4.19) and expand in terms of $\mathcal{B}$ in order to obtain

$$E^h[U^h(\cdot,t)] = \frac{1}{2}\sum_{\kappa=1}^{N_{\mathrm{el}}} J^{(\kappa)} \sum_{i=1}^{N^*}\sum_{j=1}^{N^*} \tilde{u}_i^{(h,\kappa)}(t)\Bigg(\underbrace{\left\langle \phi^{(i)}, \phi^{(j)} \right\rangle_{\underline{\underline{W}}}}_{=M_{ij}} + \underbrace{\sum_{\underline{\alpha}\in\mathcal{M}} \frac{c^{(\underline{\alpha})}}{|\hat{\Omega}|}\left\langle \partial^{(\underline{\alpha})}\phi^{(i)}, \partial^{(\underline{\alpha})}\phi^{(j)} \right\rangle_{\underline{\underline{W}}}}_{=K_{ij}}\Bigg)\tilde{u}_j^{(h,\kappa)}(t), \tag{4.21}$$

which recovers the expression in (4.20) since we see from (3.36) that $\underline{\underline{M}}^{(\kappa)}(\underline{\underline{F}}^{(\kappa)})^{-1} = J^{(\kappa)}(\underline{\underline{M}} + \underline{\underline{K}})$, which is SPD under Assumptions 2.1 and 3.4. This implies that the block-diagonal matrix on the right-hand side of (4.20) is also SPD and therefore defines a positive-definite quadratic form. □

Remark 4.6 Since the expansion of $U^h(\cdot,t) \circ \underline{X}^{(\kappa)}$ terms of the basis $\mathcal{B}$ defines an isomorphism between the vector spaces $\mathbb{P}_{\mathcal{N}}(\mathcal{T}_h)$ and $\mathbb{R}^{N_{\mathrm{el}}\cdot N^*}$, the positive-definiteness of the matrix in (4.20) implies that the energy functional in (4.19) induces a norm for the global approximation space. When (3.26) is satisfied, we accordingly recover

$$E^h[V] = \frac{1}{2}\sum_{\kappa=1}^{N_{\mathrm{el}}} \int_{\hat{\Omega}} \left(\left(V(\underline{X}^{(\kappa)}(\underline{\hat{x}}))\right)^2 + \sum_{\underline{\alpha}\in\mathcal{M}} \frac{c^{(\underline{\alpha})}}{|\hat{\Omega}|}\left(\partial^{(\underline{\alpha})} V(\underline{X}^{(\kappa)}(\underline{\hat{x}}))\right)^2 \right) J^{(\kappa)}(\underline{\hat{x}})\, d\underline{\hat{x}}, \tag{4.22}$$

which, under Assumption 3.4, is one-half of the square of the broken Sobolev-type norm introduced by Jameson [41] in the context of one-dimensional spectral-difference methods and adopted in several stability proofs for FR schemes (e.g. those [10, §3], [11, §5], and [12, §3]). The existing stability theory for VCJH schemes is therefore recovered within the present framework as a consequence of Theorem 3.1, in which we present an algebraic formulation of the family of correction fields associated with energy functionals in the form of (4.22).

In order for advection problems on bounded domains to be well posed in the continuous case, boundary conditions must be prescribed on portions of $\partial\Omega$ such that only incoming waves are specified. In the discrete case, we follow a similar approach and impose boundary and interface conditions weakly in a manner consistent with the physics of the problem through an appropriate choice of numerical flux. This is analogous to the use of simultaneous approximation terms (SATs), which are penalties added to strong-form discretizations at boundaries and interfaces in order to obtain stable approximations of initial-boundary-value problems with SBP operators,[8] and were introduced in the finite-difference context by Carpenter *et al.* [42], extending the penalty approach developed for spectral methods by Funaro and Gottlieb [43].

[8] Within the present framework, the third term of (3.23) may be viewed as a SAT, and serves the same purpose in the analysis of energy stability.

Assumption 4.3 *For the constant-coefficient linear advection equation in (2.2), the (non-periodic) boundary $\partial\Omega$ with an outward unit normal vector denoted by $\underline{n} : \partial\Omega \to \mathbb{S}^{d-1}$ may be partitioned into disjoint subsets $\partial\Omega^- := \{\underline{x} \in \partial\Omega : \underline{a}\cdot\underline{n}(\underline{x}) \leq 0\}$ and $\partial\Omega^+ := \partial\Omega \setminus \partial\Omega^-$, the former on which we prescribe $U(\underline{x}, t) = B(\underline{x}, t)$. The set $\mathcal{F}_{\partial\Omega}$ is then further partitioned into the disjoint subsets $\mathcal{F}_{\partial\Omega^-} := \{(\kappa, \zeta) \in \mathcal{F}_{\partial\Omega} : \Gamma^{(\kappa,\zeta)} \subseteq \partial\Omega^-\}$ and $\mathcal{F}_{\partial\Omega^+} := \mathcal{F}_{\partial\Omega} \setminus \mathcal{F}_{\partial\Omega^-}$, and the numerical flux takes the form*

$$F^*(U^-, U^+, \underline{n}) := \frac{1}{2}\sum_{m=1}^{d} n_m \big(F^{(m)}(U^-) + F^{(m)}(U^+)\big) - \frac{\lambda|\underline{a}\cdot\underline{n}|}{2}(U^+ - U^-) \quad (4.23)$$

for $U^-, U^+ \in \mathbb{R}$ and $\underline{n} \in \mathbb{S}^{d-1}$, where the parameter $\lambda \in \mathbb{R}_0^+$ assumes the same value on either side of an interior interface, and is taken to be unity for any facet on $\partial\Omega$.

Remark 4.7 The numerical flux in (4.23) satisfies the conservation property assumed in Theorem 4.3, recovering a fully upwind flux for $\lambda = 1$ and a central flux for $\lambda = 0$, which correspond to upwind and symmetric SATs, respectively, in SBP terminology.

We are now equipped to present the following generalized energy stability result for DG and FR schemes in strong and weak form, in which (as discussed, for example, in Gustafsson *et al.* [44, Ch. 11]) we show that the energy functional defined in (4.19) may only increase in time due to contributions arising from the inflow boundary.

Theorem 4.5 *Under Assumptions 2.1, 2.2, 3.1, 3.2, 3.4, 4.1, 4.2, and 4.3, the DG and FR methods in (3.17) and (3.23), respectively, are energy stable for the constant-coefficient linear advection equation in (2.2), satisfying a bound on the rate of energy growth given for all $t \in (0, T)$ by*

$$\frac{dE^h[U^h(\cdot,t)]}{dt} \leq \sum_{(\kappa,\zeta)\in\mathcal{F}_{\partial\Omega^-}} \frac{|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\Big\langle B(\cdot,t)\circ\underline{X}^{(\kappa)}, B(\cdot,t)\circ\underline{X}^{(\kappa)}\Big\rangle_{\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}}. \quad (4.24)$$

Proof Energy stability will be proven for the strong-form FR method in (3.23), which is equivalent to the filtered weak-form DG scheme in (4.2) under the present assumptions as a result of Theorem 4.2, recovering the standard DG method in (3.17) with $\underline{\underline{K}} = \underline{\underline{0}}$. Beginning with a single element and pre-multiplying (3.23) by $(\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}}(\underline{\underline{M}} + \underline{\underline{K}})$, we may use (3.36) and apply the chain rule in time to obtain

$$\begin{aligned}
&\frac{d}{dt}\left(\frac{1}{2}\big(\tilde{\underline{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\underline{\underline{M}}^{(\kappa)}\big(\underline{\underline{F}}^{(\kappa)}\big)^{-1}\tilde{\underline{u}}^{(h,\kappa)}(t)\right) = \\
&\qquad \underbrace{-\sum_{m=1}^{d}\big(\tilde{\underline{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{M}} + \underline{\underline{K}}\big)\underline{\underline{D}}^{(m)}\tilde{\underline{f}}^{(h,\kappa,m)}(t)}_{=:\,\Sigma_{\mathrm{I}}^{(\kappa)}(t)} \\
&\quad + \sum_{\zeta=1}^{N_{\mathrm{fac}}}\underbrace{\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\big(\tilde{\underline{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{f}}^{(h,\kappa,m)}(t)}_{=:\,\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)} \\
&\qquad \underbrace{-\sum_{\zeta=1}^{N_{\mathrm{fac}}}\big(\tilde{\underline{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta)}(t)}_{=:\,\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)},
\end{aligned} \quad (4.25)$$

where the contributions $\Sigma_{\mathrm{I}}^{(\kappa)}(t)$, $\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)$, and $\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)$ are defined as above. Noting that the SBP property in (3.15) and the relation $(\underline{\underline{M}}+\underline{\underline{K}})\underline{\underline{D}}^{(m)}=\underline{\underline{S}}^{(m)}$ hold for all $m\in\{1,\dots d\}$ under Assumptions 3.1 and 3.4, respectively, we obtain

$$\big(\underline{\underline{M}}+\underline{\underline{K}}\big)\underline{\underline{D}}^{(m)}=\frac{1}{2}\sum_{\zeta=1}^{N_{\mathrm{fac}}}\hat{n}_m^{(\zeta)}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}+\frac{1}{2}\Big(\underline{\underline{S}}^{(m)}-\big(\underline{\underline{S}}^{(m)}\big)^{\mathrm{T}}\Big) \tag{4.26}$$

by splitting $\underline{\underline{S}}^{(m)}$ into its symmetric and skew-symmetric parts. Using the skew-symmetry of the second term on the right-hand side of (4.26), we then have

$$\begin{aligned}&\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{M}}+\underline{\underline{K}}\big)\underline{\underline{D}}^{(m)}\underline{\tilde{u}}^{(h,\kappa)}(t)\\&\quad=\frac{1}{2}\sum_{\zeta=1}^{N_{\mathrm{fac}}}\hat{n}_m^{(\zeta)}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t).\end{aligned} \tag{4.27}$$

Since the mapping $\underline{X}^{(\kappa)}$ is affine and the flux is linear with constant coefficients, we may substitute $F^{(n)}(U^{(h,\kappa)}(\underline{\hat{x}},t))=a_nU^{(h,\kappa)}(\underline{\hat{x}},t)$ into (2.10) and use the relation $\underline{\underline{P}}\underline{\underline{V}}=\underline{\underline{I}}$ in order to express the first two terms on the right-hand side of (4.25) as

$$\Sigma_{\mathrm{I}}^{(\kappa)}(t)=\sum_{m=1}^{d}\sum_{n=1}^{d}a_nJ^{(\kappa)}\big[(\nabla\underline{X}^{(\kappa)})^{-1}\big]_{mn}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{M}}+\underline{\underline{K}}\big)\underline{\underline{D}}^{(m)}\underline{\tilde{u}}^{(h,\kappa)}(t) \tag{4.28}$$

and

$$\begin{aligned}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)&=\sum_{m=1}^{d}\sum_{n=1}^{d}a_nJ^{(\kappa)}\big[(\nabla\underline{X}^{(\kappa)})^{-1}\big]_{mn}\hat{n}_m^{(\zeta)}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)\\&=\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{J}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t),\end{aligned} \tag{4.29}$$

where the second equality in (4.29) results from (2.5). Applying (4.27) to (4.28) and using the first equality in (4.29) then results in $\Sigma_{\mathrm{I}}^{(\kappa)}(t)=\sum_{\zeta=1}^{N_{\mathrm{fac}}}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)/2$. The local energy balance in (4.25) may therefore be expressed as

$$\frac{d}{dt}\left(\frac{1}{2}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\underline{\underline{M}}^{(\kappa)}\big(\underline{\underline{F}}^{(\kappa)}\big)^{-1}\underline{\tilde{u}}^{(h,\kappa)}(t)\right)=\sum_{\zeta=1}^{N_{\mathrm{fac}}}\left(\frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)-\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)\right), \tag{4.30}$$

where we note that the right-hand side contains only facet contributions. Summing (4.30) over all elements and splitting the net contribution $\Sigma_{\mathrm{net}}^{(\kappa,\zeta)}(t):=(\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)+\Sigma_{\mathrm{II}}^{(\nu,\eta)}(t))/2-(\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)+\Sigma_{\mathrm{III}}^{(\nu,\eta)}(t))$ due to each shared interface equally between the index pairs $(\kappa,\zeta),(\nu,\eta)\in\mathcal{F}_\Omega$ corresponding to the coincident or periodically connected facets $\Gamma^{(\kappa,\zeta)}\subset\partial\Omega^{(\kappa)}$ and $\Gamma^{(\nu,\eta)}\subset\partial\Omega^{(\nu)}$ results in the global balance

$$\frac{dE^h[U^h(\cdot,t)]}{dt}=\sum_{(\kappa,\zeta)\in\mathcal{F}_\Omega}\frac{1}{2}\Sigma_{\mathrm{net}}^{(\kappa,\zeta)}(t)+\sum_{(\kappa,\zeta)\in\mathcal{F}_{\partial\Omega}}\left(\frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)-\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)\right), \tag{4.31}$$

where the energy functional appears on left-hand side as a consequence of Lemma 4.1. As we will now demonstrate, the contributions to the right-hand side of (4.31) arising from interior or periodic interfaces and those arising from facets on the inflow and outflow portions of the boundary may be analyzed separately in order to establish an upper bound for the time rate of change in energy.

Interior or periodic interfaces

Similarly to the proof of Theorem 4.3, it follows from Assumption 4.1 that for any interior or periodic interface corresponding to $(\kappa,\zeta)\in\mathcal{F}_\Omega$, the contributions arising from a numerical flux function taking the form of (4.23) may be expressed as

$$\begin{aligned}\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)=\frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)+\frac{\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}}{2}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t)\\-\frac{\lambda|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t)\\+\frac{\lambda|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)\end{aligned}\tag{4.32}$$

and

$$\begin{aligned}\Sigma_{\mathrm{III}}^{(\nu,\eta)}(t)=\frac{1}{2}\Sigma_{\mathrm{II}}^{(\nu,\eta)}(t)\\-\frac{\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}}{2}\big(\underline{\tilde{u}}^{(h,\nu)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)\\-\frac{\lambda|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{\tilde{u}}^{(h,\nu)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)\\+\frac{\lambda|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{\tilde{u}}^{(h,\nu)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t).\end{aligned}\tag{4.33}$$

Noting that the first terms of (4.32) and (4.33) cancel out the contributions of $\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)$ and $\Sigma_{\mathrm{II}}^{(\nu,\eta)}(t)$ to $\Sigma_{\mathrm{net}}^{(\kappa,\zeta)}(t)$, respectively, and that the matrix $\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}$ is SPSD under Assumptions 2.1, 3.2, and 4.2, the net contribution to the energy balance in (4.31) corresponding to any $(\kappa,\zeta)\in\mathcal{F}_\Omega$ is then given by

$$\begin{aligned}\Sigma_{\mathrm{net}}^{(\kappa,\zeta)}(t)&=-\frac{\lambda|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\Big(\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)\\&\quad-2\big(\underline{\tilde{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t)\\&\quad+\big(\underline{\tilde{u}}^{(h,\nu)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t)\Big)\\&=-\frac{\lambda|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{d}^{(\kappa,\zeta)}(t)\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t),\end{aligned}\tag{4.34}$$

where we define $\underline{d}^{(\kappa,\zeta)}(t):=\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)-\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t)$. It therefore follows that $\Sigma_{\mathrm{net}}^{(\kappa,\zeta)}(t)\le 0$ if $\lambda\ge 0$, where for $\lambda=0$, the interior or periodic interfaces do not contribute to the energy balance, whereas for $\lambda>0$, the upwind bias of the numerical flux in (4.23) dissipates energy at a rate depending quadratically on the size of the inter-element jump. We have therefore demonstrated that the interior/periodic interface coupling procedure is energy stable, and now turn our attention to the contributions to (4.31) resulting from the inflow and outflow portions of the boundary.

Inflow boundary

For a given facet $\Gamma^{(\kappa,\zeta)} \subset \partial\Omega^-$, the contribution $\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)/2 - \Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)$ to the energy balance in (4.31) corresponding to $(\kappa,\zeta) \in \mathcal{F}_{\partial\Omega^-}$ may be obtained by expressing $\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)$ and $\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)$ as in (4.29) and (4.32), respectively, with $\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\tilde{\underline{u}}^{(h,\nu)}(t)$ replaced by $\underline{b}^{(\kappa,\zeta)}(t) := [B(\underline{X}^{(\kappa)}(\hat{\underline{x}}^{(\zeta,1)},t)),\ldots,B(\underline{X}^{(\kappa)}(\hat{\underline{x}}^{(\zeta,N_\zeta)},t))]^{\mathrm{T}}$. We may then take $\lambda = 1$ as per Assumption 4.3 and note that $\underline{a}\cdot\underline{n}^{(\kappa,\zeta)} = -|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|$, resulting in

$$\begin{aligned}
\frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t) - \Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t) &= \frac{|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\Big(2\big(\tilde{\underline{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t) \\
&\qquad - \big(\tilde{\underline{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t)\Big) \\
&= \frac{|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\Big(\big(\underline{b}^{(\kappa,\zeta)}(t)\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t) \\
&\qquad - \big(\underline{d}^{(\kappa,\zeta)}(t)\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t)\Big) \\
&\leq \frac{|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{b}^{(\kappa,\zeta)}(t)\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t),
\end{aligned} \tag{4.35}$$

where the second equality follows from completing the square and defining $\underline{d}^{(\kappa,\zeta)}(t) := \underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t) - \underline{b}^{(\kappa,\zeta)}(t)$ and the last line follows from the fact that $\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}$ is SPSD.

Outflow boundary

Using the fact that $\Gamma^{(\kappa,\zeta)} \subset \partial\Omega^+$ implies that $\underline{a}\cdot\underline{n}^{(\kappa,\zeta)} = |\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|$, it follows from taking $\lambda = 1$ in (4.32) that $\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t) = \Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)$. The resulting contribution to the second sum on the right-hand side of (4.31) for any $(\kappa,\zeta) \in \mathcal{F}_{\partial\Omega^+}$ is therefore simply $-\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)/2$, which is non-positive as a result of (4.29) and the positive-semidefiniteness of $\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}$, and thus does not contribute to growth in the solution energy. The final energy balance is then given by

$$\begin{aligned}
\frac{dE^h[U^h(\cdot,t)]}{dt} &= \sum_{(\kappa,\zeta)\in\mathcal{F}_{\partial\Omega^-}} \frac{|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{b}^{(\kappa,\zeta)}(t)\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t) \\
&\quad - \sum_{(\kappa,\zeta)\in\mathcal{F}_{\partial\Omega^+}} \frac{|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\tilde{\underline{u}}^{(h,\kappa)}(t)\big)^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t) \\
&\quad - \sum_{(\kappa,\zeta)\in\mathcal{F}_{\partial\Omega^-}} \frac{|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{2}\big(\underline{d}^{(\kappa,\zeta)}(t)\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t) \\
&\quad - \sum_{(\kappa,\zeta)\in\mathcal{F}_{\Omega}} \frac{\lambda|\underline{a}\cdot\underline{n}^{(\kappa,\zeta)}|}{4}\big(\underline{d}^{(\kappa,\zeta)}(t)\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t),
\end{aligned} \tag{4.36}$$

where the first term on the right-hand side is recognized as the right-hand side of (4.24), corresponding to the energy entering the domain through the inflow boundary, while the second corresponds to the energy exiting the domain through the outflow boundary. The third and fourth terms correspond to the energy dissipation resulting from the facet jumps, which arise due to the discontinuous nature of the approximation

space and the weak imposition of the boundary conditions. Since all but the first of such contributions to the energy balance are non-positive, we obtain the inequality in (4.24), thereby proving that the discretization is energy stable. □

Remark 4.8 For homogeneous (i.e. $B(\underline{x}, t) = 0$ for all $\underline{x} \in \partial\Omega^-$) or periodic boundary conditions, the energy balance in (4.24) reduces to

$$\frac{dE^h[U^h(\cdot,t)]}{dt} \leq 0, \quad \forall\, t \in (0,T), \tag{4.37}$$

implying that $E^h[U^h(\cdot,t)]$ is invariant for periodic problems when a non-dissipative (i.e. central) numerical flux is chosen for all interfaces, as given by (4.23) with $\lambda = 0$.

5 Numerical experiments

5.1 Design choices and implementation

The methods described in §3 are implemented within a free and open-source Python code[9] offering substantial flexibility in the design choices afforded for DG and FR methods, which are detailed in this section.

Reference element and approximation space

The numerical experiments in this work involve formulations in two space dimensions on a triangular reference element given by

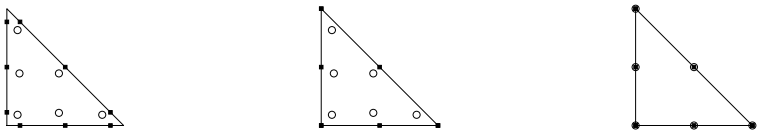
$$\hat{\Omega} := \left\{ \underline{\hat{x}} \in \mathbb{R}^2 : \hat{x}_1 \geq -1;\ \hat{x}_2 \geq -1;\ \hat{x}_1 + \hat{x}_2 \leq 0 \right\}, \tag{5.1}$$

where $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ is taken to be the total-degree polynomial space $\mathbb{P}_p(\hat{\Omega})$, which results from (2.6) with the choice of $\mathcal{N} := \{\underline{\alpha} \in \mathbb{N}_0^2 : \alpha_1 + \alpha_2 \leq p\}$. Such a space is of dimension $N^* = (p+1)(p+2)/2$, and we report results for values of $p \in \{2, 3, 4\}$.

Discrete inner products and nodal sets

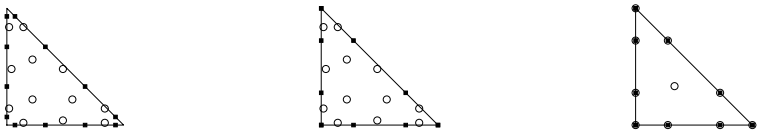
Three options for discrete inner products are employed for the numerical experiments, which are referred to as *Quadrature I*, *Quadrature II*, and *Collocation* schemes. The first two make use of the quadrature-based approach described in Appendix B.1, while the third employs the collocation-based approach in Appendix B.2. For the quadrature-based schemes, all volume terms are computed using positive quadrature rules of degree $2p$ due to Xiao and Gimbutas [45]. To compute the facet terms, we use LG quadrature rules with $p+1$ nodes on each facet (i.e. degree $2p+1$) for the Quadrature I schemes and LGL quadrature rules with $p+1$ nodes on each facet (i.e. degree $2p-1$) for the Quadrature II schemes. Although both approaches result in Assumption 3.2 being satisfied (where, as discussed in Remark B.1, the rank of $\underline{\underline{V}}$ was verified numerically), the insufficient accuracy of the facet integration for the Quadrature II schemes precludes the use of Lemma B.1. As a result, Assumption 3.1 is satisfied for the Quadrature I schemes, but not for the Quadrature II schemes,

[9] Available at https://github.com/tristanmontoya/GHOST.



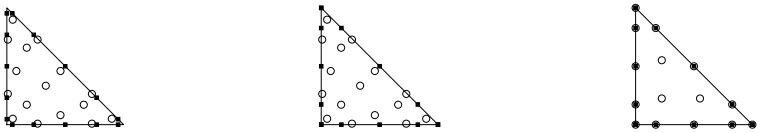
(a) Quadrature I (b) Quadrature II (c) Collocation

Fig. 1: Nodal sets for $p = 2$



(a) Quadrature I (b) Quadrature II (c) Collocation

Fig. 2: Nodal sets for $p = 3$



(a) Quadrature I (b) Quadrature II (c) Collocation

Fig. 3: Nodal sets for $p = 4$

leading to the loss of the SBP property in the latter case. The Collocation schemes interpolate the volume terms on the "warp & blend" nodes introduced by Warburton [46], which form unisolvent sets for the corresponding polynomial spaces $\mathbb{P}_p(\hat{\Omega})$, and include $p+1$ LGL nodes on each facet, which are used to interpolate the facet terms. Therefore, as discussed in Appendix B.2, the Collocation schemes satisfy Assumptions 3.1 and 3.2 by construction. The nodal sets corresponding to each choice of discrete inner product for polynomial degrees 2, 3, and 4 are shown in Figures 1, 2, and 3, respectively, where the symbols ○ and ■ represent nodes in $\mathcal{S}$ and $\mathcal{S}^{(\zeta)}$, respectively.

Polynomial basis

For the quadrature-based schemes, we take $\mathcal{B}$ to be an orthonormal modal basis defined as in (A.3), while for the collocation-based schemes, nodal bases given as in (A.6) are used with $\tilde{\mathcal{S}} = \mathcal{S}$, in this case corresponding to the "warp & blend" nodes in [46], which also satisfy $\mathcal{S}^{(\zeta)} \subset \tilde{\mathcal{S}}$ for all $\zeta \in \{1, \ldots, N_{\text{fac}}\}$.

Semi-discrete residual formulation

We employ the strong formulation in (3.23) as well as the weak formulation in (4.2), in the case of $\underline{\underline{K}} = \underline{\underline{0}}$ recovering the standard unfiltered strong-form and weak-form DG schemes in (3.33) and (3.17), respectively.

Lifting and filter matrices

The lifting and filter matrices are given as in (3.28) and (3.34), respectively, with $\underline{\underline{K}} = \underline{\underline{K}}^{\mathrm{2D}}$ as defined in (C.1). The coefficient $c \in \mathbb{R}$ scaling the matrix $\underline{\underline{K}}$ is chosen either as $c = c_{\mathrm{DG}} = 0$, recovering a DG scheme in strong or weak form, or as $c = c_+$, corresponding to the VCJH scheme found in [11, §6] to allow for the largest stable time step for a given polynomial degree when solving the linear advection equation.[10] The numerical experiments in [11, §7] employ values of c_+ given by 4.3×10^{-2}, 6.0×10^{-4}, and 5.6×10^{-6} for polynomial degrees 2, 3, and 4, respectively; we define $\underline{\underline{K}}$ for $c = c_+$ using the same values, and, as a result of the conditions of Theorem 3.1 (in addition to Assumption 3.4) being satisfied for all discretizations employed in this section, we therefore recover identical correction fields to those in their work.

Numerical flux function

The numerical flux function for the linear advection equation is given as in (4.23), where we report results for $\lambda \in \{0, 1\}$, corresponding to central and upwind fluxes. In the case of the Euler equations, we employ Roe's approximate Riemann solver [47].

Mesh and coordinate transformation

As the test problems considered in this work are posed on the square domain $\Omega := (0, L)^2$ with $L \in \mathbb{R}^+$, we begin with a regular Cartesian grid consisting of $M \in \mathbb{N}$ equal intervals in each direction, subdividing each quadrilateral into two right-angled triangles in order to obtain a total of $N_{\mathrm{el}} = 2M^2$ elements. Using an affine transformation $\underline{X}^{(\kappa)}_{\mathrm{affine}} \in \mathbb{P}_1(\hat{\Omega})^2$ to map the interpolation nodes $\mathcal{S}_{\mathrm{map}} := \{\underline{\hat{x}}^{(1)}_{\mathrm{map}}, \ldots, \underline{\hat{x}}^{(N_{\mathrm{map}})}_{\mathrm{map}}\} \subset \hat{\Omega}$ associated with a Lagrange basis $\mathcal{B}_{\mathrm{map}} := \{\ell^{(1)}_{\mathrm{map}}, \ldots, \ell^{(N_{\mathrm{map}})}_{\mathrm{map}}\}$ of degree $p_{\mathrm{map}} \in \mathbb{N}$ onto each straight-sided element of the original mesh, we then interpolate the warping

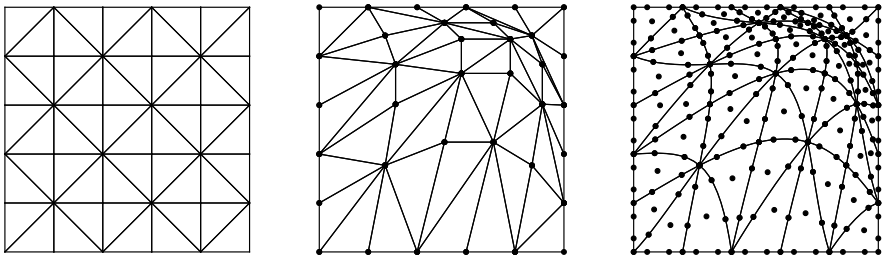
$$\underline{\Theta}(\underline{x}) := \begin{bmatrix} x_1 + \frac{1}{5} L \sin(\pi x_1 / L) \sin(\pi x_2 / L) \\ x_2 + \frac{1}{5} L \exp(1 - x_2/L) \sin(\pi x_1 / L) \sin(\pi x_2 / L) \end{bmatrix} \tag{5.2}$$

considered by Del Rey Fernández *et al.* [48], resulting in a mapping given by

$$\underline{X}^{(\kappa)}(\underline{\hat{x}}) := \sum_{i=1}^{N_{\mathrm{map}}} \underline{\Theta}\big(\underline{X}^{(\kappa)}_{\mathrm{affine}}(\underline{\hat{x}}^{(i)}_{\mathrm{map}})\big) \ell^{(i)}_{\mathrm{map}}(\underline{\hat{x}}). \tag{5.3}$$

In this work, the interpolation nodes in $\mathcal{S}_{\mathrm{map}}$ are taken to be the "warp and blend" nodes in [46], which for $p = p_{\mathrm{map}}$ coincide with those pictured for the Collocation schemes in Figures 1, 2, and 3. Examples of meshes obtained through the above warping procedure with $M = 5$ are shown in Figure 4, where the symbol $\bullet$ is used to represent the image of each node in $\mathcal{S}_{\mathrm{map}}$ under the mapping given in (5.3).

[10] The latter choice is merely used as a reference value in order to demonstrate that the results in §4 hold for $\underline{\underline{K}} \neq \underline{\underline{0}}$; the optimality of such a choice in terms of time step size is highly dependent on the time-marching method, the mesh, and the particular problem being solved.



(a) Original mesh (b) $p_{\text{map}} = 1$ (c) $p_{\text{map}} = 3$

Fig. 4: Warping of a split Cartesian mesh through Lagrange interpolation

Temporal discretization

The standard explicit fourth-order Runge-Kutta method is used to advance the solution in time, where the time step is chosen as $\Delta t \approx C_t h/a$,[11] with a characteristic element size of $h = L/M$ and a characteristic wave speed $a \in \mathbb{R}^+$ taken in §5.2 to be the magnitude of the advection velocity and in §5.3 to be the magnitude of the background velocity for the isentropic vortex. Motivated by the Courant-Friedrichs-Lewy (CFL) condition described for the DG method by Cockburn and Shu [49, §2.2], we choose $C_t = \beta/(2p+1)$ with $\beta = 2.5 \times 10^{-3}$ in order to ensure that the error due to the temporal discretization is dominated by the error due to the spatial discretization.

Quantities of interest

In order to numerically verify the theoretical results established in §4, the following quantities of interest are evaluated for each scheme under consideration.

1. *Equivalence.* Denoting the numerical solutions at time $t = T$ obtained using the strong-form and weak-form implementations of a given method applied to a particular problem as $\underline{U}^h_{\text{strong}}(\cdot, T)$ and $\underline{U}^h_{\text{weak}}(\cdot, T)$, respectively, we evaluate the L^2 norm of their difference, which is given for $e \in \{1, \ldots, N_{\text{eq}}\}$ by

$$\left\| U^h_{e,\text{strong}}(\cdot,T) - U^h_{e,\text{weak}}(\cdot,T) \right\|_\Omega := \sqrt{\int_\Omega \left(U^h_{e,\text{strong}}(\underline{x},T) - U^h_{e,\text{weak}}(\underline{x},T) \right)^2 d\underline{x}}$$
$$= \sqrt{\sum_{\kappa=1}^{N_{\text{el}}} \int_{\hat{\Omega}} \left(U^{(h,\kappa)}_{e,\text{strong}}(\underline{\hat{x}},T) - U^{(h,\kappa)}_{e,\text{weak}}(\underline{\hat{x}},T) \right)^2 J^{(\kappa)}(\underline{\hat{x}})\, d\underline{\hat{x}}} \quad (5.4)$$

where the integrals on the reference element are evaluated exactly (up to machine precision) using a quadrature rule of degree $4p$ from [45]. Such a quantity is expected to be zero for the Quadrature I and Collocation schemes (but not necessarily for the Quadrature II schemes) as a result of Theorems 4.1 and 4.2.

[11] Specifically, we take $\Delta t = T/N_t$, where the total number of time steps is given by $N_t = \lfloor T/\Delta t^* \rfloor$ in terms of the target time step $\Delta t^* := C_t h/a$, with $\lfloor \cdot \rfloor$ denoting the floor operator.

2. *Conservation.* We compute $I^h[U_e^h(\cdot,T)] - I^h[U_e^h(\cdot,0)]$ for $e \in \{1,\ldots,N_{\text{eq}}\}$, corresponding to the net production of each conserved quantity, which, as discussed in Remark 4.4, is expected to be zero as a result of Theorems 4.3 and 4.4, the conditions of which are satisfied for all discretizations described in this section.
3. *Energy difference.* In the case of the linear advection equation, we also evaluate the energy difference $E^h[U^h(\cdot,T)] - E^h[U^h(\cdot,0)]$, which, as discussed in Remark 4.8, is ensured by Theorem 4.5 to be negative with $\lambda = 1$ and zero with $\lambda = 0$ for the Quadrature I and Collocation schemes (but not necessarily for the Quadrature II schemes) applied the periodic problem considered here.

Remark 5.1 In the results tabulated in §5.2 and §5.3, values corresponding to properties established theoretically in §4 that are violated (i.e. due to design choices leading to the stated assumptions not being satisfied) are shown in italic type. Furthermore, a horizontal line denotes a computation for which the scheme was found to be unstable.

Remark 5.2 All computations in this work are performed using double-precision arithmetic. Hence, due to accumulated roundoff error as well as the influence of the temporal discretization (which is not accounted for in the theoretical analysis), we consider a quantity to be zero if it is on the order of 10^{-12} or less. This is many orders of magnitude smaller than the solution error for the grids considered in this work, which are deliberately chosen to be coarse in order to unambiguously demonstrate that the properties proven in §4 hold discretely regardless of the grid resolution, rather than asymptotically.

5.2 Linear advection equation

We solve the constant-coefficient linear advection equation in (2.2) on a square domain of side length $L \in \mathbb{R}^+$, with periodic boundary conditions applied in both directions. The initial condition is given by $U^0(\underline{x}) := \sin(2\pi x_1/L)\sin(2\pi x_2/L)$ and the advection velocity is prescribed as $\underline{a} := a[\cos(\theta),\sin(\theta)]^{\mathrm{T}}$ with $a \in \mathbb{R}^+$ and $\theta \in [0,2\pi)$. The mesh is constructed as described in §5.1, where we use $p_{\text{map}} = 1$ such that Assumption 4.2 is satisfied, and the solution is advanced in time for one period, corresponding to $T = L/(a\max(|\cos(\theta)|,|\sin(\theta)|))$. For the numerical experiments in this section, the solver is run with parameter values of $L = 1$, $M = 8$, $a = \sqrt{2}$ and $\theta = \pi/4$.

The results for polynomial degrees 2, 3, and 4 are shown in Tables 1, 2, and 3, respectively. In all cases, the results for the Quadrature I and Collocation schemes are consistent with the theory, satisfying the equivalence, conservation, and stability (i.e. energy conservation for $\lambda = 0$ and energy dissipation for $\lambda = 1$) properties which were proven in §4. As discussed in §5.1, the Quadrature II schemes violate Assumption 3.1, and thus the theory in §4.1 and §4.3 does not apply; accordingly, the strong-form and weak-form numerical solutions are different, and both the strong and weak forms are found to be unstable when a central flux (i.e. $\lambda = 0$) is used. Although the Quadrature II schemes are seen to be stable in practice for an upwind flux (i.e. $\lambda = 1$), no *a priori* energy estimate exists due to the violation of the SBP property. The use of an upwind flux in such a case can therefore be interpreted as the addition of numerical dissipation to stabilize an unstable baseline scheme.

Table 1: Linear advection equation, $p = 2$

Scheme	c	λ	Equivalence	Conservation		Energy difference	
				Strong	Weak	Strong	Weak
Quadrature I	c_{DG}	0	8.937e-15	-4.710e-16	-5.130e-16	-1.665e-15	-1.610e-15
		1	1.257e-15	-1.126e-16	-2.004e-16	-5.847e-03	-5.847e-03
	c_+	0	6.789e-15	-5.609e-16	-7.048e-16	-7.772e-16	6.106e-16
		1	1.002e-15	-3.223e-16	-2.406e-16	-5.328e-02	-5.328e-02
Quadrature II	c_{DG}	0	—	—	—	—	—
		1	*1.741e-02*	-1.380e-15	4.121e-16	-5.308e-03	-4.728e-03
	c_+	0	—	—	—	—	—
		1	*1.949e-02*	-5.756e-16	-6.061e-17	-4.864e-02	-5.315e-02
Collocation	c_{DG}	0	4.130e-15	1.477e-16	-3.348e-16	-7.550e-15	-9.992e-15
		1	3.251e-15	-9.151e-17	-5.074e-16	-5.775e-03	-5.775e-03
	c_+	0	3.911e-15	-1.095e-16	-3.338e-16	1.776e-15	-1.221e-15
		1	2.994e-15	-7.058e-17	-2.102e-16	-5.156e-02	-5.156e-02

Table 2: Linear advection equation, $p = 3$

Scheme	c	λ	Equivalence	Conservation		Energy difference	
				Strong	Weak	Strong	Weak
Quadrature I	c_{DG}	0	2.518e-14	-1.212e-16	-1.637e-16	1.887e-15	6.106e-16
		1	2.639e-15	-1.269e-17	-1.054e-16	-1.942e-04	-1.942e-04
	c_+	0	1.463e-14	-1.366e-17	-2.243e-16	2.220e-15	7.772e-16
		1	2.339e-15	-6.904e-16	-7.605e-16	-2.315e-03	-2.315e-03
Quadrature II	c_{DG}	0	—	—	—	—	—
		1	*2.736e-03*	2.267e-16	-1.434e-16	-1.831e-04	-1.710e-04
	c_+	0	—	—	—	—	—
		1	*2.697e-03*	-6.112e-16	-7.589e-19	-2.255e-03	-2.282e-03
Collocation	c_{DG}	0	6.416e-15	-7.264e-18	-3.464e-16	-3.886e-16	-1.721e-15
		1	2.043e-15	-3.193e-16	-4.105e-16	-1.970e-04	-1.970e-04
	c_+	0	4.335e-15	2.244e-16	-4.142e-17	0.000e00*	1.665e-16
		1	1.209e-15	4.001e-17	-1.709e-16	-2.302e-03	-2.302e-03

* Indicates a difference of two floating-point quanitities with the same finite-precision representation

5.3 Euler equations

The propagation of an isentropic vortex is commonly used as a test case for numerical methods applied to nonlinear hyperbolic systems. Numerous versions of this problem have been posed in the literature, with the form considered here being a modification of that presented by Shu [50, §5.1] following the general formulation described by Spiegel *et al.* [51]. Considering the Euler equations in (2.3) and denoting the Mach number and direction of the background flow as $\mathrm{Ma}_\infty \in \mathbb{R}_0^+$ and $\theta \in [0, 2\pi)$, respectively, the initial velocity field is given for a vortex of strength $\varepsilon \in \mathbb{R}^+$ centred at $\underline{x}^0 \in \Omega$ as

$$\underline{V}^0(\underline{x}) := \mathrm{Ma}_\infty \left(\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} + \varepsilon \exp\left(1 - \|\underline{x} - \underline{x}^0\|^2\right) \begin{bmatrix} -(x_2 - x_2^0) \\ x_1 - x_1^0 \end{bmatrix} \right), \tag{5.5}$$

Table 3: Linear advection equation, $p = 4$

Scheme	c	λ	Equivalence	Conservation		Energy Difference	
				Strong	Weak	Strong	Weak
Quadrature I	c_{DG}	0	1.281e-14	4.943e-16	-2.732e-16	-1.776e-15	-7.216e-16
		1	2.301e-15	4.210e-16	1.444e-16	-4.326e-06	-4.326e-06
	c_+	0	8.634e-15	-1.134e-15	-1.394e-15	6.106e-16	-4.996e-16
		1	2.400e-15	-3.072e-16	-5.458e-16	-5.770e-05	-5.770e-05
Quadrature II	c_{DG}	0	—	—	—	—	—
		1	*4.366e-04*	-4.139e-16	-4.358e-16	-4.219e-06	-3.999e-06
	c_+	0	—	—	—	—	—
		1	*3.523e-04*	-6.670e-16	-8.674e-19	-5.974e-05	-5.569e-05
Collocation	c_{DG}	0	1.018e-14	-8.815e-17	9.780e-17	-2.720e-15	2.054e-15
		1	5.212e-15	-1.828e-16	-3.014e-17	-4.360e-06	-4.360e-06
	c_+	0	8.760e-15	-2.413e-16	-6.336e-16	-2.165e-15	3.886e-15
		1	6.157e-15	1.253e-16	-3.226e-16	-5.702e-05	-5.702e-05

the initial temperature field is prescribed as

$$T^0(\underline{x}) := 1 - \frac{(\gamma-1)\varepsilon^2 \mathrm{Ma}_\infty^2}{2} \exp\bigl(1 - \|\underline{x} - \underline{x}^0\|^2\bigr), \tag{5.6}$$

and constant entropy is enforced throughout the spatial domain. Using (5.5) and (5.6), as well as the equation of state and isentropic relations for an ideal gas with constant specific heat, the initial condition is then given by

$$\underline{U}^0(\underline{x}) := \begin{bmatrix} (T^0(\underline{x}))^{1/(\gamma-1)} \\ (T^0(\underline{x}))^{1/(\gamma-1)} \underline{V}^0(\underline{x}) \\ \frac{1}{\gamma-1}(T^0(\underline{x}))^{\gamma/(\gamma-1)} + \frac{1}{2}(T^0(\underline{x}))^{1/(\gamma-1)} \|\underline{V}^0(\underline{x})\|^2 \end{bmatrix}. \tag{5.7}$$

Similarly to the advection problem described in §5.2, such an initial condition results in an exact solution on an infinite domain given by the advection of the vortex at the constant background velocity. For the numerical experiments in this section, periodic boundary conditions are applied in both directions on a square domain of side length L, where the meshes are constructed as in §5.1 with $p_{\mathrm{map}} = p$, and the solver is run for one period such that $T = L/(\mathrm{Ma}_\infty \max(|\cos(\theta)|, |\sin(\theta)|))$, with parameter values of $L = 10$, $M = 16$, $\gamma = 1.4$, $\mathrm{Ma}_\infty = 0.4$, $\theta = \pi/4$, and $\varepsilon = 1$.

The results for polynomial degrees 2, 3, and 4 are shown in Tables 4, 5, and 6, respectively, where we see that the strong and weak forms are equivalent to one another for the Quadrature I and Collocation schemes, and the discrete integrals of all four solution variables are invariant for all three approaches, as expected from the analysis in §4.1 and §4.2. For the Quadrature II schemes, the strong and weak forms again result in different numerical solutions due to Assumption 3.1 not being satisfied. We do not investigate energy balances for this problem, as the theory in §4.3 is not applicable as a result of the flux being nonlinear and the mapping not being affine.[12] Hence, the numerical results in this section for the Euler equations, as with those in §5.2 for the linear advection equation, are consistent with the theory in §4.

[12] The focus of this paper is on standard DG and FR methods in conservation form, which are, in general, not provably stable for nonlinear problems or curvilinear coordinates. While outside the scope of this manuscript, entropy-stable or split-form schemes would be required in order to ensure that auxiliary balances analogous to (4.24) are satisfied for such problems.

Table 4: Euler equations, $p = 2$

Scheme	c	Equation	Equivalence	Conservation	
				Strong	Weak
Quadrature I	c_{DG}	ρ	6.471e-14	3.695e-13	1.421e-14
		ρV_1	1.470e-13	2.593e-13	-7.567e-13
		ρV_2	1.698e-13	1.030e-12	1.148e-12
		E	1.606e-13	1.023e-12	3.979e-13
	c_+	ρ	5.382e-14	3.979e-13	8.527e-14
		ρV_1	1.398e-13	2.984e-13	-8.100e-13
		ρV_2	1.589e-13	1.020e-12	1.066e-12
		E	1.378e-13	9.663e-13	-2.842e-14
Quadrature II	c_{DG}	ρ	*3.384e-02*	7.105e-14	-2.842e-14
		ρV_1	*6.001e-02*	-8.846e-13	-2.132e-12
		ρV_2	*6.388e-02*	1.148e-12	1.759e-12
		E	*7.306e-02*	0.000e00	-6.253e-13
	c_+	ρ	*2.529e-02*	0.000e00	-1.563e-13
		ρV_1	*3.513e-02*	-8.811e-13	-2.316e-12
		ρV_2	*3.853e-02*	1.123e-12	1.808e-12
		E	*5.652e-02*	1.137e-13	-6.253e-13
Collocation	c_{DG}	ρ	1.284e-13	3.553e-13	-3.126e-13
		ρV_1	3.083e-13	5.222e-13	-1.616e-12
		ρV_2	3.230e-13	1.343e-12	1.034e-12
		E	3.352e-13	1.393e-12	-4.547e-13
	c_+	ρ	1.501e-13	5.400e-13	-3.695e-13
		ρV_1	5.104e-13	5.365e-13	-3.244e-12
		ρV_2	5.586e-13	1.329e-12	2.135e-12
		E	4.469e-13	1.592e-12	-8.811e-13

6 Conclusions

We have proposed a unifying framework enabling the algebraic formulation of a broad class of high-order DG and FR methods applied to systems of conservation laws, facilitating the unified theoretical analysis of such schemes through matrix-based techniques. The role of the multidimensional SBP property is examined in detail for methods within the proposed framework, revealing new insights and generalizing existing results pertaining to the equivalence, conservation, and stability properties of DG and FR methods. Potential avenues for future research include the application of similar unifying principles to nonlinearly stable or split-form schemes and problems with diffusive terms as well as the comparative evaluation of methods within and outside of the present framework. Specifically, it would be interesting to examine the difference in computational expense between strong-form and weak-form implementations of methods which we have shown in this paper to be mathematically equivalent, and to compare the DG and FR methods considered in this paper with collocated multidimensional SBP schemes without an analytical basis (e.g. those described in [21] and [52]). Additionally, the theory could be extended to enable proofs of convergence, establishing *a priori* error estimates for linear problems with smooth solutions as consequences of the SBP property, which was discussed by Svärd and Nordström [53] in the context of finite-difference schemes in one space dimension.

Acknowledgements The authors are grateful to Gianmarco Mengaldo, Masayuki Yano, David Del Rey Fernández, and Alex Bercik for their feedback. The polynomial basis functions, interpolation nodes, and quadrature rules (with the exception of the LGL quadrature) employed

Table 5: Euler equations, $p = 3$

Scheme	c	Equation	Equivalence	Conservation	
				Strong	Weak
Quadrature I	c_{DG}	ρ	7.946e-14	4.263e-13	9.948e-14
		ρV_1	1.920e-13	-7.248e-13	-8.349e-13
		ρV_2	1.637e-13	2.373e-12	1.165e-12
		E	1.803e-13	1.535e-12	4.547e-13
	c_+	ρ	7.618e-14	5.116e-13	5.684e-14
		ρV_1	1.817e-13	-6.537e-13	-8.065e-13
		ρV_2	1.612e-13	2.380e-12	1.208e-12
		E	1.811e-13	1.648e-12	4.547e-13
Quadrature II	c_{DG}	ρ	*1.653e-02*	7.105e-14	8.527e-14
		ρV_1	*2.712e-02*	-1.343e-12	-1.599e-12
		ρV_2	*3.173e-02*	2.103e-12	1.830e-12
		E	*2.774e-02*	6.537e-13	1.421e-13
	c_+	ρ	*1.235e-02*	1.421e-13	1.847e-13
		ρV_1	*1.730e-02*	-1.332e-12	-1.474e-12
		ρV_2	*2.003e-02*	2.096e-12	1.748e-12
		E	*2.060e-02*	5.684e-13	4.263e-13
Collocation	c_{DG}	ρ	2.505e-13	4.974e-13	8.242e-13
		ρV_1	6.482e-13	-1.975e-12	1.322e-12
		ρV_2	6.931e-13	3.634e-12	1.219e-12
		E	5.879e-13	1.364e-12	2.103e-12
	c_+	ρ	1.691e-13	6.111e-13	6.679e-13
		ρV_1	5.552e-13	-1.972e-12	7.248e-13
		ρV_2	5.716e-13	3.617e-12	1.037e-12
		E	4.905e-13	9.948e-13	1.364e-12

for the computations were generated using `modepy`, developed by Andreas Klöckner. The uniform triangular grids and the LGL quadrature rules were generated using `meshzoo` and `quadpy`, respectively, both by Nico Schlömer. All plots were produced using Matplotlib [54].

Statements

Code availability All computations described in this work were performed using the Generalized High-Order Solver Toolbox (`GHOST`), which was developed by the first author and is available under the GNU General Public License at https://github.com/tristanmontoya/GHOST.

Availability of data and material The data generated from the numerical experiments described in §5 as well as the scripts used to process and tabulate the results are available within the above repository.

Funding The authors acknowledge the financial support provided by the University of Toronto, the Natural Sciences and Engineering Research Council of Canada, and the Government of Ontario. Computations were performed on the Niagara supercomputer at the SciNet HPC Consortium [55], which is funded by the Canada Foundation for Innovation, the Government of Ontario, the Ontario Research Fund – Research Excellence, and the University of Toronto.

Conflicts of interest/Competing interests The authors declare that there are no conflicts of interest or competing interests associated with this work.

Table 6: Euler equations, $p = 4$

Scheme	c	Equation	Equivalence	Conservation	
				Strong	Weak
Quadrature I	c_{DG}	ρ	6.169e-14	1.705e-13	1.563e-13
		ρV_1	8.503e-14	-2.167e-13	-4.832e-13
		ρV_2	7.916e-14	7.674e-13	9.770e-13
		E	9.730e-14	4.263e-13	3.126e-13
	c_+	ρ	5.551e-14	1.421e-13	2.132e-13
		ρV_1	8.555e-14	-2.380e-13	-5.471e-13
		ρV_2	8.252e-14	7.923e-13	1.013e-12
		E	8.694e-14	2.842e-13	2.558e-13
Quadrature II	c_{DG}	ρ	*1.150e-02*	5.969e-13	7.105e-14
		ρV_1	*9.686e-03*	-3.624e-13	-2.121e-12
		ρV_2	*9.717e-03*	2.167e-12	2.093e-12
		E	*7.893e-03*	1.535e-12	2.842e-13
	c_+	ρ	*4.837e-03*	4.690e-13	1.421e-13
		ρV_1	*5.590e-03*	-3.091e-13	-2.107e-12
		ρV_2	*6.339e-03*	2.188e-12	1.993e-12
		E	*5.901e-03*	1.819e-12	1.137e-13
Collocation	c_{DG}	ρ	3.886e-13	-1.435e-12	1.393e-12
		ρV_1	8.213e-13	-3.755e-12	5.258e-13
		ρV_2	8.456e-13	-1.005e-12	3.457e-12
		E	9.621e-13	-3.922e-12	3.212e-12
	c_+	ρ	3.765e-13	-1.293e-12	1.563e-12
		ρV_1	8.498e-13	-3.787e-12	5.294e-13
		ρV_2	8.323e-13	-1.016e-12	3.581e-12
		E	9.764e-13	-3.922e-12	3.325e-12

Appendix A Polynomial bases

Although the theoretical analysis in this work is independent of the choice of basis $\mathcal{B}$ employed for a given discretization, the concrete implementation of a DG or FR method nevertheless requires *some* basis to be chosen. Two families of such bases, both of which are used for the computations in §5, are thus described in this appendix.

A.1 Modal (L^2-orthogonal) basis

A basis $\mathcal{B}_0 := \{\phi_0^{(1)}, \ldots, \phi_0^{(N^*)}\}$ is said to be orthogonal with respect to the L^2 inner product on the reference element if any two basis functions $\phi_0^{(i)}, \phi_0^{(j)} \in \mathcal{B}_0$ satisfy

$$\int_{\hat{\Omega}} \phi_0^{(i)}(\underline{\hat{x}})\phi_0^{(j)}(\underline{\hat{x}})\, d\underline{\hat{x}} = 0, \quad \forall\, i \neq j. \tag{A.1}$$

Analytical expressions for such bases are well known for certain spaces and element types; for example, an orthogonal basis for the space $\mathbb{P}_p(\hat{\Omega})$ on the reference triangle given in (5.1) was introduced by Proriol [56] (see also Koornwinder [57, §3.3] and Dubiner [58, §5]) as a two-variable generalization of the Legendre polynomial system. Defining $P_r^{(a,b)} \in \mathbb{P}_r([-1,1])$ as the degree $r \in \mathbb{N}_0$ Jacobi polynomial satisfying

$$\int_{-1}^{1} P_q^{(a,b)}(\xi)P_r^{(a,b)}(\xi)(1-\xi)^a(1+\xi)^b\, d\xi = 0, \quad \forall\, q \in \{0, \ldots, r-1\}, \tag{A.2}$$

for $a, b > -1$, such basis functions, often referred to as the Proriol-Koornwinder-Dubiner (PKD) polynomials, are given in their normalized form by

$$\phi_0^{(\pi(\underline{\alpha}))}(\underline{\chi}(\underline{\xi})) := \sqrt{\left(\alpha_1 + \tfrac{1}{2}\right)\left(\alpha_1 + \alpha_2 + 1\right)} P_{\alpha_1}^{(0,0)}(\xi_1)\left(\tfrac{1-\xi_2}{2}\right)^{\alpha_1} P_{\alpha_2}^{(2\alpha_1+1,0)}(\xi_2), \quad \text{(A.3)}$$

where $\pi : \mathcal{N} \to \{1, \ldots, (p+1)(p+2)/2\}$ represents a scalar ordering[13] of the set $\mathcal{N} := \{\underline{\alpha} \in \mathbb{N}_0^2 : \alpha_1 + \alpha_2 \leq p\}$, and we define the transformation $\underline{\chi} : [-1, 1]^2 \to \hat{\Omega}$ as

$$\underline{\chi}(\underline{\xi}) := \begin{bmatrix} \frac{1}{2}(1+\xi_1)(1-\xi_2) - 1 \\ \xi_2 \end{bmatrix}. \quad \text{(A.4)}$$

Remark A.1 Similar bases may be constructed on quadrilateral, hexahedral, tetrahedral, prismatic, and pyramidal elements, as described in [39, Ch. 3].

A.2 Nodal (Lagrange) basis

While it is possible to directly employ a modal basis to represent the numerical solution such that $\mathcal{B} = \mathcal{B}_0$, we may alternatively choose $\mathcal{B}$ to be a nodal basis associated with the set $\tilde{\mathcal{S}} := \{\underline{\tilde{\hat{x}}}^{(1)}, \ldots, \underline{\tilde{\hat{x}}}^{(N^*)}\} \subset \hat{\Omega}$ such that $\phi^{(i)}(\underline{\tilde{\hat{x}}}^{(j)}) = \delta_{ij}$. Considering the space $\mathbb{P}_p([-1, 1])$, the corresponding Lagrange basis functions are given by

$$\phi^{(i)}(\hat{x}) := \prod_{j=1, j\neq i}^{p+1} \frac{\hat{x} - \tilde{\hat{x}}^{(j)}}{\tilde{\hat{x}}^{(i)} - \tilde{\hat{x}}^{(j)}} \quad \text{(A.5)}$$

for any set $\tilde{\mathcal{S}} \subset [-1, 1]$ consisting of $p+1$ distinct nodes. In the case of $d \geq 2$, analytical expressions do not necessarily exist for nodal bases on arbitrary non-tensorial nodal sets; we therefore take a more general approach by constructing a generalized Vandermonde matrix $\underline{\underline{\tilde{V}}} \in \mathbb{R}^{N^* \times N^*}$ for a modal basis $\mathcal{B}_0$ as described in Appendix A.1, which is nonsingular when $\tilde{\mathcal{S}}$ is unisolvent for $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$. The corresponding nodal basis functions for such a space are then given in terms of $\mathcal{B}_0$ as

$$\phi^{(i)}(\underline{\hat{x}}) := \sum_{j=1}^{N^*} \left[\underline{\underline{\tilde{V}}}^{-\mathrm{T}}\right]_{ij} \phi_0^{(j)}(\underline{\hat{x}}), \quad \text{where} \quad \tilde{V}_{ij} := \phi_0^{(j)}\left(\underline{\tilde{\hat{x}}}^{(i)}\right). \quad \text{(A.6)}$$

The entries of the corresponding differentiation matrix in (3.12) are therefore

$$D_{ij}^{(m)} = \frac{\partial \phi^{(j)}}{\partial \hat{x}_m}\left(\underline{\tilde{\hat{x}}}^{(i)}\right) = \sum_{k=1}^{N^*} \left[\underline{\underline{\tilde{V}}}^{-\mathrm{T}}\right]_{jk} \frac{\partial \phi_0^{(k)}}{\partial \hat{x}_m}\left(\underline{\tilde{\hat{x}}}^{(i)}\right), \quad \forall i, j \in \{1, \ldots, N^*\}, \quad \text{(A.7)}$$

defining a nodal SBP operator on $\tilde{\mathcal{S}}$ in the sense of [21, Definition 2.1] in the case of discrete inner products satisfying Assumptions 3.1 and 3.2.

Remark A.2 In the case of $d \geq 2$, unisolvency is not guaranteed for arbitrary nodal sets $\tilde{\mathcal{S}}$ containing N^* distinct nodes. Special care is therefore needed to ensure that $\underline{\underline{\tilde{V}}}$ is invertible, where we refer to Marchildon and Zingg [59] for the derivation of necessary conditions for obtaining unisolvent symmetrical nodal sets for total-degree polynomial spaces on triangles and tetrahedra. Further details regarding nodal bases for high-order methods may be found, for example, in [38, §3.1, §6.1, §10.1].

[13] Such an ordering may be chosen arbitrarily, provided that $\underline{\alpha} \mapsto \pi(\underline{\alpha})$ is bijective; for example, Hesthaven and Warburton [38, §6.1] consider $\pi(\underline{\alpha}) := \alpha_2 + (p+1)\alpha_1 + 1 - \alpha_1(\alpha_1 - 1)/2$.

Appendix B Discrete inner products

Implementations of the DG and FR methods presented in §2.5 can often be characterized by the techniques employed for numerical integration or projection, which may be formalized in terms of discrete inner products defined as in (3.1) and (3.2). In this appendix, we situate the standard quadrature-based integration approach commonly employed for DG methods (i.e. those in [4–7]) as well as the collocation-based approach employed for the nodal DG formulations described in [38] and the FR schemes in [8–12] within the context of the present framework.

B.1 Quadrature-based approximation

Considering quadrature rules on $\mathcal{S}$ and $\mathcal{S}^{(\zeta)}$ with positive weights given by $\{\omega^{(i)}\}_{i=1}^{N}$ and $\{\omega^{(\zeta,i)}\}_{i=1}^{N_\zeta}$, respectively, we may define the diagonal weight matrices

$$\underline{\underline{W}} := \operatorname{diag}\big(\omega^{(1)},\dots,\omega^{(N)}\big) \quad \text{and} \quad \underline{\underline{B}}^{(\zeta)} := \operatorname{diag}\big(\omega^{(\zeta,1)},\dots,\omega^{(\zeta,N_\zeta)}\big), \tag{B.1}$$

such that the double sums in (3.1) and (3.2) reduce to

$$\big\langle U,V\big\rangle_{\underline{\underline{W}}} := \sum_{i=1}^{N} U(\underline{\hat{x}}^{(i)})V(\underline{\hat{x}}^{(i)})\,\omega^{(i)} \tag{B.2}$$

and

$$\big\langle U,V\big\rangle_{\underline{\underline{B}}^{(\zeta)}} := \sum_{i=1}^{N_\zeta} U(\underline{\hat{x}}^{(\zeta,i)})V(\underline{\hat{x}}^{(\zeta,i)})\,\omega^{(\zeta,i)}, \tag{B.3}$$

respectively. The following lemma provides sufficient conditions on such quadrature rules for Assumption 3.1 to be satisfied.

Lemma B.1 *Assumption 3.1 is satisfied with $\mathbb{P}_{\mathcal{N}}(\hat{\Omega}) = \mathbb{P}_p(\hat{\Omega})$ for any $p \in \mathbb{N}_0$ whenever the discrete inner products are computed as in (B.2) and (B.3) using volume and facet quadrature rules of at least total degree $2p-1$ and $2p$, respectively.*

Proof For any $U,V \in \mathbb{P}_p(\hat{\Omega})$, the volume integrals in (3.3) contain functions in $\mathbb{P}_{2p-1}(\hat{\Omega})$, while the facet integrals contain functions in $\mathbb{P}_{2p}(\hat{\Gamma}^{(\zeta)})$. For volume and facet quadrature rules of total degree $2p-1$ or greater and $2p$ or greater, respectively, all terms in (3.3) are computed exactly, and hence (3.4) holds for $m \in \{1,\dots,d\}$. □

Remark B.1 In addition to the positive-definiteness of $\underline{\underline{W}}$ and $\underline{\underline{B}}^{(\zeta)}$, which is clearly the case if and only if the volume and facet quadrature weights are strictly positive, Assumption 3.2 requires $\underline{\underline{V}}$ to be of rank N^*, which is difficult to ensure theoretically, particularly in the case of $N > N^*$. We have nevertheless numerically verified that such a property is satisfied for all schemes employed for the computations in §5.

Remark B.2 Since Assumptions 2.1 and 3.2 imply that the matrices and $\underline{\underline{J}}^{(\kappa)}$ and $\underline{\underline{W}}$ are both SPD, it follows by a similar argument to that used in the proof of Lemma 3.1 that $\underline{\underline{M}}^{(\kappa)}$ is also SPD under such conditions when $\underline{\underline{W}}$ is diagonal, as given in (B.1), implying that such conditions are sufficient for Assumption 3.3 to be satisfied.

B.2 Collocation-based approximation

Assuming that $\mathcal{S}$ is unisolvent for a space $\mathbb{P}_{\mathcal{K}}(\hat{\Omega}) \supseteq \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ of dimension $N \geq N^*$, we may define a nodal basis $\mathcal{B}_{\mathrm{c}} := \{\ell^{(1)}, \dots, \ell^{(N)}\}$ for $\mathbb{P}_{\mathcal{K}}(\hat{\Omega})$ such that the collocation projection $\mathit{\Pi}_{\mathrm{c}} U \in \mathbb{P}_{\mathcal{K}}(\hat{\Omega})$ of a function $U : \hat{\Omega} \to \mathbb{R}$ is given explicitly as

$$(\mathit{\Pi}_{\mathrm{c}} U)(\underline{\hat{x}}) := \sum_{i=1}^{N} U(\underline{\hat{x}}^{(i)}) \ell^{(i)}(\underline{\hat{x}}), \quad \text{where} \quad \ell^{(i)}(\underline{\hat{x}}^{(j)}) = \delta_{ij}, \tag{B.4}$$

recovering the projection in (2.13) when $\mathbb{P}_{\mathcal{K}}(\hat{\Omega}) = \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$. Similarly, if $\mathcal{S}^{(\zeta)}$ is unisolvent for a polynomial space on $\hat{\Gamma}^{(\zeta)} \subset \partial\hat{\Omega}$ which is of dimension $N_\zeta \geq N_\zeta^*$ and contains $\mathbb{P}_{\mathcal{N}}(\hat{\Gamma}^{(\zeta)})$, we may define a projection operator analogously to (B.4) as

$$(\mathit{\Pi}_{\mathrm{c}}^{(\zeta)} U)(\underline{\hat{x}}) := \sum_{i=1}^{N_\zeta} U(\underline{\hat{x}}^{(\zeta,i)}) \ell^{(\zeta,i)}(\underline{\hat{x}}), \quad \text{where} \quad \ell^{(\zeta,i)}(\underline{\hat{x}}^{(\zeta,j)}) = \delta_{ij}. \tag{B.5}$$

The corresponding discrete inner products are then given as

$$\left\langle U, V \right\rangle_{\underline{\underline{W}}} := \int_{\hat{\Omega}} (\mathit{\Pi}_{\mathrm{c}} U)(\underline{\hat{x}}) (\mathit{\Pi}_{\mathrm{c}} V)(\underline{\hat{x}}) \, d\underline{\hat{x}} \tag{B.6}$$

and

$$\left\langle U, V \right\rangle_{\underline{\underline{B}}^{(\zeta)}} := \int_{\hat{\Gamma}^{(\zeta)}} (\mathit{\Pi}_{\mathrm{c}}^{(\zeta)} U)(\underline{\hat{x}}) (\mathit{\Pi}_{\mathrm{c}}^{(\zeta)} V)(\underline{\hat{x}}) \, d\hat{s}, \tag{B.7}$$

which take the forms in (3.1) and (3.2), respectively, with

$$W_{ij} := \int_{\hat{\Omega}} \ell^{(i)}(\underline{\hat{x}}) \ell^{(j)}(\underline{\hat{x}}) \, d\underline{\hat{x}} \quad \text{and} \quad B_{ij}^{(\zeta)} := \int_{\hat{\Gamma}^{(\zeta)}} \ell^{(\zeta,i)}(\underline{\hat{x}}) \ell^{(\zeta,j)}(\underline{\hat{x}}) \, d\hat{s}, \tag{B.8}$$

where the above integrals may be evaluated as part of a preprocessing stage.

Remark B.3 Noting that the matrices $\underline{\underline{W}}$ and $\underline{\underline{B}}^{(\zeta)}$ defined in (B.8) are generally dense, and are SPD due to the linear independence of the nodal bases used to define the projections (see, for example, Horn and Johnson [60, Theorem 7.2.10]), Assumptions 3.1 and 3.2 (as well as Assumption 3.3 for meshes satisfying Assumption 2.1, at least in the case of $\mathbb{P}_{\mathcal{K}}(\hat{\Omega}) = \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ considered for the numerical experiments in §5) are therefore satisfied by construction. Such a quadrature-free formulation enables the construction of conservative and energy-stable schemes on arbitrary unisolvent nodal sets for which the requirements of Lemma B.1 may not be met.

Remark B.4 Recalling Remark 4.3, the weights of the interpolatory quadrature rule on the abscissae $\mathcal{S}$ under which conservation may be proven for collocation-based approximations may be expressed in terms of the nodal basis $\mathcal{B}_{\mathrm{c}}$ as

$$\omega^{(i)} = \sum_{j=1}^{N} W_{ij} = \int_{\hat{\Omega}} \ell^{(i)}(\underline{\hat{x}}) \, d\underline{x}, \quad \forall\, i \in \{1, \dots, N\}, \tag{B.9}$$

where such a quadrature is exact (at least) for all functions in $\mathbb{P}_{\mathcal{K}}(\hat{\Omega})$. If, however, the nodes in $\mathcal{S}$ are chosen such that the quadrature rule with weights given as in

(B.9) is exact for all products of *two* functions in $\mathbb{P}_{\mathcal{K}}(\hat{\Omega})$,[14] it can be shown (see, for example, [37]) that $\underline{\underline{W}}$ reduces to a diagonal matrix of quadrature weights, recovering an approximation identical to that in Appendix B.1. An analogous equivalence holds for the matrices $\underline{\underline{B}}^{(\zeta)}$ and the associated facet quadrature rules.

Appendix C Parametrization of the $\underline{\underline{K}}$ matrix

The following lemma establishes suitable choices of multi-index sets $\mathcal{M} \subset \mathcal{N}$ satisfying the first part of Assumption 3.4 for schemes employing total-degree polynomial spaces (the case of tensor-product elements is discussed by Cicchino and Nadarajah in [61]).

Lemma C.1 *Taking $\mathbb{P}_{\mathcal{N}}(\hat{\Omega}) = \mathbb{P}_p(\hat{\Omega})$ for any $p \in \mathbb{N}_0$, the choice of multi-indices $\mathcal{M} := \{\underline{\alpha} \in \mathbb{N}_0^d : |\underline{\alpha}| = p\}$ in (3.28) ensures that $\underline{\underline{K}}\,\underline{\underline{D}}^{(m)} = \underline{\underline{0}}$ for all $m \in \{1, \dots, d\}$.*

Proof Noting from (3.28) that $\underline{\underline{K}}\,\underline{\underline{D}}^{(m)} = \underline{\underline{0}}$ is satisfied for a given $m \in \{1, \dots, d\}$ if $\underline{\underline{D}}^{(\underline{\alpha})}\underline{\underline{D}}^{(m)} = \underline{\underline{0}}$ for all $\underline{\alpha} \in \mathcal{M}$, it is sufficient to require the image of any $V \in \mathbb{P}_p(\hat{\Omega})$ under the corresponding differential operator $\partial^{|\underline{\alpha}|+1}/\partial\hat{x}_1^{\alpha_1} \cdots \partial\hat{x}_m^{\alpha_m+1} \cdots \partial\hat{x}_d^{\alpha_d}$ to be the zero function. Since differentiating $V \in \mathbb{P}_p(\hat{\Omega})$ a total of $p+1$ times always yields the zero function, and the action of the above operator consists of differentiating a total of $|\underline{\alpha}|+1$ times, constraining $\mathcal{M}$ to satisfy $|\underline{\alpha}| = p$ leads to the desired result. □

Following [11, §5.3] and [12, §4.1], the set $\mathcal{M}$ defined above may be parametrized in terms of $d-1$ scalar indices to obtain

$$\underline{\underline{K}}^{\mathrm{1D}} := \frac{c}{|\hat{\Omega}|}\left(\underline{\underline{D}}^{p}\right)^{\mathrm{T}}\underline{\underline{M}}\,\underline{\underline{D}}^{p}, \qquad \underline{\underline{K}}^{\mathrm{2D}} := \frac{c}{|\hat{\Omega}|}\sum_{q=0}^{p}\binom{p}{q}\left(\underline{\underline{D}}^{(p-q,q)}\right)^{\mathrm{T}}\underline{\underline{M}}\,\underline{\underline{D}}^{(p-q,q)}, \tag{C.1}$$

and

$$\underline{\underline{K}}^{\mathrm{3D}} := \frac{c}{|\hat{\Omega}|}\sum_{q_1=0}^{p}\sum_{q_2=0}^{q_1}\binom{p}{q_1}\binom{q_1}{q_2}\left(\underline{\underline{D}}^{(p-q_1,q_1-q_2,q_2)}\right)^{\mathrm{T}}\underline{\underline{M}}\,\underline{\underline{D}}^{(p-q_1,q_1-q_2,q_2)}, \tag{C.2}$$

defining one-parameter families of symmetric correction fields, where $c \in \mathbb{R}$ determines the properties of the resulting schemes. With respect to the formalism of Theorem 3.1, the parametrizations used to define $\underline{\underline{K}}^{\mathrm{2D}}$ and $\underline{\underline{K}}^{\mathrm{3D}}$ in (C.1) and (C.2) correspond to $c^{(\underline{\alpha})} = c\binom{p}{\alpha_2}$ and $c^{(\underline{\alpha})} = c\binom{p}{p-\alpha_1}\binom{p-\alpha_1}{\alpha_3}$, respectively, and lead to $\underline{\underline{K}}$ being SPSD for all $c \geq 0$, which is sufficient under Assumption 3.2 for $\underline{\underline{M}} + \underline{\underline{K}}$ to be SPD, implying that the second part of Assumption 3.4 is satisfied under such conditions.

Remark C.1 While $c \geq 0$ is typically assumed for simplicial elements with $d \geq 2$, it has been shown for $d = 1$ (see, for example, [10, §3.3], [13, §3.2], and [23, §3.6]) that there exists $c_- < 0$ depending only on the polynomial degree p (and on the volume quadrature rule, if (3.26) is not satisfied) such that $\underline{\underline{M}} + \underline{\underline{K}}$ is SPD for $c_- < c < \infty$.

[14] Taking $\mathbb{P}_{\mathcal{K}}(\hat{\Omega}) = \mathbb{P}_q(\hat{\Omega})$ or $\mathbb{P}_{\mathcal{K}}(\hat{\Omega}) = \mathbb{Q}_q(\hat{\Omega})$ for $q \in \mathbb{N}_0$, this would require the nodes in $\mathcal{S}$ to be associated with a volume quadrature rule of at least (total or tensor-product) degree $2q$, as is the case, for example, when $\mathcal{S}$ consists of $N = q+1$ LG nodes in one dimension.

References

1. H.-O. Kreiss and J. Oliger, "Comparison of accurate methods for the integration of hyperbolic equations," *Tellus*, vol. 24, pp. 199–215, June 1972.
2. Z. J. Wang, K. Fidkowski, R. Abgrall, F. Bassi, D. Caraeni, A. Cary, H. Deconinck, R. Hartmann, K. Hillewaert, H. T. Huynh, N. Kroll, G. May, P.-O. Persson, B. van Leer, and M. Visbal, "High-order CFD methods: Current status and perspective," *International Journal for Numerical Methods in Fluids*, vol. 72, pp. 811–845, Jan. 2013.
3. W. H. Reed and T. R. Hill, "Triangular mesh methods for the neutron transport equation," Tech. Rep. LA-UR-73-479, Los Alamos Scientific Laboratory, Apr. 1973.
4. B. Cockburn and C.-W. Shu, "TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws II: General framework," *Mathematics of Computation*, vol. 52, pp. 411–435, Apr. 1989.
5. B. Cockburn, S.-Y. Lin, and C.-W. Shu, "TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws III: One-dimensional systems," *Journal of Computational Physics*, vol. 84, pp. 90–113, Sept. 1989.
6. B. Cockburn, S. Hou, and C.-W. Shu, "The Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV: The multidimensional case," *Mathematics of Computation*, vol. 54, pp. 545–581, Apr. 1990.
7. B. Cockburn and C.-W. Shu, "The Runge-Kutta discontinuous Galerkin method for conservation laws V: Multidimensional systems," *Journal of Computational Physics*, vol. 141, pp. 199–224, Apr. 1998.
8. H. T. Huynh, "A flux reconstruction approach to high-order schemes including discontinuous Galerkin methods," *18th AIAA Computational Fluid Dynamics Conference*, June 2007.
9. Z. J. Wang and H. Gao, "A unifying lifting collocation penalty formulation including the discontinuous Galerkin, spectral volume/difference methods for conservation laws on mixed grids," *Journal of Computational Physics*, vol. 228, pp. 8161–8186, Nov. 2009.
10. P. E. Vincent, P. Castonguay, and A. Jameson, "A new class of high-order energy stable flux reconstruction schemes," *Journal of Scientific Computing*, vol. 47, pp. 50–72, Sept. 2010.
11. P. Castonguay, P. E. Vincent, and A. Jameson, "A new class of high-order energy stable flux reconstruction schemes for triangular elements," *Journal of Scientific Computing*, vol. 51, pp. 224–256, June 2011.
12. D. M. Williams and A. Jameson, "Energy stable flux reconstruction schemes for advection–diffusion problems on tetrahedra," *Journal of Scientific Computing*, vol. 59, pp. 721–759, Sept. 2013.
13. Y. Allaneau and A. Jameson, "Connections between the filtered discontinuous Galerkin method and the flux reconstruction approach to high order discretizations," *Computer Methods in Applied Mechanics and Engineering*, vol. 75, pp. 3628–3636, Dec. 2011.
14. D. De Grazia, G. Mengaldo, D. Moxey, P. E. Vincent, and S. J. Sherwin, "Connections between the discontinuous Galerkin method and high-order flux reconstruction schemes," *International Journal for Numerical Methods in Fluids*, vol. 75, pp. 860–877, May 2014.
15. G. Mengaldo, D. De Grazia, P. E. Vincent, and S. J. Sherwin, "On the connections between discontinuous Galerkin and flux reconstruction schemes: Extension to curvilinear meshes," *Journal of Scientific Computing*, vol. 67, pp. 1272–1292, June 2015.
16. P. Zwanenburg and S. Nadarajah, "Equivalence between the energy stable flux reconstruction and filtered discontinuous Galerkin schemes," *Journal of Computational Physics*, vol. 306, pp. 343–369, Feb. 2016.
17. D. C. Del Rey Fernández, J. E. Hicken, and D. W. Zingg, "Review of summation-by-parts operators with simultaneous approximation terms for the numerical solution of partial differential equations," *Computers & Fluids*, vol. 95, pp. 171–196, May 2014.
18. M. Svärd and J. Nordström, "Review of summation-by-parts schemes for initial-boundary-value problems," *Journal of Computational Physics*, vol. 268, pp. 17–38, July 2014.
19. H.-O. Kreiss and G. Scherer, "Finite element and finite difference methods for hyperbolic partial differential equations," in *Mathematical Aspects of Finite Elements in Partial Differential Equations* (C. de Boor, ed.), pp. 195–212, Academic Press, 1974.
20. D. C. Del Rey Fernández, P. D. Boom, and D. W. Zingg, "A generalized framework for nodal first derivative summation-by-parts operators," *Journal of Computational Physics*, vol. 266, pp. 214–239, June 2014.

21. J. E. Hicken, D. C. Del Rey Fernández, and D. W. Zingg, "Multidimensional summation-by-parts operators: General theory and application to simplex elements," *SIAM Journal on Scientific Computing*, vol. 38, no. 4, pp. A1935–A1958, 2016.
22. G. J. Gassner, "A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods," *SIAM Journal on Scientific Computing*, vol. 35, pp. A1233–A1253, May 2013.
23. H. Ranocha, P. Öffner, and T. Sonar, "Summation-by-parts operators for correction procedure via reconstruction," *Journal of Computational Physics*, vol. 311, pp. 299–328, Apr. 2016.
24. G. J. Gassner, A. R. Winters, and D. A. Kopriva, "Split form nodal discontinuous Galerkin schemes with summation-by-parts property for the compressible Euler equations," *Journal of Computational Physics*, vol. 327, pp. 39–66, Dec. 2016.
25. J. Chan, "On discretely entropy conservative and entropy stable discontinuous Galerkin methods," *Journal of Computational Physics*, vol. 362, pp. 346–374, June 2018.
26. D. A. Kopriva and G. J. Gassner, "On the quadrature and weak form choices in collocation type discontinuous Galerkin spectral element methods," *Journal of Scientific Computing*, vol. 44, pp. 136–155, May 2010.
27. R. M. Kirby and G. E. Karniadakis, "De-aliasing on non-uniform grids: Algorithms and applications," *Journal of Computational Physics*, vol. 191, pp. 249–264, Oct. 2003.
28. G. Mengaldo, D. De Grazia, D. Moxey, P. E. Vincent, and S. J. Sherwin, "Dealiasing techniques for high-order spectral element methods on regular and irregular grids," *Journal of Computational Physics*, vol. 299, pp. 56–81, Oct. 2015.
29. X. Zhang and C.-W. Shu, "Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: Survey and new developments," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 467, pp. 2752–2776, May 2011.
30. M. Vinokur, "Conservation equations of gasdynamics in curvilinear coordinate systems," *Journal of Computational Physics*, vol. 14, pp. 105–125, Feb. 1974.
31. M. E. Gurtin, E. Fried, and L. Anand, *The Mechanics and Thermodynamics of Continua*. Cambridge University Press, 2010.
32. N. Dyn and M. S. Floater, "Multivariate polynomial interpolation on lower sets," *Journal of Approximation Theory*, vol. 177, pp. 34–42, Jan. 2014.
33. M. Yu, Z. J. Wang, and Y. Liu, "On the accuracy and efficiency of discontinuous Galerkin, spectral difference and correction procedure via reconstruction methods," *Journal of Computational Physics*, vol. 259, pp. 70–95, Feb. 2014.
34. E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer Berlin Heidelberg, 3rd ed., 2009.
35. C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains*. Springer Berlin Heidelberg, 2006.
36. W. R. Boland and C. S. Duris, "Product type quadrature formulas," *BIT*, vol. 11, pp. 139–158, June 1971.
37. D. R. Hunkins, "Product type multiple integration formulas," *BIT*, vol. 13, pp. 408–414, Dec. 1973.
38. J. S. Hesthaven and T. Warburton, *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*. Springer New York, 2008.
39. G. E. Karniadakis and S. J. Sherwin, *Spectral/hp Element Methods for Computational Fluid Dynamics*. Oxford University Press, 2nd ed., 2005.
40. T. Chen and C.-W. Shu, "Review of entropy stable discontinuous Galerkin methods for systems of conservation laws on unstructured simplex meshes," *CSIAM Transactions on Applied Mathematics*, vol. 1, pp. 1–52, June 2020.
41. A. Jameson, "A proof of the stability of the spectral difference method for all orders of accuracy," *Journal of Scientific Computing*, vol. 45, pp. 348–358, Jan. 2010.
42. M. H. Carpenter, D. Gottlieb, and S. Abarbanel, "Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: Methodology and application to high-order compact schemes," *Journal of Computational Physics*, vol. 111, pp. 220–236, Apr. 1994.
43. D. Funaro and D. Gottlieb, "A new method of imposing boundary conditions in pseudospectral approximations of hyperbolic equations," *Mathematics of Computation*, vol. 51, pp. 599–599, Oct. 1988.

44. B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time-Dependent Problems and Difference Methods*. John Wiley & Sons, Inc., 2nd ed., 2013.
45. H. Xiao and Z. Gimbutas, "A numerical algorithm for the construction of efficient quadrature rules in two and higher dimensions," *Computers & Mathematics with Applications*, vol. 59, pp. 663–676, Jan. 2010.
46. T. Warburton, "An explicit construction of interpolation nodes on the simplex," *Journal of Engineering Mathematics*, vol. 56, pp. 247–262, Sept. 2006.
47. P. L. Roe, "Approximate Riemann solvers, parameter vectors, and difference schemes," *Journal of Computational Physics*, vol. 43, pp. 357–372, Sept. 1981.
48. D. C. Del Rey Fernández, P. D. Boom, M. Shademan, and D. W. Zingg, "Numerical investigation of tensor-product summation-by-parts discretization strategies and operators," in *55th AIAA Aerospace Sciences Meeting*, American Institute of Aeronautics and Astronautics, Jan. 2017.
49. B. Cockburn and C.-W. Shu, "Runge-Kutta discontinuous Galerkin methods for convection-dominated problems," *Journal of Scientific Computing*, vol. 16, Sept. 2001.
50. C.-W. Shu, "Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws," in *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations: Lectures given at the 2nd Session of the Centro Internazionale Matematico Estivo (C.I.M.E.) held in Cetraro, Italy, June 23–28, 1997* (A. Quarteroni, ed.), pp. 325–432, Springer Berlin Heidelberg, 1998.
51. S. C. Spiegel, H. T. Huynh, and J. R. DeBonis, "A survey of the isentropic Euler vortex problem using high-order methods," in *22nd AIAA Computational Fluid Dynamics Conference*, American Institute of Aeronautics and Astronautics, June 2015.
52. D. C. Del Rey Fernández, J. E. Hicken, and D. W. Zingg, "Simultaneous approximation terms for multi-dimensional summation-by-parts operators," *Journal of Scientific Computing*, vol. 75, pp. 83–110, Apr. 2018.
53. M. Svärd and J. Nordström, "On the convergence rates of energy-stable finite-difference schemes," *Journal of Computational Physics*, vol. 397, Nov. 2019.
54. J. D. Hunter, "Matplotlib: A 2D graphics environment," *Computing in Science & Engineering*, vol. 9, pp. 90–95, June 2007.
55. M. Ponce, R. van Zon, S. Northrup, D. Gruner, J. Chen, F. Ertinaz, A. Fedoseev, L. Groer, F. Mao, B. C. Mundim, M. Nolta, J. Pinto, M. Saldarriaga, V. Slavnic, E. Spence, C.-H. Yu, and W. R. Peltier, "Deploying a top-100 supercomputer for large parallel workloads," in *Proceedings of the Practice and Experience in Advanced Research Computing on Rise of the Machines (Learning)*, ACM, July 2019.
56. J. Proriol, "Sur une famille de polynomes à deux variables orthogonaux dans un triangle," *Comptes Rendus Hebdomadaires des Séances de l'Académie des Sciences*, vol. 245, pp. 2459–2461, Dec. 1957.
57. T. Koornwinder, "Two-variable analogues of the classical orthogonal polynomials," in *Theory and Application of Special Functions* (R. Askey, ed.), pp. 435–495, Elsevier, 1975.
58. M. Dubiner, "Spectral methods on triangles and other domains," *Journal of Scientific Computing*, vol. 6, pp. 345–390, Dec. 1991.
59. A. L. Marchildon and D. W. Zingg, "Unisolvency for polynomial interpolation in simplices with symmetrical nodal distributions," Submitted to *Journal of Scientific Computing*, 2021.
60. R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2nd ed., 2013.
61. A. Cicchino and S. Nadarajah, "A new norm and stability condition for tensor product flux reconstruction schemes," *Journal of Computational Physics*, vol. 429, Mar. 2021.