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 with respect to suitable quadrature rules and discrete norms 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

The Version of Record is available at `https://doi.org/10.1007/s10915-022-01935-3`.

T. Montoya
University of Toronto Institute for Aerospace Studies, Toronto, Canada
E-mail: `tristan.montoya@mail.utoronto.ca`

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 systems of partial differential equations (PDEs). High-order methods (i.e. those of third order or higher) have shown significant promise in these contexts, particularly for computations involving the propagation of waves over long distances (an early example being the work of Kreiss and Oliger [1]) or 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 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 Fernandez *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 novel 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 by subsequent extensions to entropy-stable and kinetic-energy-preserving DG-type schemes exploiting the SBP property in the context of compressible flows (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 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.

- 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 of the 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.
- The Vincent-Castonguay-Jameson-Huynh (VCJH) family of FR schemes (as described in [10–12]), is recast in a generalized matrix form, allowing for the use of arbitrary nodal or modal bases on any element type as well as over-integration of the volume and/or 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 efficiency.
- 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.
- Energy stability is established with respect to suitable discrete norms 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 as special cases of a more general algebraic 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 DG and FR methods and illustrate the main ideas underlying the application of the SBP approach to such schemes in the context of a one-dimensional advection problem. In §3, we develop a unified matrix formulation for a broad class of DG and FR methods applied to linear or nonlinear conservation laws in one or more space dimensions, 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

Symbols bearing single underlines denote vectors (treated as column matrices), whereas symbols bearing double underlines denote matrices. Symbols in bold are used specifically to denote Cartesian (i.e. spatial) vectors, for which we employ the usual del operator, dot product, and Euclidean norm. Tensors containing indices of both

types (e.g. fluxes for systems of conservation laws) are bolded as well as underlined, and we assume that coordinate transformations and differential operators act over the spatial indices only (i.e. treating them as Cartesian vectors). The symbols $\mathbb{R}$, $\mathbb{R}^+$, $\mathbb{R}_0^+$, $\mathbb{N}$, $\mathbb{N}_0$, and $\mathbb{S}^{d-1}$ denote the real numbers, the positive real numbers, the non-negative real numbers, the natural numbers (excluding zero), the natural numbers including zero, and the unit $(d-1)$-sphere, 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 identity matrices, respectively, with their dimensions inferred from the context. Given any domain $\mathcal{D} \subset \mathbb{R}^d$ of dimension $d \in \mathbb{N}$, we use the symbol $\partial\mathcal{D}$ to denote its boundary, $\bar{\mathcal{D}} := \mathcal{D} \cup \partial\mathcal{D}$ to denote its closure, and $|\mathcal{D}|$ to denote its Lebesgue measure (i.e. d-dimensional volume); the interior of a closed domain $\mathcal{D}$ is then given by $\mathring{\mathcal{D}} := \mathcal{D} \setminus \partial\mathcal{D}$. Other important 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 $N_c \in \mathbb{N}$ conservative variables $\underline{U}(\boldsymbol{x}, t) \in \mathcal{Y} \subseteq \mathbb{R}^{N_c}$, where $\mathcal{Y}$ denotes the set of admissible solution states, and $\boldsymbol{x}$ and t denote the spatial and temporal coordinates, respectively. Such systems of PDEs take the general form

$$\begin{aligned} \frac{\partial \underline{U}(\boldsymbol{x}, t)}{\partial t} + \nabla_{\boldsymbol{x}} \cdot \underline{\boldsymbol{F}}(\underline{U}(\boldsymbol{x}, t)) = \underline{0}, \qquad & \forall\, (\boldsymbol{x}, t) \in \Omega \times (0, T), \\ \underline{U}(\boldsymbol{x}, 0) = \underline{U}^0(\boldsymbol{x}), \qquad & \forall\, \boldsymbol{x} \in \Omega, \end{aligned} \tag{2.1}$$

subject to appropriate boundary conditions, if necessary, where $\Omega \subset \mathbb{R}^d$ denotes a fixed domain of dimension $d \in \mathbb{N}$ with a piecewise smooth boundary, $T \in \mathbb{R}^+$ denotes the final time, $\underline{\boldsymbol{F}}(\underline{U}(\boldsymbol{x}, t))$ denotes the flux tensor, and $\underline{U}^0(\boldsymbol{x})$ denotes the initial data. Although 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 constant-coefficient linear advection equation,

$$\frac{\partial U(\boldsymbol{x}, t)}{\partial t} + \nabla_{\boldsymbol{x}} \cdot \big(\boldsymbol{a} U(\boldsymbol{x}, t)\big) = 0, \tag{2.2}$$

which governs the transport of a scalar quantity $U(\boldsymbol{x}, t) \in \mathbb{R} =: \mathcal{Y}$ at a velocity given by $\boldsymbol{a} \in \mathbb{R}^d$. We are also interested in the Euler equations, which constitute a system of $N_c = d + 2$ coupled PDEs governing the conservation of mass, momentum, and energy for a compressible, inviscid, and adiabatic fluid. Such a system is given by

$$\frac{\partial}{\partial t} \begin{bmatrix} \rho(\boldsymbol{x}, t) \\ \rho(\boldsymbol{x}, t) V_1(\boldsymbol{x}, t) \\ \vdots \\ \rho(\boldsymbol{x}, t) V_d(\boldsymbol{x}, t) \\ E(\boldsymbol{x}, t) \end{bmatrix} + \sum_{m=1}^{d} \frac{\partial}{\partial x_m} \begin{bmatrix} \rho(\boldsymbol{x}, t) V_m(\boldsymbol{x}, t) \\ \rho(\boldsymbol{x}, t) V_1(\boldsymbol{x}, t) V_m(\boldsymbol{x}, t) + P(\boldsymbol{x}, t) \delta_{1m} \\ \vdots \\ \rho(\boldsymbol{x}, t) V_d(\boldsymbol{x}, t) V_m(\boldsymbol{x}, t) + P(\boldsymbol{x}, t) \delta_{dm} \\ V_m(\boldsymbol{x}, t)(E(\boldsymbol{x}, t) + P(\boldsymbol{x}, t)) \end{bmatrix} = \underline{0}, \tag{2.3}$$

where $\rho(\boldsymbol{x}, t) \in \mathbb{R}$ denotes the fluid density, $\boldsymbol{V}(\boldsymbol{x}, t) \in \mathbb{R}^d$ denotes the flow velocity, $E(\boldsymbol{x}, t) \in \mathbb{R}$ denotes the total energy per unit volume, and $P(\boldsymbol{x}, t) \in \mathbb{R}$ denotes the pressure, which is related to the other variables through the equation of state given by $P(\boldsymbol{x}, t) = (\gamma - 1)(E(\boldsymbol{x}, t) - \frac{1}{2}\rho(\boldsymbol{x}, t)\|\boldsymbol{V}(\boldsymbol{x}, t)\|^2)$ 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 $\mathcal{Y} := \{\underline{U}(\boldsymbol{x}, t) \in \mathbb{R}^{d+2} : P(\boldsymbol{x}, t), \rho(\boldsymbol{x}, t) > 0\}$.

Remark 2.1 Throughout this paper, we tacitly assume that the values of numerical solutions to (2.1) remain within $\mathcal{Y}$. 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 [27]), 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 $\mathcal{T}^h := \{\Omega^{(1)}, \ldots, \Omega^{(N_e)}\}$ consisting of $N_e \in \mathbb{N}$ 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 $\bigcup_{\kappa=1}^{N_e} \Omega^{(\kappa)} = \bar{\Omega}$ as well as $\mathring{\Omega}^{(\kappa)} \cap \mathring{\Omega}^{(\nu)} = \emptyset$ for any $\kappa \neq \nu$. Likewise, the boundary of each element $\Omega^{(\kappa)} \in \mathcal{T}^h$ consists of $N_f \in \mathbb{N}$ smooth facets $\{\Gamma^{(\kappa,1)}, \ldots, \Gamma^{(\kappa,N_f)}\}$ satisfying $\bigcup_{\zeta=1}^{N_f} \Gamma^{(\kappa,\zeta)} = \partial\Omega^{(\kappa)}$ and $\mathring{\Gamma}^{(\kappa,\zeta)} \cap \mathring{\Gamma}^{(\kappa,\eta)} = \emptyset$ for any $\zeta \neq \eta$.*[1] *Furthermore, each element is the image of a closed, polytopal reference element $\hat{\Omega} \subset \mathbb{R}^d$ under a smooth, time-invariant mapping $\boldsymbol{X}^{(\kappa)} : \hat{\Omega} \to \Omega^{(\kappa)}$ satisfying $J^{(\kappa)}(\boldsymbol{\xi}) := \det(\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi})) > 0$, where the entries of the Jacobian matrix $\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}) \in \mathbb{R}^{d\times d}$ are given by $[\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi})]_{mn} := \partial X_m^{(\kappa)}(\boldsymbol{\xi})/\partial \xi_n$.*

We now recall from Vinokur [28] that systems in the form of (2.1) retain a conservative formulation under such a transformation, which is given in the present context by

$$\frac{\partial J^{(\kappa)}(\boldsymbol{\xi}) \underline{U}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}), t)}{\partial t} + \nabla_{\boldsymbol{\xi}} \cdot \left(J^{(\kappa)}(\boldsymbol{\xi}) (\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}))^{-1} \underline{\boldsymbol{F}}(\underline{U}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}), t)) \right) = \underline{0}. \quad (2.4)$$

Defining $J^{(\kappa,\zeta)}(\boldsymbol{\xi}) := \| J^{(\kappa)}(\boldsymbol{\xi}) (\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}))^{-\mathrm{T}} \hat{\boldsymbol{n}}^{(\zeta)} \|$, where $\hat{\boldsymbol{n}}^{(\zeta)} \in \mathbb{S}^{d-1}$ is the (constant) outward unit normal vector to the facet $\hat{\Gamma}^{(\zeta)} \subset \partial\hat{\Omega}$ of the reference element, we note that the outward unit normal vector to the corresponding physical facet is given by $\boldsymbol{n}^{(\kappa,\zeta)} : \Gamma^{(\kappa,\zeta)} \to \mathbb{S}^{d-1}$ according to Nanson's formula (see, for example, Gurtin *et al.* [29, §8.1] for a derivation in the context of continuum mechanics),

$$J^{(\kappa,\zeta)}(\boldsymbol{\xi})\, \boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi})) = J^{(\kappa)}(\boldsymbol{\xi}) (\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}))^{-\mathrm{T}} \hat{\boldsymbol{n}}^{(\zeta)}. \quad (2.5)$$

Furthermore, since periodic boundary conditions may be treated identically to interior interfaces through mesh connectivity, we hereinafter abuse notation by using the symbol $\partial\Omega$ to refer only to non-periodic portions of the domain boundary.

2.4 Polynomial Approximation Spaces

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

$$\mathbb{P}_{\mathcal{N}}(\hat{\Omega}) := \operatorname{span}\left\{ \hat{\Omega} \ni \boldsymbol{\xi} \mapsto \xi_1^{\alpha_1} \cdots \xi_d^{\alpha_d} : \boldsymbol{\alpha} \in \mathcal{N} \right\}, \quad (2.6)$$

[1] 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.

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 the choice of the multi-index set $\mathcal{N}$, which is standard in the theory of multivariate polynomial approximation (see, for example, Cohen and Migliorati [30] and references therein) and, importantly, ensures that $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ is closed under partial differentiation.

Assumption 2.2 *The set $\mathcal{N}$ is downward closed in the sense that for any given $\boldsymbol{\beta} \in \mathcal{N}$, every $\boldsymbol{\alpha} \in \mathbb{N}_0^d$ satisfying $\alpha_m \leq \beta_m$ 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 $\underline{U}^h(\boldsymbol{x}, t) := \underline{U}^{(h,\kappa)}((\boldsymbol{X}^{(\kappa)})^{-1}(\boldsymbol{x}), t)$ for $\boldsymbol{x} \in \Omega^{(\kappa)}$, with each component belonging to the space

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

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

Remark 2.2 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 $\mathcal{N} := \{\boldsymbol{\alpha} \in \mathbb{N}_0^d : |\boldsymbol{\alpha}| \leq p\}$ for the former and $\mathcal{N} := \{\boldsymbol{\alpha} \in \mathbb{N}_0^d : \max_{m=1}^d \alpha_m \leq p\}$ for the latter, where we have defined $|\boldsymbol{\alpha}| := \alpha_1 + \cdots + \alpha_d$. These are the natural choices for triangular/tetrahedral and quadrilateral/hexahedral elements, respectively.

Remark 2.3 Rather than constructing approximations of $\underline{U}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}), t)$ such that $\underline{U}^{(h,\kappa)}(\cdot, t) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_c}$, it is possible to instead approximate $J^{(\kappa)}(\boldsymbol{\xi})\underline{U}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}), t)$ such that $J^{(\kappa)}\underline{U}^{(h,\kappa)}(\cdot, t) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_c}$. This modification of the approximation space in (2.7) has no effect when $\boldsymbol{X}^{(\kappa)}$ is an affine mapping, 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.* [31, §3], such schemes have been observed to suffer from degraded accuracy on curvilinear meshes and are not considered further in this work.

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 the form of (2.4) and introduce some relevant notation.

2.5.1 Discontinuous Galerkin Method

Integrating (2.4) by parts over the reference element against a suitable test function and transforming the resulting facet integrals using (2.5) leads to a local weak

formulation of the system, which is given for all $\kappa \in \{1, \dots, N_e\}$ and $t \in (0, T)$ by

$$\int_{\hat{\Omega}} \Bigg(V(\boldsymbol{\xi}) \frac{\partial J^{(\kappa)}(\boldsymbol{\xi}) \underline{U}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}), t)}{\partial t} - \nabla_{\boldsymbol{\xi}} V(\boldsymbol{\xi}) \cdot \big(J^{(\kappa)}(\boldsymbol{\xi}) (\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}))^{-1} \underline{\boldsymbol{F}}(\underline{U}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}), t)) \big) \Bigg) \, \mathrm{d}\boldsymbol{\xi} + \sum_{\zeta=1}^{N_f} \int_{\hat{\Gamma}^{(\zeta)}} V(\boldsymbol{\xi}) J^{(\kappa,\zeta)}(\boldsymbol{\xi}) \big(\boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi})) \cdot \underline{\boldsymbol{F}}(\underline{U}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}), t)) \big) \, \mathrm{d}\hat{s} = \underline{0}. \quad (2.8)$$

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 $\underline{F}^* : \Upsilon \times \Upsilon \times \mathbb{S}^{d-1} \to \mathbb{R}^{N_c}$, typically corresponding to an approximate Riemann solver developed in the context of finite-volume methods (see, for example, Toro [32]), we obtain a weak-form DG discretization of the system in (2.4), in which, for all $\kappa \in \{1, \dots, N_e\}$ and $t \in (0, T)$, we seek a function $\underline{U}^{(h,\kappa)}(\cdot, t) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_c}$ satisfying

$$\int_{\hat{\Omega}} \left(V(\boldsymbol{\xi}) \frac{\partial J^{(\kappa)}(\boldsymbol{\xi}) \underline{U}^{(h,\kappa)}(\boldsymbol{\xi}, t)}{\partial t} - \nabla_{\boldsymbol{\xi}} V(\boldsymbol{\xi}) \cdot \underline{\boldsymbol{F}}^{(h,\kappa)}(\boldsymbol{\xi}, t) \right) \mathrm{d}\boldsymbol{\xi} + \sum_{\zeta=1}^{N_f} \int_{\hat{\Gamma}^{(\zeta)}} V(\boldsymbol{\xi}) J^{(\kappa,\zeta)}(\boldsymbol{\xi}) \underline{F}^{(*,\kappa,\zeta)}(\boldsymbol{\xi}, t) \, \mathrm{d}\hat{s} = \underline{0}, \quad \forall V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}). \quad (2.9)$$

In the above, the transformed flux is evaluated in terms of the numerical solution as

$$\underline{\boldsymbol{F}}^{(h,\kappa)}(\boldsymbol{\xi}, t) := J^{(\kappa)}(\boldsymbol{\xi}) (\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}))^{-1} \underline{\boldsymbol{F}}(\underline{U}^{(h,\kappa)}(\boldsymbol{\xi}, t)), \quad (2.10)$$

and the numerical flux function is evaluated on each facet as

$$\underline{F}^{(*,\kappa,\zeta)}(\boldsymbol{\xi}, t) := \underline{F}^*(\underline{U}^{(h,\kappa)}(\boldsymbol{\xi}, t), \underline{U}^{(+,\kappa,\zeta)}(\boldsymbol{\xi}, t), \boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}))), \quad (2.11)$$

where $\underline{U}^{(+,\kappa,\zeta)}(\boldsymbol{\xi}, t) \in \Upsilon$ 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 in reference coordinates, we therefore obtain $\underline{U}^{(h,\kappa)}(\cdot, 0) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_c}$ for $\kappa \in \{1, \dots, N_e\}$ as the solution to

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

Remark 2.4 As the integrals in (2.9) and (2.12) are often impractical (if not impossible) to evaluate analytically, implementations of DG methods typically employ some form of numerical integration, which we discuss in §3 within the context of the algebraic formulation presented therein.

2.5.2 Flux Reconstruction Method

While DG and FR methods both employ discontinuous polynomial approximation spaces as described in §2.4, the FR formulation 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 $\mathcal{S} := \{\boldsymbol{\xi}^{(1)}, \dots, \boldsymbol{\xi}^{(N)}\} \subset \hat{\Omega}$ of cardinality $N = N^*$ which is unisolvent for $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, implying that for any function[2] $U : \hat{\Omega} \to \mathbb{R}$, there exists a unique collocation projection $I_{\mathcal{N}} U \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ satisfying the interpolation condition

$$(I_{\mathcal{N}} U)(\boldsymbol{\xi}^{(i)}) = U(\boldsymbol{\xi}^{(i)}), \quad \forall\, i \in \{1, \dots, N\}. \tag{2.13}$$

We also require *correction functions* $\boldsymbol{G}^{(\zeta,j)} : \hat{\Omega} \to \mathbb{R}^d$ associated with the nodal set $\mathcal{S}^{(\zeta)} := \{\boldsymbol{\xi}^{(\zeta,1)}, \dots, \boldsymbol{\xi}^{(\zeta,N_\zeta)}\} \subseteq \hat{I}^{(\zeta)}$ for $\zeta \in \{1, \dots, N_f\}$, satisfying

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

where it is typically assumed that $\mathcal{S}^{(\zeta)}$ is of cardinality $N_\zeta = N_\zeta^*$ and is unisolvent for the space $\mathbb{P}_{\mathcal{N}}(\hat{I}^{(\zeta)})$. 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 $\underline{U}^{(h,\kappa)}(\cdot,t) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_c}$ as

$$\begin{aligned} &\frac{\partial I_{\mathcal{N}}(J^{(\kappa)}\underline{U}^{(h,\kappa)}(\cdot,t))}{\partial t} + \nabla_{\boldsymbol{\xi}} \cdot \Bigg(I_{\mathcal{N}} \underline{\boldsymbol{F}}^{(h,\kappa)}(\cdot,t) \\ &\quad + \sum_{\zeta=1}^{N_f} \sum_{j=1}^{N_\zeta} \Big(J^{(\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,j)}) \underline{F}^{(*,\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,j)},t) \\ &\quad - \hat{\boldsymbol{n}}^{(\zeta)} \cdot \big(I_{\mathcal{N}} \underline{\boldsymbol{F}}^{(h,\kappa)}(\cdot,t) \big)(\boldsymbol{\xi}^{(\zeta,j)}) \Big) \boldsymbol{G}^{(\zeta,j)} \Bigg) = \underline{0} \end{aligned} \tag{2.15}$$

for all $\kappa \in \{1, \dots, N_e\}$ and $t \in (0, T)$, where the operator $I_{\mathcal{N}}$ is taken to act componentwise on vector-valued or tensor-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.

Remark 2.5 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 employ 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. modal bases, as described in Appendix A.1 and used alongside nodal bases for the computations in §5) are possible.

[2] For the purposes of the present analysis, the regularity of such functions is is not of concern; we simply require data to be available at all nodes in $\mathcal{S}$ in order to form an interpolant.

2.6 Basic Algebraic Concepts in One Dimension

To introduce the main ideas which will be employed for the matrix-based analysis in §4, we first provide a simple one-dimensional example; readers familiar with the theory in [22–25] are invited to skip ahead to §3. Here, we consider the linear advection equation with a constant wave speed $a \in \mathbb{R}^+$, which is given by

$$\frac{\partial U(x,t)}{\partial t} + a\frac{\partial U(x,t)}{\partial x} = 0, \quad \forall\, (x,t) \in (x_L, x_R) \times (0,T), \tag{2.16}$$

supplemented with the initial condition $U(x,0) = U^0(x)$ and the boundary condition $U(x_L,t) = B(t)$. Considering a single element for simplicity, the solution to (2.16) is approximated as $U^h(x,t) = \sum_{i=1}^{p+1} u_i^h(t)\ell^{(i)}(x)$ in terms of a Lagrange basis $\mathcal{B} := \{\ell^{(1)}, \dots, \ell^{(p+1)}\}$ for the polynomial space $\mathbb{P}_p([x_L, x_R])$ of degree $p \in \mathbb{N}$, which is given for any set of $p+1$ (distinct) nodes $\mathcal{S} := \{x^{(1)}, \dots, x^{(p+1)}\} \subset [x_L, x_R]$ by

$$\ell^{(i)}(x) := \prod_{j=1, j\neq i}^{p+1} \frac{x - x^{(j)}}{x^{(i)} - x^{(j)}}. \tag{2.17}$$

Taking an upwind numerical flux and noting that it is sufficient to test with each basis function $\ell^{(i)} \in \mathcal{B}$ due to the linearity of all terms with respect to the test function, a one-dimensional DG method analogous to (2.9) is given for all $t \in (0,T)$ as

$$\begin{aligned} &\int_{x_L}^{x_R} \left(\ell^{(i)}(x)\frac{\partial U^h(x,t)}{\partial t} - a\frac{\mathrm{d}\ell^{(i)}(x)}{\mathrm{d}x} U^h(x,t) \right) \mathrm{d}x \\ &\quad + a\ell^{(i)}(x_R)U^h(x_R,t) - a\ell^{(i)}(x_L)B(t) = 0, \quad \forall\, i \in \{1, \dots, p+1\}. \end{aligned} \tag{2.18}$$

The above formulation may then be expressed algebraically in terms of the vector $\underline{u}^h(t) \in \mathbb{R}^{p+1}$ of nodal expansion coefficients $u_i^h(t) = U^h(x^{(i)}, t)$ as

$$\underline{\underline{M}}\frac{\mathrm{d}\underline{u}^h(t)}{\mathrm{d}t} - a\underline{\underline{D}}^{\mathrm{T}}\underline{\underline{M}}\underline{u}^h(t) + a\underline{t}_R\underline{t}_R^{\mathrm{T}}\underline{u}^h(t) - a\underline{t}_L B(t) = \underline{0}, \tag{2.19}$$

where the entries of the derivative matrix $\underline{\underline{D}} \in \mathbb{R}^{(p+1)\times(p+1)}$ are given by

$$D_{ij} := \frac{\mathrm{d}\ell^{(j)}}{\mathrm{d}x}(x^{(i)}), \tag{2.20}$$

and the vectors $\underline{t}_L := [\ell^{(1)}(x_L), \dots, \ell^{(p+1)}(x_L)]^{\mathrm{T}}$ and $\underline{t}_R := [\ell^{(1)}(x_R), \dots, \ell^{(p+1)}(x_R)]^{\mathrm{T}}$ are used to extrapolate the solution to the left and right endpoints of the domain, respectively. For the nodal schemes considered here, the entries of the mass matrix $\underline{\underline{M}} \in \mathbb{R}^{(p+1)\times(p+1)}$ in (2.19) may either be evaluated exactly as

$$M_{ij} := \int_{x_L}^{x_R} \ell^{(i)}(x)\ell^{(j)}(x)\,\mathrm{d}x, \tag{2.21}$$

or approximated (i.e. lumped) using collocated quadrature as

$$M_{ij} := \delta_{ij}\omega^{(i)}, \quad \text{where} \quad \omega^{(i)} := \int_{x_L}^{x_R} \ell^{(i)}(x)\,\mathrm{d}x, \tag{2.22}$$

completing the algebraic formulation of the DG method.

The FR approach is simplified for this problem by noting that the projection operator $I_{\mathcal{N}}$ in (2.15) vanishes and that the correction procedure need only be applied at the left boundary in the case of an upwind numerical flux and a positive wave speed. The resulting strong formulation is then given for all $t \in (0,T)$ by

$$\frac{\partial U^h(\cdot,t)}{\partial t} + a\frac{\partial}{\partial x}\left(U^h(\cdot,t) + \big(B(t) - U^h(x_L,t)\big)G_L\right) = 0, \tag{2.23}$$

where, adopting the original sign convention introduced for one-dimensional FR schemes in [8], the left correction function $G_L \in \mathbb{P}_{p+1}([x_L,x_R])$ is assumed to satisfy $G_L(x_L) = 1$ and $G_L(x_R) = 0$. Evaluating (2.23) at each node in $\mathcal{S}$ to obtain an expansion in terms of the nodal basis $\mathcal{B}$, the FR method may be expressed algebraically as

$$\frac{\mathrm{d}\underline{u}^h(t)}{\mathrm{d}t} + a\underline{\underline{D}}\,\underline{u}^h(t) + \big(aB(t) - a\underline{t}_L^{\mathrm{T}}\underline{u}^h(t)\big)\underline{g}_L' = \underline{0}, \tag{2.24}$$

where the influence of the correction function is contained within the vector $\underline{g}_L' := [(\mathrm{d}G_L/\mathrm{d}x)(x^{(1)}),\ldots,(\mathrm{d}G_L/\mathrm{d}x)(x^{(p+1)})]^{\mathrm{T}}$ consisting of the nodal values of its derivative.

The following lemma, which was proven by Carpenter and Gottlieb [33, Lemmas 2.1 and 2.2] for one-dimensional Lagrange bases including nodes at both endpoints, is essential for our analysis of the schemes in (2.19) and (2.24). More general nodal sets were considered in [20] and employed within the context of FR methods in [23].

Lemma 2.1 *Suppose that the entries of the mass matrix $\underline{\underline{M}}$ are computed exactly as in (2.21), or, alternatively, that the interpolatory quadrature rule in (2.22) is of at least degree $2p-1$ with positive weights. The mass matrix is then symmetric positive-definite (SPD), and the derivative matrix in (2.20) satisfies the SBP property*

$$\underline{\underline{M}}\,\underline{\underline{D}} + \underline{\underline{D}}^{\mathrm{T}}\underline{\underline{M}} = \underline{t}_R\underline{t}_R^{\mathrm{T}} - \underline{t}_L\underline{t}_L^{\mathrm{T}}. \tag{2.25}$$

Remark 2.6 For one-dimensional discretizations of any order employing the derivative matrix in (2.20) and the lumped mass matrix in (2.22), the SBP property is satisfied on the LGL nodes as well as the LG nodes,[3] as the corresponding collocated quadrature rules are of degree $2p-1$ and $2p+1$, respectively, and have positive weights, meeting the requirements of Lemma 2.1. If the exact mass matrix in (2.21) is used instead of (2.22), the SBP property is satisfied on *any* set of $p+1$ distinct nodes. Considering, for example, an approximation of degree $p=2$ on the LG quadrature nodes $\mathcal{S} := \{-\frac{1}{5}\sqrt{15}, 0, \frac{1}{5}\sqrt{15}\}$, one can easily verify that (2.25) is satisfied with

$$\underline{\underline{D}} := \begin{bmatrix} -\frac{1}{2}\sqrt{15} & \frac{2}{3}\sqrt{15} & --\frac{1}{6}\sqrt{15} \\ -\frac{1}{6}\sqrt{15} & 0 & \frac{1}{6}\sqrt{15} \\ \frac{1}{6}\sqrt{15} & --\frac{2}{3}\sqrt{15} & \frac{1}{2}\sqrt{15} \end{bmatrix} \quad \text{and} \quad \underline{\underline{M}} := \begin{bmatrix} \frac{5}{9} & 0 & 0 \\ 0 & \frac{8}{9} & 0 \\ 0 & 0 & \frac{5}{9} \end{bmatrix}, \tag{2.26}$$

where the extrapolation vectors are given by $\underline{t}_L := [\frac{1}{6}(5+\sqrt{15}), -\frac{2}{3}, \frac{1}{6}(5-\sqrt{15})]^{\mathrm{T}}$ and $\underline{t}_R := [\frac{1}{6}(5-\sqrt{15}), -\frac{2}{3}, \frac{1}{6}(5+\sqrt{15})]^{\mathrm{T}}$, and we note that (2.21) and (2.22) are equivalent in this case, since a collocated LG quadrature rule exactly integrates any product of two basis functions in $\mathcal{B}$.

[3] This is also true for the Legendre-Gauss-Radau (LGR) nodes, which contain one endpoint of the domain and support a quadrature rule of degree $2p$ (see, for example, [20, §7.2]).

The equivalence, conservation, and energy stability properties which will be established more generally in §4 are illustrated in the present context (i.e. considering a one-dimensional linear problem on a single element) with the following theorem.

Theorem 2.1 *Consider an FR discretization in the form of (2.24), satisfying*

$$(\underline{\underline{M}} + \underline{\underline{K}})\underline{g}'_L = -\underline{t}_L, \quad \text{where} \quad \underline{\underline{K}} := \frac{c}{x_R - x_L}\big(\underline{\underline{D}}^p\big)^{\mathrm{T}}\underline{\underline{M}}\underline{\underline{D}}^p, \tag{2.27}$$

where $\underline{\underline{M}}$ is SPD, and $c \in \mathbb{R}$ is chosen such that $\underline{\underline{M}} + \underline{\underline{K}}$ is also SPD. Provided that the SBP property in (2.25) is satisfied, such a strong-form discretization of the linear advection equation in (2.16) is equivalent to the weak-form discretization

$$\underline{\underline{\tilde{M}}}\frac{\mathrm{d}\underline{u}^h(t)}{\mathrm{d}t} - a\underline{\underline{D}}^{\mathrm{T}}\underline{\underline{M}}\underline{u}^h(t) + a\underline{t}_R\underline{t}_R^{\mathrm{T}}\underline{u}^h(t) - a\underline{t}_L B(t) = \underline{0}, \tag{2.28}$$

where the mass matrix in (2.19) is replaced by $\underline{\underline{\tilde{M}}} := \underline{\underline{M}}\underline{\underline{F}}^{-1}$, corresponding to pre-multiplication of the semi-discrete DG residual by a constant filter matrix given by $\underline{\underline{F}} := (\underline{\underline{I}} + \underline{\underline{M}}^{-1}\underline{\underline{K}})^{-1}$. Moreover, such DG and FR methods with the SBP property are discretely conservative, satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}\underline{1}^{\mathrm{T}}\underline{\underline{M}}\underline{u}^h(t) + a\underline{t}_R^{\mathrm{T}}\underline{u}^h(t) - aB(t) = 0, \tag{2.29}$$

and are energy stable with respect to the discrete norm induced by $\underline{\underline{\tilde{M}}}$, satisfying

$$\frac{\mathrm{d}}{\mathrm{d}t}(\underline{u}^h(t))^{\mathrm{T}}\underline{\underline{\tilde{M}}}\underline{u}^h(t) \le a(B(t))^2. \tag{2.30}$$

Proof Pre-multiplying (2.24) by $\underline{\underline{M}} + \underline{\underline{K}}$ and using $\underline{\underline{\tilde{M}}} = \underline{\underline{M}}(\underline{\underline{I}} + \underline{\underline{M}}^{-1}\underline{\underline{K}}) = \underline{\underline{M}} + \underline{\underline{K}}$ as well as $\underline{\underline{K}}\underline{\underline{D}} = \underline{\underline{0}}$, the latter of which follows from the form of $\underline{\underline{K}}$ in (2.27) due to the fact that differentiating a degree p polynomial $p+1$ times always yields the zero function, we obtain the strong-form filtered DG scheme recovered via FR in [13, §3.2],

$$\underline{\underline{\tilde{M}}}\frac{\mathrm{d}\underline{u}^h(t)}{\mathrm{d}t} + a\underline{\underline{S}}\underline{u}^h(t) - \big(aB(t) - a\underline{t}_L^{\mathrm{T}}\underline{u}^h(t)\big)\underline{t}_L = \underline{0}, \tag{2.31}$$

where $\underline{\underline{S}} := \underline{\underline{M}}\underline{\underline{D}}$ is the (advective) stiffness matrix. Applying the SBP property in (2.25) to the second term on the left-hand side of (2.31), we obtain

$$\underline{\underline{\tilde{M}}}\frac{\mathrm{d}\underline{u}^h(t)}{\mathrm{d}t} + a\big(\underline{t}_R\underline{t}_R^{\mathrm{T}} - \underline{t}_L\underline{t}_L^{\mathrm{T}} - \underline{\underline{D}}^{\mathrm{T}}\underline{\underline{M}}\big)\underline{u}^h(t) - \big(aB(t) - a\underline{t}_L^{\mathrm{T}}\underline{u}^h(t)\big)\underline{t}_L = \underline{0}, \tag{2.32}$$

where, observing that $a\underline{t}_L\underline{t}_L^{\mathrm{T}}\underline{u}^h(t)$ cancels from the second and third terms, we recover the weak formulation in (2.28).

To establish conservation, we pre-multiply (2.28) by $\underline{1}^{\mathrm{T}}$ and use the fact that constant functions are differentiated and interpolated exactly to obtain $\underline{1}^{\mathrm{T}}\underline{\underline{D}}^{\mathrm{T}} = \underline{0}^{\mathrm{T}}$ and $\underline{1}^{\mathrm{T}}\underline{t}_L = 1$, resulting in

$$\underline{1}^{\mathrm{T}}\underline{\underline{\tilde{M}}}\frac{\mathrm{d}\underline{u}^h(t)}{\mathrm{d}t} + a\underline{t}_R^{\mathrm{T}}\underline{u}^h(t) - aB(t) = \underline{0}. \tag{2.33}$$

Using $\underline{1}^{\mathrm{T}}\underline{\underline{K}} = \underline{0}^{\mathrm{T}}$ to obtain $\underline{1}^{\mathrm{T}}\underline{\underline{\tilde{M}}} = \underline{1}^{\mathrm{T}}\underline{\underline{M}}$ and moving the time derivative outside of the first term in (2.33) results in (2.29), where the discrete integral on the left-hand side is exact under the conditions of Lemma 2.1, such that $\underline{1}^{\mathrm{T}}\underline{\underline{M}}\underline{u}^h(t) = \int_{x_L}^{x_R} U^h(x,t)\,\mathrm{d}x$.

Energy stability follows from a similar analysis to that described in the general context of [20, §6]. Pre-multiplying the strong formulation in (2.24) by $(\underline{u}^h(t))^{\mathrm{T}}(\underline{\underline{M}} + \underline{\underline{K}})$, applying the chain rule in time, and using $\underline{\underline{K}}\,\underline{\underline{D}} = \underline{\underline{0}}$, we obtain

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\underline{u}^h(t))^{\mathrm{T}}\underline{\underline{\tilde{M}}}\,\underline{u}^h(t) = -a(\underline{u}^h(t))^{\mathrm{T}}\underline{\underline{S}}\,\underline{u}^h(t) + (\underline{u}^h(t))^{\mathrm{T}}\big(aB(t) - a\underline{t}_L^{\mathrm{T}}\underline{u}^h(t)\big)\underline{t}_L. \quad (2.34)$$

Separating the matrix $\underline{\underline{S}}$ into its symmetric and skew-symmetric parts and again applying the SBP property in (2.25) results in

$$\begin{aligned}(\underline{u}^h(t))^{\mathrm{T}}\underline{\underline{S}}\,\underline{u}^h(t) &= \frac{1}{2}(\underline{u}^h(t))^{\mathrm{T}}\big(\underline{t}_R\underline{t}_R^{\mathrm{T}} - \underline{t}_L\underline{t}_L^{\mathrm{T}}\big)\underline{u}^h(t)\\ &= \frac{1}{2}\big((\underline{t}_R^{\mathrm{T}}\underline{u}^h(t))^2 - (\underline{t}_L^{\mathrm{T}}\underline{u}^h(t))^2\big),\end{aligned} \quad (2.35)$$

which, when applied to the first term on the right-hand side of (2.34), leads to

$$\begin{aligned}\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\underline{u}^h(t))^{\mathrm{T}}\underline{\underline{\tilde{M}}}\,\underline{u}^h(t) &= -\frac{a}{2}\big((\underline{t}_L^{\mathrm{T}}\underline{u}^h(t))^2 - 2\underline{t}_L^{\mathrm{T}}\underline{u}^h(t)B(t)\big) - \frac{a}{2}(\underline{t}_R^{\mathrm{T}}\underline{u}^h(t))^2\\ &= \frac{a}{2}(B(t))^2 - \frac{a}{2}(\underline{t}_L^{\mathrm{T}}\underline{u}^h(t) - B(t))^2 - \frac{a}{2}(\underline{t}_R^{\mathrm{T}}\underline{u}^h(t))^2,\end{aligned} \quad (2.36)$$

where we have completed the square in order to obtain the second equality. Noting the signs of each term in (2.36), we therefore have the bound in (2.30). □

Remark 2.7 Assuming that the entries of the mass matrix are evaluated exactly as in (2.21) and requiring the correction function to satisfy (2.27), we recover the one-parameter family of VCJH FR methods introduced in [10]. More generally, energy-stable FR schemes can be constructed by finding a matrix $\underline{\underline{K}}$ such that the SBP property is satisfied when the matrix $\underline{\underline{M}}$ in (3.11) is replaced by $\underline{\underline{M}} + \underline{\underline{K}}$, which must be SPD to produce an energy estimate. Vincent *et al.* [34, Theorem 1] achieve this by requiring $\underline{\underline{K}}$ to be symmetric, $\underline{\underline{M}} + \underline{\underline{K}}$ to be SPD, and $\underline{\underline{K}}\,\underline{\underline{D}}$ to be skew-symmetric, leading to an extended range of energy-stable schemes.

Theorem 2.1 provides a simple illustration of the utility of the SBP property for the analysis of high-order discretizations of conservation laws, in which the equivalence, stability, and conservation properties of various schemes are established algebraically as consequences of the matrix properties of the operators constituting such discretizations. This approach enables an analysis that is highly general (as many schemes are amenable to matrix-based formulations with similar algebraic properties) yet leads to simple proofs requiring very few additional theoretical abstractions or idealizations (as the analysis is based directly on the computational form of the schemes and explicitly considers the influence of numerical integration). The basic concepts considered in this subsection apply much more broadly, however, extending to discretizations of multidimensional and nonlinear problems as well as to modal and over-integrated formulations. Such generalizations are the focus of the remainder of this paper.

3 Algebraic Formulation

3.1 Approximation on the Reference Element

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.

3.1.1 Discrete Inner Products

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 over-integration, which in the present context refers to the evaluation of the semi-discrete residual for each element using $N > N^*$ nodes to compute the volume terms and/or $N_\zeta > N_\zeta^*$ nodes to compute the facet terms. The L^2 inner products on the reference element and each of its facets may then be approximated as

$$\int_{\hat{\Omega}} U(\boldsymbol{\xi})V(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \approx \big\langle U, V\big\rangle_W := \sum_{i=1}^{N}\sum_{j=1}^{N} U(\boldsymbol{\xi}^{(i)})W_{ij}V(\boldsymbol{\xi}^{(j)}) \tag{3.1}$$

and

$$\int_{\hat{\Gamma}^{(\zeta)}} U(\boldsymbol{\xi})V(\boldsymbol{\xi})\,\mathrm{d}\hat{s} \approx \big\langle U, V\big\rangle_{B^{(\zeta)}} := \sum_{i=1}^{N_\zeta}\sum_{j=1}^{N_\zeta} U(\boldsymbol{\xi}^{(\zeta,i)})B_{ij}^{(\zeta)}V(\boldsymbol{\xi}^{(\zeta,j)}), \tag{3.2}$$

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 $\langle\cdot,\cdot\rangle_W$ and $\langle\cdot,\cdot\rangle_{B^{(\zeta)}}$ 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 $m \in \{1,\ldots,d\}$ as

$$\underbrace{\int_{\hat{\Omega}} U(\boldsymbol{\xi})\frac{\partial V(\boldsymbol{\xi})}{\partial \xi_m}\,\mathrm{d}\boldsymbol{\xi}}_{\approx \langle U, \partial V/\partial\xi_m\rangle_W} + \underbrace{\int_{\hat{\Omega}} \frac{\partial U(\boldsymbol{\xi})}{\partial \xi_m}V(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi}}_{\approx \langle \partial U/\partial\xi_m, V\rangle_W} = \sum_{\zeta=1}^{N_f}\underbrace{\int_{\hat{\Gamma}^{(\zeta)}} U(\boldsymbol{\xi})V(\boldsymbol{\xi})\hat{n}_m^{(\zeta)}\,\mathrm{d}\hat{s}}_{\approx \hat{n}_m^{(\zeta)}\langle U,V\rangle_{B^{(\zeta)}}}, \tag{3.3}$$

where we make the following assumption regarding the above approximations.

Assumption 3.1 *For any $m \in \{1,\ldots,d\}$, 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*

$$\left\langle U, \frac{\partial V}{\partial \xi_m}\right\rangle_W + \left\langle \frac{\partial U}{\partial \xi_m}, V\right\rangle_W = \sum_{\zeta=1}^{N_f}\hat{n}_m^{(\zeta)}\big\langle U, V\big\rangle_{B^{(\zeta)}}. \tag{3.4}$$

3.1.2 Polynomial Bases

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

$$V_{ij} := \phi^{(j)}(\boldsymbol{\xi}^{(i)}) \quad \text{and} \quad R_{ij}^{(\zeta)} := \phi^{(j)}(\boldsymbol{\xi}^{(\zeta,i)}), \tag{3.5}$$

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 bases, we make the following additional assumption.

[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.

Assumption 3.2 *The matrix $\underline{\underline{V}}$ is of rank N^*. Additionally, $\underline{\underline{W}}$ is SPD, and $\underline{\underline{B}}^{(\zeta)}$ is (at least) symmetric positive-semidefinite (SPSD) for all $\zeta \in \{1, \dots, N_f\}$.*

Remark 3.1 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

$$U(\boldsymbol{\xi}) = \sum_{i=1}^{N^*} \tilde{u}_i \phi^{(i)}(\boldsymbol{\xi}), \tag{3.6}$$

which may be evaluated on $\mathcal{S}$ and $\mathcal{S}^{(\zeta)}$ through the matrix-vector products

$$\underline{\underline{V}}\underline{\tilde{u}} = \left[U(\boldsymbol{\xi}^{(1)}), \dots, U(\boldsymbol{\xi}^{(N)})\right]^{\mathrm{T}} \quad \text{and} \quad \underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}} = \left[U(\boldsymbol{\xi}^{(\zeta,1)}), \dots, U(\boldsymbol{\xi}^{(\zeta,N_\zeta)})\right]^{\mathrm{T}}. \tag{3.7}$$

The positive-definiteness of the reference mass matrix with entries consisting of the discrete inner products $\langle \phi^{(i)}, \phi^{(j)} \rangle_W$ 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. □

3.1.3 Projection Operators

Associated with the discrete inner product in (3.1) is an orthogonal projection operator approximating the L^2 projection of a function onto $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, 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 $\Pi_{\mathcal{N}} U \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ of a given function $U : \hat{\Omega} \to \mathbb{R}$ with respect to the discrete inner product in (3.1), satisfying*

$$\left\langle V, \Pi_{\mathcal{N}} U - U \right\rangle_W = 0, \quad \forall\, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}), \tag{3.8}$$

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

$$(\Pi_{\mathcal{N}} U)(\boldsymbol{\xi}) = \sum_{i=1}^{N^*} \left(\sum_{j=1}^{N} P_{ij} U(\boldsymbol{\xi}^{(j)}) \right) \phi^{(i)}(\boldsymbol{\xi}), \tag{3.9}$$

where we define $\underline{\underline{P}} := \underline{\underline{M}}^{-1} \underline{\underline{V}}^{\mathrm{T}} \underline{\underline{W}}$. The projection then satisfies $\Pi_{\mathcal{N}} U = U$ 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 (3.6) into (3.9) 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 (3.9) 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 $\underline{\underline{V}}\,\underline{\underline{P}} = \underline{\underline{V}}\,\underline{\underline{V}}^{-1} = \underline{\underline{I}}$ in the case of $N = N^*$. □

3.1.4 Summation-by-Parts Property

Recalling that Assumption 2.2 implies that the reference approximation space in (2.6) is closed under partial differentiation, the operator $\partial/\partial\xi_m$ on $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ can be represented algebraically in 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 (3.6), the corresponding derivative may be expanded as

$$\frac{\partial U(\boldsymbol{\xi})}{\partial \xi_m} = \sum_{i=1}^{N^*}\left(\sum_{j=1}^{N^*} D_{ij}^{(m)}\tilde{u}_j\right)\phi^{(i)}(\boldsymbol{\xi}). \tag{3.10}$$

The following lemma establishes that the derivative matrix given by (3.10) satisfies a multidimensional generalization of the SBP property in (2.25), which serves as a discrete analogue of the integration-by-parts relation in (3.3).

Lemma 3.3 *Under Assumptions 2.2 and 3.2, the derivative matrices given as in (3.10) in terms of any basis $\mathcal{B}$ for $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ satisfy the multidimensional SBP property*

$$\underline{\underline{M}}\,\underline{\underline{D}}^{(m)} + \left(\underline{\underline{D}}^{(m)}\right)^{\mathrm{T}}\underline{\underline{M}} = \sum_{\zeta=1}^{N_f}\hat{n}_m^{(\zeta)}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}, \quad \forall\, m \in \{1,\dots,d\}. \tag{3.11}$$

Proof As noted in [25, Eq. (26)], the result follows directly from expanding the functions $U, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ in (3.4) in terms of $\mathcal{B}$, using (3.10) to obtain expansions of the partial derivatives $\partial U/\partial\xi_m$, and $\partial V/\partial\xi_m$, which belong to $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ as a consequence of Assumption 2.2, and expressing the discrete inner products in matrix form. □

Remark 3.2 As the operators in (3.11) 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, Eq. (3.5)]), 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 (3.6) in terms of the expansion coefficients $\underline{\tilde{u}}^{(h,\kappa,\ell)}(t) \in \mathbb{R}^{N^*}$ as $U_\ell^{(h,\kappa)}(\boldsymbol{\xi},t) = \sum_{i=1}^{N^*}\tilde{u}_i^{(h,\kappa,\ell)}(t)\phi^{(i)}(\boldsymbol{\xi})$. In the particular case of the DG scheme in (2.9), we may apply such an expansion

to the numerical solution, test with each basis function $\phi_i \in \mathcal{B}$, and approximate the volume and facet integrals using (3.1) and (3.2), respectively, resulting in an inner-product formulation given by

$$\sum_{j=1}^{N^*} \left\langle \phi^{(i)}, J^{(\kappa)}\phi^{(j)} \right\rangle_W \frac{\mathrm{d}\tilde{u}_j^{(h,\kappa,\ell)}(t)}{\mathrm{d}t} - \sum_{m=1}^{d} \left\langle \frac{\partial \phi^{(i)}}{\partial \xi_m}, F_{\ell,m}^{(h,\kappa)}(\cdot,t) \right\rangle_W + \sum_{\zeta=1}^{N_f} \left\langle \phi^{(i)}, J^{(\kappa,\zeta)} F_\ell^{(*,\kappa,\zeta)}(\cdot,t) \right\rangle_{B^{(\zeta)}} = 0 \tag{3.12}$$

for all $\kappa \in \{1,\dots,N_e\}$, $\ell \in \{1,\dots,N_c\}$, $i \in \{1,\dots,N^*\}$, and $t \in (0,T)$. Defining the diagonal matrices

$$\underline{\underline{J}}^{(\kappa)} := \mathrm{diag}\left(J^{(\kappa)}(\boldsymbol{\xi}^{(1)}),\dots,J^{(\kappa)}(\boldsymbol{\xi}^{(N)})\right) \tag{3.13}$$

and

$$\underline{\underline{J}}^{(\kappa,\zeta)} := \mathrm{diag}\left(J^{(\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,1)}),\dots,J^{(\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,N_\zeta)})\right), \tag{3.14}$$

the local mass matrix with entries given by $\langle \phi^{(i)}, J^{(\kappa)}\phi^{(j)}\rangle_W$ for all pairs of basis functions $\phi^{(i)}, \phi^{(j)} \in \mathcal{B}$ may be expressed as $\underline{\underline{M}}^{(\kappa)} := \underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}$. An algebraic formulation of (3.12) may then be obtained using the operators introduced in §3.1 as

$$\underline{\underline{M}}^{(\kappa)} \frac{\mathrm{d}\underline{\tilde{u}}^{(h,\kappa,\ell)}(t)}{\mathrm{d}t} - \sum_{m=1}^{d} \left(\underline{\underline{D}}^{(m)}\right)^{\mathrm{T}} \underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{f}^{(h,\kappa,m,\ell)}(t) + \sum_{\zeta=1}^{N_f} \left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t) = \underline{0} \tag{3.15}$$

for all $\kappa \in \{1,\dots,N_e\}$, $\ell \in \{1,\dots,N_c\}$, and $t \in (0,T)$, where we define

$$\underline{f}^{(h,\kappa,m,\ell)}(t) := \left[F_{\ell,m}^{(h,\kappa)}(\boldsymbol{\xi}^{(1)},t),\dots,F_{\ell,m}^{(h,\kappa)}(\boldsymbol{\xi}^{(N)},t)\right]^{\mathrm{T}} \tag{3.16}$$

and

$$\underline{f}^{(*,\kappa,\zeta,\ell)}(t) := \left[F_\ell^{(*,\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,1)},t),\dots,F_\ell^{(*,\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,N_\zeta)},t)\right]^{\mathrm{T}}. \tag{3.17}$$

Remark 3.3 In order to obtain the vectors $\underline{f}^{(h,\kappa,m,\ell)}(t)$ and $\underline{f}^{(*,\kappa,\zeta,\ell)}(t)$, the fluxes in (2.10) and (2.11) may be evaluated pointwise following the pre-multiplication of $\underline{\tilde{u}}^{(h,\kappa,\ell)}(t)$ by $\underline{\underline{V}}$ and $\underline{\underline{R}}^{(\zeta)}$ as in (3.7) 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 DG residual may be obtained from (3.15) 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 $\kappa \in \{1,\dots,N_e\}$.*

Remark 3.4 Assumption 3.3 is satisfied automatically as a consequence of Assumptions 2.1 and 3.2 when $N = N^*$, in which case $\underline{\underline{M}}^{(\kappa)}$ is a product of invertible matrices. Alternatively, Assumptions 2.1 and 3.2 lead to Assumption 3.3 being satisfied in the general case of $N \geq N^*$ if we additionally require $\underline{\underline{W}}$ to be diagonal, as the resulting $\underline{\underline{M}}^{(\kappa)}$ can be shown to be SPD in a similar manner to the proof of Lemma 3.1.

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 (3.8) as

$$\left\langle V, J^{(\kappa)}\left[\underline{U}^{(h,\kappa)}(\cdot,0) - \underline{U}^0 \circ \boldsymbol{X}^{(\kappa)}\right]_\ell \right\rangle_W = 0, \quad \forall\, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}), \tag{3.18}$$

for all $\kappa \in \{1,\dots,N_e\}$ and $\ell \in \{1,\dots,N_c\}$. A unique solution $\underline{U}^{(h,\kappa)}(\cdot,0) \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^{N_c}$ to (3.18) then exists under Assumption 3.3, with coefficients given by

$$\underline{\tilde{u}}^{(h,\kappa,\ell)}(0) = \left(\underline{\underline{M}}^{(\kappa)}\right)^{-1} \underline{\underline{V}}^{\mathrm{T}} \underline{\underline{W}}\, \underline{\underline{J}}^{(\kappa)} \left[U_\ell^0(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(1)})), \dots, U_\ell^0(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(N)}))\right]^{\mathrm{T}}. \tag{3.19}$$

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 $H^{(\zeta,j)}(\boldsymbol{\xi}) := \nabla_{\boldsymbol{\xi}} \cdot \boldsymbol{G}^{(\zeta,j)}(\boldsymbol{\xi})$, which belongs to $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ as a result of the first condition in (2.14), and using the projection operator $\mathit{\Pi}_{\mathcal{N}}$ as opposed to $I_{\mathcal{N}}$ in order to allow for over-integration, the FR method in (2.15) may be expressed more generally as

$$\begin{aligned} &\frac{\partial \mathit{\Pi}_{\mathcal{N}}(J^{(\kappa)}\underline{U}^{(h,\kappa)}(\cdot,t))}{\partial t} + \nabla_{\boldsymbol{\xi}} \cdot \mathit{\Pi}_{\mathcal{N}} \underline{\boldsymbol{F}}^{(h,\kappa)}(\cdot,t) \\ &\quad + \sum_{\zeta=1}^{N_f} \sum_{j=1}^{N_\zeta} \Bigg(J^{(\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,j)}) \underline{F}^{(*,\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,j)},t) \\ &\quad - \hat{\boldsymbol{n}}^{(\zeta)} \cdot \left(\mathit{\Pi}_{\mathcal{N}} \underline{\boldsymbol{F}}^{(h,\kappa)}(\cdot,t)\right)(\boldsymbol{\xi}^{(\zeta,j)}) \Bigg) H^{(\zeta,j)} = \underline{0}. \end{aligned} \tag{3.20}$$

An algebraic formulation then follows from expanding all functions belonging to $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ appearing in (3.20) in terms of $\mathcal{B}$ and using the operators in §3.1 to obtain

$$\begin{aligned} &\underline{\underline{P}}\, \underline{\underline{J}}^{(\kappa)} \underline{\underline{V}} \frac{\mathrm{d}\underline{\tilde{u}}^{(h,\kappa,\ell)}(t)}{\mathrm{d}t} + \sum_{m=1}^{d} \underline{\underline{D}}^{(m)} \underline{\underline{P}}\, \underline{f}^{(h,\kappa,m,\ell)}(t) \\ &\quad + \sum_{\zeta=1}^{N_f} \underline{\underline{L}}^{(\zeta)} \left(\underline{\underline{J}}^{(\kappa,\zeta)} \underline{f}^{(*,\kappa,\zeta,\ell)}(t) - \sum_{m=1}^{d} \hat{n}_m^{(\zeta)} \underline{\underline{R}}^{(\zeta)} \underline{\underline{P}}\, \underline{f}^{(h,\kappa,m,\ell)}(t) \right) = \underline{0} \end{aligned} \tag{3.21}$$

for all $\kappa \in \{1,\dots,N_e\}$, $\ell \in \{1,\dots,N_c\}$, and $t \in (0,T)$. The correction fields for the facet $\hat{\Gamma}^{(\zeta)}$ are encoded within the lifting matrix $\underline{\underline{L}}^{(\zeta)} \in \mathbb{R}^{N^* \times N_\zeta}$, which satisfies

$$H^{(\zeta,j)}(\boldsymbol{\xi}) = \sum_{i=1}^{N^*} L_{ij}^{(\zeta)} \phi^{(i)}(\boldsymbol{\xi}), \quad \forall\, j \in \{1,\dots,N_\zeta\}, \tag{3.22}$$

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 space dimensions.

Remark 3.5 Noting that $\underline{\underline{M}}\underline{\underline{P}}\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}} = \underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}} = \underline{\underline{M}}^{(\kappa)}$, we see for general (i.e. including non-collocated) nodal sets that the matrix pre-multiplying the time derivative in (3.21) has an inverse given by $(\underline{\underline{P}}\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}})^{-1} = (\underline{\underline{M}}^{(\kappa)})^{-1}\underline{\underline{M}}$ under Assumption 3.3, ensuring that the 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 (3.21), 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 polynomial space $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ satisfying Assumption 2.2, suppose that there exist correction functions $\boldsymbol{G}^{(\zeta,j)}$ and correction fields $H^{(\zeta,j)}$ parametrized by the coefficients $\mathcal{C} := \{c_{\boldsymbol{\alpha}}\}_{\boldsymbol{\alpha}\in\mathcal{M}}$ with $\mathcal{M} \subset \mathcal{N}$ and $c_{\boldsymbol{\alpha}} \in \mathbb{R}$ for all $\boldsymbol{\alpha} \in \mathcal{M}$, satisfying the conditions in (2.14) as well as*

$$\int_{\hat{\Omega}} \left(\boldsymbol{G}^{(\zeta,j)}(\boldsymbol{\xi}) \cdot \nabla_{\boldsymbol{\xi}} V(\boldsymbol{\xi}) - \sum_{\boldsymbol{\alpha}\in\mathcal{M}} \frac{c_{\boldsymbol{\alpha}}}{|\hat{\Omega}|} D^{\boldsymbol{\alpha}} V(\boldsymbol{\xi}) D^{\boldsymbol{\alpha}} H^{(\zeta,j)}(\boldsymbol{\xi}) \right) \mathrm{d}\boldsymbol{\xi} = 0 \tag{3.23}$$

for all $V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$, $\zeta \in \{1,\dots,N_f\}$, and $j \in \{1,\dots,N_\zeta\}$, where we define the differential operator $D^{\boldsymbol{\alpha}} := \partial^{|\alpha|}/\partial\xi_1^{\alpha_1}\cdots\partial\xi_d^{\alpha_d}$. Then, given any basis $\mathcal{B}$ for $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ and discrete inner products defined as in (3.1) and (3.2), satisfying

$$\left\langle U, V \right\rangle_W = \int_{\hat{\Omega}} U(\boldsymbol{\xi}) V(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}, \quad \forall U, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}), \tag{3.24}$$

and

$$\left\langle U, V \right\rangle_{B^{(\zeta)}} = \int_{\hat{\Gamma}^{(\zeta)}} U(\boldsymbol{\xi}) V(\boldsymbol{\xi}) \, \mathrm{d}\hat{\mathrm{s}}, \quad \forall U, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Gamma}^{(\zeta)}), \tag{3.25}$$

such correction fields may be expanded as in (3.22) for any $\zeta \in \{1,\dots,N_f\}$ with

$$\underline{\underline{L}}^{(\zeta)} := (\underline{\underline{M}} + \underline{\underline{K}})^{-1} \left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \quad \textit{and} \quad \underline{\underline{K}} := \sum_{\boldsymbol{\alpha}\in\mathcal{M}} \frac{c_{\boldsymbol{\alpha}}}{|\hat{\Omega}|} \left(\underline{\underline{D}}^{\boldsymbol{\alpha}}\right)^{\mathrm{T}} \underline{\underline{M}}\,\underline{\underline{D}}^{\boldsymbol{\alpha}}, \tag{3.26}$$

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

Proof Applying integration by parts to the first term of the integrand in (3.23) 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 (3.24) and (3.25) as well as the fact that $H^{(\zeta,j)}$ is defined as the divergence of $\boldsymbol{G}^{(\zeta,j)}$ to obtain

$$\left\langle V, H^{(\zeta,j)} \right\rangle_W + \sum_{\boldsymbol{\alpha}\in\mathcal{M}} \frac{c_{\boldsymbol{\alpha}}}{|\hat{\Omega}|} \left\langle D^{\boldsymbol{\alpha}} V, D^{\boldsymbol{\alpha}} H^{(\zeta,j)} \right\rangle_W = \sum_{\eta=1}^{N_f} \left\langle V, \boldsymbol{G}^{(\zeta,j)} \cdot \hat{\boldsymbol{n}}^{(\eta)} \right\rangle_{B^{(\eta)}} \tag{3.27}$$

for all $V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$. Since all terms in (3.27) are linear with respect to V, we can test with $\phi^{(k)} \in \mathcal{B}$ and apply the expansion in (3.22) to $H^{(\zeta,j)} \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ in order to obtain

$$\begin{aligned}\sum_{i=1}^{N^*}\left(\left\langle\phi^{(k)},\phi^{(i)}\right\rangle_W+\sum_{\boldsymbol{\alpha}\in\mathcal{M}}\frac{c_{\boldsymbol{\alpha}}}{|\hat{\Omega}|}\left\langle D^{\boldsymbol{\alpha}}\phi^{(k)},D^{\boldsymbol{\alpha}}\phi^{(i)}\right\rangle_W\right)L_{ij}^{(\zeta)}\\ =\sum_{\eta=1}^{N_f}\left\langle\phi^{(k)},\boldsymbol{G}^{(\zeta,j)}\cdot\hat{\boldsymbol{n}}^{(\eta)}\right\rangle_{B^{(\eta)}},\quad\forall\,k\in\{1,\dots,N^*\}.\end{aligned}\tag{3.28}$$

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

$$\sum_{i=1}^{N^*}(M_{ki}+K_{ki})L_{ij}^{(\zeta)}=\sum_{i=1}^{N_\zeta}R_{ik}^{(\zeta)}B_{ij}^{(\zeta)},\quad\forall\,k\in\{1,\dots,N^*\},\tag{3.29}$$

for which a unique solution in the form of (3.26) 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 (3.23) for all $j \in \{1,\dots,N_\zeta\}$ and $\zeta \in \{1,\dots,N_f\}$, the correction fields are then fully specified under the present assumptions by defining the lifting matrices as in (3.26). □

Remark 3.6 While the conditions on the correction functions in Theorem 3.1 imply the existence of correction fields given by (3.22) 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 (3.22) with a unique correction function $\boldsymbol{G}^{(\zeta,j)}$ 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 (3.21) 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 in (3.24) and (3.25) 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, which may be satisfied under slightly weaker accuracy requirements. Discrete inner products satisfying such assumptions may therefore be used within the present framework to construct lifting matrices in the form of (3.26) even when the conditions of Theorem 3.1 are not met, although (as discussed in [23, §3.6] for collocated LGL quadrature) the corresponding correction fields would differ in such a case from those of the standard VCJH scheme associated with a given set of coefficients $\mathcal{C}$.

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

[5] Recalling the formulation in (3.20), 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 $\boldsymbol{G}^{(\zeta,j)}$, that appear explicitly in the implementation of a multidimensional FR scheme.

Assumption 3.4 *The correction fields are given in terms of (3.22) with the matrix* $\underline{\underline{L}}^{(\zeta)}$ *defined as in (3.26), 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 (3.23), the first condition in Assumption 3.4 amounts to choosing $\mathcal{M}$ such that the second term of the integrand in (3.23) is constant. In such a case, we recover the fundamental assumption for the energy stability analysis of the VCJH schemes in [10, Eqs. (3.31), (3.32)] and [11, Eq. (5.37)],

$$\int_{\hat{\Omega}} \boldsymbol{G}^{(\zeta,j)}(\boldsymbol{\xi}) \cdot \nabla_{\boldsymbol{\xi}} V(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi} = \sum_{\boldsymbol{\alpha} \in \mathcal{M}} c_{\boldsymbol{\alpha}} (D^{\boldsymbol{\alpha}} V)(D^{\boldsymbol{\alpha}} H^{(\zeta,j)}), \quad \forall \, V \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega}), \tag{3.30}$$

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

3.5 Nodal Formulations and Multidimensional SBP-SAT Discretizations

In the case of a collocated nodal basis, which, as discussed in Remark 3.3, corresponds to the choice of $\mathcal{S} = \tilde{\mathcal{S}}$, where $\tilde{\mathcal{S}}$ contains the nodal points associated with a Lagrange basis (see, for example, Appendix A.2), we obtain $\underline{\underline{V}} = \underline{\underline{P}} = \underline{\underline{I}}$ and $\underline{\underline{M}} = \underline{\underline{W}}$. Employing such a simplification, which was also used for the one-dimensional discretizations presented in §2.6, the formulations in (3.15) and (3.21) are given in terms of the vector $\underline{u}^{(h,\kappa,\ell)}(t) \in \mathbb{R}^N$ of nodal expansion coefficients $u_i^{(h,\kappa,\ell)}(t) = U_\ell^{(h,\kappa)}(\boldsymbol{\xi}^{(i)}, t)$ as

$$\begin{aligned} \underline{\underline{M}}\underline{\underline{J}}^{(\kappa)} \frac{\mathrm{d}\underline{u}^{(h,\kappa,\ell)}(t)}{\mathrm{d}t} - \sum_{m=1}^{d} \left(\underline{\underline{D}}^{(m)}\right)^{\mathrm{T}} \underline{\underline{M}}\, \underline{f}^{(h,\kappa,m,\ell)}(t) \\ + \sum_{\zeta=1}^{N_f} \left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{J}}^{(\kappa,\zeta)} \underline{f}^{(*,\kappa,\zeta,\ell)}(t) = \underline{0} \end{aligned} \tag{3.31}$$

and

$$\begin{aligned} \underline{\underline{J}}^{(\kappa)} \frac{\mathrm{d}\underline{u}^{(h,\kappa,\ell)}(t)}{\mathrm{d}t} + \sum_{m=1}^{d} \underline{\underline{D}}^{(m)} \underline{f}^{(h,\kappa,m,\ell)}(t) \\ + \sum_{\zeta=1}^{N_f} \underline{\underline{L}}^{(\zeta)} \left(\underline{\underline{J}}^{(\kappa,\zeta)} \underline{f}^{(*,\kappa,\zeta,\ell)}(t) - \sum_{m=1}^{d} \hat{n}_m^{(\zeta)} \underline{\underline{R}}^{(\zeta)} \underline{f}^{(h,\kappa,m,\ell)}(t) \right) = \underline{0}, \end{aligned} \tag{3.32}$$

respectively, where it is usually assumed that $N_\zeta = N_\zeta^*$. The formulation in (3.31) then recovers the weak form of the DGSEM when a quadrature-based approximation is used (see, for example, Appendix B.1) and the weak form of the nodal DG approach described in [38] when a collocation-based approximation is used (see, for example, Appendix B.2). The former employs a lumped (i.e. diagonal) mass matrix, whereas the latter employs an exact mass matrix, as given in one dimension by (2.22) and (2.21), respectively, which differ when the quadrature is of insufficient accuracy to exactly integrate all products of two basis functions. Referring to §4.1 for a discussion

of the conditions under which the strong and weak formulations are equivalent, we also note that the standard VCJH FR methods are recovered from (3.32) under Assumption 3.4 when a collocation-based approximation is used.

Moreover, we are not restricted to the use of discretizations employing polynomial expansions; the theory in §4 immediately extends to methods using *any* nodal approximation $\underline{\underline{D}}^{(m)}$ of each partial derivative $\partial/\partial\xi_m$ satisfying the matrix-based SBP property in (3.11), which is typically supplemented with polynomial exactness conditions (c.f. [21, Definition 2.1]) to ensure high-order accuracy. Within the SBP context, boundary or interface penalties such as the third term in (3.32) are referred to as *simultaneous approximation terms* (SATs). Such terms are added to strong-form discretizations employing SBP operators in order to weakly enforce boundary or interface conditions in a stable and conservative manner, and were introduced in the finite-difference context by Carpenter *et al.* [41], extending the penalty approach developed for spectral methods by Funaro and Gottlieb [42]. A more detailed discussion of SATs for multidimensional SBP operators as well as an analysis of the resulting SBP-SAT schemes for variable-coefficient advection problems in curvilinear coordinates is provided by Del Rey Fernández *et al.* [43].

3.6 Connection Between FR Methods and Strong-Form Filtered DG Schemes

Consider an FR method in the form of (3.21) for which $\underline{\underline{L}}^{(\zeta)}$ is defined as in (3.26) with $\underline{\underline{K}} = \underline{\underline{0}}$, as is obtained when the conditions of Theorem 3.1 are satisfied with $c_{\boldsymbol{\alpha}} = 0$ for all $\boldsymbol{\alpha} \in \mathcal{M}$. Pre-multiplying such a formulation by the reference mass matrix $\underline{\underline{M}}$ and recalling Remark 3.5, we recover a DG method in strong conservation form (see, for example, [10, §3.5.1], [11, §5.4], [12, §5], and [16, §2.3]) given by

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

where we define $\underline{\underline{S}}^{(m)} := \underline{\underline{M}}\,\underline{\underline{D}}^{(m)}$ and note that the components of the flux in (2.10) are projected onto $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ prior to differentiating on the reference element as well as prior to computing the normal trace of the flux on each facet through extrapolation.

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 exactly using the chain rule and integrating or projecting afterwards, and to compute the facet terms by evaluating the normal trace 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 (3.21) 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.

[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 $\zeta \in \{1, \ldots, N_f\}$.

For more general choices of correction functions satisfying (3.30), it was demonstrated in [13] for one-dimensional collocated discretizations (and extended to certain multidimensional schemes in [12, §5] and [16, §2.4]) that the resulting FR formulation recovers a linearly filtered DG method in the sense that the semi-discrete FR residual may be recovered by pre-multiplying the strong-form DG residual, which we obtain through isolating the time derivative in (3.33), 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 this equivalence within the context of the present framework.

Lemma 3.4 *Under Assumptions 2.2, 3.2, 3.3, and 3.4, the FR method in (3.21) is equivalent to the strong-form filtered DG scheme given for all $\kappa \in \{1,\dots,N_e\}$, $\ell \in \{1,\dots,N_c\}$, and $t \in (0,T)$ by*

$$\begin{aligned}&\tilde{\underline{\underline{M}}}^{(\kappa)}\frac{\mathrm{d}\tilde{\underline{u}}^{(h,\kappa,\ell)}(t)}{\mathrm{d}t}+\sum_{m=1}^{d}\underline{\underline{S}}^{(m)}\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)\\&\quad+\sum_{\zeta=1}^{N_f}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\left(\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t)-\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)\right)=\underline{0},\end{aligned}\tag{3.34}$$

where $\tilde{\underline{\underline{M}}}^{(\kappa)} := \underline{\underline{M}}^{(\kappa)}(\underline{\underline{F}}^{(\kappa)})^{-1}$ is defined in terms of the filter matrix on the physical element $\underline{\underline{F}}^{(\kappa)} := (\underline{\underline{P}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}})\underline{\underline{F}}(\underline{\underline{P}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}})^{-1}$ for a reference filter $\underline{\underline{F}} := (\underline{\underline{I}} + \underline{\underline{M}}^{-1}\underline{\underline{K}})^{-1}$.

Proof Pre-multiplying (3.21) 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 $\underline{\underline{M}}^{(\kappa)} = \underline{\underline{M}}\,\underline{\underline{P}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}$ to obtain

$$(\underline{\underline{M}} + \underline{\underline{K}})(\underline{\underline{P}}\,\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}) = \underline{\underline{M}}^{(\kappa)}\left(\underline{\underline{F}}^{(\kappa)}\right)^{-1} = \tilde{\underline{\underline{M}}}^{(\kappa)},\tag{3.35}$$

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

$$\begin{aligned}&\tilde{\underline{\underline{M}}}^{(\kappa)}\frac{\mathrm{d}\tilde{\underline{u}}^{(h,\kappa,\ell)}(t)}{\mathrm{d}t}+\sum_{m=1}^{d}\underline{\underline{S}}^{(m)}\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)+\sum_{m=1}^{d}\underline{\underline{K}}\,\underline{\underline{D}}^{(m)}\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)\\&\quad+\sum_{\zeta=1}^{N_f}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\left(\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t)-\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)\right)=\underline{0}.\end{aligned}\tag{3.36}$$

Noting that the third term on the left-hand side of (3.36) 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 recover (3.34), which is identical to (3.33) with the exception of the mass matrix. □

Remark 3.9 If $\underline{\underline{K}}\,\underline{\underline{D}}^{(m)} \neq \underline{\underline{0}}$, in which case the third term of (3.36) does not vanish, or if the lifting matrix takes a different form from that in (3.26), the resulting FR discretization has no known equivalence to a filtered or unfiltered DG method. Moreover, as shown in [16, §2.4], the filtered DG scheme recovered in (3.34) may be expressed equivalently with the filter matrix only pre-multiplying the facet contributions to the residual. In this work, we retain the filter on the entire semi-discrete residual (i.e. as a modification to the mass matrix, whose inverse must pre-multiply all terms in order to isolate the time derivative), which facilitates 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 PDE to be obtained from the strong form, the SBP property in (3.11) may be used analogously to transform a strong-form discretization into a weak-form discretization, and vice versa (as discussed, for example, in [17, §8.1] and [22, §2.1]). Although not relying explicitly on the matrix form of the SBP property in (3.11), 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 (3.15). 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 (3.15).*

Proof Applying the SBP property in (3.11) to the second term of (3.33), we obtain

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

Recognizing that $\sum_{\zeta=1}^{N_f}\sum_{m=1}^{d}\hat{n}_m^{(\zeta)}(\underline{\underline{R}}^{(\zeta)})^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)$ cancels from the second and third terms on the left-hand side of (4.1) and noting the relation $(\underline{\underline{S}}^{(m)})^{\mathrm{T}}\underline{\underline{P}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)=(\underline{\underline{D}}^{(m)})^{\mathrm{T}}\underline{\underline{V}}^{\mathrm{T}}\underline{\underline{W}}\,\underline{f}^{(h,\kappa,m,\ell)}(t)$, we therefore recover (3.15). □

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 (3.21), 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 (3.21) is equivalent to the weak-form filtered DG scheme given by*

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

for all $\kappa\in\{1,\dots,N_e\}$, $\ell\in\{1,\dots,N_c\}$, *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 (3.34), 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 (3.11) to the second term of (3.34) in order to recover the weak formulation in (4.2). □

Although the weak-form filtered and unfiltered DG methods in (4.2) and (3.15) are mathematically equivalent (under the stated assumptions) to the strong-form FR and DG schemes in (3.21) and (3.33), respectively, the use of the weak form as opposed to the strong form results in the elimination of facet contributions of the form $\sum_{m=1}^{d} \hat{n}_m^{(\zeta)} \underline{\underline{R}}^{(\zeta)} \underline{\underline{P}}\, \underline{f}^{(h,\kappa,m,\ell)}(t)$. Although this reduction in the required number of floating-point operations for evaluating the semi-discrete residual would ostensibly result in the weak form being more efficient than the strong form (at least in the context of explicit temporal integration), the cost of additional local matrix-vector products could be relatively small if data such as the components of $\underline{\underline{P}}\, \underline{f}^{(h,\kappa,m,\ell)}(t)$ are already available in cache, and may not substantially affect the overall run time, particularly if the simulation is not compute bound. Moreover, as discussed in [26, §7] in the context of the DGSEM, certain choices of nodes may significantly reduce the cost of performing the flux extrapolation. For example, if a nodal basis as in Appendix A.2 is used for the strong formulation, with $\mathcal{S}^{(\zeta)} \subset \tilde{\mathcal{S}}$ for all $\zeta \in \{1, \ldots, N_f\}$, the vector $\underline{\underline{P}}\, \underline{f}^{(h,\kappa,m,\ell)}(t)$, which is needed to compute the volume terms, already contains the flux values at all facet nodes, and thus no additional matrix-vector products would be required.

Remark 4.1 The weak form of the VCJH family of FR methods proposed in [16, Eq. (3.7)] may be expressed in the present notation as

$$\underline{\underline{M}}^{(\kappa)} \frac{\mathrm{d}\underline{\tilde{u}}^{(h,\kappa,\ell)}(t)}{\mathrm{d}t} - \sum_{m=1}^{d} \left(\underline{\underline{S}}^{(m)}\right)^{\mathrm{T}} \underline{\underline{P}}\, \underline{f}^{(h,\kappa,m,\ell)}(t) + \sum_{\zeta=1}^{N_f} \underline{\underline{F}}^{\mathrm{T}} \left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{J}}^{(\kappa,\zeta)}\, \underline{f}^{(*,\kappa,\zeta,\ell)}(t)$$
$$+ \sum_{\zeta=1}^{N_f} \sum_{m=1}^{d} \hat{n}_m^{(\zeta)} \left(\underline{\underline{I}} - \underline{\underline{F}}^{\mathrm{T}}\right) \left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{R}}^{(\zeta)} \underline{\underline{P}}\, \underline{f}^{(h,\kappa,m,\ell)}(t) = \underline{0}, \tag{4.3}$$

which is algebraically equivalent to the strong formulation in (3.21) under the assumptions of Theorem 4.2, but requires a larger number of floating-point operations than the weak-form DG method in (3.15) when $\underline{\underline{K}} \neq \underline{\underline{0}}$. Theorem 4.2 demonstrates that by instead using the weak formulation in (4.2), we are in fact able to achieve equal computational cost to (3.15) for general choices of $\underline{\underline{K}}$ satisfying Assumption 3.4. It can also be concluded from the above analysis that any implementation of a DG method in strong or weak form can be easily modified to recover the full range of VCJH schemes through a simple modification to the mass matrix (or, equivalently, through the application of a filter to the entire semi-discrete residual).

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 indices $\kappa, \nu \in \{1, \ldots, N_e\}$ with $\kappa \neq \nu$ such that $\partial\Omega^{(\kappa)} \cap \partial\Omega^{(\nu)} \neq \emptyset$, there exist $\zeta, \eta \in \{1, \ldots, N_f\}$ such that either $\Gamma^{(\kappa,\zeta)} = \Gamma^{(\nu,\eta)}$, or the facets are connected via periodic boundary conditions. Moreover, the nodes on such facets align in physical space up to some permutation $\tau^{(\kappa,\zeta)}$, satisfying*

$$\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,i)}) = \boldsymbol{X}^{(\nu)}(\boldsymbol{\xi}^{(\eta,\tau^{(\kappa,\zeta)}(i))}), \quad \forall\, i \in \{1, \ldots, N_\zeta\}, \tag{4.4}$$

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 (see, for example, Horn and Johnson [44, Definition 4.5.4]) via the permutation matrix* $\underline{\underline{T}}^{(\kappa,\zeta)} \in \mathbb{R}^{N_\zeta \times N_\zeta}$ *with entries* $T_{ij}^{(\kappa,\zeta)} := \delta_{\tau^{(\kappa,\zeta)}(i),j}$*, as given by*

$$\underline{\underline{B}}^{(\eta)}\underline{\underline{J}}^{(\nu,\eta)} = \big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}. \tag{4.5}$$

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, and 3.3, the weak-form unfiltered and filtered DG methods in (3.15) and (4.2), respectively, the latter with a filter matrix associated with correction fields satisfying Assumption 3.4, are locally conservative with respect to a quadrature rule that is independent of the filter matrix, satisfying*

$$\frac{\mathrm{d}}{\mathrm{d}t}\underline{1}^{\mathrm{T}}\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}\underline{u}^{(h,\kappa,\ell)}(t) + \sum_{\zeta=1}^{N_f}\underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t) = 0 \tag{4.6}$$

for all $\kappa \in \{1,\dots,N_e\}$*,* $\ell \in \{1,\dots,N_c\}$*, and* $t \in (0,T)$*, where the nodal values of the numerical solution are given by* $\underline{u}^{(h,\kappa,\ell)}(t) := \underline{\underline{V}}\underline{\tilde{u}}^{(h,\kappa,\ell)}(t)$*. Moreover, such schemes are globally conservative with the addition of Assumption 4.1, satisfying*

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{\kappa=1}^{N_e}\underline{1}^{\mathrm{T}}\underline{\underline{W}}\underline{\underline{J}}^{(\kappa)}\underline{u}^{(h,\kappa,\ell)}(t) \; + \sum_{\Gamma^{(\kappa,\zeta)}\subset\partial\Omega}\underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t) = 0 \tag{4.7}$$

for all $\ell \in \{1,\dots,N_c\}$ *and* $t \in (0,T)$*, provided that* $\underline{F}^*$ *is conservative in the sense that* $\underline{F}^*(\underline{U}^-,\underline{U}^+,\boldsymbol{n}) = -\underline{F}^*(\underline{U}^+,\underline{U}^-,-\boldsymbol{n})$ *for all* $\underline{U}^-,\underline{U}^+ \in \mathcal{Y}$ *and* $\boldsymbol{n} \in \mathbb{S}^{d-1}$*.*

Proof Beginning with the weak-form filtered DG scheme in (4.2), which recovers the standard DG method in (3.15) when $\underline{\underline{K}} = \underline{\underline{0}}$, we pre-multiply by $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}$, which, by Assumption 2.2 and Lemma 3.2, contains the expansion coefficients for the constant function $\hat{\Omega} \mapsto 1$ in terms of the basis $\mathcal{B}$. Recalling (3.35), we therefore obtain

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

where the second term vanishes due to the fact that $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}(\underline{\underline{D}}^{(m)})^{\mathrm{T}} = \underline{0}^{\mathrm{T}}$ 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 (3.26), such a property also implies that $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}\underline{\underline{K}} = \underline{0}^{\mathrm{T}}$, which may be used along with the relations $\underline{\underline{M}}^{(\kappa)} = \underline{\underline{M}}\underline{\underline{P}}\underline{\underline{J}}^{(\kappa)}\underline{\underline{V}}$ and $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}(\underline{\underline{R}}^{(\zeta)})^{\mathrm{T}} = \underline{1}^{\mathrm{T}}$ to obtain the statement of local conservation in (4.6).

To establish global conservation, we sum (4.6) over all $\kappa \in \{1,\dots,N_e\}$ and split the net rate of production $\Phi_e^{(\kappa,\zeta)}(t) := \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t) + \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\eta)}\underline{\underline{J}}^{(\nu,\eta)}\underline{f}^{(*,\nu,\eta,\ell)}(t)$ arising from each shared interface equally between the oppositely oriented 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

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\sum_{\kappa=1}^{N_e} \underline{1}^{\mathrm{T}}\underline{\underline{W}}\,\underline{\underline{J}}^{(\kappa)}\underline{u}^{(h,\kappa,\ell)}(t) \;+ \sum_{\Gamma^{(\kappa,\zeta)}\not\subset\partial\Omega} \frac{1}{2}\Phi_e^{(\kappa,\zeta)}(t) \\ + \sum_{\Gamma^{(\kappa,\zeta)}\subset\partial\Omega} \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t) = 0,\end{aligned} \tag{4.9}$$

where the second sum must be taken over both sides of any interface (i.e. considering coincident but oppositely oriented facets to be distinct). 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

$$\begin{aligned}\Phi_e^{(\kappa,\zeta)}(t) &= \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{f}^{(*,\kappa,\zeta,\ell)}(t) + \underline{1}^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{f}^{(*,\nu,\eta,\ell)}(t) \\ &= \underline{1}^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\big(\underline{f}^{(*,\kappa,\zeta,\ell)}(t) + \underline{\underline{T}}^{(\kappa,\zeta)}\underline{f}^{(*,\nu,\eta,\ell)}(t)\big).\end{aligned} \tag{4.10}$$

Using (4.4) and the fact that the outward unit normal vectors for an interior or periodic interface satisfy

$$\boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,i)})) = -\boldsymbol{n}^{(\nu,\eta)}(\boldsymbol{X}^{(\nu)}(\boldsymbol{\xi}^{(\eta,\tau^{(\kappa,\zeta)}(i))})), \quad \forall\, i \in \{1,\dots,N_\zeta\}, \tag{4.11}$$

the terms in parentheses on the second line of (4.10) are given by

$$\underline{f}^{(*,\kappa,\zeta,\ell)}(t) = \begin{bmatrix} F_\ell^*\big(\underline{U}^{(h,\kappa)}(\boldsymbol{\xi}^{(\zeta,1)},t), \underline{U}^{(h,\nu)}(\boldsymbol{\xi}^{(\eta,\tau^{(\kappa,\zeta)}(1))},t), \\ \qquad\qquad \boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,1)}))\big) \\ \vdots \\ F_\ell^*\big(\underline{U}^{(h,\kappa)}(\boldsymbol{\xi}^{(\zeta,N_\zeta)},t), \underline{U}^{(h,\nu)}(\boldsymbol{\xi}^{(\eta,\tau^{(\kappa,\zeta)}(N_\zeta))},t), \\ \qquad\qquad \boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,N_\zeta)}))\big) \end{bmatrix} \tag{4.12}$$

and

$$\underline{\underline{T}}^{(\kappa,\zeta)}\underline{f}^{(*,\nu,\eta,\ell)}(t) = \begin{bmatrix} F_\ell^*\big(\underline{U}^{(h,\nu)}(\boldsymbol{\xi}^{(\eta,\tau^{(\kappa,\zeta)}(1))},t), \underline{U}^{(h,\kappa)}(\boldsymbol{\xi}^{(\zeta,1)},t), \\ \qquad\qquad -\boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,1)}))\big) \\ \vdots \\ F_\ell^*\big(\underline{U}^{(h,\nu)}(\boldsymbol{\xi}^{(\eta,\tau^{(\kappa,\zeta)}(N_\zeta))},t), \underline{U}^{(h,\kappa)}(\boldsymbol{\xi}^{(\zeta,N_\zeta)},t), \\ \qquad\qquad -\boldsymbol{n}^{(\kappa,\zeta)}(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,N_\zeta)}))\big) \end{bmatrix}, \tag{4.13}$$

which sum to the zero vector due to the conservation property of the numerical flux. Each term of the second sum in (4.9) then vanishes, and hence we obtain (4.7). □

The local conservation property in (4.6) may be expressed in terms of quadrature as

$$\begin{aligned}&\underbrace{\frac{\mathrm{d}}{\mathrm{d}t}\sum_{i=1}^{N}\underline{U}^h(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(i)}),t)J^{(\kappa)}(\boldsymbol{\xi}^{(i)})\,\omega^{(i)}}_{\approx \int_{\Omega^{(\kappa)}}\underline{U}^h(\boldsymbol{x},t)\,\mathrm{d}\boldsymbol{x}} \\ &+ \sum_{\zeta=1}^{N_f}\underbrace{\sum_{i=1}^{N_\zeta}\underline{F}^{(*,\kappa,\zeta)}(\boldsymbol{\xi}^{(i)},t)J^{(\kappa,\zeta)}(\boldsymbol{\xi}^{(\zeta,i)})\,\omega^{(\zeta,i)}}_{\approx \int_{\Gamma^{(\kappa,\zeta)}}\underline{F}^{(*,\kappa,\zeta)}((\boldsymbol{X}^{(\kappa)})^{-1}(\boldsymbol{x}),t)\,\mathrm{d}s} = \underline{0},\end{aligned} \tag{4.14}$$

where, regardless of the coefficients $\mathcal{C}$ used to form the matrix $\underline{\underline{K}}$ in (3.26), the reference volume and facet quadrature weights in (4.14) depend only on the discrete inner products in (3.1) and (3.2), and are given in the general case (i.e. allowing for dense weight matrices such as those constructed as in Appendix B.2) by

$$\omega^{(i)} := \sum_{j=1}^{N} W_{ij} \quad \text{and} \quad \omega^{(\zeta,i)} := \sum_{j=1}^{N_\zeta} B_{ij}^{(\zeta)}. \tag{4.15}$$

To establish discrete conservation for strong-form DG and FR methods, the divergence theorem on must be satisfied on the reference element under the discrete inner products, or equivalently, under the quadrature rules used in (4.14), for vector fields with components belonging to the approximation space. Such a condition is given by

$$\left\langle 1, \nabla_{\boldsymbol{\xi}} \cdot \boldsymbol{V} \right\rangle_W = \sum_{\zeta=1}^{N_f} \left\langle 1, \boldsymbol{V} \cdot \hat{\boldsymbol{n}}^{(\zeta)} \right\rangle_{B^{(\zeta)}}, \quad \forall\, \boldsymbol{V} \in \mathbb{P}_{\mathcal{N}}(\hat{\Omega})^d, \tag{4.16}$$

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 (3.21) and for the strong-form DG scheme in (3.33) with the additional requirement that the discrete divergence theorem in (4.16) holds, as is the case for methods satisfying Assumption 3.1.*

Proof We begin by pre-multiplying (3.21) by $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}(\underline{\underline{M}} + \underline{\underline{K}})$, which for $\underline{\underline{K}} = \underline{\underline{0}}$ is equivalent to pre-multiplying (3.33) by $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}$. Using the fact that $(\underline{\underline{M}} + \underline{\underline{K}})\underline{\underline{D}}^{(m)} = \underline{\underline{S}}^{(m)}$ holds for all $m \in \{1, \ldots, d\}$ under Assumption 3.4 and noting that the relation

$$\sum_{m=1}^{d} (\underline{\underline{P}}\underline{1})^{\mathrm{T}} \underline{\underline{S}}^{(m)} = \sum_{\zeta=1}^{N_f} \sum_{m=1}^{d} \hat{n}_m^{(\zeta)} (\underline{\underline{P}}\underline{1})^{\mathrm{T}} \left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{R}}^{(\zeta)} \tag{4.17}$$

may obtained by expanding (4.16) in terms of $\mathcal{B}$, or, if Assumption 3.1 is satisfied, by pre-multiplying (3.11) by $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}$, using $(\underline{\underline{P}}\underline{1})^{\mathrm{T}}(\underline{\underline{D}}^{(m)})^{\mathrm{T}}\underline{\underline{M}} = \underline{0}^{\mathrm{T}}$, and summing over $m \in \{1, \ldots, d\}$, we obtain the relation in (4.8), with the second term already eliminated. The remainder of the proof is then identical to that of Theorem 4.3. □

Remark 4.2 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 $\boldsymbol{G}^{(\zeta,j)}$, 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 (3.21), wherein the influence of the correction fields is contained within the lifting matrices.

[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).

4.3 Energy Stability

While the theory in §4.1 and §4.2 applies to DG and FR methods for linear or nonlinear systems of conservation laws 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 $\boldsymbol{X}^{(\kappa)}$ is affine for all $\kappa \in \{1, \dots, N_e\}$, and hence we may abuse notation to treat $\nabla_{\boldsymbol{\xi}} \boldsymbol{X}^{(\kappa)}$, $J^{(\kappa)}$, $J^{(\kappa,\zeta)}$, and $\boldsymbol{n}^{(\kappa,\zeta)}$ as constants.*

The solution energy and associated norm employed for the stability analysis in this section are introduced with the following lemma, where we note that the equation index is suppressed as we are considering a scalar conservation law (i.e. $N_c = 1$).

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

$$\mathcal{E}^h(t) := \frac{1}{2} \sum_{\kappa=1}^{N_e} (\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}} \underline{\underline{\tilde{M}}}^{(\kappa)} \underline{\tilde{u}}^{(h,\kappa)}(t) \tag{4.18}$$

may be expressed as a positive-definite quadratic form in terms of the global degrees of freedom $\underline{\tilde{u}}^h(t) := [(\underline{\tilde{u}}^{(h,1)}(t))^{\mathrm{T}}, \dots, (\underline{\tilde{u}}^{(h,N_e)}(t))^{\mathrm{T}}]^{\mathrm{T}}$, inducing a norm given by $\|U^h(\cdot,t)\|_{\Omega,h} := \sqrt{2\mathcal{E}^h(t)}$ for the approximation space $\mathbb{P}_{\mathcal{N}}(\mathcal{T}^h)$ defined in (2.7).

Proof Since $J^{(\kappa)}$ is constant under Assumption 4.2, the relation in (3.35) reduces to $\underline{\underline{\tilde{M}}}^{(\kappa)} = J^{(\kappa)}(\underline{\underline{M}} + \underline{\underline{K}})$, and hence $\underline{\underline{\tilde{M}}}^{(\kappa)}$ is SPD for all $\kappa \in \{1, \dots, N_e\}$ under Assumptions 2.1 and 3.4. As the right-hand side of (4.18) may be expressed equivalently as $\frac{1}{2}(\underline{\tilde{u}}^h(t))^{\mathrm{T}} \tilde{\mathbb{M}} \underline{\tilde{u}}^h(t)$, where $\tilde{\mathbb{M}} \in \mathbb{R}^{N_e \cdot N^* \times N_e \cdot N^*}$ is a block-diagonal matrix with SPD blocks $\underline{\underline{\tilde{M}}}^{(\kappa)}$ along its diagonal, such a quadratic form is also positive definite, and thus induces a norm for $\mathbb{R}^{N_e \cdot N^*}$. Since the expansion of $U^h(\cdot,t) \circ \boldsymbol{X}^{(\kappa)}$ terms of the basis $\mathcal{B}$ to obtain $\underline{\tilde{u}}^h(t)$ defines an isomorphism between the vector spaces $\mathbb{P}_{\mathcal{N}}(\mathcal{T}^h)$ and $\mathbb{R}^{N_e \cdot N^*}$, it therefore follows that (4.18) also induces a norm for $\mathbb{P}_{\mathcal{N}}(\mathcal{T}^h)$. □

Remark 4.3 When the condition in (3.24) is satisfied, we accordingly obtain

$$\|U^h(\cdot,t)\|_{\Omega,h}^2 = \sum_{\kappa=1}^{N_e} \int_{\hat{\Omega}} \left(\left(U^h(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}),t) \right)^2 + \sum_{\boldsymbol{\alpha} \in \mathcal{M}} \frac{c_{\boldsymbol{\alpha}}}{|\hat{\Omega}|} \left(D^{\boldsymbol{\alpha}} U(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}),t) \right)^2 \right) J^{(\kappa)}(\boldsymbol{\xi}) \, \mathrm{d}\boldsymbol{\xi}, \tag{4.19}$$

which, under Assumption 3.4, recovers the broken Sobolev-type norm introduced by Jameson [45] 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 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.

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 $\boldsymbol{n}(\boldsymbol{x}) \in \mathbb{S}^{d-1}$ may be partitioned into disjoint subsets $\partial\Omega^- := \{\boldsymbol{x} \in \partial\Omega : \boldsymbol{a}\cdot\boldsymbol{n}(\boldsymbol{x}) \leq 0\}$ and $\partial\Omega^+ := \partial\Omega \setminus \partial\Omega^-$, the former on which we prescribe the inflow boundary condition $U(\boldsymbol{x},t) = B(\boldsymbol{x},t)$. The numerical flux takes the general form*

$$F^*(U^-,U^+,\boldsymbol{n}) := \frac{1}{2}\big(\boldsymbol{F}(U^-)+\boldsymbol{F}(U^+)\big)\cdot\boldsymbol{n} - \frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}|}{2}(U^+-U^-) \tag{4.20}$$

for $U^-, U^+ \in \mathbb{R}$ and $\boldsymbol{n} \in \mathbb{S}^{d-1}$, 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 4.4 The numerical flux in (4.20) 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.* [46, Ch. 11]) we show that the energy defined in (4.18) 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 (3.15) and (3.21), respectively, are energy stable for the constant-coefficient linear advection equation, satisfying a bound given for all $t \in (0,T)$ by*

$$\frac{\mathrm{d}\mathcal{E}^h(t)}{\mathrm{d}t} \leq \sum_{\Gamma^{(\kappa,\zeta)} \subseteq \partial\Omega^-} \frac{|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2} (\underline{b}^{(\kappa,\zeta)}(t))^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{J}}^{(\kappa,\zeta)} \underline{b}^{(\kappa,\zeta)}(t), \tag{4.21}$$

where we define $\underline{b}^{(\kappa,\zeta)}(t) := [B(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,1)},t)),\ldots,B(\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}^{(\zeta,N_\zeta)},t))]^{\mathrm{T}}$ and note that the right-hand side vanishes for homogeneous or periodic boundary conditions.

Proof Energy stability will be proven for the strong-form FR method in (3.21), 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 (3.15) with $\underline{\underline{K}} = \underline{\underline{0}}$. Beginning with a single element and pre-multiplying (3.21) by $(\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}}(\underline{\underline{M}} + \underline{\underline{K}})$, we may use (3.35) and apply the chain rule in time to obtain

$$\begin{aligned}
\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}} \underline{\underline{\tilde{M}}}^{(\kappa)} \tilde{\underline{u}}^{(h,\kappa)}(t) = &\underbrace{-\sum_{m=1}^{d} (\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}} (\underline{\underline{M}} + \underline{\underline{K}}) \underline{\underline{D}}^{(m)} \underline{\underline{P}}\, \underline{f}^{(h,\kappa,m)}(t)}_{=:\,\Sigma_{\mathrm{I}}^{(\kappa)}(t)} \\
&+ \sum_{\zeta=1}^{N_f} \underbrace{\sum_{m=1}^{d} \hat{n}_m^{(\zeta)} (\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}} \big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{R}}^{(\zeta)} \underline{\underline{P}}\, \underline{f}^{(h,\kappa,m)}(t)}_{=:\,\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)} \\
&- \sum_{\zeta=1}^{N_f} \underbrace{(\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}} \big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}} \underline{\underline{B}}^{(\zeta)} \underline{\underline{J}}^{(\kappa,\zeta)}\, \underline{f}^{(*,\kappa,\zeta)}(t)}_{=:\,\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)},
\end{aligned} \tag{4.22}$$

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. Substituting $\boldsymbol{F}(U^{(h,\kappa)}(\boldsymbol{\xi},t)) = \boldsymbol{a}U^{(h,\kappa)}(\boldsymbol{\xi},t)$ into (2.10) as well as using $\underline{\underline{P}}\underline{\underline{V}} = \underline{\underline{I}}$ and $(\underline{\underline{M}}+\underline{\underline{K}})\underline{\underline{D}}^{(m)} = \underline{\underline{S}}^{(m)}$, which follow from Lemma 3.2 and Assumption 3.4, respectively, the first two terms on the right-hand side of (4.22) may be expressed as

$$\Sigma_{\mathrm{I}}^{(\kappa)}(t) = \sum_{m=1}^{d}\sum_{n=1}^{d} a_n J^{(\kappa)}\left[(\nabla_{\boldsymbol{\xi}}\boldsymbol{X}^{(\kappa)})^{-1}\right]_{mn}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\underline{\underline{S}}^{(m)}\underline{\tilde{u}}^{(h,\kappa)}(t) \tag{4.23}$$

and

$$\begin{aligned}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t) &= \sum_{m=1}^{d}\sum_{n=1}^{d} a_n J^{(\kappa)}\left[(\nabla_{\boldsymbol{\xi}}\boldsymbol{X}^{(\kappa)})^{-1}\right]_{mn}\hat{n}_m^{(\zeta)}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)\\ &= \boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t),\end{aligned} \tag{4.24}$$

where the second equality in (4.24) results from (2.5). Separating $\underline{\underline{S}}^{(m)}$ into its symmetric and skew-symmetric parts and using (3.11), we obtain

$$(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\underline{\underline{S}}^{(m)}\underline{\tilde{u}}^{(h,\kappa)}(t) = \frac{1}{2}\sum_{\zeta=1}^{N_f}\hat{n}_m^{(\zeta)}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t), \tag{4.25}$$

which recovers (2.35) in the one-dimensional case. Applying (4.25) to (4.23) and using the first equality in (4.24) then results in $\Sigma_{\mathrm{I}}^{(\kappa)}(t) = \sum_{\zeta=1}^{N_f}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)/2$. The local energy balance in (4.22) may therefore be expressed as

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\underline{\underline{\tilde{M}}}^{(\kappa)}\underline{\tilde{u}}^{(h,\kappa)}(t) = \sum_{\zeta=1}^{N_f}\left(\frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t) - \Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)\right), \tag{4.26}$$

where we note that the right-hand side contains only facet contributions. Summing (4.26) 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 facets $\Gamma^{(\kappa,\zeta)} \subset \partial\Omega^{(\kappa)}$ and $\Gamma^{(\nu,\eta)} \subset \partial\Omega^{(\nu)}$ which are coincident or connected via periodicity results in the global balance

$$\frac{\mathrm{d}\mathcal{E}^h(t)}{\mathrm{d}t} = \sum_{\Gamma^{(\kappa,\zeta)}\not\subset\partial\Omega}\frac{1}{2}\Sigma_{\mathrm{net}}^{(\kappa,\zeta)}(t) + \sum_{\Gamma^{(\kappa,\zeta)}\subset\partial\Omega}\left(\frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t) - \Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)\right). \tag{4.27}$$

As we will now demonstrate, the contributions to the right-hand side of (4.27) 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.

Similarly to the proof of Theorem 4.3, it follows from Assumption 4.1 that for any interior or periodic interface, the contributions arising from a numerical flux function taking the form of (4.20) may be expressed as

$$\begin{aligned}\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t) = \frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t) &+ \frac{\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}}{2}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t)\\ &- \frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\underline{\tilde{u}}^{(h,\nu)}(t)\\ &+ \frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\underline{\tilde{u}}^{(h,\kappa)}(t)\end{aligned} \tag{4.28}$$

and

$$\begin{aligned}\Sigma_{\mathrm{III}}^{(\nu,\eta)}(t) &= \frac{1}{2}\Sigma_{\mathrm{II}}^{(\nu,\eta)}(t)\\ &- \frac{\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}}{2}(\tilde{\underline{u}}^{(h,\nu)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t)\\ &- \frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\tilde{\underline{u}}^{(h,\nu)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t)\\ &+ \frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\tilde{\underline{u}}^{(h,\nu)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\tilde{\underline{u}}^{(h,\nu)}(t). \qquad (4.29)\end{aligned}$$

Noting that the first terms of (4.28) and (4.29) 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 (4.27) is then given by

$$\begin{aligned}\Sigma_{\mathrm{net}}^{(\kappa,\zeta)}(t) &= -\frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}\Big((\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t)\\ &\qquad - 2(\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\tilde{\underline{u}}^{(h,\nu)}(t)\\ &\qquad + (\tilde{\underline{u}}^{(h,\nu)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\eta)}\big)^{\mathrm{T}}\big(\underline{\underline{T}}^{(\kappa,\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\tilde{\underline{u}}^{(h,\nu)}(t)\Big)\\ &= -\frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\underline{d}^{(\kappa,\zeta)}(t))^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t), \qquad (4.30)\end{aligned}$$

where we define $\underline{d}^{(\kappa,\zeta)}(t) := \underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t) - \underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\tilde{\underline{u}}^{(h,\nu)}(t)$. 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 (4.20) 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 (4.27) resulting from the inflow and outflow portions of the boundary.

For any 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 (4.27) may be obtained by expressing $\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t)$ and $\Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t)$ as in (4.24) and (4.28), respectively, with $\underline{\underline{T}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\eta)}\tilde{\underline{u}}^{(h,\nu)}(t)$ replaced by $\underline{b}^{(\kappa,\zeta)}(t)$. We may then take $\lambda = 1$ as per Assumption 4.3 and note that $\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)} = -|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|$, resulting in

$$\begin{aligned}\frac{1}{2}\Sigma_{\mathrm{II}}^{(\kappa,\zeta)}(t) - \Sigma_{\mathrm{III}}^{(\kappa,\zeta)}(t) &= \frac{|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}\Big(2(\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t)\\ &\qquad - (\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\big(\underline{\underline{R}}^{(\zeta)}\big)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t)\Big)\\ &= \frac{|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}\Big((\underline{b}^{(\kappa,\zeta)}(t))^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t)\\ &\qquad - (\underline{d}^{(\kappa,\zeta)}(t))^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t)\Big)\\ &\le \frac{|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\underline{b}^{(\kappa,\zeta)}(t))^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t), \qquad (4.31)\end{aligned}$$

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.

Using the fact that $\mathit{\Gamma}^{(\kappa,\zeta)} \subset \partial\Omega^+$ implies $\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)} = |\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|$, it follows from taking $\lambda = 1$ in (4.28) that $\mathit{\Sigma}_{\mathrm{III}}^{(\kappa,\zeta)}(t) = \mathit{\Sigma}_{\mathrm{II}}^{(\kappa,\zeta)}(t)$. The resulting contribution to the second sum on the right-hand side of (4.27) is therefore simply $-\mathit{\Sigma}_{\mathrm{II}}^{(\kappa,\zeta)}(t)/2$, which is non-positive as a result of (4.24) and the positive-semidefiniteness of $\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}$, and hence we see that any facet lying on the outflow boundary does not contribute to growth in the solution energy. The final energy balance is then given by

$$\begin{aligned}\frac{\mathrm{d}\mathcal{E}^h(t)}{\mathrm{d}t} &= \sum_{\mathit{\Gamma}^{(\kappa,\zeta)}\subseteq\partial\Omega^-} \frac{|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\underline{b}^{(\kappa,\zeta)}(t))^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{b}^{(\kappa,\zeta)}(t)\\ &- \sum_{\mathit{\Gamma}^{(\kappa,\zeta)}\subseteq\partial\Omega^+} \frac{|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\tilde{\underline{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\left(\underline{\underline{R}}^{(\zeta)}\right)^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{\underline{R}}^{(\zeta)}\tilde{\underline{u}}^{(h,\kappa)}(t)\\ &- \sum_{\mathit{\Gamma}^{(\kappa,\zeta)}\subseteq\partial\Omega^-} \frac{|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{2}(\underline{d}^{(\kappa,\zeta)}(t))^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t)\\ &- \sum_{\mathit{\Gamma}^{(\kappa,\zeta)}\not\subseteq\partial\Omega} \frac{\lambda|\boldsymbol{a}\cdot\boldsymbol{n}^{(\kappa,\zeta)}|}{4}(\underline{d}^{(\kappa,\zeta)}(t))^{\mathrm{T}}\underline{\underline{B}}^{(\zeta)}\underline{\underline{J}}^{(\kappa,\zeta)}\underline{d}^{(\kappa,\zeta)}(t),\end{aligned} \tag{4.32}$$

where the first term corresponds 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 (4.21), thereby proving that the discretization is energy stable. □

Remark 4.5 For periodic boundary conditions and a central numerical flux, all terms on the right-hand side of (4.32) vanish, leading to discrete energy conservation.

Remark 4.6 The focus of this paper is on DG and FR methods in standard conservation (i.e. divergence) form. These discretizations are generally not provably stable for nonlinear or variable-coefficient PDEs, nor are they provably stable for meshes consisting of elements for which the transformation from the reference element is not affine, despite often being stable in practice and nevertheless used extensively in such contexts. While outside the scope of this paper, entropy-stable or split-form schemes allow for bounds analogous to (4.21) to be established for more complex problems. In the case of the Euler equations in (2.3), which are discretized using standard DG and FR methods in §5, entropy-stable nodal formulations based on the SBP property were introduced on tensor-product elements by Fisher *et al.* [47] and Carpenter *et al.* [48], which were extended to simplicial elements by Crean *et al.* [49] and Chen and Shu [50], and to modal formulations by Chan [25]. In all of the above formulations, it is required that the matrix $\underline{\underline{W}}$ be diagonal and that the facet terms (i.e. the SATs) be constructed using a lifting operator defined as in (3.26) with $\underline{\underline{K}} = \underline{\underline{0}}$. A subset of the discretizations discussed in this work (including quadrature-based unfiltered DG methods as well as nodal SBP schemes with diagonal mass matrices) therefore readily extend to entropy-stable formulations.

5 Numerical Experiments

5.1 Design Choices and Implementation

The methods described in §3 are implemented within a free and open-source Python code[8] offering substantial flexibility in the design choices afforded for DG and FR methods, which are detailed in this section.

5.1.1 Reference Element and Approximation Spaces

The numerical experiments in this work involve formulations in two space dimensions on a triangular reference element given by $\hat{\mathcal{T}}^2 := \left\{\boldsymbol{\xi} \in [-1,1]^2 : \xi_1 + \xi_2 \leq 0\right\}$, where $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$ is taken to be the total-degree polynomial space $\mathbb{P}_p(\hat{\mathcal{T}}^2)$, which results from (2.6) with the choice of $\mathcal{N} := \{\boldsymbol{\alpha} \in \mathbb{N}_0^2 : \alpha_1 + \alpha_2 \leq p\}$. Such a space is of dimension $N^* = (p+1)(p+2)/2$, and we report results for values of $p \in \{2, 3, 4\}$.

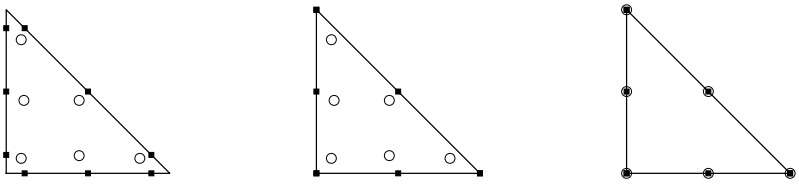
5.1.2 Discrete Inner Products and Nodal Sets

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 [51]. 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, 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 [52], 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 Figs. 1, 2, and 3, respectively, where the symbols ○ and ■ represent nodes in $\mathcal{S}$ and $\mathcal{S}^{(\zeta)}$, respectively.

5.1.3 Polynomial Bases

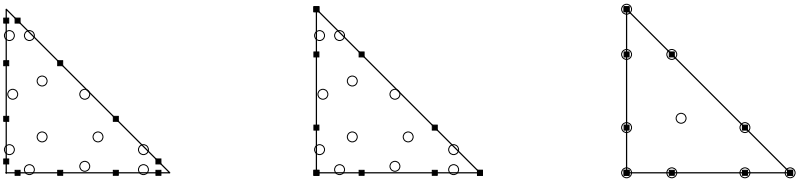
For the quadrature-based schemes, we take $\mathcal{B}$ to be an orthonormal modal basis defined as in (A.3), while for the collocation-based schemes, nodal bases given as in (A.6) are used with $\tilde{\mathcal{S}} = \mathcal{S}$, in this case corresponding to the "warp & blend" nodes in [52], which also satisfy $\mathcal{S}^{(\zeta)} \subset \tilde{\mathcal{S}}$ for all $\zeta \in \{1, \ldots, N_f\}$.

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



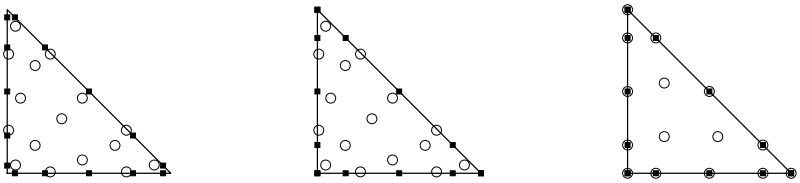
(a) Quadrature I (b) Quadrature II (c) Collocation

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 $p = 4$

5.1.4 Semi-Discrete Residual Formulation

We employ the strong formulation in (3.21) 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 (3.15), respectively.

5.1.5 Lifting and Filter Matrices

The matrix $\underline{\underline{K}}$ used to form the lifting and filter matrices is given as in (C.1), where 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.[9] 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

[9] 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.

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.

5.1.6 Numerical Flux Function

The numerical flux function for the linear advection equation is given as in (4.20), where we report results for $\lambda \in \{0, 1\}$, corresponding to central and upwind fluxes. In the case of the Euler equations, we use Roe's approximate Riemann solver [53].

5.1.7 Mesh and Coordinate Transformation

As the test problems considered in this work are posed on the 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 $N_e = 2M^2$ elements of equal size. Using an affine transformation $\boldsymbol{X}^{(\kappa)}_{\text{affine}} \in \mathbb{P}_1(\hat{\Omega})^2$ to map the nodes $\mathcal{S}_{\text{map}} := \{\boldsymbol{\xi}^{(1)}_{\text{map}}, \ldots, \boldsymbol{\xi}^{(N_{\text{map}})}_{\text{map}}\} \subset \hat{\Omega}$ associated with a Lagrange basis $\mathcal{B}_{\text{map}} := \{\ell^{(1)}_{\text{map}}, \ldots, \ell^{(N_{\text{map}})}_{\text{map}}\}$ of degree $p_{\text{map}} \in \mathbb{N}$ onto each straight-sided element of the original mesh, we then interpolate the warping

$$\boldsymbol{\Theta}(\boldsymbol{x}) := \begin{bmatrix} x_1 + \frac{1}{5} L \sin(\pi x_1 / L) \sin(\pi x_2 / L) \\ x_2 + \frac{1}{5} L \exp(1 - x_2/L) \sin(\pi x_1 / L) \sin(\pi x_2 / L) \end{bmatrix} \tag{5.1}$$

considered by Del Rey Fernández *et al.* [54], resulting in a mapping given by

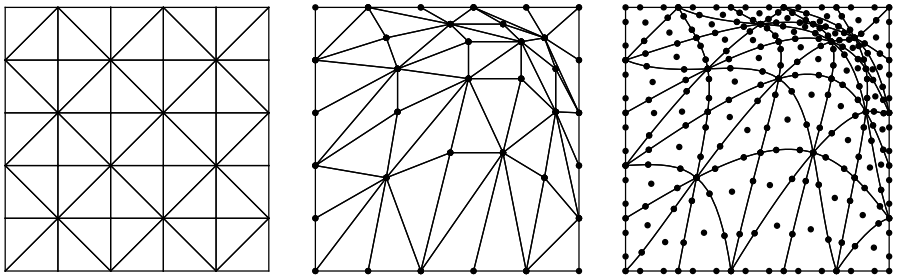
$$\boldsymbol{X}^{(\kappa)}(\boldsymbol{\xi}) := \sum_{i=1}^{N_{\text{map}}} \boldsymbol{\Theta}(\boldsymbol{X}^{(\kappa)}_{\text{affine}}(\boldsymbol{\xi}^{(i)}_{\text{map}})) \ell^{(i)}_{\text{map}}(\boldsymbol{\xi}). \tag{5.2}$$

In this work, the interpolation nodes in $\mathcal{S}_{\text{map}}$ are taken to be the "warp and blend" nodes in [52], which for $p = p_{\text{map}}$ coincide with those pictured for the Collocation schemes in Figs. 1, 2, and 3. Examples of meshes obtained through the above warping procedure with $M = 5$ are shown in Fig. 4, where the symbol $\bullet$ is used to represent the image of each node in $\mathcal{S}_{\text{map}}$ under the mapping given in (5.2).

5.1.8 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 $\Delta t \approx C_t h/a$,[10] 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 [55, §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 that of the spatial discretization.

[10] 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.



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

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

5.1.9 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.

- 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 $\ell \in \{1, \ldots, N_c\}$ by

$$\textbf{Equivalence Metric} := \sqrt{\int_\Omega \left(U^h_{\ell,\text{strong}}(\boldsymbol{x}, T) - U^h_{\ell,\text{weak}}(\boldsymbol{x}, T)\right)^2 \mathrm{d}\boldsymbol{x}}. \tag{5.3}$$

  This metric is expected to be zero for the Quadrature I and Collocation schemes (but not necessarily for the Quadrature II schemes) due to Theorems 4.1 and 4.2.
- For both the strong and weak forms, we compute the net change in the discrete integral of each conservative variable, which is given for $\ell \in \{1, \ldots, N_c\}$ by

$$\textbf{Conservation Metric} := \sum_{\kappa=1}^{N_e} \underline{1}^{\mathrm{T}} \underline{\underline{W}}\, \underline{\underline{J}}^{(\kappa)}\, \underline{\underline{V}}\, \underline{\tilde{u}}^{(h,\kappa,\ell)}(t)\Big|_0^T. \tag{5.4}$$

  This metric 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.
- In the case of the linear advection equation, we also evaluate the net change in the discrete solution energy defined in (4.21), which is given by

$$\textbf{Energy Metric} := \frac{1}{2} \sum_{\kappa=1}^{N_e} (\underline{\tilde{u}}^{(h,\kappa)}(t))^{\mathrm{T}}\, \underline{\underline{\tilde{M}}}^{(\kappa)}\, \underline{\tilde{u}}^{(h,\kappa)}(t)\Big|_0^T. \tag{5.5}$$

  This metric is expected from 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. A horizontal line denotes a computation for which the scheme was found to be unstable.

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

Scheme	c	λ	Equivalence Metric	Conservation Metric		Energy Metric	
				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

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.

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 $U^0(\boldsymbol{x}) := \sin(2\pi x_1/L)\sin(2\pi x_2/L)$ and the advection velocity is prescribed as $\boldsymbol{a} := a[\cos(\theta), \sin(\theta)]^{\mathrm{T}}$ with $a \in \mathbb{R}^+$ and $\theta \in [0, 2\pi)$. The mesh is constructed as described in §5.1, where we use $p_{\mathrm{map}} = 1$ 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 3: Linear advection equation, $p = 4$

Scheme	c	λ	Equivalence Metric	Conservation Metric		Energy Metric	
				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

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

Scheme	c	λ	Equivalence Metric	Conservation Metric		Energy Metric	
				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	0.000e+00*	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 [56, §5.1] following the general formulation described by Spiegel *et al.* [57]. 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 $\boldsymbol{x}^0 \in \Omega$ as

$$\boldsymbol{V}^0(\boldsymbol{x}) := \mathrm{Ma}_\infty \left(\begin{bmatrix} \cos(\theta) \\ \sin(\theta) \end{bmatrix} + \varepsilon \exp\left(1 - \|\boldsymbol{x} - \boldsymbol{x}^0\|^2\right) \begin{bmatrix} -(x_2 - x_2^0) \\ x_1 - x_1^0 \end{bmatrix} \right), \tag{5.6}$$

the initial temperature field is prescribed as

$$T^0(\boldsymbol{x}) := 1 - \frac{(\gamma-1)\varepsilon^2 \mathrm{Ma}_\infty^2}{2} \exp\bigl(1 - \|\boldsymbol{x} - \boldsymbol{x}^0\|^2\bigr), \tag{5.7}$$

and constant entropy is enforced throughout the spatial domain. Using (5.6) and (5.7), 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

$$\underline{U}^0(\boldsymbol{x}) := \begin{bmatrix} (T^0(\boldsymbol{x}))^{1/(\gamma-1)} \\ (T^0(\boldsymbol{x}))^{1/(\gamma-1)} \boldsymbol{V}^0(\boldsymbol{x}) \\ \frac{1}{\gamma-1}(T^0(\boldsymbol{x}))^{\gamma/(\gamma-1)} + \frac{1}{2}(T^0(\boldsymbol{x}))^{1/(\gamma-1)} \|\boldsymbol{V}^0(\boldsymbol{x})\|^2 \end{bmatrix}. \tag{5.8}$$

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. 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.

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 was 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, which were confirmed in numerical experiments. Based on this analysis, the following results were demonstrated theoretically for discretizations of linear as well as nonlinear conservation laws on unstructured grids employing curvilinear elements of any type and polynomial approximations of any order.

– If the components of the transformed flux in (2.10) are projected onto the polynomial approximation space on the reference element before evaluating the divergence and normal trace, the strong formulation in (3.33) of any DG method with the SBP property, regardless of the choice of basis or quadrature, is algebraically equivalent to the weak formulation in (3.15), which, unlike the strong formulation, does not require the extrapolation of the projected flux to the facet nodes.

Table 4: Euler equations, $p = 2$

Scheme	c	Equation	Equivalence Metric	Conservation Metric	
				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*	0.000e+00	–6.253e–13
	c_+	ρ	*2.529e–02*	0.000e+00	–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

- Any VCJH FR method satisfying the SBP property may be expressed equivalently as a filtered (or, in the case of $\underline{\underline{K}} = \underline{\underline{0}}$, unfiltered) DG method in strong form, as in (3.34), or in weak form, as in (4.2), where, unlike in [16], the filter matrix is applied to the entire semi-discrete residual and not only to the facet terms. These equivalences extend to FR methods employing modal bases and to those for which over-integration is employed for the volume and/or facet terms.
- The VCJH filter matrix defined in Lemma 3.4 does not alter the conservation properties of a given DG method. Consequently, any VCJH FR method satisfying the discrete divergence theorem is locally conservative with respect to the same quadrature rule as the corresponding DG scheme, with global conservation resulting from a suitable choice of numerical flux.

We have also demonstrated that all of the above DG and VCJH FR methods with the SBP property are energy stable with respect to discrete norms given as in Lemma 4.1 for linear advection problems on affine meshes, provided that a central or upwind-biased numerical flux is used for all interior or periodic interfaces, and that the boundary condition is weakly imposed on the inflow portion of the boundary. Due to the discretization-agnostic nature of the matrix-based analysis employed in this work, the equivalence, conservation, and energy stability results described above extend directly to any method employing derivative matrices satisfying the SBP property in (3.11), provided that the boundary and interface SATs are constructed appropriately.

Potential avenues for future research include the application of similar unifying principles to fully discrete formulations (i.e. considering the influence of the temporal discretization, which is neglected in the present analysis) and problems with diffusive

Table 5: Euler equations, $p = 3$

Scheme	c	Equation	Equivalence Metric	Conservation Metric	
				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

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 and weak formulations which we have shown in this paper to be mathematically equivalent, as well as to compare (i.e. through numerical experiments and eigensolution analysis) the DG and FR methods considered in this paper with those employing multidimensional SBP operators without an analytical basis. The theory could also be extended to enable proofs of convergence and *a priori* error estimates. More broadly, the present work may be viewed as a case study on the applicability of algebraic techniques for the unified analysis of numerical methods across multiple discretization paradigms. Although no single framework is likely to encompass *all* methods of interest for a given class of problem, we hope that the unifying perspective advocated in this paper will enable practitioners to consider a wider range of numerical methods within a common framework than would otherwise be possible, thereby facilitating the development or identification of schemes with advantageous properties for particular applications.

Acknowledgements

The authors are grateful to Gianmarco Mengaldo, Masayuki Yano, David Del Rey Fernández, and Alex Bercik for their feedback, as well as to Andreas Klöckner, Nico Schlömer, and the developers of NumPy [58], SciPy [59], and Matplotlib [60] for their free and open-source software contributions.

Table 6: Euler equations, $p = 4$

Scheme	c	Equation	Equivalence Metric	Conservation Metric	
				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

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 [61], 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

The authors declare that there are no conflicts of interest or competing interests which could have influenced this work.

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

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

Analytical expressions for such bases are well known for certain spaces and element types; for example, an orthogonal basis for the space $\mathbb{P}_p(\hat{\mathcal{T}}^2)$ was introduced by Proriol [62] (see also Koornwinder [63, §3.3] and Dubiner [64, §5]) as a triangular analogue 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

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

the normalized Proriol-Koornwinder-Dubiner (PKD) basis functions are given by

$$\phi_0^{(\sigma(\alpha))}(\boldsymbol{\chi}(\boldsymbol{\eta})) := \sqrt{\big(\alpha_1 + \tfrac{1}{2}\big)\big(\alpha_1 + \alpha_2 + 1\big)}\, P_{\alpha_1}^{(0,0)}(\eta_1)\Big(\frac{1-\eta_2}{2}\Big)^{\alpha_1} P_{\alpha_2}^{(2\alpha_1+1,0)}(\eta_2), \tag{A.3}$$

where $\sigma : \mathcal{N} \to \{1, \dots, (p+1)(p+2)/2\}$ defines an ordering[11] of the multi-index set $\mathcal{N} := \{\boldsymbol{\alpha} \in \mathbb{N}_0^2 : \alpha_1 + \alpha_2 \leq p\}$, and the mapping $\boldsymbol{\chi} : [-1,1]^2 \to \hat{\mathcal{T}}^2$ is given by

$$\boldsymbol{\chi}(\boldsymbol{\eta}) := \begin{bmatrix} \frac{1}{2}(1+\eta_1)(1-\eta_2) - 1 \\ \eta_2 \end{bmatrix}. \tag{A.4}$$

Similar bases are described for quadrilateral, hexahedral, tetrahedral, prismatic, and pyramidal elements in [39, Ch. 3].

[11] Such an ordering may be chosen arbitrarily, provided that σ is a bijection; for example, Hesthaven and Warburton [38, §6.1] consider $\sigma(\alpha) := \alpha_2 + (p+1)\alpha_1 + 1 - \alpha_1(\alpha_1 - 1)/2$.

A.2 Nodal (Lagrange) Basis

While it is possible to directly employ a modal basis to represent the numerical solution such that $\mathcal{B} = \mathcal{B}_0$, it can often be more computationally efficient to use a Lagrange basis $\mathcal{B} := \{\phi^{(1)}, \ldots, \phi^{(N^*)}\}$ associated with a nodal set $\tilde{\mathcal{S}} := \{\tilde{\boldsymbol{\xi}}^{(1)}, \ldots, \tilde{\boldsymbol{\xi}}^{(N^*)}\} \subset \hat{\Omega}$ such that $\phi^{(i)}(\tilde{\boldsymbol{\xi}}^{(j)}) = \delta_{ij}$. Although the corresponding basis functions are given explicitly by (2.17) for $d = 1$, analytical expressions do not necessarily exist for nodal bases on arbitrary non-tensorial nodal sets in the case of $d \geq 2$. We therefore take a more general approach by constructing a matrix $\underline{\underline{\tilde{V}}} \in \mathbb{R}^{N^* \times N^*}$ with entries

$$\tilde{V}_{ij} := \phi_0^{(j)}(\tilde{\boldsymbol{\xi}}^{(i)}) \tag{A.5}$$

for a modal basis $\mathcal{B}_0$ as described in Appendix A.1, which is invertible when $\tilde{\mathcal{S}}$ is unisolvent for the space $\mathbb{P}_{\mathcal{N}}(\hat{\Omega})$. The Lagrange polynomials constituting the basis $\mathcal{B}$ may then be obtained in terms of the modal basis as

$$\phi^{(i)}(\boldsymbol{\xi}) := \sum_{j=1}^{N^*} \left[\underline{\underline{\tilde{V}}}^{-\mathrm{T}}\right]_{ij} \phi_0^{(j)}(\boldsymbol{\xi}). \tag{A.6}$$

The entries of the corresponding derivative matrix satisfying (3.10) are therefore

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

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.1 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 [65] 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].

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 (e.g. 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 general 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

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

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

$$\left\langle U, V \right\rangle_W := \sum_{i=1}^{N} U(\boldsymbol{\xi}^{(i)}) V(\boldsymbol{\xi}^{(i)})\, \omega^{(i)} \tag{B.2}$$

and

$$\left\langle U, V \right\rangle_{B^{(\zeta)}} := \sum_{i=1}^{N_\zeta} U(\boldsymbol{\xi}^{(\zeta,i)}) V(\boldsymbol{\xi}^{(\zeta,i)})\, \omega^{(\zeta,i)}, \tag{B.3}$$

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 $p \in \mathbb{N}_0$, on any polytopal reference element, if the discrete inner products are computed as in (B.2) and (B.3) using 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 $\mathbb{P}_{2p}(\hat{I}^{(\zeta)})$. 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.

B.2 Collocation-Based Approximation

If $\mathcal{S}$ 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 $\{\ell^{(1)}, \dots, \ell^{(N)}\}$ for $\mathbb{P}_{\mathcal{K}}(\hat{\Omega})$ satisfying $\ell^{(i)}(\boldsymbol{\xi}^{(j)}) = \delta_{ij}$ as in Appendix A.2. The collocation projection $I_{\mathcal{K}} U \in \mathbb{P}_{\mathcal{K}}(\hat{\Omega})$ of a function $U : \hat{\Omega} \to \mathbb{R}$ is then given by

$$(I_{\mathcal{K}} U)(\boldsymbol{\xi}) := \sum_{i=1}^{N} U(\boldsymbol{\xi}^{(i)}) \ell^{(i)}(\boldsymbol{\xi}), \tag{B.4}$$

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 $\hat{I}^{(\zeta)} \subset \partial\hat{\Omega}$ which is of dimension $N_\zeta \geq N_\zeta^*$ and contains $\mathbb{P}_{\mathcal{N}}(\hat{I}^{(\zeta)})$, we may define a projection operator analogously to (B.4) as

$$(I^{(\zeta)} U)(\boldsymbol{\xi}) := \sum_{i=1}^{N_\zeta} U(\boldsymbol{\xi}^{(\zeta,i)}) \ell^{(\zeta,i)}(\boldsymbol{\xi}), \tag{B.5}$$

where the nodal basis functions $\{\ell^{(\zeta,1)},\dots,\ell^{(\zeta,N_\zeta)}\}$ satisfy $\ell^{(\zeta,i)}(\boldsymbol{\xi}^{(\zeta,j)})=\delta_{ij}$. Integrating the products of such projections, we obtain discrete inner products given by

$$\big\langle U,V\big\rangle_W := \int_{\hat{\Omega}} (I_{\mathcal{K}}U)(\boldsymbol{\xi})(I_{\mathcal{K}}V)(\boldsymbol{\xi})\,\mathrm{d}\boldsymbol{\xi} \tag{B.6}$$

and

$$\big\langle U,V\big\rangle_{B^{(\zeta)}} := \int_{\hat{\Gamma}^{(\zeta)}} (I^{(\zeta)}U)(\boldsymbol{\xi})(I^{(\zeta)}V)(\boldsymbol{\xi})\,\mathrm{d}\hat{s}, \tag{B.7}$$

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

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

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

Remark B.2 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 projection operators (see, for example, [44, 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 computations 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 quadrature accuracy requirements of Lemma B.1 may not be met.

Remark B.3 Recalling (4.15), 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 Lagrange basis functions as

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

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 $\mathbb{P}_{\mathcal{K}}(\hat{\Omega})$,[12] 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.

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 [66]).

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 $\mathcal{M}:=\{\boldsymbol{\alpha}\in\mathbb{N}_0^d : |\boldsymbol{\alpha}|=p\}$ in (3.26) ensures that $\underline{\underline{K}}\,\underline{\underline{D}}^{(m)}=\underline{0}$ is satisfied for all $m\in\{1,\dots,d\}$.*

[12] Taking $\mathbb{P}_{\mathcal{K}}(\hat{\Omega})=\mathbb{P}_q(\hat{\Omega})$, this would require the nodes in $\mathcal{S}$ to be associated with a volume quadrature rule of at least degree $2q$, as is the case for the collocated LG quadrature rule in Remark 2.6, for which the mass matrix is the same whether computed as in (2.21) or (2.22).

Proof Noting from (3.26) that $\underline{\underline{K}}\underline{\underline{D}}^{(m)} = \underline{\underline{0}}$ is satisfied for a given $m \in \{1, \dots, d\}$ if $\underline{\underline{D}}^{\boldsymbol{\alpha}}\underline{\underline{D}}^{(m)} = \underline{\underline{0}}$ for all $\boldsymbol{\alpha} \in \mathcal{M}$, it is sufficient to require the image of any $V \in \mathbb{P}_p(\hat{\Omega})$ under the corresponding differential operator $\partial^{|\boldsymbol{\alpha}|+1}/\partial\xi_1^{\alpha_1} \cdots \partial\xi_m^{\alpha_m+1} \cdots \partial\xi_d^{\alpha_d}$ 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 $|\boldsymbol{\alpha}|+1$ times, constraining $\mathcal{M}$ to satisfy $|\boldsymbol{\alpha}| = p$ leads to the desired result. □

Based on symmetry considerations discussed in [11, §5.3] and [12, §4.1], the set of multi-indices $\mathcal{M}$ defined in Lemma C.1 may be parametrized in terms of $d-1$ scalar indices. The matrix $\underline{\underline{K}}$ in (3.26) is then given in two and three space dimensions by

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

and

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

respectively, recovering (2.27) in the one-dimensional case, where $c \in \mathbb{R}$ determines the properties of the resulting schemes. With respect to the formalism of Theorem 3.1, the parametrizations defining $\underline{\underline{K}}$ in (C.1) and (C.2) correspond to $c_{\boldsymbol{\alpha}} = c\binom{p}{\alpha_2}$ and $c_{\boldsymbol{\alpha}} = c\binom{p}{p-\alpha_1}\binom{p-\alpha_1}{\alpha_3}$, 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 (3.24) is not satisfied) such that $\underline{\underline{M}} + \underline{\underline{K}}$ is SPD for $c_- < c < \infty$.

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," in *18th AIAA Computational Fluid Dynamics Conference*, American Institute of Aeronautics and Astronautics, 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, pp. A1935–A1958, July 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, Aug. 2010.
27. 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.
28. M. Vinokur, "Conservation equations of gasdynamics in curvilinear coordinate systems," *Journal of Computational Physics*, vol. 14, pp. 105–125, Feb. 1974.
29. M. E. Gurtin, E. Fried, and L. Anand, *The Mechanics and Thermodynamics of Continua*. Cambridge University Press, 2010.
30. A. Cohen and G. Migliorati, "Multivariate approximation in downward closed polynomial spaces," in *Contemporary Computational Mathematics – A Celebration of the 80th Birthday of Ian Sloan* (J. Dick, F. Y. Kuo, and H. Woźniakowski, eds.), pp. 233–282, Springer, 2018.
31. 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.
32. E. F. Toro, *Riemann Solvers and Numerical Methods for Fluid Dynamics: A Practical Introduction*. Springer, 3rd ed., 2009.
33. M. H. Carpenter and D. Gottlieb, "Spectral methods on arbitrary grids," *Journal of Computational Physics*, vol. 129, pp. 74–86, Nov. 1996.
34. P. E. Vincent, A. M. Farrington, F. D. Witherden, and A. Jameson, "An extended range of stable-symmetric-conservative flux reconstruction correction functions," *Computer Methods in Applied Mechanics and Engineering*, vol. 296, pp. 248–272, Nov. 2015.
35. C. Canuto, M. Y. Hussaini, A. Quarteroni, and T. A. Zang, *Spectral Methods: Fundamentals in Single Domains*. Springer, 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, 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. 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.
42. 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.
43. 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.
44. R. A. Horn and C. R. Johnson, *Matrix Analysis*. Cambridge University Press, 2nd ed., 2013.
45. 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.
46. B. Gustafsson, H.-O. Kreiss, and J. Oliger, *Time-Dependent Problems and Difference Methods*. John Wiley & Sons, Inc., 2nd ed., 2013.
47. T. C. Fisher and M. H. Carpenter, "High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains," *Journal of Computational Physics*, vol. 252, pp. 518–557, Nov. 2013.
48. M. H. Carpenter, T. C. Fisher, E. J. Nielsen, and S. H. Frankel, "Entropy stable spectral collocation schemes for the Navier-Stokes equations: Discontinuous interfaces," *SIAM Journal on Scientific Computing*, vol. 36, pp. B835–B867, Oct. 2014.

49. J. Crean, J. E. Hicken, D. C. Del Rey Fernández, D. W. Zingg, and M. H. Carpenter, "Entropy-stable summation-by-parts discretization of the Euler equations on general curved elements," *Journal of Computational Physics*, vol. 356, pp. 410–438, Mar. 2018.
50. T. Chen and C.-W. Shu, "Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws," *Journal of Computational Physics*, vol. 345, pp. 427–461, Sept. 2017.
51. 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.
52. T. Warburton, "An explicit construction of interpolation nodes on the simplex," *Journal of Engineering Mathematics*, vol. 56, pp. 247–262, Sept. 2006.
53. P. L. Roe, "Approximate Riemann solvers, parameter vectors, and difference schemes," *Journal of Computational Physics*, vol. 43, pp. 357–372, Sept. 1981.
54. 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.
55. B. Cockburn and C.-W. Shu, "Runge-Kutta discontinuous Galerkin methods for convection-dominated problems," *Journal of Scientific Computing*, vol. 16, Sept. 2001.
56. 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, 1998.
57. 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.
58. C. R. Harris, K. J. Millman, S. J. van der Walt, R. Gommers, P. Virtanen, D. Cournapeau, E. Wieser, J. Taylor, S. Berg, N. J. Smith, R. Kern, M. Picus, S. Hoyer, M. H. van Kerkwijk, M. Brett, A. Haldane, J. F. del Río, M. Wiebe, P. Peterson, P. Gérard-Marchant, K. Sheppard, T. Reddy, W. Weckesser, H. Abbasi, C. Gohlke, and T. E. Oliphant, "Array programming with NumPy," *Nature*, vol. 585, pp. 357–362, Sept. 2020.
59. P. Virtanen, R. Gommers, T. E. Oliphant, M. Haberland, T. Reddy, D. Cournapeau, E. Burovski, P. Peterson, W. Weckesser, J. Bright, S. J. van der Walt, M. Brett, J. Wilson, K. J. Millman, N. Mayorov, A. R. J. Nelson, E. Jones, R. Kern, E. Larson, C. J. Carey, İ. Polat, Y. Feng, E. W. Moore, J. VanderPlas, D. Laxalde, J. Perktold, R. Cimrman, I. Henriksen, E. A. Quintero, C. R. Harris, A. M. Archibald, A. H. Ribeiro, F. Pedregosa, P. van Mulbregt, and SciPy 1.0 Contributors, "SciPy 1.0: Fundamental algorithms for scientific computing in Python," *Nature Methods*, vol. 17, pp. 261–272, Mar. 2020.
60. J. D. Hunter, "Matplotlib: A 2D graphics environment," *Computing in Science & Engineering*, vol. 9, pp. 90–95, June 2007.
61. 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)*, Association for Computing Machinery, July 2019.
62. 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.
63. T. Koornwinder, "Two-variable analogues of the classical orthogonal polynomials," in *Theory and Application of Special Functions* (R. Askey, ed.), pp. 435–495, Academic Press, 1975.
64. M. Dubiner, "Spectral methods on triangles and other domains," *Journal of Scientific Computing*, vol. 6, pp. 345–390, Dec. 1991.
65. A. L. Marchildon and D. W. Zingg, "Unisolvency for polynomial interpolation in simplices with symmetrical nodal distributions," *Journal of Scientific Computing*, vol. 92, Aug. 2022.
66. 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.