IMA Journal of Numerical Analysis (2016) Page 1 of 28
doi:10.1093/imanum/drnxxx

Divided difference estimates and accuracy enhancement of discontinuous Galerkin methods for nonlinear symmetric systems of hyperbolic conservation laws

XIONG MENG†
School of Mathematics, University of East Anglia,
Norwich Research Park, Norwich NR4 7TJ, United Kingdom,
and Department of Mathematics, Harbin Institute of Technology,
Harbin, Heilongjiang 150001, China

AND

JENNIFER K. RYAN‡
School of Mathematics, University of East Anglia,
Norwich Research Park, Norwich NR4 7TJ, United Kingdom

[Received on 17 May 2016; revised on 2 December 2016]

In this paper, we investigate the accuracy-enhancement for the discontinuous Galerkin (DG) method for solving one-dimensional nonlinear symmetric systems of hyperbolic conservation laws. For nonlinear equations, the divided difference estimate is an important tool that allows for superconvergence of the post-processed solutions in the local $\boldsymbol{L}^2$ norm. Therefore, we first prove that the $\boldsymbol{L}^2$ norm of the α-th order $(1 \leqslant \alpha \leqslant k+1)$ divided difference of the DG error with upwind fluxes is of order $k+\frac{3}{2}-\frac{\alpha}{2}$, provided that the flux Jacobian matrix, $\boldsymbol{f}'(\boldsymbol{u})$, is symmetric positive definite. Furthermore, using the duality argument, we are able to derive superconvergence estimates of order $2k+\frac{3}{2}-\frac{\alpha}{2}$ for the negative-order norm, indicating that some particular compact kernels can be used to extract at least $(\frac{3}{2}k+1)$th order superconvergence for nonlinear systems of conservation laws. Numerical experiments are shown to demonstrate the theoretical results.

Keywords: discontinuous Galerkin method; nonlinear symmetric systems of hyperbolic conservation laws; negative-order norm estimates; post-processing; divided difference.

1. Introduction

Smoothness-Increasing Accuracy-Conserving (SIAC) filtering allows for extracting a higher-order accurate solution from the discontinuous Galerkin (DG) approximation, which can aid in reducing approximation errors. The motivation for this study is that the accuracy enhancing capabilities of the SIAC filter (Ryan *et al.*, 2005; Mirzaee *et al.*, 2011) for the DG method requires establishing convergence characteristics for the divided difference of the errors; see Theorem 2.1 below. In Meng & Ryan (2017), this was done for nonlinear scalar hyperbolic conservation laws. However, extending these estimates to nonlinear hyperbolic systems is more challenging. A nonlinear system of hyperbolic conservation laws is a more general model arising from fluid dynamics. One such model is the Euler equations in gas

†Email: xiongmeng@hit.edu.cn
‡Corresponding author. Email: Jennifer.Ryan@uea.ac.uk

dynamics. In this paper we concentrate on the theoretical and computational aspects of the accuracy-enhancement of DG methods for solving one-dimensional nonlinear systems of conservation laws of the form

$$\boldsymbol{u}_t + \boldsymbol{f}(\boldsymbol{u})_x = 0, \qquad (x,t) \in (a,b) \times (0,T], \tag{1.1a}$$

$$\boldsymbol{u}(x,0) = \boldsymbol{u}_0(x), \qquad x \in \Omega = (a,b), \tag{1.1b}$$

where $\boldsymbol{u}_0(x)$ is a given smooth initial function. Here $\boldsymbol{u} = (u_1,\ldots,u_m)^{\mathrm{T}}$ is the unknown vector-valued solution, and $\boldsymbol{f}(\boldsymbol{u}) = (f_1(\boldsymbol{u}),\ldots,f_m(\boldsymbol{u}))^{\mathrm{T}}$ with $f_i(\boldsymbol{u}) = f_i(u_1,\ldots,u_m)$ $(i = 1,\ldots,m)$ is the given flux function. The nonlinear flux function $\boldsymbol{f}(\boldsymbol{u})$ is assumed to be sufficiently smooth with respect to the the exact solution $\boldsymbol{u}$, and $\boldsymbol{u}$ is a smooth function of x. In this paper, periodic boundary conditions are assumed, which, however, is not essential. We show that the $\boldsymbol{L}^2$ norm of the α-th order $(1 \leqslant \alpha \leqslant k+1)$ divided difference of the DG error achieves $(k+\frac{3}{2}-\frac{\alpha}{2})$th order using upwind fluxes, provided that the flux Jacobian matrix $\boldsymbol{f}'(\boldsymbol{u}) := \partial \boldsymbol{f}/\partial \boldsymbol{u}$ is positive definite. By a duality argument, a superconvergent negative-order norm estimate of order $2k+\frac{3}{2}-\frac{\alpha}{2}$ is further obtained. This allows for extracting the hidden accuracy of at least $(\frac{3}{2}k+1)$th order for nonlinear systems of conservation laws, indicating that it is possible to use a more compact kernel with fewer B-splines.

The DG method has an inherent superconvergence property, which has attracted the attention of many researchers for solving the first-order hyperbolic equations (see, e.g., Adjerid *et al.*, 2002; Adjerid & Massey, 2006; Adjerid & Weinhart, 2009, 2011; Cao *et al.*, 2014; Cheng & Shu, 2010; Cockburn *et al.*, 2003; Guo *et al.*, 2013; Ryan *et al.*, 2005; Steffen *et al.*, 2008; Yang & Shu, 2012), high order equations (see, e.g., Celiker & Cockburn, 2007; Ji *et al.*, 2012; Hufford & Xing, 2014; Meng *et al.*, 2012b) and elliptic problems (see, e.g., Adjerid & Baccouch, 2012; Cockburn *et al.*, 2009). One of the superconvergence properties that allows for superconvergence extraction through SIAC filtering is the negative-order norm estimates. The post-processing technique makes use of information contained in the negative-order norm entailing that a special convolution kernel can be constructed to extract the hidden accuracy. This is performed only at the very end of the computation. Some superconvergent post-processing results of DG methods for hyperbolic equations are available in the literature. Motivated by the work of Bramble and Schatz for elliptic equations in Bramble & Schatz (1977), Cockburn *et al.* (2003) established the post-processing theory for DG methods for hyperbolic equations that expresses the post-processed solutions in the $\boldsymbol{L}^2$ norm in terms of the divided difference error estimates in the negative-order norm. Later, Ryan et al. investigated different aspects of the SIAC filters (see, e.g., Ryan & Shu, 2003; Curtis *et al.*, 2007; Steffen *et al.*, 2008).

From the post-processing theory in Bramble & Schatz (1977) and Cockburn *et al.* (2003), it is evident that negative-order norm error estimates of the divided differences are essential tools that allow for extracting superconvergent estimates of the post-processed solutions in the $\boldsymbol{L}^2$ norm. We note that, unlike purely linear equations (Cockburn *et al.*, 2003; Ji *et al.*, 2012), superconvergent estimates about the post-processed solution for quasi-linear/nonlinear equations require establishing both the L^2 norm and negative-order norm estimates of divided differences of the DG error. For example, for linear hyperbolic equations with variable coefficient, negative-order norm error estimates of the divided differences are shown in Mirzaee *et al.* (2011), and the corresponding L^2 norm estimates are provided in Meng & Ryan (2017).

Let us now mention a particular work that investigates accuracy enhancement and divided difference error estimates of DG methods for scalar nonlinear hyperbolic conservation laws (Meng & Ryan, 2017). Specifically, the analysis starts from a superconvergence result of the DG solution towards a particular projection of the exact solution (supercloseness). Then, by establishing important relations between the spatial derivatives and time derivatives of a particular projection of divided differences of DG errors and further by analyzing L^2 estimates of the time derivatives of the error, we were able to derive a useful L^2 norm error estimates for the divided difference. Next, superconvergent negative-order norm error estimates for the divided difference are obtained which depend on a suitable construction of the dual problem for the divided difference of the nonlinear scalar hyperbolic conservation laws.

To set a solid theoretical foundation of the post-processing technique for more general problems that are useful in computational fluid dynamics, it is therefore necessary to study the accuracy enhancement of DG methods for nonlinear (symmetric) systems of hyperbolic conservation laws. The generalization from the scalar nonlinear case to systems of nonlinear conservation laws in this paper involves both similarities and further difficulties and thus some new techniques are needed. As for the similarities, we would like to mention that an energy analysis is used and Taylor expansion is employed to deal with the nonlinearity of the flux function. Another similarity is that the superconvergence analyses both indicate a possible link between supercloseness and negative-order norm estimates; see the detailed proof below and also in Meng & Ryan (2017).

As indicated in Meng & Ryan (2017), the first main difficulty arising from L^2 norm estimates of the divided difference of the particular projection of the DG error can be handled by establishing important relations between the spatial derivatives and time derivatives of a particular projection of divided differences of DG errors. However, another essential difficulty in this paper is treating estimates of the divided difference of the projection error as the projection for the nonlinear systems is no longer linear. Note that the projection for the system case is constructed based on the local characteristic decomposition, and therefore, by Leibniz rule, the main difficulty is switched to estimating the divided difference of $\boldsymbol{R}$, whose columns are the right eigenvectors of the flux Jacobian $\boldsymbol{f}'(\boldsymbol{u})$ linearized at the center of each cell. To this end, we propose to analyze the eigenstructures of $\boldsymbol{f}'(\boldsymbol{u})$ and find that $\boldsymbol{R}$ can be expressed in terms of the components of $\boldsymbol{f}'(\boldsymbol{u})$ as well as its eigenvalues. Further, noting that the entries of $\boldsymbol{R}$ are compositions of some smooth functions, and using the chain rule for divided differences (see, e.g., Floater & Lyche, 2007) as well as the chain rule for derivatives (Faà di Bruno's Formula), we conclude that the leading term of the divided difference of $\boldsymbol{R}$ is a constant matrix. This finding together with the fact that the divided difference of the projection error of the characteristic variable is in possession of optimal approximation error estimate leads to the desired results in Corollary 2.2 and Corollary 3.1.

There are some other difficulties in deriving superconvergent error estimates of DG methods for nonlinear systems of conservation laws. As mentioned before, a supercloseness result about a special projection of the DG error (denoted by $\boldsymbol{\xi} := \mathbb{P}\boldsymbol{u} - \boldsymbol{u}_h = \mathbb{P}\boldsymbol{e}$) needs to be established, which is a starting point in advancing $\boldsymbol{L}^2$ norm estimates for high order divided differences. In order to do this, unlike Meng & Ryan (2017) or Meng *et al.* (2012a), we express the $\boldsymbol{L}^2$ norm of $\boldsymbol{\xi}_t$ in terms of the jump seminorm of $\boldsymbol{\xi}$ rather than the $\boldsymbol{L}^2$ norm of $\boldsymbol{\xi}$; see Lemma 3.3 below and Lemma 3.7 in Meng *et al.* (2012a). Additionally, to perform error estimates for a nonlinear system of hyperbolic conservation laws, the properties of the divided difference for composite functions and clear definitions of the special Gauss–Radau projection as well as the upwind numerical flux should also be illustrated. Finally, we would like to point out that it is not trivial for the two-dimensional extension, especially for establishing the relations between spatial derivatives and time derivatives of the errors that are used to derive a sharp bound for the $\boldsymbol{L}^2$ norm of divided differences of the DG error.

This paper is organized as follows. In Section 2, we give the DG scheme for the divided differences of nonlinear systems of hyperbolic conservation laws, and present some preliminaries especially for the properties of divided differences as well as the DG spatial operator. In Section 3, we state and discuss the $\boldsymbol{L}^2$ norm error estimates for divided differences of nonlinear systems of hyperbolic conservation laws, and then display the main proofs for a supercloseness result and divided difference estimates. Further, superconvergent negative-order norm error estimates are given in Section 4. In Section 5, numerical experiments are shown to demonstrate the theoretical results. Concluding remarks and comments on future work are given in section 6. Finally, in the appendix we provide the proofs for some of the more technical lemmas.

2. The DG scheme and preliminaries

2.1 *The DG scheme*

In this section, we follow Cockburn *et al.* (1989), Meng & Ryan (2017) and present the DG scheme for divided differences of nonlinear system of hyperbolic conservation laws (1.1).

The standard notation of the DG method is used here. We use the mesh $a = x_{\frac{1}{2}} < x_{\frac{3}{2}} < \cdots < x_{N+\frac{1}{2}} = b$ to cover the domain $\Omega = (a,b)$, and set $x_j = (x_{j-\frac{1}{2}} + x_{j+\frac{1}{2}})/2$. To facilitate analysis of divided difference estimates, we introduce two overlapping uniform meshes for Ω, denoted by $I_j = (x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}})$ and $I_{j+\frac{1}{2}} = (x_j, x_{j+1})$ with mesh size $h = x_{j+\frac{1}{2}} - x_{j-\frac{1}{2}}$. Associated with these meshes, the following discontinuous finite element space are defined

$$\boldsymbol{V}_h^{\alpha,k} = \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega^\alpha) : \boldsymbol{v}|_{I_{j'}} \in \boldsymbol{\mathcal{P}}^k(I_{j'}),\ \forall j' = j + \frac{\ell}{2}, \ell = \alpha \bmod 2, j = 1, \ldots, N\},$$

where $\boldsymbol{L}^2(\Omega^\alpha) := [L^2(\Omega^\alpha)]^m$ with $\Omega^\alpha = (a + \frac{\ell}{2}h, b + \frac{\ell}{2}h)$, $\boldsymbol{\mathcal{P}}^k(I_{j'}) := [\mathcal{P}^k(I_{j'})]^m$, and $\mathcal{P}^k(I_{j'})$ is the space of polynomials of degree at most k on the cell $I_{j'} := (x_{j'-\frac{1}{2}}, x_{j'+\frac{1}{2}})$. Here and in what follows, α denotes the α-th order divided difference of a smooth or piecewise function, that is

$$\partial_h^\gamma \boldsymbol{w}(x) = \frac{1}{h^\gamma} \sum_{i=0}^{\gamma} (-1)^i \binom{\gamma}{i} \boldsymbol{w}\left(x + \left(\frac{\gamma}{2} - i\right) h\right). \tag{2.1}$$

In particular, if α is even, we set $\boldsymbol{V}_h^{\alpha,k} = \boldsymbol{V}_h^k$. Noting that functions in $\boldsymbol{V}_h^{\alpha,k}$ are allowed to have discontinuities across cell interfaces, we use $\boldsymbol{w}_i^-$ and $\boldsymbol{w}_i^+$ to represent the left and right limits of $\boldsymbol{w}(x)$ at the discontinuity point x_i. Furthermore, at each element boundary point, the jump and the mean of $\boldsymbol{w}(x)$ are denoted by $[\![\boldsymbol{w}]\!] = \boldsymbol{w}^+ - \boldsymbol{w}^-$ and $\{\!\{\boldsymbol{w}\}\!\} = \frac{1}{2}(\boldsymbol{w}^+ + \boldsymbol{w}^-)$, respectively.

The α-th order divided difference of the nonlinear systems of conservation laws (1.1) is

$$\partial_h^\alpha \boldsymbol{u}_t + \partial_h^\alpha \boldsymbol{f}(\boldsymbol{u})_x = 0, \qquad (x,t) \in \Omega^\alpha \times (0,T], \tag{2.2a}$$
$$\partial_h^\alpha \boldsymbol{u}(x,0) = \partial_h^\alpha \boldsymbol{u}_0(x), \qquad x \in \Omega^\alpha. \tag{2.2b}$$

We are now ready to define the DG scheme for (2.2). That is, find $\partial_h^\alpha \boldsymbol{u}_h \in \boldsymbol{V}_h^{\alpha,k}$ such that the following *weak* formulation

$$\left((\partial_h^\alpha \boldsymbol{u}_h)_t, \boldsymbol{v}_h\right)_{j'} = \mathcal{H}_{j'}\left(\partial_h^\alpha \boldsymbol{f}(\boldsymbol{u}_h), \boldsymbol{v}_h\right) \tag{2.3}$$

is satisfied for all $\boldsymbol{v}_h \in \boldsymbol{V}_h^{\alpha,k}$ and $j = 1, \ldots, N$, where $\mathcal{H}_{j'}(\cdot,\cdot)$ represents the DG spatial discretization operator defined on each cell $I_{j'}$, i.e.,

$$\mathcal{H}_{j'}(\boldsymbol{w}, \boldsymbol{v}) = (\boldsymbol{v}_x, \boldsymbol{w})_{j'} - \left((\boldsymbol{v}^-)^{\mathrm{T}} \hat{\boldsymbol{w}}\right)_{j'+\frac{1}{2}} + \left((\boldsymbol{v}^+)^{\mathrm{T}} \hat{\boldsymbol{w}}\right)_{j'-\frac{1}{2}}.$$

As usual, $(\cdot,\cdot)_{j'}$ denotes the standard inner product in $\boldsymbol{L}^2(I_{j'})$, i.e., $(\boldsymbol{w}, \boldsymbol{v})_{j'} = \int_{I_{j'}} \boldsymbol{w}^{\mathrm{T}} \boldsymbol{v}\, dx$.

Similar to the scalar nonlinear conservation laws in Meng & Ryan (2017), the numerical flux $\hat{\boldsymbol{f}}_{j'+\frac{1}{2}}$ is chosen to be an upwind flux. For completeness, in what follows we shall present the detailed definition of upwind flux for (2.3). The idea is based on the local characteristic decomposition. Following Cockburn *et al.* (1989) and Zhang & Shu (2006), consider the Jacobian flux $\boldsymbol{f}'(\boldsymbol{u}_{j'+1/2}) := \boldsymbol{f}'(\boldsymbol{u})|_{\boldsymbol{u}=\boldsymbol{u}_{j'+1/2}}$. The corresponding eigenvalues, left and right eigenvectors are denoted by $\lambda_i, \boldsymbol{\ell}_i, \boldsymbol{r}_i$ $(i = 1,\ldots,m)$, normalized so that $\boldsymbol{\ell}_m \boldsymbol{r}_n = \delta_{m,n}$. Further, at each cell boundary point $x_{j'+\frac{1}{2}}$, the numerical flux $\hat{\boldsymbol{f}}_{j'+\frac{1}{2}} = \hat{\boldsymbol{f}}\big((\boldsymbol{u}_h)^-_{j'+\frac{1}{2}}, (\boldsymbol{u}_h)^+_{j'+\frac{1}{2}}\big)$ is determined by the following procedure.

1. Transform $\boldsymbol{f}(\boldsymbol{u}_h^{\pm})$ to the eigenspace of $\boldsymbol{f}'(\boldsymbol{u}_{j'+1/2})$, i.e.,

$$v_i^{\pm} = \boldsymbol{\ell}_i\, \boldsymbol{f}(\boldsymbol{u}_h^{\pm}), \quad i = 1,\ldots,m.$$

2. Apply the scalar upwind setting to $v_i^{\pm}$ in the ith characteristic field $(i = 1,\ldots,m)$, and the numerical flux $\hat{v}_i$ depends on the sign of λ_i, i.e.,

$$\hat{v}_i = \begin{cases} v_i^-, & \text{if } \lambda_i \geqslant 0, \\ v_i^+, & \text{if } \lambda_i < 0. \end{cases}$$

3. The result is transformed back to the physical field to get $\hat{\boldsymbol{f}}_{j'+\frac{1}{2}}$, namely

$$\hat{\boldsymbol{f}}_{j'+\frac{1}{2}} = \sum_{i=1}^{m} \hat{v}_i\, \boldsymbol{r}_i.$$

Moreover, analysis of L^2 norm error estimates of divided differences requires that the flux Jacobian matrix $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite. That is, eigenvalues of $\boldsymbol{f}'(\boldsymbol{u}_{j'+1/2})$ are all positive. It follows from the above procedure that $\hat{\boldsymbol{f}}_{j'+\frac{1}{2}} = \boldsymbol{f}\big((\boldsymbol{u}_h)^-_{j'+\frac{1}{2}}\big)$. Consequently,

$$\mathcal{H}_{j'}(\boldsymbol{w},\boldsymbol{v}) = (\boldsymbol{v}_x,\boldsymbol{w})_{j'} - \big((\boldsymbol{v}^-)^{\mathrm{T}}\boldsymbol{w}^-\big)_{j'+\frac{1}{2}} + \big((\boldsymbol{v}^+)^{\mathrm{T}}\boldsymbol{w}^-\big)_{j'-\frac{1}{2}} \tag{2.4a}$$

$$= -(\boldsymbol{v},\boldsymbol{w}_x)_{j'} - \big((\boldsymbol{v}^+)^{\mathrm{T}}[\![\boldsymbol{w}]\!]\big)_{j'-\frac{1}{2}}. \tag{2.4b}$$

For periodic boundary conditions, the removal of j' in $\mathcal{H}_{j'}$ denotes the sum of all $I_{j'}$, i.e.,

$$\mathcal{H}(\boldsymbol{w},\boldsymbol{v}) = (\boldsymbol{v}_x,\boldsymbol{w}) + \sum_{j=1}^{N} \big([\![\boldsymbol{v}]\!]^{\mathrm{T}}\boldsymbol{w}^-\big)_{j'+\frac{1}{2}} \tag{2.5a}$$

$$= -(\boldsymbol{v},\boldsymbol{w}_x) - \sum_{j=1}^{N} \big((\boldsymbol{v}^+)^{\mathrm{T}}[\![\boldsymbol{w}]\!]\big)_{j'-\frac{1}{2}}, \tag{2.5b}$$

where $(\boldsymbol{v},\boldsymbol{w}) = \sum_{j=1}^{N} (\boldsymbol{v},\boldsymbol{w})_{j'}$ denotes the inner product in $\boldsymbol{L}^2(\Omega^\alpha)$. Here and below, in order to distinguish two overlapping meshes the summation is calculated with respect to j rather than j'.

2.2 *Preliminaries*

In this section, we introduce the necessary norms, projections and inequalities that will be useful in our analysis. We begin by noting that C is used to denote a generic positive constant which is independent of h but may depend on the exact solution $\boldsymbol{u}$ as well as its time and spatial derivatives. Moreover, we denote by $C_\star$ a nonnegative constant that depends on higher order (at least second order) derivatives of $\boldsymbol{f}(\boldsymbol{u})$.

2.2.1 *Sobolev spaces and norms.* For systems of conservation laws discussed in this paper, we would like to use $\|\cdot\|_M$ to represent the 2-norm (length) of a vector, or the spectral norm of a real matrix, respectively. Specifically, $\|\boldsymbol{v}\|_M = \sqrt{\sum_{i=1}^m v_i^2}$ for any vector $\boldsymbol{v} = (v_1, v_2, \ldots, v_m)^{\mathrm{T}}$, and $\|\boldsymbol{A}\|_M = \sqrt{\rho(\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A})}$ for any real matrix $\boldsymbol{A}$, i.e., $\|\boldsymbol{A}\|_M$ is the square root of the largest eigenvalue of the positive-semidefinite matrix $\boldsymbol{A}^{\mathrm{T}}\boldsymbol{A}$. Furthermore, if $\boldsymbol{A}$ is symmetric, then $\|\boldsymbol{A}\|_M = \rho(\boldsymbol{A})$. For any matrix-valued function $\boldsymbol{A}$ and vector-valued functions $\boldsymbol{w}, \boldsymbol{v}$, the following Cauchy–Schwarz inequality is helpful in our analysis

$$|\boldsymbol{w}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{v}| \leqslant \|\boldsymbol{A}\|_M \|\boldsymbol{w}\|_M \|\boldsymbol{v}\|_M. \tag{2.6}$$

The Sobolev spaces can be easily defined for the vector-valued function space. To be more specific, for any integer $s \geqslant 0$, we use $\boldsymbol{W}^{s,p}(D) := [W^{s,p}(D)]^m$ to denote the vector-valued Sobolev space on subdomain $D \subset \Omega$ with the norm $\|\cdot\|_{s,p,D}$. In particular, if $p = 2$, we set $\boldsymbol{W}^{s,p}(D) = \boldsymbol{H}^s(D)$, and $\|\cdot\|_{s,p,D} = \|\cdot\|_{s,D}$, and further if $s = 0$, we set $\|\cdot\|_{s,D} = \|\cdot\|_D$ with $\|\boldsymbol{v}\|_D = \sqrt{\int_D \|\boldsymbol{v}\|_M^2\, dx}$. If $p = \infty, s = 0$, we set $\boldsymbol{W}^{s,p}(D) = \boldsymbol{L}^\infty(D)$, and $\|\cdot\|_{s,p,D} = \|\cdot\|_{\infty,D}$ with $\|\boldsymbol{v}\|_{\infty,D} = \operatorname{ess\,sup}_{x\in D} \|\boldsymbol{v}(x)\|_M$. For simplicity, when $D = \Omega$, we will omit the index D. The norms of matrix-valued Sobolev space can be defined in the same way. Moreover, we use Ω_h to denote the union of all elements, i.e., $\Omega_h = \{D\}$, and the norm of *broken* Sobolev spaces $\boldsymbol{W}^{s,p}(\Omega_h) := \{\boldsymbol{v} \in \boldsymbol{L}^2(\Omega) : \boldsymbol{v}|_D \in \boldsymbol{W}^{s,p}(D),\ \forall D \subset \Omega_h\}$ can be easily defined, which is a formal sum of the contribution to each element D. Besides, for $\boldsymbol{v} \in \boldsymbol{H}^1(\Omega_h)$, the $\boldsymbol{L}^2$ norm at cell boundaries and the *jump seminorm* are defined as follows

$$\|\boldsymbol{v}\|_{\Gamma_h} = \left(\sum_{j=1}^N \left(\|\boldsymbol{v}^+_{j'-1/2}\|_M^2 + \|\boldsymbol{v}^-_{j'+1/2}\|_M^2\right)\right)^{\frac{1}{2}}, \quad [\![\boldsymbol{v}]\!] = \left(\sum_{j=1}^N \|[\![\boldsymbol{v}]\!]_{j'-1/2}\|_M^2\right)^{\frac{1}{2}}.$$

Finally, the negative-order norm is defined as

$$\|\boldsymbol{v}\|_{-\ell,\Omega} = \sup_{\boldsymbol{\Phi}\in C_0^\infty(\Omega)} \frac{(\boldsymbol{v}, \boldsymbol{\Phi})}{\|\boldsymbol{\Phi}\|_\ell}. \tag{2.7}$$

Note that the negative-order norms can be used to detect the oscillations of a function around zero; for more details, see Cockburn *et al.* (2003).

2.2.2 *Local focus shifting (linearization).* Since the linearization technique is repeatedly used in analysis for nonlinear problems, we present the following inequality regarding local focus shifting (linearization) for nonlinear systems. Let $\boldsymbol{B}$ be a matrix-valued function, for example $\boldsymbol{B} = \boldsymbol{f}'(\boldsymbol{u})$, or $\partial_t \boldsymbol{f}'(\boldsymbol{u})$, which is assumed to be smooth enough with respect to $\boldsymbol{u}$. Then their focus shifting (i.e., change of the vector at which the function is evaluated) satisfies the following Lipschitz continuity

$$\|\boldsymbol{B}(\boldsymbol{w}) - \boldsymbol{B}(\boldsymbol{v})\|_M \leqslant C_\star \|\boldsymbol{w} - \boldsymbol{v}\|_M \tag{2.8}$$

due to the well-known Wielandt–Hoffman Theorem (Golub & Van Loan, 2012), where $\boldsymbol{w}$ and $\boldsymbol{v}$ are two local focuses. Note that (2.8) will be useful in our later analysis, especially for the estimates to the projection errors.

2.2.3 *Properties for the divided differences.* As indicated in Meng & Ryan (2017), one of the most important tools in deriving L^2 and negative-order norm error estimates of the divided difference for nonlinear equations is the properties of divided differences. Note that it is straightforward to extend the properties of divided differences from the scalar case to the vector/matrix case. In what follows, we only list these properties without proof and refer the readers to Meng & Ryan (2017) for more details. Specifically, we would like to list the Leibniz rule and the relation between divided differences and derivatives.

For any vector-valued functions $\boldsymbol{w}$ and $\boldsymbol{v}$, the following Leibniz rule holds

$$\partial_h^\gamma(\boldsymbol{w}(x)\boldsymbol{v}(x)) = \sum_{i=0}^{\gamma}\binom{\gamma}{i}\partial_h^i\boldsymbol{w}\left(x+\frac{\gamma-i}{2}h\right)\partial_h^{\gamma-i}\boldsymbol{v}\left(x-\frac{i}{2}h\right). \tag{2.9}$$

Note that (2.9) is still valid even if $\boldsymbol{w}$ and $\boldsymbol{v}$ are piecewise functions with possible discontinuities at cell interfaces or they are composite functions. If $\boldsymbol{w}$ is the composition of a smooth matrix-valued function $\boldsymbol{G}$ and a smooth vector-valued function $\boldsymbol{u}$, i.e., $\boldsymbol{w}(x) := \boldsymbol{G}(\boldsymbol{u}(x))$, we can prove the following property

$$\partial_h^\gamma \boldsymbol{G}(\boldsymbol{u}(x)) = \partial_x^\gamma \boldsymbol{G}(\boldsymbol{u}(x)) + C_\gamma h \boldsymbol{\Psi}_\gamma(x), \tag{2.10}$$

where C_γ is a positive constant and $\boldsymbol{\Psi}_\gamma$ is a smooth matrix-valued function. This is because the divided difference of a matrix-valued function is a matrix resulting from applying the divided difference operator to its each component, and the scalar/componentwise version of (2.10) has already been proved in Meng & Ryan (2017). We would like to remark that the property (2.10) is very useful in proving Corollary 2.1.

2.2.4 *Projections and interpolation properties.* Prior to giving the definition of Gauss–Radau projections for the system case, let us recall two kinds of scalar Gauss–Radau projections into $V_h^{\alpha,k} = \{v \in L^2(\Omega^\alpha) : v|_{I_{j'}} \in \mathcal{P}^k(I_{j'}),\ \forall j' = j + \frac{\ell}{2}, \ell = \alpha \bmod 2, j = 1,\ldots,N\}$. That is, for $q \in H^1(\Omega_h)$, the local Gauss–Radau projection of q is the unique function in $\mathcal{P}^k(I_{j'})$ such that, for each j'

$$\left(q - \mathsf{P}^-q, z_h\right)_{j'} = 0, \quad \forall z_h \in \mathcal{P}^{k-1}(I_{j'}), \qquad (q-\mathsf{P}^-q)^-_{j'+\frac{1}{2}} = 0; \tag{2.11a}$$

$$\left(q - \mathsf{P}^+q, z_h\right)_{j'} = 0, \quad \forall z_h \in \mathcal{P}^{k-1}(I_{j'}), \qquad (q-\mathsf{P}^+q)^+_{j'-\frac{1}{2}} = 0. \tag{2.11b}$$

To define the projection for the system case, we consider the Jacobian matrix $\boldsymbol{f}'(\boldsymbol{u}_{j'}) := \boldsymbol{f}'(\boldsymbol{u})|_{\boldsymbol{u}=\boldsymbol{u}_{j'}}$ with $\boldsymbol{u}_{j'} = \boldsymbol{u}(x_{j'},t)$. The corresponding eigenvalues, left and right eigenvectors are denoted by $\lambda_i, \boldsymbol{\ell}_i, \boldsymbol{r}_i$ $(i = 1,\ldots,m)$, normalized so that $\boldsymbol{\ell}_m \boldsymbol{r}_n = \delta_{m,n}$. Thus, on each cell $I_{j'}$, the Gauss–Radau projection of a vector-valued function $\boldsymbol{u}$, denoted by $\mathbb{P}\boldsymbol{u}$, is the unique function in $\boldsymbol{\mathcal{P}}^k(I_{j'})$ determined by the following procedure.

1. Transform $\boldsymbol{u}$ to the eigenspace of $\boldsymbol{f}'(\boldsymbol{u}_{j'})$, i.e.,

$$v_i = \boldsymbol{\ell}_i \boldsymbol{u}, \quad i = 1,\ldots,m.$$

2. Apply the scalar Gauss–Radau projection (2.11) to v_i in the ith characteristic field $(i = 1, \dots, m)$, and the projection $\mathsf{P}v_i$ depends on the sign of λ_i, i.e.,

$$\mathsf{P}v_i = \begin{cases} \mathsf{P}^- v_i, & \text{if } \lambda_i \geqslant 0, \\ \mathsf{P}^+ v_i, & \text{if } \lambda_i < 0. \end{cases}$$

3. The result is transformed back to the physical field to get $\mathbb{P}\boldsymbol{u}$:

$$\mathbb{P}\boldsymbol{u} = \sum_{i=1}^{m} \mathsf{P}v_i \boldsymbol{r}_i.$$

Note that the above Gauss–Radau projection has been used to derive optimal convergence results of the fully-discrete DG scheme for nonlinear systems of conservation laws, when the upwind flux is considered; see Luo *et al.* (2015).

In particular, if the flux Jacobian matrix $\boldsymbol{f}'(\boldsymbol{u})$ is always positive definite for $\boldsymbol{u}$ and x, then $\mathbb{P}\boldsymbol{u} = \boldsymbol{R}\mathsf{P}\boldsymbol{v}$ with $\boldsymbol{R} = (\boldsymbol{r}_1, \cdots, \boldsymbol{r}_m)$ and $\mathsf{P}\boldsymbol{v} = \mathsf{P}^-\boldsymbol{v} = (\mathsf{P}^- v_1, \dots, \mathsf{P}^- v_m)^{\mathrm{T}}$. Further, denoting by $\boldsymbol{\eta}_{\boldsymbol{v}} = \boldsymbol{v} - \mathsf{P}\boldsymbol{v}$ and $\boldsymbol{\eta}_{\boldsymbol{u}} = \boldsymbol{u} - \mathbb{P}\boldsymbol{u}$, we have that $\boldsymbol{\eta}_{\boldsymbol{u}} = \boldsymbol{R}\boldsymbol{\eta}_{\boldsymbol{v}}$, since $\boldsymbol{u} = \boldsymbol{R}\boldsymbol{v}$. Note that $\boldsymbol{R}$ is a constant matrix in each element $I_{j'}$ due to the local linearization $\boldsymbol{f}'(\boldsymbol{u}_{j'})$, we conclude, by the definition of scalar Gauss–Radau projection P^- in (2.11a), that for each j',

$$(\boldsymbol{u} - \mathbb{P}\boldsymbol{u}, \boldsymbol{z}_h)_{j'} = 0, \quad \forall \boldsymbol{z}_h \in \mathfrak{P}^{k-1}(I_{j'}), \qquad (\boldsymbol{u} - \mathbb{P}\boldsymbol{u})^-_{j'+\frac{1}{2}} = 0. \tag{2.12}$$

Moreover, for $\boldsymbol{u} \in \boldsymbol{W}^{k+1,\infty}(\Omega_h)$, by a standard scaling argument (Ciarlet, 1978; Brenner & Scott, 2007), we have

$$\|\boldsymbol{\eta}_{\boldsymbol{u}}\| + h\|(\boldsymbol{\eta}_{\boldsymbol{u}})_x\| + h^{1/2}\|\boldsymbol{\eta}_{\boldsymbol{u}}\|_{\Gamma_h} \leqslant Ch^{k+1}\|\boldsymbol{u}\|_{k+1}, \tag{2.13a}$$

$$\|\boldsymbol{\eta}_{\boldsymbol{u}}\|_\infty \leqslant Ch^{k+1}\|\boldsymbol{u}\|_{k+1,\infty}, \tag{2.13b}$$

where C is independent of h.

Finally, we list some inverse properties of the finite element space $\boldsymbol{V}_h^{\alpha,k}$ for the one-dimensional case. For any $\boldsymbol{q} \in \boldsymbol{V}_h^{\alpha,k}$, there exists a positive inverse constant C independent of $\boldsymbol{q}$ and h, such that

$$\text{(i) } \|\partial_x \boldsymbol{q}\| \leqslant Ch^{-1}\|\boldsymbol{q}\|; \quad \text{(ii) } \|\boldsymbol{q}\|_{\Gamma_h} \leqslant Ch^{-\frac{1}{2}}\|\boldsymbol{q}\|; \quad \text{(iii) } \|\boldsymbol{q}\|_\infty \leqslant Ch^{-\frac{1}{2}}\|\boldsymbol{q}\|.$$

2.2.5 *Properties of the DG discretization operator.* The following properties of the DG discretization operator are useful in the proof of $\boldsymbol{L}^2$ norm divided difference estimates.

LEMMA 2.1 (Meng & Ryan, 2017) Suppose that the matrix-valued function $\boldsymbol{G}(\boldsymbol{u}(x,t))$ ($\boldsymbol{G} = \boldsymbol{f}'(\boldsymbol{u}), \partial_t \boldsymbol{f}'(\boldsymbol{u})$ etc) is smooth with respect to each variable. Then, for any $\boldsymbol{w}, \boldsymbol{v} \in \boldsymbol{V}_h^{\alpha,k}$, there holds the following inequality

$$\mathcal{H}(\boldsymbol{G}\boldsymbol{w}, \boldsymbol{v}) \leqslant C_\star \left(\|\boldsymbol{w}\| + \|\boldsymbol{w}_x\| + h^{-\frac{1}{2}}\|[\![\boldsymbol{w}]\!]\| \right) \|\boldsymbol{v}\|, \tag{2.14a}$$

and in particular, if $\boldsymbol{G} = \boldsymbol{f}'(\boldsymbol{u})$ is real positive definite (and thus, $\boldsymbol{G} \geqslant \delta \boldsymbol{I}$ with $\delta > 0$ being the smallest eigenvalue of $\boldsymbol{G}$ and $\boldsymbol{I}$ the identity matrix), there holds

$$\mathcal{H}(\boldsymbol{G}\boldsymbol{w}, \boldsymbol{w}) \leqslant C_\star\|\boldsymbol{w}\|^2 - \frac{\delta}{2}\|[\![\boldsymbol{w}]\!]\|^2. \tag{2.14b}$$

Proof. The proof of (2.14a) follows by considering the equivalent *strong* form of $\mathcal{H}$, (2.5b). To prove (2.14b), we apply integration by parts to each diagonal and non-diagonal term of the quadratic form $(\boldsymbol{w}_x, \boldsymbol{G}\boldsymbol{w})$ to get the following compact form

$$
\begin{aligned}
\mathcal{H}(\boldsymbol{G}\boldsymbol{w},\boldsymbol{w}) &= -\frac{1}{2}(\boldsymbol{w},\partial_x \boldsymbol{G}\boldsymbol{w}) + \sum_{j=1}^{N}\left([\![\boldsymbol{w}]\!]^{\mathrm{T}}\boldsymbol{G}(\boldsymbol{w}^- - \{\!\{\boldsymbol{w}\}\!\})\right)_{j'-\frac{1}{2}} \\
&= -\frac{1}{2}(\boldsymbol{w},\partial_x \boldsymbol{G}\boldsymbol{w}) - \frac{1}{2}\sum_{j=1}^{N}\left([\![\boldsymbol{w}]\!]^{\mathrm{T}}\boldsymbol{G}[\![\boldsymbol{w}]\!]\right)_{j'-\frac{1}{2}} \\
&\leqslant C_\star\|\boldsymbol{w}\|^2 - \frac{\delta}{2}\|[\boldsymbol{w}]\|^2,
\end{aligned}
$$

where we have also used the Cauchy–Schwarz inequality (2.6) in the last step. □

COROLLARY 2.1 (Meng & Ryan, 2017) Under the same conditions as in Lemma 2.1, we have, for small enough h,

$$
\mathcal{H}((\partial_h^\alpha \boldsymbol{G})\boldsymbol{w},\boldsymbol{v}) \leqslant C_\star\left(\|\boldsymbol{w}\| + \|\boldsymbol{w}_x\| + h^{-\frac{1}{2}}\|[\boldsymbol{w}]\|\right)\|\boldsymbol{v}\|, \ \ \forall \alpha \geqslant 0. \tag{2.15}
$$

Proof. The proof follows by combining the relation (2.10) and (2.14a) in Lemma 2.1. □

LEMMA 2.2 Suppose that the matrix-valued function $\boldsymbol{G}(\boldsymbol{u}(x,t))$ ($\boldsymbol{G} = \boldsymbol{f}'(\boldsymbol{u}), \partial_t \boldsymbol{f}'(\boldsymbol{u})$ etc) is smooth with respect to each variable. Then, for any $\boldsymbol{u} \in \boldsymbol{H}^{k+1}(\Omega_h)$ and $\boldsymbol{z} \in \boldsymbol{V}_h^{\alpha,k}$ there holds

$$
\mathcal{H}(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u}),\boldsymbol{z}) \leqslant C_\star h^{k+2}\|\boldsymbol{z}_x\|, \tag{2.16a}
$$

$$
\mathcal{H}(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u}),\boldsymbol{z}) \leqslant C_\star h^{k+1}\|\boldsymbol{z}\|. \tag{2.16b}
$$

Proof. We need only to prove (2.16a), since, by inverse inequality (i), (2.16b) is a direct consequence. Using the exact collocation property of the projection $\mathbb{P}$ in (2.12), we have

$$
\mathcal{H}(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u}),\boldsymbol{z}) = (\boldsymbol{z}_x, \boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u})).
$$

Next, on each cell $I_{j'}$, we use the local linearization approach to rewrite $\boldsymbol{G}(\boldsymbol{u}(x,t))$ as $\boldsymbol{G}(\boldsymbol{u}) = \boldsymbol{G}(\boldsymbol{u}_{j'}) + \big(\boldsymbol{G}(\boldsymbol{u}) - \boldsymbol{G}(\boldsymbol{u}_{j'})\big)$ with $\boldsymbol{u}_{j'} = \boldsymbol{u}(x_{j'},t)$. Clearly, on each element $I_{j'}$, by (2.8), we have $\|\boldsymbol{G}(\boldsymbol{u}) - \boldsymbol{G}(\boldsymbol{u}_{j'})\|_\infty = \operatorname{ess\,sup}_{x\in\Omega}\|\boldsymbol{G}(\boldsymbol{u}) - \boldsymbol{G}(\boldsymbol{u}_{j'})\|_M \leqslant C_\star h$ due to the smoothness of $\boldsymbol{G}$ and $\boldsymbol{u}$. Using the orthogonality property of the projections $\mathbb{P}$ and P, (2.12), we arrive at

$$
\begin{aligned}
\mathcal{H}(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u}),\boldsymbol{z}) &= \big(\boldsymbol{z}_x, (\boldsymbol{G}(\boldsymbol{u}) - \boldsymbol{G}(\boldsymbol{u}_{j'}))(\boldsymbol{u} - \mathbb{P}\boldsymbol{u})\big) \\
&\leqslant \|\boldsymbol{G}(\boldsymbol{u}) - \boldsymbol{G}(\boldsymbol{u}_{j'})\|_\infty \|\boldsymbol{z}_x\| \|\boldsymbol{u} - \mathbb{P}\boldsymbol{u}\| \\
&\leqslant C_\star h^{k+2}\|\boldsymbol{z}_x\|,
\end{aligned}
$$

where we have used the Cauchy–Schwarz inequality (2.6) and the approximation property (2.13a). □

COROLLARY 2.2 Under the same conditions as in Lemma 2.2, we have, for small enough h,

$$
\mathcal{H}(\partial_h^\alpha(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u})),\boldsymbol{z}) \leqslant C_\star h^{k+1}\|\boldsymbol{z}\|, \ \ \forall \alpha \geqslant 0. \tag{2.17}
$$

Proof. The case $\alpha = 0$ has been proved in Lemma 2.2. For $\alpha \geqslant 1$, by the Leibniz rule (2.9) and using the fact that $\boldsymbol{u} - \mathbb{P}\boldsymbol{u} = \boldsymbol{R}(\boldsymbol{v} - \mathsf{P}^-\boldsymbol{v})$ with $\boldsymbol{R}$ being the matrix composed of the right eigenvectors of $\boldsymbol{f}'(\boldsymbol{u}_{j'})$ and that both the divided difference operator and the projection operator P^- are linear, we rewrite $\partial_h^{\alpha}(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u}))$ as

$$\begin{aligned}\partial_h^{\alpha}(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u})) &= \sum_{\ell=0}^{\alpha} \binom{\alpha}{\ell} \partial_h^{\ell}\boldsymbol{G}\left(x + \frac{\alpha-\ell}{2}h\right) \partial_h^{\alpha-\ell}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u})\left(x - \frac{\ell}{2}h\right) \\ &\triangleq \sum_{\ell=0}^{\alpha} \binom{\alpha}{\ell} \sum_{\gamma=0}^{\alpha-\ell} \binom{\alpha-\ell}{\gamma} \check{\boldsymbol{G}}\check{\boldsymbol{R}}(\check{\boldsymbol{v}} - \mathsf{P}^-\check{\boldsymbol{v}})\end{aligned}$$

with

$$\check{\boldsymbol{G}} = \partial_h^{\ell}\boldsymbol{G}\left(x + \frac{\alpha-\ell}{2}h\right),\ \check{\boldsymbol{R}} = \partial_h^{\gamma}\boldsymbol{R}\left(x + \frac{\alpha-2\ell-\gamma}{2}h\right),\ \check{\boldsymbol{v}} = \partial_h^{\alpha-\ell-\gamma}\boldsymbol{v}\left(x - \frac{\ell+\gamma}{2}h\right).$$

Note that we have a uniform mesh as these operators don't commute for a nonuniform mesh. Note also that $\boldsymbol{R}$ is a piecewise constant matrix in each $I_{j'}$ that depends on $\boldsymbol{f}'(\boldsymbol{u}_{j'})$. Thus,

$$\mathcal{H}\left(\partial_h^{\alpha}(\boldsymbol{G}(\boldsymbol{u} - \mathbb{P}\boldsymbol{u})), \boldsymbol{z}\right) = \sum_{\ell=0}^{\alpha} \binom{\alpha}{\ell} \sum_{\gamma=0}^{\alpha-\ell} \binom{\alpha-\ell}{\gamma} \mathcal{H}\left(\check{\boldsymbol{G}}\check{\boldsymbol{R}}(\check{\boldsymbol{v}} - \mathsf{P}\check{\boldsymbol{v}}), \boldsymbol{z}\right). \tag{2.18}$$

Clearly, by (2.10), $\check{\boldsymbol{G}}$ is also a smooth matrix-valued function with respect to each variable with the leading term $\partial_x^{\ell}\boldsymbol{G}\left(x + \frac{\alpha-\ell}{2}h\right)$. Moreover, the properties (2.12) and (2.13a) are still valid for $\check{\boldsymbol{v}} - \mathsf{P}^-\check{\boldsymbol{v}}$, since it can be regarded as the projection error of the function $\check{\boldsymbol{v}}$. However, obtaining a sharp estimate to the term $\partial_h^{\gamma}\boldsymbol{R}$ involved in $\check{\boldsymbol{R}}$ is intractable, which requires a deeper analysis. Otherwise, by directly using the definition of the divided difference (2.1), $\partial_h^{\gamma}\boldsymbol{R}$ would be of order $h^{-\gamma}$, which would inhibit any superconvergence results. Indeed, by considering eigenstructures of the matrix $\boldsymbol{G}$ and the smoothness of $\boldsymbol{f}'(\boldsymbol{u})$, we are able to prove, after careful analysis, that

$$\lim_{h\to 0} \partial_h^{\gamma}\boldsymbol{R}(x) = \partial_x^{\gamma}\boldsymbol{R}(\boldsymbol{u}(x_{j'})), \tag{2.19}$$

and thus the leading term of $\check{\boldsymbol{R}}$ is a constant matrix, which is of order h^0. To clearly display the proof of (2.19), let us consider the 2×2 ($m = 2$) matrix $\boldsymbol{G} = \boldsymbol{f}'(\boldsymbol{u})$, whose entries $g_{p,q} = \frac{\partial f_p}{\partial u_q}$ are also smooth scalar functions due to the smoothness of $\boldsymbol{f}$. It follows from the construction of the projection $\mathbb{P}$ that $\boldsymbol{R}$ is the matrix whose columns are the right eigenvectors of $\boldsymbol{f}'(\boldsymbol{u}_{j'})$, which can be expressed in terms of $g_{p,q}(\boldsymbol{u}_{j'})$ and the corresponding eigenvalues. Specifically,

$$\boldsymbol{R} = \begin{bmatrix} g_{1,2} & \lambda_2 - g_{2,2} \\ \lambda_1 - g_{1,1} & g_{2,1} \end{bmatrix},$$

where $\lambda_{1,2} = \frac{\lambda^{(1)} \pm \lambda^{(2)}}{2}$ with

$$\begin{aligned}\lambda^{(1)} &= g_{1,1} + g_{2,2}, \\ \lambda^{(2)} &= \sqrt{(g_{1,1} + g_{2,2})^2 - 4(g_{1,1}g_{2,2} - g_{1,2}g_{2,1})}.\end{aligned}$$

Therefore, in order to analyze $\partial_h^\gamma \boldsymbol{R}$, it is sufficient to consider $\partial_h^\gamma g_{p,q}$ and $\partial_h^\gamma \lambda_{1,2}$. By (2.10), we have that

$$\lim_{h\to 0} \partial_h^\gamma g_{p,q}(\boldsymbol{u}_{j'}) = \partial_x^\gamma g_{p,q}(\boldsymbol{u}_{j'}), \quad p,q=1,2, \tag{2.20a}$$

$$\lim_{h\to 0} \partial_h^\gamma \lambda^{(1)}(\boldsymbol{u}_{j'}) = \partial_x^\gamma (g_{1,1}+g_{2,2})(\boldsymbol{u}_{j'}). \tag{2.20b}$$

It remains to consider $\partial_h^\gamma \lambda^{(2)}$ if the term inside the square root of $\lambda^{(2)}$ is always positive. Otherwise, $\lambda_{1,2} = \frac{\lambda^{(1)}}{2}$. Note that $\lambda^{(2)}$ can be expressed in terms of the composition of three smooth functions, namely, $\lambda^{(2)} = z(w(\boldsymbol{u}(x)))$ with $z(w) = \sqrt{w}$ $(w > 0)$, $w(\boldsymbol{u}) = (g_{1,1}(\boldsymbol{u}) + g_{2,2}(\boldsymbol{u}))^2 - 4(g_{1,1}(\boldsymbol{u})g_{2,2}(\boldsymbol{u}) - g_{1,2}(\boldsymbol{u})g_{2,1}(\boldsymbol{u}))$ and $\boldsymbol{u} = \boldsymbol{u}(x)$. Thus, by (2.10),

$$\lim_{h\to 0} \partial_h^\gamma \lambda^{(2)}(\boldsymbol{u}_{j'}) = \partial_x^\gamma z(w(\boldsymbol{u}(x_{j'}))). \tag{2.20c}$$

The property (2.19) follows by collecting the results in (2.20a)–(2.20c). Finally, to complete the proof of this Corollary, we need only to apply the same procedure as that in the proof of Lemma 2.2 to each $\mathcal{H}$ term on the right side of (2.18). $\square$

2.2.6 *Smoothness-Increasing Accuracy-Conserving (SIAC) filters.* SIAC filters represent a family of filters designed to at least conserve the order of accuracy of the DG solution. It is a post-processing procedure. For the symmetric SIAC filter, the post-processing procedure for scalar equations was given, for example, in Cockburn *et al.* (2003); Mirzaee *et al.* (2012); Meng & Ryan (2017). Here, we concentrate on the symmetric filter. To apply the SIAC filter to systems of conservation laws, we need only to apply the filter corresponding to the scalar case to each component of the approximation vector.

The following theorem shows the relation between negative-order norm error estimates for divided differences and $\boldsymbol{L}^2$ norm of the post-processed error.

THEOREM 2.1 (Bramble & Schatz, 1977) For $0 < T < T^\star$, where $T^\star$ is the maximal time of existence of the smooth solution, let $\boldsymbol{u} \in L^\infty([0,T];\boldsymbol{H}^\nu(\Omega))$ be the exact solution of (1.1). Let $\Omega_0 + 2\mathrm{supp}(K_h^{\nu,k+1}(x)) \Subset \Omega$ and $\boldsymbol{U}$ is any approximation to $\boldsymbol{u}$, then

$$\|\boldsymbol{u}(T) - K_h^{\nu,k+1} \star \boldsymbol{U}\|_{\Omega_0} \leqslant \frac{h^\nu}{\nu!} C_1 |\boldsymbol{u}|_\nu + C_1 C_2 \sum_{\alpha \leqslant k+1} \|\partial_h^\alpha (\boldsymbol{u} - \boldsymbol{U})\|_{-(k+1),\Omega},$$

where C_1 and C_2 depend on Ω_0, k, but is independent of h.

As we can see from the above theorem, in order to have the ability to extract a superconvergent approximation using the B-spline convolution filter, we must be able to demonstrate that higher order convergence exists in the negative-order norm for not only the solution, but the divided differences as well. Since the duality argument is an important tool in deriving superconvergent negative-order norm estimates and the dual problem for nonlinear systems is a variable coefficient problem, in what follows we recall a regularity result.

LEMMA 2.3 (Hörmander, 1997; Ji *et al.*, 2013) Consider the variable coefficient system of conservation laws with a periodic boundary condition for all $t \in [0,T]$

$$\boldsymbol{\varphi}_t(x,t) + \boldsymbol{A}(x,t)\boldsymbol{\varphi}_x(x,t) = 0, \tag{2.21a}$$

$$\boldsymbol{\varphi}(x,0) = \boldsymbol{\varphi}_0(x), \tag{2.21b}$$

where $\boldsymbol{A}(x,t)$ is a given smooth matrix-valued periodic function. For any $\ell \geqslant 0$, fixed time t and $\boldsymbol{A}(x,t) \in L^{\infty}([0,T];\boldsymbol{W}^{2\ell+1,\infty}(\Omega))$, then the solution of (2.21) satisfies the following regularity property

$$\|\boldsymbol{\varphi}(x,t)\|_{\ell} \leqslant C\|\boldsymbol{\varphi}(x,0)\|_{\ell},$$

where C is a constant depending on $\|\boldsymbol{A}\|_{L^{\infty}([0,T];\boldsymbol{W}^{2\ell+1,\infty}(\Omega))}$.

3. $\boldsymbol{L}^2$ norm estimates for divided differences

In this section, we provide an analysis to the $\boldsymbol{L}^2$ norm estimates for the divided differences of the DG error, which is useful to derive superconvergent negative-order norm estimates.

3.1 *The main results in the $\boldsymbol{L}^2$ norm*

As usual, we split the DG error $\boldsymbol{e} = \boldsymbol{u} - \boldsymbol{u}_h$ into two parts, namely $\boldsymbol{e} = \boldsymbol{\eta} + \boldsymbol{\xi}$ with $\boldsymbol{\eta} = \boldsymbol{u} - \mathbb{P}\boldsymbol{u}$ being the projection error and $\boldsymbol{\xi} = \mathbb{P}\boldsymbol{u} - \boldsymbol{u}_h := \mathbb{P}\boldsymbol{e} \in \boldsymbol{V}_h^{\alpha,k}$. Here the projection $\mathbb{P}$ is defined on each cell $I_{j'}$ corresponding to the sign variation of the eigenvalues of $\boldsymbol{f}'(\boldsymbol{u})$; specifically, for any $t \in [0,T]$ and $x \in \Omega$, assuming that $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite, then on each element $I_{j'}$, we choose $\mathbb{P}\boldsymbol{u} = \boldsymbol{R}\,\mathsf{P}^{-}\boldsymbol{v}$, and thus $\boldsymbol{\eta} = \boldsymbol{R}\,\boldsymbol{\eta}_{\boldsymbol{v}}$ with $\boldsymbol{\eta}_{\boldsymbol{v}} = \boldsymbol{v} - \mathsf{P}^{-}\boldsymbol{v}$ and $\boldsymbol{v} = \boldsymbol{R}^{-1}(\boldsymbol{u}_{j'})\,\boldsymbol{u}$.

We are now ready to state the main theorem for the $\boldsymbol{L}^2$ norm error estimates.

THEOREM 3.1 For any $0 \leqslant \alpha \leqslant k+1$, let $\partial_h^{\alpha}\boldsymbol{u}$ be the exact solution of equation (2.2), which is assumed to be sufficiently smooth with bounded derivatives, and assume that $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite. Let $\partial_h^{\alpha}\boldsymbol{u}_h$ be the numerical solution of scheme (2.3) with initial condition $\partial_h^{\alpha}\boldsymbol{u}_h(0) = \mathbb{P}(\partial_h^{\alpha}\boldsymbol{u}_0)$ when the upwind flux is used. For a uniform mesh of $\Omega = (a,b)$, if the finite element space $\boldsymbol{V}_h^{\alpha,k}$ of piecewise polynomials with arbitrary degree $k \geqslant 1$ is used, then for small enough h and any $T > 0$ there holds the following error estimate

$$\|\partial_h^{\alpha}\boldsymbol{\xi}(T)\|^2 + \int_0^T [\![\partial_h^{\alpha}\boldsymbol{\xi}]\!]^2\,\mathrm{d}t \leqslant C_{\star}h^{2k+3-\alpha}, \tag{3.1}$$

where the positive constant $C_{\star}$ depends on $\boldsymbol{u}$, T and $\boldsymbol{f}$, but is independent of h.

COROLLARY 3.1 Under the same conditions as in Theorem 3.1, if in addition $\alpha \geqslant 1$ we have the following error estimates:

$$\|\partial_h^{\alpha}(\boldsymbol{u} - \boldsymbol{u}_h)(T)\| \leqslant C_{\star}h^{k+\frac{3}{2}-\frac{\alpha}{2}}. \tag{3.2}$$

Proof. Using similar argument in Corollary 2.2, we have that

$$\partial_h^{\alpha}\boldsymbol{\eta} = \sum_{\ell=0}^{\alpha}\binom{\alpha}{\ell}\check{\boldsymbol{R}}(\check{\boldsymbol{v}} - \mathsf{P}^{-}\check{\boldsymbol{v}})$$

where $\check{\boldsymbol{v}} = \partial_h^{\alpha-\ell}\boldsymbol{v}\left(x - \frac{\ell}{2}h\right)$ and $\check{\boldsymbol{R}} = \partial_h^{\ell}\boldsymbol{R}\left(x + \frac{\alpha-\ell}{2}h\right)$, and thus

$$\|\partial_h^{\alpha}\boldsymbol{\eta}\| \leqslant Ch^{k+1}\|\partial_h^{\alpha}\boldsymbol{u}\|_{k+1} \tag{3.3}$$

by the interpolation error estimate (2.13a) and the fact that the leading term of $\check{\boldsymbol{R}}$ is a constant matrix (2.19), due to the smoothness of $\boldsymbol{f}'(\boldsymbol{u})$. To complete the proof, we need only to combine (3.1) and (3.3) and use the triangle inequality. □

REMARK 3.1 We would like to point out that if we combine the error estimates from this paper with the typical divided differences for unstructured meshes, the accuracy enhancement can also be obtained for unstructured meshes. Indeed, for linear equations, Cockburn *et al.* (2003) suggests that the divided difference estimate for unstructured meshes is of order $2k+1+m-\alpha$ with $m=(2k+1)/(3k+2)$. Moreover, numerical tests for linear hyperbolic equations and unstructured meshes were carried out in Mirzaee *et al.* (2013).

To prove high order divided difference estimates in Theorem 3.1, we need first to establish a supercloseness result with $\alpha=0$. The superconvergence result for $\boldsymbol{\xi}$ (zeroth order divided difference) is given in the following proposition, which generalizes the supercloseness result from the scalar nonlinear conservation laws in Meng *et al.* (2012a) to the system case.

PROPOSITION 3.2 Let $\boldsymbol{u}$ be the exact solution of the system (1.1), which is assumed to be sufficiently smooth with bounded derivatives, and assume that $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite. Let $\boldsymbol{u}_h$ be the numerical solution of scheme (2.3) ($\alpha=0$) with initial condition $\boldsymbol{u}_h(0)=\mathbb{P}\boldsymbol{u}_0$ when the upwind flux is used. For a quasi-uniform mesh of $\Omega=(a,b)$, if the finite element space $\boldsymbol{V}_h^k$ of piecewise polynomials with arbitrary degree $k\geqslant 1$ is used, then for small enough h and any $t\in(0,T]$ there holds the following error estimates

$$\|\boldsymbol{\xi}\|+\left(\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2\,\mathrm{d}\tau\right)^{\frac{1}{2}}\leqslant C_\star h^{k+\frac{3}{2}}, \tag{3.4a}$$

$$\|\boldsymbol{\xi}_x\|\leqslant C_\star(\|\boldsymbol{\xi}_t\|+h^{k+1}), \tag{3.4b}$$

$$\|\boldsymbol{\xi}_t\|+\left(\int_0^t [\![\boldsymbol{\xi}_t]\!]^2\,\mathrm{d}\tau\right)^{\frac{1}{2}}\leqslant Ch^{k+1}+C_\star h^{-\frac{1}{2}}\left(\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2\,\mathrm{d}\tau\right)^{\frac{1}{2}}, \tag{3.4c}$$

where C and $C_\star$ depend on $\boldsymbol{u}$, t and $\boldsymbol{f}$, but is independent of h.

The proof of this proposition is given in Section 3.3.

3.2 *The proof of Theorem 3.1*

As mentioned in the introduction, the main difficulties come from estimates of $\|\partial_h^\alpha\boldsymbol{\xi}\|$ and $\|\partial_h^\alpha\boldsymbol{\eta}\|$. Using an energy analysis together with the properties of the DG discretization operator established in Section 2.2.5, we can see that the proof of Theorem 3.1 for the system case mainly follows along the same line as that for the scalar nonlinear case in Meng & Ryan (2017). Therefore, we omit detailed proofs and only point out the following two main differences

1. *Estimate of* $\|\partial_h^\alpha\boldsymbol{\eta}\|$. For scalar nonlinear equations, the estimate of $\|\partial_h^\alpha\eta\|$ is trivial, as both the divided difference operator ∂_h and the projection operator P^- are linear and thus commute with each other. However, for the system case, the projection $\mathbb{P}$ does not commute with ∂_h. As discussed in Corollary 2.2, this difficulty can be addressed by analyzing the eigenstructures of $\boldsymbol{f}'(\boldsymbol{u})$ and by using the property of the divided difference for composite functions in (2.10).

2. *Taylor expansion.* For nonlinear systems of conservation laws, in order to write out the nonlinear terms, namely $\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h)$ and $\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h^-)$, we need to use the following second order Taylor expansion

$$\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h)=\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}+\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}-\boldsymbol{e}^{\mathrm{T}}\mathsf{H}\boldsymbol{e}, \tag{3.5a}$$

$$\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h^-)=\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}^-+\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}^- -(\boldsymbol{e}^-)^{\mathrm{T}}\tilde{\mathsf{H}}\boldsymbol{e}^-. \tag{3.5b}$$

Here and below, $\boldsymbol{e}^{\mathrm{T}}\mathsf{H}\,\boldsymbol{e} := (\boldsymbol{e}^{\mathrm{T}}\mathsf{H}_1\,\boldsymbol{e},\ldots,\boldsymbol{e}^{\mathrm{T}}\mathsf{H}_m\,\boldsymbol{e})^{\mathrm{T}}$ with H_i being the Hessian matrix in the integral form of the remainders of the second order Taylor expansion, and the (p,q)-th entry of H_i given by $(\mathsf{H}_i)_{p,q} = \int_0^1 \frac{\partial^2 f_i(\boldsymbol{u}^s)}{\partial u_p \partial u_q}(1-s)\,\mathrm{d}s$ with $\boldsymbol{u}^s = \boldsymbol{u} + s(\boldsymbol{u}_h - \boldsymbol{u})$. Likewise for $(\boldsymbol{e}^-)^{\mathrm{T}}\tilde{\mathsf{H}}\,\boldsymbol{e}^-$. We would like to emphasize that the various order spatial derivatives, time derivatives and divided differences of each components of H and $\tilde{\mathsf{H}}$ are all bounded uniformly due to the smoothness of $\boldsymbol{f}$ and $\boldsymbol{u}$. Without loss of generality, we take the first order divided difference estimate $\|\partial_h \boldsymbol{\xi}\|$ for example. In order to obtain optimal $(k+1)$th order, we need only to choose $\boldsymbol{v}_h = \partial_h \boldsymbol{\xi}$ in the error equation involving the first order divided differences and use properties of the DG discretization operator in Section 2.2.5 in combination with the superconvergence error estimates in Proposition 3.2.

3.3 *The proof of Proposition 3.2*

The original DG scheme with $\alpha = 0$ is

$$((\boldsymbol{u}_h)_t, \boldsymbol{v}_h)_j = \mathcal{H}_j\left(\boldsymbol{f}(\boldsymbol{u}_h), \boldsymbol{v}_h\right), \tag{3.6}$$

which holds for all $\boldsymbol{v}_h \in \boldsymbol{V}_h^k$ and $j = 1,\ldots,N$. For periodic boundary conditions under consideration in this paper, by Galerkin orthogonality and summing over all j, we get the error equation

$$(\boldsymbol{e}_t, \boldsymbol{v}_h) = \mathcal{H}\left(\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{u}_h), \boldsymbol{v}_h\right) \tag{3.7}$$

for all $\boldsymbol{v}_h \in \boldsymbol{V}_h^k$. Letting $\boldsymbol{v}_h = \boldsymbol{\xi} = \mathbb{P}\boldsymbol{u} - \boldsymbol{u}_h$, we arrive at the following identity

$$LHS = RHS, \tag{3.8}$$

where

$$LHS = (\boldsymbol{e}_t, \boldsymbol{\xi}), \tag{3.9a}$$
$$RHS = \mathcal{H}\left(\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{u}_h), \boldsymbol{\xi}\right). \tag{3.9b}$$

Clearly,

$$LHS = \frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\xi}\|^2 + (\boldsymbol{\eta}_t, \boldsymbol{\xi}). \tag{3.10a}$$

If we now denote by $\boldsymbol{\xi}_j^c = \frac{1}{h_j}\int_{I_j} \boldsymbol{\xi}\,\mathrm{d}x$ the cell average of $\boldsymbol{\xi}$ on each element I_j, and further define piecewise constant polynomial $\boldsymbol{\xi}^c$ whose restriction on I_j is $\boldsymbol{\xi}_j^c$, then we can easily obtain a bound for $(\boldsymbol{\eta}_t, \boldsymbol{\xi})$,

$$|(\boldsymbol{\eta}_t, \boldsymbol{\xi})| = |\left(\boldsymbol{\eta}_t, \boldsymbol{\xi} - \boldsymbol{\xi}^c\right)| \leqslant Ch^{k+2}\|\boldsymbol{\xi}_x\|, \tag{3.10b}$$

since, by (2.12), $\boldsymbol{\eta}$ and thus $\boldsymbol{\eta}_t$ are orthogonal to piecewise constant functions, where in the last step we have also used the approximation error estimates (2.13a) and the Poincaré–Wirtinger inequality $\|\boldsymbol{\xi} - \boldsymbol{\xi}^c\| \leqslant Ch\|\boldsymbol{\xi}_x\|$.

In what follows, we shall estimate RHS, which is given in the following lemma.

LEMMA 3.1 Suppose that the interpolation property (2.13a) is satisfied. Then we have

$$RHS \leqslant (\mathsf{C}(e) + C_\star h^{-3}\|\boldsymbol{e}\|_\infty^2)\|\boldsymbol{\xi}\|^2 - \frac{\delta}{2}[\![\boldsymbol{\xi}]\!]^2 + C_\star h^{k+2}\|\boldsymbol{\xi}_x\| + Ch^{2k+3} \tag{3.10c}$$

with $\mathsf{C}(e) = C + C_\star h^{-1}\|\boldsymbol{e}\|_\infty$, where C and $C_\star$ are independent of h and $\boldsymbol{u}_h$.

Proof. Using the second order Taylor expansion (3.5)

$$\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h)=\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}+\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}-\boldsymbol{e}^{\mathrm{T}}\mathsf{H}\boldsymbol{e}\triangleq\boldsymbol{\theta}_1+\boldsymbol{\theta}_2+\boldsymbol{\theta}_3, \tag{3.11a}$$

$$\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h^-)=\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}^-+\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}^--(\boldsymbol{e}^-)^{\mathrm{T}}\tilde{\mathsf{H}}\boldsymbol{e}^-\triangleq\boldsymbol{\theta}_1^-+\boldsymbol{\theta}_2^-+\boldsymbol{\theta}_3^-, \tag{3.11b}$$

we rewrite *RHS* as

$$RHS=\Theta_1+\Theta_2+\Theta_3$$

with Θ_i given by

$$\Theta_i=\mathcal{H}(\boldsymbol{\theta}_i,\boldsymbol{\xi})=(\boldsymbol{\xi}_x,\boldsymbol{\theta}_i)+\sum_{j=1}^{N}([\![\boldsymbol{\xi}]\!]^{\mathrm{T}}\boldsymbol{\theta}_i^-)_{j+\frac{1}{2}},\quad (i=1,2,3),$$

which will be estimated one by one below.

By the same argument as that in the proof of (2.14b) in Lemma 2.1, we have that

$$\Theta_1\leqslant C_\star\|\boldsymbol{\xi}\|^2-\frac{\delta}{2}[\![\boldsymbol{\xi}]\!]^2. \tag{3.12a}$$

A direct application of (2.16a) in Lemma 2.2 leads to a bound for Θ_2

$$\Theta_2\leqslant C_\star h^{k+2}\|\boldsymbol{\xi}_x\|. \tag{3.12b}$$

It follows from the Cauchy–Schwarz inequality, the inverse properties (i) as well as (ii), and the approximation error estimate (2.13a), that

$$\begin{aligned}\Theta_3&\leqslant C_\star\|\boldsymbol{e}\|_\infty\left(\|\boldsymbol{e}\|\|\boldsymbol{\xi}_x\|+\|\boldsymbol{e}\|_{\Gamma_h}\|\boldsymbol{\xi}\|_{\Gamma_h}\right)\\&\leqslant C_\star h^{-1}\|\boldsymbol{e}\|_\infty\left(\|\boldsymbol{\xi}\|+\|\boldsymbol{\eta}\|+h^{\frac{1}{2}}\|\boldsymbol{\eta}\|_{\Gamma_h}\right)\|\boldsymbol{\xi}\|\\&\leqslant C_\star h^{-1}\|\boldsymbol{e}\|_\infty\|\boldsymbol{\xi}\|^2+C_\star h^k\|\boldsymbol{e}\|_\infty\|\boldsymbol{\xi}\|\\&\leqslant (C_\star h^{-1}\|\boldsymbol{e}\|_\infty+C_\star h^{-3}\|\boldsymbol{e}\|_\infty^2)\|\boldsymbol{\xi}\|^2+Ch^{2k+3},\end{aligned} \tag{3.12c}$$

where Young's inequality is used in the last step. To finish the proof of Lemma 3.1, we need only to combine (3.12a)–(3.12c). □

We now insert the estimates (3.10a)–(3.10c) into (3.8) to get

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\xi}\|^2+\frac{\delta}{2}[\![\boldsymbol{\xi}]\!]^2\leqslant(\mathsf{C}(e)+C_\star h^{-3}\|\boldsymbol{e}\|_\infty^2)\|\boldsymbol{\xi}\|^2+C_\star h^{k+2}\|\boldsymbol{\xi}_x\|+Ch^{2k+3}. \tag{3.13}$$

To deal with the nonlinearity of $\boldsymbol{f}(\boldsymbol{u})$ we make an a priori assumption that, for small enough h

$$\|\mathbb{P}\boldsymbol{u}-\boldsymbol{u}_h\|\leqslant h^2. \tag{3.14}$$

This a priori assumption can be verified by using the same argument as that in Meng *et al.* (2012a) for piecewise polynomials of degree $k\geqslant 1$, and is useful to derive a crude bound for $\boldsymbol{\xi}$, which is necessary in the proof of $\boldsymbol{\xi}_t$ in Lemma 3.2.

COROLLARY 3.2 Suppose that the interpolation property (2.13b) is satisfied, then the a priori assumption (3.14) implies that

$$\|\boldsymbol{e}\|_\infty\leqslant Ch^{\frac{3}{2}}\quad\textit{and}\quad\|\boldsymbol{\xi}\|_\infty\leqslant Ch^{\frac{3}{2}}. \tag{3.15}$$

Proof. This follows from the inverse property (iii), the interpolation property (2.13b) and triangle inequality. □

COROLLARY 3.3 Under the same conditions as in Lemma 3.1, if the a priori assumption (3.14) holds, we have the following error estimates

$$\|\boldsymbol{e}\| \leqslant Ch^{k+1} \quad \textit{and} \quad \|\boldsymbol{\xi}\| \leqslant Ch^{k+1}. \tag{3.16}$$

Proof. We first apply inverse inequality (i) to (3.10b) and (3.12b) to obtain $|(\boldsymbol{\eta}_t,\boldsymbol{\xi})| + \Theta_2 \leqslant C_\star h^{k+1}\|\boldsymbol{\xi}\|$. Then, noting (3.13), the results in Corollary 3.3 follow by using (3.15) implied by the a priori assumption (3.14) and a simple application of Gronwall's inequality together with the fact that $\boldsymbol{\xi}(\cdot,0)=0$ due to the special choice of the initial condition. □

From (3.13), one can see that the supercloseness result of $\|\boldsymbol{\xi}\|$ depends heavily on the estimate of $\|\boldsymbol{\xi}_x\|$ and further $\|\boldsymbol{\xi}_t\|$, which are given in the following two lemmas.

LEMMA 3.2 Under the same conditions as in Proposition 3.2, if, in addition, the a priori assumption (3.14) holds, we have

$$\|\boldsymbol{\xi}_x\| \leqslant C(\|\boldsymbol{\xi}_t\| + h^{k+1}), \tag{3.17}$$

for any $t \in [0,T]$, where C is independent of h and $\boldsymbol{u}_h$.

The proof of this lemma is postponed to Appendix A.1.

LEMMA 3.3 Under the same conditions as in Proposition 3.2, if, in addition, the a priori assumption (3.14) holds, we have

$$\|\boldsymbol{\xi}_t\| + \left(\int_0^t [\![\boldsymbol{\xi}_t]\!]^2 \,\mathrm{d}\tau\right)^{\frac{1}{2}} \leqslant Ch^{k+1} + C_\star h^{-\frac{1}{2}}\left(\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2 \,\mathrm{d}\tau\right)^{\frac{1}{2}}, \tag{3.18}$$

for any $t \in [0,T]$, where C and $C_\star$ are independent of h and $\boldsymbol{u}_h$.

The proof of this lemma is deferred to Appendix A.2. It is worth pointing out that, unlike the scalar case, $\|\boldsymbol{\xi}_t\|$ is bounded by $[\![\boldsymbol{\xi}]\!]$ instead of $\|\boldsymbol{\xi}\|$ in Meng *et al.* (2012a). This enables us to fully make use of properties of the DG operator established in Section 2.2.5 to deal with the mixed integral term K_1 (see Appendix A.2), which simplifies the proof a lot, and the technique based on integration by parts with respect to time as that in Meng *et al.* (2012a) is no longer needed.

Collecting the estimates (3.17) and (3.18) into (3.13) and using (3.15), we have

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\xi}\|^2 + \frac{\delta}{2}[\![\boldsymbol{\xi}]\!]^2 \leqslant C_1\|\boldsymbol{\xi}\|^2 + C_2\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2 \,\mathrm{d}\tau + C_3 h^{2k+3}, \tag{3.19}$$

where C_1, C_2 and C_3 are positive constants independent of h. Note that there holds the following identity

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2 \,\mathrm{d}\tau = [\![\boldsymbol{\xi}(t)]\!]^2.$$

Then, (3.19) becomes

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\|\boldsymbol{\xi}(t)\|^2 + \delta\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2 \,\mathrm{d}\tau\right) \leqslant C_0\left(\|\boldsymbol{\xi}(t)\|^2 + \delta\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2 \,\mathrm{d}\tau\right) + Ch^{2k+3}, \tag{3.20}$$

where $C_0 = \max(2C_1, 2C_2/\delta)$, $C = 2C_3$ are positive constants independent of h.

An application of Gronwall's inequality together with the fact that $\boldsymbol{\xi}(\cdot,0)=0$ gives us the desired result (3.4a), namely

$$\|\boldsymbol{\xi}\| + \left(\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2 \,\mathrm{d}\tau\right)^{\frac{1}{2}} \leqslant C_\star h^{k+\frac{3}{2}}. \tag{3.21}$$

To complete the proof of Proposition 3.2, we need only to combine Lemma 3.2 and Lemma 3.3.

4. Superconvergent error estimates

Although a superconvergent result about the negative-order norm estimates for the DG error itself to the scalar nonlinear conservation laws has been given in Ji *et al.* (2013), this paper goes further in that it addresses nonlinear systems and treats the estimates for both the equation itself and the divided differences of the equation. It is worth emphasizing that compared to Ji *et al.* (2013) the following superconvergent estimate about the negative-order norm of the divided differences of the DG error is more complicated and technical, as it not only needs to use the duality argument but also requires establishing the corresponding $\boldsymbol{L}^2$ norm error estimates of the divided difference as shown in Section 3.

THEOREM 4.1 For any $1 \leqslant \alpha \leqslant k+1$, let $\partial_h^\alpha \boldsymbol{u}$ be the exact solution of the problem (2.2), which is assumed to be sufficiently smooth with bounded derivatives, and assume that $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite. Let $\partial_h^\alpha \boldsymbol{u}_h$ be the numerical solution of the scheme (2.3) with initial condition $\partial_h^\alpha \boldsymbol{u}_h(0) = \mathbb{P}(\partial_h^\alpha \boldsymbol{u}_0)$ when the upwind flux is used. For a uniform mesh of $\Omega = (a,b)$, if the finite element space $\boldsymbol{V}_h^{\alpha,k}$ of piecewise polynomials with arbitrary degree $k \geqslant 1$ is used, then for small enough h and any $T>0$ there holds the following error estimate

$$\|\partial_h^\alpha(\boldsymbol{u}-\boldsymbol{u}_h)(T)\|_{-(k+1),\Omega} \leqslant C h^{2k+\frac{3}{2}-\frac{\alpha}{2}}, \tag{4.1}$$

where the positive constant C depends on $\boldsymbol{u}$, δ and T, but is independent of h.

The above negative-order norm error estimate together with Theorem 2.1 leads to a superconvergent result for the post-processed solution.

COROLLARY 4.1 Under the same conditions as in Theorem 4.1, if in addition $K_h^{\nu,k+1}$ is a convolution kernel consisting of $\nu = 2k+1+\omega$ $(\omega \geqslant \lceil -\frac{k}{2} \rceil)$ B-splines of order $k+1$ such that it reproduces polynomials of degree $\nu - 1$, then we have

$$\|\boldsymbol{u} - K_h^{\nu,k+1} \star \boldsymbol{u}_h\| \leqslant C h^{\frac{3}{2}k+1}. \tag{4.2}$$

4.1 *Proof of the main results in the negative-order norm*

As mentioned before, the negative-order norm estimates for the divided differences of the DG error depend on both the corresponding L^2 norm estimates and the duality argument. On the one hand, it is highly nontrivial to derive $\boldsymbol{L}^2$ norm error estimates of the divided differences from the standard $\boldsymbol{L}^2$ error estimates (see, e.g., Zhang & Shu, 2010; Luo *et al.*, 2015) and some delicate supercloseness results needs to be established; see Section 3. On the other hand, to perform the duality analysis, we follow the same line as that for the scalar case in Ji *et al.* (2013) and Meng & Ryan (2017). First, by (2.7), we need to concentrate on the estimate of

$$(\partial_h^\alpha(\boldsymbol{u}-\boldsymbol{u}_h)(T), \boldsymbol{\Phi}) \tag{4.3}$$

for $\boldsymbol{\Phi} \in C_0^\infty(\Omega)$. Then, define the dual problem as: find a function $\boldsymbol{\varphi}$ such that $\boldsymbol{\varphi}(\cdot,t)$ is periodic for all $t \in [0,T]$ and

$$\partial_h^\alpha \boldsymbol{\varphi}_t + \boldsymbol{f}'(\boldsymbol{u})\partial_h^\alpha \boldsymbol{\varphi}_x = 0, \qquad (x,t) \in \Omega \times [0,T), \tag{4.4a}$$

$$\boldsymbol{\varphi}(x,T) = \boldsymbol{\Phi}(x), \quad x \in \Omega. \tag{4.4b}$$

A combination of (2.2a) and (4.4a) gives us

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\partial_h^\alpha \boldsymbol{u}, \boldsymbol{\varphi}\right) + \mathcal{F}(\boldsymbol{u};\boldsymbol{\varphi}) = 0, \tag{4.5}$$

where $\mathcal{F}(\boldsymbol{u};\boldsymbol{\varphi}) = (-1)^\alpha \left(\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{u} - \boldsymbol{f}(\boldsymbol{u}), \partial_h^\alpha \boldsymbol{\varphi}_x\right)$. Thus,

$$\left(\partial_h^\alpha \boldsymbol{u}, \boldsymbol{\varphi}\right)(T) = \left(\partial_h^\alpha \boldsymbol{u}, \boldsymbol{\varphi}\right)(0) - \int_0^T \mathcal{F}(\boldsymbol{u};\boldsymbol{\varphi})\,\mathrm{d}t. \tag{4.6}$$

Consequently, for any $\boldsymbol{\kappa} \in \boldsymbol{V}_h^{\alpha,k}$, we deduce that

$$(\partial_h^\alpha(\boldsymbol{u}-\boldsymbol{u}_h)(T), \boldsymbol{\Phi}) = \mathbb{G}_1 + \mathbb{G}_2 + \mathbb{G}_3,$$

where

$$\begin{aligned}
\mathbb{G}_1 &= \left(\partial_h^\alpha(\boldsymbol{u}-\boldsymbol{u}_h), \boldsymbol{\varphi}\right)(0),\\
\mathbb{G}_2 &= -\int_0^T \left[\left((\partial_h^\alpha \boldsymbol{u}_h)_t, \boldsymbol{\varphi}-\boldsymbol{\kappa}\right) - \mathcal{H}\left(\partial_h^\alpha \boldsymbol{f}(\boldsymbol{u}_h), \boldsymbol{\varphi}-\boldsymbol{\kappa}\right)\right]\mathrm{d}t,\\
\mathbb{G}_3 &= -\int_0^T \left[\left(\partial_h^\alpha \boldsymbol{u}_h, \boldsymbol{\varphi}_t\right) + \mathcal{H}\left(\partial_h^\alpha \boldsymbol{f}(\boldsymbol{u}_h), \boldsymbol{\varphi}\right) + \mathcal{F}(\boldsymbol{u},\boldsymbol{\varphi})\right]\mathrm{d}t.
\end{aligned}$$

The estimates to $\mathbb{G}_1, \mathbb{G}_2, \mathbb{G}_3$ can be obtained essentially following the same arguments as those for the scalar case in Meng & Ryan (2017). Thus, we will only present the results here and omit detailed proofs.

LEMMA 4.1 (Projection estimate) There exists a positive constant C, independent of h, such that

$$|\mathbb{G}_1| \leqslant Ch^{2k+1}\|\partial_h^\alpha \boldsymbol{u}_0\|_{k+1}\|\boldsymbol{\varphi}(0)\|_{k+1}. \tag{4.7}$$

LEMMA 4.2 (Residual) There exists a positive constant C, independent of h, such that

$$|\mathbb{G}_2| \leqslant Ch^{2k+\frac{3}{2}-\frac{\alpha}{2}}\|\boldsymbol{\varphi}\|_{L^1([0,T];\boldsymbol{H}^{k+1})}. \tag{4.8}$$

LEMMA 4.3 (Consistency) There exists a positive constant C, independent of h, such that

$$|\mathbb{G}_3| \leqslant Ch^{2k+3-\frac{\alpha}{2}}\|\boldsymbol{\varphi}\|_{L^1([0,T];\boldsymbol{H}^{k+1})}. \tag{4.9}$$

Collecting the estimates in Lemmas 4.1–4.3 and using the regularity result in Lemma 2.3, namely $\|\boldsymbol{\varphi}\|_{k+1} \leqslant C\|\boldsymbol{\Phi}\|_{k+1}$, we get a bound for $\left(\partial_h^\alpha(\boldsymbol{u}-\boldsymbol{u}_h)(T), \boldsymbol{\Phi}\right)$

$$(\partial_h^\alpha(\boldsymbol{u}-\boldsymbol{u}_h)(T), \boldsymbol{\Phi}) \leqslant Ch^{2k+\frac{3}{2}-\frac{\alpha}{2}}\|\boldsymbol{\Phi}\|_{k+1}.$$

Thus, by (2.7), we have the bound for the negative-order norm

$$\|\partial_h^\alpha(\boldsymbol{u}-\boldsymbol{u}_h)(T)\|_{-(k+1),\Omega} \leqslant Ch^{2k+\frac{3}{2}-\frac{\alpha}{2}}.$$

This finishes the proof of Theorem 4.1.

5. Numerical examples

The superconvergent result in Corollary 4.1 suggests that a more compact kernel with fewer B-splines can achieve the theoretical superconvergence order, and the standard full kernel (a kernel function composed of a linear combination of $2k+1$ B-splines of order $k+1$) is no longer necessary. Therefore, in this section, we show the effect of using different total number of B-splines (denoted by $\nu = 2k+1+\omega$ with $\omega \geqslant \lceil -\frac{k}{2} \rceil$) of the kernel in our numerical experiments. To reduce time errors, we consider the third-order Runge–Kutta time discretization and choose a small time step. The numerical errors and convergence orders using P^2 and P^3 polynomials are given, and a specific value of $\omega = -2$ is chosen to match the superconvergence order. It is worth pointing out that a quadruple precision package is used for the post-processing procedure for P^3 polynomials in Example 5.1 and Example 5.2, which helps us to get rid of the effect of round off errors in our calculations. The numerical results are only shown for the density to save space.

EXAMPLE 5.1 Consider the one-dimensional Euler equations of compressible gas dynamics

$$\boldsymbol{u}_t + \boldsymbol{f}(\boldsymbol{u})_x = 0 \tag{5.1a}$$

with

$$\boldsymbol{u} = \begin{pmatrix} \rho \\ \rho v \\ E \end{pmatrix}, \quad \boldsymbol{f}(\boldsymbol{u}) = \begin{pmatrix} \rho v \\ \rho v^2 + p \\ v(E+p) \end{pmatrix}, \tag{5.1b}$$

where $E = \frac{p}{\gamma-1} + \frac{1}{2}\rho v^2$ and $\gamma = 1.4$ with periodic boundary conditions and the following initial conditions: $\rho(x,0) = 1 + 0.5\sin(x), v(x,0) = 1, p(x,0) = 1,\ x \in [0, 2\pi]$.

The numerical errors and orders at $T = 1$ are given in Table 5.1. From the table, we can see that the standard full kernel ($\omega = 0$) could yield at least $(2k+1)$th order superconvergence, which is similar to the results for linear hyperbolic systems in Cockburn *et al.* (2003). For the compact kernel with $\omega = -2$, superconvergence of order $2k$ can be observed. The pointwise errors are plotted in Figure 5.1, which show that the post-processed filter with the standard or the more compact kernel can both remove oscillations in the errors.

Table 5.1. L^2- and L^∞ errors for Example 5.1 (Euler equation with smooth solution). Before post-processing (left), after post-processing (middle) and post-processing with the more compact kernel (right). $T = 1$.

		Before post-processing				Post-processed ($\omega = 0$)				Post-processed ($\omega = -2$)			
	Mesh	L^2 error	Order	L^∞ error	Order	L^2 error	Order	L^∞ error	Order	L^2 error	Order	L^∞ error	Order
P^2	20	5.35E-05	–	1.83E-04	–	1.28E-06	–	1.82E-06	–	6.58E-05	–	9.30E-05	–
	40	6.69E-06	3.00	2.31E-05	2.99	2.24E-08	5.83	3.19E-08	5.83	4.14E-06	3.99	5.86E-06	3.99
	80	8.36E-07	3.00	2.89E-06	3.00	4.24E-10	5.73	6.02E-10	5.73	2.59E-07	4.00	3.67E-07	4.00
	160	1.04E-07	3.00	3.61E-07	3.00	8.91E-12	5.57	1.26E-11	5.57	1.62E-08	4.00	2.29E-08	4.00
P^3	20	1.03E-06	–	2.74E-06	–	4.94E-08	–	6.98E-08	–	1.82E-06	–	2.57E-06	–
	40	6.52E-08	3.99	1.93E-07	3.82	2.54E-10	7.60	3.60E-10	7.60	2.88E-08	5.98	4.07E-08	5.98
	80	4.03E-09	4.01	1.19E-08	4.02	1.45E-12	7.45	2.06E-12	7.45	4.50E-10	6.00	6.37E-10	6.00
	160	2.52E-10	4.00	7.43E-10	4.00	9.25E-15	7.30	1.31E-14	7.30	7.04E-12	6.00	9.95E-12	6.00

EXAMPLE 5.2 Consider the Euler equation with a source term

$$\boldsymbol{u}_t + \boldsymbol{f}(\boldsymbol{u})_x = \boldsymbol{g}(x,t)$$

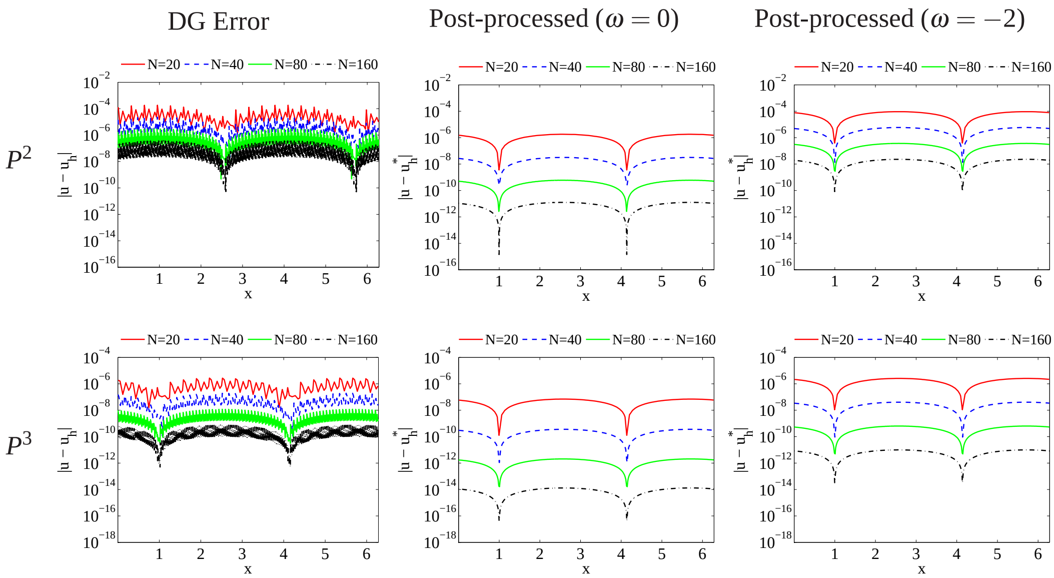

FIG. 5.1. The errors in absolute value and in logarithmic scale for P^2 (top) and P^3 (bottom) polynomials with $N = 20, 40, 80$ and 160 elements for Example 5.1 (Euler equation with smooth solution). Before post-processing (left), after post-processing (middle) and post-processing with the more compact kernel (right). $T = 1$.

with periodic boundary conditions and the following initial condition: $\rho(x,0) = 2+0.5\sin(x), v(x,0) = 1-0.1\cos(x), p(x,0) = 1$. Here, $\boldsymbol{u}$ and $\boldsymbol{f}(\boldsymbol{u})$ has been given in (5.1b), and $\boldsymbol{g}(x,t)$ is suitably chosen such that the exact solution is $\rho(x,t) = 2+0.5\sin(x+t), v(x,t) = 1-0.1\cos(x+t), p(x,t) = 1$.

The numerical errors and orders at $T = 1$ are given in Table 5.2. From the table, we can see that the orders of convergence for the standard kernel ($\omega = 0$) and the more compact kernel ($\omega = -2$) are $2k+1$ and $2k$, respectively. The pointwise errors are plotted in Figure 5.2, which show that the post-processed errors with both kernels are less oscillatory and much smaller in magnitude, and that the errors of our more compact kernel are less oscillatory than that for the standard kernel. This example demonstrates that the SIAC filter is also effective for nonlinear systems of conservation laws with source terms.

Table 5.2. L^2- and L^∞ errors for Example 5.2 (Euler equation with source terms). Before post-processing (left), after post-processing (middle) and post-processing with the more compact kernel (right). $T = 1$.

		Before post-processing				Post-processed ($\omega = 0$)				Post-processed ($\omega = -2$)			
	Mesh	L^2 error	Order	L^∞ error	Order	L^2 error	Order	L^∞ error	Order	L^2 error	Order	L^∞ error	Order
P^2	20	5.37E-05	–	1.79E-04	–	1.16E-06	–	1.80E-06	–	6.56E-05	–	9.28E-05	–
	40	6.71E-06	3.00	2.40E-05	2.90	1.92E-08	5.92	3.29E-08	5.78	4.14E-06	3.99	5.85E-06	3.99
	80	8.34E-07	3.01	3.16E-06	2.93	3.35E-10	5.84	5.69E-10	5.85	2.59E-07	4.00	3.66E-07	4.00
	160	1.04E-07	3.00	3.87E-07	3.03	6.63E-12	5.66	1.26E-11	5.50	1.62E-08	4.00	2.29E-08	4.00
P^3	20	1.10E-06	–	3.90E-06	–	9.16E-08	–	1.41E-07	–	1.80E-06	–	2.57E-06	–
	40	6.56E-08	4.07	2.30E-07	4.08	6.79E-10	7.08	1.06E-09	7.05	2.86E-08	5.98	4.06E-08	5.98
	80	4.03E-09	4.03	1.31E-08	4.13	5.24E-12	7.02	8.41E-12	6.98	4.49E-10	5.99	6.37E-10	6.00
	160	2.52E-10	4.00	8.05E-10	4.03	4.10E-14	7.00	6.63E-14	6.99	7.03E-12	6.00	9.95E-12	6.00

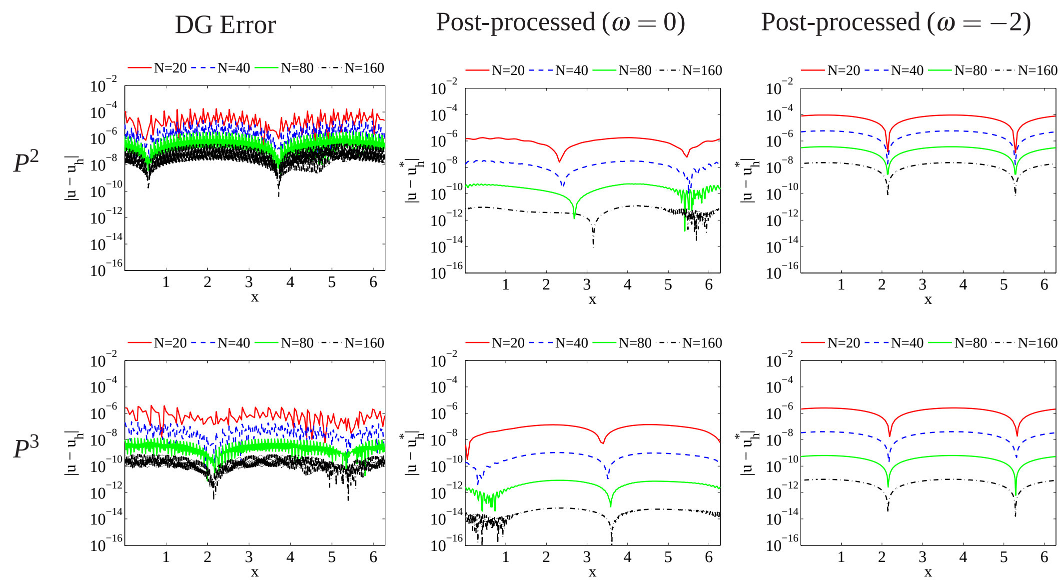

FIG. 5.2. The errors in absolute value and in logarithmic scale for P^2 (top) and P^3 (bottom) polynomials with $N = 20, 40, 80$ and 160 elements for Example 5.2 (Euler equation with source terms). Before post-processing (left), after post-processing (middle) and post-processing with the more compact kernel (right). $T = 1$.

EXAMPLE 5.3 In this example we consider the Sod problem, namely the system (5.1) with the following initial condition: $\rho(x,0) = 1, v(x,0) = 0, p(x,0) = 1$ for $x \leqslant 0$ and $\rho(x,0) = 0.125, v(x,0) = 0, p(x,0) = 0.1$ for $x > 0$, $x \in [-5,5]$.

We test the Example 5.3 at $T = 2$, when the solution contains shock and rarefaction. We measure the errors on the smooth region, $[-5,-2.6] \cup [4,5]$. The orders of convergence with different kernels are listed in Table 5.3 and pointwise errors are plotted in Figure 5.3. We can see that the post-processed errors are smaller in magnitude for most of elements. This example demonstrates that the accuracy enhancement technique is also useful for nonlinear systems of hyperbolic conservation laws with complex discontinuous solutions.

Table 5.3. L^2- and L^∞ errors in smooth regions for Example 5.3 (Sod problem with complex discontinuous solution). Before post-processing (left), after post-processing (middle) and post-processing with the more compact kernel (right). $T = 2$.

		Before post-processing				Post-processed ($\omega = 0$)				Post-processed ($\omega = -2$)			
	Mesh	L^2 error	Order	L^∞ error	Order	L^2 error	Order	L^∞ error	Order	L^2 error	Order	L^∞ error	Order
P^2	50	1.12E-03	–	8.75E-03	–	9.06E-04	–	6.84E-03	–	8.59E-04	–	6.54E-03	–
	100	3.13E-04	1.84	3.29E-03	1.41	2.35E-04	1.95	1.90E-03	1.85	1.89E-04	2.19	1.52E-03	2.11
	200	3.91E-05	3.00	3.60E-04	3.19	2.50E-05	3.23	2.41E-04	2.98	1.97E-05	3.26	1.94E-04	2.96
	400	1.28E-06	4.93	2.70E-05	3.74	9.89E-07	4.66	1.61E-05	3.91	8.01E-07	4.62	1.27E-05	3.93
P^3	50	6.14E-04	–	3.47E-03	–	2.56E-04	–	2.16E-03	–	2.66E-04	–	2.46E-03	–
	100	1.32E-04	2.22	1.44E-03	1.27	1.84E-05	3.80	1.58E-04	3.78	1.08E-05	4.62	1.59E-04	3.95
	200	1.47E-05	3.16	2.37E-04	2.60	6.81E-07	4.76	1.17E-05	3.75	5.84E-07	4.21	9.57E-06	4.05
	400	3.32E-07	5.47	5.43E-06	5.45	1.25E-08	5.77	2.64E-07	5.47	1.18E-08	5.63	2.54E-07	5.23

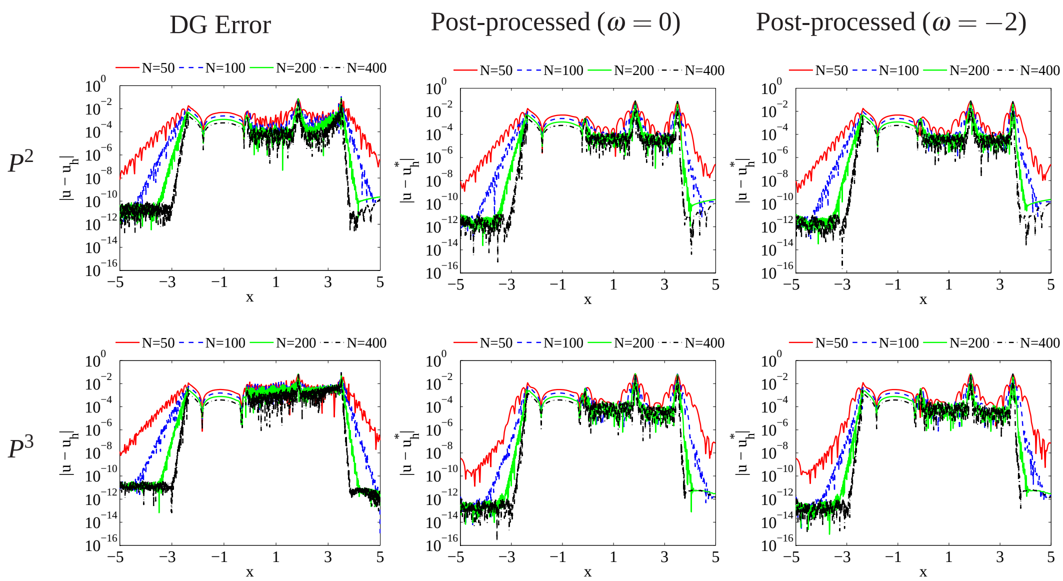

FIG. 5.3. The errors in absolute value and in logarithmic scale for P^2 (top) and P^3 (bottom) polynomials with $N = 50, 100, 200$ and 400 elements for Example 5.3 (Sod problem with discontinuous solution). Before post-processing (left), after post-processing (middle) and post-processing with the more compact kernel (right). $T = 2$.

6. Concluding remarks

In this paper, we investigate divided difference estimates and accuracy enhancement of DG methods for nonlinear symmetric systems of hyperbolic conservation laws. These estimates are essential for theoretically proving that it is possible to draw out extra orders of accuracy using a SIAC filter. The main technical difficulties come from the estimates to the divided difference of the projection error as well as the supercloseness property. By using properties of the DG discretization operator and properties of the divided differences, we are able to prove that the $\boldsymbol{L}^2$ norm of the α-th order divided difference of the DG error achieves $(k+\frac{3}{2}-\frac{\alpha}{2})$th order when upwind fluxes are used, under the condition that flux Jacobian matrix $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite. The $\boldsymbol{L}^2$ norm estimates together with a duality argument produce superconvergent negative-order norm estimates of order $2k+\frac{3}{2}-\frac{\alpha}{2}$, allowing for that the post-processed solution to be of at least $(\frac{3}{2}k+1)$th order superconvergent to the exact solution in the $\boldsymbol{L}^2$ norm. Thus, some computationally efficient more compact kernels can be used to match the proved superconvergence order in practice. A series of numerical experiments are given, showing that oscillations can be removed a lot using our more compact kernels and that the accuracy enhancement holds true for general nonlinear systems of conservation laws with different initial conditions and complex structure of solutions.

Future work consists of the study of accuracy enhancement of the DG method for nonlinear scalar and systems of conservation laws in multi-dimensional cases on structured as well as unstructured meshes. Investigation of some suitable numerical examples would also be carried out.

Funding

The first author was supported by the EU under the 7th Framework Programme Marie Curie International Incoming Fellowship FP7-PEOPLE-2013-IIF (GA-622845). The second author is supported by the Air Force Office of Scientific Research, Air Force Material Command under grant number FA8655-09-1-3055.

REFERENCES

ADJERID, S., DEVINE, K. D., FLAHERTY, J. E. & KRIVODONOVA, L. (2002) A posteriori error estimation for discontinuous Galerkin solutions of hyperbolic problems. *Comput. Methods Appl. Mech. Engrg.*, **191**, 1097–1112.

ADJERID, S. & BACCOUCH, M. (2012) A superconvergent local discontinuous Galerkin method for elliptic problems. *J. Sci. Comput.*, **52**, 113–152.

ADJERID, S. & MASSEY, T. C. (2006) Superconvergence of discontinuous Galerkin solutions for a nonlinear scalar hyperbolic problem. *Comput. Methods Appl. Mech. Engrg.*, **195**, 3331–3346.

ADJERID, S. & WEINHART, T. (2009) Discontinuous Galerkin error estimation for linear symmetric hyperbolic systems. *Comput. Methods Appl. Mech. Engrg.*, **198**, 3113–3129.

ADJERID, S. & WEINHART, T. (2011) Discontinuous Galerkin error estimation for linear symmetrizable hyperbolic systems. *Math. Comp.*, **80**, 1335–1367.

BRAMBLE, J. & SCHATZ, A. (1977) Higher order local accuracy by averaging in the finite element method. *Math. Comp.*, **31**, 94–111.

BRENNER, S. & SCOTT, R. (2007) *The mathematical theory of finite element methods.* New York: Springer.

CAO, W., ZHANG, Z. & ZOU, Q. (2014) Superconvergence of discontinuous Galerkin methods for linear hyperbolic equations. *SIAM J. Numer. Anal.*, **52**, 2555–2573.

CELIKER, F. & COCKBURN, B. (2007) Superconvergence of the numerical traces of discontinuous Galerkin and hybridized methods for convection-diffusion problems in one space dimension. *Math. Comp.*, **76**, 67–96.

CHENG, Y. & SHU, C.-W. (2010) Superconvergence of discontinuous Galerkin and local discontinuous Galerkin schemes for linear hyperbolic and convection-diffusion equations in one space dimension. *SIAM J. Numer. Anal.*, **47**, 4044–4072.

CIARLET, P. G. (1978) *The finite element method for elliptic problems*. Amsterdam: The Netherlands: North-Holland.

COCKBURN, B., LIN, S.-Y. & SHU, C.-W. (1989) TVB Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws III: one-dimensional systems. *J. Comput. Phys.*, **84**, 90–113.

COCKBURN, B., LUSKIN, M., SHU, C.-W. & SÜLI, E. (2003) Enhanced accuracy by post-processing for finite element methods for hyperbolic equations. *Math. Comp.*, **72**, 577–606.

COCKBURN, B., GUZMÁN, J. & WANG, H. (2009) Superconvergent discontinuous Galerkin methods for second-order elliptic problems. *Math. Comp.*, **78**, 1–24.

CURTIS, S., KIRBY, R. M., RYAN, J. K. & SHU, C.-W. (2007) Postprocessing for the discontinuous Galerkin method over nonuniform meshes. *SIAM J. Sci. Comput.*, **30**, 272–289.

FLOATER, M. & LYCHE, T. (2007) Two chain rules for divided differences and Faà di Bruno's formula. *Math. Comp.*, **76**, 867–877.

GOLUB, G. H. & VAN LOAN, C. F. (2012) *Matrix computations*. Baltimore: The Johns Hopkins University Press.

GUO, W., ZHONG, X. & QIU, J.-M. (2013) Superconvergence of discontinuous Galerkin and local discontinuous Galerkin methods: Eigen-structure analysis based on Fourier approach. *J. Comput. Phys.*, **235**, 458–485.

HÖRMANDER, L. (1997) *Lectures on nonlinear hyperbolic differential equations*. Berlin: Springer.

HUFFORD, C. & XING, Y. (2014) Superconvergence of the local discontinuous Galerkin method for the linearized Korteweg–de Vries equation. *J. Comput. Appl. Math.*, **255**, 441–455.

JI, L., XU, Y. & RYAN, J. K. (2012) Accuracy-enhancement of discontinuous Galerkin solutions for convection-diffusion equations in multiple-dimensions. *Math. Comp.*, **81**, 1929–1950.

JI, L., XU, Y. & RYAN, J. K. (2013) Negative-order norm estimates for nonlinear hyperbolic conservation laws. *J. Sci. Comput.*, **54**, 531–548.

LUO, J., SHU, C.-W. & ZHANG, Q. (2015) A priori error estimates to smooth solutions of the third order Runge–Kutta discontinuous Galerkin method for symmetrizable systems of conservation laws. *ESAIM: Math. Modelling Numer. Anal.*, **49**, 991–1018.

MENG, X., SHU, C.-W., ZHANG, Q. & WU, B. (2012a) Superconvergence of discontinuous Galerkin methods for scalar nonlinear conservation laws in one space dimension. *SIAM J. Numer. Anal.*, **50**, 2336–2356.

MENG, X., SHU, C.-W. & WU, B. (2012b) Superconvergence of the local discontinuous Galerkin method for linear fourth-order time-dependent problems in one space dimension. *IMA J. Numer. Anal.*, **32**, 1294–1328.

MENG, X. & RYAN, J. K. (2017) Discontinuous Galerkin methods for nonlinear scalar hyperbolic conservation laws: divided difference estimates and accuracy enhancement. *Numer. Math.*, to appear, doi:10.1007/s00211–016–0833–y.

MIRZAEE, H., JI, L., RYAN, J. K. & KIRBY, R. M. (2011) Smoothness-Increasing Accuracy-Conserving (SIAC) postprocessing for discontinuous Galerkin solutions over structured triangular meshes. *SIAM J. Numer. Anal.*, **49**, 1899–1920.

MIRZAEE, H., RYAN, J. K. & KIRBY, R. M. (2012) Efficient implementation of Smoothness-Increasing Accuracy-Conserving (SIAC) filters for discontinuous Galerkin solutions. *J. Sci. Comput.*, **52**, 85–112.

MIRZAEE, H., KING, J., RYAN, J. K. & KIRBY, R. M. (2013) Smoothness-Increasing Accuracy-Conserving filters for discontinuous Galerkin solutions over unstructured triangular meshes. *SIAM J. Sci. Comput.*, **35**, A212–A230.

RYAN, J. K., SHU, C.-W. & ATKINS, H. (2005) Extension of a post processing technique for the discontinuous Galerkin method for hyperbolic equations with application to an aeroacoustic problem. *SIAM J. Sci. Comput.*, **26**, 821–843.

RYAN, J. K. & SHU, C.-W. (2003) On a one-sided post-processing technique for the discontinuous Galerkin methods. *Methods Appl. Anal.*, **10**, 295–308.

STEFFEN, M., CURTIS, S., KIRBY, R. M. & RYAN, J. K. (2008) Investigation of Smoothness-Increasing Accuracy-Conserving filters for improving streamline integration through discontinuous fields. *IEEE–TVCG*, **14**, 680–692.

YANG, Y. & SHU, C.-W. (2012) Analysis of optimal superconvergence of discontinuous Galerkin method for linear hyperbolic equations. *SIAM J. Numer. Anal.*, **50**, 3110–3133.

ZHANG, Q. & SHU, C.-W. (2006) Error estimates to smooth solutions of Runge–Kutta discontinuous Galerkin method for symmetrizable systems of conservation laws. *SIAM J. Numer. Anal.*, **44**, 1703–1720.

ZHANG, Q. & SHU, C.-W. (2010) Stability analysis and a priori error estimates of the third order explicit Runge–Kutta discontinuous Galerkin method for scalar conservation laws. *SIAM J. Numer. Anal.*, **48**, 1038–1063.

A. Appendix: Proof of several lemmas

In this appendix, we give the proofs for some of the more technical lemmas.

A.1 *The proof of Lemma 3.2*

Let us prove the relation between $\|\boldsymbol{\xi}_x\|$ and $\|\boldsymbol{\xi}_t\|$ in Lemma 3.2. Consider the error equation (3.7), namely

$$(\boldsymbol{e}_t, \boldsymbol{v}_h) = \mathcal{H}\left(\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{u}_h), \boldsymbol{v}_h\right) \tag{A.1}$$

which holds for all $\boldsymbol{v}_h \in \boldsymbol{V}_h^k$. To deal with the nonlinearity of the flux function $\boldsymbol{f}(\boldsymbol{u})$, we use the second order Taylor expansion (3.11a) and (3.11b) to rewrite (A.1) as

$$(\boldsymbol{e}_t, \boldsymbol{v}_h) = \Theta_1 + \Theta_2 + \Theta_3 \tag{A.2}$$

with Θ_i given by

$$\Theta_i = \mathcal{H}\left(\boldsymbol{\theta}_i, \boldsymbol{v}_h\right) = ((\boldsymbol{v}_h)_x, \boldsymbol{\theta}_i) + \sum_{j=1}^{N} \left([\![\boldsymbol{v}_h]\!]^{\mathrm{T}} \boldsymbol{\theta}_i^- \right)_{j+\frac{1}{2}}, \quad (i = 1, 2, 3),$$

which will be estimated one by one below.

First consider Θ_1. We begin by using the *strong* form of $\mathcal{H}$, (2.5b), to get

$$\Theta_1 = \mathcal{H}\left(\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}, \boldsymbol{v}_h\right) = -\left(\boldsymbol{v}_h, \partial_x(\boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi})\right) - \sum_{j=1}^{N} \left((\boldsymbol{v}_h^+)^{\mathrm{T}} \boldsymbol{f}'(\boldsymbol{u})[\![\boldsymbol{\xi}]\!] \right)_{j-\frac{1}{2}}.$$

Next, let L_k be the standard Legendre polynomial of degree k in $[-1,1]$, so $L_k(-1) = (-1)^k$, and L_k is orthogonal to any polynomials of degree at most $k-1$. If we now let $\boldsymbol{v}_h = \boldsymbol{\xi}_x - \boldsymbol{d}L_k(s)$ with $\boldsymbol{d} = (-1)^k(\boldsymbol{\xi}_x)^+_{j-\frac{1}{2}}$ being a constant vector and $s = \frac{2(x-x_j)}{h_j} \in [-1,1]$, we obtain

$$\Theta_1 = -\left(\boldsymbol{v}_h, \partial_x \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}\right) - \left(\boldsymbol{\xi}_x - \boldsymbol{d}L_k(s), \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}_x\right) \triangleq -W - Z, \tag{A.3}$$

since $(\boldsymbol{v}_h)^+_{j-\frac{1}{2}} = 0$. On each element I_j, by the linearization $\boldsymbol{f}'(\boldsymbol{u}) = \boldsymbol{f}'(\boldsymbol{u}_j) + (\boldsymbol{f}'(\boldsymbol{u}) - \boldsymbol{f}'(\boldsymbol{u}_j))$ and noting $\left(\boldsymbol{d}L_k, \boldsymbol{f}'(\boldsymbol{u}_j)\boldsymbol{\xi}_x\right)_j = 0$, we arrive at an equivalent form of Z

$$Z = Z_1 + Z_2, \tag{A.4}$$

where

$$Z_1 = \left(\boldsymbol{\xi}_x, \boldsymbol{f}'(\boldsymbol{u}_j)\boldsymbol{\xi}_x\right),$$
$$Z_2 = \left(\boldsymbol{\xi}_x - \boldsymbol{d}L_k, (\boldsymbol{f}'(\boldsymbol{u}) - \boldsymbol{f}'(\boldsymbol{u}_j))\boldsymbol{\xi}_x\right).$$

By the inverse property (ii), it is easy to show, for $\boldsymbol{v}_h = \boldsymbol{\xi}_x - \boldsymbol{d}L_k(s)$, that

$$\|\boldsymbol{v}_h\| \leqslant C\|\boldsymbol{\xi}_x\|.$$

Inserting the above results into (A.2) and using the assumption that $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite (and thus $\boldsymbol{f}'(\boldsymbol{u}) \geqslant \delta \boldsymbol{I}$), we obtain

$$\delta\|\boldsymbol{\xi}_x\|^2 \leqslant Z_1 = \Theta_2 + \Theta_3 - W - Z_2 - (\boldsymbol{e}_t, \boldsymbol{\xi}_x - \boldsymbol{d}L_k). \tag{A.5}$$

We will estimate the terms on the right side of (A.5) one by one below.

A direct application of (2.16b) in Lemma 2.2 leads to a bound for Θ_2

$$|\Theta_2| \leqslant C_\star h^{k+1}\|\boldsymbol{\xi}_x\|. \tag{A.6a}$$

By an analysis similar to that in the proof of (3.12c), we get

$$|\Theta_3| \leqslant C_\star h^{-1}\|\boldsymbol{e}\|_\infty \left(\|\boldsymbol{\xi}\| + h^{k+1}\right)\|\boldsymbol{\xi}_x\|, \tag{A.6b}$$

where we have also used the approximation error estimate (2.13a). By the Cauchy–Schwarz inequality, we have

$$|W| \leqslant C_\star\|\boldsymbol{\xi}\|\|\boldsymbol{\xi}_x\|. \tag{A.6c}$$

Using the Cauchy–Schwarz inequality as well as the inverse property (i), and taking into account the fact that $\|\boldsymbol{f}'(\boldsymbol{u}) - \boldsymbol{f}'(\boldsymbol{u}_j)\|_M \leqslant C_\star h$ on each element I_j, we obtain

$$|Z_2| \leqslant C_\star\|\boldsymbol{\xi}\|\|\boldsymbol{\xi}_x\|. \tag{A.6d}$$

The triangle inequality and the approximation error estimate (2.13a) yield

$$|(\boldsymbol{e}_t, \boldsymbol{v}_h)| \leqslant C(\|\boldsymbol{\xi}_t\| + h^{k+1})\|\boldsymbol{\xi}_x\|. \tag{A.6e}$$

Finally, the error estimate (3.17) follows by collecting the estimates (A.6a)–(A.6e) into (A.5) and by using the estimates (3.15) and (3.16) in Corollary 3.2 and Corollary 3.3, respectively. This finishes the proof of Lemma 3.2.

A.2 *The proof of Lemma 3.3*

Let us first prove the initial error estimate for $\|\boldsymbol{\xi}_t(0)\|$. We start by noting that the error equation (3.7) still holds at $t = 0$ for any $\boldsymbol{v}_h \in \boldsymbol{V}_h^k$. Since $\boldsymbol{\xi}(\cdot, 0) = 0$, the nonlinear terms in (3.11a) and (3.11b) on the right-hand side of (3.7) reduce to

$$\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{u}_h) = \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta} - \boldsymbol{\eta}^{\mathrm{T}}\mathsf{H}\,\boldsymbol{\eta}, \tag{A.7a}$$
$$\boldsymbol{f}(\boldsymbol{u}) - \boldsymbol{f}(\boldsymbol{u}_h^-) = \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}^- - (\boldsymbol{\eta}^-)^{\mathrm{T}}\tilde{\mathsf{H}}\,\boldsymbol{\eta}^-. \tag{A.7b}$$

By an analysis similar to that in the proof of Lemma 3.1, we can easily get a bound for the right-hand side of (3.7) at $t=0$, denoted by $\mathcal{RHS}$; it reads

$$\mathcal{RHS} \leqslant C_\star(h^{k+1}+h^k\|\boldsymbol{\eta}(\cdot,0)\|_\infty)\|\boldsymbol{v}_h\|, \tag{A.8}$$

which holds for any $\boldsymbol{v}_h \in \boldsymbol{V}_h^k$. If we now let $\boldsymbol{v}_h = \boldsymbol{\xi}_t(\cdot,0)$ in (3.7) as well as in (A.8), we get that

$$\|\boldsymbol{\xi}_t(\cdot,0)\| \leqslant \|\boldsymbol{\eta}_t(\cdot,0)\| + C_\star(h^{k+1}+h^k\|\boldsymbol{\eta}(\cdot,0)\|_\infty) \leqslant Ch^{k+1}, \tag{A.9}$$

by the interpolation properties (2.13a) and (2.13b).

We then move on to the estimate of $\|\boldsymbol{\xi}_t(\cdot,t)\|$ for $t>0$. To do that, we proceed as follows. We take the time derivative of the error equation (3.7) and let $\boldsymbol{v}_h = \boldsymbol{\xi}_t$ to get

$$(\boldsymbol{e}_{tt},\boldsymbol{\xi}_t) = \mathcal{H}\left((\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h))_t,\boldsymbol{\xi}_t\right). \tag{A.10}$$

To estimate the right-hand side of (A.10), we use the Taylor expansion (3.11) to split the nonlinear terms as follows

$$\begin{aligned}(\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h))_t &= \partial_t \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi} + \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}_t + \partial_t \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta} + \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}_t - \boldsymbol{e}^{\mathrm{T}}\partial_t \mathsf{H}\boldsymbol{e} - 2\boldsymbol{e}^{\mathrm{T}}\mathsf{H}\,\partial_t\boldsymbol{e}\\ &\triangleq \boldsymbol{\rho}_1+\cdots+\boldsymbol{\rho}_6, \end{aligned}\tag{A.11a}$$

$$\begin{aligned}(\boldsymbol{f}(\boldsymbol{u})-\boldsymbol{f}(\boldsymbol{u}_h^-))_t &= \partial_t \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}^- + \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\xi}_t^- + \partial_t \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}^-\\ &\quad + \boldsymbol{f}'(\boldsymbol{u})\boldsymbol{\eta}_t^- - (\boldsymbol{e}^-)^{\mathrm{T}}\partial_t\tilde{\mathsf{H}}\boldsymbol{e}^- - 2(\boldsymbol{e}^-)^{\mathrm{T}}\tilde{\mathsf{H}}\,\partial_t\boldsymbol{e}^-\\ &\triangleq \boldsymbol{\rho}_1^-+\cdots+\boldsymbol{\rho}_6^-, \end{aligned}\tag{A.11b}$$

since H and $\tilde{\mathsf{H}}$ are symmetric matrices. Therefore, the right-hand side of (A.10), denoted by Υ, can be formulated as

$$\Upsilon = \mathsf{K}_1+\cdots+\mathsf{K}_6 \tag{A.12}$$

with $\mathsf{K}_i = \mathcal{H}(\boldsymbol{\rho}_i,\boldsymbol{\xi}_t)$ $(i=1,\ldots,6)$. Consequently, (A.10) can be represented by

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\xi}_t\|^2 \leqslant \Upsilon + \|\boldsymbol{\eta}_{tt}\|\|\boldsymbol{\xi}_t\| \leqslant \Upsilon + Ch^{k+1}\|\boldsymbol{\xi}_t\|, \tag{A.13}$$

by the interpolation error estimate (2.13a).

We estimate the term K_1 first. A simple application of (2.14a) in Lemma 2.1 gives us a bound for K_1; it reads

$$\begin{aligned}\mathsf{K}_1 &\leqslant C_\star\left(\|\boldsymbol{\xi}\|+\|\boldsymbol{\xi}_x\|+h^{-\frac12}\|[\boldsymbol{\xi}]\|\right)\|\boldsymbol{\xi}_t\|\\ &\leqslant C_\star\left(\|\boldsymbol{\xi}_t\|+h^{k+1}+h^{-\frac12}\|[\boldsymbol{\xi}]\|\right)\|\boldsymbol{\xi}_t\|\\ &\leqslant C_\star\|\boldsymbol{\xi}_t\|^2+h^{-1}\|[\boldsymbol{\xi}]\|^2+Ch^{2k+2}, \end{aligned}\tag{A.14a}$$

where we have used (3.16) and (3.17) in the second step and Young's inequality in the last step. Next, a direct application of (2.14b) in Lemma 2.1 leads to a bound for K_2

$$\mathsf{K}_2 \leqslant C_\star\|\boldsymbol{\xi}_t\|^2 - \frac{\delta}{2}\|[\boldsymbol{\xi}_t]\|^2, \tag{A.14b}$$

where we have used the assumption that $\boldsymbol{f}'(\boldsymbol{u})$ is positive definite with the smallest eigenvalue δ. To estimate K_3 and K_4, we need only to employ the property (2.16b) in Lemma 2.2, which is given as follows

$$\mathsf{K}_3 + \mathsf{K}_4 \leqslant C_\star h^{k+1} \|\boldsymbol{\xi}_t\|, \tag{A.14c}$$

where we have also used the fact that $\mathbb{P}$ is a linear operator with respect to t, namely $(\mathbb{P}\boldsymbol{u})_t = \mathbb{P}(\boldsymbol{u}_t)$, and thus $\|(\boldsymbol{u} - \mathbb{P}\boldsymbol{u})_t\| \leqslant Ch^{k+1}\|\boldsymbol{u}_t\|_{k+1}$ by the approximation error estimate (2.13a). It is easy to show, for high order terms K_5 and K_6, that

$$\mathsf{K}_5 \leqslant C_\star h^{-1}\|e\|_\infty(\|\boldsymbol{\xi}\| + h^{k+1})\|\boldsymbol{\xi}_t\| \leqslant C_\star h^{k+1}\|\boldsymbol{\xi}_t\|, \tag{A.14d}$$

$$\mathsf{K}_6 \leqslant C_\star h^{-1}\|e\|_\infty(\|\boldsymbol{\xi}_t\| + h^{k+1})\|\boldsymbol{\xi}_t\| \leqslant C_\star \|\boldsymbol{\xi}_t\|^2 + Ch^{k+1}\|\boldsymbol{\xi}_t\|, \tag{A.14e}$$

where we have also employed (3.15) and (3.16) in Corollary 3.2 and Corollary 3.3 in the last inequality. Therefore, by collecting the estimates (A.14a)–(A.14e) into (A.12) and (A.13), we get, after a simple application of Young's inequality, that

$$\frac{1}{2}\frac{\mathrm{d}}{\mathrm{d}t}\|\boldsymbol{\xi}_t\|^2 + \frac{\delta}{2}[\![\boldsymbol{\xi}_t]\!]^2 \leqslant C_\star\|\boldsymbol{\xi}_t\|^2 + h^{-1}[\![\boldsymbol{\xi}]\!]^2 + Ch^{2k+2}, \tag{A.15}$$

where C and $C_\star$ are positive constants independent of h. Finally, a direct application of Gronwall's inequality together with the initial error estimate (A.9) leads to the desired result

$$\|\boldsymbol{\xi}_t\| + \left(\int_0^t [\![\boldsymbol{\xi}_t]\!]^2 \,\mathrm{d}\tau\right)^{\frac{1}{2}} \leqslant Ch^{k+1} + C_\star h^{-\frac{1}{2}}\left(\int_0^t [\![\boldsymbol{\xi}(\tau)]\!]^2 \,\mathrm{d}\tau\right)^{\frac{1}{2}}.$$

This completes the proof of Lemma 3.3.