INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN FLUIDS
Int. J. Numer. Meth. Fluids 2004; **46**:457–484 (DOI: 10.1002/fld.766)

Non-oscillatory relaxation methods for the shallow-water equations in one and two space dimensions

Mohammed Seaïd*,†

Fachbereich Mathematik, TU Darmstadt, 64289 Darmstadt, Germany

SUMMARY

In this paper, a new family of high-order relaxation methods is constructed. These methods combine general higher-order reconstruction for spatial discretization and higher order implicit-explicit schemes or TVD Runge–Kutta schemes for time integration of relaxing systems. The new methods retain all the attractive features of classical relaxation schemes such as neither Riemann solvers nor characteristic decomposition are needed. Numerical experiments with the shallow-water equations in both one and two space dimensions on flat and non-flat topography demonstrate the high resolution and the ability of our relaxation schemes to better resolve the solution in the presence of shocks and dry areas without using either Riemann solvers or front tracking techniques. Copyright © 2004 John Wiley & Sons, Ltd.

KEY WORDS: shallow-water equations; relaxation methods; higher order non-oscillatory schemes; Runge–Kutta methods

1. INTRODUCTION

During the last decades there has been an enormous amount of activity related to the construction of approximate solutions for the shallow-water equation written in conservative form as

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{S}(\mathbf{U}) \tag{1}$$

where

$$\mathbf{U} = \begin{pmatrix} h \\ hu \end{pmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{pmatrix} hu \\ hu^2 + \dfrac{g}{2}h^2 \end{pmatrix}, \quad \mathbf{S}(\mathbf{U}) = \begin{pmatrix} 0 \\ -ghZ_x \end{pmatrix}$$

*Correspondence to: M. Seaïd, Fachbereich Mathematik, TU Darmstadt, 64289 Darmstadt, Germany.
†E-mail: seaid@mathematik.tu-darmstadt.de

Contract/grant sponsor: Deutsche Forschungsgemeinschaft (DFG); contract/grant number: 1105/9-1

Published online 16 August 2004

Received 29 June 2003
Revised 30 January 2004

where $Z(x)$ is the function characterizing the bottom topography, $h(t,x)$ is the height of the water above the bottom, g is the acceleration due to gravity, u is the flow velocity. The two-dimensional shallow-water equations in conservative form read,

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x + \mathbf{G}(\mathbf{U})_y = \mathbf{S}(\mathbf{U}) \tag{2}$$

where

$$\mathbf{U} = \begin{pmatrix} h \\ hu \\ hv \end{pmatrix}, \quad \mathbf{F}(\mathbf{U}) = \begin{pmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ huv \end{pmatrix}$$

$$\mathbf{G}(\mathbf{U}) = \begin{pmatrix} hv \\ huv \\ hv^2 + \frac{1}{2}gh^2 \end{pmatrix}, \quad \mathbf{S}(\mathbf{U}) = \begin{pmatrix} 0 \\ -ghZ_x \\ -ghZ_y \end{pmatrix}$$

Here, the variables $h(t,x,y)$, g, $Z(x,y)$ are the same as in the one-dimensional case, u and v are flow velocity in the x and y direction, respectively.

Equations (1) and (2) have been widely used to model water flows, flood waves, dam-break problems, and have been studied in a number of books and papers in, among others [1–8]. Computing their numerical solutions is not trivial due to non-linearity, the presence of the convective term and the coupling of the equations through $\mathbf{S}(\mathbf{U})$. Hence, in many problems (1) and (2), the convective terms are distinctly more important than the source terms; particularly when certain non-dimensional parameters reach high values (as example of these parameters the Froude number), these convective terms are a source of computational difficulties and oscillations. It is well known that the solutions of Equations (1) and (2) present steep fronts and even shock discontinuities, which need to be resolved accurately in applications and often cause severe numerical difficulties [2, 9].

Relaxation schemes have recently been extensively applied and studied, see for example References [10–15]. The original relaxation model in Reference [10] was firstly proposed for the homogeneous hyperbolic system,

$$\mathbf{U}_t + \mathbf{F}(\mathbf{U})_x = \mathbf{0} \tag{3}$$

where $\mathbf{U}(t,x) \in \mathbb{R}^N$ is a N-vector of conserved quantities, $\mathbf{F}(\mathbf{U}) \in \mathbb{R}^N$ is non-linear flux function such that the Jacobian $\partial\mathbf{F}(\mathbf{U})/\partial\mathbf{U}$ is diagonalizable with real eigenvalues. In Reference [10], the conservation law is replaced by the system (known as relaxation system),

$$\begin{aligned} \mathbf{U}_t + \mathbf{V}_x &= 0 \\ \mathbf{V}_t + \mathbf{A}^2\mathbf{U}_x &= -\frac{1}{\tau}(\mathbf{V} - \mathbf{F}(\mathbf{U})) \end{aligned} \tag{4}$$

where $\mathbf{V} \in \mathbb{R}^N$, $\mathbf{A}^2 \in \mathbb{R}^{N\times N}$ is a diagonal matrix with positive diagonal elements A_k^2, $k = 1,\ldots,N$, and τ is the relaxation time. We use $\mathbf{A}^2$ to denote the matrix $\mathbf{A}$ in Reference [10] to avoid

square roots in formulas below. In the above and in what follows bold face type denotes vector quantities.

The relaxation system (4) has a typical semilinear structure with the two linear characteristic variables

$$\mathbf{V} + \mathbf{AU} \quad \text{and} \quad \mathbf{V} - \mathbf{AU} \tag{5}$$

The main feature in considering this model lies essentially on the semilinear structure of the relaxation system, which can be solved numerically without using Riemann solvers. Moreover it can be shown analytically, see for example [16–18], that solution to (4) approaches solution to the original problem (3) as $\tau \longrightarrow 0$ if the subcharacteristic condition

$$\frac{\lambda_1^2}{A_1^2} + \frac{\lambda_2^2}{A_2^2} + \cdots + \frac{\lambda_N^2}{A_N^2} \leqslant 1 \tag{6}$$

is satisfied in (4), where $\lambda_1, \ldots, \lambda_N$ are the eigenvalues of $\partial \mathbf{F}(\mathbf{U})/\partial \mathbf{U}$.

The relaxation schemes have been designed typically in Reference [10], such that a first-order upwind scheme and second-order MUSCL scheme used for the space discretization and a second-order implicit–explicit TVD Runge–Kutta scheme for the time integration. In fact, relaxation schemes are a combination of non-oscillatory upwind space discretization and a TVD implicit–explicit time integration of the resulting semi-discrete system, see for instance References [10, 14, 16]. The fully discrete system of Equation (4) is referred to as a relaxing system while that of the limiting system as the relaxation rate tends to zero, $\tau \to 0$, is called relaxed scheme. For systems of conservation laws, these schemes offer a very attractive alternative for standard integration schemes, consult [10, 19] for numerical illustrations. The application of relaxation scheme for the solution of hyperbolic conservation laws with source terms was also discussed in References [12, 13, 20, 21]. In all these works, only first- and second-order schemes have been discussed.

Despite the performance and competitive features of relaxations methods, there is not, to our knowledge, any attempt to solve the two-dimensional shallow-water equations using relaxation methods. However, there is a work that is related to this problem. Authors in Reference [22] have applied the first- and second-order relaxation schemes to the one-dimensional shallow-water equations. High-order accurate methods are important in scientific computing because they offer a mean to obtain accurate solutions with less work that may be required for methods of lower accuracy. In this paper, following the same ideas, we first extend the relaxation schemes of Reference [10] to higher-order by combining a third-order central weighted essentially non-oscillatory (CWENO) reconstruction and a third-order TVD implicit–explicit Runge–Kutta scheme. Then, we use the resulting scheme to compute the solutions of some test problems on shallow-water flows. The obtained results demonstrate good shock resolution with high accuracy in smooth regions and without any non-physical oscillations near the shock areas. From a practical point of view, the performance of our relaxation scheme is very attractive since the computed solutions remain, stable, monotone and highly accurate even on coarse meshes without solving Riemann problems or requiring special front tracking techniques.

The organization of the paper is as follows: Section 2 is devoted to the construction of third-order semi-discrete relaxation schemes for both one- and two-dimensional problems. In Section 3, we introduce a third-order implicit–explicit TVD Runge–Kutta scheme for time

integration. Section 4 illustrates the performance and accuracy of the schemes through experiments with several benchmark tests on shallow-water problems in both one and two space-dimensional cases. In the last section some conclusions are listed.

2. HIGHER ORDER RELAXATION SCHEME

2.1. The one-dimensional shallow-water equations

The relaxation system we propose for Equation (1) is

$$\begin{aligned} \mathbf{U}_t + \mathbf{V}_x &= \mathbf{S}(\mathbf{U}) \\ \mathbf{V}_t + \mathbf{A}^2\mathbf{U}_x &= -\frac{1}{\tau}(\mathbf{V} - \mathbf{F}(\mathbf{U})) \end{aligned} \tag{7}$$

At the limit ($\tau \longrightarrow 0$) Equation (7) are reduced to the original system (1) by the local equilibrium $\mathbf{V} = \mathbf{F}(\mathbf{U})$. The two linear characteristic variables of (7) are

$$\mathbf{V} + \mathbf{A}\mathbf{U} \quad \text{and} \quad \mathbf{V} - \mathbf{A}\mathbf{U} \tag{8}$$

To discretize the system of equations (7) we assume, for simplicity, an equally spaced grid with grid space size $\Delta x = x_{i+1/2} - x_{i-1/2}$ and we consider a cell in the spatial domain which we denote $\mathscr{I}_i = [x_{i-1/2}, x_{i+1/2}]$ containing the gridpoint x_i. We use

$$\mathbf{U}_{i+1/2} = \mathbf{U}(x_{i+1/2}, t) \quad \text{and} \quad \mathbf{U}_i = \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} \mathbf{U}(x,t)\,\mathrm{d}x$$

to denote the point-value and the approximate cell-average of the function $\mathbf{U}$ at $(x_{i+1/2}, t)$, and (x_i, t), respectively. We also use the following difference notation:

$$\mathscr{D}_x \mathbf{U}_i = \frac{\mathbf{U}_{i+1/2} - \mathbf{U}_{i-1/2}}{\Delta x} \tag{9}$$

Then, a semi-discrete approximation for the system of equations (7) can be written as

$$\begin{aligned} \frac{\mathrm{d}\mathbf{U}_i}{\mathrm{d}t} + \mathscr{D}_x \mathbf{V}_i &= \mathbf{S}(\mathbf{U})_i \\ \frac{\mathrm{d}\mathbf{V}_i}{\mathrm{d}t} + \mathbf{A}^2 \mathscr{D}_x \mathbf{U}_i &= -\frac{1}{\tau}(\mathbf{V}_i - \mathbf{F}(\mathbf{U})_i) \end{aligned} \tag{10}$$

The kth component of the approximate solution is reconstructed by a piecewise polynomial over the grid points as

$$U_k(x,t) = \sum_i \mathscr{P}_i(x; \mathbf{U})\chi_i(x), \quad \chi_i = \mathbb{I}_{\mathscr{I}_i} \tag{11}$$

where $\mathscr{P}_i$'s are polynomials defined in $\mathscr{I}_i$. The values of U_k at the cell boundary point between cells $\mathscr{I}_i$ and $\mathscr{I}_{i+1}$, $x_{i+1/2}$, are denoted as

$$U_k^+(x_{i+1/2};\mathbf{U}) = \mathscr{P}_{i+1}(x_{i+1/2};\mathbf{U}) \quad \text{and} \quad U_k^-(x_{i+1/2};\mathbf{U}) = \mathscr{P}_i(x_{i+1/2};\mathbf{U})$$

Now and henceforth, the subscript k will be dropped. The degree of the polynomial $\mathscr{P}_i$ is determined by the required order of accuracy of the method. In this paper, we consider the third-order CWENO reconstruction in Reference [23] which is also the compact central scheme reconstruction [24]. Thus,

$$\mathscr{P}_i(x;\mathbf{U}) = \omega_L P_L(x) + \omega_R P_R(x) + \omega_C P_C(x)$$

where

$$\omega_l = \frac{\alpha_l}{\sum_m \alpha_m}, \quad l,m \in \{L,R,C\}, \quad \alpha_l = \frac{c_l}{(\varepsilon + IS_l)^p}, \quad c_L = c_R = \frac{1}{4}, \quad c_C = \frac{1}{2} \tag{12}$$

Note that the normalizing factor $\sum_m \alpha_m$ is used here to guarantee $\sum_l \omega_l = 1$. The smoothness indicators IS_l and the polynomials $P_l(x)$ are given by

$$IS_L = (U_i - U_{i-1})^2, \quad IS_R = (U_{i+1} - U_i)^2, \quad IS_C = 13/3(U_{i+1} - 2U_i + U_{i-1})^2 + 1/4(U_{i+1} - U_{i-1})^2$$

$$P_L(x) = U_i + \frac{U_i - U_{i-1}}{\Delta x}(x - x_i), \quad P_R(x) = U_i + \frac{U_{i+1} - U_i}{\Delta x}(x - x_i)$$

$$P_C(x) = U_i - 1/12(U_{i+1} - 2U_i + U_{i-1}) + \frac{U_{i+1} - U_{i-1}}{2(\Delta x)}(x - x_i) + \frac{(U_{i+1} - 2U_i + U_{i-1})}{(\Delta x)^2}(x - x_i)^2$$

The constant ε in (12) guarantees that the denominator does not vanish and is empirically taken to be 10^{-6}. Likewise the value of p is chosen to provide the highest accuracy in smooth areas and ensures the non-oscillatory nature of the solution near the discontinuities and is selected to be $p = 2$.

With this background we can now reconstruct the characteristic variables (8) as follows:

$$(V + A_k U)_{i+1/2} = (V + A_k U)^-_{i+1/2} = \mathscr{P}_i(x_{i+1/2};\mathbf{V} + \mathbf{A}\mathbf{U})$$

$$(V - A_k U)_{i+1/2} = (V - A_k U)^+_{i+1/2} = \mathscr{P}_{i+1}(x_{i+1/2};\mathbf{V} - \mathbf{A}\mathbf{U})$$

Hence

$$U_{i+1/2} = \frac{1}{2A_k}(\mathscr{P}_i(x_{i+1/2}; \mathbf{V} + \mathbf{AU}) - \mathscr{P}_{i+1}(x_{i+1/2}; \mathbf{V} - \mathbf{AU}))$$

$$V_{i+1/2} = \tfrac{1}{2}(\mathscr{P}_i(x_{i+1/2}; \mathbf{V} + \mathbf{AV}) + \mathscr{P}_{i+1}(x_{i+1/2}; \mathbf{V} - \mathbf{AU}))$$

For completeness we write down the explicit expressions of the flux variables below

$$\begin{aligned}
V_{i+1/2} &= \tfrac{1}{2}(\mathscr{P}_i(x_{i+1/2}; \mathbf{V} + \mathbf{AU}) + \mathscr{P}_{i+1}(x_{i+1/2}; \mathbf{V} - \mathbf{Au})) \\
&= \tfrac{1}{2}\left\{\omega_L^+\left[(V + A_k u)_i + \tfrac{1}{2}((V + A_k U)_i - (V + A_k U)_{i-1})\right]\right. \\
&\quad + \omega_R^+\left[(V + A_k U)_i + \tfrac{1}{2}((V + A_k U)_{i+1} - (V + A_k U)_i)\right] \\
&\quad + \omega_C^+\left[(V + A_k U)_i - \tfrac{1}{12}((V + A_k U)_{i+1} - 2(V + A_k U)_i + (V + A_k U)_{i-1})\right. \\
&\quad + \tfrac{1}{4}((V + A_k U)_{i+1} - (V + A_k U)_{i-1}) \\
&\quad \left. + \tfrac{1}{4}((V + A_k u)_{i+1} - 2(V + A_k u)_i + (V + A_k U)_{i-1})\right] \\
&\quad + \omega_L^-\left[(V - A_k U)_{i+1} + \tfrac{1}{2}((V - A_k U)_{i+1} - (V - A_k U)_i)\right] \\
&\quad + \omega_R^-\left[(V - A_k U)_{i+1} + \tfrac{1}{2}((V - A_k U)_{i+2} - (V - A_k U)_{i+1})\right] \\
&\quad + \omega_C^-\left[(V - A_k U)_{i+1} - \tfrac{1}{12}((V - A_k U)_{i+2} - 2(V - A_k U)_{i+1} + (V - A_k U)_i)\right. \\
&\quad + \tfrac{1}{4}((V - A_k U)_{i+2} - (V - A_k U)_i) \\
&= \left.\tfrac{1}{4}((V - A_k U)_{i+2} - 2(V - A_k U)_{i+1} + (V - A_k U)_i)\right]\Big\}
\end{aligned}$$

where

$$\omega_l^{\pm} = \frac{\alpha_l^{\pm}}{\sum_m \alpha_m^{\pm}}, \quad l, m \in \{L, R, C\}, \quad \alpha_l^{\pm} = \frac{c_l}{(\varepsilon + IS_l^{\pm})^p}, \quad c_L = c_R = \frac{1}{4}, \quad c_C = \frac{1}{2} \tag{13}$$

$$IS_L^{\pm} = ((V \pm A_k U)_i - (V \pm A_k U)_{i-1})^2, \quad IS_R^{\pm} = ((V \pm A_k U)_{i+1} - (V \pm A_k U)_i)^2$$

$$\begin{aligned}
IS_C^{\pm} &= \tfrac{13}{3}((V \pm A_k U)_{i+1} - 2(V \pm A_k U)_i + (V \pm A_k U)_{i-1})^2 \\
&\quad + \tfrac{1}{4}((V \pm A_k U)_{i+1} - (V \pm A_k U)_{i-1})^2
\end{aligned}$$

The expression for $U_{i+1/2}$ can be derived analogously. Therefore, we obtain the following expressions for $U_{i+1/2}$ and analogously for $V_{i+1/2}$:

$$U_{i+1/2} = \frac{U_i + U_{i+1}}{2} - \frac{V_{i+1} - V_i}{2A_k} + \frac{\boldsymbol{\sigma}_i^+ + \boldsymbol{\sigma}_{i+1}^-}{4A_k}$$

$$V_{i+1/2} = \frac{V_i + V_{i+1}}{2} - A_k \frac{U_{i+1} - U_i}{2} + \frac{\boldsymbol{\sigma}_i^+ - \boldsymbol{\sigma}_{i+1}^-}{4}$$

where

$$\begin{aligned}\boldsymbol{\sigma}_i^{\pm} &= \omega_L^{\pm}((V \pm A_k U)_i - (V \pm A_k U)_{i-1}) + \omega_R^{\pm}((V \pm A_k U)_{i+1} - (V \pm A_k U)_i) \\ &+ \frac{\omega_C^{\pm}}{2}((V \pm A_k U)_{i+1} - (V \pm A_k U)_{i-1}) \\ &+ \frac{\omega_C^{\pm}}{3}((V \pm A_k U)_{i+1} - 2(V \pm A_k U)_i + (V \pm A_k U)_{i-1})\end{aligned}$$

In (10) we approximate the flux $\mathbf{F}(\mathbf{U})_i$ and the source $\mathbf{S}(\mathbf{U})_i$ using the fourth-order Simpson quadrature rule,

$$\mathbf{F}(\mathbf{U})_i = \tfrac{1}{6}(\mathbf{F}(\mathbf{U}_{i+1/2}) + 4\mathbf{F}(\mathbf{U}_i) + \mathbf{F}(\mathbf{U}_{i-1/2})) \tag{14}$$

and similarly for $\mathbf{S}(\mathbf{U})_i$ with the reconstruction given above is used

$$U_{i+1/2} = U_{i+1/2}^-, \quad U_i = \mathscr{P}_i(x_i; \mathbf{U}), \quad U_{i-1/2} = U_{i-1/2}^+$$

Thus, for the third-order reconstruction we obtain the following approximations:

$$\begin{aligned}U_{i+1/2} &= \omega_L P_L(x_{i+1/2}) + \omega_R P_R(x_{i+1/2}) + \omega_C P_C(x_{i+1/2}) \\ U_i &= \omega_L P_L(x_i) + w_R P_R(x_i) + \omega_C P_C(x_i) \\ U_{i-1/2} &= \omega_L P_L(x_{i-1/2}) + \omega_R P_R(x_{i-1/2}) + \omega_C P_C(x_{i-1/2})\end{aligned} \tag{15}$$

with the corresponding weights, ω_L, ω_R, ω_C of the polynomials defined accordingly.

2.2. The two-dimensional shallow-water equations

The two-dimensional relaxation system associated with Equation (2) reads,

$$\begin{aligned}\mathbf{U}_t + \mathbf{V}_x + \mathbf{W}_y &= \mathbf{S}(\mathbf{U}) \\ \mathbf{V}_t + \mathbf{A}^2\mathbf{U}_x &= -\frac{1}{\tau}(\mathbf{V} - \mathbf{F}(\mathbf{U})) \\ \mathbf{W}_t + \mathbf{B}^2\mathbf{U}_y &= -\frac{1}{\tau}(\mathbf{W} - \mathbf{G}(\mathbf{U}))\end{aligned} \tag{16}$$

Consequently, in the limit system (16) approaches the original system (2) by the local equilibrium $\mathbf{V}=\mathbf{F}(\mathbf{U})$ and $\mathbf{W}=\mathbf{G}(\mathbf{U})$. System (16) has linear characteristic variables given by

$$\mathbf{V}\pm\mathbf{A}\mathbf{U} \quad \text{and} \quad \mathbf{W}\pm\mathbf{B}\mathbf{U} \tag{17}$$

For the space discretization of Equation (16), we cover the spatial domain with rectangular cells $\mathscr{C}_{i,j}=[x_{i-1/2},x_{i+1/2}]\times[y_{j-1/2},y_{j+1/2}]$ of uniform sizes Δx and Δy. The cells, $\mathscr{C}_{i,j}$, are centred at $(x_i=i\Delta x, y_j=j\Delta y)$. We use the notations $\mathbf{U}_{i\pm1/2,j}(t)=\mathbf{U}(x_{i\pm1/2},y_j,t)$, $\mathbf{U}_{i,j\pm1/2}(t)=\mathbf{U}(x_i,\ y_{j\pm1/2},\ t)$ and

$$\mathbf{U}_{i,j}(t)=\frac{1}{\Delta x}\frac{1}{\Delta y}\int_{x_{i-1/2}}^{x_{i+1/2}}\int_{y_{-1/2}}^{y_{i+1/2}}\mathbf{U}(x,y,t)\,\mathrm{d}x\,\mathrm{d}y$$

to denote the point-values and the approximate cell-average of $\mathbf{U}$ at $(x_{i\pm1/2},\ y_j,\ t)$, $(x_i,\ y_{j\pm1/2},\ t)$, and (x_i,y_j,t), respectively. Then, the semi-discrete approximation of (16) is,

$$\begin{aligned}
&\frac{\mathrm{d}\mathbf{U}_{i,j}}{\mathrm{d}t}+\mathscr{D}_x\mathbf{V}_{i,j}+\mathscr{D}_y\mathbf{W}_{i,j}=\mathbf{S}(\mathbf{U})_{i,j}\\
&\frac{\mathrm{d}\mathbf{V}_{i,j}}{\mathrm{d}t}+\mathbf{A}^2\mathscr{D}_x\mathbf{U}_{i,j}=-\frac{1}{\tau}(\mathbf{V}_{i,j}-\mathbf{F}(\mathbf{U})_{i,j})\\
&\frac{\mathrm{d}\mathbf{W}_{i,j}}{\mathrm{d}t}+\mathbf{B}^2\mathscr{D}_y\mathbf{U}_{i,j}=-\frac{1}{\tau}(\mathbf{W}_{i,j}-\mathbf{G}(\mathbf{U})_{i,j})
\end{aligned} \tag{18}$$

where $\mathscr{D}_x$ and $\mathscr{D}_y$ are difference operators defined by

$$\mathscr{D}_x\mathbf{U}_{i,j}=\frac{\mathbf{U}_{i+1/2,j}-\mathbf{U}_{i-1/2,j}}{\Delta x},\quad \mathscr{D}_y\mathbf{U}_{i,j}=\frac{\mathbf{U}_{i,j+1/2}-\mathbf{U}_{i,j-1/2}}{\Delta y}$$

The approximate solution is reconstructed by a piecewise polynomial over the grid points as

$$U(x,y,t)=\sum_{i,j}\mathscr{P}_{i,j}(x,y;\mathbf{U})\chi_{i,j}(x,y),\quad \chi_{i,j}=\mathbb{I}_{\mathscr{C}_{i,j}} \tag{19}$$

where $\mathscr{P}_{i,j}$'s are polynomials defined in $\mathscr{C}_{i,j}$ and reconstructed dimension by dimension as

$$\mathscr{P}_{i,j}(x,y;\mathbf{U})=\mathscr{P}_i(x;\mathbf{U})+\mathscr{P}_j(y;\mathbf{U})$$

In the following, we formulate the x- direction polynomial $\mathscr{P}_i(x;\mathbf{U})$, the formulation of $\mathscr{P}_j(y;\mathbf{U})$ can be done analogously. Hence

$$\mathscr{P}_i(x;\mathbf{U})=\omega_L P_L(x)+\omega_R P_R(x)+\omega_C P_C(x)$$

where the weights $\omega_l,\ l\in\{L,R,C\}$ are the same as in (12) with

$$IS_L=(U_{i,j}-U_{i-1,j})^2,\quad IS_R=(U_{i+1,j}-U_{i,j})^2$$

$$IS_C = \tfrac{13}{3}(U_{i+1,j} - 2U_{i,j} + U_{i-1,j})^2 + \tfrac{1}{4}(U_{i+1,j} - U_{i-1,j})^2$$

$$P_L(x) = \frac{U_{i,j}}{2} + \frac{U_{i,j} - U_{i-1,j}}{\Delta x}(x - x_i), \quad P_R(x) = \frac{U_{i,j}}{2} + \frac{U_{i+1,j} - U_{i,j}}{\Delta x}(x - x_i)$$

$$P_C(x) = \frac{U_{i,j}}{2} - \tfrac{1}{24}(U_{i+1,j} - 2U_{i,j} + U_{i-1,j}) - \tfrac{1}{24}(U_{i,j+1} - 2U_{i,j} + U_{i,j-1}) + \frac{U_{i+1,j} - U_{i-1,j}}{2(\Delta x)}(x - x_i) + \frac{(U_{i+1,j} - 2U_{i,j} + U_{i-1,j})}{(\Delta x)^2}(x - x_i)^2$$

We can now discretize the characteristic variables (17) as follows:

$$\begin{aligned} (V + A_k U)_{i+1/2,j} &= \mathscr{P}_i(x_{i+1/2}; \mathbf{V} + \mathbf{AU}), \quad (V - A_k U)_{i+1/2,j} = \mathscr{P}_{i+1}(x_{i+1/2}; \mathbf{V} - \mathbf{AU}) \\ (W + B_k U)_{i,j+1/2} &= \mathscr{P}_j(y_{j+1/2}; \mathbf{W} + \mathbf{BU}), \quad (W - B_k U)_{i,j+1/2} = \mathscr{P}_{j+1}(y_{j+1/2}; \mathbf{W} - \mathbf{BU}) \end{aligned} \tag{20}$$

Recall that U, V, W, A_k and B_k are the kth ($k = 1, \ldots, N$) components of $\mathbf{U}$, $\mathbf{V}$, $\mathbf{W}$, $\mathbf{A}$ and $\mathbf{B}$, respectively. Solving (20) for the unknowns $U_{i+1/2,j}$, $V_{i+1/2,j}$, $U_{i,j+1/2}$, and $W_{i,j+1/2}$ gives,

$$U_{i+1/2,j} = \frac{1}{2A_k}(\mathscr{P}_i(x_{i+1/2}; \mathbf{V} + \mathbf{AU}) - \mathscr{P}_{i+1}(x_{i+1/2}; \mathbf{V} - \mathbf{AU}))$$

$$v_{i+1/2,j} = \tfrac{1}{2}\left(\mathscr{P}_i(x_{i+1/2}; \mathbf{V} + \mathbf{AU}) + \mathscr{P}_{i+1}(x_{i+1/2}; \mathbf{V} - \mathbf{AU})\right)$$

$$U_{i,j+1/2} = \frac{1}{2B_k}(\mathscr{P}_j(y_{j+1/2}; \mathbf{W} + \mathbf{BU}) - \mathscr{P}_{j+1}(x_{j+1/2}; \mathbf{W} - \mathbf{BU}))$$

$$W_{i,j+1/2} = \tfrac{1}{2}(\mathscr{P}_j(x_{j+1/2}; \mathbf{W} + \mathbf{BU}) + \mathscr{P}_{j+1}(x_{j+1/2}; \mathbf{W} - \mathbf{BU}))$$

Therefore, we obtain the following expressions for the numerical fluxes in (18):

$$U_{i+1/2,j} = \frac{U_{i,j} + U_{i+1,j}}{2} - \frac{V_{i+1,j} - V_{i,j}}{2A_k} + \frac{\boldsymbol{\sigma}^{x,+}_{i,j} + \boldsymbol{\sigma}^{x,-}_{i+1,j}}{4A_k}$$

$$V_{i+1/2,j} = \frac{V_{i,j} + V_{i+1,j}}{2} - A_k\frac{U_{i+1,j} - U_{i,j}}{2} + \frac{\boldsymbol{\sigma}^{x,+}_{i,j} - \boldsymbol{\sigma}^{x,-}_{i+1,j}}{4}$$

$$U_{i,j+1/2} = \frac{U_{i,j} + U_{i,j+1}}{2} - \frac{W_{i,j+1} - W_{i,j}}{2B_k} + \frac{\boldsymbol{\sigma}^{y,+}_{i,j} + \boldsymbol{\sigma}^{y,-}_{i,j+1}}{4B_k}$$

$$W_{i,j+1/2} = \frac{W_{i,j} + W_{i+1,j}}{2} - B_k\frac{U_{i,j+1} - U_{i,j}}{2} + \frac{\boldsymbol{\sigma}^{y,+}_{i,j} - \boldsymbol{\sigma}^{y,-}_{i,j+1}}{4}$$

where $\boldsymbol{\sigma}_{i,j}^{x,\pm}$ and $\boldsymbol{\sigma}_{i,j}^{y,\pm}$ are the slopes of $\mathbf{V} \pm \mathbf{AU}$ and $\mathbf{W} \pm \mathbf{BU}$ on the cell $\mathscr{C}_{i,j}$, respectively. They are defined by

$$\begin{aligned}
\boldsymbol{\sigma}_{i,j}^{x,\pm} &= \omega_L^{\pm}((V \pm A_k U)_{i,j} - (V \pm A_k U)_{i-1,j}) + \omega_R^{\pm}((V \pm A_k U)_{i+1,j} - (V \pm A_k U)_{i,j}) \\
&\quad + \frac{\omega_C^{\pm}}{2}((V \pm A_k U)_{i+1,j} - (V \pm A_k U)_{i-1,j}) \\
&\quad + \frac{\omega_C^{\pm}}{3}((V \pm A_k U)_{i+1,j} - 2(V \pm A_k U)_{i,j} + (V \pm A_k U)_{i-1,j}) \\
&\quad - \frac{\omega_C^{\pm}}{6}((V \pm A_k U)_{i,j+1} - 2(V \pm A_k U)_{i,j} + (V \pm A_k U)_{i,j-1}) \\
\boldsymbol{\sigma}_{i,j}^{y,\pm} &= \omega_L^{\pm}((W \pm B_k U)_{i,j} - (W \pm B_k U)_{i,j-1}) + \omega_R^{\pm}((W \pm B_k U)_{i,j+1} - (W \pm B_k U)_{i,j}) \\
&\quad + \frac{\omega_C^{\pm}}{2}((W \pm B_k U)_{i,j+1} - (W \pm B_k U)_{i,j-1}) \\
&\quad + \frac{\omega_C^{\pm}}{3}((W \pm B_k U)_{i,j+1} - 2(W \pm B_k U)_{i,j} + (W \pm B_k U)_{i,j-1}) \\
&\quad - \frac{\omega_C^{\pm}}{6}((W \pm B_k U)_{i+1,j} - 2(W \pm B_k U)_{i,j} + (W \pm B_k U)_{i-1,j})
\end{aligned}$$

The weight parameters $\omega_L^{\pm}$, $\omega_R^{\pm}$ and $\omega_C^{\pm}$ for $\boldsymbol{\sigma}_{i,j}^{x,\pm}$ are given by (13) with,

$$IS_L^{\pm} = ((V \pm A_k U)_{i,j} - (V \pm A_k U)_{i-1,j})^2, \quad IS_R^{\pm} = ((V \pm A_k U)_{i+1,j} - (V \pm A_k U)_{i,j})^2$$

$$\begin{aligned}
IS_C^{\pm} &= \tfrac{13}{3}((V \pm A_k U)_{i+1,j} - 2(V \pm A_k U)_{i,j} + (V \pm A_k U)_{i-1,j})^2 \\
&\quad + \tfrac{1}{4}((V \pm A_k U)_{i+1,j} - (V \pm A_k U)_{i-1,j})^2
\end{aligned}$$

The corresponding weight parameters for $\boldsymbol{\sigma}_{i,j}^{y,\pm}$ are obtained by changing $V \pm A_k U$ to $W \pm B_k U$ in the above formulas and differentiating respect to the y- direction. We close by pointing out that in this higher-order scheme we approximate $\mathbf{F}(\mathbf{U})_{i,j}$, $\mathbf{G}(\mathbf{U})_{i,j}$ and $\mathbf{S}(\mathbf{U})_{i,j}$ in (18) using the fourth-order Simpson quadrature rule dimension by dimension as in (14).

Remark

Another way to construct relaxation system that gives at the limit Equation (2) is to incorporate the source term into the flux function and use straightforwardly the scheme as in [10], i.e.

$$\begin{aligned}
&\mathbf{U}_t + \mathbf{V}_x + \mathbf{W}_y = \mathbf{0} \\
&\mathbf{V}_t + \tilde{\mathbf{A}}^2 \mathbf{U}_x = -\frac{1}{\tau}(\mathbf{V} - \mathbf{F}(\mathbf{U})) + \frac{1}{2\tau}\int^x \mathbf{S}(\mathbf{U})\,\mathrm{d}x \qquad (21) \\
&\mathbf{W}_t + \tilde{\mathbf{B}}^2 \mathbf{U}_y = -\frac{1}{\tau}(\mathbf{W} - \mathbf{G}(\mathbf{U})) + \frac{1}{2\tau}\int^y \mathbf{S}(\mathbf{U})\,\mathrm{d}y
\end{aligned}$$

The above semi-discretization remains valid for (21) with the only difference that the relaxation system (21) approaches, in the limit, the original system (2) by the local equilibrium,

$$\mathbf{V} = \tilde{\mathbf{F}}(\mathbf{U}) = \mathbf{F}(\mathbf{U}) - \frac{1}{2}\int^{x} \mathbf{S}(\mathbf{U})\,\mathrm{d}x, \quad \text{and} \quad \mathbf{W} = \tilde{\mathbf{G}}(\mathbf{U}) = \mathbf{G}(\mathbf{U}) - \frac{1}{2}\int^{y} \mathbf{S}(\mathbf{U})\,\mathrm{d}y \tag{22}$$

We would like to point out that construction (21) has been experimented in Reference [22] for the first- and second-order schemes. In our computational test problems, the results obtained by the construction (21) are not presented because they overlap those obtained by the construction (16). This would not be the case if same characteristic speed $\mathbf{A} = \tilde{\mathbf{A}}$ and $\mathbf{B} = \tilde{\mathbf{B}}$ are used in both constructions, compare the results in Reference [22].

2.3. TVD Runge–Kutta methods

The semi-discrete formulations (10) or (18) can be rewritten in common ordinary differential equations notation as

$$\frac{\mathrm{d}\mathscr{Y}}{\mathrm{d}t} = \mathscr{F}(\mathscr{Y}) - \frac{1}{\tau}\mathscr{G}(\mathscr{Y}) \tag{23}$$

where the time-dependent vector functions

$$\mathscr{Y} = \begin{pmatrix} \mathbf{U}_i \\ \mathbf{V}_i \end{pmatrix}, \quad \mathscr{F}(\mathscr{Y}) = \begin{pmatrix} \mathbf{S}(\mathbf{U})_i - \mathscr{D}_x \mathbf{V}_i \\ -\mathbf{A}^2 \mathscr{D}_x \mathbf{U}_i \end{pmatrix}, \quad \mathscr{G}(\mathscr{Y}) = \begin{pmatrix} \mathbf{0} \\ \mathbf{V}_i - \mathbf{F}(\mathbf{U})_i \end{pmatrix}$$

for the one-dimensional case (10) or

$$\mathscr{Y} = \begin{pmatrix} \mathbf{U}_{i,j} \\ \mathbf{V}_{i,j} \\ \mathbf{W}_{i,j} \end{pmatrix}, \quad \mathscr{F}(\mathscr{Y}) = \begin{pmatrix} \mathbf{S}(\mathbf{U})_{i,j} - \mathscr{D}_x \mathbf{V}_{i,j} - \mathscr{D}_y \mathbf{W}_{i,j} \\ -\mathbf{A}^2 \mathscr{D}_x \mathbf{U}_{i,j} \\ -\mathbf{B}^2 \mathscr{D}_y \mathbf{U}_{i,j} \end{pmatrix}$$

$$\mathscr{G}(\mathscr{Y}) = \begin{pmatrix} \mathbf{0} \\ \mathbf{V}_{i,j} - \mathbf{F}(\mathbf{U})_{i,j} \\ \mathbf{W}_{i,j} - \mathbf{G}(\mathbf{U})_{i,j} \end{pmatrix}$$

for the two-dimensional formulation (18). Due to the presence of stiff term in (23), one cannot use fully explicit schemes to integrate Equation (23), particularly when $\tau \longrightarrow 0$. On the other hand, integrating Equation (23) by fully implicit scheme, either linear or non-linear algebraic equations have to be solved at every time step of the computational process. To find solutions of such systems is computationally very demanding. In this paper we consider an alternative approach based on implicit–explicit (IMEX) Runge–Kutta splitting. The non-stiff stage of the splitting for $\mathscr{F}$ is straightforwardly treated by an explicit Runge–Kutta scheme, while the stiff stage for $\mathscr{G}$ is approximated by a diagonally implicit Runge–Kutta scheme. Compare References [25, 26] for more details.

Let Δt be the time step and $\mathscr{Y}^n$ denotes the approximate solution at $t = n\Delta t$. We formulate the IMEX scheme for system (23) as,

$$\begin{aligned} K_l &= \mathscr{Y}^n + \Delta t \sum_{m=1}^{l-1} \tilde{a}_{lm} \mathscr{F}(K_m) - \frac{\Delta t}{\tau} \sum_{m=1}^{s} a_{lm} \mathscr{G}(K_m), \quad l = 1, 2, \ldots, s \\ \mathscr{Y}^{n+1} &= \mathscr{Y}^n + \Delta t \sum_{l=1}^{s} \tilde{b}_l \mathscr{F}(K_l) - \frac{\Delta t}{\tau} \sum_{l=1}^{s} b_l \mathscr{G}(K_l) \end{aligned} \tag{24}$$

The $s \times s$ matrices $\tilde{A} = (\tilde{a}_{lm})$, $\tilde{a}_{lm} = 0$ for $m \geqslant l$ and $A = (a_{lm})$ are chosen such that the resulting scheme is explicit in $\mathscr{F}$, and implicit in $\mathscr{G}$. The s-vectors $\tilde{b}$ and b are the canonical coefficients which characterize the IMEX s-stage Runge–Kutta scheme [26]. They can be given by the standard double *tableau* in Butcher notation as

0	0	0	0	0	0	0	0	0	0	0	0	0
$\tilde{c}_2$	$\tilde{a}_{21}$	0	0	0	0	c_2	a_{21}	a_{22}	0	0	0
$\tilde{c}_3$	$\tilde{a}_{31}$	$\tilde{a}_{32}$	0	0	0	c_3	a_{31}	a_{32}	a_{33}	0	0
$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$	$\vdots$
$\tilde{c}_s$	$\tilde{a}_{s1}$	$\tilde{a}_{s2}$	$\cdots$	$\tilde{a}_{ss-1}$	0	c_s	a_{s1}	a_{s2}	$\cdots$	a_{ss-1}	a_{ss}
	$\tilde{b}_1$	$\tilde{b}_2$	$\cdots$	$\tilde{b}_{s-1}$	$\tilde{b}_s$		b_1	b_2	$\cdots$	b_{s-1}	b_s

Here, $\tilde{c}$ and c are s-vectors used in the non-autonomous cases. The left and right tables represent the explicit and the implicit Runge–Kutta methods, respectively. The implementation of the IMEX algorithm to solve (23) can be carried out in the following steps:

1. For $l = 1, \ldots, s$,
   (a) Evaluate K_l^* as: $K_l^* = \mathscr{Y}^n + \Delta t \sum_{m=1}^{l-2} \tilde{a}_{lm} \mathscr{F}(K_m) + \Delta t \tilde{a}_{ll-1} \mathscr{F}(K_{l-1})$
   (b) Solve for K_l : $K_l = K_l^* - \frac{\Delta t}{\varepsilon} \sum_{m=1}^{l-1} a_{lm} \mathscr{G}(K_m) - \frac{\Delta t}{\tau} a_{ll} \mathscr{G}(K_l)$
2. Update $\mathscr{Y}^{n+1}$ as: $\mathscr{Y}^{n+1} = \mathscr{Y}^n + \Delta t \sum_{l=1}^{s} \tilde{b}_l \mathscr{F}(K_l) - \frac{\Delta t}{\tau} \sum_{l=1}^{s} b_l \mathscr{G}(K_l)$

Notice that, using the above relaxation scheme neither linear algebraic equation nor non-linear source terms can arise. In addition, the high-order relaxation scheme is stable independently of τ, so that the choice of Δt is based only on the usual CFL condition

$$\mathrm{CFL} = \max_{1 \leqslant k \leqslant N} \left(\frac{\Delta t}{\Delta x}, A_k \frac{\Delta t}{\Delta x} \right) \leqslant 1 \tag{25}$$

for the one-dimensional problems or

$$\mathrm{CFL} = \max_{1 \leqslant k \leqslant N} \left(\frac{\Delta t}{\delta}, A_k \frac{\Delta t}{\Delta x}, B_k \frac{\Delta t}{\Delta y} \right) \leqslant 1 \tag{26}$$

for the two-dimensional problems. In (26), δ denotes the maximum cell size, $\delta = \max(\Delta x, \Delta y)$.

In our numerical computations we use the third-order IMEX scheme proposed in Reference [25], the associated double Butcher tables can be represented as

$$\begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & \gamma & 0 & 0 \\ 1-\gamma & \gamma-1 & 2-2\gamma & 0 \\ \hline & 0 & \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{c|ccc} 0 & 0 & 0 & 0 \\ \gamma & 0 & \gamma & 0 \\ 1-\gamma & 0 & 2-2\gamma & \gamma \\ \hline & 0 & \frac{1}{2} & \frac{1}{2} \end{array} \tag{27}$$

where $\gamma = (3+\sqrt{3})/6$. Other IMEX schemes of third and higher-order are also discussed in Reference [26]. However, most of them used more intermediate stages than the IMEX method (27). In practice, less stages in IMEX methods require less computational cost and also avoid order reduction in the overall method. The sensibility of relaxation schemes on the selection of IMEX methods has been addressed in Reference [27]. The authors in Reference [27] have performed several numerical tests on one-dimensional hyperbolic systems with relaxation. Obviously, at the limit ($\tau \longrightarrow 0$) the time integration procedure tends to a time integration scheme of the limit equations based on the explicit scheme given by the left table in (27).

Remarks

(i) Note that the first- and second-order relaxation schemes studied earlier in Reference [10] can be interpreted as (19) by taking

$$\mathscr{P}_{i,j}(x,y;\mathbf{U}) = U_{i,j} \quad \text{and} \quad \mathscr{P}_{i,j}(x,y;\mathbf{U}) = U_{i,j} + \frac{U'_{i,j}}{\Delta x}(x-x_i) + \frac{U^{\cdot}_{i,j}}{\Delta y}(y-y_j)$$

respectively. Here $U'_{i,j}/\Delta x$ and $U^{\cdot}_{i,j}/\Delta y$ are discrete slopes in the x and y directions approximated in Reference [10] by MUSCL method along with a chosen slop limiter function. The second-order time integration procedure in Reference [10] can be also represented as (24) where the explicit and implicit Runge–Kutta tables are given by

$$\begin{array}{c|cc} 0 & 0 & 0 \\ 1 & 1 & 0 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array} \qquad \begin{array}{c|cc} -1 & -1 & 0 \\ 2 & 1 & 1 \\ \hline & \frac{1}{2} & \frac{1}{2} \end{array}$$

(ii) In order to avoid initial and boundary layers in (16), initial and boundary conditions are chosen to be consistent to the associated local equilibrium. For instance, if initial data $\mathbf{U} = \mathbf{U}_0$ and Dirichlet boundary condition $\mathbf{U} = \mathbf{U}_b$ are given, then boundary and initial conditions for (16) are given by

$$\begin{aligned} \mathbf{V}(t,\mathbf{x}) &= \mathbf{F}(\mathbf{U}_b), \quad \mathbf{W}(t,\mathbf{x}) = \mathbf{G}(\mathbf{U}_b) \\ \mathbf{V}(0,\mathbf{x}) &= \mathbf{F}(\mathbf{U}_0(\mathbf{x})), \quad \mathbf{W}(0,\mathbf{x}) = \mathbf{G}(\mathbf{U}_0(\mathbf{x})) \end{aligned} \tag{28}$$

If instead, the relaxation system (21) is used, then boundary and initial conditions for $\mathbf{V}$ and $\mathbf{W}$ are chosen according to equilibrium (22). In general, any choice that leads at the limit to the associated boundary and initial equilibrium can be also used.

(iii) In practice, the characteristic speeds A_k and B_k in (16) can be chosen based on rough estimates of eigenvalues of $\partial\mathbf{F}(\mathbf{U})/\partial\mathbf{U}$ and $\partial\mathbf{G}(\mathbf{U})/\partial\mathbf{U}$, respectively. Another choice is to calculate A and B locally at every cell as

$$\begin{aligned} A_{i+1/2,j} &= \max_{U\in\{U^-_{i+1/2,j},U^+_{i+1/2,j}\}} \left|\frac{\partial\mathbf{F}}{\partial U_k}(U)\right| \\ B_{i,j+1/2} &= \max_{U\in\{U^-_{i,j+1/2},U^+_{i,j+1/2}\}} \left|\frac{\partial\mathbf{G}}{\partial U_k}(U)\right| \end{aligned} \tag{29}$$

where

$$U^+_{i+1/2,j} = \mathscr{P}_{i+1}(x_{i+1/2};\mathbf{U}), \quad U^-_{i+1/2,j} = \mathscr{P}_i(x_{i+1/2};\mathbf{U})$$

$$U^+_{i,j+1/2} = \mathscr{P}_{j+1}(y_{j+1/2};\mathbf{U}), \quad U^-_{i,j+1/2} = \mathscr{P}_j(y_{j+1/2};\mathbf{U})$$

A global choice is simply to take the maximum over the grid points in (29).

$$A_k = B_k = \max_{i,j}(A_{i+1/2,j}, B_{i,j+1/2}), \quad k = 1,\ldots,N$$

It is worth saying that, larger A_k and B_k values usually add more numerical dissipation.

3. NUMERICAL EXPERIMENTS

In this section, we perform numerical tests with our third-order relaxation scheme on the one- and two-dimensional shallow-water equations. In all our computations the CFL number is fixed and time steps Δt are calculated according to condition (25) or (26) that depend on either one- or two-dimensional problem. Furthermore, the relaxation rate τ is set to 10^{-6} in all the test examples presented in this section.

3.1. One-dimensional examples

The one-dimensional relaxation system of shallow-water equations is given by (7) with $\mathbf{A}^2 = \mathrm{diag}\{A_1^2, A_2^2\}$. In all the test cases presented here we choose

$$A_1 = A_2 = \max(\sup|u + \sqrt{gh}|, \sup|u - \sqrt{gh}|) \tag{30}$$

Note that $u \pm \sqrt{gh}$ are the two eigenvalues of the one-dimensional shallow-water equations. The following test cases are selected:

3.1.1. Dam-break problem. First we consider the dam-break problem in a rectangular channel with flat bottom, $Z(x) = 0$. The channel is of length 2000 m, the space discretization $\Delta x = 10$ m and the initial conditions are given by

$$h(0,x) = \begin{cases} h_\mathrm{l} & \text{if } x \leqslant 1000\,\mathrm{m}, \\ h_\mathrm{r} & \text{if } x > 1000\,\mathrm{m}, \end{cases} \qquad u(0,x) = 0\,\mathrm{m/s}$$

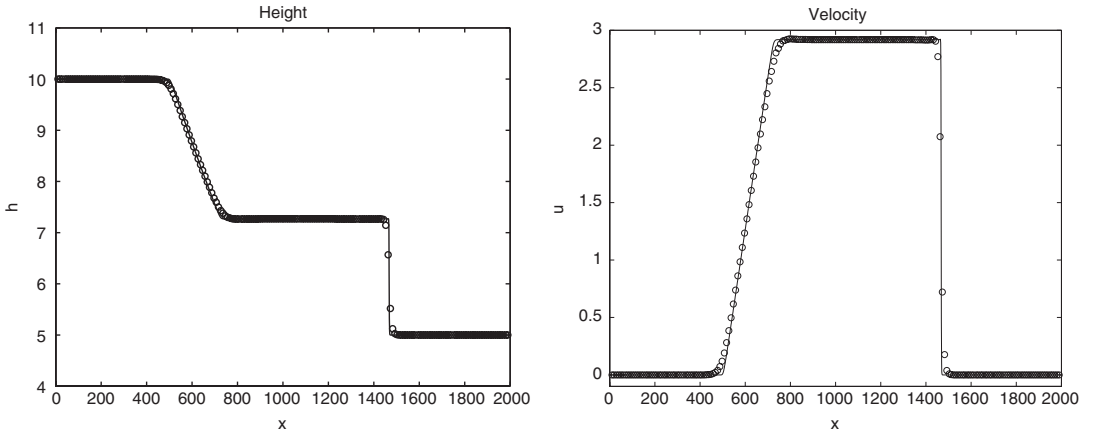


Figure 1. Water height and velocity plots for dam-break on wet bed, $h_r/h_l = 0.5$, at $t = 50$ s.

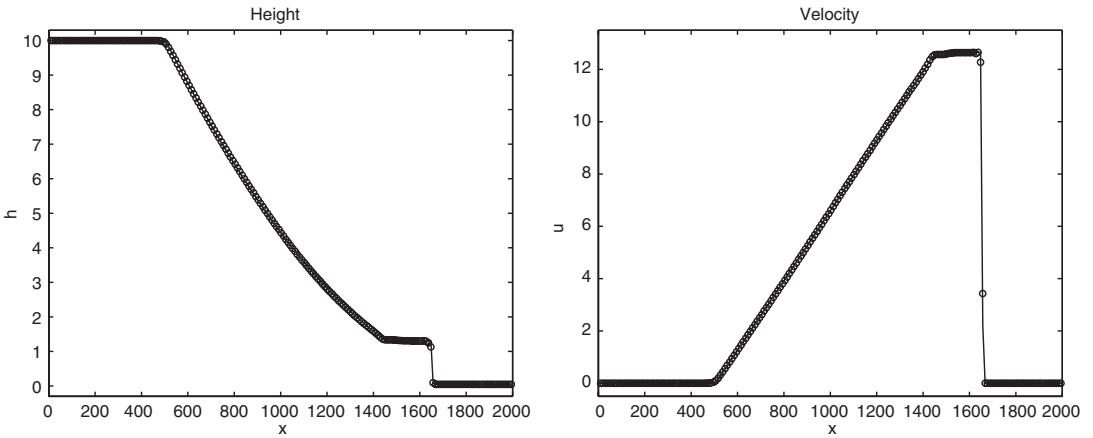


Figure 2. Water height and velocity plots for dam-break on wet bed, $h_r/h_l = 0.005$, at $t = 50$ s.

where $h_l > h_r$ in order to be consistent with the physical phenomenon of a dam-break from the left to the right. At $t = 0$ the dam collapses and the flow problem consists of a shock wave traveling downstream and a rarefaction wave traveling upstream. We start by assuming that in both sides of the dam there are water with corresponding heights $h_l = 10\,\text{m}$ and $h_r = 5\,\text{m}$ (depth ratio $h_r/h_l = 0.5$), for the second test $h_l = 10\,\text{m}$ and $h_r = 0.05\,\text{m}$ (depth ratio $h_r/h_l = 0.005$). Note that the flow structure in the channel is subcritical for depth ratio greater than 0.5, and is supercritical when depth ratio becomes smaller than 0.5. In Figures 1 and 2 we plot water height and velocity field at $t = 50$ s using CFL $= 0.5$ for both tests. The relaxation scheme captured correctly the discontinuity and the shock without need for very fine mesh, compare Reference [4].

Next, we consider a dry bed in the downstream of the dam, $h_l = 10\,\text{m}$ and $h_r = 0\,\text{m}$ (depth ratio $h_r/h_l = \infty$). This example is more difficult than the previous one because of the singularity that occurs at the transition point of the advancing front. The water height and the velocity

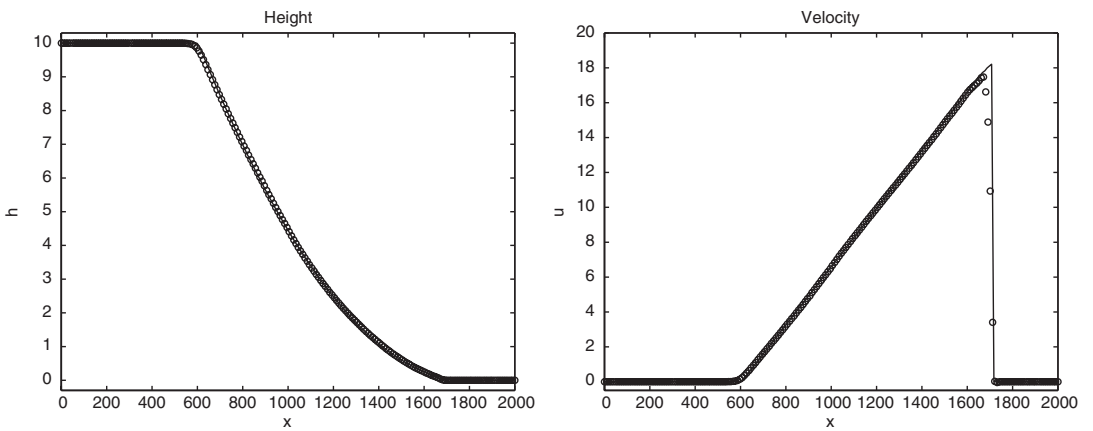


Figure 3. Water height and velocity plots for dam-break on dry bed, $h_r/h_l = \infty$, at $t = 40$ s.

flow at $t = 40$ s are shown in Figure 3 using CFL $= 0.5$. These results are in very close agreement with the exact solution and with small conservation error in velocity variable due to the fact that the third-order polynomial reconstruction (11) reduces to first-order accuracy in area where the velocity peak is localized. No oscillations or negative values have been observed in the solution. The analytical reference solutions for these test problems are due to Reference [1].

To compare the accuracy of our relaxation scheme to the first- and second-order schemes originally introduced in Reference [10], we have reproduced the results for the dam-break on wet bed (depth ratio $h_r/h_l = 0.5$) using first-, second- and third-order relaxation schemes. In Table I we display the relative errors at $t = 25$ s measured in term of L^1- norm by the difference between the pointvalues of the exact solution and the reconstructed pointvalues of the computed solution. We compare errors in both variables height h, and velocity u. In Table I, M stands for the number of gridpoints used in the computation. These results show that the relaxation schemes achieve their designed order of accuracy for this test problem. The high accuracy of our scheme over the relaxation schemes from Reference [10] is clearly demonstrated in both, height and velocity variables. Notice that the number of gridpoints for the second-order relaxation scheme to have the same error as the third-order scheme is almost two times that of the third-order scheme. The decay rate is slow for the velocity variable.

3.1.2. Flow over a hump. To investigate the ability of our algorithm to preserve the correct steady-state solutions, we apply the relaxation scheme to a series of benchmark test problems for lake at rest, subcritical, and transcritical flows. They are widely used to test numerical algorithms for the shallow-water equations. In all these test examples the channel length is 25 m and the bottom topography Z is defined as,

$$Z(x) = \begin{cases} 0.2 - 0.05(x - 10)^2 & \text{if } 8 \leqslant x \leqslant 12 \\ 0 & \text{otherwise} \end{cases} \tag{31}$$

Table I. Error-norms for dam-break on wet bed, $h_r/h_l = 0.5$, with CFL = 0.5.

		Errors in h		Errors in v	
M	Scheme	L^1-error	Rate	L^1-error	Rate
200	1st order	7.905E-1	—	9.982E-1	—
	2nd order	3.311E-1	—	6.036E-1	—
	3rd order	1.035E-1	—	3.713E-1	—
400	1st order	6.288E-1	0.33	8.872E-1	0.17
	2nd order	1.392E-1	1.25	2.997E-1	1.01
	3rd order	2.641E-2	1.97	1.259E-1	1.56
800	1st order	3.871E-1	0.70	6.317E-1	0.49
	2nd order	4.405E-2	1.66	1.398E-1	1.10
	3rd order	4.418E-3	2.58	2.539E-2	2.31
1600	1st order	2.031E-1	0.93	3.432E-1	0.88
	2nd order	1.041E-2	2.08	3.825E-2	1.87
	3rd order	4.774E-4	3.21	3.218E-3	2.98
3200	1st order	9.812E-2	1.05	1.801E-1	0.93
	2nd order	1.987E-3	2.39	9.629E-3	1.99
	3rd order	3.598E-5	3.73	3.832E-4	3.07

The initial conditions are given by

$$h(0,x) + Z(x) = H \text{ m}, \quad u(0,x) = 0 \text{ m/s} \tag{32}$$

Numerical results are shown for the flow at rest, the subcritical flow, the transcritical flow without shock, and the transcritical flow with shock. All these results are displayed at $t = 200$s using 200 gridpoints and a CFL number fixed to 0.5. The analytical solutions are also plotted (by solid lines) within the obtained numerical results.

Flow at rest: This test problem consists of Equations (1) and (31) augmented with initial condition (32), where $H = 2$ m. Figure 4 presents the computed water level and discharge, hu. As can be seen in this figure, the relaxation scheme preserves the correct steady flow, to almost machine accuracy. On the hump region, the scheme produces small errors in the velocity component due to the spatial reconstruction (11).

Subcritical flow: We solve Equations (1) and (31)–(32) subject to an upstream boundary condition on the discharge $hu = 4.42\,\text{m}^2/\text{s}$ and a downstream condition on the height $H = 2$ m. The results are shown in Figure 5 for the water level and discharge. Once again, the relaxation scheme resolves accurately this test problem with small errors in the discharge plot over the hump area.

Transcritical flow without shock: In this test case, we solve the Equations (1) and (31)–(32) using an upstream boundary condition for the discharge $hu = 1.53\,\text{m}^2/\text{s}$ and a downstream boundary condition for the water level $H = 0.66$ m only if the flow is subcritical. If the flow is supercritical, no condition is imposed. Figure 6 shows the water level and the discharge plots. As mentioned by many authors, the correct capturing of the discharge is more difficult than the water height in these test problems.

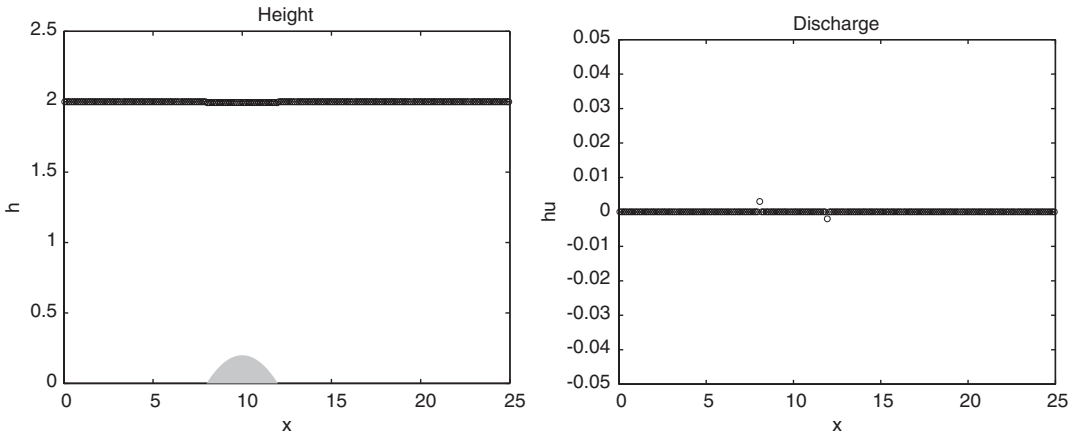


Figure 4. Water height and discharge plots for flow at rest.

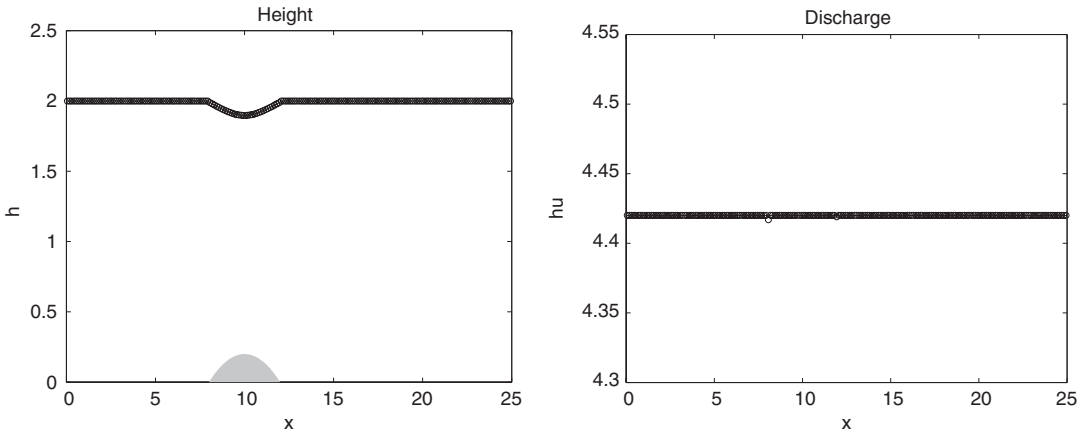


Figure 5. Water height and discharge plots for subcritical flow.

Transcritical flow with shock: This test problem is similar to the previous one but with different boundary conditions. Here, a discharge of $hu = 0.18\,\mathrm{m}^2/\mathrm{s}$ is imposed as the upstream boundary condition and a water level of $H = 0.33\,\mathrm{m}$ as the downstream boundary condition. The obtained results for this test are displayed in Figure 7. In comparison with the results reported in Reference [22] for the flow over a hump, the present relaxation scheme provides a good accuracy, such as the solution of flow discharge.

3.1.3. Drain on a non-flat bottom. This is a challenging numerical test example as it involves the calculation of dry areas in the computational domain. As proposed in Reference [6], the length of the channel is 25 m and the bed profile Z is mathematically defined by (31). The initial conditions are,

$$h(0,x) + Z(x) = 0.5\,\mathrm{m}, \quad u(0,x) = 0\,\mathrm{m/s}$$

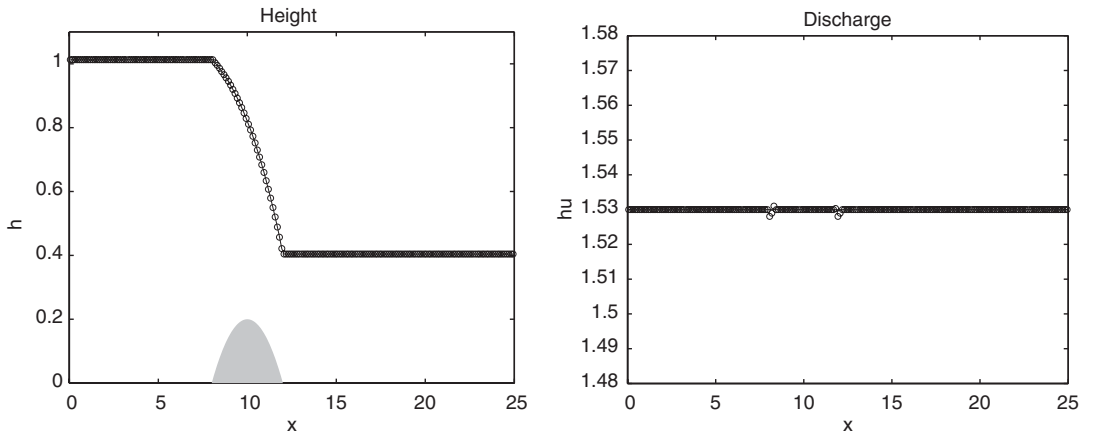


Figure 6. Water height and discharge plots for transcritical flow without shock.

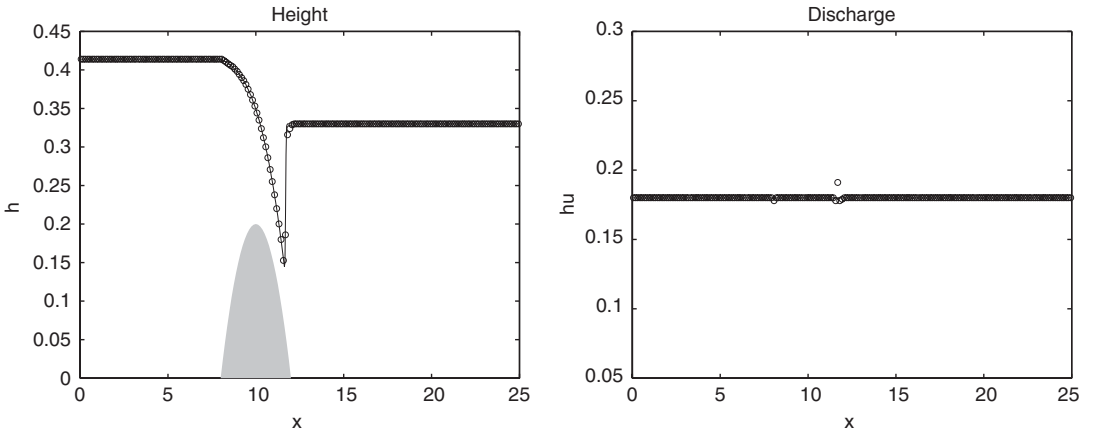


Figure 7. Water height and discharge plots for transcritical flow with shock.

Reflective conditions are used on the upstream boundary and the downstream boundary condition is that of a dry bed. The steady state solution of this problem is a flow at rest, in the left part of the hump $h+Z=0.2\,\mathrm{m}$, $u=0\,\mathrm{m/s}$ and dry flow, $h=0\,\mathrm{m}$, $u=0\,\mathrm{m/s}$ in the right of the hump.

We discretize the space domain into 250 uniform gridpoints and $\mathrm{CFL}=0.5$. In Figure 8, we present the evolution of water depth and the discharge plots at several times $t=10, 20, 100$, and 1000 s. The relaxation scheme performs very well for this case and gives results which converge to the expected steady-state solution. These results compare well with those published in Reference [6].

Notice that no modifications have been added to the method to overcome the dry areas in the computational domain. However, since for such area one or both eigenvalues of the Jacobian

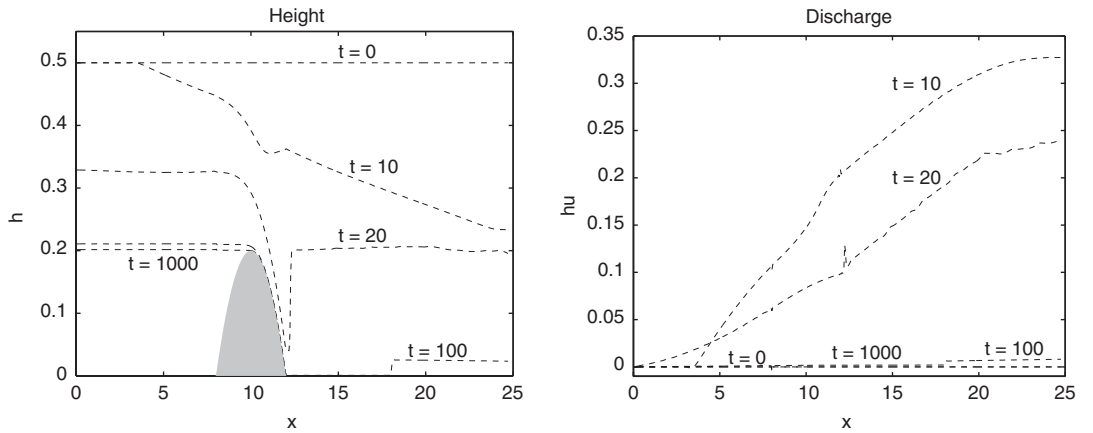


Figure 8. Water height and discharge plots for drain on non-flat bottom at different times.

can pass through zero, the characteristic speeds A_1 and A_2 in (30) are not valid any more. To overcome the zero speeds in the relaxation system (7), we perturb these characteristic speeds by 10^{-2} far from zero. The monotonicity of the scheme is preserved and no non-physical oscillations or extra numerical diffusion have been detected during the computations.

3.2. *Two-dimensional examples*

The two-dimensional relaxation system of shallow-water equations is constructed as (16) with $\mathbf{A}^2 = \mathrm{diag}\{A_1^2, A_2^2, A_3^2\}$ and $\mathbf{B}^2 = \mathrm{diag}\{B_1^2, B_2^2, B_3^2\}$. The selection of these parameters is made based on the eigenvalues $\{u, u \pm \sqrt{gh}\}$ and $\{v, v \pm \sqrt{gh}\}$ of the Jacobian matrices $\partial \mathbf{F}(\mathbf{U})/\partial \mathbf{U}$ and $\partial \mathbf{G}(\mathbf{U})/\partial \mathbf{U}$, as

$$A_1 = A_2 = A_3 = \max(\sup|u|, \sup|u - \sqrt{gh}|, \sup|u + \sqrt{gh}|)$$

$$B_1 = B_2 = B_3 = \max(\sup|v|, \sup|v - \sqrt{gh}|, \sup|v + \sqrt{gh}|)$$

A threshold of 10^{-2} is added to these characteristic speeds wherever they vanish. This perturbation is needed to treat dry cells in the computational domain. We perform the following test cases:

3.2.1. Circular dam-break problem. This example was first proposed in Reference [3]. The space domain is a 50 m long square with a cylindrical dam with radius 11 m and centred in the square. The initial water height is 10 m inside the dam and 1 m outside the dam and water is initially at rest. At $t = 0$, the cylindrical wall forming the dam collapses and time evolution of water level is computed. As in Reference [3], we discretize the domain uniformly in 50×50 gridpoints and the solution is displayed at $t = 0.69$ s. The contour plot of water height is shown in Figure 9. In this figure we have also included the surface plot for a better insight. As can be seen a bore has formed and the water drains from the deepest region as a rarefaction wave progresses outwards. The flow in that region becomes supercritical. The

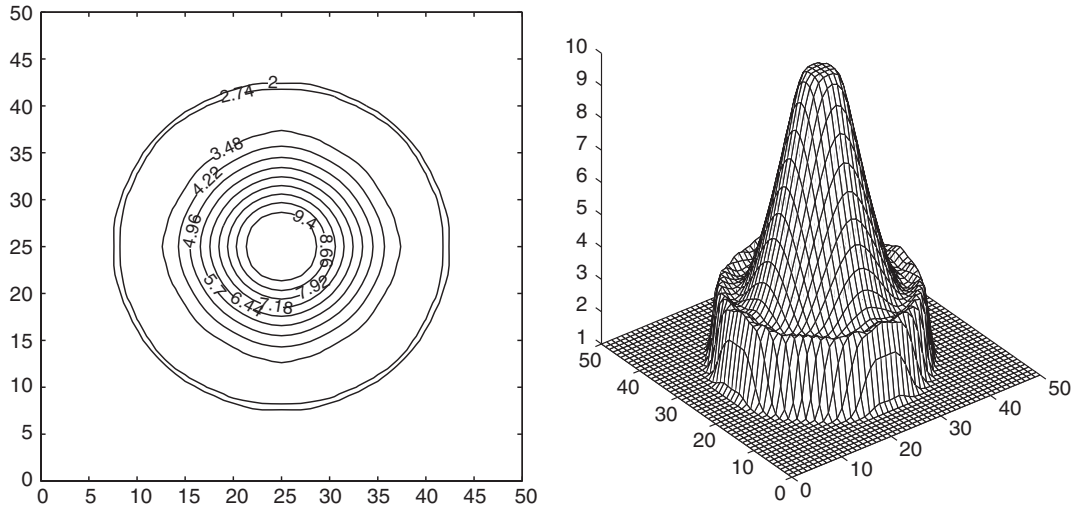


Figure 9. Contour and surface plots of water height for the circular dam-break problem on wet bed.

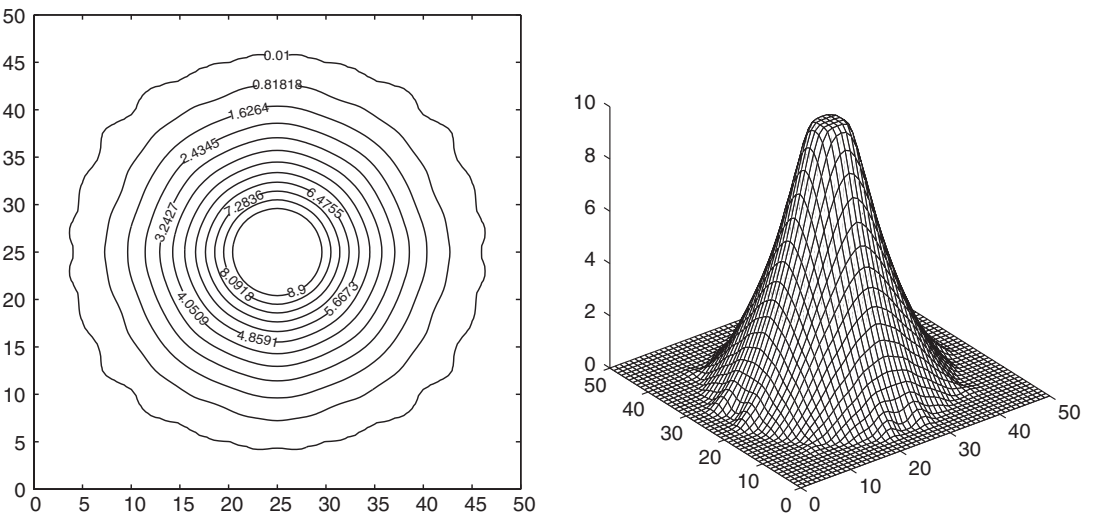
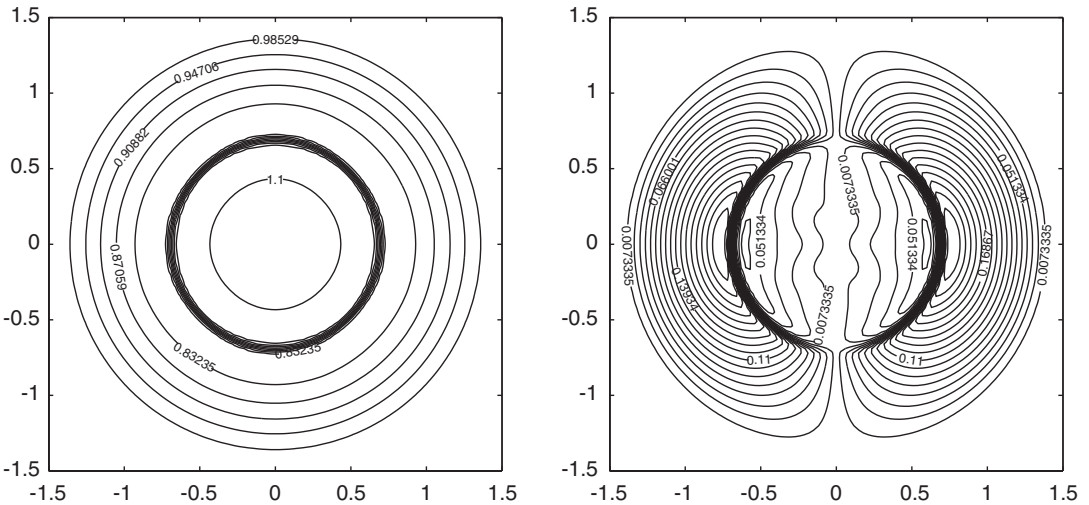


Figure 10. Contour and surface plots of water height for the circular dam-break problem on dry bed.

results show that the circular symmetry is preserved very well by our relaxation scheme. The results agree with Reference [3].

We now turn our attention to the presence of dry areas in the computational domain. In order to illustrate the performance of the relaxation scheme we set the downstream water depth to 0 m and we solve the problem using the same mesh as for the previous test. Figure 10 shows the results at $t = 0.69$ s. We see that no bore forms, instead a rarefaction wave extends


0.98529
0.94706
0.90882
0.87059
0.83235
0.79412
1.1
0.11
0.13924
0.16848
0.19771
0.22695
0.25619
0.28543
0.31467
0.34391
0.37314
0.40238
0.43162
0.46086
0.49009
0.51933
0.54857
0.57781
0.60705
0.63628
0.66552
0.69476
0.72400
0.75323
0.78247
0.81171
0.84095
0.87018
0.89942
0.92866
0.95790
0.98713
1.01637
1.04561
1.07485
1.10408
1.13332
1.16256
1.19180
1.22103
1.25027
1.27951
1.30875
1.33798
1.36722
1.39646
1.42570
1.45493
1.48417
1.51341
1.54265
1.57188
1.60112
1.63036
1.65960

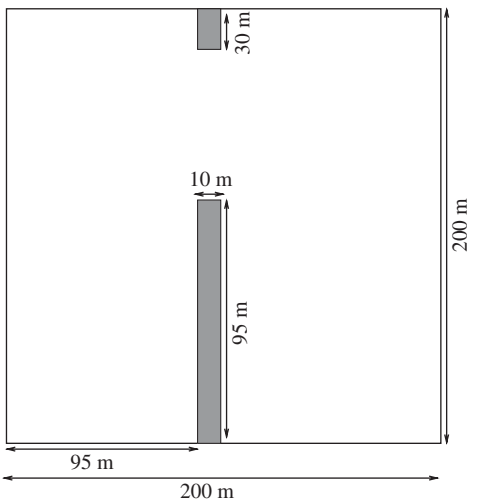


Figure 12. Geometry of the squared dam-break problem.

the computed solution remains, stable, monotone and highly accurate even on coarse grids without solving Riemann problems or requiring special front tracking procedures.

3.2.3. Flow over a smooth bump. This test problem is proposed and analysed in [28, 29] to check the property of preserving the initial steady state for numerical methods. It consists of Equation (2) defined in unit square domain with a symmetric bump localized at the centre of the domain and defined by

$$Z(x, y) = \max\left[0, 0.25 - 5\left(\left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2\right)\right]$$

The domain is closed with wall boundaries. Two situations are considered: (i) initial water level cover totally the bump and, (ii) initial water cover partially the bump. In both situations, the spatial domain is discretized in 50 cells in each direction, the CFL number is set to 0.5 and the solutions are computed for $t = 1$ min.

In the first test case, the initial conditions are

$$h + Z = 0.5\ \text{m}, \quad u = 0\ \text{m}^2/\text{s}, \quad v = 0\ \text{m}^2/\text{s}$$

Figure 14 represents the numerical results obtained by the relaxation scheme for the water depth. For the second test case, the initial conditions are,

$$h + Z = 0.1\ \text{m}, \quad u = 0\ \text{m}^2/\text{s}, \quad v = 0\ \text{m}^2/\text{s}$$

The computed results for this test case are shown in Figure 15. In both test cases, the relaxation scheme maintains correctly the steady flow structure without any kind of spurious oscillations over the hump.

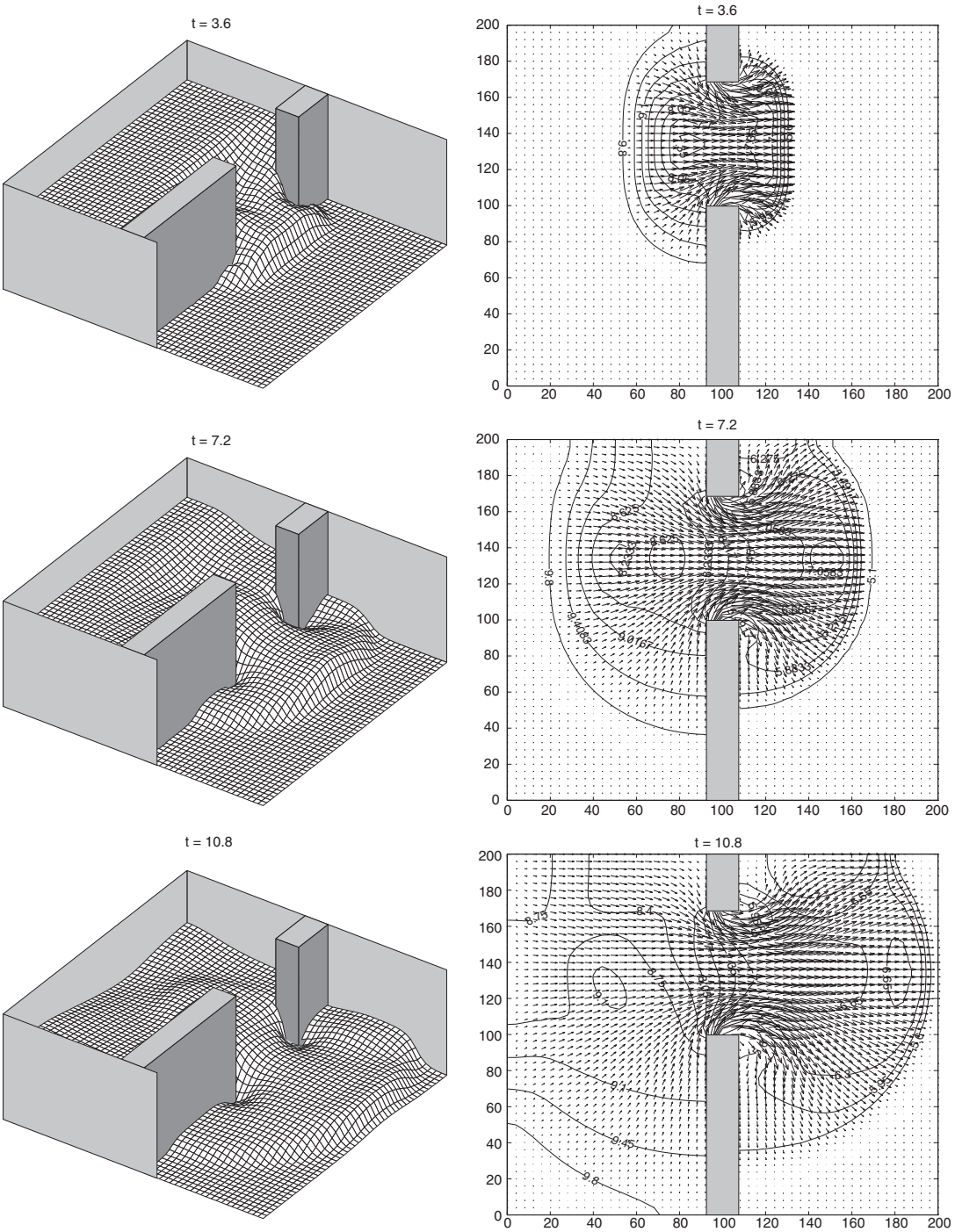


Figure 13. Water surface elevation and velocity field for squared dam-break problem.

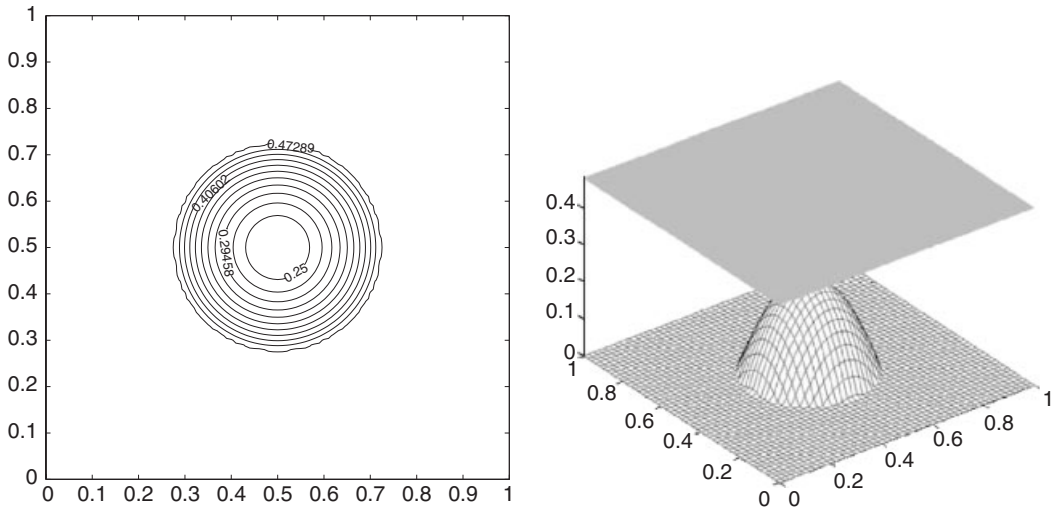


Figure 14. Contour and surface plots of water depth for flow over a totally covered bump.

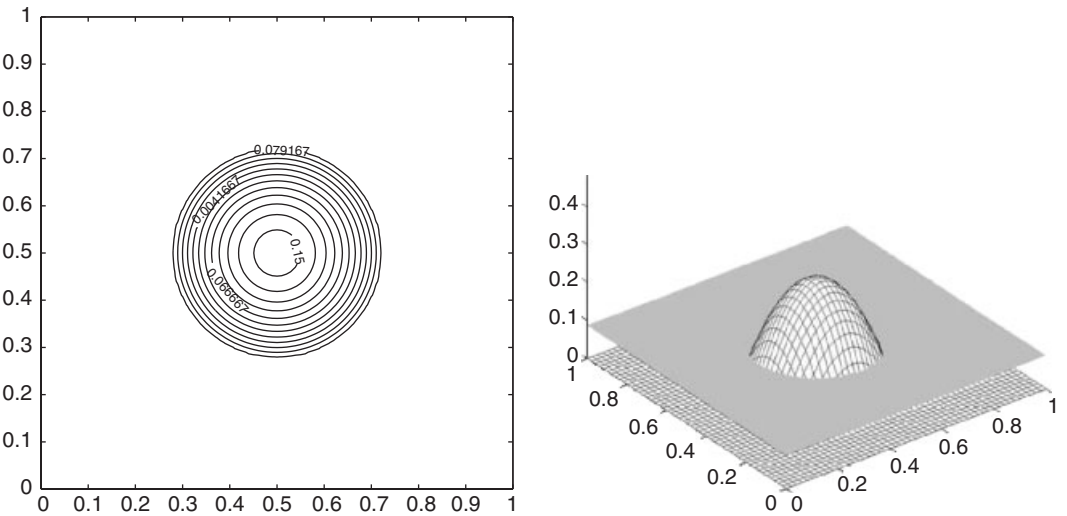


Figure 15. Contour and surface plots of water depth for flow over a partially covered bump.

3.2.4. Flow over a shaped bump. Our purpose in the following test problem is to examine the ability of our relaxation scheme to handle the two-dimensional shallow-water equations on non-flat bottom. In this academic test, we consider the Equation (2) in the rectangular channel $[0,2]\times[0,1]$ with an elliptical-shaped hump defined by [8],

$$Z(x,y)=0.8\exp(-5(x-0.9)^2-50(y-0.5)^2) \tag{33}$$

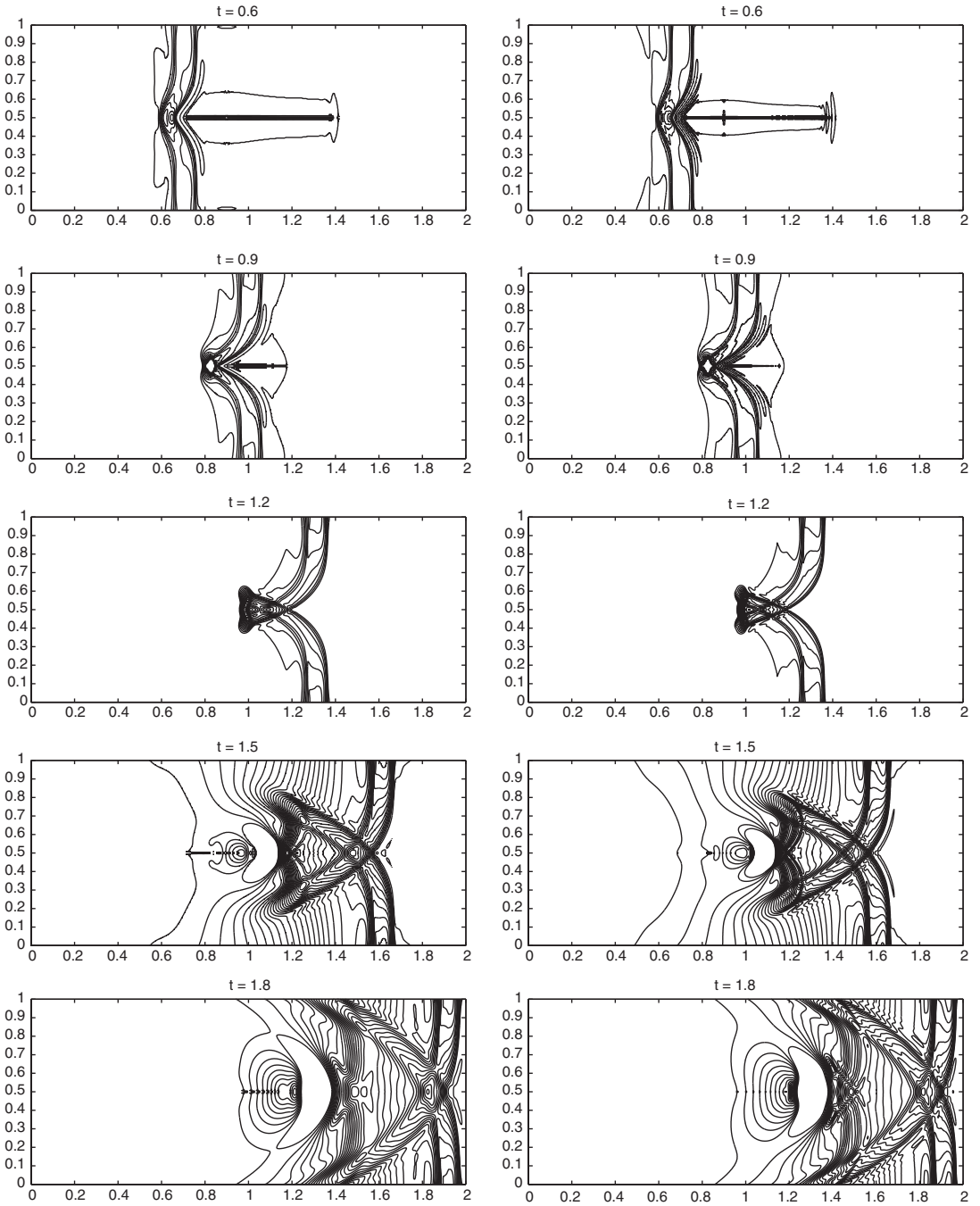


Figure 16. Contours of water surface for flow over a shaped hump using 200×100 gridpoints (left column) and 400×200 gridpoints (right column) at times 0.6, 0.9, 1.2, 1.5 and 1.8.

Note that hump (33) is centred at $(0.9, 0.5)$ with a minimum and maximum heights of 0.01 and 0.8, respectively. Initially, the flow is at rest and the water surface is flat with a small perturbation in the vertical slab $[0.05, 0.15]$ as

$$h(0,x,y) = \begin{cases} 1.01 - Z(x,y) & \text{if } 0.05 < x < 0.15 \\ 1 - Z(x,y) & \text{otherwise} \end{cases}$$

Outflow boundary conditions are imposed on all domain boundaries, and we use two meshes of 200×100 and 400×200 gridpoints for comparison. In our computation we set the dimensionless $g = 1$. In Figure 16, we display contour plots of the water distribution as it flows past the hump. Note that the speed of water flow is slower above the hump than elsewhere in the channel, producing a distortion of the initially flat distribution of the water. We can see that the small complex structures of the water flow being captured by our relaxation scheme.

4. CONCLUSIONS

Relaxation schemes of first- and second-order accuracy were introduced in Reference [10]. In this paper we have reconstructed high-order relaxation schemes by using WENO ideas and a class of TVD high-order Runge–Kutta time integration methods. We have further generalized and extended the relaxation schemes for the two-dimensional shallow-water equations. This procedure combines the attractive attributes of the two methods to yield a procedure for either flat or non-flat topography. The new method retains all the attractive features of central schemes such as neither Riemann solvers nor characteristic decomposition are needed. Furthermore, the scheme does not require either non-linear solution or special front tracking techniques.

The third-order relaxation method have been tested on systems of shallow-water equations in one and two space dimensions. The obtained results indicate good shock resolution with high accuracy in smooth regions and without any non-physical oscillations near the shock areas. The convergence to the correct steady-state solution has been clearly verified in flow at rest and drain on non-flat bottom. Although we have restricted our numerical computations to the frictionless shallow-water problems, the current relaxation scheme can be straightforwardly extended to more general shallow-water flows with bottom friction and Coriolis forces.

ACKNOWLEDGEMENT

The author thanks an anonymous reviewer for his helpful comments which subsequently improved the current manuscript, such as the inclusion of the flow at rest in test problems.

REFERENCES

1. Stoker JJ. *Water Waves*. Interscience Publishers, Inc.: New York, 1986.
2. Toro EF. *Shock-Capturing Methods for Free-Surface Shallow Flows*. Wiley: New York, 2001.
3. Alcrudo F, Navarro PG. A high resolution Godunov-type scheme in finite volumes for the 2D shallow water equations. *International Journal for Numerical Methods in Fluids* 1993; **16**:489–505.
4. Delis AI, Skeels CP. TVD schemes for open channel flow. *International Journal for Numerical Methods in Fluids* 1998; **26**:791–809.

5. Fennema RJ, Chaudhry MH. Explicit methods for 2D-transient free-surface flows. *Journal of Hydraulic Engineering* 1990; **116**:1013–1034.
6. Gallouët T, Herard JM, Seguin N. Some approximate Godunov schemes to compute shallow water equations with topography. *Computers and Fluids* 2003; **32**:479–513.
7. Levy D, Puppo G, Russo G. Central-Upwind schemes for the Saint–Venant system. *Mathematical Modelling and Numerical Analysis* 2002; **36**:397–425.
8. LeVeque Randall J. Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm. *Journal of Computational Physics* 1998; **146**:346–365.
9. LeVeque Randall J. *Numerical Methods for Conservation Laws* (2nd edn). Lectures in Mathematics. ETH: Zürich, 1992.
10. Jin S, Xin Z. The relaxation schemes for systems of conservation laws in arbitrary space dimensions. *Communications on Pure and Applied Mathematics* 1995; **48**:235–276.
11. Aregba-Driollet D, Natalini R. Convergence of relaxation scheme for conservation laws. *Applicable Analysis* 1996; **61**:163–193.
12. Chalabi A. Convergence of relaxation scheme for hyperbolic conservation laws with stiff source terms. *Mathematics of Computation* 1999; **68**:955–970.
13. Chen GQ, Levermore CD, Liu TP. Hyperbolic conservation laws with stiff relaxation terms and entropy. *Communications on Pure and Applied Mathematics* 1994; **47**:787–830.
14. Liu HL, Warnecke G. Convergence rates for relaxation schemes approximating conservation laws. *SIAM Journal on Numerical Analysis* 2000; **37**:1316–1337.
15. Klar A. Relaxation schemes for a Lattice–Boltzmann type discrete velocity model and numerical Navier–Stokes limit. *Journal of Computational Physics* 1999; **148**:1–17.
16. Natalini R. Convergence to equilibrium for relaxation approximations of conservation laws. *Communications on Pure and Applied Mathematics* 1996; **49**:795–823.
17. Xu WQ. Relaxation limit for piecewise smooth solutions to conservation laws. *Journal of Differential Equations* 2000; **162**:140–173.
18. Lattanzio C, Serre D. Convergence of a relaxation scheme for hyperbolic systems of conservation laws. *Numerische Mathematik* 2001; **88**:121–134.
19. Banda MK, Seaïd M. A class of the relaxation schemes for two-dimensional Euler systems of gas dynamics. *Lecture Notes in Computer Science* 2002; **2329**:930–939.
20. Fan H, Jin S, Teng Z. Zero reaction limit for hyperbolic conservation laws with source terms. *Journal of Differential Equations* 2000; **168**:270–294.
21. Jin S, Pareschi L, Toscani G. Uniformly accurate diffusive relaxation schemes for multiscale transport equations. *SIAM Journal on Numerical Analysis* 2000; **38**:913–936.
22. Delis AI, Katsaounis Th. Relaxation schemes for the shallow water equations. *International Journal for Numerical Methods in Fluids* 2003; **41**:695–719.
23. Kurganov A, Levy D. A third-order semi-discrete central scheme for conservation laws and convection–diffusion equations. *SIAM Journal on Scientific Computing* 2000; **22**:1461–1488.
24. Levy D, Puppo G, Russo G. Compact central WENO schemes for multidimensional conservation laws. *SIAM Journal on Scientific Computing* 2000; **22**:656–672.
25. Ascher U, Ruuth S, Spiteri R. Implicit–explicit Runge–Kutta methods for time-dependent partial differential equations. *Applied Numerical Mathematics* 1997; **25**:151–167.
26. Pareschi L, Russo G. Implicit–Explicit Runge–Kutta schemes for stiff systems of differential equations. *Recent Trends in Numerical Analysis* 2000; **3**:269–289.
27. Pareschi L, Russo G. Implicit–Explicit Runge–Kutta schemes and applications to hyperbolic systems with relaxation. *Journal of Scientific Computing*, in press.
28. Brufau P, Vázquez-Cendón ME, García-Navarro P. A numerical model for the flooding and drying of irregular domains. *International Journal for Numerical Methods in Fluids* 2002; **39**:247–275.
29. Brufau P, García-Navarro P. Unsteady free surface flow simulation over complex topography with a multidimensional upwind technique. *Journal of Computational Physics* 2003; **186**:503–526.
30. Banda MK, Seaïd M. Higher-order relaxation schemes for hyperbolic systems of conservation laws. Submitted.
31. Bermudez A, Vázquez-Cendón ME. Upwind methods for hyperbolic conservation laws with source terms. *Computers & Fluids* 1994; **23**:1049–1071.
32. Lax P, Liu X. Solution of two-dimensional Riemann problems of gas dynamics by positive schemes. *SIAM Journal on Scientific Computing* 1998; **19**:319–340.
33. Weinan E, Liu JG. Essentially compact schemes for unsteady viscous incompressible flows. *Journal of Computational Physics* 1996; **126**:122–138.

