ANZIAM J. **66** (2024), 181–200
doi:10.1017/S1446181124000191

ON WELL-POSED BOUNDARY CONDITIONS AND ENERGY STABLE FINITE-VOLUME METHOD FOR THE LINEAR SHALLOW WATER WAVE EQUATION

RUDI PRIHANDOKO✉[1], KENNETH DURU[1], STEPHEN ROBERTS[1] and CHRISTOPHER ZOPPOU[1]

(Received 28 February, 2024; accepted 9 July, 2024; first published online 22 October, 2024)

Abstract

We derive and analyse well-posed boundary conditions for the linear shallow water wave equation. The analysis is based on the energy method and it identifies the number, location and form of the boundary conditions so that the initial boundary value problem is well-posed. A finite-volume method is developed based on the summation-by-parts framework with the boundary conditions implemented weakly using penalties. Stability is proven by deriving a discrete energy estimate analogous to the continuous estimate. The continuous and discrete analysis covers all flow regimes. Numerical experiments are presented verifying the analysis.

2020 *Mathematics subject classification*: 65M12.

Keywords and phrases: numerical analysis, finite volume, boundary condition, energy method, shallow water wave equation.

1. Introduction

Numerical models that solve the shallow water wave equations (SWWEs) have become a common tool for modelling environmental problems. This system of nonlinear hyperbolic partial differential equations (PDEs) represent the conservation of mass and momentum of unsteady free surface flow subject to gravitational forces. The SWWEs assume that the fluid is inviscid, incompressible and the wavelength of the wave is much greater than its height. Typically, these waves are associated with flows caused, for example, by tsunamis, storm surges and floods in riverine systems. The SWWEs are

[1] Mathematical Science Institute, Australian National University, Canberra 2600, Australia;
e-mail: rudi.prihandoko@anu.edu.au, kenneth.duru@anu.edu.au, stephen.roberts@anu.edu.au, christopher.zoppou@anu.edu.au

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

also a fundamental component for predicting a range of aquatic processes, including sediment transport and the transport of pollutants. All these processes can have a significant impact on the environment, vulnerable communities and infrastructure. Therefore, making accurate predictions using the SWWEs is crucial for urban, rural and environmental planners.

For practical problems, the SWWEs have been solved numerically using finite-difference methods [9], finite-volume methods [14], discontinuous Galerkin method [13] and the method of characteristics [2]. Although, the SWWEs are in common use, a rigorous theoretical investigation of boundary conditions necessary for their solution is still an area of active research [4].

In this paper, we investigate well-posed boundary conditions for the linearized SWWEs using the energy method [5, 6] and develop a provably stable numerical method for the model. The SWWEs that we use are written in conservative form, where the mass h and momentum uh are conserved, and energy is lost through shocks. This is physically reasonable and validated by experimental data. We wish to extended this work to the nonlinear form of the equations in the future. Since, the nonlinear equations admit shocks, addressing the shock speed appropriately is necessary and, therefore, involve the conservative quantities not the primitive variables.

Following Ghader and Nordström [4], our analysis identifies the type, location and number of boundary conditions that are required to yield a well-posed initial boundary value problem (IBVP). More importantly, we formulate the boundary conditions so that they can be readily implemented in a stable manner for numerical approximations that obey the summation-by-parts (SBP) principle [7]. We demonstrate this by deriving a stable finite-volume method using the SBP framework and impose the boundary conditions weakly using the simultaneous approximation term (SAT) method [1]. This SBP-SAT approach enables us to prove that the numerical scheme satisfies the discrete counterparts of energy estimates required for well-posedness of the IBVP, resulting in a provably stable and conservative numerical scheme.

The continuous and discrete analysis covers all flow regimes, namely sub-critical, critical and super-critical flows. Numerical experiments are performed to verify the theoretical analysis of the continuous and discrete models.

In the following section, we derive stable boundary conditions for the linearised SWWE in the continuous level and in Section 3, in the discrete level. In Section 4, we verify the stability of the numerical model using numerical tests. We conclude that the SBP-SAT finite-volume scheme is stable for a variety of boundary conditions in Section 5.

2. Continuous analysis

The one-dimensional (1D) SWWEs are

$$\frac{\partial h}{\partial t} + \frac{\partial (uh)}{\partial x} = 0, \qquad \frac{\partial (uh)}{\partial t} + \frac{\partial (u^2 h + gh^2/2)}{\partial x} = 0, \tag{2.1}$$

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

where $x \in \mathbb{R}$ is a spatial variable, $t \geq 0$ is time, $h(x,t) > 0$ and $u(x,t)$ are the water depth and the depth averaged fluid velocity, respectively, and $g > 0$ is the gravitational acceleration.

To make our analysis tractable, we linearise the SWWEs by substituting $h = H + \widetilde{h}$ and $u = U + \widetilde{u}$ into (2.1), where $\widetilde{h}$ and $\widetilde{u}$ denote perturbations of the constant water depth $H > 0$ and fluid velocity U, respectively.

After simplifying, the linearised SWWEs are

$$\frac{\partial \tilde{h}}{\partial t} + U\frac{\partial \tilde{h}}{\partial x} + H\frac{\partial \tilde{u}}{\partial x} = 0, \qquad \frac{\partial \tilde{u}}{\partial t} + g\frac{\partial \tilde{h}}{\partial x} + U\frac{\partial \tilde{u}}{\partial x} = 0. \tag{2.2}$$

Introducing the unknown vector field $\mathbf{p} = \begin{bmatrix}\widetilde{h} & \widetilde{u}\end{bmatrix}^{\top}$, the linear equation (2.2) can be rewritten in a more compact form as

$$\frac{\partial \mathbf{p}}{\partial t} = \mathcal{D}\mathbf{p}, \quad \mathcal{D} = -\mathcal{M}\frac{\partial}{\partial x}, \quad \mathbf{p} = \begin{bmatrix}\widetilde{h} & \widetilde{u}\end{bmatrix}^{\top}, \quad \mathcal{M} = \begin{bmatrix} U & H \\ g & U \end{bmatrix}. \tag{2.3}$$

To avoid inconsistencies in the units of the matrix entries, we rescale the variables as follows:

$$\mathbf{q} = \begin{bmatrix} \widetilde{\widetilde{h}} \\ \widetilde{\widetilde{u}} \end{bmatrix} = \begin{bmatrix} \widetilde{h}/H \\ \widetilde{u}/c \end{bmatrix} = W\mathbf{p}, \quad W = \begin{bmatrix} 1/H & 0 \\ 0 & 1/c \end{bmatrix}. \tag{2.4}$$

First, we multiply (2.3) by W, and introduce the dimensionless variable $\mathbf{q}$ and the symmetric matrix $M = W\mathcal{M}W^{-1}$. After simplifying and dropping the double-tilde, for simplicity, the dimensionless equation can be written as

$$\frac{\partial \mathbf{q}}{\partial t} = D\mathbf{q}, \quad D = -M\frac{\partial}{\partial x}, \quad \mathbf{q} = \begin{bmatrix} h & u \end{bmatrix}^{\top}, \quad M = \begin{bmatrix} U & c \\ c & U \end{bmatrix}. \tag{2.5}$$

Note that in (2.5), the constant coefficient matrix M is symmetric. This will simplify the following analysis.

We will consider (2.5) in a bounded domain, and augment it with initial and boundary conditions. Let our domain be $\Omega = [0, 1]$ and $\Gamma = \{0, 1\}$ be the boundary points. We consider the IBVP

$$\frac{\partial \mathbf{q}}{\partial t} = D\mathbf{q}, \quad x \in \Omega, \; t \geq 0, \tag{2.6a}$$

$$\mathbf{q}(x, 0) = \mathbf{f}(x), \quad x \in \Omega, \tag{2.6b}$$

$$\mathcal{B}\mathbf{q} = \mathbf{b}(t), \quad x \in \Gamma, \; t \geq 0, \tag{2.6c}$$

where $\mathcal{B}$ is a linear boundary operator, $\mathbf{b}$ is the boundary data and $\mathbf{f} \in L^2(\Omega)$ is the initial condition. One objective of this study is to investigate the choice of the boundary operator $\mathcal{B}$ which ensures that the IBVP (2.6) is well-posed. To simplify the coming analysis, we will consider zero boundary data $\mathbf{b} = 0$, but the results can be extended to nontrivial boundary data $\mathbf{b} \neq 0$. Furthermore, numerical experiments performed later in this paper confirm that the analysis is valid for nonzero boundary data.

Let $\mathbf{p}$ and $\mathbf{q}$ be real-valued vector functions, and define the standard $L^2(\Omega)$ scalar product and the norm

$$(\mathbf{p}, \mathbf{q}) = \int_\Omega \mathbf{p}^\top \mathbf{q}\, dx, \quad \|\mathbf{q}\|^2 = (\mathbf{q}, \mathbf{q}) > 0, \tag{2.7}$$

for all nonzero $\mathbf{q} \in \mathbb{R}^2$.

DEFINITION 2.1. The IBVP (2.6) is *well-posed* if a unique solution $\mathbf{q}$ satisfies

$$\|\mathbf{q}(\cdot, t)\| \le \kappa e^{\nu t} \|\mathbf{f}\|, \quad \|\mathbf{f}\| < \infty$$

for some constants $\kappa > 0$ and $\nu \in \mathbb{R}$ independent of $\mathbf{f}$.

The well-posedness of the IBVP (2.6) can be related to the boundedness of the differential operator D. We introduce the function space

$$\mathbb{V} = \{\mathbf{p} |\ \mathbf{p}(x) \in \mathbb{R}^2,\ \|\mathbf{p}\| < \infty,\ 0 \le x \le 1,\ \{\mathcal{B}\mathbf{p} = 0,\ x \in \Gamma\}\}.$$

The following two definitions are useful.

DEFINITION 2.2. The operator D is said to be *semi-bounded* in the function space $\mathbb{V}$ if it satisfies

$$(\mathbf{q}, D\mathbf{q}) \le \nu \|\mathbf{q}\|^2, \quad \nu \in \mathbb{R}.$$

DEFINITION 2.3. The differential operator D is *maximally semi-bounded* if it is semi-bounded in the function space $\mathbb{V}$, but not semi-bounded in any space with fewer boundary conditions.

It is well known that the maximally semi-boundedness of differential operator D is a necessary and sufficient condition for the well-posedness of the IBVP (2.6) [6].

Thus, to ensure that the IBVP (2.6) is well-posed, we need: (a) the differential operator D to be semi-bounded; and (b) the minimal number of boundary conditions such that D is maximally semi-bounded.

To begin with, we will show that the differential operator D is semi-bounded in $L_2(\Omega)$.

LEMMA 2.4. *Consider the differential operator D with the constant coefficients and symmetric matrix M given in* (2.5) *and the L_2 scalar product defined in* (2.7)*, where* $\boldsymbol{q}^\top \boldsymbol{q} > 0$ *for all nonzero* $\boldsymbol{q} \in \mathbb{R}^2$*. If* $(\boldsymbol{q}^\top M\boldsymbol{q})|_0^1 \ge 0$*, then D is semi-bounded.*

PROOF. We consider $(\mathbf{q}, D\mathbf{q})$ and use integration-by-parts, then

$$(\mathbf{q}, D\mathbf{q}) = -\int_\Omega \mathbf{q}^\top M \frac{\partial \mathbf{q}}{\partial x}\, dx = -\frac{1}{2}\int_\Omega \frac{\partial}{\partial x}(\mathbf{q}^\top M\mathbf{q})\, dx = -\frac{1}{2}(\mathbf{q}^\top M\mathbf{q})|_0^1.$$

Thus, if the boundary term $(\mathbf{q}^\top M\mathbf{q})|_0^1 \ge 0$, then $(\mathbf{q}, \mathbf{Dq}) \le 0$. In particular, the upper bound $(\mathbf{q}, \mathbf{Dq}) = 0$ satisfies Definition 2.2 with $\nu = 0$. □

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

The next step will be to derive boundary operators $\{\mathcal{B}\mathbf{p} = 0,\ x \in \Gamma\}$ with minimal number of boundary conditions such that the boundary term is never negative, $(\mathbf{q}^\top M\mathbf{q})|_0^1 \geq 0$.

Noting that the norm is related to the dimensionless mechanical energy $\tilde{E}$, that is,

$$\frac{1}{2}\|\mathbf{q}\|^2 = E := \int_\Omega \frac{1}{2}(h^2 + u^2)\,dx > 0 \quad \text{for all } \mathbf{q} \in \mathbb{R}^2\backslash\{\mathbf{0}\}.$$

The mechanical energy in the physical units can be recovered through the scaling $\widetilde{E} = c^2HE$.

We introduce the boundary term

$$\mathrm{BT} := -(\mathbf{q}^\top M\mathbf{q})|_0^1.$$

By using the eigen-decomposition $M = S\Lambda S^T$ given by

$$S = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}, \quad \Lambda = \begin{bmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{bmatrix}, \quad \lambda_1 = U + c, \quad \lambda_2 = U - c, \tag{2.8}$$

with the linear transformation

$$\begin{bmatrix} w_1 \\ w_2 \end{bmatrix} = S^\top\mathbf{q} = \frac{1}{\sqrt{2}}\begin{bmatrix} h + u \\ h - u \end{bmatrix}, \tag{2.9}$$

the boundary term can be re-written as

$$\mathrm{BT} = (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_{x=0} - (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_{x=1}. \tag{2.10}$$

The number of boundary conditions will depend on the signs of the eigenvalues λ_1, λ_2, which in turn depend on the magnitude of the flow U relative to the characteristic wave speed c, and are determined by the Froude number $Fr = |U|/c$. If $0 \leq Fr < 1$, then $\lambda_1 > 0$ and $\lambda_2 < 0$ for any U. In this case, we have one boundary condition each in the inflow and outflow. For $Fr > 1$, the sign of the eigenvalues λ_1 and λ_2 will take the sign of U. In this case, we have either two boundary conditions on the left if $U > 0$ (inflow at $x = 0$) or two boundary conditions on the right if $U < 0$ (inflow at $x = 1$). The case where $Fr = 1$, we have $\lambda_1 > 0$ and $\lambda_2 = 0$ if $U > 0$, and $\lambda_1 = 0$ and $\lambda_2 < 0$ if $U < 0$. That is, we only have one boundary condition at the inflow at $x = 0$ if $U > 0$ or at $x = 1$ if $U < 0$.

Sub-critical flow. The flow is sub-critical when $\mathrm{Fr} < 1$, which implies $\lambda_1 > 0$ and $\lambda_2 < 0$. We need one boundary condition at $x = 0$ and another boundary condition at $x = 1$. Therefore, for the sub-critical flow regime, we always need an inflow boundary condition and an outflow boundary condition for any U. We formulate the boundary conditions by

$$\{\mathcal{B}\mathbf{p} = \mathbf{b},\ x \in \Gamma\} \equiv \{w_1 - \gamma_0 w_2 = b_1(t),\ x = 0;\ w_2 - \gamma_N w_1 = b_2(t),\ x = 1\}, \tag{2.11}$$

where $\gamma_0, \gamma_N \in \mathbb{R}$ are boundary reflection coefficients. The following lemma constrains the parameters γ_0, γ_N.

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

LEMMA 2.5. *Consider the boundary term* BT *defined in* (2.10) *and the boundary condition* (2.11) *with* $\boldsymbol{b} = 0$ *for sub-critical flows* $\mathrm{Fr} < 1$ *with* $\lambda_1 > 0$ *and* $\lambda_2 < 0$. *If* $0 \leq \gamma_0^2 \leq -\lambda_2/\lambda_1$ *and* $0 \leq \gamma_N^2 \leq -\lambda_1/\lambda_2$, *then the boundary term is never positive, that is,* $\mathrm{BT} \leq 0$.

PROOF. Let $w_1 = \gamma_0 w_2$ at $x = 0$ and $w_2 = \gamma_N w_1$ at $x = 1$, and consider

$$(\lambda_1 w_1^2 + \lambda_2 w_2^2)|_0 - (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_1 = w_2^2(\lambda_1 \gamma_0^2 + \lambda_2)|_0 - w_1^2(\lambda_1 + \lambda_2 \gamma_N^2)|_1.$$

Thus, if $0 \leq \gamma_0^2 \leq -\lambda_2/\lambda_1$ and $0 \leq \gamma_N^2 \leq -\lambda_1/\lambda_2$, then $\lambda_1\gamma_0^2 + \lambda_2 \leq 0$ and $\lambda_1 + \lambda_2\gamma_N^2 \geq 0$, and then

$$\mathrm{BT} = w_2^2(\lambda_1 \gamma_0^2 + \lambda_2)|_0 - w_1^2(\lambda_1 + \lambda_2 \gamma_N^2)|_1 \leq 0. \qquad \square$$

Super-critical flow. When $\mathrm{Fr} > 1$, the flow is super-critical, then λ_1 and λ_2 both take the sign of the average flow velocity U. That is, if $U > 0$, then $\lambda_1 > 0$, $\lambda_2 > 0$, and if $U < 0$, then $\lambda_1 < 0$, $\lambda_2 < 0$. Thus, when $U > 0$ and $\mathrm{Fr} > 1$, we need two boundary conditions at $x = 0$, and no boundary condition at $x = 1$. Similarly, when $U < 0$ and $\mathrm{Fr} > 1$, we need two boundary conditions at $x = 1$ and no boundary conditions at $x = 0$. Therefore, for super-critical flows, there are no outflow boundary conditions for any U. We formulate the boundary conditions by

$$\{\mathcal{B}\mathbf{q} = \mathbf{b},\ x \in \Gamma\} \equiv \{w_1 = b_1(t),\ w_2 = b_2(t),\ x = 0 \text{ if } U > 0 \text{ and } \mathrm{Fr} > 1\}, \tag{2.12a}$$

$$\{\mathcal{B}\mathbf{q} = \mathbf{b},\ x \in \Gamma\} \equiv \{w_1 = b_1(t),\ w_2 = b_2(t),\ x = 1 \text{ if } U < 0 \text{ and } \mathrm{Fr} > 1\}. \tag{2.12b}$$

LEMMA 2.6. *Consider the boundary term* BT *defined in* (2.10) *and the boundary condition* (2.12) *with* $\boldsymbol{b} = 0$ *for super-critical flows* $\mathrm{Fr} > 1$, *we have* $\mathrm{BT} \leq 0$.

PROOF. Let $U > 0$ with $\lambda_1 > 0$, $\lambda_2 > 0$ if $w_1 = 0,\ w_2 = 0$, at $x = 0$, then

$$\mathrm{BT} = (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_0 - (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_1 = -\tfrac{1}{2}(\lambda_1 w_1^2 + \lambda_2 w_2^2)|_1 \leq 0.$$

If $U < 0$ with $\lambda_1 < 0$, $\lambda_2 < 0$ and $w_1 = 0,\ w_2 = 0$, at $x = 1$, then

$$\mathrm{BT} = (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_0 - (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_1 = \tfrac{1}{2}(\lambda_1 w^2 + \lambda_2 w_2^2)|_0 \leq 0. \qquad \square$$

Critical flow. The flow is critical when $\mathrm{Fr} = 1$. Note that this case is degenerate, since there is only one nonzero eigenvalue, that is, $U > 0$ implies $\lambda_1 > 0$, $\lambda_2 = 0$ and $U < 0$ implies $\lambda_1 = 0$, $\lambda_2 < 0$. However, it can also be treated by prescribing only one boundary condition for the system. The location of the boundary condition will be determined by the sign of U, similar to the super-critical flow regime. We prescribe the boundary conditions by

$$\{\mathcal{B}\mathbf{q} = \mathbf{b},\ x \in \Gamma\} \equiv \{w_1 = b_1(t),\ x = 0 \text{ if } U > 0 \text{ and } \mathrm{Fr} = 1\}, \tag{2.13a}$$

$$\{\mathcal{B}\mathbf{q} = 0,\ x \in \Gamma\} \equiv \{w_2 = b_2(t),\ x = 1 \text{ if } U < 0 \text{ and } \mathrm{Fr} = 1\}. \tag{2.13b}$$

LEMMA 2.7. *Consider the boundary term* BT *defined in* (2.10) *and the boundary condition* (2.13) *with* $\boldsymbol{b} = 0$ *for critical flows* $U^2 = gH$, *we have* BT ≤ 0.

PROOF. Let $U > 0$ with $\lambda_1 > 0$, $\lambda_2 = 0$ if $w_1 = 0$, at $x = 0$,

$$\mathrm{BT} = (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_0 - (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_1 = -\tfrac{1}{2}\lambda_1 w_1^2|_1 \le 0.$$

If $U < 0$ with $\lambda_1 = 0$, $\lambda_2 < 0$ and $w_2 = 0$, at $x = 1$, then also

$$\mathrm{BT} = (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_0 - (\lambda_1 w_1^2 + \lambda_2 w_2^2)|_1 = \tfrac{1}{2}\lambda_2 w_2^2|_0 \le 0.$$

The proof is complete. □

We will conclude this section with the theorem that proves the well-posedness of the IBVP (2.6).

THEOREM 2.8. *Consider the IBVP* (2.6) *where the boundary operator* $\mathcal{B}\boldsymbol{q} = 0$ *is defined by* (2.11) *with* $\gamma_0^2 \le -\lambda_2/\lambda_1$ *and* $\gamma_N^2 \le -\lambda_1/\lambda_2$ *for sub-critical flows,* Fr < 1*; by* (2.12) *for the super-critical flow regime,* Fr > 1*; and by* (2.13) *for critical flows,* Fr $= 1$*; we have the energy estimate*

$$\frac{1}{2}\frac{d}{dt}\|\boldsymbol{q}\|_W^2 = \mathrm{BT} \le 0. \tag{2.14}$$

PROOF. We use the energy method, that is, from the left, we multiply (2.6a) with $\mathbf{q}^\top W$ and integrate over the domain. As above, integration-by-parts gives

$$\frac{1}{2}\frac{d}{dt}\|\mathbf{q}\|_W^2 = \left(\mathbf{q}, \frac{\partial \mathbf{q}}{\partial t}\right)_W = (\mathbf{q}, \mathbf{D}\mathbf{q})_W = \mathrm{BT}.$$

Using Lemmas 2.5–2.7 for each flow regime gives BT ≤ 0. Then the proof is complete. □

This energy estimate (2.14) is what a stable numerical method should emulate.

3. Numerical scheme

We derive a stable finite-volume method for the IBVP (2.6) encapsulated in the SBP framework. Numerical stability is proved by deriving discrete energy estimates analogous to Theorem 2.8.

3.1. The finite-volume method To begin, the domain, $\Omega = [0, 1]$, is subdivided into $N + 1$ computational nodes having $x_i = x_{i-1} + \Delta x_i$ for $i = 1, 2, \ldots N$, with $x_0 = 0$, $\Delta x_i > 0$ and $\sum_{i=1}^{N} \Delta x_i = 1$. We consider the control cell $I_i = [x_{i-1/2}, x_{i+1/2}]$ for each interior node $1 \le i \le N - 1$, and for the boundary nodes $\{x_0, x_N\}$, the control cells are $I_0 = [x_0, x_{1/2}]$ and $I_N = [x_{N-1/2}, x_N]$, see Figure 1. Note that $|I_i| = \Delta x_i/2 + \Delta x_{i+1}/2$ for the interior nodes $1 \le i \le N - 1$, and for the boundary nodes $i \in \{0, N\}$, we have

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

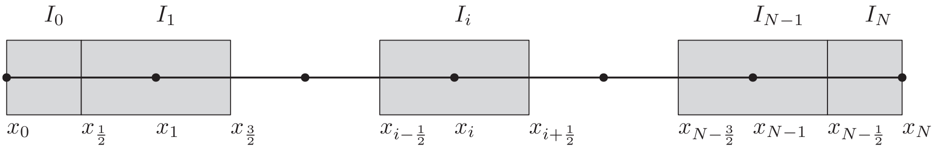


FIGURE 1. Finite-volume nodes x_i and control cells I_i.

$|I_0| = \Delta x_1/2$ and $|I_N| = \Delta x_N/2$. The control cells I_i are connected and do not overlap, and $\sum_{i=0}^{N} |I_i| = \sum_{i=1}^{N} \Delta x_i = 1$.

Consider the integral form of (2.2) over the control cells I_i:

$$\frac{d}{dt}\int_{I_0} \mathbf{p}(x,t)\,dx + \mathcal{M}\mathbf{p}(x_{1/2},t) - \mathcal{M}\mathbf{p}(x_0,t) = 0,$$
$$\frac{d}{dt}\int_{I_i} \mathbf{p}(x,t)\,dx + \mathcal{M}\mathbf{p}(x_{i+1/2},t) - \mathcal{M}\mathbf{p}(x_{i-1/2},t) = 0, \quad 1 \le i \le N-1,$$
$$\frac{d}{dt}\int_{I_N} \mathbf{p}(x,t)\,dx + \mathcal{M}\mathbf{p}(x_N,t) - \mathcal{M}\mathbf{p}(x_{N-1/2},t) = 0$$

with $\mathcal{M} = \begin{bmatrix} U & H \\ g & U \end{bmatrix}$ and $\mathbf{p} = \begin{bmatrix} \tilde{h} & \tilde{u} \end{bmatrix}^{\top}$.

We introduce the cell-average

$$\bar{\mathbf{p}}_i = \frac{1}{|I_i|}\int_{I_i} \mathbf{p}(x,t)\,dx,$$

and approximate the PDE flux $\mathcal{M}\mathbf{p}$ with the local Lax–Friedrich flux

$$\mathcal{M}\mathbf{p}(x_{i+1/2},t) \approx \frac{\mathcal{M}\bar{\mathbf{p}}_{i+1} + \mathcal{M}\bar{\mathbf{p}}_i}{2} - \frac{\alpha}{2}(\bar{\mathbf{p}}_{i+1} - \bar{\mathbf{p}}_i), \quad \alpha \ge 0, \tag{3.1}$$

and

$$\mathcal{M}\mathbf{p}(x_0,t) \approx \mathcal{M}\bar{\mathbf{p}}_0, \quad \mathcal{M}\mathbf{p}(x_N,t) \approx \mathcal{M}\bar{\mathbf{p}}_N.$$

The evolution of the cell-average is governed by the semi-discrete system

$$|I_0|\frac{d\bar{\mathbf{p}}_0}{dt} + \mathcal{M}\frac{\bar{\mathbf{p}}_1 - \bar{\mathbf{p}}_0}{2} - \frac{\alpha}{2}(\bar{\mathbf{p}}_1 - \bar{\mathbf{p}}_0) = 0, \tag{3.2a}$$
$$|I_i|\frac{d\bar{\mathbf{p}}_i}{dt} + \mathcal{M}\frac{\bar{\mathbf{p}}_{i+1} - \bar{\mathbf{p}}_{i-1}}{2} - \frac{\alpha}{2}(\bar{\mathbf{p}}_{i+1} - 2\bar{\mathbf{p}}_i + \bar{\mathbf{p}}_{i-1}) = 0,\ 1 \le i \le N-1, \tag{3.2b}$$
$$|I_N|\frac{d\bar{\mathbf{p}}_N}{dt} + \mathcal{M}\frac{\bar{\mathbf{p}}_N - \bar{\mathbf{p}}_{N-1}}{2} - \frac{\alpha}{2}(\bar{\mathbf{p}}_{N-1} - \bar{\mathbf{p}}_N) = 0. \tag{3.2c}$$

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

Introducing the discrete solution vector $\bar{\mathbf{p}} = [\bar{\mathbf{p}}_0, \bar{\mathbf{p}}_1, \ldots, \bar{\mathbf{p}}_N]^\top$ and rewriting (3.2) in a more compact form,

$$(I \otimes P)\frac{d\bar{\mathbf{p}}}{dt} + (\mathcal{M} \otimes Q)\bar{\mathbf{p}} - \frac{\alpha}{2}(I \otimes A)\bar{\mathbf{p}} = 0, \tag{3.3}$$

where $\otimes$ denotes the Kronecker product and

$$Q = \begin{pmatrix} -\frac{1}{2} & \frac{1}{2} & 0 & \cdots & 0 & 0 & 0 \\ -\frac{1}{2} & 0 & \frac{1}{2} & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & -\frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & \cdots & 0 & -\frac{1}{2} & \frac{1}{2} \end{pmatrix}, \quad A = \begin{pmatrix} -1 & 1 & 0 & \cdots & 0 & 0 & 0 \\ 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix},$$

and $P = \text{diag}([|I_0|, |I_1|, \ldots, |I_N|])$. The matrix Q is related to the spatial derivative operator, A is a numerical dissipation operator and $\alpha \geq 0$ controls the amount of numerical dissipation applied. Note that A is symmetric and negative semi-definite, that is, $A = A^\top$ and $\mathbf{p}^\top A\mathbf{p} \leq 0$ for all $\mathbf{p} \in \mathbb{R}^{N+1}$. The important stability property of the semi-discrete approximation (3.3) is that the associated discrete derivative operator satisfies the SBP property. To see this, we rewrite (3.3) as

$$\frac{d\bar{\mathbf{p}}}{dt} + (\mathcal{M} \otimes D_x)\bar{\mathbf{p}} - \frac{\alpha}{2}(I \otimes P^{-1}A)\bar{\mathbf{p}} = 0,$$

where I is the 2×2 identity matrix and

$$D_x = P^{-1}Q, \quad Q + Q^\top = \text{diag}([-1, 0, \ldots, 0, 1]). \tag{3.4}$$

Equation (3.4) is the so-called SBP property [5, 7] for the first derivative d/dx, which can be useful in proving numerical stability of the discrete approximation (3.3). Note that we have not enforced any boundary condition yet, the boundary condition (2.6c) will be implemented weakly using penalties.

3.2. Numerical boundary treatment and stability In this section, we will implement the boundary conditions and prove numerical stability. The boundary conditions are implemented using the SAT method, similar terms used as in [1]; by appending the boundary operators (2.11)–(2.13) to the right-hand side of (3.3) with penalty weights,

$$(I \otimes P)\frac{d\bar{\mathbf{p}}}{dt} + (\mathcal{M} \otimes Q)\bar{\mathbf{p}} - \frac{\alpha}{2}(I \otimes A)\bar{\mathbf{p}} = \text{SAT}. \tag{3.5}$$

With $\mathbf{e}_0 = [1, 0, \ldots 0]^T$ and $\mathbf{e}_N = [0, 0, \ldots 1]^T$, the SAT for sub-critical flow is

$$\text{SAT} = -\frac{1}{2}(W^{-1}S \otimes \mathbf{I})\begin{bmatrix} \tau_{01}\mathbf{e}_0(\bar{w}_1 - \gamma_0\bar{w}_2 - b_1(t)) + \tau_{N1}\mathbf{e}_N(\bar{w}_2 - \gamma_N\bar{w}_1 - b_2(t)) \\ \tau_{02}\mathbf{e}_0(\bar{w}_1 - \gamma_0\bar{w}_2 - b_1(t)) + \tau_{N2}\mathbf{e}_N(\bar{w}_2 - \gamma_N\bar{w}_1 - b_2(t)) \end{bmatrix}, \tag{3.6}$$

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

and for critical/super-critical flow regimes,

$$\mathrm{SAT} = -\frac{1}{2}(W^{-1}S \otimes \mathbf{I})\begin{bmatrix}\tau_{01}\mathbf{e}_0(\bar{w}_1 - b_1(t))\\ \tau_{02}\mathbf{e}_0(\bar{w}_2 - b_2(t))\end{bmatrix}, \quad U > 0, \tag{3.7a}$$

$$\mathrm{SAT} = -\frac{1}{2}(W^{-1}S \otimes \mathbf{I})\begin{bmatrix}\tau_{N1}\mathbf{e}_N(\bar{w}_1 - b_1(t))\\ \tau_{N2}\mathbf{e}_N(\bar{w}_2 - b_2(t))\end{bmatrix}, \quad U < 0. \tag{3.7b}$$

Here, S is the orthonormal eigenvector matrix given in (2.8) and W is the diagonal weight matrix given in (2.4). The coefficients τ_{01}, τ_{02}, τ_{N1}, τ_{N2} are real penalty parameters to be determined by requiring stability. Note that (3.5) is a consistent semi-discrete approximation of the IBVP (2.6) for all nontrivial choices of the penalty parameters. The semi-discrete approximation (3.5), given that the discrete derivative operator satisfies the SBP property (3.4), is often referred to as the SBP-SAT scheme [3, 8]. We introduce the discrete weighted L_2-norm

$$||\bar{\mathbf{q}}||^2 := \bar{\mathbf{q}}^T(\mathbf{I} \otimes P)\bar{\mathbf{q}} \geq 0$$

for some weighted matrix $\mathcal{W}$. The semi-discrete approximation (3.5) is stable if the discrete energy norm $||\bar{\mathbf{q}}||^2$ is nonincreasing with time. We will now prove the stability of the semi-discrete approximation (3.5) for sub-critical flows.

THEOREM 3.1. *Consider the semi-discrete finite-volume approximation* (3.5) *with the SAT* (3.6) *and* $\boldsymbol{b} = 0$ *for sub-critical flow regimes, where* $\lambda_1 > 0$, $\lambda_2 < 0$ *and* $\gamma_0^2 \leq -\lambda_2/\lambda_1$, $\gamma_N^2 \leq -\lambda_1/\lambda_2$. *If the penalty parameters are chosen such that*

$$\tau_{01} = \lambda_1, \quad \tau_{02} = \gamma_0\lambda_1; \quad \tau_{N2} = -\lambda_2, \quad \tau_{N1} = -\gamma_N\lambda_2,$$

then

$$\frac{1}{2}\frac{d}{dt}||\bar{\boldsymbol{q}}||^2 \leq 0 \quad \textit{for all } t \geq 0.$$

PROOF. We use the energy method, that is, from the left, we multiply (3.5) with $(W \otimes \mathbf{I})$, and using identity $\bar{\mathbf{q}} = W\bar{\mathbf{p}}$ and $M = W\mathcal{M}W^{-1}$,

$$(I \otimes P)\frac{d\bar{\mathbf{q}}}{dt} + (M \otimes Q)\bar{\mathbf{q}} - \frac{\alpha}{2}(I \otimes A)\bar{\mathbf{q}} = (W \otimes \mathbf{I})\mathrm{SAT}.$$

We multiply this equation with $\mathbf{q}^\top$, add its transpose and then simplify further, then:

$$\frac{1}{2}\frac{d}{dt}||\bar{\mathbf{q}}||^2 + \frac{1}{2}\bar{\mathbf{q}}^T(M \otimes (Q + Q^T))\bar{\mathbf{q}} - \frac{\alpha}{2}\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} = \bar{\mathbf{q}}^T(W \otimes \mathbf{I})\mathrm{SAT}.$$

Using the SBP property (3.4) and the eigen-decomposition of M,

$$\frac{1}{2}\frac{d}{dt}||\bar{\mathbf{q}}||^2 - \frac{\alpha}{2}\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} = \frac{1}{2}\mathrm{BT}_{\mathrm{num}},$$

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

where

$$\begin{aligned}\mathrm{BT}_{\mathrm{num}} &= (\lambda_1\bar{w}_1^2 + \lambda_2\bar{w}_2^2 - (\tau_{01}\bar{w}_1(\bar{w}_1 - \gamma_0\bar{w}_2) + \tau_{02}\bar{w}_2(\bar{w}_1 - \gamma_0\bar{w}_2)))|_{i=0}\\ &\quad - (\lambda_1\bar{w}_1^2 + \lambda_2\bar{w}_2^2 + (\tau_{N1}\bar{w}_1(\bar{w}_2 - \gamma_N\bar{w}_1) + \tau_{N2}\bar{w}_2(\bar{w}_2 - \gamma_N\bar{w}_1)))|_{i=N}.\end{aligned}$$

Thus, if $\tau_{01} = \lambda_1$, $\tau_{02} = \gamma_0\lambda_1$; $\tau_{N2} = -\lambda_2$, $\tau_{N1} = -\gamma_N\lambda_2$, then

$$\mathrm{BT}_{\mathrm{num}} = (\lambda_2 + \lambda_1\gamma_0^2)\bar{w}_2^2|_{i=0} - (\lambda_1 + \lambda_2\gamma_N^2)\bar{w}_1^2|_{i=N}.$$

Since $\lambda_1 > 0$, $\lambda_2 < 0$ and

$$(\lambda_2 + \lambda_1\gamma_0^2) \le 0 \iff \gamma_0^2 \le -\lambda_2/\lambda_1; \quad (\lambda_1 + \lambda_2\gamma_N^2) \ge 0 \iff \gamma_N^2 \le -\lambda_1/\lambda_2,$$

then it must be true that $\mathrm{BT}_{\mathrm{num}} \le 0$. Since A is negative semi-definite, then for $\alpha \ge 0$, we have $\alpha/2\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} \le 0$ and

$$\frac{1}{2}\frac{d}{dt}\|\bar{\mathbf{q}}\|^2 = \frac{\alpha}{2}\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} + \mathrm{BT}_{\mathrm{num}} \le 0.$$

This completes the proof. □

The next theorem will prove the stability of the semi-discrete approximation (3.5) for super-critical flows.

THEOREM 3.2. *Consider the semi-discrete finite-volume approximation* (3.5) *with the SAT* (3.7) *and* $\boldsymbol{b} = 0$ *for super-critical flows. If the penalty parameters are chosen such that* $\tau_{01} \ge \lambda_1$, $\tau_{02} \ge \lambda_2$; $\tau_{N1} \ge -\lambda_1$, $\tau_{N2} \ge -\lambda_2$, *then*

$$\frac{1}{2}\frac{d}{dt}\|\bar{\boldsymbol{q}}\|^2 \le 0 \quad \textit{for all } t \ge 0.$$

PROOF. As above, the energy method with the SBP property (3.4) and the eigen-decomposition of M yields

$$\frac{1}{2}\frac{d}{dt}\|\bar{\mathbf{q}}\|^2 - \frac{\alpha}{2}\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} = \mathrm{BT}_{\mathrm{num}},$$

where

$$\begin{aligned}\mathrm{BT}_{\mathrm{num}} &= ((\lambda_1 - \tau_{01})\bar{w}_1^2 + (\lambda_2 - \tau_{02})\bar{w}_2^2)|_{i=0} - (\lambda_1\bar{w}_1^2 + \lambda_2\bar{w}_2^2)|_{i=N}, \quad U > 0,\\ \mathrm{BT}_{\mathrm{num}} &= (\lambda_1\bar{w}_1^2 + \lambda_2\bar{w}_2^2)|_{i=0} - ((\lambda_1 + \tau_{N1})\bar{w}_1^2 + (\lambda_2 + \tau_{N2})\bar{w}_2^2)|_{i=N}, \quad U < 0.\end{aligned}$$

Therefore, if $\tau_{01} \ge \lambda_1, \tau_{02} \ge \lambda_2; \tau_{N1} \ge -\lambda_1, \tau_{N2} \ge -\lambda_2$, then we have $\mathrm{BT}_{\mathrm{num}} \le 0$. Noting that $\alpha \ge 0$ and, as previous, we have $\alpha/2\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} \le 0$, which gives us

$$\frac{1}{2}\frac{d}{dt}\|\bar{\mathbf{q}}\|^2 = \frac{\alpha}{2}\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} + \mathrm{BT}_{\mathrm{num}} \le 0,$$

so the proof is complete. □

Finally, we will prove the stability of the semi-discrete approximation (3.5) for critical flows.

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

THEOREM 3.3. *Consider the semi-discrete finite-volume approximation* (3.5) *with the SAT* (3.7) *and* $\boldsymbol{b} = 0$ *for critical flows. If the penalty parameters are chosen such that* $\tau_{01} \geq \lambda_1$, $\tau_{02} = 0$; $\tau_{N1} = 0$, $\tau_{N2} \geq -\lambda_2$, *then*

$$\frac{1}{2}\frac{d}{dt}\|\bar{\boldsymbol{q}}\|^2 \leq 0 \quad \text{for all } t \geq 0.$$

PROOF. The zero penalties ensure consistency of the SAT, that is, $\tau_{02} = 0$ and $\tau_{N1} = 0$ give

$$\text{SAT} = -\frac{1}{2}(W^{-1}S \otimes \mathbf{I})\begin{bmatrix}\tau_{01}H\mathbf{e}_0\bar{w}_1 \\ 0\end{bmatrix}, \quad U > 0,$$
$$\text{SAT} = -\frac{1}{2}(W^{-1}S \otimes \mathbf{I})\begin{bmatrix}0 \\ \tau_{N2}g\mathbf{e}_N\bar{w}_2\end{bmatrix}, \quad U < 0.$$

Again the energy method with the SBP property (3.4) and the eigen-decomposition of M yield

$$\frac{1}{2}\frac{d}{dt}\|\bar{\mathbf{q}}\|^2 - \frac{\alpha}{2}\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} = \text{BT}_{\text{num}},$$

where

$$\text{BT}_{\text{num}} = (\lambda_1 - \tau_{01})\bar{w}_1^2|_{i=0} - \lambda_1\bar{w}_1^2|_{i=N}, \quad U > 0,\ \lambda_1 > 0,\ \lambda_2 = 0,$$
$$\text{BT}_{\text{num}} = \lambda_2\bar{w}_2^2|_{i=0} - (\lambda_2 + \tau_{N2})\bar{w}_2^2|_{i=N}, \quad U < 0,\ \lambda_1 = 0,\ \lambda_2 < 0.$$

Therefore, if $\tau_{01} \geq \lambda_1$ and $\tau_{N2} \geq -\lambda_2$, then we have $\text{BT}_{\text{num}} \leq 0$. Using the fact that $\alpha \geq 0$ and $\alpha/2\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} \leq 0$ again gives the result that we wanted:

$$\frac{1}{2}\frac{d}{dt}\|\bar{\mathbf{q}}\|^2 = \frac{\alpha}{2}\bar{\mathbf{q}}^T(\mathbf{I} \otimes A)\bar{\mathbf{q}} + \text{BT}_{\text{num}} \leq 0.$$

This completes the proof. □

4. Numerical experiments

In this section, we perform numerical experiments to verify the analysis undertaken in the previous sections. Similar to the theoretical analysis, the numerical experiments cover the three flow regimes, namely the sub-critical, critical and super-critical flow regimes. We used $H = 1$ m, $g = 9.8$ m/s^2, and $U \in \{\frac{1}{2}\sqrt{gH}, \sqrt{gH}, 2\sqrt{gH}\}$, which correspond to the three different flow regimes. The interval of interest is $[0, L]$ with $L > 0$. Note that $U > 0$ so that $x = 0$ is the inflow boundary and $x = L$ is the outflow boundary. The locations and the number of boundary conditions required are given in Table 1, and the explicit forms of the boundary conditions considered here are given in Table 2, where $g_1(t)$ and $g_2(t)$ are the boundary data.

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

TABLE 1. The number and location of the boundary condition in all regime. The boundary at $x = 0$ ($x = 1$) is inflow (outflow) boundary if $U > 0$ and outflow (inflow) boundary if $U < 0$.

Regime	Type of boundary	Number of boundary conditions
sub-critical	inflow	1
	outflow	1
critical	inflow	1
	outflow	0
super-critical	inflow	2
	outflow	0

TABLE 2. Transmissive boundary conditions in all regimes with $U > 0$.

Regime	U	Boundaries	Boundary conditions
sub-critical	$< \sqrt{gH}$	$x = 0$	$\frac{1}{2}(\widetilde{h} + \sqrt{H/g}\,\widetilde{u}) = g_1$
		$x = L$	$\frac{1}{2}(\widetilde{h} - \sqrt{H/g}\,\widetilde{u}) = g_2$
critical	$= \sqrt{gH}$	$x = 0$	$\frac{1}{2}(\widetilde{h} + \sqrt{H/g}\,\widetilde{u}) = g_1$
super-critical	$> \sqrt{gH}$	$x = 0$	$\frac{1}{2}(\widetilde{h} + \sqrt{H/g}\,\widetilde{u}) = g_1$
			$\frac{1}{2}(\widetilde{h} - \sqrt{H/g}\,\widetilde{u}) = g_2$

The semi-discrete system (3.5) is integrated in time using the classical fourth-order accurate explicit Runge–Kutta method with the time step

$$\Delta t = \mathrm{Cr}\,\frac{\Delta x}{|U| + \sqrt{gH}}, \qquad \mathrm{Cr} = 0.25,$$

where Cr is the Courant–Friedrichs–Lewy number, $\Delta x = L/N$ is the uniform cell width and N is the number of finite-volume cells. We will consider a centred numerical flux with $\alpha = 0$ and the local Lax–Friedrich's numerical flux (3.1) with $\alpha > 0$, and verify numerical accuracy. Note that the semi-discrete approximation is energy stable for all $\alpha \geq 0$.

Nonhomogeneous boundary data. We consider zero initial conditions, that is, $u(x, 0) = 0$ and $h(x, 0) = 0$, and send a wave into the domain through the inflow boundary at $x = 0$. We will consider specifically $g_1(t) \neq 0$ and $g_2(t) = 0$ for the boundary conditions given in Table 2, so that the corresponding IBVP has the exact solution

$$\widetilde{h}(x,t) = g_1\Big(t - \frac{x}{U + \sqrt{gH}}\Big), \qquad \widetilde{u}(x,t) = \frac{1}{\sqrt{H/g}}\,g_1\Big(t - \frac{x}{U + \sqrt{gH}}\Big).$$

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

We will consider a smooth boundary data given by

$$g_1(t) = \begin{cases} (\sin(\pi t))^4 & \text{if } 0 \leq t \leq 1, \\ 0 & \text{otherwise,} \end{cases} \qquad g_2(t) = 0 \quad \text{for all } t \geq 0,$$

and non-smooth boundary data given by

$$g_1(t) = \begin{cases} 1 & \text{if } 0 < t \leq 1, \\ 0 & \text{otherwise,} \end{cases} \qquad g_2(t) = 0 \quad \text{for all } t \geq 0.$$

The boundary data for the boundary conditions in Table 2 can be rewritten as $b_1(t)$ and/or $b_2(t)$, and in terms of $w_1 = b_1(t) = \sqrt{2}/Hg_1(t)$ and $w_2 = b_2(t) = \sqrt{2}/Hg_2(t)$ for the given boundary. In the sub-critical case, by using the linear transformation (2.9), the boundary condition can be rewritten in the form (2.11) with $\gamma_0 = 0$ and $\gamma_N = 0$.

Using the fact that λ_1 and λ_2 have different signs, we have $\gamma_0^2 \leq -\lambda_2/\lambda_1$ and $\gamma_N^2 \leq -\lambda_1/\lambda_2$ for all $\mathrm{Fr} < 1$, that is, the condition of Lemma 2.5 is satisfied.

For the critical flow regime, we have $\mathrm{Fr} = 1$ and $\lambda_2 = 0$. Only one boundary condition is imposed at the inflow, $\bar{w}_1 = b_1(t) = (\sqrt{2}/H)g_1(t)$. This condition can be rewritten to match the condition in Theorem 3.3 by using the linear transformation (2.9) and the fact that $U^2 = gH$.

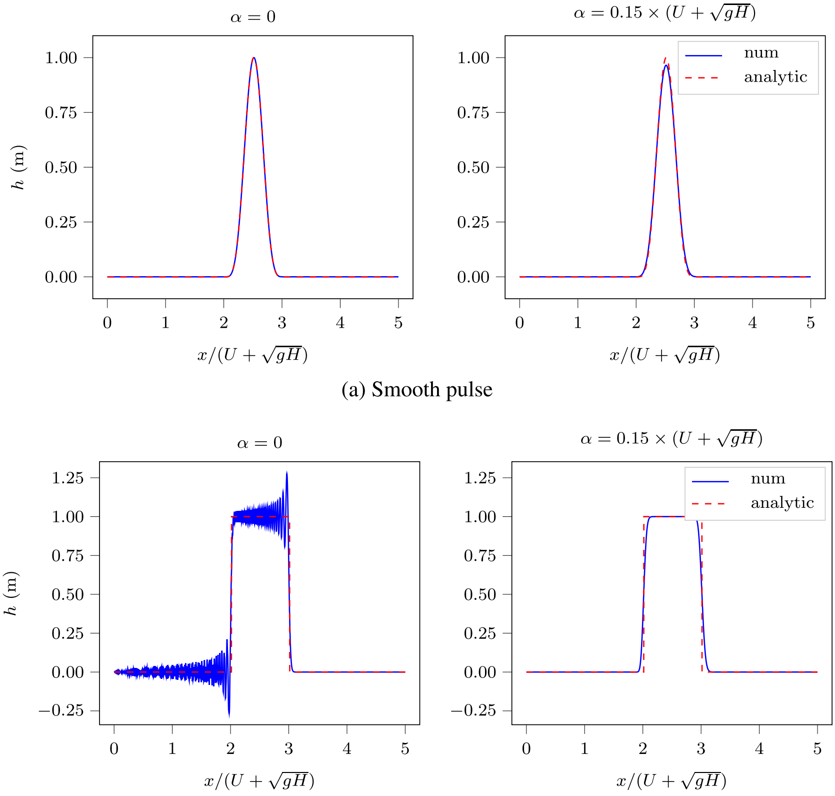
For the super-critical flow regime, two boundary conditions are imposed at the inflow boundary. That is, $\bar{w}_1 = b_1(t) = (\sqrt{2}/H)g_1(t)$, $\bar{w}_2 = b_2(t) = (\sqrt{2}/H)g_2(t)$ as the boundary condition at $x = 0$. These boundary conditions are equivalent to (2.12a).

The boundary data will generate a pulse from the left boundary at $x = 0$, which will propagate through the domain and leave the domain through $x = L$.

Figure 2 shows the snapshot of the sub-critical flow solutions at $t = 3.02$ s for both smooth and nonsmooth boundary data, with $\alpha = 0$ and $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$. In the plots, we have scaled the horizontal axis by the wave speed $(U + \sqrt{gH})$ so that the solution is spatially invariant for all flow regimes. Note that for the smooth pulse, the numerical solution matches the exact solution excellently well for $\alpha = 0$ and $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$, although with $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$, the peak of the numerical is slightly dissipated. For the nonsmooth pulse, when $\alpha = 0$, the propagation speed of the pulse is well approximated by the numerical solution. However, there are numerical oscillations generated by the propagating discontinuities. When $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$, the numerical solution is nonoscillatory, but the discontinuous edges of the solutions are smeared.

The evolution of the numerical solutions and the exact solutions, at all flow regimes, are shown in Figure 3 for the smooth pulse and in Figure 4 for the nonsmooth pulse. The pulses enter the domain through the inflow boundary at $x = 0$ and leave the domain through the outflow at $x = L = (U + \sqrt{gH}) \times 5$. Note that because of the re-scaling of the x-axis to $x/(U + \sqrt{gH})$, the solutions are invariant for all three flow regimes.

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press



(a) Smooth pulse

(b) Non-smooth pulse

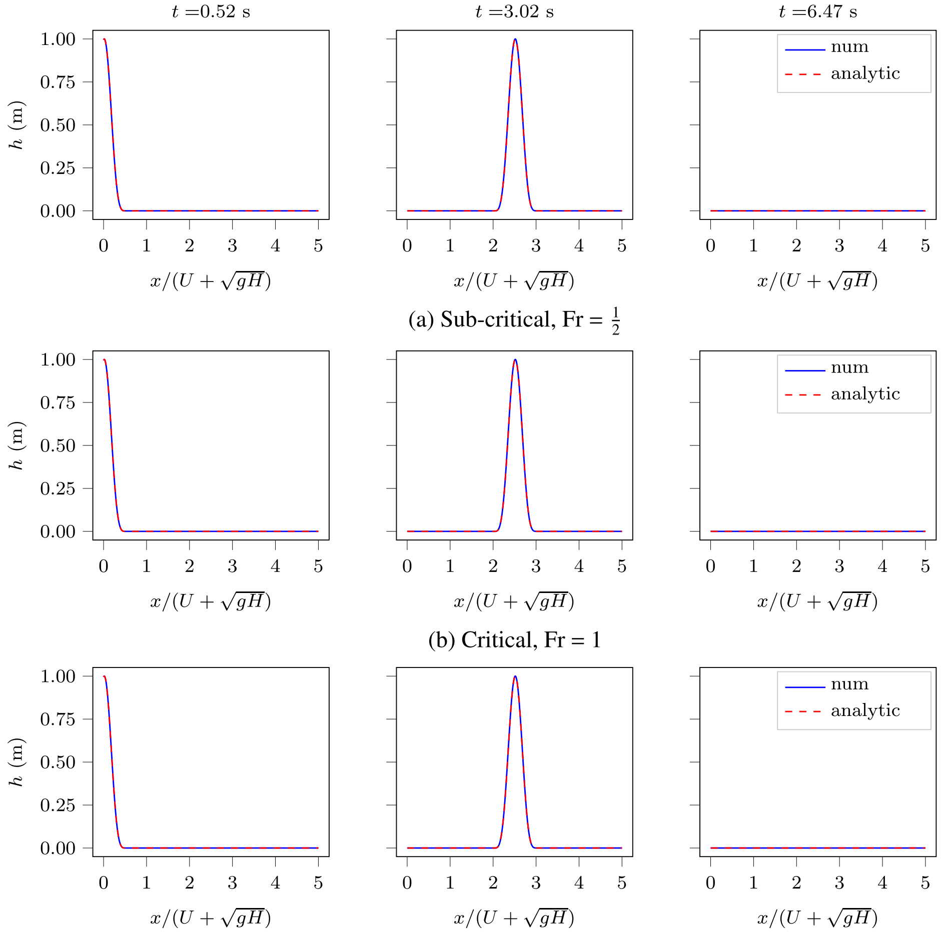
FIGURE 2. The snapshots of the numerical and exact solutions with $\Delta x = L \times 2^{-11}$ m at time $t = 3.02$ s for a sub-critical flow regime with smooth and nonsmooth boundary data. For the smooth boundary data, the numerical solution matches the exact solution well for $\alpha = 0$ and $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$. Note, however, with $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$, the peak of the numerical solution is slightly dissipated. For the nonsmooth boundary data, when $\alpha = 0$, the propagation speed of the pulse is well approximated by the numerical solution. However, there are numerical oscillations generated by the propagating discontinuities. When $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$, the numerical solution is nonoscillatory, but propagating discontinuities are smoothed out.

Convergence test. Here, we verify the convergence properties of the numerical method. We will use the method of the manufactured solution [11]. That is, we force the system to have the exact smooth solution

$$h(x,t) = \cos(2\pi t)\sin(6\pi x), \qquad u(x,t) = \sin(2\pi t)\cos(4\pi x). \tag{4.1}$$

The initial conditions $h(x,0)$, $u(x,0)$ and the boundary data $g_1(t)$ and $g_2(t)$ are chosen to match the analytical solution (4.1). We compute the numerical solution on a sequence

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press



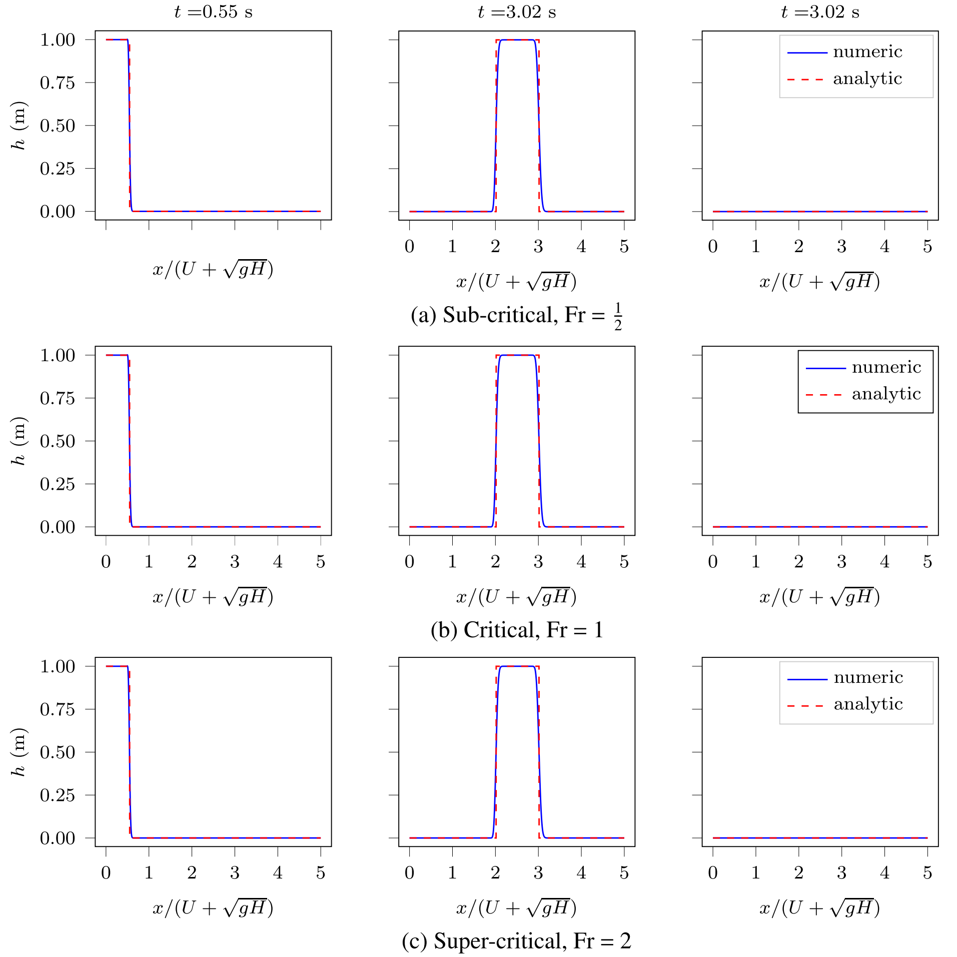
(a) Sub-critical, Fr = $\frac{1}{2}$

(b) Critical, Fr = 1

(c) Super-critical, Fr = 2

FIGURE 3. The evolution of the numerical solutions and the exact solutions at all the three flow regimes with smooth boundary data, $\Delta x = L \times 2^{-11}$ m and $\alpha = 0$. The solutions enter the domain through the inflow boundary at $x = 0$ and leave the domain through the outflow at $x = L = (U + \sqrt{gH}) \times 5$. Note that because of the re-scaling of the x-axis to $x/(U + \sqrt{gH})$, the solutions are invariant for all three flow regimes.

of increasing number of finite-volume cells, $N = 64, 128, 256, 512, 1024, 2048$. The L_2-error and convergence rates of the error are shown in Figure 5 and also presented in Table 3. We have performed numerical experiments with no dissipation $\alpha = 0$ and with numerical dissipation set at $\alpha = 0.05$. From Table 3, we see that the method is second-order accurate $O(\Delta x^2)$ when $\alpha = 0$, and first-order accurate $O(\Delta x)$ when $\alpha > 0$. These are in agreement with the theory.



(a) Sub-critical, Fr = $\frac{1}{2}$

(b) Critical, Fr = 1

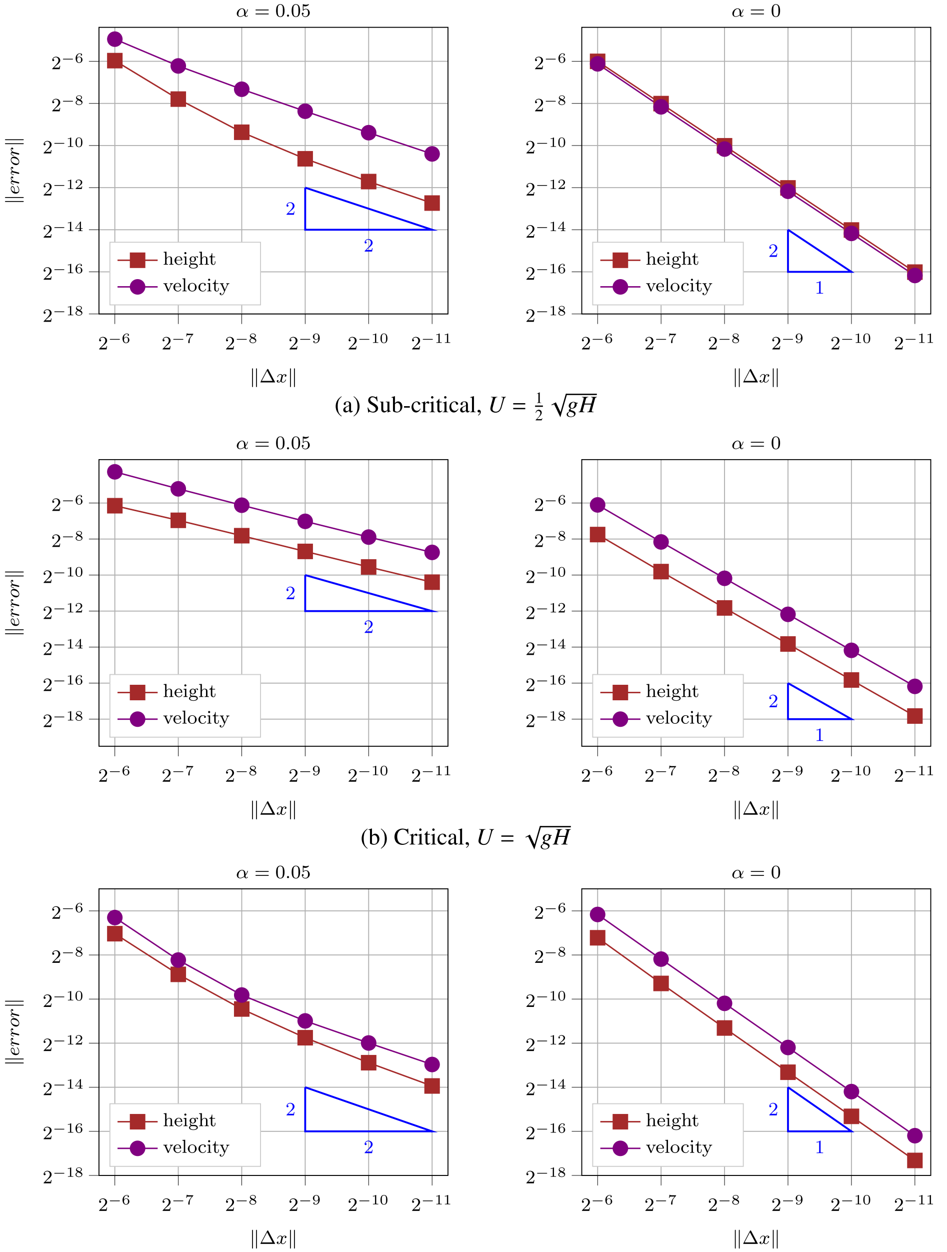
(c) Super-critical, Fr = 2

FIGURE 4. The evolution of the numerical solutions and the exact solutions at all three flow regimes with nonsmooth boundary data, $\Delta x = L \times 2^{-11}$ m and $\alpha = 0.15 \times (U + \sqrt{gH}) > 0$. The discontinuous solutions enter the domain through the inflow boundary at $x = 0$ and leave the domain through the outflow at $x = L = (U + \sqrt{gH}) \times 5$. Note that because of the re-scaling of the x-axis to $x/(U + \sqrt{gH})$, the solutions are invariant for all three flow regimes.

5. Conclusion

Well-posed boundary conditions are crucial for accurate numerical solutions of IBVPs. In this study, we have analysed well-posed boundary conditions for the linear SWWE in 1D. The analysis is based on the energy method and prescribes the number, location and form of the boundary conditions so that the IBVP is well-posed. A summary of the results is shown in Table 1 and covers all flow regimes. We formulate the boundary conditions such that they can be readily implemented in a stable manner

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press



(a) Sub-critical, $U = \frac{1}{2}\sqrt{gH}$

(b) Critical, $U = \sqrt{gH}$

(c) Super-critical, $U = 2\sqrt{gH}$

FIGURE 5. The error and convergence of the error at final time $t = 0.1$ using the manufactured solution for all flow regimes.

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press

TABLE 3. The error and convergence of the error at final time $t = 0.1$ using the manufactured solution for all flow regimes.

(a) Sub-critical, $U = \frac{1}{2}\sqrt{gH}$

N	$\alpha = 0.0$				$\alpha = 0.05$			
	h error	rate	u error	rate	h error	rate	u error	rate
64	1.56×10^{-02}		1.44×10^{-02}		1.60×10^{-02}		3.24×10^{-02}	
128	3.88×10^{-03}	2.01	3.49×10^{-03}	2.06	4.50×10^{-03}	1.77	1.34×10^{-02}	1.21
256	9.68×10^{-04}	2.00	8.69×10^{-04}	2.01	1.51×10^{-03}	1.49	6.23×10^{-03}	1.08
512	2.42×10^{-04}	2.00	2.17×10^{-04}	2.01	6.31×10^{-04}	1.20	3.02×10^{-03}	1.03
1024	6.05×10^{-05}	2.00	5.41×10^{-05}	2.00	2.98×10^{-04}	1.06	1.49×10^{-03}	1.01
2048	1.51×10^{-05}	2.00	1.35×10^{-05}	2.00	1.47×10^{-04}	1.01	7.41×10^{-04}	1.01

(b) Critical, $U = \sqrt{gH}$

N	$\alpha = 0.0$				$\alpha = 0.05$			
	h error	rate	u error	rate	h error	rate	u error	rate
64	4.64×10^{-03}		1.45×10^{-02}		1.41×10^{-02}		5.19×10^{-02}	
128	1.12×10^{-03}	2.08	3.50×10^{-03}	2.08	8.03×10^{-03}	0.88	2.70×10^{-02}	0.96
256	2.76×10^{-04}	2.03	8.63×10^{-04}	2.03	4.44×10^{-03}	0.90	1.44×10^{-02}	0.94
512	6.88×10^{-05}	2.00	2.15×10^{-04}	2.00	2.43×10^{-03}	0.92	7.72×10^{-03}	0.93
1024	1.72×10^{-05}	2.00	5.39×10^{-05}	2.00	1.33×10^{-03}	0.91	4.21×10^{-03}	0.92
2048	4.30×10^{-06}	2.00	1.35×10^{-05}	2.00	7.43×10^{-04}	0.90	2.34×10^{-03}	0.90

(c) Super-critical, $U = 2\sqrt{gH}$

N	$\alpha = 0.0$				$\alpha = 0.05$			
	h error	rate	u error	rate	h error	rate	u error	rate
64	6.71×10^{-03}		1.40×10^{-02}		7.63×10^{-03}		1.27×10^{-02}	
128	1.60×10^{-03}	2.10	3.43×10^{-03}	2.03	2.12×10^{-03}	1.80	3.33×10^{-03}	1.90
256	3.93×10^{-04}	2.03	8.54×10^{-04}	2.01	7.15×10^{-04}	1.49	1.11×10^{-03}	1.50
512	9.79×10^{-05}	2.01	2.13×10^{-04}	2.00	2.90×10^{-04}	1.23	4.91×10^{-04}	1.13
1024	2.45×10^{-05}	2.00	5.32×10^{-05}	2.00	1.32×10^{-04}	1.10	2.45×10^{-04}	1.00
2048	6.11×10^{-06}	2.00	1.33×10^{-05}	2.00	6.35×10^{-05}	1.04	1.25×10^{-04}	0.98

using the SBP-SAT method. We propose a finite-volume method formulated in the SBP framework and implement the boundary conditions weakly using SAT. Stable penalty parameters and proof of numerical stability are derived via discrete energy estimates analogous to the continuous estimate. Numerical experiments are performed to verify the analysis. The error rates comply with the methods that we use. Our continuous and numerical analysis covers all flow regimes and can be extended to the nonlinear problem. The next step in our study will extend the 1D theory and results to two dimensions, and implement our scheme in open source software [10, 12] for efficient and accurate simulations of the nonlinear shallow water equations.

Acknowledgements

This research is conducted as part of doctoral study funded by Indonesian Endowment Fund for Education (LPDP). We would like to express our gratitude to Larry Forbes for his insightful feedback, which has enriched our understanding and interpretation of the results presented in this paper.

References

[1] M. H. Carpenter, D. Gottlieb and S. Abarbanel, "Time-stable boundary conditions for finite-difference schemes solving hyperbolic systems: methodology and application to high-order compact schemes", *J. Comput. Phys.* **111** (1994) 220–236; doi:10.1006/jcph.1994.1057.

[2] J. A. Cunge, F. M. Holly and A. Verwey, *Practical aspects of computational river hydraulics* (Pitman Advanced Publishing Program, Boston, MA, 1980).

[3] G. J. Gassner, "A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods", *SIAM J. Sci. Comput.* **35** (3) (2013) A1233–A1253; doi:10.1137/120890144.

[4] S. Ghader and J. Nordström, "Revisiting well-posed boundary conditions for the shallow water equations", *Dyn. Atmospheres Oceans* **66** (2014) 1–9; doi:10.1016/j.dynatmoce.2014.01.002.

[5] B. Gustafsson, *High order difference methods for time dependent PDE*, Volume 38 of *Springer Series in Computational Mathematics* (Springer-Verlag, Berlin–Heidelberg, 2007); doi:https://doi.org/10.1007/978-3-540-74993-6.

[6] B. Gustafsson, H.-O. Kreiss and J. Oliger, *Time dependent problems and difference methods*, Volume 24 of *Pure and Applied Mathematics: A Wiley-Interscience series of Texts, Monographs, and Tracts* (John Wiley & Sons, New York, 1995).

[7] H.-O. Kreiss and G. Scherer, "Finite element and finite difference methods for hyperbolic partial differential equations", in: *Mathematical aspects of finite elements in partial differential equations*, Proceedings of a Symposium Conducted by the Mathematics Research Center, University of Wisconsin–Madison, April 1–3, 1974 (ed. C. de Boor) (Academic Press, Cambridge, MA, 1974) 195–212; doi:10.1016/B978-0-12-208350-1.50012-1.

[8] T. Lundquist and J. Nordström, "The SBP-SAT technique for initial value problems", *J. Comput. Phys.* **270** (2014) 86–104; doi:10.1016/j.jcp.2014.03.048.

[9] K. Mahmood, *Unsteady flow in open channels*, Volume 2 (eds. V. M. Yevjevich and W. A. Miller) (Water Resources Publications, CO, 1975);
https://books.google.com.au/books?id=wApSAAAAMAAJ.

[10] O. M. Nielsen, S. G. Roberts, D. Gray, A. McPherson and A. Hitchman, "Hydrodynamic modelling of coastal inundation", in: *MODSIM 2005 International Congress on Modelling and Simulation* (eds. A. Zerger and R. M. Argent) (Modelling and Simulation Society of Australia & New Zealand, Canberra, 2005) 521–523; http://www.mssanz.org.au/modsim05/papers/nielsen.pdf.

[11] P. J. Roache, "Code verification by the method of manufactured solutions", *J. Fluids Eng.* **124** (2001) 4–10; doi: 10.1115/1.1436090.

[12] S. Roberts, G. Davies and O. Nielsen, ANUGA github repository, 6 (2022);
https://github.com/anuga-community/anuga_core.

[13] A. R. Winters and G. J. Gassner, "A comparison of two entropy stable discontinuous Galerkin spectral element approximations for the shallow water equations with non-constant topography", *J. Comput. Phys.* **301** (2015) 357–376; doi:10.1016/j.jcp.2015.08.034.

[14] C. Zoppou and S. Roberts, "Explicit schemes for dam-break simulations", *J. Hydraul. Eng.* **129** (2003) 11–34; doi:10.1061/(ASCE)0733-9429(2003)129:1(11).

https://doi.org/10.1017/S1446181124000191 Published online by Cambridge University Press