Energy-Preserving and Stable Approximations for the Two-Dimensional Shallow Water Equations

Eitan Tadmor* and Weigang Zhong

Department of Mathematics, Center for Scientific Computation and Mathematical Modeling (CSCAMM) and Institute for Physical Science and Technology (IPST), University of Maryland, MD 20742, USA, `tadmor@cscamm.umd.edu`

Summary. We present a systematic development of energy-stable approximations of the two-dimensional shallow water (SW) equations, which are based on the general framework of *entropy conservative* schemes introduced in [Tad03, TZ06]. No artificial numerical viscosity is involved: stability is dictated solely by eddy viscosity. In particular, in the absence of any dissipative mechanism, the resulting numerical schemes *precisely* preserve the total energy, which serves as an entropy function for the SW equations. We demonstrate the dispersive nature of such entropy conservative schemes with a series of scalar examples, interesting for their own sake. We then turn to the SW equations. Numerical experiments of the partial-dam-break problem with energy-preserving and energy stable schemes, successfully simulate the propagation of circular shock and the vortices formed on the both sides of the breach.

1 Introduction

Consider a three-dimensional domain in which the homogenous fluid flows with a free-surface under the influence of gravity. One of the widely used approaches for the description of such unsteady free-surface flows is that of shallow water. Under the shallow-water approximation that refers to the fact that a horizontal scale is in excess of the depth of the fluid, the 3D Navier–Stokes equations can be simplified to the shallow water equations with the depth-averaged continuity equation and momentum equations. Neglecting diffusion of momentum due to wind effects and Coriolis terms, we consider two-dimensional shallow water (SW) equations in the conservative form for free-surface compressible flow with flat frictionless bottom on two dimensional x–y plane,

$$\frac{\partial}{\partial t}\begin{bmatrix} h \\ uh \\ vh \end{bmatrix} + \frac{\partial}{\partial x}\begin{bmatrix} uh \\ u^2h + gh^2/2 \\ uvh \end{bmatrix} + \frac{\partial}{\partial y}\begin{bmatrix} vh \\ uvh \\ v^2h + gh^2/2 \end{bmatrix} = \zeta\frac{\partial}{\partial x}\begin{bmatrix} 0 \\ u_xh \\ v_xh \end{bmatrix} + \zeta\frac{\partial}{\partial y}\begin{bmatrix} 0 \\ u_yh \\ v_yh \end{bmatrix}.$$

(1.1)

Here, $h = h(x, y, t)$ is the total water depth which plays the role of density, and $(u(x, y, t), v(x, y, t))$ are the depth-averaged velocity components along x and y direction. The three equations express, respectively, conservation laws of mass and momentum in x and y direction for the shallow water flow, driven by convective fluxes on the LHS together with eddy viscous fluxes on the RHS. These fluxes involve the constant gravity acceleration $g > 0$, and $\zeta > 0$ is the eddy viscosity. By ignoring the small scale vortices in the motion, we calculate a large-scale flow motion with eddy viscosity ζ that characterizes the transport and dissipation of energy into the smaller scales of the flow.

If we turn off the eddy viscosity ($\zeta = 0$), system (1.1) is reduced to the inviscid shallow water equations,

$$\frac{\partial}{\partial t} \begin{bmatrix} h \\ uh \\ vh \end{bmatrix} + \frac{\partial}{\partial x} \begin{bmatrix} uh \\ u^2h + gh^2/2 \\ uvh \end{bmatrix} + \frac{\partial}{\partial y} \begin{bmatrix} vh \\ uvh \\ v^2h + gh^2/2 \end{bmatrix} = 0. \qquad (1.2)$$

The SW equations (1.1) constitute an incompletely parabolic system, whose solutions can exhibit discontinuities associated with hydraulic jumps and bores in flows or the propagation of sharp fronts. In this paper, we are concerned with construction of energy-stable numerical methods for simulating two dimensional flows, in which initial discontinuities associated with partial-dam-break need to be evolved in time. The conservation of the total energy, $E = (gh^2 + u^2h + v^2h)/2$, guarantees that such numerical simulations of shallow water flows are nonlinearly stable and free of *artificial* numerical viscosity, which may dramatically change the profiles of the solutions in long time integration. In our computation, conservation of the total energy is enforced by utilizing *entropy conservative* fluxes which are tailored to preserve the energy, being an entropy function for the SW equations. The resulting numerical scheme is energy-stable, free of artificial numerical viscosity in the sense that energy dissipation is driven solely by the eddy viscous fluxes. In the particular case that eddy viscosity is absent, $\zeta = 0$, our scheme *precisely* preserves the total energy E.

A general framework for the construction of entropy-conservative schemes for 1D nonlinear conservation laws is introduced in Sect. 3, following [Tad03, TZ06]. We then test these entropy-conservative schemes for 1D Burgers' equation being the prototype of scalar nonlinear conservation laws in Sect. 4. In Sect. 5, we generalize the recipe for the entropy-stable approximations of two dimensional shallow water equations with the energy playing the role of entropy. The extension is carried out dimension by dimension. The algorithm along each dimension follows the same recipe outlined in the one-dimensional setup. The key ingredient behind these schemes is the construction of energy-preserving numerical fluxes. Our main results on the 2D shallow water equations are summarized in Theorem 5.1. To illustrate the performance of the new schemes, we test a two-dimensional partial-dam-break problem in Sect. 6.

The numerical results, especially those of the fine meshes, successfully simulate both the circular shock water wave propagations and the vortices formed on both sides of the breach.

2 Entropy Dissipation: The General Framework

2.1 Entropy Variables

We consider a two-dimensional hyperbolic system,

$$\frac{\partial}{\partial t}\mathbf{u} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{u}) + \frac{\partial}{\partial y}\mathbf{g}(\mathbf{u}) = 0. \tag{2.1}$$

We assume that it obeys an additional conservation law where a convex entropy function $U(\mathbf{u})$ is balanced by entropy fluxes $F(\mathbf{u})$ and $G(\mathbf{u})$,

$$\frac{\partial}{\partial t}U(\mathbf{u}) + \frac{\partial}{\partial x}F(\mathbf{u}) + \frac{\partial}{\partial y}G(\mathbf{u}) = 0. \tag{2.2}$$

Note that (2.2) holds if the entropy function $U(\mathbf{u})$ is linked to the entropy fluxes $F(\mathbf{u})$ and $G(\mathbf{u})$ through the compatibility relations,

$$U_{\mathbf{u}}^{\top}\mathbf{f}_{\mathbf{u}} = F_{\mathbf{u}}^{\top}, \quad U_{\mathbf{u}}^{\top}\mathbf{g}_{\mathbf{u}} = G_{\mathbf{u}}^{\top}. \tag{2.3}$$

In fact, multiplying (2.1) by $U_{\mathbf{u}}^{\top}$ on the left, one recovers the equivalence between (2.1) and (2.3) for all classical solutions $\mathbf{u}$'s of (2.1). These formal manipulations are valid only under the smooth region. To justify these steps in the presence of shock discontinuities, the conservation laws (2.1) are realized as appropriate vanishing viscosity limits, $\mathbf{u} = \lim_{\zeta \downarrow 0} \mathbf{u}^{\zeta}$, where $\mathbf{u}^{\zeta}$ is governed by the (possibly incompletely) parabolic system

$$\frac{\partial}{\partial t}\mathbf{u}^{\zeta} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{u}^{\zeta}) + \frac{\partial}{\partial y}\mathbf{g}(\mathbf{u}^{\zeta}) = \zeta\frac{\partial}{\partial x}\left(Q\frac{\partial}{\partial x}\mathbf{u}^{\zeta}\right) + \zeta\frac{\partial}{\partial y}\left(Q\frac{\partial}{\partial y}\mathbf{u}^{\zeta}\right), \quad \zeta \geq 0. \tag{2.4}$$

Here, $\zeta \downarrow 0$ stands for the vanishing viscosity amplitude such as the eddy viscosity coefficient in the SW equations (1.1)), and $Q = Q(\mathbf{u})$ is any *admissible* viscosity coefficient which is H-symmetric positive-definite,

$$QH = (QH)^{\top} \geq 0, \qquad H := (U_{\mathbf{uu}})^{-1}. \tag{2.5}$$

The passage from vanishing viscosity limits to weak entropy solutions of (2.1) is classical, [Lax73], and we refer to the more comprehensive recent books of e.g., [Ser99, Daf00]. Here, we shall study these limits in terms of the *entropy variables*, $\mathbf{v}(\mathbf{u}) := U_{\mathbf{u}}(\mathbf{u})$. We assume that the entropy $U(\mathbf{u})$ is convex, so that the nonlinear mapping $\mathbf{u} \mapsto \mathbf{v}$ is one-to-one. Following [God61, Moc80], we

claim that the change of variables, $\mathbf{u} = \mathbf{u}(\mathbf{v})$, puts the system (2.1) into the equivalent *symmetric form*,

$$\frac{\partial}{\partial x}\mathbf{u}(\mathbf{v}) + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{u}(\mathbf{v})) + \frac{\partial}{\partial y}\mathbf{g}(\mathbf{u}(\mathbf{v})) = 0.$$

The above system is symmetric in the sense that the Jacobian matrices fluxes are,

$$\mathbf{u_v}(\mathbf{v}) = (\mathbf{u_v}(\mathbf{v}))^\top, \quad \mathbf{f_v}(\mathbf{v}) = (\mathbf{f_v}(\mathbf{v}))^\top, \text{ and } \mathbf{g_v}(\mathbf{v}) = (\mathbf{g_v}(\mathbf{v}))^\top. \quad (2.6)$$

Indeed, a straightforward computation using the compatibility relations (2.3) shows that $\mathbf{u}(\mathbf{v})$, $\mathbf{f}(\mathbf{v})$, and $\mathbf{g}(\mathbf{v})$ are, respectively, the gradients of the corresponding potential functions ϕ, ψ^x, and ψ^y,

$$\mathbf{u}(\mathbf{v}) = \phi_\mathbf{v}(\mathbf{v}), \quad \phi(\mathbf{v}) := \langle \mathbf{v}, \mathbf{u}(\mathbf{v}) \rangle - U(\mathbf{u}(\mathbf{v})), \quad (2.7)$$

$$\mathbf{f}(\mathbf{v}) = \psi^x_\mathbf{v}(\mathbf{v}), \quad \psi^x(\mathbf{v}) := \langle \mathbf{v}, \mathbf{f}(\mathbf{v}) \rangle - F(\mathbf{u}(\mathbf{v})), \quad (2.8)$$

$$\mathbf{g}(\mathbf{v}) = \psi^y_\mathbf{v}(\mathbf{v}), \quad \psi^y(\mathbf{v}) := \langle \mathbf{v}, \mathbf{g}(\mathbf{v}) \rangle - G(\mathbf{u}(\mathbf{v})). \quad (2.9)$$

Hence the Jacobian matrices $H(\mathbf{v}) := \mathbf{u_v}(\mathbf{v})$, $A^x(\mathbf{v}) := \mathbf{f_v}(\mathbf{v})$, and $A^y(\mathbf{v}) := \mathbf{g_v}(\mathbf{v})$ in (2.6) are symmetric, being Hessians of the potentials $\phi(\mathbf{v})$, $\psi^x(\mathbf{v})$, and $\psi^y(\mathbf{v})$. Moreover, the convexity of $U(\cdot)$ implies that H is positive definite, $H = (U_{\mathbf{uu}})^{-1} > 0$.

We now introduce the same entropy change of variables, $\mathbf{u} = \mathbf{u}(\mathbf{v})$, into the associated parabolic system (2.4), which reads

$$\frac{\partial}{\partial t}\mathbf{u}(\mathbf{v}^\varsigma) + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{v}^\varsigma) + \frac{\partial}{\partial y}\mathbf{g}(\mathbf{v}^\varsigma) = \varsigma\frac{\partial}{\partial x}\left(S(\mathbf{v}^\varsigma)\frac{\partial}{\partial x}\mathbf{v}^\varsigma\right) + \varsigma\frac{\partial}{\partial y}\left(S(\mathbf{v}^\varsigma)\frac{\partial}{\partial y}\mathbf{v}^\varsigma\right). \quad (2.10\text{a})$$

By (2.6) and admissibility condition (2.5), the system (2.10a) is symmetric in the sense that the Jacobian matrices involved are all symmetric, namely, (2.6) holds and

$$S(\mathbf{u}(\mathbf{v})) = S^\top(\mathbf{u}(\mathbf{v})) > 0, \quad S(\mathbf{v}) := Q(\mathbf{u}(\mathbf{v}))\mathbf{u_v}(\mathbf{v}). \quad (2.10\text{b})$$

Integrate (2.4) against the *entropy variable* $\mathbf{v} := U_\mathbf{u}$, employ the compatibility relations (2.3) and use 'differentiation by parts' on the dissipation terms on the RHS to find the following entropy balance statement,

$$\frac{\partial}{\partial t}U(\mathbf{u}^\varsigma) + \frac{\partial}{\partial x}\left(F(\mathbf{u}^\varsigma) - \varsigma\langle \mathbf{v}^\varsigma, Q\mathbf{u}^\varsigma_x \rangle\right) + \frac{\partial}{\partial y}\left(G(\mathbf{u}^\varsigma) - \varsigma\langle \mathbf{v}^\varsigma, Q\mathbf{u}^\varsigma_y \rangle\right) =$$
$$-\varsigma\left[\langle \mathbf{v}^\varsigma_x, S(\mathbf{v}^\varsigma)\mathbf{v}^\varsigma_x \rangle + \langle \mathbf{v}^\varsigma_y, S(\mathbf{v}^\varsigma)\mathbf{v}^\varsigma_y \rangle\right] \leq 0. \quad (2.11)$$

Letting $\varsigma \downarrow 0$, we obtain the *entropy inequality*, [God61, Kru70, Lax71]

$$\frac{\partial}{\partial t}U(\mathbf{u}) + \frac{\partial}{\partial x}F(\mathbf{u}) + \frac{\partial}{\partial y}G(\mathbf{u}) \leq 0. \quad (2.12)$$

This shows that weak solution dissipates entropy. The precise amount of entropy decay is dictated by the specific dissipation: spatial integration of (2.11) yields the entropy decay statement,

$$\frac{d}{dt} \int_y \int_x U(\mathbf{u}^\varsigma)\, dxdy = -\varsigma \int_y \int_x \left[\langle \mathbf{v}_x^\varsigma, S(\mathbf{v}^\varsigma)\mathbf{v}_x^\varsigma \rangle + \langle \mathbf{v}_y^\varsigma, S(\mathbf{v}^\varsigma)\mathbf{v}_y^\varsigma \rangle \right]\, dxdy \leq 0.$$

$$(2.13)$$

2.2 The Example of the Shallow Water Equations

We consider the 2D shallow water equations (1.1) for the conservative variables $\mathbf{u} := (h, uh, vh)^\top$ where h is the water-depth and u, v are depth-averaged velocity components along x and y-direction. The total energy is given by the depth-averaged sum of the potential and kinetic energies,

$$E(\mathbf{u}) := \frac{gh^2 + u^2h + v^2h}{2}.$$

$$(2.14a)$$

The total energy plays the role of an entropy function for the SW equations. Straightforward computation gives us the following entropy fluxes, entropy variables and potentials.

- Entropy fluxes

$$F(\mathbf{u}) = guh^2 + \frac{u^3h + uv^2h}{2}, \quad G(\mathbf{u}) = gvh^2 + \frac{u^2vh + v^3h}{2}.$$

$$(2.14b)$$

- Entropy variable

$$\mathbf{v}(\mathbf{u}) = \begin{bmatrix} gh - \dfrac{u^2 + v^2}{2} \\ u \\ v \end{bmatrix}$$

$$(2.14c)$$

with the Jacobian matrices, $H := \mathbf{u_v}$ and $H^{-1} = \mathbf{v_u}$, given by

$$H = \frac{1}{g} \begin{bmatrix} 1 & u & v \\ u & c^2 + u^2 & uv \\ v & uv & c^2 + v^2 \end{bmatrix}, \quad H^{-1} = \frac{1}{h} \begin{bmatrix} c^2 + u^2 + v^2 & -u & -v \\ -u & 1 & 0 \\ -v & 0 & 1 \end{bmatrix},$$

$$(2.14d)$$

where $c := \sqrt{gh}$ is the 'sound' speed, or wave celerity.

- The potentials of the temporal and spatial fluxes $\mathbf{u}(\mathbf{v})$, $\mathbf{f}(\mathbf{u}(\mathbf{v}))$ and $\mathbf{g}(\mathbf{u}(\mathbf{v}))$ are given, respectively, by

$$\phi(\mathbf{v}) = \frac{gh^2}{2}, \quad \psi^x(\mathbf{v}) = \frac{guh^2}{2}, \quad \psi^y(\mathbf{v}) = \frac{gvh^2}{2}.$$

$$(2.14e)$$

The general statement of entropy balance, (2.13), amounts to

$$\frac{d}{dt}\int_y\int_x E(\mathbf{u})\,dxdy = -\zeta\int_y\int_x h(u_x^2+u_y^2+v_x^2+v_y^2)\,dxdy, \quad E(\mathbf{u}) = \frac{gh^2+u^2h+v^2h}{2}. \tag{2.15}$$

Since $h \geq 0$, we conclude that the total energy is decreasing in time, thus recovering energy stability. In fact, the expression on the RHS of (2.2) specifies the precise decay rate, which is dictated solely by the viscous fluxes through their dependence on the nonnegative eddy viscosity ζ. Our objective in this paper is to construct "faithful" approximations to the 2D shallow water equations, which precisely reproduce the energy balance (2.2).

3 Entropy Conservative Schemes: The 1D Setup

Setting $\mathbf{g} \equiv 0$ in (2.1), we consider the one-dimensional system of hyperbolic conservation laws,

$$\frac{\partial\mathbf{u}}{\partial t} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{u}) = 0, \quad x \in \mathbb{R},\ t > 0, \tag{3.1}$$

governing the N-vector of conserved variables $\mathbf{u} = [u_1,\cdots,u_N]^\top$ and balanced by the flux functions $\mathbf{f} = [f_1,\cdots,f_N]^\top$. We assume it is endowed with an entropy pair, (U,F), such that every strong solution of (3.1) satisfies the entropy equality

$$\frac{\partial}{\partial t}U(\mathbf{u}) + \frac{\partial}{\partial x}F(\mathbf{u}) = 0, \tag{3.2}$$

whereas weak solutions are sought to satisfy the entropy inequality, $U(\mathbf{u})_t + F(\mathbf{u})_x \leq 0$.

We now turn our attention to consistent approximations of (3.1), (3.2), based on semi-discrete conservative schemes of the form

$$\frac{d}{dt}\mathbf{u}_\nu(t) = -\frac{1}{\Delta x}\left(\mathbf{f}_{\nu+\frac{1}{2}} - \mathbf{f}_{\nu-\frac{1}{2}}\right). \tag{3.3}$$

Here, $\mathbf{u}_\nu(t)$ denotes the discrete solution along the equally spaced grid lines, $(x_\nu := \nu\Delta x, t)$, and $\mathbf{f}_{\nu+\frac{1}{2}}$ is the Lipschitz-continuous numerical flux which occupies a stencil of $2p$-gridvalues,

$$\mathbf{f}_{\nu+\frac{1}{2}} = \mathbf{f}(\mathbf{u}_{\nu-p+1},\cdots,\mathbf{u}_{\nu+p}).$$

The scheme is *consistent* with the system (3.1) if $\mathbf{f}(\mathbf{u},\mathbf{u},\cdots,\mathbf{u}) = \mathbf{f}(\mathbf{u})$, $\forall\mathbf{u} \in \mathbb{R}^N$. Making the change of variables $\mathbf{u}_\nu = \mathbf{u}(\mathbf{v}_\nu)$, we obtain the equivalent form of (3.3)

$$\frac{d}{dt}\mathbf{u}(\mathbf{v}_\nu(t)) = -\frac{1}{\Delta x}\left(\mathbf{f}_{\nu+\frac{1}{2}} - \mathbf{f}_{\nu-\frac{1}{2}}\right). \tag{3.4}$$

The essential difference lies with the numerical flux, $\mathbf{f}_{\nu+\frac{1}{2}}$, which is now expressed in terms of the entropy variables,

$$\mathbf{f}_{\nu+\frac{1}{2}} = \mathbf{f}\left(\mathbf{v}_{\nu-p+1}, \cdots, \mathbf{v}_{\nu+p}\right) := \mathbf{f}\left(\mathbf{u}\left(\mathbf{v}_{\nu-p+1}\right), \cdots, \mathbf{u}\left(\mathbf{v}_{\nu+p}\right)\right),$$

consistent with the differential flux, $\mathbf{f}(\mathbf{v}, \mathbf{v}, \cdots, \mathbf{v}) = \mathbf{f}(\mathbf{v}) \equiv \mathbf{f}(\mathbf{u}(\mathbf{v}))$. The semi-discrete schemes (3.3) and (3.4) are completely identical. The entropy variables-based formula (3.4) has the advantage that it provides a natural ordering of symmetric matrices, which in turn enables us to *compare* the numerical viscosities of different schemes, consult [Tad87] for details. In particular, we will be able to utilize the so called entropy conservative discretization of [Tad03] for the convective part of the system of conservation laws (3.1), and thus recover the precise entropy balance dictated by physical dissipative terms of the underlying original systems.

The scheme (3.3) is called *entropy-conservative* if it satisfies a discrete *entropy equality*,

$$\frac{d}{dt}U\left(\mathbf{u}_{\nu}(t)\right) + \frac{1}{\Delta x}\left(F_{\nu+\frac{1}{2}} - F_{\nu-\frac{1}{2}}\right) = 0, \tag{3.5}$$

where $F_{\nu+\frac{1}{2}} = F\left(\mathbf{u}_{\nu-p+1}, \cdots, \mathbf{u}_{\nu+p}\right)$ is a consistent numerical entropy flux, $F(\mathbf{u}, \mathbf{u}, \cdots, \mathbf{u}) = F(\mathbf{u})$, $\forall \mathbf{u} \in \mathbb{R}^{N}$. Entropy conservative schemes will play an essential role in the construction of entropy stable schemes, by adding a judicious amount of physical viscosity.

The key step in the construction of entropy conservative schemes for the systems of conservation laws is the choice of *an arbitrary* piecewise-constant path in phase space. We shall use the phase space of the entropy variable $\mathbf{v}$ to connect two neighboring gridvalues, $\mathbf{v}_{\nu}$ and $\mathbf{v}_{\nu+1}$, at the spatial cell $[x_{\nu}, x_{\nu+1}]$, through the intermediate states $\{\mathbf{v}_{\nu+\frac{1}{2}}^{j}\}_{j=1}^{N}$. To this end, let $\{\mathbf{r}_{j} \equiv \mathbf{r}_{\nu+\frac{1}{2}}^{j}\}_{j=1}^{N}$ be an arbitrary set of N linearly independent N-vectors, and let $\{\boldsymbol{\ell}_{j} \equiv \boldsymbol{\ell}_{\nu+\frac{1}{2}}^{j}\}_{j=1}^{N}$ be the corresponding orthogonal set. We introduce the intermediate gridvalues, $\{\mathbf{v}_{\nu+\frac{1}{2}}^{j}\}_{j=1}^{N}$, which define a piecewise constant path in phase space across the jump $\Delta\mathbf{v}_{\nu+\frac{1}{2}} := \mathbf{v}_{\nu+1} - \mathbf{v}_{\nu}$,

$$\begin{cases} \mathbf{v}_{\nu+\frac{1}{2}}^{1} = \mathbf{v}_{\nu} \\ \mathbf{v}_{\nu+\frac{1}{2}}^{j+1} = \mathbf{v}_{\nu+\frac{1}{2}}^{j} + \left\langle \boldsymbol{\ell}_{j}, \Delta\mathbf{v}_{\nu+\frac{1}{2}} \right\rangle \mathbf{r}_{j}, \quad j = 1, 2, \cdots, N-1, . \\ \mathbf{v}_{\nu+\frac{1}{2}}^{N+1} = \mathbf{v}_{\nu+1} \end{cases} \tag{3.6}$$

Theorem 3.1 (Tadmor [Tad03, Theorem 6.1]). *Consider the system of conservation laws (3.1). Given the entropy pair (U, F), then the conservative scheme*

$$\frac{d}{dt}\mathbf{u}_{\nu}(t) = -\frac{1}{\Delta x_{\nu}}\left(\mathbf{f}_{\nu+\frac{1}{2}}^{*} - \mathbf{f}_{\nu-\frac{1}{2}}^{*}\right) \tag{3.7}$$

with a numerical flux $\mathbf{f}^*_{\nu+\frac{1}{2}}$

$$\mathbf{f}^*_{\nu+\frac{1}{2}} = \sum_{j=1}^{N} \frac{\psi\left(\mathbf{v}^{j+1}_{\nu+\frac{1}{2}}\right) - \psi\left(\mathbf{v}^{j}_{\nu+\frac{1}{2}}\right)}{\left\langle \boldsymbol{\ell}_j, \Delta\mathbf{v}_{\nu+\frac{1}{2}} \right\rangle} \boldsymbol{\ell}_j \tag{3.8}$$

is an entropy-conservative approximation, consistent with (3.1) and (3.2). Here, $\mathbf{v} = U_{\mathbf{u}}(\mathbf{u})$ *are the entropy variables associated with the entropy* U, *and* $\psi(\mathbf{v}) := \langle \mathbf{v}, \mathbf{f}(\mathbf{u}(\mathbf{v})) \rangle - F(\mathbf{u}(\mathbf{v}))$ *is the entropy potential.*

The proof is based on the fact that the entropy equality (3.5) holds if and only if $\left\langle \Delta\mathbf{v}_{\nu+\frac{1}{2}}, \mathbf{f}^*_{\nu+\frac{1}{2}} \right\rangle$ equals a *conservative difference*,

$$\left\langle \Delta\mathbf{v}_{\nu+\frac{1}{2}}, \mathbf{f}^*_{\nu+\frac{1}{2}} \right\rangle = \Delta\psi_{\nu+\frac{1}{2}}, \qquad \Delta\psi_{\nu+\frac{1}{2}} := \psi(\mathbf{v}_{\nu+1}) - \psi(\mathbf{v}_\nu). \tag{3.9}$$

Indeed, (3.9) is equivalent to (3.5),

$$\left\langle \mathbf{v}_\nu, \mathbf{f}^*_{\nu+\frac{1}{2}} - \mathbf{f}^*_{\nu-\frac{1}{2}} \right\rangle = F_{\nu+\frac{1}{2}} - F_{\nu-\frac{1}{2}}, \tag{3.10a}$$

where the numerical entropy flux $F_{\nu+\frac{1}{2}}$ is given by

$$F_{\nu+\frac{1}{2}} = \frac{1}{2}\left[\left\langle \mathbf{v}_\nu + \mathbf{v}_{\nu+1}, \mathbf{f}^*_{\nu+\frac{1}{2}} \right\rangle - \left(\psi(\mathbf{v}_\nu) + \psi(\mathbf{v}_{\nu+1}) \right) \right] \tag{3.10b}$$

A straightforward manipulation of the numerical flux (3.8) confirms the desired equality (3.9),

$$\left\langle \Delta\mathbf{v}_{\nu+\frac{1}{2}}, \mathbf{f}^*_{\nu+\frac{1}{2}} \right\rangle = \sum_{j=1}^{N} \frac{\psi\left(\mathbf{v}^{j+1}_{\nu+\frac{1}{2}}\right) - \psi\left(\mathbf{v}^{j}_{\nu+\frac{1}{2}}\right)}{\left\langle \boldsymbol{\ell}_j, \Delta\mathbf{v}_{\nu+\frac{1}{2}} \right\rangle} \left\langle \boldsymbol{\ell}_j, \Delta\mathbf{v}_{\nu+\frac{1}{2}} \right\rangle$$

$$= \sum_{j=1}^{N} \psi\left(\mathbf{v}^{j+1}_{\nu+\frac{1}{2}}\right) - \psi\left(\mathbf{v}^{j}_{\nu+\frac{1}{2}}\right) = \psi\left(\mathbf{v}^{N+1}_{\nu+\frac{1}{2}}\right) - \psi\left(\mathbf{v}^{1}_{\nu+\frac{1}{2}}\right)$$

$$= \Delta\psi_{\nu+\frac{1}{2}}.$$

Although the recipe for constructing entropy-conservative fluxes in (3.8) allows an *arbitrary* choice of a path in phase space, inappropriate choices of the path may cause the computed intermediate values to lie outside the physical space, say $h < 0$. A 'physically relevant' choice is offered by a Riemann path which consists of $\{\mathbf{u}^{j}_{\nu+\frac{1}{2}}\}_{j=1}^{N}$, stationed along an (approximate) set of right eigenvectors, $\{\hat{\mathbf{r}}_j\}$, of the Jacobian $\mathbf{f}_{\mathbf{u}}(\mathbf{u}_{\nu+\frac{1}{2}})$. Set $\mathbf{v}^{j}_{\nu+\frac{1}{2}} = \mathbf{v}(\mathbf{u}^{j}_{\nu+\frac{1}{2}})$, $j = 1, 2, \ldots, N$, and let $\boldsymbol{\ell}_j$'s be the orthogonal system to $\{\mathbf{v}^{j+1} - \mathbf{v}^j\}_{j=1}^{N}$. This will be our choice of a path for computing entropy stable approximations of shallow water equations in Sect. 5 below. The construction of the entropy conservative flux $\mathbf{f}^*_{\nu+\frac{1}{2}}$ follows [TZ06, Algorithm 1] which states,

Algorithm 1 *If* $\quad \mathbf{u}_\nu = \mathbf{u}_{\nu+1}$ *then* $\quad \mathbf{f}^*_{\nu+\frac{1}{2}} = \mathbf{f}(\mathbf{v}_\nu)$; *else*

- *Set* $\mathbf{u}^1_{\nu+\frac{1}{2}} := \mathbf{u}_\nu$ *and compute recursively the intermediate states,*

$$\mathbf{u}^{j+1}_{\nu+\frac{1}{2}} = \mathbf{u}^{j}_{\nu+\frac{1}{2}} + \left\langle \widehat{\boldsymbol{\ell}}_j, \Delta\mathbf{u}_{\nu+\frac{1}{2}} \right\rangle \widehat{\mathbf{r}}_j, \quad j = 1, 2, 3. \tag{3.11}$$

 Here, $\{\widehat{\boldsymbol{\ell}}_j\}$ *and* $\{\widehat{\mathbf{r}}_j\}$ *are the left and right eigensystems of an averaged Jacobian* $\widetilde{A}_{\nu+\frac{1}{2}}$, *given by the Roe matrix,* $\widetilde{A}_{\nu+\frac{1}{2}} = \widetilde{A}(\mathbf{u}_\nu, \mathbf{u}_{\nu+1})$ *(see [Roe81]).*
- *Set* $\mathbf{r}_j := \mathbf{v}(\mathbf{u}^{j+1}_{\nu+\frac{1}{2}}) - \mathbf{v}(\mathbf{u}^{j}_{\nu+\frac{1}{2}})$ *and compute* $\{\boldsymbol{\ell}_j\}^3_{j=1}$ *as the corresponding orthogonal system. (Note that* $\{\mathbf{r}_j, \boldsymbol{\ell}_j\}$ *is the eigen-path in* $\mathbf{v}$*-space, corresponding to the eigen-path in* $\mathbf{u}$*-space,* $\{\widehat{\mathbf{r}}_j, \widehat{\boldsymbol{\ell}}_j\}$.*)*
- *Compute the entropy-conservative numerical flux,*

$$\mathbf{f}^*_{\nu+\frac{1}{2}} = \sum_{j=1}^{3} \frac{\psi(\mathbf{v}^{j+1}_{\nu+\frac{1}{2}}) - \psi(\mathbf{v}^{j}_{\nu+\frac{1}{2}})}{\left\langle \boldsymbol{\ell}_j, \Delta\mathbf{v}_{\nu+\frac{1}{2}} \right\rangle} \boldsymbol{\ell}_j. \tag{3.12}$$

4 Scalar Problems

We test our entropy stable schemes with the prototype example of inviscid Burgers' equation. Though very simple, the inviscid Burgers' equation is often used as the testing ground for numerical approximations of nonlinear conservation laws.

4.1 Entropy Conservative Schemes

We consider the inviscid Burgers' equation,

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}f(u) = 0, \quad f(u) = \frac{1}{2}u^2. \tag{4.1}$$

Any convex function $U(u)$ serves as an entropy function for the scalar Burgers equation. The solutions of (4.1) satisfy, at the *formal* level,

$$\frac{\partial}{\partial t}U(u) + \frac{\partial}{\partial x}F(u) = 0. \tag{4.2}$$

These are additional conservation laws balanced by the corresponding entropy flux functions $F(u)$ satisfying the compatibility relation $U'f' = F'$. Spatial integration then yields the total entropy conservation (ignoring boundary contributions)

$$\int_x U(x, t)\, dx = \int_x U(x, 0)\, dx. \tag{4.3}$$

We now turn to the discrete framework. Discretization in space yields the semi-discrete scheme,

$$\frac{d}{dt}u_\nu(t) + \frac{1}{\Delta x}\left(f_{\nu+\frac{1}{2}} - f_{\nu-\frac{1}{2}}\right) = 0. \tag{4.4}$$

Clearly, $\sum u_\nu(t)\Delta x$ is conserved. We seek a consistent numerical flux $f_{\nu+\frac{1}{2}}$, that is entropy conservative in the sense of satisfying the discrete analogue of (4.2),

$$\frac{d}{dt}U(u_\nu(t)) + \frac{1}{\Delta x}(F_{\nu+\frac{1}{2}} - F_{\nu-\frac{1}{2}}) = 0,$$

so that we have the additional conservation of entropy $\sum U(u_\nu(t))\Delta x$. According to Theorem 3.1, consult (3.9), such 2-point *scalar* entropy conservative fluxes are *uniquely* determined, $f_{\nu+\frac{1}{2}} = f^*_{\nu+\frac{1}{2}}$, by

$$f_{\nu+\frac{1}{2}} = f^*_{\nu+\frac{1}{2}} := \frac{\psi(u_{\nu+1}) - \psi(u_\nu)}{v(u_{\nu+1}) - v(u_\nu)}. \tag{4.5}$$

Recall that $v(u) := U'(u)$ is the entropy variable associated with the entropy pair (U, F), and $\psi(u) := v(u)f(u) - F(u)$ is the potential function of the flux $f(u(v))$. We demonstrate the constructions of above entropy conservative numerical flux with two different choices of entropy functions:

- We begin with the logarithmic entropy $U(u) = -\ln u$ together with the entropy flux $F(u) = -u$. We use the entropy variable $v(u) = -1/u$. The entropy flux potential in this case is $\psi(u) = -1/2v = u/2$. The entropy conservative numerical flux (4.5) then reads,

$$f^*_{\nu+\frac{1}{2}} := \frac{\psi(u_{\nu+1}) - \psi(u_\nu)}{v(u_{\nu+1}) - v(u_\nu)} = \frac{1}{2}u_\nu u_{\nu+1}.$$

This numerical flux yields the entropy conservative schemes

$$\frac{d}{dt}u_\nu(t) = u_\nu(t)\frac{u_{\nu+1}(t) - u_{\nu-1}(t)}{2\Delta x}.$$

This scheme was discussed by Goodman and Lax in [GL88], Hou and Lax in [HL91], and Levermore and Liu in [LL96] in their study of the *dispersive oscillations* arising in numerical solutions of the conservative schemes for the inviscid Burgers' equation.
- Next, we consider the family of entropy functions,

$$U_p(u) = u^{2p}, \quad p = 1, 2, \cdots, \tag{4.6}$$

with the corresponding entropy flux functions $F_p(u) = 2pu^{2p+1}/(2p + 1)$. Using the entropy variable $v(u) := U'(u) = 2pu^{2p-1}$ and the potential function $\psi(u) := v(u)f(u) - F(u) = \frac{p(2p-1)}{2p+1}u^{2p+1}$, we compute the entropy conservative flux

$$f^*_{\nu+\frac{1}{2}} := \frac{\psi(u_{\nu+1}) - \psi(u_\nu)}{v(u_{\nu+1}) - v(u_\nu)} = \frac{2p-1}{2(2p+1)} \cdot \frac{u_{\nu+1}^{2p+1} - u_\nu^{2p+1}}{u_{\nu+1}^{2p-1} - u_\nu^{2p-1}}. \tag{4.7}$$

The resulting scheme (4.4), (4.7) is entropy conservative in the sense that the discrete analogue of total entropy conservation (4.3) is satisfied,

$$\sum_\nu u_\nu^{2p}(t)\, \Delta x = \sum_\nu u_\nu^{2p}(0)\, \Delta x.$$

Thus, for each p we obtain its own U_p-entropy conservative scheme.

Remark 4.1. Although these schemes with the entropy-conservative flux (4.7) admit the dispersive oscillations shown in the numerical results of Sect. 4.3, we expect the amplitude of these oscillations to be reduced for increasing p's, as the conservation of entropies U_p,

$$\left[\sum_\nu u_\nu^{2p}(t)\, \Delta x \right]^{\frac{1}{2p}} = \left[\sum_\nu u_\nu^{2p}(0)\, \Delta x \right]^{\frac{1}{2p}} \tag{4.8}$$

approaches the maximum principle, $||u_\nu(t)||_{L^\infty} \le ||u_\nu(0)||_{L^\infty}$ (the inequality reflects the small amount of dissipation due to time discretization). Indeed, as $p \uparrow \infty$, the entropy-conservative schemes based on (4.7) approach the first-order entropy stable Engquist–Osher scheme [EO80].

4.2 Entropy Dissipation

To recover the physical relevant *entropy inequality*, that is

$$\partial_t U_p(u) + \partial_x F_p(u) \le 0,$$

one can add numerical dissipation,

$$\frac{d}{dt} u_\nu(t) + \frac{1}{\Delta x} \left(f^*_{\nu+\frac{1}{2}} - f^*_{\nu-\frac{1}{2}} \right) = \frac{\epsilon}{(\Delta x)^2} \Big(d(u_{\nu+1}) - 2d(u_\nu) + d(u_{\nu-1}) \Big), \quad \epsilon > 0. \tag{4.9}$$

This serves as an approximation to the vanishing viscosity regularization

$$u_t + f(u)_x = \epsilon d(u)_{xx}, \quad d'(u) > 0, \ \epsilon > 0.$$

Sum this scheme (4.9) against the entropy variable v_ν to find

$$\frac{d}{dt} \sum_\nu U_p(u_\nu(t))\Delta x + \sum_\nu v_\nu \left(f^*_{\nu+\frac{1}{2}} - f^*_{\nu-\frac{1}{2}} \right)$$
$$= \epsilon \sum_\nu v_\nu \frac{d(u_{\nu+1}) - 2d(u_\nu) + d(u_{\nu-1})}{\Delta x}. \tag{4.10}$$

According to (3.10a), the second term on the left of (4.10) vanishes, $\sum \left(F_{\nu+\frac{1}{2}} - F_{\nu-\frac{1}{2}} \right) \Delta x = 0$. Summation by parts on the RHS of (4.10) yields

$$\epsilon \sum_\nu v_\nu \frac{d(u_{\nu+1}) - 2d(u_\nu) + d(u_{\nu-1})}{\Delta x} = -\frac{\epsilon}{\Delta x} \sum_\nu \left(v_{\nu+1} - v_\nu \right)$$
$$\times \left(d(u_{\nu+1}) - d(u_\nu) \right) \le 0,$$

since $d'(v) = d'(u)u'(v) > 0$, and hence $(v_{\nu+1} - v_\nu) \cdot (d(u_{\nu+1}) - d(u_\nu)) > 0$. The resulting entropy balance that follows reads,

$$\frac{d}{dt} \sum_\nu U_p(u_\nu(t)) \Delta x = -\frac{\epsilon}{\Delta x} \sum_\nu \Delta v_{\nu+\frac{1}{2}} \Delta d_{\nu+\frac{1}{2}} \le 0. \qquad (4.11)$$

Observe that the amount of entropy dissipation on the right is completely determined by the dissipation term $\epsilon d(u)$. No artificial viscosity is introduced by the convective term. If we exclude any dissipative mechanism ($\epsilon = 0$), then we are back at the entropy conservative schemes of Sect. 4.1.

4.3 Numerical Experiments

Time Discretization

To complete the computation of a semi-discrete scheme, the semi-discrete entropy conservative scheme (4.4), (4.7) needs to be augmented with a proper time discretization. To enable a large time-stability region and maintain simplicity, the explicit three-stage third-order Runge–Kutta (RK3) method will be used, Consult [GST01] for more detail of its strong stability-preserving property,

$$\begin{cases} u^{(1)} = u^n + \Delta t \mathcal{L}(u^n) \\ u^{(2)} = \frac{3}{4}u^n + \frac{1}{4}u^{(1)} + \frac{1}{4}\Delta t \mathcal{L}(u^{(1)}) \\ u^{n+1} = \frac{1}{3}u^n + \frac{2}{3}u^{(2)} + \frac{2}{3}\Delta t \mathcal{L}(u^{(2)}) \end{cases} \qquad (4.12a)$$

where

$$[\mathcal{L}(u)]_\nu := -\frac{1}{\Delta x}(f^*_{\nu+\frac{1}{2}} - f^*_{\nu-\frac{1}{2}}). \qquad (4.12b)$$

We note that this explicit RK3 time discretization produces a negligible amount of entropy dissipation. For a general framework of entropy conservative fully discrete schemes, consult [LMR02].

Continuous Initial Conditions

We first solve the inviscid Burgers equation (4.1) in the domain $x \in [0, 1]$ with initial condition, $u(0, x) = \sin(2\pi x)$ and subject to periodic boundary conditions $u(t, 1) = u(t, 0)$. In Fig. 1 we display the numerical solutions for (4.12a) and (4.12b) with the numerical flux (4.7) for different choices of p. For small values of p, the dispersive oscillations become noticeable after the shock is generated due to the absence of any dissipative mechanism in the entropy-conservative scheme. As p increases, the amplitude of the spurious dispersive oscillations decreases, which reflects the control of the increasing L^{2p}-norms in (4.8).
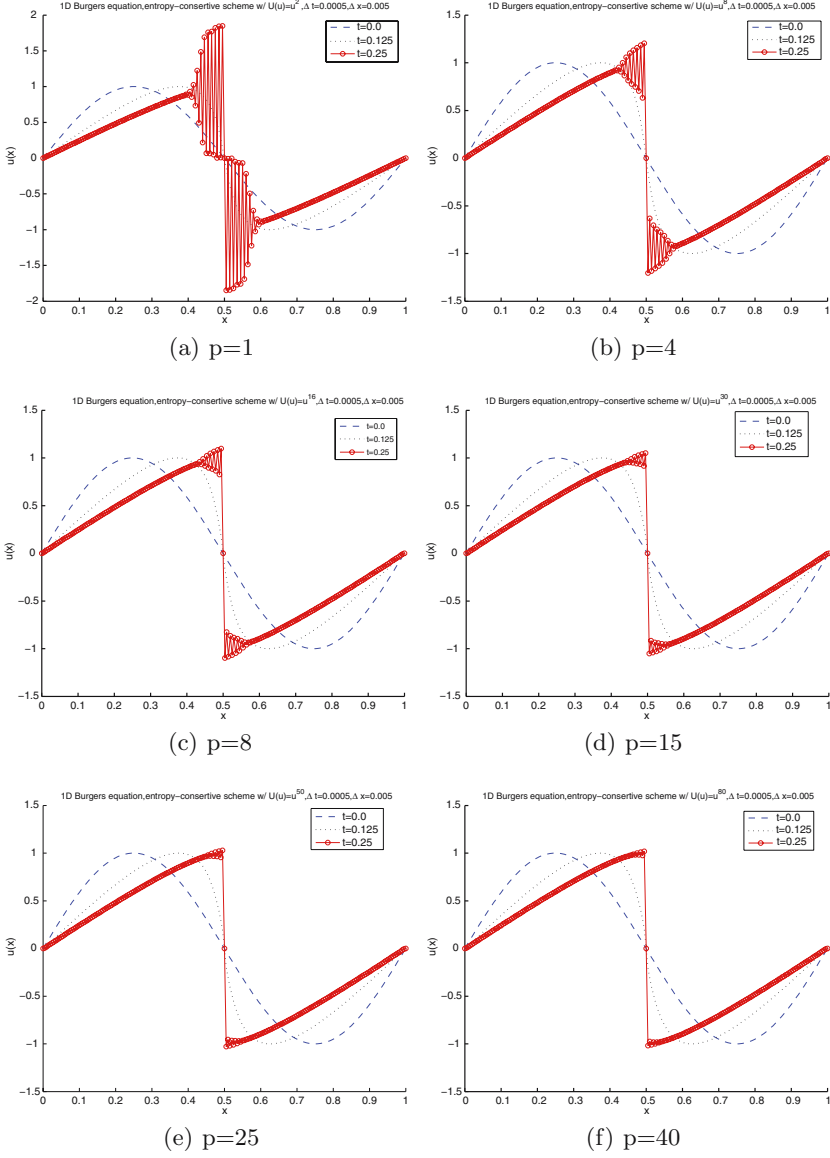
Fig. 1. 1D Burger's equation, sine initial condition, entropy-conservative schemes, 200 spatial grids, $U(u) = u^{2p}$

Discontinuous Initial Conditions

We solve the 1D inviscid Burgers equation (4.1) in the domain $x \in [0, 1]$ with the discontinuous initial condition,

$$u(0, x) = \begin{cases} 2, & x \in [0,\, 0.5] \\ 1, & x \in (0.5,\, 1] \end{cases}$$

The boundary values are extrapolated from the interior points. Since we are only interested in the propagation of the shock wave in the computational domain $[0,\, 1]$, there is interaction with the boundary values which do not vary in the time interval under consideration. In Fig. 2, we display the numerical solutions for (4.12a) and (4.12b) with the numerical flux (4.7) for different choices of p. Those solutions show the same pattern as the $\sin(2\pi x)$ initial condition. Diminishing amplitude of the dispersive oscillations demonstrates the control of the L^{2p}-norm of the solution with each p.

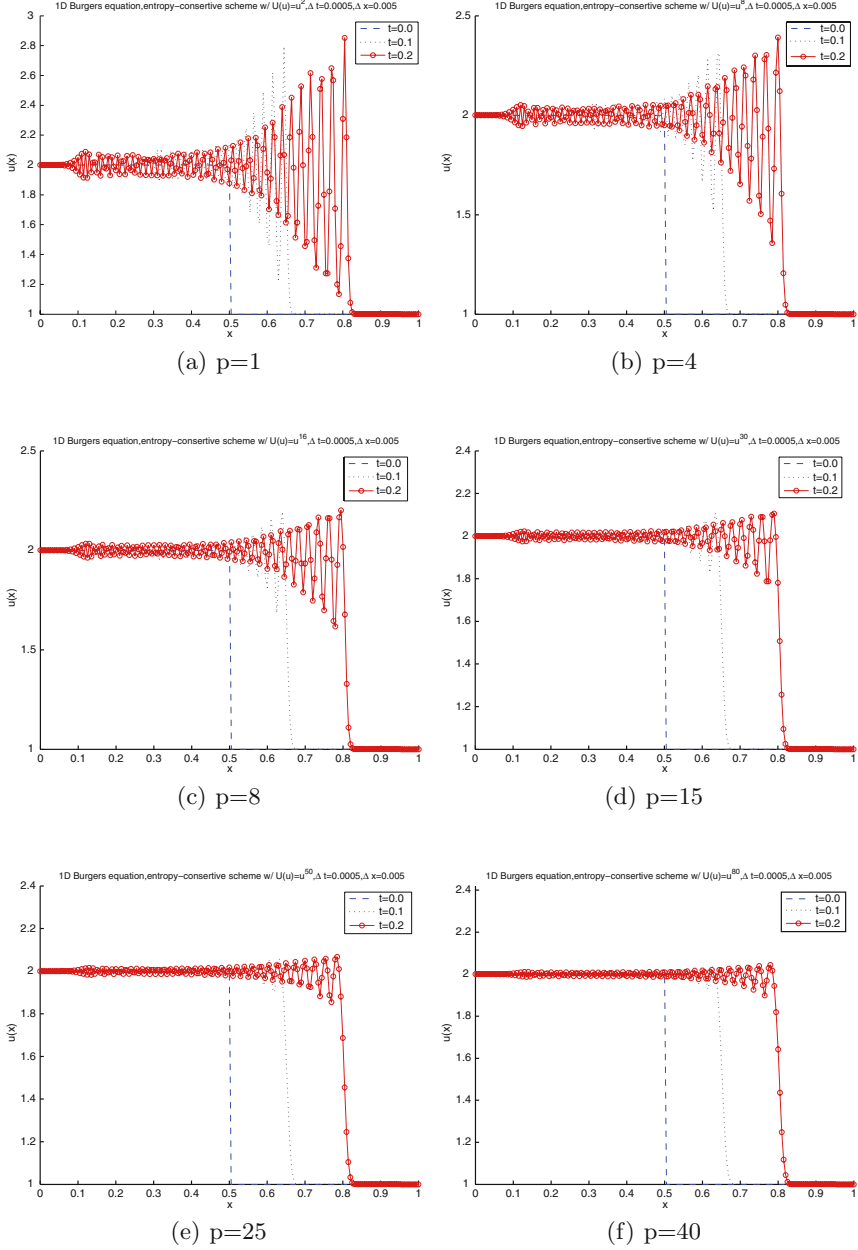# 5 2D Shallow Water Equations

5.1 Energy Stable Schemes

We turn to the construction of entropy/energy-stable schemes for the 2D shallow water equations,

$$\frac{\partial}{\partial t}\mathbf{u} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{u}) + \frac{\partial}{\partial y}\mathbf{g}(\mathbf{u}) = \zeta \frac{\partial}{\partial x}\left(h\frac{\partial}{\partial x}\mathbf{d}(\mathbf{u})\right) + \zeta \frac{\partial}{\partial y}\left(h\frac{\partial}{\partial y}\mathbf{d}(\mathbf{u})\right), \quad \mathbf{u} = \begin{bmatrix} h \\ uh \\ vh \end{bmatrix},$$

(5.1)

with convective fluxes $\mathbf{f} = [uh,\, u^2 h + gh^2/2,\, uvh]^\top$, $\mathbf{g} = [vh,\, uvh,\, v^2 h + gh^2/2]^\top$, and additional diffusive terms $\mathbf{d} = [0,\, u,\, v]^\top$.

The second-order semi-discrete entropy conservative schemes (3.7), (3.8) can be extended to two dimensional shallow water equations (5.1) in a straightforward manner. Recall that E denotes the total energy which is serving as an admissible entropy function with the corresponding entropy fluxes (F, G) associated with the two dimensional shallow water equations, $\mathbf{v} := U_{\mathbf{v}}$ are the corresponding entropy variables (2.14c), and (ψ^x, ψ^y) are the potential pair (2.14e). We discretize the convective fluxes on the LHS using the entropy-conservative differences indicated in 1D setup dimension by dimension. For the dissipative terms on the RHS, we employ the centered differences, while the intermediate h-values are taken to be the arithmetic mean of two neighboring grid-points, $\widehat{h}_{\nu + \frac{1}{2}, \mu} := (h_{\nu + 1, \mu} + h_{\nu, \mu})/2$. We then obtain the entropy stable semi-discrete schemes

$$\frac{d}{dt}\mathbf{u}_{\nu,\mu}(t) + \frac{1}{\Delta x}(\mathbf{f}^*_{\nu + \frac{1}{2}, \mu} - \mathbf{f}^*_{\nu - \frac{1}{2}, \mu}) + \frac{1}{\Delta y}(\mathbf{g}^*_{\nu, \mu + \frac{1}{2}} - \mathbf{g}^*_{\nu, \mu - \frac{1}{2}})$$

$$= \frac{\zeta}{\Delta x}\left(\widehat{h}_{\nu + \frac{1}{2}, \mu}\frac{\mathbf{d}_{\nu + 1, \mu} - \mathbf{d}_{\nu, \mu}}{\Delta x} - \widehat{h}_{\nu - \frac{1}{2}, \mu}\frac{\mathbf{d}_{\nu, \mu} - \mathbf{d}_{\nu - 1, \mu}}{\Delta x}\right)$$

$$+ \frac{\zeta}{\Delta y}\left(\widehat{h}_{\nu, \mu + \frac{1}{2}}\frac{\mathbf{d}_{\nu, \mu + 1} - \mathbf{d}_{\nu, \mu}}{\Delta x} - \widehat{h}_{\nu, \mu - \frac{1}{2}}\frac{\mathbf{d}_{\nu, \mu} - \mathbf{d}_{\nu, \mu - 1}}{\Delta x}\right), \quad (5.2a)$$

(a) p=1

(b) p=4

(c) p=8

(d) p=15

(e) p=25

(f) p=40

Fig. 2. 1D Burger's equation, shock initial condition, entropy-conservative schemes, 200 spatial grids, $U(u) = u^{2p}$

with the entropy-conservative fluxes $\mathbf{f}^*_{\nu+\frac{1}{2},\mu}$ and $\mathbf{g}^*_{\nu,\mu+\frac{1}{2}}$ outlined in (3.12) along x and y direction, respectively,

$$
\begin{aligned}
\mathbf{f}^*_{\nu+\frac{1}{2},\mu} &= \sum_{j=1}^{3} \frac{\psi^x\left(\mathbf{v}^{j+1}_{\nu+\frac{1}{2},\mu}\right) - \psi^x\left(\mathbf{v}^{j}_{\nu+\frac{1}{2},\mu}\right)}{\left\langle \boldsymbol{\ell}_{x_j}, \Delta\mathbf{v}_{\nu+\frac{1}{2},\mu}\right\rangle} \boldsymbol{\ell}_{x_j} \\
&= \frac{g}{2}\sum_{j=1}^{3} \frac{(h^{j+1}_{\nu+\frac{1}{2},\mu})^2 u^{j+1}_{\nu+\frac{1}{2},\mu} - (h^{j}_{\nu+\frac{1}{2},\mu})^2 u^{j}_{\nu+\frac{1}{2},\mu}}{\left\langle \boldsymbol{\ell}_{x_j}, \Delta\mathbf{v}_{\nu+\frac{1}{2},\mu}\right\rangle} \boldsymbol{\ell}_{x_j}, \qquad (5.2b)
\end{aligned}
$$

$$
\begin{aligned}
\mathbf{g}^*_{\nu,\mu+\frac{1}{2}} &= \sum_{j=1}^{3} \frac{\psi^y\left(\mathbf{v}^{j+1}_{\nu,\mu+\frac{1}{2}}\right) - \psi^y\left(\mathbf{v}^{j}_{\nu,\mu+\frac{1}{2}}\right)}{\left\langle \boldsymbol{\ell}_{y_j}, \Delta\mathbf{v}_{\nu,\mu+\frac{1}{2}}\right\rangle} \boldsymbol{\ell}_{y_j} \\
&= \frac{g}{2}\sum_{j=1}^{3} \frac{(h^{j+1}_{\nu,\mu+\frac{1}{2}})^2 v^{j+1}_{\nu,\mu+\frac{1}{2}} - (h^{j}_{\nu,\mu+\frac{1}{2}})^2 v^{j}_{\nu,\mu+\frac{1}{2}}}{\left\langle \boldsymbol{\ell}_{y_j}, \Delta\mathbf{v}_{\nu,\mu+\frac{1}{2}}\right\rangle} \boldsymbol{\ell}_{y_j}, \qquad (5.2c)
\end{aligned}
$$

Here, $\mathbf{u}_{\nu,\mu}(t)$ denotes the discrete solution at the grid point (x_ν, y_μ, t) with $x_\nu := \nu\Delta x$, $y_\mu := \mu\Delta y$, Δx and Δy being the uniform mesh sizes, and $\mathbf{d}_{\nu,\mu} := \mathbf{d}(\mathbf{u}_{\nu,\mu})$. The numerical flux $\mathbf{f}^*_{\nu+\frac{1}{2},\mu}$ and $\mathbf{g}^*_{\nu,\mu+\frac{1}{2}}$ are constructed separately along two different phase paths dictated by two sets of vectors $\{\boldsymbol{\ell}_{x_j}\}$ and $\{\boldsymbol{\ell}_{y_j}\}$. Finally, $\{u^j\}$, $\{v^j\}$, and $\{h^j\}$ are intermediate values of height and velocities along paths in the phase space. The physical relevance of the intermediate solutions along the paths needs to be maintained. To this end, we choose to work along the paths which are determined by (approximate) Riemann solvers. Specifically, we use the eigensystems of the Roe matrix in the x and y directions, [Roe81, Gla87],

$$
\widetilde{A}^x = \begin{bmatrix} 0 & 1 & 0 \\ \bar{c}^2_{\nu+\frac{1}{2},\mu} - \bar{u}^2_{\nu+\frac{1}{2},\mu} & 2\bar{u}_{\nu+\frac{1}{2},\mu} & 0 \\ -\bar{u}_{\nu+\frac{1}{2},\mu}\bar{v}_{\nu+\frac{1}{2},\mu} & \bar{v}_{\nu+\frac{1}{2},\mu} & \bar{u}_{\nu+\frac{1}{2},\mu} \end{bmatrix},
$$

$$
\widetilde{A}^y = \begin{bmatrix} 0 & 0 & 1 \\ -\bar{u}_{\nu,\mu+\frac{1}{2}}\bar{v}_{\nu,\mu+\frac{1}{2}} & \bar{v}_{\nu,\mu+\frac{1}{2}} & \bar{u}_{\nu,\mu+\frac{1}{2}} \\ \bar{c}^2_{\nu,\mu+\frac{1}{2}} - \bar{v}^2_{\nu,\mu+\frac{1}{2}} & 0 & 2\bar{v}_{\nu,\mu+\frac{1}{2}} \end{bmatrix}. \qquad (5.3a)
$$

Here $\bar{u}$, $\bar{v}$, and $\bar{c}$ are the average values of the velocities u, v and the sound speed $c := \sqrt{gh}$ at Roe-average state,

$$
\bar{u} = \frac{u_R\sqrt{h_R} + u_L\sqrt{h_L}}{\sqrt{h_R} + \sqrt{h_L}}, \quad \bar{v} = \frac{v_R\sqrt{h_R} + v_L\sqrt{h_L}}{\sqrt{h_R} + \sqrt{h_L}}, \quad \bar{c} = \sqrt{\frac{g(h_R + h_L)}{2}},
$$
$$\qquad (5.3b)$$

where the subscripts $(\cdot)_R$ and $(\cdot)_L$ represent two neighboring spatial grid-points. The vector sets $\{\widehat{\mathbf{r}}_{x_j}\}_{j=1}^{3}$ and $\{\widehat{\mathbf{r}}_{y_j}\}_{j=1}^{3}$ are chosen to be the right eigen-vectors of the x- and y-Roe matrices (5.3a) (omitting the sub/superscripts of all averaged variables)

$$\widehat{\mathbf{r}}_{x_1} = \begin{bmatrix} 1 \\ \bar{u} - \bar{c} \\ \bar{v} \end{bmatrix}, \ \widehat{\mathbf{r}}_{x_2} = \begin{bmatrix} 0 \\ 0 \\ \bar{c} \end{bmatrix}, \ \widehat{\mathbf{r}}_{x_3} = \begin{bmatrix} 1 \\ \bar{u} + \bar{c} \\ \bar{v} \end{bmatrix}, \qquad (5.3c)$$

$$\widehat{\mathbf{r}}_{y_1} = \begin{bmatrix} 1 \\ \bar{u} \\ \bar{v} - \bar{c} \end{bmatrix}, \ \widehat{\mathbf{r}}_{y_2} = \begin{bmatrix} 0 \\ -\bar{c} \\ 0 \end{bmatrix}, \ \widehat{\mathbf{r}}_{y_3} = \begin{bmatrix} 1 \\ \bar{u} \\ \bar{v} + \bar{c} \end{bmatrix}, \qquad (5.3d)$$

with the corresponding left eigenvector sets $\{\widehat{\boldsymbol{\ell}}_{x_j}\}_{j=1}^3$ and $\{\widehat{\boldsymbol{\ell}}_{y_j}\}_{j=1}^3$ given by

$$\widehat{\boldsymbol{\ell}}_{x_1} = \begin{bmatrix} \dfrac{\bar{u} + \bar{c}}{2\bar{c}} \\ -\dfrac{1}{2\bar{c}} \\ 0 \end{bmatrix}, \ \widehat{\boldsymbol{\ell}}_{x_2} = \begin{bmatrix} -\dfrac{\bar{v}}{\bar{c}} \\ 0 \\ \dfrac{1}{\bar{c}} \end{bmatrix}, \ \widehat{\boldsymbol{\ell}}_{x_3} = \begin{bmatrix} \dfrac{-\bar{u} + \bar{c}}{2\bar{c}} \\ \dfrac{1}{2\bar{c}} \\ 0 \end{bmatrix}, \qquad (5.3e)$$

$$\widehat{\boldsymbol{\ell}}_{y_1} = \begin{bmatrix} \dfrac{\bar{v} + \bar{c}}{2\bar{c}} \\ 0 \\ -\dfrac{1}{2\bar{c}} \end{bmatrix}, \ \widehat{\boldsymbol{\ell}}_{y_2} = \begin{bmatrix} \dfrac{\bar{u}}{\bar{c}} \\ -\dfrac{1}{\bar{c}} \\ 0 \end{bmatrix}, \ \widehat{\boldsymbol{\ell}}_{y_3} = \begin{bmatrix} \dfrac{-\bar{v} + \bar{c}}{2\bar{c}} \\ 0 \\ \dfrac{1}{2\bar{c}} \end{bmatrix}. \qquad (5.3f)$$

We now are able to form the intermediate paths along x and y directions in $\mathbf{u}$-space as in (3.6): starting with $\mathbf{u}^1_{\nu + \frac{1}{2}, \mu} = \mathbf{u}^1_{\nu, \mu + \frac{1}{2}} = \mathbf{u}_{\nu, \mu}$, we proceed with

$$\mathbf{u}^{j+1}_{\nu + \frac{1}{2}, \mu} = \mathbf{u}^j_{\nu + \frac{1}{2}, \mu} + \left\langle \widehat{\boldsymbol{\ell}}_{x_j}, \Delta\mathbf{u}_{\nu + \frac{1}{2}, \mu} \right\rangle \widehat{\mathbf{r}}_{x_j}, \ j = 1, 2, 3, \ \Delta\mathbf{u}_{\nu + \frac{1}{2}, \mu} := \mathbf{u}_{\nu + 1, \mu} - \mathbf{u}_{\nu, \mu},$$

$$\mathbf{u}^{j+1}_{\nu, \mu + \frac{1}{2}} = \mathbf{u}^j_{\nu, \mu + \frac{1}{2}} + \left\langle \widehat{\boldsymbol{\ell}}_{y_j}, \Delta\mathbf{u}_{\nu, \mu + \frac{1}{2}} \right\rangle \widehat{\mathbf{r}}_{y_j}, \ j = 1, 2, 3, \ \Delta\mathbf{u}_{\nu, \mu + \frac{1}{2}} := \mathbf{u}_{\nu, \mu + 1} - \mathbf{u}_{\nu, \mu}.$$

The construction of the entropy-conservative numerical fluxes $\mathbf{f}^*_{\nu + \frac{1}{2}, \mu}$ and $\mathbf{g}^*_{\nu, \mu + \frac{1}{2}}$ follows the algorithm indicated in Algorithm 1.

Remark 5.1. We point out that in the case $\left\langle \widehat{\boldsymbol{\ell}}_j, \Delta\mathbf{u} \right\rangle = 0$ for certain j's in $\mathbf{u}$-space, which may cause $\langle \boldsymbol{\ell}_j, \Delta\mathbf{v} \rangle = 0$ in $\mathbf{v}$-space, hence fail Algorithm 1. Arguing along the same line as [TZ06, Remark 3.5], we compute the corresponding entropy-conservative numerical fluxes using the alternate formulas,

$$\mathbf{f}^*_{\nu + \frac{1}{2}, \mu} = \sum_{\{j | \xi_{x_j} \neq 0\}} \frac{\psi^x(\mathbf{v}^j_{\nu + \frac{1}{2}, \mu} + \xi_{x_j}\mathbf{r}_{x_j}) - \psi^x(\mathbf{v}^j_{\nu + \frac{1}{2}, \mu})}{\xi_{x_j}} \boldsymbol{\ell}_{x_j}, \ \xi_{x_j} := \left\langle \boldsymbol{\ell}_{x_j}, \Delta\mathbf{v}_{\nu + \frac{1}{2}, \mu} \right\rangle,$$

$$\mathbf{g}^*_{\nu, \mu + \frac{1}{2}} = \sum_{\{j | \xi_{y_j} \neq 0\}} \frac{\psi^y(\mathbf{v}^j_{\nu, \mu + \frac{1}{2}} + \xi_{y_j}\mathbf{r}_{y_j}) - \psi^y(\mathbf{v}^j_{\nu, \mu + \frac{1}{2}})}{\xi_{y_j}} \boldsymbol{\ell}_{y_j}, \ \xi_{y_j} := \left\langle \boldsymbol{\ell}_{y_j}, \Delta\mathbf{v}_{\nu, \mu + \frac{1}{2}} \right\rangle,$$

where the right and left eigensystems $\{\mathbf{r}_{x_j}\}_{j=1}^3 \{\mathbf{r}_{y_j}\}_{j=1}^3$ and $\{\boldsymbol{\ell}_{x_j}\}_{j=1}^3 \{\boldsymbol{\ell}_{y_j}\}_{j=1}^3$ are constructed as the precise mirror images of the Roe-paths in $\mathbf{v}$-space,

$$\mathbf{r}_j^x := [H]_{\nu+\frac{1}{2},\,\mu}^{-1} \widehat{\mathbf{r}}_{x_j}, \qquad \boldsymbol{\ell}_j^x := [H]_{\nu+\frac{1}{2},\,\mu} \widehat{\boldsymbol{\ell}}_{x_j}, \qquad j = 1, 2, 3$$

$$\mathbf{r}_j^y := [H]_{\nu,\,\mu+\frac{1}{2}}^{-1} \widehat{\mathbf{r}}_{y_j}, \qquad \boldsymbol{\ell}_j^y := [H]_{\nu,\,\mu+\frac{1}{2}} \widehat{\boldsymbol{\ell}}_{y_j}, \qquad j = 1, 2, 3$$

where $[H]_{\nu+\frac{1}{2},\,\mu}$ and $[H]_{\nu,\,\mu+\frac{1}{2}}$ denote the averaged symmetrizers such that $\Delta\mathbf{u}_{\nu+\frac{1}{2},\,\mu} = [H]_{\nu+\frac{1}{2},\,\mu}\Delta\mathbf{v}_{\nu+\frac{1}{2},\,\mu}$ and $\Delta\mathbf{u}_{\nu,\,\mu+\frac{1}{2}} = [H]_{\nu,\,\mu+\frac{1}{2}}\Delta\mathbf{v}_{\nu,\,\mu+\frac{1}{2}}$.

We summarize our main result on 2D shallow water equations in the following theorem.

Theorem 5.1. *Let $E = (gh^2 + u^2h + v^2h)/2$ be the total energy of the 2D shallow water equations (5.1). Then, the semi-discrete approximation (5.2a) with entropy conservative fluxes $\mathbf{f}_{\nu+\frac{1}{2},\,\mu}^*$ and $\mathbf{g}_{\nu,\,\mu+\frac{1}{2}}^*$ given in (5.2b), (5.2c), (5.3), is energy stable, and the following discrete energy balance is satisfied,*

$$\frac{d}{dt}\sum_{\nu,\,\mu} E(\mathbf{u}_{\nu,\,\mu}(t))\Delta x \Delta y = -\zeta \sum_{\nu,\,\mu} \left\{ \widehat{h}_{\nu+\frac{1}{2},\,\mu} \left[\left(\frac{\Delta u_{\nu+\frac{1}{2},\,\mu}}{\Delta x}\right)^2 + \left(\frac{\Delta v_{\nu+\frac{1}{2},\,\mu}}{\Delta x}\right)^2 \right] \right.$$

$$\left. + \widehat{h}_{\nu,\,\mu+\frac{1}{2}} \left[\left(\frac{\Delta u_{\nu,\,\mu+\frac{1}{2}}}{\Delta y}\right)^2 + \left(\frac{\Delta v_{\nu,\,\mu+\frac{1}{2}}}{\Delta y}\right)^2 \right] \right\} \Delta x \Delta y. \quad (5.4)$$

Observe that no artificial viscosity is introduced in the sense that the energy dissipation statement (5.4) is the precise discrete analogue of the energy balance statement (2.2).

Proof. Multiply (5.2a) by $[U_{\mathbf{u}}]_{\nu,\,\mu}^\top = \mathbf{v}_{\nu,\,\mu}^\top$, and sum up all spatial cells to get the balance of the total entropy,

$$\frac{d}{dt}\sum_{\nu,\,\mu} E(\mathbf{u}_{\nu,\,\mu}(t))\Delta x \Delta y + \sum_{\nu,\,\mu}\left\langle \mathbf{v}_{\nu,\,\mu}, \mathbf{f}_{\nu+\frac{1}{2},\,\mu}^* - \mathbf{f}_{\nu-\frac{1}{2},\,\mu}^* \right\rangle \Delta y$$

$$+ \sum_{\nu,\,\mu}\left\langle \mathbf{v}_{\nu,\,\mu}, \mathbf{g}_{\nu,\,\mu+\frac{1}{2}}^* - \mathbf{g}_{\nu,\,\mu-\frac{1}{2}}^* \right\rangle \Delta x$$

$$= \zeta \sum_{\nu,\,\mu}\left\langle \mathbf{v}_{\nu,\,\mu}, \widehat{h}_{\nu+\frac{1}{2},\,\mu}\Delta\mathbf{d}_{\nu+\frac{1}{2},\,\mu} - \widehat{h}_{\nu-\frac{1}{2},\,\mu}\Delta\mathbf{d}_{\nu-\frac{1}{2},\,\mu} \right\rangle \frac{\Delta y}{\Delta x}$$

$$+ \zeta \sum_{\nu,\,\mu}\left\langle \mathbf{v}_{\nu,\,\mu}, \widehat{h}_{\nu,\,\mu+\frac{1}{2}}\Delta\mathbf{d}_{\nu,\,\mu+\frac{1}{2}} - \widehat{h}_{\nu,\,\mu-\frac{1}{2}}\Delta\mathbf{d}_{\nu,\,\mu-\frac{1}{2}} \right\rangle \frac{\Delta x}{\Delta y} \quad (5.5)$$

Since the numerical fluxes $\mathbf{f}_{\nu+\frac{1}{2},\,\mu}^*$ and $\mathbf{g}_{\nu,\,\mu+\frac{1}{2}}^*$ are chosen as the entropy conservative fluxes in x and y directions respectively, they satisfy the entropy conservative requirement (3.10a), so that their $\mathbf{v}$-moments on the left of (5.5) amount to perfect differences,

$$\left\langle \mathbf{v}_{\nu,\,\mu}, \mathbf{f}_{\nu+\frac{1}{2},\,\mu}^* - \mathbf{f}_{\nu-\frac{1}{2},\,\mu}^* \right\rangle = F_{\nu+\frac{1}{2},\,\mu} - F_{\nu-\frac{1}{2},\,\mu}, \qquad (5.6a)$$

$$\left\langle \mathbf{v}_{\nu,\,\mu}, \mathbf{g}_{\nu,\,\mu+\frac{1}{2}}^* - \mathbf{g}_{\nu,\,\mu-\frac{1}{2}}^* \right\rangle = G_{\nu,\,\mu+\frac{1}{2}} - G_{\nu,\,\mu-\frac{1}{2}}, \qquad (5.6b)$$

with consistent entropy fluxes given by (consult (3.10b)),

$$2F_{\nu+\frac{1}{2},\,\mu} = \left\langle (\mathbf{v}_{\nu,\,\mu} + \mathbf{v}_{\nu+1,\,\mu}),\, \mathbf{f}^*_{\nu+\frac{1}{2},\,\mu} \right\rangle - (\psi^x(\mathbf{v}_{\nu,\,\mu}) + \psi^x(\mathbf{v}_{\nu+1,\,\mu}))$$

$$2G_{\nu,\,\mu+\frac{1}{2}} = \left\langle (\mathbf{v}_{\nu,\,\mu} + \mathbf{v}_{\nu,\,\mu+1}),\, \mathbf{g}^*_{\nu,\,\mu+\frac{1}{2}} \right\rangle - (\psi^y(\mathbf{v}_{\nu,\,\mu}) + \psi^y(\mathbf{v}_{\nu,\,\mu+1})).$$

On the other hand, summation by parts and explicit computation using the entropy variable (2.14c) on the RHS of (5.5) yield

$$\zeta \sum_{\nu,\,\mu} \left\langle \mathbf{v}_{\nu,\,\mu},\, \widehat{h}_{\nu+\frac{1}{2},\,\mu} \Delta \mathbf{d}_{\nu+\frac{1}{2},\,\mu} - \widehat{h}_{\nu-\frac{1}{2},\,\mu} \Delta \mathbf{d}_{\nu-\frac{1}{2},\,\mu} \right\rangle \frac{\Delta y}{\Delta x}$$

$$= -\zeta \sum_{\nu,\,\mu} \left\langle \Delta \mathbf{v}_{\nu+\frac{1}{2},\,\mu},\, \widehat{h}_{\nu+\frac{1}{2},\,\mu} \Delta \mathbf{d}_{\nu+\frac{1}{2},\,\mu} \right\rangle \frac{\Delta y}{\Delta x}$$

$$= -\zeta \sum_{\nu,\,\mu} \left[\frac{1}{(\Delta x)^2} \widehat{h}_{\nu+\frac{1}{2},\,\mu} \left(\left(\Delta u_{\nu+\frac{1}{2},\,\mu}\right)^2 + \left(\Delta v_{\nu+\frac{1}{2},\,\mu}\right)^2 \right) \right] \Delta x \Delta y$$

$$\text{(5.7a)}$$

$$\zeta \sum_{\nu,\,\mu} \left\langle \mathbf{v}_{\nu,\,\mu},\, \widehat{h}_{\nu,\,\mu+\frac{1}{2}} \Delta \mathbf{d}_{\nu,\,\mu+\frac{1}{2}} - \widehat{h}_{\nu,\,\mu-\frac{1}{2}} \Delta \mathbf{d}_{\nu,\,\mu-\frac{1}{2}} \right\rangle \frac{\Delta x}{\Delta y}$$

$$= -\zeta \sum_{\nu,\,\mu} \left\langle \Delta \mathbf{v}_{\nu,\,\mu+\frac{1}{2}},\, \widehat{h}_{\nu,\,\mu+\frac{1}{2}} \Delta \mathbf{d}_{\nu,\,\mu+\frac{1}{2}} \right\rangle \frac{\Delta x}{\Delta y}$$

$$= -\zeta \sum_{\nu,\,\mu} \left[\frac{1}{(\Delta y)^2} \widehat{h}_{\nu,\,\mu+\frac{1}{2}} \left(\left(\Delta u_{\nu,\,\mu+\frac{1}{2}}\right)^2 + \left(\Delta v_{\nu,\,\mu+\frac{1}{2}}\right)^2 \right) \right] \Delta x \Delta y$$

$$\text{(5.7b)}$$

By (5.6) and (5.7), the semi-discrete energy balance statement (5.4) now follows,

$$\frac{d}{dt} \sum_{\nu,\,\mu} E(\mathbf{u}_{\nu,\,\mu}(t)) \Delta x \Delta y = -\zeta \sum_{\nu,\,\mu} \left\{ \widehat{h}_{\nu+\frac{1}{2},\,\mu} \left[\left(\frac{\Delta u_{\nu+\frac{1}{2},\,\mu}}{\Delta x}\right)^2 + \left(\frac{\Delta v_{\nu+\frac{1}{2},\,\mu}}{\Delta x}\right)^2 \right] \right.$$

$$\left. + \widehat{h}_{\nu,\,\mu+\frac{1}{2}} \left[\left(\frac{\Delta u_{\nu,\,\mu+\frac{1}{2}}}{\Delta y}\right)^2 + \left(\frac{\Delta v_{\nu,\,\mu+\frac{1}{2}}}{\Delta y}\right)^2 \right] \right\} \Delta x \Delta y. \quad \square$$

5.2 Energy Preserving Schemes

In the case that the eddy viscosity is absent, $\zeta = 0$, all the dissipation terms on the RHS of the difference scheme (5.2a) vanish,

$$\frac{d}{dt} \mathbf{u}_{\nu,\,\mu}(t) + \frac{1}{\Delta x} (\mathbf{f}^*_{\nu+\frac{1}{2},\,\mu} - \mathbf{f}^*_{\nu-\frac{1}{2},\,\mu}) + \frac{1}{\Delta y} (\mathbf{g}^*_{\nu,\,\mu+\frac{1}{2}} - \mathbf{g}^*_{\nu,\,\mu-\frac{1}{2}}) = 0. \qquad \text{(5.8)}$$

The resulting scheme serves as an *energy preserving* approximation to the inviscid shallow water equations (1.2) with the discrete energy *equality*,

$$\frac{d}{dt} \sum_{\nu, \mu} E(\mathbf{u}_{\nu, \mu}(t)) \Delta x \Delta y = 0.$$

Remark 5.2. We note that energy preserving semi-discrete scheme (5.2), (5.3) may allow a substantial increase of the potential *enstrophy*, $\frac{1}{2} \sum \eta_{\nu, \mu}^2 / h_{\nu, \mu}$, especially for the flow over steep topography, due to spurious energy cascade into smaller scales, consult [AL77, AL81]. Here, η is the sum of the relative vorticity $v_x - u_y$ and the Coriolis parameter at that latitude. After a long term integration, a significant amount of energy is transferred into the smallest resolvable scales, where truncation error becomes relevant. It would be desirable to adapt our energy stable discretization to retain the additional conservation of enstrophy, advocated in [Ara97, AL81].

6 Numerical Experiments for 2D Shallow Water Equations

6.1 Boundary Conditions

The numerical treatment of boundaries is intended to be as physically relevant as possible. We describe two basic types of boundary conditions that are applicable to the two dimensional shallow water problems: the first type simulates a boundary at infinity or a transmissive boundary; the second type applies in the presence of solid fixed walls.

Transmissive Boundaries

These are cases in which boundaries are supposed to be transparent in the sense that waves are allowed to pass through. The inflow and outflow conditions need to be described, hence the method of characteristics in two dimension follows. The local value of the Froude number $Fr := V/\sqrt{gL}$ determines the flow regime and, accordingly, the number of boundary conditions to apply. Here V and L denote the characteristic velocity and length scales of the phenomenon, respectively. For subcritical flow, two external boundary conditions are required at inflow boundaries, whereas only one boundary condition is required at outflow boundaries. Two dimensional supercritical flow requires three inflow boundary conditions and no boundary condition at outflow boundaries where the flow is only influenced by the information coming from the interior nodes.

Reflective Boundaries

This is a particular case in which the flow is confined inside a fixed field by solid walls where we impose the reflective boundary conditions. Since our
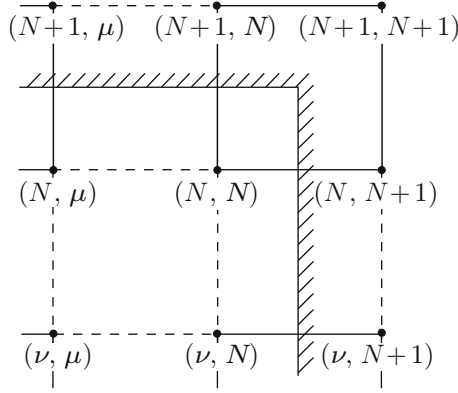
Fig. 3. Right-hand boundary

testing problems in next section are concerned with the flow in a square basin, without losing generality, we consider the computational domain in the upper-right corner with the solid boundaries along x and y-direction as shown in Fig. 3. By the three-point stencil used in our semi-discrete scheme, we try to impose the value of one computational grid point added outside boundary.

The reflection is incorporated by changing the sign of the normal component of the velocity, while the water depth is unaltered. The values at all the $(\nu, N + 1)$ points on the right-hand side of the wall are replaced by the values at interior (ν, N) points and sign of the normal velocity component u is switched,

$$h_{\nu, N+1} = h_{\nu, N}, \quad u_{\nu, N+1} = -u_{\nu, N}, \quad v_{\nu, N+1} = v_{\nu, N};$$

the values at all the $(N + 1, \mu)$ points on the top of the wall are replaced by the values at interior (N, μ) points and sign of the normal velocity component v is switched

$$h_{N+1, \mu} = h_{N, \mu}, \quad u_{N+1, \mu} = u_{N, \mu}, \quad v_{N+1, \mu} = -v_{N, \mu};$$

the values at all the $(N + 1, N + 1)$ point in the upper-right corner are given by

$$h_{\nu+1, N+1} = h_{\nu, \mu}, \quad u_{\nu+1, N+1} = -u_{\nu, \mu}, \quad v_{\nu+1, N+1} = -v_{\nu, \mu}.$$

6.2 Time Discretization

Similar to the time discretizations of the Burgers' equation, we integrate the entropy stable scheme (5.2) and (5.3) with the explicit three-stage Runge–Kutta method (4.12a) by its high-order accuracy, large stability region and simplicity.

$$\begin{cases} \mathbf{u}^{(1)} &= \mathbf{u}^n + \Delta t \mathcal{L}(\mathbf{u}^n) \\ \mathbf{u}^{(2)} &= \dfrac{3}{4}\mathbf{u}^n + \dfrac{1}{4}\mathbf{u}^{(1)} + \dfrac{1}{4}\Delta t \mathcal{L}(\mathbf{u}^{(1)}) \\ \mathbf{u}^{n+1} &= \dfrac{1}{3}\mathbf{u}^n + \dfrac{2}{3}\mathbf{u}^{(2)} + \dfrac{2}{3}\Delta t \mathcal{L}(\mathbf{u}^{(2)}) \end{cases} \tag{6.1a}$$

where

$$[\mathcal{L}(\mathbf{u})]_{\nu,\mu} = -\frac{1}{\Delta x}(\mathbf{f}_{\nu+\frac{1}{2},\mu} - \mathbf{f}_{\nu-\frac{1}{2},\mu}) - \frac{1}{\Delta y}(\mathbf{g}_{\nu,\mu+\frac{1}{2}} - \mathbf{g}_{\nu,\mu-\frac{1}{2}})$$

$$+ \frac{\zeta}{\Delta x}(\widehat{h}_{\nu+\frac{1}{2},\mu}\frac{\mathbf{d}_{\nu+1,\mu} - \mathbf{d}_{\nu,\mu}}{\Delta x} - \widehat{h}_{\nu-\frac{1}{2},\mu}\frac{\mathbf{d}_{\nu,\mu} - \mathbf{d}_{\nu-1,\mu}}{\Delta x})$$

$$+ \frac{\zeta}{\Delta y}(\widehat{h}_{\nu,\mu+\frac{1}{2}}\frac{\mathbf{d}_{\nu,\mu+1} - \mathbf{d}_{\nu,\mu}}{\Delta x} - \widehat{h}_{\nu,\mu-\frac{1}{2}}\frac{\mathbf{d}_{\nu,\mu} - \mathbf{d}_{\nu,\mu-1}}{\Delta x}). \tag{6.1b}$$

6.3 Numerical Results

We test our entropy-stable schemes with the two dimensional frictionless partial-dam-break problem originally studied by Fennema and Chaudhry in [FC90]. It imposes computational difficulties due to the discontinuous initial conditions. It also involves other computational issues like boundary treatments and positive-water-depth preserving solver.

As shown in Fig. 4, the simplified geometry of the problem consists of a $1{,}400 \times 1{,}400$ m² basin with a idealized dam in the middle. Water is limited by the fixed, solid, frictionless walls in this square basin. To prevent any damping by the source terms, a frictionless, horizontal bottom is used. All walls are assumed to be reflective. The initial water level of the dam is 10 m and the tail
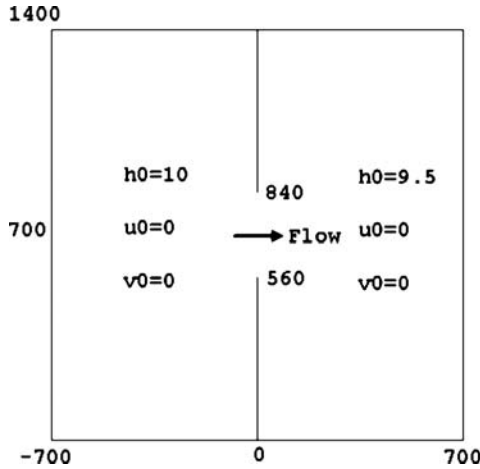
Fig. 4. Geometry configuration and initial setting of 2D Partial-Dam-Break problem

water is 9.5 m high. Central part of the dam is assumed to fail instantaneously or the gate in the middle of the dam is opened instantly. Water is released into the downstream side through a breach 280 m wide, located between $y = 560$ and $y = 840$, forming a wave that propagates while spreading laterally. A negative wave propagates upstream at the same time. For simplicity, the Coriolis force is ignored in the computation. The acceleration due to gravity is taken to be $9.8\,\mathrm{m\,s^{-2}}$. Although there is no analytical reference solution for this test problem, other numerical results of similar problems are available in [FC90, CK04].
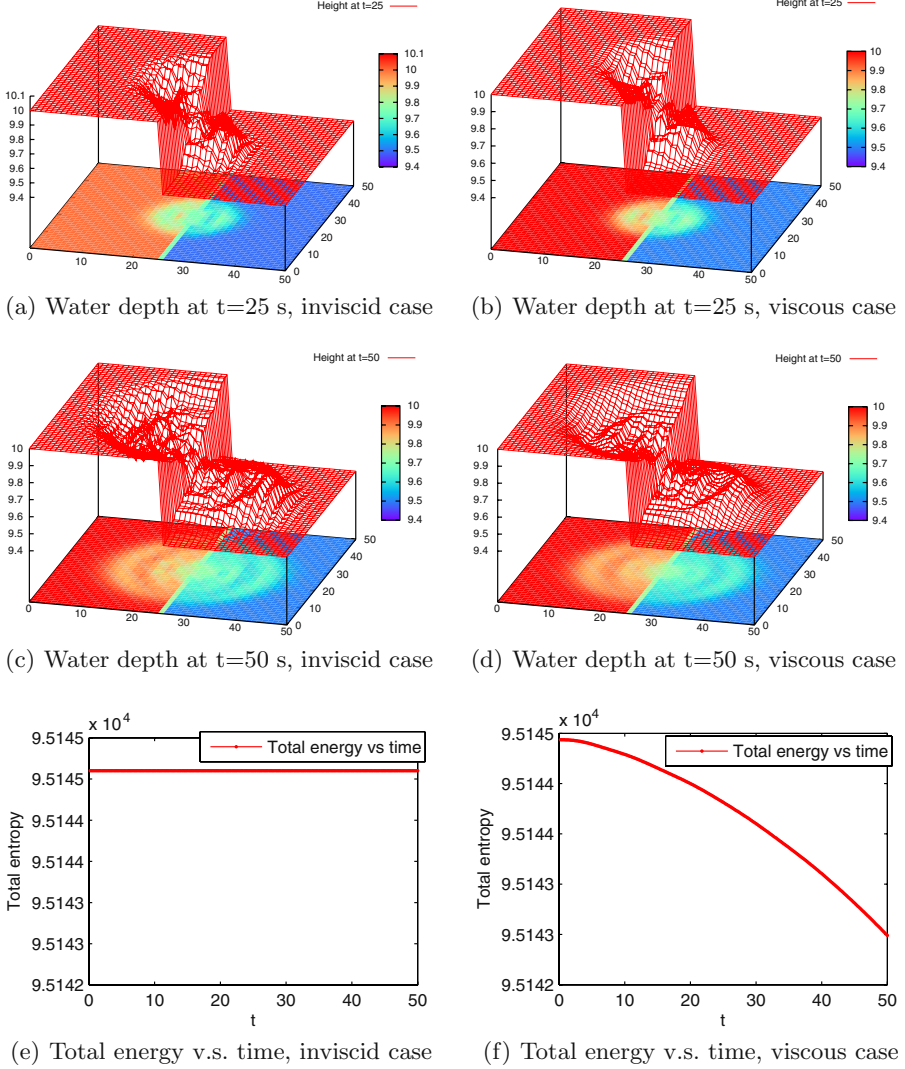
In the following figures, we display the numerical solutions for the fully discrete scheme (6.1a)–(6.1b) with the numerical fluxes (5.2b)–(5.2c). The sum of potential and kinetic energy serves as the generalized entropy function in the design of our numerical schemes,

$$E(\mathbf{u}) = \frac{gh^2 + u^2h + v^2h}{2}.$$

Uniform space and time grid sizes, $\Delta x = \Delta y$ and Δt are used. The computational model is run for up to 50 s after the dam broke when the water waves haven't reached the boundaries. Both inviscid and viscous cases are explored. For the viscous cases, the eddy viscosity is taken to be $10\,\mathrm{m^2s^{-1}}$. We use different spatial resolutions for the same problem, and adjust time step according to the CFL condition.

We first solve the inviscid and viscous shallow water equations on the computational domain consisting of a 50×50 cell square grid with $\Delta x = \Delta y = 28$ m. We group our numerical results of inviscid shallow water equations along the left column of Fig. 5. For comparison, the results of viscous shallow water equations with eddy viscosity $\zeta = 10\,\mathrm{m^2s^{-1}}$ are summarized on the right column. The first and second row of Fig. 5 depict the perspective plots of water surface profiles at $t=25$ s and $t=50$ s respectively. Remnants of the dam are represented by jumps near the middle of the plot. The vertical scale is exaggerated with respect to the horizontal scales. We observe that the numerical solutions of the water depth in Fig. 5a, c successfully simulate both the circular shock water wave propagations and the vortices formed on the both sides of the breach. The undershoots are also developed near sharp corners of the remanent dam. These steep degressions in the water surface are noticeable downstream of the breach at $t = 50$ s. Similar numerical tests were done in [CK04] by the second-order central-upwind schemes, which were originally proposed in [KT00].
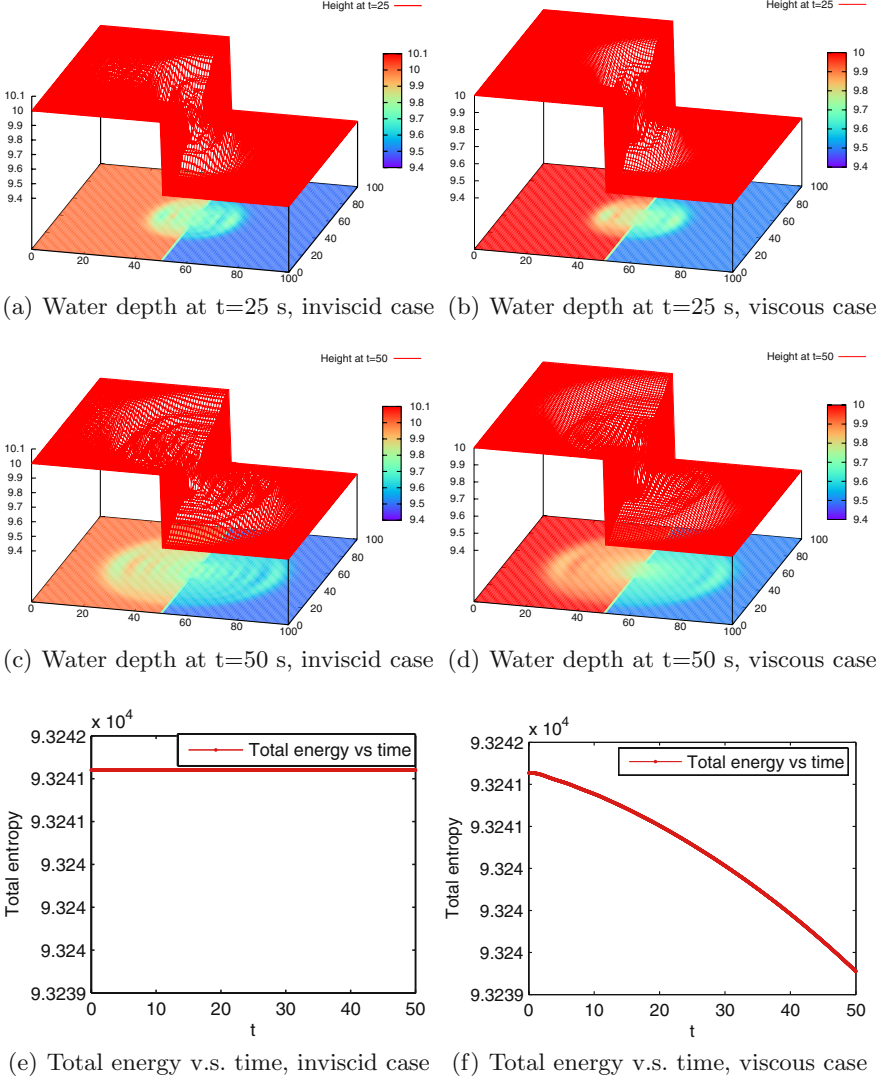
For the inviscid shallow water equations, dispersive errors of the numerical schemes, in the form of spurious oscillations in the mesh scale, are noticeable near the breach in Fig. 5a, c. For the viscous shallow water equations, as shown in Fig. 5b, d, the presence of eddy viscosity causes the oscillations to be dramatically reduced around the breach. In addition to eliminating the wiggles, the eddy viscosity terms also single out the undershoot near sharp corners of the remnants of dam without damping it.

(a) Water depth at t=25 s, inviscid case

(b) Water depth at t=25 s, viscous case

(c) Water depth at t=50 s, inviscid case

(d) Water depth at t=50 s, viscous case

(e) Total energy v.s. time, inviscid case

(f) Total energy v.s. time, viscous case

Fig. 5. Shallow water equations, $\zeta = 10\,\mathrm{m^2 s^{-1}}$, Dam-Break, $1{,}400 \times 1{,}400\,\mathrm{m^2}$ basin, reflective-slip boundary, $\Delta x = \Delta y = 28\,\mathrm{m}$, $\Delta t = 0.2\,\mathrm{s}$
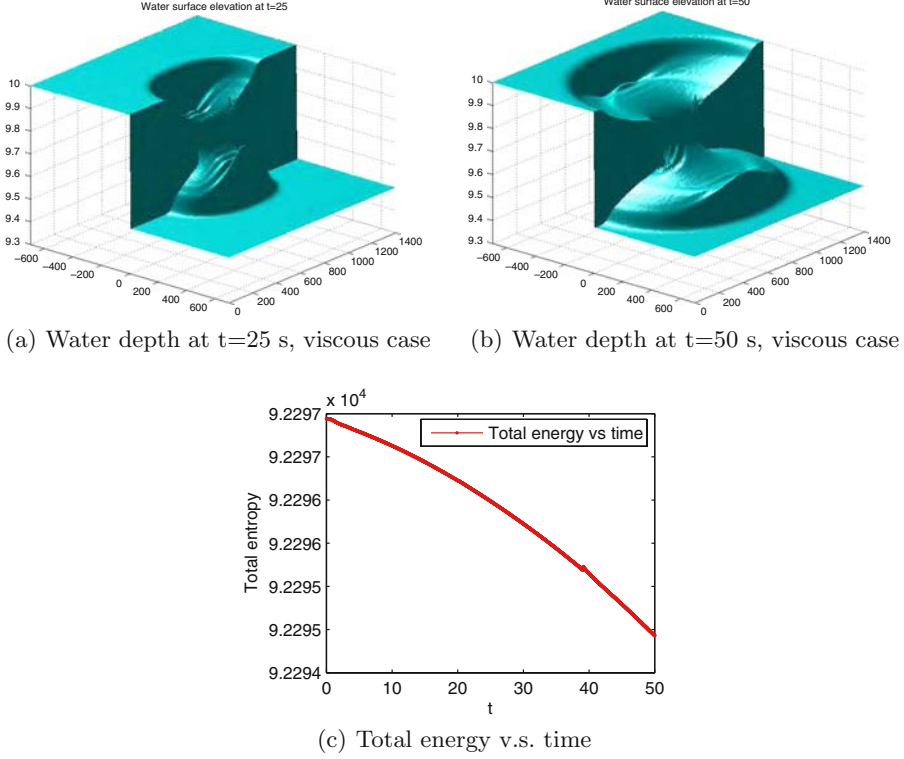
We display the total entropy scaled by 10^4 versus time in Fig. 5f. Compared with the same entropy plot of the inviscid problem in Fig. 5c, the plot of total energy in Fig. 5f reveals a $\mathcal{O}(1)$ energy decay due to the presence of eddy viscosity, while the negligible amount of energy decay introduced by RK3 time discretization for the inviscid shallow water equations is not detectable under the same scale in Fig. 5c.

(a) Water depth at t=25 s, inviscid case   (b) Water depth at t=25 s, viscous case

(c) Water depth at t=50 s, inviscid case   (d) Water depth at t=50 s, viscous case

(e) Total energy v.s. time, inviscid case   (f) Total energy v.s. time, viscous case

Fig. 6. Shallow water equations, $\zeta = 10\,\mathrm{m}^2\mathrm{s}^{-1}$, Dam-Break, $1{,}400 \times 1{,}400\,\mathrm{m}^2$ basin, reflective-slip boundary, $\Delta x = \Delta y = 14\,\mathrm{m}$, $\Delta t = 0.01\,\mathrm{s}$

Next, in Fig. 6, we display the numerical solutions of the same problem in the refined spatial mesh with $\Delta x = \Delta y = 14\,\mathrm{m}$. Following the same pattern as in Fig. 5, Fig. 6 presents the perspective plots and total energy versus time. For the inviscid case, the profiles of the water elevation in Fig. 6a, c demonstrate smoother numerical solutions due to the decrease of the grid size, while the spurious oscillations in the mesh scale are still detectable near the breach

Water surface elevation at t=25

Water surface elevation at t=50

(a) Water depth at t=25 s, viscous case   (b) Water depth at t=50 s, viscous case

(c) Total energy v.s. time

Fig. 7. Viscous shallow water equations, $\zeta = 10 \, \mathrm{m}^2\mathrm{s}^{-1}$, Dam-Break, $1{,}400 \times 1{,}400 \, \mathrm{m}^2$ basin, reflective-slip boundary, $\Delta x = \Delta y = 7 \, \mathrm{m}$, $\Delta t = 0.002 \, \mathrm{s}$

because of the energy-preserving shallow water solver with the increase of the total enstrophy. For the viscous case with $\zeta = 10 \, \mathrm{m}^2\mathrm{s}^{-1}$, Fig. 6b, d show the smoother solutions than inviscid solutions in Fig. 6a, c. The amplitude of those wiggles near the breach are significantly reduced though they are still detectable. Further refinement of the mesh from (100×100) to (200×200) generates very smooth solutions of the water depth h in Fig. 7a, b, when the oscillations are limited in the very small mesh scale.

Acknowledgments

References

[AL77]   Akio Arakawa and Vivian R. Lamb. Computational design of the basic dynamical process of the ucla general circulation model. *Meth. Comput. Phys.*, 17:173–265, 1977.

[AL81]   Akio Arakawa and Vivian R. Lamb. A potential enstrophy and energy conserving scheme for the shallow water equations. *Mont. Weat. Rev.*, 109:18–36, 1981.

[Ara97] Akio Arakawa. Computational design for long-term numerical integration of the equations of fluid motion: two-dimensional incompressible flow. I [J. Comput. Phys. **1** (1966), no. 1, 119–143]. *J. Comput. Phys.*, 135(2):101–114, 1997. With an introduction by Douglas K. Lilly, Commemoration of the 30th anniversary {of J. Comput. Phys.}.

[CK04]   Alina Chertock and Alexander Kurganov. On a hybrid finite-volume-particle method. *M2AN Math. Model. Numer. Anal.*, 38(6):1071–1091, 2004.

[Daf00] Constantine M. Dafermos. *Hyperbolic conservation laws in continuum physics*, volume 325 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 2000.

[EO80]   Bjorn Engquist and Stanley Osher. Stable and entropy satisfying approximations for transonic flow calculations. *Math. Comp.*, 34(149):45–75, 1980.

[FC90]   Robert J. Fennema and M. Hanif Chaudhry. Explicit methods for 2d transient free-surface flows. *J. Hydraul. Eng. ASCE*, 116(8):1013–1034, 1990.

[GL88]   Jonathan Goodman and Peter D. Lax. On dispersive difference schemes. I. *Comm. Pure Appl. Math.*, 41(5):591–613, 1988.

[Gla87] Paul Glaister. Difference schemes for the shallow water equations. Numerical Analysis Report 9/87, University of Reading, Department of Mathematics, 1987.

[God61] Sergei K. Godunov. An interesting class of quasi-linear systems. *Dokl. Akad. Nauk SSSR*, 139:521–523, 1961.

[GST01] Sigal Gottlieb, Chi-Wang Shu, and Eitan Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM Rev.*, 43(1):89–112 (electronic), 2001.

[HL91]   Thomas Y. Hou and Peter D. Lax. Dispersive approximations in fluid dynamics. *Comm. Pure Appl. Math.*, 44(1):1–40, 1991.

[Kru70] Stanislav N. Kružkov. First order quasilinear equations with several independent variables. *Mat. Sb. (N.S.)*, 81 (123):228–255, 1970.

[KT00]   Alexander Kurganov and Eitan Tadmor. New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations. *J. Comput. Phys.*, 160(1):241–282, 2000.

[Lax71] Peter D. Lax. *Shock waves and entropy.* Acaemic Press, New York, 1971. Contributions to Nonlinear Functional Analysis, (E.A.Zarantonello, ed.).

[Lax73] Peter D. Lax. *Hyperbolic systems of conservation laws and the mathematical theory of shock waves.* Society for Industrial and Applied Mathematics, Philadelphia, Pa., 1973. Conference Board of the Mathematical Sciences Regional Conference Series in Applied Mathematics, No. 11.

[LL96]   C. David Levermore and Jian-Guo Liu. Large oscillations arising in a dispersive numerical scheme. *Phys. D*, 99(2-3):191–216, 1996.

[LMR02] Philippe G. Lefloch, J. M. Mercier, and C. Rohde. Fully discrete, entropy conservative schemes of arbitrary order. *SIAM J. Numer. Anal.*, 40(5):1968–1992 (electronic), 2002.

[Moc80] Michael S. Mock. Systems of conservation laws of mixed type. *J. Differential Equations*, 37(1):70–88, 1980.

[Roe81] Phil L. Roe. Approximate Riemann solvers, parameter vectors, and difference schemes. *J. Comput. Phys.*, 43(2):357–372, 1981.

[Ser99] Denis Serre. *Systems of Conservation Laws, 1: Hyperbolicity, Entropies, Shock Waves.* Cambridge University Press, 1999.

[Tad87] Eitan Tadmor. The numerical viscosity of entropy stable schemes for systems of conservation laws. I. *Math. Comp.*, 49(179):91–103, 1987.

[Tad03] Eitan Tadmor. Entropy stability theory for difference approximations of nonlinear conservation laws and related time-dependent problems. *Acta Numer.*, 12:451–512, 2003.

[TZ06] Eitan Tadmor and Weigang Zhong. Entropy stable approximations of Navier-Stokes equations with no artificial numerical viscosity. *J. Hyperbolic Differ. Equ.*, 3(3):529–559, 2006.