Numerical simulation of chemical kinetics, transport and flow processes

by

Emmanuel Adoliwine Amikiya

Submitted in partial fulfillment of the requirements for the degree

Philosophiae Doctor

In the Department of Mathematics and Applied Mathematics

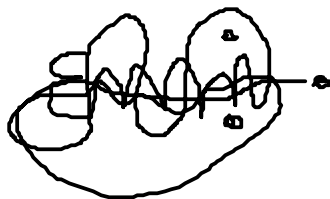In the Faculty of Natural and Agricultural Sciences

University of Pretoria

Pretoria

January 29, 2020

Declaration

I, Amikiya, Emmanuel Adoliwine declare that the thesis, which I hereby submit for the degree Doctor of Philosophy at the University of Pretoria, is my own work and has not previously been submitted by me for a degree at this or any other tertiary institution.

SIGNATURE:

DATE:

FEBRUARY 3, 2020

Dedication

I dedicate this thesis to my late father, Akampie Sylvester Amikiya.

Acknowledgement

Title   Numerical simulation of chemical kinetics, transport and flow processes

Name   Emmanuel Adoliwine Amikiya

Supervisor   Prof. Mapundi K. Banda

Department   Mathematics & Applied Mathematics

Degree   Doctor of Philosophy

Abstract

In this thesis, numerical solution procedures are developed for simulating chemical phenomena. Mathematical models for phenomena involving flow, transport and reaction of chemical species are computationally challenging to simulate due to stiffness, high degrees of freedom and spatial dependence. Such challenges are resolved (in this thesis) by combining model decoupling techniques with compatible efficient numerical schemes. Chemical phenomena is decomposed into well-mixed chemical systems, poorly-mixed systems (or spatial dependent kinetics) and flow with reactive transport systems. Mathematical models for the systems are Ordinary Differential Equations (ODEs), parabolic Partial Differential Equations (PDEs) and hyperbolic PDEs, respectively. In the ODE model, stiffness is resolved by positivity-preserving implicit schemes while the large degrees of freedom is reduced by stoichiometric and continuous-time iteration methods. In the parabolic model, model decoupling techniques are employed to reduce the degrees of freedom while Implicit-Explicit numerical schemes are presented for resolving stiffness. Further, numerical schemes that have dispersion-dissipation-preserving properties have also been discussed. In the hyperbolic model, model decoupling techniques have been presented for reducing the degrees of freedom while shock-capturing, well-balanced numerical schemes have been presented for resolving nonlinear hyperbolic effects. The results from experiments show that the proposed numerical solution procedures can efficiently resolve the challenges in simulating chemical phenomena.

Contents

List of Figures

x

List of Tables

xiv

Chapter 1

Background to chemical reaction phenomena

Chemical equilibrium and chemical kinetics are the two main interesting research themes in chemical reaction phenomena and have received considerable attention over many decades. Chemical kinetics is a research discipline that tries to quantify the time-dependent dynamics of a chemical system before equilibrium is established [37]. Chemical reaction processes can be found in several scientific disciplines (including Environmental Science, Pharmacology, Biology, Biomedical Sciences and Medical Science [81]), thus, chemical kinetic problems have received considerable attention from scientists with different backgrounds.

Mathematical modelling has been a corner stone in the development of chemical kinetic theory, for more than two centuries. In 1850, a German scientist published an article on hydrolysis of sucrose, in which a mathematical model was used to adequately establish a link between thermodynamics and chemical reactions [37]. Differential models have been employed to accurately describe the dynamics of simple chemical systems such as bimolecular reactions. Exponential and polynomial models have aided in the development of collision theory, transition state theory, Arrhenius equation, mass action law and many more significant theories in chemical reactions [37].

Over the past two centuries, the chemical systems that have been considered in kinetic studies have been relatively simple, thus, accurate models were simple and could easily be analysed without computers. However, the systems in modern times are complex and

large (e.g. refinement of petroleum involves thousands of molecules engaged in thousands of reactions [37, 95]). Consequently, mathematical modelling of modern chemical systems is challenging. Research questions can still be answered by means of computer simulation, however, the approach is constrained by scarcity of resources (e.g. memory). Thus, it is necessary to develop efficient computer simulation approaches for chemical phenomena. Generally, kinetic studies can be classified into detailed and reduction kinetic approaches (that encompass decomposition and lumping approaches).

Decomposition approaches include the Quasi-Steady-State (QSS) approximation [37, 125, 149], Partial Equilibrium approximation (PEA) [149, 46, 124, 70, 68], rate-determining [37], Computational Singular Perturbation (CSP) [125, 91, 93] and others can be found in [172, 47, 55, 131, 92, 69, 80, 152]. Decomposition approaches are developed based on a fact that intermediate reactions proceed faster than overall reactions, thus, resulting in a wide time-scale variation within the system [37, 125, 124]. Consequently, the system is decomposed into fast, medium and slow sub-systems where further assumptions are employed to obtain smaller models. Sensitivity analysis has been employed to distinguish the fast groups from the slow ones, some research [125, 27, 136, 51] have reported success while others reported short falls [125, 1, 22, 40]. With decomposition approaches, only few parameters are needed to build models and simulation is less expensive. However, the approaches oversimplify the kinetics to an extent that accuracy is lost and the models are not consistent with the reality. Other properties of the decomposition methods have been discussed in [125, 104, 165, 58, 59].

Lumping approaches are developed based on the fact that there are several species or reactions (within the large chemical system) that have a common property. Those individual reactions/species that have the common property are grouped together and studied as one component called a *lump* [37]. All model reduction techniques (including lumping techniques) are not exact since the resulting models do not describe the detailed kinetics.

The detailed kinetic approaches are developed to study the time-dependent behaviour of the entire chemical system. Resulting models are non-linear differential equations that characteristically lack analytical solutions, except for a few simple chemical systems. As a result, numerical schemes are employed to approximate or simulate the solution.

The detailed kinetic approaches are consistent and accurate but are however, expensive (sometimes infeasible) to implement.

The focus of this thesis is on simulating the kinetics of chemical reactions (i.e. simulation of the transient aspect of chemical reaction phenomena). The law of mass action and other physical laws are applied to derive differential equations that model the kinetics of chemical reactions. The resulting differential equations are nonlinear, stiff and have high degrees of freedom [119, 188, 120, 146]. Due to lack of exact solutions in the general case, numerical schemes are developed as alternative methods of solution. However, combining accurate numerical schemes with the already high degrees of freedom yields an expensive simulation. This thesis combines model decoupling methods with compatible accurate numerical schemes, for the purpose of efficiently simulating chemical kinetic phenomena.

Numerical simulation of the kinetics of three groups of chemical reaction systems will be considered in this study. The first group consists of chemical reactions that occur in well-stirred (also known as well-mixed) environments. Time is the only independent variable of such systems. The second group (called reactive transport systems or poorly-mixed systems) consists of chemical reactions that occur alongside physical processes such as diffusion and advection. The third group (called flow and reactive systems) consists of chemical systems that occur in a fluid flow environment (specifically, shallow water flow environment). The second and third groups are spatially-dependent in addition to the time-dependence.

The rest of the thesis is organised into six chapters. In Chapter 2, modelling of reactive flow systems is presented and analysed. In Chapter 3, modelling and simulation of well-stirred chemical systems is discussed. In Chapter 4, modelling and simulation of reactive transport systems is discussed. Dispersion preserving schemes for reactive transport systems are discussed in Chapter 5. Numerical simulation of flow and reactive systems is discussed in Chapter 6, followed by the conclusion in Chapter 7.

Chapter 2

Preliminary: reactive flow modelling

In this chapter, the kinetics of flow and reactive transport processes are modelled and analysed, using physical conservation laws and the shallow water principle. Properties such as hyperbolicity and nonlinear effects of the derived shallow water model are presented. Riemann problems and elementary wave solutions have also been discussed. Riemann problems of Euler's models (in gas dynamics) and shallow water model (in hydrodynamics) have been discussed in literature. The discussion here, on the Riemann problem is an application of known results in [100, 161], thus, the results are not novel.

2.1   Shallow water flow model

In this section, a mathematical model is derived using physico-chemical laws and shallow water principle. A rectangular reference frame is considered, where the x and y axes frame the horizontal plane and z axis is vertical. The three dimensional spatial differential operator is denoted by $\nabla = \left(\frac{\partial}{\partial x}, \frac{\partial}{\partial y}, \frac{\partial}{\partial z}\right)$ and the velocity vector is denoted by $\mathbf{v}_e = (u_e, v_e, w_e)$.

The governing equations for fluid flow and associated phenomena are derived from a physical principle, which says that the rate of change of an extensive property of a fluid occupying a volume, is the balance between fluxes and the rate of creation or destruction of the property in the volume [168, 43, 76]. In a smooth domain, the principle is a

4

differential equation that can be written as follows:

$$\frac{\partial \Phi}{\partial t} + \nabla \cdot \mathbf{J} - S_\Phi = 0, \tag{2.1}$$

where Φ is density of the extensive property, $\mathbf{J}$ is a flux vector and S_Φ is the source/sink term for the extensive property. If the mass of the carrier fluid (with density ρ_e and fluxes $\mathbf{J} = \rho_e \mathbf{v_e}$ flowing without sources or sinks) is considered as the extensive property, the conservation principle (2.1) yields [76]:

$$\frac{\partial \rho_e}{\partial t} + \nabla \cdot \rho_e \mathbf{v_e} = 0, \tag{2.2}$$

known as the continuity equation. If momentum is the extensive property of the fluid, the density $\Phi = \rho_e \mathbf{v}_e$ quantifies momentum density, the flux is a balance of convective processes, fluid pressure (denoted p_e) and stresses ($\mathcal{T}$) (i.e. $\mathbf{J} = \rho_e \mathbf{v_e} \mathbf{v_e} + p_e \delta - \mathcal{T}$ where δ is a unit tensor) and the source terms are body forces such as gravity (i.e. $S_\Phi = \rho_e \mathbf{g}$, $\mathbf{g} = (0, 0, -g)$ and g is the acceleration due to gravity) [168, 43]. Thus, balancing momentum in an elementary volume (i.e. applying principle (2.1)) yields:

$$\frac{\partial \rho_e \mathbf{v_e}}{\partial t} + \nabla \cdot \rho_e \mathbf{v_e} \mathbf{v_e} = -\nabla p_e + \nabla \cdot \mathcal{T} + \rho_e \mathbf{g}. \tag{2.3}$$

The discussion here is limited to a Newtonian fluid that is inviscid, incompressible and has constant density. Under such flow conditions, the continuity and momentum equations (2.2) and (2.3) reduce in terms, to the incompressible Euler model [168, 76]:

$$\frac{\partial u_e}{\partial x} + \frac{\partial v_e}{\partial y} + \frac{\partial w_e}{\partial z} = 0, \tag{2.4}$$

$$\frac{\partial u_e}{\partial t} + \frac{\partial u_e^2}{\partial x} + \frac{\partial u_e v_e}{\partial y} + \frac{\partial w_e u_e}{\partial z} = -\frac{1}{\rho_e} \frac{\partial p_e}{\partial x}, \tag{2.5}$$

$$\frac{\partial v_e}{\partial t} + \frac{\partial u_e v_e}{\partial x} + \frac{\partial v_e^2}{\partial y} + \frac{\partial v_e w_e}{\partial z} = -\frac{1}{\rho_e} \frac{\partial p_e}{\partial y}, \tag{2.6}$$

$$\frac{\partial w_e}{\partial t} + \frac{\partial u_e w_e}{\partial x} + \frac{\partial v_e w_e}{\partial y} + \frac{\partial w_e^2}{\partial z} = -\frac{1}{\rho_e} \frac{\partial p_e}{\partial z} - g. \tag{2.7}$$

If the fluid is carrying chemical species with concentrations denoted by $\mathbf{U}_e$ and diffusivities in the x, y, z directions, respectively denoted by $\Gamma^x, \Gamma^y, \Gamma^z$, then, the extensive property has density $\Phi = \mathbf{U}_e$ and the flux $\mathbf{J}$ is a balance of diffusion and advection. Thus, balancing species concentration in an elementary volume (i.e. applying principle (2.1)) yields:

$$
\begin{aligned}
&\frac{\partial \mathbf{U}_e}{\partial t} + \frac{\partial u_e \mathbf{U}_e}{\partial x} + \frac{\partial v_e \mathbf{U}_e}{\partial y} + \frac{\partial w_e \mathbf{U}_e}{\partial z} \\
&= \frac{\partial}{\partial x} \Gamma^x \frac{\partial \mathbf{U}_e}{\partial x} + \frac{\partial}{\partial y} \Gamma^y \frac{\partial \mathbf{U}_e}{\partial y} + \frac{\partial}{\partial z} \Gamma^z \frac{\partial \mathbf{U}_e}{\partial z} + \mathbf{S}_U,
\end{aligned} \tag{2.8}
$$

where $\mathbf{S}_U$ denotes chemical reactions.

In the system (2.4)-(2.8) there are as many equations as unknowns, thus, a solution may be obtained if appropriate boundary and initial conditions are specified. However, analytical solutions are not always feasible and numerical solutions are challenging and expensive due to the high space dimension and degrees of freedom of the concentration vector. Consequently, the incompressible Euler model is approximated (using approprite assumptions) to make it more tractable. Moreover, some flow problems have negligible vertical scales compared to horizontal scales, such problems are classified as shallow water flow problems [76, 153, 100, 161, 160]. It is assumed in this section, that the reactive flow problem under consideration obeys the shallow water principle.

Further, the underlying principle that vertical scales are negligible (as compared to the horizontal scales) imply that vertical acceleration (i.e the left hand side of (2.7)) is negligible [76, 153, 100, 161, 160]. Thus, the Right Hand Side (RHS) of (2.7) is zero. That is:

$$
-\frac{1}{\rho_e} \frac{\partial p_e}{\partial z} - g = 0. \tag{2.9}
$$

Integrating (2.9) from the reference water depth $h(x, y, t)$, to any water depth z yields the hydrostatic pressure function:

$$
p_e = \rho_e g(h - z). \tag{2.10}
$$

Further, from the hydrostatic pressure function (2.10), one can obtain the following pressure gradients across x and y directions:

$$\frac{\partial p_e}{\partial x} = \rho_e g \frac{\partial h}{\partial x}. \tag{2.11}$$

and

$$\frac{\partial p_e}{\partial y} = \rho_e g \frac{\partial h}{\partial y}. \tag{2.12}$$

The pressure gradients (2.11) and (2.12) are independent of z, which imply that the x, y velocity components u_e and v_e are independent of the z coordinate. Applying the pressure derivatives (2.11) and (2.12) in the Euler model (2.4)-(2.8) can be restated as follows:

$$\frac{\partial u_e}{\partial x} + \frac{\partial v_e}{\partial y} + \frac{\partial w_e}{\partial z} = 0, \tag{2.13}$$

$$\frac{\partial u_e}{\partial t} + \frac{\partial u_e^2}{\partial x} + \frac{\partial v_e u_e}{\partial y} + \frac{\partial w_e u_e}{\partial z} = -g \frac{\partial h}{\partial x}, \tag{2.14}$$

$$\frac{\partial v_e}{\partial t} + \frac{\partial u_e v_e}{\partial x} + \frac{\partial v_e^2}{\partial y} + \frac{\partial v_e w_e}{\partial z} = -g \frac{\partial h}{\partial y}, \tag{2.15}$$

and

$$\frac{\partial \mathbf{U}_e}{\partial t} + \frac{\partial \mathbf{J}^x}{\partial x} + \frac{\partial \mathbf{J}^y}{\partial y} + \frac{\partial \mathbf{J}^z}{\partial z} = \mathbf{S}_U, \tag{2.16}$$

where $\mathbf{J}^x = u_e \mathbf{U}_e - \Gamma^x \frac{\partial \mathbf{U}_e}{\partial x}$, $\mathbf{J}^y = v_e \mathbf{U}_e - \Gamma^y \frac{\partial \mathbf{U}_e}{\partial y}$ and $\mathbf{J}^z = w_e \mathbf{U}_e - \Gamma^z \frac{\partial \mathbf{U}_e}{\partial z}$.

The system (2.13)-(2.16) has one equation less than system (2.4)-(2.8) and therefore, is less expensive to solve. However, there are more unknowns than equations in system (2.13)-(2.16). This is due to the presence of w_e, the z velocity component (that is assumed insignificant under the shallow water assumption). To eliminate the redundant unknown, the system is averaged over depth and boundary conditions are employed to simplify the resulting averaged equations.

There are two relevant boundaries (located at z_b and z_s along the vertical axis) in the three dimensional domain where conditions necessary for deriving the shallow water

model must be specified. The bottom topography of the channel denoted by $\eta(z_b, x, y)$, is assumed fixed in time and impermeable, and the free-surface denoted by $H(z_s, x, y, t)$ that is under the influence of gravity and atmospheric pressure. This variable quantifies the total depth (i.e. $H = h + \eta$) and is also known as water surface [187].

The following no-normal flow and diffusion conditions are imposed at the bottom boundary denoted $\eta(z_b, x, y)$:

$$u_e \frac{\partial \eta}{\partial x} + v_e \frac{\partial \eta}{\partial y} + w_e = 0, \quad \Gamma_x \frac{\partial \eta}{\partial x} = \Gamma_y \frac{\partial \eta}{\partial y} = 0 \quad \text{and}$$

$$\mathbf{J}_x(\eta) \frac{\partial \eta}{\partial x} + \mathbf{J}_y(\eta) \frac{\partial \eta}{\partial y} + \mathbf{J}_z(\eta) = 0. \tag{2.17}$$

Further, no-normal flow and diffusion conditions are also imposed at the free boundary denoted $H(x, y, t)$:

$$\frac{\partial H}{\partial t} + u_e \frac{\partial H}{\partial x} + v_e \frac{\partial H}{\partial y} - w_e = 0, \quad \Gamma_x \frac{\partial H}{\partial x} = \Gamma_y \frac{\partial H}{\partial y} = 0 \quad \text{and}$$

$$\mathbf{U}_e(H) \frac{\partial H}{\partial t} + \mathbf{J}^x(H) \frac{\partial H}{\partial x} + \mathbf{J}^y(H) \frac{\partial H}{\partial y} - \mathbf{J}^z(H) = 0. \tag{2.18}$$

Moreover, the depth-averaged unknown variables are defined as follows:

$$\overline{u_e} = u = \frac{1}{h} \int_{\eta}^{H} u_e dz, \quad \overline{v_e} = v = \frac{1}{h} \int_{\eta}^{H} v_e dz,$$

$$\overline{\Gamma_x} = \Gamma_{xs} = \frac{1}{h} \int_{\eta}^{H} \Gamma_x dz, \quad \overline{\Gamma_y} = \Gamma_{ys} = \frac{1}{h} \int_{\eta}^{H} \Gamma_y dz, \tag{2.19}$$

$$\mathbf{S}_h = \int_{\eta}^{H} \mathbf{S}_U dz \quad \text{and} \quad \mathbf{U} = \int_{\eta}^{H} \mathbf{U}_e dz.$$

Further, the original parameters/variables $u_e, v_e, \Gamma_x, \Gamma_y$ and $\mathbf{U}_e$ are related to the averaged quantities by the following expressions:

$$u_e = \overline{u_e} + u'_e, \quad v_e = \overline{v_e} + v'_e,$$

$$\Gamma_x = \overline{\Gamma_x} + \Gamma'_x, \quad \Gamma_y = \overline{\Gamma_y} + \Gamma'_y, \quad \text{and} \tag{2.20}$$

$$\mathbf{U}_e = \overline{\mathbf{U}_e} + \mathbf{U}'_e, \tag{2.21}$$

where $u'_e, v'_e, \Gamma'_x, \Gamma'_y$ and $\mathbf{U}'_e$ are mean deviations from the quantities $u_e, v_e, \Gamma_x, \Gamma_y$ and $\mathbf{U}_e$ respectively. Employing the depth-averaged quantities given by (2.19) and their

relations given by (2.20), the unknown quantity w_e in the system (2.13)-(2.16) can easily be eliminated.

Firstly, integrate the continuity equation (2.13) over the depth as follows:

$$\int_\eta^H \frac{\partial u_e}{\partial x}dz + \int_\eta^H \frac{\partial v_e}{\partial y}dz + \int_\eta^H \frac{\partial w_e}{\partial z}dz = 0. \tag{2.22}$$

Apply Leibniz rule of integration to the integrals in (2.22) to obtain:

$$\frac{\partial}{\partial x}\int_\eta^H u_e dz + \frac{\partial}{\partial y}\int_\eta^H v_e dz - \left(u_e(H)\frac{\partial H}{\partial x} + v_e(H)\frac{\partial H}{\partial y} - w_e(H)\right)$$
$$- \left(u_e(\eta)\frac{\partial \eta}{\partial x} + v_e(\eta)\frac{\partial \eta}{\partial y} + w_e(\eta)\right) = 0. \tag{2.23}$$

Apply boundary conditions (2.17) and (2.18) in Equation (2.23) to obtain:

$$\frac{\partial}{\partial x}\int_\eta^H u_e dz + \frac{\partial}{\partial y}\int_\eta^H v_e dz + \frac{\partial h}{\partial t} = 0. \tag{2.24}$$

Use definition (2.19) in (2.24) to obtain:

$$\frac{\partial hu}{\partial x} + \frac{\partial hv}{\partial y} + \frac{\partial h}{\partial t} = 0. \tag{2.25}$$

Secondly, integrate the $x-$momentum equation (2.14) over the water column and apply Leibniz rule of integration to obtain:

$$\frac{\partial}{\partial t}\int_\eta^H u_e dz + \frac{\partial}{\partial x}\int_\eta^H u_e^2 dz + \frac{\partial}{\partial y}\int_\eta^H u_e v_e dz - u_e(H)\left(\frac{\partial H}{\partial t} + u_e(H)\frac{\partial H}{\partial x}\right.$$
$$\left. + v_e(H)\frac{\partial H}{\partial y} - w_e(H)\right) - u_e(\eta)\left(u_e(\eta)\frac{\partial \eta}{\partial x} + v_e(\eta)\frac{\partial \eta}{\partial y} + w_e(\eta)\right)$$
$$= -gh\frac{\partial h}{\partial x}. \tag{2.26}$$

Apply boundary conditions (2.17) and (2.18) in Equation (2.26) to obtain:

$$\frac{\partial}{\partial t}\int_\eta^H u_e dz + \frac{\partial}{\partial x}\int_\eta^H u_e^2 dz + \frac{\partial}{\partial y}\int_\eta^H u_e v_e dz = -gh\frac{\partial h}{\partial x}. \tag{2.27}$$

Apply definitions (2.19) and (2.20) in the integral equation (2.27) to obtain:

$$\frac{\partial hu}{\partial t} + \frac{\partial}{\partial x}\int_\eta^H (\overline{u_e} + u_e')(\overline{u_e} + u_e')dz + \frac{\partial}{\partial y}\int_\eta^H (\overline{u_e} + u_e')(\overline{v_e} + v_e')dz$$
$$= -gh\frac{\partial h}{\partial x}. \tag{2.28}$$

9

Simplify (2.28) to obtain:

$$\frac{\partial hu}{\partial t} + \frac{\partial huu}{\partial x} + \frac{\partial huv}{\partial y}$$
$$= -gh\frac{\partial h}{\partial x} - \frac{\partial}{\partial x}\int_\eta^H u_e'u_e'dz - \frac{\partial}{\partial y}\int_\eta^H u_e'v_e'dz. \tag{2.29}$$

Using arguments similar to those used to obtain (2.29), the depth-averaged $y-$momentum equation (2.15) can be written as follows:

$$\frac{\partial hv}{\partial t} + \frac{\partial huv}{\partial x} + \frac{\partial hvv}{\partial y}$$
$$= -gh\frac{\partial h}{\partial y} - \frac{\partial}{\partial x}\int_\eta^H u_e'v_e'dz - \frac{\partial}{\partial y}\int_\eta^H v_e'v_e'dz. \tag{2.30}$$

Thirdly, integrate the species transport equation (2.16) over the water column and apply Leibniz rule of integration to obtain:

$$\frac{\partial}{\partial t}\int_\eta^H \mathbf{U}_edz + \frac{\partial}{\partial x}\int_\eta^H \mathbf{J}_xdz + \frac{\partial}{\partial y}\int_\eta^H \mathbf{J}_ydz - \left(\mathbf{U}_e(H)\frac{\partial H}{\partial t} + \mathbf{J}_x(H)\frac{\partial H}{\partial x}\right.$$
$$\left. + \mathbf{J}_y(\eta)\frac{\partial \eta}{\partial y} - \mathbf{J}_z(\eta)\right) - \left(\mathbf{J}_x(\eta)\frac{\partial \eta}{\partial x} + \mathbf{J}_y(\eta)\frac{\partial \eta}{\partial y} + \mathbf{J}_z(\eta)\right)$$
$$= \mathbf{S}_h. \tag{2.31}$$

Apply boundary conditions (2.17) and (2.18) in Equation (2.31) to obtain:

$$\frac{\partial}{\partial t}\int_\eta^H \mathbf{U}_edz + \frac{\partial}{\partial x}\int_\eta^H \mathbf{J}_xdz + \frac{\partial}{\partial y}\int_\eta^H \mathbf{J}_ydz = \mathbf{S}_h. \tag{2.32}$$

Use definitions (2.19) in (2.32) to obtain:

$$\frac{\partial h\mathbf{U}}{\partial t} + \frac{\partial}{\partial x}\int_\eta^H u_e\mathbf{U}_edz + \frac{\partial}{\partial y}\int_\eta^H v_e\mathbf{U}_edz$$
$$= \frac{\partial}{\partial x}\int_\eta^H \Gamma_x\frac{\partial \mathbf{U}_e}{\partial x}dz + \frac{\partial}{\partial y}\int_\eta^H \Gamma_y\frac{\partial \mathbf{U}_e}{\partial y}dz + \mathbf{S}_h. \tag{2.33}$$

Apply definitions (2.19) and (2.20), the zero-diffusion boundary conditions (2.17) and (2.18) in (2.33) and simplify to obtain:

$$\frac{\partial h\mathbf{U}}{\partial t} + \frac{\partial hu\mathbf{U}}{\partial x} + \frac{\partial hv\mathbf{U}}{\partial y} = \frac{\partial}{\partial x}\Gamma_{xs}\frac{\partial h\mathbf{U}}{\partial x} + \frac{\partial}{\partial y}\Gamma_{ys}\frac{\partial h\mathbf{U}}{\partial y}$$
$$- \frac{\partial}{\partial x}\int_{z_l}^\eta (u_e'\mathbf{U}_e' - \Gamma_x'\frac{\partial \mathbf{U}_e}{\partial x})dz - \frac{\partial}{\partial y}\int_{z_l}^\eta (v_e'\mathbf{U}_e' - \Gamma_y'\frac{\partial \mathbf{U}_e}{\partial y})dz$$
$$+ \mathbf{S}_h. \tag{2.34}$$

10

The integrals in (2.29), (2.30) and (2.34) are corrections for the averaged terms. These integrals may be approximated or determined through experimental data. A popular procedure is to replace the correction terms by empirical expressions that represent Coriolis force, wind forces, bottom friction etc [182, 114, 111]. In this discussion, an empirical expression that quantifies bottom friction is used to close the momentum equations and the correction terms in (2.34) are considered negligible. Consequently, a one dimensional form of the closed reactive-shallow model states that:

$$\frac{\partial h}{\partial t} + \frac{\partial hu}{\partial x} = 0, \tag{2.35}$$

$$\frac{\partial hu}{\partial t} + \frac{\partial}{\partial x}(hu^2 + \frac{1}{2}gh^2) = -gh\frac{\partial \eta}{\partial x} - \frac{gC^2|u|u}{h^{4/3}}, \tag{2.36}$$

$$\frac{\partial h\mathbf{U}}{\partial t} + \frac{\partial}{\partial x}hu\mathbf{U} = \frac{\partial}{\partial x}\Gamma\frac{\partial h\mathbf{U}}{\partial x} + \mathbf{S}_h, \tag{2.37}$$

where C is the manning coefficient.

2.2 Hyperbolicity and non-linear effects

In this section, properties such as hyperbolicity and consequence of the non-linearity of the reactive-shallow water model are presented. The one dimensional form of the model is a balanced law and the homogeneous part is a conservation law. The model is analysed in both primitive and conservative formulations. Definitions relevant to the presentation are presented first, followed by the main results.

Definition 2.1 (Balanced laws). *Balanced laws are partial differential equations that can be expressed as follows (in one space dimension) [100, 161, 160, 53]:*

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{Q}) = \mathbf{S}(\mathbf{Q}), \tag{2.38}$$

where

$$\mathbf{Q} = \begin{bmatrix} q_1 \\ q_2 \\ \vdots \\ q_M \end{bmatrix}, \quad \mathbf{f}(\mathbf{Q}) = \begin{bmatrix} f_1(\mathbf{Q}) \\ f_2(\mathbf{Q}) \\ \vdots \\ f_M(\mathbf{Q}) \end{bmatrix}, \quad and \quad \mathbf{S}(\mathbf{Q}) = \begin{bmatrix} s_1(\mathbf{Q}) \\ s_2(\mathbf{Q}) \\ \vdots \\ s_M(\mathbf{Q}) \end{bmatrix}$$

11

$\mathbf{Q}$ *denotes the vector of* M *conserved variables,* $\mathbf{f}(\mathbf{Q})$ *denotes the vector of fluxes and* $\mathbf{S}(\mathbf{Q})$ *denotes a vector of sources for the conserved variables. The Jacobian of the flux vector function* $\mathbf{f}(\mathbf{Q})$ *is a matrix defined as follows:*

$$\mathbb{A}(\mathbf{Q}) = \begin{bmatrix} \frac{\partial f_1}{\partial q_1} & \frac{\partial f_1}{\partial q_2} & \cdots & \frac{\partial f_1}{\partial q_M} \\ \frac{\partial f_2}{\partial q_1} & \frac{\partial f_2}{\partial q_2} & \cdots & \frac{\partial f_2}{\partial q_M} \\ \vdots & \vdots & \vdots & \vdots \\ \frac{\partial f_M}{\partial q_1} & \frac{\partial f_M}{\partial q_2} & \cdots & \frac{\partial f_M}{\partial q_M} \end{bmatrix}. \tag{2.39}$$

Definition 2.2 (Eigenvalues and Eigenvectors)**.** *The eigenvalues* λ_i *of a matrix* $\mathbb{A}$ *are the zeros of the characteristic polynomial [100, 161, 160]:*

$$|\mathbb{A} - \lambda\mathbb{I}| = det(\mathbb{A} - \lambda\mathbb{I}) \tag{2.40}$$

where $\mathbb{I}$ *is the identity matrix. A right eigenvector of a matrix* $\mathbb{A}$ *corresponding to an eigenvalue* λ_i *of* $\mathbb{A}$ *is a vector* $\mathbf{K}^{(i)} = [k_1, k_2, \ldots, k_M]^T$ *satisfying*

$$\mathbb{A}\mathbf{K}^{(i)} = \lambda_i\mathbf{K}^{(i)}. \tag{2.41}$$

Similarly, a left eigenvector of a matrix $\mathbb{A}$ *corresponding to an eigenvalue* λ_i *of* $\mathbb{A}$ *is a vector* $\mathbf{L}^{(i)} = [l_1, l_2, \ldots, l_M]$ *such that*

$$\mathbf{L}^{(i)}\mathbb{A} = \lambda_i\mathbf{L}^{(i)}. \tag{2.42}$$

Definition 2.3 (Hyperbolic system [100, 161, 53])**.** *The balanced law* (2.38) *is said to be hyperbolic if* $\mathbb{A}$ *has* M *real eigenvalues* $\lambda_1, \ldots, \lambda_M$ *and a corresponding set of* M *linearly independent right eigenvectors* $\mathbf{K}^{(1)}, \ldots, \mathbf{K}^{(M)}$. *The system is said to be strictly hyperbolic if the eigenvalues* λ_i *are all distinct.*

Proposition 2.1. *The Shallow water model* (2.35)-(2.36) *with chemical transport* (2.37) *is hyperbolic with eigenvalues given by:*

$$\lambda_1 = u - c, \quad \lambda_2 = u, \quad \lambda_3 = u + c$$

where $c = \sqrt{hg}$ *is called celerity, and corresponding linearly independent eigenvectors are*

12

given by:

$$\mathbf{K}^{(1)} = \alpha_1 \begin{bmatrix} 1 \\ u - c \\ U \end{bmatrix}, \quad \mathbf{K}^{(2)} = \alpha_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad and \; \mathbf{K}^{(3)} = \alpha_3 \begin{bmatrix} 1 \\ u + c \\ U \end{bmatrix},$$

where U is any element of vector $\mathbf{U}$ and $\alpha_1, \alpha_2, \alpha_3$ are real numbers. The model is strictly hyperbolic for a wet bed problem.

Proof. The shallow water model with reactive transport given by Equations (2.35)-(2.37) can be written in the form:

$$\frac{\partial}{\partial t} \begin{bmatrix} h \\ hu \\ h\mathbf{U} \end{bmatrix} + \frac{\partial}{\partial x} \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ hu\mathbf{U} \end{bmatrix} = \begin{bmatrix} 0 \\ -gh\frac{\partial \eta}{\partial x} - \frac{g\gamma C^2 |u| u}{h^{4/3}} \\ \frac{\partial}{\partial x}\Gamma\frac{\partial h\mathbf{U}}{\partial x} + \mathbf{S}_h \end{bmatrix} \tag{2.43}$$

Thus, the vectors for conserved quantities, fluxes and sources are respectively given by:

$$\mathbf{Q} = \begin{bmatrix} h \\ hu \\ h\mathbf{U} \end{bmatrix}, \quad \mathbf{f}(\mathbf{Q}) = \begin{bmatrix} hu \\ hu^2 + \frac{1}{2}gh^2 \\ hu\mathbf{U} \end{bmatrix}, \quad and \; \mathbf{S}(\mathbf{Q}) = \begin{bmatrix} 0 \\ -gh\frac{\partial \eta}{\partial x} - \frac{g\gamma C^2 |u| u}{h^{4/3}} \\ \frac{\partial}{\partial x}\Gamma\frac{\partial h\mathbf{U}}{\partial x} + \mathbf{S}_h \end{bmatrix}. \tag{2.44}$$

For any component U of the unknown vector $\mathbf{U}$, the Jacobian of the flux vector $\mathbf{f}(\mathbf{Q})$ is given by:

$$\mathbb{A}(\mathbf{Q}) = \begin{bmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uU & U & u \end{bmatrix}. \tag{2.45}$$

The characteristic polynomial of $\mathbb{A}(\mathbf{Q})$ is given by:

$$det(\mathbb{A} - \lambda\mathbb{I}) = (\lambda - u)(\lambda^2 - 2u\lambda - (c^2 - u^2)) \tag{2.46}$$

$$= (\lambda - u)(\lambda - (u + c))(\lambda - (u - c)). \tag{2.47}$$

Thus, by definition the eigenvalues of $\mathbb{A}(\mathbf{Q})$ are $\lambda_1 = u - c$, $\lambda_2 = u$, and $\lambda_3 = u + c$. The linearly independent vectors

$$\mathbf{K}^{(1)} = \alpha_1 \begin{bmatrix} 1 \\ u - c \\ U \end{bmatrix}, \quad \mathbf{K}^{(2)} = \alpha_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{K}^{(3)} = \alpha_3 \begin{bmatrix} 1 \\ u + c \\ U \end{bmatrix},$$

and eigenvalues λ_1, λ_2 and λ_3 satisfy definition (2.2), that is,

$$\begin{bmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uU & U & u \end{bmatrix} \begin{bmatrix} \alpha_1 \\ \alpha_1(u - c) \\ \alpha_1 U \end{bmatrix} = \alpha_1(u - c) \begin{bmatrix} 1 \\ u - c \\ U \end{bmatrix} \quad \text{implying} \quad \mathbb{A}\mathbf{K}^{(1)} = \lambda_1 \mathbf{K}^{(1)}, \quad (2.48)$$

$$\begin{bmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uU & U & u \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \alpha_2 \end{bmatrix} = u\alpha_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix} \quad \text{implying} \quad \mathbb{A}\mathbf{K}^{(2)} = \lambda_2 \mathbf{K}^{(2)} \quad (2.49)$$

and

$$\begin{bmatrix} 0 & 1 & 0 \\ c^2 - u^2 & 2u & 0 \\ -uU & U & u \end{bmatrix} \begin{bmatrix} \alpha_3 \\ \alpha_3(u + c) \\ \alpha_3 U \end{bmatrix} = (u + c)\alpha_3 \begin{bmatrix} 1 \\ u + c \\ U \end{bmatrix} \quad \text{implying} \quad \mathbb{A}\mathbf{K}^{(3)} = \lambda_3 \mathbf{K}^{(3)}. \quad (2.50)$$

Thus $\mathbf{K}^{(1)}, \mathbf{K}^{(2)}$ and $\mathbf{K}^{(3)}$ are right eigenvectors corresponding to the eigenvalues λ_1, λ_2 and λ_3 respectively. For a wet bed problem the water depth is positive (i.e. $h > 0$) implying a positive celerity (i.e. $c > 0$) and completely distinct eigenvalues. Hence, the system is strictly hyperbolic. $\qquad \square$

Moreover, model (2.35)-(2.37) can also be formulated using primitive or physical variables h, u, and $\mathbf{U}$. Simple expansion of the derivatives and algebraic manipulations yield the following quasi-linear system [161, 160]:

$$\frac{\partial \mathbf{W}}{\partial t} + \mathbb{B}(\mathbf{W})\frac{\partial \mathbf{W}}{\partial x} = \mathbf{S}(\mathbf{W}), \quad (2.51)$$

where

$$
\mathbf{W} = \begin{bmatrix} h \\ u \\ \mathbf{U} \end{bmatrix}, \quad \mathbb{B}(\mathbf{W}) = \begin{bmatrix} u & h & 0 \\ g & u & 0 \\ 0 & 0 & u \end{bmatrix} \quad \text{and} \quad \mathbf{S}(\mathbf{W}) = \begin{bmatrix} 0 \\ -g\frac{\partial \eta}{\partial x} - \frac{g\gamma C^2 |u|u}{h^{7/3}} \\ \frac{1}{h}\frac{\partial}{\partial x}\Gamma\frac{\partial \mathbf{U}}{\partial x} + \frac{1}{h}\mathbf{S}_h \end{bmatrix}.
$$

Proposition 2.2. *The quasi-linear system* (2.51) *is hyperbolic with eigenvalues given by:*

$$
\lambda_1 = u - c, \quad \lambda_2 = u, \quad \lambda_3 = u + c
$$

where $c = \sqrt{hg}$ *and corresponding linearly independent eigenvectors are given by:*

$$
\mathbf{R}^{(1)} = \beta_1 \begin{bmatrix} h \\ -c \\ 0 \end{bmatrix}, \quad \mathbf{R}^{(2)} = \beta_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{R}^{(3)} = \beta_3 \begin{bmatrix} h \\ c \\ 0 \end{bmatrix},
$$

where U *is a component of* $\mathbf{U}$ *and* β_1, β_2 *and* β_3 *are real numbers. The system is strictly hyperbolic for a wet bed problem.*

Proof. For any component U of the concentration vector $\mathbf{U}$ the characteristic equation of the Jacobian matrix $\mathbb{B}(\mathbf{W})$ is given by:

$$
det(\mathbb{B} - \lambda\mathbb{I}) = (\lambda - u)\big((\lambda - u)^2 - gh\big) \tag{2.52}
$$

$$
= (\lambda - u)(\lambda - (u + c))(\lambda - (u - c)). \tag{2.53}
$$

Thus, by definition the eigenvalues of $\mathbb{A}(\mathbf{W})$ are $\lambda_1 = u - c$, $\lambda_2 = u$, and $\lambda_3 = u + c$. The linearly independent vectors

$$
\mathbf{R}^{(1)} = \beta_1 \begin{bmatrix} h \\ -c \\ 0 \end{bmatrix}, \quad \mathbf{R}^{(2)} = \beta_2 \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \text{and} \quad \mathbf{R}^{(3)} = \beta_3 \begin{bmatrix} h \\ c \\ 0 \end{bmatrix},
$$

and eigenvalues λ_1, λ_2 and λ_3 satisfy definition (2.2), that is,

$$
\begin{bmatrix} u & h & 0 \\ g & u & 0 \\ 0 & 0 & u \end{bmatrix} \begin{bmatrix} h\beta_1 \\ -c\beta_1 \\ 0 \end{bmatrix} = (u - c)\beta_1 \begin{bmatrix} h \\ -c \\ 0 \end{bmatrix} \quad \text{implying} \quad \mathbb{B}\mathbf{R}^{(1)} = \lambda_1 \mathbf{R}^{(1)}, \tag{2.54}
$$

15

$$\begin{bmatrix} u & h & 0 \\ g & u & 0 \\ 0 & 0 & u \end{bmatrix} \begin{bmatrix} 0 \\ 0 \\ \beta_2 \end{bmatrix} = u\beta_2 \begin{bmatrix} 0 \\ 0 \\ \beta_2 \end{bmatrix} \quad \text{implying} \quad \mathbb{B}\mathbf{R}^{(2)} = \lambda_2\mathbf{R}^{(2)} \tag{2.55}$$

and

$$\begin{bmatrix} u & h & 0 \\ g & u & 0 \\ 0 & 0 & u \end{bmatrix} \begin{bmatrix} h\beta_3 \\ c\beta_3 \\ 0 \end{bmatrix} = (u+c)\beta_3 \begin{bmatrix} h \\ c \\ 0 \end{bmatrix} \quad \text{implying} \quad \mathbb{B}\mathbf{R}^{(3)} = \lambda_3\mathbf{R}^{(3)}. \tag{2.56}$$

Thus $\mathbf{R}^{(1)}, \mathbf{R}^{(2)}$ and $\mathbf{R}^{(3)}$ are right eigenvectors corresponding to the eigenvalues λ_1, λ_2 and λ_3 respectively. For a wet bed problem the water depth is positive (i.e. $h > 0$) implying a positive celerity (i.e. $c > 0$) and completely distinct eigenvalues. Hence, the system is strictly hyperbolic. $\qquad\square$

Furthermore, small-amplitude waves such as gravity waves can be investigated using a linearised form of the shallow water model. Suppose the fluid is flowing at a constant velocity u_0 and has a constant depth h_0, then the gravity waves propagate at speeds $\lambda_1 = u_0 - c_0$, $\lambda_2 = u_0$, and $\lambda_3 = u_0 + c_0$ with celerity $c_0 = \sqrt{gh_0}$. Note that λ_1 and λ_2 families of waves propagate with speeds $\pm c_0$ relative to the fluid velocity and can assume any sign (i.e. positive or negative) depending on a dimensionless critical number called *Froude number* defined by:

$$\mathsf{F} = \frac{|u|}{c}. \tag{2.57}$$

The depth of a fluid varies across a wave as the depth of a trough is shallow compared with the depth of a crest [100, 161, 53]. Since the celerity (which measures speed relative to the fluid velocity) depends on depth, the speed varies across a particular wave as the crest propagates faster than the trough. If the depth of the fluid is exceedingly greater than the amplitude of the wave (i.e. flows involving small amplitude waves), then variations in speed across the wave are negligible and the linearised model can adequately capture the physics of the flow phenomenon. However, the speed variations across the wave are not negligible if the amplitude is larger, the non-linear model is employed in such flow cases. The consequence of non-linearity is expansion and compression of the wave, that can lead to the development of nearly discontinuous parts (known in literature as *shocks* or *hydraulic jumps*), discontinuities and rarefactions [100, 161].

16

2.3 Characteristic fields

Each eigenvalue $\lambda_i(\mathbf{Q})$ or eigenvector $\mathbf{K}^{(i)}(\mathbf{Q})$ defines a characteristic field called λ_i-field or $\mathbf{K}^{(i)}-$field [100]. The characteristic fields are either linearly degenerate or genuinely nonlinear. Given an M dimensional vector $\mathbf{Q} = (q_1, q_2, \ldots, q_M)$, the gradient of each eigenvalue $\lambda_i(\mathbf{Q})$ in the M dimensional vector space is given by [100, 161, 160, 53]:

$$\nabla \lambda_i(\mathbf{Q}) = \left(\frac{\partial}{\partial q_1} \lambda_i(\mathbf{Q}), \frac{\partial}{\partial q_2} \lambda_i(\mathbf{Q}), \ldots, \frac{\partial}{\partial q_M} \lambda_i(\mathbf{Q}) \right)^T. \tag{2.58}$$

Definition 2.4 (Genuinely non-linear and linearly degenerate fields)**.** *Characteristic fields (corresponding to the eigenvalue-eigenvector pair $\lambda_i - \mathbf{K}^{(i)}$) that satisfy [100, 161, 53]:*

$$\nabla \lambda_i(\mathbf{Q}) \cdot \mathbf{K}^{(i)}(\mathbf{Q}) = 0, \tag{2.59}$$

(where $'\cdot'$ denotes dot product in phase space) are said to be linearly degenerate, otherwise, are said to be genuinely non-linear.

Proposition 2.3 (Fields in conservative formulation)**.** *For the conservation law (i.e. homogeneous part of the balanced law):*

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial}{\partial x} \mathbf{f}(\mathbf{Q}) = 0 \tag{2.60}$$

where vector $\mathbf{Q}$ and flux function $\mathbf{f}(\mathbf{Q})$ are defined as in (2.43), *the characteristic field corresponding to λ_2 is linearly degenerate while λ_1 and λ_3-fields are genuinely non-linear.*

Proof. The gradient of the eigenvalue $\lambda_2(\mathbf{Q})$ is given by:

$$\nabla \lambda_2(\mathbf{Q}) = \left(-\frac{u}{h}, \frac{1}{h}, 0 \right)^T. \tag{2.61}$$

Thus, the dot product in phase space of the gradient $\nabla \lambda_2(\mathbf{Q})$ and corresponding eigenvector $\mathbf{K}^{(2)}(\mathbf{Q})$ is given by:

$$\nabla \lambda_2(\mathbf{Q}) \cdot \mathbf{K}^{(2)}(\mathbf{Q}) = \begin{bmatrix} -\frac{u}{h} \\ \frac{1}{h} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} 0 \\ 0 \\ \alpha_2 \end{bmatrix} = 0. \tag{2.62}$$

Moreover, the gradient of the eigenvalue $\lambda_1(\mathbf{Q})$ is given by:

$$\nabla \lambda_1(\mathbf{Q}) = \left(-\frac{u}{h} - \frac{1}{2}\sqrt{g/h}, \frac{1}{h}, 0 \right)^T. \tag{2.63}$$

Thus the dot product in phase space of the gradient $\nabla \lambda_1(\mathbf{Q})$ and corresponding eigenvector $\mathbf{K}^{(1)}(\mathbf{Q})$ is given by:

$$\nabla \lambda_1(\mathbf{Q}) \cdot \mathbf{K}^{(1)}(\mathbf{Q}) = \begin{bmatrix} -\frac{u}{h} - \frac{1}{2}\sqrt{g/h} \\ \frac{1}{h} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \alpha_1 \\ \alpha_1(u-c) \\ \alpha_1 U \end{bmatrix} = \alpha_1 \left(\frac{-c}{h} - \frac{1}{2}\sqrt{g/h} \right) \neq 0. \tag{2.64}$$

Further,

$$\nabla \lambda_3(\mathbf{Q}) = \left(-\frac{u}{h} + \frac{1}{2}\sqrt{g/h}, \frac{1}{h}, 0 \right)^T. \tag{2.65}$$

Thus, the dot product in phase space of the gradient $\nabla \lambda_3(\mathbf{Q})$ and corresponding eigenvector $\mathbf{K}^{(3)}(\mathbf{Q})$ is given by:

$$\nabla \lambda_3(\mathbf{Q}) \cdot \mathbf{K}^{(3)}(\mathbf{Q}) = \begin{bmatrix} -\frac{u}{h} + \frac{1}{2}\sqrt{g/h} \\ \frac{1}{h} \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \alpha_3 \\ \alpha_3(u+c) \\ \alpha_3 U \end{bmatrix} = \alpha_3 \left(\frac{c}{h} + \frac{1}{2}\sqrt{g/h} \right) \neq 0. \tag{2.66}$$

Hence, λ_1 and λ_3-fields are genuinely non-linear.

$\square$

Proposition 2.4 (Fields in primitive formulation)**.** *Given the quasi-linear system:*

$$\frac{\partial \mathbf{W}}{\partial t} + \mathbb{B}(\mathbf{W})\frac{\partial \mathbf{W}}{\partial x} = 0, \tag{2.67}$$

with vector $\mathbf{W}$ and Jacobian matrix $\mathbb{B}(\mathbf{W})$ defined as in (2.51), the characteristic field corresponding to λ_2 is linearly degenerate while λ_1 and λ_3-fields are genuinely non-linear.

Proof. Using the eigenvalues (i.e. $\lambda_1(\mathbf{W}), \lambda_2(\mathbf{W})$ and $\lambda_3(\mathbf{W})$) and right eigenvectors (i.e. $\mathbf{R}^{(1)}(\mathbf{W}), \mathbf{R}^{(2)}(\mathbf{W})$ and $\mathbf{R}^{(3)}(\mathbf{W})$) of the Jacobian matrix $\mathbb{B}(\mathbf{W})$, the proof follows analogously as in the conservative formulation (2.60) presented above. $\square$

2.4 Riemann problem and wave solutions

Existence and uniqueness of solutions to partial differential equations depend on initial and boundary conditions. In this section, a well-posed reactive-shallow water model is presented and general solutions are constructed. As shown in Section 2.2 above, waves that are initially smooth soon develop discontinuities, it is therefore, appropriate to analyse the model with discontinuous initial data. For the purpose of the analysis here, the conservative formulation of the reactive-shallow water model is considered on a boundless flow domain and discontinuous initial data set is assumed. The resulting initial-value problem (IVP) is generally known as Riemann problem [100, 161, 160, 53].

Using the conservative formulation, the reactive-shallow water Riemann problem states that, find $\mathbf{Q}$ such that:

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{Q}) = 0, \quad \text{in} \quad -\infty \leq x \leq \infty,$$

$$(2.68)$$

$$\mathbf{Q}(x,0) = \mathbf{Q}^0(x) = \begin{cases} \mathbf{Q}_L & \text{for } x < 0 \\ \mathbf{Q}_R & \text{for } x > 0 \end{cases}$$

where vector $\mathbf{Q}$ and flux function $\mathbf{f}(\mathbf{Q})$ are defined as in (2.43). Symbols $\mathbf{Q}_L$ and $\mathbf{Q}_R$ respectively denote left and right data states which are written explicitly as follows (for each U_L in $\mathbf{U}_L$ and U_R in $\mathbf{U}_R$):

$$\mathbf{Q}_L = \begin{bmatrix} h_L \\ h_L u_L \\ h_L U_L \end{bmatrix} \quad \text{and} \quad \mathbf{Q}_R = \begin{bmatrix} h_R \\ h_R u_R \\ h_R U_R \end{bmatrix}. \tag{2.69}$$

The general solution of the Riemann problem (2.69) consists of four states (i.e. $\mathbf{Q}_L, \mathbf{Q}_{*L}, \mathbf{Q}_{*R}, \mathbf{Q}_R$) separated by three waves that include contact and shock or rarefaction waves, see illustration in Figure 2.1. Consequently, there are four possible wave patterns in the general solution, as depicted in Figure 2.2. The main task of the Riemann problem is to find the intermediate states using the left and right data states available. However,

connections between data states are established using wave relations such as Riemann invariants, Rankine Hugoniot jump conditions and entropy conditions [100, 161, 160, 53].
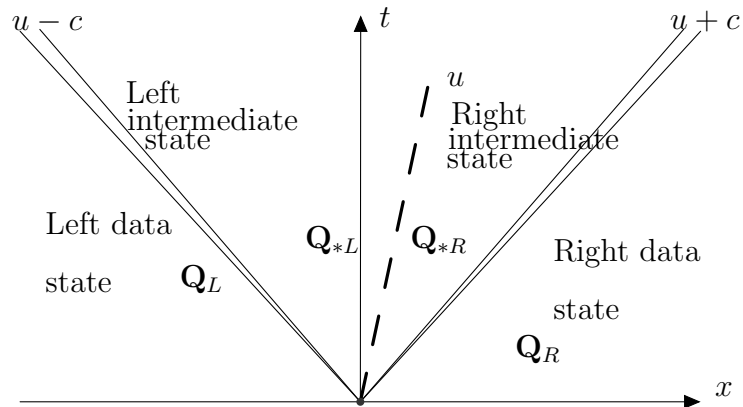
Figure 2.1: Structure of the general solution of the Riemann problem
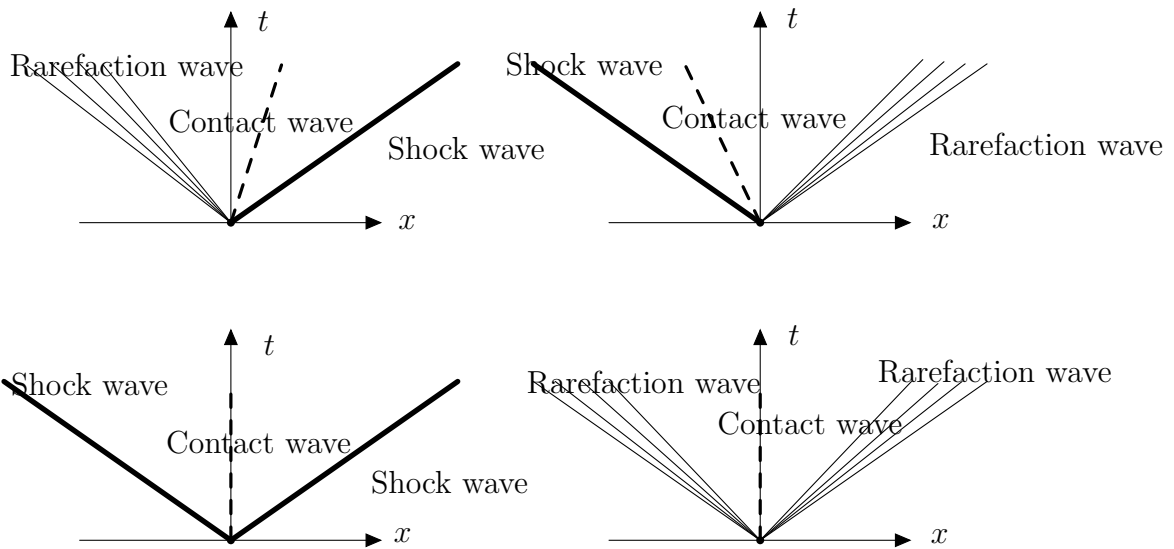
Figure 2.2: Possible wave patterns of the Riemann problem

Definition 2.5 (General Riemann invariants). *Given a general hyperbolic system of the*

20

quasi-linear form [100, 161, 160, 53]:

$$\frac{\partial \mathbf{W}}{\partial t} + \mathbb{B}(\mathbf{W})\frac{\partial \mathbf{W}}{\partial x} = 0, \tag{2.70}$$

with $\mathbf{W} = (W_1, W_2, \ldots, W_M)^T$ and an eigenvector corresponding to the eigenvalue λ_i denoted by $\mathbf{R}^{(i)} = (R_1^{(i)}, R_2^{(i)}, \ldots, R_M^{(i)})^T$, then general Riemann invariants are relations that satisfy the following $M - 1$ ordinary differential equations:

$$\frac{dW_1}{R_1^{(i)}} = \frac{dW_2}{R_2^{(i)}} = \cdots = \frac{dW_M}{R_M^{(i)}}. \tag{2.71}$$

Definition 2.6 (Rankine Hugoniot jump condition). *Given a general hyperbolic system of the conservative form [100, 161, 160, 53]:*

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{Q}) = 0, \tag{2.72}$$

where $\mathbf{Q}$ is the vector of unknowns and $\mathbf{f}(\mathbf{Q})$ is a vector of fluxes, the Rankine Hugoniot jump condition states that at a discontinuity:

$$\mathbf{f}_R - \mathbf{f}_L = S(\mathbf{Q}_R - \mathbf{Q}_L) \tag{2.73}$$

where S denotes speed of the discontinuity, $\mathbf{f}_R = \mathbf{f}(\mathbf{Q}_R)$, $\mathbf{f}_R = \mathbf{f}(\mathbf{Q}_R)$ and subscripts R and L denote right and left states of the discontinuity, respectively.

2.4.1 Rarefaction waves

Across a rarefaction wave, the left and right data states are connected through a smooth region in a genuinely non-linear field where characteristics diverge and Riemann invariants apply [100, 161, 160, 53]. Since the λ_1 and λ_3 characteristic fields of the hyperbolic system under consideration are genuinely non-linear, waves associated with them include rarefactions. A rarefaction wave associated with the λ_1 characteristic field is called a left rarefaction (depicted in Figure 2.4) while a rarefaction associated with λ_3 is called a right rarefaction (see Figure 2.3) [160].

Proposition 2.5 (Right rarefaction wave relations). *Given the left and right data states of a right rarefaction in primitive variables, i.e.,*

$$\mathbf{W}_{*R} = \begin{bmatrix} h_{*R} \\ u_{*R} \\ U_{*R} \end{bmatrix} \quad and \quad \mathbf{W}_R = \begin{bmatrix} h_R \\ u_R \\ U_R \end{bmatrix} \tag{2.74}$$

respectively, then:

$$u_{*R} - 2c_{*R} = u_R - 2c_R \quad and \quad U_{*R} = U_R, \tag{2.75}$$

where $c_{*R} = \sqrt{gh_{*R}}$ and $c_R = \sqrt{gh_R}$.
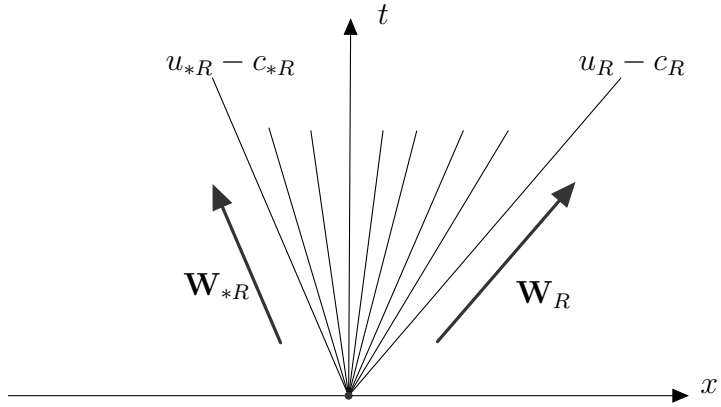
Figure 2.3: Right rarefaction wave

Proof. The right eigenvector of the λ_3 characteristic field in the primitive formulation is given by:

$$\mathbf{R}^{(3)} = \alpha_3(h, c, 0)^T, \tag{2.76}$$

where $c = \sqrt{gh}$. Thus, the generalised Riemann invariants associated with the λ_3 eigenvalue satisfy:

$$\frac{dh}{\alpha_3 h} = \frac{du}{\alpha_3 c} = \frac{dU}{0}, \tag{2.77}$$

Firstly, by solving the differential equation:

$$\frac{dh}{\alpha_3 h} = \frac{du}{\alpha_3 c}, \tag{2.78}$$

22

the following expression is obtained:

$$u - 2c = constant. \tag{2.79}$$

Thus, by applying initial data (2.74) in (2.79) yields:

$$u_{*R} - 2c_{*R} = u_R - 2c_R. \tag{2.80}$$

Secondly, solving the differential equation:

$$\frac{du}{\alpha_3 c} = \frac{dU}{0}, \tag{2.81}$$

yields:

$$U = constant. \tag{2.82}$$

Thus, applying initial data (2.74) in (2.82) yields:

$$U_{*R} = U_R. \tag{2.83}$$

$\square$

Proposition 2.6 (Left rarefaction wave relations). *Given the left and right data states of a left rarefaction in primitive variables, i.e.,*

$$\mathbf{W}_L = \begin{bmatrix} h_L \\ u_L \\ U_L \end{bmatrix} \quad and \quad \mathbf{W}_{*L} = \begin{bmatrix} h_{*L} \\ u_{*L} \\ U_{*L} \end{bmatrix} \tag{2.84}$$

respectively, then:

$$u_L + 2c_L = u_{*L} + 2c_{*L} \quad and \quad U_L = U_{*L}, \tag{2.85}$$

where $c_L = \sqrt{gh_L}$ and $c_{*L} = \sqrt{gh_{*L}}$.

Proof. The right eigenvector of the λ_1 characteristic field in the primitive formulation is given by:

$$\mathbf{R}^{(3)} = \alpha_1(h, -c, 0)^T, \tag{2.86}$$
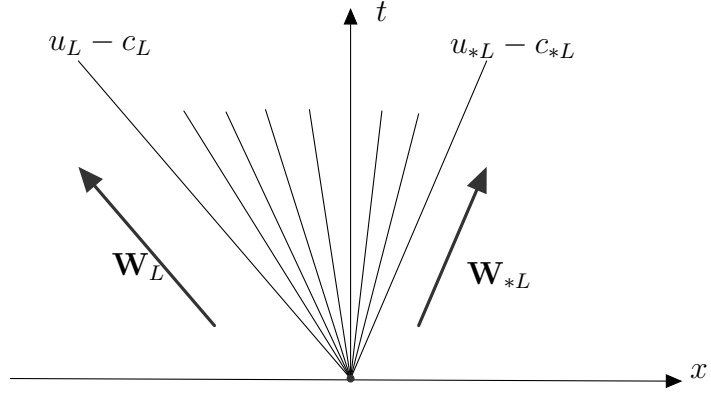
Figure 2.4: Left rarefaction wave

where $c = \sqrt{gh}$. Thus, the generalised Riemann invariants associated with the λ_1 eigenvalue satisfy:

$$\frac{dh}{\alpha_1 h} = -\frac{du}{\alpha_1 c} = \frac{dU}{0}. \tag{2.87}$$

Thus, algebraic manipulations of the differential equations (2.87) yield:

$$u + 2c = constant \quad \text{and} \quad U = constant. \tag{2.88}$$

Furthermore, by applying initial data (2.84) in (2.88) yields:

$$u_L + 2c_L = u_{*L} + 2c_{*L} \quad \text{and} \quad U_L = U_{*L}. \tag{2.89}$$

$\square$

2.4.2 Contact waves

Across a contact wave, characteristics are parallel, Rankine Hugoniot condition and Riemann invariants hold and the data states are connected by a single jump discontinuity in a linearly degenerate field [100, 161, 160, 53]. Since the λ_2 characteristic of the hyperbolic

system under consideration is linearly degenerate, the wave associated with it is a contact wave (depicted in Figure 2.5).

Proposition 2.7 (Contact wave relations). *Given the left and right data states of a contact wave in primitive variables, i.e.,*

$$\mathbf{W}_{*L} = \begin{bmatrix} h_{*L} \\ u_{*L} \\ U_{*L} \end{bmatrix} \quad and \quad \mathbf{W}_{*L} = \begin{bmatrix} h_{*R} \\ u_{*R} \\ U_{*R} \end{bmatrix} \tag{2.90}$$

respectively, then:

$$h_{*L} = h_{*R} \quad and \quad u_{*L} = u_{*R}. \tag{2.91}$$
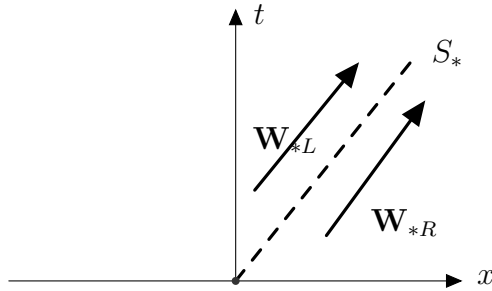
Figure 2.5: Middle contact wave

Proof. The right eigenvector of the λ_2 characteristic field in the primitive formulation is given by:

$$\mathbf{R}^{(3)} = \alpha_2 (0, 0, 1)^T, \tag{2.92}$$

Generalised Riemann invariants associated with the λ_2 eigenvalue satisfy:

$$\frac{dh}{0} = \frac{du}{0} = \frac{dU}{\alpha_2}. \tag{2.93}$$

Thus, algebraic manipulations of the differential equations (2.93) yield:

$$h = constant \quad \text{and} \quad u = constant.$$ (2.94)

Moreover, applying initial data (2.84) in (2.94) yields:

$$h_{*L} = h_{*R} \quad \text{and} \quad u_{*L} = u_{*R}.$$ (2.95)

□

2.4.3 Shock waves

Across a shock wave, the data states are connected by a single jump discontinuity in a genuinely non-linear field where the Rankine Hugoniot and entropy conditions apply [100, 161, 160, 53]. Solutions of hyperbolic systems involving shocks are not unique, thus, the entropy inequality:

$$\lambda_i(\mathbf{Q}_L) > S_i > \lambda_i(\mathbf{Q}_R)$$ (2.96)

where S_i is the shock speed, determines a physically relevant solution. Since shock waves are associated with genuinely non-linear characteristic fields, shocks can be associated with the λ_1 and λ_3 characteristics of the hyperbolic system under consideration. A shock wave associated with the λ_1 characteristic field is called a left shock (shown in Figure 2.6) while the one associated with λ_3 is called a right shock (depicted in Figure 2.7).

Proposition 2.8 (Left shock wave relations). *Given the left and right data states in primitive variables:*

$$\mathbf{W}_L = \begin{bmatrix} h_L \\ u_L \\ U_L \end{bmatrix} \quad \text{and} \quad \mathbf{W}_{*L} = \begin{bmatrix} h_{*L} \\ u_{*L} \\ U_{*L} \end{bmatrix}$$ (2.97)

respectively, separated by a left shock wave of speed S_L, the following relations hold across the shock:

$$U_{*L} = U_L, \ h_{*L} = \frac{-h_L + h_L\sqrt{1 + 8(\mathsf{F}_L - \mathsf{F}_S)^2}}{2}, \ u_{*L} = u_L - c_L\left(1 - \frac{h_L}{h_{*L}}\right)(\mathsf{F}_L - \mathsf{F}_S),$$ (2.98)

where $c_L = \sqrt{gh_L}$ is the celerity and $\mathsf{F}_L = \frac{u_L}{c_L}, \mathsf{F}_S = \frac{S_L}{c_L}$ are Froude numbers.
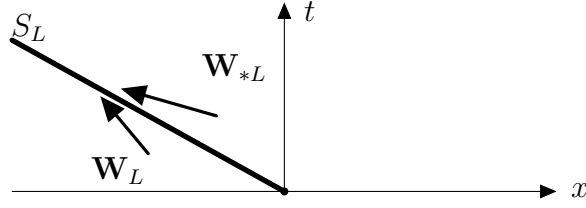
Figure 2.6: Left shock wave

Proof. Given the shock speed S_L, velocities at the left and right side of the shock can be expressed in terms of the shock speed as follows:

$$u_L = \hat{u}_L + S_L \quad \text{and} \quad u_{*L} = \hat{u}_{*L} + S_L, \tag{2.99}$$

where $\hat{u}_L$ and $\hat{u}_{*L}$ are variables in a frame of reference where the shock speed is zero. Thus, by using (2.99) variables in the conservative formulation can be transformed into conservative variables in the frame of reference where the shock speed is zero, as follows:

$$\mathbf{W}_L = \begin{bmatrix} h_L \\ h_L \hat{u}_L \\ h_L U_L \end{bmatrix} \quad \text{and} \quad \mathbf{W}_{*L} = \begin{bmatrix} h_{*L} \\ h_{*L} \hat{u}_{*L} \\ h_{*L} U_{*L} \end{bmatrix}. \tag{2.100}$$

Applying the transformed left and right data states (2.100) and the flux function $\mathbf{f}(\mathbf{Q})$ defined in (2.43), the Rankine Hugoniot jump condition states that:

$$h_{*L}\hat{u}_{*L} = h_L\hat{u}_L, \tag{2.101}$$

$$h_{*L}\hat{u}_{*L}^2 + \frac{1}{2}gh_{*L}^2 = h_L\hat{u}_L^2 + \frac{1}{2}gh_L^2, \tag{2.102}$$

$$h_{*L}\hat{u}_{*L}U_{*L} = h_L\hat{u}_L U_L. \tag{2.103}$$

It follows immediately by the mass flux Equations (2.101) and (2.103) that:

$$U_{*L} = U_L. \tag{2.104}$$

Further, denoting the mass flux across the left shock by ϕ_L, it follows from Equation (2.101) that the mass flux is constant across the shock, i.e.:

27

$$h_{*L}\hat{u}_{*L} = h_L\hat{u}_L = \phi_L. \tag{2.105}$$

Thus,

$$\hat{u}_{*L} = \frac{\phi_L}{h_{*L}}, \quad \text{and} \quad \hat{u}_L = \frac{\phi_L}{h_L}. \tag{2.106}$$

Expressing the momentum flux equality (2.102) in terms of the mass flux ϕ_L yields:

$$\phi_L(\hat{u}_L - \hat{u}_{*L}) = \frac{1}{2}g(h_{*L}^2 - h_L^2). \tag{2.107}$$

Applying the velocity expressions (2.106) in the momentum flux equality (2.107) and manipulating algebraically yields:

$$\phi_L = \sqrt{\frac{1}{2}gh_{*L}h_L(h_{*L} + h_L)}. \tag{2.108}$$

From the transformation (2.99), $\hat{u}_{*L} - \hat{u}_L = u_{*L} - u_L$, thus, applying original velocity variables with the mass flux expression (2.108) in the momentum flux equality (2.107), and manipulating algebraically yields:

$$u_{*L} = u_L - (h_{*L} - h_L)\sqrt{\frac{1}{2}g\frac{(h_{*L} + h_L)}{h_{*L}h_L}}. \tag{2.109}$$

Moreover, the velocity transformation (2.99) can also be expressed in terms of the mass flux using (2.106) as follows:

$$u_L = \frac{\phi_L}{h_L} + S_L. \tag{2.110}$$

Applying the mass flux expression (2.108) in (2.110) and manipulating algebraically yields the following expression for the shock speed:

$$S_L = u_L - c_L\sqrt{\frac{1}{2}h_{*L}\frac{(h_{*L} + h_L)}{h_L^2}}, \tag{2.111}$$

where $c_L = \sqrt{gh_L}$. Writing the shock expression (2.111) in terms of Froude numbers yields:

$$\mathsf{F}_L - \mathsf{F}_S = \sqrt{\frac{1}{2}h_{*L}\frac{(h_{*L} + h_L)}{h_L^2}}. \tag{2.112}$$

28

Equation (2.112) is quadratic in h_{*L}, and has roots given by:

$$h_{*L} = \frac{-h_L \pm h_L \sqrt{1 + 8(\mathsf{F}_L - \mathsf{F}_S)^2}}{2}. \tag{2.113}$$

However, $h_{*L} \geq h_L$ for a left shock, thus, expression (2.112) implies $\mathsf{F}_L - \mathsf{F}_S \geq 1$. Therefore, the relevant root of the quadratic (2.112) is given by:

$$h_{*L} = \frac{-h_L + h_L \sqrt{1 + 8(\mathsf{F}_L - \mathsf{F}_S)^2}}{2}. \tag{2.114}$$

Furthermore, applying the shock expression (2.111) and using (2.112) in the velocity expression (2.109) yields:

$$u_{*L} = u_L - c_L \left(1 - \frac{h_L}{h_{*L}}\right)(\mathsf{F}_L - \mathsf{F}_S). \tag{2.115}$$

$\square$

Proposition 2.9 (Right shock wave relations). *Given the left and right data states in primitive variables:*

$$\mathbf{W}_{*R} = \begin{bmatrix} h_{*R} \\ u_{*R} \\ U_{*R} \end{bmatrix} \quad and \quad \mathbf{W}_R = \begin{bmatrix} h_R \\ u_R \\ U_R \end{bmatrix} \tag{2.116}$$

respectively, separated by a right shock wave of speed S_R, the following relations hold across the shock:

$$U_{*R} = U_R, \ h_{*R} = \frac{-h_R + h_R \sqrt{1 + 8(\mathsf{F}_L - \mathsf{F}_S)^2}}{2}, \ u_{*R} = u_R + c_R \left(1 - \frac{h_R}{h_{*R}}\right)(\mathsf{F}_L - \mathsf{F}_S), \tag{2.117}$$

where $c_R = \sqrt{g h_R}$ is the celerity and $\mathsf{F}_R = \frac{u_R}{c_R}, \mathsf{F}_S = \frac{S_R}{c_R}$ are Froude numbers.

Proof. Given the shock speed S_R, velocities at the left and right side of the shock can be expressed in terms of the shock speed as follows:

$$u_R = \hat{u}_R + S_R \ \text{ and } \ u_{*R} = \hat{u}_{*R} + S_R, \tag{2.118}$$

where $\hat{u}_R$ and $\hat{u}_{*R}$ are variables in a frame of reference where the shock speed is zero. Thus, by using (2.118) variables in the conservative formulation can be transformed into conservative variables in the frame of reference where the shock speed is zero, as follows:
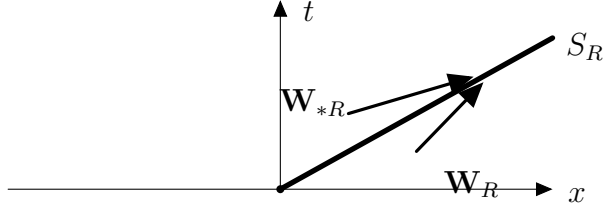
Figure 2.7: Right shock wave

$$\mathbf{W}_R = \begin{bmatrix} h_R \\ h_R \hat{u}_R \\ h_R U_R \end{bmatrix} \quad \text{and} \quad \mathbf{W}_{*R} = \begin{bmatrix} h_{*R} \\ h_{*R} \hat{u}_{*R} \\ h_{*R} U_{*R} \end{bmatrix}. \tag{2.119}$$

Applying the transformed left and right data states (2.119) and the flux function $\mathbf{f}(\mathbf{Q})$ defined in (2.43), the Rankine Hugoniot jump condition states that:

$$h_{*R} \hat{u}_{*R} = h_R \hat{u}_R, \tag{2.120}$$

$$h_{*R} \hat{u}_{*R}^2 + \frac{1}{2} g h_{*R}^2 = h_R \hat{u}_R^2 + \frac{1}{2} g h_R^2, \tag{2.121}$$

$$h_{*R} \hat{u}_{*R} U_{*R} = h_R \hat{u}_R U_R. \tag{2.122}$$

It follows immediately by the mass flux Equations (2.120) and (2.122) that:

$$U_{*R} = U_R. \tag{2.123}$$

Further, denoting the mass flux across the left shock by ϕ_R, it follows from Equation (2.120) that the mass flux is constant across the shock, i.e.,:

$$-h_{*R} \hat{u}_{*R} = -h_R \hat{u}_R = \phi_R. \tag{2.124}$$

Thus,

$$\hat{u}_{*R} = -\frac{\phi_R}{h_{*R}}, \quad \text{and} \quad \hat{u}_R = -\frac{\phi_R}{h_R}. \tag{2.125}$$

Expressing the momentum flux equality (2.121) in terms of the mass flux ϕ_R yields:

$$\phi_R(\hat{u}_{*R} - \hat{u}_R) = \frac{1}{2}g(h_{*R}^2 - h_R^2). \tag{2.126}$$

Applying velocity expressions (2.125) in the momentum flux equality (2.126) and manipulating algebraically yields:

$$\phi_R = \sqrt{\frac{1}{2}gh_{*R}h_R(h_{*R} + h_R)}. \tag{2.127}$$

From the transformation (2.118), $\hat{u}_{*R} - \hat{u}_R = u_{*R} - u_R$, thus, applying original velocity variables with the mass flux expression (2.127) in the momentum flux equality (2.127), and manipulating algebraically yields:

$$u_{*R} = u_R + (h_{*R} - h_R)\sqrt{\frac{1}{2}g\frac{(h_{*R} + h_R)}{h_{*R}h_R}}. \tag{2.128}$$

Moreover, the velocity transformation (2.118) can also be expressed in terms of the mass flux using (2.125) as follows:

$$u_R = -\frac{\phi_R}{h_R} + S_R. \tag{2.129}$$

Applying the mass flux expression (2.127) in (2.129) and manipulating algebraically yields the following expression for the shock speed:

$$S_R = u_R + c_R\sqrt{\frac{1}{2}h_{*R}\frac{(h_{*R} + h_R)}{h_R^2}}, \tag{2.130}$$

where $c_R = \sqrt{gh_R}$. Writing the shock expression (2.130) in terms of Froude numbers yields:

$$\mathsf{F}_R - \mathsf{F}_S = \sqrt{\frac{1}{2}h_{*R}\frac{(h_{*R} + h_R)}{h_R^2}}. \tag{2.131}$$

Equation (2.131) is quadratic in h_{*R}, and has roots given by:

$$h_{*R} = \frac{-h_R \pm h_R\sqrt{1 + 8(\mathsf{F}_R - \mathsf{F}_S)^2}}{2}. \tag{2.132}$$

However, $h_{*R} \geq h_R$ for a right shock, thus, expression (2.131) implies $\mathsf{F}_R - \mathsf{F}_S \geq 1$. Therefore, the relevant root of the quadratic (2.132) is given by:

$$h_{*R} = \frac{-h_R + h_R\sqrt{1 + 8(\mathsf{F}_R - \mathsf{F}_S)^2}}{2}. \tag{2.133}$$

31

Furthermore, applying the shock expression (2.130) and using (2.131) in the velocity expression (2.128) yields:

$$u_{*R} = u_R + c_R\Big(1 - \frac{h_R}{h_{*R}}\Big)(\mathsf{F}_R - \mathsf{F}_S).$$
(2.134)

$\square$

2.5 Summary: reactive flow modelling

Reactive flow refers to fluid flow that involves chemical reactions. Mathematical models for reactive flow processes are differential equations that are derived by the application of physical laws such as conservation of mass, momentum and energy.

In Section 2.1, a continuity equation, species transport and momentum equations were derived by applying a physical principle that governs mass and momentum transport in an elementary volume. Simplifying assumptions were applied to the flow model to derive Euler's model that is coupled with chemical transport equations. Further, the shallow water principle was applied to the Euler model and averaged over depth to yield reactive-shallow water flow model.

In Section 2.2, the hyperbolicity of the shallow water flow with chemical transport has been discussed. Due to non-linearity and hyperbolicity of the model, the solution of the model admits shocks (also known as hydraulic jumps), discontinuities and rarefactions, which are not usually present in linear flow problems.

In Section 2.3, characteristic fields were discussed where two of the three eigenvalues and corresponding eigenvectors have been shown to be linearly degenerate and one was shown to be genuinely non-linear.

In Section 2.4, a Riemann problem and its solution structure have been discussed. The general solution of the Riemann problem was constructed using wave relations such as the Rankine Hugoniot condition, entropy inequality and Riemann invariants. Shock-capturing numerical schemes will be presented in subsequent chapters for the efficient simulation of reactive flow processes.

Chapter 3

Modelling and simulating chemical kinetics

The kinetics of well-stirred chemical reactions are modelled and simulated in this chapter. Existence of nonnegative and conservative solutions of the resulting model have been discussed. A stoichiometric method for reducing the high degrees of freedom (of the model) and other model decoupling methods have been discussed. Stiffness-resolving numerical schemes that preserve nonnegativity and conservativity, and are compatible with the decoupling methods have been presented. The goal here, is to develop an efficient modelling and simulation procedure for simulating the kinetics of well-stirred chemical reaction systems. Numerical experiments have been performed to validate the model decoupling methods and to verify the numerical schemes.

Remark: *The discussion in this chapter has been published, see [3].*

3.1 Chemical kinetic modelling

Spatially independent chemical phenomena (referred to as chemical kinetic processes) characteristically occur in several steps called elementary reactions. A collection of all the elementary steps in a chemical process is called a mechanism. In general, chemical kinetics studies mechanisms and speed with which a particular chemical phenomenon proceeds [6, 7, 72, 81]. Balanced stoichiometric equations are used to represent mechanisms and polynomials called rate laws are used to quantify the speed with which the reaction

occurs. In some chemical processes, some elementary steps are reversible (thus involve forward and backward directions) while others are irreversible (involve only a forward direction). Since the irreversible steps are reversible steps where the backward direction vanishes, the reversible approach is adapted here to simplify the presentation, without loss of generality.

If there are N chemical species participating in M elementary steps of a particular chemical process, where $N_{f,r}$ species participate in the forward direction and $N_{b,r}$ species participate in the backward direction of the rth elementary step, then the mechanism can be expressed as follows [37, 3]:

$$\sum_{i=1}^{N_{f,r}} a_{r,i} A_{r,i} \rightleftarrows \sum_{j=1}^{N_{b,r}} b_{r,j} B_{r,j}, \quad r = 1, \ldots, M, \tag{3.1}$$
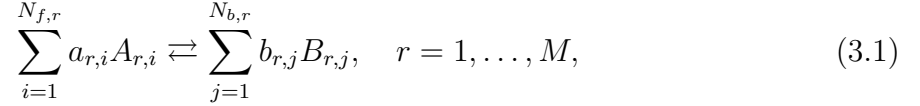
where $a_{r,i}$ and $b_{r,j}$ are stoichiometric coefficients for the species $A_{r,i}$ and $B_{r,j}$ in the rth elementary reaction. In general, $\sum_{r=1}^{M}(N_{f,r}+N_{b,r}) \neq N$ due to the presence of networking species (i.e. species engaged in more than one elementary reaction).

Denote the global set (i.e. set of all the species in the chemical process) whose cardinality is N by $\mathbf{C}$ and the vector of species concentrations by $\mathbf{U} \in \mathcal{R}_+^N$ (where $\mathcal{R}_+$ denotes non-negative real numbers). Denote the set of species in the rth elementary step whose cardinality is $N_r = N_{f,r} + N_{b,r}$, by $\mathbf{C}_r$. This set $\mathbf{C}_r \subseteq \mathbf{C}$ is defined by putting the species in the forward direction in the first $N_{f,r}$ positions of $\mathbf{C}_r$, followed by the species in the backward direction. Thus, by the stoichiometric equation (3.1),

$$\mathbf{C}_r = \{A_{r,1}, A_{r,2}, \ldots A_{r,N_{f,r}}, B_{r,1}, B_{r,2}, \ldots, B_{r,N_{b,r}}\},$$

and corresponding vector of concentrations $\mathbf{U}_r \in \mathcal{R}_+^{N_r}$, is given by:

$$\mathbf{U}_r = ([A]_{r,1}, [A]_{r,2}, \ldots [A]_{r,N_{f,r}}, [B]_{r,1}, [B]_{r,2}, \ldots, [B]_{r,N_{b,r}}), \quad r = 1, 2, \ldots, M,$$

where $[A]_{r,1}$ denotes concentration of species $A_{r,1}$. Moreover, if $\boldsymbol{\alpha}_r = (\alpha_{r,1}, \alpha_{r,2}, \ldots, \alpha_{N_r})$ is the set of orders for reactants and $\boldsymbol{\sigma}_r = (\sigma_{r,1}, \sigma_{r,2}, \ldots, \sigma_{r,N_r})$ is the set of stoichiometric numbers for the chemical species in $\mathbf{C}_r$, then the rate law for the rth elementary step (3.1) is defined by [37, 3]:

$$R_r(\mathbf{U}) = R_r(\mathbf{U}_r) = K_{fr} \prod_{i=1}^{N_{f,r}} U_{r,i}^{\alpha_{r,i}} - K_{b,r} \prod_{j=N_{f,r}+1}^{N_r} U_{r,j}^{\alpha_{r,j}}, \quad r = 1, \ldots, M \tag{3.2}$$

where $K_{f,r}$, $K_{b,r}$ are the forward and backward reaction constants, respectively. With rate law (3.2), the expression that governs the evolution of any species in the global set **C**, is given by [37, 3]:

$$\frac{dU_k}{dt} = \sum_{r=1}^{M} \sigma_{kr} R_r(\mathbf{U}), \quad k = 1, \ldots, N, \tag{3.3}$$

where t is the time variable, σ_{kr} is the stoichiometric number of the kth species in the rth elementary step. Thus, system (3.3) is a coupled system of N ordinary differential equations (ODEs) that must be solved to obtain species concentration profiles over time. A complete initial value problem (IVP) can be stated that, find $\mathbf{U} : \mathcal{R}_+ \longrightarrow \mathcal{R}_+^N$ such that:

$$\frac{d\mathbf{U}}{dt} = \sigma \mathbf{R}(\mathbf{U}(t), t), \quad t \in [0, T),$$
$$\mathbf{U}(t = 0) = \mathbf{U}_0, \tag{3.4}$$

where T is final time, $\sigma \in \mathcal{R}^{N \times M}$ is the stoichiometric matrix and $\mathbf{R} : \mathcal{R}^N \times \mathcal{R}_+ \longrightarrow \mathcal{R}^M$ is a vector of the M rate laws. The rest of the discussion in this section will assume that the stoichiometric matrix is non-trivial (i.e. each row or column contains at least one non-zero entry), is conservative (i.e. the vector $\mathbf{e} = (1, 1, \ldots, 1)^{Tr}$ (where Tr indicates transpose) is orthogonal to $range(\sigma)$) and has full rank (i.e. $rank(\sigma) = M$). The following conditions are imposed on the rate laws $\mathbf{R}$:

1. $\mathbf{U} = \mathbf{0}$ implies $R_r(\mathbf{U}, t) = 0$ and $\mathbf{U} > \mathbf{0}$ implies $R_r(\mathbf{U}, t) > 0$, for $t > 0$ and $r = 1, 2, \ldots, M$.

2. There exist $Z_r(\mathbf{U}, t) \in \mathcal{C}^0(\bar{\mathcal{R}}_+^N, \bar{\mathcal{R}}_+)$ if $\sigma_{kr} < 0$, such that $Z_r(\mathbf{U}, t) = 0$ if $\mathbf{U} = \mathbf{0}$, $Z_r(\mathbf{U}, t) > 0$ if $\mathbf{U} > \mathbf{0}$, and $R_r(\mathbf{U}, t) = Z_r(\mathbf{U}, t)U_k$, for $k = 1, 2, \ldots, N$.

3. $R_r(\mathbf{U}, t) \in \mathcal{C}^0(\bar{\mathcal{R}}_+^N, \bar{\mathcal{R}}_+)$, $R_r(\cdot, t)$ is locally Lipschitz continuous in $\mathcal{R}^N$, uniformly in t, and finally $R_r(\bar{\mathbf{U}}, \cdot) \in \mathcal{L}^\infty(\mathcal{R}_+)$ for $\bar{\mathbf{U}} \in \bar{\mathcal{R}}_+^N$.

Furthermore, with the conditions imposed on the stoichiometric matrix and reaction rate law, the existence, uniqueness and non-negativity of the solution of the IVP (3.4) can now be discussed as in [45, 71, 129]. The discussion here will make use of the following Lemma that presents a classical global existence result (given in [61]):

Lemma 3.1. *Assume that a function $\mathbf{G}(\mathbf{x}, t)$ is defined in a closed extended domain $\bar{Y}$ where $Y = \Omega \times (t_1, t_2)$ and $\Omega \in \mathcal{R}^N$. If $\mathbf{G}$ is continuous in an open subset $D \subset Y$, and that $\mathbf{G}$ is uniformly continuous in t and locally Lipschitz continuous with respect to $\mathbf{x}$ contained in D, further, if there are constants $K_1, K_2 \in \mathcal{R}_+$ such that:*

$$\|\mathbf{G}(\mathbf{x}, t)\| \leq K_1 + K_2 \|\mathbf{x}\|, \quad for \ all \ \ (\mathbf{x}, t) \in \bar{Y},$$

then for all initial conditions $(\mathbf{V}_0, t_0) \in Y$, there exist at least one solution to the IVP:

$$\frac{d\mathbf{V}(t)}{dt} = \mathbf{G}(\mathbf{V}(t), t), \quad t \in (t_1, t_2),$$

$$\mathbf{V}(t_0) = \mathbf{V}_0.$$

Furthermore, in the following discussion, the $p-$norm of a vector $\mathbf{V} \in \mathcal{R}^N$ will be denoted by:

$$\|\mathbf{V}\|_p = \Big(\sum_{k=1}^N |V_k|^p \Big)^{\frac{1}{p}},$$

and for any matrix $\mathbb{A} \in \mathcal{R}^{N \times M}$, the $p-$norm will be denoted by:

$$\|\mathbb{A}\|_p = \sup_{\|\mathbf{V}\|_p = 1, \mathbf{V} \in \mathcal{R}^M} \|\mathbb{A}\mathbf{V}\|_p.$$

Moreover, let a slope function be defined as $\mathbf{F}(\mathbf{U}, t) = \sigma \mathbf{R}(\mathbf{U}(t), t)$ and a constant be defined as $\Lambda = \|\mathbf{U}_0\|_1$ such that a convex set $\Omega \subset \mathcal{R}^N$ is defined as:

$$\Omega = \{\mathbf{V} \in \bar{\mathcal{R}}_+^N : \|\mathbf{V}\|_1 \leq \Lambda\}.$$

Trivially, the initial data $\mathbf{U}(t = 0) = \mathbf{U}_0 = \mathbf{0}$ if $\Lambda = 0$, thus by the first condition imposed on the rate law, $\mathbf{U}(t) = \mathbf{0}$ for all t. Therefore, we focus on the non-trivial case where $\Lambda > 0$. Consequently, the slope function is modified as follows:

$$\tilde{\mathbf{F}}(\mathbf{U}, t) = \mathbf{F}(\mathbf{H}(\mathbf{U}), t), \tag{3.5}$$

where $\mathbf{H}(\mathbf{U})$ is a unique projector of $\mathbf{U}$ on Ω defined as:

$$\mathbf{H}(\mathbf{U}) = \begin{cases} \mathbf{U}, & \text{if } \mathbf{U} \in \Omega, \\ \frac{\Lambda}{\|\mathbf{U})\|_1} \mathbf{U}, & \text{Otherwise.} \end{cases} \tag{3.6}$$

The uniqueness of $\mathbf{H}(\mathbf{U})$ is due to the convexity of Ω.

Proposition 3.1. *If the conditions imposed on the stoichiometric matrix and rate laws are satisfied, the projector* $\mathbf{H}$ *in* (3.6) *is Lipschitz continuous while the modified slope function* (3.5) *is bounded, Lipschitz continuous in the first argument and uniformly in* t.

Proof. Firstly, if $\mathbf{U}_1, \mathbf{U}_2 \in \Omega$ then

$$\mathbf{H}(\mathbf{U}_1) - \mathbf{H}(\mathbf{U}_2) = \mathbf{U}_1 - \mathbf{U}_2,$$

which implies

$$\|\mathbf{H}(\mathbf{U}_1) - \mathbf{H}(\mathbf{U}_2)\|_1 = \|\mathbf{U}_1 - \mathbf{U}_2\|_1.$$

And if $\mathbf{U}_1, \mathbf{U}_2 \notin \Omega$ then $\|\mathbf{U}_1\|_1 > \Lambda$ and $\|\mathbf{U}_2\|_1 > \Lambda$, thus,

$$\mathbf{H}(\mathbf{U}_1) - \mathbf{H}(\mathbf{U}_2) = \Lambda\Big(\frac{\mathbf{U}_1}{\|\mathbf{U}_1\|_1} - \frac{\mathbf{U}_2}{\|\mathbf{U}_2\|_1}\Big)$$

$$< \mathbf{U}_1 - \mathbf{U}_2.$$

Hence for any $\mathbf{U}_1, \mathbf{U}_2 \in \mathcal{R}^N$,

$$\|\mathbf{H}(\mathbf{U}_1) - \mathbf{H}(\mathbf{U}_2)\|_1 \leq \|\mathbf{U}_1 - \mathbf{U}_2\|_1$$

Secondly, for any $\mathbf{U} \in \mathcal{R}^N$, we have:

$$\sup_{t\in\mathcal{R}_+,\mathbf{U}\in\mathcal{R}^N} \|\tilde{\mathbf{F}}(\mathbf{U},t)\|_1 = \sup_{t\in\mathcal{R}_+,\mathbf{U}\in\mathcal{R}^N} \|\mathbf{F}(\mathbf{U},t)\|_1$$

$$\leq \|\sigma\|_1 \sup_{t\in\mathcal{R}_+,\mathbf{U}\in\mathcal{R}^N} \|\mathbf{R}(\mathbf{U}(t),t)\|_1 < \infty.$$

Further, for any $\mathbf{U}_1, \mathbf{U}_2 \in \mathcal{R}^N$,

$$\tilde{\mathbf{F}}(\mathbf{U}_1,t) - \tilde{\mathbf{F}}(\mathbf{U}_2,t) = \sigma\Big(\mathbf{R}(\mathbf{H}(\mathbf{U}_1),t) - \mathbf{R}(\mathbf{H}(\mathbf{U}_2),t)\Big).$$

Since $\mathbf{R}(\mathbf{U},t)$ is Lipschitz continuous there exists a Lipschitz constant ϱ, such that:

$$\|\mathbf{R}(\mathbf{U}_1,t) - \mathbf{R}(\mathbf{U}_2,t)\|_1 \leq \varrho\|\mathbf{U}_1 - \mathbf{U}_2.\|_1$$

Therefore, we have :

$$\|\tilde{\mathbf{F}}(\mathbf{U}_1,t) - \tilde{\mathbf{F}}(\mathbf{U}_2,t)\|_1 \leq \|\sigma\|_1 \|\mathbf{R}(\mathbf{H}(\mathbf{U}_1),t) - \mathbf{R}(\mathbf{H}(\mathbf{U}_2),t)\|_1$$

$$\leq \varrho\|\sigma\|_1 \|\mathbf{H}(\mathbf{U}_1) - \mathbf{H}(\mathbf{U}_2)\|_1.$$

Hence

$$\|\tilde{\mathbf{F}}(\mathbf{U}_1,t) - \tilde{\mathbf{F}}(\mathbf{U}_2,t)\|_1 \leq \varrho\|\sigma\|_1 \|\mathbf{U}_1 - \mathbf{U}_2\|_1.$$

$\square$

Proposition 3.2 (Existence, uniqueness and positivity of solution). *If the conditions imposed on the stoichiometric matrix σ and rate laws $\mathbf{R}$ in (3.4) are satisfied, then for any initial condition $\mathbf{U}(0) = \mathbf{U}_0$, the IVP:*

$$\frac{d\mathbf{U}(t)}{dt} = \tilde{\mathbf{F}}(\mathbf{U}(t), t), \quad t \in (0, T)$$

$$\mathbf{U}(0) = \mathbf{U}_0,$$

has a unique non-negative solution $\mathbf{U} \in [\mathcal{C}^1(\mathcal{R}_+)]^N$, where $\tilde{\mathbf{F}}$ is the modified slope function defined in (3.5). Moreover, for all $t \geq 0$, $\mathbf{U}(t) > \mathbf{0}$ if $\mathbf{U}_0 > \mathbf{0}$.

Proof. Firstly, the modified slope function $\tilde{\mathbf{F}}$ is continuous in the open domain $\mathcal{R}^N \times \mathcal{R}_+$, thus, setting $D = \mathcal{R}^N \times \mathcal{R}_+$ satisfies the first condition in Lemma 3.1. By Proposition 3.1, $\tilde{\mathbf{F}}$ is Lipschitz continuous with respect to $\mathbf{U}$ and uniformly in t, in the domain $\mathcal{R}^N \times \mathcal{R}_+$, thus satisfies the second condition of Lemma 3.1. Further, the boundedness of $\tilde{\mathbf{F}}$ (shown in Proposition 3.1) shows that there exists constants $K_1 = 0$ and $\|\sigma\|_1 \sup_{t \in \mathcal{R}_+, \mathbf{U} \in \mathcal{R}^N} \|\mathbf{R}(\mathbf{U}(t), t)\|_1 < \infty$, thus, the third condition is satisfied.

Secondly, it can be shown that the solution remains non-negative for all non-negative initial conditions. Trivially, if $\mathbf{U}(0) = \mathbf{0}$ the concentration profiles for all the species remain $\mathbf{U}(t) = \mathbf{0}$ for all $t \geq 0$, since the first condition on the rate law $\mathbf{R}$ in (3.4) is satisfied. To prove the non-trivial case, let J_k^- be the set of indices r for which $\sigma_{kr} < 0$ and J_k^+ be the set of indices r for which $\sigma_{kr} \geq 0$. Since the second condition on $\mathbf{R}$ is satisfied, there exist $Z_r(\mathbf{U}, t) \in \mathcal{C}^0(\bar{\mathcal{R}}_+^N, \bar{\mathcal{R}}_+)$ such that the governing evolution equations for the species can be reformulated as:

$$\frac{dU_k(t)}{dt} = \sum_{r \in J_k^-} \sigma_{kr} Z_r(\mathbf{U}(t), t) U_k + \sum_{r \in J_k^+} \sigma_{kr} R_r(\mathbf{U}(t), t), \quad k = 1, 2, \ldots, N. \tag{3.7}$$

Further, by setting $\varphi_k(t) = -\sum_{r \in J_k^-} \sigma_{kr} Z_r(\mathbf{U}(t), t)$ and $\phi_k(t) = \sum_{r \in J_k^+} \sigma_{kr} R_r(\mathbf{U}(t), t)$, the reformulation (3.7) becomes:

$$\frac{dU_k(t)}{dt} = -\varphi_k(t) U_k(t) + \phi_k(t), \quad k = 1, 2, \ldots, N. \tag{3.8}$$

Since the conditions imposed on the rate laws $\mathbf{R}$, are satisfied, the parameters $\varphi_k(t)$ and $\phi_k(t)$ in (3.8) (that are continuous function of t) are non-negative whenever $\mathbf{U}$ is

non-negative. Moreover, from (3.8), a further reformulation can be obtained as:

$$\int_0^t \frac{d}{d\theta}\left(e^{\int_0^\theta \varphi_k(\tau)d\tau}U_k(\theta)\right)d\theta = \int_0^t e^{\int_0^\theta \varphi_k(\tau)d\tau}\phi_k(\theta)d\theta, \quad k = 1, 2, \ldots, N. \qquad (3.9)$$

By setting

$$W_k(t) = e^{\int_0^t \varphi_k(\tau)d\tau}U_k(t), \quad k = 1, 2, \ldots, N. \qquad (3.10)$$

we obtain from (3.9) that:

$$W_k(t) = W_k(0) + \int_0^t e^{\int_0^\theta \varphi_k(\tau)d\tau}\phi_k(\theta)d\theta, \quad t > 0, \quad k = 1, 2, \ldots, N. \qquad (3.11)$$

Due to the non-negativity of the parameters (i.e. $\varphi_k \geq 0$ and $\phi_k \geq 0$ for $k = 1, 2, \ldots, N$) in expressions (3.10) and (3.11), the solution remains non-negative (i.e. $U_k(t) \geq 0$ for $t \in [0, T)$). Otherwise, there exist some species with index k, such that the concentration $U_k(0) \geq 0$ while $U_k(T) < 0$. If $U_k(0) \geq 0$ it follows from expression (3.10) that $W_k(T) < 0$, however, from expression (3.11) it follows that $W_k(T) > 0$, which is a contradiction. Further, all components whose initial concentrations were positive remain positive. Otherwise, there exist some component k, such that $U_k(0) > 0$ while $U_k(T) = 0$. If $U_k(0) \geq 0$ it follows from expression (3.10) that $W_k(T) = 0$, while from expression (3.11) it follows that $W_k(T) > 0$, which is a contradiction.

Finally, since $\tilde{\mathbf{F}}(\mathbf{U}, t) = \mathbf{F}(\mathbf{H}(\mathbf{U}), t)$ we have $\frac{d}{dt}\mathbf{H}(\mathbf{U}(t)) = \sigma\mathbf{R}(\mathbf{H}(\mathbf{U}), t)$, thus, due to the conservation property of the stoichiometric matrix, $\mathbf{e}^{Tr}\frac{d}{dt}\mathbf{H}(\mathbf{U}(t)) = 0$. Consequently, it follows that $\mathbf{e}^{Tr}\mathbf{H}(\mathbf{U}(t)) = \Lambda$, where $\Lambda = \mathbf{e}^{Tr}\mathbf{H}(\mathbf{U}(0)) \geq 0$. The non-negativity property implies that for all $t \geq 0$, $\|\mathbf{H}(\mathbf{U}(t))\|_1 = \Lambda$. Further, it follows that $\tilde{\mathbf{F}} = \mathbf{F}$ and $\mathbf{H}(\mathbf{U}(t)) = \mathbf{U}(t)$ along the solution curve. The solution is unique since it is contained in a compact set of $\mathcal{R}^N$. $\qquad \square$

Proposition 3.3 (Conservativity of solution). *The solution $\mathbf{U}$ of the IVP (3.4) conserves mass (i.e. $\mathbf{e}^{Tr}\mathbf{U} = $ constant) if the stoichiometric matrix is conservative (i.e. $\mathbf{e} \in Ker(\sigma^{Tr})$). Conversely, if the rate law conditions (i.e. three conditions state under (3.4)) are satisfied, $\mathbf{e}^{Tr}\mathbf{U}$ is constant and $rank(\sigma) = M$, then $\mathbf{e} \in Ker(\sigma^{Tr})$.*

Proof. If the stoichiometric matrix is conservative (i.e. $\mathbf{e} \in Ker(\sigma^{Tr})$) we have $\mathbf{e}^{Tr}\sigma = \mathbf{0}$.

Thus, from the IVP (3.4), we obtain:

$$\mathbf{e}^{Tr}\frac{d\mathbf{U}}{dt} = \mathbf{e}^{Tr}\sigma\mathbf{R}(\mathbf{U}(t), t),$$

$$= 0.$$

It then follows that the total mass $\mathbf{e}^{Tr}\mathbf{U} = constant$ is preserved in all time $t \geq 0$. Conversely, if the solution preserves mass, then we have $\mathbf{e}^{Tr}\mathbf{U}(t) = \Lambda$, for all $t \geq 0$, where Λ is a constant. Thus, we have $\mathbf{e}^{Tr}\frac{d}{dt}\mathbf{U} = 0$, which implies from the IVP (3.4) that $\mathbf{e}^{Tr}\sigma\mathbf{R}(\mathbf{U}(t), t) = 0$. Since the stoichiometric matrix has full rank (i.e. $rank(\sigma) = M$) and all the conditions imposed on the rate law are satisfied $\mathbf{R} \neq 0$, it follows that $\mathbf{e}^{Tr}\sigma = 0$. Hence, we have $\mathbf{e} \in Ker(\sigma^{Tr})$. □

3.2 Model decoupling methods

The total number of chemical species N that participate in a particular chemical kinetic process is usually very large, thus, the rate law (3.3) is an $N-$dimensional polynomial where N is very large. Further, in many chemical processes several elementary processes are involved, thus, the total number of elementary reactions (i.e M) is also very large. The large number of species coupled with the large number of elementary processes result in models with large degrees of freedom and subsequently, result in complicated numerical algorithms. Moreover, the large degrees of freedom is due to coupling of the N species. Thus, although very few (chemical species) profiles are usually of interest in laboratory experiments and numerical simulation studies, the governing IVP for the species of interest can not be solved without applying decoupling methods. In this section, some decoupling methods (that are more appropriate for chemical kinetic processes) are presented.

3.2.1 Stoichiometric decoupling method

During chemical kinetic processes, mass (concentration) of a particular species increases when the species is formed from other species, and decreases when it is converted into other species. The mass increment or loss of a particular species is referred to as mass (or concentration) transformed, and an expression that balances the mass of a particular

chemical species at any given time is referred to as mass balanced expression [45, 171]. Let $\mathbf{U}_0$ represent a vector of initial concentrations, $\mathbf{U}_T$ represent the vector of transformed species concentrations and $\mathbf{U}_S$ quantify mass added or removed from the system. Mass balance expression for a chemical system can be written as follows:

$$\mathbf{U} = \mathbf{U}_0 + \mathbf{U}_S + \mathbf{U}_T. \tag{3.12}$$

If mass of a species increase, the transformed concentrations has a positive value for that particular species, but has a negative value for all species that contributed to the mass increment of that particular species. A species transformed concentration is the product of the its stoichiometric number and the extent of reaction [7]. Denote the extent of reaction by χ, then the mass balance expression for any species i, in a single reaction can be written as follows:

$$U_i = U_{0_i} + \sigma_i \chi + U_{S_i}, \quad i = 1, 2, \ldots, N. \tag{3.13}$$

If a constant concentration of the species is introduced or removed during the reaction process, U_{S_i} is constant. Thus, applying the mass balance expression (3.13) in the ODE (3.4) and manipulating algebraically yields:

$$\frac{d\chi}{dt} = K_f \prod_{i=1}^{N_f} \left(U_{0_i} + \sigma_i \chi + U_{S_i} \right)^{\alpha_i} - K_b \prod_{j=N_f+1}^{N} \left(U_{0_j} + \sigma_j \chi + U_{S_j} \right)^{\alpha_j}. \tag{3.14}$$

Observe in ODE (3.14) that the extent of reaction χ, is the only variable that must be solved for. Once the extent of reaction is known, the mass balance expression (3.13) is employed to compute concentration profiles for all the species. It is sometimes feasible to solve ODE (3.14) analytically, however, numerical procedures are employed in the general case due to model non-linearity issues.

Most natural systems are not closed, thus, chemical systems are able to exchange materials with their surroundings at any time. Due to the exchange of materials U_{S_k} is time-dependent, thus, ODE (3.14) does not hold in the general case. In such time dependent case, an expression for the extent of reaction can be obtained from the mass

41

balance expression (3.13) as follows [37, 3]:

$$\chi = -\frac{1}{\sigma_k}\left(U_{0_k} + U_{S_k}\right) + \frac{1}{\sigma_k}U_k, \ k \in \{1, \dots, N\}. \tag{3.15}$$

Applying the extent of reaction (3.15) in the mass balance expression (3.13), the following holds:

$$U_i = d_i + \frac{\sigma_i}{\sigma_k}U_k, \ i = 1, 2, \dots, N, \ k \in \{1, \dots, N\}, \ k \neq i, \tag{3.16}$$

where

$$d_i = U_{0_i} + U_{S_i} - \frac{\sigma_i}{\sigma_k}\left(U_{0_k} + U_{S_k}\right).$$

Using expression (3.16) in ODE (3.3) yields a decoupled set of ODEs that govern the evolution of chemical species in a mechanism that involves one elementary step. The decoupled ODEs have slope functions that are polynomials in one variable (i.e. concentration U_k of the kth chemical species) that are written explicitly as follows:

$$\frac{1}{\sigma_k}\frac{dU_k}{dt} = K_f \prod_{i=1}^{N_f}\left(d_i + \frac{\sigma_i}{\sigma_k}U_k\right)^{\alpha_i} - K_b \prod_{j=N_f+1}^{N}\left(d_j + \frac{\sigma_j}{\sigma_k}U_k\right)^{\alpha_j}, \ k \in \{1, \dots, N\}. \tag{3.17}$$

Moreover, it is an easy task to extend the reformulation to multi-step reactions. Since the reaction proceeds in multi-steps, species mass balances are obtained for each species i, in the rth step as follows:

$$U_{r,i} = U_{0_{r,i}} + \sigma_{r,i} \ \chi_r + U_{S_{r,i}}, \ i = 1, 2, \dots, N_r, \ r = 1, \dots, M \tag{3.18}$$

where χ_r is extent of the rth reaction, $U_{0_{r,i}}$ denotes initial concentration and $U_{S_{r,i}}$ denotes sources/sinks. If n denotes the index of the species of interest in the subset $\mathbf{U}_r$ of chemical species in the rth reaction, then it follows from mass balance expression (3.18), that the extent of the rth reaction is given by:

$$\chi_r = -\frac{1}{\sigma_{r,n}}\left(U_{0_{r,n}} + U_{S_{r,n}}\right) + \frac{1}{\sigma_{r,n}}U_{r,n}, \ n \in \{1, \dots, N_r\}. \tag{3.19}$$

Applying the extent of reaction (3.19) in the mass balance expression (3.18) yields:

$$U_{r,i} = d_{r,i} + \frac{\sigma_{r,i}}{\sigma_{r,n}}U_{r,n}, \ i = 1, \dots, N_r, \ r = 1, \dots, M, \ n \in \{1, \dots, N_r\} \tag{3.20}$$

where

$$d_{r,i} = U_{0_{r,i}} + U_{S_{r,i}} - \frac{\sigma_{r,i}}{\sigma_{r,n}}\Big(U_{0_{r,n}} + U_{S_{r,n}}\Big).$$

Consequently, it follows from the rate law (3.2) and expression (3.20) that the stoichiometrically decoupled set of ODEs is given by:

$$\frac{dU_k}{dt} = \sum_{r=1}^{M} \sigma_{kr}\Big(K_{f,r} \prod_{i=1}^{N_{f,r}} \Big(d_{r,i} + \frac{\sigma_{r,i}}{\sigma_{kr}}U_k\Big)^{\alpha_{r,i}} - K_{b,r} \prod_{j=1+N_{f,r}}^{N_r} \Big(d_{r,j} + \frac{\sigma_{r,j}}{\sigma_{kr}}U_k\Big)^{\alpha_{r,j}}\Big),$$

$$k \in \{1, \dots, N\}. \quad (3.21)$$

3.2.2 Other decoupling methods

Mathematical formulations of most dynamical systems often lead to ODEs (whether linear or non-linear) in the generalized form [176]:

$$\frac{d\mathbf{U}}{dt} = \mathbf{F}(\mathbf{U}, t), \qquad t \in [0, T), \qquad \mathbf{U}(0) = \mathbf{U}_0, \qquad (3.22)$$

where $\mathbf{U} \in \mathcal{R}^N$ and $\mathbf{F} : \mathcal{R}^N \times \mathcal{R}_+$. Several research efforts have been made to determine how the generalized time dependent ODE (3.22) can be solved efficiently [176]. Solution procedures for (3.22) can be classified into two, namely, direct and decomposition/iterative methods [176, 115, 109]. With the direct approach, all the differential equations in the system are discretized identically with the same integration method and simulated within common constraints [176]. However large systems usually introduce further constraints such as stiffness, where some components propagate faster than others. Thus, the direct approach is inefficient for large systems [176, 115, 109]. Instead, the system can be decomposed (decoupled) into subsystems where appropriate techniques are employed for each subsystem [96]. According to the authors in [96] decomposition can be achieved by *tearing* or through *relaxation*. With the tearing method, decomposition takes advantage of the structure (e.g. block diagonal, sparsity etc.) of the Jacobian of the slope function $\mathbf{F}$, thus, computational complexity largely depends on the structure of the Jacobian [96]. Relaxation approaches on the other hand, reduce complexity without regards to the structure of the Jacobian [96].

With the relaxation approach, the ODEs are decoupled from each other before numerical schemes are applied. This approach (also known as Waveform relaxation) has been

applied extensively to decouple models in circuit theory and wave propagation problems [176, 115, 109]. In this section, the waveform relaxation methods are employed to decouple the system (3.22), that models chemical processes. Since the approach decouples the individual ODEs in the system, it is convenient to express ODE system (3.22) in component form as follows:

$$\frac{dU_k}{dt} = F_k(\mathbf{U}, t), \qquad t \in [0, T), \qquad \mathbf{U}(0) = \mathbf{U}_0, \quad k = 1, 2, \ldots, N. \qquad (3.23)$$

Observe in (3.23) that concentrations $\mathbf{U}$ of all the species must be available in order to evaluate the slope function F_k. Thus, ODE (3.23) is still coupled with the other ODEs in the system (3.22). Among several techniques developed to decouple (3.23), Picard-Lindelov methods that include Successive-Over-Relaxation, Gauss-Jacobi and Gauss-Seidel techniques, will be considered here. Applying Gauss-Jacobi relaxation approach to (3.23) yields the following continuous-time iteration [14]:

$$\begin{cases} \frac{d}{dt} U_k^{i+1} = F_k(U_1^i, \ldots, U_{k-1}^i, U_k^{i+1}, U_{k+1}^i, \ldots, U_N^i, t), \\ U_k^{i+1}(0) = U_{0,i}, \quad k = 1, 2, \ldots, N, \quad t \in [0, T), \quad i = 0, 1, \ldots. \end{cases} \qquad (3.24)$$

Similarly, by employing the Gauss-Seidel waveform relaxation to decouple, ODE (3.23) yields the following continuous-time iteration [14]:

$$\begin{cases} \frac{d}{dt} U_k^{i+1} = F_k(U_1^{i+1}, \ldots, U_{k-1}^{i+1}, U_k^{i+1}, U_{k+1}^i, \ldots, U_N^i, t), \\ U_k^{i+1}(0) = U_{0,i}, \quad k = 1, 2, \ldots, N, \quad t \in [0, T), \quad i = 0, 1, \ldots. \end{cases} \qquad (3.25)$$

Finally, by employing Successive-Over-Relaxation (SOR) to decouple, ODE (3.23) yields the following continuous-time iteration [14]:

$$\begin{cases} \frac{d}{dt} \bar{U}_k^{i+1} = F_k(U_1^{i+1}, \ldots, U_{k-1}^{i+1}, \bar{U}_k^{i+1}, U_{k+1}^i, \ldots, U_N^i, t), \\ \bar{U}_k^{i+1}(0) = U_{0,i}, \\ U_k^{i+1} = \omega U_k^i + (1 - \omega) \bar{U}_k^{i+1}, \quad k = 1, 2, \ldots, N, \quad t \in [0, T), \quad i = 0, 1, \ldots, \end{cases} \qquad (3.26)$$
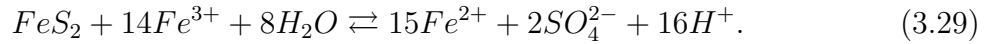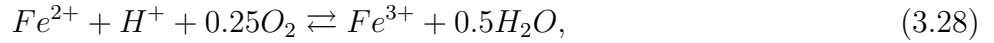
where ω is a real parameter. If the parameter $\omega = 0$ the Successive-Over-Relaxation (SOR) approach reverts to the Gauss-Seidel waveform relaxation.

3.3 Application: acid mine generation and neutralization

Among many other chemical systems acidic mine generation and neutralization will be presented here and used as application problems throughout this thesis.

3.3.1 Acid generation (pyrite oxidation)

During mining of minerals, pyrite is exposed in large quantities to air and water. This exposure leads to oxidation and subsequent acidification of water bodies. Further, the acidic water has high solubility which enables it to dissociate other minerals that contain metal ions. The final result is an environmental pollutant that has severe adverse effects on ecological systems. The balanced stoichiometric equations for the pollutant generation are as follows [94, 50, 147]:

$$FeS_2 + 3.5O_2 + H_2O \rightleftarrows Fe^{2+} + 2SO_4^{2-} + 4H^+, \tag{3.27}$$

$$Fe^{2+} + H^+ + 0.25O_2 \rightleftarrows Fe^{3+} + 0.5H_2O, \tag{3.28}$$

$$FeS_2 + 14Fe^{3+} + 8H_2O \rightleftarrows 15Fe^{2+} + 2SO_4^{2-} + 16H^+. \tag{3.29}$$

Thus, the respective rate laws for the balanced stoichiometric equations (3.27)-(3.29) can be written out as follows:

$$R_{O_2} = K_{f1}[FeS_2][O_2] - K_{b1}[Fe^{2+}][H^+][SO_4^{2-}], \tag{3.30}$$

$$R_{Fe^{3+}} = K_{f2}[FeS_2][Fe^{3+}] - K_{b2}[Fe^{2+}][H^+][SO_4^{2-}], \tag{3.31}$$

$$R_{Fe^{2+}} = K_{f3}[O_2][Fe^{2+}]. \tag{3.32}$$

It follows from stoichiometric equations (3.27)-(3.29) and rate laws (3.30)-(3.32) that the governing equations for the species that are participating in the acid generation, are given by the following autonomous system of ODEs:

$$\frac{d\mathbf{U}}{dt} = \mathbf{F}(\mathbf{U}), \tag{3.33}$$

where $\mathbf{U} = \left([H^+], [FeS_2], [O_2], [Fe^{2+}], [SO_4^{2-}], [Fe^{3+}]\right)^{Tr}$ and $\mathbf{F}(\mathbf{U}) = \left(2R_{O_2} + 16R_{Fe^{3+}} - R_{Fe^{2+}}, \quad -R_{O_2} - R_{Fe^{3+}}, \quad -3.5R_{O_2} - 0.25R_{Fe^{2+}}, \quad R_{O_2} + 15R_{Fe^{3+}} - R_{Fe^{2+}}, \quad 2R_{O_2} + 2R_{Fe^{3+}}, \quad -14R_{Fe^{3+}} + R_{Fe^{2+}}\right)^{Tr}$.

The amount of acidity of the water is determined by the amount of hydrogen ion generated, thus, the species of interest here is the hydrogen ion H^+. However, the ODE for the hydrogen ion in system (3.33) cannot be isolated and solved alone, due to coupling (with other species concentrations) through the slope function $\mathbf{F}(\mathbf{U})$. Even if the waveform decoupling methods are employed the entire system (3.33) (with six degrees of freedom) must be solved in each time step (this has computational cost implications). This problem can be avoided if the stoichiometric method is applied to decouple (3.33) instead of the waveform approach. If the stoichiometric method is applied to system (3.33), the ODE for the hydrogen ion becomes:

$$\frac{d[H^+]}{dt} = D_1 + D_2[H^+] + D_3[H^+]^2 + D_4[H^+]^3, \tag{3.34}$$

where $\mathbf{U}_0 = \left([H^+]_0, [FeS_2]_0, [O_2]_0, [Fe^{2+}]_0, [SO_4^{2-}]_0, [Fe^{3+}]_0\right)^{Tr}$ is initial data, $\mathbf{U}_S = \left([H^+]_S, [FeS_2]_S, [O_2]_S, [Fe^{2+}]_S, [SO_4^{2-}]_S, [Fe^{3+}]_S\right)^{Tr}$ is secondary sources,

$C_{20S} = [FeS_2]_0 + [FeS_2]_S + 0.5([H^+]_0 + [H^+]_S)$, $C_{30S} = [O_2]_0 + [O_2]_S + \frac{7}{4}([H^+]_0 + [H^+]_S)$, $C_{40S} = [Fe^{2+}]_0 + [Fe^{2+}]_S - 0.5([H^+]_0 + [H^+]_S)$, $C_{50S} = [SO_4^{2-}]_0 + [SO_4^{2-}]_S - ([H^+]_0 + [H^+]_S)$, $C_{202S} = [FeS_2]_0 + [FeS_2]_S + \frac{1}{16}([H^+]_0 + [H^+]_S)$, $C_{404S} = [Fe^{2+}]_0 + [Fe^{2+}]_S - \frac{15}{16}([H^+]_0 + [H^+]_S)$, $C_{505S} = [SO_4^{2-}]_0 + [SO_4^{2-}]_S - \frac{2}{16}([H^+]_0 + [H^+]_S)$, $C_{606S} = [[Fe^{3+}]_0 + [[Fe^{3+}]_S + \frac{14}{16}([H^+]_0 + [H^+]_S)$, $C_{03S} = [O_2]_0 + [O_2]_S - 0.25([H^+]_0 + [H^+]_S)$, $C_{04S} = [Fe^{2+}]_0 + [Fe^{2+}]_S - ([H^+]_0 + [H^+]_S)$,

$$D_1 = 2.9K_{f1}C_{20S}C_{30S} + 23.2K_{f2}C_{202S}C_{606S} - 1.45K_{f3}C_{04S}C_{03S}$$
$$D_2 = -K_{f1}\left(\frac{7}{2}C_{20S} + C_{30S}\right) - 2K_{b1}C_{40S}C_{50S} - K_{f2}\left(14C_{202S} + C_{606S}\right)$$
$$\quad - K_{b2}\left(2C_{404S} + 15C_{505S}\right) - K_{f3}\left(0.25C_{04S} + C_{03S}\right)$$
$$D_3 = \frac{7}{4}K_{f1} - K_{b1}\left(2C_{40S} + C_{50S}\right) + \frac{14}{16}K_{f2} - K_{b2}\left(2C_{404S} + 15C_{505S}\right) - 0.25K_{f3}$$
$$D_4 = -K_{b1} - \frac{30}{16}K_{b2}.$$

It can be observed that ODE (3.34) is completely decoupled and thus, can be solved independently to obtain the hydrogen ion profile at all times. Thus, instead of solving

six ODEs in each time step, only one ODE will be solved if the stoichiometric method is applied to decouple.

3.3.2 Acid neutralization: limestone

Due to the abundance of limestone, several neutralization techniques have been developed for remediating acidic mines. The calcite in limestone neutralizes the hydrogen ions if the limestone comes into contact with the acidic water. The balanced stoichiometric equations for the calcite neutralization are as follows [130, 126]:

$$CaCO_3 + H^+ \rightleftarrows Ca^{2+} + HCO_3^-, \tag{3.35}$$

$$CaCO_3 + H_2CO_3 \rightleftarrows Ca^{2+} + 2HCO_3^-, \tag{3.36}$$

$$CaCO_3 + H_2O \rightleftarrows Ca^{2+} + HCO_3^+ + OH^-. \tag{3.37}$$
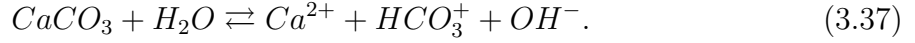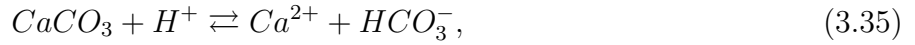
Let K_{H^+} denote forward reaction constant in (3.35), $K_{H_2CO_3}$ denotes forward reaction constant in (3.36) and K_{H_2O} denotes forward reaction constant in (3.36). If $K_{Ca^{2+}}$ is the backward reaction rate constant then the experimentally determined rate law for the entire calcite neutralization is given by [130, 126]:

$$\begin{aligned} R_{CaCO_3} = K_{H^+}[H^+]^* + K_{H_2CO_3}[H_2CO_3]^* + K_{H_2O}[H_2O]^* \\ - K_{Ca^{2+}}[Ca^{2+}]^*[H_2CO_3^-]^*, \end{aligned} \tag{3.38}$$

where symbol $[]^*$ indicates ion activity. Since the earlier presentation involved concentration (but not activities), it is appropriate to convert the activities into concentration. Further, data available in [126] shows that $K_{H_2CO_3}[H_2CO_3]^* + K_{H_2O}[H_2O]^*$ is negligible, thus will be neglected here. Thus, by neglecting the terms $[H_2CO_3]^* + K_{H_2O}[H_2O]^*$ in (3.38) and following discussion in [156, 155, 94] on activity-concentration conversion yields:

$$R_{CaCO_3} = F([H^+], [Ca^{2+}], [HCO_3^-]) = K_f[H^+] - K_b[Ca^{2+}][HCO_3^-], \tag{3.39}$$

where $K_f = K_{H^+}\eta_{H^+}$, $K_b = K_{Ca^{2+}}\eta_{HCO_3^-}\eta_{Ca^{2+}}$ and $\eta_{HCO_3^-}, \eta_{Ca^{2+}}\ \eta_{H^+}$ are activity coefficients.

Therefore, the relevant balanced stoichiometric equation for the calcite neutralization is (3.35) and the rate at which the neutralization occurs is given by (3.39). The following information can be extracted from (3.35) and (3.39):

$$\mathbf{C} = (CaCO_3, \ H^+, \ Ca^{2+}, \ HCO_3^-), \ \mathbf{U} = ([CaCO_3], \ [H^+], \ [Ca^{2+}], \ [HCO_3^-]),$$

$$N_f = 2, \ N = 4, \ \mathbf{U}_0 = ([CaCO_3]_0, \ [H^+]_0, \ [Ca^{2+}]_0, \ [HCO_3^-]_0), \ \mathbf{U}_S = ([CaCO_3]_S,$$

$$[H^+]_S, \ [Ca^{2+}]_S, \ [HCO_3^-]_S), \alpha = (0, \ 1, \ 1, \ 1), \ \text{and} \ \beta = (-1, \ -1, \ 1, \ 1). \tag{3.40}$$

Applying data (3.40) in the generalized species reaction ODE system (3.3) with rate law (3.3), yields the governing equations for the species $CaCO_3, \ H^+, \ Ca^{2+}, \ HCO_3^-$, given by:

$$-\frac{d[CaCO_3]}{dt} = K_f[H^+] - K_b[Ca^{2+}][HCO_3^-], \tag{3.41}$$

$$-\frac{d[H^+]}{dt} = K_f[H^+] - K_b[Ca^{2+}][HCO_3^-], \tag{3.42}$$

$$\frac{d[Ca^{2+}]}{dt} = K_f[H^+] - K_b[Ca^{2+}][HCO_3^-], \tag{3.43}$$

$$\frac{d[HCO_3^-]}{dt} = K_f[H^+] - K_b[Ca^{2+}][HCO_3^-]. \tag{3.44}$$

The goal in the neutralization process is to reduce to acceptable levels the high concentration of hydrogen ions in the water. Thus, the species of interest in the calcite neutralization, is the hydrogen ion which is governed by ODE (3.42) in the system (3.41) -(3.44). However, ODE (3.42) cannot be isolated and solved due to coupling in the slope function. Instead of solving ODE system (3.41) -(3.44) (with four degrees of freedom) in each time step, the stoichiometric method can be applied to significantly reduce computational cost.

If it is assumed (just for simplification) that there is no secondary source for any species (i.e. $[CaCO_3]_S = 0, \ [H^+]_S = 0, \ [Ca^{2+}]_S = 0$ and $[HCO_3^-]_S = 0$), then applying the stoichiometric method (3.21) using data (3.40) yields the following decoupled

set of ODEs governing the time evolution of species:

$$-\frac{d[CaCO_3]}{dt} = -K_b[CaCO_3]^2 + \lambda_1[CaCO_3] - \lambda_2, \tag{3.45}$$

$$-\frac{d[H^+]}{dt} = -K_b[H^+]^2 + \lambda_3[H^+] - \lambda_4, \tag{3.46}$$

$$\frac{d[Ca^{2+}]}{dt} = \lambda_5 - \lambda_6[Ca^{2+}] - K_b[Ca^{2+}]^2, \tag{3.47}$$

$$\frac{d[HCO_3^-]}{dt} = \lambda_7 - \lambda_8[HCO_3^-] - K_b[HCO_3^-]^2. \tag{3.48}$$

where

$[CaCO_3]_H = [H^+]_0 - [CaCO_3]_0, \; [CaCO_3]_C = [H^+]_0 + [CaCO_3]_0,$
$[CaCO_3]_{HC} = [H^+]_0 + [CaCO_3]_0, \quad \lambda_1 = K_f + K_b\Big([CaCO_3]_C + [CaCO_3]_{HC}\Big),$
$\lambda_2 = K_f[CaCO_3]_H - K_b\Big([CaCO_3]_C + [CaCO_3]_{HC}\Big), \lambda_3 = K_f + K_b\Big([H^+]_0 + [HCO_3^-]_0\Big) +$
$K_b\Big([H^+]_0 + [Ca^{2+}]_0\Big), \lambda_4 = K_b\Big([H^+]_0 + [HCO_3^-]_0\Big)\Big([H^+]_0, + [Ca^{2+}]_0\Big), \quad \lambda_5 = K_f\Big([H^+]_0 +$
$[Ca^{2+}]_0\Big), \quad \lambda_6 = \Big(K_f + K_b[H^+]_0 - K_b[Ca^{2+}]_0,\Big), \lambda_7 = K_f\Big([H^+]_0 + [HCO_3^-]_0\Big),$
$\lambda_8 = \Big(K_f + K_b[H^+]_0 - K_b[HCO_3^-]_0\Big).$

It can be observed that all ODEs (3.45)-(3.48) are completely uncoupled and thus, each ODE can be solved independently to obtain profiles for all species at all times. Instead of solving four ODEs in each time step, only one ODE will be solved if the stoichiometric method is applied to decouple (3.41)-(3.44).

3.3.3 Analytical results: calcite system

Eventually, the forward (i.e. dissolution) and backward (i.e. precipitation) reactions of the calcite system (3.35) reaches equilibrium, where the rate of precipitation balances with the rate of dissolution. The ratio of a forward and backward reaction of a system in equilibrium is a constant called equilibrium constant. Let K_{eq} be the equilibrium constant of (3.35), then from rate law (3.41):

$$K_{\text{eq}} = \frac{[Ca^{2+}][HCO_3^-]}{[H^+]} = \frac{K_f}{K_b}. \tag{3.49}$$

The analytical expression (3.49) is very useful for analysing the system that is in equilibrium.

Moreover,, with data (3.40), an exact solution for the calcite system (3.45)-(3.30) can

easily be derived. Define some constants μ_1 and μ_2 as follows:

$$\mu_1 = K_f[H^+]_0 - K_b[Ca^{2+}]_0[HCO_3^-]_0 \text{ and } \mu_2 = -(K_f + K_b[Ca^{2+}]_0 + K_b[HCO_3^-]_0).$$

The exact solution of system (3.45)-(3.30) is given by:

$$[CaCO_3] = [CaCO_3]_0 - \frac{\Theta_1\Theta_2\left(1 - \exp^{-K_b(\Theta_1-\Theta_2)t}\right)}{\Theta_2 - \Theta_1\exp^{-K_b(\Theta_1-\Theta_2)t}}, \tag{3.50}$$

$$[H^+] = [H^+]_0 - \frac{\Theta_1\Theta_2\left(1 - \exp^{-K_b(\Theta_1-\Theta_2)t}\right)}{\Theta_2 - \Theta_1\exp^{-K_b(\Theta_1-\Theta_2)t}}, \tag{3.51}$$

$$[Ca^{2+}] = [Ca^{2+}]_0 + \frac{\Theta_1\Theta_2\left(1 - \exp^{-K_b(\Theta_1-\Theta_2)t}\right)}{\Theta_2 - \Theta_1\exp^{-K_b(\Theta_1-\Theta_2)t}}, \tag{3.52}$$

$$[HCO_3^-] = [HCO_3^-]_0 + \frac{\Theta_1\Theta_2\left(1 - \exp^{-K_b(\Theta_1-\Theta_2)t}\right)}{\Theta_2 - \Theta_1\exp^{-K_b(\Theta_1-\Theta_2)t}}, \tag{3.53}$$

where

$$\Theta_1 = \frac{-\mu_2 - \sqrt{\mu_2^2 + 4K_b\mu_1}}{-2K_b} \text{ and } \Theta_2 = \frac{-\mu_2 + \sqrt{\mu_2^2 + 4K_b\mu_1}}{-2K_b}.$$

Furthermore, in pure water the calcite rate law (3.39) can be approximated by polynomials in two variables. In pure water, the calcite precipitation rate has second order dependence on the species Ca^{2+} and HCO_3^-, thus [4]:

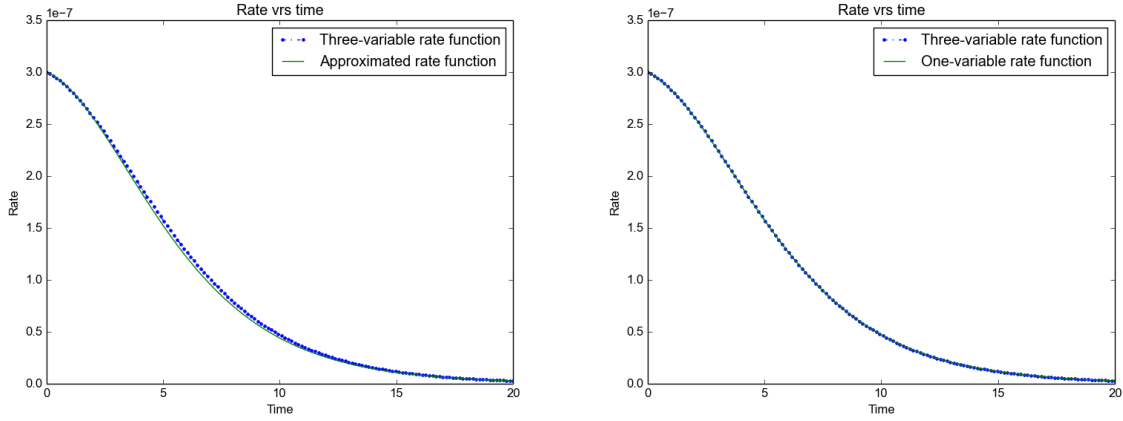$$K_b[Ca^{2+}][HCO_3^-] \approx 2K_b\left([Ca^{2+}]\right)^2, \tag{3.54}$$

or

$$K_b[Ca^{2+}][HCO_3^-] \approx 2K_b\left([HCO_3^-]\right)^2. \tag{3.55}$$

Consequently, applying approximations (3.54) and (3.55) in the rate law (3.38) of the calcite dissolution-precipitation reaction (3.35) yields the following two variable polynomials:

$$R_{CaCO_3} = K_f[H^+] - 2K_b\left([Ca^{2+}]\right)^2, \tag{3.56}$$

and

$$R_{CaCO_3} = K_f[H^+] - 2K_b\left([HCO_3^-]\right)^2. \tag{3.57}$$

(a) Two-variable rate law   (b) One-variable rate law

Figure 3.1: Rate profiles for the three-variable rate law (3.38), two-variable rate laws (3.56) -(3.57) and one variable rate law (3.45)-(3.48). These profiles were computed with the aid of the analytical solution (3.50)-(3.53).

In Figure 3.1, the original calcite rate law (3.39) that is a three-variable polynomial is compared with approximations (3.56) - (3.57) (that are two-variable polynomials) and rate laws (3.45)-(3.48) (that are one-variable polynomials obtianed with the stoichiometric method). It is clear from Figure 3.1 that, the one-variable rate law that has been obtained using the stoichiometric decoupling method, is more accurate than the other approximations. Moreover, the pure water case where the two-variable rate law applies, rarely occurs and since the acid mine drainage problem occurs in the natural environment, approximations (3.55) - (3.56) are not always useful. Further, apart from accuracy issues it is computationally expensive to employ the pure water approximations in simulation studies, due to the fact that models resulting from such approximations have more degrees of freedom.

3.4 Numerical schemes

In general, analytical solution to the IVP (3.4) are not possible due to complications (e.g. non-linearity and stiffness [95, 28, 174, 35]) associated with the rate laws. Numerical schemes serve as alternatives to exact solutions, however, only numerical schemes with implicit [110] features are able to resolve stiffness efficiently [135]. Suitable implicit schemes

for IVP (3.4) must preserve the non-negativity and conservation properties as well as bound the model reduction errors introduced by model reduction methods [135, 133, 64]. In this section, fully implicit schemes are considered and presented as efficient schemes for simulating the reduced model.

Even after the application of model reduction methods to (3.3), the equations governing the evolution of species with time, can be written in the general form:

$$\frac{d\mathbf{U}}{dt} = \mathbf{F}(\mathbf{U}, t), \qquad t \in [0, T), \tag{3.58}$$
$$\mathbf{U}(0) = \mathbf{U}_0.$$

If N_t is the total number of time grid points, then the time step size $\Delta t = \frac{T}{N_t}$, the nth time in the discrete time interval is $t^n = n\Delta t$, and $t^{n+1} = t^n + \Delta t$. Thus, the first fully implicit class of schemes (called theta schemes) for (3.58) states that for $\theta \in [0, 1]$:

$$\mathbf{U}^{n+1} - \Delta t\theta\mathbf{F}(\mathbf{U}^{n+1}, t^{n+1}) = \mathbf{U}^n + \Delta t(1 - \theta)\mathbf{F}(\mathbf{U}^n, t^n) \tag{3.59}$$

where $\mathbf{U}^{n+1} = \mathbf{U}(t^n + \Delta t)$ and $\mathbf{U}^n = \mathbf{U}(t^n)$.

If $\theta = 0$ theta scheme (3.59) is the first order consistent and conditionally stable explicit Euler scheme also known as forward Euler scheme. If $\theta = 1$, the theta scheme (3.59) is the first order consistent and unconditionally stable implicit Euler scheme also know as backward Euler scheme. If $\theta = \frac{1}{2}$ the theta scheme (3.59) is the second order consistent and unconditionally stable Crank-Nicholson scheme. Convergence of these schemes have been discussed extensively in [71].

Furthermore, another robust class of scheme for problem (3.58) is the diagonally implicit Runge Kutta (DIRK) class of schemes. A second order diagonally implicit Runge-Kutta (named DIRK2) that is parametrized by α states that:

$$\mathbf{U}^* = \mathbf{U}^n + \Delta t\mathbf{F}(\mathbf{U}^*, t^*),$$
$$\mathbf{U}^{n+1} - \alpha\Delta t\mathbf{F}(\mathbf{U}^{n+1}, t^{n+1}) = \frac{2\alpha - 1}{\alpha}\mathbf{U}^n + \frac{1 - \alpha}{\alpha}\mathbf{U}^*. \tag{3.60}$$

Among other values the parameter $\alpha = 1 \pm \frac{\sqrt{2}}{2}$ will be considered in this discussion. The following lemma will be useful in the positivity discussions (involving theta and DIRK2 schemes) to proceed.

Proposition 3.4. *If the conditions imposed on the stoiciometric matrix σ and rate laws* **R** *in* (3.4) *are satisfied and $\sigma\mathbf{R}(\mathbf{U}, t) = \mathbb{Z}(\mathbf{U}, t)\mathbf{U}$, where $\mathbb{Z}(\mathbf{U}, t)$ is an $N \times N$ matrix that satisfies*

$$\mathbb{Z}_{kk} \leq 0, \quad \mathbb{Z}_{kr} \geq 0 \text{ if } k \neq r, \quad and \quad \sum_{k=1}^{N} \mathbb{Z}_{kr} = 0,$$

then the matrix $\mathbb{B} = \mathbb{I} - \kappa_a \mathbb{Z}$ (where $\kappa_a > 0$) is an invertible M-matrix.

Proof. Since all the non-diagonal elements are non-positive, matrix $\mathbb{B}$ is a Z-matrix. Further, all the diagonal elements are positive. Also, Matrix $\mathbb{B}$ is diagonally dominant implying that, the real parts of its eigenvalues are positive (strictly). Hence $\mathbb{B}$ is an invertible M-matrix. $\qquad\square$

Proposition 3.5 (Convervativity of schemes)**.** *If the conditions imposed on the stoiciometric matrix σ and rate laws* **R** *in* (3.4) *are satisfied then the theta scheme* (3.59) *and DIRK2* (3.60) *are conservative.*

Proof. Firstly, since $\mathbf{F}(\mathbf{U}, t) = \sigma\mathbf{R}(\mathbf{U}, t)$ in the IVP (3.4), the theta scheme states that in one time step:

$$\mathbf{U}^{n+1} - \Delta t \theta \sigma \mathbf{R}(\mathbf{U}^{n+1}, t^{n+1}) = \mathbf{U}^n + \Delta t (1 - \theta) \sigma \mathbf{R}(\mathbf{U}^n, t^n), \quad \text{for } n = 0, 1, \ldots. \quad (3.61)$$

Since the conservative condition imposed on the stoichiometric matrix σ is satisfied, $\mathbf{e}^{Tr}\sigma = 0$, thus, we obtain from the theta scheme (3.60) that $\mathbf{e}^{Tr}\mathbf{U}^{n+1} = \mathbf{e}^{Tr}\mathbf{U}^n$ for $n = 0, 1, \ldots$. Implying that $\mathbf{e}^{Tr}\mathbf{U}^n = \mathbf{e}^{Tr}\mathbf{U}_0$ for any $n \geq 0$. Hence, the theta class of schemes for the IVP (3.4) are conservative for any value of θ.

Secondly, DIRK2 scheme for IVP (3.4) states that in one time step:

$$\mathbf{U}^* = \mathbf{U}^n + \Delta t \sigma \mathbf{R}(\mathbf{U}^*, t^*),$$
$$\mathbf{U}^{n+1} - \alpha \Delta t \sigma \mathbf{R}(\mathbf{U}^{n+1}, t^{n+1}) = \frac{2\alpha - 1}{\alpha}\mathbf{U}^n + \frac{1 - \alpha}{\alpha}\mathbf{U}^*. \quad (3.62)$$

By conservativity of the stoichiometric matrix, we obtain from (3.62) that:

$$\mathbf{e}^{Tr}\mathbf{U}^* = \mathbf{e}^{Tr}\mathbf{U}^n,$$
$$\mathbf{e}^{Tr}\mathbf{U}^{n+1} = \frac{2\alpha - 1}{\alpha}\mathbf{e}^{Tr}\mathbf{U}^n + \frac{1 - \alpha}{\alpha}\mathbf{e}^{Tr}\mathbf{U}^*. \quad (3.63)$$

It follows that $\mathbf{e}^{Tr}\mathbf{U}^n = \mathbf{e}^{Tr}\mathbf{U}_0$, for $n \geq 0$. Hence, DIRK2 conserves the total mass in all time. $\qquad\square$

Proposition 3.6 (Positivity of theta schemes). *Assume the conditions imposed on the stoiciometric matrix σ and rate laws $\mathbf{R}$ in (3.4) are satisfied and that there exist an $N \times N$ matrix $\mathbb{Z}(\mathbf{U}, t)$ satisfying the properties in proposition 3.4, such that $\mathbf{F}(\mathbf{U}, t) = \mathbb{Z}(\mathbf{U}, t)\mathbf{U}$. If the minimum component of the slope function $\mathbf{F}$ evaluated at time t^n is given by $F_i = \sum_r \mathbb{Z}_{ir} U_r^n = \min_{1 \le k \le N} \sum_r \mathbb{Z}_{kr} U_r^n$ such that some threshold time step is defined by:*

$$
\Delta t_c = \begin{cases}
\frac{U_i}{|F_i|} & \theta = 0 \ \text{and} \ F_i < 0 \\
\infty & \theta = 1 \ \text{or} \ F_i \ge 0 \\
\frac{U_i}{(1-\theta)|F_i|} & \text{otherwise}
\end{cases}
$$

then the theta scheme (3.59) preserves non-negative solution if $\Delta t < \Delta t_c$.

Proof. Firstly, if the $N \times N$ matrix $\mathbb{Z}(\mathbf{U}, t)$ exists $\mathbf{F}(\mathbf{U}, t) = \mathbb{Z}(\mathbf{U}, t)\mathbf{U}$, thus, the theta class of schemes for IVP (3.4) state that:

$$
\left(\mathbb{I} - \Delta t \theta \mathbb{Z}(\mathbf{U}^{n+1}, t^{n+1}) \right) \mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t (1 - \theta) \mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n, \quad \text{for } n = 0, 1, \ldots. \quad (3.64)
$$

A scheme is non-negativity preserving if $\mathbf{U}^{n+1} \ge 0$ whenever $\mathbf{U}^n \ge 0$. Consider a case when $\theta = 0$ (called explicit Euler or forward Euler scheme), scheme (3.64) reduces to:

$$
\mathbf{U}^{n+1} = \mathbf{U}^n + \Delta t \mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n, \quad \text{for } n = 0, 1, \ldots. \quad (3.65)
$$

Consequently, if $\mathbf{U}^n \ge 0$ and $\mathbf{F}(\mathbf{U}^n, t^n) = \mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n \ge 0$ it follows from the right hand side of (3.65) that $\mathbf{U}^{n+1} \ge 0$ for any time step $\Delta t < \Delta t_c = \infty$. However, if $\mathbf{U}^n \ge 0$ and $\mathbf{F}(\mathbf{U}^n, t^n) = \mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n \not\ge 0$ then there exist $F_i = \mathbb{Z}_{ir} U_r^n < 0$ which is the minimun component of $\mathbf{F}$. Expressing (3.65) with regards to the minimum component yields:

$$
U_i^{n+1} = U_i^n + \Delta t F_i, \quad \text{for } n = 0, 1, \ldots. \quad (3.66)
$$

Therefore, for non-negativity of (3.66) the time step must satisfy the constraint $\Delta t < \frac{U_i}{|F_i|}$.

Secondly, the case where $\theta \ne 0$ yields a scheme where a positive solution is sought for the non-linear problem (3.64). Let $\Lambda_\theta = \|\mathbf{U}^n\|_1$ be a constant such that a convex compact set is defined by:

$$
\Omega_\theta = \{ \mathbf{V} \in \mathcal{R}^N : \|\mathbf{V}\|_1 \le \Lambda_\theta, \ \mathbf{V} \ge 0 \}.
$$

Further, let $\mathbb{B}_\theta(\mathbf{V}) = \mathbb{I} - \theta \Delta t \mathbb{Z}(\mathbf{V}, t^{n+1})$ matrix that is defined on Ω_θ. Due to convexity and compactness of Ω_θ, matrix $\mathbb{B}_\theta(\mathbf{V})$ is well-defined. Moreover, by Proposition (3.4) $\mathbb{B}_\theta(\mathbf{V})$ is a non-singular M-matrix. Since $\mathbb{B}_\theta(\mathbf{V})$ is well-defined, the following fixed point function is well-defined in Ω_θ :

$$\Phi_\theta(\mathbf{V}) = \mathbb{B}_\theta^{-1}(\mathbf{V})\Big(\mathbf{U}^n + \Delta t(1-\theta)\mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n\Big), \quad \mathbf{V} \in \Omega_\theta, \quad \text{for } n = 0, 1, \dots \quad (3.67)$$

Any fixed point of (3.67) is also a solution to the non-linear problem (3.64). Moreover, the constructed fixed point function (3.67) satisfies:

$$\Phi_\theta(\mathbf{V}) - \theta \mathbb{Z}(\mathbf{V}, t^n)\Phi_\theta(\mathbf{V}) = \mathbf{U}^n + \Delta t(1-\theta)\mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n. \quad (3.68)$$

By exploiting the properties of matrix $\mathbb{Z}$, it follows from (3.68) that $\mathbf{e}^{Tr}\Phi_\theta(\mathbf{V}) = \mathbf{e}^{Tr}\mathbf{U}^n = \Lambda_\theta$, thus, $\|\Phi_\theta(\mathbf{V})\|_1 = \Lambda_\theta$. Furthermore, if $\mathbf{U}^n \geq 0$ and $\theta = 0$ (i.e. the implicit or backward Euler scheme) from (3.67) it follows that $\Phi_\theta(\mathbf{V}) \geq 0$ for any time step $\Delta t < \infty$, since $\mathbb{B}_\theta(\mathbf{V})$ is well-defined. If $\mathbf{U}^n \geq 0$, $\theta \neq 0$ and $\mathbf{F} = \mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n \geq 0$ then $\mathbf{U}^n + \Delta t(1-\theta)\mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n \geq 0$ for any time step $\Delta t < \infty$. Otherwise, if $\mathbf{F} = \mathbb{Z}(\mathbf{U}^n, t^n)\mathbf{U}^n < 0$ then there exist a negative minimum component of $\mathbf{F}$ (i.e. $F_i = \mathbb{Z}_{ir}U_r^n < 0$). Thus, the corresponding expression $U_i^n + \Delta t(1-\theta)\mathbb{Z}_{ir}U_r^n$ is non-negative if $\mathbf{U}^n \geq 0$ and $\Delta t < \frac{U_i}{(1-\theta)|F_i|}$. It follows from (3.67) that $\Phi_\theta(\mathbf{V}) \geq 0$ for any time step $\Delta t < \frac{U_i}{(1-\theta)|F_i|}$, whenever $\theta \neq 0$ and $\mathbf{U}^n \geq 0$.

Therefore, it is clear that $\Phi_\theta : \Omega_\theta \longrightarrow \Omega_\theta$, thus, employing Brouwer's fixed point theorem it can be concluded that there exist a fixed point $\mathbf{V} \in \Omega_\theta$ that is a non-negative solution to the non-linear problem (3.64). Hence, the theta class of schemes (3.59) preserve non-negativity if $\Delta t < \Delta t_c$. $\qquad\square$

Proposition 3.7 (Positivity of DIRK2 schemes)**.** *If there exist an $N \times N$ matrix $\mathbb{Z}(\mathbf{U}, t)$ satisfying the properties in proposition 3.4 such that $\mathbf{F}(\mathbf{U}, t) = \mathbb{Z}(\mathbf{U}, t)\mathbf{U}$ then the DIRK2 scheme (3.60) preserves non-negativity for any time step $\Delta t < \infty$.*

Proof. Since the $N \times N$ matrix $\mathbb{Z}(\mathbf{U}, t)$ exists and $\mathbf{F}(\mathbf{U}, t) = \mathbb{Z}(\mathbf{U}, t)\mathbf{U}$, the DIRK2 class of schemes for IVP (3.4) can be written as follows:

$$\mathbf{U}^* - \Delta t \mathbb{Z}(\mathbf{U}^*, t^*)\mathbf{U}^* = \mathbf{U}^n,$$

$$\mathbf{U}^{n+1} - \alpha \Delta t \mathbb{Z}(\mathbf{U}^{n+1}, t^{n+1})\mathbf{U}^{n+1} = \frac{2\alpha - 1}{\alpha}\mathbf{U}^n + \frac{1-\alpha}{\alpha}\mathbf{U}^{,*} \quad \text{for } n = 0, 1, \dots \quad (3.69)$$

By proposition 3.4, the matrix $\mathbb{I} - \Delta t \mathbb{Z}(\mathbf{U}^*, t^*)$ is a non-singular M-matrix. As a result $\mathbf{U}^* \geq 0$ whenever $\mathbf{U}^n \geq 0$. Further, $\mathbb{I} - \alpha \Delta t \mathbb{Z}(\mathbf{U}^{n+1}, t^{n+1})$ is also a non-singular M-matrix, thus, $\mathbf{U}^{n+1} \geq 0$ for any time step $\Delta t < \infty$ and for any value of the parameter α whenever $\mathbf{U}^n > 0$. Hence the DIRK2 scheme (3.60) preserves non-negativity without constraints on the time step size. □

3.5   Numerical experiments

Numerical experiments are conducted in this section to verify the numerical schemes and model decoupling techniques that have been presented in the chapter. Accuracy tests are performed on the numerical schemes to check convergence and then followed by simulation cost experiments to check the robustness of the decoupling schemes.

3.5.1   Convergence test

All the decoupling schemes (i.e. stoichiometric, Gauss-Jacobi, Successive-Over-Relaxation (SOR) and Gauss-Seidel decoupling methods) were applied to decouple the acid generation model (3.33) and calcite model (3.41)-(3.44), and then followed by an application of the theta schemes (3.59) and DIRK2 schemes (3.60). Since the calcite model has an exact solution, results are presented for the calcite model only. If the schemes are compatible with the decoupling methods, the schemes must converge with grid refinement. An excellent combination of the numerical schemes and decoupling methods must converge with the appropriate order of the numerical scheme.

Since the Successive-Over-Relaxation (SOR) has a parameter ω, an experiment has been conducted to determine suitable values for the SOR method. Table (3.1) shows the results of the experiments, it can be observed that any value of $\omega \leq 0.09$ yields accurate results for the SOR method. In the rest of the experiments the parameter is fixed at $\omega = 0.09$. *Theoretical results for $\omega \in [0, 1]$ have been discussed in [176], therefore, the experimental behaviour here is not surprising.*

Firstly, convergence tests were performed on a combination of Crank-Nicholson scheme (i.e. $\theta = \frac{1}{2}$) and all the decoupling schemes. Figures 3.2a and 3.2a shows the results of the experiments. The results indicate a decreasing hydrogen ion concentration across all

the decoupling schemes. However, compared with the exact solution it is observed that the stoichiometric method is more accurate than the other methods. This observation is confirmed by the L_∞ and L_2 error profiles displayed by Figures 3.3a and 3.3b.

Secondly, convergence tests were performed on a combination of Backward Euler scheme (i.e. $\theta = 1$) and all the decoupling schemes. Figures 3.2c and 3.2d display the results of the experiments. The results indicate a decreasing hydrogen ion concentration across all the decoupling schemes. However, it can be observed that the stoichiometric method is more accurate than the other methods. This observation is confirmed by the L_∞ and L_2 error profiles displayed by Figures 3.3c and 3.3d.

Thirdly, convergence tests were performed on a combination of DIRK2 scheme (with $\alpha = 1 + 0.5\sqrt{2}$) and all the decoupling schemes. Results of the experiments are displayed in Figures 3.2e and 3.2f. The results indicate a decreasing hydrogen ion concentration across all the decoupling schemes. However, it can be observed that the stoichiometric method is more accurate than the other methods. This observation is confirmed by the L_∞ and L_2 error profiles displayed by Figures 3.3e and 3.3f.

Furthermore, orders of convergence have been computed for the schemes as applied across the decoupling methods. Taylor expansion shows that the Crank-Nicholson scheme and DIRK2 schemes are second order consistent while the Backward Euler scheme is first order. Table 3.2 shows the results for the combination of each numerical scheme and the stoichiometric method. It can be observed that all the schemes converge according to their appropriate orders when combined with the stoichiometric method. However, Tables 3.3 -3.5 show that the schemes do not converge according to their appropriate orders when combined with the other decoupling methods (i.e. Gauss-Jacobi, Gauss-Seidel and SOR).

3.5.2   Cost of simulation

Experiments were conducted to check cost of simulating the calcite model with all combinations of the numerical schemes and the decoupling methods. CPU time, CPU time differences and relative CPU time were used to measure cost of simulations across different grid resolutions. The CPU time differences and relative CPU time were computed using the expressions:

ω	0.9	0.7	0.5	0.09	0.001	0.005	0.0001
DIRK2	5.0999	2.7917	1.6003	0.6807	0.6562	0.6563	0.6562
Backward Euler	5.1278	2.8275	1.6120	0.5877	0.5345	0.5357	0.5342
Crank-Nicholson	5.010	2.7917	1.6001	0.6824	0.6582	0.6583	0.6582

Table 3.1: Errors ($\| \cdot \|_2$) of the *backward Euler, DIRK2, and Crank-Nicholson* schemes applied to SOR decoupled calcite model (3.41)-(3.44). The errors where measured across varying values of the parameter ω, $N_t = 20$, $T = 25$ with all other model parameters held constant. The actual values in the table are 7 orders of magnitude smaller.

$$\text{CPU time difference} = CPU_C - CPU_S \qquad (3.70)$$

and

$$\text{Relative CPU time} = \frac{CPU_C - CPU_S}{CPU_C} \qquad (3.71)$$

where CPU_S is CPU time for the stoichiometric method and CPU_C is CPU time for any of the other methods. The final time was fixed at ($T = 20$) and a very fine grid (8000 time steps) were used to conduct the experiments. Note that the fine grid resolution was selected on purpose to reduce the effect of discretization errors on the final simulated results.

Crank-Nicholson scheme was applied to discretize after all the decoupling methods (i.e. Gauss-Seidel, SOR, stoichiometric and Gauss-Jacobi methods) have been applied to decoupled the calcite model (3.41) -(3.44). CPU time and CPU time differences were computed using (3.70) and (3.71). Figures 3.4e and 3.4f display CPU time for all the decoupling schemes against time steps. The general observation is that CPU time for all the decoupling methods increased across increasing time steps. However, the CPU time of the stoichiometric decoupling method is far smaller than the other methods. Further, the observed increasing profile of the CPU difference in Figure 3.4f indicates that the cost incurred by the stoichiometric–Crank-Nicholson approach reduces with grid refinement, relative to the other methods.

Secondly, the Backward Euler scheme was applied to discretize after all the decoupling methods (i.e. Gauss-Seidel, SOR, stoichiometric and Gauss-Jacobi methods) have been

Stoichiometric method

Numerical scheme	N_x	$\| \cdot \|_\infty$		$\| \cdot \|_2$	
		Error	Order	Error	Order
	10	0.0452	-	0.0809	-
	20	0.0110	2.0388	0.0273	1.5674
DIRK2	40	0.0027	2.0265	0.0095	1.5229
	80	0.0007	1.9475	0.0033	1.5255
	160	0.0002	2.0418	0.0012	1.4594
	10	1.1901	-	2.2598	-
	20	0.6448	0.8841	1.6520	0.6182
Backward Euler	40	0.3331	0.9533	1.1892	0.4743
	80	0.1694	0.9756	0.8487	0.4867
	160	0.0854	0.9881	0.6030	0.4931
	10	0.0830	-	0.1213	-
	20	0.0206	2.0100	0.0425	1.5130
Crank-Nicholson	40	0.0052	2.0141	0.0150	1.5025
	80	0.0013	2.0000	0.0053	1.5009
	160	0.0003	2.1155	0.0019	1.4800

Table 3.2: Errors and Orders of *DIRK2, backward Euler and Crank-Nicholson* schemes applied to stoichiometrically decoupled model (3.46), computed across norms and time steps. The actual error values in the table are 7 orders of magnitude smaller

applied to decoupled the calcite model (3.41)-(3.44). CPU time and CPU time differences were computed using (3.70). Results of the experiments are displayed in Figures 3.4c and 3.4d. The observations are similar to those obtained with the Crank-Nicholson scheme.

Thirdly, Figures 3.4a and 3.4b display the CPU times and CPU time differences measured in the simulation involving combinations of DIRK2 scheme and all the decoupling methods. The observations are very different from those obtained in the Crank-Nicholson and Backward Euler cases.

Finally, relative CPU times were computed with expression (3.71) for all the combina-

Gauss-Jacobi method

Numerical scheme	N_x	$\lVert \cdot \rVert_\infty$		$\lVert \cdot \rVert_2$	
		Error	Order	Error	Order
	10	3.2968	-	8.6763	-
	20	1.6860	0.9675	6.1866	0.4879
DIRK2	40	0.8528	0.9833	4.3950	0.4932
	80	0.4289	0.9915	3.1154	0.4964
	160	0.2151	0.9957	2.2057	0.4981
	10	3.8491	-	9.4241	-
	20	2.0191	0.9308	6.9436	0.6182
Backward Euler	40	0.6742	1.5825	5.0227	0.4743
	80	0.5246	0.3620	3.5942	0.4867
	160	0.2641	0.3432	2.5571	0.4931
	10	3.2776	-	8.6408	-
	20	1.6809	0.9996	6.1728	0.4852
Crank-Nicholson	40	0.8515	0.9811	4.3898	0.4918
	80	0.4286	0.9904	3.1135	0.4972
	160	0.2150	0.9967	2.2050	0.4977

Table 3.3: Errors and Orders of *DIRK2, backward Euler and Crank-Nicholson* schemes applied to Gauss-Jacobi decoupled calcite model (3.41)-(3.44), computed across norms and time steps. The actual error values in the table are 7 orders of magnitude smaller.

tions of numerical schemes and decoupling methods. Figure 3.5 displays the results of the experiments. It can be observed in Figures 3.5a and 3.5b that more than *76 percent* CPU time will be saved if stoichiometric decoupling method is applied with *DIRK2* scheme to solve the calcite model. Over *73 percent* of CPU time will be saved if the stoichiometric method is combined with Crank-Nicholson scheme. Figures 3.5c and 3.5d show that over *69 percent* of CPU time will be saved if the stoichiometric method is combined with the *Backward Euler* scheme.

Gauss-Seidel method

Numerical scheme	N_x	$\| \cdot \|_\infty$		$\| \cdot \|_2$	
		Error	Order	Error	Order
DIRK2	10	3.6351	-	9.7662	-
	20	1.7694	1.0387	6.5624	0.5736
	40	0.8736	1.0182	4.5266	0.5358
	80	0.4341	1.0089	3.1618	0.5177
	160	0.2164	1.0043	2.2228	0.5083
Backward Euler	10	2.6798	-	7.6458	-
	20	1.3459	0.9936	5.3422	0.5172
	40	0.6742	0.9973	3.7529	0.5094
	80	0.3374	0.9987	2.6448	0.5049
	160	0.1688	0.9991	1.8670	0.5024
Crank-Nicholson	10	3.6661	-	9.8386	-
	20	1.7759	1.0457	6.5824	0.5827
	40	0.8751	1.0211	4.5329	0.5381
	80	0.4964	0.8180	3.1639	0.5187
	160	0.2165	1.1971	2.2280	0.5060

Table 3.4: Errors and Orders of *DIRK2, backward Euler and Crank-Nicholson* schemes applied to Gauss-Seidel decoupled calcite model (3.41)-(3.44), computed across norms and time steps. The actual error values in the table are 7 orders of magnitude smaller.

3.6   Chapter summary

Numerical simulation of chemical kinetic processes is complicated by large numbers of chemical species participating in several elementary reactions. The large number of chemical species result in models with high degrees of freedom that are expensive to simulate. Further, large numbers also result in stiffness that pose a severe constraint to some numerical schemes. Concentrations of chemical species are non-negative thus, governing models and solution procedures must ensure non-negativity. Solution procedures should be constructed to reduce cost of simulation significantly without lost of accuracy. This is

Successive-Over-Relaxation method

Numerical scheme	N_x	$\|\cdot\|_\infty$		$\|\cdot\|_2$	
		Error	Order	Error	Order
DIRK2	10	3.2090	-	9.1567	-
	20	1.6931	1.7483	6.8069	0.4278
	40	1.1076	0.6122	5.6656	0.2647
	80	0.8807	0.3307	5.4667	0.0519
	160	0.7847	0.1665	5.1357	0.0902
Backward Euler	10	2.6576	−	7.3989	−
	20	1.6142	0.7193	5.8768	0.3323
	40	1.1254	0.5203	5.1733	0.1840
	80	0.9020	0.3193	5.0287	0.0410
	160	0.7977	0.1773	5.0267	0.0006
Crank-Nicholson	10	3.2370	-	9.2198	-
	20	1.6958	0.9327	6.8237	0.4341
	40	1.1078	0.6143	5.6704	0.2671
	80	0.8807	0.3310	5.4680	0.0524
	160	0.7848	0.1663	5.1360	0.0903

Table 3.5: Errors and Orders of *DIRK2, backward Euler and Crank-Nicholson* schemes applied to Successive-Over-Relaxation decoupled calcite model (3.41)-(3.44), computed across norms and time steps. The actual error values in the table are 7 orders of magnitude smaller.

achieved by combining numerical schemes with model reduction/decoupling techniques.

In Section 3.1, mathematical modelling of chemical kinetics has been provided. A system of Ordinary Differential Equations (ODEs) have been derived to describe the evolution of chemical species with time. Further, with the conditions provided in the section, discussion on existence, uniqueness and non-negativity of solution for the ODE system have been provided.

In Section 3.2, Model reduction/decoupling techniques have been presented for the ODE system derived in Section 3.1. The stoichiometric method has been developed based

on stoichiometry and mass balances. This method is most suitable when few of the many species profiles are of interest. Other decoupling techniques such as the Successive-Over-Relaxation (SOR), Gauss-Seidel and Gauss-Jacobi methods have also been presented.

In Section 3.3, chemical kinetic systems called acid mine generation and neutralization have been presented as application problems. The stoichiometric method was applied to both kinetic systems to significantly reduce their degrees of freedom. Further, the accuracy of the method has been tested (using the calcite rate law). Results showed that the stoichiometric method is very accurate.

In Section 3.4, numerical schemes for simulating the ODE system presented in Section 3.3 has been discussed. Due to the stiffness constraint, only implicit schemes are considered. Theta class of schemes and a second order diagonally implicit Runge Kutta class of schemes (DIRK2) have been shown to be non-negativity preserving and satisfy the conservation principle of the ODE system in Section 3.1.

In Section 3.5, numerical experiments have been conducted to verify the numerical schemes, validate the decoupling methods and to check efficiency of the decoupling methods. One notable observation is that, the solution procedures do not converge to their theoretical orders in the $\| \cdot \|_2$ norm. We suspect that the $\| \cdot \|_2$ norm is incompatible with the solution procedures (i.e. schemes coupled with decoupling methods), however, further studies will be conducted in future to confirm. However, the results showed that all the schemes are compatible with the stoichiometric decoupling method and that the stoichiometric method can significantly reduce simulation cost and maintain high accuracy.

Therefore, a combination of the non-negativity preserving schemes (i.e. implicit theta and DIRK2 schemes) with the stoichiometric decoupling method provides an efficient alternative tool for modelling and simulating chemical kinetic processes.
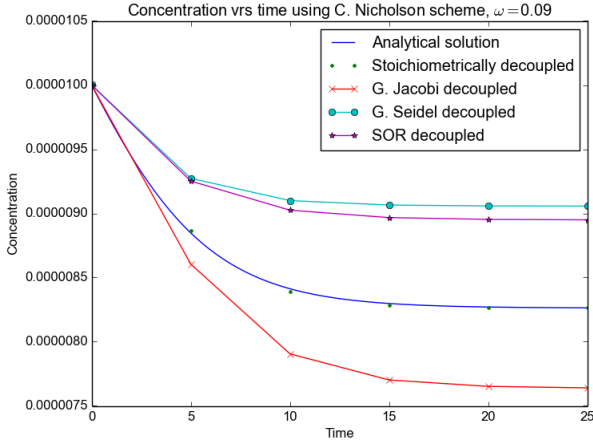
(a) Crank-Nicholson/five time steps.
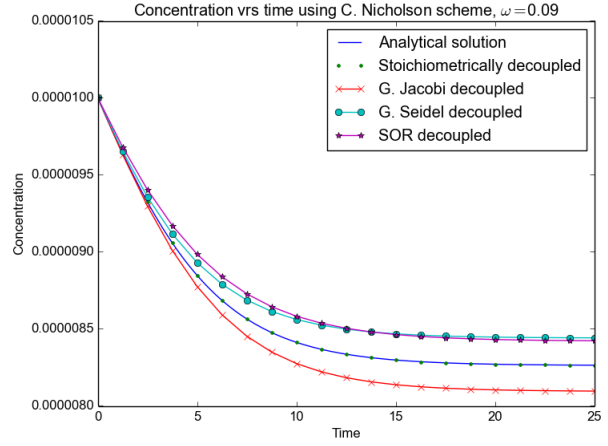
(b) Crank-Nicholson/twenty time steps.
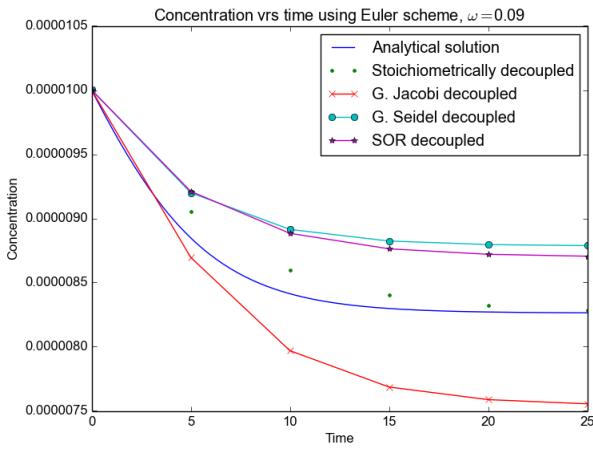
(c) Backward Euler/five time steps.
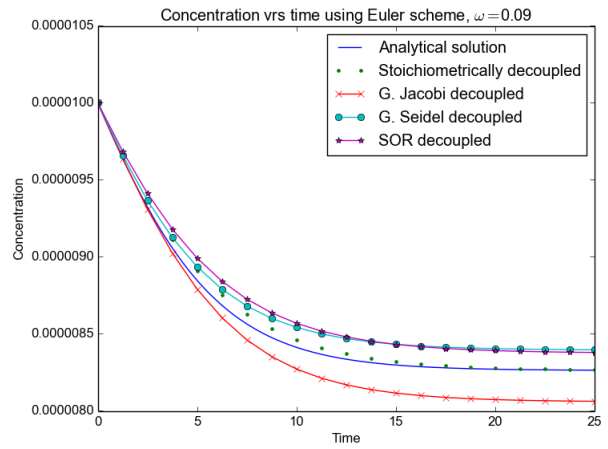
(d) Backward Euler/twenty time steps.
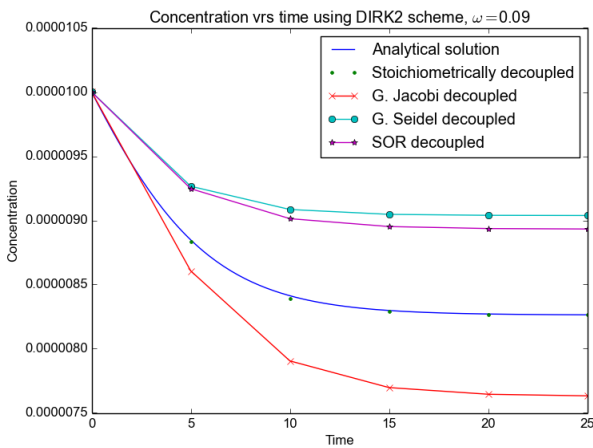
(e) DIRK2/five time steps.
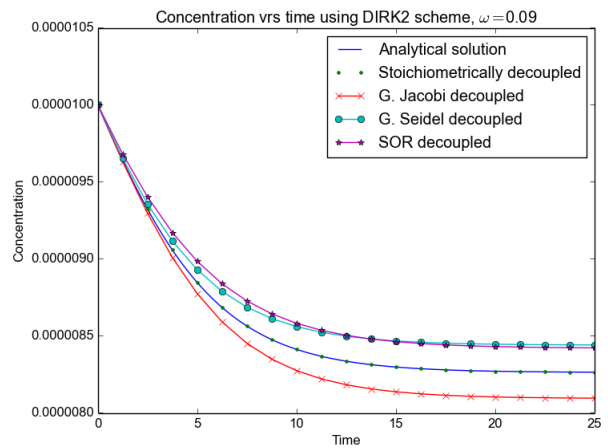
(f) DIRK2/twenty time steps.

Figure 3.2: Numerical and analytical solutions for ODE (3.41), where $U_k = [H^+], T = 25, [H^+]_0 = 10^{-5}$, and $F_k(t, [H^+]) = -K_b[H^+]^2 + \lambda_3[H^+] - \lambda_4$. The analytical solution is Equation (3.51).
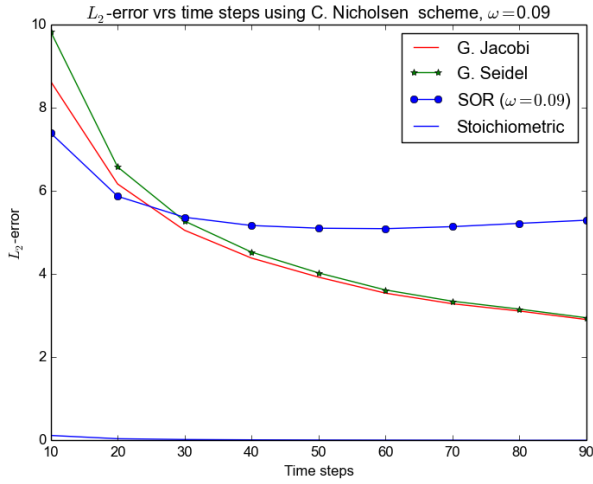
(a) L_2- errors/Crank-Nicholson scheme.
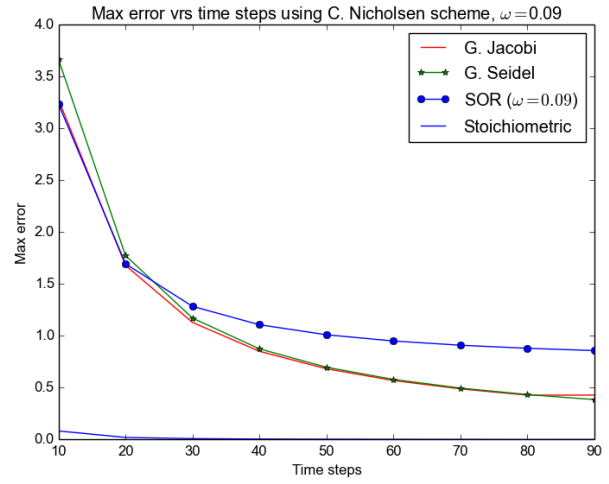
(b) Max errors/Crank-Nicholson scheme.
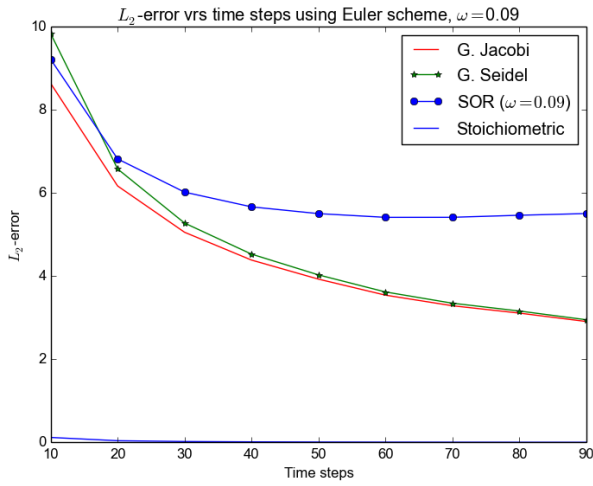
(c) L_2- errors/backward Euler scheme.
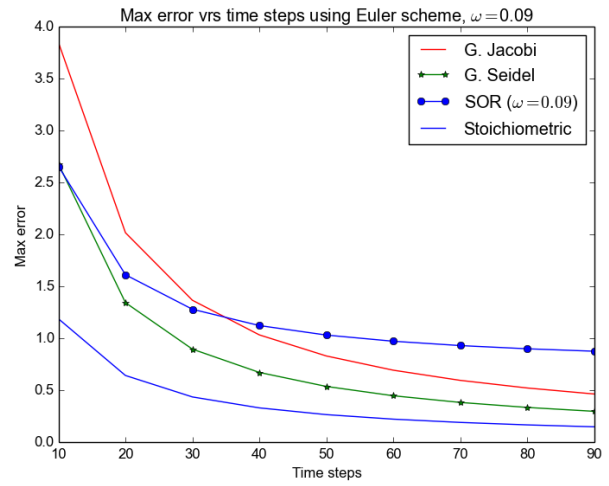
(d) Max errors/backward Euler scheme.
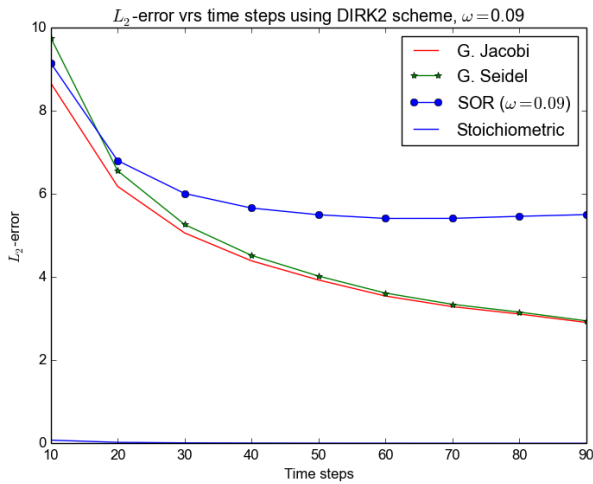
(e) L_2- errors/*DIRK2* scheme.
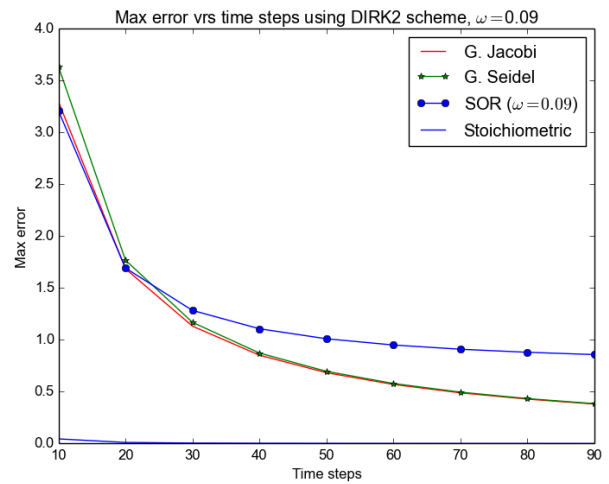
(f) Max errors/*DIRK2* scheme.

Figure 3.3: Errors (L_∞ and L_2) of the numerical schemes, measured across decoupled models and time steps. The actual error values in the figures are 7 orders of magnitude smaller.
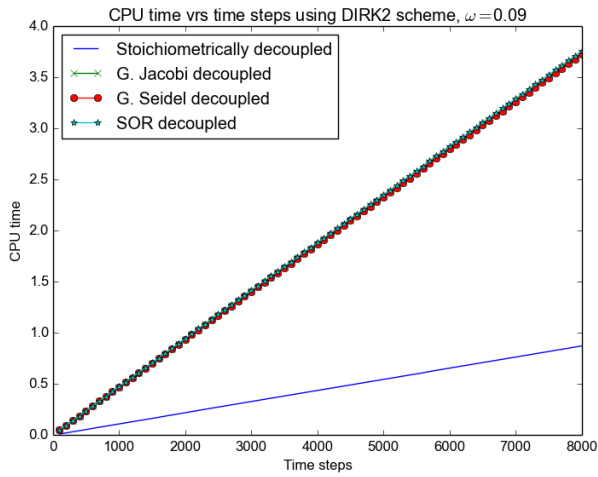
(a) CPU time/*DIRK2* scheme.
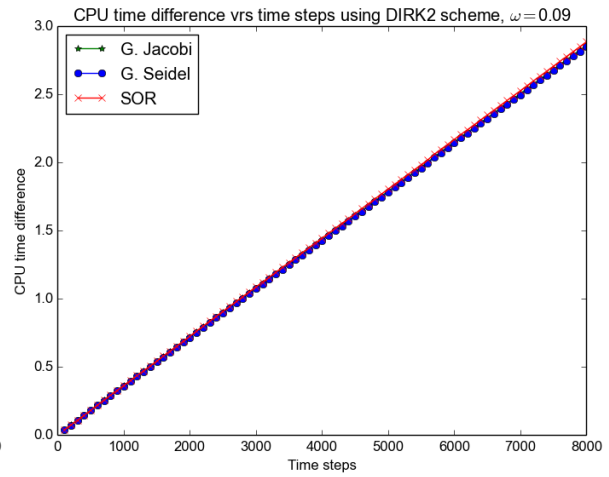
(b) CPU time difference/*DIRK2* scheme.
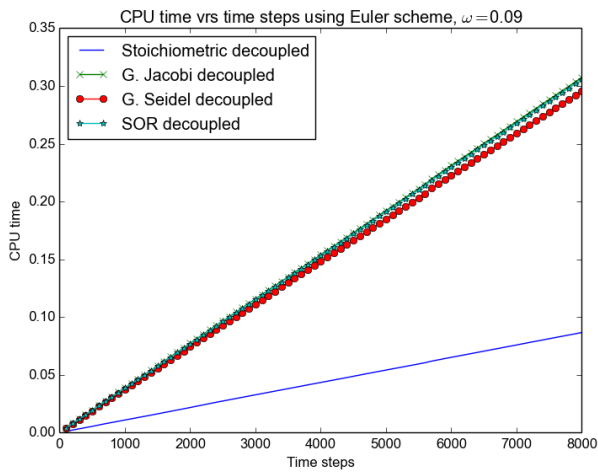
(c) CPU time/B. Euler scheme.
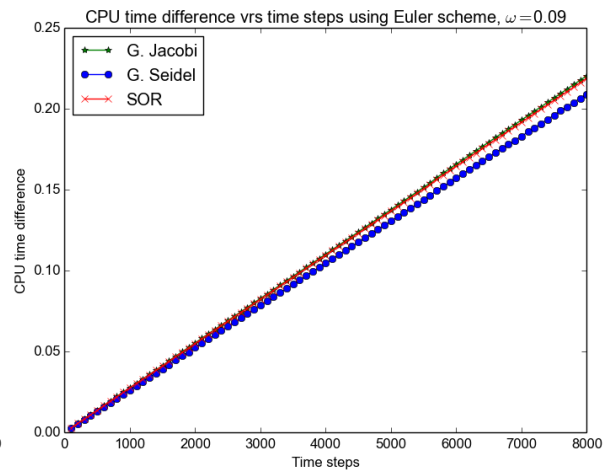
(d) CPU time difference/B. Euler scheme.
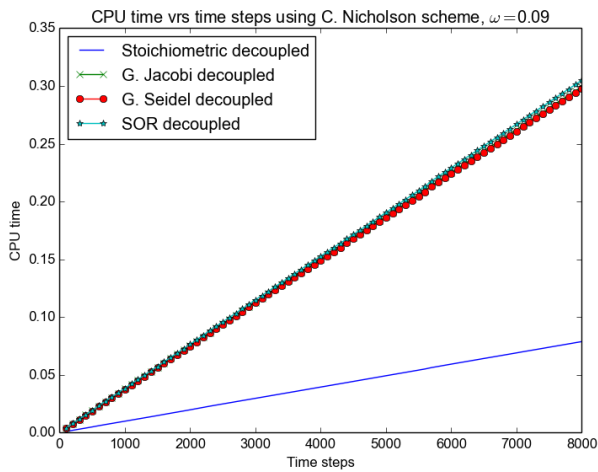
(e) CPU time/C. Nicholson scheme.
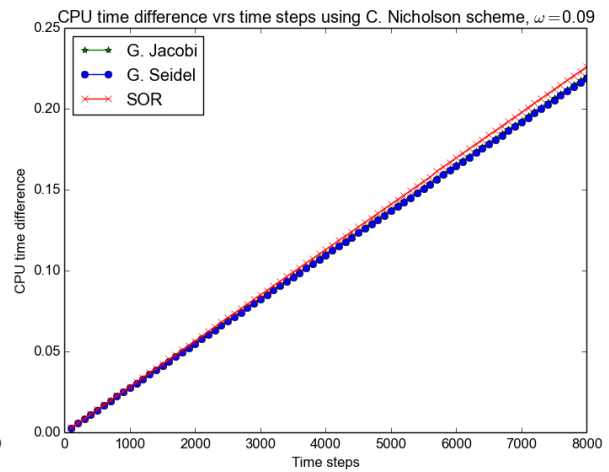
(f) CPU time difference/C. Nicholson scheme.

Figure 3.4: CPU time and CPU time differences for *Gauss Jacobi, Gauss Seidel and SOR,* using *backward Euler,* DIRK2 *and Crank-Nicholson*, measured across time steps (in the calcite model).
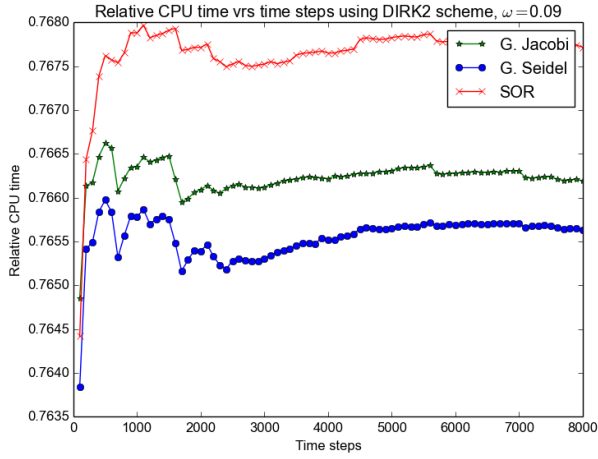
(a) *DIRK2* scheme/8000 steps
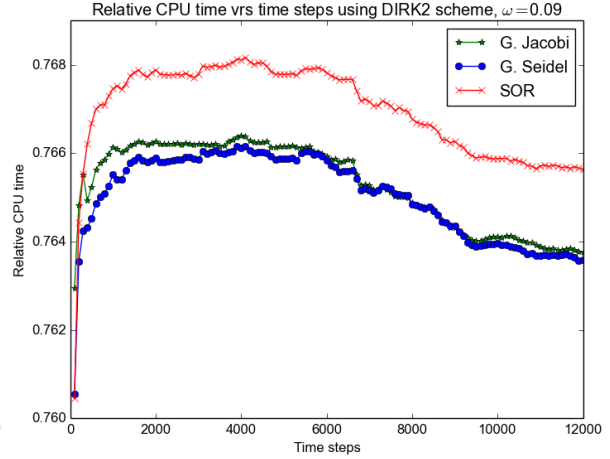
(b) *DIRK2* scheme/12000 steps
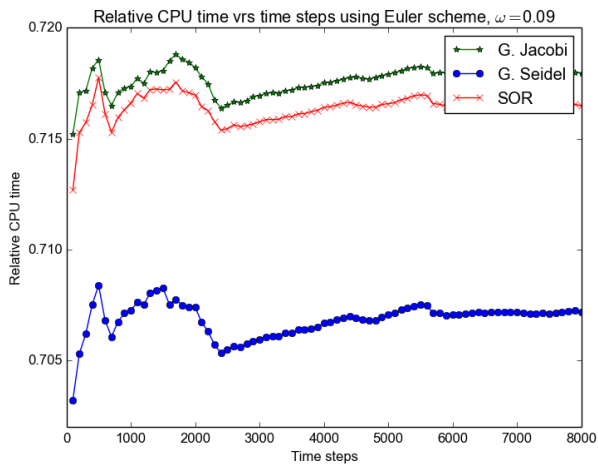
(c) Backward Euler scheme/8000 steps
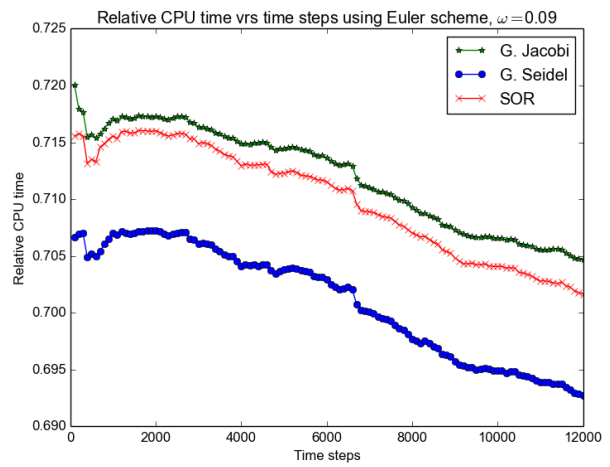
(d) Backward Eulerscheme/12000 steps
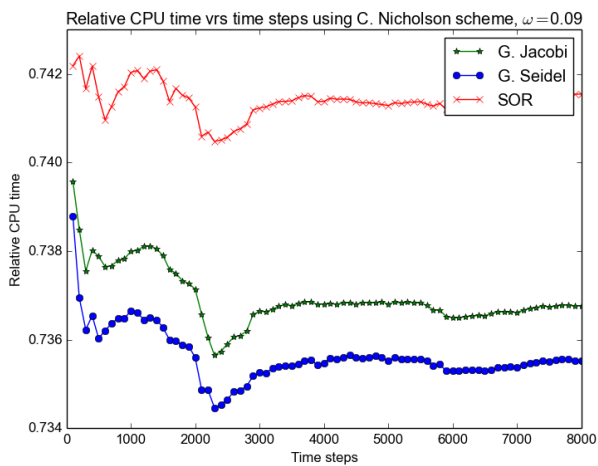
(e) Crank-Nicholson scheme/8000 steps.
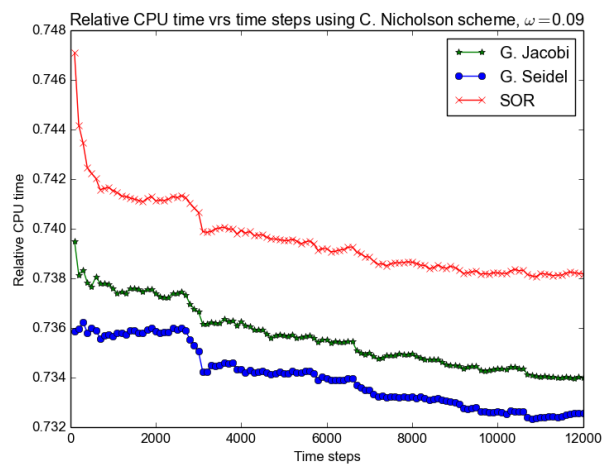
(f) Crank-Nicholson scheme/12000 steps.

Figure 3.5: Relative CPU time for *Gauss-Jacobi, Gauss-Seidel and SOR* methods using *DIRK2,* Backward Euler *and Crank-Nicholson* schemes, measured across time steps in the calcite model.

Chapter 4

Reactive transport modelling and simulation

Models for chemical reactions that occur in transport processes (such as advection and diffusion), are characteristically nonlinear and challenging to simulate due to stiffness, high degrees of freedom and spatial heterogeneity. In this chapter, the stoichiometric procedure is extended to non-linear reactive transport models to reduce their high degrees of freedom. Second order accurate numerical schemes that are compatible with the stoichiometric method have been presented.

Remark: *Simulation studies on linear reactive transport modelling has been published in [2], however, the results presented here are under review.*

4.1 Introduction

Over the past five decades, there has been a growing interest in spatio-temporal reactions due to the ability of such reactions to produce fascinating patterns [62, 78]. The first reported chemical oscillatory system was in the 1950s [15, 44]. In 1952, Turing [164] showed that spatio-temporal patterns can form from a coupled set of non-linear reactions and diffusion. The origin of chemical instability has been studied by [116]. Stability of first-order exothermic reactions in a continuous stirred tank reactor has been discussed in [5], and advance dynamics that are due to chemical reactions in industrial processes have been discussed in [42].

Furthermore, in 1978 it was shown (theoretically in [84] and experimentally in [29]) that diffusive transport could induce instability in isothermal chemical systems. The authors in [52, 33], have discussed chemical instabilities in the non-isothermal systems. Further, in [122, 121, 159] several chemical reactors have been designed to study instabilities in both temporal and spatio-temporal chemical systems. Instabilities (pattern formation) can also occur in convecting fluids that involve chemical reactions and have been discussed in [106, 108, 134]. For Convection-diffusion-reaction systems, instabilities have been reported by [62].

Over the years, research (both theoretical and experimental) has been conducted on relatively smaller systems where analytical results are available. However, modern chemical systems are large, thus, require other approaches for analyses. The focus of this study is numerical simulation of spatio-temporal systems that involve non-linear convection, diffusion and reactions. The models for such systems are nonlinear, stiff and have high degrees of freedom, thus, are computationally expensive to simulate. The goal is to reduce simulation cost using model reduction methods and compatible high order numerical schemes that can resolve stiffness and non-linearities.

In Section 4.2, a nonlinear reactive transport model is presented and the stoichimetric method is applied to reduce the degrees of freedom to any degree of interest. In Section 4.3, a fully nonlinear model that describes neutralization of acidic effluent water is presented, and the stoichiometric method is applied to reduce the degrees of freedom. In Section 4.4, semi-implicit numerical schemes are proposed for the fully nonlinear models. In Section 4.5, numerical experiments are conducted to verify the numerical schemes and to validate the stoichiometric decoupling method, and the chapter is concluded in Section 4.6.

4.2 Modeling reactive transport systems

General fully nonlinear models for reactive transport processes are presented here. The models are composed of nonlinear terms that describe convection, diffusion and reactions that occur simultaneously in spatially varying chemical processes. The models are derived by applying the law of Mass action (discussed in Chapter 3) to close the general transport

equation that was presented in Chapter 2.

Firstly, a two-dimensional non-porous rectangular domain Ω_d (with boundaries located at L_x and L_y units in the x, y directions, respectively) is considered, in which a fluid flows uniformly with hydraulic conductivity denoted K_h, in a time period T. Further, denote the fluid velocities in the x, y directions, respectively by $\mathbf{v} = (u, v)$. If the fluid (here taken as water) carries a chemical species whose concentration is denoted by U_k, and that, the Darcy flow approximation holds, then by conserving species mass across an elementary volume yields a generalised non-linear convection-diffusion-reaction model written as follows:

$$\mathbf{v} = -K_h \nabla p \tag{4.1}$$

$$\frac{\partial U_k}{\partial t} + \frac{\partial}{\partial x} f^x(U_k, \mathbf{v}) + \frac{\partial}{\partial y} f^y(U_k, \mathbf{v}) = \frac{\partial}{\partial x}(\Gamma^x \frac{\partial}{\partial x} U_k) + \frac{\partial}{\partial y}(\Gamma^y \frac{\partial}{\partial y} U_k) + S_k, \tag{4.2}$$

$$k = 1, \ldots, N_s, \quad ((x, y), t) \in \Omega_d \times [0, T)$$

where p is the fluid pressure, $f^x(U_k, \mathbf{v}), f^y(U_k, \mathbf{v})$ denote advective fluxes in the x, y directions, Γ^x, Γ^y denote nonlinear diffusivities of the species in the fluid and S_k is a real-valued function that quantifies the rate of addition/reduction of species k.

The source term S_k in the reactive transport model 4.2 must be defined, in order to complete the modelling. For the spatio-temporal chemical systems (under consideration here) where the mass action law applies, the kinetic model for well-mixed chemical systems (i.e. Equation (3.3)) is used for the closure. Thus, closure of the model yields:

$$\mathbf{v} = -K_h \nabla p \tag{4.3}$$

$$\frac{\partial U_k}{\partial t} + \frac{\partial}{\partial x} f^x(U_k, \mathbf{v}) + \frac{\partial}{\partial y} f^y(U_k, \mathbf{v}) = \frac{\partial}{\partial x}(\Gamma^x \frac{\partial}{\partial x} U_k) + \frac{\partial}{\partial y}(\Gamma^y \frac{\partial}{\partial y} U_k)$$

$$+ \sum_{r=1}^{M} \sigma_{kr} \left(K_{fr} \prod_{i=1}^{N_{f,r}} U_{r,i}^{\alpha_{r,i}} - K_{b,r} \prod_{j=N_{f,r}+1}^{N_r} U_{r,j}^{\alpha_{r,j}} \right),$$

$$\tag{4.4}$$

$$k = 1, \ldots, N_s, \quad ((x, y), t) \in \Omega_d \times [0, T).$$

However, the Partial Differential Equations (PDEs) in the closed model (4.3)-(4.4) are coupled through the source terms. A standard simulation procedure includes the use of

waveform relaxation methods discussed in Chapter 3. However, the standard procedure is computationally expensive especially if concentration profiles of few species are of interest. Thus, stoichiometric decoupling is proposed here for closure of the Equation (4.2). Consequently, the stoichiometric closure (i.e. using Equation (3.21)) yields:

$$\mathbf{v} = -K_h \nabla p \tag{4.5}$$

$$\frac{\partial U_k}{\partial t} + \frac{\partial}{\partial x} f^x(U_k, \mathbf{v}) + \frac{\partial}{\partial y} f^y(U_k, \mathbf{v}) = \frac{\partial}{\partial x}(\Gamma^x \frac{\partial}{\partial x} U_k) + \frac{\partial}{\partial y}(\Gamma^y \frac{\partial}{\partial y} U_k)$$

$$+ \sum_{r=1}^{M} \sigma_{kr} \Big(K_{f,r} \prod_{i=1}^{N_{f,r}} \Big(d_{r,i} + \frac{\sigma_{r,i}}{\sigma_{kr}} U_k \Big)^{\alpha_{r,i}} - K_{b,r} \prod_{j=1+N_{f,r}}^{N_r} \Big(d_{r,j} + \frac{\sigma_{r,j}}{\sigma_{kr}} U_k \Big)^{\alpha_{r,j}} \Big), \tag{4.6}$$

$$k = 1, \ldots, N_s, \quad ((x, y), t) \in \Omega_d \times [0, T].$$

In subsequent sections of this chapter, the large model (4.3)-(4.4) will be simulated using Gauss-Jacobi decoupling, thus, we refer to the model as Gauss-Jacobi decoupled model or large model, and the model resulting from stoichiometric closure is referred to as stoichiometric model or reduced model.

4.3 Modelling acid mine drainage

The closed transport equations above are general models, however, specific models can be obtained by using kinetic data that are specific to a particular chemical kinetic system. The particular system considered here is from Environmental Chemical Engineering and the motivation for this choice follows the work in [48], where the processes that occur in laboratory experiments involving neutralization of acidic effluents in cartridges have been accurately analysed using a reactive transport model. Here, the efficiency of the two modelling/closure procedures are validated using the acid drainage problem.

As mentioned in Chapter 3, experimental data in [126] shows that reactions (3.36) and (3.37) contribute insignificantly compared with reaction (3.35), thus reaction (3.35) is considered in the spatio-temporal case. Consequently, the chemical species present are:

$$\mathbf{C} = \Big(CaCO_3, H^+, Ca^{2+}, HCO_3^- \Big),$$

corresponding concentration and source/sink vectors are given respectively, by:

$$\mathbf{U} = \Big(U_1, U_2, U_3, U_4 \Big) \quad \text{and} \quad \mathbf{S} = \Big(S_1, S_2, S_3, S_4 \Big).$$

Except for calcite $CaCO_3(k = 1)$ that has a zero reaction order, the reaction orders for all the other species are ones (these orders have been obtained from the rate law provided in [126]). Thus, by the law of mass action the rate of reaction is a polynomial in four variables, which can be expressed as follows:

$$R(U_1, U_2, U_3, U_4) = K_f U_2 - K_b U_3 U_4. \tag{4.7}$$

Hydrogen ions ($k = 2$) are responsible for acidity of the effluent water, therefore, treatment of the effluent water implies neutralization of the hydrogen ions. Thus, hydrogen ion (H^+) is the species of interest in the investigation. Using stoichiometry and equation (4.7), the reactive transport equation for hydrogen ions is given by:

$$\frac{\partial U_2}{\partial t} + \frac{\partial}{\partial x} f^x(U_2) + \frac{\partial}{\partial y} f^y(U_2) = \frac{\partial}{\partial x}(\Gamma^x \frac{\partial}{\partial x} U_2) + \frac{\partial}{\partial y}(\Gamma^y \frac{\partial}{\partial y} U_2) - K_f U_2 + K_b U_3 U_4, \tag{4.8}$$

The solution of PDE (4.8) can provide sufficient information about the limestone method of neutralizing the acid (hydrogen ions). However, current concentrations of calcium ions (U_3) and hydrocarbonate ions (U_4) must be available to enable solution of PDE (4.8). Two transport equations must be solved simultaneously with PDE (4.8) after applying Gauss-Jacobi decoupling. The extra two transport equations are as follows:

$$\frac{\partial U_3}{\partial t} + \frac{\partial}{\partial x} f^x(U_3) + \frac{\partial}{\partial y} f^y(U_3) = \frac{\partial}{\partial x}(\Gamma^x \frac{\partial}{\partial x} U_3) + \frac{\partial}{\partial y}(\Gamma^y \frac{\partial}{\partial y} U_3) + K_f U_2 - K_b U_3 U_4, \tag{4.9}$$

and

$$\frac{\partial U_4}{\partial t} + \frac{\partial}{\partial x} f^x(U_4) + \frac{\partial}{\partial y} f^y(U_4) = \frac{\partial}{\partial x}(\Gamma^x \frac{\partial}{\partial x} U_4) + \frac{\partial}{\partial y}(\Gamma^y \frac{\partial}{\partial y} U_4) + K_f U_2 - K_b U_3 U_4. \tag{4.10}$$

However the high degrees of freedom of PDE system (4.8)-(4.10) will increase computational cost especially for stiff problems. A remedy is to apply the stoichiometric method to reduce the system's degrees of freedom. Firstly, define the initial data of the species by the vector:

$$\mathbf{U}_0 = \left(U_{10}, \ U_{20}, U_{30}, U_{40} \right),$$

and assume that there are no sources or sinks (a closed reactor) so that the source/sink vector is given by:

$$\mathbf{U}_S = \Big(0, \ 0, \ 0, \ 0\Big).$$

There is only one stoichiometric equation in the limestone neutralization process, thus $M = r = 1$. In this case, the global variables (i.e. $\alpha, \mathbf{C}, \sigma, K_b$) for species information are the same as the local variables (i.e. $\alpha_1, \mathbf{C}_1, \sigma_1, K_{b,1}$). Given the neutralization reaction (3.35), we have the following input data for the stoichiometric decoupling procedure:

$$\mathbf{C}_1 = \mathbf{C} = \Big(CaCO_3, H^+, Ca^{2+}, HCO_3^-\Big), \quad \alpha_1 = \alpha = \Big(0, 1, 1, 1\Big),$$
$$\sigma_1 = \sigma = \Big(-1, -1, 1, 1\Big) \ N_{f,1} = 2, \ N_1 = 4, \ K_{f,1} = K_f, \ K_{b,1} = K_b. \tag{4.11}$$

The species of interest is H^+ which corresponds to the index $k = 2$ in the global vector $\mathbf{C}$, and index $n = 2$ in the local vector $\mathbf{C}_1$. Thus using input data (4.11) in (3.2) and manipulating algebraically yields a single-variable polynomial rate law in U_2 only, given by:

$$\frac{\partial U_2}{\partial t} = S_2(U_2) = \alpha_{H1}U_2^2 + \alpha_{H2}U_2 + \alpha_{H3}, \tag{4.12}$$

where $\alpha_{H1} = K_b$, $\alpha_{H2} = -K_f + K_b\Big(2U_{20} + U_{40} + U_{30}\Big)$ and $\alpha_{H3} = K_b\Big(U_{10} + U_{40}\Big)\Big(U_{20} + U_{30}\Big)$.

Thus, by replacing the source/sink term $-K_f U_2 + K_b U_3 U_4$ in PDE (4.8) with (4.12), we obtain a decoupled transport equation for the limestone neutralization method, as follows:

$$\frac{\partial U_2}{\partial t} + \frac{\partial}{\partial x}f^x(U_2) + \frac{\partial}{\partial y}f^y(U_2) = \frac{\partial}{\partial x}(\Gamma^x\frac{\partial}{\partial x}U_2) + \frac{\partial}{\partial y}(\Gamma^y\frac{\partial}{\partial y}U_2)$$
$$+ \alpha_{H1}U_2^2 + \alpha_{H2}U_2 + \alpha_{H3}. \tag{4.13}$$

Observe that PDE (4.13) is free from the current concentrations of the other species except in the coefficients α_{H1}, α_{H2} and α_{H3}, that have been calculated from the initial and source data for all the species.

4.4 Numerical schemes

The simultaneous occurrence of advection, diffusion and reaction in reactive transport models poses an interesting stiffness challenge in numerical simulation [73, 169]. In such models, stiffness result from/across the three processes unlike the well-mixed models where stiffness occurs among reaction terms only. Similar to the well-mixed models, stiffness may be resolved by implicit numerical schemes, however, high degrees of freedom and dependence on space variables make fully implicit schemes expensive [73]. An efficient numerical procedure is to apply explicit discretizations to non-stiff terms and implicit discretizations to stiff terms, without explicitly splitting the PDE. Such an approach is generally known as additive splitting. Another approach is to first apply operator splitting techniques to the PDEs, thus, explicitly splitting the advection, diffusion and reactions terms. In both approaches, the resulting numerical schemes are called IMplicit-EXplicit (IMEX) schemes. However, explicit splitting techniques have accuracy issues and do not preserve transient balances [169]. Therefore, the IMEX schemes that result from additive splitting are considered here.

The goal here is to present second order low cost numerical schemes that are suitable for stiff spatio-temporal chemical systems and are compatible with the stoichiometric decoupling method. The method of lines (MOL) discretization procedure where the spatial derivatives are discretized first, will be employed here.

4.4.1 Spatial discretization

The stoichiometrically decoupled model will be compared with the original large model in order to establish efficacy, thus, we find it appropriate to reduce the spatial dimension to one, to enable visual comparison of numerical solutions. Thus, using vector notations, the 1-D version of the non-linear convection-diffusion-reaction model (4.1)-(4.2) is written as follows:

$$u = -K_h \frac{dp}{dx}, \tag{4.14}$$

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial}{\partial x}\mathbf{f} = \frac{\partial}{\partial x}(\Gamma \frac{\partial}{\partial x}\mathbf{U}) + \mathbf{S}, \quad \in [0, L_x] \times [0, T) \tag{4.15}$$

where $\mathbf{f}(\mathbf{U}, u)$ is the species advective flux that could be linear or non-linear and Γ is the diffusivity tensor (assumed diagonal) that depends on the concentration vector $\mathbf{U}$. For the purposes of numerical investigations, we denote diffusivity of the kth chemical species by Γ_{kk}, and define Buckley-Leverett, Burgers and linear advective fluxes for the kth species respectively as follows:

$$f_k(U_k, u) = 4u \frac{U_k^2}{4U_k^2 - (1 - U_k)}, \quad f_k(U_k, u) = \frac{u}{2}U_k^2 \text{ and } f_k(U_k, u) = uU_k,$$

where U_k is k^{th} components of the vector $\mathbf{U}$. Diffusivity functions (i.e. Γ_{kk}) will be defined later in the discussion.

Furthermore, we apply the method of lines (MOL) discretization procedure. Spurious oscillations could develop near reaction fronts and boundaries if suitable techniques are not employed to discretize the advection term. These oscillations could result in negative concentration values that have no physical and chemical meaning. Thus we apply finite difference WENO (third order) procedure with Lax-Friedrichs flux splitting to discretize the advection term. Applying a conservative finite difference approximation to the advection term in (4.15) at the ith node x_i, yields:

$$\frac{\partial}{\partial x}\mathbf{f}|_{x_i} \approx \frac{1}{\Delta x}\left(\hat{\mathbf{f}}_{i+\frac{1}{2}} - \hat{\mathbf{f}}_{i-\frac{1}{2}}\right), \tag{4.16}$$

where Δx is the spatial step size and $\hat{\mathbf{f}}_{i+1}$ is a numerical flux. A third order WENO reconstruction/approximation of the numerical flux is considered, thus:

$$\hat{\mathbf{f}}_{i+\frac{1}{2}} = w_0\left(\frac{1}{2}\mathbf{f}(\mathbf{U}_i) + \frac{1}{2}\mathbf{f}(\mathbf{U}_{i+1})\right) + w_1\left(-\frac{1}{2}\mathbf{f}(\mathbf{U}_{i-1}) + \frac{3}{2}\mathbf{f}(\mathbf{U}_i)\right), \tag{4.17}$$

where

$$w_r = \frac{\alpha_r}{\alpha_1 + \alpha_2}, \quad \alpha_r = \frac{d_r}{(\epsilon + \beta_r)^2}, \quad r = 0, 1,$$

the linear weights are $d_0 = \frac{2}{3}$, $d_1 = \frac{1}{3}$, the smoothness indicators are $\beta_0 = \left[\mathbf{f}(\mathbf{U}_{i+1}) - \mathbf{f}(\mathbf{U}_i)\right]^2$, $\beta_1 = \left[\mathbf{f}(\mathbf{U}_i) - \mathbf{f}(\mathbf{U}_{i-1})\right]^2$ and $\epsilon = 10^{-3}$ is a parameter taken to ensure a non-zero denominator. Further, a Lax-Friedrich splitting of the flux is performed to incorporate upwinding in the final numerical schemes, thus:

$$\mathbf{f}^+(\mathbf{U}) = \frac{1}{2}\big(\mathbf{f}(\mathbf{U}) + \alpha\mathbf{U}\big), \quad \mathbf{f}^-(\mathbf{U}) = \frac{1}{2}\big(\mathbf{f}(\mathbf{U}) - \alpha\mathbf{U}\big), \tag{4.18}$$

where $\alpha = \max_{\mathbf{U}} |\mathbf{f}'(\mathbf{U})|$, $\mathbf{f}^+(\mathbf{U})$ is the positive wind part and $\mathbf{f}^-(\mathbf{U})$ is the negative part.

The derivative corresponding to diffusion is discretized using the second order central differencing approximation as follows:

$$\frac{\partial}{\partial x}(\Gamma \frac{\partial}{\partial x}\mathbf{U})|_{x_i} \approx \frac{\Gamma\left(\frac{\mathbf{U}_{i+1}+\mathbf{U}_i}{2}\right)}{(\Delta x)^2}\mathbf{U}_{i+1} - \left(\frac{\Gamma\left(\frac{\mathbf{U}_{i+1}+\mathbf{U}_i}{2}\right)}{(\Delta x)^2} + \frac{\Gamma\left(\frac{\mathbf{U}_i+\mathbf{U}_{i-1}}{2}\right)}{(\Delta x)^2}\right)\mathbf{U}_i$$
$$+ \frac{\Gamma\left(\frac{\mathbf{U}_i+\mathbf{U}_{i-1}}{2}\right)}{(\Delta x)^2}\mathbf{U}_{i-1}. \tag{4.19}$$

Applying approximations (4.16) and (4.19) in equation (4.15), the semi-discrete form of the reactive transport model states that, find $\tilde{\mathbf{U}}$ such that:

$$\frac{\partial \tilde{\mathbf{U}}}{\partial t} = \mathbf{F}_d(\tilde{\mathbf{U}}) + \mathbf{F}_a(\tilde{\mathbf{U}}) + \mathbf{S}(\tilde{\mathbf{U}}), \quad \in [0, T) \tag{4.20}$$

where $\tilde{\mathbf{U}} = (\mathbf{U}_1, \mathbf{U}_2, \ldots, \mathbf{U}_{N_x})$, $\mathbf{F}_d(\tilde{\mathbf{U}})$ is the discretized diffusion term, $\mathbf{F}_a(\tilde{\mathbf{U}})$ is the discretized advection term and N_x is the total number of grid points. Further, discretized diffusion derivative can be linearised as follows:

$$\mathbf{F}_d(\tilde{\mathbf{U}}) \approx \mathbf{F}_d(\tilde{\mathbf{U}}^n) + \mathcal{C}_n(\tilde{\mathbf{U}} - \tilde{\mathbf{U}}^n), \tag{4.21}$$

where $\tilde{\mathbf{U}}^n = \tilde{\mathbf{U}}(t_n)$ and $\mathcal{C}_n = \mathcal{C}(\tilde{\mathbf{U}}^n)$ is the Jacobian of $\mathbf{F}_d(\tilde{\mathbf{U}})$ evaluated at $\tilde{\mathbf{U}}^n$. Applying the linearised discrete diffusion function (4.21) in the semi-discrete transport equation (4.20) yields:

$$\frac{\partial \tilde{\mathbf{U}}}{\partial t} = \mathbf{F}_d(\tilde{\mathbf{U}}^n) + \mathcal{C}_n\tilde{\mathbf{U}} - \mathcal{C}_n\tilde{\mathbf{U}}^n + \mathbf{F}_a(\tilde{\mathbf{U}}) + \mathbf{S}(\tilde{\mathbf{U}}), \quad t \in [0, T). \tag{4.22}$$

A complete discretization will be obtained in the following subsections using different time integrators. The resulting complete discretizations are all IMEX schemes for fully nonlinear reactive transport models.

4.4.2 Implicit integration factor scheme

The terms in the semi-discrete transport equation (4.22) can be grouped into stiff and non-stiff terms. The stiff group (that consists of reaction and diffusion terms) pose severe time step restrictions on some time integrators. Implicit integration factor class of schemes use exact integration procedures to remove time step restrictions posed by

the linear part of the discretized diffusion derivative [18, 34, 105, 79, 73]. Denote the combined discrete diffusion and advection flux functions by:

$$\mathbf{F}(\tilde{\mathbf{U}}(t_n + \tau)) = \mathbf{F}_d(\tilde{\mathbf{U}}(t_n + \tau)) - \mathcal{C}_n\tilde{\mathbf{U}}(t_n + \tau) + \mathbf{F}_a(\tilde{\mathbf{U}}(t_n + \tau)), \qquad (4.23)$$

where τ is the time step. Further, let $r > 0$ be an integer and denote the time interpolation points by $t_{n+i} = t_n + \tau_i$, $i = 1, 0, -1, \ldots, 1 - r$, where $\tau_1 = \Delta t_n$, $\tau_0 = 0$, $\tau_i = -\sum_{k=i}^{-1} \Delta t_{n+k}$, for $i = -1, -2, -3, \ldots, 1 - r$. Then by multiplying (4.22) by an appropriate integration factor and manipulating algebraically, yields the generalized r^{th} order implicit integration factor scheme:

$$\tilde{\mathbf{U}}^{n+1} = e^{-\mathcal{C}_n \Delta t_n}\tilde{\mathbf{U}}^n$$

$$+ \Delta t_n\Big(\sum_{i=2-r}^{1} \alpha_{n+i}e^{\mathcal{C}_n(\Delta t_n - \tau_i)}\mathbf{S}(\tilde{\mathbf{U}}^{n+i}) + \sum_{i=1-r}^{0} \beta_{n+i}e^{\mathcal{C}_n(\Delta t_n - \tau_i)}\mathbf{F}(\tilde{\mathbf{U}}^{n+i}) \Big), \qquad (4.24)$$

where the coefficients are defined as follows:

$$\alpha_{n+i} = \frac{1}{\Delta t_n} \int_0^{\Delta t_n} \prod_{j=2-r,\ j\neq i}^{1} \frac{\tau - \tau_j}{\tau_i - \tau_j}d\tau, \quad i = 1, 0, -1, \ldots, 2 - r.$$

$$\beta_{n+i} = \frac{1}{\Delta t_n} \int_0^{\Delta t_n} \prod_{j=1-r,\ j\neq i}^{0} \frac{\tau - \tau_j}{\tau_i - \tau_j}d\tau, \quad i = 0, -1, \ldots, 1 - r.$$

Only the second order implicit integration factor scheme with $r = 2$ (here named IMEX-IIF2) scheme will be included in further discussions. For $r = 2$, the IMEX-IIF2 scheme is expression (4.24) with coefficients given by:

$$\alpha_{n+1} = \frac{1}{2}, \quad \alpha_n = \frac{1}{2}, \quad \beta_n = \frac{\left(\frac{\Delta t_n}{2} + \Delta t_{n-1}\right)}{\Delta t_{n-1}}, \quad \text{and} \quad \beta_{n-1} = -\frac{\Delta t_n}{2\Delta t_{n-1}}. \qquad (4.25)$$

Furthermore, linear analyses of numerical schemes reveal convergence properties of the numerical schemes and has been used to analyse several schemes (see for example [132, 151, 185, 117]). Thus, issues relating to convergence (i.e. consistency and stability) of the IMEX-IIF2 will be studied next, using linear analysis. Since second order central differencing and third order WENO reconstruction have been used to discretize

the diffusion and advection terms, respectively, the overall spatial discretization is second order consistent. Thus, the focus is now on consistency of the temporal discretization. To proceed further, a linear transport equation is spatially discretized to obtain:

$$\frac{\partial \tilde{\mathbf{U}}}{\partial t} = \mathbb{D}\tilde{\mathbf{U}} + \mathbb{A}\tilde{\mathbf{U}} + \mathbb{R}\tilde{\mathbf{U}}, \quad t \in [0, T). \tag{4.26}$$

where $\mathbb{D}, \mathbb{A}$ and $\mathbb{R}$ are constant coefficient matrices corresponding to diffusion, advection and reaction terms, respectively.

Proposition 4.1. *Given the semi-discrete problem* (4.26), *the integration factor scheme* (4.24) *with coefficients given by* (4.25), *is second order consistent.*

Proof. Applying Scheme (4.24) with coefficients (4.25) to discretize (4.26) yields:

$$\tilde{\mathbf{U}}^{n+1} = \Big(\mathbb{I} - \frac{\mathbb{R}}{2}\Delta t\Big)^{-1} e^{\mathbb{D}\Delta t}\Big(\mathbb{I} + \frac{\mathbb{R}}{2}\Delta t + \frac{3\mathbb{A}}{2}\Delta t\Big)\tilde{\mathbf{U}}^n - \frac{\Delta t}{2}\Big(\mathbb{I} - \frac{\mathbb{R}}{2}\Delta t\Big)^{-1} e^{2\mathbb{D}\Delta t}\tilde{\mathbf{U}}^{n-1}. \tag{4.27}$$

However,

$$\Big(\mathbb{I} - \frac{\mathbb{R}}{2}\Delta t\Big)^{-1} e^{\mathbb{D}\Delta t}\Big(\mathbb{I} + \frac{\mathbb{R}}{2}\Delta t + \frac{3\mathbb{A}}{2}\Delta t\Big) = \mathbb{I} + \Big(\mathbb{R} + \mathbb{D} + \frac{3\mathbb{A}}{2}\Big)\Delta t$$
$$+ \Big(\frac{\mathbb{R}^2}{2} + \frac{\mathbb{D}^2}{2} + \frac{\mathbb{R}\mathbb{D}}{2} + \frac{3\mathbb{R}\mathbb{A}}{2} + \frac{\mathbb{D}\mathbb{R}}{2} + \frac{3\mathbb{D}\mathbb{A}}{2}\Big)\Delta t^2 + \cdots, \tag{4.28}$$

and

$$\frac{\Delta t}{2}\Big(\mathbb{I} - \frac{\mathbb{R}}{2}\Delta t\Big)^{-1} e^{2\mathbb{D}\Delta t}\tilde{\mathbf{U}}^{n-1} = \Big(\frac{\mathbb{A}}{2}\Delta t + \Big(\frac{\mathbb{R}\mathbb{A}}{4} + \mathbb{D}\mathbb{A} - \frac{\mathbb{A}^2}{2} - \frac{\mathbb{A}\mathbb{D}}{2} - \frac{\mathbb{A}\mathbb{R}}{2}\Big)\Delta t^2 + \cdots .\Big)\tilde{\mathbf{U}}^n. \tag{4.29}$$

Expressions (4.28) and (4.29) imply that:

$$\tilde{\mathbf{U}}^{n+1} = \Big(\mathbb{I} + (\mathbb{A} + \mathbb{D} + \mathbb{R})\Delta t + \frac{(\mathbb{A} + \mathbb{D} + \mathbb{R})^2}{2}\Delta t^2 + \mathcal{O}(\Delta t^3)\Big)\tilde{\mathbf{U}}^n. \tag{4.30}$$

Thus, since the exact solution of ODE (4.26) is $\tilde{\mathbf{U}}(t_n + \Delta t) = e^{(\mathbb{A}+\mathbb{D}+\mathbb{R})\Delta t}\tilde{\mathbf{U}}(t_n)$, it follows from (4.30) that the IMEX-IIF2 scheme (i.e Scheme (4.24) with coefficients (4.25)) is second order consistent.

$\square$

Moreover, the stability of the IMEX-IIF2 can easily be established following the linear stability procedures in [117, 185, 73]. To proceed with the analysis, the following scalar linear problem is considered:

$$\partial_t U = -dU + aU + rU, \quad \in [0, T). \tag{4.31}$$

where ∂_t denotes derivative with respect to t, d is a positive real coefficient corresponding to diffusion, a is a real coefficient corresponding to advection and r is complex coefficient corresponding to chemical reactions.

Applying Scheme (4.24) with coefficients (4.25) to discretize (4.31) yields:

$$\left(1 - \frac{r}{2}\Delta t\right)U^{n+1} = e^{-d\Delta t}\left(1 + \frac{1}{2}r\Delta t + \frac{3a}{2}\Delta t\right)U^n - \frac{a\Delta t}{2}e^{-2d\Delta t}U^{n-1}. \tag{4.32}$$

Let $\lambda_r = r\Delta t$, then substituting $U^n = e^{in\theta}$ in (4.32) and rearranging yields:

$$\lambda_r = \frac{e^{-d\Delta t}\left(-e^{i\theta}\left(2 + 3a\Delta t\right) + a\Delta t e^{-d\Delta t}\right)}{e^{i\theta}\left(e^{i\theta} + e^{-d\Delta t}\right)}. \tag{4.33}$$

Figure 4.1 shows stability regions for different values of $a\Delta t$ and $d\Delta t$. Clearly, it can be observed in Figures 4.1a - 4.1d that, for fixed values of $a\Delta t$ the region of instability shrinks with increasing values of $d\Delta t$. Moreover, for a fixed value of $d\Delta t$, the region of instability grows larger for increasing values of $|a|\Delta t$. This observation is due to the explicit discretization of the advection term, thus, the time step is constrained by the explicit part of the scheme. However, the IMEX-IIF2 scheme is stable and has larger time step size when compared to fully explicit schemes (of the same order) for reactive transport equations.

4.4.3 Other schemes

For comparison, other standard implicit-explicit schemes that are of the Runge-Kutta and multistep types are presented here. Diffusion and reaction terms are considered stiff, thus, from the semi-discrete equation (4.22), the discretized stiff parts together form a slope function that is given by:

$$\mathbf{F}_I(\tilde{\mathbf{U}}) = \mathbf{F}_d(\tilde{\mathbf{U}}^n) + \mathcal{C}_n\tilde{\mathbf{U}} - \mathcal{C}_n\tilde{\mathbf{U}}^n + \mathbf{S}(\tilde{\mathbf{U}}).$$

Thus, a generalized linear multistep (k-step) implicit-explicit numerical scheme for (4.22) states that [71]:

$$\sum_{j=0}^{k} \alpha_j \tilde{\mathbf{U}}^{n+j} = \Delta t \sum_{j=0}^{k} \beta_j \mathbf{F}_I(\tilde{\mathbf{U}}^{n+j}) + \Delta t \sum_{j=0}^{k-1} \beta_j^* \mathbf{F}_a(\tilde{\mathbf{U}}^{n+j}), \tag{4.34}$$

where α_j, β_j, and β_j^* are coefficients that determine a particular method in this class. One particular multistep scheme of interest is derived from implicit and explicit two-step backward differentiation formulas (BDF) (henceforth, this scheme will be referred to as IMEX-BDF). It is second order convergent, the implicit part is A-stable and it is stated as follows [71]:

$$\frac{3}{2}\tilde{\mathbf{U}}^{n+1} - 2\tilde{\mathbf{U}}^n + \frac{1}{2}\tilde{\mathbf{U}}^{n-1} = \Delta t \mathbf{F}_I(\tilde{\mathbf{U}}^{n+1}) + 2\Delta t \mathbf{F}_a(\tilde{\mathbf{U}}^n) - \Delta t \mathbf{F}_a(\tilde{\mathbf{U}}^{n-1}). \tag{4.35}$$

Another particular multistep scheme involves two steps, is second order consistent and is A-stable in the implicit part. The scheme uses Adams-Bashforth's discretization for the nonstiff part and Crank-Nicholson's discretization for the stiff part. The scheme states that [71]:

$$\tilde{\mathbf{U}}^{n+1} = \tilde{\mathbf{U}}^n + \frac{3}{2}\Delta t \mathbf{F}_a(\tilde{\mathbf{U}}^n) - \frac{1}{2}\Delta t \mathbf{F}_a(\tilde{\mathbf{U}}^{n-1}) + \frac{1}{2}\Delta t \mathbf{F}_I(\tilde{\mathbf{U}}^{n+1}) + \frac{1}{2}\Delta t \mathbf{F}_I(\tilde{\mathbf{U}}^n). \tag{4.36}$$

Furthermore, for second order consistent schemes of the Runge-Kutta type, we consider a scheme that involves a combination of both implicit and explicit discretizations. It has two stages where the first stage is explicit Euler and the second stage is IMEX. The scheme (henceforth referred to as IMEX-RK1) can be written as follows:

$$\tilde{\mathbf{U}}^* = \tilde{\mathbf{U}}^n + \Delta t \mathbf{F}(\tilde{\mathbf{U}}^n)$$
$$\tilde{\mathbf{U}}^{n+1} = \tilde{\mathbf{U}}^n + \frac{1}{4}\Delta t \mathbf{F}(\tilde{\mathbf{U}}^n) + \frac{1}{4}\Delta t \mathbf{F}_a(\tilde{\mathbf{U}}^*) + \frac{1}{2}\Delta t \mathbf{F}_I(\tilde{\mathbf{U}}^{n+1}). \tag{4.37}$$

The last second order consistent scheme of the Runge-Kutta type to be considered in this discussion, also have two-stages where the first stage is explicit and the second stage is IMEX, we refer to this scheme as IMEX-RK2 and state it as follows:

$$\tilde{\mathbf{U}}^* = \tilde{\mathbf{U}}^n + 0.5\Delta t \mathbf{F}_a(\tilde{\mathbf{U}}^n)$$
$$\tilde{\mathbf{U}}^{n+1} = \tilde{\mathbf{U}}^n + \frac{1}{2}\Delta t \mathbf{F}_I(\tilde{\mathbf{U}}^n) + \frac{1}{2}\Delta t \mathbf{F}_a(\tilde{\mathbf{U}}^*) + \frac{1}{2}\Delta t \mathbf{F}_I(\tilde{\mathbf{U}}^{n+1}). \tag{4.38}$$

In the following section, all the schemes presented in this section will be verified with a stiff chemical transport model.

4.5   Numerical experiments

In this section, numerical results on suitability of the numerical schemes, efficiency (i.e. accuracy, compatibility and cost of simulation) of the stoichiometric decoupling method are presented. The numerical schemes will be applied to solve a stiff transport model, L_∞ errors and orders will be measured in order to check convergence. Three best-performing schemes will be selected and applied to discretize the Gauss-Jacobi and stoichiometric models for the acid drainage problem. Accuracy of the stoichiometric procedure will be determined by computing differences between numerical solutions of the stoichiometric and Gauss-Jacobi models. Results on simulation cost of the stoichiometric method will be presented finally.

4.5.1   Convergence test: numerical schemes

Here, grid independent tests to check convergence of the numerical schemes are performed. A system of advection-diffusion transport with stiffness (presented in [73]) is discretized by all the schemes. The one dimensional form of the stiff model in [73] is given by:

$$\frac{\partial U}{\partial t} + \frac{\partial uU}{\partial x} = \Gamma\frac{\partial^2 U}{\partial x^2} - K_u U + V, \qquad 0 < x < 2\pi, \tag{4.39}$$

$$\frac{\partial V}{\partial t} + \frac{\partial uV}{\partial x} = \Gamma\frac{\partial^2 V}{\partial x^2} - K_v V, \qquad 0 < x < 2\pi. \tag{4.40}$$

Using periodic boundary conditions, the exact solution for the system (4.39)-(4.40) is:

$$U(x,t) = \left(e^{-(K_u+\Gamma)t} + e^{-(K_v+\Gamma)t}\right)\cos(x - ut), \tag{4.41}$$

$$V(x,t) = (K_u - K_v)e^{-(K_v+\Gamma)t}\cos(x - ut). \tag{4.42}$$

The input data are as follows:

$$\Gamma = 0.001, \quad u = 0.001, \quad K_v = 1, \quad K_u = 100, \tag{4.43}$$

and the initial data are as follows:

$$U(x,0) = 2\cos(x),$$ (4.44)

$$V(x,0) = (K_u - K_v)\cos(x).$$ (4.45)

Furthermore, we define Peclet number as follows:

$$P_e = \frac{u\Delta x}{\Gamma}.$$

When Peclet number is high ($P_e > 0$) advection dominates the transport process and when Peclet number is low ($P_e < 0$) diffusion dominates. In critical Peclet number ($P_e = 0$) transport both advection and diffusion processes occur simultaneously, we refer to the critical Peclet number case as a normal transport process. Convergence of the numerical schemes will be investigated in the three transport cases (i.e. advection dominated, diffusion dominated and normal transport cases) while maintaining stiffness in the reaction.

Figures 4.2 and 4.3 compare the numerical solutions (i.e. IMEX-IIF2, IMEX-CNAB, IMEX-RK1 and IMEX-RK2 solutions) with the analytical solution for the PDE system (4.39)-(4.40). One can observe that the numerical solutions mimic the analytical solution very well without oscillations for all the transport cases, across space and time. Figures 4.4 and 4.5 compare IMEX-BDF solution with the analytical solution. One can observe that the IMEX-BDF solutions are very inaccurate for all the transport cases and across space and time. Further investigations reveal that the IMEX-RK2 scheme performed better than IMEX-RK1 scheme, thus, we only consider IMEX-RK2, IMEX-IIF2 and IMEX-CNAB schemes in further investigations to follow. Moreover, errors and orders (measured with the L_∞ norm) for the IMEX-RK2, IMEX-IIF2 and IMEX-CNAB schemes are provided in Table 4.1. One can observe that the IMEX-RK2, IMEX-IIF2 and IMEX-CNAB solutions converged to the analytical solution with second order in all the flow cases (i.e. low, high and critical Peclet number cases).

4.5.2 Accuracy test: stoichiometric method

In this section, we present results on accuracy of the stoichiometric decoupling procedure, that were obtained using IMEX-RK2, IMEX-IIF2 and IMEX-CNAB to discretize

the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13) for hydrogen ions. Accuracy is measured by the error (called model reduction error) between the numerical solution of a PDE in the stoichiometric model and the numerical solution of a PDE in the the Gauss-Jacobi model. The PDEs in both models are evolution equations for hydrogen ions (the chemical species of interest). We define the model reduction error as follows:

$$\text{Model reduction error} = \|\text{SOL}_G - \text{SOL}_S\|_\infty \tag{4.46}$$

and where SOL_S is a numerical solution of the stoichiometric model and SOL_G is numerical solution of the Gauss-Jacobi model. Some input data used in the experiments are given by:

$$L_x = 2\pi, \ T = 1, \frac{d}{dx}p = -0.004, K_f = 0.13, \ K_b = 0.0025, K_h = 1,$$
$$U_{20} = 0.01\cos(x), \ U_{30} = 0.0001\cos(x), \ U_{40} = 0.0001\cos(x).$$

Figures 4.6 and 4.7 show profiles of hydrogen concentrations in the stoichiometric and Gauss-Jacobi models, across numerical schemes, space and time. One can observe that the concentration of hydrogen ions reduces in space (relative to the initial concentration) and time. This observation is expected in an AMD remediation reactor. One can also observe that the solution from the stoichiometric model mimic the solution from the Gauss-Jacobi model very well without oscillations (across numerical schemes, space and time) for all the transport cases. Clearly, it can be observed that the transport cases (advection, diffusion and advection-diffusion) have no effect on the accuracy of the solution of the stoichiometric model.

Figures 4.8 and 4.9 show model reduction error profiles and Tables 4.2, 4.3 and 4.4 contain model reduction errors across numerical schemes and transport cases. For all the transport cases the model reduction errors measured with IMEX-IIF2 and IMEX-CNAB solutions, oscillated in a very small neighbourhood of 9×10^{-7}. Thus the stoichiometric method is accurate since the model reduction error is negligible. However, the model reduction errors computed with IMEX-RK2 solutions behaved slightly different in the fully nonlinear and advection transport cases. The reduction errors decreased monotonically into the neighbourhood of 9×10^{-7} and then oscillated in a small neighbourhood. This

83

observation can be seen clearly in Figures 4.9b, 4.9c and 4.9d, where Buckley-Leverett and linear fluxes are used in the advection and fully nonlinear transport cases.

4.5.3 Simulation cost: stoichiometric method

In this Section, we present and discuss results on cost of simulating the stoichiometric and Gauss-Jacobi models. We measure cost by the CPU time, CPU time differences and relative CPU time that the numerical schemes require to solve the Gauss-Jacobi model and the stoichiometric model. We consider a fully nonlinear flow case where Buckley-Leverett flux and nonlinear diffusivity ($\Gamma_{kk} = 1E - 4\cos(U_k)$) are used. The CPU time differences and relative CPU time are defined as follows:

$$\text{CPU time difference} = \text{CPU}_L - \text{CPU}_S \tag{4.47}$$

and

$$\text{Relative CPU time} = \frac{\text{CPU}_L - \text{CPU}_S}{\text{CPU}_L} \tag{4.48}$$

where CPU_S is CPU time for the stoichiometric model and CPU_L is CPU time for the Gauss-Jacobi model. In the experiments, we fixed the final time at ten ($T = 1$) and used a fine spatial resolution (600-900 spatial steps with $\Delta t = 0.5\Delta x$) to ensure that results are not affected much by numerical discretization errors.

The final time was set at $T = 1$, IMEX-IIF2, IMEX-CNAB and IMEX-RK2 discretizations were applied to the fully nonlinear flow cases in both models presented in Section 4.3 and the CPU times for both models were measured. Figures 4.10a, 4.10c and 4.10e show plots of CPU time against time steps. For all the schemes and for both models, the CPU time generally increased with increasing number of time steps, however, the CPU time for the stoichiometric model recorded the least CPU time. Another observation is that, the CPU time difference between the stoichiometric model and the Gauss-Jacobi model increased with increasing time steps. This observation can be seen clearly in Figures 4.10b, 4.10d and 4.10f.

Furthermore, to determine the CPU time saved by solving the stoichiometric model instead of the Gauss-Jacobi model, the CPU time differences for both models (measured

across all discretizations) were normalized by the CPU time for the Gauss-Jacobi model, (i.e using Equation (4.48)). Figure 4.11 shows the results of the investigation. About *67 percent* of CPU time will be saved if IMEX-IIF2 discretization is used, *71 percent* will be saved if IMEX-CNAB or IMEX-RK2 is used to solve the stoichiometric model instead of the Gauss-Jacobi model.

4.6   Chapter summary

Numerical simulation of chemical transport phenomena is expensive and challenging due to nonlinearity, high degrees of freedom, heterogeneity and stiffness. Model reduction methods can be employed to make simulation possible or reduce the high degrees of freedom and simulation cost.

In Section 4.1, the importance of spatio-temporal reactions and some interesting research on such reactions have been highlighted. Most of the studies linearised the processes to enable analytical studies on them, however, most spatio-temporal reaction systems that occur in nature are nonlinear.

In Section 4.2, we presented a nonlinear reactive transport model (that is composed of N_s partial differential equations) and applied the stoichiometric method to reduce the degrees of freedom to any degree less or equal to N_s.

Further, a spatio-temporal nonlinear model for acid neutralization has been discussed as an application problem in Section 4.3. The original full nonlinear model that has three degrees of freedom was reduced to one degree of freedom, using the stoichiometric method.

In Section 4.4, numerical schemes for discretizing the nonlinear reactive transport models have been presented. Second order central differencing was used to discretize the diffusion operator and conservative differencing with third order WENO reconstruction was used to discretize the advection operator. Five second order time integrators (namely, IMEX-IIF2, IMEX-CNAB, IMEX-RK1 and IMEX-RK2 schemes) that have stiffness-resolving properties have been presented.

In Section 4.5, numerical experiments were conducted to verify the suitability of the numerical schemes and to validate the efficiency/efficacy of the stoichiometric method for

reactive transport systems. Results indicated that IMEX-IIF2, IMEX-CNAB and IMEX-RK2 were most suitable for stiff problems. Results also show that the model reduction errors induced by the stoichiometric method were negligible across space, time, schemes and flow cases. Further, CPU time, differences and relative CPU time showed that the stoichiometric method coupled with the numerical schemes can significantly reduce simulation cost.

Therefore, by coupling the stoichiometric method with the numerical schemes provided, simulation cost of reactive transport models (both linear and nonlinear models) can significantly be reduced without compromising accuracy.
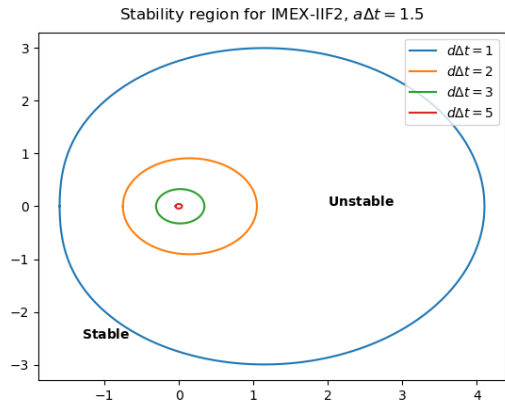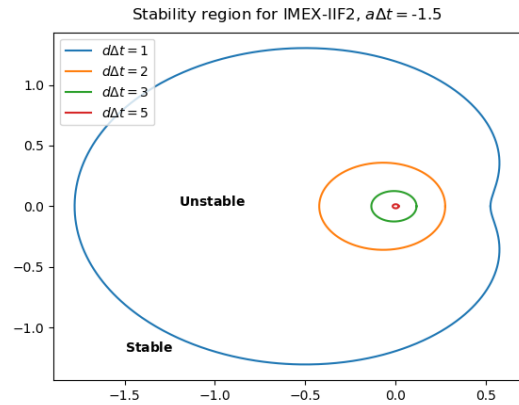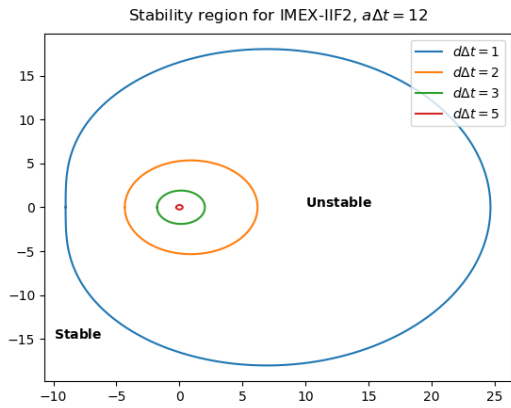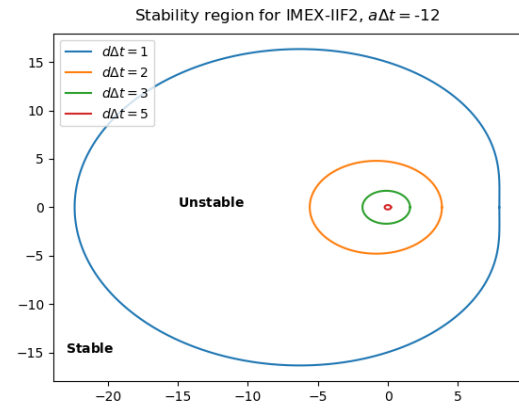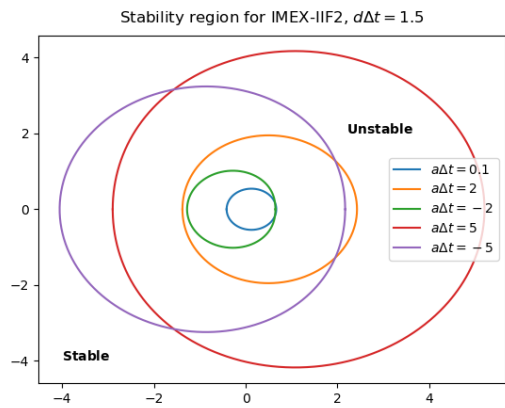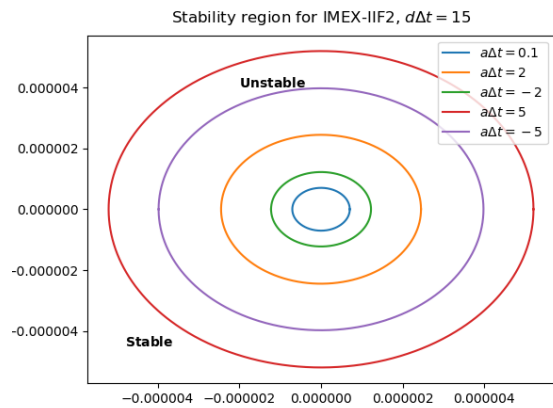
Figure 4.1: Stability regions for IMEX-IIF2, computed with different values of $d\Delta t$ and $a\Delta t$

Numerical schemes

Transport	N_x	IMEX-IIF2		IMEX-CNAB		IMEX-RK2	
		Error	Order	Error	Order	Error	Order
	10	0.3315	-	0.3325	-	0.3325	-
Diffusion ($P_e = 0$)	20	0.0817	2.0206	0.0819	2.0214	0.0819	2.0214
with	40	0.0195	2.0595	0.0196	2.0630	0.0196	2.0630
reaction	80	0.0048	2.0297	0.0048	2.0297	0.0048	2.0297
	160	0.0012	2.0000	0.0012	2.0000	0.0012	2.0000
	10	0.3389	-	0.3390	-	0.3399	-
Advection ($P_e = \infty$)	20	0.0829	2.0314	0.0829	2.0318	0.0825	2.0426
with	40	0.0198	2.0659	0.0198	2.0659	0.0199	2.0516
reaction	80	0.0050	1.9855	0.0050	1.9855	0.0050	1.9928
	160	0.0012	2.0589	0.0012	2.0589	0.0012	2.0589
	10	0.3374	-	0.3384	-	0.3393	-
Normal flow ($P_e = \Delta x$)	20	0.0825	2.0320	0.0828	2.0310	0.0823	2.0436
(advection-diffusion)	40	0.0197	2.0662	0.0197	2.0714	0.0196	2.070
with reaction	80	0.0048	2.0371	0.0048	2.0371	0.0048	2.0297
	160	0.0012	2.0000	0.0012	2.0000	0.0012	2.0000

Table 4.1: Errors and orders (L_∞) of the numerical schemes applied to solve system (4.39)-(4.40), errors were computed using analytical solution (4.42). The solutions and errors were computed for diffusion dominated, advection dominated and fully nonlinear transport cases using $\Delta t = 0.5\Delta x$.
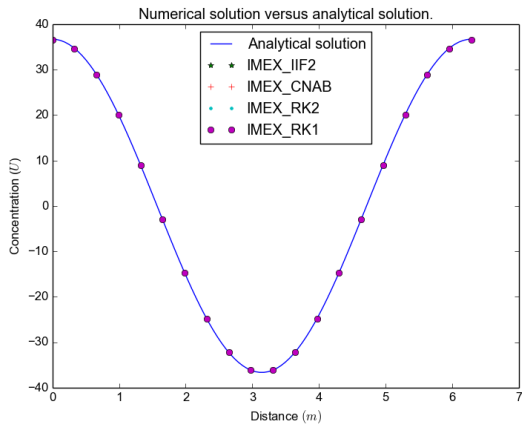
(a) Diffusion-reaction/$N_x = 20$
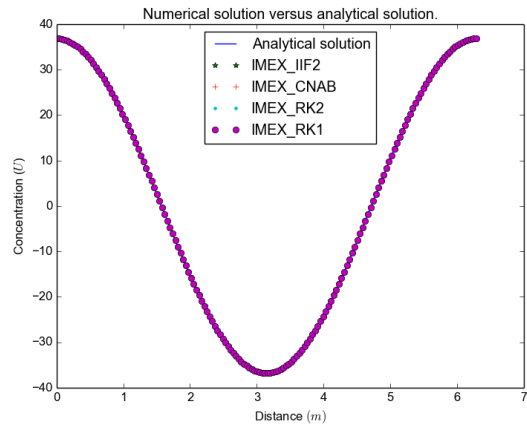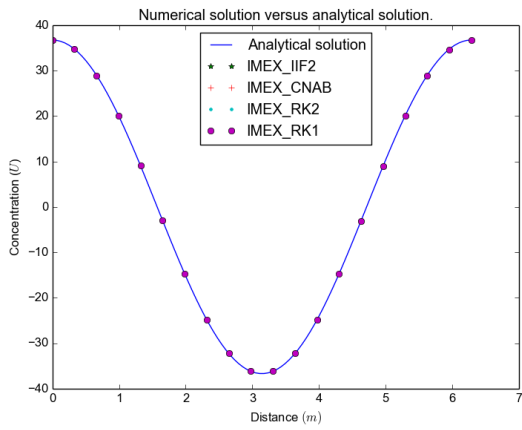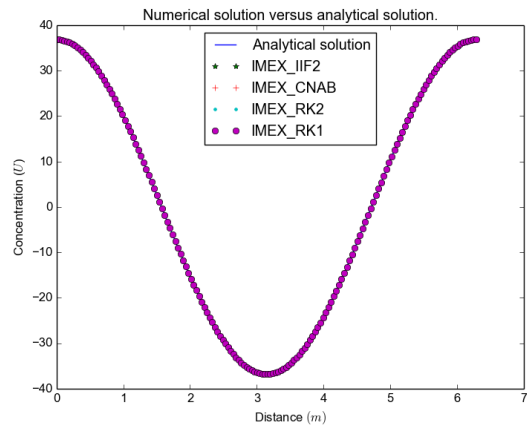
(b) Diffusion-reaction/ $N_x = 160$

(c) Advection-reaction/$N_x = 20$

(d) Advection-reaction/$N_x = 160$
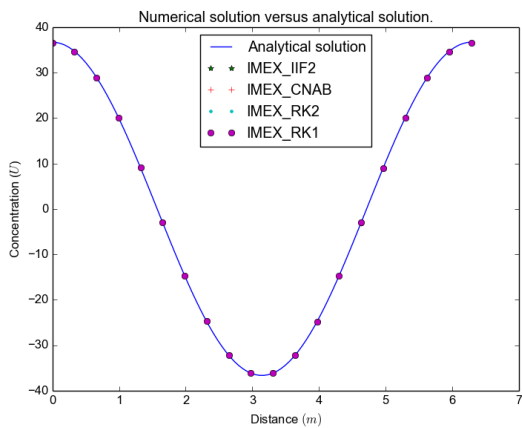
(e) Advection-diffusion-reaction/$N_x = 20$
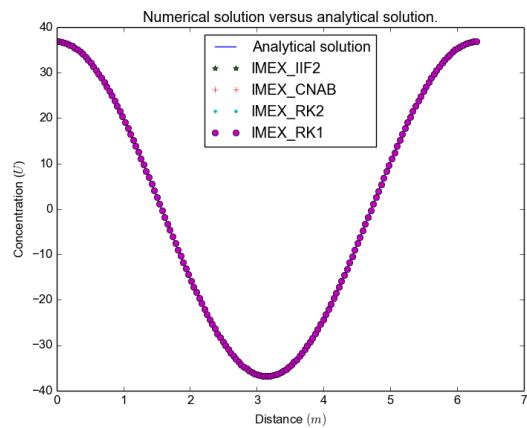
(f) Advection-diffusion-reaction/$N_x = 160$

Figure 4.2: Numerical and analytical solutions for system (4.39)-(4.40). The solutions were computed in diffusion dominated, advection dominated and normal transport cases
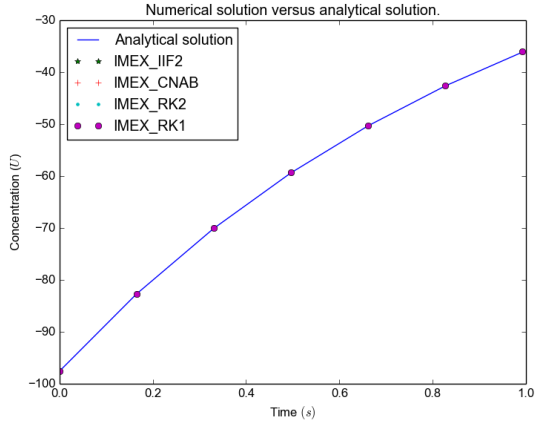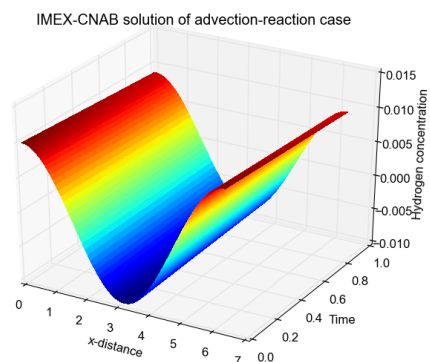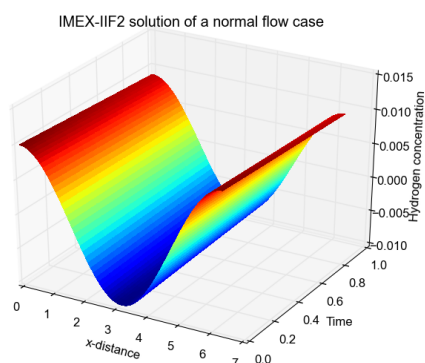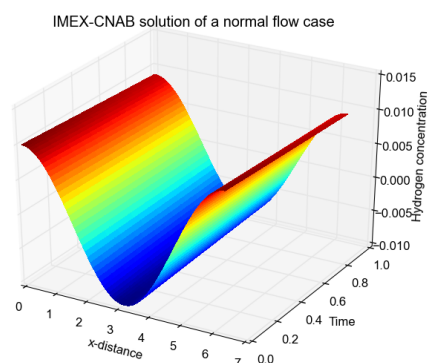
(a) Diffusion-reaction/$N_x = 20$
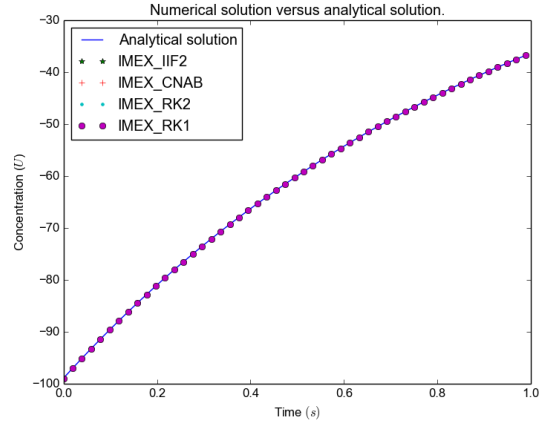
(b) Diffusion-reaction/ $N_x = 160$
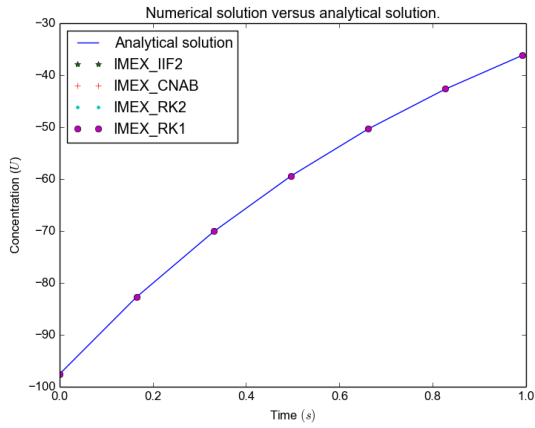
(c) Advection-reaction/$N_x = 20$
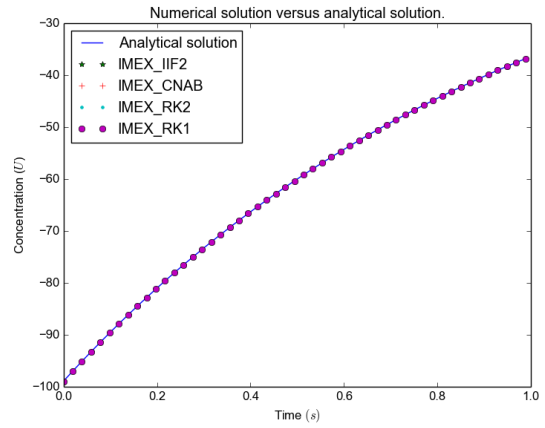
(d) Advection-reaction/$N_x = 160$
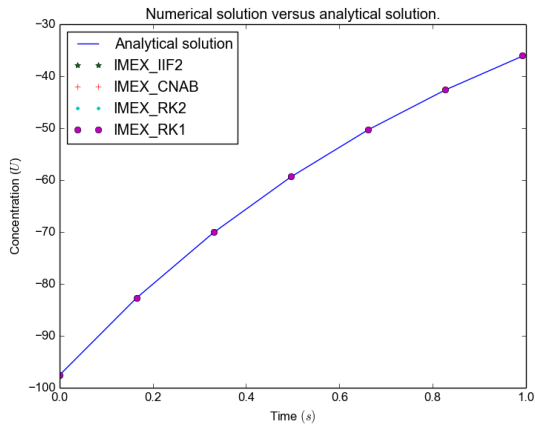
(e) Advection-diffusion-reaction/$N_x = 20$
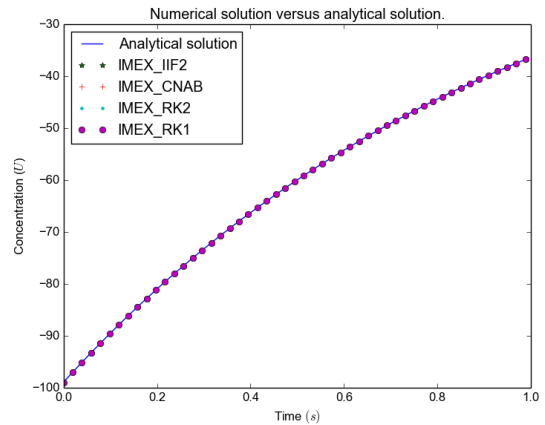
(f) Advection-diffusion-reaction/$N_x = 160$

Figure 4.3: Numerical and analytical solutions for system (4.39)-(4.40). The solutions were computed for diffusion, advection dominated and normal transport cases
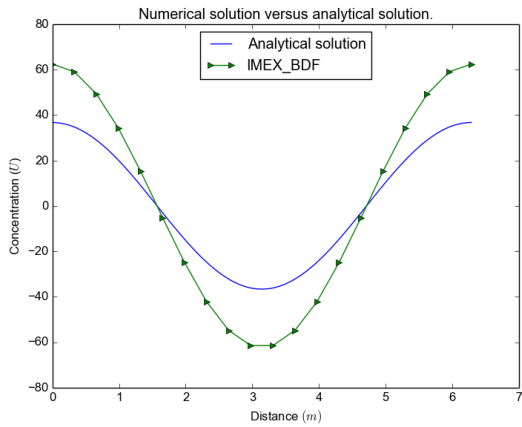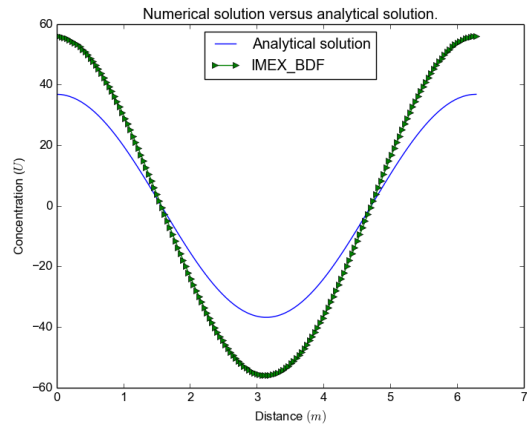
(a) Diffusion-reaction/$N_x = 160$

(b) Diffusion-reaction/ $N_x = 160$
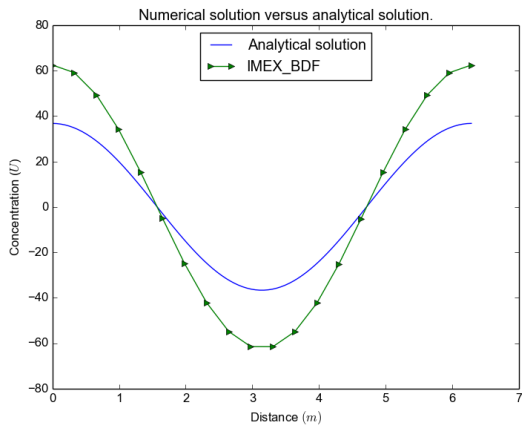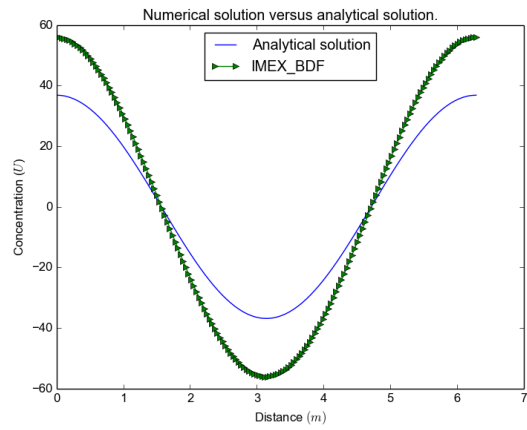
(c) Advection-reaction/$N_x = 20$

(d) Advection-reaction/$N_x = 160$
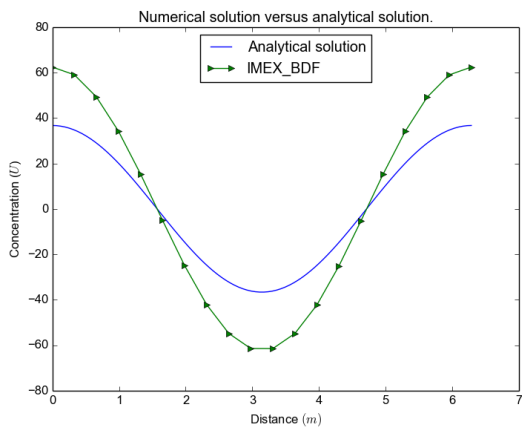
(e) Advection-diffusion-reaction/$N_x = 20$.
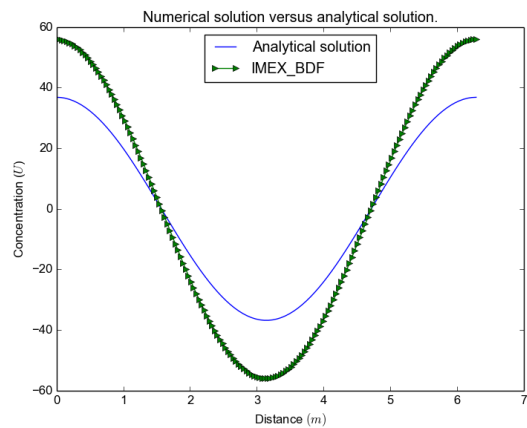
(f) Advection-diffusion-reaction/$N_x = 160$

Figure 4.4: IMEX-BDF solution and analytical solutions for system (4.39)-(4.40). The solutions were computed for diffusion, advection, and normal transport cases.
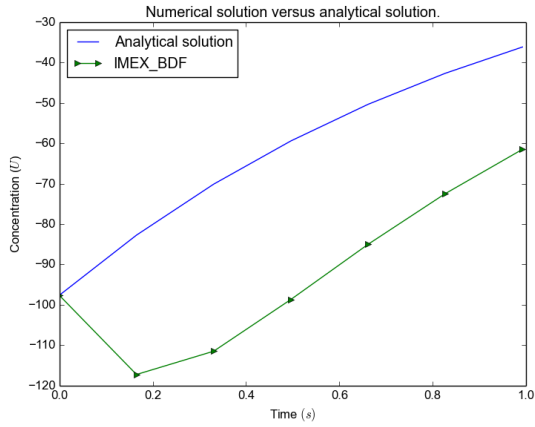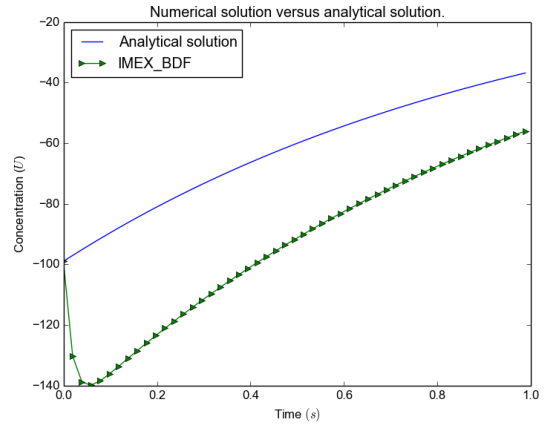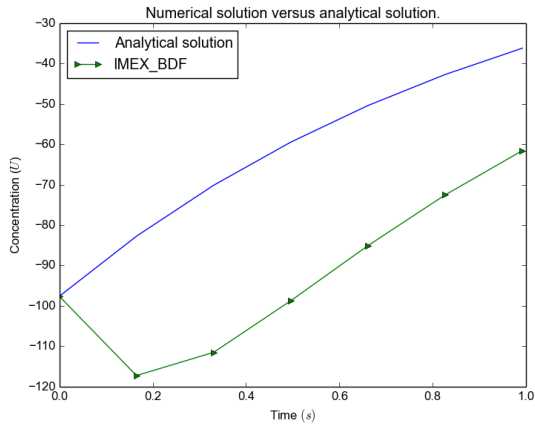
(a) Diffusion-reaction/$N_x = 20$

(b) Diffusion-reaction/ $N_x = 160$

(c) Advection-reaction/$N_x = 20$
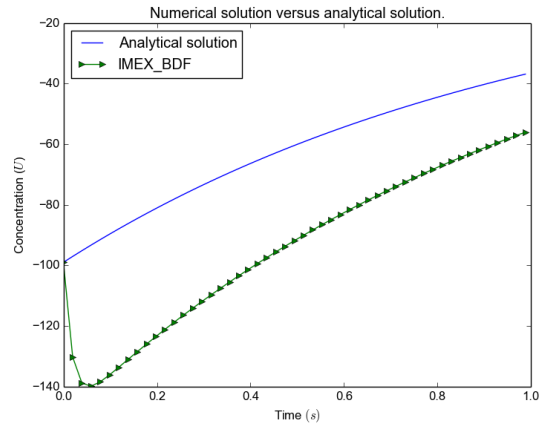
(d) Advection-reaction/$N_x = 160$
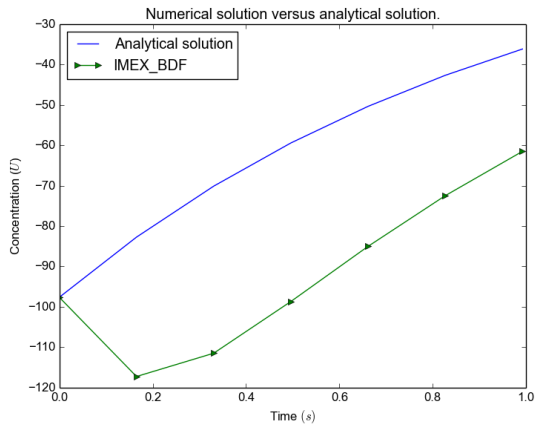
(e) Advection-diffusion-reaction/$N_x = 20$
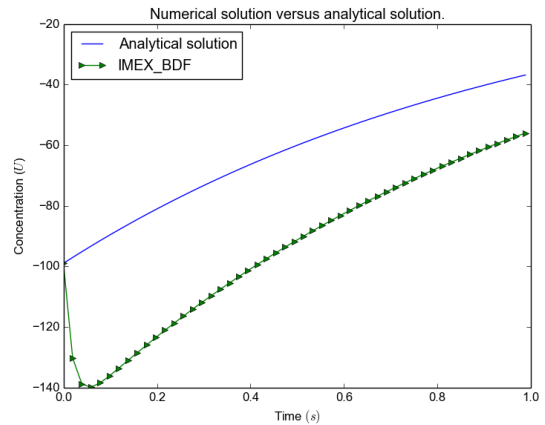
(f) Advection-diffusion-reaction/$N_x = 160$

Figure 4.5: IMEX-BDF solution and analytical solutions for system (4.39)-(4.40). The solutions are obtained for diffusion dominated, advection dominated and normal transport cases.
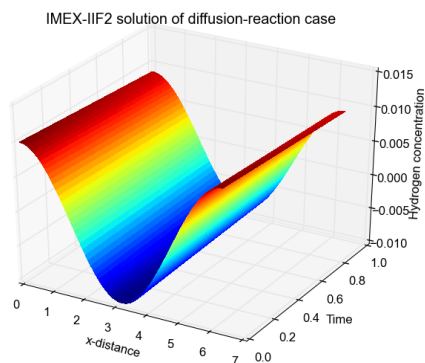
(a) Diffusion-reaction/IMEX-IIF2.
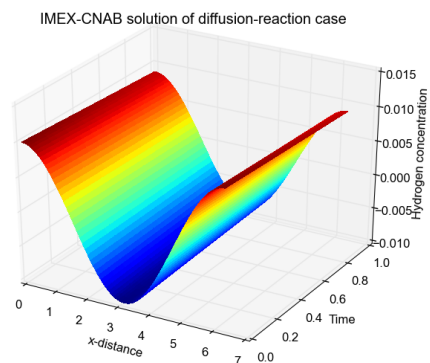
(b) Diffusion-reaction/ IMEX-CNAB.
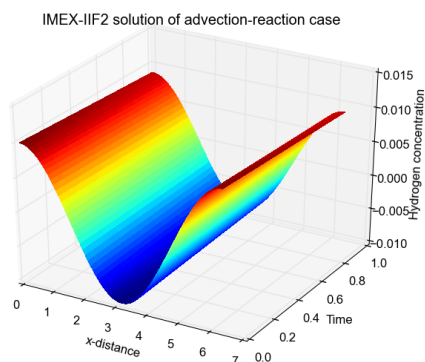
(c) Advection-reaction/IMEX-IIF2.

(d) Advection-reaction/IMEX-CNAB.

(e) Advection-diffusion-reaction/IMEX-IIF2

(f) Advection-diffusion-reaction/IMEX-CNAB

Figure 4.6: Hydrogen concentration (across both time and space) in the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13). Solutions were computed for diffusion dominated, advection dominated and fully nonlinear transport cases with reaction.
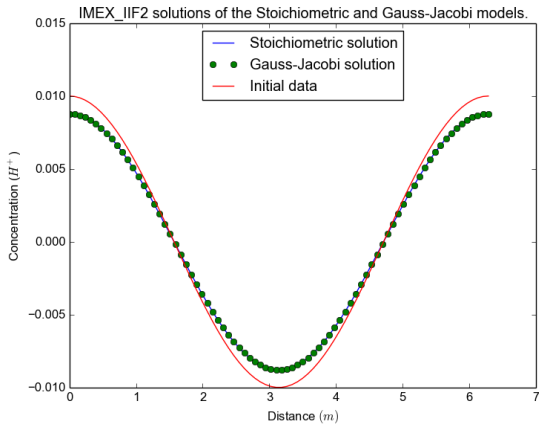
(a) Diffusion-reaction/Space.
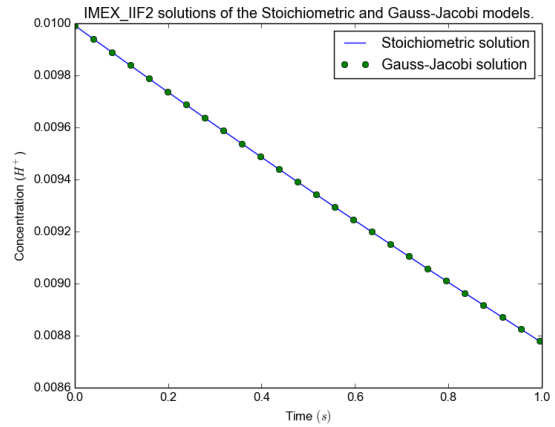
(b) Diffusion-reaction/ Time.
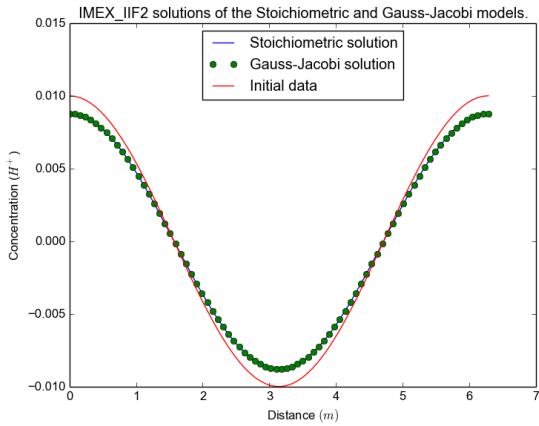
(c) Advection-reaction/Space.
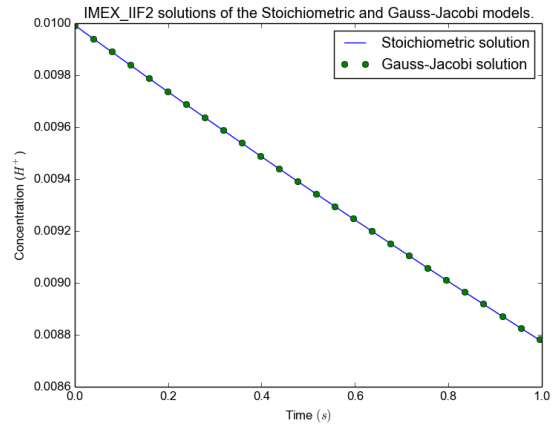
(d) Advection-reaction/Time.
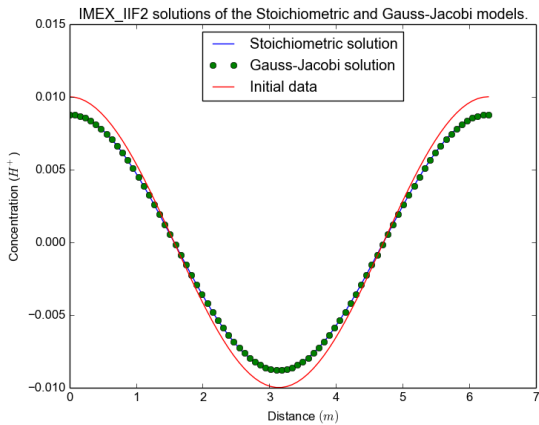
(e) Advection-diffusion-reaction/Space
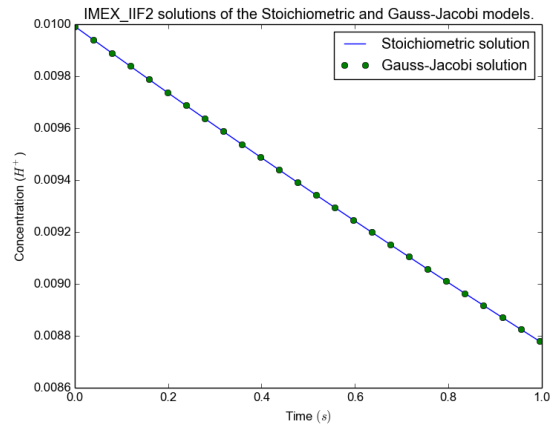
(f) Advection-diffusion-reaction/Time

Figure 4.7: Hydrogen concentration (across time or space) in the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13). Solutions were computed for diffusion dominated, advection dominated and fully nonlinear transport cases with reaction.

Diffusion dominated transport

| Diffusivity (Γ_{kk})	N_x	Numerical schemes		
		IMEX-IIF2	IMEX-CNAB	IMEX-RK2
	20	9.045526	9.045521	9.045521
	30	8.972541	8.972537	8.972537
$\Gamma_{kk} = 0.001U_k$	40	8.926316	8.926316	8.926316
	80	9.167396	9.167396	9.167396
	90	9.121991	9.121991	9.121991
	100	9.085575	9.085575	9.085575
	20	9.044596	9.044597	9.044598
	30	8.971614	8.971615	8.971615
$\Gamma_{kk} = 0.0001\cos(U_k)$	40	8.925498	8.925498	8.925498
	80	9.166528	9.166528	9.166528
	90	9.121132	9.121132	9.121132
	100	9.084723	9.084723	9.084723
	20	9.045416	9.045416	9.045416
	30	8.972428	8.972428	8.972428
$\Gamma_{kk} = 0.00001U_k^2$	40	8.926305	8.926305	8.926305
	80	9.167384	9.167384	9.167384
	90	9.121979	9.121979	9.121979
	100	9.085563	9.085563	9.085563

Table 4.2: Model reduction errors (L_∞) computed using the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13), using IMEX-IIF2, IMEX-CNAB and IMEX-RK2 discretizations. Solutions and errors were computed for a diffusion dominated nonlinear transport case. The actual values in this table are 7 orders of magnitude smaller.

Advection dominated transport

Flux	N_x	Numerical schemes		
		IMEX-IIF2	IMEX-CNAB	IMEX-RK2
Buckley-Leverett	20	9.046959	9.046959	10.119124
	30	8.973472	8.973472	9.456461
	40	8.927083	8.927083	9.068835
	80	9.167792	9.167792	9.233987
	90	9.122338	9.122338	9.174521
	100	9.085883	9.085883	9.128068
Linear	20	9.051138	9.051138	32.822316
	30	8.976056	8.976056	22.545088
	40	8.928938	8.928938	17.822199
	80	9.168702	9.168702	11.804073
	90	9.123127	9.123127	11.245699
	100	9.086579	9.086579	10.829188
Burgers	20	9.045613	9.045613	9.180696
	30	8.972561	8.972561	9.033514
	40	8.926404	8.926404	8.960603
	80	9.167436	9.167436	9.175810
	90	9.122025	9.122025	9.128628
	100	9.085604	9.085604	9.090944

Table 4.3: Model reduction errors (L_∞) computed using the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13), using IMEX-IIF2, IMEX-CNAB and IMEX-RK2 discretizations. Solutions and errors were computed for advection dominated nonlinear transport cases. The actual values in this table are 7 orders of magnitude smaller.

Fully nonlinear transport

Parameters	N_x	Numerical schemes		
		IMEX-IIF2	IMEX-CNAB	IMEX-RK2
	20	9.046969	9.046969	10.119138
Buckley-Leverett flux	30	8.973483	8.973482	9.456473
$\left(f_k(U_k) = \frac{0.016U_k^2}{(4U_k^2+(1-U_k)^2)} \right)$	40	8.927094	8.927095	9.197874
with	80	9.167804	9.167804	9.233999
$\Gamma_{kk} = 0.0001\cos(U_k)$	90	9.122349	9.122349	9.174533
	100	9.085894	9.085894	9.128079
	20	9.045624	9.045623	9.180769
Burgers flux	30	8.972572	8.972572	9.033525
$\left(f_k(U_k) = 0.002U_k^2 \right)$	40	8.926416	8.926415	8.960614
with	80	9.167448	9.167448	9.175822
$\Gamma_{kk} = 0.0001U_k$	90	9.122026	9.122026	9.128629
	100	9.085605	9.085605	9.090945
	20	9.044792	9.044794	9.179845
Burgers flux	30	8.971747	8.971748	9.032685
$\left(f_k(U_k) = 0.002U_k^2 \right)$	40	8.925597	8.925597	8.959787
with	80	9.166580	9.166580	9.174952
$\Gamma_{kk} = 0.0001\cos(U_k)$	90	9.121178	9.121178	9.127779
	100	9.084764	9.084764	9.090102

Table 4.4: Model reduction errors (L_∞) computed using the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13), using IMEX-IIF2, IMEX-CNAB and IMEX-RK2 discretizations. Solutions and errors were computed for fully nonlinear transport cases. The actual values in the table are 7 orders of magnitude smaller.
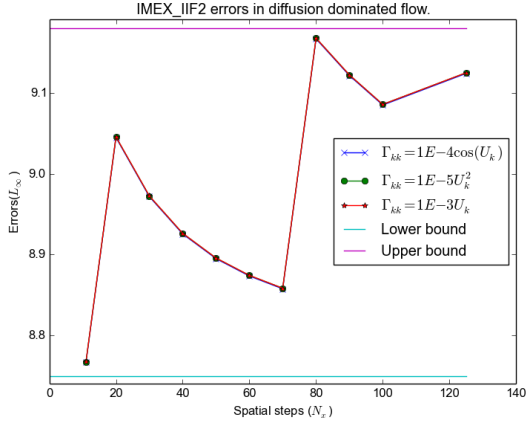
(a) Diffusion-reaction/IMEX-IIF2 errors.
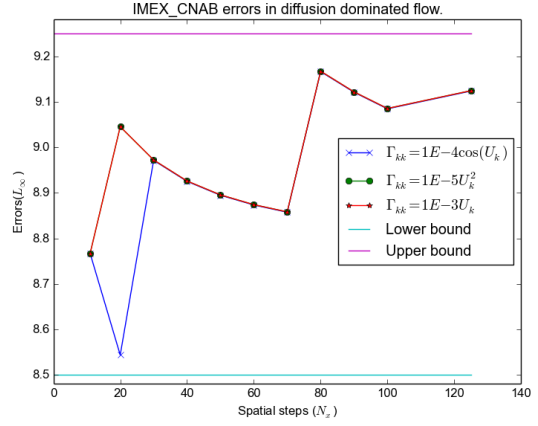
(b) Diffusion-reaction/IMEX-CNAB errors.
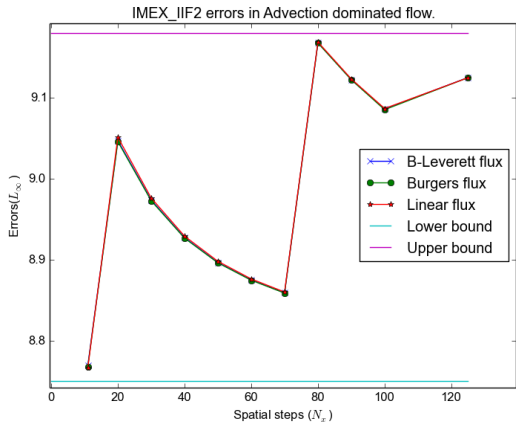
(c) Advection-reaction/IMEX-IIF2 errors.
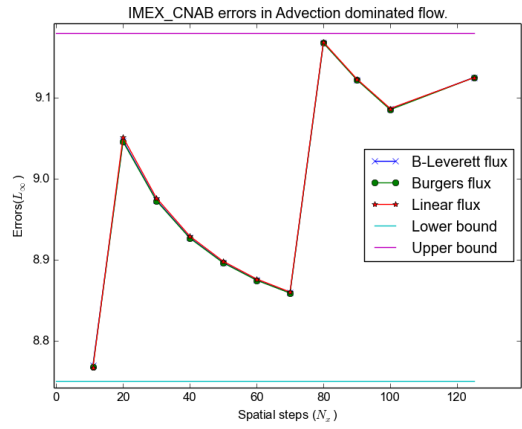
(d) Advection-reaction/IMEX-CNAB errors.
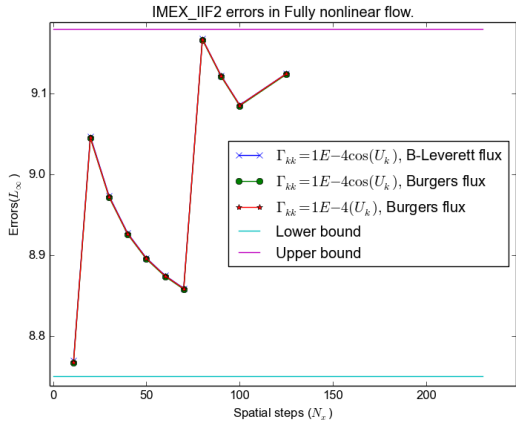
(e) Fully nonlinear model/IMEX-IIF2 errors
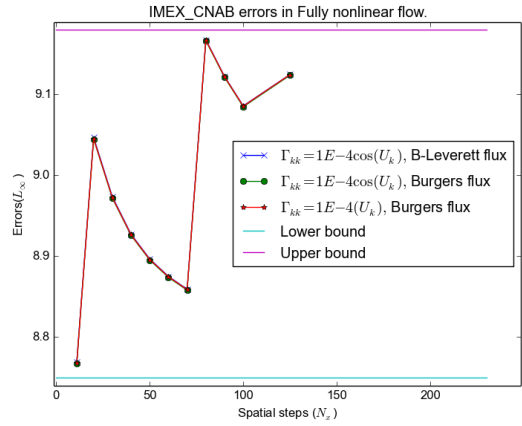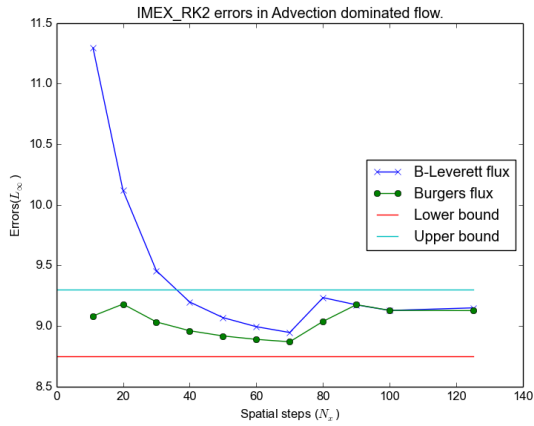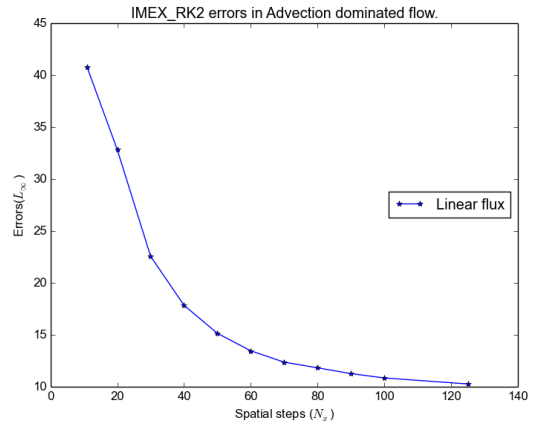
(f) Fully nonlinear model/IMEX-CNAB errors

Figure 4.8: Model reduction errors (L_∞) computed using the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13), using IMEX-IIF2 and IMEX-CNAB discretizations. Solutions and errors were computed for a diffusion, advection and fully nonlinear transport cases. The actual values in this figure are 7 orders of magnitude smaller.
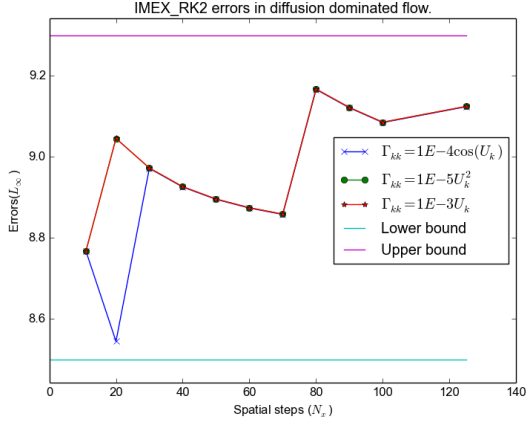
(a) Diffusion-reaction/IMEX-RK2 errors.
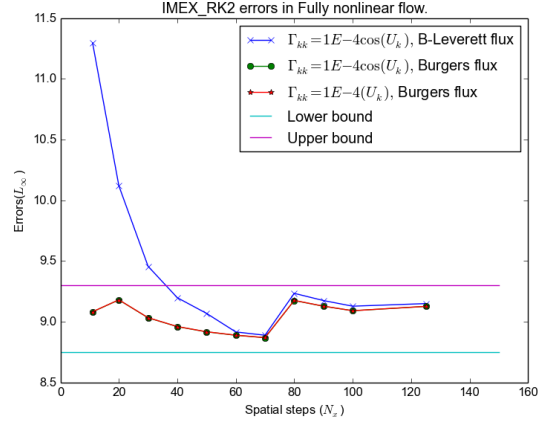
(b) Fully nonlinear model/ IMEX-RK2.

(c) Advection-reaction/IMEX-RK2 errors.

(d) Advection-reaction/IMEX-RK2 errors.

Figure 4.9: Model reduction errors (L_∞) computed using the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13), using IMEX-RK2 discretization. Solutions and errors were computed for diffusion, advection and fully nonlinear transport cases. The actual values in this figure are 7 orders of magnitude smaller.
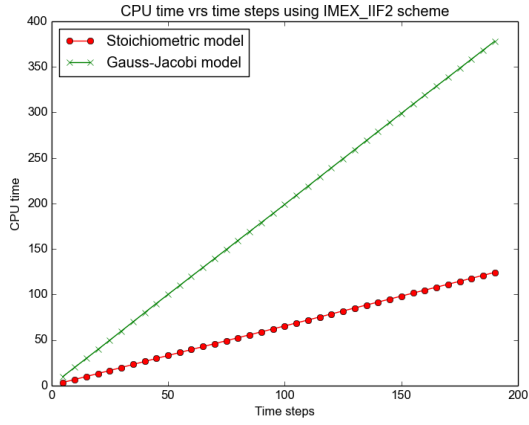
(a) CPU time/IMEX-IIF2 scheme.
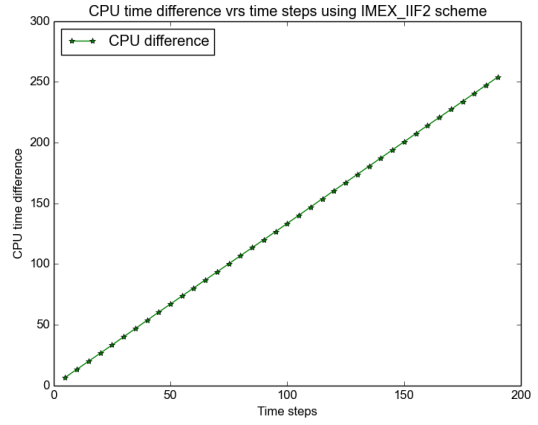
(b) CPU difference/IMEX-IIF2 scheme.
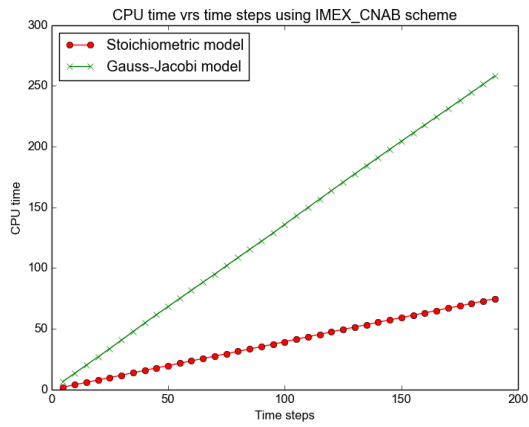
(c) CPU time/IMEX-CNAB scheme.
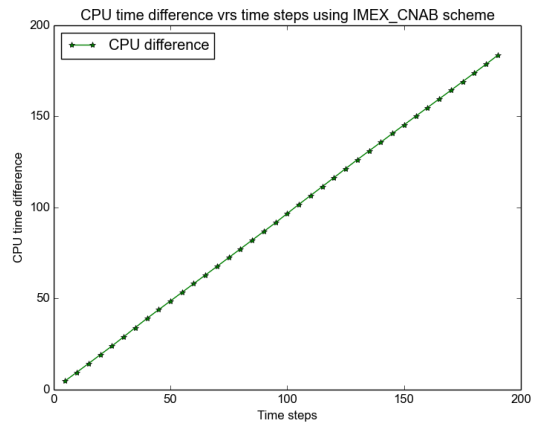
(d) CPU difference/IMEX-CNAB scheme.
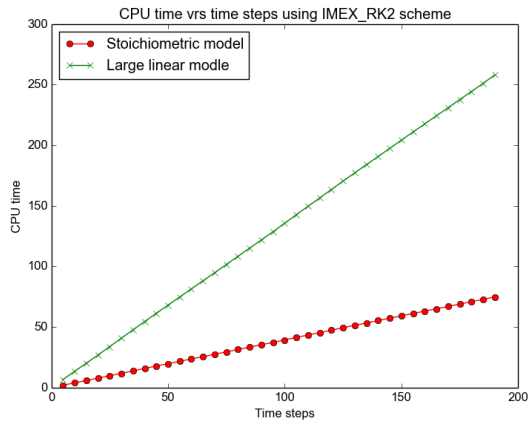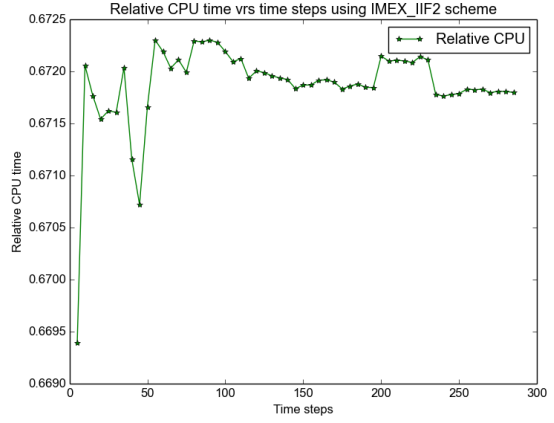
(e) CPU time /IMEX-RK2 scheme
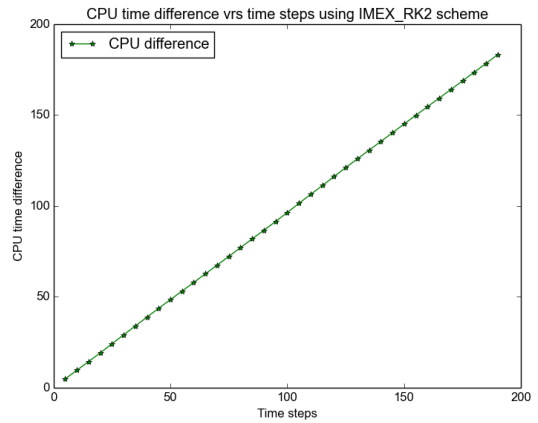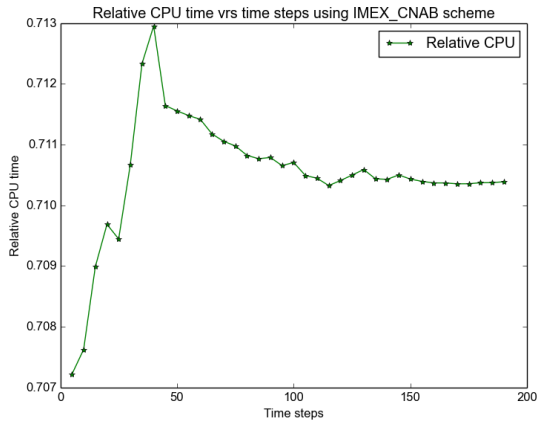
(f) CPU difference/IMEX-RK2 scheme

Figure 4.10: CPU time and differences for the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13), using IMEX-IIF2, IMEX-CNAB and IMEX-RK2 discretizations. The solutions were computed for a fully nonlinear flow case, using Burgers flux, $\Gamma_{kk} = 0.0001 \cos(U_k)$, $N_x = 600$, and $\Delta t = 0.5 \Delta x$.
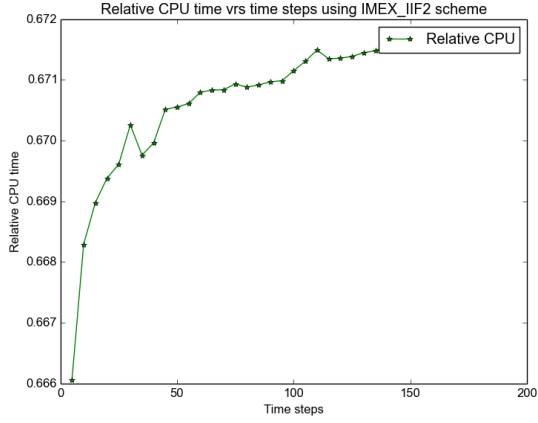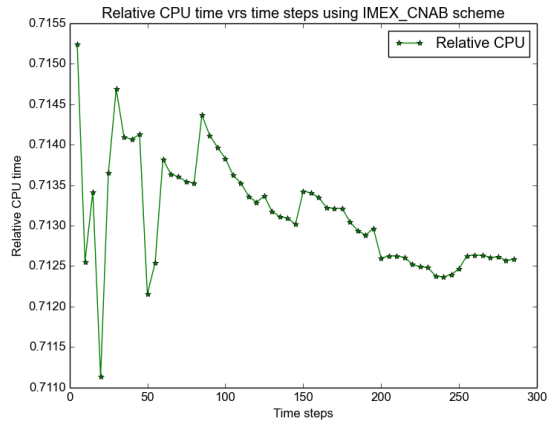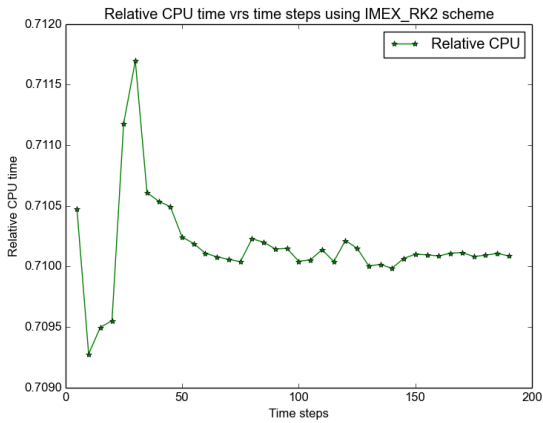
(a) Relative CPU/IMEX-IIF2/$N_x = 600$.
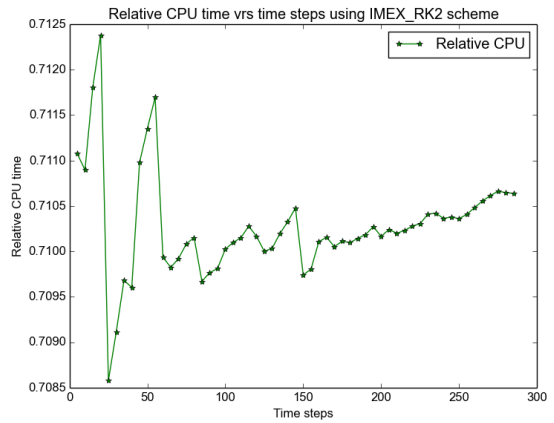
(b) Relative CPU /IMEX-IIF2/$N_x = 900.$.

(c) Relative CPU /IMEX-CNAB/$N_x = 600$.

(d) Relative CPU /IMEX-CNAB/$N_x = 900$.

(e) Relative CPU /IMEX-RK2/$N_x = 600$.

(f) Relative CPU time/IMEX-RK2/$N_x = 900$.

Figure 4.11: Relative CPU time for the Gauss-Jacobi model (4.8)-(4.10) and stoichiometric model (4.13), using IMEX-IIF2, IMEX-CNAB and IMEX-RK2 discretizations. The solutions were computed for a fully nonlinear flow case, using Burgers flux, $\Gamma_{kk} = 0.0001 \cos(U_k)$, $N_x = 600$, or 900, and $\Delta t = 0.5 \Delta x$.

Chapter 5

Dispersion preserving schemes for reactive transport processes

In this chapter, global spectral analysis of reactive transport is discussed. Dispersion relations are used to analyse the dispersion-preserving abilities of some temporal numerical schemes, in order to establish their suitability for simulating reactive transport processes.

Remark: *The results presented here are under review for possible publication.*

5.1 Introduction

Numerical schemes are alternative approaches to analytical approaches in simulating complicated phenomena. However, numerical schemes are approximations to exact solutions, thus, the level of accuracy of numerical schemes depend on the ability of the scheme to closely mimic properties of governing equations [39, 32, 157]. Inaccurate numerical schemes introduce non-physical behaviours in the solution, thus, it is necessary to develop accurate numerical schemes as alternatives to the exact solutions [157]. This requires rigorous error analysis of numerical schemes for a particular model.

However, the non-availability of exact solutions for most complex phenomena (especially non-linear problems) makes error analysis of numerical schemes a difficult task and sometimes impossible without simplifying assumptions. Models are often linearised and then followed by analyses to obtain analytical relations (e.g. exact solutions, dissipation and dispersion expressions), which are relevant for analysing numerical schemes [157].

There are many techniques for analysing numerical schemes, the popular techniques include Taylor expansions, Von Neumann stability analysis [157, 163, 154], GKS stability theory [157, 60] and time-stability analysis [26, 186].

Further, each technique has merits and limitations depending on the problem under consideration. Spectral analysis has been shown in [157, 137] to be efficient for analysing parabolic and hyperbolic problems. Spectral analysis of parabolic or hyperbolic problems result in dispersion/dissipation relations which can be used to analyse the numerical schemes for corresponding problems. Numerical schemes that satisfy dispersion relations for a particular propagation problem are said to be dispersion preserving (DP) schemes for that problem.

Global spectral analysis has recently gained popularity in research involving propagation problems [137, 143, 145, 144]. The authors in [141] have analysed numerical schemes for linear advection equation, by adapting the spectral analysis presented in [170]. By performing spectral analysis on the linear advection equation, the authors in [140] have shown that numerical methods have different error and signal dynamics. Finite difference schemes for linear diffusion equation have been analysed in [138]. One notable observation in the linear advection and diffusion cases is that, numerical phase speed and numerical diffusion are not constant. Positivity preserving Galerkin method for advection-diffusion-reaction have also been analysed in [77] using the spectral approach. In [139] spectral analysis of Galerkin finite element method for the advection equation have been discussed. One observation with the finite element methods is good error dispersion behaviour.

Furthermore, compact finite difference and finite volume methods for Euler equations have been discussed in [142]. The compact finite difference methods proved superior in the case of Riemann problems (for the Euler equations). Finite difference methods for Navier-Stokes equations have been analysed by traditional techniques in [150, 89, 31, 175]. In [157], the authors extended the spectral approach to compact finite difference methods for linear advection-diffusion equation in order to ascertain accurate schemes for the Navier-Stokes equations.

In this discussion, time integrators for the linear advection-diffusion-reaction equation are analysed in order to ascertain accurate integrators for chemical transport models. As stated in previous chapters (and discussed in [19, 169, 132, 21, 39, 107, 184]), not

all numerical schemes are suitable for chemical models, due to high degrees of freedom, stiffness and positivity constraints. Theta integrators and second order diagonally implicit Runge-Kutta (DIRK2) integrators have been shown to be robust in well-mixed chemical problems (discussed in Chapter 3), and IMEX schemes have been shown to be robust in chemical transport problems (discussed in Chapter 4). In this chapter, dispersion properties of the Theta, DIRK2 and some IMEX schemes will be discussed.

In Section 5.2, the unknown concentration variable and linear transport equation are transformed to spectral variable and equation respectively, and exact dispersion relations and solutions are derived. In Section 5.3, general numerical dispersion relations for the linear transport equation are derived. In Section 5.4, dispersion analysis of the Theta, DIRK2, and some IMEX integrators are discussed. In Section 5.5, numerical experiments are conducted to confirm the analyses. Finally, the chapter is concluded in Section 5.6.

5.2   Spectral analysis of reactive transport equation

Chemical transport phenomena involves physical processes (such as advection and diffusion) and chemical reactions. A simple model for such phenomena is the linear 1D advection-diffusion-reaction equation given by:

$$\frac{\partial U}{\partial t} + u\frac{\partial U}{\partial x} = \Gamma\frac{\partial^2 U}{\partial^2 x} + K_r U, \tag{5.1}$$

where U denotes concentration of a chemical species, x is a space variable, t is a time variable, u denotes constant transport velocity, K_r is chemical reactivity and Γ denotes constant diffusivity. Three special cases of the linear equation (5.1) occur when the reactivity K_r assumes zero, positive or negative values. Equation (5.1) models pure diffusive, diffusive-destructive and diffusive-productive physico-chemical problems when the reactivity K_r assumes zero, negative or positive values respectively.

The exact dispersion relations for the linear transport equation are firstly, determined by applying global spectral analysis. Thus, by transforming the concentration function $U(x,t)$, from the $x-t$ plane to the hybrid spectral plane [157, 137, 141], we obtain:

$$U(x,t) = \int \mathcal{U}(\kappa,t)e^{i\kappa x}d\kappa, \tag{5.2}$$

where $\mathcal{U}(\kappa, t)$ is the Fourier amplitude and κ is the wave number. Using the spectral transform (5.2), the linear chemical transport model (5.1) can be represented in the spectral plane as follows:

$$\frac{\partial \mathcal{U}}{\partial t} + iu\kappa \mathcal{U} = -\Gamma \kappa^2 \mathcal{U} + K_r \mathcal{U}, \tag{5.3}$$

Given a general initial condition $U(x, 0) = \int \mathcal{U}_0(\kappa) e^{i\kappa x} d\kappa$, the exact solution of (5.3) can be obtained as follows:

$$\mathcal{U}(\kappa, t) = \mathcal{U}_0(\kappa) e^{-(i\kappa u + \kappa^2 \Gamma - K_r)t}. \tag{5.4}$$

Moreover, the concentration of the chemical species can also be represented in the bi-dimensional Fourier-Laplace transform [157]:

$$U(x, t) = \iint \mathcal{U}(\kappa, \omega) e^{i(\kappa x - \omega t)} d\kappa d\omega, \tag{5.5}$$

where ω denotes circular frequencies. Applying the Fourier-Laplace transform (5.5) directly in the linear convection-diffusion-reaction equation (5.1), yields the physico-chemical dispersion relation:

$$\omega = u\kappa - i(\kappa^2 \Gamma - K_r). \tag{5.6}$$

Wave propagation problems are characterised by dispersion relations, that yield information about the phase and group velocities of propagating signals [157]. Dispersion relation (5.6) is different from the one obtained in [157] for advection-diffusion equation, as it indicates that chemical reactions also affect signal propagation. Any numerical algorithm for solving a wave propagation problem can only be suitable/accurate if it satisfies the dispersion relation for the problem [157, 137, 140]. In [144], multiple time level schemes have been analysed based on correct dispersion relations for convection problems and in [157], the idea has been adapted for convection-diffusion problems. In this discussion, the same idea will be used to analyse the linear chemical transport problem.

From the dispersion relation (5.6), one can obtain the complex physico-chemical phase speed for the linear chemical transport equation as follows:

$$u_{ex} = \frac{\omega}{\kappa} = u - i\left(\kappa \Gamma - \frac{K_r}{\kappa}\right). \tag{5.7}$$

And from the same relation (i.e. Equation (5.6)) the physical group velocity is given by [157]:

$$V_{ex} = \frac{d\omega}{d\kappa} = u - i2\kappa\Gamma. \tag{5.8}$$

Denoting the real and complex parts of the physical group velocities by $Re(V_{ex})$ and $Im(V_{ex})$ respectively, and using the relation (5.8) the diffusivity can be expressed as follows:

$$\Gamma = \frac{i}{2\kappa}\big[Re(V_{ex}) - u\big] - \frac{Im(V_{ex})}{2\kappa} \tag{5.9}$$

Since diffusivity is a real quantity for physical systems, the imaginary part of (5.9) must vanish, thus:

$$Re(V_{ex}) = u. \tag{5.10}$$

Furthermore, the diffusivity in the linear transport equation (5.1) is positive, the real part of the complex quantity (5.9) must be positive, implying:

$$Im(V_{ex}) < 0. \tag{5.11}$$

Denote the time step size by Δt, spatial step size by Δx, Peclet number by $P_e = \frac{\Gamma \Delta t}{\Delta x^2}$, Courant-Friedrich-Lewis (CFL) number by $C_f = \frac{u\Delta t}{\Delta x}$, $\lambda_r = K_r \Delta t$, and the physico-chemical amplification factor by G_{ex}. The amplification factor of a numerical method or the exact solution is defined as follows [157]:

$$G_{ex} = \frac{\mathcal{U}(k, t + \Delta t)}{\mathcal{U}(\kappa, t)} \tag{5.12}$$

Using the analytical solution (5.4) in definition (5.12) yields the following physico-chemical amplification factor:

$$G_{ex} = e^{-(i\kappa u + \kappa^2 \Gamma - K_r)\Delta t}, \tag{5.13}$$

Using Peclet and CFL number relations, the amplification factor (5.13) can be re-expressed as follows:

$$G_{ex} = e^{-i\omega\Delta t}$$
$$= e^{-P_e(\kappa\Delta x)^2 + \lambda_r} e^{-iC_f(\kappa\Delta x)}. \tag{5.14}$$

The magnitude of the physico-chemical amplification factor G_{ex} given by (5.12) will always be less than one (i.e. the usual stability criterion for numerical methods) if:

$$-P_e(\kappa\Delta x)^2 + \lambda_r < 0. \tag{5.15}$$

One can observe that condition (5.15) is always satisfied by pure diffusive problems (where $\lambda_r = 0$) and diffusive-destructive problems (where $\lambda_r < 0$) for the entire range of $\kappa\Delta x$. However, diffusive-productive problems (where $\lambda_r > 0$) will satisfy condition (5.15) only if:

$$\kappa\Delta x > \sqrt{\frac{\lambda_r}{P_e}}. \tag{5.16}$$

Thus $(\kappa\Delta x)_c = \sqrt{\frac{\lambda_r}{P_e}}$ is a critical dimensionless number that determines the constraint $|G_{ex}| < 1$ for diffusive-productive problems.

The absolute part of G_{ex} in (5.14) is the amplification factor of the exact solution. Contours (in $\kappa\Delta x - C_f$ plane) of the absolute part of G_{ex} are shown in figure 5.1 for the pure diffusive case using Peclet numbers; $P_e = 0.01, 0.05, 0.25$ and 0.5. The amplification factor decreases across increasing wave numbers but does not depend on CFL numbers. Figures 5.2b-5.2f show contours of $|G_{ex}|$ in the diffusive-destructive case. Figures 5.2a-5.2e show contours of $|G_{ex}|$ in the diffusive-productive case. The observations in the productive and destructive cases are not very different from the pure diffusive case. Numerical integrator for the linear problem are therefore, required to closely mimic this behaviour.

5.3 Numerical dispersion relations

In the spectral analysis of the 1D equation, exact derivatives of (5.2) were used, however, numerical approximations can be applied instead. If the spectral representation (5.2) is substituted into the advection-diffusion reaction equation and the advection and diffusion derivatives are numerically approximated, a semi-discrete equation is obtained as follows:

$$\frac{\partial \mathcal{U}}{\partial t} + iu_{apx}\kappa\mathcal{U} = -\Gamma_{apx}\kappa^2\mathcal{U} + K_r\mathcal{U}, \tag{5.17}$$
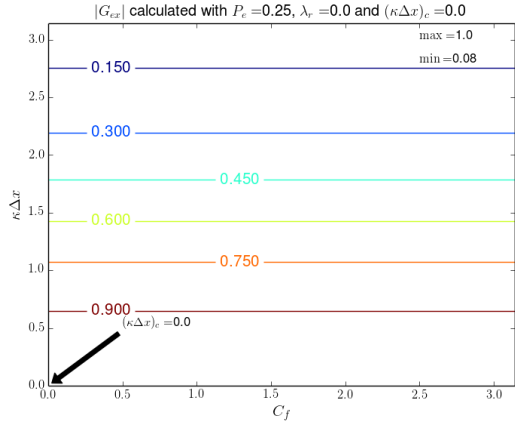
where u_{apx} and Γ_{apx} result from the numerical approximations of the advection and diffusion terms, respectively. The quantities u_{apx} and Γ_{apx} represent numerical phase speed and numerical diffusion, respectively.
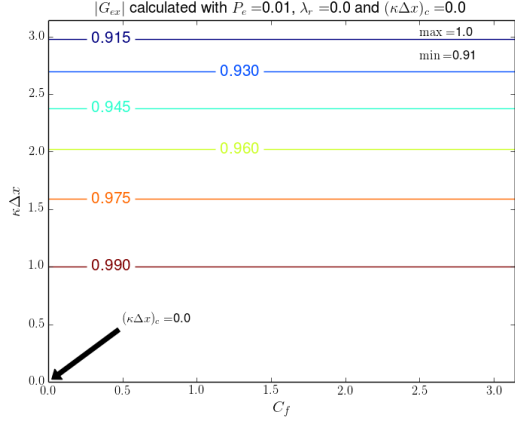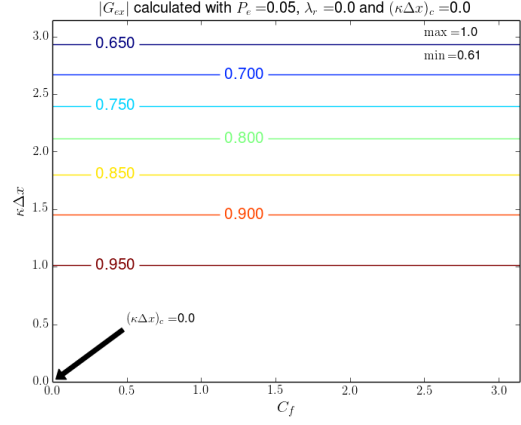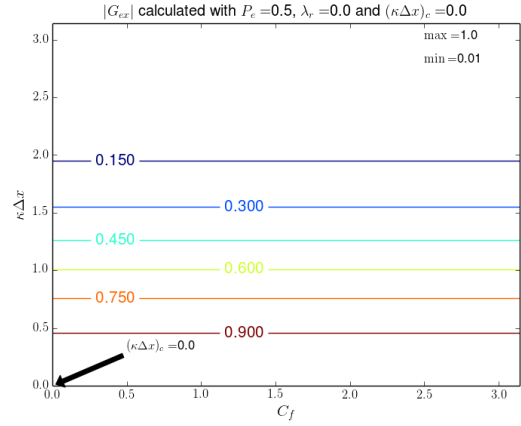
(a) $P_e = 0.01, \lambda_r = 0.0$ (b) $P_e = 0.05, \lambda_r = 0.0$

(c) $P_e = 0.25, \lambda_r = 0.0$ (d) $P_e = 0.5, \lambda_r = 0.0$

Figure 5.1: Contour for the physico-chemical amplification factor $|G_{ex}|$ of the analytical solution for 1D linear convection-diffusion-reaction equation. The contours display the behaviour of $|G_{ex}|$ for varying Peclet numbers in a pure diffusive problem.

The analytical solution of the semi-discrete equation (5.17) with initial data $U(x,0) = \int \mathcal{U}_0(\kappa)e^{i\kappa x}d\kappa$ is given by:

$$\mathcal{U}(\kappa, t) = \mathcal{U}_0(\kappa)e^{-(iu_{apx}\kappa + \Gamma_{apx}\kappa^2 - K_r)t}. \tag{5.18}$$

Further, applying the Fourier-Laplace transform (5.5) to the linear convection-diffusion-reaction equation (5.1) and numerically approximating the space-time derivatives yields the following numerical physical-chemical dispersion relation:

$$\omega_{apx} = u_{apx}\kappa - i(\kappa^2\Gamma_{apx} - K_r). \tag{5.19}$$

Analogously, numerical approximations of the linear problem (5.1) possess amplifica-

tion factors that control the behaviour of numerical solutions as time evolves. To achieve acceptable accuracy it is necessary for the numerical schemes to preserve some vital properties of the linear problem [157]. Thus, the amplification factor of the numerical schemes should closely mimic the behaviour of the analytical amplification factor. Applying the numerical solution (5.18) in definition (5.12) yields the following numerical amplification factor:

$$G_{apx} = e^{-(\kappa^2 \Gamma_{apx} - K_r)\Delta t} e^{-i\kappa u_{apx}\Delta t}. \tag{5.20}$$

In general, the numerical phase speed u_{apx} and numerical diffusion Γ_{apx} vary across numerical schemes. According to the numerical amplification factor (5.20), the shift in phase for each time step is given by:

$$\phi = \kappa u_{apx}\Delta t, \tag{5.21}$$

and is related to the amplification factor by the expression:

$$\tan(\phi) = -\left(\frac{Im(G_{apx})}{Re(G_{apx})}\right). \tag{5.22}$$

Thus, a normalized numerical phase velocity u_{apx}, relating the amplification factor can be obtained from (5.21) and (5.22) as follows:

$$\frac{u_{apx}}{u} = -\frac{1}{C_f(k\Delta x)} \arctan\left(\frac{Im(G_{apx})}{Re(G_{apx})}\right). \tag{5.23}$$

The numerical group velocity V_{apx}, is the derivative of the numerical dispersion relation (5.19) with respect to wave number, i.e. $V_{apx} = \frac{d\omega_{apx}}{d\kappa}$ [157]. Normalizing and simplifying the resulting expression yields:

$$\frac{V_{apx}}{V} = -\frac{1}{C_f}\frac{d}{d(k\Delta x)}\left(\arctan\left(\frac{Im(G_{apx})}{Re(G_{apx})}\right)\right). \tag{5.24}$$

An expression relating numerical diffusion Γ_{apx}, with the magnitude of the amplification factor can be obtained from (5.20) as follows:

$$\frac{\Gamma_{apx}}{\Gamma} = \frac{\lambda_r - \ln|G_{apx}|}{P_e(\kappa\Delta x)^2}. \tag{5.25}$$

The normalized numerical diffusion relation (5.25) of a particular scheme is unity if the scheme has the same diffusion as the exact solution. If the relation is greater than one,

then the scheme is highly diffusive and if the relation is less than one, then the method has low diffusion. If the diffusion relation is negative then the numerical scheme is anti-diffusive for a diffusive problem, thus the scheme is unstable. Similar interpretations are given to the normalized relations for the phase speed, group velocity and amplification factors. Thus, to achieve high accurate numerical simulation of the time-dependent advection-diffusion-reaction problem, it is important that numerical integrators preserve most of the physical and chemical dispersion properties including amplification factor, numerical diffusion, phase speed and group velocity. The Theta and DIRK2 schemes that have been discussed earlier, will be analysed using the normalized relations (to determine their dispersion-preserving abilities), in order to establish their suitability for reactive transport problems.

Proposition 5.1 (Theta scheme). *The amplification factor of the theta scheme (with θ in the range $0 \leq \theta \leq 1$) for the 1D advection-diffusion reaction equation* (5.1) *is given by:*

$$G_\theta(\theta) = \frac{(1-\theta P)(1+(1-\theta)P) + \theta(1-\theta)(C_f \kappa \Delta x)^2}{\left(1-\theta P\right)^2 + (\theta C_f \kappa \Delta)^2} - i \frac{C_f \kappa \Delta x}{\left(1-\theta P\right)^2 + (\theta C_f \kappa \Delta)^2},$$

where $P = -P_e(\kappa \Delta x)^2 + \lambda_r$.

Proof. Applying Fourier-spectral method to the space derivatives of the 1D problem (5.1) yields the ODE (5.3), implying:

$$F(\mathcal{U}) = -i\omega \mathcal{U}. \tag{5.26}$$

Therefore, the theta-Fourier spectral scheme states that:

$$\left(1 + \theta i\omega \Delta x\right)\mathcal{U}^{n+1} = \left(1 - \theta i\omega \Delta x\right)\mathcal{U}^n. \tag{5.27}$$

Thus using (5.27), the theta amplification factor $G_\theta = \frac{\mathcal{U}^{n+1}}{\mathcal{U}^n}$, can be derived as follows:

$$G_\theta = \frac{1 - \theta i\omega \Delta x}{1 + \theta i\omega \Delta x}. \tag{5.28}$$

Substituting,

$$-i\omega \Delta x = -iC_f(\kappa \Delta x) - P_e(\kappa \Delta x)^2 + \lambda_r, \tag{5.29}$$

and manipulating algebraically yields:

$$G_\theta = \frac{(1-\theta P)(1+(1-\theta)P) + \theta(1-\theta)(C_f \kappa \Delta x)^2}{\left(1-\theta P\right)^2 + (\theta C_f \kappa \Delta)^2} - i\frac{C_f \kappa \Delta x}{\left(1-\theta P\right)^2 + (\theta C_f \kappa \Delta)^2}, \quad (5.30)$$

where $P = -P_e(\kappa \Delta x)^2 + \lambda_r$. $\qquad\qquad\qquad\qquad\qquad\qquad\square$

Proposition 5.2 (DIRK2 scheme). *The second order diagonally implicit Runge-Kutta scheme (DIRK2) with $\alpha = 1 \pm \frac{\sqrt{2}}{2}$, that states:*

$$\mathcal{U}^* = \mathcal{U}^n + \Delta t F(\mathcal{U}^*),$$
$$\mathcal{U}^{n+1} = \frac{2\alpha - 1}{\alpha}\mathcal{U}^n + \frac{1-\alpha}{\alpha}\mathcal{U}^* + \alpha \Delta t F(\mathcal{U}^{n+1}), \quad (5.31)$$

has amplification function given by:

$$G_{DIRK2}(\alpha) = \frac{\frac{2\alpha-1}{\alpha}\left((1-\alpha P)^3 + (C_f\kappa\Delta x)^2(1-\alpha P)\right) + \frac{1-\alpha}{\alpha}(1-\alpha P)^2 - (\alpha C_f\kappa\Delta x)^2}{\left((1-\alpha P)^2 + (C_f\kappa\Delta x)^2\right)^2}$$

$$- i\frac{(2\alpha-1)C_f\kappa\Delta x\left((1-\alpha P)^2 + (C_f\kappa\Delta x)^2\right) + 2(1-\alpha)C_f\kappa\Delta x(1-\alpha P)}{\left((1-\alpha P)^2 + (C_f\kappa\Delta x)^2\right)^2}.$$

Proof. Apply the slope function given by Expression (5.26) in the ODE (5.31) and manipulate the result algebraically to yield :

$$G_{DIRK2}(\alpha) = \frac{\frac{2\alpha-1}{\alpha}\left((1-\alpha P)^3 + (C_f\kappa\Delta x)^2(1-\alpha P)\right) + \frac{1-\alpha}{\alpha}(1-\alpha P)^2 - (\alpha C_f\kappa\Delta x)^2}{\left((1-\alpha P)^2 + (C_f\kappa\Delta x)^2\right)^2}$$

$$(5.32)$$

$$- i\frac{(2\alpha-1)C_f\kappa\Delta x\left((1-\alpha P)^2 + (C_f\kappa\Delta x)^2\right) + 2(1-\alpha)C_f\kappa\Delta x(1-\alpha P)}{\left((1-\alpha P)^2 + (C_f\kappa\Delta x)^2\right)^2}.$$

$$\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$$

Furthermore, as shown earlier in Chapter 4, IMEX schemes admit larger time step sizes and are less expensive for simulating advection-diffusion reaction problems, thus, IMEX schemes that are composed of Backward Euler and Forward Euler schemes will be

analysed to establish their dispersion preserving abilities. Let $H(\mathcal{U})$ denote a non-stiff function and $G(\mathcal{U})$ denote a stiff function such that the following ODE holds:

$$\frac{\partial \mathcal{U}}{\partial t} = G(\mathcal{U}) + H(\mathcal{U}). \tag{5.33}$$

The non-stiff function is the discretized diffusion and reaction terms while the discretized advection term is the nonstiff function in the 1D reactive transport problem (5.1).

Proposition 5.3. *Given the 1D reactive transport problem* (5.1), *one IMEX combination of Euler schemes (named IMEX-FBE1) given by:*

$$\mathcal{U}^{n+1} = \mathcal{U}^n + \Delta t \Big(G(\mathcal{U}^{n+1}) + H(\mathcal{U}^n) \Big) \tag{5.34}$$

has amplification function given by:

$$G_{FBE1} = \frac{1}{1-P} - i\frac{C_f \kappa \Delta x}{1-P}.$$

Proof. Using Fourier-spectral method for the space derivatives yields the following functions:

$$H(\mathcal{U}) = -iu\kappa\mathcal{U}, \quad G(\mathcal{U}) = (-\kappa^2\Gamma + K_r)\mathcal{U}. \tag{5.35}$$

Using expression (5.35) in expression (5.34), the IMEX-FBE1-Fourier spectral scheme for the 1D linear problem states that:

$$(1 - P)\mathcal{U}^{n+1} = (1 - iC_f\kappa\Delta x)\mathcal{U}^n. \tag{5.36}$$

Thus, the amplification factor for IMEX-FBE1 is given by:

$$G_{FBE1} = \frac{1}{1-P} - i\frac{C_f \kappa \Delta x}{1-P}. \tag{5.37}$$

$\square$

Proposition 5.4. *Given the 1D reactive transport problem* (5.1), *another IMEX combination of Euler schemes (named IMEX-FBE2) that states:*

$$\mathcal{U}^* = \mathcal{U}^n + \Delta t \Big(G(\mathcal{U}^*) + H(\mathcal{U}^n) \Big)$$
$$\mathcal{U}^{n+1} = \mathcal{U}^n + \Delta t \Big(G(\mathcal{U}^*) + H(\mathcal{U}^*) \Big) \tag{5.38}$$

has amplification function given by:

$$G_{FBE2} = \frac{1 - (C_f \kappa \Delta x)^2}{1 - P} - i \frac{C_f \kappa \Delta x (1 + P)}{1 - P}.$$

Proof. The slope function for the 1D reactive transport problem is given by expression (5.35), thus, substituting into IMEX-FBE2 scheme (5.38) yields the IMEX-FBE2-Fourier spectral scheme:

$$(1 - P)\mathcal{U}^* = (1 - iC_f \kappa \Delta x)\mathcal{U}^n$$
$$\mathcal{U}^{n+1} = \mathcal{U}^n + (P - iC_f \kappa \Delta x)\mathcal{U}^* \tag{5.39}$$

It follows from (5.39) that, the amplification function for IMEX-FBE2 is given by:

$$G_{FBE2} = \frac{1 - (C_f \kappa \Delta x)^2}{1 - P} - i \frac{C_f \kappa \Delta x (1 + P)}{1 - P}. \tag{5.40}$$

$\square$

5.4 Analysis of the time integrators

Given the amplification functions for the Theta, DIRK2 and IMEX schemes, it is a trivial task to compute the normalized dispersion relations (5.23) - (5.25), for each of the schemes and for each of the transport conditions (pure diffusive, diffusive-destructive and diffusive productive transport cases). Among the theta schemes, only results for Crank-Nicholson are presented here. However, results for the first order theta schemes indicate that, backward Euler (also known as implicit Euler) scheme has better dispersion-preserving abilities than the forward Euler (also known as explicit Euler) scheme.

Contours of all the normalized dispersion relations for all the schemes (and for all the three transport cases of the 1D linear advection-diffusion reaction problem) have been displayed in Figures 5.3-5.8. The results will be analysed according to schemes and transport cases.

However, before the analysis, a preamble is introduced to simplify the presentation of results. When a normalized relation assumes a value of one, the corresponding integrator is consistent with regards to that relation, while negative values indicate instability. To

proceed with the analysis, a closed interval containing the range of values (for a particular normalized relation) is defined. If the interval contain one then the scheme is consistent, and if the interval contains negative values then the scheme is conditionally stable. The region of instability is the area between the minimum value and zero.

5.4.1 Pure diffusive transport

Pure diffusive transport problems are those that do not involve reactions or any form of sources or sinks. Thus, dispersion properties of the numerical schemes are analysed here, for such transport problems. Contours of all the normalized relations for the schemes (applied to the pure diffusive problem) have been displayed in Figures 5.3 and 5.4.

The normalized amplification factor for Crank-Nicholson integrator decreased in the closed interval $[0.31, 1]$, across CFL and wave numbers (see Figure 5.3a), while the normalized amplification factor for DIRK2 increased in the closed interval $[0.21, 4.62]$ across CFL and wave numbers (see Figure 5.3b). Further, the normalized amplification factor for IMEX-FBE1 integrator decreased in the closed interval $[1, 9.98]$, across CFL and wave numbers (see Figure 5.4a), while the normalized amplification factor for IMEX-FBE2 increased in the closed interval $[0.87, 97.87]$ across CFL and wave numbers (see Figure 5.4b). Since all the intervals contain one, then all the schemes are consistent with regards to amplification factor. However, Crank-Nicholson has the most accurate amplification factor, followed by DIRK2 scheme (since DIRK2 and Crank-Nicholson are second order schemes, the observation here is not surprising). Among the first order IMEX schemes, the amplification factor for IMEX-FBE1 is more accurate than the amplification factor of IMEX-FBE2 scheme. Since all the intervals do not contain negative numbers, then all the schemes are stable with regards to amplification factor.

The normalized phase speed for Crank-Nicholson integrator decreased in the closed interval $[0.04, 1]$, across CFL and wave numbers (see Figure 5.3c), while the normalized phase speed for DIRK2 decreased in the closed interval $[-1.91, 2]$ across CFL and wave numbers (see Figure 5.3d). Further, the normalized phase speed for IMEX-FBE1 integrator decreased in the closed interval $[0.15, 1]$, across CFL and wave numbers (see Figure 5.4c), while the normalized phase speed for IMEX-FBE2 decreased in the closed interval

$[-1.57, 1.57]$ across CFL and wave numbers (see Figure 5.4d). Since all the intervals contain one, then all the schemes are consistent with regards to phase speed. However, IMEX-FBE1 has the most accurate phase speed, followed by Crank-Nicholson scheme. IMEX-FBE1 and Crank-Nicholson schemes are unconditionally stable with regards to phase speed (since their intervals do not contain negative values), while IMEX-FBE2 and DIRK2 are conditionally stable.

The normalized group velocity for Crank-Nicholson integrator decreased in the closed interval $[0.1, 1]$, across CFL and wave numbers (see Figure 5.3e), while the normalized group velocity for DIRK2 decreased in the closed interval $[-2.53, 4]$ across CFL and wave numbers (see Figure 5.3f). Further, the normalized group velocity for IMEX-FBE1 integrator decreased in the closed interval $[0.01, 1]$, across CFL and wave numbers (see Figure 5.4e), while the normalized group velocity for IMEX-FBE2 increased in the closed interval $[0, 1.33]$ across CFL and wave numbers (see Figure 5.4f). Since all the intervals contain one, then all the schemes are consistent with regards to group velocity. However, Crank-Nicholson has the most accurate group velocity, followed by IMEX-FBE1 scheme. All except DIRK2 scheme, are unconditionally stable with regards to group velocity, since their intervals do not contain negative values.

5.4.2   Diffusive-productive transport

Diffusive-productive transport problems are those that chemical reactions produce concentration sources. Thus, dispersion properties of the numerical schemes are analysed here, for such transport problems. Contours of all the normalized relations for the schemes (applied to the pure diffusive-productive problem) have been displayed in Figures 5.5 and 5.6.

The normalized amplification factor for Crank-Nicholson integrator decreased in the closed interval $[0.25, 1]$, across CFL and wave numbers (see Figure 5.5a), while the normalized amplification factor for DIRK2 increased in the closed interval $[0.11, 5.89]$ across CFL and wave numbers (see Figure 5.5b). Further, the normalized amplification factor for IMEX-FBE1 integrator decreased in the closed interval $[1, 9.96]$, across CFL and wave numbers (see Figure 5.6a), while the normalized amplification factor for

IMEX-FBE2 increased in the closed interval $[0.75, 97.18]$ across CFL and wave numbers (see Figure 5.6b). Since all the intervals contain one, all the schemes are consistent with regards to amplification factor. However, Crank-Nicholson has the most accurate amplification factor, followed by DIRK2 scheme. Among the first order IMEX schemes, the amplification factor for IMEX-FBE1 is more accurate than that for IMEX-FBE2 scheme. Since all the intervals do not contain negative numbers, all the schemes are stable with regards to amplification factor, in diffusive-productive transport problems.

The normalized phase speed for Crank-Nicholson integrator decreased in the closed interval $[0.04, 1]$ across CFL and wave numbers (see Figure 5.5c), while the normalized phase speed for DIRK2 decreased in the closed interval $[-1.91, 2]$ across CFL and wave numbers (see Figure 5.5d). Further, the normalized phase speed for IMEX-FBE1 integrator decreased in the closed interval $[0.15, 1]$, across CFL and wave numbers (see Figure 5.6c), while the normalized phase speed for IMEX-FBE2 decreased in the closed interval $[-1.57, 1.57]$ across CFL and wave numbers (see Figure 5.6d). Since all the intervals contain one, then all the schemes are consistent with regards to phase speed. However, IMEX-FBE1 has the most accurate phase speed, followed by Crank-Nicholson scheme. IMEX-FBE1 and Crank-Nicholson schemes are unconditionally stable with regards to phase speed (since their intervals do not contain negative values), while IMEX-FBE2 and DIRK2 are conditionally stable, if diffusive-productive problems are considered.

The normalized group velocity for Crank-Nicholson integrator decreased in the closed interval $[0.1, 1.01]$, across CFL and wave numbers (see Figure 5.5e), while the normalized group velocity for DIRK2 decreased in the closed interval $[-10.49, 12.34]$ across CFL and wave numbers (see Figure 5.5f). Further, the normalized group velocity for IMEX-FBE1 integrator decreased in the closed interval $[0.01, 1]$, across CFL and wave numbers (see Figure 5.6e), while the normalized group velocity for IMEX-FBE2 increased in the closed interval $[-0.03, 2.49]$ across CFL and wave numbers (see Figure 5.6f). Since all the intervals contain one, then all the schemes are consistent with regards to group velocity. However, Crank-Nicholson has the most accurate group velocity, followed by IMEX-FBE1 scheme. DIRK2 and IMEX-FBE2 are conditionally stable with regards to group velocity since their intervals contain negative values, while Crank-Nicholson and IMEX-FBE1 are unconditionally stable.

5.4.3 Diffusive-destructive case

Diffusive-destructive transport problems are those that chemical reactions produce concentration sinks. Thus, dispersion properties of the numerical schemes are analysed here, for such transport problems. Contours of all the normalized relations for the schemes (applied to the pure diffusive-destructive problem) have been displayed in Figures 5.7 and 5.8.

The normalized amplification factor for Crank-Nicholson integrator decreased in the closed interval $[0.38, 1]$ across CFL and wave numbers (see Figure 5.7a), while the normalized amplification factor for DIRK2 increased in the closed interval $[0.31, 4.23]$ across CFL and wave numbers (see Figure 5.7b). Further, the normalized amplification factor for IMEX-FBE1 integrator decreased in the closed interval $[1.02, 10.35]$, across CFL and wave numbers (see Figure 5.8a), while the normalized amplification factor for IMEX-FBE2 increased in the closed interval $[0.98, 101.46]$ across CFL and wave numbers (see Figure 5.8b). Since all the intervals contain one, then all the schemes are consistent with regards to amplification factor. However, Crank-Nicholson has the most accurate amplification factor followed by DIRK2 scheme. Among the first order IMEX schemes, the amplification factor for IMEX-FBE1 is more accurate than the amplification factor of IMEX-FBE2 scheme. Since all the intervals do not contain negative numbers, then all the schemes are stable with regards to amplification factor, if diffusive-destructive transport problems are considered.

The normalized phase speed for Crank-Nicholson integrator decreased in the closed interval $[0.04, 1.03]$ across CFL and wave numbers (see Figure 5.7c), while the normalized phase speed for DIRK2 decreased in the closed interval $[-0.95, 1.26]$ across CFL and wave numbers (see Figure 5.7d). Further, the normalized phase speed for IMEX-FBE1 integrator decreased in the closed interval $[0.15, 1]$, across CFL and wave numbers (see Figure 5.8c), while the normalized phase speed for IMEX-FBE2 decreased in the closed interval $[-1.57, 1.57]$ across CFL and wave numbers (see Figure 5.8d). Since all the intervals contain one, then all the schemes are consistent with regards to phase speed. However, IMEX-FBE1 has the most accurate phase speed, followed by Crank-Nicholson scheme. IMEX-FBE1 and Crank-Nicholson schemes are unconditionally stable with regards to

phase speed (since their intervals do not contain negative values), while IMEX-FBE2 and DIRK2 are conditionally stable, if diffusive-destructive problems are considered.

The normalized group velocity for Crank-Nicholson integrator decreased in the closed interval $[0.1, 1.01]$, across CFL and wave numbers (see Figure 5.7e), while the normalized group velocity for DIRK2 decreased in the closed interval $[-1, 2.51]$ across CFL and wave numbers (see Figure 5.7f). Further, the normalized group velocity for IMEX-FBE1 integrator decreased in the closed interval $[0.01, 1]$, across CFL and wave numbers (see Figure 5.8e), while the normalized group velocity for IMEX-FBE2 increased in the closed interval $[-0.01, 0.61]$ across CFL and wave numbers (see Figure 5.8f). Since all the intervals contain one, then all the schemes are consistent with regards to group velocity. However, Crank-Nicholson has the most accurate group velocity, followed by IMEX-FBE1 scheme. All except DIRK2 scheme, are unconditionally stable with regards to group velocity, if diffusive-destructive problems are considered.

5.5 Numerical experiments

The exact solution for the ODE (5.3) can be expressed in terms of Peclet and wave numbers as follows:

$$\mathcal{U}(\kappa, t^n) = \mathcal{U}_0(\kappa)e^{-n(iC_f\kappa\Delta x + P_e(\kappa\Delta x)^2 - \lambda_r)}, \tag{5.41}$$

where the nth time point is given by $t^n = n\Delta t$. The solutions of the recurrence relations (5.34) and (5.38) can be written as:

$$\mathcal{U}^n = \mathcal{U}_0(\kappa)G_{apx}^n, \tag{5.42}$$

where G_{apx} is given by G_{FBE1} in (5.37) for IMEX-FBE1 or G_{FBE2} in (5.40) for IMEX-FBE2.

Figures 5.9 and 5.10 compare (in a spectral plane) the analytical solution and numerical solutions for the 1D linear advection-diffusion reaction problem. It can be observed from 5.9 that the Crank-Nicholson solution is more accurate than DIRK2 solution, for all the three transport cases. Similarly, it can be observed in Figure 5.10 that IMEX-FBE1

118

is more accurate than IMEX-FBE2, in all the three transport cases. *Although numerical schemes do not have exact dispersion-dissipation properties, the observation here shows that Crank-Nicholson and IMEX-FBE1 schemes are more suitable for the reactive transport model.* These observations validate the dispersion analyses presented above.

5.5.1 Accuracy tests with 1D system

Numerical simulation of a 1D transport problem that consists of two PDEs with stiff source terms is discussed here. Since, numerical results for the theta and DIRK2 schemes were discussed extensively in Chapter 2, only results for the IMEX schemes are presented here. The test example states that find $U(x, t)$ and $V(x, t)$ such that:

$$\frac{\partial U}{\partial t} + \frac{\partial uU}{\partial x} = \Gamma \frac{\partial^2 U}{\partial x^2} - K_u U + V, \qquad 0 < x < 2\pi, \tag{5.43}$$

$$\frac{\partial V}{\partial t} + \frac{\partial uV}{\partial x} = \Gamma \frac{\partial^2 V}{\partial x^2} - K_v V, \qquad 0 < x < 2\pi. \tag{5.44}$$

Using periodic boundary conditions, the exact solution for the system (5.43)-(5.44) is:

$$U(x, t) = \left(e^{-(K_u + \Gamma)t} + e^{-(K_v + \Gamma)t} \right) \cos(x - ut), \tag{5.45}$$

$$V(x, t) = (K_u - K_v) e^{-(K_v + \Gamma)t} \cos(x - ut). \tag{5.46}$$

The input data are as follows:

$$\Gamma = 0.01, \quad u = 0.01, \quad K_v = 1, \quad K_u = 100, \tag{5.47}$$

and the initial data are as follows:

$$U(x, 0) = 2\cos(x), \tag{5.48}$$

$$V(x, 0) = (K_u - K_v)\cos(x). \tag{5.49}$$

Figure 5.11 shows the solution and error profiles for IMEX-FBE1 and IMEX-FBE2 applied to the stiff system. It is clear that both numerical schemes are accurate and

convergent for the stiff system. Table 5.1 shows that both schemes have the same order of convergence. Two finite difference schemes (that are order one and have been discussed in [157]) have also been applied to solve this system. The first finite difference scheme is derived by applying forward-differencing to the time derivatives and central differencing to all spatial derivatives (the resulting scheme is named Euler-CD2-CD2) and the second finite difference scheme is obtained by applying forward differencing to the time derivatives, forward differencing to advection terms and central differencing to diffusion terms (the resulting scheme is named Euler-UP1-CD2). Figure 5.12, shows that the IMEX schemes are superior in accuracy to both finite difference schemes. Although, the dispersion analysis have been performed using a single transport equation, the results in this test problem shows that the results of the analysis are not limited to single transport equations.

5.5.2 Accuracy tests with 1D nonlinear problem

Numerical simulation of a 1D nonlinear reactive transport equation (where all the processes including diffusion and reaction are nonlinear) is discussed here. Again, only results for the IMEX schemes are presented here. The nonlinear transport problem states that, find $U(x,t)$ such that [73]:

$$\frac{\partial U}{\partial t} + \frac{\partial U}{\partial x} = \frac{\partial}{\partial x} U \frac{\partial U}{\partial x} - 1 - 0.25 \cos(2(x-t)) + U, \qquad 0 < x < 2\pi. \tag{5.50}$$

With initial condition $U(x,0) = 1 + 0.5\sin(x)$ and periodic boundary conditions, the analytical solution is:

$$U(x,t) = 1 + 0.5\sin(x-t). \tag{5.51}$$

Some computed solutions for this 1D fully nonlinear problem are displayed in Figure 5.13. The results show that the IMEX schemes (especially IMEX-FBE1) are accurate and have better stiffness resolution capabilities than the finite difference schemes. Results for this test problem indicate that the observations in the dispersion analysis are not limited to only linear transport problems.

5.5.3 Accuracy tests with 2D problems

Numerical simulation of a 2D reactive transport problem is discussed here. The goal is to show that the observations of the dispersion analysis also apply in 2D transport problems. The first 2D test example is a 2D version of system (5.43)-(5.44) given by [73]:

$$\frac{\partial U}{\partial t} + \frac{u}{2}\Big(\frac{\partial U}{\partial x} + \frac{\partial U}{\partial y}\Big) = \frac{\Gamma}{2}\Big(\frac{\partial^2 U}{\partial x^2} + \frac{\partial^2 U}{\partial y^2}\Big) - K_u U + V, \qquad 0 < x, y < 2\pi, \qquad (5.52)$$

$$\frac{\partial V}{\partial t} + \frac{u}{2}\Big(\frac{\partial V}{\partial x} + \frac{\partial V}{\partial y}\Big) = \frac{\Gamma}{2}\Big(\frac{\partial^2 V}{\partial x^2} + \frac{\partial^2 V}{\partial y^2}\Big) - K_v V, \qquad 0 < x, y < 2\pi. \qquad (5.53)$$

With periodic boundary conditions, the exact solution is:

$$U(x, y, t) = \Big(e^{-(K_u + \Gamma)t} + e^{-(K_v + \Gamma)t}\Big)\cos(x + y - ut), \qquad (5.54)$$

$$V(x, y, t) = (K_u - K_v)e^{-(K_v + \Gamma)t}\cos(x + y - ut). \qquad (5.55)$$

The second 2D problem is fully nonlinear and it is given by [73]:

$$\frac{\partial U}{\partial t} + 0.5\frac{\partial U^2}{\partial x} + 0.5\frac{\partial U^2}{\partial y} = \frac{\partial}{\partial x}U\frac{\partial U}{\partial x} + \frac{\partial}{\partial y}U\frac{\partial U}{\partial y} - U^2$$
$$+ 1.125 - 0.625\cos(2(x + y - t)) + 0.25\sin(2(x + y - t))$$
$$+ 0.5\cos(x + y - t) + 2\sin(x + y - t), \qquad 0 < x, y < 2\pi. \qquad (5.56)$$

With periodic boundaries, the analytical solution is:

$$U(x, y, t) = 1 + 0.5\sin(x + y - t). \qquad (5.57)$$

The results are displayed in Figure 5.14. Both IMEX schemes are efficient in resolving stiffness in the problem involving a system, while IMEX-FBE1 is better in the fully nonlinear case. The results here indicate that the findings in the dispersion analysis are not limited to 1D problems.

5.5.4 Application: environmental chemical engineering system

Numerical simulation of acidic mine effluents are discussed here. One method for attenuating the acidity in effluents is limestone neutralization. As discussed in earlier chapters,

the stoichiometry of the acid-calcite reaction (following the discussions in [155, 11, 126, 130]) is given by:

$$CaCO_3 + H^+ \rightleftharpoons Ca^{2+} + HCO_3^-. \tag{5.58}$$

$$CaCO_3 + H_2CO_3 \rightleftharpoons Ca^{2+} + 2HCO_3^-, \tag{5.59}$$

$$CaCO_3 + H_2O \rightleftharpoons Ca^{2+} + HCO_3^+ + OH^- \tag{5.60}$$
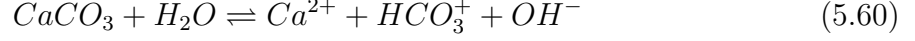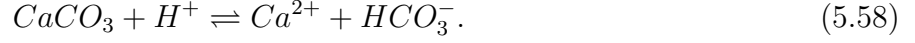
Experimental data shows that reactions (5.59) and (5.60) contribute insignificantly compared with reaction (5.58), thus, the system reduces to reaction (5.58) which contains the chemical species [2]:

$$\mathbf{C} = \left(CaCO_3, H^+, Ca^{2+}, HCO_3^- \right),$$

whose corresponding concentrations (i.e. current and initial) and source/sink vectors are denoted respectively by [2]:

$$\mathbf{U} = \left(U_1, U_2, U_3, U_4 \right), \quad \mathbf{U}_0 = \left(U_{10}, U_{20}, U_{30}, U_{40} \right)$$

and

$$\mathbf{S} = \left(0, \quad -K_f U_2 + K_b U_3 U_4, \quad K_f U_2 - K_b U_3 U_4, \quad K_f U_2 - K_b U_3 U_4 \right),$$

where K_f and K_b denote reaction constants.

Hydrogen ions ($k = 2$) are responsible for acidity of the effluent water, therefore, neutralization of the hydrogen ions implies attenuation of the pollutant. The relevant system of semi-linear transport equations to solve for hydrogen ions is given by:

$$\frac{\partial \mathbf{U}}{\partial t} + \frac{\partial u \mathbf{U}}{\partial x} = \Gamma \frac{\partial^2 \mathbf{U}}{\partial x^2} + \mathbf{S}. \tag{5.61}$$

System (5.61) has four (i.e. high) degrees of freedom thus, is computationally expensive to solve since only hydrogen concentration (i.e. one degree of freedom) is required. A reduced semi-linear equation for hydrogen ions transport (using the stoichiometric method discussed in [3] and in Chapter 3) is given by:

$$\frac{\partial U_2}{\partial t} + \frac{\partial u U_2}{\partial x} = \Gamma \frac{\partial^2 U_2}{\partial x^2} + \alpha_{H1} U_2^2 + \alpha_{H2} U_2 + \alpha_{H3}, \tag{5.62}$$

where $\alpha_{H1} = K_b$, $\alpha_{H2} = -K_f + K_b\left(2U_{20} + U_{40} + U_{30}\right)$ and $\alpha_{H3} = K_b\left(U_{10} + U_{40}\right)\left(U_{20} + U_{30}\right)$.

Input data used in the numerical experiments are:

$$L_x = 2\pi, \ T = 1, u = 0.01, \Gamma = 0.01, K_f = 0.13, \ K_b = 0.0025, K_h = 1,$$
$$U_{20} = 0.01\cos(x), \ U_{30} = 0.0001\cos(x), \ U_{40} = 0.0001\cos(x).$$

Figure 5.15 shows the solutions of IMEX-FBE1 and IMEX-FBE2 (applied to solve both large and reduced models) in all the transport cases (i.e. advection, diffusion and advection-diffusion cases). It can be observed that the solution of both schemes are decreasing with time thus, reproducing the chemistry of neutralization. Moreover, Figure 5.16, shows decreasing error profile for IMEX-FBE1 and increasing error profile for IMEX-FBE2, thus IMEX-FBE1 is more compatible with the model reduction procedure. Furthermore, CPU time given by Figure 5.17 shows that IMEX-FBE1 is less expensive to implement.

5.6 Chapter summary

Numerical schemes are approximations to exact solutions and can exhibit non-physical behaviours. Analysis of numerical schemes is therefore, necessary to eliminate inaccurate schemes. Global spectral analysis on theta, DIRK2 and IMEX schemes has been performed to ascertain accurate schemes for reactive transport problems.

Global spectral analysis of linear reactive transport equation has been discussed in Section 5.2. The spectral analysis yielded a physico-chemical dispersion relation (5.6), from which other dispersion relations (i.e. phase speed (5.7) and group velocity (5.8)) and a well-posedness constraint (5.11) have been derived.

In Section 5.3, numerical dispersion relations (including normalized amplification factor, phase speed and group velocity) have been derived. Amplification functions for theta, DIRK2 and IMEX schemes have been presented.

In Section 5.4, the normalized dispersion relations were used to analyse the numerical schemes for all the three transport problems (i.e. pure diffusive, diffusive-productive and

diffusive-destructive problems). The analyses showed that all the schemes are dispersion-preserving, however, Crank-Nicholson and IMEX-FBE1 schemes have the best properties.

In Section 5.5, numerical experiments were conducted to validate the dispersion analysis. The results showed that the implicit and IMEX schemes are more accurate for simulating stiff problems than some finite differencing schemes of the same order. The numerical experiments also showed that, observations from the dispersion analysis are not limited to single linear reactive transport problems.

Therefore, the theta, DIRK2 and IMEX schemes are dispersion preserving schemes, but Crank-Nicholson and IMEX-FBE1 schemes have better properties and are more suitable for reactive transport problems.
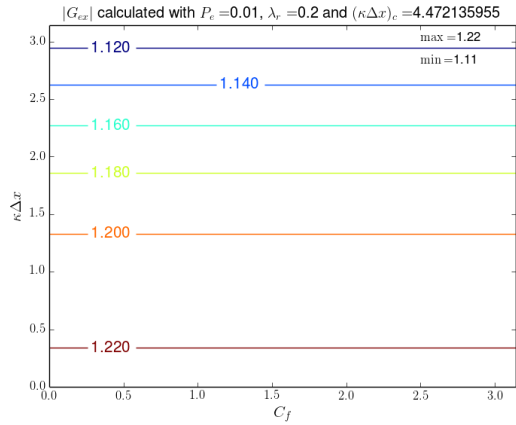
(a) $P_e = 0.01, \lambda_r = 0.2$
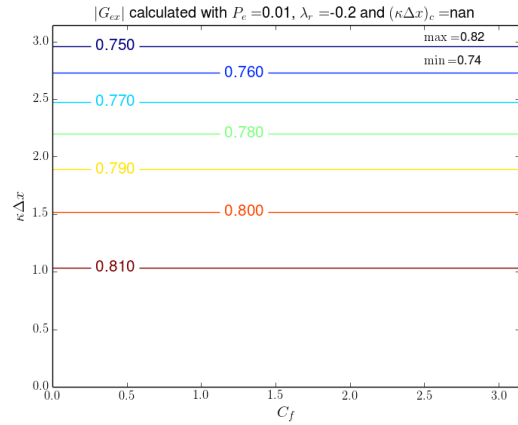
(b) $P_e = 0.01, \lambda_r = -0.2$
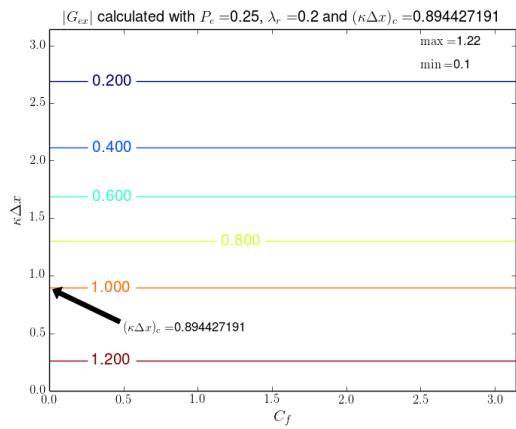
(c) $P_e = 0.25, \lambda_r = 0.2$
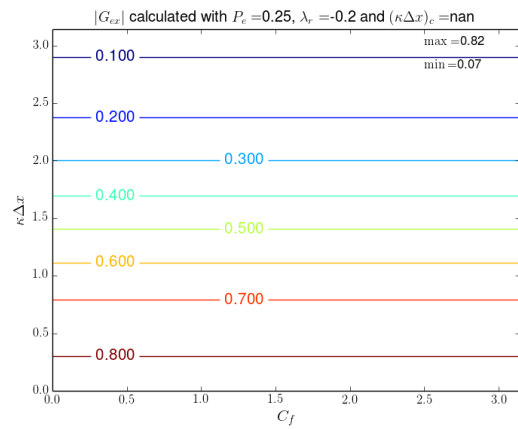
(d) $P_e = 0.25, \lambda_r = -0.2$
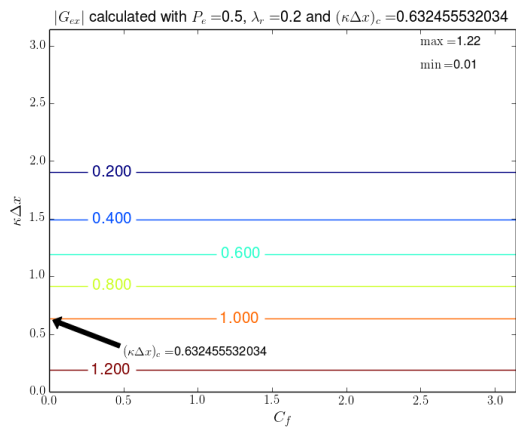
(e) $P_e = 0.5, \lambda_r = 0.2$
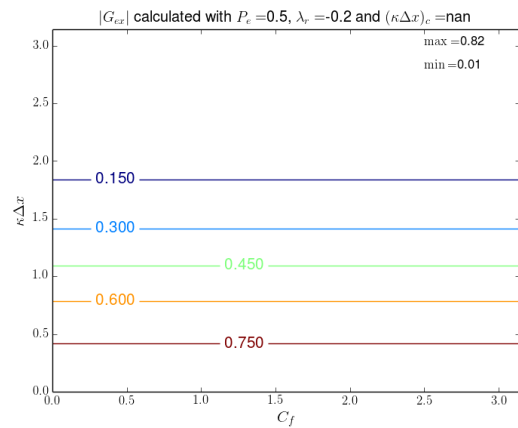
(f) $P_e = 0.5, \lambda_r = -0.2$

Figure 5.2: Contour for the physico-chemical amplification factor $|G_{ex}|$ of the analytical solution for 1D linear convection-diffusion-reaction equation. The contours display the behaviour of $|G_{ex}|$ for varying Peclet numbers in diffusive-destructive (right column) and diffusive-productive (left column) problems.
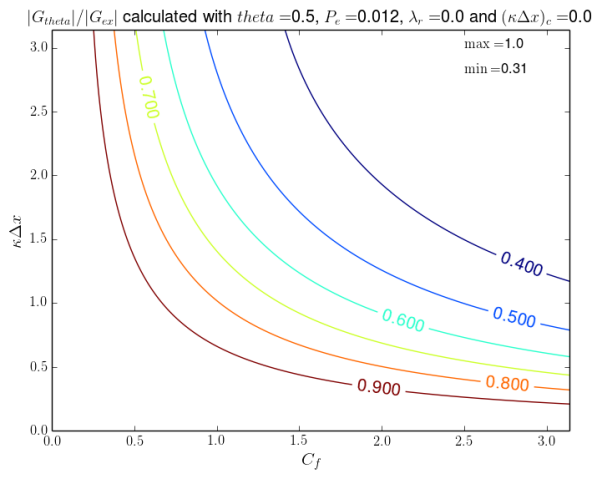
(a) Crank-Nicholson amplification
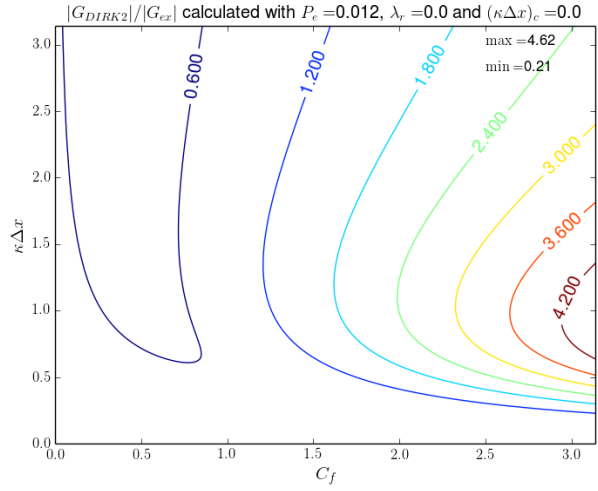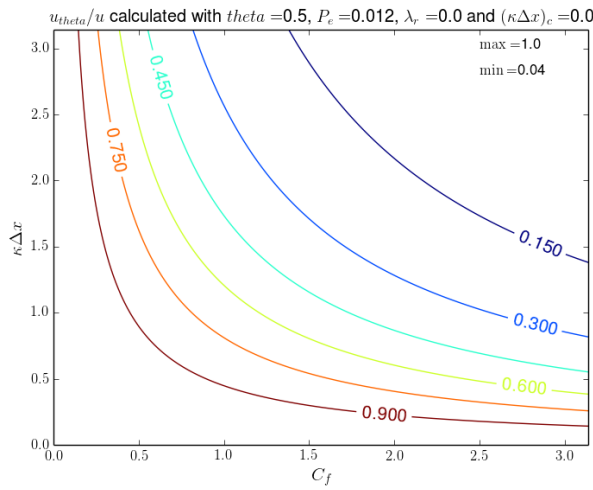
(b) DIRK2 amplification

(c) Crank-Nicholson phase speed
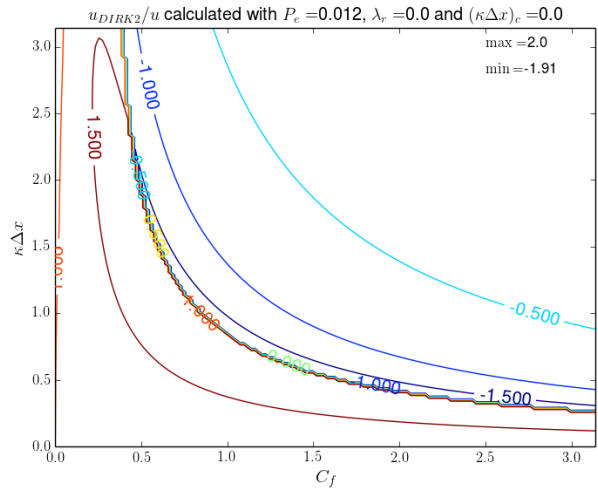
(d) DIRK2 phase speed
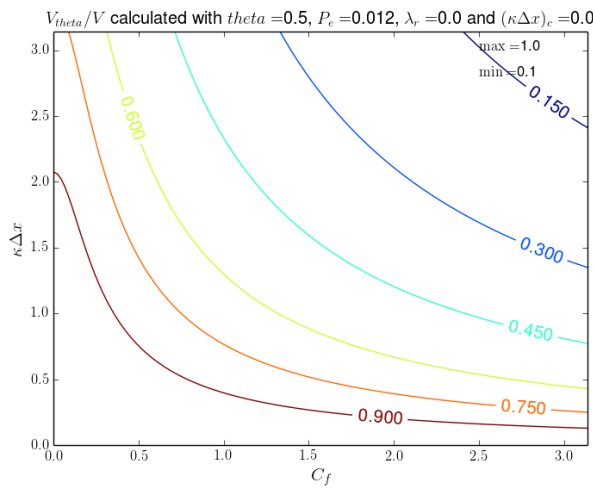
(e) Crank-Nicholson group velocity
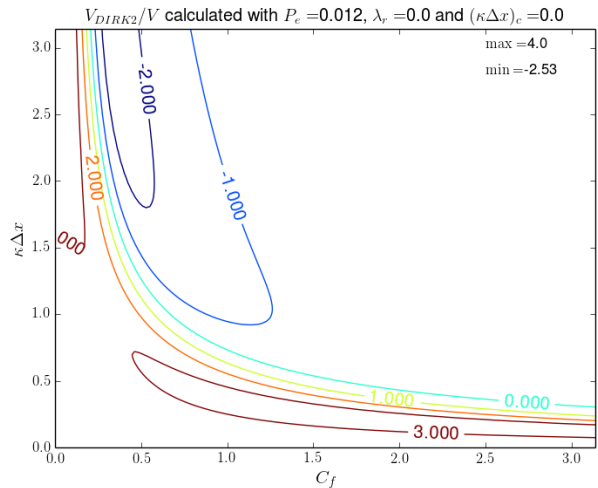
(f) DIRK2 group velocity

Figure 5.3: Contours for Crank-Nicholson and DIRK2 schemes applied in a pure diffusive 1D linear convection-diffusion-reaction problem.
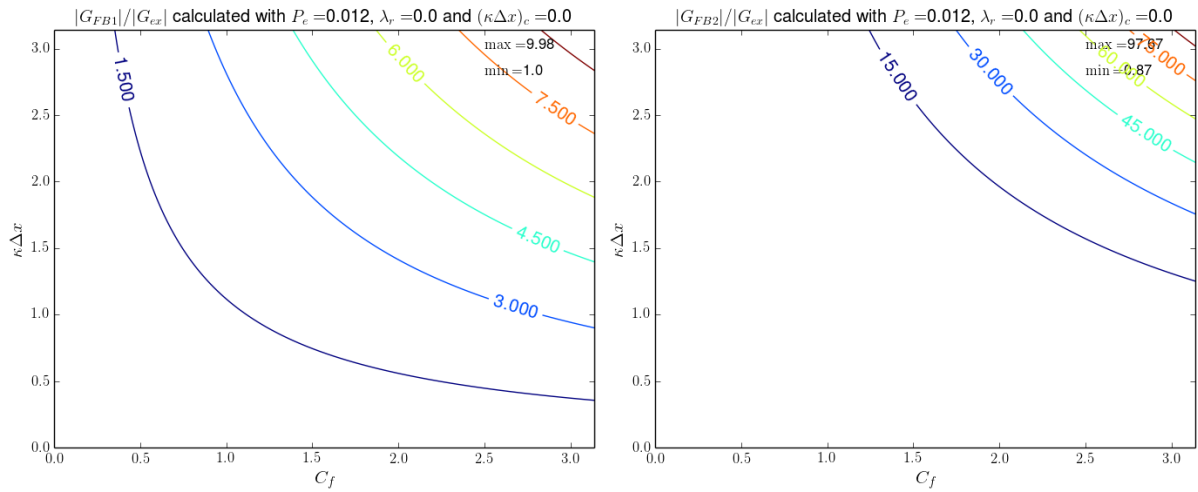
(a) IMEX-FBE1 amplification

(b) IMEX-FBE2 amplification
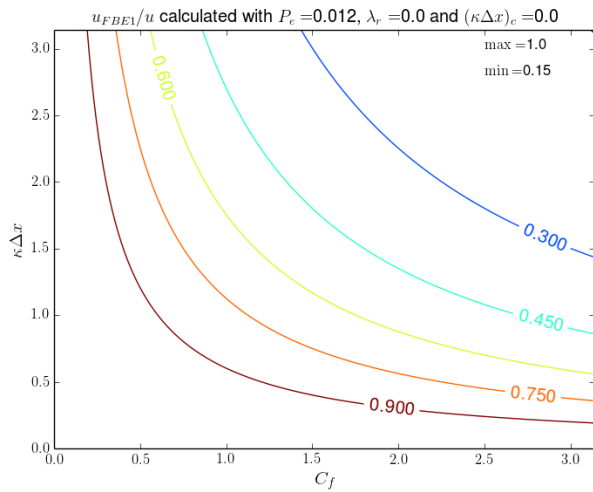
(c) IMEX-FBE1 phase speed
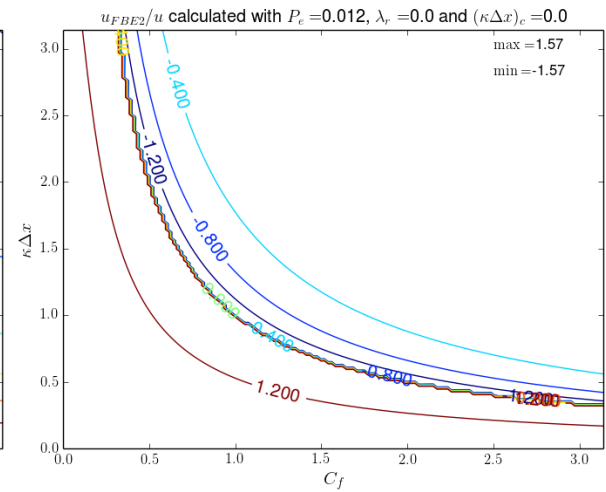
(d) IMEX-FBE2 phase speed
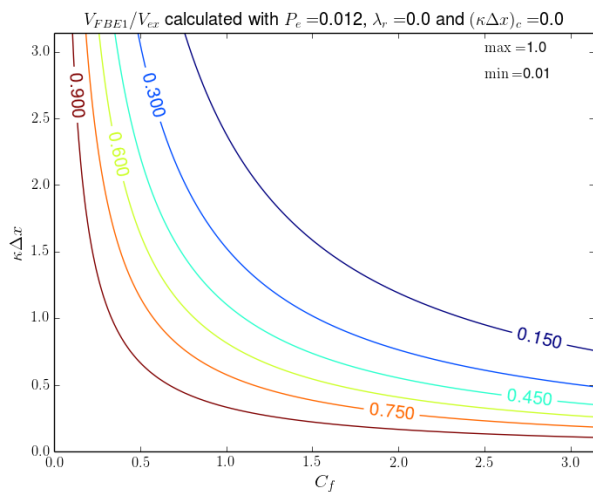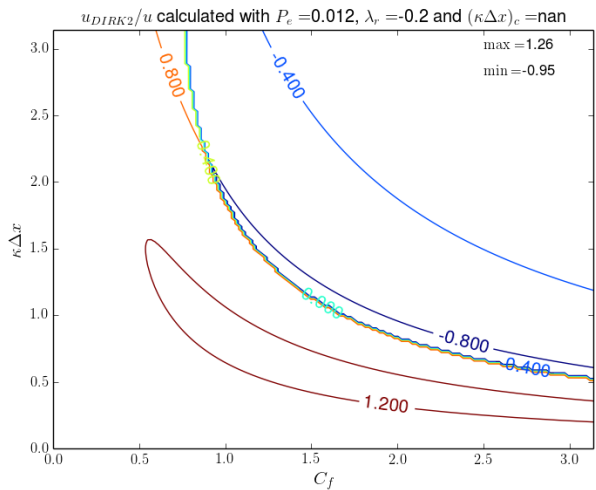
(e) IMEX-FBE1 group velocity
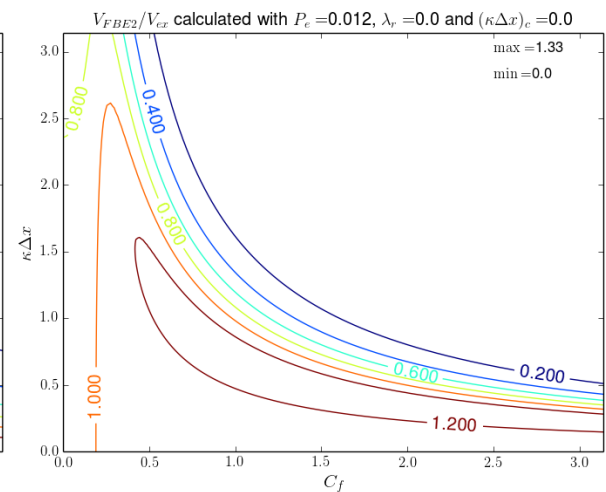
(f) IMEX-FBE2 group velocity

Figure 5.4: Contours for implicit-explicit Euler schemes applied in a pure diffusive 1D linear convection-diffusion-reaction problem.
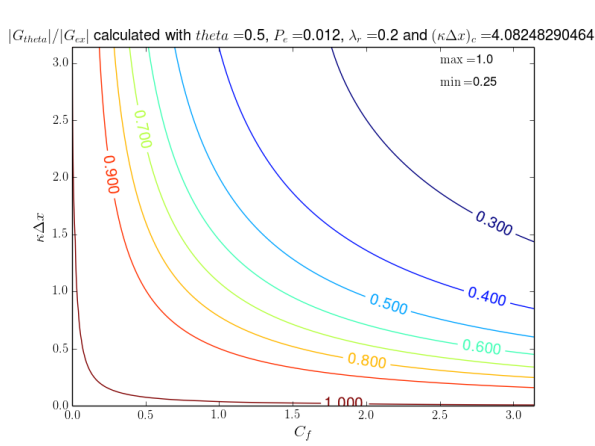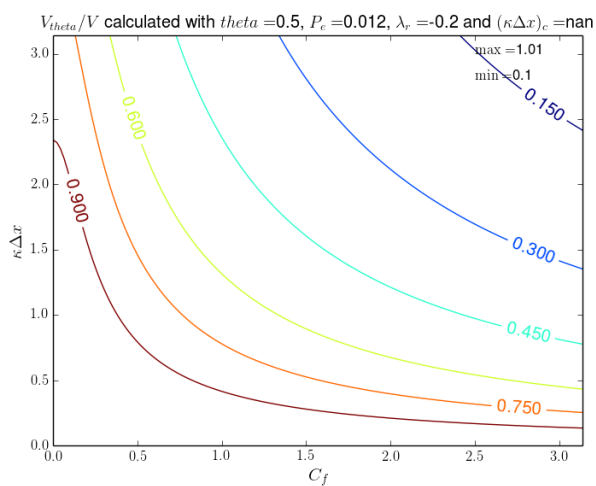
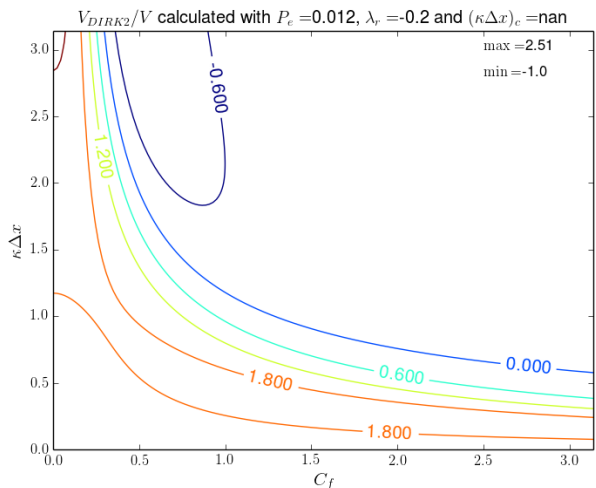(a) Crank-Nicholson amplification
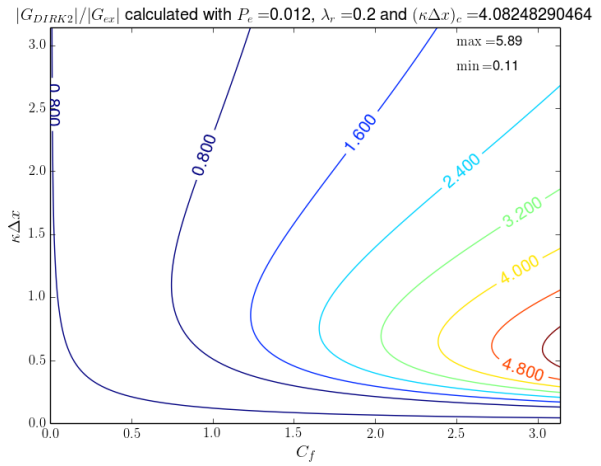
(b) DIRK2 amplification
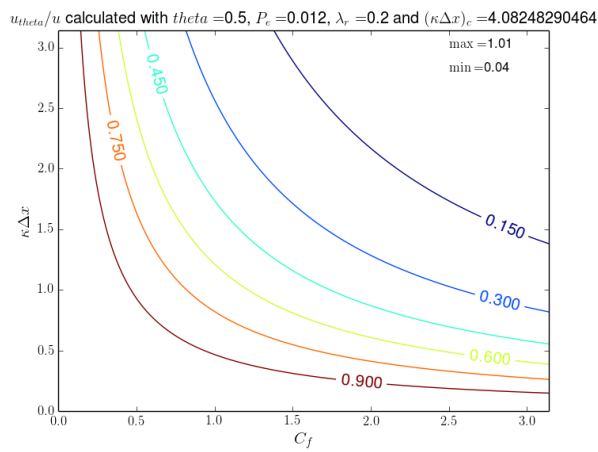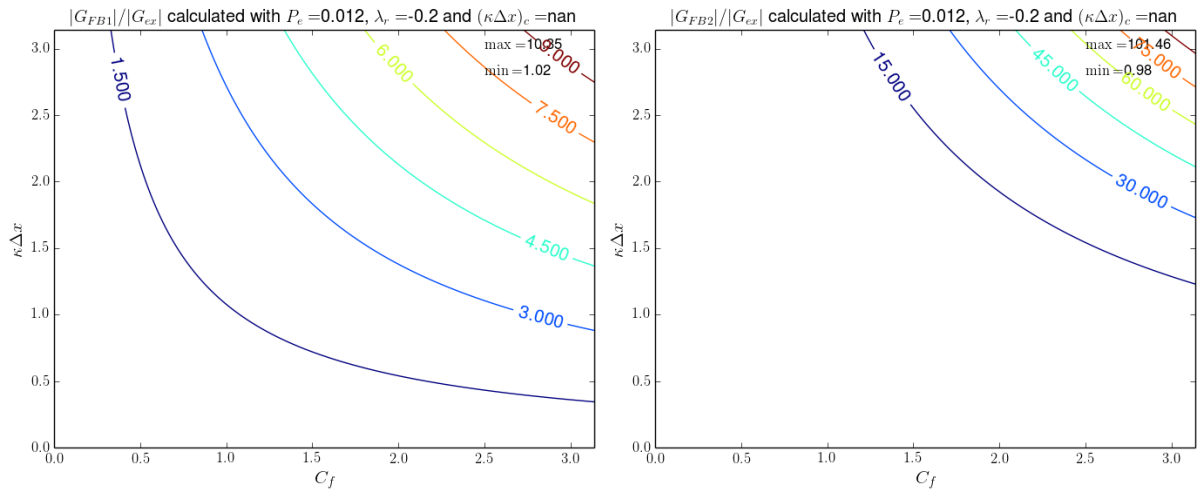
(c) Crank-Nicholson phase speed
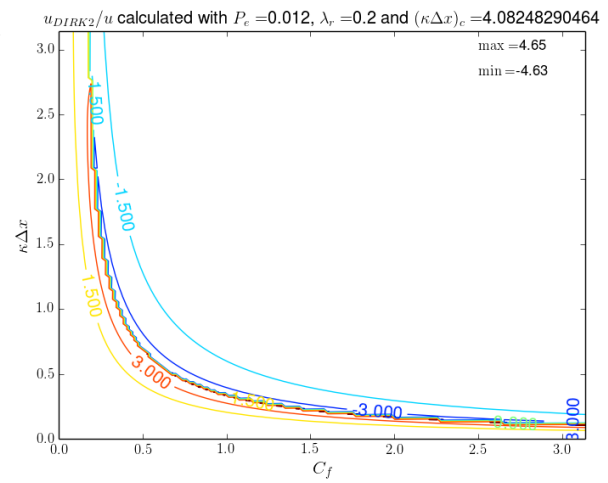
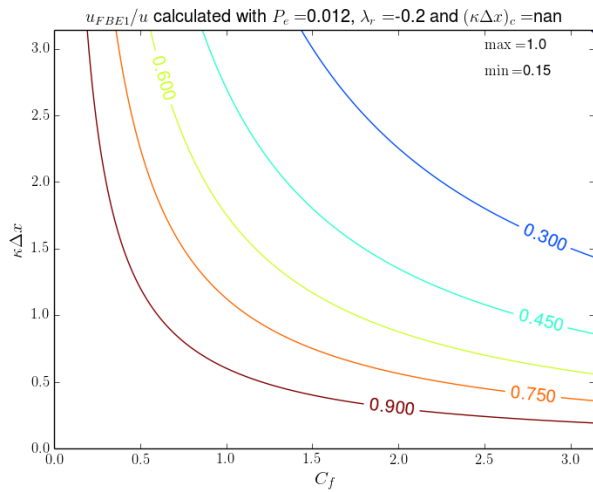(d) DIRK2 phase speed
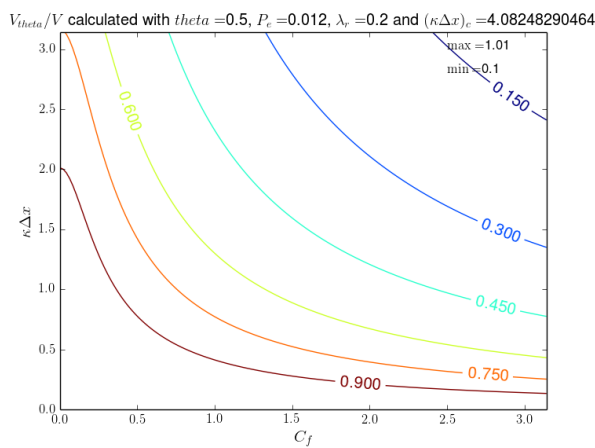
(e) Crank-Nicholson group velocity
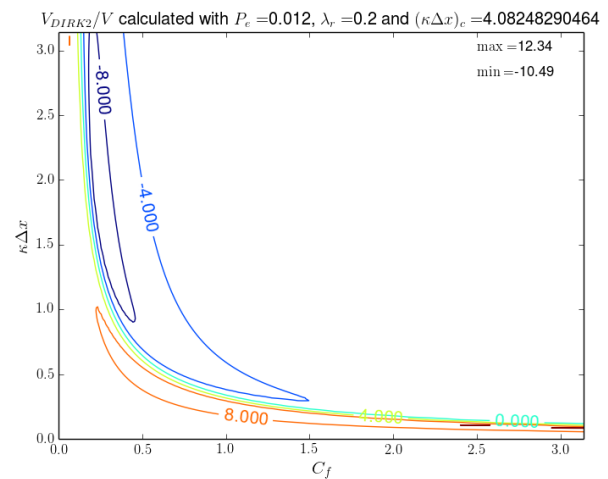
(f) DIRK2 group velocity

Figure 5.5: Contours for Crank-Nicholson and DIRK2 schemes applied in a diffusive-productive 1D linear convection-diffusion-reaction problem.
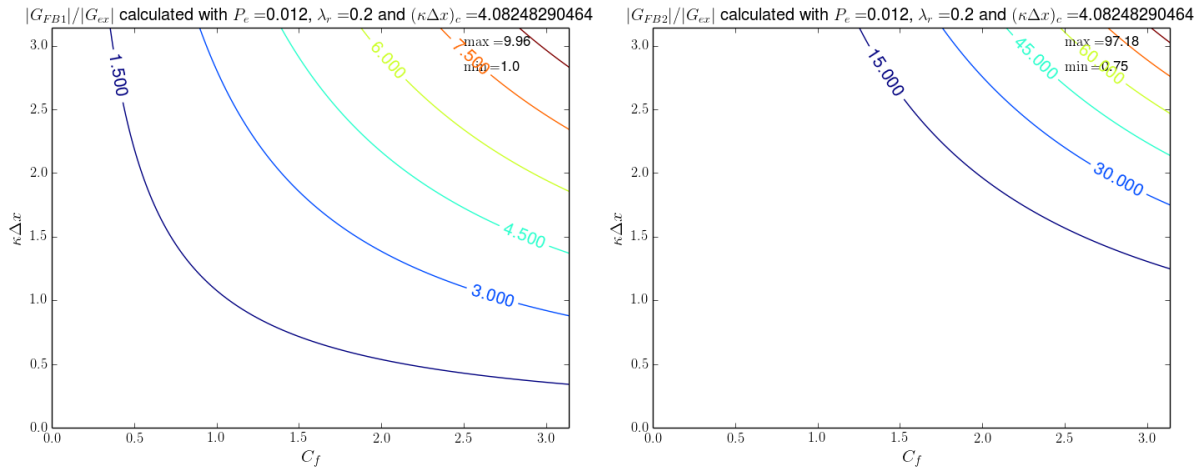
(a) IMEX-FBE1 amplification

(b) IMEX-FBE2 amplification
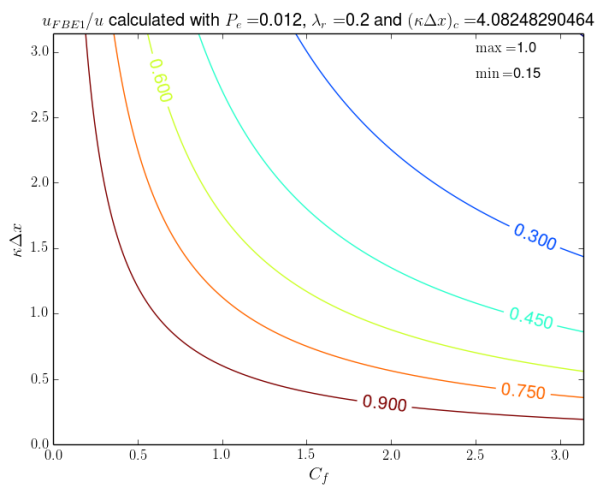
(c) IMEX-FBE1 phase speed
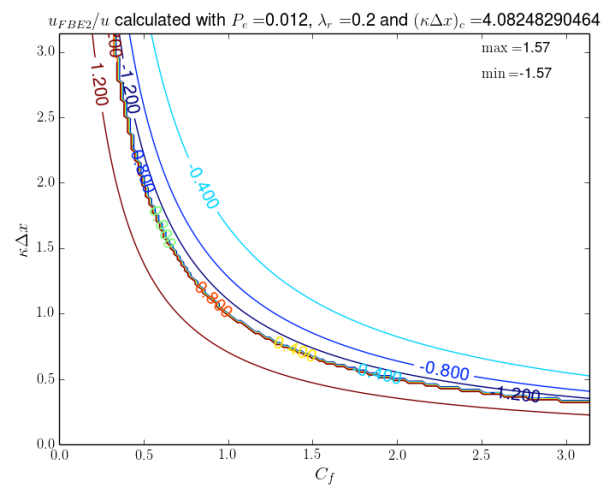
(d) IMEX-FBE2 phase speed
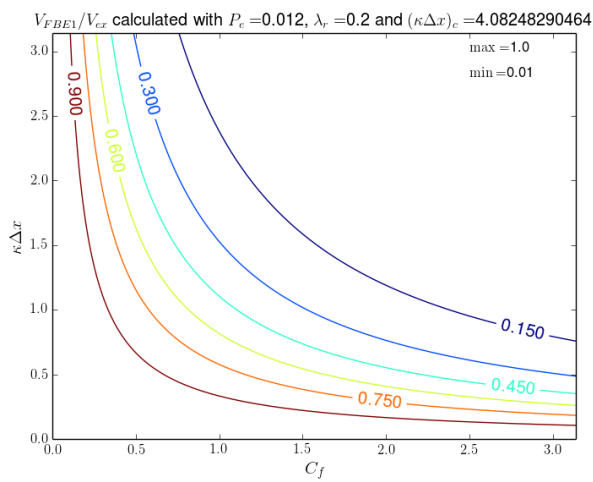
(e) IMEX-FBE1 group velocity
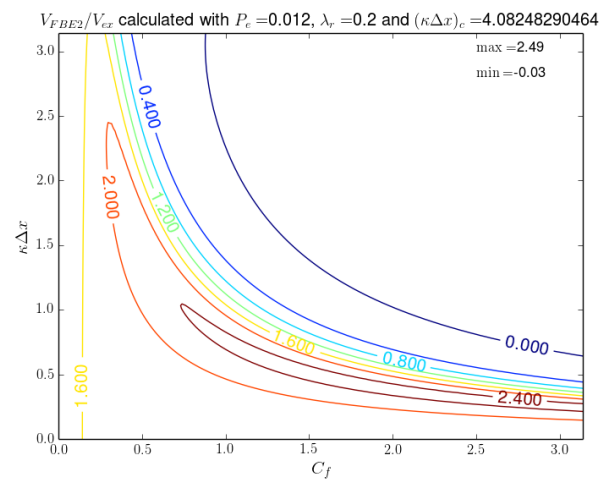
(f) IMEX-FBE2 group velocity

Figure 5.6: Contours for implicit-explicit Euler schemes applied in a diffusive-productive 1D linear convection-diffusion-reaction problem.
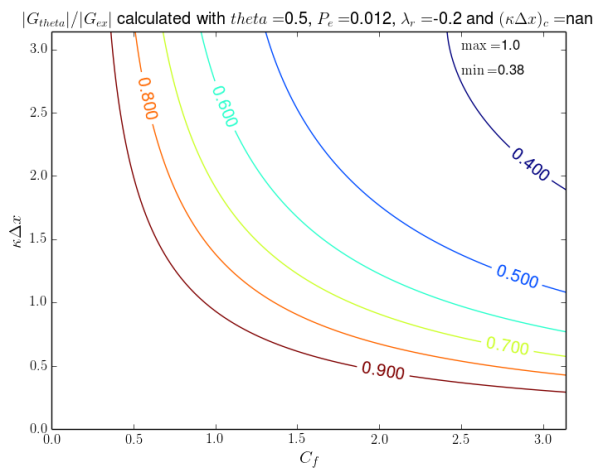
(a) Crank-Nicholson amplification
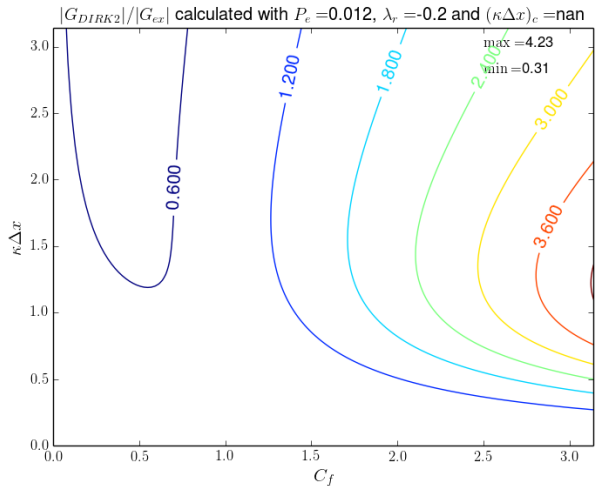
(b) DIRK2 amplification
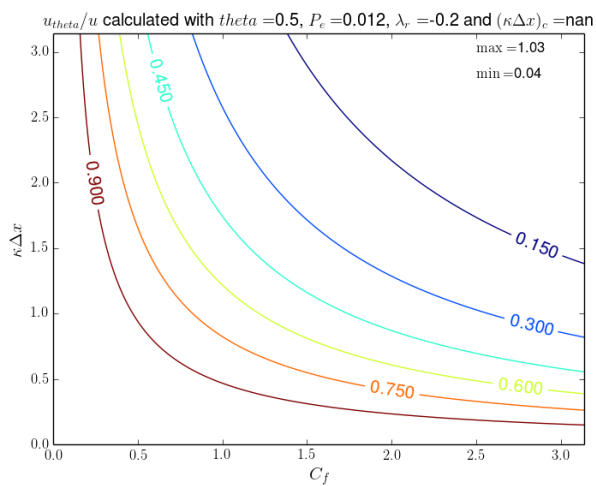
(c) Crank-Nicholson phase speed

(d) DIRK2 phase speed

(e) Crank-Nicholson group velocity

(f) DIRK2 group velocity

Figure 5.7: Contours for Crank-Nicholson and DIRK2 schemes applied in a diffusive-destructive 1D linear convection-diffusion-reaction problem.
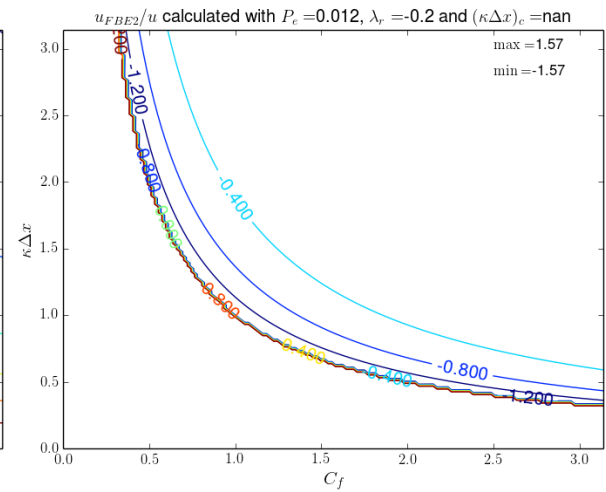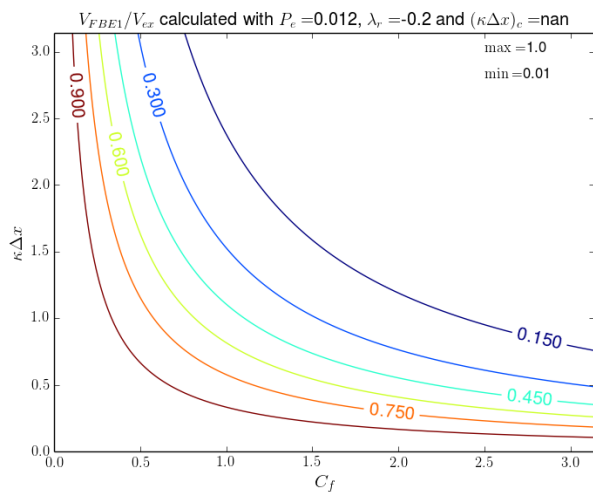
(a) IMEX-FBE1 amplification
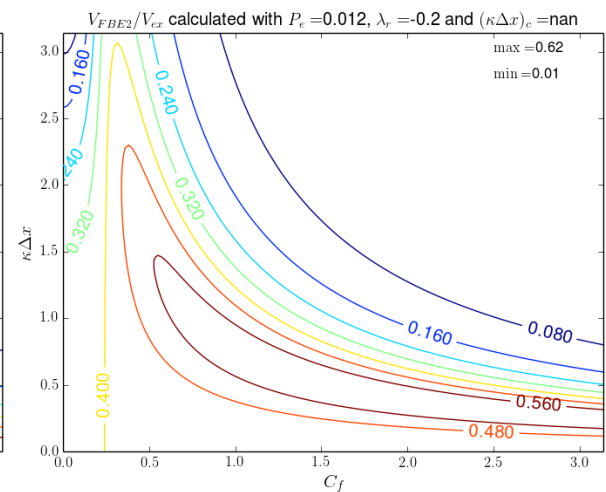
(b) IMEX-FBE2 amplification

(c) IMEX-FBE1 phase speed
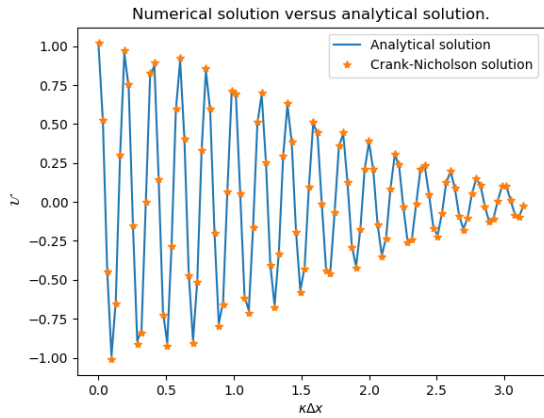
(d) IMEX-FBE2 phase speed
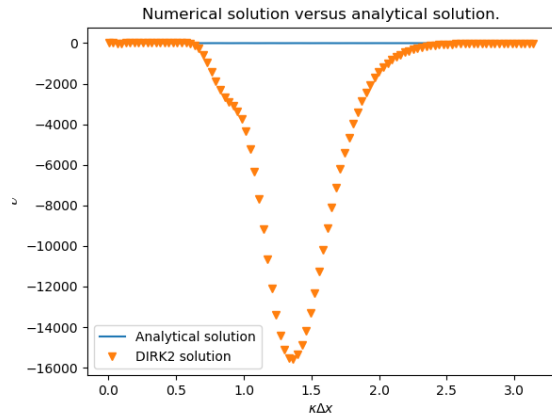
(e) IMEX-FBE1 group velocity
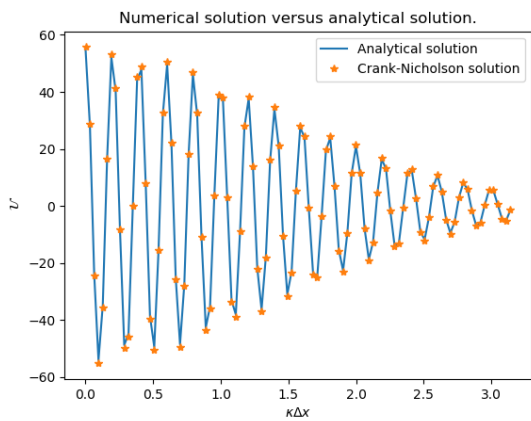
(f) IMEX-FBE2 group velovcity

Figure 5.8: Contours of the implicit-explicit Euler schemes applied in a diffusive-destructive 1D linear convection-diffusion-reaction problem.
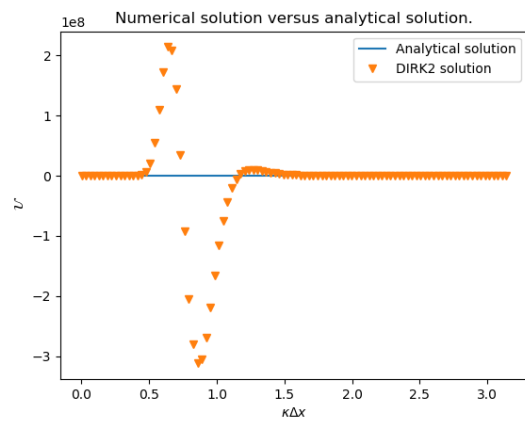
(a) Pure diffusive/Crank-Nicholson

(b) Pure diffusive/DIRK2

(c) Diffusive-productive/Crank-Nicholson

(d) Diffusive-productive/DIRK2
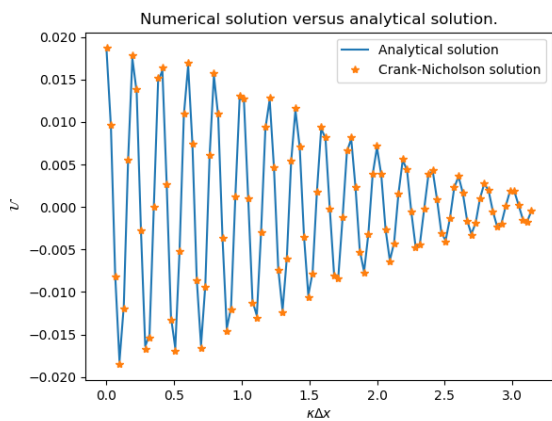
(e) Diffusive-destructive/Crank-Nicholson
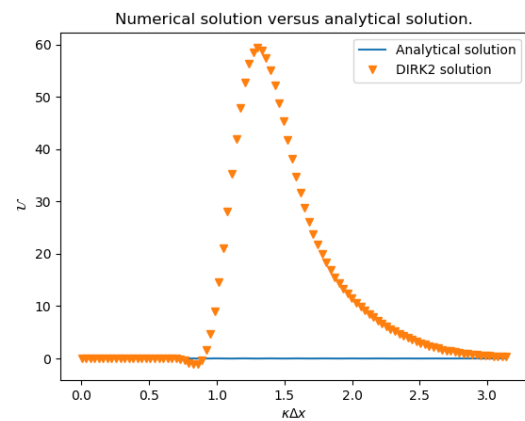
(f) Diffusive-destructive/DIRK2

Figure 5.9: Crank-Nicholson and DIRK2 solutions of the 1D linear convection-diffusion-reaction problem, computered with $P_e = 0.25$.

(a) Pure diffusive /IMEX-FBE1
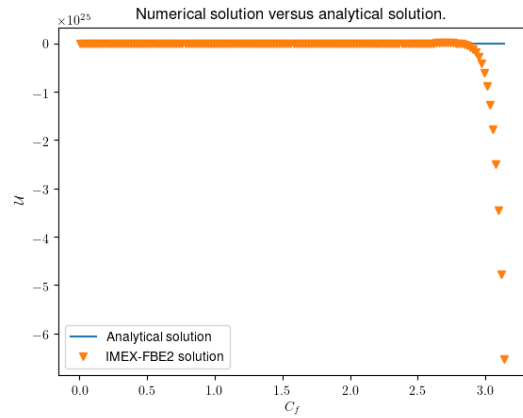
(b) Pure diffusive /IMEX-FBE2
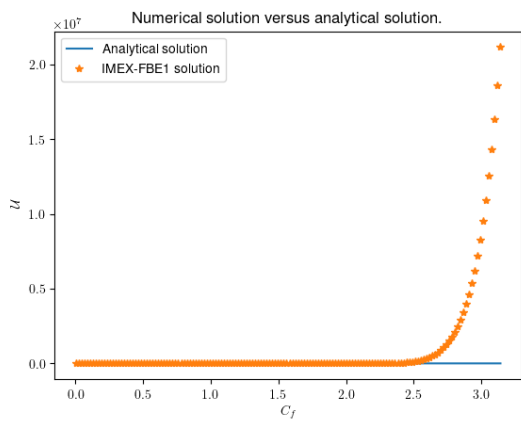
(c) Diffusive-productive/IMEX-FBE1

(d) Diffusive-productive/IMEX-FBE2
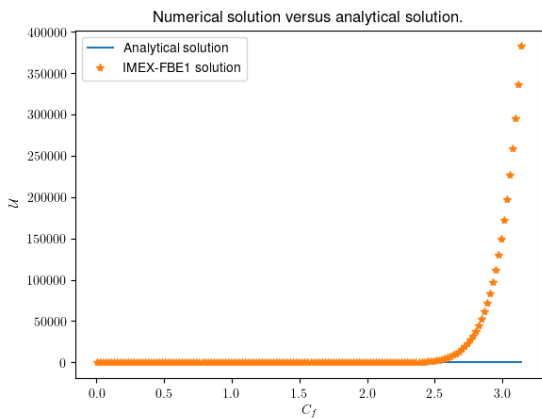
(e) Diffusive-destructive/IMEX-FBE1
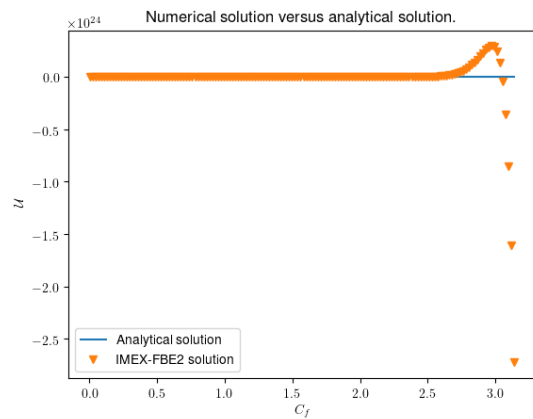
(f) Diffusive-destructive/IMEX-FBE1
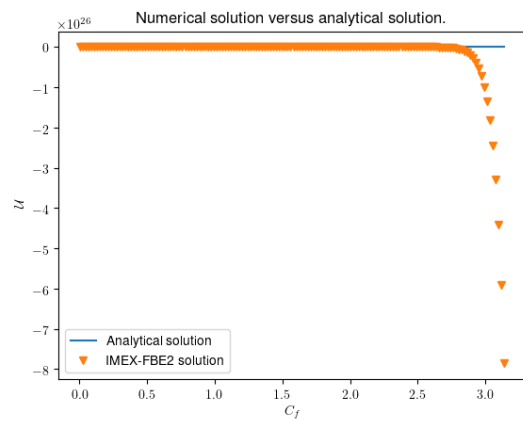
Figure 5.10: IMEX-FBE1 and IMEX-FBE2 solutions of the 1D linear convection-diffusion-reaction problem computed with $P_e = 0.25$.

Convergence test					
		IMEX-FBE1		IMEX-FBE2	
Transport	Spatial steps (N_x)	Error	Order	Error	Order
Diffusion	10	4.9012	-	4.9012	-
	20	2.8100	0.8026	2.8100	0.8026
	40	1.4371	0.9674	1.4371	0.9674
	80	0.7190	0.9991	0.7190	0.9991
	160	0.3660	0.9742	0.3660	0.9742
Advection	10	4.8363	-	4.8757	-
	20	2.7788	0.7994	2.7886	0.8060
	40	1.4103	0.9785	1.4129	0.9807
	80	0.7119	0.9862	0.7119	0.9889
	160	0.3568	0.9965	0.3568	0.9965
Advection-diffusion	10	4.9040	-	4.9436	-
	20	2.8065	0.8052	2.8166	0.8116
	40	1.4375	0.9652	1.4355	0.9724
	80	0.7192	0.9991	0.7192	0.9971
	160	0.3681	0.9663	0.3669	0.9710

Table 5.1: Errors and orders (L_∞) of the numerical schemes applied to solve system (5.43)-(5.44), errors were computed using analytical solution (5.46). The solutions and errors were computed for diffusion dominated, advection dominated and semi-linear transport cases using $\Delta t = 0.5\Delta x$.
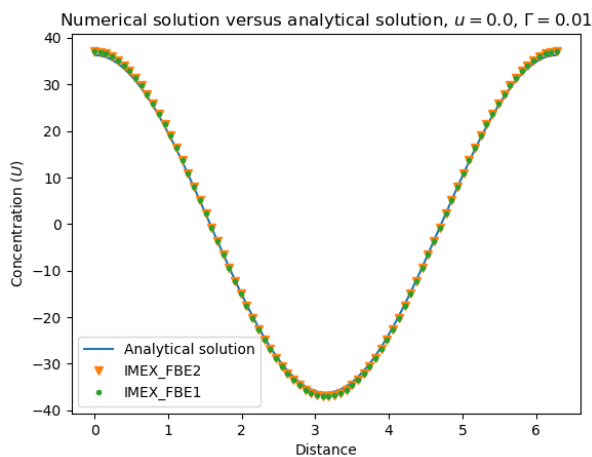
(a) Solutions/Diffusion transport
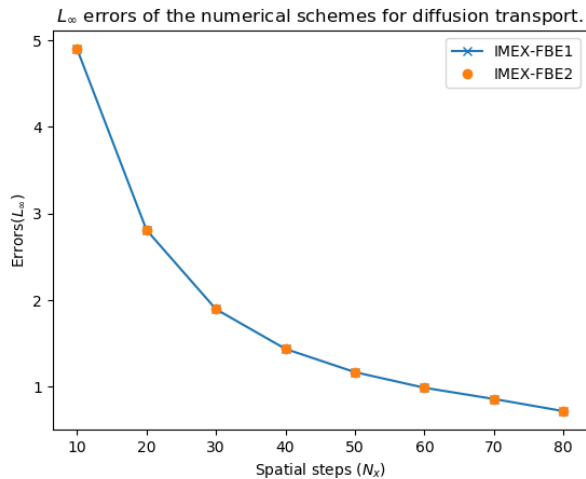
(b) L_∞ error/diffusion transport
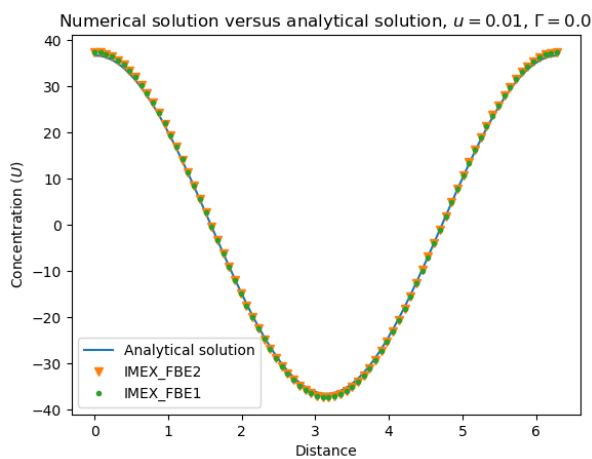
(c) Solutions/advective transport

(d) L_∞ error/advective transport
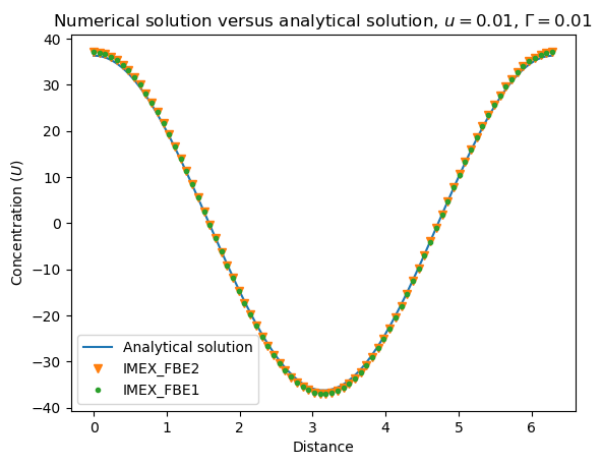
(e) Solution/advection-diffusion transport

(f) L_∞ error/advection-diffusion transport

Figure 5.11: Errors (L_∞) and numerical solutions of the 1D stiff system (5.43)-(5.44), errors were computed using analytical solution (5.46), computed with $\Delta t = 0.5\Delta x$.
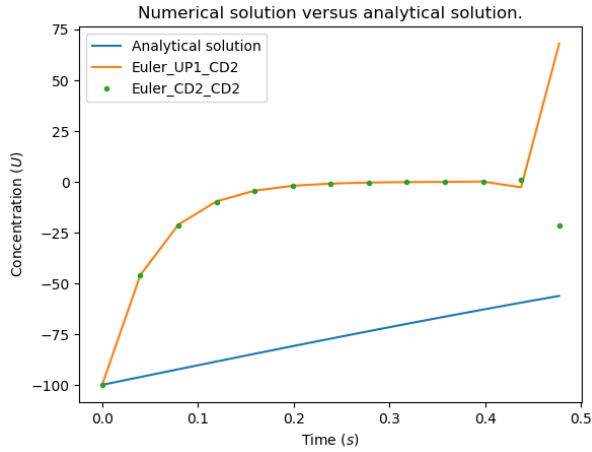
(a) Time/finite difference schemes
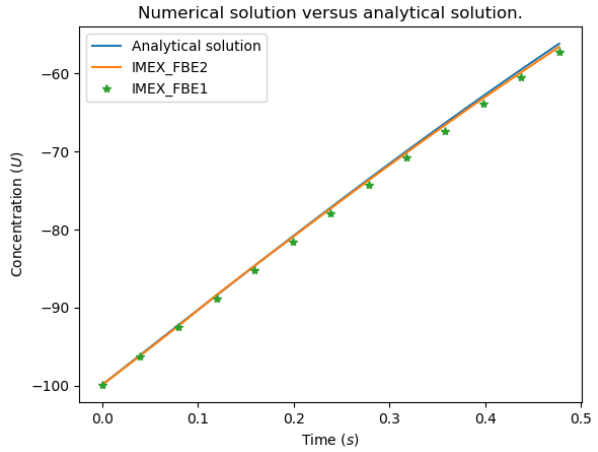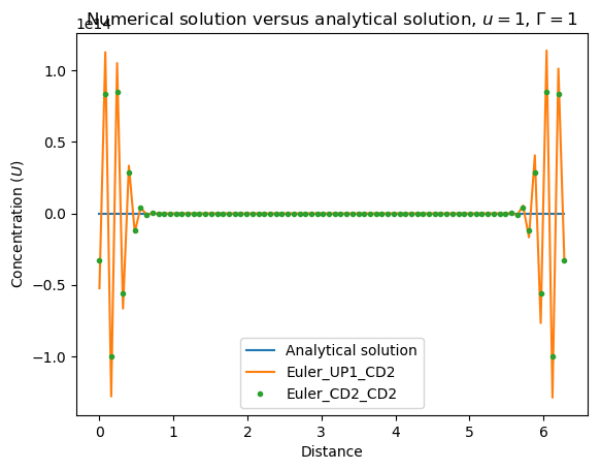
(b) Time/IMEX schemes

(c) Space/finite difference schemes
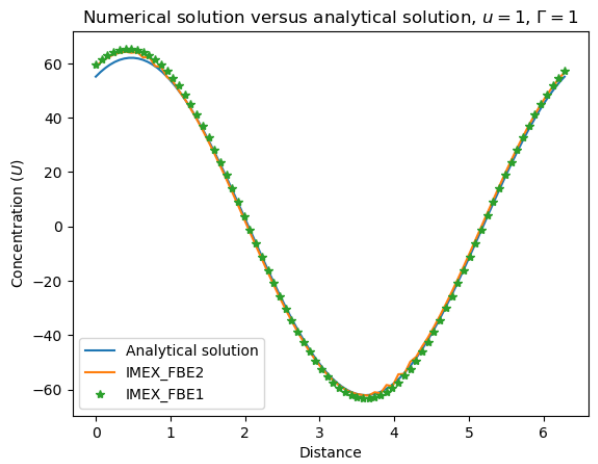
(d) Space/IMEX schemes
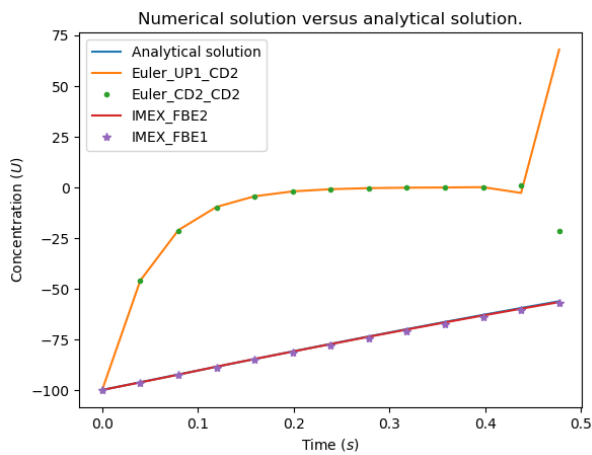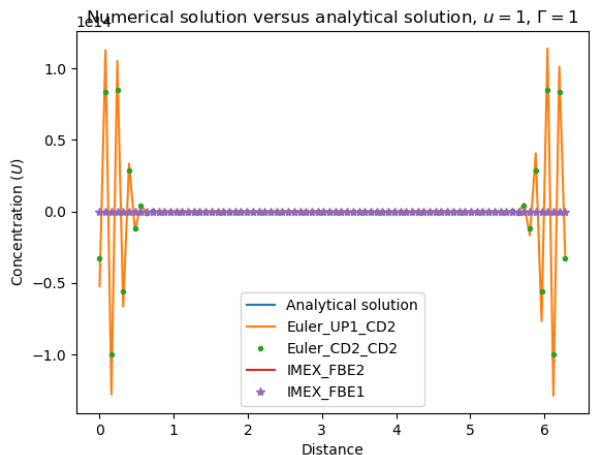
(e) Time/all schemes

(f) Space/all schemes

Figure 5.12: Numerical and analytical solutions for the stiff system of semi-linear transport problem (5.43)-(5.44), computed with time step size $\Delta t = 0.5\Delta x$. The IMEX schemes are compared (in space and time) with finite difference schemes of the same order.
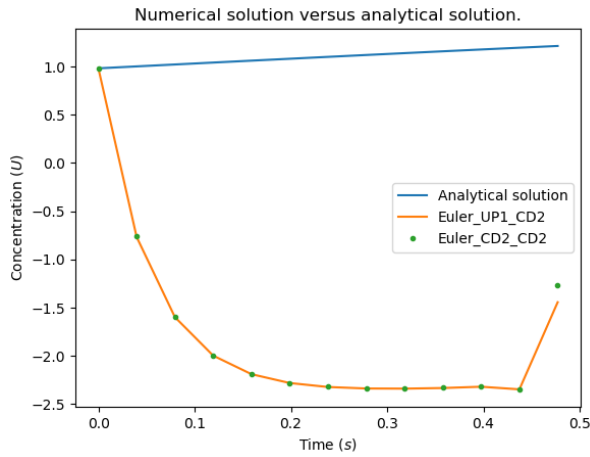
(a) Time/finite difference schemes
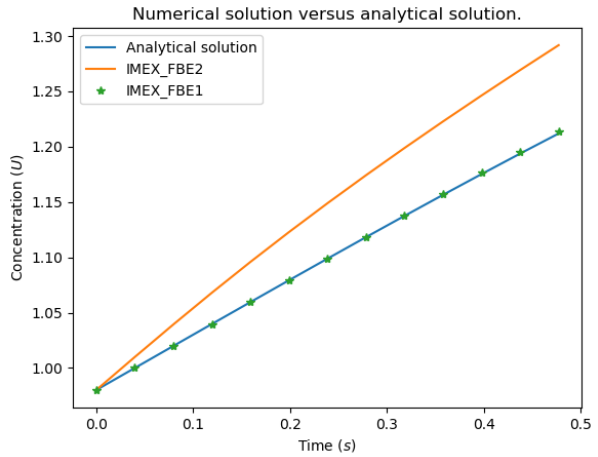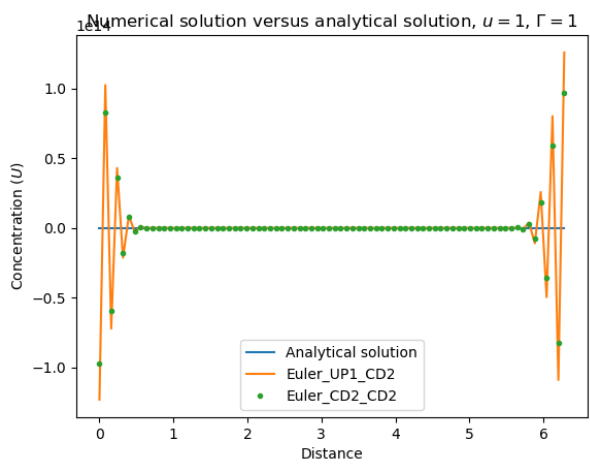
(b) Time/IMEX schemes

(c) Space/finite difference schemes
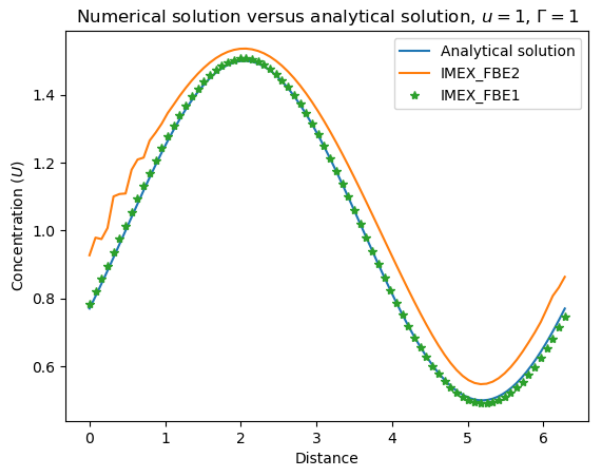
(d) Space/IMEX schemes
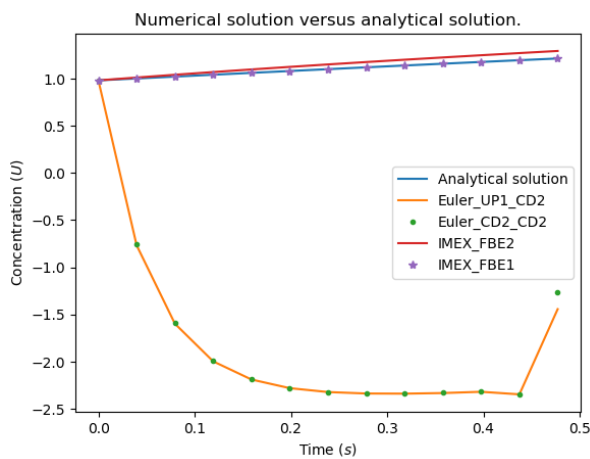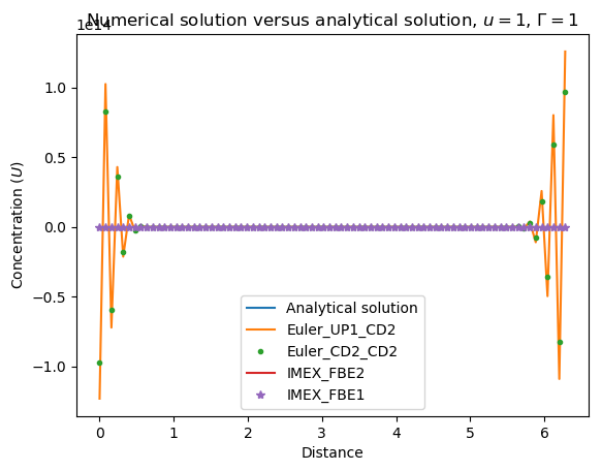
(e) Time/all schemes

(f) Space/all schemes

Figure 5.13: Numerical and analytical solutions for the 1D stiff nonlinear transport problem (5.50)-(5.51), computed with time step size $\Delta t = 0.3\Delta x$. The IMEX schemes are compared (in space and time) with finite difference schemes of the same order.
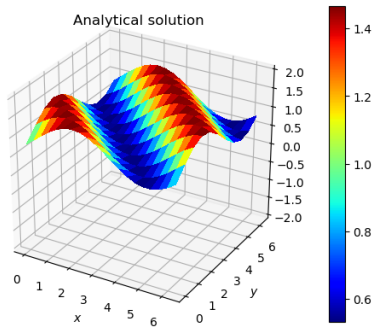
(a) Exact solution/nonlinear 2D problem
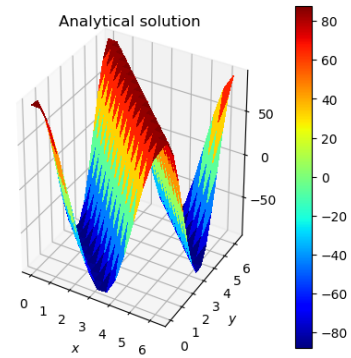
(b) Exact solution/stiff 2D linear system
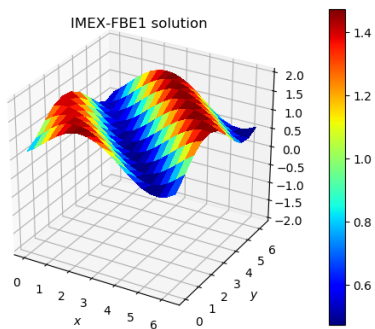
(c) IMEX-FBE1 /nonlinear 2D problem
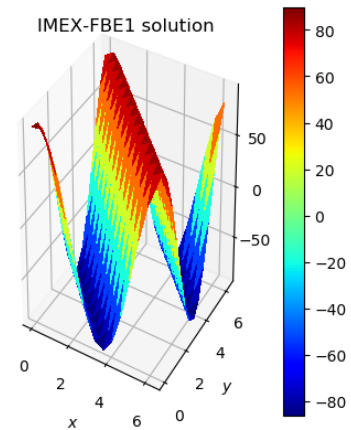
(d) IMEX-FBE1 /stiff 2D linear system
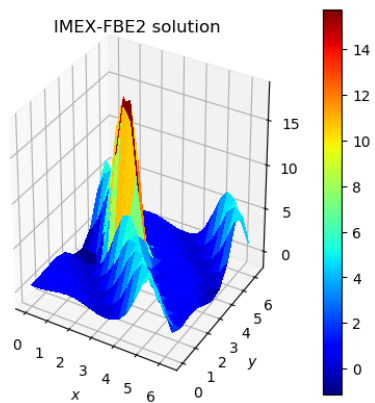
(e) IMEX-FBE2/nonlinear 2D problem
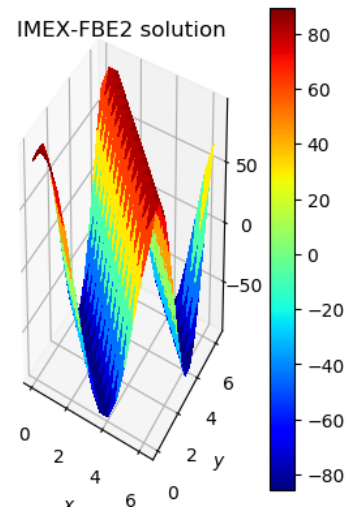
(f) IMEX-FBE2/stiff 2D linear system

Figure 5.14: Numerical and analytical solutions for the 2D system (5.52)-(5.53) and nonlinear transport problem (5.56), computed with $\Delta t = 0.2\Delta x$.
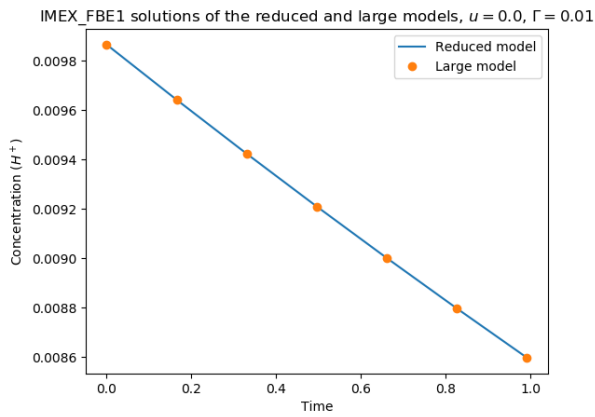
(a) IMEX-FBE1/diffusion transport
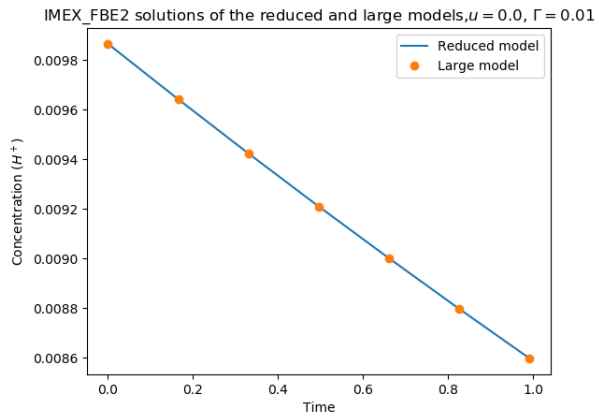
(b) IMEX-FBE2/diffusion transport
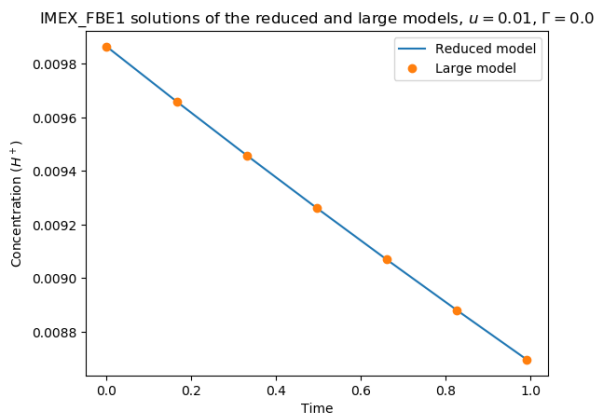
(c) IMEX-FBE1/advective transport
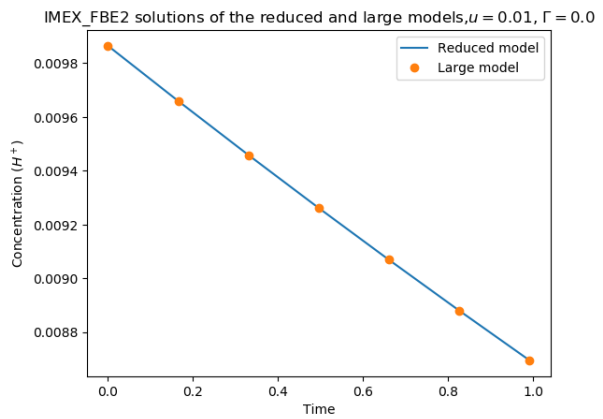
(d) IMEX-FBE2/advective transport
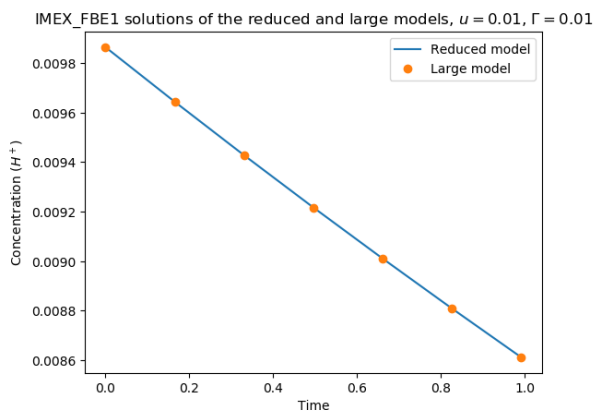
(e) IMEX-FBE1/advection-diffusion
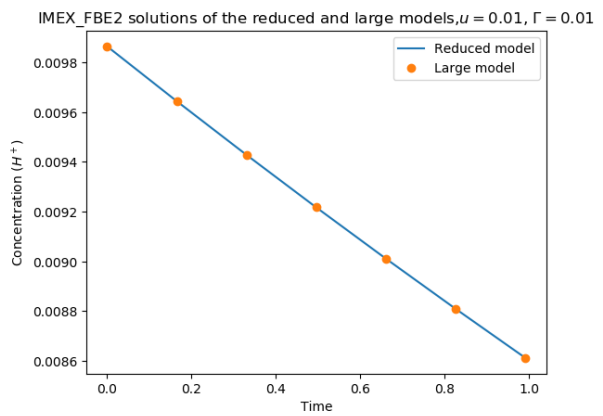
(f) IMEX-FBE2/advection-diffusion

Figure 5.15: IMEX-FBE1 and IMEX-FBE2 solutions of the acid drainage system
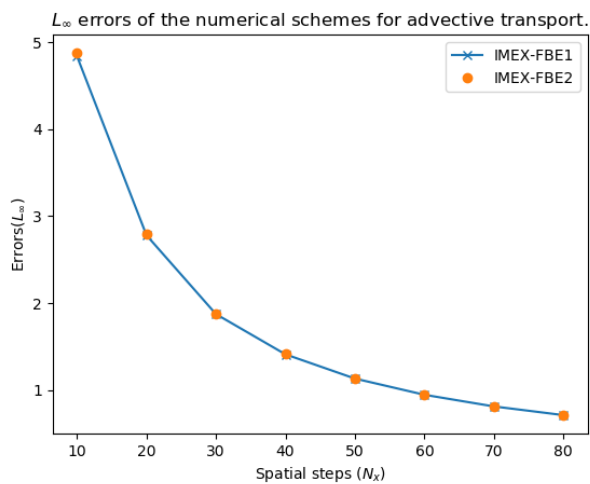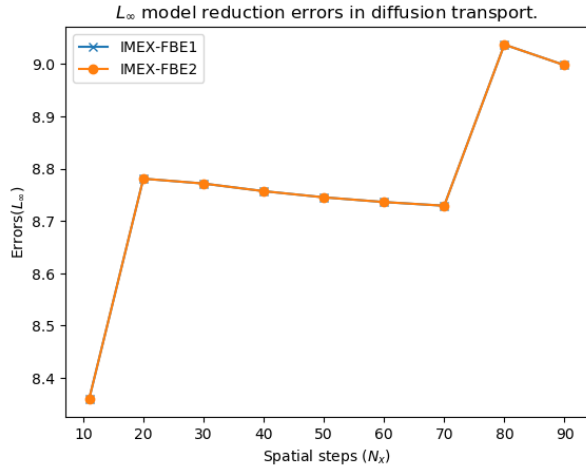
(a) L_∞ errors/diffusion transport
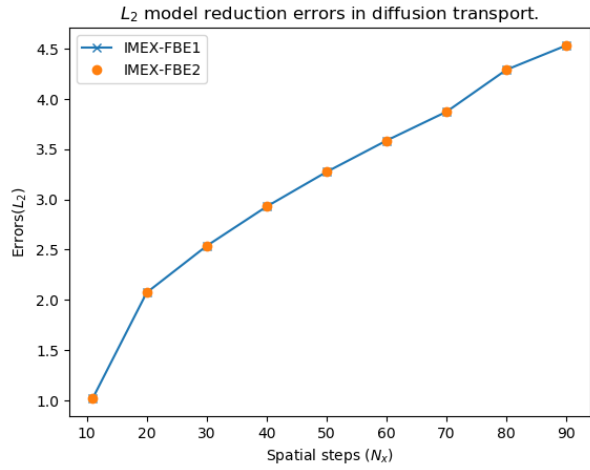
(b) L_2 errors/diffusion transport
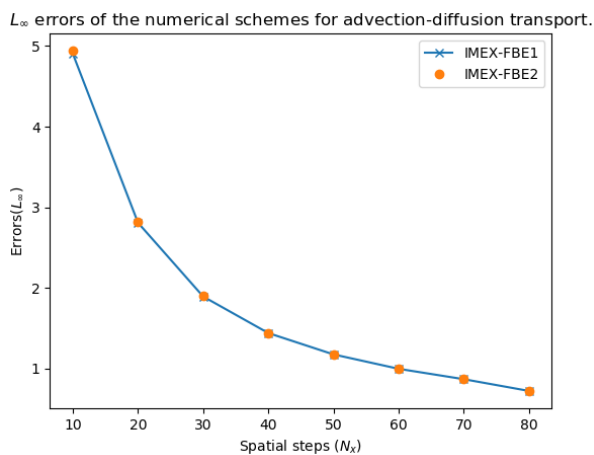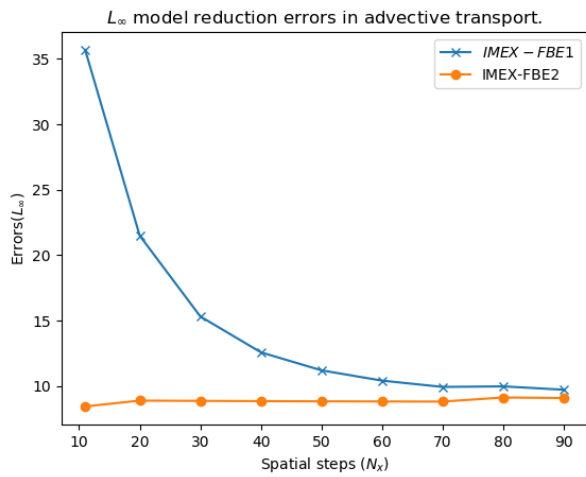

(c) L_∞ errors/advective transport
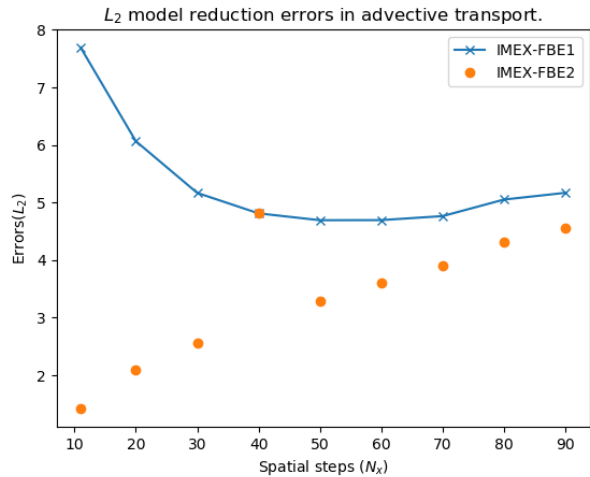
(d) L_2 error/advective transport
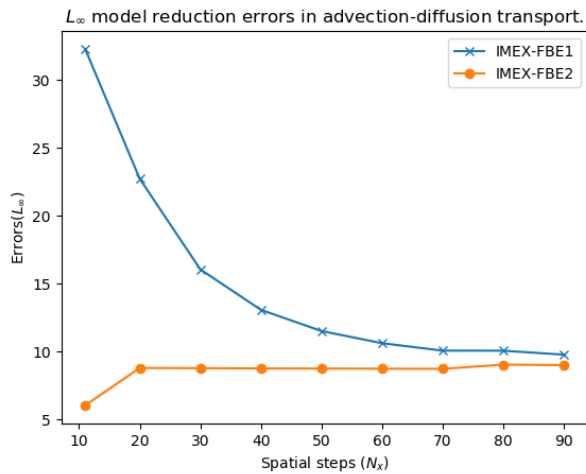
(e) L_∞ errors/advection-diffusion transport
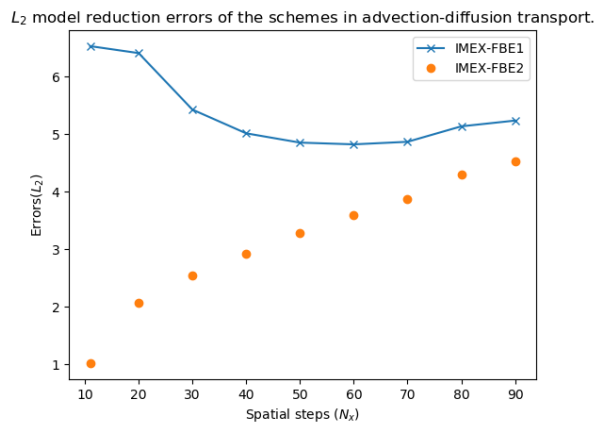
(f) L_2 errors/advection-diffusion transport

Figure 5.16: IMEX-FBE1 and IMEX-FBE2 solutions of the acid drainage system using 80 spatial steps.
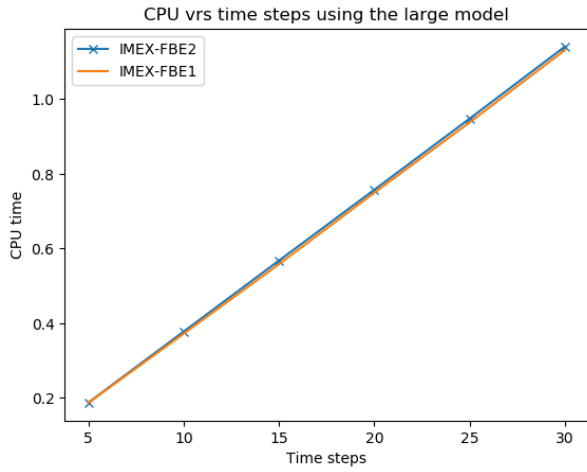
(a) CPU/large model/diffusion
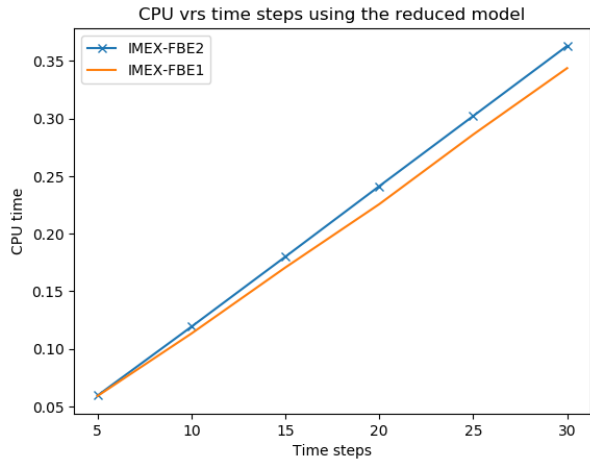
(b) CPU/reduced model/diffusion
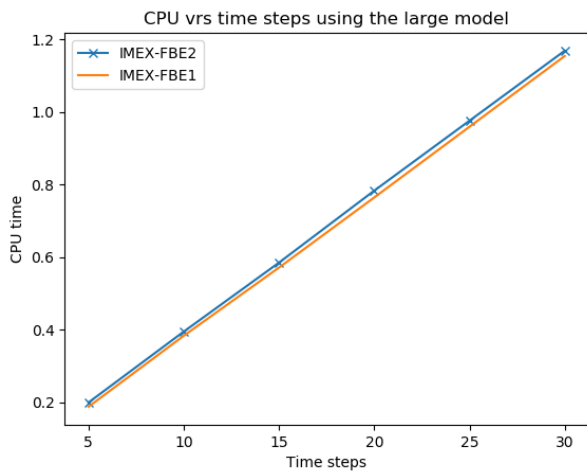
(c) CPU/large model/advection
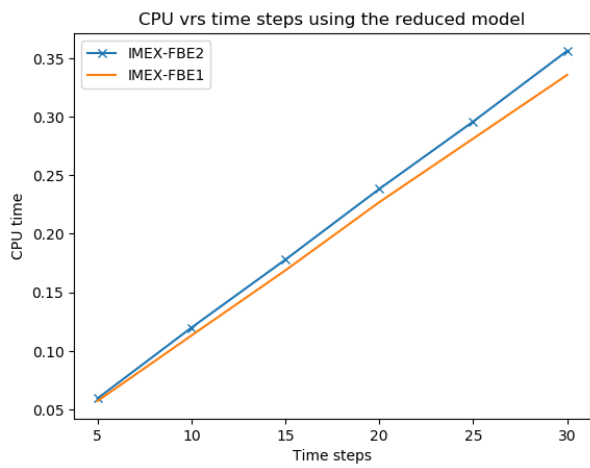
(d) CPU/reduced model/advection
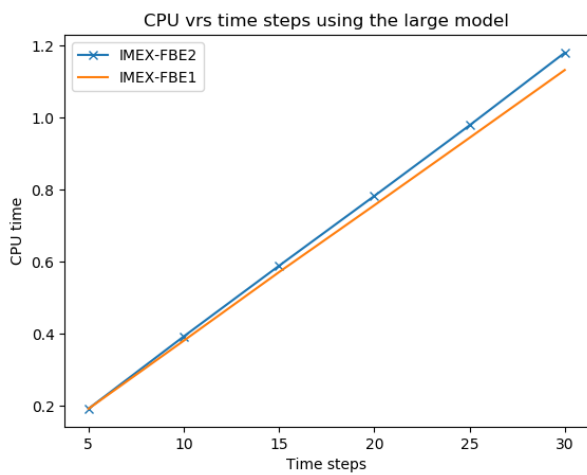
(e) CPU/large model/advection-diffusion
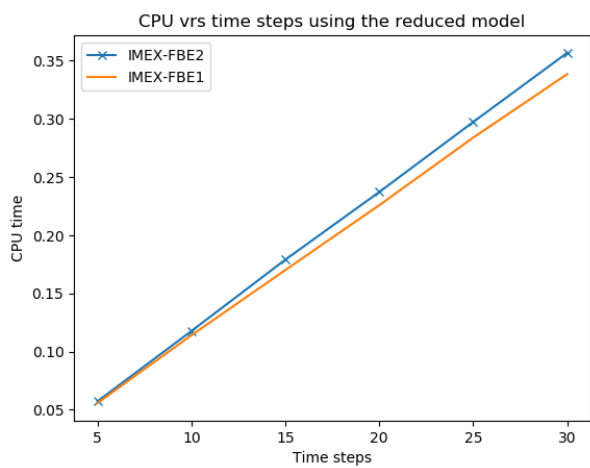
(f) CPU/reduced model/advection-diffusion

Figure 5.17: CPU time for IMEX-FBE1 and IMEX-FBE2 computed with 80 spatial steps in the acid drainage system

Chapter 6

Numerical simulation of reactive flows

6.1 Introduction

Generally, numerical simulation of reacting flows is challenging due to the presence of stiff nonlinear reaction terms present in the models [173, 166, 158, 90, 97, 183]. Research in the area of reacting flow simulation is not extensive, however, some progress has been made. It has been shown in [90, 97], that numerical schemes that are unsuitable for reacting flows show spurious oscillations due to their inability to resolve the stiff nonlinear terms. Some numerical simulation studies of reactive flows involving scalar advection have been discussed in [10, 30, 12], those involving advection and diffusion can be found in [16, 24, 113] and those involving advection-diffusion-reaction can be found in [148, 102]. In this chapter, the focus is on numerical simulation of shallow water flows with reactive transport.

Shallow-water equations have been employed to model water flows in lakes, reservoirs, oceans and many other flow processes where vertical length scales are insignificant, compared with horizontal length scales [85, 103]. Consequently, shallow water equations have contributed significantly to understanding dangerous phenomena such as hurricanes, dam breaks and tsunamis in the fields of oceanography, hydraulic engineering and climate studies [85, 103].

The homogeneous set of shallow water equations (i.e. shallow water equations with a flat bottom topography) form a nonlinear hyperbolic system of PDEs that admit non-smooth solutions (rarefactions, contact discontinuities and shocks) even when initial data is smooth [85, 162] (also see discussion in Chapter 2). In general, exact solutions do not exist for this non-linear hyperbolic problem and numerical simulation is non-trivial due to the non-smoothness of the solution. However, several numerical schemes have been developed and successfully used to simulate shallow water equations. Some finite volume schemes include Godunov schemes, Lax-Wendroff, Lax-Friedrichs, and Richtmyer schemes discussed in [100, 161, 162]. The authors in [86, 88] developed central-upwind schemes for conservation laws with application to shallow water equations. The central-upwind and higher orders schemes however, show oscillations near discontinuities. In general, it has been established theoretically in [54], that linear second order schemes show oscillations near discontinuities. However, very accurate schemes have been designed and successfully applied to simulate the homogeneous shallow water model.

Further, the homogeneous shallow water equations are not accurate for modelling flow processes that occur in a domain that has a non-flat bottom topography [101]. Thus, a source term is introduced to account for the non-flat bottom topography. The resulting non-homogeneous model is a non-linear system of PDEs called balanced laws [85, 8, 101]. A special property (called C-property) associated with balanced laws is a balance between convective momentum fluxes and geometric source terms when the flow becomes steady [17]. The C-property poses a constraint for some numerical schemes, however, other numerical schemes called well-balanced schemes [8, 85, 101], satisfy the C-property. Numerical schemes for balanced laws that are not well-balanced are highly inaccurate, thus, efforts have been made to construct well-balanced schemes [8, 87, 99, 167, 75]. Notably, pioneering research on well balanced schemes include the works of [57, 56]. Among several techniques used to construct well balanced schemes include the quasi-steady wave propagation technique in [99], surface gradient reconstruction in [187], source term decomposition with Weighted Essential Non-oscillatory (WENO) reconstruction in [177, 178, 179, 180], hydrostatic reconstruction in [8, 9] and quadrature with WENO reconstruction in [118]. Well balanced Hermite WENO schemes have been constructed by [25, 128] and well balanced cell-centered schemes have been constructed in [65]. Schemes

that are not well balanced but are convergent can in some cases be used to simulate (using very fine grids) however, such an approach has cost implications [85]. Further, well-balanced schemes have been shown to resolve small perturbations at steady state better than the non-well-balanced schemes [127, 181].

Furthermore, the water depth in shallow water model is a non-negative quantity and must remain non-negative in time and space. However, not all schemes can guarantee the non-negativity of the water depth, especially in problems that involve dry/wet areas in the flow domain [8, 85, 103]. Numerical schemes for shallow water models that satisfy the non-negativity constraint are known as positivity preserving schemes [85]. Some well balanced positivity preserving schemes can be found in [13, 20, 23, 49]. If the non-negativity constraint is not satisfied by a particular numerical scheme, spurious oscillations appear in the solution.

In this chapter, shallow water equations are coupled with reactive transport equations for modelling flow processes that involve chemical/sediment transport with reactions, in rivers, lakes, channels and other physico-chemical systems (that satisfy the shallow water assumption) with complex topographies in the flow domain. The resulting model is a nonlinear set of parabolic Partial Differential Equations (PDEs) that is stiff (i.e. the PDEs govern advective, diffusive and reactive processes where there is a wide time scale difference in information propagation) [173, 166, 106] and have high degrees of freedom due to many chemical species participating in several reactions. Since *IMEX constructions are efficient for resolving stiffness* (as discussed in Chapters 4 and 5), *conservative well-balanced constructions can resolve complex topographies and non-smoothness issues* (see discussions in [8, 85, 101]), *the stoichiometric decoupling method can reduce the degrees of freedom* (discussed in Chapters 3 and 4), *second (or higher) order consistent schemes can reduce simulation cost* (as discussed in [73, 85, 63]) and finally, *strong stability-preserving schemes can significantly reduce spurious oscillations* (as discussed in [63, 82]), the goal here is to present a scheme that possesses almost all the properties listed.

6.2 Modelling reactive flow processes

Following the discussion in Chapter 2, if a flow with reactive transport through a horizontal rectangular domain Ω_h (with non-flat bottom topography) satisfies the shallow water assumption, then the governing equations for momentum and mass transport in 1D are given by the balance law:

$$\frac{\partial}{\partial t}\mathbf{Q} + \frac{\partial}{\partial x}\mathbf{f}(\mathbf{Q}) = \mathbf{S}(\mathbf{Q}), \quad \in [0, T) \times \Omega_h \tag{6.1}$$

$$\mathbf{Q}(0, x) = \mathbf{Q}_0(x)$$

where T is final time, $\mathbf{Q} = (h, hu, h\mathbf{U})^{Tr}$, $\mathbf{Q}_0 = (h_0, h_0 u_0, \mathbf{U}_0)^{Tr}$, $\mathbf{f}(\mathbf{Q}) = (hu, \ hu^2 + \frac{1}{2}gh^2, \ hu\mathbf{U})^{Tr}$ and $\mathbf{S}(\mathbf{Q}) = \left(0, \ -gh\frac{\partial \eta}{\partial x} - \frac{gC^2|u|u}{h^{1/3}}, \ \frac{\partial}{\partial x}\Gamma\frac{\partial \mathbf{U}}{\partial x} + \mathbf{S}_h\right)^{Tr}$.

For closure of the governing equations (6.1), the term $\mathbf{S}_h(\mathbf{Q})$ (that quantifies sources or sinks of the species being transported) must be defined. In sediment transport studies, $\mathbf{S}_h$ is a sink known as isotherm. For reactive flow problems that satisfy the law of mass action, $\mathbf{S}_h$ is a polynomial that measures either a sink or source. If the fluid is transporting N chemical species that are engaged in M reactions, the term $\mathbf{S}_h$ is defined by the mass action rate law (3.4) [173], thus:

$$\mathbf{S}_h(\mathbf{Q}) = \sigma\mathbf{R}(\mathbf{Q}), \tag{6.2}$$

where σ is a stoichiometric matrix, $\mathbf{R}(\mathbf{Q}) = \left(R_1(\mathbf{Q}), R_2(\mathbf{Q}), \dots, R_M(\mathbf{Q})\right)^{Tr}$ and $R_r(\mathbf{Q})$, is given by (3.2) for all $r = 1, 2, \dots, M$. However, the standard procedure of closure (i.e. closure with Expressions (3.2) and (3.4)) is expensive. Thus, the stoichiometric method (i.e. Equation (3.21)) is proposed here for closure of (6.1). Hence, $\mathbf{S}_h$ is defined by (6.2) with $R_r(\mathbf{Q})$ given by (3.21), for any index r.

A classical solution of PDE system (6.1), is any set of C^1 functions $\mathbf{Q}$, that satisfies the system. Following discussions in Chapter 2 and [161, 53, 100], it is clear that the nonlinear PDE system (6.1) admits non-smooth solutions (also known as *weak solutions*) even if the initial data is smooth, thus, classical solutions do not exist for all time.

A challenge faced in simulation studies is that, weak solutions are not unique for general nonlinear conservation laws [53, 98]. Due to the non-uniqueness, additional constraints known as entropy conditions must be imposed on the weak solutions to yield

physically correct solutions [53, 98]. The entropy conditions can be imposed explicitly or implicitly, the explicit procedures are known as shock tracking methods while the implicit ones are known as shock capturing methods [98].

Another interesting challenge is encountered when simulating steady flows (equilibrium flows with asymptotic solutions have been discussed in [66]). When the flow achieves equilibrium, the momentum and continuity equations assume the form:

$$\frac{\partial}{\partial x} uh = 0, \tag{6.3}$$

$$\frac{\partial}{\partial x} \left(\frac{1}{2} u^2 + gh + g\eta \right) = 0. \tag{6.4}$$

Thus, from ODEs (6.3) and (6.4), it is clear that the mass flux and mechanical energy of the fluid are constant. One particular flow equilibrium of interest is the lake at rest, where the fluid velocity vanishes. For such an equilibrium, it can be deduced from the momentum equation (6.4) that the total depth of the fluid is constant (i.e. $h + \eta = constant$). The challenge, therefore, is to find asymptotic solutions or shock-capturing numerical schemes that can preserve the invariants in the steady flows.

6.3 Numerical schemes

In Section 6.3, some of the simulation challenges of the reactive flow model were highlighted. In this section, numerical schemes that are shock capturing, stable, at least second order consistent in smooth regions, monotone and computationally less expensive are presented here. Numerical schemes for conservation laws that possess the above listed properties are known as high resolution schemes [100, 98]. Since the model under consideration is a balance law, the schemes presented here must also satisfy the C-property, thus, the challenge here is to construct well-balanced high resolution schemes that are compatible with the stoichiometric decoupling method. Construction of the well-balanced suitable schemes are elaborated in the following subsections.

6.3.1 Spatial discretization

Firstly, a uniform discretization of the domain into control volumes is considered, where the ith control volume has left and right interfaces denoted by $x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}$ respectively,

a center located at $x_i = \frac{x_{i-\frac{1}{2}} + x_{i+\frac{1}{2}}}{2}$ and a size denoted by $\Delta x = x_{i+\frac{1}{2}} - x_{i-\frac{1}{2}}$. The time interval is also discretized uniformly with step size Δt (i.e. $t^{n+1} = t^n + \Delta t$ where t^n denotes the nth time step). Further, the averages of the unknowns $\mathbf{Q}(t, x)$ and source terms $\mathbf{S}(\mathbf{Q})$, over the control volume are respectively denoted by:

$$\mathbf{Q}_i(t) = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{Q}(t, x) dx, \quad \text{and} \quad \mathbf{S}_i = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathbf{S}(\mathbf{Q}) dx. \tag{6.5}$$

Thus, integrating (6.1) over the control volume and applying (6.5) yields:

$$\Delta x \frac{d}{dt} \mathbf{Q}_i(t) = -\left(\mathbf{f}_{i+\frac{1}{2}} - \mathbf{f}_{i-\frac{1}{2}} \right) + \mathbf{S}_i, \tag{6.6}$$

where $\mathbf{f}_{i+\frac{1}{2}}$ is the value of the flux at the $i + \frac{1}{2}$ interface of the control volume. The value of the flux is not known at the interfaces in advance, thus, must be approximated using consistent numerical fluxes and cell center values (e.g. $\mathbf{f}_{i+\frac{1}{2}} = \mathbf{f}^{num}(\mathbf{Q}_{i+1}, \mathbf{Q}_i)$ where $\mathbf{f}^{num}$ denotes a numerical flux function). Instead of the cell center values, reconstructed interface values can also be used (and will be used here) to approximate the flux (i.e. $\mathbf{f}_{i+\frac{1}{2}} = \mathbf{f}^{num}(\mathbf{Q}_{i+\frac{1}{2}}^R, \mathbf{Q}_{i+\frac{1}{2}}^L)$, where $\mathbf{Q}_{i+\frac{1}{2}}^R, \mathbf{Q}_{i+\frac{1}{2}}^L$ are reconstructed cell interface values). It is an established fact that linear and piecewise constant reconstructions yield second order and first order schemes, respectively, [161, 100, 187]. Thus, linear reconstruction is employed to obtain slope limited values at the right and left side of the $i + \frac{1}{2}$ interface, i.e.:

$$\mathbf{Q}_{i+\frac{1}{2}}^R = \mathbf{Q}_{i+1} - 0.5 \Delta x \delta \mathbf{Q}_{i+1} \quad \text{and} \quad \mathbf{Q}_{i+\frac{1}{2}}^L = \mathbf{Q}_i + 0.5 \Delta x \delta \mathbf{Q}_i, \tag{6.7}$$

where $\delta \mathbf{Q}_i$ is a gradient of the unknown that is computed from a slope limiter function. Among many limiter functions, the *minmod* limiter function of [187] is considered here, i.e.:

$$\delta \mathbf{Q}_i = \max(0, \min(a_i, b_i)), \tag{6.8}$$

where

$$a_i = \frac{\mathbf{Q}_{i+1} - \mathbf{Q}_i}{\Delta x} \quad \text{and} \quad b_i = \frac{\mathbf{Q}_i - \mathbf{Q}_{i-1}}{\Delta x}.$$

It is clear from (6.7) that if the gradients $\delta \mathbf{Q}_i$ and $\delta \mathbf{Q}_{i+1}$ assume zero values, the reconstruction becomes piecewise constant and the values of $\mathbf{Q}_{i+\frac{1}{2}}^R$ and $\mathbf{Q}_{i+\frac{1}{2}}^L$ are simply the cell

center values. Henceforth, schemes that result from linear reconstruction will be referred to as *minmod* schemes and those that result from piecewise constant reconstruction will be referred to as *upwind* schemes.

Furthermore, motivated by the work in [8], hydrostatic reconstruction is applied and non-negativity constraints are imposed on the water depth after the limiter reconstruction (6.7). Consequently, the interface values for the fluid velocity and species concentration are given by (6.7), however, the water depth is given by:

$$h^+_{i+\frac{1}{2}} = \max\left(0, h^R_{i+\frac{1}{2}} + \eta^R_{i+\frac{1}{2}} - \eta_{i+\frac{1}{2}}\right) \text{ and } h^-_{i+\frac{1}{2}} = \max\left(0, h^L_{i+\frac{1}{2}} + \eta^L_{i+\frac{1}{2}} - \eta_{i+\frac{1}{2}}\right), \quad (6.9)$$

where $\eta_{i+\frac{1}{2}} = \max(\eta^R_{i+\frac{1}{2}}, \eta^L_{i+\frac{1}{2}})$. Thus, the interface values of the conserved variables (for evaluating the numerical flux) are given by:

$$\mathbf{Q}^+_{i+\frac{1}{2}} = (h^+_{i+\frac{1}{2}}, h^+_{i+\frac{1}{2}} u^R_{i+\frac{1}{2}}, h^+_{i+\frac{1}{2}} \mathbf{U}^R_{i+\frac{1}{2}})^{Tr} \text{ and } \mathbf{Q}^-_{i+\frac{1}{2}} = (h^-_{i+\frac{1}{2}}, h^-_{i+\frac{1}{2}} u^L_{i+\frac{1}{2}}, h^-_{i+\frac{1}{2}} \mathbf{U}^L_{i+\frac{1}{2}})^{Tr}.$$
$$(6.10)$$

6.3.2 Numerical fluxes and wave speed estimates

Given the reconstructed interface values $\mathbf{Q}^+_{i+\frac{1}{2}}$ and $\mathbf{Q}^-_{i+\frac{1}{2}}$ in (6.7), another task in the spatial discretization is to find consistent numerical fluxes. In this section, Godunov's procedure (that involves solving a local Riemann problem and applying the solution to compute the fluxes) is used to derive the fluxes for the reactive shallow water flow problem under consideration.

Firstly, consider a Riemann problem of the homogeneous part of the reactive flow model (6.1) given by:

$$\frac{\partial \mathbf{Q}}{\partial t} + \frac{\partial}{\partial x} \mathbf{f}(\mathbf{Q}) = 0, \text{ in } [x_l, x_r] \times [0, T),$$

$$(6.11)$$

$$\mathbf{Q}(x,0) = \mathbf{Q}^0(x) = \begin{cases} \mathbf{Q}_L & \text{for } x < 0, \\ \mathbf{Q}_R & \text{for } x > 0. \end{cases}$$

It was shown in Chapter 2, that the similarity solution to the Riemann problem consists of three states (denoted by $\mathbf{Q}_L, \mathbf{Q}_{*L}, \mathbf{Q}_{*R}$ and $\mathbf{Q}_R$) that are separated by three

waves (with speeds denoted by S_L, S_* and S_R), i.e.:

$$\mathbf{Q}\left(\frac{x}{t}\right) = \begin{cases} \mathbf{Q}_L & \text{if } \frac{x}{t} \leq S_L, \\ \mathbf{Q}_{*L} & \text{if } S_L \leq \frac{x}{t} \leq S_*, \\ \mathbf{Q}_{*R} & \text{if } S_* \leq \frac{x}{t} \leq S_R, \\ \mathbf{Q}_R & \text{if } \frac{x}{t} \geq S_R. \end{cases} \tag{6.12}$$

Thus, applying (6.12) to evaluate the natural flux $\mathbf{f}$, (defined in (6.1)) yields the Godunov's flux:

$$\mathbf{f}^{hllc}\left(\mathbf{Q}\left(\frac{x}{t} = 0\right)\right) = \begin{cases} \mathbf{f}(\mathbf{Q}_L) & \text{if } 0 \leq S_L, \\ \mathbf{f}(\mathbf{Q}_{*L}) & \text{if } S_L \leq 0 \leq S_*, \\ \mathbf{f}(\mathbf{Q}_{*R}) & \text{if } S_* \leq 0 \leq S_R, \\ \mathbf{f}(\mathbf{Q}_R) & \text{if } 0 \geq S_R. \end{cases} \tag{6.13}$$

However, application of the Godunov's procedure using an exact Riemann solver is expensive. Thus, we present an approximate Godunov flux (known as HLLC flux [161]) for the reactive shallow water flow model, that is less expensive.

The wave speeds can be estimated directly or indirectly by a depth average approach [161, 100]. In [36], direct estimates for the fastest and slowest signal speeds are provided for the purpose of simulating the Euler equations in gas dynamics. For the shallow water flow problem under consideration, similar estimates are given by [161]:

$$S_L = u_L - c_L, \quad S_R = u_R + c_R, \tag{6.14}$$

and

$$S_L = \min(u_L - c_L, u_R - c_R), \quad S_R = \max(u_L + c_L, u_R + c_R), \tag{6.15}$$

where c_L and c_R are the left and right celerities defined in Chapter 2. We refer to estimates (6.14) and (6.15) as Davis estimates. Further, in [36, 41] Roe averaged quantities are proposed for the purpose of computing the fastest and slowest signals (in the Euler Equations). A similar estimate for the model under consideration is given by:

$$S_L = \tilde{u} - \tilde{c} \quad \text{and} \quad S_R = \tilde{u} + \tilde{c} \tag{6.16}$$

where

$$\tilde{u} = \frac{\sqrt{h_L}u_L + \sqrt{h_R}u_R}{\sqrt{h_L} + \sqrt{h_R}} \quad \text{and} \quad \tilde{c} = \sqrt{\frac{g}{2}(h_L + h_R)}. \tag{6.17}$$

We refer to estimates (6.16) and (6.17) as Roe estimates. In [67], the depth estimate approach has been used to estimate the wave speeds. If the intermediate fluid depth h_*, is known, the fastest and slowest signal speeds are computed using:

$$S_L = u_L - c_L q_L, \quad S_R = u_R + c_R q_R, \tag{6.18}$$

where for index $k = L, R$:

$$q_k = \begin{cases} \sqrt{\frac{1}{2}h_* \frac{(h_* + h_k)}{h_k^2}} & \text{if } h_* > h_k, \\ 1 & \text{if } h_* \leq h_k. \end{cases} \tag{6.19}$$

In addition to the estimates for the fastest and slowest wave speeds, the speed of the middle contact wave must also be estimated. The intermediate depth and contact wave speed are taken as follows [67]:

$$S_* = \frac{1}{2}(u_L + u_R) - (h_R - h_L)\frac{c_L + c_R}{h_L + h_R} \tag{6.20}$$

and

$$h_* = \frac{1}{2}(h_L + h_R) - \frac{1}{4}(u_R - u_L)\frac{h_L + h_R}{c_L + c_R}. \tag{6.21}$$

We refer to estimates (6.20) and (6.21) as depth estimates. Given the estimates for the wave speeds, we can then reformulate the Riemann problem (6.11) into a form, suitable for further discussions.

Proposition 6.1 (Consistency formulation). *Assuming that the wave speed $S_L < S_* < S_R$ are defined and that the intermediate state variables are defined by the integral averages:*

$$\mathbf{Q}_{*L} = \frac{1}{T(S_* - S_L)}\int_{TS_L}^{TS_*} \mathbf{Q}(x, T)dx, \quad \mathbf{Q}_{*R} = \frac{1}{T(S_R - S_*)}\int_{TS_*}^{TS_R} \mathbf{Q}(x, T)dx \tag{6.22}$$

then the integral formulation of the Riemann problem (6.11) yields:

$$\frac{S_* - S_L}{S_R - S_L}\mathbf{Q}_{*L} + \frac{S_R - S_*}{S_R - S_L}\mathbf{Q}_{*R} = \frac{S_R\mathbf{Q}_R - S_L\mathbf{Q}_L + \mathbf{f}_L - \mathbf{f}_R}{S_R - S_L}, \tag{6.23}$$

where $\mathbf{f}_L = \mathbf{f}(\mathbf{Q}_L)$ and $\mathbf{f}_R = \mathbf{f}(\mathbf{Q}_R)$.

Proof. Firstly the integral formulation of the Riemann problem (6.11) is given by:

$$\int_{x_l}^{x_r} \mathbf{Q}(x,T)dx = \int_{x_l}^{x_r} \mathbf{Q}(x,0)dx + \int_0^T \mathbf{f}(\mathbf{Q}(x_l,t))dt - \int_0^T \mathbf{f}(\mathbf{Q}(x_r,t))dt$$

$$= x_r \mathbf{Q}_R - x_l \mathbf{Q}_L + T(\mathbf{f}_L - \mathbf{f}_R), \tag{6.24}$$

where $\mathbf{f}_L = \mathbf{f}(\mathbf{Q}_L)$ and $\mathbf{f}_R = \mathbf{f}(\mathbf{Q}_R)$. However, at any chosen final time T, the distances covered by the waves satisfy the following inequalities:

$$x_l \le TS_L \le TS_* \le TS_R \le x_R. \tag{6.25}$$

Thus, by using (6.25), the integral of the function $\mathbf{Q}(x,t)$ evaluated at the final time T, is given by:

$$\int_{x_l}^{x_r} \mathbf{Q}(x,T)dx = \int_{x_l}^{TS_L} \mathbf{Q}(x,T)dx + \int_{TS_L}^{TS_*} \mathbf{Q}(x,T)dx + \int_{TS_*}^{TS_R} \mathbf{Q}(x,T)dx + \int_{TS_R}^{x_r} \mathbf{Q}(x,T)dx,$$

$$= \int_{TS_L}^{TS_*} \mathbf{Q}(x,T)dx + \int_{TS_*}^{TS_R} \mathbf{Q}(x,T)dx$$

$$+ (TS_L - x_l)\mathbf{Q}_L + (x_r - TS_R)\mathbf{Q}_R. \tag{6.26}$$

Applying integral expression (6.24) in (6.26) and simplifying the result yields:

$$\int_{TS_L}^{TS_*} \mathbf{Q}(x,T)dx + \int_{TS_*}^{TS_R} \mathbf{Q}(x,T)dx = T(S_R \mathbf{Q}_R - S_L \mathbf{Q}_L + \mathbf{f}_L - \mathbf{f}_R). \tag{6.27}$$

Dividing through by the distance $T(S_R - S_L)$ between the fastest and slowest signals yields [161]:

$$\frac{1}{T(S_R - S_L)} \int_{TS_L}^{TS_*} \mathbf{Q}(x,T)dx + \frac{1}{T(S_R - S_L)} \int_{TS_*}^{TS_R} \mathbf{Q}(x,T)dx = \frac{S_R \mathbf{Q}_R - S_L \mathbf{Q}_L + \mathbf{f}_L - \mathbf{f}_R}{S_R - S_L}.$$

$$\tag{6.28}$$

Applying the integral averages (6.22) in (6.28) yields the desired results. $\qquad\square$

Proposition 6.2 (Sufficient condition for consistency). *Given the wave speeds $S_L < S_* < S_R$, Rankine-Hugoniot jump conditions are sufficient for consistency (i.e. satisfy Proposition 6.1) and the intermediate state variables are given by (for index $k = L, R$):*

$$h_{*k} = \frac{h_k(u_k - S_k)}{S_* - S_k}, \quad h_{*k} u_{*k} = \frac{h_k(u_k - S_k)S_*}{S_* - S_k}, \quad h_{*k} U_{*k} = \frac{h_k(u_k - S_k)U_K}{S_* - S_k}. \tag{6.29}$$

151

Proof. By applying Rankine-Hugoniot jump conditions across the three waves moving with speeds S_L, S_* and S_R yields [100, 161]:

$$\mathbf{f}_{*L} - \mathbf{f}_L = S_L(\mathbf{Q}_{*L} - \mathbf{Q}_L), \tag{6.30}$$

$$\mathbf{f}_{*R} - \mathbf{f}_{*L} = S_*(\mathbf{Q}_{*R} - \mathbf{Q}_{*L}), \tag{6.31}$$

$$\mathbf{f}_{*R} - \mathbf{f}_R = S_R(\mathbf{Q}_{*R} - \mathbf{Q}_R), \tag{6.32}$$

where $\mathbf{f}_{*L}$ and $\mathbf{f}_{*R}$ are fluxes corresponding to the intermediate state variables $\mathbf{f}_{*L}$ and $\mathbf{f}_{*R}$ respectively.

Rearranging terms in (6.30) and (6.32) yield expressions for the left and right intermediate fluxes:

$$\mathbf{f}_{*L} = \mathbf{f}_L + S_L(\mathbf{Q}_{*L} - \mathbf{Q}_L), \tag{6.33}$$

$$\mathbf{f}_{*R} = \mathbf{f}_R + S_R(\mathbf{Q}_{*R} - \mathbf{Q}_R). \tag{6.34}$$

Substituting the flux expressions (6.33) and (6.34) in the jump condition (6.31) and manipulating algebraically yields :

$$(S_* - S_L)\mathbf{Q}_{*L} + (S_R - S_*)\mathbf{Q}_{*R} = S_R\mathbf{Q}_R - S_L\mathbf{Q}_L + \mathbf{f}_L - \mathbf{f}_R, \tag{6.35}$$

which is identical to the consistency condition (6.23). Furthermore, expressions for the intermediate states $\mathbf{Q}_{*L}$ and $\mathbf{Q}_{*R}$ can be determined from the intermediate flux expressions (6.33)-(6.34). For the index $k = L, R$, the intermediate fluxes can be expressed as follows:

$$\mathbf{f}_{*k} = \mathbf{f}_k + S_k(\mathbf{Q}_{*k} - \mathbf{Q}_k), \quad k = L, R. \tag{6.36}$$

Thus, from the general flux expressions (6.36), the intermediate fluxes for the reactive shallow water model under consideration become:

$$h_{*k}u_{*k} = h_k u_k + S_k(h_{*k} - h_k), \tag{6.37}$$

$$h_{*k}u_{*k}^2 + \frac{1}{2}gh_{*k}^2 = h_k u_k^2 + \frac{1}{2}gh_k^2 + S_k(h_{*k}u_{*k} - h_k u_k), \tag{6.38}$$

$$h_{*k}u_{*k}U_{*k} = h_k u_k U_k + S_k(h_{*k}U_{*k} - h_k U_k). \tag{6.39}$$

Since the middle wave is a contact wave, the fluid velocity must be equal to speed of the contact wave. Consequently, substituting $u_{*k} = S_*$ into (6.37) - (6.39) and manipulating algebraically yields (6.29). □

Therefore, by employing Proposition 6.2, a consistent approximate HLLC flux is obtained as follows:

$$\mathbf{f}^{hllc} = \begin{cases} \mathbf{f}_L & \text{if } 0 \leq S_L, \\ \mathbf{f}_{*L} & \text{if } S_L \leq 0 \leq S_*, \\ \mathbf{f}_{*R} & \text{if } S_* \leq 0 \leq S_R, \\ \mathbf{f}_R & \text{if } 0 \geq S_R. \end{cases} \tag{6.40}$$

where $\mathbf{f}_L = \mathbf{f}\left(\mathbf{Q}_L\right)$, $\mathbf{f}_R = \mathbf{f}\left(\mathbf{Q}_R\right)$ and $\mathbf{f}_{*L}, \mathbf{f}_{*R}$ are defined as in (6.33) and (6.34).

Furthermore, if the middle contact wave S_* is ignored and arguments similar to those used to obtain the HLLC flux (6.40) are made, another approximate Godunov flux known as HLL flux (denoted $\mathbf{f}^{hll}$) is obtained. Given the wave speeds $S_L < S_R$, then the HLL flux states that:

$$\mathbf{f}^{hll} = \begin{cases} \mathbf{f}_L & \text{if } 0 \leq S_L, \\ \frac{S_R \mathbf{f}_L - S_L \mathbf{f}_R + S_L S_R \left(\mathbf{Q}_R - \mathbf{Q}_L\right)}{S_R - S_L} & \text{if } S_L \leq 0 \leq S_R, \\ \mathbf{f}_R & \text{if } 0 \geq S_R. \end{cases} \tag{6.41}$$

However, the presence of positive wave speeds in a Riemann problem can further simplify the fluxes. Let S_α be a positive wave in the Riemann problem, then by setting $S_L = -S_\alpha$ and $S_R = S_\alpha$, the HLL flux (6.41) simplifies to:

$$\mathbf{f}^{rus} = \frac{1}{2}\left(\mathbf{f}_L + \mathbf{f}_R\right) - \frac{1}{2}S_\alpha\left(\mathbf{Q}_R - \mathbf{Q}_L\right) \tag{6.42}$$

known as Rusanov flux [161]. Among several estimates for the positive wave speed, we consider:

$$S_\alpha = \max(|u_L| + c_L, |u_R| + c_R) \tag{6.43}$$

(named *RuDa* estimate) in subsequent discussions. The relationship between wave speed estimates and numerical fluxes will also be discussed in subsequent sections.

6.3.3 Source term discretization and well-balancing

In a frictionless steady flow the advective flux balances the topography term in the momentum equation. Specifically, under frictionless quiescent flow conditions (i.e. $u = 0$ and $h + \eta = H = constant$) the momentum equation states that:

$$\frac{d}{dx}\left(\frac{1}{2}gh^2\right) = -gh\frac{\partial \eta}{\partial x} \tag{6.44}$$

Thus, integrating (6.44) over the control volume $(x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}})$ yields the discretized topography term, i.e.:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} -gh\frac{\partial \eta}{\partial x}dx = \frac{1}{2}g\left(\left(h_{i+\frac{1}{2}}^-\right)^2 - \left(h_{i-\frac{1}{2}}^+\right)^2\right). \tag{6.45}$$

Moreover, the integral of the friction term is approximated as follows [85]:

$$\int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \frac{gC^2|u|u}{h^{1/3}}dx \approx gC^2\Delta x\left(\frac{2h_i}{h_i + \max(h_i^2, 0)}\right)^{\frac{7}{3}}|u_i|u_i. \tag{6.46}$$

The chemical diffusion derivatives at the control volume interfaces are discretized using forward differencing, thus:

$$\left(\frac{\partial \mathbf{U}}{\partial x}\right)_{i+\frac{1}{2}} \approx \frac{\mathbf{U}_{i+1} - \mathbf{U}_i}{\Delta x} \quad \text{and} \quad \left(\frac{\partial \mathbf{U}}{\partial x}\right)_{i-\frac{1}{2}} \approx \frac{\mathbf{U}_i - \mathbf{U}_{i-1}}{\Delta x}. \tag{6.47}$$

Therefore, from (6.45), (6.46) and (6.47) the discretized source term in (6.6) is given by:

$$\mathbf{S}_i = \begin{bmatrix} 0 \\[2mm] \frac{1}{2}g\left(\left(h_{i+\frac{1}{2}}^-\right)^2 - \left(h_{i-\frac{1}{2}}^+\right)^2\right) - gC^2\Delta x\left(\frac{2h_i}{h_i + \max(h_i^2, 0)}\right)^{\frac{7}{3}}|u_i|u_i \\[4mm] \frac{\Gamma_{i-\frac{1}{2}}}{\Delta x}\mathbf{U}_{i-1} - \left(\frac{\Gamma_{i-\frac{1}{2}}}{\Delta x} + \frac{\Gamma_{i+\frac{1}{2}}}{\Delta x}\right)\mathbf{U}_i + \frac{\Gamma_{i+\frac{1}{2}}}{\Delta x}\mathbf{U}_{i+1} + \Delta x\mathbf{S}_h(\mathbf{U}_i) \end{bmatrix} \tag{6.48}$$

Proposition 6.3 (C-property preservation)**.** *The semi-discrete scheme* (6.6) *(with numerical flux defined as* $\mathbf{f}_{i+\frac{1}{2}} = \mathbf{f}^{rus}(\mathbf{Q}_{i+\frac{1}{2}}^+, \mathbf{Q}_{i+\frac{1}{2}}^-)$ *or* $\mathbf{f}_{i+\frac{1}{2}} = \mathbf{f}^{hllc}(\mathbf{Q}_{i+\frac{1}{2}}^+, \mathbf{Q}_{i+\frac{1}{2}}^-)$ *and with interface values* $\mathbf{Q}_{i+\frac{1}{2}}^+, \mathbf{Q}_{i+\frac{1}{2}}^-$ *defined in* (6.10)*) preserves the C-property.*

Proof. A well-balanced or a C-property preserving scheme is one that preserves the flux-topography balance under frictionless quiescent flow conditions. Let $L_i\big(\mathbf{Q}_i(t)\big)$ denote the right hand side of the momentum component of the semi-discrete scheme (6.6), then:

$$L_i\big(\mathbf{Q}_i(t)\big) = -\left(f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}\right) + \frac{1}{2}g\left(\left(h_{i+\frac{1}{2}}^-\right)^2 - \left(h_{i-\frac{1}{2}}^+\right)^2\right)$$
$$- gC^2\Delta x\left(\frac{2h_i}{h_i + \max(h_i^2, 0)}\right)^{\frac{7}{3}}|u_i|u_i, \tag{6.49}$$

154

where $f_{i+\frac{1}{2}} = f^{rus}(\mathbf{Q}^+_{i+\frac{1}{2}}, \mathbf{Q}^-_{i+\frac{1}{2}})$ or $f_{i+\frac{1}{2}} = f^{hllc}(\mathbf{Q}^+_{i+\frac{1}{2}}, \mathbf{Q}^-_{i+\frac{1}{2}})$ and the natural flux of momentum is given by:

$$f(\mathbf{Q}) = hu^2 + \frac{1}{2}gh^2. \tag{6.50}$$

The manning coefficient C in (6.49) vanishes in a frictionless flow, thus, the last term in (6.49) vanishes, yielding:

$$L_i\big(\mathbf{Q}_i(t)\big) = -\left(f_{i+\frac{1}{2}} - f_{i-\frac{1}{2}}\right) + \frac{1}{2}g\left(\big(h^-_{i+\frac{1}{2}}\big)^2 - \big(h^+_{i-\frac{1}{2}}\big)^2\right). \tag{6.51}$$

Thus, to prove Proposition 6.3 we show that (6.51) vanishes under quiescent flow conditions. However, under the quiescent conditions, $u^L_{i+\frac{1}{2}} = u^R_{i+\frac{1}{2}} = u^L_{i-\frac{1}{2}} = u^R_{i-\frac{1}{2}} = u_i = 0$, evaluating and exploiting the continuity/consistency properties of the numerical fluxes yields:

$$f_{i+\frac{1}{2}} = f^{rus}(\mathbf{Q}^+_{i+\frac{1}{2}}, \mathbf{Q}^-_{i+\frac{1}{2}}) = \frac{1}{2}g\big(h^-_{i+\frac{1}{2}}\big)^2 \quad \text{and} \quad f_{i+\frac{1}{2}} = f^{hllc}(\mathbf{Q}^+_{i+\frac{1}{2}}, \mathbf{Q}^-_{i+\frac{1}{2}}) = \frac{1}{2}g\big(h^-_{i+\frac{1}{2}}\big)^2. \tag{6.52}$$

Similarly,

$$f_{i-\frac{1}{2}} = f^{rus}(\mathbf{Q}^+_{i-\frac{1}{2}}, \mathbf{Q}^+_{i-\frac{1}{2}}) = \frac{1}{2}g\big(h^+_{i-\frac{1}{2}}\big)^2 \quad \text{and} \quad f_{i-\frac{1}{2}} = f^{hllc}(\mathbf{Q}^+_{i-\frac{1}{2}}, \mathbf{Q}^-_{i-\frac{1}{2}}) = \frac{1}{2}g\big(h^+_{i-\frac{1}{2}}\big)^2. \tag{6.53}$$

Hence, substituting (6.52) and (6.53) into (6.51) completes the proof. $\qquad\square$

6.3.4 Temporal discretization

Strong stability preserving (SSP) time marching schemes (also known as *Total Variation Diminishing* or *monotonicity preserving* schemes have been shown to be robust in simulating stiff problems in Astrophysics [83, 123, 63]. Since the current model has characteristics similar to the Astrophysical models, we present strong stability preserving schemes here, to march the semi-discrete scheme (6.6).

Further, due to the presence of stiff terms (e.g. chemical reactions and diffusion) in the current model, the semi-discrete scheme (6.6) is first split into the flow part (non-stiff part) and the chemical transport/reaction part (the stiff part). Let $\mathbf{W} = (h, uh)^{Tr}$, then

the semi-discrete scheme (6.6) is split as follows:

$$\frac{d}{dt}\mathbf{W}_i(t) = \mathbf{L}_W(\mathbf{W}_i(t)) \tag{6.54}$$

$$\frac{d}{dt}\mathbf{U}_i(t) = \mathbf{L}_I(\mathbf{U}_i(t)) + \mathbf{L}_E(\mathbf{U}_i(t)), \tag{6.55}$$

where $\mathbf{Q} = (\mathbf{W}, \mathbf{U})^{Tr}$. The slope functions $\mathbf{L}_I(\mathbf{Q}_i(t)), \mathbf{L}_E(\mathbf{Q}_i(t))$ are respectively the stiff and non-stiff parts of the chemical transport components and $\mathbf{L}_W(\mathbf{Q}_i(t))$ is the flow components of the slope function:

$$\mathbf{L}(\mathbf{Q}_i(t)) = -\frac{1}{\Delta x}\left(\mathbf{f}_{i+\frac{1}{2}} - \mathbf{f}_{i-\frac{1}{2}}\right) + \mathbf{S}_i. \tag{6.56}$$

Furthermore, since the flow component of the reactive shallow water model is non-stiff, we apply the third order, two-stage strong stability preserving scheme of [74] to the semi-discrete scheme (6.54). Consequently, the scheme is stated as follows:

$$\mathbf{W}_i^1 = \mathbf{W}_i^n + \Delta t \mathbf{L}_W(\mathbf{W}_i^n)$$

$$\mathbf{W}_i^2 = \frac{3}{4}\mathbf{W}_i^n + \frac{1}{4}\mathbf{W}_i^1 + \frac{1}{4} + \Delta t \mathbf{L}_W(\mathbf{W}_i^1)$$

$$\mathbf{W}_i^{n+1} = \frac{1}{3}\mathbf{W}_i^n + \frac{2}{3}\mathbf{W}_i^2 + \frac{1}{4} + \Delta t \mathbf{L}_W(\mathbf{W}_i^2) \tag{6.57}$$

where Δt is the time step, $\mathbf{W}_i^n = \mathbf{W}_i(t^n)$, $\mathbf{W}_i^{n+1} = \mathbf{W}_i(t^n + \Delta t)$ and t^n denotes the nth time. Moreover, since the chemical transport components of the model contain stiff and non-stiff terms, we apply the Implicit-Explicit (IMEX) Runge-Kutta scheme of [63, 83, 123] (that is second order consistent, $A(\frac{\pi}{2})$-stable and strong stability preserving) on the semi-discrete scheme (6.55). Consequently, with the parameter $\gamma = 1 - \frac{1}{\sqrt{2}}$, the SSP-IMEX scheme is expressed as follows:

$$\mathbf{U}_i^1 = \mathbf{U}_i^n + \Delta t \mathbf{L}_I(\mathbf{U}_i^1)$$

$$\mathbf{U}_i^2 = \mathbf{U}_i^n + \Delta t \mathbf{L}_E(\mathbf{U}_i^1) + (1 - 2\gamma)\Delta t \mathbf{L}_I(\mathbf{U}_i^1) + \gamma \Delta t \mathbf{L}_I(\mathbf{U}_i^2)$$

$$\mathbf{U}_i^{n+1} = \mathbf{U}_i^n + \frac{1}{2}\Delta t \left(\mathbf{L}_E(\mathbf{U}_i^1) + \mathbf{L}_E(\mathbf{U}_i^2) + \mathbf{L}_I(\mathbf{U}_i^1) + \mathbf{L}_I(\mathbf{U}_i^2)\right). \tag{6.58}$$

Since Scheme (6.58) is second order and Scheme (6.57) is third order, the final time-marching scheme is second order. However, the complete scheme (including both spatial and temporal parts) is second order if the *Minmod* limiter is used in the spatial discretization, but reduces to first order if piecewise constants are used instead. The accuracy of the numerical schemes will be verified in Section 6.4.

6.4 Numerical experiments

Numerical experiments are conducted in this Section 6.4 to verify the numerical scheme that has been presented in previous sections. The nonlinear equations that involve root-finding are solved using Newton-Raphson's method. Experiments are conducted under the CFL condition:

$$\max(\sup |S_R|, \ \sup |S_L|)\frac{\Delta t}{\Delta x} \leq 1. \tag{6.59}$$

The accuracy of the scheme will be verified with two benchmark problems that involve flat and non-flat bottom topographies (specifically, dam break and tidal flow problems). Compatibility with the stoichiometric decoupling method will be assessed using the acid drainage and neutralization problem.

6.4.1 Shallow water flow experiments

Firstly, the accuracy of the scheme is assessed using a flat bottom topography dam break problem. The bench mark is composed of two water columns of different depths that are separated by a dam. Denoting the water depths of the two water columns by h_L and h_R (with $h_L > h_R$), if the dam is positioned at the origin, the initial dam break problem states that [38]:

$$hu(x,0) = 0 \ \text{ and } \ h(x,0) = \begin{cases} h_L & \text{for } x < 0, \\ h_R & \text{for } x > 0. \end{cases}, \tag{6.60}$$

where $h_L > h_R \geq 0$. If the right column depth h_R is vanishes, the benchmark is referred to as *Dry bed dam break* problem, otherwise, it is referred to as *Wet bed dam break* problem [38, 112]. Initially (i.e. at time $t = 0$), the dam is broken and the water upstream flows downstream. The analytical profile of the dry dam problem is given by [112]:

$$h(x,t) = \begin{cases} h_L & \text{for } \frac{x}{t} \leq 0, \\ \frac{4}{9g}\left(\sqrt{gh_L} - \frac{x}{2t}\right)^2 & \text{for } -\sqrt{gh_L} < \frac{x}{t} < 2\sqrt{gh_L}, \\ 0 & \text{for } \frac{x}{t} \geq 2\sqrt{gh_L}. \end{cases} \tag{6.61}$$

$$u(x,t) = \begin{cases} 0 & \text{for } \frac{x}{t} \leq 0, \\ \frac{2}{3}\left(\sqrt{gh_L} + \frac{x}{t}\right) & \text{for } -\sqrt{gh_L} < \frac{x}{t} < 2\sqrt{gh_L}, \\ 0 & \text{for } \frac{x}{t} \geq 2\sqrt{gh_L}. \end{cases} \qquad (6.62)$$

The numerical challenge associated with the dry bed dam break problem is how to capture (without oscillations) the singularity that occurs at the transition point between the wet and dry zone [38]. The current Scheme (6.57) was applied to the dry bed dam break problem, using $h_L = 2$, $T = 0.1$, $C = 0$, $g = 9.8$ and cells $N_x = 750$ cells in the domain $[-5, 5]$. Figures 6.1a and 6.1b show results for the upwind and Minmod schemes. It is clear from the results that the current schemes are non-oscillatory even at the points of transition.

Secondly, the solution to the wet bed dam break problem consists of a shock wave propagating downstream and a rarefaction propagating upstream [38, 112]. The challenge for numerical schemes therefore, is to capture the shocks and transition points without oscillations. Scheme (6.57) was applied to the dry bed dam break problem, using $h_L = 2$, $h_R = 0.7$, $T = 0.51$, $C = 0$ and $N_x = 750$ in the same domain. The reference solution (representing the analytical solution) was computed with a fine grid (7000 cells using the upwind scheme). Figures 6.1c and 6.1d show results for the upwind and Minmod schemes in the wet bed dam break problem. It can be observed that the rarefaction and shock have been accurately captured without oscillations at the points of transition.

Thirdly, the schemes are tested on a tidal wave flow problem that involves a non-flat bottom topography. The initial condition is given by [101]:

$$uh(x,t) = 0 \quad \text{and} \quad h(x,t) = 60.5 - \eta(x), \qquad (6.63)$$

where $\eta(x) = 10 + \frac{40x}{L_x} + 10\sin\left(\pi\left(\frac{40x}{L_x} - \frac{1}{2}\right)\right)$ and $L_x = 14000$. With boundary conditions given by:

$$h(0,t) = 64.5 - 4\sin\left(\pi\left(\frac{4t}{86,400} + \frac{1}{2}\right)\right) \quad \text{and} \quad hu(L_x,t) = 0, \qquad (6.64)$$

the asymptotic solution is given by [101]:

$$h(x,t) = 64.5 - \eta(x) - 4\sin\left(\pi\left(\frac{4t}{86,400} + \frac{1}{2}\right)\right) \tag{6.65}$$

$$uh(x,t) = \frac{(x - L_x)\pi}{5400}\cos\left(\pi\left(\frac{4t}{86,400} + \frac{1}{2}\right)\right) \tag{6.66}$$

Simulation studies were conducted on the tidal wave problem using Scheme (6.57) in the domain $[0, 14000]$ using $N_x = 150$ grid cells for a period $T = 7552.13$. Figures 6.1e and 6.1f show results for the upwind and Minmod discretizations. The depth profiles are free from oscillations and mimic the analytical solution well.

Furthermore, the accuracy of the schemes did not vary significantly across numerical fluxes, if chosen wave speed estimates are appropriate for the chosen numerical flux. Figure 6.2 shows that Davies' estimates (i.e. (6.14) and (6.15)), Roe's estimates (i.e. (6.16) and (6.17)) and RuDa estimates (i.e. (6.43)) are compatible with the Rusanov flux while depth estimates (i.e. (6.20) and (6.21)) are compatible with the HLLC flux. Further, the error profiles of the schemes have been computed using the dry bed dam break problem, Figures 6.3e and 6.3f depict the results. The error profiles are decreasing monotonically with increasing grid cells. Further, Table 6.1 shows that the schemes converge in the expected orders.

6.4.2 Chemical reaction experiments

In this section, experiments are conducted to verify the suitability of the schemes for simulating reactions that occur within shallow water flow problems. The efficiency of the schemes are verified based on stiffness-resolving ability and compatibility with the stoichiometric decoupling method.

In Section 5.5 of Chapter 5, stiff problems (involving Darcy flow and reactive transport) have been used to assess the accuracy of the schemes presented in that chapter. Here, the strong stability preserving IMEX Scheme (6.58) is tested using the stiff System (5.43) and the 1D nonlinear transport bench mark (5.50) of that section. Figure 6.3 shows the concentration profiles in both Darcy and shallow water flow environments. It clear that the results are accurate and non-oscillatory. Particularly, Figures 6.3a and 6.3c show that the SSP-IMEX Scheme (6.58) can resolve stiffness very well. Figures 6.3b

Convergence test

Benchmark	Spatial steps (N_x)	Upwind		Minmod	
		Error	Order	Error	Order
	80	0.3472	-	0.3057	-
Dam	160	0.1640	1.0823	0.0719	2.0885
break	320	0.0860	0.9315	0.0185	1.9589
	640	0.0354	1.2809	0.0046	2.0082
	1280	0.0176	1.0083	0.0011	2.0646
	50	0.2299	-	0.1472	-
Tidal	100	0.0994	1.2100	0.0306	2.2667
wave	200	0.0464	1.0993	0.0075	2.0290
flow	400	0.0231	1.0065	0.0019	1.9813
	800	0.0114	1.0191	0.0005	1.9264

Table 6.1: Errors (L_∞) and orders of the *Upwind* and *Minmod* schemes, computed using the tidal wave flow and dry bed dam break problem.

and 6.3d display shocks that developed as a consequence of the nonlinear hyperbolicity of the reactive shallow water model. Although, the initial concentration profiles were smooth, the final concentration profiles involved shocks and transition points that have been captured by the proposed scheme.

In Sections 4.5 and 5.5, the acid drainage and neutralization (in a Darcy flow environment) bench mark has been used to test numerical schemes on their compatibility with the stoichiometric decoupling method. Here, we couple the shallow water flow model with the original large chemical transport model (5.61) (resulting model has seven degrees of freedom) and with the stoichiometrically reduced model 5.62 (resulting model has three degrees of freedom) and use to assess the schemes proposed. The large model that uses Gauss-Jacobi decoupling is the reference model for measuring the accuracy of the stoichiometrically reduced model. Figure 6.4 shows some results when Schemes (6.57) and (6.58) were applied to the resulting models (i.e. reduced model with three degrees of freedom and large model with seven degrees of freedom). The velocity profiles 6.4a and

6.4b show that the proposed scheme is accurate. The concentration profiles in Figures 6.4c-6.4f show that the stoichiometrically reduced model is accurate in time and space. Further, the model reduction (decoupling) error profiles are given in Table 6.2 and in Figures 6.5a-6.5b. The error (values are 10^{-8} smaller than shown) oscillates in the L_∞ norm but has an upper bound of $9.7E - 8$. However, the model reduction error diverges in the L_2 norm (the error values are 10^{-7} smaller).

Decoupling errors				
Grid cells (N_x)	Minmod scheme		Upwind scheme	
	$\|\cdot\|_\infty$	$\|\cdot\|_2$	$\|\cdot\|_\infty$	$\|\cdot\|_2$
100	9.5109	5.0283	9.5473	5.0475
200	9.2742	6.9393	9.2926	6.9538
300	9.5180	8.7279	9.5270	8.7386
400	9.3815	9.9360	9.3847	9.9430
500	9.5026	11.2557	9.5026	11.2607
600	9.5826	12.4368	9.5807	12.4404
700	9.4939	13.3100	9.4908	13.3122
800	9.5545	14.3221	9.5505	14.3234
900	9.6015	15.2676	9.5969	15.2682
1000	9.5372	15.9862	9.5371	16.1558

Table 6.2: Model decoupling errors of Scheme (6.57), computed across norms and spatial steps. The error values in this table are 10^{-8} smaller than shown.

Further, experiments on simulation cost have also been conducted using the large and reduced shallow water model. The CPU time and CPU time differences have been measured for both models and the results are displayed in Figures 6.5c-6.5f. It is clear from the results that it is less expensive to simulate the stoichiometrically reduced model with three degrees of freedom than to simulate the original large model with seven degrees of freedom. Since the errors in Table 6.2 and in Figures 6.5a-6.5b are negligible, it can be concluded that, the combined stoichiometric decoupling method and the proposed schemes are efficient for simulating reactive shallow water flows.

6.5 Chapter summary

The presence of shocks, high degrees of freedom, stiffness and non-flat bottom topographies in reactive shallow water flow problems make numerical simulation of such problems challenging. The goal of this chapter includes the construction of a high resolution well-balanced scheme that is compatible with the stoichiometric decoupling method presented in Chapter 3.

In Section 6.1, a review of research conducted in the broad area of reactive flows has been presented. The review reveals that shallow water equations have a broad area of application and therefore, a lot of studies have been conducted on shallow water equation. The review also outlined some simulation studies conducted on conservation laws, and some techniques for developing well-balanced numerical schemes for balanced laws.

In Section 6.2, a model for flows with reactive transport processes that satisfy the shallow water principle has been presented, and some of the simulation challenges have been outlined.

In Section 6.3, numerical schemes for the reactive shallow water flow model has been constructed. Numerical fluxes and wave speed estimates have been provided and the schemes have been shown to preserve C-property. Strong stability preserving time-marching schemes have been used to discretize the semi-discrete scheme.

In Section 6.4, numerical experiments have been conducted to verify the schemes. The results show that the proposed numerical schemes are shock-capturing, stiffness-resolving, well-balanced and are compatible with the stoichiometric method.

Therefore, the combination of the present schemes and the stoichiometric decoupling method provides an efficient tool for simulating reactive shallow water flow problems.
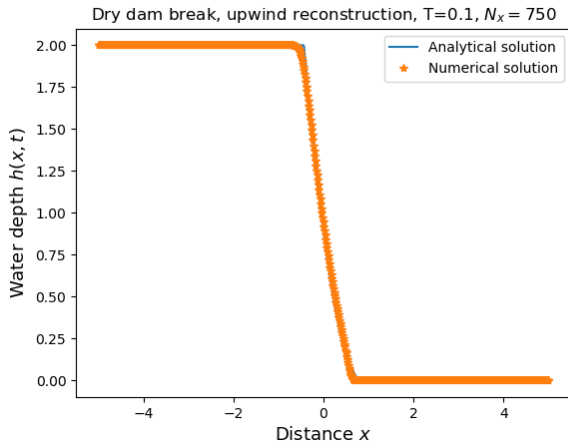
(a) Water depth/Dry/Upwind.
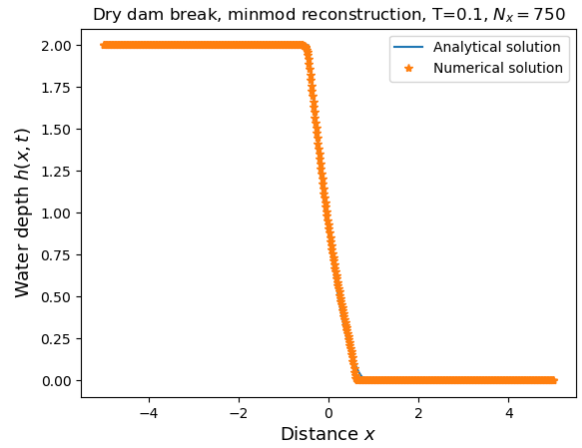
(b) Water depth/Dry/Minmod.
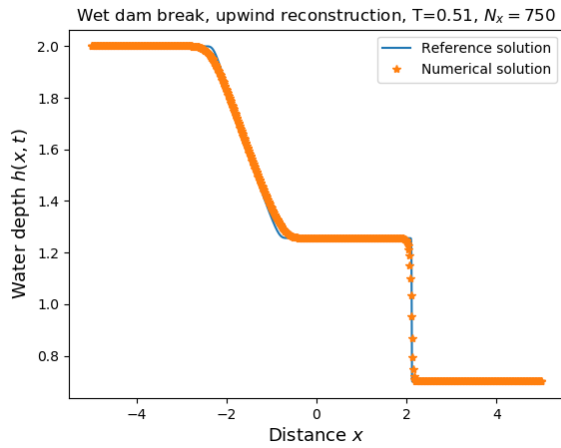
(c) Water depth/Wet/Upwind.
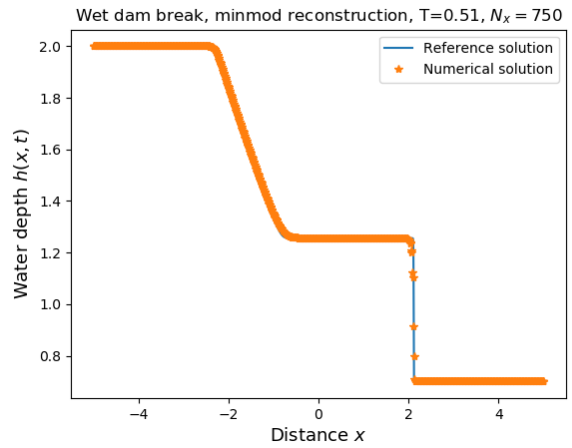
(d) Water depth/Wet/Minmod.
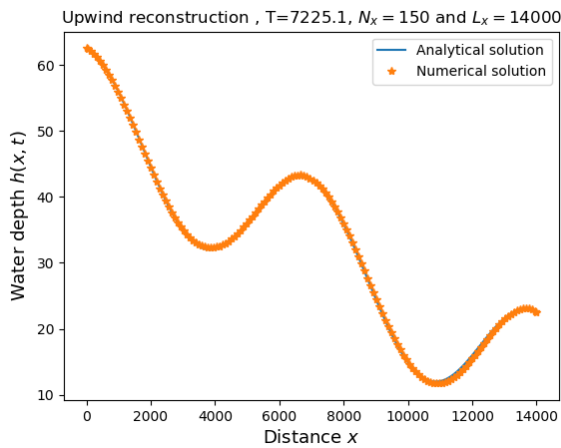
(e) Water depth/Tidal/Upwind
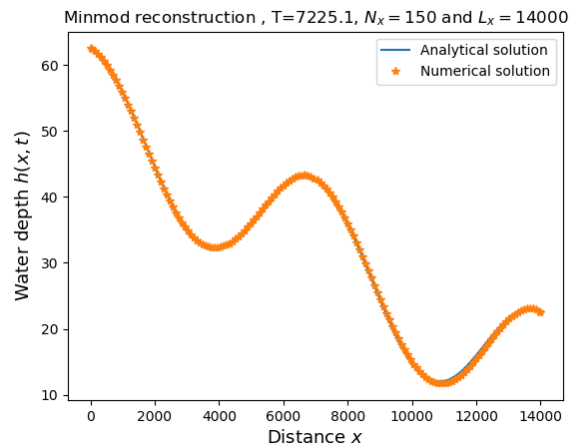
(f) Water depth/Wet/Minmod.

Figure 6.1: Numerical and analytical solutions for the dry and wet dam break problems, where $T = 0.1$, and $N_x = 750$. The left and right columns display results for dry and wet dam break problems, respectively.
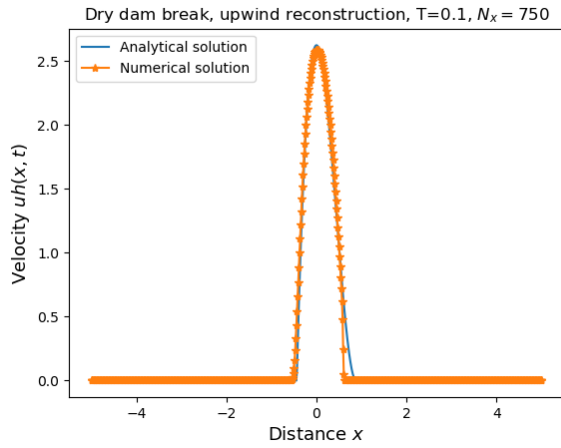
(a) Rusanov flux/Roe estimate.
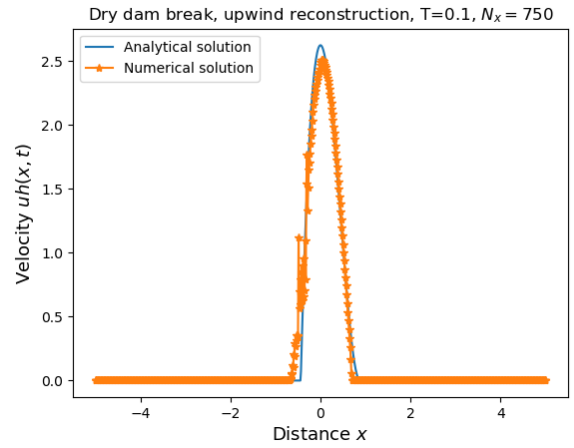
(b) HLLC flux/Roe estimate.
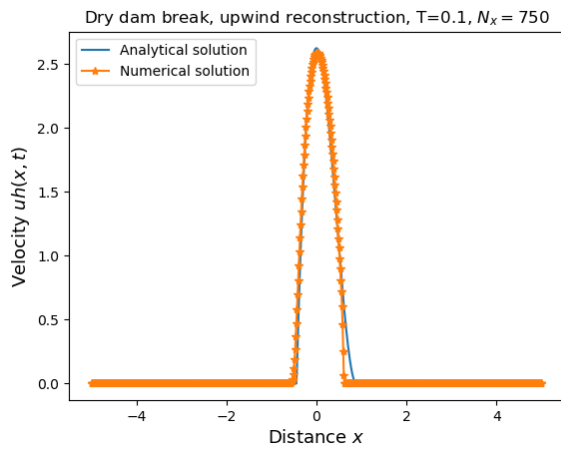
(c) Rusanov flux/Davis estimate.
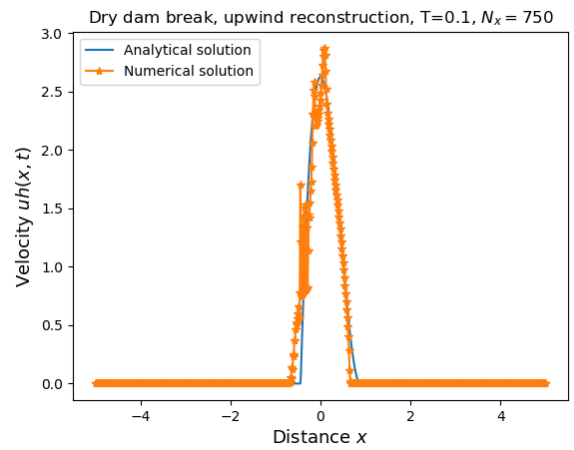
(d) HLLC flux/Davis estimate.
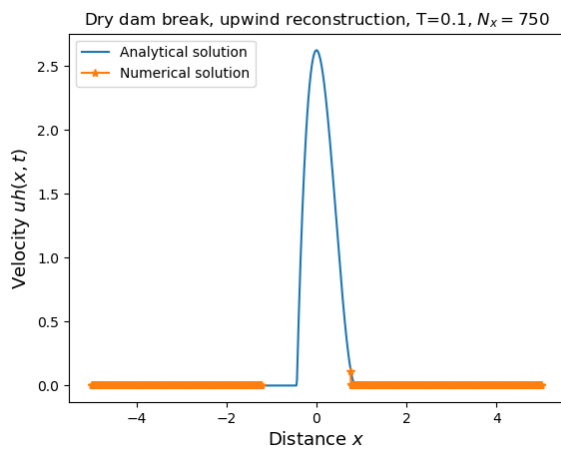
(e) Rusanov flux/Depth estimate.
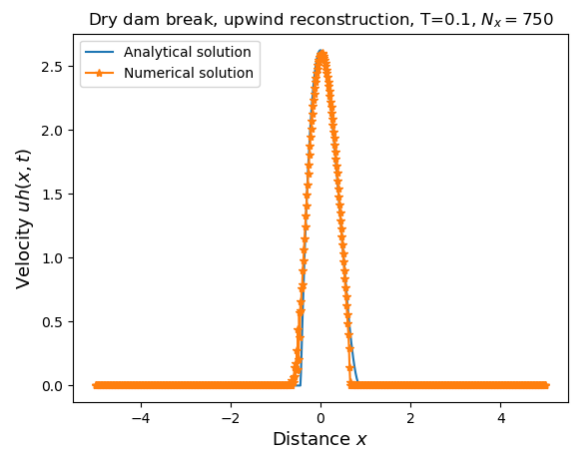
(f) HLLC flux/Depth estimate.

Figure 6.2: Numerical and analytical solutions for the dry dam break problem, computed with different wave estimates and fluxes. The left and right columns display results for Rusanov and HLLC fluxes, respectively.
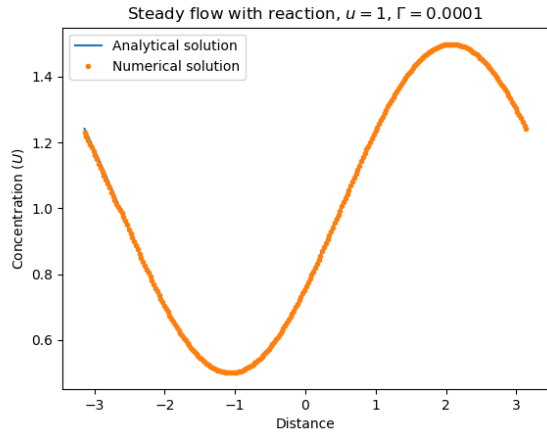
(a) Darcy flow/single reaction.
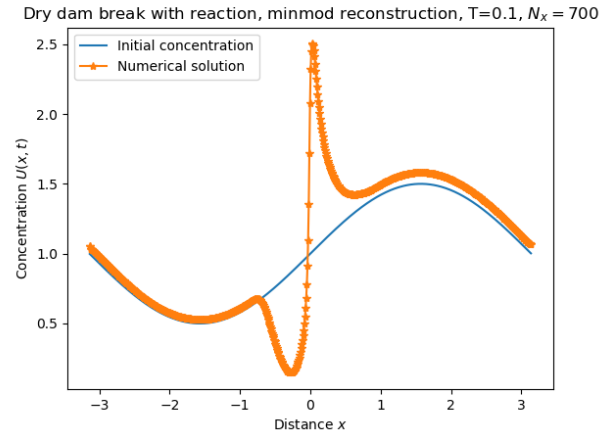
(b) Shallow water flow/single reaction.
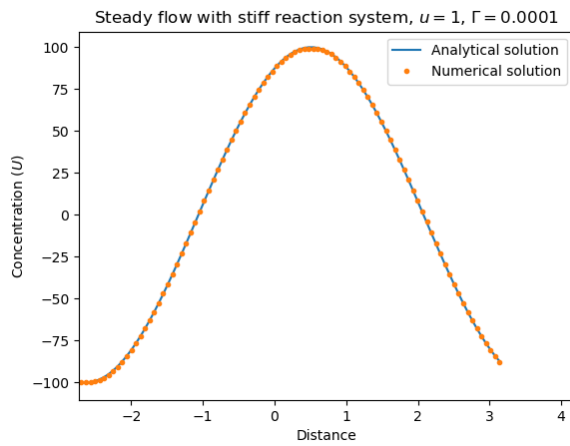
(c) Darcy flow/stiff system.
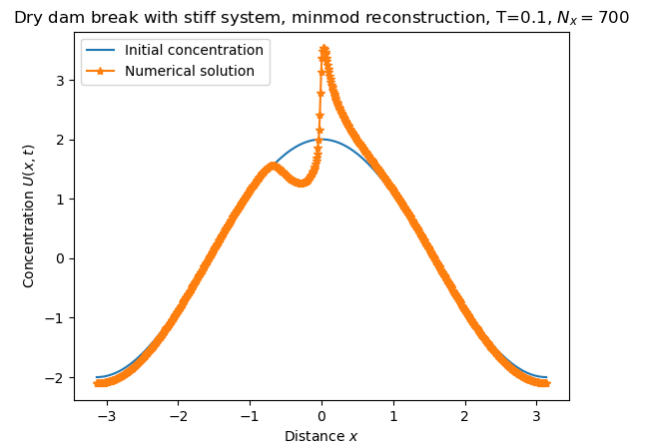
(d) Shallow water flow/stiff system.
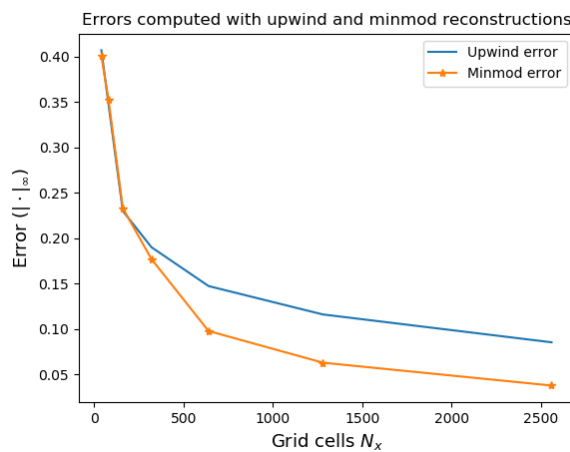
(e) Rusanov flux/L_∞ errors.
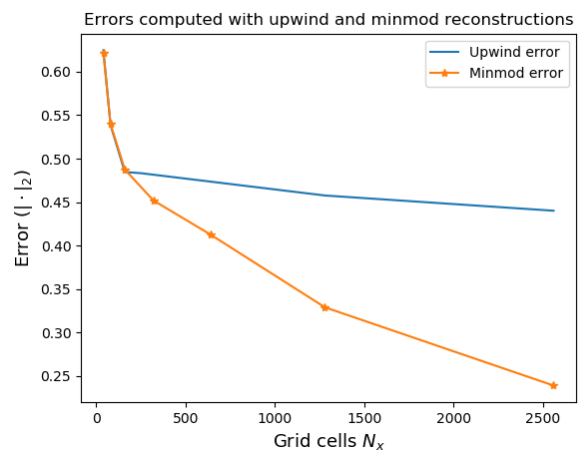
(f) HLLC flux/L_2 errors.

Figure 6.3: Concentration profiles in single and stiff systems of reactions that occur in shallow water flow and Darcy flow environments. The errors (max and L_2) of the numerical scheme were computed using the analytical solution of the dry dam problem.
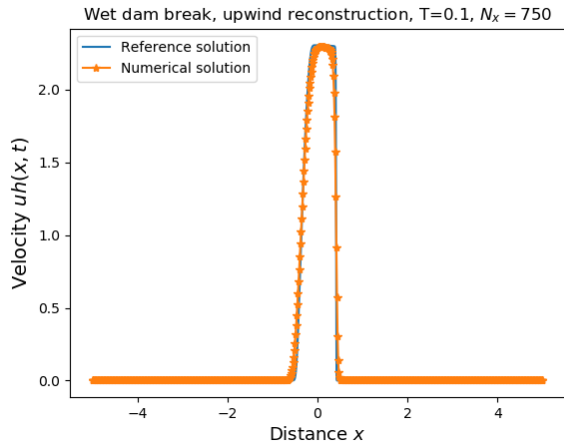
(a) Upwind/flow rate.
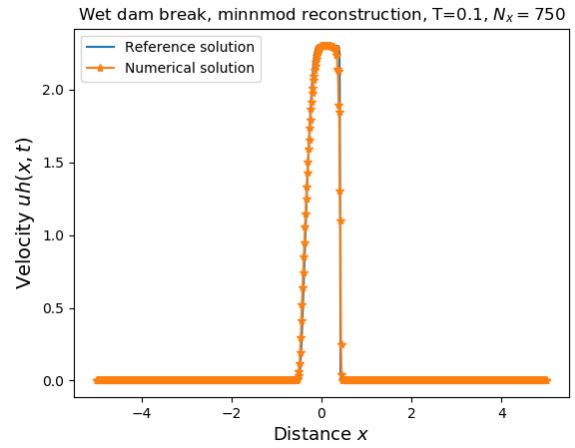
(b) Minmod/flow rate.
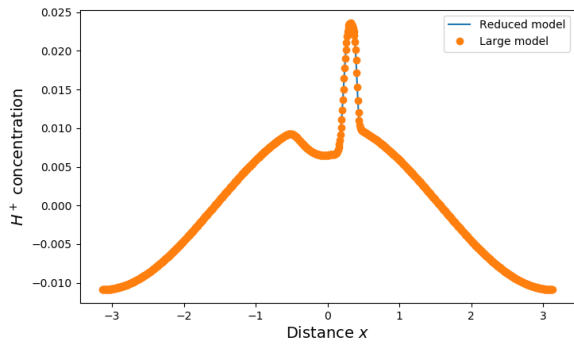
(c) Upwind/concentration/distance.
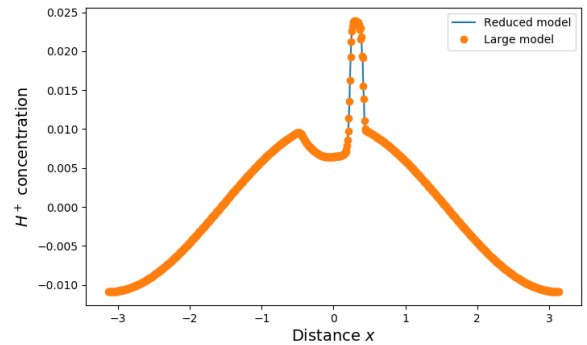
(d) Minmod/concentration/distance.
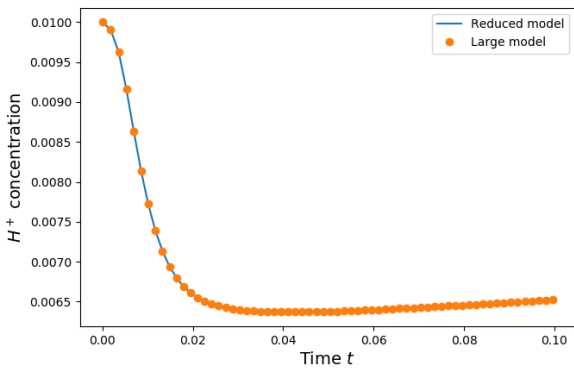
(e) Upwind/concentration/time.
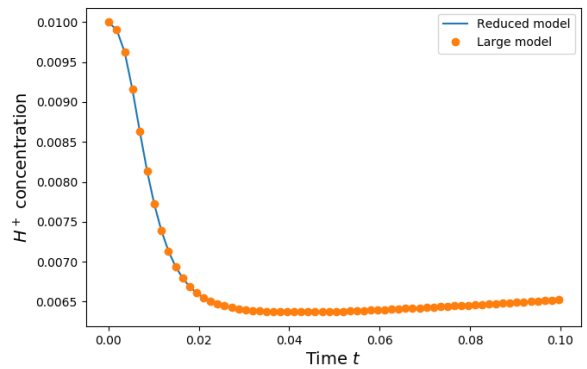
(f) Minmod/concentration/time.

Figure 6.4: Flow rates and concentration profiles in a reactive wet dam break problem involving acid neutralization. The profiles were obtained by using both upwind and minmod reconstructions.
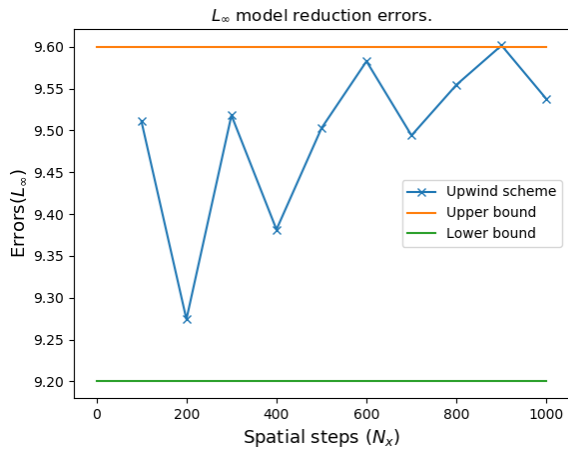
(a) Model reduction error/Upwind.
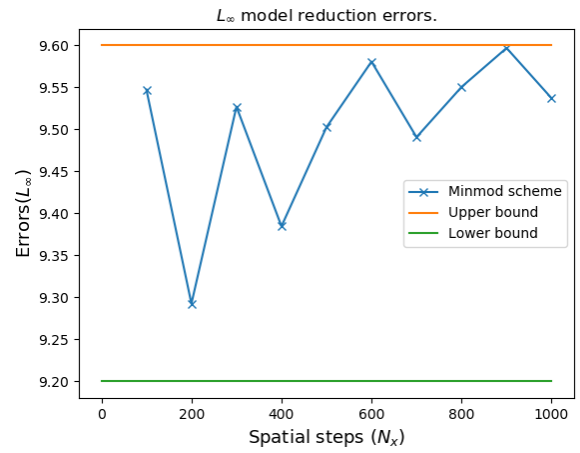
(b) Model reduction error/Minmod.
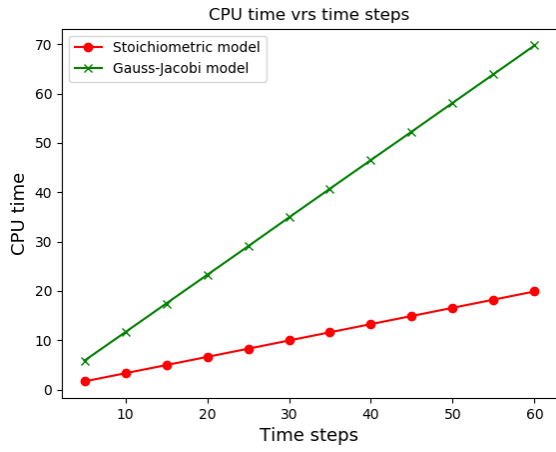
(c) CPU time/Upwind.
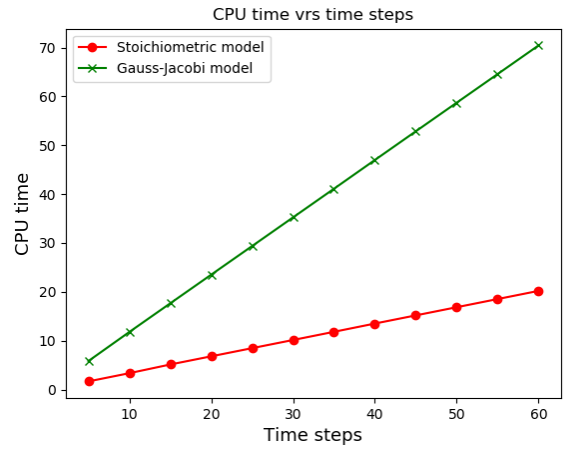
(d) CPU time/Minmod.
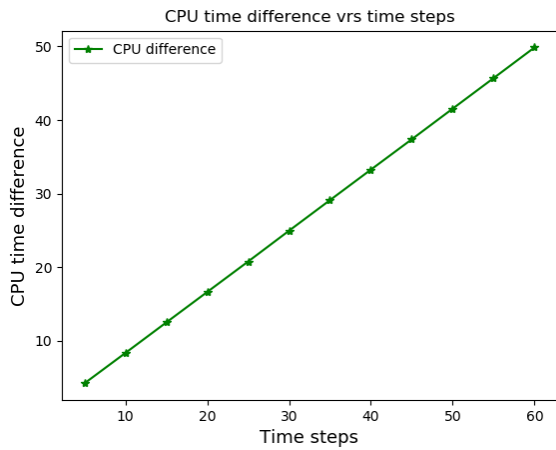
(e) CPU time difference/Upwind.
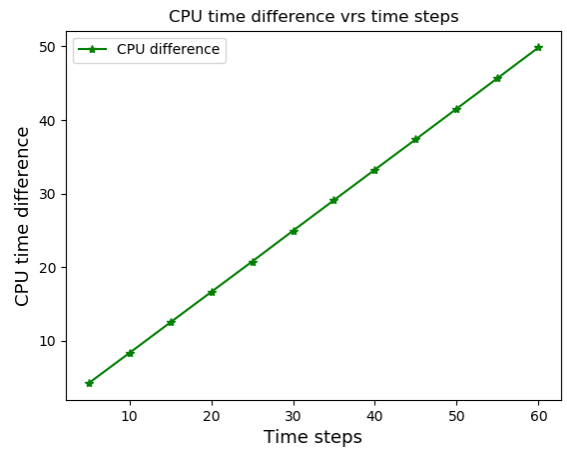
(f) CPU time difference/Minmod.

Figure 6.5: Model reduction errors, CPU time and CPU time differences of the Stoichiometrically decoupled model (reduced model) and Gauss Jacobi decoupled model (large model).

167

Chapter 7

Conclusion and future studies

Numerical simulation of chemical phenomena has been discussed in this thesis. Due to the presence of many species in chemical systems, that participate in a complex network of chemical reactions, mathematical modelling often result in models that have high degrees of freedom and possess certain properties (i.e. stiffness, positivity, conservativity, dissipation/dispersion properties, C-property, shocks etc) that must be preserved or captured in simulations. While the high degrees of freedom makes numerical simulation expensive, the properties of resulting models pose as constraints for some numerical schemes. The challenge therefore, is to develop efficient procedures (i.e. accurate and less expensive procedures) for simulating chemical phenomena. The goal of this thesis was to develop efficient solution procedures by combining model decoupling/reduction methods with compatible accurate numerical schemes, for the purpose of simulating the kinetics of chemical reactions.

Moreover, the simulation challenge has been tackled in three parts which include simulation of: ODEs (that model well-mixed chemical systems that depend on time only) parabolic PDEs (that models reactive transport systems or poorly-mixed chemical systems that depend on space and time) and hyperbolic PDEs (that models flow and reactive transport processes). In Chapter 2, the reactive shallow water model has been derived and analysed to establish nonlinear effects such as shocks, rarefactions and discontinuities. In Chapter 3, modelling and simulation of chemical kinetics (i.e. the ODE part) has been discussed. Model decoupling methods and compatible accurate numerical schemes (that satisfy stiffness, positivity and conservativity constraints) have been dis-

cussed. Numerical simulation of the parabolic part has been discussed in Chapters 4-5. Numerical schemes for the parabolic part have been shown to resolve stiffness, preserve dissipation/dispersion properties and are compatible with the model decoupling method. Shock-capturing, well-balanced schemes have been discussed in Chapter 6, for simulating reactive shallow water problems.

While the presented numerical schemes can resolve stiffness and other constraints (thus, are accurate), the stoichiometric decoupling method can significantly reduce the degrees of freedom. A combination of the schemes and the stoichiometric method, therefore, provides an efficient tool for simulating chemical phenomena.

In further studies, we will explore the stoichiometric method to establish its limitations. *Throughout the study, the solution procedures have not converged to the expected orders in the L_2 norm, we will investigate this unsual behaviour in future studies.* Other model decoupling/reduction algorithms will be developed based on graph theory for simulating chemical phenomena. Further studies will also go into the construction of more accurate well-balanced high resolution schemes for reactive shallow water flow problems.

However, the main contribution of this thesis include the discussion on the stoichiometric decoupling method, discussion involving dissipation-dispersion properties of numerical schemes and the construction of well-balanced high resolution schemes for reactive flow problems.

Bibliography

[1] D. L. Allara and D. Edelson. A computational modeling study of the low-temperature pyrolysis of n-alkanes; mechanisms of propane, n-butane, and n-pentane pyrolyses. *International Journal of Chemical Kinetics*, 7(4):479–507, 1975.

[2] E. A. Amikiya and M. K. Banda. Modelling and simulation of reactive transport phenomena. *Journal of Computational Science*, 28:155–167, 2018.

[3] E. A. Amikiya and M. K. Banda. A stoichiometric method for reducing simulation cost of chemical kinetic models. *Computers & Chemical Engineering*, 112:274–291, 2018.

[4] C. A. J. Appelo, E. Verweij, and H. Schäfer. A hydrogeochemical transport model for an oxidation experiment with pyrite/calcite/exchangers/organic matter containing sand. *Applied Geochemistry*, 13(2):257–268, 1998.

[5] R. Aris and N. R. Amundson. An analysis of chemical reactor stability and control-I: The possibility of local control, with perfect or imperfect control mechanisms. *Chemical Engineering Science*, 7(3):121–131, 1958.

[6] M. Arthur, S. H. Lam, and D. A. Goussis. *Introduction to physical Chemistry*. Prentice-Hall, Inc, 1st edition, 1982.

[7] P. W. Atkins. *Physical Chemistry*. Oxford University Press, 3rd edition, 1987.

[8] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. Perthame. A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows. *SIAM Journal of Scientific Computing*, 25(6):2050–2065, 2004.

[9] E. Audusse and M.-O. Bristeau. Transport of pollutant in shallow water: a two time-step kinetic method. *ESAIM: Mathematical Modelling and Numerical Analysis*, 37(2):389–416, 2003.

[10] E. Audusse and M.-O. Bristeau. A well-balanced positivity preserving "second-order" scheme for shallow water flows on unstructured meshes. *Journal of Computational Physics*, 206(1):311–333, 2005.

[11] C. Ayora, M. A. Caraballo, F. Macias, T. S. Rötting, J. Carrera, and J.-M. Nieto. Acid mine drainage in the Iberian Pyrite Belt: 2. Lessons learned from recent passive remediation experiences. *Environmental Science and Pollution Research*, 20(11):7837–7853, 2013.

[12] L. Begnudelli and B. F. Sanders. Unstructured grid finite-volume algorithm for shallow-water flow and scalar transport with wetting and drying. *Journal of Hydraulic Engineering*, 132(4):371–384, 2006.

[13] A. Beljadid, A. Mohammadian, and A. Kurganov. Well-balanced positivity preserving cell-vertex central-upwind scheme for shallow water flows. *Computers & Fluids*, 136:193–206, 2016.

[14] A. Bellen, Z. Jackiewicz, and M. Zennaro. Contractivity of waveform relaxation Runge-Kutta iterations and related limit methods for dissipative systems in the maximum norm. *SIAM Journal on Numerical Analysis*, 31(2):499–523, 1994.

[15] B. P. Belousov. A periodic reaction and its mechanism. *Oscillation and Travelling Waves in Chemical Systems*, 1951.

[16] F. Benkhaldoun, I. Elmahi, M. Seaïd, et al. Well-balanced finite volume schemes for pollutant transport by shallow water equations on unstructured meshes. *Journal of Computational Physics*, 226(1):180–203, 2007.

[17] A. Bermudez and M. E. Vazquez. Upwind methods for hyperbolic conservation laws with source terms. *Computers & Fluids*, 23(8):1049–1071, 1994.

[18] G. Beylkin, J. M. Keiser, and L. Vozovoi. A new class of time discretization schemes for the solution of nonlinear PDEs. *Journal of Computational Physics*, 147(2):362–387, 1998.

[19] J. Blom and J. Verwer. A comparison of integration methods for atmospheric transport-chemistry problems. *Journal of Computational and Applied Mathematics*, 126(1-2):381–396, 2000.

[20] A. Bollermann, S. Noelle, and M. Lukáčová-Medvid'ová. Finite volume evolution Galerkin methods for the shallow water equations with dry beds. *Communications in Computational Physics*, 10(2):371–404, 2011.

[21] A. Bourlioux, A. T. Layton, and M. L. Minion. High-order multi-implicit spectral deferred correction methods for problems of reactive flow. *Journal of Computational Physics*, 189(2):651–675, 2003.

[22] N. J. Brown, G. Li, and M. L. Koszykowski. Mechanism reduction via principal component analysis. *International Journal of Chemical Kinetics*, 29(6):393–414, 1997.

[23] S. Bryson, Y. Epshteyn, A. Kurganov, and G. Petrova. Well-balanced positivity preserving central-upwind scheme on triangular grids for the Saint-Venant system. *ESAIM: Mathematical Modelling and Numerical Analysis*, 45(3):423–446, 2011.

[24] V. Caleffi. A new well-balanced Hermite weighted essentially non-oscillatory scheme for shallow water equations. *International Journal for Numerical Methods in Fluids*, 67(9):1135–1159, 2011.

[25] V. Caleffi. A new well-balanced Hermite weighted essentially non-oscillatory scheme for shallow water equations. *International Journal for Numerical Methods in Fluids*, 67(9):1135 – 1159, 2011.

[26] M. H. Carpenter, D. Gottlieb, and S. Abarbanel. The stability of numerical boundary treatments for compact high-order finite-difference schemes. *Journal of Computational Physics*, 108(2):272–295, 1993.

[27] R. W. Carr Jr., D. G. Peterson, and F. K. Smith. Flash photolysis of 1, 3-dichlorotetrafluoroacetone in the presence of oxygen. Kinetics and mechanism of the oxidation of the chlorodifluoromethyl radicals. *Journal of Physical Chemistry;(United States)*, 90(4), 1986.

[28] J. Cash. Efficient numerical methods for the solution of stiff initial-value problems and differential algebraic equations. *Proceedings of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, 459(2032):797–815, 2003.

[29] V. Castets, E. Dulos, J. Boissonade, and P. De Kepper. Experimental evidence of a sustained standing Turing-type nonequilibrium chemical pattern. *Physical Review Letters*, 64(24):2953, 1990.

[30] L. Cea and M. E. Vázquez-Cendón. Unstructured finite volume discretisation of bed friction and convective flux in solute transport models linked to the shallow water equations. *Journal of Computational Physics*, 231(8):3317–3339, 2012.

[31] T. F. Chan. Stability analysis of finite difference schemes for the advection-diffusion equation. *SIAM Journal on Numerical Analysis*, 21(2):272–284, 1984.

[32] B. M. Chen-Charpentier and H. V. Kojouharov. An unconditionally positivity preserving scheme for advection-diffusion reaction equations. *Mathematical and Computer Modelling*, 57(9-10):2177–2185, 2013.

[33] G. Continillo. *Analysis, Simulation, Dynamics of Chemical Reactors*. CUEN srl, Naples, Italy, 1st edition, 1995.

[34] S. M. Cox and P. C. Matthews. Exponential time differencing for stiff systems. *Journal of Computational Physics*, 176(2):430–455, 2002.

[35] C. F. Curtiss and J. O. Hirschfelder. Integration of stiff equations. *Proceedings of the National Academy of Sciences of the United States of America*, 38(3):235, 1952.

[36] S. Davis. Simplified second-order Godunov-type methods. *SIAM Journal on Scientific and Statistical Computing*, 9(3):445–473, 1988.

[37] L. P. de Oliveira, D. Hudebine, D. Guillaume, and J. J. Verstraete. A review of kinetic modeling methodologies for complex processes. *Oil & Gas Science and Technology–Revue d'IFP Energies nouvelles*, 71(3):45, 2016.

[38] A. Delis and T. Katsaounis. Relaxation schemes for the shallow water equations. *International Journal for Numerical Methods in Fluids*, 41(7):695–719, 2003.

[39] R. Djouad and B. Sportisse. Solving reduced chemical models in air pollution modelling. *Applied Numerical Mathematics*, 44(1-2):49–61, 2003.

[40] D. Edelson. Computer simulation in chemical kinetics. *Science*, 214(4524):981–986, 1981.

[41] B. Einfeldt. On Godunov-type methods for gas dynamics. *SIAM Journal on Numerical Analysis*, 25(2):294–318, 1988.

[42] S. S. E. H. Elnashaie. *Dynamic modelling, bifurcation and chaotic behaviour of gas-solid catalytic reactors*, volume 9. CRC Press, 1996.

[43] J. H. Ferziger, M. Perić, and R. L. Street. *Computational methods for fluid dynamics*. Springer, 3rd edition, 2002.

[44] J. Field and M. Burger. *Oscillations and traveling waves in chemical systems*. Wiley, New York, 1985.

[45] L. Formaggia and A. Scotti. Positivity and conservation properties of some integration schemes for mass action kinetics. *SIAM Journal on Numerical Analysis*, 49(3):1267–1288, 2011.

[46] A. W. Forman. *Combustion Theory: The Fundamental Theory of Chemically Reacting flow Systems*. Benjamin Cummings Publishing Co., 2nd edition, 1985.

[47] N. Fricker, J. Beaudouin, P. Richter, R. Eils, P. H. Krammer, and I. N. Lavrik. Model-based dissection of CD95 signaling dynamics reveals both a pro-and anti-apoptotic role of c-FLIPL. *Journal of Cell Biology*, 190(3):377–389, 2010.

[48] L. Fusi, M. Primicerio, and A. Monti. A model for calcium carbonate neutralization in the presence of armoring. *Applied Mathematical Modelling*, 39(1):348–362, 2015.

[49] J. M. Gallardo, C. Parés, and M. Castro. On a well-balanced high-order finite volume scheme for shallow water equations with topography and dry areas. *Journal of Computational Physics*, 227(1):574–601, 2007.

[50] R. M. Garrels and M. E. Thompson. Oxidation of pyrite by iron sulfate solutions. *American Journal of Science*, 258:57–67, 1960.

[51] O. Gautier, R. W. Carr Jr, and C. Seigneur. Variational sensitivity analysis of a photochemical smog mechanism. *International Journal of Chemical Kinetics*, 17(12):1347–1364, 1985.

[52] B. Giuseppe, M. Giona, and G. A. Romano. *Chaos And Fractals In Chemical Engineering-Proceedings Of The First National Conference*. World Scientific, 1995.

[53] E. Godlewski and P.-A. Raviart. *Numerical approximation of hyperbolic systems of conservation laws*, volume 118. Springer Science & Business Media, 2013.

[54] S. K. Godunov. A difference method for numerical calculation of discontinuous solutions of the equations of hydrodynamics. *Matematicheskii Sbornik*, 89(3):271–306, 1959.

[55] D. A. Goussis. Quasi steady state and partial equilibrium approximations: their relation and their validity. *Combustion Theory and Modelling*, 16(5):869–926, 2012.

[56] J. Greenberg, A. Leroux, R. Baraille, and A. Noussair. Analysis and approximation of conservation laws with source terms. *SIAM Journal of Numerical Analysis*, 34(5):1980–2007, 1997.

[57] J. M. Greenberg and A.-Y. LeRoux. A well-balanced scheme for the numerical processing of source terms in hyperbolic equations. *SIAM Journal on Numerical Analysis*, 33(1):1–16, 1996.

[58] E. Grimme. An implicitly restarted Lanczos method for the model reduction of stable large-scale systems. Technical report, University of of Illinois at Urbana-Champaign, 1993.

[59] E. Grimme, D. Sorensen, and P. Van Dooren. Stable partial realizations via an implicitly restarted Lanczos method. In *Proceedings of 1994 American Control Conference - ACC '94*, volume 3, pages 2814–2818, 1994.

[60] B. Gustafsson, H.-O. Kreiss, and A. Sundström. Stability theory of difference approximations for mixed initial boundary value problems. II. *Mathematics of Computation*, pages 649–686, 1972.

[61] E. Hairer, S. Nørsett, and G. Wanner. *Solving Ordinary Differential Equations I: Nonstiff Problems.* Springer Berlin, 2nd edition, 1993.

[62] R. Henda and K. Alhumaizi. Spatiotemporal patterns in a two-dimensional reaction-diffusion-convection system: Effect of transport parameters. *Mathematical and Computer Modelling*, 36(11-13):1361–1373, 2002.

[63] I. Higueras, N. Happenhofer, O. Koch, and F. Kupka. Optimized strong stability preserving IMEX Runge–Kutta methods. *Journal of Computational and Applied Mathematics*, 272:116–140, 2014.

[64] Z. Horváth. Positivity of Runge-Kutta and diagonally split Runge-Kutta methods. *Applied Numerical Mathematics*, 28(2-4):309–326, 1998.

[65] J. Hou, F. Simons, M. Mahgoub, and R. Hinkelmann. A robust well-balanced model on unstructured grids for shallow water flows with wetting and drying over complex topography. *Computer Methods in Applied Mechanics and Engineering*, 257:126–149, 2013.

[66] D. D. Houghton and A. Kasahara. Nonlinear shallow fluid flow over an isolated ridge. *Communications on Pure and Applied Mathematics*, 21(1):1–23, 1968.

[67] J. Hu. A simple numerical scheme for the 2D shallow-water system. Technical report, The National Hydraulics and Environmental Laboratory, 78400 Chatou, France, 2018.

[68] Y.-J. Huang, L. Hong, and W.-A. Yong. Partial equilibrium approximations in apoptosis II. The death-inducing signaling complex subsystem. *Mathematical Biosciences*, 270:126–134, 2015.

[69] Y.-J. Huang and W.-A. Yong. A stable simplification of a FAS-signaling pathway model for apoptosis. In *2012 IEEE 6th International Conference on Systems Biology (ISB)*, pages 125–134. IEEE, 2012.

[70] Y.-J. Huang and W.-A. Yong. Partial equilibrium approximations in apoptosis. I. The intracellular-signaling subsystem. *Mathematical Biosciences*, 246(1):27–37, 2013.

[71] W. Hundsdorfer and J. G. Verwer. *Numerical solution of time-dependent advection-diffusion-reaction equations*, volume 33. Springer Science & Business Media, 2013.

[72] N. L. Ira. *Physical Chemistry*. McGraw-Hill, Inc, 3rd edition, 1987.

[73] T. Jiang and Y.-T. Zhang. Krylov implicit integration factor WENO methods for semilinear and fully nonlinear advection-diffusion-reaction equations. *Journal of Computational Physics*, 253:368–388, 2013.

[74] Y. Jiang and H. Liu. Invariant-region-preserving DG methods for multi-dimensional hyperbolic conservation law systems, with an application to compressible Euler equations. *Journal of Computational Physics*, 373:385–409, 2018.

[75] S. Jin. A steady-state capturing method for hyperbolic systems with geometrical source terms. *ESAIM: Mathematical Modelling and Numerical Analysis*, 35(4):631–645, 2001.

[76] R. S. Johnson. *A modern introduction to the mathematical theory of water waves*, volume 19. Cambridge University Press, 2nd edition, 1997.

[77] V. Joshi and R. K. Jaiman. A positivity preserving variational method for multi-dimensional convection-diffusion-reaction equation. *Journal of Computational Physics*, 339:247–284, 2017.

[78] R. Kapral and K. Showalter. *Chemical waves and patterns*, volume 10. Springer Science & Business Media, 2012.

[79] A.-K. Kassam and L. N. Trefethen. Fourth-order time-stepping for stiff PDEs. *SIAM Journal on Scientific Computing*, 26(4):1214–1233, 2005.

[80] J. P. Keener and J. Sneyd. *Mathematical physiology*, volume 1. Springer, 1998.

[81] P. Klappa. *Kinetics for Bioscientist*. Peter Klappa & Ventus Publishing ApS, 2009.

[82] T. Koto. IMEX Runge-Kutta schemes for reaction-diffusion equations. *Journal of Computational and Applied Mathematics*, 215(1):182–195, 2014.

[83] F. Kupka, N. Happenhofer, I. Higueras, and O. Koch. Total-variation-diminishing implicit-explicit Runge–Kutta methods for the simulation of double-diffusive convection in astrophysics. *Journal of Computational Physics*, 231(9):3561–3586, 2012.

[84] Y. Kuramoto. Diffusion-induced chaos in reaction systems. *Progress of Theoretical Physics Supplement*, 64:346–367, 1978.

[85] A. Kurganov. Finite-volume schemes for shallow-water equations. *Acta Numerica*, 27:289–351, 2018.

[86] A. Kurganov and D. Levy. A third-order semidiscrete central scheme for conservation laws and convection-diffusion equations. *SIAM Journal on Scientific Computing*, 22(4):1461–1488, 2000.

[87] A. Kurganov and D. Levy. Central-upwind schemes for the Saint-Venant system. *ESAIM: Mathematical Modelling and Numerical Analysis*, 36(3):397–425, 2002.

[88] A. Kurganov, S. Noelle, and G. Petrova. Semidiscrete central-upwind schemes for hyperbolic conservation laws and Hamilton-Jacobi equations. *SIAM Journal on Scientific Computing*, 23(3):707–740, 2001.

[89] Y.-K. Kwok. Stability analysis of six-point finite difference schemes for the constant coefficient convective-diffusion equation. *Computers & Mathematics with Applications*, 23(12):3–11, 1992.

[90] A. Lafon and H. Yee. Dynamical Approach Study of Spurious Steady-State Numerical Solutions of Nonlinear Differential Equations Part IV. Stability vs. Methods of Discretizing Nonlinear Source Terms in Reaction-Convection Equations. *International Journal of Computational Fluid Dynamics*, 6(2):89–123, 1996.

[91] S. H. Lam. Using CSP to understand complex chemical kinetics. *Combustion Science and Technology*, 89(5-6):375–404, 1993.

[92] S. H. Lam and D. A. Goussis. Understanding complex chemical kinetics with computational singular perturbation. In *Symposium (International) on Combustion*, volume 22, pages 931–941, 1989.

[93] S. H. Lam and D. A. Goussis. Conventional asymptotics and computational singular perturbation for simplified kinetics modelling. In *Reduced kinetic mechanisms and asymptotic approximations for methane-air flames*, pages 227–242. Springer, 1991.

[94] D. Langmuir. *Aqueous Environmental Chemistry*. Prentice-Hall, Inc., New York, 1st edition, 1997.

[95] A. M. M. Leal, M. J. Blunt, and T. C. LaForce. A chemical kinetics algorithm for geochemical modelling. *Applied Geochemistry*, 29(55):46–61, 2015.

[96] E. Lelarasmee. *The waveform relaxation method for time domain analysis of large scale integrated circuits: Theory and Applications.* Electronics Research Laboratory, College of Engineering, University of California, 1982.

[97] R. LeVeque and H. Yee. A study of numerical methods for hyperbolic conservation laws with stiff source terms. *Journal of Computational Physics*, 86(1):187–210, 1990.

[98] R. J. LeVeque. *Numerical methods for Conservation Laws*. Birkhàuser, Verlag, Switzerland, 2nd edition, 1992.

[99] R. J. LeVeque. Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm. *Journal of Computational Physics*, 146(1):346–365, 1998.

[100] R. J. LeVeque et al. *Finite volume methods for hyperbolic problems*, volume 31. Cambridge University Press, 2002.

[101] G. Li, V. Caleffi, and Z. Qi. A well-balanced finite difference WENO scheme for shallow water flow model. *Applied Mathematics and Computation*, 265:1–16, 2015.

[102] D. Liang, X. Wang, R. A. Falconer, and B. N. Bockelmann-Evans. Solving the depth-integrated solute transport equation with a TVD-MacCormack scheme. *Environmental Modelling & Software*, 25(12):1619–1629, 2010.

[103] X. Liu, J. Albright, Y. Epshteyn, and A. Kurganov. Well-balanced positivity preserving central-upwind scheme with a novel wet/dry reconstruction on triangular grids for the Saint-Venant system. *Journal of Computational Physics*, 374:213–236, 2018.

[104] U. Maas. Simplifying chemical kinetics using intrinsic low-dimensional manifolds. In *Gas Phase Chemical Reaction Systems*, pages 334–342. Springer, 1996.

[105] Y. Maday, A. T. Patera, and E. M. Rønquist. An operator-integration-factor splitting method for time-dependent problems: application to incompressible fluid flow. *Journal of Scientific Computing*, 5(4):263–292, 1990.

[106] B. S. Martincigh and R. H. Simoyi. Convective instabilities induced by an exothermic autocatalytic chemical reaction. *Physical Review E*, 52(2):1606–1613, 1995.

[107] G. J. McRae, W. R. Goodin, and J. H. Seinfeld. Numerical solution of the atmospheric diffusion equation for chemically reacting flows. *Journal of Computational Physics*, 45(1):1–42, 1982.

[108] M. Menzinger and A. B. Rovinsky. *The differential flow instabilities In: Kapral R., Showalter K. (eds) Chemical Waves and Patterns. Understanding Chemical Reactivity*, volume 10, pages 365–397. Springer, Dordrecht, 1995.

[109] U. Miekkala and O. Nevanlinna. Convergence of dynamic iteration methods for initial value problems. *SIAM Journal on Scientific and Statistical Computing*, 8(4):459–482, 1987.

[110] M. Milev and A. Tagliani. Efficient implicit scheme with positivity preserving and smoothing properties. *Journal of Computational and Applied Mathematics*, 243:1–9, 2013.

[111] R. Moussa and C. Bocquillon. Approximation zones of the Saint-Venant equations for flood routing with overbank. *Hydrology & Earth System Science*, 4(2):251–261, 2000.

[112] S. Mungkasi. Finite Volume Methods for the One-Dimensional Shallow Water Equations. Master's thesis, The Australian National University, Australia, 2008.

[113] J. Murillo and P. García-Navarro. Improved Riemann solvers for complex transport in two-dimensional unsteady shallow flow. *Journal of Computational Physics*, 230(19):7202–7239, 2011.

[114] F. T. Namio, E. Ngondiep, R. Ntchantcho, and J. C. Ntonga. Mathematical Model of Complete Shallow Water Problem with Source Terms, Stability Analysis of Lax-Wendroff Scheme. *Journal of Theoretical and Computational Science*, 2(132), 2015.

[115] O. Nevanlinna. Remarks on Picard-Lindelöv iteration. *BIT Numerical Mathematics*, 29(3):535–562, 1989.

[116] G. Nicolis and I. Prigogine. *Self-Organization in Non-Equilibrium Systems*. Wiley, New York, 1977.

[117] Q. Nie, Y.-T. Zhang, and R. Zhao. Efficient semi-implicit schemes for stiff systems. *Journal of Computational Physics*, 214(2):521–537, 2006.

[118] S. Noelle, N. Pankratz, G. Puppo, and J. R. Natvig. Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows. *Journal of Computational Physics*, 213(2):474–499, 2006.

[119] D. K. Nordstrom. Hydrogeochemical processes governing the origin, transport and fate of major and trace elements from mine wastes and mineralized rock to surface waters. *Applied Geochemistry*, 26(11):1777–1791, 2011.

[120] D. K. Nordstrom, D. W. Blowes, and C. J. Ptacek. Hydrogeochemistry and microbiology of mine drainage: an update. *Applied Geochemistry*, 57:3–16, 2015.

[121] Z. Noszticzius, W. Horsthemke, W. D. McCormick, H. L. Swinney, and W. Y. Tam. Sustained chemical waves in an annular gel reactor: a chemical pinwheel. *Nature*, 329(6140):619–620, 1987.

[122] A. Pacault, P. Hanusse, P. De Kepper, C. Vidal, and J. Boissonade. Phenomena in homogeneous chemical systems far from equilibrium. *Accounts of Chemical Research*, 9(12):438–445, 1976.

[123] L. Pareschi and G. Russo. Implicit-explicit Runge-Kutta schemes and applications to hyperbolic systems with relaxation. *Journal of Scientific Computing*, 25(1):129–155, 2005.

[124] N. Peters. Numerical and asymptotic analysis of systematically reduced reaction schemes for hydrocarbon flames. In *Glowinski R., Larrouturou B., Temam R. (eds) Numerical simulation of combustion phenomena. Lecture Notes in Physics*, volume 241, pages 90–109. Springer, 1985.

[125] L. Petzold and W. Zhu. Model reduction for chemical kinetics: An optimization approach. *AIChE Journal*, 45(4):869–886, 1999.

[126] L. N. Plummer, T. M. L. Wigley, and D. L. Parkhurst. The kinetics of calcite dissolution in CO_2-water systems at 5^oC to 60^oC and 0.0 to 1.0 atm CO_2. *American Journal of Science*, 278(2):179–216, 1978.

[127] G. Puppo and G. Russo. *Numerical methods for balance laws.* Dipartimento di matematica della Seconda Universitá di Napoli, 2009.

[128] J. Qiu and C.-W. Shu. Hermite WENO schemes and their application as limiters for Runge-Kutta discontinuous Galerkin method: one-dimensional case. *Journal of Computational Physics*, 193(1):115–135, 2004.

[129] A. Quarteroni, R. Sacco, and F. Saleri. *Numerical Mathematics*. Springer-Verlag, Berlin, 1st edition, 2007.

[130] M. M. Reddy, L. N. Plummer, and E. Busenberg. Crystal growth of calcite from calcium bicarbonate solution at constant P_{CO_2} and 25^oC : a test of calcite dissolution model. *Geochimica et Cosmochimica Acta*, 45(8):1281–1289, 1981.

[131] Z. Ren and S. B. Pope. Second-order splitting schemes for a class of reactive systems. *Journal of Computational Physics*, 227(17):8165–8176, 2008.

[132] D. L. Ropp and J. N. Shadid. Stability of operator splitting methods for systems with indefinite operators: Advection-diffusion-reaction systems. *Journal of Computational Physics*, 228(9):3508–3516, 2009.

[133] A. Sandu. Positive numerical integration methods for chemical kinetic systems. *Journal of Computational Physics*, 170(2):589–602, 2001.

[134] R. A. Satnoianu, J. H. Merkin, and S. K. Scott. Differential-flow-induced instability in a cubic autocatalator system. *Journal of Engineering Mathematics*, 33(1):77–102, 1998.

[135] B. Schippmann and H. Burchard. Rosenbrock methods in biogeochemical modelling–a comparison to Runge-Kutta methods and modified Patankar schemes. *Ocean Modelling*, 37(3-4):112–121, 2011.

[136] C. Seigneur, G. Stephanopoulos, and R. W. Carr Jr. Dynamic sensitivity analysis of chemical reaction systems: a variational method. *Chemical Engineering Science*, 37(6):845–853, 1982.

[137] T. K. Sengupta. *High accuracy computing methods: fluid flows and wave phenomena*. New York, USA: Cambridge University Press, 2003.

[138] T. K. Sengupta and A. Bhole. Error dynamics of diffusion equation: effects of numerical diffusion and dispersive diffusion. *Journal of Computational Physics*, 266:240–251, 2014.

[139] T. K. Sengupta, Y. G. Bhumkar, M. K. Rajpoot, V. K. Suman, and S. Saurabh. Spurious waves in discrete computation of wave phenomena and flow problems. *Applied Mathematics and Computing*, 218(18):9035–9065, 2012.

[140] T. K. Sengupta, A. Dipankar, and P. Sagaut. Short notes: Error dynamics: Beyond von Neumann analysis. *Journal of Computational Physics*, 226(2):1211–1218, 2007.

[141] T. K. Sengupta, G. Ganeriwal, and S. De. Analysis of central and upwind compact schemes. *Journal of Computational Physics*, 192(2):677–694, 2003.

[142] T. K. Sengupta, R. Jain, and A. Dipankar. A new flux-vector splitting compact finite volume scheme. *Journal of Computational Physics*, 207(18):261–281, 2005.

[143] T. K. Sengupta, V. Lakshmanan, and V. Vijay. A new combined stable and dispersion relation preserving compact scheme for non-periodic problems. *Journal of Computational Physics*, 228(8):3048–3071, 2009.

[144] T. K. Sengupta, A. Sengupta, and K. Saurabh. Global spectral analysis of multilevel time integration schemes: Numerical properties for error analysis. *Applied Mathematics and Computation*, 304:41–57, 2017.

[145] T. K. Sengupta, V. Vijay, and S. Bhaumik. Further improvement and analysis of CCD scheme: dissipation discretization and de-aliasing properties. *Journal of Computational Physics*, 228(17):6150–6168, 2009.

[146] K. S. Shabani, F. D. Aredejani, R. N. Singh, R. Marandi, and H. Soleimanyfar. Numerical modeling of Cu^{2+} and Mn^{2+} ions biosorption by Aspergillus Niger Fungal biomass in Continuous Reactor. *Archives of Mining Sciences*, 56(3):461–476, 2011.

[147] P. C. Singer and W. Stumm. Acidic mine drainage: the rate-determining step. *Science*, 167(3921):1121–1123, 1970.

[148] A. Siviglia and E. F. Toro. WAF method and splitting procedure for simulating hydro-and thermal-peaking waves in open-channel flows. *Journal of Hydraulic Engineering*, 135(8):651–662, 2009.

[149] M. D. Smooke. *Reduced kinetic mechanisms and asymptotic approximations for methane-air flames: a topical volume.* Springer, 1991.

[150] E. Sousa. *Finite differences for the convection-diffusion equation: On stability and boundary conditions.* PhD thesis, University of Oxford, 2001.

[151] B. Sportisse. An analysis of operator splitting techniques in the stiff case. *Journal of Computational Physics*, 161(1):140–168, 2000.

[152] B. Sportisse and R. Djouad. Reduction of chemical kinetics in air pollution modeling. *Journal of Computational Physics*, 164(2):354–376, 2000.

[153] J. J. Stoker. *Water Waves: The Mathematical Theory with Applications.* John Wiley & Sons, Inc., 1st edition, 1992.

[154] J. C. Strikwerda. *Finite difference schemes and partial differential equations*, volume 88. SIAM, 2004.

[155] W. Stumm and G. F. Lee. Oxygenation of ferrous iron. *Industrial & Engineering Chemistry*, 53(2):143–146, 1961.

[156] W. Stumm and J. J. Morgan. *Aquatic Chemistry: An introduction emphasizing chemical equilibria in natural waters.* New York, NY: Wiley-Interscience, 1970.

[157] V. K. Suman, T. K. Sengupta, C. J. D. Prasad, K. S. Mohan, and D. Sanwalia. Spectral analysis of finite difference schemes for convection diffusion equation. *Computers & Fluids*, 150:95–114, 2017.

[158] M. Svärd and S. Mishra. Implicit-explicit schemes for flow equations with stiff source terms. *Journal of Computational and Applied Mathematics*, 235(6):1564–1577, 2011.

[159] W. Y. Tam, W. Horsthemke, Z. Noszticzius, and H. L. Swinney. Sustained spiral waves in a continuously fed unstirred chemical reactor. *The Journal of Chemical Physics*, 88(5):3395–3396, 1988.

[160] E. F. Toro. *Shock-capturing methods for free-surface shallow flows*, volume 868. Wiley New York, 2001.

[161] E. F. Toro. *Riemann Solvers and Numerical Methods for Fluid Dynamics*. Springer-Verlag Berlin Heidelberg, 3rd edition, 2009.

[162] E. F. Toro and P. Garcia-Navarro. Godunov-type methods for free-surface shallow flows: A review. *Journal of Hydraulic Research*, 45(6):736–751, 2007.

[163] L. N. Trefethen. Group velocity in finite difference schemes. *SIAM Review*, 24(2):113–136, 1982.

[164] A. M. Turing. The chemical basis of morphogenesis. *Bulletin of Mathematical Biology*, 52(1-2):153–197, 1990.

[165] S. Vajda, P. Valko, and T. Turanyi. Principal component analysis of kinetic models. *International Journal of Chemical Kinetics*, 17(1):55–81, 1985.

[166] D. Vanzo, A. Siviglia, and E. F. Toro. Pollutant transport by shallow water equations on unstructured meshes: hyperbolization of the model and numerical solution via a novel flux splitting scheme. *Journal of Computational Physics*, 321:1–20, 2016.

[167] M. E. Vázquez-Cendón. Improved treatment of source terms in upwind schemes for the shallow water equations in channels with irregular geometry. *Journal of Computational Physics*, 148(2):497–526, 1999.

[168] H. K. Versteeg and W. Malalasekera. *An Introduction to Computational Fluid Dynamics: The Finite Volume Method*. Pearson Education Limited, 2nd edition, 2007.

[169] J. G. Verwer, B. P. Sommeijer, and W. Hundsdorfer. RKC time-stepping for advection-diffusion-reaction problems. *Journal of Computational Physics*, 201(1):61–79, 2004.

[170] R. Vichnevetsky and J. B. Bowles. Fourier analysis of numerical approximation of hyperbolic equations. *Mathematics of Computation*, 43(168):615, 1984.

[171] J. M. Walter. *Physical Chemistry*. Prentice-Hall, Inc., 5th edition, 1972.

[172] S. Walton, O. Hassan, and K. Morgan. Reduced order modelling for unsteady fluid flow using proper orthogonal decomposition and radial basis functions. *Applied Mathematical Modelling*, 37(20-21):8930–8945, 2013.

[173] W. Wang, C.-W. Shu, H. C. Yee, and B. Sjögreen. High-order well-balanced schemes and applications to non-equilibrium flow. *Journal of Computational Physics*, 228(18):6682–6702, 2009.

[174] G. Wanner and E. Hairer. *Solving ordinary differential equations II. Stiff and Differential Algebraic Problems*. Springer Berlin Heidelberg, 2nd edition, 1996.

[175] P. Wesseling. von Neumann stability conditions for the convection-diffusion eqation. *IMA Journal of Numerical Analysis*, 16(4):583–598, 1996.

[176] J. A. White, A. Sangiovanni-Vincentelli, F. Odeh, and A. Ruehli. *Waveform relaxation: Theory and practice*. Electronics Research Laboratory, College of Engineering, UCB, 1985.

[177] Y. Xing and C.-W. Shu. High order finite difference WENO schemes with the exact conservation property for the shallow water equations. *Journal of Computational Physics*, 208(1):206–227, 2005.

[178] Y. Xing and C.-W. Shu. High-order well-balanced finite difference WENO schemes for a class of hyperbolic systems with source terms. *Journal of Scientific Computing*, 27(1-3):477–494, 2006.

[179] Y. Xing and C.-W. Shu. High order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *Journal of Computational Physics*, 214(2):567–598, 2006.

[180] Y. Xing and C.-W. Shu. A new approach of high order well-balanced finite volume WENO schemes and discontinuous Galerkin methods for a class of hyperbolic systems with source terms. *Journal of Computational Physics*, 1(1):100–134, 2006.

[181] Y. Xing, C.-W. Shu, and S. Noelle. On the advantage of well-balanced schemes for moving-water equilibria of the shallow water equations. *Journal of Scientific Computing*, 48(1-3):339–349, 2011.

[182] B. Yuan, D. Yuan, J. Sun, and J. Tao. A finite volume model for coupling surface and subsurface flows. *Procedia Engineering*, 31:62–67, 2012.

[183] B. Zhang and J.-H. Wang. A short note on the counter-intuitive spurious behaviors in stiff reacting flow. *Journal of Computational Physics*, 291:52–59, 2015.

[184] Y. Zhang, C. H. Bischof, R. C. Easter, and P.-T. Wu. Sensitivity analysis of a mixed-phase chemical mechanism using automatic differentiation. *Journal of Geophysical Research: Atmospheres*, 103(D15):18953–18979, 1998.

[185] S. Zhao, J. Ovadia, X. Liu, Y.-T. Zhang, and Q. Nie. Operator splitting implicit integration factor methods for stiff reaction-diffusion-advection systems. *Journal of Computational Physics*, 230(15):5996–6009, 2011.

[186] X. Zhong. High-order finite-difference schemes for numerical simulation of hypersonic boundary-layer transition. *Journal of Computational Physics*, 144(2):662–709, 1998.

[187] J. G. Zhou, D. M. Causon, C. G. Mingham, and D. M. Ingram. The surface gradient method for the treatment of source terms in the shallow-water equations. *Journal of Computational Physics*, 168(1):1–25, 2001.

[188] C. Zhu, F. Q. Hu, and D. S. Burden. Multi-component reactive transport modeling of natural attenuation of an acid groundwater plume at a uranium mill tailings site. *Journal of Contaminant Hydrology*, 52(1-4):85–108, 2001.