# UNIVERSIDAD DE CONCEPCIÓN

## Centro de Investigación en Ingeniería Matemática (CI$^2$MA)


**A second-order invariant-region-preserving scheme for a transport-flow model of polydisperse sedimentation**

Juan Barajas-Calonge, Raimund Bürger,
Pep Mulet, Luis M. Villada

PREPRINT 2025-16

SERIE DE PRE-PUBLICACIONES

# A SECOND-ORDER INVARIANT-REGION-PRESERVING SCHEME FOR A TRANSPORT-FLOW MODEL OF POLYDISPERSE SEDIMENTATION*

JUAN BARAJAS-CALONGE†, RAIMUND BÜRGER‡, PEP MULET§, AND LUIS MIGUEL VILLADA†

**Abstract.** A polydisperse suspension is a mixture of a number $N$ of species of small solid particles, which may differ in size or density, dispersed in viscous fluid. The sedimentation of such a mixture gives rise to the segregation of species and flow of the mixture due to density fluctuations. In two space dimensions, and for equal-density particles, this process can be described by a hyperbolic system of $N$ nonlinear conservation laws for the particle volume fractions coupled with a version of the Stokes system for the volume-averaged flow field of the mixture. A second-order numerical scheme for this transport-flow model is formulated by combining a finite-difference approximation of the Stokes system with a finite volume (FV) scheme for the transport equations, both defined on a Cartesian grid on a rectangular domain. The FV scheme is based on a central weighted essentially non-oscillatory (CWENO) reconstruction [M. J. Castro and M. Semplice, Int. J. Numer. Methods Fluids, 89 (2019), pp. 304–325] applied to the first-order local Lax-Friedrichs (LLF) numerical flux. By the application of scaling limiters to the CWENO reconstruction polynomials (following [X. Zhang and C.-W. Shu, J. Comput. Phys., 229 (2010), pp. 3091–3120]) and utilizing that the Stokes solver generates a discretely divergence-free (DDF) velocity field, one can prove that the FV scheme has the invariant region preserving (IRP) property, i.e., the volume fractions are nonnegative and sum up at most to a set maximum value. Numerical examples illustrate the model and the scheme.

**Key words.** Coupled transport-flow problem, system of conservation laws, Stokes system, sedimentation, CWENO reconstruction, discretely divergence-free velocity, invariant region preservation.

**MSC codes.** 76T20, 65M06

## 1. Introduction.

**1.1. Scope.** We are interested in high-order finite volume schemes for coupled transport-flow models of the following form, where $\Phi := (\phi_1, \dots, \phi_N)^{\mathrm{T}}$:

$$\partial_t \phi_l + \nabla \cdot \big(\phi_l \boldsymbol{q} + f_l(\Phi)\boldsymbol{k}\big) = 0, \quad f_l(\Phi) = \phi_l v_l(\Phi), \quad l = 1, \dots, N; \tag{1.1a}$$

$$-\nabla \cdot \big(\mu(\phi)\boldsymbol{\varepsilon}(\boldsymbol{q})\big) + \nabla p = \boldsymbol{g}(\phi), \quad \boldsymbol{\varepsilon}(\boldsymbol{q}) := \frac{1}{2}(\nabla \boldsymbol{q} + (\nabla \boldsymbol{q})^{\mathrm{T}}), \tag{1.1b}$$

$$\nabla \cdot \boldsymbol{q} = 0, \tag{1.1c}$$

where $\phi := \phi_1 + \cdots + \phi_N$, posed on a rectangular domain $\Omega := (x_a, x_b) \times (y_a, y_b)$ for $t > 0$, along with the initial and boundary conditions

$$\Phi(\boldsymbol{x}, 0) = \Phi_0(\boldsymbol{x}) \quad \text{for all } \boldsymbol{x} \in \Omega, \text{ where } \Phi_0 := (\phi_{1,0}, \dots, \phi_{N,0})^{\mathrm{T}}, \tag{1.2}$$

$$\boldsymbol{q} = \boldsymbol{0} \quad \text{and} \quad (f_l(\Phi)\boldsymbol{k}) \cdot \boldsymbol{n} = 0, \quad l = 1, \dots, N, \quad \text{on } \partial\Omega, \tag{1.3}$$

$$\int_\Omega p(\boldsymbol{x}, t)\, \mathrm{d}\boldsymbol{x} = 0 \quad \text{for all } t \geq 0. \tag{1.4}$$

*Submitted to the editors DATE.

†GIMNAP-Departamento de Matemáticas, Universidad del Bío-Bío, Casilla 5-C, Concepción, Chile (juan.barajas2001@alumnos.ubiobio.cl, lvillada@ubiobio.cl).

‡CI²MA and Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Casilla 160-C, Concepción, Chile (rburger@ing-mat.udec.cl).

§Departament de Matemàtiques, Universitat de València, Av. Vicent Andrés Estellés s/n, E-46100 Burjassot, Spain (mulet@uv.es).

The system (1.1) arises as a model of sedimentation of a polydisperse suspension of $N$ species of solid particles with diameters $d_1 \geq \cdots \geq d_N$ and densities $\rho_1, \ldots, \rho_N$ dispersed in a viscous fluid. Here $\phi_l$ is the sought volume fraction of species $l$ (having diameter $d_l$ and density $\rho_l$) as a function of spatial position $\boldsymbol{x}$ and time $t$. Other unknowns are the volume-averaged velocity of the mixture $\boldsymbol{q} = \boldsymbol{q}(\boldsymbol{x}, t)$ and the pressure $p = p(\boldsymbol{x}, t)$. The vector $\boldsymbol{k}$ is the downward-pointing unit vector, and $v_1(\Phi), \ldots, v_N(\Phi)$ are given velocity functions that describe segregation, that is, the local movement of particle species $l$ with respect to the mixture. A common choice of these functions or equivalently, of the flux vector $\boldsymbol{f}(\Phi) \coloneqq (f_1(\Phi), \ldots, f_N(\Phi))^{\mathrm{T}}$, is the Masliyah-Lockett-Bassoon (MLB) model of polydisperse sedimentation [26,27]. Furthermore, we assume that the viscosity $\mu = \mu(\phi)$ of the mixture is a function of the total solids fraction $\phi$, and $\boldsymbol{g}(\phi)$ is a term describing the density fluctuations of the mixture. (The model ingredients will be stated precisely in Section 2.) Notice that the system of conservation laws (1.1a) is a transport equation for the solid species and (1.1b), (1.1c) is a version of the Stokes system describing the motion of the mixture. The transport and the flow equations are strongly coupled since $\boldsymbol{q}$ arises as a transport velocity in (1.1a), and at the same time, the functions $\mu$ and $\boldsymbol{g}$ arising in (1.1b) depend on $\phi$.

It is the purpose of this contribution to introduce a numerical scheme for the approximation of (1.1)–(1.4) that is second-order accurate in space and time, and that has the so-called invariant region preservation (IRP) property, which means that if the discretized initial datum $\Phi_0$ belongs to the set of physically relevant values

$$\mathcal{D} \coloneqq \left\{(\phi_1, \ldots, \phi_N)^{\mathrm{T}} \in \mathbb{R}^N \,:\, \phi_1 \geq 0, \ldots, \phi_N \geq 0, \phi \coloneqq \phi_1 + \cdots + \phi_N \leq \phi_{\max}\right\},$$

then the discrete solution vectors $\Phi$ take values in $\mathcal{D}$ at all times. The scheme is defined on a Cartesian grid on $\Omega$ and is based on alternating between a finite-difference (FD) discretization for (1.1b), (1.1c) and a finite volume (FV) scheme for (1.1a). The FV scheme employs the first-order local Lax-Friedrichs (LLF) numerical flux combined with a central weighted essentially non-oscillatory (CWENO) reconstruction [10]. The IRP property is proven directly for the LLF scheme and remains valid for the CWENO reconstruction through the application of scaling limiters (akin to those introduced in [34, 35]). The main novelty is the application of positivity-preserving and bound-preserving high-order discretizations to a coupled transport-flow problem where the transport part is a nonlinear system of conservation laws of arbitrary size. Moreover, to the authors' knowledge this is the first time the MLB model is directly used for a two-dimensional simulation of polydisperse suspension flow and segregation.

**1.2. Related work.** Weighted essentially non-oscillatory (WENO) reconstructions (cf., e.g., [23, 24, 31]), including the particular CWENO procedure [10] used herein, lead to schemes that sharply resolve discontinuities of solutions of systems of conservation laws but in some cases produce unphysical solution values, such as negative densities. In fact, for the model (1.1) or its one-dimensional version (discussed below), WENO-based reconstructions usually fail to preserve the invariant region $\mathcal{D}$ (see [7, 18] or Examples 2–4 in Section 5). This general shortcoming of FV schemes with WENO-based reconstructions, as well as other high-order discretizations of conservation laws such as discontinuous Galerkin (DG) schemes, have led to the construction of so-called property-preserving numerical schemes for conservation laws (see [25] for an overview). Specifically, the method proposed herein is based on the approach by Zhang and Shu [34, 35], who applied linear scaling limiters to WENO reconstruction polynomials to ensure the IRP property in conjunction with Legendre Gauss-Lobatto quadrature to express cell averages in terms of bound-preserving intermediate states

(cf. [25, Remark 2.46].) These techniques were subsequently extended to compressible flow problems including high-order DG schemes for compressible Euler equations with source terms [36], high-order FV Hermite WENO schemes for compressible Euler equations [9], high-order DG schemes for compressible Navier-Stokes equations [34], bound- and positivity preserving FV WENO schemes for a five-equation model of two-medium flows [32], and arbitrary Lagrangian-Eulerian (ALE) FV schemes for compressible two-medium flow equations with stiffened gas equation of state [33].

To prove conservation and IRP properties of the FV scheme for (1.1a), we need to ensure that the discretized flow field $\boldsymbol{q}$ satisfies a certain discretely divergence-free (DDF) property. This property is achieved by a marker-and-cell (MAC) arrangement of unknowns and the particular FD approximation of the Stokes system (1.1b), (1.1c). Thus, the respective discretizations of the transport and flow parts should be *compatible* in the sense of [15]. Approaches to handle the intricate relationship between the DDF constraint and the IRP property include a new high-order FV method for ideal magnetohydrodynamics equations [17] and positivity and maximum-principle-preserving DG finite element schemes for coupled flow-transport problems [19].

The MLB model is based on the assumption [26,27] that the motion of one particle of species $i$ is driven by the density difference $\bar{\rho}_i(\Phi) \coloneqq \rho_i - \rho(\Phi)$, where $\rho(\Phi) = \rho_1\phi_1 + \cdots + \rho_N\phi_N + (1-\phi)\rho_\mathrm{f}$ is the local density of the mixture, and $\rho_\mathrm{f}$ is the density of the fluid. For the derivation of the model from mass and linear momentum balances and constitutive assumptions and its hyperbolicity analysis, we refer to [4,6,8,18]. For this model the authors recently advanced WENO methods with an IRP property [1] for the one-dimensional zero-flux initial-boundary value problem

$$\partial_t \Phi + \partial_x \boldsymbol{f}(\Phi) = \boldsymbol{0}, \quad x \in (0,L), \quad t > 0, \tag{1.5}$$

$$\Phi(x, t=0) = \Phi_0(x) \in \mathcal{D}, \quad x \in (0,L); \quad \boldsymbol{f}|_{x=0} = \boldsymbol{0}, \quad \boldsymbol{f}|_{x=L} = \boldsymbol{0}, \quad t > 0. \tag{1.6}$$

Other multispecies kinematic flow models that in one space dimension give rise to systems similar to (1.5), and which can be handled by similar WENO methods with IRP property, include multiclass vehicular traffic [1,3,18] and the separation of liquid-liquid dispersions [30]. On the other hand, two well-known phenomena of sedimentation require description in two space dimensions, and are captured by (1.1)–(1.4). One of them is the so-called "Boycott effect" [5] that refers to the increase of settling rates in a cylinder that is inclined compared with a vertical orientation. The phenomenon is associated with a rapidly upwards-streaming layer of clear liquid beneath an inclined wall. The other one, the so-called "Diehl test" [16], originally devised for identifying the flux function of a scalar conservation law modelling sedimentation, involves an initial condition where suspension is initially located above clear liquid.

**1.3. Outline of the paper.** The remainder of this work is organized as follows. In section 2 we state preliminaries, starting in section 2.1 with the assumptions on the functions $f_l(\Phi) = \phi_l v_l(\Phi)$ that allow us to prove the IRP property. In section 2.2 we show that the velocity functions $v_l(\Phi)$ of the MLB model and equal-density particles satisfy these assumptions. In section 3 we describe the scheme to numerically solve the transport-flow problem (1.1)–(1.3), starting with the (simple) discretization of $\Omega$ by a Cartesian grid. In section 3.1 we discretize the Stokes system (1.1b), (1.1c) by second-order accurate finite differences that yield a linear system with a symmetric and positive definite matrix block, as we prove in section 3.2. (Achieving a higher order of accuracy while maintaining symmetry is not possible with finite differences, due to boundary conditions.) Furthermore we demonstrate that this system for $\boldsymbol{q}$ and $p$ (for given $\Phi$) is uniquely solvable. Next, in section 3.3, we outline the first-

order FV scheme for the solution of (1.1a) over one time step assuming that $\boldsymbol{q}$ is given (by the method of section 3.2). This FV scheme is based on the LLF numerical flux, and we prove that it satisfies the IRP property under a suitable Courant-Friedrichs-Lewy (CFL) condition. Section 4 is devoted to the CWENO reconstruction (based on the one proposed in [10]) of concentration vectors $\Phi$ on the Cartesian grid. We limit ourselves to third-order reconstruction, which is described in section 4.1. The reconstructions define polynomial traces of $\Phi$ along the boundaries of each cell. The numerical fluxes across the cell boundaries depend on integrals of the reconstructions along the corresponding boundary. These integrals are approximated by a suitable Gauss quadrature rule that is introduced in section 4.2 along with another family of Gauss-Lobatto quadrature rules. The construction (which involves the Gauss-Lobatto quadrature) of scaling limiters for the high-order CWENO reconstructions to ensure that these have the IRP property is detailed in section 4.3. In section 4.4 we outline the high-order time integration and state the complete scheme for (1.1)–(1.4) in algorithmic form. In section 5 four numerical examples are presented: an accuracy test with $N = 2$, two scenarios for $N = 2$ including the Boycott effect and the Diehl test, and settling of $N = 4$ species. Some conclusions are collected in section 6.

## 2. Preliminaries.

**2.1. Assumptions.** We assume that there exist a piecewise differentiable function $w = w(\phi)$ which is differentiable in $[0, \phi_{\max}]$ and parameters $\kappa_1 \geq \kappa_2 \geq \cdots \geq \kappa_N > 0$ such that

$$\boldsymbol{v}(\Phi)^{\mathrm{T}}\Phi = w(\phi)\boldsymbol{\kappa}^{\mathrm{T}}\Phi, \quad \text{where } \boldsymbol{v}(\Phi) \coloneqq (v_1(\Phi), \ldots, v_N(\Phi))^{\mathrm{T}} \text{ and } \boldsymbol{\kappa} \coloneqq (\kappa_1, \ldots, \kappa_N)^{\mathrm{T}},$$

where $w(\phi) \geq 0$ for all $\phi \in [0, \phi_{\max}]$, $w(\phi_{\max}) = 0$, and $w'(\phi) \leq w'(\tilde{\phi})$ if $0 \leq \phi \leq \tilde{\phi} \leq \phi_{\max}$. Note that this implies that $w'(\phi) \leq 0$ for all $\phi \in (0, \phi_{\max})$. Furthermore, we assume that the system (1.5) is hyperbolic on $\mathcal{D}$. Let $\lambda_1(\Phi) \geq \cdots \geq \lambda_N(\Phi)$ denote the eigenvalues of the Jacobian matrix $\boldsymbol{\mathcal{J}}(\Phi)$ of $\boldsymbol{f}(\Phi)$ at $\Phi \in \mathcal{D}$. It is assumed there exists a piecewise continuous function $\psi = \psi(\phi)$ such that $\psi(\phi)\boldsymbol{\kappa}^{\mathrm{T}}\Phi \leq \lambda_N(\Phi)$ and $w'(\phi) \geq \psi(\phi)$ for all $\Phi \in \mathcal{D}$, that $v_1(\Phi) \geq 0$ for all $\Phi \in \mathcal{D}$, and that there exists a known upper bound $M_2 = M_2(\Phi)$ such that $M_2(\Phi) \geq \max\{v_1(\Phi), \lambda_1(\Phi)\}$ for all $\Phi \in \mathcal{D}$. Summarizing, we assume that

$$\begin{aligned} M_1(\Phi) \coloneqq \psi(\phi)\boldsymbol{\kappa}^{\mathrm{T}}\Phi \leq \lambda_N(\Phi) \leq \cdots \leq \lambda_1(\Phi) \\ \leq \max\{v_1(\Phi), \lambda_1(\Phi)\} \leq M_2(\Phi) \quad \text{for all } \Phi \in \mathcal{D}. \end{aligned} \tag{2.1}$$

**2.2. The MLB model of polydisperse sedimentation.** For particles of diameters $d_1 \geq \cdots \geq d_N$, different densities, parameters $\delta_l \coloneqq d_l^2/d_1^2$, $\boldsymbol{\delta} \coloneqq (\delta_1 = 1, \delta_2, \ldots, \delta_N)^{\mathrm{T}}$, and $\tilde{\mu}_{\mathrm{sed}} \coloneqq g d_1^2/(18\mu_{\mathrm{f}})$, where $g$ is the acceleration of gravity and $\mu_{\mathrm{f}}$ is the viscosity of the pure fluid, the MLB model states that

$$v_l(\Phi) = \tilde{\mu}_{\mathrm{sed}} V(\phi)\big(\delta_l\bar{\rho}_l(\Phi) - (\delta_1\phi_1\bar{\rho}_1(\Phi) + \cdots + \delta_N\phi_N\bar{\rho}_N(\Phi))\big), \quad l = 1, \ldots, N. \tag{2.2}$$

Here $V(\phi)$ is a hindrance factor that is assumed to satisfy $V(0) = 1$, $V'(\phi) \leq 0$, and $V(\phi_{\max}) = 0$. Since for different-density particles ($\rho_i \neq \rho_j$ for at least one choice of $i \neq j$), the system (1.5) with $v_l(\Phi)$ given by (2.2) is in general not hyperbolic on $\mathcal{D}$ [8], we assume equal solid densities, i.e., $\rho_1 = \cdots = \rho_N =: \rho_{\mathrm{s}}$. Then (2.2) reduces to

$$v_l(\Phi) = \mu_{\mathrm{sed}}(1-\phi)V(\phi)(\delta_l - \boldsymbol{\delta}^{\mathrm{T}}\Phi), \quad l = 1, \ldots, N; \quad \mu_{\mathrm{sed}} \coloneqq (\rho_{\mathrm{s}} - \rho_{\mathrm{f}})\tilde{\mu}_{\mathrm{sed}}. \tag{2.3}$$

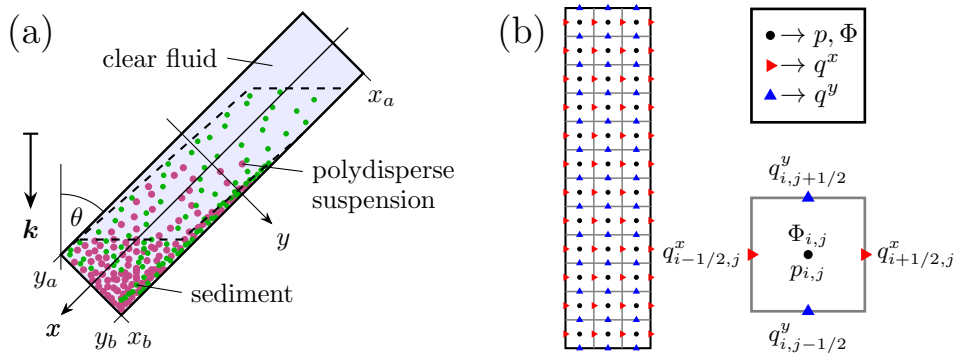


FIG. 1. *(a) Settling of a polydisperse suspension in an inclined vessel at some time $t > 0$ (magenta and green dots represent large and small particles, respectively), (b) approximation of $\Phi$, $p$, and $\boldsymbol{q}$ on the cell boundaries and cell center according to the marker-and-cell (MAC) strategy.*

One can prove that if $v_l(\Phi)$, $l = 1, \dots, N$ are given by (2.3), then (1.5) is strictly hyperbolic on $\mathcal{D}$ (see [1,6,18]), and the eigenvalues $\lambda_1, \dots, \lambda_N$ and velocities $v_1, \dots, v_N$ (all depending on $\Phi$) satisfy the so-called interlacing property; specifically,

$$\begin{aligned}\tilde{M}_1(\Phi) &:= \mu_{\text{sed}}\big(\delta_N V(\phi) + ((1-\phi)V'(\phi) - 2V(\phi))\boldsymbol{\delta}^{\mathrm{T}}\Phi\big)\\ &< \lambda_N < v_N < \lambda_{N-1} < \cdots < \lambda_1 < v_1 =: M_2(\Phi) \quad \text{for all } \Phi \in \mathcal{D}.\end{aligned}$$

Notice that for $\rho_{\text{s}} > \rho_{\text{f}}$, there holds $v_1(\Phi) \geq 0$ for all $\Phi \in \mathcal{D}$.

A standard choice of $V(\phi)$ is the modified Richardson-Zaki function [29]

$$V(\phi) := \begin{cases} \tilde{V}(\phi) := (1-\phi)^{n_{\text{RZ}}-2} & \text{for } 0 < \phi < \phi_*,\\ \tilde{V}(\phi_*) + \tilde{V}'(\phi_*)(\phi - \phi_*) & \text{for } \phi_* \leq \phi \leq \phi_{\max}, \quad n_{\text{RZ}} > 3.\\ 0 & \text{otherwise,}\end{cases}$$

Here $\tau(\phi) := \tilde{V}(\phi_*) + \tilde{V}'(\phi_*)(\phi - \phi_*)$ is the tangent to $\tilde{V}(\phi)$ at $(\phi_*, \tilde{V}(\phi_*))$, where $\phi_*$ is chosen such that $\tau(\phi_{\max}) = 0$, i.e., $\phi_* = ((n_{\text{RZ}} - 2)\phi_{\max} - 1)/(n_{\text{RZ}} - 3)$.

It can be proven that with this function, the MLB model with equal-density particles satisfies the assumptions of section 2.1 for $w(\phi) := \mu_{\text{sed}}(1-\phi)^2 V(\phi)$, $\boldsymbol{\kappa} := \boldsymbol{\delta}$, and $\psi(\phi) := \mu_{\text{sed}}((1-\phi)V'(\phi) - 2V(\phi))$. Note that $w'(\phi) = (1-\phi)\psi(\phi)$.

**3. Numerical scheme.** We discretize $\Omega$ by $km$ control volumes of size $h_x h_y$. We define the intervals $I_i^x := [x_{i-1/2}, x_{i+1/2}]$ and $I_j^y := [y_{j-1/2}, y_{j+1/2}]$ and the cells $K_{i,j} := I_i^x \times I_j^y$, where $x_{i+1/2} := x_a + ih_x$ for $i = 0, \dots, k$ and $y_{j+1/2} := y_a + jh_y$ for $j = 0, \dots, m$. The concentration $\Phi$ and the pressure $p$ are approximated in the center of each cell while the velocity $\boldsymbol{q}$ is discretized on a staggered grid by a marker-and-cell (MAC) strategy, see [21] and Figure 1(b). Consequently, and partly motivated by the Dirichlet boundary conditions for $\boldsymbol{q}$ and the absence of boundary conditions for $p$,

$$\begin{aligned}\Phi_{i,j}(t) &\approx \frac{1}{h_x h_y}\int_{K_{i,j}} \Phi(\boldsymbol{x}, t)\,\mathrm{d}\boldsymbol{x}, \quad p_{i,j} \approx p(x_i, y_j), \quad i = 1, \dots, k, \quad j = 1, \dots, m,\\ q^x_{i+1/2,j} &\approx q^x(x_{i+1/2}, y_j), \quad i = 1, \dots, k-1, \quad j = 1, \dots, m,\\ q^y_{i,j+1/2} &\approx q^y(x_i, y_{j+1/2}), \quad i = 1, \dots, k, \quad j = 1, \dots, m-1.\end{aligned}$$

For later use we introduce the following difference operators for expressions of the type $\omega(a^1_{i,j}, a^2_{i,j}, \dots, a^\nu_{i,j})$, where in each line either the upper or the lower sign holds:

$$\Delta^x_\pm \omega(a^1_{i,j}, a^2_{i,j}, \dots, a^\nu_{i,j}) \coloneqq \mp\big(\omega(a^1_{i,j}, a^2_{i,j}, \dots, a^\nu_{i,j}) - \omega(a^1_{i\pm1,j}, a^2_{i\pm1,j}, \dots, a^\nu_{i\pm1,j})\big),$$
$$\Delta^y_\pm \omega(a^1_{i,j}, a^2_{i,j}, \dots, a^\nu_{i,j}) \coloneqq \mp\big(\omega(a^1_{i,j}, a^2_{i,j}, \dots, a^\nu_{i,j}) - \omega(a^1_{i,j\pm1}, a^2_{i,j\pm1}, \dots, a^\nu_{i,j\pm1})\big).$$

**3.1. Finite-difference approximation of the Stokes system.** The system (1.1b), (1.1c) for the velocity $\boldsymbol{q} = (q^x, q^y)^{\mathrm{T}}$ and the pressure $p$ has the structure of a Stokes problem. By writing out in detail the partial derivatives in $\nabla \cdot (\mu \nabla \boldsymbol{q})$ and $\nabla \cdot (\mu (\nabla \boldsymbol{q})^{\mathrm{T}})$ and setting $\boldsymbol{g}(\phi) = (g_1(\phi), g_2(\phi))^{\mathrm{T}}$ we may rewrite (1.1b), (1.1c) as

$$-(\mu q^x_x)_x - \tfrac{1}{2}(\mu q^y_x)_y - \tfrac{1}{2}(\mu q^x_y)_y + p_x = g_1(\phi), \tag{3.1}$$
$$-(\mu q^y_y)_y - \tfrac{1}{2}(\mu q^y_x)_x - \tfrac{1}{2}(\mu q^x_y)_x + p_y = g_2(\phi), \tag{3.2}$$
$$-q^x_x - q^y_y = 0, \tag{3.3}$$

We henceforth write $u \coloneqq q^x, v \coloneqq q^y$ for simplicity of notation. Equations (3.1), (3.2), and (3.3) will be evaluated at $(x_{i+1/2}, y_j)$, $(x_i, y_{j+1/2})$, and $(x_i, y_j)$, respectively.

The pressure gradient $\nabla p$ is approximated by centered finite differences; these do not involve boundary conditions:

$$p_x(x_{i+1/2}, y_j) \approx \Delta^x_+ p_{i,j}/h_x, \quad i = 1, \dots, k-1, \quad j = 1, \dots, m, \tag{3.4}$$
$$p_y(x_i, y_{j+1/2}) \approx \Delta^y_+ p_{i,j}/h_y, \quad i = 1, \dots, k, \quad j = 1, \dots, m-1. \tag{3.5}$$

To describe these and other approximations in usable form, we define the $k \times m$ matrix $[p] \coloneqq (p_{i,j})$ (of approximate values of $p$), the $(k-1) \times m$ matrix $[p_x]$ and the $k \times (m-1)$ matrix $[p_y]$ of the approximate values of $p_x$ and $p_y$, respectively. Analogous notation will be used for other quantities. Thus, we may cast (3.4) and (3.5) in matrix form as

$$[p_x] = -h_x^{-1} \boldsymbol{D}_k^{\mathrm{T}} [p], \quad [p_y] = -h_y^{-1} [p] \boldsymbol{D}_m; \tag{3.6}$$

here and for later use we define the $n \times (n-1)$ difference matrices

$$\boldsymbol{D}_n \coloneqq \begin{bmatrix} 1 & 0 & \dots & 0 \\ -1 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 1 \\ 0 & \cdots & 0 & -1 \end{bmatrix}, \quad \boldsymbol{D}_n^* \coloneqq \begin{bmatrix} 2 & 0 & \cdots & 0 \\ -1 & 1 & \ddots & \vdots \\ 0 & \ddots & \ddots & 0 \\ \vdots & \ddots & -1 & 1 \\ 0 & \cdots & 0 & -2 \end{bmatrix}.$$

We approximate $(\mu u_x)_x(x_{i+1/2}, y_j) \approx h_x^{-2} \Delta^x_+ (\mu_{i,j} \Delta^x_- u_{i+1/2,j})$ for $i = 1, \dots, k-1$ and $j = 1, \dots, m$, where $\mu_{i,j} \coloneqq \mu(\phi^n_{i,j})$. If we define the $(k-1) \times m$ matrix $[u] = (u_{i+1/2,j})$ and impose the boundary conditions by $u_{1/2,j} = u_{k+1/2,j} = 0$, then the approximation of $(\mu u_x)_x$ can be expressed by the $(k-1) \times m$ matrix

$$[(\mu u_x)_x] = -h_x^{-2} \boldsymbol{D}_k^{\mathrm{T}} \big([\mu] * (\boldsymbol{D}_k [u])\big).$$

(We recall that for matrices $\boldsymbol{A} = (A_{i,j})$ and $\boldsymbol{B} = (B_{i,j})$ of the same size, the Hadamard product $\boldsymbol{A} * \boldsymbol{B}$ is defined by $(A * B)_{i,j} \coloneqq A_{i,j} B_{i,j}$.) Moreover, we approximate $(\mu u_y)_y(x_{i+1/2}, y_j) \approx h_y^{-2} \Delta^y_- (\mu_{i+1/2,j+1/2} \Delta^y_+ u_{i+1/2,j})$ for $i = 1, \dots, k-1$ and $j = 1, \dots, m$. The boundary conditions $u_{i+1/2,1/2} = u_{i+1/2,m+1/2} = 0$ are implemented by setting $u_{i+1/2,0} = -u_{i+1/2,1}$ and $u_{i+1/2,m+1} = -u_{i+1/2,m+1}$. Next, we define

$$\mu_{i+1/2,j-1/2} \coloneqq \begin{cases} \operatorname{avg}(\mu_{i+1,j}, \mu_{i,j}, \mu_{i+1,j-1}, \mu_{i,j-1}), & j = 2, \dots, m, \\ \operatorname{avg}(\frac{3}{2}\mu_{i+1,1} - \frac{1}{2}\mu_{i+1,2}, \frac{3}{2}\mu_{i,1} - \frac{1}{2}\mu_{i,2}), & j = 1, \\ \operatorname{avg}(\frac{3}{2}\mu_{i+1,m} - \frac{1}{2}\mu_{i+1,m-1}, \frac{3}{2}\mu_{i,m} - \frac{1}{2}\mu_{i,m-1}), & j = m+1 \end{cases}$$

for $i = 1, \dots, k-1$ and $j = 1, \dots, m+1$ and the corresponding $(k-1) \times (m+1)$ matrix $[\bar{\mu}] \coloneqq (\mu_{i+1/2,j-1/2})$, where the average function avg can be linear or harmonic. We may now represent the approximation of $(\mu u_y)_y$ by the $(k-1) \times m$ matrix

$$[(\mu u_y)_y] = -h_y^{-2}\big([\bar{\mu}] * ([u](\boldsymbol{D}_{m+1}^*)^{\mathrm{T}})\big)\boldsymbol{D}_{m+1}.$$

This approximation is first-order accurate for $j = 1$ and $j = m$. The term $(\mu v_x)_y$ is approximated by $(\mu v_x)_y(x_{i+1/2}, y_j) \approx h_x^{-1}h_y^{-1}\Delta_-^y(\mu_{i+1/2,j+1/2}\Delta_-^x v_{i+1,j+1/2})$ for $i = 1, \dots, k-1$ and $j = 1, \dots, m$. Analogously to $[u]$, we define the $k \times (m-1)$ matrix $[v] = (v_{i,j+1/2})$. The latter approximation can be written as a $(k-1) \times m$ matrix

$$[(\mu v_x)_y] = -h_x^{-1}h_y^{-1}\big([\bar{\mu}](:, 2:m) * (\boldsymbol{D}_k^{\mathrm{T}}[v])\big)\boldsymbol{D}_m^{\mathrm{T}},$$

where $[\bar{\mu}](:, 2:m)$ denotes the $(k-1) \times (m-1)$ matrix formed by columns 2 to $m$ of $[\bar{\mu}]$. Next, we approximate $(\mu v_x)_x(x_i, y_{j+1/2}) \approx h_x^{-2}\Delta_-^x(\mu_{i+1/2,j+1/2}\Delta_-^x v_{i+1,j+1/2})$ for $i = 1, \dots, k$ and $j = 1, \dots, m-1$. To express this in matrix form, we define

$$\bar{\bar{\mu}}_{i-1/2,j+1/2} \coloneqq \begin{cases} \operatorname{avg}(\mu_{i,j+1}, \mu_{i,j}, \mu_{i-1,j+1}, \mu_{i-1,j}), & 2 \le i \le k, \\ \operatorname{avg}(\frac{3}{2}\mu_{1,j+1} - \frac{1}{2}\mu_{2,j+1}, \frac{3}{2}\mu_{1,j} - \frac{1}{2}\mu_{2,j}), & i = 1, \\ \operatorname{avg}(\frac{3}{2}\mu_{k,j+1} - \frac{1}{2}\mu_{k-1,j+1}, \frac{3}{2}\mu_{k,j} - \frac{1}{2}\mu_{k-1,j}), & i = k+1 \end{cases}$$

for $1 \le i \le k+1$ and $1 \le j \le m-1$ along with the $(k+1) \times (m-1)$ matrix $[\bar{\bar{\mu}}] \coloneqq (\tilde{\mu}_{i-1/2,j+1/2})$. The approximation of $(\mu v_x)_x$ is given by the $k \times (m-1)$ matrix

$$[(\mu v_x)_x] = -h_x^{-2}\boldsymbol{D}_{k+1}^{\mathrm{T}}\big([\bar{\bar{\mu}}] * (\boldsymbol{D}_{k+1}^*[v])\big).$$

(This approximation is first-order accurate for $i = 1$ and $i = k$.) In addition, we approximate the term $(\mu v_y)_y$ by $(\mu v_y)_y(x_i, y_{j+1/2}) \approx h_y^{-2}\Delta_-^y(\mu_{i,j+1}\Delta_+^y v_{i,j+1/2}))$ for $i = 1, \dots, k$ and $j = 1, \dots, m-1$, which is expressed as a $k \times (m-1)$ matrix:

$$[(\mu v_y)_y] = -h_y^{-2}\big([\mu] * ([v]\boldsymbol{D}_m^{\mathrm{T}})\big)\boldsymbol{D}_m.$$

Similarly, $(\mu u_y)_x(x_i, y_{j+1/2}) \approx h_x^{-1}h_y^{-1}\Delta_-^x(\mu_{i+1/2,j+1/2}\Delta_-^y u_{i+1/2,j+1})$ for $i = 1, \dots, k$ and $j = 1, \dots, m-1$, which can be written as the $k \times (m-1)$ matrix

$$[(\mu u_y)_x] = -h_x^{-1}h_y^{-1}\boldsymbol{D}_k\big([\bar{\mu}](2:k, :) * ([u]\boldsymbol{D}_m)\big)$$

(notation is similar to that of $[(\mu v_x)_y]$). Finally, we approximate $\nabla \cdot \boldsymbol{q}$ by the sum of $u_x(x_i, y_j) \approx h_x^{-1}\Delta_-^x u_{i+1/2,j}$ and $v_y(x_i, y_j) \approx h_y^{-1}\Delta_-^y v_{i,j+1/2}$ for $i = 1, \dots, k$ and $j = 1, \dots, m$. Thus, we obtain the $k \times m$ matrices

$$[u_x] = h_x^{-1}\boldsymbol{D}_k[u], \quad [v_y] = h_y^{-1}[v]\boldsymbol{D}_m^{\mathrm{T}}.$$

**3.2. Linear system for the Stokes problem.** We first recall that for a $p \times q$ matrix $\boldsymbol{A} = (a_{i,j})$ and an $r \times s$ matrix $\boldsymbol{B}$, the Kronecker product $\boldsymbol{A} \otimes \boldsymbol{B}$ is

$$\boldsymbol{A} \otimes \boldsymbol{B} \coloneqq \begin{bmatrix} a_{1,1}\boldsymbol{B} & \cdots & a_{1,q}\boldsymbol{B} \\ \vdots & \ddots & \vdots \\ a_{p,1}\boldsymbol{B} & \cdots & a_{p,q}\boldsymbol{B} \end{bmatrix} \quad \text{(an } rp \times sq \text{ block matrix)}.$$

If $\operatorname{col}(\boldsymbol{M})$ denotes the column vector obtained by stacking each column of a matrix $\boldsymbol{M}$ and $\operatorname{diag}(\boldsymbol{v})$ is the diagonal matrix of the entries of a vector $\boldsymbol{v}$, then $\operatorname{col}(\boldsymbol{B}\boldsymbol{M}\boldsymbol{A}^{\mathrm{T}}) =$

$(\boldsymbol{A}\otimes\boldsymbol{B})\,\mathrm{col}(\boldsymbol{M})$ and $\mathrm{col}(\boldsymbol{B}*\boldsymbol{M}) = \mathrm{diag}(\mathrm{col}(\boldsymbol{B}))\,\mathrm{col}(\boldsymbol{M})$. To identify the linear system to be solved, we now use these identities to convert the formulas for $[p_x]$, $[p_y]$, $[(\mu u_x)_x]$ etc., which depend linearly on $[u]$, $[v]$, or $[p]$, into expressions for $\mathrm{col}([p_x])$, $\mathrm{col}([p_y])$, $\mathrm{col}([(\mu u_x)_x])$ etc. that depend linearly on $\mathrm{col}([u])$, $\mathrm{col}([v])$ and $\mathrm{col}([p])$. For simplicity, we write $\mathrm{col}(p_x)$, $\mathrm{col}(p_y)$, $\mathrm{col}((\mu u_x)_x)$, instead of $\mathrm{col}([p_x])$, $\mathrm{col}([p_y])$, $\mathrm{col}([(\mu u_x)_x])$, etc.

In what follows, $\boldsymbol{I}_n$ denotes the $n\times n$ identity matrix. Clearly, (3.6) implies

$$\mathrm{col}(p_x) = -h_x^{-1}(\boldsymbol{I}_m\otimes\boldsymbol{D}_k^{\mathrm{T}})\,\mathrm{col}(p),\quad \mathrm{col}(p_y) = -h_y^{-1}(\boldsymbol{D}_m^{\mathrm{T}}\otimes\boldsymbol{I}_k)\,\mathrm{col}(p)$$

for the gradient of pressure $\nabla p$. For the terms $\nabla\cdot(\mu(\phi)\nabla\boldsymbol{q})$ and $\nabla\cdot\boldsymbol{q}$ we get

$$\begin{aligned}
\mathrm{col}((\mu u_x)_x) &= -h_x^{-2}(\boldsymbol{I}_m\otimes\boldsymbol{D}_k^{\mathrm{T}})\,\mathrm{diag}(\mathrm{col}(\mu))(\boldsymbol{I}_m\otimes\boldsymbol{D}_k)\,\mathrm{col}(u),\\
\mathrm{col}((\mu u_y)_y) &= -h_y^{-2}(\boldsymbol{D}_{m+1}^{\mathrm{T}}\otimes\boldsymbol{I}_{k-1})\,\mathrm{diag}(\mathrm{col}(\bar\mu))(\boldsymbol{D}_{m+1}^{*}\otimes\boldsymbol{I}_{k-1})\,\mathrm{col}(u),\\
\mathrm{col}((\mu v_x)_y) &= -h_x^{-1}h_y^{-1}(\boldsymbol{D}_m\otimes\boldsymbol{I}_{k-1})\,\mathrm{diag}\big(\mathrm{col}(\bar\mu(:,2:m))\big)(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_k^{\mathrm{T}})\,\mathrm{col}(v),\\
\mathrm{col}((\mu v_x)_x) &= -h_x^{-2}(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_{k+1}^{\mathrm{T}})\,\mathrm{diag}(\mathrm{col}(\bar{\bar\mu}))(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_{k+1}^{*})\,\mathrm{col}(v),\\
\mathrm{col}((\mu v_y)_y) &= -h_y^{-2}(\boldsymbol{D}_m^{\mathrm{T}}\otimes\boldsymbol{I}_k)(\mathrm{diag}(\mathrm{col}(\mu))(\boldsymbol{D}_m\otimes\boldsymbol{I}_k)\,\mathrm{col}(v),\\
\mathrm{col}((\mu u_y)_x) &= -h_x^{-1}h_y^{-1}(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_k)\,\mathrm{diag}(\mathrm{col}(\bar{\bar\mu}(2:k,:)))(\boldsymbol{D}_m^{\mathrm{T}}\otimes\boldsymbol{I}_{k-1})\,\mathrm{col}(u),\\
\mathrm{col}(u_x) &= h_x^{-1}(\boldsymbol{I}_m\otimes\boldsymbol{D}_k)\,\mathrm{col}(u),\quad \mathrm{col}(v_y) = h_y^{-1}(\boldsymbol{D}_m\otimes\boldsymbol{I}_k)\,\mathrm{col}(v).
\end{aligned}$$

This allows us to formulate the FD approximation of (3.1)-(3.3) as a linear system

$$\begin{bmatrix}\boldsymbol{A}^{x,x} & \boldsymbol{A}^{x,y} & \boldsymbol{B}^{x}\\ \boldsymbol{A}^{y,x} & \boldsymbol{A}^{y,y} & \boldsymbol{B}^{y}\\ (\boldsymbol{B}^{x})^{\mathrm{T}} & (\boldsymbol{B}^{y})^{\mathrm{T}} & \boldsymbol{0}\end{bmatrix}\begin{pmatrix}\mathrm{col}(u)\\ \mathrm{col}(v)\\ \mathrm{col}(p)\end{pmatrix} = \begin{pmatrix}\mathrm{col}(G_1(\phi))\\ \mathrm{col}(G_2(\phi))\\ \boldsymbol{0}\end{pmatrix}, \tag{3.7}$$

where the blocks of the matrix are given by

$$\begin{aligned}
\boldsymbol{A}^{x,x} &= h_x^{-2}(\boldsymbol{I}_m\otimes\boldsymbol{D}_k^{\mathrm{T}})\,\mathrm{diag}(\mathrm{col}(\mu))(\boldsymbol{I}_m\otimes\boldsymbol{D}_k)\\
&\quad + \tfrac{1}{2}h_y^{-2}(\boldsymbol{D}_{m+1}^{\mathrm{T}}\otimes\boldsymbol{I}_{k-1})\,\mathrm{diag}(\mathrm{col}(\bar\mu))(\boldsymbol{D}_{m+1}^{*}\otimes\boldsymbol{I}_{k-1}),\\
\boldsymbol{A}^{x,y} &= \tfrac{1}{2}h_x^{-1}h_y^{-1}(\boldsymbol{D}_m\otimes\boldsymbol{I}_{k-1})\,\mathrm{diag}(\mathrm{col}(\bar\mu(:,2:m)))(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_k^{\mathrm{T}}),\\
\boldsymbol{A}^{y,x} &= \tfrac{1}{2}h_x^{-1}h_y^{-1}(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_k)\,\mathrm{diag}(\mathrm{col}(\bar{\bar\mu}(2:k,:)))(\boldsymbol{D}_m^{\mathrm{T}}\otimes\boldsymbol{I}_{k-1}),\\
\boldsymbol{A}^{y,y} &= h_y^{-2}(\boldsymbol{D}_m^{\mathrm{T}}\otimes\boldsymbol{I}_k)\,\mathrm{diag}(\mathrm{col}(\mu))(\boldsymbol{D}_m\otimes\boldsymbol{I}_k)\\
&\quad + \tfrac{1}{2}h_x^{-2}(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_{k+1}^{\mathrm{T}})\,\mathrm{diag}(\mathrm{col}(\bar{\bar\mu}))(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_{k+1}^{*}),\\
\boldsymbol{B}^{x} &= -h_x^{-1}(\boldsymbol{I}_m\otimes\boldsymbol{D}_k^{\mathrm{T}}),\quad \boldsymbol{B}^{y} = -h_y^{-1}(\boldsymbol{D}_m^{\mathrm{T}}\otimes\boldsymbol{I}_k).
\end{aligned}$$

The right-hand side is defined via $[G_1(\phi)] = (G_1(\phi)_{i,j})$ for $i = 1,\dots,k-1$ and $j = 1,\dots,m$ and $[G_2(\phi)] = (G_2(\phi)_{i,j})$ for $i = 1,\dots,k$ and $j = 1,\dots,m-1$, where

$$\begin{aligned}
g_1(\phi)(x_{i+1/2},y_j) &\approx G_1(\phi)_{i,j} \coloneqq g_1\big((\phi_{i,j}+\phi_{i+1,j})/2\big),\\
g_2(\phi)(x_i,y_{j+1/2}) &\approx G_2(\phi)_{i,j} \coloneqq g_2\big((\phi_{i,j}+\phi_{i,j+1})/2\big).
\end{aligned}$$

Let us write (3.7) as

$$\boldsymbol{M}\begin{pmatrix}\boldsymbol{c_q}\\ \boldsymbol{c}_p\end{pmatrix} = \begin{pmatrix}\boldsymbol{g}\\ \boldsymbol{0}\end{pmatrix},\quad \boldsymbol{M}\coloneqq\begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B}^{\mathrm{T}} & \boldsymbol{0}\end{bmatrix},\quad \boldsymbol{A}\coloneqq\begin{bmatrix}\boldsymbol{A}^{x,x} & \boldsymbol{A}^{x,y}\\ \boldsymbol{A}^{y,x} & \boldsymbol{A}^{y,y}\end{bmatrix},\ \boldsymbol{B}\coloneqq\begin{bmatrix}\boldsymbol{B}^{x}\\ \boldsymbol{B}^{y}\end{bmatrix}, \tag{3.8}$$

where $\boldsymbol{c_q}\coloneqq(\mathrm{col}(u)^{\mathrm{T}},\mathrm{col}(v)^{\mathrm{T}})^{\mathrm{T}}$, $\boldsymbol{c}_p\coloneqq\mathrm{col}(p)$, and $\boldsymbol{g}\coloneqq(\mathrm{col}(G_1(\phi))^{\mathrm{T}},\mathrm{col}(G_2(\phi))^{\mathrm{T}})^{\mathrm{T}}$.

LEMMA 1. *The matrix $\boldsymbol{A}$ is symmetric and positive definite.*

*Proof.* We have that $\boldsymbol{A} = \boldsymbol{P} + \boldsymbol{Q} + \frac{1}{2}\boldsymbol{R}$, where $\boldsymbol{P}$ is a $2\times 2$ block diagonal matrix:

$$\begin{aligned}\boldsymbol{P} = \operatorname{blockdiag}\big(&\tfrac{1}{2}h_y^{-2}(\boldsymbol{D}_{m+1}^{\mathrm{T}}\otimes\boldsymbol{I}_{k-1})\operatorname{diag}(\operatorname{col}(\bar\mu))((\boldsymbol{D}_{m+1}^{*}-\boldsymbol{D}_{m+1})\otimes\boldsymbol{I}_{k-1}),\\ &\tfrac{1}{2}h_x^{-2}(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_{k+1}^{\mathrm{T}})\operatorname{diag}(\operatorname{col}(\bar{\bar\mu}))(\boldsymbol{I}_{m-1}\otimes(\boldsymbol{D}_{k+1}^{*}-\boldsymbol{D}_{k+1}))\big),\end{aligned}$$

which turns out to be diagonal with nonnegative entries, since $\boldsymbol{D}_{\ell+1}^{\mathrm{T}}\operatorname{diag}(\boldsymbol{a})(\boldsymbol{D}_{\ell+1}^{*}-\boldsymbol{D}_{\ell+1}) = \operatorname{diag}(a_1,0,\dots,0,a_{\ell+1})$ for all $\boldsymbol{a}=(a_1,\dots,a_{\ell+1})^{\mathrm{T}}$; moreover,

$$\begin{aligned}\boldsymbol{Q} = \operatorname{blockdiag}\big(&h_x^{-2}(\boldsymbol{I}_m\otimes\boldsymbol{D}_k^{\mathrm{T}})\operatorname{diag}(\operatorname{col}(\mu))(\boldsymbol{I}_m\otimes\boldsymbol{D}_k),\\ &h_y^{-2}(\boldsymbol{D}_m^{\mathrm{T}}\otimes\boldsymbol{I}_k)\operatorname{diag}(\operatorname{col}(\mu))(\boldsymbol{D}_m\otimes\boldsymbol{I}_k)\big),\end{aligned}$$

where both blocks are symmetric and positive definite, since $\mu_{i,j}>0$ and the rank of $\boldsymbol{I}_m\otimes\boldsymbol{D}_k$ and $\boldsymbol{D}_m\otimes\boldsymbol{I}_k$ is $mk$, i.e., the columns of both blocks are linearly independent; and $\boldsymbol{R}$ is a $2\times 2$ block matrix with blocks

$$\begin{aligned}\boldsymbol{R}_{11} &= h_y^{-2}(\boldsymbol{D}_{m+1}^{\mathrm{T}}\otimes\boldsymbol{I}_{k-1})\operatorname{diag}(\operatorname{col}(\bar\mu))(\boldsymbol{D}_{m+1}\otimes\boldsymbol{I}_{k-1}),\\ \boldsymbol{R}_{21}^{\mathrm{T}} = \boldsymbol{R}_{12} &= h_x^{-1}h_y^{-1}(\boldsymbol{D}_m\otimes\boldsymbol{I}_{k-1})\operatorname{diag}(\operatorname{col}(\bar\mu(:,2:m)))(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_k^{\mathrm{T}}),\\ \boldsymbol{R}_{22} &= h_x^{-2}(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_{k+1}^{\mathrm{T}})\operatorname{diag}(\operatorname{col}(\bar{\bar\mu}))(\boldsymbol{I}_{m-1}\otimes\boldsymbol{D}_{k+1}).\end{aligned}$$

Since $[\tilde\mu] \coloneqq [\bar{\bar\mu}](2:k,:) = [\bar\mu](:,2:m)$, the matrix $\boldsymbol{R}$ is symmetric. Moreover, since $\boldsymbol{D}_\ell^{\mathrm{T}} = -\boldsymbol{D}_{\ell+1}(2:\ell,:)$ for $\ell = k,m$, there holds

$$\boldsymbol{R}_{12} = \big(h_y^{-1}\boldsymbol{D}_{m+1}^{\mathrm{T}}\otimes\boldsymbol{I}_{k-1}\big)\tilde{\boldsymbol{L}}\big(\boldsymbol{I}_{m-1}\otimes h_x^{-1}\boldsymbol{D}_{k+1}\big)$$

for the $(k-1)(m+1)\times(k+1)(m-1)$ matrix $\tilde{\boldsymbol{L}}$ given by

$$\tilde L_{(i-1)(k-1)+j,(p-1)(k+1)+q} \coloneqq \begin{cases}\tilde\mu_{j,i-1}, & 2\le i\le m,\ 2\le q\le k,\ i-1=p,\ j=q-1,\\ 0 & \text{otherwise,}\end{cases}$$

$$i=1,\dots,m+1,\quad j=1,k-1,\quad p=1,\dots,m-1,\quad q=1,\dots,k+1.$$

We have $\boldsymbol{R} = \boldsymbol{S}^{\mathrm{T}}\tilde{\boldsymbol{R}}\boldsymbol{S}$ with $\boldsymbol{S} \coloneqq \operatorname{blockdiag}(h_y^{-1}\boldsymbol{D}_{m+1}\otimes\boldsymbol{I}_{k-1}, \boldsymbol{I}_{m-1}\otimes h_x^{-1}\boldsymbol{D}_{k+1}))$ and

$$\tilde{\boldsymbol{R}} \coloneqq \begin{bmatrix}\operatorname{diag}(\operatorname{col}(\bar\mu)) & \tilde{\boldsymbol{L}}\\ \tilde{\boldsymbol{L}}^{\mathrm{T}} & \operatorname{diag}(\operatorname{col}(\bar{\bar\mu}))\end{bmatrix}.$$

The matrix $\tilde{\boldsymbol{R}}$ is positive semidefinite since for $\boldsymbol{C}\in\mathbb{R}^{(k-1)\times(m+1)}$, $\boldsymbol{D}\in\mathbb{R}^{(k+1)\times(m-1)}$,

$$\begin{aligned}&\big(\operatorname{col}(\boldsymbol{C})^{\mathrm{T}},\operatorname{col}(\boldsymbol{D})^{\mathrm{T}}\big)\tilde{\boldsymbol{R}}\begin{pmatrix}\operatorname{col}(C)\\ \operatorname{col}(D)\end{pmatrix}\\ &= \sum_{j=1}^{k-1}\sum_{i\in\{1,m+1\}}\bar\mu_{j,i}C_{j,i}^2 + \sum_{i=1}^{m-1}\sum_{j\in\{1,k+1\}}\bar{\bar\mu}_{ji}D_{ji}^2 + \sum_{j=1}^{k-1}\sum_{i=1}^{m-1}\tilde\mu_{ji}(C_{j,i+1}+D_{i+1,j})^2 \ge 0.\end{aligned}$$

Consequently, as the sum of two symmetric semipositive definite matrices and one positive definite matrix, the matrix $\boldsymbol{A}$ is symmetric and positive definite. ∎

To see that (3.8) is solvable for any $\boldsymbol{g}$, we consider the homogeneous system

$$\boldsymbol{M}\begin{pmatrix}\boldsymbol{c_q}\\ \boldsymbol{c}_p\end{pmatrix} = \begin{bmatrix}\boldsymbol{A} & \boldsymbol{B}\\ \boldsymbol{B}^{\mathrm{T}} & \boldsymbol{0}\end{bmatrix}\begin{pmatrix}\boldsymbol{c_q}\\ \boldsymbol{c}_p\end{pmatrix} = \begin{pmatrix}\boldsymbol{0}\\ \boldsymbol{0}\end{pmatrix}.$$

The first block of equations yields $\boldsymbol{c_q} = -\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{c}_p$. Inserting this into the second block we get $\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{c}_p = \boldsymbol{0}$, hence $\boldsymbol{c}_p^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}}\boldsymbol{A}^{-1}\boldsymbol{B}\boldsymbol{c}_p = 0$, which holds exactly for those $\boldsymbol{c}_p$ such that $\boldsymbol{B}\boldsymbol{c}_p = \boldsymbol{0}$, which by the particular structure of $\boldsymbol{B}$ holds if and only if $\boldsymbol{c}_p = \alpha\boldsymbol{e}$, where $\alpha \in \mathbb{R}$ and $\boldsymbol{e} \coloneqq (1, \dots, 1)^{\mathrm{T}}$. For these vectors $\boldsymbol{c}_p$ we get $\boldsymbol{c_q} = \boldsymbol{0}$. Consequently, the image of $\boldsymbol{M}$, $\mathrm{im}(\boldsymbol{M})$, is given by $\mathrm{im}(\boldsymbol{M}) = \ker(\boldsymbol{M}^{\mathrm{T}})^{\perp} = \ker(\boldsymbol{M})^{\perp} = \{\alpha(\boldsymbol{0}^{\mathrm{T}}, \boldsymbol{e}^{\mathrm{T}})^{\mathrm{T}} : \alpha \in \mathbb{R}\}^{\perp}$. However, for any vector $\boldsymbol{g}$, there holds $(\boldsymbol{g}^{\mathrm{T}}, \boldsymbol{0}^{\mathrm{T}})(\boldsymbol{0}^{\mathrm{T}}, \boldsymbol{e}^{\mathrm{T}})^{\mathrm{T}} = 0$, and therefore $(\boldsymbol{g}^{\mathrm{T}}, \boldsymbol{0}^{\mathrm{T}})^{\mathrm{T}} \in \mathrm{im}(\boldsymbol{M})$, so that the Stokes system is compatible. Since

$$(\boldsymbol{0}^{\mathrm{T}}, \boldsymbol{e}^{\mathrm{T}}) \begin{bmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{B}^{\mathrm{T}} & \boldsymbol{0} \end{bmatrix} = \begin{bmatrix} \boldsymbol{e}^{\mathrm{T}}\boldsymbol{B}^{\mathrm{T}} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix} = \boldsymbol{0},$$

one can remove any of the equations in the last block of (3.7), since it depends linearly on the rest, and set the corresponding pressure (component of $\boldsymbol{c}_p$) to zero, and the resulting systems has only one solution.

**3.3. Approximation of the concentration equations.** To discretize the concentration equations (1.1a) we can view the system in vector form

$$\partial_t \Phi + \partial_x \boldsymbol{F}^x(q^x, \Phi) + \partial_y \boldsymbol{F}^y(q^y, \Phi) = \boldsymbol{0}, \tag{3.9}$$

where $\boldsymbol{F}^x(q^x, \Phi) = q^x\Phi + k^x\boldsymbol{f}(\Phi)$ and $\boldsymbol{F}^y(q^y, \Phi) = q^y\Phi + k^y\boldsymbol{f}(\Phi)$, where $\boldsymbol{k} = (k^x, k^y)^{\mathrm{T}}$. We denote by $\Phi^n_{i,j} = \Phi_{i,j}(t_n)$ the average of $\Phi$ on a control volume $K_{i,j}$ at time $t_n = n\tau$, $n \in \mathbb{N}$, where $\tau > 0$ is a time step specified later. For the computation of the numerical flux we use the FD approximation of $\boldsymbol{q}$ obtained from solving (3.7). This discretization implies the discretely divergence-free (DDF) property

$$h_x^{-1}\big(q^{x,n}_{i+1/2,j} - q^{x,n}_{i-1/2,j}\big) + h_y^{-1}\big(q^{y,n}_{i,j+1/2} - q^{y,n}_{i,j-1/2}\big) = 0, \tag{3.10}$$

to be used in the proof of Theorem 4. For an explicit approximation in time we get

$$\begin{aligned}
&\int_{t_n}^{t_{n+1}} \int_{\partial K_{i,j}} \mathsf{F}(\boldsymbol{q}, \Phi) \cdot \boldsymbol{n}_{K_{i,j}} \mathrm{d}S \\
&\approx \tau\big(h_y \boldsymbol{\mathcal{H}}(q^{x,n}_{i+1/2,j}, \Phi^n_{i,j}, \Phi^n_{i+1,j}, \boldsymbol{e}_1) + h_x \boldsymbol{\mathcal{H}}(q^{y,n}_{i,j+1/2}, \Phi^n_{i,j}, \Phi^n_{i,j+1}, \boldsymbol{e}_2) \\
&\qquad + h_y \boldsymbol{\mathcal{H}}(q^{x,n}_{i-1/2,j}, \Phi^n_{i,j}, \Phi^n_{i-1,j}, -\boldsymbol{e}_1) + h_x \boldsymbol{\mathcal{H}}(q^{y,n}_{i,j-1/2}, \Phi^n_{i,j}, \Phi^n_{i,j-1}, -\boldsymbol{e}_2)\big),
\end{aligned}$$

where $\mathsf{F} \coloneqq (\boldsymbol{F}^x, \boldsymbol{F}^y)^{\mathrm{T}}$, $\boldsymbol{e}_1$ and $\boldsymbol{e}_2$ are the unit vectors in $\mathbb{R}^2$, and $\boldsymbol{\mathcal{H}}$ is some numerical flux function associated to the projection of system (3.9) in the direction $\pm\boldsymbol{e}_i$, $i = 1, 2$. We follow [11, sect. 3] to construct this projection and formally define $[\nabla\Phi]_i \coloneqq (\partial_1\Phi_i, \partial_2\Phi_i)^{\mathrm{T}}$. For a generic vector $\boldsymbol{\eta} = (\eta_1, \eta_2)^{\mathrm{T}} \in \{\pm\boldsymbol{e}_1, \pm\boldsymbol{e}_2\}$ and denoting $\boldsymbol{\eta}^{\perp} \coloneqq (-\eta_2, \eta_1)^{\mathrm{T}}$, let us assume that the solution $\Phi(\boldsymbol{x}, t)$ of (3.9) does not vary in the direction $\boldsymbol{\eta}^{\perp}$, that is, $\nabla\Phi \cdot \boldsymbol{\eta}^{\perp} = 0$. If we define $\boldsymbol{\xi}(z, t) \coloneqq \Phi(\boldsymbol{x}(z), t)$ with $\boldsymbol{x}(z) \coloneqq z\boldsymbol{\eta} + \boldsymbol{x}_0$ for some $\boldsymbol{x}_0 \in \mathbb{R}^2$, then $\partial_z\boldsymbol{\xi}(z, t) = \nabla\Phi(\boldsymbol{x}(z), t) \cdot \boldsymbol{\eta}$ and

$$\begin{aligned}
&\partial_z\big(\boldsymbol{F}^x(q^x, \boldsymbol{\xi})\eta_1 + \boldsymbol{F}^y(q^y, \boldsymbol{\xi})\eta_2\big)(z, t) \\
&= \sum_{i=1}^{2} \big(\partial_1\boldsymbol{F}^x(q^x, \Phi)\eta_1 + \partial_i\boldsymbol{F}^y(q^y, \Phi)\eta_2\big)[\nabla\Phi]_i\big(\boldsymbol{x}(z), t\big) \cdot \boldsymbol{\eta} \\
&= \big(\partial_x\boldsymbol{F}^x(q^x, \Phi) + \partial_y\boldsymbol{F}^y(q^y, \Phi)\big)\big(\boldsymbol{x}(z), t\big),
\end{aligned}$$

hence $\boldsymbol{\xi}(z, t)$ is a solution of the one-dimensional system of conservation laws

$$\partial_t\boldsymbol{\xi} + \partial_z\big(\mathsf{F}(\boldsymbol{q}, \boldsymbol{\xi}) \cdot \boldsymbol{\eta}\big) = 0, \quad \text{where} \quad \mathsf{F}(\boldsymbol{q}, \boldsymbol{\xi}) \cdot \boldsymbol{\eta} = (\boldsymbol{q} \cdot \boldsymbol{\eta})\boldsymbol{\xi} + (\boldsymbol{k} \cdot \boldsymbol{\eta})\boldsymbol{f}(\boldsymbol{\xi}). \tag{3.11}$$

Let us analyze the eigenvalues $\tilde{\lambda}_i(\boldsymbol{q},\boldsymbol{\xi},\boldsymbol{\eta})$ of the Jacobian matrix $\tilde{\boldsymbol{\mathcal{J}}}(\boldsymbol{q},\boldsymbol{\xi},\boldsymbol{\eta})$ associated to (3.11). Clearly, $\tilde{\boldsymbol{\mathcal{J}}}(\boldsymbol{q},\boldsymbol{\xi},\boldsymbol{\eta}) = (\boldsymbol{q}\cdot\boldsymbol{\eta})\boldsymbol{I} + (\boldsymbol{k}\cdot\boldsymbol{\eta})\boldsymbol{\mathcal{J}}(\boldsymbol{\xi})$, where $\boldsymbol{\mathcal{J}}(\boldsymbol{\xi})$ is the Jacobian matrix associated with (1.5). If $\lambda_i(\boldsymbol{\xi})$, $i=1,\dots,N$, are the eigenvalues of $\boldsymbol{\mathcal{J}}(\boldsymbol{\xi})$, then

$$\tilde{\lambda}_i(\boldsymbol{q},\boldsymbol{\xi},\boldsymbol{\eta}) = \boldsymbol{q}\cdot\boldsymbol{\eta} + (\boldsymbol{k}\cdot\boldsymbol{\eta})\lambda_i(\boldsymbol{\xi}), \quad i=1,\dots,N.$$

According to (2.1), $\tilde{M}_1(\boldsymbol{q}\cdot\boldsymbol{\eta},\boldsymbol{\xi},\boldsymbol{\eta}) < \tilde{\lambda}_i < \tilde{M}_2(\boldsymbol{q}\cdot\boldsymbol{\eta},\boldsymbol{\xi},\boldsymbol{\eta})$ for all $i=1,\dots,N$, where

$$\tilde{M}_p(q,\boldsymbol{\xi},\boldsymbol{\eta}) \coloneqq \begin{cases} q + (\boldsymbol{k}\cdot\boldsymbol{\eta})M_p(\boldsymbol{\xi}) & \text{if } \boldsymbol{k}\cdot\boldsymbol{\eta} > 0, \\ q + (\boldsymbol{k}\cdot\boldsymbol{\eta})M_{3-p}(\boldsymbol{\xi}) & \text{if } \boldsymbol{k}\cdot\boldsymbol{\eta} \le 0, \end{cases} \quad p=1,2. \tag{3.12}$$

Combining all approximations yields the marching formula

$$\Phi_{i,j}^{n+1} = \Phi_{i,j}^n - \lambda_x\Delta_-^x\boldsymbol{\mathcal{H}}^x\big(q_{i+1/2,j}^{x,n},\Phi_{i,j}^n,\Phi_{i+1,j}^n\big) - \lambda_y\Delta_-^y\boldsymbol{\mathcal{H}}^y\big(q_{i,j+1/2}^{y,n},\Phi_{i,j}^n,\Phi_{i,j+1}^n\big) \tag{3.13}$$

for the concentration vectors $\Phi$, where $\lambda_x \coloneqq \tau/h_x$, $\lambda_y \coloneqq \tau/h_y$, and

$$\boldsymbol{\mathcal{H}}^\sigma(q,\boldsymbol{\xi},\boldsymbol{\nu}) = \tfrac{1}{2}\big(\boldsymbol{F}^\sigma(q,\boldsymbol{\xi}) + \boldsymbol{F}^\sigma(q,\boldsymbol{\nu})\big) - \tfrac{1}{2}\alpha^\sigma(q,\boldsymbol{\xi},\boldsymbol{\nu})(\boldsymbol{\nu}-\boldsymbol{\xi}), \quad \sigma = x,y, \tag{3.14}$$

are the three-point local Lax-Friedrichs (LLF) numerical fluxes, where we define

$$\begin{aligned}\alpha^\sigma(q,\boldsymbol{\xi},\boldsymbol{\nu}) \coloneqq \max\Big\{&\Big|\min_{0\le z\le 1}\tilde{M}_1\big(q, z\boldsymbol{\xi}+(1-z)\boldsymbol{\nu},\eta^\sigma\big)\Big|,\\ &\Big|\max_{0\le z\le 1}\tilde{M}_2\big(q, z\boldsymbol{\xi}+(1-z)\boldsymbol{\nu},\eta^\sigma\big)\Big|\Big\}, \quad \sigma = x,y,\end{aligned} \tag{3.15}$$

for $\boldsymbol{\eta}^x = \boldsymbol{e}_1$ and $\boldsymbol{\eta}^y = \boldsymbol{e}_2$.

The time step $\tau = \tau_n$ is computed after each iteration from the CFL condition

$$\alpha_n(\lambda_x^n + \lambda_y^n) \le 1, \quad \alpha_n = \max_{i,j}\{\alpha_{i\pm1/2,j}^n, \alpha_{i,j\pm1/2}^n\}. \tag{3.16}$$

with $\alpha_{i\pm1/2,j}^n = \alpha^x(q_{i\pm1/2,j}^{x,n},\Phi_{i,j}^n,\Phi_{i\pm1,j}^n)$ and $\alpha_{i,j\pm1/2}^n = \alpha^y(q_{i,j\pm1/2}^{y,n},\Phi_{i,j}^n,\Phi_{i,j\pm1}^n)$.

*Remark* 2. Using the DDF property (3.10) we get from (3.13)

$$\begin{aligned}\Phi_{i,j}^{n+1} &= \Phi_{i,j}^n - \lambda_x\Delta_-^x\boldsymbol{\mathcal{F}}^x\big(q_{i+1/2,j}^{x,n},\Phi_{i,j}^n,\Phi_{i+1,j}^n\big) - \tfrac{1}{2}\lambda_x\big(q_{i+1/2,j}^{x,n}\Phi_{i+1,j}^n - q_{i-1/2,j}^{x,n}\Phi_{i-1,j}^n\big)\\ &\quad - \lambda_y\Delta_-^y\boldsymbol{\mathcal{F}}^y\big(q_{i,j+1/2}^{y,n},\Phi_{i,j}^n,\Phi_{i,j+1}^n\big) - \tfrac{1}{2}\lambda_y\big(q_{i,j+1/2}^{y,n}\Phi_{i,j+1}^n - q_{i,j-1/2}^{y,n}\Phi_{i,j-1}^n\big)\\ &\quad + \tfrac{1}{2}\tau\big(h_x^{-1}\Delta_-^x q_{i+1/2,j}^{x,n} + h_y^{-1}\Delta_-^y q_{i,j+1/2}^{y,n}\big)\Phi_{i,j}^n\\ &= \Phi_{i,j}^n - \lambda_x\Delta_-^x\boldsymbol{\mathcal{F}}^x\big(q_{i+1/2,j}^{x,n},\Phi_{i,j}^n,\Phi_{i+1,j}^n\big) - \tfrac{1}{2}\lambda_x\big(q_{i+1/2,j}^{x,n}\Phi_{i+1,j}^n - q_{i-1/2,j}^{x,n}\Phi_{i-1,j}^n\big)\\ &\quad - \lambda_y\Delta_-^y\boldsymbol{\mathcal{F}}^y\big(q_{i,j+1/2}^{y,n},\Phi_{i,j}^n,\Phi_{i,j+1}^n\big) - \tfrac{1}{2}\lambda_y\big(q_{i,j+1/2}^{y,n}\Phi_{i,j+1}^n - q_{i,j-1/2}^{y,n}\Phi_{i,j-1}^n\big),\end{aligned}$$

where we define the numerical fluxes

$$\boldsymbol{\mathcal{F}}^\sigma(q,\Phi_{\mathrm{L}},\Phi_{\mathrm{R}}) \coloneqq \tfrac{1}{2}k^\sigma\big(\boldsymbol{f}(\Phi_{\mathrm{L}}) + \boldsymbol{f}(\Phi_{\mathrm{R}})\big) - \tfrac{1}{2}\alpha^\sigma(q,\Phi_{\mathrm{L}},\Phi_{\mathrm{R}})(\Phi_{\mathrm{R}}-\Phi_{\mathrm{L}}), \quad \sigma = x,y.$$

This notation allows us to prove the IRP property of the first order scheme and also it will help us in the description of the high-order reconstruction procedure in section 4.

The proof of the IRP property for the LLF scheme relies on the following theorem.

THEOREM 3. *Let us consider the "one-dimensional" scheme, $\sigma = x, y$,*

$$\boldsymbol{\xi}_0^{n+1} = \boldsymbol{\xi}_0^n - \lambda\big(\boldsymbol{\mathcal{F}}^\sigma(q_+, \boldsymbol{\xi}_0^n, \boldsymbol{\xi}_1^n) - \boldsymbol{\mathcal{F}}^\sigma(q_-, \boldsymbol{\xi}_{-1}^n, \boldsymbol{\xi}_0^n)\big) - \tfrac{1}{2}\lambda(q_+\boldsymbol{\xi}_1^n - q_-\boldsymbol{\xi}_{-1}^n), \tag{3.17}$$

*where $\lambda \in \mathbb{R}^+$, $q_-, q_+ \in \mathbb{R}$. If the CFL condition*

$$\lambda \max\big\{\alpha^\sigma(q_+, \boldsymbol{\xi}_0, \boldsymbol{\xi}_1), \alpha^\sigma(q_-, \boldsymbol{\xi}_{-1}, \boldsymbol{\xi}_0)\big\} \le 1 \tag{3.18}$$

*is in effect, $\boldsymbol{\xi}_{-1}, \boldsymbol{\xi}_0, \boldsymbol{\xi}_1 \in \mathcal{D}$ and $\boldsymbol{\xi}_0^{n+1} = (\xi_{1,0}^{n+1}, \dots, \xi_{N,0}^{n+1})^{\mathrm{T}}$, then*

$$\xi_{l,0}^{n+1} \ge 0, \quad l = 1, \dots, N, \tag{3.19}$$

$$\xi_{1,0}^{n+1} + \cdots + \xi_{N,0}^{n+1} \le \phi_{\max} - \tfrac{1}{2}\lambda(q_+ - q_-)\phi_{\max}. \tag{3.20}$$

*Proof.* We omit in this proof the superscript $n$ and let $\xi \coloneqq \xi_1 + \cdots + \xi_N$ and

$$\alpha_\pm \coloneqq \alpha^\sigma(q_\pm, \boldsymbol{\xi}_0, \boldsymbol{\xi}_{\pm 1}), \tag{3.21}$$

where either the upper or the lower sign holds here and in the rest of the proof.

To show that $\xi_{l,0}^{n+1} \ge 0$ for all $l$, we start from (3.17) to get for all $l = 1, \dots, N$

$$\begin{aligned} \xi_{l,0}^{n+1} &= \big(1 - \tfrac{1}{2}\lambda(\alpha_+ + \alpha_-)\big)\xi_{l,0} + \tfrac{1}{2}\lambda\alpha_+ G_l^+(\boldsymbol{\xi}_1) + \tfrac{1}{2}\lambda\alpha_- G_l^-(\boldsymbol{\xi}_{-1}), \\ G_l^\pm(\boldsymbol{\xi}) &\coloneqq \xi_l\big(1 \mp \alpha_\pm^{-1}(q_\pm + k^\sigma v_l(\boldsymbol{\xi}))\big). \end{aligned} \tag{3.22}$$

From (2.1), (3.21) and (3.15) we get that $|q_\pm + k^\sigma v_l(\boldsymbol{\xi}_{\pm 1})| \le \alpha_\pm$ and therefore $G_l^\pm(\boldsymbol{\xi}_{\pm 1}) \ge 0$ from $\xi_{l,j} \ge 0$, $j \in \{-1, 1\}$. We thus obtain (3.19) from $\xi_{l,0} \ge 0$, (3.18), and (3.22).

To prove (3.20), we sum (3.22) over $l = 1, \dots, N$, use that $\boldsymbol{v}(\boldsymbol{\xi})^{\mathrm{T}}\boldsymbol{\xi} = w(\xi)\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}$ (cf. section 2.1) and $\xi_0 \le \phi_{\max}$ to get

$$\begin{aligned} \xi_0^{n+1} &= \big(1 - \tfrac{1}{2}\lambda(\alpha_+ + \alpha_-)\big)\xi_0 + \tfrac{1}{2}\lambda\alpha_+\xi_1 + \tfrac{1}{2}\lambda\alpha_-\xi_{-1} \\ &\quad - \tfrac{1}{2}\lambda\big(q_+\xi_1 + k^\sigma w(\xi_1)\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_1 - q_-\xi_{-1} + k^\sigma w(\xi_{-1})\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{-1}\big) \\ &= \big(1 - \tfrac{1}{2}\lambda(\alpha_+ + \alpha_-)\big)\xi_0 + \tfrac{1}{2}\lambda(\alpha_+ + \alpha_-)\phi_{\max} - \tfrac{1}{2}\lambda(q_+ - q_-)\phi_{\max} \\ &\quad + \tfrac{1}{2}(\mathcal{Y}_+ + \mathcal{Y}_-) \le \phi_{\max} - \tfrac{1}{2}\lambda(q_+ - q_-)\phi_{\max} + \tfrac{1}{2}(\mathcal{Y}_+ + \mathcal{Y}_-), \end{aligned}$$

where

$$\mathcal{Y}_\pm \coloneqq \alpha_\pm(\xi_{\pm 1} - \phi_{\max}) \mp \big(q_\pm(\xi_{\pm 1} - \phi_{\max}) + k^\sigma w(\xi_{\pm 1})\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{\pm 1}\big),$$

so the proof is complete when showing that $\mathcal{Y}_\pm \le 0$.

Since $w(\phi_{\max}) = 0$, we observe that

$$\mathcal{Y}_\pm = \mathcal{W}_\pm(\xi_{\pm 1}) - \mathcal{W}_\pm(\phi_{\max}), \quad \mathcal{W}_\pm(\xi) \coloneqq \alpha_\pm\xi \mp \big(q_\pm\xi + k^\sigma w(\xi)\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{\pm 1}\big).$$

By the mean value theorem, we obtain

$$\mathcal{Y}_\pm = \mathcal{W}'_\pm(\vartheta_{\pm 1})(\xi_{\pm 1} - \phi_{\max}), \quad \mathcal{W}'_\pm(\xi) = \alpha_\pm \mp \big(q_\pm + k^\sigma w'(\xi)\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{\pm 1}\big),$$

for suitable $\vartheta_{\pm 1} \in [\xi_{\pm 1}, \phi_{\max}]$, therefore it suffices to show that $\mathcal{W}'_\pm(\vartheta_{\pm 1}) \ge 0$, since $\xi_{\pm 1} - \phi_{\max} \le 0$. Since, by assumption, $w'$ is increasing in $(0, \phi_{\max})$, $\psi(\xi) \le w'(\xi) \le 0$ for all $\xi \in (0, \phi_{\max})$, $M_2(\boldsymbol{\xi}) \ge 0$ for all $\boldsymbol{\xi} \in \mathcal{D}$, and $\vartheta_\pm \in (\xi_{\pm 1}, \phi_{\max})$, it follows that

$$M_1(\boldsymbol{\xi}_\pm) = \psi(\xi_{\pm 1})\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{\pm 1} \le w'(\xi_{\pm 1})\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{\pm 1} \le w'(\vartheta_\pm)\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{\pm 1} \le 0 \le M_2(\boldsymbol{\xi}_{\pm 1}),$$

therefore $\alpha_\pm = \alpha^\sigma(q_\pm, \boldsymbol{\xi}_0, \boldsymbol{\xi}_{\pm 1})$ and the definition of $\alpha^\sigma$ in (3.15) and (3.12) imply

$$\alpha_\pm \ge \big|q_\pm + k^\sigma w'(\vartheta_\pm)\boldsymbol{\kappa}^{\mathrm{T}}\boldsymbol{\xi}_{\pm 1}\big|,$$

which yields $\mathcal{W}'_\pm(\vartheta_{\pm 1}) \ge 0$. This concludes the proof of Theorem 3. ∎

Now we are in a position to establish the following result.

THEOREM 4. *If the CFL condition* (3.16) *is in effect and the velocity functions* $v_i(\Phi)$ *satisfy the assumptions of section 2.1, then the LLF scheme defined by* (3.13) *with the numerical flux* (3.14) *satisfies the invariant region preservation property*

$$\Phi^n_{i,j} \in \mathcal{D} \text{ for all } K_{i,j} \quad \Rightarrow \quad \Phi^{n+1}_{i,j} \in \mathcal{D} \text{ for all } K_{i,j}; \text{ for all } n = 0, 1, 2, \ldots. \tag{3.23}$$

*Proof.* According to Remark 2, $\Phi^{n+1}_{i,j} = \tilde{\lambda}_x \boldsymbol{H}^x_{i,j} + \tilde{\lambda}_y \boldsymbol{H}^y_{i,j}$ for a given cell $K_{i,j}$, where we define $\lambda := \lambda_x + \lambda_y$, $\tilde{\lambda}_x := \lambda_x/\lambda$, $\tilde{\lambda}_y := \lambda_y/\lambda$, and the vectors

$$\begin{aligned}
\boldsymbol{H}^x_{i,j} &:= \Phi^n_{i,j} - \lambda_x \Delta^x_- \boldsymbol{\mathcal{F}}^x\big(q^{x,n}_{i+1/2,j}, \Phi^n_{i,j}, \Phi^n_{i+1,j}\big) - \tfrac{1}{2}\lambda_x\big(q^{x,n}_{i+1/2,j}\Phi^n_{i+1,j} - q^{x,n}_{i-1/2,j}\Phi^n_{i-1,j}\big),\\
\boldsymbol{H}^y_{i,j} &:= \Phi^n_{i,j} - \lambda_y \Delta^y_- \boldsymbol{\mathcal{F}}^y\big(q^{y,n}_{i,j+1/2}, \Phi^n_{i,j}, \Phi^n_{i,j+1}\big) - \tfrac{1}{2}\lambda_y\big(q^{y,n}_{i,j+1/2}\Phi^n_{i,j+1} - q^{y,n}_{i,j-1/2}\Phi^n_{i,j-1}\big).
\end{aligned}$$

By the CFL condition (3.16), Theorem 3, and denoting $\boldsymbol{H}^\sigma_{i,j} := (H^\sigma_{1,i,j}, \ldots, H^\sigma_{N,i,j})^{\mathrm{T}}$, we get $H^\sigma_{l,i,j} \geq 0$, $\sigma = x, y$, for $l = 1, \ldots, N$, and

$$\sum_{l=1}^N H^x_{l,i,j} \leq \phi_{\max} - \frac{\lambda}{2}\big(\Delta^x_- q^{x,n}_{i+1/2,j}\big)\phi_{\max}, \quad \sum_{l=1}^N H^y_{l,i,j} \leq \phi_{\max} - \frac{\lambda}{2}\big(\Delta^y_- q^{y,n}_{i,j+1/2}\big)\phi_{\max}.$$

These inequalities imply that $\phi^{n+1}_{l,i,j} = \tilde{\lambda}_x H^x_{l,i,j} + \tilde{\lambda}_y H^y_{l,i,j} \geq 0$ for $l = 1, \ldots, N$. Using the DDF property (3.10) we may complete the proof of $\Phi^{n+1}_{i,j} \in \mathcal{D}$ by noticing that

$$\begin{aligned}
&\phi^{n+1}_{1,i,j} + \cdots + \phi^{n+1}_{N,i,j}\\
&= \tilde{\lambda}_x\big(H^x_{1,i,j} + \cdots + H^x_{N,i,j}\big) + \tilde{\lambda}_y\big(H^y_{1,i,j} + \cdots + H^y_{N,i,j}\big)\\
&\leq \tilde{\lambda}_x\phi_{\max} - \tfrac{1}{2}\lambda_x\big(\Delta^x_- q^{x,n}_{i+1/2,j}\big)\phi_{\max} + \tilde{\lambda}_y\phi_{\max} - \tfrac{1}{2}\lambda_y\big(\Delta^y_- q^{y,n}_{i,j+1/2}\big)\phi_{\max}\\
&= \phi_{\max} - \tfrac{1}{2}\tau\big(h_x^{-1}\Delta^x_- q^{x,n}_{i+1/2,j} + h_y^{-1}\Delta^y_- q^{y,n}_{i,j+1/2}\big)\phi_{\max} = \phi_{\max}. \qquad \square
\end{aligned}$$

**4. CWENO-based reconstruction on Cartesian grids.** Since the order of accuracy of the solution of the Stokes problem is two, an order of accuracy of at least two is needed in the transport part. We nevertheless pursue third-order accuracy.

**4.1. Third-order reconstruction.** We assume that the (at most) second-degree so-called optimal polynomial $P_{\mathrm{opt}}(x, y)$ approximates the nine cell averages of the stencil $\mathcal{S}_0 := \Omega_0 \cup \cdots \cup \Omega_8$ (see Figure 2(a)) in the least-squares sense. Then, four first-degree polynomials $P_r(x, y)$, $r = 1, \ldots, 4$, are computed as least-squares approximations of the cell averages over the respective sub-stencils $\mathcal{S}_1$ to $\mathcal{S}_4$. A second-degree central polynomial $P_0(x, y)$ is then defined such that $P_{\mathrm{opt}}(x, y) = \mathtt{d}_0 P_0(x, y) + \mathtt{d}_1 P_1(x, y) + \cdots + \mathtt{d}_4 P_4(x, y)$, where $\mathtt{d}_r > 0$ are the so-called linear weights which satisfy $\mathtt{d}_0 + \cdots + \mathtt{d}_4 = 1$. The final reconstruction polynomial $P_{i,j}(x, y)$ over the cell $K_{i,j}$ is a convex combination $P_{i,j}(x, y) := \omega_0 P_0(x, y) + \omega_1 P_1(x, y) + \cdots + \omega_4 P_4(x, y)$ with the so-called nonlinear weights $\omega_r := \tilde{\omega}_r/(\tilde{\omega}_0 + \cdots + \tilde{\omega}_4)$, where $\tilde{\omega}_r := \mathtt{d}_r/(\mathrm{IS}_r + \varepsilon)^{\mathtt{p}}$, $r = 0, \ldots, 4$. Here $\varepsilon > 0$ is a small parameter, $\mathtt{p} \geq 2$, and $\mathrm{IS}_r$ is a smoothness indicator (see [10] for details on the polynomials and smoothness indicators).

Let $\boldsymbol{P}_{i,j} := (P^{(1)}_{i,j}, \ldots, P^{(N)}_{i,j})^{\mathrm{T}}$ be the reconstruction polynomial vector, where $P^{(l)}_{i,j}$ is the second-degree polynomial for component $l$ of $\Phi^n_{i,j}$ obtained by the CWENO-based procedure described above. In terms of the traces of $\boldsymbol{P}_{i,j}(x, y)$ on the edges

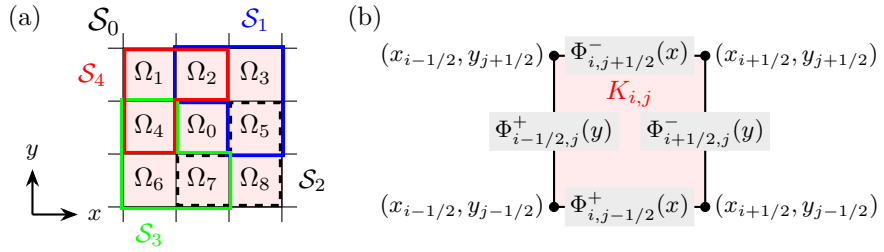


FIG. 2. *(a) Stencil* $\mathcal{S}_0 := \Omega_0 \cup \Omega_1 \cup \cdots \cup \Omega_8$ *and sub-stencils* $\mathcal{S}_1 := \Omega_2 \cup \Omega_3 \cup \Omega_5$, $\mathcal{S}_2 := \Omega_5 \cup \Omega_7 \cup \Omega_8$, $\mathcal{S}_3 := \Omega_4 \cup \Omega_6 \cup \Omega_7$, *and* $\mathcal{S}_4 := \Omega_1 \cup \Omega_2 \cup \Omega_4$, *(b) illustration of the traces* (4.2).

of $K_{i,j}$ (see Figure 2(b)) the FV scheme (3.13) becomes

$$\begin{aligned}\Phi^{n+1}_{i,j} = \Phi^n_{i,j} - \frac{\tau}{h_x h_y}\Bigg( &\int_{I^y_j} \Delta^x_- \boldsymbol{\mathcal{F}}^x\big(q^x_{i+1/2,j}(y), \Phi^-_{i+1/2,j}(y), \Phi^+_{i+1/2,j}(y)\big)\,\mathrm{d}y \\ &+ \frac{1}{2}\int_{I^y_j} \big(q^x_{i+1/2,j}(y)\Phi^+_{i+1/2,j}(y) - q^x_{i-1/2,j}(y)\Phi^-_{i-1/2,j}(y)\big)\,\mathrm{d}y \\ &+ \int_{I^x_i} \Delta^y_- \boldsymbol{\mathcal{F}}^y\big(q^y_{i,j+1/2}(x), \Phi^-_{i,j+1/2}(x), \Phi^+_{i,j+1/2}(x)\big)\,\mathrm{d}x \\ &+ \frac{1}{2}\int_{I^x_i} \big(q^y_{i,j+1/2}(x)\Phi^+_{i,j+1/2}(x) - q^y_{i,j-1/2}(x)\Phi^-_{i,j-1/2}(x)\big)\,\mathrm{d}x\Bigg),\end{aligned} \tag{4.1}$$

where the integrals are to be approximated with sufficient accuracy. The traces are given by

$$\begin{aligned}&\Phi^+_{i-1/2,j}(y) = \boldsymbol{P}_{i,j}(x_{i-1/2}, y), \quad \Phi^-_{i+1/2,j}(y) = \boldsymbol{P}_{i,j}(x_{i+1/2}, y) \quad \text{for } y \in I^y_j, \\ &\Phi^+_{i,j-1/2}(x) = \boldsymbol{P}_{i,j}(x, y_{j-1/2}), \quad \Phi^-_{i,j+1/2}(x) = \boldsymbol{P}_{i,j}(x, y_{j+1/2}) \quad \text{for } x \in I^x_i, \\ &q^x_{i+1/2,j}(y) = q^x_{i+1/2,j}, \quad q^x_{i-1/2,j}(y) = q^x_{i-1/2,j} \quad \text{for } y \in I^y_j, \\ &q^y_{i,j+1/2}(x) = q^y_{i,j+1/2}, \quad q^y_{i,j-1/2}(x) = q^y_{i,j-1/2} \quad \text{for } x \in I^x_i,\end{aligned} \tag{4.2}$$

see Figure 2(b). This approach has been used successfully for a DG discretization (cf., e.g., [35, Th. 4.3]) and for a FV compact WENO scheme [20, sect. 4].

**4.2. Quadrature formulas.** To approximate the integrals in (4.1) with sufficient accuracy, we apply a $G$-point Gauss quadrature rule, which is exact for single variable polynomials of degree $2G-1$. We denote by

$$S^x_i := \{x^\beta_i : \beta = 1, \ldots, G\} \quad \text{and} \quad S^y_j := \{y^\beta_j : \beta = 1, \ldots, G\} \tag{4.3}$$

the Gauss quadrature points (nodes) on $I^x_i$ and $I^y_j$, respectively, and by $w_\beta$ the corresponding weights on the interval $[-\frac{1}{2}, \frac{1}{2}]$, such that their sum equals one. In the subsequent proofs we will also use the $L$-point Gauss–Lobatto quadrature rule, where

$$\hat{S}^x_i := \{\hat{x}^\alpha_i : \alpha = 1, \ldots, L\} \quad \text{and} \quad \hat{S}^y_j := \{\hat{y}^\alpha_j : \alpha = 1, \ldots, L\} \tag{4.4}$$

are the Gauss-Lobatto nodes on $I^x_i$ and $I^y_j$, respectively. Moreover, we use the Cartesian-product notation $S^x_i \times S^y_j := \{(x, y) : x \in S^x_i, y \in S^y_j\}$ to define the set

$$S_{i,j} := (S^x_i \times \hat{S}^y_j) \cup (\hat{S}^x_i \times S^y_j). \tag{4.5}$$

We may now replace (4.1) by the effective scheme

$$
\begin{aligned}
\Phi_{i,j}^{n+1} &= \Phi_{i,j}^{n} - \big(w_1\tilde{\boldsymbol{H}}_{i,j}^{x,1} + \cdots + w_G\tilde{\boldsymbol{H}}_{i,j}^{x,G}\big) - \big(w_1\tilde{\boldsymbol{H}}_{i,j}^{y,1} + \cdots + w_G\tilde{\boldsymbol{H}}_{i,j}^{y,G}\big),\\
\tilde{\boldsymbol{H}}_{i,j}^{x,\beta} &\coloneqq \lambda_x\Delta_-^x\boldsymbol{\mathcal{F}}^x\big(q_{i+1/2,j}^x, \Phi_{i+1/2,j}^-(y_j^\beta), \Phi_{i+1/2,j}^+(y_j^\beta)\big)\\
&\quad + \tfrac{1}{2}\lambda_x\big(q_{i+1/2,j}^x\Phi_{i+1/2,j}^+(y_j^\beta) - q_{i-1/2,j}^x\Phi_{i-1/2,j}^-(y_j^\beta)\big),\\
\tilde{\boldsymbol{H}}_{i,j}^{y,\beta} &\coloneqq \lambda_y\Delta_-^y\boldsymbol{\mathcal{F}}^y\big(q_{i,j+1/2}^y, \Phi_{i,j+1/2}^-(x_i^\beta), \Phi_{i,j+1/2}^+(x_i^\beta)\big)\\
&\quad + \tfrac{1}{2}\lambda_y\big(q_{i,j+1/2}^y\Phi_{i,j+1/2}^+(x_i^\beta) - q_{i,j-1/2}^y\Phi_{i,j-1/2}^-(x_i^\beta)\big), \quad \beta = 1,\dots,G.
\end{aligned}
\tag{4.6}
$$

In section 5 we use $G = 2$ with $S_i^x = \{x_i - (\sqrt{3}/6)h_x, x_i + (\sqrt{3}/6)h_x\}$, $S_j^y = \{y_j - (\sqrt{3}/6)h_y, y_j + (\sqrt{3}/6)h_y\}$ and weights $w_1, w_2 = 1/2$; and $L = 3$ with $\hat{S}_i^x = \{x_i - h_x/2, x_i, x_i + h_x/2\}$, $\hat{S}_j^y = \{y_j - h_y/2, y_j, y_j + h_y/2\}$ and weights $\hat{w}_1, \hat{w}_3 = 1/6$ and $\hat{w}_2 = 2/3$. Both quadrature rules are exact for polynomials of degree 3.

**4.3. Scaling limiters.** We herein slightly modify the Zhang and Shu [34, 35] scaling limiter to handle the equation (1.1). The first step is to limit each concentration $\phi_l$. Let us assume that $\Phi_{i,j}^n \in \mathcal{D}$ and $\tilde{K}_{i,j}$ is a prescribed subset of $K_{i,j}$. We then replace the polynomials $P_{i,j}^{(l)}(x,y)$ by

$$
\begin{aligned}
&\tilde{P}_{i,j}^{(l)}(x,y) \coloneqq \theta_l\big(P_{i,j}^{(l)}(x,y) - \phi_{l,i,j}^n\big) + \phi_{l,i,j}^n,\\
&\theta_l \coloneqq \min\left\{\frac{\phi_{l,i,j}^n}{\phi_{l,i,j}^n - m_{i,j}^{(l)}}, 1\right\}, \quad m_{i,j}^{(l)} \coloneqq \min_{(x,y)\in\tilde{K}_{i,j}} P_{i,j}^{(l)}(x,y), \quad l = 1,\dots,N.
\end{aligned}
\tag{4.7}
$$

Then the cell average of $\tilde{P}_{i,j}^{(l)}(x,y)$ over $K_{i,j}$ is still $\phi_{l,i,j}^n$ and

$$
\tilde{P}_{i,j}^{(l)}(x,y) \geq 0 \quad \text{for all } (x,y)\in\tilde{K}_{i,j} \text{ and } l = 1,\dots,N.
\tag{4.8}
$$

Next, we define the polynomial

$$
\begin{aligned}
&\hat{P}_{i,j}(x,y) \coloneqq \hat{\theta}\big(\tilde{P}_{i,j}^{(1)}(x,y) + \cdots + \tilde{P}_{i,j}^{(N)}(x,y) - \phi_{i,j}^n\big) + \phi_{i,j}^n,\\
&\hat{\theta} \coloneqq \min\left\{\left|\frac{\phi_{\max} - \phi_{i,j}^n}{M_{i,j} - \phi_{i,j}^n}\right|, 1\right\}, \quad M_{i,j} \coloneqq \max_{(x,y)\in\tilde{K}_{i,j}}\left(\sum_{l=1}^N \tilde{P}_{i,j}^{(l)}(x,y)\right).
\end{aligned}
\tag{4.9}
$$

Thus, $\hat{P}_{i,j}(x,y) \leq \phi_{\max}$ for all $(x,y)\in\tilde{K}_{i,j}$. We now define the modified polynomials

$$
\bar{P}_{i,j}^{(l)}(x,y) \coloneqq \hat{\theta}\big(\tilde{P}_{i,j}^{(l)}(x,y) - \phi_{l,i,j}^n\big) + \phi_{l,i,j}^n, \quad l = 1,\dots,N,
\tag{4.10}
$$

and setting $\bar{\boldsymbol{P}}_{i,j} \coloneqq (\bar{P}_{i,j}^{(1)},\dots,\bar{P}_{i,j}^{(N)})^{\mathrm{T}}$, replace the traces of the reconstruction polynomials (4.2) by pointwise evaluations to be sampled by the Gauss quadrature formula:

$$
\Phi_{i\pm1/2,j}^{\mp}(y_j^\beta) \approx \bar{\boldsymbol{P}}_{i,j}(x_{i\pm1/2}, y_j^\beta), \quad \Phi_{i,j\pm1/2}^{\mp}(x_i^\beta) \approx \bar{\boldsymbol{P}}_{i,j}(x_i^\beta, y_{j\pm1/2}).
\tag{4.11}
$$

*Remark* 5. The accuracy analysis of [1, sect. 4.4] shows that the limiters defined in (4.7) and (4.9) do not destroy the order of accuracy of the reconstruction.

We first prove the following lemma.

LEMMA 6. *Consider $\bar{\boldsymbol{P}}_{i,j}(x,y)$ defined by (4.10). If $\Phi_{i,j}^n \in \mathcal{D}$, then $\bar{\boldsymbol{P}}_{i,j}(x,y) \in \mathcal{D}$ for all $(x,y)\in\tilde{K}_{i,j}$.*

*Proof.* Let $(x,y) \in \tilde{K}_{i,j}$. By definition (4.10), (4.8), and since $\hat{\theta} \in [0,1]$, we have

$$\bar{P}_{i,j}^{(l)}(x,y) = \hat{\theta}\tilde{P}_{i,j}^{(l)}(x,y) + (1-\hat{\theta})\phi_{l,i,j}^n \geq 0 \quad l = 1,\dots,N, \tag{4.12}$$

$$\begin{aligned}\bar{P}_{i,j}^{(1)}(x,y) + \cdots + \bar{P}_{i,j}^{(N)}(x,y) &= \hat{\theta}\big(\tilde{P}_{i,j}^{(1)}(x,y) + \cdots + \tilde{P}_{i,j}^{(N)}(x,y) - \phi_{i,j}^n\big) + \phi_{i,j}^n \\ &= \hat{P}_{i,j}(x,y) \leq \phi_{\max}.\end{aligned} \tag{4.13}$$

Combining (4.12) and (4.13) we deduce that $\bar{\boldsymbol{P}}_{i,j}(x,y) \in \mathcal{D}$ for all $(x,y) \in \tilde{K}_{i,j}$. ∎

Now, we are in position to state the following result.

THEOREM 7. *Consider the finite volume scheme (4.6) associated with the reconstruction polynomials $\bar{\boldsymbol{P}}_{i,j}(x,y)$ defined by (4.10) in the sense that (4.11) is used with $\tilde{K}_{i,j} = S_{i,j}$, where $S_{i,j}$ is the stencil (4.5). If the CFL condition*

$$\max_{i,j,\beta}\big\{\alpha_{i+1/2,\beta}, \alpha_{\beta,j+1/2}\big\}(\lambda_x^n + \lambda_y^n) \leq \min_{\gamma=1,\dots,L} \hat{w}_\gamma \tag{4.14}$$

*is in effect and $\Phi_{i,j}^n \in \mathcal{D}$ for all $i$ and $j$, then $\Phi_{i,j}^{n+1} \in \mathcal{D}$ for all $i$ and $j$.*

*Proof.* Utilizing the Gauss quadrature rule (4.3) for $\bar{\boldsymbol{P}}_{i,j}$ on $I_i^x$, we get

$$\begin{aligned}\Phi_{i,j}^n &= \frac{1}{h_x h_y}\int_{I_j^y}\int_{I_i^x} \bar{\boldsymbol{P}}_{i,j}(x,y)\,\mathrm{d}x\,\mathrm{d}y = \frac{1}{h_x h_y}\int_{I_j^y}\left(\sum_{\beta=1}^{G} \bar{\boldsymbol{P}}_{i,j}(x_i^\beta, y) w_\beta h_x\right)\mathrm{d}y \\ &= \sum_{\beta=1}^{G} w_\beta\left(\frac{1}{h_y}\int_{I_j^y} \bar{\boldsymbol{P}}_{i,j}(x_i^\beta, y)\mathrm{d}y\right) = w_1\Phi_{i,j}^{n,y,1} + \cdots + w_G\Phi_{i,j}^{n,y,G},\end{aligned}$$

where $\Phi_{i,j}^{y,\beta}$ denotes the average of $\bar{\boldsymbol{P}}_{i,j}(x^\beta, y)$ over $I_j^y$. Analogously, applying the Gauss quadrature rule (4.3) to the interval $I_j^y$ yields $\Phi_{i,j}^n = w_1\Phi_{i,j}^{x,1} + \cdots + w_G\Phi_{i,j}^{x,G}$, where $\Phi_{i,j}^{x,\beta}$ denotes the average of $\bar{\boldsymbol{P}}_{i,j}(x, y^\beta)$ over $I_i^x$. Combining both representations of $\Phi_{i,j}^n$ and considering that $\tilde{\lambda}_x + \tilde{\lambda}_y = 1$, we obtain

$$\Phi_{i,j}^n = \tilde{\lambda}_x\big(w_1\Phi_{i,j}^{x,1} + \cdots + w_G\Phi_{i,j}^{x,G}\big) + \tilde{\lambda}_y\big(w_1\Phi_{i,j}^{y,1} + \cdots + w_G\Phi_{i,j}^{y,G}\big). \tag{4.15}$$

Replacing $\Phi_{i,j}^n$ on the right-hand side of (4.6) by the right-hand side of (4.15) we get

$$\Phi_{i,j}^{n+1} = \tilde{\lambda}_x\big(w_1\boldsymbol{H}_{i,j}^{x,1} + \cdots + w_G\boldsymbol{H}_{i,j}^{x,G}\big) + \tilde{\lambda}_y\big(w_1\boldsymbol{H}_{i,j}^{y,1} + \cdots + w_G\boldsymbol{H}_{i,j}^{y,G}\big), \tag{4.16}$$

where $\boldsymbol{H}_{i,j}^{\sigma,\beta} \coloneqq \Phi_{i,j}^{\sigma,\beta} - \lambda\tilde{\boldsymbol{H}}_{i,j}^{\sigma,\beta}$ for $\sigma = x, y$ and $\beta = 1,\dots,G$. To prove that all vectors $\boldsymbol{H}_{i,j}^{\sigma,\beta}$ assume values in $\mathcal{D}$ we focus on $\sigma = x$. Next, define the vectors

$$\begin{aligned}&\boldsymbol{s}_{i,j}^{\gamma,\beta} \coloneqq \bar{\boldsymbol{P}}_{i,j}(\hat{x}_i^\gamma, y_j^\beta) \quad \text{for } \gamma = 1,\dots,L, \\ &\boldsymbol{s}_{i,j}^{0,\beta} \coloneqq \bar{\boldsymbol{P}}_{i-1,j}(x_{i-1/2}, y_j^\beta), \quad \text{and} \quad \boldsymbol{s}_{i,j}^{L+1,\beta} \coloneqq \bar{\boldsymbol{P}}_{i+1,j}(x_{i+1/2}, y_j^\beta)\end{aligned}$$

for the Gauss-Lobatto points (4.4). By Lemma 6, $\boldsymbol{s}_{i,j}^{\gamma,\beta} \in \mathcal{D}$ for these $\gamma$ and $\beta$. The quadrature rule implies that $\Phi_{i,j}^{x,\beta} = \hat{w}_1\boldsymbol{s}_{i,j}^{1,\beta} + \cdots + \hat{w}_L\boldsymbol{s}_{i,j}^{L,\beta}$. Moreover (cf. (4.6)),

$$\begin{aligned}&\Delta_-^x\boldsymbol{\mathcal{F}}^x\big(q_{i+1/2,j}^x, \Phi_{i+1/2,j}^-(y_j^\beta), \Phi_{i+1/2,j}^+(y_j^\beta)\big) \\ &= \boldsymbol{\mathcal{F}}^x\big(q_{i+1/2,j}^x, \boldsymbol{s}_{i,j}^{L,\beta}, \boldsymbol{s}_{i,j}^{L+1,\beta}\big) - \boldsymbol{\mathcal{F}}^x\big(q_{i-1/2,j}^x, \boldsymbol{s}_{i,j}^{0,\beta}, \boldsymbol{s}_{i,j}^{1,\beta}\big) \\ &= \boldsymbol{\mathcal{F}}^x\big(q_{i+1/2,j}^x, \boldsymbol{s}_{i,j}^{L,\beta}, \boldsymbol{s}_{i,j}^{L+1,\beta}\big) - \boldsymbol{\mathcal{F}}^x\big(0, \boldsymbol{s}_{i,j}^{L-1,\beta}, \boldsymbol{s}_{i,j}^{L,\beta}\big)\end{aligned}$$

$$+\sum_{\gamma=2}^{L-1}\big(\boldsymbol{\mathcal{F}}^x\big(0,\boldsymbol{s}_{i,j}^{\gamma,\beta},\boldsymbol{s}_{i,j}^{\gamma+1,\beta}\big)-\boldsymbol{\mathcal{F}}^x\big(0,\boldsymbol{s}_{i,j}^{\gamma-1,\beta},\boldsymbol{s}_{i,j}^{\gamma,\beta}\big)\big)$$
$$+\boldsymbol{\mathcal{F}}^x\big(0,\boldsymbol{s}_{i,j}^{1,\beta},\boldsymbol{s}_{i,j}^{2,\beta}\big)-\boldsymbol{\mathcal{F}}^x\big(q^x_{i-1/2,j},\boldsymbol{s}_{i,j}^{0,\beta},\boldsymbol{s}_{i,j}^{1,\beta}\big).$$

We may now rewrite $\boldsymbol{H}_{i,j}^{x,\beta}$ as follows, where $\boldsymbol{s}^\gamma \coloneqq \boldsymbol{s}_{i,j}^{\gamma,\beta}$ in this particular computation:

$$\begin{aligned}
\boldsymbol{H}_{i,j}^{x,\beta} &= \hat{w}_1\boldsymbol{s}^1+\cdots+\hat{w}_L\boldsymbol{s}^L-\lambda\big(\boldsymbol{\mathcal{F}}^x(q^x_{i+1/2,j},\boldsymbol{s}^L,\boldsymbol{s}^{L+1})-\boldsymbol{\mathcal{F}}^x(q^x_{i-1/2,j}\boldsymbol{s}^0,\boldsymbol{s}^1)\big)\\
&\quad-\tfrac{1}{2}\lambda(q^x_{i+1/2,j}\boldsymbol{s}^{L+1}-q^x_{i-1/2,j}\boldsymbol{s}^0)\\
&= \hat{w}_1\boldsymbol{s}^1+\cdots+\hat{w}_L\boldsymbol{s}^L-\lambda\big(\boldsymbol{\mathcal{F}}^x(q^x_{i+1/2,j},\boldsymbol{s}^L,\boldsymbol{s}^{L+1})-\boldsymbol{\mathcal{F}}^x(0,\boldsymbol{s}^{L-1},\boldsymbol{s}^L)\big)\\
&\quad-\lambda\sum_{\gamma=2}^{L-1}\big(\boldsymbol{\mathcal{F}}^x(0,\boldsymbol{s}^\gamma,\boldsymbol{s}^{\gamma+1})-\boldsymbol{\mathcal{F}}^x(0,\boldsymbol{s}^{\gamma-1},\boldsymbol{s}^\gamma)\big)\\
&\quad-\lambda\big(\boldsymbol{\mathcal{F}}^x(0,\boldsymbol{s}^1,\boldsymbol{s}^2)-\boldsymbol{\mathcal{F}}^x(q^x_{i-1/2,j},\boldsymbol{s}^0,\boldsymbol{s}^1)\big)-\tfrac{1}{2}\lambda\big(q^x_{i+1/2,j}\boldsymbol{s}^{L+1}-q^x_{i-1/2,j}\boldsymbol{s}^0\big)\\
&= \hat{w}_L\big\{\boldsymbol{s}^L-\hat{w}_L^{-1}\lambda\big(\boldsymbol{\mathcal{F}}^x(q^x_{i+1/2,j},\boldsymbol{s}^L,\boldsymbol{s}^{L+1})-\boldsymbol{\mathcal{F}}^x(0,\boldsymbol{s}^{L-1},\boldsymbol{s}^L)-\tfrac{1}{2}q^x_{i+1/2,j}\boldsymbol{s}^{L+1}\big)\big\}\\
&\quad+\sum_{\gamma=2}^{L-1}\hat{w}_\gamma\big\{\boldsymbol{s}^\gamma-\hat{w}_\gamma^{-1}\lambda\big(\boldsymbol{\mathcal{F}}^x(0,\boldsymbol{s}^\gamma,\boldsymbol{s}^{\gamma+1})-\boldsymbol{\mathcal{F}}^x\big(0,\boldsymbol{s}^{\gamma-1},\boldsymbol{s}^\gamma\big)\big)\big\}\\
&\quad+\hat{w}_1\big\{\boldsymbol{s}^1-\hat{w}_1^{-1}\lambda\big(\boldsymbol{\mathcal{F}}^x(0,\boldsymbol{s}^1,\boldsymbol{s}^2)-\boldsymbol{\mathcal{F}}^x(q^x_{i-1/2,j},\boldsymbol{s}^0,\boldsymbol{s}^1)+\tfrac{1}{2}q^x_{i-1/2,j}\boldsymbol{s}^0\big)\big\}\\
&= \hat{w}_1\boldsymbol{H}_{i,j}^{x,1,\beta}+\cdots+\hat{w}_L\boldsymbol{H}_{i,j}^{x,L,\beta},
\end{aligned}$$

where $\boldsymbol{H}_{i,j}^{x,\gamma,\beta}$ denotes the corresponding term in curled brackets multiplying $\hat{w}_\gamma$, $\gamma=1,\dots,L$. If $H_{l,i,j}^{x,\beta}$ and $H_{l,i,j}^{x,\gamma,\beta}$, $l=1,\dots,N$, denote the components of the vectors $\boldsymbol{H}_{i,j}^{x,\beta}$ and $\boldsymbol{H}_{i,j}^{x,\gamma,\beta}$, respectively, then Theorem 3 implies that $H_{l,i,j}^{x,\gamma,\beta}\geq 0$ for $l=1,\dots,N$, $\beta=1,\dots,G$ and $\gamma=1,\dots,L$, along with

$$\sum_{l=1}^N H_{l,i,j}^{x,\gamma,\beta}\leq\begin{cases}\phi_{\max}+\frac{1}{2}\hat{w}_1^{-1}\lambda q^x_{i-1/2,j}\phi_{\max}, & \gamma=1,\\ \phi_{\max}, & \gamma=2,\dots,L-1,\\ \phi_{\max}-\frac{1}{2}\hat{w}_L^{-1}\lambda q^x_{i+1/2,j}\phi_{\max}, & \gamma=L.\end{cases}$$

Consequently, $H_{l,i,j}^{x,\beta}\geq 0$ and

$$\begin{aligned}\sum_{l=1}^N H_{l,i,j}^{x,\beta}&=\sum_{l=1}^N\sum_{\gamma=1}^L\hat{w}_\gamma H_{l,i,j}^{x,\gamma,\beta}=\sum_{\gamma=1}^L\hat{w}_\gamma\sum_{l=1}^N H_{l,i,j}^{x,\gamma,\beta}\\&\leq\phi_{\max}(\hat{w}_1+\cdots+\hat{w}_L)-\tfrac{1}{2}\lambda\phi_{\max}\Delta_-^x q^x_{i+1/2,j}\\&=\phi_{\max}\big(1-\tfrac{1}{2}\lambda\Delta_-^x q^x_{i+1/2,j}\big),\quad \beta=1,\dots,G.\end{aligned}\tag{4.17}$$

Analogously we get $H_{l,i,j}^{y,\beta}\geq 0$ for $l=1,\dots,N$, $\beta=1,\dots,G$ and $\gamma=1,\dots,L$, and

$$\sum_{l=1}^N H_{l,i,j}^{y,\beta}\leq\phi_{\max}\big(1-\tfrac{1}{2}\lambda\Delta_-^y q^y_{i,j+1/2}\big),\quad \beta=1,\dots,G.\tag{4.18}$$

Then, from (4.16) it is clear that $\phi_{l,i,j}^{n+1}\geq 0$, for $l=1,\dots,N$. Finally, by using (4.17), (4.18) and the DDF property (3.10) we obtain

$$\sum_{l=1}^N\phi_{l,i,j}^{n+1}=\tilde{\lambda}_x\sum_{\beta=1}^G w_\beta\left(\sum_{l=1}^N H_{l,i,j}^{x,\beta}\right)+\tilde{\lambda}_y\sum_{\beta=1}^G w_\beta\left(\sum_{l=1}^N H_{l,i,j}^{y,\beta}\right)$$

**Algorithm 4.1** Coupling algorithm

**Input:** $\rho_{\mathrm{f}}, \rho_{\mathrm{s}}, \mu_{\mathrm{f}}, n_{\mathrm{RZ}}, \upsilon, \theta, N, d_1, \ldots, d_N, x_a, x_b, y_a, y_b, k, m$

$\Phi^n \leftarrow \Phi_0, \quad \tau \leftarrow \tau_0, \quad n \leftarrow 0$

$[\boldsymbol{q}^n, p] \leftarrow \mathrm{STOKES}(\Phi^n)$, where STOKES means (3.7) given $\Phi^n$

**while** $t < t_{\mathrm{end}}$ **do**

  **for** $i = 1, \ldots, k$ **do**

    **for** $j = 1, \ldots, m$ **do**

      $[\Phi^{\pm}_{i\pm1/2,\beta}, \Phi^{\pm}_{\beta,j\pm1/2}] \leftarrow \mathrm{CWENO3}(\Phi^n)$

      compute $\tilde{\boldsymbol{H}}^{\sigma,\beta}_{i,j}$, $\sigma = x, y$ from (4.6) given $\alpha_{i\pm1/2,\beta}, \alpha_{\beta,j\pm1/2}$

      $\boldsymbol{H}^{\sigma,\beta}_{i,j} \leftarrow \Phi^{\sigma,\beta}_{i,j} - \lambda \tilde{\boldsymbol{H}}^{\sigma,\beta}_{i,j}$, $\sigma = x, y$

      $\mathcal{L}(\Psi^n) \leftarrow \tilde{\lambda}_x \sum_\beta w_\beta \boldsymbol{H}^{x,\beta}_{i,j} + \tilde{\lambda}_y \sum_\beta w_\beta \boldsymbol{H}^{y,\beta}_{i,j}$, where $\Psi^n = (\boldsymbol{q}^n, \Phi^n)$

    **end for**

  **end for**

  $\Phi^{(1)} \leftarrow \Phi^n + \tau \mathcal{L}(\Psi^n)$

  $[\boldsymbol{q}^{(1)}, \sim] \leftarrow \mathrm{STOKES}(\Phi^{(1)})$

  compute $\mathcal{L}(\Psi^{(1)})$, where $\Psi^{(1)} = (\boldsymbol{q}^{(1)}, \Phi^{(1)})$ by processes described above

  $\Phi^{(2)} \leftarrow \frac{3}{4}\Phi^n + \frac{1}{4}\big(\Phi^{(1)} + \tau \mathcal{L}(\Psi^{(1)})\big)$

  $[\boldsymbol{q}^{(2)}, p] \leftarrow \mathrm{STOKES}(\Phi^{(1)})$

  compute $\mathcal{L}(\Psi^{(2)})$, where $\Psi^{(2)} = (\boldsymbol{q}^{(2)}, \Phi^{(2)})$ by processes described above

  $\Phi^{n+1} \leftarrow \frac{1}{3}\Phi^n + \frac{2}{3}\big(\Phi^{(2)} + \tau \mathcal{L}(\Psi^{(2)})\big)$, where $\Psi^{(2)} = (\boldsymbol{q}^{(2)}, \Phi^{(2)})$

  $\boldsymbol{q}^{n+1} \leftarrow \boldsymbol{q}^{(2)}, \quad p^{n+1} \leftarrow p$

  $\Delta t \leftarrow \mathrm{CFL} \cdot \frac{h_x h_y}{h_x + h_y} \cdot \frac{\min_\gamma \hat{w}_\gamma}{\alpha}$, where $\alpha \coloneqq \max_{i,j,\beta}\big\{\alpha_{i+1/2,\beta}, \alpha_{\beta,j+1/2}\big\}$

  $t \leftarrow t + \tau, \quad n \leftarrow n + 1$

**end while**

**Output:** $\{(\Phi^1, \boldsymbol{q}^1, p^1), \ldots, (\Phi^n, \boldsymbol{q}^n, p^n)\}$

$$
\begin{aligned}
&\leq (\tilde{\lambda}_x + \tilde{\lambda}_y)(w_1 + \cdots + w_G)\phi_{\max} - \tfrac{1}{2}\phi_{\max}\lambda_x(w_1 + \cdots + w_G)\Delta^x_- q^x_{i+1/2,j} \\
&\quad - \tfrac{1}{2}\phi_{\max}\lambda_y(w_1 + \cdots + w_G)\Delta^y_- q^y_{i,j+1/2} \\
&= \phi_{\max}\big(1 - \tfrac{1}{2}\tau(h_x^{-1}\Delta^x_- q^x_{i+1/2,j} + h_y^{-1}\Delta^y_- q^y_{i,j+1/2})\big) = \phi_{\max}.
\end{aligned}
$$

Hence it follows that $\Phi^{n+1}_{i,j} \in \mathcal{D}$. ∎

**4.4. Time discretization.** We here employ a strong-stability preserving (SSP) third-order TVD Runge-Kutta time discretization. Due to convexity in the intermediate stages, this time discretization preserves the IRP property. To satisfy the CFL condition (4.14) the time step $\tau$ is computed deceptively for each time step $n$, see Algorithm 4.1 for a description of the complete numerical scheme.

**5. Numerical examples.** In part following [28], we define dimensionless (starred) variables via $\boldsymbol{x} = X\boldsymbol{x}^*$, $t = Tt^*$, and $\boldsymbol{q} = (X/T)\boldsymbol{q}^*$, where $X$ is a reference length and $T \coloneqq X/v_{\mathrm{St}}$ is a reference time, where $v_{\mathrm{St}} \coloneqq (\rho_{\mathrm{s}} - \rho_{\mathrm{f}})g d_1^2/(18\mu_{\mathrm{f}})$ is the Stokes velocity, i.e., the settling velocity of a single particle of the largest species, where $\mu_{\mathrm{f}}$ is the viscosity of the fluid. Furthermore, we define the parameter $\mu_0 \coloneqq (\rho_{\mathrm{f}}/(18(\rho_{\mathrm{s}} - \rho_{\mathrm{f}})))(X/d_1)^2$. To simplify notation, we omit the asterisk and obtain the dimensionless Stokes equation (1.1b), where $\mu(\phi) = (1/\mu_0)(1 - \phi/\phi_{\max})^{-\upsilon}$ and $\boldsymbol{g}(\phi) \coloneqq (\rho_{\mathrm{s}} - \rho_{\mathrm{f}})\phi\boldsymbol{k}/\rho_{\mathrm{f}}$. In all examples we use the constants $n_{\mathrm{RZ}} = 4.6$, $\upsilon = 2$, $X = 1\,\mathrm{m}$, $\phi_{\max} = 0.6$, $\rho_{\mathrm{s}} = 2790\,\mathrm{kg/m^3}$, $\rho_{\mathrm{f}} = 1208\,\mathrm{kg/m^3}$, and $\mu_{\mathrm{f}} = 0.02416\,\mathrm{Pa\,s}$, such that $\mu_0 = 4.086 \times 10^3$. We use the initial time step $\tau_0 = 0.0003$ and CFL $= 1.0$.

TABLE 1
*Example 1: $L^1$ error, numerical order, and CPU time for IRP-LFCW and IRP-LF schemes at $t = 0.01$ or $0.2$ with $\theta = 0^\circ$ or $10^\circ$. Reference solutions are computed on $k_{\text{ref}} \times m_{\text{ref}} = 409600$ cells.*

| | $k \times m$ | $e_1^{k,m}$ | $\theta_1^{k,m}$ | $e_2^{k,m}$ | $\theta_1^{k,m}$ | $e_{\text{tot}}^{k,m}$ | $\theta_{\text{tot}}^{k,m}$ | cpu [s] |
|---|---|---|---|---|---|---|---|---|
| IRP-LFCW, $t = 0.01$, $\theta = 0^\circ$ | $5 \times 5$ | 2.097e-03 | — | 2.093e-03 | — | 4.191e-03 | — | 2.719e-02 |
| | $10 \times 10$ | 6.282e-04 | 1.73 | 6.316e-04 | 1.72 | 1.259e-03 | 1.73 | 3.981e-02 |
| | $20 \times 20$ | 9.571e-05 | 2.71 | 9.604e-05 | 2.71 | 1.917e-03 | 2.71 | 2.451e-01 |
| | $40 \times 40$ | 1.715e-05 | 2.48 | 1.635e-05 | 2.55 | 3.351e-04 | 2.51 | 1.681 |
| IRP-LFCW, $t = 0.2$, $\theta = 0^\circ$ | $5 \times 5$ | 9.165e-03 | — | 1.015e-02 | — | 1.932e-02 | — | 3.80e-01 |
| | $10 \times 10$ | 4.109e-03 | 1.15 | 4.880e-03 | 1.05 | 8.990e-03 | 1.10 | 9.63e-01 |
| | $20 \times 20$ | 2.259e-03 | 0.86 | 2.854e-03 | 0.77 | 5.114e-03 | 0.81 | 3.85 |
| | $40 \times 40$ | 1.160e-03 | 0.96 | 8.565e-04 | 1.73 | 2.017e-03 | 1.34 | 27.81 |
| IRP-LFCW, $t = 0.01$, $\theta = 10^\circ$ | $5 \times 5$ | 2.003e-03 | — | 2.006e-03 | — | 4.009e-03 | — | 1.736e-02 |
| | $10 \times 10$ | 5.798e-04 | 1.78 | 5.833e-04 | 1.78 | 1.163e-03 | 1.78 | 3.223e-02 |
| | $20 \times 20$ | 8.911e-05 | 2.70 | 8.949e-05 | 2.70 | 1.786e-04 | 2.70 | 2.372e-01 |
| | $40 \times 40$ | 1.632e-05 | 2.44 | 1.553e-05 | 2.52 | 3.185e-05 | 2.48 | 1.602 |
| IRP-LF, $t = 0.01$, $\theta = 10^\circ$ | $5 \times 5$ | 2.974e-03 | — | 2.942e-03 | — | 5.916e-03 | — | 9.041e-03 |
| | $10 \times 10$ | 1.708e-03 | 0.79 | 1.709e-03 | 0.78 | 3.417e-03 | 0.79 | 1.330e-02 |
| | $20 \times 20$ | 9.152e-04 | 0.90 | 9.174e-04 | 0.89 | 1.832e-03 | 0.89 | 7.980e-02 |
| | $40 \times 40$ | 4.585e-04 | 1.00 | 4.598e-04 | 1.00 | 9.183e-04 | 1.00 | 0.526 |

For the CWENO reconstruction we employ the same, linear, non-linear weights, $\varepsilon$, and $\mathtt{p}$ as in [10]. We denote by "LFCW" and "IRP-LFCW" the numerical scheme without and with the limiters, respectively and by "IRP-LF" the first order method (3.13). To compute approximate $L^1$ errors at different times we denote by $(\phi_{l,i,j}^{k,m}(t))$ for $i = 1, \dots, k$ and $j = 1, \dots, m$ and $(\phi_{l,i,j}^{k_{\text{ref}},m_{\text{ref}}}(t))$ for $i = 1, \dots, k_{\text{ref}}$ and $j = 1, \dots, m_{\text{ref}}$, $l = 1, \dots, N$, the numerical solution at time $t$ calculated with $k \times m$ and $k_{\text{ref}} \times m_{\text{ref}}$ cells, respectively. We compute the projection of the reference solution by

$$\tilde{\phi}_{l,i,j}^{\text{ref},k,m}(t) = \frac{1}{R_x R_y} \sum_{p=1}^{R_x} \sum_{q=1}^{R_y} \phi_{l,R_x(i-1)+p,R_y(j-1)+q}^{n_{\text{ref}}}(t), \quad l = 1, \dots, N,$$

where $R_x \coloneqq k_{\text{ref}}/k$, and $R_y \coloneqq m_{\text{ref}}/m$. The approximate $L^1$ errors of the numerical solution on the mesh with $k \times m$ cells at time $t$ is then given by

$$e_l^{k,m}(t) \coloneqq h_x h_y \sum_{i=1}^{k} \sum_{j=1}^{m} \left| \tilde{\phi}_{l,i,j}^{\text{ref},k,m}(t) - \phi_{l,i,j}^{k,m}(t) \right|, \ l = 1, \dots, N; \ e_{\text{tot}}^{k,m}(t) \coloneqq \sum_{l=1}^{N} e_l^{k,m}(t).$$

Then we calculate a numerical order of convergence by

$$\theta_l^{k,m}(t) \coloneqq \log_2 \frac{e_l^{k,m}(t)}{e_l^{2k,2m}(t)}, \quad l = 1, \dots, N; \quad \theta_{\text{tot}}^{k,m}(t) \coloneqq \log_2 \frac{e_{\text{tot}}^{k,m}(t)}{e_{\text{tot}}^{2k,2m}(t)}.$$

We measure $\underline{\phi}_l \coloneqq \min_{i,j,n}\{\phi_{l,i,j}^n\}$, $l = 1, \dots, N$, and $\bar{\phi} \coloneqq \max_{i,j,n}\{\phi_{i,j}^n\}$ to verify satisfaction of the IRP property. To verify the DDF property, we evaluate

$$e_{\operatorname{div} \boldsymbol{q}}^{k,m}(t_n) \coloneqq \max\left\{ |h_x^{-1} \Delta_-^x q_{i+1/2,j}^{x,n} + h_y^{-1} \Delta_-^y q_{i,j+1/2}^{y,n}| : 1 \le i \le k-1, \, 1 \le j \le m-1 \right\}.$$

**5.1. Example 1: $N = 2$, numerical order of accuracy.** We verify numerically the convergence rate of the IRP-LFCW scheme on $\Omega \coloneqq (0,1) \times (0,1)$ for $N = 2$

TABLE 2
*Example 2: $L^1$ errors, numerical order, and CPU time for IRP-LFCW schemes at $t = 1$. The reference solution is computed with $k_{\mathrm{ref}} = 1280$, $m_{\mathrm{ref}} = 320$, for a total of 409600 cells.*

| $k \times m$ | $e_1^{k,m}$ | $\theta_1^{k,m}$ | $e_2^{k,m}$ | $\theta_1^{k,m}$ | $e_{\mathrm{tot}}^{k,m}$ | $\theta_{\mathrm{tot}}^{k,m}$ | cpu [s] |
|---|---|---|---|---|---|---|---|
| $40 \times 10$ | 3.352e-02 | — | 1.421e-02 | — | 4.774e-02 | — | 1.46 |
| $80 \times 20$ | 1.829e-02 | 0.87 | 8.792e-03 | 0.69 | 2.708e-02 | 0.82 | 9.65 |
| $160 \times 40$ | 8.590e-03 | 1.09 | 5.382e-03 | 0.71 | 1.397e-02 | 0.95 | 96.2 |
| $320 \times 80$ | 4.645e-03 | 0.88 | 2.899e-03 | 0.89 | 7.545e-03 | 0.88 | 942.8 |

TABLE 3
*Example 2: minimum of the solutions $\phi_{l,i,j}^n$, $l = 1, 2$, and maximum of the solution $\phi_{i,j}^n$ obtained by schemes LFCW and IRP-LFCW with $k = 160$, $m = 40$, until $t = 1.15$.*

| | $\theta$ | $\underline{\phi}_{,1}$ | $\underline{\phi}_{,2}$ | $\bar{\phi}$ | | $\theta$ | $\underline{\phi}_{,1}$ | $\underline{\phi}_{,2}$ | $\bar{\phi}$ |
|---|---|---|---|---|---|---|---|---|---|
| LFCW | 0° | -3.98e-03 | -1.92e-03 | 0.585346 | | 0° | 1.27e-31 | 3.69e-20 | 0.585263 |
| | 10° | -7.25e-03 | -4.19e-03 | 0.600110 | IRP- | 10° | 6.34e-22 | 1.07e-17 | 0.599487 |
| | 20° | -1.61e-02 | -7.13e-03 | 0.603366 | LFCW | 20° | 8.82e-20 | 1.08e-17 | 0.599877 |
| | 30° | -1.81e-02 | -1.34e-02 | 0.603352 | | 30° | 4.29e-20 | 6.46e-18 | 0.599906 |

particle species with diameters $d_1 = 2.9\times10^{-3}$ m and $d_2 = 2.0\times10^{-3}$ m. The (smooth) initial datum is $\Phi_0(x, y) = (\phi_0(x, y), \phi_0(x, y))^{\mathrm{T}}$ with $k_{\mathrm{ref}} = m_{\mathrm{ref}} = 5 \times 2^7$. The approximate $L^1$ errors and corresponding numerical orders for $\theta = 0°$ are displayed in the upper two blocks of Table 1 for $t = 0.01$ (before shock formation) and $t = 0.2$ (after shock formation). The convergence rates for increasing values of $k \times m$ confirm second-order accuracy of the IRP-LFCW scheme when the solution is still smooth. In the lower two blocks of Table 1 we compare the behavior of the first-order IRP-LF and the second-order IRP-LFCW schemes at $t = 0.01$ for $\theta = 10°$ and observe a substantial improvement of accuracy of the IRP-LFCW scheme compared with the IRP-LF scheme.

**5.2. Example 2: $N = 2$, Boycott effect.** Here and in Examples 3 and 4 we consider $\Omega = (0, 4) \times (0, 1)$, and in this example $N = 2$ species with diameters as in Example 1 and $\Phi_0(x, y) = (0.06, 0.02)^{\mathrm{T}}$. We vary the inclination angle by $\theta = 0°, 10°, 20°, 30°$ and set $k = 320$ and $m = 40$. In Figure 3 we observe an increase in the settling efficiency for $\phi_1$, $\phi_2$, and $\phi$, respectively, with increasing $\theta$. In Table 2 we show approximate $L^1$ errors and CPU time for the scheme at $t = 0.01$. We observe the convergence of method with order smaller than one due to the presence of shocks in the numerical solutions. The quantities $\underline{\phi}_{,l}$ for $l = 1, 2$ and $\bar{\phi}$, with and without the use of limiters, are presented in Table 3, until $t = 1.15$. For the scheme without limiters it occurs that $\underline{\phi}_{,l} < 0$ or $\bar{\phi} > \phi_{\max}$, while in the case of the scheme with limiters the numerical solution remains within $\mathcal{D}$ for all $\theta$, as expected.

**5.3. Example 3: $N = 2$, Diehl test.** We start from $\Phi_0(x, y) = (0.12, 0.08)^{\mathrm{T}}$ if $0 \le x \le 2$ and $\Phi_0(x, y) = (0, 0)^{\mathrm{T}}$ if $2 < x \le 4$. We set an angle of inclination of $\theta = 30°$ and $k = 320$ and $m = 40$. Figures 4 and 5 show the simulation of the Diehl test at four time points. This process forms a characteristic "tongue" of mixture that settles until all solids accumulate at the bottom. We also plot the flow vectors of the fluid velocity, which illustrate the circulation effect. Moreover, we observe that as the particles settle, an increase in pressure develops near the bottom due to the growing concentration of sediment. Figure 6(a) shows the evolution of $\|\boldsymbol{q}\|_\infty$ over time, illustrating the decrease in fluid velocity as the particles approach a steady-

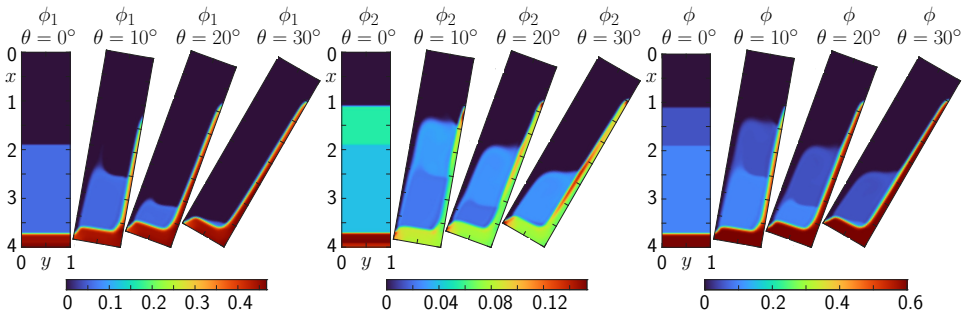


FIG. 3. *Example 2: simulation of the settling of an initially homogeneous bi-disperse suspension in vessels with various angles of inclination at simulated time $t = 3$.*

TABLE 4

*Example 3: $L^1$ errors, numerical order, and CPU time for IRP-LFCW schemes at $t = 0.01$. The reference solution is computed with $k_{\text{ref}} = 1280$, $m_{\text{ref}} = 320$, for a total of 409600 cells.*

| $k \times m$ | $e_1^{k,m}$ | $\theta_1^{k,m}$ | $e_2^{k,m}$ | $\theta_1^{k,m}$ | $e_{\text{tot}}^{k,m}$ | $\theta_{\text{tot}}^{k,m}$ | cpu [s] |
|---|---|---|---|---|---|---|---|
| $20 \times 5$ | 4.997e-03 | — | 3.132e-03 | — | 8.129e-03 | — | 2.158e-02 |
| $40 \times 10$ | 3.512e-03 | 0.51 | 2.091e-03 | 0.58 | 5.603e-03 | 0.53 | 6.102e-02 |
| $80 \times 20$ | 2.741e-03 | 0.35 | 1.572e-03 | 0.41 | 4.314e-03 | 0.37 | 5.057e-01 |
| $160 \times 40$ | 2.081e-03 | 0.40 | 1.104e-03 | 0.51 | 3.186e-03 | 0.43 | 3.961 |

state solution. Then, to study the numerical $L^1$ error we use $m = 5 \times 2^k$ and $k = 4m$, $k = 0, \dots, 3$, and employ a reference solution with $k_{\text{ref}} = 1280$, $m_{\text{ref}} = 320$. In Table 4 we show approximate $L^1$ errors and CPU times at $t = 0.01$. Again we observe orders of convergence smaller than one (for the same reason as in Example 2). The numerical extrema $\underline{\phi}_{,l}$ for $l = 1, 2$ and $\bar{\phi}$ until $t = 2.5$ are displayed in Table 5. For the scheme without limiters we observe some cases of $\underline{\phi}_{,l} < 0$ or $\bar{\phi} > \phi_{\max}$, while for the scheme with limiters the numerical solution remains in $\mathcal{D}$.

**5.4. Example 4 ($N = 4$).** We consider $N = 4$ particle species with diameters $d_1 = 290 \times 10^{-5}$ m, $d_2 = 250 \times 10^{-5}$ m, $d_3 = 200 \times 10^{-5}$ m, $d_4 = 130 \times 10^{-5}$ m, and $\theta = 30°$. We choose $\Phi_0(x, y) = (0.05, 0.05, 0.05, 0.05)^{\mathrm{T}}$. The values $\underline{\phi}_{,l}$ for $l = 1, 2$ and $\bar{\phi}$ are presented in Table 6 up to $t = 3$. In the case of the scheme without limiters, we observe some instances of $\underline{\phi}_{,l} < 0$ and $\bar{\phi} > \phi_{\max}$. On the contrary, for the scheme with limiters, the numerical solution remains within $\mathcal{D}$, as expected. We observe that the scheme preserves the invariant region $\mathcal{D}$ regardless of the number of different particle diameters $N$, as stated in Theorem 7. A key element in the proof is the DDF property (3.10), which we verify numerically by computing $e_{\operatorname{div} \boldsymbol{q}}^{k,m}$ at several times, see Table 7. Finally, Figure 6(b) shows the time evolution of $\|\boldsymbol{q}\|_\infty$, illustrating the decrease in fluid velocity as the particles approach a steady-state solution.

**6. Conclusions.** Many mathematical models impose some requirements for the admissibility of their state variables that call for numerical methods that satisfy these requirements, which usually come as invariant-region properties. In the case of multidimensional polydisperse sedimentation, with concentrations of each solid species and bulk velocity as state variables, individual concentrations are bound to be non-negative, the total solid concentration to be below a prescribed threshold, and the bulk velocity to be divergence-free. We herein propose a FV/FD solver on rotated Cartesian meshes for the coupled transport-flow model for two-dimensional polydis-

TABLE 5
*Example 3: minimum of the solutions $\phi^n_{l,i,j}$, $l = 1, 2$, and maximum of the solution $\phi^n_{i,j}$ obtained by schemes LFCW (without limiters), IRP-LFCW (with limiters), until $t = 2.5$ .*

| | $k \times m$ | $\underline{\phi}_{,1}$ | $\underline{\phi}_{,2}$ | $\bar{\phi}$ | | $k \times m$ | $\underline{\phi}_{,1}$ | $\underline{\phi}_{,2}$ | $\bar{\phi}$ |
|---|---|---|---|---|---|---|---|---|---|
| LFCW | $20 \times 5$ | -8.07e-03 | -6.21e-03 | 0.503233 | | $20 \times 5$ | 0.000000 | 0.000000 | 0.502076 |
| | $40 \times 10$ | -9.35e-03 | -8.70e-03 | 0.566773 | IRP- | $40 \times 10$ | 0.000000 | 0.000000 | 0.566619 |
| | $80 \times 20$ | -1.40e-02 | -1.45e-02 | 0.590730 | LFCW | $80 \times 20$ | 0.000000 | 0.000000 | 0.591358 |
| | $160 \times 40$ | -1.58e-02 | -1.67e-02 | 0.606257 | | $160 \times 40$ | 0.000000 | 0.000000 | 0.600000 |

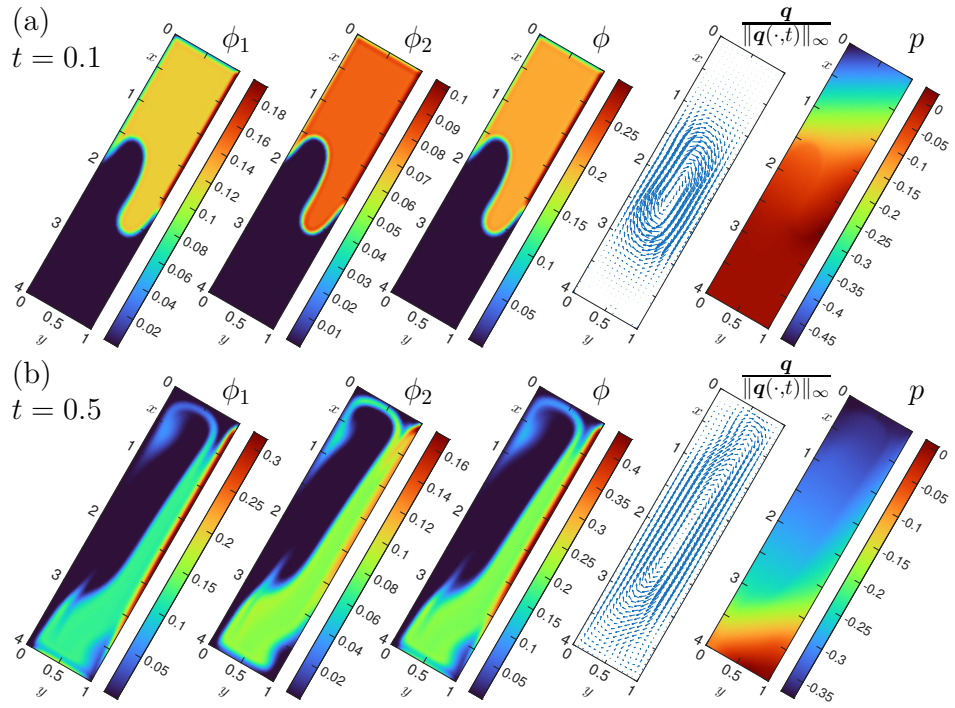


FIG. 4. *Example 3: numerical solution at simulated times (a) $t = 0.1$, (b) $t = 0.5$.*

perse sedimentation that achieves second-order accuracy through central WENO reconstructions of the solid concentration designed with a filtering (the application of scaling limiters) to achieve the cited invariant-region property. The FD solver for the bulk velocity provably yields velocity fields that are discretely divergence free (DDF), an essential feature for the proof of the IRP property. The numerical results reconfirm that the scheme has the advertised properties, and agree qualitatively with experimental information (e.g., in [14]) on batch settling, in particular concerning the formation of $N$ layers of different composition in the bottom of the vessel with nearly horizontal interfaces in conjunction with the upward-streaming layer of clear liquid beneath the inclined upper wall (see Figures 3 and 7). Future work should aim at comparing numerical simulations, with appropriate parameters, with experimental data such as the observed descent of these interfaces (cf., e.g., [14, Fig. 4]).

While the numerical examples have been limited to the MLB model, we empha-

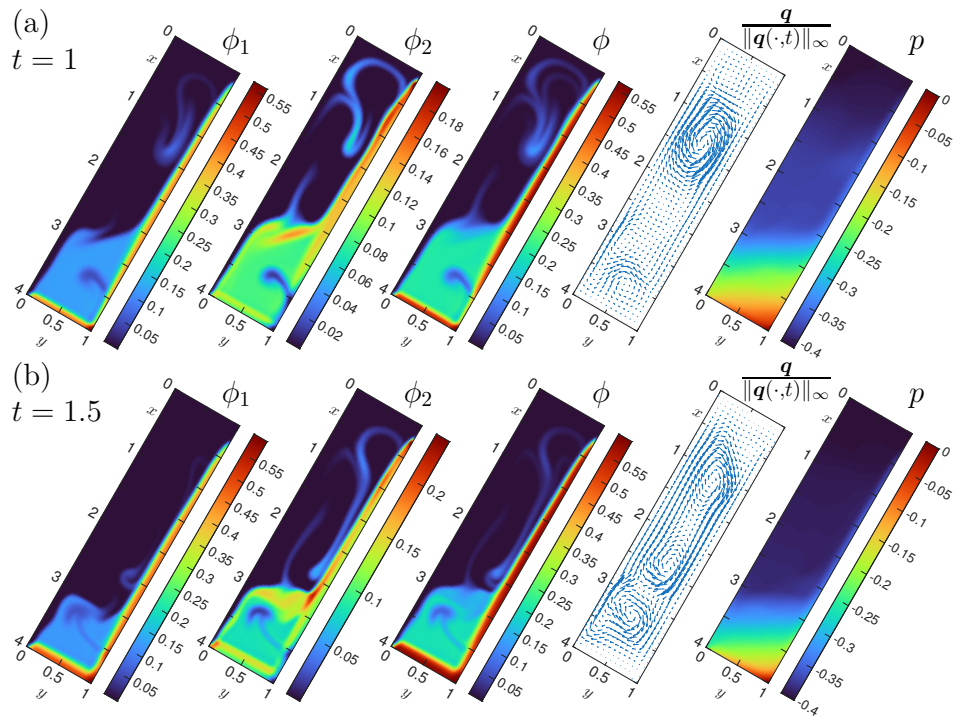


FIG. 5. *Example 3: numerical solution at simulated times (a) $t = 1.0$, (b) $t = 1.5$.*

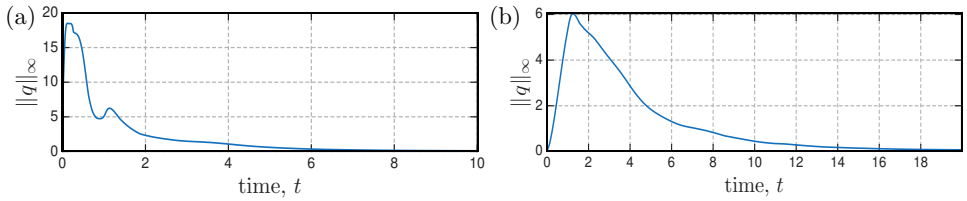


FIG. 6. *Examples 3 and 4: simulated evolution of $\|\boldsymbol{q}\|_\infty$ for (a) Example 3, (b) Example 4.*

size that the properties utilized for the analysis in sections 3 and 4 are not specific for that model but only rely on the generic assumptions of section 2.1. Other systematically constructed models could therefore replace the MLB model, if we consider for instance liquid-liquid dispersions [30]. That said, the assumptions of section 2.1 were inspired by the authors' previous work [1] where they are also verified for a model of multiclass vehicular traffic [3] (not considered herein). A closer inspection of the proof of Theorem 3 will reveal that these can possibly be relaxed, for example to include hindrance factors $V_i(\phi)$ (instead of $V(\phi)$) specific to each species. Such a generalization would make the present approach applicable to additional models (e.g., those reviewed in [8], which include [12]; or the one proposed in [2]).

With respect to numerical schemes we first comment that the LLF scheme has been chosen as a first-order scheme with IRP property because the proof of Theorem 3 takes a relatively simple form. Instead of the LLF numerical flux one could also employ the Harten-Lax-van Leer (HLL) numerical flux [22], for which the IRP property is

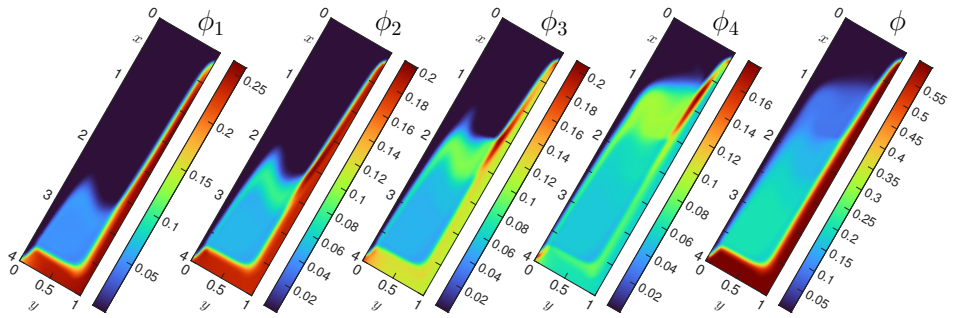


FIG. 7. *Example 4: numerical solution at simulated time* $t = 3$.

TABLE 6
*Example 4: minimum of the solutions* $\phi^n_{l,i,j}$, $l = 1, \ldots, 4$, *and maximum of the solution* $\phi^n_{i,j}$ *obtained by schemes LFCW (without limiters) and IRP-LFCW (with limiters) until* $t = 3$.

| | $k \times m$ | $\underline{\phi}_{,1}$ | $\underline{\phi}_{,2}$ | $\underline{\phi}_{,3}$ | $\underline{\phi}_{,4}$ | $\bar{\phi}$ |
|---|---|---|---|---|---|---|
| LFCW | $20 \times 5$ | -1.11e-03 | -9.42e-04 | -1.20e-03 | 6.25e-03 | 0.520070 |
| | $40 \times 10$ | -4.45e-03 | -3.03e-03 | -1.53e-03 | -1.03e-03 | 0.581633 |
| | $80 \times 20$ | -7.05e-03 | -7.33e-03 | -5.70e-03 | -2.88e-03 | 0.605112 |
| | $160 \times 40$ | -8.96e-03 | -1.13e-02 | -9.75e-03 | -1.07e-03 | 0.641303 |
| IRP-LFCW | $20 \times 5$ | 9.56e-11 | 6.10e-09 | 1.77e-06 | 6.35e-03 | 0.520119 |
| | $40 \times 10$ | 2.62e-18 | 1.08e-16 | 3.55e-12 | 1.46e-06 | 0.581665 |
| | $80 \times 20$ | 8.56e-21 | 3.05e-19 | 4.47e-18 | 5.53e-12 | 0.599993 |
| | $160 \times 40$ | 5.80e-26 | 7.17e-23 | 3.75e-20 | 1.46e-17 | 0.600000 |

also proven in [1]; the proof is, however, slightly more involved than for the LLF numerical flux. Furthermore, although rectangular domains allow us to simulate configurations of theoretical and practical interest (besides the Boycott effect and Diehl test, an inclined rectangular channel has been proposed as a continuous-flow device for classification of polydisperse suspensions [13]), our limitation to rectangular domains is essentially imposed by the FD Stokes solver that will be removed by using DG schemes on unstructured meshes. It remains to be evaluated whether a DG approach would allow for higher than second order of global accuracy.

**Acknowledgments.** We acknowledge support by ANID (Chile) via projects Fondecyt 1250676, CRHIAM (ANID/Fondap/1523A0001; to RB), Centro de Modelamiento Matemático (CMM; BASAL project FB210005), and Becas Doctorado Nacional 2022, folio 21221387 (to JBC). PM is supported by project PID2023-146836NB-I00, granted by MCIN/ AEI /10.13039/ 501100011033.

## 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 (2025), 114081 (29 pages).

[2] D. K. BASSON, S. BERRES, AND R. BÜRGER, *On models of polydisperse sedimentation with particle-size-specific hindered-settling factors*, Appl. Math. Modelling, 33 (2009), pp. 1815–1835.

[3] S. BENZONI-GAVAGE AND R. M. COLOMBO, *An n-populations model for traffic flow*, Eur. J.

TABLE 7
*Example 4: Indicator of the DDF property $e^{k,m}_{\mathrm{div}\,\boldsymbol{q}}$ for various discretizations and times.*

| $k \times m$ | $e^{k,m}_{\mathrm{div}\,\boldsymbol{q}}(t=0.5)$ | $e^{k,m}_{\mathrm{div}\,\boldsymbol{q}}(t=1.5)$ | $e^{k,m}_{\mathrm{div}\,\boldsymbol{q}}(2.5)$ | $e^{k,m}_{\mathrm{div}\,\boldsymbol{q}}(t=3)$ |
|---|---|---|---|---|
| $40 \times 10$ | 6.439e-15 | 8.567e-15 | 4.996e-15 | 3.885e-15 |
| $80 \times 20$ | 8.826e-15 | 2.398e-14 | 2.331e-14 | 1.582e-14 |
| $160 \times 40$ | 1.685e-14 | 5.051e-14 | 3.286e-14 | 2.664e-14 |

Appl. Math., 14 (2003), pp. 587–612.
[4] S. Berres, R. Bürger, K. H. Karlsen, and E. M. Tory, *Strongly degenerate parabolic-hyperbolic systems modeling polydisperse sedimentation with compression*, SIAM J. Appl. Math., 64 (2003), pp. 41–80.
[5] A. E. Boycott, *Sedimentation of blood corpuscles*, Nature, 104 (1920), p. 532.
[6] R. Bürger, R. Donat, P. Mulet, and C. A. Vega, *Hyperbolicity analysis of polydisperse sedimentation models via a secular equation for the flux Jacobian*, SIAM J. Appl. Math., 70 (2010), pp. 2186–2213.
[7] R. Bürger, R. Donat, P. Mulet, and C. A. Vega, *On the implementation of WENO schemes for a class of polydisperse sedimentation models*, J. Comput. Phys., 230 (2011), pp. 2322–2344.
[8] R. Bürger, K. H. Karlsen, E. M. Tory, and W. L. Wendland, *Model equations and instability regions for the sedimentation of polydisperse suspensions of spheres*, ZAMM Z. Angew. Math. Mech., 82 (2002), pp. 699–722.
[9] X. Cai, X. Zhang, and J. Qiu, *Positivity-preserving high order finite volume HWENO schemes for compressible Euler equations*, J. Sci. Comput., 68 (2016), pp. 464–483.
[10] M. J. Castro and M. Semplice, *Third-and fourth-order well-balanced schemes for the shallow water equations based on the CWENO reconstruction*, Int. J. Numer. Methods Fluids, 89 (2019), pp. 304–325.
[11] M. J. Castro Díaz, T. Chacón Rebollo, E. D. Fernández-Nieto, J. M. González Vida, and C. Parés, *Well-balanced finite volume schemes for 2D non-homogeneous hyperbolic systems. Application to the dam break of Aznalcóllar*, Comput. Methods Appl. Mech. Engrg., 197 (2008), pp. 3932–3950.
[12] R. H. Davis and H. Gecol, *Hindered settling function with no empirical parameters for polydisperse suspensions*, AIChE J., 40 (1994), pp. 570–575.
[13] R. H. Davis and H. Gecol, *Classification of concentrated suspensions using inclined settlers*, Int. J. Multiphase Flow, 22 (1996), pp. 563–574.
[14] R. H. Davis, E. Herbolzheimer, and A. Acrivos, *The sedimentation of polydisperse suspensions in vessels having inclined walls*, Int. J. Multiphase Flow, 8 (1983), pp. 571–585.
[15] C. Dawson, S. Sun, and M. F. Wheeler, *Compatible algorithms for coupled flow and transport*, Comput. Methods Appl. Mech. Eng., 193 (2004), pp. 2565–2580.
[16] S. Diehl, *Estimation of the batch-settling flux function for an ideal suspension from only two experiments*, Chem. Eng. Sci., 62 (2007), pp. 4589–4601.
[17] S. Ding and K. Wu, *A new discretely divergence-free positivity-preserving high-order finite volume method for ideal MHD equations*, SIAM J. Sci. Comput., 46 (2024), pp. A50–A79.
[18] R. Donat and P. Mulet, *A secular equation for the Jacobian matrix of certain multispecies kinematic flow models*, Numer. Meth. Partial Diff. Eqns., 26 (2010), pp. 159–175.
[19] S. Gong, Y.-J. Lee, Y. Li, and Y. Yu, *Positivity and maximum principle preserving discontinuous Galerkin finite element schemes for a coupled flow and transport*, J. Sci. Comput., 103 (2025), 20 (20 pages).
[20] Y. Guo, T. Xiong, and Y. Shi, *A maximum-principle-satisfying high-order finite volume compact WENO scheme for scalar conservation laws with applications in incompressible flows*, J. Sci. Comput., 65 (2015), pp. 83–109.
[21] F. Harlow and J. Welch, *Numerical calculation of time-dependent viscous incompressible flow of fluid with free surface*, Phys. Fluids, 8 (1965), pp. 2182–2189.
[22] A. Harten, P. D. Lax, and B. van Leer, *On upstream differencing and Godunov-type schemes for hyperbolic conservation laws*, SIAM Rev., 25 (1983), pp. 35–61.
[23] J. S. Hesthaven, *Numerical methods for conservation laws*, vol. 18 of Computational Science & Engineering, SIAM, Philadelphia, PA, 2018.
[24] G.-S. Jiang and C.-W. Shu, *Efficient implementation of weighted ENO schemes*, J. Comput. Phys., 126 (1996), pp. 202–228.
[25] D. Kuzmin and H. Hajduk, *Property-preserving numerical schemes for conservation laws*,

World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2024.

[26] M. J. Lockett and K. S. Bassoon, *Sedimentation of binary particle mixtures*, Powder Technol., 24 (1979), pp. 1–7.

[27] J. H. Masliyah, *Hindered settling in a multi-species particle system*, Chem. Eng. Sci., 34 (1979), pp. 1166–1168.

[28] S. J. McCaffery, L. Elliott, and D. B. Ingham, *Two-dimensional enhanced sedimentation in inclined fracture channels*, Math. Engrg. Indust., 7 (1998), pp. 97–125.

[29] J. F. Richardson and W. N. Zaki, *Sedimentation and fluidization I*, Trans. Instn. Chem. Engrs. (London), 32 (1954), pp. 35–53.

[30] F. Rosso and G. Sona, *Gravity-driven separation of oil-water dispersions*, Adv. Math. Sci. Appl., 11 (2001), pp. 127–151.

[31] C.-W. Shu, *High order weighted essentially nonoscillatory schemes for convection dominated problems*, SIAM Review, 51 (2009), pp. 82–126.

[32] F. Zhang and J. Cheng, *A bound-preserving and positivity-preserving finite volume WENO scheme for solving five-equation model of two-medium flows*, Commun. Nonlinear Sci. Numer. Simul., 114 (2022), 106649 (22 pages).

[33] F. Zhang and J. Cheng, *Bound-preserving and positivity-preserving analysis of direct arbitrary Lagrangian-Eulerian finite volume scheme for compressible two-medium flows with stiffened gas equation of state*, SIAM J. Sci. Comput., 47 (2025), pp. B899–B924.

[34] X. Zhang, *On positivity-preserving high order discontinuous Galerkin schemes for compressible Navier–Stokes equations*, J. Comput. Phys., 328 (2017), pp. 301–343.

[35] X. Zhang and C.-W. Shu, *On maximum-principle-satisfying high order schemes for scalar conservation laws*, J. Comput. Phys., 229 (2010), pp. 3091–3120.

[36] X. Zhang and C.-W. Shu, *Positivity-preserving high order discontinuous Galerkin schemes for compressible Euler equations with source terms*, J. Comput. Phys., 230 (2011), pp. 1238–1248.

# Centro de Investigación en Ingeniería Matemática (CI$^2$MA)

## PRE-PUBLICACIONES 2025

2025-05 Fahim Aslam, Zayd Hajjej, Jianghao Hao, Mauricio Sepúlveda: *Global Existence and Asymptotic Profile of an Infinite Memory Logarithmic Wave Equation with Fractional Derivative and Strong Damping*

2025-06 Jessika Camaño, Ricardo Oyarzúa, Katherine Rojo: *A momentum and mass conservative pseudostress-based mixed finite element method for the Stokes problem*

2025-07 Aníbal Coronel, Fernando Huancas, Mauricio Sepúlveda: *Theoretical and numerical approaches of reaction term identification in a SIS reaction diffusion system*

2025-08 Isaac Bermudez, Gabriel N. Gatica, Juan P. Silva: *A new Banach spaces-based mixed finite element method for the coupled Navier–Stokes and Darcy equations*

2025-09 Sergio Caucao, Gabriel N. Gatica, Saulo Medrado, Yuri D. Sobral: *A posteriori error analysis of mixed finite element methods for a regularized mu(I)-rheology model of granular materials*

2025-10 Veronica Anaya, Gerardo Chowell, Felipe Jara, Mauricio Sepúlveda: *Optimal Control of Inter-Population Disease Spread via Reaction–Diffusion Models*

2025-11 Ramiro Acevedo, Rommel Bustinza, Christian Gómez: *A Transient Eddy Current Problem via Potential Formulation with Current Excitation*

2025-12 Abraham J. Arenas, Juan Barajas-Calonge, Gilberto González-Parra, Luis M. Villada: *A second-order nonstandard finite difference scheme for eco-epidemiological predator-prey models*

2025-13 Alonso Bustos, Sergio Caucao, Gabriel N. Gatica: *Mixed-primal and fully-mixed formulations for the convection-diffusion-reaction system based upon Brinkman–Forchheimer equations*

2025-14 Gabriel N. Gatica, Zeinab Gharibi, Ricardo Oyarzúa: *Banach spaces-based fully mixed finite element methods for the n-dimensional Boussinesq problem with temperature-dependent parameters*

2025-15 Sergio Caucao, Ricardo Oyarzúa, Segundo Villa-Fuentes: *A priori and a posteriori error analyses of a fully-mixed finite element method for the coupled Navier-Stokes/Darcy problem*

2025-16 Juan Barajas-Calonge, Raimund Bürger, Pep Mulet, Luis M. Villada: *A second-order invariant-region-preserving scheme for a transport-flow model of polydisperse 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