Journal of Scientific Computing manuscript No.
(will be inserted by the editor)

Interior Penalties for Summation-by-Parts Discretizations of Linear Second-Order Differential Equations

Jianfeng Yan · Jared Crean · Jason E. Hicken

Abstract This work focuses on simultaneous approximation terms (SATs) for multidimensional summation-by-parts (SBP) discretizations of linear second-order partial differential equations with variable coefficients. Through the analysis of adjoint consistency and stability, we present several conditions on the SAT penalties for general operators, including those operators that do not have nodes on their boundary or do not correspond with a collocation discontinuous Galerkin method. Based on these conditions, we generalize the modified scheme of Bassi and Rebay (BR2) and the symmetric interior penalty Galerkin (SIPG) method to SBP-SAT discretizations. Numerical experiments are carried out on unstructured grids with triangular elements to verify the theoretical results.

1 Introduction

Recently, Fisher and Carpenter [1] showed that diagonal-norm summation-by-parts (SBP) operators can be used to construct *provably entropy-stable* semi-discretizations of the Euler and Navier-Stokes equations; see also [2]. This result is significant because it opens the door to nonlinearly-stable, high-order discretizations that do not rely on exact integration. Such schemes have the potential to be efficient and robust for industrially relevant problems.

The original theory in [1,2] was developed for tensor-product operators based on classical SBP finite-difference methods [3,4], so there have been several efforts to generalize the results [5–9]. In particular, given the possible limitations of hexahedral mesh generation, we are interesting in generalizing the theory to include

Jianfeng Yan, Graduate Student
Rensselaer Polytechnic Institute
E-mail: yanj4@rpi.edu

Jared Crean, Graduate Student
Rensselaer Polytechnic Institute
E-mail: creanj@rpi.edu

Jason E. Hicken, Assistant Professor
Rensselaer Polytechnic Institute
E-mail: hickej2@rpi.edu

multidimensional SBP operators [10], which can support more diverse element shapes and nodal distributions for simulations on unstructured grids.

In order to develop entropy-stable discretizations of the Navier-Stokes equations for general SBP operators, both the inviscid and viscous terms must be considered. For treatment of the inviscid terms in the context of multidimensional SBP discretizations, we direct the interested reader to [8] and [9]. The present work, which focuses on linear elliptic and parabolic operators, is motivated by the entropy-stable treatment of the viscous terms in the Navier-Stokes equations. Specifically, we are interested in the precise conditions for constructing accurate, stable, and adjoint-consistent interior penalties that enforce boundary conditions and inter-element coupling.

Interior penalties are known as simultaneous approximation terms (SATs) in the SBP literature [11]. Penalties for second-order PDEs have been well studied by both the SBP community [12–15] and the finite-element community (see the review [16] and the references therein). Nevertheless, multidimensional SBP-SAT discretizations introduce generalizations that have not, to the best of our knowledge, been considered in the either the DG or the SBP-SAT literature.

- The DG literature assumes explicit basis functions, and several results in the FE context rely on this assumption, e.g. the inverse trace inequalities of Warburton and Hesthaven [17]. SBP operators do not have a unique underlying basis, in general.
- The SBP-SAT literature typically assumes the interface nodes of adjacent elements coincide. In those cases when nonconforming nodes are considered, e.g. [13], the nodes are usually assumed to lie on the interface. In [15], the authors consider tensor-product discretizations without nodes on the interfaces, but only the Baumann-Oden [18] penalty is investigated. Furthermore, adjoint consistency is rarely addressed in the SBP literature and has not been considered for operators whose nodes are strictly interior to the element.

Based on the above gaps, the objective of this work is to identify the conditions on the SAT coefficient matrices to obtain multidimensional SBP-SAT discretizations that are simultaneously consistent, conservative, adjoint consistent, and stable. In the process of meeting this goal, we generalize the modified scheme of Bassi and Rebay (BR2) [19] and the symmetric interior penalty Galerkin (SIPG) method [16, 20, 21] to multidimensional SBP discretizations. In addition, we show how, in the SBP-SAT context, SIPG can be derived from BR2 using matrix analysis.

The remaining sections are organized as follows. We introduce our notation and the model PDE in Section 2. Section 3 reviews the multidimensional SBP definition and describes the matrices used to discretize various continuous operations. Section 4 then presents the SBP-SAT discretization of the model PDE. Section 5 investigates the adjoint consistency of the discretization and delineates the necessary adjoint-consistency conditions on the SAT penalties. The penalties are further constrained by the energy-stability analysis in Section 6. The resulting conditions are used to generalize the BR2 and SIPG methods to multidimensional SBP discretizations in Section 7. Verification studies are provided in Section 8, and a summary is provided in Section 9.

2 Preliminaries

2.1 Notation

Functions are denoted with capital letters in calligraphic font; for example $\mathcal{U} \in L^2(\Omega)$ is a square-integrable function on the domain Ω. A function evaluated on a node set is denoted by a lowercase letter in bold font. For example, the function $\mathcal{U}$ evaluated at the nodes $X = \{(x_i, y_i)\}_{i=1}^{n}$ is given by

$$\boldsymbol{u} = \left[\mathcal{U}(x_1, y_1)\; \mathcal{U}(x_2, y_2)\; \cdots\; \mathcal{U}(x_n, y_n)\right]^T .$$

The space of polynomials of total degree p, or less, in x and y on Ω is denoted by $\mathbb{P}_p(\Omega)$. As with generic functions, a polynomial that is evaluated at the points of X will be represented using its corresponding lowercase letter in bold font; for example, for $\mathcal{P} \in \mathbb{P}_p(\Omega)$ we would have

$$\boldsymbol{p} = \left[\mathcal{P}(x_1, y_1)\; \mathcal{P}(x_2, y_2)\; \cdots\; \mathcal{P}(x_n, y_n)\right]^T .$$

Matrices are represented with an uppercase sans-serif type, for example $\mathsf{A} \in \mathbb{R}^{n\times m}$. Unless indicated otherwise, a subscript indicates a vector or matrix evaluated on a particular element or face. For example $\boldsymbol{u}_\kappa$ and $(\mathsf{D}_x)_\kappa$ are the solution and derivative operator on element Ω_κ, respectively.

2.2 The model parabolic PDE

We consider the following linear second-order parabolic PDE — or the corresponding steady Poisson PDE — defined on the compact domain $\Omega \subset \mathbb{R}^2$:

$$\frac{\partial \mathcal{U}}{\partial t} = \nabla \cdot (\Lambda \nabla \mathcal{U}) + \mathcal{F}, \quad \forall\, (x, y) \in \Omega,\ t \in [0, T], \tag{1}$$

where $\mathcal{F} \in L^2(\Omega)$ is a given source term, and

$$\Lambda \equiv \begin{bmatrix} \lambda_{xx} & \lambda_{xy} \\ \lambda_{yx} & \lambda_{yy} \end{bmatrix}$$

is a symmetric, positive-definite tensor that is a smooth function of (x, y). The parabolic PDE is provided with the initial condition

$$\mathcal{U}(0, x, y) = \mathcal{U}_0(x, y), \quad \forall\, (x, y) \in \Omega, \tag{2}$$

where $\mathcal{U}_0 \in L^2(\Omega)$. Finally, the PDE is supplied with the steady Dirichlet and Neumann boundary conditions,

$$\begin{aligned} \mathcal{U}(t, x, y) &= \mathcal{U}_{\mathcal{D}}(x, y), && \forall\, (x, y) \in \Gamma^{\mathcal{D}}, \\ (\Lambda \nabla \mathcal{U}(t, x, y)) \cdot \boldsymbol{n} &= \mathcal{U}_{\mathcal{N}}(x, y), && \forall\, (x, y) \in \Gamma^{\mathcal{N}}, \end{aligned} \tag{3}$$

respectively, where $\boldsymbol{n} = [n_x, n_y]^T$ is the outward pointing unit normal on the boundary $\partial\Omega = \Gamma^{\mathcal{D}} \cup \Gamma^{\mathcal{N}}$, with $\partial\Omega \setminus \Gamma^{\mathcal{D}} = \Gamma^{\mathcal{N}}$. Finally, we assume that the data

— $\mathcal{F}$, $\Lambda, \mathcal{U}_0, \mathcal{U}_{\mathcal{D}}$, and $\mathcal{U}_{\mathcal{N}}$ — and the geometry are such that initial-boundary-value problem (1)–(3) is well posed.

In addition to the above strong form of the PDE, we will also refer to an associated weak formulation, specifically

$$\int_{\Omega} \mathcal{V} \frac{\partial \mathcal{U}}{\partial t} \, \mathrm{d}\Omega = \mathcal{R}(\mathcal{U}, \mathcal{V}), \qquad \forall\, \mathcal{V} \in \mathbb{H}^1(\Omega), \tag{4}$$

where the spatial residual $\mathcal{R}(\cdot,\cdot) : \mathbb{H}^1(\Omega) \times \mathbb{H}^1(\Omega) \to \mathbb{R}$ is defined by

$$\begin{aligned} \mathcal{R}(\mathcal{U}, \mathcal{V}) \equiv & -\int_{\Omega} (\nabla \mathcal{V})^T \Lambda (\nabla \mathcal{U}) \, \mathrm{d}\Omega + \int_{\Omega} \mathcal{V} \mathcal{F} \, \mathrm{d}\Omega \\ & + \int_{\Gamma^{\mathcal{N}}} \mathcal{V} \mathcal{U}_{\mathcal{N}} \, \mathrm{d}\Gamma + \int_{\Gamma^{\mathcal{D}}} \mathcal{V} (\Lambda \nabla \mathcal{U}) \cdot \mathbf{n} \, \mathrm{d}\Gamma + \int_{\Gamma^{\mathcal{D}}} (\mathcal{U} - \mathcal{U}_{\mathcal{D}}) (\Lambda \nabla \mathcal{V}) \cdot \mathbf{n} \, \mathrm{d}\Gamma. \end{aligned}$$

The first four terms in $\mathcal{R}$ are obtained by multiplying (1) by an arbitrary test function $\mathcal{V} \in \mathbb{H}^1(\Omega)$, integrating over the domain Ω, applying integration by parts, and then imposing the Neumann boundary conditions on $\Gamma^{\mathcal{N}}$. The last term in $\mathcal{R}$ is the residual associated with the Dirichlet boundary conditions.

3 Discrete operators

3.1 multidimensional SBP operators

We adopt the definition of multidimensional SBP operators proposed in [10]. To keep the presentation self-contained, the definition for an operator approximating $\partial/\partial x$ on a two dimensional domain is provided below. The definition for the SBP operator approximating $\partial/\partial y$ is analogous.

Definition 1 Two-dimensional summation-by-parts operator: Consider an open and bounded subdomain $\Omega_\kappa \subset \Omega$ with a piecewise-smooth boundary $\partial\Omega_\kappa$, and n_κ interior nodes $X_\kappa = \{(x_i, y_i)\}_{i=1}^{n_\kappa}$. The matrix D_x is a degree p SBP approximation to the first derivative $\frac{\partial}{\partial x}$ on the nodes X_κ if

1. for all $\mathcal{P} \in \mathbb{P}_p(\Omega_\kappa)$, the vector $\mathsf{D}_x \boldsymbol{p}_\kappa$ is equal to $\partial\mathcal{P}/\partial x$ at the nodes X_κ;
2. $\mathsf{D}_x = \mathsf{H}^{-1}\mathsf{Q}_x$, where H is symmetric positive-definite, and;
3. $\mathsf{Q}_x = \mathsf{S}_x + \frac{1}{2}\mathsf{E}_x$, where $\mathsf{S}_x^T = -\mathsf{S}_x$, $\mathsf{E}_x^T = \mathsf{E}_x$, and E_x satisfies

$$\boldsymbol{p}^T \mathsf{E}_x \boldsymbol{q} = \oint_{\partial\Omega_\kappa} \mathcal{P}\mathcal{Q} n_x \mathrm{d}\Gamma,$$

for all polynomials $\mathcal{P}, \mathcal{Q} \in \mathbb{P}_r(\Omega_\kappa)$, where $r \geq p$, and n_x is the x component of $\boldsymbol{n} = [n_x, n_y]^{\mathrm{T}}$, the outward pointing unit normal on $\partial\Omega_\kappa$.

The subsequent analysis is restricted to so-called diagonal-norm SBP operators, that is, SBP operators for which H is a diagonal matrix with positive entries. In this case, it was shown in [10] that the nodes X_κ and diagonal entries of H define a cubature rule that is exact for polynomials of total degree $2p-1$. Thus, we have the following approximation for sufficiently smooth functions $\mathcal{U}$ and $\mathcal{V}$:

$$\boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \boldsymbol{u}_\kappa = \int_{\Omega_\kappa} \mathcal{V}\mathcal{U} \, \mathrm{d}\Omega + \mathrm{O}(h^{2p}), \tag{5}$$

where h is a linear measure of the size of Ω_κ. Furthermore, the matrix operator Q_x can also be interpreted as an approximate integral [10]:

$$\boldsymbol{v}_\kappa^T \mathsf{Q}_x \boldsymbol{u}_\kappa = \int_{\Omega_\kappa} \mathcal{V} \frac{\partial \mathcal{U}}{\partial x} \, \mathrm{d}\Omega + \mathrm{O}(h^{\min(2p,r+1)}). \tag{6}$$

This interpretation of Q_x, together with the accuracy of E_x in Definition 1, leads directly to a high-order approximation of integration-by-parts:

$$\begin{aligned} &\boldsymbol{v}_\kappa^T \mathsf{Q}_x \boldsymbol{u}_\kappa + \boldsymbol{v}_\kappa^T \mathsf{Q}_x^T \boldsymbol{u}_\kappa = \boldsymbol{v}_\kappa^T \mathsf{E}_x \boldsymbol{u}_\kappa \\ \Rightarrow &\int_{\Omega_\kappa} \mathcal{V} \frac{\partial \mathcal{U}}{\partial x} \, \mathrm{d}\Omega + \int_{\Omega_\kappa} \mathcal{U} \frac{\partial \mathcal{V}}{\partial x} \, \mathrm{d}\Omega = \oint_{\partial\Omega_\kappa} \mathcal{V}\mathcal{U} n_x \, \mathrm{d}\Gamma + \mathrm{O}(h^{\min(2p,r+1)}). \end{aligned}$$

These relationships between the SBP matrices and integral bilinear forms will be helpful when we relate the SBP discretization to weak-form finite-element discretizations.

3.2 Face-based operators

In order to define SATs for multidimensional SBP operators, we follow References [22,23] and introduce interpolation/extrapolation operators from the SBP element nodes to cubature nodes on the faces of the elements.

Consider an element Ω_κ with a piecewise smooth boundary $\partial\Omega_\kappa$, and let $\gamma \subset \partial\Omega_\kappa$ denote one of its faces. Let $X_\gamma = \{(x_j, y_j)\}_{j=1}^{n_\gamma} \subset \gamma$ be a set of cubature nodes with corresponding positive weights $\{b_j\}_{j=1}^{n_\gamma}$ that is exact for polynomials of degree $2r$, where $r \geq p$ is the same integer appearing in Definition 1. The matrix $\mathsf{R}_{\gamma\kappa} \in \mathbb{R}^{n_\gamma \times n_\kappa}$ is a degree r interpolation/extrapolation operator from the SBP nodes X_κ to the face nodes X_γ if, for all $\mathcal{P} \in \mathbb{P}_r(\Omega_\kappa)$,

$$\left(\mathsf{R}_{\gamma\kappa} \boldsymbol{p}_\kappa\right)_j = \sum_{i=1}^{n_\kappa} (\mathsf{R}_{\gamma\kappa})_{ji} \mathcal{P}(x_i, y_i) = \mathcal{P}(x_j, y_j), \qquad \forall j = 1, 2, \ldots, n_\gamma. \tag{7}$$

In other words, $\mathsf{R}_{\gamma\kappa}$ exactly interpolates/extrapolates polynomials of degree r, where $r \geq p$, from the volume nodes of element Ω_κ to the nodes of its face γ.

For a given (strong) cubature rule of degree $2p-1$ defined on Ω_κ, it was shown in [23] that there exists at least one SBP operator whose corresponding matrix E_x has the decomposition

$$\mathsf{E}_x = \sum_{\gamma \subset \partial\Omega_\kappa} \mathsf{R}_{\gamma\kappa}^T \mathsf{N}_{x,\gamma} \mathsf{B}_\gamma \mathsf{R}_{\gamma\kappa}, \tag{8}$$

where $\mathsf{B}_\gamma = \mathrm{diag}\left(b_1, b_2, \ldots, b_{n_\gamma}\right)$ is an $n_\gamma \times n_\gamma$ diagonal matrix holding the cubature weights for γ along its diagonal, and $\mathsf{N}_{x,\gamma} = \mathrm{diag}\left(n_{x,1}, n_{x,2}, \ldots, n_{x,n_\gamma}\right)$ is an $n_\gamma \times n_\gamma$ diagonal matrix holding the x component of the outward unit normal with respect to Ω_κ at the cubature points of γ. We will assume in the following analysis that the SBP operators are such that E_x has the decomposition (8), and that the operators in the y direction have analogous decompositions.

Finally, we need to discretize the normal derivative operator, $\mathbf{n} \cdot (\Lambda\nabla)$, at the nodes of face γ. To this end, we introduce the diagonal matrices $\mathsf{\Lambda}_{xx}$, $\mathsf{\Lambda}_{xy}$, $\mathsf{\Lambda}_{yx}$

and $\mathsf{\Lambda}_{yy}$, which store the Cartesian elements of the tensor $\mathit{\Lambda}$ evaluated at the SBP nodes. For example,

$$\mathsf{\Lambda}_{xx} = \operatorname{diag}\left(\lambda_{xx}(x_1, y_1), \lambda_{xx}(x_2, y_2), \ldots, \lambda_{xx}(x_n, y_n)\right). \tag{9}$$

With these matrices, we can discretize the normal derivative operator as

$$\mathsf{D}_{\gamma\kappa} = \mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\kappa}\left(\mathsf{\Lambda}_{xx}\mathsf{D}_x + \mathsf{\Lambda}_{xy}\mathsf{D}_y\right)_\kappa + \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\kappa}\left(\mathsf{\Lambda}_{yx}\mathsf{D}_x + \mathsf{\Lambda}_{yy}\mathsf{D}_y\right)_\kappa. \tag{10}$$

Based on the accuracy of D_x, D_y, and $\mathsf{R}_{\gamma\kappa}$, the above discretization is exact for $\mathcal{U} \in \mathbb{P}_p(\mathit{\Omega}_\kappa)$ and $\mathit{\Lambda}\nabla\mathcal{U} \in \mathbb{P}_r(\mathit{\Omega}_\kappa)$; therefore, since $r \geq p$, $\mathsf{D}_{\gamma\kappa}$ gives an order h^{p+1} approximation to the normal derivative at the nodes of γ.

4 SBP discretization of parabolic PDEs

This section describes the SBP-SAT discretization of (1). This discretization uses the matrix operators we have already defined, as well as several yet-to-be-defined operators. For future reference, Table 1 summarizes all the matrix operators with a brief description of their purpose and the location in the text that provides a more detailed definition.

Table 1: Summary of matrix operators used in the SBP-SAT discretization.

Matrix	**Description**	**Refer to...**
H_κ	diagonal norm/mass matrix for element $\mathit{\Omega}_\kappa$	Def. 1
D_x (D_y)	SBP approximation of $\partial/\partial x$ (resp. $\partial/\partial y$)	Def. 1
$\mathsf{\Lambda}_{xx}$, $\mathsf{\Lambda}_{xy}$, $\mathsf{\Lambda}_{yx}$, $\mathsf{\Lambda}_{yy}$	hold λ_{xx}, λ_{xy}, λ_{yx}, λ_{yy} along their diagonal	Eq. (9)
D_κ	discretization of $\nabla \cdot (\mathit{\Lambda}\nabla)$	Eq. (12)
M_κ	used to discretize $\int_{\mathit{\Omega}_\kappa}(\nabla\mathcal{V})^T\mathit{\Lambda}(\nabla\mathcal{U})\, d\mathit{\Omega}$	Prop. 1
B_γ	diagonal matrix of cubature weights for face γ	Eq. (8)
$\mathsf{N}_{x,\gamma}$ ($\mathsf{N}_{y,\gamma}$)	diagonal matrix of n_x (resp. n_y) for face γ	Eq. (8)
$\mathsf{R}_{\gamma\kappa}$ ($\mathsf{R}_{\gamma\nu}$)	interpolate from nodes of κ (resp. ν) to nodes of γ	Eq. (7)
$\mathsf{D}_{\gamma\kappa}$ ($\mathsf{D}_{\gamma\nu}$)	approximation of $\boldsymbol{n} \cdot (\mathit{\Lambda}\nabla)$ on γ w.r.t. κ (resp. ν)	Eq. (10)
$\mathsf{T}^{(i)}_{\gamma\kappa}$, $i = 1, 2, 3, 4$	penalty coefficients for interface γ of element $\mathit{\Omega}_\kappa$	Eq. (16)
$\mathsf{T}^{\mathcal{D}}_{\gamma}$	penalty coefficients for Dirichlet face γ of element $\mathit{\Omega}_\kappa$	Eq. (17)

4.1 Discretization of spatial derivatives

Let $\mathcal{T}_h = \bigcup_{\kappa=1}^{K} \mathit{\Omega}_\kappa$ denote a partition of the domain $\mathit{\Omega}$ into K SBP elements, where $\mathit{\Omega}_\kappa$ denotes the domain of the κth element. The discrete solution on element $\mathit{\Omega}_\kappa$ will be represented by the vector $\boldsymbol{u}_\kappa \in \mathbb{R}^{n_\kappa}$ whose entries are the discrete solution at the SBP nodes X_κ. The global discrete solution, denoted $\boldsymbol{u}_h \in \mathbb{R}^{\sum n_\kappa}$, is the concatenation of all elementwise solutions.

Ignoring boundary conditions for the time being, a consistent SBP semi-discretization of (1) on element $\mathit{\Omega}_\kappa$ is given by

$$\frac{\mathrm{d}\boldsymbol{u}_\kappa}{\mathrm{d}t} = \mathsf{D}_\kappa \boldsymbol{u}_\kappa + \boldsymbol{f}_\kappa, \tag{11}$$

where $\boldsymbol{f}_\kappa$ is $\mathcal{F}$ evaluated at the nodes of element Ω_κ, and

$$\mathsf{D}_\kappa = \left\{ \begin{bmatrix} \mathsf{D}_x & \mathsf{D}_y \end{bmatrix} \begin{bmatrix} \mathsf{\Lambda}_{xx} & \mathsf{\Lambda}_{xy} \\ \mathsf{\Lambda}_{yx} & \mathsf{\Lambda}_{yy} \end{bmatrix} \begin{bmatrix} \mathsf{D}_x \\ \mathsf{D}_y \end{bmatrix} \right\}_\kappa, \tag{12}$$

is the SBP approximation of $\nabla \cdot (\Lambda \nabla)$ on element Ω_κ, with $\mathsf{D}_x \in \mathbb{R}^{n_\kappa \times n_\kappa}$ and $\mathsf{D}_y \in \mathbb{R}^{n_\kappa \times n_\kappa}$ being the first-derivative SBP operators in the x and y directions, respectively; see Definition 1.

Remark 1 Based on the form of D_κ in (12), our discretization falls in the class of "first-derivative twice" SBP approximations of the second-derivative. For classical finite-difference methods with repeating interior stencils, applying the first derivative twice approximately doubles the stencil size and is typically less accurate [24]; however, the SBP operators that we intend to use are dense matrices similar to spectral operators, for which applying the first derivative twice is "equivalent to differentiation with an explicitly formed second-derivative operator" [11].

Before incorporating boundary conditions, it is worth pausing to draw the connection between the strong-form discretization (11) and the integral weak form. To relate the SBP discretization to the weak form, we will need the following proposition, which is a straightforward consequence of the properties in Definition 1 and is stated without proof.

Proposition 1 *Let D_κ be defined as in (12). Then, $\forall\, \boldsymbol{u}_\kappa, \boldsymbol{v}_\kappa \in \mathbb{R}^{n_\kappa}$,*

$$\boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \mathsf{D}_\kappa \boldsymbol{u}_\kappa = -\boldsymbol{v}_\kappa^T \mathsf{M}_\kappa \boldsymbol{u}_\kappa + \sum_{\gamma \subset \partial\Omega_\kappa} \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{D}_{\gamma\kappa} \boldsymbol{u}_\kappa, \tag{13}$$

where M_κ is the symmetric semi-definite matrix

$$\mathsf{M}_\kappa = \begin{bmatrix} \mathsf{D}_x^T & \mathsf{D}_y^T \end{bmatrix}_\kappa \begin{bmatrix} \mathsf{H}\mathsf{\Lambda}_{xx} & \mathsf{H}\mathsf{\Lambda}_{xy} \\ \mathsf{H}\mathsf{\Lambda}_{yx} & \mathsf{H}\mathsf{\Lambda}_{yy} \end{bmatrix}_\kappa \begin{bmatrix} \mathsf{D}_x \\ \mathsf{D}_y \end{bmatrix}_\kappa.$$

Remark 2 Identity (13) is the SBP analog of integration by parts in the context of the PDE (1); specifically, it is the discrete form of

$$\int_{\Omega_\kappa} \mathcal{V}\, \nabla \cdot (\Lambda \nabla \mathcal{U})\, \mathrm{d}\Omega = -\int_{\Omega_\kappa} (\nabla \mathcal{V})^T \Lambda (\nabla \mathcal{U})\, \mathrm{d}\Omega + \oint_{\partial\Omega_\kappa} \mathcal{V}\, (\Lambda \nabla \mathcal{U}) \cdot \boldsymbol{n}\, \mathrm{d}\Gamma.$$

This also demonstrates that SBP operators are closely related to mimetic finite-difference methods; see, for example, [25] and the references therein.

To obtain the SBP weak form, we left multiply the strong form (11) by $\boldsymbol{v}_\kappa^T \mathsf{H}_\kappa$ and apply identity (13):

$$\boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \frac{\mathrm{d}\boldsymbol{u}_\kappa}{\mathrm{d}t} = -\boldsymbol{v}_\kappa^T \mathsf{M}_\kappa \boldsymbol{u}_\kappa + \sum_{\gamma \subset \partial\Omega_\kappa} \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{D}_{\gamma\kappa} \boldsymbol{u}_\kappa + \boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \boldsymbol{f}_\kappa. \tag{14}$$

Each term in the above discretization can be related to an integral bilinear form using the approximation properties of the matrices H_κ, $(\mathsf{Q}_x)_\kappa$, and $(\mathsf{Q}_y)_\kappa$ — see (5) and (6) — as well as the accuracy of $\mathsf{R}_{\gamma\kappa}$, B_{γ}, and $\mathsf{D}_{\gamma\kappa}$. Indeed, for sufficiently smooth functions $\mathcal{U}$ and $\mathcal{V}$, (14) implies that

$$\int_{\Omega_\kappa} \mathcal{V}\frac{\partial \mathcal{U}}{\partial t}\,\mathrm{d}\Omega = -\int_{\Omega_\kappa} (\nabla\mathcal{V})^T \Lambda\,(\nabla\mathcal{U})\,\mathrm{d}\Omega \\ + \sum_{\gamma\subset\partial\Omega_\kappa}\int_{\gamma} \mathcal{V}\,(\Lambda\nabla\mathcal{U})\cdot\mathbf{n}\,\mathrm{d}\Gamma + \int_{\Omega_\kappa}\mathcal{V}\mathcal{F}\,\mathrm{d}\Omega + \mathrm{O}(h^{p+1}),$$

In other words, satisfying the discrete weak form (14) implies that the continuous weak form is satisfied to order h^{p+1}.

4.2 Penalty terms enforcing continuity and boundary conditions

Boundary conditions and interelement continuity are enforced weakly by introducing penalty terms on the right-hand side of (11):

$$\frac{\mathrm{d}\boldsymbol{u}_\kappa}{\mathrm{d}t} = \mathsf{D}_\kappa\boldsymbol{u}_\kappa + \boldsymbol{f}_\kappa - \mathsf{H}_\kappa^{-1}\boldsymbol{s}_\kappa^{\mathcal{I}}(\boldsymbol{u}_h) - \mathsf{H}_\kappa^{-1}\boldsymbol{s}_\kappa^{\mathcal{B}}(\boldsymbol{u}_h,\boldsymbol{u}_{\mathcal{D}},\boldsymbol{u}_{\mathcal{N}})\,. \tag{15}$$

The vectors $\boldsymbol{s}_\kappa^{\mathcal{I}}$ and $\boldsymbol{s}_\kappa^{\mathcal{B}}$ are the interface and boundary SAT penalties, respectively, which we define below. Briefly, these penalties involve linear combinations of the (approximate) jumps in the function and its normal derivative across elements and at the boundary. These jumps vanish for sufficiently smooth solutions ensuring that the discretization is consistent.

For element Ω_κ the SAT interface penalties are defined by

$$\boldsymbol{s}_\kappa^{\mathcal{I}}(\boldsymbol{u}_h) = \sum_{\gamma\subset\Gamma_\kappa^{\mathcal{I}}} \begin{bmatrix}\mathsf{R}_{\gamma\kappa}^T & \mathsf{D}_{\gamma\kappa}^T\end{bmatrix} \begin{bmatrix}\mathsf{T}_{\gamma\kappa}^{(1)} & \mathsf{T}_{\gamma\kappa}^{(3)} \\ \mathsf{T}_{\gamma\kappa}^{(2)} & \mathsf{T}_{\gamma\kappa}^{(4)}\end{bmatrix} \begin{bmatrix}\mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa - \mathsf{R}_{\gamma\nu}\boldsymbol{u}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa + \mathsf{D}_{\gamma\nu}\boldsymbol{u}_\nu\end{bmatrix}, \tag{16}$$

where we use ν to denote the generic index of the element sharing face γ with the κth element, i.e., $\gamma = \Omega_\kappa \cap \Omega_\nu$. Note that all the matrix operators defined for Ω_κ are defined analogously for Ω_ν. For example, $\mathsf{R}_{\gamma\nu} \in \mathbb{R}^{n_\gamma\times n_\nu}$ is an interpolation/extrapolation operator from the nodes of Ω_ν to the nodes of γ, and

$$\mathsf{D}_{\gamma\nu} = -\mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\nu}\left(\mathsf{\Lambda}_{xx}\mathsf{D}_x + \mathsf{\Lambda}_{xy}\mathsf{D}_y\right)_\nu - \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\nu}\left(\mathsf{\Lambda}_{yx}\mathsf{D}_x + \mathsf{\Lambda}_{yy}\mathsf{D}_y\right)_\nu\,,$$

is an approximation to the normal derivative at the nodes of γ with respect to Ω_ν. Recall that $\mathsf{N}_{x,\gamma}$ and $\mathsf{N}_{y,\gamma}$ hold the x and y components of $\boldsymbol{n}$ with respect to Ω_κ, so the sign of these matrices must be reversed for $\mathsf{D}_{\gamma\nu}$.

The matrices $\mathsf{T}_{\gamma\kappa}^{(i)} \in \mathbb{R}^{n_\gamma\times n_\gamma}$, $i = 1,2,3,4$ appearing in the definition of $\boldsymbol{s}_\kappa^{\mathcal{I}}$ denote the SAT coefficient matrices for element Ω_κ on face γ. We will assume that these coefficient matrices are symmetric but are otherwise unspecified; *it is the primary objective of the subsequent analysis to determine the constraints on these matrices that lead to adjoint consistency and stability.* Note that $\mathsf{T}_{\gamma\kappa}^{(i)} \neq \mathsf{T}_{\gamma\nu}^{(i)}$ in general; that is, we do not assume *ab initio* that the coefficient matrices of two adjacent elements are necessarily equal.

The SAT boundary penalties are defined by

$$\begin{aligned}\boldsymbol{s}_\kappa^{\mathcal{B}}\left(\boldsymbol{u}_h,\boldsymbol{u}_{\mathcal{D}},\boldsymbol{u}_{\mathcal{N}}\right) = \sum_{\gamma\subset\Gamma_\kappa^{\mathcal{D}}} \begin{bmatrix}\mathsf{R}_{\gamma\kappa}^T & \mathsf{D}_{\gamma\kappa}^T\end{bmatrix}\begin{bmatrix}\mathsf{T}_\gamma^{\mathcal{D}}\\ -\mathsf{B}_\gamma\end{bmatrix}\left(\mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa - \boldsymbol{u}_{\gamma\mathcal{D}}\right)&\\ + \sum_{\gamma\subset\Gamma_\kappa^{\mathcal{N}}} \mathsf{R}_{\gamma\kappa}^T\mathsf{B}_\gamma(\mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa - \boldsymbol{u}_{\gamma\mathcal{N}}),&\end{aligned} \tag{17}$$

where $\Gamma_\kappa^{\mathcal{D}} = \partial\Omega_\kappa \cap \Gamma^{\mathcal{D}}$ and $\Gamma_\kappa^{\mathcal{N}} = \partial\Omega_\kappa \cap \Gamma^{\mathcal{N}}$. The vectors $\boldsymbol{u}_{\gamma\mathcal{D}}$ and $\boldsymbol{u}_{\gamma\mathcal{N}}$ in the boundary penalties denote the functions $\mathcal{U}_{\mathcal{D}}$ and $\mathcal{U}_{\mathcal{N}}$, respectively, evaluated at the cubature nodes of face γ. $\mathsf{T}_\gamma^{\mathcal{D}}$ is a coefficient matrix for the SAT on a Dirichlet boundary face of Ω_κ, and, as with the $\mathsf{T}_{\gamma\kappa}^{(i)}$, it will be constrained by the subsequent analysis.

For those more familiar with finite-element discretizations, it may be helpful to relate the penalties $\boldsymbol{s}_\kappa^{\mathcal{I}}$ and $\boldsymbol{s}_\kappa^{\mathcal{B}}$ to integral forms. To do so, we adopt the standard definitions for the scalar and vector jump operators on interface γ; thus, if $\mathcal{U}_\kappa$ and $\mathcal{U}_\nu$ denote the traces of $\mathcal{U}$ taken from the interior of Ω_κ and Ω_ν, respectively, and $\boldsymbol{n}_\kappa$ and $\boldsymbol{n}_\nu$ denote the outward pointing normals with respective to Ω_κ and Ω_ν, then

$$[\![\mathcal{U}]\!] \equiv \mathcal{U}_\kappa\boldsymbol{n}_\kappa + \mathcal{U}_\nu\boldsymbol{n}_\nu, \qquad \text{and} \qquad [\![\Lambda\nabla\mathcal{U}]\!] \equiv (\Lambda\nabla\mathcal{U}_\kappa)\cdot\boldsymbol{n}_\kappa + (\Lambda\nabla\mathcal{U}_\nu)\cdot\boldsymbol{n}_\nu.$$

To simplify the comparison between the SAT and integral penalties, we will assume that $\mathsf{T}_{\gamma\kappa}^{(i)} = \mathsf{B}_\gamma\tau_\kappa^{(i)}$, where the $\tau_\kappa^{(i)}$, $i=1,2,3,4$, are scalars. However, *this diagonal assumption is not used in the subsequent analysis.*

As we did for the weak form (14), we let $\boldsymbol{v}_\kappa \in \mathbb{R}^{n_\kappa}$ be an arbitrary vector. Then, under the assumption $\mathsf{T}_{\gamma\kappa}^{(i)} = \mathsf{B}_\gamma\tau_\kappa^{(i)}$, the product between $\boldsymbol{v}_\kappa$ and the interface SAT can be interpreted as

$$\boldsymbol{v}_\kappa^T\boldsymbol{s}_\kappa^{\mathcal{I}}\left(\boldsymbol{u}_h\right) \approx \sum_{\gamma\subset\Gamma_\kappa^{\mathcal{I}}}\int_\gamma \begin{bmatrix}\mathcal{V}\\ (\Lambda\nabla\mathcal{V})\cdot\boldsymbol{n}_\kappa\end{bmatrix}^T \begin{bmatrix}\tau_\kappa^{(1)} & \tau_\kappa^{(3)}\\ \tau_\kappa^{(2)} & \tau_\kappa^{(4)}\end{bmatrix}\begin{bmatrix}[\![\mathcal{U}]\!]\cdot\boldsymbol{n}_\kappa\\ [\![\Lambda\nabla\mathcal{U}]\!]\end{bmatrix}\,\mathrm{d}\Gamma,$$

and the product between $\boldsymbol{v}_\kappa$ and the boundary SAT can be interpreted as

$$\begin{aligned}\boldsymbol{v}_\kappa^T\boldsymbol{s}_\kappa^{\mathcal{B}}\left(\boldsymbol{u}_h,\boldsymbol{u}_{\mathcal{D}},\boldsymbol{u}_{\mathcal{N}}\right) \approx \sum_{\gamma\subset\Gamma_\kappa^{\mathcal{D}}}\int_\gamma \begin{bmatrix}\mathcal{V}\\ (\Lambda\nabla\mathcal{V})\cdot\boldsymbol{n}_\kappa\end{bmatrix}^T \begin{bmatrix}\tau^{(\mathcal{D})}\\ -1\end{bmatrix}\left(\mathcal{U}_\kappa - \mathcal{U}_{\mathcal{D}}\right)\,\mathrm{d}\Gamma&\\ + \sum_{\gamma\subset\Gamma_\kappa^{\mathcal{N}}}\int_\gamma \mathcal{V}\left((\Lambda\nabla\mathcal{U}_\kappa)\cdot\boldsymbol{n}_\kappa - \mathcal{U}_{\mathcal{N}}\right)\,\mathrm{d}\Gamma.&\end{aligned}$$

4.3 Face-based weak forms of the discretization

Left multiplying the strong-form discretization (15) by $\boldsymbol{v}_\kappa^T\mathsf{H}_\kappa$ and using (13), we arrive at

$$\begin{aligned}\boldsymbol{v}_\kappa^T\mathsf{H}_\kappa\frac{\mathrm{d}\boldsymbol{u}_\kappa}{\mathrm{d}t} = -\boldsymbol{v}_\kappa^T\mathsf{M}_\kappa\boldsymbol{u}_\kappa + \sum_{\gamma\subset\partial\Omega_\kappa}\boldsymbol{v}_\kappa^T\mathsf{R}_{\gamma\kappa}^T\mathsf{B}_\gamma\mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa + \boldsymbol{v}_\kappa^T\mathsf{H}_\kappa\boldsymbol{f}_\kappa&\\ - \boldsymbol{v}_\kappa^T\boldsymbol{s}_\kappa^{\mathcal{I}}\left(\boldsymbol{u}_h\right) - \boldsymbol{v}_\kappa^T\boldsymbol{s}_\kappa^{\mathcal{B}}\left(\boldsymbol{u}_h,\boldsymbol{u}_{\mathcal{D}},\boldsymbol{u}_{\mathcal{N}}\right).&\end{aligned} \tag{18}$$

The above equation is the element-based weak form of the discretization. For the subsequent analysis, two equivalent face-based weak forms will prove more useful. To obtain the first face-based weak formulation, we sum the element-based weak form over all Ω_κ. After rearrangement, this gives the SBP-SAT version of (4):

$$\sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \frac{\mathrm{d}\boldsymbol{u}_\kappa}{\mathrm{d}t} = R_h(\boldsymbol{u}_h, \boldsymbol{v}_h), \qquad \forall\, \boldsymbol{v}_h \in \mathbb{R}^{\sum n_\kappa},$$

where the spatial residual on the right is defined by

$$\begin{aligned} R_h(\boldsymbol{u}_h, \boldsymbol{v}_h) \equiv &- \sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{v}_\kappa^T \mathsf{M}_\kappa \boldsymbol{u}_\kappa + \sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \boldsymbol{f}_\kappa \\ &- \sum_{\gamma \subset \Gamma^{\mathcal{I}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{v}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{v}_\nu \end{bmatrix}^T \begin{bmatrix} \mathsf{T}_{\gamma\kappa}^{(1)} & -\mathsf{T}_{\gamma\kappa}^{(1)} & \mathsf{T}_{\gamma\kappa}^{(3)} - \mathsf{B}_\gamma & \mathsf{T}_{\gamma\kappa}^{(3)} \\ -\mathsf{T}_{\gamma\nu}^{(1)} & \mathsf{T}_{\gamma\nu}^{(1)} & \mathsf{T}_{\gamma\nu}^{(3)} & \mathsf{T}_{\gamma\nu}^{(3)} - \mathsf{B}_\gamma \\ \mathsf{T}_{\gamma\kappa}^{(2)} & -\mathsf{T}_{\gamma\kappa}^{(2)} & \mathsf{T}_{\gamma\kappa}^{(4)} & \mathsf{T}_{\gamma\kappa}^{(4)} \\ -\mathsf{T}_{\gamma\nu}^{(2)} & \mathsf{T}_{\gamma\nu}^{(2)} & \mathsf{T}_{\gamma\nu}^{(4)} & \mathsf{T}_{\gamma\nu}^{(4)} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{u}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{u}_\nu \end{bmatrix} \\ &- \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \end{bmatrix}^T \begin{bmatrix} \mathsf{T}_\gamma^{\mathcal{D}} & -\mathsf{B}_\gamma \\ -\mathsf{B}_\gamma & 0 \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa - \boldsymbol{u}_{\gamma\mathcal{D}} \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa \end{bmatrix} + \sum_{\gamma \subset \Gamma^{\mathcal{N}}} \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \boldsymbol{u}_{\gamma\mathcal{N}}. \end{aligned} \tag{19}$$

The residual definition (19) will be our starting point for the energy stability analysis in Section 6.

Next, we derive an equivalent face-based residual that will be useful for the adjoint analysis. Swapping the roles of $\boldsymbol{u}_\kappa$ and $\boldsymbol{v}_\kappa$ in the identity (13), and then transposing and rearranging the result, we obtain

$$-\boldsymbol{v}_\kappa^T \mathsf{M}_\kappa \boldsymbol{u}_\kappa = \boldsymbol{v}_\kappa^T \mathsf{D}_\kappa^T \mathsf{H}_\kappa \boldsymbol{u}_\kappa - \sum_{\gamma \subset \partial\Omega_\kappa} \boldsymbol{v}_\kappa^T \mathsf{D}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa,$$

where we have used the symmetry of M_κ. Substituting this expression for $-\boldsymbol{v}_\kappa^T \mathsf{M}_\kappa \boldsymbol{u}_\kappa$ into (19) produces

$$\begin{aligned} R_h(\boldsymbol{u}_h, \boldsymbol{v}_h) \equiv &\sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{v}_\kappa^T \mathsf{D}_\kappa^T \mathsf{H}_\kappa \boldsymbol{u}_\kappa + \sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \boldsymbol{f}_\kappa + \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \boldsymbol{v}_\kappa^T \mathsf{D}_{\gamma\kappa}^T \mathsf{B}_\gamma \boldsymbol{u}_{\gamma\mathcal{D}} \\ &- \sum_{\gamma \subset \Gamma^{\mathcal{I}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{v}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{v}_\nu \end{bmatrix}^T \begin{bmatrix} \mathsf{T}_{\gamma\kappa}^{(1)} & -\mathsf{T}_{\gamma\kappa}^{(1)} & \mathsf{T}_{\gamma\kappa}^{(3)} - \mathsf{B}_\gamma & \mathsf{T}_{\gamma\kappa}^{(3)} \\ -\mathsf{T}_{\gamma\nu}^{(1)} & \mathsf{T}_{\gamma\nu}^{(1)} & \mathsf{T}_{\gamma\nu}^{(3)} & \mathsf{T}_{\gamma\nu}^{(3)} - \mathsf{B}_\gamma \\ \mathsf{T}_{\gamma\kappa}^{(2)} + \mathsf{B}_\gamma & -\mathsf{T}_{\gamma\kappa}^{(2)} & \mathsf{T}_{\gamma\kappa}^{(4)} & \mathsf{T}_{\gamma\kappa}^{(4)} \\ -\mathsf{T}_{\gamma\nu}^{(2)} & \mathsf{T}_{\gamma\nu}^{(2)} + \mathsf{B}_\gamma & \mathsf{T}_{\gamma\nu}^{(4)} & \mathsf{T}_{\gamma\nu}^{(4)} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{u}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{u}_\nu \end{bmatrix} \\ &- \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \left[\mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa\right]^T \left[\; \mathsf{T}_\gamma^{\mathcal{D}} \;\; -\mathsf{B}_\gamma\right] \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa - \boldsymbol{u}_{\gamma\mathcal{D}} \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa \end{bmatrix} + \sum_{\gamma \subset \Gamma^{\mathcal{N}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \end{bmatrix}^T \begin{bmatrix} \mathsf{B}_\gamma \boldsymbol{u}_{\gamma\mathcal{N}} \\ -\mathsf{B}_\gamma \mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa \end{bmatrix}. \end{aligned} \tag{20}$$

5 Adjoint consistency analysis

It is well known in the finite-element community that adjoint, or dual, consistency is necessary for obtaining optimal error rates in the L^2 norm [16]. More generally, adjoint consistency leads to superconvergent (integral) functional estimates [21, 26–31], which can significantly improve the accuracy of outputs like lift and drag when using high-order methods. Given the close connection between SBP finite-difference methods and FE methods, it is perhaps not surprising that tensor-product SBP discretizations also exhibit superconvergent functionals when discretized in a dual consistent manner [32, 33].

For the reasons listed above, adjoint consistency is a property that we would like our multi-dimensional SBP discretizations to satisfy. Therefore, in the following subsections, we investigate the constraints on the SAT penalties in (15) that guarantee adjoint consistency. We begin by briefly reviewing the dual problem associated with the steady version of (1).

5.1 A generic adjoint PDE: the continuous case

The adjoint depends on the primal PDE and a particular functional of interest. For the following adjoint-consistency analysis, we consider the linear functional

$$\mathcal{J}(\mathcal{U}) = \int_{\Omega} \mathcal{G}\mathcal{U} \,\mathrm{d}\Omega + \int_{\Gamma^{\mathcal{N}}} \mathcal{V}_{\mathcal{N}}\mathcal{U} \,\mathrm{d}\Gamma - \int_{\Gamma^{\mathcal{D}}} \mathcal{V}_{\mathcal{D}} \left(\Lambda\nabla\mathcal{U}\right)\cdot\boldsymbol{n}\,\mathrm{d}\Gamma, \tag{21}$$

where $\mathcal{G} \in L^2(\Omega)$, $\mathcal{V}_{\mathcal{D}} \in L^2(\Gamma^{\mathcal{D}})$ and $\mathcal{V}_{\mathcal{N}} \in L^2(\Gamma^{\mathcal{N}})$.

The adjoint, which we will denote by $\mathcal{V}$, is the sensitivity of $\mathcal{J}(\mathcal{U})$ to perturbations in the residual. This definition is made precise by the following variational statement: find $\mathcal{V} \in \mathbb{H}^1(\Omega)$ such that

$$\mathcal{R}^*(\mathcal{V}, \delta\mathcal{U}) \equiv \mathcal{J}'[\mathcal{U}](\delta\mathcal{U}) + \mathcal{R}'[\mathcal{U}](\delta\mathcal{U}, \mathcal{V}) = 0, \qquad \forall\, \delta\mathcal{U} \in \mathbb{H}^1(\Omega), \tag{22}$$

where the prime denotes Fréchet differentiation with respect to the quantity in the brackets. For the linear output and residual under consideration,

$$\mathcal{J}'[\mathcal{U}](\delta\mathcal{U}) = \int_{\Omega} \mathcal{G}\delta\mathcal{U} \,\mathrm{d}\Omega + \int_{\Gamma^{\mathcal{N}}} \mathcal{V}_{\mathcal{N}}\delta\mathcal{U} \,\mathrm{d}\Gamma - \int_{\Gamma^{\mathcal{D}}} \mathcal{V}_{\mathcal{D}} \left(\Lambda\nabla\delta\mathcal{U}\right)\cdot\boldsymbol{n}\,\mathrm{d}\Gamma$$

and

$$\mathcal{R}'[\mathcal{U}](\delta\mathcal{U}, \mathcal{V}) = -\int_{\Omega} \left(\nabla\mathcal{V}\right)^T \Lambda \left(\nabla\delta\mathcal{U}\right)\,\mathrm{d}\Omega + \int_{\Gamma^{\mathcal{D}}} \mathcal{V}\left(\Lambda\nabla\delta\mathcal{U}\right)\cdot\mathbf{n}\,\mathrm{d}\Gamma + \int_{\Gamma^{\mathcal{D}}} \delta\mathcal{U}\left(\Lambda\nabla\mathcal{V}\right)\cdot\mathbf{n}\,\mathrm{d}\Gamma.$$

To obtain the strong-form of the adjoint PDE, we assume that $\mathcal{V}$ is sufficiently smooth and apply integration by parts to the appropriate terms in (22). This produces a sum of integrals over Ω, $\Gamma^{\mathcal{N}}$, and $\Gamma^{\mathcal{D}}$ that vanishes for all $\delta\mathcal{U}$:

$$\mathcal{R}^*(\mathcal{V}, \delta\mathcal{U}) = \int_{\Omega} \delta\mathcal{U}\left[\nabla\cdot\left(\Lambda\nabla\mathcal{V}\right) + \mathcal{G}\right]\,\mathrm{d}\Omega + \int_{\Gamma^{\mathcal{N}}} \delta\mathcal{U}\left[\mathcal{V}_{\mathcal{N}} - \left(\Lambda\nabla\mathcal{V}\right)\cdot\mathbf{n}\right]\,\mathrm{d}\Gamma + \int_{\Gamma^{\mathcal{D}}} \left(\mathcal{V} - \mathcal{V}_{\mathcal{D}}\right)\left(\Lambda\nabla\delta\mathcal{U}\right)\cdot\mathbf{n}\,\mathrm{d}\Gamma = 0. \tag{23}$$

The integrands in the above integrals must be zero, since $\delta\mathcal{U} \in \mathbb{H}^1(\Omega)$ is otherwise arbitrary, and we conclude that the adjoint satisfies the following PDE:

$$\begin{aligned}
&\nabla \cdot (\Lambda \nabla \mathcal{V}) + \mathcal{G} = 0, && \forall\,(x,y) \in \Omega, \\
&\mathcal{V} = \mathcal{V}_{\mathcal{D}}, && \forall\,(x,y) \in \Gamma^{\mathcal{D}}, \\
&(\Lambda \nabla \mathcal{V}) \cdot \boldsymbol{n} = \mathcal{V}_{\mathcal{N}}, && \forall\,(x,y) \in \Gamma^{\mathcal{N}}.
\end{aligned} \tag{24}$$

5.2 Functional discretization and the discrete adjoint equation

We discretize the functional (21) as

$$\begin{aligned}
J_h(\boldsymbol{u}_h) := \sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{g}_\kappa^T \mathsf{H}_\kappa \boldsymbol{u}_\kappa + \sum_{\gamma \subset \Gamma^{\mathcal{N}}} \boldsymbol{v}_{\gamma\mathcal{N}}^T \mathsf{B}_\gamma \mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \boldsymbol{v}_{\gamma\mathcal{D}}^T \mathsf{B}_\gamma \mathsf{D}_{\gamma\kappa} \boldsymbol{u}_\kappa \\
+ \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \boldsymbol{v}_{\gamma\mathcal{D}}^T \mathsf{T}_\gamma^{\mathcal{D}} (\mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \boldsymbol{u}_{\gamma\mathcal{D}}),
\end{aligned} \tag{25}$$

where $\boldsymbol{v}_{\gamma\mathcal{N}}$ and $\boldsymbol{v}_{\gamma\mathcal{D}}$ denote $\mathcal{V}_{\mathcal{N}}$ and $\mathcal{V}_{\mathcal{D}}$, respectively, evaluated at the cubature nodes of the generic face γ, and

$$\boldsymbol{g}_\kappa^T = [\mathcal{G}(x_0)\ \mathcal{G}(x_1)\ \dots\ \mathcal{G}(x_{n_\kappa})]. \tag{26}$$

Remark 3 The first three terms in (25) are direct discretizations of the first three terms in (21). The fourth term in (25) is an order h^{r+1} term; the interpolation/extrapolation operators are exact for degree $r \geq p$ polynomials, so $\mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa = \boldsymbol{u}_{\gamma\mathcal{D}} + \mathrm{O}(h^{r+1})$. This last term in J_h is included for adjoint consistency [21].

To derive the discrete adjoint equation, we will follow a process analogous to the continuous adjoint derivation. To this end, we begin with the SBP-SAT version of (22). Note that in the finite-dimensional case, the Fréchet derivatives are equivalent to the usual derivative, so we have

$$\begin{aligned}
R_h^*(\boldsymbol{v}_h, \boldsymbol{\delta u}_h) &= J_h'[\boldsymbol{u}_h](\boldsymbol{\delta u}_h) + R_h'[\boldsymbol{u}_h](\boldsymbol{\delta u}_h, \boldsymbol{v}_h) \\
&= \boldsymbol{\delta u}_h^T \frac{\partial}{\partial \boldsymbol{u}_h} J_h(\boldsymbol{u}_h) + \boldsymbol{\delta u}_h^T \frac{\partial}{\partial \boldsymbol{u}_h} R_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = 0, \qquad \forall\, \boldsymbol{\delta u}_h \in \mathbb{R}^{\sum n_\kappa}.
\end{aligned}$$

The term involving the partial derivative of J_h is

$$\begin{aligned}
\boldsymbol{\delta u}_h^T \frac{\partial}{\partial \boldsymbol{u}_h} J_h(\boldsymbol{u}_h) = \sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{\delta u}_\kappa^T \mathsf{H}_\kappa \boldsymbol{g}_\kappa + \sum_{\gamma \subset \Gamma^{\mathcal{N}}} \boldsymbol{\delta u}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \boldsymbol{v}_{\gamma\mathcal{N}} \\
+ \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa} \boldsymbol{\delta u}_\kappa \\ \mathsf{D}_{\gamma\kappa} \boldsymbol{\delta u}_\kappa \end{bmatrix}^T \begin{bmatrix} \mathsf{T}_\gamma^{\mathcal{D}} \\ -\mathsf{B}_\gamma \end{bmatrix} \boldsymbol{v}_{\gamma\mathcal{D}}.
\end{aligned}$$

Next, for the term involving the partial derivative of R_h, we make use of (20). We replace $\boldsymbol{u}_\kappa$ with $\boldsymbol{\delta u}_\kappa$, since $\boldsymbol{u}_\kappa$ appears linearly in $\mathcal{R}_h$, eliminate constant terms, and transpose the result:

$$\begin{aligned}
\boldsymbol{\delta u}_h^T \frac{\partial}{\partial \boldsymbol{u}_h} R_h(\boldsymbol{u}_h, \boldsymbol{v}_h) &= \sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{\delta u}_\kappa^T \mathsf{H}_\kappa \mathsf{D}_\kappa \boldsymbol{v}_\kappa \\
&- \sum_{\gamma \subset \Gamma^{\mathcal{I}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{\delta u}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{\delta u}_\nu \end{bmatrix}^T \begin{bmatrix} \mathsf{T}^{(1)}_{\gamma\kappa} & -\mathsf{T}^{(1)}_{\gamma\nu} & \mathsf{T}^{(2)}_{\gamma\kappa} + \mathsf{B}_\gamma & -\mathsf{T}^{(2)}_{\gamma\nu} \\ -\mathsf{T}^{(1)}_{\gamma\kappa} & \mathsf{T}^{(1)}_{\gamma\nu} & -\mathsf{T}^{(2)}_{\gamma\kappa} & \mathsf{T}^{(2)}_{\gamma\nu} + \mathsf{B}_\gamma \\ \mathsf{T}^{(3)}_{\gamma\kappa} - \mathsf{B}_\gamma & \mathsf{T}^{(3)}_{\gamma\nu} & \mathsf{T}^{(4)}_{\gamma\kappa} & \mathsf{T}^{(4)}_{\gamma\nu} \\ \mathsf{T}^{(3)}_{\gamma\kappa} & \mathsf{T}^{(3)}_{\gamma\nu} - \mathsf{B}_\gamma & \mathsf{T}^{(4)}_{\gamma\kappa} & \mathsf{T}^{(4)}_{\gamma\nu} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{v}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{v}_\nu \end{bmatrix} \\
&- \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \end{bmatrix}^T \begin{bmatrix} \mathsf{T}^{\mathcal{D}}_\gamma \\ -\mathsf{B}_\gamma \end{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa - \sum_{\gamma \subset \Gamma^{\mathcal{N}}} \boldsymbol{\delta u}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{D}_{\gamma\kappa} \boldsymbol{v}_\kappa.
\end{aligned}$$

Summing these terms, we obtain the discrete, SBP-SAT version of (23):

$$\begin{aligned}
R_h^*(\boldsymbol{v}_h, \boldsymbol{\delta u}_h) &= \sum_{\Omega_\kappa \in \mathcal{T}_h} \boldsymbol{\delta u}_\kappa^T \mathsf{H}_\kappa \left[\mathsf{D}_\kappa \boldsymbol{v}_\kappa + \boldsymbol{g}_\kappa \right] \\
&- \sum_{\gamma \subset \Gamma^{\mathcal{I}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{\delta u}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{\delta u}_\nu \end{bmatrix}^T \begin{bmatrix} \mathsf{T}^{(1)}_{\gamma\kappa} & -\mathsf{T}^{(1)}_{\gamma\nu} & \mathsf{T}^{(2)}_{\gamma\kappa} + \mathsf{B}_\gamma & -\mathsf{T}^{(2)}_{\gamma\nu} \\ -\mathsf{T}^{(1)}_{\gamma\kappa} & \mathsf{T}^{(1)}_{\gamma\nu} & -\mathsf{T}^{(2)}_{\gamma\kappa} & \mathsf{T}^{(2)}_{\gamma\nu} + \mathsf{B}_\gamma \\ \mathsf{T}^{(3)}_{\gamma\kappa} - \mathsf{B}_\gamma & \mathsf{T}^{(3)}_{\gamma\nu} & \mathsf{T}^{(4)}_{\gamma\kappa} & \mathsf{T}^{(4)}_{\gamma\nu} \\ \mathsf{T}^{(3)}_{\gamma\kappa} & \mathsf{T}^{(3)}_{\gamma\nu} - \mathsf{B}_\gamma & \mathsf{T}^{(4)}_{\gamma\kappa} & \mathsf{T}^{(4)}_{\gamma\nu} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{v}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{v}_\nu \end{bmatrix} \\
&- \sum_{\gamma \subset \Gamma^{\mathcal{D}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{\delta u}_\kappa \end{bmatrix}^T \begin{bmatrix} \mathsf{T}^{\mathcal{D}}_\gamma \\ -\mathsf{B}_\gamma \end{bmatrix} (\mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa - \boldsymbol{v}_{\gamma\mathcal{D}}) - \sum_{\gamma \subset \Gamma^{\mathcal{N}}} \boldsymbol{\delta u}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma (\mathsf{D}_{\gamma\kappa} \boldsymbol{v}_\kappa - \boldsymbol{v}_{\gamma\mathcal{N}}) \\
&= 0
\end{aligned}$$

Since $\boldsymbol{\delta u}_h$ is arbitrary in $R_h^*(\boldsymbol{v}_h, \boldsymbol{\delta u}_h)$, we can set $\boldsymbol{\delta u}_\nu = \boldsymbol{0}$ and $\boldsymbol{\delta u}_\kappa = \boldsymbol{e}_i$, where $\boldsymbol{e}_i$ is the ith column of the $n_\kappa \times n_\kappa$ identity. Making these choices, and multiplying the result by H_κ^{-1}, we obtain the SBP-SAT discretization of the strong form of the adjoint equation on element Ω_κ:

$$\mathsf{D}_\kappa \boldsymbol{v}_\kappa + \boldsymbol{g}_\kappa - \mathsf{H}_\kappa^{-1} (\boldsymbol{s}_\kappa^{\mathcal{I}})^*(\boldsymbol{v}_h) - \mathsf{H}_\kappa^{-1} (\boldsymbol{s}_\kappa^{\mathcal{B}})^*(\boldsymbol{v}_h, \boldsymbol{v}_{\mathcal{D}}, \boldsymbol{v}_{\mathcal{N}}) = \boldsymbol{0}, \tag{27}$$

where the adjoint SAT penalties for the interfaces are

$$\left(\boldsymbol{s}_\kappa^{\mathcal{I}}\right)^* (\boldsymbol{v}_h) = \sum_{\gamma \subset \Gamma^{\mathcal{I}}_\kappa} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}^T & \mathsf{D}_{\gamma\kappa}^T \end{bmatrix} \begin{bmatrix} \mathsf{T}^{(1)}_{\gamma\kappa} & -\mathsf{T}^{(1)}_{\gamma\nu} & \mathsf{T}^{(2)}_{\gamma\kappa} + \mathsf{B}_\gamma & -\mathsf{T}^{(2)}_{\gamma\nu} \\ \mathsf{T}^{(3)}_{\gamma\kappa} - \mathsf{B}_\gamma & \mathsf{T}^{(3)}_{\gamma\nu} & \mathsf{T}^{(4)}_{\gamma\kappa} & \mathsf{T}^{(4)}_{\gamma\nu} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{v}_\nu \\ \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{v}_\nu \end{bmatrix}, \tag{28}$$

and the penalties for the boundaries are

$$\begin{aligned}
\left(\boldsymbol{s}_\kappa^{\mathcal{B}}\right)^* (\boldsymbol{u}_h, \boldsymbol{u}_{\mathcal{D}}, \boldsymbol{u}_{\mathcal{N}}) &= \sum_{\gamma \subset \Gamma^{\mathcal{D}}_\kappa} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}^T & \mathsf{D}_{\gamma\kappa}^T \end{bmatrix} \begin{bmatrix} \mathsf{T}^{\mathcal{D}}_\gamma \\ -\mathsf{B}_\gamma \end{bmatrix} (\mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa - \boldsymbol{v}_{\gamma\mathcal{D}}) \\
&\quad + \sum_{\gamma \subset \Gamma^{\mathcal{N}}_\kappa} \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma (\mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa - \boldsymbol{v}_{\gamma\mathcal{N}}).
\end{aligned}$$

5.3 Adjoint consistency

The question we are interested in answering is, under what conditions is (27) an accurate discretization of (24)? This is answered by the following theorem.

Theorem 1 *The primal discretization* (15) *and functional discretization* (25) *produce an adjoint discretization,* (27)*, that has a truncation error of order h^{p+1} provided the exact adjoint $\mathcal{V}$ is sufficiently smooth on Ω and the SAT penalty matrices satisfy*

$$\begin{aligned} \mathsf{T}^{(1)}_{\gamma\kappa} &= \mathsf{T}^{(1)}_{\gamma\nu}, & \mathsf{T}^{(2)}_{\gamma\kappa} + \mathsf{T}^{(2)}_{\gamma\nu} &= -\mathsf{B}_\gamma, \\ \mathsf{T}^{(3)}_{\gamma\kappa} + \mathsf{T}^{(3)}_{\gamma\nu} &= \mathsf{B}_\gamma, & \mathsf{T}^{(4)}_{\gamma\kappa} &= \mathsf{T}^{(4)}_{\gamma\nu}. \end{aligned} \tag{29}$$

Proof Clearly the sum $\mathsf{D}_\kappa \boldsymbol{v}_\kappa + \boldsymbol{g}_\kappa$ in (27) is an order h^{p+1} discretization of the adjoint PDE in (24). Indeed, D_κ is the same operator used in the primal discretization.

The boundary SAT, $(\boldsymbol{s}^{\mathcal{B}}_\kappa)^*$, introduces an error that is also $\mathrm{O}(h^{p+1})$. To see this, recall that $\mathsf{R}_{\gamma\kappa}$ and $\mathsf{D}_{\gamma\kappa}$ are exact for polynomials of degree p, and $\boldsymbol{v}_{\gamma\mathcal{D}}$ and $\boldsymbol{v}_{\gamma\mathcal{N}}$ are the exact boundary values evaluated at the nodes of γ. Thus, the differences $\mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa - \boldsymbol{v}_{\gamma\mathcal{D}}$ and $\mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa - \boldsymbol{v}_{\gamma\mathcal{N}}$ vanish for polynomial solutions of degree p or less.

To show that the interface SAT is order h^{p+1}, it is sufficient to show that $(\boldsymbol{s}^{\mathcal{I}}_\kappa)^*(\boldsymbol{v}_h) = \mathbf{0}$ for polynomial solutions $\mathcal{V} \in \mathbb{P}_p(\Omega)$. For these polynomials, the interpolation/extrapolation and normal-derivative operators are exact, so the interpolated values on either side of face γ are equal:

$$\mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa = \mathsf{R}_{\gamma\nu}\boldsymbol{v}_\nu \equiv \boldsymbol{v}_\gamma, \qquad \text{and} \qquad \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa = -\mathsf{D}_{\gamma\nu}\boldsymbol{v}_\nu \equiv \boldsymbol{v}'_\gamma.$$

Substituting these identities into the adjoint interface SATs (28) gives

$$(\boldsymbol{s}^{\mathcal{I}}_\kappa)^*(\boldsymbol{v}_h) = \sum_{\gamma \subset \Gamma^{\mathcal{I}}_\kappa} \begin{bmatrix} \mathsf{R}^T_{\gamma\kappa} & \mathsf{D}^T_{\gamma\kappa} \end{bmatrix} \begin{bmatrix} \mathsf{T}^{(1)}_{\gamma\kappa} - \mathsf{T}^{(1)}_{\gamma\nu} & \mathsf{T}^{(2)}_{\gamma\kappa} + \mathsf{T}^{(2)}_{\gamma\nu} + \mathsf{B}_\gamma \\ \mathsf{T}^{(3)}_{\gamma\kappa} + \mathsf{T}^{(3)}_{\gamma\nu} - \mathsf{B}_\gamma & \mathsf{T}^{(4)}_{\gamma\kappa} - \mathsf{T}^{(4)}_{\gamma\nu} \end{bmatrix} \begin{bmatrix} \boldsymbol{v}_\gamma \\ \boldsymbol{v}'_\gamma \end{bmatrix}.$$

The 2×2 block matrix in the above sum vanishes under the conditions (29), and we have adjoint consistency to order h^{p+1}. □

Remark 4 (Conservation) As shown in [16], adjoint consistency implies conservation. Therefore, if the conditions (29) are satisfied the SBP-SAT discretization will be elementwise conservative, in the sense that $\sum_{\Omega_\kappa \in \mathcal{T}'_h} \mathbf{1}^T \mathsf{H}_\kappa \mathrm{d}\boldsymbol{u}_\kappa/\mathrm{d}t$ depends only on the boundary faces of $\mathcal{T}'_h$ when $\boldsymbol{f}_\kappa = \mathbf{0}$, for any subset of elements $\mathcal{T}'_h \subset \mathcal{T}_h$. To see this, take $\boldsymbol{v}_\kappa = \mathbf{1}$ and $\boldsymbol{v}_\nu = \mathbf{1}$ in (20), and make use of $\mathsf{D}_\kappa \mathbf{1} = \mathbf{0}$, $\mathsf{D}_{\gamma\kappa}\mathbf{1} = \mathbf{0}$, $\mathsf{D}_{\gamma\nu}\mathbf{1} = \mathbf{0}$, and $\mathsf{R}_{\gamma\kappa}\mathbf{1} = \mathsf{R}_{\gamma\nu}\mathbf{1} = \mathbf{1}$.

6 Energy analysis

The objective of this section is to further constrain the SAT penalty matrices based on the conditions for discrete energy stability. Analogous conditions can then be used to obtain entropy-stable discretizations of the viscous terms in the Navier-Stokes equations.

6.1 Energy analysis of the discrete homogeneous problem

Before presenting the conditions for energy stability, we first simplify the penalty matrices based on the adjoint consistency conditions (29). In particular, we will drop the dependence of the $\mathsf{T}^{(1)}$ and $\mathsf{T}^{(4)}$ matrices on the elements:

$$\mathsf{T}^{(1)}_{\gamma\kappa} = \mathsf{T}^{(1)}_{\gamma\nu} \equiv \mathsf{T}^{(1)}_{\gamma}, \qquad \text{and} \qquad \mathsf{T}^{(4)}_{\gamma\kappa} = \mathsf{T}^{(4)}_{\gamma\nu} \equiv \mathsf{T}^{(4)}_{\gamma}.$$

In addition, we will also assume that

$$\mathsf{T}^{(3)}_{\gamma\kappa} - \mathsf{T}^{(2)}_{\gamma\kappa} = \mathsf{B}_{\gamma}. \tag{30}$$

This is not strictly required by the adjoint-consistency analysis, but by the desire to make the 4x4 block matrix in (19) symmetric; symmetric discretizations have been shown to improve the accuracy with which discontinuous Galerkin methods approximate the eigenvalues of the Laplacian [34]. Note that (30), together with the conditions of Theorem 1, implies that $\mathsf{T}^{(3)}_{\gamma\kappa} = -\mathsf{T}^{(2)}_{\gamma\nu}$, $\mathsf{T}^{(3)}_{\gamma\nu} = -\mathsf{T}^{(2)}_{\gamma\kappa}$, and $\mathsf{T}^{(3)}_{\gamma\nu} - \mathsf{T}^{(2)}_{\gamma\nu} = \mathsf{B}_{\gamma}$, which, in turn, imply the symmetry of the 4x4 block matrix in (19).

We will need the following lemma for the stability analysis. The purpose of the lemma is to shift the volume terms in the residual R_h to the faces, so that these terms can contribute to the semi-definiteness of the interface terms. This idea generalizes the "borrowing trick" employed in [11] to multidimensional SBP operators.

Lemma 1 *For each face γ of element Ω_κ, let $\alpha_{\gamma\kappa} > 0$ be given such that $\sum_{\gamma\subset\Gamma_\kappa} \alpha_{\gamma\kappa} = 1$. Then the SBP-SAT residual R_h corresponding to the homogeneous version of the initial-boundary-value problem (1)–(3) — that is, with $\mathcal{F} = 0$, $\mathcal{U}_{\mathcal{D}} = 0$, and $\mathcal{U}_{\mathcal{N}} = 0$ — can be written as*

$$\begin{aligned}
&R_h(\boldsymbol{u}_h, \boldsymbol{v}_h) = \\
&- \sum_{\gamma\subset\Gamma^{\mathcal{I}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{v}_\nu \\ \mathsf{F}_{\kappa}\boldsymbol{v}_\kappa \\ \mathsf{F}_{\nu}\boldsymbol{v}_\nu \end{bmatrix}^T \begin{bmatrix} \mathsf{T}^{(1)}_{\gamma} & -\mathsf{T}^{(1)}_{\gamma} & \mathsf{T}^{(2)}_{\gamma\kappa}\mathsf{C}_{\gamma\kappa} & -\mathsf{T}^{(2)}_{\gamma\nu}\mathsf{C}_{\gamma\nu} \\ -\mathsf{T}^{(1)}_{\gamma} & \mathsf{T}^{(1)}_{\gamma} & -\mathsf{T}^{(2)}_{\gamma\kappa}\mathsf{C}_{\gamma\kappa} & \mathsf{T}^{(2)}_{\gamma\nu}\mathsf{C}_{\gamma\nu} \\ \mathsf{C}^T_{\gamma\kappa}\mathsf{T}^{(2)}_{\gamma\kappa} & -\mathsf{C}^T_{\gamma\kappa}\mathsf{T}^{(2)}_{\gamma\kappa} & \alpha_{\gamma\kappa}\mathsf{\Lambda}^*_\kappa & \\ -\mathsf{C}^T_{\gamma\nu}\mathsf{T}^{(2)}_{\gamma\nu} & \mathsf{C}^T_{\gamma\nu}\mathsf{T}^{(2)}_{\gamma\nu} & & \alpha_{\gamma\nu}\mathsf{\Lambda}^*_\nu \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa \\ \mathsf{R}_{\gamma\nu}\boldsymbol{u}_\nu \\ \mathsf{F}_{\kappa}\boldsymbol{u}_\kappa \\ \mathsf{F}_{\nu}\boldsymbol{u}_\nu \end{bmatrix} \\
&\qquad - \sum_{\gamma\subset\Gamma^{\mathcal{I}}} \begin{bmatrix} \mathsf{D}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{v}_\nu \end{bmatrix}^T \begin{bmatrix} \mathsf{T}^{(4)}_{\gamma} & \mathsf{T}^{(4)}_{\gamma} \\ \mathsf{T}^{(4)}_{\gamma} & \mathsf{T}^{(4)}_{\gamma} \end{bmatrix} \begin{bmatrix} \mathsf{D}_{\gamma\kappa}\boldsymbol{u}_\kappa \\ \mathsf{D}_{\gamma\nu}\boldsymbol{u}_\nu \end{bmatrix} \\
&\qquad - \sum_{\gamma\subset\Gamma^{\mathcal{D}}} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{v}_\kappa \\ \mathsf{F}_{\kappa}\boldsymbol{v}_\kappa \end{bmatrix}^T \begin{bmatrix} \mathsf{T}^{\mathcal{D}}_{\gamma} & -\mathsf{B}_{\gamma}\mathsf{C}_{\gamma\kappa} \\ -\mathsf{C}^T_{\gamma\kappa}\mathsf{B}_{\gamma} & \alpha_{\gamma\kappa}\mathsf{\Lambda}^*_\kappa \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}\boldsymbol{u}_\kappa \\ \mathsf{F}_{\kappa}\boldsymbol{u}_\kappa \end{bmatrix},
\end{aligned} \tag{31}$$

where we have introduced the matrices

$$\mathsf{F}_\kappa = \left\{ \begin{bmatrix} \mathsf{\Lambda}_{xx} & \mathsf{\Lambda}_{xy} \\ \mathsf{\Lambda}_{yx} & \mathsf{\Lambda}_{yy} \end{bmatrix} \begin{bmatrix} \mathsf{D}_x \\ \mathsf{D}_y \end{bmatrix} \right\}_\kappa, \qquad \mathsf{C}_{\gamma\kappa} = \left[\mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\kappa} \;\; \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\kappa} \right],$$

$$\mathsf{F}_\nu = \left\{ \begin{bmatrix} \mathsf{\Lambda}_{xx} & \mathsf{\Lambda}_{xy} \\ \mathsf{\Lambda}_{yx} & \mathsf{\Lambda}_{yy} \end{bmatrix} \begin{bmatrix} \mathsf{D}_x \\ \mathsf{D}_y \end{bmatrix} \right\}_\nu, \qquad \mathsf{C}_{\gamma\nu} = - \left[\mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\nu} \;\; \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\nu} \right],$$

and

$$\mathsf{\Lambda}^*_\kappa = \left\{ \begin{bmatrix} \mathsf{\Lambda}_{xx} & \mathsf{\Lambda}_{xy} \\ \mathsf{\Lambda}_{yx} & \mathsf{\Lambda}_{yy} \end{bmatrix}^{-1} \begin{bmatrix} \mathsf{H} & \\ & \mathsf{H} \end{bmatrix} \right\}_\kappa, \qquad \mathsf{\Lambda}^*_\nu = \left\{ \begin{bmatrix} \mathsf{\Lambda}_{xx} & \mathsf{\Lambda}_{xy} \\ \mathsf{\Lambda}_{yx} & \mathsf{\Lambda}_{yy} \end{bmatrix}^{-1} \begin{bmatrix} \mathsf{H} & \\ & \mathsf{H} \end{bmatrix} \right\}_\nu$$

Proof The full proof follows from straightforward algebra and is omitted; however, we will highlight two observations that make the connection between (31) and (19) clearer. First, we note that

$$\mathsf{C}_{\gamma\kappa}\mathsf{F}_{\kappa} = \mathsf{D}_{\gamma\kappa} \qquad \text{and} \qquad \mathsf{C}_{\gamma\nu}\mathsf{F}_{\nu} = \mathsf{D}_{\gamma\nu}.$$

Second, the elemental matrix M_{κ} can be decomposed as

$$\mathsf{M}_{\kappa} = \begin{bmatrix} \mathsf{D}_x \\ \mathsf{D}_y \end{bmatrix}_{\kappa}^{T} \begin{bmatrix} \mathsf{H}\Lambda_{xx} & \mathsf{H}\Lambda_{xy} \\ \mathsf{H}\Lambda_{yx} & \mathsf{H}\Lambda_{yy} \end{bmatrix}_{\kappa} \begin{bmatrix} \mathsf{D}_x \\ \mathsf{D}_y \end{bmatrix}_{\kappa} = \sum_{\gamma\subset\Gamma_\kappa} \alpha_{\gamma\kappa}\mathsf{F}_{\kappa}^{T}\Lambda_{\kappa}^{*}\mathsf{F}_{\kappa}.$$

□

We will now state and prove the main energy-stability result.

Theorem 2 *Consider the homogeneous version of (1)–(3), where $\mathcal{F}=0, \mathcal{U}_{\mathcal{D}}=0$, and $\mathcal{U}_{\mathcal{N}}=0$. The SBP-SAT discretization of this initial-boundary-value problem has a non-increasing solution norm, with respect to the H matrix, provided*

$$\mathsf{T}_{\gamma}^{(1)} - \mathsf{T}_{\gamma\kappa}^{(2)}\mathsf{C}_{\gamma\kappa}\left(\alpha_{\gamma\kappa}\Lambda_{\kappa}^{*}\right)^{-1}\mathsf{C}_{\gamma\kappa}^{T}\mathsf{T}_{\gamma\kappa}^{(2)} - \mathsf{T}_{\gamma\nu}^{(2)}\mathsf{C}_{\gamma\nu}\left(\alpha_{\gamma\nu}\Lambda_{\nu}^{*}\right)^{-1}\mathsf{C}_{\gamma\nu}^{T}\mathsf{T}_{\gamma\nu}^{(2)} \succeq 0, \tag{32}$$

$$\mathsf{T}_{\gamma}^{\mathcal{D}} - \mathsf{B}_{\gamma}\mathsf{C}_{\gamma\kappa}\left(\alpha_{\gamma\kappa}\Lambda_{\kappa}^{*}\right)^{-1}\mathsf{C}_{\gamma\kappa}^{T}\mathsf{B}_{\gamma} \succeq 0, \tag{33}$$

$$\textit{and} \qquad \mathsf{T}_{\gamma}^{(4)} \succeq 0, \tag{34}$$

where $\mathsf{A}\succeq 0$ indicates that A is positive semi-definite.

Proof The SBP-SAT discretization of the homogeneous equation is given by

$$\sum_{\kappa\in\mathcal{T}_h} \boldsymbol{v}_{\kappa}^{T}\mathsf{H}_{\kappa}\frac{\mathrm{d}\boldsymbol{w}_{\kappa}}{\mathrm{d}t} = R_h(\boldsymbol{w}_h, \boldsymbol{v}_h),$$

where $R_h(\boldsymbol{w}_h, \boldsymbol{v}_h)$ is defined in (31). If we can show that $R_h(\boldsymbol{w}_h, \boldsymbol{w}_h) \leq 0$ for all $\boldsymbol{w}_h$, then we will have $\sum_{\kappa\in\mathcal{T}_h} \boldsymbol{w}_{\kappa}^{T}\mathsf{H}_{\kappa}\mathrm{d}\boldsymbol{w}_{\kappa}/\mathrm{d}t \leq 0$ and the desired result will follow.

The scalar $R_h(\boldsymbol{w}_h, \boldsymbol{w}_h)$ is nonpositive if the symmetric matrices in the three sums of (31) are positive semi-definite. We begin by considering the matrix that appears in the sum over the faces of the Dirichlet boundary:

$$\begin{bmatrix} \mathsf{T}_{\gamma}^{\mathcal{D}} & -\mathsf{B}_{\gamma}\mathsf{C}_{\gamma\kappa} \\ -\mathsf{C}_{\gamma\kappa}^{T}\mathsf{B}_{\gamma} & \alpha_{\gamma\kappa}\Lambda_{\kappa}^{*} \end{bmatrix} \succeq 0.$$

Since, $\alpha_{\gamma\kappa}\Lambda_{\kappa}^{*}$ is positive definite, the above matrix is positive semi-definite if the associated Schur complement is positive semi-definite:

$$\mathsf{T}_{\gamma}^{\mathcal{D}} - \mathsf{B}_{\gamma}\mathsf{C}_{\gamma\kappa}\left(\alpha_{\gamma\kappa}\Lambda_{\kappa}^{*}\right)^{-1}\mathsf{C}_{\gamma\kappa}^{T}\mathsf{B}_{\gamma} \succeq 0,$$

which is precisely the condition (33).

Next, consider the matrix involving $\mathsf{T}_{\gamma}^{(4)}$ in (31):

$$\begin{bmatrix} \mathsf{T}_{\gamma}^{(4)} & \mathsf{T}_{\gamma}^{(4)} \\ \mathsf{T}_{\gamma}^{(4)} & \mathsf{T}_{\gamma}^{(4)} \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \otimes \mathsf{T}_{\gamma}^{(4)},$$

where $\otimes$ denotes the Kronecker product. Since the eigenvalues of $\left[\begin{smallmatrix} 1 & 1 \\ 1 & 1 \end{smallmatrix}\right]$ are zero and two, it follows from the spectral theory of Kronecker products that the eigenvalues of the above matrix are twice the eigenvalues of $\mathsf{T}_{\gamma}^{(4)}$ and n_{γ} zeros. Thus, we require that $\mathsf{T}_{\gamma}^{(4)} \succeq \mathsf{0}$.

Finally, we analyze the matrix containing $\mathsf{T}_{\gamma}^{(1)}$. Similar to the matrix in the boundary-face sum, we make use of the fact that $\alpha_{\gamma\kappa}\mathsf{\Lambda}_{\kappa}^{*}$ and $\alpha_{\gamma\nu}\mathsf{\Lambda}_{\nu}^{*}$ are positive definite to conclude that the 4×4 block matrix is positive semi-definite if the Schur complement is also positive semi-definite, i.e.

$$\begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \otimes \left\{ \mathsf{T}_{\gamma}^{(1)} - \begin{bmatrix} \mathsf{T}_{\gamma\kappa}^{(2)}\mathsf{C}_{\gamma\kappa} & \mathsf{T}_{\gamma\nu}^{(2)}\mathsf{C}_{\gamma\nu} \end{bmatrix} \begin{bmatrix} (\alpha_{\gamma\kappa}\mathsf{\Lambda}_{\kappa}^{*})^{-1} & \\ & (\alpha_{\gamma\nu}\mathsf{\Lambda}_{\nu}^{*})^{-1} \end{bmatrix} \begin{bmatrix} \mathsf{C}_{\gamma\kappa}^{T}\mathsf{T}_{\gamma\kappa}^{(2)} \\ \mathsf{C}_{\gamma\nu}^{T}\mathsf{T}_{\gamma\nu}^{(2)} \end{bmatrix} \right\} \succeq \mathsf{0}.$$

The eigenvalues of $\left[\begin{smallmatrix} 1 & -1 \\ -1 & 1 \end{smallmatrix}\right]$ are zero and two; thus, to ensure that the above Kronecker product is positive semi-definite, we must require that

$$\mathsf{T}_{\gamma}^{(1)} - \mathsf{T}_{\gamma\kappa}^{(2)}\mathsf{C}_{\gamma\kappa}\left(\alpha_{\gamma\kappa}\mathsf{\Lambda}_{\kappa}^{*}\right)^{-1}\mathsf{C}_{\gamma\kappa}^{T}\mathsf{T}_{\gamma\kappa}^{(2)} - \mathsf{T}_{\gamma\nu}^{(2)}\mathsf{C}_{\gamma\nu}\left(\alpha_{\gamma\nu}\mathsf{\Lambda}_{\nu}^{*}\right)^{-1}\mathsf{C}_{\gamma\nu}^{T}\mathsf{T}_{\gamma\nu}^{(2)} \succeq \mathsf{0},$$

which is condition (32). □

7 Generalization of existing methods

In Sections 5 and 6 we obtained sufficient conditions that allow us to construct different schemes with adjoint consistency and energy stability. In this section we show that these conditions can be used to recover two popular interior penalty methods used in FE methods, namely, the modified scheme of Bassi and Rebay (BR2) [19] and the symmetric interior penalty method (SIPG) [21, 35].

While the stability conditions of Theorem 2 depend on $\mathsf{T}_{\gamma\kappa}^{(2)}$ and $\mathsf{T}_{\gamma\nu}^{(2)}$, there remains considerable flexibility in the values adopted for these matrices, provided they satisfy (30) and the conditions in Theorem 1. Additionally, although a positive semi-definite $\mathsf{T}_{\gamma\kappa}^{(4)}$ may influence the accuracy and continuity of solutions, it is not necessary nor is it sufficient to guarantee coercivity of the bilinear form. Accordingly, a simple and effective choice for the penalty matrices is

$$\begin{aligned} \mathsf{T}_{\gamma\kappa}^{(3)} &= -\mathsf{T}_{\gamma\kappa}^{(2)} = \frac{1}{2}\mathsf{B}_{\gamma}, \\ \mathsf{T}_{\gamma\kappa}^{(4)} &= \mathsf{T}_{\gamma\nu}^{(4)} = 0, \end{aligned}$$

which are the values used for the remainder of the paper. Note that other choices are possible that lead to asymmetric or one-sided schemes, such as the compact discontinuous Galerkin scheme [36], but these are not considered in this paper.

We now investigate two specific expressions for $\mathsf{T}_{\gamma}^{(1)}$ and $\mathsf{T}_{\gamma}^{\mathcal{D}}$ and show how these are related to BR2 and SIPG.

7.1 The modified scheme of Bassi and Rebay (BR2)

Based on the stability analysis in Section 6, specifically Theorem 2, a straightforward choice for the SAT penalties is

$$\mathsf{T}_{\gamma}^{(1)} = \frac{1}{4}\mathsf{B}_{\gamma}\left[\mathsf{C}_{\gamma\kappa}\left(\alpha_{\gamma\kappa}\mathsf{\Lambda}_{\kappa}^{*}\right)^{-1}\mathsf{C}_{\gamma\kappa}^{T} + \mathsf{C}_{\gamma\nu}\left(\alpha_{\gamma\nu}\mathsf{\Lambda}_{\nu}^{*}\right)^{-1}\mathsf{C}_{\gamma\nu}^{T}\right]\mathsf{B}_{\gamma}, \tag{35}$$

$$\mathsf{T}_{\gamma}^{\mathcal{D}} = \mathsf{B}_{\gamma}\mathsf{C}_{\gamma\kappa}\left(\alpha_{\gamma\kappa}\mathsf{\Lambda}_{\kappa}^{*}\right)^{-1}\mathsf{C}_{\gamma\kappa}^{T}\mathsf{B}_{\gamma}. \tag{36}$$

We now show that the above penalty matrices generalize the modified scheme of Bassi and Rebay [19] to multidimensional SBP discretizations. For ease of exposition, we will consider the scalar constant-coefficient diffusion case, that is

$$\mathit{\Lambda} = \begin{bmatrix}\lambda_{xx} & \lambda_{xy} \\ \lambda_{yx} & \lambda_{yy}\end{bmatrix} = \lambda \begin{bmatrix}1 & \\ & 1\end{bmatrix}.$$

A similar analysis with a spatially varying tensor $\mathit{\Lambda}$ gives the same conclusion. In addition, we will only focus on the interface penalty of BR2, since the relationship to $\mathsf{T}_\gamma^{\mathcal{D}}$ is similar.

The penalties in the BR2 method that correspond with the matrix $\mathsf{T}_\gamma^{(1)}$ are of the form

$$C_{\mathrm{BR2}} \int_{\partial \mathit{\Omega}_\kappa \cap \mathit{\Gamma}^{\mathcal{I}}} \mathcal{V}_\kappa \frac{1}{2} \left[n_x \left(\mathcal{L}_{x,\kappa}^\gamma + \mathcal{L}_{x,\nu}^\gamma \right) + n_y \left(\mathcal{L}_{y,\kappa}^\gamma + \mathcal{L}_{y,\nu}^\gamma \right) \right] \mathrm{d}\mathit{\Gamma}, \tag{37}$$

where C_{BR2} is a positive constant. Here, the *scalar* lifting operators, $\mathcal{L}_{x,\kappa}^\gamma$ and $\mathcal{L}_{y,\kappa}^\gamma$, are defined by the variational statements

$$\begin{aligned} & \int_{\mathit{\Omega}_\kappa} \mathcal{V}_\kappa \mathcal{L}_{x,\kappa}^\gamma \,\mathrm{d}\mathit{\Omega} = \frac{1}{2} \int_\gamma \mathcal{V}_\kappa \lambda (\mathcal{U}_\kappa - \mathcal{U}_\nu) n_x \,\mathrm{d}\mathit{\Gamma}, \qquad \forall \mathcal{V}_\kappa \in \mathbb{P}_p(\mathit{\Omega}_\kappa), \\ \text{and} \quad & \int_{\mathit{\Omega}_\kappa} \mathcal{V}_\kappa \mathcal{L}_{y,\kappa}^\gamma \,\mathrm{d}\mathit{\Omega} = \frac{1}{2} \int_\gamma \mathcal{V}_\kappa \lambda (\mathcal{U}_\kappa - \mathcal{U}_\nu) n_y \,\mathrm{d}\mathit{\Gamma}, \qquad \forall \mathcal{V}_\kappa \in \mathbb{P}_p(\mathit{\Omega}_\kappa), \end{aligned}$$

where $\mathcal{U}_\kappa$ and $\mathcal{U}_\nu$ denote the finite-dimensional solution on the elements $\mathit{\Omega}_\kappa$ and $\mathit{\Omega}_\nu$, respectively, and $\mathcal{V}_\kappa$ denotes the test function on $\mathit{\Omega}_\kappa$.

The multidimensional SBP discretizations of the lifting-operator variational statements are

$$\begin{aligned} & \boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \boldsymbol{l}_{x,\kappa}^\gamma = \frac{\lambda}{2} \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{N}_{x,\gamma} (\mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \mathsf{R}_{\gamma\nu} \boldsymbol{u}_\nu), \qquad \forall\, \boldsymbol{v}_\kappa \in \mathbb{R}^{n_\kappa}, \\ \text{and} \quad & \boldsymbol{v}_\kappa^T \mathsf{H}_\kappa \boldsymbol{l}_{y,\kappa}^\gamma = \frac{\lambda}{2} \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{N}_{y,\gamma} (\mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \mathsf{R}_{\gamma\nu} \boldsymbol{u}_\nu), \qquad \forall\, \boldsymbol{v}_\kappa \in \mathbb{R}^{n_\kappa}, \end{aligned}$$

where $\boldsymbol{l}_{x,\kappa}^\gamma \in \mathbb{R}^{n_\kappa}$ and $\boldsymbol{l}_{y,\kappa}^\gamma \in \mathbb{R}^{n_\kappa}$ are the discrete lifting operators. Choosing $\boldsymbol{v}_\kappa$ appropriately (i.e. as elements of the identity matrix), we obtain the explicit expressions

$$\begin{aligned} & \boldsymbol{l}_{x,\kappa}^\gamma = \frac{\lambda}{2} \mathsf{H}_\kappa^{-1} \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{N}_{x,\gamma} (\mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \mathsf{R}_{\gamma\nu} \boldsymbol{u}_\nu), \\ \text{and} \quad & \boldsymbol{l}_{y,\kappa}^\gamma = \frac{\lambda}{2} \mathsf{H}_\kappa^{-1} \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{N}_{y,\gamma} (\mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \mathsf{R}_{\gamma\nu} \boldsymbol{u}_\nu). \end{aligned}$$

Next, we turn to the SBP discretization of the BR2 penalty (37). Using the above expressions for $\boldsymbol{l}_{x,\kappa}^\gamma$ and $\boldsymbol{l}_{y,\kappa}^\gamma$, and the analogous ones for $\boldsymbol{l}_{x,\nu}^\gamma$ and $\boldsymbol{l}_{y,\nu}^\gamma$, we obtain the discretization

$$\begin{aligned} & \frac{C_{\mathrm{BR2}}}{2} \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \left[\mathsf{N}_{x,\gamma} (\mathsf{R}_{\gamma\kappa} \boldsymbol{l}_{x,\kappa}^\gamma + \mathsf{R}_{\gamma\nu} \boldsymbol{l}_{x,\nu}^\gamma) + \mathsf{N}_{y,\gamma} (\mathsf{R}_{\gamma\kappa} \boldsymbol{l}_{y,\kappa}^\gamma + \mathsf{R}_{\gamma\nu} \boldsymbol{l}_{y,\nu}^\gamma) \right] \\ & = \frac{C_{\mathrm{BR2}}}{4} \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \lambda \Big[(\mathsf{N}_{x,\gamma} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-1} \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{N}_{x,\gamma} + \mathsf{N}_{y,\gamma} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-1} \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma \mathsf{N}_{y,\gamma}) \\ & \quad + (\mathsf{N}_{x,\gamma} \mathsf{R}_{\gamma\nu} \mathsf{H}_\nu^{-1} \mathsf{R}_{\gamma\nu}^T \mathsf{B}_\gamma \mathsf{N}_{x,\gamma} + \mathsf{N}_{y,\gamma} \mathsf{R}_{\gamma\nu} \mathsf{H}_\nu^{-1} \mathsf{R}_{\gamma\nu}^T \mathsf{B}_\gamma \mathsf{N}_{y,\gamma}) \Big] (\mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \mathsf{R}_{\gamma\nu} \boldsymbol{u}_\nu) \\ & = \boldsymbol{v}_\kappa^T \mathsf{R}_{\gamma\kappa}^T \mathsf{T}_{\mathrm{BR2}} (\mathsf{R}_{\gamma\kappa} \boldsymbol{u}_\kappa - \mathsf{R}_{\gamma\nu} \boldsymbol{u}_\nu), \end{aligned}$$

where

$$\begin{aligned}
\mathsf{T}_{\mathrm{BR2}} &= \frac{C_{\mathrm{BR2}}}{4}\mathsf{B}_\gamma \lambda \Big[(\mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\kappa}\mathsf{H}_\kappa^{-1}\mathsf{R}_{\gamma\kappa}^T\mathsf{B}_\gamma\mathsf{N}_{x,\gamma} + \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\kappa}\mathsf{H}_\kappa^{-1}\mathsf{R}_{\gamma\kappa}^T\mathsf{B}_\gamma\mathsf{N}_{y,\gamma}) \\
&\qquad + (\mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\nu}\mathsf{H}_\nu^{-1}\mathsf{R}_{\gamma\nu}^T\mathsf{B}_\gamma\mathsf{N}_{x,\gamma} + \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\nu}\mathsf{H}_\nu^{-1}\mathsf{R}_{\gamma\nu}^T\mathsf{B}_\gamma\mathsf{N}_{y,\gamma}) \Big] \\
&= \frac{C_{\mathrm{BR2}}}{4}\mathsf{B}_\gamma \Bigg\{ \begin{bmatrix} \mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\kappa} & \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\kappa} \end{bmatrix} \begin{bmatrix} \lambda\mathsf{H}_\kappa^{-1} & \\ & \lambda\mathsf{H}_\kappa^{-1} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}^T\mathsf{N}_{x,\gamma} \\ \mathsf{R}_{\gamma\kappa}^T\mathsf{N}_{y,\gamma} \end{bmatrix} \\
&\qquad + \begin{bmatrix} \mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\nu} & \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\nu} \end{bmatrix} \begin{bmatrix} \lambda\mathsf{H}_\nu^{-1} & \\ & \lambda\mathsf{H}_\nu^{-1} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\nu}^T\mathsf{N}_{x,\gamma} \\ \mathsf{R}_{\gamma\nu}^T\mathsf{N}_{y,\gamma} \end{bmatrix} \Bigg\} \mathsf{B}_\gamma \\
&= \frac{C_{\mathrm{BR2}}}{4}\mathsf{B}_\gamma \Big[\mathsf{C}_{\gamma\kappa}(\Lambda_\kappa^*)^{-1}\mathsf{C}_{\gamma\kappa}^T + \mathsf{C}_{\gamma\nu}(\Lambda_\nu^*)^{-1}\mathsf{C}_{\gamma\nu}^T \Big] \mathsf{B}_\gamma.
\end{aligned}$$

In the above derivation, we reversed the direction of $\mathsf{N}_{x,\gamma}$ and $\mathsf{N}_{y,\gamma}$ for element Ω_ν, and we used the fact that diagonal matrices commute to express $\mathsf{B}_\gamma\mathsf{N}_{x,\gamma} = \mathsf{N}_{x,\gamma}\mathsf{B}_\gamma$ and $\mathsf{B}_\gamma\mathsf{N}_{y,\gamma} = \mathsf{N}_{y,\gamma}\mathsf{B}_\gamma$.

From the above expression for $\mathsf{T}_{\mathrm{BR2}}$, we see that (35) is indeed the SBP generalization of the BR2 penalty (37) with $C_{\mathrm{BR2}} = \alpha_{\gamma\kappa}^{-1}$. We will henceforth refer to this scheme as SAT-BR2.

Remark 5 To the best of our knowledge, this is the first time the SBP-SAT generalization of the BR2 scheme has been presented. This is significant, because it provides a means of implementing the popular BR2 scheme with multidimensional SBP operators that do not have underlying basis functions.

7.2 The symmetric interior penalty method (SIPG)

A disadvantage of the SAT-BR2 penalties is that their $\mathsf{T}_\gamma^{(1)}$ and $\mathsf{T}_\gamma^{\mathcal{D}}$ matrices can be computationally expensive to evaluate. This is not an issue for linear problems — since these matrices can be precomputed and stored if sufficient memory is available — but it can be an issue in nonlinear problems when the diffusion coefficient(s) depend on the state.

In contrast to dense penalty matrices, the symmetric interior penalty method (SIPG) [21,35,37] uses diagonal (or block diagonal for systems) $\mathsf{T}_\gamma^{(1)}$ and $\mathsf{T}_\gamma^{\mathcal{D}}$ with a single parameter that is chosen to be sufficiently large to ensure stability. In this section, we demonstrate how the multidimensional SBP-SAT generalization of SIPG can be derived from the conditions in Theorem 2. First, we need the following lemma.

Lemma 2 *Let $(\lambda_{\max})_\kappa$ be the largest eigenvalue of $\begin{bmatrix} \Lambda_{xx} & \Lambda_{xy} \\ \Lambda_{yx} & \Lambda_{yy} \end{bmatrix}_\kappa$ and let $\|\mathsf{A}\|_2 = \sqrt{\rho(\mathsf{A}\mathsf{A}^T)}$ denote the matrix 2-norm. Then*

$$\mathsf{B}_\gamma\mathsf{C}_{\gamma\kappa}\left(\alpha_{\gamma\kappa}\Lambda_\kappa^*\right)^{-1}\mathsf{C}_{\gamma\kappa}^T\mathsf{B}_\gamma \preceq \frac{(\lambda_{\max})_\kappa \|\mathsf{B}_\gamma^{\frac{1}{2}}\mathsf{R}_{\gamma\kappa}\mathsf{H}_\kappa^{-\frac{1}{2}}\|_2^2}{\alpha_{\gamma\kappa}}\mathsf{B}_\gamma. \tag{38}$$

Proof We recall a few facts that will be useful. The matrices B_γ, $\mathsf{N}_{x,\gamma}$, and $\mathsf{N}_{y,\gamma}$ are diagonal; therefore, they commute with one another. Furthermore, the diagonal matrix B_γ holds positive cubature weights on its diagonal, so it can be factored as $\mathsf{B}_\gamma = \mathsf{B}_\gamma^{\frac{1}{2}}\mathsf{B}_\gamma^{\frac{1}{2}}$. For similar reasons we can write $\mathsf{H}_\kappa = \mathsf{H}_\kappa^{\frac{1}{2}}\mathsf{H}_\kappa^{\frac{1}{2}}$. Finally, $\mathsf{N}_{x,\gamma}\mathsf{N}_{x,\gamma}^T + \mathsf{N}_{y,\gamma}\mathsf{N}_{y,\gamma}^T = \mathsf{I}$, since $\mathsf{N}_{x,\gamma}$ and $\mathsf{N}_{y,\gamma}$ hold the x and y components of the unit normal along γ.

Now, let $\boldsymbol{u}_\gamma \in \mathbb{R}^{n_\gamma}$ be an arbitrary solution on the nodes of the face γ. Then products with $\boldsymbol{u}_\gamma$ and the matrix on the left of (38) can be bounded as follows:

$$\begin{aligned}
&\boldsymbol{u}_\gamma^T \mathsf{B}_\gamma \mathsf{C}_{\gamma\kappa} \left(\alpha_{\gamma\kappa}\mathsf{\Lambda}_\kappa^*\right)^{-1} \mathsf{C}_{\gamma\kappa}^T \mathsf{B}_\gamma \boldsymbol{u}_\gamma \\
&= \frac{1}{\alpha_{\gamma\kappa}} \boldsymbol{u}_\gamma^T \mathsf{B}_\gamma \begin{bmatrix} \mathsf{N}_{x,\gamma}\mathsf{R}_{\gamma\kappa} & \mathsf{N}_{y,\gamma}\mathsf{R}_{\gamma\kappa} \end{bmatrix} \begin{bmatrix} \mathsf{H}_\kappa^{-1} & \\ & \mathsf{H}_\kappa^{-1} \end{bmatrix} \begin{bmatrix} \mathsf{\Lambda}_{xx} & \mathsf{\Lambda}_{xy} \\ \mathsf{\Lambda}_{yx} & \mathsf{\Lambda}_{yy} \end{bmatrix} \begin{bmatrix} \mathsf{R}_{\gamma\kappa}^T \mathsf{N}_{x,\gamma}^T \\ \mathsf{R}_{\gamma\kappa}^T \mathsf{N}_{y,\gamma}^T \end{bmatrix} \mathsf{B}_\gamma \boldsymbol{u}_\gamma \\
&\leq \frac{(\lambda_{\max})_\kappa}{\alpha_{\gamma\kappa}} \boldsymbol{u}_\gamma^T \mathsf{B}_\gamma^{\frac{1}{2}} \begin{bmatrix} \mathsf{N}_{x,\gamma} & \mathsf{N}_{y,\gamma} \end{bmatrix} \left\{ \begin{bmatrix} 1 & \\ & 1 \end{bmatrix} \otimes \left(\mathsf{B}_\gamma^{\frac{1}{2}} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-1} \mathsf{R}_{\gamma\kappa}^T \mathsf{B}_\gamma^{\frac{1}{2}} \right) \right\} \begin{bmatrix} \mathsf{N}_{x,\gamma}^T \\ \mathsf{N}_{y,\gamma}^T \end{bmatrix} \mathsf{B}_\gamma^{\frac{1}{2}} \boldsymbol{u}_\gamma \\
&\leq \frac{(\lambda_{\max})_\kappa \|\mathsf{B}_\gamma^{\frac{1}{2}} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-\frac{1}{2}}\|_2^2}{\alpha_{\gamma\kappa}} \boldsymbol{u}_\gamma^T \mathsf{B}_\gamma^{\frac{1}{2}} \begin{bmatrix} \mathsf{N}_{x,\gamma} & \mathsf{N}_{y,\gamma} \end{bmatrix} \begin{bmatrix} \mathsf{N}_{x,\gamma}^T \\ \mathsf{N}_{y,\gamma}^T \end{bmatrix} \mathsf{B}_\gamma^{\frac{1}{2}} \boldsymbol{u}_\gamma \\
&\leq \frac{(\lambda_{\max})_\kappa \|\mathsf{B}_\gamma^{\frac{1}{2}} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-\frac{1}{2}}\|_2^2}{\alpha_{\gamma\kappa}} \boldsymbol{u}_\gamma^T \mathsf{B}_\gamma \boldsymbol{u}_\gamma .
\end{aligned}$$

The desired result follows from the above inequality, since $\boldsymbol{u}_\gamma$ is arbitrary. □

We can now state the SAT-SIPG penalties that lead to energy stability.

Theorem 3 *The discretization* (15) *is energy stable if*

$$\mathsf{T}_\gamma^{(1)} = \delta_\gamma^{(1)} \mathsf{B}_\gamma, \qquad \textit{and} \qquad \mathsf{T}_\gamma^{\mathcal{D}} = \delta_\gamma^{\mathcal{D}} \mathsf{B}_\gamma, \tag{39}$$

where

$$\begin{aligned}
\delta_\gamma^{(1)} &= \frac{(\lambda_{\max})_\kappa \|\mathsf{B}_\gamma^{\frac{1}{2}} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-\frac{1}{2}}\|_2^2}{4\alpha_{\gamma\kappa}} + \frac{(\lambda_{\max})_\nu \|\mathsf{B}_\gamma^{\frac{1}{2}} \mathsf{R}_{\gamma\nu} \mathsf{H}_\nu^{-\frac{1}{2}}\|_2^2}{4\alpha_{\gamma\nu}}, \\
\delta_\gamma^{\mathcal{D}} &= \frac{(\lambda_{\max})_\kappa \|\mathsf{B}_\gamma^{\frac{1}{2}} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-\frac{1}{2}}\|_2^2}{\alpha_{\gamma\kappa}}.
\end{aligned}$$

Proof The proof follows from Lemma 2, the conditions in Theorem 2, and the aforementioned choice $\mathsf{T}_{\gamma\kappa}^{(3)} = -\mathsf{T}_{\gamma\kappa}^{(2)} = \frac{1}{2}\mathsf{B}_\gamma$. □

Remark 6 The approximations that lead to SAT-SIPG produce a more conservative bound, on the one hand, but a cheaper penalty (than SAT-BR2), on the other hand. However, this assumes that we can precompute $(\lambda_{\max})_\kappa$ on each element. For nonlinear problems, we recommend replacing this value with an estimate for the upper bound of the spectral radius of the tensor $\mathit{\Lambda}$ over all nodes of κ; otherwise the computational advantage of SAT-SIPG over SAT-BR2 will be compromised.

Remark 7 To the best of our knowledge, this is the first time the SIPG penalty has been related to BR2 using straightforward matrix analysis.

The SIPG penalty parameters $\delta_\gamma^{(1)}$ and $\delta_\gamma^{\mathcal{D}}$ are similar to those given by Shahbazi [37]. Indeed, we have verified that they are identical for degree p operators on simplex elements with constant-coefficient scalar diffusion, provided

1. the SBP matrices H_κ and B_γ and their corresponding nodes define cubature rules that are exact for polynomials of degree $2p$, and;
2. the number of SBP nodes is equal to the number of basis functions for $\mathbb{P}_p$.

When these two conditions are satisfied the SBP cubatures reproduce the L^2 norm on the volume and face exactly, so the inverse trace inequalities of Warburton and Hesthaven apply [17]. However, in general, H_κ is only exact for polynomials of degree $2p-1$ and there are more SBP nodes than basis functions in $\mathbb{P}_p$, so the penalties given here differ from [37]. Furthermore, the penalties $\delta_\gamma^{(1)}$ and $\delta_\gamma^{\mathcal{D}}$ are more general than those provided in [37], because they are applicable to spatially varying tensor diffusion and elements other than simplices.

In practice, we define SBP operators on a reference element and employ a coordinate transformation for each element in the physical domain. Therefore, some remarks are warranted regarding the implementation of the SAT-SIPG penalties when coordinate transformations are used. Let $\boldsymbol{x}(\boldsymbol{\xi})$ be an affine and bijective coordinate transformation from reference space, $\boldsymbol{\xi} = [\xi, \eta]^T \in \Omega_\xi$, to physical space. When such a coordinate mapping is used with SAT-SIPG penalties, $(\lambda_{\max})_\kappa$ corresponds to the largest eigenvalue of

$$\begin{bmatrix} \mathsf{J}\mathsf{\Lambda}_{xx} & \mathsf{J}\mathsf{\Lambda}_{xy} \\ \mathsf{J}\mathsf{\Lambda}_{yx} & \mathsf{J}\mathsf{\Lambda}_{yy} \end{bmatrix}_\kappa,$$

where J is a diagonal matrix holding the determinant of the mapping Jacobian at each node of Ω_κ. In addition, the penalties matrices $\mathsf{T}_\gamma^{(1)}$ and $\mathsf{T}_\gamma^{\mathcal{D}}$ in (39) must be multiplied by the squared norm of the scaled contravariant basis vectors at the face nodes, i.e., the diagonal matrix whose jth entry is

$$\left\| \left[\mathcal{J} (n_\xi \nabla \xi + n_\eta \nabla \eta) \right]_j \right\|^2, \qquad \forall j = 1, 2, \dots, n_\gamma,$$

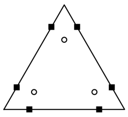
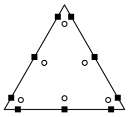
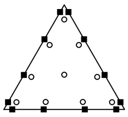
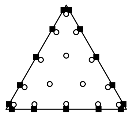
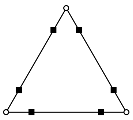
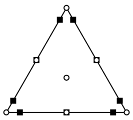
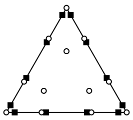
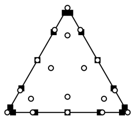
where $\mathcal{J} = \det(\partial \boldsymbol{x} / \partial \boldsymbol{\xi})$ is the determinant of the mapping Jacobian, $\nabla\xi$ and $\nabla\eta$ are the contravariant basis vectors, and n_ξ and n_η are the components of the unit normal on face γ in reference space. Note that the squared norm $\|\mathsf{B}_\gamma^{\frac{1}{2}} \mathsf{R}_{\gamma\kappa} \mathsf{H}_\kappa^{-\frac{1}{2}}\|_2^2$ can be pre-computed in reference space.

8 Numerical experiments

This section presents some numerical experiments to verify the theory developed in Sections 5, 6, and 7. For these experiments, we consider two families of SBP operators developed for simplex elements.

SBP-Ω: These operators have strictly internal nodes, and the number of nodes is equal to the number of basis functions in $\mathbb{P}_p(\Omega_\xi)$; therefore, the SBP-SAT discretizations based on these operators are essentially equivalent to collocation discontinuous-Galerkin finite-element methods. For the degree $p = 1$ and $p = 2$ operators, the SBP norm is a $2p$ degree cubature, while for $p = 3$ and $p = 4$, the norm is a degree $2p - 1$ cubature. Thus, for constant coefficient-diffusion, the SIPG penalty is identical to Shahbazi's for $p = 1$ and $p = 2$, while it is different for $p = 3$ and $p = 4$ (see the discussion in Section 7.2). The SBP-Ω operators were first presented in [23]; see also [22].

Table 2: The SBP-Ω and SBP-Γ operators for the triangle. The open circles denote the locations of the SBP nodes, while the black squares denote the locations of the face cubature points (for a given degree p SBP operator, the face cubatures are the same for both families).

	degree			
family	$p = 1$	$p = 2$	$p = 3$	$p = 4$
SBP-Ω	 3 nodes	 6 nodes	 10 nodes	 15 nodes
SBP-Γ	 3 nodes	 7 nodes	 12 nodes	 18 nodes

SBP-Γ: These operators were designed to have $p + 1$ nodes on each face; consequently, the interpolation operator $\mathsf{R}_{\gamma\kappa}$ uses only those nodes that lie on face γ. With the exception of $p = 1$, the SBP-Γ operators have more SBP nodes than basis functions in $\mathbb{P}_p(\Omega_\xi)$. For this reason, there are no (known) basis functions associated with these operators for $p > 1$; they are finite-difference operators but not finite-element operators. The SBP-Γ operators were presented in [10].

Table 2 summarizes the SBP-Ω and SBP-Γ operators considered in this work. For further details on the construction of these operators, please see [10] and [23].

The SAT-SIPG and SAT-BR2 generalizations in Section 7 are implemented with face-weight parameters, $\alpha_{\gamma\kappa}$, computed using the face area as follows:

$$\alpha_{\gamma\kappa} = \begin{cases} \dfrac{\mathcal{A}(\gamma)}{\mathcal{A}(\Gamma_\kappa^{\mathcal{I}}) + 2\mathcal{A}(\Gamma_\kappa^{\mathcal{D}})}, & \gamma \in \Gamma^{\mathcal{I}}, \\ \dfrac{2\mathcal{A}(\gamma)}{\mathcal{A}(\Gamma_\kappa^{\mathcal{I}}) + 2\mathcal{A}(\Gamma_\kappa^{\mathcal{D}})}, & \gamma \in \Gamma, \end{cases}$$

where the function $\mathcal{A}(\gamma)$ computes the size of face γ, i.e., length in 2D and area in 3D. The condition $\sum_{\gamma\subset\Gamma_\kappa} \alpha_{\gamma\kappa} = 1$ required in Lemma 1 is clearly satisfied by the above definition.

Remark 8 As noted above, the interpolation operator $\mathsf{R}_{\gamma\kappa}$ for the SBP-Γ discretizations depends only on the nodes on the boundary of γ, so the cost of applying $\mathsf{R}_{\gamma\kappa}$ or $\mathsf{R}_{\gamma\kappa}^T$ is $\mathrm{O}(p)$ in two-dimensions. In three dimensions, the SBP-Γ operators have $(p+1)(p+2)/2$ nodes on each face and the cost of applying the interpolation operator and its transpose is $\mathrm{O}(p^2)$. A consequence of this sparsity structure of $\mathsf{R}_{\gamma\kappa}$, as well as the norm H_κ being diagonal, is that the SAT-BR2 penalty matrix $\mathsf{T}_{\mathrm{BR2}}$ has an asymptotic cost of $\mathrm{O}(p)$ and $\mathrm{O}(p^2)$ in two and three dimensions, respectively, when SBP-Γ operators are used. In contrast, the SBP-Ω operators require dense extrapolation operators, so the cost of applying $\mathsf{R}_{\gamma\kappa}$ and $\mathsf{R}_{\gamma\kappa}^{T}$ — and, consequently, the cost of $\mathsf{T}_{\mathrm{BR2}}$ — scales as $\mathrm{O}(p^2)$ in two dimensions and $\mathrm{O}(p^3)$ in three dimensions.

In all cases the time derivative is discretized using the second-order accurate Crank-Nicolson method.

8.1 Description of continuous problem

We use the method of manufactured solutions to construct an analytical solution to the problem (1)–(3). We consider a unit-square domain, $\Omega = [0,1]^2$, and define the manufactured solution and tensor diffusion to be

$$\begin{aligned} \mathcal{U}(t,x,y) &= e^{-t}\sin(2\pi x)\sin(2\pi y) \\ \text{and} \qquad \Lambda &= \begin{bmatrix} x^2+1 & xy \\ xy & y^2+1 \end{bmatrix}, \end{aligned} \tag{40}$$

respectively. The source term $\mathcal{F}$ is found by substituting Λ and $\mathcal{U}$ into (1). Homogeneous Dirichlet boundary conditions are applied along the boundaries of Ω.

We will assess adjoint consistency indirectly by verifying that we achieve functional superconvergence. For this purpose we define the functional

$$\mathcal{J}(t) = \int_{\Omega} \mathcal{U}(t,x,y)^2 \mathrm{d}\Omega = \frac{1}{4}\mathrm{e}^{-2t}, \tag{41}$$

which is a special case of (21) with $\mathcal{V}_{\mathcal{N}} = 0$ and $\mathcal{V}_{\mathcal{D}} = 0$.

8.2 Accuracy study

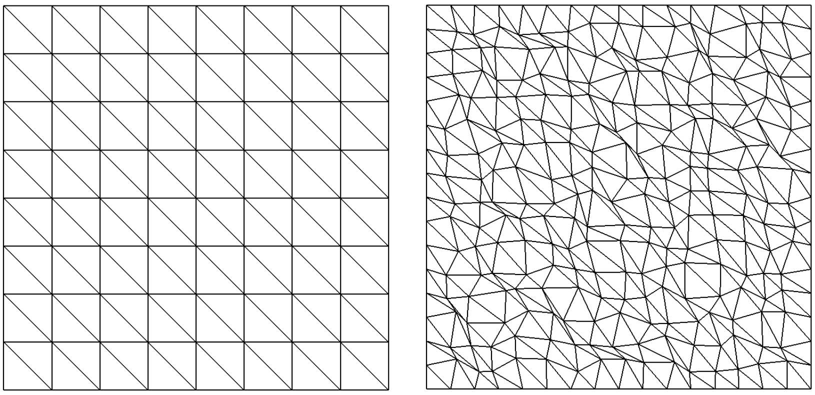
The first experiment is intended to study the accuracy of the primal discretization as well as verifying the adjoint consistency analysis; for the former we examine solution accuracy while for the latter we examine functional accuracy.

Remark 9 The primal and adjoint solution must be accurate to order h^{p+1} to obtain $\mathrm{O}(h^{2p})$ accurate functionals [16,21]. Therefore, obtaining $2p$-rate superconvergent functionals provides an indirect verification of adjoint consistency.

To assess the solution error at some time t, we use the L^2 error with the element integrals approximated using the SBP matrix H_κ scaled appropriately by the (affine) mapping Jacobian. For the functional error we use $\epsilon_{\mathcal{J}}(t) \equiv |\mathcal{J}(t) - J_h(t)|$, where $J_h(t)$ is discretized as shown in (25).

The discretization is advanced in time using Crank-Nicolson with a step size of $\Delta t = 10^{-5}$ until $t = 0.01$ units (i.e. 1000 time steps). We consider four grids with $K = 128$, 288, 648, and 1458 uniform triangular elements. The coarsest mesh is shown in Figure 1a. The nominal element size is given by $h \equiv 1/\sqrt{K/2}$.

The solution error at the final time, $t = 0.01$ units, is plotted versus h in Figures 2a and 2b for the SBP-Ω and SBP-Γ operators, respectively. In all cases the discretizations display $p+1$ convergence rates. For this particular problem,



(a) Coarsest mesh, $K = 128$

(b) Perturbed mesh, 16×16

Fig. 1: Example meshes used for numerical experiments.

the discretizations based on SBP-$\varGamma$ operators are somewhat more accurate using SAT-SIPG penalties, and the SBP-$\varOmega$ operators have slightly lower errors with SAT-BR2 penalties.

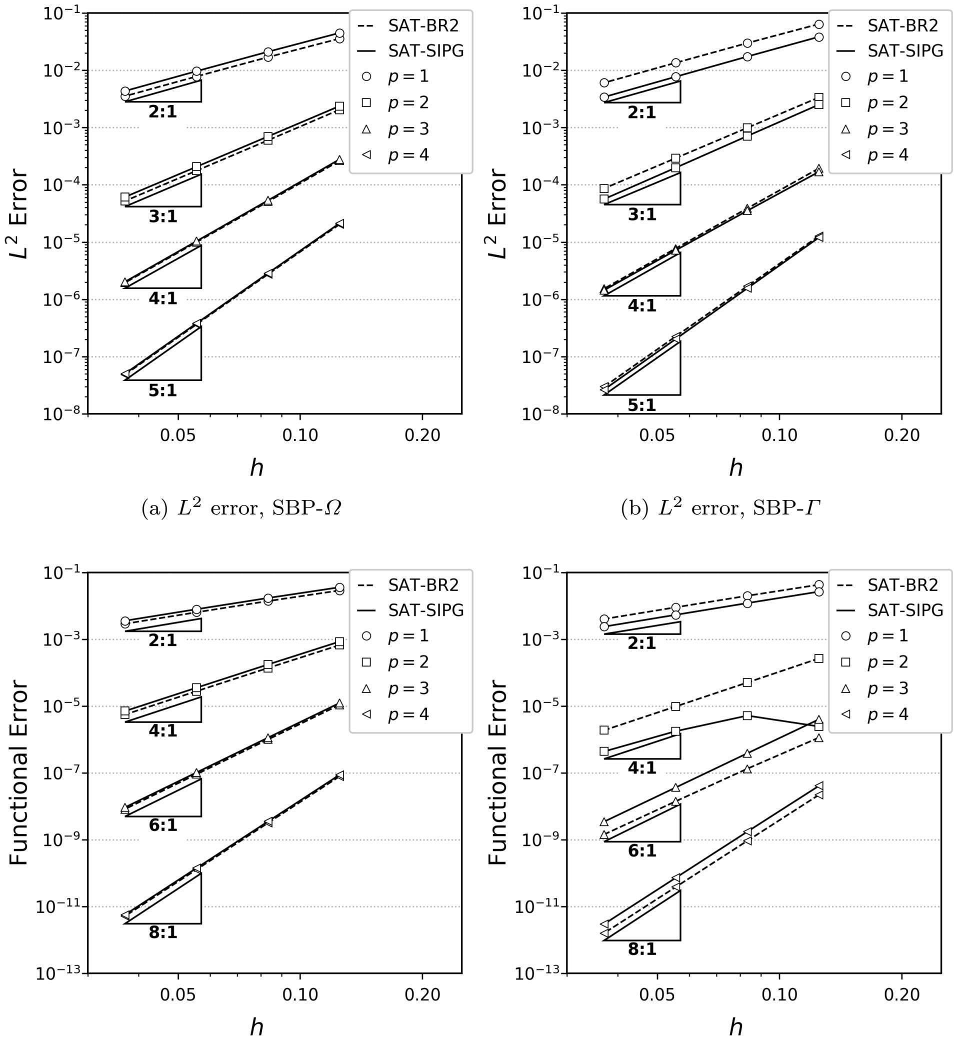
Figures 2c and 2d plot the functional errors at $t = 0.01$ for the SBP-$\varOmega$ and SBP-$\varGamma$ operators, respectively. Here we observe rates of approximately $2p$, which agrees with the theoretical order of convergence for adjoint-consistent discretizations of integral functionals [21, 32]. The only significant outlier is the $p = 2$ SBP-$\varGamma$ discretization with SAT-SIPG penalties, which does not display its asymptotic-error behavior on the grids considered; however, it is interesting that this scheme is significantly more accurate than the other $p = 2$ discretizations.

Finally, we note that, for $p > 1$, the functional errors of the SBP-$\varGamma$ discretizations are noticeably smaller than the corresponding SBP-$\varOmega$ discretizations. This is opposite the trend observed for multi-dimensional SBP discretizations of advection problems [22, 23]. Further study and analysis is necessary to explain why one operator performs better for advection problems while the other performs better for diffusion problems.

8.3 Tightness of the stability bound and energy stability

In Section 6 we proved the sufficiency of the stability conditions, but not the necessity. Therefore, scaling $\mathsf{T}_{\gamma}^{(1)}$ and $\mathsf{T}^{\mathcal{D}}$ by a relaxation factor $\sigma \in (0, 1]$ may still yield a stable bilinear form. To some degree, such a relaxation factor can serve as a measure of the tightness of the stability conditions. For example, overly conservative SAT penalties will allow for a relaxation factor $\sigma \ll 1$; on the other hand, a necessary and sufficient stability condition would only permit $\sigma \geq 1$.

To study the tightness of the stability conditions, we investigate the effect of scaling the penalties on the spectra of the SBP-SAT discretizations. In particular, we compute the eigenvalues of the global stiffness matrix $\mathsf{A}(\sigma)$, which is based on the bilinear form (19) with $\varLambda$ defined by (40); the dependence of $\mathsf{A}(\sigma)$ on the



(a) L^2 error, SBP-Ω

(b) L^2 error, SBP-Γ

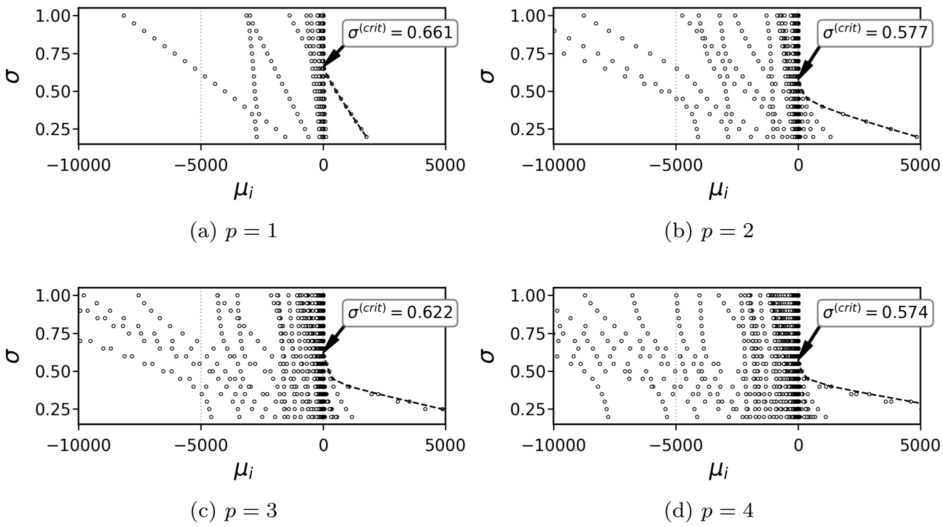
(c) functional error, SBP-Ω

(d) functional error, SBP-Γ

Fig. 2: L^2 solution and functional errors at the final time versus element size h.

relaxation factor comes about because $\mathsf{T}_{\gamma}^{(1)}$ and $\mathsf{T}^{\mathcal{D}}$ are scaled by σ, as described above. The penalties are mesh dependent, so we consider the less ideal but more realistic mesh shown in Figure 1b for this study. This mesh is extremely rough and almost tangled; indeed, the largest angle in the mesh is 179.90°.

We will denote the eigenvalues of $\mathsf{A}(\sigma)$ by $\mu_i(\sigma)$, $i = 1, 2, \dots, (\sum_{\kappa=1}^{K} n_\kappa)$. According to the theory developed in Sections 6 and 7 the maximum eigenvalue should be non-positive, $\max_i \mu_i(\sigma = 1) \leq 0$, since the discretizations are symmetric negative-semi-definite. However, as we decrease σ we expect some eigenvalues to eventually become positive. The question we are interested in answering is, how small does σ need to be for this to happen?



(a) $p = 1$

(b) $p = 2$

(c) $p = 3$

(d) $p = 4$

Fig. 3: Spectra of the SBP-Γ, SAT-BR2 discretizations as a function of the penalty relaxation factor σ.

Figure 3 plots the spectra for the discretizations based on the SBP-Γ operators with scaled SAT-BR2 penalties. The spectra are plotted in horizontal sets adjacent to the corresponding value of the relaxation parameter σ, which is varied from $\sigma = 1$ to $\sigma = 0.2$ in increments of 0.05. As σ is decreased the spurious eigenvalues shift right until one or more of the eigenvalues becomes positive; see Section 8.4 for further discussion of the spurious eigenvalues and the spectrum more generally.

The dashed lines in the subfigures of Figure 3 connect the largest (positive) eigenvalue from each spectrum, and the last two points on this line are extrapolated to estimate the critical relaxation factor, $\sigma^{(\mathrm{crit})}$, below which the scheme is unstable. For the SBP-Γ operators with SAT-BR2 penalties we see that the critical relaxation factor is approximately 0.6 for this rough mesh, which suggests that the stability bound is relatively tight — in the sense that it is not orders of magnitude larger than necessary.

Table 3 lists the critical relaxation factors for all combinations of operators and penalties considered. The critical relaxation factors are less than one for all the discretizations, which is consistent with the conditions in Theorem 2. In general, $\sigma^{(\mathrm{crit})}$ is larger for the SBP-Γ operators, indicating that the bound is tighter for these operators. Furthermore, as expected, the critical relaxation factor is larger for SAT-BR2 than it is for SAT-SIPG, since the latter is essentially a conservative bound on the former.

To complement the above study of $\sigma^{(\mathrm{crit})}$, we solve the homogeneous problem ($\mathcal{F} = 0$, and $\mathcal{U}_{\mathcal{D}} = 0$) using two different relaxation factors for each family of SBP operators: $\sigma = 1$, 0.3 for SBP-Ω and $\sigma = 1$, 0.45 for SBP-Γ. The smaller, unstable values of σ are chosen based on the results in Table 3. As mentioned in Section 6, the solution of the SBP-SAT discretization will have a decreasing energy, provided the stability conditions of Theorem 2 are satisfied.

Table 3: Critical relaxation factors, $\sigma^{(\text{crit})}$, that indicate the tightness of the stability bound for the four discretizations considered.

	SBP-Ω		SBP-Γ	
degree (p)	SAT-BR2	SAT-SIPG	SAT-BR2	SAT-SIPG
1	0.325	0.320	0.661	0.637
2	0.369	0.333	0.577	0.455
3	0.364	0.315	0.622	0.455
4	0.416	0.354	0.574	0.450

For this experiment, we again use the 16×16 perturbed mesh shown in Figure 1b and a time step of $\Delta t = 10^{-3}$. For the initial condition, we generate random values at the collocation nodes, and both the scaled and unscaled cases use the same initial condition for a given SBP operator and degree p.

Figure 4 plots the L^2 solution energy versus time: the upper and lower figures correspond to scaled and unscaled penalties, respectively. As can be seen, all solutions based on scaled penalties (i.e. $\sigma = 0.3, 0.45$) diverge eventually, even those cases for which the energy reduces below 10^{-5}. In contrast, the energy for solutions corresponding to unscaled penalties (i.e., $\sigma = 1$) is monotonically decreasing, as expected.

8.4 Accuracy of the spectra and conditioning

We conclude the results by investigating the accuracy of the spectra and the related conditioning of the discretizations. In order to compare with analytical eigenvalues, we consider the Laplace operator with unit diffusion, $\Lambda = \left[\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right]$, on the domain $[0,1]^2$. In this case, the analytical eigenvalues for the Laplacian are given by

$$\mu_{i,j}^{(\text{exact})} = \pi^2(i^2 + j^2), \qquad \forall\, i,j = 1,2,3,\ldots.$$

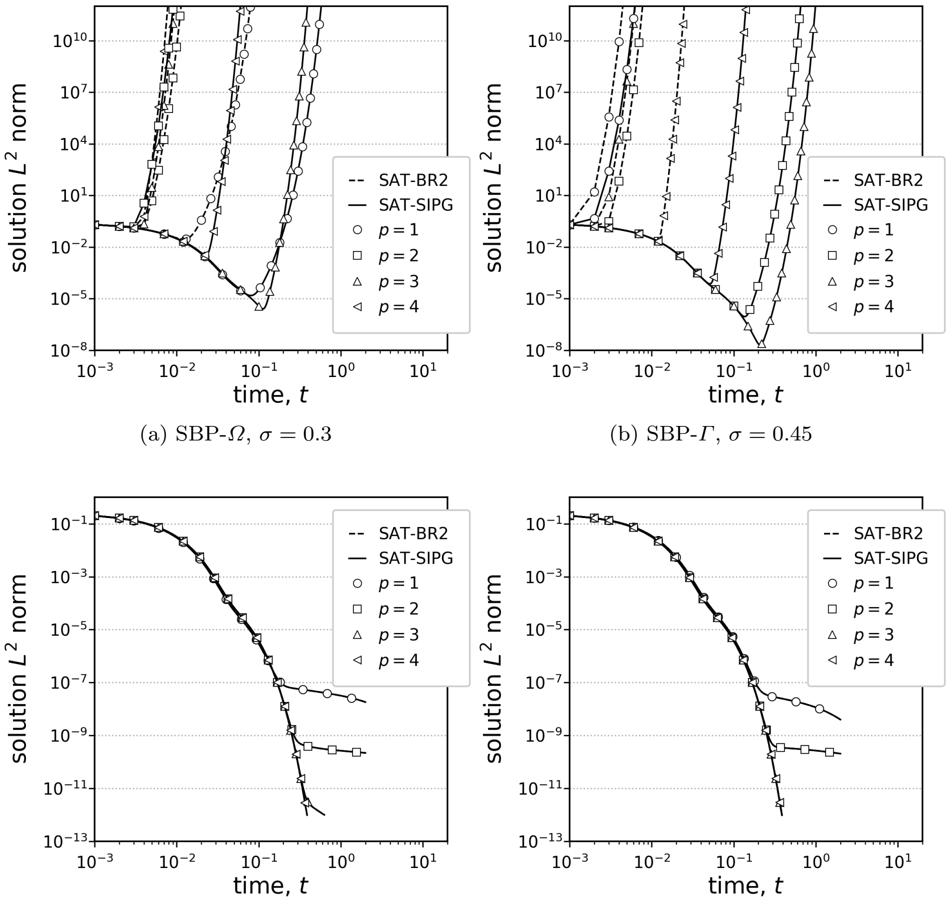
For this study, the eigenvalues for the discretizations are based on the matrix[1] $\mathsf{H}^{-1}\mathsf{A}$, where H denotes the global mass matrix and A denotes the global stiffness matrix defined by the bilinear form in (19). The discretizations are formed on the course uniform mesh shown in Figure 1a. In all cases we use unscaled penalties ($\sigma = 1$).

Figure 5 plots the relative error in the eigenvalues, defined below, for the four discretizations under consideration:

$$\text{relative eigenvalue error} \equiv \left| \mu_{i,j}/\mu_{i,j}^{(\text{exact})} - 1 \right|,$$

where $\mu_{i,j}$ denotes an eigenvalue of $\mathsf{H}^{-1}\mathsf{A}$. The errors are plotted versus eigenvalue index, which corresponds to ordering the eigenvalues in non-decreasing magnitude.

[1] Equivalently, we can consider the generalized eigenvalue problem $\mathsf{A}\boldsymbol{v}_{i,j} = \mu_{i,j}\mathsf{H}\boldsymbol{v}_{i,j}$; see, for example, [38, Chapter 8]



(a) SBP-Ω, $\sigma = 0.3$

(b) SBP-Γ, $\sigma = 0.45$

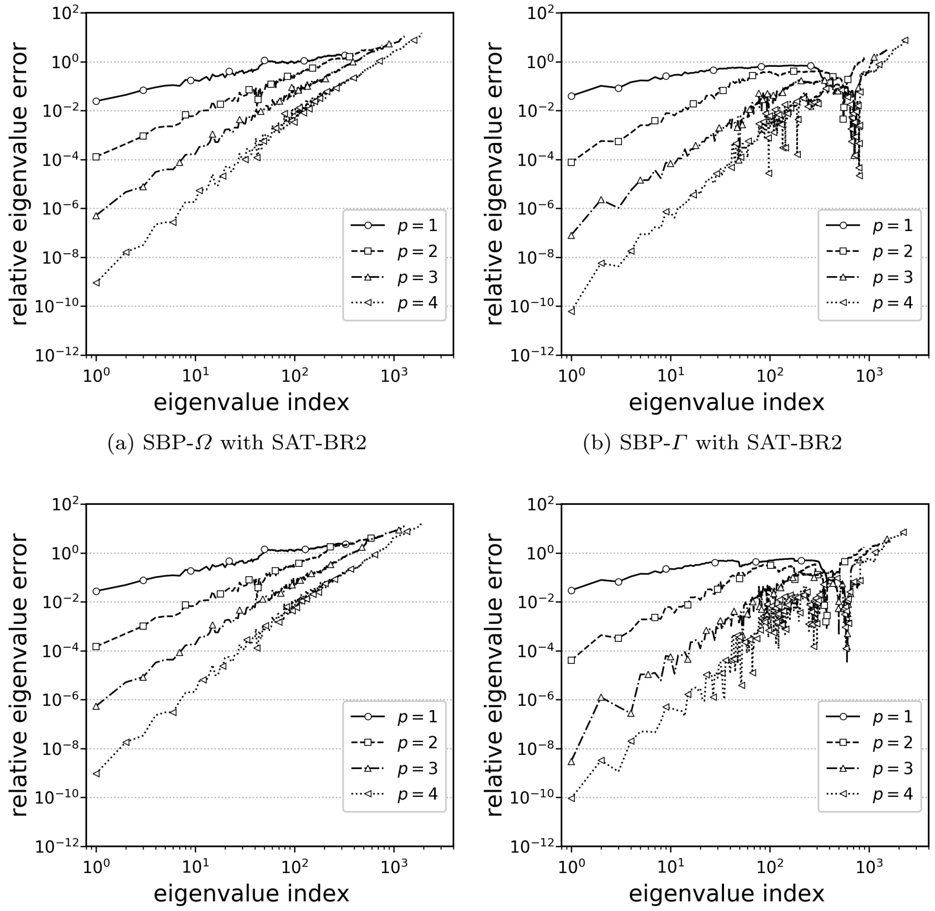
(c) SBP-Ω, $\sigma = 1$

(d) SBP-Γ, $\sigma = 1$

Fig. 4: Energy history of homogeneous problem advanced in time using Crank-Nicolson. Symbols are spaced logarithmically for clarity.

As expected, we see that all the schemes accurately capture the low-frequency modes, and that the error in these modes is decreased significantly by using higher-order schemes. At the other extreme, the high-frequency modes are poorly estimated. For analogous DG schemes, it is known that the high-frequency eigenvalues of $\mathsf{H}^{-1}\mathsf{A}$ are associated with spurious discontinuous modes [34, 38, 39].

There is only one notable difference between the subplots in Figure 5. There is a range of eigenvalues between index 600 and 1000 that the SBP-Γ discretizations capture accurately relative to the SBP-Ω discretizations. The difference is not related to the number of nodes per element, because even the $p = 1$ discretizations exhibit this difference; recall, both the SBP-Γ and SBP-Ω operators have 3 nodes per element. This may, in part, explain why the SBP-Γ discretizations produce smaller functional and solution errors for second-order PDEs — as discussed in



(a) SBP-Ω with SAT-BR2

(b) SBP-Γ with SAT-BR2

(c) SBP-Ω with SAT-SIPG

(d) SBP-Γ with SAT-SIPG

Fig. 5: Relative error in the spectra of the spatial discretizations on the uniform 8×8 mesh in Fig. 1a. Symbols are spaced logarithmically for clarity.

Section 8.2 — but this merely shifts the burden to explaining why these operators have better spectral-approximation properties.

Table 4 lists the condition numbers of the stiffness matrix, A, for the discretizations considered. As with conventional DG schemes, the condition number grows rapidly with p, which is a consequence of the spurious eigenvalues associated with the high-frequency, discontinuous modes. The resulting semi-discrete problem is stiff and is best solved using an implicit time-marching method. Of course, large condition numbers are also a concern for implicit solution methods. However, the linear system that arises in implicit methods is a linear combination of the diagonal mass matrix and the stiffness matrix, and the combination has more favorable conditioning; see [39] for the DG case. Furthermore, in the context of iterative Krylov solvers, the largest eigenvalues can be effectively eliminated using a smoother, e.g.

Table 4: Condition numbers of the Laplace-equation bilinear forms on the uniform 8×8 grid with scalar unit diffusion.

	SBP-Ω		SBP-Γ	
degree (p)	SAT-BR2	SAT-SIPG	SAT-BR2	SAT-SIPG
1	782.84	887.51	256.14	254.29
2	2870.19	3253.98	1463.22	1601.83
3	10175.81	11604.45	4670.38	5066.73
4	19249.23	22339.10	14541.89	15305.40

Gauss-Seidel or ILU preconditioners, since these eigenvalues are associated with high-frequency modes.

9 Summary and Conclusions

We generalized the SAT methodology to accommodate multi-dimensional SBP discretizations of second-order PDEs, including SBP operators whose volume nodes do not coincide with a boundary cubature. We considered a general form of SAT that uses dense penalty coefficient matrices on each face of the SBP elements. Starting with this general framework, we carried out analyses of adjoint consistency and energy stability, and, based on these analyses, we determined parameter-free conditions on the coefficient matrices that guarantee a conservative, energy-stable, primal-consistent, and adjoint-consistent discretization.

In contrast with previous finite-element analyses of interior penalties, the SAT conditions given here apply to general (tensor) diffusion coefficients and arbitrary elements. Furthermore, the conditions are entirely algebraic. Using the properties of SBP operators, our analysis accounts for inexact integration explicitly from the beginning.

Two popular interior penalty methods used in the FE community, BR2 and SIPG, were generalized to multi-dimensional SBP-SAT discretizations. We demonstrated that the SIPG penalty can be obtained from BR2 using straightforward matrix analysis; to the best of our knowledge, this algebraic connection has not been previously reported.

Several numerical test cases were carried out to verify the analysis and compare the performance of SAT-BR2 and SAT-SIPG when applied in conjunction with two families of SBP operators: the so-called SBP-Γ and SBP-Ω operators. Mesh refinement studies confirmed that the discretizations achieve design order and that they produce superconvergent functionals; the latter was used to establish adjoint consistency. Energy stability was demonstrated using an extremely skewed mesh. Furthermore, our stability bound was shown to be relatively tight in the sense that a scaling factor applied to one of the SATs could not be reduced below one order of magnitude without causing instability. Finally, the spectra of the spatial discretization was shown to be consistent with analogous discontinuous Galerkin methods.

References

1. Fisher, T. C. and Carpenter, M. H., "High-order entropy stable finite difference schemes for nonlinear conservation laws: Finite domains," *Journal of Computational Physics*, Vol. 252, No. 1, 2013, pp. 518–557.
2. Fisher, T. C., *High-order L2 stable multi-domain finite difference method for compressible flows*, Ph.D. thesis, Purdue University, 2012.
3. Kreiss, H.-O. and Scherer, G., "Finite element and finite difference methods for hyperbolic partial differential equations," *Mathematical aspects of finite elements in partial differential equations*, Academic Press, New York/London, 1974, pp. 195–212.
4. Strand, B., "Summation by parts for finite difference approximations for d/dx," *Journal of Computational Physics*, Vol. 110, No. 1, 1994, pp. 47–67.
5. Carpenter, M. H., Fisher, T. C., Nielsen, E. J., and Frankel, S. H., "Entropy Stable Spectral Collocation Schemes for the Navier–Stokes Equations: Discontinuous Interfaces," *SIAM Journal on Scientific Computing*, Vol. 36, No. 5, 2014, pp. B835–B867.
6. Parsani, M., Carpenter, M. H., and Nielsen, E. J., "Entropy stable wall boundary conditions for the three-dimensional compressible Navier-Stokes equations," *Journal of Computational Physics*, Vol. 292, No. C, 2015, pp. 88–113.
7. Gassner, G. J., Winters, A. R., and Kopriva, D. A., "Split Form Nodal Discontinuous Galerkin Schemes with Summation-by-parts Property for the Compressible Euler Equations," *Journal Computational Physics*, Vol. 327, No. C, Dec. 2016, pp. 39–66.
8. Chen, T. and Shu, C.-W., "Entropy stable high order discontinuous Galerkin methods with suitable quadrature rules for hyperbolic conservation laws," *Journal of Computational Physics*, Vol. 345, 2017, pp. 427–461.
9. Crean, J., Hicken, J. E., Del Rey Fernández, D. C., Zingg, D. W., and Carpenter, M. H., "Entropy-Stable Summation-By-Parts Discretization of the Euler Equations on General Curved Elements," *Journal of Computational Physics (in revision)*, 2017.
10. Hicken, J. E., Del Rey Fernández, D. C., and Zingg, D. W., "Multi-dimensional Summation-By-Parts Operators: General Theory and Application to Simplex Elements," *SIAM Journal on Scientific Computing*, Vol. 38, No. 4, 2016, pp. A1935–A1958.
11. Carpenter, M. H., Nordström, J., and Gottlieb, D., "A stable and conservative interface treatment of arbitrary spatial accuracy," *Journal of Computational Physics*, Vol. 148, No. 2, 1999, pp. 341–365.
12. Carpenter, M. H., Nordström, J., and Gottlieb, D., "Revisiting and extending interface penalties for multi-domain summation-by-parts operators," *Journal of Scientific Computing*, Vol. 45, No. 1, June 2010, pp. 118–150.
13. Mattsson, K. and Carpenter, M. H., "Stable and accurate interpolation operators for high-order multiblock finite difference methods," *SIAM Journal on Scientific Computing*, Vol. 32, No. 4, 2010, pp. 2298–2320.
14. Gong, J. and Nordström, J., "Interface procedures for finite difference approximations of the advection-diffusion equation," *Journal of Computational and Applied Mathematics*, Vol. 236, No. 5, 2011, pp. 602–620.
15. Del Rey Fernández, D. C. and Zingg, D. W., "Generalized summation-by-parts operators for the second derivative with a variable coefficient," *SIAM Journal on Scientific Computing*, Vol. 37, No. 6, 2015, pp. A2840–A2864.
16. Arnold, D. N., Brezzi, F., Cockburn, B., and Marini, L. D., "Unified Analysis of Discontinuous Galerkin Methods for Elliptic Problems," *SIAM Journal on Numerical Analysis*, Vol. 39, No. 5, 2002, pp. 1749–1779.
17. Warburton, T. and Hesthaven, J., "On the constants in hp-finite element trace inverse inequalities," *Computer Methods in Applied Mechanics and Engineering*, Vol. 192, No. 25, 2003, pp. 2765 – 2773.
18. Baumann, C. E. and Oden, J. T., "A discontinuous hp finite element method for convection–diffusion problems," *Computer Methods in Applied Mechanics and Engineering*, Vol. 175, No. 3, 1999, pp. 311–341.
19. Bassi, F., Crivellini, A., Rebay, S., and Savini, M., "Discontinuous Galerkin solution of the Reynolds-averaged Navier–Stokes and k–w turbulence model equations," *Computers & Fluids*, Vol. 34, No. 4–5, 2005, pp. 507 – 540.
20. Douglas, J. and Dupont, T., "Interior Penalty Procedures for Elliptic and Parabolic Galerkin Methods Computing Methods in Applied Sciences," *Computing Methods in Applied Sciences*, edited by R. Glowinski and J. L. Lions, Vol. 58 of *Lecture Notes in Physics*, chap. 6, Springer Berlin / Heidelberg, Berlin, Heidelberg, 1976, pp. 207–216.

21. Hartmann, R., “Adjoint Consistency Analysis of Discontinuous Galerkin Discretizations,” *SIAM Journal on Numerical Analysis*, Vol. 45, No. 6, 2007, pp. 2671–2696.
22. Hicken, J. E., Del Rey Fernández, D. C., and Zingg, D. W., “Simultaneous approximation terms for multi-dimensional summation-by-parts operators,” *46th AIAA Fluid Dynamics Conference*, Washington, DC, June 2016, AIAA–2016–3971.
23. Del Rey Fernández, D. C., Hicken, J. E., and Zingg, D. W., “Simultaneous Approximation Terms for Multi-dimensional Summation-by-Parts Operators,” *Journal of Scientific Computing*, Aug. 2017, pp. 1–28.
24. Mattsson, K., “Summation by Parts Operators for Finite Difference Approximations of Second-Derivatives with Variable Coefficients,” *Journal of Scientific Computing*, Vol. 51, No. 3, June 2012, pp. 650–682.
25. Lipnikov, K., Manzini, G., and Shashkov, M., “Mimetic finite difference method,” *Journal of Computational Physics*, Vol. 257, Jan. 2014, pp. 1163–1227.
26. Babuška, I. and Miller, A., “The post-processing approach in the finite element method—part 1: Calculation of displacements, stresses and other higher derivatives of the displacements,” *International Journal for Numerical Methods in Engineering*, Vol. 20, No. 6, June 1984, pp. 1085–1109.
27. Babuška, I. and Miller, A., “The post-processing approach in the finite element method—Part 2: The calculation of stress intensity factors,” *International Journal for Numerical Methods in Engineering*, Vol. 20, No. 6, June 1984, pp. 1111–1129.
28. Pierce, N. A. and Giles, M. B., “Adjoint recovery of superconvergent functionals from PDE approximations,” *SIAM Review*, Vol. 42, No. 2, 2000, pp. 247–264.
29. Giles, M. B. and Süli, E., “Adjoint methods for PDEs: a posteriori error analysis and postprocessing by duality,” *Acta Numerica*, Vol. 11, 2002, pp. 145–236.
30. Lu, J. C., *An a posteriori error control framework for adaptive precision optimization using discontinuous Galerkin finite element method*, Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, 2005.
31. Fidkowski, K. J. and Darmofal, D. L., “Review of output-based error estimation and mesh adaptation in computational fluid dynamics,” *AIAA Journal*, Vol. 49, No. 4, April 2011, pp. 673–694.
32. Hicken, J. E. and Zingg, D. W., “Superconvergent functional estimates from summation-by-parts finite-difference discretizations,” *SIAM Journal on Scientific Computing*, Vol. 33, No. 2, 2011, pp. 893–922.
33. Hicken, J. E. and Zingg, D. W., “Dual consistency and functional accuracy: a finite-difference perspective,” *Journal of Computational Physics*, Vol. 256, Jan. 2014, pp. 161–182.
34. Antonietti, P. F., Buffa, A., and Perugia, I., “Discontinuous Galerkin approximation of the Laplace eigenproblem,” *Computer Methods in Applied Mechanics and Engineering*, Vol. 195, No. 25, 2006, pp. 3483–3503.
35. Shahbazi, K., Mavriplis, D. J., and Burgess, N. K., “Multigrid algorithms for high-order discontinuous Galerkin discretizations of the compressible Navier-Stokes equations,” *Journal of Computational Physics*, Vol. 228, No. 21, Nov. 2009, pp. 7917–7940.
36. Peraire, J. and Persson, P.-O., “The Compact Discontinuous Galerkin (CDG) Method for Elliptic Problems,” *SIAM Journal on Scientific Computing*, Vol. 30, No. 4, 2008, pp. 1806–1824.
37. Shahbazi, K., “An explicit expression for the penalty parameter of the interior penalty method,” *Journal of Computational Physics*, Vol. 205, No. 2, May 2005, pp. 401–407.
38. Hesthaven, J. S. and Warburton, T., *Nodal discontinuous Galerkin methods: algorithms, analysis, and applications*, Springer-Verlag, New York, 2008.
39. Kirby, R. M. and Karniadakis, G. E., “Selecting the Numerical Flux in Discontinuous Galerkin Methods for Diffusion Problems,” *Journal of Scientific Computing*, Vol. 22-23, No. 1-3, Jan. 2005, pp. 385–411.