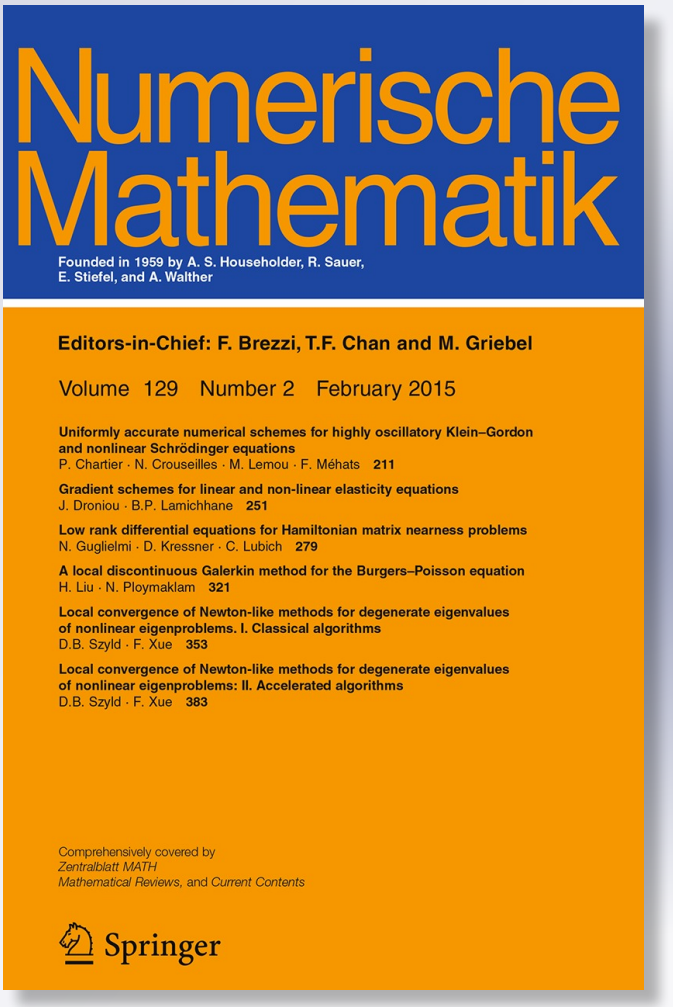
# *A local discontinuous Galerkin method for the Burgers–Poisson equation*

**Hailiang Liu & Nattapol Ploymaklam**

**Numerische Mathematik**

ISSN 0029-599X
Volume 129
Number 2

Numer. Math. (2015) 129:321-351
DOI 10.1007/s00211-014-0641-1




Springer

Numer. Math. (2015) 129:321–351
DOI 10.1007/s00211-014-0641-1

Numerische Mathematik

# A local discontinuous Galerkin method for the Burgers–Poisson equation

**Hailiang Liu · Nattapol Ploymaklam**

Received: 20 September 2013 / Revised: 26 March 2014 / Published online: 8 June 2014


**Abstract** In this work, we design, analyze and test a local discontinuous Galerkin method for solving the Burgers–Poisson equation. This model, proposed by Whitham [Linear and nonlinear waves, 1974] as a simplified model for shallow water waves, admits conservation of both momentum and energy as two invariants. The proposed numerical method is high order accurate and preserves two invariants, hence producing solutions with satisfying long time behavior. The $L^2$-stability of the scheme for general solutions is a consequence of the energy preserving property. The optimal order of accuracy for polynomial elements of even degree is proven. A series of numerical tests is provided to illustrate both accuracy and capability of the method.

**Mathematics Subject Classification (2000)** 65M60 · 65M12 · 35Q53

## 1 Introduction

In this paper, we are interested in numerical approximations to the Burgers–Poisson (BP) equation of the form

$$u_t + \left(\frac{u^2}{2} - \phi\right)_x = 0, \tag{1a}$$

$$\phi_{xx} - \phi = u. \tag{1b}$$

H. Liu (✉) · N. Ploymaklam
Mathematics Department, Iowa State University, Ames, IA 50011, USA
e-mail: hliu@iastate.edu

N. Ploymaklam
e-mail: nattapol@iastate.edu

The subscript $t$ (or $x$, respectively) denotes the differentiation with respect to time variable $t$ (or spatial variable $x$), where $u$ and $\phi$ depend on $(t, x) \in (0, \infty) \times \mathbb{R}$. System (1) can be rewritten as a nonlocal equation

$$u_t + \left(\frac{u^2}{2} + G * u\right)_x = 0 \tag{2}$$

with the kernel $G(x) = \frac{1}{2}e^{-|x|}$. This nonlocal model was found as a simplified shallow water model by Whitham [34] to approximate the model with a singular kernel

$$G(x) = \frac{1}{2\pi}\int_{\mathbb{R}} \left(\frac{\tanh k}{k}\right)^{1/2} e^{ikx}\, dk.$$

For (2) with initial data $u_0 \in BV(\mathbb{R})$, it is shown in [16] that there exists a unique global weak solution $u \in L^\infty_{\text{loc}}([0, \infty); BV(\mathbb{R}))$. For smooth initial data $u_0 \in H^s(\mathbb{R})$ with $s > 3/2$, there exists a unique smooth solution $u \in L^\infty((0, T); H^s(\mathbb{R})) \cap C((0, T); H^{s-1}(\mathbb{R}))$, at least for some finite time $T$. Furthermore, analysis of traveling waves in [16] shows that there are three generic cases of wave patterns, including solitary waves, peaked periodic waves, and shock waves, and the set of pairs $(u_{-\infty}, u_\infty)$ can be connected by a shock wave only when $u_\infty - u_{-\infty} \geq 2$.

In this work, we develop a local discontinuous Galerkin (LDG) method to solve this nonlinear BP equation with initial data $u_0(x)$, posed on a bounded domain $[0, L]$, with periodic boundary conditions. For other type of boundary conditions, the method can be modified to incorporate the specified boundary condition through suitable boundary numerical fluxes, while still using the conserved numerical fluxes for other cell interfaces.

Our proposed scheme is high order accurate, and preserves two invariants of momentum and energy, hence producing solutions with satisfying long time behavior. The $L^2$-stability of the scheme for general solutions is a consequence of the energy preserving property.

In the context of water waves, one of the best known local models is probably the Korteweg de Vries (KdV)-equation,

$$u_t + uu_x + u_{xxx} = 0.$$

This equation possesses soliton solutions' coherent structures that interact nonlinearly among themselves, then reemerge, retaining their identity and showing particle-like scattering behavior. In shallow water wave theory, the nonlinear shallow water equations which neglect dispersion altogether lead to the finite time wave breaking. On the other hand the third order derivative term in the KdV equation will prevent this ever happening in its solutions. In reality, some water waves appear to break, if the wave height is above certain threshold. Therefore in [34] Whitham raised an intriguing question: what kind of mathematical equation can describe waves with breaking? He suggested equation (2) with the above two kernels; many competing models have since

been suggested to capture one aspect or another of the classical water-wave problem, see e.g., [4,5,14,15,17,21,22,28,29].

One common feature of these models is the associated global invariants, infinitely many or finitely many. The BP equation preserves both the momentum and the energy, that is, it has the following two global invariants

$$\int u(0,x)\,dx = \int u(t,x)\,dx =: E_1(t), \tag{3a}$$

$$\int u^2(0,x)\,dx = \int u^2(t,x)\,dx =: E_2(t). \tag{3b}$$

It is desirable to design stable and high order accurate numerical schemes which preserve two invariants for solving the BP equation. It is believed that numerical methods preserving more invariants are advantageous: besides the high accuracy of numerical solutions, an invariant preserving scheme can preserve good stability properties after long-time numerical integration. Much more effort has been devoted in this topic for different integrable PDEs recently [7,18,30,31].

The goal of this paper is to develop a discontinuous Galerkin (DG) method to preserve both momentum and energy at the discrete setting. The DG method is a class of finite element methods using completely discontinuous piecewise-polynomials for the numerical solution and the test functions. It was first designed and has been successful for solving first order PDEs such as nonlinear conservation laws [9–11,13,32]. The local discontinuous Galerkin (LDG) method is an extension of the DG method for solving higher order PDEs. It was first designed for convection-diffusion equations [12], and has been extended to other higher order wave equations, including the KdV equation [25,37–39] and the Camassa–Holm equation [35], see also the recent review paper [36] on the LDG methods for higher order PDEs. The idea of the LDG method is to rewrite higher order equations into a first order system, and then apply the DG method on the system. In contrast, the direct discontinuous Galerkin (DDG) methods, proposed in [26,27] primarily for diffusion equations, aimed at directly solving higher order PDEs by the DG discretization, see e.g., [2,40] for energy preserving DG methods for KdV type equations, and [24] for the Degasperis–Procesi equation. The DDG method, as another class of DG methods for higher order partial differential equations, is to directly force the weak solution formulation of the PDE into the DG function space for both the numerical solution and test functions. Unlike the traditional LDG method, the DDG method does not rewrite the original equation into a larger first order system. The main novelty in the DDG schemes proposed in [26,27] lies in numerical flux choices for the solution gradient, which involves higher order derivatives evaluated crossing cell interfaces.

In this work we propose an LDG method based on formulation (1), for which the second equation was rewritten into a first order system for applying the LDG discretization. In the algorithm we update the solution in two steps: (1) given $u$, obtain $\phi$ by solving (1b) with the LDG method; (2) with the obtained $\phi$, update $u$ by solving (1a) with a standard DG method using a conservative numerical flux so that the resulting scheme preserves two integrals $E_1$ and $E_2$ in smooth region.

As for error estimates, we define a global projection dictated by the selected numerical flux and obtain the needed projection error, following the strategy of error estimates carried out in [23] for the DDG method to solve convection-diffusion equations. Through careful estimates using this global projection, we obtain the optimal order of accuracy for polynomial elements of even degree. This is confirmed by the numerical tests with $k = 2, 4$. Numerical tests also show that for $k$ odd, only $k$-th order of accuracy is observed. Such an optimal error estimate only for $k =$ even was also shown in [2], and numerically observed in [2,40] for KdV type equations. The main feature of the scheme presented in this work is its capability to produce wave solutions with satisfying long time behavior.

We want to point out that our estimates apply only for smooth solutions. However, for some initial configuration, the BP equation may admit discontinuous solutions at finite time, and beyond that time weak solutions need to be considered. A question is that in what sense our high order LDG methods mean for weak solutions in large times. Some rigorous $C_t^0(L_x^1)$ estimate would be desirable to understand this issue. A recent example of this type of estimates can be found in [1] for well-balance schemes on non-resonant scalar balance laws.

The rest of the paper is organized as follows. In Sect. 2, we formulate our LDG method with a class of numerical fluxes. In Sect. 3, we show that the LDG method for solving the Poisson equation (1b) is well defined and stable, and the method is shown to conserve both momentum and energy for the given numerical fluxes. In Sect. 4, we obtain the optimal order of error between the numerical solution and smooth solutions for the conservative scheme when using polynomial elements of even degree. Finally, in Sect. 5, we present numerical examples to illustrate the capacity of the LDG scheme to preserve two invariants after a long-time simulation.

## 2 The discontinuous Galerkin method

### 2.1 LDG formulation

We develop a local discontinuous Galerkin (LDG) method for the BP equation subject to initial data $u_0(x)$, posed on $I = [0, L]$ with periodic boundary conditions. Let us partition the interval $I$ into $0 = x_{1/2}, x_{3/2}, \ldots, x_{N+1/2} = L$ to get $N$ equal subintervals and denote each cell by $I_j = [x_{j-1/2}, x_{j+1/2}]$, $j = 1, \ldots, N$. The center of the cell is $x_j = \frac{1}{2}(x_{j-1/2} + x_{j+1/2})$. Here the uniform mesh is taken just for simplicity of the analysis; one may well use non-uniform meshes in the implementation of the method.

The piecewise polynomial space $V_h^k$ is defined as the space of polynomials of degree up to $k$ in each cell $I_j$, that is,

$$V_h^k = \{v : v|_{I_j} \in P^k(I_j), j = 1, 2, \ldots, N\}. \tag{4}$$

Note that functions in $V_h^k$ are allowed to have discontinuities across the interfaces. The solution of the DG method is denoted by $u_h$, which belongs to the finite element space $V_h^k$. We denote the limit values of $u_h$ at $x_{j+1/2}$ from the right and from the left by $(u_h)^+_{j+1/2}$ and $(u_h)^-_{j+1/2}$ respectively. Let $\omega$ be a piecewise smooth function, its

jump across the cell interface be denoted by $[\omega] := \omega^+ - \omega^-$, and its average at the cell interface, $\frac{\omega^+ + \omega^-}{2}$, be denote by $\{\omega\}$.

To define the LDG method, we introduce an auxiliary variable $p = \phi_x$ and rewrite (1a)–(1b) as follows:

$$u_t + \left(\frac{u^2}{2}\right)_x - p = 0, \tag{5a}$$

$$p - \phi_x = 0, \tag{5b}$$

$$p_x - \phi - u = 0. \tag{5c}$$

Then, the scheme is defined as follows: find $u_h, p_h, \phi_h \in V_h^k$ such that

$$\int_{I_j} (u_h)_t \rho \, dx - \int_{I_j} \frac{u_h^2}{2} \rho_x \, dx + \widehat{\frac{u_h^2}{2}} \rho \,|_{\partial I_j} - \int_{I_j} p_h \rho \, dx = 0, \tag{6a}$$

$$\int_{I_j} p_h \gamma \, dx + \int_{I_j} \phi_h \gamma_x \, dx - \widehat{\phi_h} \gamma \,|_{\partial I_j} = 0, \tag{6b}$$

$$-\int_{I_j} p_h q_x \, dx + \widehat{p_h} q \,|_{\partial I_j} - \int_{I_j} (\phi_h + u_h) q \, dx = 0, \tag{6c}$$

$$\int_{I_j} (u_h - u)|_{t=0} v \, dx = 0, \tag{6d}$$

for all test functions $\rho, \gamma, q, v$ in the finite element space $V_h^k$. The choice for numerical fluxes $\widehat{u_h^2}, \widehat{\phi_h}, \widehat{p_h}$ is given by

$$\widehat{u_h^2} = \frac{1}{3}\left((u_h^+)^2 + u_h^+ u_h^- + (u_h^-)^2\right), \tag{7a}$$

$$\widehat{\phi_h} = \theta \phi_h^+ + (1-\theta)\phi_h^-, \tag{7b}$$

$$\widehat{p_h} = (1-\theta) p_h^+ + \theta p_h^-, \tag{7c}$$

where $\theta \in [0, 1/2]$. Here, the numerical fluxes at the endpoints of $I$ can be defined using $U_{1/2}^- := U_{N+1/2}^-$ and $U_{N+1/2}^+ := U_{1/2}^+$ where $U$ represents $u_h^2$, $\phi_h$, or $p_h$. The resulting LDG scheme (6) subject to the fluxes (7) with $\theta = 1/2$ is called LDG-C.

For discontinuous solutions, an entropy flux for $\widehat{u_h^2}$ is needed in order to capture the entropy solution. One well-known choice is the Lax–Friedrich flux of the form

$$\widehat{u_h^2} = \frac{1}{2}\left((u_h^-)^2 + (u_h^+)^2 - \sigma(u_h^+ - u_h^-)\right), \ \sigma = 2 \max_{u \in [u_h^-, u_h^+]} |u|, \tag{8}$$

with which the resulting LDG scheme is called LDG-D.

In practice, one may adopt an adaptive numerical flux

$$\widehat{u_h^2} = \begin{cases} \frac{1}{3}\left((u_h^+)^2 + u_h^+ u_h^- + (u_h^-)^2\right) & \text{if } |u^+ - u^-| < 10^{-2} \\ \frac{1}{2}\left((u_h^-)^2 + (u_h^+)^2 - 2\sigma(u_h^+ - u_h^-)\right) & \text{otherwise.} \end{cases} \tag{9}$$

Here, $10^{-2}$ may vary as long as it can serve as a shock detector. The resulting scheme is called LDG-Ad.

*Remark 2.1* Such a dissipative numerical flux is sufficient for the scheme to capture shocks at the cell interfaces. In practice, shock may well occur in the interior of computational cells, and a limiter is necessary to be imposed, as a result the approximation degenerates to first-order around shocks. In this work we use the TVBM limiter introduced by Cockburn and Shu [11].

Before concluding this section, we outline the algorithm to compute the numerical solution.

## 2.2 Algorithm

1. We use $U_h$ to denote the vector containing the degree of freedom for $u_h$. We compute both $\psi$ and $p$ from solving the coupled system (6b), (6c)

$$\Phi_h = A_{1-\theta}\Phi_h - U_h, \quad P_h = A_\theta \Phi_h. \tag{10}$$

2. Given $u_h$ only, the coupled system is wellposed for $\theta \in [0, 1]$ and leads to

$$\Phi_h = -(I - A_{1-\theta}A_\theta)^{-1}U_h, \quad P_h = -A_\theta(I - A_{1-\theta}A_\theta)^{-1}U_h,$$

which when substituted into (6a) gives a closed ODE system for $u_h$:

$$\frac{d}{dt}U_h = -\frac{1}{2}D(u_h^2) + A_\theta(I - A_{1-\theta}A_\theta)^{-1}U_h, \tag{11}$$

where $D(u_h^2)$ denotes the vector containing the degree of freedom of the DG differentiation of $u_h^2$ with the numerical flux (7a).
3. We use a time discretization method to solve the obtained semi-discrete system for $u_h$.

This algorithm indicates that it is important that the coupled system (6b), (6c) is well-posed, and we will show this in Sect. 3.

**Notation.** We use $\|\cdot\|_{m,\Omega}$ as the $H^m$-norm over domain $\Omega$, and $|\cdot|_{m,\Omega}$ as its semi-norm. For $m = 0$, we simply use $\|\cdot\|_\Omega$ to denote the $L^2$-norm over domain $\Omega$. We also use the notation $\|\cdot\|_{\infty,\Omega}$ to denote the $L^\infty$ norm over domain $\Omega$. The domain $\Omega$ could be a computational cell $I_j$ or a master domain $\hat{I} := [-1, 1]$. If $\Omega$ is the whole domain,

we do not specify the domain unless necessary. For piecewise smooth function we use the same notation to denote contributions from all cells, for example

$$\|\omega\|_m^2 = \sum_{j=1}^{N} \|\omega\|_{m,I_j}^2.$$

## 3 Analytical properties of the scheme

### 3.1 Existence, uniqueness, and stability

In this section, we prove the existence, uniqueness, and stability of $p_h$, $\phi_h$ obtained from (6b)–(6c) with numerical fluxes (7b)–(7c), given $u_h$.

**Lemma 3.1** *The numerical scheme* (6b)–(6c) *with the numerical flux* (7b)–(7c) *for any* $\theta \in [0, 1]$ *satisfies*

$$2\|p_h\|^2 + \|\phi_h\|^2 \le \|u_h\|^2. \tag{12}$$

*Proof* We choose $\gamma = p_h$ and $q = \phi_h$. Then (6b)–(6c) gives

$$\int_{I_j} p_h^2\,dx + \int_{I_j} \phi_h (p_h)_x\,dx - \widehat{\phi_h} p_h \,|_{\partial I_j} = 0,$$
$$-\int_{I_j} p_h (\phi_h)_x\,dx + \widehat{p_h}\phi \,|_{\partial I_j} - \int_{I_j} \phi_h^2\,dx = \int_{I_j} u_h \phi_h\,dx.$$

Subtracting the two equations above gives

$$\int_{I_j} \left(p_h^2 + \phi_h^2\right) dx$$
$$= -\int_{I_j} \phi_h (p_h)_x\,dx + \widehat{\phi_h} p_h \,|_{\partial I_j} - \int_{I_j} p_h (\phi_h)_x\,dx + \widehat{p_h}\phi \,|_{\partial I_j} - \int_{I_j} u_h \phi_h\,dx.$$

Take summation over $j$ and use the periodic boundary condition to get

$$\int_I \left(p_h^2 + \phi_h^2\right) dx = -\int_I \phi_h (p_h)_x\,dx - \sum_j \left(\widehat{\phi_h}[p_h]\right)_{j+\frac{1}{2}} - \int_I p_h (\phi_h)_x\,dx$$
$$-\sum_j \left(\widehat{p_h}[\phi]\right)_{j+\frac{1}{2}} - \int_I u_h \phi_h\,dx.$$

The first four terms on the right-hand side can be simplified to

$$-\int_I (\phi_h p_h)_x \, dx - \sum_j \left(\widehat{\phi_h}[p_h] + \widehat{p_h}[\phi]\right)_{j+\frac{1}{2}} = \sum_j \left([\phi_h p_h] - \widehat{\phi_h}[p_h] - \widehat{p_h}[\phi]\right)_{j+\frac{1}{2}} = 0,$$

because of the choice of numerical fluxes (7b)–(7c). Therefore, we have that

$$\int_I \left(p_h^2 + \phi_h^2\right) dx \le \int_I |u_h \phi_h| \, dx \le \frac{1}{2}\|u_h\|^2 + \frac{1}{2}\|\phi_h\|^2,$$

which proves (12). □

*Remark 3.1* The inequality (12) shows that (6b) and (6c) produce a unique pair $(p_h, \phi_h)$ for any given $u_h$.

### 3.2 Discrete conservation laws

In this section, we look at the properties of the numerical solution $u_h$ that are analogous to (3a)–(3b).

**Theorem 3.2** *For the LDG scheme* (6) *subject to numerical fluxes* (7) *with any* $\theta \in [0, 1/2]$*, the following relations hold for all* $t > 0$*:*

$$\int_0^L u_h(t,x)dx = \int_0^L u_h(0,x)dx, \tag{13}$$

$$\int_0^L u_h^2(t,x)dx = \int_0^L u_h^2(0,x)dx + (2\theta - 1)\sum_j \int_0^t \left([\phi_h]^2 + [p_h]^2\right)_{j+\frac{1}{2}} d\tau. \tag{14}$$

*Hence the scheme is conservative for* $\theta = 1/2$*, and the scheme is energy stable for* $0 \le \theta < 1/2$.

*Remark 3.2* For solutions with discontinuities, we use the numerical flux (8) or (9) together with (7b), (7c) with $\theta = 1/2$ so that the quadratic entropy dissipates at admissible discontinuities. Our numerical tests indicate that the choice with $\theta \in (0, 1/2)$ works as well.

*Proof* Because (6) holds for any test function in $V_h^k$, we choose $\rho = 1$ and $\gamma = 1$ in (6a)–(6b), respectively, to obtain

$$\int_{I_j} (u_h)_t \, dx + \widehat{\frac{u_h^2}{2}}\Big|_{\partial I_j} - \widehat{\phi_h}\Big|_{\partial I_j} = 0.$$

Take summation over all $j$ and use the periodic boundary condition, we have

$$\frac{d}{dt}\int_I u_h\,dx = -\sum_j \left(\widehat{\frac{u_h^2}{2}}\,|_{\partial I_j} - \widehat{\phi_h}\,|_{\partial I_j}\right) = 0,$$

This proves (13).

Next, we choose the test functions $\rho = u_h$, $\gamma = -\phi_h$, and $q = -p_h$ in (6a)–(6c) to obtain

$$\int_{I_j} (u_h)_t\, u_h\,dx - \int_{I_j} \frac{u_h^2}{2}\,(u_h)_x\,dx + \widehat{\frac{u_h^2}{2}} u_h\,|_{\partial I_j} - \int_{I_j} p_h u_h\,dx = 0, \tag{15}$$

$$-\int_{I_j} p_h\phi_h\,dx - \int_{I_j} \phi_h\,(\phi_h)_x\,dx + \widehat{\phi_h}\phi_h\,|_{\partial I_j} = 0, \tag{16}$$

$$\int_{I_j} p_h\,(p_h)_x\,dx - \widehat{p_h} p_h\,|_{\partial I_j} + \int_{I_j} (\phi_h + u_h) p_h\,dx = 0, \tag{17}$$

Integrating some terms out and adding the above three relations together, we get

$$\int_{I_j} (u_h)_t\, u_h\,dx + \left(\widehat{\frac{u_h^2}{2}} u_h - \frac{u_h^3}{6}\right)|_{\partial I_j} + \left(\widehat{\phi_h}\phi_h - \frac{\phi_h^2}{2}\right)|_{\partial I_j} - \left(\widehat{p_h} p_h - \frac{p_h^2}{2}\right)|_{\partial I_j} = 0.$$

Summing the terms above for all $j = 1, \ldots, N$ and using the periodic boundary condition, we get

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\int u_h^2\,dx &= \sum_j -\left(\widehat{\frac{u_h^2}{2}} u_h - \frac{u_h^3}{6}\right)|_{\partial I_j} - \left(\widehat{\phi_h}\phi_h - \frac{\phi_h^2}{2}\right)|_{\partial I_j} + \left(\widehat{p_h} p_h - \frac{p_h^2}{2}\right)|_{\partial I_j} \\
&= \sum_j \left(\widehat{\frac{u_h^2}{2}}[u_h] - \frac{1}{6}[u_h^3]\right)_{j+\frac{1}{2}} + \left(\widehat{\phi_h}[\phi_h] - \frac{1}{2}[\phi_h^2]\right)_{j+\frac{1}{2}} - \left(\widehat{p_h}[p_h] - \frac{1}{2}[p_h^2]\right)_{j+\frac{1}{2}} \\
&= \sum_j \left(\widehat{\phi_h} - \{\phi_h\}\right)[\phi_h] + \left(\{p_h\} - \widehat{p_h}\right)[p_h] \\
&= \sum_j \left(\theta - \frac{1}{2}\right)\left([\phi_h]^2 + [p_h]^2\right)
\end{aligned}$$

which proves (14). □

## 4 Error estimations

In this section we estimate the error from approximating $p_h, \phi_h, u_h$. We proceed by defining a global projection with some established properties to prove that the errors from approximating $p_h$ and $\phi_h$ can be controlled by the errors from approximating $u_h$. Then, we show that the error from approximating $u_h$ is of optimal order. In the use of other boundary conditions, there is a need to refine the proof by carefully estimating the errors induced from boundary terms.

### 4.1 The global projection

Let $\omega \in L^2(I)$, and be smooth on each $I_j$, say $\omega|_{I_j} \in H^s(I_j)$ for $s \geq k+1$, we define the projection $Q_\theta$ such that it satisfies the following properties:

$$\int_{I_j} (Q_\theta \omega)\, v\, dx = \int_{I_j} \omega v\, dx, \quad \forall v \in P^{k-1}, \ j = 1, \cdots, N, \tag{18a}$$

$$\widehat{Q_\theta \omega}_{j+\frac{1}{2}} = \widehat{\omega}_{j+1/2}, \quad j = 1, \ldots, N, \tag{18b}$$

where

$$\widehat{V} := \theta V^+ + (1-\theta) V^-.$$

For $j = N$, we use the periodic extension to define $(Q_\theta \omega)^+_{N+1/2}$, in order to be consistent with the numerical flux defined in (7).

We first show that the projection $Q_\theta w$ is well defined.

**Lemma 4.1** *The projection $Q_\theta$ satisfying (18) is uniquely defined for either $\theta \neq \frac{1}{2}$ or $\theta = \frac{1}{2}$ with $k$ even and $N$ odd.*

*Proof* Let $\{\psi_l\}_{l=0}^k$ be a set of orthogonal Legendre polynomials on $[-1, 1]$ of degree up to $k$. We can write the projection $Q_\theta$ of $\omega \in H^{k+1}(I)$ on each cell $I_j$ as

$$(Q_\theta \omega)\left(x_j + \frac{h}{2}\xi\right) = \sum_{l=0}^{k} a_l^j \psi_l(\xi), \quad \xi \in [-1, 1].$$

With $v = \psi_i$, the condition (18a) gives

$$a_i^j = \frac{2i+1}{2} \int_{-1}^{1} \omega\left(x_j + \frac{h}{2}\xi\right) \psi_i(\xi)\, d\xi, \quad i = 0, \ldots, k-1, \tag{19}$$

where we have used $\int_{-1}^1 \psi_i^2(\xi)\, d\xi = \frac{2}{2i+1}$.

It remains to determine $a_k^j$ for $j = 1, \ldots, N$. Since $\psi_l(\pm 1) = (\pm 1)^l$, the condition (18b) gives

$$\theta\left(\sum_{l=0}^{k} a_l^{j+1}(-1)^l\right)+(1-\theta)\left(\sum_{l=0}^{k} a_l^{j}\right)=\hat{\omega}(x_{j+\frac{1}{2}}),\quad j=1,\ldots,N. \tag{20}$$

Because $\omega$ is periodic, we require that

$$\sum_{l=0}^{k} a_l^{N+1}\psi_l(\xi)=\sum_{l=0}^{k} a_l^{1}\psi_l(\xi),\quad \forall \xi\in[-1,1].$$

which allows us to write the system (20) as

$$\begin{bmatrix} 1-\theta & (-1)^k\theta & 0 & \cdots & 0\\ 0 & 1-\theta & (-1)^k\theta & \cdots & 0\\ \vdots & \ddots & \ddots & \ddots & \vdots\\ (-1)^k\theta & 0 & 0 & \cdots & 1-\theta \end{bmatrix}\cdot\begin{pmatrix} a_k^1\\ a_k^2\\ \vdots\\ a_k^N \end{pmatrix}=\begin{pmatrix} b_1\\ b_2\\ \vdots\\ b_N \end{pmatrix}, \tag{21}$$

where $b_j=\hat{\omega}(x_{j+\frac{1}{2}})-\theta\left(\sum_{l=0}^{k-1} a_l^{j+1}(-1)^l\right)-(1-\theta)\left(\sum_{l=0}^{k-1} a_l^{j}\right)$. The determinant of the coefficient matrix $A$ above is given by $(1-\theta)^N+(-1)^{N+1+kN}\theta^N$, which is non-zero for all $\theta\neq\frac{1}{2}$. When $\theta=\frac{1}{2}$, the determinant is non-zero whenever $N$ is odd and $k$ is even. This proves the lemma. □

**Lemma 4.2** *For $\omega|_{I_j}\in H^{k+1}(I_j)$ for $j=1\cdots,N$, we have the following projection error*

$$\|Q_\theta\omega-\omega\|\le Ch^{k+1}|\omega|_{k+1}, \tag{22}$$

*where $C$ depends on $k\ge 1$ and $\theta$.*

*Proof* The proof is carried out in two steps.

Step 1. We first establish the following inequality

$$\|(Q_\theta-I)\omega\|^2\le Ch\sum_{j=1}^{N}\|\tilde{\omega}^j\|_{1,\hat{I}}^2, \tag{23}$$

where

$$\tilde{\omega}^j(\xi):=\omega\left(x_j+\frac{h}{2}\xi\right),\quad \xi\in[-1,1]=\hat{I}. \tag{24}$$

By the Cauchy–Schwarz inequality, we have from (19) that for $j=1,\ldots,N$,

$$|a_i^j|^2\le\frac{2i+1}{2}\|\tilde{w}^j\|_{0,\hat{I}}^2,\quad i=0,\ldots,k-1. \tag{25}$$

Hence

$$\sum_{j=0}^{N}\sum_{i=0}^{k-1}|a_i^j|^2 \le (k-1/2)\sum_{j=1}^{N}\|\tilde{\omega}^j\|_{0,\hat{I}}^2.$$

From (21) of the form

$$a_k = A^{-1}b,$$

where $a_k = [a_k^1, \ldots, a_k^N]^T$ and $b = [b_1, \ldots, b_N]^T$, it follows that

$$\begin{aligned}\sum_{j=1}^{N}(a_k^j)^2 = b^T(A^{-1})^T A^{-1}b &\le C\sum_{j=1}^{N}(b_j)^2 \\ &\le C\sum_{j=1}^{N}\left[(\tilde{\omega}^j(1))^2 + \sum_{l=0}^{k-1}(a_l^{j+1})^2 + \sum_{l=0}^{k-1}(a_l^j)^2\right] \le C\sum_{j=1}^{N}\|\tilde{\omega}^j\|_{1,\hat{I}}^2,\end{aligned}$$

where we have used the Sobolev inequality $|\tilde{\omega}^j|_{\infty,\hat{I}} \le C\|\tilde{\omega}^j\|_{1,\hat{I}}$. Hence,

$$\begin{aligned}\|Q_\theta\omega\|^2 &= \sum_{j=1}^{N}\|Q_\theta\omega\|_{0,I_j}^2 \\ &= \sum_{j=1}^{N}\left[\frac{h}{2}\sum_{l=0}^{k-1}(a_l^j)^2\|\psi_l\|_{0,\hat{I}}^2 + \frac{h}{2}(a_k^j)^2\|\psi_k\|_{0,\hat{I}}^2\right] \\ &\le h\sum_{j=1}^{N}\left[\sum_{i=0}^{k-1}|a_i^j|^2 + (a_k^j)^2\right] \le Ch\sum_{j=1}^{N}\|\tilde{\omega}^j\|_{1,\hat{I}}^2.\end{aligned}$$

Step 2. For any $v \in V_h^k(I)$, we have that $Q_\theta v = v$. Therefore, using (23) we have

$$\begin{aligned}\|Q_\theta\omega - \omega\|^2 &= \|(Q_\theta - I)(\omega - v)\|^2 \\ &\le Ch\sum_{j=1}^{N}\|\tilde{\omega}^j - \hat{v}^j\|_{1,\hat{I}}^2.\end{aligned}$$

The left hand sides does not depend on $v$ at all, we then have

$$\begin{aligned}\|Q_\theta\omega - \omega\|^2 &\le Ch\sum_{j=1}^{N}\inf_{\hat{v}^j\in P^k[-1,1]}\|\tilde{\omega}^j - \hat{v}^j\|_{1,\hat{I}}^2 \\ &\le Ch\sum_{j=1}^{N}|\tilde{\omega}^j|_{k+1,\hat{I}}^2 =: Ch^{2k+2}|\omega|_{k+1}^2,\end{aligned}$$

where the Bramble–Hilbert lemma (Ref. [8]) has been used. The proof of (22) is complete. □

**Lemma 4.3** *For $k \geq 1$ the following inequality holds,*

$$\sum_{j=1}^{N} \left| (\omega - Q_\theta \omega)(x^-_{j+\frac{1}{2}}) \right|^2 \leq C|\omega|^2_{k+1} h^{2k+1}. \tag{26}$$

*The constant $C$ depends on $k$ and $\theta$.*

*Proof* On each interval $I_j$, using the orthogonality relation (18a), we have

$$\omega(x)|_{I_j} := \tilde{\omega}^j(\xi) = \sum_{l=0}^{\infty} \omega_l^j \psi_l(\xi),$$

$$Q_\theta \omega(x)|_{I_j} := \widetilde{(Q_\theta \omega)}^j(\xi) = \sum_{l=0}^{k-1} \omega_l^j \psi_l(\xi) + \alpha_k^j \psi_k(\xi).$$

Hence, by $\psi_l(1) = 1$, we have

$$\left| (\omega - Q_\theta \omega)(x^-_{j+\frac{1}{2}}) \right|^2 \leq 2 \left| \sum_{l=k}^{\infty} \omega_l^j \right|^2 + 2|\alpha_k^j|^2. \tag{27}$$

To control the first term on the right-hand side of (27), we consider the following expression

$$\partial_\xi \tilde{\omega}^j(\xi) = \sum_{l=0}^{\infty} \beta_l^j \psi_l(\xi). \tag{28}$$

Following the idea in [6], we integrate (28) with respect to $\xi$ to get

$$\tilde{\omega}^j(\xi) = \tilde{\omega}^j(-1) + \sum_{l=0}^{\infty} \beta_l^j \int_{-1}^{\xi} \psi_l(v) dv.$$

Using the property of Legendre polynomials

$$\int_{-1}^{\xi} \psi_i(v) dv = \frac{1}{2i+1} \left( \psi_{i+1}(\xi) - \psi_{i-1}(\xi) \right),$$

we can write

$$\tilde{\omega}^j(\xi) = \tilde{\omega}^j(-1) + \left(\beta_0^j - \frac{\beta_1^j}{3}\right)\psi_0(\xi) + \sum_{l=1}^{\infty}\left(\frac{\beta_{l-1}^j}{2l-1} - \frac{\beta_{l+1}^j}{2l+3}\right)\psi_l(\xi).$$

Therefore,

$$\omega_i^j = \left(\frac{\beta_{i-1}^j}{2i-1} - \frac{\beta_{i+1}^j}{2i+3}\right), \quad i \geq 1.$$

Thus,

$$\sum_{l=k}^{\infty}\omega_l^j = \left(\frac{\beta_{k-1}^j}{2k-1} + \frac{\beta_k^j}{2k+1}\right),$$
$$\sum_{l=k}^{\infty}\omega_l^{j+1}(-1)^l = (-1)^k\left(\frac{\beta_{k-1}^{j+1}}{2k-1} - \frac{\beta_k^{j+1}}{2k+1}\right).$$

These ensure the following estimates

$$\left|\sum_{l=k}^{\infty}\omega_l^j\right|^2 \leq \frac{1}{2k-1}\left(\frac{2(\beta_{k-1}^j)^2}{2k-1} + \frac{2(\beta_{k+1}^j)^2}{2k+1}\right) \leq \frac{1}{2k-1}\|\partial_\xi\tilde{\omega}^j\|_{0,\hat{I}}^2 \quad (29a)$$

$$\left|\sum_{l=k}^{\infty}\omega_l^{j+1}(-1)^l\right|^2 \leq \frac{1}{2k-1}\|\partial_\xi\tilde{\omega}^{j+1}\|_{0,\hat{I}}^2. \quad (29b)$$

The second term on the right-hand side of (27) is determined by (18b), i.e.,

$$\theta\alpha_k^{j+1}(-1)^k + (1-\theta)\alpha_k^j = \theta\left(\sum_{l=k}^{\infty}\omega_l^{j+1}(-1)^l\right) + (1-\theta)\left(\sum_{l=k}^{\infty}\omega_l^j\right), \quad (30)$$

where we have used $\hat{\omega}(x_{j+1/2}) = \theta\tilde{\omega}^{j+1}(-1) + (1-\theta)\tilde{\omega}^j(1)$. We then have from (29) and (30) that

$$\begin{aligned}\sum_{j=1}^{N}|\alpha_k^j|^2 &\leq C\sum_{j=1}^{N}\left(\left|\sum_{l=k}^{\infty}\omega_l^{j+1}(-1)^l\right|^2 + \left|\sum_{l=k}^{\infty}\omega_l^j\right|^2\right)\\ &\leq C\sum_{j=1}^{N}\|\partial_\xi\tilde{\omega}^{j+1}\|_{0,\hat{I}}^2 + \|\partial_\xi\tilde{\omega}^j\|_{0,\hat{I}}^2\\ &\leq C\sum_{j=1}^{N}\|\partial_\xi\tilde{\omega}^j\|_{0,\hat{I}}^2.\end{aligned}$$

Insertion of these estimates back into (27) yields

$$\sum_{j=1}^{N}\left|(\omega-Q_\theta\omega)(x^-_{j+\frac{1}{2}})\right|^2 \le C\sum_{j=1}^{N}\|\partial_\xi\tilde{\omega}^j\|^2_{0,\hat{I}}.$$

Recall that $Q_\theta v = v$ for any $v \in P^k$, we proceed

$$\begin{aligned}\sum_{j=1}^{N}\left|(\omega-Q_\theta\omega)(x^-_{j+\frac{1}{2}})\right|^2 &\le C\sum_{j=1}^{N}\inf_{\tilde{v}\in P^k}\|\partial_\xi\tilde{\omega}^j-\partial_\xi\tilde{v}\|^2_{0,\hat{I}}\\ &= C\sum_{j=1}^{N}\inf_{\tilde{p}\in P^{k-1}}\|\partial_\xi\tilde{\omega}^j-\tilde{p}\|^2_{0,\hat{I}}\\ &\le C\sum_{j=1}^{N}|\partial_\xi\tilde{\omega}^j|^2_{k,\hat{I}}\\ &= \left(\frac{h}{2}\right)^{2k+2}\left(\frac{h}{2}\right)^{-1}C|\omega|^2_{k+1} \le C|\omega|^2_{k+1}h^{2k+1}.\end{aligned}$$

□

We will use the error estimates obtained in Lemmas 4.2–4.3 to estimate the error of the computed solution. Moreover, for any $w \in V_h^k$, we utilize the following inverse properties which can be easily derived from the classical ones (see e.g., [8]),

$$\|\partial_x w\| \le Ch^{-1}\|w\|, \tag{31a}$$

$$\|w\|_{\Gamma_h} \le Ch^{-1/2}\|w\|, \tag{31b}$$

$$\|w\|_\infty \le Ch^{-1/2}\|w\|, \tag{31c}$$

where

$$\|w\|^2_{\Gamma_h} := \sum_{j=1}^{N}\left(\left|w^-_{j+1/2}\right|^2+\left|w^+_{j+1/2}\right|^2\right).$$

The constant $C$ is independent of $w$ and $h$.

*Remark 4.1* Here the estimates for inverse inequalities are valid for piecewise polynomials; the proof usually uses the equivalence of norms for finite dimensional problems and some scaling techniques. It is often desired that the constant is as small as possible [33]. Here we just list these results without any further specification of the constants.

### 4.2 Approximating $p$ and $\phi$

**Lemma 4.4** *Let $(u, p, \phi)$ be the exact solution of the system (5). Let $(u_h, p_h, \phi_h)$ be obtained from (6) with the choice of fluxes (7). Let $\theta \in [0, 1]$ be such that both $Q_\theta$ and $Q_{1-\theta}$ are uniquely defined, then the following inequality holds for all $t > 0$.*

$$\|Q_{1-\theta}p - p_h\|^2 + \|Q_\theta\phi - \phi_h\|^2 \le \|Q_{1-\theta}p - p\|^2 + 2\|Q_\theta\phi - \phi\|^2 + 2\|u - u_h\|^2. \tag{32}$$

*Proof* Since the scheme with fluxes (7) is consistent, (6b)–(6c) also hold for $(u, p, \phi)$. In other words,

$$\int_{I_j} p\gamma + \int_{I_j} \phi\gamma_x - \phi\gamma\,|_{\partial I_j} = 0, \tag{33a}$$

$$-\int_{I_j} pq_x + pq\,|_{\partial I_j} - \int_{I_j} (\phi + u)q = 0, \tag{33b}$$

Subtracting (6b)–(6c) from (33a)–(33b), we get the error equations

$$\int_{I_j} (p - p_h)\gamma + \int_{I_j} (\phi - \phi_h)\gamma_x - \left(\phi - \widehat{\phi_h}\right)\gamma\,|_{\partial I_j} = 0, \tag{34a}$$

$$-\int_{I_j} (p - p_h)q_x + (p - \widehat{p_h})\,q\,|_{\partial I_j} - \int_{I_j} (\phi - \phi_h)q = \int_{I_j} (u - u_h)q. \tag{34b}$$

Define $\epsilon_p = Q_{1-\theta}p - p$, $w_p = Q_{1-\theta}p - p_h$, $\widehat{\epsilon_p} = \widehat{Q_{1-\theta}p} - p$, and $\widehat{w_p} = \widehat{Q_{1-\theta}p} - \widehat{p_h}$. (Similar definition can be given for $\epsilon_\phi$, $w_\phi$, $\widehat{\epsilon_\phi}$, and $\widehat{w_\phi}$ associated with $Q_\theta$.) We then choose $\gamma = w_p$, $q = w_\phi$ and take the summation of (34) over $j$ to get

$$\begin{aligned}
&\sum_j \int_{I_j} (w_p - \epsilon_p)w_p + \sum_j \int_{I_j} (w_\phi - \epsilon_\phi)(w_p)_x + \sum_j \left(\widehat{w_\phi} - \widehat{\epsilon_\phi}\right)[w_p]_{j+\frac{1}{2}} = 0,\\
&\quad - \sum_j \int_{I_j} (w_p - \epsilon_p)(w_\phi)_x - \sum_j \left(\widehat{w_p} - \widehat{\epsilon_p}\right)[w_\phi]_{j+\frac{1}{2}} - \sum_j \int_{I_j} (w_\phi - \epsilon_\phi)w_\phi\\
&= \int_I (u - u_h)w_\phi.
\end{aligned}$$

Take the difference of both equations, we get

$$\int_I w_p^2 + \int_I w_\phi^2 = \int_I \epsilon_p w_p + \int_I \epsilon_\phi w_\phi + \varrho_1 + \varrho_2 + \varrho_3 - \int_I (u - u_h)w_\phi, \tag{35}$$

where

$$\varrho_1 = -\sum_j \int_{I_j} \left(w_\phi (w_p)_x + w_p (w_\phi)_x\right) - \sum_j \widehat{w_\phi}[w_p]_{j+\frac{1}{2}} - \sum_j \widehat{w_p}[w_\phi]_{j+\frac{1}{2}},$$

$$\varrho_2 = \sum_j \int_{I_j} \epsilon_\phi (w_p)_x + \sum_j \widehat{\epsilon_\phi}[w_p]_{j+\frac{1}{2}},$$

$$\varrho_3 = \sum_j \int_{I_j} \epsilon_p (w_\phi)_x + \sum_j \widehat{\epsilon_p}[w_\phi]_{j+\frac{1}{2}}.$$

First, note that $w_\phi (w_p)_x + w_p (w_\phi)_x = (w_p w_\phi)_x$, so

$$\varrho_1 = \sum_j [w_p w_\phi]_{j+\frac{1}{2}} - \sum_j \widehat{w_\phi}[w_p]_{j+\frac{1}{2}} - \sum_j \widehat{w_p}[w_\phi]_{j+\frac{1}{2}} = 0,$$

with the choice of numerical fluxes (7b)–(7c). As for $\varrho_2$, the property (18a) of $Q_\theta$ gives

$$\sum_j \int_{I_j} \epsilon_\phi (w_p)_x = \sum_j \int_{I_j} (Q_\theta \phi - \phi)(w_p)_x = 0,$$

since $(w_p)_x$ is in $P^{k-1}$. On the other hand, the property (18b) of $Q_\theta$ gives

$$\sum_j \widehat{\epsilon_\phi}[w_p]_{j+\frac{1}{2}} = \sum_j \left(\widehat{Q_\theta \phi} - \phi\right)[w_p]_{j+\frac{1}{2}} = 0.$$

Similarly, the term $\varrho_3$ vanishes by the properties (18) of $Q_{1-\theta}$.

Using $\varrho_i = 0$ for $i = 1, 2, 3$ and the Young's inequality $ab \le \frac{a^2}{2\mu} + \frac{\mu b^2}{2}$ with $\mu = 1$ for the first term and $\mu = \frac{1}{2}$ for the last two terms in (35), we get

$$\frac{1}{2}\|w_p\|^2 + \frac{1}{2}\|w_\phi\|^2 \le \frac{1}{2}\|\epsilon_p\|^2 + \|\epsilon_\phi\|^2 + \|u - u_h\|^2, \tag{36}$$

which proves (32). □

### 4.3 Approximating $u$

**Theorem 4.5** *Let $u \in L^\infty((0,T]; H^s(I))$, $s \ge k+1$, be the smooth solution to (1), for $0 < t < T$. If $k$ is even, then the numerical solution, $u_h$, obtained from the scheme (6) and the numerical fluxes (7) satisfies*

$$\sup_{t\in[0,T]} \|u(t,\cdot) - u_h(t,\cdot)\| \le C\|u\|_{L^\infty((0,T];H^{k+1}(I))} h^{k+1}. \tag{37}$$

*The constant $C$ may depend on $T$ and the data given, but is independent of the mesh size.*

*Proof* Since the scheme (6) with fluxes (7) is consistent, (6a) also holds for $(u, p, \phi)$. In other words,

$$\int_{I_j} u_t \rho - \int_{I_j} \frac{u^2}{2} \rho_x + \widehat{\frac{u^2}{2}} \rho \,|_{\partial I_j} - \int_{I_j} p\rho = 0. \tag{38}$$

Define $w = Q_{1/2}u - u_h$ and $\epsilon = Q_{1/2}u - u$. We have that $u - u_h = w - \epsilon$. Subtracting (6a) from (38) and choose $\rho = w$, we get

$$\int_{I_j} w_t w = \int_{I_j} \epsilon_t w + \int_{I_j} \left( \frac{u^2}{2} - \frac{u_h^2}{2} \right) w_x - \left( \frac{u^2}{2} - \widehat{\frac{u_h^2}{2}} \right) w \,|_{\partial I_j} + \int_{I_j} e_p w,$$

where $e_p = p - p_h$.

Take summation over all $j$ and introduce $\{u_h\}^2/2$ into the third term on the right-hand side to get

$$\begin{aligned} \int_I w_t w = \int_I \epsilon_t w + \sum_j \int_{I_j} \left( \frac{u^2}{2} - \frac{u_h^2}{2} \right) w_x + \sum_j \left( \frac{u^2}{2} - \frac{\{u_h\}^2}{2} \right) [w]_{j+\frac{1}{2}} \\ + \sum_j \left( \frac{\{u_h\}^2}{2} - \widehat{\frac{u_h^2}{2}} \right) [w]_{j+\frac{1}{2}} + \int_I e_p w. \end{aligned}$$

Using the identity $A^2/2 - B^2/2 = A(A-B) - (A-B)^2/2$, we get

$$\begin{aligned} \int_I w_t w = \int_I \epsilon_t w + \sum_j \int_{I_j} u\,(u - u_h)\, w_x - \frac{1}{2} \sum_j \int_{I_j} (u - u_h)^2 w_x \\ + \sum_j u\,(u - \{u_h\})\,[w]_{j+\frac{1}{2}} - \frac{1}{2} \sum_j (u - \{u_h\})^2\,[w]_{j+\frac{1}{2}} \\ + \sum_j \left( \frac{\{u_h\}^2}{2} - \widehat{\frac{u_h^2}{2}} \right) [w]_{j+\frac{1}{2}} + \int_I e_p w. \end{aligned}$$

Let $\{w\} = \{Q_{1/2}u\} - \{u_h\}$ and $\{\epsilon\} = \{Q_{1/2}u\} - u$, we can write

$$\int_I w_t w = \int_I \epsilon_t w + \tau_1 + \tau_2 + \tau_3 + \tau_4 + \tau_5 + \int_I e_p w, \tag{39}$$

where

$$\tau_1 = \sum_j \int_{I_j} uww_x + \sum_j u\{w\}[w]_{j+\frac{1}{2}},$$

$$\tau_2 = -\sum_j \int_{I_j} u\epsilon w_x - \sum_j u\{\epsilon\}[w]_{j+\frac{1}{2}},$$

$$\tau_3 = -\frac{1}{2}\sum_j \int_{I_j} w^2 w_x - \frac{1}{2}\sum_j \{w\}^2[w]_{j+\frac{1}{2}},$$

$$\tau_4 = \sum_j \int_{I_j} w\epsilon w_x - \frac{1}{2}\sum_j \int_{I_j} \epsilon^2 w_x + \sum_j \{w\}\{\epsilon\}[w]_{j+\frac{1}{2}} - \frac{1}{2}\sum_j \{\epsilon\}^2[w]_{j+\frac{1}{2}},$$

$$\tau_5 = \sum_j \left(\frac{\{u_h\}^2}{2} - \frac{\widehat{u_h^2}}{2}\right)[w]_{j+\frac{1}{2}}.$$

Note that

$$\begin{aligned}\tau_1 &= \sum_j \int_{I_j} u\left(\frac{w^2}{2}\right)_x + \sum_j u\{w\}[w]_{j+\frac{1}{2}} \\ &= -\sum_j u\left[\frac{w^2}{2}\right]_{j+\frac{1}{2}} - \sum_j \int_{I_j} u_x\left(\frac{w^2}{2}\right) + \sum_j u\{w\}[w]_{j+\frac{1}{2}} \\ &= -\sum_j \int_{I_j} u_x\left(\frac{w^2}{2}\right) \le \frac{1}{2}\|u_x\|_\infty \|w\|^2.\end{aligned}$$

As for $\tau_2$, we write $u(x) = u(x_j) + u'(x_j^*)(x - x_j)$ for all $x \in I_j$ where $x_j^*$ is between $x$ and $x_j$. Therefore,

$$\begin{aligned}\left|-\sum_j \int_{I_j} u\epsilon w_x - \sum_j u\{\epsilon\}[w]_{j+\frac{1}{2}}\right| &\le \left|-\sum_j u(x_j)\int_{I_j} \epsilon w_x \right. \\ &\qquad \left. -\sum_j u'(x_j^*)\int_{I_j} \epsilon(x - x_j)w_x\right| + |0| \\ &\le \left|\sum_j u(x_j)\int_{I_j} \epsilon w_x\right| + \sum_j \left|u'(x_j^*)\right| \int_{I_j} |\epsilon h w_x|\end{aligned}$$

$$\leq |0| + \frac{1}{2}\|u_x\|_\infty \left(\|\epsilon\|^2 + h^2\|w_x\|^2\right)$$
$$\leq C(h^{2k+2} + \|w\|^2),$$

because of the projection properties (18), the inverse property (31a), and Lemma 4.2.

For $\tau_3$, we can show that

$$-\frac{1}{2}\sum_j \int_{I_j} \left(\frac{w^3}{3}\right)_x - \frac{1}{2}\sum_j \{w\}^2[w]_{j+\frac{1}{2}} = \frac{1}{2}\sum_j \left[\frac{w^3}{3}\right] - \frac{1}{2}\sum_j \{w\}^2[w]_{j+\frac{1}{2}}$$
$$= \frac{1}{24}\sum_j [w]^3_{j+\frac{1}{2}} \leq Ch^{-1}\|w\|_\infty\|w\|^2 \leq Ch^{-3/2}\|w\|^3,$$

by the inverse properties (31b)–(31c).

From the inverse properties (31a) and (31c), it follows that

$$\|w_x\|_\infty \leq Ch^{-3/2}\|w\|,$$

with which we are able to estimate terms in $\tau_4$ by using Lemma 4.2,

$$\sum_j \int_{I_j} w\epsilon w_x \leq C\|w\|\|\epsilon\|\|w_x\|_\infty \leq Ch^{k+1-3/2}\|w\|^2 = Ch^{k-1/2}\|w\|^2,$$
$$\frac{1}{2}\sum_j \int_{I_j} \epsilon^2 w_x \leq \frac{1}{2}\|w_x\|_\infty\|\epsilon\|^2 \leq Ch^{2k+1/2}\|w\|.$$

As for the remaining terms in $\tau_4$,

$$\sum_j \{w\}\{\epsilon\}[w]_{j+\frac{1}{2}} - \frac{1}{2}\sum_j \{\epsilon\}^2[w]_{j+\frac{1}{2}} = 0$$

because of the projection property (18b).

Finally, using the fact that $\{u_h\}^2/2 - \widehat{u_h^2}/2 = -[u_h]^2/24$, and $[u_h] = [u - u_h] = [w] - [\epsilon]$, we have

$$\tau_5 = \sum_j \left(\frac{\{u_h\}^2}{2} - \frac{\widehat{u_h^2}}{2}\right)[w]_{j+\frac{1}{2}} = \sum_j -\frac{1}{24}[u_h]^2[w]_{j+\frac{1}{2}},$$
$$= \sum_j -\frac{1}{24}[w]^3_{j+\frac{1}{2}} + \frac{1}{12}[\epsilon][w]^2_{j+\frac{1}{2}} - \frac{1}{24}[\epsilon]^2[w]_{j+\frac{1}{2}}$$
$$\leq C\|w\|_\infty(\|w\|^2_{\Gamma_h} + \|\epsilon\|_{\Gamma_h}\|w\|_{\Gamma_h} + \|\epsilon\|^2_{\Gamma_h})$$
$$\leq Ch^{-1/2}\|w\|(h^{-1}\|w\|^2$$

$$
\begin{aligned}
&+ h^{k+1/2} h^{-1/2} \|w\| + h^{2k+1}) \\
&\leq C(h^{2k+2} + \|w\|^2 + h^{-3/2}\|w\|^3),
\end{aligned}
$$

where we have used the inverse properties (31b)–(31c) and Lemma 4.3.

The results from $\tau_1$ to $\tau_5$ and (39) give

$$
\frac{d}{dt}\|w\|^2 \leq C_1 \left( h^{2k+2} + \|w\|^2 + h^{-3/2}\|w\|^3 \right). \tag{40}
$$

We note that

$$
\|w(t=0,\cdot)\|^2 \leq C_2 h^{2k+2}, \tag{41}
$$

because $w(0,\cdot) = \epsilon(0,\cdot) + (u_0 - u_h(0,\cdot))$, where $u_h(0,\cdot)$ is prepared using a standard $L^2-$projection from the given initial data. To solve (40) with initial data (41), we introduce

$$
G(t) = h^{2k+2} + \int_0^t \|w(\tau,\cdot)\|^2 + h^{-3/2}\|w(\tau,\cdot)\|^3 \, d\tau. \tag{42}
$$

With this and (41), we can write

$$
\|w(t,\cdot)\|^2 \leq C G(t). \tag{43}
$$

Hence, for $C_* = C \max\{1, \sqrt{C}\}$,

$$
G'(t) \leq C_* \left( G(t) + h^{-3/2} G(t)^{3/2} \right). \tag{44}
$$

Integrate (44) to get

$$
F\left( \frac{G}{G(0)} \right) \leq C_* T, \tag{45}
$$

where

$$
F(\eta) = \int_1^\eta \frac{1}{\xi + h^{-3/2}\sqrt{G(0)}\xi^{3/2}} \, d\xi = \int_1^\eta \frac{1}{\xi + h^{k-1/2}\xi^{3/2}} \, d\xi. \tag{46}
$$

For $k \geq 1$, we have that $F'(\eta)$ is uniformly (with respect to $h$) positive and bounded above by 1 for all $\eta > 1$. Thus, there exists $\tilde{C}$ such that $F(\tilde{C}) = C_* T$ for given $T > 0$. Therefore, we have that $F(\frac{G}{G(0)}) \leq F(\tilde{C})$ which implies $\frac{G}{G(0)} \leq \tilde{C}$. Using this and (43), we prove (37) as desired. □

### 4.4 Time discretization

We partition the time interval $[0, T]$ into $M$ equal subintervals with boundaries $\{t^n\}, n = 0, 1, 2, \ldots, M$. Set $\Delta t = T/M$ as the time step. In order to preserve both mass and energy at the fully discrete level, we may use the Crank–Nicolson time discretization to find

$$u_h^{n+1} = 2u_h^* - u_h^n,$$

where $u_h^*$ is determined by

$$\begin{aligned} &2\int_{I_j} \frac{u_h^* - u_h^n}{\Delta t}\rho - \int_{I_j} \frac{(u_h^*)^2}{2}\rho_x + \widehat{\frac{(u_h^*)^2}{2}}\rho\,|_{\partial I_j} - \int_{I_j} p_h^*\rho = 0, \\ &\int_{I_j} p_h^*\gamma + \int_{I_j} \phi_h^*\gamma_x - \widehat{\phi_h^*}\gamma\,|_{\partial I_j} = 0, \\ &-\int_{I_j} p_h^* q_x + \widehat{p_h^*} q\,|_{\partial I_j} - \int_{I_j} (\phi_h^* + u_h^*)\, q = 0. \end{aligned} \tag{47}$$

Indeed, this time discretization has the desired and provable properties.

**Theorem 4.6** *The fully-discrete scheme* (47) *gives solution* $u_h^n$ *that satisfies*

$$\int_0^L u_h^{n+1}\,dx = \int_0^L u_h^n\,dx, \tag{48}$$

$$\int_0^L \left(u_h^{n+1}\right)^2 dx = \int_0^L (u_h^n)^2\,dx + \Delta t(2\theta - 1)\sum_j \left([\phi_h^*]^2 + [p_h^*]^2\right)_{j+\frac{1}{2}}, \tag{49}$$

*for all* $0 \le n < M$. *Here,* $\phi_h^* = (\phi_h^{n+1} + \phi_h^n)/2$, *and* $p_h^* = (p_h^{n+1} + p_h^n)/2$.

*Proof* Take the test functions $\rho = 1$, $\gamma = 1$ in (47), adding together, to get

$$\int_{I_j} \frac{u_h^{n+1} - u_h^n}{\Delta t} + \widehat{\frac{(u_h^*)^2}{2}}\,|_{\partial I_j} - \widehat{\phi_h^*}\,|_{\partial I_j} = 0,$$

which upon summation over $j$ proves (48). Next, we choose the test functions $\rho = u_h^*$, $\gamma = -\phi_h^*$, and $q = -p_h^*$ so that

$$\int_{I_j} \frac{u_h^{n+1}-u_h^n}{\Delta t}u_h^* - \int_{I_j} \frac{\left(u_h^*\right)^2}{2}\left(u_h^*\right)_x + \widehat{\frac{\left(u_h^*\right)^2}{2}}u_h^*\,|_{\partial I_j} - \int_{I_j} p_h^* u_h^* = 0,$$

$$-\int_{I_j} p_h^*\phi_h^* - \int_{I_j} \phi_h^*\left(\phi_h^*\right)_x + \widehat{\phi_h^*}\phi_h^*\,|_{\partial I_j} = 0,$$

$$\int_{I_j} p_h^*\left(p_h^*\right)_x - \widehat{p_h^*}p_h^*\,|_{\partial I_j} + \int_{I_j}(\phi_h^* + u_h^*)p_h^* = 0.$$

Summation of the above three equations over $j$ gives

$$\sum_j \int_{I_j} \frac{\left(u_h^{n+1}\right)^2 - \left(u_h^n\right)^2}{2\Delta t} = \int_0^L \frac{\left(u_h^{n+1}\right)^2 - \left(u_h^n\right)^2}{2\Delta t} = \left(\theta - \frac{1}{2}\right)\sum_j \left([\phi_h]^2 + [p_h]^2\right),$$

which leads to (49). □

Note that the above time discretization is fully nonlinear and requires the costly iteration solver. In practice, one would prefer to use some explicit solver with high order accuracy for time discretization. In our numerical simulation we choose to use the TVD third-order Runge–Kutta method [19] to solve the ODE system of the form $\dot{\mathbf{a}} = \mathfrak{L}(\mathbf{a})$:

$$\begin{aligned}
\mathbf{a}^{(1)} &= \mathbf{a}^n + \Delta t\,\mathfrak{L}(\mathbf{a}^n),\\
\mathbf{a}^{(2)} &= \frac{3}{4}\mathbf{a}^n + \frac{1}{4}\mathbf{a}^{(1)} + \frac{1}{4}\Delta t\,\mathfrak{L}(\mathbf{a}^{(1)}),\\
\mathbf{a}^{n+1} &= \frac{1}{3}\mathbf{a}^n + \frac{2}{3}\mathbf{a}^{(2)} + \frac{2}{3}\Delta t\,\mathfrak{L}(\mathbf{a}^{(2)}),
\end{aligned} \tag{50}$$

where $\mathbf{a}^n$ is the coefficient vector of $u_h^n$.

## 5 Numerical tests

It is known [16] that one steady solution of system (1a)–(1b) is given by

$$U_1(x) = \frac{4}{3}\left(e^{-|x|/2} - 1\right). \tag{51}$$

The system also has a steady periodic solution of the form

$$U_2(x) = \frac{4}{3}\left(\frac{\cosh\left(\frac{x}{2}\right)}{\cosh\left(\frac{p}{2}\right)} - 1\right), \tag{52}$$

**Table 1** Errors for Example 1 (accuracy test) with $\theta = 1/2$

| $k$ | $N$ | $\theta = 1/2$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | $L_1$ | Order | $L_2$ | Order | $L_\infty$ | Order |
| 1 | 10 | 5.5907e−02 | | 3.3360e−02 | | 4.8122e−02 | |
| | 20 | 2.8223e−02 | 0.9862 | 1.6765e−02 | 0.9926 | 2.6199e−02 | 0.8772 |
| | 40 | 1.4137e−02 | 0.9974 | 8.3909e−03 | 0.9986 | 1.3877e−02 | 0.9168 |
| | 80 | 7.0694e−03 | 0.9998 | 4.1959e−03 | 0.9999 | 7.1781e−03 | 0.9510 |
| 2 | 10 | 1.6478e−03 | | 1.0558e−03 | | 2.6115e−03 | |
| | 20 | 2.0461e−04 | 3.0096 | 1.3282e−04 | 2.9908 | 3.8008e−04 | 2.7805 |
| | 40 | 2.5457e−05 | 3.0067 | 1.6615e−05 | 2.9990 | 5.3250e−05 | 2.8355 |
| | 80 | 3.1778e−06 | 3.0019 | 2.0768e−06 | 3.0000 | 7.2587e−06 | 2.8750 |
| 3 | 10 | 2.5013e−04 | | 1.7960e−04 | | 6.4039e−04 | |
| | 20 | 3.1133e−05 | 3.0062 | 2.2535e−05 | 2.9945 | 9.9254e−05 | 2.6897 |
| | 40 | 3.8845e−06 | 3.0026 | 2.8200e−06 | 2.9984 | 1.4970e−05 | 2.7290 |
| | 80 | 4.8514e−07 | 3.0013 | 3.5262e−07 | 2.9995 | 2.2034e−06 | 2.7643 |
| 4 | 10 | 1.2182e−06 | | 9.6189e−07 | | 4.5036e−06 | |
| | 20 | 3.7401e−08 | 5.0256 | 3.0251e−08 | 4.9908 | 1.8170e−07 | 4.6314 |
| | 40 | 1.1545e−09 | 5.0178 | 9.4598e−10 | 4.9990 | 7.1522e−09 | 4.6670 |
| | 80 | 3.5841e−11 | 5.0095 | 2.9535e−11 | 5.0013 | 2.7574e−10 | 4.6970 |

for $-p < x < p$ and by periodic continuation with period $2p$. Because the system (1a)–(1b) is Galilean invariant, a family of traveling-wave solutions (1a)–(1b) may be obtained from the steady solutions as

$$u(t, x) = U(x - u_0 t) + u_0, \tag{53}$$

where $U$ is the steady state solution (51) or (52). We will use both steady and traveling wave solutions to test our scheme.

*Example 1* (Accuracy test) We run the semi-discrete scheme (6) and the numerical flux (7) with $\theta = 1/2, 0$, along with the third order Runge–Kutta method (50) on the steady state problem which has (52) as its exact solution. The results for $k = 1, 2, 3, 4$ are given in Tables 1 and 2 below. Here, we use $\Delta t = 0.001$, final time $t_{\max} = 2$, and $p = 2$. The norms of the error were computed by using the sixteen-point Gauss quadrature rule.

The results show that the optimal order of accuracy is achieved only when $k =$ even, which is consistent with our theoretical result on the optimal error estimates for $k =$ even. Also such an observation seems unaffected by the choice of $\theta \in [0, 1]$, though we only display results for $\theta = 1/2$ and $\theta = 0$.

*Example 2* (Energy-preserving test) We compare the performance of the LDG-C scheme and the LDG-D scheme with $\theta = 1/2$ on the traveling wave version of

**Table 2** Errors for Example 1 (accuracy test) with $\theta = 0$

| $k$ | $N$ | $\theta = 1/2$ | | | | | |
|---|---|---|---|---|---|---|---|
| | | $L_1$ | Order | $L_2$ | Order | $L_\infty$ | Order |
| 1 | 10 | 4.4335e−02 | | 3.0487e−02 | | 4.5356e−02 | |
| | 20 | 2.3380e−02 | 0.9232 | 1.5731e−02 | 0.9546 | 2.3925e−02 | 0.9227 |
| | 40 | 1.2089e−02 | 0.9516 | 8.0245e−03 | 0.9711 | 1.1393e−02 | 1.0704 |
| | 80 | 6.1588e−03 | 0.9730 | 4.0569e−03 | 0.9841 | 5.6451e−03 | 1.0131 |
| 2 | 10 | 1.9403e−04 | | 1.2987e−04 | | 3.0732e−04 | |
| | 20 | 1.7491e−05 | 3.4716 | 1.2618e−05 | 3.3635 | 3.9729e−05 | 2.9515 |
| | 40 | 1.4935e−06 | 3.5498 | 1.2035e−06 | 3.3903 | 5.0581e−06 | 2.9735 |
| | 80 | 1.3537e−07 | 3.4637 | 1.1552e−07 | 3.3809 | 6.3888e−07 | 2.9850 |
| 3 | 10 | 1.0986e−05 | | 1.0699e−05 | | 4.5712e−05 | |
| | 20 | 1.4206e−06 | 2.9510 | 1.3271e−06 | 3.0111 | 6.4713e−06 | 2.8204 |
| | 40 | 1.9201e−07 | 2.8872 | 1.6821e−07 | 2.9800 | 8.9441e−07 | 2.8551 |
| | 80 | 2.5077e−08 | 2.9368 | 2.1246e−08 | 2.9850 | 1.2020e−07 | 2.8955 |
| 4 | 10 | 8.5328e−08 | | 5.4694e−08 | | 1.4438e−07 | |
| | 20 | 1.9935e−09 | 5.4197 | 1.3489e−09 | 5.3415 | 5.4551e−09 | 4.7261 |
| | 40 | 4.3824e−11 | 5.5074 | 3.1240e−11 | 5.4323 | 1.9264e−10 | 4.8236 |
| | 80 | 1.2288e−12 | 5.1564 | 8.4612e−13 | 5.2064 | 6.3020e−12 | 4.9339 |

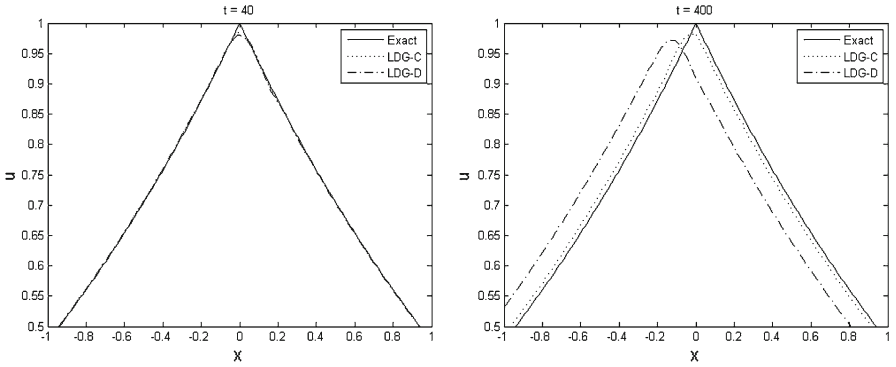



**Fig. 1** Example 2: comparison between the LDG-C and LDG-D scheme with $\theta = 1/2$. *Left* $t = 40$. *Right* $t = 400$

(51), with velocity $u_0 = 1$. The simulation is done on 160 elements with polynomials of degree $k = 4$ over the domain $[-20, 20]$. The time step for the third order Runge–Kutta method (50) is $\Delta t = 0.001$, and $t_{\max} = 400$.

In Fig. 1, we see that both schemes perform well over a short time. However, after a long time ($t = 400$), the LDG-C scheme performs clearly better as we can observe that it produces a smaller phase shift. In terms of $L^2$-energy, initially, $\|u_h(0, \cdot)\|$ is 2.108223389275528. At $t = 400$, the numerical solution obtained by the LDG-C scheme has $L^2$ energy $\|u_h(400, \cdot)\| = \mathbf{2.108223}389275528$, which agrees with the

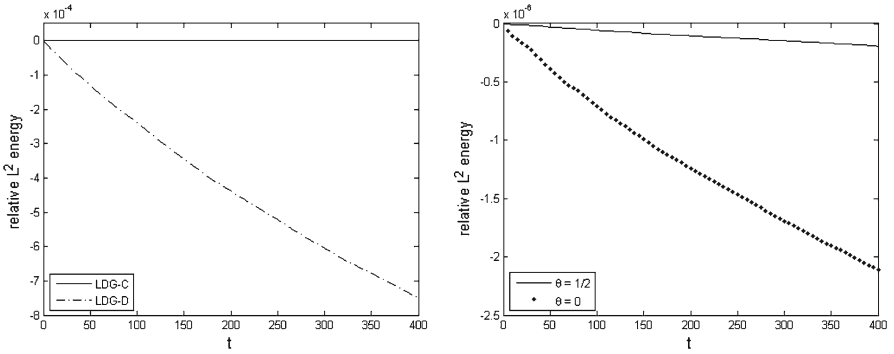



**Fig. 2** Example 2: the evolution of the relative $L^2$ energy over long-time simulation. *Left* comparison between LDG-C and LDG-D with $\theta = 1/2$. *Right* comparison between the flux (7) with $\theta = 1/2$ (LDG-C) and the flux (7) with $\theta = 0$

initial energy up to 6th decimal place. On the other hand, the LDG-D scheme with $\theta = 1/2$ yields $\|u_h(400, \cdot)\| = \mathbf{2.10}7474302191212$, which agrees with the initial energy up to only 2nd decimal place. Here, the $L^2$ norms were computed by using the sixteen-point Gauss quadrature rule.

We plot the evolution of the relative energy $\|u_h(0, \cdot)\| - \|u_h(400, \cdot)\|$ in Fig. 2. In addition to the comparison between LDG-C and LDG-D, we also compare the performance of the flux (7) with $\theta = 1/2$ (i.e. LDG-C) and with $\theta = 0$. The result is as we expected: when $\theta = 1/2$, the energy is conserved better than when $\theta \in [0, 1/2)$.

The numerical tests indicate that after long time simulation, phase shift is a main source of error, while the shape of waves remains stable. In order to quantify the shape error we use the formula introduced in [3]

$$\hat{e}(t, x) = \min_{\xi \in [-0.5, 0.5]} \|u_h(t, x) - u(t, x + \xi)\|,$$

for the numerical solution $u_h$ obtained from the LDG-C scheme with $\theta = 1/2$, while $u(t, x)$ is the exact solution. The shape error defined above compares how good the approximation is, modulo the translation group on the periodic domain, and it minimizes the difference between the numerical approximation and the spatially shifted exact solution. In Fig. 3, we see that the shape error fluctuates around a constant in time, with a visible periodic behavior. In contrast, the absolute $L^2$-error grows in time.

In next three examples, we use LDG-Ad on polynomial elements of degree $k = 2$ along with the TVBM limiter introduced in [11]. Here, we use the mesh size $h = 1/16$ for Examples 3 and 4, and $h = 1/4$ for Example 5. The threshold $10^{-2}$ in (9) depends on the data. It is obtained from numerical experiment. As for the choice of $\theta$, we use $\theta = 1/2$, and we also observe similar results from tests using $\theta \in (0, 1/2)$.

*Example 3* We test the conservative scheme for initial data

$$u_0(x) = \begin{cases} 0.5 & 0 \le x \le 1, \\ -x + 1.5, & 1 \le x \le 4, \\ -2.5, & x \ge 4. \end{cases}$$

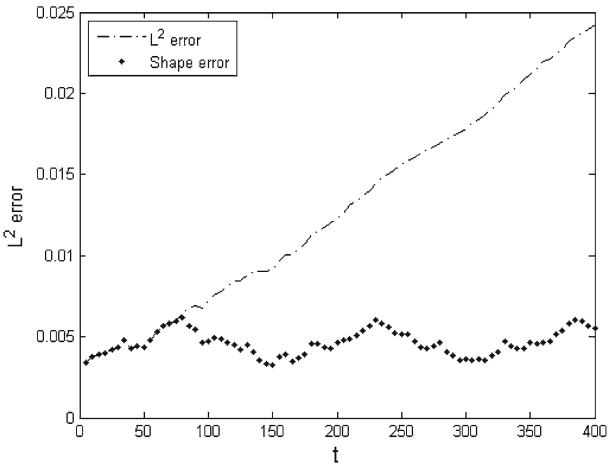



**Fig. 3** Example 2: the evolution of the $L^2$ error and the shape error obtained from the LDG-C scheme

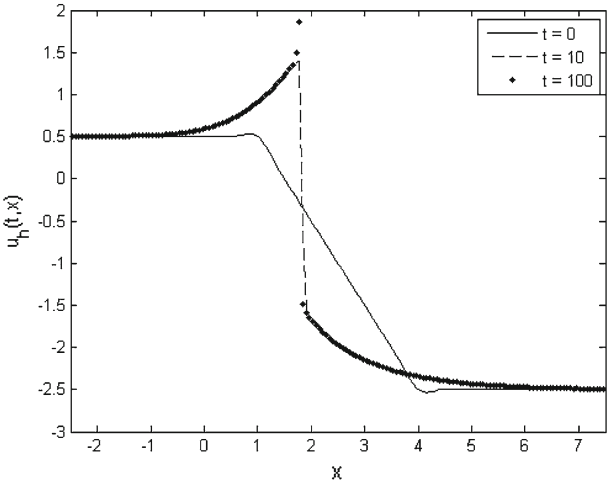



**Fig. 4** Example 3: the computed solution at $t = 0, 10, 100$

This initial data has a downward ramp of height 3 and the constant states lying symmetric with respect to $u = -1$, the solution is expected to converge to a stationary solution. In Fig. 4, we observe a stable pattern formation as analyzed in [16]. In our experiment we use a modified initial data in $C^2$, which agrees with the original data everywhere except for near $x = 1, 4$, so that we can apply directly the TVBM limiter introduced in [11]. Our goal is to observe the stable wave pattern, so the choice of initial modification is not essential.

*Example 4* We consider another initial data of the form

$$u_0(x) = \begin{cases} -0.5 & x \le 8, \\ 15.5 - 2x, & 8 \le x \le 8.5, \\ -1.5, & x \ge 8.5. \end{cases}$$

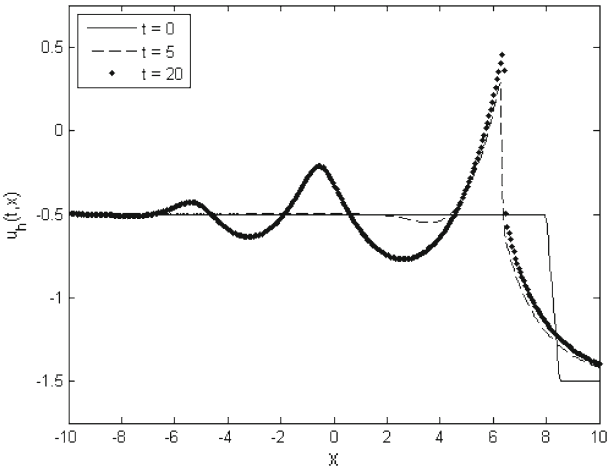



**Fig. 5** Example 4: the computed solution at $t = 0, 5, 20$

This example with smaller jump has no stable stationary solution to converge. We plot the computed solution at the different times in Fig. 5, from which we can see that dispersive effects with oscillations propagate to the left of the ramp as analyzed in [16].

*Example 5* In this example we test interaction of traveling waves. It is known that the interaction of solitons for the KdV equation

$$u_t + uu_x + u_{xxx} = 0,$$

can be illustrated through a family of solutions derived in [20]. One of them reads as follows

$$u(t,x) = 12\frac{k_1^2 e^{\theta_1} + k_2^2 e^{\theta_2} + 2(k_2-k_1)^2 e^{\theta_1+\theta_2} + a^2(k_2^2 e^{\theta_1} + k_1^2 e^{\theta_2})e^{\theta_1+\theta_2}}{(1 + e^{\theta_1} + e^{\theta_2} + a^2 e^{\theta_1+\theta_2})^2},$$

where

$$k_1 = 0.4, \quad k_2 = 0.6, \quad a^2 = \left(\frac{k_1 - k_2}{k_1 + k_2}\right) = \frac{1}{25},$$
$$\theta_1 = k_1 x - k_1^3 t + x_1, \quad \theta_2 = k_2 x - k_2^3 t + x_2, \quad x_1 = 4, \quad x_2 = 5.$$

Since the BP system is dispersive, and close to the KdV equation in some regime of physical parameters, we use $u(0, x)$ as the initial data for the BP system and run the simulation to observe the interaction of two traveling waves. The result is similar to the KdV case [40]: the two peaks travel from left to right, and the speed of the tall one is greater than that of the short one; the taller one eventually passes the shorter one. In addition, oscillations develop on the left as time increases. This is similar to

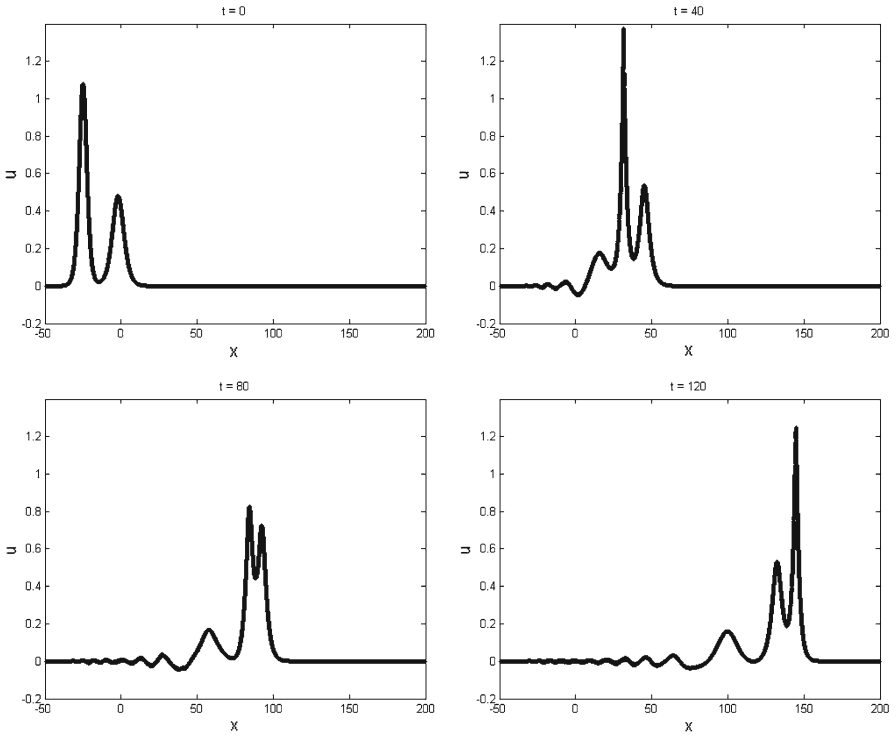



**Fig. 6** Example 5: the evolution of two traveling waves at $t = 0, 40, 80, 120$

the downward ramp of height 2 in Example 4, as shown in Fig. 6. In fact, if we rescale $(t, x)$ by $(\epsilon t, \epsilon(x + t))$ in the BP system, we obtain

$$\partial_t u + uu_x - u_x + (1 - \epsilon^2 \partial_x^2)^{-1} u_x = 0,$$

which to the first order leads to

$$u_t + uu_x + \epsilon^2 u_{xxx} = 0.$$

**Acknowledgments** The authors wish to thank the referees for their careful reading of the manuscript and valuable suggestions resulting in improvements in this paper. This research was partially supported by the National Science Foundation under grant DMS-1312636 and the KI-Net research network.

## References

1. Amadori, D., Gosse, L.: Transient $L^1$ error estimates for well-balanced schemes on non-resonant scalar balance laws. J. Differ. Equ. **255**(3), 469–502 (2013)
2. Bona, J.L., Chen, H., Karakashian, O.A., Xing, Y.: Conservative, discontinuous Galerkin methods for the generalized Korteweg-de Vries equation. Math. Comput. **82**(283), 1401–1432 (2013)

3. Bona, J.L., Dougalis, V.A., Karakashian, O.A., McKinney, W.R.: Conservative, high-order numerical schemes for the generalized Korteweg–de Vries equation. Philos. Trans. R. Soc. Lond. Ser. A **351**(1695), 107–164 (1995)
4. Camassa, R., Holm, D.D.: An integrable shallow water equation with peaked solitons. Phys. Rev. Lett. **71**(11), 1661–1664 (1993)
5. Camassa, R., Holm, D.D., Hyman, J.M.: A new integrable shallow water equation. Adv. Appl. Mech. **31**(31), 1–33 (1994)
6. Castillo, P., Cockburn, B., Schötzau, D., Schwab, C.: Optimal a priori error estimates for the $hp$-version of the local discontinuous Galerkin method for convection-diffusion problems. Math. Comput. **71**(238), 455–478 (2002)
7. Celledoni, E., Grimm, V., McLachlan, R.I., McLaren, D.I., O'Neale, D., Owren, B., Quispel, G.R.W.: Preserving energy resp. dissipation in numerical PDEs using the "average vector field" method. J. Comput. Phys. **231**(20), 6770–6789 (2012)
8. Ciarlet, P.G.: The Finite Element Method for Elliptic Problems. North-Holland, New York (1978)
9. Cockburn, B., Hou, S., Shu, C.-W.: The Runge–Kutta local projection discontinuous Galerkin finite element method for conservation laws. IV. The multidimensional case. Math. Comput. **54**(190), 545–581 (1990)
10. Cockburn, B., Lin, S.Y., Shu, C.-W.: TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. III. One-dimensional systems. J. Comput. Phys. **84**(1), 90–113 (1989)
11. Cockburn, B., Shu, C.-W.: TVB Runge-Kutta local projection discontinuous Galerkin finite element method for conservation laws. II. General framework. Math. Comput. **52**(186), 411–435 (1989)
12. Cockburn, B., Shu, C.-W.: The local discontinuous Galerkin method for time-dependent convection-diffusion systems. SIAM J. Numer. Anal. **35**(6), 2440–2463 (1998). (electronic)
13. Cockburn, B.: The Runge–Kutta discontinuous Galerkin method for conservation laws. V. Multidimensional systems. J. Comput. Phys. **141**(2), 199–224 (1998)
14. Constantin, A., Lannes, D.: The hydrodynamical relevance of the Camassa–Holm and Degasperis–Procesi equations. Arch. Ration. Mech. Anal. **192**(1), 165–186 (2009)
15. Degasperis, A., Kholm, D.D., Khon, A.N.I.: A new integrable equation with peakon solutions. Teoret. Mat. Fiz. **133**(2), 170–183 (2002)
16. Fellner, K., Schmeiser, C.: Burgers-Poisson: a nonlinear dispersive model equation. SIAM J. Appl. Math. **64**(5), 1509–1525 (2004). (electronic)
17. Fuchssteiner, B., Fokas, A.S.: Symplectic structures, their Bäcklund transformations and hereditary symmetries. Phys. D **4**(1), 47–66 (1981/1982)
18. Furihata, D., Mori, M.: General derivation of finite difference schemes by means of a discrete variation. Trans. Japan Soc. Ind. Appl. Math. **8**, 317–340 (1998)
19. Gottlieb, S., Shu, C.-W.: Total variation diminishing Runge–Kutta schemes. Math. Comput. **67**(221), 73–85 (1998)
20. Hirota, R.: Exact solution of the Korteweg–de Vries equation for multiple collisions of solitons. Phys. Rev. Lett. **27**(18), 1192–1194 (1971)
21. Holm, D.D., Staley, M.F.: Wave structure and nonlinear balances in a family of evolutionary PDEs. SIAM J. Appl. Dyn. Syst. **2**(3), 323–380 (2003). (electronic)
22. Johnson, R.S., Camassa-Holm.: Korteweg–de Vries and related models for water waves. J. Fluid Mech. **455**, 63–82 (2002)
23. Liu, H.: Optimal error estimates of the direct discontinuous Galerkin method for convection–diffusion equations. Math. Comput. (2014, accepted)
24. Liu, H., Huang, Y., Yi, N.: A direct discontinuous Galerkin method for the Degasperis–Procesi equation. Methods Appl. Anal. **21**(1), 83–106 (2014)
25. Liu, H., Yan, J.: A local discontinuous Galerkin method for the Korteweg–de Vries equation with boundary effect. J. Comput. Phys. **215**(1), 197–218 (2006)
26. Liu, H., Yan, J.: The direct discontinuous Galerkin (DDG) methods for diffusion problems. SIAM J. Numer. Anal. **47**(1), 675–698 (2008/2009)
27. Liu, H., Yan, J.: The direct discontinuous Galerkin (DDG) method for diffusion with interface corrections. Commun. Comput. Phys. **8**(3), 541–564 (2010)
28. Liu, H., Yin, Z.: Global regularity, and wave breaking phenomena in a class of nonlocal dispersive equations. In Nonlinear partial differential equations and hyperbolic wave phenomena, Contemp. Math., vol. 526, pp. 273–294. Amer. Math. Soc., Providence (2010)

29. Liu, Y., Yin, Z.: Global existence and blow-up phenomena for the Degasperis–Procesi equation. Commun. Math. Phys. **267**(3), 801–820 (2006)
30. Matsuo, T.: Dissipative/conservative Galerkin method using discrete partial derivatives for nonlinear evolution equations. J. Comput. Appl. Math. **218**(2), 506–521 (2008)
31. McLachlan, R.I., Quispel, G.R.W., Robidoux, N.: Geometric integration using discrete gradients. R. Soc. Lond. Philos. Trans. Ser. A Math. Phys. Eng. Sci. **357**(1754), 1021–1045 (1999)
32. Reed, W.H., Hill, T.R.: Triangular mesh methods for the neutron transport equation. Los Alamos, Report LA-UR-73-479 (1973)
33. Warburton, T., Hesthaven, J.S.: On the constants in $hp$-finite element trace inverse inequalities. Comput. Methods Appl. Mech. Eng. **192**(25), 2765–2773 (2003)
34. Whitham, G.B.: Linear and Nonlinear Waves. Wiley, New York (1974)
35. Xu, Y., Shu, C.-W.: A local discontinuous Galerkin method for the Camassa–Holm equation. SIAM J. Numer. Anal. **46**(4), 1998–2021 (2008)
36. Xu, Y., Shu, C.-W.: Local discontinuous Galerkin methods for high-order time-dependent partial differential equations. Commun. Comput. Phys. **7**(1), 1–46 (2010)
37. Xu, Y., Shu, C.W.: Local discontinuous Galerkin methods for two classes of two-dimensional nonlinear wave equations. Phys. D Nonlinear Phenom. **208**(1–2), 21–58 (2005)
38. Xu, Y., Shu, C.W.: Error estimates of the semi-discrete local discontinuous Galerkin method for nonlinear convection–diffusion and KdV equations. Comput. Methods Appl. Mech. Eng. **196**(37–40), 3805–3822 (2007)
39. Yan, J., Shu, C.W.: A local discontinuous Galerkin method for KdV type equations. SIAM J. Numer. Anal. **40**, 769–791 (2003)
40. Yi, N., Huang, Y., Liu, H.: A direct discontinuous Galerkin method for the generalized Korteweg–de Vries equation: energy conservation and boundary effect. J. Comput. Phys. **242**, 351–366 (2013)