CSIAM Trans. Appl. Math.
doi: 10.4208/csiam-am.2020-0003

Vol. **1**, No. 1, pp. 1-52
March 2020

REVIEW ARTICLE

Review of Entropy Stable Discontinuous Galerkin Methods for Systems of Conservation Laws on Unstructured Simplex Meshes

Tianheng Chen and Chi-Wang Shu*

Division of Applied Mathematics, Brown University, Providence, RI 02912, USA.

Received 31 May 2019; Accepted 29 January 2020

Abstract. In this paper, we will build a roadmap for the growing literature of high order quadrature-based entropy stable discontinuous Galerkin (DG) methods, trying to elucidate the motivations and emphasize the contributions. Compared to the classic DG method which is only provably stable for the square entropy, these DG methods can be tailored to satisfy an arbitrary given entropy inequality, and do not require exact integration. The methodology is within the summation-by-parts (SBP) paradigm, such that the discrete operators collocated at the quadrature points should satisfy the SBP property. The construction is relatively easy for quadrature rules with collocated surface nodes. We use the flux differencing technique to ensure entropy balance within elements, and the simultaneous approximation terms (SATs) to produce entropy dissipation on element interfaces. The further extension to general quadrature rules is achieved through careful modifications of SATs.

AMS subject classifications: 65M12, 65M60

Key words: System of conservation laws, entropy stability, discontinuous Galerkin method, summation-by-parts.

1 Introduction

Systems of conservation laws describe the phenomena that the production of a conserved quantity in any domain is balanced by a flux through the boundary [22]. Entropy inequalities, which help to single out the "physically relevant" solution, are crucial to the well-posedness of conservation laws. Therefore, when designing numerical methods, we hope that entropy inequalities are satisfied at certain discrete level. Such property is

*Corresponding author. *Email addresses:* tianheng_chen@alumni.brown.edu (T. Chen), chi-wang_shu@brown.edu (C.-W. Shu)

called entropy stability. Entropy stability analysis is well-developed for the first order (finite volume) method. The key concepts are Tadmor's entropy conservative fluxes and entropy stable fluxes [81,82]. For high order entropy stable finite volume methods, a major result is the *TeCNO* scheme, proposed by Fjordholm, Mishra and Tadmor [31] as a version of ENO schemes [45]. The authors used high order linear combinations of entropy conservative fluxes in [61], along with the sign property of ENO reconstruction [32].

Discontinuous Galerkin (DG) methods [13–15, 17], due to their local conservation, great parallel efficiency and flexibility for dealing with unstructured meshes, constitute another popular category of high order numerical methods for solving conservation laws. It is well known that the classic DG method satisfies a discrete entropy inequality with respect to the square entropy (i.e., L^2 stability), for scalar conservation laws [58] and symmetric systems [54]. However, the stability result is only valid for the square entropy function. There is no provable stability for problems such that the square function does not define an entropy function. Moreover, we implicitly assume that all integrals in the DG formulation are evaluated exactly. In practice, numerical quadrature rules are usually applied, and the method we actually code up might not be stable. One possible remedy to accomplish entropy stability for an arbitrary given entropy function is to approximate the entropy variables of that entropy function directly (see [2,53,56,87]). This approach is computationally expensive, as we need to solve nonlinear systems at each time step, even for explicit time discretization. Besides, the stability proof still relies on the assumption of exact integration.

Over the past decade, there have been rapid developments on entropy stable quadrature-based DG methods. These DG methods are often characterized in the matrix-vector nodal formulation collocated at the quadrature points [49, 60]. Because of the approximation error induced by quadrature, we no longer have the integration by parts property and the chain rule. The methodology was first developed for the Legendre-Gauss-Lobatto quadrature rule in one space dimension. The corresponding discrete operators (i.e., matrices) were shown to satisfy the summation-by-parts (SBP) property [24, 26, 80], which is the discrete analogue of integration by parts. The distinctive feature of the Gauss-Lobatto quadrature rule is that it contains the two boundary points. Then we make sure that the boundary matrices are diagonal, and neighboring cells can be coupled in a natural way through penalty type terms, usually called simultaneous approximation terms (SATs) in the literature. In order to deal with the loss of chain rule, ad hoc split form methods have been provided for the Burgers equation [36], shallow water equations [39] and Euler equations [38] (for kinetic energy stability). In [4, 5], Carpenter et al. revealed the generic logic behind the splitting procedure by demonstrating the *flux differencing* technique. Flux differencing is essentially high order difference operations on Tadmor's entropy conservative fluxes, and is applicable to any system with any given entropy function.

The one-dimensional Gauss-Lobatto nodal methodology can be easily generalized to multi-dimensional Cartesian meshes through tensor product. In [12], Chen and Shu proposed the entropy stable DG method on unstructured simplex meshes, by intro-

ducing special Gauss-Lobatto type quadrature rules with collocated surface quadrature points, and establishing discrete operators with the multi-dimensional summation-by-parts property [27, 52]. The further extension to general quadrature rules is highly non-trivial. Although we are still able to produce SBP operators in the general setting, the boundary matrices are dense, which makes the treatment of element coupling terms (i.e., SAT) more involved. In [65, 69, 71], the authors again used the idea of splitting to construct ad hoc entropy stable SATs for some special problems. Then for general systems, two different entropy stable DG methods were given by Chan in [7, 8] and Crean et al. in [19, 20]. In both approaches, effectively an augmented set of SBP operators with diagonal boundary matrices was invented, so that the flux differencing term and the SAT were built on those newly defined operators. They will be named hybridized SBP operators approach and global SBP operators approach in this paper. We also remark that in [1], Abgrall recommended a "brute force" type approach that eliminates entropy error and enforces chain rule directly. It arrived at the same goal as flux differencing, without necessitating entropy conservative fluxes.

There have been numerous contributions improving the framework in many other aspects. To name a few, entropy stable DG methods were devised for convection-diffusion equations [4,5,12,37], MHD equations [3,62], shallow water equations [39,69,83–85] (with the well-balancedness property), gradient flow problems [77, 78], two-phase flow problems [73] and stochastic problems [64] (via the generalized polynomial chaos approach in [86]). The staggered-grid variant was discussed in [25, 67], and by using this idea, *modal* DG formulations (evolving polynomials instead of nodal values) were recovered in [7, 8]. Continuous SBP operators and the corresponding entropy stable continuous finite element method was developed in [50]. The assumption of conforming simplex meshes can also be greatly relaxed. People have studied the generalization to curvilinear meshes [3, 4, 9, 20, 37], non-conforming meshes [35], moving meshes [74] and space-time meshes [34].

The objective of this paper is to systematically review and reinterpret the existing quadrature-based entropy stable DG methods, primarily in the context of unstructured simplex meshes. The rest of this paper is organized as follows. In Section 2, we briefly present some necessary background materials, including continuous entropy analysis for systems of conservation laws, and discrete entropy analysis for the first order method and the classic DG method. In Section 3, we introduce quadrature rules on simplex elements and the corresponding summation-by-parts operators, deriving the matrix-vector nodal representation of the classic DG method. In Section 4, we derive the entropy stable DG method for quadrature rules with collocated surface nodes (and diagonal boundary matrices), which is followed by the extension to general quadrature rules (and dense boundary matrices) in Section 5. We check the accuracy of these DG methods by carrying out a simple numerical test for the two-dimensional Burgers equation. Several additional topics are explained in Section 6. Concluding remarks and future research directions are given in Section 7. Finally in the appendices, we demonstrate the equivalence of flux differencing and splitting in certain cases, as well as the proofs of some theorems.

2 Background: systems of conservation laws

2.1 Continuous entropy analysis

The general form of systems of conservation laws is

$$\frac{\partial \mathbf{u}}{\partial t}+\sum_{m=1}^{d}\frac{\partial \mathbf{f}_m(\mathbf{u})}{\partial x_m}=0,\quad (t,\mathbf{x})\in[0,\infty)\times\mathbb{R}^d, \tag{2.1}$$

where $\mathbf{u}\in\mathbb{R}^p$ are vector-valued conservative variables, and $\mathbf{f}_m\in\mathbb{R}^p$ are flux functions. A scalar convex function $U(\mathbf{u})$ is called an entropy function for (2.1) if there exist entropy fluxes $\{F_m(\mathbf{u})\}_{m=1}^d$, such that the following integrability condition holds

$$U'(\mathbf{u})\mathbf{f}_m'(\mathbf{u})=F_m'(\mathbf{u}),\quad 1\le m\le d. \tag{2.2}$$

$U'(\mathbf{u})$ and $F_m'(\mathbf{u})$ are understood as row vectors, and $\mathbf{f}_m'(\mathbf{u})$ is the $p\times p$ Jacobian matrix. Given a strictly convex entropy function U, let $\mathbf{v}=U'(\mathbf{u})^T$ be the entropy variables. Then $\mathbf{v}'(\mathbf{u})=U''(\mathbf{u})$ is symmetric positive-definite, and the mapping $\mathbf{u}\mapsto\mathbf{v}$ is invertible. We also define the potential fluxes

$$\psi_m(\mathbf{v})=\mathbf{v}^T\mathbf{f}_m(\mathbf{u}(\mathbf{v}))-F_m(\mathbf{u}(\mathbf{v})),\quad 1\le m\le d. \tag{2.3}$$

One can verify that (see e.g. [42])

$$\psi_m'(\mathbf{v})=\mathbf{f}_m(\mathbf{u}(\mathbf{v}))^T. \tag{2.4}$$

In addition, for a unit vector $\mathbf{n}\in\mathbb{R}^d$, we set

$$\mathbf{f}_{\mathbf{n}}(\mathbf{u})=\sum_{m=1}^{d}n_m\mathbf{f}_m(\mathbf{u}),\quad F_{\mathbf{n}}(\mathbf{u})=\sum_{m=1}^{d}n_mF_m(\mathbf{u}),\quad \psi_{\mathbf{n}}(\mathbf{v})=\sum_{m=1}^{d}n_m\psi_m(\mathbf{v}).$$

If $\mathbf{u}$ is a smooth solution of (2.1), by (2.2), $U(\mathbf{u})$ satisfies a secondary conservation law

$$\frac{\partial U(\mathbf{u})}{\partial t}+\sum_{m=1}^{d}\frac{\partial F_m(\mathbf{u})}{\partial x_m}=0. \tag{2.5}$$

At discontinuities, we require the entropy to dissipate, a weak solution $\mathbf{u}$ of (2.1) is called an entropy solution if for all entropy functions, we have the following entropy inequality

$$\frac{\partial U(\mathbf{u})}{\partial t}+\sum_{m=1}^{d}\frac{\partial F_m(\mathbf{u})}{\partial x_m}\le 0\quad\text{(in the weak sense)}. \tag{2.6}$$

Formally integrating (2.6) in space, and assuming that $\mathbf{u}$ is compactly supported, we obtain the bound

$$\frac{d}{dt}\int_\Omega U(\mathbf{u})d\mathbf{x}\le 0. \tag{2.7}$$

In other words, the total amount of entropy is non-increasing with respect to time.

For scalar conservation laws ($p=1$), any convex function U defines an entropy function, with entropy fluxes $F_m(u)=\int^u U'(s)f_m'(s)ds$. Due to the abundance of entropy functions, there exists a unique "physically correct" entropy solution. For general systems, existence of entropy function is no longer guaranteed, and both existence and uniqueness of entropy solutions are much more challenging. Fortunately, in almost all systems we encounter in practice (e.g. shallow water equations, Euler equations, MHD equations), we are able to find entropy functions with physical meaning. We refer interested readers to [22, 41] for more details on the entropy analysis of systems of conservation laws.

2.2 First order method

Now we start to look into the numerical aspects. We will mostly conduct semi-discrete analysis, i.e., temporal discretization is not taken into account. For spatial discretization, suppose that $\Omega\in\mathbb{R}^d$ is some polygonal computational domain equipped with periodic boundary condition. Let $\mathcal{T}_h=\{T_\kappa\}_{\kappa=1}^K$ be some conforming partition of Ω, and h be the characteristic length of $\mathcal{T}_h$. We assume that each element T_κ is a simplex, so that ∂T_κ consists of $(d-1)$-dimensional simplex faces. The collection of faces is denoted by

$$\Gamma_h=\{\gamma:\gamma=\partial T_\kappa\cap\partial T_\nu,\ 1\le\kappa,\nu\le K,\ \kappa\neq\nu\}. \tag{2.8}$$

Given $T_\kappa\in\mathcal{T}_h$ and $\gamma\in\Gamma_h$ such that $\gamma\in\partial T_\kappa$, we use the notation $\mathbf{n}^{\gamma\kappa}$ to represent the unit outward normal vector at γ. We will often omit the superscripts γ and κ if they can be inferred from the context.

The first order (finite volume) method evolves the piecewise constant function $\mathbf{u}_h(t,\mathbf{x})=\sum_{\kappa=1}^K\mathbf{u}^\kappa(t)\mathbb{1}_{T_\kappa}(\mathbf{x})$, and is written in the conservative form

$$\frac{d\mathbf{u}^\kappa}{dt}+\frac{1}{|T_\kappa|}\Big(\sum_{\gamma\in\partial T_\kappa}|\gamma|\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}^\kappa,\mathbf{u}^\nu)\Big)=0,\quad 1\le\kappa\le K, \tag{2.9}$$

where Ω_ν is the adjacent element on the opposite side of γ, and $\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)$ is some directional interface numerical flux function, satisfying

1. Consistency: $\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u},\mathbf{u})=\mathbf{f}_{\mathbf{n}}(\mathbf{u})$.
2. Single-valuedness: $\widehat{\mathbf{f}}_{-\mathbf{n}}(\mathbf{u}_R,\mathbf{u}_L)=-\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)$.

It actually approximates the following integral form of (2.1):

$$\frac{d}{dt}\Big(\int_{T_\kappa}\mathbf{u}d\mathbf{x}\Big)+\int_{\partial T_\kappa}\mathbf{f}_{\mathbf{n}}(\mathbf{u})dS=0. \tag{2.10}$$

Entropy stability of (2.9) is thoroughly studied by Tadmor in [81, 82]. For an entropy function U, the rate of change of the total entropy is

$$\begin{aligned}\frac{d}{dt}\int_\Omega U(\mathbf{u}_h(t,\mathbf{x}))d\mathbf{x} &= \frac{d}{dt}\Big(\sum_{\kappa=1}^{K}|T_\kappa|U^\kappa\Big) = -\sum_{\kappa=1}^{K}(\mathbf{v}^\kappa)^T\Big(\sum_{\gamma\in\partial T_\kappa}|\gamma|\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}^\kappa,\mathbf{u}^\nu)\Big)\\ &= -\sum_{\gamma\in\Gamma_h}|\gamma|\Big((\mathbf{v}^\kappa)^T\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}^\kappa,\mathbf{u}^\nu)+(\mathbf{v}^\nu)^T\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\nu}}(\mathbf{u}^\nu,\mathbf{u}^\kappa)\Big)\quad(\gamma=\partial T_\kappa\cap\partial T_\nu)\\ &= \sum_{\gamma\in\Gamma_h}|\gamma|(\mathbf{v}^\nu-\mathbf{v}^\kappa)^T\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}^\kappa,\mathbf{u}^\nu)\quad(\text{since }\mathbf{n}^{\gamma\nu}=-\mathbf{n}^{\gamma\kappa}),\end{aligned}\tag{2.11}$$

where we use the short hand notation $U^\kappa=U(\mathbf{u}^\kappa)$ and $\mathbf{v}^\kappa=\mathbf{v}(\mathbf{u}^\kappa)$. This motivates us to define the concepts of entropy conservative flux and entropy stable flux.

Definition 2.1. *For $1\le m\le d$, a numerical flux function $\mathbf{f}_{m,S}(\mathbf{u}_L,\mathbf{u}_R)$ is called entropy conservative with respect to some entropy U if it is consistent, symmetric and satisfies the following equality:*

$$(\mathbf{v}_R-\mathbf{v}_L)^T\mathbf{f}_{m,S}(\mathbf{u}_L,\mathbf{u}_R)=\psi_{m,R}-\psi_{m,L},\tag{2.12}$$

where we again set $\mathbf{v}_{L,R}=\mathbf{v}(\mathbf{u}_{L,R})$ and $\psi_{m,(L,R)}=\psi_m(\mathbf{v}_{L,R})$. $\{\psi_m\}_{m=1}^d$ are the potential fluxes given in (2.3). *Given entropy conservative fluxes in all space dimensions, we also define the directional entropy conservative flux*

$$\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_L,\mathbf{u}_R)=\sum_{m=1}^{d}n_m\mathbf{f}_{m,S}(\mathbf{u}_L,\mathbf{u}_R).$$

Definition 2.2. *A directional numerical flux function $\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)$ is called entropy stable with respect to some entropy U if it is consistent, single-valued and satisfies the following inequality:*

$$(\mathbf{v}_R-\mathbf{v}_L)^T\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)\le\psi_{\mathbf{n},R}-\psi_{\mathbf{n},L}.\tag{2.13}$$

Recall (2.11). If $\widehat{\mathbf{f}}_{\mathbf{n}}$ is entropy stable,

$$\begin{aligned}\frac{d}{dt}\int_\Omega U(\mathbf{u}_h(t,\mathbf{x}))d\mathbf{x} &\le \sum_{\gamma\in\Gamma_h}|\gamma|(\psi^\nu_{\mathbf{n}^{\gamma\kappa}}-\psi^\kappa_{\mathbf{n}^{\gamma\kappa}}) = -\sum_{\gamma\in\Gamma_h}|\gamma|(\psi^\kappa_{\mathbf{n}^{\gamma\kappa}}+\psi^\nu_{\mathbf{n}^{\gamma\nu}})\\ &= \sum_{\kappa=1}^{K}\Big(\sum_{\gamma\in\partial T_\kappa}|\gamma|\psi^\kappa_{\mathbf{n}^{\gamma\kappa}}\Big)=0\quad\Big(\text{since }\sum_{\gamma\in\partial T_\kappa}|\gamma|\mathbf{n}^{\gamma\kappa}=0\Big).\end{aligned}$$

We accordingly say that (2.9) is entropy stable with respect to U. Similarly, if $\widehat{\mathbf{f}}_{\mathbf{n}}$ is entropy conservative, the total entropy does not change and the scheme is said to be entropy conservative.

In the scalar case, the entropy conservative fluxes are uniquely determined

$$f_{m,S}(u_L,u_R)=\begin{cases}\frac{\psi_{m,R}-\psi_{m,L}}{v_R-v_L}, & u_L\ne u_R,\\ f_m(u_L), & u_L=u_R.\end{cases}\tag{2.14}$$

For general systems, (2.12) is underdetermined and $\mathbf{f}_{m,S}(\mathbf{u}_L,\mathbf{u}_R)$ is not unique. Various computationally affordable entropy conservative fluxes have been provided for shallow water equations [30], Euler equations [10,51,57,70] and MHD equations [11,23]. As for the construction of entropy stable fluxes, we can prove that the monotone fluxes [18,46] for scalar conservation laws and Godunov-type fluxes [47] for general systems are stable with respect to *all* entropy functions. Another common practice in the literature [5,7,10, 37,57] is to simply add some entropy dissipation to the entropy conservative flux:

$$\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)=\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_L,\mathbf{u}_R)+\widehat{\mathbf{d}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R),$$

where the entropy dissipation function $\widehat{\mathbf{d}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)$ satisfies the following conditions:

1. Consistency: $\widehat{\mathbf{d}}_{\mathbf{n}}(\mathbf{u},\mathbf{u})=0$.
2. Single-valuedness: $\widehat{\mathbf{d}}_{-\mathbf{n}}(\mathbf{u}_R,\mathbf{u}_L)=-\widehat{\mathbf{d}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)$.
3. Entropy dissipation: $(\mathbf{v}_R-\mathbf{v}_L)^T\widehat{\mathbf{d}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)\leq 0$.

For example, the local Lax-Friedrichs dissipation function is a popular choice of $\widehat{\mathbf{d}}_{\mathbf{n}}$:

$$\widehat{\mathbf{d}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)=-\lambda_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)(\mathbf{u}_R-\mathbf{u}_L), \tag{2.15}$$

where $\lambda_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)\geq 0$ is some estimate of the largest absolute eigenvalue in $\mathbf{f}'_{\mathbf{n}}(\mathbf{u})$. Notice that in this approach, $\widehat{\mathbf{f}}_{\mathbf{n}}$ is only stable with respect to a *single given* entropy function, as entropy conservative fluxes are specific to entropy functions.

2.3 Classic DG method

Unlike in the first order method, generally speaking, entropy stability with respect to all entropy functions can not be accomplished in high order methods. Osher [66] suggested the concept of E-schemes to characterize numerical methods supporting all entropy inequalities, and proved that E-schemes are at most first order accurate. Therefore we have to make a compromise, i.e., to expect entropy stability with respect to a *single given* entropy function.

In the classic DG method, we keep the locality of the first order formulation, and evolve high order piecewise discontinuous polynomials. Given polynomial degree $k\geq 0$, we define the DG space

$$\mathbf{V}_h^k=\{\mathbf{w}_h:\mathbf{w}_h^\kappa\in[\mathcal{P}^k(T_\kappa)]^p,\,1\leq\kappa\leq K\}, \tag{2.16}$$

where $\mathbf{w}_h^\kappa$ is the restriction of $\mathbf{w}_h$ on T_κ. We seek $\mathbf{u}_h(t,\cdot)\in\mathbf{V}_h^k$ such that for each $\mathbf{w}_h\in\mathbf{V}_h^k$ and $1\leq\kappa\leq K$,

$$\int_{T_\kappa}\left(\frac{\partial\mathbf{u}_h^\kappa}{\partial t}\right)^T\mathbf{w}_h^\kappa d\mathbf{x}-\sum_{m=1}^d\int_{T_\kappa}\mathbf{f}_m(\mathbf{u}_h^\kappa)^T\frac{d\mathbf{w}_h^\kappa}{dx_m}d\mathbf{x}=-\sum_{\gamma\in\partial T_\kappa}\int_\gamma\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^\kappa,\mathbf{u}_h^\nu)^T\mathbf{w}_h^\kappa dS. \tag{2.17}$$

Again, $\widehat{\mathbf{f}}_{\mathbf{n}}$ is some consistent and single-valued directional numerical flux function, and $\gamma=\partial T_\kappa\cap\partial T_\nu$. Eq. (2.17) is usually called the *weak form* of the DG method as it approximates the weak problem

$$\int_{\mathbb{R}^d}\frac{\partial\mathbf{u}(t,\mathbf{x})^T}{\partial t}\mathbf{w}(\mathbf{x})d\mathbf{x}-\sum_{m=1}^{d}\int_{\mathbb{R}^d}\mathbf{f}_m(\mathbf{u}(t,\mathbf{x}))^T\frac{d\mathbf{w}(\mathbf{x})}{dx_m}d\mathbf{x}=0, \tag{2.18}$$

for all smooth and compactly supported $\mathbf{w}$. The *strong form* of the DG method is obtained after a simple integration by parts

$$\int_{T_\kappa}\Big(\frac{\partial\mathbf{u}_h^\kappa}{\partial t}+\sum_{m=1}^{d}\frac{\partial\mathbf{f}_m(\mathbf{u}_h^\kappa)}{\partial x_m}\Big)^T\mathbf{w}_h^\kappa d\mathbf{x}=\sum_{\gamma\in\partial T_\kappa}\int_\gamma(\mathbf{f}_{\mathbf{n}}(\mathbf{u}_h^\kappa)-\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^\kappa,\mathbf{u}_h^\nu))^T\mathbf{w}_h^\kappa dS, \tag{2.19}$$

which corresponds to Eq. (2.1) itself. The classic DG method is L^2 stable if we have a square entropy function, e.g. in scalar problems [58] and symmetric systems [54].

Theorem 2.1. *If $U=\frac{1}{2}\mathbf{u}^T\mathbf{u}$ is an entropy function of* (2.1)*, and $\widehat{\mathbf{f}_{\mathbf{n}}}$ is entropy stable with respect to U, then the DG method* (2.17) *and* (2.19) *is L^2 stable in the sense that*

$$\frac{d}{dt}\int_\Omega U(\mathbf{u}_h)d\mathbf{x}=\frac{d}{dt}\Big(\frac{1}{2}\|\mathbf{u}_h\|^2_{L^2(\Omega)}\Big)\le 0. \tag{2.20}$$

Proof. Since $U=\frac{1}{2}\mathbf{u}^T\mathbf{u}$, $\mathbf{v}=\mathbf{u}$, and $\psi_m'(\mathbf{u})=\mathbf{f}_m(\mathbf{u})$. We set $\mathbf{w}_h=\mathbf{u}_h$ in (2.17) and get

$$\begin{aligned}\frac{d}{dt}\Big(\frac{1}{2}\|\mathbf{u}_h\|^2_{L^2}\Big)&=\sum_{\kappa=1}^{K}\int_{T_\kappa}\Big(\frac{\partial\mathbf{u}_h^\kappa}{\partial t}\Big)^T\mathbf{u}_h^\kappa d\mathbf{x}\\&=\sum_{\kappa=1}^{K}\Big(\sum_{m=1}^{d}\mathbf{f}_m(\mathbf{u}_h^\kappa)^T\frac{\partial\mathbf{u}_h^\kappa}{\partial x_m}d\mathbf{x}-\sum_{\gamma\in\partial T_\kappa}\int_\gamma\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^\kappa,\mathbf{u}_h^\nu)^T\mathbf{u}_h^\kappa dS\Big)\\&=\sum_{\kappa=1}^{K}\sum_{\gamma\in\partial T_\kappa}\int_\gamma\Big(\psi_{\mathbf{n}}(\mathbf{u}_h^\kappa)-\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^\kappa,\mathbf{u}_h^\nu)^T\mathbf{u}_h^\kappa dS\Big)\\&=\sum_{\gamma\in\Gamma_h}\int_\gamma\Big(\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}_h^\kappa,\mathbf{u}_h^\nu)^T(\mathbf{u}_h^\nu-\mathbf{u}_h^\kappa)-(\psi_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}_h^\nu)-\psi_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}_h^\kappa))\Big)\le 0.\end{aligned}$$

The last inequality results from the entropy stability of $\widehat{\mathbf{f}}_{\mathbf{n}}$. □

The stability result is limited to the square entropy function. For a general entropy U, the mapping $\mathbf{u}\mapsto\mathbf{v}$ is nonlinear, and $\mathbf{v}(\mathbf{u}_h)$ does not live in the piecewise polynomial space $\mathbf{V}_h^k$. We can not use $\mathbf{v}(\mathbf{u}_h)$ as the test function. An alternative approach, originally found by Hughes, Franca and Mallet [56] in the context of continuous finite element

method, is to approximate $\mathbf{v}$ directly. We evolve $\mathbf{v}_h(t,\cdot)\in\mathbf{V}_h^k$ such that for each $\mathbf{w}_h\in\mathbf{V}_h^k$ and $1\le\kappa\le K$,

$$\begin{aligned}&\int_{T_\kappa}\Big(\frac{\partial\mathbf{u}(\mathbf{v}_h^\kappa)}{\partial t}\Big)^T\mathbf{w}_h^\kappa d\mathbf{x}-\sum_{m=1}^d\int_{T_\kappa}\mathbf{f}_m(\mathbf{u}(\mathbf{v}_h^\kappa))^T\frac{d\mathbf{w}_h^\kappa}{dx_m}d\mathbf{x}\\ =&-\sum_{\gamma\in\partial T_\kappa}\int_\gamma\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}(\mathbf{v}_h^\kappa),\mathbf{u}(\mathbf{v}_h^\nu))^T\mathbf{w}_h^\kappa dS.\end{aligned}\tag{2.21}$$

Then we can prove that (2.21) is entropy stable with respect to U, by simply taking $\mathbf{w}_h=\mathbf{v}_h$ and repeating the proof of Theorem 2.1. This approach has the drawback of solving a nonlinear system at each time step (nonlinear solvers could be avoided by writing $\partial_t\mathbf{u}(\mathbf{v}_h^\kappa)=\mathbf{u}'(\mathbf{v}_h^\kappa)\partial_t\mathbf{v}_h^\kappa$ and inverting the Jacobian matrix $\mathbf{u}'(\mathbf{v}_h^\kappa)$, but at the cost of violating primary conservation [21]). We will not concentrate on it in this paper.

The entropy stable DG methods we are going to discuss do not incur nonlinear solvers. They are based on quadrature points and nodal formulation, so that we can perform nonlinear mapping freely. Actually, quadrature rules are necessary for the implementation of the DG method. If the flux functions $\{\mathbf{f}_m\}_{m=1}^d$ are not polynomials (e.g. in Euler equations), it is costly or even impossible to evaluate the second integral in (2.17) exactly. There are two technical challenges related to the nodal form. We need discrete versions of integration by parts and the chain rule $\mathbf{f}_m(\mathbf{u}(\mathbf{v}))^T\partial_{x_m}\mathbf{v}=\partial_{x_m}\psi_m(\mathbf{v})$, which are crucial to the proof of entropy stability. In subsequent sections, we will bring into the summation-by-parts operators, and the flux differencing technique, to handle these difficulties.

3 Summation-by-parts operators

Summation-by-parts (SBP) operators mimic integration by parts at the discrete level. One can check [24,26,80] for the review of SBP operators in one space dimension, and [27,52] for the generalization to higher space dimensions. SBP operators are widely used in designing high order and provably stable numerical methods [44], in particular, entropy stable DG type methods [5,36,39]. In a nutshell, by specifying suitable volume and surface quadrature rules, we construct SBP operators based on quadrature points, and by applying those quadrature rules, we are able to rewrite the DG method under the SBP framework.

3.1 Quadrature rules

The degree k SBP operators are built on at least degree $(2k-1)$ volume quadrature rules and at least degree $2k$ surface quadrature rules. Let $\{p_l(\mathbf{x})\}_{l=1}^{\mathcal{N}_{P,k}}$ be a set of basis functions of $\mathcal{P}^k(\mathbb{R}^d)$, such that

$$\mathcal{N}_{P,k}=\dim\mathcal{P}^k(\mathbb{R}^d)=\binom{k+d}{d}.$$

For each $1\leq\kappa\leq K$, suppose that there is an at least degree $(2k-1)$ quadrature rule on T_κ, associated with $\mathcal{N}_{Q,k}\geq\mathcal{N}_{P,k}$ nodes $\{\mathbf{x}_j^\kappa\}_{j=1}^{\mathcal{N}_{Q,k}}$, and positive weights $\{\omega_j^\kappa\}_{j=1}^{\mathcal{N}_{Q,k}}$. For each $\gamma\in\Gamma_h$, we also choose some at least degree $2k$ (surface) quadrature rule on γ, associated with $\mathcal{N}_{B,k}$ nodes $\{\mathbf{x}_s^\gamma\}_{s=1}^{\mathcal{N}_{B,k}}$, and positive weights $\{\tau_s^\gamma\}_{s=1}^{\mathcal{N}_{B,k}}$. We introduce the vector notation of nodal functions. For some scalar function u on Ω,

$$\overrightarrow{u^\kappa}=\begin{bmatrix}u(\mathbf{x}_1^\kappa) & \cdots & u(\mathbf{x}_{\mathcal{N}_{Q,k}}^\kappa)\end{bmatrix}^T,\quad \overrightarrow{u^\gamma}=\begin{bmatrix}u(\mathbf{x}_1^\gamma) & \cdots & u(\mathbf{x}_{\mathcal{N}_{B,k}}^\gamma)\end{bmatrix}^T.$$

Then the continuous and discrete inner products on T_κ and γ are defined as

$$(u,v)_{T_\kappa}=\int_{T_\kappa}uvd\mathbf{x},\quad (u,v)_{T_\kappa,\omega}=\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j u(\mathbf{x}_j^\kappa)v(\mathbf{x}_j^\kappa)=\left(\overrightarrow{u^\kappa}\right)^T M^\kappa\overrightarrow{v^\kappa},\tag{3.1}$$

$$\langle u,v\rangle_\gamma=\int_\gamma uvdS,\quad \langle u,v\rangle_{\gamma,\tau}=\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s u(\mathbf{x}_s^\gamma)v(\mathbf{x}_s^\gamma)=\left(\overrightarrow{u^\gamma}\right)^T B^\gamma\overrightarrow{v^\gamma},\tag{3.2}$$

where the volume mass matrix (M^κ) and the surface mass matrix (B^γ) are diagonal matrices of quadrature weights:

$$M^\kappa=\mathrm{diag}\{\omega_1^\kappa,\cdots,\omega_{\mathcal{N}_{Q,k}}^\kappa\},\quad B^\gamma=\mathrm{diag}\{\tau_1^\gamma,\cdots,\tau_{\mathcal{N}_{B,k}}^\gamma\}.\tag{3.3}$$

We also define the Vandermonde matrices, whose columns are nodal values of $\{p_l(\mathbf{x})\}_{l=1}^{\mathcal{N}_{P,k}}$:

$$V^\kappa=\begin{bmatrix}\overrightarrow{p_1^\kappa} & \cdots & \overrightarrow{p_{\mathcal{N}_{P,k}}^\kappa}\end{bmatrix},\quad V^\gamma=\begin{bmatrix}\overrightarrow{p_1^\gamma} & \cdots & \overrightarrow{p_{\mathcal{N}_{P,k}}^\gamma}\end{bmatrix}.\tag{3.4}$$

Derivatives of polynomials in $\mathcal{P}^k(\mathbb{R}^d)$ still belong to $\mathcal{P}^k(\mathbf{R}^d)$. We set $\mathcal{N}_{P,k}\times\mathcal{N}_{P,k}$ polynomial (modal) differentiation matrices $\widehat{D}_m$ for $1\leq m\leq d$, such that

$$\frac{\partial p_l}{\partial x_m}(\mathbf{x})=\sum_{r=1}^{\mathcal{N}_{P,k}}\widehat{D}_{m,rl}p_r(\mathbf{x}).$$

Then $V^\kappa\widehat{D}_m$ is the Vandermonde matrix of $\{\partial_{x_m}p_l(\mathbf{x})\}_{l=1}^{\mathcal{N}_{P,k}}$ on T_κ. According to integration by parts and the algebraic accuracy of $(\cdot,\cdot)_{T_\kappa,\omega}$ and $\langle\cdot,\cdot\rangle_{\gamma,\tau}$,

$$\begin{aligned}
&\left((V^\kappa)^TM^\kappa(V^\kappa\widehat{D}_m)\right)_{lr}+\left((V^\kappa\widehat{D}_m)^TM^\kappa V^\kappa\right)_{lr}\\
=&\left(\overrightarrow{p_l^\kappa}\right)^TM^\kappa\overrightarrow{(\partial_{x_m}p_r)^\kappa}+\left(\overrightarrow{(\partial_{x_m}p_l)^\kappa}\right)^TM^\kappa\overrightarrow{p_r^\kappa}\\
=&(p_l,\partial_{x_m}p_r)_{T_\kappa,\omega}+(\partial_{x_m}p_l,p_r)_{T_\kappa,\omega}=(p_l,\partial_{x_m}p_r)_{T_\kappa}+(\partial_{x_m}p_l,p_r)_{T_\kappa}\\
=&\sum_{\gamma\in\partial T^\kappa}n_m^{\gamma\kappa}\langle p_l,p_r\rangle_\gamma=\sum_{\gamma\in\partial T^\kappa}n_m^{\gamma\kappa}\langle p_l,p_r\rangle_{\gamma,\tau}\\
=&\sum_{\gamma\in\partial T^\kappa}n_m^{\gamma\kappa}\left(\overrightarrow{p_l^\gamma}\right)^TB^\gamma\overrightarrow{p_r^\gamma}=\sum_{\gamma\in\partial T^\kappa}n_m^{\gamma\kappa}\left((V^\gamma)^TB^\gamma V^\gamma\right)_{lr}.
\end{aligned}$$

In other words,

$$\widehat{M}^\kappa \widehat{D}_m + \widehat{D}_m^T \widehat{M}^\kappa = \sum_{\gamma\in\partial T^\kappa} n_m^{\gamma\kappa} \widehat{B}^\gamma, \tag{3.5}$$

which is the *modal* summation-by-parts property. The modal mass matrices are

$$\widehat{M}^\kappa = (V^\kappa)^T M^\kappa V^\kappa, \quad \widehat{M}^\kappa_{lr} = (p_l, p_r)_{T_\kappa,\omega}, \quad \widehat{B}^\gamma = (V^\gamma)^T B^\gamma V^\gamma, \quad \widehat{B}^\gamma_{lr} = \langle p_l, p_r\rangle_{\gamma,\tau}. \tag{3.6}$$

3.2 The SBP property

Inspired by (3.5), we come up with the definition of *nodal* SBP operators.

Definition 3.1. *For $1\le\kappa\le K$, D_m^κ (of size $\mathcal{N}_{Q,k}\times\mathcal{N}_{Q,k}$) and $\{R^{\gamma\kappa}\}_{\gamma\in\partial T^\kappa}$ (of size $\mathcal{N}_{B,k}\times\mathcal{N}_{Q,k}$) are called the degree k difference matrix approximating ∂_{x_m}, and extrapolation matrices mapping data from T_κ to γ, if the following conditions hold:*

1. *Exactness: both D_m^κ and $R^{\gamma\kappa}$ should be exact for polynomials of degree $\le k$; that is,*

$$D_m^\kappa V^\kappa = V^\kappa \widehat{D}_m, \quad R^{\gamma\kappa} V^\kappa = V^\gamma. \tag{3.7}$$

2. *Summation-by-parts: setting $S_m^\kappa = M^\kappa D_m^\kappa$ and $E^{\gamma\kappa} = (R^{\gamma\kappa})^T B^\gamma R^{\gamma\kappa}$, we have*

$$S_m^\kappa + (S_m^\kappa)^T = M^\kappa D_m^\kappa + (D_m^\kappa)^T M^\kappa = \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa} E^{\gamma\kappa} = \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa} (R^{\gamma\kappa})^T B^\gamma R^{\gamma\kappa}. \tag{3.8}$$

It is called the diagonal-norm SBP property as M^κ is a diagonal matrix.

A simple choice of extrapolation matrices is $R^{\gamma\kappa} = V^\gamma P^\kappa$, where P^κ is the L^2 projection matrix [7] with respect to the discrete inner product $(\cdot,\cdot)_{T_\kappa,\omega}$:

$$P^\kappa = (\widehat{M}^\kappa)^{-1} (V^\kappa)^T M^\kappa. \tag{3.9}$$

Then $P^\kappa V^\kappa = (\widehat{M}^\kappa)^{-1}\widehat{M}^\kappa = I_{\mathcal{N}_{Q,k}}$ and $R^{\gamma\kappa}V^\kappa = V^\gamma$. The existence of SBP difference matrices is ensured by the following theorem [12, 27, 52].

Theorem 3.1. *Assume that we have an extrapolation matrices $R^{\gamma\kappa}$ with the exactness property. Then the difference matrices, given by the formula*

$$D_m^\kappa = \frac{1}{2}(M^\kappa)^{-1} \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa} (R^{\gamma\kappa} + V^\gamma P^\kappa)^T B^\gamma (R^{\gamma\kappa} - V^\gamma P^\kappa) + V^\kappa \widehat{D}_m P^\kappa, \tag{3.10}$$

also satisfy the exactness property and SBP property.

Proof. Since $P^\kappa V^\kappa = I_{N_{Q,k}}$, $(R^{\gamma\kappa} - V^\gamma P^\kappa)V^\kappa = R^{\gamma\kappa}V^\kappa - V^\gamma = 0$, and

$$D_m^\kappa V^\kappa = V^\kappa \widehat{D}_m P^\kappa V^\kappa = V^\kappa \widehat{D}_m.$$

As for the SBP property,

$$S_m^\kappa = M^\kappa D_m^\kappa = \frac{1}{2}\sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa}(R^{\gamma\kappa}+V^\gamma P^\kappa)^T B^\gamma (R^{\gamma\kappa}-V^\gamma P^\kappa) + M^\kappa V^\kappa \widehat{D}_m P^\kappa$$

$$= \frac{1}{2}\sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa}(E^{\kappa\gamma}+(V^\gamma P^\kappa)^T B^\gamma R^{\gamma\kappa} - (R^{\gamma\kappa})^T B^\gamma V^\gamma P^\kappa - (P^\kappa)^T \widehat{B}^\gamma P^\kappa) + (P^\kappa)^T \widehat{M}^\kappa \widehat{D}_m P^\kappa.$$

Summing S_m^κ and its transpose yields

$$\begin{aligned} S_m^\kappa + (S_m^\kappa)^T &= \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa}(E^{\kappa\gamma} - (P^\kappa)^T \widehat{B}^\gamma P^\kappa) + (P^\kappa)^T \widehat{M}^\kappa \widehat{D}_m P^\kappa + (P^\kappa)^T \widehat{D}_m^T \widehat{M}^\kappa P^\kappa \\ &= \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa} E^{\kappa\gamma} \quad \text{(by the modal SBP property)}, \end{aligned}$$

which completes the proof. □

Remark 3.1. By the exactness property and SBP property,

$$S_m^\kappa \overrightarrow{1^\kappa} = D_m^\kappa \overrightarrow{1^\kappa} = \overrightarrow{0^\kappa}, \quad R^{\gamma\kappa}\overrightarrow{1^\kappa} = \overrightarrow{1^\gamma}, \tag{3.11a}$$

$$(S_m^\kappa)^T \overrightarrow{1^\kappa} = \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa} E^{\gamma\kappa}\overrightarrow{1^\kappa} = \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa}(R^{\gamma\kappa})^T B^\gamma \overrightarrow{1^\gamma}, \tag{3.11b}$$

where $\overrightarrow{0^\kappa}$ ($\overrightarrow{0^\gamma}$) and $\overrightarrow{1^\kappa}$ ($\overrightarrow{1^\gamma}$) represent the vector of 0s and 1s evaluated on T_κ (γ).

Remark 3.2. We would like to highlight some special cases of $R^{\kappa\gamma}$ and D_m^κ:

1. If $\mathcal{N}_{P,k} = \mathcal{N}_{Q,k}$ (e.g. the one-dimensional Legendre-Gauss quadrature rule and Legendre-Gauss-Lobatto quadrature rule with $(k+1)$ points), the Vandermonde matrix V^κ is invertible. Then both $R^{\kappa\gamma}$ and D_m^κ are uniquely determined:
$$R^{\gamma\kappa} = V^\gamma (V^\kappa)^{-1}, \quad D_m^\kappa = V^\kappa \widehat{D}_m (V^\kappa)^{-1}.$$
2. If $R^{\gamma\kappa} = V^\gamma P^\kappa$, the first term of (3.10) vanishes, and $D_m^\kappa = V^\kappa \widehat{D}_m P^\kappa$.
3. If the volume quadrature rule has collocated surface quadrature points (e.g. the one-dimensional Legendre-Gauss-Lobatto quadrature rule), given $\gamma \in \partial T_\kappa$, without loss of generality we assume that $\mathbf{x}_s^\gamma = \mathbf{x}_s^\kappa$ for each $1 \le s \le \mathcal{N}_{B,k}$. Then we can choose $R^{\gamma\kappa} = \begin{bmatrix} I_{\mathcal{N}_{B,k}} & 0 \end{bmatrix}$, a simple restriction, such that
$$E^{\gamma\kappa} = \begin{bmatrix} B^\gamma & 0 \\ 0 & 0 \end{bmatrix}$$
is a diagonal matrix, and
$$D_m^\kappa = \frac{1}{2}(M^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa}(I_{\mathcal{N}_{Q,k}} + V^\kappa P^\kappa)^T E^{\gamma\kappa}(I_{\mathcal{N}_{Q,k}} - V^\kappa P^\kappa) + V^\kappa \widehat{D}_m P^\kappa. \tag{3.12}$$

We also define the extended vector of nodal values to incorporate vector-valued functions $\mathbf{u}$:

$$\overrightarrow{\mathbf{u}^{\kappa}}=\begin{bmatrix}\mathbf{u}(\mathbf{x}_1^{\kappa})\\ \vdots \\ \mathbf{u}(\mathbf{x}_{\mathcal{N}_{Q,k}}^{\kappa})\end{bmatrix},\quad \overrightarrow{\mathbf{u}^{\gamma}}=\begin{bmatrix}\mathbf{u}(\mathbf{x}_1^{\gamma})\\ \vdots \\ \mathbf{u}(\mathbf{x}_{\mathcal{N}_{B,k}}^{\gamma})\end{bmatrix},$$

as well as the Kronecker products

$$\mathbf{M}^{\kappa}=M^{\kappa}\otimes I_p,\quad \mathbf{B}^{\gamma}=B^{\gamma}\otimes I_p,\quad \mathbf{D}_m^{\kappa}=D_m^{\kappa}\otimes I_p,\quad \mathbf{R}^{\gamma\kappa}=R^{\gamma\kappa}\otimes I_p.$$

We still have the SBP property

$$\mathbf{S}_m^{\kappa}=\mathbf{M}^{\kappa}\mathbf{D}_m^{\kappa},\quad \mathbf{E}^{\gamma\kappa}=(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^{\gamma}\mathbf{R}^{\gamma\kappa},\quad \mathbf{S}_m^{\kappa}+(\mathbf{S}_m^{\kappa})^T=\sum_{\gamma\in\partial T_{\kappa}}n_m^{\gamma\kappa}\mathbf{E}^{\gamma\kappa}. \tag{3.13}$$

Remark 3.3. Conceptually, the SBP framework can be further generalized to arbitrary polygonal meshes without any difficulty. We stick to simplex meshes for practical purposes. We only need to store one set of matrices on some reference simplex. Then the local matrices can be acquired through the affine mapping between the reference element and the local element. This is efficient in terms of space complexity, especially for meshes with a large number of elements.

3.3 Nodal DG formulation

Recall the classic DG method (2.17), written as inner products:

$$\Big(\frac{\partial\mathbf{u}_h^{\kappa}}{\partial t},\mathbf{w}_h\Big)_{T_{\kappa}}-\sum_{m=1}^{d}\Big(\mathbf{f}_m(\mathbf{u}_h^{\kappa}),\frac{d\mathbf{w}_h^{\kappa}}{dx_m}\Big)_{T_{\kappa}}=-\sum_{\gamma\in\partial T_{\kappa}}\Big\langle\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^{\kappa},\mathbf{u}_h^{\nu}),\mathbf{w}_h^{\kappa}\Big\rangle_{\gamma}. \tag{3.14}$$

We use the volume quadrature rule to approximate the left hand side, and the surface quadrature rule to approximate the right hand side, replacing the continuous inner products with discrete inner products:

$$\Big(\frac{\partial\mathbf{u}_h^{\kappa}}{\partial t},\mathbf{w}_h\Big)_{T_{\kappa},\omega}-\sum_{m=1}^{d}\Big(\mathbf{f}_m(\mathbf{u}_h^{\kappa}),\frac{d\mathbf{w}_h^{\kappa}}{dx_m}\Big)_{T_{\kappa},\omega}=-\sum_{\gamma\in\partial T_{\kappa}}\Big\langle\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^{\kappa},\mathbf{u}_h^{\nu}),\mathbf{w}_h^{\kappa}\Big\rangle_{\gamma,\tau}. \tag{3.15}$$

Specific to the DG method, we expand $\mathbf{u}_h^{\kappa}$ and $\mathbf{w}_h^{\kappa}$ under the basis $\{p_l(\mathbf{x})\}_{l=1}^{\mathcal{N}_{P,k}}$:

$$\mathbf{u}_h^{\kappa}(t,\mathbf{x})=\sum_{l=1}^{\mathcal{N}_{P,k}}\widehat{\mathbf{u}}_l^{\kappa}(t)p_l(\mathbf{x}),\quad \mathbf{w}_h^{\kappa}(t,\mathbf{x})=\sum_{l=1}^{\mathcal{N}_{P,k}}\widehat{\mathbf{w}}_l^{\kappa}(t)p_l(\mathbf{x}).$$

Define the vectors of polynomial coefficients

$$\overrightarrow{\widehat{\mathbf{u}}^{\kappa}}=\begin{bmatrix}\widehat{\mathbf{u}}_1^{\kappa}\\ \vdots \\ \widehat{\mathbf{u}}_{\mathcal{N}_{P,k}}^{\kappa}\end{bmatrix},\quad \overrightarrow{\widehat{\mathbf{w}}^{\kappa}}=\begin{bmatrix}\widehat{\mathbf{w}}_1^{\kappa}\\ \vdots \\ \widehat{\mathbf{w}}_{\mathcal{N}_{P,k}}^{\kappa}\end{bmatrix},$$

and the vectors of nodal values

$$\overrightarrow{\mathbf{u}^\kappa}=\begin{bmatrix}\mathbf{u}_h(\mathbf{x}_1^\kappa)\\ \vdots \\ \mathbf{u}_h(x^\kappa_{\mathcal{N}_{Q,k}})\end{bmatrix},\quad \overrightarrow{\mathbf{w}^\kappa}=\begin{bmatrix}\mathbf{w}_h(\mathbf{x}_1^\kappa)\\ \vdots \\ \mathbf{w}_h(x^\kappa_{\mathcal{N}_{Q,k}})\end{bmatrix},\quad \overrightarrow{\mathbf{f}_m^\kappa}=\begin{bmatrix}\mathbf{f}_m(\mathbf{u}_1^\kappa)\\ \vdots \\ \mathbf{f}_m(\mathbf{u}^\kappa_{\mathcal{N}_{Q,k}})\end{bmatrix}.$$

Then $\overrightarrow{\mathbf{u}^\kappa}=\mathbf{V}^\kappa\overrightarrow{\widehat{\mathbf{u}}^\kappa}$ and $\overrightarrow{\mathbf{w}^\kappa}=\mathbf{V}^\kappa\overrightarrow{\widehat{\mathbf{w}}^\kappa}$. Likewise we can also define $\overrightarrow{\mathbf{v}^\kappa}$ and $\overrightarrow{U^\kappa}$. On a face $\gamma\in\partial T_\kappa$, let the superscript $\gamma\kappa$ represent the vector of extrapolated nodal values:

$$\begin{aligned}&\overrightarrow{\mathbf{u}^{\gamma\kappa}}=\mathbf{R}^{\gamma\kappa}\overrightarrow{\mathbf{u}^\kappa}=\mathbf{V}^\gamma\overrightarrow{\widehat{\mathbf{u}}^\kappa},\quad \overrightarrow{\mathbf{w}^{\gamma\kappa}}=\mathbf{R}^{\gamma\kappa}\overrightarrow{\mathbf{w}^\kappa}=\mathbf{V}^\gamma\overrightarrow{\widehat{\mathbf{w}}^\kappa},\\ &\overrightarrow{\mathbf{f}_m^{\gamma\kappa}}=\mathbf{R}^{\gamma\kappa}\overrightarrow{\mathbf{f}_m^\kappa},\quad \overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}}=\sum_{m=1}^d n_m^{\gamma\kappa}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}}.\end{aligned}$$

We also put nodal values of the interface numerical flux into a vector

$$\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}=\begin{bmatrix}\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^\kappa(\mathbf{x}_1^\gamma),\mathbf{u}_h^\nu(\mathbf{x}_1^\gamma))\\ \vdots \\ \widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_h^\kappa(\mathbf{x}^\gamma_{\mathcal{N}_{B,k}}),\mathbf{u}_h^\nu(\mathbf{x}^\gamma_{\mathcal{N}_{B,k}}))\end{bmatrix}=\begin{bmatrix}\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_1^{\gamma\kappa},\mathbf{u}_1^{\gamma\nu})\\ \vdots \\ \widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}^{\gamma\kappa}_{\mathcal{N}_{B,k}},\mathbf{u}^{\gamma\nu}_{\mathcal{N}_{B,k}})\end{bmatrix}.$$

Using these notations and discrete operators in Section 3.1 and Section 3.2, we are able to recast (3.15) into a compact matrix-vector formulation:

$$\left(\overrightarrow{\widehat{\mathbf{w}}^\kappa}\right)^T\widehat{\mathbf{M}}^\kappa\frac{d\overrightarrow{\widehat{\mathbf{u}}^\kappa}}{dt}-\sum_{m=1}^d\left(\mathbf{V}^\kappa\widehat{\mathbf{D}}_m\overrightarrow{\widehat{\mathbf{w}}^\kappa}\right)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^\kappa}=-\sum_{\gamma\in\partial T_\kappa}\left(\mathbf{V}^\gamma\overrightarrow{\widehat{\mathbf{w}}^\kappa}\right)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}.$$

Since $\overrightarrow{\widehat{\mathbf{w}}^\kappa}$ can be arbitrary, we obtain

$$\frac{d\overrightarrow{\widehat{\mathbf{u}}^\kappa}}{dt}-(\widehat{\mathbf{M}}^\kappa)^{-1}\sum_{m=1}^d(\mathbf{V}^\kappa\widehat{\mathbf{D}}_m)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^\kappa}=-(\widehat{\mathbf{M}}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}(\mathbf{V}^\gamma)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}.\tag{3.16}$$

This is called weak modal formulation as we evolve the vector $\overrightarrow{\widehat{\mathbf{u}}^\kappa}$. Applying $\mathbf{V}^\kappa$ to (3.16), we come up with the weak nodal formulation that describes the evolution of $\overrightarrow{\mathbf{u}^\kappa}$:

$$\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt}-(\mathbf{M}^\kappa)^{-1}\sum_{m=1}^d(\mathbf{V}^\kappa\widehat{\mathbf{D}}_m\mathbf{P}^\kappa)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^\kappa}=-(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}(\mathbf{V}^\gamma\mathbf{P}^\kappa)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}},\tag{3.17}$$

where we use the relation $\mathbf{V}^\kappa(\widehat{\mathbf{M}}^k)^{-1}=(\mathbf{M}^\kappa)^{-1}(\mathbf{P}^\kappa)^T$. It is a special case of the more general weak nodal DG formulation

$$\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt}-(\mathbf{M}^\kappa)^{-1}\sum_{m=1}^d(\mathbf{D}_m^\kappa)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^\kappa}=-(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}},\tag{3.18}$$

by choosing $\mathbf{R}^{\gamma\kappa}=\mathbf{V}^{\gamma}\mathbf{P}^{\kappa}$ and $\mathbf{D}_m^{\kappa}=\mathbf{V}^{\kappa}\widehat{\mathbf{D}}_m\mathbf{P}^{\kappa}$. According to the SBP property (3.8), we also deduce the equivalent strong nodal DG formulation:

$$\begin{aligned}\frac{d\overrightarrow{\mathbf{u}^{\kappa}}}{dt}+\underbrace{\sum_{m=1}^{d}\mathbf{D}_m^{\kappa}\overrightarrow{\mathbf{f}_m^{\kappa}}}_{\text{Difference term}}&=\underbrace{(\mathbf{M}^{\kappa})^{-1}\sum_{\gamma\in\partial T_{\kappa}}\Big(\sum_{m=1}^{d}n_m^{\gamma\kappa}\mathbf{E}^{\gamma\kappa}\overrightarrow{\mathbf{f}_m^{\kappa}}-(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big)}_{\text{Simultaneous approximation term}}\\&=(\mathbf{M}^{\kappa})^{-1}\sum_{\gamma\in\partial T_{\kappa}}(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\Big(\sum_{m=1}^{d}n_m^{\gamma\kappa}\overrightarrow{\mathbf{f}_m^{\gamma\kappa}}-\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big)\\&=(\mathbf{M}^{\kappa})^{-1}\sum_{\gamma\in\partial T_{\kappa}}(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\Big(\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}}-\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big).\end{aligned}\tag{3.19}$$

It can be viewed as a spectral collocation method with penalty type terms on element interfaces [48]. These penalty terms are called simultaneous approximation terms (SATs) by the SBP community.

Remark 3.4. We should emphasize the caveats concerning the link between modal and nodal DG formulations. By taking interpolation, the modal formulation only implies a specific nodal formulation (with particular choices of $\mathbf{R}^{\gamma\kappa}$ and $\mathbf{D}_m^{\kappa}$). On the other hand, by taking projection, all nodal formulations (with any $\mathbf{R}^{\gamma\kappa}$ and $\mathbf{D}_m^{\gamma}$ satisfying the exactness property and the SBP property) will lead to the modal formulation. In fact, applying $\mathbf{P}^{\kappa}$ to (3.18) and setting $\widehat{\mathbf{u}}^{\kappa}=\mathbf{P}^{\kappa}\mathbf{u}^{\kappa}$ yield

$$\frac{d\overrightarrow{\mathbf{u}^{\kappa}}}{dt}-\mathbf{P}^{\kappa}(\mathbf{M}^{\kappa})^{-1}\sum_{m=1}^{d}(\mathbf{D}_m^{\kappa})^{T}\mathbf{M}^{\kappa}\overrightarrow{\mathbf{f}_m^{\kappa}}=-\mathbf{P}^{\kappa}(\mathbf{M}^{\kappa})^{-1}\sum_{\gamma\in\partial T_{\kappa}}(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}.$$

This reduces to (3.16) because of the exactness properties:

$$\begin{aligned}&\mathbf{P}^{\kappa}(\mathbf{M}^{\kappa})^{-1}(\mathbf{D}_m^{\kappa})^{T}=(\widehat{\mathbf{M}}^{\kappa})^{-1}(\mathbf{D}_m^{\kappa}\mathbf{V}^{\kappa})^{T}=(\widehat{\mathbf{M}}^{\kappa})^{-1}(\mathbf{V}^{\kappa}\widehat{\mathbf{D}}_m)^{T},\\&\mathbf{P}^{\kappa}(\mathbf{M}^{\kappa})^{-1}(\mathbf{R}^{\gamma\kappa})^{T}=(\widehat{\mathbf{M}}^{\kappa})^{-1}(\mathbf{R}^{\gamma\kappa}\mathbf{V}^{\kappa})^{T}=(\widehat{\mathbf{M}}^{\kappa})^{-1}(\mathbf{V}^{\gamma})^{T}.\end{aligned}$$

The reason for such asymmetric relation is the fact that $\mathcal{N}_{Q,k}\geq\mathcal{N}_{P,k}$. Then $\mathbf{V}^{\kappa}$ is not surjective, and $\mathbf{P}^{\kappa}$ is not injective.

4 Entropy stable DG method with collocated surface nodes

The nodal DG formulation (3.19) (and (3.18)) does not satisfy any entropy inequality (even the L^2 stability). For an entropy function U, the discrete total entropy is given by

$$\sum_{\kappa=1}^{K}(U(\mathbf{u}_h),1)_{T_{\kappa},\omega}=\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^{\kappa}}\Big)^{T}M^{\kappa}\overrightarrow{U^{\kappa}}=\sum_{\kappa=1}^{K}\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^{\kappa}U_j^{\kappa}.$$

Then the entropy growth rate of (3.18) is

$$\begin{aligned}
&\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa \overrightarrow{U^\kappa}\Big) = \sum_{\kappa=1}^{K}\sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa \frac{dU_j^\kappa}{dt} \\
&= \sum_{\kappa=1}^{K}\sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa (\mathbf{v}_j^\kappa)^T \frac{d\mathbf{u}_j^\kappa}{dt} = \sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{v}^\kappa}\Big)^T \mathbf{M}^\kappa \frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt} \\
&= \sum_{\kappa=1}^{K}\Big(\sum_{m=1}^{d}\Big(\overrightarrow{\mathbf{v}^\kappa}\Big)^T (\mathbf{D}_m^\kappa)^T \mathbf{M}^\kappa \overrightarrow{\mathbf{f}_m^\kappa} - \sum_{\gamma\in\partial T_\kappa}\Big(\overrightarrow{\mathbf{v}^\kappa}\Big)^T (\mathbf{R}^{\gamma\kappa})^T \mathbf{B}^\gamma \overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big).
\end{aligned}$$

However, we can not characterize the first term as the chain rule

$$\Big(\mathbf{f}_m(\mathbf{u}_h), \frac{\partial \mathbf{v}(\mathbf{u}_h)}{\partial x_m}\Big)_{T_\kappa} = \Big(\frac{\partial \psi_m(\mathbf{v}(\mathbf{u}_h))}{\partial x_m}, 1\Big)_{T_\kappa} \quad \text{(for the continuous solution polynomial)}$$

is no longer valid at the discrete level. That is,

$$\Big(\overrightarrow{\mathbf{v}^\kappa}\Big)^T (\mathbf{D}_m^\kappa)^T \mathbf{M}^\kappa \overrightarrow{\mathbf{f}_m^\kappa} \neq \Big(\overrightarrow{\psi_m^\kappa}\Big)^T D_m^\kappa M^\kappa \overrightarrow{1^\kappa} \quad \text{(for the discrete nodal values).}$$

In this section, we will modify the scheme (3.19) and make it entropy stable, in the special case that the volume quadrature rule has collocated surface quadrature nodes, and that the boundary matrix $E^{\gamma\kappa}$ is diagonal (i.e. the third case in Remark 3.2). The modification amounts to the flux differencing technique in [4, 5, 12, 29].

4.1 Flux differencing

Identity (2.12) satisfied by entropy conservative fluxes serves as the discrete analogue of the chain rule. The flux differencing technique, in which we replace the difference term in (3.19) with high order difference operation of entropy conservative fluxes, is the key to entropy balance within an element. The modified nodal DG method reads

$$\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt} + \underbrace{2\sum_{m=1}^{d} \mathbf{D}_m^\kappa \circ \mathbf{F}_{m,S}\Big(\overrightarrow{\mathbf{u}^\kappa}, \overrightarrow{\mathbf{u}^\kappa}\Big)\overrightarrow{\mathbf{1}^\kappa}}_{\text{Flux differencing term}} = \underbrace{(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}(\mathbf{R}^{\gamma\kappa})^T \mathbf{B}^\gamma\Big(\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}} - \overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big)}_{\text{Simultaneous approximation term}}, \tag{4.1}$$

where $\circ$ denotes the Hadamard (pointwise) product of vectors and matrices, and $\mathbf{F}_{m,S}(\cdot,\cdot)$ is the matrix of pairwise combinations of entropy conservative fluxes [19, 20]:

$$\mathbf{F}_{m,S}(\overrightarrow{\mathbf{u}_L}, \overrightarrow{\mathbf{u}_R}) = \begin{bmatrix} \operatorname{diag}(\mathbf{f}_{m,S}(\mathbf{u}_{L,1}, \mathbf{u}_{R,1})) & \cdots & \operatorname{diag}(\mathbf{f}_{m,S}(\mathbf{u}_{L,1}, \mathbf{u}_{R,\mathcal{N}_R})) \\ \vdots & \ddots & \vdots \\ \operatorname{diag}(\mathbf{f}_{m,S}(\mathbf{u}_{L,\mathcal{N}_L}, \mathbf{u}_{R,1})) & \cdots & \operatorname{diag}(\mathbf{f}_{m,S}(\mathbf{u}_{L,\mathcal{N}_L}, \mathbf{u}_{R,\mathcal{N}_R})) \end{bmatrix},$$

for $\overrightarrow{\mathbf{u}_L^\kappa}\in\mathbb{R}^{p\mathcal{N}_L}$ and $\overrightarrow{\mathbf{u}_R^\kappa}\in\mathbb{R}^{p\mathcal{N}_R}$. We can clarify the involved flux differencing term by writing down the evolution of the nodal values

$$\frac{d\mathbf{u}_j^\kappa}{dt}+2\sum_{m=1}^{d}\sum_{l=1}^{\mathcal{N}_{Q,k}}D_{m,jl}^\kappa\mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)=\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}R_{sj}^{\gamma\kappa}\frac{\tau_s^\gamma}{\omega_j^\kappa}(\mathbf{f}_{\mathbf{n},s}^{\gamma\kappa}-\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu})). \tag{4.2}$$

We will take a deeper look into the flux differencing term in Appendix A, showing that if the entropy conservative fluxes are *separable*, flux differencing is actually equivalent to the splitting technique in [36,38,39,65,69,71]. Before proving the main result of this section, we first give a lemma indicating the effects of flux differencing on primary conservation and entropy growth.

Lemma 4.1. *If for each $1\le m\le d$, $\mathbf{f}_{m,S}$ is an entropy conservative flux with respect to some entropy function U, then*

$$\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{M}^\kappa\left(2\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\right)=\sum_{\gamma\in\partial T_\kappa}n_m^{\gamma\kappa}\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}, \tag{4.3}$$

$$\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{M}^\kappa\left(2\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\right) \tag{4.4}$$

$$=\sum_{\gamma\in\partial T_\kappa}n_m^{\gamma\kappa}\left(\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}-\left(\overrightarrow{\psi_m^\kappa}\right)^T E^{\kappa\gamma}\overrightarrow{1^\kappa}\right). \tag{4.5}$$

Moreover, if $E^{\kappa\gamma}$ is diagonal, we have the simplified result

$$\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{M}^\kappa\left(2\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\right)=\sum_{\gamma\in\partial T_\kappa}n_m^{\gamma\kappa}\left(\overrightarrow{\mathbf{1}^\gamma}\right)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_m^{\gamma\kappa}}, \tag{4.6}$$

$$\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{M}^\kappa\left(2\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\right)=\sum_{\gamma\in\partial T_\kappa}n_m^{\gamma\kappa}\left(\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_m^{\gamma\kappa}}-\left(\overrightarrow{\psi_m^{\gamma\kappa}}\right)^T B^\gamma\overrightarrow{1^\gamma}\right). \tag{4.7}$$

Proof. Since $\mathbf{M}^\kappa$ is diagonal,

$$\mathbf{M}^\kappa\left(\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\right)=\mathbf{S}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right),$$

and by the symmetry of $\mathbf{f}_{m,S}$, $\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)$ is a symmetric matrix. Then

$$\begin{aligned}
&\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{M}^\kappa\left(2\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\right)\\
=&2\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{S}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\\
=&\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T(\mathbf{S}_m^\kappa+(\mathbf{S}_m^\kappa)^T)\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\quad\text{(by symmetry of }\mathbf{f}_{m,S})\\
=&\sum_{\gamma\in\partial T_\kappa}n_m^{\gamma\kappa}\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}\quad\text{(by the SBP property),}
\end{aligned}$$

and

$$\begin{aligned}
&\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{M}^\kappa \left(2\mathbf{D}_m^\kappa \circ \mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa}, \overrightarrow{\mathbf{u}^\kappa}\right) \overrightarrow{\mathbf{1}^\kappa}\right) \\
=&2\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{S}_m^\kappa \circ \mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa}, \overrightarrow{\mathbf{u}^\kappa}\right) \overrightarrow{\mathbf{1}^\kappa} \\
=&\sum_{\gamma \in \partial T_\kappa} n_m^{\gamma\kappa} \left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{E}^{\gamma\kappa} \circ \mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa}, \overrightarrow{\mathbf{u}^\kappa}\right) \overrightarrow{\mathbf{1}^\kappa} + \left(\overrightarrow{\mathbf{v}^\kappa}\right)^T (\mathbf{S}_m^\kappa - (\mathbf{S}_m^\kappa)^T) \circ \mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa}, \overrightarrow{\mathbf{u}^\kappa}\right) \overrightarrow{\mathbf{1}^\kappa}.
\end{aligned}$$

The second term equals

$$\begin{aligned}
&\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T (\mathbf{S}_m^\kappa - (\mathbf{S}_m^\kappa)^T) \circ \mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa}, \overrightarrow{\mathbf{u}^\kappa}\right) \overrightarrow{\mathbf{1}^\kappa} \\
=&\sum_{j=1}^{\mathcal{N}_{Q,k}} \sum_{l=1}^{\mathcal{N}_{Q,k}} (\mathbf{v}_j^\kappa)^T (S_{m,jl}^\kappa - S_{m,lj}^\kappa) \mathbf{f}_{m,S}(\mathbf{u}_j^\kappa, \mathbf{u}_l^\kappa) \\
=&\sum_{j=1}^{\mathcal{N}_{Q,k}} \sum_{l=1}^{\mathcal{N}_{Q,k}} S_{m,jl}^\kappa (\mathbf{v}_j^\kappa - \mathbf{v}_l^\kappa)^T \mathbf{f}_{m,S}(\mathbf{u}_j^\kappa, \mathbf{u}_l^\kappa) \quad \text{(by symmetry of } \mathbf{f}_{m,S}) \\
=&\sum_{j=1}^{\mathcal{N}_{Q,k}} \sum_{l=1}^{\mathcal{N}_{Q,k}} S_{m,jl}^\kappa (\psi_{m,j}^\kappa - \psi_{m,l}^\kappa) \quad \text{(by entropy conservation of } \mathbf{f}_{m,S}) \\
=&\left(\overrightarrow{\psi_m^\kappa}\right)^T (S_m^\kappa - (S_m^\kappa)^T) \overrightarrow{1^\kappa} = -\sum_{\gamma \in \partial T_\kappa} n_m^{\gamma\kappa} \left(\overrightarrow{\psi_m^\kappa}\right)^T E^{\kappa\gamma} \overrightarrow{1^\kappa} \quad \text{(by relation (3.11)).}
\end{aligned}$$

Hence (4.3) and (4.4) are proved. Moreover, if $E^{\gamma\kappa}$ is diagonal,

$$\mathbf{E}^{\gamma\kappa} \circ \mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa}, \overrightarrow{\mathbf{u}^\kappa}\right) \overrightarrow{\mathbf{1}^\kappa} = \mathbf{E}^{\gamma\kappa} \overrightarrow{\mathbf{f}_m^\kappa} = (\mathbf{R}^{\gamma\kappa})^T \mathbf{B}^\gamma \overrightarrow{\mathbf{f}_m^{\gamma k}} \quad \text{(by consistency of } \mathbf{f}_{m,S}), \tag{4.8}$$

which implies (4.6) and (4.7). □

We are ready to provide the main theorem, which states that the nodal DG method (4.1) is conservative, entropy stable, and maintains high order accuracy, under the assumptions that

1. The volume quadrature rule has collocated surface quadrature nodes ($\{\mathbf{x}_s^\gamma\}_{s=1}^{\mathcal{N}_{B,k}}$ is a subset of $\{\mathbf{x}_j^\kappa\}_{j=1}^{\mathcal{N}_{Q,k}}$ for $\gamma \in \partial T_\kappa$), $R^{\kappa\gamma}$ is a simple restriction onto γ, and $E^{\kappa\gamma}$ is diagonal.
2. The simplex meshes $\{\mathcal{T}_h\}$, parameterized by h, are shape regular and quasi-uniform.
3. All mappings and numerical fluxes (e.g. $\mathbf{v}(\mathbf{u})$, $\mathbf{f}_m(\mathbf{u})$, $\mathbf{f}_{m,S}(\mathbf{u}_L, \mathbf{u}_R)$, etc) are smooth and Lipschitz continuous.
4. $\mathbf{f}_{m,S}$ is entropy conservative, and $\widehat{\mathbf{f}}_\mathbf{n}$ is entropy stable with respect to an arbitrary given entropy function U.

Theorem 4.1. *If all assumptions above hold, then the scheme* (4.1) *is consistent in that for a smooth solution* $\mathbf{u}$ *of* (2.1)*, the local truncation error is of high order:*

$$\frac{d\mathbf{u}_j^\kappa}{dt}+2\sum_{m=1}^{d}\sum_{l=1}^{\mathcal{N}_{Q,k}}D_{m,jl}^\kappa \mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)-\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}R_{sj}^{\gamma\kappa}\frac{\tau_s^\gamma}{\omega_j^\kappa}(\mathbf{f}_{\mathbf{n},s}^{\gamma\kappa}-\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu}))=\mathcal{O}(h^k). \tag{4.9}$$

It is also conservative and entropy stable with respect to U *in that*

$$\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{1}^\kappa}\Big)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{u}^\kappa}\Big)=0,\quad \frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa\overrightarrow{U^\kappa}\Big)\le 0. \tag{4.10}$$

Proof. Consistency: $\mathbf{u}$ is single-valued at interfaces. Since $R^{\gamma\kappa}$ and $R^{\gamma\nu}$ are simple restrictions, $\mathbf{u}_s^{\gamma\kappa}=\mathbf{u}_s^{\gamma\nu}=\mathbf{u}(\mathbf{x}_s^\gamma)$ and $\mathbf{f}_{\mathbf{n},s}^{\gamma\kappa}=\mathbf{f}_{\mathbf{n}}(\mathbf{u}(\mathbf{x}_s^\gamma))$. Then the simultaneous approximation term vanishes. It suffices to show that the flux differencing term is of high order, i.e.,

$$2\sum_{l=1}^{\mathcal{N}_{Q,k}}D_{m,jl}^\kappa \mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)-\frac{\partial \mathbf{f}_m(\mathbf{u})}{\partial x_m}(\mathbf{x}_j^\kappa)=\mathcal{O}(h^k).$$

Let $\widetilde{\mathbf{f}}_{m,S}(\mathbf{x},\mathbf{y})=\mathbf{f}_{m,S}(\mathbf{u}(x),\mathbf{u}(y))$ and $\widetilde{\mathbf{f}}_m(\mathbf{x})=\mathbf{f}_m(\mathbf{u}(\mathbf{x}))$. Then $\widetilde{\mathbf{f}}_{m,S}$ is also symmetric and consistent such that $\widetilde{\mathbf{f}}_{m,S}(\mathbf{x},\mathbf{x})=\widetilde{\mathbf{f}}_m(\mathbf{x})$. Therefore

$$\frac{\partial\widetilde{\mathbf{f}}_m}{\partial x_m}(\mathbf{x})=\frac{\partial\widetilde{\mathbf{f}}_{m,S}}{\partial x_m}(\mathbf{x},\mathbf{x})+\frac{\partial\widetilde{\mathbf{f}}_{m,S}}{\partial y_m}(\mathbf{x},\mathbf{x})=2\frac{\partial\widetilde{\mathbf{f}}_{m,S}}{\partial y_m}(\mathbf{x},\mathbf{x}).$$

According to the approximation property of local difference matrix D_m^κ,

$$\begin{aligned}&2\sum_{l=1}^{\mathcal{N}_{Q,k}}D_{m,jl}^\kappa \mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)=2\frac{\partial\widetilde{\mathbf{f}}_{m,S}}{\partial y_m}(\mathbf{x}_j^\kappa,\mathbf{x}_j^\kappa)+\mathcal{O}(h^k)\\ =&\frac{\partial\widetilde{\mathbf{f}}_m}{\partial x_m}(\mathbf{x}_j^\kappa)+\mathcal{O}(h^k)=\frac{\partial \mathbf{f}_m(\mathbf{u})}{\partial x_m}(\mathbf{x}_j^\kappa)+\mathcal{O}(h^k).\end{aligned}$$

Conservation and entropy stability: according to (4.6) and (4.7),

$$\begin{aligned}&\frac{d}{dt}\Big(\Big(\overrightarrow{\mathbf{1}^\kappa}\Big)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{u}^\kappa}\Big)\\ =&\sum_{\gamma\in\partial T_\kappa}\Big(\overrightarrow{\mathbf{1}^\gamma}\Big)^T\mathbf{B}^\gamma\Big(\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}}-\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big)-\sum_{m=1}^{d}\sum_{\gamma\in\partial T_\kappa}n_m^{\gamma\kappa}\Big(\overrightarrow{\mathbf{1}^\gamma}\Big)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_m^{\gamma\kappa}}\end{aligned}$$

$$=-\sum_{\gamma\in\partial T_\kappa}\left(\overrightarrow{\mathbf{1}^\gamma}\right)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}=-\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu}),$$

$$\begin{aligned}
&\frac{d}{dt}\left(\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{u}^\kappa}\right)\\
&=\sum_{\gamma\in\partial T_\kappa}\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T\mathbf{B}^\gamma\left(\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}}-\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\right)-\sum_{m=1}^{d}\sum_{\gamma\in\partial T_\kappa}n_m^{\gamma\kappa}\left(\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_m^{\gamma\kappa}}-\left(\overrightarrow{\psi_m^{\gamma\kappa}}\right)^TB^\gamma\overrightarrow{1^\gamma}\right)\\
&=\sum_{\gamma\in\partial T_\kappa}\left(\left(\overrightarrow{\psi_{\mathbf{n}}^{\gamma\kappa}}\right)^TB^\gamma\overrightarrow{1^\gamma}-\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\right)\\
&=\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma(\psi_{\mathbf{n},s}^{\gamma\kappa}-(\mathbf{v}_s^{\gamma\kappa})^T\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu})).
\end{aligned}$$

We are only left with interface terms. Summing over κ gives us

$$\begin{aligned}
&\frac{d}{dt}\left(\sum_{\kappa=1}^{K}\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{u}^\kappa}\right)=-\sum_{\kappa=1}^{K}\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu})\\
=&-\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma(\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu})+\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\nu}}(\mathbf{u}_s^{\gamma\nu},\mathbf{u}_s^{\gamma\kappa}))=0,\\
&\frac{d}{dt}\left(\sum_{\kappa=1}^{K}\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{u}^\kappa}\right)=\sum_{\kappa=1}^{K}\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma(\psi_{\mathbf{n},s}^{\gamma\kappa}-(\mathbf{v}_s^{\gamma\kappa})^T\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu}))\\
=&\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma((\mathbf{v}_s^{\gamma\nu}-\mathbf{v}_s^{\gamma\kappa})^T\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu})-(\psi_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\kappa}-\psi_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\nu}))\le 0,
\end{aligned}$$

by entropy stability of $\widehat{\mathbf{f}}_{\mathbf{n}}$. We again use the fact that $R^{\gamma\kappa}$ and $R^{\gamma\nu}$ are simple restrictions, so that $\mathbf{v}_s^{\gamma\kappa}=\mathbf{v}(\mathbf{u}_s^{\gamma\kappa})$ and $\psi_{\mathbf{n},s}^{\gamma\kappa}=\psi_{\mathbf{n}}(\mathbf{v}_s^{\gamma\kappa})$. □

Remark 4.1. The computational cost of (4.1) is dominated by the number of entropy conservative flux evaluations, which is of the order $\mathcal{O}(\mathcal{N}_{Q,k}^2)=\mathcal{O}(k^{2d})$ in each element (by assuming that $\mathcal{N}_{Q,k}=\mathcal{O}(k^d)$). In contrast, for Cartesian meshes, the difference matrices are sparse (due to tensor product), and the number of flux evaluations is of the order $\mathcal{O}(k^{d+1})$.

4.2 Accuracy test

We test the accuracy of scheme (4.1) for the two-dimensional Burgers equation

$$\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial u^2}{\partial x_1}+\frac{1}{2}\frac{\partial u^2}{\partial x_2}=0,\quad \mathbf{x}\in[0,1]^2, \tag{4.11}$$

with periodic boundary condition and initial data $u(0,\mathbf{x})=0.5\sin(2\pi(x_1+x_2))$. We can easily compute the exact solution by tracing back characteristic lines along the diagonal direction. The entropy function is taken be the hyperbolic cosine function $U=\cosh u$, such that

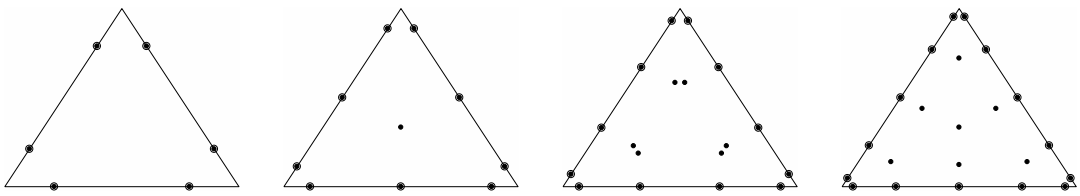
$$v=\sinh u,\quad F_1=F_2=u\cosh u-\sinh u,\quad \psi_1=\psi_2=\Big(\frac{u^2}{2}+1\Big)\cosh u-u\sinh u.$$

The entropy conservative fluxes are given by

$$f_{1,S}(u_L,u_R)=f_{2,S}(u_L,u_R)=\frac{(\frac{u_L^2}{2}+1)\cosh u_L-u_L\sinh u_L-(\frac{u_R^2}{2}+1)\cosh u_R+u_R\sinh u_R}{\sinh u_R-\sinh u_L}.$$

If $|u_L-u_R|$ is small, the division suffers from the round-off effect, and we use the first 5 terms of Taylor series to approximate the numerator and the denominator. The cutoff value for $|u_L-u_R|$ is 10^{-3}. The local Lax-Friedrichs flux will be employed on element interfaces.

The test is performed on a hierarchy of unstructured triangular meshes generated by the Gmsh software [40]. To implement SBP operators on those triangles, we need to find a quadrature rule that achieves volume and surface accuracy simultaneously. For the surface accuracy, we put $(k+1)$ Legendre-Gauss points along each edge, and for the volume accuracy, we use the numerical package in [88] to obtain degree $(2k-1)$ quadrature rules with collocated Legendre-Gauss edge nodes. The distribution of quadrature points on the equilateral triangle is illustrated in Fig. 1. Then we assemble the mass matrices and difference matrices using (3.12).



(a) $k=1,\mathcal{N}_{Q,k}=6$ (b) $k=2,\mathcal{N}_{Q,k}=10$ (c) $k=3,\mathcal{N}_{Q,k}=18$ (d) $k=4,\mathcal{N}_{Q,k}=22$

Figure 1: Degree $2k-1$ quadrature rules on triangles with collocated Legendre-Gauss edge nodes for $k=1,2,3,4$. Dots are quadrature points for the triangle, and circles are quadrature points for the edges. The symbols overlap as edge nodes play both roles.

The scheme (4.1) is evolved in time with the third order strong stability preserving (SSP) Runge-Kutta method [43,76]. We would like to compute with $k=2,3,4$. The time step is proportional to $h^{(k+1)/3}$, so that the time error will be dominated by the space error. Numerical errors and orders of convergence at $t=0.1$ (before the shock wave emerges) are displayed in Table 1. We observe reduced rate of convergence (less than the optimal $(k+1)$-th order), especially for the L^∞ error. It is probably due to the algebraic accuracy of quadrature rules. It was shown in [13,55] that under the standard assumptions of smooth

Table 1: Errors and orders of convergence of (4.1) for the two-dimensional Burgers equation at $t=0.1$. Degree $(2k-1)$ volume quadrature rules are used.

k	h	L^1 error	order	L^2 error	order	L^∞ error	order
2	1/16	1.324e-03		3.182e-03		5.708e-02	
	1/32	2.337e-04	2.503	6.825e-04	2.221	1.577e-02	1.856
	1/64	3.800e-05	2.620	1.362e-04	2.326	4.701e-03	1.746
	1/128	5.628e-06	2.756	2.380e-05	2.516	1.222e-03	1.944
	1/256	8.219e-07	2.776	3.986e-06	2.578	2.329e-04	2.391
3	1/16	1.811e-04		5.932e-04		1.851e-02	
	1/32	2.376e-05	2.930	1.010e-04	2.555	4.591e-03	2.012
	1/64	2.225e-06	3.417	1.055e-05	3.258	6.936e-04	2.727
	1/128	1.977e-07	3.492	1.106e-06	3.253	1.135e-04	2.611
	1/256	1.818e-08	3.443	1.181e-07	3.228	1.238e-05	3.196
4	1/8	3.494e-04		1.130e-03		2.753e-02	
	1/16	3.514e-05	3.314	1.367e-04	3.047	5.646e-03	2.286
	1/32	2.609e-06	3.752	1.334e-05	3.357	9.824e-04	2.523
	1/64	1.264e-07	4.368	7.157e-07	4.220	8.201e-05	3.582
	1/128	5.226e-09	4.596	3.490e-08	4.358	5.968e-06	3.780

solution and shape regular meshes, we need an at least degree $2k$ volume quadrature and an at least degree $(2k+1)$ surface quadrature to attain optimal convergence. Both requirements are one degree higher than the minimal requirements of SBP operators. Here the surface (Legendre-Gauss) quadrature rule is accurate enough, but the volume quadrature rule is not.

5 Entropy stable DG method on general set of nodes

In this section, we would like to extend the entropy stable DG methodology to arbitrary volume and surface quadrature rules. Without the collocated surface nodes assumption, the scheme (4.1) is not entropy stable, as we are facing some new obstacles:

1. The extrapolation matrix $R^{\gamma\kappa}$ is not a restriction. Then $\mathbf{v}_s^{\gamma\kappa} \neq \mathbf{v}(\mathbf{u}_s^{\gamma\kappa})$ and $\psi_{\mathbf{n},s}^{\gamma\kappa} \neq \psi_{\mathbf{n}}(\mathbf{v}_s^{\gamma\kappa})$. The sign of $(\mathbf{v}_s^{\gamma\nu}-\mathbf{v}_s^{\gamma\kappa})^T\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\mathbf{u}_s^{\gamma\kappa},\mathbf{u}_s^{\gamma\nu})-(\psi_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\kappa}-\psi_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\nu})$ is indeterminate.
2. The boundary matrix $E^{\gamma\kappa}$ is dense, and the identity (4.8) is not valid. We are not able to simplify the term $\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^k},\overrightarrow{\mathbf{u}^k}\right)\overrightarrow{\mathbf{1}^k}$.

We solve the first issue by defining the *entropy-extrapolated* nodal values. Set $\widetilde{\overrightarrow{\mathbf{u}}^{\gamma k}}$ and $\widetilde{\overrightarrow{\psi}_m^{\gamma k}}$ such that $\widetilde{\mathbf{u}}_s^{\gamma\kappa}=\mathbf{u}(\mathbf{v}_s^{\gamma\kappa})$ and $\widetilde{\psi}_{m,s}^{\gamma\kappa}=\psi_m(\mathbf{v}_s^{\gamma\kappa})$. We require the interface numerical flux to

depend on those entropy-extrapolated values:

$$\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}=\begin{bmatrix}\widehat{\mathbf{f}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_1^{\gamma\kappa},\widetilde{\mathbf{u}}_1^{\gamma\nu})\\ \vdots\\ \widehat{\mathbf{f}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_{\mathcal{N}_{B,k}}^{\gamma\kappa},\widetilde{\mathbf{u}}_{\mathcal{N}_{B,k}}^{\gamma\nu})\end{bmatrix}.$$

In order to cope with the second issue, we design certain augmented discrete operators that satisfy SBP property with diagonal boundary matrices. Two possible approaches, i.e., the hybridized SBP operators in [7,8] and the global SBP operators in [19,20] will be covered. Besides, we also consider the "brute force" approach in [1].

5.1 Approach 1: hybridized SBP operators

We start to analyze the hybridized SBP operators approach in [7,8] by Chan. The key idea is to combine volume nodes and surface nodes together. Given $1\leq\kappa\leq K$, for simplicity of notations, we assume that T_κ is a triangular element with three edges (it will certainly work on the higher dimensional simplex with $(d+1)$ faces):

$$\partial T_\kappa=\gamma\cup\sigma\cup\eta.$$

We define the hybridized vector of nodal values, by adding entropy-extrapolated values on faces:

$$\overrightarrow{\mathbf{u}^{\kappa,h}}=\begin{bmatrix}\overrightarrow{\mathbf{u}^{\kappa}}\\ \overrightarrow{\widetilde{\mathbf{u}}^{\gamma\kappa}}\\ \overrightarrow{\widetilde{\mathbf{u}}^{\sigma\kappa}}\\ \overrightarrow{\widetilde{\mathbf{u}}^{\eta\kappa}}\end{bmatrix},\quad \overrightarrow{\mathbf{v}^{\kappa,h}}=\begin{bmatrix}\overrightarrow{\mathbf{v}^{\kappa}}\\ \overrightarrow{\mathbf{v}^{\gamma\kappa}}\\ \overrightarrow{\mathbf{v}^{\sigma\kappa}}\\ \overrightarrow{\mathbf{v}^{\eta\kappa}}\end{bmatrix}.$$

The hybridized mass matrix and boundary matrices on T_κ are all diagonal:

$$M^{\kappa,h}=\begin{bmatrix}M^\kappa&0&0&0\\0&B^\gamma&0&0\\0&0&B^\sigma&0\\0&0&0&B^\tau\end{bmatrix},\quad E^{\gamma\kappa,h}=\begin{bmatrix}0&0&0&0\\0&B^\gamma&0&0\\0&0&0&0\\0&0&0&0\end{bmatrix},\quad E^{\sigma\kappa,h}=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&B^\sigma&0\\0&0&0&0\end{bmatrix},\quad E^{\eta\kappa,h}=\begin{bmatrix}0&0&0&0\\0&0&0&0\\0&0&0&0\\0&0&0&B^\eta\end{bmatrix},$$

and for each $1\leq m\leq d$, the hybridized difference matrix is

$$D_m^{\kappa,h}=\begin{bmatrix}D_m^\kappa-\frac{1}{2}(M^\kappa)^{-1}(n_m^{\gamma\kappa}E^{\gamma\kappa}+n_m^{\sigma\kappa}E^{\sigma\kappa}+n_m^{\eta\kappa}E^{\eta\kappa}) & \frac{1}{2}n_m^{\gamma\kappa}(M^\kappa)^{-1}(R^{\gamma\kappa})^TB^\gamma & \frac{1}{2}n_m^{\sigma\kappa}(M^\kappa)^{-1}(R^{\sigma\kappa})^TB^\sigma & \frac{1}{2}n_m^{\eta\kappa}(M^\kappa)^{-1}(R^{\eta\kappa})^TB^\eta\\ -\frac{1}{2}n_m^{\gamma\kappa}R^{\gamma\kappa} & \frac{1}{2}n_m^{\gamma\kappa}I_{\mathcal{N}_{B,k}} & 0 & 0\\ -\frac{1}{2}n_m^{\sigma\kappa}R^{\sigma\kappa} & 0 & \frac{1}{2}n_m^{\sigma\kappa}I_{\mathcal{N}_{B,k}} & 0\\ -\frac{1}{2}n_m^{\eta\kappa}R^{\eta\kappa} & 0 & 0 & \frac{1}{2}n_m^{\eta\kappa}I_{\mathcal{N}_{B,k}}\end{bmatrix}.$$

The next theorem follows immediately from the exactness property and SBP property of the original operators. Proof will be omitted.

Theorem 5.1. *The hybridized discrete operators satisfy the following conditions:*

1. *Exactness:*

$$D_m^{\kappa,h}\begin{bmatrix} V^\kappa \\ V^\gamma \\ V^\sigma \\ V^\eta \end{bmatrix} = \begin{bmatrix} V^\kappa \widehat{D}_m \\ 0 \\ 0 \\ 0 \end{bmatrix}. \quad \textit{In particular, } D_m^{\kappa,h}\overrightarrow{1^{\kappa,h}} = \overrightarrow{0^{\kappa,h}}. \tag{5.1}$$

2. *Summation-by-parts:*

$$M^{\kappa,h}D_m^{\kappa,h} + (D_m^{\kappa,h})^T M^{\kappa,h} = \sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa} E^{\gamma\kappa,h}. \tag{5.2}$$

With the hybridized SBP operators at hand, we develop the nodal DG method

$$\begin{aligned} &\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt} + \underbrace{2\mathbf{L}^{\kappa,h}\Big(\sum_{m=1}^{d}\mathbf{D}_m^{\kappa,h}\circ\mathbf{F}_{m,S}\Big(\overrightarrow{\mathbf{u}^{\kappa,h}},\overrightarrow{\mathbf{u}^{\kappa,h}}\Big)\overrightarrow{\mathbf{1}^{\kappa,h}}\Big)}_{\text{Hybridized flux differencing term}} \\ &= \underbrace{(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\Big(\overrightarrow{\widetilde{\mathbf{f}}_{\mathbf{n}}^{\gamma\kappa}} - \overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big)}_{\text{Simultaneous approximation term}}, \end{aligned} \tag{5.3}$$

where

$$L^{\kappa,h} = \begin{bmatrix} I_{\mathcal{N}_{Q,k}} & (M^\kappa)^{-1}(R^{\gamma\kappa})^T B^\gamma & (M^\kappa)^{-1}(R^{\sigma\kappa})^T B^\sigma & (M^\kappa)^{-1}(R^{\eta\kappa})^T B^\eta \end{bmatrix}$$

help decouple the hybridized vector. We can also write down the formulation solely in terms of the original SBP operators:

$$\begin{aligned} &\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt} + \underbrace{2\sum_{m=1}^{d}\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\Big(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\Big)\overrightarrow{\mathbf{1}^\kappa}}_{\text{Flux differencing term}} \\ &= \underbrace{(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}\Big(\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa})\overrightarrow{\mathbf{1}^\kappa} - (\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}}_{\text{Simultaneous approximation term}} \\ &\quad + \underbrace{(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\Big(\mathbf{R}^{\gamma\kappa}\circ\mathbf{F}_{\mathbf{n},S}\Big(\overrightarrow{\widetilde{\mathbf{u}}^{\gamma\kappa}},\overrightarrow{\mathbf{u}^\kappa}\Big)\Big)\overrightarrow{\mathbf{1}^\kappa} - ((\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma)\circ\mathbf{F}_{\mathbf{n},S}\Big(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\widetilde{\mathbf{u}}^{\gamma\kappa}}\Big)\overrightarrow{\mathbf{1}^\gamma}\Big)}_{\text{Skew-symmetric correction term}}, \end{aligned} \tag{5.4}$$

where $\mathbf{F}_{\mathbf{n},S}(\cdot,\cdot)$ is the matrix of pairwise combinations of $\mathbf{f}_{\mathbf{n},S}(\cdot,\cdot)$. Comparing it with the unmodified DG method (3.19), we have not only applied the flux differencing technique, but also tuned the SAT by adding some skew-symmetric correction term. The

component-wise representation is

$$\begin{aligned}
&\frac{d\mathbf{u}_j^\kappa}{dt}+2\sum_{m=1}^{d}\sum_{l=1}^{\mathcal{N}_{Q,k}} D_{m,jl}^{\kappa}\mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)\\
&=\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}} R_{sj}^{\gamma\kappa}\frac{\tau_s^\gamma}{\omega_j^\kappa}\Big(\sum_{l=1}^{\mathcal{N}_{Q,k}} R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)\\
&\quad-\widehat{\mathbf{f}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})+\sum_{l=1}^{\mathcal{N}_{Q,k}} R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\mathbf{u}_l^\kappa)-\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\widetilde{\mathbf{u}}_s^{\gamma\kappa})\Big).
\end{aligned}\tag{5.5}$$

We proceed to prove the consistency, primary conservation, and entropy stability of (5.3). Similar to Lemma 4.1, we have the following result for the hybridized flux differencing term.

Lemma 5.1. *If for each $1\le m\le d$, $\mathbf{f}_{m,S}$ is an entropy conservative flux with respect to some entropy function U, then*

$$\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{L}^{\kappa,h}\mathbf{M}^\kappa\left(2\mathbf{D}_m^{\kappa,h}\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^{\kappa,h}},\overrightarrow{\mathbf{u}^{\kappa,h}}\right)\overrightarrow{\mathbf{1}^{\kappa,h}}\right)=\sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa}\left(\overrightarrow{\mathbf{1}^\gamma}\right)^T\mathbf{B}^\gamma\overrightarrow{\widetilde{\mathbf{f}}_m^{\gamma\kappa}},\tag{5.6}$$

$$\begin{aligned}
&\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{L}^{\kappa,h}\mathbf{M}^\kappa\left(2\mathbf{D}_m^{\kappa,h}\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^{\kappa,h}},\overrightarrow{\mathbf{u}^{\kappa,h}}\right)\overrightarrow{\mathbf{1}^{\kappa,h}}\right)\\
&\quad=\sum_{\gamma\in\partial T_\kappa} n_m^{\gamma\kappa}\left(\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T\mathbf{B}^\gamma\overrightarrow{\widetilde{\mathbf{f}}_m^{\gamma\kappa}}-\left(\overrightarrow{\widetilde{\psi}_m^{\gamma\kappa}}\right)^T B^\gamma\overrightarrow{1^\gamma}\right).
\end{aligned}\tag{5.7}$$

Proof. By the definition of $\mathbf{L}^{\kappa,h}$,

$$\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{L}^{\kappa,h}\mathbf{M}^\kappa=\left[\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T\mathbf{M}^\kappa\quad\left(\overrightarrow{\mathbf{1}^\gamma}\right)^T\mathbf{B}^\gamma\quad\left(\overrightarrow{\mathbf{1}^\sigma}\right)^T\mathbf{B}^\sigma\quad\left(\overrightarrow{\mathbf{1}^\eta}\right)^T\mathbf{B}^\eta\right]=\left(\overrightarrow{\mathbf{1}^{\kappa,h}}\right)^T\mathbf{M}^{\kappa,h},$$

$$\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{L}^{\kappa,h}\mathbf{M}^\kappa=\left[\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T\mathbf{M}^\kappa\quad\left(\overrightarrow{\mathbf{v}^\gamma}\right)^T\mathbf{B}^\gamma\quad\left(\overrightarrow{\mathbf{v}^\sigma}\right)^T\mathbf{B}^\sigma\quad\left(\overrightarrow{\mathbf{v}^\eta}\right)^T\mathbf{B}^\eta\right]=\left(\overrightarrow{\mathbf{v}^{\kappa,h}}\right)^T\mathbf{M}^{\kappa,h}.$$

The rest of proof is the same as Lemma 4.1. We make use of the SBP property of hybridized operators, and the identity $D_m^{\kappa,h}\overrightarrow{1^{\kappa,h}}=\overrightarrow{0^{\kappa,h}}$. □

Theorem 5.2. *Under the same assumptions as in Theorem* 4.1 *(except for the collocated surface nodes assumption), the scheme* (5.3) *is consistent in that for a smooth solution* $\mathbf{u}$ *of* (2.1)*, the local truncation error*

$$\begin{aligned}
&\frac{d\mathbf{u}_j^\kappa}{dt}+2\sum_{m=1}^{d}\sum_{l=1}^{\mathcal{N}_{Q,k}} D_{m,jl}^{\kappa}\mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)-\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}} R_{sj}^{\gamma\kappa}\frac{\tau_s^\gamma}{\omega_j^\kappa}\Big(\sum_{l=1}^{\mathcal{N}_{Q,k}} R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)\\
&\quad-\widehat{\mathbf{f}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})+\sum_{l=1}^{\mathcal{N}_{Q,k}} R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\mathbf{u}_l^\kappa)-\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\widetilde{\mathbf{u}}_s^{\gamma\kappa})\Big)=\mathcal{O}(h^k),
\end{aligned}\tag{5.8}$$

as well as conservative and entropy stable with respect to U in that

$$\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{1}^{\kappa}}\Big)^{T}\mathbf{M}^{\kappa}\overrightarrow{\mathbf{u}^{\kappa}}\Big)=0,\quad \frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^{\kappa}}\Big)^{T}M^{\kappa}\overrightarrow{U^{\kappa}}\Big)\leq 0. \tag{5.9}$$

Proof. For consistency, we already know that the truncation error of the flux differencing term is of high order. As a result of shape regular and quasi-uniform mesh, the quadrature weights have the scales $\omega_j^{\kappa}=\Theta(h^d)$ and $\tau_s^{\gamma}=\Theta(h^{d-1})$, and since the extrapolation matrices are invariant under affine mapping, the coefficients have the scale $R_{sj}^{\gamma\kappa}=\Theta(1)$. Hence it suffices to show that

$$\sum_{l=1}^{\mathcal{N}_{Q,k}}R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^{\kappa},\mathbf{u}_l^{\kappa})-\widehat{\mathbf{f}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})+\sum_{l=1}^{\mathcal{N}_{Q,k}}R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\mathbf{u}_l^{\kappa})-\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^{\kappa},\widetilde{\mathbf{u}}_s^{\gamma\kappa})=\mathcal{O}(h^{k+1}).$$

By the approximation property of extrapolation and Lipschitz continuity of $\mathbf{u}(\mathbf{v})$,

$$\mathbf{v}_s^{\gamma\kappa}-\mathbf{v}(\mathbf{x}_s^{\gamma})=\mathcal{O}(h^{k+1}),\quad \widetilde{\mathbf{u}}_s^{\gamma\kappa}-\mathbf{u}(\mathbf{x}_s^{\gamma})=\mathcal{O}(h^{k+1}).$$

We check each term separately:

$$\begin{aligned}
&\sum_{l=1}^{\mathcal{N}_{Q,k}}R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^{\kappa},\mathbf{u}_l^{\kappa})=\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^{\kappa},\mathbf{u}(\mathbf{x}_s^{\gamma}))+\mathcal{O}(h^{k+1}),\\
&\widehat{\mathbf{f}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})=\mathbf{f}_{\mathbf{n}}(\mathbf{u}(\mathbf{x}_s^{\gamma}))+\mathcal{O}(h^{k+1}),\\
&\sum_{l=1}^{\mathcal{N}_{Q,k}}R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\mathbf{u}_l^{\kappa})=\mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\mathbf{u}(\mathbf{x}_s^{\gamma}))+\mathcal{O}(h^{k+1})=\mathbf{f}_{\mathbf{n}}(\mathbf{u}(\mathbf{x}_s^{\gamma}))+\mathcal{O}(h^{k+1}),\\
&\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^{\kappa},\widetilde{\mathbf{u}}_s^{\gamma\kappa})=\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^{\kappa},\mathbf{u}(\mathbf{x}_s^{\gamma}))+\mathcal{O}(h^{k+1}).
\end{aligned}$$

Then the truncation error of boundary terms is also of high order. The proof of conservation and entropy stability is straightforward. We insert (5.6) and (5.7) and get

$$\begin{aligned}
&\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{1}^{\kappa}}\Big)^{T}\mathbf{M}^{\kappa}\overrightarrow{\mathbf{u}^{\kappa}}\Big)=-\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^{\gamma}(\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})+\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\nu}}(\widetilde{\mathbf{u}}_s^{\gamma\nu},\widetilde{\mathbf{u}}_s^{\gamma\kappa}))=0,\\
&\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{v}^{\kappa}}\Big)^{T}\mathbf{M}^{\kappa}\overrightarrow{\mathbf{u}^{\kappa}}\Big)=\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^{\gamma}((\mathbf{v}_s^{\gamma\nu}-\mathbf{v}_s^{\gamma\kappa})^{T}\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})-(\widetilde{\psi}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\kappa}-\widetilde{\psi}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\nu}))\leq 0.
\end{aligned}$$

Now the last inequality is valid for entropy-extrapolated values. □

Remark 5.1. In fact, scheme (4.1) is a special case of (5.3) (and (5.4)). If we assume collocated surface nodes, since $R^{\gamma\kappa}$ is a simple restriction, $\overrightarrow{\widetilde{\mathbf{u}}^{\gamma\kappa}}=\overrightarrow{\mathbf{u}^{\gamma\kappa}}$, and

$$\begin{aligned}
&\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^{\kappa}},\overrightarrow{\mathbf{u}^{\kappa}})\overrightarrow{\mathbf{1}^{\kappa}}=\mathbf{E}^{\gamma\kappa}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\kappa}}=(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}},\\
&(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\Big(\mathbf{R}^{\gamma\kappa}\circ\mathbf{F}_{\mathbf{n},S}\Big(\overrightarrow{\widetilde{\mathbf{u}}^{\gamma\kappa}},\overrightarrow{\mathbf{u}^{\kappa}}\Big)\Big)\overrightarrow{\mathbf{1}^{\kappa}}=(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}},\\
&((\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma})\circ\mathbf{F}_{\mathbf{n},S}\Big(\overrightarrow{\mathbf{u}^{\kappa}},\overrightarrow{\widetilde{\mathbf{u}}^{\gamma\kappa}}\Big)\overrightarrow{\mathbf{1}^{\gamma}}=(\mathbf{R}^{\gamma\kappa})^{T}\mathbf{B}^{\gamma}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}}.
\end{aligned}$$

Therefore the skew-symmetric correction term in (5.4) vanishes, and we recover (4.1).

5.2 Approach 2: global SBP operators

The global SBP operators approach was found by Crean et al. in [19,20]. We consider the nodal values on different elements as a whole, grouping them into a single global vector:

$$\overrightarrow{\mathbf{u}^g}=\begin{bmatrix}\overrightarrow{\mathbf{u}^1}\\ \vdots \\ \overrightarrow{\mathbf{u}^K}\end{bmatrix},\quad \overrightarrow{\mathbf{v}^g}=\begin{bmatrix}\overrightarrow{\mathbf{v}^1}\\ \vdots \\ \overrightarrow{\mathbf{v}^K}\end{bmatrix}.$$

The global mass matrix is

$$M^g=\begin{bmatrix}M^1 & & \\ & \ddots & \\ & & M^K\end{bmatrix},$$

and the global difference matrices are assembled as

$$D_m^g=\begin{bmatrix}D_m^{g,11} & \cdots & D_m^{g,1K}\\ \vdots & \ddots & \vdots \\ D_m^{g,K1} & \cdots & D_m^{g,KK}\end{bmatrix},$$

where we set

$$D_m^{g,\kappa\nu}=\begin{cases}D_m^\kappa-\frac{1}{2}(M^\kappa)^{-1}\sum\limits_{\gamma\in\partial T_\kappa}n_m^\gamma E^{\gamma\kappa} & \text{if } \kappa=\nu,\\ 0 & \text{if } \kappa\neq\nu \text{ and } \partial T_\kappa\cap\partial T_\nu=\emptyset,\\ \frac{1}{2}n_m^{\gamma\kappa}(M^\kappa)^{-1}(R^{\gamma\kappa})^T B^\gamma R^{\gamma\nu}:=\frac{1}{2}n_m^{\gamma\kappa}(M^\kappa)^{-1}E^{\kappa\nu} & \text{if } \kappa\neq\nu \text{ and } \partial T_\kappa\cap\partial T_\nu=\gamma.\end{cases}$$

Theorem 5.3. *The global mass matrix and global difference matrices satisfy the following conditions:*

1. *Exactness:*

$$D_m^g V^g=V^g\widehat{D}_m,\quad \text{where } V^g=\begin{bmatrix}V^1\\ \vdots \\ V^K\end{bmatrix}.\quad \text{In particular, } D_m^g\overrightarrow{1^g}=\overrightarrow{0^g}. \tag{5.10}$$

2. *Summation-by-parts:*

$$M^gD_m^g+(D_m^g)^TM^g=0. \tag{5.11}$$

The proof is again a direct application of the local exactness property and the local SBP property. Here the right hand side of (5.11) is zero, as on each interface, the contributions from its two sides will cancel out with each other (and by periodic boundary condition, all faces are interfaces). We produce the following nodal DG method, using global SBP operators:

$$\frac{d\overrightarrow{\mathbf{u}^g}}{dt}+\underbrace{2\sum_{m=1}^{d}\mathbf{D}_m^g\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^g},\overrightarrow{\mathbf{u}^g}\right)\overrightarrow{\mathbf{1}^g}}_{\text{Global flux differencing term}}=0. \tag{5.12}$$

The SATs on element interfaces are implied by the global flux differencing term. Plugging the definition of D_m^g, we derive its element-wise formulation

$$\begin{aligned}&\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt}+\underbrace{2\sum_{m=1}^{d}\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}}_{\text{Flux differencing term}}\\&=\underbrace{(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}\left(\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa})\overrightarrow{\mathbf{1}^\kappa}-\mathbf{E}^{\kappa\nu}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\nu})\overrightarrow{\mathbf{1}^\nu}\right)}_{\text{Simultaneous approximation term}},\end{aligned} \tag{5.13}$$

and component-wise formulation

$$\begin{aligned}&\frac{d\mathbf{u}_j^\kappa}{dt}+2\sum_{m=1}^{d}\sum_{l=1}^{\mathcal{N}_{Q,k}}D_{m,jl}^\kappa\mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)\\&=\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}R_{sj}^{\gamma\kappa}\frac{\tau_s^\gamma}{\omega_j^\kappa}\Big(\sum_{l=1}^{\mathcal{N}_{Q,k}}(R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)-R_{sl}^{\gamma\nu}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\nu))\Big).\end{aligned} \tag{5.14}$$

Theorem 5.4. *Under the same assumptions as in Theorem* 5.2*, the scheme* (5.12) *is consistent in that for a smooth solution* $\mathbf{u}$ *of* (2.1)*,*

$$\begin{aligned}&\frac{d\mathbf{u}_j^\kappa}{dt}+2\sum_{m=1}^{d}\sum_{l=1}^{\mathcal{N}_{Q,k}}D_{m,jl}^\kappa\mathbf{f}_{m,S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)\\&\quad-\sum_{\gamma\in\partial T_\kappa}\sum_{s=1}^{\mathcal{N}_{B,k}}R_{sj}^{\gamma\kappa}\frac{\tau_s^\gamma}{\omega_j^\kappa}\Big(\sum_{l=1}^{\mathcal{N}_{Q,k}}(R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)-R_{sl}^{\gamma\nu}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\nu))\Big)=\mathcal{O}(h^k),\end{aligned} \tag{5.15}$$

as well as conservative and entropy conservative with respect to U *in that*

$$\frac{d}{dt}\left(\left(\overrightarrow{\mathbf{1}^g}\right)^T\mathbf{M}^g\overrightarrow{\mathbf{u}^g}\right)=-\left(\overrightarrow{\mathbf{1}^g}\right)^T\mathbf{M}^g\left(2\mathbf{D}_m^g\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^g},\overrightarrow{\mathbf{u}^g}\right)\overrightarrow{\mathbf{1}^g}\right)=0, \tag{5.16}$$

$$\frac{d}{dt}\left(\left(\overrightarrow{1^g}\right)^T M^g\overrightarrow{U^g}\right)=-\left(\overrightarrow{\mathbf{v}^g}\right)^T\mathbf{M}^g\left(2\mathbf{D}_m^g\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^g},\overrightarrow{\mathbf{u}^g}\right)\overrightarrow{\mathbf{1}^g}\right)=0. \tag{5.17}$$

Proof. Conservation and entropy conservation are actually global versions of (4.6) and (4.7), and can be proved in the same way as in Lemma 4.1. For consistency, since

$$\sum_{l=1}^{\mathcal{N}_{Q,k}} R_{sl}^{\gamma\kappa}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\kappa)=\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}(\mathbf{x}_s^\gamma))+\mathcal{O}(h^{k+1}),$$

and

$$\sum_{l=1}^{\mathcal{N}_{Q,k}} R_{sl}^{\gamma\nu}\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}_l^\nu)=\mathbf{f}_{\mathbf{n},S}(\mathbf{u}_j^\kappa,\mathbf{u}(\mathbf{x}_s^\gamma))+\mathcal{O}(h^{k+1}),$$

the truncation error is of high order. □

In practice, the entropy conservative scheme (5.12) will generate strong spurious oscillations in the vicinity of shock waves as entropy should be dissipated at discontinuities. It is necessary to impose entropy dissipation on element interfaces to make it entropy stable,. For $1\le\kappa\le K$ and $\gamma\in\partial T_\kappa$, we define

$$\overrightarrow{\mathbf{d}_\mathbf{n}^{\gamma\kappa,*}}=\begin{bmatrix}\widehat{\mathbf{d}}_\mathbf{n}(\widetilde{\mathbf{u}}_1^{\gamma\kappa},\widetilde{\mathbf{u}}_1^{\gamma\nu})\\ \vdots\\ \widehat{\mathbf{d}}_\mathbf{n}(\widetilde{\mathbf{u}}_{\mathcal{N}_{B,k}}^{\gamma\kappa},\widetilde{\mathbf{u}}_{\mathcal{N}_{B,k}}^{\gamma\nu})\end{bmatrix},$$

where the $\mathbf{d}_\mathbf{n}$ is some entropy dissipation function with respect to U (see Section 2.2). We create the entropy stable scheme

$$\begin{aligned}&\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt}+\underbrace{2\sum_{m=1}^{d}\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa}\right)\overrightarrow{\mathbf{1}^\kappa}}_{\text{Flux differencing term}}\\ &=\underbrace{(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}\left(\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\kappa})\overrightarrow{\mathbf{1}^\kappa}-\mathbf{E}^{\kappa\nu}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^\kappa},\overrightarrow{\mathbf{u}^\nu})\overrightarrow{\mathbf{1}^\nu}-(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\overrightarrow{\mathbf{d}_\mathbf{n}^{\gamma\kappa,*}}\right)}_{\text{Simultaneous approximation term}}.\end{aligned}\tag{5.18}$$

Corollary 5.1. *If $\widehat{\mathbf{d}}_\mathbf{n}$ is an entropy dissipation function with respect to U, then the scheme* (5.18) *is consistent, conservative and entropy stable with respect to U.*

Proof. For a smooth solution $\mathbf{u}$, because of consistency of $\widehat{\mathbf{d}}_\mathbf{n}$,

$$\widehat{\mathbf{d}}_\mathbf{n}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})=\widehat{\mathbf{d}}_\mathbf{n}(\mathbf{u}(\mathbf{x}_s^\gamma),\mathbf{u}(\mathbf{x}_s^\gamma))+\mathcal{O}(h^{k+1})=\mathcal{O}(h^{k+1}).$$

Hence entropy dissipation does not affect consistency. The effects on primary conservation and entropy stability are

$$-\sum_{\kappa=1}^{K}\sum_{\gamma\in\partial T_\kappa}\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\overrightarrow{\mathbf{d}_\mathbf{n}^{\gamma\kappa,*}}=-\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma(\widehat{\mathbf{d}}_{\mathbf{n}^{\gamma\kappa}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})+\widehat{\mathbf{d}}_{\mathbf{n}^{\gamma\nu}}(\widetilde{\mathbf{u}}_s^{\gamma\nu},\widetilde{\mathbf{u}}_s^{\gamma\kappa}))=0,$$

$$-\sum_{\kappa=1}^{K}\sum_{\gamma\in\partial T_\kappa}\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\overrightarrow{\mathbf{d}_\mathbf{n}^{\gamma\kappa,*}}=\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma(\mathbf{v}_s^{\gamma\nu}-\mathbf{v}_s^{\gamma\kappa})^T\widehat{\mathbf{d}}_{\mathbf{n}^{\gamma\kappa}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})\le 0.$$

As a consequence, the scheme is clearly conservative and entropy stable. □

Remark 5.2. In the case of collocated surface nodes, the boundary terms in (5.18) are

$$\mathbf{E}^{\gamma\kappa}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^k},\overrightarrow{\mathbf{u}^k})\overrightarrow{\mathbf{1}^k}=(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma k}},\quad \mathbf{E}^{\kappa\nu}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^k},\overrightarrow{\mathbf{u}^\nu})\overrightarrow{\mathbf{1}^\nu}=(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\overrightarrow{\mathbf{f}_{\mathbf{n},S}^{\gamma\kappa,*}},$$

where

$$\overrightarrow{\mathbf{f}_{\mathbf{n},S}^{\gamma\kappa,*}}=\begin{bmatrix}\mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_1^{\gamma\kappa},\widetilde{\mathbf{u}}_1^{\gamma\nu})\\ \vdots\\ \mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_{\mathcal{N}_{B,k}}^{\gamma\kappa},\widetilde{\mathbf{u}}_{\mathcal{N}_{B,k}}^{\gamma\nu})\end{bmatrix}$$

is the vector of entropy conservative fluxes on the interface. Then the scheme (5.18) reduces to

$$\frac{d\overrightarrow{\mathbf{u}^k}}{dt}+2\sum_{m=1}^{d}\mathbf{D}_m^\kappa\circ\mathbf{F}_{m,S}\left(\overrightarrow{\mathbf{u}^k},\overrightarrow{\mathbf{u}^k}\right)\overrightarrow{\mathbf{1}^k}=(\mathbf{M}^\kappa)^{-1}\sum_{\gamma\in\partial T_\kappa}(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\left(\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma k}}-\overrightarrow{\mathbf{f}_{\mathbf{n},S}^{\gamma\kappa,*}}-\overrightarrow{\mathbf{d}_{\mathbf{n}}^{\gamma\kappa,*}}\right).$$

We again recover (4.1), by setting $\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}=\overrightarrow{\mathbf{f}_{\mathbf{n},S}^{\gamma\kappa,*}}+\overrightarrow{\mathbf{d}_{\mathbf{n}}^{\gamma\kappa,*}}$, i.e.,

$$\widehat{\mathbf{f}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})=\mathbf{f}_{\mathbf{n},S}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})+\widehat{\mathbf{d}}_{\mathbf{n}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu}).$$

Remark 5.3. The element coupling term $\mathbf{E}^{\kappa\nu}\circ\mathbf{F}_{\mathbf{n},S}(\overrightarrow{\mathbf{u}^k},\overrightarrow{\mathbf{u}^\nu})\overrightarrow{\mathbf{1}^\nu}$ depends on all nodal values on the neighboring element T_ν. This will harm the locality of the DG formulation, and make the implementation of non-periodic boundary conditions (inflow, outflow, solid wall, etc.) more difficult.

5.3 Approach 3: directly enforcing entropy balance

The method in [1] was written in the more general residual distribution framework. We will focus on the version for nodal DG formulations. We start with the unmodified nodal DG method (3.19). For $1\le\kappa\le K$, the local entropy error of (3.19) on T_κ is defined as

$$\mathcal{E}^\kappa=\sum_{m=1}^{d}\left(\overrightarrow{\mathbf{v}^k}\right)^T(\mathbf{D}_m^\kappa)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^k}-\sum_{\gamma\in\partial T_\kappa}\left(\overrightarrow{1^\gamma}\right)^T B^\gamma\overrightarrow{\widetilde{\psi}_{\mathbf{n}}^{\gamma k}}.\tag{5.19}$$

Due to the lack of discrete chain rule, the entropy error is nonzero. However, we are able to demonstrate that for smooth solutions, $\mathcal{E}^\kappa$ is of high order.

Theorem 5.5. *Under the same assumptions as in Theorem* 5.2, *for a smooth solution* $\mathbf{u}$ *of* (2.1), *the local entropy error* $\mathcal{E}^\kappa=\mathcal{O}(h^{k+d})$.

Proof. For a smooth solution $\mathbf{u}$, the following identity holds at the continuous level:

$$\sum_{m=1}^{d}\Big(\mathbf{f}_m(\mathbf{u}),\frac{\partial \mathbf{v}(\mathbf{u})}{\partial x_m}\Big)_{T_\kappa}=\sum_{m=1}^{d}\Big(\frac{\partial \psi_m(\mathbf{v}(\mathbf{u}))}{\partial x_m},1\Big)_{T_\kappa}=\sum_{\gamma\in\partial T_\kappa}\langle\psi_{\mathbf{n}}(\mathbf{v}(\mathbf{u})),1\rangle_\gamma.$$

By the approximation property of difference and extrapolation matrix, and the algebraic accuracy of volume and surface quadrature rule,

$$\begin{aligned}
&\sum_{m=1}^{d}\Big(\mathbf{f}_m(\mathbf{u}),\frac{\partial \mathbf{v}(\mathbf{u})}{\partial x_m}\Big)_{T_\kappa}=\sum_{m=1}^{d}\Big(\mathbf{f}_m(\mathbf{u}),\frac{\partial \mathbf{v}(\mathbf{u})}{\partial x_m}\Big)_{T_\kappa,\omega}+\mathcal{O}(h^{2k+d})\\
=&\sum_{m=1}^{d}\Big(\overrightarrow{\partial_{x_m}\mathbf{v}^\kappa}\Big)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^\kappa}+\mathcal{O}(h^{2k+d})=\sum_{m=1}^{d}\Big(\mathbf{D}_m^\kappa\overrightarrow{\mathbf{v}^\kappa}\Big)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^\kappa}+\mathcal{O}(h^{k+d}),\\
&\sum_{\gamma\in\partial T_\kappa}\langle\psi_{\mathbf{n}}(\mathbf{v}(\mathbf{u})),1\rangle_\gamma=\sum_{\gamma\in\partial T_\kappa}\langle\psi_{\mathbf{n}}(\mathbf{v}(\mathbf{u})),1\rangle_{\gamma,\tau}+\mathcal{O}(h^{2k+d})\\
=&\sum_{\gamma\in\partial T_\kappa}\Big(\overrightarrow{1^\gamma}\Big)^TB^\gamma\overrightarrow{\psi_{\mathbf{n}}^\gamma}+\mathcal{O}(h^{2k+d})=\sum_{\gamma\in\partial T_\kappa}\Big(\overrightarrow{1^\gamma}\Big)^TB^\gamma\overrightarrow{\widetilde{\psi}_{\mathbf{n}}^{\gamma\kappa}}+\mathcal{O}(h^{k+d}).
\end{aligned}$$

Hence

$$\mathcal{E}^\kappa=\sum_{m=1}^{d}\Big(\overrightarrow{\mathbf{v}^\kappa}\Big)^T(\mathbf{D}_m^\kappa)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{f}_m^\kappa}-\sum_{\gamma\in\partial T_\kappa}\Big(\overrightarrow{1^\gamma}\Big)^TB^\gamma\overrightarrow{\widetilde{\psi}_{\mathbf{n}}^{\gamma\kappa}}=\mathcal{O}(h^{k+d}).$$

The proof is completed. □

In order to neutralize the entropy error, a simple linear correction term will be introduced to (3.19), resulting in the scheme

$$\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt}+\underbrace{\sum_{m=1}^{d}\mathbf{D}_m^\kappa\overrightarrow{\mathbf{f}_m^\kappa}}_{\text{Difference term}}=\underbrace{(\mathbf{M}^\kappa)^{-1}\Big(\sum_{\gamma\in\partial T_\kappa}(\mathbf{R}^{\gamma\kappa})^T\mathbf{B}^\gamma\Big(\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa}}-\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\Big)}_{\text{Simultaneous approximation term}}-\underbrace{\frac{\mathcal{E}^\kappa}{\Big(\overrightarrow{\mathbf{v}^{\kappa,b}}\Big)^T\overrightarrow{\mathbf{v}^{\kappa,b}}}\overrightarrow{\mathbf{v}^{\kappa,b}}}_{\text{Linear correction term}}\Big),\tag{5.20}$$

where $\overrightarrow{\mathbf{v}^{\kappa,b}}$ is the vector of normalized nodal values of $\mathbf{v}$:

$$\overrightarrow{\mathbf{v}^{\kappa,b}}=\begin{bmatrix}\mathbf{v}_0^\kappa-\overline{\mathbf{v}}^\kappa\\ \vdots\\ \mathbf{v}_{\mathcal{N}_{Q,k}}^\kappa-\overline{\mathbf{v}}^\kappa\end{bmatrix},\quad \overline{\mathbf{v}}^\kappa=\frac{1}{\mathcal{N}_{Q,k}}\sum_{j=1}^{\mathcal{N}_{Q,k}}\mathbf{v}_j^\kappa.$$

Theorem 5.6. *Under the same assumptions as in Theorem* 5.2, *the scheme* (5.20) *is conservative and entropy stable.*

Proof. From the definition of $\overrightarrow{\mathbf{v}^{\kappa,o}}$, $\left(\overrightarrow{\mathbf{1}^{\kappa}}\right)^T\overrightarrow{\mathbf{v}^{\kappa,o}}=0$ and

$$\left(\overrightarrow{\mathbf{v}^{\kappa}}\right)^T\overrightarrow{\mathbf{v}^{\kappa,o}}=\sum_{j=1}^{\mathcal{N}_{Q,k}}(\mathbf{v}_j^{\kappa})^T(\mathbf{v}_j^{\kappa}-\overline{\mathbf{v}}^{\kappa})=\sum_{j=1}^{\mathcal{N}_{Q,k}}(\mathbf{v}_j^{\kappa}-\overline{\mathbf{v}}^{\kappa})^T(\mathbf{v}_j^{\kappa}-\overline{\mathbf{v}}^{\kappa})=\left(\overrightarrow{\mathbf{v}^{\kappa,o}}\right)^T\overrightarrow{\mathbf{v}^{\kappa,o}}.$$

Then conservation and entropy stability can be easily proved:

$$\begin{aligned}
\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\left(\overrightarrow{\mathbf{1}^{\kappa}}\right)^T\mathbf{M}^{\kappa}\overrightarrow{\mathbf{u}^{\kappa}}\Big)&=-\sum_{\kappa=1}^{K}\sum_{\gamma\in\partial T_{\kappa}}\overrightarrow{\mathbf{1}^{\gamma}}\mathbf{B}^{\gamma}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}\\
&=-\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^{\gamma}(\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})+\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\nu}}(\widetilde{\mathbf{u}}_s^{\gamma\nu},\widetilde{\mathbf{u}}_s^{\gamma\kappa}))=0,\\
\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\left(\overrightarrow{1^{\kappa}}\right)^TM^{\kappa}\overrightarrow{U^{\kappa}}\Big)&=\sum_{\kappa=1}^{K}\Big(\sum_{\gamma\in\partial T_{\kappa}}\left(\overrightarrow{1^{\gamma}}\right)^TB^{\gamma}\overrightarrow{\widetilde{\psi}_{\mathbf{n}}^{\gamma\kappa}}+\mathcal{E}^{\kappa}-\sum_{\gamma\in\partial T_{\kappa}}\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T\mathbf{B}^{\gamma}\overrightarrow{\mathbf{f}_{\mathbf{n}}^{\gamma\kappa,*}}-\mathcal{E}^{\kappa}\Big)\\
&=\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^{\gamma}((\mathbf{v}_s^{\gamma\nu}-\mathbf{v}_s^{\gamma\kappa})^T\widehat{\mathbf{f}}_{\mathbf{n}^{\gamma\kappa}}(\widetilde{\mathbf{u}}_s^{\gamma\kappa},\widetilde{\mathbf{u}}_s^{\gamma\nu})-(\widetilde{\psi}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\kappa}-\widetilde{\psi}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\nu}))\leq 0.
\end{aligned}$$

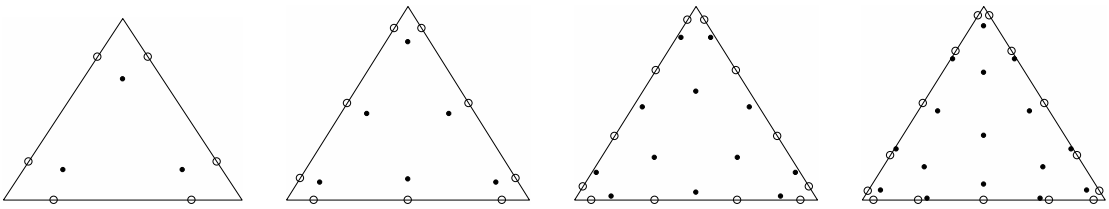
The proof is completed. □

Remark 5.4. $\overline{\mathbf{v}}^{\kappa}$ does not have to be the arithmetic mean. The extension to a weighted average of nodal values is straightforward.

Remark 5.5. Although we have proved that $\mathcal{E}^{\kappa}$ is of high order, this does not guarantee the consistency of (5.20). The main reason is that $\overrightarrow{\mathbf{v}^{\kappa,o}}=\mathcal{O}(h)$, and we are not able to control the truncation error of linear correction term, which is of the order $\mathcal{O}(h^{k+d})/\mathcal{O}(h^2)$ with the coefficients of the two $\mathcal{O}$ terms in the denominator and in the numerator not necessarily related, hence there is the danger of the coefficient in the denominator going to zero faster than that of the numerator, which might lead to a degeneracy of accuracy.

5.4 Accuracy test

We test the numerical convergence rates of three entropy stable DG methods in this section, for the two-dimensional Burgers equation associated with hyperbolic cosine entropy function. The settings are the same as in Section 4.2. We use the local Lax-Friedrichs flux in (5.3) and (5.20), and the local Lax-Friedrichs entropy dissipation function in the implementation of (5.18). SBP operators are built on degree $2k$ volume quadrature rules on triangles, exhibited in Fig. 2. Compared to Fig. 1, these quadrature rules have better algebraic accuracy with fewer degrees of freedom. This is a major benefit of removing the collocated surface nodes constraint. The extrapolation matrices and difference matrices are simply chosen as $R^{\gamma\kappa}=V^{\gamma}P^{\kappa}$ and $D_m^{\kappa}=V^{\kappa}\widehat{D}_mP^{\kappa}$.



(a) $k=1, \mathcal{N}_{Q,k}=3$ (b) $k=2, \mathcal{N}_{Q,k}=6$ (c) $k=3, \mathcal{N}_{Q,k}=12$ (d) $k=4, \mathcal{N}_{Q,k}=16$

Figure 2: Degree $2k$ quadrature rules triangles for $k=1,2,3,4$. Dots are quadrature points for the triangle, and circles are quadrature points for the edges.

Table 2: Errors and orders of convergence of (5.3) for the two-dimensional Burgers equation at $t=0.1$. Degree $2k$ volume quadrature rules are used.

k	h	L^1 error	order	L^2 error	order	L^∞ error	order
2	1/16	2.244e-04	-	4.597e-04	-	6.226e-03	-
	1/32	3.391e-05	2.726	8.188e-05	2.489	1.367e-03	2.187
	1/64	4.386e-06	2.951	1.074e-05	2.931	1.793e-04	2.931
	1/128	5.717e-07	2.939	1.557e-06	2.786	2.992e-05	2.583
	1/256	7.511e-08	2.928	2.238e-07	2.798	5.511e-06	2.440
3	1/16	5.445e-05	-	1.913e-04	-	3.067e-03	-
	1/32	4.526e-06	3.589	1.921e-05	3.316	5.347e-04	2.520
	1/64	3.019e-07	3.906	1.365e-06	3.815	3.096e-05	4.110
	1/128	1.920e-08	3.975	9.184e-08	3.894	3.538e-06	3.130
	1/256	1.275e-09	3.913	6.148e-09	3.901	2.941e-07	3.589
4	1/8	1.476e-04	-	4.757e-04	-	6.202e-03	-
	1/16	1.092e-05	3.757	4.482e-05	3.408	1.132e-03	2.454
	1/32	4.984e-07	4.454	2.637e-06	4.087	1.150e-04	3.300
	1/64	1.528e-08	5.028	7.728e-08	5.093	5.355e-06	4.424
	1/128	4.818e-10	4.987	2.472e-09	4.966	1.641e-07	5.028

Numerical results at $t=0.1$ are presented in Table 2 for scheme (5.3), Table 3 for scheme (5.18), and Table 4 for scheme (5.20). Since the volume quadrature rule is of degree $2k$, the accuracy requirements in [13, 55] are met, and there is some hope to recover optimal $(k+1)$-th order convergence. We do see optimal convergence in Table 2 and Table 4, despite the fact that the truncation error of (5.20) is not fully understood. However, the convergence is still below optimal for the global SBP scheme (5.18) in Table 3.

6 Additional topics

In this section, we treat different entropy stable nodal DG discretizations (including (4.1), (5.3), (5.12) and (5.20)) in the same manner, using the generic representation

Table 3: Errors and orders of convergence of (5.18) for the two-dimensional Burgers equation at $t=0.1$. Degree $2k$ volume quadrature rules are used.

k	h	L^1 error	order	L^2 error	order	L^∞ error	order
2	1/16	5.300e-04	-	1.367e-03	-	1.451e-02	-
	1/32	7.057e-05	2.909	2.154e-04	2.666	2.813e-03	2.366
	1/64	8.973e-06	2.975	3.028e-05	2.831	4.761e-04	2.563
	1/128	1.089e-06	3.042	3.872e-06	2.967	8.605e-05	2.468
	1/256	1.363e-07	2.998	5.082e-07	2.930	1.447e-05	2.572
3	1/16	8.324e-05	-	2.611e-04	-	4.398e-03	-
	1/32	8.824e-06	3.238	3.621e-05	2.850	8.565e-04	2.361
	1/64	7.546e-07	3.548	3.655e-06	3.308	1.463e-04	2.550
	1/128	5.944e-08	3.666	3.537e-07	3.370	2.444e-05	2.581
	1/256	4.908e-09	3.598	3.530e-08	3.325	3.024e-06	3.015
4	1/8	1.945e-04	-	5.587e-04	-	6.280e-03	-
	1/16	1.768e-05	3.459	6.770e-05	3.045	1.607e-03	1.966
	1/32	1.123e-06	3.977	5.577e-06	3.602	2.089e-04	2.943
	1/64	4.482e-08	4.648	2.583e-07	4.432	1.709e-05	3.612
	1/128	1.549e-09	4.855	9.283e-09	4.798	9.867e-07	4.114

Table 4: Errors and orders of convergence of (5.20) for the two-dimensional Burgers equation at $t=0.1$. Degree $2k$ volume quadrature rules are used.

k	h	L^1 error	order	L^2 error	order	L^∞ error	order
2	1/16	2.321e-04	-	4.779e-04	-	6.456e-03	-
	1/32	3.456e-05	2.748	8.407e-05	2.507	1.431e-03	2.173
	1/64	4.432e-06	2.963	1.088e-05	2.950	1.888e-04	2.923
	1/128	5.748e-07	2.947	1.566e-06	2.797	3.105e-05	2.604
	1/256	7.532e-08	2.932	2.244e-07	2.803	5.511e-06	2.494
3	1/16	6.136e-05	-	2.201e-04	-	3.608e-03	-
	1/32	5.050e-06	3.603	2.199e-05	3.323	6.367e-04	2.503
	1/64	3.266e-07	3.951	1.513e-06	3.861	3.610e-05	4.141
	1/128	2.031e-08	4.007	9.928e-08	3.929	3.837e-06	3.234
	1/256	1.330e-09	3.933	6.533e-09	3.926	3.184e-07	3.591
4	1/8	1.787e-04	-	5.778e-04	-	7.072e-03	-
	1/16	1.298e-05	3.783	5.361e-05	3.430	1.356e-03	2.382
	1/32	5.751e-07	4.497	3.081e-06	4.121	1.346e-04	3.334
	1/64	1.693e-08	5.086	8.712e-08	5.144	6.191e-06	4.442
	1/128	5.198e-10	5.026	2.711e-09	5.006	1.902e-07	5.024

$$\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt} = \mathbf{r}^\kappa\left(\overrightarrow{\mathbf{u}^\mathcal{S}}\right), \tag{6.1}$$

such that

$$\sum_{\kappa=1}^{K}\left(\overrightarrow{\mathbf{1}^\kappa}\right)^T \mathbf{M}^\kappa \mathbf{r}^\kappa\left(\overrightarrow{\mathbf{u}^\mathcal{S}}\right) = 0, \quad \sum_{\kappa=1}^{K}\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{M}^\kappa \mathbf{r}^\kappa\left(\overrightarrow{\mathbf{u}^\mathcal{S}}\right) \le 0.$$

For clarity, proofs of the theorems in this section will be provided in Appendix B.

6.1 Compatibility with limiters

For the classic DG method, people have developed a wide class of limiters, such as the TVD/TVB (total variation diminishing/bounded) limiter [75], the bound-preserving limiter [92, 93] and the WENO limiter [68], to enable extra stabilization. The idea of limiters can certainly be transferred to nodal DG formulations. Generally speaking, after applying some limiter, we compute the modified set of nodal values, denoted by $\overrightarrow{\mathbf{u}^{\kappa,\text{new}}}$, for each $1 \le \kappa \le K$. We require that the average value of $\mathbf{u}$ on T_κ is unchanged, as the primary conservation should be maintained.

$$\frac{1}{|T_\kappa|}\sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa \mathbf{u}_j^{\kappa,\text{new}} = \frac{1}{|T_\kappa|}\sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa \mathbf{u}_j^\kappa = \overline{\mathbf{u}}^\kappa.$$

In [12], the authors proved that if the limiter squeezes the data towards the average value, the total amount of entropy will not increase.

Theorem 6.1. *Suppose that the modified values are given by*

$$\mathbf{u}_j^{\kappa,new} = \overline{\mathbf{u}}^\kappa + \lambda_j^\kappa(\mathbf{u}_j^\kappa - \overline{\mathbf{u}}^\kappa), \tag{6.2}$$

where $0 \le \lambda_j^\kappa \le 1$ for each $1 \le j \le \mathcal{N}_{Q,k}$ and $1 \le \kappa \le K$, then for any convex entropy function U, we have

$$\left(\overrightarrow{1^\kappa}\right)^T M^\kappa \overrightarrow{U^{\kappa,new}} \le \left(\overrightarrow{1^\kappa}\right)^T M^\kappa \overrightarrow{U^\kappa}, \quad \textit{i.e.,} \quad \sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa U(\mathbf{u}_j^{\kappa,new}) \le \sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa U(\mathbf{u}_j^\kappa). \tag{6.3}$$

For instance, in the bound-preserving limiter, we perform a simple linear scaling procedure $\mathbf{u}_j^{\kappa,\text{new}} = \overline{\mathbf{u}}^\kappa + \lambda^\kappa(\mathbf{u}_j^\kappa - \overline{\mathbf{u}}^\kappa)$ with $0 \le \lambda^\kappa \le 1$, making sure the modified nodal values are within some physical bound. Therefore the bound-preserving limiter will not violate entropy stability. A special entropy stable TVD/TVB limiter was also designed for one-dimensional scalar conservation laws in [12]. These limiters make the most sense only for quadrature rules with collocated surface nodes. On general set of nodes, due to the emergence of entropy-extrapolated values, proving the bound-preserving property or the TVD/TVB property is very challenging, despite the fact the proof of Theorem 6.1 still holds.

Remark 6.1. Limiters only work for fully discrete schemes. The argument is incomplete unless we prove the entropy stability of the fully discrete version of (6.1). Time discretization will be discussed later.

6.2 Convection-diffusion equations

We add viscous diffusive terms to the conservation law (2.1):

$$\frac{\partial \mathbf{u}}{\partial t}+\sum_{m=1}^{d}\frac{\partial}{\partial x_m}\Big(\mathbf{f}_m(\mathbf{u})-\sum_{r=1}^{d}C_{mr}(\mathbf{v})\frac{\partial \mathbf{v}}{\partial \mathbf{x}_r}\Big)=0, \tag{6.4}$$

where $\mathbf{v}$ is the entropy variable of some entropy function U, and $C_{mr}(\mathbf{v})$ are $p\times p$ matrix-valued functions. One typical examples is the compressible Navier-Stokes equations. We assume that the matrix

$$\begin{bmatrix} C_{11}(\mathbf{v}) & \cdots & C_{1d}(\mathbf{v}) \\ \vdots & \ddots & \vdots \\ C_{d1}(\mathbf{v}) & \cdots & C_{dd}(\mathbf{v}) \end{bmatrix}$$

is symmetric semi-positive-definite. Then (6.4) supports the entropy inequality with respect to U. Entropy stable discretization of (6.4) is investigated in [4,5,37], where a nodal version of the local discontinuous Galerkin (LDG) method [6,16] is introduced. We recast (6.4) into the mixed form

$$\frac{\partial \mathbf{u}}{\partial t}+\sum_{m=1}^{d}\frac{\partial}{\partial x_m}(\mathbf{f}_m(\mathbf{u})-\mathbf{q}_m)=0,\quad \mathbf{q}_m=\sum_{r=1}^{d}C_{mr}(\mathbf{v})\boldsymbol{\theta}_r,\quad \boldsymbol{\theta}_r=\frac{\partial \mathbf{v}}{\partial x_r}. \tag{6.5}$$

The LDG method evolves the approximations of $\mathbf{u}$ and $\{\boldsymbol{\theta}_r\}_{r=1}^{d}$ simultaneously. Once again for each $1\le\kappa\le K$, $\overrightarrow{\mathbf{u}^{\kappa}}$ and $\overrightarrow{\boldsymbol{\theta}_r^{\kappa}}$ denote the vector of nodal values in T_κ. We further define that

$$\mathbf{C}_{mr}^{\kappa}=\mathrm{diag}\{C_{mr}(\mathbf{v}_1^{\kappa}),\cdots,C_{mr}(\mathbf{v}_{\mathcal{N}_{Q,k}}^{\kappa})\},\quad \overrightarrow{\mathbf{q}_m^{\kappa}}=\sum_{r=1}^{d}\mathbf{C}_{mr}^{\kappa}\overrightarrow{\boldsymbol{\theta}_r^{\kappa}}.$$

Neighboring elements are coupled via $\widehat{\mathbf{f}}_{\mathbf{n}}(\mathbf{u}_L,\mathbf{u}_R)$, as well as single-valued numerical fluxes of $\mathbf{v}$ and $\mathbf{q}_{\mathbf{n}}$:

$$\widehat{\mathbf{v}}=\widehat{\mathbf{v}}(\mathbf{v}_L,\mathbf{v}_R),\quad \widehat{\mathbf{q}}_{\mathbf{n}}=\widehat{\mathbf{q}}_{\mathbf{n}}(\mathbf{v}_L,\mathbf{v}_R,\mathbf{q}_{\mathbf{n},L},\mathbf{q}_{\mathbf{n},R}). \tag{6.6}$$

On the face $\gamma\in\partial T_\kappa$, we also let $\overrightarrow{\mathbf{v}^{\gamma\kappa,*}}$ and $\overrightarrow{\mathbf{q}_{\mathbf{n}}^{\gamma\kappa,*}}$ describe the vectors of the nodal values of corresponding numerical fluxes:

$$\overrightarrow{\mathbf{v}^{\gamma\kappa,*}}=\begin{bmatrix}\widehat{\mathbf{v}}(\mathbf{v}_1^{\gamma\kappa},\mathbf{v}_1^{\gamma\nu})\\ \vdots \\ \widehat{\mathbf{v}}(\mathbf{v}_{\mathcal{N}_{B,k}}^{\gamma\kappa},\mathbf{v}_{\mathcal{N}_{B,k}}^{\gamma\nu})\end{bmatrix},\quad \overrightarrow{\mathbf{q}_{\mathbf{n}}^{\gamma\kappa,*}}=\begin{bmatrix}\widehat{\mathbf{q}}_{\mathbf{n}}(\mathbf{v}_1^{\gamma\kappa},\mathbf{v}_1^{\gamma\nu},\mathbf{q}_{\mathbf{n},1}^{\gamma\kappa},\mathbf{q}_{\mathbf{n},1}^{\gamma\nu})\\ \vdots \\ \widehat{\mathbf{q}}_{\mathbf{n}}(\mathbf{v}_{\mathcal{N}_{B,k}}^{\gamma\kappa},\mathbf{v}_{\mathcal{N}_{B,k}}^{\gamma\nu},\mathbf{q}_{\mathbf{n},\mathcal{N}_{B,k}}^{\gamma\kappa},\mathbf{q}_{\mathbf{n},\mathcal{N}_{B,k}}^{\gamma\nu})\end{bmatrix}.$$

The LDG discretization of (6.5) is

$$\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt} = \mathbf{r}^\kappa\left(\overrightarrow{\mathbf{u}^g}\right) + \underbrace{\sum_{m=1}^{d} \mathbf{D}_m^\kappa \overrightarrow{\mathbf{q}_m^\kappa}}_{\text{Difference term}} - \underbrace{\sum_{\gamma\in\partial T_\kappa} (\mathbf{M}^\kappa)^{-1}(\mathbf{R}^{\gamma\kappa})^T \mathbf{B}^\gamma \left(\overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa}} - \overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa,*}}\right)}_{\text{Simultaneous approximation term}}, \tag{6.7a}$$

$$\overrightarrow{\boldsymbol{\theta}_r^\kappa} = \underbrace{\mathbf{D}_r^\kappa \overrightarrow{\mathbf{v}^\kappa}}_{\text{Difference term}} - \underbrace{\sum_{\gamma\in\partial T_\kappa} n_r^{\gamma\kappa}(\mathbf{M}^\kappa)^{-1}(\mathbf{R}^{\gamma\kappa})^T \mathbf{B}^\gamma \left(\overrightarrow{\mathbf{v}^{\gamma\kappa}} - \overrightarrow{\mathbf{v}^{\gamma\kappa,*}}\right)}_{\text{Simultaneous approximation term}}, \quad 1\le r\le d. \tag{6.7b}$$

Theorem 6.2. *Given parameters $\alpha\ge 0$ and $\beta\in\mathbb{R}$, if we choose the LDG fluxes*

$$\widehat{\mathbf{v}}(\mathbf{v}_L,\mathbf{v}_R) = \frac{1}{2}(\mathbf{v}_L+\mathbf{v}_R) + \beta(\mathbf{v}_L-\mathbf{v}_R), \tag{6.8a}$$

$$\widehat{\mathbf{q}}_\mathbf{n}(\mathbf{v}_L,\mathbf{v}_R,\mathbf{q}_{\mathbf{n},L},\mathbf{q}_{\mathbf{n},R}) = \frac{1}{2}(\mathbf{q}_{\mathbf{n},L}+\mathbf{q}_{\mathbf{n},R}) - \beta(\mathbf{q}_{\mathbf{n},L}-\mathbf{q}_{\mathbf{n},R}) - \alpha(\mathbf{v}_L-\mathbf{v}_R), \tag{6.8b}$$

then (6.7) *is entropy stable with respect to U.*

6.3 Modal formulation

We have only considered nodal DG formulations up to now. The recovery of modal formulations was explored in [7–9]. The idea is similar to the staggered-grid DG methods in [25, 67]. The polynomial basis functions $\{p_l(\mathbf{x})\}_{l=1}^{\mathcal{N}_{P,k}}$ play the role of *solution points* (say, we can assume that they are Lagrangian interpolation polynomials), where the numerical solution is stored; while the quadrature nodes $\{\mathbf{x}_j^\kappa\}_{j=1}^{\mathcal{N}_{Q,k}}$ are regarded as *flux points*, where the function evaluations take place. The communication between these two sets of points is via entropy variables, which brings us the concept of *entropy-projected* values. Vandermonde matrix V^κ and projection matrix P^κ are the corresponding interpolation operators.

Recall the notations in Section 3.3. Let $\mathbf{u}_h^\kappa(\mathbf{x})$ be the numerical solution, and $\overrightarrow{\widehat{\mathbf{u}}^\kappa}$ be the vector of polynomial coefficients on T_κ. The vector of nodal values is $\overrightarrow{\mathbf{u}^\kappa} = \mathbf{V}^\kappa \overrightarrow{\widehat{\mathbf{u}}^\kappa}$. For the entropy variables $\mathbf{v}$, we define the projected polynomial:

$$\overrightarrow{\widehat{\mathbf{v}}^\kappa} = \mathbf{P}^\kappa \overrightarrow{\mathbf{v}^\kappa}, \quad \mathbf{v}_h^\kappa(\mathbf{x}) = \sum_{l=1}^{\mathcal{N}_{Q,k}} \widehat{\mathbf{v}}_l^\kappa p_l(\mathbf{x}),$$

as well as the entropy-projected values $\overrightarrow{\widetilde{\mathbf{v}}^\kappa}$ and $\overrightarrow{\widetilde{\mathbf{u}}^\kappa}$, such that

$$\overrightarrow{\widetilde{\mathbf{v}}^\kappa} = \mathbf{V}^\kappa \overrightarrow{\widehat{\mathbf{v}}^\kappa} = \mathbf{V}^\kappa \mathbf{P}^\kappa \overrightarrow{\mathbf{v}^\kappa} = \begin{bmatrix} \mathbf{v}_h(\mathbf{x}_1^\kappa) \\ \vdots \\ \mathbf{v}_h(\mathbf{x}_{\mathcal{N}_{Q,k}}^\kappa) \end{bmatrix}, \quad \overrightarrow{\widetilde{\mathbf{u}}^\kappa} = \begin{bmatrix} \mathbf{u}(\widetilde{\mathbf{v}}_1^\kappa) \\ \vdots \\ \mathbf{u}(\widetilde{\mathbf{v}}_{\mathcal{N}_{Q,k}}^\kappa) \end{bmatrix}.$$

Now given the generic entropy stable nodal DG formulation (6.1), its modal counterpart is derived through projection and inserting entropy-projected values:

$$\frac{d\overrightarrow{\widehat{\mathbf{u}}^\kappa}}{dt}=\mathbf{P}^\kappa\mathbf{r}^\kappa\Big(\overrightarrow{\widetilde{\mathbf{u}}^\mathcal{S}}\Big),\quad \overrightarrow{\widetilde{\mathbf{u}}^\mathcal{S}}=\begin{bmatrix}\overrightarrow{\widetilde{\mathbf{u}}^1}\\ \vdots\\ \overrightarrow{\widetilde{\mathbf{u}}^K}\end{bmatrix}. \tag{6.9}$$

Theorem 6.3. *Under standard assumptions, if* (6.1) *is conservative and entropy stable, then the modal formulation* (6.9) *is also conservative and entropy stable, in the sense that*

$$\frac{d}{dt}\Big(\int_\Omega \mathbf{u}_h(t,\mathbf{x})d\mathbf{x}\Big)=\frac{d}{dt}\Big(\sum_{\kappa=1}^K\Big(\overrightarrow{\mathbf{1}^\kappa}\Big)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{u}^\kappa}\Big)=0,\quad \frac{d}{dt}\Big(\sum_{\kappa=1}^K\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa\overrightarrow{U^\kappa}\Big)\le 0. \tag{6.10}$$

6.4 Curvilinear meshes

Curvilinear meshes are usually preferred in the decomposition of domains with complex geometry. For a curvilinear mesh, still denoted by $\{T_\kappa\}_{\kappa=1}^K$, suppose that there exists a reference simplex element T (with reference coordinates $\boldsymbol{\xi}$), such that T_κ is the image of T under some invertible mapping $\boldsymbol{\xi}\mapsto\mathbf{x}_\kappa(\boldsymbol{\xi})$. We define the Jacobian factor $J^\kappa=\det(\mathbf{x}'_\kappa(\boldsymbol{\xi}))$, and the metric terms

$$G^\kappa_{mr}=J^\kappa\frac{\partial\xi_r}{\partial x^\kappa_m},\quad 1\le m,r\le d.$$

The metric terms satisfy the following geometric conservation law [59]

$$\sum_{r=1}^d\frac{\partial G^\kappa_{mr}}{\partial\xi_r}=0,\quad \text{for each } 1\le m\le d. \tag{6.11}$$

Then we rewrite the conservation law (2.1) in terms of reference coordinates:

$$J^\kappa\frac{\partial\mathbf{u}}{\partial t}+\sum_{r=1}^d\frac{\partial}{\partial\xi_r}\Big(\sum_{m=1}^d G^\kappa_{mr}\mathbf{f}_m(\mathbf{u})\Big)=0. \tag{6.12}$$

This is actually a problem with variable coefficients, as both J^κ and G^κ_{mr} are non-constant functions (they are constant only for simplex meshes where all mappings are affine). In the case that the quadrature rule has collocated surface nodes, an entropy stable DG method for (6.12) was created by Fisher in [28], and applied to different problems in [3,4, 37]. The main difficulty lies in the treatment of the metric terms. Roughly speaking, the nodal values of the metric terms must satisfy the discrete geometric conservation law (i.e., the discrete version of (6.11)). Exact evaluation of the metric terms will in general fail. One possible procedure for computing two-dimensional and three-dimensional metric

terms was uncovered in [59]. Moreover, metric terms should be averaged in the flux differencing term, which corresponds to the split form

$$J^\kappa \frac{\partial \mathbf{u}}{\partial t} + \frac{1}{2}\sum_{r=1}^{d}\Big(\frac{\partial}{\partial \xi_r}\Big(\sum_{m=1}^{d} G_{mr}^\kappa \mathbf{f}_m(\mathbf{u})\Big) + \sum_{m=1}^{d} G_{mr}^\kappa \frac{\partial \mathbf{f}_m(\mathbf{u})}{\partial \xi_r}\Big) = 0. \tag{6.13}$$

It is a well-known splitting technique for problems with variable coefficients (see e.g. [48]). For general quadrature rules, the same idea was used in [9] to derive a curvilinear variant of the hybridized SBP operators, such that the discrete geometric conservation law is written in terms of volume metric terms and surface metric terms. The idea also works for global SBP operators, but at the cost of requiring a global discrete geometric conservation law. A slightly different approach was presented in [20] to maintain the locality of the geometric conservation law.

6.5 Time discretization

One important motivation for quadrature-based DG formulations is the pursuit of entropy stable methods that can be exactly implemented. This goal is only partly accomplished due to the assumption of semi-discrete analysis. Fully discrete entropy stability is mostly established for implicit time stepping schemes. For example, applying the Euler backward scheme to (6.1) yields

$$\overrightarrow{\mathbf{u}^{\kappa,(n+1)}} = \overrightarrow{\mathbf{u}^{\kappa,(n)}} + \Delta t \mathbf{r}^\kappa \Big(\overrightarrow{\mathbf{u}^{g,(n+1)}}\Big), \tag{6.14}$$

where $\overrightarrow{\mathbf{u}^{\kappa,(n)}}$ is the solution vector on T_κ at the n-th step. By the convexity of U,

$$\begin{aligned}
&\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa \overrightarrow{U^{\kappa,(n+1)}} \le \sum_{\kappa=1}^{K}\Big(\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa \overrightarrow{U^{\kappa,h}} + \Big(\overrightarrow{\mathbf{v}^{\kappa,(n+1)}}\Big)^T \mathbf{M}^\kappa\Big(\overrightarrow{\mathbf{u}^{\kappa,(n+1)}} - \overrightarrow{\mathbf{u}^{\kappa,(n)}}\Big)\Big)\\
=&\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa \overrightarrow{U^{\kappa,h}} + \Delta t\sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{v}^{\kappa,(n+1)}}\Big)^T \mathbf{M}^\kappa \mathbf{r}^\kappa\Big(\overrightarrow{\mathbf{u}^{g,(n+1)}}\Big) \le \sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa \overrightarrow{U^{\kappa,h}}.
\end{aligned}$$

Hence the Euler backward time stepping is entropy stable. A general framework of high order entropy stable implicit time stepping schemes was discussed in [61]. Time discretization can also be handled by the so-called space-time DG technique, in which we regard the time variable as an extra dimension, and the equation (2.1) as a steady state conservation law in $(d+1)$ dimensions. Then we directly adopt existing entropy stable methods to discretize the steady state problem. See [34] for the space-time version of quadrature-based DG methods, and [2, 53, 87] for the space-time version of (2.21) (DG method that approximates entropy variables and assumes exact integration). Clearly, the space-time DG methods are also implicit in time.

In contrast, the entropy stability of explicit time discretization is by large an open problem. For the first order method using monotone fluxes (in scalar problems) or Godunov type fluxes, it is well-known that the entropy stability result is still valid in the

fully discrete case with Euler forward time stepping (see e.g. Chapter 3 of [41]). In the context of high order DG methods, the L^2 stability of linear equations is proved for Runge-Kutta time stepping [91], and Lax-Wendroff time stepping [79]. However, relatively little is known for the nonlinear extension (in the sense of both flux functions and entropy function). Entropy stable relaxation Runge-Kutta methods were demonstrated in [72], where a (multiplicative) relaxation parameter was applied to existing Runge-Kutta methods at each time step. In order to achieve entropy stability, the value of the relaxation parameter was determined by solving a nonlinear equation. See also [63] for an experimental study on the entropy stability of explicit Runge-Kutta methods.

7 Concluding remarks

Starting from the pioneering work in [5,36], high order entropy stable quadrature-based DG methods have been developed into an exuberant research area. These DG methods can be stable with respect to an arbitrary given entropy function. Therefore we circumvent the limitation of the L^2 stability result for the classic DG method. There are three main ingredients contributing to entropy stability:

1. Discrete operators with the summation-by-parts property. The existence of difference matrices is clinched by Eq. (3.10).
2. Flux differencing technique. We apply difference matrices to bivariate entropy conservative fluxes, instead of the univariate flux functions, to enable chain rule. The "brute force" method in Section 5.3 provides another option of enforcing entropy balance.
3. Entropy stable SATs that couple adjacent elements. For quadrature rules with collocated surface nodes, simply inserting entropy stable fluxes on interfaces is enough. For general set of nodes, we need to put more effort. Two possible constructions of SATs, implicitly implemented in augmented SBP operators, are reviewed in Section 5.1 and Section 5.2.

The entropy stable DG framework is of great versatility in that a variety of concepts can be incorporated into it. We have only discussed a few topics in this paper, including the generalization to convection-diffusion equations, the transformation between nodal and modal formulations, the handling of curvilinear meshes and the development of fully-discrete methods. The bound-preserving limiter and TVD/TVB limiter can be imposed on quadrature points with collocated surface nodes. On the other hand, the main advantage of general quadrature rules is the possibility of attaining better accuracy with smaller number of degrees of freedom.

We speculate some possible directions for future research:

1. Rigorous error analysis for smooth problems. There are positive results for the classic DG method in [55,89,90].

2. Establishing convergence of numerical solutions. Since the DG methods only satisfy a single entropy condition, we might not be able to show convergence to the entropy solution. The paradigm of measure-valued solutions was used in [33, 53].
3. Bound-preserving limiter for general quadrature rules. This is of practical importance. For problems with strong shocks, the code is likely to crash due to non-physical values (e.g. negative density or negative pressure in Euler equations), and bound-preserving limiter is usually desired.
4. Entropy stable explicit time stepping schemes.

Acknowledgments

This research was supported by NSF grant DMS-1719410 and ARO grant W911NF-16-1-0103.

Appendices

A Equivalence of flux differencing and splitting

In this section, we build the link between flux differencing and splitting, and also present some special examples to further illustrate the equivalence. For the sake of simplicity, let us consider the one-dimensional conservation law:

$$\frac{\partial \mathbf{u}}{\partial t}+\frac{\partial \mathbf{f}(\mathbf{u})}{\partial x}=0. \tag{A.1}$$

We assume that the entropy conservative flux $\mathbf{f}_S(\mathbf{u}_L,\mathbf{u}_R)$ is separable. That is, there exists a finite sequence of functions $\{\mathbf{g}_i(\mathbf{u})\}_{i=1}^n$, such that $\mathbf{f}_S$ has the symmetric decomposition

$$\mathbf{f}_S(\mathbf{u}_L,\mathbf{u}_R)=\sum_{i=1}^{n}\mathbf{g}_i(\mathbf{u}_L)\circ\mathbf{g}_{n+1-i}(\mathbf{u}_R). \tag{A.2}$$

By consistency of $\mathbf{f}_S$,

$$\mathbf{f}(\mathbf{u})=\sum_{i=1}^{n}\mathbf{g}_i(\mathbf{u})\circ\mathbf{g}_{n+1-i}(\mathbf{u}). \tag{A.3}$$

As a result, if $\mathbf{u}$ is a smooth solution,

$$\begin{aligned}\frac{\partial \mathbf{f}(\mathbf{u})}{\partial x}&=\sum_{i=1}^{n}\Big(\mathbf{g}_i(\mathbf{u})\circ\frac{\partial \mathbf{g}_{n+1-i}(\mathbf{u})}{\partial x}+\frac{\partial \mathbf{g}_i(\mathbf{u})}{\partial x}\circ\partial \mathbf{g}_{n+1-i}(\mathbf{u})\Big)\\&=2\sum_{i=1}^{n}\mathbf{g}_i(\mathbf{u})\circ\frac{\partial \mathbf{g}_{n+1-i}(\mathbf{u})}{\partial x}.\end{aligned}$$

We have the split form of (A.1)

$$\frac{\partial \mathbf{u}}{\partial t} = 2\sum_{i=1}^{n} \mathbf{g}_i(\mathbf{u}) \circ \frac{\partial \mathbf{g}_{n+1-i}(\mathbf{u})}{\partial x}. \tag{A.4}$$

Theorem A.1. *Under the assumption of separable* $\mathbf{f}_S$*, the flux differencing term is actually the discretization of split form* (A.4)*:*

$$2\mathbf{D}^\kappa \circ \mathbf{F}_S\left(\overrightarrow{\mathbf{u}^k}, \overrightarrow{\mathbf{u}^k}\right)\overrightarrow{\mathbf{1}^k} = 2\sum_{i=1}^{n} \overrightarrow{\mathbf{g}_i^k} \circ \left(\mathbf{D}^\kappa \overrightarrow{\mathbf{g}^\kappa_{n+1-i}}\right). \tag{A.5}$$

Proof. We simply examine each component of the flux differencing term:

$$\begin{aligned}\left(2\mathbf{D}^\kappa \circ \mathbf{F}_S\left(\overrightarrow{\mathbf{u}^k}, \overrightarrow{\mathbf{u}^k}\right)\overrightarrow{\mathbf{1}^k}\right)_j &= 2\sum_{l=1}^{\mathcal{N}_{Q,k}} D^\kappa_{jl} \mathbf{f}_S(\mathbf{u}^\kappa_j, \mathbf{u}^\kappa_l) = 2\sum_{i=1}^{n} \mathbf{g}_i(\mathbf{u}^\kappa_j) \circ \Big(\sum_{l=1}^{\mathcal{N}_{Q,k}} D^\kappa_{jl} \mathbf{g}_{n+1-i}(\mathbf{u}^\kappa_l)\Big) \\ &= 2\sum_{i=1}^{n} \mathbf{g}^\kappa_{i,j} \circ \left(\mathbf{D}^\kappa \overrightarrow{\mathbf{g}^\kappa_{n+1-i}}\right)_j.\end{aligned}$$

The proof is completed. □

A.1 Linear symmetric system

Consider the one-dimensional linear symmetric system

$$\frac{\partial \mathbf{u}}{\partial t} + \frac{\partial (A\mathbf{u})}{\partial x} = 0, \tag{A.6}$$

where A is some constant symmetric $p \times p$ matrix. The square function $U = \frac{1}{2}\mathbf{u}^T\mathbf{u}$ defines an entropy function, with

$$\mathbf{v} = \mathbf{u}, \quad F = \frac{1}{2}\mathbf{u}^T A\mathbf{u}, \quad \psi = \mathbf{v}^T\mathbf{f} - F = \frac{1}{2}\mathbf{u}^T A\mathbf{u}.$$

We simply take $\mathbf{f}_S$ to be the arithmetic mean

$$\mathbf{f}_S(\mathbf{u}_L, \mathbf{u}_R) = \frac{1}{2}(A\mathbf{u}_L + A\mathbf{u}_R) = \frac{1}{2}(\mathbf{f}(\mathbf{u}_L) + \mathbf{f}(\mathbf{u}_R)). \tag{A.7}$$

Hence the flux differencing term reduces to

$$2\mathbf{D}^\kappa \circ \mathbf{F}_S\left(\overrightarrow{\mathbf{u}^k}, \overrightarrow{\mathbf{u}^k}\right) = \overrightarrow{\mathbf{1}^k} \circ \left(\mathbf{D}^\kappa \overrightarrow{\mathbf{f}^k}\right) + \overrightarrow{\mathbf{f}^k} \circ \left(\mathbf{D}^\kappa \overrightarrow{\mathbf{1}^k}\right) = \mathbf{D}^\kappa \overrightarrow{\mathbf{f}^k}.$$

We recover the difference term in the unmodified DG method (3.19).

A.2 Burgers equation

For the one-dimensional Burgers equation

$$\frac{\partial u}{\partial t}+\frac{1}{2}\frac{\partial u^2}{\partial x}=0,$$

we still use the square entropy function $U=\frac{u^2}{2}$, with

$$v=u,\quad F=\psi=\frac{u^3}{3},\quad \psi=\frac{u^3}{6}.$$

The entropy conservative flux is

$$f_S(u_L,u_R)=\frac{\psi_R-\psi_L}{v_R-v_L}=\frac{1}{6}(u_L^2+u_Lu_R+u_R^2). \tag{A.8}$$

Then

$$\begin{aligned}2D^\kappa\circ F_S\left(\overrightarrow{u^\kappa},\overrightarrow{u^\kappa}\right)\overrightarrow{1^\kappa}&=\frac{2}{3}\overrightarrow{1^\kappa}\circ\left(D^\kappa\overrightarrow{f^\kappa}\right)+\frac{1}{3}\overrightarrow{u^\kappa}\circ\left(D^\kappa\overrightarrow{u^\kappa}\right)+\frac{2}{3}\overrightarrow{f^\kappa}\circ\left(D^\kappa\overrightarrow{1^\kappa}\right)\\&=\frac{2}{3}D^\kappa\overrightarrow{f^\kappa}+\frac{1}{3}\overrightarrow{u^\kappa}\circ\left(D^\kappa\overrightarrow{u^\kappa}\right),\end{aligned}$$

which is the discretization of the split form

$$\frac{\partial u}{\partial t}+\frac{1}{3}\frac{\partial u^2}{\partial x}+\frac{1}{3}u\frac{\partial u}{\partial x}=0. \tag{A.9}$$

It is called the skew-symmetric splitting technique in [36, 71].

A.3 Shallow water equations

The one-dimensional shallow water equations read

$$\frac{\partial}{\partial t}\begin{bmatrix}h\\hw\end{bmatrix}+\frac{\partial}{\partial x}\begin{bmatrix}hw\\hw^2+\frac{1}{2}gh^2\end{bmatrix}=0. \tag{A.10}$$

Here h and w are the water depth and velocity, and g is the gravity acceleration constant. The total energy function $U=\frac{1}{2}hw^2+\frac{1}{2}gh^2$ serves as an entropy function with

$$\mathbf{v}=\begin{bmatrix}gh-\frac{1}{2}w^2\\w\end{bmatrix},\quad F=\frac{1}{2}hw^3+gh^2w,\quad \psi=\frac{1}{2}gh^2w.$$

We can construct the separable entropy conservative flux

$$\mathbf{f}_S(\mathbf{u}_L,\mathbf{u}_R)=\begin{bmatrix}\frac{1}{2}(h_Lw_L+h_Rw_R)\\\frac{1}{4}(h_Lw_L+h_Rw_R)(w_L+w_R)+\frac{1}{2}gh_Lh_R\end{bmatrix}. \tag{A.11}$$

The corresponding flux differencing term is equivalent to the discretization of

$$\frac{\partial h}{\partial t}+\frac{\partial (hw)}{\partial x}=0, \tag{A.12a}$$

$$\frac{\partial (hw)}{\partial t}+\frac{1}{2}\frac{\partial (hw^2)}{\partial x}+\frac{1}{2}w\frac{\partial (hw)}{\partial x}+\frac{1}{2}hw\frac{\partial w}{\partial x}+gh\frac{\partial h}{\partial x}=0, \tag{A.12b}$$

which is the skew-symmetric splitting procedure in [39].

B Proofs of the theorems in Section 6

B.1 Theorem 6.1

We will prove Theorem 6.1. Since

$$\frac{1}{|T_\kappa|}\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa \mathbf{u}_j^{\kappa,\text{new}}=\frac{1}{|T_\kappa|}\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(\overline{\mathbf{u}}^\kappa+\lambda_j^\kappa(\mathbf{u}_j^\kappa-\overline{\mathbf{u}}^\kappa))=\overline{\mathbf{u}}^\kappa,$$

we have

$$\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa\lambda_j^\kappa\mathbf{u}_j^\kappa=\Big(\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa\lambda_j^\kappa\Big)\overline{\mathbf{u}}^\kappa,\quad \sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(1-\lambda_j^\kappa)\mathbf{u}_j^\kappa=\Big(\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(1-\lambda_j^\kappa)\Big)\overline{\mathbf{u}}^\kappa.$$

By the convexity of U,

$$\begin{aligned}&U(\mathbf{u}_j^{\kappa,\text{new}})\le\lambda_j^\kappa U(\mathbf{u}_j^\kappa)+(1-\lambda_j^\kappa)U(\overline{\mathbf{u}}^\kappa),\\ &\Big(\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(1-\lambda_j^\kappa)\Big)U(\overline{\mathbf{u}}^\kappa)\le\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(1-\lambda_j^\kappa)U(\mathbf{u}_j^\kappa).\end{aligned}$$

Therefore

$$\begin{aligned}\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa U(\mathbf{u}_j^{\kappa,\text{new}})&\le\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(\lambda_j^\kappa U(\mathbf{u}_j^\kappa)+(1-\lambda_j^\kappa)U(\overline{\mathbf{u}}^\kappa))\\ &=\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa\lambda_j^\kappa U(\mathbf{u}_j^\kappa)+\Big(\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(1-\lambda_j^\kappa)\Big)U(\overline{\mathbf{u}}^\kappa)\\ &\le\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa\lambda_j^\kappa U(\mathbf{u}_j^\kappa)+\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa(1-\lambda_j^\kappa)U(\mathbf{u}_j^\kappa)\\ &=\sum_{j=1}^{\mathcal{N}_{Q,k}}\omega_j^\kappa U(\mathbf{u}_j^\kappa).\end{aligned}$$

B.2 Theorem 6.2

We will prove Theorem 6.2, i.e., the entropy stability of LDG method (6.7) for convection-diffusion equations. We neglect the convective term $\mathbf{r}^\kappa\left(\overrightarrow{\mathbf{u}^\kappa}\right)$ as it is already entropy stable. Then the entropy growth rate in T_κ is

$$\frac{d}{dt}\left(\left(\overrightarrow{1^\kappa}\right)^T M^\kappa \overrightarrow{U^\kappa}\right) = \sum_{m=1}^{d}\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{S}_m^\kappa \overrightarrow{\mathbf{q}_m^\kappa} - \sum_{\gamma\in\partial T_\kappa}\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T \mathbf{B}^\gamma\left(\overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa}} - \overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa,*}}\right).$$

We left multiply (6.7b) by $\left(\overrightarrow{\mathbf{q}_r^\kappa}\right)^T \mathbf{M}^\kappa$:

$$\left(\overrightarrow{\mathbf{q}_r^\kappa}\right)^T \mathbf{M}^\kappa \overrightarrow{\boldsymbol{\theta}_r^\kappa} = \left(\overrightarrow{\mathbf{q}_r^\kappa}\right)^T \mathbf{S}_r^\kappa \overrightarrow{\mathbf{v}^\kappa} - \sum_{\gamma\in\partial T_\kappa} n_r^{\gamma\kappa}\left(\overrightarrow{\mathbf{q}_r^{\gamma\kappa}}\right)^T \mathbf{B}^\gamma\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}} - \overrightarrow{\mathbf{v}^{\gamma\kappa,*}}\right), \quad 1\le r\le d.$$

Summing up the two identities above yields:

$$\begin{aligned}\frac{d}{dt}\left(\left(\overrightarrow{1^\kappa}\right)^T M^\kappa \overrightarrow{U^\kappa}\right) = &-\sum_{r=1}^{d}\left(\overrightarrow{\mathbf{q}_r^\kappa}\right)^T \mathbf{M}^\kappa \overrightarrow{\boldsymbol{\theta}_r^\kappa} + \sum_{r=1}^{d}\left(\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{S}_r^\kappa \overrightarrow{\mathbf{q}_r^\kappa} + \left(\overrightarrow{\mathbf{q}_r^\kappa}\right)^T \mathbf{S}_r^\kappa \overrightarrow{\mathbf{v}^\kappa}\right)\\ &- \sum_{\gamma\in\partial T_\kappa}\left(\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T \mathbf{B}^\gamma\left(\overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa}} - \overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa,*}}\right) + \left(\overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa}}\right)^T \mathbf{B}^\gamma\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}} - \overrightarrow{\mathbf{v}^{\gamma\kappa,*}}\right)\right).\end{aligned}$$

The first sum is non-positive since

$$\begin{aligned}-\sum_{r=1}^{d}\left(\overrightarrow{\mathbf{q}_r^\kappa}\right)^T \mathbf{M}^\kappa \overrightarrow{\boldsymbol{\theta}_r^\kappa} &= -\sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa\left(\sum_{r=1}^{d}(\mathbf{q}_{r,j}^\kappa)^T \boldsymbol{\theta}_{r,j}^\kappa\right)\\ &= -\sum_{j=1}^{\mathcal{N}_{Q,k}} \omega_j^\kappa\left(\sum_{m=1}^{d}\sum_{r=1}^{d}(\boldsymbol{\theta}_{m,j}^\kappa)^T C_{mr}(\mathbf{v}_j^\kappa)\boldsymbol{\theta}_{r,j}^\kappa\right) \le 0.\end{aligned}$$

The second sum, according to the SBP property, equals

$$\begin{aligned}&\sum_{r=1}^{d}\left(\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{S}_r^\kappa \overrightarrow{\mathbf{q}_r^\kappa} + \left(\overrightarrow{\mathbf{q}_r^\kappa}\right)^T \mathbf{S}_r^\kappa \overrightarrow{\mathbf{v}^\kappa}\right)\\ =&\sum_{r=1}^{d}\sum_{\gamma\in\partial T_\kappa} n_r^{\kappa\gamma}\left(\overrightarrow{\mathbf{v}^\kappa}\right)^T \mathbf{E}^{\kappa\gamma}\overrightarrow{\mathbf{q}_r^\kappa} = \sum_{\gamma\in\partial T_\kappa}\left(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\right)^T \mathbf{B}^\gamma \overrightarrow{\mathbf{q}_\mathbf{n}^{\gamma\kappa}}.\end{aligned}$$

Now there are only interface terms. We sum over κ and get

$$\begin{aligned}
&\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa \overrightarrow{U^\kappa}\Big)\\
\le&\sum_{\kappa=1}^{K}\sum_{\gamma\in\partial T_\kappa}\Big(\Big(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\Big)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{q}_{\mathbf{n}}^{\gamma\kappa,*}}+\Big(\overrightarrow{\mathbf{q}_{\mathbf{n}}^{\gamma\kappa}}\Big)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{v}^{\gamma\kappa,*}}-\Big(\overrightarrow{\mathbf{v}^{\gamma\kappa}}\Big)^T\mathbf{B}^\gamma\overrightarrow{\mathbf{q}_{\mathbf{n}}^{\gamma\kappa}}\Big)\\
=&\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma\Big((\mathbf{v}_s^{\gamma\kappa}-\mathbf{v}_s^{\gamma\nu})^T\widehat{\mathbf{q}}_{\mathbf{n}^{\gamma\kappa}}(\mathbf{v}_s^{\gamma\kappa},\mathbf{v}_s^{\gamma\nu},\mathbf{q}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\kappa},\mathbf{q}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\nu})+(\mathbf{q}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\kappa}-\mathbf{q}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\nu})^T\widehat{\mathbf{v}}(\mathbf{v}_s^{\gamma\kappa},\mathbf{v}_s^{\gamma\nu})\\
&\qquad-((\mathbf{v}_s^{\gamma\kappa})^T\mathbf{q}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\kappa}-(\mathbf{v}_s^{\gamma\nu})^T\mathbf{q}_{\mathbf{n}^{\gamma\kappa},s}^{\gamma\nu})\Big).
\end{aligned}$$

By the definition of LDG fluxes (6.8) and the identity

$$\mathbf{v}_L^T\mathbf{q}_{\mathbf{n},L}-\mathbf{v}_R^T\mathbf{q}_{\mathbf{n},R}=\frac{1}{2}(\mathbf{v}_L+\mathbf{v}_R)^T(\mathbf{q}_{\mathbf{n},L}-\mathbf{q}_{\mathbf{n},R})+\frac{1}{2}(\mathbf{v}_L-\mathbf{v}_R)^T(\mathbf{q}_{\mathbf{n},L}+\mathbf{q}_{\mathbf{n}_R}),$$

we are left with

$$\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa\overrightarrow{U^\kappa}\Big)\le-\sum_{\gamma\in\Gamma_h}\sum_{s=1}^{\mathcal{N}_{B,k}}\tau_s^\gamma\alpha(\mathbf{v}_s^{\gamma\kappa}-\mathbf{v}_s^{\gamma\nu})^T(\mathbf{v}_s^{\gamma\kappa}-\mathbf{v}_s^{\gamma\nu})\le 0.$$

Hence the LDG method is entropy stable with respect to U.

B.3 Theorem 6.3

We will prove Theorem 6.3, i.e., the conservation and entropy stability of the modal formulation (6.9). The evolution of nodal values is

$$\frac{d\overrightarrow{\mathbf{u}^\kappa}}{dt}=\mathbf{V}^\kappa\mathbf{P}^\kappa\mathbf{r}^\kappa\Big(\overrightarrow{\widetilde{\mathbf{u}}^\xi}\Big).$$

Since $\mathbf{M}^\kappa\mathbf{V}^\kappa\mathbf{P}^\kappa=(\mathbf{P}^\kappa)^T\widehat{\mathbf{M}}^\kappa\mathbf{P}^\kappa=(\mathbf{V}^\kappa\mathbf{P}^\kappa)\mathbf{M}^\kappa$,

$$\begin{aligned}
&\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{1}^\kappa}\Big)^T\mathbf{M}^\kappa\overrightarrow{\mathbf{u}^\kappa}\Big)=\sum_{\kappa=1}^{K}\Big(\mathbf{V}^\kappa\mathbf{P}^\kappa\overrightarrow{\mathbf{1}^\kappa}\Big)^T\mathbf{M}^\kappa\mathbf{r}^\kappa\Big(\overrightarrow{\widetilde{\mathbf{u}}^\xi}\Big)=\sum_{\kappa=1}^{K}\Big(\overrightarrow{\mathbf{1}^\kappa}\Big)^T\mathbf{M}^\kappa\mathbf{r}^\kappa\Big(\overrightarrow{\widetilde{\mathbf{u}}^\xi}\Big),\\
&\frac{d}{dt}\Big(\sum_{\kappa=1}^{K}\Big(\overrightarrow{1^\kappa}\Big)^T M^\kappa\overrightarrow{U^\kappa}\Big)=\sum_{\kappa=1}^{K}\Big(\mathbf{V}^\kappa\mathbf{P}^\kappa\overrightarrow{\mathbf{v}^\kappa}\Big)^T\mathbf{M}^\kappa\mathbf{r}^\kappa\Big(\overrightarrow{\widetilde{\mathbf{u}}^\xi}\Big)=\sum_{\kappa=1}^{K}\Big(\overrightarrow{\widetilde{\mathbf{v}}^\kappa}\Big)^T\mathbf{M}^\kappa\mathbf{r}^\kappa\Big(\overrightarrow{\widetilde{\mathbf{u}}^\xi}\Big).
\end{aligned}$$

Then from the conservation and entropy stability of (6.1), we see that (6.9) is also conservative and entropy stable.

References

[1] R. Abgrall. A general framework to construct schemes satisfying additional conservation relations. Application to entropy conservative and entropy dissipative schemes. *Journal of Computational Physics*, 372:640–666, 2018.

[2] T. J. Barth. Numerical methods for gasdynamic systems on unstructured meshes. In *An Introduction to Recent Developments in Theory and Numerics for Conservation Laws*, Springer, 1999, pp.195–285.

[3] M. W. Bohm, A. R. Winters, G. J. Gassner, D. Derigs, F. J. Hindenlang and J. Saur. An entropy stable nodal discontinuous Galerkin method for the resistive MHD equations. Part I: Theory and numerical verification. *Journal of Computational Physics*, Article 108076, 2018. DOI: 10.1016/j.jcp.2018.06.027

[4] M. H. Carpenter, T. C. Fisher, E. J. Nielsen, M. Parsani, M. Svärd and N. Yamaleev. Entropy stable summation-by-parts formulations for compressible computational fluid dynamics. In *Handbook of Numerical Analysis*, volume 17, Elsevier, 2016, pp.495–524.

[5] M. H. Carpenter, T. C. Fisher, E. J. Nielsen and S. H. Frankel. Entropy stable spectral collocation schemes for the Navier–Stokes equations: discontinuous interfaces. *SIAM Journal on Scientific Computing*, 36:B835–B867, 2014.

[6] P. Castillo, B. Cockburn, I. Perugia, and D. Schötzau. An a priori error analysis of the local discontinuous Galerkin method for elliptic problems *SIAM Journal on Numerical Analysis*, 38:1676–1706, 2000.

[7] J. Chan. On discretely entropy conservative and entropy stable discontinuous Galerkin methods. *Journal of Computational Physics*, 362:346–374, 2018.

[8] J. Chan, D. C. Del Rey Fernández and M. H. Carpenter. Efficient entropy stable Gauss collocation methods. *SIAM Journal on Scientific Computing*, 41:A2938–A2966, 2019.

[9] J. Chan and L. C. Wilcox. On discretely entropy stable weight-adjusted discontinuous Galerkin methods: curvilinear meshes. *Journal of Computational Physics*, 378:366–393, 2019.

[10] P. Chandrashekar. Kinetic energy preserving and entropy stable finite volume schemes for compressible Euler and Navier-Stokes equations. *Communications in Computational Physics*, 14:1252–1286, 2013.

[11] P. Chandrashekar and C. Klingenberg. Entropy stable finite volume scheme for ideal compressible MHD on 2-D Cartesian meshes. *SIAM Journal on Numerical Analysis*, 54:1313–1340, 2016.

[12] 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*, 345:427–461, 2017.

[13] 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*, 54:545–581, 1990.

[14] 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*, 84:90–113, 1989.

[15] 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*, 52:411–435, 1989.

[16] B. Cockburn and C.-W. Shu. The local discontinuous Galerkin method for time-dependent convection-diffusion systems. *SIAM Journal on Numerical Analysis*, 35:2440–2463, 1998.

[17] B. Cockburn and C.-W. Shu. The Runge–Kutta discontinuous Galerkin method for conservation laws V: multidimensional systems. *Journal of Computational Physics*, 141:199–224, 1998.
[18] M. G. Crandall and A. Majda. Monotone difference approximations for scalar conservation laws. *Mathematics of Computation*, 34:1–21, 1980.
[19] J. Crean, J. E. Hicken, D. C. Del Rey Fernández, D. W. Zingg and M. H. Carpenter. High-order, entropy-stable discretizations of the Euler equations for complex geometries. In *23rd AIAA Computational Fluid Dynamics Conference. American Institute of Aeronautics and Astronautics*, 2017.
[20] 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*, 356:410-438, 2018.
[21] J. Crean, K. Panda, A. Ashley and J. E. Hicken. Investigation of stabilization methods for multi-dimensional summation-by-parts discretizations of the Euler equations. In *54th AIAA Aerospace Sciences Meeting*, 2016.
[22] C. M. Dafermos. *Hyperbolic Conservation Laws in Continuum Physics*, volume 325, Springer, 2010.
[23] D. Derigs, A. R. Winters, G. J. Gassner, S. Walch and M. W. Bohm. Ideal GLM-MHD: About the entropy consistent nine-wave magnetic field divergence diminishing ideal magnetohydrodynamics equations. *Journal of Computational Physics*, 364:420–467, 2018.
[24] 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*, 266:214–239, 2014.
[25] D. C. Del Rey Fernández, J. Crean, M. H. Carpenter and J. E. Hicken. Staggered-grid entropy-stable multidimensional summation-by-parts discretizations on curvilinear coordinates. *Journal of Computational Physics*, 392:161–186, 2019.
[26] 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*, 95:171–196, 2014.
[27] 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*, 75:83–110, 2018.
[28] T. C. Fisher. *High-order L2 stable multi-domain finite difference method for compressible flows*, Ph.D. Thesis, Purdue University, 2012.
[29] T. C. Fisher and M. H. Carpenter. High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains. *Journal of Computational Physics*, 252:518–557, 2013.
[30] U. S. Fjordholm, S. Mishra and E. Tadmor. Well-balanced and energy stable schemes for the shallow water equations with discontinuous topography. *Journal of Computational Physics*, 230:5587–5609, 2011.
[31] U. S. Fjordholm, S. Mishra and E. Tadmor. Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws. *SIAM Journal on Numerical Analysis*, 50:544–573, 2012.
[32] U. S. Fjordholm, S. Mishra and E. Tadmor. ENO reconstruction and ENO interpolation are stable. *Foundations of Computational Mathematics*, 13:139–159, 2013.
[33] U. S. Fjordholm, S. Mishra and E. Tadmor. On the computation of measure-valued solutions. *Acta Numerica*, 25:567–679, 2016.
[34] L. Friedrich, G. Schnücke, A. R. Winters, D. C. Del Rey Fernández, G. J. Gassner

and M. H. Carpenter. Entropy stable space–time discontinuous Galerkin schemes with summation-by-parts property for hyperbolic conservation laws. *Journal of Scientific Computing*, 80:175–222, 2018.

[35] L. Friedrich, A. R. Winters, D. C. Del Rey Fernández, G. J. Gassner, M. Parsani and M. H. Carpenter. An entropy stable h/p non-conforming discontinuous Galerkin method with the summation-by-parts property. *Journal of Scientific Computing*, 77:689–725, 2018.

[36] G. J. Gassner. A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods. *SIAM Journal on Scientific Computing*, 35:A1233–A1253, 2013.

[37] G. J. Gassner, A. R. Winters, F. J. Hindenlang, and D. A. Kopriva. The BR1 scheme is stable for the compressible Navier–Stokes equations. *Journal of Scientific Computing*, 77:154–200, 2018.

[38] G. J. Gassner, A. R. Winters and D. A. Kopriva. Split form nodal discontinuous Galerkin schemes with summation-by-parts property for the compressible Euler equations. *Journal of Computational Physics*, 327:39–66, 2016.

[39] G. J. Gassner, A. R. Winters and D. A. Kopriva. A well balanced and entropy conservative discontinuous Galerkin spectral element method for the shallow water equations. *Applied Mathematics and Computation*, 272:291–308, 2016.

[40] C. Geuzaine and J.-F. Remacle. Gmsh: A 3-D finite element mesh generator with built-in pre-and post-processing facilities. *International Journal for Numerical Methods in Engineering*, 79:1309–1331, 2009.

[41] E. Godlewski and P.-A. Raviart. *Hyperbolic Systems of Conservation Laws*, Ellipses, 1991.

[42] E. Godlewski and P.-A. Raviart. *Numerical Approximation of Hyperbolic Systems of Conservation Laws*, volume 118, Springer, 2013.

[43] S. Gottlieb, C.-W. Shu and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Review*, 43:89–112, 2001.

[44] B. Gustafsson, H.-O. Kreiss, and J. Oliger. *Time Dependent Problems and Difference Methods*, volume 24, John Wiley & Sons, 1995.

[45] A. Harten, B. Engquist, S. Osher and S. R. Chakravarthy. Uniformly high order accurate essentially non-oscillatory schemes, III. In *Upwind and High-resolution Schemes*, Springer, 1987, pp.218–290.

[46] A. Harten, J. M. Hyman, P. D. Lax and B. Keyfitz. On finite-difference approximations and entropy conditions for shocks. *Communications on Pure and Applied Mathematics*, 29:297–322, 1976.

[47] A. Harten, P. D. Lax and B. Van Leer. On upstream differencing and Godunov-type schemes for hyperbolic conservation laws. *SIAM Review*, 25:35–61, 1983.

[48] J. S. Hesthaven, S. Gottlieb and D. Gottlieb. *Spectral Methods for Time-dependent Problems*, volume 21, Cambridge University Press, 2007.

[49] J. S. Hesthaven and T. Warburton. *Nodal Discontinuous Galerkin Methods: Algorithms, Analysis, and Applications*, Springer, 2007.

[50] J. E. Hicken. Entropy-stable, high-order discretizations using continuous summation-by-parts operators. In *AIAA Aviation 2019 Forum*, 2019.

[51] J. E. Hicken and J. Crean. A family of entropy-conservative flux functions for the Euler equations. *arXiv preprint*, arXiv:1807.03832, 2018.

[52] 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*, 38:A1935–A1958, 2016.

[53] A. Hiltebrand and S. Mishra. Entropy stable shock capturing space–time discontinuous Galerkin schemes for systems of conservation laws. *Numerische Mathematik*, 126:103–151, 2014.
[54] S. Hou and X.-D. Liu. Solutions of multi-dimensional hyperbolic systems of conservation laws by square entropy condition satisfying discontinuous Galerkin method. *Journal of Scientific Computing*, 31:127–151, 2007.
[55] J. Huang and C.-W. Shu. Error estimates to smooth solutions of semi-discrete discontinuous Galerkin methods with quadrature rules for scalar conservation laws. *Numerical Methods for Partial Differential Equations*, 33:467–488, 2017.
[56] T. J. R. Hughes, L. P. Franca, and M. Mallet. A new finite element formulation for computational fluid dynamics: I. Symmetric forms of the compressible Euler and Navier–Stokes equations and the second law of thermodynamics. *Computer Methods in Applied Mechanics and Engineering*, 54:223–234, 1986.
[57] F. Ismail and P. L. Roe. Affordable, entropy-consistent Euler flux functions II: Entropy production at shocks. *Journal of Computational Physics*, 228:5410–5436, 2009.
[58] G.-S. Jiang and C.-W. Shu. On a cell entropy inequality for discontinuous Galerkin methods. *Mathematics of Computation*, 62:531–538, 1994.
[59] D. A. Kopriva. Metric identities and the discontinuous spectral element method on curvilinear meshes. *Journal of Scientific Computing*, 26:301–327, 2006.
[60] 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*, 44:136–155, 2010.
[61] P. G. Lefloch, J.-M. Mercier and C. Rohde. Fully discrete, entropy conservative schemes of arbitrary order. *SIAM Journal on Numerical Analysis*, 40:1968–1992, 2002.
[62] Y. Liu, C.-W. Shu and M. Zhang. Entropy stable high order discontinuous Galerkin methods for ideal compressible MHD on structured meshes. *Journal of Computational Physics*, 354:163–178, 2018.
[63] C. Lozano. Entropy production by explicit Runge–Kutta schemes. *Journal of Scientific Computing*, 76:521–564, 2018.
[64] P. Öffner, J. Glaubitz and H. Ranocha. Stability of correction procedure via reconstruction with summation-by-parts operators for Burgers' equation using a polynomial chaos approach. *ESAIM: M2AN*, 52:2215–2245, 2018.
[65] S. Ortleb. A kinetic energy preserving DG scheme based on Gauss–Legendre points. *Journal of Scientific Computing*, 71:1135–1168, 2017.
[66] S. Osher Riemann solvers, the entropy condition, and difference. *SIAM Journal on Numerical Analysis*, 21:217–235, 1984.
[67] M. Parsani, M. H. Carpenter, T. C. Fisher and E. J. Nielsen. Entropy stable staggered grid discontinuous spectral collocation methods of any order for the compressible Navier–Stokes equations. *SIAM Journal on Scientific Computing*, 38:A3129–A3162, 2016.
[68] J. Qiu and C.-W. Shu. Runge–Kutta discontinuous Galerkin method using WENO limiters. *SIAM Journal on Scientific Computing*, 26:907–929, 2005.
[69] H. Ranocha. Shallow water equations: split-form, entropy stable, well-balanced, and positivity preserving numerical methods. *GEM-International Journal on Geomathematics*, 8:85–133, 2017.
[70] H. Ranocha. *Generalised summation-by-parts operators and entropy stability of numerical methods for hyperbolic balance laws*, Ph.D. Thesis, Technische Universitt Braunschweig, 2018.
[71] H. Ranocha, P. Öffner and T. Sonar. Extended skew-symmetric form for summation-by-parts

operators and varying Jacobians. *Journal of Computational Physics*, 342:13–28, 2017.
[72] H. Ranocha, M. Sayyari, L. Dalcin, M. Parsani and D. I. Ketcheson. Relaxation Runge-Kutta methods: fully-discrete explicit entropy-stable schemes for the Euler and Navier-Stokes equations. *SIAM Journal on Scientific Computing*, 42:A612–A638, 2020.
[73] F. Renac. Entropy stable DGSEM for nonlinear hyperbolic systems in nonconservative form with application to two-phase flows. *Journal of Computational Physics*, 382:1–26, 2019.
[74] G. Schnücke, N. Krais, T. Bolemann and G. J. Gassner. Entropy stable discontinuous Galerkin schemes on moving meshes for hyperbolic conservation laws. *Journal of Scientific Computing*, 82:69, 2020.
[75] C.-W. Shu. TVB uniformly high-order schemes for conservation laws. *Mathematics of Computation*, 49:105–121, 1987.
[76] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *Journal of Computational Physics*, 77:439–471, 1988.
[77] Z. Sun, J. A. Carillo and C.-W. Shu. A discontinuous Galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials. *Journal of Computational Physics*, 352:76–104, 2018.
[78] Z. Sun, J. A. Carillo and C.-W. Shu. An entropy stable high-order discontinuous Galerkin method for cross-diffusion gradient flow systems. *Kinetic and Related Models*, 12:885–908, 2019.
[79] Z. Sun and C.-W. Shu. Stability analysis and error estimates of Lax–Wendroff discontinuous Galerkin methods for linear conservation laws. *ESAIM: Mathematical Modelling and Numerical Analysis*, 51:1063–1087, 2017.
[80] M. Svärd and J. Nordström. Review of summation-by-parts schemes for initial–boundary-value problems. *Journal of Scientific Computing*, 58:61–89, 2014.
[81] E. Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. I. *Mathematics of Computation*, 49:91–103, 1987.
[82] E. Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numerica*, 12:451–512, 2003.
[83] N. Wintermeyer, A. R. Winters, G. J. Gassner and D. A. Kopriva. An entropy stable nodal discontinuous Galerkin method for the two dimensional shallow water equations on unstructured curvilinear meshes with discontinuous bathymetry. *Journal of Computational Physics*, 340:200–242, 2017.
[84] N. Wintermeyer, A. R. Winters, G. J. Gassner and T. Warburton. An entropy stable discontinuous Galerkin method for the shallow water equations on curvilinear meshes with wet/dry fronts accelerated by GPUs. *Journal of Computational Physics*, 375:447–480, 2018.
[85] A. R. Winters and G. J. Gassner. A comparison of two entropy stable discontinuous Galerkin spectral element approximations for the shallow water equations with non-constant topography. *Journal of Computational Physics*, 301:357–376, 2015.
[86] D. Xiu. *Numerical Methods for Stochastic Computations: a Spectral Method Approach*, Princeton University Press, 2010.
[87] M. Zakerzadeh and G. May. Entropy stable discontinuous Galerkin scheme for the compressible Navier-Stokes equations. In *55th AIAA Aerospace Sciences Meeting*, 2017, pp.0084.
[88] L. Zhang, T. Cui and H. Liu. A set of symmetric quadrature rules on triangles and tetrahedra. *Journal of Computational Mathematics*, 27:89–96, 2009.
[89] Q. Zhang and C.-W. Shu. Error estimates to smooth solutions of Runge–Kutta discontinuous Galerkin methods for scalar conservation laws. *SIAM Journal on Numerical Analysis*, 42:641–666, 2004.

[90] Q. Zhang and C.-W. Shu. Error estimates to smooth solutions of Runge–Kutta discontinuous Galerkin method for symmetrizable systems of conservation laws. *SIAM Journal on Numerical Analysis*, 44:1703–1720, 2006.

[91] Q. Zhang and C.-W. Shu. Stability analysis and a priori error estimates of the third order explicit Runge–Kutta discontinuous Galerkin method for scalar conservation laws. *SIAM Journal on Numerical Analysis*, 48:1038–1063, 2010.

[92] X. Zhang and C.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *Journal of Computational Physics*, 229:3091–3120, 2010.

[93] X. Zhang and C.-W. Shu. On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *Journal of Computational Physics*, 229:8918–8934, 2010.