

# Unconditionally stable space-time discontinuous residual distribution for shallow-water flows

Domokos Sarmany, Matthew Hubbard, Mario Ricchiuto

# ▶ To cite this version:

Domokos Sarmany, Matthew Hubbard, Mario Ricchiuto. Unconditionally stable space-time discontinuous residual distribution for shallow-water flows. [Research Report] RR-7958, 2012. hal-00696083v1

# HAL Id: hal-00696083 https://inria.hal.science/hal-00696083v1

Submitted on 10 May 2012 (v1), last revised 14 May 2012 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers. L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



# Unconditionally stable space-time discontinuous residual distribution for shallow-water flows

D. Sarmany, M. Hubbard, M. Ricchiuto

RESEARCH REPORT N° 7958 May 2012 Project-Team BACCHUS



# Unconditionally stable space-time discontinuous residual distribution for shallow-water flows

D. Sarmany<sup>\*</sup>, M. Hubbard<sup>\*</sup>, M. Ricchiuto<sup>†</sup>

Project-Team BACCHUS

Research Report n° 7958 — May 2012 — 33 pages

**Abstract:** This article describes a discontinuous implementation of residual distribution for shallow-water flows. The emphasis is put on the space-time implementation of residual distribution for the time-dependent system of equations with discontinuity in time only. This lifts the time-step restriction that even implicit continuous residual distribution schemes invariably suffer from, and thus leads to an unconditionally stable discretisation. The distributions are the space-time variants of the upwind distributions for the steady-state system of equations and are designed to satisfy the most important properties of the original mathematical equations: positivity, linearity preservation, conservation and hydrostatic balance. The purpose of the several numerical examples presented in this article is twofold. First, to show that the discontinuous numerical discretisation does indeed exhibit all the desired properties when applied to the shallow-water equations. Second, to investigate how much the time step can be increased without adversely affecting the accuracy of the scheme and whether this translates into gains in computational efficiency. Comparison to other existing residual distribution schemes is also provided to demonstrate the improved performance of the scheme.

**Key-words:** hyperbolic conservation laws, shallow-water equations, space-time discontinuous representation, residual distribution

\* School of Computing, University of Leeds, Leeds, United Kingdom - d.sarmany@leeds.ac.uk, m.e.hubbard@leeds.ac.uk † INRIA Bordeaux Sud-Ouest, 351 Cours de la Libération, 33405 Talence Cedex, France - mario.ricchiuto@inria.fr

#### RESEARCH CENTRE BORDEAUX – SUD-OUEST

351, Cours de la Libération Bâtiment A 29 33405 Talence Cedex

# Unconditionally stable space-time discontinuous residual distribution for shallow-water flows

**Résumé :** Cet article décrit une méthode Residual Distribution avec une représentation discontinue en temps pour les équations Shallow Water. L'approche espace-temps avec une représentation discontinue permet de se débarrasser de la condition de cfl associée à la preservation de la positivité des coefficient du schéma. Une méthode inconditionnellement stable est obtenue. L'approche residual distribution permet de satisfare un nombre important de proprités : positivité, lineraity preservation, conservativité, well balancedness. Ce travail a comme but: de montrer que l'approche proposée effectivement a toutes les propriétés mentionnées; de vérifier jusque à quelle limite on peut augmenter la taille du pas de temps sans observer des pertes de précision importantes. Des comparaisons avec les approches existantes permettent de valider ultérieurement la méthode

**Mots-clés :** hyperbolic conservation laws, shallow-water equations, space-time discontinuous representation, residual distribution

# Contents

т,

T	Introduction	3
<b>2</b>	Space-time discontinuous residual distribution	4
	2.1 Scalar equations	5
	2.2 Discontinuity in time for the scalar equation	7
	2.3 Systems of equations	8
	2.4 Conservative linearisation for systems of equations	8
	2.5 Upwind RD schemes for systems of equations	8
	2.6 Discontinuity in time for systems of equations	9
3	Application to the shallow-water system	10
	3.1 Conservative schemes	10
	3.1.1 A conservative linearisation	10
	3.1.2 The CRD approach	10
	3.2 Hydrostatic balance: the C-property	11
	3.2.1 The CRD approach	11
	3.2.2 Conservative linearisation	11
4	Implementation details	12
5	Numerical results	12
	5.1 Circular dam break over wet bed	13
	5.2 Improving the iterative convergence	13
	5.3 Lake at rest: the preservation of the hydrostatic balance	15
	5.4 Travelling vortex	16
	5.5 Partial dam break	19
	5.6 Circular dam break over nonsmooth bed	19
6	Comparison with existing schemes	20
	6.1 Travelling vortex	21
	6.2 Partial dam break	21
	6.3 Circular dam break over nonsmooth bed	23
7	Concluding remarks and outlook	<b>25</b>

# 1 Introduction

The framework of residual distribution (RD) has a nearly thirty-year-old history, having first been introduced in [?] as an alternative to finite volume schemes for the numerical discretisation of hyperbolic conservation laws. They can more naturally represent the most important underlying physical properties of steady-flow problems – especially avoiding spurious oscillations around discontinuities – while still providing second-order accurate approximations. For time-dependent problems, however, it initially proved to be rather more difficult to construct an efficient, second-order accurate, RD scheme that is also free of spurious oscillations. Much of the recent research in the field, therefore, has been aimed at devising truly time-dependent discretisations that retain all the desired properties of the underlying physical equations. Recent reviews of the field can be found in [?, ?, ?].

Almost all of these schemes, however, use continuous discretisations of the physical variables. This leads to numerical methods that are both implicit, because of the presence of a global mass matrix, and require a time-step restriction to obtain stability, positivity or both. For many practical applications in physics or engineering, these methods can prove to be computationally inefficient and uncompetitive with more established discretisation methods, such as finite volumes [?, ?], finite elements [?] or finite

•

differences [?]. To remedy this shortcoming, two different approaches have been proposed in the literature for RD schemes that improve the efficiency of the time-step algorithm as well as keep the discrete representation continuous. One is to make the discretisation genuinely explicit and thus remove the need to solve an algebraic system at each time step [?]. The other is to introduce double layers in a space-time implicit scheme, which removes the restrictions on the time step in the second layer [?].

In contrast, it is also possible to allow the discrete representation to be discontinuous across cell/prism interfaces. It is only relatively recently that this concept has been introduced – at first for steady-state equations only [?, ?]. The potential benefits of the discontinuous RD (DRD) framework in space include: a) the possibility to implement h-adaptivity more easily; b) a mass-matrix structure that may render the inversion of the mass matrix computationally less expensive. As a further development, the work in [?] took the first steps towards unconditionally stable space-time RD schemes by introducing discontinuous representation in time. One can also view this construction as a special, and slightly simpler, case of the space-time double-layer scheme [?, ?], where the first layer is of zero height, while there is still no formal condition on the height of the second layer.

In this article, we extend the concept of space-time DRD to the shallow-water equations with possibly non-flat bottom topography. Mathematically, the non-flat bottom topography manifests itself as a source term in the governing equations. This means that apart from the usual requirements – positivity, linearity preservation, conservation and upwinding – a successful numerical scheme should also preserve the hydrostatic balance. Various RD schemes have been developed for shallow-water flows in the past one and a half decades. Both wave decomposition schemes [?, ?] and matrix distribution schemes [?, ?] proved to be successful in representing the underlying physical properties. However, they are often computationally rather expensive owing partly to a time-step restriction mentioned above, and partly to poor iterative convergence. As a result and more recently, the explicit RD approach has been applied to the shallow-water system [?].

We take a different approach in this work and instead of replacing an intrinsically implicit scheme with an explicit one, we aim to eliminate the time-step restriction and construct an unconditionally stable space-time RD scheme. The focus of this article is therefore on discontinuous representation in time, which in this way promises to improve the efficiency of the implicit space-time discretisation. Our spacetime DRD discretisation is also designed to be upwind, conservative and hydrostatically well-balanced. Through a number of numerical experiments, we investigate whether these properties are satisfied, as well as positivity and linearity preservation – two other properties that are formally lost but experimentally often observed. Crucially, we also address the question of how far the time step can be increased before the discretisation adversely affects accuracy and positivity. Comparison with other implicit RD schemes is used to demonstrate the improved performance of this approach.

The remaining part of the article is organised as follows. In Section 2, we describe the framework of the space-time DRD for general hyperbolic conservation laws in the case of both scalar and system variables. Section 3 applies the scheme to the shallow-water equations, while Section 4 outlines details about the numerical implementation of the discretisation. An extensive study of the scheme via numerous computational examples is carried out in Section 5. Section 6 compares the space-time DRD scheme developed in this work with other existing RD schemes. Finally, we draw conclusions and provide outlook in Section 7.

# 2 Space-time discontinuous residual distribution

In this section, we give a general description for space-time DRD schemes for hyperbolic conservation laws with homogeneous right-hand side. Assume a two-dimensional spatial domain  $\Omega \subset \mathbb{R}^2$ , its triangular tessellation  $\Omega_h$ , and the corresponding prismatic space-time mesh  $\Omega_h^t$ . Let, furthermore, E denote a given triangle,  $E_t$  denote the space-time prism defined over this triangle,  $D_i = \bigcup_{i \in E} E$ ,  $d\Omega_h = dx dy$ ,  $d\Omega_h^t = dx dy dt$ ,  $d(\partial \Omega) = ds(x, y)$  and  $d(\partial \Omega_t) = ds_t(x, y, t)$ . In the last two definitions, s and  $s_t$  denote the mappings  $s : \Omega \to \partial \Omega$  and  $\Omega_t \to \partial \Omega_t$ , respectively.

#### 2.1 Scalar equations

Consider the scalar conservation law

$$\partial_t u + \nabla \cdot \boldsymbol{f} = 0 \quad \text{or} \quad \partial_t u + \boldsymbol{a}(u) \cdot \nabla u = 0$$
 (1)

with appropriate initial conditions and Dirichlet boundary conditions at the inflow part of the domain. Here  $\mathbf{f}$  represents the conservative flux vector and  $\mathbf{a}(u) = \partial \mathbf{f}/\partial u$  is the wave speed. The associated residual over each space-time prism  $E_t$  is given by

$$\phi_{E_t} = \int_{t^n}^{t^{n+1}} \int_E \left(\partial_t u + \nabla \cdot \boldsymbol{f}\right) \mathrm{d}x \,\mathrm{d}y \,\mathrm{d}t.$$
<sup>(2)</sup>

For the original (non-discretised) equation (1), it therefore holds that

$$\phi = \int_{t^n}^{t^{n+1}} \int_{\Omega_h} \left( \partial_t u + \nabla \cdot \boldsymbol{f} \right) \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t = \sum_{E \in \Omega_h} \phi_{E_t},$$

provided that the solution u is bounded and piecewise differentiable.

We carry out the RD numerical discretisation in the spirit of [?] and it consists of the following general steps.

- 1. In every space-time prism, replace the unknown u with an approximation  $u_h$  that is linear in space and linear in time.
- 2. Transform the prism residual into a space-time boundary integral

$$\phi_{E_t} = \int_{t^n}^{t^{n+1}} \int_E \left(\partial_t u_h + \nabla \cdot \boldsymbol{f}\right) \mathrm{d}x \, \mathrm{d}y \, \mathrm{d}t = \int_{E_t} \nabla_t \cdot \boldsymbol{f}_t \, \mathrm{d}\Omega_t = \int_{\partial E_t} \boldsymbol{f}_t \cdot \boldsymbol{n}_t \, \mathrm{d}(\partial\Omega_t), \tag{3}$$

where  $\nabla_t = (\partial t, \partial x, \partial y)$ ,  $\boldsymbol{f}_t = (u, \boldsymbol{f})$ , and  $\boldsymbol{n}_t$  is the outward-pointing unit vector normal to the space-time prism.

3. Let L denote the space of linear one-dimensional Lagrangian basis functions on the interval  $[t^n, t^{n+1}]$ and W denote the space of linear two-dimensional basis functions associated with a triangle. Assume furthermore that  $u_h \in L \times W$ . Then the mid-point rule in time results in

$$\phi_{E_t} = \int_{\partial E_t} \boldsymbol{f}_t \cdot \boldsymbol{n}_t \, \mathrm{d}(\partial \Omega_t)$$

$$\approx \int_E \left( u_h^{n+1} - u_h^n \right) \mathrm{d}\Omega + \frac{\Delta t}{2} \left( \int_{\partial E} \boldsymbol{f}(u^n) \cdot \boldsymbol{n} \, \mathrm{d}(\partial \Omega) + \int_{\partial E} \boldsymbol{f}(u^{n+1}) \cdot \boldsymbol{n} \, \mathrm{d}(\partial \Omega) \right)$$

$$= \int_E \left( u_h^{n+1} - u_h^n \right) \mathrm{d}\Omega + \frac{\Delta t}{2} (\phi_E^n + \phi_E^{n+1}), \quad (4)$$

where  $\Delta t = [t^n, t^{n+1}]$  is the 'height' of the prism. This is a second-order accurate approximation of the prism residual, as long as the spatial residuals,  $\phi_E^n$  and  $\phi_E^{n+1}$ , are computed to be at least one order more accurate than the discretisation itself.

4. Distribute the prism fluctuation  $\phi_{E_t}$  (4) to the six vertices of the prism in a conservative manner. That is, the fractions of the residual sent to vertex *i* at time levels *n* and *n* + 1 are defined as

$$\phi_{i,n}^E = \beta_{i,n}^E \phi_{E_t} \quad \text{and} \quad \phi_{i,n+1}^E = \beta_{i,n+1}^E \phi_{E_t},$$
(5)

with  $\sum_{i \in E} \beta_{i,n}^E + \sum_{i \in E} \beta_{i,n+1}^E = 1.$ 

5. Solve the algebraic system

$$\sum_{E \in D_i} \phi_{i,n}^E = 0,$$

$$\sum_{E \in D_i} \phi_{i,n+1}^E = 0,$$
(6)

 $\forall i \in \Omega_h$  at each time step.

How to precisely distribute the residual does greatly depend on the direction of the flow, which is typically indicated by 'inflow' parameters. For the space-time scheme, these are defined as

$$k_{i,n} = -\frac{\Delta t}{4} \overline{a} \cdot n_i - \frac{|E|}{3}, \quad k_{i,n}^+ = \max(0, k_{i,n}), \quad k_{i,n}^- = \min(0, k_{i,n}),$$

$$k_{i,n+1} = -\frac{\Delta t}{4} \overline{a} \cdot n_i + \frac{|E|}{3}, \quad k_{i,n+1}^+ = \max(0, k_{i,n+1}), \quad k_{i,n+1}^- = \min(0, k_{i,n+1}), \quad (7)$$

$$N_t = \frac{1}{\sum_{i \in E} k_{i,n}^+ + \sum_{i \in E} k_{i,n+1}^+},$$

where  $n_i$  is the outward pointing normal vector opposite node *i* with length that of the edge opposite node *i*. The quantity  $\overline{a}$  is a prism-averaged state of the values of *a* at the vertices of the prism.

If the discrete representation of the solution  $u_h$  is continuous over the whole space-time domain, the scheme is only consistent if there is no residual sent back to the previous time level  $t^n$ . This condition results in a CFL-type restriction on the height of the space-time prism  $\Delta t$ ,

$$k_{i,n} = -\frac{\Delta t}{4} \boldsymbol{a} \cdot \boldsymbol{n}_i - \frac{|E|}{3} \le 0.$$
(8)

Although this time step is generally larger than a more traditional CFL condition in an upwind finitevolume method – as we show it in the Appendix – it is an undesirable condition for a numerical discretisation that is inherently implicit.

The properties of a given distribution depend to a great degree on the precise definition of the coefficients. The typical requirements for a broadly successful RD scheme include [?]

- *positivity*, which warrants that the numerical approximations are free of spurious oscillations;
- *linearity preservation*, which ensures that a (k-1)th-order polynomial representation leads to a kth-order accurate scheme;
- conservation, which guarantees that discontinuities are captured correctly;
- *compactness*, which is primarily for computational efficiency and requires that the cell/prism residual be distributed to its own vertices only;
- continuous dependence of the coefficients, which enhances the iterative convergence of the algebraic solver; and
- *upwinding*, which dictates that the discretised model propagates information in the same direction and at the same velocity as its non-discretised counterpart.

By Godunov's theorem [?, ?], only nonlinear schemes can satisfy the conditions for both positivity and linearity preservation. Very often, though, linear schemes are used as the basis for constructing these nonlinear schemes. If  $\bar{a}$  is a conservative linearisation of  $a(u) = \partial f/\partial u$  over the space-time prism, then we can compute the discrete residual as

$$\phi_{E_t} = \sum_{i \in E} k_{i,n} u_i^n + \sum_{i \in E} k_{i,n+1} u_i^{n+1}, \tag{9}$$

and the resulting scheme will be conservative as long as the constraint following (5) is satisfied [?]. Otherwise, the prism residual needs to be computed by a sufficiently accurate quadrature rule in order to obtain conservation, even when (5) holds. This latter version is also known as the conservative RD (CRD) formulation [?, ?].

There are a large number of numerical schemes that can be (re)cast in the RD framework. We now describe only the space-time upwind schemes that are the focus of our investigation in this article. For this, the parameters defined in (7) are used.

• The space-time N (STN) scheme [?] is here defined so that it is suitable for the CRD formulation [?],

$$u_t^{\text{in}} = N_t \left( \sum_{j \in E} k_{j,n}^+ u_j^n + \sum_{j \in E} k_{j,n+1}^+ u_j^{n+1} - \phi_{E_t} \right),$$

$$\left(\phi_{i,n}^E\right)^N = k_{i,n}^+ \left(u_i^n - u_t^{\text{in}}\right), \qquad \left(\phi_{i,n+1}^E\right)^N = k_{i,n+1}^+ \left(u_i^{n+1} - u_t^{\text{in}}\right).$$
(10)

This is a linear scheme that has all the desired properties except linearity preservation. We note, however, that positivity is only formally guaranteed if we use (9) for the computation of  $\phi_{E_t}$ ; see [?] for space-time DRD schemes. Nevertheless, oscillation-free behaviour is also often observed in computational experiments [?] for the CRD scheme, when a quadrature rule is used to evaluate  $\phi_{E_t}$ .

• The space-time LDA (STLDA) [?, ?] scheme is defined as

$$\left(\phi_{i,n}^{E}\right)^{LDA} = k_{i,n}^{+} N_t \phi_{E_t}, \qquad \left(\phi_{i,n+1}^{E}\right)^{LDA} = k_{i,n+1}^{+} N_t \phi_{E_t}, \tag{11}$$

which is also a linear scheme and has all the desired properties except positivity.

• *The space-time blended* (STB) scheme combines two linear schemes, typically the STN and STLDA schemes, through a nonlinear blending coefficient,

$$\left(\phi_{i,n}^{E}\right)^{B} = \theta \left(\phi_{i,n}^{E}\right)^{N} + (1-\theta) \left(\phi_{i,n}^{E}\right)^{LDA},$$

$$\left(\phi_{i,n+1}^{E}\right)^{B} = \theta \left(\phi_{i,n+1}^{E}\right)^{N} + (1-\theta) \left(\phi_{i,n+1}^{E}\right)^{LDA}.$$

$$(12)$$

The blending coefficient determines how 'well' the required properties, especially positivity, are satisfied. The definition adopted in this article [?] is

$$\theta = \frac{|\phi_{E_t}|}{\sum_{i \in E} \left| \left( \phi_{i,n}^E \right)^N \right| + \sum_{i \in E} \left| \left( \phi_{i,n+1}^E \right)^N \right|}.$$
(13)

#### 2.2 Discontinuity in time for the scalar equation

When the representation of the discrete solution is allowed to be discontinuous in time, one needs to introduce additional residuals. These can be viewed as prism residuals with  $\Delta t \rightarrow 0$ , see [?, ?], and thus read

$$\psi_E = \lim_{\Delta t \to 0} \phi_{E_t} = \lim_{\Delta t \to 0} \int_{\partial E_t} \boldsymbol{f}_t \cdot \boldsymbol{n}_t \, \mathrm{d}(\partial \Omega_t) = \int_{\partial E} [u_h^n] \, \mathrm{d}\Omega = \frac{|E|}{3} \sum_{i \in E} [u_i^n], \tag{14}$$

where  $[\cdot]$  represents the jump through the face. The simple vertex-centred distribution,

$$\psi_{i,n}^{E} = \frac{|E|}{3} \left[ u_{i}^{n} \right], \tag{15}$$

was shown in [?, ?] to be positive and linearity preserving, and this is the formulation we use throughout this article. As a result, in the case when only discontinuities in time are introduced, the algebraic system takes the form

$$\sum_{E \in D_i} \left( \phi_{i,n}^E + \psi_{i,n}^E \right) = 0,$$

$$\sum_{E \in D_i} \phi_{i,n+1}^E = 0,$$
(16)

 $\forall i \in \Omega_h.$ 

RR n° 7958

### 2.3 Systems of equations

Much of the RD framework described for the scalar equation can be applied to the nonlinear hyperbolic system of conservation laws,

$$\partial_t U + \nabla \cdot \boldsymbol{F} = 0 \quad \text{or} \quad \partial_t U + \boldsymbol{\mathcal{A}}(U) \cdot \nabla U = 0,$$
(17)

where  $\mathcal{A}(U) = [\mathcal{A}_x, \mathcal{A}_y] = [\partial F_x/\partial U, \partial F_y/\partial U] = \partial F/\partial U$  is the wave-speed tensor. The two key elements where the extension is not so straightforward are positivity and upwinding. Positivity of the general system (17) is not clearly defined; instead, an algorithmic property that preserves the nonnegativity of certain physical quantities – water height, density, pressure, concentration, etc. – are required.

The definitions of the upwind directions for (17) are not immediately apparent, either. Different approaches have been proposed, the most popular of which is matrix distribution [?], which we also adopt in the current work. For space-time prisms, the inflow parameters used for the prism distribution are then defined as

$$\mathcal{K}_{i,n} = -\frac{\Delta t}{4} \overline{\mathcal{A}} \cdot \mathbf{n}_i - \frac{|E|}{3} \mathcal{I},$$

$$\mathcal{K}_{i,n+1} = -\frac{\Delta t}{4} \overline{\mathcal{A}} \cdot \mathbf{n}_i + \frac{|E|}{3} \mathcal{I},$$
(18)

where  $\mathcal{I}$  is the identity matrix and  $\overline{\mathcal{A}}$  represents a prism-averaged state of the flux Jacobian  $\mathcal{A}$ .

#### 2.4 Conservative linearisation for systems of equations

As in the scalar case,  $\overline{\mathcal{A}}$  is preferably derived from a conservative linearisation. For systems of equations, however, the exact form of the conservative linearisation is rarely straightforward. It often assumes linear variation in quantities other than the conservative variables U. The set of linearly varying quantities is traditionally called the Roe-parameter vector. The (vector-valued) discrete residuals of the system (17) can then be computed as

$$\Phi_{E_t} = \sum_{i \in E} \mathcal{K}_{i,n} \widetilde{U}_i^n + \sum_{i \in E} \mathcal{K}_{i,n+1} \widetilde{U}_i^{n+1} \quad \text{with} \quad \widetilde{U}_i^n = \frac{\partial U}{\partial Z} Z_i^n, \quad \widetilde{U}_i^{n+1} = \frac{\partial U}{\partial Z} Z_i^{n+1}, \tag{19}$$

where Z is the vector of the Roe-parameter variables [?, ?] and the derivative matrix  $\frac{\partial U}{\partial Z}$  is evaluated at the arithmetic mean states of the parameter-vector variables over the space-time prism.

If the conservative linearisation is not known or its implementation is not practical, it is also possible to achieve conservation by using a sufficiently accurate quadrature rule to compute the prism residual [?, ?].

#### 2.5 Upwind RD schemes for systems of equations

Assuming that the inflow matrices defined in (18) are diagonalisable, as is the case in almost all physical applications, we have  $\mathcal{K}_{i,n} = \mathcal{RDR}^{-1}$ , where  $\mathcal{D}$  is the diagonal matrix of eigenvalues,  $\mathcal{R}^{-1}$  is the matrix of the left and  $\mathcal{R}$  of the right eigenvectors. Defining  $\mathcal{D}^{\pm} = \frac{1}{2} (\mathcal{D} \pm |\mathcal{D}|)$  with  $|\mathcal{D}|$  denoting the absolute values of the entries, we can introduce the other inflow matrix parameters (cf. (7)),

$$\mathcal{K}_{i,n}^{+} = (\mathcal{R}\mathcal{D}^{+}\mathcal{R}^{-1})_{i,n}, \quad \mathcal{K}_{i,n+1}^{+} = (\mathcal{R}\mathcal{D}^{+}\mathcal{R}^{-1})_{i,n+1}, \\
\mathcal{K}_{i,n}^{-} = (\mathcal{R}\mathcal{D}^{-}\mathcal{R}^{-1})_{i,n}, \quad \mathcal{K}_{i,n+1}^{-} = (\mathcal{R}\mathcal{D}^{-}\mathcal{R}^{-1})_{i,n+1}, \\
\mathcal{N}_{t} = \left(\sum_{i \in E} \mathcal{K}_{i,n}^{+} + \sum_{i \in E} \mathcal{K}_{i,n+1}^{+}\right)^{-1}.$$
(20)

These, in turn, are used to define the upwind RD schemes along the lines of the scalar case (10)-(12).

• The space-time N (STN) scheme for systems is now defined as

$$\widetilde{U}_{t}^{\mathrm{in}} = \mathcal{N}_{t} \left( \sum_{j \in E} \mathcal{K}_{j,n}^{+} \widetilde{U}_{j}^{n} + \sum_{j \in E} \mathcal{K}_{j,n+1}^{+} \widetilde{U}_{j}^{n+1} - \Phi_{E_{t}} \right),$$

$$\left( \Phi_{i,n}^{E} \right)^{N} = \mathcal{K}_{i,n}^{+} \left( \widetilde{U}_{i}^{n} - \widetilde{U}_{t}^{\mathrm{in}} \right), \qquad \left( \Phi_{i,n+1}^{E} \right)^{N} = \mathcal{K}_{i,n+1}^{+} \left( \widetilde{U}_{i}^{n+1} - \widetilde{U}_{t}^{\mathrm{in}} \right),$$

$$(21)$$

where  $\tilde{U}$  is defined as in (19) if a conservative linearisation exists and thus  $\Phi_{E_t}$  is computed as in (19). On the other hand,  $\tilde{U} = U$  if the CRD formulation is used and  $\Phi_{E_t}$  is computed by means of a quadrature rule. In the system case, positivity should still be understood as positivity of the underlying linear advection. For nonlinear systems, positivity in the strict sense is not and should not be satisfied. Instead, the more generally property that the solution is free of spurious oscillations is required.

• The space-time LDA (STLDA) scheme for systems is defined as

$$\left(\Phi_{i,n}^{E}\right)^{LDA} = \mathcal{K}_{i,n}^{+} \mathcal{N}_{t} \Phi_{E_{t}}, \qquad \left(\Phi_{i,n+1}^{E}\right)^{LDA} = \mathcal{K}_{i,n+1}^{+} \mathcal{N}_{t} \Phi_{E_{t}}.$$
(22)

• *The space-time blended* (STB) scheme for systems is still defined as a combination of STN and the STLDA schemes,

$$\left(\Phi_{i,n}^{E}\right)^{B} = \Theta\left(\Phi_{i,n}^{E}\right)^{N} + \left(\mathcal{I} - \Theta\right)\left(\Phi_{i,n}^{E}\right)^{LDA},$$

$$\left(\Phi_{i,n+1}^{E}\right)^{B} = \Theta\left(\Phi_{i,n+1}^{E}\right)^{N} + \left(\mathcal{I} - \Theta\right)\left(\Phi_{i,n+1}^{E}\right)^{LDA}.$$
(23)

The matrix blending parameter  $\Theta$  can now be computed in a number of different ways [?, ?]. The direct analogue of (13) is

$$\Theta_{1} = \operatorname{diag}\left(\frac{|\Phi_{E_{t}}|}{\sum_{i \in E} \left|\left(\Phi_{i,n}^{E}\right)^{N}\right| + \sum_{i \in E} \left|\left(\Phi_{i,n+1}^{E}\right)^{N}\right|}\right),\tag{24}$$

where the division should be understood as an elementwise operation. Another possibility, proposed in [?], is to choose a particular direction  $\boldsymbol{\xi} = (\xi_x, \xi_y)$  and compute the decomposition  $\boldsymbol{\mathcal{A}} \cdot \boldsymbol{\xi} = \mathcal{R}_{\xi} \mathcal{D}_{\xi} \mathcal{R}_{\xi}^{-1}$ . The blending (23) is then carried out on the 'characteristic' residuals

$$\Phi_{i,n}^{N} = \mathcal{R}_{\xi}^{-1} \Phi_{i,n}^{N}, \quad \Phi_{i,n+1}^{N} = \mathcal{R}_{\xi}^{-1} \Phi_{i,n+1}^{N}, \quad \Phi_{i,n}^{LDA} = \mathcal{R}_{\xi}^{-1} \Phi_{i,n}^{LDA}, \quad \Phi_{i,n+1}^{LDA} = \mathcal{R}_{\xi}^{-1} \Phi_{i,n+1}^{LDA}$$

with the blending parameter computed as

$$\Theta_{2} = \operatorname{diag}\left(\frac{\left|\sum_{i \in E} \Phi_{i,n}^{N} + \sum_{i \in E} \Phi_{i,n+1}^{N}\right|}{\sum_{i \in E} \left|\Phi_{i,n}^{N}\right| + \sum_{i \in E} \left|\Phi_{i,n+1}^{N}\right|}\right),\tag{25}$$

where we drop the superscript 'E' to avoid clutter. Finally, we calculate the blended residuals based on the original variables by  $\Phi^B_{i,n} = \mathcal{R}_{\xi} \Phi^B_{i,n}$  and  $\Phi^B_{i,n+1} = \mathcal{R}_{\xi} \Phi^B_{i,n+1}$ .

#### 2.6 Discontinuity in time for systems of equations

The treatment of the discontinuity in time is directly analogous to (15), and is defined as [?]

$$\Psi_i^E = \frac{|E|}{3} \left[ U_{i,n}^n \right],\tag{26}$$

which leads to the algebraic equations

$$\sum_{E \in D_i} \left( \Phi_{i,n}^E + \Psi_{i,n}^E \right) = 0,$$

$$\sum_{E \in D_i} \Phi_{i,n+1}^E = 0,$$
(27)

 $\forall i \in \Omega_h$ , which needs to be solved at each time step.

## 3 Application to the shallow-water system

In this article, we apply the space-time RD framework for systems of equations to the frictionless shallowwater equations with non-flat bottom topography, i.e. we seek solution to the system

$$\partial_t U + \nabla \cdot \boldsymbol{F}(U) + S(U) = 0 \quad \text{on} \quad \Omega_t,$$
(28)

where

$$U = \begin{bmatrix} d \\ du \\ dv \end{bmatrix}, \quad \mathbf{F} = \begin{bmatrix} F_x & F_y \end{bmatrix} = \begin{bmatrix} du & dv \\ du^2 + \frac{gd^2}{2} & duv \\ duv & dv^2 + \frac{gd^2}{2} \end{bmatrix}, \quad S = -\begin{bmatrix} 0 \\ gd\frac{\partial b(x,y)}{\partial x} \\ gd\frac{\partial b(x,y)}{\partial y} \end{bmatrix}.$$
 (29)

Here d is the water height,  $\mathbf{u} = (u, v)$  is the flow velocity and b is the height of the bottom topography. The level of the free surface is defined as  $\eta = d + b$ . In the case of  $b \equiv 0$ , the nonlinear system (28), with the variables (29), expresses the conservation of water height d and discharge  $d\mathbf{u}$  – it can, therefore, admit discontinuous solutions (hydraulic jumps). In order to capture these discontinuities correctly, the discretisation of (28) should also be done in a conservative manner.

#### 3.1 Conservative schemes

There are essentially two different ways to achieve a conservative RD scheme for (28). One is to derive a conservative linearisation for the flux Jacobian by assuming linear variation of the Roe-parameter vector [?, ?]. The other is to use the CRD formulation [?, ?], where a nonconservative linearisation of the conservative variables is used and conservation is achieved by computing the prism residuals via quadrature rules.

#### 3.1.1 A conservative linearisation

In the steady state, conservative linearisation of (28) can be obtained by assuming the linear variation of the Roe-parameter vector  $Z = \frac{1}{\sqrt{d}}U$ . Then

$$\overline{\frac{\partial U}{\partial Z}} = \begin{bmatrix} 2\overline{\sqrt{d}} & 0 & 0\\ \overline{\sqrt{du}} & \overline{\sqrt{d}} & 0\\ \overline{\sqrt{dv}} & 0 & \overline{\sqrt{d}} \end{bmatrix}$$

so that the evaluation of  $\overline{\mathcal{A}}$  at the following cell-averaged states [?], together with (19), result in a conservative discretisation,

$$\widetilde{u} = \frac{\overline{\sqrt{d}u}}{\overline{\sqrt{d}}}, \qquad \widetilde{v} = \frac{\overline{\sqrt{d}v}}{\overline{\sqrt{d}}}, \qquad \widetilde{d} = 0.9\overline{d} + 0.1\frac{\sqrt{d_1d_2d_3}}{\overline{\sqrt{d}}}.$$
(30)

Here  $\overline{\cdot}$  signifies the arithmetic mean of the values at the vertices (represented by the subscripts 1, 2, 3) of the triangle.

For the space-time prism, one can use the linearisation (30) at the bottom and top of the prism, and apply the trapezium rule in time to achieve conservation.

#### 3.1.2 The CRD approach

As an alternative to conservative linearisation, one can instead use the arithmetic means  $\overline{d}$ ,  $\overline{u}$ ,  $\overline{v}$  (over either triangles or prisms), and apply the CRD formulation. In this case, however, the positivity of the N and STN schemes is formally lost, even though it is often observed in numerical experiments.

#### 3.2 Hydrostatic balance: the C-property

The shallow-water system (28)–(29) is, by construction, in hydrostatic balance as this is one of the main assumptions in its derivation from the Navier-Stokes equation [?]. Almost all numerical discretisations preserve this property exactly over a flat bed. Over a general (even smooth) non-flat bed, however, the exact preservation of the hydrostatic balance (also called the C-property) is not always straightforward.

#### 3.2.1 The CRD approach

As it was proved in [?], as long as the water-height d is assumed to be linearly varying and so is the bottom topography b, linearity preserving CRD schemes also satisfy the C-property [?, ?]. This simply means that the source term's contribution to the prism residual needs to be computed as

$$\Phi^{b}_{E_{t}} = \Delta t \frac{g\overline{d}}{2} \sum_{i \in E} \begin{bmatrix} 0\\ b_{i}\boldsymbol{n}_{i} \end{bmatrix}.$$
(31)

The conventional definition of the STN scheme (21), however, does not satisfy the hydrostatic balance, which means that the corresponding B and STB schemes will also lack this property. So instead of (21) we need to apply a slightly modified version of the scheme. For the CRD formulation, it means that  $U = [d, du, dv]^T$  is replaced by  $V = [\eta, du, dv]^T$  so that we have

$$V_{t}^{\text{in}} = \mathcal{N}_{t} \left( \sum_{j \in E} \mathcal{K}_{j,n}^{+} V_{j}^{n} + \sum_{j \in E} \mathcal{K}_{j,n+1}^{+} V_{j}^{n+1} - \Phi_{E_{t}} \right),$$

$$\left( \Phi_{i,n}^{E} \right)^{N} = \mathcal{K}_{i,n}^{+} \left( V_{i}^{n} - V_{t}^{\text{in}} \right), \qquad \left( \Phi_{i,n+1}^{E} \right)^{N} = \mathcal{K}_{i,n+1}^{+} \left( V_{i}^{n+1} - V_{t}^{\text{in}} \right).$$
(32)

**Proposition.** The STN scheme (32) exactly satisfies the hydrostatic balance (C-property). As a result, the STB scheme defined in (12) also exactly satisfies the same property.

*Proof.* In each space-time prism, with dropping the superscript Es, the STN scheme (32) can be rewritten as [?]

$$\Phi_{i,n}^{N} = \Phi_{i,n}^{LDA} + \mathcal{K}_{i,n}^{+} \mathcal{N}_{t} \sum_{j \in E} \mathcal{K}_{j,n}^{+} \left( V_{i}^{n} - V_{j}^{n} \right) + \mathcal{K}_{i,n}^{+} \mathcal{N}_{t} \sum_{j \in E} \mathcal{K}_{j,n+1}^{+} \left( V_{i}^{n} - V_{j}^{n+1} \right),$$
  
$$\Phi_{i,n+1}^{N} = \Phi_{i,n+1}^{LDA} + \mathcal{K}_{i,n+1}^{+} \mathcal{N}_{t} \sum_{j \in E} \mathcal{K}_{j,n}^{+} \left( V_{i}^{n+1} - V_{j}^{n} \right) + \mathcal{K}_{i,n}^{+} \mathcal{N}_{t} \sum_{j \in E} \mathcal{K}_{j,n+1}^{+} \left( V_{i}^{n+1} - V_{j}^{n+1} \right).$$

The fluctuations  $\Phi_{i,n}^{LDA}$  and  $\Phi_{i,n+1}^{LDA}$  are exactly zero for  $V = [\eta, du, dv]^T = [\text{const}, 0, 0]^T$  as long as (31) is used to compute the source term's contribution to the prism residual. Each of the other terms is exactly zero because the assumed linear variation of  $U = [d, du, dv]^T$  and b implies the linear variation of  $V = [\eta, du, dv]^T$ , i.e.  $V = [\eta, du, dv]^T = [\text{const}, 0, 0]^T$ .

#### 3.2.2 Conservative linearisation

We are not aware of a conservative linearisation for the shallow-water equations that preserves the hydrostatic balance exactly as well as conserves water height and discharge.

Since the linear variation of  $\sqrt{d}$  is assumed in the conservative linearisation proposed in this work, it is natural to assume the linear variation of  $\sqrt{b}$ , too, and compute the source term's contribution in the prism residual as

$$\Phi^{b}_{E_{t}} = \Delta t g \widetilde{d} \sum_{i \in E} \left[ \frac{0}{\sqrt{b} \sqrt{b_{i}} \, \boldsymbol{n}_{i}} \right].$$
(33)

This formulation introduces a small error in the hydrostatic balance because in general  $\eta_i = d_i + b_i = \text{const} \Rightarrow \overline{\sqrt{d}}\sqrt{d_i} + \overline{\sqrt{b}}\sqrt{b_i} = \text{const}.$ 

Conversely, it is possible to achieve exact hydrostatic balance by assuming the linear variation of  $\sqrt{d}$ and  $\sqrt{\eta}$  (instead of  $\sqrt{d}$  and  $\sqrt{b}$ ) and compute the source term's contribution in the prism residual as

$$\Phi_{E_t}^b = \Delta t \frac{g\widetilde{d}}{2} \sum_{i \in E} \left[ \left( \sqrt{\eta} \left( \sqrt{\eta_i^n} + \sqrt{\eta_i^{n+1}} \right) - \sqrt{d} \left( \sqrt{d_i^n} + \sqrt{d_i^{n+1}} \right) \right) \boldsymbol{n}_i \right].$$
(34)

This formulation preserves the lake at rest,  $V = [\eta, du, dv]^T = [\text{const}, 0, 0]^T$ , exactly but in general introduces an error in the conservation of water height and discharge because the term

$$\sum_{i \in E} \left[ \left( \overline{\sqrt{\eta}} \left( \sqrt{\eta_i^n} + \sqrt{\eta_i^{n+1}} \right) - \overline{\sqrt{d}} \left( \sqrt{d_i^n} + \sqrt{d_i^{n+1}} \right) \right) \boldsymbol{n}_i \right]$$

does not stay constant in time even when b does.

We emphasise, however, that in both cases the errors are rather small and in most computational experiments they do not affect the results substantially.

## 4 Implementation details

In the numerical examples presented in this article, we assume the discrete representation to be discontinuous in time only. In this case, we need to solve the algebraic system (27), which is the space-time DRD discretisation of (28), at each time step. For this a simple pseudo-time-stepping algorithm is used

$$\begin{pmatrix} U_i^n \\ U_i^{n+1} \end{pmatrix}_{m+1} = \begin{pmatrix} U_i^n \\ U_i^{n+1} \end{pmatrix}_m - \frac{\tau}{s_i^t} \sum_{E \in D_i} \begin{pmatrix} \Phi_{i,n}^E + \Psi_{i,n}^E \\ \Phi_{i,n+1}^E \end{pmatrix}_m,$$
(35)

where  $s_i^t = \Delta t s_i$  is the volume of the dual space-time cell (with  $s_i$  being the volume of the spatial dual cell). The pseudo-time step  $\tau$  is given by

$$\tau = 0.9 \min_{i} \frac{s_i^t}{\sum_{E \in D_i} \rho(\mathcal{K}_{i,n+1}^+)}, \qquad \rho(\mathcal{K}_{i,n+1}^+) = \max \operatorname{diag} \mathcal{D}_{i,n+1}^+, \tag{36}$$

with  $\rho(\mathcal{M})$  denoting the spectral radius of a given matrix  $\mathcal{M}$ . In all the computations, we use the stopping criterion

$$\texttt{rel\_tol} = \frac{\| \left( \Psi^n, \Psi^{n+1} \right)_m^T \|_1}{\| \left( \Psi^n, \Psi^{n+1} \right)_0^T \|_1} < 10^{-3}.$$

Often, though, we also apply an intermediate criterion to freeze the blending parameter, i.e. to stop its recomputation and continue the iteration with a constant value. In our numerical experiments, this happens once rel  $tol = 10^{-1.5}$  is reached.

Only results with the CRD formulation are reported. We note, however, that for flat bottom topography, the scheme with the conservative linearisation provides identical results. Even when the bottom topography is not flat, we only observe significant difference between the results provided by the two different formulations when the exact satisfaction of the hydrostatic balance is investigated (cf. Section 5.3).

## 5 Numerical results

In this section, we present a number of time-dependent test cases – most of which are nonlinear – to validate the discontinuous schemes presented in the previous sections. Throughout the section, the representation is discontinuous in time, which results in an unconditionally stable and globally positive scheme.

#### 5.1 Circular dam break over wet bed

Dam-break examples are used to experimentally assess whether the numerical solution is free of spurious oscillations. The first dam-break example we consider is that of a circular dam with radius  $r = \sqrt{x^2 + y^2} = 60$  separating water levels d = 10 and d = 0.5 in a square basin. The computational domain is the top right quarter,  $\Omega = [0, 100]^2$ , of the entire basin with solid-wall boundary conditions. The blending parameter  $\Theta_2$  is used in this example.

The velocity field tends to be more sensitive to spurious oscillations than the water height so we plot Froude-number contours for the STB scheme.

Figure 1 shows that the STB scheme captures the discontinuity well without introducing spurious oscillations for CFL = 1, CFL = 2 and CFL = 4. Interestingly, however, the schemes with CFL = 2 and CFL = 4 seem to be less diffusive than the scheme with CFL = 1. Three-dimensional plots of the water height are given in Figure 2 and they reveal no discernible oscillations for CFL = 1, 2, 4, either.



Figure 1: Circular dam break. 30 Froude-number contour plots of the discontinuous STB scheme for CFL = 1 (top left), CFL = 2 (top right), CFL = 4 (bottom left) and CFL = 8 (bottom right).

#### 5.2 Improving the iterative convergence

Any increase in the time step for any implicit space-time scheme will only translate into gains in computational work if the schemes with different CFL numbers perform similarly in terms of iterative convergence at each time step. We use this test to investigate this property when the STB scheme is applied. At each time step, we integrate until a relative tolerance  $rel_tol = 10^{-3}$  is reached. This is typically more than enough to achieve the accuracy required by a space-time RD discretisation because we can use the result from the previous time step as an initial guess. We plot the convergence histories in Figure 3 for different CFL numbers for two slightly different blending strategies. In both cases the blending parameter  $\Theta_2$  is applied, but in the second instance we freeze the value after reaching the relative tolerance



Figure 2: Circular dam break. 3D water-height plots of the discontinuous STB scheme for CFL = 1 (top left), CFL = 2 (top right), CFL = 4 (bottom left) and CFL = 8 (bottom right).

 $rel_tol = 10^{-1.5}$  and carry on with the iteration using this constant value.



Figure 3: Circular dam break. Convergence of pseudo-time iteration for the space-time DRD scheme with CFL = 1, 2, 4, 8 when the blending parameter is computed at every pseudo-time step (left) and when it is frozen once rel\_tol =  $10^{-1.5}$  is reached (right).

CFL = 1.0	$\ \eta-1\ $	$\ u\ $	$\ v\ $
$L^1$	5.405934e - 19	1.255336e - 16	5.870068e - 17
$L^2$	8.000922 e - 18	5.165644e - 16	2.647004 e - 16
$L^{\infty}$	2.220446e - 16	1.034175e - 14	6.178231e - 15
$\overline{\text{CFL}} = 2.0$			
$L^1$	1.534435e - 18	9.505484e - 17	4.305477e - 17
$L^2$	1.309467 e - 17	4.054003e - 16	1.890289e - 16
$L^{\infty}$	2.220446e - 16	8.781880e - 15	4.230023e - 15
CFL = 4.0			
$L^1$	3.744022e - 18	1.078974e - 16	5.057157e - 17
$L^2$	2.054486e - 17	4.354319e - 16	2.091099e - 16
$L^{\infty}$	2.220446e - 16	8.131857e - 15	4.450119e - 15
CFL = 8.0			
$L^1$	6.180234e - 18	1.177755e - 16	5.675335e - 17
$L^2$	2.657800e - 17	4.540011e - 16	2.285096e - 16
$L^{\infty}$	2.220446e - 16	7.692615e - 15	5.239349e - 15

Table 1: Lake at rest. Errors at time t = 0.5 for the discontinuous STB scheme with different CFL numbers.

It is clear from the results that the step from CFL = 4 to CFL = 8 does not overall lead to any reduction in computational work. Doubling the time step is obviously beneficial, and moving to CFL = 4 can still provide efficiency gains as long as we freeze the blending parameter (at rel\_tol =  $10^{-1.5}$  in this example).

#### 5.3 Lake at rest: the preservation of the hydrostatic balance

This time-dependent example is to experimentally verify that the modified STN scheme (32) does indeed preserve the C-property, i.e. it is hydrostatically well-balanced [?, ?, ?, ?]. The blended scheme should inherit this property since it is a linear interpolation between the STN and the STLDA schemes. Also, the actual form of the blending parameter does not influence the lake-at-rest property so for the sake of brevity we only show results for the STB scheme with  $\Theta_1$ , which we freeze after rel\_tol =  $10^{-1.5}$ . In the numerical tests we assume a smooth bathymetric function,

$$b(x,y) = 0.8e^{-5(x-0.9)^2 - 50(y-0.5)^2},$$
(37)

and 'still-water' initial conditions,  $[\eta^0, u^0, v^0] = [1, 0, 0]$ , over the domain  $\Omega = [0, 2] \times [0, 1]$ , using a mesh with a typical edge resolution of h = 0.01. Table 1 shows the errors after integrating until t = 0.5. The errors remain around machine precision until the end of the time marching, which indicates a well-balanced scheme. The numbers are for test runs with solid-wall boundary conditions but weak characteristic boundary conditions provide near identical results.

As a second variant of this test case, we put a perturbation on the initial, 'still-water' condition,

$$\eta^{0} = \begin{cases} 1.01 & \text{if } 0.05 < x < 0.15 \\ 1 & \text{otherwise} \end{cases}$$
(38)

For this example, the boundary conditions are weakly enforced everywhere: symmetric (i.e. solid wall) for the bottom and top boundaries, and characteristic freestream for the left and right boundaries including all corners.

Figures 4–7 plot 50 contours of the free surface  $\eta$  for CFL = 1, 2, 4, 8, respectively, at times t = 0.24and t = 0.48, while Figure 8 shows the same along the line y = 0.5. They reveal that qualitatively the scheme with CFL = 1 and CFL = 2 appear to be the most accurate while increasing the CFL number further results in the solution being more diffusive. This phenomenon is partly because of the loss of accuracy associated with a large time step relative to the mesh size, and partly because of the lack of local positivity of space-time implicit schemes.

Nevertheless, some of the most important qualitative properties are satisfied for all CFL numbers: a) they preserve the C-property in front of the perturbation; and b) they capture the interaction between the gravitational wave and the non-flat bottom; c) they settle back to lake at rest after the wave has passed.



Figure 4: Perturbation to the lake at rest when solved with the discontinuous STB scheme with CFL = 1. 50 water-height contours between values 0.992 and 1.012 at t = 0.24 (left) and t = 0.48 (right) are plotted.



Figure 5: Perturbation to the lake at rest when solved with the discontinuous STB scheme with CFL = 2. 50 water-height contours between values 0.992 and 1.012 at t = 0.24 (left) and t = 0.48 (right) are plotted.

#### 5.4 Travelling vortex

To evaluate the accuracy and (grid) convergence properties of the STLDA and STB schemes, we include the example of a travelling vortex with known exact solution [?, ?]. Given a flat bottom topography, the exact velocity field is expressed as  $u_{\infty} + u'$ , with

$$\boldsymbol{u}' = \begin{cases} \Gamma \left( 1 + \cos(\omega r_c) \right) \left( y_c - y, x - x_c \right) & \text{if } \omega r_c < \pi \\ (0, 0) & \text{otherwise} \end{cases},$$

and  $u_{\infty}$  being constant. The constant  $\Gamma$  is the vortex intensity parameter,  $(x_c, y_c)$  are the coordinates of the centre of the vortex,  $r_c$  is the distance from the centre of the vortex, and  $\omega$  is the angular wave



Figure 6: Perturbation to the lake at rest when solved with the discontinuous STB scheme with CFL = 4. 50 water-height contours between values 0.992 and 1.012 at t = 0.24 (left) and t = 0.48 (right) are plotted.



Figure 7: Perturbation to the lake at rest when solved with the discontinuous STB scheme with CFL = 8. 50 water-height contours between values 0.992 and 1.012 at t = 0.24 (left) and t = 0.48 (right) are plotted.

frequency associated with the diameter of the vortex. The water height is then given as

$$d(r_c) = d_{\infty} + \begin{cases} \frac{1}{g} \left(\frac{\Gamma}{\omega}\right)^2 \left(\kappa(\omega r_c) - \kappa(\pi)\right) & \text{if } \omega r_c < \pi\\ 0 & \text{otherwise} \end{cases},$$

with

$$\kappa(x) = 2\cos(x) + 2x\sin(x) + \frac{1}{8}\cos(2x) + \frac{x}{4}\sin(2x) + \frac{3}{4}x^2$$

and  $d_{\infty} = 1$ .

For the grid-convergence study, we set  $u_{\infty} = (6,0)$ ,  $\Gamma = 15$ ,  $\omega = 4\pi$ , g = 9.80665 and use a sequence of five unstructured triangulations of the domain  $\Omega = [0,2] \times [0,1]$  with characteristic mesh sizes h = 1/10, 1/20, 1/40, 1/80, 1/160, respectively. At the initial state the centre of the vortex is at  $(x_c, y_c) = (0.5, 0.5)$  and the time marching stops at t = 1/6, when  $(x_c, y_c) = (1.5, 0.5)$ . Freestream characteristic boundary conditions are used everywhere.

Figure 9 shows grid convergence of the STLDA and STB schemes with CFL = 1, 2, 4, 8. Second-order accuracy is observed for the STLDA scheme, although in the case of CFL = 8 this is only reached at the finest mesh. Also, it is clearly between CFL = 4 and CFL = 8 that the larger time step has a significant effect on the accuracy of the scheme.

The convergence rate for the STB scheme is slightly suboptimal – at around 1.8 – but still better than existing results of nonlinear space-time RD schemes [?, ?]. The blending in this example is applied to the residuals of the characteristic variables  $\Theta_2$ .



Figure 8: Perturbation to the lake at rest when solved with the discontinuous STB scheme with CFL = 1, CFL = 2, CFL = 4 and CFL = 8. Slice plots are shown along the line y = 0.5 at t = 0.24 and t = 0.48.



Figure 9: Travelling vortex. Grid convergence for the discontinuous STLDA (left) and the discontinuous STB (right) schemes is plotted.

#### 5.5 Partial dam break

This example is similar to the one in Section 5.1 but has a more complex geometry. We consider the domain  $[0, 200]^2$  with a dam that separates water levels d = 10 and d = 5. The dam is situated in the region  $[95, 105] \times [0, 200]$  and it breaks between y = 95 and y = 170 at initial time t = 0. The computational domain is thus  $\Omega = [0, 200] \setminus \Omega_{dam}$ , where  $\Omega_{dam} = ((95, 105) \times (0, 95)) \cup ((95, 105) \times (170, 200))$ . An unstructured mesh with characteristic mesh size of  $h \approx 2$  is used and solid-wall boundary conditions are imposed everywhere. The blending parameter is defined on the characteristic values,  $\Theta_2$ , and we also freeze the parameter once rel\_tol =  $10^{-1.5}$  is reached. Figure 10 shows water-height contours at the end of the time integration t = 7.2, while Figure 11 shows slice plots along the line y = 135. The schemes capture both the rarefaction wave (left of the break) and the shock wave (right of the break) with acceptable accuracy and compare favourably to published results in the literature [?, ?]. As in the previous examples, however, the scheme with CFL = 8 is markedly less accurate than the schemes with lower CFL numbers.

#### 5.6 Circular dam break over nonsmooth bed

This test case considers a two-dimensional variant of the Riemann problem over discontinuous bottom topography, proposed in [?]. The computational domain is now  $\Omega = [0, 30]^2$  with bathymetric function,

$$b(x,y) = \begin{cases} 0 & \text{if } x+y < 30\\ 0.2 & \text{otherwise} \end{cases},$$

and initial condition,

$$\eta(x,y) = \begin{cases} 1.461837 & \text{if } r < 15\\ 0.308732 & \text{otherwise} \end{cases},$$

where  $r = \sqrt{(x^2 + y^2)}$  is the radius of the dam. Solid-wall boundary conditions are used at the left and bottom boundaries while homogeneous Neumann at the right and top ones. The characteristic mesh size is  $h \approx 0.3$  and the time integration stops at time t = 10. The largest value of the blending parameter,  $\theta_2^{max} = \max \Theta_2$ , is applied to all variables of the residual in order to achieve an additional stabilising effect. The simulation follows the wave hitting the underwater wall, then partially reflecting from it and partially moving forward and exiting the domain. There is also a stationary shock wave along the discontinuity of the bed.

30 contours of the free surface  $\eta$  are depicted in Figure 12 for CFL = 4. The figure shows four snapshots of the solution at intervals of exactly 2.5 in time. All three waves – the outgoing, the reflected



Figure 10: Partial dam break. 30 water-height contours between 4 and 9.95 for the discontinuous STB scheme with CFL = 1, 2, 4 and 8.

and the stationary – are well captured. The contour plots for the STB schemes with other CFL numbers are omitted for this example because they show very similar behaviour to what we observe in the previous dam-break problems: the schemes with CFL = 2 and CFL = 4 provide the best qualitative results while the one with CFL = 1 is more diffusive and the one with CFL = 8 is both more diffusive and more oscillatory. This general pattern is also apparent from Figure 13, which shows slice plots for all CFL numbers along the diagonal, x = y, of the domain.

# 6 Comparison with existing schemes

This section compares the space-time scheme developed in this work with two other RD discretisations that exist in the literature. Both of these schemes are implicit but also require a time-step restriction because the representation in continuous in time as well as in space. The development of the explicit RKRD scheme [?] for the shallow-water equations is ongoing work [?]. That scheme and its comparison to some implicit formulations will be reported elsewhere. The two other schemes we consider from the literature are the following.

- 1. The blended LDA-N scheme where the LDA scheme is defined as in [?]. This is a similar upwind blended scheme to what is developed in this article but it is not a space-time formulation.
- 2. The stabilised LLF scheme in [?]. This is currently one of the most robust RD schemes for timedependent shallow-water simulations. It is essentially a central scheme and it is computationally cheaper than an upwind schemes that is continuous in both space and time.



Figure 11: Partial dam break. Slice plots of water height at final time t = 7.2 along the line y = 135 for the discontinuous STB scheme with CFL = 1, 2, 4 and 8.

We choose three test cases to compare the space-time scheme investigated here with the two other schemes. The first is the travelling vortex in Section 5.4 to compare the computational work against accuracy. The second is the partial dam break in Section 5.5 and the third is the circular dam break over non-smooth bed topography in Section 5.6. We measure computational work in the total number of pseudo-time iterations over the entire time integration. This, however, does not reflect the fact that the upwind blended schemes are computationally more expensive per space-time prism than the LLFs scheme. The blended scheme requires about four times as much computational work per prism as the LLFs schemes, while the discontinuous STB schemes requires about eight times as much. To reflect this, we define one work unit as being the amount of computational work the LLFs scheme needs per space-time prism.

#### 6.1 Travelling vortex

This test case is used to compare different RD schemes based on the computational work needed to achieve a given accuracy. The set-up of the test case is the same as in Section 5.4. Tables 2, 3 and 4 show the computational performance of the blended scheme, the LLFs scheme and the discontinuous STB scheme with CFL = 4. The computational work is measured both in the total number of pseudo-time iterations and in work units defined above. Figure 14 plots accuracy achieved for computational work. The results indicate the discontinuous STB scheme (with CFL = 4 in this case) is the most computationally efficient overall.

#### 6.2 Partial dam break

We use this example, which is the same as the one in Section 5.5, to compare the computational performance of the space-time scheme with the two other schemes when discontinuities in the solution have to



22

Figure 12: Circular dam break over discontinuous bed. 30 free-surface contours for the discontinuous STB scheme with CFL = 4.

be captured. Figure 15 shows the same water-height contours for these two schemes as Figure 10 for the space-time blended scheme. The blended LDA-N scheme of Ferrante appears to give the better result of the two. The STB scheme with CFL = 1, 2, 4 are of comparable quality to these but the one with CFL = 8 is clearly inferior. Figure 16 shows slice plots to directly compare two of the discontinuous STB schemes with these methods.

The computational performance of the two schemes and that of the STB scheme with CFL = 1, 2, 4, 8 is listed in Table 5. The computational work is measured in three ways until the final time of the simulation, t = 7.2, is reached: as the number of physical-time steps; as the number of total pseudo-time iterations; and as the number of work units defined at the beginning of this section. By the first two measures, the discontinuous space-time scheme outperforms all other schemes even with CFL = 1. This is in part because the past-shield condition is often (but not always) larger than the CFL-type restrictions used in other schemes. But it is also because the iterative convergence of the STB scheme is relatively good. In terms of work units, the LLFs scheme is still less expensive but it is also less accurate, though accuracy can only be assessed qualitatively for this example.

In particular, moving from CFL = 1 to CFL = 2 provides the most gains by more than halving the total computational cost. The benefits of increasing the time step from CFL = 2 to CFL = 4 are less obvious if the mild deterioration in the quality of the numerical simulation is also taken into account. The performance of the STB scheme with CFL = 8 is clearly poor. In this test case, it seems to generate spurious modes that spoil both the iterative performance of the scheme and quality of the results.



Figure 13: Plots of the free surface along the line x = y. The slices are taken at t = 2.5, t = 5.0, t = 7.5 and t = 10.0. In each figure the results for CFL = 1, CFL = 2, CFL = 4 and CFL = 8 are shown.

#### 6.3 Circular dam break over nonsmooth bed

We use this relatively difficult dam break problem to compare the performances of the different schemes on both a quasi-uniform and a locally refined mesh. First, we compare the schemes on the same quasi-uniform mesh as used in Section 5.6. Table 6 lists the computational work associated with the continuous-in-time blended scheme, the LLFs scheme, and the space-time schemes with CFL = 1, 2, 4, 8. It shows that the computational cost, measured as the total number of pseudo-time iterations, is reduced to about a third of the LLFs scheme and to about half of blended scheme with the LDA part defined as in [?].

Second, since there is a stationary shock wave over the discontinuous bed topography, it is natural to use local refinement there. The typical edge length in the region where the refinement takes place is 0.1, which is one-third of the typical edge length elsewhere in the domain. We show results for the same test case but with a mesh that is locally refined where the stationary shock occurs. The space-time scheme is now run with CFL = 9, which would approximately correspond to CFL = 3 in the region where the mesh is not refined. Figures 17–19 show 30 free-surface contour plots for the blended scheme of Ferrante [?], for the LLFs scheme [?], and for the discontinuous space-time scheme developed here with CFL = 9, respectively. Slice plots of the free surface along the line x = y are shown in Figure 21, while 3D free-surface plots at t = 7.5 are shown in Figure 20.

The space-time scheme seems to be less oscillatory than the two other schemes studied in this test case. In particular, the LLFs scheme develops rather large spurious spikes along the stationary shock in the second half of the integration time interval (see the bottom two subfigures in Figure 18 and also the 3D plots in Figure 20). By contrast, the space-time DRD schemes do not suffer from these artifacts,

Mesh size	$\ d_{ex} - d_h\ _{L^2}$	No. time steps	Total no. iterations	Work units
1/10	3.9667e-01	41	1061	$2.01\mathrm{e}{+06}$
1/20	2.7576e-01	77	1857	$1.38\mathrm{e}{+07}$
1/40	9.2256e-02	163	3973	$1.18\mathrm{e}{+08}$
1/80	2.2355e-02	330	3957	$4.69\mathrm{e}{+08}$
1/160	5.7876e-03	649	6485	$3.07\mathrm{e}{+09}$

Table 2: Travelling vortex. Error and computational work for the blended LDAN scheme with Ferrante's LDA.

Table 3: Travelling vortex. Error and computational work for the LLFs scheme.

Mesh size	$\ d_{ex}-d_h\ _{L^2}$	No. time steps	Total no. iterations	Work units
1/10	4.1190e-01	40	1223	$5.80 e{+}05$
1/20	3.1413e-01	77	2297	$4.26\mathrm{e}{+06}$
1/40	1.0290e-01	163	3797	$2.82\mathrm{e}{+07}$
1/80	2.6203 e-02	330	4340	$1.29\mathrm{e}{+08}$
1/160	6.6451 e- 03	649	7788	$9.23\mathrm{e}{+08}$



Figure 14: Travelling vortex. Computational work needed to achieve a given accuracy for the LLFs scheme, for the blended non-space-time scheme and for the discontinuous STB scheme with CFL = 4.

albeit at the price of a somewhat larger numerical diffusion. Overall, the space-time DRD schemes appear to represent the physical processes most accurately.

Table 7 lists the computational work for five schemes, including three space-time DRD schemes. The space-time schemes need a relatively large number of iterations to converge in each physical-time step

Mesh size	$\ d_{ex} - d_h\ _{L^2}$	No. time steps	Total no. iterations	Work units
1/10	4.2018e-02	10	376	$1.43e{+}06$
1/20	3.5085e-02	18	781	$1.16\mathrm{e}{+07}$
1/40	1.2608e-02	38	1784	$1.06\mathrm{e}{+08}$
1/80	3.3537e-03	73	3578	$8.49\mathrm{e}{+08}$
1/160	8.6824 e-04	155	7414	$7.03\mathrm{e}{+}09$

Table 4: Travelling vortex. Error and computational work for the discontinuous STB with CFL = 4.



Figure 15: Partial dam break. 30 water-height contours between 4 and 9.95 for the blended scheme and the LLFs scheme at the end of the integration t = 7.2. This figure is directly comparable to Figure 10 for the space-time DRD scheme.

so the benefits to computational cost are less pronounced in this example. They are still about twice as computationally efficient as the other schemes when measured in total number of pseudo-time iterations. When measured in number of physical-time steps, the gain is about 7-10 times. This opens up the possibility of further efficiency gains if the performance of the pseudo-iterative algorithm is improved. Also, this example shows a bigger difference in accuracy between the space-time scheme and the two other schemes (in favour of the space-time schemes) than other dam break problems.

## 7 Concluding remarks and outlook

This article applies the framework of discontinuous residual distribution (DRD) to the shallow-water equations with non-flat bottom topography. The focus is on the space-time representation that is discontinuous in time only. This choice is motivated by the fact that discontinuity in time lifts the time-step restriction on the size of the space-time prism and thus results in an unconditionally stable discretisation.

As the numerical experiments demonstrate, we can indeed increase the time step of discontinuous space-time schemes and still retain the most important properties of the shallow-water system: conservation, linearity preservation, upwinding, hydrostatic balance and local positivity. In particular, we emphasise that our interest here is restricted to 'pure' upwinding and therefore no stabilisation term [?] is included in these discretisations.

Comparison to two other implicit RD schemes show that the space-time DRD algorithm developed in this work provides more accurate results (measured either quantitatively or qualitatively) than the currently available RD schemes for the shallow-water equations. The comparisons were carried out on both quasi-uniform and locally refined meshes. Depending on the particular test case and implementation, the best-performing space-time DRD schemes require around 7-10 times fewer physical-time steps than



Figure 16: Partial dam break. Slice plots of water height at final time t = 7.2 along the line y = 135 for the blended non-space-time scheme, for the LLFs scheme and for the discontinuous STB scheme with CFL = 2, 4.

the other schemes. This, on current implementation, translates into 2-4 times fewer pseudo-time iterations over the entire integration. In terms of estimated total computational work, the discontinuous space-time schemes are still less expensive for a given level of accuracy.

Although more accurate than other currently available RD schemes, it has also emerged from the study that the benefits from increasing the time step in the space-time DRD scheme has practical limitations. One of these is the loss of accuracy associated with larger time steps. Another is the relatively large number of pseudo-time iterations needed per physical-time step. It is possible to mitigate these drawbacks by focusing on improving the iterative convergence of the schemes with medium-sized time steps, say between CFL = 3 and CFL = 5 (or higher in case of local mesh refinement) for these genuinely time-varying cases. Improving the iterative convergence by applying a more advanced pseudo-time-stepping algorithm [?] is among the most attractive options. In particular, a further study should compare the computational performance of the space-time DRD scheme with that of the explicit RKRD scheme [?, ?] when applied to the same or similar test cases presented in this work.

Overall, the likeliest areas of application where the proposed scheme could prove advantageous will probably include a relatively large degree of stiffness that comes from the physical equations when the viscous term is included.

## Acknowledgement

We thank Philip Roe for his contribution to the conservative linearisation of the shallow-water equations.

Scheme	No. time steps	Average no. iterations	Total no. iterations	Work units
B–Ferrante	134	29.44	3945	$3.59\mathrm{e}{+08}$
LLFs	127	66.39	8432	$1.92\mathrm{e}{+08}$
STDRD1	105	34.39	3611	$6.57\mathrm{e}{+08}$
STDRD2	54	30.80	1663	$3.02\mathrm{e}{+08}$
STDRD4	28	52.32	1465	$2.66\mathrm{e}{+08}$
STDRD8	14	132.64	1857	$3.38\mathrm{e}{+08}$

Table 5: Partial dam break on a quasi-uniform mesh. Computational work is shown for two existing implicit RD schemes and the space-time scheme with CFL = 1, 2, 4, 8.



Figure 17: Circular dam break over discontinuous bed using a locally refined mesh. 30 free-surface contours for the blended scheme where the LDA part is computed as in [?].

# Appendix

Based on Figure 22, in this Appendix we give an estimate for the CFL-condition of the vertex-centred upwind FV scheme with explicit Euler time discretisation on structured triangulation. This is to be compared with the past-shield condition (8) of the space-time scheme.



Figure 18: Circular dam break over discontinuous bed using a locally refined mesh. 30 free-surface contours for the LLFs scheme taken from [?].

#### Case 1: advection along the axis

The upwind FV scheme gets contributions from edges 1, 2, and 3 (red, blue, and yellow in Figure 22). For a vertical speed, the angles between the speed and normals are

$$\theta_1 = \frac{\pi}{4}, \qquad \theta_2 = \frac{\pi}{8}, \qquad \theta_3 = \frac{\pi}{2} - \frac{\pi}{8}$$

One can similarly show that the edge lengths are

$$l_1 = l_4 = \frac{2}{3} \frac{\Delta x}{\cos(\frac{\pi}{4})}, \qquad l_2 = l_3 = l_5 = l_6 = \frac{2}{3} \frac{\Delta x}{\cos(\frac{\pi}{8})}.$$

As a consequence, the limiting explicit Euler time step from positivity analysis reads

$$\Delta t(al_1\cos(\theta_1) + al_2\cos(\theta_2) + al_3\cos(\theta_3)) = |C| = 6\left(\frac{1}{3}\frac{\Delta x^2}{2}\right),$$
$$a\Delta t\Delta x\left(\frac{2}{3} + \frac{2}{3} + \frac{2}{3}\tan(\frac{\pi}{8})\right) = \Delta x^2,$$
$$\frac{a\Delta t}{\Delta x} = \frac{3}{2(2 + \tan(\frac{\pi}{8}))} \approx 0.62132.$$

In the same case, one easily shows that the past-shield condition gives

$$\frac{a\Delta x}{2}\Delta t = 2\frac{|T|}{3} = \frac{1}{3}\Delta x^2 \Rightarrow \frac{a\Delta t}{\Delta x} = \frac{2}{3} \approx 0.666667.$$



Figure 19: Circular dam break over discontinuous bed using a locally refined mesh. 30 free-surface contours for the space-time blended scheme with CFL = 9.

### Case 2: advection along the diagonal

The upwind FV scheme gets contributions from edges 1, 2, and 6 (red, blue, and green in Figure 22). In this case, the angles between the speed and normals are

$$\theta_1 = 0, \qquad \theta_2 = \frac{\pi}{8}, \qquad \theta_6 = \frac{\pi}{2} - \frac{\pi}{8}$$

As a consequence, the limiting explicit Euler time step from positivity analysis reads

$$\Delta t(al_1 \cos(\theta_1) + al_2 \cos(\theta_2) + al_6 \cos(\theta_6)) = |C| = 6\left(\frac{1}{3}\frac{\Delta x^2}{2}\right)$$
$$a\Delta t\Delta x \left(\frac{2}{3\cos(\frac{\pi}{4})} + \frac{2}{3} + \frac{2}{3}\tan(\frac{\pi}{8})\right) = \Delta x^2$$
$$\frac{a\Delta t}{\Delta x} = \frac{3}{2(\frac{1}{\cos(\frac{\pi}{4})} + 1 + \tan(\frac{\pi}{8}))} \approx 0.53033$$

In the same case, one again shows that the past-shield condition still gives

$$\frac{a\Delta x}{2}\Delta t = 2\frac{|T|}{3} = \frac{1}{3}\Delta x^2 \Rightarrow \frac{a\Delta t}{\Delta x} = \frac{2}{3} \approx 0.666667.$$

Table 6: Circular dam break over non-smooth bed using a quasi-uniform mesh. Computational work is shown for two existing implicit RD schemes and the space-time scheme with CFL = 1, 2, 4, 8.

Scheme	No. time steps	Average no. iterations	Total no. iterations	Work units
B–Ferrante	438	20.80	9112	$8.48e{+}08$
LLFs	436	36.14	15757	$3.67\mathrm{e}{+08}$
STDRD1	343	26.78	9185	$1.71\mathrm{e}{+09}$
STDRD2	173	29.41	5088	$9.47\mathrm{e}{+08}$
STDRD4	88	55.48	4882	$9.08\mathrm{e}{+08}$
STDRD8	45	135.38	6092	$1.13\mathrm{e}{+09}$

Table 7: Circular dam break over non-smooth bed using a locally refined mesh. Computational work is shown for two existing implicit RD schemes and the space-time scheme with CFL = 3, 6, 9.

Scheme	No. time steps	Average no. iterations	Total no. iterations	Work units
B–Ferrante	1260	17.34	21849	$2.41\mathrm{e}{+09}$
LLFs	1202	17.52	21064	$5.82\mathrm{e}{+08}$
STDRD3	332	52.77	17518	$3.87\mathrm{e}{+09}$
STDRD6	169	66.30	11205	$2.48\mathrm{e}{+09}$
STDRD9	114	86.15	9821	$2.17\mathrm{e}{+09}$



Figure 20: Circular dam break over discontinuous bed using a locally refined mesh. 3D free-surface plots at t = 7.5 for the blended LDAN scheme, the LLFs scheme and the discontinuous STB scheme (CFL = 9).



Figure 21: Circular dam break over discontinuous bed using a locally refined mesh. Plots of the free surface along the line x = y. The slices are taken at t = 2.5, t = 5.0, t = 7.5 and t = 10.0. In each figure the results for the blended LDAN scheme, the LLFs scheme and the discontinuous STB scheme (CFL = 9) are shown.



Figure 22: Sketch of the vertex-centred finite volume scheme



#### RESEARCH CENTRE BORDEAUX – SUD-OUEST

351, Cours de la Libération Bâtiment A 29 33405 Talence Cedex Publisher Inria Domaine de Voluceau - Rocquencourt BP 105 - 78153 Le Chesnay Cedex inria.fr

ISSN 0249-6399