A FULLY DISCRETE, MASS-CONSERVATIVE ADAPTIVE DISCONTINUOUS SPECTRAL-ELEMENT METHOD FOR THE ONE-DIMENSIONAL RICHARDS EQUATION IN MIXED FORM

TRISTAN MONTOYA* AND RAYMOND J. SPITERI*

Abstract. We present and analyze a spatially and temporally adaptive high-order method for the one-dimensional Richards equation describing vertical water transport in unsaturated soil, with an emphasis on resolving sharp wetting fronts while satisfying a fully discrete water mass balance. The spatial discretization uses an arbitrary-order local discontinuous Galerkin spectral-element formulation with collocated Legendre–Gauss–Lobatto quadrature, and the time integration is based on a high-order Rosenbrock–Wanner method. We focus primarily on the mixed formulation of the Richards equation, which advances water content as the conserved variable while evaluating the nonlinear diffusion operator in terms of pressure head; the resulting fully discrete scheme satisfies the water mass balance in exact arithmetic, independent of spatial resolution and time step size, including under spatial and temporal adaptivity. In contrast, the pressure-head formulation advances pressure head through a nonlinear capacity function and satisfies the corresponding conservation identity only at the semi-discrete level. Numerical experiments examine convergence through spatial and temporal refinement studies, with mass conservation assessed under spatial and temporal adaptivity using the recently proposed solver flux output method (SFOM). The SFOM results confirm machine-precision mass conservation for the mixed formulation, whereas the pressure-head formulation exhibits mass balance errors above roundoff. Moreover, in the temporal refinement studies, the mixed formulation yields more accurate solutions than the pressure-head formulation for a prescribed time step size. The combination of high-order accuracy, spatial and temporal adaptivity, and fully discrete mass conservation makes the proposed mixed-form discontinuous spectral-element method a promising approach for simulating unsaturated flow in next-generation land models.

Key words. Richards equation, spectral-element method, local discontinuous Galerkin method, Rosenbrock–Wanner method, spatial and temporal adaptivity, fully discrete mass conservation

MSC codes. 65M50, 65M60, 65M70, 35K55, 86A05

1. Introduction.

The one-dimensional Richards equation [35, 34] describes the vertical transport of water in variably saturated soil under gravitational and capillary forces. It models conservation of water mass by relating the temporal change in water content to the spatial variation of the hydraulic flux, which depends nonlinearly on the pressure head and its gradient through the Darcy–Buckingham law. Despite its apparent simplicity, the Richards equation poses significant challenges for numerical methods due to the strong nonlinearity of the flux and storage terms as well as its transition from parabolic behaviour to an elliptic limit as the soil nears saturation. These features give rise to a wide range of spatial and temporal scales, particularly as sharp wetting fronts develop and propagate, making accurate resolution computationally expensive with fixed grids and constant time step sizes. Furthermore, the semi-discrete systems obtained from discretizing the Richards equation in space typically result in stiff systems of ordinary differential equations (ODEs) or differential-algebraic equations (DAEs), which require careful treatment to maintain stability and accuracy without resorting to prohibitively small time steps. The review papers by Farthing and Ogden [13], Zha *et al.* [41], and Chatzikamaris *et al.* [3] provide comprehensive discussions of the challenges encountered in the numerical solution of the Richards equation.

Numerical methods that combine adaptive spatial and temporal discretization strategies can significantly reduce the computational cost of subsurface flow simulations

*Department of Computer Science, University of Saskatchewan, Saskatoon, SK, Canada (tristan.montoya@usask.ca, spiteri@cs.usask.ca).

while maintaining the accuracy needed to resolve sharp wetting fronts and other localized phenomena (see, for example, Miller *et al.* [26]). Relative to the second-order spatial discretization schemes commonly used in hydrologic modelling, high-order methods can achieve more accurate solutions for a given number of degrees of freedom in smooth solution regimes, with discontinuous Galerkin methods showing particular promise in the context of the Richards equation due to their local conservation properties and compact stencils (see, for example, Li *et al.* [23, 24], Dolejší *et al.* [10, 12], Clément *et al.* [5], and Tian *et al.* [38]). However, ensuring an exact fully discrete water mass balance that is independent of the time step size and preserved through adaptive mesh refinement and coarsening remains a significant challenge, even for spatial discretizations that are locally conservative at the semi-discrete level. This work addresses these challenges by developing and analyzing a spatially and temporally adaptive high-order method that ensures exact fully discrete mass conservation over the complete simulation, including both implicit time integration and mesh adaptation. The method combines the following key components:

- An arbitrary-order spectral-element approximation with collocated Legendre–Gauss–Lobatto (LGL) quadrature (see, for example, Kopriva [22]) is used together with a local discontinuous Galerkin (LDG) formulation [7], resulting in a compact spatial stencil and a diagonal element mass matrix.
- The mixed formulation of the Richards equation is used, in which water content is advanced as the conserved variable while the pressure head is retained within the spatial operator to evaluate the nonlinear flux. For comparison, we also consider the pressure-head formulation, in which pressure head is advanced through the nonlinear capacity function using the same spatial discretization; a purely water-content-based formulation is left for future work.
- Tree-based mesh adaptivity implemented in `Trixi.jl` [31] is used to concentrate resolution around sharp wetting fronts, where we employ a refinement indicator based on the total variation of the water content within each element together with a mass-conservative solution transfer procedure.
- A linearly implicit Rosenbrock–Wanner method [16] is used for time integration of the resulting stiff system of DAEs, providing high-order accuracy while avoiding iterative nonlinear solves and implementing adaptive time step control based on an embedded error estimate.
- The fully discrete conservation analysis uses the solver flux output method (SFOM) introduced by Ireson *et al.* [19] to track cumulative boundary fluxes with the same time integration method used to advance the solution. Under standard stage solvability assumptions, the resulting augmented DAE system preserves a linear invariant that guarantees an exact water mass balance in exact arithmetic for every accepted time step, which is preserved across mesh updates.
- Numerical experiments demonstrate high-order spatial and temporal convergence for a manufactured solution as well as mass conservation to roundoff under spatial and temporal adaptivity for canonical infiltration problems.

To our knowledge, the method presented in this work appears to be the first spatially and temporally adaptive high-order discretization for the Richards equation for which an exact fully discrete water mass balance is established, including time integration as well as mesh refinement and coarsening.

The remainder of this paper is organized as follows. Section 2 presents the mixed and pressure-head formulations of the Richards equation. Section 3 describes the spatial discretization, including the spectral-element approximation, LDG formulation,

spatial adaptivity, and mass-conservative solution transfer. Section 4 introduces the adaptive Rosenbrock–Wanner time integration and establishes the fully discrete water mass balance. Section 5 presents numerical experiments using an open-source Julia implementation of the proposed approach, in which we assess spatial and temporal convergence as well as mass conservation. Section 6 summarizes the main findings and discusses future work.

2. Mathematical formulation. We consider vertical flow in a one-dimensional soil column of length $L > 0$, with depth measured downward from the soil surface. Denoting the pressure head at depth $z \in [0, L]$ and time $t \geq 0$ by $\psi(z,t)$, the Richards equation is defined as

$$\partial_t \vartheta(\psi(z,t)) = \partial_z f(\psi(z,t), \partial_z \psi(z,t)), \quad z \in [0,L], \quad t > 0, \tag{2.1}$$

where the nonlinear flux function is defined by $f(\psi, \partial_z\psi) \coloneqq \kappa(\psi)(\partial_z\psi - 1)$, with ϑ and κ denoting the water content and hydraulic conductivity functions provided by a constitutive model. The unit gravitational-gradient contribution arises from the derivative of the total hydraulic head $\psi - z$, neglecting velocity head, under the convention that z is positive downward. In this work, we restrict our attention to pressure-head values in the unsaturated regime $\psi < 0$. Over this range, we assume that ϑ is twice continuously differentiable with a positive first derivative and that κ is continuously differentiable and non-negative.

We supplement (2.1) with an initial pressure-head profile $\psi(z,0)$. A Dirichlet boundary condition (i.e., a prescribed pressure head) or a Neumann boundary condition (i.e., a prescribed flux) may be imposed at the top or bottom of the soil column, and we assume that all boundary data are continuously differentiable in time. Integrating (2.1) in space and time, we see that the change in total volumetric water storage per unit cross-sectional area (hereafter referred to as the total water storage) is equal to the time-integrated net flux of water through its boundaries:

$$\begin{aligned} &\int_0^L \vartheta(\psi(z,t))\,\mathrm{d}z - \int_0^L \vartheta(\psi(z,0))\,\mathrm{d}z \\ &\quad = \int_0^t f(\psi(L,s), \partial_z\psi(L,s))\,\mathrm{d}s - \int_0^t f(\psi(0,s), \partial_z\psi(0,s))\,\mathrm{d}s. \end{aligned} \tag{2.2}$$

For constant water density and column cross-sectional area, the total water storage is proportional to the total water mass; accordingly, we refer to this balance as water mass conservation.

2.1. Mixed formulation. As written in (2.1), the conservative variable is the water content $\vartheta(\psi(z,t))$, not the pressure head $\psi(z,t)$. Because water content depends nonlinearly on pressure head, a time integration method that advances the pressure head and preserves only linear invariants of its state variables does not automatically conserve mass. Celia *et al.* [2] sought to address this issue by adopting a mixed formulation of the Richards equation, in which the water content is advanced in time as a conserved variable while the flux is evaluated in terms of the pressure head. To facilitate time integration using general-purpose DAE solvers, we retain water content and pressure head as separate solution variables, similarly to Kees and Miller [20], Cumming *et al.* [8], and Bootsma *et al.* [1]. Denoting the conserved water-content variable by $\theta(z,t)$, we express the mixed-form Richards equation as

$$\partial_t \theta(z,t) = \partial_z f(\psi(z,t), \partial_z\psi(z,t)), \tag{2.3a}$$

$$\theta(z,t) = \vartheta(\psi(z,t)), \tag{2.3b}$$

where (2.3b) is a pointwise algebraic constraint that relates the water content to the pressure head through the constitutive model.

2.2. Pressure-head formulation. For comparison, we also consider the pressure-head formulation of the Richards equation, which results from applying the chain rule to the time derivative in (2.1) to obtain

$$c(\psi(z,t))\partial_t\psi(z,t) = \partial_z f(\psi(z,t), \partial_z\psi(z,t)), \tag{2.4}$$

where the capacity function is given by $c(\psi) \coloneqq \vartheta'(\psi)$. Because $c(\psi) > 0$ over the range of pressure-head values considered, we can divide (2.4) by $c(\psi(z,t))$ to solve explicitly for the time derivative of the pressure head. Although the mixed and pressure-head formulations are equivalent at the continuous level for smooth solutions, the distinction between evolving $\theta(z,t)$ as a conserved quantity in (2.3) and evolving $\psi(z,t)$ according to (2.4) is central to the fully discrete mass conservation result discussed in subsection 4.4.

3. Spatial discretization. In this section, we describe the LDG spectral-element discretization used for both the mixed and pressure-head formulations.

3.1. Spectral-element approximation. Following a standard nodal spectral-element construction (as described, for example, in [22]), let the column $[0, L]$ be partitioned into K non-overlapping elements $[z_{k-1}, z_k]$ with $0 = z_0 < \cdots < z_K = L$. With the element size defined as $h_k \coloneqq z_k - z_{k-1}$, each element can be mapped from the reference interval $[-1, 1]$ through the affine transformation

$$X_k(\xi) \coloneqq \frac{h_k}{2}\xi + \frac{z_k + z_{k-1}}{2}, \tag{3.1}$$

which has a constant Jacobian given by $J_k = h_k/2$. Under the mapping $z = X_k(\xi)$, we approximate the pressure head and its gradient on each element as

$$\psi(X_k(\xi), t) \approx \psi_k^N(\xi, t) \coloneqq \sum_{i=0}^{N} \psi_{k,i}(t)\ell_i(\xi), \tag{3.2a}$$

$$\frac{\partial\psi}{\partial z}(X_k(\xi), t) \approx g_k^N(\xi, t) \coloneqq \sum_{i=0}^{N} g_{k,i}(t)\ell_i(\xi), \tag{3.2b}$$

where $N \geq 1$ is the chosen polynomial degree of the spatial approximation, and the Lagrange basis functions $\{\ell_i\}_{i=0}^N$ for the space $\mathbb{P}_N([-1, 1])$ of real-valued polynomials of degree at most N are defined for $N + 1$ distinct interpolation nodes $\{\xi_i\}_{i=0}^N$ as

$$\ell_i(\xi) \coloneqq \prod_{\substack{m=0 \\ m\neq i}}^{N} \frac{\xi - \xi_m}{\xi_i - \xi_m}, \quad i \in \{0, \ldots, N\}. \tag{3.3}$$

In this work, we choose $\{\xi_i\}_{i=0}^N$ to be the LGL nodes, which are the roots of $(1 - \xi^2)P_N'(\xi)$, where P_N is the Legendre polynomial of degree N. The corresponding positive quadrature weights $\{\omega_i\}_{i=0}^N$ are then given by

$$\omega_i \coloneqq \int_{-1}^{1} \ell_i(\xi)\,\mathrm{d}\xi = \frac{2}{N(N+1)}\frac{1}{P_N(\xi_i)^2}, \quad i \in \{0, \ldots, N\}, \tag{3.4}$$

resulting in a quadrature rule that is exact for all polynomials of degree at most $2N-1$. Adopting a zero-based indexing convention, we construct the collocation derivative matrix $\boldsymbol{D}$ and diagonal weight matrix $\boldsymbol{W}$ associated with the LGL quadrature rule as

$$D_{ij} \coloneqq \frac{\mathrm{d}\ell_j}{\mathrm{d}\xi}(\xi_i), \quad W_{ij} \coloneqq \omega_i \delta_{ij}, \quad i,j \in \{0,\dots,N\}. \tag{3.5}$$

Because the approximation space contains the constant function, the derivative matrix satisfies $\boldsymbol{D}\mathbf{1} = \mathbf{0}$, where $\mathbf{1} \coloneqq (1,\dots,1)^{\mathrm{T}}$ is a vector of ones and $\mathbf{0} \coloneqq (0,\dots,0)^{\mathrm{T}}$ is a vector of zeros, both of length $N+1$. Furthermore, as shown by Gassner [14], the derivative and quadrature weight matrices satisfy the summation-by-parts property

$$\boldsymbol{W}\boldsymbol{D} + \boldsymbol{D}^{\mathrm{T}}\boldsymbol{W} = \boldsymbol{e}_N \boldsymbol{e}_N^{\mathrm{T}} - \boldsymbol{e}_0 \boldsymbol{e}_0^{\mathrm{T}}, \tag{3.6}$$

where we define the vectors $\boldsymbol{e}_0 \coloneqq (1,0,\dots,0)^{\mathrm{T}}$ and $\boldsymbol{e}_N \coloneqq (0,\dots,0,1)^{\mathrm{T}}$.

3.2. Local discontinuous Galerkin method. We now construct the LDG discretization of the spatial terms shared by the mixed and pressure-head formulations, leaving the choice of evolved unknown to subsection 3.3. Following Cockburn and Shu [7], we first rewrite the second-order problem as a first-order system given by

$$g(z,t) = \partial_z \psi(z,t), \tag{3.7a}$$

$$\partial_t \vartheta(\psi(z,t)) = \partial_z f(\psi(z,t), g(z,t)). \tag{3.7b}$$

Expressing the system in reference coordinates using (3.1), inserting the polynomial approximations in (3.2), and integrating by parts against a test function then gives the following weak formulation in space: find $\psi_k^N(\cdot,t), g_k^N(\cdot,t) \in \mathbb{P}_N([-1,1])$ satisfying

$$\begin{aligned}\int_{-1}^{1} J_k w(\xi) g_k^N(\xi,t)\,\mathrm{d}\xi = &-\int_{-1}^{1} \partial_\xi w(\xi)\psi_k^N(\xi,t)\,\mathrm{d}\xi \\ &- w(-1)\psi_{k-1}^\star(t) + w(1)\psi_k^\star(t),\end{aligned} \tag{3.8a}$$

$$\begin{aligned}\int_{-1}^{1} J_k v(\xi)\partial_t \vartheta(\psi_k^N(\xi,t))\,\mathrm{d}\xi = &-\int_{-1}^{1} \partial_\xi v(\xi) f(\psi_k^N(\xi,t), g_k^N(\xi,t))\,\mathrm{d}\xi \\ &- v(-1)f_{k-1}^\star(t) + v(1)f_k^\star(t),\end{aligned} \tag{3.8b}$$

for all test functions $v, w \in \mathbb{P}_N([-1,1])$. The quantities $\psi_{k-1}^\star(t)$ and $\psi_k^\star(t)$ are unique numerical traces of the pressure head at the upper endpoint z_{k-1} and the lower endpoint z_k, respectively, of element k, and $f_{k-1}^\star(t)$ and $f_k^\star(t)$ are the corresponding numerical fluxes. Using a minimally dissipative LDG interface treatment with alternating traces and no interior penalty (see, for example, Cockburn and Dong [6]), we take the numerical pressure-head trace from the upper side of each interior interface as

$$\psi_k^\star(t) \coloneqq \psi_k^N(1,t), \quad k \in \{1,\dots,K-1\}, \tag{3.9}$$

and the numerical flux from the lower side as

$$f_k^\star(t) \coloneqq f(\psi_{k+1}^N(-1,t), g_{k+1}^N(-1,t)), \quad k \in \{1,\dots,K-1\}. \tag{3.10}$$

The numerical traces and fluxes at the top and bottom boundaries are determined by the chosen boundary conditions, which we assume to be of Dirichlet or Neumann

type. For instance, a Dirichlet boundary condition of the form $\psi(0,t) = \psi_{\mathrm{T}}(t)$ can be weakly imposed at the top boundary as

$$\psi_0^\star(t) \coloneqq \psi_{\mathrm{T}}(t), \tag{3.11a}$$

$$f_0^\star(t) \coloneqq f(\psi_1^N(-1,t), g_1^N(-1,t)) + \tau(\psi_{\mathrm{T}}(t), h_1, N)\left(\psi_1^N(-1,t) - \psi_{\mathrm{T}}(t)\right), \tag{3.11b}$$

where $\tau(\psi, h, N) \coloneqq N(N+1)\kappa(\psi)/h$ is a penalty parameter that provides a standard scaling of the pressure-head jump with respect to the polynomial degree N and element size h under collocated LGL quadrature (see, for example, Manzanero *et al.* [25, Appendix A]). At the bottom boundary, we can similarly impose $\psi(L,t) = \psi_{\mathrm{B}}(t)$ as

$$\psi_K^\star(t) \coloneqq \psi_{\mathrm{B}}(t), \tag{3.12a}$$

$$f_K^\star(t) \coloneqq f(\psi_K^N(1,t), g_K^N(1,t)) - \tau(\psi_{\mathrm{B}}(t), h_K, N)\left(\psi_K^N(1,t) - \psi_{\mathrm{B}}(t)\right), \tag{3.12b}$$

where we note that the penalty term is subtracted to account for the reversed orientation of the outward normal. Neumann boundary conditions of the form $f(\psi(0,t), g(0,t)) = f_{\mathrm{T}}(t)$ or $f(\psi(L,t), g(L,t)) = f_{\mathrm{B}}(t)$ are imposed by setting the numerical flux to the prescribed value and taking the numerical pressure-head trace from the adjacent interior state.

3.3. Semi-discrete formulation. To obtain a system of ODEs or DAEs, we take the test functions in (3.8) to be the Lagrange basis functions and approximate the integrals using the LGL quadrature rule. Evaluating the approximate pressure head and its gradient on element k at the LGL quadrature nodes as $\boldsymbol{\psi}_k(t) \coloneqq (\psi_{k,0}(t), \dots, \psi_{k,N}(t))^{\mathrm{T}}$ and $\boldsymbol{g}_k(t) \coloneqq (g_{k,0}(t), \dots, g_{k,N}(t))^{\mathrm{T}}$, respectively, we compute the nonlinearities pointwise as

$$\boldsymbol{f}_k(t) \coloneqq \left(f(\psi_{k,0}(t), g_{k,0}(t)), \dots, f(\psi_{k,N}(t), g_{k,N}(t))\right)^{\mathrm{T}}, \tag{3.13a}$$

$$\boldsymbol{C}_k(t) \coloneqq \operatorname{diag}\left(c(\psi_{k,0}(t)), \dots, c(\psi_{k,N}(t))\right). \tag{3.13b}$$

The remaining step is to choose the evolved unknown. The semi-discrete mixed formulation approximates the water content on element k as

$$\theta(X_k(\xi), t) \approx \theta_k^N(\xi, t) \coloneqq \sum_{i=0}^{N} \theta_{k,i}(t)\ell_i(\xi), \tag{3.14}$$

and evolves the vector $\boldsymbol{\theta}_k(t) \coloneqq (\theta_{k,0}(t), \dots, \theta_{k,N}(t))^{\mathrm{T}}$ of water-content values at the LGL quadrature nodes, resulting in a semi-discrete system given by

$$\boldsymbol{g}_k(t) = \frac{1}{J_k}\left(-\boldsymbol{W}^{-1}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\psi}_k(t) - \boldsymbol{W}^{-1}\boldsymbol{e}_0\psi_{k-1}^\star(t) + \boldsymbol{W}^{-1}\boldsymbol{e}_N\psi_k^\star(t)\right), \tag{3.15a}$$

$$\dot{\boldsymbol{\theta}}_k(t) = \frac{1}{J_k}\left(-\boldsymbol{W}^{-1}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{f}_k(t) - \boldsymbol{W}^{-1}\boldsymbol{e}_0 f_{k-1}^\star(t) + \boldsymbol{W}^{-1}\boldsymbol{e}_N f_k^\star(t)\right), \tag{3.15b}$$

subject to the algebraic constraint $\boldsymbol{\theta}_k(t) = (\vartheta(\psi_{k,0}(t)), \dots, \vartheta(\psi_{k,N}(t)))^{\mathrm{T}}$. The semi-discrete pressure-head formulation likewise follows from evolving the vector $\boldsymbol{\psi}_k(t)$ of pressure-head values through the capacity function as

$$\boldsymbol{g}_k(t) = \frac{1}{J_k}\left(-\boldsymbol{W}^{-1}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\psi}_k(t) - \boldsymbol{W}^{-1}\boldsymbol{e}_0\psi_{k-1}^\star(t) + \boldsymbol{W}^{-1}\boldsymbol{e}_N\psi_k^\star(t)\right), \tag{3.16a}$$

$$\boldsymbol{C}_k(t)\dot{\boldsymbol{\psi}}_k(t) = \frac{1}{J_k}\left(-\boldsymbol{W}^{-1}\boldsymbol{D}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{f}_k(t) - \boldsymbol{W}^{-1}\boldsymbol{e}_0 f_{k-1}^\star(t) + \boldsymbol{W}^{-1}\boldsymbol{e}_N f_k^\star(t)\right). \tag{3.16b}$$

The formulations in (3.15) and (3.16) therefore share the same LDG approximation of the nonlinear flux divergence, but they differ in the variable to which the time derivative is applied. We then obtain the following semi-discrete counterpart of (2.2).

THEOREM 3.1. *If the numerical flux is single-valued at every interior interface, then the mixed and pressure-head formulations satisfy the semi-discrete water mass balance*

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=1}^{K} J_k \mathbf{1}^{\mathrm{T}} \boldsymbol{W}\boldsymbol{\theta}_k(t) = f_K^\star(t) - f_0^\star(t), \tag{3.17}$$

where $\boldsymbol{\theta}_k(t)$ denotes the nodal water-content vector. In the mixed formulation, this vector is evolved directly; in the pressure-head formulation, it is evaluated from the evolved pressure-head vector according to $\boldsymbol{\theta}_k(t) \coloneqq (\vartheta(\psi_{k,0}(t)), \dots, \vartheta(\psi_{k,N}(t)))^{\mathrm{T}}$.

Proof. For the mixed formulation, left-multiplication of (3.15b) by $J_k \mathbf{1}^{\mathrm{T}} \boldsymbol{W}$ gives the local water mass balance

$$\begin{aligned} J_k \mathbf{1}^{\mathrm{T}} \boldsymbol{W}\dot{\boldsymbol{\theta}}_k(t) &= -\mathbf{1}^{\mathrm{T}} \boldsymbol{D}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{f}_k(t) - \mathbf{1}^{\mathrm{T}} \boldsymbol{e}_0 f_{k-1}^\star(t) + \mathbf{1}^{\mathrm{T}} \boldsymbol{e}_N f_k^\star(t) \\ &= -f_{k-1}^\star(t) + f_k^\star(t), \end{aligned} \tag{3.18}$$

where the second equality follows from the identities $\boldsymbol{D}\mathbf{1} = \mathbf{0}$ and $\mathbf{1}^{\mathrm{T}}\boldsymbol{e}_0 = \mathbf{1}^{\mathrm{T}}\boldsymbol{e}_N = 1$. For the pressure-head formulation, applying the chain rule in time to the nodal definition of $\boldsymbol{\theta}_k(t)$ gives $\dot{\boldsymbol{\theta}}_k(t) = \boldsymbol{C}_k(t)\dot{\boldsymbol{\psi}}_k(t)$, so (3.16b) yields the same elementwise water mass balance obtained in (3.18). For either formulation, summing this balance over all elements and moving the time derivative outside the sum yields

$$\frac{\mathrm{d}}{\mathrm{d}t}\sum_{k=1}^{K} J_k \mathbf{1}^{\mathrm{T}} \boldsymbol{W}\boldsymbol{\theta}_k(t) = \sum_{k=1}^{K} \left(f_k^\star(t) - f_{k-1}^\star(t) \right), \tag{3.19}$$

and hence the global semi-discrete balance in (3.17) follows from the fact that the single-valued numerical fluxes cancel at the interior interfaces. ∎

3.4. Spatial adaptivity. We use the hierarchical tree-based mesh structure and adaptivity mechanics implemented in `Trixi.jl` [31], although the indicator and conservative transfer procedure employed in this work are specifically tailored to the Richards equation. An h-adaptive approach is used, wherein refinement bisects an element into two children, coarsening merges two adjacent siblings into their parent, and the polynomial degree remains fixed across all elements. In order to track moving fronts in the unsaturated regime, the elementwise refinement indicator is computed as the discrete total variation of the water content within a given element, normalized by the global nodal water-content range. We compute the refinement indicator for each element as

$$\eta_k(t) \coloneqq \frac{\sum_{i=0}^{N} \omega_i \left| \sum_{j=0}^{N} D_{ij} \vartheta(\psi_{k,j}(t)) \right|}{\theta_{\max}(t) - \theta_{\min}(t) + \epsilon}, \tag{3.20}$$

where $\epsilon > 0$ is a small regularization parameter that prevents division by zero, and we compute the minimum and maximum nodal water-content values over all elements as

$$\theta_{\min}(t) \coloneqq \min_{k=1}^{K} \min_{i=0}^{N} \vartheta(\psi_{k,i}(t)), \quad \theta_{\max}(t) \coloneqq \max_{k=1}^{K} \max_{i=0}^{N} \vartheta(\psi_{k,i}(t)). \tag{3.21}$$

A simple two-threshold rule is used to determine whether an element should be refined, coarsened, or left unchanged. Let l_k denote the current refinement level of element k, such that the size of the element is given by $h_k = 2^{-l_k} L$, and let $l_{\min}$ and $l_{\max}$ denote the prescribed minimum and maximum refinement levels, respectively, with $l_{\min} < l_{\max}$. For a coarsening threshold η_{c} and a refinement threshold η_{r}, with $\eta_{\mathrm{c}} < \eta_{\mathrm{r}}$, we mark element k for refinement if $l_k < l_{\max}$ and $\eta_k(t) > \eta_{\mathrm{r}}$, and we mark it for coarsening if $l_k > l_{\min}$ and $\eta_k(t) \leq \eta_{\mathrm{c}}$. We split each element that is marked for refinement, and we merge a pair of sibling elements only if both are marked for coarsening. Additional refinements are performed to ensure that the difference in refinement levels between neighbouring elements is at most one, and coarsening is only performed if the resulting mesh would satisfy this same condition.

3.5. Mass-conservative solution transfer. When an element is refined to produce two children, the parent polynomial on the reference element, which we denote for an arbitrary transfer variable as $\phi_k^N \in \mathbb{P}_N([-1,1])$, is split into two child polynomials given by

$$\phi_{k,\mathrm{L}}^N(\zeta) \coloneqq \phi_k^N\left(\frac{\zeta-1}{2}\right), \quad \phi_{k,\mathrm{R}}^N(\zeta) \coloneqq \phi_k^N\left(\frac{\zeta+1}{2}\right), \tag{3.22}$$

which are evaluated at the reference LGL nodes to obtain the nodal degrees of freedom after the mesh update. When two sibling elements are combined to coarsen the mesh, we compute the L^2 projection of the child transfer-variable polynomials onto the parent element. Specifically, given the child polynomials $\phi_{k,\mathrm{L}}^N, \phi_{k,\mathrm{R}}^N \in \mathbb{P}_N([-1,1])$, we solve for a parent polynomial $\phi_k^N \in \mathbb{P}_N([-1,1])$ such that

$$\begin{aligned}\int_{-1}^{1} \phi_k^N(\xi) v(\xi)\,\mathrm{d}\xi = \frac{1}{2}\int_{-1}^{1} \phi_{k,\mathrm{L}}^N(\zeta) v\left(\frac{\zeta-1}{2}\right)\mathrm{d}\zeta \\ + \frac{1}{2}\int_{-1}^{1} \phi_{k,\mathrm{R}}^N(\zeta) v\left(\frac{\zeta+1}{2}\right)\mathrm{d}\zeta\end{aligned} \tag{3.23}$$

holds for all $v \in \mathbb{P}_N([-1,1])$. This projection is evaluated using an $(N+1)$-point Legendre–Gauss quadrature rule, which is exact for the polynomial products involved. We now consider the case in which water content is the transfer variable, which was proposed in [26] as a means of conserving mass across mesh updates for an adaptive finite-difference spatial discretization. The following result shows that every such update preserves the discrete total water mass for our discontinuous spectral-element method.

THEOREM 3.2. *Let the transfer variable be the water-content polynomial defined in (3.14). For the mixed formulation, its nodal coefficients $\theta_{k,i}(t)$ are evolved directly according to (3.15b), whereas for the pressure-head formulation, the polynomial is obtained through Lagrange interpolation with nodal values $\theta_{k,i}(t) = \vartheta(\psi_{k,i}(t))$. If the solution transfer procedures in (3.22) and (3.23) are applied for a mesh update at time t with $\phi_k^N(\xi) = \theta_k^N(\xi,t)$, $\phi_{k,\mathrm{L}}^N(\zeta) = \theta_{k,\mathrm{L}}^N(\zeta,t)$, and $\phi_{k,\mathrm{R}}^N(\zeta) = \theta_{k,\mathrm{R}}^N(\zeta,t)$, every adaptive mesh update then preserves the total water storage in (3.17).*

Proof. First, we consider a parent element that is refined into two child elements. Denoting the Jacobian of the parent-element mapping by J_k, noting that the Jacobians of the child-element mappings are given by $J_{k,\mathrm{L}} = J_{k,\mathrm{R}} = J_k/2$, and using the fact that the polynomials $\theta_k^N(\cdot,t), \theta_{k,\mathrm{L}}^N(\cdot,t), \theta_{k,\mathrm{R}}^N(\cdot,t) \in \mathbb{P}_N([-1,1])$ are all integrated exactly by the LGL quadrature rule, we obtain the total water storage within the two child

elements after refinement as

$$
\begin{aligned}
&J_{k,\mathrm{L}}\mathbf{1}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\theta}_{k,\mathrm{L}}(t) + J_{k,\mathrm{R}}\mathbf{1}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\theta}_{k,\mathrm{R}}(t)\\
&\qquad = \frac{J_k}{2}\int_{-1}^{1}\theta_{k,\mathrm{L}}^{N}(\zeta,t)\,\mathrm{d}\zeta + \frac{J_k}{2}\int_{-1}^{1}\theta_{k,\mathrm{R}}^{N}(\zeta,t)\,\mathrm{d}\zeta\\
&\qquad = \frac{J_k}{2}\int_{-1}^{1}\theta_k^{N}\left(\frac{\zeta-1}{2},t\right)\mathrm{d}\zeta + \frac{J_k}{2}\int_{-1}^{1}\theta_k^{N}\left(\frac{\zeta+1}{2},t\right)\mathrm{d}\zeta\\
&\qquad = J_k\int_{-1}^{0}\theta_k^{N}(\xi,t)\,\mathrm{d}\xi + J_k\int_{0}^{1}\theta_k^{N}(\xi,t)\,\mathrm{d}\xi\\
&\qquad = J_k\int_{-1}^{1}\theta_k^{N}(\xi,t)\,\mathrm{d}\xi\\
&\qquad = J_k\mathbf{1}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\theta}_k(t),
\end{aligned}
\tag{3.24}
$$

which is the total water storage within the parent element before refinement. Second, we consider a pair of sibling elements that are combined under coarsening to form a single parent element. Setting $v = 1$ and taking the water content as the transfer variable in (3.23) gives

$$
\int_{-1}^{1}\theta_k^{N}(\xi,t)\,\mathrm{d}\xi = \frac{1}{2}\int_{-1}^{1}\theta_{k,\mathrm{L}}^{N}(\zeta,t)\,\mathrm{d}\zeta + \frac{1}{2}\int_{-1}^{1}\theta_{k,\mathrm{R}}^{N}(\zeta,t)\,\mathrm{d}\zeta. \tag{3.25}
$$

Multiplying both sides by J_k and using the exactness of the LGL quadrature for the polynomial integrands, we therefore obtain

$$
J_k\mathbf{1}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\theta}_k(t) = J_{k,\mathrm{L}}\mathbf{1}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\theta}_{k,\mathrm{L}}(t) + J_{k,\mathrm{R}}\mathbf{1}^{\mathrm{T}}\boldsymbol{W}\boldsymbol{\theta}_{k,\mathrm{R}}(t), \tag{3.26}
$$

which shows that the water storage within the parent element obtained after coarsening is equal to the sum of the child-element storage contributions before coarsening. Because the contributions from elements that are neither refined nor coarsened remain unchanged by the mesh update, summing the contributions over all elements before and after the mesh update demonstrates that the total water mass is preserved. ∎

Remark 3.3. Because $\vartheta(\psi)$ is strictly increasing for $\psi < 0$, its inverse is well defined on the corresponding range of water-content values. However, conservation alone does not ensure that the transferred nodal water-content values lie in the domain of the inverse constitutive relation. We assume that this admissibility condition holds after each mesh update and reconstruct the nodal pressure-head vector as $\boldsymbol{\psi}_k(t) = (\vartheta^{-1}(\theta_{k,0}(t)),\ldots,\vartheta^{-1}(\theta_{k,N}(t)))^{\mathrm{T}}$ for every element k on the updated mesh.

4. Temporal discretization. By concatenating the elementwise unknowns in (3.15) or (3.16) to obtain a global state vector $\boldsymbol{y}(t)$, we obtain a system of the form

$$
\boldsymbol{M}\dot{\boldsymbol{y}}(t) = \boldsymbol{\mathcal{F}}(\boldsymbol{y}(t),t), \tag{4.1}
$$

where, on each fixed mesh between adaptive mesh updates, $\boldsymbol{M}$ is a constant mass matrix, which may be singular, and $\boldsymbol{\mathcal{F}}$ is a nonlinear function of the state and time. A mesh update changes the dimensions of the state vector and mass matrix consistently with the updated mesh, which, strictly speaking, yields a new system of the form in (4.1) that is initialized through the transfer procedure described in subsection 3.5. When defining mappings on the state space, we omit the time dependence of the state variables; thus, $\boldsymbol{y}$ denotes a generic state, whereas $\boldsymbol{y}(t)$ denotes a solution trajectory.

The global state vector for the mixed formulation concatenates the global water-content and pressure-head vectors, which are defined, respectively, as

$$\boldsymbol{\Theta}(t) \coloneqq \begin{pmatrix} \boldsymbol{\theta}_1(t) \\ \vdots \\ \boldsymbol{\theta}_K(t) \end{pmatrix}, \quad \boldsymbol{\Psi}(t) \coloneqq \begin{pmatrix} \boldsymbol{\psi}_1(t) \\ \vdots \\ \boldsymbol{\psi}_K(t) \end{pmatrix}. \tag{4.2}$$

Letting $\boldsymbol{I}$ denote the identity matrix of size $K(N+1)$ and $\boldsymbol{0}$ denote the zero matrix of the same dimensions, the resulting DAE system is then given by (4.1) with

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{\Theta} \\ \boldsymbol{\Psi} \end{pmatrix}, \quad \boldsymbol{M} = \begin{pmatrix} \boldsymbol{I} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}, \quad \boldsymbol{\mathcal{F}}(\boldsymbol{y}, t) = \begin{pmatrix} \boldsymbol{\mathcal{R}}(\boldsymbol{\Psi}, t) \\ \boldsymbol{\Theta} - \boldsymbol{\vartheta}(\boldsymbol{\Psi}) \end{pmatrix}, \tag{4.3}$$

where $\boldsymbol{\mathcal{R}}(\boldsymbol{\Psi}, t)$ is the spatial residual obtained by computing the gradient approximation on each element using (3.15a) and using the gradient values to evaluate the elementwise residuals in (3.15b). The vector $\boldsymbol{\vartheta}(\boldsymbol{\Psi})$ is obtained by applying the constitutive relation to each entry in the global pressure-head vector $\boldsymbol{\Psi}$, and we initialize the water-content vector as $\boldsymbol{\Theta}(0) = \boldsymbol{\vartheta}(\boldsymbol{\Psi}(0))$, so that the algebraic constraint is satisfied exactly at the beginning of the simulation. Because the Jacobian of the algebraic block of $\boldsymbol{\mathcal{F}}(\boldsymbol{y}, t)$ with respect to the vector $\boldsymbol{\Psi}$ of algebraic variables is given by $-\boldsymbol{C}(\boldsymbol{\Psi})$, the DAE system is of index one under the assumption that all nodal capacity values are strictly positive.

For the pressure-head formulation, the state contains only $\boldsymbol{\Psi}(t)$. Under the assumptions in section 2, all nodal capacity values are strictly positive, so dividing both sides of (3.16b) by the nodal values of the capacity function and concatenating contributions from all elements results in an ODE system given by (4.1) with

$$\boldsymbol{y} = \boldsymbol{\Psi}, \quad \boldsymbol{M} = \boldsymbol{I}, \quad \boldsymbol{\mathcal{F}}(\boldsymbol{y}, t) = \boldsymbol{C}(\boldsymbol{\Psi})^{-1} \boldsymbol{\mathcal{R}}(\boldsymbol{\Psi}, t), \tag{4.4}$$

where $\boldsymbol{C}(\boldsymbol{\Psi})$ is a diagonal matrix with its diagonal entries obtained by evaluating the capacity function for each entry of $\boldsymbol{\Psi}$, and $\boldsymbol{\mathcal{R}}(\boldsymbol{\Psi}, t)$ is the same spatial residual vector as in the mixed formulation. It is also possible to retain the capacity function on the left-hand side (see, for example, Tocci *et al.* [39]), although we do not pursue this approach because it introduces a state-dependent mass matrix.

4.1. Rosenbrock–Wanner methods. We integrate (4.1) in time using an adaptive Rosenbrock–Wanner method, which is a linearly implicit high-order method for stiff systems of ODEs and index-one DAEs (see, for example, Hairer and Wanner [16, Sections IV.7 and VI.4]). An s-stage embedded Rosenbrock–Wanner method is characterized by the coefficients α_{ij}, γ_{ij}, b_i, and $\hat{b}_i$, with

$$\alpha_{ij} = 0 \quad \text{for } j \geq i, \quad \gamma_{ij} = 0 \quad \text{for } j > i, \quad \gamma_{ii} = \gamma \neq 0. \tag{4.5}$$

Given the solution $\boldsymbol{y}^n$ at time t^n, the method computes a trial step (which may be accepted or rejected) to advance the solution to time $t^n + \Delta t^n$, where Δt^n is the trial step size and $t^0 = 0$, as

$$\boldsymbol{y}^{n+1} = \boldsymbol{y}^n + \sum_{i=1}^{s} b_i \boldsymbol{k}_i^n, \quad \hat{\boldsymbol{y}}^{n+1} = \boldsymbol{y}^n + \sum_{i=1}^{s} \hat{b}_i \boldsymbol{k}_i^n, \tag{4.6}$$

where $\boldsymbol{y}^{n+1}$ is the primary approximation and $\hat{\boldsymbol{y}}^{n+1}$ is an embedded approximation used for error estimation, both of which are defined on the mesh at time t^n. At each

stage $i \in \{1, \dots, s\}$, the stage solution vector $\boldsymbol{k}_i^n$ is obtained by solving the linear system of stage equations

$$\begin{aligned}\left(\boldsymbol{M} - \gamma \Delta t^n \frac{\partial \boldsymbol{\mathcal{F}}}{\partial \boldsymbol{y}}(\boldsymbol{y}^n, t^n)\right)\boldsymbol{k}_i^n &= \Delta t^n \boldsymbol{\mathcal{F}}\left(\boldsymbol{y}^n + \sum_{j=1}^{i-1} \alpha_{ij}\boldsymbol{k}_j^n, t^n + \Delta t^n \sum_{j=1}^{i-1}\alpha_{ij}\right) \\ &\quad + \Delta t^n \frac{\partial \boldsymbol{\mathcal{F}}}{\partial \boldsymbol{y}}(\boldsymbol{y}^n, t^n)\sum_{j=1}^{i-1}\gamma_{ij}\boldsymbol{k}_j^n + (\Delta t^n)^2 \sum_{j=1}^{i}\gamma_{ij}\frac{\partial \boldsymbol{\mathcal{F}}}{\partial t}(\boldsymbol{y}^n, t^n),\end{aligned} \tag{4.7}$$

where the right-hand side $\boldsymbol{\mathcal{F}}$ is defined using the spatial discretization at time t^n, and we note that the Jacobian $\partial \boldsymbol{\mathcal{F}}/\partial \boldsymbol{y}$ and the time derivative $\partial \boldsymbol{\mathcal{F}}/\partial t$ are held fixed across all stages of each trial step. Although the analysis in this paper is presented in terms of a general s-stage method, all numerical experiments use the eight-stage, fifth-order `Rodas5P` scheme [36] implemented within the `OrdinaryDiffEqRosenbrock.jl` package from the `DifferentialEquations.jl` ecosystem [30].

4.2. Linear system and Jacobian evaluation. The Jacobian matrix for the mixed formulation in (4.3) has the block structure

$$\frac{\partial \boldsymbol{\mathcal{F}}}{\partial \boldsymbol{y}}(\boldsymbol{y}^n, t^n) = \begin{pmatrix} \boldsymbol{0} & (\partial \boldsymbol{\mathcal{R}}/\partial \boldsymbol{\Psi})(\boldsymbol{\Psi}^n, t^n) \\ \boldsymbol{I} & -\boldsymbol{C}(\boldsymbol{\Psi}^n) \end{pmatrix}. \tag{4.8}$$

The matrix on the left-hand side of (4.7) is therefore given by

$$\boldsymbol{M} - \gamma \Delta t^n \frac{\partial \boldsymbol{\mathcal{F}}}{\partial \boldsymbol{y}}(\boldsymbol{y}^n, t^n) = \begin{pmatrix} \boldsymbol{I} & -\gamma \Delta t^n (\partial \boldsymbol{\mathcal{R}}/\partial \boldsymbol{\Psi})(\boldsymbol{\Psi}^n, t^n) \\ -\gamma \Delta t^n \boldsymbol{I} & \gamma \Delta t^n \boldsymbol{C}(\boldsymbol{\Psi}^n) \end{pmatrix}, \tag{4.9}$$

where all blocks are diagonal except for the upper-right block, which remains sparse due to the local nature of the LDG discretization, with its particular sparsity pattern determined by the elementwise couplings arising from the LDG numerical trace and numerical flux. The sparse Jacobian is computed using forward-mode automatic differentiation with graph colouring [33, 27], where the sparsity pattern and colouring are specified according to the LDG discretization. The resulting matrix in (4.9) is factorized once per trial step using KLU, a sparse direct LU solver for non-symmetric linear systems [9]. The resulting factorization is then stored and used to compute $\boldsymbol{k}_i^n$ for each right-hand side in (4.7). The same sparse factorization approach is used for the pressure-head formulation, in which $\partial \boldsymbol{\mathcal{F}}/\partial \boldsymbol{y}$ retains the same sparsity pattern as $\partial \boldsymbol{\mathcal{R}}/\partial \boldsymbol{\Psi}$, because derivatives of the capacity function contribute only to the diagonal entries of the Jacobian.

Remark 4.1. Taking the Schur complement of the upper-left identity block in (4.9), we obtain the matrix $\gamma \Delta t^n \left(\boldsymbol{C}(\boldsymbol{\Psi}^n) - \gamma \Delta t^n (\partial \boldsymbol{\mathcal{R}}/\partial \boldsymbol{\Psi})(\boldsymbol{\Psi}^n, t^n)\right)$, which is nonsingular when the condition

$$|\gamma| \Delta t^n \left\| \boldsymbol{C}(\boldsymbol{\Psi}^n)^{-1} \frac{\partial \boldsymbol{\mathcal{R}}}{\partial \boldsymbol{\Psi}}(\boldsymbol{\Psi}^n, t^n) \right\| < 1 \tag{4.10}$$

holds under any consistent matrix norm (see, for example, Stewart and Sun [37, Chapter III, Theorem 2.5]). Because (4.10) can be ensured for sufficiently small Δt^n when the capacity function is strictly positive, the linear system in (4.7) is therefore solvable under the same conditions. A similar argument can be used to show that the linear system for the pressure-head formulation is solvable for a strictly positive, continuously differentiable capacity function and a sufficiently small time step, although the time step restriction may differ from that of the mixed form.

4.3. Adaptive time integration. After computing the trial-step update in (4.6), we estimate the temporal error in water content. For the mixed formulation, let $\boldsymbol{\Theta}^n$ and $\boldsymbol{\Theta}^{n+1}$ denote the water-content blocks of the primary approximation at the corresponding time levels, and let $\hat{\boldsymbol{\Theta}}^{n+1}$ denote the water-content block of the embedded approximation, all defined on the mesh at time t^n, before any possible mesh adaptation at time t^{n+1}. For the pressure-head formulation, we analogously define $\boldsymbol{\Theta}^n := \boldsymbol{\vartheta}(\boldsymbol{\Psi}^n)$, $\boldsymbol{\Theta}^{n+1} := \boldsymbol{\vartheta}(\boldsymbol{\Psi}^{n+1})$, and $\hat{\boldsymbol{\Theta}}^{n+1} := \boldsymbol{\vartheta}(\hat{\boldsymbol{\Psi}}^{n+1})$, where the constitutive relation is applied componentwise. As in Ranocha *et al.* [32, Section 2.3], we compute the normalized error estimate for the trial step as

$$E^{n+1} := \sqrt{\frac{1}{K^n(N+1)} \sum_{j=1}^{K^n(N+1)} \left| \frac{\Theta_j^{n+1} - \hat{\Theta}_j^{n+1}}{\texttt{atol} + \texttt{rtol} \max\left(\left|\Theta_j^n\right|, \left|\Theta_j^{n+1}\right|\right)} \right|^2}, \tag{4.11}$$

where `atol` and `rtol` are user-specified absolute and relative tolerances, respectively, and K^n denotes the number of elements in the mesh at time t^n. We accept the trial step and advance the solution to time t^{n+1} only if $E^{n+1} \leq 1$ and every nodal pressure-head value in the primary approximation is negative. After an accepted step, a proportional-integral controller (see, for example, Gustafsson *et al.* [15]) determines a ratio for the step-size update as

$$r^{n+1} := \begin{cases} r_{\max}^{n+1}, & E^{n+1} = 0, \\ \min\left(r_{\max}^{n+1}, \max\left(r_{\min}^{n+1}, \sigma (E^{n+1})^{-\beta_1} \max(E^n, E^0)^{\beta_2}\right)\right), & \text{otherwise}, \end{cases} \tag{4.12}$$

where $r_{\min}^{n+1} \in (0,1)$ and $r_{\max}^{n+1} \geq 1$ are the minimum and maximum permitted ratios between time steps, $\beta_1, \beta_2 > 0$ are the current-error and previous-error exponents, $\sigma \in (0,1]$ is a safety factor, and $E^0 > 0$ is a prescribed initial value for the error estimate. The time step size is then updated as

$$\Delta t^{n+1} := \begin{cases} \Delta t^n, & r_{\mathrm{db}}^- \leq r^{n+1} \leq r_{\mathrm{db}}^+, \\ r^{n+1} \Delta t^n, & \text{otherwise}, \end{cases} \tag{4.13}$$

where r_{db}^- and r_{db}^+ are the lower and upper thresholds, respectively, of a controller deadband that suppresses small changes in the step size, with $r_{\mathrm{db}}^- \leq 1 \leq r_{\mathrm{db}}^+$. If the final time would be overshot by the next time step, the step size is accordingly reduced to reach the final time exactly. If any nodal pressure-head value in the primary approximation is non-negative (i.e., a saturated state is reached), we reject the trial step and repeat the update with a smaller step size obtained by multiplying Δt^n by $r_{\min}^{n+1}$. If the pressure-head values are all negative but the normalized error estimate satisfies $E^{n+1} > 1$, we similarly reject the trial step but instead multiply Δt^n by $\max(r_{\min}^{n+1}, \sigma(E^{n+1})^{-\beta_1})$. An adaptive mesh update is performed after every n_{adapt} accepted time steps, with the solution transferred to the new mesh following the procedure described in subsection 3.5, preserving the current time step size and error estimate across the mesh update for use in the step-size controller. Although the controller accepts user-specified absolute minimum and maximum step sizes, we do not impose either bound in this work.

4.4. Fully discrete mass conservation. The proof of fully discrete mass conservation relies on the preservation of linear invariants by the Rosenbrock–Wanner method, which is well known in the context of ODEs (see, for example, Hundsdorfer and Verwer [18, Section II.1.5]). The following lemma demonstrates that this property straightforwardly extends to the DAE system in (4.1) for invariants of the form $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{y}$.

LEMMA 4.2. *Consider a system of the form (4.1) with a constant, possibly singular mass matrix $\boldsymbol{M}$ and a right-hand side $\boldsymbol{\mathcal{F}}(\boldsymbol{y}, t)$ that is continuously differentiable with respect to $\boldsymbol{y}$ and t. Suppose that there exists a constant vector $\boldsymbol{c}$ such that*

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{\mathcal{F}}(\boldsymbol{y}, t) = 0 \tag{4.14}$$

for all states $\boldsymbol{y}$ and times t within the domain of $\boldsymbol{\mathcal{F}}$. The Rosenbrock–Wanner method defined by the update formulas in (4.6) and the stage equations in (4.7) then preserves the linear invariant $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{y}$, provided that the stage equations are solvable.

Proof. Following [18], we differentiate (4.14) with respect to $\boldsymbol{y}$ and t to obtain

$$\boldsymbol{c}^{\mathrm{T}} \frac{\partial \boldsymbol{\mathcal{F}}}{\partial \boldsymbol{y}}(\boldsymbol{y}, t) = \boldsymbol{0}^{\mathrm{T}}, \quad \boldsymbol{c}^{\mathrm{T}} \frac{\partial \boldsymbol{\mathcal{F}}}{\partial t}(\boldsymbol{y}, t) = 0. \tag{4.15}$$

Left-multiplying the stage equations in (4.7) by $\boldsymbol{c}^{\mathrm{T}}$ and using (4.14) and (4.15), we see that each stage solution satisfies

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{k}_i^n = 0. \tag{4.16}$$

Left-multiplying the update formula for $\boldsymbol{y}^{n+1}$ in (4.6) by $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{M}$ and substituting (4.16), we therefore obtain $\boldsymbol{c}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{y}^{n+1} = \boldsymbol{c}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{y}^n$. ∎

The following theorem establishes the fully discrete counterpart of (2.2) for the mixed-form discontinuous spectral-element discretization with Rosenbrock–Wanner time integration, including nonzero boundary fluxes.

THEOREM 4.3. *Consider the mixed formulation under the assumptions of Theorem 3.1, advanced in time using the Rosenbrock–Wanner method defined by (4.6) and (4.7), with the adaptivity performed using the water-content solution transfer described in subsection 3.5, and assume that the stage equations are solvable at each time step (see Remark 4.1). To track the cumulative boundary fluxes, we follow [19] and introduce the passive SFOM variables $F_{\mathrm{T}}(t)$ and $F_{\mathrm{B}}(t)$, which, at the semi-discrete level, satisfy*

$$\dot{F}_{\mathrm{T}}(t) = f_{\mathrm{T}}^{\star}(\boldsymbol{\Psi}(t), t), \quad F_{\mathrm{T}}(0) = 0, \quad \dot{F}_{\mathrm{B}}(t) = f_{\mathrm{B}}^{\star}(\boldsymbol{\Psi}(t), t), \quad F_{\mathrm{B}}(0) = 0, \tag{4.17}$$

where $f_{\mathrm{T}}^{\star}(\boldsymbol{\Psi}(t), t)$ and $f_{\mathrm{B}}^{\star}(\boldsymbol{\Psi}(t), t)$ are the numerical flux values at the top and bottom boundaries expressed as functions of the state $\boldsymbol{\Psi}(t)$ and time t. We integrate (4.17) in time using the same Rosenbrock–Wanner stages, time steps, and step-acceptance decisions used to advance $\boldsymbol{y}(t)$ to obtain the discrete cumulative flux values F_{T}^n and F_{B}^n at time t^n, which are not affected by adaptive mesh updates. In exact arithmetic, every accepted time step then satisfies

$$\sum_{k=1}^{K^{n+1}} J_k^{n+1} \mathbf{1}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{\theta}_k^{n+1} - \sum_{k=1}^{K^n} J_k^n \mathbf{1}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{\theta}_k^n = \left(F_{\mathrm{B}}^{n+1} - F_{\mathrm{B}}^n\right) - \left(F_{\mathrm{T}}^{n+1} - F_{\mathrm{T}}^n\right), \tag{4.18}$$

where we use $\boldsymbol{\theta}_k^n$ to denote the contribution to the evolved water-content block $\boldsymbol{\Theta}^n$ from element k on the mesh obtained after any adaptive mesh update at time t^n.

Proof. Appending the cumulative boundary fluxes in (4.17) to the mixed-form DAE in (4.3) gives an augmented system that takes the form of (4.1) with

$$\boldsymbol{y} = \begin{pmatrix} \boldsymbol{\Theta} \\ \boldsymbol{\Psi} \\ F_{\mathrm{T}} \\ F_{\mathrm{B}} \end{pmatrix}, \quad \boldsymbol{M} = \operatorname{diag}(\boldsymbol{I}, \boldsymbol{0}, 1, 1), \quad \boldsymbol{\mathcal{F}}(\boldsymbol{y}, t) = \begin{pmatrix} \boldsymbol{\mathcal{R}}(\boldsymbol{\Psi}, t) \\ \boldsymbol{\Theta} - \boldsymbol{\vartheta}(\boldsymbol{\Psi}) \\ f_{\mathrm{T}}^{\star}(\boldsymbol{\Psi}, t) \\ f_{\mathrm{B}}^{\star}(\boldsymbol{\Psi}, t) \end{pmatrix}. \tag{4.19}$$

Consider first the time-integration update from t^n to t^{n+1} on the current mesh, before any scheduled adaptive mesh update at t^{n+1}. Suppressing time step index superscripts on mesh-related quantities, we define the constant vector

$$\boldsymbol{c} = \begin{pmatrix} J_1 \mathbf{1}^{\mathrm{T}} \boldsymbol{W} & \cdots & J_K \mathbf{1}^{\mathrm{T}} \boldsymbol{W} & \mathbf{0}^{\mathrm{T}} & 1 & -1 \end{pmatrix}^{\mathrm{T}}, \tag{4.20}$$

where, in this case, $\mathbf{0}^{\mathrm{T}}$ denotes a row vector of length $K(N+1)$, indicating the absence of direct contributions from the pressure-head block $\boldsymbol{\Psi}$. Contracting with the right-hand side $\boldsymbol{\mathcal{F}}(\boldsymbol{y}, t)$, we obtain

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{\mathcal{F}}(\boldsymbol{y}, t) = \begin{pmatrix} J_1 \mathbf{1}^{\mathrm{T}} \boldsymbol{W} & \cdots & J_K \mathbf{1}^{\mathrm{T}} \boldsymbol{W} \end{pmatrix} \boldsymbol{\mathcal{R}}(\boldsymbol{\Psi}, t) + f_{\mathrm{T}}^{\star}(\boldsymbol{\Psi}, t) - f_{\mathrm{B}}^{\star}(\boldsymbol{\Psi}, t), \tag{4.21}$$

which vanishes for every admissible state $\boldsymbol{y}$ and time t due to the elementwise balance and interface cancellation established in the proof of Theorem 3.1. Therefore, Lemma 4.2 guarantees that the quantity

$$\boldsymbol{c}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{y}(t) = \sum_{k=1}^{K} J_k \mathbf{1}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{\theta}_k(t) + F_{\mathrm{T}}(t) - F_{\mathrm{B}}(t) \tag{4.22}$$

is preserved when advancing the solution from t^n to t^{n+1} on a fixed mesh. Expanding and rearranging the resulting invariant identity yields (4.18) when the solutions at t^n and t^{n+1} are both represented on the mesh at t^n, such that $K^{n+1} = K^n$ and $J_k^{n+1} = J_k^n$. If a mesh update is performed after advancing the solution to t^{n+1}, the mass-conservative solution transfer in Theorem 3.2 preserves the water-content sum, and the cumulative boundary fluxes F_{T}^{n+1} and F_{B}^{n+1} remain unchanged. Therefore, (4.18) also holds after the solution transfer to the updated mesh, with K^{n+1} and J_k^{n+1} denoting the number of elements and the Jacobian for element k, respectively, on the updated mesh. ∎

5. Numerical experiments. The methods described in this paper are implemented in `HydroTrixi.jl`,[1] a Julia package that extends the `Trixi.jl` numerical simulation framework [31] to nonlinear parabolic problems for applications in hydrology. To examine the accuracy, convergence, and mass conservation properties of the mixed and pressure-head formulations described in this work, `HydroTrixi.jl` is used to solve a series of benchmark problems for the one-dimensional Richards equation, with all simulations using double-precision arithmetic. The pressure-head formulation is included for comparison purposes; the proposed fully discrete mass-conservative method uses the mixed formulation with water-content solution transfer.

5.1. Constitutive models. We consider both the Haverkamp [17] and van Genuchten–Mualem [40, 29] constitutive models. For the Haverkamp model, the water content and the hydraulic conductivity are given as functions of the pressure head by

$$\vartheta(\psi) = \theta_{\mathrm{r}} + \frac{\theta_{\mathrm{s}} - \theta_{\mathrm{r}}}{1 + (a_{\mathrm{H}}|\psi|)^{\beta_{\mathrm{H}}}}, \quad \kappa(\psi) = \frac{\kappa_{\mathrm{s}}}{1 + (b_{\mathrm{H}}|\psi|)^{\gamma_{\mathrm{H}}}}, \tag{5.1}$$

where θ_{r} and θ_{s} are the residual and saturated water-content values, respectively, with $\theta_{\mathrm{r}} < \theta_{\mathrm{s}}$, and $\kappa_{\mathrm{s}} > 0$ is the saturated hydraulic conductivity. The parameters $a_{\mathrm{H}} > 0$ and $b_{\mathrm{H}} > 0$ are inverse pressure-head scales, and $\beta_{\mathrm{H}} > 0$ and $\gamma_{\mathrm{H}} > 0$ are

[1] The source code is available at https://github.com/uofs-simlab/HydroTrixi.jl, and the scripts used to generate the numerical results are included in this paper's reproducibility repository.

Table 1: Parameters for the Haverkamp and van Genuchten–Mualem constitutive models used in the numerical experiments.

Symbol	Description	Value
Haverkamp constitutive model		
θ_r	Residual water content	0.075
θ_s	Saturated water content	0.287
κ_s	Saturated conductivity	$9.44 \times 10^{-5}\,\mathrm{m\,s^{-1}}$
a_H	Inverse pressure-head scale	$2.70740\,\mathrm{m^{-1}}$
b_H	Inverse pressure-head scale	$5.24084\,\mathrm{m^{-1}}$
β_H	Water-retention exponent	3.96
γ_H	Conductivity exponent	4.74
van Genuchten–Mualem constitutive model		
θ_r	Residual water content	0.102
θ_s	Saturated water content	0.368
κ_s	Saturated conductivity	$9.22 \times 10^{-5}\,\mathrm{m\,s^{-1}}$
α_V	Inverse pressure-head scale	$3.35\,\mathrm{m^{-1}}$
ℓ_V	Pore-connectivity parameter	0.5
n_V	Water-retention exponent	2

the water-retention and hydraulic-conductivity exponents, respectively. The van Genuchten–Mualem model is given by

$$\begin{aligned} \vartheta(\psi) &= \theta_r + (\theta_s - \theta_r)\left[1 + (\alpha_V|\psi|)^{n_V}\right]^{-m_V}, \\ \kappa(\psi) &= \kappa_s S_e(\psi)^{\ell_V}\left[1 - (1 - S_e(\psi)^{1/m_V})^{m_V}\right]^2, \end{aligned} \tag{5.2}$$

where $S_e(\psi) \coloneqq (\vartheta(\psi) - \theta_r)/(\theta_s - \theta_r)$ is the effective saturation, $\alpha_V > 0$ is an inverse pressure-head scale, $n_V > 1$ is a dimensionless water-retention exponent, ℓ_V is the dimensionless pore-connectivity parameter, and $m_V \coloneqq 1 - 1/n_V$ is a derived shape parameter. The parameter values used in this paper for the Haverkamp and van Genuchten–Mualem models are listed in Table 1. The Haverkamp values are chosen to reproduce the first infiltration benchmark considered by Celia *et al.* [2], whereas the van Genuchten–Mualem values correspond to the New Mexico soil properties used in their second infiltration benchmark.

5.2. Spatial and temporal convergence for a manufactured solution. We assess spatial and temporal convergence for fixed meshes and time step sizes using two variants of the manufactured-solution problem introduced by Keita *et al.* [21]. Both variants use the Haverkamp model with the parameters in Table 1 and prescribe the exact pressure-head profile

$$\psi(z,t) = 0.204\,\mathrm{m}\,\tanh\left[\frac{1}{2}\left(100\,\mathrm{m^{-1}}\,z + \frac{t}{12\,\mathrm{s}} - 15\right)\right] - 0.411\,\mathrm{m}, \tag{5.3}$$

on a column of length $L = 0.2\,\mathrm{m}$ for $t \in (0\,\mathrm{s}, 120\,\mathrm{s}]$. To produce the manufactured solution, a source term given by

$$s(z,t) \coloneqq c(\psi(z,t))\partial_t\psi(z,t) - \partial_z\left[\kappa(\psi(z,t))\left(\partial_z\psi(z,t) - 1\right)\right], \tag{5.4}$$

is added to the right-hand side of (2.1), where $\psi(z,t)$ is the manufactured solution profile given by (5.3). In both variants, we obtain the initial condition by evaluating this profile at $t = 0\,\mathrm{s}$, and we specify the corresponding exact pressure head $\psi(0\,\mathrm{m}, t)$ as a Dirichlet condition at the top boundary. At the bottom boundary, the first variant prescribes the exact pressure head $\psi(L,t)$, whereas the second prescribes the exact flux $f(\psi(L,t), \partial_z\psi(L,t))$. We refer to these variants as the *Dirichlet–Dirichlet* and *Dirichlet–Neumann* problems, respectively.

To examine spatial convergence, we consider fixed polynomial degrees $N \in \{3,4\}$ and perform successive uniform refinements from $l = 4$ to $l = 8$, corresponding to $K \in \{16, 32, 64, 128, 256\}$. We use a fixed time step size of $\Delta t = 0.1\,\mathrm{s}$ to ensure that temporal errors remain small compared to spatial errors. The nominal mesh size is defined by $\Delta z \coloneqq L/(KN)$, corresponding to the average distance between the LGL quadrature nodes in physical space, and we define the normalized quadrature-based L^2 error and the nodal L^∞ error in pressure head at $t = 120\,\mathrm{s}$ as

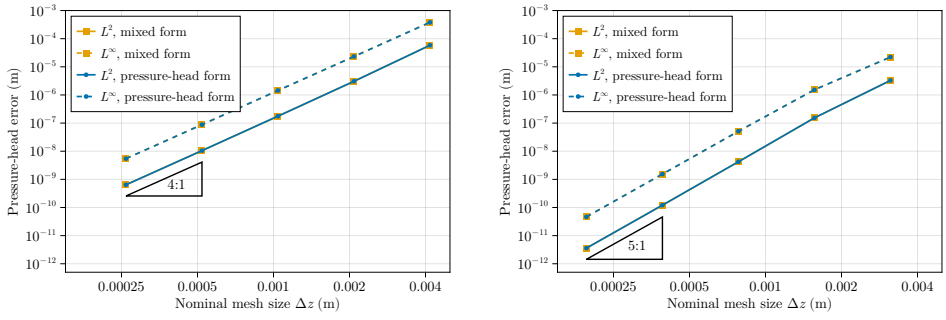
$$L^2 \text{ pressure-head error} \coloneqq \sqrt{\frac{1}{L}\sum_{k=1}^{K} J_k \sum_{i=0}^{N} \omega_i \left(\psi(X_k(\xi_i),t) - \psi_k^N(\xi_i,t)\right)^2}, \tag{5.5a}$$

$$L^\infty \text{ pressure-head error} \coloneqq \max_{k=1}^{K} \max_{i=0}^{N} \left|\psi(X_k(\xi_i),t) - \psi_k^N(\xi_i,t)\right|, \tag{5.5b}$$

respectively, where both norms return values in the same units as the pressure head (i.e., metres for the examples considered in this paper) under this normalization. In Figure 1, we plot both error norms against Δz for each combination of boundary condition configuration and polynomial degree, using the mixed and pressure-head formulations. In all cases considered, the L^2 and L^∞ error norms both converge at an order of about $N+1$, which is optimal when discretizing a smooth problem using a polynomial approximation of degree N. Moreover, the mixed and pressure-head formulations produce nearly identical pressure-head errors at every refinement level, which is to be expected because they make use of the same semi-discrete spatial operator but differ in the variables that they evolve in time.

To assess the effect of the choice between the mixed formulation and the pressure-head formulation on the accuracy of the scheme, we consider a fixed spatial discretization with $N = 3$ and $K = 4096$, which is fine enough to ensure that the spatial error is negligible compared to the temporal error, and we successively halve the time step to observe the temporal convergence for both formulations. The convergence of the pressure-head error for both boundary condition configurations is plotted in Figure 2 for fixed time step sizes ranging from $\Delta t = 4\,\mathrm{s}$ to $\Delta t = 0.25\,\mathrm{s}$. Between the smallest and second-smallest time step sizes, both formulations exhibit temporal rates of convergence between fourth and fifth order, consistent with the reduction in order noted in [36] when applying the nominally fifth-order `Rodas5P` scheme to stiff problems. Although the order reduction is stronger for the mixed formulation than for the pressure-head formulation at smaller time step sizes, the mixed formulation is nevertheless more accurate for the entire range of time step sizes considered.

5.3. Adaptive simulations of infiltration benchmark problems. To assess mass conservation under spatial and temporal adaptivity, we consider the two infiltration problems described in [2]. For the first problem, adapted from the sand infiltration case considered in [17], we use the Haverkamp parameters in Table 1 for a soil column of length $L = 0.4\,\mathrm{m}$ and advance the solution until a final time of $t = 360\,\mathrm{s}$. The initial condition is given by $\psi(z, 0\,\mathrm{s}) = -0.615\,\mathrm{m}$ for $z \in [0\,\mathrm{m}, 0.4\,\mathrm{m}]$, and the



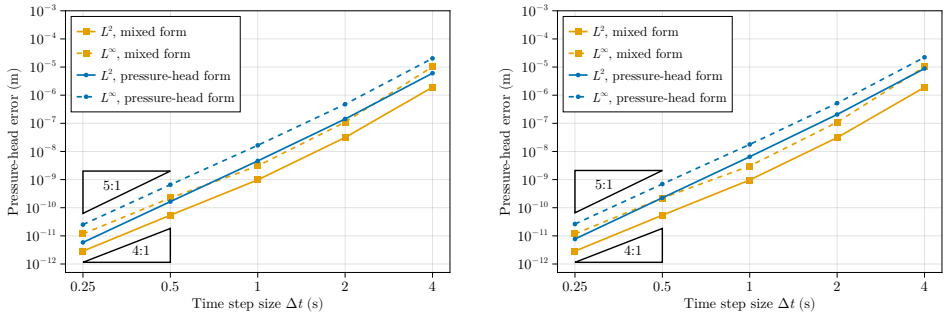
(a) Dirichlet–Dirichlet, $N = 3$

(b) Dirichlet–Dirichlet, $N = 4$

(c) Dirichlet–Neumann, $N = 3$

(d) Dirichlet–Neumann, $N = 4$

Fig. 1: Spatial convergence of the pressure-head error at $t = 120\,\mathrm{s}$ for the manufactured-solution problem using a fixed time step size of $\Delta t = 0.1\,\mathrm{s}$.



(a) Dirichlet–Dirichlet

(b) Dirichlet–Neumann

Fig. 2: Temporal convergence of the pressure-head error at $t = 120\,\mathrm{s}$ for the manufactured-solution problem using a fixed spatial discretization with $N = 3$ and $K = 4096$.

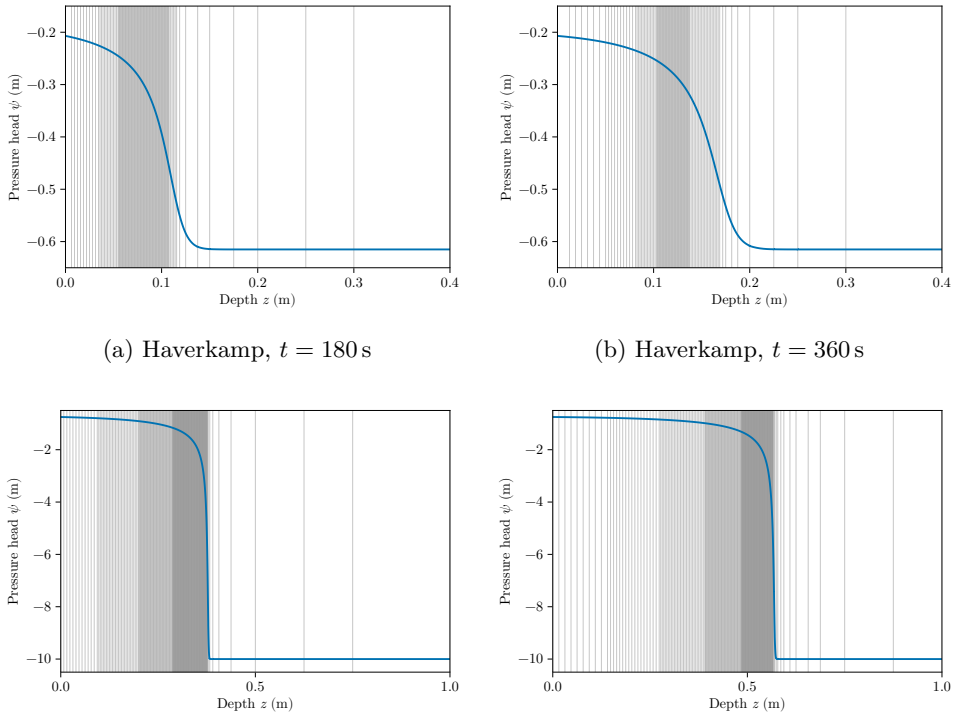
Table 2: Discretization and adaptivity parameters for the Haverkamp and New Mexico infiltration problems.

Symbol	Description	Value
Spatial discretization and adaptivity		
N	Polynomial degree	3
l_{init}	Initial refinement level	6
$l_{\min}$	Minimum refinement level	2
$l_{\max}$	Maximum refinement level	10
η_{c}	Coarsening threshold	0.003
η_{r}	Refinement threshold	0.03
ϵ	Indicator regularization parameter	10^{-11}
n_{adapt}	Accepted time steps between mesh updates	10
Temporal discretization and adaptivity		
β_1	Current-error exponent	0.14
β_2	Previous-error exponent	0.08
σ	Safety factor	0.9
E^0	Initial error value	10^{-4}
$r^n_{\min}$	Minimum step ratio ($n \geq 1$)	0.2
$r^1_{\max}$	Initial maximum step ratio	10^4
$r^n_{\max}$	Maximum step ratio ($n \geq 2$)	10
r^-_{db}	Lower deadband threshold	5/6
r^+_{db}	Upper deadband threshold	1
Δt^0	Initial time step	$10^{-2}\,\text{s}$
`atol`	Absolute temporal error tolerance	10^{-11}
`rtol`	Relative temporal error tolerance	10^{-5}, 10^{-7}, 10^{-9}

time-independent Dirichlet boundary conditions are given by $\psi(0\,\text{m}, t) = -0.207\,\text{m}$ and $\psi(L, t) = -0.615\,\text{m}$ for $t \in (0\,\text{s}, 360\,\text{s}]$. We also consider the second infiltration problem from [2], which uses the van Genuchten–Mualem parameters in Table 1 for a soil column of length $L = 1\,\text{m}$,[2] advancing until a final time of $t = 86400\,\text{s}$. The initial condition is given by $\psi(z, 0\,\text{s}) = -10\,\text{m}$ for $z \in [0\,\text{m}, 1\,\text{m}]$, and the time-independent Dirichlet boundary conditions are given by $\psi(0\,\text{m}, t) = -0.75\,\text{m}$ and $\psi(1\,\text{m}, t) = -10\,\text{m}$ for $t \in (0\,\text{s}, 86400\,\text{s}]$. For ease of reference, we call these the *Haverkamp infiltration problem* and the *New Mexico infiltration problem*, respectively, with the latter name reflecting that its soil parameters approximate the properties of a field site in New Mexico [2]. The discretization and adaptivity parameters used for both infiltration problems are summarized in Table 2. We do not claim that these choices are optimal for either problem, but they nevertheless provide realistic settings for assessing mass conservation under spatial and temporal adaptivity, which, as a consequence of Theorem 4.3, is expected to hold for the mixed formulation with water-content solution transfer independently of the specific adaptivity parameters chosen.

Figure 3 shows the solution profiles for the Haverkamp and New Mexico infiltration problems at intermediate and final times, using the mixed formulation with water-

[2]The domain used here matches the depth range shown in [2, Figure 3], whereas the accompanying text places the lower boundary at a depth of 0.6 m.



(a) Haverkamp, $t = 180\,\mathrm{s}$

(b) Haverkamp, $t = 360\,\mathrm{s}$

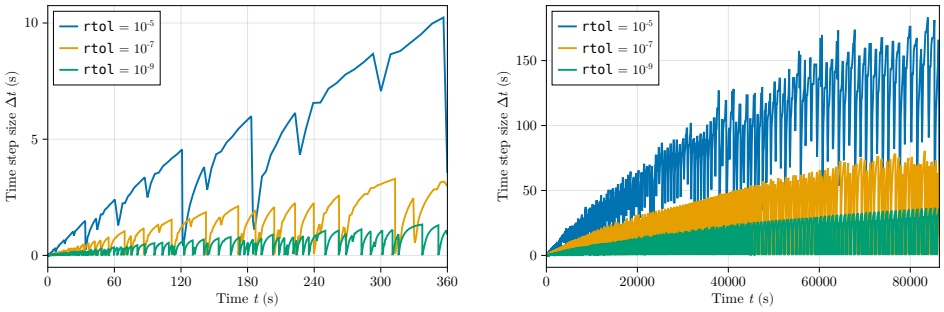
(c) New Mexico, $t = 43200\,\mathrm{s}$

(d) New Mexico, $t = 86400\,\mathrm{s}$

Fig. 3: Pressure-head profiles and element boundaries for the Haverkamp and New Mexico infiltration problems using the mixed formulation with water-content solution transfer and a relative temporal error tolerance of 10^{-7}.

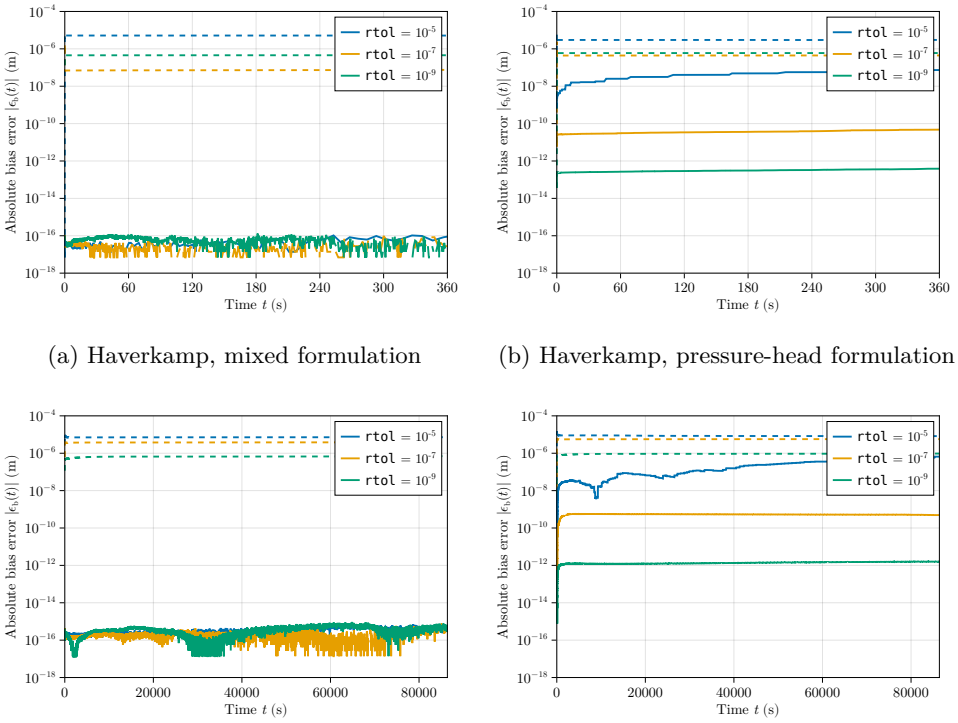
content solution transfer and a relative temporal error tolerance of 10^{-7}. The grey vertical lines indicate element boundaries, and we see that using the total-variation-based refinement indicator results in a mesh that adapts to the evolving solution and effectively tracks the moving front as infiltration progresses. The time step histories for both problems are shown in Figure 4 for relative tolerances of 10^{-5}, 10^{-7}, and 10^{-9}. As expected, the accepted time step sizes are smaller for tighter relative tolerances. Sharp decreases in accepted time step sizes are observed immediately following mesh adaptation events. These decreases are consistent with transient temporal error introduced by transferring the solution between meshes. Similar interactions between temporal error control and mesh adaptivity have been noted in other investigations, with Dolejší and May [11] introducing an automatic reduction of the time step following mesh adaptation to avoid rejected steps.

Following [19], we quantify mass conservation using the *bias error* and plot its absolute value in Figure 5 for each combination of formulation (i.e., mixed form or pressure-head form) and solution transfer variable (i.e., water content or pressure head). Using the SFOM variables and notation of subsection 4.4, the fully discrete



(a) Haverkamp (b) New Mexico

Fig. 4: Accepted time step sizes for the Haverkamp and New Mexico infiltration problems using the mixed formulation with water-content solution transfer and three different relative temporal error tolerances.



(a) Haverkamp, mixed formulation (b) Haverkamp, pressure-head formulation

(c) New Mexico, mixed formulation (d) New Mexico, pressure-head formulation

Fig. 5: Evolution of the absolute bias error for the Haverkamp and New Mexico infiltration problems with three different relative temporal error tolerances. Solid lines denote water-content solution transfer, with gaps indicating bias errors that are zero in double precision; dashed lines denote pressure-head solution transfer.

bias error at time t^n is given by

$$\epsilon_{\mathrm{b}}(t^n) \coloneqq \left(F_{\mathrm{B}}^n - F_{\mathrm{B}}^0\right) - \left(F_{\mathrm{T}}^n - F_{\mathrm{T}}^0\right) - \left(\sum_{k=1}^{K^n} J_k^n \mathbf{1}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{\theta}_k^n - \sum_{k=1}^{K^0} J_k^0 \mathbf{1}^{\mathrm{T}} \boldsymbol{W} \boldsymbol{\theta}_k^0\right), \tag{5.6}$$

where the water-content vectors $\boldsymbol{\theta}_k^0$ and $\boldsymbol{\theta}_k^n$ are evaluated on the respective meshes at initialization and at time t^n, respectively. For the mixed formulation, these vectors are obtained from the evolved water-content block of the DAE solution, whereas for the pressure-head formulation, they are computed from the pressure-head ODE solution using the constitutive relation. As expected from Theorem 4.3, the mixed formulation preserves the mass balance to roundoff in both problems for all three relative tolerances, provided that the solution transfer is performed using the water content following the conservative transfer procedure described in Theorem 3.2; transferring the pressure head instead of the water content results in a loss of mass conservation. For the pressure-head formulation, transferring the water content reduces the mass bias relative to transferring the pressure head, but the bias error does not reach roundoff and remains dependent on the temporal error tolerance.

6. Conclusions. In this work, we presented a high-order discretization of the one-dimensional Richards equation that supports adaptivity in space and time and ensures mass conservation throughout the simulation. The method combines a collocated LDG spatial discretization of the mixed-form Richards equation with a high-order Rosenbrock–Wanner time integration scheme and a mass-conservative solution transfer approach. Numerical experiments based on an open-source Julia implementation confirm high-order spatial and temporal convergence, assess the influence of the temporal formulation on accuracy, and demonstrate the importance of evolving and transferring water content for mass conservation.

This paper provides a foundation for further development of adaptive high-order numerical methods for land models that conserve key physical invariants. Future work will extend the methodology to more complex hydrologic scenarios, including those involving saturated and near-saturated states, heterogeneous soils, solution-dependent boundary conditions, lateral flow, and coupled mass–energy transport. Representative test problems, including those proposed by Clark *et al.* [4], will provide benchmarks for assessing these extensions, with the central methodological challenge being to retain high-order accuracy and fully discrete mass conservation under spatial and temporal adaptivity while preserving the additional invariants and balance laws of the extended models.

Acknowledgments. Funding for this work was provided by the National Oceanic and Atmospheric Administration (NOAA), awarded to the Cooperative Institute for Research to Operations in Hydrology (CIROH) through the NOAA Cooperative Agreement with The University of Alabama, NA22NWS4320003. This work was also supported by the Natural Sciences and Engineering Research Council of Canada (NSERC) through a Canada Postdoctoral Research Award. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of NOAA or NSERC. OpenAI Codex and GitHub Copilot were used to assist with the development of the Julia code used for the numerical experiments and with locating relevant literature, typesetting, and manuscript editing. The authors checked every incorporated suggestion for correctness and take full responsibility for the code and manuscript content.

Reproducibility of computational results. The scripts used to generate the results and figures in this paper are provided in the reproducibility repository [28].

REFERENCES

[1] H. P. Bootsma, M. J. van der Ploeg, and A. H. Weerts, *Why modified Picard works: Mass-conservative and higher-order time integration for the Richardson–Richards equation*, Advances in Water Resources, 213 (2026), 105327, https://doi.org/10.1016/j.advwatres.2026.105327.

[2] M. A. Celia, E. T. Bouloutas, and R. L. Zarba, *A general mass-conservative numerical solution for the unsaturated flow equation*, Water Resources Research, 26 (1990), pp. 1483–1496, https://doi.org/10.1029/WR026i007p01483.

[3] A. Chatzikamaris, G. Bourantas, V. N. Burganos, and N. Malamos, *Numerical solvers for the Richards equation: A comparative review of stability, efficiency and mass conservation*, Eng, 7 (2026), 467, https://doi.org/10.3390/eng7090467.

[4] M. P. Clark, R. Zolfaghari, K. R. Green, S. Trim, W. J. M. Knoben, A. Bennett, B. Nijssen, A. Ireson, and R. J. Spiteri, *The numerical implementation of land models: Problem formulation and laugh tests*, Journal of Hydrometeorology, 22 (2021), pp. 1627–1648, https://doi.org/10.1175/JHM-D-20-0175.1.

[5] J.-B. Clément, F. Golay, M. Ersoy, and D. Sous, *An adaptive strategy for discontinuous Galerkin simulations of Richards' equation: application to multi-materials dam wetting*, Advances in Water Resources, 151 (2021), 103897, https://doi.org/10.1016/j.advwatres.2021.103897.

[6] B. Cockburn and B. Dong, *An analysis of the minimal dissipation local discontinuous Galerkin method for convection-diffusion problems*, Journal of Scientific Computing, 32 (2007), pp. 233–262, https://doi.org/10.1007/s10915-007-9130-3.

[7] B. Cockburn and C.-W. Shu, *The local discontinuous Galerkin method for time-dependent convection-diffusion systems*, SIAM Journal on Numerical Analysis, 35 (1998), pp. 2440–2463, https://doi.org/10.1137/S0036142997316712.

[8] B. Cumming, T. Moroney, and I. Turner, *A mass-conservative control volume-finite element method for solving Richards' equation in heterogeneous porous media*, BIT Numerical Mathematics, 51 (2011), pp. 845–864, https://doi.org/10.1007/s10543-011-0335-3.

[9] T. A. Davis and E. Palamadai Natarajan, *Algorithm 907: KLU, a direct sparse solver for circuit simulation problems*, ACM Transactions on Mathematical Software, 37 (2010), pp. 36:1–36:17, https://doi.org/10.1145/1824801.1824814.

[10] V. Dolejší, M. Kuraz, and P. Solin, *Adaptive higher-order space-time discontinuous Galerkin method for the computer simulation of variably-saturated porous media flows*, Applied Mathematical Modelling, 72 (2019), pp. 276–305, https://doi.org/10.1016/j.apm.2019.02.037.

[11] V. Dolejší and G. May, *An anisotropic hp-mesh adaptation method for time-dependent problems based on interpolation error control*, Journal of Scientific Computing, 95 (2023), 36, https://doi.org/10.1007/s10915-023-02153-1.

[12] V. Dolejší, H.-G. Shin, and M. Vlasák, *Error estimates and adaptivity of the space-time discontinuous Galerkin method for solving the Richards equation*, Journal of Scientific Computing, 101 (2024), 11, https://doi.org/10.1007/s10915-024-02650-x.

[13] M. W. Farthing and F. L. Ogden, *Numerical solution of Richards' equation: A review of advances and challenges*, Soil Science Society of America Journal, 81 (2017), pp. 1257–1269, https://doi.org/10.2136/sssaj2017.02.0058.

[14] G. J. Gassner, *A skew-symmetric discontinuous Galerkin spectral element discretization and its relation to SBP-SAT finite difference methods*, SIAM Journal on Scientific Computing, 35 (2013), pp. A1233–A1253, https://doi.org/10.1137/120890144.

[15] K. Gustafsson, M. Lundh, and G. Söderlind, *A PI stepsize control for the numerical solution of ordinary differential equations*, BIT Numerical Mathematics, 28 (1988), pp. 270–287, https://doi.org/10.1007/BF01934091.

[16] E. Hairer and G. Wanner, *Solving ordinary differential equations II: Stiff and differential-algebraic problems*, Springer, 2nd ed., 1996, https://doi.org/10.1007/978-3-642-05221-7.

[17] R. Haverkamp, M. Vauclin, J. Touma, P. J. Wierenga, and G. Vachaud, *A comparison of numerical simulation models for one-dimensional infiltration*, Soil Science Society of America Journal, 41 (1977), pp. 285–294, https://doi.org/10.2136/sssaj1977.03615995004100020024x.

[18] W. Hundsdorfer and J. G. Verwer, *Numerical solution of time-dependent advection-diffusion-reaction equations*, vol. 33 of Springer Series in Computational Mathematics, Springer, 2003,

https://doi.org/10.1007/978-3-662-09017-6.

[19] A. M. Ireson, R. J. Spiteri, M. P. Clark, and S. A. Mathias, *A simple, efficient, mass-conservative approach to solving Richards' equation (openRE, v1.0)*, Geoscientific Model Development, 16 (2023), pp. 659–677, https://doi.org/10.5194/gmd-16-659-2023.

[20] C. E. Kees and C. T. Miller, *Higher order time integration methods for two-phase flow*, Advances in Water Resources, 25 (2002), pp. 159–177, https://doi.org/10.1016/S0309-1708(01)00054-9.

[21] S. Keita, A. Beljadid, and Y. Bourgault, *Implicit and semi-implicit second-order time stepping methods for the Richards equation*, Advances in Water Resources, 148 (2021), 103841, https://doi.org/10.1016/j.advwatres.2020.103841.

[22] D. A. Kopriva, *Implementing spectral methods for partial differential equations: Algorithms for scientists and engineers*, Springer, 2009, https://doi.org/10.1007/978-90-481-2261-5.

[23] H. Li, M. W. Farthing, C. N. Dawson, and C. T. Miller, *Local discontinuous Galerkin approximations to Richards' equation*, Advances in Water Resources, 30 (2007), pp. 555–575, https://doi.org/10.1016/j.advwatres.2006.04.011.

[24] H. Li, M. W. Farthing, and C. T. Miller, *Adaptive local discontinuous Galerkin approximation to Richards' equation*, Advances in Water Resources, 30 (2007), pp. 1883–1901, https://doi.org/10.1016/j.advwatres.2007.02.007.

[25] J. Manzanero, A. M. Rueda-Ramírez, G. Rubio, and E. Ferrer, *The Bassi Rebay 1 scheme is a special case of the symmetric interior penalty formulation for discontinuous Galerkin discretisations with Gauss–Lobatto points*, Journal of Computational Physics, 363 (2018), pp. 1–10, https://doi.org/10.1016/j.jcp.2018.02.035.

[26] C. T. Miller, C. Abhishek, and M. W. Farthing, *A spatially and temporally adaptive solution of Richards' equation*, Advances in Water Resources, 29 (2006), pp. 525–545, https://doi.org/10.1016/j.advwatres.2005.06.008.

[27] A. Montoison, G. Dalle, and A. Gebremedhin, *Revisiting sparse matrix coloring and bicoloring*, 2025, https://arxiv.org/abs/2505.07308.

[28] T. Montoya and R. J. Spiteri, *Reproducibility repository for "A fully discrete, mass-conservative adaptive discontinuous spectral-element method for the one-dimensional Richards equation in mixed form"*. Zenodo, 2026, https://doi.org/10.5281/zenodo.23270657. Version 0.1.2.

[29] Y. Mualem, *A new model for predicting the hydraulic conductivity of unsaturated porous media*, Water Resources Research, 12 (1976), pp. 513–522, https://doi.org/10.1029/WR012i003p00513.

[30] C. Rackauckas and Q. Nie, *DifferentialEquations.jl – a performant and feature-rich ecosystem for solving differential equations in Julia*, Journal of Open Research Software, 5 (2017), 15, https://doi.org/10.5334/jors.151.

[31] H. Ranocha, M. Schlottke-Lakemper, A. R. Winters, E. Faulhaber, J. Chan, and G. J. Gassner, *Adaptive numerical simulations with Trixi.jl: A case study of Julia for scientific computing*, Proceedings of the JuliaCon Conferences, 1 (2022), 77, https://doi.org/10.21105/jcon.00077.

[32] H. Ranocha, A. R. Winters, H. G. Castro, L. Dalcin, M. Schlottke-Lakemper, G. J. Gassner, and M. Parsani, *On error-based step size control for discontinuous Galerkin methods for compressible fluid dynamics*, Communications on Applied Mathematics and Computation, 7 (2025), pp. 3–39, https://doi.org/10.1007/s42967-023-00264-y.

[33] J. Revels, M. Lubin, and T. Papamarkou, *Forward-mode automatic differentiation in Julia*, 2016, https://arxiv.org/abs/1607.07892.

[34] L. A. Richards, *Capillary conduction of liquids through porous mediums*, Physics, 1 (1931), pp. 318–333, https://doi.org/10.1063/1.1745010.

[35] L. F. Richardson, *Weather prediction by numerical process*, Cambridge University Press, 1922.

[36] G. Steinebach, *Construction of Rosenbrock–Wanner method Rodas5P and numerical benchmarks within the Julia Differential Equations package*, BIT Numerical Mathematics, 63 (2023), 27, https://doi.org/10.1007/s10543-023-00967-x.

[37] G. W. Stewart and J.-g. Sun, *Matrix perturbation theory*, Academic Press, 1990.

[38] L. Tian, L. Lu, H. Guo, J. Cheng, and Y. Yang, *An h-adaptive local discontinuous Galerkin method for solving Richards' equation in unsaturated soils*, Journal of Computational and Applied Mathematics, 491 (2027), 117953, https://doi.org/10.1016/j.cam.2026.117953.

[39] M. D. Tocci, C. T. Kelley, and C. T. Miller, *Accurate and economical solution of the pressure-head form of Richards' equation by the method of lines*, Advances in Water Resources, 20 (1997), pp. 1–14, https://doi.org/10.1016/S0309-1708(96)00008-5.

[40] M. T. van Genuchten, *A closed-form equation for predicting the hydraulic conductivity of unsaturated soils*, Soil Science Society of America Journal, 44 (1980), pp. 892–898,

https://doi.org/10.2136/sssaj1980.03615995004400050002x.

[41] Y. Zha, J. Yang, J. Zeng, C.-H. M. Tso, W. Zeng, and L. Shi, *Review of numerical solution of Richardson–Richards equation for variably saturated flow in soils*, WIREs Water, 6 (2019), e1364, https://doi.org/10.1002/wat2.1364.