ON LINEARLY IMPLICIT IMEX RUNGE-KUTTA METHODS FOR DEGENERATE CONVECTION-DIFFUSION PROBLEMS MODELING POLYDISPERSE SEDIMENTATION

SEBASTIANO BOSCARINO[A], RAIMUND BÜRGER[B], PEP MULET[C], GIOVANNI RUSSO[A], AND LUIS MIGUEL VILLADA[D]

Abstract. Implicit-explicit (IMEX) Runge-Kutta (RK) methods are suitable for the solution of nonlinear, possibly strongly degenerate, convection-diffusion problems, since the stability restrictions, coming from the explicitly treated convective part, are much less severe than those that would be deduced from an explicit treatment of the diffusive term. A particularly efficient variant of these schemes, so-called linearly explicit IMEX-RK schemes, arise from discretizing the diffusion terms in a way that more carefully distinguishes between stiff and nonstiff dependence, such that in each time step only a linear system needs to be solved. These schemes provide an efficient tool for the numerical exploration of sediment formation and composition under a strongly degenerate polydisperse sedimentation model.

1. Introduction

This contribution is concerned with numerical methods for systems of nonlinear convection-diffusion equations of the type

$$\partial_t \Phi + \partial_x \boldsymbol{f}(\Phi) = \partial_x \big(\boldsymbol{B}(\Phi) \partial_x \Phi \big), \tag{1.1}$$

where $\Phi = (\phi_1, \ldots, \phi_N)^{\mathrm{T}}$ is the sought solution as a function of spatial position x and time t, $\boldsymbol{f}(\Phi) = (f_1(\Phi), \ldots, f_N(\Phi))^{\mathrm{T}}$ is a vector of flux density functions, and $\boldsymbol{B}(\Phi)$ is a given $N \times N$ matrix function expressing a diffusive correction, where we allow that $\boldsymbol{B}(\Phi) = \boldsymbol{0}$ on a set of nonzero N-dimensional measure, so that (1.1) is possibly strongly degenerate. The system (1.1) is supplied with initial and boundary conditions. Its applications include a model of polydisperse sedimentation, where the diffusion term accounts for sediment compressibility [1], and multiclass traffic flow with anticipation lengths and reaction times (not considered herein; see [2, 6]).

This work is focused on a class of semi-implicit finite difference methods for the solution of (1.1), namely so-called linearly implicit implicit-explicit Runge-Kutta

Date: March 10, 2015.

[A]Department of Mathematics and Computer Science, University of Catania, 95125 Catania, Italy. E-Mail: boscarino@dmi.unict.it, russo@dmi.unict.it.

[B]CI^{2}MA and Departamento de Ingeniería Matemática, Facultad de Ciencias Físicas y Matemáticas, Universidad de Concepción, Casilla 160-C, Concepción, Chile. E-mail: rburger@ing-mat.udec.cl.

[C]Departament de Matemàtica Aplicada, Universitat de València, Av. Dr. Moliner 50, E-46100 Burjassot, Spain. E-Mail: mulet@uv.es.

[D]Departamento de Matemática, Facultad de Ciencias, Universidad del Bío-Bío, Casilla 5-C, Concepción, Chile. E-Mail: lvillada@ubiobio.cl.

(LI-IMEX-RK) methods (see [2, 3]), which were recently presented in [2] as an alternative to nonlinearly implicit implicit-explicit (NI-IMEX-RK) methods introduced in [6] for the same problem. The latter involve the solution of nonlinear algebraic equations by iterative methods. Roughly speaking, it turned out in [2] that for a variety of IMEX pairs of Runge-Kutta schemes with the same mesh size, LI-IMEX methods produce slightly larger errors, but are easier to implement than their NI-IMEX counterparts, and are more efficient in error reduction per CPU time. The purpose of this contribution is twofold: firstly, we summarize the methods introduced in [2] and provide further support of the findings of that paper by numerical experiments with a four-stage IMEX-RK pair of schemes. Secondly, we employ for the first time a four-stage LI-IMEX-RK scheme for original numerical experiments that provide new insight into variants of a model of polydisperse sedimentation introduced in [1].

2. Numerical schemes

2.1. **Semi-discrete formulation and new approach.** Assume that the semi-discrete formulation of (1.1) can be written in vector form as

$$\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}t} = -\frac{1}{\Delta x}(\Delta^- \boldsymbol{f})(\boldsymbol{\Phi}) + \frac{1}{\Delta x^2}\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})\boldsymbol{\Phi}, \tag{2.1}$$

where $\boldsymbol{\Phi} = (\Phi_1(t), \ldots, \Phi_M(t))^{\mathrm{T}}$ is the sought solution vector, where $\Phi_j(t)$ is the approximate solution at spatial position x_j, $\Delta x := x_{j+1} - x_j$ for $j = 1, \ldots, M$ is the uniform grid spacing, $(\Delta^- \boldsymbol{f})(\boldsymbol{\Phi}) \in \mathbb{R}^{NM}$ denotes the vector of numerical flux vector differences associated with the discretization of $\partial_x \boldsymbol{f}(\Phi)$, and $\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi}) \in \mathbb{R}^{(NM)\times(NM)}$ is a block tridiagonal matrix arising from the discretization of $\partial_x(\boldsymbol{B}(\Phi)\partial_x\Phi)$. The precise algebraic forms of $(\Delta^- \boldsymbol{f})(\boldsymbol{\Phi})$ and $\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})$ are provided in [6]. Note that the matrix $\boldsymbol{\mathcal{B}}$ inherits its discontinuous dependence on $\boldsymbol{\Phi}$ from that of $\boldsymbol{B}$ on Φ.

The new approach is based on distinguishing in (2.1) between stiff and non-stiff dependence on the solution vector $\boldsymbol{\Phi}$, and in choosing the time discretization by an implicit and an explicit RK scheme, respectively, of an IMEX pair of schemes accordingly. In the product $\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})\boldsymbol{\Phi}$ the occurrence of the solution $\boldsymbol{\Phi}$ within $\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})$ is considered nonstiff, while that of the factor $\boldsymbol{\Phi}$ is considered stiff. Within LI-IMEX schemes the implicit treatment is applied only to that second factor while NI-IMEX schemes are based on treating implicitly the whole expression $\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})\boldsymbol{\Phi}$. The LI-IMEX approach does not require solutions of *nonlinear* systems (in contrast to NI-IMEX methods), since these methods require only solving a discretized convection-diffusion equation with a *linear* diffusion term in which the matrix $\boldsymbol{\mathcal{B}}$ is given.

2.2. **Spatial discretization and time integrator.** We denote by $\frac{1}{\Delta x}(\Delta^- \boldsymbol{f})(\boldsymbol{\Phi})$ the discrete version of the convective term $\partial_x \boldsymbol{f}(\Phi)$, which is computed using the WENO5-GHLL finite difference scheme [8]. For the diffusive term $\partial_x(\boldsymbol{B}(\Phi)\partial_x\Phi)$ we use the same discretization as in [6]. Denoting $\boldsymbol{\Phi} = (\boldsymbol{\Phi}_1^{\mathrm{T}}, \ldots \boldsymbol{\Phi}_M^{\mathrm{T}})^{\mathrm{T}} \in \mathbb{R}^{MN}$, where $\boldsymbol{\Phi}_i(t) \approx \Phi(x_i, t) \in \mathbb{R}^N$, we can define the $M \times M$ block tridiagonal matrix $\boldsymbol{\mathcal{B}} = \boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})$, with blocks of size $N \times N$, by

$$\big(\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})\boldsymbol{\Phi}\big)_i(t) = \big(\boldsymbol{B}_{i-1/2}\boldsymbol{\Phi}_{i-1} - (\boldsymbol{B}_{i-1/2} + \boldsymbol{B}_{i+1/2})\boldsymbol{\Phi}_i + \boldsymbol{B}_{i+1/2}\boldsymbol{\Phi}_{i+1}\big)(t)$$

for $i = 1, \ldots, M$, where $\boldsymbol{B}_{i+1/2} := \frac{1}{2}(\boldsymbol{B}(\boldsymbol{\Phi}_i) + \boldsymbol{B}(\boldsymbol{\Phi}_{i+1}))$. The terms $\boldsymbol{B}_{i\pm1/2}\boldsymbol{\Phi}_{i\pm1}$ for $i = 1$ and $i = M$ are modified according to the boundary conditions.

The pair of Butcher arrays of IMEX-RK methods for the time integration of (2.1) is given by

$$\mathcal{A}_{\mathrm{ERK}} := \begin{array}{c|c} \tilde{\boldsymbol{c}} & \tilde{\boldsymbol{A}} \\ \hline & \tilde{\boldsymbol{b}}^{\mathrm{T}} \end{array}, \qquad \mathcal{A}_{\mathrm{DIRK}} := \begin{array}{c|c} \boldsymbol{c} & \boldsymbol{A} \\ \hline & \boldsymbol{b}^{\mathrm{T}} \end{array},$$

where $\tilde{\boldsymbol{A}} = (\tilde{a}_{ij})$ (with $\tilde{a}_{ij} = 0$ for all $j \geq i$) and $\boldsymbol{A} = (a_{ij})$ (with $a_{ij} = 0$ for all $j > i$) are the $s \times s$ matrices of the explicit (ERK) and (diagonally) implicit (DIRK) parts of the method, respectively, while $\tilde{\boldsymbol{b}} = (\tilde{b}_1, \ldots, \tilde{b}_s)^{\mathrm{T}}$, $\tilde{\boldsymbol{c}} = (\tilde{c}_1, \ldots, \tilde{c}_s)^{\mathrm{T}}$, $\boldsymbol{b} = (b_1, \ldots, b_s)^{\mathrm{T}}$ and $\boldsymbol{c} = (c_1, \ldots, c_s)^{\mathrm{T}}$ are s-dimensional vectors of real coefficients, with $\tilde{\boldsymbol{c}}$ and $\boldsymbol{c}$ given by the usual relations $\tilde{c}_i = \tilde{a}_{i1} + \cdots + \tilde{a}_{i,i-1}$ and $c_i = a_{i1} + \cdots + a_{ii}$.

Now we rewrite the semidiscrete formulation (2.1) in the form

$$\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}t} = C(\boldsymbol{\Phi}) + D(\boldsymbol{\Phi}), \tag{2.2}$$

$$C(\boldsymbol{\Phi}) := -\frac{1}{\Delta x}(\Delta^- \boldsymbol{f})(\boldsymbol{\Phi}), \quad D(\boldsymbol{\Phi}) := \frac{1}{\Delta x^2}\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi})\boldsymbol{\Phi}. \tag{2.3}$$

An IMEX-RK scheme applied to (2.2) reads as

$$\boldsymbol{\Phi}^{(i)} = \boldsymbol{\Phi}^n + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} C(\boldsymbol{\Phi}^{(j)}) + \Delta t \sum_{j=1}^{i} a_{ij} D(\boldsymbol{\Phi}^{(j)}), \tag{2.4}$$

$$\boldsymbol{\Phi}^{n+1} = \boldsymbol{\Phi}^n + \Delta t \sum_{j=1}^{s} \tilde{b}_i C(\boldsymbol{\Phi}^{(i)}) + \Delta t \sum_{j=1}^{s} b_i D(\boldsymbol{\Phi}^{(i)}). \tag{2.5}$$

We observe that in IMEX-RK schemes there is only one family of stages values. Furthermore, the fact that the implicit part of the method is a *diagonally implicit* RK (DIRK) scheme makes the implementation of an IMEX-RK scheme simpler, and ensures that C is effectively computed explicitly.

Notice that the system (2.2) can be treated by another family of methods, called Additive Runge-Kutta (ARK) schemes [12], and which have the following form:

$$K_i = C\left(\boldsymbol{\Phi}^n + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} K_j\right) + D\left(\boldsymbol{\Phi}^n + \Delta t \sum_{j=1}^{i-1} a_{ij} K_j + \Delta t a_{ii} K_i\right),\ i = 1, \ldots, s, \tag{2.6}$$

$$\boldsymbol{\Phi}^{n+1} = \boldsymbol{\Phi}^n + \Delta t \sum_{i=1}^{s} b_i K_i. \tag{2.7}$$

Next we describe the nonlinearly implicit IMEX-RK methods and the linearly implicit IMEX-RK ones where the first are based on the IMEX-RK framework (2.4), (2.5), while the second are a semi-implicit variant of the ARK approach (2.6), (2.7).

2.3. **Nonlinearly implicit IMEX-RK methods.** To summarize the NI-IMEX-RK method for (1.1) introduced in [6], we consider the semidiscrete formulation (2.2), (2.3). The simplest IMEX scheme for the approximation of (2.1) is

$$\boldsymbol{\Phi}^{n+1} = \boldsymbol{\Phi}^n - \frac{\Delta t}{\Delta x}(\Delta^- \boldsymbol{f})(\boldsymbol{\Phi}^n) + \frac{\Delta t}{\Delta x^2}\boldsymbol{\mathcal{B}}(\boldsymbol{\Phi}^{n+1})\boldsymbol{\Phi}^{n+1}, \tag{2.8}$$

where $\boldsymbol{\Phi}^n$ denotes the approximate value of $\boldsymbol{\Phi}(t)$ at $t = t^n$. In general, the various steps of a NI-IMEX-RK scheme necessary to advance $\boldsymbol{\Phi}^n$ from time t^n to $t^{n+1} = t^n + \Delta t$ in system (2.2), (2.3) are given in [6, Algorithms 3.1 and 4.1]. For the present

discussion it is sufficient to point out that as part of the NI-IMEX-RK framework, (2.4) requires for each stage i, $i = 1, \ldots, s$, of the diagonally implicit part of the underlying IMEX-RK scheme that one solves for the vector $\boldsymbol{u} = \boldsymbol{\Phi}^{(i)} \in \mathbb{R}^{MN}$ a nonlinear system of NM scalar equations of the following form:

$$\boldsymbol{u} - a_{ii}\Delta t D\big(\boldsymbol{u}\big) - \boldsymbol{r}_i = \boldsymbol{0}, \quad i = 1, \ldots, s, \tag{2.9}$$

where $\boldsymbol{r}_i \in \mathbb{R}^{MN}$ is given by

$$\boldsymbol{r}_i = \boldsymbol{\Phi}^n + \Delta t \left(\sum_{j=1}^{i-1} a_{ij} K_j + \sum_{j=1}^{i-1} \tilde{a}_{ij} \tilde{K}_j \right).$$

To approximately solve (2.9) by the Newton-Raphson method we must require the coefficients of $\boldsymbol{B}$, and therefore those of $\boldsymbol{\mathcal{B}}$, to be at least continuously differentiable. However, the model of interest herein does not naturally satisfy this assumption. Therefore $\boldsymbol{B}$ is replaced by a smooth approximation, and instead of attempting to solve (2.9) directly we solve first a smoothed version of (2.9), reduce the smoothing parameter appropriately, use the approximate solution as a starting value for solving the same problem with the new smoothing parameter and so on until that parameter has reached a minimum value [6]. This procedure is, of course, computationally costly but has turned out robust. Note that the smoothing procedure does not remove the possibly strongly degenerate nature of the problem.

2.4. **Linearly implicit IMEX-RK methods.** To overcome the excessive numerical work for the solution of (2.9), an essential gain is obtained by the following approach. We rewrite the semidiscrete formulation (2.1) in the form

$$\frac{\mathrm{d}\boldsymbol{\Phi}}{\mathrm{d}t} = \mathcal{K}(\boldsymbol{\Phi}^*, \boldsymbol{\Phi}), \qquad \mathcal{K}(\boldsymbol{\Phi}^*, \boldsymbol{\Phi}) := C(\boldsymbol{\Phi^*}) + \frac{1}{\Delta x^2} \boldsymbol{\mathcal{B}}(\boldsymbol{\Phi}^*)\boldsymbol{\Phi}, \tag{2.10}$$

with C given in (2.3), and observe that the only stiff term is the linear term $\boldsymbol{\Phi}$ that multiplies $\mathcal{B}(\boldsymbol{\Phi}^*)$, we therefore treat $\boldsymbol{\Phi}^*$ explicitly as argument of $\boldsymbol{f}$ and $\boldsymbol{\mathcal{B}}$, while $\boldsymbol{\Phi}$ is implicit in the term to which $\boldsymbol{\mathcal{B}}$ is applied.

The schemes (2.6) can be extended to the more general equation (2.10). Consider an autonomous equation of the form $\mathrm{d}\boldsymbol{y}/\mathrm{d}t = \mathcal{K}(\boldsymbol{y}^*, \boldsymbol{y})$, where the function $\mathcal{K} : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}^m$ is supposed to be sufficiently differentiable. Suppose that the dependence on the first argument of $\mathcal{K}$ is nonstiff, while that on the second argument is stiff. Such a system can be rewritten in the form

$$\frac{\mathrm{d}\boldsymbol{y}^*}{\mathrm{d}t} = \mathcal{K}(\boldsymbol{y}^*, \boldsymbol{y}), \quad \frac{\mathrm{d}\boldsymbol{y}}{\mathrm{d}t} = \mathcal{K}(\boldsymbol{y}^*, \boldsymbol{y}), \tag{2.11}$$

with $\boldsymbol{y}^*(t_0) = \boldsymbol{y}(t_0) = \boldsymbol{y}_0$. This is a particular case of *partitioned system*, [7], but with an additional computational cost since we double the number of variables.

Now let us apply an IMEX scheme to (2.11). This results in the formulas

$$\boldsymbol{Y}_i^* = \boldsymbol{y}_n^* + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} \mathcal{K}(\boldsymbol{Y}_j^*, \boldsymbol{Y}_j), \quad \boldsymbol{Y}_i = \boldsymbol{y}_n + \Delta t \sum_{j=1}^{i} a_{ij} \mathcal{K}(\boldsymbol{Y}_j^*, \boldsymbol{Y}_j),$$

$$\boldsymbol{y}_{n+1}^* = \boldsymbol{y}_n^* + \Delta t \sum_{i=1}^{s} \tilde{b}_i \mathcal{K}(\boldsymbol{Y}_i^*, \boldsymbol{Y}_i), \quad \boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \Delta t \sum_{i=1}^{s} b_i \mathcal{K}(\boldsymbol{Y}_i^*, \boldsymbol{Y}_i).$$

Observe that if $\tilde{\boldsymbol{b}} = \boldsymbol{b}$, then $\boldsymbol{y}_n^* = \boldsymbol{y}_n$ for all $n \geq 0$ and therefore the duplication of variables is not necessary if we adopt the RK fluxes $K_i = \mathcal{K}(\boldsymbol{Y}_i^*, \boldsymbol{Y}_i)$ as basic

unknowns, so that one can rewrite the scheme in the form

$$K_i = \mathcal{K}\left(\boldsymbol{y}_n + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} K_j, \boldsymbol{y}_n + \Delta t \sum_{j=1}^{i-1} a_{ij} K_j + \Delta t a_{ii} K_i\right), \quad i = 1, \ldots, s, \tag{2.12}$$

with the numerical solution

$$\boldsymbol{y}_{n+1} = \boldsymbol{y}_n + \Delta t \sum_{i=1}^{s} b_i K_i. \tag{2.13}$$

Notice that with the choice $\mathcal{K}(\Phi^*, \Phi) = C(\Phi^*) + D(\Phi)$, the approach reduces to the one proposed by Zhong [12], requiring the solution of nonlinear systems.

Due to the formal equivalence with partitioned systems, order conditions for ARK schemes can be derived from those for IMEX-RK schemes. High order in time can be obtained by adopting IMEX-RK schemes with $\tilde{\boldsymbol{b}} = \boldsymbol{b}$. For a more general description of the relation between IMEX and ARK methods see [4]. A limited linear stability anlysis for LI-IMEX-RK schemes is provided in [2].

The simplest first-order LI-IMEX-RK scheme for (2.12), (2.13) is

$$\boldsymbol{\Phi}^{n+1} = \boldsymbol{\Phi}^n - \frac{\Delta t}{\Delta x}(\Delta^- \boldsymbol{f})(\boldsymbol{\Phi}^n) + \frac{\Delta t}{\Delta x^2} \boldsymbol{\mathcal{B}}(\boldsymbol{\Phi}^n) \boldsymbol{\Phi}^{n+1}.$$

By (2.12) and (2.13) the step from t^n to $t^{n+1} = t^n + \Delta t$ of an LI-IMEX-RK scheme is given in general given by the following algorithm.

Algorithm 2.1 (Linearly implicit IMEX-RK (LI-IMEX-RK) scheme)**.**

Input: approximate solution vector $\boldsymbol{\Phi}^n$ *for* $t = t^n$

do $i = 1, \ldots, s$

compute the stage values:

$$\boldsymbol{\Phi}^{*(i)} \leftarrow \boldsymbol{\Phi}^n + \Delta t \sum_{j=1}^{i-1} \tilde{a}_{ij} K_j, \quad \hat{\boldsymbol{\Phi}}^{(i)} \leftarrow \boldsymbol{\Phi}^n + \Delta t \sum_{j=1}^{i-1} a_{ij} K_j$$

solve for K_i *the linear system*

$$K_i = C\big(\boldsymbol{\Phi}^{*(i)}\big) + \frac{1}{\Delta x^2} \boldsymbol{\mathcal{B}}\big(\boldsymbol{\Phi}^{*(i)}\big)\Big(\hat{\boldsymbol{\Phi}}^{(i)} + \Delta t a_{ii} K_i\Big) \tag{2.14}$$

enddo

$$\boldsymbol{\Phi}^{n+1} \leftarrow \boldsymbol{\Phi}^n + \Delta t \sum_{j=1}^{s} b_j K_j$$

Output: approximate solution vector $\boldsymbol{\Phi}^{n+1}$ *for* $t = t^{n+1} = t^n + \Delta t$.

In this approach, the system (2.14) is linear in K_i and the numerical solution can be obtained by solving a convection-diffusion equation with a linear diffusion term in which the matrix function $\boldsymbol{B}$, and therefore $\boldsymbol{\mathcal{B}}$, is computed explicitly.

For our computations we use the Strongly-Stability Preserving (SSP) four-stage third-order IMEX scheme SSP3(4,3,3) [9] described by the following Butcher arrays:

$$\boldsymbol{\mathcal{A}}_{\text{ERK}} = \begin{array}{c|cccc} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 \\ \frac{1}{2} & 0 & \frac{1}{4} & \frac{1}{4} & 0 \\ \hline & 0 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{array}, \quad \boldsymbol{\mathcal{A}}_{\text{DIRK}} = \begin{array}{c|cccc} \alpha & \alpha & 0 & 0 & 0 \\ 0 & -\alpha & \alpha & 0 & 0 \\ 1 & 0 & 1-\alpha & \alpha & 0 \\ \frac{1}{2} & \beta & \eta & \frac{1}{2}-\beta-\eta-\alpha & \alpha \\ \hline & 0 & \frac{1}{6} & \frac{1}{6} & \frac{2}{3} \end{array},$$

where $\alpha = 0.24169426078821$, $\beta = \alpha/4$, $\eta = 0.12915286960590$.

3. Model and variants

3.1. **Basic framework.** We briefly outline model of sedimentation of polydisperse suspensions forming compressible sediments described in detail in [1] (see also [2, 6]). We consider a suspension of equal-density particles belonging to N species with sizes $d_1 > d_2 > \ldots > d_N$. We let ϕ_i denote the local volume fraction of species i having size d_i. The evolution of $\Phi = \Phi(x,t)$ as a function of depth x and time t in a column of height $\mathcal{L}$ is then determined by (1.1) posed on the x-interval $(0,\mathcal{L})$ for $t > 0$, along with $\Phi(x,0) = \Phi_0(x)$ for $0 \le x \le \mathcal{L}$, where Φ_0 is the given initial concentration distribution, and zero-flux boundary conditions, i.e.,

$$\boldsymbol{f}(\Phi) - \boldsymbol{B}(\Phi)\partial_x\Phi = \boldsymbol{0} \quad \text{for } x = 0 \text{ and } x = \mathcal{L}, \quad t > 0. \tag{3.1}$$

The flux density functions $f_1, \ldots, f_N$ are given by

$$f_i(\Phi) = \mu\bar{\varrho}_\mathrm{s}\phi_i V(\phi)(1-\phi)(\delta_i - \boldsymbol{\delta}^\mathrm{T}\Phi), \quad i = 1, \ldots, N, \tag{3.2}$$

where $\mu > 0$ is a constant, $\bar{\varrho}_\mathrm{s} > 0$ is the solid mass density minus the fluid density, $\delta_i := d_i^2/d_1^2$, $\boldsymbol{\delta} := (\delta_1, \delta_2, \ldots, \delta_N)^\mathrm{T}$, $\delta_1 = 1$, $\phi := \phi_1 + \cdots + \phi_N$, and $V(\phi)$ is a hindered settling function that can be chosen as $V(\phi) = (1-\phi)^{n_\mathrm{RZ}-2}$ for $0 \le \phi \le \phi_\mathrm{max}$ and $V(\phi) = 0$ otherwise, where $n_\mathrm{RZ} > 2$ is a material-dependent exponent.

The diffusion matrix is given by $\boldsymbol{B}(\Phi) := (\alpha_{ij})_{1\le i,j\le N}$ where

$$\begin{aligned}\alpha_{ij} = \frac{\mu V(\phi)}{g\phi}\Bigg\{&(1-\phi)\phi_i(\delta_i - \boldsymbol{\delta}^\mathrm{T}\Phi)\sigma_\mathrm{e}'(\phi) \\ &- \left[\delta_i\delta_{ij} - \delta_j\phi_i - \frac{\phi_i}{\phi}(\delta_i - \boldsymbol{\delta}^\mathrm{T}\Phi)\right]p(\phi)\Bigg\}, \quad i,j = 1,\ldots,N,\end{aligned} \tag{3.3}$$

where δ_{ij} is the standard Kronecker symbol and σ_e denotes the effective solid stress function, and σ_e' is its derivative. This function can be chosen as

$$\sigma_\mathrm{e}(\phi) = \begin{cases} 0 & \text{for } \phi \le \phi_\mathrm{c}, \\ \sigma_0\big((\phi/\phi_\mathrm{c})^k - 1\big) & \text{for } \phi > \phi_\mathrm{c}, \end{cases} \quad \sigma_0, k > 0, \tag{3.4}$$

where ϕ_c is a critical concentration at which the particles touch each other.

3.2. **Model variants.** The function $p(\phi)$ can be chosen as $p(\phi) = \sigma_\mathrm{e}$. This is precisely the model introduced and analyzed in [1], and to which we refer as "Model A". Roughly speaking, the term involving σ_e' models the effect of sediment compressibility, while that involving $p(\phi)$ describes differential diffusion of particle species within the sediment layer. The numerical results presented in [1] and re-calculated in [2] by an IMEX method showed that Model A, although it is rigorously derived from balance equations of continuum mechanics, presents an unrealistically vigorous relative differential movement of particles within the sediment. As a consequence of this excessive sediment diffusivity, the final (large-time) stationary composition of the sediment exhibits a gently downward-increasing total volume fraction ϕ, but the relative volume fraction ϕ_i/ϕ of each species is the same at each position.

In [1, Sect. 7.5] it was speculated, though not proved rigorously nor explored numerically, that several variants of Model A, denoted here Models B and C, would not exhibit this phenomenon. We here consider "Model B", corresponding to $p \equiv 0$ and which is in part motivated by the treatment in [10], and "Model C", corresponding to $p(\phi) = \sigma_\mathrm{e}(\phi)$ for $\phi \le \phi^*$ and $p(\phi) = 0$ for $\phi > \phi^*$, where $\phi^* > \phi_\mathrm{c}$ is a second critical concentration. Here ϕ^* plays the role of a concentration at which there is a change in sedimentation behaviour from differential sedimentation to "en

masse settling of the entire suspension" [11]. Note that for a rigorous implementation of the latter idea we would also have to change the convective flux at $\phi = \phi^*$; unfortunately this is outside the scope of this paper.

4. Numerical examples

4.1. **Preliminaries.** We solve (1.1) numerically for $0 \le t \le T$ and $0 \le x \le 1$ for Models A, B and C. Numerical results are obtained by the LI-SSP3(4,3,3) scheme. The x-interval $[0,1]$, which is aligned in vertical direction and corresponds to a column of height 1 m, is subdivided into M subintervals of length $\Delta x = 1/M$. We denote by Δt the time step used to advance the numerical solution from $t = t^n$ to $t^{n+1} = t^n + \Delta t$ and by Φ_j^n the vector of numerical solutions associated with cell $[j\Delta x, (j+1)\Delta x]$, $j = 0, \ldots, M-1$, at time t^n. For each iteration, Δt is determined by the following formula (derived from a linearized CFL condition, where $\varrho(\cdot)$ is the spectral radius: $(\Delta t/\Delta x) \max_{1\le j\le M} \varrho(\mathcal{J}_{\boldsymbol{f}}(\Phi_j^n)) = 0.7$.

4.2. **Example 1-2: Comparing Models A, B, C.** We compare Models A, B and C simulating the settling of a tridisperse ($N = 3$) suspension forming a compressible sediment [1]. The mixture is described by the model functions (3.2)–(3.4) with $\phi_{\max} = 0.66$, $n_{\mathrm{RZ}} = 4.7$, $\sigma_0 = 180\,\mathrm{Pa}$, $\phi_{\mathrm{c}} = 0.2$, $\phi^* = 0.25$, $k = 2$, $\mu = gd^2/(18\mu_{\mathrm{f}})$ where $\mu_{\mathrm{f}} = 10^{-3}\,\mathrm{Pa\,s}$ is the viscosity of the fluid, $d = 1.19\times10^{-5}\,\mathrm{m}$, $\bar{\rho}_{\mathrm{s}} = 1800\,\mathrm{kg/m^3}$ is the solid-fluid density difference, and $g = 9.81\,\mathrm{m/s^2}$ is the acceleration of gravity. The normalized squared particle sizes are $\boldsymbol{\delta} = (1, 0.5, 0.25)^{\mathrm{T}}$, and we employ zero-flux boundary conditions (3.1) in a vessel of height 1 m. In Example 1, the initial concentration is $\Phi_0(x,0) = (0.04, 0.04, 0.04)^{\mathrm{T}}$ for $0 \le x \le 1\,\mathrm{m}$, and in Example 2, we consider $\hat{\Phi}_0(x,0) = (0.2, 0.2, 0.2)^{\mathrm{T}}/3$ for $0 \le x \le 0.6\,\mathrm{m}$ and $\hat{\Phi}_0 = 0$ otherwise.

In Figure 1 we display the numerical solution obtained with the LI-SSP3(4,3,3) scheme with $\Delta x = 1/3200\,\mathrm{m}$ with initial condition $\Phi(x,0)$ for each models at three simulated times. Figure 2 we display the numerical solution for the alternative initial composition $\hat{\Phi}(x,0)$.

4.3. **Example 3: Efficiency of NI/LI-SSP3(4,3,3) schemes for Model C.** In this example we compare the efficiency of NI/LI-SSP3(4,3,3) schemes for obtaining the numerical solution in model C at simulated times $T_1 = 5000\,\mathrm{s}$ and $T_2 = 30000\,\mathrm{s}$. We consider the same datas as in Example 1. As reference solution we consider the solution computed with NI-SSP2(3,3,2) scheme with $\Delta x = 1/6400\,\mathrm{m}$. For the NI schemes we solve the nonlinear system by Algorithm 4.1 of [6] that uses a variant of the Newton-Raphson method with a prescribed relative tolerance tol, where the regularization $\boldsymbol{B}_\varepsilon$ of the original diffusion matrix $\boldsymbol{B}$ is achieved by replacing the function σ_{e} in (3.3) by $\sigma_{\mathrm{e}}(\phi;\varepsilon) = \sigma_{\mathrm{e}}(\phi)\exp(-\varepsilon/(\phi-\phi_{\mathrm{c}})^2 - \varepsilon/(\phi-\phi^*)^2)$, where $\varepsilon > 0$ decreases gradually from $\varepsilon_0 = 10^{-3}$ to $\varepsilon_{\min} = 10^{-6}$, $\mathrm{tol} = 10^{-8}$. Figures 3 (a) and (b) provide the corresponding efficiency plots (approximate L^1 errors versus CPU time).

4.4. **Conclusions.** In [2] it is concluded, in light of the numerical experiments conducted in that paper, that the new LI-IMEX-RK schemes yield very similar results, independently of the choice of the underlying Runge-Kutta schemes, that they approximate the same solutions as their NI counterparts (introduced in [6]), and that in many cases they are more efficient. The present provides further support of these findings since these schemes are tested herein for new models (Models B and C), an alternative WENO-type discretization of the convective term (we here

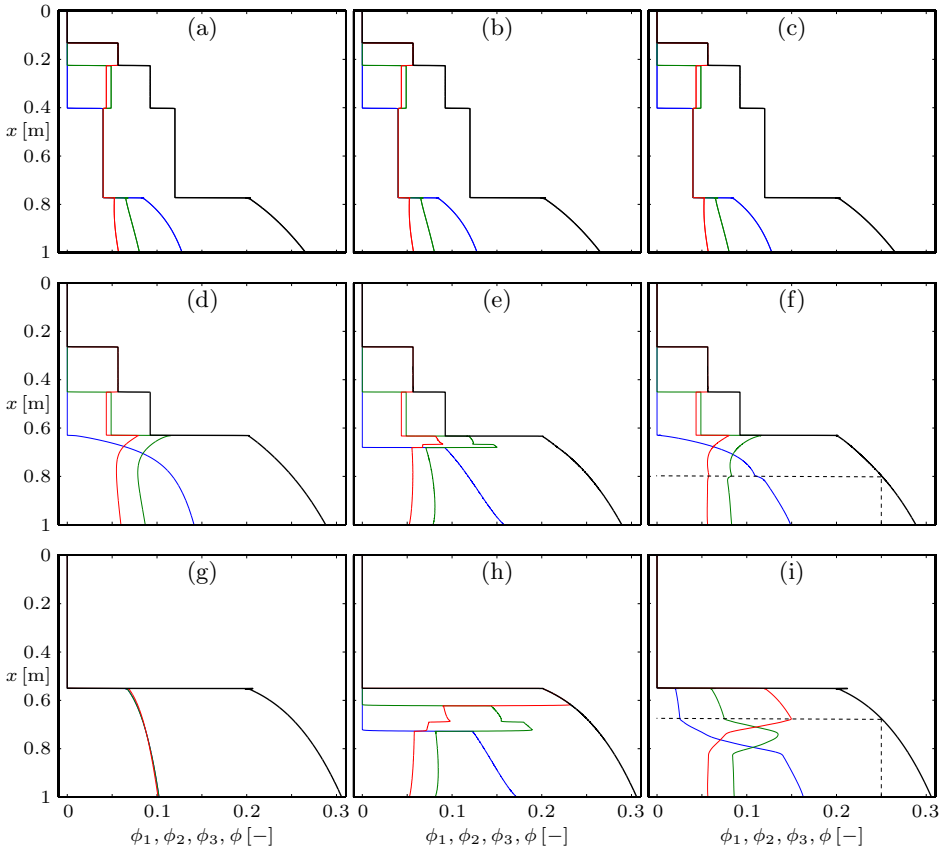

FIGURE 1. Example 1: numerical solution by LI-SSP3(4,3,3) scheme with $\Delta x = 1/3200$ m for solving Model A (a, d, g), Model B (b, e, h) and Model C (c, f, i) at simulated time (a, b, c) $T = 5000$ s, (d, e, f) $T = 10000$ s and (g, h, i) $T = 50000$ s.

employ the scheme WENO5-GHLL of [8] while in [2] the scheme WENO-SPEC of [5] is utilized), and a fourth-order IMEX-RK scheme not considered before in [2, 6]. The results of Figure 3 indicate that the error produced by LI schemes for a given discretization is larger than that of the corresponding NI counterpart (as expected), however they are still competitive since they are at least as efficient in error reduction per CPU time than their corresponding NI counterparts. In fact, LI schemes have been designed to execute faster than their respective NI counterparts since the former need to solve only one linear system per RK stage, whereas the latter have to solve many during the nonlinear solves in Algorithm 4.1 of [6]. Thus, the findings of Figure 3 are consistent with those of [2].

Concerning Model A and its variants, we have seen that several semi-empirical model approaches [1, 10, 11] postulate critical values of ϕ associated with abrupt changes in the definition of $\boldsymbol{B}(\Phi)$. Both NI-IMEX-RK [6] and LI-IMEX-RK [2] schemes are tailored to handle the numerical solution of (1.1) under this assumption. However, efficiency is a critical issue in this problem class. In fact, one wishes not only to capture the transient dynamics of the settling process (typically governed

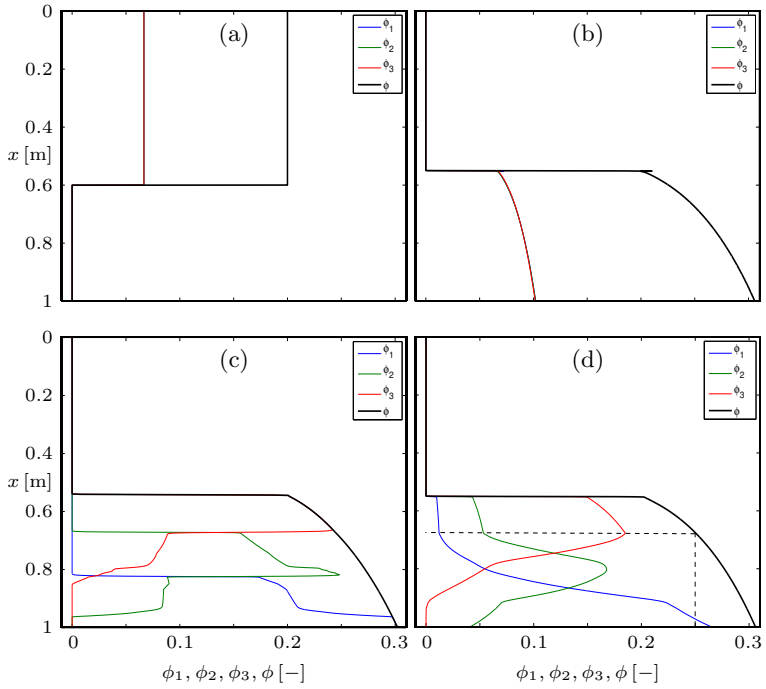

FIGURE 2. Example 2: (a) initial condition, (b, c, d) numerical solution with LI-SSP3(4,3,3) scheme with $\Delta x = 1/1600$ m for solving (b) Model A, (c) Model B and (d) Model C at simulated time $T = 60000$ s.

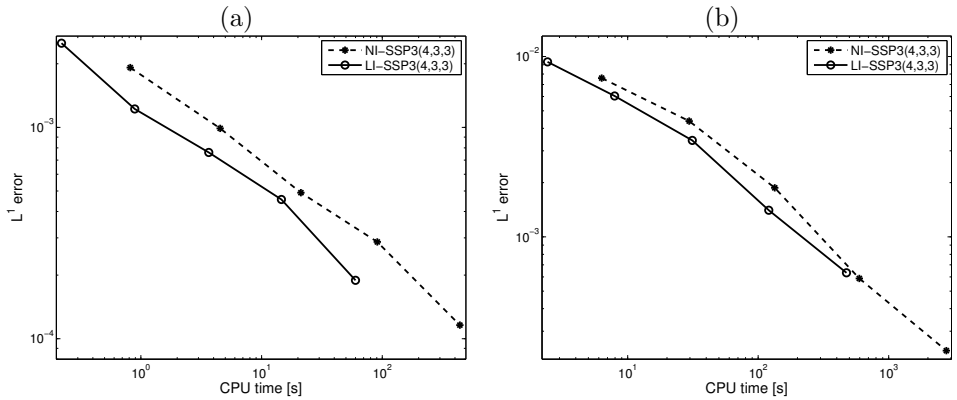

FIGURE 3. Example 3 (Model C): approximate L^1 errors vs. CPU time for NI/LI-SSP3(4,3,3) schemes with $\Delta x = 1/M$ m with $M =$ 100, 200, 400, 800 and 1600. (a) $T = 5000$ s, (b) $T = 30000$ s.

for small times by the behaviour of (1.1) where $\phi < \phi_c$, i.e., where the governing equation degenerates into first-order hyperbolic type), but also the composition of the final sediment, which usually requires long-time simulations. We observe

that the profiles generated by Models A, B and C at $T = 5000\,\mathrm{s}$ are nearly the same (Figure 1 (a), (b) and (c)), but that the sediment composition after long time is different for each of the models (Figures 1 (g), (h) and (i)). (We have also performed tests for $T = 200000\,\mathrm{s}$. Results virtually coincides with those for $T = 60000\,\mathrm{s}$ so we can say that Figures 1 (g), (h) and (i) practically display a final state.) Another remarkable result is that while the final sediment composition under Model A seems to depend only on the initial total mass of ϕ_i for each species i (Figures 1 (a)) and 2 (b) almost coincide), for Models B and C we obtain different sediment compositions for different choices of $\phi_i(\cdot, 0)$ having the same respective total mass. Clearly these phenomena should be further explored.

Acknowledgements

RB is supported by projects Fondecyt 1130154; BASAL CMM, U. de Chile and CI^{2}MA, U. de Concepción; CRHIAM, CONICYT/FONDAP/15130015; Conicyt Anillo ACT1118 (ANANUM); Red Doctoral REDOC.CTA, MINEDUC UCO1202. PM is partially supported by Spanish MINECO project MTM2011-22741. LMV is supported by Fondecyt project 11140708. GR and SB have been partially supported by project PRIN 2009, "Innovative numerical methods of hyperbolic problems with aplications to fluid dynamics, kinetic theory and computational biology", prot. n. 2009588FHJ, of the Italian Ministry of Education, University and Research.

References

[1] 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), 41–80.

[2] S. Boscarino, R. Bürger, P. Mulet, G. Russo and L.M. Villada. *Linearly implicit IMEX Runge-Kutta Methods for a class of degenerate convection-diffusion problems.* SIAM J. Sci. Comput., to appear.

[3] S. Boscarino, P.G. LeFloch, and G. Russo. *High order asymptotic-preserving methods for fully nonlinear relaxation problems.* SIAM J. Sci. Comput., **36** (2014), A377–A395.

[4] S. Boscarino, F. Filbet and G. Russo. *High order semi-implicit schemes for time dependent partial differential equations.* In revision.

[5] 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), 2322–2344.

[6] R. Bürger, P. Mulet, and L.M. Villada. *Regularized nonlinear solvers for IMEX methods applied to diffusively corrected multi-species kinematic flow models.* SIAM J. Sci. Comput., **35** (2013), B751–B777.

[7] E. Hairer, S.P. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations. I. Non-stiff problems*, Springer Series in Comput. Mathematics, Vol. 8, Springer Verlag, (2nd edition), 1993.

[8] M.C. Martí and P. Mulet. *Some techniques for improving the resolution of finite difference component-wise WENO schemes for polydisperse sedimentation models.* Appl. Numer. Math., **78** (2014), 1–13.

[9] L. Pareschi and G. Russo. *Implicit-Explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation.* J. Sci. Comput., **25** (2005), 129–155.

[10] Y.T. Shih, D. Gidaspow and D.T. Wasan. *Hydrodynamics of sedimentation of multisized particles.* Powder Technol., **50** (1987), 201–215.

[11] J. Zeng and D.R. Lowe. *A numerical method for sedimentation from highly-concentrated multi-sized suspensions.* Math. Geol., **24** (1992) 393–415.

[12] X. Zhong. *Additive semi-implicit Runge-Kutta methods for computing high-speed nonequilibrium reactive flows.* J. Comput. Phys., **128** (1996), 19–31.