FULLY COUPLED METHODS FOR MULTIPHASE MОРPHODYNAMICS

C. Michoski, C. Dawson,
Institute for Computational Engineering and Sciences (ICES), Computational Hydraulics Group (CHG)
University of Texas at Austin, Austin, TX 78712

C. Mirabito,
Department of Mechanical Engineering,
Massachusetts Institute of Technology, Cambridge MA 02139

E. J. Kubatko,
Department of Civil and Environmental Engineering and Geodetic Sciences
The Ohio State University, Columbus, OH 43210

D. Wirasaet, J. J. Westerink
Computational Hydraulics Laboratory, Department of Civil Engineering and Geological Sciences
University of Notre Dame, Notre Dame, IN 46556

Abstract

We present numerical methods for a system of equations consisting of the two dimensional Saint–Venant shallow water equations (SWEs) fully coupled to a completely generalized Exner formulation of hydrodynamically driven sediment discharge. This formulation is implemented by way of a discontinuous Galerkin (DG) finite element method, using a Roe Flux for the advective components and the unified form for the dissipative components. We implement a number of Runge–Kutta time integrators, including a family of strong stability preserving (SSP) schemes, and Runge–Kutta Chebyshev (RKC) methods. A brief discussion is provided regarding implementational details for generalizable computer algebra tokenization using arbitrary algebraic fluxes. We then run numerical experiments to show standard convergence rates, and discuss important mathematical and numerical nuances that arise due to prominent features in the coupled system, such as the emergence of nondifferentiable and sharp zero crossing functions, radii of convergence in manufactured solutions, and nonconservative product (NCP) formalisms. Finally we present a challenging application model concerning hydrothermal venting across metalliferous muds in the presence of chemical reactions occurring in low pH environments.

Keywords: Saint-Venant, shallow water, transport, morphodynamic, geophysical flows, layered multiphase, sediment, fully coupled, SSPRK, RKC, discontinuous Galerkin, hydrothermal vent, nondifferentiable flux, multicomponent reactive flows.

Contents

§1 INTRODUCTION ................................................................. 2
§2 GENERAL GOVERNING EQUATIONS ..................................... 3
§3 Numerical formulation

3.1 Computational numerics ................................................. 7
3.2 Implementing generalizable strategies ......................................... 11

§4 Example systems

4.1 Convergent solutions .................................................. 12
4.2 Nuanced analytic behavior of solutions ........................................ 13
4.3 Hydrothermal vents ..................................................... 18

§5 Conclusion

22

§6 Acknowledgements

22

APPENDIX

22

§1 Introduction

Discontinuous Galerkin finite element methods for modeling coastal, oceanic and inland flows have substantially matured in recent years [17, 19, 31, 43, 45, 58, 66, 67]. Among the many modeling challenges that the physics of these models impart to the mathematical and numerical subsystems, is the question of how one should appropriately represent and couple sedimentary transport that is driven by the dominant flow properties of the associated hydrodynamic wave characteristics [12, 22, 36, 45, 55, 56].

The first challenge that presents itself might be said to exist at the level of the morphodynamic geophysics of the sedimentary transport representation. In this area, the general theoretical underpinnings lie in the form of the Exner equation [15, 48, 62, 63], which can be viewed as a simple balance law for the conservation of mass in fluvial processes. For layered bathymetric bed loads $b_i$, where $i \leq \ell$ corresponds to layer $i$ of the possible $\ell$ strata in the sedimentary structure (i.e., multiple phases), these equations satisfy the seemingly straightforward $\ell$ conservation laws, $\partial_t b_i + \nabla_x \cdot \tilde{q}_i = 0$ for every positive $i \leq \ell$, when diffusive forces are neglected (note that we use $\nabla_x$ to denote the spatial gradient).

This superficially simple conservation law, it turns out, is anything but simple, as the discharge fluxes $\tilde{q}_i$ end up being not only highly nonlinear functions of the state space, but in fact highly irregular mathematical objects that take on different mathematical forms for different types of sediment (e.g. the stratigraphic granulometry of the bed), different grades of bed slopes, different interaction strengths between the hydrodynamic forcings in the systems — for example, mudslides, debris flows, avalanches, flowslides, sturzstrom, slesches, sullage, gyttja, etc. — and so on [20].

This complicated array of “bed evolution types and forms” leads to a delicate framework for the strongly coupled system of partial differential equations in question. For example, when coupling the Exner-type models to a set of shallow water equations, it is not clear if the generalized transported quantity $(\nabla_x \cdot \tilde{q}_i)$ has a hyperbolic signature and should be treated as a primarily advective operator, or whether it has a strong nonlinear elliptic signature that transform its basic behavior, etc. Of course this question cannot be answered in the generalized setting. Moreso, as we discuss below, the situation is substantially more delicate even than simply trying to determine the signature behavior of the system of equations, as the discharge functions are in most common representations given by nondifferentiable functions.

Regardless of the mathematical features of these solutions, the underlying geophysical theories are carefully and thoroughly derived and represent often times extremely accurate empirical models tracking sedimentary evolution in specific morphodynamic contexts [11, 20, 50, 51, 63, 68]. Moreover, the uncoupled shallow water equations, which have a relatively solid mathematical foundation, introduce a number of their own delicate features into the system; such as nonconservative products arising in the
predominantly hyperbolic convective subsystem. This nonconservative product formulation that arises in the classical derivation of the most standard form of the shallow water equations, has been thoroughly studied and is known to present quite subtle features into the numerical solution space of the shallow water hydrodynamics, even without any sedimentary coupling.

In this analysis, we present a strongly coupled solution to such systems of sedimentary flow that are being primarily driven by shallow water hydrodynamics. In our approach, we assume a “linearizable” form for the advective flux, such that the resulting Jacobian matrices of the system are computationally and algebraically “well-posed,” inasmuch as it can be cast into forms with nonsingular functional representations. When such systems exist, we derive the eigenproblem of the fully coupled and fully generalized multicomponent two-dimensional system in its most general form.

This analytic decomposition of the problem into advective and diffusive subsystems is then used to recast the solution in its discrete form. We project the system of equations into a discontinuous basis, and utilize a Roe flux formulation to arbitrary order accuracy. We should note that recent work [15] was performed in this general direction, where the solution was restricted to third order accuracy for only the Grass equation and where a slightly more diffuse Harten–Lax-van Leer numerical flux was implemented. We additionally utilize the unification framework of Arnold, Brezzi, Cockburn and Marini [3] for our solution to the parabolic subsystem, and implement a family of strong stability preserving Runge–Kutta (SSPRK) and Chebyshev (RKC) time discretization schemes [57, 64] to recover potentially truncated eigenmodes in the discrete solution space.

Finally we show some numerical test cases of the fully coupled system. First in §4.1 we perform an ideal test case to demonstrate the expected convergence rates and orders that the system is expected to satisfy. Then in §4.2 we discuss some of the important underlying subtleties that the ideal test case sufficiently conceals. Namely, we discuss the delicate balance between truncation error, sharp analytic zero crossing functions with steep gradients, path convergence, slolalimiting, and model formulation. Finally in §4.3 we present an idealized hydrothermal vent application model. Here we couple the morphodynamic shallow water system to a chemically active kinetic model for shallow water systems, and explore its behavior when venting concentrated protons in the presence of pliable metalliferous muds.

§2 General governing equations

We are primarily concerned with the two-dimensional Saint-Venant system, sometimes referred to as the two-dimensional shallow water equations, fully coupled to a generalized form of the Exner equation for sedimentary transport that comprise the following coupled nonlinear system over the domain \((t, \mathbf{x}) \in (0, T) \times \Omega, \text{ for } \mathbf{x} = (x, y) \text{ and } \Omega \subset \mathbb{R}^2\),

\[
\begin{align*}
\partial_t H + \nabla_x \cdot \mathbf{q} &= 0, \\
\partial_t \mathbf{q} + \nabla_x \cdot (\mathbf{q} \otimes \mathbf{u} + \frac{1}{2} \mathbf{g} H^2) &= \mathbf{g} H \nabla_x b + \nabla_x \cdot (\eta \nabla_x \mathbf{q}) + S, \\
\partial_t b_i + \nabla_x \cdot \mathbf{q}_i - \nabla_x \cdot (\mathbf{D} \nabla_x b_i) &= 0,
\end{align*}
\tag{2.1}
\]

where \(\mathbf{q} = \mathbf{q}(t, \mathbf{x})\) is chosen to satisfy a fairly general and inclusive flux formulation for the sediment discharge \(\mathbf{q} = \mathbf{q}(H, \mathbf{q}, b)\), with the momentum flux \(\mathbf{q} = H \mathbf{u}\). The total height of the water column \(H = H(t, \mathbf{x})\) is a linear combination of the bathymetric bed load \(b = b(t, \mathbf{x})\) with layered strata \(b = \sum_i b_i\) and the free surface height \(\zeta = \zeta(t, \mathbf{x})\), such that \(H = \zeta + b\); \(\mathbf{g}\) is the gravitational constant, \(\eta\) is the eddy viscosity tensor (often treated as a constant \(\eta \in \mathbb{R}^+\)), and \(\mathbf{D}\) the sedimentary eddy viscosity tensor (also often treated as a constant \(\mathbf{D} \in \mathbb{R}^+\)). The remaining source term \(S = S(\mathbf{x}, t)\) accounts for all remaining first order forcings in the system, and often, such as when \(S\) contains the wind forcings in hurricane storm surge models for example [10, 13], is the dominant term. Note that below when we write \(b\) and its corresponding flux in the Exner equation with no index, we mean to restrict to the case of the single stratum, \(\ell := 1\).
Now, let $U = (\zeta, H u, b)^\top$ be the state vector of the system (see below, we will use this state vector $U$ and $U = (H, H u, b)^\top$ interchangeably by way of (3.15)), and $\mathbf{F} = \mathbf{F}(U) = (H u, H u \otimes u + \frac{1}{2} g H^2, \tilde{q})^\top$ the nonlinear flux. For the single layer $\ell = 1$ this can be split into components and written in matrix form as,

$$
\mathbf{F} = (f_x, f_y) = \begin{pmatrix}
\frac{u H}{H u^2 + \frac{1}{2} g H^2} & \frac{v H}{H u v} \\
\frac{H u^2 + \frac{1}{2} g H^2}{H u v} & \frac{H v^2 + \frac{1}{2} g H^2}{\tilde{q}_x} \\
A_g H^{-3} (\frac{q_x^3}{2} + \frac{1}{2} q_x q_y + \frac{1}{2} q_y^2 q_x) & A_g H^{-3} (\frac{q_y^3}{2} + \frac{1}{2} q_y q_x + \frac{1}{2} q_x q_y^2)
\end{pmatrix}
$$

(2.2)

where we have let the components of the velocity vector be $u = (u, v)$, and $\tilde{q}_x$ and $\tilde{q}_y$ denote the abstract $x$ and $y$ components of the sedimentary flux vector $\tilde{q}$.

The formula that the discharge equation $\tilde{q}$ takes can be quite complicated, so much so that frequently one comes across fluxes that lead to ill-posed Jacobian matrices. For example, consider the standard Grass equation, where $\tilde{q} = A_g H^{-1} |H|^{-1} q^{m-1} q$, for $A_g, m \in \mathbb{R}$, where $A_g = A_g(s^2/m)$ is a constant factor that includes the prefactor $1/(1 - \phi)$, for $\phi = \phi(\ell)$ the layer specified sediment porosity. Then as $q_x = H u$ and $q_y = H v$, when $m = 3$ the flux components can be algebraically balanced over $x$ and $y$,

$$
\mathbf{F} = (f_x, f_y) = \begin{pmatrix}
0 & 1 \\
\frac{g H - u^2}{-u v} & 0 \\
3A_g H^{-4} q_x & A_g H^{-3} (3q_x^2 + q_x q_y + \frac{q_y^2}{2}) \\
-3A_g H^{-4} q_y & A_g H^{-3} (3q_y^2 + q_y q_x + \frac{q_x^2}{2})
\end{pmatrix}
$$

(2.3)

Setting $q_x = (\frac{q_x^3}{2} + \frac{1}{2} q_x q_y + \frac{1}{2} q_y^2 q_x)$ and $q_y = (\frac{q_y^3}{2} + \frac{1}{2} q_y q_x + \frac{1}{2} q_x q_y^2)$ the Jacobian matrices then satisfy the following two equations, first in $x$

$$
\mathbf{\Gamma}^x n_x = n_x \begin{pmatrix}
0 & 1 & 0 & 0 \\
g H - u^2 & 2 u & 0 & 0 \\
-uv & 0 & u & 0 \\
-3A_g H^{-4} q_x & A_g H^{-3} (3q_x^2 + q_x q_y + \frac{q_y^2}{2}) & A_g H^{-3} (\frac{q_x^2}{2} + q_x q_y) & -3A_g H^{-4} q_x
\end{pmatrix},
$$

(2.4)

and then in $y$

$$
\mathbf{\Gamma}^y n_y = n_y \begin{pmatrix}
0 & 1 & 0 & 0 \\
-uv & 0 & 1 & 0 \\
g H - v^2 & 0 & u & 0 \\
-3A_g H^{-4} q_y & A_g H^{-3} (\frac{q_y^2}{2} + q_y q_x) & A_g H^{-3} (3q_y^2 + q_y q_x + \frac{q_x^2}{2}) & -3A_g H^{-4} q_y
\end{pmatrix}.
$$

(2.5)

This is an algebraically well-balanced formulation (i.e., there is no first order singular behavior in the Jacobian representation, and there is algebraic symmetry relative to $x$ and $y$) even if, as we shall see below, this formulation is largely driven by stiff nonlinear forcings. However, it should be noted that even for the fairly common Grass equation, the Jacobian terms can lead to singular behavior, e.g. signum functions, Dirac delta functions, etc. For example, consider the simple constant in $y$ vector field $q = (x, 0)$. Then when $m = 2$ and $A_g = 1$, one obtains the vector field $\tilde{q} = H^{-1} |H|^{-1} x |x|$, such that $\partial_x \tilde{q}$ is formally nondifferentiable at the origin. These concerns take on a significant practical precedence numerically, and lead to the necessity of linearization about the formal flux in order to achieve a stable and robust numerical method.

Though the Grass equation [20] demonstrates first order differential instabilities, it still might be viewed as one of the simplest of the possible forms that the classical Exner fluxes take, and also one of the relatively more “regularized” forms. For example, there is also the Meyer-Peter & Müller equation [14] [20]...
for median grain diameter flows in rivers and channels with gentle slopes \( i.e., \) slopes less than 2\%, which is given to satisfy
\[
\bar{q} = A q \text{sgn}(u) \left( c_1 + c_2 \frac{q^2}{H^2 R^{1/3}} \right),
\]
for \( c_1, c_2 \in \mathbb{R} \) constants and \( R = R(H) \) the hydraulic ratio. Note that in these equations, we assume \( q^2 := |q|^2 \). Furthermore, this formula demonstrates rudimentary functional irregularity with the (mathematically) alarming appearence of the signum function \( \text{sgn}(u) \) containing a zeroth order discontinuity at the origin. The appearence of the signum in (2.6) is nothing but an empirical representative of the componentwise directionality of the water-particle semi-excursion in the bed layer \([14, 47]\), yet, from a mathematical point of view introduces jumps in a vector field where zero is frequently achieved (for example, in a tidal vector field that oscillates through two opposing directions). In this sense then, \( \text{sgn}(u) \) is a vector valued function, and can be computed as \( \text{sgn}(u) = (\text{sgn}(u), \text{sgn}(v)) \).

In cases such as (2.6) it is frequently easy in practice to replace the signum function with a smooth analogue, such as the hyperbolic tangent, and thus completely sidestep the question of mathematical irregularity at the outset. These type of "easily adjustable representations" appear in many popular bed transport formulas, such as the Fernández Luque \& Van Beek equation [20], the Nielsen formula [17] and even the Bagnold equation for wind saltation, that all demonstrate comparable "adjustable" irregularity to the zeroth order. The van Rijn equation [59–61], in contrast, is a fractional polynomial equation in \( H, u \) and \( R \) \( i.e., \) functions comprised of linear combinations of fractional monomials of any degree \( d \in \mathbb{R} \), and as such displays relatively greater nonlinear regularity with respect to its arguments, than do many other formulas.

An example of a more modern approach to sediment transport is provided by the Restrepo sand-ridge evolution model [50] [51] for fully differentiable data. Here we are tasked with solving a coupled system of empirically determined auxiliary equations, first for \( b_0 \), and then, to second order for example, for \( b \), \( i.e., \)
\[
\partial_t b_0 + c_0 \nabla_x b_0 = c_1 b_0 + c_2, \quad \text{and} \quad \partial_t b + b_0 \nabla_x b = c_1 b + \frac{c_2}{2} \nabla_x b_0^2,
\]
and so forth to arbitrary asymptotic order. The frozen terms \( c_0, c_1 \) and \( c_2 \) serve as parameters here, and represent time-varying wave characteristics from the eigendecomposition of the flow-coupled variables, \( H \) and \( q \), and for accuracy require spectrally resolved short-wave dispersion models. Of the many notable features of the model, one of its most important is that it is designed to recover the semiperiodic dynamic dendritic sand-bar formations observed in coastal sand-ridges, while additionally preserving a zero mean slope condition throughout the domain.

Another important and prevalent approach to sidesteps many of the complications that the above models introduce, are those driven almost entirely by accumulated empirical relations, such as the Universal Soil Loss Equation (USLE) in use in the sediment retention model used by the U.S. Department of the Interior, the National Aeronautics and Space Administration (NASA), the Soil Conservation Service of the U.S. Department of Agriculture, the Association of American Geographers, and the International Geographical Union [2]. These models are quite dynamic in the range of features they can address, though they remain largely driven by large datasets of experimentally determined constants, such as large datasets of land use and land cover (LULC) maps, which, though appealing from physical arguments and validation studies, make them relatively difficult to perform careful code verification on, as the error ranges in the empirical datasets end up eclipsing the measurable error bounds in the deterministic features of the continuum models. Likely these empirical models coupled to, for example, parameter estimation methods [17] in such a way as to quantify the stochasticity in the parameter variation, could lead to powerfully predictive methods with stable theoretical underpinnings.

Nevertheless, for the remainder of this paper, we will work on a formalism that assumes that \( \bar{q} \) conforms to a conventional Exner formalism, and moreso can be linearized about a well-formed Jacobian. In the general formulation, the Jacobian matrices can then be written in terms of the split components
\[ \Gamma^* = \Gamma^*(\mathbf{u}) = (J_U f), \] relative to the decomposition of the unit outward normal \( \mathbf{n} = (n_x, n_y) \), such that generally the \( x \)-component Jacobian for any such coupled sediment system is given to satisfy

\[
\Gamma^x n_x = n_x \begin{pmatrix} 0 & 1 & 0 & 0 \\ gH - u^2 & 2u & 0 & 0 \\ -uv & v & u & 0 \\ \partial H \tilde{q}_x & \partial H_\alpha \tilde{q}_x & \partial H \nu \tilde{q}_x & \partial \nu \tilde{q}_x \end{pmatrix},
\]

and the \( y \)-component to satisfy

\[
\Gamma^y n_y = n_y \begin{pmatrix} 0 & 0 & 1 & 0 \\ gH - v^2 & 0 & 2v & 0 \\ -uv & v & u & 0 \\ \partial H \tilde{q}_y & \partial H_\nu \tilde{q}_y & \partial \nu \tilde{q}_y & \partial \nu \tilde{q}_y \end{pmatrix}.
\]

Now taking the nonlinear system (2.1) and solving the characteristic equation \( \det (\sum_i \Gamma^i \cdot \mathbf{n} - \mathbb{I} \varsigma) = 0 \) on the boundary \( \partial \Omega_i \), the vector form of the eigenvalues are found to be

\[
\varsigma_1 = \mathbf{u} \cdot \mathbf{n} + c, \quad \varsigma_2 = \mathbf{u} \cdot \mathbf{n}, \quad \varsigma_3 = \mathbf{u} \cdot \mathbf{n} - c, \quad \varsigma_4 = \partial \nu \tilde{q} \cdot \mathbf{n}.
\]

The characteristic wave celerity \( c = \sqrt{gH} \), and in many shallow water applications it is \( c \) that is the dominate eigenmode of the system. The corresponding eigenvector matrix \( V \) is then given by

\[
V = \begin{pmatrix} \frac{(c^2 - \varsigma_1^2)(c^2 + \varsigma_1)(c^2 n_y^2 - v^2)}{\varsigma_1 \alpha \gamma_1} & 0 & \frac{(\varsigma_1^2 - c^2)(\varsigma_2 - c - \varsigma_4)(c^2 n_y^2 - v^2)}{\varsigma_1 \beta \gamma_2} & 0 \\ \frac{\alpha \gamma_1}{\varsigma_1} & \frac{\varsigma_1 \gamma_1 n_y}{\varsigma_1 \beta 2} & \frac{\varsigma_1 \beta \gamma_2}{\varsigma_1 (\varsigma_2 - \varsigma_4)} & 1 \\ \frac{\gamma_3}{\varsigma_1} & \frac{\varsigma_1 \gamma_1}{\varsigma_1 \beta 2} & \frac{\varsigma_1 \beta \gamma_2}{\varsigma_1 (\varsigma_2 - \varsigma_4)} & 0 \\ \frac{1}{\varsigma_1} & \frac{1}{\varsigma_1} & \frac{1}{\varsigma_1} & 1 \end{pmatrix},
\]

where \( V = V(U) = (c_1 | c_2 | c_3 | c_4) \) such that \( c_1 \) is the column eigenvector associated to the eigenvalue \( \varsigma_1 \), \( \text{etc.} \), and the supplemental variables are all provided in the appendix. For the sake of completion, the inverse of this matrix is given by

\[
V^{-1} = \begin{pmatrix} -\frac{\alpha \gamma_1 \gamma_2 \gamma_3 (n_y^2 + n_x^2 + n_x n_x)}{\varsigma_1 \beta 2} & -\frac{\alpha \gamma_1 \gamma_2 n_y}{\varsigma_1 \beta 2} & -\frac{\alpha \gamma_1 \gamma_2 n_x}{\varsigma_1 \beta 2} & 0 \\ \frac{\alpha \gamma_1 \gamma_2 \gamma_3 (n_y^2 + n_x^2 + n_x n_x)}{\varsigma_1 \beta 2} & \frac{\alpha \gamma_1 \gamma_2 n_y}{\varsigma_1 \beta 2} & \frac{\alpha \gamma_1 \gamma_2 n_x}{\varsigma_1 \beta 2} & 0 \\ \frac{\alpha \gamma_1 \gamma_2 \gamma_3 (n_y^2 + n_x^2 + n_x n_x)}{\varsigma_1 \beta 2} & \frac{\alpha \gamma_1 \gamma_2 n_y}{\varsigma_1 \beta 2} & \frac{\alpha \gamma_1 \gamma_2 n_x}{\varsigma_1 \beta 2} & 0 \\ \frac{\alpha \gamma_1 \gamma_2 \gamma_3 (n_y^2 + n_x^2 + n_x n_x)}{\varsigma_1 \beta 2} & \frac{\alpha \gamma_1 \gamma_2 n_y}{\varsigma_1 \beta 2} & \frac{\alpha \gamma_1 \gamma_2 n_x}{\varsigma_1 \beta 2} & 0 \end{pmatrix},
\]

where again the variables definitions are provided in the appendix.

Finally, let us just mention that the Exner equation is not so much an equation in the standard sense of the word, as it is a family of equations indexed not only by sediment layers \( i \leq \ell \), but also by the discharge laws \( \tilde{q}_i \). This way, in order to perform the standard analysis over the complete system (2.1) (for example, computing an \( L^p \)-stability result on the system) we, in principle, must work over the entire category of functions satisfying the empirical laws \( \tilde{q}_i \). Since the Exner equation is the functorial object that associates the category of solutions of \( \tilde{b} \) to the category of functions \( \tilde{q}_i \), it is perhaps most appropriate to view the Exner equation as, instead of a specified equation, rather the functor that connects the objects of these two categories to each other.
§3 NUMERICAL FORMULATION

§3.1 Computational numerics

We recast (2.1) using the state vector $U$ and convective $F$ and diffusive $G$ fluxes from §2 as

$$U_t + F_x - G_x = g,$$

given initial conditions $U|_{t=0} = U_0$, (3.1)

with Robin boundary constraints,

$$a_iU_i + \nabla_x U_i \left( d_i \cdot n + c_i \cdot \tau \right) - f_i = 0, \quad \text{on } \partial \Omega. \quad (3.2)$$

Here we have defined the viscous flux matrix as $G = G(U, U_x)$ with a general source term $g = g(t, x) = (g_1, \ldots, g_m)$, where $x \in \mathbb{R}^2$ and $t \in (0, T)$. The vectors $a$, $d$, $c$ and $f$ are comprised of the four functions, $a_i = a_i(t, x)$, $d_i = d_i(t, x)$, $c_i = c_i(t, x)$ and $f_i = f_i(t, x)$ for $i = 1, \ldots, 4$, where $n = (n_x, n_y)$ denotes the unit outward pointing normal, and $\tau = (\tau_x, \tau_y)$ the unit tangent vector.

We utilize the unified auxiliary flux formalism of [3] for the parabolic subsystem, so the auxiliary variable $\Sigma$ allows for (3.1) to be recast as

$$U_t + F_x - G_x = g, \quad \text{and } \Sigma = U_x, \quad (3.3)$$

with $G = G(U, \Sigma)$.

Let us discretize our domain $\Omega$. Consider the open set $\Omega \subset \mathbb{R}^2$ with boundary $\partial \Omega$, given $T > 0$ such that $Q_T = (0, T) \times \Omega$. Let $T_h$ denote the partition of the closure of the polygonal triangulation of $\Omega$, which we denote $\Omega_h$, into a finite number of polygonal elements denoted $\Omega_e$, such that $T_h = \{\Omega_{e_1}, \Omega_{e_2}, \ldots, \Omega_{e_n}\}$, for $n \in \mathbb{N}$ the number of elements in $\Omega_h$. Here and below the mesh diameter $h$ is chosen to satisfy $h = \min_{ij} (d_{ij})$ for the distance function $d_{ij} = d(x_i, x_j)$ and elementwise face vertices $x_i, x_j \in \partial \Omega$, when the mesh is structured and regular. For unstructured meshes we provide a range and average over the mesh.

Now, let $\Gamma_{ij}$ denote the face shared by two neighboring elements $\Omega_e_i$ and $\Omega_e_j$, and for $i \in I \subset \mathbb{Z}^+ = \{1, 2, \ldots\}$ define the indexing set $r(i) = \{j \in I : \Omega_{e_j} \text{ is a neighbor of } \Omega_{e_i}\}$. Let us denote all $\Omega_e_i$ containing the boundary $\partial \Omega_h$ by $S_i$ and letting $I_B \subset \mathbb{Z}^- = \{-1, -2, \ldots\}$ define $s(i) = \{j \in I_B : S_j \text{ is a face of } \Omega_{e_i}\}$ such that $\Gamma_{ij} = S_j$ for $\Omega_{e_i} \in \Omega_h$ when $S_j \in \partial \Omega_{e_i}, j \in I_B$. Then for $\Xi_i = r(i) \cup s(i)$, we have

$$\partial \Omega_{e_i} = \bigcup_{j \in \Xi(i)} \Gamma_{ij}, \quad \text{and } \partial \Omega_{e_i} \cap \partial \Omega_h = \bigcup_{j \in s(i)} \Gamma_{ij}.$$

We are interested in obtaining an approximate solution to $U$ at time $t$ on the finite dimensional space of discontinuous piecewise polynomial functions over $\Omega$ restricted to $T_h$, given as

$$S^p_h(\Omega_h, T_h) = \{v : v|_{\Omega_{e_i}} \in \mathcal{P}^p(\Omega_{e_i}) \quad \forall \Omega_{e_i} \in T_h\}$$

for $\mathcal{P}^p(\Omega_{e_i})$ the space of degree of (at most) $p$ polynomials over $\Omega_{e_i}$.

Choosing a set of degree $p$ polynomial basis functions $N_p \in \mathcal{P}^p(\Omega_{e_i})$ for $p = 1, \ldots, n_p$ the corresponding degrees of freedom, we can denote the state vector at time $t$ over $\Omega_{e_i}$, by

$$U_{hp}(t, x) = \sum_{p=1}^{n_p} U^i_p(t) N^i_p(x), \quad \forall x \in \Omega_{e_i}, \quad (3.4)$$

where the $N^i_p$’s are the finite element shape functions in the DG setting, and the $U^i_p$’s correspond to the unknowns. We characterize the finite dimensional test functions

$$v_{hp}, w_{hp} \in W^{k,q}(\Omega_h, T_h), \quad \text{by } v_{hp}(x) = \sum_{p=1}^{n_p} v^i_p N^i_p(x) \quad \text{and } w_{hp}(x) = \sum_{p=1}^{n_p} w^i_p N^i_p(x)$$
where \( v_i^p \) and \( w_k^p \) are the coordinates in each \( \Omega_{e_i} \), with the broken Sobolev space over the partition \( \mathcal{T}_h \) defined by
\[
W^{k,q}(\Omega_h, \mathcal{T}_h) = \{ \omega : \omega|_{\Omega_{e_i}} \in W^{k,q}(\Omega_{e_i}) \quad \forall \Omega_{e_i} \in \mathcal{T}_h \}.
\]

Thus, for \( U \) a classical solution to (3.3), multiplying by \( v_{hp} \) or \( w_{hp} \) and integrating elementwise by parts yields the coupled system:
\[
\frac{d}{dt} \int_{\Omega_{e_i}} U \cdot v_{hp} dx + \int_{\Omega_{e_i}} (F \cdot v_{hp})_x dx - \int_{\Omega_{e_i}} F : v_{hp}^2 dx \\
- \int_{\Omega_{e_i}} (G \cdot v_{hp})_x dx + \int_{\Omega_{e_i}} G : v_{hp}^2 dx = \int_{\Omega_{e_i}} v_{hp} \cdot g dx,
\]
(3.5)
\[
\int_{\Omega_{e_i}} \Sigma \cdot w_{hp} dx - \int_{\Omega_{e_i}} (U \cdot w_{hp})_x dx + \int_{\Omega_{e_i}} U : w_{hp}^2 dx = 0,
\]
where \((\cdot\cdot)\) denotes the scalar product.

Now, let \( n_{ij} \) be the unit outward normal to \( \partial \Omega_{e_i} \) on \( \Gamma_{ij} \), and let \( v|_{\Gamma_{ij}} \) and \( v|_{\Gamma_{ji}} \) denote the values of \( v \) on \( \Gamma_{ij} \) considered from the interior and the exterior of \( \Omega_{e_i} \), respectively. Then by choosing componentwise approximations in (3.5) by substituting in (3.4), we arrive with the approximate form of the first term of (3.5):
\[
\frac{d}{dt} \int_{\Omega_{e_i}} U_{hp} \cdot v_{hp} dx \approx \frac{d}{dt} \int_{\Omega_{e_i}} U \cdot v_{hp} dx,
\]
(3.6)
the second term using a convective numerical flux \( \Phi \), by
\[
\Phi_i(U_{hp}|_{\Gamma_{ij}}, U_{hp}|_{\Gamma_{ji}}, v_{hp}) = \sum_{j \in \Xi(i)} \int_{\Gamma_{ij}} \Phi(U_{hp}|_{\Gamma_{ij}}, U_{hp}|_{\Gamma_{ji}}, n_{ij}) \cdot v_{hp}|_{\Gamma_{ij}} d\Xi \\
\approx \sum_{j \in \Xi(i)} \int_{\Gamma_{ij}} \sum_{l=1}^2 (F)_l \cdot (n_{ij})_l v_{hp}|_{\Gamma_{ij}} d\Xi,
\]
(3.7)
and the third term in (3.5) by,
\[
\Theta_i(U_{hp}, v_{hp}) = \int_{\Omega_{e_i}} F_{hp} : v_{hp}^2 dx \approx \int_{\Omega_{e_i}} F : v_{hp}^2 dx.
\]
(3.8)
The remaining source term is then given by
\[
\mathcal{H}_i(g_{hp}, v_{hp}) = \int_{\Omega_{e_i}} v_{hp} \cdot g_{hp} dx \approx \int_{\Omega_{e_i}} v_{hp} \cdot g dx.
\]
(3.9)

The numerical flux \( \Phi \) is constructed for the purposes of this study as a Roe flux \( \Phi_{Roe} \) using the eigendecomposition from §2. Note that the source term \( g \) in the momentum equation contains the non-conservative product, \( gH \nabla_x b \). This fact has a nontrivial impact on the behavior of the hyperbolic subsystem. Most notably, the finite-time formation of discontinuous solutions in the bed \( b \) leads to non-unique paths in the weak formulation of the non-conservative product. Let us reserve this nuance for the discussion in §4, while here we proceed by characterizing the standard Roe flux formalism.

First we need some definitions. The standard jump condition, relative to the traces on the edges, is given by \( [v_{hp}] = v_{hp}|_{\Gamma_{ij}} - v_{hp}|_{\Gamma_{ji}} \). We further make use of the following diagonal matrix
\[
|\Lambda| = \begin{pmatrix}
|s_1| & 0 & 0 & 0 \\
0 & |s_2| & 0 & 0 \\
0 & 0 & |s_3| & 0 \\
0 & 0 & 0 & |s_4|
\end{pmatrix}.
\]
as well as the set of Roe-averaged “state variables” \( w_{Roe} \), and the set of Roe-averaged “derived primitives” \( v_{Roe} \) of the state variables \( w_{hp} \) (such as the velocity which is derived from the state momentum and column height, \( q_x = H u_x \)), respectively as

\[
\begin{align*}
\text{Roe average:} & \quad w_{Roe} = \frac{1}{2} (w_{hp} | \Gamma_{ij} + w_{hp} | \Gamma_{ji}), \quad \text{and} \quad v_{Roe} = \frac{1}{2} \left( \frac{v_{hp} | \Gamma_{ij} \sqrt{\frac{\Omega_{hp}}{\Omega_{hp} \Gamma_{ij}}} + v_{hp} | \Gamma_{ji} \sqrt{\frac{\Omega_{hp}}{\Omega_{hp} \Gamma_{ji}}}}\right)
\end{align*}
\]

Then the Roe flux \( \Phi_{Roe} \) is defined by

\[
\Phi_{Roe} = \frac{1}{2} \left( \sum_{l=1}^{2} (F_{hp})_l \cdot (n_{ij}) \right) v_{hp} | \Gamma_{ij} - \frac{1}{2} \left( \frac{\nabla v_{Roe} | \hat{A}_{Roe} \hat{V}^{-1}}{U_{hp}} \right) [U_{hp}],
\]

where each matrix \((\hat{\cdot})_{Roe}\) indicates a linearized form of the corresponding matrix expressed in terms of the Roe-averaged variables.

Next we approximate the boundary diffusive term of (3.5) using a generalized diffusive flux \( \mathcal{G} \) such that,

\[
\mathcal{G}(\Sigma_{hp}, U_{hp}, v_{hp}) = \sum_{j \in \Xi(i)} \int_{\Gamma_{ij}} \mathcal{G}(\Sigma_{hp} | \Gamma_{ij}, \Sigma_{hp} | \Gamma_{ji}, U_{hp} | \Gamma_{ij}, U_{hp} | \Gamma_{ji}, n_{ij}) \cdot v_{hp} | \Gamma_{ij} d\Xi
\]

\[
\approx \sum_{j \in \Xi(i)} \int_{\Gamma_{ij}} \sum_{l=1}^{N} (G)_l \cdot (n_{ij}) v_{hp} | \Gamma_{ij} d\Xi,
\]

which the second diffusion term is approximated by

\[
\mathcal{N}_t(\Sigma_{hp}, U_{hp}, v_{hp}) = \int_{\Omega_{ei}} G_{hp} \cdot v_{x}^{hp} dx \approx \int_{\Omega_{ei}} G \cdot v_{x}^{hp} dx.
\]

For the auxiliary equation in (3.5) we expand it such that the approximate solution satisfies,

\[
\mathcal{B}(\hat{U}, \Sigma_{hp}, U_{hp}, w_{hp}, w_{2}^{hp}) = \int_{\Omega_{ei}} \Sigma_{hp} \cdot w_{hp} dx + \int_{\Omega_{ei}} U_{hp} \cdot w_{x}^{hp} dx
\]

\[
- \sum_{j \in \Xi(i)} \int_{\Gamma_{ij}} \hat{U}(U_{hp} | \Gamma_{ij}, U_{hp} | \Gamma_{ji}, w_{hp} | \Gamma_{ij}, n_{ij}) d\Xi,
\]

where,

\[
\sum_{i \in I} \sum_{j \in \Xi(i)} \int_{\Gamma_{ij}} \hat{U}(U_{hp} | \Gamma_{ij}, U_{hp} | \Gamma_{ji}, w_{hp} | \Gamma_{ij}, n_{ij}) d\Xi \approx \sum_{i \in I} \sum_{j \in \Xi(i)} \sum_{l=1}^{N} \sum_{ij} \int_{\Gamma_{ij}} (U)_l \cdot (n_{ij}) q_{hp} | \Gamma_{ij} d\Xi
\]

given a generalized numerical flux \( \hat{U} \), such that

\[
\int_{\Omega_{ei}} \Sigma_{hp} \cdot w_{hp} dx \approx \int_{\Omega_{ei}} \Sigma \cdot w_{hp} dx, \quad \text{and} \quad \int_{\Omega_{ei}} U_{hp} \cdot w_{x}^{hp} dx \approx \int_{\Omega_{ei}} U \cdot w_{x}^{hp} dx.
\]

Combining the above approximations and setting \( \mathcal{B}^* = \sum_{\Omega_{ei} \in h} \mathcal{B}_{ei} \), while defining the inner product

\[
(a_{hp}^{\Sigma}, b_{hp})_{\Omega_{ei}} = \sum_{\Omega_{ei} \in h} \int_{\Omega_{ei}} a_{hp}^{\Sigma} \cdot b_{hp} dx,
\]

we arrive at our approximate solution to (3.3) as the pair of functions \((U_{hp}, \Sigma_{hp})\) for all \( t \in (0, T) \) satisfying:
The semidiscrete discontinuous Galerkin formulation

a) \( U_{hp} \in C^1([0,T); S_h^p) \), \( \Sigma_{hp} \in S_h^p \),

b) \( \frac{d}{dt}(U_{hp}, v_{hp})_{\Omega_c} + \tilde{\Phi}(U_{hp}, v_{hp}) - \Theta(U_{hp}, v_{hp}) \)

\[- \mathcal{G}(\Sigma_{hp}, U_{hp}, v_{hp}) + \mathcal{N}(\Sigma_{hp}, U_{hp}, v_{hp}) = \mathcal{H}(g_{hp}, v_{hp}), \tag{3.13} \]

c) \( \mathcal{Q}(U, \Sigma_{hp}, U_{hp}, w_{hp}, w_{hp}) = 0 , \)

d) \( U_{hp}(0) = \Pi_{hp}U_0, \)

where \( \Pi_{hp} \) is a projection operator onto the space of discontinuous piecewise polynomials \( S_h^p \). Below we utilize the standard \( L^2 \)-projection, given for a function \( f_0 \in L^2(\Omega_c) \) such that our approximate projection \( f_{0,h} \in L^2(\Omega_c) \) is obtained by solving, \( \int_{\Omega_c} f_{0,h} v_{hp} dx = \int_{\Omega_c} f_0 v_{hp} dx. \)

Before moving to the fully discrete form, we note that a common feature of shallow water models is that whenever the bathymetry \( b \) is independent of time, the mass conservation equation from (2.1) naturally takes a form that can be written only in terms of the free surface \( \zeta \). As such, many codes are written utilizing the state vector \( U = (\zeta, q, b)^T \) rather than the state vector \( U = (H, q, b)^T \). In such a case, one can easily recover the state vector \( U = (\zeta, q, b)^T \) by utilizing the following trick. Neglecting viscous and source terms for the sake of transparency, notice that we can rewrite the mass equation after multiplying by a test function and integrating as

\[
\frac{d}{dt} \int_{\Omega_c} v_{hp} (\zeta_{hp} + b_{hp}) \, dx + \int_{\Omega_c} v_{hp} \nabla_x \cdot q_{hp} \, dx = 0, \tag{3.14}
\]

\[
\frac{d}{dt} \int_{\Omega_c} v_{hp} b_{hp} \, dx = - \int_{\Omega_c} v_{hp} \nabla_x \cdot \hat{q} \, dx,
\]

such that by substitution we equivalently have:

\[
\frac{d}{dt} \int_{\Omega_c} v_{hp} \zeta_{hp} \, dx + \int_{\Omega_c} v_{hp} \nabla_x \cdot (q_{hp} - \hat{q}_{hp}) \, dx = 0 \tag{3.15}
\]

This is a convenient way of adapting legacy solvers that have built-in static bathymetry, since now all that needs to be updated to move to the dynamic bathymetry model along are the adapted forms that the fluxes take.

The discretization in time follows now directly from (3.13), where we first employ a family of SSP (strong stability preserving, or often “total variation diminishing (TVD)” Runge–Kutta schemes as discussed in [53, 54]. That is, for the generalized SSP Runge–Kutta scheme we rewrite (3.13) in the form: \( MU_t = \mathcal{L} \), where \( U = (U_1, \ldots, U_p) \) for each element from (3.4), where \( \mathcal{L} = \mathcal{L}(U, \Sigma) \) is the advection-diffusion contribution along with the source term, and where \( M \) is the usual mass matrix. Then the generalized \( s \) stage of order \( \gamma \) SSP Runge–Kutta method (denoted SSP\((s, \gamma)\), SSPRK\((s, \gamma)\), or RKSSP\((s, \gamma)\)) may be written to satisfy:

\[
U^{(0)} = U^n, \quad U^{(i)} = \sum_{r=0}^{i-1} (\alpha_{ir} U^r + \Delta t \beta_{ir} M^{-1} \mathcal{L}^r), \quad \text{for } i = 1, \ldots, s \tag{3.16}
\]

\[
U^{n+1} = U^{(s)},
\]

where \( \mathcal{L}^r = \mathcal{L}(U^r, \Sigma^r) = \mathcal{L}(U^r, \Sigma^r, t^n + \delta_t \Delta t) \) and the solution at the \( n \)-th timestep is given as \( U^n = U_{|t=t^n} \) and at the \( (n+1) \)-st timestep by \( U^{n+1} = U_{|t=t^{n+1}} \), with \( t^{n+1} = t^n + \Delta t \). The \( \alpha_{ir} \) and \( \beta_{ir} \).
Finally the operator \( L^\alpha \) corresponds to the time-lag complication arising in the constraints of the TVD formalism. That is \( \delta_r = \sum_{i=0}^{r-1} \mu_{ir} \), where \( \mu_{ir} = \beta_{ir} + \sum_{i=r+1}^{r-1} \mu_{ir} \alpha_{ir} \), where we have taken that \( \alpha_{ir} \geq 0 \) satisfying \( \sum_{i=0}^{r-1} \alpha_{ir} = 1 \).

To recover the optimal thin region stability (see \cite{57, 34}) we alternatively adopt the finite damped RKC method of second order, where (3.16) is replaced by

\[
\begin{align*}
U^{(0)} &= U^n, \\
U^{(1)} &= U^{(0)} + \Delta t^n \tilde{\mu}_1 M^{-1} L^0 \\
U^{(j)} &= (1 - \hat{\mu}_j - \hat{\nu}_j) U^{(0)} + \hat{\mu}_j U^{(j-1)} + \hat{\nu}_j U^{(j-2)} \\
&+ \Delta t^n \tilde{\mu}_j M^{-1} L^{j-1} + \Delta t^n \tilde{\gamma}_j M^{-1} L^0 \quad \text{for} \ j \in \{2, \ldots, \chi\} \\
U^{n+1} &= U^{(\chi)}.
\end{align*}
\]

Here, \( \tilde{\mu}_1 = \omega_1 \omega_0^{-1} \) and for each \( j \in \{2, \ldots, \chi\} \):

\[
\hat{\mu}_j = \frac{2 \hat{b}_j \omega_0}{b_{j-1}}, \quad \hat{\nu}_j = \frac{-\hat{b}_j}{b_{j-2}} \quad \text{with} \quad a_j = 1 - b_j T_j(\omega_0), \quad \hat{b}_j = b_2, \quad \hat{b}_1 = \omega_0^{-1} \quad \hat{b}_j = T_j''(\omega_0) T_j'(\omega_0)^{-2} \quad \text{for} \ j \in \{2, \ldots, \chi\},
\]

and \( \hat{\gamma}_j = a_{j-1} \tilde{\mu}_j \), where \( \omega_0 = 1 + \epsilon \chi^{-2} \), \( \omega_1 = T'_\chi(\omega_0) T''_\chi(\omega_0)^{-1} \), where the \( T_j \) are the Chebyshev polynomials of the first kind, and \( U_j \) the Chebyshev polynomials of the second kind which define the derivatives, given by the recursion relations:

\[
\begin{align*}
T_0(x) &= 1, \quad T_1(x) = x, \quad T_j(x) = 2x T_{j-1}(x) - T_{j-2}(x) \quad \text{for} \ j \in \{2, \ldots, \chi\}, \\
U_0(x) &= 1, \quad U_1(x) = 2x, \quad U_j(x) = 2x U_{j-1}(x) - U_{j-2}(x) \quad \text{for} \ j \in \{2, \ldots, \chi\},
\end{align*}
\]

\[
T'_j(x) = j U_{j-1}, \quad T''_j(x) = \left( \frac{j(n+1)T_j - U_j}{x^2 - 1} \right) \quad \text{for} \ j \in \{2, \ldots, \chi\}.
\]

Finally the operator \( L^\beta \) is evaluated at time \( L^\beta(t^n + \tilde{\epsilon}_j \Delta t^n) \), where the \( \tilde{\epsilon}_j \) are given by:

\[
c_0 = 0, \quad c_1 = \frac{1}{4} c_2 \omega_0^{-1}, \quad c_j = T'_\chi(\omega_0) T''_\chi(\omega_0) \approx \frac{j^2 - 1}{\chi^2 - 1} \quad \text{for} \ j \in \{2, \ldots, \chi - 1\}, \quad c_\chi = 1.
\]

Notice that in contrast to the SSPRK schemes where the stage expansion is used to thicken the stability region along the admissible imaginary axis while reducing the number of stable negative real eigenvalues along the real axis, in the RKC methods the stage expansion is used to lengthen the stability region along the real axis, as discussed at length in \cite{57}. Such temporal discretizations can always be performed, but in the explicit methodology the timestep restriction often becomes too severe to efficiently model realistic systems.

Our examples in this paper will all be given in the context of the discontinuous Galerkin shallow water code described in \cite{9, 17, 32, 35, 45, 58}, which employs a fully coupled system of (3.13). For the polynomial basis we choose the hierarchical Dubiner basis, and our meshes are comprised of triangular elements. Also, please note that due to the proliferation of variable indices, we will frequently drop the \((\cdot)_{hp}\) subscript below, especially when it is clear from the context that the discrete solution is the object of study.

### §3.2 Implementing generalizable strategies

The code is implemented using an optimized Fortran code (both Fortran 77 and Fortran 90) built around the aforementioned discontinuous Galerkin shallow water code described in \cite{9, 17, 32, 35, 45, 58}. The
primary features added to this code include a fully parallelized implementation of the system outlined in §2 and §3.1. In order to work over a “generalized” form of the Exner equation that relies on solving Jacobian-based Reimann problems, the numerical implementation has to be made “algebraic.” In other words, in order to solve a fully coupled Reimann solver at the edge boundaries of each element relative to “any” algebraic Exner form, we must be able to pass around variables as either algebraic symbols (i.e., logical tokens) or as type-specified Fortran variables.

Along these lines, we choose to implement a Python-based preprocessing wrapper in order to exploit Python’s advanced (though relatively simple) computer algebra capabilities. In this way we are able to feed in the “form” of the Exner equation in a parameter list at the beginning of the simulation. For example, assuming the Grass equation is satisfied for $m = 3$, we can merely load the following two character strings with respect to a splitting over the “Roe variables,” $(\zeta_{\text{Roe}}, b_{\text{Roe}}, q_{\text{Roe}})$:

$$(\text{ZE}_\text{ROE} + \text{bed}_\text{ROE})^{-3} \times ((QX_{\text{ROE}})^3 + \frac{1}{2}(QX_{\text{ROE}}^2)QY_{\text{ROE}} + (QY_{\text{ROE}}^2)QX_{\text{ROE}}))$$

$$(\text{ZE}_\text{ROE} + \text{bed}_\text{ROE})^{-3} \times ((QY_{\text{ROE}})^3 + \frac{1}{2}(QY_{\text{ROE}}^2)QX_{\text{ROE}} + (QX_{\text{ROE}}^2)QY_{\text{ROE}}))$$

while the prefactor coefficient $A_g$ is loaded as a separate parameter.

This string is then read into a Python function parser, where we utilize Python’s built-in SymPy (symbolic Python) library to compute the corresponding Jacobian matrices in the $x$ and $y$ components for any algebraically well-formed Exner construction. These matrices are then stored as character strings and sent to two files; again in $x$ and $y$, respectively. Finally, the character strings are loaded into Fortran and tokenized using a Fortran function parser that converts the character strings back into bytecode. It is then trivial to evaluate, for example, the eigenvalue summation (i.e., dot product) $\varsigma_4 = \partial \tilde{q} \cdot \mathbf{n}$, as well as to construct the eigenvector matrices, etc.

In order to parallelize the tokenization, the $n$ MPI processes must independently initialize the algebraic subsystem (e.g., load SymPy using a local script). This is done only once for the first timestep of the computation on each parallel subdomain, which effectively globally tokenizes the “algebraic form” of the system for the remainder of the simulation. It should be noted, the algebraic differentiation, even for a complicated Exner form, is extremely fast (e.g. requires minimal computational cost) relative even to the adjacent preprocessing steps. What this accomplishes is a fast, or essentially computationally free, way of implementing a versatile and general form of the coupled Exner equation in order to accommodate a large array of “algebraic types” in layered sedimentary transport. In this way, we are able to avoid restricting to a single representational form of the equations beyond selecting a parameter input setting, or possible a linearization and/or mollification/smoothing of the Exner flux formula. This formalism can also be extended to “domain-dependent sediment modeling,” where the particular Exner form chosen can have a spatial or temporal dependence (a type of mortaring),

$$\tilde{q} = \{q|_{\Omega_1 \times (0, T_1)}, \tilde{q}|_{\Omega_2 \times (0, T_2)}, \ldots \}, \quad \text{for} \quad T = \bigcup T_i \quad \text{and} \quad \Omega = \bigcup \Omega_i.$$

For example, when mudslides or sturzstrom occur near coastal regions due to strong storm surge forcings such as wind and rain in one domain $\Omega_1$ at some time point $T_1$, while in areas of deeper water $\Omega_2$ simpler and more mild forms of the Exner equation can be maintained over long timeframes, e.g. $\forall t \in T$.

§4 EXAMPLE SYSTEMS

§4.1 Convergent solutions

In order to test the numerical convergence of (2.1), we construct a pair of manufactured solutions over the two dimensional domain $[-0.5, 0.5]^2$. First we assume vanishing viscosity and sediment diffusion,
such that \( (2.1) \) becomes:

\[
\begin{align*}
\partial_t H + \nabla_x \cdot (Hu) &= S_1, \\
\partial_t (Hu) + \nabla_x \cdot (Hu \otimes u + \frac{3}{2}gH^2) &= gH\nabla_x b + S_2, \\
\partial_t b + \nabla_x \cdot \tilde{q} &= S_3,
\end{align*}
\]

(4.1)

where the source terms \( S_1, S_2, \) and \( S_3 \) are chosen to satisfy the manufactured system.

Now, as a simple test, let us assign the following analytic solution:

\[
H = H_0, \quad q = \omega (y, -x), \quad b = b_0 + \varpi \cos(t),
\]

(4.2)

where \( \zeta = H - b \) and \( f = e^{(-x^2-y^2)/\sigma} \). Here we set \( \omega = 0.001, \varpi = 0.01 \) and \( \sigma = 0.02 \), and use \( \tilde{q} = A_\varpi \theta \) in place of the Grass equation. The initial water column is \( H_0 = 9 \) and \( b_0 = 5 \). Multiplying this through, and noticing that \( \partial_t H \) and \( \partial_t q \) both vanish by construction, then we only require the source term, \( S_3 = -\varpi f \sin(t) \) in order to test the convergence properties of the numerical method. Here we assume the natural Dirichlet boundary conditions — where as the Gaussian function \( f \) approximately vanishes to machine precision on the boundaries, is set to constant values of \( b_0 \) and \( \zeta_0 = H_0 - b_0 \) at the boundary. The fluxes at the boundary are given by transmissive conditions \( q|_{\partial \Omega_{ij}} = q|_{\partial \Omega_{ji}} \).

Let us note here that the reason for choosing the Gaussian function \( f \) is that both its gradient \( \nabla_x b|_{\partial \Omega} \approx 0 \) and its value \( b|_{\partial \Omega} \approx 0 \) approximately vanish at the boundary. The effective consequence is that then the boundary error is screened from polluting the convergence rate. The issue of weakly imposed, or \( L^\infty \), first order boundary layer formation has been extensively addressed in [39, 41].

As seen in Table 1, the error is well-behaved for this simple example problem. It should however be noted, that even in this highly idealized setting, the bathymetric jumps \( [b_{hp}] \) across interior cells are far from zero even prior to the first timestep, and even though the initial conditions are quite smooth. This is of course due to the truncation error in the \( L^2 \)-projection onto a discontinuous polynomial basis. In fact, this feature of the basis introduces some subtleties into the solution space whenever the non-conservative product takes precedence, which we discuss next.

### §4.2 Nuanced analytic behavior of solutions

Now let us consider a more complicated manufactured solution, and take the opportunity to address some of the subtleties that underlie the full system. In this section our aim, rather than revisiting the strict convergence behavior already shown in §4.1, is to expand our understanding of the space of admissible coupled sedimentary systems that (4.1) subsumes. As our goal is to develop a robust and accurate model for a generalized system, we recognize a need to develop an organized sense for the many subtle underlying features present in the system, which, when ignored and/or understated, can make modeling these coupled systems difficult, leading to both unstable and inaccurate results. It is in this

<table>
<thead>
<tr>
<th>( p )</th>
<th>( L^2 )-error (( h_{hp} ))</th>
<th>( h^{-1} )</th>
<th>Rate</th>
<th>( p )</th>
<th>( L^2 )-error (( h_{hp} ))</th>
<th>( h^{-1} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( 6.42 \times 10^{-8} )</td>
<td>16</td>
<td>-</td>
<td>1*</td>
<td>( 6.42 \times 10^{-8} )</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>( 1.53 \times 10^{-8} )</td>
<td>32</td>
<td>2.06</td>
<td>2</td>
<td>( 1.73 \times 10^{-8} )</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>( 3.72 \times 10^{-9} )</td>
<td>64</td>
<td>2.04</td>
<td>3</td>
<td>( 9.72 \times 10^{-10} )</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>( 8.73 \times 10^{-10} )</td>
<td>128</td>
<td>2.09</td>
<td>4</td>
<td>( 7.79 \times 10^{-11} )</td>
<td>16</td>
</tr>
</tbody>
</table>

Table 1: Here we show the \( L^2 \)-error and convergence rates of the simple analytic model from (4.1) and (4.2) using SSP(6,4) and \( \Delta t = 1.e^{-5} \). The * denotes the pre-asymptotic \( p \)-convergent behavior.
Figure 1: Here we show the steep velocity contours of $|u|$ from (4.3) in the upper left, where the vector field is shown in the upper right. On the bottom image the difference map $|b_{hp} - b_{exact}|$ is superimposed over a slice of the velocity profile. All plots were made using $h = 1/256$ and $p = 1$.

spirit that we explore some of the more salient hazards present in the system, as well as some popular alternative ways of viewing these types of systems.

As above in §4.1 we assume the form of (4.1) and again construct a second order approximate vanishing boundary treatment, which is just to say that we choose a solution for the model state vector to satisfy $U|_{\partial \Omega} \approx 0$, as well as fluxes that approximate $\nabla_x U|_{\partial \Omega} \approx 0$. Under these restrictions, we assign the following analytic representation:

$$
H = H_0, \quad Hu = f \varpi \cos t (\xi_x, \xi_y), \quad b = b_0 + f \varpi \cos t, \quad (4.3)
$$

where $\zeta = H - b$, $f = e^{(-x^2-y^2)/\sigma_1}$, $\xi_x = \tanh(x)$, and $\xi_y = \tanh(y)$. Here again we let $\varpi = 0.01$ and $\sigma_1 = 0.02$, but use instead that $H_0 = 1$ and $b_0 = 0.5$.

Now, let us choose the Grass-like equation for the sedimentary flux $\bar{q}$, but let us assume that $m = 1$...
Table 2: Here we show the $L^2$-error and convergence rates for both the “exact analytic model” from (4.1) and (4.4) as well as the linearized form from (4.5), using SSP(6,4) and $\Delta t = 1e^{-5}$.

<table>
<thead>
<tr>
<th>$p$</th>
<th>$L^2$-error ($b_{hp}$)</th>
<th>$h^{-1}$</th>
<th>Rate</th>
<th>$p$</th>
<th>$L^2$-error ($b_{hp}$)</th>
<th>$h^{-1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>$1.43 \times 10^{-6}$</td>
<td>16</td>
<td>–</td>
<td>1</td>
<td>$1.43 \times 10^{-6}$</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>$4.11 \times 10^{-7}$</td>
<td>32</td>
<td>1.80</td>
<td>2</td>
<td>$2.61 \times 10^{-7}$</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>$1.26 \times 10^{-7}$</td>
<td>64</td>
<td>1.71</td>
<td>3</td>
<td>$2.03 \times 10^{-8}$</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>$3.91 \times 10^{-8}$</td>
<td>128</td>
<td>1.69</td>
<td>4</td>
<td>$1.10 \times 10^{-9}$</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>$1.09 \times 10^{-8}$</td>
<td>256</td>
<td>1.84</td>
<td>5</td>
<td>$1.45 \times 10^{-10}$</td>
<td>16</td>
</tr>
<tr>
<td>1</td>
<td>$2.69 \times 10^{-9}$</td>
<td>512</td>
<td>2.02</td>
<td>–</td>
<td>–</td>
<td>–</td>
</tr>
</tbody>
</table>

Concurrently, consider a slightly adapted form of our system (4.1) such that the Grass equation with $m = 2$ reads, $\tilde{q} = A_g H^{-1} q$. If we linearize about the modulus, then (4.4) is the same, except the third order term becomes:

$$S_3 = -A_g \frac{H}{H^3} q|\tilde{q}|(2 \tanh(x) x - 2\sigma + \sigma \tanh(x)^2 + \sigma \tanh(y)^2 + 2\tanh(y)y)/(H^3\alpha).$$

In this section we will refer to (4.4) as the exact solution, and (4.5) as the linearized form.

One of the first features of these solutions that stand out, is the sharp dynamic spike in the initial velocity profile near the origin, as seen in Figure 1. This particular feature of the solution represents, in the discontinuous basis, a fairly steep discontinuity relative to the jump condition (as captured in the large gradients in (4.3)), as opposed to the example from §4.1 that has a static rotating vector field with a constant gradient field. Nevertheless $q$ is an analytic function, which is just to say that it is equal to its Taylor series expansion locally, or about the origin its Maclaurin series expansion. In fact, all of the initial state variables are analytic functions of $x$, $y$ and $t$. Further note that the initial conditions for the exact (4.4) and linearized (4.5) fluxes are equivalent.

Naively, we might expect at this point to observe the full $p + 1$ rate of convergence in our solution space, as our initial conditions are analytic functions that satisfy manufactured frameworks in either exact or fully linearized forms. However, observe Table 2. It is immediately apparent that the “exact analytic manufactured solution” to our system (4.1) using (4.4) does not converge at the expected rate. If we perform a Maclaurin series expansion of the initial condition in $u$ in one direction (i.e., just in $x$ and holding the $y$ component constant), we can see that to twelfth order, we have a polynomial of the form:

$$u = x - (1 \times 10^2)x^3 + (5 \times 10^4)x^5 - (2 \times 10^5)x^7 + (4 \times 10^6)x^9 - (8 \times 10^7)x^{11} + O(12).$$
If we construct the Mercer-Roberts plot of this expansion (see the appendix to [40]), we find that to twelfth order the approximate radius of convergence for this function is \( r \simeq 0.05 \) (viz., for low polynomial order the radius of convergence vanishes). For linear basis functions then, this suggests that to recover this stability disc we need a minimum of 12 elements to span the interval \((0, r)\), which on the regular mesh domain \([-\frac{1}{2}, \frac{1}{2}]\) implies that we cannot expect proper convergence for linear elements before we pass the threshold \( h = 1/228\); the first stable configuration, with the expected convergence rate thereafter (notice that this is exactly the behavior we observe in [2]). It is interesting to note that the difference map from Figure 1 is maximized across the boundary of this disc \([0, r]\), and that we have also found that the type I dioristic energy regularity measure from [43] achieves a maximum across the boundary of the disc as well. This illustrates the first analytic subtlety that arises frequently in sedimentary systems predicated on formulas with sharp zero crossings (i.e., points in mathematical functions where the sign changes), which is: far from all fully analytic representations converge over practical ranges of \( h \), or at all over certain subdomains at fixed orders of \( p \).

Now let us look at the second case (4.5). Here we have the same conditions on the initial data, but presumably to make matters worse, our linearization about the modulus \(|H^{-1}q|\) is analytically ill-posed. That is, the Maclaurin series expansion of the full flux function is not analytic at all. For example, consider the solution along the line \( y = 0 \) in two dimensions, then \( \partial_y \tilde{q} = \partial_x (A_y h^{-1} H^{-1} q_x |q_x|) \), which is a nondifferentiable function at \((x, y) = (0, 0)\). Nevertheless, the existence of this type I jump discontinuity at the origin, as we see in Table 2 has no impact on the convergence rate of the solution, which, it turns out, is completely saturated rather by the error from the “envelope of the oscillatory structure” [40] of the analytic function itself, as expressed in the disc of convergence of the initial conditions above.

This brings into focus a number of additional nuances of the system (4.1). First, it should be noted as a general feature of nonlinear systems, our formulation is constructed to handle arbitrary forms of the Exner flux \( \tilde{q} \) and thus we cannot provide a guarantee that there are no elliptic modes present in the eigenstructure of the system. Nevertheless, we might be tempted to infer the hyperbolicity of (4.1) by linearization, which raises the important question: “how does one handle convergence of a solution with large local gradients managed with and relative to transport via nonconservative products?”

With regards to “large gradients,” the accuracy concern has been addressed by our two examples above, and really can not be overstated. Complementary to the accuracy concerns, the stability of the solution is addressed by the ubiquitous problem of providing slope and flux limited solutions in the discontinuous Galerkin framework [37, 42]. We have addressed this problem in detail in [42], and will primarily defer any discussion in this paper to our previous work. Let us simply recall that the salient features of that analysis were that: (1) generic solutions possessing “large local gradients” often demonstrate first order truncation error in the discontinuous basis (as further demonstrated in [5]) even when the solution is only discontinuous relative to the radius of convergence of a smooth function, (2) most slope limiting regimes saturate near first order for sharp gradient solutions, even when specifically constructed for higher order accuracy, and (3) solutions with varying local spectral order are more stably convergent near large gradients when reducing their local order (i.e., in a polynomial basis, the local degree).

Nevertheless there exists a second interesting feature in the context of sediment transport, which is that of how can one approach or view the convergent behavior in a “nonconservative product” (NCP) formalism. The standard answer to this question has been to create “path consistent” schemes that rely on a correction term to the standard numerical flux [16, 45, 49]. In the case of the Roe flux \( \Phi_{\text{Roe}} \) the nonconservative product form of the flux \( \Phi_{\text{NCP}} \), can be taken to be:

\[
\Phi_{\text{NCP}} = \begin{cases} 
(F_{hp})_1 \cdot (n_{ij})_1 \psi_{hp} |r_{ij} - v_{nc}|, & S |r_{ij}| > 0 \\
\Phi_{\text{Roe}} - \frac{(S |r_{ij} + S_{ji}|)}{(S_{ji} - S_{ij})} \psi_{nc} & S |r_{ij}| < 0 < S |r_{ji}| , \\
(F_{hp})_2 \cdot (n_{ij})_2 \psi_{hp} |r_{ji} + v_{nc}| & S |r_{ji}| < \end{cases}
\]
with interfacial wavespeeds in terms of the eigenvalues (2.9),

\[ S|\Gamma_{ij} = \min \left\{ \min_{i \leq 4} \varsigma_i |\Gamma_{ij}, \min_{i \leq 4} \varsigma_i |\Gamma_{ji} \right\}, \quad \text{and} \quad S|\Gamma_{ji} = \max \left\{ \max_{i \leq 4} \varsigma_i |\Gamma_{ij}, \max_{i \leq 4} \varsigma_i |\Gamma_{ji} \right\}. \]

The signature behavior of the correction term is determined by the inter-element jumps, as \( \nu_{nc} = \frac{1}{2}(0, g\{H_{hp}\}[b_{hp}][n^1,0]^T) \), and here we have just defined the average at the interface by, \( \{H_{hp}\} = \frac{1}{2}(H_{hp}|\Gamma_{ij} + H_{hp}|\Gamma_{ji}) \).

The elementary observation in a nonconservative hyperbolic system, is that the nonconservative products mathematical behavior is that of a Borel-measure source term relative to an inhomogeneous conservation law \( 16 27 15 49 \). This measure does not inherit the nascent convergence behavior implicit in the standard conservation law formalism; which begs the question, “when might we expect these types of systems to actually converge?” As shown in the recent papers of \( 1 \), \( 24 \), \( 38 \), the experimental/numerical convergence of “path-consistent nonconservative systems” in one dimension, where the precise path is exactly computable, still may not, in general, converge to the exact (i.e. correct) solution. In these cases, the convergence of the “consistent path” does not uniquely approach the “exact path” in the limit. This means, of course, that the error about and around local jumps has a rate of convergence that is predictably nonasymptotic to the anticipated order, since the path as anticipated by the interior boundaries is either constant or divergent from the exact solution. This behavior is seen even when the systems are constructed using a posteriori entropy consistent methods on the set of admissible solutions. This is a somewhat disheartening observation, since the jump conditions \( \{b_{hp}\} \) are nonvanishing even when projected over a relatively smooth initial condition.

Though largely beyond the scope of the present paper, it is worth mentioning that these difficulties in the shallow water setting have been well-established for some time. What seems to be a somewhat conventional suggestion, is that the natural dimensional extension of the Saint-Venant system (2.1) is only heuristically suitable in systems where the “altitude of the relief” (i.e., the slope of the local bathymetry \( b \)) admits only smooth representations with “relatively small” slopes \( \nabla_x b \) \( 6 8 23 28 29 \). By inference then, even just in the case of the Exner equation representation (nevermind the complications in the fully coupled hydrodynamic system), this is of course not always strictly the case, as many popular and common sedimentary flux formulations are in-and-of-themselves presented in terms of nondifferentiable zero crossing functions, which at the very least evolve by way of sharp gradient forcings. These functions can be easily smoothed however (even explicitly mollified), as long as one adopts a flexibility in the approach one takes to the empirical form of the sedimentary flux representation. This is a particularly beautiful lesson that the engineering provides the mathematics.

Moreover, in cases like these, derivations such as J. Restrepo’s Sand Ridge model might offer distinct advantages, due to the careful physical arguments that balance admissible bathymetric slopes before collapsing into smoothly varying sand avalanches with stabilizing viscous terms over characteristic internal wavelengths \( 50 \). Another general suggestion that seems to have found an audience in the mathematical modeling community, is to consider adapted two-dimensional shallow water type systems that are well-behaved under the extended requisite conditions of the models (such as when bathymetric slopes become steep) \( 23 \). Incidentally, these new derivations are constructed as to lead to fully conservative systems, sidestepping the difficulties introduced by the classical NCP-type formulations in the Saint-Venant model. They do not, however, come without a cost, as the conserved state variables are written in terms of, for example, the so-called bottom profile variable \( \tilde{v} = -\partial_t b - u \cdot \nabla_x b \), and somewhat physically non-intuitive scalar potential functions, such as \( V = u - \tilde{v} \partial_x b \). Likewise, it is not clear that such strongly coupled models are particularly well-suited to handle the divergent classes of models arising in the geophysics underlying the systems; as, for example, is encoded in the many different forms \( \tilde{q} \) takes in the Exner formalism. It is certainly not ideal (and possibly not even realistic) to have to perform a complicated mathematical reformulation each time the form of \( \tilde{q} \) changes.

At present, it seems unclear what the “best” solution to these difficult questions might be, particularly
with respect to the complex interactions between the underlying mathematics, the numerical methods, the scientific engineering, and the physics that all come to bear on the solution to the problem. However, let us suggest, at present at least, that the upshot of this discussion from the composite perspective might be summarized as follows: though it is immediately apparent that numerical jumps in the solution affect both the path-consistent behavior of the system as well as introducing truncation error, the following numerical observation persists in our analysis, and seemingly largely trumps the antecedent concerns; and that is — NCP formulations generally do and can converge as seen in §4.1 for smoothly varying bed morphologies when truncation error does not dominate, and only start to fail to the methods convergence order as the forcings tend towards numerically steep gradients. From the point of view of the numerical methods involved, this is an extremely important observation, that seems to strongly suggest that regardless of whether the system itself is fully conservative or not, the numerical method will quite likely become limited by first a TVD-like principle after projecting the initial state into a discontinuous basis, and second from error issuing from truncation due to linearizations in the nonlinear fluxes, long before concerns relating to NCP path-consistency start to dominate. Thus, unless one wants to use substantially more complicated numerical techniques to preserve discontinuities interior to discrete elements or track discontinuities along moving mesh faces (such as in generalized finite element methods [4]), being overly concerned with path-consistency and the error introduced due to the non-uniqueness of the NCP formalism is likely to lead to somewhat wasted effort in the generalized application-directed setting. Rather, it seems that a more productive approach to the problem might be in developing patching functions between domain-dependent flux representations \( \tilde{q} \) in the dynamic setting, such that more realistic application domains can be directly addressed.

§4.3 Hydrothermal vents

The application model that we present here, is that of seafloor hydrothermal vents that frequently develop along submarine landforms, such as seamounts, oceanic trenches and submarine ridges. These structures are of course observed from epipalegic regions (viz. sunlit or upper subsurface regions) all the way down into hadal zones (viz. deep ocean), where they are more common due to the relative abundance of upwelling magma in the subsurface. What characterizes hydrothermal vents from other oceanic chemical seeps (such as brine pools), is primarily their unusually high temperatures. Hydrothermal vents can approach \( \sim 400^\circ \text{C} \) near volcanic vents, which of course introduce interesting nonhydrostatic feature into the thermodynamic systems. These systems are of particular interest in that they lead to the formation of exotic and delicate aquatic ecosystems, that have been used for broad applications, such as aquatic mining of sulfide deposits, all the way to quests for understanding extremophile biodiversity (i.e., the biodiversity of organisms living under extreme thermodynamic conditions) and the formation of primitive atmospheres in understanding the origins of life.

For simplicity, here we simply consider a model hydrothermal vent at constant temperature, which is to say, we only consider the region proximal to the vent itself. Our model is meant to elicit the parameters characteristic of a “black smoker,” which spews hot plumes of iron monosulfide that leads to its “black smoke” appearance as seen in Figure 2. The vent velocity is held constant \( u \in \mathbb{R} \) well below pyroclastic levels, while the bathymetry \( b \) and free surface \( \zeta \) vary according to the following chemically active form (see [42, 44] for more background) of the two-dimensional Saint-Venant system:

\[
\begin{align*}
\partial_t (H \beta_j) + \nabla_x \cdot (H \beta_j u) - H \omega_j &= 0, \\
\partial_t (Hu) + \nabla_x \cdot (Hu \otimes u + \frac{1}{2} g H^2) &= g H \nabla_x b + \nabla_x \cdot (\eta \nabla_x H u) + S, \\
\partial_t b + \nabla_x \cdot \tilde{q} - \nabla_x \cdot (\mathcal{D} \nabla_x b) &= 0.
\end{align*}
\]
Figure 2: Here we show a black smoker from the Mariner vent site in the Pacific Ocean’s Eastern Lau Spreading Center, reproduced with permission [69].

with mass action law,

$$\mathcal{J}_j = \sum_{r \in \mathcal{R}} \left( \nu^b_{jr} - \nu^f_{jr} \right) \left( k_f \prod_{i=1}^n \beta^f_i r_i - k_b \prod_{i=1}^n \beta^b_i r_i \right)$$

given initial data, $H_0 = 2,

$$\zeta_0 = H_0 - b_0, \quad b_0 = 1 + f \varpi, \quad u_0 = f \varpi (\xi_x, \xi_y) / H_0, \quad \beta_{j,0} = k \varpi$$

where all shared variables are the same as in §4.1, except for the fact that there are now two reactive chemical constituents $\beta_j = \beta_j(t, \mathbf{x})$ for $j \in \{1, 2\}$. These constituents obey the law of mass action $\mathcal{J}_j = \mathcal{J}_j(\beta)$, with $\mathcal{D} = \eta = 0.001$, while the source term is chosen such that $S = 0$. Additionally we have set $k = e(-x^2-y^2)/0.002$ and $\varpi = 0.1$. The porosity is chosen for the sake of illustration such that the metaliferous mud $\rho_b$ and seawater $\rho_w$ densities are approximate $\rho_b \approx \rho_w$. For simplicity we set fully transparent boundary conditions on the domain boundary, such that for any unknown $v$ we have $v|_{\Gamma_{ij}} = v|_{\Gamma_{ji}}$ with the unit outward pointing normal preserving direction.

The chemical mass action $\mathcal{J}_j$ may also be viewed as a reaction term in the transport equation where we have neglected the usual Fickian diffusion to a first approximation, as discussed in our paper. Here, $k_f, k_b \in \mathbb{R}^+$ are the forward and backward reaction rate constants, and $\nu^f_{jr}, \nu^b_{jr} \in \mathbb{Z}$ are the corresponding constant stoichiometric coefficients given for reaction $r$ in the reaction space $\mathcal{R}$. See [13, 25, 26, 44, 46] for more details on these basic equations. Further note that the mass equation splits into the coupled transport system,

$$\partial_t H + \nabla_x \cdot (H \mathbf{u}) = 0, \quad \partial_t \beta_j + \mathbf{u} \cdot \nabla_x \beta_j - \mathcal{J}_j = 0. \quad (4.7)$$

As a test bed, we choose our remaining parameters based on the rainbow vent field located on the mid-Atlantic ridge [21], comprised of black smokers. These particular “black smokers” are hydrothermal vents characterized by mineral and chemical transport across a collection of superheated fissures that materialize in the form of relatively symmetric mineralized “chimneys” (again, see the inset in Figure 2 for example).
Figure 3: Here we show the evolved metaliferous mud bathymetry and corresponding concentrations [H+] after 1.5 days. On the top we see the RKC solution, and on the bottom we have the RKSSP solution both at second stage and second order.

As discussed in [21], two important reactions that occur in these chimneys are iron hydroxychloride formation reactions and serpentinization reactions [30]. Both of these reactions lead to large hydron concentrations near the plume center, which have the effect of generating relatively low pH environments. Two such reactions are given by:

\[
\begin{align*}
2\text{MgSiO}_3 \text{(enstatite)} + \text{Mg}^{2+} + 3\text{H}_2\text{O} & \xrightarrow{k_{f1}} \text{Mg}_3\text{Si}_2\text{O}_5(\text{OH})_4(\text{serpentine}) + 2\text{H}^+
\protect\tag{4.8} \\
2\text{FeCl}_2^0 + 3\text{H}_2\text{O} & \xrightarrow{k_{f2}} \text{Fe}_2(\text{OH})_3\text{Cl}(s) + 3\text{H}^+ + 3\text{Cl}^-
\end{align*}
\]

Restricting to the second reaction, the equilibrium constant has been empirically found to approximately satisfy

\[K_{eq} \approx [\text{Cl}]^3[\text{Fe}]^{-2}e^{-3pH},\]

with the measured rainbow vent pH \(\approx 2.8\), [Cl] \(\approx 750\ \text{mM}\), and [Fe] \(\approx 24000\ \mu\text{M}\), yielding the forward rate constant \(k_{f2}\). We rewrite (4.8),

\[
\mathfrak{N}_1 + \mathfrak{N}_2 \xrightarrow{k_{f2}} \mathfrak{N}_3 + 3\text{H}^+ + 3\text{Cl}^-
\]

with each \(\mathfrak{N}\) treated as excess constant bath constituents for purposes of demonstration (note this is chosen for simplicity and due to the relative stoichiometric weights). This leads to the coupled kinetic equations

\[
\partial_t \beta_i = 9k_{f2}[\mathfrak{N}_3]\beta_i^3 \beta_j^3, \quad \text{for}\ i \neq j,
\]
Figure 4: The time evolution of the metaliferous mud bathymetry using the RKC(2,2) scheme, starting at $t=0$ and using timesteps of $\sim 4$ hours, such that the final step is $\sim 32$ hours into the simulation.

which represent the $\partial_t \beta_j = \mathcal{A}_j$ part of the second equation in (4.7). The $\beta$'s are measured in molarity, such that the maximal initial pH determines that $\max[H^+] = 0.01M$. The concentration of iron rich constituents, on the other hand, are not treated chemically here, but rather as metaliferous muds that move via the sedimentary aggradation/degradation at the chimney mouth. In this way, we might view this particular model system as demonstrating the early development and formation of such a chimney structure from a seamount initial state.

With this simplified setup we are able to easily track the local oceanic acidification near and around the mouth of the hydrothermal vent (at least that produced by the single reaction pathway we consider). In Figure 3 we show the proton concentration around the chimney, which demonstrates how strongly coupled these concentrations are to the local bathymetry. Of course, as we have used the Grass equation here, the discharge flux is largely determined by the local velocity field, which is evident. All of our examples were run using an $h = 1/64$ mesh. Further it is worth noting that both solutions in Figure 3 are run to second order, but the RKC(2,2) solution exhibits substantially less numerical diffusion than the RKSSP(2,2) solution. This is entirely expected, as discussed and shown in detail in [57, 64], and is a recurring observation [44] in Runge-Kutta based DG methods. Finally, in Figure 4 we show the time evolution of the iron rich mud relatively to the velocity field forcing using the RKC(2,2) time integrator. The solution is remarkably stable even under these relatively large forcings (no slope limiting has been used here), and serves to illustrate the potential of the numerical method for studying complicated multicomponent reactive ocean dynamics.
§5 Conclusion

We have introduced a fully coupled shallow water system that couples sediment evolution to the dominant underlying hydrodynamic forcings in §2. Our solution is based on the two-dimensional extension of the Saint–Venant formulation, and is formulated to accommodate any multilayer sedimentary flux that both: a) satisfies the Exner formalism of geophysics, and b) can be linearized in such a way as to have a well-posed numerical representation. When these assumptions are met, we provide the exact eigendecomposition of the nonlinear flux matrix of the system, in its most generalized form.

Next in §3 we implemented this general multiphase system into a discrete framework. The solution is projected onto a degree $p$ discontinuous polynomial basis using a standard discontinuous Galerkin method, where a Roe flux is chosen to solve the strongly coupled Riemann problem for the quasi-hyperbolic subsystem, and we use the standard unified formulation for the parabolic subsystem. We implemented a number of Runge-Kutta time integrators to fully discretize our solution space.

In §4 we introduced some example experiments to probe the numerical accuracy and systemic nuance of the discrete solution to (2.1). Our first example is a simplified ideal system, chosen merely to demonstrate the convergence rates and expected behavior of such a method. In §4.2 we addressed one of the most important complications that arise in the system of PDEs. Namely, we emphasized how delicate the interplay between the mathematics (modelisation), numerics, and scientific applications are in these systems, and how even slight forcings in any one direction can lead to potential degradation in the solutions comprehensive veracity. Finally we presented an application model that couples an additional chemical subsystem in order to study the difficult problem of the behavior and evolution of hydrothermal vents on the ocean floor. We found that our solutions are surprisingly robust, even given extreme forcings and sharp initial gradients. This seems to indicate that strongly coupled systems of the form of (2.1) solved using DG methods are quite well-suited for a diverse and wide-range of application studies.

§6 Acknowledgements

The first author would like to acknowledge the support of the National Science Foundation grant NSF OCI-0749015. The second author would like to acknowledge the support of the National Science Foundation grant NSF DMS-0915223. The fourth author would like to acknowledge the support of the National Science Foundation grant, DMS-0915118 and DMS-1045151. The fifth and sixth authors would like to acknowledge the support of the National Science Foundation grant NSF OCI07-46232.

Appendix

The variables from the eigendecomposition problem are:

\[
\begin{align*}
\beta &= n_x n_y (c^2 \partial_{H_u} \tilde{q}_x + c^2 \partial_{H_u} \tilde{q}_y + v^2 \partial_{H_u} \tilde{q}_y + n_y (c^2 \partial_{H_v} \tilde{q}_y - uv \partial_{H_u} \tilde{q}_y - c^2 \partial_{H_u} \tilde{q}_x) \\
&+ n_y (c^2 \partial_{H_u} \tilde{q}_x - c^2 \partial_{H_u} \tilde{q}_x - v \partial_{H_u} \tilde{q}_y - v^2 \partial_{H_u} \tilde{q}_y + v^2 \partial_{H_u} \tilde{q}_x) + s_3 \partial_{H_u} \tilde{q}_x (n_y s_2 - n_y^2 v) \\
&+ (v \partial_{H_u} \tilde{q}_y - \nu s_2 \partial_{H_u} \tilde{q}_y - c \partial_{H_u} \tilde{q}_x) n_y n_x - (v \partial_{H_u} \tilde{q}_x + v^2 \partial_{H_u} \tilde{q}_x) n_x \\
&+ (\nu s_2 \partial_{H_u} \tilde{q}_x - c \partial_{H_u} \tilde{q}_x - c \partial_{H_u} \tilde{q}_x - c \partial_{H_u} \tilde{q}_x) n_y n_x - (c \partial_{H_u} \tilde{q}_x + v \partial_{H_u} \tilde{q}_x) n_x \\
\alpha &= n_x n_y (c^2 \partial_{H_v} \tilde{q}_x + c^2 \partial_{H_v} \tilde{q}_y + v^2 \partial_{H_v} \tilde{q}_y + n_y (c^2 \partial_{H_v} \tilde{q}_y - uv \partial_{H_v} \tilde{q}_y - c^2 \partial_{H_v} \tilde{q}_x) \\
&+ n_y (c^2 \partial_{H_v} \tilde{q}_x - c^2 \partial_{H_v} \tilde{q}_x - v \partial_{H_v} \tilde{q}_y - v^2 \partial_{H_v} \tilde{q}_y + v^2 \partial_{H_v} \tilde{q}_x) + s_1 \partial_{H_v} \tilde{q}_x (n_y s_2 - n_y^2 v) \\
&+ (c \partial_{H_v} \tilde{q}_y + \nu s_2 \partial_{H_v} \tilde{q}_y - \nu s_2 \partial_{H_v} \tilde{q}_y - (v \partial_{H_v} \tilde{q}_x + v^2 \partial_{H_v} \tilde{q}_x) n_x \\
&+ (c \partial_{H_v} \tilde{q}_y + \nu s_2 \partial_{H_v} \tilde{q}_y + c \partial_{H_v} \tilde{q}_y + c \partial_{H_v} \tilde{q}_y) n_y n_x - (v \partial_{H_v} \tilde{q}_x + v^2 \partial_{H_v} \tilde{q}_x) n_x \\
&+ (c \partial_{H_v} \tilde{q}_y + \nu s_2 \partial_{H_v} \tilde{q}_y + c \partial_{H_v} \tilde{q}_y + c \partial_{H_v} \tilde{q}_y) n_y n_x - (v \partial_{H_v} \tilde{q}_x + v^2 \partial_{H_v} \tilde{q}_x) n_x.
\end{align*}
\]


\[ \gamma_1 = n_y(c^3 - v^2s_2 - v^2c - s_2c^2) + vc^2 - s_1(s_2v - n_yv^2) + s_2v, \]
\[ \gamma_2 = n_y(-c^3 - v^2s_2 + v^2c - s_2c^2) + vc^2 - s_3(s_2v - n_yv^2) + s_2v, \]
\[ \gamma_3 = n_y^2uc^2 - n_xvngc^2, \quad \gamma_4 = csu - c^3nx, \]
\[ \chi = n_y^2\partial H_u\hat{q}_y + \partial H_v\hat{q}_x n_y^2 - \partial H_v\hat{q}_x + (\partial H_u\hat{q}_x - \partial H_v\hat{q}_y)n_y n_x, \]
\[ \iota = \gamma_2s_3n_y + n_y\gamma_1s_1 + n_xs_3\gamma_4 + n_xs_3\gamma_3 - n_xs_1\gamma_4 + n_xs_1\gamma_3, \]
\[ \nu_2 = v^2 - n_y^2c^2, \quad \nu_3 = -\chi^2 - \chi^2 + 2\chi s_2s_4, \quad \nu_4 = \beta(\gamma^2 - 2\gamma^2 - cs + s_2 + s_2), \]
\[ \nu_5 = -\alpha\gamma^2 + \alpha\gamma_2 - \alpha\gamma_2 + 2\alpha s_2s_4 - \alpha s_4, \quad \nu_6 = n_y^2c^2 + \nu_2c^2, \quad \nu_7 = \nu_3c^2 + \chi c_4, \]
\[ J_8 = \beta s_3 - \alpha\gamma_2 - \alpha\gamma_2 - 2\beta s_2s_4 + 2s_2s_4 + \alpha s_2 + \beta s_2 - \alpha s_4 + \beta s_4 - \beta s_4, \]
\[ J_9 = \chi J_6 + \nu_3n_y^2c^2 + \nu_2c^3, \quad B = \gamma_2 + c - \gamma_4, \quad A = \gamma_4 + c - \gamma_2, \]
\[ D = -\gamma_2\gamma_4J_9 + \gamma_2\gamma_3J_9 - \gamma_3\gamma_1J_9 - \gamma_3\gamma_1J_9 + \gamma_1\gamma_2n_y J_8 + (\gamma_3\gamma_1J_7 + \gamma_4\gamma_1J_7 - \gamma_2\gamma_3J_7 - \gamma_2\gamma_3J_7)n_y^2 + \gamma_3n_\gamma_1t_5 + \gamma_3n_\gamma_1t_5 + n_x\gamma_2\gamma_4t_4 - n_x\gamma_4\gamma_2t_4, \]
\[ E = -\gamma_1\gamma_1J_9 - \gamma_3\gamma_2J_9 + (-n_x\gamma_1\gamma_3J_7 - n_x\gamma_3\gamma_3J_7 + n_x\gamma_1\gamma_4J_7 - n_x\gamma_3\gamma_4J_7 + tJ_7)n_y + n_x\gamma_1\gamma_5 - n_x\gamma_3\gamma_4t_4, \]
\[ F = \gamma_1\gamma_3J_9 + \gamma_3\gamma_3J_9 - \gamma_1\gamma_4J_9 - \gamma_3\gamma_4J_9 - \gamma_2s_3n_y J_8 + (-\gamma_1\gamma_3J_7 - \gamma_3\gamma_3J_7 + \gamma_1\gamma_4J_7 - \gamma_3\gamma_4J_7)n_y^2 + n_x\gamma_1\gamma_5 - n_x\gamma_3\gamma_4t_5 + n_x\gamma_3\gamma_4t_4 + tJ_5 \]

References


