

UNIVERSIDAD DE CONCEPCIÓN



CENTRO DE INVESTIGACIÓN EN INGENIERÍA MATEMÁTICA (CI²MA)



Invariant-region-preserving high-order schemes for a model of
reactive sedimentation

JUAN BARAJAS-CALONGE, JULIO CAREAGA,
LUIS M. VILLADA

PREPRINT 2026-28

SERIE DE PRE-PUBLICACIONES

Invariant-region-preserving high-order schemes for a model of reactive sedimentation

JUAN BARAJAS-CALONGE[†] JULIO CAREAGA^{*‡} LUIS-MIGUEL VILLADA[§]

Abstract

Reactive sedimentation of mixtures composed by biological and inert solid matter including liquid substrates is modelled by a system of convection-diffusion-reaction equations. In this process, the small solid particles dispersed in the viscous fluid settle due to gravity force while simultaneous chemical reactions take place between solid and liquid components (substrates). Among the applications of reactive sedimentation processes, the principal use is in the simulation and control of secondary settling tanks (SSTs) in water resource recovery facilities (WRRFs). The spatially 1D governing equations include a nonlinear and non-homogeneous strongly degenerate parabolic equation for the total concentration of solids, and transport equations including reaction terms for the solid components and substrates. To accurately approximate the model equations, a high-order finite volume scheme is developed making use of polynomial reconstructions. Two reconstruction methods are used, a second-order monotonic upstream-centered scheme (MUSCL), and a third-order central weighted essentially non-oscillatory (CWENO3) method. In addition, for the time discretization, we employ strong stability preserving Runge-Kutta (SSPRK) schemes of second- and third-order. The CWENO3 spatial discretization is complemented with limiters, and in the case of zero diffusion, the designed high-order schemes are shown to preserve an invariant-region property. This property is particularly relevant as it ensures that physically relevant numerical solutions are obtained at each time iteration. In addition, high-order approximations for the diffusion terms are proposed to approximate the complete model equations. Simulations of denitrification in SSTs under a continuously operated regime in WRRFs demonstrate the performance of the model and its discretization. Finally, convergence tests experimentally show the theoretical order of accuracy of the developed schemes and the invariant-region preservation with and without diffusion.

Keywords: systems of conservation laws, reactive sedimentation, invariant region preservation, finite volume methods, polynomial reconstructions

Mathematics Subject Classification: 35L65, 65L06, 65M08, 76T20

1 Introduction

1.1 Scope

Reactive sedimentation describes the process in which a mixture of solid particles and a liquid phase, including dissolved substrates, begin to settle due to gravity force, while reactions between the solid and liquid components take place. Motivated by applications in water resource recovery facilities

*Corresponding author

[†]GIMNAP-Departamento de Matemática, Universidad del Bío-Bío, Chile and Departamento de Ciencias Naturales y Tecnología, Universidad de Aysén, Chile, email: juan.barajas2001@alumnos.ubiobio.cl.

[‡]Bernoulli Institute, University of Groningen, The Netherlands, email: j.c.careaga.solis@rug.nl.

[§]GIMNAP-Departamento de Matemática, Universidad del Bío-Bío, Chile and CI²MA, Universidad de Concepción, Casilla 160-C, Concepción, Chile, email: lvillada@ubiobio.cl.

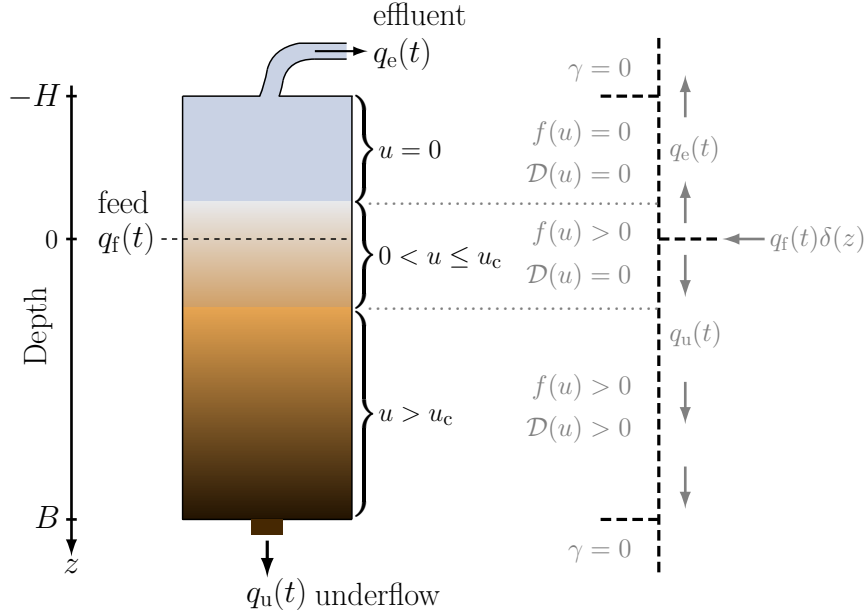


Figure 1.1: Schematic of a secondary settling tank (SST). The depth z corresponds to the spatial coordinate, ranging from $z = -H$ to $z = B$ for the inner tank, and $z < -H$ and $z \geq B$ for the pipes. The colours in the SST represent a profile for the total solids concentration u . The diagram to the right of the SST shows the flux f , compression function $\mathcal{D}$ and bulk velocities q_u (underflow), q_e (effluent) and q_f (feed) for the particular concentration profile u along the z -coordinate.

(WRRFs), reactive sedimentation has been shown to be particularly relevant in the simulation and control of secondary settling tanks (SSTs) where biological and inert matter interact with substrates dissolved in water [14, 16, 21]. The suspension within the SST, also called activated sludge, comprises a large amount of heterotrophic bacteria, typically considered as solid particles in mathematical models, which consume dissolved substrates and produce various compounds during and after metabolism. Among the biokinetic reactions between bacteria and substrates is the denitrification process, in which bacteria use nitrate (NO_3) or nitrite (NO_2) to produce nitrogen (N_2). In an SST, the aim is to recover clean water from the overflow (effluent) located at the top of the tank, and remove the concentrated sediment from the underflow at the bottom, see Figure 1.1. Then, simulations of reactive sedimentation in SSTs have to ensure a balance between the bulk flows (feed, effluent and underflow), while all concentrations remain positive, and the total concentration of solids does not exceed a maximum packing threshold. This motivates the use of the so-called invariant-region-preserving schemes [1, 4], which ensure that the approximated quantities remain physically relevant. Further applications of reactive sedimentation in the wastewater treatment field can be found in the so-called sequencing batch reactors (SBRs), in which a series of settling tanks are operated in a continuous process of feed and extraction of solid and liquid matter [6, 7].

The reactive sedimentation model in [4], which finds its basis on the underlying interaction between physical and biochemical processes, takes into account the highly nonlinear evolution of sedimentation and the interconnected reactions between solid and liquid components. In turn, the spatially 1D convection-diffusion-reaction modelling equations [4], in terms of the depth coordinate $z \in \mathbb{R}$ and time

$t > 0$, can be written as follows:

$$\frac{\partial \mathbf{c}}{\partial t} + \frac{\partial}{\partial z}(\mathcal{V}_c(u, z, t)\mathbf{c}) = \frac{\partial}{\partial z}\left(\gamma(z)\frac{\partial}{\partial z}\mathcal{B}(u)\mathbf{c}\right) + \mathbf{b}_c(\mathbf{c}, \mathbf{s}, z, t), \quad (1.1a)$$

$$\frac{\partial \mathbf{s}}{\partial t} + \frac{\partial}{\partial z}(\mathcal{V}_s(u, z, t)\mathbf{s}) = -\frac{\partial}{\partial z}\left(\frac{\gamma(z)\partial_z \mathcal{B}(u)u}{\rho - u}\mathbf{s}\right) + \mathbf{b}_s(\mathbf{c}, \mathbf{s}, z, t), \quad (1.1b)$$

$$\mathbf{c} = (c^{(1)}, \dots, c^{(n_c)})^T, \quad \mathbf{s} = (s^{(1)}, \dots, s^{(n_s)})^T, \quad u = u(z, t) = \sum_{i=1}^{n_c} c^{(i)}(z, t), \quad (1.1c)$$

where the unknowns are the vector of solid components $\mathbf{c} = \mathbf{c}(z, t) \in \mathbb{R}^{n_c}$, vector of substrate components $\mathbf{s} = \mathbf{s}(z, t) \in \mathbb{R}^{n_s}$, with $n_c, n_s \in \mathbb{N}$, and the total concentration of solids $u = u(z, t)$ which is a scalar quantity. Scalar functions $\mathcal{V}_c$ and $\mathcal{V}_s$ are related to convective velocities for the solids and liquid phases, respectively, while $\mathcal{B}$ accounts for sediment compressibility. Functions $\mathbf{b}_c \in \mathbb{R}^{n_c}$ and $\mathbf{b}_s \in \mathbb{R}^{n_s}$ group both reaction and feed-related terms (including a point source), γ is a characteristic function, which equals one inside the vessel and zero otherwise, and ρ is the density of the solid phase. Suitable initial conditions $\mathbf{c}(z, 0) = \mathbf{c}_0(z)$, $\mathbf{s}(z, 0) = \mathbf{s}_0(z)$ and $u(z, 0) = u_0(z)$ satisfying (1.1c) for $z \in \mathbb{R}$ are supplemented to (1.1), while no boundary conditions need to be imposed. All model ingredients, including the specific definition of functions, parameters, and constants of the system of partial differential equations (PDEs) (1.1), are described in detail in Section 2.

Although the convective terms in Equations (1.1a) and (1.1b) are linear with respect to the respective vector of unknowns, the total solids concentration u defined by (1.1c) is modelled by a non-linear convection-diffusion-reaction equation which is the result of the addition of each component of (1.1a). Furthermore, this equation is parabolic strongly degenerate, which becomes hyperbolic for u below a critical value, leading to solutions which may exhibit discontinuities, posing an additional complication to the numerical approximation. Furthermore, the low regularity related to discontinuous solutions also affects the order of accuracy of a numerical scheme. Therefore, special techniques need to be implemented to achieve high-order such as the so-called polynomial reconstructions [34, 23]. Polynomial reconstructions like the monotonic upstream-centered scheme (MUSCL) [24, 40] and weighted essentially non-oscillatory (WENO) reconstructions [29] allow to maintain a high-order accuracy, while keeping a good approximation of discontinuities.

The main aim of this work is to develop a high-order numerical scheme to approximate (1.1) making use of polynomial reconstructions for $\mathbf{c}$, $\mathbf{s}$ and u , and that at the same time satisfies an invariant-region-preserving property for the case $\mathcal{B} = 0$. For this, we will focus on the second-order MUSCL and a third-order WENO reconstruction, which combines linear and quadratic polynomials. The reconstructed polynomials of each solution component will be specially designed such that the invariant-region property is achieved, including the use of a modified version of the well known Zhang and Shu limiters [47, 44], which allow to preserve positivity and the maximum principle. Furthermore, suitable strong stability preserving Runge-Kutta (SSPRK) methods will be employed for the time approximations.

1.2 Related works

Models of continuous sedimentation of monodispersed particles in the non-reactive case and in one spatial dimension can be found in [11, 15, 12, 10]. For a review on models of non-reactive sedimentation for SSTs, including applications in water science, we refer to [30]. A common constitutive assumption in these models is the settling velocity which nonlinearly depends on the volume fraction of solids particles in agreement with Kynch's theory [26]. Furthermore, compression effects where sediment particles get into contact with each other and form a network for concentrations above a critical value $u_c > 0$ are incorporated as a degenerate diffusion term into the model equation [10]. Another

important complication, as described in [15], is the discontinuous flux function, which arises as a result of the continuously operated regime.

The biokinetic reactions that take place in SSTs are typically described by the comprehensive framework of activated sludge models (ASM) [20, 21], which characterize the reactions between the solids and substrates by means of systems of ordinary differential equations (ODEs). The first ODE-based activated sludge model is the ASM1 [20] and further standardised models are ASM2, ASM2d, and ASM3 [21], that differ according to the level of biological complexity they represent [16]. Reduced models focusing on the biokinetic reactions related to nitrification and denitrification processes can be found in [14, 22, 33, 27, 41]. For these reduced models, the key variables are the concentrations of diluted nitrate (NO_3), nitrite (NO_2) and produced nitrogen (N_2). One of the first PDE based models for reactive (batch) sedimentation was developed in [5], where the governing equations correspond to a convection-diffusion-reaction system of equations written for a reduced biokinetic model of denitrification taking place in the SSTs. This model was later extended to include an arbitrary number of solids and substrate components for the case of continuous sedimentation [4, 9], which corresponds to the model studied in this work. Applications to sequencing batch reactors can be found in the series of works [6, 7, 8], where the main difficulty is to correctly approximate a moving boundary. Although the numerical methods developed to simulate these reactive sedimentation models are shown to satisfy an invariant-region preserving property, they are only first-order accurate. Furthermore, the presence of shocks and discontinuities in the solutions require a special treatment in order to accomplish high-order approximations.

For high-order numerical approximations, some of the most widely used techniques for nonlinear conservation laws are the so-called polynomial reconstruction methods [28], which compute the numerical flux by locally using polynomial representations for the left and right interface states from neighbouring solution values. The weighted essentially non-oscillatory (WENO) reconstructions were introduced in [34], which are methods designed to avoid spurious oscillations near discontinuities by adaptively combining several candidate reconstructions. Several extensions to WENO schemes have been developed in the literature [13, 31, 44, 46, 48], see also the review provided in [37]. In [13], a central WENO (CWENO) scheme is introduced, whose main characteristic is that the polynomial reconstructions may include polynomials of different degrees. Multi-resolution WENO schemes are developed in [31, 48] designed such that high-order accuracy is achieved through adaptive reconstruction across multiple spatial scales.

Invariant-region properties, which may combine positivity and maximum-principle preservation for more than one variable in a PDE system, plays a crucial role to ensure that numerical solutions are physically relevant in many contexts. For instance, numerical schemes that preserve the positivity of density and pressure have been developed for the Euler equations of compressible gas dynamics [45, 47], while positivity-preserving methods have also been proposed for the shallow water equations [42, 43] and for polydisperse sedimentation models [1, 2]. Finally, a comprehensive overview of property-preserving schemes for conservation laws, together with an extensive list of references, can be found in the monograph by Kuzmin and Hajduk [25].

1.3 Outline of the paper

The paper is organised as follows. In Section 2, we introduce the model of reactive sedimentation, including the governing equations, functions and parameters. The spatial approximation of the model equations is introduced in Section 3. The polynomial reconstructions are presented in Section 4, where in particular the MUSCL scheme is discussed in Section 4.1 and the CWENO scheme is treated in Section 4.2. Additional limiters for the CWENO reconstructions are included in Section 4.3, and the time approximations for the second- and third-order schemes are introduced in Section 5. The

theoretical results related to the invariant-region preserving of the high-order scheme for the case without compression effects are included in Section 6. In Section 7, we present numerical examples to show the theoretical order of accuracy of the scheme, the invariant-region property and reliability of the developed high-order scheme. Finally, concluding remarks are included in Section 8.

2 Model equations

To model the one-dimensional reactive sedimentation process, we first consider that the sedimentation takes place in a tank whose depth z varies from $z = -H$ (height) to $z = B$ (bottom), with $H, B \geq 0$. Furthermore, we assume that the process is primarily influenced by gravity, so horizontal variations are negligible. For the mixture, we consider $n_c \in \mathbb{N}$ solid components and $n_s \in \mathbb{N}$ substrates, defined in vector notation as $\mathbf{c} := (c^{(1)}, c^{(2)}, \dots, c^{(n_c)})^\mathbf{t}$ and $\mathbf{s} := (s^{(1)}, s^{(2)}, \dots, s^{(n_s)})^\mathbf{t}$, respectively, both considered vectors of unknowns. The vector of solid concentrations $\mathbf{c} \in \mathbb{R}^{n_c}$ typically contains concentrations of heterotrophic organisms and non-degradable matter, while the vector of substrates $\mathbf{s} \in \mathbb{R}^{n_s}$ contains concentrations of substrates dissolved in water. The total solids' concentration u , defined as the sum of components of $\mathbf{c}$, is also considered a quantity of interest, therefore there are $n_c + n_s + 1$ unknowns in total, all of them functions depending on the spatial coordinate $z \in \mathbb{R}$ and time $t > 0$. The system of PDEs governing this process is given by [4]

$$\frac{\partial c^{(i)}}{\partial t} + \frac{\partial}{\partial z} \left((q(z, t) + \gamma(z)(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u))) c^{(i)} \right) = b_c^{(i)}(\mathbf{c}, \mathbf{s}, z, t), \quad \text{for } i = 1, \dots, n_c, \quad (2.1a)$$

$$\frac{\partial s^{(j)}}{\partial t} + \frac{\partial}{\partial z} \left(\left(q(z, t) - \frac{\gamma(z)(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u))u}{\rho - u} \right) s^{(j)} \right) = b_s^{(j)}(\mathbf{c}, \mathbf{s}, z, t), \quad \text{for } j = 1, \dots, n_s, \quad (2.1b)$$

$$u(z, t) = c^{(1)}(z, t) + \dots + c^{(n_c)}(z, t), \quad (2.1c)$$

where $v_{\text{hs}} := v_{\text{hs}}(u)$ is the hindered settling velocity, defined such that $v_{\text{hs}} \in C^1([0, u_{\text{max}}])$, with $v'_{\text{hs}}(u) \leq 0$ and $v_{\text{hs}}(u_{\text{max}}) = 0 \leq v_{\text{hs}}(u) \leq v_0 = v_{\text{hs}}(0)$, for $0 \leq u \leq u_{\text{max}}$, and there exists a positive constant L_v for which

$$|v_{\text{hs}}(u) - v_{\text{hs}}(\tilde{u})| \leq L_v |u - \tilde{u}|, \quad \text{for all } 0 \leq u, \tilde{u} \leq u_{\text{max}}, \quad (2.2)$$

which corresponds to the Lipschitz continuity condition of v_{hs} . Furthermore, $q := q(z, t)$ is the volume-average bulk velocity defined as

$$q(z, t) = \begin{cases} -q_e(t) & \text{if } z \leq 0, \\ q_u(t) & \text{if } z > 0, \end{cases}$$

where $q_u + q_e = q_f$ with q_f , q_e and q_u being the feed, underflow and effluent bulk flows, time-dependent functions respectively. Function $\mathcal{B} := \mathcal{B}(u)$ describes the sediment compressibility, which is zero for $u < u_c$ where $u_c \in \mathbb{R}^+$ is a critical concentration above which the particles come into contact, creating a network capable of bearing stress. The constant density of the solid phase is $\rho > 0$, and the source terms $\mathbf{b}_c = (b_c^{(1)}, \dots, b_c^{(n_c)})^\mathbf{t} \in \mathbb{R}^{n_c}$ and $\mathbf{b}_s = (b_s^{(1)}, \dots, b_s^{(n_s)})^\mathbf{t} \in \mathbb{R}^{n_s}$ are vector functions that involve feed and reaction terms corresponding to equations (2.1a) and (2.1b), respectively defined by

$$\begin{aligned} \mathbf{b}_c(\mathbf{c}, \mathbf{s}, z, t) &= \delta(z) q_f(t) \mathbf{c}_f(t) + \gamma(z) \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{c}, \mathbf{s}), \\ \mathbf{b}_s(\mathbf{c}, \mathbf{s}, z, t) &= \delta(z) q_f(t) \mathbf{s}_f(t) + \gamma(z) \boldsymbol{\sigma}_s \mathbf{r}(\mathbf{c}, \mathbf{s}). \end{aligned}$$

Here δ is the Dirac distribution, $\mathbf{c}_f$ and $\mathbf{s}_f$ are the vectors of feed concentration of solids and liquid components, respectively, $0 \leq \mathbf{r}(\mathbf{c}, \mathbf{s}) \in \mathbb{R}^{n_r}$ is a vector of nonlinear functions modelling the reaction

processes, with n_r a positive integer, $\boldsymbol{\sigma}_c \in \mathbb{R}^{n_c \times n_r}$ and $\boldsymbol{\sigma}_s \in \mathbb{R}^{n_s \times n_r}$ are constant dimensionless stoichiometric matrices. We further define b_u as the sum of the components of $\mathbf{b}_c$, that is

$$b_u(\mathbf{c}, \mathbf{s}, z, t) := \sum_{i=1}^{n_c} b_c^{(i)}(\mathbf{c}, \mathbf{s}, z, t) = \delta(z)q_f(t)u_f(t) + \gamma(z)\mathbf{1}^\top \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{c}, \mathbf{s}),$$

where $\mathbf{1}$ is the vector of ones in $\mathbb{R}^{n_c}$. Moreover, we define the maximum concentration of solids $u_{\max} \in \mathbb{R}^+$ and assume that $u_{\max} < \rho$. We observe that adding up (2.1a) and using (2.1c) one obtains the following non-linear scalar equation for the total concentration u :

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial z} \left((q(z, t) + \gamma(z)(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u)))u \right) = b_u(\mathbf{c}, \mathbf{s}, z, t), \quad (2.3)$$

where the product $v_{\text{hs}}(u)u = f(u)$ corresponds to the nonlinear batch flux function [26], and the compression function for the total solids concentration is defined by a function $\mathcal{D} := \mathcal{D}(u)$ such that $\partial_z \mathcal{D}(u) = (\partial_z \mathcal{B}(u))u$. In the non-reactive case [10], function $\mathcal{D}$ is defined depending on v_{hs} and the effective solids stress $\sigma_e := \sigma_e(u)$, which is positive only when $u > u_c$ and $\sigma_e(u) = 0$ for $u \leq u_c$. These compression-related functions are then defined by the integrals

$$\mathcal{D}(u) := \int_{u_c}^u d(s)ds \quad \text{and} \quad \mathcal{B}(u) = \int_{u_c}^u \frac{d(s)}{s}ds \quad \text{with} \quad d(u) := \frac{\rho}{g(\rho - \rho_L)} v_{\text{hs}}(u) \sigma_e'(u),$$

where $0 < \rho_L < \rho$ is the density of the liquid phase. Note that v_{hs} is typically assumed to be zero for values above $u_{\max}$. Equation (2.3) is then strongly degenerate parabolic due to the fact that $\mathcal{D}$ vanishes in the u -interval $[0, u_c]$ and it degenerates into a first-order hyperbolic equation. The final system of equations for the unknowns u , $\mathbf{c}$ and $\mathbf{s}$ is given by

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial z} \left((q(z, t) + \gamma(z)(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u)))u \right) = b_u(\mathbf{c}, \mathbf{s}, z, t), \quad (2.4a)$$

$$\frac{\partial \mathbf{c}}{\partial t} + \frac{\partial}{\partial z} \left((q(z, t) + \gamma(z)(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u)))\mathbf{c} \right) = \mathbf{b}_c(\mathbf{c}, \mathbf{s}, z, t), \quad (2.4b)$$

$$\frac{\partial \mathbf{s}}{\partial t} + \frac{\partial}{\partial z} \left(\left(q(z, t) - \gamma(z) \frac{(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u))u}{\rho - u} \right) \mathbf{s} \right) = \mathbf{b}_s(\mathbf{c}, \mathbf{s}, z, t). \quad (2.4c)$$

Note that in the system above, condition (2.1c) has been used to obtain Equation (2.4a), hence (2.1) implies (2.4). On the contrary, the opposite implication to go from (2.4) to (2.1) can be understood in the following way. Adding up (2.4b) by components and subtracting (2.4a) we get:

$$\frac{\partial}{\partial t} \vartheta + \frac{\partial}{\partial z} \left((q(z, t) + \gamma(z)(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u)))\vartheta \right) = 0, \quad (2.5)$$

with $\vartheta := u - (c^{(1)} + \dots + c^{(n_c)})$. The new variable ϑ satisfies then the linear advection equation (2.5) with homogeneous initial and boundary conditions. Now, under the assumption that u is a given function and that the unique solution to (2.4) is the homogeneous solution $\vartheta = 0$, then condition (2.1c) follows directly, and therefore in this case system (2.4) would imply (2.1). A proper proof requires to study the entropy solutions to (2.1) and regularity of $q(z, t) + \gamma(z)(v_{\text{hs}}(u) - \partial_z \mathcal{B}(u))$, which is out of the scope of this work. In [4], it is shown that choosing an appropriate first-order numerical scheme for (2.4), condition (2.1c) is fulfilled at every time iteration. For the high-order numerical scheme described in Section 3, we are going to also show that condition (2.1c) is fulfilled for the case when the compression effects are not considered.

To simplify the notation used in the spatial discretization of system (2.4) in Section 3, we introduce

the convective velocity, diffusion-related fluxes and total flux of (2.4a):

$$\mathcal{V} := \mathcal{V}(u, z, t) = q(z, t) + \gamma(z)v_{\text{hs}}(u), \quad (2.6)$$

$$\mathcal{J} := \mathcal{J}(u, \partial_z u, z) = \gamma(z)\partial_z \mathcal{D}(u), \quad (2.7)$$

$$\mathcal{E} := \mathcal{E}(u, \partial_z u, z) = \gamma(z)\partial_z \mathcal{B}(u), \quad (2.8)$$

$$\mathcal{F} := \mathcal{F}(u, z, t) = (\mathcal{V}(u, z, t) - \mathcal{E}(u, \partial_z u, z))u. \quad (2.9)$$

In addition, the vector fluxes corresponding to (2.4b) and (2.4c), are respectively defined by

$$\begin{aligned} \mathcal{G}^c(\mathbf{c}, u, \partial_z u, z, t) &:= (\mathcal{V}(u, z, t) - \mathcal{E}(u, \partial_z u, z))\mathbf{c}, \\ \mathcal{G}^s(\mathbf{s}, u, \partial_z u, z, t) &:= \left(\rho q(z, t) - \mathcal{V}(u, z, t)u + \mathcal{E}(u, \partial_z u, z)u \right) \frac{\mathbf{s}}{\rho - u}. \end{aligned}$$

The set of physically relevant solutions (invariant-region set) is defined as:

$$\mathcal{W} := \left\{ (u, \mathbf{c}, \mathbf{s}) \in \mathbb{R} \times \mathbb{R}^{n_c} \times \mathbb{R}^{n_s} : 0 \leq u \leq u_{\max}, \quad \mathbf{1}^t \mathbf{c} = u, \quad \mathbf{c} \geq \mathbf{0}, \quad \mathbf{s} \geq \mathbf{0} \right\}. \quad (2.10)$$

Finally, we summarise all the hypotheses related to the reaction terms required for our model in the following assumption.

Assumption 2.1. *Let $\boldsymbol{\xi} \in \{\mathbf{c}, \mathbf{s}\}$ and consider the corresponding number of components $n_{\boldsymbol{\xi}}$ (n_c or n_s depending on $\boldsymbol{\xi}$). Given $k = 1, \dots, n_{\boldsymbol{\xi}}$, let $\mathcal{J}_{\boldsymbol{\xi},k}^-$ and $\mathcal{J}_{\boldsymbol{\xi},k}^+$ be respectively the set of negative and non-negative coefficients related to the k row of the stoichiometric matrix $\sigma_{\boldsymbol{\xi}}$:*

$$\mathcal{J}_{\boldsymbol{\xi},k}^- := \left\{ j \in \mathbb{N} : \sigma_{\boldsymbol{\xi}}^{(k,j)} < 0 \right\}, \quad \mathcal{J}_{\boldsymbol{\xi},k}^+ := \left\{ j \in \mathbb{N} : \sigma_{\boldsymbol{\xi}}^{(k,j)} \geq 0 \right\}.$$

The stoichiometric matrices σ_c and σ_s , and vector of reaction rates $\mathbf{r}$ satisfy the following hypotheses

1. The vector of ones belongs to $\ker(\sigma_c^t, \sigma_s^t)$.
2. For each $j = 1, \dots, n_r$, there holds $r^{(j)} \in C(\bar{\mathbb{R}}_+^{n_c}, \bar{\mathbb{R}}_+^{n_s})$, with $r^{(j)}$ locally Lipschitz continuous.
3. Given $k = 1, \dots, n_{\boldsymbol{\xi}}$, for every $j \in \mathcal{J}_{\boldsymbol{\xi},k}^-$ there exists a bounded function $\bar{r}_{\boldsymbol{\xi}}^{(j)} \in C(\bar{\mathbb{R}}_+^{n_c}, \bar{\mathbb{R}}_+^{n_s})$ with $\bar{r}_{\boldsymbol{\xi}}^{(j)}(\mathbf{c}, \mathbf{s}) > 0$ if $\mathbf{c} > \mathbf{0}$ and $\mathbf{s} > \mathbf{0}$, and $\bar{r}_{\boldsymbol{\xi}}^{(j)}(\mathbf{c}, \mathbf{s}) = 0$ if $\mathbf{c} = \mathbf{0}$ and $\mathbf{s} = \mathbf{0}$, such that

$$r^{(j)}(\mathbf{c}, \mathbf{s}) = \bar{r}_{\boldsymbol{\xi}}^{(j)}(\mathbf{c}, \mathbf{s})\xi^{(k)}.$$

There exists an upper bound $R > 0$ such that $\bar{r}_{\boldsymbol{\xi}}^{(j)}(\mathbf{c}, \mathbf{s}) \leq R$ for all $j \in \mathcal{J}_{\boldsymbol{\xi},k}^-$ and all $(u, \mathbf{c}, \mathbf{s}) \in \mathcal{W}$.

4. There exists an $R_{\max} > 0$ such that each component of $|\sigma_{\boldsymbol{\xi}} \mathbf{r}(\mathbf{c}, \mathbf{s})|$ is bounded by $R_{\max}$ for all $(u, \mathbf{c}, \mathbf{s}) \in \mathcal{W}$.
5. There exists an $\varepsilon > 0$ such that

$$\sigma_c \mathbf{r}(\mathbf{c}, \mathbf{s}) = \mathbf{0} \quad \text{for all } u = \mathbf{1}^t \mathbf{c} = c^{(1)} + \dots + c^{(n_c)} \geq u_{\max} - \varepsilon.$$

Remark. The first assumption is made in order to guarantee conservation of mass, while the second one is to secure that in absence of spatial variations, the resulting ODE system is well-posed. The third assumption ensures that the reaction terms tend to zero when the vectors of concentration are zero and gives a bound to control the reaction terms having negative stoichiometric indexes. Assumptions 4 and 5 are made to properly bound the reaction terms when showing the invariant-region property in Section 6.

3 Spatial approximation

We begin by discretizing the spatial domain within the tank $\Omega := [-H, B]$ into $N \in \mathbb{N}$ cells of length $h = (H + B)/N$ defined by the intervals $[z_{j-1/2}, z_{j+1/2}] =: I_j$, where the cell-limits are $z_{j+1/2} = -H + jh$, with $j = 0, \dots, N$. The cell centre of each j -cell is denoted by z_j . The unknowns at each j -cell are approximated by cell sliding averages of the form

$$u_j(t) = \frac{1}{h} \int_{I_j} u(z, t) dz, \quad j = 0, \dots, N,$$

where $u_j(t) \approx u(z, t)$ for all $z \in I_j$ and time $t > 0$. Similarly, vectors $\mathbf{c}_j(t)$ and $\mathbf{s}_j(t)$ are defined component-wise as cell averages approximating $\mathbf{c}(z, t)$ and $\mathbf{s}(z, t)$ for all $z \in I_j$ and time $t > 0$, respectively. Then, system (2.4) is approximated by a finite volume numerical scheme based on piecewise constant solutions. The discretized indicator function is given by

$$\gamma_{j+1/2} = \gamma(z_{j+1/2}) := \begin{cases} 1 & \text{if } z_{j+1/2} \in (-H, B), \\ 0 & \text{otherwise,} \end{cases}$$

and the discretized Kronecker delta is

$$\delta_{j, j_f} := \begin{cases} 1 & \text{if } j = j_f, \\ 0 & \text{otherwise,} \end{cases}$$

where $j_f = \lceil H/h \rceil$. Furthermore, we approximate the known spatially discontinuous and time-dependent function q as

$$q_{j+1/2}(t) = \begin{cases} -q_e(t) & j \leq j_f, \\ q_u(t) & j > j_f. \end{cases} \quad (3.1)$$

To approximate the convective fluxes in the upcoming sections, we introduce the upwind operator

$$\text{Upw}(a; b, c) := \max\{a, 0\}b + \min\{a, 0\}c \quad \text{for } a, b, c \in \mathbb{R}, \quad (3.2)$$

where a in this case plays the role of the convective velocity, while b and c would correspond to the left and right states in a standard discretization, respectively. Note that operator Upw corresponds to a two-point Lipschitz monotone flux, which is nondecreasing in the second argument and nonincreasing in the third argument. We will also use the following notation

$$a^+ = \max\{a, 0\}, \quad a^- = \min\{a, 0\}, \quad \text{for } a \in \mathbb{R}.$$

Note that with this notation, the upwind operator can be written as

$$\text{Upw}(a; b, c) = a^+ b + a^- c.$$

3.1 Total concentration equation

We begin by describing the finite volume approximation of Equation (2.4a) which combined with a monotonic upstream-centered scheme (MUSCL) or weighted essentially non-oscillatory (WENO) reconstructions [36, 29] gives rise to a high-order scheme for the total concentration u . To do so, we integrate Equation (2.4a) over each cell interval I_j and apply integration by parts to obtain:

$$\int_{I_j} \partial_t u \, dz + \widehat{\mathcal{F}}_{j+1/2} - \widehat{\mathcal{F}}_{j-1/2} = \int_{I_j} b_u(\mathbf{c}, \mathbf{s}, z, t) \, dz, \quad (3.3)$$

for all $j = 1, \dots, N$, where $\widehat{\mathcal{F}}_{j+1/2}$ corresponds to the discretization of the total flux (2.9) at the cell boundary $z = z_{j+1/2}$ and time $t > 0$, respectively, whose precise definitions will be described next. To define $\widehat{\mathcal{F}}_{j+1/2}$, we follow the approach from [4] and make use of an upwind scheme with discrete velocity $\widehat{\mathcal{V}}_{j+1/2} - \mathcal{E}_{j+1/2}$ corresponding to the approximation of $\mathcal{V}$ in (2.6) and $\mathcal{E}$ in (2.8), respectively, which we then combine with polynomial reconstructions. To achieve the desired higher-order approximation, we consider a polynomial reconstruction of u on each interval I_j , given by polynomials $p_j^u := p_j^u(z)$ of degree $\ell - 1$, with ℓ a positive integer. Then, for each cell interface $j + 1/2$, the left and right polynomial reconstructions are respectively denoted by $\widehat{u}_{j+1/2}^L$ and $\widehat{u}_{j+1/2}^R$, and the numerical flux is defined by

$$\widehat{\mathcal{F}}_{j+1/2} := \text{Upw}\left(\widehat{\mathcal{V}}_{j+1/2} - \mathcal{E}_{j+1/2}; \widehat{u}_{j+1/2}^L, \widehat{u}_{j+1/2}^R\right), \quad (3.4)$$

where the discrete velocity $\widehat{\mathcal{V}}_{j+1/2}$ is given by

$$\widehat{\mathcal{V}}_{j+1/2} := q_{j+1/2} + \gamma_{j+1/2} v_{\text{hs}}(\widehat{u}_{j+1/2}^R).$$

The left and right reconstructed values of u at $z_{j+1/2}$ are computed using the polynomial reconstructions at each side of the interface by $\widehat{u}_{j+1/2}^L := p_j^u(z_{j+1/2})$ and $\widehat{u}_{j+1/2}^R := p_{j+1}^u(z_{j+1/2})$, respectively. One of the main properties that p_j^u satisfies, which is going to be widely employed in the next sections, is the cell-average preservation:

$$u_j = \frac{1}{h} \int_{I_j} p_j^u(z) dz, \quad \text{for } j = 1, \dots, N. \quad (3.5)$$

A full description of how to determine the polynomials p_j^u is going to be provided in Section 4.

To obtain a fully second- and third-order discretization, the diffusive term $\mathcal{E}$ must be approximated with at least the same order of accuracy as the convective flux, which is going to be of order two for the MUSCL scheme and of order three for the WENO scheme in Section 4. Furthermore, as there exist two functions related to the compression effects, we propose two approaches to compute $\mathcal{E}_{j+1/2}$, one based on the standard compression function $\mathcal{D}$, called “ $\mathcal{D}$ -method”, and a second option using directly function $\mathcal{B}$, called “ $\mathcal{B}$ -method”. For the $\mathcal{D}$ -method, we first discretize $\partial_z \mathcal{D}$ through the following second and fourth order related formulas

$$\mathcal{J}_{j+1/2}^{\text{2nd}} := \gamma_{j+1/2} \frac{\mathcal{D}(u_{j+1}) - \mathcal{D}(u_j)}{h}, \quad (3.6)$$

$$\mathcal{J}_{j+1/2}^{\text{4th}} := \gamma_{j+1/2} \frac{-\mathcal{D}(\tilde{u}_{j+2}) + 27\mathcal{D}(\tilde{u}_{j+1}) - 27\mathcal{D}(\tilde{u}_j) + \mathcal{D}(\tilde{u}_{j-1})}{24h}, \quad (3.7)$$

where we use the following fourth-order auxiliary reconstruction of u at the cell centre z_j from the cell averages:

$$\tilde{u}_j := u_j - \frac{1}{24}(u_{j+1} - 2u_j + u_{j-1}). \quad (3.8)$$

Then, given $\mathcal{J}_{j+1/2} = \mathcal{J}_{j+1/2}^{\text{2nd}}$ for the second-order scheme or $\mathcal{J}_{j+1/2} = \mathcal{J}_{j+1/2}^{\text{4th}}$ for a fourth-order approximation of $\mathcal{J}$, we discretize $\mathcal{E}$ making use of the relation $(\partial_z \mathcal{B}(u))u = \partial_z \mathcal{D}(u)$, which connects functions $\mathcal{D}$ and $\mathcal{B}$. Therefore, we can define $\mathcal{E}_{j+1/2}$ by the following quotient:

$$\mathcal{E}_{j+1/2} = \begin{cases} \frac{\mathcal{J}_{j+1/2}}{\bar{u}_{j+1/2}} & \text{if } \bar{u}_{j+1/2} := \frac{1}{2}(\widehat{u}_{j+1/2}^L + \widehat{u}_{j+1/2}^R) \geq u_c, \\ 0 & \text{otherwise,} \end{cases} \quad (3.9)$$

where we have also used the reconstructions $\widehat{u}_{j+1/2}^L$ and $\widehat{u}_{j+1/2}^R$.

For the $\mathcal{B}$ -method, we directly discretize $\partial_z \mathcal{B}$ using the second- and fourth-order formulas introduced for $\partial_z \mathcal{D}$ in (3.6) and (3.7), respectively, such that we get the two following approximations for $\mathcal{E}$:

$$\mathcal{E}_{j+1/2}^{2\text{nd}} := \gamma_{j+1/2} \frac{\mathcal{B}(u_{j+1}) - \mathcal{B}(u_j)}{h}, \quad (3.10)$$

$$\mathcal{E}_{j+1/2}^{4\text{th}} := \gamma_{j+1/2} \frac{-\mathcal{B}(\tilde{u}_{j+2}) + 27\mathcal{B}(\tilde{u}_{j+1}) - 27\mathcal{B}(\tilde{u}_j) + \mathcal{B}(\tilde{u}_{j-1})}{24h}, \quad (3.11)$$

where $\tilde{u}_j$ is defined in (3.8). Then, depending on the order of the approximation, we choose either the second-order $\mathcal{E}_{j+1/2} = \mathcal{E}_{j+1/2}^{2\text{nd}}$ or the fourth-order approach $\mathcal{E}_{j+1/2} = \mathcal{E}_{j+1/2}^{4\text{th}}$.

For the right-hand side of (3.3), we use the definition of the Dirac distribution and a quadrature rule for the integral involving the reaction terms. In turn, we approximate the integral of the reaction terms making use of the polynomial reconstructions of $\mathbf{c}$ and $\mathbf{s}$ over the cell $I_j = [z_{j-1/2}, z_{j+1/2}]$, respectively denoted by $\mathbf{p}_j^c$ and $\mathbf{p}_j^s$, and employing a G -point Legendre-Gauss-Lobatto quadrature rule with weights $w_1, w_2, \dots, w_G$ for the interval $[-\frac{1}{2}, \frac{1}{2}]$ such that $w_1 + \dots + w_G = 1$, and with quadrature nodes given by

$$S_j := \left\{ z_{j-1/2} = z_j^1, z_j^2, \dots, z_j^{G-1}, z_j^G = z_{j+1/2} \right\}, \quad \text{for all } j = 0, 1, \dots, N. \quad (3.12)$$

Then, we have

$$\int_{I_j} b_u(\mathbf{c}, \mathbf{s}, z, t) dz \approx \delta_{j,j_f} q_f(t) u_f(t) + \gamma_j h \sum_{\alpha=1}^G w_\alpha \mathbf{1}^t \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)),$$

with $\gamma_j = \gamma(z_j)$.

3.2 Component concentration equations

We first observe that Equations (2.4b) and (2.4c) are both linear in their respective vectors of unknowns, $\mathbf{c}$ and $\mathbf{s}$, therefore an upwind scheme is the natural choice to approximate the related vector fluxes. Furthermore, as mentioned earlier in Section 2, the approximation of (2.4b) should be designed such that the sum over each component equation equals the approximation of Equation (2.4a). This condition is required in order to ensure that the sum of the components of the discrete solution match the approximation of the total concentration of solids. To establish the higher-order accuracy, we proceed as for the total concentration of solids and make use of polynomial reconstructions on each component of the vectors $\mathbf{c}$ and $\mathbf{s}$, respectively. Then, given the vector functions $\mathbf{p}_j^c$ and $\mathbf{p}_j^s$, each one composed of n_c and n_s polynomials of degree $\ell - 1$, respectively, we denote the polynomial reconstructions of $\mathbf{c}$ and $\mathbf{s}$ respectively by $\hat{\mathbf{c}}_{j+1/2}^L$ (resp. $\hat{\mathbf{c}}_{j+1/2}^R$) and $\hat{\mathbf{s}}_{j+1/2}^L$ (resp. $\hat{\mathbf{s}}_{j+1/2}^R$). Hence, the approximation of the (convective) fluxes are given by

$$\hat{\mathbf{g}}_{j+1/2}^c := \text{Upw} \left(\hat{\mathbf{v}}_{j+1/2} - \mathcal{E}_{j+1/2}; \hat{\mathbf{c}}_{j+1/2}^L, \hat{\mathbf{c}}_{j+1/2}^R \right), \quad (3.13)$$

$$\hat{\mathbf{g}}_{j+1/2}^s := \text{Upw} \left(\rho q_{j+1/2} - \hat{\mathcal{F}}_{j+1/2}; \frac{\hat{\mathbf{s}}_{j+1/2}^L}{\rho - \hat{u}_{j+1/2}^L}, \frac{\hat{\mathbf{s}}_{j+1/2}^R}{\rho - \hat{u}_{j+1/2}^R} \right), \quad (3.14)$$

where $\hat{\mathbf{v}}_{j+1/2}$, $\mathcal{E}_{j+1/2}$ and $\hat{\mathcal{F}}_{j+1/2}$ are defined in Section 3.1 and the polynomial reconstructions at the cell boundaries are given, as in the case of u , by

$$\hat{\mathbf{c}}_{j+1/2}^L = \mathbf{p}_j^c(z_{j+1/2}), \quad \hat{\mathbf{c}}_{j+1/2}^R = \mathbf{p}_{j+1}^c(z_{j+1/2}), \quad \hat{\mathbf{s}}_{j+1/2}^L = \mathbf{p}_j^s(z_{j+1/2}) \quad \text{and} \quad \hat{\mathbf{s}}_{j+1/2}^R = \mathbf{p}_{j+1}^s(z_{j+1/2}).$$

The right-hand sides of Equations (2.4b) and (2.4c), which involve feed and reaction terms, are approximated in the same way as for the total concentration equation by

$$\begin{aligned}\int_{I_j} \mathbf{b}_c(\mathbf{c}, \mathbf{s}, z, t) dz &\approx \delta_{j,j_f} q_f(t) \mathbf{c}_f(t) + \gamma_j h \sum_{\alpha=1}^G w_\alpha \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)), \\ \int_{I_j} \mathbf{b}_s(\mathbf{c}, \mathbf{s}, z, t) dz &\approx \delta_{j,j_f} q_f(t) \mathbf{s}_f(t) + \gamma_j h \sum_{\alpha=1}^G w_\alpha \boldsymbol{\sigma}_s \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)),\end{aligned}$$

where the quadrature weights w_α and nodes z_j^α for all $\alpha = 1, \dots, G$ have been defined in Section 3.1.

3.3 Semi-discrete scheme

Collecting the spatial approximations described in Sections 3.1 and 3.2 for the equations in the reactive sedimentation model (2.4), we can write the spatially discretized (semi-discrete) scheme as the following ODE system for each $j = 0, 1, \dots, N+1$, and $0 < t \leq T$:

$$\frac{du_j}{dt}(t) = -\frac{1}{h}(\widehat{\mathcal{F}}_{j+1/2} - \widehat{\mathcal{F}}_{j-1/2}) + \frac{1}{h}\delta_{j,j_f} q_f u_f + \gamma_j \sum_{\alpha=1}^G w_\alpha \mathbf{1}^t \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)), \quad (3.15a)$$

$$\frac{d\mathbf{c}_j}{dt}(t) = -\frac{1}{h}(\widehat{\mathcal{G}}_{j+1/2}^c - \widehat{\mathcal{G}}_{j-1/2}^c) + \frac{1}{h}\delta_{j,j_f} q_f \mathbf{c}_f + \gamma_j \sum_{\alpha=1}^G w_\alpha \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)), \quad (3.15b)$$

$$\frac{d\mathbf{s}_j}{dt}(t) = -\frac{1}{h}(\widehat{\mathcal{G}}_{j+1/2}^s - \widehat{\mathcal{G}}_{j-1/2}^s) + \frac{1}{h}\delta_{j,j_f} q_f \mathbf{s}_f + \gamma_j \sum_{\alpha=1}^G w_\alpha \boldsymbol{\sigma}_s \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)), \quad (3.15c)$$

where the ghost cells $j = 0$ and $j = N+1$ are included to properly compute the fluxes in the numerical scheme. Note that in principle, as γ controls the inner and outer part of the sedimentation tank, the calculation at the ghost cells is straightforward from the scheme. This method-of-lines formulation, is going to be particularly helpful for the implementation of a higher-order time approximation method in Section 5. In addition, to simplify the notation in the upcoming sections, we introduce the vectors containing the cell values related to the stencil of the scheme of the discretized unknowns $\mathbf{u}_{\text{st},j}(t) \in \mathbb{R}^3$, $\mathbf{c}_{\text{st},j}(t) \in \mathbb{R}^{n_c \times 3}$ and $\mathbf{s}_{\text{st},j}(t) \in \mathbb{R}^{n_s \times 3}$ for $t > 0$, such that

$$\mathbf{u}_{\text{st},j}(t) := (u_{j-1}(t), u_j(t), u_{j+1}(t)), \quad (3.16)$$

$$\mathbf{c}_{\text{st},j}(t) := (\mathbf{c}_{j-1}(t), \mathbf{c}_j(t), \mathbf{c}_{j+1}(t)), \quad (3.17)$$

$$\mathbf{s}_{\text{st},j}(t) := (\mathbf{s}_{j-1}(t), \mathbf{s}_j(t), \mathbf{s}_{j+1}(t)), \quad (3.18)$$

for all $j = 0, 1, \dots, N+1$. Moreover, we define the functions corresponding to the discretized differential operators and source terms of each equation of system (3.15) respectively as

$$\begin{aligned}\mathcal{H}_j^u(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n) &:= -\frac{1}{h}(\widehat{\mathcal{F}}_{j+1/2} - \widehat{\mathcal{F}}_{j-1/2}) + \frac{1}{h}\delta_{j,j_f} q_f u_f + \gamma_j \sum_{\alpha=1}^G w_\alpha \mathbf{1}^t \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)), \\ \mathcal{H}_j^c(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n) &:= -\frac{1}{h}(\widehat{\mathcal{G}}_{j+1/2}^c - \widehat{\mathcal{G}}_{j-1/2}^c) + \frac{1}{h}\delta_{j,j_f} q_f \mathbf{c}_f + \gamma_j \sum_{\alpha=1}^G w_\alpha \boldsymbol{\sigma}_c \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)), \\ \mathcal{H}_j^s(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n) &:= -\frac{1}{h}(\widehat{\mathcal{G}}_{j+1/2}^s - \widehat{\mathcal{G}}_{j-1/2}^s) + \frac{1}{h}\delta_{j,j_f} q_f \mathbf{s}_f + \gamma_j \sum_{\alpha=1}^G w_\alpha \boldsymbol{\sigma}_s \mathbf{r}(\mathbf{p}_j^c(z_j^\alpha), \mathbf{p}_j^s(z_j^\alpha)).\end{aligned}$$

4 Polynomial reconstructions

In this section, we address the polynomial reconstructions appearing at the numerical fluxes of the space discretization in system (3.15). We describe two cases, a second-order reconstruction based on a MUSCL scheme [24, 40], and a third-order central WENO scheme (CWENO3) proposed in [29]. Moreover, for all $j = 1, \dots, N$, the vectors of polynomials $\mathbf{p}_j^c$ and $\mathbf{p}_j^s$ associated with (3.13) and (3.14), respectively, are such that

$$\mathbf{p}_j^c(z) := (p_j^{c,(1)}(z), p_j^{c,(2)}(z), \dots, p_j^{c,(n_c)}(z))^t, \quad \mathbf{p}_j^s(z) := (p_j^{s,(1)}(z), p_j^{s,(2)}(z), \dots, p_j^{s,(n_s)}(z))^t,$$

where $p_j^{c,(i)}$ is related to the $c_j^{(i)}$ concentration component for all $i \in \{1, \dots, n_c\}$, and analogously, $p_j^{s,(k)}$ is related to the $s_j^{(k)}$ substrate component for all $k \in \{1, \dots, n_s\}$. For notational convenience, the time dependence is suppressed in the polynomials above and will likewise be omitted throughout the remainder of this section.

Every polynomial $p_j^{c,(i)}$ and $p_j^{s,(k)}$ is assumed to have the same polynomial degree, and will be computed employing a similar procedure as for the polynomial $p_j^u(z)$ associated with (3.4). Therefore, in what follows, we will concentrate the description on the polynomial reconstructions corresponding to $p_j^u(z)$ and provide the key ingredients for the construction of the other polynomials.

4.1 Second-order reconstruction

The second-order accuracy, for $\ell = 2$, is obtained by employing a linear reconstruction of u over each interval I_j . To do so, we use a standard MUSCL scheme [40] combined with a minmod slope limiter so that the linear polynomial reconstruction at each j -cell is given by

$$p_j^u(z) := u_j + \sigma_j^u(z - z_j), \quad \text{for } z \in I_j, \quad (4.1)$$

where the slope σ_j^u is computed by

$$\sigma_j^u := \frac{1}{h} \text{minmod}\left(\alpha_M(u_j - u_{j-1}), \frac{1}{2}(u_{j+1} - u_{j-1}), \alpha_M(u_{j+1} - u_j)\right),$$

with parameter $\alpha_M \in (1, 2)$ and minmod function defined as

$$\text{minmod}(a, b, c) = \begin{cases} \text{sign}(a) \min\{|a|, |b|, |c|\} & \text{if } \text{sign}(a) = \text{sign}(b) = \text{sign}(c), \\ 0 & \text{otherwise,} \end{cases}$$

for all $a, b, c \in \mathbb{R}$, where $\text{sign}(a)$ returns the sign of the real number a . For the vector of solid and substrate components, we define the MUSCL polynomial reconstructions as follows

$$p_j^{c,(i)}(z) = c_j^{(i)} + \sigma_j^{c,i}(z - z_j), \quad \text{for } i = 1, \dots, n_c, \quad (4.2)$$

$$p_j^{s,(k)}(z) = s_j^{(k)} + \sigma_j^{s,k}(z - z_j), \quad \text{for } k = 1, \dots, n_s, \quad (4.3)$$

for all $z \in I_j$, where

$$\sigma_j^{c,i} := \frac{1}{h} \begin{cases} \alpha_M(c_j^{(i)} - c_{j-1}^{(i)}) & \text{if } \sigma_j^u = \alpha_M(u_j - u_{j-1}), \\ \frac{1}{2}(c_{j+1}^{(i)} - c_{j-1}^{(i)}) & \text{if } \sigma_j^u = \frac{1}{2}(u_{j+1} - u_{j-1}), \\ \alpha_M(c_{j+1}^{(i)} - c_j^{(i)}) & \text{if } \sigma_j^u = \alpha_M(u_{j+1} - u_j), \\ 0 & \text{otherwise,} \end{cases}$$

$$\sigma_j^{s,k} := \frac{1}{h} \text{minmod}\left(\alpha_M(s_j^{(k)} - s_{j-1}^{(k)}), \frac{1}{2}(s_{j+1}^{(k)} - s_{j-1}^{(k)}), \alpha_M(s_{j+1}^{(k)} - s_j^{(k)})\right).$$

From the definition of p_j^u , $\mathbf{p}_j^c$, and $\mathbf{p}_j^s$, it is straightforward to show that the average-preserving property (3.5) is fulfilled for each polynomial, that is

$$u_j = \frac{1}{h} \int_{I_j} p_j^u(z) dz, \quad \mathbf{c}_j = \frac{1}{h} \int_{I_j} \mathbf{p}_j^c(z) dz, \quad \mathbf{s}_j = \frac{1}{h} \int_{I_j} \mathbf{p}_j^s(z) dz, \quad \text{for all } j.$$

The following lemma states that the MUSCL reconstruction satisfies an invariant-region property.

Lemma 4.1. *Consider the polynomials p_j^u , $\mathbf{p}_j^c$ and $\mathbf{p}_j^s$ defined in (4.1), (4.2) and (4.3), respectively. Assume that $(u_j, \mathbf{c}_j, \mathbf{s}_j) \in \mathcal{W}$. Then, there holds*

$$(p_j^u(z), \mathbf{p}_j^c(z), \mathbf{p}_j^s(z)) \in \mathcal{W}, \quad \text{for all } z \in I_j.$$

Proof. We first assume that $\sigma_j^u > 0$, from which we directly have

$$\sigma_j^u \leq \frac{1}{h} \alpha_M (u_j - u_{j-1}), \quad \sigma_j^u \leq \frac{1}{2h} (u_{j+1} - u_{j-1}), \quad \sigma_j^u \leq \frac{1}{h} \alpha_M (u_{j+1} - u_j).$$

Then, we can conclude that $0 \leq u_{j-1} < u_j < u_{j+1} \leq u_{\max}$. Now, since p_j^u is a linear and increasing polynomial with respect to z , we only need to show the bounds at the limits $z_{j\pm 1/2}$, that is

$$\begin{aligned} p_j^u(z_{j-1/2}) &= u_j - \frac{h}{2} \sigma_j^u \geq (1 - \tilde{\alpha}_M) u_j + \tilde{\alpha}_M u_{j-1} \geq 0, \\ p_j^u(z_{j+1/2}) &= u_j + \frac{h}{2} \sigma_j^u \leq (1 - \tilde{\alpha}_M) u_j + \tilde{\alpha}_M u_{j+1} \leq u_{\max}, \end{aligned}$$

where $\tilde{\alpha}_M = \alpha_M/2$. The case $\sigma_j^u < 0$ is analogous, and $\sigma_j^u = 0$ is trivial. The bounds $p_j^{c,i} \geq 0$ and $p_j^{s,k} \geq 0$ can be shown in the same way as for p_j^u for all $i = 1, \dots, n_c$ and $k = 1, \dots, n_s$, and thanks to the definition of the slope $\sigma_j^{c,i}$, we have

$$\mathbf{1}^t \mathbf{p}_j^c(z) = \sum_{i=1}^{n_c} c_j^{(i)} + \sigma_j^{c,i} (z - z_j) = u_j + \sum_{i=1}^{n_c} \sigma_j^{c,i} (z - z_j) = u_j + \sigma_j^u (z - z_j) = p_j^u(z),$$

which concludes the proof. $\square$

4.2 Third-order reconstruction

To achieve third-order accuracy, we set $\ell = 3$ and consider the compact central WENO scheme introduced in [29], in which the polynomial reconstruction combines linear and quadratic polynomials. We will refer to this third-order scheme as CWENO3. For ease of presentation, we drop the sub-index j when describing the linear and quadratic polynomials involved in this reconstruction. In turn, we define the left and right linear interpolation polynomials as

$$p_L^u(z) := u_{j-1} + \frac{u_j - u_{j-1}}{h} (z - z_{j-1}), \quad p_R^u(z) := u_j + \frac{u_{j+1} - u_j}{h} (z - z_j), \quad (4.4)$$

respectively, and the so-called *central polynomial* denoted by p_C^u , is defined as a quadratic polynomial computed with the values of u_{j+1} , u_j and u_{j-1} , see Figure 4.2. The polynomial reconstruction of u at the cell I_j is then determined by the combination of these three polynomials in the following way

$$p_j^u(z) := w_L p_L^u(z) + w_C p_C^u(z) + w_R p_R^u(z), \quad (4.5)$$

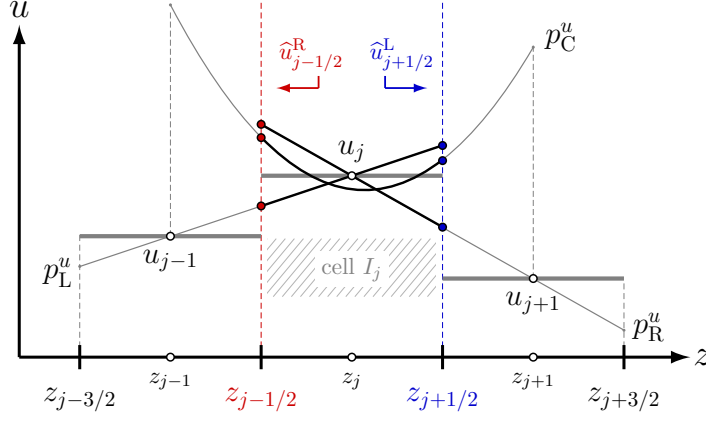


Figure 4.2: Illustration of the polynomials p_L^u , p_R^u and p_C^u related to the reconstruction p_j^u defined in (4.5) at the cell $I_j = [z_{j-1/2}, z_{j+1/2}]$ for the CWENO3 scheme from Section 4.2. The coloured dots are used in the construction of $\hat{u}_{j-1/2}^R$ (red) and $\hat{u}_{j+1/2}^L$ (blue).

where w_L , w_C and w_R are positive weights that add up to one. The central polynomial p_C^u depends on p_L^u and p_R^u , and the quadratic optimal polynomial:

$$p_{\text{opt}}^u(z) := u_j - \frac{u_{j+1} - 2u_j + u_{j-1}}{24} + \frac{u_{j+1} - u_{j-1}}{2h}(z - z_j) + \frac{u_{j+1} - 2u_j + u_{j-1}}{2h^2}(z - z_j)^2,$$

such that

$$\begin{aligned} p_C^u(z) &:= 2p_{\text{opt}}^u(z) - \frac{1}{2}(p_L^u(z) + p_R^u(z)) \\ &= u_j - \frac{u_{j+1} - 2u_j + u_{j-1}}{12} + \frac{u_{j+1} - u_{j-1}}{2h}(z - z_j) + \frac{u_{j+1} - 2u_j + u_{j-1}}{h^2}(z - z_j)^2. \end{aligned}$$

The above formula is given for the specific choice of constants $C_L = C_R = 1/4$ and $C_C = 1/2$, which appear in the following definition of the weights:

$$w_K := \frac{\alpha_K}{\alpha_L + \alpha_C + \alpha_R}, \quad \alpha_K = \frac{C_K}{(\varepsilon + \mathcal{S}_K^u)^2} \quad \text{for } K \in \{L, C, R\}, \quad (4.6)$$

where the indicators of smoothness of p_L^u , p_R^u and p_C^u are respectively given by

$$\mathcal{S}_L^u := (u_j - u_{j-1})^2, \quad (4.7a)$$

$$\mathcal{S}_R^u := (u_{j+1} - u_j)^2, \quad (4.7b)$$

$$\mathcal{S}_C^u := \frac{13}{12}(u_{j+1} - 2u_j + u_{j-1})^2 + \frac{1}{4}(u_{j+1} - u_{j-1})^2. \quad (4.7c)$$

The constant $\varepsilon > 0$ in the definition of the weights (4.6) is taken to prevent the denominator from vanishing. In addition, the values of ε should be such that $\varepsilon + h^2 u_x^2 \gg |\mathcal{S}_K^u - h^2 u_x^2|$ in smooth regions, and that $\varepsilon \ll \mathcal{S}_K^u$ near discontinuities for $K \in \{L, C, R\}$.

To highlight the functional dependencies in the calculation of the polynomial reconstructions, we introduce $\mathcal{P}_j$ as the function that given u_{j-1} , u_j and u_{j+1} defines the quadratic polynomial (4.5) described above, that is

$$p_j^u(z) =: \mathcal{P}_j\left(z, u_{j-1}, u_j, u_{j+1}; \mathcal{S}_L^u, \mathcal{S}_C^u, \mathcal{S}_R^u\right). \quad (4.8)$$

Then, we can directly introduce the polynomial reconstructions for the vectors of solid concentrations and substrates as follows

$$p_j^{c,(i)} = \mathcal{P}_j\left(z, c_{j-1}^{(i)}, c_j^{(i)}, c_{j+1}^{(i)}; \mathcal{S}_L^u, \mathcal{S}_C^u, \mathcal{S}_R^u\right), \quad \text{for } i = 1, \dots, n_c, \quad (4.9)$$

$$p_j^{s,(k)} = \mathcal{P}_j\left(z, s_{j-1}^{(k)}, s_j^{(k)}, s_{j+1}^{(k)}; \mathcal{S}_L^{s,k}, \mathcal{S}_C^{s,k}, \mathcal{S}_R^{s,k}\right), \quad \text{for } k = 1, \dots, n_s, \quad (4.10)$$

with $\mathcal{S}_L^{s,k}$, $\mathcal{S}_C^{s,k}$ and $\mathcal{S}_R^{s,k}$ defined as in (4.7) by

$$\begin{aligned} \mathcal{S}_L^{s,k} &:= (s_j^{(k)} - s_{j-1}^{(k)})^2, \\ \mathcal{S}_R^{s,k} &:= (s_{j+1}^{(k)} - s_j^{(k)})^2, \\ \mathcal{S}_C^{s,k} &:= \frac{13}{12}(s_{j+1}^{(k)} - 2s_j^{(k)} + s_{j-1}^{(k)})^2 + \frac{1}{4}(s_{j+1}^{(k)} - s_{j-1}^{(k)})^2, \end{aligned}$$

where we have also omitted index j as in the previous description for the total concentration. Since by construction the left, right and central interpolation polynomials are linear in terms of the discretized values of the unknowns, and since for $\mathbf{p}_j^c$ we have evaluated $\mathcal{S}_L$, $\mathcal{S}_C$, and $\mathcal{S}_R$ on terms depending on the total concentration, we can show that

$$p_j^u = \sum_{i=1}^{n_c} p_j^{c,(i)} = \mathbf{1}^\top \mathbf{p}_j^c. \quad (4.11)$$

4.3 Positivity preserving scaling limiters

It is well known that the reconstruction procedure described above leads to schemes that resolve discontinuities sharply but in some cases fail to preserve the positivity of solutions. This shortcoming motivated the development of the linear scaling limiter by Zhang and Shu [44, 46]. We begin describing this limiter for the polynomial reconstructions of u_j and $\mathbf{c}_j$, respectively. To implement this limiter, we replace the polynomials in (4.8) and (4.9) by

$$\tilde{p}_j^u(z) := \theta_j^{\text{solid}}(p_j^u(z) - u_j) + u_j, \quad (4.12)$$

$$\tilde{p}_j^{c,(i)}(z) := \theta_j^{\text{solid}}(p_j^{c,(i)}(z) - c_j^{(i)}) + c_j^{(i)}, \quad \text{for } i = 1, \dots, n_c, \quad (4.13)$$

where θ_j^{solid} is defined such that $0 \leq \tilde{p}_j^u \leq u_{\max}$, and in addition, due to the connection between vector $\mathbf{c}$ and u , we impose the stronger condition that $p_j^{c,(i)} \geq 0$ for all i in the definition of θ_j^{solid} , which thanks to Equation (4.11) also implies $p_j^u \geq 0$. Then, we set

$$\theta_j^{\text{solid}} = \min \left\{ 1, \frac{u_{\max} - u_j}{M_j^u - u_j}, \min_{1 \leq i \leq n_c} \frac{c_j^{(i)}}{c_j^{(i)} - m_j^{c,i}} \right\}, \quad (4.14)$$

with

$$m_j^{c,i} := \min_{z \in I_j} p_j^{c,(i)}(z), \quad M_j^u := \max_{z \in I_j} p_j^u(z). \quad (4.15)$$

Then, using the modified polynomial including the limiter, the reconstructed variables in (3.4) are now computed as $\hat{u}_{j+1/2}^L := \tilde{p}_j^u(z_{j+1/2})$ and $\hat{u}_{j+1/2}^R := \tilde{p}_{j+1}^u(z_{j+1/2})$, and analogously for the reconstruction of vector $\mathbf{c}_j$. Observe that from the definition of θ_j^{solid} and Equation (4.11), we can show that

$$\mathbf{1}^\top \tilde{\mathbf{p}}_j^c(z) = \sum_{i=1}^{n_c} p_j^{c,(i)}(z) = \sum_{i=1}^{n_c} \left(\theta_j^{\text{solid}}(p_j^{c,(i)}(z) - c_j^{(i)}) + c_j^{(i)} \right) = \theta_j^{\text{solid}} \sum_{i=1}^{n_c} (p_j^{c,(i)}(z) - c_j^{(i)}) + u_j = \tilde{p}_j^u(z). \quad (4.16)$$

For the polynomial reconstruction related to the vector of substrates, unlike the case of the solids related variables, we only require non-negativity of each component, therefore we define

$$\tilde{p}_j^{\mathbf{s},(k)}(z) := \theta_{k,j}^{\text{fluid}} (p_j^{\mathbf{s},(k)}(z) - s_j^{(k)}) + s_j^{(k)}, \quad \text{for all } k = 1, \dots, n_{\mathbf{s}}, \quad (4.17)$$

where

$$\theta_{k,j}^{\text{fluid}} = \min \left\{ 1, \frac{s_j^{(k)}}{s_j^{(k)} - m_j^{\mathbf{s},k}} \right\}, \quad \text{with } m_j^{\mathbf{s},k} = \min_{z \in I_j} p_j^{\mathbf{s},(k)}(z). \quad (4.18)$$

Next, we recall some useful inequalities for the minimum and maximum values related to the polynomial reconstructions involved in the definition of the slopes in (4.14) and (4.18), respectively. The polynomial reconstruction of the total solids concentration satisfies

$$\min_{z \in I_j} p_j^u(z) =: m_j^u \leq u_j = \frac{1}{h} \int_{I_j} p_j^u(z) dz \leq M_j^u, \quad (4.19)$$

Furthermore, for the minimum values of the reconstructions for the vectors of solids and substrates components, we have

$$m_j^{\mathbf{c},i} \leq c_j^{(i)} \leq u_j \leq M_j^u, \quad m_j^{\mathbf{c},i} \leq p_j^{\mathbf{c},(i)}, \quad \text{for } i = 1, \dots, n_{\mathbf{c}}, \quad (4.20)$$

$$m_j^{\mathbf{s},k} \leq s_j^{(k)}, \quad m_j^{\mathbf{s},k} \leq p_j^{\mathbf{s},(k)}, \quad \text{for } k = 1, \dots, n_{\mathbf{s}}, \quad (4.21)$$

where from the second equation in (4.20), we conclude that

$$\sum_{i=1}^{n_{\mathbf{c}}} m_j^{\mathbf{c},i} \leq \mathbf{1}^{\mathbf{t}} p_j^{\mathbf{c}} = p_j^u. \quad (4.22)$$

The above inequalities allow us to state that $0 \leq \theta_j^{\text{solid}}, \theta_{j,k}^{\text{fluid}} \leq 1$ for all k , which implies that (4.12), (4.13) and (4.17) are in fact, convex combinations between the averaged values and the reconstructed polynomials. The next lemma shows that the limiters preserve the invariant-region property.

Lemma 4.2. *Consider the polynomial reconstructions $\tilde{p}_j^u$, $\tilde{p}_j^{\mathbf{c}}$ and $\tilde{p}_j^{\mathbf{s}}$ for the third-order CWENO3 scheme defined in (4.12), (4.13) and (4.17), respectively. Assume that $(u_j, \mathbf{c}_j, \mathbf{s}_j) \in \mathcal{W}$ for all j . Then, there holds*

$$(\tilde{p}_j^u(z), \tilde{p}_j^{\mathbf{c}}(z), \tilde{p}_j^{\mathbf{s}}(z)) \in \mathcal{W}, \quad \text{for all } z \in I_j.$$

Proof. Given $z \in I_j$, if $p_j^u(z) - u_j < 0$, then $p_j^u(z) < u_j \leq u_{\max}$ and therefore $\tilde{p}_j^u(z) \leq u_{\max}$, as $\tilde{p}_j^u$ is a convex combination between u_j and p_j^u . In addition, if $p_j^u(z) - u_j \geq 0$, we have

$$\tilde{p}_j^u(z) \leq \frac{u_{\max} - u_j}{M_j^u - u_j} (p_j^u(z) - u_j) + u_j \leq u_{\max}.$$

For the lower bounds, we observe that for all $1 \leq i \leq n_{\mathbf{c}}$, $\theta_j^{\text{solid}} \leq c_j^{(i)} / (c_j^{(i)} - m_j^{\mathbf{c},i})$, which is equivalent to $\theta_j^{\text{solid}} (c_j^{(i)} - m_j^{\mathbf{c},i}) \leq c_j^{(i)}$ as $c_j^{(i)} - m_j^{\mathbf{c},i} \geq 0$ thanks to (4.20). Then, adding over i and using (4.22), we get

$$0 \leq (1 - \theta_j^{\text{solid}})u_j + \theta_j^{\text{solid}} \sum_{i=1}^{n_{\mathbf{c}}} m_j^{\mathbf{c},i} \leq (1 - \theta_j^{\text{solid}})u_j + \theta_j^{\text{solid}} p_j^u = \tilde{p}_j^u.$$

Similarly, for the vector of solids and substrates we have

$$\begin{aligned} 0 &\leq (1 - \theta_j^{\text{solid}})c_j^{(i)} + \theta_j^{\text{solid}}m_j^{\mathbf{c},i} \leq (1 - \theta_j^{\text{solid}})c_j^{(i)} + \theta_j^{\text{solid}}p_j^{\mathbf{c},(i)} = \tilde{p}_j^{\mathbf{c},(i)}, \quad \text{for } i = 1, \dots, n_{\mathbf{c}}, \\ 0 &\leq (1 - \theta_{k,j}^{\text{fluid}})s_j^{(k)} + \theta_{k,j}^{\text{fluid}}m_j^{\mathbf{s},k} \leq (1 - \theta_{k,j}^{\text{fluid}})s_j^{(k)} + \theta_{k,j}^{\text{fluid}}p_j^{\mathbf{s},(k)} = \tilde{p}_j^{\mathbf{s},(k)}, \quad \text{for } k = 1, \dots, n_{\mathbf{s}}, \end{aligned}$$

and finally, $\tilde{p}_j^{\mathbf{c},(1)} + \tilde{p}_j^{\mathbf{c},(2)} + \dots + \tilde{p}_j^{\mathbf{c},(n_{\mathbf{c}})} = \tilde{p}_j^u$ is fulfilled thanks to Equation (4.16), concluding in this way the proof. $\square$

The minimum and maximum values $m_j^{\mathbf{c},i}$, $m_j^{\mathbf{s},k}$ and M_j^u in (4.14) and (4.18) require evaluating the extrema of polynomials on each cell. This inconvenience can be avoided if we define simplified limiters (as in [44]) that reduce the evaluations of each polynomial to a finite number of nodes of a G -point Legendre-Gauss-Lobatto quadrature rule on the interval $I_j = [z_{j-1/2}, z_{j+1/2}]$. This formula is exact for the integral of polynomials of degree up to $2G - 3$, such as the quadrature rule described in Section 3.2, and the main gain is that the local extrema can be computed using the polynomials at the limits $z_{j\pm 1/2}$, average on I_j , and weights of the quadrature rule. To explain the derivation of this compact formula, we focus on p_j^u , and compute its minimum, denoted by m_j^u , and its maximum M_j^u . We consider the weights w_α and nodes z_j^α for $\alpha = 1, \dots, G$ introduced in Section 3.2 to approximate the reaction terms, and further define $\nu_\alpha := p_j^u(z_j^\alpha)$ with which we have

$$u_j = \frac{1}{h} \int_{I_j} p_j^u(z) dz = \sum_{\alpha=1}^G w_\alpha \nu_\alpha = w_1 \nu_1 + w_G \nu_G + \sum_{\alpha=2}^{G-1} w_\alpha \nu_\alpha, \quad (4.23)$$

where thanks to the mean value theorem, there exists $\xi \in I_j$ such that the last sum above can be written as

$$\sum_{\alpha=2}^{G-1} w_\alpha \nu_\alpha = (1 - w_1 - w_G) \sum_{\alpha=2}^{G-1} \frac{w_\alpha}{1 - w_1 - w_G} p_j^u(z_j^\alpha) = (1 - w_1 - w_G) p_j^u(\xi), \quad (4.24)$$

where we have used the fact the coefficients in the middle sum are such that

$$\sum_{\alpha=2}^{G-1} \frac{w_\alpha}{1 - w_1 - w_G} = 1.$$

Then, replacing (4.24) into (4.23), we have

$$u_j = w_1 p_j^u(z_{j-1/2}) + w_G p_j^u(z_{j+1/2}) + (1 - w_1 - w_G) p_j^u(\xi),$$

from which we can determine the value of $p_j^u(\xi)$ as follows:

$$p_j^u(\xi) = \frac{1}{1 - w_1 - w_G} \left(u_j - w_1 p_j^u(z_{j-1/2}) - w_G p_j^u(z_{j+1/2}) \right) =: \eta_j^u.$$

Note that the formula above can be computed even though ξ is a priori unknown. Thus the minimum and maximum values of $p_j^u(z)$ over I_j can be computed through the compact formulas:

$$m_j^u = \min \{ p_j^u(z_{j-1/2}), \eta_j^u, p_j^u(z_{j+1/2}) \}, \quad M_j^u = \max \{ p_j^u(z_{j-1/2}), \eta_j^u, p_j^u(z_{j+1/2}) \}. \quad (4.25)$$

In particular, for a third-order reconstruction we employ a three-point ($G = 3$) quadrature rule with weights $w_1 = 1/6$, $w_2 = 2/3$, and $w_3 = 1/6$. In the remainder of the paper, unless otherwise specified, we will only consider the modified polynomial reconstructions including the limiters in (4.12), (4.13) and (4.17), for which we will drop the tilde, and simply denote them as p_j^u , $\mathbf{p}_j^{\mathbf{c}}$ and $\mathbf{p}_j^{\mathbf{s}}$ instead.

5 Time approximation

As pointed out in [38], a suitable time discretization for higher order approximations is a ℓ -th total variation diminishing (TVD) Runge-Kutta method, also known as strong stability preserving Runge-Kutta (SSPRK) methods [18, 19]. Due to convexity in the intermediate stages, in particular for the second and third-order SSPRK methods, these time discretizations preserve the positivity of numerical solutions. To begin describing the time approximation, we consider a time step $\Delta t > 0$, discrete time points $t^n := n\Delta t$ with $n \in \mathbb{N}$ and each time dependent variable and unknown evaluated at t^n will be denoted with the superscript n , for example $u_j(t^n) = u_j^n$, etc. We employ second- and third-order SSPRK schemes for the MUSCL and CWENO3 spatial reconstructions, respectively.

5.1 Second order time approximation

The second-order explicit SSPRK method is given by the following Butcher tableau:

$$\begin{array}{c|c} 0 & \\ 1 & 1 \\ \hline & 1/2 \quad 1/2 \end{array}$$

Then, we sequentially solve the three equations in (3.15) in two stages.

- **Stage 1:** Solve for the first intermediate variables $\mathbf{u}_{\text{st},j}^{(1)}$, $\mathbf{c}_{\text{st},j}^{(1)}$ and $\mathbf{s}_{\text{st},j}^{(1)}$ for all $j = 1, \dots, N$:

$$\begin{aligned} u_j^{(1)} &= u_j^n + \Delta t \mathcal{H}_j^u(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n), \\ \mathbf{c}_j^{(1)} &= \mathbf{c}_j^n + \Delta t \mathcal{H}_j^c(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n), \\ \mathbf{s}_j^{(1)} &= \mathbf{s}_j^n + \Delta t \mathcal{H}_j^s(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n). \end{aligned}$$

- **Stage 2:** Update the solution at time t^{n+1} for all $j = 1, \dots, N$:

$$\begin{aligned} u_j^{n+1} &= \frac{1}{2}u_j^n + \frac{1}{2}u_j^{(1)} + \frac{1}{2}\Delta t \mathcal{H}_j^u(\mathbf{u}_{\text{st},j}^{(1)}, \mathbf{c}_{\text{st},j}^{(1)}, \mathbf{s}_{\text{st},j}^{(1)}), \\ \mathbf{c}_j^{n+1} &= \frac{1}{2}\mathbf{c}_j^n + \frac{1}{2}\mathbf{c}_j^{(1)} + \frac{1}{2}\Delta t \mathcal{H}_j^c(\mathbf{u}_{\text{st},j}^{(1)}, \mathbf{c}_{\text{st},j}^{(1)}, \mathbf{s}_{\text{st},j}^{(1)}), \\ \mathbf{s}_j^{n+1} &= \frac{1}{2}\mathbf{s}_j^n + \frac{1}{2}\mathbf{s}_j^{(1)} + \frac{1}{2}\Delta t \mathcal{H}_j^s(\mathbf{u}_{\text{st},j}^{(1)}, \mathbf{c}_{\text{st},j}^{(1)}, \mathbf{s}_{\text{st},j}^{(1)}). \end{aligned}$$

5.2 Third order time approximation

The third-order explicit SSPRK method is given by the following Butcher tableau:

$$\begin{array}{c|ccc} 0 & & & \\ 1 & 1 & & \\ 1/2 & 1/4 & 1/4 & \\ \hline & 1/6 & 1/6 & 2/3 \end{array}$$

Then, we sequentially solve the three equations in (3.15) in three stages.

- **Stage 1:** Solve for the first intermediate variables $\mathbf{u}_{\text{st},j}^{(1)}$, $\mathbf{c}_{\text{st},j}^{(1)}$ and $\mathbf{s}_{\text{st},j}^{(1)}$ for all $j = 1, \dots, N$:

$$\begin{aligned} u_j^{(1)} &= u_j^n + \Delta t \mathcal{H}_j^u(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n), \\ \mathbf{c}_j^{(1)} &= \mathbf{c}_j^n + \Delta t \mathcal{H}_j^c(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n), \\ \mathbf{s}_j^{(1)} &= \mathbf{s}_j^n + \Delta t \mathcal{H}_j^s(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n). \end{aligned}$$

- **Stage 2:** Solve for the second intermediate variables $\mathbf{u}_{\text{st},j}^{(2)}$, $\mathbf{c}_{\text{st},j}^{(2)}$ and $\mathbf{s}_{\text{st},j}^{(2)}$ for all $j = 1, \dots, N$:

$$\begin{aligned} u_j^{(2)} &= \frac{3}{4}u_j^n + \frac{1}{4}u_j^{(1)} + \frac{1}{4}\Delta t \mathcal{H}_j^u(\mathbf{u}_{\text{st},j}^{(1)}, \mathbf{c}_{\text{st},j}^{(1)}, \mathbf{s}_{\text{st},j}^{(1)}), \\ \mathbf{c}_j^{(2)} &= \frac{3}{4}\mathbf{c}_j^n + \frac{1}{4}\mathbf{c}_j^{(1)} + \frac{1}{4}\Delta t \mathcal{H}_j^c(\mathbf{u}_{\text{st},j}^{(1)}, \mathbf{c}_{\text{st},j}^{(1)}, \mathbf{s}_{\text{st},j}^{(1)}), \\ \mathbf{s}_j^{(2)} &= \frac{3}{4}\mathbf{s}_j^n + \frac{1}{4}\mathbf{s}_j^{(1)} + \frac{1}{4}\Delta t \mathcal{H}_j^s(\mathbf{u}_{\text{st},j}^{(1)}, \mathbf{c}_{\text{st},j}^{(1)}, \mathbf{s}_{\text{st},j}^{(1)}). \end{aligned}$$

- **Stage 3:** Update the solution at time t^{n+1} for all $j = 1, \dots, N$:

$$\begin{aligned} u_j^{n+1} &= \frac{1}{3}u_j^n + \frac{2}{3}u_j^{(1)} + \frac{2}{3}\Delta t \mathcal{H}_j^u(\mathbf{u}_{\text{st},j}^{(2)}, \mathbf{c}_{\text{st},j}^{(2)}, \mathbf{s}_{\text{st},j}^{(2)}), \\ \mathbf{c}_j^{n+1} &= \frac{1}{3}\mathbf{c}_j^n + \frac{2}{3}\mathbf{c}_j^{(1)} + \frac{2}{3}\Delta t \mathcal{H}_j^c(\mathbf{u}_{\text{st},j}^{(2)}, \mathbf{c}_{\text{st},j}^{(2)}, \mathbf{s}_{\text{st},j}^{(2)}), \\ \mathbf{s}_j^{n+1} &= \frac{1}{3}\mathbf{s}_j^n + \frac{2}{3}\mathbf{s}_j^{(1)} + \frac{2}{3}\Delta t \mathcal{H}_j^s(\mathbf{u}_{\text{st},j}^{(2)}, \mathbf{c}_{\text{st},j}^{(2)}, \mathbf{s}_{\text{st},j}^{(2)}). \end{aligned}$$

6 Invariant region property

In this section we are interested in showing that the spatially high-order scheme (3.15) with MUSCL and CWENO3 reconstructions, for the case without diffusion, namely $\mathcal{D} = \mathcal{B} = 0$, combined with a suitable time discretization as described in Section 5, satisfies an invariant-region-preserving property. For the invariant-region property, one first assumes that at a given time t^n , the solution $(u_j^n, \mathbf{c}_j^n, \mathbf{s}_j^n)$ belongs to $\mathcal{W}$, and then the aim is to show that $(u_j^{n+1}, \mathbf{c}_j^{n+1}, \mathbf{s}_j^{n+1})$ also belongs to $\mathcal{W}$. The standard approach followed in the literature [1, 2, 46] is to establish an invariant-region property for the first order scheme using forward Euler approximation in time, and then using this result, state the invariant-region property for the high-order scheme.

Before studying the first-order scheme, we collect some useful bounds related to the reaction terms. Given $(u_j^n, \mathbf{c}_j^n, \mathbf{s}_j^n) \in \mathcal{W}$, the following inequality can be obtained as a direct consequence of Assumption 2.1:

$$\mathbf{1}^\top \boldsymbol{\sigma}_c \mathbf{r}_j^n = \sum_{k=1}^{n_c} \sum_{i=1}^{n_r} \sigma_c^{(k,i)} r_j^{(i),n} \geq - \sum_{k=1}^{n_c} \sum_{i \in \mathcal{J}_{c,k}^-} (-\sigma_c^{(k,i)}) \bar{r}_{c,j}^{(i)} c_j^{(k),n} \geq -\Lambda_{c,1} R u_j^n, \quad (6.1)$$

$$(\boldsymbol{\sigma}_s \mathbf{r}_j^n)^{(k)} = \sum_{i=1}^{n_r} \sigma_s^{(k,i)} r_j^{(i),n} \geq - \sum_{i \in \mathcal{J}_{s,k}^-} (-\sigma_s^{(k,i)}) \bar{r}_{s,j}^{(i)} s_j^{(k),n} \geq -\Lambda_{s,1} R s_j^{(k),n}, \quad (6.2)$$

where $\Lambda_{\boldsymbol{\xi},1} \geq 0$ is defined as

$$\Lambda_{\boldsymbol{\xi},1} := \max_k \sum_{i \in \mathcal{J}_{\boldsymbol{\xi},k}^-} (-\sigma_{\boldsymbol{\xi}}^{(k,i)}) \quad \text{for } \boldsymbol{\xi} \in \{\mathbf{c}, \mathbf{s}\}.$$

For an upper bound of the reaction term for the solid components, we let $L_R > 0$ be the constant related to the Lipschitz continuity of each function $r^{(i)}$ for $i = 1, \dots, n_r$, see Assumption 2.1. Then, using the Lipschitz continuity condition of $r^{(i)}$ and the fact that $r^{(i)}(\mathbf{c}, \mathbf{s}) = \mathbf{0}$ for $\mathbf{c}$ vectors such that $\mathbf{1}^\top \mathbf{c} = u_{\max}$, and defining $\bar{\mathbf{c}} = \mathbf{c}_j^n + \frac{1}{n_c}(u_{\max} - u_j^n)\mathbf{1}$, we get

$$|r^{(i)}(\bar{\mathbf{c}}, \mathbf{s}_j^n) - r_j^{(i),n}| = r_j^{(i),n} \leq L_R \sum_{i=1}^{n_c} |c_j^{(i),n} + \frac{1}{n_c}(u_{\max} - u_j^n) - c_j^{(i),n}| = L_R(u_{\max} - u_j^n),$$

which can be used to obtain the following bound

$$\mathbf{1}^\top \boldsymbol{\sigma}_c \mathbf{r}_j^n \leq \sum_{k=1}^{n_c} \sum_{i \in \mathcal{J}_{c,k}^+} \sigma_c^{(k,i)} r_j^{(i),n} \leq L_R(u_{\max} - u_j^n) \sum_{k=1}^{n_c} \sum_{i \in \mathcal{J}_{c,k}^+} \sigma_c^{(k,i)} \leq L_R \Lambda_{c,2}(u_{\max} - u_j^n), \quad (6.3)$$

where $\Lambda_{c,2} \geq 0$ is defined as

$$\Lambda_{c,2} := \sum_{k=1}^{n_c} \sum_{i \in \mathcal{J}_{c,k}^+} \sigma_c^{(k,i)} \geq 0.$$

In what follows, we are going to consider the Courant–Friedrichs–Lewy (CFL) condition

$$\Delta t \leq \left(\frac{1}{h} \left(\frac{\rho \|q_f\|_\infty}{\rho - u_{\max}} + \max \{v_0, L_v\} \right) + \max \{ \Lambda_{c,1} R, \Lambda_{s,1} R, \Lambda_{c,2} L_R \} \right)^{-1}, \quad (6.4)$$

where we have that the ratio $\lambda := \Delta t/h$ is bounded. Observe that to properly cover the diffusive case, with $\mathcal{D} \neq 0$, an additional term of the type $\|d\|_\infty/h^2$ must be included inside the parentheses on the right-hand side of (6.4).

To show the invariant-region property of the first order scheme [4], we take advantage of the structure of the PDE system (2.4), which can be solved sequentially. Therefore, we first establish the bound $0 \leq u_j^{n+1} \leq u_{\max}$, then $\mathbf{c}_j^{n+1} \geq \mathbf{0}$ and $\mathbf{1}^\top \mathbf{c}_j^{n+1} = u_j^{n+1}$, and lastly $\mathbf{s}_j^{n+1}$ respectively in the three lemmas that follow.

Lemma 6.1. *Assume $(u_j^n, \mathbf{c}_j^n, \mathbf{s}_j^n) \in \mathcal{W}$ for all $j = 1, \dots, N$, and consider the following three points first-order scheme for the total solids concentration*

$$u_j^{n+1} = u_j^n - \lambda(\mathcal{F}_{j+1/2}^n - \mathcal{F}_{j-1/2}^n) + \lambda \delta_{j,j_f} q_{f_f}^n u_f^n + \Delta t \mathbf{1}^\top \boldsymbol{\sigma}_c \mathbf{r}_j^n =: \Phi_j^u(u_{j-1}^n, u_j^n, u_{j+1}^n), \quad (6.5)$$

where

$$\mathcal{F}_{j+1/2}^n = \mathcal{V}_{j+1/2}^{n,+} u_j^n + \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n, \quad \text{and} \quad \mathcal{V}_{j+1/2}^n = q_{j+1/2}^n + \gamma_{j+1/2} v_{hs}(u_{j+1}^n). \quad (6.6)$$

Then, under CFL condition (6.4), there holds $0 \leq u_j^{n+1} \leq u_{\max}$ for all $j = 1, \dots, N$.

Proof. For the lower bound, we use the definition of $\mathcal{V}_{j\pm 1/2}^n$ in (6.6), including the definition of $q_{j\pm 1/2}^n$ in (3.1) and the fact that $0 \leq v_{hs}(u) \leq v_0$ for all $0 \leq u \leq u_{\max}$, the lower bound (6.1) and CFL condition (6.4) to obtain

$$\begin{aligned} u_j^{n+1} &= u_j^n - \lambda \mathcal{V}_{j+1/2}^{n,+} u_j^n + \lambda \mathcal{V}_{j-1/2}^{n,-} u_j^n - \lambda \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n + \lambda \mathcal{V}_{j-1/2}^{n,+} u_{j-1}^n + \lambda \delta_{j,j_f} q_{f_f}^n u_f^n + \Delta t \mathbf{1}^\top \boldsymbol{\sigma}_c \mathbf{r}_j^n \\ &\geq u_j^n - \lambda (\max \{q_u(t)\} + v_0 + \max \{q_e(t)\}) u_j^n - \Delta t \Lambda_c R u_j^n + \lambda \delta_{j,j_f} q_{f_f}^n u_f^n \\ &\geq (1 - \lambda (\|q_f\|_\infty + v_0) - \Delta t \Lambda_c R) u_j^n \geq 0. \end{aligned}$$

For the upper bound, we proceed as follows. First, we subtract u_j^{n+1} to $u_{\max}$ such that we bound the following equation instead

$$\begin{aligned} u_{\max} - u_j^{n+1} &= u_{\max} - u_j^n + \lambda(\mathcal{V}_{j+1/2}^{n,+} u_j^n - \mathcal{V}_{j-1/2}^{n,-} u_j^n + \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n - \mathcal{V}_{j-1/2}^{n,+} u_{j-1}^n) \\ &\quad - \lambda \delta_{j,j_f} q_f^n u_f^n - \Delta t \mathbf{1}^t \sigma_c \mathbf{r}_j^n. \end{aligned}$$

Then, the convective terms can be lower-bounded by

$$\begin{aligned} &\mathcal{V}_{j+1/2}^{n,+} u_j^n - \mathcal{V}_{j-1/2}^{n,-} u_j^n + \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n - \mathcal{V}_{j-1/2}^{n,+} u_{j-1}^n - \delta_{j,j_f} q_f^n u_f^n \\ &\geq (q_{j+1/2}^{n,+} - q_{j-1/2}^{n,-}) u_j^n - (q_{j-1/2}^{n,+} - q_{j+1/2}^{n,-} + \delta_{j,j_f} q_f^n) u_{\max} - v_{\text{hs}}(u_j^n) u_{\max} \end{aligned}$$

where we have used that $u_f^n \leq u_{\max}$, and the following estimates

$$\begin{aligned} \mathcal{V}_{j+1/2}^{n,+} u_j^n &\geq q_{j+1/2}^{n,+} u_j^n, \quad -\mathcal{V}_{j-1/2}^{n,-} u_j^n \geq -q_{j-1/2}^{n,-} u_j^n, \quad \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n \geq q_{j+1/2}^{n,-} u_{\max}, \\ -\mathcal{V}_{j-1/2}^{n,+} u_{j-1}^n &\geq -q_{j-1/2}^{n,+} u_{\max} - (\gamma_{j-1/2} v_{\text{hs}}(u_j^n))^+ u_{\max}. \end{aligned}$$

The term related to v_{hs} can be bounded using its Lipschitz continuity at $u = u_{\max}$ in (2.2), i.e., $v_{\text{hs}}(u_j^n) \leq L_v(u_{\max} - u_j^n)$, together with $u_j^n - u_{\max} \leq 0$. Furthermore, using the following identity

$$\|q_f\|_{\infty} \geq q_{j+1/2}^{n,+} - q_{j-1/2}^{n,-} = q_{j-1/2}^{n,+} - q_{j+1/2}^{n,-} + \delta_{j,j_f} q_f^n = \begin{cases} q_e^n & \text{if } j < j_f - 1, \\ q_f^n & \text{if } j = j_f, \\ q_u^n & \text{if } j > j_f + 1, \end{cases}$$

we have

$$\begin{aligned} &\mathcal{V}_{j+1/2}^{n,+} u_j^n - \mathcal{V}_{j-1/2}^{n,-} u_j^n + \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n - \mathcal{V}_{j-1/2}^{n,+} u_{j-1}^n - \delta_{j,j_f} q_f^n u_f^n \\ &\geq \|q_f\|_{\infty} (u_j^n - u_{\max}) - L_v(u_{\max} - u_j^n). \end{aligned}$$

Finally, combining the above inequality with (6.3), and CFL condition (6.4), we get

$$u_{\max} - u_j^{n+1} \geq \left(1 - \lambda(\|q_f\|_{\infty} + L_v) - \Delta t L_R \Lambda_{c,2}\right) (u_{\max} - u_j^n) \geq 0,$$

concluding with this the proof. $\square$

Remark. Following the proof of Lemma 6.1, we can straightforwardly show that under the CFL condition (6.4), there holds

$$0 \leq u_j^n + \Delta t \mathbf{1}^t \sigma_c \mathbf{r}_j^n \leq u_{\max}. \quad (6.7)$$

Lemma 6.2. Assume $(u_j^n, \mathbf{c}_j^n, \mathbf{s}_j^n) \in \mathcal{W}$ for all $j = 1, \dots, N$, and consider the following three points first-order scheme for the vector of solid concentrations

$$\mathbf{c}_j^{n+1} = \mathbf{c}_j^n - \lambda(\mathcal{G}_{j+1/2}^{c,n} - \mathcal{G}_{j-1/2}^{c,n}) + \lambda \delta_{j,j_f} q_f^n \mathbf{c}_f^n + \Delta t \sigma_c \mathbf{r}_j^n =: \Phi_j^c(\mathbf{c}_{j-1}^n, \mathbf{c}_j^n, \mathbf{c}_{j+1}^n), \quad (6.8)$$

where

$$\mathcal{G}_{j+1/2}^{c,n} := \text{Upw}(\mathcal{V}_{j+1/2}^n; \mathbf{c}_j^n, \mathbf{c}_{j+1}^n).$$

Then, under CFL condition (6.4), there holds $\mathbf{c}_j^{n+1} \geq \mathbf{0}$ and

$$\mathbf{1}^t \mathbf{c}_j^{n+1} = c_j^{(1),n+1} + c_j^{(2),n+1} + \dots + c_j^{(n_c),n+1} = u_j^{n+1}, \quad \text{for all } j = 1, \dots, N. \quad (6.9)$$

Proof. For the lower bound, we follow the same bounds as in the proof of Lemma 6.1, such that for $i = 1, \dots, n_c$, we have

$$\begin{aligned} c_j^{(i),n+1} &= c_j^{(i),n} - \lambda \mathcal{V}_{j+1/2}^{n,+} c_j^{(i),n} + \lambda \mathcal{V}_{j-1/2}^{n,-} c_j^{(i),n} - \lambda \mathcal{V}_{j+1/2}^{n,-} c_{j+1}^{(i),n} + \lambda \mathcal{V}_{j-1/2}^{n,+} c_{j-1}^{(i),n} \\ &\quad + \lambda \delta_{j,j_f} q_f^n c_f^{(i),n} + \Delta t (\boldsymbol{\sigma}_c \mathbf{r}_j^n)^{(i)} \\ &\geq \left(1 - \lambda(\|q_f\|_\infty + v_0) - \Delta t \Lambda_c R\right) c_j^{(i),n} \geq 0, \end{aligned}$$

where the last inequality is a consequence of the CFL condition (6.4). Note that the bound for the reaction term can be relaxed, but we keep Λ_c as this is already needed in the CFL condition due to Lemma 6.1. Finally, Equation (6.9) is trivially fulfilled as each term on Φ_j^c adds up the corresponding term in (6.5). $\square$

Lemma 6.3. Assume $(u_j^n, c_j^n, s_j^n) \in \mathcal{W}$ for all $j = 1, \dots, N$, and consider the following three points first-order scheme for the vector of substrates

$$s_j^{n+1} = s_j^n - \lambda(\mathcal{G}_{j+1/2}^{s,n} - \mathcal{G}_{j-1/2}^{s,n}) + \lambda \delta_{j,j_f} q_f^n s_f^n + \Delta t \boldsymbol{\sigma}_s \mathbf{r}_j^n =: \Phi_j^s(s_{j-1}^n, s_j^n, s_{j+1}^n), \quad (6.10)$$

where

$$\mathcal{G}_{j+1/2}^{s,n} = \text{Upw} \left(\rho q_{j+1/2}^n - \mathcal{F}_{j+1/2}^n; \frac{s_j^n}{\rho - u_j^n}, \frac{s_{j+1}^n}{\rho - u_{j+1}^n} \right).$$

Then, under CFL condition (6.4), there holds $s_j^{n+1} \geq 0$ for all $j = 1, \dots, N$.

Proof. Taking into account the positive and negative parts from the definition of $\mathcal{G}_{j\pm 1/2}^{s,n}$, for $i = 1, \dots, n_s$, we have

$$s_j^{(i),n+1} \geq s_j^{(i),n} - \lambda \frac{s_j^{(i),n}}{\rho - u_j^n} \left((\rho q_{j+1/2}^n - \mathcal{F}_{j+1/2}^n)^+ - \lambda (\rho q_{j-1/2}^n - \mathcal{F}_{j-1/2}^n)^- \right) + \Delta t (\boldsymbol{\sigma}_s \mathbf{r}_j^n)^{(i)},$$

where each convective term $(\rho q_{j\pm 1/2}^n - \mathcal{F}_{j\pm 1/2}^n)^\pm$ can be lower-bounded using $u_j^n \leq u_{\max} \leq \rho$ and $0 \leq v_{\text{hs}}(u_j^n) \leq v_0$ as follows

$$\begin{aligned} (\rho q_{j+1/2}^n - \mathcal{F}_{j+1/2}^n)^+ &= \left(\rho q_{j+1/2}^n - \mathcal{V}_{j+1/2}^{n,+} u_j^n - \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n \right)^+ \leq \left(\rho q_{j+1/2}^n - \mathcal{V}_{j+1/2}^{n,-} u_{j+1}^n \right)^+ \\ &\leq (\rho q_{j+1/2}^n - q_{j+1/2}^{n,-} u_{j+1}^n)^+ \leq (q_{j+1/2}^{n,-} (\rho - u_{j+1}^n) + \rho q_{j+1/2}^{n,+})^+ \\ &\leq \rho q_{j+1/2}^{n,+}, \end{aligned}$$

and

$$\begin{aligned} (\rho q_{j-1/2}^n - \mathcal{F}_{j-1/2}^n)^- &= \left(\rho q_{j-1/2}^n - \mathcal{V}_{j-1/2}^{n,+} u_{j-1}^n - \mathcal{V}_{j-1/2}^{n,-} u_j^n \right)^- \geq \left(\rho q_{j-1/2}^n - \mathcal{V}_{j-1/2}^{n,+} u_{j-1}^n \right)^+ \\ &\geq (\rho q_{j-1/2}^n - q_{j-1/2}^{n,+} u_{j-1}^n - v_{\text{hs}}(u_j^n) u_{j-1}^n)^- \\ &\geq (q_{j-1/2}^{n,+} (\rho - u_{j-1}^n) + \rho q_{j-1/2}^{n,-} - v_{\text{hs}}(u_j^n) u_{j-1}^n)^- \\ &\geq \rho q_{j-1/2}^{n,-} - v_{\text{hs}}(u_j^n) u_{j-1}^n \geq \rho (q_{j-1/2}^{n,-} - v_0). \end{aligned}$$

Finally, using $\rho/(\rho - u_{\max}) \geq \rho/(\rho - u_j^n)$ and the lower estimate of the reaction term (6.2), we have

$$\begin{aligned} s_j^{(i),n+1} &\geq s_j^{(i),n} - \lambda \frac{\rho s_j^{(i),n}}{\rho - u_j^n} \left(q_{j+1/2}^{n,+} - q_{j-1/2}^{n,-} + v_0 \right) - \Delta t \Lambda_s R s_j^{(i),n} \\ &\geq s_j^{(i),n} - \lambda s_j^{(i),n} \left(\frac{\rho \|q_f\|_\infty}{\rho - u_{\max}} + v_0 \right) - \Delta t \Lambda_s R s_j^{(i),n} \\ &= \left(1 - \lambda \left(\frac{\rho \|q_f\|_\infty}{\rho - u_{\max}} + v_0 \right) - \Delta t \Lambda_s R \right) s_j^{(i),n}, \end{aligned}$$

which together with the CFL condition (6.4), implies the desired result $s_j^{(i),n+1} \geq 0$. $\square$

Combining Lemmas 6.1, 6.2 and 6.3, we can conclude that the first-order scheme, given by equations (6.5), (6.8) and (6.10), is invariant-region-preserving over the set $\mathcal{W}$ for every time iteration t^n provided the CFL condition (6.4) holds. The next theorem states the invariant-region property for the MUSCL scheme (Section 4.1) and CWENO3 including limiters (Section 4.2), respectively, combined with forward Euler time approximations.

Theorem 6.4. *Assume $(u_j^n, c_j^n, s_j^n) \in \mathcal{W}$ for all $j = 1, \dots, N$, and that the CFL condition (6.4) holds with λ/w_1 instead of λ . Consider the full discretization of (3.15), with MUSCL or CWENO3 reconstructions including limiters, and forward Euler in time, given by*

$$u_j^{n+1} = u_j^n + \Delta t \mathcal{H}_j^u(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n), \quad (6.11a)$$

$$\mathbf{c}_j^{n+1} = \mathbf{c}_j^n + \Delta t \mathcal{H}_j^c(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n), \quad (6.11b)$$

$$\mathbf{s}_j^{n+1} = \mathbf{s}_j^n + \Delta t \mathcal{H}_j^s(\mathbf{u}_{\text{st},j}^n, \mathbf{c}_{\text{st},j}^n, \mathbf{s}_{\text{st},j}^n), \quad (6.11c)$$

for $j = 1, \dots, N$, where $\mathbf{u}_{\text{st},j}^n$, $\mathbf{c}_{\text{st},j}^n$ and $\mathbf{s}_{\text{st},j}^n$ are the vectors defined in (3.16). Then, there holds $(u_j^{n+1}, \mathbf{c}_j^{n+1}, \mathbf{s}_j^{n+1}) \in \mathcal{W}$.

Proof. We begin by recalling that since $(u_j^n, \mathbf{c}_j^n, \mathbf{s}_j^n) \in \mathcal{W}$, thanks to Lemma 4.1 in the case of MUSCL scheme or Lemma 4.2 in the case of CWENO3 including limiters, the polynomial reconstructions $(p_j^u, \mathbf{p}_j^c, \mathbf{p}_j^s)$ belong to $\mathcal{W}$ for all $j = 1, \dots, N$. Furthermore, the following identities hold

$$u_j^n = \sum_{\alpha=1}^G w_\alpha p_j^{u,n}(z_j^\alpha), \quad \mathbf{c}_j^n = \sum_{\alpha=1}^G w_\alpha \mathbf{p}_j^{c,n}(z_j^\alpha), \quad \mathbf{s}_j^n = \sum_{\alpha=1}^G w_\alpha \mathbf{p}_j^{s,n}(z_j^\alpha), \quad (6.12)$$

with z_j^α and w_j^α , the nodes and quadrature weights defined in Section 3.1. Without loss of generality, we assume $w_1 = w_G$, and introduce the following short-hand notation for the reaction terms

$$\hat{\mathbf{r}}_{j-1/2}^{n,R} = \sigma_c \mathbf{r}(\hat{\mathbf{c}}_{j-1/2}^{n,R}, \hat{\mathbf{s}}_{j-1/2}^{n,R}), \quad \hat{\mathbf{r}}_{j+1/2}^{n,L} = \sigma_c \mathbf{r}(\hat{\mathbf{c}}_{j+1/2}^{n,L}, \hat{\mathbf{s}}_{j+1/2}^{n,L}), \quad \mathbf{r}_j^{n,\alpha} = \mathbf{r}(\mathbf{p}_j^{c,n}(z_j^\alpha), \mathbf{p}_j^{s,n}(z_j^\alpha)),$$

for all $\alpha = 1, \dots, G$. Then, we first show the bounds related to u_j^{n+1} , for which we use (6.12), add and

subtract $\lambda \mathcal{F}(\widehat{u}_{j-1/2}^{n,R}, \widehat{u}_{j+1/2}^{n,L})$ and reorganise the resulting terms in a convex combination as follows:

$$\begin{aligned}
u_j^{n+1} &= \sum_{\alpha=2}^{G-1} w_\alpha p_j^{u,n}(z_j^\alpha) + w_1 \widehat{u}_{j-1/2}^{n,R} + w_1 \widehat{u}_{j+1/2}^{n,L} - \lambda \left(\mathcal{F}(\widehat{u}_{j+1/2}^{n,L}, \widehat{u}_{j+1/2}^{n,R}) - \mathcal{F}(\widehat{u}_{j-1/2}^{n,L}, \widehat{u}_{j-1/2}^{n,R}) \right) \\
&\quad + \lambda \delta_{j,j_f} q_f^n u_f^n + \Delta t w_1 \mathbf{1}^t \sigma_c \widehat{r}_{j-1/2}^{n,R} + \Delta t w_1 \mathbf{1}^t \sigma_c \widehat{r}_{j+1/2}^{n,L} + \Delta t \sum_{\alpha=2}^{G-1} w_\alpha \mathbf{1}^t \sigma_c r_j^{n,\alpha} \\
&= w_1 \left(\widehat{u}_{j+1/2}^{n,L} - \frac{\lambda}{w_1} (\mathcal{F}(\widehat{u}_{j+1/2}^{n,L}, \widehat{u}_{j+1/2}^{n,R}) - \mathcal{F}(\widehat{u}_{j-1/2}^{n,R}, \widehat{u}_{j+1/2}^{n,L})) \right) + \frac{\lambda}{w_1} \delta_{j,j_f} q_f^n \left(\frac{1}{2} u_f^n \right) + \Delta t \mathbf{1}^t \sigma_c \widehat{r}_{j+1/2}^{n,L} \\
&\quad + \widehat{u}_{j-1/2}^{n,R} - \frac{\lambda}{w_1} (\mathcal{F}(\widehat{u}_{j-1/2}^{n,R}, \widehat{u}_{j+1/2}^{n,L}) - \mathcal{F}(\widehat{u}_{j-1/2}^{n,L}, \widehat{u}_{j-1/2}^{n,R})) + \frac{\lambda}{w_1} \delta_{j,j_f} q_f^n \left(\frac{1}{2} u_f^n \right) + \Delta t \mathbf{1}^t \sigma_c \widehat{r}_{j-1/2}^{n,R} \\
&\quad + \sum_{\alpha=2}^{G-1} w_\alpha (p_j^{u,n}(z_j^\alpha) + \Delta t \mathbf{1}^t \sigma_c r_j^{n,\alpha}) \\
&= w_1 \Phi_j^u(\widehat{u}_{j-1/2}^{n,R}, \widehat{u}_{j+1/2}^{n,L}, \widehat{u}_{j+1/2}^{n,R}) + \sum_{\alpha=2}^{G-1} w_\alpha (p_j^{u,n}(z_j^\alpha) + \Delta t \mathbf{1}^t \sigma_c r_j^{n,\alpha}) \\
&\quad + w_G \Phi_j^u(\widehat{u}_{j-1/2}^{n,L}, \widehat{u}_{j-1/2}^{n,R}, \widehat{u}_{j+1/2}^{n,L}),
\end{aligned}$$

which from Lemma 6.1, and Equation (6.7) as $0 \leq p_j^u \leq u_{\max}$, with CFL condition (6.4) using λ/w_1 instead of λ , we conclude that this is in fact a convex combination of quantities bounded between 0 and $u_{\max}$, therefore $0 \leq u_j^{n+1} \leq u_{\max}$. Similarly, for the vector of solid concentrations and substrates, we have

$$\begin{aligned}
c_j^{n+1} &= w_1 \Phi_j^c(\widehat{c}_{j-1/2}^{n,R}, \widehat{c}_{j+1/2}^{n,L}, \widehat{c}_{j+1/2}^{n,R}) + \sum_{\alpha=2}^{G-1} w_\alpha (p_j^{c,n}(z_j^\alpha) + \Delta t \sigma_c r_j^{n,\alpha}) + w_G \Phi_j^c(\widehat{c}_{j-1/2}^{n,L}, \widehat{c}_{j-1/2}^{n,R}, \widehat{c}_{j+1/2}^{n,L}), \\
s_j^{n+1} &= w_1 \Phi_j^s(\widehat{s}_{j-1/2}^{n,R}, \widehat{s}_{j+1/2}^{n,L}, \widehat{s}_{j+1/2}^{n,R}) + \sum_{\alpha=2}^{G-1} w_\alpha (p_j^{s,n}(z_j^\alpha) + \Delta t \sigma_s r_j^{n,\alpha}) + w_G \Phi_j^s(\widehat{s}_{j-1/2}^{n,L}, \widehat{s}_{j-1/2}^{n,R}, \widehat{s}_{j+1/2}^{n,L}),
\end{aligned}$$

which thanks to Lemmas 6.2 and 6.3, we can conclude that $c_j^{n+1} \geq 0$, $s_j^{n+1} \geq 0$ and $\mathbf{1}^t c_j^{n+1} = u_j^{n+1}$, hence $(u_j^{n+1}, c_j^{n+1}, s_j^{n+1}) \in \mathcal{W}$. □

In the next theorem, we show that discretizations of (2.4) with the MUSCL and CWENO3 reconstructions including limiters, combined with SSPRK approximations of order two and three respectively, satisfy the invariant-region property on $\mathcal{W}$.

Theorem 6.5. *Let Δt satisfy the CFL condition of Theorem 6.4 and assume that the polynomial reconstructions (MUSCL or CWENO3 including limiters) are recomputed, at every Runge–Kutta stage, from the cell averages of that stage. Then the second- and third-order SSPRK methods of Section 5 are invariant-region preserving, that is, given $(u_j^n, c_j^n, s_j^n) \in \mathcal{W}$ then $(u_j^{n+1}, c_j^{n+1}, s_j^{n+1}) \in \mathcal{W}$ for all j and $n \in \mathbb{N}$.*

Proof. Following the ideas from [18, 19], we show that the updated solution of the multi step scheme can be written as a convex combination of terms belonging to $\mathcal{W}$. To do so, we define the vector of solutions $\vec{u}_j = (u_j, c_j, s_j)$ and $\vec{u}_{\text{st},j} = (u_{\text{st},j}, c_{\text{st},j}, s_{\text{st},j})$, and write the SSPRK methods of Section 5 as

$$\vec{u}_j^{(0)} = \vec{u}_j^n, \quad \vec{u}_j^{(i)} = \sum_{k=0}^{i-1} \left(\alpha_{ik} \vec{u}_j^{(k)} + \Delta t \beta_{ik} \vec{\mathcal{H}}_j(\vec{u}_{\text{st},j}^{(k)}) \right), \quad \text{for } i \leq m, \quad \vec{u}_j^{n+1} = \vec{u}_j^{(m)}, \quad (6.13)$$

for all j , where $\vec{\mathcal{H}}_j := (\mathcal{H}_j^u, \mathcal{H}_j^c, \mathcal{H}_j^s)$, and $\alpha_{ik}, \beta_{ik} \geq 0$ are coefficients satisfying

$$\sum_{k=0}^{i-1} \alpha_{ik} = 1 \quad \text{and} \quad \beta_{ik} \leq \alpha_{ik}, \quad \text{for all } i = 1, \dots, m.$$

Now, setting $\tau_{ik} := (\beta_{ik}/\alpha_{ik})\Delta t$ whenever $\alpha_{ik} > 0$, we observe that the i -th stage of (6.13) can be rewritten as the convex combination

$$\vec{\mathbf{u}}_j^{(i)} = \sum_{k=0}^{i-1} \alpha_{ik} \mathcal{E}_{ikj}^{(k)}, \quad \text{with} \quad \mathcal{E}_{ikj}^{(k)} := \vec{\mathbf{u}}_j^{(k)} + \tau_{ik} \vec{\mathcal{H}}_j(\vec{\mathbf{u}}_{\text{st},j}^{(k)}), \quad \mathcal{E}_{ikj}^{(0)} := \vec{\mathbf{u}}_j^m, \quad (6.14)$$

where $\mathcal{E}_{ikj}^{(k)}$ denotes the forward Euler operator with time step $\tau_{ik} > 0$ associated with the semi-discrete scheme (3.15). Since $\tau_{ik} \leq \Delta t$, then τ_{ik} fulfils the CFL condition (6.4) and therefore thanks to Theorem 6.4, if $\vec{\mathbf{u}}_j^{(k)} \in \mathcal{W}$, we have $\mathcal{E}_{\tau_{ik}}^{(k)} \in \mathcal{W}$ for $k = 0, \dots, i-1$ and $i \leq m$. The rest of the proof follows straightforwardly by induction over index i , starting with $\vec{\mathbf{u}}^{(0)} \in \mathcal{W}$, and using the fact that for each i , $\vec{\mathbf{u}}^{(i)}$ is given by a convex combination of $\mathcal{E}_{ikj}^{(k)}$ from $k = 1, \dots, i-1$. $\square$

Remark. Note that condition $\beta_{ik} \leq \alpha_{ik}$ in the proof above holds for the two schemes presented in Section 5, but does not hold automatically for all SSPRK methods.

Remark. The semi-discrete scheme (3.15a)–(3.15c) is advanced in time only for the interior cells indexed by $j \in \{1, \dots, N\}$. The n_g layers of ghost cells on either side of the computational domain, indexed by $j \in \{1 - n_g, \dots, 0\}$ and $j \in \{N + 1, \dots, N + n_g\}$, are not evolved in time. Instead, their values are reset after each stage of the time integrator described in Section 5 according to

$$u_{1-j} := u_1, \quad u_{N+j} := u_N, \quad \text{for } j = 1, \dots, n_g, \quad (6.15)$$

with analogous component-wise assignments for $\mathbf{c}$ and $\mathbf{s}$. For the reconstructions described in Section 4 and the two-point approximations $\mathcal{J}_{j+1/2}^{\text{2nd}}$ (3.6) and $\mathcal{E}_{j+1/2}^{\text{2nd}}$ (3.10) of the diffusive-related terms, we set $n_g = 1$, whereas the fourth-order approximations (3.7) and (3.11) require $n_g = 2$.

In addition, we observe that at the top and bottom of the tank one has $\gamma_{1/2} = \gamma_{N+1/2} = 0$, so that $\mathcal{J}_{1/2} = \mathcal{J}_{N+1/2} = 0$, and consequently $\mathcal{E}_{1/2} = \mathcal{E}_{N+1/2} = 0$. Hence $\hat{\mathcal{V}}_{1/2} = q_{1/2} = -q_e \leq 0$ by (3.9), and from the definition of upwind operator we have $\hat{\mathcal{F}}_{1/2} = -q_e \hat{u}_{1/2}^R$ and $\hat{\mathcal{G}}_{1/2}^c = -q_e \hat{\mathcal{C}}_{1/2}^R$. Furthermore, since $\hat{u}_{1/2}^R \leq u_{\max} < \rho$, then $\rho q_{1/2} - \hat{\mathcal{F}}_{1/2} = -q_e(\rho - \hat{u}_{1/2}^R) \leq 0$, so

$$\hat{\mathcal{G}}_{1/2}^s = -q_e(\rho - \hat{u}_{1/2}^R) \frac{\hat{\mathbf{s}}_{1/2}^R}{\rho - \hat{u}_{1/2}^R} = -q_e \hat{\mathbf{s}}_{1/2}^R.$$

In the same way we obtain that $\hat{\mathcal{V}}_{N+1/2} = q_u \geq 0$ and

$$\hat{\mathcal{F}}_{N+1/2} = q_u \hat{u}_{N+1/2}^L, \quad \hat{\mathcal{G}}_{N+1/2}^c = q_u \hat{\mathcal{C}}_{N+1/2}^L, \quad \hat{\mathcal{G}}_{N+1/2}^s = q_u \hat{\mathbf{s}}_{N+1/2}^L.$$

Therefore, we observe that traces $\hat{u}_{1/2}^L$, $\hat{\mathcal{C}}_{1/2}^L$, $\hat{\mathbf{s}}_{1/2}^L$ and their counterparts at $z_{N+1/2}$ are multiplied by $(-q_e)^+ = 0$ and by $q_u^- = 0$, respectively. Thus, no polynomial reconstruction is therefore needed in the ghost cells. Besides, the ghost values enter the scheme solely through the reconstruction stencils of the two extreme cells I_1 and I_N . Since (6.15) is a copy of an interior state which belongs to $\mathcal{W}$, then hypothesis of Lemmas 6.1 to 6.3 and of Theorem 6.4 hold at the ghost indices with no additional requirement. Moreover, since (6.15) is enforced after each stage of (6.13), every forward Euler operator $\mathcal{E}_{ikj}^{(k)}$ appearing in the proof of Theorem 6.5 acts on a state whose ghost values already belong to $\mathcal{W}$. The induction over the stage index is therefore unaffected.

Table 6.1: Example 1: L^1 spatial errors (7.2) and order of convergence (7.3) computed with the First-order, MUSCL and IRP-CWENO3 schemes at $t = 0.1$ h. The reference solution is computed using the IRP-CWENO3 scheme with $N_{\text{ref}} = 10240$ cells.

First-order												
N	e_N^u	$\mathcal{R}_N^u$	$e_N^{c,1}$	$\mathcal{R}_N^{c,1}$	$e_N^{c,2}$	$\mathcal{R}_N^{c,2}$	$e_N^{s,1}$	$\mathcal{R}_N^{s,1}$	$e_N^{s,2}$	$\mathcal{R}_N^{s,2}$	$e_N^{s,3}$	$\mathcal{R}_N^{s,3}$
160	9.86e-2	—	7.08e-2	—	2.78e-2	—	1.76e-4	—	1.14e-3	—	1.22e-4	—
320	5.14e-2	0.94	3.69e-2	0.94	1.45e-2	0.94	9.08e-5	0.95	5.90e-4	0.96	6.30e-5	0.96
640	2.63e-2	0.97	1.89e-2	0.97	7.41e-3	0.97	4.62e-5	0.97	3.00e-4	0.98	3.20e-5	0.98
1280	1.33e-2	0.98	9.55e-3	0.98	3.75e-3	0.98	2.33e-5	0.99	1.51e-4	0.99	1.61e-5	0.99
2560	6.69e-3	0.99	4.80e-3	0.99	1.89e-3	0.99	1.17e-5	0.99	7.59e-5	0.99	8.09e-6	0.99

MUSCL												
N	e_N^u	$\mathcal{R}_N^u$	$e_N^{c,1}$	$\mathcal{R}_N^{c,1}$	$e_N^{c,2}$	$\mathcal{R}_N^{c,2}$	$e_N^{s,1}$	$\mathcal{R}_N^{s,1}$	$e_N^{s,2}$	$\mathcal{R}_N^{s,2}$	$e_N^{s,3}$	$\mathcal{R}_N^{s,3}$
160	3.09e-3	—	2.22e-3	—	8.72e-4	—	3.47e-6	—	2.10e-5	—	2.40e-6	—
320	7.06e-4	2.13	5.07e-4	2.13	1.99e-4	2.13	7.69e-7	2.17	4.50e-6	2.22	4.95e-7	2.28
640	1.66e-4	2.09	1.19e-4	2.09	4.69e-5	2.09	1.79e-7	2.11	1.02e-6	2.14	1.12e-7	2.15
1280	3.87e-5	2.10	2.78e-5	2.10	1.09e-5	2.10	4.31e-8	2.05	2.44e-7	2.07	2.64e-8	2.08
2560	9.11e-6	2.09	6.54e-6	2.09	2.57e-6	2.09	1.05e-8	2.04	5.95e-8	2.03	6.43e-9	2.04

IRP-CWENO3												
N	e_N^u	$\mathcal{R}_N^u$	$e_N^{c,1}$	$\mathcal{R}_N^{c,1}$	$e_N^{c,2}$	$\mathcal{R}_N^{c,2}$	$e_N^{s,1}$	$\mathcal{R}_N^{s,1}$	$e_N^{s,2}$	$\mathcal{R}_N^{s,2}$	$e_N^{s,3}$	$\mathcal{R}_N^{s,3}$
160	1.99e-3	—	1.43e-3	—	5.62e-4	—	2.52e-6	—	1.63e-5	—	1.80e-6	—
320	2.49e-4	3.00	1.79e-4	3.00	7.02e-5	3.00	3.03e-7	3.06	1.96e-6	3.06	2.16e-7	3.05
640	3.03e-5	3.04	2.17e-5	3.04	8.54e-6	3.04	3.75e-8	3.01	2.42e-7	3.02	2.67e-8	3.02
1280	3.72e-6	3.02	2.67e-6	3.02	1.05e-6	3.02	4.65e-9	3.01	3.00e-8	3.01	3.32e-9	3.01
2560	4.49e-7	3.05	3.22e-7	3.05	1.27e-7	3.05	5.62e-10	3.05	3.63e-9	3.05	4.01e-10	3.05

7 Numerical examples

We consider a reduced biokinetic model, the denitrification process of converting nitrate (NO_3) into nitrogen gas (N_2) presented in [5]. In this model the vector of solids concentration is given by $\mathbf{c} = (c^{(1)}, c^{(2)})$, where $c^{(1)}$ and $c^{(2)}$ correspond to the concentration of ordinary heterotrophic organisms X_{OHO} , and non-degradable organics X_{U} , respectively. The vector of substrates is $\mathbf{s} = (s^{(1)}, s^{(2)}, s^{(3)})$ with $s^{(1)}$, $s^{(2)}$ and $s^{(3)}$ corresponding to the concentration of nitrate S_{NO_3} , readily biodegradable substrate S_{S} , and nitrogen S_{N_2} , respectively. The vector of nonlinear reactions and stoichiometric matrices are given by

$$\mathbf{r}(\mathbf{c}, \mathbf{s}) = \Psi(u) c^{(1)} \begin{pmatrix} \mu(s^{(1)}, s^{(2)}) \\ b \end{pmatrix}, \quad \boldsymbol{\sigma}_{\mathbf{c}} = \begin{pmatrix} 1 & -1 \\ 0 & f_p \end{pmatrix}, \quad \boldsymbol{\sigma}_{\mathbf{s}} = \begin{pmatrix} -\bar{Y} & -1/Y & \bar{Y} \\ 0 & 0.8 & 0 \end{pmatrix},$$

where $b = 6.94 \times 10^{-6} \text{s}^{-1}$ is the decay rate of heterotrophic organisms, $f_p = 0.2$ is the portion of these that decays to non-degradable organics, and the parameters $Y = 0.67$ and $\bar{Y} = 0.172216$ are (dimensionless) yield factors. Function $\Psi := \Psi(u)$ is included such that the fourth condition in Assumption 2.1 is fulfilled, which is assumed to be $\Psi(u) = 1$ for $u < u_{\text{max}} - \tilde{\varepsilon}$ for a small $0 < \tilde{\varepsilon} < \varepsilon$, and $\Psi(u) = 0$ for $u \geq u_{\text{max}} - \varepsilon$. As in all examples shown in this section, the total concentration of solids remains considerably smaller than u_{max} , we simply set $\Psi \equiv 1$. The specific growth rate function is defined by

$$\mu(s^{(1)}, s^{(2)}) = \mu_{\text{max}} \frac{s^{(1)}}{\kappa_1 + s^{(1)}} \frac{s^{(2)}}{\kappa_2 + s^{(2)}},$$

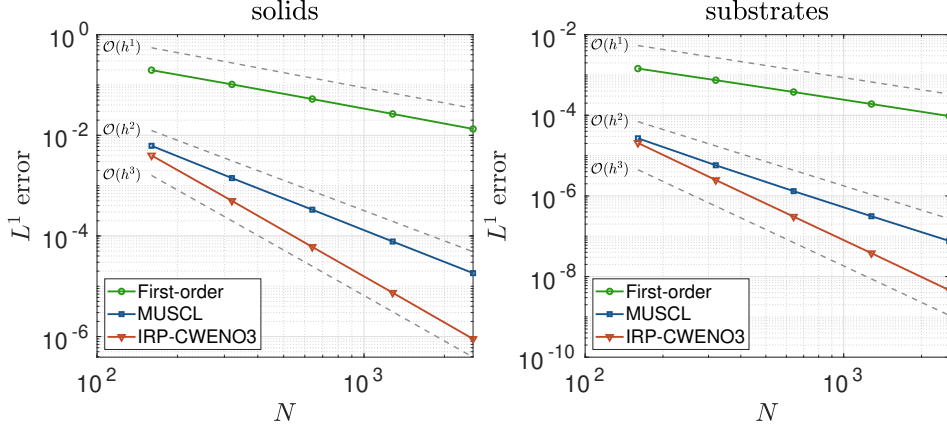


Figure 7.3: Example 1: combined L^1 spatial error of the solids $e_N^u + e_N^{c,1} + e_N^{c,2}$ (left) and substrates $e_N^{s,1} + e_N^{s,2} + e_N^{s,3}$ (right) computed with the First-order, MUSCL and IRP-CWENO3 schemes varying with respect to the number of cells N at $t = 0.1$ h. The dashed lines are the corresponding exact order of convergence curve in each case. The reference solution is computed using the IRP-CWENO3 scheme with $N_{\text{ref}} = 10240$ cells.

with $\mu_{\max} = 5.56 \times 10^{-5} \text{ s}^{-1}$ being the maximum growth rate, and $\kappa_1 = 5 \times 10^{-4} \text{ kg/m}^3$ and $\kappa_2 = 0.02 \text{ kg/m}^3$ are saturation constants. Then, the reaction terms read as follows

$$\sigma_c \mathbf{r} = c^{(1)} \begin{bmatrix} \mu(s^{(1)}, s^{(2)}) - b \\ f_p b \end{bmatrix}, \quad \sigma_s \mathbf{r} = c^{(1)} \begin{bmatrix} -\bar{Y} \mu(s^{(1)}, s^{(2)}) \\ -\mu(s^{(1)}, s^{(2)})/Y + 0.8b \\ \bar{Y} \mu(s^{(1)}, s^{(2)}) \end{bmatrix},$$

The maximal total solids concentration is set to $u_{\max} = 30 \text{ kg/m}^3$. The constitutive functions used in all simulations are

$$v_{\text{hs}}(u) = \frac{v_0}{1 + (u/\tilde{u})^\eta} - \frac{v_0}{1 + (u_{\max}/\tilde{u})^\eta}, \quad \sigma_e(u) := \beta \chi_{\{u \geq u_c\}}(u - u_c),$$

with $\chi_{\{u \geq u_c\}}$ defined as the indicator function over the set $\{u \geq u_c\}$, and constants $v_0 = 1.76 \times 10^{-3} \text{ m/s}$, $\tilde{u} = 3.87 \text{ kg/m}^3$, $u_c = 5 \text{ kg/m}^3$, $\eta = 3.58$ and $\beta = 0.2 \text{ m}^2/\text{s}^2$. Other parameters are $\rho = 1050 \text{ kg/m}^3$, $\rho_L = 998 \text{ kg/m}^3$, $g = 9.81 \text{ m/s}^2$, cross-sectional area $A = 400 \text{ m}^2$, $H = 1 \text{ m}$ and $B = 3 \text{ m}$. The parameters employed in reconstructions are $\alpha_M = 1.4$ and $\varepsilon = h^2$. Furthermore, in what follows, given $J \in \mathbb{N}$, we are going to use the following short-hand notation $\mathbb{Z}_J := \{1, 2, \dots, J\}$.

The numerical scheme and all simulations presented in this work have been implemented in Matlab, and the current solver can be found in the following Github repository:

<https://github.com/ReacSedim/mainsolver>.

In what follows, the first-order scheme combined with forward Euler time discretization used in Lemma 6.1, 6.2 and 6.3, which has also been developed in [4] is going to be denoted as “First-order”. The second-order scheme combining MUSCL reconstruction (Section 4.1) and SSPRK2 in time will be simply referred to as “MUSCL”. The third-order WENO scheme developed in Section 4.2 including the limiters from Section 4.3 and SSPRK3 in time will be termed “IRP-CWENO3”, while the same scheme without limiters will be called “CWENO3”.

7.1 Example 1: Accuracy tests

Since no closed-form solution of the proposed model is available in the literature, we compute the numerical errors using a reference solution obtained from the here developed IRP-CWENO3 scheme

on a fine mesh of N_{ref} cells. To this end, we let $(u_l^{\text{ref}})_{l \in \mathbb{Z}_{N_{\text{ref}}}}$ be the reference solution at the time T under consideration, associated with the cells $I_l^{\text{ref}} = [z_{l-1/2}^{\text{ref}}, z_{l+1/2}^{\text{ref}}]$, for all $l \in \mathbb{Z}_{N_{\text{ref}}}$. Then, given a discretization with N number of cells, being N a divisor of N_{ref} , and a specific cell I_j , we define $\mathcal{K}_j$ as the set of $M := N_{\text{ref}}/N \in \mathbb{N}$ fine cells contained in the coarse I_j :

$$\mathcal{K}_j := \{ l \in \mathbb{Z}_{N_{\text{ref}}} : I_l^{\text{ref}} \subseteq I_j \}.$$

Then, the cell average of the reference solution over I_j is given by

$$\tilde{u}_j^{\text{ref}}(T) := \frac{1}{M} \sum_{l \in \mathcal{K}_j} u_l^{\text{ref}}(T), \quad \text{for all } j \in \mathbb{Z}_N, \quad (7.1)$$

and the average of the solid and substrate components $\tilde{c}_j^{\text{ref},(i)}(T)$ and $\tilde{s}_j^{\text{ref},(k)}(T)$ are defined analogously for each component $i = 1, \dots, n_c$ and $k = 1, \dots, n_s$. Furthermore, for a fixed number of cells N , we denote by $u_j^N(T)$, $c_j^{N,(i)}(T)$ and $s_j^{N,(k)}(T)$ the numerical solution at time T computed on the mesh of N cells. The approximate L^1 errors of each individual unknown are then defined by

$$\begin{aligned} e_N^u(T) &:= h \sum_{j \in \mathbb{Z}_N} |\tilde{u}_j^{\text{ref}}(T) - u_j^N(T)|, \\ e_N^{c,i}(T) &:= h \sum_{j \in \mathbb{Z}_N} |\tilde{c}_j^{\text{ref},(i)}(T) - c_j^{N,(i)}(T)|, \quad i = 1, \dots, n_c, \\ e_N^{s,k}(T) &:= h \sum_{j \in \mathbb{Z}_N} |\tilde{s}_j^{\text{ref},(k)}(T) - s_j^{N,(k)}(T)|, \quad k = 1, \dots, n_s, \end{aligned} \quad (7.2)$$

where $h = (H + B)/N$. We also measure the convergence rates by

$$\mathcal{R}_N^u := \log_2 \left(\frac{e_N^u}{e_{2N}^u} \right), \quad \mathcal{R}_N^{c,i} := \log_2 \left(\frac{e_N^{c,i}}{e_{2N}^{c,i}} \right), \quad \mathcal{R}_N^{s,k} := \log_2 \left(\frac{e_N^{s,k}}{e_{2N}^{s,k}} \right), \quad (7.3)$$

for all $i = 1, \dots, n_c$ and $k = 1, \dots, n_s$.

For this test, we consider a sequence of spatial discretizations with $N_l = 10 \cdot 2^l$, for $l = 4, \dots, 8$, and compute a reference solution using $N_{\text{ref}} = 10240$ grid points. We set $q_f(t) = q_u(t) = q_e(t) = 0$, which implies that $u_f(t) = 0$, $c_f(t) = \mathbf{0}$, and $s_f(t) = \mathbf{0}$. Furthermore, we prescribe the following smooth initial condition:

$$u_0(z) = 5 + 2\psi(z) \text{ kg/m}^3, \quad c_0(z) = u_0(z) \begin{pmatrix} 5/7 \\ 2/7 \end{pmatrix}, \quad s_0(z) = \begin{pmatrix} 0.5 + 0.05\psi(z) \\ 1.0 + 0.1\psi(z) \\ 0.1 + 0.01\psi(z) \end{pmatrix} \text{ kg/m}^3,$$

where $\psi(z) := \varphi(z - 1)$, and φ is the function defined by

$$\varphi(\zeta) := \begin{cases} (1 - \zeta^2)^6 & \text{if } |\zeta| < 1, \\ 0 & \text{if } |\zeta| \geq 1. \end{cases}$$

The approximate L^1 spatial errors defined by (7.2) and their corresponding orders of convergence given by (7.3) are displayed in Table 6.1 at time $t = 0.1$ h for the three schemes, First-order, MUSCL and IRP-CWENO3. We observe that the numerical errors effectively decrease upon mesh refinement for all components of the solution, and the order of convergence in all unknowns confirm first-, second- and third-order for the First-order, MUSCL, and IRP-CWENO3 schemes, respectively. In Figure 7.3 we show the error plots for the combined L^1 errors of the solid components given by $e_N^u + e_N^{c,1} + e_N^{c,2}$, and the combined errors for the substrates $e_N^{s,1} + e_N^{s,2} + e_N^{s,3}$ for the three already mentioned schemes. In all methods, we observe that the error curves are approximately parallel to the actual curves related to the exact orders of convergence $\mathcal{O}(h^1)$, $\mathcal{O}(h^2)$ and $\mathcal{O}(h^3)$, respectively.

Table 7.2: Example 2: invariant-region-preserving without diffusion for the CWENO3 and IRP-CWENO3 schemes, respectively. The minimum and maximum values related to set $\mathcal{W}$ at the specified times t , denoted by underlining and overlining, respectively, are defined in (7.4).

$t = 3 \text{ h}$						IRP-CWENO3				
N	<u>$\underline{u}$</u>	<u>$\bar{u}$</u>	<u>$\underline{c}$</u>	<u>$\underline{s}$</u>	<u>$u - \mathbf{1}^t \mathbf{c}$</u>	<u>$\underline{u}$</u>	<u>$\bar{u}$</u>	<u>$\underline{c}$</u>	<u>$\underline{s}$</u>	<u>$u - \mathbf{1}^t \mathbf{c}$</u>
40	-1.08e-2	19.35	-7.72e-3	-5.93e-3	4.44e-12	0.0	19.35	0.0	0.0	4.38e-12
80	-8.48e-3	21.18	-6.05e-3	-3.15e-3	5.30e-12	0.0	21.18	0.0	0.0	5.24e-12
160	-4.45e-3	22.34	-3.18e-3	-8.31e-4	5.81e-12	0.0	22.34	0.0	0.0	5.62e-12
320	-2.77e-3	22.82	-1.97e-3	-7.68e-4	5.91e-12	0.0	22.82	0.0	0.0	5.92e-12
640	-1.68e-3	23.00	-1.20e-3	-5.80e-4	7.25e-12	0.0	23.00	0.0	0.0	7.25e-12

$t = 6 \text{ h}$						IRP-CWENO3				
N	<u>$\underline{u}$</u>	<u>$\bar{u}$</u>	<u>$\underline{c}$</u>	<u>$\underline{s}$</u>	<u>$u - \mathbf{1}^t \mathbf{c}$</u>	<u>$\underline{u}$</u>	<u>$\bar{u}$</u>	<u>$\underline{c}$</u>	<u>$\underline{s}$</u>	<u>$u - \mathbf{1}^t \mathbf{c}$</u>
40	-1.08e-2	19.91	-7.72e-3	-5.93e-3	5.30e-12	0.0	19.91	0.0	0.0	4.93e-12
80	-8.48e-3	21.18	-6.05e-3	-3.15e-3	6.47e-12	0.0	21.18	0.0	0.0	6.43e-12
160	-4.45e-3	22.34	-3.18e-3	-8.31e-4	7.76e-12	0.0	22.34	0.0	0.0	7.65e-12
320	-2.77e-3	22.82	-1.97e-3	-7.68e-4	9.83e-12	0.0	22.82	0.0	0.0	9.76e-12
640	-1.68e-3	23.00	-1.20e-3	-5.80e-4	1.30e-11	0.0	23.00	0.0	0.0	1.31e-11

7.2 Example 2: Invariant-region-preservation without diffusion

To show that the invariant-region results from Section 6 hold for the IRP-CWENO3 scheme, i.e., the computed (u_j^n, c_j^n, s_j^n) belongs to $\mathcal{W}$ (2.10) for all cells I_j and time t^n , we consider a standard example of reactive sedimentation from [4, 5]. For this purpose, we specify the following piecewise-constant-in-time feed concentrations:

$$s_f(t) = \begin{pmatrix} 0.006 \\ 0.0009 \\ 0 \end{pmatrix} \text{ kg/m}^3, \quad c_f(t) = u_f(t) \begin{pmatrix} 5/7 \\ 2/7 \end{pmatrix}, \quad u_f(t) = \begin{cases} 1.0 \text{ kg/m}^3 & \text{if } 0 \text{ h} \leq t < 2 \text{ h}, \\ 0.5 \text{ kg/m}^3 & \text{if } 2 \text{ h} \leq t < 4 \text{ h}, \\ 3.0 \text{ kg/m}^3 & \text{if } 4 \text{ h} \leq t < 7 \text{ h}, \\ 4.0 \text{ kg/m}^3 & \text{if } t \geq 7 \text{ h}, \end{cases}$$

and the feed and underflow bulk velocities

$$q_f(t) = \begin{cases} 1.125 \text{ m/h} & \text{if } 0 \text{ h} \leq t < 2 \text{ h}, \\ 0.325 \text{ m/h} & \text{if } 2 \text{ h} \leq t < 4 \text{ h}, \\ 1.625 \text{ m/h} & \text{if } t \geq 4 \text{ h}, \end{cases} \quad q_u(t) = \begin{cases} 0.075 \text{ m/h} & \text{if } 0 \text{ h} \leq t < 2 \text{ h}, \\ 0.25 \text{ m/h} & \text{if } 2 \text{ h} \leq t < 4 \text{ h}, \\ 0.0875 \text{ m/h} & \text{if } 4 \text{ h} \leq t < 7 \text{ h}, \\ 0.125 \text{ m/h} & \text{if } t \geq 7 \text{ h}, \end{cases}$$

where the effluent velocity is given by $q_e(t) = q_f(t) - q_u(t)$. In addition, we also set the following initial functions:

$$c_0(z) = u_0(z) \begin{pmatrix} 5/7 \\ 2/7 \end{pmatrix}, \quad u_0(z) = \begin{cases} 0 \text{ kg/m}^3 & \text{if } z < 0.5 \text{ m}, \\ 3.8z + 1.6 \text{ kg/m}^3 & \text{if } z \geq 0.5 \text{ m}, \end{cases}$$

$$s_0(z) = \begin{cases} (0.006, 0, 0)^t \text{ kg/m}^3 & \text{if } z < 0.5 \text{ m}, \\ (0, 0.12(z - 0.5), 0.006)^t \text{ kg/m}^3 & \text{if } z \geq 0.5 \text{ m}. \end{cases}$$

In order to assess the effect of the scaling limiters described in section 4.3, we monitor the minimum and maximum values related to set $\mathcal{W}$ attained by each discrete unknown over the whole space-time

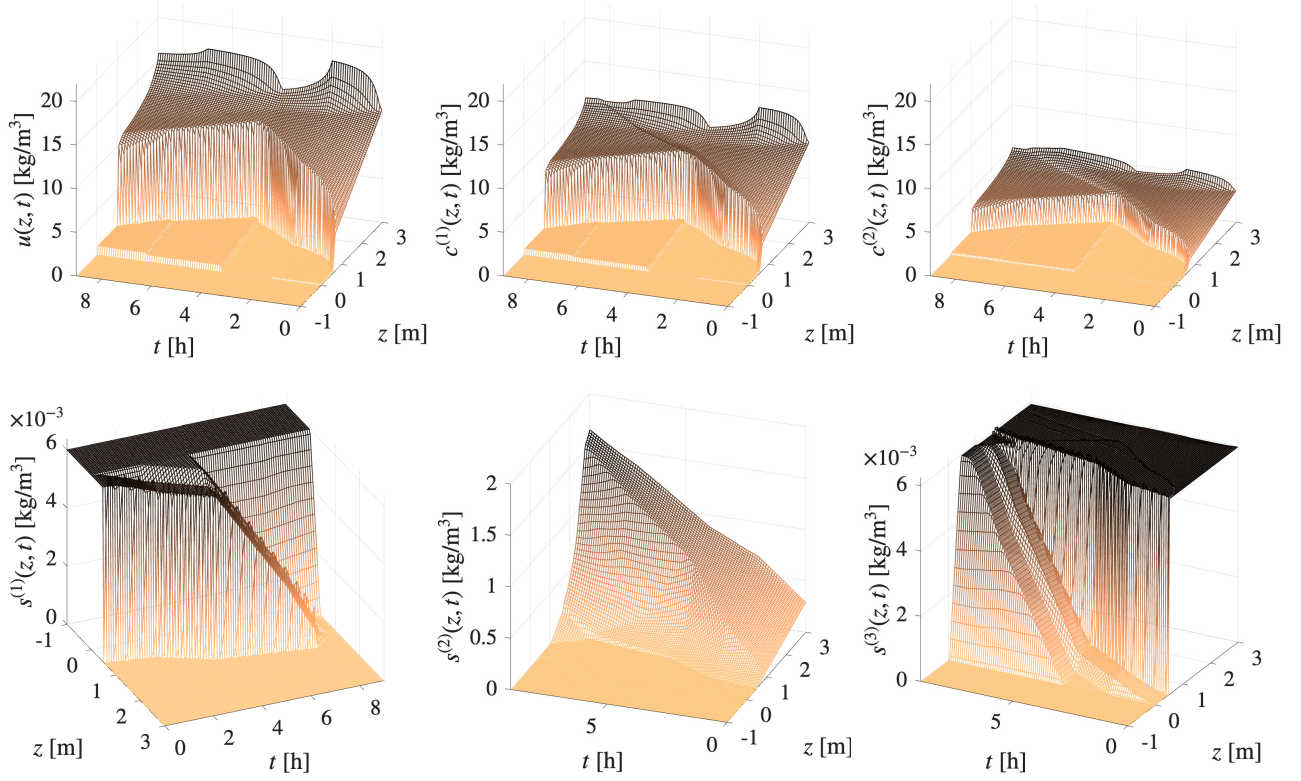


Figure 7.4: Example 2: numerical solution with $N = 640$ produced with IRP-CWENO3 scheme at $t = 9$ h. The solution has been projected onto a coarse grid.

computational grid, that is, we compute

$$\begin{aligned}
 \underline{u} &:= \min_{(j,n) \in \mathbb{Z}_N \times \mathbb{Z}_{N_T}} \{u_j^n\}, & \bar{u} &:= \max_{(j,n) \in \mathbb{Z}_N \times \mathbb{Z}_{N_T}} \{u_j^n\}, & \underline{c} &:= \min_{\substack{(j,n) \in \mathbb{Z}_N \times \mathbb{Z}_{N_T} \\ i \in \{1, \dots, n_c\}}} \{c_j^{n,(i)}\}, \\
 \underline{s} &:= \min_{\substack{(j,n) \in \mathbb{Z}_N \times \mathbb{Z}_{N_T} \\ l \in \{1, \dots, n_s\}}} \{s_j^{n,(l)}\}, & \overline{|u - \mathbf{1}^\top \mathbf{c}|} &:= \max_{(j,n) \in \mathbb{Z}_N \times \mathbb{Z}_{N_T}} \{|u_j^n - \mathbf{1}^\top \mathbf{c}_j^n|\},
 \end{aligned} \tag{7.4}$$

where both $\underline{c}$ and $\underline{s}$ are scalars and N_T stands for the number of time steps. The quantities defined in (7.4), computed with the CWENO3 scheme without limiters, and with the invariant-region-preserving scheme IRP-CWENO3, are reported in Table 7.2 for different mesh sizes at two simulation times, namely, $t = 3$ h and $t = 6$ h. For the schemes without limiters, negative values of $\underline{u}$, $\underline{c}$, and $\underline{s}$ occur in the numerical solutions for all mesh sizes and at both simulation times. By contrast, when the limiters are applied, the numerical solutions remain within $\mathcal{W}$, as shown in Section 6. In addition, in Figure 7.4, we show the numerical solution obtained with the IRP-CWENO3 and observe that the scheme accurately reproduce the test case from [4]. Figure 7.5 compares the first-, second- and third-order schemes on a fixed grid with $N = 160$ cells against a reference solution computed with the IRP-CWENO3 scheme on a finer mesh (with $N_{\text{ref}} = 640$) and we observe that the numerical solution approaches the reference one as the order of the scheme increases.

7.3 Example 3: Diffusive case with compression effects

In this example, we consider the inclusion of the diffusion-related term $\partial_z \mathcal{B}$ (resp. $\partial_z \mathcal{D}$) in Equation (2.4a), which accounts for the compression effects above a critical concentration u_c , and its

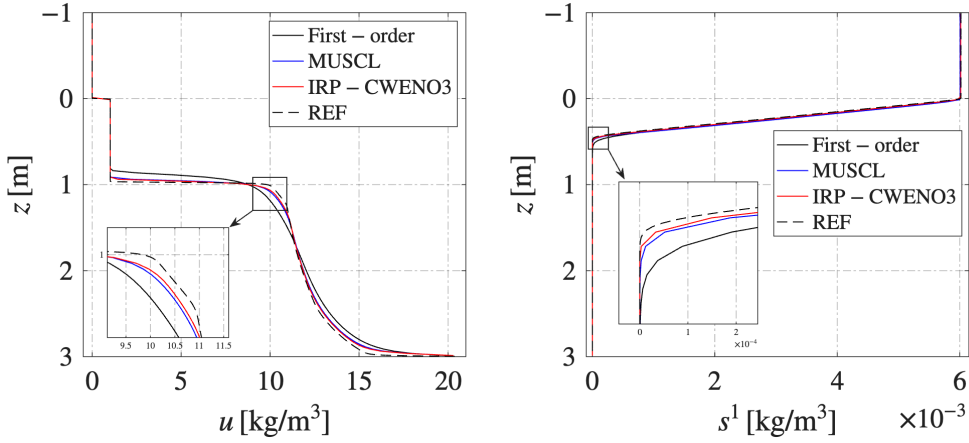


Figure 7.5: Example 2: numerical solutions with $N = 160$ produced with first-order, MUSCL and IRP-CWENO3 schemes at $t = 9$ h. The reference solution is computed with $N_{\text{ref}} = 640$ using IRP-CWENO3 scheme.

corresponding flux term $\mathcal{E}$ (resp. $\mathcal{J}$). We study the MUSCL and IRP-CWENO3 schemes combined with approximations of the diffusion-related operator given by $\mathcal{E}_{j+1/2}$ following the $\mathcal{D}$ -method and $\mathcal{B}$ -method from Section 3. For this we consider the second- and fourth-order approximations of $\mathcal{E}$ (resp. $\mathcal{J}$) defined in (3.10) and (3.11) (resp. (3.6) and (3.7)).

First, we use the functions and initial conditions from Example 1 to assess the convergence rates of both schemes in combination with the two proposed methods to approximate the diffusion terms. We then compute a reference solution with the IRP-CWENO3 scheme using the $\mathcal{D}$ -method with $N_{\text{ref}} = 2560$ and $\mathcal{J}_{j+1/2} = \mathcal{J}_{j+1/2}^{4\text{th}}$ until time $t = 0.1$ h. The spatial numerical errors are reported in Table 7.3 and show that both schemes, MUSCL and IRP-CWENO3, achieve their expected orders of convergence with both approaches, $\mathcal{D}$ -method and $\mathcal{B}$ -method, respectively. Next, we use the functions and initial conditions from Example 2 and simulate the continuous reactive sedimentation problem up to $t = 9$ h with $N = 640$. The corresponding numerical solutions are presented in Figure 7.6, where we also observe the evolution of the expected discontinuity produced due to the boundary and initial conditions, and the corresponding smoothing effect in the plot of u for concentrations above u_c . In addition, using this setting, we investigate the invariant-region-preserving property of the MUSCL and IRP-CWENO3 schemes when combined with the proposed discretizations of the diffusion-related terms. The results reported in Table 7.4 show that the invariant region is preserved, that is, $\underline{u} \approx 0$, $\underline{c} \approx 0$, $\underline{s} \approx 0$, $\underline{u} < u_{\text{max}}$ and $|u - \mathbf{1}^T \mathbf{c}| \approx 0$, in both approaches, with the $\mathcal{D}$ -method and $\mathcal{B}$ -method, respectively.

8 Conclusions

In this work, we have developed high-order finite volume numerical schemes for the approximation of a model of reactive sedimentation, which consists of a system of convection-diffusion-reaction equations. Two component-wise polynomial reconstructions have been proposed, a second-order MUSCL and a third-order central WENO scheme (CWENO3), which in both cases, the polynomial reconstructions for the solid components are designed such that the sum of each solid component gives the total concentration of solids. Positivity of each concentration component, and boundedness of the total concentration of solids are established directly from the definition of the MUSCL scheme, while limiters were implemented for the CWENO3 scheme. In this way, we adapted the standard Zhang and Shu

Table 7.3: Example 3: numerical errors (7.2) and approximate order of convergence (7.3) at time $t = 0.1$ h computed with the MUSCL and IRP-CWENO3 schemes, respectively. The diffusion function is approximated using (3.6). The reference solution is computed using the IRP-CWENO3 scheme with $N_{\text{ref}} = 2560$ cells.

MUSCL & $\mathcal{D}$ -method with $\mathcal{J}_{j+1/2}^{2\text{nd}}$												
N	e_N^u	$\mathcal{R}_N^u$	$e_N^{c,1}$	$\mathcal{R}_N^{c,1}$	$e_N^{c,2}$	$\mathcal{R}_N^{c,2}$	$e_N^{s,1}$	$\mathcal{R}_N^{s,1}$	$e_N^{s,2}$	$\mathcal{R}_N^{s,2}$	$e_N^{s,3}$	$\mathcal{R}_N^{s,3}$
80	1.42e-3	—	1.02e-3	—	4.00e-4	—	3.43e-6	—	2.21e-5	—	2.68e-6	—
160	3.99e-4	1.83	2.87e-4	1.83	1.13e-4	1.83	1.00e-6	1.78	6.31e-6	1.81	7.62e-7	1.81
320	1.13e-4	1.82	8.10e-5	1.82	3.18e-5	1.82	2.73e-7	1.88	1.80e-6	1.81	2.14e-7	1.84
640	2.99e-5	1.92	2.15e-5	1.92	8.44e-6	1.92	7.17e-8	1.93	4.78e-7	1.91	5.62e-8	1.93
MUSCL & $\mathcal{B}$ -method with $\mathcal{E}_{j+1/2}^{2\text{nd}}$												
80	1.44e-3	—	1.03e-3	—	4.05e-4	—	3.51e-6	—	2.29e-5	—	2.75e-6	—
160	3.93e-4	1.87	2.82e-4	1.87	1.11e-4	1.87	9.94e-7	1.82	6.24e-6	1.87	7.45e-7	1.88
320	1.11e-4	1.83	7.94e-5	1.83	3.12e-5	1.83	2.68e-7	1.89	1.75e-6	1.83	2.08e-7	1.84
640	2.93e-5	1.92	2.10e-5	1.92	8.27e-6	1.92	7.01e-8	1.93	4.65e-7	1.91	5.47e-8	1.93
IRP-CWENO3 & $\mathcal{D}$ -method with $\mathcal{J}_{j+1/2}^{4\text{th}}$												
80	1.59e-3	—	1.14e-3	—	4.48e-4	—	4.01e-6	—	2.84e-5	—	3.24e-6	—
160	1.93e-4	3.04	1.39e-4	3.04	5.45e-5	3.04	4.81e-7	3.06	3.39e-6	3.07	3.87e-7	3.06
320	2.49e-5	2.95	1.79e-5	2.95	7.04e-6	2.95	5.99e-8	3.01	4.19e-7	3.02	4.75e-8	3.03
640	3.56e-6	2.81	2.56e-6	2.81	1.00e-6	2.81	7.78e-9	2.95	5.33e-8	2.97	5.95e-9	3.00
IRP-CWENO3 & $\mathcal{B}$ -method with $\mathcal{E}_{j+1/2}^{4\text{th}}$												
80	1.60e-3	—	1.15e-3	—	4.52e-4	—	4.04e-6	—	2.87e-5	—	3.28e-6	—
160	1.94e-4	3.05	1.39e-4	3.05	5.47e-5	3.05	4.84e-7	3.06	3.41e-6	3.07	3.90e-7	3.07
320	2.50e-5	2.96	1.79e-5	2.96	7.05e-6	2.96	6.01e-8	3.01	4.20e-7	3.02	4.77e-8	3.03
640	3.56e-6	2.81	2.56e-6	2.81	1.01e-6	2.81	7.79e-9	2.95	5.34e-8	2.98	5.96e-9	3.00

[44, 46] limiters to achieve that the CWENO3 polynomials lie in the set of physically relevant solutions $\mathcal{W}$ (2.10).

As a first result, we show that under an appropriate CFL condition (6.4), each component of the numerical solution computed with the related first-order upwind-based scheme without compression effects (no diffusion) and forward Euler in time satisfies an invariant-region-preserving property in $\mathcal{W}$. Some ingredients used to obtain this result were the decreasing and Lipschitz continuity condition of the hindered-settling velocity v_{hs} and the hypothesis on the reaction terms given in Assumption 2.1. Unlike previous works [4, 7, 8], these invariant-region properties for the first-order scheme were achieved without using a monotonicity argument. The main theoretical results are the invariant-region-preserving property in $\mathcal{W}$ for the high-order schemes with the MUSCL and CWENO3 schemes, respectively, with forward Euler time approximation, and also with the strong-stability preserving (SSP) second- and third-order TVD Runge-Kutta time discretizations. The drawback, in the case of the higher-order time discretization, is a more restrictive CFL condition. For the approximation of the diffusion-related terms, we proposed two approaches, one discretizing the standard compression function $\mathcal{D}$ [10] which corresponds to a degenerate diffusion function for the total concentration u , and a second approach for based on the factor function $\mathcal{B}$. In both cases, the diffusion-related flux $\mathcal{E}$ (2.8) was included in the approximation of the total flux $\mathcal{F}$ (2.9) of the equation for the total concentration u , which was based on an upwind scheme combined with the corresponding polynomial reconstructions. This particular approximation of $\mathcal{F}$ (2.9) including the diffusion-related flux $\mathcal{E}$, and the proposed approximation for the equation of the vector of solid components $\mathbf{c}$, was devised such that the sum of the

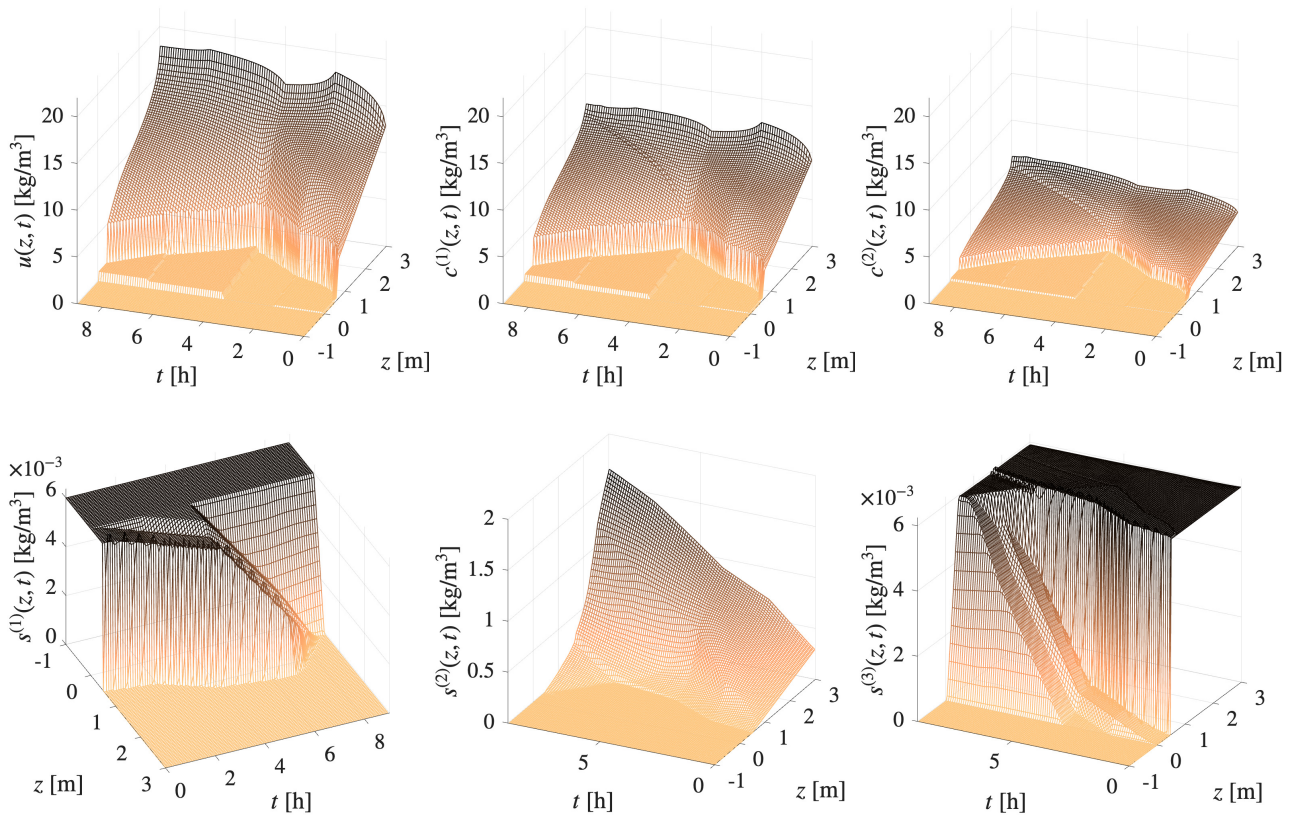


Figure 7.6: Example 3: numerical solution with $N = 640$ produced with IRP-CWENO3 scheme at $t = 9$ h. The solution has been projected onto a coarse grid.

solid components equals u . The numerical results corroborate the theoretical findings, in particular the invariant-region property, and also the order of accuracy in the case without diffusion. Notably, the numerical examples show that the proposed approaches to discretize the diffusion-related terms preserve the desired order of convergence, and also preserve the invariant-region property.

Regarding future extensions of this work, the main line of research is to establish the theoretical orders of accuracy of the proposed schemes for the diffusion terms, and to show that these schemes also preserve the invariant-region property. An analysis of the stability of the proposed numerical schemes also deserves to be the subject of future research. Further studies of the reactive sedimentation model in multidimensional context, either in axisymmetric domains or three dimensional ones, are of research interest. However, multidimensional models often require solving additional Navier-Stokes equations for the bulk velocity, which becomes a vector, and the so-called excess pore-pressure. Therefore, the bulk velocity is no longer a given function, leading to the need of additional polynomial reconstructions for this bulk velocity. Finally, as the model is written in terms of an arbitrary number of solid and substrate components, this work is suitable for the simulation of more general activated sludge processes involving several concentrations of solid particles and substrates, such as the ASM1 [21] or general eutrophication processes [3, 17, 32, 35, 39].

Funding

This work was partially supported by ANID-Chile through Fondecyt project No. 3230553. LMV is supported by ANID - Chile through of Centro de Modelamiento Matemático (CMM), project FB210005

Table 7.4: Example 3: invariant-region-preserving test including diffusion for the MUSCL and IRP-CWENO3 schemes with the $\mathcal{D}$ -method and $\mathcal{B}$ -method from Section 3, respectively. The minimum and maximum values related to set $\mathcal{W}$ at the specified time t , denoted by underlining and overlining, respectively, are defined in (7.4).

MUSCL										
$t = 9 \text{ h}$	$\mathcal{D}$ -method with $\mathcal{J}_{j+1/2}^{2\text{nd}}$					$\mathcal{B}$ -method with $\mathcal{E}_{j+1/2}^{2\text{nd}}$				
N	<u>u</u>	$\bar{u}$	<u>c</u>	<u>s</u>	$ u - \mathbf{1}^t c $	<u>u</u>	$\bar{u}$	<u>c</u>	<u>s</u>	$ u - \mathbf{1}^t c $
40	0.0	18.50	0.0	0.0	1.37e-12	0.0	18.49	0.0	0.0	1.44e-12
80	0.0	18.88	0.0	0.0	1.35e-12	0.0	18.88	0.0	0.0	9.63e-13
160	0.0	19.04	0.0	0.0	2.28e-12	0.0	19.04	0.0	0.0	9.63e-13
320	0.0	19.11	0.0	0.0	6.45e-12	0.0	19.11	0.0	0.0	1.04e-12

IRP-CWENO3										
$t = 9 \text{ h}$	$\mathcal{D}$ -method with $\mathcal{J}_{j+1/2}^{4\text{th}}$					$\mathcal{B}$ -method with $\mathcal{E}_{j+1/2}^{4\text{th}}$				
N	<u>u</u>	$\bar{u}$	<u>c</u>	<u>s</u>	$ u - \mathbf{1}^t c $	<u>u</u>	$\bar{u}$	<u>c</u>	<u>s</u>	$ u - \mathbf{1}^t c $
40	0.0	18.45	0.0	0.0	5.26e-12	0.0	18.44	0.0	0.0	5.60e-12
80	0.0	18.85	0.0	0.0	8.51e-12	0.0	18.85	0.0	0.0	8.64e-12
160	0.0	19.03	0.0	0.0	1.15e-11	0.0	19.03	0.0	0.0	1.28e-11
320	0.0	19.10	0.0	0.0	2.40e-11	0.0	19.10	0.0	0.0	3.15e-11

of BASAL funds for Centers of Excellence. JBC is supported by the National Agency for Research and Development, ANID-Chile through the Scholarship Program, Becas Doctorado Nacional 2022, 21221387.

Declarations

Conflicts of Interest. The authors declare that they have no conflicts of interest.

Data Availability. This manuscript has no associated data.

References

- [1] J. Barajas-Calonge, R. Bürger, P. Mulet, and L.M. Villada. Invariant-region-preserving WENO schemes for one-dimensional multispecies kinematic flow models. *J. Comput. Phys.*, 537:114081, 2025.
- [2] J. Barajas-Calonge, R. Bürger, P. Mulet, and L.M. Villada. A Second-Order Invariant-Region-Preserving Scheme for a Transport-Flow Model of Polydisperse Sedimentation. *J. Sci. Comput.*, 108(1):29, 2026.
- [3] A.N. Blauw, H.F.J. Los, M. Bokhorst, and P.L.A. Erftemeijer. Gem: a generic ecological model for estuaries and coastal waters. *Hydrobiologia*, 618(1), 2008.
- [4] R. Bürger, J. Careaga, and S. Diehl. A method-of-lines formulation for a model of reactive settling in tanks with varying cross-sectional area. *IMA J. Appl. Math.*, 86(3):514–546, 2021.

- [5] R. Bürger, J. Careaga, S. Diehl, C. Mejías, I. Nopens, E. Torfs, and P. A. Vanrolleghem. Simulations of reactive settling of activated sludge with a reduced biokinetic model. *Computers Chem. Eng.*, 92:216–229, 2016.
- [6] R. Bürger, J. Careaga, S. Diehl, and R. Pineda. A moving-boundary model of reactive settling in wastewater treatment. Part 1: Governing equations. *Appl. Math. Modelling*, 106:390–401, 2022.
- [7] R. Bürger, J. Careaga, S. Diehl, and R. Pineda. A moving-boundary model of reactive settling in wastewater treatment. Part 2: Numerical scheme. *Appl. Math. Modelling*, 111:247–269, 2022.
- [8] R. Bürger, J. Careaga, S. Diehl, and R. Pineda. Numerical schemes for a moving-boundary convection-diffusion-reaction model of sequencing batch reactors. *ESAIM: Math. Model. Numer. Anal.*, 57(5):2931–2976, 2023.
- [9] R. Bürger, S. Diehl, and C. Mejías. A difference scheme for a degenerating convection-diffusion-reaction system modelling continuous sedimentation. *ESAIM: Math. Modelling Num. Anal.*, 52(2):365–392, 2018.
- [10] R. Bürger, K.H. Karlsen, and J.D. Towers. A model of continuous sedimentation of flocculated suspensions in clarifier-thickener units. *SIAM J. Appl. Math.*, 65(3):882–940, 2005.
- [11] J.-Ph. Chancelier, M. Cohen De Lara, and F. Pacard. Analysis of a Conservation PDE With Discontinuous Flux: A Model of Settler. *SIAM J. Appl. Math.*, 54(4):954–995, 1994.
- [12] J. De Clercq, I. Nopens, J. Defrancq, and P.A. Vanrolleghem. Extending and calibrating a mechanistic hindered and compression settling model for activated sludge using in-depth batch experiments. *Water Res.*, 42(3):781–791, 2008.
- [13] I. Cravero, G. Puppo, M. Semplice, and G. Visconti. Cweno: Uniformly accurate reconstructions for balance laws. *Math. Comput.*, 87(312):1689–1719, 2017.
- [14] M.N. Cruz Bournazou, H. Arellano-Garcia, G. Wozny, G. Lyberatos, and C. Kravaris. ASM3 extended for two-step nitrification–denitrification: a model reduction for sequencing batch reactors. *J. Chem. Technol. Biotechnol.*, 87(7):887–896, 2012.
- [15] S. Diehl. A conservation Law with Point Source and Discontinuous Flux Function Modelling Continuous Sedimentation. *SIAM J. Appl. Math.*, 56(2):388–419, 1996.
- [16] X. Flores-Alsina, K.V. Gernaey, and U. Jeppsson. Benchmarking biological nutrient removal in wastewater treatment plants: influence of mathematical model assumptions. *Water Sci. Technol.*, 65(8):1496–1505, 2012.
- [17] S. Gargallo, M. Martín, N. Oliver, and C. Hernández-Crespo. Sedimentation and resuspension modelling in free water surface constructed wetlands. *Ecol. Eng.*, 98:318–329, 2017.
- [18] S. Gottlieb and C.-W. Shu. Total variation diminishing Runge-Kutta schemes. *Math. Comput.*, 67(221):73–85, 1998.
- [19] S. Gottlieb, C.-W. Shu, and E. Tadmor. Strong stability-preserving high-order time discretization methods. *SIAM review*, 43(1):89–112, 2001.
- [20] M. Henze, C.P.L. Grady, W. Gujer, G.V.R. Marais, and T. Matsuo. Activated Sludge Model No. 1. Sci. tech. rep. no. 1., IAWQ, London, UK, 1987. No DOI assigned.
- [21] M. Henze, W. Gujer, T. Mino, and M.C.M. van Loosdrecht. *Activated Sludge Models ASM1, ASM2, ASM2d and ASM3*. IWA Publishing, London, UK, 2000. No DOI assigned.

- [22] U. Jeppsson and G. Olsson. Reduced order models for on-line parameter identification of the activated sludge process. *Water Sci. Technol.*, 28(11-12):173–183, 1993.
- [23] G.-S. Jiang and C.-W. Shu. Efficient implementation of weighted ENO schemes. *J. Comput. Phys.*, 126(1):202–228, 1996.
- [24] V.P. Kolgan. Application of the minimum-derivative principle in the construction of finite-difference schemes for numerical analysis of discontinuous solutions in gas dynamics. *Uch. Zap. TsAGI [Sci. Notes Central Inst. Aerodyn]*, 3(6):68–77, 1972.
- [25] D. Kuzmin and H. Hajduk. *Property-preserving numerical schemes for conservation laws*. World Scientific, 2024.
- [26] G.J. Kynch. A theory of sedimentation. *Trans. Faraday Soc.*, 48:166, 1952.
- [27] P. Lessard S. Julien, J.P. Babary. Theoretical and practical identifiability of a reduced order model in an activated sludge process doing nitrification and denitrification. *Water Sci. Technol.*, 37(12), 1998.
- [28] R.J. LeVeque. *Numerical Methods for Conservation Laws*. Birkhäuser Basel, 1990.
- [29] D. Levy, G. Puppo, and G. Russo. Compact central WENO schemes for multidimensional conservation laws. *SIAM J. Sci. Comput.*, 22(2):656–672, 2000.
- [30] B. Li and M.K. Stenstrom. Research advances and challenges in one-dimensional modeling of secondary settling Tanks – A critical review. *Water Res.*, 65:40–63, 2014.
- [31] J. Li, C.-W. Shu, and J. Qiu. Multi-resolution HWENO schemes for hyperbolic conservation laws. *J. Comput. Phys.*, 446:110653, 2021.
- [32] I. Markelov, R.-M. Couture, R. Fischer, S. Haande, and P. Van Cappellen. Coupling Water Column and Sediment Biogeochemical Dynamics: Modeling Internal Phosphorus Loading, Climate Change Responses, and Mitigation Measures in Lake Vansjø, Norway. *J. Geophys. Res. Biogeosci.*, 124(12):3847–3866, 2019.
- [33] B.-J. Ni and H.-Q. Yu. An approach for modeling two-step denitrification in activated sludge systems. *Chem. Eng. Sci.*, 63(6):1449–1459, 2008.
- [34] S. Osher X.-D. Liu and T. Chan. Weighted essentially non-oscillatory schemes. *J. Comput. Phys.*, 115(1):200–212, 1994.
- [35] J. Pauer. Nitrification in the water column and sediment of a hypereutrophic lake and adjoining river system. *Water Res.*, 34(4):1247–1254, 2000.
- [36] C.-W. Shu. *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*, pages 325–432. Springer Berlin Heidelberg, 1998.
- [37] C.-W. Shu. High order weighted essentially nonoscillatory schemes for convection dominated problems. *SIAM Review*, 51(1):82–126, 2009.
- [38] C.-W. Shu and S. Osher. Efficient implementation of essentially non-oscillatory shock-capturing schemes. *J. Comput. Phys.*, 77(2):439–471, 1988.
- [39] E. Tipping, J.F. Boyle, D.N. Schillereff, B.M. Spears, and G. Phillips. Macronutrient processing by temperate lakes: A dynamic model for long-term, large-scale application. *Sci. Total Environ.*, 572:1573–1585, 2016.

- [40] B. Van Leer. Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method. *J. Comput. Phys.*, 32(1):101–136, 1979.
- [41] A. van Haandel, G. Ekama, and G. Marais. The activated sludge process—3 single sludge denitrification. *Water Res.*, 15(10):1135–1152, 1981.
- [42] Y. Xing and X. Zhang. Positivity-preserving well-balanced discontinuous Galerkin methods for the shallow water equations on unstructured triangular meshes. *J. Sci. Comput.*, 57(1):19–41, 2013.
- [43] Y. Xing, X. Zhang, and C.-W. Shu. Positivity-preserving high order well-balanced discontinuous Galerkin methods for the shallow water equations. *Adv. Water Resour.*, 33(12):1476–1493, 2010.
- [44] X. Zhang and C.-W. Shu. On maximum-principle-satisfying high order schemes for scalar conservation laws. *J. Comput. Phys.*, 229(9):3091–3120, 2010.
- [45] X. Zhang and C.-W. Shu. On positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations on rectangular meshes. *J. Comput. Phys.*, 229(23):8918–8934, 2010.
- [46] X. Zhang and C.-W. Shu. Maximum-principle-satisfying and positivity-preserving high-order schemes for conservation laws: survey and new developments. *Proc. R. Soc. A: Math. Phys. Eng. Sci.*, 467(2134):2752–2776, 2011.
- [47] X. Zhang and C.-W. Shu. Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms. *J. Comput. Phys.*, 230(4):1238–1248, 2011.
- [48] J. Zhu and C.-W. Shu. A new type of multi-resolution WENO schemes with increasingly higher order of accuracy. *J. Comput. Phys.*, 375:659–683, 2018.

Centro de Investigación en Ingeniería Matemática (CI²MA)

PRE-PUBLICACIONES 2026

- 2026-17 TOMÁS BARRIOS, EDWIN BEHRENS, ROMMEL BUSTINZA, JOSE M. CASCON: *A stabilized displacement - stress formulation for a linear elasticity problema with mixed boundary conditions*
- 2026-18 GONZALO A. BENAVIDES, SERGIO CAUCAO, GABRIEL N. GATICA, YURI D. SOBRAL: *New Banach spaces-based mixed finite element methods for steady-state flows of magnetic fluids*
- 2026-19 FAHIM ASLAM, MARCELO CAVALCANTI, ZAYD HAJJEJ, YOU HAIYANG LIN BO, MAURICIO SEPÚLVEDA: *Blow-up of a Lamé system with fractional damping and infinite memory: theoretical and numerical study*
- 2026-20 GABRIEL N. GATICA, SALIM MEDDAHI: *A hybridizable discontinuous Galerkin method for the elasto-acoustic transmission problem*
- 2026-21 IDULFO ARROCHA, RAIMUND BÜRGER, MARÍA GÓMEZ, YOLANDA VÁSQUEZ: *Modeling and numerical simulation of the colonization of plant growth-promoting bacteria in the rice rhizosphere*
- 2026-22 ABRAHAM J. ARENAS, JUAN BARAJAS-CALONGE, GILBERTO GONZÁLEZ-PARRA, LUIS M. VILLADA: *Convergence analysis of a nonstandard finite volume method for a diffusion-production-destruction model*
- 2026-23 JUAN BARAJAS-CALONGE, RAIMUND BÜRGER, STEFAN DIEHL, LUIS M. VILLADA: *A network of conservation laws with discontinuous fluxes and source terms modeling countercurrent decantation*
- 2026-24 RAIMUND BÜRGER, BAKHTIYOR KHUZHAYOROV, USMONALI SAYDULLAEV, IRODA BEKNAZAROVA: *Numerical simulation of filtration of a suspension through a radial filter with given flow rate*
- 2026-25 FERNANDO MUÑOZ GARCÍA, FERNANDO ROLDÁN: *Forward-reflected-backward algorithm with linesearch*
- 2026-26 EIDER ALDANA, RICARDO OYARZÚA: *A Lagrange multiplier approach for fluid-flow problem with inhomogeneous mixed boundary conditions*
- 2026-27 JEAN-CHRISTOPHE PESQUET, FERNANDO ROLDÁN: *Nonlinear forward-backward algorithm for solving non-monotone+Lipschitz inclusions with applications to adjoint mismatch problems*
- 2026-28 JUAN BARAJAS-CALONGE, JULIO CAREAGA, LUIS M. VILLADA: *Invariant-region-preserving high-order schemes for a model of reactive sedimentation*

Para obtener copias de las Pre-Publicaciones, escribir o llamar a: DIRECTOR, CENTRO DE INVESTIGACIÓN EN INGENIERÍA MATEMÁTICA, UNIVERSIDAD DE CONCEPCIÓN, CASILLA 160-C, CONCEPCIÓN, CHILE, TEL.: 41-2661324, o bien, visitar la página web del centro: <http://www.ci2ma.udec.cl>



**CENTRO DE INVESTIGACIÓN EN
INGENIERÍA MATEMÁTICA (CI²MA)
Universidad de Concepción**



Casilla 160-C, Concepción, Chile
Tel.: 56-41-2661324/2661554/2661316
<http://www.ci2ma.udec.cl>

