

UNIVERSIDAD DE CONCEPCIÓN
DIRECCIÓN DE POSTGRADO
CONCEPCIÓN-CHILE



HIPERBOLICIDAD Y SOLUCIÓN NUMÉRICA DE CIERTOS MODELOS
DE SEDIMENTACIÓN POLIDISPERSA

*Tesis para optar al grado de
Doctor en Ciencias Aplicadas con mención en Ingeniería Matemática*

Carlos Arturo Vega Fuentes

FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

2010

HIPERBOLICIDAD Y SOLUCIÓN NUMÉRICA DE CIERTOS MODELOS DE SEDIMENTACIÓN POLIDISPERSA

Carlos Arturo Vega Fuentes

Directores de Tesis: Prof. Raimund Bürger, Universidad de Concepción, Chile.

Prof. Pep Mulet, Universitat de València, España.

Director de Programa: Prof. Raimund Bürger, Universidad de Concepción, Chile.

COMISIÓN EVALUADORA

Dr. Manuel J. Castro, Universidad de Málaga, España.

Dr. Rosa Donat, Universidad de Valencia, España.

Dr. Carlos E. Mejía, Universidad Nacional de Colombia, Sede Medellín.

Dr. Christian Schaerer, Universidad Nacional de Asunción, Paraguay.

COMISIÓN EXAMINADORA

Firma: _____

Dr. Raimund Bürger (Profesor Guía)

Universidad de Concepción, Chile.

Firma: _____

Dr. Gabriel Gatica

Universidad de Concepción, Chile.

Firma: _____
Dr. Oscar Link
Universidad de Concepción, Chile.

Firma: _____
Dr. Marko Rojas-Medar
Universidad del Bío-Bío, Sede Chillán.

Firma: _____
Dr. Mauricio Sepulveda
Universidad del Concepción, Chile.

Fecha Examen de Grado: _____

Calificación: _____

AGRADECIMIENTOS

Han sido años de muchas experiencias positivas en Chile, tanto en el plano académico, como en el ámbito personal. Ha llegado el momento de agradecer a muchas personas e instituciones que de una u otra manera aportaron para que mis estudios doctorales llegaran a feliz término.

Expreso primero mi especial agradecimiento a los profesores Raimund Bürger y Pep Mulet, quienes han dirigido con acierto, dedicación y paciencia esta tesis doctoral y a la vez han compartido conmigo sus conocimientos durante todo este proceso.

A los miembros del Departamento de Ingeniería Matemática, en particular, a los profesores que orientaron los cursos y seminarios en el doctorado.

A la familia del CI²MA y a su director el profesor Gabriel Gatica. A los compañeros del programa de doctorado, muy especialmente, a los amigos con quienes comencé mis estudios.

A los profesores Rosa Donat de la Universitat de València, quien junto con Pep me permitieron realizar un par de estancias en Valencia; Guillermo Cervantes de la Universidad del Norte, por su decidido apoyo al inicio de este proceso.

Mi agradecimiento también a las diferentes instituciones financiadoras: la Comisión Nacional de Investigación Científica y Tecnológica de Chile (CONICYT), Universidad del Norte (Colombia), Proyecto MECESUP UCO013 y el CMM de la Universidad de Chile Proyecto Basal PFB 03 C.C. 2420.

Por último, pero no menos importante, expreso mi gratitud a mi familia: mi esposa Maria Emilia y mi hijo Rafael José, mis padres Paulina y Rafael, mis sobrinos Maria Lucía y Rafael Augusto y a mis hermanas Narlys y Corina.

Al Creador, por todo lo que me ha brindado.

A la memoria de mis abuelos Amelia, Rosa, José María y Luis Claudio y del amigo Pedro René Quintrel (Don René).

Resumen

En éste trabajo de tesis se realiza un estudio de la hiperbolicidad estricta para algunos modelos de sedimentación polidispersa que conducen a un sistema unidimensional de N leyes de conservación no lineales y fuertemente acopladas. A partir del hecho de que la función de flujo para los modelos considerados, se puede expresar en términos de un número pequeño (con respecto al número de especies N) de funciones escalares que dependen sólo del vector de concentraciones, se obtiene que la matriz Jacobiana del sistema posee una estructura particular, que permite identificar sus valores propios con las raíces de una función racional $R(\lambda)$ estudiada previamente en las referencias [1] y [26].

Además de obtener información cualitativa acerca de los valores propios, se obtiene una manera de localizarlos y aproximarlos numéricamente. De hecho, se provee toda la información característica necesaria para realizar las simulaciones numéricas con métodos robustos de alta resolución, en particular, el popular método WENO (*Weighted Essentially Non-oscillatory*) de quinto orden. La pertinencia y las ventajas de éste método, implementado utilizando la información característica en forma intensiva, se ilustra con una considerable cantidad de ejemplos numéricos.

Contents

| | |
|---|-----------|
| Resumen | xi |
| 1 Introducción | 1 |
| 1.1 Versión en español. | 1 |
| 1.2 English version. | 6 |
| 2 General framework | 11 |
| 3 Hyperbolicity analysis of polydisperse sedimentation models via a secular equation for the flux Jacobian | 15 |
| 3.1 Introduction | 15 |
| 3.1.1 Scope | 15 |
| 3.1.2 Related work | 17 |
| 3.1.3 Outline of this chapter | 19 |
| 3.2 The secular equation | 20 |
| 3.3 Kinematic models of polydisperse sedimentation | 24 |
| 3.3.1 The Masliyah-Lockett-Bassoon (MLB) model | 24 |
| 3.3.2 Models based on Batchelor’s approach | 25 |
| 3.4 Hyperbolicity analysis via the secular equation | 27 |
| 3.4.1 The Masliyah-Lockett-Bassoon (MLB) model | 27 |
| 3.4.2 Preliminaries for the BW and HS models | 30 |
| 3.4.3 The Batchelor and Wen (BW) model | 32 |
| 3.4.4 The Höfler and Schwarzer (HS) model | 35 |
| 3.5 Numerical examples | 41 |
| 3.5.1 The Roe and KT schemes | 41 |
| 3.5.2 Numerical examples | 44 |

| | | |
|----------|---|------------|
| 3.6 | Conclusions of Chapter 3 | 45 |
| 4 | On the implementation of WENO schemes for a class of polydisperse sedimentation models | 49 |
| 4.1 | Introduction | 49 |
| 4.1.1 | Scope | 49 |
| 4.1.2 | Related work | 50 |
| 4.2 | Hyperbolicity analysis for a variation of the HS model | 53 |
| 4.3 | Numerical schemes | 54 |
| 4.3.1 | Semi-discrete schemes | 54 |
| 4.3.2 | Boundary conditions | 55 |
| 4.3.3 | Time discretization | 56 |
| 4.3.4 | Flux vector splittings and viscosity coefficients | 56 |
| 4.3.5 | The SPEC and COMP schemes | 58 |
| 4.4 | Numerical results | 61 |
| 4.4.1 | Example 4 (MLB model, $N = 2$) | 62 |
| 4.4.2 | Example 5 (MLB model, $N = 4$) | 66 |
| 4.4.3 | Example 6 (MLB model, $N = 11$) | 68 |
| 4.4.4 | Example 7 (MLB model, $N = 2$) | 69 |
| 4.4.5 | Example 8: HS model with $N = 2$ and $N = 4$ | 71 |
| 4.5 | Conclusions of Chapter 4 | 72 |
| 5 | Additional results of Hyperbolicity | 79 |
| 5.1 | Introduction | 79 |
| 5.2 | The Davis and Gecol (DG) model | 79 |
| 5.3 | BW and HS models for $\beta_3 < 0$ | 87 |
| 5.3.1 | The Batchelor and Wen (BW) model | 89 |
| 5.3.2 | The Höfler and Schwarzer (HS) model | 94 |
| 6 | Conclusiones generales y trabajo futuro | 103 |
| 6.1 | Conclusiones | 103 |
| 6.2 | Limitaciones del enfoque de este trabajo | 104 |
| 6.3 | Trabajo futuro | 104 |

List of Figures

| | | |
|-----|---|----|
| 3.1 | Maximum total concentrations ϕ^* for which hyperbolicity of the BW model is ensured (a) for the coefficients (3.3.13) (with $\beta_3 = 0$), (b) for $\beta_0 = -3.5$, $\beta_1 = -1.1$, $\beta_2 = -1.02$ (according to (3.3.12) for large Péclet numbers) and $\beta_3 = 0$, and (c) for $\beta_0 = -3.42$, $\beta_1 = -1.96$, $\beta_2 = -1.21$ (according to (3.3.12) for small Péclet numbers) and $\beta_3 = 0$ | 36 |
| 3.2 | Region of hyperbolicity ($H(\phi, \boldsymbol{\beta}, d_N) < 0$) for the HS model (a) for the coefficients (3.3.13) (with $\beta_3 = 0$), (b) for $\beta_0 = -3.5$, $\beta_1 = -1.1$, $\beta_2 = -1.02$ (according to (3.3.12) for large Péclet numbers) and $\beta_3 = 0$, and (c) for $\beta_0 = -3.42$, $\beta_1 = -1.96$, $\beta_2 = -1.21$ (according to (3.3.12) for small Péclet numbers) and $\beta_3 = 0$ | 41 |
| 3.3 | Example 1 (MLB and HS models with $N = 2$): numerical solution at $t = 50$ s (top), $t = 300$ s (middle) and $t = 1000$ s (bottom). The horizontal axis in each plot denotes concentration, and the vertical axis denotes normalized depth. Here and in Figure 3.4 the left and middle column correspond to the Roe and KT scheme, respectively, applied to the MLB model and the right column corresponds to the KT scheme applied to the HS model. | 42 |
| 3.4 | Example 2 (MLB and HS models, $N = 4$): numerical solution at $t = 50$ s (top row), $t = 200$ s (second row), $t = 300$ s (third row) and $t = 1000$ s (bottom row) with $\Delta x = 0.0005$ | 43 |
| 3.5 | Example 3 (MLB and HS models, $N = 11$): numerical solution at $t = 600$ s (first column) and $t = 1000$ s (second column) with $\Delta x = 0.0005$. The top and middle rows corresponds to MLB model using Roe and KT schemes respectively. The bottom row corresponds to HS model with KT scheme. | 46 |

| | | |
|------|--|----|
| 4.1 | Example 4: numerical solution for ϕ_1, ϕ_2 (a, c) and ϕ (b, d) at $t = 50$ s computed by SPEC-INT with $M = 6400$ (a, b) and COMP-GLF with $M = 25600$ (c, d). | 60 |
| 4.2 | Example 4: numerical solution for ϕ_1, ϕ_2 (a, c) and ϕ (b, d) at $t = 300$ s computed by SPEC-INT with $M = 6400$ (a, b) and COMP-GLF with $M = 25600$ (c, d). | 61 |
| 4.3 | Example 4: numerical solution at $t = 50$ s with $M = 400$ (a) and enlarged views (b–f). The reference solution is computed by SPEC-INT with $M = 6400$. | 63 |
| 4.4 | Example 4: numerical solution at $t = 300$ s with $M = 400$ (a) and enlarged views (b–f). The reference solution is computed by SPEC-INT with $M = 6400$. | 64 |
| 4.5 | Example 5: reference solution for ϕ_1, \dots, ϕ_4 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c–f), at $t = 50$ s. | 66 |
| 4.6 | Example 5: reference solution for ϕ_1, \dots, ϕ_4 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c–f), at $t = 300$ s. | 67 |
| 4.7 | Example 5: details of numerical solutions obtained by SPEC-INT and SPEC-LLF with $M = 400$ at $t = 300$ s. | 68 |
| 4.8 | Example 6: reference solution for ϕ_1, \dots, ϕ_{11} and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c–f), at $t = 300$ s. | 69 |
| 4.9 | Example 7: reference solution for ϕ_1, ϕ_2 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c–f), at $t = 50$ s. | 70 |
| 4.10 | Example 7: reference solution for ϕ_1, ϕ_2 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c–f), at $t = 250$ s. | 71 |
| 4.11 | Example 8: numerical solution for ϕ_1, ϕ_2 with $M = 400$ (a) and ϕ (b) at $t = 50$ s and enlarged views (c–f) of zones indicated by rectangles in plot (a). The reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$ | 74 |
| 4.12 | Example 8 (HS variant with $N = 2$): numerical solution for ϕ_1, ϕ_2 with $M = 400$ (a) at $t = 250$ s and enlarged views (b–d). The reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$ | 75 |

| | | |
|------|--|-----|
| 4.13 | Example 8 (HS variant with $N = 4$): numerical solution for ϕ_1, \dots, ϕ_4 with $M = 400$ (a) and ϕ (b) at $t = 250$ s and enlarged views (c–f), where the reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$ | 76 |
| 4.14 | Example 8 (HS with $N = 4$): numerical solution for ϕ_1, \dots, ϕ_4 with $M = 400$ (a) and ϕ (b) at $t = 250$ s and enlarged views (c–f), where the reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$ | 77 |
| 5.1 | Region of hyperbolicity ($\mathcal{S}(\phi, \boldsymbol{\beta}, d_N) < 0$) for the DG model (left) and the diagram with the region where the condition (5.2.6) holds (right). The components of $\boldsymbol{\beta}$ used in the calculations are given by (3.3.13). | 83 |
| 5.2 | Davis-Gecol model with $N = 2$: numerical solution for ϕ_1, ϕ_2 and ϕ at $t = 50$ s (a, b), and the corresponding solutions at $t = 300$ s (c, d). | 85 |
| 5.3 | Numerical solutions for DG model with $N = 2$ (top row) at $t = 50$ s (a) and $t = 300$ s (b). Bottom row corresponds to approximate solutions for $N = 4$ at $t = 50$ s (c) and $t = 300$ s (d). | 86 |
| 5.4 | Maximum total concentrations ϕ^* for which hyperbolicity of the BW model is ensured with the coefficients (3.3.12): for large Péclet numbers (left) and for small Péclet numbers (right). | 94 |
| 5.5 | Regions of hyperbolicity ($H(\phi, \boldsymbol{\beta}, d_N) < 0$ and $M(\phi, \boldsymbol{\beta}, d_N) < 0$) for the HS model: with the coefficients (3.3.12) for large Péclet numbers (left), and with the coefficients (3.3.12) for small Péclet numbers(right). | 101 |

List of Tables

| | | |
|-----|--|----|
| 3.1 | MLB and HS models, $N = 11$: initial concentrations ϕ_i^0 , real particle sizes D_i , and normalized particle sizes d_i | 45 |
| 4.1 | Example 4: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr). The reference solution is computed by SPEC-INT with $M = 6400$ | 62 |
| 4.2 | Example 4: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr). The reference solution is computed by COMP-GLF with $M = 25600$ | 65 |
| 4.3 | Example 5: approximate L^1 -errors ($\times 10^{-5}$) and convergence rates (cr). | 66 |
| 4.4 | Example 6: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr). | 69 |
| 4.5 | Example 7: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr) for Riemann problem. The reference solution is computed by SPEC-INT with $M_{\text{ref}} = 6400$ | 72 |
| 4.6 | Example 8: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr) for HS model. The reference solution is computed by SPEC-INT with $M_{\text{ref}} = 6400$ | 73 |

Chapter 1

Introducción

1.1 Versión en español.

En términos generales los modelos para la sedimentación de suspensiones polidispersas contemplan pequeñas partículas de forma esférica dispersas en un fluido viscoso, de tal forma que cada partícula pertenece a una de las N especies, las cuales difieren por su tamaño o densidad. La sedimentación polidispersa aparece en múltiples aplicaciones tales como: procesamiento de minerales, tratamiento de aguas residuales, diseño de espesadores y clarificadores, la industria petrolera, ingeniería química y medicina. Muchos investigadores [4, 5, 6, 12, 17, 23, 35, 36, 38, 42, 45, 46, 47, 49, 52, 58, 64, 68] han estudiado los procesos de sedimentación tanto desde el punto de vista teórico como experimental, en particular, las suspensiones polidispersas.

Bajo la hipótesis de que el diámetro de las partículas es pequeño (en comparación con el área de la sección transversal del recipiente en el cual se encuentra la suspensión) se puede identificar cada especie $i \in \{1, \dots, N\}$ con una fase continua, donde la especie i tiene densidad ρ_i y diámetro d_i . Supondremos también el siguiente orden de los tamaños de las partículas: $d_1 = 1 \geq d_2 \geq \dots \geq d_N$, con $d_i \neq d_j$ o bien $\rho_i \neq \rho_j$ para $i \neq j$.

Si denotamos por ϕ_i y v_i la fracción de volumen y velocidad de asentamiento de la especie i respectivamente, entonces, usando la ecuación de continuidad para cada una de las N especies, se obtiene que, el proceso de sedimentación polidispersa se puede describir mediante un sistema de primer orden, no-lineal y acoplado de N leyes de conservación, esto es,

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = 0, \quad (1.1.1)$$

donde $\Phi := \Phi(x, t) := (\phi_1(x, t), \dots, \phi_N(x, t))^T$ es el vector de incógnitas (en este caso,

concentraciones locales) y $\mathbf{f}(\Phi) := (f_1(\Phi), \dots, f_N(\Phi))^T$ es el vector *función de flujo* con $f_i(\Phi) := \phi_i v_i(\Phi)$, $i = 1, \dots, N$.

Para el sistema (1.1.1) se buscan soluciones $\Phi = \Phi(x, t)$ que asumen valores en el conjunto $\Phi \in \bar{\mathcal{D}}_{\phi_{\max}}$, donde $\bar{\mathcal{D}}_{\phi_{\max}}$ es la clausura del conjunto

$$\mathcal{D}_{\phi_{\max}} := \{ \Phi \in \mathbb{R}^N : \phi_1 > 0, \dots, \phi_N > 0, \phi := \phi_1 + \dots + \phi_N < \phi_{\max} \}.$$

Aquí $0 < \phi_{\max} \leq 1$ es la máxima concentración de sólidos permitida.

Para una aplicación típica a la *sedimentación batch* de una suspensión en una columna de altura L , el sistema (1.1.1) está definido en $\Omega_T := \{(x, t) \in \mathbb{R}^2 \mid 0 \leq x \leq L, 0 \leq t \leq T\}$ para un tiempo final dado $T > 0$ junto con la condición inicial

$$\Phi(x, 0) = \Phi^0(x) = (\phi_1^0(x), \dots, \phi_N^0(x))^T, \quad \Phi^0(x) \in \bar{\mathcal{D}}_{\phi_{\max}}, \quad x \in [0, L] \quad (1.1.2)$$

y condiciones de borde con flujo cero

$$\mathbf{f}|_{x=0} = \mathbf{f}|_{x=L} = 0, \quad (1.1.3)$$

las cuales indican que no hay flujo en las partes superior e inferior de la columna.

Recordemos que el sistema de leyes de conservación (1.1.1) es *hiperbólico* si para cualquier Φ todos los valores propios de la matriz Jacobiana

$$\mathcal{J}_{\mathbf{f}}(\Phi) = (f_{ij}) = \left(\frac{\partial f_i}{\partial \phi_j} \right), \quad i, j = 1 \dots, N \quad (1.1.4)$$

de la función de flujo son reales; si además los valores propios son distintos se dice que el sistema es *estrictamente hiperbólico*.

En la referencia [5] se dedujo el siguiente criterio para predecir la *estabilidad* de suspensiones bidispersas ($N = 2$)

$$I_2(\Phi) := (f_{11}(\Phi) - f_{22}(\Phi))^2 - 4f_{12}(\Phi)f_{21}(\Phi) < 0. \quad (1.1.5)$$

Si (1.1.5) se cumple, entonces el proceso de sedimentación es *inestable* lo cual, según se menciona en [12], produce una disminución de la calidad de la separación. En el otro caso, es decir, si $I_2(\Phi) \geq 0$ entonces se predice un comportamiento *estable* de la sedimentación.

Por otra parte, en [17] se mostró que el criterio (1.1.5) está estrechamente relacionado con la pérdida de hiperbolicidad del modelo (para el caso $N = 2$); más aun, en dicha referencia se demostró que, en el caso general, es decir, para un número arbitrario de especies, la hiperbolicidad puede considerarse como un criterio de estabilidad.

Por lo tanto, el análisis de los valores propios, y en general de la información característica de la matriz (1.1.4) es fundamental tanto en el estudio teórico de los modelos de sedimentación polidispersa como también en la implementación numérica de los mismos. Infortunadamente, para el caso general, la matriz Jacobiana $\mathcal{J}_f(\Phi)$ no posee una estructura que permita inferir a priori información acerca de la naturaleza de los valores propios.

Además, el uso directo del polinomio característico para extraer información relativa a los valores propios puede resultar un procedimiento dispendioso (ver por ejemplo ref. [11] para un modelo particular). Una alternativa factible propuesta en las referencias [26, 27] es explotar el hecho de que para algunos de los modelos de sedimentación de amplio uso en la literatura, la velocidad de asentamiento para cada especie puede escribirse en función de un número pequeño (con respecto al número de especies) m de variables auxiliares, esto es,

$$v_i = v_i(p_1, \dots, p_m), \quad p_l = p_l(\Phi), \quad i = 1, \dots, N, \quad l = 1, \dots, m.$$

En tal caso, la matriz Jacobiana de la función de flujo puede expresarse como una perturbación de rango m de una matriz diagonal, a saber,

$$\mathcal{J}_f(\Phi) = \mathbf{D} + \mathbf{B}\mathbf{A}^T, \quad (1.1.6)$$

donde $\mathbf{D} = \mathbf{D}(\Phi)$ es una matriz diagonal cuyas entradas son las velocidades de asentamiento y $\mathbf{A} = \mathbf{A}(\Phi)$, $\mathbf{B} = \mathbf{B}(\Phi)$ son matrices de rango m . Esta estructura particular, permite (usando un resultado del algebra lineal demostrado en las referencias [1], [26] conocido como *ecuación secular* -Teorema 3.2.1 del capítulo 3-) reducir el análisis del carácter hiperbólico del sistema (1.1.1) al estudio de las raíces de la ecuación

$$R(\lambda) := \det [\mathbf{I} + \mathbf{A}^T(\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{B}] = 1 + \sum_{i=1}^N \frac{\gamma_i}{v_i - \lambda} = 0.$$

donde los coeficientes γ_i , $i = 1, \dots, N$, se pueden calcular a partir de determinantes de orden $s \leq m$. Por ende, la eficiencia computacional de esta estrategia está basada en el hecho de que m sea pequeño y que las raíces de $R(\lambda)$ se puedan calcular con un esfuerzo moderado, en este sentido, un resultado muy útil que se obtiene a partir de la *ecuación secular* (ver ref. [26] o Corolario 3.2.1 del capítulo 3 de este trabajo) afirma que si efectivamente la matriz Jacobiana tiene la forma (1.1.6) y todos los coeficientes γ 's tienen el mismo signo, entonces dicha matriz es diagonalizable con valores propios reales y además se cumple la siguiente *propiedad de entrelazamiento*

$$\begin{cases} v_N + \sum_{i=1}^N \gamma_i < v_N < \lambda_{N-1} < \dots < \lambda_1 < v_1, & \text{si } \gamma_1, \dots, \gamma_N < 0, \\ v_N < \lambda_N < v_{N-1} < \lambda_{N-1} < \dots < v_1 < \lambda_1 < v_1 + \sum_{i=1}^N \gamma_i, & \text{si } \gamma_1, \dots, \gamma_N > 0. \end{cases}$$

Aparte de caracterizar los valores propios, la relevancia de la *propiedad de entrelazamiento* radica en que permite localizarlos y calcularlos mediante un buscador de raíces adecuado. Este último aspecto es fundamental en la simulación numérica, puesto que usualmente los esquemas numéricos precisan al menos de una buena aproximación de las velocidades características.

Este trabajo de tesis está organizado de la siguiente manera. El capítulo 2 describe brevemente un marco general para poner en contexto ésta memoria.

En el capítulo 3 se estudia la hiperbolicidad de algunos modelos para un número arbitrario de especies y cuyas velocidades de asentamiento v_i permiten aplicar la teoría descrita brevemente arriba. Entre los modelos de este tipo se incluyen los desarrollados por Masliyah [46] y Lockett y Bassoon [45] (“MLB model”), Batchelor [4] y Batchelor y Wen [6] (“BW model”), Höfler y Schwarzer [15, 37, 38] (“HS model”). En el caso del modelo MLB (para el cual $m = 2$) dicho análisis se realiza tanto para especies de igual densidad y diferentes tamaños, como para especies de igual diámetro y distintas densidades. En la primera situación la hiperbolicidad estricta se prueba en el conjunto de soluciones admisibles sin incluir ninguna restricción adicional, mientras que en el segundo caso, se muestra que el modelo es estrictamente hiperbólico pero en un subconjunto de $\mathcal{D}_{\phi_{\max}}$.

El análisis de hiperbolicidad para los modelos BW y HS se hace sólo para especies de igual densidad y diferentes diámetros. En el caso del modelo BW se muestra la estabilidad para suspensiones diluidas (concentraciones cercanas a cero). En cambio para el modelo HS, se muestra que es estrictamente hiperbólico para situaciones más realistas que permiten asumir valores de ϕ_{\max} próximos a 1. Este capítulo incluye en su parte final simulaciones numéricas usando esquemas de bajo orden que ilustran la utilidad de la información característica obtenida como subproducto del estudio teórico de la hiperbolicidad. El contenido de este capítulo corresponde al artículo [14]:

- R. Bürger, R. Donat, P. Mulet, C.A. Vega, *Hyperbolicity analysis of polydisperse sedimentation models via a secular equation for the flux Jacobian*, SIAM J. Appl. Math. **70** (2010), 2186–2213.

El capítulo 4 se refiere al uso intensivo de la información característica completa (valores y vectores propios), en la implementación de esquemas de alta resolución tipo WENO (*Weighted essentially non-oscillatory*) para la discretización espacial, los cuales combinados con métodos de discretización temporal que preservan fuertemente la estabilidad (o *SSP* por sus siglas en inglés) resultan ser muy robustos. Se muestra a partir de experimentos numéricos

la ventajas de utilizar la información espectral en el esquema WENO en vez de su contraparte por componentes (*component-wise*). En particular, la propiedad de entrelazamiento permite aproximar más convenientemente y en forma local los *coeficientes de viscosidad numérica*. Las simulaciones, que incluyen también tablas de errores y tasas de convergencia, se desarrollan para el modelo MLB y una variación del modelo HS cuya hiperbolicidad estricta se prueba a partir de la hiperbolicidad del modelo original.

El contenido de este capítulo corresponde al siguiente trabajo el cual ha sido aceptado para publicación en la revista *Journal of Computational Physics*

- R. Bürger, R. Donat, P. Mulet, C.A. Vega, *On the implementation of WENO schemes for a class of polydisperse sedimentation models.*

En el capítulo 5 se estudia primero el problema de la hiperbolicidad del modelo de Davis-Gecol (DG). Se observará que, las hipótesis que garantizan la estabilidad de éste modelo son muy restrictivas en el sentido de que, o bien las diferencias entre el diámetro de las partículas de las diferentes especies (y con la misma densidad) debe ser pequeño, o bien la concentración máxima permitida debe ser pequeña, lo cual contrasta con los resultados más generales obtenidos para el modelo HS. Se incluyen en este capítulo algunas simulaciones con dos especies, tanto con datos que satisfacen las hipótesis bajo las cuales se logra probar la hiperbolicidad, como para datos que no satisfacen dichas hipótesis. Termina dicho capítulo con algunos resultados adicionales para los modelos BW y HS. Este material forma parte de un trabajo en preparación.

Finalmente en el último breve capítulo se presentan unas conclusiones generales, perspectivas para trabajos futuros y algunas limitaciones del enfoque abordado en esta tesis.

1.2 English version.

Polydisperse suspensions consist of small solid particles dispersed in a viscous fluid, where the particles are assumed to belong to a number N of species that differ in size or density. Sedimentation processes are involved in several applications such as mineral processing, wastewater treatment, petrology, chemical engineering and medicine. Sedimentation models have been widely studied in the literature [4, 5, 6, 12, 17, 23, 35, 36, 38, 42, 45, 46, 47, 49, 52, 58, 64, 68] by considering theoretical and experimental aspects.

If the particles are small compared with the cross-sectional area of the settling vessel, then the N species can be treated as superimposed continuous phases, where species i is associated with the volume fraction ϕ_i , the phase velocity v_i , size (diameter) d_i , and density ϱ_i , where we assume that $d_1 = 1 \geq d_2 \geq \dots \geq d_N$ and $d_i \neq d_j$ or $\varrho_i \neq \varrho_j$ for $i \neq j$. The continuity equation applied to each species, leads to a strongly coupled, nonlinear systems of first-order, spatially one-dimensional conservation laws:

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = 0, \quad (1.2.1)$$

where $\Phi := \Phi(x, t) := (\phi_1(x, t), \dots, \phi_N(x, t))^T$ is the vector of unknown (in this case, *local concentrations*) and $\mathbf{f}(\Phi) := (f_1(\Phi), \dots, f_N(\Phi))^T$ is the *flux function* vector with $f_i(\Phi) := \phi_i v_i(\Phi)$, $i = 1, \dots, N$.

For the system (1.2.1), the solutions $\Phi = \Phi(x, t)$ are assumed to take values in $\Phi \in \bar{\mathcal{D}}_{\phi_{\max}}$, where $\bar{\mathcal{D}}_{\phi_{\max}}$ is the closure of the set

$$\mathcal{D}_{\phi_{\max}} := \{ \Phi \in \mathbb{R}^N : \phi_1 > 0, \dots, \phi_N > 0, \phi := \phi_1 + \dots + \phi_N < \phi_{\max} \}.$$

Here $0 < \phi_{\max} \leq 1$ is the maximum solid concentration. For a typical application to *batch sedimentation* of a polydisperse suspension in a column of height L , the system (1.2.1) is defined in $\Omega_T := \{(x, t) \in \mathbb{R}^2 \mid 0 \leq x \leq L, 0 \leq t \leq T\}$ for a given final time final $T > 0$ along with the initial condition

$$\Phi(x, 0) = \Phi^0(x) = (\phi_1^0(x), \dots, \phi_N^0(x))^T, \quad \Phi^0(x) \in \bar{\mathcal{D}}_{\phi_{\max}}, \quad x \in [0, L]$$

and zero-flux boundary conditions corresponding to a closed column, that is,

$$\mathbf{f}|_{x=0} = \mathbf{f}|_{x=L} = 0.$$

We recall that the system of conservation laws (1.2.1) is *hyperbolic* if for all Φ the Jacobian matrix

$$\mathcal{J}_{\mathbf{f}}(\Phi) = (f_{ij}) = \left(\frac{\partial f_i}{\partial \phi_j} \right), \quad i, j = 1, \dots, N \quad (1.2.2)$$

has only real eigenvalues, and *strictly hyperbolic* if these are in addition pairwise distinct.

In the reference [5] was derived the following criterion for predicting the *stability* of bidisperse suspensions ($N = 2$)

$$I_2(\Phi) := (f_{11}(\Phi) - f_{22}(\Phi))^2 - 4f_{12}(\Phi)f_{21}(\Phi) < 0. \quad (1.2.3)$$

If (1.2.3) holds, then the sedimentation process is *unstable* which could affect the quality of the separation [12]. Otherwise, *i.e.*, if $I_2(\Phi) \geq 0$ then a *stable* behavior is predicted.

On the other hand, in [17] was shown (for $N = 2$) that the criterion (1.2.3) is closely related with the loss of hyperbolicity; moreover they claimed that in general (N arbitrary), the hyperbolicity can be considered as a stability criterion. Therefore, the analysis not only of the eigenvalues but the full spectral information of the Jacobian (1.2.2) is a relevant issue for the theoretical and numerical analysis of polydisperse sedimentation models.

Unfortunately, in the general case, the Jacobian $\mathcal{J}_f(\Phi)$ has no a structure that allow us to conclude a priori something about the nature of the eigenvalues.

In addition, the direct using of the eigenpolynomial in order to obtain information about the eigenvalues can be a bulky procedure (e.g. see [11] for a particular model). A feasible alternative proposed in references [26, 27] is to exploit the fact that for some sedimentation models widely used in the literature, the settling velocity of each species can be expressed as function of a small number (compare with the number of species) m of scalar auxiliary variables, that is,

$$v_i = v_i(p_1, \dots, p_m), \quad p_l = p_l(\Phi), \quad i = 1, \dots, N, \quad l = 1, \dots, m.$$

In this case, the Jacobian matrix can be written as a rank- m perturbation of a diagonal matrix, namely,

$$\mathcal{J}_f(\Phi) = \mathbf{D} + \mathbf{B}\mathbf{A}^T, \quad (1.2.4)$$

where $\mathbf{D} = \mathbf{D}(\Phi)$ is a diagonal matrix which entries are the settling velocities and $\mathbf{A} = \mathbf{A}(\Phi)$, $\mathbf{B} = \mathbf{B}(\Phi)$ are rank- m matrix. This interesting particular structure let us (by using a result from linear algebra proved in [1] and [26] and so-called *secular equation* -Theorem 3.2.1 of chapter 3-) to reduce the hyperbolicity analysis of the system (1.2.1) to the problem of finding the roots of the equation

$$R(\lambda) := \det [\mathbf{I} + \mathbf{A}^T(\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{B}] = 1 + \sum_{i=1}^N \frac{\gamma_i}{v_i - \lambda} = 0.$$

where the coefficients γ_i , $i = 1, \dots, N$, can be computed by means of determinant of low order $s \leq m$. The computational efficiency of this strategy is based on the fact that m is small and the roots of $R(\lambda)$ can be computed with a moderate effort, for this purpose a very useful result obtained from the *secular equation* (see ref. [26] or Corollary 3.2.1 of Chapter 3 in this work) which allow us to conclude that if the Jacobian matrix has the form given by (1.2.4) and all coefficients γ 's are either positives or negative, then the Jacobian is diagonalizable with real eigenvalues. Moreover the following *interlacing property* holds

$$\begin{cases} v_N + \sum_{i=1}^N \gamma_i < v_N < \lambda_{N-1} < \dots < \lambda_1 < v_1, & \text{si } \gamma_1, \dots, \gamma_N < 0, \\ v_N < \lambda_N < v_{N-1} < \lambda_{N-1} < \dots < v_1 < \lambda_1 < v_1 + \sum_{i=1}^N \gamma_i, & \text{si } \gamma_1, \dots, \gamma_N > 0. \end{cases}$$

The *interlacing property* besides being important to characterize the eigenvalues, can also be used to localize them by employing a convenient root finder. This last matter is relevant for numerical simulations due to the numerical schemes usually require a good approximation of the characteristic speeds.

This thesis is organized with the following structure. Chapter 2 contents a brief general framework of this memoir.

In Chapter 3 an analysis of the hyperbolicity for some sedimentation models is provided for an arbitrary number of species taking into account that the form of the settling velocities permits to apply the arguments described above. We consider the models by Masliyah [46] and Lockett and Bassoon [45] (“MLB model”), Batchelor [4] and Batchelor and Wen [6] (“BW model”), Höfler and Schwarzer [15, 37, 38] (“HS model”).

In the case of the MLB model (where $m = 2$) the analysis is performed for the equal-density and different diameters case and for equal-size and different densities ones. In the former situation, strict hiperbolicity is established in the whole set of admissible solutions and the second case, the hyperbolicity is shown but in a subset of $\mathcal{D}_{\phi_{\max}}$. For the BW and HS models the hyperbolicity analysis is only done in the case of particles with the same density and different size. The stability for the BW model is restricted to dilute suspensions (concentrations close to zero). However, the hyperbolicity (strict) for the HS model is showed for values of ϕ_{\max} close to 1. This chapter also includes numerical simulations with low order schemes that illustrate the usefulness of the information obtained as a by-product of the theoretical analysis. The content of this chapter corresponds to the paper [14]:

- R. Bürger, R. Donat, P. Mulet, C.A. Vega, *Hyperbolicity analysis of polydisperse sedimentation models via a secular equation for the flux Jacobian*, SIAM J. Appl. Math. **70** (2010), 2186–2213.

Chapter 4 deals with employing the full characteristic information (eigenvalues and eigenvectors) intensively in order to implement robust numerical schemes as WENO (*Weighted essentially non-oscillatory*) for the spatial discretization which together with strong-stability preserving methods (*SSP*) are very robust. Numerical experiments will show the advantages of using a *characteristic-wise* WENO method instead of the *component-wise version*. In particular, the interlacing property is a convenient tool for approximating (locally) the viscosity coefficients. The simulations also include tables of error and convergence rates.

The material of this chapter corresponds to the following paper which has been accepted for publication in *Journal of Computational Physics*:

- R. Bürger, R. Donat, P. Mulet, C.A. Vega, *On the implementation of WENO schemes for a class of polydisperse sedimentation models.*

Chapter 5 is devoted in the first part to the stability of the Davis-Gecol model. It will be noticed that the hypothesis for concluding hyperbolicity are somewhat restrictive since the species must not differ so much in size, which contrasts with results obtained for the HS model. Some numerical experiments are included even for case where the hypothesis that assure hyperbolicity does not hold. This chapter ends with additional results for BW and HS models. These results are part of work in preparation.

Finally, the last short chapter is dedicated to some concluding remarks, limitation of this approach and future work.

Chapter 2

General framework

The theory of polydisperse sedimentation processes is part of a more general theory of multi-species kinematic flow models, which also include multi-class traffic flow models. These models are governed by nonlinear, strongly coupled systems of first-order, spatially one-dimensional conservation laws of the form

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = 0, \quad \mathbf{f}(\Phi) := (f_1(\Phi), \dots, f_N(\Phi))^T, \quad f_i(\Phi) := \phi_i v_i(\Phi), \quad i = 1, \dots, N, \quad (2.0.1)$$

where $\Phi := \Phi(x, t) := (\phi_1(x, t), \dots, \phi_N(x, t))^T$ is the vector of unknown (whose components represent volume fractions in the case of polydisperse sedimentation models and densities for traffic flow models) and v_i is the phase velocity associated with the class or species i which only depends on Φ . The hyperbolicity of multi-species kinematic flow models and their numerical approximation were studied recently, see [7, 27, 66, 70, 71, 72, 73] for traffic flow models and [11, 13, 15, 17, 19, 68] for polydisperse sedimentation models.

The multi-class traffic flow model, which is an extension of the well-known Lighthill-Whitham-Richards scalar traffic model, was developed independently by Wong and Wong [66], and Benzoni-Gavage and Colombo [7]. In the last reference, the hyperbolicity was shown by exhibiting a symmetriser for the Jacobian matrix $\mathcal{J}_{\mathbf{f}}(\Phi) := (\partial f_i / \partial \phi_k)_{1 \leq i, k \leq N}$ in the particular case that the velocities v_i , $i = 1, \dots, N$ depend on the total density ϕ in a linear fashion. Subsequently, Zhang *et al.* [70] analyzed the hyperbolicity for the general case by deriving a concise expression of the characteristic polynomial

$$P_N(\lambda) = \det(\mathcal{J}_{\mathbf{f}}(\Phi) - \lambda \mathbf{I});$$

under the assumption that the velocities v_i are pairwise distinct and $\phi_i \neq 0$ for all $i = 1, \dots, N$, they could conclude that the aforementioned polynomial has N different bounded

roots λ_i , $i = 1, \dots, N$, such that

$$v_N + \sum_{i=1}^N c_i < \lambda_1 < v_1 < \lambda_2 < \dots < v_{N-1} < \lambda_N < v_N, \quad (2.0.2)$$

with $c_i = \phi_i v_i'(\phi)$.

Since the eigenvalues λ_i cannot be expressed explicitly in closed algebraic form, the property (2.0.2) has been used, for instance, in references [19, 73] with the purpose of approximating the parameter

$$\alpha = \max_{0 \leq j \leq M-1} \max_{1 \leq i \leq N} |\lambda_i(\Phi_j)|$$

where $\Phi_j \approx \Phi(x_j, t)$ and x_j , $j = 0, \dots, M-1$ is a grid point of the computational domain. The parameter α is necessary in designing high resolution shock capturing (HRSC) schemes (e.g. WENO schemes) that involve the Lax-Friedrichs flux splitting. In references [73] and [19] (where adaptive multiresolution WENO technique is employed), the viscosity coefficient α was approximated by

$$\alpha = \max_{0 \leq j \leq M-1} \max_{1 \leq i \leq N} |v_i(\Phi_j)|.$$

However, as was pointed out by Zhang et al. [72], this choice of α can be inadequate and therefore adds numerical viscosity. They chose instead

$$\alpha = \max_{0 \leq j \leq M-1} \max \left(\left| v_N(\Phi_j) + \sum_{i=1}^N c_i(\Phi_j) \right|, |v_1(\Phi_j)| \right).$$

In either case, the WENO schemes implemented in the last references were applied in a component-wise fashion, which can lead to oscillatory behavior. To overcome this drawback, Donat and Mulet [26] provided the complete spectral decomposition for multi-class traffic flow models, with the help of this information the authors could implement characteristic-based version of HRSC schemes and compare it with the component-wise counterpart. In [26], the strategy employed to obtain the eigenstructure and analyze previously the hyperbolicity was based on viewing the Jacobian of the system as a rank-one perturbation of a diagonal matrix. The same authors used a similar procedure for a model of polydisperse suspension [27].

Most of polydisperse sedimentation models originate from chemical engineering literature [4, 5, 23, 24, 35, 38, 49, 52, 58]. In particular, we mention the following models which can be classified according to the form of the phase velocity: the Masliyah-Lockett-Bassoon (MLB model), the models based on Batchelor's approach such as the Batchelor-Wen (BW), Höfler-Schwarzer (HS) and the Davis-Gecol (DG) models, and the Patwardhan-Tien model. Related

to the hyperbolicity analysis of some of these models, we first recall here the criterion applied in [17] to the case of two species. In this case is possible to obtain information from a convenient writing of the characteristic polynomial:

$$P_2(\lambda) = \det(\mathcal{J}_f(\Phi) - \lambda \mathbf{I}) = \left(\lambda - \frac{1}{2} \left[\frac{\partial f_1(\Phi)}{\partial \phi_1} + \frac{\partial f_2(\Phi)}{\partial \phi_2} \right] \right)^2 - \frac{1}{4} I_2, \quad (2.0.3)$$

where

$$I_2 = \left(\frac{\partial f_1(\Phi)}{\partial \phi_1} - \frac{\partial f_2(\Phi)}{\partial \phi_2} \right)^2 + 4 \frac{\partial f_1(\Phi)}{\partial \phi_2} \frac{\partial f_2(\Phi)}{\partial \phi_1}. \quad (2.0.4)$$

Hence, P_2 has one pair of complex conjugate roots, that is, the system is elliptic, if and only if $I_2 < 0$, otherwise the corresponding system is hyperbolic. In [17] the hyperbolicity of the MLB model with particles of the same densities and arbitrary diameters was proved, as well for the HS model but only with $N = 2$, the hyperbolicity for $N = 3$ is conjectured from numerical evidence. For the DG model (bidisperse case) the stability is provided under the assumption that the particles do not differ too much in diameter, indeed, it was noticed that the size of instability region increases as the size ratio D_1/D_2 also increases. Expression (2.0.4) is essentially in agreement with the stability criterion for bidisperse sedimentation derived in [12] which is evaluated for the MLB and PT models. We then recall that the hyperbolicity (strict) is closely related with stability of the models. Some consequences of lack of stability include (see [12]), the formation of blobs and “fingers” in bidisperse sedimentation, increased sedimentation rates, decreased separation quality, and nonhomogeneous sediments in material manufacturing by suspension processing. On the other hand, strict hyperbolicity, and thus stability for equal-density spheres agrees with experimental evidence, since instabilities have only been observed with particles of different densities [64]. In general, hyperbolicity is a desirable property that guarantees that the solution of the conservation law (2.0.1) involves simple waves, where each eigenvalue represents a finite propagation speed of solution information.

For $N > 2$ it seems to be difficult to obtain a simple expression like (2.0.3) that allow us to conclude stability for a given model. However Berres *et al.* [11] proved the hyperbolicity of the MLB model (with equal-density particles and different diameters) for arbitrary N analyzing directly the characteristic polynomial. The same result was proved in [26] (without considering others models as BW, HS or DG), the authors did not use directly the characteristic polynomial but the strategy performed in [27] for multi-class traffic models. On the other hand, the possibility of studying the hyperbolicity of polydisperse sedimentation models aforementioned, by means of a symmetriser seems to be a complicated task (even for the simple case of $N = 2$) due to the form of the velocities functions v_i .

Numerical solutions of polydisperse sedimentation models is also important due to the nonlinear character of the system (2.0.1) does not provide exact entropy weak solutions even if $N = 2$ for an initially homogeneous suspension. Some schemes that have been applied to these sedimentation models (as well for multi-species traffic flow models) include central difference schemes (Riemman solver free) such as the Kurganov-Tadmor [41] implemented in [15, 68], where it is necessary to provide a local speed of propagation from the (approximated) eigenvalues of the Jacobian matrix of the flux function. Other simulations have been done using the WENO approach see ref. [19] and [26]. In the former reference, the procedure was performed without the whole spectral information, while in [26], simulations allow to compare the component-wise and the characteristic-wise versions but only for the MLB model. In this context, the current thesis work considers the analysis of the problem of hyperbolicity and numerical simulations for some of the polydisperse sedimentation models mentioned above and not considered in previous literature for an arbitrary number of species.

Chapter 3

Hyperbolicity analysis of polydisperse sedimentation models via a secular equation for the flux Jacobian

3.1 Introduction

3.1.1 Scope

Polydisperse suspensions consist of small solid particles dispersed in a viscous fluid, where the particles are assumed to belong to a number N of species that differ in size or density. The sedimentation of such mixtures is frequently described by spatially one-dimensional models. If the particles are small compared with the cross-sectional area of the settling vessel, then the N species can be treated as superimposed continuous phases, where species i is associated with the volume fraction ϕ_i , the phase velocity v_i , size (diameter) d_i , and density ϱ_i , where we assume that $d_1 = 1 \geq d_2 \geq \dots \geq d_N$ and $d_i \neq d_j$ or $\varrho_i \neq \varrho_j$ for $i \neq j$. The continuity equations of the N species are then $\partial_t \phi_i + \partial_x(\phi_i v_i) = 0$, $i = 1, \dots, N$, where t is time and x is depth. The velocities v_1, \dots, v_N are assumed to be given functions of the vector $\Phi := \Phi(x, t) := (\phi_1(x, t), \dots, \phi_N(x, t))^T$ of local concentrations. This yields systems of conservation laws of the type

$$\partial_t \Phi + \partial_x \mathbf{f}(\Phi) = 0, \quad \mathbf{f}(\Phi) := (f_1(\Phi), \dots, f_N(\Phi))^T, \quad f_i(\Phi) := \phi_i v_i(\Phi), \quad i = 1, \dots, N. \quad (3.1.1)$$

The one-dimensional model (3.1.1), where the concentrations are the only unknown flow variables, is called *kinematic*. We are interested in the hyperbolicity analysis of (3.1.1) for

arbitrary N under the assumption that the velocities v_1, \dots, v_N are functions of a small number m ($m \ll N$) of scalar functions of Φ , i.e.,

$$v_i = v_i(p_1, \dots, p_m), \quad p_l = p_l(\Phi), \quad i = 1, \dots, N, \quad l = 1, \dots, m. \quad (3.1.2)$$

Under the present assumptions, the entries $f_{ij}(\Phi) := \partial f_i(\Phi)/\partial \phi_j$ of the Jacobian $\mathcal{J}_f(\Phi)$ are given by

$$f_{ij} = \frac{\partial(\phi_i v_i)}{\partial \phi_j} = v_i \delta_{ij} + \sum_{l=1}^m \phi_i \frac{\partial v_i}{\partial p_l} \frac{\partial p_l}{\partial \phi_j}, \quad i, j = 1, \dots, N, \quad (3.1.3)$$

i.e., $\mathcal{J}_f(\Phi)$ is a rank- m perturbation of a diagonal matrix. Models of this type include those by Masliyah [46] and Lockett and Bassoon [45] (“MLB model”), Batchelor [4] and Batchelor and Wen [6] (“BW model”), Davis and Gecol [23] (“DG model”), and Höfler and Schwarzer [15, 37, 38] (“HS model”).

Hyperbolicity is an important property for polydisperse models, since it is often related to the range of validity of the models. However, the analysis of the characteristic polynomial of the Jacobian matrix of the system, in order to determine its eigenvalues, is rarely an easy task. Strict hyperbolicity for any N has only been proven for the MLB model, under certain restrictions (for equal-density particles, see [3, 11]).

In [26], the authors provide a proof of the hyperbolicity of the MLB model for equal-density spheres that does not involve an explicit computation of $\det(\mathcal{J}_f(\Phi) - \lambda \mathbf{I})$. It exploits the algebraic structure of the Jacobian matrix, and makes use of the fact that the eigenvalues of a rank- m perturbation of a diagonal matrix can be characterized as the roots of the so-called secular equation [1]. The analysis is based on a rational function, $R(\lambda)$, that satisfies

$$\det(\mathcal{J}_f(\Phi) - \lambda \mathbf{I}) = R(\lambda) \prod_{i=1}^N (v_i - \lambda) \quad (3.1.4)$$

for a fixed vector Φ , under appropriate circumstances. For (3.1.1), $R(\lambda)$ is of the form

$$R(\lambda) = \sum_{i=1}^N \frac{\gamma_i}{v_i - \lambda},$$

and its coefficients γ_i , $i = 1, \dots, N$ can be calculated with acceptable effort for moderate values of m . The key result is that if these coefficients are of the same sign, then the existence of N different eigenvalues of $\mathcal{J}_f(\Phi)$ is ensured. Moreover, these eigenvalues can be localized since they interlace with v_1, \dots, v_N . This is also an important property from the numerical

point of view, since no explicit formulas for the eigenvalues are available, and its computation must be always carried out by root finders.

This approach has proven to be more convenient than the explicit computation of $\det(\mathcal{J}_{\mathbf{f}}(\Phi) - \lambda \mathbf{I})$ by successive row and column eliminations done for polydisperse models in e.g. in [11, 53] (see [26] for the “secular” approach), or for kinematic traffic flow models in [70] (see [27] for the “secular” approach).

It is the purpose of this chapter to employ this calculus to provide a new proof of hyperbolicity for variants of the MLB model, and to derive new hyperbolicity results for the BW and HS models. In particular, we identify conditions on the smallest particle size, the maximum solids concentration and certain model parameters under which these models are strictly hyperbolic for arbitrary N . Numerical simulations illustrate the MLB and HS models, and demonstrate how the hyperbolicity analysis provides characteristic information required by numerical schemes.

3.1.2 Related work

For particles that have the same density, and after suitably rescaling the time variable, the components $f_i(\Phi)$, $i = 1, \dots, N$ of the flux vector $\mathbf{f}(\Phi)$ of the MLB model can be stated as

$$f_i(\Phi) = \phi_i(1 - \phi)V(\phi)(d_i^2 - (\phi_1 d_1^2 + \dots + \phi_N d_N^2)), \quad i = 1, \dots, N, \quad (3.1.5)$$

where $\phi = \phi_1 + \dots + \phi_N$ and the so-called hindered settling factor $V(\phi)$ is assumed to satisfy

$$V(0) = 1, \quad V(\phi_{\max}) = 0, \quad V'(\phi) \leq 0 \quad \text{for } \phi \in [0, \phi_{\max}], \quad (3.1.6)$$

where ϕ_{\max} is the maximum total solids concentration. We consider vectors $\Phi \in \bar{\mathcal{D}}_{\phi_{\max}}$, where $\bar{\mathcal{D}}_{\phi_{\max}}$ is the closure of the set

$$\mathcal{D}_{\phi_{\max}} := \{\Phi \in \mathbb{R}^N : \phi_1 > 0, \dots, \phi_N > 0, \phi := \phi_1 + \dots + \phi_N < \phi_{\max}\}.$$

In [17] it was shown that loss of hyperbolicity, that is, the occurrence of pairs of complex-conjugate eigenvalues of $\mathcal{J}_{\mathbf{f}}(\Phi)$, is an instability criterion for polydisperse suspensions. For $N = 2$ this criterion requires evaluating the discriminant $I_2(\Phi) := (f_{11}(\Phi) - f_{22}(\Phi))^2 - 4f_{12}(\Phi)f_{21}(\Phi)$; for vectors Φ with $I_2(\Phi) < 0$, the system (3.1.1) is unstable (elliptic) [5]. In [2, 12, 17], instability regions for $N = 2, 3$ and different choices of $\mathbf{f}(\Phi)$ are determined, while in [11] it is proven that for equal-density particles ($\varrho_1 = \dots = \varrho_N$), arbitrary N and $d_i \neq d_j$

for $i \neq j$, (3.1.1) with the MLB flux vector (3.1.5) is strictly hyperbolic for all $\Phi \in \mathcal{D}_1$. The consequences of instability include the formation of blobs and “fingers” in bidisperse sedimentation and the formation of nonhomogeneous sediments [12]. These phenomena have been observed in experiments (e.g., in [64]) under the circumstances predicted by the instability criterion. For one-dimensional kinematic models, loss of hyperbolicity sometimes predicts anomalous numerical solutions, for example, heavy and buoyant particles block each other within the vessel [10], and the steady-state sediment composition varies continuously [8].

On the other hand, strict hyperbolicity, and thus stability for equal-density spheres agrees with experimental evidence, since instabilities have only been observed with particles of different densities [64]. Consequently, a sound model should be strictly hyperbolic for equal-density particles, at least if the parameter d_N is sufficiently close to one. Thus, there is interest in determining a region of guaranteed hyperbolicity of a given model in dependence of d_N and ϕ_{\max} . This region should be independent of N , since only d_N can be controlled in real applications, for example by sieving. This work outlines a calculus that provides such a criterion for a number of models. While the results for the MLB model have been obtained by other methods (but at considerably more effort, see [11]), the analysis of the BW and HS models is new. In contrast to the MLB model, within the BW model the hindrance of all species to a given species i is not described by a factor $V = V(\phi)$, but by a factor that depends on $1 + \mathbf{s}_i^T \Phi$, where $\mathbf{s}_i^T = (S_{i1}, \dots, S_{iN})$ is a vector of non-positive coefficients and $S_{ij} < 0$ is a function of the size ratio d_j/d_i . The BW model is valid for dilute suspensions only (i.e., for suspensions of small concentration), and the HS and DG models were both proposed as extensions of the BW model to the whole range of concentrations from the dilute limit to packed sediments. The BW, HS and DG models are algebraically more complicated than the MLB model, and the results of [11], based on deriving the characteristic polynomial of $\mathcal{J}_f(\Phi)$, are difficult to apply in this case. However, if one employs a cubic polynomial dependence of S_{ij} on d_j/d_i , then the BW, HS and DG models become cases of (3.1.1) and (3.1.2) for $m = 4$, and a hyperbolicity analysis becomes feasible via the secular equation. Nevertheless, to make this chapter concise and to focus on the main ideas, we herein set the coefficient of the cubic term to zero. This assumption is also made a priori in part of the literature [15, 37, 38], and is otherwise justified by the observation that this coefficient is usually very small. Consequently, in this chapter we limit the present analysis to the case $m = 3$. In addition, we do not analyze herein the DG model, since previous work for $N = 2$ and $N = 3$ [17] showed that this model is hyperbolic for fairly narrow size distributions only. In a later chapter, however, we will employ the secular equation to explicitly derive bounds

of the hyperbolicity region for the DG model.

Although hyperbolicity in the equal-density case is a criterion that helps to decide whether a given polydisperse sedimentation model is formally sound and generates characteristic information important for the implementation of numerical schemes, we do not attempt to judge which model is preferable or more accurate. Rather, we highlight the relevance of our analysis by mentioning that recent works that employ either the MLB model or the BW, DG or HS model include [34, 47, 51, 63, 68, 69] and [21, 24, 36], respectively.

Clearly, this analysis should be extended to additional sedimentation models (see Section 3.6). Other multi-species kinematic flow models of the type (3.1.1), (3.1.2), which are amenable to a similar hyperbolicity analysis, include multi-class vehicular traffic [7, 18, 19, 27, 66, 73, 70] and the creaming of emulsions [18, 53].

3.1.3 Outline of this chapter

In Section 3.2 we outline the secular equation and its application to (3.1.1), (3.1.2), stating the basic hyperbolicity theorem, the “interlacing property” (i.e., the separation of eigenvalues by the velocities), and the computation of eigenvectors. In Section 3.3 the MLB, BW and HS models of polydisperse sedimentation are stated. In Section 3.4, which is at the core of this chapter, the secular equation is applied to analyze the hyperbolicity of each of these models. For the MLB model (where $m = 2$), we first present in Section 3.4.1 a more compact proof of hyperbolicity for equal-density spheres than in [3, 11], and then estimate the hyperbolicity region for particles that differ in density only. Then, in Section 3.4.3, we present a new analysis of the BW model for equal-density spheres by means of the secular equation, which results in a characterization of the parameter range within which the equations are hyperbolic, and in an upper bound of the total concentration up to which this property can be guaranteed. This bound is fairly small, in accordance with the limitation of the BW model to dilute suspensions. Then, in Section 3.4.4, we analyze the HS model in a similar manner. It turns out that the HS model is strictly hyperbolic if d_N is not too small. Section 3.5 presents some simulations of the MLB and HS models made by the Roe [43] and Kurganov-Tadmor (KT) [41] schemes to illustrate the sedimentation processes and the use of characteristic information provided by the calculus of the secular equation. Section 3.6 collects some conclusions of our analysis.

3.2 The secular equation

For a general flux vector $\mathbf{f}(\Phi) = (f_1(\Phi), \dots, f_N(\Phi))^T$ with $f_i(\Phi) = \phi_i v_i(\Phi)$ for $i = 1, \dots, N$, and where v_i satisfies the functional dependence given by (3.1.2), and a fixed vector Φ the Jacobian $\mathcal{J}_f = \mathcal{J}_f(\Phi)$ is a matrix of the form

$$\begin{aligned} \mathcal{J}_f &= \mathbf{D} + \mathbf{B}\mathbf{A}^T, \quad \mathbf{D} := \text{diag}(v_1, \dots, v_N), \\ \begin{cases} \mathbf{B} := (B_{il}) = (\phi_i \partial v_i / \partial p_l), \\ \mathbf{A} := (A_{jl}) = (\partial p_l / \partial \phi_j), \end{cases} & \quad 1 \leq i, j \leq N, \quad 1 \leq l \leq m. \end{aligned} \quad (3.2.1)$$

(This is a matrix representation of (3.1.3).)

As a motivation for the secular equation, we briefly summarize the analysis in [26]. To this end, let us assume that λ is an eigenvalue of $\mathcal{J}_f(\Phi)$, with eigenvector $\mathbf{x} \neq \mathbf{0}$, i.e., $(\mathbf{D} + \mathbf{B}\mathbf{A}^T)\mathbf{x} = \lambda\mathbf{x}$ and such that

$$\lambda \neq v_i \quad \text{for all } i = 1, \dots, N. \quad (3.2.2)$$

Since $(\mathbf{D} - \lambda\mathbf{I})$ is invertible, we can also write

$$\mathbf{x} + (\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{B}(\mathbf{A}^T\mathbf{x}) = \mathbf{0}, \quad (3.2.3)$$

and multiplying this relation by \mathbf{A}^T , we get

$$\mathbf{A}^T\mathbf{x} + \mathbf{A}^T(\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{B}(\mathbf{A}^T\mathbf{x}) = \mathbf{0}, \quad (3.2.4)$$

i.e., the vector $\boldsymbol{\xi} := \mathbf{A}^T\mathbf{x} \in \mathbb{R}^m$ satisfies $\mathbf{M}_\lambda\boldsymbol{\xi} = \mathbf{0}$, where

$$\mathbf{M}_\lambda := \mathbf{I} + \mathbf{A}^T(\mathbf{D} - \lambda\mathbf{I})^{-1}\mathbf{B}. \quad (3.2.5)$$

Clearly, we must have $\boldsymbol{\xi} \neq \mathbf{0}$, since otherwise the second term in the left-hand side of (3.2.3) would be zero, and we would have $\mathbf{x} = \mathbf{0}$, in contradiction to our assumption that \mathbf{x} is a non-zero eigenvector. Hence, any eigenvalue $\lambda \neq v_i$ for all i must be a root of the equation $\det \mathbf{M}_\lambda = 0$, and we obtain a direct relation between the equation $\det \mathbf{M}_\lambda = 0$ and the eigenvalues of \mathcal{J}_f . The optimal situation is when $R(\lambda) := \det \mathbf{M}_\lambda$ has N different real roots, since in this case these must be all the eigenvalues of \mathcal{J}_f , which ensures strict hyperbolicity of the system (3.1.1).

The secular equation

$$R(\lambda) = \det \mathbf{M}_\lambda = 0 \quad (3.2.6)$$

provides, hence, relevant information on the eigenvalues of \mathcal{J}_f . Rather than forming explicitly the matrix \mathbf{M}_λ and compute its determinant, we use a form of the function $R(\lambda)$ that can be traced back to Anderson [1], obtained after using certain algebraic results concerning eigenvalues of rank- m perturbations of a diagonal matrix.

We introduce now the relevant notation. We denote by S_r^p the set of all (ordered) subsets of r elements taken from a set of p elements. Assuming that \mathbf{X} is an $m \times N$ matrix, and given the index sets $I := \{i_1 < \dots < i_k\} \in S_k^N$ and $J := \{j_1 < \dots < j_l\} \in S_l^m$, we denote by $\mathbf{X}^{I,J}$ the $k \times l$ submatrix of \mathbf{X} given by $(\mathbf{X}^{I,J})_{p,q} = X_{i_p,j_q}$. The following theorem can be found in [1], but we give here the form in [26], which provides the explicit formulas to be used in the applications.

Theorem 3.2.1 (The secular equation, [1, 26]) *Assume that \mathbf{D} is a diagonal matrix as given by (3.2.1) with $v_i > v_j$ for $i < j$, and that \mathbf{A} and \mathbf{B} have the formats specified in (3.2.1). Let $\lambda \neq v_i$ for $i = 1, \dots, N$. Then λ is an eigenvalue of $\mathbf{D} + \mathbf{B}\mathbf{A}^T$ if and only if*

$$R(\lambda) := \det \mathbf{M}_\lambda = 1 + \sum_{i=1}^N \frac{\gamma_i}{v_i - \lambda} = 0. \quad (3.2.7)$$

The coefficients γ_i , $i = 1, \dots, N$, are given by the following expression:

$$\gamma_i = \sum_{r=1}^{\min\{N,m\}} \sum_{i \in I \in S_r^N, J \in S_r^m} \frac{\det \mathbf{A}^{I,J} \det \mathbf{B}^{I,J}}{\prod_{l \in I, l \neq i} (v_l - v_i)}. \quad (3.2.8)$$

The importance of the secular equation is elucidated by the following corollary, which is an extended version of [26, Corollary 1].

Corollary 3.2.1 *With the notation of Theorem 3.2.1, assume that $\gamma_i \cdot \gamma_j > 0$ for $i, j = 1, \dots, N$. Then $\mathbf{D} + \mathbf{B}\mathbf{A}^T$ is diagonalizable with real eigenvalues $\lambda_1, \dots, \lambda_N$. If $\gamma_1, \dots, \gamma_N < 0$, the interlacing property*

$$M_1 := v_N + \gamma_1 + \dots + \gamma_N < \lambda_N < v_N < \lambda_{N-1} < \dots < \lambda_1 < v_1 \quad (3.2.9)$$

holds, while for $\gamma_1, \dots, \gamma_N > 0$, the following analogous property holds:

$$v_N < \lambda_N < v_{N-1} < \lambda_{N-1} < \dots < v_1 < \lambda_1 < M_2 := v_1 + \gamma_1 + \dots + \gamma_N. \quad (3.2.10)$$

Proof. If $v_i \neq v_j$ for $i \neq j$ and $\gamma_1, \dots, \gamma_N$ have the same sign s , then $R(\lambda) \rightarrow (\mp s)\infty$ as $\lambda \rightarrow v_i^\pm$, hence R changes sign between two consecutive poles. Since $R(\lambda) \rightarrow 1$ for $\lambda \rightarrow \pm\infty$,

there is another change of sign to the left of v_N if $s = -1$ or to the right of v_1 if $s = 1$. Thus, noting that $M_1 < v_N$ and $M_2 > v_1$ in the respective cases $\gamma_i < 0$ and $\gamma_i > 0$, the property of M_1 being a lower bound and M_2 being an upper bound for $\lambda_1, \dots, \lambda_N$, respectively, is established if we can show that $R(M_1) \geq 0$ and $R(M_2) \geq 0$, respectively. In the case $\gamma_i < 0$ we have that

$$M_1 - v_i = v_N - v_i + \gamma_1 + \dots + \gamma_N \leq \gamma_1 + \dots + \gamma_N, \quad \text{for } i = 1, \dots, N,$$

which implies that $\gamma_i/(M_1 - v_i) \leq \gamma_i/(\gamma_1 + \dots + \gamma_N)$ for $i = 1, \dots, N$, and therefore

$$R(M_1) = 1 + \sum_{i=1}^N \frac{\gamma_i}{v_i - M_1} \geq 1 - \sum_{i=1}^N \frac{\gamma_i}{\gamma_1 + \dots + \gamma_N} = 0.$$

The proof of $R(M_2) \geq 0$ is analogous. \square

As a consequence of Corollary 3.2.1, strict hyperbolicity of (3.1.1) follows whenever the coefficients γ_i of the associated secular equation of the system have a uniform sign. The interlacing property is important for numerical schemes, since the actual eigenvalues may be computed conveniently by a root finder. The bounds for the eigenvalues, i.e. the characteristic speeds of the system, are also important for numerical purposes.

Remark 3.2.1 *The conditions of Corollary 3.2.1 are sufficient for the strict hyperbolicity of the models, but are far from necessary. A slightly weaker set of hypotheses leading to strict hyperbolicity would be that the coefficients γ_i either have a definite sign or, if $\gamma_i = 0$, then v_i is not a root of the secular equation, for Theorem 3.2.1 would then anyway provide N different roots of the characteristic polynomial. The condition $R(v_i) \neq 0$ whenever $\gamma_i = 0$ is hard to analyze in general, but there is a situation where it can be neatly ensured, given, for instance, that $v_1 > \dots > v_N$ and $\gamma_j \leq 0$ for all j : $\gamma_i = 0$ implies $\gamma_j = 0$ for all $j < i$, which yields*

$$R(v_i) = 1 + \sum_{j=i+1}^N \frac{\gamma_j}{v_j - v_i} > 0. \quad (3.2.11)$$

Remark 3.2.2 *In the context of models of polydisperse sedimentation, the situation addressed in Remark 3.2.1 occurs if we choose $\Phi = (\phi_1, \dots, \phi_N)^T$ such that $\phi_i = 0$ for $1 \leq i \leq i_0$ and $\phi_i \in (0, \phi_{\max})$ for $i_0 + 1 \leq i \leq N$ for an index $i_0 \in \{1, \dots, N\}$. Consequently, suppose that we are able to establish strict hyperbolicity on all of $\mathcal{D}_{\phi_{\max}}$, then this property also holds on certain parts of the boundary of $\mathcal{D}_{\phi_{\max}}$; on the remaining parts of that boundary, the model (3.1.1) is still hyperbolic, but not necessarily strictly hyperbolic.*

Remark 3.2.3 *Anderson [1] mentions that for the case of a self-adjoint rank-1 perturbation of a diagonal matrix, Theorem 3.2.1 was proved first by Golub [33], who also utilizes the expression “secular equation” [33, p. 327] for the algebraic form of $R(\lambda)$ in that case. Nevertheless, this expression is in fact much older, and appeared at least as early as in the work of Cauchy [20], where the adjective secular has the meaning of “existing or continuing through the centuries”.*

Finally, another important by-product of this derivation is the possibility of having an explicit expression of the spectral decomposition of the Jacobian matrix of the system, which is also a useful asset in numerical simulations. Assume that λ is a root of the secular equation, i.e., λ is an eigenvalue of \mathcal{J}_f that satisfies (3.2.2). Then $\boldsymbol{\xi} = \mathbf{A}^T \mathbf{x}$ is a solution of $\mathbf{M}_\lambda \boldsymbol{\xi} = \mathbf{0}$. But \mathbf{M}_λ is an $m \times m$ matrix that can easily be computed. Given two vectors $\mathbf{g} = (g_1, \dots, g_N)^T$, $\mathbf{h} = (h_1, \dots, h_N)^T \in \mathbb{R}^N$, if we use the notation

$$[\mathbf{g}, \mathbf{h}] := [\mathbf{g}, \mathbf{h}]_\lambda := \mathbf{g}^T (\mathbf{D} - \lambda \mathbf{I})^{-1} \mathbf{h} = \sum_{k=1}^N \frac{g_k h_k}{v_k - \lambda} \quad (3.2.12)$$

and denote by $\mathbf{a}_1, \dots, \mathbf{a}_m$ and $\mathbf{b}_1, \dots, \mathbf{b}_m$ the columns of \mathbf{A} and \mathbf{B} , respectively, then

$$\mathbf{M}_\lambda = \mathbf{I} + ([\mathbf{a}_i, \mathbf{b}_j])_{1 \leq i, j \leq m}.$$

Assuming that we can compute a non-zero solution $\boldsymbol{\xi}$ of

$$\mathbf{M}_\lambda \boldsymbol{\xi} = \mathbf{0}, \quad (3.2.13)$$

we can use the relation $\mathbf{x} + (\mathbf{D} - \lambda \mathbf{I})^{-1} \mathbf{B}(\mathbf{A}^T \mathbf{x}) = \mathbf{0}$ to compute a right eigenvector of \mathcal{J}_f as

$$\mathbf{x} = -(\mathbf{D} - \lambda \mathbf{I})^{-1} \mathbf{B} \boldsymbol{\xi}. \quad (3.2.14)$$

The same procedure may be employed to calculate the left eigenvectors of \mathcal{J}_f , since they are the right eigenvectors of $\mathcal{J}_f^T = \mathbf{D} + \mathbf{A} \mathbf{B}^T$. In other words, the roles of \mathbf{A} and \mathbf{B} and corresponding columns need to be interchanged. This will be further illustrated for the MLB model.

For the case given by (3.2.1) and assuming that $m < N$, we note that with \mathbf{A} and \mathbf{B} defined in (3.2.1) we can write

$$\det \mathbf{A}^{I,J} = \det \left(\frac{\partial p_J}{\partial \phi_I} \right), \quad \det \mathbf{B}^{I,J} = \det \left(\frac{\partial v_I}{\partial p_J} \right) \prod_{l \in I} \phi_l,$$

where the notation above should be self-explanatory. Then, we can write

$$\gamma_i = \phi_i \sum_{r=1}^m \gamma_{r,i}, \quad \gamma_{r,i} = \sum_{i \in I \in S_r^N} \prod_{l \in I, l \neq i} \frac{\phi_l}{v_l - v_i} \sum_{J \in S_r^m} \det \left(\frac{\partial v_I}{\partial p_J} \right) \det \left(\frac{\partial p_J}{\partial \phi_I} \right). \quad (3.2.15)$$

When $m = 1$ or $m = 2$, these quantities can be easily computed and the hyperbolicity analysis via the secular equation is much less involved than the study of $\det(\mathcal{J}_f(\Phi) - \lambda \mathbf{I})$. For $m = 3$ or $m = 4$, the computations in the secular equation are more involved, but they can be useful in providing at least partial results concerning hyperbolicity, where the theoretical analysis of $\det(\mathcal{J}_f(\Phi)) - \lambda \mathbf{I}$ is essentially out of reach.

3.3 Kinematic models of polydisperse sedimentation

3.3.1 The Masliyah-Lockett-Bassoon (MLB) model

The MLB model is based on the following explicit equation for the solid-fluid relative velocity $u_i := v_i - v_f$ of species i , also called slip velocity:

$$u_i = \frac{gd_i^2(\bar{\varrho}_i - \bar{\boldsymbol{\varrho}}^T \Phi)}{18\mu_f} V_i(\Phi), \quad (3.3.1)$$

where g is the acceleration of gravity, $\bar{\varrho}_i := \varrho_i - \varrho_f$ is the reduced density of particle species i , where ϱ_i and ϱ_f are the mass densities of particle species i and the fluid, respectively, μ_f is the viscosity of the fluid, $\bar{\boldsymbol{\varrho}} := (\bar{\varrho}_1, \dots, \bar{\varrho}_N)^T$, and $V_i(\Phi)$ is the hindered settling factor for species i . In most previous works, this factor is assumed to be the same for all species, and is assumed to depend on $\phi := \phi_1 + \dots + \phi_N$ only, i.e. $V_i(\Phi) = V(\phi)$ for $i = 1, \dots, N$, and may be given by the Richardson-Zaki [52] expression

$$V(\phi) = (1 - \phi)^{n-2} \quad \text{for } \phi \in \bar{\mathcal{D}}_{\phi_{\max}}, \quad V(\phi) = 0 \quad \text{otherwise}, \quad n > 2. \quad (3.3.2)$$

Expressing the velocities v_1, \dots, v_N and v_f in terms of the volume average velocity $q := (1 - \phi)v_f + \phi_1 v_1 + \dots + \phi_N v_N$ and the slip velocities u_1, \dots, u_N we obtain the flux functions

$$f_i(\Phi) = \phi_i v_i = q\phi_i + \phi_i(u_i - (\phi_1 u_1 + \dots + \phi_N u_N)), \quad i = 1, \dots, N. \quad (3.3.3)$$

while summing the continuity equations for the solids species and that of the fluid yields $\partial_x q = 0$, i.e., $q = 0$ for batch settling in a closed vessel. Then, inserting (3.3.1) into (3.3.3), assuming that (for example, after rescaling x) the constant $g/(18\mu_f)$ equals one and recalling the ordering

$$d_1 = 1 \geq d_2 \geq \dots \geq d_{N-1} \geq d_N, \quad (3.3.4)$$

we obtain

$$f_i(\Phi) = \phi_i v_i(\Phi), \quad v_i(\Phi) = \left[(\bar{\varrho}_i - \bar{\boldsymbol{\varrho}}^T \Phi) d_i^2 V_i(\Phi) - \sum_{m=1}^N \phi_m d_m^2 (\bar{\varrho}_m - \bar{\boldsymbol{\varrho}}^T \Phi) V_m(\Phi) \right], \quad (3.3.5)$$

$$i = 1, \dots, N.$$

Under the assumption $V_i(\Phi) = V(\phi)$ for $i = 1, \dots, N$, the equation for v_i in (3.3.5) assumes the form

$$v_i(\Phi) = V(\phi) \left[(\bar{\varrho}_i - \bar{\boldsymbol{\varrho}}^T \Phi) d_i^2 - \sum_{m=1}^N \phi_m d_m^2 (\bar{\varrho}_m - \bar{\boldsymbol{\varrho}}^T \Phi) \right], \quad i = 1, \dots, N. \quad (3.3.6)$$

On the other hand, for equal-density particles we have $\bar{\varrho}_i - \bar{\boldsymbol{\varrho}}^T \Phi = (1 - \phi)(\varrho_s - \varrho_f)$, where ϱ_s is the uniform solids density. Hence, we assume that t has been rescaled so that (3.3.5) becomes

$$v_i(\Phi) = (1 - \phi) \left[d_i^2 V_i(\Phi) - \sum_{m=1}^N \phi_m d_m^2 V_m(\Phi) \right], \quad i = 1, \dots, N. \quad (3.3.7)$$

Clearly, combining the assumption $V_i(\Phi) = V(\phi)$ with that of equal particle densities leads to the flux function (3.1.5).

Finally, we may also consider a suspension of particles of equal size d of species that differ in density only. Assuming $V_i(\Phi) = V(\phi)$ and that the factor $gd^2/(18\mu_f)$ is set to one, we then obtain

$$v_i(\Phi) = V(\phi) [\bar{\varrho}_i + (\phi - 2)\bar{\boldsymbol{\varrho}}^T \Phi], \quad i = 1, \dots, N. \quad (3.3.8)$$

The secular equation can be employed for the hyperbolicity analysis of the MLB model in various particular instances, as we shall see in Section 3.4.

3.3.2 Models based on Batchelor's approach

The MLB model is derived from the mass and linear momentum balance equations of the particle species and the fluid (see [11]). A different approach is due to Batchelor [4], which is based on the following expression for the settling velocity v_i of spheres of species i , having diameter d_i , in a dilute suspension:

$$v_i(\Phi) = v_i(\mathbf{0})(1 + \mathbf{s}_i^T \Phi), \quad i = 1, \dots, N. \quad (3.3.9)$$

Here, $v_i(\mathbf{0})$ is the settling velocity of a single sphere of species i in pure fluid, that is, $v_i(\mathbf{0})$ is the Stokes velocity $v_i(\mathbf{0}) = d_i^2 \bar{\varrho}_i / (18\mu_f)$, and $\mathbf{s}_i^T := (S_{i1}, \dots, S_{iN})$ is the i -th row of the matrix

$\mathbf{S} = (S_{ij})_{1 \leq i, j \leq N}$. The dimensionless sedimentation coefficients S_{ij} are in general negative functions of $\lambda_{ij} := d_j/d_i$ and $\varrho_{ij} := \bar{\varrho}_j/\bar{\varrho}_i$, of the Péclet number

$$\mathcal{P}_{ij} := \frac{(d_i + d_j)|v_j(\mathbf{0}) - v_i(\mathbf{0})|}{(4\mathcal{D}_{ij})},$$

and of interparticle attractive-repulsive forces. Here, $\mathcal{D}_{ij} := (kT)(3\pi\mu_f)^{-1}(d_i^{-1} + d_j^{-1})$ is the so-called relative diffusivity, where T is temperature and k is the Boltzmann constant [4, 6]. The coefficients S_{ij} can be calculated from the pair distribution function, which represents the statistical structure of the suspension [4]. This was done numerically by Batchelor and Wen [6] for a range of values of $\lambda = \lambda_{ij}$ and $\varrho = \varrho_{ij}$, considering the limits of either a large ($\mathcal{P}_{ij} \gg 1$) or a small ($\mathcal{P}_{ij} \ll 1$) Péclet number, and neglecting Brownian diffusion.

The secular equation can be employed for the hyperbolicity analysis of several models based on Batchelor's approach with equal-density particles ($\varrho_{ij} = 1$ for $1 \leq i, j \leq N$). In this case, after rescaling time, we may express (3.3.9) as

$$v_i(\Phi) = d_i^2(1 + \mathbf{s}_i^T \Phi), \quad i = 1, \dots, N, \quad (3.3.10)$$

and the coefficients S_{ij} can be reasonably approximated by a formula of the type

$$S_{ij} = \sum_{l=0}^3 \beta_l \left(\frac{d_j}{d_i}\right)^l, \quad 1 \leq i, j \leq N. \quad (3.3.11)$$

We will refer to (3.3.10), (3.3.11) as the Batchelor and Wen (BW) model.

Davis and Gecol [23] were the first to approximate the numerical values of S_{ij} , tabulated in [6] for $\varrho_{ij} = 1$ for eight different values of λ_{ij} , by an expression of the type (3.3.11); they obtained the coefficients

$$\begin{aligned} \boldsymbol{\beta}^T &:= (\beta_0, \dots, \beta_3) \\ &= \begin{cases} (-3.5, -1.1, -1.02, -0.002) & \text{for large Péclet numbers } (\mathcal{P}_{ij} \gg 1), \\ (-3.42, -1.96, -1.21, -0.013) & \text{for small Péclet numbers } (\mathcal{P}_{ij} \ll 1). \end{cases} \end{aligned} \quad (3.3.12)$$

We observe that in both cases, $\beta_i < 0$ for $i = 0, \dots, 3$, and that $|\beta_3|$ is very small. In fact, some authors utilize $\beta_3 = 0$ a priori; for example, Höfler and Schwarzer [38] fit the data from [6] for large Péclet numbers to a second-order polynomial corresponding to

$$\boldsymbol{\beta}^T = (\beta_0, \dots, \beta_3) = (-3.52, -1.04, -1.03, 0). \quad (3.3.13)$$

That $|\beta_3|$ should be small while $\beta_0, \beta_1, \beta_2 \leq 0$ is also supported by theoretical asymptotical result [4] stating that $S_{ij} + \varrho_{ij}(\lambda_{ij}^2 + 3\lambda_{ij} + 1) \rightarrow 0$ as $\lambda_{ij} \rightarrow \infty$, which is relevant here only

for $\varrho_{ij} = 1$. For a detailed discussion of the coefficients S_{ij} and further data we refer to [4, 6, 54, 65]. Our further analysis will indeed rely on the negativity of the β_i . As we shall see shortly, setting $\beta_3 = 0$ simplifies greatly the computations involved in the hyperbolicity analysis via the secular equation.

To overcome the limitation of (3.3.10), and the BW approach, to dilute suspensions, other models have been proposed. Davis and Gecol [23] replace (3.3.10) by

$$v_i(\Phi) = d_i^2(1 + \mathbf{s}_i^T \Phi - S_{ii}\phi)(1 - \phi)^{-S_{ii}}, \quad (3.3.14)$$

and claimed that (3.3.14) could be used for size ratios $d_N \geq 1/8$. However, in [17] it is shown that for $N = 2$ and $d_2 \approx 1/6$, the system (3.1.1) based on using (3.3.14) exhibits unphysical instability regions for equal-density spheres. Another velocity equation that formally extends (3.3.10) to the whole range of concentrations was suggested by Höfler and Schwarzer [15, 37, 38]:

$$v_i(\Phi) = d_i^2 \exp(\mathbf{s}_i^T \Phi + n\phi)(1 - \phi)^n, \quad n \geq 0. \quad (3.3.15)$$

For $\Phi \rightarrow \mathbf{0}$, (3.3.14) and (3.3.15) have the same partial derivatives as (3.3.10), while for $\phi \rightarrow 1$, the velocities v_i given by (3.3.14) and (3.3.15) vanish. Moreover, for the HS model it is straightforward to verify (see [17]) that $I_2 > 0$ for arbitrary non-positive Batchelor matrices \mathbf{S} and $N = 2$. Furthermore, based on numerical tests, it was conjectured in [17] that the model based on (3.3.15) would be stable also for $N = 3$. The present work confirms this conjecture and shows that the model is stable for arbitrary N , provided that for a given vector of coefficients β , the quantities d_N and Φ satisfy some mild conditions.

3.4 Hyperbolicity analysis via the secular equation

Since our analysis should be general with respect to the number of species N and the particle size classes d_1, \dots, d_N , we employ the smallest normalized particle size d_N as the only scalar parameter that characterizes the width of the particle size distribution. We always assume the ordering of particle sizes (3.3.4). This means that for equal-density particles, $i > j$ for $i, j \in \{1, \dots, N\}$ is equivalent to $d_i < d_j$.

3.4.1 The Masliyah-Lockett-Bassoon (MLB) model

The MLB model for equal-density spheres (3.1.1), (3.1.5) is known to be strictly hyperbolic for all N and all $\Phi \in \mathcal{D}_{\phi_{\max}}$. A proof of this fact was obtained in [11], by deriving an

explicit formula for the characteristic polynomial of the Jacobian matrix $\mathcal{J}_f(\Phi)$. In [26], a considerably shorter proof of this same fact was provided by using the secular equation.

To illustrate the use of the secular equation to show strict hyperbolicity, we consider the MLB model for species of identical density in a more general case, considered in [3]. Basson et al. [3] recently analyzed the variant of the MLB model for equal-density spheres based on (3.3.7) with the following hindered settling factors:

$$\begin{aligned} V_i(\phi) &= (1 - \phi)^{n_i - 2} \quad \text{for } \phi \in \bar{\mathcal{D}}_{\phi_{\max}}, \quad V_i(\phi) = 0 \quad \text{otherwise,} \\ n_i &> 2, \quad n_i \leq n_j \text{ for } i < j. \end{aligned} \quad (3.4.1)$$

In [3], an analysis of the characteristic polynomial of $\mathcal{J}_f(\Phi)$ similar to that of [11] led to the conclusion that the model is still strictly hyperbolic for all N and $\Phi \in \mathcal{D}_{\phi_{\max}}$. Here we provide a much shorter proof of this fact using the secular equation.

Lemma 3.4.1 *The MLB model for equal-density spheres (3.1.1), (3.3.7) and the hindered settling factors (3.4.1) is strictly hyperbolic for all $\Phi \in \mathcal{D}_{\phi_{\max}}$. Its eigenvalues $\lambda_i = \lambda_i(\Phi)$ satisfy the interlacing property*

$$\begin{aligned} M_1(\Phi) &< \lambda_N(\Phi) < v_N(\Phi) < \lambda_{N-1}(\Phi) < v_{N-1}(\Phi) < \cdots < \lambda_1(\Phi) < v_1(\Phi), \\ M_1(\Phi) &:= d_N^2 V_N(\Phi) + \sum_{j=1}^N d_j^2 \phi_j ((1 - \phi) V_j'(\phi) - 2V_j(\phi)). \end{aligned}$$

The right and left eigenvectors of $\mathcal{J}_f(\Phi)$, denoted by \mathbf{x} and \mathbf{y} , respectively, that correspond to a root λ of the secular equation are

$$\mathbf{x}_i = \frac{1}{v_i - \lambda} \left[b_{i,1} \sum_{k=1}^N \frac{a_{k,1} b_{k,2}}{v_k - \lambda} - b_{i,2} \left(1 + \sum_{k=1}^N \frac{a_{k,1} b_{k,1}}{v_k - \lambda} \right) \right], \quad i = 1, \dots, N, \quad (3.4.2)$$

$$\mathbf{y}_i = \frac{1}{v_i - \lambda} \left[a_{i,1} \sum_{k=1}^N \frac{b_{k,1} a_{k,2}}{v_k - \lambda} - a_{i,2} \left(1 + \sum_{k=1}^N \frac{b_{k,1} a_{k,1}}{v_k - \lambda} \right) \right], \quad i = 1, \dots, N, \quad (3.4.3)$$

where

$$b_{i,1} = \phi_i d_i^2 V_i'(\phi), \quad b_{i,2} = -\phi_i, \quad a_{i,1} = 1, \quad a_{i,2} = \sum_{j=1}^N d_j^2 V_j'(\phi) \phi_j + d_i^2 V_i(\phi). \quad (3.4.4)$$

Proof. In this case, we have $m = 2$, and the velocities can be expressed by $v_i = V_i(p_1) d_i^2 - p_2$, where $p_1 := \phi = \phi_1 + \cdots + \phi_N$ and $p_2 = d_1^2 V_1(\phi) \phi_1 + \cdots + d_N^2 V_N(\phi) \phi_N$. To compute the expressions in (3.2.15), we need

$$\begin{aligned}
v_l - v_i &= V_l(\phi)d_l^2 - V_i(\phi)d_i^2, & \frac{\partial v_i}{\partial p_1} &= d_i^2 V_i'(\phi), & \frac{\partial v_i}{\partial p_2} &= -1, \\
\frac{\partial p_1}{\partial \phi_i} &= 1, & \frac{\partial p_2}{\partial \phi_i} &= \sum_{j=1}^N d_j^2 V_j'(\phi)\phi_j + d_i^2 V_i(\phi); \\
\det\left(\frac{\partial v_{i,l}}{\partial p_{1,2}}\right) &= d_l^2 V_l'(\phi) - d_i^2 V_i'(\phi), & \det\left(\frac{\partial p_{1,2}}{\partial \phi_{i,l}}\right) &= d_l^2 V_l(\phi) - d_i^2 V_i(\phi).
\end{aligned}$$

In this way, we obtain

$$\begin{aligned}
\gamma_{1,i} &= \frac{\partial v_i}{\partial p_1} \frac{\partial p_1}{\partial \phi_j} + \frac{\partial v_i}{\partial p_2} \frac{\partial p_2}{\partial \phi_j} = d_i^2 (V_i'(\phi) - V_i(\phi)) - \sum_{j=1}^N d_j^2 V_j'(\phi)\phi_j, \\
\gamma_{2,i} &= \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\phi_j}{v_j - v_i} \det\left(\frac{\partial v_{i,j}}{\partial p_{1,2}}\right) \det\left(\frac{\partial p_{1,2}}{\partial \phi_{i,j}}\right) = \sum_{\substack{j=1 \\ j \neq i}}^N \phi_j (d_j^2 V_j'(\phi) - d_i^2 V_i'(\phi)),
\end{aligned}$$

which finally leads to $\gamma_i = (V_i'(\phi)(1 - \phi) - V_i(\phi))\phi_i d_i^2$. Due to (3.1.6), we obtain $\gamma_i < 0$ for all $i = 1, \dots, N$ and $\Phi \in \mathcal{D}_{\phi_{\max}}$. The interlacing property and the form of M_1 follow from Corollary 3.2.1. Therefore, we deduce (3.4.2) by considering $\lambda = \lambda_j$, and also taking $\boldsymbol{\xi} = (-[\mathbf{a}_1, \mathbf{b}_2], 1 + [\mathbf{a}_1, \mathbf{b}_1]) \in \mathbb{R}^2$ as solution of the first of the two (equivalent) equations in (3.2.13) and substituting into (3.2.14). Since the left eigenvectors of the matrix $\mathcal{J}_{\mathbf{f}} = \mathbf{D} + \mathbf{B}\mathbf{A}^T$ are the right eigenvectors of $\mathcal{J}_{\mathbf{f}}^T = \mathbf{D} + \mathbf{A}\mathbf{B}^T$, (3.4.3) can be deduced from (3.4.2) by interchanging the roles of \mathbf{A} and \mathbf{B} . \square

For equal-sized particles, which differ in density only, $\mathcal{D}_{\phi_{\max}}$ has in general a sub-region with lack of hyperbolicity [12, 17]. In this case, the quantities γ_i will in general not have a definite sign, but we may still employ the secular equation to estimate the size of the hyperbolicity region of $\mathcal{D}_{\phi_{\max}}$.

Lemma 3.4.2 *The MLB model for equal-sized heavy spheres (3.1.1), (3.3.8), where $\bar{\varrho}_1 > \bar{\varrho}_2 > \dots > \bar{\varrho}_N > 0$ and the hindered settling factor $V(\phi)$ is given by (3.3.2), is strictly hyperbolic for all $\Phi \in \mathcal{D}_{\phi_*} \subset \mathcal{D}_1$, where*

$$\phi_* = \frac{n\bar{\varrho}_N}{\bar{\varrho}_1 + \bar{\varrho}_N(n-1)} = \frac{n}{n + \gamma - 1}, \quad \gamma = \bar{\varrho}_1/\bar{\varrho}_N > 1. \quad (3.4.5)$$

Proof. In this case, we have $v_i = V(p_1)(\bar{\varrho}_i + (p_1 - 2)p_2)$, where $p_1 = \phi$ and $p_2 = \bar{\boldsymbol{\varrho}}^T \Phi$. This implies

$$v_j - v_i = V(\phi)(\bar{\varrho}_j - \bar{\varrho}_i), \quad \frac{\partial v_i}{\partial p_1} = V'(\phi)(\bar{\varrho}_i + (\phi - 2)p_2) + V(\phi)p_2,$$

$$\frac{\partial v_i}{\partial p_2} = V(\phi)(\phi - 2), \quad \frac{\partial p_1}{\partial \phi_i} = 1, \quad \frac{\partial p_2}{\partial \phi_i} = \bar{\varrho}_i;$$

$$\det \left(\frac{\partial v_{ij}}{\partial p_{12}} \right) = V(\phi)V'(\phi)(\phi - 2)(\bar{\varrho}_i - \bar{\varrho}_j), \quad \det \left(\frac{\partial p_{12}}{\partial \phi_{ij}} \right) = \bar{\varrho}_j - \bar{\varrho}_i.$$

In this way, we get

$$\gamma_{1,i} = V'(\phi)(\bar{\varrho}_i + (\phi - 2)p_2) + V(\phi)(p_2 + (\phi - 2)\bar{\varrho}_i), \quad \gamma_{2,i} = V'(\phi)(\phi - 2)(\phi\bar{\varrho}_i - p_2),$$

which for $V(\phi)$ given by (3.3.2) yields

$$\gamma_i = -\phi_i(1 - \phi)^{n-2}[\bar{\varrho}_i(1 + (n - 1)(1 - \phi)) - p_2].$$

Taking into account the ordering $\bar{\varrho}_1 > \dots > \bar{\varrho}_N$, we have $\gamma_i < 0$ if $\phi < 1$ and

$$p_2 < \bar{\varrho}_i(1 + (n - 1)(1 - \phi)), \quad \text{for all } i = 1, \dots, N,$$

or equivalently,

$$0 < \psi(\Phi) := -\bar{\varrho}^T \Phi + \bar{\varrho}_N(1 + (n - 1)(1 - \phi)) \quad \text{for all } \Phi \in \bar{\mathcal{D}}_{\phi_*}.$$

In order to find conditions on ϕ_* for ensuring that $\psi(\phi) > 0$ for all $\phi \in [0, \phi_{\max}]$, we may consider the linear optimization programme

$$\min_{\phi_j \geq 0, \phi \leq \phi_*} \psi(\Phi). \quad (3.4.6)$$

It is straightforward to see that the solution of (3.4.6) is attained at a vertex of $\bar{\mathcal{D}}_{\phi_*}$. We then obtain

$$\begin{aligned} \min_{\Phi \in \bar{\mathcal{D}}_{\phi_*}} \psi(\Phi) &= \min \{ \psi((0, \dots, 0)^T), \min \{ \psi(\phi_* \mathbf{e}_1), \dots, \psi(\phi_* \mathbf{e}_N) \} \} \\ &= \min_{1 \leq j \leq N} \{ -\bar{\varrho}_j \phi_* + \bar{\varrho}_N(1 + (n - 1)(1 - \phi_*)) \} \\ &= -\bar{\varrho}_1 \phi_* + \bar{\varrho}_N(1 + (n - 1)(1 - \phi_*)). \end{aligned}$$

This directly gives the bound (3.4.5) that ensures strict hyperbolicity in $\mathcal{D}_{\phi_{\max}}$. \square

3.4.2 Preliminaries for the BW and HS models

These two models can be expressed as

$$v_i(\Phi) = d_i^2 \varphi(\mathbf{s}_i^T \Phi + n\phi)(1 - \phi)^n, \quad i = 1, \dots, N, \quad (3.4.7)$$

where $\varphi(z) = 1 + z$, $n = 0$ for the BW model, and $\varphi(z) = \exp(z)$, $n \geq 0$, arbitrary, for the HS model. We define

$$\mathbf{a}_\nu := \mathbf{d}_{\nu-1}^T := (d_1^{\nu-1}, d_2^{\nu-1}, \dots, d_N^{\nu-1}), \quad p_\nu := \mathbf{a}_\nu^T \Phi, \quad \nu = 1, \dots, 4, \quad (3.4.8)$$

and write

$$\mathbf{s}_i^T \Phi = \sum_{j=1}^N \left(\sum_{\nu=0}^3 \beta_\nu \left(\frac{d_j}{d_i} \right)^\nu \right) \phi_j = \sum_{\nu=0}^3 \frac{\beta_\nu}{d_i^\nu} \mathbf{a}_\nu^T \Phi = \sum_{\nu=0}^3 \frac{\beta_\nu}{d_i^\nu} p_{\nu+1}, \quad i = 1, \dots, N.$$

In this chapter we shall analyze only the case $\beta_3 = 0$, for which $m = 3$. We may then express (3.4.7) as

$$\begin{aligned} v_i(\Phi) &= v_i(p_1, \dots, p_3) \\ &= d_i^2 \varphi((\beta_0 + n)p_1 + \beta_1 d_i^{-1} p_2 + \beta_2 d_i^{-2} p_3) (1 - p_1)^n, \quad i = 1, \dots, N. \end{aligned} \quad (3.4.9)$$

Let us write $p_1 = \phi$, and define

$$\eta_i := \varphi(\mathbf{s}_i^T \Phi + n\phi) \quad \text{and} \quad \eta'_i := \varphi'(\mathbf{s}_i^T \Phi + n\phi) \quad \text{for } i = 1, \dots, N,$$

where $\varphi'(z) := d\varphi(z)/dz$. Taking into account that for the BW and HS models either $n = 0$ or $\eta'_i = \eta_i$, we readily compute here that the quantities $\alpha_i^k = \partial p_k / \partial \phi_i$ and $\beta_i^k = \phi_i \partial v_i / \partial p_k$ are given by

$$\alpha_i^k = d_i^{k-1}, \quad \beta_i^k = d_i^{3-k} \phi_i (1 - \phi)^n \tilde{\beta}_{k-1} \eta'_i, \quad \tilde{\beta}_0 = \beta_0 - \frac{n\phi}{1 - \phi}, \quad \tilde{\beta}_k = \beta_k, \quad k = 1, 2.$$

We now calculate the products $\alpha_I^J \beta_I^J$ of $\alpha_I^J := \det \mathbf{A}^{I,J}$ and $\beta_I^J := \det \mathbf{B}^{I,J}$ in the formula (3.2.15) for $m = 3$,

$$\begin{aligned} \gamma_i &= \alpha_i^1 \beta_i^1 + \alpha_i^2 \beta_i^2 + \alpha_i^3 \beta_i^3 n \\ &+ \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\alpha_{ij}^{12} \beta_{ij}^{12} + \alpha_{ij}^{13} \beta_{ij}^{13} + \alpha_{ij}^{23} \beta_{ij}^{23}}{v_j - v_i} + \sum_{\substack{j,k=1 \\ i \neq j < k \neq i}}^N \frac{\alpha_{ijk}^{123} \beta_{ijk}^{123}}{(v_k - v_i)(v_j - v_i)}, \end{aligned} \quad (3.4.10)$$

which is written out here in some detail for the ease of keeping track of the terms to be evaluated. Moreover, we adopt the convention that sums over a void index range are zero, and utilize the following notation:

$$\pi_{ijk} := (d_j - d_i)(d_k - d_i)(d_k - d_j). \quad (3.4.11)$$

We then obtain

$$\begin{aligned} \alpha_i^k \beta_i^k &= d_i^2 \phi_i (1 - \phi)^n \tilde{\beta}_{k-1} \eta'_i, \quad k = 1, 2, 3, \\ \alpha_{ij}^{p,p+q} \beta_{ij}^{p,p+q} &= -(d_i d_j)^{2-q} \phi_i \phi_j (1 - \phi)^{2n} \eta'_i \eta'_j \tilde{\beta}_{p-1} \tilde{\beta}_{p+q-1} (d_i^q - d_j^q)^2, \quad q = 1, 2, \\ \alpha_{ijk}^{123} \beta_{ijk}^{123} &= -\phi_i \phi_j \phi_k (1 - \phi)^{3n} \eta'_i \eta'_j \eta'_k \tilde{\beta}_0 \beta_1 \beta_2 \pi_{ijk}^2. \end{aligned} \quad (3.4.12)$$

3.4.3 The Batchelor and Wen (BW) model

We first show that for a sufficiently dilute suspension (i.e., Φ is close to zero in a sense made precise below), the BW model is stable by demonstrating that γ_i has a definite sign, at least under certain mild restrictions on the parameters d_N and β , and if the components of Φ are sufficiently small. To this end, we rewrite (3.4.10) as $\gamma_i = \phi_i(\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i})$. Inserting the expressions (3.4.11) and (3.4.12), with $n = 0$, $\tilde{\beta}_k = \beta_k$, $\eta'_i = 1$, into (3.4.10) and defining $\hat{\eta}_i := 1 + \mathbf{s}_i^T \Phi$, we obtain

$$\begin{aligned} \mathcal{S}_{1,i} &:= d_i^2(\beta_0 + \beta_1 + \beta_2), \\ \mathcal{S}_{2,i} &:= \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\phi_j}{d_j^2 \hat{\eta}_j - d_i^2 \hat{\eta}_i} (-\beta_0 \beta_1 d_i d_j (d_j - d_i)^2 - \beta_0 \beta_2 (d_j^2 - d_i^2)^2 - \beta_1 \beta_2 d_i d_j (d_j - d_i)^2), \\ \mathcal{S}_{3,i} &:= -\beta_0 \beta_1 \beta_2 \sum_{\substack{j,k=1 \\ i \neq j < k \neq i}}^N \frac{\phi_j \phi_k \pi_{ijk}^2}{(d_k^2 \hat{\eta}_k - d_i^2 \hat{\eta}_i)(d_j^2 \hat{\eta}_j - d_i^2 \hat{\eta}_i)}. \end{aligned} \tag{3.4.13}$$

Since $\beta_i \leq 0$ for $i = 0, 1, 2$ and at least one β_i is negative, we see that $\mathcal{S}_{1,i} < 0$ for $\phi_i > 0$; moreover, here $\mathcal{S}_{1,i}$ is independent of Φ or N . Consequently, we now show that $\gamma_i < 0$ for all $i = 1, \dots, N$ by possibly imposing further conditions on the parameters d_N , β and Φ . Our strategy is based on splitting the sums of (3.4.13) into positive and negative parts (produced by summands of the corresponding sign), estimating the contributions of positive sign, and then showing that these estimates ensure that $\gamma_i < 0$. To this end, suppose that there is a constant $\theta \geq 1$ such that

$$-\mathbf{s}_i^T \Phi \leq \frac{1}{1 + \theta} \quad \text{for all } i = 1, \dots, N. \tag{3.4.14}$$

Clearly, this condition is satisfied if and only if the inequality for $i = N$ is satisfied, i.e.,

$$-\mathbf{s}_N^T \Phi = \sum_{j=1}^N \left(-\sum_{\nu=0}^2 \frac{\beta_\nu d_j^\nu}{d_N^\nu} \right) \phi_j \leq \frac{1}{1 + \theta}. \tag{3.4.15}$$

(This is a combined condition on the choices of d_N , β and Φ , which we will discuss after stating the main result for this model.) This implies that $1 + \mathbf{s}_i^T \Phi \geq -\theta \mathbf{s}_i^T \Phi$ for all $i = 1, \dots, N$, i.e., it is a positive lower bound for the velocities v_i , and we then know that for $i < j$, the following inequalities are valid:

$$0 < (d_i^2 \hat{\eta}_i - d_j^2 \hat{\eta}_j)^{-1} \leq ((1 + \mathbf{s}_j^T \Phi)(d_i^2 - d_j^2))^{-1} \leq \left(-\theta (d_i^2 - d_j^2) \sum_{\nu=0}^2 \frac{\beta_\nu}{d_j^\nu} \mathbf{d}_\nu^T \Phi \right)^{-1}. \tag{3.4.16}$$

Clearly, we may further estimate the last term in (3.4.16) by omitting some of the summands.

Lemma 3.4.3 *If (3.4.15) is satisfied then we have the following inequalities:*

$$\mathcal{S}_{2,i} \leq -d_i^2 \theta^{-1} (2\beta_0 + \beta_2), \quad (3.4.17)$$

$$\mathcal{S}_{3,i} \leq -d_i^2 \theta^{-2} (2\beta_0). \quad (3.4.18)$$

Proof. Since $\hat{\eta}_i > \hat{\eta}_j$ for $i < j$, the summands of $\mathcal{S}_{2,i}$ with $j < i$ and $j > i$ are negative and positive, respectively; let us denote the corresponding partial sums by $\mathcal{S}_{2,i}^- \leq 0$ and $\mathcal{S}_{2,i}^+ \geq 0$, with $\mathcal{S}_{2,i} = \mathcal{S}_{2,i}^- + \mathcal{S}_{2,i}^+$. Our aim is to bound $\mathcal{S}_{2,i}^+$ in such a way that this quantity is compensated by the (negative) terms of \mathcal{S}_1 .

Let us now turn to $\mathcal{S}_{2,i}^+$. We here get

$$\begin{aligned} \mathcal{S}_{2,i}^+ \leq & -\frac{1}{\theta} \sum_{j=i+1}^N \left\{ \frac{\beta_0 \beta_1 d_i (d_i - d_j)^2 d_j^2 \phi_j}{(d_i^2 - d_j^2) \beta_1 \mathbf{d}_1^T \Phi} + \frac{\beta_0 \beta_2 (d_i + d_j)^2 (d_i - d_j)^2 d_j^2 \phi_j}{(d_i^2 - d_j^2) \beta_2 \mathbf{d}_2^T \Phi} \right. \\ & \left. + \frac{\beta_1 \beta_2 d_i (d_i - d_j)^2 d_j^2 \phi_j}{\beta_1 (d_i^2 - d_j^2) \mathbf{d}_1^T \Phi} \right\}. \end{aligned} \quad (3.4.19)$$

Consequently, since $d_i > d_j$ for $j > i$, we obtain from (3.4.19) the following inequality:

$$\mathcal{S}_{2,i}^+ \leq -\frac{d_i^2}{\theta} \sum_{j=i+1}^N \left\{ \beta_0 \left(\frac{d_j \phi_j}{\mathbf{d}_1^T \Phi} + \frac{d_j^2 \phi_j}{\mathbf{d}_2^T \Phi} \right) + \beta_2 \frac{d_j \phi_j}{\mathbf{d}_1^T \Phi} \right\},$$

which implies (3.4.17), given that

$$\frac{d_{i+1}^p \phi_{i+1} + \cdots + d_N^p \phi_N}{d_1^p \phi_1 + \cdots + d_N^p \phi_N} \leq 1, \quad p = 1, 2.$$

Since only those summands of $\mathcal{S}_{3,i}$ are positive for which either $i < j$ and $i < k$ or $i > j$ and $i > k$, we rewrite $\mathcal{S}_{3,i}$ as $\mathcal{S}_{3,i} = \mathcal{S}_{3,i}^- + \mathcal{S}_{3,i}^{+,1} + \mathcal{S}_{3,i}^{+,2}$, where $\mathcal{S}_{3,i}^- < 0$, $\mathcal{S}_{3,i}^{+,1} > 0$ and $\mathcal{S}_{3,i}^{+,2} > 0$, and $\mathcal{S}_{3,i}^{+,1}$ and $\mathcal{S}_{3,i}^{+,2}$ are the partial of $\mathcal{S}_{3,i}$ for which $j > i$, $k > i$ and $k \neq j$ and $j < i$, $k < i$ and $k \neq j$, respectively.

Applying several versions of (3.4.16) to both factors in the denominator of the summands of $\mathcal{S}_{3,i}^{+,1}$, we obtain

$$\mathcal{S}_{3,i}^{+,1} \leq -\frac{1}{\theta^2} \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\beta_0 \pi_{ijk}^2 d_j \phi_j d_k^2 \phi_k}{(d_k^2 - d_i^2)(d_j^2 - d_i^2) \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi}.$$

Noting that for $j, k > i$, we have that

$$\frac{\pi_{ijk}^2}{(d_i^2 - d_j^2)(d_i^2 - d_k^2)} \leq d_i^2, \quad \sum_{\substack{j,k=i+1 \\ j < k}}^N d_j \phi_j d_k^2 \phi_k \leq \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi, \quad (3.4.20)$$

we finally obtain the inequality

$$\mathcal{S}_{3,i}^{+,1} \leq -\beta_0 \frac{d_i^2}{\theta^2}. \quad (3.4.21)$$

Furthermore, the version of (3.4.16) with the roles of i and j interchanged shows that

$$\begin{aligned} \mathcal{S}_{3,i}^{+,2} &\leq -\frac{1}{\theta^2} \sum_{\substack{j,k=1 \\ k < j}}^{i-1} \frac{\beta_0 \pi_{ijk}^2 d_i \phi_j d_i^2 \phi_k}{(d_k^2 - d_i^2)(d_j^2 - d_i^2) \mathbf{d}_1^\top \Phi \mathbf{d}_2^\top \Phi} \\ &\leq -\frac{\beta_0}{\theta^2 \mathbf{d}_1^\top \Phi \mathbf{d}_2^\top \Phi} \sum_{\substack{j,k=1 \\ k < j}}^{i-1} (d_k - d_j)^2 d_i \phi_j d_i^2 \phi_k \\ &\leq -\frac{d_i^2 \beta_0}{\theta^2 \mathbf{d}_1^\top \Phi \mathbf{d}_2^\top \Phi} \sum_{\substack{j,k=1 \\ k < j}}^{i-1} d_k^2 \phi_k d_j \phi_j \leq -\beta_0 \frac{d_i^2}{\theta^2}. \end{aligned}$$

Combining this with (3.4.21) we obtain (3.4.18). \square

Corollary 3.4.1 *For the BW model, the following inequality is valid:*

$$\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i} \leq d_i^2 M(\theta, \boldsymbol{\beta}), \quad (3.4.22)$$

where we define the function

$$M(\theta, \boldsymbol{\beta}) := (1 - 2\theta^{-1} - 2\theta^{-2})\beta_0 + \beta_1 + (1 - \theta^{-1})\beta_2. \quad (3.4.23)$$

Proof. Combining the inequalities (3.4.17) and (3.4.18) we obtain (3.4.22) and (3.4.23). Each of the inequalities (3.4.17) and (3.4.18) estimates a non-negative sum from above, and therefore remains valid if the respective sum runs over a void index range, and is therefore zero. Consequently, (3.4.22) and (3.4.23) hold for arbitrary numbers of species N . \square

We have proven the following theorem.

Theorem 3.4.4 *Assume that θ is chosen such that the inequality*

$$M(\theta, \boldsymbol{\beta}) < 0 \quad (3.4.24)$$

is satisfied, where $M(\theta, \boldsymbol{\beta})$ is defined in (3.4.23). If the maximum solids concentration ϕ_{\max} is chosen such that the inequality (3.4.15) is satisfied for all $\Phi \in \mathcal{D}_{\phi_{\max}}$ for this value of θ , then $\gamma_i < 0$ for $i = 1, \dots, N$ and $\Phi \in \mathcal{D}_{\phi_{\max}}$, i.e., the model equations are strictly hyperbolic on $\mathcal{D}_{\phi_{\max}}$.

The advantage of introducing the parameter θ becomes apparent now, and is related to the fact that the BW model is valid for dilute suspensions only. Suppose that we choose an admissible value of θ , then (3.4.15) can hold either for a dilute suspension, i.e. ϕ is small, but for a large range of coefficients β , or we consider relatively small (in absolute value) coefficients β and obtain a hyperbolicity (stability) result valid up to relatively large concentrations. Our analysis also shows that for $N = 2$, we additionally have $\mathcal{S}_{3,i} = 0$ and the terms in which we divide by θ^2 are zero.

To illustrate the predictions of Theorem 3.4.4, let us first consider the coefficients β (with $\beta_3 = 0$) given by (3.3.13). Since $M(\theta, \beta)$ is a strictly decreasing function of θ , it is sufficient to solve $M(\theta, \beta) = 0$ for θ to conclude that in this case, $M(\theta, \beta) < 0$ for $\theta > \theta_{\min} \approx 2.259847$, i.e., the set of admissible values of θ is (θ_{\min}, ∞) . In this case the hyperbolicity of the BW model is ensured for those vectors Φ that satisfy (3.4.15) with a strict inequality and $\theta = \theta_{\min}$. The sharp evaluation of this inequality requires specifying d_2, \dots, d_N . However, if we only wish to determine the largest value ϕ^* of the total concentration ϕ up to which we can guarantee hyperbolicity, then we can rewrite the left-hand side of (3.4.15) as $\sigma_1\phi_1 + \dots + \sigma_N\phi_N$, where we define $\sigma_j := -\beta_0 - \beta_1 d_j d_N^{-1} - \beta_2 d_j^2 d_N^{-2}$. Then the sought concentration ϕ^* solves the problem “minimize ϕ subject to $\sigma_1\phi_1 + \dots + \sigma_N\phi_N = (1 + \theta_{\min})^{-1}$ ”. Expressing ϕ_1 in terms of ϕ_2, \dots, ϕ_N and ϕ , we can rewrite this equation as

$$\phi = (1 - \sigma_1^{-1}\sigma_2)\phi_2 + \dots + (1 - \sigma_1^{-1}\sigma_N)\phi_N + \sigma_1^{-1}(1 + \theta_{\min})^{-1}. \quad (3.4.25)$$

Since $\sigma_1 > \sigma_2 > \dots > \sigma_N$, the coefficients of ϕ_2, \dots, ϕ_N on the right-hand side are all positive, and the minimum ϕ^* of ϕ is attained for $\phi_2 = \dots = \phi_N = 0$. Its value is $\phi^* = \sigma_1^{-1}(1 + \theta_{\min})^{-1}$. Figure 3.1 (a) shows a plot of ϕ^* as a function of d_N for this case.

Finally, for the purpose of illustration, let us consider the coefficients β given (3.3.12) for large or small Péclet numbers, but where we replace the respective values of β_3 by zero. In these cases, we obtain the respective values $\theta_{\min} \approx 2.252800$ and $\theta_{\min} \approx 2.135459$, and we show in Figures 3.1 (b) and (c) the corresponding plots of ϕ^* as a function of d_N .

3.4.4 The Höfler and Schwarzer (HS) model

Let us now analyze the HS model based on the velocity equation (3.3.15). This model is the sub-case of (3.4.7) for $\varphi(z) = \exp z$ and $n \geq 0$ arbitrary. For this model, $\eta_i = \eta'_i$ causes considerable simplification, and the quantities γ_i given by (3.4.10) can be expressed as

$$\gamma_i = \phi_i(1 - \phi)^n \eta_i (\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i}), \quad (3.4.26)$$

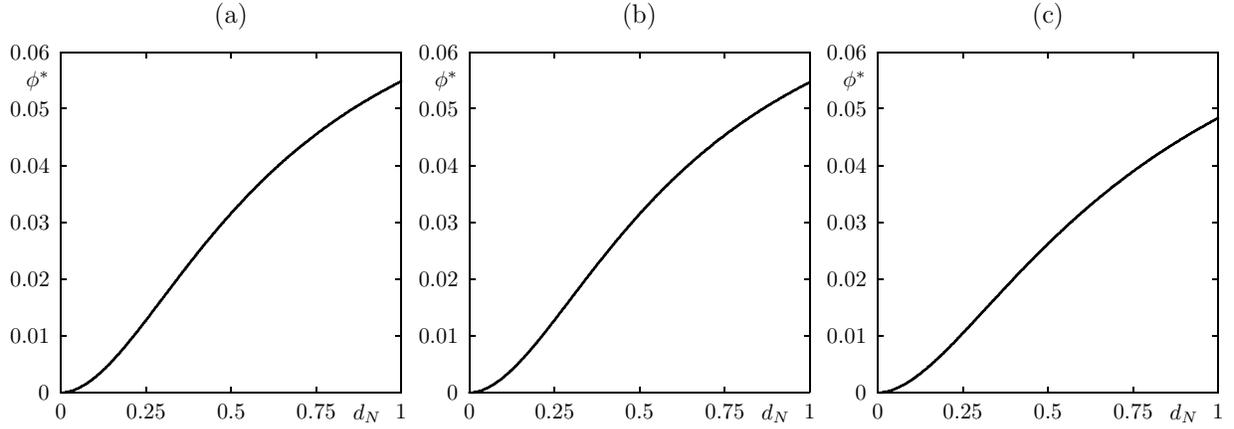


Figure 3.1: Maximum total concentrations ϕ^* for which hyperbolicity of the BW model is ensured (a) for the coefficients (3.3.13) (with $\beta_3 = 0$), (b) for $\beta_0 = -3.5$, $\beta_1 = -1.1$, $\beta_2 = -1.02$ (according to (3.3.12) for large Péclet numbers) and $\beta_3 = 0$, and (c) for $\beta_0 = -3.42$, $\beta_1 = -1.96$, $\beta_2 = -1.21$ (according to (3.3.12) for small Péclet numbers) and $\beta_3 = 0$.

where in terms of $\tilde{\eta}_i := \exp(\mathbf{s}_i^T \Phi)$ we define for the HS model

$$\begin{aligned} \mathcal{S}_{1,i} &:= d_i^2 (\tilde{\beta}_0 + \beta_1 + \beta_2), \\ \mathcal{S}_{2,i} &:= - \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\phi_j \tilde{\eta}_j}{d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i} \left\{ (d_i - d_j)^2 \tilde{\beta}_0 (\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) \right. \\ &\quad \left. + \beta_1 \beta_2 d_i d_j (d_i - d_j)^2 \right\}, \\ \mathcal{S}_{3,i} &:= -\tilde{\beta}_0 \beta_1 \beta_2 \sum_{\substack{j,k=1 \\ j < k, j,k \neq i}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)}. \end{aligned}$$

Proceeding in a similar way as for the BW model, we now show that $\gamma_i < 0$ for all vectors $\Phi \in \mathcal{D}_{\phi_{\max}}$, by possibly introducing further structural assumptions on the coefficients $\beta_0, \beta_1, \beta_2$. The decisive difference is, however, that the final result should be valid for the whole range of concentrations from the dilute to the concentrated limit since the HS model is supposed to cover this range, in contrast to the BW model.

The following lemma will be used in slight variants in several instances.

Lemma 3.4.5 *The following inequality holds for $i < j$:*

$$\begin{aligned} \frac{\tilde{\eta}_j}{d_i^2 \tilde{\eta}_i - d_j^2 \tilde{\eta}_j} &\leq - \left(e(d_i^2 - d_j^2) \sum_{s=1}^2 \beta_s \left(\frac{1}{d_j^s} - \frac{1}{d_i^s} \right) \mathbf{d}_s^T \Phi \right)^{-1} \\ &= - \frac{1}{e(d_i^2 - d_j^2)} \left[\sum_{s=1}^2 \beta_s \frac{d_i^s - d_j^s}{d_i^s d_j^s} \mathbf{d}_s^T \Phi \right]^{-1}. \end{aligned} \quad (3.4.27)$$

Proof. We first calculate for $i < j$:

$$\begin{aligned} 0 &< \frac{\tilde{\eta}_j}{d_i^2 \tilde{\eta}_i - d_j^2 \tilde{\eta}_j} = \frac{1}{d_i^2 \exp((\mathbf{s}_i^T - \mathbf{s}_j^T) \Phi) - d_j^2} \leq \frac{1}{\exp((\mathbf{s}_i^T - \mathbf{s}_j^T) \Phi) (d_i^2 - d_j^2)} \\ &= \frac{\exp((\mathbf{s}_j^T - \mathbf{s}_i^T) \Phi)}{d_i^2 - d_j^2} = \exp \left(\sum_{s=1}^2 \beta_s \left(\frac{1}{d_j^s} - \frac{1}{d_i^s} \right) \mathbf{d}_s^T \Phi \right) \frac{1}{d_i^2 - d_j^2}. \end{aligned}$$

Now, since $d_i > d_j$ for $i < j$, the argument of the exponential in the last expression is negative. Inequality (3.4.27) is now a consequence of $\exp(-\alpha) \leq e^{-1} \alpha^{-1}$ for $\alpha > 0$. \square

The expression (3.4.27) can be estimated further from above if we drop any of the three summands in the expression in squared brackets. Moreover, we first note that also for this model, $\mathcal{S}_{1,i} < 0$. Then we analyze the positive and negative parts of $\mathcal{S}_{2,i}$ and $\mathcal{S}_{3,i}$ separately, and show that we eventually obtain $\gamma_i < 0$.

Lemma 3.4.6 *Let us rewrite $\mathcal{S}_{2,i}$ as $\mathcal{S}_{2,i} = \mathcal{S}_{2,i}^+ + \mathcal{S}_{2,i}^-$, where $\mathcal{S}_{2,i}^+$ and $\mathcal{S}_{2,i}^-$ correspond to the summands of $\mathcal{S}_{2,i}$ with $j > i$ and $j < i$, respectively. Then $\mathcal{S}_{2,i}^- \leq 0$, and the following inequality holds:*

$$\mathcal{S}_{2,i}^+ \leq - \frac{d_i^2}{e} (\tilde{\beta}_0 + \beta_2). \quad (3.4.28)$$

Proof. Since $\exp(\mathbf{s}_i^T \Phi) > \exp(\mathbf{s}_j^T \Phi)$ for $i < j$ and $\exp(\mathbf{s}_i^T \Phi) < \exp(\mathbf{s}_j^T \Phi)$ for $i > j$, the factor multiplying $\{\dots\}$ in the summands of $\mathcal{S}_{2,i}^-$ is always positive, while $\{\dots\} < 0$. This confirms that $\mathcal{S}_{2,i}^- \leq 0$ (note that for $i = 1$, the sum is void, i.e. $\mathcal{S}_{2,i}^- = 0$). To estimate $\mathcal{S}_{2,i}^+$, note first that from Lemma 3.4.5 we may conclude that

$$\begin{aligned} \mathcal{S}_{2,i}^+ &\leq \frac{\tilde{\beta}_0}{e} \sum_{j=i+1}^N \frac{(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) (d_i - d_j)^2 \phi_j}{(d_i^2 - d_j^2) \left[\beta_1 \frac{d_i - d_j}{d_i d_j} \mathbf{d}_1^T \Phi + \beta_2 \frac{d_i^2 - d_j^2}{d_i^2 d_j^2} \mathbf{d}_2^T \Phi \right]} \\ &\quad - \frac{\beta_2}{e} \sum_{j=i+1}^N \frac{(d_i - d_j)^2 d_i^2 d_j^2 \phi_j}{(d_i^2 - d_j^2) (d_i - d_j) \mathbf{d}_1^T \Phi} \\ &\leq - \frac{\tilde{\beta}_0 d_i^2}{e} \sum_{j=i+1}^N \frac{d_j^2 (\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) \phi_j}{\beta_1 d_i d_j (d_i + d_j) \mathbf{d}_1^T \Phi + \beta_2 (d_i + d_j)^2 \mathbf{d}_2^T \Phi} \end{aligned}$$

$$\begin{aligned}
& - \frac{d_i^2 \beta_2}{e} \sum_{j=i+1}^N \frac{d_j^2 \phi_j}{(d_i + d_j) \mathbf{d}_1^T \Phi} \\
& \leq - \frac{d_i^2 \tilde{\beta}_0}{e} \sum_{j=i+1}^N \frac{(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) d_j^2 \phi_j}{(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) \sum_{k=i+1}^N d_k^2 \phi_k} - \frac{d_i^2 \beta_2}{e} \sum_{j=i+1}^N \frac{d_j \phi_j}{\mathbf{d}_1^T \Phi},
\end{aligned}$$

which implies (3.4.28). \square

Lemma 3.4.7 *Assume that we rewrite $\mathcal{S}_{3,i}$ as $\mathcal{S}_{3,i} = \mathcal{S}_{3,i}^- + \mathcal{S}_{3,i}^{+,1} + \mathcal{S}_{3,i}^{+,2}$, where $\mathcal{S}_{3,i}^{+,1}$ and $\mathcal{S}_{3,i}^{+,2}$ are the sums over all summands for which $j > i$, $k > i$ and $k \neq j$ and $j < i$, $k < i$ and $k \neq j$, respectively. Then we have $\mathcal{S}_{3,i}^- < 0$, $\mathcal{S}_{3,i}^{+,1} > 0$ and $\mathcal{S}_{3,i}^{+,2} > 0$. Furthermore, the following inequality holds:*

$$\mathcal{S}_{3,i}^{+,1} \leq - \frac{d_i^2 \tilde{\beta}_0}{e^2}. \quad (3.4.29)$$

Finally, let us assume that the parameters $\boldsymbol{\beta}$ are related to the sizes \mathbf{d}_1 via the condition

$$\forall 1 \leq j < i \leq N : \quad \forall \phi \in [0, \phi_{\max}] : \quad \tilde{H}_{ij}(\phi, \boldsymbol{\beta}) < 0, \quad (3.4.30)$$

where we define the functions

$$\tilde{H}_{ij}(\phi, \boldsymbol{\beta}) := -\tilde{\beta}_0 (\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) - \beta_2 \beta_1 d_i d_j - \phi (d_j - d_i)^2 \tilde{\beta}_0 \beta_1 \beta_2. \quad (3.4.31)$$

Then

$$\mathcal{S}_{2,i}^- + \mathcal{S}_{3,i}^{+,2} \leq 0. \quad (3.4.32)$$

Proof. The inequalities $\mathcal{S}_{3,i}^- < 0$, $\mathcal{S}_{3,i}^{+,1} > 0$ and $\mathcal{S}_{3,i}^{+,2} > 0$ are a simple consequence of the fact that only those summands of $\mathcal{S}_{3,i}$ are positive for which either $i < j$ and $i < k$ or $i > j$ and $i > k$, according to the ordering $d_1 > d_2 > \dots > d_N$. To deal with

$$\mathcal{S}_{3,i}^{+,1} = - \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2 \tilde{\beta}_0 \beta_1 \beta_2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)}, \quad (3.4.33)$$

note first that, based on formulas similar to (3.4.27), we get

$$\begin{aligned}
& - \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2 \tilde{\beta}_0 \beta_1 \beta_2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)} \\
& \leq - \frac{\tilde{\beta}_0}{e^2} \sum_{\substack{j,k=i+1 \\ j \neq k}}^N \frac{\phi_j \phi_k d_i^3 d_j d_k^2 (d_i - d_j)^2 (d_j - d_k)^2 (d_k - d_i)^2}{(d_i + d_j)(d_i - d_j)^2 (d_i + d_k)^2 (d_i - d_k)^2 \mathbf{d}_2^T \Phi \mathbf{d}_1^T \Phi}
\end{aligned}$$

$$\leq -\frac{d_i^2 \tilde{\beta}_0}{e^2} \sum_{\substack{j,k=i+1 \\ j \neq k}}^N \frac{\phi_j \phi_k d_j d_k^2}{\mathbf{d}_2^T \Phi \mathbf{d}_1^T \Phi} \leq -\frac{d_i^2 \tilde{\beta}_0}{e^2}.$$

Next, we see that the term

$$\mathcal{S}_{3,i}^{+,2} := - \sum_{\substack{j,k=1 \\ j < k}}^{i-1} \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2 \tilde{\beta}_0 \beta_1 \beta_2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)}$$

cannot be estimated easily. However, we may “compensate” this term with $\mathcal{S}_{2,i}^-$, as expressed in (3.4.32). (To ensure that our hyperbolicity result is also valid for $N = 3$, $\mathcal{S}_{3,i}^{+,1}$ should be compensated by one of the terms that have arisen *earlier* in our analysis.) Observe now that

$$\mathcal{S}_{2,i}^- + \mathcal{S}_{3,i}^{+,2} = \sum_{j=1}^{i-1} \frac{\phi_j \tilde{\eta}_j (d_i - d_j)^2}{d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i} \mathcal{R}_{ij}, \quad (3.4.34)$$

where we define

$$\begin{aligned} \mathcal{R}_{ij} &:= -\tilde{\beta}_0 (\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) - \beta_2 \beta_1 d_i d_j + \tilde{\mathcal{R}}_{ij}, \\ \tilde{\mathcal{R}}_{ij} &:= -\tilde{\beta}_0 \beta_1 \beta_2 \sum_{k=j+1}^{i-1} \frac{\phi_k (d_k - d_i)^2 (d_k - d_j)^2 \tilde{\eta}_k}{d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i}. \end{aligned} \quad (3.4.35)$$

Since $d_i < d_j$ and $d_i < d_k$ in these summands, and the factor multiplying \mathcal{R}_{ij} in (3.4.34) is positive, we will satisfy (3.4.32) by achieving that $\mathcal{R}_{ij} < 0$. Noting that for $j < k < i$

$$\begin{aligned} \frac{(d_k - d_i)^2 (d_k - d_j)^2 \tilde{\eta}_k}{d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i} &= \frac{(d_k - d_i)^2 (d_k - d_j)^2}{d_k^2 - d_i^2 \exp((\mathbf{s}_i^T - \mathbf{s}_k^T) \Phi)} \\ &\leq \frac{(d_k - d_j)^2 (d_k - d_i)}{d_k + d_i} \leq (d_j - d_i)^2, \end{aligned}$$

we conclude that $\tilde{\mathcal{R}}_{ij} \leq -\tilde{\beta}_0 \beta_1 \beta_2 (d_j - d_i)^2 (\phi_{j+1} + \phi_{j+2} + \cdots + \phi_{i-1})$. Thus, (3.4.32) holds if the parameters $\boldsymbol{\beta}$ are related to d_1, \dots, d_N by (3.4.30), where $\tilde{H}_{ij} := \tilde{H}_{ij}(\phi; \boldsymbol{\beta})$ is defined in (3.4.31). \square

Summarizing, and collecting the inequalities for the various terms, we see that if (3.4.30) is met then

$$\begin{aligned} \mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i} &= \mathcal{S}_{1,i} + \mathcal{S}_{2,i}^- + \mathcal{S}_{2,i}^+ + \mathcal{S}_{3,i}^- + \mathcal{S}_{3,i}^{+,1} + \mathcal{S}_{3,i}^{+,2} \\ &< \mathcal{S}_{1,i} + \mathcal{S}_{2,i}^+ + \mathcal{S}_{3,i}^{+,1} \leq d_i^2 M(\phi, \boldsymbol{\beta}) < 0, \end{aligned} \quad (3.4.36)$$

where we define the function

$$M(\phi, \boldsymbol{\beta}) := (1 - e^{-1} - e^{-2}) \tilde{\beta}_0 + \beta_1 + (1 - e^{-1}) \beta_2. \quad (3.4.37)$$

Furthermore, for the discussion of models with a large number of species N , the criterion (3.4.30) with \tilde{H}_{ij} defined by (3.4.31) is inconvenient, since it requires inspection of a large number of size pairs d_i and d_j . Thus, we propose a sufficient condition for (3.4.30) to be satisfied for all pairs $j < i$. To this end, we fix a pair $i > j$, define $\delta := \delta_{ij} = d_i/d_j$, and divide (3.4.31) by d_j^2 to obtain

$$\tilde{H}_{ij} = -\tilde{\beta}_0(\beta_1\delta + \beta_2(1 + \delta)^2) - (\beta_2\beta_1\delta) - \phi(1 - \delta)^2\tilde{\beta}_0\beta_1\beta_2. \quad (3.4.38)$$

Since δ varies between d_N and one, a sufficient condition for (3.4.30) to be satisfied is given by

$$\forall \phi \in [0, \phi_{\max}] : \quad H(\phi, \boldsymbol{\beta}, d_N) < 0, \quad (3.4.39)$$

where the following definition of $H(\phi, \boldsymbol{\beta}, d_N)$ is derived from the observation that the two terms in the first line of (3.4.38) are non-positive, while the term in the second line is non-negative:

$$H(\phi, \boldsymbol{\beta}, d_N) := -\tilde{\beta}_0(\beta_1 d_N + \beta_2(1 + d_N)^2) - \beta_2\beta_1 d_N - \phi(1 - d_N)^2\tilde{\beta}_0\beta_1\beta_2. \quad (3.4.40)$$

Theorem 3.4.8 *Assume that the vector of parameters $\boldsymbol{\beta}$, the maximum solids concentration ϕ_{\max} and the width of the particle size distribution, characterized by the value of $d_N \in (0, 1]$, are chosen such that the inequality (3.4.39) is satisfied, where the expression $H(\phi, \boldsymbol{\beta}, d_N)$ is defined by (3.4.40). Then $\gamma_i < 0$ for $i = 1, \dots, N$, i.e., the model equations are strictly hyperbolic for $\Phi \in \mathcal{D}_{\phi_{\max}}$.*

For the coefficients $\boldsymbol{\beta}$ (with $\beta_3 = 0$) given by (3.3.13) and $n = 2$, the curve $H(\phi, \boldsymbol{\beta}, d_N) = 0$ is plotted in Figure 3.2 (a) in a ϕ versus d_N plot. It turns out that for $d_N > d_N^* := 0.0078595$ (this number is a solution of $H(1, \boldsymbol{\beta}, d_N) = 0$), the HS model equations are strictly hyperbolic in $\mathcal{D}_{\phi_{\max}}$ without any restrictions on ϕ_{\max} . Note that for $0 < d_N < d_N^*$, condition (3.4.39) is violated only for values of ϕ_{\max} very close to one. In fact, Figure 3.2 (a) indicates that the HS model with the parameters (3.3.13) is strictly hyperbolic for arbitrarily small values of d_N if we set $\phi_{\max} \leq 0.96$. Given that d_N^* is already a small number, we can say that hyperbolicity holds for almost all cases of practical interest for this model. Figures 3.2 (b) and (c) show the corresponding result for the two sets of parameters given by (3.3.12) (but with $\beta_3 = 0$). In these two cases we obtain larger values of d_N^* than in Figure 3.2 (a), but hyperbolicity is still ensured for a large range of cases of practical interest.

Finally, we remark here that the value $n = 2$ in (3.3.15) was utilized in the examples of Figure 3.2, and that very similar curves are obtained for alternative values $1 \leq n \leq 5$.

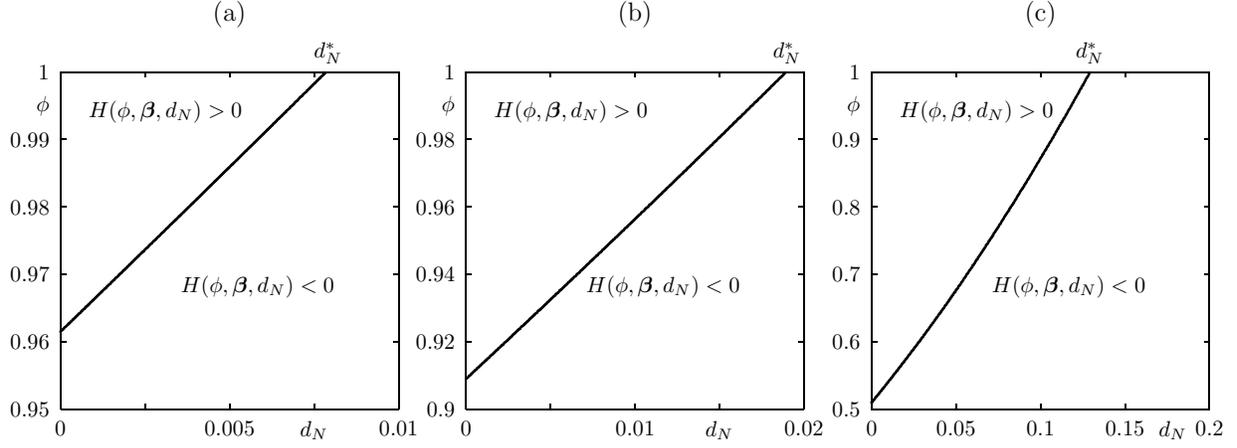


Figure 3.2: Region of hyperbolicity ($H(\phi, \beta, d_N) < 0$) for the HS model (a) for the coefficients (3.3.13) (with $\beta_3 = 0$), (b) for $\beta_0 = -3.5$, $\beta_1 = -1.1$, $\beta_2 = -1.02$ (according to (3.3.12) for large Péclet numbers) and $\beta_3 = 0$, and (c) for $\beta_0 = -3.42$, $\beta_1 = -1.96$, $\beta_2 = -1.21$ (according to (3.3.12) for small Péclet numbers) and $\beta_3 = 0$.

3.5 Numerical examples

We apply the Roe scheme and KT schemes to the MLB and HS models to simulate batch settling of a suspension with equal-density particles in a vessel of normalized depth one. We first briefly describe both schemes, and refer to [43] and [41] for the Roe and KT scheme, respectively. We discretize the spatial domain $[0, 1]$ into M cells of size $\Delta x = 1/M$. The time step is denoted by Δt , and we define $x_j := j\Delta x$ and $t_n := n\Delta t$. Furthermore, we assume that $\lambda := \Delta t/\Delta x$ is fixed by an appropriate CFL condition.

3.5.1 The Roe and KT schemes

The conservative form of the Roe scheme for (3.1.1) is given by

$$\Phi_j^{n+1} = \Phi_j^n - \lambda(\mathcal{F}_{j+1/2}^n - \mathcal{F}_{j-1/2}^n), \quad j = 1, \dots, M, \quad n = 0, 1, 2, \dots, \quad (3.5.1)$$

where $\Phi_j^n = (\phi_{1,j}^n, \dots, \phi_{N,j}^n)^\top$ and the numerical flux vector is defined as

$$\mathcal{F}_{j+1/2}^n = \frac{1}{2}[\mathbf{f}(\Phi_{j+1}^n) + \mathbf{f}(\Phi_j^n)] - \frac{1}{2}(\alpha_1|\lambda_1|\mathbf{r}_1 + \dots + \alpha_N|\lambda_N|\mathbf{r}_N). \quad (3.5.2)$$

Here, $\lambda_1, \dots, \lambda_N$ are the eigenvalues of the Jacobian \mathcal{J}_f evaluated at

$$\Phi_{j+1/2}^n = \frac{1}{2}(\Phi_{j+1}^n + \Phi_j^n),$$

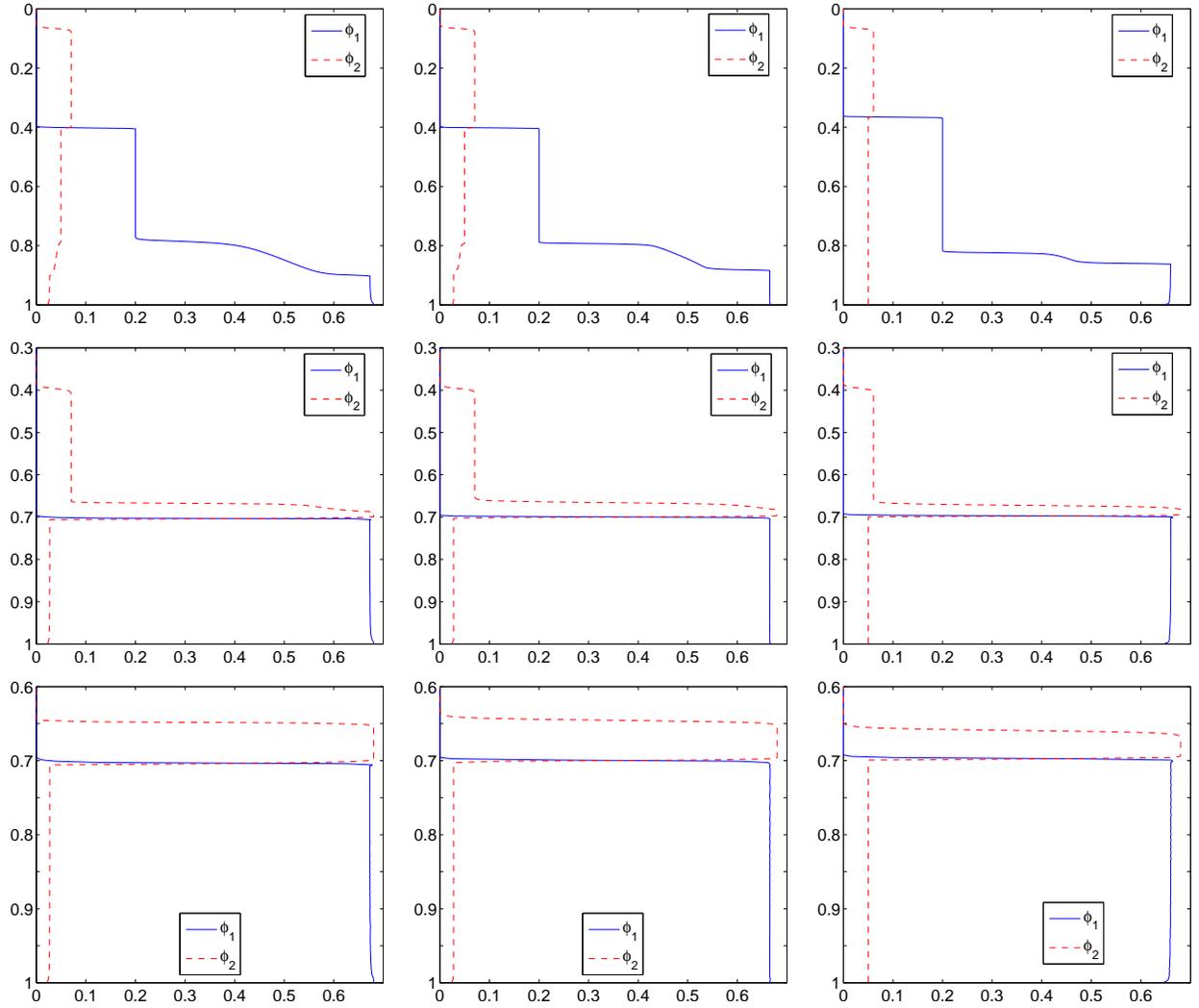


Figure 3.3: Example 1 (MLB and HS models with $N = 2$): numerical solution at $t = 50$ s (top), $t = 300$ s (middle) and $t = 1000$ s (bottom). The horizontal axis in each plot denotes concentration, and the vertical axis denotes normalized depth. Here and in Figure 3.4 the left and middle column correspond to the Roe and KT scheme, respectively, applied to the MLB model and the right column corresponds to the KT scheme applied to the HS model.

which are calculated by a root finder starting from the velocities

$$v_1(\Phi_{j+1/2}^n), \dots, v_N(\Phi_{j+1/2}^n).$$

The components of $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N)^T$ are given by $\Phi_{j+1}^n - \Phi_j^n = \alpha_1 \mathbf{r}_1 + \dots + \alpha_N \mathbf{r}_N$, or equivalently, $\boldsymbol{\alpha} = \mathbf{R}^{-1}(\Phi_{j+1}^n - \Phi_j^n)$. Here, $\mathbf{r}_1, \dots, \mathbf{r}_N$ are the normalized right eigenvectors of $\mathcal{J}_{\mathbf{f}}(\Phi_{j+1/2}^n)$, which form the columns of \mathbf{R} . The characteristic information is given by the

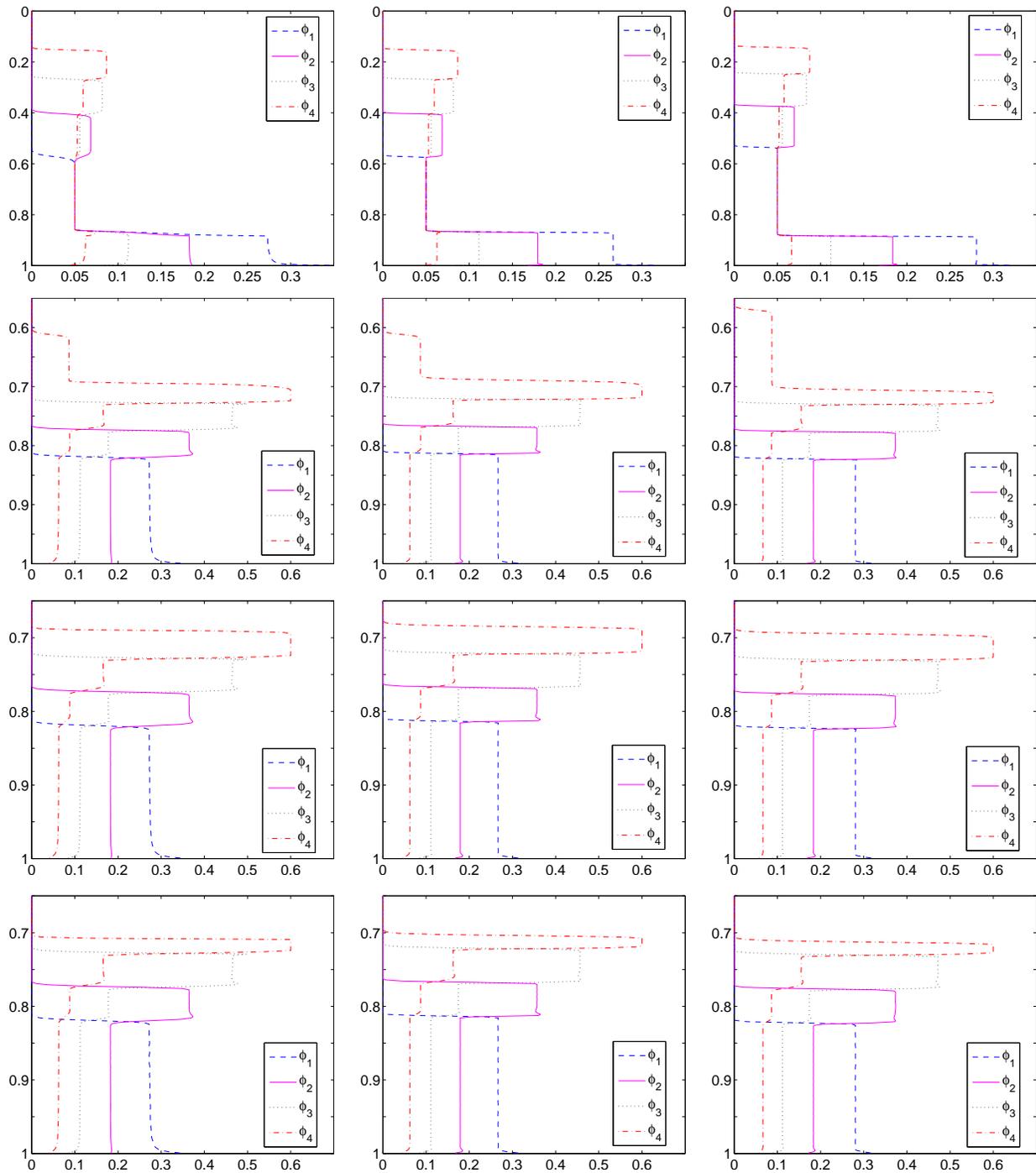


Figure 3.4: Example 2 (MLB and HS models, $N = 4$): numerical solution at $t = 50$ s (top row), $t = 200$ s (second row), $t = 300$ s (third row) and $t = 1000$ s (bottom row) with $\Delta x = 0.0005$.

secular equation and (3.4.2) and (3.4.3) for the MLB model. For a given CFL number CFL, we employ $\Delta t = \text{CFL}\Delta x/\rho$, where ρ the biggest characteristic velocity, i.e.,

$$\rho = \max_{j=1,\dots,M} \max_{i=1,\dots,N} |\lambda_i(\Phi_{j+1/2})|$$

In contrast to the just-described Roe scheme, the second-order central KT scheme [41] does not require knowledge of the complete eigenstructure of the problem. However, this method does rely on the local speed of wave propagation. The semi-discrete conservative form of the KT scheme is

$$\frac{d\Phi_j}{dt} = -\frac{1}{\Delta x}(\mathcal{H}_{j+1/2} - \mathcal{H}_{j-1/2}), \quad (3.5.3)$$

with the numerical flux vector

$$\mathcal{H}_{j+1/2} := \frac{1}{2}[\mathbf{f}(\Phi_{j+1/2}^+) - \mathbf{f}(\Phi_{j+1/2}^-)] - \frac{a_{j+1/2}}{2}(\Phi_{j+1/2}^+ - \Phi_{j+1/2}^-).$$

The extrapolated values $\Phi_{j+1/2}^\pm$ are

$$\Phi_{j+1/2}^+ = \Phi_{j+1} - (\Delta x/2)(\Phi_x)_{j+1} \quad \text{and} \quad \Phi_{j+1/2}^- = \Phi_j + (\Delta x/2)(\Phi_x)_j,$$

where $a_{j+1/2}$ is the maximal local speed, which we take as

$$a_{j+1/2} := \max\{\rho(\mathcal{J}_f(\Phi_{j+1/2}^+)), \rho(\mathcal{J}_f(\Phi_{j+1/2}^-))\},$$

where $\rho(\mathcal{J}_f(\cdot))$ denotes the spectral radius of the matrix $\mathcal{J}_f(\cdot)$. The approximate spatial derivative of $\Phi(x, t)$ is computed using a θ -minmod limiter for each component of Φ . In our implementations, we use $\theta = 1.3$.

For time discretization we use the optimal third-order TVD Runge-Kutta methods (see e.g. [61]) with CFL = 0.5. Then we take $\Delta t = \text{CFL}\Delta x/S_{\max}^n$, where S_{\max}^n denotes the biggest local propagation speed throughout the domain at time t_n . For both the Roe and KT schemes we utilize zero-flux boundary conditions, i.e. $\mathbf{f}_{|x=0} = \mathbf{f}_{|x=L} = 0$, which in the numerical approach corresponds to take numerical fluxes equal to zero at both ends of the spatial domain.

3.5.2 Numerical examples

The numerical experiments are similar to those of [19]; Examples 1, 2 and 3 (see Figs. 3.3, 3.4 and 3.5 respectively) correspond to the cases $N = 2$, $N = 4$ and $N = 11$, respectively. All examples are based on the physical parameters $g = 9.81 \text{ m/s}^2$, $\mu_f = 0.02416 \text{ Pas}$, $\varrho_f =$

| | | | | | | |
|------------------------|--------|--------|--------|--------|--------|--------|
| i | 1 | 2 | 3 | 4 | 5 | 6 |
| $\phi_i^0[10^{-3}]$ | 0.435 | 3.747 | 14.420 | 32.603 | 47.912 | 47.762 |
| $D_i[10^{-5}\text{m}]$ | 8.769 | 8.345 | 7.921 | 7.497 | 7.073 | 6.649 |
| d_i | 1.000 | 0.952 | 0.903 | 0.855 | 0.807 | 0.758 |
| i | 7 | 8 | 9 | 10 | 11 | |
| $\phi_i^0[10^{-3}]$ | 32.663 | 15.104 | 4.511 | 0.783 | 0.060 | |
| $D_i[10^{-5}\text{m}]$ | 6.225 | 5.801 | 5.377 | 4.953 | 4.529 | |
| d_i | 0.710 | 0.662 | 0.613 | 0.565 | 0.516 | |

Table 3.1: MLB and HS models, $N = 11$: initial concentrations ϕ_i^0 , real particle sizes D_i , and normalized particle sizes d_i .

1208 kg/m^3 , $\varrho_1 = \dots = \varrho_N = \varrho_s = 2790 \text{ kg/m}^3$ that correspond to a standard published experiment [57]. The function $V(\phi)$ in the MLB model has the exponent $n = 4.7$, except for $N = 11$ in which case we choose $n = 4.65$, while the parameters β_0, \dots, β_3 for the HS models are those given by (3.3.13) (with $\beta_3 = 0$).

In Example 1, the original depth of the vessel is $L = 0.3 \text{ m}$; this is also true for $N = 4$, and the unnormalized particle diameters are $D_1 = 4.96 \times 10^{-4} \text{ m}$ and $D_2 = 1.25 \times 10^{-4} \text{ m}$, corresponding to $d_1 = 1$ and $d_2 = D_2/D_1 = 0.25202$. The maximum total concentration is $\phi_{\max} = 0.68$, along with the initial concentrations $\Phi^0 = (\phi_1^0, \phi_2^0) = (0.2, 0.05)^T$. For Example 2, we choose $d_1 = 1$, $d_2 = 0.8$, $d_3 = 0.6$ and $d_4 = 0.4$, $\phi_{\max} = 0.6$, and $\phi_i^0 = 0.05$ for $i = 1, \dots, 4$. Finally, in Example 3 for $N = 11$, which is based on experimental data from [59], we consider $L = 0.935 \text{ m}$ and $\phi_{\max} = 0.641$. The initial concentrations ϕ_i^0 , diameters D_i and normalized diameters $d_i = D_i/D_1$ are given in Table 3.1.

3.6 Conclusions of Chapter 3

Our analysis illustrates the use of the secular equation as a tool for the hyperbolicity analysis for polydisperse sedimentation models, and leads to estimates of hyperbolicity regions that qualitatively agree with the ranges of validity of the MLB, BW and HS models; recall that the BW model is valid for dilute suspensions only, which is consistent with the limitations visible in Figure 3.1. For the BW and HS models, only the sign of the coefficients β_0, \dots, β_2 , but not the values, enter our analysis; results will only change quantitatively for other sets of parameters. In addition, a similar analysis could be advanced for the case that β_3

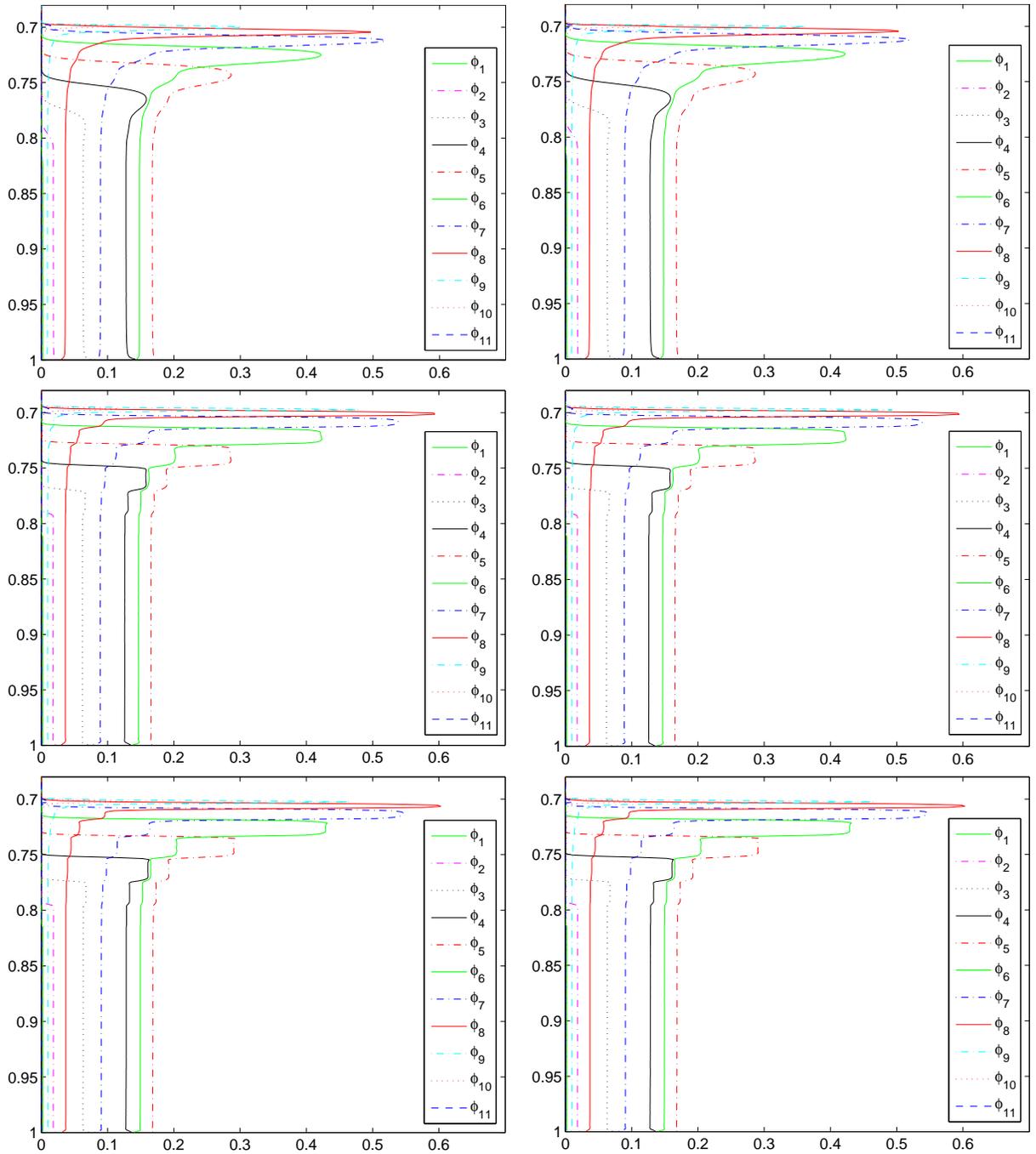


Figure 3.5: Example 3 (MLB and HS models, $N = 11$): numerical solution at $t = 600$ s (first column) and $t = 1000$ s (second column) with $\Delta x = 0.0005$. The top and middle rows correspond to the MLB model using Roe and KT schemes respectively. The bottom row corresponds to the HS model with the KT scheme.

is a small, but positive parameter. Previous works [26, 27, 70] already illustrated applications of the secular equation [1] for perturbations of diagonal matrices of ranks $m = 1$ and $m = 2$ (for the multi-class Lighthill-Whitham-Richards traffic and the MLB sedimentation model, respectively), but it is here for the first time (to our knowledge) that this result is led to practical use for $m = 3$.

Clearly, our analysis is subject to limitations in terms of the accuracy of the estimates of the hyperbolicity region. Theorems 3.4.4 and 3.4.8 state in which regions hyperbolicity is *ensured*, that is, where we can guarantee that $\gamma_i \cdot \gamma_j > 0$. However, this property is a sufficient, but not a necessary condition to ensure hyperbolicity; the models may well be hyperbolic in other sub-regions of parameter space, but with $\gamma_i \cdot \gamma_j \leq 0$ for some choices of i and j . While this is an intrinsic limitation of the secular equation, our analysis of the HS model shows that slightly larger hyperbolicity regions may be obtained for a given set of particle sizes d_1, \dots, d_N if the functions \tilde{H}_{ij} given by (3.4.38) (rather than the single function $H(\phi, \beta, d_N)$) are evaluated. Also, further realism can be added to polydisperse sedimentation models if the phase space is not simply limited by a hyperplane $\phi = \phi_{\max}$, but by a curved surface in $\bar{\mathcal{D}}_1$ which takes into account that mixtures of small and large particles permit denser packings than monodisperse sediments of any of the species involved.

Despite these limitations, the present calculus can be extended in several possible ways. First of all, we selected the MLB, BW and HS models because the computations are slightly different in each of these cases. In particular, our interest in the HS model is motivated by a result from [17] stating that the HS model is, unlike the DG model, hyperbolic for $N = 2$ without further restrictions. More advanced models that should be analyzed include the models presented in [36, 49, 58]. The model by Patwardhan and Tien [49] generalizes the MLB model, and utilizes a more involved expression for v_i . The models in [36] are further modifications of the BW model; they consider S_{ij} to be a rational (rather than polynomial) function of d_j/d_i . Finally, the difficulty associated with the model by Selim et al. [58], which is otherwise similar to the MLB model, consists in the postulated dependence of v_i on partial sums like $\phi_1 + \dots + \phi_{i-1}$; it is unclear at the moment whether this model can be transformed so that (3.1.2) is satisfied.

Let us also mention that although we focus here on spatially one-dimensional models, the present hyperbolicity calculus remains valid for the two- or three-dimensional version of (3.1.1). In fact, in that case the model equation (3.1.1) is replaced by $\partial_t \phi_i + \nabla \cdot (\mathbf{q} \phi_i + f_i(\Phi) \mathbf{k}) = 0$ for $i = 1, \dots, N$, where \mathbf{q} is the volume-averaged mixture flow velocity (for which additional equations, e.g. a version of the Navier-Stokes equations have to be solved), and \mathbf{k} is the

downwards-pointing unit vector. This equation is hyperbolic if and only if (3.1.1) with the same vector $\mathbf{f}(\Phi) = (f_1(\Phi), \dots, f_N(\Phi))^T$ is so (see [11] for details).

Concerning the numerical examples, observe that in all examples the model parameters have been chosen such that both the MLB and HS models are strictly hyperbolic on $\mathcal{D}_{\phi_{\max}}$ according to the analysis of Section 3.4. Our results illustrate that for both the MLB and HS models, solutions for equal-density spheres with a constant initial composition Φ^0 typically evolve as follows: if $\Phi^0 \in \mathcal{D}_{\phi_{\max}}$, i.e., $\phi_i^0 > 0$ for $i = 1, \dots, N$, then the bottommost zone will contain particles of all species, the next zone will contain species 2 to N (i.e., $\phi_1 = 0$), the next only species 3 to N (i.e., $\phi_1 = \phi_2 = 0$), and so on, until a zone is reached in which only (the smallest) species N is present, followed by a zone void of particles ($\Phi = 0$). The composition of each of these zones corresponds to the situation addressed in Remark 3.2.2, i.e., strict hyperbolicity is ensured, and is also obtained from a method of solution based on the construction of kinematic shocks that separate areas of constant composition [35, 57].

Having said this, we mention that the construction of exact solutions to the system of conservation laws (3.1.1) that satisfy an entropy condition is complicated since most choices of $\mathbf{f}(\Phi)$ will lead to a system of conservation laws that, in the best case, can be proven to be strictly hyperbolic on $\mathcal{D}_{\phi_{\max}}$ (as a consequence of our analysis), but whose characteristic fields in general are neither linearly degenerate nor genuinely nonlinear, which rules out, for example, the use of Lax’s shock admissibility criterion. A suitable shock admissibility criterion is Liu’s entropy condition (see e.g. [22] for details on these criteria). The construction of solutions that satisfy this condition has been undertaken so far only for $N = 2$ [9].

Moreover, in our numerical simulations, we have chosen fairly simple schemes, which nevertheless utilize characteristic information that we do now have access to thanks to the secular equation. However, the true strength of the availability of characteristic information lies in the possibility to utilize high-resolution “spectral” schemes, such as the one introduced in [27] for the multiclass LWR traffic model. In terms of resolution these schemes are a possible alternative to component-wise discretizations such as WENO [73] or WENO-multiresolution [19] schemes, which are the standard at present since the effort needed to obtain this information has hitherto been considered excessive. We will come back to spectral schemes for the present models in the next chapter.

Chapter 4

On the implementation of WENO schemes for a class of polydisperse sedimentation models

4.1 Introduction

4.1.1 Scope

This chapter deals with high-resolution numerical schemes for systems of conservation laws that arise as one-dimensional kinematic models for the sedimentation of polydisperse suspensions which were considered in Chapter 3.

The purpose is to demonstrate that very efficient high-order accurate weighted essentially non-oscillatory (WENO) schemes for the numerical solution of (1.1.1)–(1.1.3) can indeed be constructed by incorporating characteristic information. This information is available due to the recent hyperbolicity analysis made in Chapter 3 [14], and can be incorporated in various ways. Specifically, we use these results in order to provide a good estimation of the viscosity coefficient in a Lax-Friedrichs-type flux splitting. This allows to construct high resolution component-wise WENO schemes, akin to those proposed in [72] for the Multiclass Lighthill-Whitham-Richards (MCLWR) models in traffic flow. In addition, the full spectral decomposition of $\mathcal{J}_f(\Phi)$, which can be numerically computed at each cell interface thanks to the analysis in [14], can be used in order to obtain *characteristic-based* WENO schemes, for which the WENO reconstruction procedure is applied to the local characteristic variables and fluxes at each cell-interface. When combined with a strong stability preserving (SSP)

Runge-Kutta-type time discretization, the resulting SSP-WENO-SPEC schemes are shown to be extremely robust in a number of numerical experiments concerning the MLB and HS models, including several properties specific to the present application such as non-negativity of the solution, almost avoidance of overshoots of the numerical total density ϕ beyond ϕ_{\max} , and accurate rendering of stationary kinematic shocks that separate sediment layers of different composition.

4.1.2 Related work

WENO-type spatial flux reconstructions, which emerged from earlier essentially non-oscillatory (ENO) schemes, have become a well-established, versatile tool for the construction of high-resolution conservative schemes in numerous applications. The first WENO scheme, of third-order accuracy, was introduced by Liu, Osher and Chan in [44], while a general framework to construct WENO schemes of arbitrary order of accuracy was provided by Jiang and Shu [39]. We refer to Shu [61, 60] for further details, applications, and references. If applied to a system of conservation laws, the WENO procedure will produce a spatially semi-discrete system of ODE, for which a discretization in time can be chosen separately [62]. A suitable choice are total variation diminishing Runge-Kutta schemes [31, 61], also known as strong stability preserving (SSP) methods [32], because of their favorable stability properties.

While WENO-based high-resolution shock-capturing schemes have been applied successfully to a wide range of convection-dominated problems [60], the polydisperse sedimentation models considered herein present some specific challenges for numerical simulation. These models belong to the wider class of multi-species kinematic flow models [19], which are characterized by a governing system of equations of the type (1.1.1) with explicit velocity functions v_1, \dots, v_N for a number N of species. Models of this type include, besides the sedimentation model, a model of settling of oil-in-water dispersion [53] and, most notably, the multi-class Lighthill-Whitham-Richards (MCLWR) kinematic traffic model, which extends the well-known LWR model to vehicles with drivers having different preferential velocities, and which was proposed by Benzoni-Gavage and Colombo [7] and Wong and Wong [66]. Meanwhile, the MCLWR model has been studied thoroughly in a series of papers including [19, 71, 70, 72, 73]. All these models can be formulated for an arbitrary number N of species, that is, of scalar equations. The basic phenomenon of interest is the segregation of species, i.e. the formation of areas of different composition from an initially homogeneous “mixture”

(e.g., suspension or traffic platoon). Segregation is usually associated with the formation of discontinuities in Φ , so-called kinematic shocks. For the sedimentation model considered for batch settling in a column, stationary kinematic shocks separate sediment layers of different composition. The accurate numerical simulation of the model is therefore of importance for the prediction of the composition of the sediment as a final “product” or deposit e.g. in medicine, the manufacturing of functionally graded materials, volcanology, and petrology (see e.g. [3] for references to these applications).

It is well known that high-resolution shock capturing schemes can be applied to systems of conservation laws either in a component-wise or in a characteristic-wise (spectral) fashion. The latter requires a detailed knowledge of the spectral decomposition of the Jacobian matrix of the system, since the eigenstructure is used in a fundamental way in the design principles of the scheme [27]. For multi-species kinematic flow models, however, eigenvalues are not available in closed form, nevertheless it has been possible to prove strict hyperbolicity of some of these models by an explicit representation of the characteristic polynomial [11, 53, 71], as well as to obtain an interlacing property of the (unknown) eigenvalues of the Jacobian $\lambda_1, \dots, \lambda_N$ with the (known) velocities v_1, \dots, v_N , which provide excellent starting values for a root finder. For the MCLWR model, the corresponding hyperbolicity and characteristic analysis was first done by Zhang et al. [70]. In [73], solutions to this model, with the additional complication of a discontinuously varying coefficient modeling variable road surface conditions, were compared with solutions generated by a component-wise WENO scheme which provides good overall accuracy. The first implementation of a component-wise WENO scheme for that model had been done by Zhang in [71], and pre-dates the hyperbolicity analysis in [70]. An improvement of the component-wise scheme in [70] is presented in [72], and used in subsequent papers. It amounts to using a more appropriate choice of the viscosity coefficient in a Lax-Friedrichs flux splitting, and is based on sharper bound for the smallest eigenvalue obtained from the hyperbolicity analysis of the MCLWR model carried out in [70]. We shall see that the results in [14] easily lead to an analogous estimation of the viscosity coefficient for the polydisperse sedimentation models considered in this chapter.

The hyperbolicity analysis for the MCLWR model is, in fact, fairly straightforward since that model gives rise to a Jacobian which is a rank-1 perturbation of a diagonal (see Section 3.1.1). In [27], this feature was exploited in order to give a much simpler proof of the hyperbolicity of the MCLWR model. The full spectral decomposition of the Jacobian matrix was then used to construct a characteristic-based version of the schemes utilized in [71], and it was demonstrated that the resolution of the characteristic-wise WENO schemes is

superior to that of their component-wise counterpart. Most notably, solutions are much less oscillatory.

In a later paper [26], Donat and Mulet showed that the hyperbolicity calculus of multi-species kinematic flow models can be greatly simplified if one employs the so-called secular equation due to Anderson [1]. Roughly speaking, the secular equation provides a systematic algebraic framework to determine the eigenvalues, and eventually the eigenvectors if the Jacobian is a rank- m , $m \ll N$, perturbation of the diagonal, and most importantly avoids the explicit representation of the characteristic polynomial. Donat and Mulet [26] showed that via the secular approach, hyperbolicity of the MLB model for equal-density spheres (a case of $m = 2$) can be proved in a few lines, which contrasts with several pages of computation necessary to exhibit the characteristic polynomial in [11]. In [14] we showed that the secular approach can also be used to estimate the region of hyperbolicity of the HS model, which corresponds to $m = 3$ or $m = 4$. In this chapter, we use the results of Chapter 3 [14] to provide a counterpart of [27] for the MLB and HS models, namely we show that the results in [14] permit to implement characteristic-wise WENO schemes, and that these are robust and have favorable properties analogous to those in [27].

This chapter is organized as follow. In Sections 4.2 present a variation of the HS model from the original version discussed in the last chapter. We shall prove the hyperbolicity (strictly) of this variation from the theorem 3.4.8 and derive the expressions for computing the eigenvectors. The numerical schemes are described in Section 4.3, starting with a spatially semi-discrete formulation and the implementation of the boundary conditions (Sects. 4.3.1 and 4.3.2), which is converted into a fully discrete scheme by a strong stability preserving Runge-Kutta scheme (SSPRK; Sect. 4.3.3). We then proceed with a general discussion of flux vector splitting (Sect. 4.3.4). Then, in Sections 4.3.4, and 4.3.5, which are at the core of this chapter, we describe how the explicit algebraic form of the velocities v_1, \dots, v_N in conjunction with the characteristic information, namely the interlacing property of eigenvalues with phase velocities and the left eigenvectors, can be used to define viscosity coefficients for the characteristic-wise computation of the flux vectors. The resulting scheme is addressed by SPEC-INT scheme; its counterpart based on less involved component-wise flux vector splitting is referred to as COMP-GLF scheme. In Section 4.4 we present a series of numerical examples for the MLB model with $N = 2, 4$ and 11, along with error histories, that illustrate the superiority of SPEC-INT (compared with COMP-GLF) in terms of accuracy. Additional examples suggest that the scheme is equally suitable for the HS model. Finally, in Section 4.5 we list some conclusions, address limitations of the applicability of the scheme and point out

possible extensions.

4.2 Hyperbolicity analysis for a variation of the HS model

In order to ensure that $v_i \rightarrow 0$ continuously for the HS model when $\phi \rightarrow \phi_{\max}$ we consider the following variant (see [38]) of the HS model:

$$v_i(\Phi) = d_i^2 \exp\left(\mathbf{s}_i^T \Phi + n \frac{\phi}{\phi_{\max}}\right) \left(1 - \frac{\phi}{\phi_{\max}}\right)^n, \quad n \geq 0, \quad (4.2.1)$$

which is defined for $\Phi \in \mathcal{D}_{\phi_{\max}}$. We will refer to this variation as the “modified HS model”.

Introducing the variable $\hat{\Phi} := (\phi_{\max}^{-1})\Phi$ we can write (4.2.1) as follows:

$$v_i(\hat{\Phi}) = d_i^2 \exp(\hat{\mathbf{s}}_i^T \hat{\Phi} + n\hat{\phi}) (1 - \hat{\phi})^n = d_i^2 \exp((\hat{\beta}_0 + n)\hat{p}_1 + \hat{\beta}_1 d_i^{-1} \hat{p}_2 + \hat{\beta}_2 d_i^{-2} \hat{p}_3) (1 - \hat{p}_1)^n$$

where $\hat{\mathbf{s}}_i^T = (\hat{S}_{i1}, \dots, \hat{S}_{iN})^T = \phi_{\max} \mathbf{s}_i^T$, $i = 1, \dots, N$. The available hyperbolicity analysis for the HS model can be applied to analyze the hyperbolicity of the modified HS model if we define the coefficients $\hat{\beta}_k := \phi_{\max} \beta_k$ for $k = 0, 1, 2$ and the quantities $\hat{p}_\nu = \mathbf{a}_\nu^T \hat{\Phi}$, $\nu = 1, 2, 3$. We can now apply Theorem 3.4.8 of chapter 3 to deduce that the modified HS model is strictly hyperbolic if

$$H(\hat{\phi}, \hat{\boldsymbol{\beta}}, d_N) := -\tilde{\beta}_0 (\hat{\beta}_1 d_N + \hat{\beta}_2 (1 + d_N)^2) - \hat{\beta}_2 \hat{\beta}_1 d_N - \hat{\phi} (1 - d_N)^2 \tilde{\beta}_0 \hat{\beta}_1 \hat{\beta}_2 < 0, \quad (4.2.2)$$

where we define

$$\tilde{\beta}_0 = \hat{\beta}_0 - \frac{n\hat{\phi}}{1 - \hat{\phi}}.$$

Then a simple algebraic computation shows (using the coefficients $\boldsymbol{\beta}$ given by (3.3.13)) that (4.2.2) holds if $d_N > 0.0078595$, that is, with the same mild restriction of the original HS model. This is the matter of the next lemma.

Lemma 4.2.1 *Assume that the coefficients $\boldsymbol{\beta}$ are given by (3.3.13). Then the modified HS model specified by phase velocities v_i given by (4.2.1) is strictly hyperbolic on $\mathcal{D}_{\phi_{\max}}$ if $d_N > 0.0078595$.*

Proof. To make the basic idea transparent, let us write $\hat{\phi}(\phi) = \phi/\phi_{\max}$ and $\hat{\boldsymbol{\beta}}(\boldsymbol{\beta}) = \phi_{\max} \boldsymbol{\beta}$. Then it is sufficient to notice that we can write

$$H(\hat{\phi}(\phi), \hat{\boldsymbol{\beta}}(\boldsymbol{\beta}), d_N) = \phi_{\max}^2 (H(\phi, \boldsymbol{\beta}, d_N) - C\tilde{H}(\phi, \boldsymbol{\beta}, d_N)), \quad (4.2.3)$$

where

$$\tilde{H}(\phi, \boldsymbol{\beta}, d_N) := (\beta_1 d_N + \beta_2(1 + d_N)^2 + \phi(1 - d_N)^2 \beta_1 \beta_2),$$

and

$$C := n\phi \left(\frac{1}{1 - \phi} - \frac{1}{\phi_{\max}(\phi_{\max} - \phi)} \right) < 0.$$

Since the variant of the model is strictly hyperbolic if $H(\hat{\phi}(\phi), \hat{\boldsymbol{\beta}}(\boldsymbol{\beta}), d_N) < 0$, it is sufficient to show that $\tilde{H}(\phi, \boldsymbol{\beta}, d_N) < 0$, but this statement is true if $d_N > 0.0078595$. \square

The spectral decomposition of $\mathcal{J}_{\mathbf{f}}(\Phi)$ is not provided in chapter 3 [14] in a form similar to that in Lemma 3.4.1, but it is easy to obtain from Theorem 3.2.1. According to Theorem 3.2.1, if $\lambda \neq v_i$ is an eigenvalue of $\mathcal{J}_{\mathbf{f}}(\Phi)$, then $\det \mathbf{M}_\lambda = 0$ and therefore $\boldsymbol{\xi} = (\xi_1, \xi_2, \xi_3)$ is a non-trivial solution of $\mathbf{M}_\lambda \boldsymbol{\xi} = \mathbf{0}$ for

$$\begin{aligned} \xi_1 &= (1 + [\mathbf{a}_2, \mathbf{b}_2])(1 + [\mathbf{a}_3, \mathbf{b}_3]) - [\mathbf{a}_2, \mathbf{b}_3][\mathbf{a}_3, \mathbf{b}_2], \\ \xi_2 &= [\mathbf{a}_2, \mathbf{b}_3][\mathbf{a}_3, \mathbf{b}_1] - [\mathbf{a}_2, \mathbf{b}_1](1 + [\mathbf{a}_3, \mathbf{b}_3]), \\ \xi_3 &= [\mathbf{a}_2, \mathbf{b}_1][\mathbf{a}_3, \mathbf{b}_2] - (1 + [\mathbf{a}_2, \mathbf{b}_2])[\mathbf{a}_3, \mathbf{b}_1]. \end{aligned}$$

Hence, by using 3.2.14 we obtain the following right eigenvector $\mathbf{x} = (x_1, \dots, x_N)^T$ for the HS model:

$$x_i = -\frac{1}{v_i - \lambda} (b_{i,1}\xi_1 + b_{i,2}\xi_2 + b_{i,3}\xi_3), \quad i = 1, \dots, N. \quad (4.2.4)$$

The left eigenvectors can be obtained by interchanging the roles of \mathbf{A} and \mathbf{B} .

4.3 Numerical schemes

4.3.1 Semi-discrete schemes

The schemes considered herein are based on the finite difference paradigm due to Shu and Osher [62] of first setting up a conservative spatial semi-discretization of the term $\partial_x \mathbf{f}(\Phi)$ and then to apply an SSP ODE solver to get a fully discrete conservative scheme with a high order of accuracy. Specifically, if we discretize the spatial domain $[0, 1]$ (after normalization) into M cells of size $\Delta x = 1/M$ and define the cell centers $x_j := (j + 1/2)\Delta x$, $j = 0, \dots, M - 1$ and the cell interfaces $x_{j+1/2} = (j + 1)\Delta x$, then the approximation to $\partial_x \mathbf{f}(x_j, t)$ is obtained by an essentially non-oscillatory reconstruction operator \mathcal{R} , applied to the fluxes $\mathbf{f}(\Phi)$ so that:

$$\partial_x \mathbf{f}(x_j, t) = \frac{1}{\Delta x} \left(\hat{\mathbf{f}}_{j+1/2} - \hat{\mathbf{f}}_{j-1/2} \right) + \mathcal{O}(\Delta x^r),$$

where r is the order of accuracy of the reconstruction and the numerical fluxes $\hat{\mathbf{f}}_{j+1/2}$ are given by

$$\begin{aligned}\hat{\mathbf{f}}_{j+1/2} &= \mathcal{R}\left(\mathbf{f}(\Phi(x_{j-s}, t)), \dots, \mathbf{f}(\Phi(x_{j+s+1}, t)); x_{j+1/2}\right) \\ &= \hat{\mathbf{f}}_{j+1/2}(\Phi(x_{j-s}, t), \dots, \Phi(x_{j+s+1}, t)).\end{aligned}\quad (4.3.1)$$

If we define the vector $\Phi := (\Phi_{-s}, \Phi_{-s+1}, \dots, \Phi_{M+s-2}, \Phi_{M+s-1})^T$, this procedure yields the semi-discrete scheme (method of lines)

$$\frac{d\Phi_j}{dt} = \mathcal{L}_j(\Phi) := -\frac{1}{\Delta x} (\hat{\mathbf{f}}_{j+1/2}(\Phi_{j-s}, \dots, \Phi_{j+s+1}) - \hat{\mathbf{f}}_{j-1/2}(\Phi_{j-s-1}, \dots, \Phi_{j+s})), \quad (4.3.2)$$

for approximations $\Phi_j(t) \approx \Phi(x_j, t)$, $j = 0, \dots, M-1$. Therefore, if we define the vector $\mathcal{L} := (\mathcal{L}_0, \dots, \mathcal{L}_{M-1})^T$, then (4.3.2) can be compactly written as

$$\frac{d\Phi(t)}{dt} = \mathcal{L}(\Phi(t)). \quad (4.3.3)$$

For well-known stability reasons, the reconstruction operator should be “upwind-biased”. In the scalar case, this means that \mathcal{R} should not depend on its last argument if $f' > 0$ and should not depend on its first argument if $f' < 0$. For nonlinear fluxes (mandatorily near sonic points, where $f' = 0$) a flux splitting approach, where

$$f = f^- + f^+, \quad f_u^+ > 0, \quad f_u^- < 0,$$

is used in order to define the numerical flux, so that

$$\begin{aligned}\hat{f}_{j+1/2} &= \mathcal{R}^+(f^+(\Phi_{j-s}), \dots, f^+(\Phi_{j+s}); x_{j+1/2}) \\ &\quad + \mathcal{R}^-(f^-(\Phi_{j-s+1}), \dots, f^-(\Phi_{j+s+1}); x_{j+1/2}),\end{aligned}\quad (4.3.4)$$

for upwind-biased reconstructions \mathcal{R}^\pm . In this work, \mathcal{R}^\pm is chosen as the mapped WENO5 (WENO5M) reconstruction, proposed in [40], to avoid a possible loss of accuracy around extrema. This technique can be extended to vectors of fluxes by its application either to each component of the system (“component-wise” schemes) or by local characteristic projections (“characteristic-wise” schemes).

4.3.2 Boundary conditions

The zero-flux boundary conditions are implemented by setting

$$\hat{\mathbf{f}}_{-1/2} = \hat{\mathbf{f}}_{M-1/2} = 0. \quad (4.3.5)$$

We recall that a WENO5 scheme requires to consider two additional ghost cells on each boundary of the computational domain. In order to guarantee that all the interpolatory stencils remain inside of the computational domain we employ a suggestion given in [61]: we set large values for the concentrations in the ghost cells, which produce large variations, so that the WENO procedure avoids the use of any stencil involving the ghost cells.

4.3.3 Time discretization

Among the variety of explicit SSP (*strong-stability preserving*) time discretization methods for the approximate solution of (4.3.3) we use the well known optimal third-order, three-stage Runge-Kutta method referred to as SSPRK(3,3), which for (4.3.3) is given by

$$\begin{aligned}\Phi^{(1)} &= \Phi^\nu + \Delta t \mathcal{L}(\Phi^\nu), \\ \Phi^{(2)} &= \frac{3}{4}\Phi^\nu + \frac{1}{4}\Phi^{(1)} + \frac{1}{4}\Delta t \mathcal{L}(\Phi^{(1)}), \\ \Phi^{\nu+1} &= \frac{1}{3}\Phi^\nu + \frac{2}{3}\Phi^{(2)} + \frac{2}{3}\Delta t \mathcal{L}(\Phi^{(2)}).\end{aligned}\tag{4.3.6}$$

SSP time discretization methods are widely used for hyperbolic PDE because they preserve the nonlinear stability properties which are necessary for problems with non-smooth solutions. On the other hand, due to convexity, the intermediate stages of the SSPRK methods have SSP properties (*i.e.*, $\|\Phi^{(i)}\| \leq \|\Phi^{(i-1)}\|$ for the internal stages). This feature is especially important for some applications [29]. For sedimentation problems it avoids unphysical negative concentrations in the internal stages. Notice that it is necessary to evaluate three times the operator $\mathcal{L}(\cdot)$ in order to move forward one time step, in fact, the *effective SSP coefficient* for SSPRK(3,3) (which is defined as [29, 55, 56] the SSP coefficient of the method divided by the number of stages) is equal to 1/3.

To satisfy the CFL condition the value of Δt is computed adaptively for each step ν . More exactly, the solution $\Phi^{\nu+1}$ at $t_{\nu+1} = t_\nu + \Delta t$ is calculated from Φ^ν by using the time step $\Delta t = \text{CFL} * \Delta x / \rho'_{\max}$, where ρ'_{\max} is an estimate of the maximal characteristic velocity for Φ^ν .

4.3.4 Flux vector splittings and viscosity coefficients

As mentioned before, a flux splitting of the type $f = f^+ + f^-$ with $f_u^+ > 0$ and $f_u^- < 0$ is required when the flux function is nonlinear. A standard recipe is provided by the Lax-

Friedrichs flux vector splitting,

$$f^+(u) = \frac{1}{2}(f(u) + \alpha u), \quad f^-(u) = \frac{1}{2}(f(u) - \alpha u) \quad (4.3.7)$$

where the viscosity coefficient α has to verify that all eigenvalues of $f_u + \alpha I$ are ≥ 0 and all eigenvalues of $f_u - \alpha I$ are ≤ 0 . Obviously, a choice such as

$$\alpha = \max_{j=0, \dots, M-1} \max_{1 \leq k \leq N} |\lambda_j^k|, \quad (4.3.8)$$

guarantees these inequalities, and we remark that α above can be easily computed for the polydisperse models being studied, since the necessary eigenvalues can be computed in an efficient manner by applying a root finder. However, we can readily apply the results in Lemmas 3.4.1 and 3.4.8 (or 4.2.1) in order to provide an estimate for (4.3.8) which does not require the computation of the eigenvalues, and is 'optimal', in the sense specified in [73],

$$\alpha = \max_{j=0, \dots, M-1} \max \left\{ \left| v_1(\Phi_j) + \sum_{k=1}^N \gamma_k(\Phi_j) \right|, |v_N(\Phi_j)| \right\}. \quad (4.3.9)$$

The choice of the viscosity coefficients (4.3.8) and (4.3.9) is *global*, hence it can be used at each cell interface, however, the resulting schemes tend to be too dissipative, even when using a characteristic-wise high resolution shock capturing scheme (see the results in [26]). In order to reduce the dissipation effects associated to the global choice of viscosity coefficient described above, a Local Lax Friedrichs (LLF) approach was proposed in [62]. The original viscosity coefficient for the computation of the numerical flux at the $i + 1/2$ interface by the LLF flux splitting approach is given in [62] by

$$\alpha_{j+1/2}^k = \max_{\Phi \in \Gamma} |\lambda_k(\Phi)|, \quad k = 1, \dots, N, \quad (4.3.10)$$

where $\Gamma := \Gamma(\Phi_j, \Phi_{j+1}) \subset \mathbb{R}^N$ is a path in phase space connecting Φ_j and Φ_{j+1} , for example a straight line. Since the characteristic fields are neither genuinely nonlinear nor linearly degenerate, the standard choice

$$\alpha_{j+1/2}^k = \max \{ |\lambda_k(\Phi_j)|, |\lambda_k(\Phi_{j+1})| \}. \quad (4.3.11)$$

will not be appropriate. Indeed, in the numerical experiments section we shall see that (4.3.11) produces numerical oscillations which do not disappear upon mesh refinement. Hence, the extrema of $\lambda_k(\Phi)$ over Γ in (4.3.10) needs to be computed. Since there is no closed form for the eigenvalues, this is not an easy task. However, we notice that the interlacing property (3.2.9) showed in Corollary 3.2.1 implies that

$$\max_{\Phi \in \Gamma} |\lambda_k(\Phi)| \leq \max_{\Phi \in \Gamma} \{ |v_{k-1}(\Phi)|, |v_k(\Phi)| \}, \quad k = 1, \dots, N, \quad (4.3.12)$$

where we set $v_0 := M_1$. Let us consider Γ as the straight line joining Φ_j and Φ_{j+1} , the minimum and the maximum of

$$g_k(a) := v_k(a\Phi_j + (1-a)\Phi_{j+1}), \quad a \in [0, 1],$$

for each value of $k \in \{1, \dots, N\}$ can be computed as the minimum or maximum of the extremal set

$$\mathcal{E}_k(\Phi_j, \Phi_{j+1}) := \{g_k(0), g_k(1)\} \cup \{g_k(a) : g'_k(a) = 0, a \in (0, 1)\}. \quad (4.3.13)$$

For the MLB model, we obtain

$$a = a_k = \frac{(n-1)(d_k^2 - \mathbf{d}_2^T(\Phi_{j+1}))(p_j - p_{j+1}) + (1-p_{j+1})\mathbf{d}_2^T(\Phi_j - \Phi_{j+1})}{n(p_j - p_{j+1})\mathbf{d}_2^T(\Phi_j - \Phi_{j+1})}, \quad (4.3.14)$$

where p_j is the value of $p = \phi$ associated with node j .

For the modified HS model we have

$$a = a_k = \frac{\phi_{\max} - p_{j+1}}{p_j - p_{j+1}} \frac{n}{\left(\beta_0 + \frac{n}{\phi_{\max}}\right)(p_j - p_{j+1}) + \beta_1 d_k^{-1} \mathbf{d}^T(\Phi_j - \Phi_{j+1}) + \beta_2 d_k^{-2} \mathbf{d}_2^T(\Phi_j - \Phi_{j+1})}. \quad (4.3.15)$$

Hence, the viscosity coefficient

$$\alpha_{j+1/2}^k = \max_{\Phi \in [\Phi_j, \Phi_{j+1}]} \{|v_{k-1}(\Phi)|, |v_k(\Phi)|\} \quad (4.3.16)$$

where $[\Phi_j, \Phi_{j+1}]$ denotes the straight line joining Φ_j and Φ_{j+1} can be readily computed at each cell interface. As we shall see in the numerical experiments section, (4.3.16) provides an adequate recipe for the local viscosity coefficient required by the LLF approach.

4.3.5 The SPEC and COMP schemes

A component-wise WENO5 scheme is defined by the numerical flux

$$\hat{f}_{j+1/2,k} = \mathcal{R}^+(f_{j-2,k}^+, \dots, f_{j+2,k}^+; x_{j+1/2}) + \mathcal{R}^-(f_{j-1,k}^-, \dots, f_{j+3,k}^-; x_{j+1/2}), \quad (4.3.17)$$

where \mathcal{R} is the mapped WENO5 reconstruction operator [40] and $f_{j,k}^\pm$ are given by the global Lax-Friedrichs flux splitting

$$(f_{j,1}^\pm, \dots, f_{j,N}^\pm)^\top = \mathbf{f}(\Phi_j) \pm \alpha \Phi_j, \quad j \in \mathbb{Z},$$

with α as defined in (4.3.9). Notice that this globally defined viscosity coefficient does not require the spectral information of the Jacobian matrix. We remark that the viscosity coefficient relies on the computation of the γ_k coefficients provided in Lemmas 3.4.1 and 4.2.1 which is consistent with the observations in [72], about the need to have a proper estimate of the minimal viscosity coefficient given by (4.3.8).

We use

$$\rho_{\max}^{\nu} = \max_{j=0,\dots,M-1} \max \left\{ \left| v_1(\Phi_j^{\nu}) + \sum_{k=1}^N \gamma_k(\Phi_j^{\nu}) \right|, |v_N(\Phi_j^{\nu})| \right\}$$

to estimate the maximal characteristic velocity for Φ^{ν} . The resulting scheme will be referred to as ‘‘COMP-GLF’’.

In order to implement a characteristic-wise scheme, we need the complete eigenstructure of $\mathcal{J}_{\mathbf{f}}(\Phi)$, which is provided by the results of the hyperbolicity analysis performed in Subsection 3.4.1 of Chapter 3 and Section 4.2 of current Chapter. The normalized left eigenvectors $(\mathbf{l}_{j+1/2}^k)^{\mathsf{T}}$ and right eigenvectors $\mathbf{r}_{j+1/2}^k$, $k = 1, \dots, N$, of

$$\mathcal{J}_{\mathbf{f}}(\Phi_{j+1/2}), \quad \Phi_{j+1/2} := \frac{1}{2}(\Phi_j + \Phi_{j+1}),$$

are computed using (3.4.2) and (3.4.3) for the MLB model and (4.2.4) for the HS model. The matrices

$$\mathbf{R}_{j+1/2} = [\mathbf{r}_{j+1/2}^1, \dots, \mathbf{r}_{j+1/2}^N], \quad \mathbf{R}_{j+1/2}^{-1} = [\mathbf{l}_{j+1/2}^1, \dots, \mathbf{l}_{j+1/2}^N]^{\mathsf{T}},$$

are needed in order to compute the local characteristic variables and fluxes around the $j+1/2$ interface as follows:

$$g_{j+1/2,i,k} := (\mathbf{l}_{j+1/2}^k)^{\mathsf{T}} \mathbf{f}(\Phi_{j+i}), \quad g_{j+1/2,i,k}^{\pm} := \frac{1}{2} (\mathbf{l}_{j+1/2}^k)^{\mathsf{T}} (\mathbf{f}(\Phi_{j+i}) \pm \alpha_{j+1/2}^k \Phi_{j+i}),$$

$$i = -2, \dots, 3, \quad j \in \mathbb{Z}, \quad k = 1, \dots, N,$$

with $\alpha_{j+1/2}^k$ given by (4.3.16). For the spectral scheme we compute the numerical fluxes as

$$\hat{\mathbf{f}}_{j+1/2} = (\hat{f}_{j+1/2,1}, \dots, \hat{f}_{j+1/2,N})^{\mathsf{T}} = \mathbf{R}_{j+1/2} \hat{\mathbf{g}}_{j+1/2}, \quad j \in \mathbb{Z}, \quad (4.3.18)$$

where $\hat{\mathbf{g}}_{j+1/2} = (\hat{g}_{j+1/2,1}, \dots, \hat{g}_{j+1/2,N})^{\mathsf{T}}$ is defined as follows. If $\lambda_j^k \cdot \lambda_{j+1}^k \leq 0$ (Case 1), we set

$$\hat{g}_{j+1/2,k} = \mathcal{R}^+(g_{j+1/2,-2,k}^+, \dots, g_{j+1/2,2,k}^+; x_{j+1/2})$$

$$+ \mathcal{R}^-(g_{j+1/2,-1,k}^-, \dots, g_{j+1/2,3,k}^-; x_{j+1/2}), \quad (4.3.19)$$

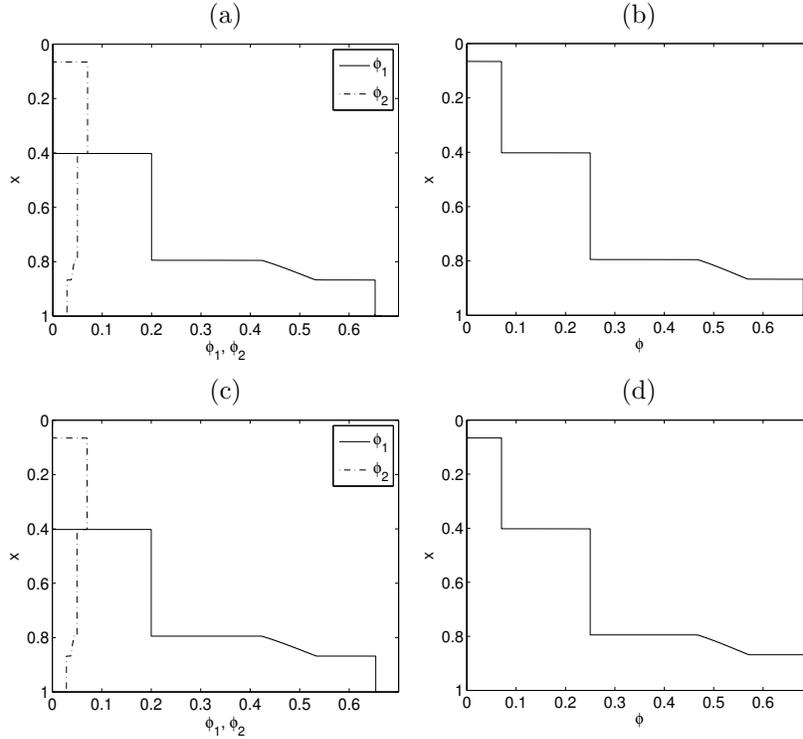


Figure 4.1: Example 4: numerical solution for ϕ_1 , ϕ_2 (a, c) and ϕ (b, d) at $t = 50$ s computed by SPEC-INT with $M = 6400$ (a, b) and COMP-GLF with $M = 25600$ (c, d).

while for $\lambda_j^k \cdot \lambda_{j+1}^k > 0$ (Case 2), we set for $k = 1, \dots, N$.

$$\hat{g}_{j+1/2,k} = \begin{cases} \mathcal{R}^+(g_{j+1/2,-2,k}, \dots, g_{j+1/2,2,k}; x_{j+1/2}) & \text{if } \lambda_j^k > 0 \text{ and } \lambda_{j+1}^k > 0, \\ \mathcal{R}^-(g_{j+1/2,-1,k}, \dots, g_{j+1/2,3,k}; x_{j+1/2}) & \text{if } \lambda_j^k < 0 \text{ and } \lambda_{j+1}^k < 0. \end{cases} \quad (4.3.20)$$

We estimate the maximal characteristic velocity for Φ^ν by:

$$\rho_{\max}^\nu = \max_{j=0, \dots, M-1} \max_{k=1, \dots, N} |\lambda_k(\Phi_{j+1/2}^\nu)|.$$

In what follows, we will address by “SPEC-INT” the characteristic-wise mapped fifth-order WENO scheme whose numerical fluxes are calculated by (4.3.18)–(4.3.20), and where the viscosity coefficient is calculated by (4.3.16) based on the interlacing property. Alternatively, for comparison purposes we will in one case employ the same scheme with the viscosity coefficient given by the usual choice (4.3.11) (instead of (4.3.16)). This scheme will be referred to as “SPEC-LLF”.

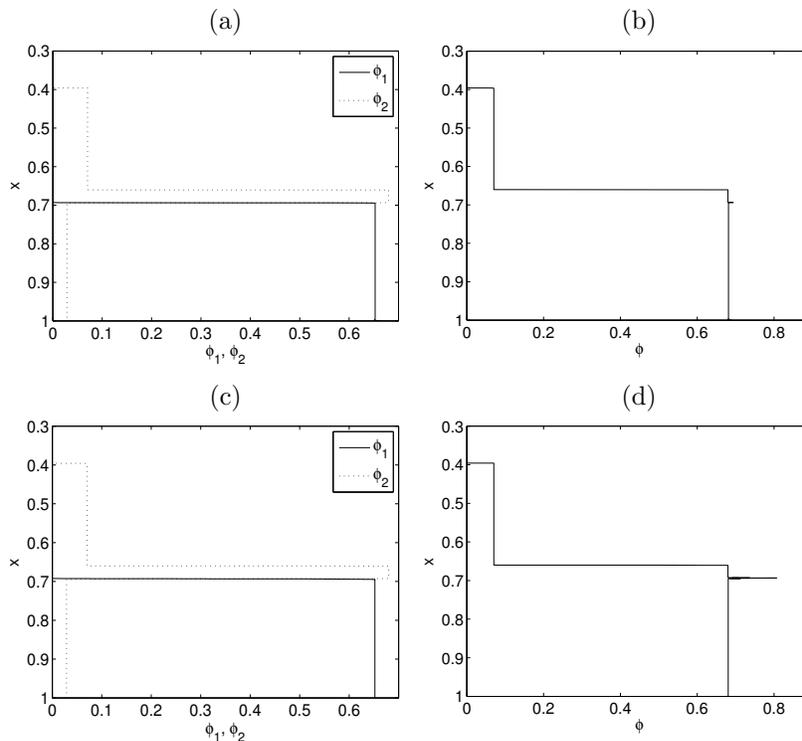


Figure 4.2: Example 4: numerical solution for ϕ_1 , ϕ_2 (a, c) and ϕ (b, d) at $t = 300$ s computed by SPEC-INT with $M = 6400$ (a, b) and COMP-GLF with $M = 25600$ (c, d).

4.4 Numerical results

In this section we perform a series of numerical experiments to highlight the numerical issues brought up earlier in the chapter. In particular, we shall see that characteristic based WENO schemes are indeed more robust than their component-wise counterparts, and that the choice of viscosity is important in the overall performance of the scheme: an incorrect choice of the viscosity coefficient in the splitting strategy can lead to an oscillatory behavior that remains under mesh refinement. In this section, we take $\text{CFL} = 0.5$ for all examples with two species and $\text{CFL} = 0.2$ for $N = 4, 11$. The data and physical parameters are the same as those used in Chapter 3 (Subsection 3.5.2), however, for the sake of readability we will re-write them for each case.

| M | e_1 | cr | e_2 | cr | e_{tot} | cr | e_1 | cr | e_2 | cr | e_{tot} | cr |
|----------------------|--------|-------|-------|-------|------------------|-----------------------|--------|-------|--------|-------|------------------|-------|
| SPEC-INT, $t = 50$ s | | | | | | SPEC-INT, $t = 300$ s | | | | | | |
| 100 | 126.04 | - | 22.41 | - | 131.10 | - | 121.20 | - | 204.41 | - | 113.24 | - |
| 200 | 63.37 | 0.992 | 10.62 | 1.077 | 64.72 | 1.018 | 65.40 | 0.890 | 107.38 | 0.929 | 56.80 | 0.995 |
| 400 | 30.54 | 1.053 | 5.42 | 0.970 | 31.60 | 1.035 | 33.76 | 0.954 | 55.53 | 0.951 | 30.07 | 0.918 |
| 800 | 16.03 | 0.930 | 2.62 | 1.051 | 16.35 | 0.950 | 14.94 | 1.177 | 26.79 | 1.052 | 15.75 | 0.932 |
| 1600 | 6.94 | 1.207 | 1.21 | 1.113 | 7.19 | 1.185 | 7.71 | 0.953 | 12.99 | 1.045 | 7.10 | 1.149 |
| COMP-GLF, $t = 50$ s | | | | | | COMP-GLF, $t = 300$ s | | | | | | |
| 100 | 169.64 | - | 43.37 | - | 187.63 | - | 175.79 | - | 530.19 | - | 423.23 | - |
| 200 | 87.06 | 0.962 | 20.52 | 1.080 | 94.26 | 0.993 | 86.40 | 1.025 | 255.21 | 1.055 | 219.27 | 0.949 |
| 400 | 44.76 | 0.960 | 9.69 | 1.082 | 47.56 | 0.987 | 45.40 | 0.928 | 186.98 | 0.449 | 174.58 | 0.329 |
| 800 | 23.94 | 0.903 | 4.82 | 1.009 | 25.10 | 0.923 | 33.31 | 0.447 | 64.09 | 1.545 | 54.82 | 1.671 |
| 1600 | 13.18 | 0.860 | 2.41 | 0.998 | 13.60 | 0.884 | 25.05 | 0.411 | 53.33 | 0.265 | 43.67 | 0.328 |

Table 4.1: Example 4: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr). The reference solution is computed by SPEC-INT with $M = 6400$.

4.4.1 Example 4 (MLB model, $N = 2$)

The first example [19, 57] corresponds to two species with density $\varrho_s = 2790 \text{ kg/m}^3$ and different diameters $D_1 = 4.96 \times 10^{-4} \text{ m}$ and $D_2 = 1.25 \times 10^{-4} \text{ m}$, corresponding to $d_1 = 1$ and $d_2 = D_2/D_1 = 0.25202$. The (unnormalized) depth of the vessel in the original experiment [57] is $L = 0.3 \text{ m}$. The maximum total concentration is $\phi_{\text{max}} = 0.68$, and the initial concentrations are $\Phi^0 = (\phi_1^0, \phi_2^0) = (0.2, 0.05)^T$. The hindered settling factor $V(\phi)$ is chosen according to (3.3.2) with the exponent $n = 4.7$. The remaining parameters are $g = 9.81 \text{ m/s}^2$, $\mu_f = 0.02416 \text{ Pa s}$ and $\varrho_f = 1208 \text{ kg/m}^3$. Moreover, here and in the following examples, the spatial coordinate x refers to normalized depth, and varies between $x = 0$ (meniscus of the suspension) and $x = 1$ (bottom of the settling column). The solution of Example 4 is well known, and has been used as a test case for a variety of methods (see, e.g., [9, 14, 15, 19]).

To compare the performance of SPEC-INT with that of COMP-GLF, we calculate numerical solutions for a sequence of spatial discretizations $\Delta x = 1/M$, and compare the solutions with two alternative reference solutions that have been computed with $M = M_{\text{ref}} = 6400$ and $M = M_{\text{ref}} = 25600$ by SPEC-INT and COMP-GLF, respectively. These solutions

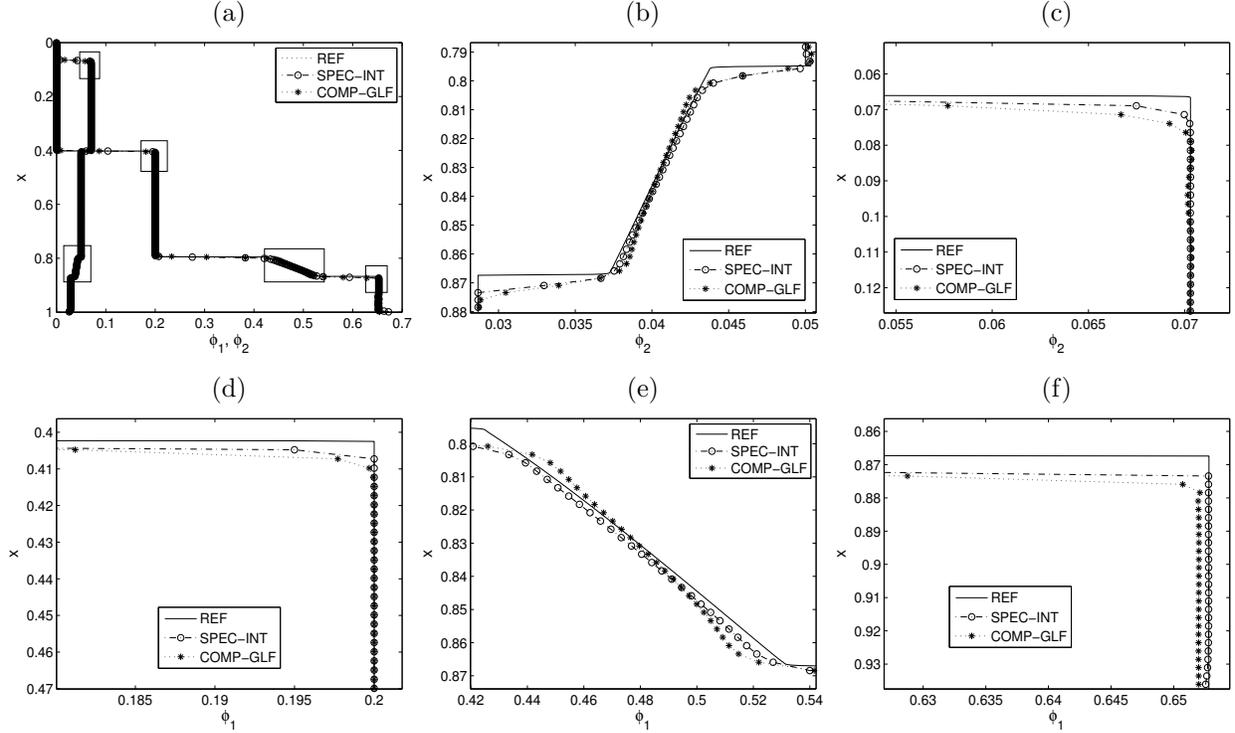


Figure 4.3: Example 4: numerical solution at $t = 50$ s with $M = 400$ (a) and enlarged views (b–f). The reference solution is computed by SPEC-INT with $M = 6400$.

are shown in Figures 4.1 and 4.2 for the simulated times $t = 50$ s and $t = 300$ s, respectively. In Table 4.1 we show approximate L^1 errors for both schemes at two selected times. These approximate errors are computed as follows. Let us denote by $\phi_i^M(\cdot, t)$ and $\phi_i^{\text{ref}}(\cdot, t)$ denote the numerical solution for the i -th component at time t calculated for the discretization $M \in \{100, 200, 400, 800, 1600\}$ and the reference discretization M_{ref} , respectively ($M_{\text{ref}} = 6400$ and $M_{\text{ref}} = 25600$ for the SPEC-INT and COMP-GLF schemes, respectively). Assume that $\phi_i^M(x, t) = \phi_{j,i}^M(t) = \text{const.}$ for $x \in [(j - 1/2)\Delta x, (j + 1/2)\Delta x)$; assume, moreover, that $\phi_i^{\text{ref}}(\cdot, t)$ is piecewise constant on the mesh with meshwidth $1/M_{\text{ref}}$. For a given time t and $r := M_{\text{ref}}/M \in \mathbb{N}$ we then calculate the approximate L^1 error in species i by

$$e_i = e_i(t) = \|\phi_i^{\text{ref}}(\cdot, t) - \phi_i^M(\cdot, t)\|_1 = \frac{1}{M_{\text{ref}}} \sum_{j=0}^{M_{\text{ref}}-1} |\phi_{j,i}^{\text{ref}}(t) - \phi_{[j/r],i}^M(t)|, \quad i = 1, \dots, N.$$

If we define $\phi_j^M(t) := \phi_{j,1}^M(t) + \dots + \phi_{j,N}^M(t)$ (and analogously, $\phi_j^{\text{ref}}(t)$), then the total approx-

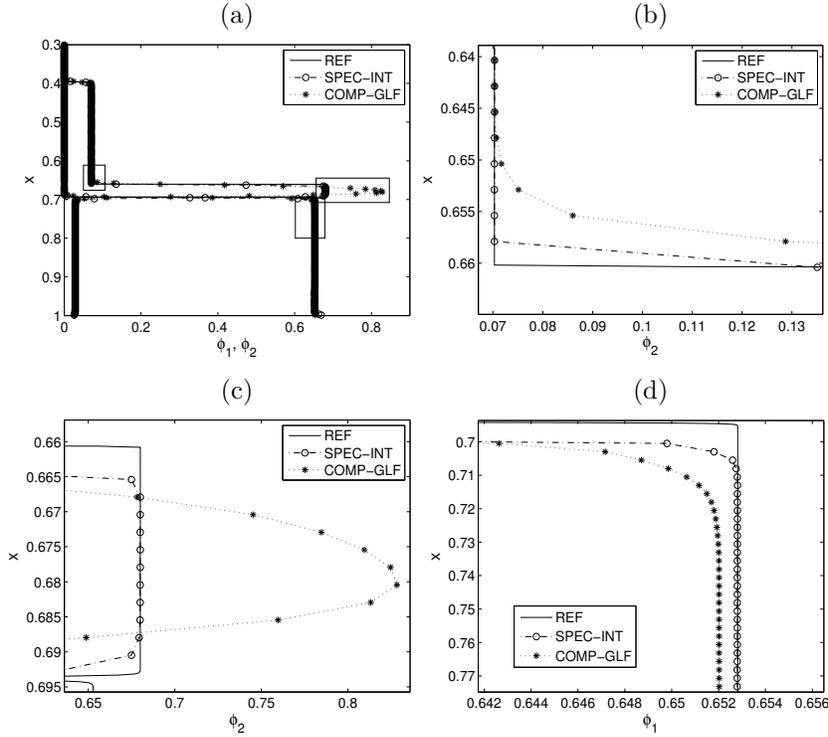


Figure 4.4: Example 4: numerical solution at $t = 300$ s with $M = 400$ (a) and enlarged views (b–f). The reference solution is computed by SPEC-INT with $M = 6400$.

imate L^1 error at that time is given by

$$e_{\text{tot}} = e_{\text{tot}}(t) = \frac{1}{M_{\text{ref}}} \sum_{j=0}^{M_{\text{ref}}-1} |\phi_j^{\text{ref}}(t) - \phi_{[j/r]}^M(t)|.$$

Note that $e_{\text{tot}}(t) \leq e_1(t) + \dots + e_N(t)$.

Table 4.1 shows that SPEC-INT produces smaller values of the error than COMP-GLF, with respect to its 'converged' solution. The difference is significant in the case of species 2 at $t = 50$ s. In the plot of the complete solutions in Figure 4.3 (a), no difference between both solutions becomes apparent, so we present enlarged views of portions of the numerical simulation (Figures 4.3 (b–f)) in which the greater accuracy of the solutions generated by SPEC-INT is appreciable. Plot (a) of Figure 4.4, which corresponds to $t = 300$ s, shows the difference of behaviour of both solutions even without the necessity to enlarge the view; nevertheless we present in Figure 4.4 (b–d) enlarged views to make local differences clearly visible. In Table 4.1 the reference solution is computed by SPEC-INT with $M_{\text{ref}} = 6400$. To exclude that our conclusion of superiority of SPEC-INT is based on a bias due to the

| M | e_1 | cr | e_2 | cr | e_{tot} | cr | e_1 | cr | e_2 | cr | e_{tot} | cr |
|----------------------|--------|-------|-------|-------|------------------|-----------------------|--------|-------|--------|-------|------------------|-------|
| SPEC-INT, $t = 50$ s | | | | | | SPEC-INT, $t = 300$ s | | | | | | |
| 100 | 125.15 | - | 22.40 | - | 131.11 | - | 115.13 | - | 195.99 | - | 110.08 | - |
| 200 | 64.58 | 0.954 | 10.77 | 1.056 | 66.36 | 0.982 | 66.21 | 0.798 | 102.20 | 0.939 | 58.07 | 0.923 |
| 400 | 33.81 | 0.934 | 5.76 | 0.902 | 34.94 | 0.926 | 33.87 | 0.967 | 48.62 | 1.072 | 33.25 | 0.804 |
| 800 | 20.63 | 0.713 | 3.03 | 0.927 | 20.74 | 0.752 | 21.04 | 0.687 | 25.26 | 0.945 | 21.64 | 0.620 |
| 1600 | 13.98 | 0.561 | 1.82 | 0.737 | 13.85 | 0.582 | 15.34 | 0.456 | 13.64 | 0.889 | 15.97 | 0.439 |
| COMP-GLF, $t = 50$ s | | | | | | COMP-GLF, $t = 300$ s | | | | | | |
| 100 | 167.08 | - | 43.21 | - | 188.63 | - | 166.51 | - | 519.44 | - | 416.07 | - |
| 200 | 87.17 | 0.939 | 20.62 | 1.067 | 96.30 | 0.970 | 78.01 | 1.094 | 243.39 | 1.094 | 211.38 | 0.977 |
| 400 | 47.05 | 0.890 | 9.96 | 1.051 | 50.80 | 0.923 | 33.50 | 1.220 | 178.31 | 0.449 | 169.73 | 0.317 |
| 800 | 28.93 | 0.702 | 5.25 | 0.922 | 30.39 | 0.741 | 22.18 | 0.594 | 56.16 | 1.667 | 49.53 | 1.777 |
| 1600 | 21.40 | 0.435 | 3.10 | 0.761 | 21.65 | 0.489 | 12.30 | 0.851 | 43.48 | 0.369 | 36.55 | 0.439 |

Table 4.2: Example 4: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr). The reference solution is computed by COMP-GLF with $M = 25600$.

choice of this scheme for the reference solution, we present a second table of errors for this example, Table 4.2, in which the numerical solutions for $M = 100, \dots, 1600$ are the same as in Table 4.1, but we utilize a reference solution obtained by COMP-GLF with $M_{\text{ref}} = 25600$. As a general observation, throughout a rather extense testing process, we may say that a numerical solution obtained by COMP-GLF agrees in quality and resolution power with the solution obtained by SPEC-INT if the meshwidth for COMP-GLF is roughly a fourth of the one used for SPEC-INT.

Of course, for a given value of M the COMP-GLF scheme is faster than the SPEC-INT scheme, since COMP-GLF does not require the complete spectral information, which avoids many computations. Nevertheless, if we seek a fixed level of resolution in the numerical simulation, then SPEC-INT turns out to be computationally more efficient. For instance, in Example 4 the CPU time is 21.01 s and 87.15 s for providing the solutions at the respective simulated times $t = 50$ s and $t = 300$ s, respectively, with SPEC-INT and $M = 400$, while to obtain a numerical solution of comparable quality (smallness of errors) by COMP-GLF we need to use $M = 1600$ points, and the corresponding CPU times are 29.15 s for $t = 50$ s and 160.80 s for $t = 300$ s.

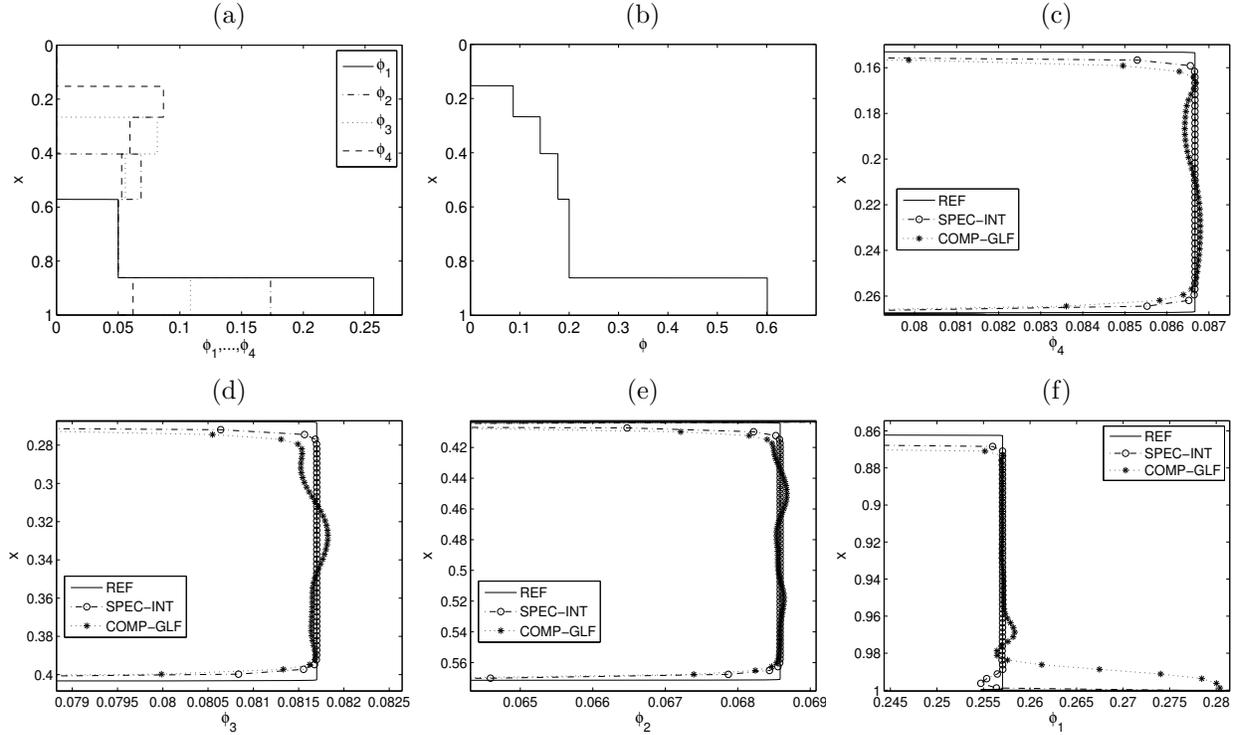


Figure 4.5: Example 5: reference solution for ϕ_1, \dots, ϕ_4 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c-f), at $t = 50$ s.

| M | SPEC-INT, $t = 50$ s | | SPEC-INT, $t = 300$ s | | COMP-GLF, $t = 50$ s | | COMP-GLF, $t = 300$ s | |
|------|----------------------|-------|-----------------------|-------|----------------------|-------|-----------------------|-------|
| | e_{tot} | cr | e_{tot} | cr | e_{tot} | cr | e_{tot} | cr |
| 100 | 103.45 | - | 158.14 | - | 200.42 | - | 212.03 | - |
| 200 | 53.68 | 0.946 | 72.64 | 1.122 | 104.16 | 0.944 | 122.90 | 0.787 |
| 400 | 23.81 | 1.173 | 33.95 | 1.097 | 50.07 | 1.057 | 79.54 | 0.628 |
| 800 | 11.95 | 0.995 | 12.08 | 1.491 | 25.02 | 1.001 | 40.90 | 0.959 |
| 1600 | 5.29 | 1.174 | 7.08 | 0.771 | 12.13 | 1.044 | 23.54 | 0.797 |

Table 4.3: Example 5: approximate L^1 -errors ($\times 10^{-5}$) and convergence rates (cr).

4.4.2 Example 5 (MLB model, $N = 4$)

We consider $d_1 = 1$, $d_2 = 0.8$, $d_3 = 0.6$ and $d_4 = 0.4$, $\phi_{\max} = 0.6$, and $\phi_i^0 = 0.05$ for $i = 1, \dots, 4$. The other parameters are the same as in Example 4. This example goes back to Greenspan and Ungarish [35], and was solved numerically in [13] with the slightly different

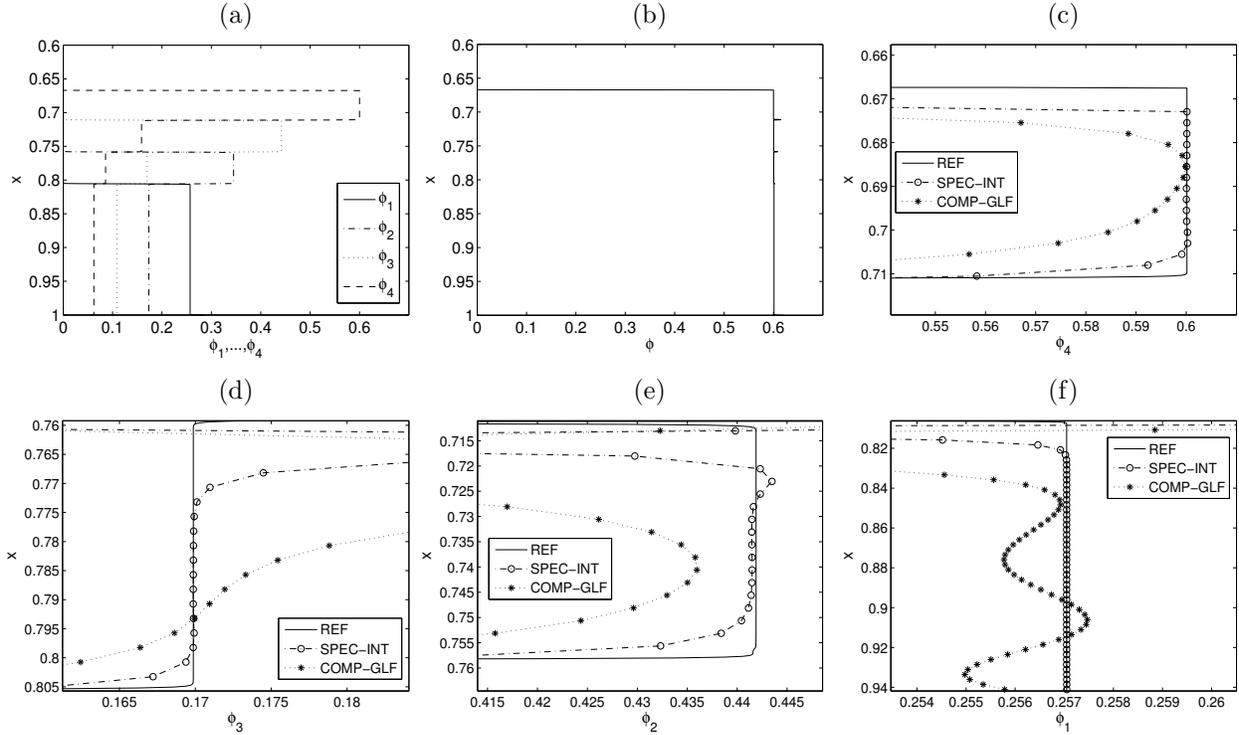


Figure 4.6: Example 5: reference solution for ϕ_1, \dots, ϕ_4 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c–f), at $t = 300$ s.

hindered settling factor $V(\phi) = (1 - (5/3)\phi)^{2.7}$ in [13]. Figures 4.5 (a, b) and 4.6 (a, b) display the reference solution obtained with SPEC-INT and $M_{\text{ref}} = 6400$ for $t = 50$ s and $t = 300$ s respectively, while plots (c–f) of both figures are enlarged views of the corresponding numerical solutions obtained with SPEC-INT and COMP-GLF with $M = 400$. Both series of plots show that at $M = 400$ the quality of approximation of piecewise constant portions of the solution and the resolution of kinematic shocks by SPEC-INT is superior to that of COMP-GLF. Table 4.3 displays the approximate total L^1 error and convergence rates for this case. For the times considered the average convergence rate using the SPEC-INT method is close to one. On the other hand, as time increases, the errors increase considerably.

We select this case to compare the performance of SPEC-INT with that of SPEC-LLF, the method based on the simpler viscosity coefficient (4.3.11). Both choices approximate the same solution globally (not shown here), and a few enlarged views of relevant parts of the numerical solution shown in Figure 4.7 indicate that the resolution of kinematic shocks by SPEC-LLF is even slightly better than by SPEC-INT. However, we observe spurious oscillations produced by SPEC-LLF in the piecewise constant parts of the solution. These

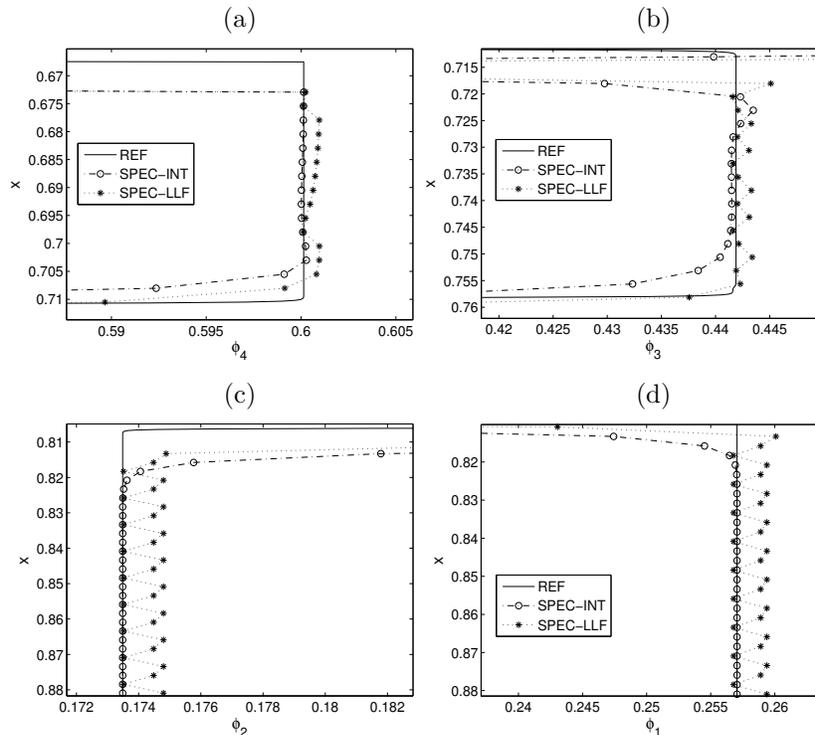


Figure 4.7: Example 5: details of numerical solutions obtained by SPEC-INT and SPEC-LLF with $M = 400$ at $t = 300$ s.

oscillations do not disappear upon mesh refinement, and indicate that the amount of viscosity introduced by (4.3.11) is not appropriate and possibly insufficient.

4.4.3 Example 6 (MLB model, $N = 11$)

This example is based on experimental data from [59], where the settling of a suspension in a column of height $L = 0.935$ m was considered. The initial concentrations ϕ_i^0 , diameters D_i and normalized diameters $d_i = D_i/D_1$ are given in Table 3.1; the maximum total concentration is $\phi_{\max} = 0.641$ [59]. Figures 4.8 (a) and (b) show the concentration profiles of the reference solution, obtained by SPEC-INT with $M_{\text{ref}} = 6400$ at $t = 300$ s. Figures 4.8 (c–f) display enlarged views of portions of the SPEC-INT and COMP-GLF solutions with $M = 400$ at the same time. Again, the superiority of the quality of approximation by SPEC-INT becomes apparent. This observation is also confirmed by the errors displayed in Table 4.4.

| M | SPEC-INT, $t = 50$ s | | SPEC-INT, $t = 300$ s | | COMP-GLF, $t = 50$ s | | COMP-GLF, $t = 300$ s | |
|------|----------------------|-------|-----------------------|-------|----------------------|-------|-----------------------|-------|
| | e_{tot} | cr | e_{tot} | cr | e_{tot} | cr | e_{tot} | cr |
| 100 | 291.87 | - | 351.60 | - | 617.68 | - | 733.96 | - |
| 200 | 135.51 | 1.107 | 182.85 | 0.943 | 304.39 | 1.021 | 393.88 | 0.898 |
| 400 | 66.22 | 1.033 | 96.86 | 0.917 | 164.93 | 0.884 | 212.10 | 0.893 |
| 800 | 36.48 | 0.860 | 44.93 | 1.108 | 89.51 | 0.882 | 112.20 | 0.919 |
| 1600 | 17.74 | 1.040 | 21.07 | 1.093 | 46.61 | 0.941 | 63.38 | 0.824 |

Table 4.4: Example 6: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr).

4.4.4 Example 7 (MLB model, $N = 2$)

We consider the MLB model with $N = 2$ and the same parameters as Example 4, but now start from the initial datum

$$\Phi(x, 0) = \begin{cases} 0.15 & \text{if } x \leq 0.5, \\ 0 & \text{if } x > 0.5, \end{cases} \quad (4.4.1)$$

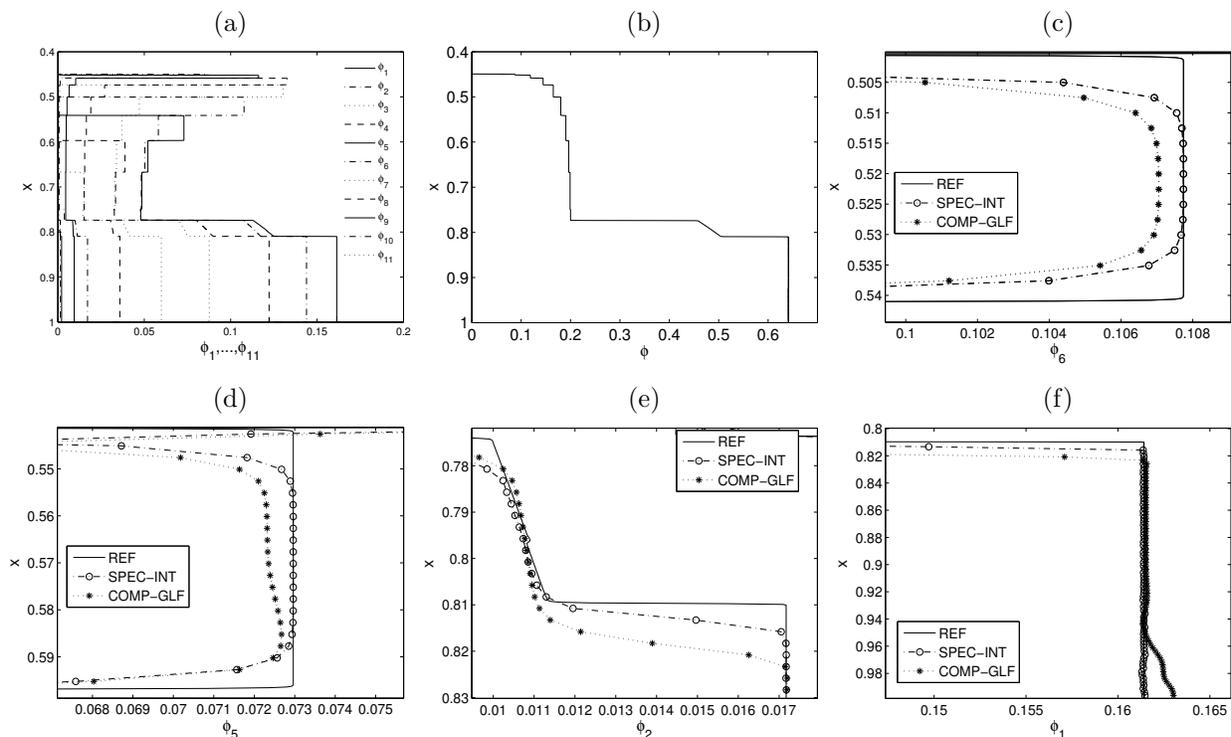


Figure 4.8: Example 6: reference solution for ϕ_1, \dots, ϕ_{11} and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c-f), at $t = 300$ s.

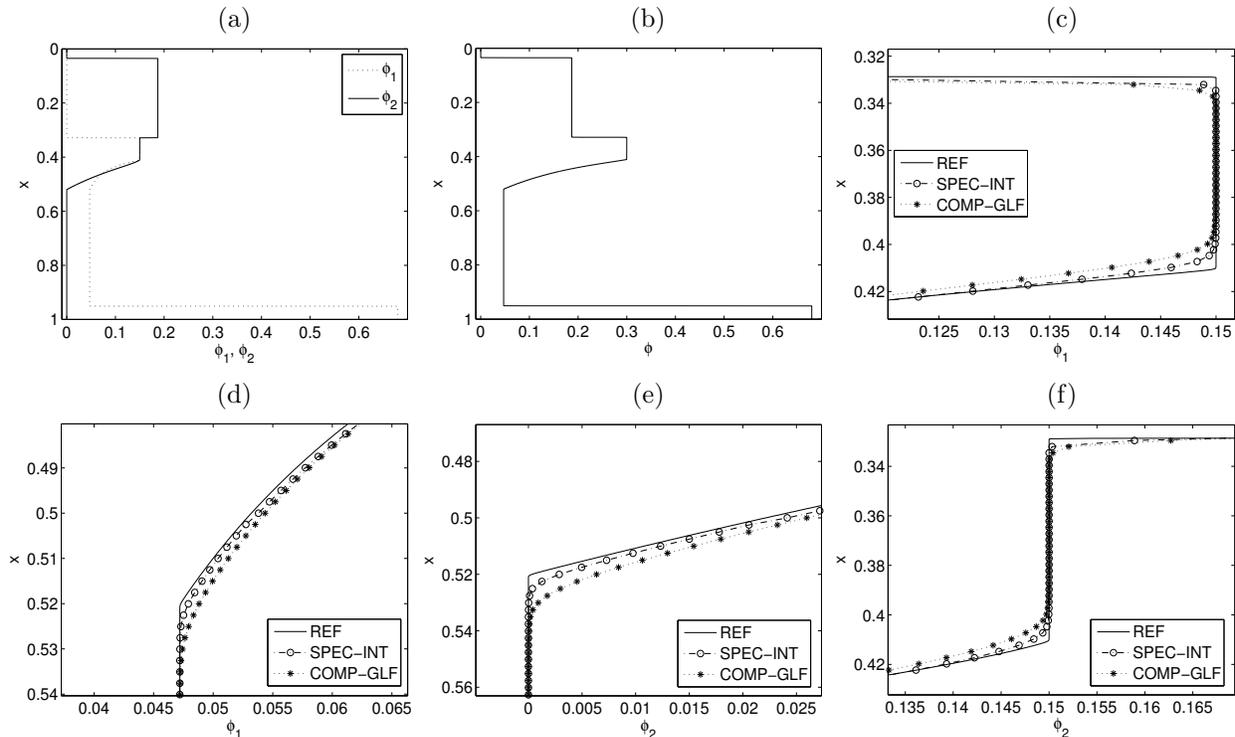


Figure 4.9: Example 7: reference solution for ϕ_1, ϕ_2 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c–f), at $t = 50$ s.

corresponding to a settling column whose upper half is initially filled with a suspension, which is separated from the lower half by a “membrane”, that is removed at $t = 0$. The suspension pouring into the lower half will then gradually dilute, and usually a transient rarefaction wave centered at $x = 0.5$ will form. (The rarefaction wave will, however, soon start to interact with concentration information traveling downwards and upwards from the suspension meniscus and column bottom, respectively.) As was shown in [25], this configuration can be realized experimentally (with some effort), and the expanding concentration gradient reveals properties of the function $V(\phi)$ which at least for $N = 1$ can be used for flux identification. A similar configuration, but on an unbounded domain, was solved for $N = 2, 4, 8$ and 32 in [16].

Figures 4.9 and 4.10, which correspond to the respective simulated times $t = 50$ s and $t = 250$ s, show the reference solution obtained by SPEC-INT with $M_{\text{ref}} = 6400$ and details illustrating the difference in solutions obtained by SPEC-INT and COMP-GLF with $M = 400$. Table 4.5 displays the errors observed for this example.

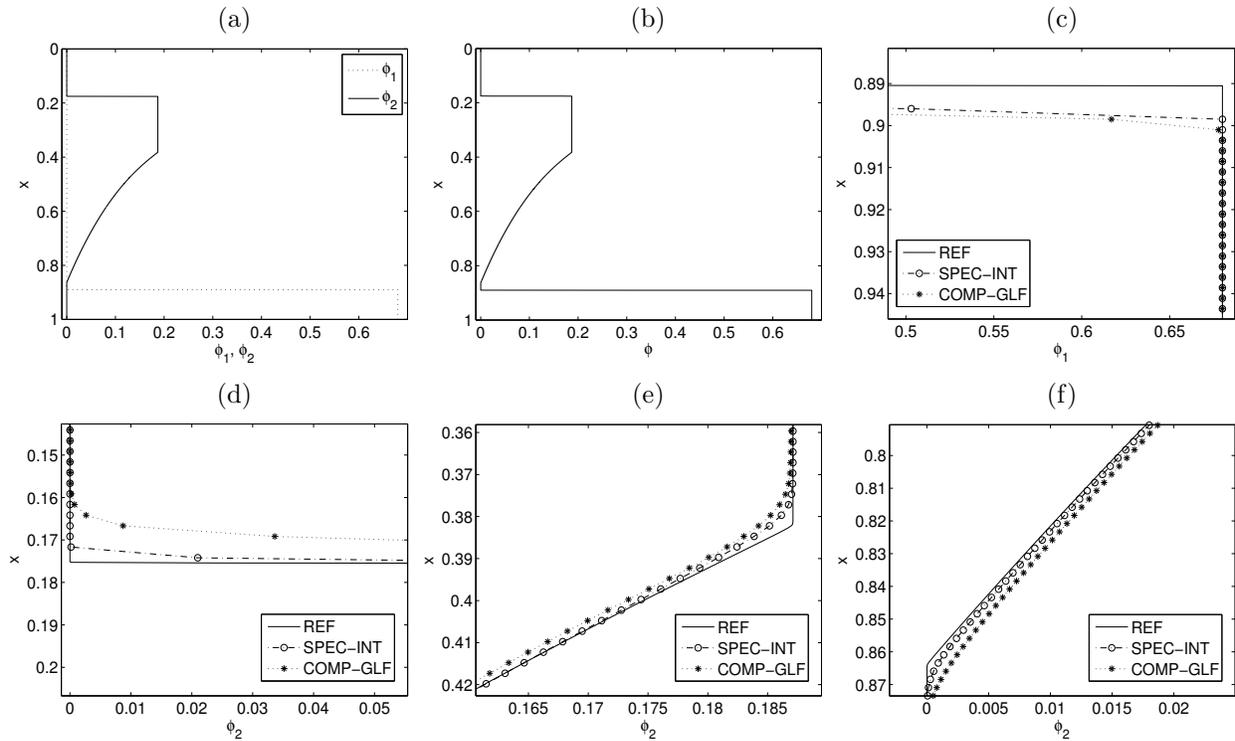


Figure 4.10: Example 7: reference solution for ϕ_1, ϕ_2 and ϕ computed by SPEC-INT with $M_{\text{ref}} = 6400$ (a, b), and details of numerical solutions with $M = 400$ (c-f), at $t = 250$ s.

4.4.5 Example 8: HS model with $N = 2$ and $N = 4$

For these examples we implement the variant of HS model described in Section 4.2. Numerical simulations are shown in Figs. 4.11 and 4.12 for $N = 2$, and Figs. 4.13 and 4.14 for $N = 4$. A noticeable difference with the MLB model (where the flux function is cut abruptly for $\phi \geq \phi_{\text{max}}$) is the profile at the rightmost part of the solution. We claim that the fact of cutting the flux function for values greater than ϕ_{max} could be produce small oscillations in zones where the concentrations must be constant, this effect cannot be observed from the global figure in the whole computational domain showed in Fig. 4.14 (a, b) but can be seen from enlarging views which are showed in subplots (c, d, e, f), compare it with Figs. 4.6 and 4.13.

Table 4.6 displays the errors observed for Example 8 for the case of the variant of the HS model with two species.

| M | e_1 | cr | e_2 | cr | e_{tot} | cr | e_1 | cr | e_2 | cr | e_{tot} | cr |
|-----------------------------|--------|-------|--------|-------|------------------|------------------------------|--------|-------|--------|-------|------------------|-------|
| SPEC-INT, $t = 50\text{ s}$ | | | | | | SPEC-INT, $t = 250\text{ s}$ | | | | | | |
| 100 | 210.30 | - | 83.10 | - | 282.82 | - | 124.22 | - | 70.15 | - | 194.19 | - |
| 200 | 107.67 | 0.966 | 34.57 | 1.265 | 136.26 | 1.054 | 64.25 | 0.951 | 29.39 | 1.255 | 93.64 | 1.052 |
| 400 | 66.79 | 0.689 | 9.42 | 1.875 | 73.25 | 0.895 | 55.27 | 0.217 | 10.08 | 1.544 | 65.35 | 0.519 |
| 800 | 37.37 | 0.838 | 4.53 | 1.057 | 40.74 | 0.846 | 32.97 | 0.745 | 5.55 | 0.860 | 38.53 | 0.762 |
| 1600 | 19.11 | 0.967 | 4.29 | 0.075 | 22.78 | 0.838 | 14.02 | 1.233 | 4.04 | 0.456 | 18.07 | 1.092 |
| COMP-GLF, $t = 50\text{ s}$ | | | | | | COMP-GLF, $t = 250\text{ s}$ | | | | | | |
| 100 | 281.79 | - | 125.25 | - | 392.34 | - | 236.18 | - | 160.41 | - | 394.76 | - |
| 200 | 161.65 | 0.802 | 73.05 | 0.778 | 226.85 | 0.790 | 121.40 | 0.960 | 77.05 | 1.058 | 198.27 | 0.994 |
| 400 | 88.24 | 0.873 | 35.53 | 1.040 | 119.74 | 0.922 | 73.96 | 0.715 | 35.46 | 1.120 | 109.42 | 0.858 |
| 800 | 47.80 | 0.884 | 18.30 | 0.957 | 64.29 | 0.897 | 41.71 | 0.826 | 18.30 | 0.954 | 60.01 | 0.866 |
| 1600 | 24.91 | 0.940 | 10.47 | 0.806 | 34.50 | 0.898 | 20.56 | 1.020 | 9.84 | 0.894 | 30.41 | 0.980 |

Table 4.5: Example 7: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr) for Riemann problem. The reference solution is computed by SPEC-INT with $M_{\text{ref}} = 6400$.

4.5 Conclusions of Chapter 4

In this chapter we have shown that the implementation of efficient WENO schemes for polydisperse sedimentation models can be accomplished by using the recent hyperbolicity analysis carried out in [14]. In addition, we have been able to characterize the viscosity coefficients to be used in Global-Lax-Friedrichs flux-splitting procedures, as well as in the Local-Lax-Friedrichs flux-splitting procedure. The particular algebraic structure of the velocities of the MLB and HS models permits to exactly determine the extremal set $\mathcal{E}_k(\Phi_j, \Phi_{j+1})$ defined in (4.3.13), and hence the specific viscosity coefficient to be used at each cell interface.

We have constructed component-wise and characteristic-based WENO5 schemes for two polydisperse sedimentation models, and have compared their performance. As in the case of the MCLWR kinematic traffic models, the characteristic-based schemes, which use the full spectral decomposition of the Jacobian matrix at each cell-interface, are more robust and lead to numerical solutions which are essentially oscillation free. We remark that this situation is absolutely similar to what is observed in the better known case of the Euler equations for gas dynamics simulations, where the superiority of characteristic-based schemes is a well

| M | e_1 | cr | e_2 | cr | e_{tot} | cr | e_1 | cr | e_2 | cr | e_{tot} | cr |
|----------------------|--------|-------|-------|-------|------------------|-----------------------|--------|-------|--------|-------|------------------|-------|
| SPEC-INT, $t = 50$ s | | | | | | SPEC-INT, $t = 250$ s | | | | | | |
| 100 | 135.94 | - | 18.45 | - | 144.90 | - | 113.12 | - | 105.16 | - | 108.60 | - |
| 200 | 70.57 | 0.946 | 7.86 | 1.231 | 73.50 | 0.979 | 59.41 | 0.929 | 53.30 | 0.980 | 56.05 | 0.954 |
| 400 | 34.34 | 1.039 | 3.76 | 1.063 | 35.80 | 1.038 | 29.82 | 0.994 | 29.01 | 0.877 | 27.37 | 1.034 |
| 800 | 16.49 | 1.058 | 1.78 | 1.073 | 17.21 | 1.057 | 11.26 | 1.404 | 12.05 | 1.267 | 14.29 | 0.937 |
| 1600 | 6.71 | 1.297 | 0.76 | 1.226 | 7.05 | 1.286 | 4.86 | 1.212 | 5.74 | 1.068 | 6.77 | 1.077 |
| COMP-GLF, $t = 50$ s | | | | | | COMP-GLF, $t = 250$ s | | | | | | |
| 100 | 138.48 | - | 26.99 | - | 160.72 | - | 171.39 | - | 245.80 | - | 194.93 | - |
| 200 | 70.17 | 0.981 | 11.93 | 1.177 | 79.71 | 1.012 | 86.78 | 0.982 | 138.36 | 0.829 | 114.99 | 0.761 |
| 400 | 35.96 | 0.964 | 5.31 | 1.166 | 40.08 | 0.992 | 45.29 | 0.938 | 70.90 | 0.964 | 56.92 | 1.014 |
| 800 | 17.02 | 1.079 | 2.52 | 1.072 | 19.06 | 1.072 | 22.87 | 0.986 | 39.53 | 0.843 | 32.69 | 0.800 |
| 1600 | 8.39 | 1.020 | 1.20 | 1.073 | 9.35 | 1.028 | 11.40 | 1.004 | 22.39 | 0.820 | 19.68 | 0.732 |

Table 4.6: Example 8: approximate L^1 errors ($\times 10^{-5}$) and convergence rates (cr) for HS model. The reference solution is computed by SPEC-INT with $M_{\text{ref}} = 6400$.

known fact. For gas dynamics, the spectral decomposition of the Jacobian matrix is given in closed form, hence the use of a characteristic-based scheme poses no special difficulties. For polydisperse models, the spectral decomposition can only be computed numerically. In addition, the characteristic fields are neither genuinely nonlinear nor linearly degenerate, hence the determination of the viscosity coefficients in flux-vector splitting schemes becomes a non-trivial task.

According to the numerical tests shown in this chapter, our proposed characteristic-based scheme (SPEC-INT) is very robust, although it is certainly very costly in terms of computational resources, since it involves an intensive usage of the characteristic information. The interlacing property allows other simplifications to be implemented. For example, in Section 4.3.5, the discrimination between Cases 1 and 2 corresponding to the use of either (4.3.19) or (4.3.20) is made in dependence of the sign of the product of eigenvalues $\lambda_j^k \cdot \lambda_{j+1}^k$. The interlacing property (3.2.9) or (3.2.10) can be used to compute this sign in terms of velocities rather than eigenvalues.

Nevertheless, we have shown that the SPEC-INT gives a good resolution on the numerical approximation with a relative small number of mesh points, hence it is competitive with

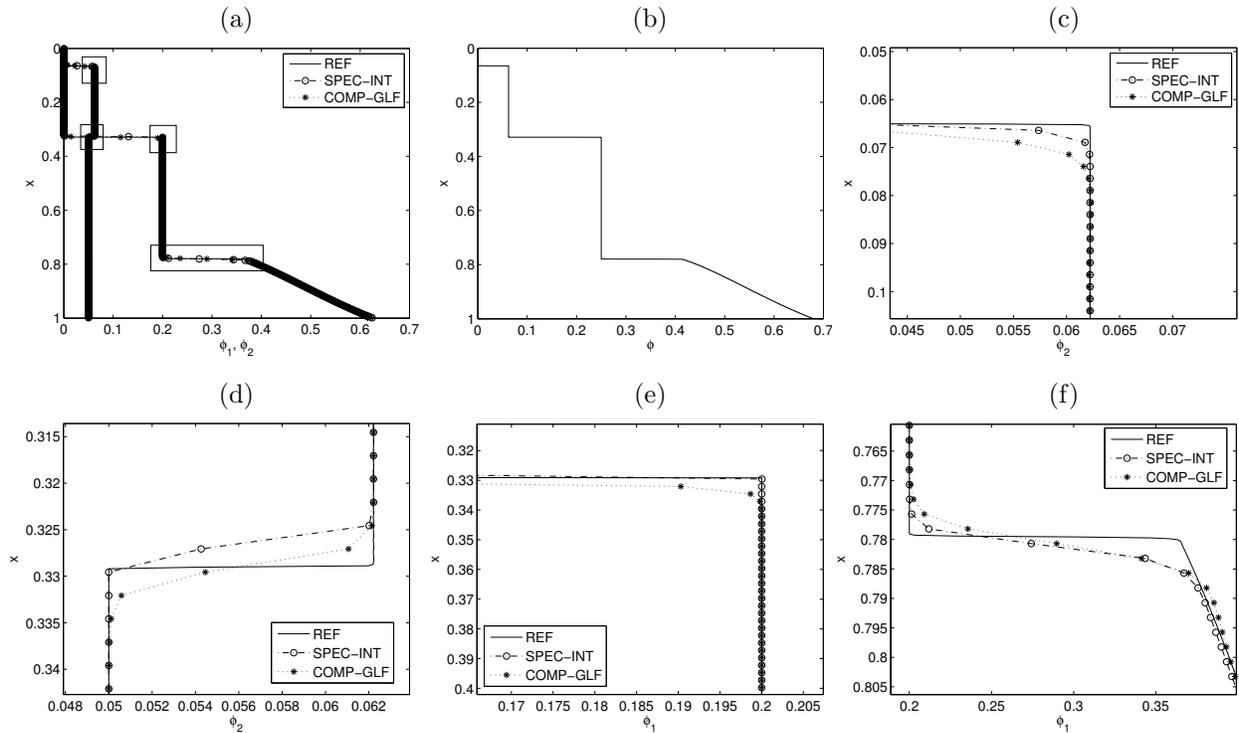


Figure 4.11: Example 8: numerical solution for ϕ_1 , ϕ_2 with $M = 400$ (a) and ϕ (b) at $t = 50$ s and enlarged views (c–f) of zones indicated by rectangles in plot (a). The reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$.

respect to the simpler component-wise schemes. We expect the SPEC-INT scheme to be even more competitive than cheaper component-wise schemes, such as COMP-GLF, in an Adaptive Mesh Refinement (AMR) framework, since its non-oscillatory properties will help to avoid unnecessary refinement in regions of constant concentration.

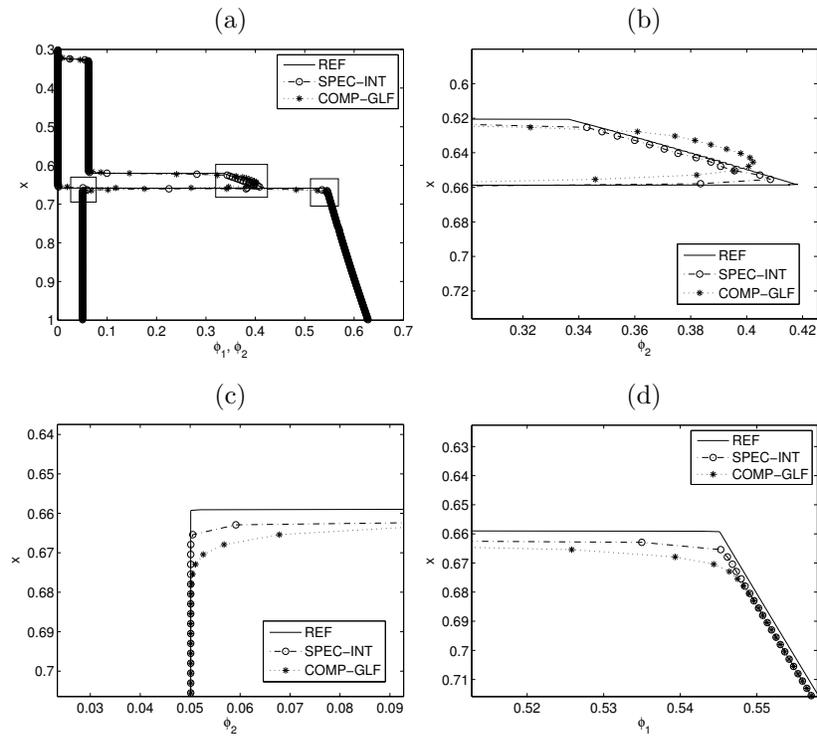


Figure 4.12: Example 8 (HS variant with $N = 2$): numerical solution for ϕ_1, ϕ_2 with $M = 400$ (a) at $t = 250$ s and enlarged views (b–d). The reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$.

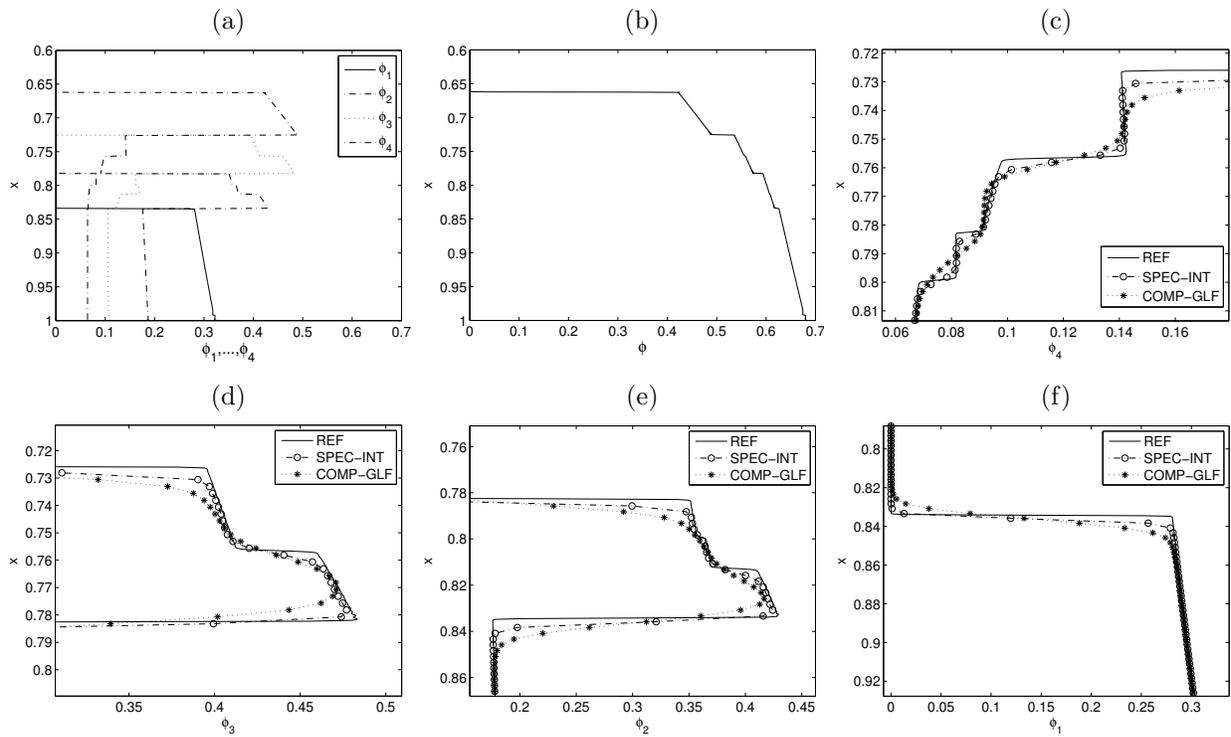


Figure 4.13: Example 8 (HS variant with $N = 4$): numerical solution for ϕ_1, \dots, ϕ_4 with $M = 400$ (a) and ϕ (b) at $t = 250$ s and enlarged views (c-f), where the reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$.

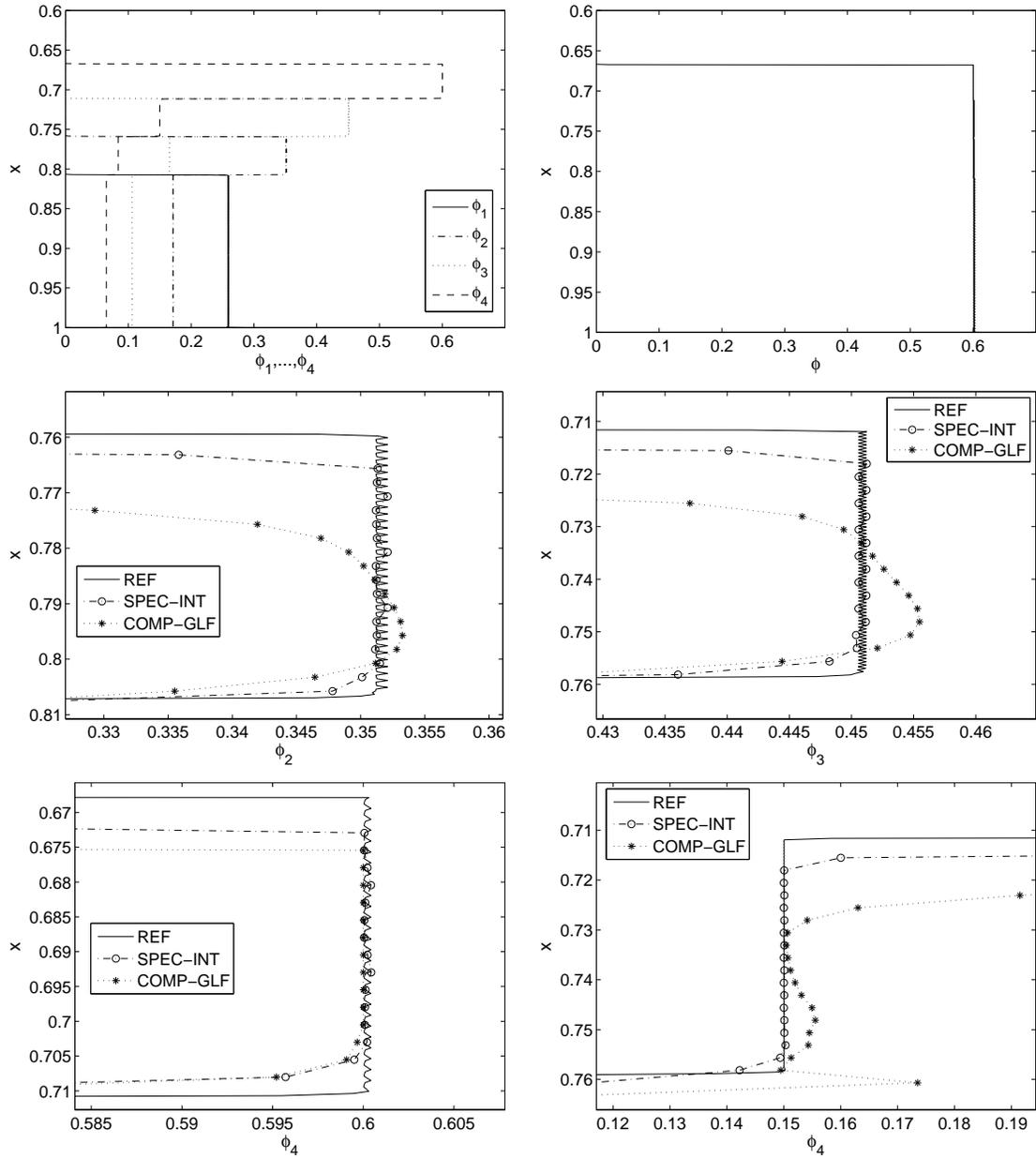


Figure 4.14: Example 8 (HS with $N = 4$): numerical solution for ϕ_1, \dots, ϕ_4 with $M = 400$ (a) and ϕ (b) at $t = 250$ s and enlarged views (c–f), where the reference solution is computed using SPEC-INT with $M_{\text{ref}} = 6400$.

Chapter 5

Additional results of Hyperbolicity

5.1 Introduction

In the first part of this chapter we perform the analysis of hyperbolicity for the Davis-Gecol model under the assumption that $\beta_3 = 0$. For this model, it is still possible to obtain conditions that guarantee strictly hyperbolicity on $\mathcal{D}_{\phi_{\max}}$ for d_N (the smallest diameter) close to one. A result in this sense was got for the particular case of $N = 2$ in the reference [17], by using directly the characteristic polynomial, in that paper was showed that DG model with $N = 2$ is stable if the solid particles differ only moderately in size, more exactly, if the size ratio is not larger than about 5.5. They also pointed out that the size of region of instability increases when the ratio D_1/D_2 increases too. Unlike previous work, the remarkable point is that the result obtained here does not depend on the number of species.

The remaining of this chapter is devoted to the analysis of hyperbolicity considering $\beta_3 < 0$ for the BW and HS model. Most of the computations imply algebraic procedures similar to the calculations performed in Chapter 3.

5.2 The Davis and Gecol (DG) model

We first recall (see equation (3.3.14)) that the settling velocity for DG model [23] is given by

$$v_i(\Phi) = d_i^2(1 + \mathbf{s}_i^T \Phi - S_{ii}\phi)(1 - \phi)^{-S_{ii}}, \quad (5.2.1)$$

where $\mathbf{s}_i^T \Phi$ and $S_{ii} = (\beta_0 + \beta_1 + \beta_2)$, for $i = 1, \dots, N$ were considered in Subsection 3.3.2. Notice that DG and BW models agree in the dilute limit.

Similarly to the analysis of the BW and HS model, we will conduct the analysis of the DG model under the a priori assumption that $\beta_3 = 0$. For $i = 1, \dots, N$ we consider $\eta_i = 1 + \mathbf{s}_i^T \Phi + n\phi$, with $n = -S_{ii}$. For this model the auxiliary scalar variables p_k , $k = 1, 2, 3$, are given by equation (3.4.8), on the other hand, the entries of matrix \mathbf{A} and \mathbf{B} are given by

$$\begin{aligned} \alpha_i^k &= d_i^{k-1}, \\ \beta_i^k &= \begin{cases} d_i^2 \phi_i (1 - \phi)^{n-1} ((1 - \phi)(\beta_0 + n) - n\eta_i) & \text{for } k = 1, \\ d_i^{3-k} \phi_i (1 - \phi)^n \beta_{k-1}, & \text{for } k = 2, 3 \end{cases} \end{aligned} \quad (5.2.2)$$

Now, taking into account that

$$\begin{aligned} v_j - v_i &= (1 - \phi)^n (d_j^2 \eta_j - d_i^2 \eta_i) \\ &= (1 - \phi)^n (d_j^2 - d_i^2) \left(1 - (\beta_1 + \beta_2)\phi + \frac{\beta_1 \mathbf{d}_1^T \Phi}{d_i + d_j} \right), \end{aligned}$$

we obtain the following coefficients for the secular equation

$$\begin{aligned} \gamma_i &= \alpha_i^1 \beta_i^1 + \alpha_i^2 \beta_i^2 + \alpha_i^3 \beta_i^3 n \\ &+ \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\alpha_{ij}^{12} \beta_{ij}^{12} + \alpha_{ij}^{13} \beta_{ij}^{13} + \alpha_{ij}^{23} \beta_{ij}^{23}}{v_j - v_i} + \sum_{\substack{j,k=1 \\ i \neq j < k \neq i}}^N \frac{\alpha_{ijk}^{123} \beta_{ijk}^{123}}{(v_k - v_i)(v_j - v_i)} \\ &= \phi_i (1 - \phi)^{n-1} (\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i}), \quad i = 1, \dots, N, \end{aligned} \quad (5.2.3)$$

where

$$\begin{aligned} \mathcal{S}_{1,i} &= -n d_i^2 \eta_i = d_i^2 (\beta_0 + \beta_1 + \beta_2) \left(1 - (\beta_1 + \beta_2)\phi + \frac{\beta_1 \mathbf{d}_1^T \Phi}{d_i} + \frac{\beta_2 \mathbf{d}_2^T \Phi}{d_i^2} \right), \\ \mathcal{S}_{2,i} &= \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\phi_j (d_i - d_j)}{(d_i + d_j) \left[1 - (\beta_1 + \beta_2)\phi + \frac{\beta_1 \mathbf{d}_1^T \Phi}{d_i + d_j} \right]} \mathcal{S}_{2,i,j}, \\ \mathcal{S}_{2,i,j} &:= (\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) [(1 - \phi)(\beta_0 + n) - n(1 - (\beta_1 + \beta_2)\phi)] \\ &+ \beta_1 \beta_2 [n(\mathbf{d}_2^T \Phi - (d_i + d_j) \mathbf{d}_1^T \Phi) + (1 - \phi) d_i d_j], \\ \mathcal{S}_{3,i} &= \sum_{\substack{j,k=1 \\ i \neq j < k \neq i}}^N \frac{\phi_j \phi_k \pi_{ijk}^2 \beta_1 \beta_2 [\beta_0 (1 - \phi) - n\phi (1 - (\beta_1 + \beta_2))]}{(d_k^2 - d_i^2)(d_j^2 - d_i^2) \left[1 - (\beta_1 + \beta_2)\phi + \frac{\beta_1 \mathbf{d}_1^T \Phi}{d_i + d_k} \right] \left[1 - (\beta_1 + \beta_2)\phi + \frac{\beta_1 \mathbf{d}_1^T \Phi}{d_i + d_j} \right]}. \end{aligned} \quad (5.2.4)$$

Unfortunately, this model does not allow for term cancellations as for the BW and HS models. However, it is still possible to deduce that the model is strictly hyperbolic on $\mathcal{D}_{\phi_{\max}}$

for realistically large values of ϕ_{\max} provided that d_N is sufficiently close to one. Our analysis leads here to a narrow size distribution only. The salient point is, however, that our bounds for $d_N < 1$ are independent of N . Here, we can prove the following result.

Lemma 5.2.1 *Consider the DG model, where the parameters d_N , ϕ_{\max} and $\boldsymbol{\beta} = (\beta_0, \beta_1, \beta_2)^T$ are assumed to satisfy*

$$d_N > 1/2, \quad (5.2.5)$$

$$1 + [\beta_1(d_N^{-1} - 1) + \beta_2(d_N^{-2} - 1)]\phi_{\max} > 0. \quad (5.2.6)$$

Then the model is strictly hyperbolic for all $\phi \in \mathcal{D}_{\phi_{\max}}$ provided that

$$\mathcal{S}(\phi, d_N; \boldsymbol{\beta}) < 0 \quad \text{for all } \phi \in (0, \phi_{\max}], \quad (5.2.7)$$

where we define

$$\mathcal{S}(\phi, d_N; \boldsymbol{\beta}) := -n \left(1 + [\beta_1(d_N^{-1} - 1) + \beta_2(d_N^{-2} - 1)]\phi \right) + \frac{1 - d_N}{d_N} \cdot \frac{C_1\phi}{2} + \frac{(1 - d_N)^4}{d_N^4} \cdot \frac{C_2\phi^2}{4}. \quad (5.2.8)$$

The constants are given by

$$\begin{aligned} C_1 &:= (\beta_1 + 4\beta_2)(\beta_0 + (\beta_1 + \beta_2)\phi_{\max}) + n\phi_{\max}[\beta_1(\beta_1 + \beta_2) + 4\beta_2^2] \\ &\quad + \beta_1\beta_2(\phi_{\max}(n(4 + d_N^{-2}) - 1) + 1), \\ C_2 &:= \beta_1\beta_2(\beta_0 - n\phi_{\max}(1 - (\beta_1 + \beta_2))). \end{aligned}$$

Proof. We first note that for all $i = 1, \dots, N$ and all $\Phi \in \mathcal{D}_{\phi_{\max}}$ the following inequality holds:

$$\eta_i = 1 + \mathbf{s}_i^T \Phi + n\phi = 1 + \sum_{j=1}^N \sum_{k=1}^2 \beta_k \left(\frac{d_j^k}{d_i^k} - 1 \right) \phi_j \geq 1 + [\beta_1(d_N^{-1} - 1) + \beta_2(d_N^{-2} - 1)]\phi_{\max}, \quad (5.2.9)$$

so (5.2.6) ensures that always $\eta_i > 0$, and therefore $\mathcal{S}_{1,i} < 0$. Observe that (5.2.6) holds if d_N is chosen sufficiently close to one, or ϕ_{\max} is sufficiently small. Next, a straightforward calculation, and utilizing that

$$(1 - \phi)(\beta_0 + n) - n(1 - (\beta_1 + \beta_2)\phi) = \beta_0(1 - \phi) - n\phi(1 - (\beta_1 + \beta_2)),$$

yields $\mathcal{S}_{2,i,j} = d_i^2 \tilde{\mathcal{S}}_{2,i,j} = d_i^2 \tilde{\mathcal{S}}_{2,i,j}(\boldsymbol{\beta}, \mathbf{d}, \Phi)$, where

$$\begin{aligned} \tilde{\mathcal{S}}_{2,i,j} = & \left(\beta_1 \frac{d_j}{d_i} + \beta_2 \left(1 + \frac{d_j}{d_i} \right)^2 \right) (\beta_0(1 - \phi) - n\phi) + n\phi \left[\beta_1(\beta_1 + \beta_2) \frac{d_j}{d_i} + \beta_2^2 \left(1 + \frac{d_j}{d_i} \right)^2 \right] \\ & + \beta_1\beta_2 \left[n \left(1 + \frac{d_j}{d_i} \right) \left\{ \phi \left(1 + \frac{d_j}{d_i} \right) - \frac{\mathbf{d}_1^T \Phi}{d_i} \right\} + n \frac{\mathbf{d}_2^T \Phi}{d_i^2} + (1 - \phi) \frac{d_j}{d_i} \right]. \end{aligned} \quad (5.2.10)$$

Note that a sufficient condition for $\tilde{\mathcal{S}}_{2,i,j}$ to be positive for all Φ , and without further restrictions on β_0 , β_1 and β_2 , is that the expression in the curled bracket is positive, i.e.,

$$\phi \left(1 + \frac{d_j}{d_i} \right) - \frac{\mathbf{d}_1^T \Phi}{d_i} = \sum_{l=1}^N \left(1 + \frac{d_j - d_l}{d_i} \right) \phi_l > 0. \quad (5.2.11)$$

A sufficient condition for (5.2.11) to hold for all vectors Φ is that the coefficients of ϕ_l for all $i, j, l \in \{1, \dots, N\}$ are positive. This occurs if and only if $1 - (1 - d_N)/d_N > 0$, or equivalently, (5.2.5) is satisfied.

Assume now that $\mathcal{S}_{2,i,j} > 0$, and note that for $d_N > 1/2$, we have that

$$1 - (\beta_1 + \beta_2)\phi + \frac{\beta_1 \mathbf{d}_1^T \Phi}{d_i + d_j} > 1 + \left[\beta_1 \left(\frac{1}{2d_N} - 1 \right) - \beta_2 \right] \phi > 1 \quad \text{for } \phi \in [0, \phi_{\max}].$$

Then we need to estimate $\mathcal{S}_{2,i}^+$, which (as in the BW and HS models) is the partial sum of all positive summands of $\mathcal{S}_{2,i}$, that is,

$$\mathcal{S}_{2,i}^+ = d_i^2 \sum_{j=i+1}^N \frac{\phi_j(d_i - d_j) \tilde{\mathcal{S}}_{2,i,j}}{(d_i + d_j) \left(1 - (\beta_1 + \beta_2)\phi + \frac{\beta_1 \mathbf{d}_1^T \Phi}{d_i + d_j} \right)}.$$

In light of our previous assumptions and considerations, we obtain

$$\mathcal{S}_{2,i}^+ \leq \frac{d_i^2(1 - d_N)\phi}{2d_N} \max_{i < j \leq N} \tilde{\mathcal{S}}_{2,i,j}.$$

However, from (5.2.10) and (5.2.5) we get that

$$\begin{aligned} \max_{i < j \leq N} \tilde{\mathcal{S}}_{2,i,j} & \leq (\beta_1 + 4\beta_2)(\beta_0 + (\beta_1 + \beta_2)\phi) + n\phi [\beta_1(\beta_1 + \beta_2) + 4\beta_2^2] \\ & \quad + \beta_1\beta_2 [n\phi(4 + d_N^{-2}) + (1 - \phi)] \\ & \leq C_1. \end{aligned}$$

Finally, similar considerations for $\mathcal{S}_{3,i}$ and noting that

$$\frac{\pi_{ijk}^2}{(d_j^2 - d_i^2)(d_k^2 - d_i^2)} = d_i^2 \frac{(d_j - d_i)(d_k - d_i)(d_k - d_j)^2}{d_i^2(d_i + d_j)(d_i + d_k)} \leq d_i^2 \frac{(1 - d_N)^4}{4d_N^4}$$

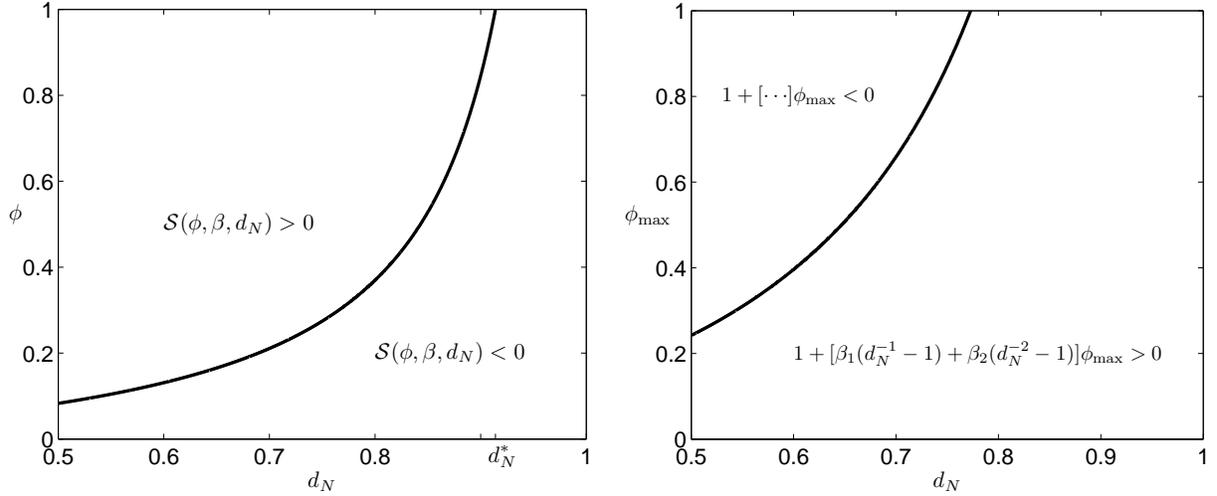


Figure 5.1: Region of hyperbolicity ($\mathcal{S}(\phi, \beta, d_N) < 0$) for the DG model (left) and the diagram with the region where the condition (5.2.6) holds (right). The components of β used in the calculations are given by (3.3.13).

lead to

$$\mathcal{S}_{3,i}^+ \leq -\frac{d_i^2(1-d_N)^4\phi_{\max}^2 C_2}{4d_N^4}.$$

Summarizing, we see that $\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i} \leq d_i^2 \mathcal{S}(\phi, d_N; \beta)$, where $\mathcal{S}(\phi, d_N; \beta)$ is defined in (5.2.8). Thus, we conclude that for given parameters d_N , ϕ_{\max} and β we have $\gamma_i < 0$ for all $i = 1, \dots, N$ on $\mathcal{D}_{\phi_{\max}}$, and therefore hyperbolicity, provided that $\mathcal{S}(\phi, d_N; \beta) < 0$ for all $\phi \in (0, \phi_{\max}]$. \square

As an example we consider the DG model with the coefficients (3.3.13) (with $\beta_3 = 0$). Figure 5.1 (left) shows the hyperbolicity region defined by condition (5.2.7) for this model. We limit the discussion here to $d_N > 1/2$, and it can be verified straightforwardly that for all pairs (d_N, ϕ_{\max}) that lie in the displayed region $\mathcal{S}(\phi, \beta, d_N) < 0$, also (5.2.6) is satisfied as can be noticed in Fig. 5.1 (right). We observe that the larger $\phi = \phi_{\max}$ is chosen, the closer d_N needs to be chosen near one, i.e., the narrower the size distribution must be to ensure hyperbolicity. In the most extreme case, for $\phi_{\max} = 1$, hyperbolicity can be observed only for $d_N > d_N^* = 0.914022$, this value is the relevant root of $\mathcal{S}(1, \beta, d_N) = 0$. Consequently, hyperbolicity, and therefore stability, can be ensured for the DG model only if the suspension is nearly monodisperse, a result that sharply contrasts with the HS model. The decisive observation is, however, that our result is independent of the number of species N .

We now present some numerical examples for two species using the method SPEC-INT

described in Chapter 4 (the Section 4.3). Since $m = 3$ (the rank of matrix \mathbf{A} and \mathbf{B}), the expressions for calculating the right and left eigenvalues of the Jacobian matrix in this model are similar to that of HS model (see equation (4.2.4)), then for each eigenvalue λ , which is approximated by a root finder that uses the interlacing property, we obtain the following right eigenvector $\mathbf{x} = (x_1, \dots, x_N)^T$ where:

$$x_i = -\frac{1}{v_i - \lambda} (b_{i,1}\xi_1 + b_{i,2}\xi_2 + b_{i,3}\xi_3), \quad i = 1, \dots, N,$$

with

$$\begin{aligned} \xi_1 &= (1 + [\mathbf{a}_2, \mathbf{b}_2])(1 + [\mathbf{a}_3, \mathbf{b}_3]) - [\mathbf{a}_2, \mathbf{b}_3][\mathbf{a}_3, \mathbf{b}_2], \\ \xi_2 &= [\mathbf{a}_2, \mathbf{b}_3][\mathbf{a}_3, \mathbf{b}_1] - [\mathbf{a}_2, \mathbf{b}_1](1 + [\mathbf{a}_3, \mathbf{b}_3]), \\ \xi_3 &= [\mathbf{a}_2, \mathbf{b}_1][\mathbf{a}_3, \mathbf{b}_2] - (1 + [\mathbf{a}_2, \mathbf{b}_2])[\mathbf{a}_3, \mathbf{b}_1], \end{aligned}$$

where

$$\mathbf{a}_l = (\alpha_1^l, \alpha_2^l, \dots, \alpha_N^l), \quad \mathbf{b}_l = (\beta_1^l, \beta_2^l, \dots, \beta_N^l), \quad l = 1, 2, 3,$$

with α_i^l, β_i^l are given by (5.2.2)

Figure 5.2 shows the approximate solutions at $t = 50$ s and $t = 300$ s for two species with the respective total concentration. The parameters are the same of Subsection 4.4.1, except that in this case we take $D_2 = 4.54 \times 10^{-4}$ and $\phi_{\max} = 0.9$, with this new data the hypothesis of the Lemma 5.2.1 are satisfied which guarantee hyperbolicity. We also show in Fig. 5.3 the numerical results obtained with the DG model with four species using the data of Example 4.4.2 which do not satisfy the sufficient conditions of Lemma 5.2.1. It can be observed overshoots for some species, which differ of results obtained along this work with other models.

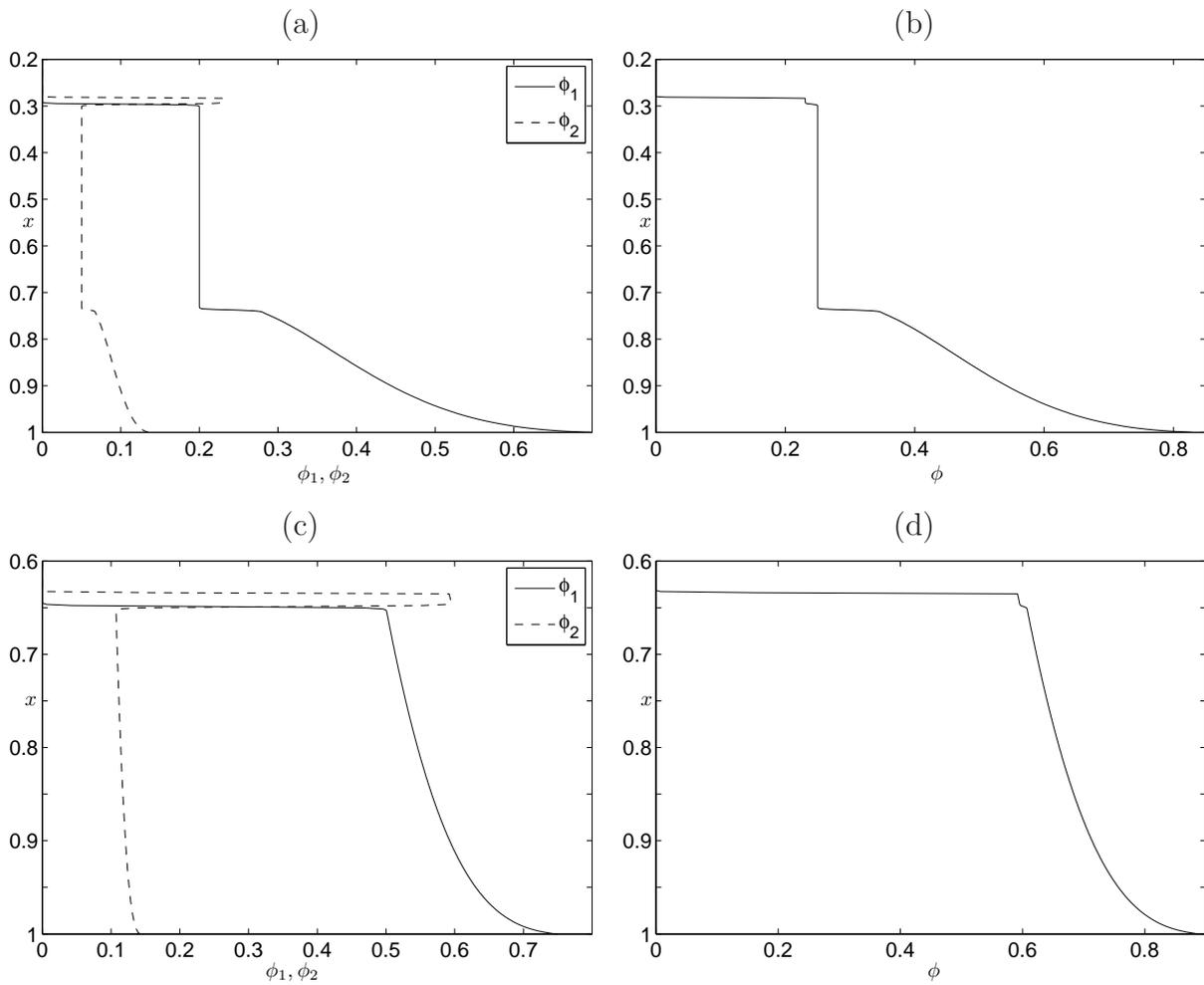


Figure 5.2: Davis-Gecol model with $N = 2$: numerical solution for ϕ_1 , ϕ_2 and ϕ at $t = 50$ s (a, b), and the corresponding solutions at $t = 300$ s (c, d).

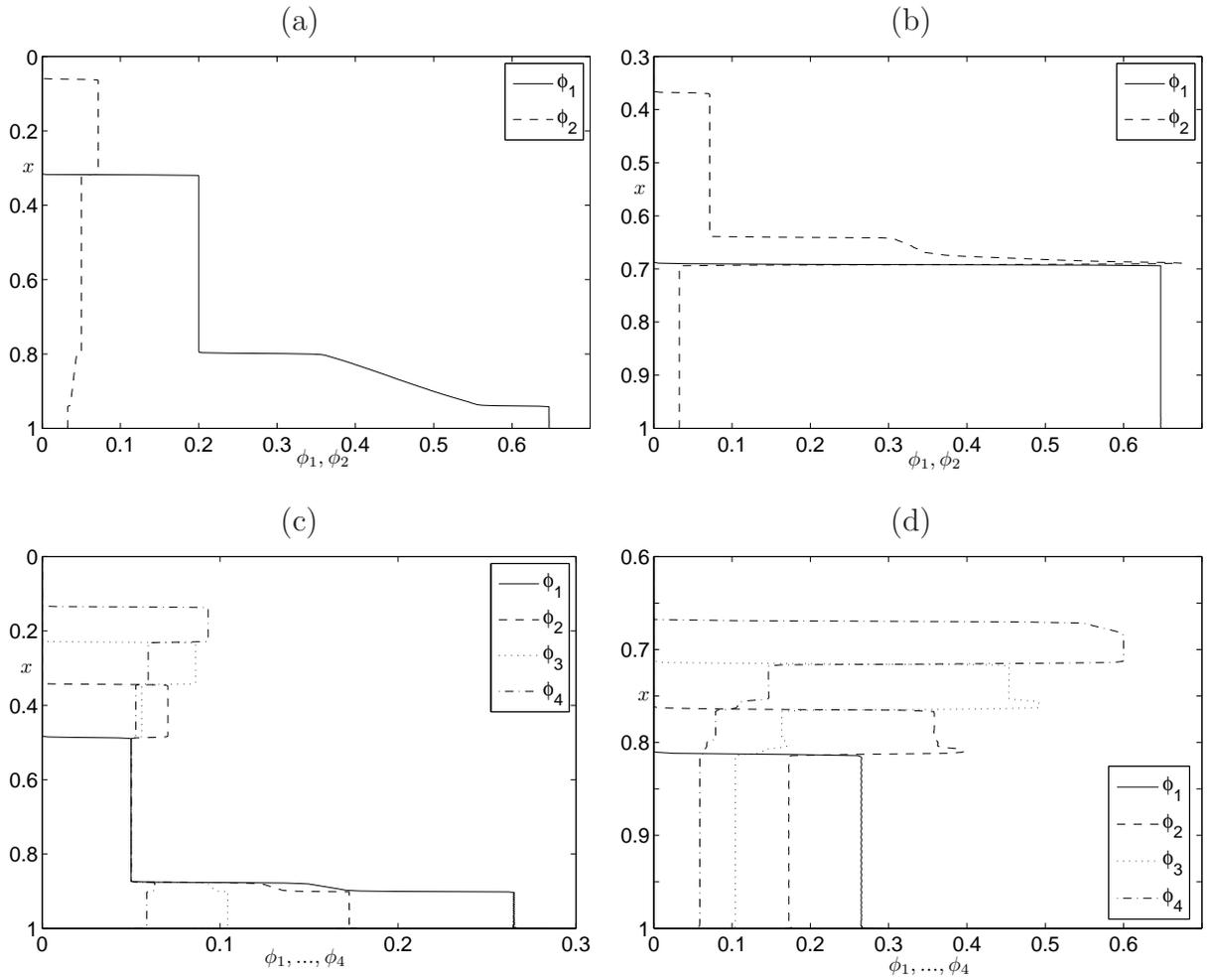


Figure 5.3: Numerical solutions for DG model with $N = 2$ (top row) at $t = 50$ s (a) and $t = 300$ s (b). Bottom row corresponds to approximate solutions for $N = 4$ at $t = 50$ s (c) and $t = 300$ s (d).

5.3 BW and HS models for $\beta_3 < 0$

In Chapter 3 we studied the hyperbolicity for the BW and HS under the assumption $\beta_3 = 0$. This mild consideration let us to hand roughly expressions in order to simplify many computations and put the main results into perspective. In this section we remove that assumption and we will see that it is still possible to obtain some sufficient conditions for hyperbolicity in this general case, but this is not free due to the algebraic expressions being more involved. However, some parts of lemma of Chapter 3 are valid with $\beta_3 < 0$ and most of the computations can be done in a straightforward manner.

Then using the notation of Subsection 3.4.2, we can write the settling velocity (3.4.7) as

$$\begin{aligned} v_i(\Phi) &= v_i(p_1, \dots, p_4) \\ &= d_i^2 \varphi((\beta_0 + n)p_1 + \beta_1 d_i^{-1} p_2 + \beta_2 d_i^{-2} p_3 + \beta_3 d_i^{-3} p_4)(1 - p_1)^n, \quad i = 1, \dots, N, \end{aligned} \quad (5.3.1)$$

where $\varphi(z) = 1 + z$, $n = 0$ for the BW model and $\varphi(z) = \exp(z)$, $n \geq 0$, arbitrary for the HS model.

By taking $\eta_i := \varphi(\mathbf{s}_i^T \Phi + n\phi)$ and $\eta'_i := \varphi'(\mathbf{s}_i^T \Phi + n\phi)$ for $i = 1, \dots, N$, where $\varphi'(z) := d\varphi(z)/dz$. The entries of matrix \mathbf{A} are $\alpha_i^k = d_i^{k-1}$, $k = 1, 2, 3, 4$, $i = 1, \dots, N$ and for the matrix \mathbf{B} are given by

$$\begin{aligned} \beta_i^1 &= d_i^2 \phi_i (1 - \phi)^{n-1} ((1 - \phi)(\beta_0 + n)\eta'_i - n\eta_i), \\ \beta_i^k &= d_i^{3-k} \phi_i (1 - \phi)^n \beta_{k-1} \eta'_i, \quad k = 2, 3, 4; \quad i = 1, \dots, N. \end{aligned}$$

We now calculate the determinants $\alpha_i^J := \det \mathbf{A}^{L,J}$ and $\beta_i^J := \det \mathbf{B}^{L,J}$ in the formula (3.2.15) for $m = 4$,

$$\begin{aligned} \gamma_i &= \alpha_i^1 \beta_i^1 + \alpha_i^2 \beta_i^2 + \alpha_i^3 \beta_i^3 + \alpha_i^4 \beta_i^4 \\ &\quad + \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\alpha_{ij}^{12} \beta_{ij}^{12} + \alpha_{ij}^{13} \beta_{ij}^{13} + \alpha_{ij}^{23} \beta_{ij}^{23} + \alpha_{ij}^{14} \beta_{ij}^{14} + \alpha_{ij}^{24} \beta_{ij}^{24} + \alpha_{ij}^{34} \beta_{ij}^{34}}{v_j - v_i} \\ &\quad + \sum_{\substack{j,k=1 \\ i \neq j < k \neq i}}^N \frac{\alpha_{ijk}^{123} \beta_{ijk}^{123} + \alpha_{ijk}^{234} \beta_{ijk}^{234} + \alpha_{ijk}^{134} \beta_{ijk}^{134} + \alpha_{ijk}^{124} \beta_{ijk}^{124}}{(v_k - v_i)(v_j - v_i)} + \sum_{\substack{j,k,l=1 \\ j < k < l \\ j,k,l \neq i}}^N \frac{\alpha_{ijkl}^{1234} \beta_{ijkl}^{1234}}{(v_k - v_i)(v_j - v_i)(v_l - v_i)}, \end{aligned} \quad (5.3.2)$$

As always we adopt the convention that sums over a void index range are zero, and utilize the following auxiliary notation:

$$\begin{aligned} \sigma_{ijk} &:= d_i + d_j + d_k, & \pi_{ijk} &:= (d_j - d_i)(d_k - d_i)(d_k - d_j), \\ \tilde{\sigma}_{ijk} &:= d_i d_j + d_i d_k + d_j d_k, & \pi_{ijkl} &:= (d_j - d_i)(d_k - d_i)(d_l - d_i)(d_l - d_j)(d_l - d_k)(d_k - d_j). \end{aligned} \quad (5.3.3)$$

We then obtain

$$\begin{aligned}
\alpha_i^1 &= 1, & \alpha_i^4 &= d_i^3, & \alpha_{ij}^{23} &= d_i d_j (d_j - d_i), & \alpha_{ij}^{34} &= d_i^2 d_j^2 (d_j - d_i), & \alpha_{ijk}^{134} &= \tilde{\sigma}_{ijk} \pi_{ijk}, \\
\alpha_i^2 &= d_i, & \alpha_{ij}^{12} &= d_j - d_i, & \alpha_{ij}^{14} &= d_j^3 - d_i^3, & \alpha_{ijk}^{123} &= \pi_{ijk}, & \alpha_{ijk}^{124} &= \sigma_{ijk} \pi_{ijk}, \\
\alpha_i^3 &= d_i^2, & \alpha_{ij}^{13} &= d_j^2 - d_i^2, & \alpha_{ij}^{24} &= d_i d_j (d_j^2 - d_i^2), & \alpha_{ijk}^{234} &= d_i d_j d_k \pi_{ijk}, & \alpha_{ijkl}^{1234} &= \pi_{ijkl},
\end{aligned} \tag{5.3.4a}$$

$$\begin{aligned}
\beta_{ij}^{12} &= \phi_i \phi_j (1 - \phi)^{2n-1} d_i d_j \beta_1 [(1 - \phi)(\beta_0 + n) \eta'_i \eta'_j (d_i - d_j) - n(\eta_i \eta'_j d_i - \eta'_i \eta_j d_j)], \\
\beta_{ij}^{13} &= \phi_i \phi_j (1 - \phi)^{2n-1} \beta_2 ((1 - \phi)(\beta_0 + n) \eta'_i \eta'_j (d_i^2 - d_j^2) - n(\eta_i \eta'_j d_i^2 - \eta'_i \eta_j d_j^2)), \\
\beta_{ij}^{14} &= \phi_i \phi_j (1 - \phi)^{2n-1} d_i^{-1} d_j^{-1} \beta_3 ((1 - \phi)(\beta_0 + n) \eta'_i \eta'_j (d_i^3 - d_j^3) - n(\eta_i \eta'_j d_i^3 - \eta'_i \eta_j d_j^3)), \\
\beta_{ij}^{23} &= \phi_i \phi_j (1 - \phi)^{2n} \beta_1 \beta_2 \eta'_i \eta'_j (d_i - d_j), \\
\beta_{ij}^{24} &= \phi_i \phi_j (1 - \phi)^{2n} \beta_1 \beta_3 \eta'_i \eta'_j d_i^{-1} d_j^{-1} (d_i^2 - d_j^2), \\
\beta_{ij}^{34} &= \phi_i \phi_j (1 - \phi)^{2n} \beta_2 \beta_3 \eta'_i \eta'_j d_i^{-1} d_j^{-1} (d_i - d_j),
\end{aligned} \tag{5.3.4b}$$

$$\begin{aligned}
\beta_{ijk}^{123} &= -(1 - \phi)^{3n-1} \phi_i \phi_j \phi_k \beta_1 \beta_2 \{ (1 - \phi)(\beta_0 + n) \pi_{ijk} \eta'_i \eta'_j \eta'_k \\
&\quad + n [d_i^2 (d_j - d_k) \eta_i \eta'_j \eta'_k - d_j^2 (d_i - d_k) \eta'_i \eta_j \eta'_k + d_k^2 (d_i - d_j) \eta'_i \eta'_j \eta_k] \}, \\
\beta_{ijk}^{124} &= -(1 - \phi)^{3n-1} \phi_i \phi_j \phi_k \beta_1 \beta_3 \{ (1 - \phi)(\beta_0 + n) d_i^{-1} d_j^{-1} d_k^{-1} \tilde{\sigma}_{ijk} \pi_{ijk} \eta'_i \eta'_j \eta'_k \\
&\quad + n [d_i^2 d_j^{-1} d_k^{-1} (d_j^2 - d_k^2) \eta_i \eta'_j \eta'_k - d_j^2 d_i^{-1} d_k^{-1} (d_i^2 - d_k^2) \eta'_i \eta_j \eta'_k \\
&\quad + d_k^2 d_i^{-1} d_j^{-1} (d_i^2 - d_j^2) \eta'_i \eta'_j \eta_k] \},
\end{aligned} \tag{5.3.4c}$$

$$\begin{aligned}
\beta_{ijk}^{134} &= -(1 - \phi)^{3n-1} \phi_i \phi_j \phi_k \beta_2 \beta_3 \{ (1 - \phi)(\beta_0 + n) d_i^{-1} d_j^{-1} d_k^{-1} \sigma_{ijk} \pi_{ijk} \eta'_i \eta'_j \eta'_k \\
&\quad + n [d_i^2 d_j^{-1} d_k^{-1} (d_j - d_k) \eta_i \eta'_j \eta'_k - d_j^2 d_i^{-1} d_k^{-1} (d_i - d_k) \eta'_i \eta_j \eta'_k \\
&\quad + d_k^2 d_i^{-1} d_j^{-1} (d_i - d_j) \eta'_i \eta'_j \eta_k] \},
\end{aligned}$$

$$\begin{aligned}
\beta_{ijk}^{234} &= -(1 - \phi)^{3n} \phi_i \phi_j \phi_k \beta_1 \beta_2 \beta_3 d_i^{-1} d_j^{-1} d_k^{-1} \pi_{ijk} \eta'_i \eta'_j \eta'_k, \\
\beta_{ijkl}^{1234} &= (1 - \phi)^{4n-1} \phi_i \phi_j \phi_k \phi_l \beta_1 \beta_2 \beta_3 \left\{ (1 - \phi)(\beta_0 + n) \frac{\pi_{ijkl}}{d_i d_j d_k d_l} \eta'_i \eta'_j \eta'_k \eta'_l \right. \\
&\quad \left. + n \left[\frac{d_i^2 \pi_{jkl}}{d_j d_k d_l} \eta_i \eta'_j \eta'_k \eta'_l - \frac{d_j^2 \pi_{ikl}}{d_i d_k d_l} \eta'_i \eta_j \eta'_k \eta'_l + \frac{d_k^2 \pi_{ijl}}{d_i d_j d_l} \eta'_i \eta'_j \eta_k \eta'_l - \frac{d_l^2 \pi_{ijk}}{d_i d_j d_k} \eta'_i \eta'_j \eta'_k \eta_l \right] \right\}.
\end{aligned} \tag{5.3.4d}$$

Substantial simplifications in the expressions β_J^J occur for the BW model, $\eta'_i = 1$ and $n = 0$, and for the HS model, where $\eta_i = \eta'_i$.

5.3.1 The Batchelor and Wen (BW) model

The coefficients γ_i , $i = 1, \dots, N$ can be rewritten now as $\gamma_i = \phi_i(\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i} + \mathcal{S}_{4,i})$. Inserting the expressions (5.3.3) and (5.3.4) into (5.3.2) and defining $\hat{\eta}_i := 1 + \mathbf{s}_i^T \Phi$, we obtain

$$\begin{aligned}
\mathcal{S}_{1,i} &:= d_i^2(\beta_0 + \beta_1 + \beta_2 + \beta_3), \\
\mathcal{S}_{2,i} &:= \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\phi_j}{d_j^2 \hat{\eta}_j - d_i^2 \hat{\eta}_i} \left\{ -\beta_0 \beta_1 d_i d_j (d_j - d_i)^2 - \beta_0 \beta_2 (d_j^2 - d_i^2)^2 - \beta_1 \beta_2 d_i d_j (d_j - d_i)^2 \right. \\
&\quad \left. - \beta_1 \beta_3 (d_j^2 - d_i^2)^2 - \beta_0 \beta_3 \frac{(d_j^3 - d_i^3)^2}{d_i d_j} - \beta_2 \beta_3 d_i d_j (d_j - d_i)^2 \right\}, \\
\mathcal{S}_{3,i} &:= \sum_{\substack{j,k=1 \\ i \neq j < k \neq i}}^N \frac{\phi_j \phi_k \pi_{ijk}^2}{(d_k^2 \hat{\eta}_k - d_i^2 \hat{\eta}_i)(d_j^2 \hat{\eta}_j - d_i^2 \hat{\eta}_i)} \left\{ -\beta_0 \left(\beta_1 \beta_2 + (\beta_1 \beta_3 + \beta_2 \beta_3) \frac{\sigma_{ijk} \tilde{\sigma}_{ijk}}{d_i d_j d_k} \right) - \beta_1 \beta_2 \beta_3 \right\}, \\
\mathcal{S}_{4,i} &:= \sum_{\substack{j,k,l=1 \\ j < k < l \\ j,k,l \neq i}}^N \frac{\phi_j \phi_k \phi_l \pi_{ijkl}^2 \beta_0 \beta_1 \beta_2 \beta_3}{(d_j^2 \hat{\eta}_j - d_i^2 \hat{\eta}_i)(d_k^2 \hat{\eta}_k - d_i^2 \hat{\eta}_i)(d_l^2 \hat{\eta}_l - d_i^2 \hat{\eta}_i) d_i d_j d_k d_l}.
\end{aligned} \tag{5.3.5}$$

Clearly, we have $\mathcal{S}_{1,i} < 0$ for $\phi_i > 0$; in addition, $\mathcal{S}_{1,i}$ is independent of Φ or N .

Related to the other terms, we suppose (see Subsection 3.4.3) that there is a constant $\theta \geq 1$ such that

$$-\mathbf{s}_N^T \Phi = \sum_{j=1}^N \left(-\sum_{\nu=0}^3 \frac{\beta_\nu d_j^\nu}{d_N^\nu} \right) \phi_j \leq \frac{1}{1 + \theta}. \tag{5.3.6}$$

This last inequality implies that for $i < j$, the following inequalities (which is a simple generalization of (3.4.16)) hold:

$$0 < (d_i^2 \hat{\eta}_i - d_j^2 \hat{\eta}_j)^{-1} \leq ((1 + \mathbf{s}_j^T \Phi)(d_i^2 - d_j^2))^{-1} \leq \left(-\theta(d_i^2 - d_j^2) \sum_{\nu=0}^3 \frac{\beta_\nu}{d_j^\nu} \mathbf{d}_\nu^T \Phi \right)^{-1}. \tag{5.3.7}$$

Of course, we can estimate the last term in (5.3.7) by omitting some of the summands. This result is the main tool for the next lemma.

Lemma 5.3.1 *The quantities $\mathcal{S}_{p,i}$, $p = 1, \dots, 4$ defined in (5.3.5) satisfy the following inequalities:*

$$\mathcal{S}_{2,i} \leq -d_i^2 \theta^{-1} (2\beta_0 + \beta_1 + \beta_2 + (7 + 3d_N^{-1})\beta_3), \tag{5.3.8}$$

$$\mathcal{S}_{3,i} \leq -d_i^2 \theta^{-2} \left[2\beta_0 + \beta_3 \left(2 + \frac{9}{2d_N} + \frac{6}{d_N^3} \right) \right], \tag{5.3.9}$$

$$\mathcal{S}_{4,i} \leq -2d_i^2\beta_1\theta^{-3}. \quad (5.3.10)$$

Proof. Since $\hat{\eta}_i > \hat{\eta}_j$ for $i < j$, the summands of $\mathcal{S}_{2,i}$ with $j < i$ and $j > i$ are negative and positive, respectively; let us denote the corresponding partial sums by $\mathcal{S}_{2,i}^- \leq 0$ and $\mathcal{S}_{2,i}^+ \geq 0$, with $\mathcal{S}_{2,i} = \mathcal{S}_{2,i}^- + \mathcal{S}_{2,i}^+$. We start finding a bound $\mathcal{S}_{2,i}^+$ in such a way that this quantity is compensated by the terms of \mathcal{S}_1 .

Let us now turn to $\mathcal{S}_{2,i}^+$. We here get

$$\begin{aligned} \mathcal{S}_{2,i}^+ &\leq -\frac{d_i^2}{\theta}(2\beta_0 + \beta_2) \\ &\quad -\frac{1}{\theta} \sum_{j=i+1}^N \left\{ \frac{\beta_1\beta_3(d_j^2 - d_i^2)^2 d_j^3 \phi_j}{\beta_3(d_i^2 - d_j^2) \mathbf{d}_3^T \Phi} + \frac{\beta_0\beta_3(d_j^3 - d_i^3)^2 \phi_j}{\beta_0 d_i d_j (d_i^2 - d_j^2) \phi} + \frac{\beta_2\beta_3 d_i (d_i - d_j)^2 d_j^3 \phi_j}{\beta_2 (d_i^2 - d_j^2) \mathbf{d}_2^T \Phi} \right\}, \end{aligned} \quad (5.3.11)$$

where the first term arises by repeating exactly the procedure performed to obtain the bound (3.4.17) in lemma 3.4.3. To deal with the second term in the summands of (5.3.11), we note that

$$\begin{aligned} \frac{(d_j^3 - d_i^3)^2}{(d_i^2 - d_j^2) d_i d_j} &= \frac{(d_i - d_j)^2 (d_i^2 + d_i d_j + d_j^2)^2}{(d_i^2 - d_j^2) d_i d_j} \leq (d_i^2 + d_i d_j + d_j^2) \left(1 + \frac{d_i}{d_j} + \frac{d_j}{d_i} \right) \\ &\leq 3d_i^2 (2 + d_N^{-1}). \end{aligned}$$

Consequently, we obtain from (5.3.11) the following inequality, which implies (5.3.8):

$$\mathcal{S}_{2,i}^+ \leq -\frac{d_i^2}{\theta}(2\beta_0 + \beta_2) - \frac{d_i^2}{\theta} \sum_{j=i+1}^N \left\{ \beta_1 \frac{d_j^3 \phi_j}{\mathbf{d}_3^T \Phi} + 3(2 + d_N^{-1}) \beta_3 \frac{\phi_j}{\phi} + \beta_3 \frac{d_j^2 \phi_j}{\mathbf{d}_2^T \Phi} \right\}.$$

Since only those summands of $\mathcal{S}_{3,i}$ are positive for which either $i < j$ and $i < k$ or $i > j$ and $i > k$, we rewrite $\mathcal{S}_{3,i}$ as $\mathcal{S}_{3,i} = \mathcal{S}_{3,i}^- + \mathcal{S}_{3,i}^{+,1} + \mathcal{S}_{3,i}^{+,2}$, where $\mathcal{S}_{3,i}^- < 0$, $\mathcal{S}_{3,i}^{+,1} > 0$ and $\mathcal{S}_{3,i}^{+,2} > 0$, and $\mathcal{S}_{3,i}^{+,1}$ and $\mathcal{S}_{3,i}^{+,2}$ are the partial of $\mathcal{S}_{3,i}$ for which $j > i$, $k > i$ and $k \neq j$ and $j < i$, $k < i$ and $k \neq j$, respectively.

Applying several versions of (5.3.7) to both factors in the denominator of the summands of $\mathcal{S}_{3,i}^{+,1}$, we obtain

$$\begin{aligned} \mathcal{S}_{3,i}^{+,1} &\leq -\frac{1}{\theta^2} \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{(\beta_0 + \beta_3) \pi_{ijk}^2 d_j \phi_j d_k^2 \phi_k}{(d_k^2 - d_i^2)(d_j^2 - d_i^2) \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \\ &\quad - \frac{\beta_3}{\theta^2} \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_k \sigma_{ijk} \tilde{\sigma}_{ijk} \pi_{ijk}^2}{(d_k^2 - d_i^2)(d_j^2 - d_i^2) d_i d_j d_k \phi} \left(\frac{d_j \phi_j}{\mathbf{d}_1^T \Phi} + \frac{d_j^2 \phi_j}{\mathbf{d}_2^T \Phi} \right) \end{aligned}$$

$$\leq -(\beta_0 + \beta_3) \frac{d_i^2}{\theta^2} - \frac{\beta_3}{\theta^2} \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_k \sigma_{ijk} \tilde{\sigma}_{ijk} (d_i - d_j)(d_i - d_k)(d_j - d_k)^2}{(d_i + d_k)(d_i + d_j)d_i d_j d_k \phi} \left(\frac{d_j \phi_j}{\mathbf{d}_1^T \Phi} + \frac{d_j^2 \phi_j}{\mathbf{d}_2^T \Phi} \right).$$

Noting that for $j, k > i$, we have that $\sigma_{ijk} \leq 3d_i$ and

$$\frac{\tilde{\sigma}_{ijk}}{(d_i + d_j)(d_i + d_k)} \leq \frac{3}{4}, \quad (d_i - d_j)(d_i - d_k) \leq d_i^2, \quad \frac{(d_j - d_k)^2}{d_i d_j d_k} \leq d_N^{-1}, \quad (5.3.12)$$

we finally obtain the inequality

$$\mathcal{S}_{3,i}^{+,1} \leq -\frac{d_i^2}{\theta^2} \left[\beta_0 + \beta_3 \left(1 + \frac{9d_i}{2d_N} \right) \right] \leq -\frac{d_i^2}{\theta^2} \left[\beta_0 + \beta_3 \left(1 + \frac{9}{2d_N} \right) \right]. \quad (5.3.13)$$

Furthermore, using the order $d_j > d_k > d_i$, and the fact that $\sigma_{ijk} \leq 3d_j$ for $j < k < i$ and the version of (5.3.7) with the roles of i and j interchanged we have that

$$\begin{aligned} \mathcal{S}_{3,i}^{+,2} &\leq -\frac{1}{\theta^2} \sum_{\substack{j,k=1 \\ j < k}}^{i-1} \frac{(\beta_0 + \beta_3) \pi_{ijk}^2 d_i \phi_j d_i^2 \phi_k}{(d_k^2 - d_i^2)(d_j^2 - d_i^2) \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \\ &\quad - \frac{\beta_3}{\theta^2} \sum_{\substack{j,k=1 \\ j < k}}^{i-1} \frac{\sigma_{ijk} \tilde{\sigma}_{ijk} \pi_{ijk}^2 \phi_j}{(d_k^2 - d_i^2)(d_j^2 - d_i^2) d_i d_j d_k \phi} \left(\frac{d_i \phi_k}{\mathbf{d}_1^T \Phi} + \frac{d_i^2 \phi_k}{\mathbf{d}_2^T \Phi} \right) \\ &\leq -\frac{\beta_0 + \beta_3}{\theta^2 \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \sum_{\substack{j,k=1 \\ j < k}}^{i-1} (d_k - d_j)^2 d_i \phi_j d_i^2 \phi_k \\ &\quad - \frac{\beta_3}{\theta^2} \sum_{\substack{j,k=1 \\ j < k}}^{i-1} \frac{\sigma_{ijk} (d_j - d_i)(d_j - d_k) \phi_j}{d_i \phi} \left(\frac{d_i \phi_k}{\mathbf{d}_1^T \Phi} + \frac{d_i^2 \phi_k}{\mathbf{d}_2^T \Phi} \right) \\ &\leq -\frac{d_i^2 (\beta_0 + \beta_3)}{\theta^2 \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \sum_{\substack{j,k=1 \\ j < k}}^{i-1} d_j^2 \phi_j d_k \phi_k - \frac{3\beta_3}{\theta^2} \sum_{\substack{j,k=1 \\ j < k}}^{i-1} \frac{d_j^2 d_j \phi_j}{d_i \phi} \left(\frac{d_i \phi_k}{\mathbf{d}_1^T \Phi} + \frac{d_i^2 \phi_k}{\mathbf{d}_2^T \Phi} \right) \\ &\leq -\frac{d_i^2}{\theta^2} \left[\beta_0 + \beta_3 \left(1 + \frac{6}{d_N^3} \right) \right]. \end{aligned}$$

Combining this with (5.3.13) we obtain (5.3.9).

Finally, we rewrite $\mathcal{S}_{4,i}$ as $\mathcal{S}_{4,i} = \mathcal{S}_{4,i}^- + \mathcal{S}_{4,i}^{+,1} + \mathcal{S}_{4,i}^{+,2}$, where $\mathcal{S}_{4,i}^{+,1}$ is the sum of all summands of $\mathcal{S}_{4,i}$ for which exactly one factor in the denominator is positive, i.e., i is the second largest species, and $\mathcal{S}_{4,i}^{+,2}$ is the sum of all summands of $\mathcal{S}_{4,i}$ for which all three factors in the denominator are positive, i.e., $i > j$, $i > k$ and $i > l$, that is, i is the smallest species. To estimate $\mathcal{S}_{4,i}^{+,1}$, we first note that

$$\mathcal{S}_{4,i}^{+,1} = \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \frac{\phi_j \phi_k \phi_l \pi_{ijkl}^2 \beta_0 \beta_1 \beta_2 \beta_3}{(d_j^2 \hat{\eta}_j - d_i^2 \hat{\eta}_i)(d_k^2 \hat{\eta}_k - d_i^2 \hat{\eta}_i)(d_l^2 \hat{\eta}_l - d_i^2 \hat{\eta}_i) d_i d_j d_k d_l}$$

$$\begin{aligned}
&\leq -\frac{\beta_1}{\theta^3 \phi \mathbf{d}_2^T \Phi \mathbf{d}_3^T \Phi} \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \frac{\phi_j \phi_k \phi_l (d_k - d_j)^2 (d_l - d_j)^2 (d_l - d_k)^2 d_k^2 d_l^3}{d_i d_j d_k d_l} \\
&\leq -\frac{\beta_1}{\theta^3 \phi \mathbf{d}_2^T \Phi \mathbf{d}_3^T \Phi} \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \frac{\phi_j \phi_k \phi_l d_j^4 d_k^2 d_l^3}{d_i d_j d_k d_l} \leq -\frac{\beta_1}{\theta^3 \phi \mathbf{d}_2^T \Phi \mathbf{d}_3^T \Phi} \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \frac{\phi_j \phi_k \phi_l d_j^3 d_k^3 d_l^2}{d_i} \\
&\leq -\frac{\beta_1 d_i^2}{\theta^3 \phi \mathbf{d}_2^T \Phi \mathbf{d}_3^T \Phi} \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \phi_j \phi_k \phi_l d_j^3 d_k^2 \leq -\frac{\beta_1 d_i^2}{\theta^3}. \tag{5.3.14}
\end{aligned}$$

By similar arguments we obtain the following estimate for $\mathcal{S}_{4,i}^{+,2}$:

$$\begin{aligned}
\mathcal{S}_{4,i}^{+,2} &= \sum_{\substack{j,k,l=1 \\ j < k < l}}^{i-1} \frac{\phi_j \phi_k \phi_l \pi_{ijkl}^2 \beta_0 \beta_1 \beta_2 \beta_3}{(d_j^2 \hat{\eta}_j - d_i^2 \hat{\eta}_i)(d_k^2 \hat{\eta}_k - d_i^2 \hat{\eta}_i)(d_l^2 \hat{\eta}_l - d_i^2 \hat{\eta}_i) d_i d_j d_k d_l} \\
&\leq -\frac{\beta_1 d_i^2}{\theta^3 \phi \mathbf{d}_2^T \Phi \mathbf{d}_3^T \Phi} \sum_{\substack{j,k,l=1 \\ j < k < l}}^{i-1} \frac{\phi_j \phi_k \phi_l d_i^2 d_j^3 d_k}{d_l} \leq -\frac{\beta_1 d_i^2}{\theta^3 \phi \mathbf{d}_2^T \Phi \mathbf{d}_3^T \Phi} \sum_{\substack{j,k,l=1 \\ j < k < l}}^{i-1} \phi_j \phi_k \phi_l d_j^3 d_k^2 \leq -\frac{\beta_1 d_i^2}{\theta^3}. \tag{5.3.15}
\end{aligned}$$

Inequality (5.3.10) is now a consequence of (5.3.14) and (5.3.15). \square

Corollary 5.3.1 *For the BW model, the following inequality is valid:*

$$\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i} + \mathcal{S}_{4,i} \leq d_i^2 M(\theta, \boldsymbol{\beta}, d_N), \tag{5.3.16}$$

where we define the function

$$\begin{aligned}
M(\theta, \boldsymbol{\beta}, d_N) &:= (1 - 2\theta^{-1} - 2\theta^{-2})\beta_0 + (1 - \theta^{-1} - 2\theta^{-3})\beta_1 + (1 - \theta^{-1})\beta_2 \\
&\quad + [1 - \theta^{-1}(7 + 3d_N^{-1}) - \theta^{-2}(2 + (9/2)d_N^{-1} + 6d_N^{-3})]\beta_3. \tag{5.3.17}
\end{aligned}$$

Proof. Combining the inequalities (5.3.8), (5.3.9) and (5.3.10) we obtain (5.3.16) and (5.3.17). Each of the inequalities (5.3.8), (5.3.9) and (5.3.10) estimates a non-negative sum from above, and therefore remains valid if the respective sum runs over a void index range, and is therefore zero. Consequently, (5.3.16) and (5.3.17) hold for arbitrary numbers of species N . \square

We have proved the following theorem.

Theorem 5.3.2 *Assume that θ is chosen such that for the smallest given particle size $d_N > 0$, the inequality*

$$M(\theta, \boldsymbol{\beta}, d_N) < 0 \tag{5.3.18}$$

is satisfied, where $M(\theta, \boldsymbol{\beta}, d_N)$ is defined in (5.3.17). If the maximum solids concentration ϕ_{\max} is chosen such that the inequality (5.3.6) is satisfied for all $\Phi \in \mathcal{D}_{\phi_{\max}}$ for this value of θ , then $\gamma_i < 0$ for $i = 1, \dots, 4$ and $\Phi \in \mathcal{D}_{\phi_{\max}}$, i.e., the model equations are strictly hyperbolic on $\mathcal{D}_{\phi_{\max}}$.

We remark first that for a given value of d_N , it is always possible to make all coefficients of beta's in (5.3.17) positive, and thereby to ensure that (5.3.18) holds, by choosing $\theta > 1$ large enough. On the other hand, the particular way in which d_N^{-1} appears in the coefficient of β_3 in (5.3.17) implies that in the case $\beta_3 < 0$, as we increase the particle size ratio, i.e. consider $d_N \rightarrow 0$ (that is, $d_N^{-1} \rightarrow \infty$), the smaller the set of admissible values of θ (that is, values of θ for which (5.3.18) holds) will become. Suppose that we choose an admissible value of θ , then (5.3.6) can hold either for a dilute suspension, i.e. ϕ is small, but for a large range of coefficients $\boldsymbol{\beta}$, or we consider relatively small (in absolute value) coefficients $\boldsymbol{\beta}$ and obtain a hyperbolicity (stability) result valid up to relatively large concentrations.

Furthermore, the strategy that has led to (5.3.17) has been motivated by the observation that $\beta_3 \leq 0$, but $|\beta_3| \ll 1$ (see our discussion in Section 3.3.2). For these reasons, we have performed the term cancellations and estimations in such a way that $1/d_N$, a potentially large number, appears only as a coefficient of β_3 . We stress that in case $\beta_3 = 0$ (Subsection 3.4.3), the set of admissible values of θ is independent of (the smallness of) d_N .

This analysis also shows that for $N = 3$ species, $\mathcal{S}_{4,i} = 0$ and the terms in which we divide by θ^3 in (5.3.17) do not appear; for $N = 2$, we additionally have $\mathcal{S}_{3,i} = 0$ and the terms in which we divide by θ^2 are zero.

Now, due to $M(\theta, \boldsymbol{\beta}, d_N)$ is a strictly decreasing function of θ , it is sufficient to solve $M(\theta, \boldsymbol{\beta}, d_N) = 0$ for θ to conclude that in this case, $M(\theta, \boldsymbol{\beta}, d_N) < 0$ for $\theta > \theta_{\min}$, but unlike the case $\beta_3 = 0$, $M(\theta, \boldsymbol{\beta}, d_N)$, and therefore θ_{\min} do depend on d_N , which we denote by $\theta_{\min} = \theta_{\min}(d_N)$. Therefore, we repeat the procedure of Subsection 3.4.3, page 35 for the purpose of determining the largest value ϕ^* of the total concentration ϕ up to which we can guarantee hyperbolicity, then we can rewrite the left-hand side of (5.3.6) as $\sigma_1\phi_1 + \dots + \sigma_N\phi_N$, where we define $\sigma_j := -\beta_0 - \beta_1 d_j d_N^{-1} - \beta_2 d_j^2 d_N^{-2} - \beta_3 d_j^3 d_N^{-3}$. Then the sought concentration ϕ^* solves the problem “minimize ϕ subject to $\sigma_1\phi_1 + \dots + \sigma_N\phi_N = (1 + \theta_{\min}(d_N))^{-1}$ ”. Expressing ϕ_1 in terms of ϕ_2, \dots, ϕ_N and ϕ , we can rewrite this equation as

$$\phi = (1 - \sigma_1^{-1}\sigma_2)\phi_2 + \dots + (1 - \sigma_1^{-1}\sigma_N)\phi_N + \sigma_1^{-1}(1 + \theta_{\min}(d_N))^{-1}.$$

Since $\sigma_1 > \sigma_2 > \dots > \sigma_N$, the coefficients of ϕ_2, \dots, ϕ_N on the right-hand side are all positive, and the minimum ϕ^* of ϕ is attained for $\phi_2 = \dots = \phi_N = 0$.

Consequently, the value ϕ^* is given here by

$$\phi^* = (\sigma_1(\beta, d_N)(1 + \theta_{\min}(d_N)))^{-1}, \quad \text{where} \quad M(\theta_{\min}(d_N), \beta, d_N) = 0.$$

As a numerical example, we consider the parameter vectors β given by (3.3.12). Figures 5.4 show plots of ϕ^* as a function of d_N for the cases of large and small Péclet numbers.

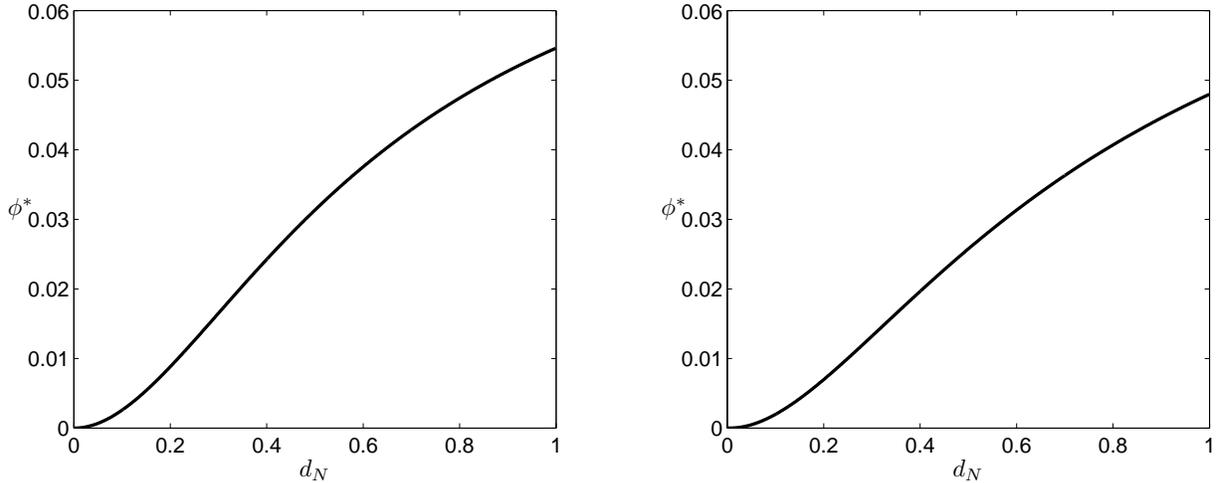


Figure 5.4: Maximum total concentrations ϕ^* for which hyperbolicity of the BW model is ensured with the coefficients (3.3.12): for large Péclet numbers (left) and for small Péclet numbers (right).

5.3.2 The Höfler and Schwarzer (HS) model

For this model we know that $\eta_i = \eta'_i$, therefore the coefficients γ_i of the secular equation given by (5.3.2) can be expressed as

$$\gamma_i = \phi_i(1 - \phi)^n \eta_i (\mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i} + \mathcal{S}_{4,i}),$$

where in terms of $\tilde{\eta}_i := \exp(\mathbf{s}_i^T \Phi)$ we define for the HS model

$$\begin{aligned} \mathcal{S}_{1,i} &:= d_i^2(\tilde{\beta}_0 + \beta_1 + \beta_2 + \beta_3), \\ \mathcal{S}_{2,i} &:= \sum_{\substack{j=1 \\ j \neq i}}^N \frac{\phi_j \tilde{\eta}_j}{d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i} \left\{ -(d_i - d_j)^2 \tilde{\beta}_0 (\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) - \tilde{\beta}_0 \beta_3 \frac{(d_i^3 - d_j^3)^2}{d_i d_j} \right. \\ &\quad \left. - (\beta_1 \beta_2 d_i d_j (d_i - d_j)^2 + \beta_1 \beta_3 (d_i^2 - d_j^2)^2 + \beta_2 \beta_3 d_i d_j (d_i - d_j)^2) \right\}, \end{aligned} \quad (5.3.19)$$

$$\mathcal{S}_{3,i} := \sum_{\substack{j,k=1 \\ j < k, j, k \neq i}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)} \left\{ -\tilde{\beta}_0 \left(\beta_1 \beta_2 + (\beta_1 \beta_3 + \beta_2 \beta_3) \frac{\sigma_{ijk} \tilde{\sigma}_{ijk}}{d_i d_j d_k} \right) \right. \\ \left. - \beta_1 \beta_2 \beta_3 \right\},$$

$$\mathcal{S}_{4,i} := \sum_{\substack{j,k,l=1 \\ j < k < l, j, k, l \neq i}} \frac{\phi_j \phi_k \phi_l \tilde{\eta}_j \tilde{\eta}_k \tilde{\eta}_l \pi_{ijkl}^2 \tilde{\beta}_0 \beta_1 \beta_2 \beta_3}{(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_l^2 \tilde{\eta}_l - d_i^2 \tilde{\eta}_i) d_i d_j d_k d_l},$$

with $\tilde{\beta}_0 = \beta_0 - \frac{n\phi}{1-\phi}$.

Next, we will prove some algebraic results that correspond to extensions of Lemma 3.4.6 and 3.4.7. In addition, the lemma 3.4.5 (or slight variants) will be used repeatedly, taking into account that this result holds by considering one more term in the sum, that is, for $i < j$ the following inequality holds

$$\frac{\tilde{\eta}_j}{d_i^2 \tilde{\eta}_i - d_j^2 \tilde{\eta}_j} \leq -\frac{1}{e(d_i^2 - d_j^2)} \left[\sum_{s=0}^3 \beta_m \frac{d_i^s - d_j^s}{d_i^s d_j^s} \mathbf{d}_s^T \Phi \right]^{-1}. \quad (5.3.20)$$

We first note that $\mathcal{S}_{1,i} < 0$. Then we analyze the positive and negative parts of $\mathcal{S}_{2,i}$, $\mathcal{S}_{3,i}$ and $\mathcal{S}_{4,i}$ separately, and show that we eventually obtain $\gamma_i < 0$.

Lemma 5.3.3 *Let us rewrite $\mathcal{S}_{2,i}$ as $\mathcal{S}_{2,i} = \mathcal{S}_{2,i}^+ + \mathcal{S}_{2,i}^-$, where $\mathcal{S}_{2,i}^+$ and $\mathcal{S}_{2,i}^-$ correspond to the summands of $\mathcal{S}_{2,i}$ with $j > i$ and $j < i$, respectively. Then $\mathcal{S}_{2,i}^- \leq 0$, and the following inequality holds:*

$$\mathcal{S}_{2,i}^+ \leq -\frac{d_i^2}{e} \left[\left(1 + 3 \frac{\beta_3}{d_N \beta_2} \right) \tilde{\beta}_0 + (\beta_1 + \beta_2 + \beta_3) \right]. \quad (5.3.21)$$

Proof. Since $\exp(\mathbf{s}_i^T \Phi) > \exp(\mathbf{s}_j^T \Phi)$ for $i < j$ and $\exp(\mathbf{s}_i^T \Phi) < \exp(\mathbf{s}_j^T \Phi)$ for $i > j$, the factor multiplying $\{\dots\}$ in the summands of $\mathcal{S}_{2,i}^-$ is always positive, while $\{\dots\} < 0$. This confirms that $\mathcal{S}_{2,i}^- \leq 0$ (note that for $i = 1$, the sum is void, i.e. $\mathcal{S}_{2,i}^- = 0$). To estimate $\mathcal{S}_{2,i}^+$, note first that from (5.3.20) we may conclude that

$$\mathcal{S}_{2,i}^+ \leq -\frac{\tilde{\beta}_0}{e} \sum_{j=i+1}^N \frac{(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) (d_i - d_j)^2 \phi_j}{(d_i^2 - d_j^2) \left[\beta_1 \frac{d_i - d_j}{d_i d_j} \mathbf{d}_1^T \Phi + \beta_2 \frac{d_i^2 - d_j^2}{d_i^2 d_j^2} \mathbf{d}_2^T \Phi \right]} - \frac{\tilde{\beta}_0 \beta_3}{\beta_2 e} \sum_{j=i+1}^N \frac{(d_i^3 - d_j^3)^2 d_i^2 d_j^2 \phi_j}{d_i d_j (d_i^2 - d_j^2)^2 \mathbf{d}_2^T \Phi} \\ - \frac{1}{e} \sum_{j=i+1}^N \left\{ \frac{\beta_2 (d_i - d_j)^2 d_i^2 d_j^2 \phi_j}{(d_i^2 - d_j^2) (d_i - d_j) \mathbf{d}_1^T \Phi} + \frac{\beta_1 (d_i^2 - d_j^2)^2 d_i^3 d_j^3 \phi_j}{(d_i^2 - d_j^2) (d_i^3 - d_j^3) \mathbf{d}_3^T \Phi} \right. \\ \left. + \frac{\beta_3 d_i d_j (d_i - d_j)^2 d_i^2 d_j^2 \phi_j}{(d_i^2 - d_j^2)^2 \mathbf{d}_2^T \Phi} \right\}$$

$$\begin{aligned}
&\leq -\frac{\tilde{\beta}_0}{e} \left(d_i^2 \sum_{j=i+1}^N \frac{d_j^2 (\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) \phi_j}{\beta_1 d_i d_j (d_i + d_j) \mathbf{d}_1^T \Phi + \beta_2 (d_i + d_j)^2 \mathbf{d}_2^T \Phi} \right. \\
&\quad \left. + \frac{\beta_3}{\beta_2} \sum_{j=i+1}^N \frac{d_i d_j (d_i^2 + d_i d_j + d_j^2)^2 \phi_j}{(d_i + d_j)^2 \mathbf{d}_2^T \Phi} \right) \\
&\quad - \frac{d_i^2}{e} \left(\sum_{j=i+1}^N \frac{\beta_2 d_j^2 \phi_j}{(d_i + d_j) \mathbf{d}_1^T \Phi} + \sum_{j=i+1}^N \frac{\beta_1 (d_i + d_j) d_i d_j^3 \phi_j}{(d_i^2 + d_i d_j + d_j^2) \mathbf{d}_3^T \Phi} + \sum_{j=i+1}^N \frac{\beta_3 d_i d_j^3 \phi_j}{(d_i + d_j)^2 \mathbf{d}_2^T \Phi} \right)
\end{aligned}$$

We may then continue estimating $\mathcal{S}_{2,i}^+$ as follows.

$$\begin{aligned}
\mathcal{S}_{2,i}^+ &\leq -\frac{d_i^2 \tilde{\beta}_0}{e} \left(\sum_{j=i+1}^N \frac{(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) d_j^2 \phi_j}{(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2) \sum_{k=i+1}^N d_k^2 \phi_k} + \frac{3\beta_3}{d_N \beta_2} \sum_{j=i+1}^N \frac{d_j^2 \phi_j}{\mathbf{d}_2^T \Phi} \right) \\
&\quad - \frac{d_i^2}{e} (\beta_1 + \beta_2 + \beta_3),
\end{aligned}$$

which implies (5.3.21). \square

Lemma 5.3.4 *Assume that we rewrite $\mathcal{S}_{3,i}$ as $\mathcal{S}_{3,i} = \mathcal{S}_{3,i}^- + \mathcal{S}_{3,i}^{+,1} + \mathcal{S}_{3,i}^{+,2}$, where $\mathcal{S}_{3,i}^{+,1}$ and $\mathcal{S}_{3,i}^{+,2}$ are the sums over all summands for which $j > i$, $k > i$ and $k \neq j$ and $j < i$, $k < i$ and $k \neq j$, respectively. Then we have $\mathcal{S}_{3,i}^- < 0$, $\mathcal{S}_{3,i}^{+,1} > 0$ and $\mathcal{S}_{3,i}^{+,2} > 0$. Furthermore, the following inequality holds:*

$$\mathcal{S}_{3,i}^{+,1} \leq -\frac{d_i^2}{e^2} \left[\left(1 + \frac{3\beta_3}{2\beta_1} + \frac{3\beta_3}{2\beta_2} \right) \tilde{\beta}_0 + \beta_3 \right]. \quad (5.3.22)$$

Finally, let us assume that the parameters $\boldsymbol{\beta}$ are related to the sizes \mathbf{d}_1 via the condition

$$\forall 1 \leq j < i \leq N : \quad \forall \phi \in [0, \phi_{\max}] : \quad \tilde{H}_{ij}(\phi, \boldsymbol{\beta}) < 0, \quad (5.3.23)$$

where we define the functions

$$\tilde{H}_{ij}(\phi; \boldsymbol{\beta}) := -\tilde{\beta}_0 \left(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2 + \beta_3 \frac{(d_i^2 + d_i d_j + d_j^2)^2}{d_i d_j} \right) \quad (5.3.24)$$

$$- (\beta_2 (\beta_1 + \beta_3) d_i d_j + \beta_1 \beta_3 (d_i + d_j)^2) - \phi G_{ij}(\phi, \boldsymbol{\beta}),$$

$$G_{ij}(\phi, \boldsymbol{\beta}) := (d_j - d_i)^2 \{ \tilde{\beta}_0 [\beta_1 \beta_2 + (\beta_1 \beta_3 + \beta_2 \beta_3) (1 + 2d_i^{-1} d_j)^2] + \beta_1 \beta_2 \beta_3 \}. \quad (5.3.25)$$

Then

$$\mathcal{S}_{2,i}^- + \mathcal{S}_{3,i}^{+,2} \leq 0. \quad (5.3.26)$$

Proof. The inequalities $\mathcal{S}_{3,i}^- < 0$, $\mathcal{S}_{3,i}^{+,1} > 0$ and $\mathcal{S}_{3,i}^{+,2} > 0$ are a simple consequence of the fact that only those summands of $\mathcal{S}_{3,i}$ are positive for which either $i < j$ and $i < k$ or $i > j$ and $i > k$, according to the ordering $d_1 > d_2 > \dots > d_N$. To deal with

$$\mathcal{S}_{3,i}^{+,1} = \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)} \left\{ -\tilde{\beta}_0 \left(\beta_1 \beta_2 + (\beta_1 \beta_3 + \beta_2 \beta_3) \frac{\sigma_{ijk} \tilde{\sigma}_{ijk}}{d_i d_j d_k} \right) - \beta_1 \beta_2 \beta_3 \right\},$$

By proceeding as in the proof of Lemma 3.4.7 (see right hand side of equation (3.4.33)), we get

$$- \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2 \tilde{\beta}_0 \beta_1 \beta_2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)} \leq -\frac{d_i^2 \tilde{\beta}_0}{e^2}.$$

We may estimate the other terms in $\mathcal{S}_{3,i}^{+,1}$ as follows:

$$\begin{aligned} \mathcal{S}_{3,i}^{+,1,1} &:= - \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \tilde{\beta}_0 \pi_{ijk}^2 (\beta_1 + \beta_2) \beta_3 \sigma_{ijk} \tilde{\sigma}_{ijk}}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i) d_i d_j d_k} \\ &\leq -\frac{\tilde{\beta}_0 (\beta_1 + \beta_2) \beta_3}{e^2 \beta_1 \beta_2 \mathbf{d}_2^T \Phi \mathbf{d}_1^T \Phi} \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{d_i^3 d_k^2 d_j \sigma_{ijk} \tilde{\sigma}_{ijk} \pi_{ijk}^2 \phi_j \phi_k}{(d_i^2 - d_j^2)(d_i^2 - d_k^2) d_i d_j d_k (d_i^2 - d_k^2) (d_i - d_j)} \\ &= -\left(\frac{\beta_3}{\beta_1} + \frac{\beta_3}{\beta_2} \right) \frac{d_i^2 \tilde{\beta}_0}{e^2 \mathbf{d}_2^T \Phi \mathbf{d}_1^T \Phi} \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{d_k \sigma_{ijk} \tilde{\sigma}_{ijk} (d_k - d_j)^2 \phi_j \phi_k}{(d_i + d_j)(d_i + d_k)^2}. \end{aligned}$$

Now, taking into account (5.3.12) and that $\sigma_{ijk}/(d_i + d_k) \leq 2$ for $i < j, k$, we get

$$\mathcal{S}_{3,i}^{+,1,1} \leq -\frac{3}{2e^2} \left(\frac{\beta_3}{\beta_1} + \frac{\beta_3}{\beta_2} \right) \frac{d_i^2 \tilde{\beta}_0}{\mathbf{d}_2^T \Phi \mathbf{d}_1^T \Phi} \sum_{\substack{j,k=i+1 \\ j < k}}^N d_k d_j^2 \phi_j \phi_k \leq -\frac{3d_i^2 \tilde{\beta}_0}{2e^2} \left(\frac{\beta_3}{\beta_1} + \frac{\beta_3}{\beta_2} \right). \quad (5.3.27)$$

Inequality (5.3.22) now follows from (5.3.27) and

$$\begin{aligned} \mathcal{S}_{3,i}^{+,1,2} &\leq -\frac{\beta_3}{e^2} \sum_{\substack{j,k=i+1 \\ j \neq k}}^N \frac{\phi_j \phi_k \pi_{ijk}^2 d_i^3 d_k d_j^2}{(d_i^2 - d_j^2)(d_i^2 - d_k^2)(d_i - d_k)(d_i^2 - d_j^2) \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \\ &\leq -\frac{\beta_3 d_i^2}{e^2}, \end{aligned}$$

where

$$\mathcal{S}_{3,i}^{+,1,2} = - \sum_{\substack{j,k=i+1 \\ j < k}}^N \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2 \beta_1 \beta_2 \beta_3}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)}$$

Next, we analyze

$$\mathcal{S}_{3,i}^{+,2} := \sum_{\substack{j,k=1 \\ j < k}}^{i-1} \frac{\phi_j \phi_k \tilde{\eta}_j \tilde{\eta}_k \pi_{ijk}^2}{(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)} \left\{ -\tilde{\beta}_0 \left(\beta_1 \beta_2 + (\beta_1 \beta_3 + \beta_2 \beta_3) \frac{\sigma_{ijk} \tilde{\sigma}_{ijk}}{d_i d_j d_k} \right) - \beta_1 \beta_2 \beta_3 \right\}.$$

As in the case of $\beta_3 = 0$, this term cannot be estimated easily and therefore will compensate it with $\mathcal{S}_{2,i}^-$, as expressed in (5.3.26). Notice that in order to ensure that our hyperbolicity result is also valid for $N = 3$, $\mathcal{S}_{3,i}^{+,1}$ should be compensated by one of the terms that have arisen *earlier* in our analysis. Observe now that

$$\mathcal{S}_{2,i}^- + \mathcal{S}_{3,i}^{+,2} = \sum_{j=1}^{i-1} \frac{\phi_j \tilde{\eta}_j (d_i - d_j)^2}{d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i} \mathcal{R}_{ij}, \quad (5.3.28)$$

where we define

$$\begin{aligned} \mathcal{R}_{ij} &:= -\tilde{\beta}_0 \left(\beta_1 d_i d_j + \beta_2 (d_i + d_j)^2 + \beta_3 \frac{(d_i^2 + d_i d_j + d_j^2)^2}{d_i d_j} \right) \\ &\quad - (\beta_2 (\beta_1 + \beta_3) d_i d_j + \beta_1 \beta_3 (d_i + d_j)^2) + \tilde{\mathcal{R}}_{ij}, \\ \tilde{\mathcal{R}}_{ij} &:= - \sum_{k=j+1}^{i-1} \frac{\phi_k (d_k - d_i)^2 (d_k - d_j)^2 \tilde{\eta}_k}{d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i} \left[\tilde{\beta}_0 \left(\beta_1 \beta_2 + (\beta_1 \beta_3 + \beta_2 \beta_3) \frac{\sigma_{ijk} \tilde{\sigma}_{ijk}}{d_i d_j d_k} \right) + \beta_1 \beta_2 \beta_3 \right]. \end{aligned} \quad (5.3.29)$$

Since $d_i < d_j$ and $d_i < d_k$ in these summands, and the factor multiplying \mathcal{R}_{ij} in (5.3.28) is positive, we will satisfy (5.3.26) by achieving that $\mathcal{R}_{ij} < 0$. Noting that for $j < k < i$

$$\begin{aligned} \frac{(d_k - d_i)^2 (d_k - d_j)^2 \tilde{\eta}_k}{d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i} &= \frac{(d_k - d_i)^2 (d_k - d_j)^2}{d_k^2 - d_i^2 \exp((\mathbf{s}_i^T - \mathbf{s}_k^T) \Phi)} \leq \frac{(d_k - d_j)^2 (d_k - d_i)}{d_k + d_i} \leq (d_j - d_i)^2, \\ \frac{\sigma_{ijk} \tilde{\sigma}_{ijk}}{d_i d_j d_k} &\leq \frac{(d_i + 2d_j)(d_i d_j + 2d_j^2)}{d_i^2 d_j} = (1 + 2d_i^{-1} d_j)^2, \end{aligned}$$

and using the function $G_{ij}(\phi, \boldsymbol{\beta})$ we have that $\tilde{\mathcal{R}}_{ij} \leq -G_{ij}(\phi, \boldsymbol{\beta})(\phi_{j+1} + \phi_{j+2} + \dots + \phi_{i-1})$. Thus, (5.3.26) holds if the parameters $\boldsymbol{\beta}$ are related to d_1, \dots, d_N by (5.3.23), where $\tilde{H}_{ij} := \tilde{H}_{ij}(\phi; \boldsymbol{\beta})$ is defined in (5.3.24). \square

Lemma 5.3.5 *Assume that we rewrite $\mathcal{S}_{4,i}$ as $\mathcal{S}_{4,i} = \mathcal{S}_{4,i}^- + \mathcal{S}_{4,i}^{+,1} + \mathcal{S}_{4,i}^{+,2}$, where $\mathcal{S}_{4,i}^{+,1}$ is the sum of all summands of $\mathcal{S}_{4,i}$ for which $j < i < k < l$, and $\mathcal{S}_{4,i}^{+,2}$ is the sum of all summands of $\mathcal{S}_{4,i}$ for which $i > j$, $i > k$ and $i > l$. Then we have $\mathcal{S}_{4,i}^- \leq 0$, $\mathcal{S}_{4,i}^{+,1} \geq 0$ and $\mathcal{S}_{4,i}^{+,2} \geq 0$, and the following inequalities hold:*

$$\mathcal{S}_{4,i}^{+,1} \leq \frac{d_i^2 \tilde{\beta}_0 \beta_3 \phi}{e^2 d_N}, \quad \mathcal{S}_{4,i}^{+,2} \leq \frac{4 \tilde{\beta}_0 \beta_1 \beta_2 \beta_3 d_i^2}{27 d_N^4} \phi^3. \quad (5.3.30)$$

Proof. Utilizing the inequality (5.3.20) and performing cancellations and using the ordering $d_j > d_i > d_l > d_k$ in the summands, we get

$$\begin{aligned} \mathcal{S}_{4,i}^{+,1} &\leq \frac{\tilde{\beta}_0 \beta_3}{e^2 \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \frac{\phi_j \phi_k \phi_l \pi_{ijkl}^2 d_i d_k d_i^2 d_l^2}{(d_j^2 - d_i^2)(d_k^2 - d_i^2)(d_l^2 - d_i^2) d_i d_j d_k d_l (d_k - d_i)(d_l^2 - d_i^2)} \\ &\leq \frac{\tilde{\beta}_0 \beta_3 d_i^2}{e^2 \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \frac{\phi_j \phi_k \phi_l d_l (d_j - d_i)(d_l - d_j)^2 (d_l - d_k)^2 (d_k - d_j)^2}{d_j (d_j + d_i)(d_k + d_i)(d_l + d_i)^2} \\ &\leq \frac{\tilde{\beta}_0 \beta_3 d_i^2}{e^2 d_N \mathbf{d}_1^T \Phi \mathbf{d}_2^T \Phi} \sum_{\substack{j,k,l=1 \\ j < i < k < l}}^N \phi_j d_j^2 \phi_k \phi_l d_l, \end{aligned}$$

which implies the first inequality in (5.3.30). Next, we employ the fact that $\tilde{\eta}_j / (d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i) \leq (d_j^2 - d_i^2)^{-1}$ when $i > j$ for calculating that

$$\begin{aligned} \mathcal{S}_{4,i}^{+,2} &= \sum_{\substack{j,k,l=1 \\ j < k < l}}^{i-1} \frac{\phi_j \phi_k \phi_l \tilde{\eta}_j \tilde{\eta}_k \tilde{\eta}_l \pi_{ijkl}^2 \tilde{\beta}_0 \beta_1 \beta_2 \beta_3}{(d_j^2 \tilde{\eta}_j - d_i^2 \tilde{\eta}_i)(d_k^2 \tilde{\eta}_k - d_i^2 \tilde{\eta}_i)(d_l^2 \tilde{\eta}_l - d_i^2 \tilde{\eta}_i) d_i d_j d_k d_l} \\ &\leq \sum_{\substack{j,k,l=1 \\ j < k < l}}^{i-1} \frac{\phi_j \phi_k \phi_l \pi_{ijkl}^2 \tilde{\beta}_0 \beta_1 \beta_2 \beta_3}{(d_j^2 - d_i^2)(d_k^2 - d_i^2)(d_l^2 - d_i^2) d_i d_j d_k d_l} \\ &\leq \sum_{\substack{j,k,l=1 \\ j < k < l}}^{i-1} \frac{\phi_j \phi_k \phi_l (d_j - d_k)^2 (d_j - d_l)(d_k - d_l) \tilde{\beta}_0 \beta_1 \beta_2 \beta_3}{d_i d_l} \\ &\leq \frac{4}{27} \frac{\tilde{\beta}_0 \beta_1 \beta_2 \beta_3 d_i^2}{d_N^4} \sum_{\substack{j,k,l=1 \\ j < k < l}}^{i-1} \phi_j \phi_k \phi_l, \end{aligned}$$

where the factor $4/27$ comes from a discussion of the maximum of the function $(d_j, d_j, d_l) \mapsto (d_j - d_k)^2 (d_j - d_l)(d_k - d_l)$ for $1 \geq d_j > d_k > d_l > 0$. This proves the second inequality in (5.3.30). \square

Summarizing, and collecting the inequalities for the various terms, we see that

$$\begin{aligned} \mathcal{S}_{1,i} + \mathcal{S}_{2,i} + \mathcal{S}_{3,i} + \mathcal{S}_{4,i} &= \mathcal{S}_{1,i} + \mathcal{S}_{2,i}^- + \mathcal{S}_{2,i}^+ + \mathcal{S}_{3,i}^- + \mathcal{S}_{3,i}^{+,1} + \mathcal{S}_{3,i}^{+,2} + \mathcal{S}_{4,i}^- + \mathcal{S}_{4,i}^{+,1} + \mathcal{S}_{4,i}^{+,2} \\ &< \mathcal{S}_{1,i} + \mathcal{S}_{2,i}^+ + \mathcal{S}_{2,i}^- + \mathcal{S}_{3,i}^{+,2} + \mathcal{S}_{3,i}^{+,1} + \mathcal{S}_{4,i}^{+,1} + \mathcal{S}_{4,i}^{+,2} \\ &\leq d_i^2 M(\phi, \boldsymbol{\beta}, d_N), \end{aligned} \tag{5.3.31}$$

where we define the function

$$\begin{aligned} M(\phi, \boldsymbol{\beta}, d_N) &:= \left[1 + \frac{4\beta_1 \beta_2 \beta_3 \phi^3}{27 d_N^4} - e^{-1} \left(1 + \frac{3\beta_3}{d_N \beta_2} \right) - e^{-2} \left(1 + \frac{3\beta_3}{2\beta_1} + \frac{3\beta_3}{2\beta_2} - \frac{\phi \beta_3}{d_N} \right) \right] \tilde{\beta}_0 \\ &\quad + (1 - e^{-1})\beta_1 + (1 - e^{-1})\beta_2 + (1 - e^{-1} - e^{-2})\beta_3. \end{aligned} \tag{5.3.32}$$

Instead of employing the criterion (5.3.23) which is not practical (for a large number of species N) as we analyzed in Chapter 3 we give a sufficient condition for (5.3.23) holds for all pairs $j < i$. Fix a pair $i > j$, define $\delta := \delta_{ij} = d_i/d_j$, and divide (5.3.24) by d_j^2 to obtain

$$\begin{aligned} \tilde{H}_{ij} = & -\tilde{\beta}_0(\beta_1\delta + \beta_2(1 + \delta)^2 + \beta_3(1 + \delta + \delta^2)(1 + \delta + \delta^{-1})) - (\beta_2(\beta_1 + \beta_3)\delta + \beta_1\beta_3(1 + \delta)^2) \\ & - \phi(1 - \delta)^2\{\tilde{\beta}_0[\beta_1\beta_2 + (\beta_1 + \beta_2)\beta_3(1 + 2\delta^{-1})] + \beta_1\beta_2\beta_3\}. \end{aligned} \quad (5.3.33)$$

Since $\delta \in (d_N, 1]$ a sufficient condition for (5.3.23) to be satisfied is given by

$$\forall \phi \in [0, \phi_{\max}] : \quad H(\phi, \boldsymbol{\beta}, d_N) < 0, \quad (5.3.34)$$

where the following definition of $H(\phi, \boldsymbol{\beta}, d_N)$ is derived from the observation that the two terms in the first line of (5.3.33) are non-positive, while the term in the second line is non-negative:

$$\begin{aligned} H(\phi, \boldsymbol{\beta}, d_N) := & -\tilde{\beta}_0(\beta_1 d_N + \beta_2(1 + d_N)^2 + \beta_3(1 + d_N + d_N^2)(2 + d_N)) \\ & - (\beta_2(\beta_1 + \beta_3)d_N + \beta_1\beta_3(1 + d_N)^2) \\ & - \phi(1 - d_N)^2\{\tilde{\beta}_0[\beta_1\beta_2 + (\beta_1 + \beta_2)\beta_3(1 + 2d_N^{-1})] + \beta_1\beta_2\beta_3\}. \end{aligned} \quad (5.3.35)$$

Theorem 5.3.6 *Assume that the vector of parameters $\boldsymbol{\beta}$, the maximum solids concentration ϕ_{\max} and the width of the particle size distribution, characterized by the value of $d_N \in (0, 1]$, are chosen such that the inequality (5.3.34) is satisfied, where the expression $H(\phi, \boldsymbol{\beta}, d_N)$ is defined by (5.3.35), and that*

$$\forall \phi \in [0, \phi_{\max}] : \quad M(\phi, \boldsymbol{\beta}, d_N) < 0, \quad (5.3.36)$$

where the function $M(\phi, \boldsymbol{\beta}, d_N)$ is defined in (5.3.32). Then $\gamma_i < 0$ for $i = 1, \dots, N$, i.e., the model equations are strictly hyperbolic for $\Phi \in \mathcal{D}_{\phi_{\max}}$.

As an example of the case $\beta_3 < 0$, consider the parameter vectors $\boldsymbol{\beta}$ given by (3.3.12); let us focus first on the case of large Péclet numbers. Figure 5.5 (left) shows in a ϕ versus d_N plot the curves $H(\phi, \boldsymbol{\beta}, d_N) = 0$ and $M(\phi, \boldsymbol{\beta}, d_N) = 0$. The region $H(\dots) < 0, M(\dots) < 0$, where the model is strictly hyperbolic, is located to the right of the curve $M(\phi, \boldsymbol{\beta}, d_N) = 0$. Here we employ a logarithmic scale since the term d_N^{-1} in (5.3.35) becomes singular. Solving $M(1, \boldsymbol{\beta}, d_N) = 0$ for d_N yields here that $M(1, \boldsymbol{\beta}, d_N) < 0$ for $d_N > d_N^* := 0.164092$, which means that for these values of d_N , the HS model is strictly hyperbolic on $\mathcal{D}_{\phi_{\max}}$ for all $\phi_{\max} \in (0, 1]$. The behaviour of the curve $M(\phi, \boldsymbol{\beta}, d_N) = 0$ indicates that this property remains valid for slightly smaller values of d_N provided that ϕ_{\max} is chosen sufficiently small.

For the parameters given by (3.3.12) for the case of small Péclet numbers, the behaviour is similar, as can be seen from Figure 5.5 (right), but the hyperbolicity region is smaller. We obtain unconditional hyperbolicity for $d_N > d_N^* := 0.328981$; this number is the solution of $M(1, \boldsymbol{\beta}, d_N) = 0$.

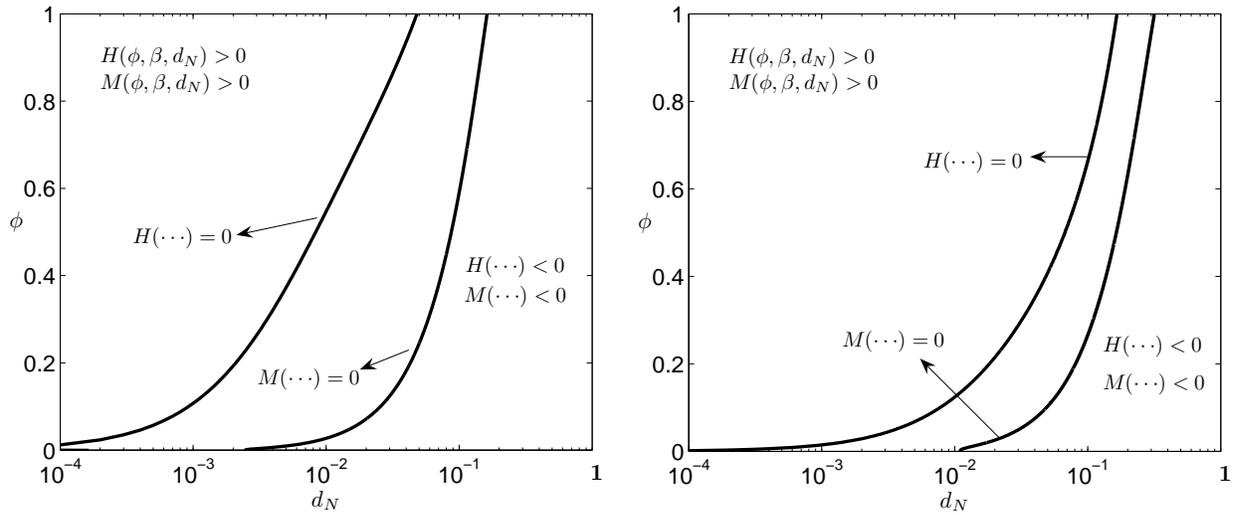


Figure 5.5: Regions of hyperbolicity ($H(\phi, \boldsymbol{\beta}, d_N) < 0$ and $M(\phi, \boldsymbol{\beta}, d_N) < 0$) for the HS model: with the coefficients (3.3.12) for large Péclet numbers (left), and with the coefficients (3.3.12) for small Péclet numbers(right).

Chapter 6

Conclusiones generales y trabajo futuro

6.1 Conclusiones

En los capítulos precedentes de este trabajo se ha utilizado un resultado del álgebra lineal (*la ecuación secular*) como una herramienta para estudiar la hiperbolicidad estricta de algunos modelos de sedimentación de suspensiones polidispersas, mas concretamente se dan condiciones suficientes que permiten determinar el tamaño de la subregión del conjunto de soluciones admisibles que garantizan estabilidad de los modelos independientemente del número de especies y considerando sólo parámetros que se pueden controlar en situaciones prácticas. Estos resultados teóricos están en concordancia con resultados obtenidos previamente en la literatura, los cuales están en muchos casos limitados a un número pequeño de especies. Los resultados teóricos obtenidos relativos a la hiperbolicidad han sido empleados en el diseño e implementación de esquemas numéricos conocidos que precisan la utilización intensiva de la información característica del sistema de leyes de conservación. Dado que para los modelos de sedimentación, tal información no está disponible en término de expresiones cerradas (como usualmente se tiene para problemas de dinámica de gases) cabe resaltar que en este trabajo se obtiene una muy buena aproximación numérica de dicha información a un costo razonable a partir de la muy útil propiedad de entrelazamiento. La calidad de la información característica se ve reflejada al comparar el comportamiento no oscilatorio y no difusivo de los esquemas numéricos que la incorporan, con aquellos esquemas que sólo usan parcialmente tal información espectral.

6.2 Limitaciones del enfoque de este trabajo

- El desarrollo teórico del análisis de hiperbolicidad mediante el uso de la ecuación secular, depende fuertemente del hecho de que las velocidades v_i puedan escribirse en la forma (3.1.2), esto implica, que el análisis expuesto no es universalmente aplicable. De hecho, como se comenta en las conclusiones del Capítulo 3 no parece claro en principio como aplicar este procedimiento a modelos como el presentado por Salim *et al.* [58] donde la velocidades dependen de sumas parciales de las concentraciones, en vez de la concentración total.
- Desde el punto de vista computacional, si bien es cierto que el uso de información característica redundante en esquemas robustos y eficientes, debe notarse que aquí también se tiene en cuenta la forma particular de la función velocidad (para aproximar máximos de dichas funciones), mas concretamente, en el cálculo del coeficiente de viscosidad (local) para el esquema SPEC-INT (ver ecuación (4.3.16)). Sin embargo, esto no constituye un obstáculo insoslayable, pues el cálculo del coeficiente de viscosidad evaluando el máximo de v_i en los extremos de la celda correspondiente produce también buenos resultados numéricos.

6.3 Trabajo futuro

- Estudiar la hiperbolicidad para los modelos involucran los coeficientes de Batchelor, para el caso en que el coeficientes β_3 es positivo y relativamente pequeño, situación que se considera en la literatura.
- Analizar el modelo de Parwardhan-Tien (PT) [49], el cual constituye una generalización y es más elaborado que el modelo MLB. El modelo PT incorpora otras cantidades en la funciones de velocidad, por ejemplo, la concentración local de partículas. La forma de las velocidades v_i para dicho modelo eventualmente permite aplicar las técnicas consideradas en este trabajo.
- Considerar en los trabajos preliminares mencionados en cada ítem anterior, métodos de discretización temporal de orden mayor que tres y a la vez eficientes desde el punto de vista computacional, descritos por ejemplo en las referencias [29, 30, 55].

Bibliography

- [1] J. Anderson, *A secular equation for the eigenvalues of a diagonal matrix perturbation*, Lin. Alg. Appl., **246** (1996), 49–70.
- [2] M. Bargiel and E.-M. Tory, *Stability of tridisperse suspensions*, Comput. Vis. Sci., **10** (2007), 163–170.
- [3] 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), 1815–1835.
- [4] G.-K. Batchelor, *Sedimentation in a dilute polydisperse system of interacting spheres. Part 1. General theory*, J. Fluid Mech., **119** (1982), 379–408.
- [5] G.-K. Batchelor and R.-W. Janse van Rensburg, *Structure formation in bidisperse sedimentation*, J. Fluid Mech., **166** (1986), 379–407.
- [6] G.-K. Batchelor and C.-S. Wen, *Sedimentation in a dilute polydisperse system of interacting spheres. Part2. Numerical results*, J. Fluid Mech., **124** (1982), 495–528.
- [7] S. Benzoni-Gavage and R.-M. Colombo, *An n -populations model for traffic flow*, Eur. J. Appl. Math., **14** (2003), 587–612.
- [8] S. Berres and R. Bürger *On the settling of a bidisperse suspension with particles having different sizes and densities*, Acta Mech., **201** (2008), 47–62.
- [9] S. Berres and R. Bürger, *On Riemann problems and front tracking for a model of sedimentation of polydisperse suspensions*, ZAMM Z. Angew. Math. Mech., **87** (2007), 665–691.

- [10] S. Berres, R. Bürger and K.-H. Karlsen, *Central schemes and systems of conservation laws with discontinuous coefficients modeling gravity separation of polydisperse suspensions*, J. Comput. Appl. Math., **164-165** (2004), 53–80.
- [11] 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.
- [12] P.-M. Biesheuvel, H. Verweij and V. Breedveld, *Evaluation of instability criterion for bidisperse sedimentation*, AIChE J., **47** (2001), 45–52.
- [13] R. Bürger, F. Concha, K.-K. Fjelde and K.-H. Karlsen, *Numerical simulation of the settling of polydisperse suspensions of spheres*, Powder Technol., **113** (2000) 30–54.
- [14] 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), 2186–2213.
- [15] R. Bürger, K.-K. Fjelde, K. Höfler and K.H. Karlsen, *Central difference solutions of the kinematic model of settling of polydisperse suspensions and three-dimensional particle-scale simulations*, J. Engrg. Math., **41** (2001), 167–187.
- [16] R. Bürger, A. García and M. Kunik, *A generalized kinetic model of sedimentation of polydisperse suspensions with a continuous particle size distribution*, Math. Models Methods Appl. Sci., **18** (2008), 1741–1785.
- [17] 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), 699–722.
- [18] R. Bürger, A. García, K.-H. Karlsen and J.-D. Towers, *A family of numerical schemes for kinematic flows with discontinuous flux*, J. Engrg. Math., **60** (2008), 387–425.
- [19] R. Bürger and A. Kozakevicius, *Adaptive multiresolution WENO schemes for multi-species kinematic flow models*, J. Comput. Phys., **224** (2007), 1190–1222.
- [20] A. Cauchy, *Mécanique céleste. Sur le mouvement de notre système planétaire. In: Œuvres Complètes d’Augustin Cauchy, Ière Sér.*, Gauthier-Villars, Paris, vol. 5, no. 98 (1885), pp. 331–341.

-
- [21] A. Cuthbertson, P. Dong, S. King and P. Davies, *Hindered settling velocity of cohesive/non-cohesive sediment mixtures*, Coastal Engrg., **55** (2008), 1197–1208.
- [22] C.-M. Dafermos, *Hyperbolic Conservation Laws in Continuum Physics, Third Ed.*, Springer Verlag, Berlin (2010).
- [23] R.-H. Davis and H. Gecol, *Hindered settling function with no empirical parameters for polydisperse suspensions*, AIChE J., **40** (1994), 570–575.
- [24] R.-H. Davis and H. Gecol, *Classification of concentrated suspensions using inclined settlers*, Int. J. Multiphase Flow, **22** (1996), 563–574.
- [25] S. Diehl, *Estimation of the batch-settling flux function for an ideal suspension from only two experiments*, Chem. Engrg. Sci., **62** (2007), 4589–4601.
- [26] R. Donat and P. Mulet, *A secular equation for the Jacobian matrix of certain multi-species kinematic flow models*, Numer. Methods Partial Differential Equations, **26** (2010), 159–175.
- [27] R. Donat and P. Mulet, *Characteristic-based schemes for multi-class Lighthill-Whitham-Richards traffic models*, J. Sci. Comput., **37** (2008), 233–250.
- [28] R. Fedkiw, B. Merriman and S. Osher, *Simplified upwind discretization of systems of hyperbolic conservation laws containing advection equations*, J. Comput. Phys., **157** (2000), 302–326.
- [29] S. Gottlieb, D.-I. Ketcheson and C.-W. Shu, *High order strong stability preserving time discretization*, J. Sci. Comput., **38** (2009), 251–289.
- [30] S. Gottlieb and S. Ruuth, *Optimal strong-stability-preserving time-stepping schemes with fast downwind spatial discretizations*, J. Sci. Comput., **27** (2006) 289–303.
- [31] S. Gottlieb and C.-W. Shu, *Total variation diminishing Runge-Kutta schemes*, Math. Comp., **67** (1998), 73–85.
- [32] S. Gottlieb, C.-W. Shu and E. Tadmor, *Strong stability-preserving high-order time discretization methods*, SIAM Rev., **43** (2001), 89–112.
- [33] G.-H. Golub, *Some modified matrix eigenvalue problems*, SIAM Rev., **15** (1973), 318–334.

- [34] M. Gray, Z. Xu and J. Masliyah, *Physics in the oil sands of Alberta*, Physics Today, **62**(3) (2009), 31–35.
- [35] H.-P. Greenspan and M. Ungarish, *On hindered settling of particles of different sizes*, Int. J. Multiphase Flow, **8** (1982), 587–604.
- [36] Z. Ha and S. Liu, *Settling velocities of polydisperse concentrated suspensions*, Canad. J. Chem. Engrg., **80** (2002), 783–790.
- [37] K. Höfler, *Simulation and Modeling of Mono- and Bidisperse Suspensions*, Doctoral Thesis, Institut für Computeranwendungen, Universität Stuttgart, Germany, 2000.
- [38] K. Höfler and S. Schwarzer, *The structure of bidisperse suspensions at low Reynolds numbers*. In: A.M. Sändig, W. Schiehlen, and W.L. Wendland (eds.), *Multifield Problems: State of the Art*, Springer Verlag, Berlin (2000), 42–49.
- [39] G.-S. Jiang and C.-W. Shu, *Efficient implementation of weighted ENO schemes*, J. Comput. Phys., **126** (1996), 202–228.
- [40] A.-K. Henrick, T.-D. Aslam and J.-M. Powers, *Mapped weighted essentially non-oscillatory schemes: Achieving optimal order near critical points*, J. Comput. Phys., **207** (2005), 542–567.
- [41] A. Kurganov and E. Tadmor, *New high-resolution central schemes for nonlinear conservation laws and convection-diffusion equations*, J. Comput. Phys., **160** (2000), 241–282.
- [42] G.J. Kynch, *A theory of sedimentation*. Trans. Faraday Soc., **48**, (1952), 166–176.
- [43] R.-J. LeVeque, *Finite Volume Methods for Hyperbolic Problems*, Cambridge University Press, Cambridge, UK (2002).
- [44] X.-D. Liu, S. Osher and T. Chan, *Weighted essentially non-oscillatory schemes*, J. Comput. Phys., **115** (1994), 200–212.
- [45] M.-J. Lockett and K.-S. Bassoon, *Sedimentation of binary particle mixtures*, Powder Technol., **24** (1979), 1–7.
- [46] J.-H. Masliyah, *Hindered settling in a multiple-species particle system*, Chem. Engrg. Sci., **34** (1979), 1166–1168.

- [47] H. Nasr-el-Din, J.-H. Masliyah and K. Nandakumar, *Continuous separation of suspensions containing light and heavy particle species*, *Canad. J. Chem. Engrg.*, **77** (1999), 1003–1012.
- [48] S. Osher and R. Fedkiw, *Level Set Methods and Dynamic Implicit Surfaces*, Springer, New York, 2003.
- [49] V.-S. Patwardhan and C. Tien, *Sedimentation and liquid fluidization of solid particles of different sizes and densities*, *Chem. Engrg. Sci.*, **40** (1985), 1051–1060.
- [50] A.-A.-I. Peer, M.-Z. Dauhoo and M. Bhuruth, *A method for improving the performance of the WENO5 scheme near discontinuities*, *Appl. Math. Lett.*, **22** (2009), 1730–1733.
- [51] S. Qian, R. Bürger and H.-H. Bau, *Analysis of sedimentation biodetectors*, *Chem. Engrg. Sci.*, **60** (2005), 2585–2598.
- [52] J.-F. Richardson and W.-N. Zaki, *Sedimentation and fluidization: Part I*, *Trans. Instn. Chem. Engrs. (London)*, **32** (1954), 35–53.
- [53] F. Rosso and G. Sona, *Gravity-driven separation of oil-water dispersions*, *Adv. Math. Sci. Appl.*, **11** (2001), 127–151.
- [54] W.-B. Russel, D.-A. Saville and W.-R. Schowalter, *Colloidal Dispersions*, Cambridge University Press, Cambridge, UK (1989).
- [55] S. Ruuth, *Global optimization of explicit strong-stability-preserving Runge-Kutta methods*, *Math. Comput.*, **75** (2006), 183–207.
- [56] S.-J. Ruuth and R.-J. Spiteri, *Two barriers on strong-stability-preserving time discretization methods*, *J. Sci. Comput.*, **17** (2002), 211–220.
- [57] W. Schneider, G. Anestis and U. Schafflinger, *Sediment composition due to settling of particles of different sizes*, *Int. J. Multiphase Flow*, **11** (1985), 419–423.
- [58] M.-S. Selim, A.-C. Kothari and R.-M. Turian, *Sedimentation of multisized particles in concentrated suspensions*, *AIChE J.*, **29** (1983), 1029–1038.
- [59] P.-T. Shannon, E. Stroupe and E.-M. Tory, *Batch and continuous thickening*, *Ind. Engrg. Chem. Fund.*, **2** (1963), 203–211.

- [60] C.-W. Shu, *High order weighted essentially nonoscillatory schemes for convection dominated problems*, SIAM Rev., **51** (2009), 82–126.
- [61] C.-W. Shu, *Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws*. In: B. Cockburn, C. Johnson, C.-W. Shu and E. Tadmor, *Advanced Numerical Approximation of Nonlinear Hyperbolic Equations* (A. Quarteroni, Ed.), Lecture Notes in Mathematics vol. 1697, Springer-Verlag, Berlin (1998), 325–432.
- [62] C.-W. Shu and S. Osher, *Efficient implementation of essentially non-oscillatory shock capturing schemes*, J. Comput. Phys., **77** (1988), 439–471.
- [63] R. Simura and K. Ozawa, *Mechanism of crystal redistribution in a sheet-like magma body: Constraints from the Nosappumisaki and other Shoshonite intrusions in the Nemuro peninsula, Northern Japan*, J. Petrology, **47** (2006), 1809–1851.
- [64] R.-H. Weiland, Y.-P. Fessas and B.-V. Ramarao, *On instabilities arising during sedimentation of two-component mixtures of solids*, J. Fluid Mech., **142** (1984), 383–389.
- [65] C.-S. Wen, *The Fundamentals of Aerosol Dynamics*. World Scientific, Singapore (1996).
- [66] G.-C.-K. Wong and S.-C.-K. Wong, *A multi-class traffic flow model—an extension of LWR model with heterogeneous drivers*, Transp. Res. A, **36** (2002), 827–841.
- [67] Z. Xu and C.-W. Shu, *Anti-diffusive flux corrections for high order finite difference WENO schemes*, J. Comput. Phys., **205** (2005), 458–485.
- [68] B. Xue and Y. Sun, *Modeling of sedimentation of polydisperse spherical beads with a broad size distribution*, Chem. Engrg. Sci., **58** (2003), 1531–1543.
- [69] A. Zeidan, S. Rohani and A. Bassi, *Dynamic and steady-state sedimentation of polydisperse suspension and prediction of outlets particle-size distribution*, Chem. Engrg. Sci., **59** (2004), 2619–2632.
- [70] P. Zhang, R.-X. Liu, S.-C. Wong and S.-Q. Dai, *Hyperbolicity and kinematic waves of a class of multi-population partial differential equations*, Eur. J. Appl. Math., **17** (2006), 171–200.

-
- [71] M. Zhang, C.-W. Shu, G.-C.-K. Wong and S.-C. Wong, *A weighted essentially non-oscillatory numerical scheme for a multi-class Lighthill-Whitham-Richards traffic flow model*, J. Comput. Phys., **191** (2003), 639–659.
- [72] P. Zhang, S.-C. Wong and S.-Q. Dai, *A note on the weighted essentially non-oscillatory numerical scheme for a multi-class Lighthill-Whitham-Richards traffic flow model*, Comm. Numer. Meth. Engrg., **25** (2009), 1120–1126.
- [73] P. Zhang, S.-C. Wong and C.-W. Shu, *A weighted essentially non-oscillatory numerical scheme for a multi-class traffic flow model on an inhomogeneous highway*, J. Comput. Phys., **212** (2006), 739–756.