Imperial College London

Modelling and Numerical Analysis of Energy-Dissipating Systems with Nonlocal Free Energy

Presented by

Sergio Pérez Pérez

Supervisors:
Prof. José A. Carrillo
Prof. Serafim Kalliadasis

A mi familia,
por quererme y apoyarme siempre

To my family,
for always loving and supporting me

"The good life is one inspired by love
and guided by knowledge.
Neither love without knowledge,
nor knowledge without love
can produce a good life."

Bertrand Russell

Declaration of Originality

I certify that this dissertation, and the research to which it refers, are the product of my own work, and that any ideas or quotations from the work of other people, published or otherwise, are fully acknowledged in accordance with the standard referencing practices of the discipline.

Sergio Pérez Pérez

Copyright Declaration

Abstract

The broad objective of this thesis is to design finite-volume schemes for a family of energy-dissipating systems. All the systems studied in this thesis share a common property: they are driven by an energy that decreases as the system evolves. Such decrease is produced by a dissipation mechanism, which ensures that the system eventually reaches a steady state where the energy is minimized. The numerical schemes presented here are designed to discretely preserve the dissipation of the energy, leading to more accurate and cost-effective simulations. Most of the material in this thesis is based on the publications [16, 54, 65, 66, 243].

The research content is structured in three parts. First, Part II presents well-balanced first-, second- and high-order finite-volume schemes for a general class of hydrodynamic systems with linear and nonlinear damping. These well-balanced schemes preserve stationary states at machine precision, while discretely preserving the dissipation of the discrete free energy for first- and second-order accuracy. Second, Part III focuses on finite-volume schemes for the Cahn-Hilliard equation that unconditionally and discretely satisfy the boundedness of the phase field and the free-energy dissipation. In addition, our Cahn-Hilliard scheme is employed as an image inpainting filter before passing damaged images into a classification neural network, leading to a significant improvement of damaged-image prediction. Third, Part IV introduces finite-volume schemes to solve stochastic gradient-flow equations. Such equations are of crucial importance within the framework of fluctuating hydrodynamics and dynamic density functional theory. The main advantages of these schemes are the preservation of non-negative densities in the presence of noise and the accurate reproduction of the statistical properties of the physical systems. All these finite-volume schemes are complemented with prototypical examples from relevant applications, which highlight the benefit of our algorithms to elucidate some of the unknown analytical results.

Acknowledgements

This thesis is only possible thanks to all the people that have supported me in the last years. First, I would like to thank my supervisors, José A. Carrillo and Serafim Kalliadasis, who have been a wonderful source of inspiration, advice and generosity since even before starting the PhD. In addition to their evident scientific excellence, I have really appreciated the numerous opportunities to participate in conferences, collaborate with international researchers, supervise students and attend summer schools. José has been an incredible mentor since my undergraduate studies, and I am deeply influenced by his lively personality, bias for action and positive attitude. Serafim has really welcomed me into his research group, and I will always bear with me his professional advices and attention to detail. It has been a true honour to work alongside you, and I will always keep you as role models.

I am also truly thankful to all the researchers with whom I have collaborated during my PhD. In particular, I am grateful to Manuel J. Castro at Universidad de Málaga and Chi-Wang Shu at Brown University for their hospitality in my research stays at their universities. Both of them are superlative experts that really inspired me to have an impact with my research, and I greatly appreciate all the tricks they taught me to make my numerical simulations more robust and reliable. I also learnt a lot by collaborating with Antonio Russo, Rafael Bailo, Miguel A. Durán-Olivencia and Peter Yatsyshin at Imperial College. In addition, I have been fortunate enough to supervise excellent students with whom I have also shared many ideas and produced publications: Fuyue Liang, Julien Mendes, Chan Huh and Ruochen Zhao. Finally, I thank my examiners Rosa Donat and Demetrios Papageorgiou, who accepted to read all these words.

There have been many more people with whom I have learnt and had fun during my PhD. From the Maths Department at Imperial I would like to thank Jeremy Wu, Rishabh Gvalani, Markus Schmidtchen, Franca Hoffmann, Urbain Vaes, Martin Chak and Matt Barker for all the seminars, summer schools, workshops and conferences attended together. I have also shared many enlightening conversations and inspiring conferences with Antonio Esposito, Havva Yoldaş, Sara Merino, Ewelina Zatorska, David Gomez-Castro, José Alfredo Cañizo, Benjamin Aymard, Marc Pradas, David Sibley, Susana Gomes, Matt Tranter and Eleftherios Kyrkinis.

My PhD studies have been funded by the Imperial College London President's PhD Scholarship. I would like to also express my gratitude to the CDT in Fluid Dynamics across Scales of Imperial College London for funding my MRes before beginning the PhD. I also acknowledge financial support from the various BAMC conferences I have attended, the SIAM Conference on Analysis of PDEs, the Gene Golub SIAM Summer School, the Society of Spanish Researchers in the UK, the Old Centralians' Trust at Imperial College London, the PDE-MANS at Granada, the Escuela Lluís Santaló at Santander and the ICIAM at Valencia.

I shall take this opportunity to also thank my previous academic supervisors: Young-Pil Choi for his inspirational role in my UROP and Erasmus exchange at Imperial College; Berend van Wachem and Thomas Curran for their guidance during my MRes thesis; and Antonio Rodríguez, María Luisa Rapún, Mariola Gómez, Juan Antonio Hernández and Ignacio Parra for their continuous support

and kindness during my undergraduate studies and beyond.

I would also like to express my gratitude to the friends that surrounded me during these years in London. Special thanks to Marc and Isaac for living with me and beating them in countless FIFA games; to Stavros, Paula, Giorgia, Fede, Dilan, Sami, Orange, Chris and Jaina for our gatherings and brunches; to Fabián, Carlota, Sofía, Edgar, Leo and Zahra for all the warm parties with delicious food; to Antonio and Virginia for the board games, dinners and plants; to my office mates Ibrahim, Amjaad, Husain and Filip for always having time for a chat. Many other friends have been amazing too: Faron, Alba from Lleida, Melissa, Sandra, Anna, Lluis, Mireia, Magda, Jazmín, Stas, Hakon, Anwar, Kamil, Wing, Giuliano, Antoine, Tanawan, Inés, Miriam and Rand, among many others. Special thanks to my friends at the Society of Spanish Researchers in the UK, with whom I have planned so many cool initiatives; to my cohort mates at the CDT in Fluid Dynamics across Scales, and to my colleagues at Amazon, specially Leonie, Laura and Manchi.

I also wish to thank my friends in Madrid, who always welcome me as if no time has passed. Special thanks to my brother Pablo, Migui, Edu López, Gonzalo, Edu Plazas and Adri for their lifelong friendship; to Borja, Catena and Gonzalo for being awesome friends, to Alex, Yajing and Diego for inspiring me and flying with me to China, to Roberto for his continuous friendship, to Lu and Cortina for our insightful conversations, to Pablo, Javichu, Juan, Mirari and Reig for our friendship since school; and to my friends from ETSIAE with whom I have shared so many moments of fun. I also cherish good memories with many other friends: Laura, Lucía, Sergio Condés, Charlie, Eugenio, Adri Sabaté, Ali, Paloma, María and Cris.

A special and emotive acknowledgement goes to Vicky, for her endless love and unconditional support. Without your cheerleading and distractions I would have written this thesis much faster. I also thank Elena, Tania and Antonio for their encouragement and all the moments together.

Finally, I feel deeply fortunate for having such an incredible family supporting me from Madrid. This thesis is dedicated to them: to my grandmother Ángeles for transmitting me her life wisdom and teaching me the right values to share my life with the people I love; to my aunt Elena for her amazing and generous spirit and for always being a constant and joyful presence in my life; to my uncle Anto for all our time together and for sharing with me many stories and technical ideas; to my grandmother Mabel, my uncle Lin and my cousins for all the unmissable birthdays, Christmas and other events together; to my family from Málaga for demonstrating that everything in life improves surrounded by loved ones; and to my siblings Pablo and Sofía for being the best friends I can imagine and for all the laughs, surprises and time together that await us in the future. Last but not least, an immense thanks is to my parents Mª Ángeles and Luis for all their endless dedication and for being extraordinary examples of a life inspired by love and guided by knowledge.

(Translation to Spanish) Finalmente, me siento muy afortunado de tener una familia maravillosa apoyándome desde Madrid. Esta tesis os la dedico a vosotros: a mi abuela Ángeles por transmitirme su sabiduría y enseñarme a compartir mi vida con la gente que quiero; a mi tía Elena por su increíble y generosa forma de ser, y por ser un apoyo constante en mi vida; a mi tío Anto por todo nuestro tiempo juntos y por compartir conmigo muchas historias e ideas sobre tecnología; a mi abuela Mabel, mi tío Lin y los primos por todas las celebraciones de cumpleaños, Navidad y demás eventos juntos; a mi familia de Málaga por demostrarme que la vida sabe mejor rodeada de seres queridos; y a mis hermanos Pablo y Sofía por ser los mejores amigos que puedo imaginar y por todas las risas, sorpresas y tiempo juntos que nos esperan en el futuro. Por último, un agradecimiento inmenso es para mis padres Mª Ángeles and Luis por su dedicación infinita y por ser ejemplos extraordinarios de cómo vivir inspirados por el amor y guiados por el conocimiento.

Publications

Chapter 5 of this thesis is based on the following publication:

[66]   J. A. Carrillo, S. Kalliadasis, S. P. Perez, and C. W. Shu. "Well-balanced finite-volume schemes for hydrodynamic equations with general free energy". In: *Multiscale Modeling & Simulation* 18.1 (2020), pp. 502-541.

Chapter 6 of this thesis is based on the following publication:

[54]   J. A. Carrillo, M. J. Castro, S. Kalliadasis, and S. P. Perez. "High-Order Well-Balanced Finite-Volume Schemes for Hydrodynamic Equations With Nonlocal Free Energy". In: *SIAM Journal on Scientific Computing* 43.2 (2021), A828-A858.

Chapter 8 of this thesis is based on the following publication:

[16]   R. Bailo, J. A. Carrillo, S. Kalliadasis, and S. P. Perez. "Unconditional bound-preserving and energy-dissipating finite-volume schemes for the Cahn-Hilliard equation". In: *arXiv preprint arXiv:2105.05351* (2021). To be submitted.

Chapter 9 of this thesis is based on the following publication:

[65]   J. A. Carrillo, S. Kalliadasis, F. Liang, and S. P. Perez. "Enhancement of damaged-image prediction through Cahn-Hilliard Image Inpainting". In: *Royal Society Open Science* 8 (2021), 201294.

Chapter 11 of this thesis is based on the following publication:

[243]   A. Russo, S. P. Perez, M. A. Durán-Olivencia, P. Yatsyshin, J. A. Carrillo, and S. Kalliadasis. "A finite-volume method for fluctuating dynamical density functional theory". In: *Journal of Computational Physics* (2020), p. 109796.

Other related publications by the author of this thesis but not included here are:

[210]   J. Mendes, A. Russo, S. P. Perez, and S. Kalliadasis. "A finite-volume scheme for gradient-flow equations with non-homogeneous diffusion". In: *Computers & Mathematics with Applications* 89 (2021), pp. 150-162.

[56]   J. A. Carrillo, Y.-P. Choi, and S. P. Perez. "A review on attractive–repulsive hydrodynamics for consensus in collective behavior". In: *Active Particles, Volume 1* (2017), pp. 259-298, Springer.

Contents

Part I

Introduction

My thesis in a nutshell

Amidst the many fascinating areas of Mathematics, this thesis belongs to the field of Numerical Analysis. Its broad objective is to design algorithms able to provide approximate solutions of mathematical problems. One may think that Numerical Analysis should be a quite young field, given that the word "algorithm" is usually associated to the development of computers in the second half of the 20th century. The truth is that an algorithm is just a set of rules to perform a numerical computation, and mathematicians had been employing them for many centuries before the advent of computers. The earliest approximate calculation that we are aware of dates back to some time between 1800 and 1600 BCE, where the square root of two is accurately computed up to six decimal digits (the square root of two is an irrational number with infinite number of decimals). The origin of the word "algorithm" is pretty old too, coming from the name of the Arabic mathematician Al-Khwarizmi (around 780-850 CE): what happened was that the Latin translators of his algebra book changed his name from "Al-Khwarizmi" to "Algoritmi".

Anyhow, mathematicians discovered quite early that, more often than not, real-world problems don't have an exact analytical solution. Many of the basic numerical methods employed today still bear the names of mathematicians whose work had a massive impact in a wide range of scientific disciplines: Newton's method, Gaussian elimination, Euler's scheme or Lagrange interpolation polynomial to name a few. The applicability of algorithms was turbocharged hand-in-hand with the rapid growth of computers and the introduction of high-level programming languages after World War II. Two main pioneers are essential to understand the birth of computer science: on the one hand, Alan Turing (1912-1954) developed the concepts of Turing machine to formalize the idea of an algorithm and the Turing test as a first step to Artificial Intelligence; on the other hand, John von Neumann (1903-1957) founded the subject of game theory and introduced the Von Neumann computer architecture. These pioneering advances in computer science built from all the ideas developed in the previous centuries, specially from the contributions of Ada Lovelace and Charles Baggage. The first simulation performed on a computer was during the Manhattan project in World War II, where the process of nuclear detonation was simulated employing 12 hard spheres and a Monte Carlo algorithm.

Nowadays Numerical Analysis finds application in most areas of engineering and physical sciences, and recently it has expanded to other fields like medicine, business or social sciences. As an example of the multidisciplinary nature of Numerical Analysis, let's just focus on some examples of applications provided in this thesis: chemotaxis of cells and organisms, image inpainting to remove damage from pictures, collective behaviour of individual agents like pedestrians or school of fish, or water flow in shallow rivers. In appearance all these applications belong to different scientific

fields, yet the algorithms that I have designed apply to all of them. The merit here also comes from the modelling of such phenomena, or in other words, the common mathematical equation that underlies beneath all these applications and allows to simulate one application or the other by just tuning some parts of such model.

The field of numerical analysis includes many subdisciplines: interpolation, regression, resolution of systems of equations, optimization or evaluation of integrals. In this thesis we focus on another really essential subfield: the computation of the solution of differential equations. More specifically, the equations solved in this thesis belong to the class of partial differential equations (PDEs) and stochastic partial differential equations (SPDEs). For instance, the airflow around a plane is simulated with the legendary Navier-Stokes equations, which is a system of PDEs that evolve the density, velocity and temperature of the airflow. The Navier-Stokes equations are so intricate that no general analytical solutions can be extracted from them (in fact, it has not yet been proven whether smooth solutions always exist in three dimensions). This means that the only way of resolving the airflow around a plane is via algorithms.

The first methods to solve differential equations date back a few centuries ago. It is believed that Newton was the first in solving differential equations manually with an algorithm, and a bit later Euler introduced the popular Euler's forward method. More recently, in the turn from the 19$^{\text{th}}$ to the 20$^{\text{th}}$ century, Runge firstly and Kutta afterwards set the foundations of the multi-stage Runge-Kutta methods of higher order. All these methods work well as long as we deal with ordinary differential equations (ODE), where the derivatives in the equation only depend on one variable (typically time). The differential equation gets trickier when we add further derivatives into it, becoming then a PDE. As an example, the Navier-Stokes equations are PDEs because they contain derivatives both in time and in space. For PDEs one needs to combine the previous numerical methods of ODEs for the time variable with a discretization technique for the spatial variables.

Such spatial discretization can be carried out in many diverse ways. The first approach was based on finite differences on rectangular grids, where the objective is to evaluate the variables only at discrete points of the domain. One of the first seminal papers of numerical analysis for PDEs precisely employed this method: in 1928 Courant, Friedrichs and Levy introduced the condition for stability of finite-difference schemes applied in hyperbolic PDEs. Their work is so fundamental that in many chapters of this thesis I prove properties of my algorithms that depend on a Courant-Friedrichs-Levy (CFL) condition to be satisfied. Finite differences were the mainstream methodology until the sixties, and at that point two other more sophisticated spacial discretizations gained traction and are still widely employed nowadays: finite volumes and finite elements. On the one hand, finite volumes have evolved to become one of the predominant methods in the area of Computational Fluid Dynamics (CFD), in part thanks to the conservative formulations of fluid-type models. On the other hand, finite elements grew as a simulation tool for structural analysis and Computational Solid Mechanics, although nowadays their applicability has spread to many other fields including CFD. Other discretization techniques that have been developed from finite volumes and finite elements are spectral or discontinuous Galerkin methods.

As the title of this thesis suggests, my research is focused on finite-volume schemes. Their objective is to transform PDEs, typically depending on space and time, into a system of ODEs depending only on time. As an example, let's say that we are studying a PDE that models the evolution of the density of a fluid. The first step is to discretize the spatial domain into a mesh of non-overlapping patches or finite volumes. Then the PDEs are transformed into a system of ODEs by integrating them over each of the finite volumes in the discretized domain. Each of the resulting ODEs describes the evolution of the average of the fluid density over that particular finite volume. Subsequently, the system of ODEs is solved by advancing in time with a temporal discretization such as Euler's forward method or Runge-Kutta. The result of the simulation is a set of discrete temporal steps for which we have computed the averages of the fluid density in each of the finite

volumes covering the spatial domain.

Finite volumes are so popular in CFD because they automatically conserve certain natural quantities of fluid-type PDEs. For instance, in a closed system the total mass of a fluid is conserved during the temporal evolution. This mass conservation property is directly preserved when simulating with finite volumes, since mass is exchanged between the finite volumes but the total mass remains constant. This occurs because, when integrating over each of the finite volumes, the resultant ODEs contain flux terms that are evaluated solely at the boundaries of the finite volume. Hence, since the flux entering a volume is identical to that leaving the adjacent volume, it follows that the total mass of the system is conserved. Another nice property of finite volumes is that the spatial discretization can be made over unstructured polygonal meshes, meaning that one can simulate complex geometries such as a Formula-1 car or the wing of a plane. A further advantage is that implementing boundary conditions is usually straightforward.

When a numerical method is designed to preserve a certain property of the system, such as the conservation of total mass, we say that such method is structure-preserving. The main reason for the success of finite volumes in CFD is that it automatically preserves some of the fundamental properties of fluid-type systems, so we could say that finite volumes are already a structure-preserving scheme per se. However, there may be other properties that the numerical method seems to satisfy in practice when running the simulation, even though there is no proof that the algorithm always preserves them. As an example, let's think about the positivity of the density: in Physics it is not possible to have negative densities, in the same way that there are no negative temperatures below 0°K. A basic finite-volume scheme with a reasonable time step is prone to keep the density positive, as long as the fluid simulation is not too complicated. The issue here is that we cannot guarantee that the density remains positive for all possible simulations – there may be particular configurations that trigger numerical artefacts resulting in negative densities. Consequently, a really appealing scientific objective is to design the finite-volume scheme so that it always keeps the density positive.

In this thesis, I focus on designing structure-preserving finite volumes with the objective of conserving specific properties of a particular family of PDEs. My finite-volume schemes, in addition to satisfying positivity, ensure the preservation of more sophisticated properties such as the decay of the free energy of the system or the maximum principle of a particular quantity (i.e. the quantity has an upper bound). There are however several difficulties to design these schemes. The first one is to translate such properties from the continuum language of PDEs to the discrete domain of algorithms. The second is to mathematically prove that the discrete version of the property is satisfied by the numerical scheme. The third is to design them general enough so that the same scheme is valid for PDEs modelling different applications but sharing a somehow common structure. Overall, structure-preserving schemes are well sought after in the current research of numerical methods, and their advantages are humongous: from greater robustness when designing general-purpose commercial simulation software to less numerical error for simulations with similar computational cost and mesh sizes.

My finite-volume schemes are designed for a general family of PDEs whose main property is the existence of a free energy that decreases when the system evolves. Such free energy accounts for how the agents of the system modelled by the PDE interact – for instance, atoms repel each other at short distances but are attracted at large distances, resulting in an intermediate equilibrium distance that is stable and where the free energy is minimum. That system of two atoms evolves until a equilibrium distance is reached, and during the evolution the free energy decreases until it gets to its minimum value. A more complicated configuration emerges when there are multiples agents all interacting with one another. The key concept here is that the interactions between two agents are modelled via a specific mathematical expression depending, for instance, on the distance between them. With many agents one just needs to compute all the possible interactions of each particle with all the other particles, sum them and proceed to the next time step.

This kind of interacting particle system has sparked a lot of interest in biological problems related to animal behaviour, pedestrian flow, cell adhesion or chemotaxis. The three basic interaction rules, denoted as the first principles of swarming, are attraction, repulsion and alignment or re-orientation of the individuals. With these three simple interactions one can create models for the interaction of particular animals such as birds, fish, ducks, and insects for instance. A really important concept here is the so-called emergent behaviour, which leads to schools of fish, flocks or birds or swarms of insects. The impressive remark here is that such collective behaviour stems just from the one-to-one interactions of each member with all the other members. Another flourishing field where interacting agents play a major role is colloidal fluids, which consist of nano- to micrometer sized particles suspended in a viscous Newtonian fluid (known as the bath). The colloidal particles are much larger than the bath particles, and this creates interesting dynamics due to the different length scales of the particles. Colloidal fluids have plenty of valuable applications such as wetting phenomena like droplet formation, aerosol deposition, cloud formation, and the transport and coagulation of nanoparticles in pulsating flow, such as in the cardiovascular system and oscillatory flow mixing.

Without a doubt these systems of interacting agents lead to a rich variety of applications. Nevertheless, this thesis is focused on algorithms applied to fluid-type equations, where instead of independent particles there is a continuous liquid or gas. What is the connection between these two apparently different systems? It turns out that, in many cases, it is possible to transform a *microscopic* system of interacting particles into a *macroscopic* equation that models the density of a fluid or gas. The mathematical tools to accomplish this are based on taking the limit of the number of particles to infinity, together with other sophisticated assumptions and statistical methods. As a result, describing a system as a fluid only makes sense if the number of particles in it is pretty high. These derivations are usually called mean-field limits, and to go from the microscopic to the macroscopic description there is a transition through an intermediate *mesoscopic* level, which consists of a kinetic equation based on probability distributions. Interestingly, for some systems the fluid-type equation has been established before understanding how to perform the mathematical derivation from the microscopic to the macroscopic level – these descriptions are usually called phenomenological or empirical, since they have not been derived from first principles of interacting particles.

Describing a particle system as a fluid is definitely an interesting abstraction from the mathematical point of view. Are there however any practical advantages in doing so? Definitely, and the reasons for it are precisely quite aligned with the objectives of this thesis: reducing the computational cost of the simulation. In a particle system of n particles there are n ODEs to solve, one per particle. In each of those ODEs we take into account the contribution of the interactions of that precise particle with all the other particles, which results in $n-1$ interactions. Consequently, the order of computations to solve a time step is $\mathcal{O}(n^2)$, and as we go to systems with higher and higher number of particles it follows that such order of computations becomes unbearable. The numerical methods to simulate these particle systems are called Molecular Dynamics, and usually alleviate the computational cost with parallelization techniques and truncations in the interaction range.

Contrary to the system of n ODEs to solve when simulating n particles, in fluid-type systems there are at most three PDES to solve: density, momentum and energy. The cost of solving these three PDEs does not depend on the number of particles in the system, since somehow we are already considering infinite particles in it. There is an additional difficulty though: while the ODEs for particles only have derivatives with respect to time, the PDEs for fluids have derivatives both in time and in space. As a result, one has to combine the temporal discretization of the equations with a spatial discretization of the domain containing the fluid. The finer the spatial mesh is, the more costly and accurate the simulation becomes. In any case, solving a fluid-type equation is always more cost-efficient than solving a system of high-number interacting particles,

and in many applications it is the only option when the order of agents escalates to thousands or millions.

After four pages without a single equation it is high time to present some of the fluid-type PDEs of this thesis. Depending on the macroscopic description, there are three alternatives for the number of PDES to consider: only one PDE for the density, two PDEs for density and momentum, or three PDEs for density, momentum and energy. When only density is considered, the typical PDE to describe the density $\rho(x, t)$ of a fluid, with x being space and t time, satisfies

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \right), \tag{1.1}$$

where I have removed the dependency of ρ with respect to x and t for simplicity. The term $\partial \rho / \partial t$ denotes the partial derivative of ρ with respect to t, and the divergence $\nabla \cdot$ and gradient ∇ contain the partial derivatives of ρ with respect to x. The term $\mathcal{F}[\rho]$ is the free-energy functional of the system, which contains the information about how the particle interaction is produced and depends on the density ρ. The expression $\delta \mathcal{F}[\rho] / \delta \rho$ is just a way of expressing the derivative of the functional $\mathcal{F}[\rho]$ with respect to ρ (usually denoted as the variation of a functional). Finally, the PDE has to be complemented with an initial density $\rho(x, t = 0)$ and some conditions for the density in the boundary of the domain (usually denoted as boundary conditions).

From the mathematical point of view, one can try to prove various properties of the above PDE describing the density evolution. For instance, many research lines are devoted to study whether the solutions are global-in-time, or on the contrary whether at some point in time there is a blow up that prevents the solution from evolving further. In this thesis, I particularly focus on another really fundamental property: the temporal decay of the free energy $\mathcal{F}[\rho]$. A system satisfying this property is usually denoted as a gradient flow or energy-dissipating, and to mathematically prove this one just needs to differentiate the free energy over time and apply boundary conditions, leading to

$$\frac{d}{dt} \mathcal{F}[\rho] = - \int_{\Omega} \rho \left| \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \right|^2 dx, \tag{1.2}$$

where we notice that the right-hand side is negative since the terms inside the integral are positive but there is a negative sign outside, and Ω just denotes that the integral is over the whole domain of the fluid. This property is precisely the one that I aim to discretely preserve with my finite-volume schemes.

The power of the density equation in (1.1) comes from the rich variety of applications it models. The choice of the free-energy functional $\mathcal{F}[\rho]$ determines the interactions between the members of a particular system, and by simply tuning a few parameters one can move from modelling biological cells to simulating the flow of a river or how people interact in an underground station. And since the density equation always satisfies (1.1), it also follows that the temporal decay of the free energy in (1.2) also holds. In this thesis I focus on general free-energies that can be applied in a wide range of multidisciplinary scientific problems. Three examples of such free energies are:

a) Ideal-gas free energy with external and interaction potentials: an ideal gas is just a theoretical gas without any interparticle interactions, but it turns out that many gases such as nitrogen, oxygen, hydrogen or noble gases behave as ideal gas in a wide range of scenarios. Their equation of state is the well-known ideal-gas law with the pressure being proportional to the product of density times temperature. Ideal gases serve as a foundation to model many other applications, which complement the ideal-gas law with customized external and interaction potentials. The free energy for this kind of systems satisfy

$$\mathcal{F}[\rho] = \int_{\Omega} \rho(\log(\rho) - 1) dx + \int_{\Omega} V(x) \rho(x) dx + \frac{1}{2} \int_{\Omega} \int_{\Omega} W(x - y) \rho(x) \rho(y) dx dy,$$

where there are three distinct integrals in it: the first one comes from the ideal-gas law, the second one contains the external potential $V(x)$, and the third one accounts for interparticle interactions with the kernel $W(x-y)$. Different choices of the potentials $V(x)$ and $W(x-y)$ lead to radically different applications. For instance, this free energy can model chemotaxis, which is usually defined as the directed movement of cells and organisms in response to chemical gradients [177]. However, this free energy also finds its applications in astrophysics and gravitation [113, 261].

b) Shallow-water free energy: the shallow-water equations model free-surface gravity waves whose wavelength is much larger than the characteristic bottom depth. These equations are applied in a wide range of engineering and scientific applications involving free-surface flows [282], such as tsunami propagation [78], dam break and flooding problems [98] and the evolution of rivers and coastal areas [93]. Their free energy corresponds to

$$\mathcal{F}[\rho] = \int_\Omega \rho^2 dx + \int_\Omega V(x)\rho(x)dx.$$

The first integral contains ρ^2, instead of the ideal-gas term $\rho(\log(\rho)-1)$. This leads to the presence of dry regions during the water-height evolution. The external potential $V(x)$ now accounts for the topography of the shallow-water system, and can be interpreted as the depth of the bottom at point x measured from a fixed level of reference.

c) Cahn-Hilliard free energy: it was firstly proposed by Cahn and Hilliard [48] to describe the process of phase separation in binary alloys. Since then it has found applicability in a wide variety of modelling problems, from capillarity and wetting phenomena [13, 258], diblock copolymer molecules [298] and tumour growth [278] to image inpainting [33, 65], topology optimization [300] and many more [180]. The free energy for all these applications satisfies

$$\mathcal{F}[\rho] = \int_\Omega \left(H(\rho) + \frac{\epsilon^2}{2} |\nabla\rho|^2 \right) d\Omega + \int_{\partial\Omega} f_w(\rho,\beta)ds.$$

In the first integral there is a competition between the hydrophobic $H(\rho)$ and hydrophilic $\frac{\epsilon^2}{2}|\nabla\rho|^2$ terms: on the one hand, $H(\rho)$ is a potential whose minima correspond to the stable phases of the system; on the other hand, $\frac{\epsilon^2}{2}|\nabla\rho|^2$ penalizes for the existence of interfaces between the stable phases. The last integral accounts for the wall free energy, which depends on the density at the wall and the equilibrium contact angle β.

The next sections of this introduction present the basic concepts of two fundamental tools of this thesis: finite-volume schemes and gradient-flow/energy-dissipating equations. First, in Chapter 2 I explain the finite-volume methodology to discretize in space, together with the temporal integrators to advance the system in time. Second, in Chapter 3 I focus on mathematically describing the PDEs solved in this thesis, placing special emphasis on their property to dissipate the free energy in time.

The rest of this section is devoted to summarizing the contributions of the main parts of this thesis. There are three core parts: Part II is focused on well-balanced finite-volume schemes for hydrodynamic systems, Part III deals with finite-volume schemes for the Cahn-Hilliard equation and its application to image inpainting, and Part IV presents finite-volume schemes for fluctuating hydrodynamics. Finally, Part V discusses some conclusions and future perspectives of the research in this thesis.

1.1 Summary of Part II: well-balanced schemes for hydrodynamic systems

We present well-balanced first-, second- and high-order finite-volume schemes for a general class of hydrodynamic systems with linear and nonlinear damping. The variation of the natural Liapunov functional of the system, given by its free energy, allows for a characterization of the stationary states by its variation. An analogue property at the discrete level enables us to preserve stationary states at machine precision, while discretely preserving the dissipation of the discrete free energy for first- and second-order accuracy. Our schemes are also equipped with a nonnegative-density reconstruction which allows for vacuum regions during the simulation. These schemes can robustly analyse the stability properties of stationary states in challenging problems such as: phase transitions in collective behaviour, generalized Euler-Poisson systems in chemotaxis and astrophysics, and models in dynamic density functional theories. We provide several prototypical examples from relevant applications highlighting the benefit of our algorithms elucidate also some of our analytical results.

1.2 Summary of Part III: the Cahn-Hilliard equation

We propose finite-volume schemes for the Cahn-Hilliard equation that unconditionally and discretely satisfy the boundedness of the phase field and the free-energy dissipation. Our numerical framework is applicable to a variety of free-energy potentials including the Ginzburg-Landau and Flory-Huggins, general wetting boundary conditions and degenerate mobilities. Its central thrust is the finite-volume upwind methodology, which we combine with a semi-implicit formulation based on the classical convex-splitting approach for the free-energy terms. Extension to an arbitrary number of dimensions is straightforward thanks to their cost-saving dimensional-splitting nature, which allows to efficiently solve higher-dimensional simulations with a simple parallelization. The numerical schemes are validated and tested in a variety of prototypical configurations with different numbers of dimensions and a rich variety of contact angles between droplets and substrates.

In addition, our Cahn-Hilliard scheme is employed as an image inpainting filter before passing damaged images into a classification neural network. The benchmark dataset employed here is MNIST, which consists of binary images of handwritten digits and is a standard dataset to validate image-processing methodologies. We train a neural network based of dense layers with the training set of MNIST, and subsequently we contaminate the test set with damage of different types and intensities. We then compare the prediction accuracy of the neural network with and without applying the Cahn-Hilliard filter to the damaged images test. Our results quantify the significant improvement of damaged-image prediction due to applying the Cahn-Hilliard filter, which for specific damages can increase up to 50% and is in general advantageous for low to moderate damage.

1.3 Summary of Part IV: finite-volume schemes for fluctuating hydrodynamics

We introduce a finite-volume numerical scheme for solving stochastic gradient-flow equations. Such equations are of crucial importance within the framework of fluctuating hydrodynamics and dynamic density functional theory. Our proposed scheme deals with general free-energy functionals, including, for instance, external fields or interaction potentials. This allows us to simulate a range of physical phenomena where thermal fluctuations play a crucial role, such as nucleation and other energy-barrier crossing transitions. A positivity-preserving algorithm for the density is derived

based on a hybrid space discretization of the deterministic and the stochastic terms and different implicit and explicit time integrators. We show through numerous applications that not only our scheme is able to accurately reproduce the statistical properties (structure factor and correlations) of the physical system, but, because of the multiplicative noise, it allows us to simulate energy barrier crossing dynamics, which cannot be captured by mean-field approaches.

Finite-volume schemes

Finite volumes are a commonplace numerical technique to solve PDEs such as conservation laws and hyperbolic systems. In finite volumes, the spatial domain is discretized in small or "finite" volumes, and in each of them we evaluate the average value of the solution. Then the PDEs are integrated over each of those finite volumes, yielding a system of ODEs for the evolution of each of the average values. Such system of ODEs is then solved in time by applying a suitable temporal discretization. Finite volumes are especially popular in fluid dynamics, where the PDEs contain divergence terms that can be evaluated at the boundaries after employing Gauss divergence theorem. Those terms can be interpreted as fluxes, and since a flux entering a cell is the same as the flux leaving the adjacent cell, it follows that finite volume schemes are conservative. They also adapt seamlessly to complex geometries or unstructured meshes, and the application of boundary conditions is usually straightforward.

In this chapter, I provide the foundations to understand the finite-volume schemes of this thesis. This is just an introduction, and I would like to point out several books that have greatly helped me to better grasp the concepts behind finite volumes and more generally numerical schemes for PDEs:

- [196] from LeVeque is an excellent and complete introductory book into finite-volume schemes, which in addition explains many advanced concepts and numerical recipes.

- [188] from Laney is another fantastic introductory book, with a strong emphasis into the fundamentals of numerical analysis.

- [254] from Shu is a clear and concise set of notes about high-order schemes based on ENO and WENO reconstructions. It also provides many numerical test cases to validate the order of the schemes.

- [269] from Toro offers a broad overview of numerical techniques for hyperbolic systems.

In this introduction, I focus on the following one-dimensional (1D) equation that models the evolution of the density ρ of a fluid:

$$\partial_t \rho(x,t) + \partial_x F(\rho(x,t)) = S(\rho(x,t), x). \tag{2.1}$$

This equation is basically an advection equation with a source term. From now on I omit the dependency of ρ with respect to x and t. $\partial_t \rho$ is the partial derivative of the density with respect to time, $\partial_x F(\rho)$ is the spatial partial derivative of a flux function $F(\rho)$, which depends on the density

ρ, and $S(\rho, x)$ is a source term depending on the density ρ and space x. Note that the steady state of ρ satisfies $\partial_t \rho = 0$, and in that scenario the advective term has to be balanced out with the source term and it follows that $\partial_x F(\rho) = S(\rho, x)$.

2.1 General formulation

The 1D finite-volume approximation of (2.1) begins by dividing the computational domain $[0, L]$ into N cells $C_i = [x_{i-1/2}, x_{i+1/2}]$, all with uniform size $\Delta x = L/N$, so that the centres of the cells satisfy $x_i = (i-1)\Delta x + \Delta x/2$, $i \in 1, \ldots, N$. A schematic of the finite-volume discretization is displayed in Figure 2.1.
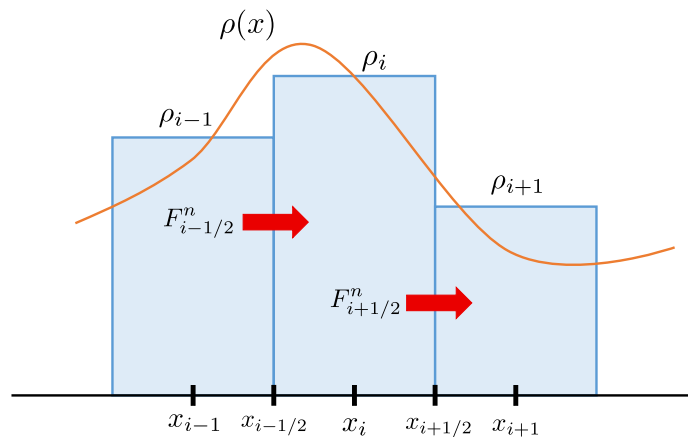
Figure 2.1: Schematic of the finite-volume discretization of a density field $\rho(x)$ with the fluxes $F_{i-1/2}^n$ and $F_{i+1/2}^n$ at the cell interfaces.

In each of the cells C_i we define the cell average ρ_i as

$$\rho_i(t) = \frac{1}{\Delta x} \int_{C_i} \rho(x, t) dx. \tag{2.2}$$

Numerically, this integral has to be computed via quadrature rules satisfying a particular order of accuracy. The simplest choice is to evaluate $\rho_i(t)$ as the value of $\rho(x, t)$ at x_i, which is the midpoint of the cell C_i. In that case the error is of order $\mathcal{O}(\Delta x^2)$ for smooth functions, and consequently for first- and second-order finite-volume schemes one need not care about applying quadrature rules to the cell-average integral: it is enough with taking the midpoint value directly. This is also why up to second order the finite-difference and finite-volume formulations are usually identical. For higher-order schemes this is not true though.

Subsequently, one has to integrate (2.1) over each of the cells C_i, resulting in

$$\int_{C_i} \partial_t \rho \, dx + \int_{C_i} \partial_x F(\rho) \, dx = \int_{C_i} S(\rho, x) \, dx. \tag{2.3}$$

Let's analyse each of the integrals. Firstly, the integral with $\partial_t \rho$ leads to

$$\int_{C_i} \partial_t \rho \, dx = \Delta x \frac{d\rho_i(t)}{dt}, \tag{2.4}$$

where the cell-average definition in (2.2) is applied. Secondly, the integral of $\partial_x F(\rho)$ yields

$$\int_{C_i} \partial_x F(\rho) \, dx = F(\rho(x_{i+1/2}, t)) - F(\rho(x_{i-1/2}, t)), \tag{2.5}$$

where the Gauss divergence theorem allows to evaluate the integral by only considering the flux values at the boundaries of the cell. Finally, the integral with $S(\rho, x)$ cannot be simplified without further assumptions, so it remains as it is. It follows that the integral equation in (2.3) results in the semi-discrete scheme

$$\frac{d\rho_i(t)}{dt} = -\frac{1}{\Delta x}\left(F(\rho(x_{i+1/2}, t)) - F(\rho(x_{i-1/2}, t))\right) + \frac{1}{\Delta x}\int_{C_i} S(\rho, x)\, dx \tag{2.6}$$

for each of the cells C_i. Now the PDE in (2.2) has turned into a set of N ODEs, one per cell and accounting for the evolution of the average density. The next step is to integrate in time from t_n to t_{n+1}, with the time step defined as $\Delta t = t_{n+1} - t_n$. The integration in time results in

$$\begin{aligned}
\rho_i^{n+1} - \rho_i^n = &-\frac{1}{\Delta x}\left(\int_{t_n}^{t_{n+1}} F(\rho(x_{i+1/2}, t))\, dt - \int_{t_n}^{t_{n+1}} F(\rho(x_{i-1/2}, t))\, dt\right) \\
&+ \frac{1}{\Delta x}\int_{t_n}^{t_{n+1}}\int_{C_i} S(\rho, x)\, dx\, dt;
\end{aligned} \tag{2.7}$$

where ρ_i^{n+1} and ρ_i^n are just the evaluations of ρ_i at t_{n+1} and t_n, respectively, and result from applying the Gauss divergence theorem.

The temporal integrals in (2.7) can be evaluated in an explicit, implicit or even semi-implicit fashion. With an explicit approximation it follows that

$$\int_{t_n}^{t_{n+1}} F(\rho(x_{i+1/2}, t))\, dt \approx \Delta t\, F(\rho(x_{i+1/2}, t_n)), \tag{2.8}$$

while for the implicit approximation one takes

$$\int_{t_n}^{t_{n+1}} F(\rho(x_{i+1/2}, t))\, dt \approx \Delta t\, F(\rho(x_{i+1/2}, t_{n+1})). \tag{2.9}$$

In semi-implicit approximations there may be an interpolation between the evaluations at t_n and t_{n+1}, or simply some terms of $F(\rho(x_{i+1/2}, t)$ are evaluated at t_n and others at t_{n+1}. Generally implicit schemes are considered more stable than explicit schemes and allow to employ larger time steps. However, their computational cost is higher since they require the resolution of a system of equations, and in addition they may introduce some numerical diffusion.

At this point let's just consider an explicit approximation in time. For simplicity the following notation is introduced:

$$F_{i+1/2}^n := F(\rho(x_{i+1/2}, t_n)), \quad F_{i-1/2}^n := F(\rho(x_{i-1/2}, t_n)). \tag{2.10}$$

The fluxes $F_{i+1/2}^n$ and $F_{i-1/2}^n$ are evaluated from the cell averages ρ_i. How many cells should be considered for their evaluations? It turns out that only the cell averages directly at the left and at the right of the boundary are needed, and the reason for this comes from the finite-speed propagation of information in hyperbolic systems such as (2.1). Since information travels at a certain speed, one can limit the time step to ensure that only the information from the cells at the left and at the right has time to arrive at the boundary. This is the essence of the CFL condition for stability, as explained in Section 2.5. Consequently, for a first-order discretization the fluxes satisfy

$$F_{i+1/2}^n = \mathcal{F}(\rho_i^n, \rho_{i+1}^n), \tag{2.11}$$

where $\mathcal{F}$ is the so-called numerical flux. More generally, the numerical flux depends on a reconstruction of the density at the left of the boundary, $\rho_{i+1/2}^{n,-}$, and at the right of the boundary, $\rho_{i+1/2}^{n,+}$, so that

$$F_{i+1/2}^n = \mathcal{F}(\rho_{i+1/2}^{n,-}, \rho_{i+1/2}^{n,+}). \tag{2.12}$$

The numerical flux $\mathcal{F}$ is usually denoted as Riemann solver, since it provides a stable resolution of the Riemann problem located at the cell interfaces. The literature concerning Riemann solvers is vast and there are different choices for it [269]: Godunov, Lax-Friedrich, kinetic, Roe, etc.

2.2 Boundary reconstructions

The numerical flux in (2.12) relies on the reconstruction of the density at the left of the boundary, $\rho_{i+1/2}^{n,-}$, and at the right of the boundary, $\rho_{i+1/2}^{n,+}$. Such reconstruction is computed from an interpolation polynomial that depends on the cell averages surrounding the boundary. The higher the order of the polynomial, the more coefficients one needs to estimate, and as a result the more cell averages ought to be considered. The spatial accuracy of a finite-volume scheme is directly linked to the order of the reconstruction polynomial. Usually, as the order of the reconstruction polynomial increases, the spatial order of the finite-volume scheme also grows.

Furthermore, the reconstruction polynomial can be tuned to ensure the preservation of discrete properties of the finite-volume scheme. For instance, let's think of the positivity of the density: it may occur that the reconstruction returns a negative value of the density at the boundary, due to a sharp local variation in the cell averages. To prevent this, the coefficients of the polynomial are modified, resulting in a polynomial that maintains its order of accuracy but ensures that all the reconstructed values are positive. The same philosophy applies to other more sophisticated properties, such as the well-balanced schemes presented in detail in Part II.

In this section, I review some relevant first-, second- and high-order reconstructions that lay the foundations for the contents of this thesis. For a more complete list of reconstruction techniques I refer the reader to [196].

2.2.1 First-order reconstruction

In a first-order reconstruction the left and right values are simply taken as

$$\rho_{i+1/2}^{n,-} = \rho_i, \quad \rho_{i+1/2}^{n,+} = \rho_{i+1}. \tag{2.13}$$

The reconstruction polynomial follows a piecewise constant distribution, with the constant value in each cell corresponding to the cell average. This is the most basic method to compute the reconstructions.

2.2.2 Second-order MUSCL reconstruction

MUSCL stands for Monotonic Upstream-centred Scheme for Conservation Laws, and this reconstruction was firstly proposed in the seminal paper of Van Leer in 1979 [272]. The reconstruction procedure relies on three steps: prediction of the gradients in each cell, linear extrapolation and limiting procedure to preserve nonnegativity of the density.

The reconstructions at the left and at the right of the boundary are computed as

$$\rho_{i+1/2}^{n,-} = \rho_i + \frac{\phi(r_i)}{2}\left(\rho_{i+1} - \rho_i\right), \quad \rho_{i+1/2}^{n,+} = \rho_{i+1} - \frac{\phi(r_{i+1})}{2}\left(\rho_{i+2} - \rho_{i+1}\right). \tag{2.14}$$

$\phi(r_i)$ is a flux limiter whose objective is to limit the slope of the reconstruction. This is necessary to avoid spurious oscillations that may arise with shocks, discontinuities of sharp variations in the density profile. The input of the flux limiter is r_i, which represents the ratio of successive gradients of the density:

$$r_i = \frac{\rho_i - \rho_{i-1}}{\rho_{i+1} - \rho_i}. \tag{2.15}$$

The ratio r_i is negative if the cell average ρ_i is a local minimum or maximum in the stencil of cell averages $\{\rho_{i-1}, \rho_i, \rho_{i+1}\}$. Another challenging scenario is when r_i has a high value due to the presence of sharp gradients. In these situations the flux limiter ensures that the reconstruction comes back to the first-order choice in (2.13). There are many proposed flux functions in the literature, and one of the most extended ones is the minmod limiter:

$$\phi(r_i) = \max[0, \min(1, r_i)]. \tag{2.16}$$

Consequently, the MUSCL reconstruction is second order in the regions of the density profile that are smooth. It however returns to a first-order reconstruction in those regions of the domain where there is a risk of generating spurious oscillations.

2.2.3 High-order WENO-like reconstruction

Weighted Essentially Non-Oscillatory (WENO) schemes are a class of high-order reconstructions that are designed to prevent the generation of spurious oscillations while maintaining a high-order accuracy. Their key insight is to employ a nonlinear adaptive procedure to automatically choose the locally smoothest stencil, hence avoiding crossing discontinuities in the interpolation procedure as much as possible. There are WENO schemes of varied orders of accuracy, and in this thesis I focus on third- and fifth-order reconstructions. In addition, certain modified WENO schemes such as the Compact Weighted Essentially Non-Oscillatory (CWENO) reconstruction operators [52, 197, 198] offer crucial advantages when applied together with finite-volume schemes. The main advantage of CWENO compared to WENO (see [253, 256, 257]) reconstruction operators is that CWENO reconstructions achieve uniform high-order approximation in the entire cell, while WENO reconstruction operators are proposed to achieve high-order approximation at the boundaries of the cell. Thus, standard WENO-5 reconstructions achieves fifth-order at the boundaries of the cell, while it is only third-order at the interior points. Therefore, CWENO reconstruction operators are specially useful in balance laws such as (2.1), where the source term has to be evaluated at inner points of the cell.

Given a family of density values $\{\rho_i(t)\}$, a WENO-type reconstruction provides at every cell $[x_{i-1/2}, x_{i+1/2}]$ a smooth function that depends on the values at some neighbour cells whose indexes belong to the so-called stencil $\mathcal{S}_i$:

$$R_i^\rho(x) = R_i^\rho(x; \{\rho_j(t)\}_{j \in \mathcal{S}_i}),$$

so that $R_i^\rho(x)$ is a high-order approximation of $\rho(x, t)$ in the cell C_i at time t. The reconstructions at the left and at the right of the boundary are computed by evaluating the reconstructed polynomial:

$$\rho_{i+1/2}^{n,-} = R_i^\rho(x_{i+1/2}), \quad \rho_{i+1/2}^{n,+} = R_{i+1}^\rho(x_{i+1/2}). \tag{2.17}$$

WENO schemes can be complemented with the positive-density limiters [299] to ensure physical admissible reconstructed values for the density. The WENO schemes applied in this thesis are thoroughly detailed in Appendix 6.B within Chapter 6, where third- and fifth-order CWENO reconstructions are employed to design high-order well-balanced schemes.

2.3 Numerical flux

A numerical flux aims to correctly approximate the physical flux by employing the cell-average density values. The design of numerical fluxes has been a thriving research field in the last decades, and the reader can find excellent summaries of the state of the art in [188, 196]. An essential property for a numerical flux is convergence, which implies that the numerical solution should converge to the true solution as the grid is refined both in time and in space. Convergence requires two conditions:

1. Consistency with the physical flux, so that $\mathcal{F}(\rho, \rho) = F(\rho)$.

2. Stability in some appropriate sense, meaning that small errors don't accumulate in time.

Another desirable property is that the numerical flux preserves the nonnegativity of the density $\rho_i(t)$ for the homogeneous problem. More sophisticated properties are Total Variation Diminishing (TVD) or monotonicity.

Numerical fluxes typically rely on two inputs: the reconstructed density values at the left and at the right of the boundary. Such reconstructed values can be computed following Section 2.2. The next step is to devise a function that, given those two values, correctly approximates the physical flux. A first naive proposal is to just evaluate the physical flux with the average of the two input values. This however turns out to be unstable even for simple physical fluxes, and one can show this by conducting a von Neumann stability analysis (see section 4.6 in [196] for further details).

For simplicity, in this section I assume that the physical flux satisfies $F(\rho) = u\rho$, where u is a constant velocity. This flux is typical in advection problems whose solutions satisfy

$$\rho(x, t) = \widetilde{\rho}(x - ut) \tag{2.18}$$

for some $\widetilde{\rho}$. This implies that the density values just translate with constant velocity u. For this choice of flux I proceed to explain two of the most versatile numerical fluxes: upwind and Lax-Friedrichs.

2.3.1 Upwind flux

Upwind methods are especially suited for hyperbolic equations such as (2.1), where the information is propagated as waves following certain characteristics. For the choice of flux $F(\rho) = u\rho$, the information is propagated at constant velocity u. The sign of the velocity plays a vital role here: if the velocity is positive, the information that arrives at the boundary $x_{i+1/2}$ comes from the cell C_i, whereas for negative velocities the information comes from the cell C_{i+1}. Upwind numerical fluxes take advantage of this knowledge about the structure of the solution to approximate the physical flux.

The upwind numerical flux when $F(\rho) = u\rho$ satisfies

$$F_{i+1/2}^n = \mathcal{F}\left(\rho_{i+1/2}^{n,-}, \rho_{i+1/2}^{n,+}\right) = u_+ \, \rho_{i+1/2}^{n,-} + u_- \, \rho_{i+1/2}^{n,+} \tag{2.19}$$

where

$$u_+ = \max(u, 0), \quad u_- = \min(u, 0).$$

Clearly only one of the summands in (2.19) is nonzero, depending on the sign of the velocity u. For more complicated systems with more than one equation there are several waves propagating at different speeds and directions, and the usual technique for upwind relies on performing a characteristic decomposition to evaluate from which side is the information coming from.

2.3.2 Lax-Friedrichs flux

The Lax-Friedrichs scheme is one of the most popular and versatile numerical fluxes. It is based on adding numerical diffusion to the naive numerical flux based on just taking the average of the flux of the two input values. As described in the introduction of this section, a numerical flux with just the average is unstable, but it turns out that the addition of numerical diffusion dampens the instabilities and renders the numerical flux stable as long as a CFL condition is satisfied.

The classical version of the Lax-Friedrichs numerical flux reads

$$F_{i+1/2} = \mathcal{F}\left(\rho_{i+1/2}^{n,-}, \rho_{i+1/2}^{n,+}\right) = \frac{1}{2}\left(F\left(\rho_{i+1/2}^{n,-}\right) + F\left(\rho_{i+1/2}^{n,+}\right)\right) - \frac{\Delta x}{2\Delta t}\left(\rho_{i+1/2}^{n,+} - \rho_{i+1/2}^{n,-}\right), \quad (2.20)$$

where the last term is added to provide numerical diffusion. Note that this numerical flux is consistent with the physical flux since the last term vanishes as the mesh is refined.

Nevertheless, one of the main drawbacks of the classical Lax-Friedrichs flux is the addition of too much diffusivity, leading to numerical results that are badly smeared unless the mesh is really fine. The numerical diffusivity can be reduced by modifying the coefficient $\Delta x/\Delta t$ to be a locally determined value, so that the numerical flux satisfies

$$F_{i+1/2} = \mathcal{F}\left(\rho_{i+1/2}^{n,-}, \rho_{i+1/2}^{n,+}\right) = \frac{1}{2}\left(F\left(\rho_{i+1/2}^{n,-}\right) + F\left(\rho_{i+1/2}^{n,+}\right)\right) - \frac{\lambda_{i+\frac{1}{2}}}{2}\left(\rho_{i+1/2}^{n,+} - \rho_{i+1/2}^{n,-}\right), \quad (2.21)$$

where $\lambda_{i+\frac{1}{2}}$ is taken as the maximum of the absolute value of the eigenvalues of the system,

$$\lambda_{i+\frac{1}{2}} = \max\left(\left|F'(\rho_{i+1/2}^{n,-})\right|, \left|F'(\rho_{i+1/2}^{n,+})\right|\right). \quad (2.22)$$

This maximum is taken locally for every node, resulting in different values of $\lambda_{i+\frac{1}{2}}$ in each of the boundaries. This method is denoted as Rusanov's method or just local Lax-Friedrichs. Choosing a local viscosity coefficient instead of a global one reduces the numerical diffusivity of the scheme.

2.4 Temporal discretization

Finite volumes turn the PDE in (2.1) into a system of ODEs, each of them accounting for the evolution of the average density in a particular cell. The next step is to discretize the ODEs in time by applying an appropriate temporal discretization. This can be achieved by several means, and the method to choose depends on the trade-off between computational cost, temporal order of accuracy and stability. In this section, I review some of the temporal discretizations employed in this thesis.

2.4.1 Forward Euler explicit scheme

This is the most basic first-order temporal integrator to solve an ODE. Due to its explicit nature all the terms at the right-hand side of (2.7) are evaluated at $t = t_n$, leading to

$$\rho_i^{n+1} - \rho_i^n = -\frac{\Delta t}{\Delta x}\left(F_{i+1/2}^n - F_{i-1/2}^n\right) + \frac{\Delta t}{\Delta x}\int_{C_i} S(\rho(x, t_n), x)\, dx\, dt. \quad (2.23)$$

All the terms at the right-hand side can be straightforwardly computed from the available values of the density cell averages at $t = t_n$. However, this scheme can become unstable for combinations of Δx and Δt that lay outside of its stability region. This can be prevented by ensuring that a CFL condition is satisfied, as explained in Section 2.5.

2.4.2 Backwards Euler implicit scheme

This method is also one of the most popular schemes to solve an ODE. Contrary to the explicit version, the implicit technique evaluates the right-hand side of (2.7) at $t = t_{n+1}$, resulting in

$$\rho_i^{n+1} - \rho_i^n = -\frac{\Delta t}{\Delta x}\left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1}\right) + \frac{\Delta t}{\Delta x}\int_{C_i} S(\rho(x, t_{n+1}), x)\, dx\, dt. \quad (2.24)$$

The inconvenience of implicit methods comes from the fact that the cell averages at $t = t_{n+1}$ are unknown, and as a result one needs to solve a (non)linear system of equations to advance in time. This increase in computational cost is compensated with much better stability properties, and in many cases one can prove that the implicit Euler scheme is unconditionally stable for all choices of Δt.

2.4.3 Multi-stage Runge-Kutta

Euler-type schemes, either explicit or implicit, are limited to first-order accuracy in time. The family of Runge-Kutta methods employs subsequent stages to achieve a desired accuracy, but relying only on the values at $t = t_n$. There are many types of Runge-Kutta methods, including explicit, implicit or even stochastic. For instance, the classical second-order Runge-Kutta applied to $\partial_t \rho = L(\rho)$, with $L(\rho)$ defined as $L(\rho) = S(\rho, x) - \partial_x F(\rho)$ following (2.1), reads

$$
\begin{cases}
\rho^{n+1/2} - \rho^n = -\frac{\Delta t}{2} L(\rho^n), \\
\rho^{n+1} - \rho^n = -\Delta t\, L(\rho^{n+1/2}).
\end{cases}
\tag{2.25}
$$

Runge-Kutta schemes are quite flexible, and have often been adapted to satisfy specific properties. In particular, a relevant property to achieve high-order accuracy with finite volumes is Total Variation Diminishing (TVD). With it, the total variation of the solution does not increase, meaning that new extrema are not generated. A related property is Total Variation Bounded (TVB), in which the growth of the total variation is bounded. These two properties enhance the stability of finite-volume schemes. As an example, the third-order TVD Runge-Kutta scheme proposed in [156] satisfies

$$
\begin{cases}
\rho^{(1)} = \rho^n + \Delta t L\left(\rho^n\right), \\
\rho^{(2)} = \frac{3}{4}\rho^n + \frac{1}{4}\rho^{(1)} + \frac{1}{4}\Delta t L\left(\rho^{(1)}\right), \\
\rho^{n+1} = \frac{1}{3}\rho^n + \frac{2}{3}\rho^{(2)} + \frac{2}{3}\Delta t L\left(\rho^{(2)}\right).
\end{cases}
\tag{2.26}
$$

2.5 CFL condition

The Courant-Friedrichs-Levy (CFL) condition is a fundamental concept when implementing finite-volume schemes in hyperbolic equations such as (2.1). Basically, it is a necessary condition that the time step has to satisfy for stability, and imposes a limit to the maximum size of the time step. The CFL condition is essential when implementing numerical fluxes such as the ones described in Section 2.3: in particular, they ensure that the information that arrives at the boundaries only comes from the directly adjacent cells, and not from farther cells.

The exact form of the CFL condition depends on the problem under consideration and the choice of numerical flux. As an example, here I prove a CFL condition for the positivity of (2.1) with $F(\rho) = u\rho$, where u is a particular constant velocity. The discretization is completed with the upwind flux in Subsection 2.3.1, the first-order reconstruction in Subsection 2.2.1 and the explicit Euler scheme in Subsection 2.4.1. For simplicity I assume that there is no source term, so that $S(\rho, x) = 0$. The finite-volume scheme for this configuration satisfies

$$
\rho_i^{n+1} = \rho_i^n - \frac{\Delta t}{\Delta x}\left(F_{i+1/2}^n - F_{i-1/2}^n\right),
\tag{2.27}
$$

with the numerical flux satisfying

$$
F_{i+1/2}^n = u_+ \,\rho_i^n + u_- \,\rho_{i+1}^n, \quad u_+ = \max(u, 0), \quad u_- = \min(u, 0).
\tag{2.28}
$$

Theorem 2.5.1 (CFL condition)**.** *Consider the discrete scheme* (2.27) *complemented with the upwind numerical flux in* (2.28)*. Then, given that* $\rho_i^n \geq 0 \; \forall i$*, it follows that* $\rho_i^{n+1} \geq 0 \; \forall i$ *provided that the following CFL condition is satisfied:*

$$\Delta t \leq \frac{\Delta x}{\max_i \left(u^+ - u^-\right)}. \tag{2.29}$$

Proof. After plugging the numerical flux (2.28) into the finite-volume scheme (2.27) it follows that

$$\begin{aligned}
\rho_i^{n+1} =& \rho_i^n - \frac{\Delta t}{\Delta x}\left(u_+ \, \rho_i^n + u_- \, \rho_{i+1}^n - u_+ \, \rho_{i-1}^n - u_- \, \rho_i^n\right) \\
=& \left(1 - \frac{\Delta t}{\Delta x}u_+ + \frac{\Delta t}{\Delta x}u_-\right)\rho_i^n - \frac{\Delta t}{\Delta x}u_- \, \rho_{i+1}^n + \frac{\Delta t}{\Delta x}u_+ \, \rho_{i-1}^n.
\end{aligned}$$

The second and third terms on the right hand side are positive since $u^+ \geq 0$ and $u^- \leq 0$ from (2.28). The first term is also positive provided that the CFL condition (2.29) is satisfied. As a result, $\rho_i^{n+1} \geq 0 \; \forall i$ as long as the CFL condition (2.29) holds. $\qquad\square$

Gradient-flow and energy-dissipating equations

All the systems studied in this thesis share a common property: they are driven by an energy that decreases as the system evolves. Such decrease is driven by a dissipation mechanism, which ensures that the system eventually reaches a steady state where the energy is minimized. The striking fact about the energy and dissipation mechanism is that they are not determined by the PDE(s) of the system: depending on the application, there may be many choices of the pair energy-dissipation mechanism that lead to the same PDE(s). When the system under consideration is formed by a single PDE, typically describing the evolution of the density, the system is said to be a gradient flow. If the system is formed by multiple PDE, such as the density and momentum equations of hydrodynamic systems, the system is said to be energy-dissipating.

For instance, let's focus on the classical heat or diffusion equation:

$$\frac{\partial \rho}{\partial t} = \Delta \rho.$$

This equation can be understood as a gradient flow that decays an energy $\mathcal{F}[\rho]$, whose form depends on the metric under consideration. Under the L^2 norm the equation can be rewritten as

$$\frac{\partial \rho}{\partial t} = -\nabla_{L^2(\mathbb{R}^d)} \mathcal{F}[\rho] = -\frac{\delta \mathcal{F}[\rho]}{\delta \rho},$$

with the choice of the free energy $\mathcal{F}[\rho]$ equal to

$$\mathcal{F}[\rho] = \frac{1}{2} \int_\Omega |\nabla \rho|^2 \, d\boldsymbol{x}.$$

But the heat equation in (3) can also be written by employing a Wasserstein norm, so that it reads

$$\frac{\partial \rho}{\partial t} = -\nabla_{\mathcal{W}^2} \mathcal{F}[\rho] = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \right),$$

with the free energy $\mathcal{F}[\rho]$ now being

$$\mathcal{F}[\rho] = \int_\Omega \rho \left(\log(\rho) - 1 \right) d\boldsymbol{x}.$$

One can check that with both metrics the result is the same, and the original heat equation (3) is recovered after plugging in the free energy and computing its variation with respect to the density.

Interestingly, the heat equation can be understood as a gradient flow under many other metrics, and I refer the reader to the excellent notes of Marc Peletier about this topic [229].

The description of the heat equation as a gradient flow under the L^2 norm and other metrics has been known for a really long time. What really sprang a new wave of interests in gradient flows was the realization that continuity equations such as $\partial_t \rho + \nabla \cdot (\rho u) = 0$ can be interpreted as gradient flows under a Wasserstein metric, with the velocity u given by $u = \nabla \delta \mathcal{F}[\rho]/\delta \rho$. This discovery is quite recent, and it was thanks to Jordan, Kinderlehrer and Otto [170, 225] in the turn from the 20$^\text{th}$ to the 21$^\text{st}$ century. Their inspiring insight became formal and general with the work of Ambrosio, Gigli and Savaré [4], and heavily relies on the theory of optimal transport [273]. A nice recent review about the topic is [244]. The number of applications relying on continuity equations is vast, and the application of the concept of Wasserstein gradient flow has fostered much progress both in analytical results and numerical schemes.

Some of the applications with a structure of Wasserstein gradient flow are:

- Convection and nonlinear diffusion [170], which appear in certain Fokker-Planck equations of the form:
$$\partial_t \rho = \nabla \left(\rho \nabla (\Pi'(\rho) + V(\boldsymbol{x})) \right).$$

- Aggregation-diffusion equations and granular flows with interactions [68]:
$$\partial_t \rho = \nabla \left(\rho \nabla (\Pi'(\rho) + W(\boldsymbol{x}) \star \rho) \right).$$

- Thin-film equations [171]:
$$\partial_t \rho = -\nabla \left(\rho \nabla (\Delta \rho) \right).$$

- The Cahn-Hilliard equation [13]:
$$\partial_t \rho = \nabla \left(M(\rho) \nabla (-\Delta \rho + H'(\rho)) \right).$$

The notation introduced here is carefully explained in the different chapters of this thesis. At this point I just remark that each of these applications possesses a distinct type of free energy, but all of them share a common PDE structure based on a gradient flow under a Wasserstein metric. Precisely, it is thanks to this common structure that it is possible to design general finite-volume schemes applicable to all of them. The schemes designed in this thesis replicate the main properties of Wasserstein gradient flows at a discrete level, such as the dissipation of the free energy.

In this thesis there are mainly two types of systems: gradient-flow overdamped equations and energy-dissipating hyperbolic systems. The former is usually derived from the latter in the strong damping limit, and in many cases both of them can be derived from many-particle systems via techniques such as mean-field limits or Density Functional Theory (DFT). The concept of gradient flow only applies to the overdamped equations. Nevertheless, the hyperbolic systems studied here also possess a free energy with the same structure and interpretation as in the overdamped equation. The dissipation mechanism is also similar, but in addition involves a kinetic energy due to the presence of momentum in the system. Another property that applies to hyperbolic systems is the existence of an entropy identity. Section 3.1 of this introduction is devoted to presenting the properties of the gradient-flow overdamped equations in this thesis, and Section 3.2 focuses on energy-dissipating hyperbolic systems.

3.1 Gradient-flow overdamped equation

Many applications are modelled with a PDE describing the time evolution of the density. In particular, in this thesis we are interested in PDEs with the following structure:

$$\frac{\partial \rho}{\partial t} = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \right), \tag{3.1}$$

where ρ is the density and $\mathcal{F}$ is the free energy of the system. Equation (3.1) possesses a variational structure and it is a gradient flow for the free energy $\mathcal{F}[\rho]$. The evolution of the free energy along a solution of (3.1) is given by

$$
\begin{aligned}
\frac{d}{dt}\mathcal{F}[\rho] &= \int_\Omega \frac{\delta\mathcal{F}[\rho]}{\delta\rho}\frac{d\rho}{dt}\,d\boldsymbol{x} \\
&= \int_\Omega \frac{\delta\mathcal{F}[\rho]}{\delta\rho}\nabla\cdot\left(\rho\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho}\right)\,d\boldsymbol{x} \\
&= -\int_\Omega \rho\left|\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho}\right|^2\,d\boldsymbol{x} \\
&\leq 0,
\end{aligned}
\tag{3.2}
$$

where we have applied integration by parts and no-flux boundary conditions. The first equality in (3.2) results from combining the chain rules of functional variation and derivation, and the spatial integral appears due to the fact that $\mathcal{F}[\rho]$ is equal to a spatial integral while $\delta\mathcal{F}[\rho]/\delta\rho$ not. To clarify this further, this is an example of free energy and its corresponding variation with respect to the density:

$$
\mathcal{F}[\rho] = \int_\Omega \rho\left(\log(\rho)-1\right)d\boldsymbol{x}, \qquad \frac{\delta\mathcal{F}[\rho]}{\delta\rho} = \log(\rho).
$$

Interestingly, the dissipation property determines the set of stationary states of (3.1): they are given by nonnegative densities where $\delta\mathcal{F}[\rho]/\delta\rho$ is constant. As a result, the free energy $\mathcal{F}[\rho]$ is a Lyapunov functional for (3.1).

In this thesis we focus on two general structures for the free energy $\mathcal{F}[\rho]$: aggregation-diffusion with external and interaction potentials and Cahn-Hilliard with wall free energy.

3.1.1 Aggregation-diffusion free energy

The free energy for nonlocal and nonlinear aggregation-diffusion equations satisfies

$$
\mathcal{F}[\rho] = \int_\Omega \Pi(\rho)d\boldsymbol{x} + \int_\Omega V(\boldsymbol{x})\rho(\boldsymbol{x})d\boldsymbol{x} + \frac{1}{2}\int_\Omega\int_\Omega W(\boldsymbol{x}-\boldsymbol{y})\rho(\boldsymbol{x})\rho(\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y},
\tag{3.3}
$$

where $\Pi(\rho)$ is the density of the internal energy, $V(\boldsymbol{x})$ is an external or confining potential, and $W(\boldsymbol{x})$ is an interactive potential. The steady states for this choice of free energy are obtained by computing the variation of the free energy with respect to the density, leading to

$$
\frac{\delta\mathcal{F}}{\delta\rho} = \Pi'(\rho) + V(\boldsymbol{x}) + W(\boldsymbol{x})\star\rho = \text{constant on each connected component of supp}(\rho).
\tag{3.4}
$$

Aggregation-diffusion free energies of this type have many applications: granular materials [44], chemotaxis and cell migration [138, 248], opinion formation [3], collective behaviour of animals [184], self-assembly of nanoparticles [168] and mathematical finance [141], to name but a few. If $V(\boldsymbol{x})$ and $W(\boldsymbol{x})$ vanish and the internal energy density satisfies $\Pi(\rho) = \rho(\log(\rho-1))$ or $\Pi(\rho) = \rho^m$, one recovers the classical heat equation or porous medium equation. Confining potentials such as $V(\boldsymbol{x})$ appear in Fokker-Planck-type equations [39]. There is a rich variety of interaction potentials $W(\boldsymbol{x})$: radial fully attractive like in chemotaxis [59] or based on power-laws as in granular materials [68]; repulsive in the short range and attractive in the large range; or compactly-supported potentials in varied biological applications [23].

3.1.2 Cahn-Hilliard free energy

The Cahn-Hilliard (CH) equation is a popular phase-field model initially proposed by Cahn and Hilliard [48] to describe the process of phase separation in binary alloys. Since then it has found applicability in a wide variety of modelling problems, from capillarity and wetting phenomena [13, 258], diblock copolymer molecules [298] and tumour growth [278] to image inpainting [33, 45, 65], topology optimization [300] and many more [180]. Similar to other phase-field models, the CH equation avoids the explicit treatment of sharp interfaces altogether by replacing them with thin transition regions through which pertinent variables and physical properties vary rapidly but continuously.

The free energy for the CH equation satisfies

$$\mathcal{F}[\rho] = \int_\Omega \left(H(\rho) + \frac{\epsilon^2}{2} |\nabla \rho|^2 \right) d\boldsymbol{x} + \int_{\partial\Omega} f_w(\rho, \beta) ds, \tag{3.5}$$

where $H(\rho)$ is a double-well potential with minima corresponding to the stable phases in the system, ϵ is a positive parameter related to the diffuse-interface width (e.g. [79]) and $f_w(\rho, \beta)$ is the wall free energy which depends on the phase field at the wall and the (equilibrium) contact angle β.

The boundary conditions imposed for the CH equation are a combination of the natural boundary condition for the wall free energy and no-flux for the chemical potential [13, 192],

$$\epsilon^2 \nabla \rho \cdot \boldsymbol{n} = -f'_w(\rho, \beta), \quad M(\rho) \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \cdot \boldsymbol{n} = 0, \tag{3.6}$$

where $\boldsymbol{n}$ is an inward-pointing unit vector normal to the wall and $M(\rho)$ is a mobility term.

The variation of the free energy (3.5) with respect to ρ follows from

$$\frac{d}{d\gamma} \mathcal{F}[\rho + \gamma \Psi] \bigg|_{\gamma=0} = \int_\Omega \left(\Psi H'(\rho) + \epsilon^2 \nabla \rho \cdot \nabla \Psi \right) d\boldsymbol{x} + \int_{\partial\Omega} \Psi f'_w(\rho, \beta) ds$$

$$= \int_\Omega \left(H'(\rho) - \epsilon^2 \Delta \rho \right) \Psi d\boldsymbol{x} + \int_{\partial\Omega} \left(\epsilon^2 \nabla \rho \cdot n + f'_w(\rho, \beta) \right) \Psi ds.$$

Now, bearing in mind that from the rules of functional variations it holds that

$$\frac{d}{d\gamma} \mathcal{F}[\rho + \gamma \Psi] \bigg|_{\gamma=0} = \int_\Omega \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \Psi d\boldsymbol{x}$$

and applying the boundary conditions (3.6) to cancel out the surface integral, it results that

$$\frac{\delta \mathcal{F}[\rho]}{\delta \rho} = H'(\rho) - \epsilon^2 \Delta \rho.$$

3.2 Energy-dissipating hyperbolic system

The hyperbolic systems studied in this thesis are formed by two coupled PDEs, modelling the evolution of the density ρ and the momentum $\rho \boldsymbol{u}$:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^d, \quad t > 0, \\ \partial_t (\rho \boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) = -\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} - \gamma \rho \boldsymbol{u}. \end{cases} \tag{3.7}$$

On the one hand, the density equation in (3.7) follows the structure of the overdamped equation (3.1), where the term $\nabla \delta \mathcal{F}[\rho]/\delta \rho$ is substituted by a velocity $\boldsymbol{u}$ modelled by the momentum

equation. On the other hand, the momentum equation contains advection terms $\nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u})$, damping terms $-\gamma \rho \boldsymbol{u}$ with γ being a positive parameter, and the free-energy term $-\rho \nabla \delta \mathcal{F}[\rho]/\delta \rho$ that leads to the pressure and potentials. The form that we assume for $\mathcal{F}[\rho]$ is the same as the aggregation-diffusion free energy in (3.3), and as a result the free-energy term $-\rho \nabla \delta \mathcal{F}[\rho]/\delta \rho$ is usually substituted as

$$-\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} = -\nabla P(\rho) - \rho \nabla \left(V(\boldsymbol{x}) + W(\boldsymbol{x}) \star \rho \right),$$

where

$$\rho \Pi''(\rho) = P'(\rho).$$

In the strong damping limit ($\gamma \to \infty$), the characteristic time scale of the momentum dynamics is much shorter than the density one. Thus, as a first approximation, the contributions of the terms $\nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u})$ and $\partial_t (\rho \boldsymbol{u})$ in (3.7) can be neglected. As a result, it follows that $\boldsymbol{u} = -\gamma^{-1} \nabla \delta \mathcal{F}[\rho]/\delta \rho$, and one recovers the overdamped equation in (3.1) where we have assumed that $\gamma = 1$.

The system (3.7) is said to be hyperbolic if the Jacobian matrix of its vector of fluxes is diagonalizable, which means that it has only real eigenvalues and a full set of eigenvectors. The fluxes in (3.7) are formed by the advection terms plus the pressure part of the free energy. For instance, in 1D they satisfy

$$F(\rho, \rho u) = \begin{pmatrix} \rho u \\ \rho u^2 + P(\rho) \end{pmatrix}.$$

The Jacobian matrix of this 1D flux has two eigenvalues: $u + \sqrt{P'(\rho)}$ and $u - \sqrt{P'(\rho)}$. Consequently, the choice of a suitable pressure is fundamental to ensure that the system (3.7) remains hyperbolic. Here we assume that the pressure is increasing, so that $P'(\rho) > 0$. With this condition of the pressure one can check that the system (3.7) is hyperbolic. This fact is essential when designing numerical schemes, and there are wide differences between schemes to solve elliptic, parabolic or hyperbolic systems. In particular, hyperbolic numerical schemes take advantage of the finite propagation speed of information across the domain and the fact that the eigenvalues are real. Elliptic systems on the contrary, have complex eigenvalues and are typically associated to a special state of a system, in principle corresponding to the minimum of the energy. Finally, in parabolic systems there is only one family of real characteristic curves.

The system (3.7) doesn't necessarily dissipate the free energy $\mathcal{F}[\rho]$, in opposition to what we prove in (3.2) about the overdamped system. Instead, the system (3.7) dissipates the so-called total energy, which is the sum of the kinetic energy and the free energy:

$$E(\rho, \boldsymbol{u}) = \int_\Omega \frac{1}{2} \rho |\boldsymbol{u}|^2 \, d\boldsymbol{x} + \mathcal{F}(\rho). \tag{3.8}$$

The strategy to prove it is based on deriving the total energy over time [60, 69, 144], in a similar

fashion as in (3.2):

$$
\begin{aligned}
\frac{dE(\rho, \boldsymbol{u})}{dt} &= \int_\Omega \frac{1}{2}\boldsymbol{u}\frac{d\left(\rho\boldsymbol{u}\right)}{dt}\, d\boldsymbol{x} + \int_\Omega \frac{1}{2}\rho\boldsymbol{u}\frac{d\boldsymbol{u}}{dt}\, d\boldsymbol{x} + \int_\Omega \frac{\delta\mathcal{F}[\rho]}{\delta\rho}\frac{d\rho}{dt}\, d\boldsymbol{x} \\
&= \int_\Omega \frac{1}{2}\boldsymbol{u}\frac{d\left(\rho\boldsymbol{u}\right)}{dt}\, d\boldsymbol{x} + \int_\Omega \frac{1}{2}\boldsymbol{u}\left(\frac{d\rho\boldsymbol{u}}{dt} - \boldsymbol{u}\frac{d\rho}{dt}\right)\, d\boldsymbol{x} + \int_\Omega \frac{\delta\mathcal{F}[\rho]}{\delta\rho}\frac{d\rho}{dt}\, d\boldsymbol{x} \\
&= \int_\Omega \boldsymbol{u}\frac{d\left(\rho\boldsymbol{u}\right)}{dt}\, d\boldsymbol{x} + \int_\Omega \frac{1}{2}\left|\boldsymbol{u}\right|^2 \nabla\cdot(\rho\boldsymbol{u})\, d\boldsymbol{x} - \int_\Omega \frac{\delta\mathcal{F}[\rho]}{\delta\rho}\nabla\cdot(\rho\boldsymbol{u})\, d\boldsymbol{x} \\
&= \int_\Omega \boldsymbol{u}\frac{d\left(\rho\boldsymbol{u}\right)}{dt}\, d\boldsymbol{x} + \int_\Omega \boldsymbol{u}\nabla\cdot(\rho\boldsymbol{u}\otimes\boldsymbol{u})\, d\boldsymbol{x} + \int_\Omega \rho\boldsymbol{u}\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho}\, d\boldsymbol{x} \\
&= \int_\Omega \boldsymbol{u}\left(\frac{d\left(\rho\boldsymbol{u}\right)}{dt} + \nabla\cdot(\rho\boldsymbol{u}\otimes\boldsymbol{u}) + \nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho}\right)\, d\boldsymbol{x} \\
&= -\gamma\int_\Omega \rho\left|\boldsymbol{u}\right|^2 d\boldsymbol{x} \\
&\le 0.
\end{aligned}
$$

Consequently, the total energy is the system is dissipated because of the inclusion of the linear damping term $\gamma\rho\boldsymbol{u}$. Without damping the total energy remains constant, even if there is dynamic evolution in the system. The term $\gamma\rho\boldsymbol{u}$ ensures that the total energy $E(\rho, \boldsymbol{u})$ keeps decreasing in time while there is kinetic energy in the system. At the same time, since the definition of the total energy in (3.8) also depends on the velocity $\boldsymbol{u}$, it follows that the velocity throughout the domain eventually vanishes. When $\boldsymbol{u} = \boldsymbol{0}$ throughout the domain, the momentum equation in (3.7) reduces to

$$
\rho\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho} = \boldsymbol{0},
$$

meaning that the same steady state relation as in the overdamped case (3.4) holds, but now including the momentum:

$$
\frac{\delta\mathcal{F}[\rho]}{\delta\rho} = \Pi'(\rho) + H(\boldsymbol{x}, \rho) = \text{constant on each connected component of supp}(\rho) \text{ and } \boldsymbol{u} = 0,
$$

where the constant can vary on different connected components of supp(ρ). This applies to the points inside the support of the density. For those points outside the support of the density and satisfying $\boldsymbol{u} = 0$, the variation of the free energy with respect to the density does not need to keep the constant value when the steady state is reached.

The system in (3.7) also satisfies an entropy identity

$$
\partial_t\eta(\rho, \rho\mathbf{u}) + \nabla\cdot\boldsymbol{G}(\rho, \rho\mathbf{u}) = -\rho\boldsymbol{u}\cdot\nabla\left(V(\boldsymbol{x}) + W(\boldsymbol{x})\star\rho\right) - \gamma\rho\left|\boldsymbol{u}\right|^2,
$$

where $\eta(\rho, \rho\boldsymbol{u})$ and $\boldsymbol{G}(\rho, \rho\boldsymbol{u})$ are the entropy and the entropy flux defined as

$$
\eta(\rho, \rho\boldsymbol{u}) = \rho\frac{\left|\boldsymbol{u}\right|^2}{2} + \Pi(\rho), \quad \boldsymbol{G}(\rho, \rho\boldsymbol{u}) = \rho\boldsymbol{u}\left(\frac{\left|\boldsymbol{u}\right|^2}{2} + \Pi'(\rho)\right).
$$

Part II

Well-balanced schemes for hydrodynamic systems

Introduction to well-balanced schemes

Well-balanced schemes have emerged as a paramount tool to simulate systems governed by balance/conservation laws. This is due to their ability to numerically preserve steady states and resolve small perturbations of those states even with coarse meshes. Well-balanced schemes were introduced nearly three decades ago, with the initial works by Bermúdez and Vázquez [29], Greenberg and Leroux [157] and Gosse [152]. One of the most popular applications for well-balanced schemes from the beginning has been the shallow-water equations. Further contributions of note are the hydrostatic reconstruction in [11, 76] together with application scenarios where well-balanced schemes have proven quite successful: tsunami propagation [78], coastal hydrodynamics [206] and irregular topographies [137], to name but a few. Inspired by the strong results for the shallow-water equations, plenty of authors have successfully employed well-balanced schemes in a plethora of balance-law problems from wave propagation in elastic media [285] and chemosensitive movement of cells [131] to flow through a nozzle [142] and the Euler equations with gravity [181, 267]. The historical evolution of well-balanced schemes is reviewed in [154]. Furthermore, the derivation of numerical schemes preserving structural properties of the evolutions under study such as dissipations or conservations of relevant physical quantities is an important line of research in hydrodynamic systems and their overdamped limits, see for instance [55, 106, 226, 264].

In the present work, we propose numerical schemes with well-balanced and free energy dissipation properties for a general class of balance laws or hydrodynamic models with attractive-repulsive interaction forces. In particular we consider interactions associated with nonlocal convolutions or functions of convolutions, which is commonplace in applications such as the Keller-Segel model [49], more general Euler-Poisson systems [162] or in dynamic-density functional theory (DDFT) [147, 148]. This class of balance laws may contain linear or nonlinear damping effects, such as the Cucker-Smale alignment term in collective behaviour [102]. The corresponding hydrodynamic systems have the general form

$$
\begin{cases}
\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^d, \quad t > 0, \\
\partial_t(\rho \boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) = -\nabla P(\rho) - \rho \nabla H(\boldsymbol{x}, \rho) - \gamma \rho \boldsymbol{u} - \rho \int_{\mathbb{R}^d} \psi(\boldsymbol{x} - \boldsymbol{y})(\boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{u}(\boldsymbol{y}))\rho(\boldsymbol{y}) \, d\boldsymbol{y},
\end{cases}
\tag{4.1}
$$

where $\rho = \rho(\boldsymbol{x}, t)$ and $\boldsymbol{u} = \boldsymbol{u}(\boldsymbol{x}, t)$ are the density and the velocity, $P(\rho)$ is the pressure, $H(\boldsymbol{x}, \rho)$ contains the attractive-repulsive effects from external V or interaction potentials W, assumed to be locally integrable, given by

$$
H(\boldsymbol{x}, \rho) = V(\boldsymbol{x}) + W(\boldsymbol{x}) \star \rho,
$$

and $\psi(\boldsymbol{x})$ is a nonnegative symmetric smooth function called the communication function in the Cucker-Smale model [102, 103] describing collective behavior of systems due to alignment [56].

The fractional-step methods [196] have been the widely-employed tool to simulate the temporal evolution of balance laws such as (4.1). They are based on a division of the problem in (4.1) into two simpler subproblems: the homogeneous hyperbolic system without source terms and the temporal evolution of density and momentum without the flux terms but including the sources. These subproblems are then resolved alternatively employing suitable numerical methods for each. This procedure introduces a splitting error which is acceptable for the temporal evolution, but becomes critical when the objective is to preserve the steady states. This is due to the fact that the steady state is reached when the fluxes are exactly balanced with the source terms in each discrete node of the domain. However, when solving alternatively the two subproblems, this discrete balance can never be achieved, since the fluxes and source terms are not resolved simultaneously.

To correct this deficiency, well-balanced schemes are designed to discretely satisfy the balance between fluxes and sources when the steady state is reached [42]. The strategy to construct well-balanced schemes relies on the fact that, when the steady state is reached, there are some constant relations of the variables that hold in the domain. These relations allow the resolution of the fluxes and sources in the same level, thus avoiding the division that the fractional-step methods introduce. Moreover, if the system enjoys a dissipative property and it has a Liapunov functional, obtaining analogous tools at the discrete level is key for the derivation of well-balanced schemes. In this work the steady-state relations and the dissipative property are obtained by means of the associated free energy, which in the case of the system in (4.1) is formulated as

$$\mathcal{F}[\rho] = \int_{\mathbb{R}^d} \Pi(\rho)d\boldsymbol{x} + \int_{\mathbb{R}^d} V(\boldsymbol{x})\rho(\boldsymbol{x})d\boldsymbol{x} + \frac{1}{2}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} W(\boldsymbol{x}-\boldsymbol{y})\rho(\boldsymbol{x})\rho(\boldsymbol{y})d\boldsymbol{x}d\boldsymbol{y}, \qquad (4.2)$$

where

$$\rho\Pi''(\rho) = P'(\rho). \qquad (4.3)$$

The pressure $P(\rho)$ and the potential term $H(\boldsymbol{x}, \rho)$ appearing in the general system in (4.1) can be gathered by considering the associated free energy. Taking into account that the variation of the free energy in (4.2) with respect to the density ρ is equal to

$$\frac{\delta\mathcal{F}}{\delta\rho} = \Pi'(\rho) + H(\boldsymbol{x}, \rho), \qquad (4.4)$$

it follows that the general system in (4.1) can be written in a compact form as

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho\boldsymbol{u}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^d, \quad t > 0, \\ \partial_t(\rho\boldsymbol{u}) + \nabla \cdot (\rho\boldsymbol{u} \otimes \boldsymbol{u}) = -\rho\nabla\frac{\delta\mathcal{F}}{\delta\rho} - \gamma\rho\boldsymbol{u} - \rho\int_{\mathbb{R}^d}\psi(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{u}(\boldsymbol{y}))\rho(\boldsymbol{y})\,d\boldsymbol{y}. \end{cases} \qquad (4.5)$$

The system in (4.5) is rather general containing a wide variety of physical problems all under the so-called density functional theory (DFT) and its dynamic extension (DDFT) see e.g. [121, 147, 148, 149, 293, 294] and the references therein. A variety of well-balanced schemes have already been constructed for specific choices of the terms $\Pi(\rho)$, $V(\boldsymbol{x})$ and $W(\boldsymbol{x})$ in the free energy in (4.2), see [11, 42, 131] for instance. Here the focus is set on the free energy and the natural structure of the system in (4.5). It is naturally advantageous to consider the concept of free energy in the construction procedure of well-balanced schemes, since they rely on relations that hold in the steady states, and moreover, the variation of the free energy with respect to the density is constant when reaching these steady states, more precisely

$$\frac{\delta\mathcal{F}}{\delta\rho} = \Pi'(\rho) + H(\boldsymbol{x}, \rho) = \text{constant on each connected component of supp}(\rho) \text{ and } u = 0, \qquad (4.6)$$

where the constant can vary on different connected components of supp(ρ). As a result, the constant relations in the steady states, which are needed for well-balanced schemes, are directly provided by the variation of the free energy with respect to the density.

The steady state relations in (4.6) hold due to the dissipation of the linear damping $-\rho\boldsymbol{u}$ or nonlinear damping in the system in (4.1), which eventually eliminates the momentum of the system. This can be justified by means of the total energy of the system, defined as the sum of kinetic and free energy,

$$E(\rho, \boldsymbol{u}) = \int_{\mathbb{R}^d} \frac{1}{2}\rho \, |\boldsymbol{u}|^2 \, d\boldsymbol{x} + \mathcal{F}(\rho), \tag{4.7}$$

since it is formally dissipated, see [60, 69, 144], as

$$\frac{dE(\rho, \boldsymbol{u})}{dt} = -\gamma \int_{\mathbb{R}^d} \rho \, |\boldsymbol{u}|^2 \, d\boldsymbol{x} - \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \psi(\boldsymbol{x} - \boldsymbol{y}) \, |\boldsymbol{u}(\boldsymbol{y}) - \boldsymbol{u}(\boldsymbol{x})|^2 \, \rho(\boldsymbol{x}) \, \rho(\boldsymbol{y}) \, d\boldsymbol{x} \, d\boldsymbol{y}. \tag{4.8}$$

The system in (4.1) also satisfies an entropy identity

$$\partial_t \eta(\rho, \rho\mathbf{u}) + \nabla \cdot \boldsymbol{G}(\rho, \rho\mathbf{u}) = -\rho\boldsymbol{u} \cdot \nabla H(\boldsymbol{x}, \rho) - \gamma\rho \, |\boldsymbol{u}|^2 - \rho \int_{\mathbb{R}^d} \psi(\boldsymbol{x} - \boldsymbol{y}) \, \boldsymbol{u}(\boldsymbol{x}) \cdot (\boldsymbol{u}(\boldsymbol{x}) - \boldsymbol{u}(\boldsymbol{y}))\rho(\boldsymbol{y}) \, d\boldsymbol{y}, \tag{4.9}$$

where $\eta(\rho, \rho\boldsymbol{u})$ and $\boldsymbol{G}(\rho, \rho\boldsymbol{u})$ are the entropy and the entropy flux defined as

$$\eta(\rho, \rho\boldsymbol{u}) = \rho\frac{|\boldsymbol{u}|^2}{2} + \Pi(\rho), \quad \boldsymbol{G}(\rho, \rho\boldsymbol{u}) = \rho\boldsymbol{u}\left(\frac{|\boldsymbol{u}|^2}{2} + \Pi'(\rho)\right). \tag{4.10}$$

From a physical point of view the entropy is always a convex function of the density [189]. As a result, from (4.10) it is justified to assume that $\Pi(\rho)$ is convex, meaning that $\Pi'(\rho)$ has an inverse function for positive densities ρ. This last fact is a necessary requirement for the construction of the well-balanced schemes of this work. Finally, notice that from the entropy identity in (4.9), one recovers the free energy dissipation in (4.8) by integration using the continuity equation to deal with the forces term $H(\boldsymbol{x}, \rho)$ and using symmetrization of the nonlinear damping term due to ψ being symmetric.

Let us also point out that the evolution of the centre of mass of the density can be computed in some particular cases. In fact, it is not difficult to deduce from (4.5) that

$$\frac{d}{dt} \int_{\mathbb{R}^d} \boldsymbol{x}\rho d\boldsymbol{x} = \int_{\mathbb{R}^d} \rho\boldsymbol{u}d\boldsymbol{x} \quad \text{and} \quad \frac{d}{dt} \int_{\mathbb{R}^d} \rho\boldsymbol{u}d\boldsymbol{x} = -\int_{\mathbb{R}^d} \nabla V(\boldsymbol{x})\rho d\boldsymbol{x} - \gamma \int_{\mathbb{R}^d} \rho\boldsymbol{u}d\boldsymbol{x} \,, \tag{4.11}$$

due to the antisymmetry of $\nabla W(\boldsymbol{x})$ and the symmetry of $\psi(\boldsymbol{x})$. Therefore, in case $V(\boldsymbol{x})$ is not present or quadratic, (4.11) is explicitly solvable. Moreover, if the potential $V(\boldsymbol{x})$ is symmetric, the initial data for the density is symmetric, and the initial data for the velocity is antisymmetric, then the solution to (4.5) keeps these symmetries in time, i.e., the density is symmetric and the velocity is antisymmetric for all times, and the centre of mass is conserved

$$\frac{d}{dt} \int_{\mathbb{R}^d} \boldsymbol{x}\rho d\boldsymbol{x} = 0 \,.$$

The steady state relations in (4.6) only hold when the linear damping term is included in the system in (4.1). When only the nonlinear damping of Cucker-Smale type is present, the system has the so-called moving steady states, see [56, 58, 69], which satisfy the more general relations

$$\frac{\delta\mathcal{F}}{\delta\rho} = \text{constant on each connected component of supp}(\rho) \text{ and } \boldsymbol{u} = \text{constant.} \tag{4.12}$$

28

However, the construction of well-balanced schemes satisfying the moving steady state relations has proven to be more difficult than for the still steady states in (4.6) without dissipation. For literature about well-balanced schemes for moving steady states without dissipation, we refer to [218, 287].

The most popular application in the literature for well-balanced schemes deals with the Saint-Venant system for shallow water flows with nonflat bottom [11, 42, 51, 199, 282, 283], for which $\Pi(\rho) = \frac{g}{2}\rho^2$, with g being the gravity constant, and $H(\boldsymbol{x}, \rho)$ depends on the bottom. Here it is important to remark the work of Audusse *et al.* in [11], where they propose a hydrostatic reconstruction that has successfully inspired more sophisticated well-balanced schemes in the area of shallow water equations [206, 217]. Another area where well-balanced schemes have been fruitful is chemosensitive movement, with the works of Filbet, Shu and their collaborators [130, 131, 153, 285]. In this case the pressure satisfies $\Pi(\rho) = \rho(\ln(\rho) - 1)$ and H depends on the chemotactic sensitivity and the chemical concentration. The list of applications of the system in (4.1) continues growing with more choices of $\Pi(\rho)$ and $H(\boldsymbol{x}, \rho)$ [285]: the elastic wave equation, nozzle flow problem, two phase flow model, etc.

The orders of accuracy from the finite volume well-balanced schemes presented before range from first- and second-order [11, 186, 195, 199, 289] to higher-order versions [137, 217, 276, 285]. Again, the most popular application has been shallow water equations, and the survey from Xing and Shu [282] provides a summary of all the shallow water methods with different accuracies. Some of the previous schemes proposed were equipped to satisfy natural properties of the systems under consideration, such as nonnegativity of the density [12, 186] or the satisfaction of a discrete entropy inequality [11, 131], enabling also the computation of dry states [137] . Theoretically the Godunov scheme satisfies all these properties [194], but its main drawback is its computationally expensive implementation. Several authors have already proposed high-order well-balanced schemes for systems where the potential terms in the free energy in (4.2) are local, such as the shallow-water equations [73, 75, 80, 90, 217], chemotaxis [131] and other applications [285].

Other well-balanced numerical approaches employed to simulate the system in (4.5) are finite differences [283, 284], which are equivalent to the finite volume methods for first-and second-order, and the discontinuous Galerkin methods [285]. There have been plenty of contributions devoted to particular configurations and scenarios, with a especial emphasis on shallow-water equations: presence of dry areas and bottom topography [137], tsunami propagation in 2D meshes [78], traffic flow model [80], moving steady states [73, 87, 218], etc. The overdamped system of (4.5) with $\psi \equiv 0$, obtained in the free inertia limit where the momentum reaches equilibrium on a much faster timescale than the density, has also been numerically resolved for general free energies of the form in (4.2), via finite volume schemes [55] or discontinuous Galerkin approaches [264]. This scheme for the overdamped system also conserves the dissipation of the free energy at the discrete approximation.

In addition to the well-balanced property, many authors have sought to construct numerical schemes that preserve the structural properties of the system in (4.1) during its temporal evolution. These endeavours have aimed to first satisfy discretely the entropy identity and second the dissipation relations for the total energy, both for the original system in (4.1) and its overdamped versions. We refer the reader to Refs. [10, 132, 266] for more information about entropy stable schemes and to Refs. [15, 55] for insights details and useful insights energy dissipating schemes. The well-balanced finite-volume scheme of our previous work [71] was designed to be at the same time well-balanced, entropy stable and energy dissipating, though only for first- and second-order accuracy.

4.1   Contribution to first- and second-order well-balanced schemes

In this work we present a finite volume scheme for a general choice of $\Pi(\rho)$ and $H(\boldsymbol{x}, \rho)$ which is first- and second-order accurate and satisfies the nonnegativity of the density, the well-balanced property, the semidiscrete entropy inequality and the semidiscrete free energy dissipation. Furthermore, as it is shown in Example 5.3.9 of Section 5.3, it can also be applied to more general free energies than the one in (4.2) and with the form

$$\mathcal{F}[\rho] = \int_{\mathbb{R}^d} \Pi(\rho) d\boldsymbol{x} + \int_{\mathbb{R}^d} V(\boldsymbol{x}) \rho(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{2} \int_{\mathbb{R}^d} K \left(W(\boldsymbol{x}) \star \rho(\boldsymbol{x}) \right) \rho(\boldsymbol{x}) d\boldsymbol{x}, \qquad (4.13)$$

where K is a function depending on the convolution of $\rho(\boldsymbol{x})$ with the kernel $W(\boldsymbol{x})$. Its variation with respect to the density satisfies

$$\frac{\delta \mathcal{F}}{\delta \rho} = \Pi'(\rho) + V(\boldsymbol{x}) + \frac{1}{2} K \left(W(\boldsymbol{x}) \star \rho \right) + \frac{1}{2} K' \left(W(\boldsymbol{x}) \star \rho \right) \left(W(\boldsymbol{x}) \star \rho \right). \qquad (4.14)$$

These free energies arise in applications related to (D)DFT [121, 147], see [70] for other related free energies and properties.

The novelty of this work is twofold. Foremost, all these previous schemes were only applicable for specific choices of $\Pi(\rho)$ and $H(\boldsymbol{x}, \rho)$, meaning that a general scheme valid for a wide range of applications is lacking. And while some previous schemes [285] could be employed in more general cases, the focus in the literature has been on the shallow water and chemotaxis equations. In addition, the function $H(\boldsymbol{x}, \rho)$, which results from summing $V(\boldsymbol{x})$ and $W(\boldsymbol{x}) \star \rho$ as in (4), has so far been taken as dependent on $\boldsymbol{x}$ only, unlike the present work where it depends on ρ by means of the convolution with an interaction potential $W(\boldsymbol{x})$.

The other novel technical aspect of this work concerns the numerical treatment of the different source terms in (4.1). In fact, in order to keep the well-balanced property and the decay of the free energy we treat source terms differently. While the dissipative terms are harmless and treated by direct approximations, the fundamental question is how to choose the discretization of the potential term given by $H(\boldsymbol{x}, \rho) = V(\boldsymbol{x}) + W(\boldsymbol{x}) \star \rho$. For this purpose we appropriately extend the ideas in [42, 131] to our case to keep the well-balanced property and the energy decay. The condition for stationary states in (4.6) is crucial in defining an approximation of the term $-\rho \nabla H(\boldsymbol{x}, \rho)$ by a discretization of $\nabla P(\rho)$ which is consistent when the new reconstructed values of the density at the interfaces taking into account the potential $H(\boldsymbol{x}, \rho)$. This general treatment includes as specific cases both the shallow-water equations [11, 42] and the hyperbolic chemotaxis problem [131].

Chapter 5 is devoted to this contribution. Sections 5.1 and 5.2 describe the first- and second-order well-balanced scheme reconstructions, and provides the proofs of their main properties. Section 5.3 contains the numerical simulations: firstly, Subsection 5.3.1 details the validation of the well-balanced property and the orders of accuracy is conducted; secondly, Subsection 5.3.2 with numerical experiments from different applications. Finally, a wide range of free energies is employed to remark the extensive nature of our well-balanced scheme.

4.2   Contribution to high-order well-balanced schemes

The main contribution of the present work is to extend our previous scheme for the system in (4.1) from first and second order to high order. Here we consider the much broader class of free energies in (4.2), which include interaction potentials leading to forces given by convolution with the density ρ and possible linear or nonlinear damping effects from the field of collective behaviour. Applications of this type include the Keller-Segel model, generalized Euler-Poisson systems [162] and DDFT [147, 148]. We complement our high-order finite volume schemes with the

desired properties of well-balancing and the nonnegativity of the density, which allows for vacuum regions in the simulations. It is worth mentioning that previous works have already addressed the construction of high-order well-balanced schemes for diverse types of steady-state relations. Nevertheless the technical challenge here is to consider non-linear and non-local steady states that involve convolutions. The well-balanced methodology for shallow-water equations, for instance, is not directly applicable to our steady states.

Our work lays the foundations for the construction of well-balanced high-order schemes that may satisfy further fundamental properties of the system in (4.1), such as the discrete versions of the energy dissipation in (4.8) and entropy identity in (4.9), or even the well-balanced property for the challenging moving steady states. Developing schemes, amongst the class of positivity-preserving high-order schemes introduced in the present work, satisfying also the entropy stability and energy dissipating properties, is a challenging open question. In Section 6.5 we however show that the numerical simulations satisfy the discrete energy dissipation in practice, even if we are not able to prove it for our high-order finite-volume scheme.

Chapter 6 is devoted to this contribution. In Section 6.1 we begin by recalling our first-order numerical scheme from Chapter 5, and then in Section 6.2 we provide an first-attempt extension of such scheme to high order. The correct well-balanced formulation for that high-order scheme is provided in Section 6.3. Section 6.4 contains the summarized algorithmic implementation of the scheme. Then, in Section 6.5 we depict a battery of simulations for relevant applications of system in (4.1). In Subsection 6.5.1 we numerically check the well-balanced property and high-order accuracy of our scheme, and subsequently in Subsection 6.5.2 we tackle applications for varied choices of the free energy, leading to interesting numerical experiments for which analytical results are limited in the literature.

First- and second-order well-balanced schemes

The contents of this chapter are based on the publication [66], done in collaboration with José A. Carrillo, Serafim Kalliadasis and Chi-Wang Shu:

[66]   J. A. Carrillo, S. Kalliadasis, S. P. Perez, and C. W. Shu. "Well-balanced finite-volume schemes for hydrodynamic equations with general free energy". In: *Multiscale Modeling & Simulation* 18.1 (2020), pp. 502-541.

The terms appearing in the one-dimensional system in (4.5) are usually gathered in the form of

$$\partial_t U + \partial_x F(U) = S_H(U, H) + S_D(x, U), \tag{5.1}$$

with

$$U = \begin{pmatrix} \rho \\ \rho u \end{pmatrix}, \quad F(U) = \begin{pmatrix} \rho u \\ \rho u^2 + P(\rho) \end{pmatrix}$$

and and

$$S_H(U, H) = \begin{pmatrix} 0 \\ -\rho \partial_x H \end{pmatrix}, \quad S_D(x, U) = \begin{pmatrix} 0 \\ -\gamma \rho u - \rho \int_{\mathbb{R}} \psi(x - y)(u(x) - u(y))\rho(y)\, dy \end{pmatrix},$$

where U are the unknown variables, $F(U)$ the fluxes, and $S_H(U, H)$ and $S_D(x, U)$ the sources related to forces with potential H and damping terms respectively. The one-dimensional finite volume approximation of (5.1) is obtained by breaking the domain into grid cells $\left(x_{i-1/2}\right)_{i \in \mathbb{Z}}$ and approximating in each of them the cell average of U. Then these cell averages are modified after each time step, depending on the flux through the edges of the grid cells and the cell average of the source term [196]. Finite volume schemes for hyperbolic systems employ an upwinding of the fluxes and in the semidiscrete case they provide a discrete version of (5.1) under the form

$$\frac{dU_i}{dt} = -\frac{F_{i+1/2} - F_{i-1/2}}{\Delta x_i} + S_{H,i} + S_{D,i}, \tag{5.2}$$

where the cell average of U in the cell $\left(x_{i-1/2}, x_{i+1/2}\right)$ is denoted as

$$U_i = \begin{pmatrix} \rho_i \\ \rho_i u_i \end{pmatrix},$$

$F_{i+1/2}$ is an approximation of the flux $F(U)$ at the point $x_{i+1/2}$, $S_{H,i}$ and $S_{D,i}$ approximations of the source terms $S_H(U, H)$ and $S_D(x, U)$ in the cell $\left(x_{i-1/2}, x_{i+1/2}\right)$, and Δx_i is the possibly variable mesh size $\Delta x_i = x_{i+1/2} - x_{i-1/2}$.

The approximation of the flux $F(U)$ at the point $x_{i+1/2}$, denoted as $F_{i+1/2}$, is achieved by means of a numerical flux $\mathcal{F}$ which depends on two reconstructed values of U at the left and right of the boundary between the cells i and $i+1$. These two values, $U_{i+1/2}^-$ and $U_{i+1/2}^+$, are computed from the cell averages following different construction procedures that seek to satisfy certain properties, such as order of accuracy or nonnegativity. Two widely-employed reconstruction procedures are the second-order finite volume monotone upstream-centred scheme for conservation laws, referred to as MUSCL [223], or the weighted-essentially non-oscillatory schemes, widely known as WENO [254].

Once these two reconstructed values are computed, $F_{i+1/2}$ is obtained from

$$F_{i+1/2} = \mathcal{F}\left(U_{i+1/2}^-, U_{i+1/2}^+\right). \tag{5.3}$$

The numerical flux $\mathcal{F}$ is usually denoted as Riemann solver, since it provides a stable resolution of the Riemann problem located at the cell interfaces, with $U_{i+1/2}^-$ denoting the reconstruction at the left of the interface and $U_{i+1/2}^+$ at the right. The literature concerning Riemann solvers is vast and there are different choices for it [269]: Godunov, Lax-Friedrich, kinetic, Roe, etc. Some usual properties of the numerical flux that are assumed [11, 42, 131] are:

1. It is consistent with the physical flux, so that $\mathcal{F}(U, U) = F(U)$.

2. It preserves the nonnegativity of the density $\rho_i(t)$ for the homogeneous problem, where the numerical flux is computed as in (5.3).

3. It satisfies a cell entropy inequality for thd entropy pair (4.10) for the homogeneous problem. Then, according to [42], it is possible to find a numerical entropy flux $\mathcal{G}$ such that

$$\begin{aligned} G(U_{i+1}) + \nabla_U\, \eta(U_{i+1}) \left(\mathcal{F}(U_i, U_{i+1}) - F(U_{i+1})\right) \\ \leq \mathcal{G}(U_i, U_{i+1}) \leq G(U_i) + \nabla_U\, \eta(U_i) \left(\mathcal{F}(U_i, U_{i+1}) - F(U_i)\right), \end{aligned} \tag{5.4}$$

where $\nabla_U\, \eta$ is the derivative of η with respect to $U = \begin{pmatrix} \rho \\ \rho u \end{pmatrix}$.

The first- and second-order well-balanced schemes described in this section propose an alternative reconstruction procedure for $U_{i+1/2}^-$ and $U_{i+1/2}^+$ which ensures that the steady state in (4.6) is discretely preserved when starting from that steady state. Sections 5.1 and 5.2 contain the first- and second-order schemes, respectively, together with their proved properties.

5.1 First-order scheme

The basic first-order schemes approximate the flux $F_{i+1/2}$ by a numerical flux $\mathcal{F}$ which depends on the cell averaged values of U at the two adjacent cells, so that the inputs for the numerical flux in (5.3) are

$$F_{i+1/2} = \mathcal{F}\left(U_i, U_{i+1}\right). \tag{5.5}$$

The resolution of the finite volume scheme in (5.2) with a numerical flux of the form in (5.5) and a cell-centred evaluation of $-\rho\partial_x H$ for the source term S_i is not generally able to preserve the steady states, as it was shown in the initial works of well-balanced schemes [155, 157]. These steady states are provided in (4.6), and satisfy that the variation of the free energy with respect

to the density has to be constant in each connected component of the support of the density. The discrete steady state is defined in a similar way,

$$\left(\frac{\delta\mathcal{F}}{\delta\rho}\right)_i = \Pi'(\rho_i) + H_i = C_\Gamma \text{ in each } \Lambda_\Gamma, \Gamma \in \mathbb{N}, \tag{5.6}$$

where Λ_Γ, $\Gamma \in \mathbb{N}$, denotes the possible infinite sequence indexed by Γ of subsets Λ_Γ of subsequent indices $i \in \mathbb{Z}$ where $\rho_i > 0$ and $u_i = 0$, and C_Γ the corresponding constant in that connected component of the discrete support.

As it was emphasized above, the preservation of these steady states for particular choices of $\Pi'(\rho)$ and $H(x,\rho)$, such as shallow water [11] or chemotaxis [131], is paramount. A solution to allow for this preservation is the so-called hydrostatic reconstruction proposed by Audusse et al. [11], where instead of evaluating the numerical flux as in (5.3), they chose

$$F_{i+1/2} = \mathcal{F}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right), \text{ where } U_{i+1/2}^{HR,\pm} = \begin{pmatrix} \rho_{i+1/2}^{HR,\pm} \\ \rho_{i+1/2}^{HR,\pm} u_{i+1/2}^{HR,\pm} \end{pmatrix}. \tag{5.7}$$

The interface values $U_{i+1/2}^{HR,\pm}$ are reconstructed from U_i and U_{i+1} by taking into account the steady state relation in (5.6). Contrary to other works in which the interface values are reconstructed to increase the order of accuracy, now the objective is to satisfy the well-balanced property. Bearing this in mind, we make use of (5.6) to the cells with centred nodes at x_i and x_{i+1} to define the interface values such that

$$\Pi'\left(\rho_{i+1/2}^{HR,-}\right) + H_{i+1/2} = \Pi'(\rho_i) + H_i,$$
$$\Pi'\left(\rho_{i+1/2}^{HR,+}\right) + H_{i+1/2} = \Pi'(\rho_{i+1}) + H_{i+1},$$

where the term $H_{i+1/2}$ is evaluated to preserve consistency and stability, with an upwind or average value obtained as

$$H_{i+1/2} = \max(H_i, H_{i+1}) \quad \text{or} \quad H_{i+1/2} = \frac{1}{2}(H_i + H_{i+1}). \tag{5.8}$$

Then, by denoting as $\xi(s)$ the inverse function of $\Pi'(s)$ for $s > 0$, we conclude that the interface values $U_{i-1/2}^{HR,\pm}$ are computed as

$$\begin{aligned} \rho_{i+1/2}^{HR,-} &= \xi\left(\Pi'(\rho_i) + H_i - H_{i+1/2}\right)_+, & u_{i+1/2}^{HR,-} &= u_i, \\ \rho_{i+1/2}^{HR,+} &= \xi\left(\Pi'(\rho_{i+1}) + H_{i+1} - H_{i+1/2}\right)_+, & u_{i+1/2}^{HR,+} &= u_{i+1}. \end{aligned} \tag{5.9}$$

The function $\xi(s)$ is well-defined for $s > 0$ since $\Pi(s)$ is strictly convex, $\Pi''(s) > 0$. This is always the case since, as mentioned in the introduction, the physical entropies are always strictly convex from (4.10). However, some physical entropies and applications allow for vacuum of the steady states, therefore we need to impose the value of $\rho_{i+1/2}^{HR,\pm}$, given that they should be nonnegative. Henceforth, $\xi(s)$ denotes the extension by zero of the inverse of $\Pi'(s)$ whenever $s > 0$.

Furthermore, the discretization of the source term with potentials $S_H(U, H)$ is taken as

$$S_{H,i} = S_{i+1/2}^{HR,-} + S_{i-1/2}^{HR,+}, \tag{5.10}$$

where

$$S_{i+1/2}^{HR,-} = \frac{1}{\Delta x_i}\begin{pmatrix} 0 \\ P\left(\rho_{i+1/2}^{HR,-}\right) - P(\rho_i) \end{pmatrix}, \quad S_{i-1/2}^{HR,+} = \frac{1}{\Delta x_i}\begin{pmatrix} 0 \\ P(\rho_i) - P\left(\rho_{i-1/2}^{HR,+}\right) \end{pmatrix}.$$

This choice is motivated by the fact that in the steady state, with $u = 0$ in (5.1), the fluxes are balanced with the sources,

$$\rho \partial_x \Pi'(\rho) = -\rho \partial_x H.$$

The source term with damping $S_D(x, U)$ is simply discretized as

$$S_{D,i} = - \begin{pmatrix} 0 \\ \gamma \rho_i u_i + \rho_i \sum_j \Delta x_j (u_i - u_j) \rho_j \psi_{ij} \end{pmatrix}.$$

Here ψ_{ij} is an approximation of the average value of ψ on the interval centred at $x_i - x_j$ of length Δx_j. From here, and integrating over the cell volume, it results that

$$\int_{x_{i-1/2}}^{x_{i+1/2}} -\rho \partial_x H \, dx = \int_{x_{i-1/2}}^{x_{i+1/2}} \rho \partial_x \Pi'(\rho) \, dx = \int_{x_{i-1/2}}^{x_{i+1/2}} \partial_x P(\rho) \, dx = P(\rho_{i+1/2}^{HR,-}) - P(\rho_{i-1/2}^{HR,+}), \quad (5.11)$$

with the relation between $\Pi'(\rho)$ and $P(\rho)$ was given in (13.6). This idea of distributing the source terms along the interfaces has already been explored in previous works [172].

The discretization of the source term in (5.10) entails that the discrete balance between fluxes and sources is accomplished when $F_{i+1/2} = P(\rho_{i+1/2}^{HR,-}) = P(\rho_{i+1/2}^{HR,+})$. The computation of the numerical fluxes expressed in (5.7), in which the interface values $U_{i-1/2}^{HR,\pm}$ are considered, enables this balance if in the steady states $U_{i+1/2}^{HR,-} = U_{i+1/2}^{HR,+} = (\rho_{i+1/2}^{HR,-}, 0) = (\rho_{i+1/2}^{HR,+}, 0)$. Moreover, the discretization of the source term as in (5.10) may seem counter-intuitive when the system is far away from the steady state, given that the balanced expressed in (5.11) only holds in those states. In spite of this, the consistency with the original system in (5.1) is not lost, as it will be proved in Subsection 5.1.1.

Let us finally discuss the discretization of the potential $H(x, \rho) = V(x) + W * \rho(x)$. We will always approximate it as

$$H_i = V_i + \sum_j \Delta x_j W_{ij} \rho_j, \text{ for all } i \in \mathbb{Z},$$

where $V_i = V(x_i)$ and $W_{ij} = W(x_i - x_j)$ in case the potential is smooth or choosing W_{ij} as an average value of W on the interval centred at $x_i - x_j$ of length Δx_j in case of general locally integrable potentials W. Let us also point out that this discretization keeps the symmetry of the discretized interaction potential $W_{ij} = W_{ji}$ for all $i, j \in \mathbb{Z}$ whenever W is smooth or solved with equal size meshes $\Delta x_i = \Delta x_j$ for all $i, j \in \mathbb{Z}$.

5.1.1 Properties of the first-order scheme

The first-order semidiscrete scheme defined in (5.2), constructed with (5.7)-(5.10), and for a numerical flux $\mathcal{F}(U_i, U_{i+1}) = (\mathcal{F}^\rho, \mathcal{F}^{\rho u})(U_i, U_{i+1})$ satisfying the properties stated in the introduction of Chapter 5, satisfies:

  (i) preservation of the nonnegativity of $\rho_i(t)$;

 (ii) well-balanced property, thus preserving the steady states given by (5.6);

(iii) consistency with the system in (4.5);

(iv) cell entropy inequality associated to the entropy pair in (4.10),

$$\Delta x_i \frac{d\eta_i}{dt} + \Delta x_i H_i \frac{d\rho_i}{dt} + G_{i+1/2} - G_{i-1/2} = -u_i \left(\gamma \Delta x_i \rho_i u_i + \Delta x_i \rho_i \sum_j \Delta x_j \rho_j (u_i - u_j) \psi_{ij} \right),$$
$$(5.12)$$

where $\eta_i = \Pi(\rho_i) + \frac{1}{2}\rho_i u_i^2$ and

$$G_{i+1/2} = \mathcal{G}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) + \mathcal{F}^\rho\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) H_{i+1/2}.$$

(v) the discrete analogue of the free energy dissipation property in (4.8) given by

$$\frac{d}{dt}E^\Delta(t) \leq -\gamma\sum_i \Delta x_i \rho_i u_i^2 - \frac{1}{2}\sum_{i,j}\Delta x_i \Delta x_j \rho_i \rho_j (u_i - u_j)^2 \psi_{ij} \tag{5.13}$$

with

$$E^\Delta = \sum_i \frac{\Delta x_i}{2}\rho_i u_i^2 + \mathcal{F}^\Delta \quad \text{and} \quad \mathcal{F}^\Delta = \sum_i \Delta x_i \left[\Pi(\rho_i) + V_i \rho_i\right] + \frac{1}{2}\sum_{i,j}\Delta x_i \Delta x_j W_{ij}\rho_i \rho_j. \tag{5.14}$$

(vi) the discrete analogue of the evolution for centre of mass in (4.11),

$$\frac{d}{dt}\left(\sum_i \Delta x_i \rho_i x_i\right) = \sum_i \Delta x_i \mathcal{F}^\rho\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right), \tag{5.15}$$

which is reduced to

$$\sum_i \Delta x_i \rho_i x_i = 0 \tag{5.16}$$

when the initial density is symmetric and the initial velocity antisymmetric. This implies that the discrete centre of mass is conserved in time and centred at 0.

Proof. Some of the following proofs follow the lines considered in [11, 131].

(i) If a first-order numerical flux $\mathcal{F}(U_i, U_{i+1}) = (\mathcal{F}^\rho, \mathcal{F}^{\rho u})(U_i, U_{i+1})$ for the homogeneous problem, such as the Lax-Friedrich scheme detailed in Appendix 5.A, satisfies the nonnegativity of the density $\rho_i(t)$, then it necessarily follows that

$$\mathcal{F}^\rho((\rho_i = 0, u_i), (\rho_{i+1}, u_{i+1})) - \mathcal{F}^\rho((\rho_{i-1}, u_{i-1}), (\rho_i = 0, u_i)) \leq 0 \quad \forall(\rho_i, u_i)_i. \tag{5.17}$$

In our case, the sources do not contribute to the continuity equation in (5.1), and for the numerical flux in (5.7) we need to check that

$$\mathcal{F}^\rho\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) - \mathcal{F}^\rho\left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+}\right) \leq 0 \tag{5.18}$$

whenever $\rho_i = 0$. When $\rho_i = 0$, the reconstruction in (5.8) and (5.9) yields $\rho_{i+1/2}^{HR,-} = \rho_{i+1/2}^{HR,+} = 0$ since $\Pi(\rho)$ is assumed to be convex, and (5.18) results in

$$\mathcal{F}^\rho((0, u_i), (\rho_{i+1/2}^{HR,+}, u_{i+1})) - \mathcal{F}^\rho((\rho_{i-1/2}^{HR,-}, u_{i-1}), (\rho_i = 0, u_i)) \leq 0 \quad \forall(\rho_{i+1/2}^{HR,+}, \rho_{i+1/2}^{HR,-}, u_i)_i. \tag{5.19}$$

Then, given that the numerical scheme is chosen so that it preserves the nonnegativity of the density for the homogeneous problem and (5.17) holds, it follows that (5.19) is satisfied too.

(ii) To preserve the steady state the discrete fluxes and source need to be balanced,

$$F_{i+1/2} - F_{i-1/2} = \Delta x S_i. \tag{5.20}$$

When the steady state holds it follows from (5.9) that $\rho_{i+1/2}^{HR,-} = \rho_{i+1/2}^{HR,+}$ and $u_{i+1/2}^{HR,-} = u_{i-1/2}^{HR,+} = 0$, and as a result $U_{i+1/2}^{HR,-} = U_{i+1/2}^{HR,+}$. Then, by consistency of the numerical flux $\mathcal{F}$,

$$
F_{i+1/2} = \mathcal{F}\left((\rho_{i+1/2}^{HR,-}, 0), (\rho_{i+1/2}^{HR,+}, 0)\right) = F(U_{i+1/2}^{HR,-}) = F(U_{i+1/2}^{HR,+}) = \begin{pmatrix} 0 \\ P(\rho_{i+1/2}^{HR,-}) \end{pmatrix} = \begin{pmatrix} 0 \\ P(\rho_{i+1/2}^{HR,+}) \end{pmatrix}.
$$
(5.21)

Concerning the source term S_i of (5.10), in the steady state it is equal to

$$
\Delta x_i S_i = \begin{pmatrix} 0 \\ P\left(\rho_{i+1/2}^{HR,-}\right) - P\left(\rho_{i-1/2}^{HR,+}\right) \end{pmatrix}.
$$
(5.22)

Then the balance in (5.20) is obtained from (5.21) and (5.22).

(iii) For the consistency with the original system of (4.5) one has to apply the criterion in [42], by which two properties concerning the consistency with the exact flux F and the consistency with the source term need to be checked. Before proceeding, the finite volume discretization in (5.2) needs to be rewritten in a non-conservative form as

$$
\frac{dU_i}{dt} = -\frac{\mathcal{F}_l(U_i, U_{i+1}, H_i, H_{i+1}) - \mathcal{F}_r(U_{i-1}, U_i, H_{i-1}, H_i)}{\Delta x_i}
$$
$$
- \begin{pmatrix} 0 \\ \gamma \rho_i u_i + \rho_i \sum_j (u_i - u_j) \rho_j \psi(x_i - x_j) \end{pmatrix}
$$
(5.23)

where

$$
\mathcal{F}_l(U_i, U_{i+1}, H_i, H_{i+1}) = F_{i+1/2} - \Delta x_i S_{i+1/2}^{HR,-},
$$
$$
\mathcal{F}_r(U_{i-1}, U_i, H_{i-1}, H_i) = F_{i-1/2} + \Delta x_i S_{i-1/2}^{HR,+}.
$$

Here the source term S_i is considered as being distributed along the cells interfaces, satisfying

$$
S_i = S_{i+1/2}^{HR,-} + S_{i-1/2}^{HR,+} - \begin{pmatrix} 0 \\ \gamma \rho_i u_i + \rho_i \sum_j (u_i - u_j) \rho_j \psi(x_i - x_j) \end{pmatrix},
$$
$$
S_{i+1/2}^{HR,-} = \frac{1}{\Delta x_i} \begin{pmatrix} 0 \\ P(\rho_{i+1/2}^{HR,-}) - P(\rho_i) \end{pmatrix} \quad \text{and} \quad S_{i-1/2}^{HR,+} = \frac{1}{\Delta x_i} \begin{pmatrix} 0 \\ P(\rho_i) - P(\rho_{i-1/2}^{HR,+}) \end{pmatrix}.
$$

The consistency with the exact flux means that $\mathcal{F}_l(U, U, H, H) = \mathcal{F}_r(U, U, H, H) = F(U)$. This is directly satisfied since $U_{i+1/2}^{HR,-} = U_i$ and $U_{i+1/2}^{HR,+} = U_{i+1}$ whenever $H_{i+1} = H_i$, due to (5.9).

For the consistency with the source term the criterion to check is

$$
\mathcal{F}_r(U_i, U_{i+1}, H_i, H_{i+1}) - \mathcal{F}_l(U_i, U_{i+1}, H_i, H_{i+1}) = \begin{pmatrix} 0 \\ -\rho(H_{i+1} - H_i) + o(H_{i+1} - H_i) \end{pmatrix}
$$

as $U_i, U_{i+1} \to U$ and $H_i, H_{i+1} \to H$. For this case,

$$
\mathcal{F}_r(U_i, U_{i+1}, H_i, H_{i+1}) - \mathcal{F}_l(U_i, U_{i+1}, H_i, H_{i+1}) = \begin{pmatrix} 0 \\ S_{i+1/2}^{HR,+} + S_{i+1/2}^{HR,-} \end{pmatrix} =
$$
$$
\begin{pmatrix} 0 \\ -(P(\xi(\Pi'(\rho_{i+1}) + H_{i+1} - H_{i+1/2}) - P(\rho_{i+1})) + (P(\xi(\Pi'(\rho_i) + H_i - H_{i+1/2}) - P(\rho_i)) \end{pmatrix},
$$
(5.24)

where $H_{i+1/2} = \max(H_i, H_{i+1})$. By assuming without loss of generality that $H_{i+1/2} = H_i$, the second term of the last matrix results in

$$-P(\xi(\Pi'(\rho_{i+1}) + H_{i+1} - H_i)) + P(\xi(\Pi'(\rho_i))) = -P(\xi(\Pi'(\rho_{i+1}) + H_{i+1} + H_i)) - P(\rho_i).$$

This term can be further approximated as

$$-(P \circ \xi)'(\Pi'(\rho_{i+1}))(H_{i+1} - H_i) + o(H_{i+1} - H_i) = -\rho_{i+1}(H_{i+1} - H_i) + o(H_{i+1} - H_i)$$

since

$$(P \circ \xi)'(\Pi'(\rho_{i+1})) = P'(\rho_{i+1})\frac{1}{\Pi''(\rho_{i+1})} = \rho_{i+1}$$

by taking derivatives in $(\xi \circ \Pi')(\rho) = \rho$ and making use of (13.6). Finally, since $\rho_{i+1} \to \rho$, the consistency with the source term is satisfied. An analogous procedure can be followed whenever $H_{i+1/2} = H_{i+1}$.

(iv) To prove (5.12) we follow the strategy from [131]. We first set $G_{i+1/2}$ to be

$$G_{i+1/2} = \mathcal{G}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) + \mathcal{F}^\rho\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right)H_{i+1/2}.$$

Subsequently, and employing the inequalities for $\mathcal{G}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right)$ in (5.4), it follows that

$$\begin{aligned}
G_{i+1/2} - G_{i-1/2} &\leq G\left(U_{i+1/2}^{HR,-}\right) + \nabla_U\eta\left(U_{i+1/2}^{HR,-}\right)\left(\mathcal{F}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) - F\left(U_{i+1/2}^{HR,-}\right)\right) \\
&\quad - G\left(U_{i-1/2}^{HR,+}\right) - \nabla_U\eta\left(U_{i-1/2}^{HR,+}\right)\left(\mathcal{F}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) - F\left(U_{i-1/2}^{HR,+}\right)\right) \\
&\quad + \mathcal{F}^\rho\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right)H_{i+1/2} - \mathcal{F}^\rho\left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+}\right)H_{i-1/2}.
\end{aligned}$$

This last inequality can be rewritten after some long computations as

$$\begin{aligned}
G_{i+1/2} - G_{i-1/2} &\leq \left(\Pi'\left(\rho_{i+1/2}^{HR,-}\right) - \frac{1}{2}u_i^2 + H_{i+1/2}\right)\mathcal{F}^\rho\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) \\
&\quad - \left(\Pi'\left(\rho_{i-1/2}^{HR,+}\right) - \frac{1}{2}u_i^2 + H_{i-1/2}\right)\mathcal{F}^\rho\left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+}\right) \\
&\quad + u_i\left(\mathcal{F}^{\rho u}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) - \mathcal{F}^{\rho u}\left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+}\right) + P\left(\rho_{i-1/2}^{HR,+}\right) - P\left(\rho_{i+1/2}^{HR,-}\right)\right).
\end{aligned}$$

From here, by bearing in mind the definition of $\rho_{i+1/2}^{HR,-}$ and $\rho_{i-1/2}^{HR,+}$ in (5.9) and the definition of the scheme in (5.2)-(5.7)-(5.10), we get

$$\begin{aligned}
G_{i+1/2} - G_{i-1/2} &\leq \left(\Pi'(\rho_i) - \frac{1}{2}u_i^2 + H_i\right)\left(\mathcal{F}^\rho\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) - \mathcal{F}^\rho\left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+}\right)\right) \\
&\quad + u_i\left(\mathcal{F}^{\rho u}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) - \mathcal{F}^{\rho u}\left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+}\right) + P\left(\rho_{i-1/2}^{HR,+}\right) - P\left(\rho_{i+1/2}^{HR,-}\right)\right) \\
&= -\left(\Pi'(\rho_i) - \frac{1}{2}u_i^2 + H_i\right)\Delta x_i\frac{d\rho_i}{dt} - \Delta x_i u_i\frac{d}{dt}(\rho_i u_i) \\
&\quad - u_i\left(\gamma\Delta x_i\rho_i u_i + \Delta x_i\rho_i\sum_j\rho_j(u_i - u_j)\psi_{ij}\right).
\end{aligned}$$

Finally, this last inequality results in the desired cell entropy inequality in (5.12) by rearranging according to (5.1), yielding

$$\Delta x_i\frac{d\eta_i}{dt} + \Delta x_i H_i\frac{d\rho_i}{dt} + G_{i+1/2} - G_{i-1/2} = -u_i\left(\gamma\Delta x_i\rho_i u_i + \Delta x_i\rho_i\sum_j\rho_j(u_i - u_j)\psi_{ij}\right). \tag{5.25}$$

(v) The last property of the scheme and formulas (5.13)-(5.14) follow by summing over the index i over identity in (5.25), collecting terms and symmetrizing the dissipation using the symmetry of ψ.

(vi) Starting from the finite volume equation for the density in (5.1),

$$\Delta x_i \frac{d\rho_i}{dt} = -\mathcal{F}^\rho \left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+} \right) + \mathcal{F}^\rho \left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+} \right),$$

one can multiply it by x_i and sum it over the index i, resulting in

$$\frac{d}{dt} \left(\sum_i \Delta x_i \rho_i x_i \right) = \sum_i x_i \left(-\mathcal{F}^\rho \left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+} \right) + \mathcal{F}^\rho \left(U_{i-1/2}^{HR,-}, U_{i-1/2}^{HR,+} \right) \right).$$

By rearranging and considering, for instance, periodic or no flux boundary conditions, we get (5.15).

On the other hand, the finite volume equation for the momentum in (5.1), after summing over the index i, becomes

$$\frac{d}{dt} \left(\sum_i \Delta x_i \rho_i u_i \right) = \sum_i \left(P \left(\rho_{i+1/2}^{HR,-} \right) - P \left(\rho_{i-1/2}^{HR,+} \right) \right) - \gamma \sum_i \Delta x_i \rho_i u_i$$
$$- \sum_{i,j} \Delta x_i \Delta x_j \rho_i \rho_j (u_i - u_j) \psi_{ij}, \tag{5.26}$$

since the numerical fluxes cancel out due to the sum over the index i. In addition, the Cucker-Smale damping term also vanishes due to the symmetry in $\psi(x)$. Finally, if the initial density is symmetric and the initial velocity antisymmetric, the sum of pressures in the RHS of (5.26) is 0, due to the symmetry in the density. This implies that the discrete solution for the density and momentum maintains those symmetries, since (5.26) is simplified as

$$\sum_i \Delta x_i \rho_i u_i = 0$$

and as a result (5.15) reduces to (5.16). This means that the discrete centre of mass is conserved in time and is centred at 0, for initial symmetric densities and initial antisymmetric velocities.

$\square$

Remark 5.1.1. As a consequence of the previous proofs, our scheme conserves all the structural properties of the hydrodynamic system (4.5) at the semidiscrete level including the dissipation of the discrete free energy (4.8) and the characterization of the steady states. These properties are analogous to those obtained for finite volume schemes in the overdamped limit [55, 264].

Remark 5.1.2. All the previous properties, which are applicable for free energies of the form (4.2), can be extended to the general free energies in (13.7). It can be shown indeed that the discrete analogue of the free energy dissipation in (5.13) still holds for a discrete total energy defined as in (5.14) and a discrete free energy of the form

$$\mathcal{F}^\Delta = \sum_i \Delta x_i \left[\Pi \left(\rho_i \right) + V_i \rho_i \right] + \frac{1}{2} \sum_i \Delta x_i \rho_i K_i, \tag{5.27}$$

where K_i is a discrete approximation of $K(W(x) \star \rho)$ at the node x_i and is evaluated as

$$K_i = K \left(\sum_j \Delta x_j W_{ij} \rho_j \right). \tag{5.28}$$

5.2   Second-order extension

The usual procedure to extend a first-order scheme to second order is by computing the numerical fluxes (5.3) from reconstructed values of the density and momentum at each side of the boundary, contrary to the cell-centred values taken for the first-order schemes (5.5). These values are classically computed in three steps: prediction of the gradients in each cell, linear extrapolation and limiting procedure to preserve nonnegativity. For instance, MUSCL [223] is a usual reconstruction procedure following these steps. From here the values $\rho_{i+1/2}^{+}$, $\rho_{i+1/2}^{-}$, $u_{i+1/2}^{+}$ and $u_{i+1/2}^{-}$ are obtained $\forall i$, where $\rho_{i+1/2}^{+}$ indicates the reconstruction at the right of the boundary $x_{i+1/2}$ and $\rho_{i+1/2}^{-}$ at the left. Then the inputs for the numerical flux in (5.3), for a usual second-order scheme, are

$$F_{i+1/2} = \mathcal{F}\left(U_{i+1/2}^{-}, U_{i+1/2}^{+}\right).$$

This procedure has already been adapted to satisfy the well-balanced property and maintain the second order for specific applications, such as shallow water [11] or chemotaxis [131]. In this subsection the objective is to extend the procedure to general free energies of the form (4.2). As it happened for the well-balanced first-order scheme, the boundary values introduced in the numerical flux, which in this case are $U_{i+1/2}^{-}$ and $U_{i+1/2}^{+}$, need to be adapted to satisfy the well-balanced property.

For the well-balanced scheme the first step is to reconstruct the boundary values $\rho_{i-1/2}^{+}$, $\rho_{i+1/2}^{-}$, $u_{i-1/2}^{+}$ and $u_{i+1/2}^{-}$ following the three mentioned steps. In addition, the reconstructed values of the potential $H(x, \rho)$ at the boundaries, $H_{i+1/2}^{-}$ and $H_{i-1/2}^{+}$ $\forall i$, have to be also computed. This is done as suggested in [11]. Instead of reconstructing directly $H_{i+1/2}^{-}$ and $H_{i-1/2}^{+}$ following the three mentioned steps, for certain applications one has to reconstruct firstly $(\Pi'(\rho) + H(x, \rho))_i$ to obtain $(\Pi'(\rho) + H(x, \rho))_{i+1/2}^{-}$ and $(\Pi'(\rho) + H(x, \rho))_{i-1/2}^{+}$, and subsequently compute $H_{i+1/2}^{-}$ and $H_{i-1/2}^{+}$ as

$$H_{i+1/2}^{-} = \left(\Pi'(\rho) + H(x, \rho)\right)_{i+1/2}^{-} - \Pi'\left(\rho_{i+1/2}^{-}\right),$$

$$H_{i-1/2}^{+} = \left(\Pi'(\rho) + H(x, \rho)\right)_{i-1/2}^{+} - \Pi'\left(\rho_{i-1/2}^{+}\right).$$

This is shown in [11] to be necessary in order to maintain nonnegativity and the steady state in applications where there is an interface between dry and wet cells. For instance, these interfaces appear when considering pressures of the form $P = \rho^m$ with $m > 0$, as it is shown in Examples 5.3.4 and 5.3.6 of Section 5.3. For other applications where vacuum regions do not occur, the values $H_{i+1/2}^{-}$ and $H_{i-1/2}^{+}$ can be directly reconstructed following the three mentioned steps.

After this first step, the inputs for the numerical flux are updated from (5.3) to satisfy the well-balanced property as

$$F_{i+1/2} = \mathcal{F}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right), \text{ where } U_{i+1/2}^{HR,-} = \begin{pmatrix} \rho_{i+1/2}^{HR,-} \\ \rho_{i+1/2}^{HR,-} u_{i+1/2}^{-} \end{pmatrix}, \quad U_{i+1/2}^{HR,+} = \begin{pmatrix} \rho_{i+1/2}^{HR,+} \\ \rho_{i+1/2}^{HR,+} u_{i+1/2}^{+} \end{pmatrix}.$$

The interface values $\rho_{i+1/2}^{HR,\pm}$ are reconstructed as in the first-order scheme, by taking into account the steady state relation in (5.6). The application of (5.6) to the cells with centred nodes x_i and x_{i+1} leads to

$$\Pi'\left(\rho_{i+1/2}^{HR,-}\right) + H_{i+1/2} = \Pi'\left(\rho_{i+1/2}^{-}\right) + H_{i+1/2}^{-},$$

$$\Pi'\left(\rho_{i+1/2}^{HR,+}\right) + H_{i+1/2} = \Pi'\left(\rho_{i+1/2}^{+}\right) + H_{i+1/2}^{+},$$

where the term $H_{i+1/2}$ is evaluated to preserve consistency and stability, with an upwind or average value obtained as

$$H_{i+1/2} = \max\left(H_{i-1/2}^{+}, H_{i+1/2}^{+}\right) \quad \text{or} \quad H_{i+1/2} = \frac{1}{2}\left(H_{i-1/2}^{+} + H_{i+1/2}^{+}\right).$$

Then, by denoting as $\xi(x)$ the inverse function of $\Pi'(x)$, the interface values $\rho_{i+1/2}^{HR,\pm}$ are computed as

$$\rho_{i+1/2}^{HR,-} = \xi\left(\Pi'\left(\rho_{i+1/2}^{-}\right) + H_{i+1/2}^{-} - H_{i+1/2}\right),$$

$$\rho_{i+1/2}^{HR,+} = \xi\left(\Pi'\left(\rho_{i+1/2}^{+}\right) + H_{i+1/2}^{+} - H_{i+1/2}\right).$$

The source term is again distributed along the interfaces,

$$S_{H,i} = S_{i+1/2}^{HR,-} + S_{i-1/2}^{HR,+} + S_i^c,$$

where

$$S_{i+1/2}^{HR,-} = \frac{1}{\Delta x_i}\begin{pmatrix} 0 \\ P\left(\rho_{i+1/2}^{HR,-}\right) - P\left(\rho_{i+1/2}^{-}\right) \end{pmatrix}, \quad S_{i-1/2}^{HR,+} = \frac{1}{\Delta x_i}\begin{pmatrix} 0 \\ P\left(\rho_{i-1/2}^{+}\right) - P\left(\rho_{i-1/2}^{HR,+}\right) \end{pmatrix}.$$

The inclusion of the central source term S_i^c is vital in order to preserve the second-order accuracy and well-balanced property of the scheme. This idea was firstly introduced in [173], where second order error estimates are derived under certain conditions for S_i^c. Further works customize this central source term S_i^c for particular applications such as shallow water equations [11, 174] or chemotaxis [131]. There is some flexibility in the choice of this term, as far as it satisfies two criteria for second-order accuracy and well-balancing. In the following remark we summarize the two criteria, which are described with more extend in Ref. [42] (specifically, (4.187) for second-order accuracy, and (4.204) for well-balancing).

Remark 5.2.1. The central source term S_i^c preserves the second-order accuracy and well-balanced property of the scheme if the following two criteria are satisfied:

(i) Second-order accuracy if

$$S_i^c\left(\rho_{i-1/2}^{+}, \rho_{i+1/2}^{-}, H_{i+1/2}^{-}, H_{i-1/2}^{+}\right) =$$

$$\begin{pmatrix} 0 \\ \left(\left(-\frac{\rho_{i-1/2}^{+} + \rho_{i+1/2}^{-}}{2} + \mathcal{O}\left(\left|\rho_{i+1/2}^{-} - \rho_{i-1/2}^{+}\right|^2 + \left|H_{i-1/2}^{+} - H_{i+1/2}^{-}\right|^2\right)\right)\left(H_{i-1/2}^{+} - H_{i+1/2}^{-}\right)\right) \end{pmatrix} \tag{5.29}$$

as $\rho_{i+1/2}^{-} - \rho_{i-1/2}^{+} \to 0$ and $H_{i-1/2}^{+} - H_{i+1/2}^{-} \to 0$.

(ii) Well-balanced property if

$$S_i^c\left(\rho_{i-1/2}^{+}, \rho_{i+1/2}^{-}, H_{i+1/2}^{-}, H_{i-1/2}^{+}\right) = F\left(\rho_{i+1/2}^{-}, H_{i-1/2}^{+}\right) - F\left(\rho_{i-1/2}^{+}, H_{i+1/2}^{-}\right), \tag{5.30}$$

meaning that the steady states are let invariant.

The objective here is to provide a general form of S_i^c which applies to general free energies of the form (4.2). Following the strategy in [42], we propose to approximate the generalized centred sources as

$$S_i^c = \frac{1}{\Delta x_i}\begin{pmatrix} 0 \\ P(\rho_{i+1/2}^{-}) - P(\rho_{i+1/2}^{-,*}) - P(\rho_{i-1/2}^{+}) + P(\rho_{i-1/2}^{+,*}) \end{pmatrix} - \begin{pmatrix} 0 \\ \gamma\rho_i u_i + \rho_i \sum_j (u_i - u_j)\rho_j \psi(x_i - x_j) \end{pmatrix},$$

where the values $\rho_{i-1/2}^{+,*}$ and $\rho_{i+1/2}^{-,*}$ are computed from the steady state relation (5.6) as

$$\rho_{i-1/2}^{+,*} = \xi\left(\Pi'\left(\rho_{i-1/2}^{+}\right) + H_{i-1/2}^{+} - H_i^*\right),$$

$$\rho_{i+1/2}^{-,*} = \xi\left(\Pi'\left(\rho_{i+1/2}^{-}\right) + H_{i+1/2}^{-} - H_i^*\right),$$

and H_i^* is a centred approximation of the potentials satisfying

$$H_i^* = \frac{1}{2}(H_{i+1/2}^- + H_{i-1/2}^+).$$

The proposed structure of S_i^c is suggested in [42] and satisfies the two criteria for second-order accuracy (5.29) and well-balanced property (5.30).

The source term with damping $S_D(x, U)$ is discretized as in the first-order scheme,

$$S_{D,i} = - \begin{pmatrix} 0 \\ \gamma \rho_i u_i + \rho_i \sum_j \Delta x_j (u_i - u_j) \rho_j \psi_{ij} \end{pmatrix}.$$

Overall, the second-order semidiscrete scheme defined in (5.2) and constructed as detailed in Section 5.2, and for a numerical flux $\mathcal{F}$ satisfying the properties stated in the introduction of Chapter 5, satisfies:

(i) preservation of the nonnegativity of $\rho_i(t)$;

(ii) well-balanced property, thus preserving the steady states given by (5.6);

(iii) consistency with the system (4.5);

(iv) second-order accuracy.

The proof of these properties is omitted here since it follows the same techniques from [11, 131], and the general procedure is very similar to the one from the first-order scheme in Subsection 5.1.1.

5.3 Numerical simulations

This section details numerical simulations in which the first- and second-order schemes from Chapter 5 are employed. Firstly, Subsection 5.3.1 contains the validation of the first- and second-order schemes: the well-balanced property and the order of accuracy of the schemes are tested in four different configurations. Secondly, Subsection 5.3.2 illustrates the application of the numerical schemes to a variety of choices of the free energy, leading to interesting numerical experiments for which analytical results are limited in the literature.

Unless otherwise stated, all simulations contain linear damping with $\gamma = 1$ and have a total unitary mass. Only the indicated ones contain the Cucker-Smale damping term, where the communication function satisfies

$$\psi(x) = \frac{1}{(1 + |x|^2)^{\frac{1}{4}}}.$$

The pressure function in the simulations has the form of $P(\rho) = \rho^m$, with $m \geq 1$. When $m = 1$ the pressure satisfies the ideal-gas relation $P(\rho) = \rho$, and the density does not develop vacuum regions during the temporal evolution. For this case the employed numerical flux is the versatile local Lax-Friedrich flux. For the simulations where $P(\rho) = \rho^m$ and $m > 1$ vacuum regions with $\rho = 0$ are generated. This implies that the hyperbolicity of the system (4.5) is lost in those regions, and the local Lax-Friedrich scheme fails. As a result, an appropriate numerical flux has to be implemented to handle the vacuum regions. In this case a kinetic solver based on [236], and already implemented in previous works [12], is employed.

The time discretization is accomplished by means of the third order TVD Runge-Kutta method [156] and the CFL number is taken as 0.7 in all the simulations. The boundary conditions are chosen to be no flux. For more details about the numerical fluxes, temporal discretization, boundary conditions and CFL number, we remit the reader to Appendix 5.A.

Videos from all the simulations displayed in this work are available at [232].

5.3.1 Validation of the numerical scheme

The validation of the schemes from Chapter 5 includes a test for the well-balanced property and a test for the order of accuracy in the transient regimes. These tests are completed in four different examples with steady states satisfying (4.6), which differ in the choice of the free energy, potentials and the inclusion of Cucker-Smale damping terms. An additional fifth example presenting moving steady states of the form (4.12) is considered to show that our schemes satisfy the order of accuracy test even for this challenging steady states.

The well-balanced property test evaluates whether the steady state solution is preserved in time up to machine precision. As a result, the initial condition of the simulation has to be directly the steady state. The results of this test for the four examples of this section are presented in Table 5.1. All the simulations are run from $t = 0$ to $t = 5$, and the number of cells is 50.

Table 5.1: Preservation of the steady state for Examples 5.3.1 to 5.3.4 with the first- and second-order schemes and double precision, at $t = 5$

	Order of the scheme	L^1 error	L^∞ error
Example 5.3.1	1$^{\text{st}}$	9.1012E-18	1.1102E-16
	2$^{\text{nd}}$	2.3191E-17	2.2843E-16
Example 5.3.2	1$^{\text{st}}$	7.8666E-18	1.1102E-16
	2$^{\text{nd}}$	1.4975E-17	1.5057E-16
Example 5.3.3	1$^{\text{st}}$	5.5020E-17	6.6613E-16
	2$^{\text{nd}}$	6.4514E-17	7.2164E-16
Example 5.3.4	1$^{\text{st}}$	1.3728E-17	2.2204E-16
	2$^{\text{nd}}$	3.4478E-18	1.1102E-16

The order of accuracy in the transient regimes test is based on evaluating the L^1 error of a numerical solution for a particular choice of Δx with respect to a reference solution, and for a time when the steady state is not reached yet. Subsequent L^1 errors are obtained after halving the Δx of the previous numerical solution, doubling in this way the total number of cells. The order of the scheme is then computed as

$$\text{Order of the scheme} = \ln_2\left(\frac{L^1\,\text{error}(\Delta x)}{L^1\,\text{error}(\Delta x/2)}\right), \tag{5.31}$$

and the Δx is halved four times.

The reference solution is frequently taken as an explicit solution of the system that is being tested. In this case, the system in (4.5) does not have an explicit solution in time for the free energies presented here, even though the steady solution can be analytically computed. Since we are interested in evaluating the order of accuracy away from equilibrium, the reference solution is computed from the same numerical scheme but with a really small Δx, so that the numerical solution can be considered as the exact one. In all cases here the reference solution is obtained from a mesh with 25600 cells, while the numerical solutions employ a number of cells between 50 and 400.

The results from the accuracy tests are shown in Tables 5.2 to 5.6. The simulations were run with the configurations specified in each example and from $t = 0$ to $t = 0.3$, unless otherwise stated. The final time of $t = 0.3$ is taken so that all examples are in the transient regime.

Example 5.3.1 (Ideal-gas pressure and attractive potential). In this example the pressure satisfies $P(\rho) = \rho$ and there is an external potential of the form $V(x) = \frac{x^2}{2}$. As a result, the relation holding in the steady state is

$$\frac{\delta\mathcal{F}}{\delta\rho} = \Pi'(\rho) + H = \ln(\rho) + \frac{x^2}{2} = \text{constant on supp}(\rho) \text{ and } u = 0. \tag{5.32}$$

The steady state, for an initial mass M_0, explicitly satisfies

$$\rho_\infty = M_0 \frac{e^{-x^2/2}}{\int_{\mathbb{R}} e^{-x^2/2} dx}. \tag{5.33}$$

For the order of accuracy test the initial conditions are

$$\rho(x, t=0) = M_0 \frac{0.2 + 5 \cos\left(\frac{\pi x}{10}\right)}{\int_{\mathbb{R}} \left(0.2 + 5 \cos\left(\frac{\pi x}{10}\right)\right) dx}, \quad \rho u(x, t=0) = -0.05 \sin\left(\frac{\pi x}{10}\right), \quad x \in [-5, 5], \tag{5.34}$$

with M_0 equal to 1 so that the total mass is unitary. The order of accuracy test from this example is shown in Table 5.2, and the evolution of the density, momentum, variation of the free energy with respect to the density, total energy and free energy are depicted in Figure 5.1. From Figure 5.1d one can notice how the discrete total energy always decreases in time, due to the discrete free energy dissipation property (5.13), and how there is an exchange between free energy and kinetic energy which makes the discrete free energy plot oscillate.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy
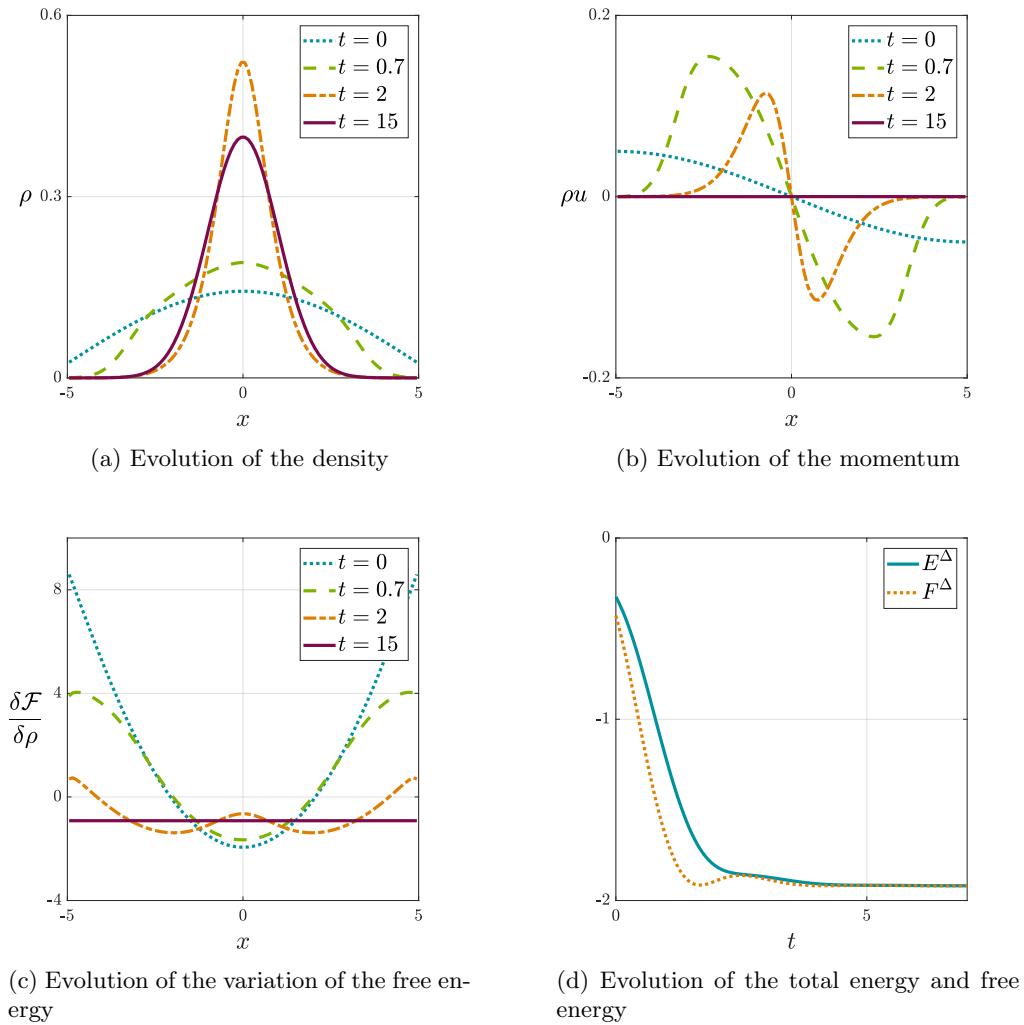
(d) Evolution of the total energy and free energy

Figure 5.1: Temporal evolution of Example 5.3.1.

Table 5.2: Accuracy test for Example 5.3.1 with the first and second-order schemes, at $t = 0.3$

Number of	First-order		Second-order	
cells	L^1 error	order	L^1 error	order
50	6.8797E-03	-	7.6166E-04	-
100	3.4068E-03	1.01	2.0206E-04	1.91
200	1.6826E-03	1.02	5.0308E-05	2.01
400	8.3104E-04	1.02	1.2879E-05	1.97

Example 5.3.2 (Ideal-gas pressure, attractive potential and Cucker-Smale damping terms). In this example the pressure satisfies $P(\rho) = \rho$ and there is an external potential of the form $V(x) = \frac{x^2}{2}$. The difference with Example 5.3.1 is that the Cucker-Smale damping terms are included, and the linear damping term $-\rho u$ excluded.

The relation holding in the steady state is expressed in (5.32) and the steady state satisfies (5.33). The initial conditions are also (5.34). The order of accuracy test from this example is shown in Table 5.3, and the evolution of the density, momentum, variation of the free energy with respect to the density, total energy and free energy are depicted in Figure 5.2. The lack of linear damping leads to higher oscillations in the momentum plots in comparison to Figure 5.1. There is also an exchange of kinetic and free energy during the temporal evolution, which could be noticed from the oscillations of the discrete free energy in Figure 5.2d.

Table 5.3: Accuracy test for Example 5.3.2 with the first and second-order schemes, at $t = 0.3$

Number of	First-order		Second-order	
cells	L^1 error	order	L^1 error	order
50	6.3195E-03	-	7.3045E-04	-
100	3.2658E-03	0.95	1.9462E-04	1.91
200	1.6373E-03	1.00	4.8629E-05	2.00
400	8.7771E-04	1.01	1.2468E-05	1.97

Example 5.3.3 (Ideal-gas pressure and attractive kernel). In this case study the pressure satisfies $P(\rho) = \rho$ and there is an interaction potential with a kernel of the form $W(x) = \frac{x^2}{2}$. The steady state for a general total mass M_0 is again equal to the steady states from Examples 5.3.1 and 5.3.2 with unit mass. The linear damping coefficient γ has been reduced, $\gamma = 0.01$, in order to compare the evolution with respect to the previous examples.

The initial conditions for the order of accuracy test are the ones from Example 5.3.1 in (5.34). The order of accuracy test from this example is shown in Table 5.4, and the evolution of the density, momentum, variation of the free energy with respect to the density, total energy and free energy are depicted in Figure 5.3. Due to the low value of γ in the linear damping, there is a repeated exchange of free energy and kinetic energy during the temporal evolution, which can be noticed from the oscillations of the free energy plot in Figure 5.3d. In the previous examples the linear damping term dissipates the momentum in a faster timescale and these exchanges only last for a few oscillations. One can also notice that the time to reach the steady state is higher than in the previous examples.

Example 5.3.4 (Pressure proportional to square of density and attractive potential). For this example the pressure satisfies $P(\rho) = \rho^2$ and there is an external potential of the form $V(x) = \frac{x^2}{2}$. Contrary to the previous Examples 5.3.1 to 5.3.3, the choice of $P(\rho) = \rho^2$ implies that regions of vacuum where $\rho = 0$ appear in the evolution and steady solution of the system. As

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy

Figure 5.2: Temporal evolution of Example 5.3.2.

Table 5.4: Accuracy test for Example 5.3.3 with the first and second-order schemes, at $t = 0.3$

Number of cells	First-order		Second-order	
	L^1 error	order	L^1 error	order
50	6.6938E-03	-	7.6135E-04	-
100	3.4702E-03	0.95	2.0207E-04	1.91
200	1.7410E-03	1.00	5.0306E-05	2.01
400	8.6890E-04	1.00	1.2879E-05	1.97

explained in the introduction of this section, the numerical flux employed for this case is a kinetic solver based on [42].

The steady state for this example with an initial mass of M_0 satisfies

$$
\rho_\infty(x) = \begin{cases} -\dfrac{1}{4}\left(x + \sqrt[3]{3M_0}\right)\left(x - \sqrt[3]{3M_0}\right) & \text{for} \quad x \in \left[-\sqrt[3]{3M_0},\, \sqrt[3]{3M_0}\right], \\ 0 & \text{otherwise.} \end{cases}
$$

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy

Figure 5.3:   Temporal evolution of Example 5.3.3.

The initial conditions taken for the order of accuracy test are

$$\rho(x, t=0) = M_0 \frac{0.1 + e^{-x^2}}{\int_{\mathbb{R}} \left(0.1 + e^{-x^2}\right) dx}, \quad \rho u(x, t=0) = -0.2 \sin\left(\frac{\pi x}{10}\right), \quad x \in [-5, 5],$$

with M_0 being the mass of the system and equal to 1. The order of accuracy test from this example is shown in Table 5.5, and the evolution of the density, momentum, variation of the free energy with respect to the density, total energy and free energy are depicted in Figure 5.4. The initial kinetic energy represents a large part of the initial total energy, and there is also an exchange between the kinetic energy and the free energy resulting in the oscillations for the plot of the discrete free energy.

As a remark, in this example the order of accuracy for the schemes with order higher than one is reduced to one both in the vacuum and interface regions, as it is also pointed out in [131]. The orders showed in Table 5.5 are computed by considering only the cells in the support of the density that are away from the interface region, and the vacuum regions are not taken into consideration.

Example 5.3.5 (Moving steady state with ideal-gas pressure, attractive kernel and Cucker-Smale damping term). The purpose of this example is to show that our scheme from

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy
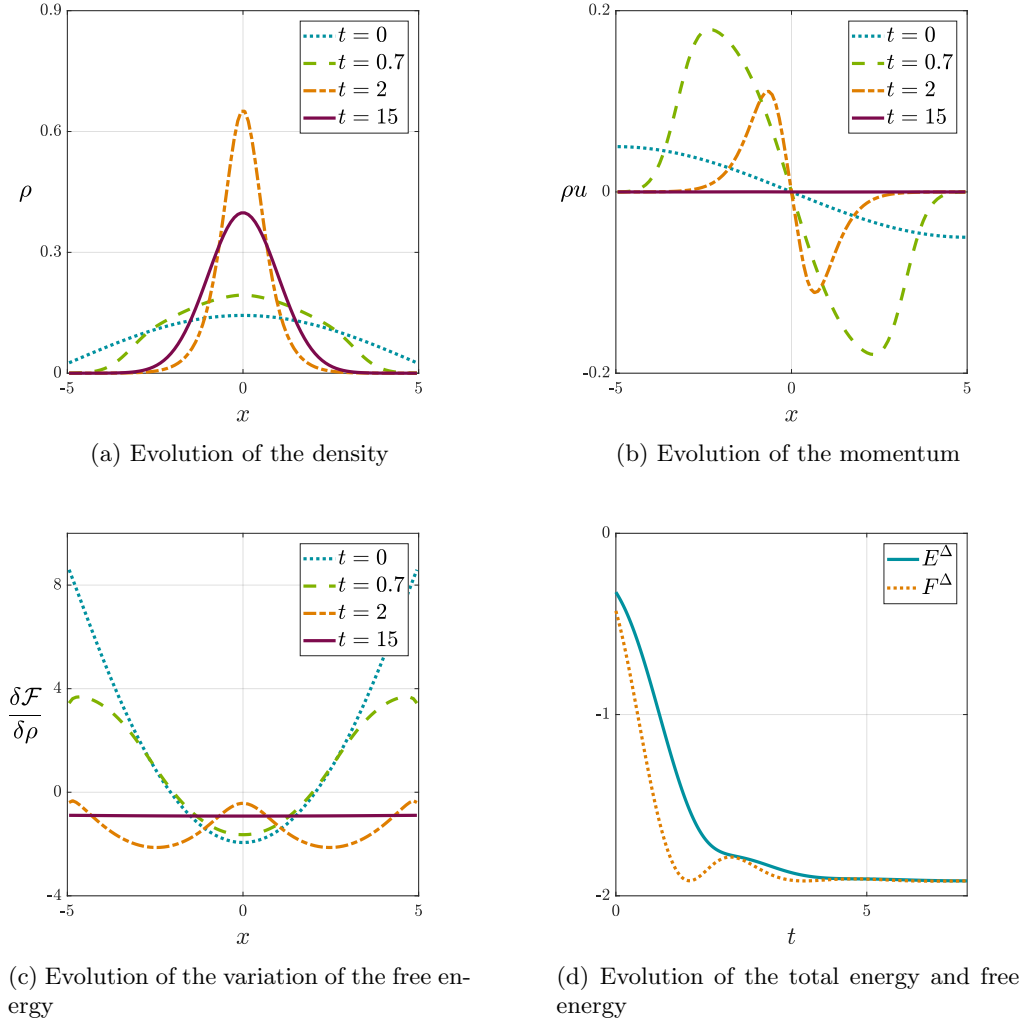
(d) Evolution of the total energy and free energy
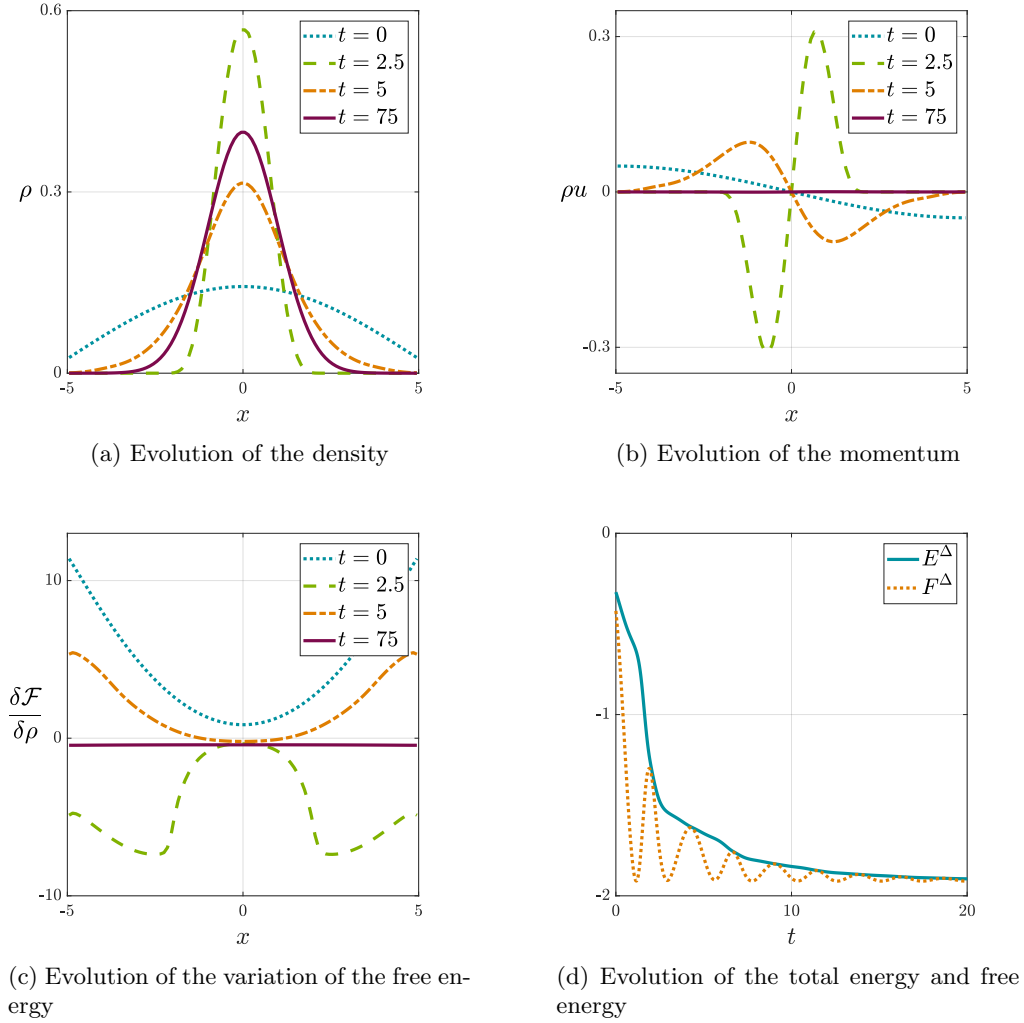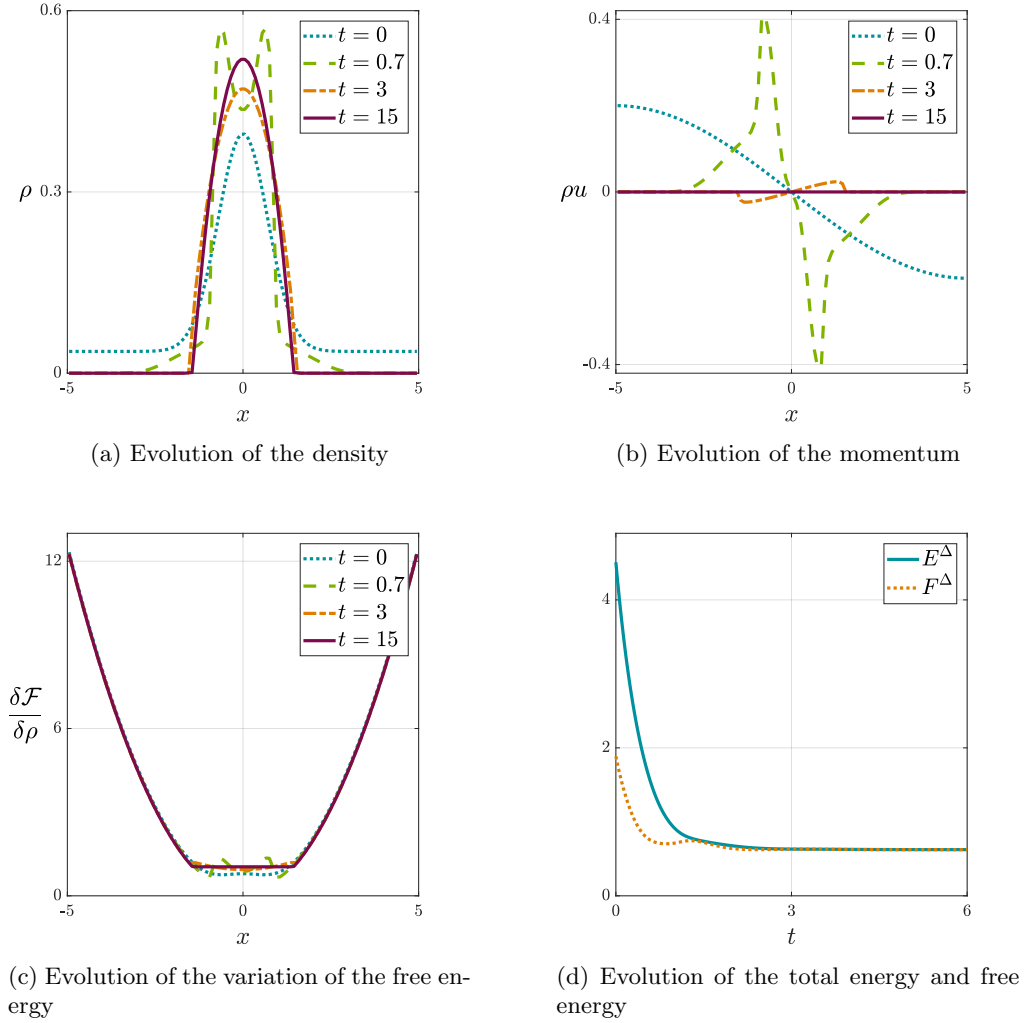
Figure 5.4:   Temporal evolution of Example 5.3.4.

Table 5.5: Accuracy test for Example 5.3.4 with the first and second-order schemes, at $t = 0.3$

Number of cells	First-order		Second-order	
	L^1 error	order	L^1 error	order
50	6.8826E-03	-	1.0735E-03	-
100	3.5106E-03	0.97	2.9188E-04	1.88
200	1.7596E-03	1.00	7.6113E-05	1.94
400	8.8184E-04	1.00	1.9103E-05	1.99

Chapter 5 preserves the order of accuracy for moving steady states of the form (4.12), where the velocity is not dissipated. As mentioned in the introduction, the generalization of well-balanced schemes to preserve moving steady states has proven to be quite complicated [218, 287], and it is not the aim of this work to construct such schemes.

For this example the pressure satisfies $P(\rho) = \rho$ and there is an interaction potential with a kernel of the form $W(x) = \frac{x^2}{2}$. The linear damping is eliminated and the Cucker-Smale damping term included. Under this configuration, there exists an explicit solution for system (4.5) consisting

in a travelling wave of the form

$$\rho(x,t) = M_0 \frac{e^{-(x-ut)^2/2}}{\int_{\mathbb{R}} e^{-x^2/2} dx}, \quad u(x,t) = 0.2, \tag{5.35}$$

with M_0 equals to 1 so that the total mass is unitary. As a result, the order of accuracy test can be accomplished by computing the error with respect to the exact reference solution, contrary to what was proposed in the previous examples. It should be remarked however that the velocity and the variation of the free energy with respect to the density profiles are not kept constant along the domain by our numerical scheme, since the well-balanced property for moving steady states is not satisfied.

The initial conditions for our simulation are (5.35) at $t = 0$, in a numerical domain with $x \in [-8, 9]$. The simulation is run until $t = 3$. The table of errors for different number of cells is showed in Table 5.6, and a depiction of the evolution of the system is illustrated in Figure 5.5. The velocity and the variation of the free energy plots are not included since they are not maintained constant with our scheme.

Table 5.6: Accuracy test for Example 5.3.5 with the first and second-order schemes, at $t = 3$

Number of cells	First-order		Second-order	
	L^1 error	order	L^1 error	order
50	9.84245E-03	-	2.78988E-03	-
100	4.92029E-03	1.00	9.09342E-04	1.62
200	2.44627E-03	1.01	2.55340E-04	1.83
400	1.21228E-03	1.01	7.47905E-05	1.77

(a) Evolution of the density

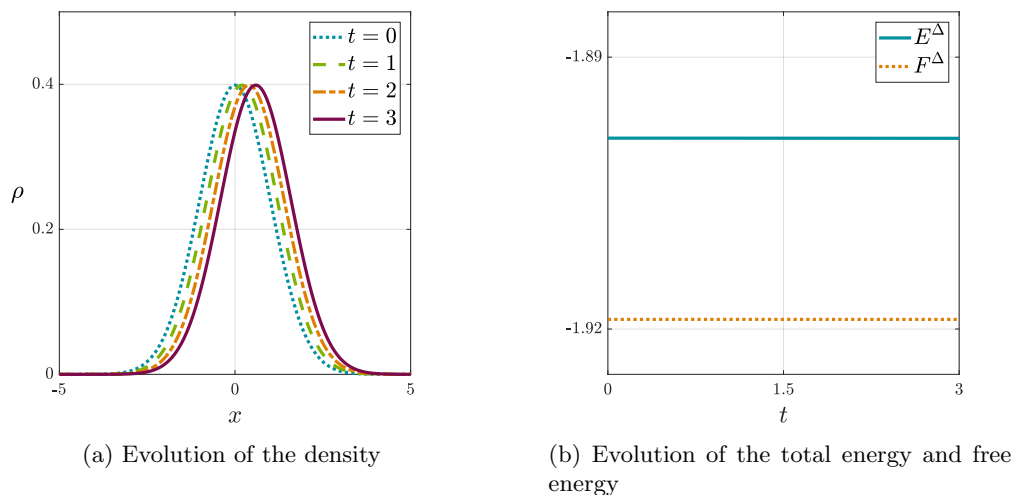(b) Evolution of the total energy and free energy

Figure 5.5: Temporal evolution of Example 5.3.5.

5.3.2 Numerical experiments and applications

This subsection applies the well-balanced scheme in Chapter 5 to a variety of free energies from systems which have acquired an important consideration in the literature. Some of these systems

have been mainly studied in their overdamped form, resulting when $\gamma \to \infty$, and as a result our well-balanced scheme can be useful in determining the role that inertia plays in those systems.

Example 5.3.6 (Pressure proportional to square of density and double-well potential). In this example the pressure is taken as in Example 5.3.4, with $P(\rho) = \rho^2$, thus leading to vacuum regions. The external potential are chosen to have a double-well shape of the form $V(x) = a\,x^4 - b\,x^2$, with $a, b > 0$. This system exhibits a variety of steady states depending on the symmetry of the initial condition, the initial mass and the shape of the external potential $V(x)$. The general expression for the steady states is

$$\rho_\infty = (C(x) - V(x))_+ = \left(C(x) - a\,x^4 + b\,x^2\right)_+,$$

where $C(x)$ is a piecewise constant function, zero outside the support of the density. Notice that $C(x)$ can attain a different value in each connected component of the support of the density.

Three different initial data are simulated in order to compare the resulting long time asymptotics, i.e., we show that different steady states are achieved corresponding to different initial data. The initial conditions are
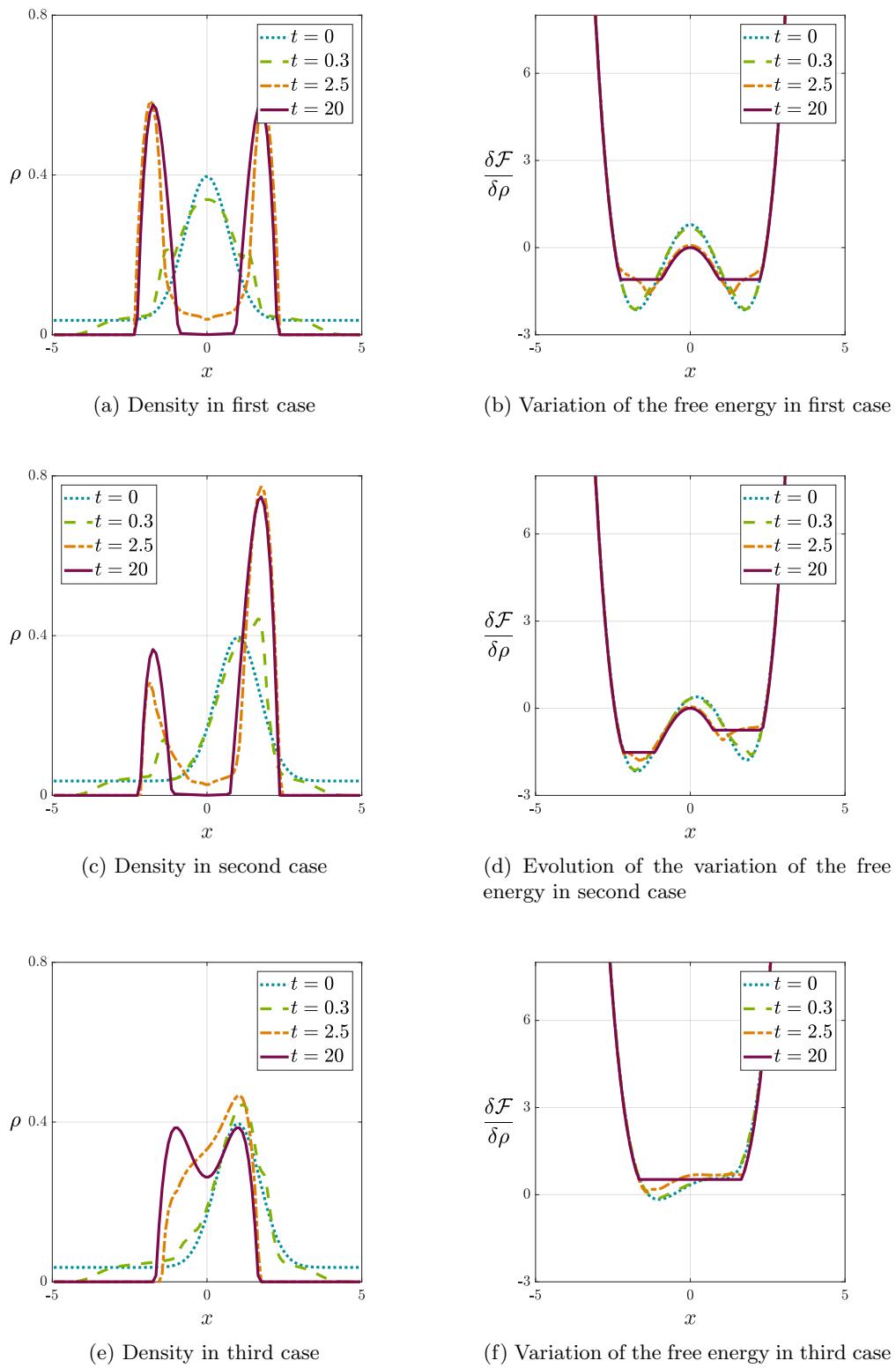
$$\rho(x, t=0) = M_0 \frac{0.1 + e^{-(x-x_0)^2}}{\int_\mathbb{R} \left(0.1 + e^{-(x-x_0)^2}\right) dx}, \quad \rho u(x, t=0) = -0.2 \sin\left(\frac{\pi x}{10}\right), \quad x \in [-10, 10],$$

with M_0 equal to 1 so that the total mass is unitary. When $x_0 = 0$, the initial density is symmetric, and when $x_0 \neq 0$ the initial density is asymmetric.

a. First case: The external potential satisfies $V(x) = \frac{x^4}{4} - \frac{3x^2}{2}$ and the initial density is symmetric with $x_0 = 0$. For this configuration the steady solution presents two disconnected bumps of density with the same mass in each of them, as it is shown in Figures 5.6a and 5.6b. The variation of the free energy with respect to the density presents the same constant value in the two disconnected supports of the density. The evolution is symmetric throughout.

b. Second case: The external potential satisfies $V(x) = \frac{x^4}{4} - \frac{3x^2}{2}$ and the initial density is asymmetric with $x_0 = 1$. The final steady density is characterised again by the two disconnected supports but for this configuration the mass in each of them varies, as shown in Figures 5.6c and 5.6d. Similarly, the variation of the free energy with respect to the density presents different constant values in the two disconnected supports of the density.

c. Third case: for this last configuration the external potential is varied and satisfies $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$, and the initial density is asymmetric with $x_0 = 1$. For this case, even though the initial density is asymmetric, the final steady density is symmetric and compactly supported due to the shape of the potential, as it is shown in Figures 5.6e and 5.6f. The variation of the free energy with respect to the density presents constant value in all the support of the density.

This behaviour shows that this problem has a complicated bifurcation diagram and corresponding stability properties depending on the parameters, for instance the coefficient on the potential well controlling the depth and support of the wells used above.

Example 5.3.7 (Ideal pressure with noise parameter and its phase transition). The model proposed for this example has a pressure satisfying $P(\rho) = \sigma\rho$, where σ is a noise parameter, and external and interaction potentials chosen to be $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$ and $W(x) = \frac{x^2}{2}$, respectively. The corresponding model in the overdamped limit has been previously studied in the context of collective behaviour [21], mean field limits [150], and systemic risk [141], see also [271] for the proof in one dimension.

(a) Density in first case

(b) Variation of the free energy in first case

(c) Density in second case

(d) Evolution of the variation of the free energy in second case

(e) Density in third case

(f) Variation of the free energy in third case

Figure 5.6: Temporal evolution of the first, second and third cases from Example 5.3.6.

We find that this hydrodynamic system exhibits a supercritical pitchfork bifurcation in the centre of mass $\hat{x}$ of the steady state when varying the noise parameter σ as its overdamped limit counterpart discussed above. For values of σ higher than a certain threshold, all steady states are symmetric and have the centre of mass $\hat{x}$ at $x = 0$. However, when σ decreases below that threshold, the pitchfork bifurcation takes place. On the one hand, if the centre of mass of the initial density is at $x = 0$, the final centre of mass in the steady state remains at $x = 0$. On the other hand, if the centre of mass of the initial density is at $x \neq 0$, the centre of mass of the steady state approaches asymptotically to $x = 1$ or $x = -1$ as $\sigma \to 0$, depending on the sign of the initial centre of mass. Finally, when $\sigma = 0$, the steady state turns into a Dirac delta at $x = 0$, $x = 1$ or $x = -1$, depending on the initial density. The pitchfork bifurcation is supercritical since
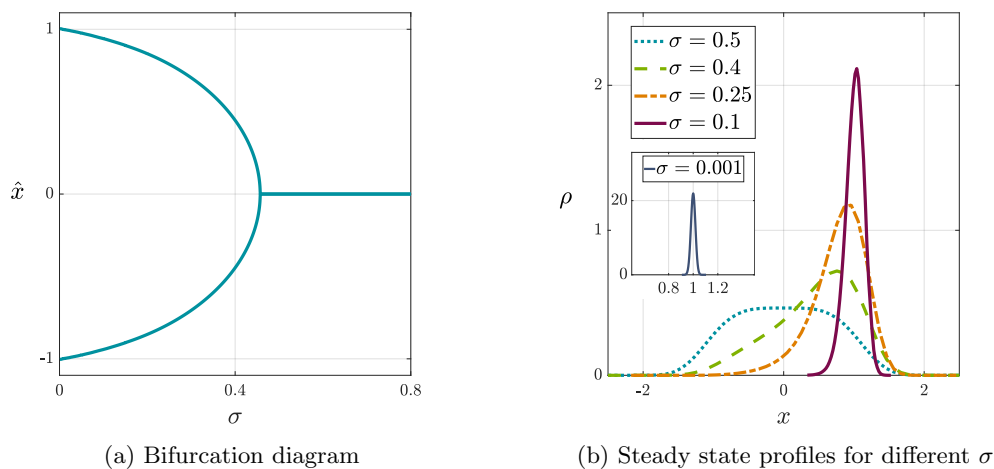

(a) Bifurcation diagram

(b) Steady state profiles for different σ

Figure 5.7: Bifurcation diagram (a) and steady states for different values of the noise parameter σ (b) from Example 5.3.7

the branch of the bifurcation corresponding to $\hat{x} = 0$ is unstable. This means that any deviation from an initial centre of mass at $x = 0$ leads to a steady centre of mass located in one of the two branches of the parabola in the bifurcation state.

The numerical scheme outlined in Chapter 5 captures this bifurcation diagram for the evolution of the hydrodynamic system. The results are shown in Figure 5.7. In it, Figure 5.7a depicts the bifurcation diagram of the final centre of mass when the noise parameter σ is varied, and for an initial centre of mass at $x \neq 0$. For a symmetric initial density and antisymmetric velocity, the centre of mass numerically remains at $x = 0$ for an adequate stopping criterion, since property (vi) in Subsection 5.1.1 holds. However, any slight error in the numerical computation unavoidably leads to a steady state deviating towards any of the two stable branches, due to the strong unstable nature of the branch with $x = 0$. In Figure 5.7a there is an illustration of the steady states resulting from an initial centre of mass located at $x > 0$, for different choices of the noise parameter σ. For $\sigma = 0.001$, which is the smallest value of σ simulated, the density profile approaches the theoretical Dirac delta expected at $x = 1$ when $\sigma \to 0$. When $\sigma = 0$ the hyperbolicity of the system in (4.5) is lost since the pressure term vanishes, and as a result the numerical approach in Chapter 5 cannot be applied.

The numerical strategy followed to recover the bifurcation diagram is based on the so-called differential continuation. It simply means that, as $\sigma \to 0$, the subsequent simulations with new and lower values of σ have as initial conditions the previous steady state from the last simulation. This allows to complete the bifurcation diagram, since otherwise the simulations with really small σ take

long time to converge for general initial conditions. In addition, to maintain sufficient resolution for the steady states close to the Dirac delta, the mesh is adapted for each simulation. This is accomplished by firstly interpolating the previous steady state with a piecewise cubic Hermite polynomial, which preserves the shape and avoids oscillations, and secondly by creating a new and narrower mesh where the interpolating polynomial is employed to construct the new initial condition for the differential continuation.

Example 5.3.8 (Hydrodynamic generalization of the Keller-Segel system - Generalized Euler-Poisson systems). The original Keller-Segel model has been widely employed in chemotaxis, which is usually defined as the directed movement of cells and organisms in response to chemical gradients [177]. These systems also find their applications in astrophysics and gravitation [113, 261]. It is a system of two coupled drift-diffusion differential equations for the density ρ and the chemoattractant concentration S,

$$\begin{cases} \partial_t \rho = \nabla \cdot (\nabla P(\rho) - \chi \rho \nabla S), \\ \partial_t S = D_s \Delta S - \theta S + \beta \rho. \end{cases}$$

In this system $P(\rho)$ is the pressure, and the biological/physical meaning of the constants χ, D_s, α and β can be reviewed in the literature [28, 167, 169]. For this example they are simplified as usual so that $\chi = D_s = \beta = 1$ and $\theta = 0$. A further assumption usually taken in the literature is that $\partial_t \rho$ is very big in comparison to $\partial_t S$ [167], leading to a simplification of the equation for the chemoattractant concentration S, which becomes the Poisson equation $-\Delta S = \rho$. Hydrodynamic extensions of the model, which include inertial effects, have also been proven to be essential for certain applications[82, 83, 138], leading to a hyperbolic system of equations with linear damping which in one dimension reads as

$$\begin{cases} \partial_t \rho + \partial_x (\rho u) = 0, \\ \partial_t (\rho u) + \partial_x (\rho u^2) = -\partial_x P(\rho) + \partial_x S - \gamma \rho u, \\ -\partial_{xx} S = \rho. \end{cases}$$

By using the fundamental solution of the Laplacian in one dimension, this equation becomes $2S = |x| \star \rho$. This term, after neglecting the constant, can be plugged in the momentum equation so that the last equation for S can be removed. As a result, the hydrodynamic Keller-Segel model is reduced to the system of equations (4.1) considered in this work, with $W(x) = |x|/2$, $V(x) = 0$ and $\psi \equiv 0$. As a final generalization [55], the original interaction potential $W(x) = |x|/2$ can be extended to be a homogeneous kernel $W(x) = |x|^\alpha / \alpha$, where $\alpha > -1$. By convention, $W(x) = \ln|x|$ for $\alpha = 0$. Further generalizations are Morse-like potentials as in [55, 64] where $W(x) = 1 - \exp(-|x|^\alpha / \alpha)$ with $\alpha > 0$.

The solution of this system can present a rich variety of behaviours due to the competition between the attraction from the local kernel $W(x)$ and the repulsion caused by the diffusion of the pressure $P(\rho)$, as reviewed in [49, 50]. By appropriately tuning the parameters α in the kernel $W(x)$ and m in the pressure $P(\rho)$, one can find compactly supported steady states, self-similar behaviour, or finite-time blow up. Three different regimes have been studied in the overdamped generalized Keller-Segel model [55]: diffusion dominated regime $(m > 1 - \alpha)$, balanced regime $(m = 1 - \alpha)$ where a critical mass separates self-similar and blow-up behaviour, and aggregation-dominated regime $(m < 1 - \alpha)$. These three regimes have not been so far analytically studied for the hydrodynamic system except for few particular cases [57, 58], and the presence of inertia indicates that the initial momentum profile plays a role together with the mass of the system to separate diffusive from blow-up behaviour.

The well-balanced scheme provided in Chapter 5 is a useful tool to effectively reach the varied steady states resulting from different values of α and m. The objective of this example is to provide some numerical experiments to show the richness of possible behaviours. This scheme can be eventually employed to numerically validate the theoretical studies concerning the existence of the different regimes for the hydrodynamic system for instance, or how the choice of the initial momentum or the total mass can lead to diffusive or blow-up behaviour. This will be explored further elsewhere.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy
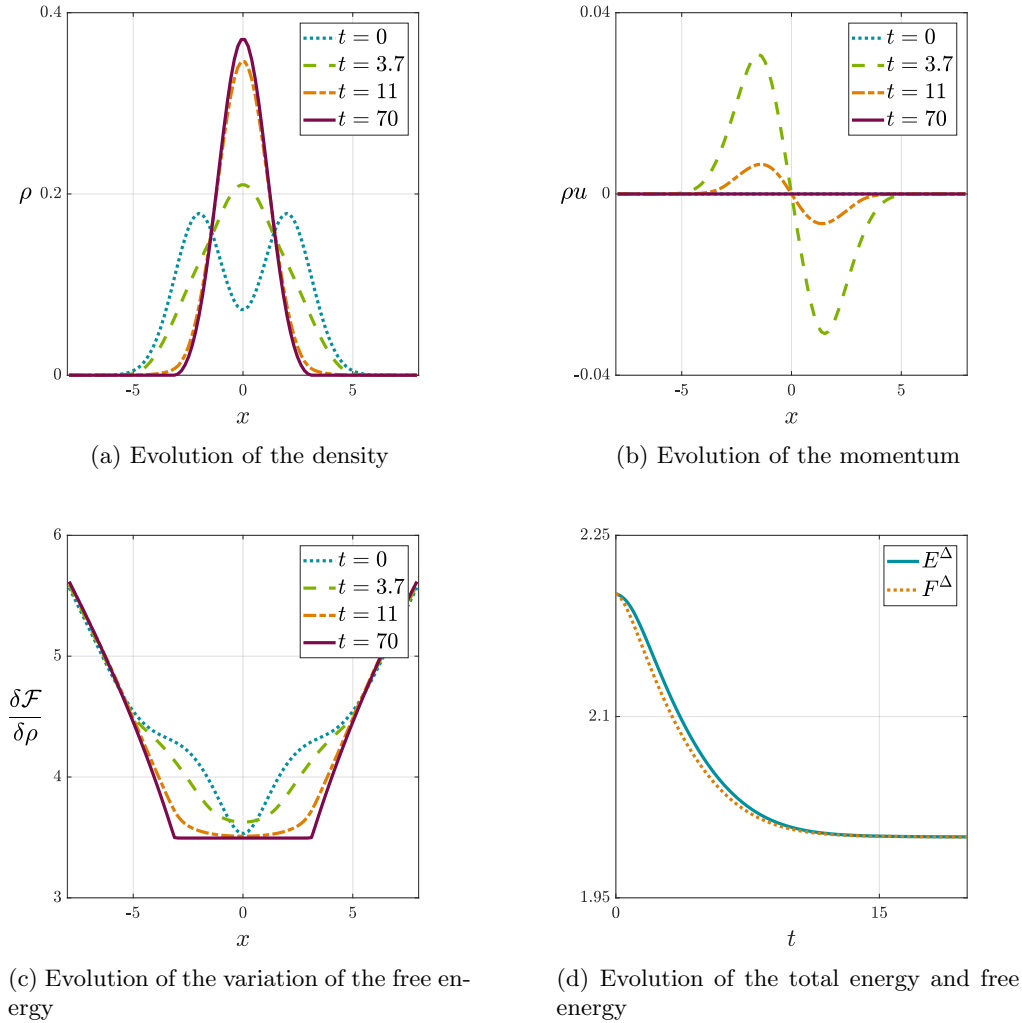
Figure 5.8:   Temporal evolution of Example 5.3.8 with compactly-supported steady state.

We have conducted two simulations with different choices of the parameters α and m. In both $m > 1$, so that a proper numerical flux able to deal with vacuum regions has to be implemented. As emphasised in the introduction of this section, the kinetic scheme developed in [235] is employed. Both of the simulations share the same initial conditions,

$$\rho(x, t=0) = M_0 \frac{e^{-\frac{4(x+2)^2}{10}} + e^{-\frac{4(x-2)^2}{10}}}{\int_{\mathbb{R}} \left(e^{-\frac{4(x-2)^2}{10}} + e^{-\frac{4(x+2)^2}{10}} \right) dx}, \quad \rho u(x, t=0) = 0, \quad x \in [-8, 8],$$

where the total mass M_0 of the system is 1.

In the first simulation the choice of parameters is $\alpha = 0.5$ and $m = 1.5$. According to the regime classification for the overdamped system, this would correspond to the diffusion-dominated regime. In the overdamped limit, solutions exist globally in time, and the steady state is compactly supported. The results are depicted in Figure 5.8 and adequately agree with this regime. In the steady state the variation of the free energy with respect to density has a constant value only in the support of the density, as expected. The total energy decreases in time and there is no exchange between the free energy and the kinetic energy since the free energy in Figure 5.8d does not oscillate.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy
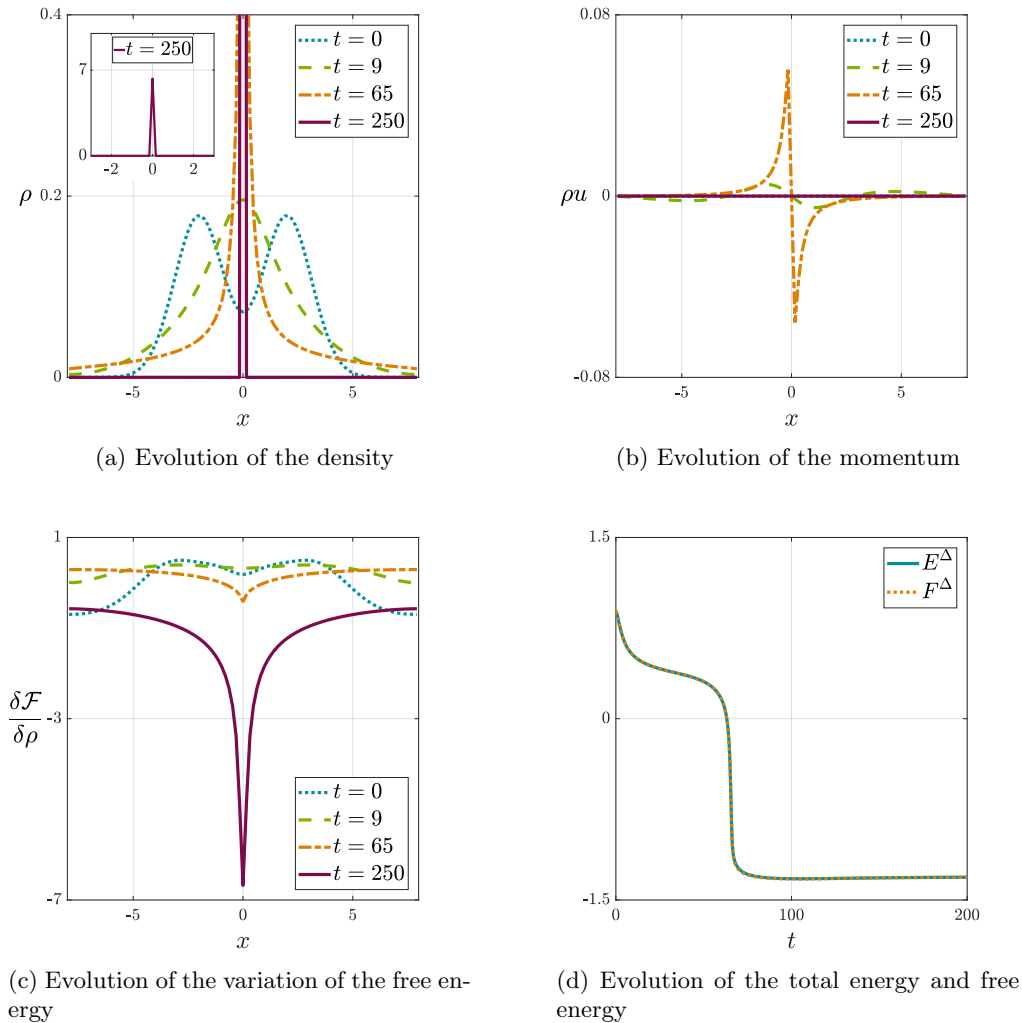
Figure 5.9: Temporal evolution of Example 5.3.8 with finite-time blow up.

The second simulation has a choice of parameters of $\alpha = -0.5$ and $m = 1.3$. In the case of the overdamped system this would correspond to the aggregation-dominated regime, where blow-up and diffusive behaviour coexist and depend on the initial density profile. The results from this simulation of the hydrodynamic system are illustrated in Figure 5.9. For this particular initial condition there is analytically finite-time blow up. Our scheme, due to the conservation of mass of the finite volume scheme, concentrates all the mass in one single cell in finite time, that is, the scheme achieves in finite time the better approximation to a Dirac Delta at a point with the

chosen mesh. Once this happens, this artificial numerical steady state depending on the mesh is kept for all times. From Figure 5.9c it is evident that the variation of the free energy with respect to density does not reach a constant value, and in Figure 5.9d the free energy presents a sharp decay when the concentration in one cell is produced (around $t \approx 65$). The value of the slope in the free energy plot theoretically tends to $-\infty$ due to the blow up, but in the simulation the decay is halted due to conservation of mass and the artificial steady state. This agrees with the fact that the expected Dirac delta profile in the density at the blow up time is obviously not reached numerically. It was also checked that this phenomena repeats for all meshes leading to more concentrated artificial steady states with more negative free energy values for more refined meshes. For other more spread initial conditions our scheme produces diffusive behaviour as expected from theoretical considerations.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy
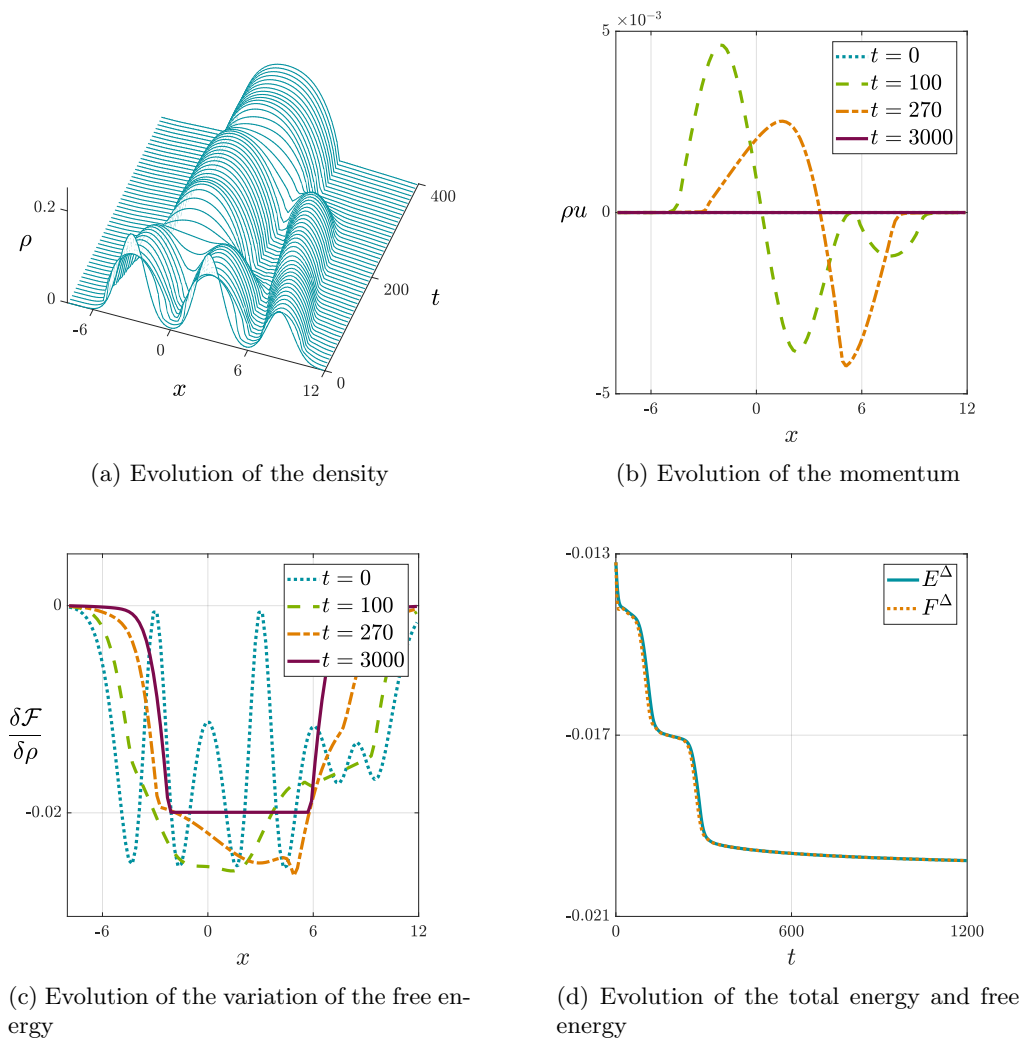
Figure 5.10: Temporal evolution of Example 5.3.8 with Morse-type potential and three initial density bumps.

A further simulation is carried out to explore the convergence in time towards equilibration with a Morse-type potential of the form $W(x) = -e^{-|x|^2/2}/\sqrt{2\pi}$. With this potential the attraction between two bumps of density separated at a considerable distance is quite small. However, when enough time has passed and the bumps get closer, they merge in an exponentially fast pace due

to the convexity of the Gaussian potential, and a new equilibrium is reached with just one bump. The interesting fact about this system is therefore the existence of two timescales: the time to get the bumps of density close enough, which could be arbitrarily slow, and the time to merge the bumps, which is exponentially fast in time.

We have set up a simulation whose initial state presents three bumps of density, with the initial conditions satisfying

$$\rho(x, t=0) = M_0 \frac{e^{-\frac{(x+3)^2}{2}} + e^{-\frac{(x-3)^2}{2}} + 0.55e^{-\frac{(x-8.5)^2}{2}}}{\int_{\mathbb{R}} \left(e^{-\frac{(x+3)^2}{2}} + e^{-\frac{(x-3)^2}{2}} + 0.55e^{-\frac{(x-8.5)^2}{2}} \right) dx}, \quad \rho u(x, t=0) = 0, \quad x \in [-8, 12],$$

and the total mass of the system equal to $M_0 = 1.2$. The parameter m in the pressure satisfies $m = 3$, and the effect of the linear damping is reduced by assigning $\gamma = 0.05$.

The results are depicted in Figure 5.10. In Figure 5.10a one can observe how the two central bumps of density merge after some time, and how the third bump, with less mass, starts getting closer in time until it also blends. This is also reflected in the evolution of the free energy in Figure 5.10d, where there are two sharp and exponential decays corresponding to the merges of the bumps.

Example 5.3.9 (DDFT for 1D hard rods). Classical (D)DFT is a theoretical framework provided by nonequilibrium statistical mechanics but has increasingly become a widely-employed method for the computational scrutiny of the microscopic structure of both uniform and non-uniform fluids [121, 147, 202, 293, 294]. The DDFT equations have the same form as in (4.5) when the hydrodynamic interactions are neglected. The starting point in (D)DFT is a functional $\mathcal{F}[\rho]$ for the fluid's free energy which encodes all microscopic information such as the ideal-gas part, short-range repulsive effects induced by molecular packing, attractive interactions and external fields. This functional can be exactly derived only for a limited number of applications, for instance the one-dimensional hard rod system from Percus [230]. However, in general it has to be approximated by making appropriate assumptions, as e.g. in the so-called fundamental-measure theory of Rosenfeld [241]. These assumptions are usually validated by carrying out appropriate test simulations (e.g. of the underlying stochastic dynamics) to compare e.g. the DDFT system with the approximate free-energy functional to the microscopic reference system [148].

The objective of this example is to show that the numerical scheme in Chapter 5 can also be applied to the physical free-energy functionals employed in (D)DFT, which satisfy the more complex expression for the free energy described in (13.7), and with a variation satisfying (13.8). For this example the focus is on the hard rods system in one dimension. Its free energy has a part depending on the local density and which satisfies the classical form for an ideal gas, with $P(\rho) = \rho$. It is therefore usually denoted as the ideal part of the free energy,

$$\mathcal{F}_{id}[\rho] = \int \Pi(\rho) dx = \int \rho(x) (\ln \rho - 1) \, dx.$$

There is also a part of general free energy in (13.7) which contains the non-local dependence of the density, and has different exact or approximative forms depending of the system under consideration. In (D)DFT it is denoted as the excessive free energy, and for the hard rods satisfies

$$\mathcal{F}_{ex}[\rho] = \frac{1}{2} \int K \left(W(\boldsymbol{x}) \star \rho(\boldsymbol{x}) \right) \rho(\boldsymbol{x}) d\boldsymbol{x}$$

$$= -\frac{1}{2} \int \rho(x + \sigma/2) \ln (1 - \eta(x)) \, dx - \frac{1}{2} \int \rho(x - \sigma/2) \ln (1 - \eta(x)) \, dx,$$

where σ is the length of a hard rod and $\eta(x)$ the local packing fraction representing the probability

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy
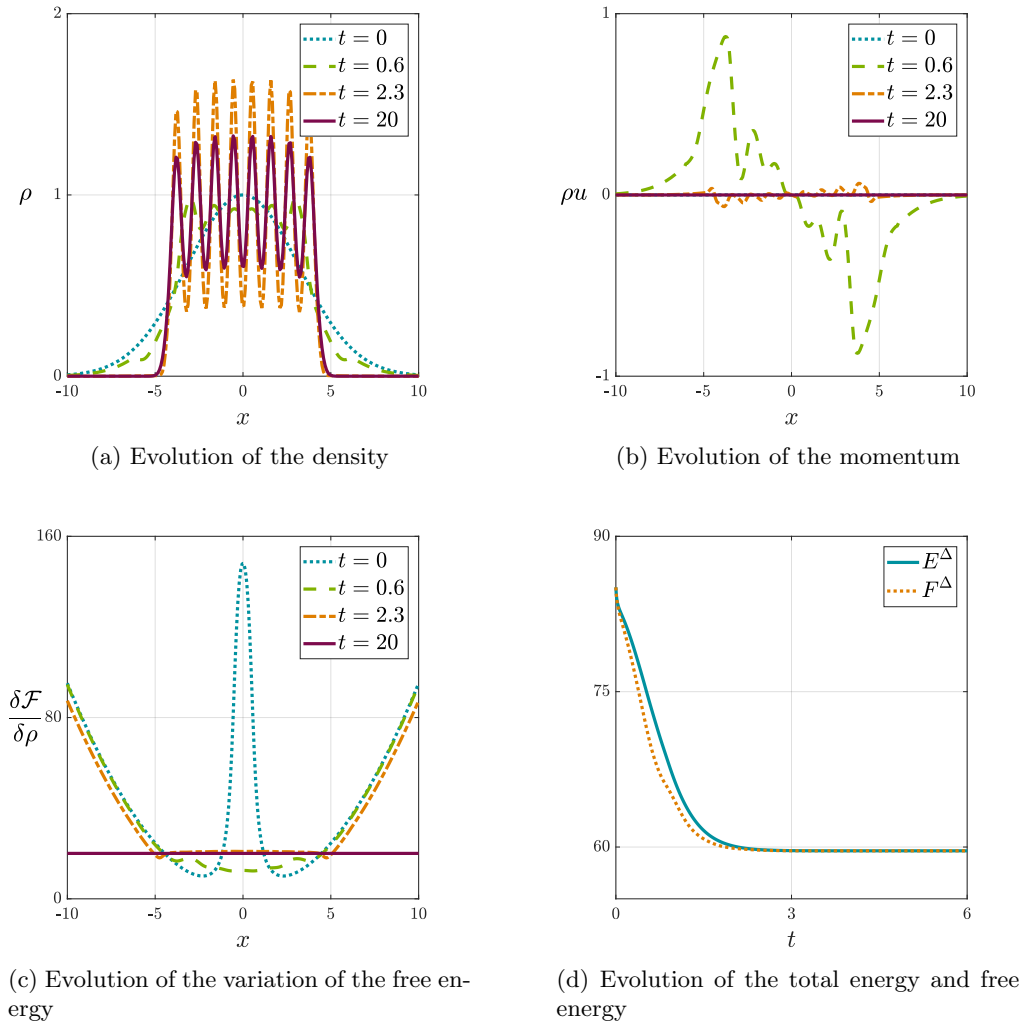
Figure 5.11:   Temporal evolution of Example 5.3.9 with 8 hard rods and a confining potential.

that a point x is covered by a hard rod,

$$\eta(x) = \int_{-\frac{\sigma}{2}}^{\frac{\sigma}{2}} \rho(x + y) dy.$$

The function $K(x)$ in this case satisfies $K(x) = \ln(1 - x)$ and the kernel $W(x)$ takes the form of a characteristic function which limits the interval of the packing function (5.3.9). To obtain the excessive free energy for the hard rods one has to also consider changes of variables in the integrals. The last part of the general free energy in (13.7) corresponds to the effect of the external potential $V(x)$. On the whole, the variation of the free energy in (13.7) with respect to the density, for the

case of hard rods, satisfies

$$\frac{\delta \mathcal{F}[\rho]}{\delta \rho} = \frac{\delta \mathcal{F}_{id}[\rho]}{\delta \rho} + \frac{\delta \mathcal{F}_{ex}[\rho]}{\delta \rho} + V(x)$$

$$= \ln(\rho) - \frac{1}{2} \ln \left(1 - \int_{x-\sigma}^{x} \rho(y) dy \right) - \frac{1}{2} \ln \left(1 - \int_{x}^{x+\sigma} \rho(y) dy \right)$$

$$+ \frac{1}{2} \int_{x-\sigma/2}^{x+\sigma/2} \left(\frac{\rho(x+\sigma/2) + \rho(x-\sigma/2)}{1 - \eta(x)} \right) dx + V(x).$$

This system can be straightforwardly simulated with the well-balanced scheme from Chapter 5 by gathering the excessive part of the free energy and the external potentials under the term $H(x, \rho)$, so that

$$H(x, \rho) = \frac{\delta \mathcal{F}_{ex}[\rho]}{\delta \rho} + V(x).$$

The first simulation seeks to capture the steady state reached by 8 hard rods of unitary mass and length $\sigma = 1$ under the presence of an external potential of the form $V(x) = x^2$. The initial conditions of the simulation are

$$\rho(x, t = 0) = e^{-\frac{x^2}{20.372}}, \quad \rho u(x, t = 0) = 0, \quad x \in [-13, 13],$$

where the density is chosen so that the total mass of the system is 8. The results are plotted in Figure 5.11. The steady state reached for the density reveals layering due to the confining effects of the external potential and the repulsion between the hard rods. These layering effects can be amplified by increasing the coefficient in the external potential. It is also observed how each of the 8 peaks has a unitary width. This is due to the fact that the length of the hard rods σ was taken as 1. The variation of the free energy with respect to the density also reaches a constant value in all the domain. For microscopic simulations of the underlying stochastic dynamics for similar examples we refer the reader to [147].

Starting from this last steady state, the second simulation performed for this example shows how the hard rods diffuse when the confining potential is removed. This simulation has as initial condition the previous steady state from Figure 5.11 and the external potential is set to $V(x) = 0$. The results are depicted in Figure 5.12, and they share the same features of the simulations in [207]. The final steady state of the density is uniform profile resultant from the diffusion of the hard rods, and in this situation the variation of the free energy with respect to the density also reaches a constant value in the steady state, as expected.

(a) Evolution of the density

(b) Evolution of the momentum
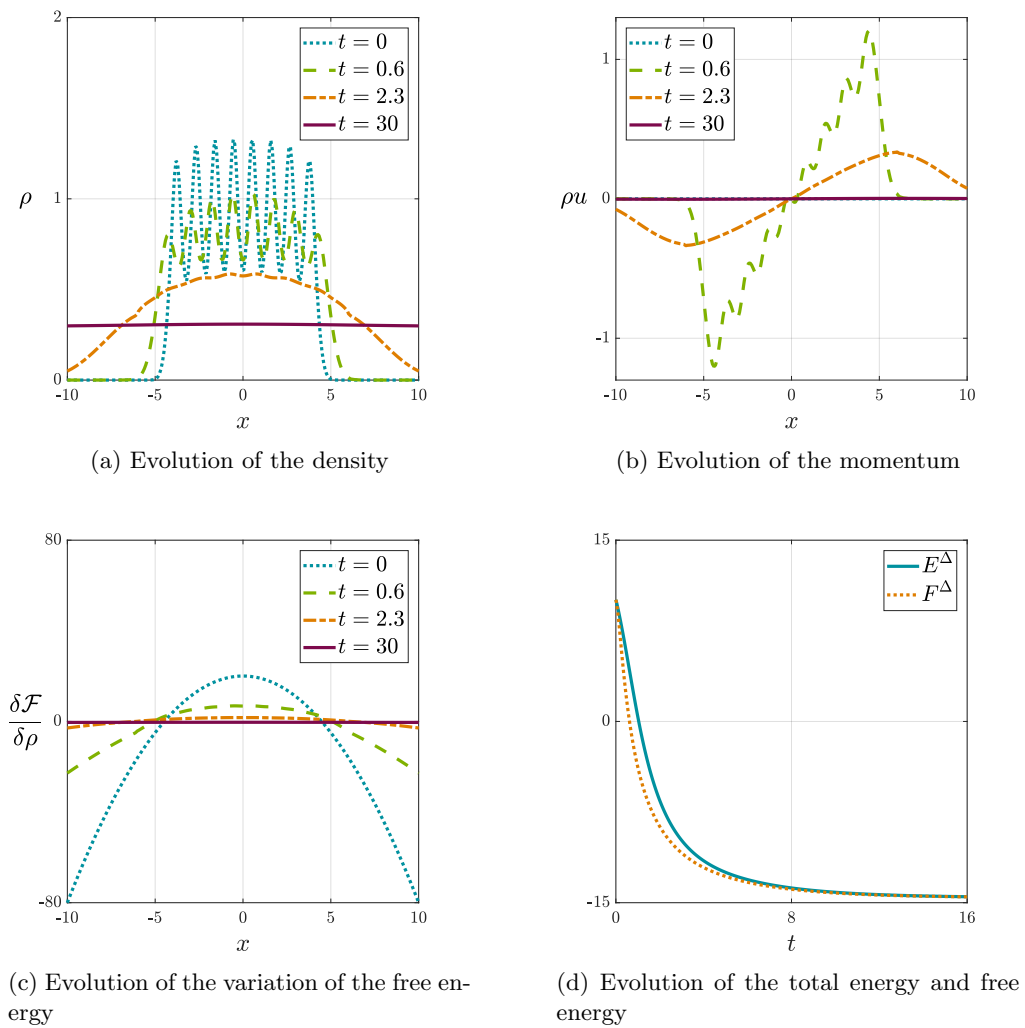

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy

Figure 5.12:   Temporal evolution of Example 5.3.9 with 8 hard rods and no potential.

Appendix

5.A Numerical flux, temporal scheme, and CFL condition

This appendix aims to present the necessary details to compute the numerical flux, boundary conditions, the CFL condition, and the temporal discretization for the simulations in Section 5.3.

The pressure function in the simulations has the form of $P(\rho) = \rho^m$, with $m \geq 1$. When $m = 0$ the pressure satisfies the ideal-gas relation $P(\rho) = \rho$, and the density does not present vacuum regions during the temporal evolution. For this case the employed numerical flux is the versatile local Lax-Friedrich flux, which approximates the flux at the boundary $F_{i+1/2}$ in (5.3) as

$$F_{i+1/2} = \mathcal{F}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) = \frac{1}{2}\left(F\left(U_{i+1/2}^{HR,-}\right) + F\left(U_{i+1/2}^{HR,+}\right) - \lambda_{i+\frac{1}{2}}\left(U_{i+1/2}^{HR,+} - U_{i+1/2}^{HR,-}\right)\right), \tag{5.36}$$

where λ is taken as the maximum of the absolute value of the eigenvalues of the system,

$$\lambda_{i+\frac{1}{2}} = \max_{\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right)} \left\{\left|u + \sqrt{P'(\rho)}\right|, \left|u - \sqrt{P'(\rho)}\right|\right\}. \tag{5.37}$$

This maximum is taken locally for every node, resulting in different values of λ along the lines of nodes. It is also possible to take the maximum globally, leading to the classical Lax-Friedrich scheme.

For the simulations where $P(\rho) = \rho^m$ and $m > 1$ vacuum regions with $\rho = 0$ are generated. This implies that the hiperbolicity of the system (4.5) is lost in those regions, and the local Lax-Friedrich scheme fails. As a result, an appropriate numerical flux has to be implemented to handle the vacuum regions. In this case a kinetic solver based on [236] is employed. This solver is constructed from kinetic formalisms applied in macroscopic models, and has already been employed in previous works for shallow-water applications [12]. The flux at the boundary $F_{i+1/2}$ in (5.3) is computed from

$$F_{i+1/2} = \mathcal{F}\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right) = A_-\left(U_{i+1/2}^{HR,-}\right) + A_+\left(U_{i+1/2}^{HR,+}\right), \tag{5.38}$$

where

$$A_-\left(\rho, \rho u\right) = \int_{\xi \geq 0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(\rho, u - \xi)\, d\xi, \quad A_+\left(\rho, \rho u\right) = \int_{\xi \leq 0} \xi \begin{pmatrix} 1 \\ \xi \end{pmatrix} M(\rho, u - \xi)\, d\xi. \tag{5.39}$$

The function $M(\rho, \xi)$ is chosen accordingly to the kinetic representation of the macroscopic system, and for this case satisfies

$$M(\rho, \xi) = \rho^{\frac{2-m}{2}} \chi\left(\frac{\xi}{\rho^{\frac{m-1}{2}}}\right). \tag{5.40}$$

The function $\chi(\omega)$ can be chosen in different ways. For this simulations we simply take it as a characteristic function,

$$\chi(\omega) = \frac{1}{\sqrt{12}} \mathbb{1}_{\{|\omega| \leq \sqrt{3}\}},$$ (5.41)

although [236] presents other possible choices for $\chi(\omega)$. Further valid numerical fluxes able to treat vacuum, such as the Rusanov flux or the Suliciu relaxation solver, are reviewed in [42].

The boundary conditions are taken to be no flux both for the density and the momentum equations. As a result, the evaluation of the numerical fluxes in (5.3) at the boundaries of the domain is taken as

$$F_{i-1/2} = 0 \;\text{ if } i = 1 \quad \text{and} \quad F_{i+1/2} = 0 \;\text{ if } i = n.$$ (5.42)

The time discretization is acomplished by means of the third order TVD Runge-Kutta method [156]. From (5.1) we can define $L(U)$ as $L(U) = S(x, U) - \partial_x F(U)$, so that $\partial_t U = L(U)$. Then, the third order TVD Runge-Kutta temporal scheme to advance from U^n to U^{n+1} with a time step Δt reads

$$U^{(1)} = U^n + \Delta t L\left(U^n\right),$$
$$U^{(2)} = \frac{3}{4}U^n + \frac{1}{4}U^{(1)} + \frac{1}{4}\Delta t L\left(U^{(1)}\right),$$
$$U^{n+1} = \frac{1}{3}U^n + \frac{2}{3}U^{(2)} + \frac{2}{3}\Delta t L\left(U^{(2)}\right).$$

The time step Δt for the case of Lax-Friedrich flux is chosen from the CFL condition,

$$\Delta t = \text{CFL} \frac{\min_i \Delta x_i}{\max_{\forall\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right)} \left\{ \left| u + \sqrt{P'(\rho)} \right|, \left| u - \sqrt{P'(\rho)} \right| \right\}},$$ (5.43)

and the Δt for the kinetic flux, with a function $\chi(\omega)$ as in (5.41), is chosen as

$$\Delta t = \text{CFL} \frac{\min_i \Delta x_i}{\max_{\forall\left(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}\right)} \left\{ |u| + 3^{\frac{m-1}{4}} \right\}}.$$ (5.44)

The CFL number is taken as 0.7 in all the simulations.

High-order well-balanced schemes

The contents of this chapter are based on the publication [54], done in collaboration with José A. Carrillo, Manuel J. Castro and Serafim Kalliadasis:

[54]   J. A. Carrillo, M. J. Castro, S. Kalliadasis, and S. P. Perez. "High-Order Well-Balanced Finite-Volume Schemes for Hydrodynamic Equations With Nonlocal Free Energy". In: *SIAM Journal on Scientific Computing* 43.2 (2021), A828-A858.

The different terms of the one-dimensional system (4.1) are usually gathered in the form of

$$\partial_t U + \partial_x F(U) = S_H(U, H) + S_D(x, U), \tag{6.1}$$

with

$$U = \begin{pmatrix} \rho \\ \rho u \end{pmatrix}, \quad F(U) = \begin{pmatrix} \rho u \\ \rho u^2 + P(\rho) \end{pmatrix}$$

and

$$S_H(U, H) = \begin{pmatrix} 0 \\ -\rho \partial_x H \end{pmatrix}, \quad S_D(x, U) = \begin{pmatrix} 0 \\ -\gamma \rho u - \rho \int_{\mathbb{R}} \psi(x - y)(u(x) - u(y))\rho(y)\, dy \end{pmatrix},$$

where U are the unknown variables, $F(U)$ the fluxes, and $S_H(U, H)$ and $S_D(x, U)$ are the sources related to forces with potential H and damping terms respectively. In what follows, we only consider the source term $S_H(U, H)$ due to the forces as we focus on the definition of a well-balanced high-order scheme for stationary solutions in (4.6). In Section 6.3 we propose a high-order discretization of the source damping term $S_D(x, U)$ that vanishes at stationary states.

We consider a mesh composed by cells $[x_{i-1/2}, x_{i+1/2}]$, $1 \leq i \leq N$, whose length Δx is supposed to be constant for simplicity. Let us denote by $U_i(t)$ the approximation of the average of the exact solution at the ith cell, $[x_{i-1/2}, x_{i+1/2}]$ at time t,

$$U_i(t) = \sum_{j=1}^{n_s} \alpha_j U(x_i^j, t) \cong \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} U(x, t)\, dx, \tag{6.2}$$

and we denote by $H_i(t)$ the approximation of the cell average of $H(x, \rho) = V(x) + W(x) \star \rho$ at time t,

$$H_i(t) = \sum_{j=1}^{n_s} \alpha_j H(x_i^j, \rho(x_i^j, t)) \cong \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} H(x, \rho(x, t)) \, dx.$$

In the previous expressions we denote as α_j and x_i^j, for $j = 1, \cdots, n_s$, the weights and quadrature points of a particular high-order quadrature formula for the cell $[x_{i-1/2}, x_{i+1/2}]$. In this work we employ the fifth-order standard Gaussian quadrature described in Appendix 6.A.

As pointed out in the introduction, one of the main contributions of this work is to construct high-order well-balanced schemes for free energies that may depend on the convolution of the density and an interaction potential, $W(x) \star \rho$. These convolutions are included in the steady state relations in (4.6), but for the discrete version of these relations one has to approximate the convolutions by a high-order quadrature formula. In the next definition we clarify the concept of well-balanced scheme for this kind of free energies.

Definition 6.0.1 (Well-balanced scheme). We consider a semi-discrete method to approximate (6.1),

$$\begin{cases} \dfrac{dU_i}{dt} = & -\dfrac{1}{\Delta x} \mathcal{L}(\Delta x, U_j(t), H_j(t), j \in \mathcal{S}_i), \\ U(0) = & U_0, \end{cases} \tag{6.3}$$

where $U(t) = \{U_i(t)\}_{i=1}^N$ represents the vector of the approximations of the averaged values of the exact solutions at time t, $U_0 = \{U_i(0)\}$ is the vector of the initial conditions, and $\mathcal{S}_i$ the stencil of the numerical scheme.

Now let us assume that $u(x) = 0$, $\rho(x)$ is a smooth function and $H_{\Delta x}(x)$ is a discrete approximation of $H = V + W \star \rho$ with the form

$$H_{\Delta x}(x) = V(x) + \Delta x \sum_{l=1}^{M} \sum_{m=1}^{n_s} \alpha_m W(x - x_l^m) \rho(x_l^m)$$

and satisfying

$$\Pi'(\rho(x)) + H_{\Delta x}(x) = C_\Gamma \text{ in each } \Lambda_\Gamma, \Gamma \in \mathbb{N}, \tag{6.4}$$

where Λ_Γ, $\Gamma \in \mathbb{N}$, denotes the possible infinite sequence indexed by Γ of subsets Λ_Γ of subsequent indices $i \in \{1, \ldots, N\}$ where $\rho(x) > 0$ and $u = 0$, and C_Γ the corresponding constant in that connected component of the discrete support.

Then it follows that the semi-discrete numerical scheme (6.3) is said to be well-balanced for

$$U = \begin{pmatrix} \rho(x) \\ 0 \end{pmatrix} \quad \text{and} \quad H_{\Delta x}$$

if the vector of their approximated averages is a critical point of (6.3), i.e.

$$\mathcal{L}(\Delta x, U_j, H_j, j \in \mathcal{S}_i) = 0, \ 1 \le i \le N,$$

where

$$U_i = \sum_{j=1}^{n_s} \alpha_j U(x_i^j) \quad \text{and} \quad H_i = \sum_{j=1}^{n_s} \alpha_j H_{\Delta x}(x_i^j).$$

In what follows we begin by briefly recalling in Section 6.1 the first-order well-balanced scheme for (6.1) introduced in [71], which serves as an starting point to construct high-order schemes by employing a high-order reconstruction operator, as described in Section 6.2. Then in Section 6.3 we describe how to adapt these high-order schemes so that they are well-balanced in the sense defined in Definition 6.0.1.

6.1 Summary of the first-order well-balanced scheme

The first-order semi-discrete well-balanced finite-volume scheme for system (6.1) introduced in [71] can be written as

$$\frac{dU_i}{dt} = -\frac{1}{\Delta x}\left(\mathbb{F}_{i+1/2}^- - \mathbb{F}_{i-1/2}^+\right), \tag{6.5}$$

where $\mathbb{F}_{i+1/2}^{\pm}$ is defined using a standard consistent numerical flux for the homogeneous system applied to the so-called hydrostatic-reconstructed states, with an extra term ensuring the consistency of the numerical scheme (6.5) applied to system (6.1), as well as its well-balanced character. In [71] the midpoint quadrature formula is used to approximate both the cell-averages of the exact solution and $H_{\Delta x}(x)$. In what follows we suppress the time-dependence in the cell averages for simplicity. We define $\mathbb{F}_{i+1/2}^{\pm}$ in terms of the cell averages U_i, U_{i+1}, H_i and H_{i+1} as

$$\mathbb{F}_{i+1/2}^{\pm}(U_i, U_{i+1}, H_i, H_{i+1}) = \mathcal{F}(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}) \pm S_{i+1/2}^{HR,\pm}, \tag{6.6}$$

where $\mathcal{F}(U,V)$ is the standard local Lax-Friedrich numerical flux for the homogeneous system,

$$\mathcal{F}(U,V) = \frac{1}{2}\left(F(U) + F(V) - |\lambda(U,V)|(V-U)\right), \tag{6.7}$$

where $|\lambda(U,V)|$ is a bound for the maximum absolute value of the wave speeds for the Riemann problem with the constant states U and V.

In [71] we employ the hydrostatic reconstruction firstly introduced in [11] in the context of shallow-water equations. Here we denote the hydrostatic-reconstructed states as $U_{i+1/2}^{HR,\pm}$, and we compute them as follows:

1) Firstly an intermediate state $H_{i+1/2}$ is computed as

$$H_{i+1/2} = \max(H_{i+1}, H_i).$$

2) Next, we define the hydrostatic-reconstructed states as

$$U_{i+1/2}^{HR,\pm} = \begin{pmatrix} \rho_{i+1/2}^{HR,\pm} \\ (\rho u)_{i+1/2}^{HR,\pm} \end{pmatrix},$$

where

$$\begin{aligned}
\rho_{i+1/2}^{HR,-} &= \xi\left(\Pi'(\rho_i) + H_i - H_{i+1/2}\right), \quad (\rho u)_{i+1/2}^{HR,-} = \rho_{i+1/2}^{HR,-} u_i, \\
\rho_{i+1/2}^{HR,+} &= \xi\left(\Pi'(\rho_{i+1}) + H_{i+1} - H_{i+1/2}\right), \quad (\rho u)_{i+1/2}^{HR,+} = \rho_{i+1/2}^{HR,+} u_{i+1},
\end{aligned} \tag{6.8}$$

with $\xi(s)$ being the inverse function of $\Pi'(s)$ for $s > 0$ and $u_i = (\rho u)_i / \rho_i$.

The last ingredients for the flux in (6.6) are the terms $S_{i+1/2}^{HR,\pm}$, which correspond to the correction introduced in the numerical scheme to guarantee consistency and well-balanced properties (see [11, 71]),

$$S_{i+1/2}^{HR,+} = \begin{pmatrix} 0 \\ P(\rho_{i+1}) - P\left(\rho_{i+1/2}^{HR,+}\right) \end{pmatrix} \quad \text{and} \quad S_{i+1/2}^{HR,-} = \begin{pmatrix} 0 \\ P\left(\rho_{i+1/2}^{HR,-}\right) - P(\rho_i) \end{pmatrix}. \tag{6.9}$$

It is straightforward to check that the semi-discrete numerical scheme (6.5)-(6.9) is well-balanced in the sense defined in Definition 6.0.1 (see [71]). It may also be surprising that the potential cell averages H_i do not appear in the source terms (6.9), but this is a consequence of the well-balanced methodology and we refer the reader to [11, 42, 71] for further details.

6.2 Ingredients to construct a high-order finite-volume scheme

The basic ingredients to design a high-order finite volume method for system (6.1), assuming $S_D(x, U) = 0$, are:

- a consistent first order numerical flux for system (6.1), like the one proposed in [71] and described in the previous Subsection;

- a high-order reconstruction operator, i.e. an operator that, given a family of cell values $\{U_i(t)\}$, provides at every cell $[x_{i-1/2}, x_{i+1/2}]$ a smooth function that depends on the values at some neighbour cells whose indices belong to the so-called stencil $\mathcal{S}_i$:

$$R_i^U(x) = R_i^U(x; \{U_j(t)\}_{j \in \mathcal{S}_i}),$$

so that $R_i^U(x)$ is a high-order approximation of $U(x, t)$ in the ith cell at time t. Here we use third- and fifth-order Compact Weighted Essentially Non-Oscillatory (CWENO) reconstruction operators [52, 197, 198]. The main advantage of CWENO compared to WENO (see [253, 256, 257]) reconstruction operators is that CWENO reconstructions achieve uniform high-order approximation in the entire cell, while WENO reconstruction operators are proposed to achieve high-order approximation at the boundaries of the cell. Thus, standard WENO-5 reconstructions achieves 5th-order at the boundaries of the cell, while it is only 3rd-order at the interior points. Therefore, CWENO reconstruction operators are specially useful in balance laws such as (6.1), where the source term has to be evaluated at inner points of the cell. We complement the CWENO reconstruction operators with the positive-density limiters from [299] to ensure physical admissible reconstructed values for the density. For further details we refer the reader to Appendix 6.B.

Using these ingredients, one could consider a high-order finite-volume semi-discrete numerical method of the form:

$$\frac{dU_i}{dt} = -\frac{1}{\Delta x}\left(\mathbb{F}_{i+1/2}^- - \mathbb{F}_{i-1/2}^+\right) + \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} S_H(R_i^U(x), R_i^H(x))\, dx, \qquad (6.10)$$

where

- $R_i^U(x)$ and $R_i^H(x)$ are the approximations of the solution $U(x, t)$ and the function $H_{\Delta x}(x)$, respectively, at the ith cell given by some high-order reconstruction operators from the sequence of cell values $\{U_i(t)\}$ and $\{H_i(t)\}$, respectively, i.e.

$$R_i^U(x) = R_i(x; \{U_j(t)\}_{j \in \mathcal{S}_i}) \quad \text{and} \quad R_i^H(x) = R_i(x; \{H_j(t)\}_{j \in \mathcal{S}_i});$$

- $\mathbb{F}_{i+1/2}^{\pm}$ is the numerical flux defined in (6.6) applied to the reconstructed states $U_{i+1/2}^{\mp}$ and $H_{i+1/2}^{\mp}$, i.e.

$$\mathbb{F}_{i+1/2}^{\pm} = \mathbb{F}(U_{i+1/2}^-, U_{i+1/2}^+, H_{i+1/2}^-, H_{i+1/2}^+)$$

with

$$U_{i+1/2}^- = R_i^U(x_{i+1/2}), \quad U_{i+1/2}^+ = R_{i+1}^U(x_{i+1/2}),$$

and

$$H_{i+1/2}^- = R_i^H(x_{i+1/2}), \quad H_{i+1/2}^+ = R_{i+1}^H(x_{i+1/2}).$$

Following the procedure described in [74, Theorem 3.2] and [75, Section 4], it is possible to prove that the semi-discrete numerical scheme (6.10) is a high-order numerical scheme of order $p > 1$, if the following three conditions are satisfied:

(i) $H_{\Delta x}(x)$ is a high-order approximation of $H(x, \rho)$ of order $p > 1$;

(ii) $R_i^U(x)$ and $R_i^H(x)$ are high-order reconstruction operators of order at least $p > 1$;

(iii) the volume integral

$$\frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} S_H(R_i^U(x), R_i^H(x)) \, dx$$

is computed exactly or approximated with a quadrature formula of order greater or equal to $p > 1$.

Unfortunately, when employing standard (CWENO, WENO, ...) reconstruction the resulting numerical scheme is, in general, not well-balanced. Indeed, if $\{U_i\}_{i=1}^N$ and $\{H_i\}_{i=1}^N$ are the cell averages of a discrete steady state satisfying (6.4), then their reconstructions do not necessary satisfy the discrete relations

$$\Pi'\left(R_i^\rho(x)\right) + R_i^H(x) = C, \; 1 \leq i \leq N. \tag{6.11}$$

In the previous expression we suppose, for simplicity, that we have only one connected component.

In what follows we aim to propose a modified reconstruction procedure which respects (6.11) for any discrete steady state satisfying (6.4). As we show in the next Subsection, thanks to this modification we can prove that our scheme is both high-order accurate and well-balanced.

6.3   High-order well-balanced numerical schemes

Let us suppose that the sequences of cell averages $\{U_i\}_{i=1}^N$ and $\{K_i\}_{i=1}^N$ are known, with

$$U_i = \begin{pmatrix} \rho_i = \displaystyle\sum_{j=1}^{n_s} \alpha_j \rho(x_i^j) \\ (\rho u)_i = \displaystyle\sum_{j=1}^{n_s} \alpha_j (\rho u)(x_i^j) \end{pmatrix}$$

and

$$K_i = \sum_{j=1}^{n_s} \alpha_j \left[\Pi'(\rho(x_i^j)) + H_{\Delta x}(x_i^j) \right]. \tag{6.12}$$

For such cell averages we propose the following reconstruction procedure:

- We consider a standard high-order reconstruction operator for the conserved variables ρ and ρu, and also applied to the sequence $\{K_i\}_{i=1}^N$,

$$\begin{aligned} R_i^\rho(x) &= R_i \left(x, \{\rho_j\}_{j \in S_i}\right), \\ R_i^{\rho u}(x) &= R_i \left(x, \{(\rho u)_j\}_{j \in S_i}\right), \\ R_i^K(x) &= R_i \left(x, \{K\}_{j \in S_i}\right); \end{aligned} \tag{6.13}$$

- the reconstruction operator for $H_{\Delta x}(x)$ is defined as

$$R_i^H(x) = R_i^K(x) - \Pi'\left(R_i^\rho(x)\right), \tag{6.14}$$

with R_i^H not being a polynomial since it depends on the function $\Pi'(\rho)$.

The previous reconstruction procedure satisfies the following property:

Theorem 6.3.1. *Let $u = 0$, $\rho(x)$ and $H_{\Delta x}(x)$ satisfying (6.4), that is ρ, $u = 0$ is a discrete stationary solution of system (6.1), then the reconstructions $R_i^\rho(x)$, $R_i^{\rho u}(x)$ and $R_i^H(x)$ are discrete stationary solutions of system (6.1) at $[x_{i-1/2}, x_{i+1/2}]$.*

Proof. Let us suppose for simplicity that the stationary solution is only defined in one connected component. Therefore, $K_i = C$, $1 \le i \le N$.

As standard reconstruction operators like CWENO are exact for constant functions, we have that $R_i^{\rho u}(x) = 0$ and $R_i^K(x) = C$. Therefore in (6.11) we have that $C = \Pi'(R_i^\rho(x)) + R_i^H(x)$, which proves the result setting $H_{\Delta x}(x) = R_i^H(x)$. $\qquad\square$

It is important to remark that, even if the reconstruction procedure satisfies the discrete steady state of system in (6.11), the semi-discrete numerical scheme (6.10) may not be in general well-balanced. This is because the integral

$$\int_{x_{i-1/2}}^{x_{i+1/2}} S_H(R_i^U(x), R_i^H(x)) \, dx$$

has to be numerically approximated, and if such integration is not exact then the well-balancing property may be destroyed (see [77]). To overcome this difficulty, we follow the strategy proposed in [77]: a local discrete stationary solution is added to the numerical scheme for every cell. We denote this solution by $U_i^*(x) = (\rho_i^*(x), 0)^T$ and $H_i^*(x)$, and it satisfies

$$\frac{1}{\Delta x} \left(F(U_i^*(x_{i+1/2})) - F(U_i^*(x_{i-1/2})) \right) = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} S_H(U_i^*(x), H_i^*(x)) \, dx. \qquad (6.15)$$

The previous steady state relation in (6.15) is satisfied if we choose $U_i^*(x) = (\rho_i^*(x), 0)^T$ and $H_i^*(x)$ as

$$U_i^*(x) = \left(\begin{array}{c} \rho_i^*(x) = R_i^\rho(x) \\ 0 \end{array} \right), \quad H_i^*(x) = K_i - \Pi'(R_i^\rho(x)). \qquad (6.16)$$

Observe that the convolution is indirectly approximated in the previous expression.

Now, we could rewrite the semi-discrete numerical scheme (6.10) by just adding the steady state expression in (6.15), yielding

$$
\begin{aligned}
\frac{dU_i}{dt} = \ & -\frac{1}{\Delta x} \left(\mathbb{F}_{i+1/2}^- - \mathbb{F}_{i-1/2}^+ - F(U_i^*(x_{i+1/2})) + F(U_i^*(x_{i-1/2})) \right) \\
& + \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} S_H(R_i^U(x), R_i^H(x)) - S_H(U_i^*(x), H_i^*(x)) \, dx.
\end{aligned}
\qquad (6.17)
$$

The advantage of this new version of the scheme relies in the fact that the integral term in (6.17) could be approximated by any high-order quadrature formula, without perturbing the well-balanced character of the numerical scheme. This comes from the fact that, for any discrete stationary solution satisfying (6.4), we have that $R_i^H(x) = H_i^*(x)$ and $R_i^\rho(x) = \rho_i^*(x)$, so that

$$S_H(R_i^U(x), R_i^H(x)) - S_H(U_i^*(x), H_i^*(x)) = 0.$$

For such integral here we follow [217], where an n-th order Richardson extrapolation formula is proposed to evaluate source terms of the form $\rho \partial_x H$. We detail the fourth- and sixth-order formulas in the Appendix 6.C. We finally conclude with the following result.

Theorem 6.3.2. *The numerical scheme (6.17) with the reconstruction operators (6.13) and (6.14) is well-balanced in the sense of Definition 6.0.1.*

Proof. Let us suppose that $u = 0$, $\rho(x)$, $H_{\Delta x}(x)$ satisfy (6.4). We also assume for simplicity that the stationary solution is defined on one connected component. Then, as proved in Theorem (6.3.1), the reconstructions $R_i^\rho(x)$, $R_i^{\rho u}(x)$ and $R_i^H(x)$ satisfy (6.11). Moreover,

$$\frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} S_H(R_i^U(x), R_i^H(x)) - S_H(U_i^*(x), H_i^*(x)) \, dx = 0,$$

and the reconstructed states at the intercells verify

$$\Pi'(\rho_{i+1/2}^+) + H_{i+1/2}^+ = \Pi'(\rho_{i+1/2}^-) + H_{i+1/2}^- = C, \quad u_{i+1/2}^- = u_{i+1/2}^+ = 0. \tag{6.18}$$

The relation (6.18) implies that the hydrostatic-reconstructed states satisfy

$$\rho_{i+1/2}^{HR,-} = \rho_{i+1/2}^{HR,+}, \quad (\rho u)_{i+1/2}^{HR,-} = (\rho u)_{i+1/2}^{HR,+} = 0. \tag{6.19}$$

Using (6.18) and (6.19) and the definition of (6.6) and (6.9), $\mathbb{F}_{i+1/2}^-$ reduces to

$$\mathbb{F}_{i+1/2}^- = \begin{pmatrix} 0 \\ P(\rho_{i+1/2}^-) \end{pmatrix}.$$

Analogously, we deduce that

$$\mathbb{F}_{i-1/2}^+ = \begin{pmatrix} 0 \\ P(\rho_{i-1/2}^+) \end{pmatrix}.$$

Now, taking into account the definition of $U_i^*(x)$ given in (6.16),

$$U_i^*(x_{i+1/2}) = \begin{pmatrix} \rho_{i+1/2}^- \\ 0 \end{pmatrix}, \quad U_i^*(x_{i-1/2}) = \begin{pmatrix} \rho_{i-1/2}^+ \\ 0 \end{pmatrix},$$

we finally conclude by noting that the numerical scheme (6.17) reduces to

$$\frac{dU_i}{dt} = \frac{1}{\Delta x} \left(\mathbb{F}_{i+1/2}^- - \mathbb{F}_{i-1/2}^+ - F(U_i^*(x_{i+1/2})) + F(U_i^*(x_{i-1/2})) \right) = 0.$$

$\square$

Remark 6.3.3. The well-balanced reconstruction operators defined in (6.13) and (6.14) employ, as expected, the approximated cell averages of the solution at each time step and an extra quantity corresponding to the cell average of the variation of the free-energy, denoted by K_i. This quantity plays an important role to achieve the well-balanced property of the reconstruction operators and the final numerical scheme. Note that the semi-discrete numerical scheme (6.17) only allows to evolve in time the cell averages of conserved variables, and as a result we should provide an extra equation to evolve the variation of the free-energy. We propose the following: suppose that $\{\rho_i^n\}$ $\{(\rho u)_i^n\}$ and $\{K_i^n\}$ are known at time $t = n\Delta t$, and suppose in addition that we use the standard explicit first order Euler scheme to evolve the conserved variables up to time $t = (n+1)\Delta t$. Then we propose to update K_i^{n+1} as

$$K_i^{n+1} = K_i^n + \sum_{j=1}^{n_s} \alpha_j \left[\Pi' \left(R_i^{\rho^{n+1}}(x_i^j) \right) - \Pi' \left(R_i^{\rho^n}(x_i^j) \right) \right]$$

$$+ \sum_{j=1}^{n_s} \alpha_j \left[\sum_{l=1}^{M} \sum_{m=1}^{n_s} \Delta x \, \alpha_m W \left(x_i^j - x_l^m \right) \left[R_l^{\rho^{n+1}}(x_l^m) - R_l^{\rho^n}(x_l^m) \right] \right], \tag{6.20}$$

where we use some high-order quadrature formula for the cell averages and convolution operator. Here we will use the fifth-order Gaussian quadrature described in Appendix 6.A. A similar procedure could be applied if a high-order RK-TVD scheme (see [156]) is used instead of the explicit Euler scheme to discretize the ODE system (6.17). This is based on the classical observation that these schemes can be written as linear combinations of explicit Euler steps. Observe also that $K_i^{n+1} = K_i^n$ on discrete stationary solutions satisfying (6.4).

Finally, if the term $S_D(x, U)$ is now added to the system, our full high-order semi-discrete well-balanced finite volume scheme can be written as

$$
\begin{aligned}
\frac{dU_i}{dt} = &-\frac{1}{\Delta x}\left(\mathbb{F}^-_{i+1/2} - \mathbb{F}^+_{i-1/2} - F(U_i^*(x_{i+1/2})) + F(U_i^*(x_{i-1/2}))\right) \\
&+ \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}} S_H(R_i^U(x), R_i^H(x)) - S_H(U_i^*(x), H_i^*(x))\, dx \\
&- \gamma(\rho u)_i + \frac{1}{\Delta x}\sum_{j=1}^{n_s}\alpha_j R_i^{\rho,t}(x_i^j)\left[\sum_{l=1}^{M}\sum_{m=1}^{n_s}\Delta x\, \alpha_m \psi(x_i^j - x_l^m)\left(R_i^{u,t}(x_i^j) - R_l^{u,t}(x_l^m)\right)R_l^{\rho,t}(x_l^m)\right],
\end{aligned}
\tag{6.21}
$$

where $R_i^{u,t} = \dfrac{R_i^{\rho u,t}}{R_i^{\rho,t}}$. Note that the new terms do not affect to the well-balance property of the scheme as they vanishes when $u = 0$.

Remark 6.3.4. In practical applications is quite important to guarantee that the numerical scheme preserves the non-negativity of the density $\rho_i(t)$. The high-order numerical scheme (6.21) preserves the non-negativity of the density as consequence of:

1. the first order numerical flux preserves the non-negativity of the density (see [71]);

2. the application of the positive-density limiter introduced in [299] and described in Subappendix 6.B.3;

3. a suitable CFL restriction (see (6.43) and [299]), described also in Subappendix 6.B.3.

6.4   Algorithmic implementation

In this section we summarize the steps to efficiently implement the high-order well-balanced finite-volume scheme of Section 6.3.

The initial conditions for system (6.1) are the initial density profile $\rho_0(x)$ and momentum profile $(\rho u)_0(x)$. These initial conditions are introduced in the numerical scheme by computing their cell averages via high-order quadrature formula,

$$
\rho_i^0 = \sum_{j=1}^{n_s}\alpha_j\rho_0(x_i^j),
$$

$$
(\rho u)_i^0 = \sum_{j=1}^{n_s}\alpha_j(\rho u)_0(x_i^j),
$$

where the coefficients α_j denote the weights of the quadrature formula that multiply the evaluation of $\rho_0(x)$ and $(\rho u)_0(x)$ at the quadrature points x_i^j, and n_s denotes the number of quadrature points. Here we employ the fifth-order Gaussian quadrature formula described in the Appendix 6.A.

The initial cell averages of the derivative of the free energy (6.4) are also required, and are similarly computed via fifth-order Gaussian quadrature

$$
K_i^0 = \sum_{j=1}^{n_s}\alpha_j\left[\Pi'\left(\rho_0(x_i^j)\right) + V(x_i^j) + \sum_{l=1}^{n}\sum_{m=1}^{n_s}\Delta x\alpha_m W\left(x_i^j - x_l^m\right)\rho_0(x_l^m)\right].
$$

The computation of the cell averages using quadrature formulas is only necessary at the initial time step of the algorithm. For further time steps, the algorithm presented here takes as inputs the cell averages ρ_i^n, $(\rho u)_i^n$ and K_i^n, evaluated at $t = n\Delta t$, and directly returns the cell averages ρ_i^{n+1}, $(\rho u)_i^{n+1}$ and K_i^{n+1} at the subsequent time step $t = (n+1)\Delta t$. The steps for such algorithm are:

1) Perform high-order reconstructions $R_i^\rho(x)$, $R_i^{\rho u}(x)$ and $R_i^K(x)$ from the sequences of cell values $\{\rho_i^n\}$, $\{(\rho u)_i^n\}$, $\{K_i^n\}$, following (6.13). In our case, such reconstructions are conducted via third- and fifth-order CWENO reconstructions [52, 197, 198], details provided in Appendix 6.B.

   For simplicity, the evaluations of the previous reconstructions at the intercells at time $t = n\Delta t$ are denoted as

   $$\rho_{i-1/2}^+ = R_i^\rho(x_{i-1/2}), \quad \rho_{i+1/2}^- = R_i^\rho(x_{i+1/2}),$$

   $$(\rho u)_{i-1/2}^+ = R_i^{\rho u}(x_{i-1/2}), \quad (\rho u)_{i+1/2}^- = R_i^{\rho u}(x_{i+1/2}), \tag{6.22}$$

   $$K_{i-1/2}^+ = R_i^K(x_{i-1/2}), \quad K_{i+1/2}^- = R_i^K(x_{i+1/2}).$$

   Furthermore, when required, the reconstruction of the velocity field u is computed as

   $$R_i^u(x) = \frac{R_i^{\rho u}(x)}{R_i^\rho(x)}, \quad u_{i-1/2}^+ = R_i^u(x_{i-1/2}), \quad u_{i+1/2}^- = R_i^u(x_{i+1/2}).$$

2) Obtain the reconstruction $R_i^H(x)$ for $H_{\Delta x}(x)$ from (6.14), and their evaluations at the intercells as

   $$H_{i-1/2}^+ = R_i^H(x_{i-1/2}), \quad H_{i+1/2}^- = R_i^H(x_{i+1/2}).$$

   In general $H_{i+1/2}^+ \neq H_{i+1/2}^-$. The average value between them is taken as

   $$H_{i+1/2} = \max\left(H_{i+1/2}^+, H_{i+1/2}^-\right).$$

3) Reconstruct the local discrete stationary solution in (6.15) for every cell, so that $U_i^*(x)$ and $H_i^*(x)$ are computed from (6.16).

4) Perform the so-called hydrostatic reconstruction described in (6.8), but now with the high-order reconstructions at the intercells. By denoting as $\xi(s)$ the inverse function of $\Pi'(s)$ for $s > 0$,

   $$\rho_{i+1/2}^{HR,-} = \xi\left(\Pi'\left(\rho_{i+1/2}^-\right) + H_{i+1/2}^- - H_{i+1/2}\right), \quad (\rho u)_{i+1/2}^{HR,-} = \rho_{i+1/2}^{HR,-} u_{i+1/2}^-,$$

   $$\rho_{i+1/2}^{HR,+} = \xi\left(\Pi'\left(\rho_{i+1/2}^+\right) + H_{i+1/2}^+ - H_{i+1/2}\right), \quad (\rho u)_{i+1/2}^{HR,+} = \rho_{i+1/2}^{HR,+} u_{i+1/2}^+.$$

5) The cell averages ρ_i^{n+1} and $(\rho u)_i^{n+1}$ at the subsequent time step $t = (n+1)\Delta t$ are updated by means of (6.21), where

   $$\mathbb{F}_{i+1/2}^\pm(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}, H_{i+1/2}^-, H_{i+1/2}^+) = \mathcal{F}(U_{i+1/2}^{HR,-}, U_{i+1/2}^{HR,+}) \pm S_{i+1/2}^{HR,\pm}, \tag{6.23}$$

   with the Lax-Friedrich flux in (6.7),

   $$S_{i+1/2}^{HR,+} = \begin{pmatrix} 0 \\ P\left(\rho_{i+1/2}^+\right) - P\left(\rho_{i+1/2}^{HR,+}\right) \end{pmatrix}, \quad S_{i+1/2}^{HR,-} = \begin{pmatrix} 0 \\ P\left(\rho_{i+1/2}^{HR,-}\right) - P\left(\rho_{i+1/2}^-\right) \end{pmatrix}, \tag{6.24}$$

$$F\left(U_i^*(x_{i-1/2})\right) = \begin{pmatrix} 0 \\ P\left(\rho_{i-1/2}^+\right) \end{pmatrix}, \quad F\left(U_i^*(x_{i+1/2})\right) = \begin{pmatrix} 0 \\ P\left(\rho_{i+1/2}^-\right) \end{pmatrix}, \tag{6.25}$$

and the integral for the high-order corrections in the source term is computed from the fourth- and sixth-order formulas in the Appendix 6.C.

6) Finally, update the value K_i^{n+1} by means of fluctuations from (6.20).

6.5 Numerical simulations

Here we employ the high-order finite volume scheme in a variety of relevant applications, taken from the fields of gas dynamics, porous media, collective behaviour and chemotaxis. First, in Subsection 6.5.1 we conduct the validation of the properties from the numerical scheme to ensure both that the high-order and the well-balanced properties are numerically satisfied. Second, in Subsection 6.5.2 we proceed to apply the scheme to challenging scenarios where analytical results are scarce.

The numerical flux for the simulations is chosen depending on the form of the pressure term, which satisfies $P(\rho) = \rho^m$ with $m \geq 1$. A local Lax-Friedrich numerical flux is employed for the examples with ideal-gas pressure, where $m = 1$ and the support of the density is not compact. On the contrary, a kinetic scheme based on [236] is employed for pressures with $m > 1$, due to the presence of vacuum regions and compactly-supported densities. For the details of these two numerical fluxes we refer the reader to our previous work [71].

For the temporal integration we implement the third-order TVD Runge-Kutta method [156], with the CFL number chosen as 0.7 in all the simulations. The CFL conditions for these two numerical fluxes are detailed in equation (6.43) of Appendix 6.B. The boundary conditions are periodic unless otherwise specified. In all the simulation we set $\gamma = 1$ in the linear damping while the nonlinear damping is in general deactivated, except for Example 6.5.4. The number of cells employed to create the plots is 200. For the figures we use the third-order time discretization scheme, unless otherwise stated.

In the following numerical simulations we focus on the temporal evolution of the density, momentum and free-energy variation in (6.12). For illustrative purposes, we also plot the evolution of the discrete versions of the total energy in (4.7) and free energy in (4.2), which are given by

$$E^\Delta = \sum_i \frac{\Delta x_i}{2} \rho_i u_i^2 + F^\Delta \quad \text{and} \quad F^\Delta = \sum_i \Delta x_i \left[\Pi(\rho_i) + V_i \rho_i\right] + \frac{1}{2} \sum_{i,j} \Delta x_i \Delta x_j W_{ij} \rho_i \rho_j. \tag{6.26}$$

It is worth mentioning that previous works have constructed finite-volume schemes that satisfy the discrete analogue of the entropy identity in (4.9) (see "entropy stable schemes" in [132, 266]) and free energy dissipation property in (4.8) (see "energy dissipating schemes" in [15, 55, 71]). The simulations satisfy the discrete energy dissipation in practice, even if we are not able to prove it for our finite-volume scheme. The extension of the present scheme to satisfy the challenging discrete properties of entropy stability and energy dissipation with high-order accuracy will be explored elsewhere.

6.5.1 Validation of the numerical scheme

The validation of the finite-volume scheme encompasses a test for the well-balanced property and a test for the high-order accuracy in the transient regimes. Both tests are conducted in two different scenarios, for which different choices of the free energy in (4.2) are taken. The details of such scenarios are written in Examples 6.5.1 and 6.5.2.

On the one hand, for the well-balanced property we show that the steady-state solution is preserved in time up to machine precision. For this we select as initial condition a density and momentum profile satisfying the steady states obtained from (4.6). Numerically this means that the discrete version of the variation of the free energy in (6.4) holds, while the momentum vanishes throughout the domain. The results for this test are depicted in Table 6.5.1, for a simulation time run from $t = 0$ to $t = 5$ and a number of cells of 50.

Table 6.5.1: Preservation of the steady state for the examples 6.5.1 and 6.5.2 with the third- and fifth-order schemes and machine precision in Matlab, at $t = 5$

	Order of the scheme	L^1 error
Example 6.5.1	3$^{\text{rd}}$	1.7082e-16
	5$^{\text{th}}$	1.7094e-16
Example 6.5.2	3$^{\text{rd}}$	5.5020E-17
	5$^{\text{th}}$	6.4514E-17

On the other hand, the spatial order of accuracy in the transient regimes is computed by evaluating the L^1 error of the numerical solution for a particular mesh-size grid Δx with respect to a reference solution. To measure the spatial high-order accuracy of the scheme we select scenarios where shock waves or sharp gradients in the density and momentum profile do not evolve, such as the ones in the examples 6.5.1 and 6.5.2, and we let the simulation run until $t = 0.1$. We repeat this procedure by halving the Δx from the previous simulation and the order of the scheme is computed as

$$\text{Order of the scheme} = \ln_2 \left(\frac{L^1 \, \text{error}(\Delta x)}{L^1 \, \text{error}(\Delta x/2)} \right).$$

For the system with a nonlocal free energy in (4.1) there are generally no explicit solutions in the transient regime. This implies that the reference solution has to be computed from the same numerical scheme with an extremely refined Δx with the aim of accepting that numerical solution as the exact one. Here we take 25600 cells to compute such reference solution, while the other numerical simulations for the order of accuracy employ 50, 100, 200 and 400 cells. The results showing the spatial third- and fifth-order of accuracy for the scenarios in examples 6.5.1 and 6.5.2 are displayed in Table 6.5.2 and 6.5.3. Note too that the timestep has to be selected as $\Delta t \approx \Delta x$ and $\Delta t \approx (\Delta x)^{5/3}$ for the third- and fifth-order spatial discretizations respectively, if one wants to obtain a correct order of convergence in space. This is because we employ a third-order TVD Runge-Kutta temporal integrator. This is a common practive in the literature, and we refer the reader to [286, Subsection 5.4] as an example with the same choice for Δt.

Example 6.5.1 (Ideal-gas pressure under an attractive potential). For this example we select an ideal-gas free energy with pressure $P(\rho) = \rho$ (or equivalently $\Pi(\rho) = \rho \ln(\rho) - 1$) and with a quadratic external potential $V(x) = x^2/2$. The steady state that we aim to preserve follows from

$$\frac{\delta \mathcal{F}}{\delta \rho} = \Pi'(\rho) + H(x, \rho) = \ln(\rho) + \frac{x^2}{2} = \text{constant on supp}(\rho) \text{ and } u = 0. \tag{6.27}$$

Free energies of this type appear in the context of chemotaxis with a fixed chemoattractant profile [131, 138, 248], where cells will typically vary their direction when reacting to the presence of a chemical substance, so that they are attracted by chemically favourable environments and dodge unfavourable ones. In chemotaxis there is a chemo-attractant function playing a role similar to the external potential $V(x)$, and in more complex models this chemo-attractant function may even have its own parabolic equation for its evolution. There are numerous well balanced schemes for chemotaxis. Amongst them we highlight the fully implicit finite-volume scheme in [129], the scheme

allowing for vacuum states in [216], the Godunov scheme in [153] and the high-order finite-volume and finite-differences schemes in [131, 215, 285].

The density profile for the steady state in (6.27) for an initial mass M_0 satisfies a Gaussian distribution of the form

$$\rho_\infty = M_0 \frac{e^{-x^2/2}}{\int_\mathbb{R} e^{-x^2/2} dx}. \tag{6.28}$$

Table 6.5.2: Spatial order-of-convergence test for Example 6.5.1 with the third- and fifth-order schemes, at $t = 0.1$

Number of cells	Third-order		Fifth-order	
	L^1 error	order	L^1 error	order
50	1.4718E-04	-	1.9260E-05	-
100	2.3726E-05	2.63	5.1254E-07	5.23
200	2.4182E-06	3.29	2.1997E-08	4.54
400	2.6708E-07	3.18	9.2613E-10	4.57

For the spatial order-of-accuracy test we take as initial condition a perturbation of the steady state in (6.28),
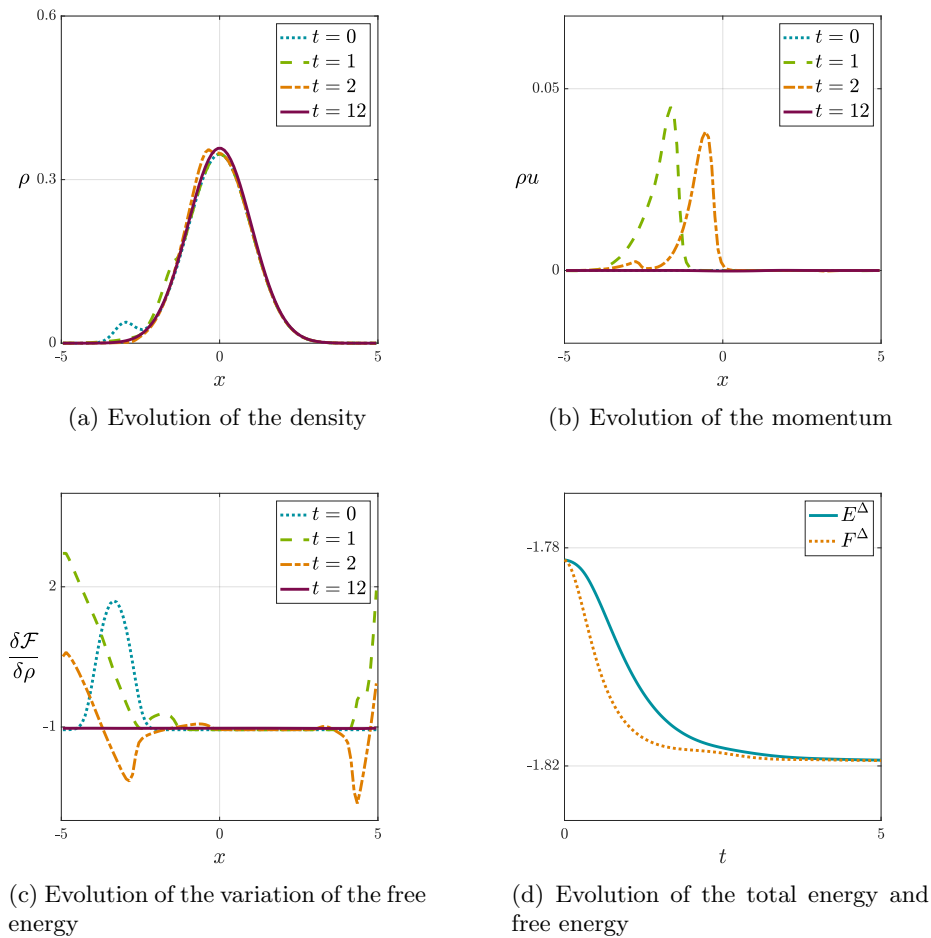
$$\rho(x, t = 0) = M_0 \frac{e^{-x^2/2} + 0.1 * e^{-5(x+3)^2}}{\int_\mathbb{R} \left(e^{-x^2/2} + 0.1 * e^{-5(x+3)^2} \right) dx}, \quad \rho u(x, t = 0) = 0, \quad x \in [-5, 5], \tag{6.29}$$

with M_0 equal to 1 so that the total mass is also 1. The spatial order-of-accuracy test from this example is shown in Table 6.5.2, while the temporal evolution of the density, momentum, free-energy with respect to the density, total energy and free energy are displayed in figure 6.5.1. The spatial third- and fifth-order of accuracy of the numerical scheme are evident from table 6.5.2. From Figure 6.5.1a we notice how the Gaussian distribution corresponding to the steady state is reached at $t = 12$, while in Figure 6.5.1c the variation of the free energy is constant throughout the domain, given that the density is not compactly supported and (6.4) is satisfied. Finally, it is evident from figure Figure 6.5.1d that the discrete analogues of the total energy and free energy (6.26) decay in time.

In Figure 6.5.2 we visually illustrate the difference in accuracy between employing the third- and fifth-order schemes of this work versus the first-order scheme in Chapter 5, which is summarized in Section 6.1. We display the density and momentum fluctuation profiles at two different times ($t = 0.2$ and $t = 0.4$), which result from substracting the initial conditions in (6.29) to the numerical profiles obtained with the same mesh of 100 cells for the three schemes. The choice of measuring the fluctuations with respect to the initial condition is motivated by capturing the transient behaviour. We also plot a reference profile obtained with the third-order scheme and 12600 cells. From Figure 6.5.2 we observe the benefit of employing the high-order schemes in comparison to the first-order one, since they provide a numerical solution much closer to the reference profile for the same number of cells.

Example 6.5.2 (Generalized Euler-Poisson system: ideal-gas pressure and attractive kernel). For this example we select an ideal-gas free energy with pressure $P(\rho) = \rho$ (or equivalently $\Pi(\rho) = \rho \ln(\rho) - 1$) together with an interaction potential with a kernel of the form $W(x) = \frac{x^2}{2}$. In this case the steady state aimed to be preserved satisfies

$$\frac{\delta \mathcal{F}}{\delta \rho} = \Pi'(\rho) + H(x, \rho) = \ln(\rho) + \frac{x^2}{2} \star \rho = \text{constant on supp}(\rho) \text{ and } u = 0. \tag{6.30}$$

74

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy

Figure 6.5.1: Temporal evolution of Example 6.5.1.

Free energies of this type are common in Euler-Poisson systems, in which the Euler equations for a compressible gas are coupled to a self-consistent force field created by the gas particles [162]. This interaction could be gravitational, leading to the modelling of Newtonian stars [37], or electrostatic with repelling forces between the particles is is the case of plasma [97, 160]. For Euler-Poisson systems the free energy contains a function $S(t, x)$ which follows a Poisson-like equation, so that

$$\frac{\delta \mathcal{F}}{\delta \rho} = \Pi'(\rho) + S(t, x) \quad \text{and} \quad \partial_{xx} S(t, x) = c\rho, \tag{6.31}$$

with c being either 1 for the gravitational case or -1 for the plasma one. The Poisson equation for $S(t, x)$ can be solved considering the fundamental solution of the Laplacian in one dimension [190], which leads to $2S = -c|x| \star \rho$. Then, by plugging this expression for S in the variation of the free energy in (6.31), one recovers the interaction potential $W(x)$ which is convoluted with the density ρ. For a $S(t, x)$ following the Poisson equation the interaction potential is $W(x) = -c|x|$, but for $c = -1$ one can generalize it to a homogeneous kernel $W(x) = |x|^\alpha/\alpha$, where $\alpha > -1$ and $W(x) = \ln|x|$ when $\alpha = 0$ for convention. A popular application of these more general kernels $W(x)$ is in the Keller-Segel system for cells and bacteria [36, 49, 63] which we explore in Example 6.5.5.

For Example 3.2 we select $\alpha = 2$, leading to the interaction potential $W(x) = \frac{x^2}{2}$ in the variation of the free energy in (6.30). The steady state for a general mass M_0 is equal to the steady
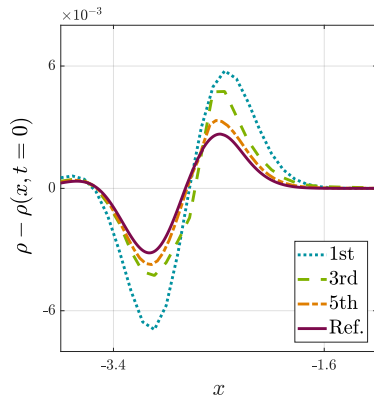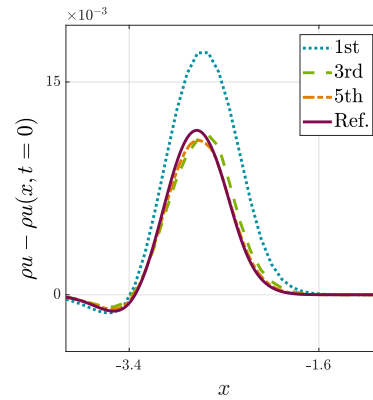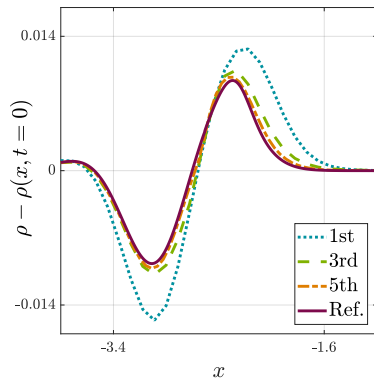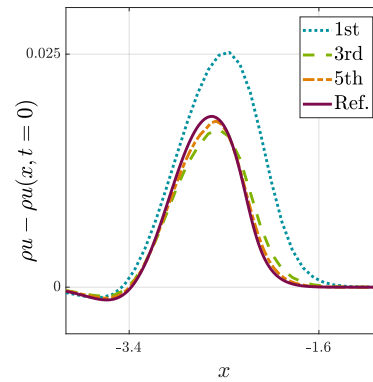
(a) Density fluctuations at $t = 0.2$

(b) Momentum fluctuations at $t = 0.2$

(c) Density fluctuations at $t = 0.4$

(d) Momentum fluctuations at $t = 0.4$

Figure 6.5.2: Density and momentum fluctuations in Example 6.5.1 for a first-, third- and fifth-order scheme with the same mesh of 100 cells. The reference solution is computed with the third-order scheme and 12600 cells.

Table 6.5.3: Spatial order-of-convergence test for Example 6.5.2 with the third- and fifth-order schemes, at $t = 0.1$

| Number of | Third-order | | Fifth-order | |
cells	L^1 error	order	L^1 error	order
50	5.0109E-04	-	1.0913E-04	-
100	1.2721E-04	1.98	5.0556E-06	4.43
200	1.7573E-05	2.86	5.3713E-08	6.56
400	2.3001E-06	2.93	2.3448E-10	4.52

state for Example 6.5.1 and satisfies (6.28). Notice that the particular choice of $W(x) = \frac{x^2}{2}$ and a symmetric initial condition makes this Example analytically equivalent to the case of external quadratic potential in Example 3.1 with the same initial data, just expand the convolution and use symmetry. However, by treating it numerically as a convolution we are able to check the order of accuracy for interaction potentials. For the spatial order-of-accuracy test the initial condition

is a symmetric perturbation of the steady state in (6.28),

$$\rho(x, t = 0) = M_0 \frac{e^{-x^2/2} + 0.05 * e^{-5(x+3)^2} + 0.05 * e^{-5(x-3)^2}}{\int_{\mathbb{R}} \left(e^{-x^2/2} + 0.05 * e^{-5(x+3)^2} + 0.05 * e^{-5(x-3)^2} \right) dx}, \quad \rho u(x, t = 0) = 0,$$

with $x \in [-10, 10]$ and M_0 equal to 1 so that the total mass is also 1. The spatial order-of-accuracy test from this example is shown in Table 6.5.3, while the temporal evolution of the density, momentum, variation of the free energy with respect to the density, total energy and free energy are depicted in Figure 6.5.3. The spatial third- and fifth-order of accuracy of the numerical scheme are evident from table 6.5.3. Figure 6.5.3a shows that the density remains symmetric at all times eventually reaching the steady state profile in (6.28). It is also evident from Figure 6.5.3c that the variation of free energy reaches a constant value in the regions where the density is non-compactly supported, while Figure 6.5.3d demonstrates that the total energy and free energy exhibit a temporal decay.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

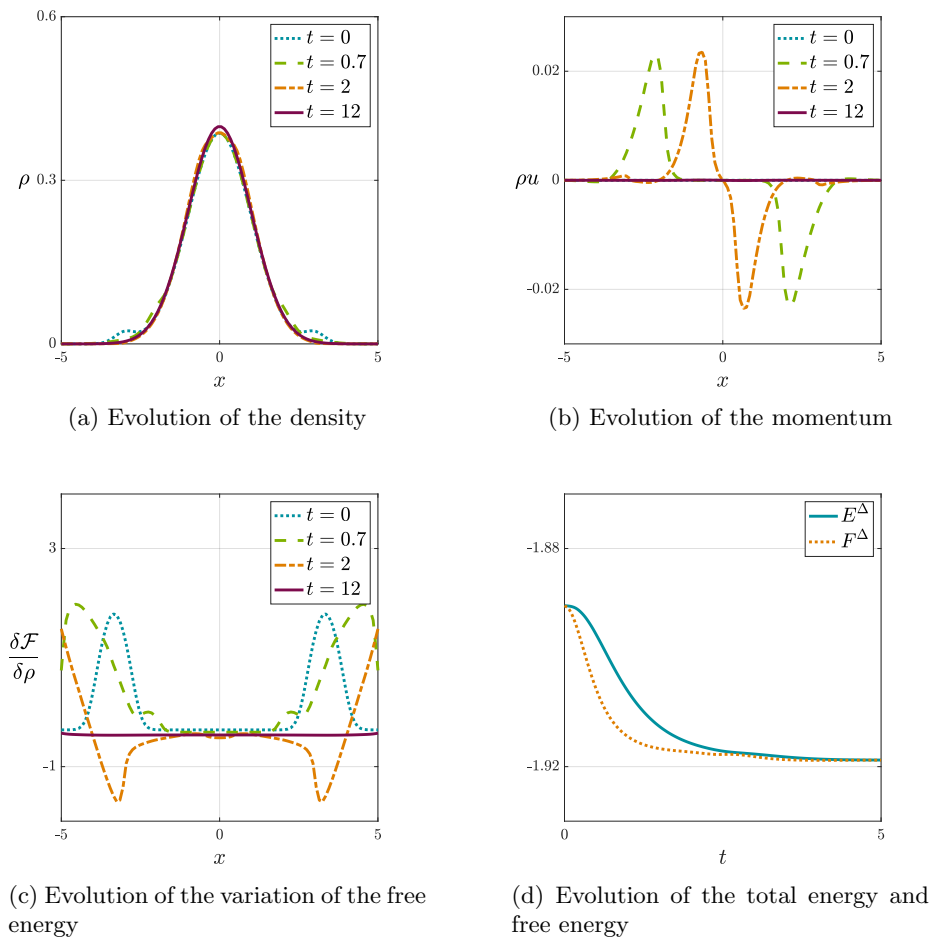(d) Evolution of the total energy and free energy

Figure 6.5.3: Temporal evolution of Example 6.5.2.

6.5.2 Numerical experiments and applications

Here we apply the finite-volume scheme to applications of the shallow-water system, a collective behaviour system with Cucker-Smale and Motsch-Tadmor damping terms, and the Keller-Segel

model. Our scheme is useful to run challenging numerical experiments for which analytical results are limited in the literature, such as in the above applications.

Example 6.5.3 (Shallow water: pressure proportional to square of density and attractive potential). In this example we select a pressure satisfying $P = \rho^2$ (or equivalently $\Pi(\rho) = \rho^2$) together with an attractive external potential $V(x)$. This scenario corresponds to the well-known shallow-water equations, which model free-surface gravity waves whose wavelength is much larger than the characteristic bottom depth. The choice of $P = \rho^2$ leads to the presence of dry regions during the water-height evolution. These equations are applied in a wide range of engineering and scientific applications involving free-surface flows [282], such as tsunami propagation [78], dam break and flooding problems [98] and the evolution of rivers and coastal areas [93].

The main three challenges to accurately simulate the shallow-water equations are the preservation of the steady states, the preservation of the water-height positivity and the transitions between wet and dry areas. Many authors have consequently proposed various numerical schemes addressing these challenges, employing methodologies ranging from finite-difference and finite-volume schemes to discontinuous Galerkin ones. The reader can find more relevant references about high-order schemes [73], well-balanced reconstructions [11], density positivity [299] and the simulation of the wet/dry front [137] in the introduction of this work and in the comprehensive survey from Xing and Shu [282].

For this example we aim to show that our numerical scheme accurately captures the dry regions during the simulation and when reaching the steady states. This is thanks to the combination of the positive-density reconstruction in Appendix 6.B and the choice of a kinetic numerical flux which is able to handle vacuum regions [236]. We show this by conducting simulations with two different choices for the external potential $V(x)$ with the following initial conditions for both cases

$$\rho(x, t = 0) = \frac{e^{-(x-x_0)^2/16}}{\int_{\mathbb{R}} e^{-(x-x_0)^2/16} dx}, \quad \rho u(x, t = 0) = -0.1 \sin\left(\frac{\pi x}{10}\right), \quad x \in [-5, 5],$$

with x_0 being the initial centre of mass. The steady states for the choice of pressure and external potentials of this example satisfy

$$\rho_\infty = (C(x) - V(x))_+, \tag{6.32}$$

where $C(x)$ is a piecewise constant function being zero outside the support of the density.

The details of each simulation are:

1) Single-well external potential and symmetric density: $V(x) = x^2/2$ and $x_0 = 0$. The results of this simulation are depicted in Figure 6.5.4. Figure 6.5.4a shows the formation of the compact support of the density during the time evolution with the steady state taking the shape of a positive parabola and satisfying (6.32). We also observe that the variation of the free energy in figure 6.5.4c reaches a constant value only in the support of the density, in agreement with the steady-state relation in (4.6). We also note that in 6.5.4d the discrete total energy decreases in time, while the discrete free energy has a slight increase around $t = 3$ due to an exchange of energy with the kinetic energy.

2) Double-well external potential and asymmetric density: $V(x) = x^4/4 - 3x^2/2$ and $x_0 = 1.5$. The results of this simulation are depicted in Figure 6.5.5. From the evolution of the density in Figure 6.5.5a it is evident that two compactly-supported bumps of density are formed when reaching the steady state. This is due to the external potential having two wells. In addition, the mass in the bumps is not the same, since the initial density is not symmetric. It is also important to remark that, when reaching the steady state, the variation of the free energy in each compacted support of the density is constant but has different values. This is

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

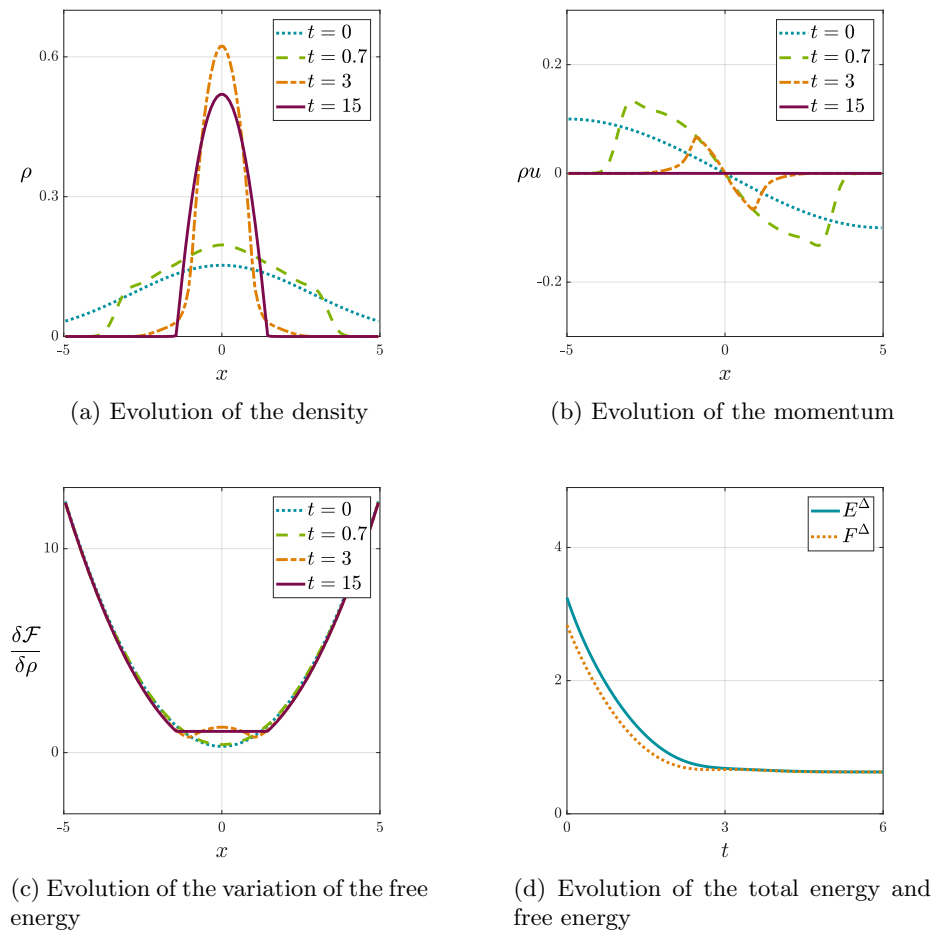(d) Evolution of the total energy and free energy

Figure 6.5.4:   Temporal evolution with single-well external potential and symmetric density in Example 6.5.3.

depicted in Figure 6.5.5c and agrees with the steady state relation (4.6). We refer the reader to our previous work [71] for similar simulations considering varied scenarios with double-well potentials.

Example 6.5.4 (Collective behaviour: comparison of linear, Cucker-Smale and Motsch–Tadmor dampings). In this example we explore the impact of adding linear and nonlinear damping terms to the general system (4.1). The motivation for the nonlinear damping comes from the field of collective behaviour, in which a large amount of interacting individuals or agents organize their dynamics by influencing each other and without the presence of a leader. Most of the literature in collective behaviour is based on individual based models (IBMs) which are particle descriptions considering the three basic effects of attraction, repulsion and alignment of the individuals. The combination of these three effects has proven to be very versatile and extends beyond the typical animal applications for schools of fish [175], herds of mammals [143] or flocks or birds [165]. Indeed, these models are now playing a critical role in understanding complex phenomena including consensus and spatio-temporal patterns in diverse problems ranging from the evolution of human languages [104] to the prediction of criminal behaviour [252] and space flight formation [231].

There are plenty of works in the literature addressing the mean-field derivation of kinetic
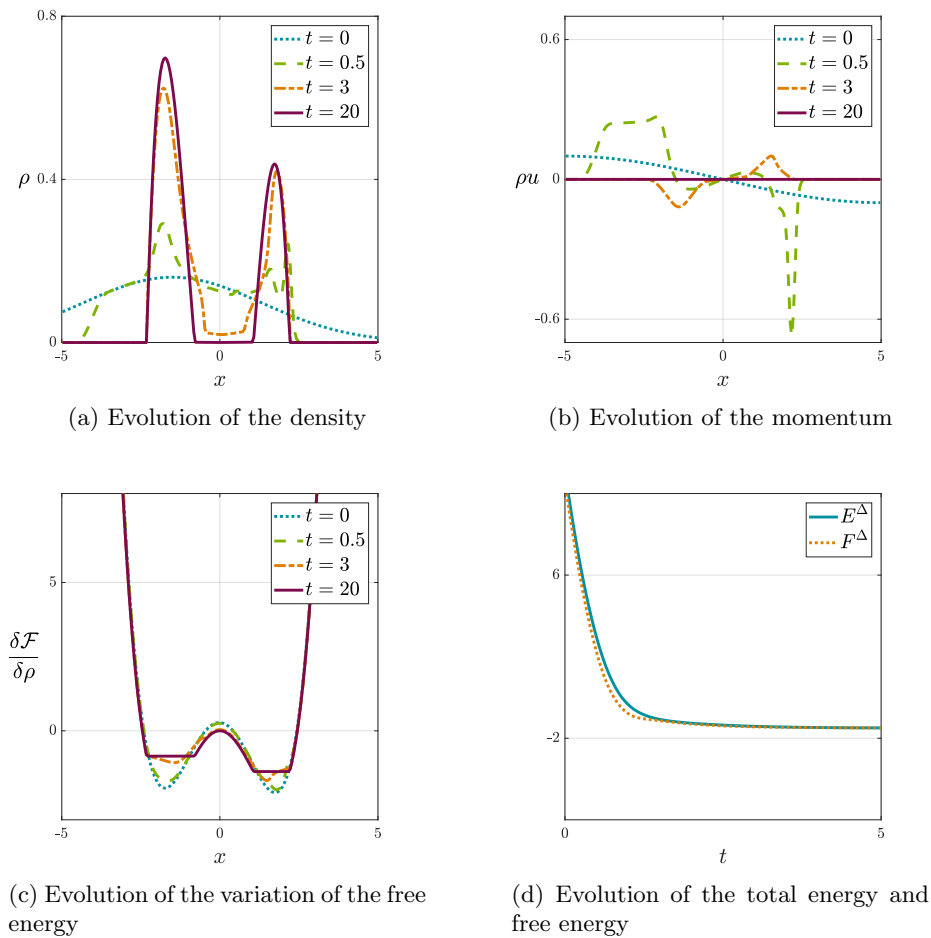
(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy

Figure 6.5.5:   Temporal evolution with double-well external potential and asymmetric density in Example 6.5.3.

and other macroscopic models from the original particle descriptions [61, 62, 161]. These derived hydrodynamic equations agree with our general system (4.1) and model the attraction and repulsion effects via the interaction potential $W(x)$. The third effect for collective behaviour is alignment which in our system (4.1) it is achieved by means of the nonlinear and nonlocal damping of the RHS of the momentum equation. The most popular approach for the velocity consensus is the Cucker-Smale (CS) model [102, 103] which adapts the momentum of a particle depending on the momentum and distance of the other particles. Several authors have proposed refined variations of the CS model, and among them we remark the weighted-normalized model by Motsch and Tadmor (MS) [214] (which will be referred to in the following as the MS model). It basically corrects the CS model by eliminating the normalization over the total number of agents, which leads to inaccurate behaviours in far-from-equilibrium scenarios. Instead, the MT model introduces the concept of relative distances between agents with the cost, however, of destroying the symmetry of the original CS model. For further details on flocking and alignment with the CS and related models we refer the reader to [56, 57, 91, 211].

The objective of this example is to illustrate the differences of adding to the general system

(4.1) linear damping, the CS or the MT model. The damping term for each of them is

$$
\begin{cases}
-\gamma\rho\boldsymbol{u} & \text{if linear damping,} \\[2mm]
-\rho\displaystyle\int_{\mathbb{R}^d}\psi(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{u}(\boldsymbol{x})-\boldsymbol{u}(\boldsymbol{y}))\rho(\boldsymbol{y})\,d\boldsymbol{y} & \text{if Cucker-Smale damping,} \\[2mm]
-\dfrac{\rho}{\psi\star\rho}\displaystyle\int_{\mathbb{R}^d}\psi(\boldsymbol{x}-\boldsymbol{y})(\boldsymbol{u}(\boldsymbol{x})-\boldsymbol{u}(\boldsymbol{y}))\rho(\boldsymbol{y})\,d\boldsymbol{y} & \text{if Motsch-Tadmor damping,}
\end{cases}
\tag{6.33}
$$

where $\psi(x)$ is a nonnegative symmetric smooth function, called the communication function, satisfying for this example

$$
\psi(x) = \frac{1}{(1+|x|^2)^{\frac{1}{4}}}.
$$

It should be noted that the CS damping term in (6.33) would reduce to linear damping if the communication function $\psi(x)$ was a constant function $\psi(x)=1$. In addition, the difference between the CS and MT models is the normalization over $\psi\star\rho$ that is added to the MT model to ensure that the damping term is independent of the total mass of the system.

The simulation for this example is chosen to specifically address a particular drawback of the CS model. This occurs in the evolution of two groups of agents separated by a certain significant distance and whose masses have different degrees of magnitude. What happens with the CS model is that the damping term for the small group of agents is negligible due to the normalization over the total number of agents in the system. This means that those agents do not seek alignment from the beginning of the simulation, and as a result the convergence towards alignment is delayed. On the contrary, with the MT model the normalization over $\psi\star\rho$ in (6.33) allows to take the relative distances between the agents into account, and the small group of agents reacts much faster to the effect of the rest of agents. In the simulation we also add a Morse-like interaction potential [55, 64] of the form $W(x)=-e^{-|x|^2/2}/\sqrt{2\pi}$, which quickly decays at large distances and does not add any attraction between the two groups of agents. Note that we are forced to add this attraction term to balance the pressure and thus allow for our well-balanced scheme. The pressure follows $P(\rho)=\rho$ (or equivalently $\Pi(\rho)=\rho\ln(\rho)-1$).

This configuration is depicted in Figure 6.5.6. Specifically, Figures 6.5.6a and 6.5.6b show with blue the initial conditions for the density and the momentum. On the one hand, in the density there are two groups of agents with mass of 0.9 and 0.1, satisfying

$$
\rho(x,t=0) = 0.9\,\frac{e^{-(x+1)^2/2}}{\int_{\mathbb{R}}e^{-(x+1)^2/2}dx} + 0.1\,\frac{e^{-(x-11)^2}}{\int_{\mathbb{R}}e^{-(x-11)^2}dx}, \quad x\in[-5,14],
$$

while on the other hand for the momentum the groups have opposite velocity signs, in agreement with

$$
\rho u(x,t=0) = \begin{cases} 2\,\rho(x,t=0) & \text{if } x < 5, \\ -2\,\rho(x,t=0) & \text{if } x \geq 5. \end{cases}
$$

What we expect to happen in this situation is that the large group imposes its velocity sign over the small group, so eventually all the agents align with positive velocity. From the momentum simulation in Figure 6.5.6b we observe that after $t=1$ the MT model has already changed the velocity sign of the small group from negative to positive, while for the CS model the velocity is still negative. In general, linear damping is the one that dissipates more momentum, as depicted in the momentum plot of Figure 6.5.6b and in the total energy plot of Figure 6.5.6d. The free energy and total energy decay are similar for both the CS and MT models. A similar numerical experiment was already conducted in [56] using particle methods.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the discrete free energy

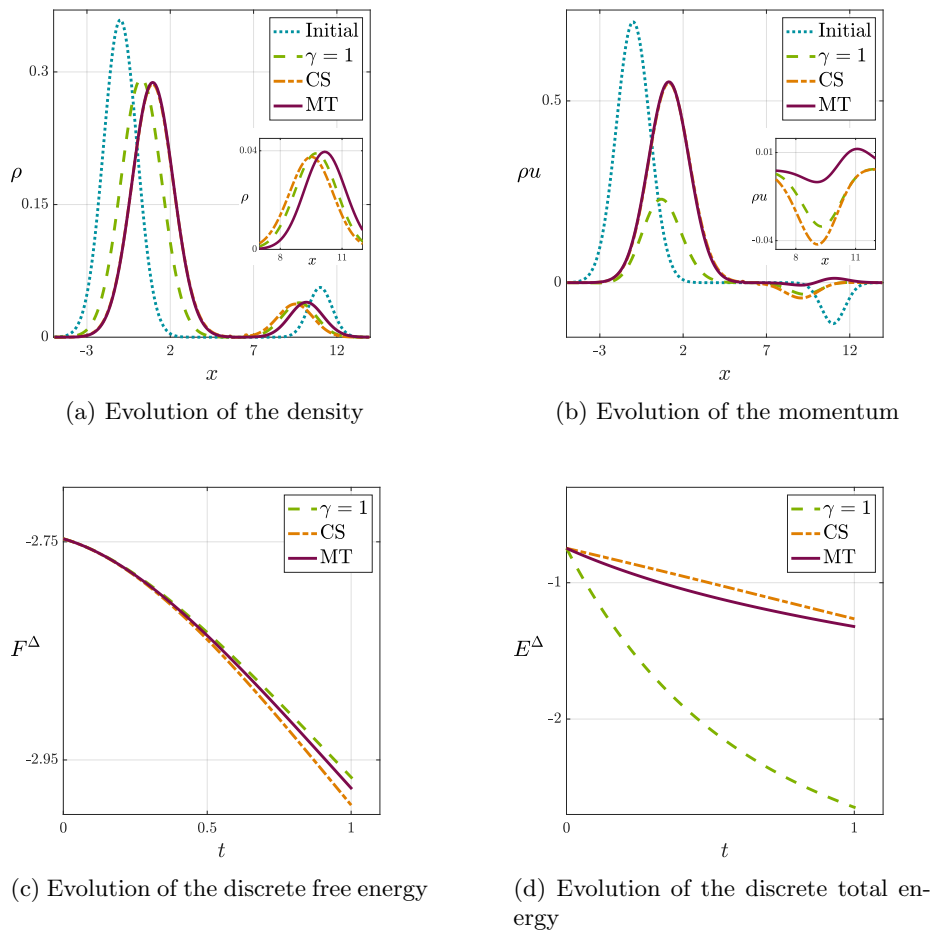(d) Evolution of the discrete total energy

Figure 6.5.6: Simulation of Example 6.5.4 until $t = 1$. $\gamma = 1$ denotes the linear damping simulation, CS the Cucker-Smale simulation and MT the Motsch-tadmor simulation.

Example 6.5.5 (Hydrodynamic Keller-Segel system). The Keller-Segel model has been widely employed for chemotactic aggregation of biological populations such as cells, bacteria or insects. It models how the production of a particular chemical by these organisms leads to long-range attraction and eventually results in self-organization. Its first formulation was proposed in [177] and consisted in a drift-diffusion equation for the density (which is obtained in the overdamped limit of our system (4.1)) coupled with a diffusion equation for the chemical concentration.

In this example we are interested in the hydrodynamic extension of the Keller-Segel model proposed in [83]. It takes into account the inertia of the biological entities and has been proposed in [83]. It follows the same structure as the generalized Euler-Poisson system in the Example 6.5.2 with the free energy satisfying (6.31) and the chemical concentration usually taken as $S = W(x) \star \rho$ (see [50, 63]). The homogeneous kernel $W(x)$ follows $W(x) = |x|^\alpha/\alpha$, where $\alpha > -1$ and $W(x) = \ln|x|$ when $\alpha = 0$ for convention. The difference with Example 6.5.2 is that here the pressure follows $P(\rho) = \rho^m$ with $m \geq 1$ (or equivalently $\Pi(\rho) = \rho^m/(m-1)$), thus allowing for compactly-supported steady states and vacuum in the density if $m > 1$. We refer the reader to [28] for more information about the Keller-Segel model and the diffusion equation for the chemical concentration.

In our previous work [71] we applied our first- and second-order well-balanced scheme to investigate the competition between the attraction from the local kernel $W(x)$ and the repulsion caused
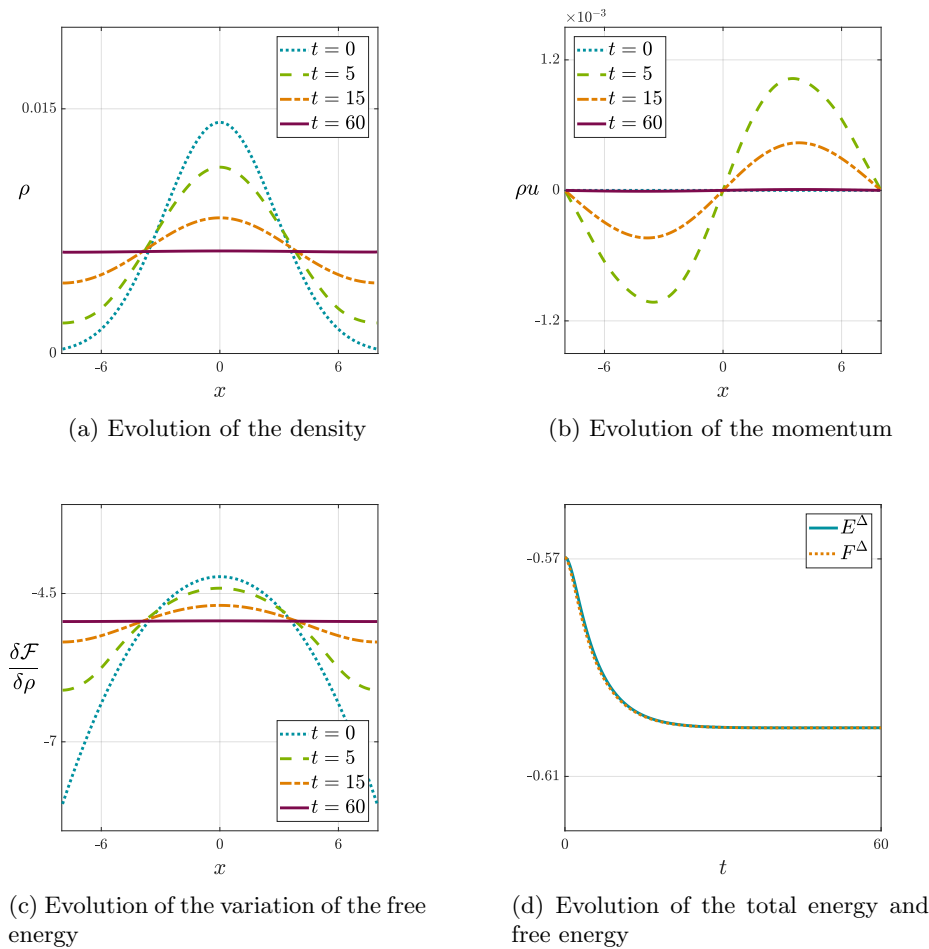
(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

(d) Evolution of the total energy and free energy

Figure 6.5.7: Temporal evolution of the Keller-Segel system with $P = \rho$, $W(x) = \ln |x|$ and initial conditions (6.34) with $M = 0.1$. Global-in-time solution.

by the diffusion of the pressure $P(\rho)$. For the overdamped Keller-Segel model there are basically three possible regimes [49, 50], which result from adequately tuning the parameters α in the kernel $W(x)$ and m in the pressure $P(\rho)$: diffusion dominated regime ($m > 1 - \alpha$), balanced regime ($m = 1 - \alpha$) where a critical mass separates self-similar and blow-up behaviour, and aggregation-dominated regime ($m < 1 - \alpha$). Results with the momentum equation included, and thus inertia, are still quite limited in the literature, with only some specific scenarios studied [57, 69]. In our previous work [71] we investigated the role of inertia for a choice of parameters of $\alpha = 0.5$, $m = 1.5$ and $\alpha = -0.5$, $m = 1.3$, which led to a diffusion-dominated and aggregation-dominated regimes, respectively.

For this example we aim to explore the case of $\alpha = 0$ which leads to the singular potential $W(x) = \ln |x|$. Initially, for the two first simulations of this example we set $m = 1$ so that $P(\rho) = \rho$. In the overdamped limit this scenario corresponds to the balanced regime since $m = 1 - \alpha$, and there is a critical mass separating the global-in-time from the finite-time blowup solution. We run two simulations with identical initial conditions which differ only in a multiplicative constant for the density which allows to set a different mass of the system. The objective is to find global-in-time and finite-time blowup solutions by only changing the mass of the system. For this we set

the initial conditions as

$$\rho(x, t = 0) = M \frac{e^{-(x)^2/16}}{\int_{\mathbb{R}} e^{-(x)^2/16} dx}, \quad \rho u(x, t = 0) = 0, \quad x \in [-8, 8], \tag{6.34}$$

with M being the mass of the system.

For the first simulation we set the mass of the system to be $M = 0.1$. The results are shown in Figure 6.5.7 where the numerical solution is clearly global-in-time and diffusion-dominated. Eventually the steady state is reached,

$$\frac{\delta \mathcal{F}}{\delta \rho} = \Pi'(\rho) + H(x, \rho) = \ln(\rho) + \ln|x| \star \rho = \text{constant on supp}(\rho) \text{ and } u = 0.$$

From Figure 6.5.7a we observe that the solution is completely diffused and the final profile for the density is uniform. From figure 6.5.7c we notice that the variation of the free energy with respect to the density is constant once the steady state is reached, and from Figure 6.5.7d we remark how the total and free energy decay during the temporal evolution.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the variation of the free energy

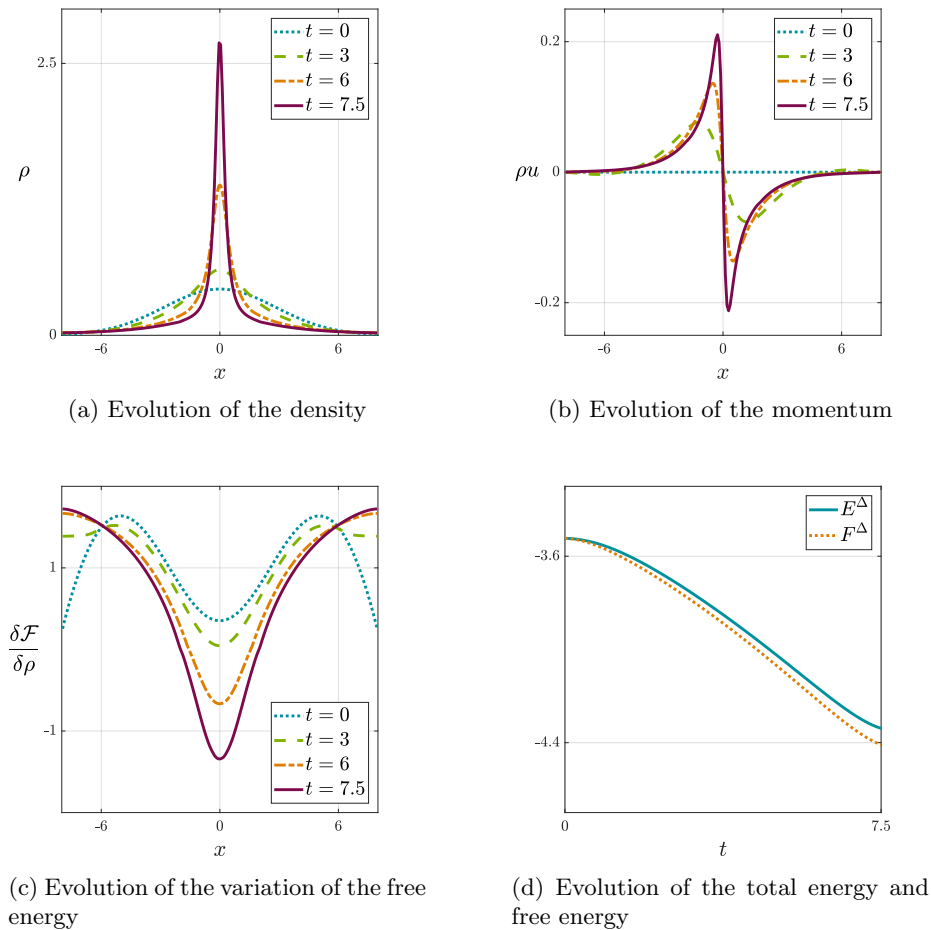(d) Evolution of the total energy and free energy

Figure 6.5.8: Temporal evolution of the Keller-Segel system with $P = \rho$, $W(x) = \ln|x|$ and initial conditions (6.34) with $M = 3$. Finite-time blowup solution.

For the second simulation we select a mass of $M = 3$ while keeping the same initial conditions as in (6.34). As displayed in figure 6.5.8, the solution now presents a finite-time blowup around

$t \approx 7.5$, leading to an aggregation-dominated behaviour. Figure 6.5.8a reveals that the density is concentrated towards the middle of the domain while Figure 6.5.8b shows that the momentum presents an infinite slope when reaching the blowup. From figure 6.5.8d we notice that the total and free energy temporally decay until the blowup occurs.

(a) Evolution of the density

(b) Evolution of the momentum

(c) Evolution of the density
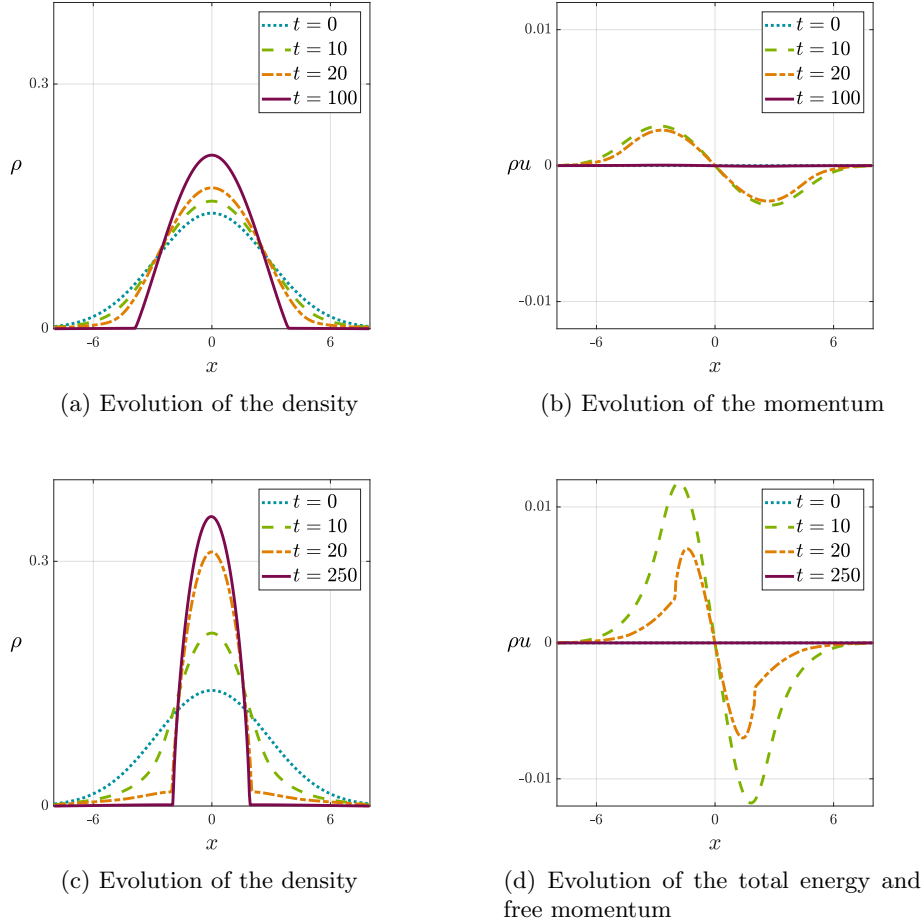
(d) Evolution of the total energy and free momentum

Figure 6.5.9:   Temporal evolution of the Keller-Segel system with $W(x) = \ln|x|$ and initial conditions (6.34) with $M = 1$. Compactly-supported steady state. (a)-(b) $P = 3\rho^2$, steady state $t = 250$, (c)-(d) $P = 3\rho^{2.5}$, steady state $t = 100$.

Finally, we also aim to compare diffusion-dominated solutions where $m > 1$ leading to steady states that are compactly supported. For this purpose we set the initial conditions to be (6.34) with a mass of $M = 1$. For comparison we look at two scenarios with $P = 3\rho^2$ and $P = 3\rho^{2.5}$ so that the exponent m is different. In Figure 6.5.9 we depict the final steady states that arise from the two choices of m. From figures Figures 6.5.9a and 6.5.9c we observe that the final compactly-supported density profiles have slightly different shapes due to the balances between the attraction from the local kernel $W(x)$ and the repulsion caused by the diffusion of the pressure $P(\rho)$.

Appendix

6.A Details about the fifth-order Gaussian quadrature

In this section we detail the procedure to approximate an integral via the fifth-order Gaussian quadrature. Briefly, Gaussian quadratures of n points yield exact values of integrals for polynomials of degree up to $2n - 1$. In our case we implement Gaussian quadratures of 3 points, so that the spatial error is of the order $\mathcal{O}(\Delta x^5)$. In this way we do not limit the order of the high-order finite volume schemes, since the order of this quadrature is always higher or equal than ones in the reconstructions of the scheme.

The approximation of a function $f(x)$ within an interval $[-1, 1]$ via a three-point Gaussian quadrature [2] satisfies

$$\int_{-1}^{1} f(x)\, dx = \frac{5}{9} f\left(-\sqrt{\frac{3}{5}}\right) + \frac{8}{9} f(0) + \frac{5}{9} f\left(\sqrt{\frac{3}{5}}\right).$$

In our case, this integration is performed within each of the finite volume cells in $[x_{i-1/2}, x_{i+1/2}]$ centred at x_i and with size Δx. The cell averages in the finite volume schemes are also divided over Δx. As a result, the transformation of weights and spatial coordinates for the Gaussian quadrature from $[-1, 1]$ to $[x_{i-1/2}, x_{i+1/2}]$ results in

$$\frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} f(x)\, dx = \frac{5}{18} f\left(x_i - \frac{\Delta x}{2}\sqrt{\frac{3}{5}}\right) + \frac{4}{9} f(x_i) + \frac{5}{18} f\left(x_i + \frac{\Delta x}{2}\sqrt{\frac{3}{5}}\right).$$

From this last expression we get that the coefficients α_j for $j \in \{1, 2, 3\}$ satisfy

$$\alpha_1 = \frac{5}{18}, \quad \alpha_2 = \frac{4}{9}, \quad \alpha_3 = \frac{5}{18},$$

with the spatial nodes x_i^j within the cell i for the evaluations of the integrand located at

$$x_i^1 = x_i - \frac{\Delta x}{2}\sqrt{\frac{3}{5}}, \quad x_i^2 = x_i, \quad x_i^3 = x_i + \frac{\Delta x}{2}\sqrt{\frac{3}{5}}.$$

6.B Details about the positive-density CWENO reconstruction

In this section we proceed to summarize the third- and fifth-order CWENO reconstructions of a generic function $g(x)$ whose cell averages $\{g_i\}$, defined as in (6.2), are taken as input. Each cell has size Δx, is centred at $\{x_i\}$ and is contained in the region $[x_{i-1/2}, x_{i+1/2}]$. These reconstructions are applied in (6.13) to compute the high-order reconstructions $R_i^\rho(x)$, $R_i^{\rho u}(x)$ and $R_i^K(x)$. For further details about the CWENO algorithm, we refer the reader to [52, 99, 197, 198].

In Subappendix 6.B.1 we detail the third-order reconstruction, in Subappendix 6.B.2 we proceed with the fifth-order reconstruction, and finally in Subappendix 6.B.3 we end up by summarizing the positive-density limiters from [299], which are essential to prove the positivity of the overall finite volume scheme.

6.B.1 Third-order CWENO reconstruction

The third-order CWENO reconstruction from [197] satisfies

$$R_i^g(x) = \underline{g_i} + \underline{g_i'}(x - x_i) + \frac{1}{2}\underline{g_i''}(x - x_i)^2, \tag{6.35}$$

with $\underline{g_i}$, $\underline{g_i'}$ and $\underline{g_i''}$ resulting from

$$
\begin{aligned}
\underline{g_i} &= w_{i-1}^i\left(\tilde{g}_{i-1} + \Delta x\, \tilde{g}_{i-1}' + \frac{1}{2}\Delta x^2\, \tilde{g}_{i-1}''\right) + w_i^i\, \tilde{g}_j + w_{i+1}^i\left(\tilde{g}_{i+1} - \Delta x\, \tilde{g}_{i+1}' + \frac{1}{2}\Delta x^2\, \tilde{g}_{i+1}''\right), \\
\underline{g_i'} &= w_{i-1}^i\left(\tilde{g}_{i-1}' + \Delta x\, \tilde{g}_{i-1}''\right) + w_i^i\, \tilde{g}_i' + w_{i+1}^i\left(\tilde{g}_{i+1}' - \Delta x\, \tilde{g}_{i+1}''\right), \\
\underline{g_i''} &= w_{i-1}^i\, \tilde{g}_{i-1}'' + w_i^i\, \tilde{g}_i'' + w_{i+1}^i\, \tilde{g}_{i+1}'',
\end{aligned}
\tag{6.36}
$$

and $\tilde{g}_k$, $\tilde{g}_k'$ and $\tilde{g}_k''$, for $k = \{i-1, i, i+1\}$, being computed as

$$\tilde{g}_k = g_k - \frac{g_{k-1} - 2g_k + g_{k+1}}{24}, \quad \tilde{g}_k' = \frac{g_{k+1} - g_{k-1}}{2\Delta x}, \quad \tilde{g}_k'' = \frac{g_{k+1} - 2g_k + g_{k-1}}{\Delta x^2}.$$

The weights w_k^i appearing in (6.36), for $k = \{i-1, i, i+1\}$, satisfy

$$w_k^i = \frac{\alpha_k^i}{\alpha_{i-1}^i + \alpha_i^i + \alpha_{i+1}^i}, \quad \text{where} \quad \alpha_k^i = \frac{C_k}{\left(\epsilon + IS_k^i\right)^p}. \tag{6.37}$$

The constants C_{i-1}, C_i, C_{i+1}, ϵ and p for α_k^i in (6.37) are

$$C_{i-1} = \frac{3}{16}, \quad C_i = \frac{5}{8}, \quad C_{i+1} = \frac{3}{16}, \quad \epsilon = 10^{-6}, \quad p = 3. \tag{6.38}$$

Finally, the smoothness indicators IS_k^i for α_k^i in (6.37), where $k = \{i-1, i, i+1\}$, result from

$$
\begin{aligned}
IS_{i-1}^i &= \frac{13}{12}\left(g_{i-2} - 2g_{i-1} + g_i\right)^2 + \frac{1}{4}\left(g_{i-2} - 4g_{i-1} + 3g_i\right)^2, \\
IS_i^i &= \frac{13}{12}\left(g_{i-1} - 2g_i + g_{i+1}\right)^2 + \frac{1}{4}\left(g_{i-1} - g_{i+1}\right)^2, \\
IS_{i+1}^i &= \frac{13}{12}\left(g_i - 2g_{i+1} + g_{i+2}\right)^2 + \frac{1}{4}\left(3g_i - 4g_{i+1} + g_{i+2}\right)^2.
\end{aligned}
\tag{6.39}
$$

The proposed choice of the constants in (6.38) and the smoothness indicators in (6.39) is based on the original work for third-order CWENO reconstruction in [197]. The reader can find about other more refined choices in [198] and later works [5, 99, 183, 246].

6.B.2 Fifth-order CWENO reconstruction

The fifth-order CWENO reconstruction from [52] satisfies

$$R_i^g(x) = g_{opt}(x) + \sum_{k \in \{1,2,3,c\}} \left(w_k^i - C_k\right) g_k(x), \tag{6.40}$$

with g_{opt}, g_1, g_2, g_3 and g_c resulting from

$$g_{opt}(x) = \sum_{j=1}^{5} a_j(x-x_i)^{j-1}, \quad g_1(x) = \sum_{j=1}^{3} b_j(x-x_i)^{j-1}, \quad g_2(x) = \sum_{j=1}^{3} c_j(x-x_i)^{j-1},$$

$$\tag{6.41}$$

$$g_3(x) = \sum_{j=1}^{3} d_j(x-x_i)^{j-1}, \quad g_c(x) = \left(g_{opt}(x) - C_1\,g_1(x) - C_2\,g_2(x) - C_3\,g_3(x)\right)/C_c.$$

The coefficients a_j for $j \in \{1,2,3,4,5\}$ which appear in the optimal polynomial $g_{opt}(x)$ in (6.41) are taken as

$$a_1 = \frac{1067}{960}g_i - \frac{29}{480}\left(g_{i+1}+g_{i-1}\right) + \frac{3}{640}\left(g_{i+2}+g_{i-2}\right), \quad a_2 = \frac{34\left(g_{i+1}-g_{i-1}\right) + 5\left(g_{i-2}-g_{i+2}\right)}{48\Delta x},$$

$$a_3 = \frac{g_{i-2} + 22\,g_i + g_{i+2} - 12\left(g_{i+1}+g_{i-1}\right)}{-16\Delta x^2}, \quad a_4 = \frac{2\left(g_{i+1}-g_{i-1}\right) + \left(g_{i-2}-g_{i+2}\right)}{-12\Delta x^3},$$

$$a_5 = \frac{g_{i-2} + 6g_i + g_{i+2} - 4\left(g_{i+1}+g_{i-1}\right)}{24\Delta x^4}.$$

The rest of the coefficients b_j, c_j and d_j for $j \in \{1,2,3\}$ which appear in the polynomials g_1, g_2 and g_3, respectively, follow from

$$b_1 = \frac{23}{24}g_i + \frac{1}{12}\left(g_{i-1} - \frac{1}{2}g_{i-2}\right), \quad b_2 = \frac{3g_i - 4g_{i-1} + g_{i-2}}{2\Delta x}, \quad b_3 = \frac{g_i - 2g_{i-1} + g_{i-2}}{2\Delta x^2},$$

$$c_1 = \frac{13}{12}g_i + \frac{1}{24}\left(g_{i-1}+g_{i+1}\right), \quad c_2 = \frac{g_{i+1}-g_{i-1}}{2\Delta x}, \quad c_3 = \frac{g_{i+1} - 2g_i + g_{i-1}}{2\Delta x^2},$$

$$d_1 = \frac{23}{24}g_i + \frac{1}{12}\left(g_{i+1} - \frac{1}{2}g_{i+2}\right), \quad d_2 = \frac{3g_i - 4g_{i+1} + g_{i+2}}{-2\Delta x}, \quad d_3 = \frac{g_i - 2g_{i+1} + g_{i+2}}{2\Delta x^2}.$$

The weights w_k^i for $k \in \{1,2,3,c\}$ in the fifth-order CWENO reconstruction (6.40) satisfy

$$w_k^i = \frac{\alpha_k^i}{\sum_{k\in\{1,2,3,c\}}\alpha_k^i}, \quad \text{where} \quad \alpha_k^i = \frac{C_k}{\left(\epsilon + IS_k^i\right)^p}. \tag{6.42}$$

The constants C_1, C_2, C_3, C_c, ϵ and p for α_k^i in (6.42) are

$$C_1 = \frac{1}{8}, \quad C_2 = \frac{1}{4}, \quad C_3 = \frac{1}{8}, \quad C_c = \frac{1}{2}, \quad \epsilon = 10^{-6}, \quad p = 2.$$

Finally, the smoothness indicators IS_k^i which appear in the computation of α_k^i in (6.42), with $k \in \{1,2,3,c\}$, result from

$$IS_1^i = b_2^2\Delta x^2 + \frac{13}{3}b_3^2\Delta x^4, \quad IS_2^i = c_2^2\Delta x^2 + \frac{13}{3}c_3^2\Delta x^4, \quad IS_3^i = d_2^2\Delta x^2 + \frac{13}{3}d_3^2\Delta x^4,$$

$$IS_4^i = a_2^2\Delta x^2 + \left(\frac{13}{3}a_3^2 + \frac{1}{2}a_2\,a_4\right)\Delta x^4.$$

6.B.3   Positive-density CWENO reconstruction

The third- and fifth-order CWENO reconstructions in (6.35) or (6.40), respectively, can be modified to yield positive values for evaluations at specific spatial points of the finite volume cell. In our case we are interested in obtaining positive values of the density at the points required by the

numerical scheme in (6.21). Namely, those points are: the boundaries of the cell, computed in (6.22) and employed for in (6.23), (6.24) and (6.25), and the quadrature points to compute the integrals in the nonlinear damping term in (6.21), the distribution of the high-order corrections in the source term in (6.21) and described in Appendix 6.C, and the update of K_i^{n+1} in (6.20).

Zhang and Shu [299] proposed a methodology to construct maximum-principle-satisfying high-order schemes. Here we apply their work in [299] to modify the CWENO reconstructions in the subsections 6.B.1 and 6.B.2, so that they preserve the positivity of the density. The procedure is the following:

1) Construct the third- or fifth-order CWENO polynomial in (6.35) or (6.40), respectively, for all the finite volume cells.

2) Evaluate the reconstructed polynomials at the spatial points required by the numerical scheme in (6.21): the boundaries of the cells and quadrature points. We denote as M the total number those spatial points.

3) For each cell i compute the minimum m_i of the evaluations of the reconstructed polynomial at the required spatial points x_i^j, for $j \in \{1, \ldots, M\}$, so that $m_i = \min_{j \in \{1, \ldots, M\}} R_i^g(x_i^j)$.

4) Modify the reconstructed polynomial in (6.35) or (6.40) so that $m_i \geq 0$, in the following way:

$$\widetilde{R_i^g}(x) = \theta \left(R_i^g(x) - g_i \right) + g_i, \quad \text{with} \quad \theta = \min \left\{ \frac{|0 - g_i|}{|m_i - g_i|}, 1 \right\}.$$

5) Evaluate all the M spatial points within each cell with the modified reconstructed polynomial $\widetilde{R_i^g}(x)$.

6) Apply the following CFL condition for the time step Δt, depending on the numerical flux employed and where α_j is the quadrature weight of the spatial point x_i^j, for $j \in \{1, \ldots, M\}$:

$$\Delta t = \begin{cases} \text{CFL} \dfrac{\Delta x \, \min_{j \in \{1, \ldots, M\}} \alpha_j}{\max\limits_{\forall \left(U_{i+\frac{1}{2}}^-, U_{i+\frac{1}{2}}^+ \right)} \left\{ \left| u + \sqrt{P'(\rho)} \right|, \left| u - \sqrt{P'(\rho)} \right| \right\}}, & \text{if Lax-Friedrich flux,} \\[3mm] \text{CFL} \dfrac{\Delta x \, \min_{j \in \{1, \ldots, M\}} \alpha_j}{\max\limits_{\forall \left(U_{i+\frac{1}{2}}^-, U_{i+\frac{1}{2}}^+ \right)} \left\{ |u| + 3^{\frac{m-1}{4}} \right\}}, & \text{if kinetic flux.} \end{cases} \tag{6.43}$$

The quadrature weights α_j employed in this work come from the fifth-order Gaussian quadrature and the integration of the source term corrections in Appendix 6.C. For the details about the numerical fluxes we refer the reader to our previous work [71]. Note too that our CFL condition is constructed to ensure the positivity of the density, but it doesn't set any constraints for the momentum. As a result, the CFL condition doesn't depend on the source terms, given that they do not appear in the density equation and our temporal integrator is explicit. This is fine for our simulations since our source terms are not stiff. It is true though that it may be challenging to simulate stiff source terms, and for that case one may have to rely on more stable implicit temporal integrators with CFL conditions depending on the source term.

6.C   Details about the integration for the high-order corrections

In this section we follow [217, 218] to propose a fourth- and sixth-order quadrature for the high-order corrections in the source term, which appear in the last term of (6.21) and satisfy

$$I_i = \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \left(S \left(R_i^\rho(x), R_i^H(x) \right) - S \left(U_i^*(x), H_i^*(x) \right) \right) \, dx. \tag{6.44}$$

The first step is to define a general trapezoidal numerical quadrature I_i^m for the integral I_i in (6.44), which employs m points $x_i^j = x_{i-1/2} + (j-1)\Delta x/m$ of the cell i, with $j \in \{1, \ldots, m\}$. Such integral yields

$$I_i^m = \sum_{j=1}^{m-1} \left[\frac{R_i^\rho(x_i^j) + R_i^\rho(x_i^{j+1})}{2} \left(R_i^H(x_i^{j+1}) - R_i^H(x_i^j) \right) - \frac{R_i^\rho(x_i^j) + R_i^\rho(x_i^{j+1})}{2} \left(H_i^*(x_i^{j+1}) - H_i^*(x_i^j) \right) \right]. \tag{6.45}$$

The integral I_i^m is a second-order approximation of I_i in (6.44), and its asymptotic form satisfies [217]

$$I_i^m = I_i + c_1 \left(\frac{\Delta x}{m} \right)^2 + c_2 \left(\frac{\Delta x}{m} \right)^4 + \ldots \tag{6.46}$$

The strategy to obtain the fourth- and sixth-order schemes relies in computing the integral I_i in (6.44) as a linear combination of I_i^m for different m, such that the desired errors in (6.46) are cancelled. The required formulas of integration are:

a) Fourth-order quadrature employing I_i^1 and I_i^2, so that

$$I_i = \frac{4I_i^2 - I_i^1}{3} + \mathcal{O}(\Delta x^4).$$

b) Sixth-order quadrature employing I_i^1, I_i^2 and I_i^3, so that

$$I_i = \frac{81}{40}I_i^3 - \frac{16}{15}I_i^2 + \frac{1}{24}I_i^1 + \mathcal{O}(\Delta x^6).$$

As a remark, the order of these quadratures is maintained as long as the order of the reconstructions for the density in (6.13) and the potential in (6.14) is greater than or at least equal to the order of the quadrature formulas. Otherwise the order of the quadrature is diminished and matches the order of the reconstruction.

Part III

An example of overdamped system: the Cahn-Hilliard equation

Introduction to the Cahn-Hilliard equation

The Cahn-Hilliard (CH) equation is a popular phase-field model initially proposed by Cahn and Hilliard [48] to describe the process of phase separation in binary alloys. Since then it has found applicability in a wide variety of modelling problems, from capillarity and wetting phenomena [13, 258], diblock copolymer molecules [298] and tumor growth [140, 278] to image inpainting [33, 45, 65], topology optimization [300] and many more [180]. Similar to other phase-field models, the CH equation avoids the explicit treatment of sharp interfaces altogether by replacing them with thin transition regions through which pertinent variables and physical properties vary rapidly but continuously. The CH equation has a gradient-flow structure of the form

$$\frac{\partial \phi}{\partial t} = \boldsymbol{\nabla} \cdot \left(M(\phi) \boldsymbol{\nabla} \frac{\delta \mathcal{F}[\phi]}{\delta \phi} \right), \tag{7.1}$$

where ϕ is the phase-field, which plays the role of an order parameter and describes the phases of the system and is a continuous function depending on time t and space x. In a binary system it takes on the limiting values $\phi = 1$ in one phase and $\phi = -1$ in the other phase.

The free energy of the system $\mathcal{F}[\phi]$ satisfies

$$\mathcal{F}[\phi] = \int_\Omega \left(H(\phi) + \frac{\epsilon^2}{2} |\nabla \phi|^2 \right) d\Omega + \int_{\partial \Omega} f_w(\phi, \beta) ds, \tag{7.2}$$

where $H(\phi)$ is a double-well potential with minima corresponding to the stable phases in the system, ϵ is a positive parameter related to the diffuse-interface width (e.g. [79]) and $f_w(\phi, \beta)$ is the wall free energy which depends on the phase field at the wall and the (equilibrium) contact angle β (e.g. [279]) – see Fig. 7.0.1 for a schematic of a droplet on a solid substrate with contact angle β). From now on $f_w'(\phi, \beta)$ denotes the derivative of $f_w(\phi, \beta)$ with respect to the phase field ϕ. The variation of the free energy with respect to the phase field $\frac{\delta \mathcal{F}[\phi]}{\delta \phi}$ is typically denoted as the chemical potential ξ. The mobility term $M(\phi)$ is usually taken as a degenerate mobility satisfying a quadratic polynomial with roots at $\phi = \pm 1$,

$$M(\phi) = M_0 (1 - \phi)(1 + \phi), \tag{7.3}$$

but it can be taken just as a constant, $M(\phi) = M_0$. The boundary conditions imposed for the CH equation in (7.1) are a combination of the natural boundary condition for the wall free energy and no-flux for the chemical potential [13, 192],

$$\epsilon^2 \boldsymbol{\nabla} \phi \cdot \boldsymbol{n} = -f_w'(\phi, \beta), \quad M(\phi) \boldsymbol{\nabla} \xi \cdot \boldsymbol{n} = 0, \tag{7.4}$$
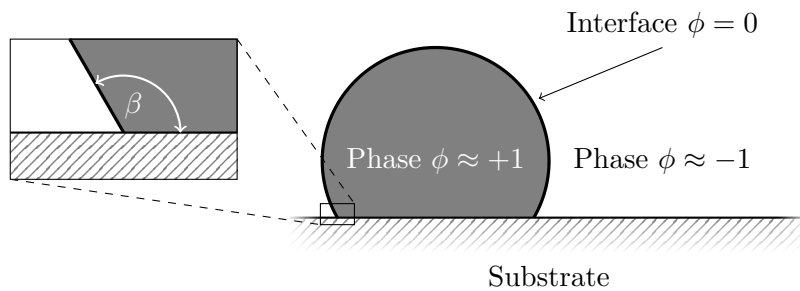
Figure 7.0.1: Schematic of a sessile droplet with $\phi \approx 1$ on a solid substrate and surrounded by a fluid with $\phi \approx -1$. The contact angle between the droplet and substrate is defined as β.

where n is an inward-pointing unit vector normal to the wall. The form of the term $f_w(\phi)$ has received considerable attention in the literature, starting from the early contributions in the field where a linear form was stipulated, e.g. [247]. Here we assume that the function f_w has bounded second derivative on $[-1, 1]$, so that it can be split into a convex part and a concave part satisfying

$$f_w(\phi, \beta) = f_{c,w}(\phi, \beta) - f_{e,w}(\phi, \beta), \tag{7.5}$$

where $f_{c,w}$ and $f_{e,w}$ are convex functions. A good choice is the cubic form [13, 258, 259, 260], which is the lowest-order polynomial to allow for the minimization of the wall free energy for the bulk densities, while at the same time preventing the formation of boundary layers,

$$f_w(\phi, \beta) = \frac{\epsilon\sqrt{2}}{2} \cos \beta \left(\frac{\phi^3}{3} - \phi \right). \tag{7.6}$$

We note that the convex-concave splitting of this cubic function depends on the choice of β. Here we employ

$$f_{c,w}(\phi, \beta) = \begin{cases} \frac{\epsilon\sqrt{2}}{2} \cos \beta \left(\frac{\phi^3}{3} - \phi + \phi^2 \right) & \text{if } \cos \beta \geq 0, \\ -\frac{\epsilon\sqrt{2}}{2} \cos \beta \phi^2 & \text{otherwise,} \end{cases}$$

$$f_{e,w}(\phi, \beta) = f_{c,w}(\phi, \beta) - f_w(\phi, \beta),$$

so that $f_{c,w}(\phi, \beta)$ and $f_{e,w}(\phi, \beta)$ are convex for $\phi \in [-1, 1]$.

The variation of the free energy (7.2) with respect to ϕ follows from

$$\frac{d}{d\gamma}\mathcal{F}[\phi + \gamma\Psi]\bigg|_{\gamma=0} = \int_{\Omega} \left(\Psi H'(\phi) + \epsilon^2 \nabla\phi \cdot \nabla\Psi \right) d\Omega + \int_{\partial\Omega} \Psi f_w'(\phi, \beta) ds$$

$$= \int_{\Omega} \left(H'(\phi) - \epsilon^2 \Delta\phi \right) \Psi d\Omega + \int_{\partial\Omega} \left(\epsilon^2 \nabla\phi \cdot n + f_w'(\phi, \beta) \right) \Psi ds,$$

from which one can apply the boundary conditions (7.4) to derive an expression for the chemical potential ξ,

$$\xi = \frac{\delta\mathcal{F}[\phi]}{\delta\phi} = H'(\phi) - \epsilon^2 \Delta\phi. \tag{7.7}$$

The free energy (7.2) experiences a temporal decay due to the gradient-flow structure of the CH equation in (7.1). This is one of the main properties of gradient flows and is obtained by differentiating the free energy $\mathcal{F}[\phi]$ with respect to time and applying the boundary conditions in (7.4), leading to

$$\frac{d}{dt}\mathcal{F}[\phi] = -\int_{\Omega} M(\phi) \left| \boldsymbol{\nabla} \frac{\delta\mathcal{F}[\phi]}{\delta\phi} \right|^2 d\Omega.$$

Mimicking this free-energy dissipation at the discrete level has been the aim of many numerical works on the CH equation. A scheme endowed with such discrete property is categorized as energy-stable, and the first unconditionally energy-stable scheme was devised by Eyre [128], who put forward the popular convex-splitting technique by which the potential $H(\phi)$ is separated in implicit contractive (convex) and explicit expansive (concave) terms. In fact this semi-implicit formulation has proven so far to be the only method of deriving unconditional energy-stable schemes, and one can show that fully implicit schemes for nonlinear systems such as the CH equation are only conditionally energy-stable depending on the time-step. This counter-intuitive fact has been recently analysed in [288], where the authors prove that the convex-splitting scheme for the CH model is exactly the same as the fully-implicit for a different model that is a nontrivial perturbation of the original CH, and the gain of stability is at the expense of a possible loss of accuracy. In any case, the convex-splitting technique is a mainstream ingredient in the construction of energy-stable schemes for the CH equation, and has been successfully applied in various discretization strategies such as finite differences [135, 159], finite volumes [105], finite elements [20, 25, 114, 278], spectral methods [164] and discontinuous Galerkin schemes [8] (see [268] for an extensive review of energy-stable schemes for the CH equation). A recent and promising strategy to design energy-stable schemes is the so-called scalar auxiliary variable [249, 250].

The convex-splitting technique has been applied for different choices of the potential $H(\phi)$. The two main ones in the literature [24, 268] that we apply in this work are: the Ginzburg-Landau double-well potential,

$$H(\phi) = \frac{1}{4}\left(\phi^2 - 1\right)^2; \tag{7.8}$$

and the logarithmic potential $H_{log}(\phi)$,

$$H_{log}(\phi) = \frac{\theta}{2}\left[(1+\phi)\ln\left(\frac{1+\phi}{2}\right) + (1-\phi)\ln\left(\frac{1-\phi}{2}\right)\right] + \frac{\theta_c}{2}(1-\phi^2) \text{ for } \phi \in (-1,1), \tag{7.9}$$

where θ and θ_c are positive constants with $\theta < \theta_c$, corresponding to the absolute temperature and absolute critical temperature, respectively. The logarithmic potential, usually referred to as the Flory-Huggins energy potential in the polymer-science community [117], is considered more physically realistic in comparison to the polynomial double-well potential in (7.8) because it can be mathematically derived from regular or ideal solution theories [117, 191]. It is, however, singular as ϕ approaches -1 or 1, in contrast to the double-well potential (7.8). The usual convex-splitting for the double-well (7.8) and logarithmic (7.9) potentials satisfy

$$H(\phi) = H_c(\phi) - H_e(\phi) = \frac{\phi^4 + 1}{4} - \frac{\phi^2}{2},$$

$$H_{log}(\phi) = H_{c,log}(\phi) - H_{e,log}(\phi) = \frac{\theta}{2}\left[(1+\phi)\ln\left(\frac{1+\phi}{2}\right) + (1-\phi)\ln\left(\frac{1-\phi}{2}\right)\right] - \left(-\frac{\theta_c}{2}(1-\phi^2)\right), \tag{7.10}$$

so that all the functions H_c, H_e, $H_{c,log}$, and $H_{e,log}$ are convex functions.

Another fundamental property that has received much attention, both in terms of PDE analysis and construction of numerical schemes, is the maximum principle or boundedness of the phase-field. On the one hand, the phase-field solution of the CH equation with logarithmic potential satisfies $|\phi| < 1$ due to the singularities of the potential (7.9) at $\phi = \pm 1$, both for degenerate and constant mobilities. This has already been proved at the PDE level in various works [1, 212], including for degenerate mobilities of the type (7.3) in [24, 125]. On the other hand, for the double-well potential (7.8) the phase-field solution might leave the interval $(-1, 1)$ in the general case [92, 268], due to the absence of singularities in the potential. There are, however, two cases when one can analytically prove a maximum principle for the double-well potential (7.8): firstly, for degenerate mobilities of the type (7.3) vanishing when $\phi = \pm 1$, the phase field is bounded in $|\phi| \leq 1$, as shown

in [125]; secondly, for general mobilities by truncating the potential (7.8) with quadratic growth at infinities, as shown in [46] and applied in [94, 251].

The construction of numerical schemes with the discrete maximum principle property has attracted considerable attention, especially in recent years. The pioneering work by Copetti and Elliot [95] proposed a fully implicit scheme that satisfies a discrete maximum principle under a condition for Δt, which depends on ϵ and the critical temperature θ_0. More recent works have sought to derive schemes that unconditionally satisfy the discrete maximum principle. An important contribution has been by Chen *et al.* [86], where a finite-difference scheme with unconditional discrete maximum principle for the logarithmic potential case is constructed for both constant and degenerate mobilities. In another recent work [134], the authors propose a flux-limiting technique based on high-order discontinuous Galerkin schemes and which unconditionally preserves global bounds for a family of PDEs including the CH equation. Many other works apply the truncation of the potential (7.8) with quadratic growth at infinities, with the objective of forcing their simulations to satisfy the discrete maximum principle when no rigorous proof can be derived [84, 251]. In spite of these remarkable efforts, there is still no scheme flexible enough to satisfy the discrete maximum principle for a general family of free energy potentials including the double-well (7.8) or the logarithmic (7.9), and at the same time allowing for wetting boundary conditions such as (7.4).

7.1 Contribution to finite-volume schemes for the Cahn-Hilliard equation

The main thrust of this work is precisely the construction of a finite-volume scheme that unconditionally satisfies both the discrete maximum principle and free-energy dissipation. The scheme maintains these two fundamental properties for general potentials including the double-well (7.8) and logarithmic (7.9), wetting conditions such as (7.5) as well as more general wall free energies, and degenerate mobilities of the type (7.3) vanishing when $\phi = \pm 1$. In contrast to previous works, the scheme is not restricted to only particular choices of the free energy potential such as the logarithmic one (7.9). In addition, for the case of the double-well potential (7.8) we do not rely on truncated potentials. Furthermore, the scheme is efficiently extended to higher-dimensional configurations due its flexible and cost-saving dimensional-splitting nature thanks to an upwind and finite-volume formulation. The computational cost can be further reduced with a straightforward parallelization resulting from the dimensional-splitting approach. The satisfaction of these unconditional properties imposes a trade-off in the order of our scheme, which in this work is limited to first-order accuracy. Nevertheless, the extension of this scheme to high order will be explored in future works. The present study builds naturally from our previous works aimed at designing structure-preserving finite-volume schemes for gradient flows and hydrodynamic systems, where a general free energy containing nonlocal interaction potentials drives the temporal evolution towards a steady state dictated by the minimizer of such free energy [15, 17, 54, 55, 66, 243].

Chapter 8 is devoted to this contribution. First, in Section 8.1 we explain the construction of the one-dimensional semi-implicit scheme, with its properties of conservation of mass, boundedness of the phase field and free-energy dissipation proved in Subsection 8.1.1. In Section 8.2 we continue with the two-dimensional semi-implicit scheme based on the dimensional-splitting formulation, and its related properties are proved in Subsection 8.2.1. Next, in Section 8.3 we depict a battery of simulations for relevant applications of the CH equation. In Subsection 8.3.1 we begin by validating the first-order spatial order of convergence of our scheme, for both one- and two-dimensional test cases. In Subsection 8.3.2 we provide one-dimensional simulations to analyze the impact of selecting different free-energy potentials and mobility functions. Finally, in Subsection 8.3.3 we depict two-dimensional simulations considering the evolution of a initial random field with different mobilities

and including wall free energy terms leading to contact angles between the droplets and substrates.

7.2   Contribution to Cahn-Hilliard image inpainting

The objective of this work is to show precisely the applicability of our numerical framework in [16] for a benchmark dataset of images in need of restoration through image inpainting. For this task we purposely add different types and intensities of damage to the popular MNIST dataset [112], and then apply image inpainting by solving the modified CH equation (9.1) with our finite-volume scheme in [16].   The MNIST database (Modified National Institute of Standards and Technology database) is a standard dataset to validate image-processing methodologies, and it contains binary handwritten images of numbers from zero to nine. We choose this dataset since the modified CH equation (9.1) is applicable to binary images such as the ones in the MNIST dataset. The extension of this work to non-binary images is also relevant and will be explored elsewhere. We also assess the improvement in pattern recognition accuracy of the restored MNIST images, and for this we construct a neural network for the task of classification. A key objective of our study is to quantify the benefits of including a CH filter before introducing a damaged image into a neural network. Our results demonstrate that accuracies in classification can increase up to 50% for particular damages in the images, and, in general, applying the CH filter improves the accuracy prediction for a wide range of low to moderate image damage. This increase in accuracy is obtained by applying our image-inpainting methodology to the MNIST dataset exclusively, and as a disclaimer we remark that other datasets or types of damage may result in different increases of accuracy. The application of other types of filters, such as texture inpainting or generative inpainting, may also result in different increases of accuracy, but overall the accuracy should be higher compared to the case where no filter is applied to the damaged image.

The robust semi-implicit finite-volume scheme for the CH equation in Chapter 8 offers crucial advantages when applied to the field of image inpainting:

- Firstly, finite volumes are a straightforward discretization when dealing with images, which often consist of rectangular-shaped pixel cells with an average color intensity. This is exactly the starting point of finite-volume schemes, and as a result it is conceptually simpler to apply finite volumes in comparison with finite elements, finite differences or discontinuous Galerkin (which would be more suitable for other more complex and rare pixel shapes such as triangular ones).

- Secondly, our scheme is based on a dimensional-splitting approach: instead of solving the full two-dimensional (2D) image altogether, this technique initially solves row by row and then column by column. This has a massive benefit in computational cost, which is reduced from $\mathcal{O}(N^{d\gamma})$ for an image with N cells in d dimensions to $\mathcal{O}(dN^{d+\gamma-1})$, with $2 < \gamma < 3$. The reason for this is that the cost of inverting a $N \times N$ matrix is $\mathcal{O}(N^{d\gamma})$, with a value of $2 < \gamma < 3$ that slightly varies depending on the inversion algorithm and matrix structure (see [96] for details). For images with N cells per dimension, the solution of the full 2D scheme involves inverting a $N^d \times N^d$ Jacobian matrix, with a subsequent cost of $\mathcal{O}(N^{d\gamma})$. In contrast, the dimensional-splitting technique requires inverting dN^{d-1} Jacobians of size $N \times N$, amounting for a total computational cost of $\mathcal{O}(dN^{d+\gamma-1})$. This is already advantageous for a 2D image, and the computational cost is further reduced for high-dimensional images, such as the ones for [41] or X-ray computed tomography [158] in medical image analysis. To add more, such dimensional-splitting technique allows for parallelization, and it is possible to half the total computational cost by solving nonadjacent rows and columns in parallel.

- Thirdly, our scheme has been extensively tested in [16] for challenging configurations of the original CH equation (7.1). In addition, in [16] we prove that the scheme unconditionally

satisfies the discrete decay of the free energy for different choices of potentials (7.8) [40], while at the same time we prove the phase-field boundedness for mobilities of the form $M(\phi) = 1 - \phi^2$. Even though the modified CH equation (9.1) is not strictly a gradient flow due to the inclusion of the fidelity term, our scheme preserves its robustness for all the image-inpainting test cases presented in this work.

The combination of these properties and reduced computational cost, together with the versatility of finite volumes, make our scheme efficient and robust for the solution of the modified CH equation in (9.1) for a variety of applications in image inpainting.

In Section 9.2 we outline the methodology: in Subsection 9.2.1 we adapt our finite-volume scheme in [16] for the modified CH equation in (9.1); in Subsection 9.2.2 we recall the two-step method for image inpainting in [33]; in Subsection 9.2.3 we detail the neural network architecture for the classification task; and lastly in Subsection 9.2.4 we explain the structure of the integrated algorithm which takes a damaged image, applies a CH filter to it, and then classifies the image through a neural network. Subsequently in Section 9.3 we present the results of the integrated algorithm applied to the MNIST dataset: in Subsection 9.3.1 we begin by identofying appropiate tunings for the values of ϵ and λ_0; in Subsection 9.3.2 we present the different types of damage introduced into the MNIST testset of images; and finally in Subsection 9.3.3 we quantify the improvement in accuracy of applying the CH filter to the damaged MNIST images before introducing them into the neural network.

Finite-volume schemes for the Cahn-Hilliard equation

The contents of this chapter are based on the publication [16], done in collaboration with Rafael Bailo, José A. Carrillo, and Serafim Kalliadasis:

[16]   R. Bailo, J. A. Carrillo, S. Kalliadasis, and S. P. Perez. "Unconditional bound-preserving and energy-dissipating finite-volume schemes for the Cahn-Hilliard equation". In: *arXiv preprint arXiv:2105.05351* (2021). To be submitted.

8.1   One-dimensional semi-implicit scheme

For 1D finite-volume approximation of the CH equation (7.1) the computational domain $[0, L]$ is divided into N cells $C_i = [x_{i-1/2}, x_{i+1/2}]$, all with uniform size $\Delta x = L/N$, so that the centres of the cells satisfy $x_i = (i - 1)\Delta x + \Delta x/2$, $i \in 1, \ldots, N$. In each of the cells C_i we define the cell average ϕ_i as

$$\phi_i(t) = \frac{1}{\Delta x} \int_{C_i} \phi(x, t)dx.$$

Subsequently, one has to integrate the CH equation (7.1) over each of the cells C_i, resulting in

$$\phi_i^{n+1} - \phi_i^n = -\frac{\Delta t}{\Delta x} \left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1} \right). \tag{8.1}$$

For the approximation of the flux at the boundary we follow an upwind approach inspired by Refs [15, 55], with the fluxes computed as

$$F_{i+1/2}^{n+1} = \left(u_{i+1/2}^{n+1} \right)^+ M(\phi_i^{n+1}, \phi_{i+1}^{n+1}) + \left(u_{i+1/2}^{n+1} \right)^- M(\phi_{i+1}^{n+1}, \phi_i^{n+1}) \tag{8.2}$$

where the velocity $u_{i+1/2}^{n+1}$ satisfies

$$u_{i+1/2}^{n+1} = -\frac{\xi_{i+1}^{n+1} - \xi_i^{n+1}}{\Delta x}$$

and the upwind is obtained from

$$\left(u_{i+1/2}^{n+1} \right)^+ = \max(u_{i+1/2}^{n+1}, 0), \quad \left(u_{i+1/2}^{n+1} \right)^- = \min(u_{i+1/2}^{n+1}, 0).$$

The discretized mobility in (8.2) is approximated with implicit density values from the left of the boundary, ϕ_i^{n+1}, and values from the right, ϕ_{i+1}^{n+1}. We construct the discretized mobility as

$$M(x,y) = M_0(1+x)^+(1-y)^+ \quad \text{or} \quad M(x,y) = M_0, \tag{8.3}$$

depending on the choice of mobility taken in (7.3). The proof of the unconditional boundedness of our numerical scheme in Subsection 8.1.1 relies heavily on the approximated mobility (8.3).

The discretized variation of the free energy ξ_i^{n+1} follows a semi-implicit formulation: the contractive part of the potential, $H_c(\rho)$, is taken as implicit; the expansive part of the potential, $H_e(\rho)$, is taken as explicit; and the Laplacian is taken as an average between the explicit and the implicit second-order discretizations. More precisely, the approximation of ξ_i^{n+1} is given by

$$\xi_i^{n+1} = H_c'(\phi_i^{n+1}) - H_e'(\phi_i^n) - \frac{\epsilon^2}{2}\left[(\Delta\phi)_i^n + (\Delta\phi)_i^{n+1}\right] + \frac{1}{\Delta x}W_i(\phi_i^{n+1},\phi_i^n), \tag{8.4}$$

where $(\Delta\phi)_i$ is the discrete 1D second-order approximation of the Laplacian appearing in (7.7), defined as

$$(\Delta\phi)_i^n = \frac{\phi_{i+1}^n - 2\phi_i^n + \phi_{i-1}^n}{\Delta x^2}, \quad (\Delta\phi)_i^{n+1} = \frac{\phi_{i+1}^{n+1} - 2\phi_i^{n+1} + \phi_{i-1}^{n+1}}{\Delta x^2},$$

and the wetting term $W(\phi_i^{n+1},\phi_i^n)$ is only evaluated at the boundaries,

$$W_i(\phi_i^{n+1},\phi_i^n) = \begin{cases} f_{c,w}'(\phi_1^{n+1},\beta) - f_{e,w}'(\phi_1^n,\beta) & \text{if } i = 1; \\ f_{c,w}'(\phi_N^{n+1},\beta) - f_{e,w}'(\phi_N^n,\beta) & \text{if } i = N; \\ 0 & \text{otherwise.} \end{cases} \tag{8.5}$$

The no-flux boundary conditions (7.4) are implemented by taking the numerical flux to vanish at the boundaries,

$$F_{i-1/2}^{n+1} = 0 \quad \text{if } i = 1; \quad F_{i+1/2}^{n+1} = 0 \quad \text{if } i = N; \tag{8.6}$$

and by computing the Laplacian terms $(\Delta\phi)_i^n$ and $(\Delta\phi)_i^{n+1}$ at the boundaries as

$$\begin{cases} (\Delta\phi)_i^n = \frac{\phi_2^n - \phi_1^n}{\Delta x^2}, \quad (\Delta\phi)_i^{n+1} = \frac{\phi_2^{n+1} - \phi_1^{n+1}}{\Delta x^2} & \text{if } i = 1; \\ (\Delta\phi)_i^n = \frac{-\phi_N^n + \phi_{N-1}^n}{\Delta x^2}, \quad (\Delta\phi)_i^{n+1} = \frac{-\phi_N^{n+1} + \phi_{N-1}^{n+1}}{\Delta x^2} & \text{if } i = N. \end{cases} \tag{8.7}$$

8.1.1   Properties of the scheme

The finite-volume scheme proposed in Section 8.1 satisfies the following properties:

(i) conservation of mass so that $\sum_{i=1}^N \phi_i^n = \sum_{i=1}^N \phi_i^{n+1}$;

(ii) boundedness of the phase-field ϕ for mobilities of the form $M(\phi) = M_0(1+\phi)(1-\phi)$, so that if $|\phi_i^n| \leq 1 \,\forall i$, then $|\phi_i^{n+1}| \leq 1 \,\forall i$;

(iii) dissipation of the discrete free energy, defined as

$$\mathcal{F}_\Delta^n = \Delta x \sum_{i=1}^N \left(H_e(\phi_i^n) - H_c(\phi_i^n)\right) + \Delta x \sum_{i=1}^{N-1} \frac{\epsilon^2}{2}|(\nabla\phi)_{i+1/2}^n|^2 + f_w(\phi_1^n) + f_w(\phi_N^n), \tag{8.8}$$

where $(\nabla\phi)_{i+1/2}^n$ is the discrete 1D approximation of the gradient at the interface, satisfying the first-order form $(\nabla\phi)_{i+1/2}^n := \frac{\phi_{i+1}^n - \phi_i^n}{\Delta x}$. The discrete dissipation is

$$\mathcal{F}_\Delta^{n+1} - \mathcal{F}_\Delta^n \leq -\Delta t \Delta x \sum_{i=1}^{N-1} \min\left(M(\phi_i^{n+1},\phi_{i+1}^{n+1}), M(\phi_{i+1}^{n+1},\phi_i^{n+1})\right)\left|u_{i+1/2}^{n+1}\right|^2 \leq 0. \tag{8.9}$$

Proof. Some of the ideas in these proofs are inspired by the studies in Refs [14, 15, 34, 55].

(i) The conservation of mass follows from summing the finite-volume scheme (8.1) for all cells C_i and applying the no-flux conditions (8.7),

$$\sum_{i=1}^{N} \left(\phi_j^{n+1} - \phi_j^n\right) = -\frac{\Delta t}{\Delta x} \sum_{i=1}^{N} \left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1}\right) = -\frac{\Delta t}{\Delta x} \left(F_{N+1/2}^{n+1} - F_{1-1/2}^{n+1}\right) = 0.$$

(ii) We follow the general proof in [14] to show the boundedness of our scheme via contradiction. Without loss of generality, assume first that there is a group of contiguous cells satisfying $\phi_i^{n+1} > 1$, for some values i. The cells belonging to such group are $\{\phi_j^{n+1}, \phi_{j+1}^{n+1}, \ldots, \phi_k^{n+1}\}$. The proof also applies if there are more groups or if the groups have only one cell. The next step is to sum the scheme (8.1) over the group of cells, resulting in

$$\frac{\Delta x}{\Delta t} \sum_{i=j}^{k} \left(\phi_i^{n+1} - \phi_i^n\right) = -\left(F_{k+1/2}^{n+1} - F_{j-1/2}^{n+1}\right). \tag{8.10}$$

Since we have assumed that $\phi_i^{n+1} > 1$ for $i \in \{j, j+1, \ldots, k\}$, it follows that the left-hand side of (8.10) is positive. As a result, the right-hand side of (8.10) also has to be positive,

$$0 < -\left(F_{k+1/2}^{n+1} - F_{j-1/2}^{n+1}\right) = -\left(u_{k+1/2}^{n+1}\right)^+ M(\phi_k^{n+1}, \phi_{k+1}^{n+1}) - \left(u_{k+1/2}^{n+1}\right)^- M(\phi_{k+1}^{n+1}, \phi_k^{n+1})$$
$$+ \left(u_{j-1/2}^{n+1}\right)^+ M(\phi_{j-1}^{n+1}, \phi_j^{n+1}) + \left(u_{j-1/2}^{n+1}\right)^- M(\phi_j^{n+1}, \phi_{j-1}^{n+1}). \tag{8.11}$$

The first and fourth terms in the right-hand side of (8.11) are negative since the mobility function returns a nonnegative value, due to (8.3), $\phi_{k+1}^{n+1} \le 1$ and $\phi_{j-1}^{n+1} \le 1$. The second and third terms are zero since $\phi_k^{n+1} > 1$ and $\phi_j^{n+1} > 1$, according to (8.3). As a result, the whole of the right-hand side of (8.11) is negative implying a contradiction that comes from assuming $\phi_i^{n+1} > 1$ for $i \in \{j, j+1, \ldots, k\}$.

Proving that $\phi_i^{n+1} > -1$ is done by following an identical strategy based on contradiction.

(iii) To show the energy dissipation we first have to take the finite-volume scheme in (8.1), multiply it over ξ_i^{n+1} in (8.4) and sum it over all cells C_i, yielding

$$\sum_{i=1}^{N} \left(\phi_i^{n+1} - \phi_i^n\right) \xi_i^{n+1} = -\frac{\Delta t}{\Delta x} \sum_{i=1}^{N} \left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1}\right) \xi_i^{n+1}.$$

Then, by substituting the expression for ξ_i^{n+1} in (8.4) and rearranging, it follows that

$$\begin{aligned}
\sum_{i=1}^{N} \left(\phi_i^{n+1} - \phi_i^n\right) \frac{\epsilon^2}{2} \left[(\Delta\phi)_i^n + (\Delta\phi)_i^{n+1}\right] = &\frac{\Delta t}{\Delta x} \sum_{i=1}^{N} \left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1}\right) \xi_i^{n+1} \\
&+ \sum_{i=1}^{N} \left(\phi_i^{n+1} - \phi_i^n\right) \left(H_c'(\phi_i^{n+1}) - H_e'(\phi_i^n)\right) \\
&+ \left(\phi_1^{n+1} - \phi_1^n\right) \frac{W_1(\phi_i^{n+1}, \phi_i^n)}{\Delta x} \\
&+ \left(\phi_N^{n+1} - \phi_N^n\right) \frac{W_N(\phi_N^{n+1}, \phi_N^n)}{\Delta x}.
\end{aligned} \tag{8.12}$$

100

Now we proceed to subtract the discrete free energies in (8.8) at subsequent times,

$$
\begin{aligned}
\frac{\mathcal{F}_\Delta^{n+1} - \mathcal{F}_\Delta^n}{\Delta x} = {} & \sum_{i=1}^N \left(H_c(\phi_i^{n+1}) - H_c(\phi_i^n) \right) - \sum_{i=1}^N \left(H_e(\phi_i^{n+1}) - H_e(\phi_i^n) \right) \\
& + \frac{\epsilon^2}{2} \sum_{i=1}^{N-1} \left(|(\nabla \phi)_{i+1/2}^{n+1}|^2 - |(\nabla \phi)_{i+1/2}^n|^2 \right) \\
& + \frac{1}{\Delta x} \left(f_w(\phi_1^{n+1}) - f_w(\phi_1^n) + f_w(\phi_N^{n+1}) - f_w(\phi_N^n) \right).
\end{aligned}
\tag{8.13}
$$

The next step is to expand the term with the discrete gradients at the interfaces. For that we apply summation by parts and the boundary conditions in (8.7), leading to

$$
\begin{aligned}
& \frac{\epsilon^2}{2} \sum_{i=1}^{N-1} \left(|(\nabla \phi)_{i+1/2}^{n+1}|^2 - |(\nabla \phi)_{i+1/2}^n|^2 \right) \\
={} & \frac{\epsilon^2}{2} \sum_{i=1}^{N-1} \left((\nabla \phi)_{i+1/2}^{n+1} + (\nabla \phi)_{i+1/2}^n \right) \left((\nabla \phi)_{i+1/2}^{n+1} - (\nabla \phi)_{i+1/2}^n \right) \\
={} & \frac{\epsilon^2}{2} \sum_{i=1}^{N-1} \left(\frac{\phi_{i+1}^{n+1} - \phi_i^{n+1}}{\Delta x} + \frac{\phi_{i+1}^n - \phi_i^n}{\Delta x} \right) \left(\frac{\phi_{i+1}^{n+1} - \phi_i^{n+1}}{\Delta x} - \frac{\phi_{i+1}^n - \phi_i^n}{\Delta x} \right) \\
={} & - \frac{\epsilon^2}{2} \sum_{i=2}^{N-1} \left(\frac{\frac{\phi_{i+1}^{n+1} - \phi_i^{n+1}}{\Delta x} - \frac{\phi_i^{n+1} - \phi_{i-1}^{n+1}}{\Delta x}}{\Delta x} + \frac{\frac{\phi_{i+1}^n - \phi_i^n}{\Delta x} - \frac{\phi_i^n - \phi_{i-1}^n}{\Delta x}}{\Delta x} \right) \left(\phi_i^{n+1} - \phi_i^n \right) \\
& + \frac{\epsilon^2}{2} \left(\frac{\phi_N^{n+1} - \phi_{N-1}^{n+1}}{\Delta x^2} + \frac{\phi_N^n - \phi_{N-1}^n}{\Delta x^2} \right) \left(\phi_N^{n+1} - \phi_N^n \right) \\
& - \frac{\epsilon^2}{2} \left(\frac{\phi_2^{n+1} - \phi_1^{n+1}}{\Delta x^2} + \frac{\phi_2^n - \phi_1^n}{\Delta x^2} \right) \left(\phi_1^{n+1} - \phi_1^n \right) \\
={} & - \sum_{i=1}^N \left(\phi_i^{n+1} - \phi_i^n \right) \frac{\epsilon^2}{2} \left[(\Delta \phi)_i^n + (\Delta \phi)_i^{n+1} \right].
\end{aligned}
\tag{8.14}
$$

The outcome of the last computations is the left-hand side of (8.12). We can then connect (8.14) and (8.12) to obtain

$$
\begin{aligned}
\frac{\epsilon^2}{2} \sum_{i=1}^{N-1} \left(|(\nabla \phi)_i^{n+1}|^2 - |(\nabla \phi)_i^n|^2 \right) = {} & - \frac{\Delta t}{\Delta x} \sum_{i=1}^N \left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1} \right) \xi_i^{n+1} \\
& - \sum_{i=1}^N \left(\phi_i^{n+1} - \phi_i^n \right) \left(H_c'(\phi_i^{n+1}) - H_e'(\phi_i^n) \right) \\
& - \left(\phi_1^{n+1} - \phi_1^n \right) \frac{W_1(\phi_i^{n+1}, \phi_i^n)}{\Delta x} - \left(\phi_N^{n+1} - \phi_N^n \right) \frac{W_N(\phi_N^{n+1}, \phi_N^n)}{\Delta x}.
\end{aligned}
\tag{8.15}
$$

Then, by taking into account (8.15), (8.13) can be rewritten as

$$\frac{\mathcal{F}_\Delta^{n+1} - \mathcal{F}_\Delta^n}{\Delta x} = \sum_{i=1}^N \left(H_c(\phi_i^{n+1}) - H_c(\phi_i^n) - \left(\phi_i^{n+1} - \phi_i^n\right) H_c'(\phi_i^{n+1}) \right)$$

$$- \sum_{i=1}^N \left(H_e(\phi_i^{n+1}) - H_e(\phi_i^n) - \left(\phi_i^{n+1} - \phi_i^n\right) H_e'(\phi_i^n) \right)$$

$$+ \frac{1}{\Delta x} \left(f_w(\phi_1^{n+1}) - f_w(\phi_1^n) - \left(\phi_1^{n+1} - \phi_1^n\right) W_1(\phi_i^{n+1}, \phi_i^n) \right)$$

$$+ \frac{1}{\Delta x} \left(f_w(\phi_N^{n+1}) - f_w(\phi_N^n) - \left(\phi_N^{n+1} - \phi_N^n\right) W_N(\phi_i^{n+1}, \phi_i^n) \right)$$

$$- \frac{\Delta t}{\Delta x} \sum_{i=1}^N \left(F_{i+1/2}^{n+1} - F_{i-1/2}^{n+1} \right) \xi_i^{n+1}$$

$$= I + II + III + IV + V.$$

Due to the convexity of both $H_c(\phi)$ and $H_e(\phi)$, which satisfy

$$H_c(\phi_i^n) - H_c(\phi_i^{n+1}) - \left(\phi_i^n - \phi_i^{n+1}\right) H_c'(\phi_i^{n+1}) \geq 0,$$
$$H_e(\phi_i^{n+1}) - H_e(\phi_i^n) - \left(\phi_i^{n+1} - \phi_i^n\right) H_e'(\phi_i^n) \geq 0,$$

it results that $I \leq 0$ and $II \leq 0$.

Due to the convex splitting of f_w in (7.5), and the construction of $W_i(\phi_i^{n+1}, \phi_i^n)$ in (8.5), it follows that

$$f_{c,w}(\phi_1^n) - f_{c,w}(\phi_1^{n+1}) - \left(\phi_1^n - \phi_1^{n+1}\right) f_{c,w}'(\phi_1^{n+1}, \beta) \geq 0,$$
$$f_{e,w}(\phi_1^{n+1}) - f_{e,w}(\phi_1^n) - \left(\phi_1^{n+1} - \phi_1^n\right) f_{e,w}'(\phi_1^n, \beta) \geq 0.$$

The same holds for $i = N$. As a result $III \leq 0$ and $IV \leq 0$.

For V we can apply the discrete summation by parts as well as the no-flux conditions in (8.6),

$$III = -\frac{\Delta t}{\Delta x} \sum_{i=1}^{N-1} F_{i+1/2}^{n+1} \left(\xi_i^{n+1} - \xi_{i+1}^{n+1} \right) = -\Delta t \sum_{i=1}^{N-1} F_{i+1/2}^{n+1} u_{i+1/2}^{n+1}$$

$$= -\Delta t \sum_{i=1}^{N-1} \left(\left(u_{i+1/2}^{n+1}\right)^+ M(\phi_i^{n+1}, \phi_{i+1}^{n+1}) + \left(u_{i+1/2}^{n+1}\right)^- M(\phi_{i+1}^{n+1}, \phi_i^{n+1}) \right) u_{i+1/2}^{n+1}$$

$$\leq -\Delta t \sum_{i=1}^{N-1} \min \left(M(\phi_i^{n+1}, \phi_{i+1}^{n+1}), M(\phi_{i+1}^{n+1}, \phi_i^{n+1}) \right) \left| u_{i+1/2}^{n+1} \right|^2 \leq 0,$$

and this is precisely the decay rate for the discrete free energy written in (8.9).

$\square$

8.2 Two-dimensional semi-implicit dimensional-splitting scheme

Here we construct a dimensional-splitting finite-volume scheme to solve the 2D CH equation (7.1). This scheme is more computationally-efficient than a full 2D scheme, and as we will demonstrate

it satisfies the unconditional properties of decay of the discrete free energy and boundedness of the phase-field.

The 2D finite-volume approximation of the CH equation (7.1) follows from initially dividing the computational domain $[0, L] \times [0, L]$ in $N \times N$ cells $C_{i,j} := [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}]$, all with uniform size $\Delta x \Delta y$ so that $x_{i+1/2} - x_{i-1/2} = \Delta x$ and $y_{j+1/2} - y_{j-1/2} = \Delta y$. In each of the cells we define the cell average ϕ_i as

$$\phi_{i,j}(t) = \frac{1}{\Delta x \Delta y} \int_{C_{i,j}} \phi(x, y, t) \, dx \, dy. \tag{8.16}$$

For the dimensional-splitting approach we firstly update the solution along the x directions, for each index j corresponding to a fixed value of y_j where $j \in [1, N]$. Subsequently, we proceed in the same way along the y directions, for each index i corresponding to a fixed value a value of x_i where $i \in [1, N]$. The index r, where $r \in [1, N]$, denotes the index j of the fixed y_j value in every x direction of the first loop, and the updated average density for each x direction with $j = r$ is $\phi_{i,j}^{n,r}$. Similarly, the index $c \in [1, N]$ denotes the index i for every fixed value of x_j in each y direction of the second loop, and the updated density for each y direction with $i = c$ is $\phi_{i,j}^{n,c}$.

To begin with we march along each of the x direction of the domain, each of them at a fixed y_j with $j = r$. The initial conditions for the scheme are $\phi^{n,0} := \phi^n$. The scheme for each x direction satisfies:

Step 1.- for each $r = 1, \ldots, N$ do:

$$\phi_{i,j}^{n,r} = \begin{cases} \phi_{i,j}^{n,r-1} - \frac{\Delta t}{\Delta x} \left(F_{i+1/2,j}^{n,r} - F_{i-\frac{1}{2},j}^{n,r} \right) & \text{if } j = r; \\ \phi_{i,j}^{n,r-1} & \text{otherwise}; \end{cases} \tag{8.17a}$$

$$F_{i+1/2,j}^{n,r} = \left(u_{i+1/2,j}^{n,r} \right)^+ M(\phi_{i,j}^r, \phi_{i+1,j}^r) + \left(u_{i+1/2,j}^{n,r} \right)^- M(\phi_{i+1,j}^r, \phi_{i,j}^r); \tag{8.17b}$$

$$\left(u_{i+1/2,j}^{n,r} \right)^+ = \max \left\{ u_{i+1/2,j}^{n,r}, 0 \right\}, \quad \left(u_{i+1/2,j}^{n,r} \right)^- = \min \left\{ u_{i+1/2,j}^{n,r}, 0 \right\}; \tag{8.17c}$$

$$u_{i+1/2,j}^{n,r} = -\frac{\xi_{i+1,j}^{n,r} - \xi_{i,j}^{n,r}}{\Delta x}; \tag{8.17d}$$

$$\xi_{i,j}^{n,r} = H_c'(\phi_{i,j}^{n,r}) - H_e'(\phi_{i,j}^{n,r-1}) - \frac{\epsilon^2}{2} \left[(\Delta \phi)_{i,j}^{n,r-1} + (\Delta \phi)_{i,j}^{n,\star} \right] \tag{8.17e}$$

$$+ \frac{1}{\Delta x} W_{i,j}^x(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1}) + \frac{1}{\Delta y} W_{i,j}^y(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1});$$

$$(\Delta \phi)_{i,j}^{n,r-1} = \frac{\phi_{i+1,j}^{n,r-1} - 2\phi_{i,j}^{n,r-1} + \phi_{i-1,j}^{n,r-1}}{\Delta x^2} + \frac{\phi_{i,j+1}^{n,r-1} - 2\phi_{i,j}^{n,r-1} + \phi_{i,j-1}^{n,r-1}}{\Delta y^2}; \tag{8.17f}$$

$$(\Delta \phi)_{i,j}^{n,\star} = \frac{\phi_{i+1,j}^{n,r} - 2\phi_{i,j}^{n,r} + \phi_{i-1,j}^{n,r}}{\Delta x^2} + \frac{\phi_{i,j+1}^{n,r-1} - 2\phi_{i,j}^{n,r} + \phi_{i,j-1}^{n,r-1}}{\Delta y^2}. \tag{8.17g}$$

The wetting terms $W_{i,j}^x(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1})$ and $W_{i,j}^y(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1})$ only apply at the boundaries and as in (8.5) satisfy,

$$W_{i,j}^x(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1}) = \begin{cases} f_{c,w}'(\phi_{1,j}^{n,r}, \beta) - f_{e,w}'(\phi_{1,j}^{n,r-1}, \beta) & \text{if } i = 1; \\ f_{c,w}'(\phi_{N,j}^{n,r}, \beta) - f_{e,w}'(\phi_{N,j}^{n,r-1}, \beta) & \text{if } i = N; \\ 0 & \text{otherwise}; \end{cases} \tag{8.18}$$

and

$$W_{i,j}^y(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1}) = \begin{cases} f_{c,w}'(\phi_{i,1}^{n,r}, \beta) - f_{e,w}'(\phi_{i,1}^{n,r-1}, \beta) & \text{if } j = 1; \\ f_{c,w}'(\phi_{i,N}^{n,r}, \beta) - f_{e,w}'(\phi_{i,N}^{n,r-1}, \beta) & \text{if } j = N; \\ 0 & \text{otherwise}. \end{cases} \tag{8.19}$$

The no-flux conditions (7.4) are numerically implemented by taking the numerical flux to vanish at the boundaries,

$$F_{i-\frac{1}{2},j}^{n,r} = 0 \quad \text{if } i = 1; \quad F_{i+1/2,j}^{n,r} = 0 \quad \text{if } i = N;$$

and by computing the Laplacian terms $(\Delta\phi)_{i,j}^{n,r-1}$ and $(\Delta\phi)_{i,j}^{n,\star}$ at the boundaries considering that

$$\begin{cases} \phi_{i,j}^{n,r-1} - \phi_{i-1,j}^{n,r-1} = 0, \quad \phi_{i,j}^{n,r} - \phi_{i-1,j}^{n,r} = 0 & \text{if } i = 1; \\ \phi_{i+1,j}^{n,r-1} - \phi_{i,j}^{n,r-1} = 0, \quad \phi_{i+1,j}^{n,r} - \phi_{i,j}^{n,r} = 0 & \text{if } i = N; \\ \phi_{i,j}^{n,r-1} - \phi_{i,j-1}^{n,r-1} = 0, \quad \phi_{i,j}^{n,r} - \phi_{i,j-1}^{n,r} = 0, \quad \phi_{i,j}^{n,r} - \phi_{i,j-1}^{n,r-1} = 0 & \text{if } j = 1; \\ \phi_{i,j+1}^{n,r-1} - \phi_{i,j}^{n,r-1} = 0, \quad \phi_{i,j+1}^{n,r} - \phi_{i,j}^{n,r} = 0, \quad \phi_{i,j+1}^{n,r} - \phi_{i,j}^{n,r-1} = 0 & \text{if } j = N. \end{cases} \quad (8.20)$$

Once the loop for the rows is completed, we define the intermediate density values as $\rho^{n+1/2} := \rho^{n,N}$. Subsequently, we continue through each of the y direction with index $c = 1, \ldots, N$, each of them at a fixed x_i with $i = c$. The initial condition for this scheme is $\phi^{n,0} := \phi^{n+1/2}$:

Step 2.- for each $c = 1, \ldots, N$ do:

$$\phi_{i,j}^{n,c} = \begin{cases} \phi_{i,j}^{n,c-1} - \frac{\Delta t}{\Delta y} \left(G_{i,j+1/2}^{n,c} - G_{i,j-\frac{1}{2}}^{n,c} \right) & \text{if } i = c; \\ \phi_{i,j}^{n,c-1} & \text{otherwise}; \end{cases} \quad (8.21a)$$

$$G_{i,j+1/2}^{n,c} = \left(v_{i,j+1/2}^{n,c} \right)^+ M(\phi_{i,j}^c, \phi_{i,j+1}^c) + \left(v_{i,j+1/2}^{n,c} \right)^- M(\phi_{i,j+1}^c, \phi_{i,j}^c); \quad (8.21b)$$

$$\left(v_{i,j+1/2}^{n,c} \right)^+ = \max\left\{ v_{i,j+1/2}^{n,c}, 0 \right\}, \quad \left(v_{i,j+1/2}^{n,c} \right)^- = \min\left\{ v_{i,j+1/2}^{n,c}, 0 \right\}; \quad (8.21c)$$

$$v_{i,j+1/2}^{n,c} = -\frac{\xi_{i,j+1}^{n,c} - \xi_{i,j}^{n,c}}{\Delta y}; \quad (8.21d)$$

$$\xi_{i,j}^{n,c} = H_c'(\phi_{i,j}^{n,c}) - H_e'(\phi_{i,j}^{n,c-1}) - \frac{\epsilon^2}{2}\left[(\Delta\phi)_{i,j}^{n,c-1} + (\Delta\phi)_{i,j}^{n,\diamond} \right] \quad (8.21e)$$

$$+ \frac{1}{\Delta x}W_{i,j}^x(\phi_{i,j}^{n,c}, \phi_{i,j}^{n,c-1}) + \frac{1}{\Delta y}W_{i,j}^y(\phi_{i,j}^{n,c}, \phi_{i,j}^{n,c-1});$$

$$(\Delta\phi)_{i,j}^{n,c-1} = \frac{\phi_{i+1,j}^{n,c-1} - 2\phi_{i,j}^{n,c-1} + \phi_{i-1,j}^{n,c-1}}{\Delta x^2} + \frac{\phi_{i,j+1}^{n,c-1} - 2\phi_{i,j}^{n,c-1} + \phi_{i,j-1}^{n,c-1}}{\Delta y^2}; \quad (8.21f)$$

$$(\Delta\phi)_{i,j}^{n,\diamond} = \frac{\phi_{i+1,j}^{n,c-1} - 2\phi_{i,j}^{n,c} + \phi_{i-1,j}^{n,c-1}}{\Delta x^2} + \frac{\phi_{i,j+1}^{n,c} - 2\phi_{i,j}^{n,c} + \phi_{i,j-1}^{n,c}}{\Delta y^2}. \quad (8.21g)$$

Here the terms $W_{i,j}^x(\phi_{i,j}^{n,c}, \phi_{i,j}^{n,c-1})$ and $W_{i,j}^y(\phi_{i,j}^{n,c}, \phi_{i,j}^{n,c-1})$ are defined as in (8.18) and (8.19). The no-flux conditions (7.4) are numerically implemented by taking the numerical flux to vanish at the boundaries,

$$G_{i,j-1/2}^{n,c} = 0 \quad \text{if } j = 1; \quad G_{i,j+1/2}^{n,c} = 0 \quad \text{if } j = N;$$

and by computing the Laplacian terms $(\Delta\phi)_{i,j}^{n,c-1}$ and $(\Delta\phi)_{i,j}^{n,\diamond}$ at the boundaries considering that

$$\begin{cases} \phi_{i,j}^{n,c-1} - \phi_{i-1,j}^{n,c-1} = 0, \quad \phi_{i,j}^{n,c-1} - \phi_{i-1,j}^{n,c} = 0, \quad \phi_{i,j}^{n,c} - \phi_{i-1,j}^{n,c-1} = 0 & \text{if } i = 1; \\ \phi_{i+1,j}^{n,c-1} - \phi_{i,j}^{n,c-1} = 0, \quad \phi_{i+1,j}^{n,c-1} - \phi_{i,j}^{n,c} = 0, \quad \phi_{i+1,j}^{n,c} - \phi_{i,j}^{n,c-1} = 0 & \text{if } i = N; \\ \phi_{i,j}^{n,c-1} - \phi_{i,j-1}^{n,c-1} = 0, \quad \phi_{i,j}^{n,c} - \phi_{i,j-1}^{n,c} = 0 & \text{if } j = 1; \\ \phi_{i,j+1}^{n,c-1} - \phi_{i,j}^{n,c-1} = 0, \quad \phi_{i,j+1}^{n,c} - \phi_{i,j}^{n,c} = 0, & \text{if } j = N. \end{cases}$$

Once the loop for the columns is completed, we define the final density values ρ^{n+1} after a discrete timestep Δt as $\rho^{n+1} := \rho^{n,N}$.

8.2.1 Properties of the dimensional-splitting 2D scheme

The dimensional-splitting 2D finite-volume scheme satisfies the following properties:

(i) conservation of mass so that $\sum_{i,j=1}^{N} \phi_{i,j}^{n} = \sum_{i,j=1}^{N} \phi_{i,j}^{n+1}$;

(ii) boundedness of the phase-field ϕ for mobilities of the form $M(\phi) = M_0(1+\phi)(1-\phi)$, so that if $|\phi_{i,j}^{n}| \leq 1 \, \forall i$, then $|\phi_{i,j}^{n+1}| \leq 1 \, \forall i$;

(iii) dissipation of the discrete free energy, $\mathcal{F}_\Delta^{n+1} - \mathcal{F}_\Delta^{n} \leq 0$, with $\mathcal{F}_\Delta^{n}$ defined as

$$
\begin{aligned}
\mathcal{F}_\Delta^{n} =& \Delta x \Delta y \sum_{i,j=1}^{N} \left(H_c(\phi_{i,j}^{n}) - H_e(\phi_{i,j}^{n}) \right) \\
&+ \Delta x \Delta y \sum_{i=1}^{N-1} \sum_{j=1}^{N} \frac{\epsilon^2}{2} \left(\frac{\phi_{i+1,j}^{n} - \phi_{i,j}^{n}}{\Delta x} \right)^2 + \Delta x \Delta y \sum_{i=1}^{N} \sum_{j=1}^{N-1} \frac{\epsilon^2}{2} \left(\frac{\phi_{i,j+1}^{n} - \phi_{i,j}^{n}}{\Delta y} \right)^2 \quad (8.22) \\
&+ \Delta x \sum_{j=1}^{N} \left(f_w(\phi_{1,j}^{n}) + f_w(\phi_{N,j}^{n}) \right) + \Delta y \sum_{i=1}^{N} \left(f_w(\phi_{i,1}^{n}) + f_w(\phi_{i,N}^{n}) \right).
\end{aligned}
$$

Proof. (i) For the conservation of mass we need to show that the mass is conserved in every row and every column, and this can be done as in the 1D case of Subsection 8.1.1.

(ii) For the unconditional boundedness of the phase-field we need to show that the phase-field is bounded in every row and every column. This is accomplished by following the contradiction strategy employed for the 1D case of Subsection 8.1.1, which in this case has to be applied independently to every row and column.

(iii) For the decay of the discrete free energy in (8.22) we show that the discrete free energy decays for every row and for every column. For this we refer the reader to Lemmas 8.A.1 and 8.A.2 in Section 8.A, where we firstly show that for every row r we have that

$$\mathcal{F}_\Delta \left(\rho^{n,r} \right) - \mathcal{F}_\Delta \left(\rho^{n,r-1} \right) \leq 0 \quad \forall r,$$

and subsequently for every column c we have that

$$\mathcal{F}_\Delta \left(\rho^{n,c} \right) - \mathcal{F}_\Delta \left(\rho^{n,c-1} \right) \leq 0 \quad \forall c.$$

Consequently, it follows that the global decay of the discrete free energy is satisfied.

$\square$

Remark 8.2.1 (Full 2D scheme). It is possible to construct a full 2D implicit scheme that satisfies the unconditional boundedness and decay of the free energy. Supposing that the cost of inverting a $N \times N$ matrix is $\mathcal{O}(N^3)$, it follows that the full 2D scheme has a computational complexity of $\mathcal{O}(N^6)$, as we need to invert a $N^2 \times N^2$ matrix. Alternatively, the cost of the dimensional-splitting scheme would be $\mathcal{O}(2N^4)$, since we need to invert a $N \times N$ matrix in each of the $2N$ rows and columns. As a result, there is a significant reduction of computational cost by employing the dimensional-splitting approach, and such cost is reduced even further in higher dimensions.

Remark 8.2.2 (Parallelization of the dimensional-splitting scheme in two and higher dimensions). The dimensional-splitting scheme in Section 8.2 can be fully parallelized in order to save computational time. This is possible due to two important observations: first, the scheme does not take notice of the order of updating the rows-columns, as long as all of them are updated; second, one row or column only depends on the values of the directly adjacent rows or columns, respectively. As a result, a strategy to parallelize the dimensional splitting scheme consists in updating at the same time all the odd rows-columns, as they do not depend on one another. At the same time the even rows-columns are also updated.

8.3   Numerical simulations

In this Section we test the applicability of our 1D and 2D schemes in a variety of configurations for the CH equation (7.1). But first, in Subsection 8.3.1 we test the spatial order of convergence of both schemes, showing that they satisfy a first-order spatial order of convergence. Subsequently, in Subsection 8.3.2 we present several test cases of 1D configurations. Finally, in Subsection 8.3.3 we focus on 2D simulations.

In all simulations the discrete free energy is computed as (8.8) and (8.22), for the 1D and 2D cases respectively. The full wetting boundary conditions are applied in Subsubsection 8.3.4, but for the rest of examples we assume $f_w(\phi, \beta) = 0$.

The code to reproduce all simulations is available in a Github repository [233]. The films of the 2D simulations are available in the figshare repository [234].

8.3.1   Order-of-convergence validation with explicit steady state

The selected test case was firstly proposed in [24], and it has subsequently served as a computational prototype adopted by many authors to validate their numerical schemes for the CH equation [166, 280]. The advantage of this test case is that an explicit steady state is known given particular initial conditions. Hence, once the steady state is reached, it can be readily utilized to measure the error of the numerical simulations for different mesh choices.

The original test case in [24] was formulated only for 1D simulations. Here, we appropriately extend it to two dimensions. We choose the logarithmic potential (7.9) in the deep quench limit with $\theta = 0$ and $\theta_c = 1$, so that $H_{log}(\phi) = (1 - \phi^2)/2$. The mobility satisfies the degenerate form in (7.3) and $\epsilon = 0.1$. The time step is $\Delta t = 0.0001$ for all simulations and we let the evolution run until $t = 0.1$.

For the 1D the spatial domain is $[0, 1]$, with the number of cells taken as 25, 50, 100 and 200 in successive simulations. For the 2D case the domain is $[0, 1] \times [0, 1]$ with the number of cells in each of the two axes as 10, 20, 40 and 80. When doubling the number of cells, the mesh size is halved and we can then compute the spatial order of the scheme as

$$\text{Order of the scheme} = \ln_2 \left(\frac{L^1 \operatorname{error}(\Delta x)}{L^1 \operatorname{error}(\Delta x/2)} \right),$$

with the L^1 error measured from the explicitly known steady state.

Following [24] we set the initial conditions of the simulation as

$$\phi_0(x, y) = \begin{cases} \cos\left(\frac{x - 1/2}{\epsilon}\right) - 1, & \text{if } \left|x - \frac{1}{2}\right| \leq \frac{\pi \epsilon}{2}, \\ -1, & \text{otherwise.} \end{cases}$$

Following [24] we set the initial conditions of the simulation as

$$\phi_0(x, y) = \begin{cases} \cos\left(\frac{x - 1/2}{\epsilon}\right) - 1, & \text{if } \left|x - \frac{1}{2}\right| \leq \frac{\pi \epsilon}{2}, \\ -1, & \text{otherwise,} \end{cases}$$

for both the 1D and 2D simulations. We notice that for the 1D simulation $y = 0\ \forall x$, while for the 2D simulation the phase-field is symmetric with respect to $x = 0.5$ (see Figure 8.3.2a for a depiction of the initial phase-field). With this choice of initial conditions the explicit steady state satisfies

$$\phi_{\text{steady}}(x, y) = \begin{cases} \frac{1}{\pi}\left[1 + \cos\left(\frac{x - 1/2}{\epsilon}\right)\right] - 1, & \text{if } \left|x - \frac{1}{2}\right| \leq \pi \epsilon, \\ -1, & \text{otherwise.} \end{cases} \tag{8.23}$$

In Table 8.3.1 we present the L^1 errors and spatial order of convergence for both the 1D and 2D tests. Clearly for both cases the spatial order of convergence stabilizes around 1 after increasing the number of cells. In Figure 8.3.1 we plot the temporal evolution of the 1D solution and the decay of the discrete free energy (8.8). For the latter we observe that the steady state is reached at $t \simeq 0.1$ as is evident from the plateau that the free energy at that time. In Figure 8.3.2 we display the initial and steady phase-fields of the 1D test. The 2D free-energy decay is omitted due to its similarity to the 1D behaviour in Figure 8.3.1b.

Number of cells	L^1 error	order		Number of cells	L^1 error	order
25	6.797E-03	-		10×10	1.112E-02	-
50	7.136E-04	3.25		20×20	6.752E-03	0.72
100	2.938E-04	1.28		40×40	2.278E-03	1.57
200	1.205E-05	1.29		80×80	7.736E-04	1.56

Table 8.3.1: Order-of-convergence test validating the first-order spatial accuracy of the 1D (left) and 2D (right) schemes. The errors are computed at $t = 0.1$ from the explicit solution (8.23) in Subsection 8.3.1

(a) Phase-field profile at different times

(b) Free-energy decay in log-x scale
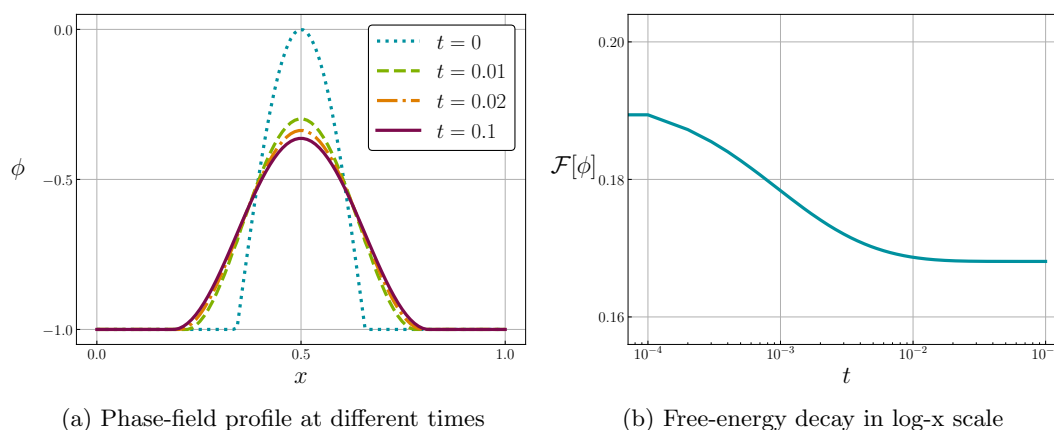
Figure 8.3.1: Temporal evolution for the one-dimensional test in Subsection 8.3.1.

8.3.2 One-dimensional simulations

In Example 8.3.1 we evolve a randomized initial phase-field to show the phase separation and coarsening for the Ginzburg-Landau double-well potential (7.8) and the logarithmic potential (7.9). Then in Example 8.3.2 we compare the impact of a constant mobility versus a degenerate mobility, for the choice of a logarithmic potential.

Example 8.3.1 (Phase separation in randomized initial phase-field). We focus on the behaviour of the CH equation (7.1) with the double-well potential (7.8) and with the logarithmic potential (7.9). As is well known, a process of phase separation with emerging clusters at $\phi = \pm 1$ takes place during the temporal evolution, and here we explore how the choice of potential affects the temporal scales and phase-field profiles at equilibrium. The initial phase-field is taken as $\phi_i(t = 0) = r$, with r being a random variable with uniform distribution in $[-0.5, 0.5]$. After
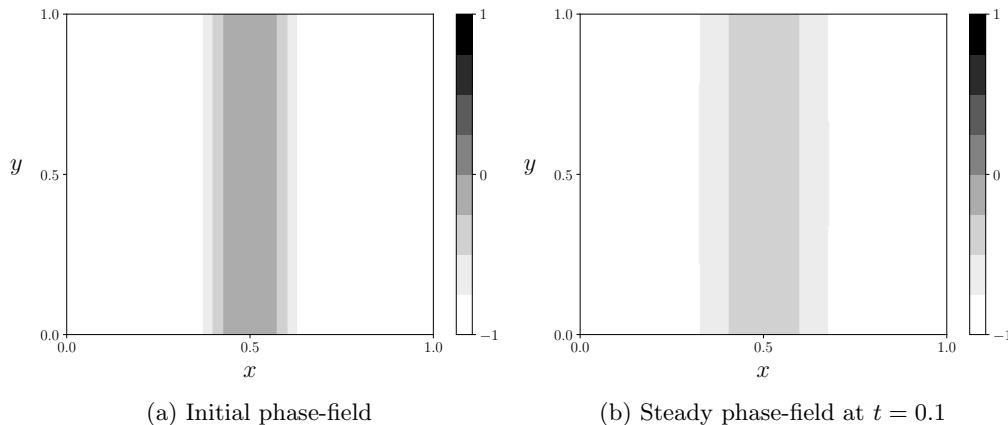
(a) Initial phase-field

(b) Steady phase-field at $t = 0.1$

Figure 8.3.2: Initial and steady phase-fields for the two-dimensional test in Subsection 8.3.1.

the sampling, we keep the same initial phase-field for the two simulations. The spatial domain is $x \in [-40, 40]$, and the number of cells is 200. The mobility is taken as the degenerate form in (7.3), and the interface parameter as $\epsilon = 1$. For the logarithmic potential we select as absolute temperature $\theta = 0.3$ and as critical temperature $\theta_c = 1$. The time step is $\Delta t = 0.01$ and we let the evolution run until $t = 30$. A similar simulation can be found in [163].

Figures 8.3.3a and 8.3.3b depict the temporal evolution of the phase-field, for the double-well and logarithmic potentials, respectively. There is a clear phase separation after starting the simulation with a randomized phase-field. Figures 8.3.3c and 8.3.3d provide a zoom of the density profile at $t = 30$, after the steady state has been reached. For the double-well potential, we observe a wider interface thickness with peaks lower than $|\phi| = 1$, while for the logarithmic potential the peaks form a plateau and have the value of $|\phi| = 1$. Lastly, in Figures 8.3.3e and 8.3.3f we illustrate the evolution of the discrete free energy (8.8), where we observe that the logarithmic potential takes slightly less time to reach the steady state.

Example 8.3.2 (Effect of mobility with logarithmic potential). The aim here is to test the effect of choosing a constant mobility $M = 1$ versus a degenerate mobility $M = (1 - \phi)(1 + \phi)$. This example was firstly proposed in [24] and later implemented in [280]. The scheme proposed in Section 8.1 satisfies the boundedness of the solution only for the degenerate mobility.

As initial condition for the simulation we select

$$\phi_0(x) = \begin{cases} 1, & \text{if } 0 \le x \le \frac{1}{3} - \frac{1}{20} \\ 20\left(\frac{1}{3} - x\right), & \text{if } \left|x - \frac{1}{3}\right| \le 20, \\ -20\left|x - \frac{41}{50}\right|, & \text{if } \left|x - \frac{41}{50}\right| \le \frac{1}{20}, \\ -1, & \text{otherwise.} \end{cases}$$

As potentials we choose the logarithmic potential (7.9) with $\theta = 0.3$ and $\theta_c = 1$. The spatial domain of the simulation is $x \in [0, 1]$, and the number of cells is 80. The time step is $\Delta t = 0.01$, $\epsilon = \sqrt{10^{-3}}$ and we let the evolution run until $t = 0.1$.

The results for the phase-field and free energy evolution are displayed in Figure 8.3.4. We notice that for the constant mobility $M = 1$ the bump at the right of the domain is quickly dissipated, while for the degenerate mobility it takes a much longer time. This happens because the bump is surrounded by a pure phase with $\phi = -1$, and the degenerate mobility is zero in pure phases. As a result, the mobility term is much higher for the constant mobility, and the exchange of mass occurs
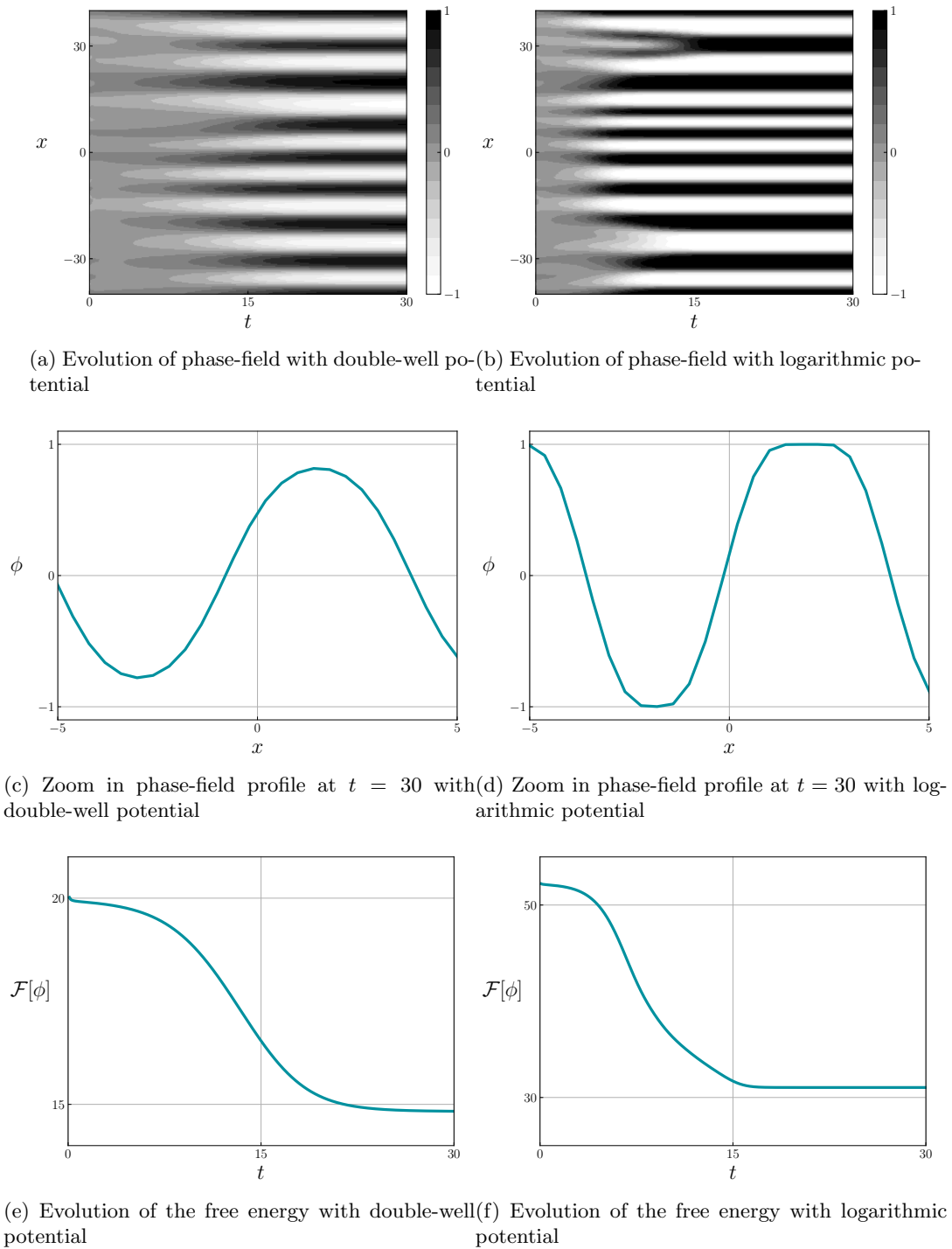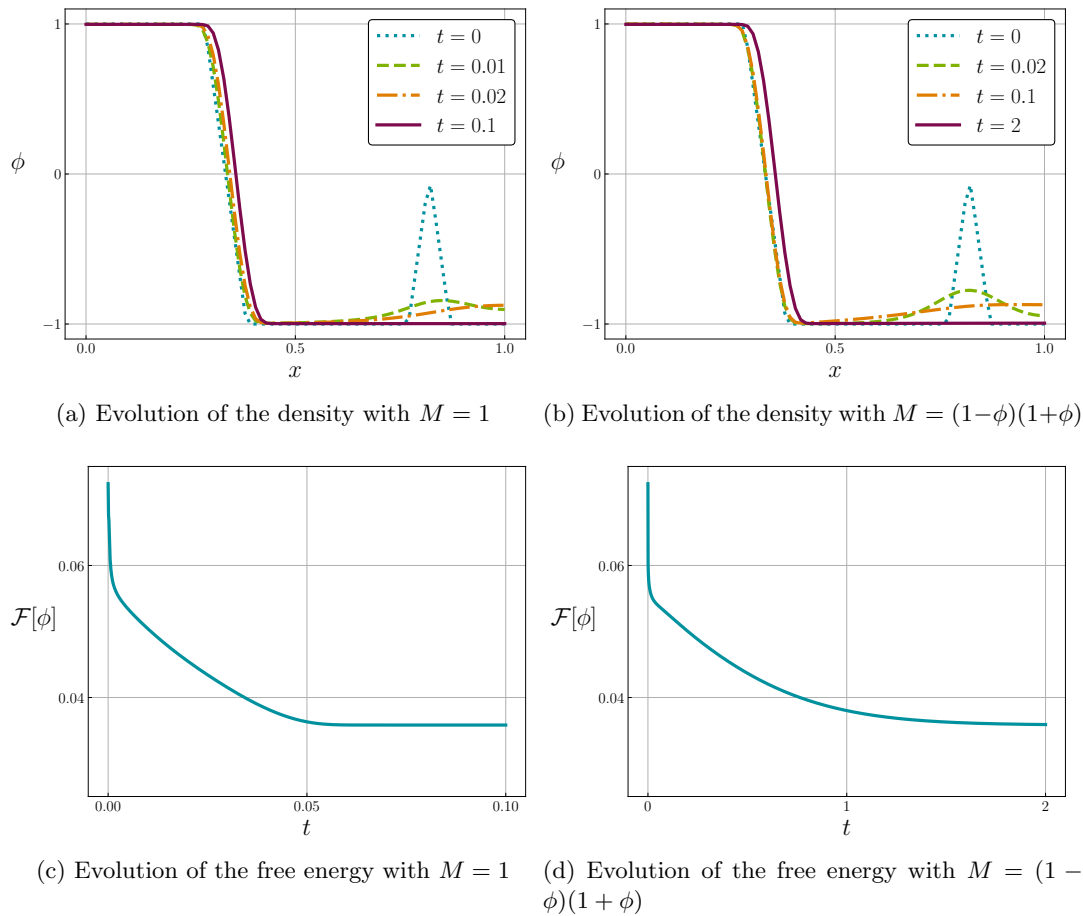
(a) Evolution of phase-field with double-well potential

(b) Evolution of phase-field with logarithmic potential

(c) Zoom in phase-field profile at $t = 30$ with double-well potential

(d) Zoom in phase-field profile at $t = 30$ with logarithmic potential

(e) Evolution of the free energy with double-well potential

(f) Evolution of the free energy with logarithmic potential

Figure 8.3.3: Temporal evolution for the initially-randomized phase-field in Example 8.3.1.

faster. In Figures 8.3.4c and 8.3.4d we depict the decay of the free energy during the evolution, and the timescale until the stationary solution is clearly much shorter for the constant mobility.

(a) Evolution of the density with $M = 1$

(b) Evolution of the density with $M = (1-\phi)(1+\phi)$

(c) Evolution of the free energy with $M = 1$

(d) Evolution of the free energy with $M = (1 - \phi)(1 + \phi)$

Figure 8.3.4: Temporal evolution for the logarithmic potential in Example 8.3.2.

8.3.3 Two-dimensional simulations

In Example 8.3.3 we analyse the impact of a constant versus a degenerate mobility when evolving a random initial phase-field. Then in Example 8.3.4 we analyse wetting phenomena with sessile droplets on a flat solid surface, by means on taking into account the wall free energy (7.6).

Example 8.3.3 (Phase separation with different mobilities). We choose the Ginzburg-Landau double-well potential in (7.8). It is a widely-employed potential for benchmarking numerical schemes of the CH equation, and we refer the reader to [24, 151, 280] for similar simulations.

As initial phase-field we select $\phi_0(x, y) = -0.4 + r$, with r being a random variable with uniform distribution in $[-0.25, 0.25]$. The 2D spatial domain of the simulation is chosen as $[-0.5, 0.5] \times [-0.5, 0.5]$ with 256 cells in each of the x and y directions. The mobility is taken as the degenerate form in (7.3), and the interface parameter as $\epsilon = 0.18$. The time step is $\Delta t = 0.0016$ and the evolution is monitored up to $t = 1$.

The results of the two simulations, for constant and degenerate mobility, are displayed in Figure 8.3.5, for three different snapshots of time. The decay of the corresponding free energy is depicted in Figure 8.3.6. For the case of constant mobility in Figures 8.3.5a, 8.3.5c and 8.3.5e we observe an initial stage of phase separation followed by a coarsening process with merging phases. From the free-energy plot in Figure 8.3.6 we remark the short time-scale of the phase separation in comparison to the long timescale of coarsening, which in turn leads to sudden decreases of free energy when distant phases merge. The log $-x$ plot in 8.3.6b illustrates the exponential decay of

the free energy during the short time-scale due to the initial phase separation from a randomised field. A similar phenomenon occurs in macroscopic systems with Morse-type interaction potentials [66], where two different timescales are also present: one for the attraction of phases and another one for the merging process. In the final snapshot at $t = 1$ only two phases remain in the domain.

The evolution with degenerate mobility is displayed in Figures 8.3.5b, 8.3.5d and 8.3.5f. The phase-separation stage is similar to the one with constant mobility, and as both simulations share the same initial condition, not surprisingly the first snapshots in Figures 8.3.5a and 8.3.5b look alike. The initial timescales also have the same order of magnitude, as displayed in the free energy plot in Figure 8.3.6. However, once the phases are formed there is little evolution in the system, due to the fact that the degenerate mobility is cancelled when $\phi = \pm 1$. As a consequence at $t = 1$ the phases have not yet merged, in sharp contrast to the case of constant mobility. In addition, the free energy reaches a much lower value for the case of constant mobility, a result of the merging of phases combined with reduction of interphases.

Example 8.3.4 (Contact angle in wetting phenomena). In this example we analyse wetting phenomena with sessile droplets on a flat solid substrate. In this application a fluid-fluid interface moves along the solid substrate, while a contact line is formed at the intersection between the interface and the substrate. Of special interest is the contact angle at the three-phase conjunction, which is determined by the wetting properties of the substrate.

For this setting the free energy in (7.2) takes into account the wetting effects at the boundary $\partial\Omega$ between the droplet and the substrate [47, 192], with $f_w(\phi)$ contributing to the wall component of the free energy. The form of the term $f_w(\phi)$ has received considerable attention in the literature, starting with early works using linear forms for liquid-gas problems [247]. Here we adopt the cubic form suggested in [13, 258, 259, 260] on the basis that it is the lowest-order polynomial to allow for the minimization of the wall free energy for the bulk densities, while at the same time preventing the formation of boundary layers. As noted in the introduction, the basic idea is that the contact line singularity can be removed without any additional physics. Details on how exactly this works, together with the associated asymptotic analysis are given in [259, 260].

Our choice for $f_w(\phi)$ satisfies (7.6), where β is the equilibrium contact angle, taken here as constant. Analogous modifications of the CH model have been applied in a variety of wetting scenarios including wetting of complex topographical geometries and chemically-heterogeneous substrates, with applications from microfluidic devices [238] to polymer films [179] and rheological systems [43].

In our first simulations we check that the equilibrium contact angle of sessile droplet on flat substrate obtained with our 2D scheme in Section 8.2 matches with the angle β imposed in the wall free-energy term $f_w(\phi, \beta)$. This is done for five choices of β: $\pi/3$, $5\pi/12$, $\pi/2$, $7\pi/12$ and $2\pi/3$. We point out that the choice of $\beta = \pi/2$ cancels $f_w(\phi, \beta)$ since $\cos\beta$ is multiplying the rest of the terms in (7.6).

As initial phase in the domain $[-0.5, 0.5] \times [0, 0.4]$ we select the semicircle

$$\phi_0(x, y) = \begin{cases} 0.97 & \text{if } x^2 + y^2 < 0.25^2; \\ -0.97 & \text{if } x^2 + y^2 \geq 0.25^2. \end{cases}$$

The number of cells both in each of the x and y directions is 256. The mobility takes the degenerate form in (7.3), and the interface parameter is chosen as $\epsilon = 0.005$. The time step is $\Delta t = 0.001$ and we follow the evolution until $t = 0.1$. We adopt the Ginzburg-Landau double-well potential in (7.8). The solid substrate boundary condition is imposed along $y = 0$. It should be noted that the initial phase-field is selected with values of ± 0.97 instead of ± 1 so that the degenerate mobility is not cancelled and the simulation can proceed uninterrupted. Similar simulations but with constant mobility were performed in [13].
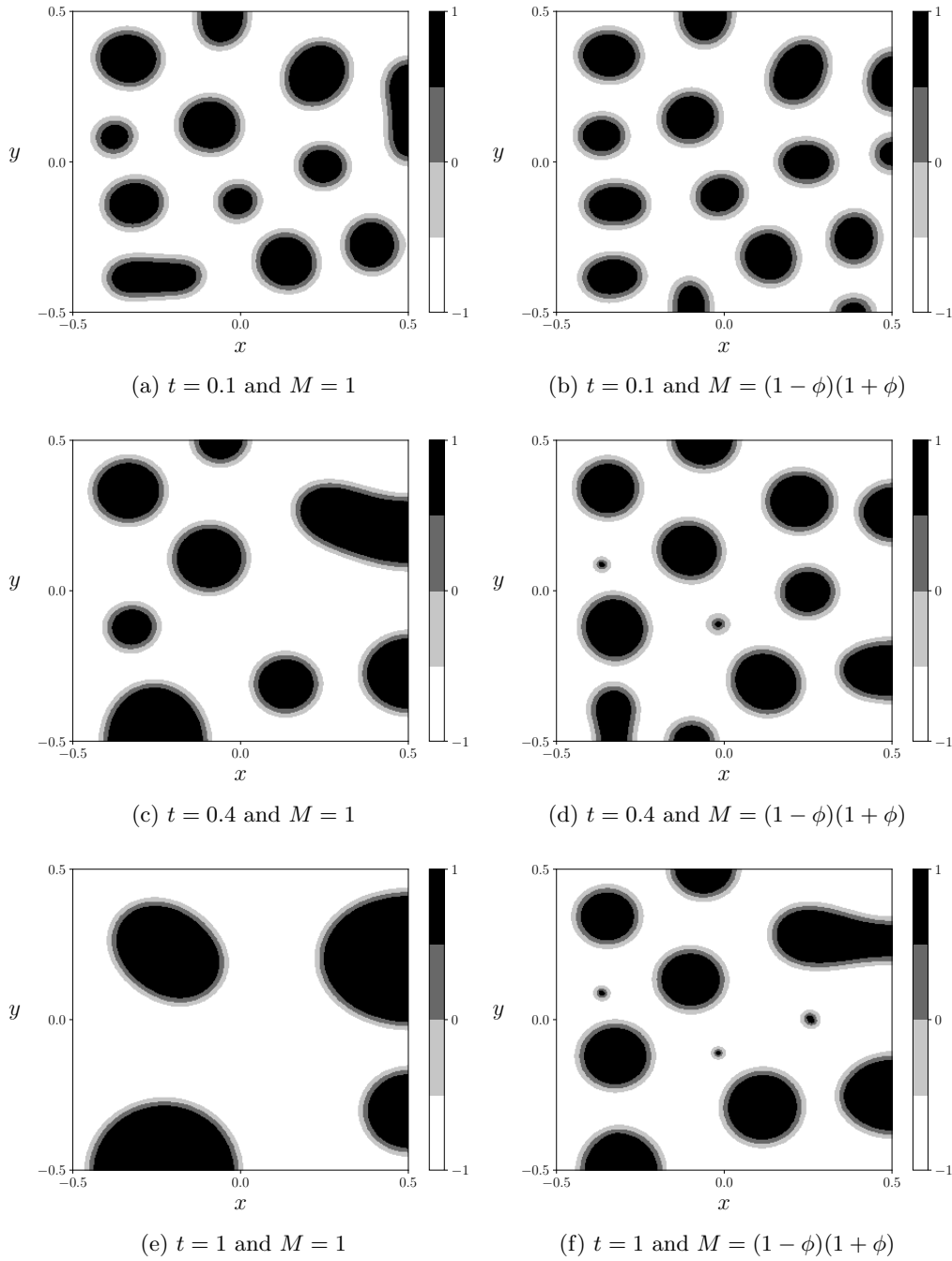
(a) $t = 0.1$ and $M = 1$

(b) $t = 0.1$ and $M = (1 - \phi)(1 + \phi)$

(c) $t = 0.4$ and $M = 1$

(d) $t = 0.4$ and $M = (1 - \phi)(1 + \phi)$

(e) $t = 1$ and $M = 1$

(f) $t = 1$ and $M = (1 - \phi)(1 + \phi)$

Figure 8.3.5: Temporal evolution for the initially-randomized phase-field in Example 8.3.3. Comparison of constant mobility (left) versus degenerate mobility (right).

The equilibrium phase-fields at $t = 0.1$ are displayed in Figures 8.3.7a to 8.3.7e, while the free-energy dissipation comparison is depicted in Figure 8.3.7f. The numerical contact angle $\hat{\beta}$ is close to the predefined angle β, and the quotient between the two is reported in the captions of Figures 8.3.7a to 8.3.7e. Such numerical angle is measured by employing the Drop Shape Analysis open-source software from [263]. From Figure 8.3.7f we observe that all the droplets stabilize after a short time of $\simeq 0.01$. It is also clear that the angle with the largest $\cos \beta$, which is $\beta = \pi/3$, has the greatest free energy overall due to the contribution of $f_w(\phi, \beta)$ in (7.6). On the contrary,

(a) Long-term evolution

(b) Log-x scale and fitting line
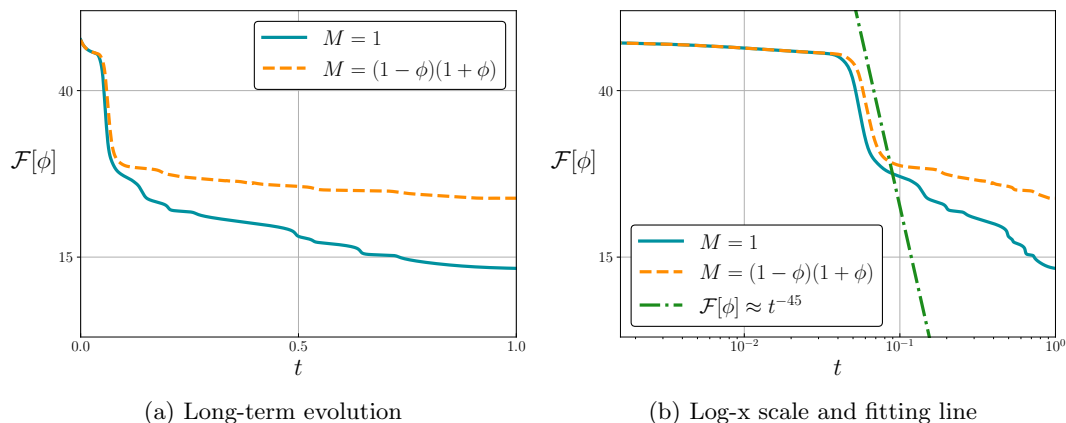
Figure 8.3.6: Free-energy decay in time comparing the constant-mobility and degenerate-mobility simulations in Example 8.3.3.

the angle $\beta = 2\pi/3$ has the lowest overall free energy. Interestingly, for the angle $\beta = \pi/2$ there is dynamical evolution even if the initial configuration already has the right contact angle. This is due to the fact that there are no interphases in the initial phase-field.

In the simulations that follow, we analyse the effect of contact angle on the merging of droplets on both hydrophilic ($\beta < \pi/2$) and hydrophobic ($\beta > \pi/2$) solid substrates. The objective is to show that on hydrophilic substrates droplets are prone to merging and, as a consequence, to forming a single phase; on the other hand, on hydrophobic substrates droplets tend to remain isolated.

As initial phase-field in the domain $[-1, 1] \times [0, 0.5]$ we select the two semicircles

$$\phi_0(x, y) = \begin{cases} 0.97 & \text{if } (x + 0.35)^2 + y^2 < 0.3^2 \text{ or } (x - 0.35)^2 + y^2 < 0.3^2; \\ -0.97 & \text{elsewhere.} \end{cases}$$

We now have 256 cells in the x direction and 64 in the y direction. The mobility assumes the degenerate form in (7.3), and the interface parameter has the value $\epsilon = 0.012$. The time step is $\Delta t = 0.0005$ and we follow the system up until $t = 15$. The employed contact angles are $\beta = \pi/4$ and $\beta = 3\pi/4$. The remaining parameters are identical to our previous simulations with single droplets.

The dynamic evolution of the droplets for the two choices of contact angle is detailed in Figures 8.3.8a to 8.3.8f, while the free-energy evolution can be found in 8.3.8g. From the plots with contact angle $\beta = \pi/4$ corresponding to a hydrophilic substrate, it can be seen that the two droplets coalesce and after $t = 15$ a single phase is clearly formed. On the contrary, the droplets with contact angle $\beta = 3\pi/4$ remain distant and do not merge during the dynamic evolution. From the free-energy plots in 8.3.8g we also find that the free energy reaches a plateau for both contact angles, meaning that the stationary state has been reached. We again employ a log $-x$ scale due to the significant free-energy decay at short timescales at the beginning of the simulation.
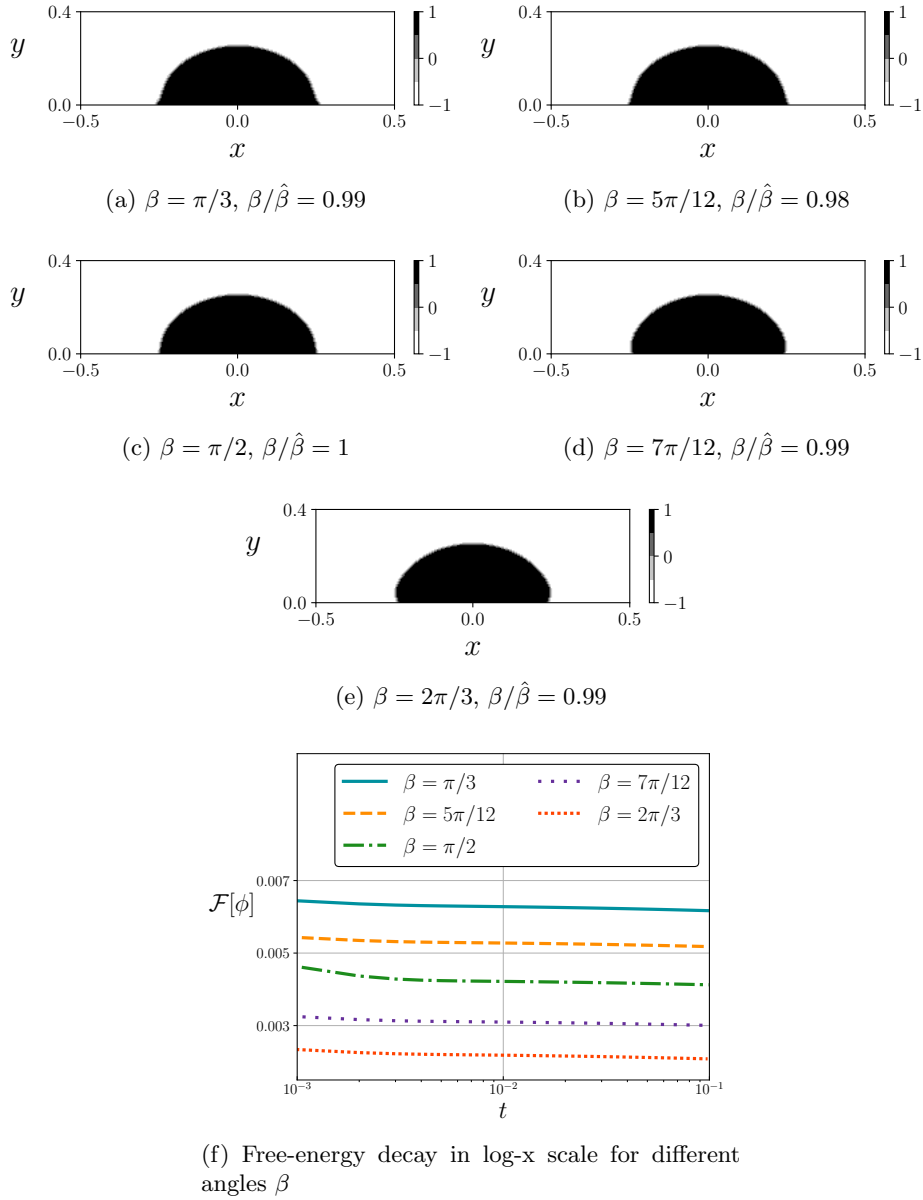
(a) $\beta = \pi/3$, $\beta/\hat{\beta} = 0.99$

(b) $\beta = 5\pi/12$, $\beta/\hat{\beta} = 0.98$

(c) $\beta = \pi/2$, $\beta/\hat{\beta} = 1$

(d) $\beta = 7\pi/12$, $\beta/\hat{\beta} = 0.99$

(e) $\beta = 2\pi/3$, $\beta/\hat{\beta} = 0.99$

(f) Free-energy decay in log-x scale for different angles β

Figure 8.3.7:   (a)-(e) Equilibrium phase-fields at $t = 2$ for different choices of angle β. (f) Dissipation of the free energy for different choices of angle β.
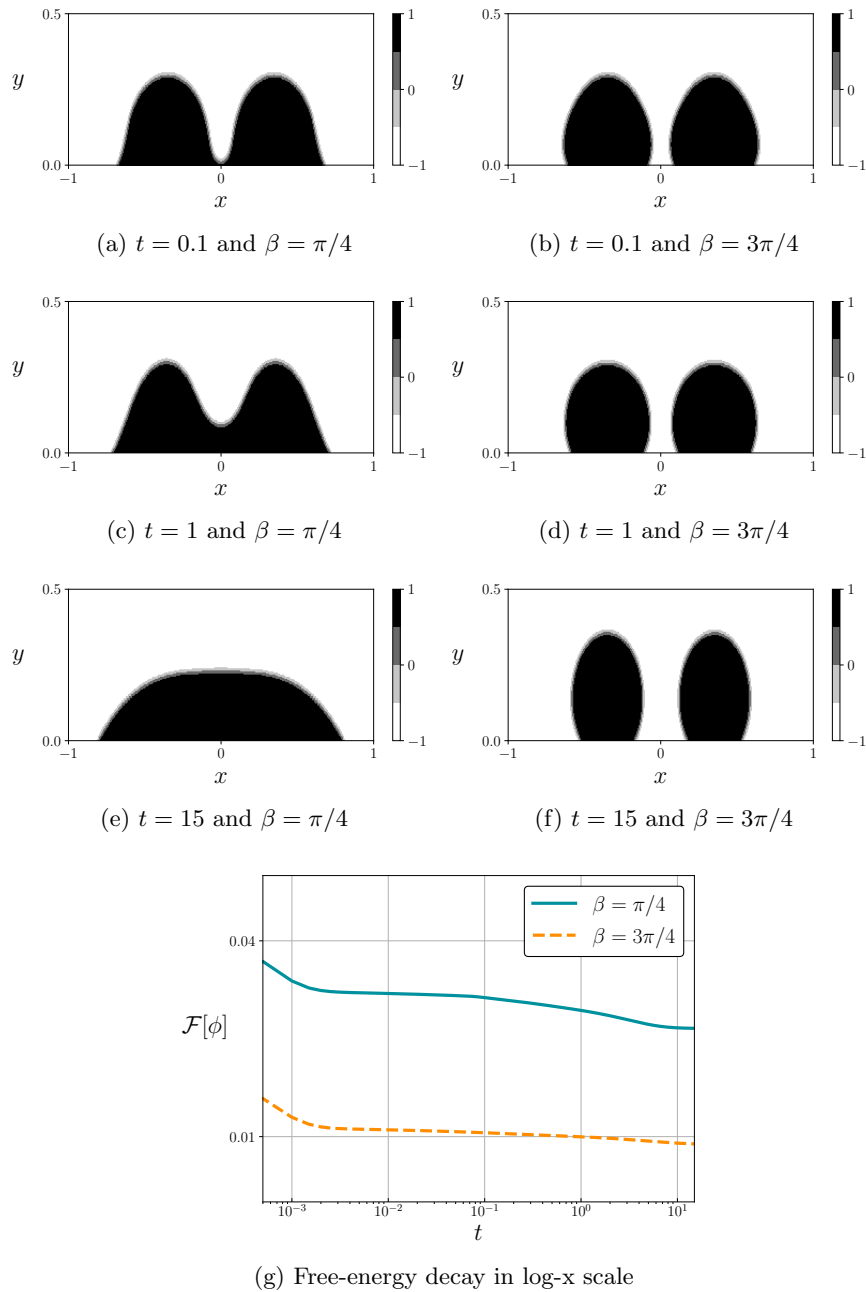
(a) $t = 0.1$ and $\beta = \pi/4$

(b) $t = 0.1$ and $\beta = 3\pi/4$

(c) $t = 1$ and $\beta = \pi/4$

(d) $t = 1$ and $\beta = 3\pi/4$

(e) $t = 15$ and $\beta = \pi/4$

(f) $t = 15$ and $\beta = 3\pi/4$

(g) Free-energy decay in log-x scale

Figure 8.3.8:   (a)-(f) Temporal evolution of two droplets with the hydrophilic angle $\beta = \pi/4$ (left) and the hydrophobic angle $\beta = 3\pi/4$ (right). (g) Comparison of free-energy decay for $\beta = \pi/4$ and $\beta = 3\pi/4$.

Appendix

8.A Dissipation of the discrete free energy in the dimensional-splitting scheme

Here we prove the dissipation of the discrete free energy in every row and every column for the dimensional-splitting 2D scheme of Section 8.2. Firstly, in Lemma 8.A.1 we prove the dissipation of the free energy in each x-direction iteration. Subsequently, in Lemma 8.A.2 we prove the dissipation of the free energy in each y-direction iteration. Both lemmas are necessary to show the global dissipation of the discrete free energy for the dimensional-splitting 2D scheme, as already highlighted in Subsection 8.2.1.

Lemma 8.A.1. *Dissipation of the discrete free energy in each x-direction iteration. Let the discrete free energy of* (7.2) *for a particular r be defined as:*

$$
\begin{aligned}
\mathcal{F}_\Delta\left(\phi^{n,r}\right) =& \Delta x \Delta y \sum_{i,j=1}^{n}\left(H_c(\phi_{i,j}^{n,r}) - H_e(\phi_{i,j}^{n,r})\right) \\
&+ \Delta x \Delta y \sum_{i=1}^{N-1}\sum_{j=1}^{N}\frac{\epsilon^2}{2}\left(\frac{\phi_{i+1,j}^{n,r} - \phi_{i,j}^{n,r}}{\Delta x}\right)^2 + \Delta x \Delta y \sum_{i=1}^{N}\sum_{j=1}^{N-1}\frac{\epsilon^2}{2}\left(\frac{\phi_{i,j+1}^{n,r} - \phi_{i,j}^{n,r}}{\Delta y}\right)^2 \quad (8.24) \\
&+ \Delta y \sum_{j=1}^{N}\left(f_w(\phi_{1,j}^{n,r}) + f_w(\phi_{N,j}^{n,r})\right) + \Delta x \sum_{i=1}^{N}\left(f_w(\phi_{i,1}^{n,r}) + f_w(\phi_{i,N}^{n,r})\right).
\end{aligned}
$$

It then follows that in the scheme (8.17) *the discrete free energy at every x-direction iteration, so that*

$$
\mathcal{F}_\Delta\left(\rho^{n,r}\right) - \mathcal{F}_\Delta\left(\rho^{n,r-1}\right) \leq 0.
$$

Proof. Multiply by $\xi_{i,r}^{n,r}$ in (8.17a) and sum up the result over the indices i and j, so that

$$
\sum_{i,j=1}^{N}\xi_{i,j}^{n,r}\left(\phi_{i,j}^{n,r} - \phi_{i,j}^{n,r-1}\right) = -\frac{\Delta t}{\Delta x}\sum_{i=1}^{N}\xi_{i,j}^{n,r}\left(F_{i+\frac{1}{2},j}^{n,r} - F_{i-\frac{1}{2},j}^{n,r}\right).
$$

By substituting in the previous equation the expression of $\xi_{i,j}^{n,r}$ in (8.17e), it follows that

$$
\begin{aligned}
\sum_{i,j=1}^{N}\left(\phi_{i,j}^{n,r} - \phi_{i,j}^{n,r-1}\right)\frac{\epsilon^2}{2}\left[(\Delta\phi)_{i,j}^{n,r-1} + (\Delta\phi)_{i,j}^{n,\star}\right] =& \frac{\Delta t}{\Delta x}\sum_{i=1}^{N}\xi_{i,j}^{n,r}\left(F_{i+\frac{1}{2},j}^{n,r} - F_{i-\frac{1}{2},j}^{n,r}\right) \\
&+ \sum_{i,j=1}^{N}\left(\phi_{i,j}^{n,r} - \phi_{i,j}^{n,r-1}\right)\left(H_c'(\phi_{i,j}^{n,r}) - H_e'(\phi_{i,j}^{n,r-1})\right)
\end{aligned}
$$

116

$$+ \frac{1}{\Delta x} \sum_{i,j=1}^{N} \left(\phi_{i,j}^{n,r} - \phi_{i,j}^{n,r-1} \right) W_{i,j}^{x}(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1})$$

$$+ \frac{1}{\Delta y} \sum_{i,j=1}^{N} \left(\phi_{i,j}^{n,r} - \phi_{i,j}^{n,r-1} \right) W_{i,j}^{y}(\phi_{i,j}^{n,r}, \phi_{i,j}^{n,r-1}). \tag{8.25}$$

Considering now the definition of the discrete free energy in (8.24), the decay between the x-direction iterations at r and $r-1$ yields

$$\frac{\mathcal{F}_\Delta\left(\phi^{n,r}\right) - \mathcal{F}_\Delta\left(\phi^{n,r-1}\right)}{\Delta x \Delta y} = \sum_{i,j=1}^{N} \left(H_c(\phi_{i,j}^{n,r}) - H_c(\phi_{i,j}^{n,r-1}) \right) - \sum_{i,j=1}^{N} \left(H_e(\phi_{i,j}^{n,r}) - H_e(\phi_{i,j}^{n,r-1}) \right)$$

$$+ \sum_{i=1}^{N-1} \sum_{j=1}^{N} \frac{\epsilon^2}{2} \left(\frac{\left(\phi_{i+1,j}^{n,r} - \phi_{i,j}^{n,r} \right)^2}{\Delta x^2} - \frac{\left(\phi_{i+1,j}^{n,r-1} - \phi_{i,j}^{n,r-1} \right)^2}{\Delta x^2} \right)$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N-1} \frac{\epsilon^2}{2} \left(\frac{\left(\phi_{i,j+1}^{n,r} - \phi_{i,j}^{n,r} \right)^2}{\Delta y^2} - \frac{\left(\phi_{i,j+1}^{n,r-1} - \phi_{i,j}^{n,r-1} \right)^2}{\Delta y^2} \right) \tag{8.26}$$

$$+ \frac{1}{\Delta x} \sum_{j=1}^{N} \left(f_w(\phi_{1,j}^{n,r}) - f_w(\phi_{1,j}^{n,r-1}) + f_w(\phi_{N,j}^{n,r}) - f_w(\phi_{N,j}^{n,r-1}) \right)$$

$$+ \frac{1}{\Delta y} \sum_{i=1}^{N} \left(f_w(\phi_{i,1}^{n,r}) - f_w(\phi_{i,1}^{n,r-1}) + f_w(\phi_{i,N}^{n,r}) - f_w(\phi_{i,N}^{n,r-1}) \right).$$

The next step is to expand the sums with the discrete gradients at the interfaces, which correspond to the second and third line of the expression in (8.26). For this we follow the same strategy as with the 1D proof in (8.14): firstly, we apply summation by parts in the x direction (for the sum depending on Δx^2) and in the y direction (for the sum depending on Δy^2); secondly, the boundary conditions in (8.20) are taken into account; and thirdly, we apply the definition of the discrete Laplacians (8.17f) and (8.17g) bearing in mind that $\phi_{i,j+1}^{n,r} = \phi_{i,j+1}^{n,r-1}$ and $\phi_{i,j-1}^{n,r} = \phi_{i,j-1}^{n,r-1}$ when $j \neq r$.

Such expansion leads to

$$\sum_{i=1}^{N-1} \sum_{j=1}^{N} \frac{\epsilon^2}{2} \left(\frac{\left(\phi_{i+1,j}^{n,r} - \phi_{i,j}^{n,r} \right)^2}{\Delta x^2} - \frac{\left(\phi_{i+1,j}^{n,r-1} - \phi_{i,j}^{n,r-1} \right)^2}{\Delta x^2} \right)$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N-1} \frac{\epsilon^2}{2} \left(\frac{\left(\phi_{i,j+1}^{n,r} - \phi_{i,j}^{n,r} \right)^2}{\Delta y^2} - \frac{\left(\phi_{i,j+1}^{n,r-1} - \phi_{i,j}^{n,r-1} \right)^2}{\Delta y^2} \right)$$

$$= \sum_{i=1}^{N-1} \sum_{j=1}^{N} \frac{\epsilon^2}{2} \left(\frac{\phi_{i+1,j}^{n,r} - \phi_{i,j}^{n,r}}{\Delta x} + \frac{\phi_{i+1,j}^{n,r-1} - \phi_{i,j}^{n,r-1}}{\Delta x} \right) \left(\frac{\phi_{i+1,j}^{n,r} - \phi_{i,j}^{n,r}}{\Delta x} - \frac{\phi_{i+1,j}^{n,r-1} - \phi_{i,j}^{n,r-1}}{\Delta x} \right)$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N-1} \frac{\epsilon^2}{2} \left(\frac{\phi_{i,j+1}^{n,r} - \phi_{i,j}^{n,r}}{\Delta y} + \frac{\phi_{i,j+1}^{n,r-1} - \phi_{i,j}^{n,r-1}}{\Delta y} \right) \left(\frac{\phi_{i,j+1}^{n,r} - \phi_{i,j}^{n,r}}{\Delta y} - \frac{\phi_{i,j+1}^{n,r-1} - \phi_{i,j}^{n,r-1}}{\Delta y} \right)$$

$$= -\sum_{i=2}^{N-1}\sum_{j=1}^{N}\frac{\epsilon^2}{2}\left(\phi_{i,j}^{n,r}-\phi_{i,j}^{n,r-1}\right)\left(\frac{\phi_{i+1,j}^{n,r}-2\phi_{i,j}^{n,r}+\phi_{i-1,j}^{n,r}}{\Delta x^2}+\frac{\phi_{i+1,j}^{n,r-1}-2\phi_{i,j}^{n,r-1}+\phi_{i-1,j}^{n,r-1}}{\Delta x^2}\right)$$

$$+\sum_{j=1}^{N}\frac{\epsilon^2}{2}\left(\phi_{N,j}^{n,r}-\phi_{N,j}^{n,r-1}\right)\left(\frac{\phi_{N,j}^{n,r}-\phi_{N-1,j}^{n,r}}{\Delta x^2}+\frac{\phi_{N,j}^{n,r-1}-\phi_{N-1,j}^{n,r-1}}{\Delta x^2}\right)$$

$$-\sum_{j=1}^{N}\frac{\epsilon^2}{2}\left(\phi_{1,j}^{n,r}-\phi_{1,j}^{n,r-1}\right)\left(\frac{\phi_{2,j}^{n,r}-\phi_{1,j}^{n,r}}{\Delta x^2}+\frac{\phi_{2,j}^{n,r-1}-\phi_{1,j}^{n,r-1}}{\Delta x^2}\right)$$

$$-\sum_{i=1}^{N}\sum_{j=2}^{N-1}\frac{\epsilon^2}{2}\left(\phi_{i,j}^{n,r}-\phi_{i,j}^{n,r-1}\right)\left(\frac{\phi_{i,j+1}^{n,r}-2\phi_{i,j}^{n,r}+\phi_{i,j-1}^{n,r}}{\Delta y^2}+\frac{\phi_{i,j+1}^{n,r-1}-2\phi_{i,j}^{n,r-1}+\phi_{i,j-1}^{n,r-1}}{\Delta y^2}\right) \quad (8.27)$$

$$+\sum_{i=1}^{N}\frac{\epsilon^2}{2}\left(\phi_{i,N}^{n,r}-\phi_{i,N}^{n,r-1}\right)\left(\frac{\phi_{i,N}^{n,r}-\phi_{i,N-1}^{n,r}}{\Delta y^2}+\frac{\phi_{i,N}^{n,r-1}-\phi_{i,N-1}^{n,r-1}}{\Delta y^2}\right)$$

$$-\sum_{i=1}^{N}\frac{\epsilon^2}{2}\left(\phi_{i,1}^{n,r}-\phi_{i,1}^{n,r-1}\right)\left(\frac{\phi_{i,2}^{n,r}-\phi_{i,1}^{n,r}}{\Delta y^2}+\frac{\phi_{i,2}^{n,r-1}-\phi_{i,1}^{n,r-1}}{\Delta y^2}\right)$$

$$= -\sum_{i,j=1}^{N}\frac{\epsilon^2}{2}\left(\phi_{i,j}^{n,r}-\phi_{i,j}^{n,r-1}\right)\left[(\Delta\phi)_{i,j}^{n,\star}+(\Delta\phi)_{i,j}^{n,r-1}\right].$$

where $\phi_{i,j-1}^{n,r}=\phi_{i,j-1}^{n,r-1}$ and $\phi_{i,j+1}^{n,r}=\phi_{i,j+1}^{n,r-1}$ due to (8.17a). The resultant expression (8.27) matches with the left-hand side in (8.25). The next step is to substitute (8.25) in (8.26), but at the same time recall that $W_{i,j}^x=0$ unless $i=1,N$ due to (8.18), and $W_{i,j}^y=0$ unless $j=1,N$ due to (8.19). It follows that

$$\frac{\mathcal{F}_\Delta\left(\phi^{n,r}\right)-\mathcal{F}_\Delta\left(\phi^{n,r-1}\right)}{\Delta x\Delta y}=\sum_{i,j=1}^{N}\left(H_c(\phi_{i,j}^{n,r})-H_c(\phi_{i,j}^{n,r-1})-H_c'(\phi_{i,j}^{n,r})\left(\phi_{i,j}^{n,r}-\phi_{i,j}^{n,r-1}\right)\right)$$

$$-\sum_{i,j=1}^{N}\left(H_e(\phi_{i,j}^{n,r})-H_e(\phi_{i,j}^{n,r-1})-H_e'(\phi_{i,j}^{n,r-1})\left(\phi_{i,j}^{n,r}-\phi_{i,j}^{n,r-1}\right)\right)$$

$$+\frac{1}{\Delta x}\sum_{j=1}^{N}\left[\left(f_w(\phi_{1,j}^{n,r})-f_w(\phi_{1,j}^{n,r-1})\right)-\left(\phi_{1,j}^{n,r}-\phi_{1,j}^{n,r-1}\right)W_{1,j}^x(\phi_{1,j}^{n,r},\phi_{1,j}^{n,r-1})\right]$$

$$+\frac{1}{\Delta x}\sum_{j=1}^{N}\left[\left(f_w(\phi_{N,j}^{n,r})-f_w(\phi_{N,j}^{n,r-1})\right)-\left(\phi_{N,j}^{n,r}-\phi_{N,j}^{n,r-1}\right)W_{N,j}^x(\phi_{N,j}^{n,r},\phi_{N,j}^{n,r-1})\right]$$

$$+\frac{1}{\Delta y}\sum_{i=1}^{N}\left[f_w(\phi_{i,1}^{n,r})-f_w(\phi_{i,1}^{n,r-1})-\left(\phi_{i,1}^{n,r}-\phi_{i,1}^{n,r-1}\right)W_{i,1}^y(\phi_{i,1}^{n,r},\phi_{i,1}^{n,r-1})\right]$$

$$+\frac{1}{\Delta y}\sum_{i=1}^{N}\left[\left(f_w(\phi_{i,N}^{n,r})-f_w(\phi_{i,N}^{n,r-1})\right)-\left(\phi_{i,N}^{n,r}-\phi_{i,N}^{n,r-1}\right)W_{i,N}^y(\phi_{i,N}^{n,r},\phi_{i,N}^{n,r-1})\right]$$

$$-\frac{\Delta t}{\Delta x}\sum_{i=1}^{N}\xi_{i,j}^{n,r}\left(F_{i+\frac{1}{2},j}^{n,r}-F_{i-\frac{1}{2},j}^{n,r}\right).$$

Following now the same reasoning as with the proof of the 1D case in Section 8.1, we get that the discrete free-energy dissipation in (8.24) is satisfied. $\qquad\square$

Lemma 8.A.2. *Dissipation of the discrete free energy in each y-direction iteration. Let the discrete free energy of* (7.2) *for a particular c be defined as:*

$$
\begin{aligned}
\mathcal{F}_\Delta\left(\phi^{n,c}\right) =& \Delta x \Delta y \sum_{i,j=1}^{n}\left(H_c(\phi_{i,j}^{n,c}) - H_e(\phi_{i,j}^{n,c})\right) \\
&+ \Delta x \Delta y \sum_{i=1}^{N-1}\sum_{j=1}^{N}\frac{\epsilon^2}{2}\left(\frac{\phi_{i+1,j}^{n,c}-\phi_{i,j}^{n,c}}{\Delta x}\right)^2 + \Delta x \Delta y \sum_{i=1}^{N}\sum_{j=1}^{N-1}\frac{\epsilon^2}{2}\left(\frac{\phi_{i,j+1}^{n,c}-\phi_{i,j}^{n,c}}{\Delta y}\right)^2 \\
&+ \Delta y \sum_{j=1}^{N}\left(f_w(\phi_{1,j}^{n,c}) + f_w(\phi_{N,j}^{n,c})\right) + \Delta x \sum_{i=1}^{N}\left(f_w(\phi_{i,1}^{n,c}) + f_w(\phi_{i,N}^{n,c})\right).
\end{aligned}
\tag{8.28}
$$

It then follows that for the scheme (8.21) *the discrete free energy decays at every y-direction iteration, which in turn implies*

$$
\mathcal{F}_\Delta\left(\rho^{n,c}\right) - \mathcal{F}_\Delta\left(\rho^{n,c-1}\right) \leq 0.
$$

Proof. The proof follows the same steps as in lemma 8.A.1, and after some algebra, we obtain that the dissipation of discrete free energy (8.28) between the y-direction iterations at c and $c-1$ satisfies

$$
\begin{aligned}
\frac{\mathcal{F}_\Delta\left(\phi^{n,c}\right) - \mathcal{F}_\Delta\left(\phi^{n,c-1}\right)}{\Delta x \Delta y} =& \sum_{i,j=1}^{N}\left(H_c(\phi_{i,j}^{n,c}) - H_c(\phi_{i,j}^{n,c-1}) - H_c'(\phi_{i,j}^{n,c})\left(\phi_{i,j}^{n,c}-\phi_{i,j}^{n,c-1}\right)\right) \\
&- \sum_{i,j=1}^{N}\left(H_e(\phi_{i,j}^{n,c}) - H_e(\phi_{i,j}^{n,c-1}) - H_e'(\phi_{i,j}^{n,c-1})\left(\phi_{i,j}^{n,c}-\phi_{i,j}^{n,c-1}\right)\right) \\
&+ \frac{1}{\Delta x}\sum_{j=1}^{N}\left[\left(f_w(\phi_{1,j}^{n,c}) - f_w(\phi_{1,j}^{n,c-1})\right) - \left(\phi_{1,j}^{n,c}-\phi_{1,j}^{n,c-1}\right)W_{1,j}^x(\phi_{1,j}^{n,c},\phi_{1,j}^{n,c-1})\right] \\
&+ \frac{1}{\Delta x}\sum_{j=1}^{N}\left[\left(f_w(\phi_{N,j}^{n,c}) - f_w(\phi_{N,j}^{n,c-1})\right) - \left(\phi_{N,j}^{n,c}-\phi_{N,j}^{n,c-1}\right)W_{N,j}^x(\phi_{N,j}^{n,c},\phi_{N,j}^{n,c-1})\right] \\
&+ \frac{1}{\Delta y}\sum_{i=1}^{N}\left[f_w(\phi_{i,1}^{n,c}) - f_w(\phi_{i,1}^{n,c-1}) - \left(\phi_{i,1}^{n,c}-\phi_{i,1}^{n,c-1}\right)W_{i,1}^y(\phi_{i,1}^{n,c},\phi_{i,1}^{n,c-1})\right] \\
&+ \frac{1}{\Delta y}\sum_{i=1}^{N}\left[\left(f_w(\phi_{i,N}^{n,c}) - f_w(\phi_{i,N}^{n,c-1})\right) - \left(\phi_{i,N}^{n,c}-\phi_{i,N}^{n,c-1}\right)W_{i,N}^y(\phi_{i,N}^{n,c},\phi_{i,N}^{n,c-1})\right] \\
&- \frac{\Delta t}{\Delta y}\sum_{i,j=1}^{N}\xi_{i,j}^{n,c}\left(F_{i+\frac{1}{2},j}^{n,c} - F_{i-\frac{1}{2},j}^{n,c}\right).
\end{aligned}
$$

Following once again the same reasoning as with the proof of the 1D case in Section 8.1, proves that the discrete free-energy dissipation in (8.28) is satisfied. $\qquad\square$

An application of the Cahn-Hilliard equation in image inpainting

The contents of this chapter are based on the publication [65], done in collaboration with José A. Carrillo, Serafim Kalliadasis, and Fuyue Liang:

[65]   J. A. Carrillo, S. Kalliadasis, F. Liang, and <u>S. P. Perez</u>. "Enhancement of damaged-image prediction through Cahn-Hilliard Image Inpainting". In: *Royal Society Open Science* 8 (2021), 201294.

9.1   Introduction to image inpainting

Image inpainting consists in filling damaged or missing areas of an image, with the ultimate objective of restoring it and making it appear as the true and original image. There are multiple applications of image inpainting, ranging from restoration of the missing areas of oil paintings and removal scratches in photographs to noisy MRI scans and blurred satellite images of the earth. Manual image inpainting techniques have been employed for many centuries by art conservators and professional restorers, but it was not until the turn of the 21$^{\text{st}}$ century that digital image inpainting models based on PDEs and variational methods were introduced [31, 72, 209]. These methods are usually referred to as non-texture, geometrical or structural inpainting since they focus on restoring the structural information in the inpainted domain such as edges, corners or curvatures. This is done by performing an image interpolation of the damaged areas based on the information collected from the surrounding environment only, leading to appealing images for the human vision system. On the contrary, texture inpainting is based on recovering global patterns of the image for the inpainted region [101], and a popular tool in this category is the exemplar-based inpainting methods[100, 193]. Associated with these developments, a field that has gained a lot of traction in recent years is the so-called generative image inpainting, where deep learning based approaches have proven to be successful even for blind inpainting in which the inpaited region is not provided a priori [281, 296, 297]. In this work we focus on non-texture image inpainting methods based on PDEs, and we refer the reader to Ref. [245] for a general review of the topic.

    There have been multiple PDE models for image inpainting proposed since the initial work of

Bertalmio *et al.* [31] nearly 20 years ago. Their trailblazing model is able to propagate isotopes, i.e. contours of uniform greyscale image intensity, through the inpainted region, a common technique employed by museum artists in restoration. As it also turns out that the original model bears close connection to fluid dynamics through the Navier-Stokes equation with the image intensity function acting as the stream function [30]. Another fluid dynamic equation that has played a pivotal role in image inpainting is the Cahn-Hilliard (CH) equation, initially proposed in [48] for phase separation in binary alloys.

The CH equation was firstly proposed in the context of image inpainting by [33]. Specifically, the authors adopted a modified CH equation for binary images with inpainting quality as accurate as the state-of-art inpainting models but with a much faster computational speed taking advantage of the efficient computational techniques already available for the CH equation [128, 274]. Since then several authors have extended the applicability of the CH equation in the field of image inpainting, for instance by taking into account greyvalue images [45, 89], nonsmooth potentials instead of the double-well potential (7.8) [40], and considering colour image inpainting [88]. The modified CH equation in [33] introduces a fidelity term $\lambda(\boldsymbol{x})$ to avoid modifying the original image outside of the inpainted region D, and the CH equation in (7.1) becomes

$$\frac{\partial \phi(\boldsymbol{x}, t)}{\partial t} = -\nabla^2 \left(\epsilon^2 \nabla^2 \phi - H'(\phi) \right) + \lambda(\boldsymbol{x}) \left(\phi(\boldsymbol{x}, t = 0) - \phi \right), \tag{9.1}$$

where

$$\lambda(\mathbf{x}) = \begin{cases} 0 & \text{if} \quad \boldsymbol{x} \in D, \\ \lambda_0 & \text{if} \quad \boldsymbol{x} \notin D, \end{cases} \tag{9.2}$$

and $\phi(\boldsymbol{x}, t = 0)$ refers to the original damaged image. The parameter ϵ plays a similar role as in the original CH equation, and here it is related to the interface between the two phases or colours presented in the image. The fidelity term $\lambda(\boldsymbol{x})$ is nonzero only in the pixels of the image that are outside the inpainted domain: the objective is to preserve those pixels in their initial values, and that's why the term $\lambda(\boldsymbol{x}) \left(\phi(\boldsymbol{x}, t = 0) - \phi \right)$ is added to introduce a penalty proportional to the difference with respect to the initial phase fields. The two parameters ϵ and λ_0 are essential to achieve an adequate image inpainting outcome, and it is usually necessary to iterate until finding appropriate tunings for their values, which typically depend on the image specifications. Furthermore, the new modified CH equation (9.1) is not stricly a gradient flow: although the original CH equation satisfies a gradient-flow structure under an H^{-1} norm and the fidelity term in (9.1) can be derived from a gradient flow under an L^2 norm, the combined modified CH equation is neither a gradient flow in H^{-1} nor L^2.

9.2 Integrated algorithm with image inpainting and pattern recognition

We detail the construction of an integrated algorithm that firstly applies image inpainting and subsequently conducts pattern recognition for the restored image. In Subsection 9.2.1 we begin by presenting the finite-volume scheme employed to solve the modified CH equation, based on the work in Chapter 8. Then in Subsection 9.2.2 we illustrate the two-step method for image inpainting, based on tuning the parameters ϵ and λ_0 of the modified CH equation. In Subsection 9.2.3 we present the neural network employed for pattern recognition, detailing its architecture and training parameters. Finally, in Subsection 9.2.4 we gather all previous elements to formulate an integrated algorithm for prediction with an image inpainting filter.

9.2.1 2D finite-volume scheme for the modified CH equation

We summarise the 2D finite-volume scheme constructed for the original CH equation in Chapter 8. This scheme satisfies an unconditional decay of the discrete free energy of the original CH equation in (7.2), for no matter what choice of the time step, and for specific choice of mobility $M(\phi) = 1 - \phi^2$ it ensures the unconditional boundedness of the phase field. The scheme can be straightforwardly extended to the modified CH equation in (9.1) proposed in [33] as we show here. Even though the modified CH equation does not possess some of the properties of the original CH equation, such as the gradient-flow structure, our finite-volume scheme preserves its robustness for all the image-inpainting test cases presented in this work. In Remark 9.2.1 we explain how to choose the time step and the mesh size, and in Remark 9.2.2 we detail how to turn the scheme into a dimensional splitting one, with promising applicability in high-dimensional images such as medical ones. We refer the reader to [17] for further details about dimensional-splitting schemes.

For simplicity, let's rewrite (9.1) in 2D by introducing $\boldsymbol{u} = (v, w)$ as the physical velocity term and ξ as the variation of the free energy with respect to the density,

$$\frac{\partial \phi(x, y, t)}{\partial t} = -\nabla \cdot \boldsymbol{u} + \lambda(x, y)\left(\phi(x, y, t = 0) - \phi(x, y, t)\right), \tag{9.3}$$

where

$$\xi = \frac{\delta \mathcal{F}[\phi]}{\delta \phi} = \epsilon^2 \nabla^2 \phi - H'(\phi), \quad \boldsymbol{u} = \nabla \xi, \quad v = \frac{\partial \xi}{\partial x}, \quad \text{and} \quad w = \frac{\partial \xi}{\partial y}. \tag{9.4}$$

For the finite-volume formulation we begin by dividing the computational domain $[0, L] \times [0, L]$ in $N \times N$ cells $C_{i,j} := [x_{i-1/2}, x_{i+1/2}] \times [y_{j-1/2}, y_{j+1/2}]$, all with uniform size $\Delta x \Delta y$ so that $x_{i+1/2} - x_{i-1/2} = \Delta x$ and $y_{j+1/2} - y_{j-1/2} = \Delta y$. The time step is denoted as Δt. In each of the cells we define the cell average $\phi_{i,j}^n$ at time $t = n\Delta t$ as

$$\phi_{i,j}^n = \frac{1}{\Delta x \Delta y} \int_{C_{i,j}} \phi(x, y, t = n\Delta t)\, dx\, dy, \tag{9.5}$$

where $\phi_{i,j}^0$ is the phase field at $t = 0$, which corresponds to the normalized pixel intensities of the initial damaged image to be inpainted.

The finite-volume scheme is then derived by integrating the modified CH equation (9.3) over each of the cells $C_{i,j}$ of the domain, leading to an implicit formulation satisfying

$$\frac{\phi_{i,j}^{n+1} - \phi_{i,j}^n}{\Delta t} = -\frac{F_{i+\frac{1}{2},j}^{n+1} - F_{i-\frac{1}{2},j}^{n+1}}{\Delta x} - \frac{G_{i,j+\frac{1}{2}}^{n+1} - G_{i,j-\frac{1}{2}}^{n+1}}{\Delta y} + \lambda_{i,j}(\phi_{i,j}^0 - \phi_{i,j}^n), \tag{9.6}$$

with $F_{i+1/2,j}^{n+1}$ and $G_{i,j+1/2}^{n+1}$ being flux approximations at the boundaries, and $\lambda_{i,j}$ the discrete version of $\lambda(\boldsymbol{x})$ in (9.2) satisfying

$$\lambda_{i,j} = \begin{cases} 0 & \text{if } (x_i, y_j) \in D, \\ \lambda_0 & \text{if } (x_i, y_j) \notin D, \end{cases} \tag{9.7}$$

with D being the inpainted domain where the image damage is located. The inpainted domain has to be determined beforehand for each image, and is formed by those finite-volume cells to be repaired during the image inpainting.

The flux terms $F_{i+1/2,j}^{n+1}$ and $G_{i,j+1/2}^{n+1}$ are the discrete approximations of the velocity components v and w at the cell interfaces $(x_{i+1/2}, y_j)$ and $(x_i, y_{j+1/2})$, respectively. Such approximations follow an upwind and implicit approach inspired by the works in Refs [15, 16, 34, 55], satisfying

$$\begin{aligned} F_{i+\frac{1}{2},j}^{n+1} &= \left(v_{i+1/2,j}^{n+1}\right)^+ + \left(v_{i+1/2,j}^{n+1}\right)^-, \\ G_{i,j+\frac{1}{2}}^{n+1} &= \left(w_{i,j+1/2}^{n+1}\right)^+ + \left(w_{i,j+1/2}^{n+1}\right)^-, \end{aligned} \tag{9.8}$$

where the velocities $v^{n+1}_{i+\frac{1}{2},j}$ and $w^{n+1}_{i,j+\frac{1}{2}}$ are discretized from (9.4) as

$$v^{n+1}_{i+\frac{1}{2},j} = -\frac{\xi^{n+1}_{i+1,j} - \xi^{n+1}_{i,j}}{\Delta x}, \quad w^{n+1}_{i,j+\frac{1}{2}} = -\frac{\xi^{n+1}_{i,j+1} - \xi^{n+1}_{i,j}}{\Delta y}, \tag{9.9}$$

where $\xi^{n+1}_{i,j}$ is the discretized variation of the free energy defined in (9.4). The upwind approach in (9.8) follows from

$$\begin{aligned}
\left(v^{n+1}_{i+1/2,j}\right)^+ &= \max(v^{n+1}_{i+1/2,j},0), \quad \left(v^{n+1}_{i+1/2,j}\right)^- = \min(v^{n+1}_{i+1/2,j},0), \\
\left(w^{n+1}_{i,j+1/2}\right)^+ &= \max(w^{n+1}_{i,j+1/2},0), \quad \left(w^{n+1}_{i,j+1/2}\right)^- = \min(w^{n+1}_{i,j+1/2},0).
\end{aligned} \tag{9.10}$$

The discretized variation of the free energy $\xi^{n+1}_{i,j}$ in (9.4) is computed with a semi-implicit scheme inspired by the ideas of [128, 274], where the so-called convexity splitting scheme is proposed to construct unconditional gradient-stable schemes (i.e. schemes that ensure the decay of the discrete version of the free energy in (7.2)). In our recent effort Chapter 8 we show that our finite-volume scheme unconditionally decreases the discrete free energy of the CH equation if the contractive part of the potential, $H_c(\phi)$, is taken as implicit; the expansive part of the potential, $H_e(\phi)$, is taken as explicit; and the Laplacian is taken as an average between the explicit and the implicit second-order discretizations, so that

$$\begin{aligned}
\xi^{n+1}_{i,j} &= H'_c(\phi^{n+1}_{i,j}) - H'_e(\phi^n_{i,j}) - \frac{\epsilon^2}{2}\left[(\Delta\phi)^n_{i,j} + (\Delta\phi)^{n+1}_{i,j}\right], \\
H(\phi) &= H_c(\phi) - H_e(\phi) = \frac{\phi^4+1}{4} - \frac{\phi^2}{2},
\end{aligned} \tag{9.11}$$

where the potential $H(\phi)$ in (7.8) is decomposed into two convex functions, $H_c(\phi)$ and $H_e(\phi)$,

$$H(\phi) = H_c(\phi) - H_e(\phi) = \frac{\phi^4+1}{4} - \frac{\phi^2}{2}.$$

The discrete two-dimensional approximation of the Laplacian $(\Delta\phi)^n_{i,j}$ is chosen to satisfy the second-order form

$$(\Delta\phi)^n_{i,j} := \frac{\phi^n_{i+1,j} - 2\phi^n_{i,j} + \phi^n_{i-1,j}}{\Delta x^2} + \frac{\phi^n_{i,j+1} - 2\phi^n_{i,j} + \phi^n_{i,j-1}}{\Delta y^2}. \tag{9.12}$$

The modified CH equation in (9.1) employs the no-flux boundary conditions defined in (7.4). The numerical implementation of the boundary conditions follows from

$$\begin{aligned}
\phi^n_{i-1,j} &= \phi^n_{i,j} \text{ for } i=1, \forall j; \quad \phi^n_{i+1,j} = \phi^n_{i,j} \text{ for } i=N, \forall j; \\
\phi^n_{i,j-1} &= \phi^n_{i,j} \text{ for } j=1, \forall i; \quad \phi^n_{i,j+1} = \phi^n_{i,j} \text{ for } j=N, \forall i; \\
F^{n+1}_{i-\frac{1}{2},j} &= 0 \text{ for } i=1, \forall j; \quad F^{n+1}_{i+\frac{1}{2},j} = 0 \text{ for } i=N, \forall j; \\
G^{n+1}_{i,j-\frac{1}{2}} &= 0 \text{ for } j=1, \forall i; \quad G^{n+1}_{i,j+\frac{1}{2}} = 0 \text{ for } j=N, \forall i.
\end{aligned} \tag{9.13}$$

Remark 9.2.1 (Choice of time step Δt and mesh size $\Delta x = \Delta y$). Explicit finite-volume schemes are often stable under a CFL condition imposed over the time step and mesh size. The implicit scheme presented in Subsection 9.2.1 does not need of a CFL condition to decay the discrete free energy of the original Cahn-Hilliard equation in (7.1), as proved in Chapter 8. This doesn't mean however that any Δt works in practice: in each time step the nonlinear system (9.6) has to be solved by iteration, and with a large Δt the convergence is likely to fail. In this work we have tested that with the choice of $\Delta t = 0.1$ and $\Delta x = \Delta y = 1$ our simulations converge to the right solution for the simulations presented in Section 9.3.

Remark 9.2.2 (Dimensional-splitting scheme and parallelisation). The full 2D finite-volume scheme presented in Subsection 9.2.1 can be reformulated by employing a dimensional-splitting methodology: instead of solving the full 2D image altogether, this technique initially solves row by row and then column by column. The detailed construction of such scheme is presented in Chapter 8 and is based on [17], and here we briefly illustrate how it works.

For the dimensional-splitting approach we firstly update the solution along the x axis, for each index j corresponding to a fix value of y_j where $j \in [1, N]$. Subsequently, we proceed in the same way along the y axis, for each index i corresponding to a fixed value a value of x_i where $i \in [1, N]$. The index r, where $r \in [1, N]$, denotes the index j of the fixed y_j value in every x axis of the first loop, and the updated average density for each x axis with $j = r$ is $\phi_{i,j}^{n,r}$. Similarly, the index $c \in [1, N]$ denotes the index i for every fixed value of x_j in each y axis of the second loop, and the updated density for each y axis with $i = c$ is $\phi_{i,j}^{n,c}$.

In the first place we go through each of the x axes of the domain at a fixed y_j with $j = r$. The initial conditions for the scheme are $\phi_{i,j}^{n,0} := \phi_{i,j}^{n}$. The scheme for each x along the loop $r = 1, \dots, N$ satisfies

$$\phi_{i,j}^{n,r} = \begin{cases} \phi_{i,j}^{n,r-1} - \frac{\Delta t}{\Delta x} \left(F_{i+1/2,j}^{n,r} - F_{i-\frac{1}{2},j}^{n,r} \right) + \lambda_{i,j}(\phi_{i,j}^{0} - \phi_{i,j}^{n,r-1}) & \text{if } j = r; \\ \phi_{i,j}^{n,r-1} & \text{otherwise;} \end{cases} \tag{9.14}$$

with $F_{i+1/2,j}^{n,r}$ computed in a similar fashion as in (9.8). Once the loop for the rows is completed, we define the intermediate density values as $\rho^{n+1/2} := \rho^{n,N}$. Subsequently, we continue through each of the y axes with index $c = 1, \dots, N$, each of them at a fixed x_i with $i = c$. The initial condition for this scheme is $\phi_{i,j}^{n,0} := \phi_{i,j}^{n+1/2}$. The scheme for each y along the loop $c = 1, \dots, N$ satisfies:

$$\phi_{i,j}^{n,c} = \begin{cases} \phi_{i,j}^{n,c-1} - \frac{\Delta t}{\Delta y} \left(G_{i,j+1/2}^{n,c} - G_{i,j-\frac{1}{2}}^{n,c} \right) + \lambda_{i,j}(\phi_{i,j}^{0} - \phi_{i,j}^{n,c-1}) & \text{if } i = c; \\ \phi_{i,j}^{n,c-1} & \text{otherwise.} \end{cases} \tag{9.15}$$

Once the loop for the columns is completed, we define the final density values $\phi_{i,j}^{n+1}$ after a discrete timestep Δt as $\phi_{i,j}^{n+1} := \phi_{i,j}^{n,N}$. This dimensional-splitting scheme can be fully parallelized in order to save computational time. This is because: (i) the scheme does not take notice the order of updating the rows/columns, as long as all of them are updated; (ii) one row or column only depends on the values of the directly adjacent rows or columns respectively. As a result, a strategy to parallelize the dimensional splitting scheme consists of updating at the same time all the odd rows/columns, since they do not depend on one another. One can then proceed with all the even rows/columns at the same time.

9.2.2 Two-step method for the modified CH equation

The two-step method followed here was firstly proposed in [33]. It basically consists of applying the finite-volume scheme in subsection 9.2.1 twice with different values of the parameter ϵ. The first stage consists of taking a large ϵ to execute a large-scale topological reconnection of shapes, leading to images with diffused edges. Subsequently, and in order to sharpen the edges after the first stage, ϵ is substantially reduced, and the final outcome becomes less blurry and diffused. We denote the corresponding values of ϵ as ϵ_1 and ϵ_2.

Adequately tuning the two values of ϵ in each stage, as well as λ, is vital to complete a successful image inpainting. Those values have to be chosen empirically and depend on the dataset and type of damage, and in Subsection 9.3.1 we conduct a study to select them. As explained there, the appropriate values for ϵ are between 0.5 and 1.5 for MNIST-like images, while $\lambda \in [1, 1000]$. The cell sizes are $\Delta x = \Delta y = 1$. The reader can find the exact values employed after the analysis of Subsection 9.3.1 in Table 9.3.2.

9.2.3 Neural network architecture for classification

The prediction of the label in the restored images is performed via a neural network constructed in TensorFlow [221]. Its architecture is defined taking into account that in this work we employ the MNIST dataset [112], which contains binary images of digits from 0 to 9 and has a resolution of 28×28 pixels. This is a benchmark dataset in the community and is the de facto "hello world" dataset of computer vision. There are consequently plenty of neural network architectures attaining extremely high accuracies for the MNIST dataset, and we refer the reader to the Kaggle competition of Digit Recognizer in [116] for examples of such architectures.

Here, however, our overarching objective aim is to quantify how the prediction of damaged images is enhanced once the CH filter is applied to the images beforehand. Hence we do not require a highly sophisticated neural network and a cutting-edge architecture as in computer vision: our images are not going to be exactly the same as in the training set due to the damage and the subsequent restoration. We then select a standard architecture for classification based on sequential dense layers. Such architecture is formed by:

1) A flatten layer that takes the 28×28 image input and turns it into an array with 784 elements. There are no weights to optimize in this layer.

2) A dense layer with 64 units and the ReLU activation function, defined as $f(x) = \max\{0, x\}$ for $x \in \mathbb{R}$. There are 784×64 weights to optimize in this layer, in addition to the bias term in each of the 64 units.

3) Another dense layer with 64 units and the ReLU activation function. There are 64×64 weights to optimize in this layer, in addition to the bias term in each of the 64 units.

4) A final dense layer with 10 units and the softmax activation function, which returns the normalized probability distribution for the 10 labels and satisfies $\sigma(z_i) = \exp(z_i)/\sum_{j=1}^{10} \exp(z_j)$, with $z = (z_1, \ldots, z_{10})$ being the output of the final dense layer with 10 units. There are 64×10 weights to optimize in this layer, in addition to the bias term in each of the 10 units.

For the training of this network we initially divide the original MNIST dataset in 60000 training images and 10000 testing images. Then we train the neural network for 10 epochs with the Adam optimizer, choosing as loss function the categorical crossentropy defined as

$$J(w) = \frac{1}{N} \sum_{i=1}^{N} \left[y_i \log(\hat{y}_i) + (1 - y_i) \log(1 - \hat{y}_i) \right],$$

with w being the weights to optimize, y_i each of the N true labels of the training dataset, and $\hat{y}_i$ each of the N predicted labels. After 10 epochs we get an accuracy for the training dataset of 99.02%, while for the test set the accuracy is 97.47%. Once the neural network is trained we keep the weights fixed for the comparison of damaged and restored images in Section 9.3. A display of the neural network is depicted in Figure 9.2.1.

9.2.4 Integrated Algorithm

The integrated algorithm proposed and tested in this work takes as input a damaged image, applies the CH filter based on Subsection 9.2.1 and Subsection 9.2.2 to restore it, and finally applies the already-trained neural network in Subsection 9.2.3 to predict its label.

To show the applicability of this integrated algorithm we initially create damage in the images of the test set in the MNIST dataset [112]. After we apply the image inpainting to the damaged test images, we introduce the restored images in the neural network. At that point, and since we
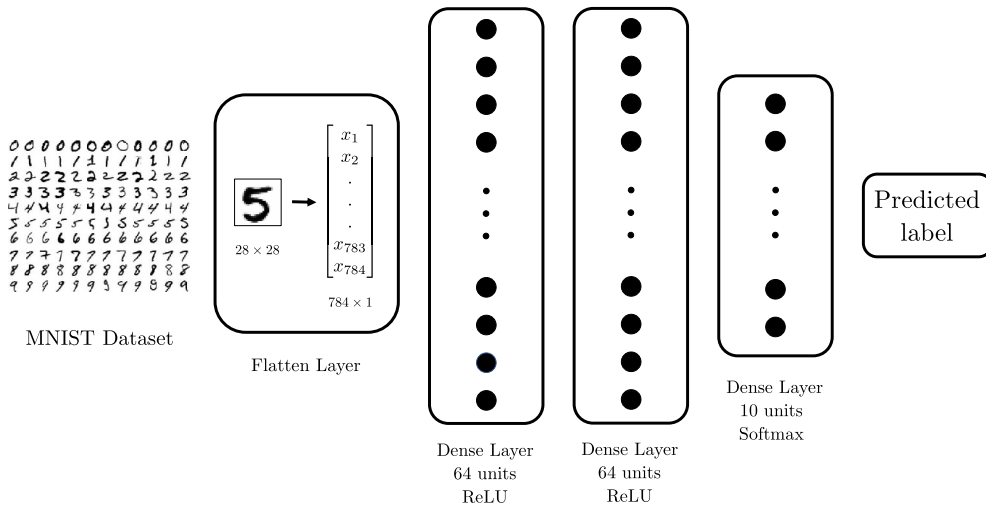
Figure 9.2.1: Diagram showing the layers of the neural network of Subsection 9.2.3.

have the true labels of the test set, we can assess the attained accuracy in comparison to directly introducing the damaged images or the original images into the neural network. This procedure is conducted for multiple types of damage in Section 9.3, and a schematic representation of all steps is depicted in Figure 9.2.2.

9.3   Application of the integrated algorithm to the MNIST dataset

Our focus here is testing the applicability of the integrated algorithm in Subsection 9.2.4 to increase predictability in damaged images. In Subsection 9.3.1 we start by analysing the impact of the parameters λ_0 and ϵ on the inpainting process, with the objective of calibrating them before employing the MNIST dataset. In Subsection 9.3.2 we detail the types of damage that we insert into the MNIST dataset, and we also show the restored outcomes of applying the CH equation as an image inpainting filter. Finally, in Subsection 9.3.3 we evaluate how the accuracy of the damaged images increases after applying the CH filter to them, for various types and degrees of damage.

9.3.1   Inpainting of a Crossline

We employ the crossline example in [33] to analyse the role of the parameters ϵ and λ_0 in the finite-volume scheme of Subsection 9.2.1. These two parameters crucially determine the success of the image inpainting procedure, and consequently appropriate calibrations for the parameters must be chosen before running the scheme. The original crossline image is depicted in Figure 9.3.1a, and we add to it a grey damage in the center, as shown in Figure 9.3.1b. This image contains 50×50 pixels or cells, each with a size of $\Delta x = \Delta y = 1$. We apply the finite-volume scheme in Subsection 9.2.1 to the damaged image in Figure 9.3.1b.

We first aim to determine λ_0 in (9.7) and we set $\epsilon = 1$ as an initial guess so that $\epsilon = \Delta x = \Delta y$. From (9.7), $\lambda_{i,j}$ is only nonzero for the predefined area of undamaged pixels. Indeed the term with $\lambda_{i,j}$ in the finite-volume scheme in (9.6) ensures that the undamaged pixels are not modified during the image inpainting, but for this λ_0 has to be sufficiently large to counterbalance the fluxes of the scheme. Bearing this in mind we run the numerical scheme with a $\Delta t = 0.1$ and until the L^1 norm between successive states is lower than a certain tolerance fixed to be 10^{-4}. Our simulation produces satisfactory results and does not break down for a range of $\lambda \in [1, 1000]$. Table 9.3.1
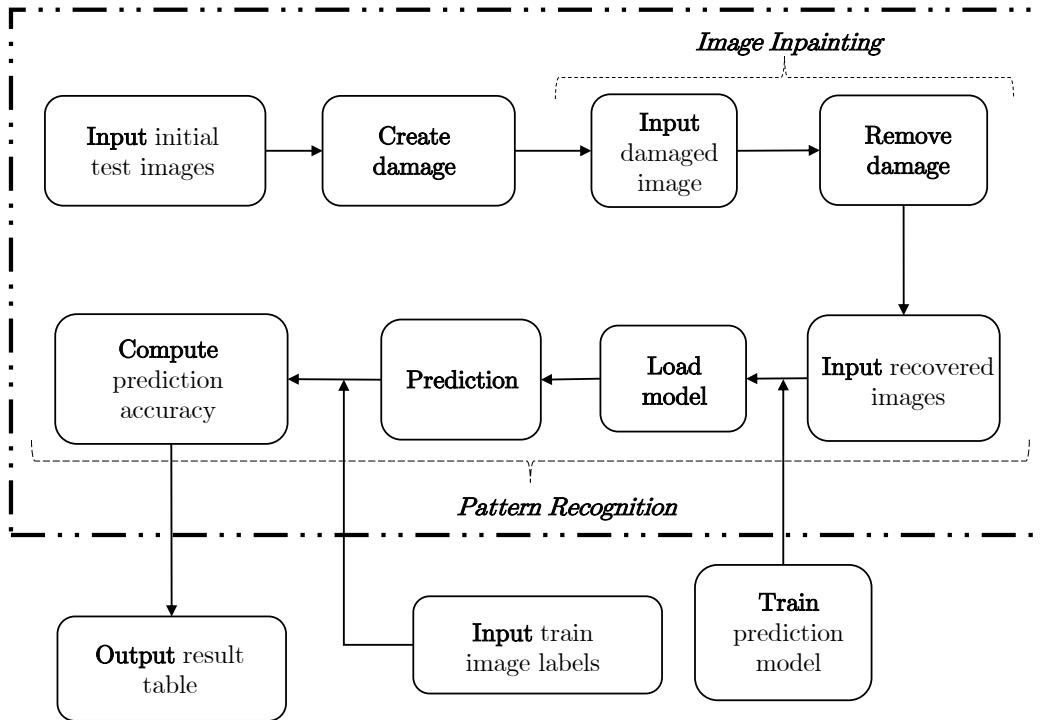
Figure 9.2.2: A schematic representation to show the applicability of the integrated algorithm.

shows that the computational time to reach the required tolerance decreases when increasing the value of λ. In addition, the L^1 norm between the final and initial state is also lower for greater λ. It is worth mentioning that different choices for Δt yield different ranges of valid λ, given that (9.6) is a singularly perturbed problem for large λ_0. Hence, greater values of λ are possible if Δt is refined. In our case, with the choice of $\Delta t = 0.1$, our finite-volume scheme does not yield any result and breaks down during the for values of $\lambda \notin [1, 1000]$.

λ	time	L^1 norm
1	489.8	61.2
10	489.7	60.8
100	484.8	60.8
1000	481.5	60.7

Table 9.3.1: Comparison for different values of λ: computational time before reaching the tolerance and L^1 norm between the final and initial state.

We next consider the tuning of the parameter ϵ which in turn is related to the pixel size Δx and Δy. For values of ϵ larger than the pixel size the outcome tends to be diffusive, while for smaller values the edges are sharpened. When applying the finite-volume scheme in (9.6) with $\lambda \in [1, 1000]$ we obtain satisfactory results for $\epsilon \in [0.5, 1.5]$, while for values outside this range the simulation breaks down because of the singular nature of (9.6) introduced by the parameter λ (see Section 9.A for more details). As a consequence, for the two-step method in Subsection 9.2.2 we first take the value $\epsilon_1 = 1.5$ for the large-scale topological reconnection of shapes, while for the second step we

choose the value $\epsilon_2 = 0.5$ to sharpen the edges. The image inpainting of the damaged image in Figure 9.3.1b resulting from applying this choice of parameters is shown in Figure 9.3.1c. More details about the choice of λ, ϵ_1 and ϵ_2 are provided in Section 9.A, where a sensitivity analysis is carried out depicting the outcome of choosing a not optimal combination of parameters.
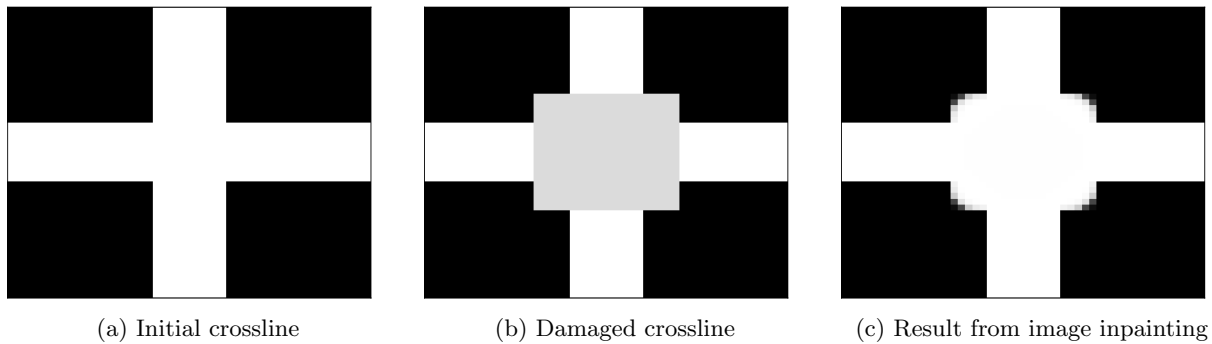

(a) Initial crossline   (b) Damaged crossline   (c) Result from image inpainting

Figure 9.3.1: Image inpainting of a crossline, inspired by [33].

The final outcome after the image inpainting in Figure 9.3.1c is not the same as the original image in Figure 9.3.1a. The reason for this is explained in the work of [32], where multiple steady-state solutions of the modified CH equation were shown to exist. As the information under the inpainting region has been destroyed, there is no way of knowing that the steady state we obtain is less accurate than other viable solutions, in comparison to Figure 9.3.1a. For further details we refer the reader to [32], where a bifurcation analysis is carried out to show that the steady state may vary depending on the choices for ϵ and Δx, Δy.

9.3.2 Damage introduced in the MNIST dataset

Here we discuss the types of damage inserted into the MNIST test set, with the objective of subsequently applying the CH filter developed in Subsection 9.2.1 for image inpainting. The varied damage employed aims to represent a mock case of damage that may be encountered in an image in need for restoration. As a result, we decide to employ two kinds of damage with different intensities: customized damage affecting particular regions of the image, and random damage selecting arbitrary pixels or horizontal lines in the image. The details of both are:

a) Customized damage: this type of damage is applied in four different fashions, as shown in Figure 9.3.2. The basic idea is to turn vertical or horizontal lines of pixels into a uniform grey intensity between black and white colour. In Figure 9.3.2 we show the outcome of applying the CH filter to the damaged images. It can be seen that our model is able to recover the images from the different types of damage, albeit with varying degrees of success. For instance, the damage introduced in Figure 9.3.2c is a thick horizontal line which implied a considerable loss of information from the original image, compared to the other types of damage. As a result, the inpainted image filter for this type of damage is not as effective as the other ones, as it can be seen from the inpaintings in Figure 9.3.2.

b) Random damage: this second type of damage is inserted in a random fashion and with different levels of intensity. Two ways of randomly creating damage are considered: one makes use of randomly selecting whole horizontal rows of pixels, while the other is obtained by randomly selecting individual pixels. In addition, for both types of random damage we employ different levels of damage intensity, so that a higher percentage of the image contains damage if the intensity rises. This allows us to test how our image inpainting algorithm
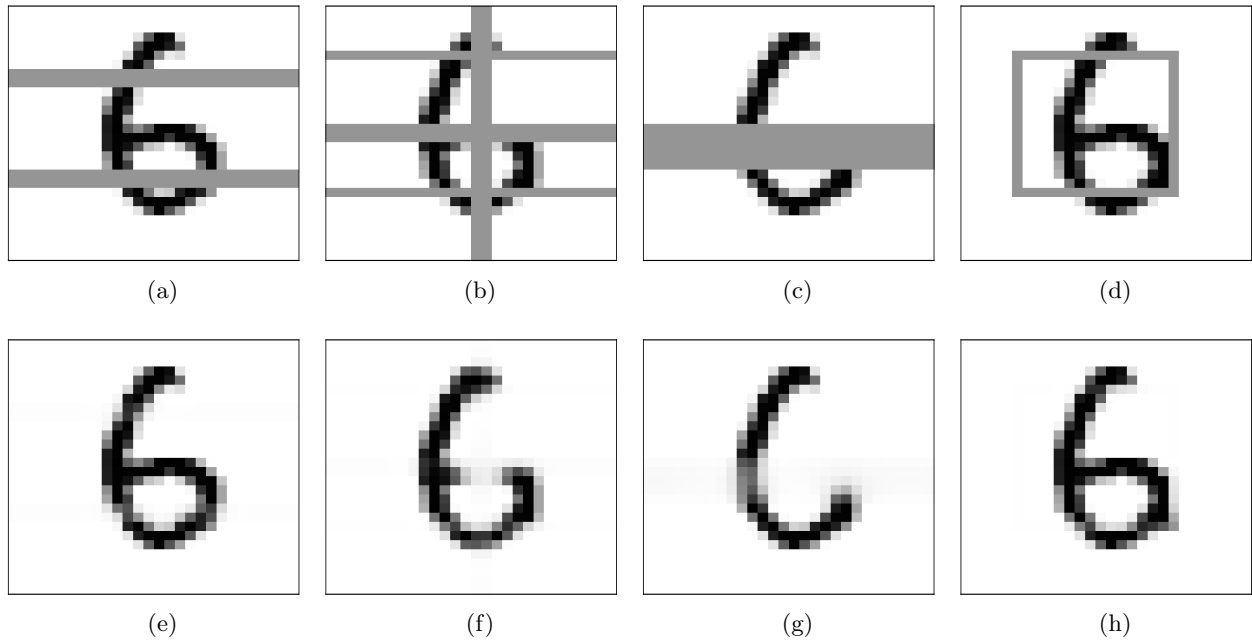
Figure 9.3.2: Customized damage applied to a particular sample of the MNIST dataset. (a)-(d): the sample with four different types of damage; (e)-(h): The outcome of applying image inpainting to the damaged samples.

behaves with increasing levels of damage in the image. Examples of these damages are shown in Figure 9.3.3. Similarly to the case of customized damage, the higher the intensity of damage the more information is lost in the inpainting, as we can see for example in the case of 80% pixel damage in Figure 9.3.3. But despite of this, our image inpainting algorithm renders recognisable images even with relative high amounts of damage.

Parameters		ϵ_1	ϵ_2	λ
Customized damage		1.5	0.5	1000
Random damage	Rows	1.5	0.5	1000
	Pixels	1.5	0.5	9000

Table 9.3.2: Optimal parameters of the two-step algorithm applied in the MNIST dataset. ϵ_1 and ϵ_2 denote the values of ϵ used in the first and second step respectively, as explained in Subsection 9.2.2.

The parameter values of λ, ϵ_1 and ϵ_2 are gathered in Table 9.3.2, and follow the reasoning discussed in Subsection 9.3.2. This choice of parameters in our finite-volume scheme in Subsection 9.2.1 leads to an effective image inpainting algorithm capable of restoring images with damage of varied nature as shown in Figure 9.3.2 and Figure 9.3.3. Further details about the selection of λ, ϵ_1 and ϵ_2 are provided in Section 9.A, where a sensitivity analysis is provided in order to justify the parameter choices in Table 9.3.2. The next step is to integrate this image inpainting algorithm within a pattern recognition framework for the MNIST dataset.

9.3.3 Pattern recognition for inpainted images

We now apply the neural network described in Subsection 9.2.3 to predict labels of damaged images with and without image inpainting, with the aim of quantifying the improvement of accuracy
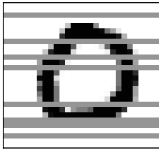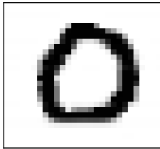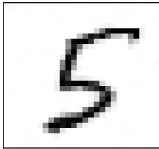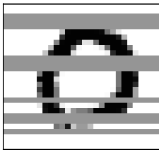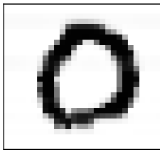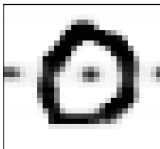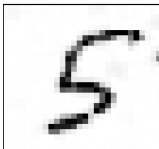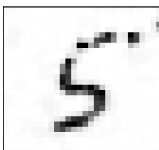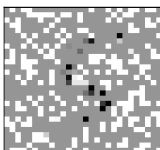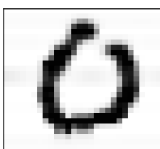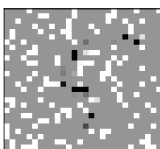
Damaged rows	Damaged	Inpainted	Damaged pixels	Damaged	Inpainted
8			30%		
10			40%		
12			50%		
14			60%		
16			70%		
18			80%		

Figure 9.3.3: Examples of image inpainting for random damage in whole horizontal rows and in individual pixels. The column entitled "Damaged rows" marks the number of rows randomly selected for damage in the 28×28 images. The column entitled "Damaged pixels" marks the percentage of randomly damaged pixels over the whole image. For higher levels of damage intensity the inpaintings lose more information.

following the application of the CH filter. This study is completed for the different types and intensities of the noise depicted in Subsection 9.3.2.

We begin by adding the types of damage in Subsection 9.3.2 to the 10,000 samples of the MNIST test dataset. The next step is to apply the CH filter and two-step method to each one of them, while also saving copies of the test images with the damage. Eventually, for each type of damage we get two batches of 10,000 images: one still with the damage, and another one with

image inpainting applied. Given that the neural network of Subsection 9.2.3 is already trained with the 60,000 samples of the MNIST training dataset, we can directly compute the accuracy of each of the two batches. This way we are able to assess the improvement in accuracy thanks to applying image inpainting to restore the damage.

Here for the validation we just employ the accuracy metric, which is defined as follows

$$\text{Accuracy} \equiv \frac{\text{Number of correct predictions}}{\text{Number of total predictions}}.$$

There are however many other metrics apart from the accuracy one that play a vital role in other classification problems: recall, precision, F1 score, true positive rate and so on. Here we believe that the accuracy metric is enough to draw conclusions about how the image inpainting is improving the predictions with respect to the damage images. This is due to the fact that the MNIST dataset is a balanced dataset, where there is generally no preference between false positives and false negatives.

The measure of improvement between the batch of samples with image inpainting and the damaged ones without it is computed as

$$\text{Improvement} \equiv \frac{\text{Accuracy}_{\text{with CH filter}} - \text{Accuracy}_{\text{without CH filter}}}{\text{Accuracy}_{\text{without CH filter}}}, \tag{9.16}$$

and it basically represents the percentage of improvement that results from adding the CH filter to the prediction process.

The results for all the types of damage under consideration are displayed in Table 9.3.3, Table 9.3.4 and Table 9.3.5. In Table 9.3.3 we gather the prediction accuracies for the four customized damages displayed in Figure 9.3.2, as well as the prediction accuracy for the unmodified MNIST test set, which for our neural network architecture is 0.97. We observe that for the types of more intense customized damage B and C the accuracy prediction for the damaged images without CH filter drops to 0.71 and 0.64, respectively. By applying the filter we find that the accuracy predictions can significantly escalate to 0.93 and 0.82, leading to improvements of 31% and 28% respectively. The other two types of customized damage A and D are not as pervasive as B and C, and as a result the accuracy predictions are high even without applying the CH filter.

Customized damage	Without CH filter	With CH filter	Improvement
A	0.84	0.96	14%
B	0.71	0.93	31%
C	0.64	0.82	28%
D	0.90	0.96	7%
Initial test images	-	0.97	-

Table 9.3.3: Accuracy for the test dataset of MNIST without and with the CH filter, for the customized damage in Figure 9.3.2. The improvement is computed following (9.16).

In Table 9.3.4 and Table 9.3.5 we test the accuracies for random damage with various levels of intensities. The objective here is to analyse how the CH filter responds when the damage occupies more and more space in the images, both for the case of rows or pixels, as displayed in Figure 9.3.3. In Table 9.3.4 we show the accuracies for a range of damaged rows between 6 and 26, bearing in mind that the dimensions of the MNIST images are 28×28. We observe that for low numbers of damaged rows the accuracy prediction even without the CH filter is high and it does not improve significantly by adding the filter. But then the improvement surges until reaching a maximum value of 47% for 16 random damaged rows, where the prediction without CH filter is 0.55 and with CH filter 0.81. From larger numbers of damaged rows the accuracies drastically drop due to the large

amount of information lost, and not even the image inpainting process is able to achieve decent accuracies. In the limit of damaged number of rows tending to 28 we observe that the accuracies are close to the ones of a dummy classifier with one out of ten chances of rightly guessing the label. In this limit there is no difference between adding the CH filter or not, and it turns out that the improvements are even negative.

Damaged rows	Without CH filter	With CH filter	Improvement
6	0.89	0.96	8%
8	0.82	0.93	13%
10	0.73	0.91	25%
12	0.66	0.87	32%
14	0.6	0.87	45%
16	0.55	0.81	47%
18	0.47	0.68	45%
20	0.40	0.48	20%
22	0.39	0.45	15%
24	0.33	0.26	-21%
26	0.20	0.12	-40%

Table 9.3.4: Accuracy for the test dataset of MNIST without and with the CH filter, for the case of random damage in rows. The improvement is computed following (9.16).

We observe a similar pattern for the case of random damaged pixels in Table 9.3.5 we observe a similar pattern. For low percentages of damaged pixels the improvement of adding the CH filter is negligible and the accuracies with and without the filter are quite high. As we increase the percentage of damaged pixels the improvement escalates until it reaches 45% for a scenario with 80% of the pixels randomly damaged. For this case the accuracy prediction without the filter is just 0.55, but thanks to the filter it significantly increases to a decent value of 0.8. For larger percentage of pixels the improvement and accuracies drop, and in the limit towards 100% of damaged pixels we get close to the accuracy of a dummy classifier. This is due to the large loss of information that the original images have suffered.

Damaged pixels	Without CH filter	With CH filter	Improvement
30%	0.93	0.93	0%
40%	0.96	0.96	0%
50%	0.91	0.95	4%
60%	0.8	0.94	18%
70%	0.75	0.93	24%
80%	0.55	0.8	45%
90%	0.39	0.46	18%
92%	0.32	0.37	16%
94%	0.33	0.34	3%
96%	0.20	0.23	15%

Table 9.3.5: Accuracy for the test dataset of MNIST without and with the CH filter, for the case of random damage in pixels. The improvement is computed following (9.16).

In Figure 9.3.4 we depict some specific examples for which the label is only predicted correctly after applying the CH filter to the damaged image. These are just some particular samples out

of the 10,000 images contained in the test dataset of MNIST, and for some of them the opposite effect can occur: that the label is correctly predicted for the damaged image but following the inpainting process it is predicted incorrectly. This only occurs for really severe damages, where there is no difference between adding the CH filter or not and the accuracies are close to the ones of a dummy classifier guessing randomly. However, we have shown in Table 9.3.4 and Table 9.3.5 that overall the CH filter increases the global accuracy for images with low to moderate damage, and consequently we expect that scenarios such as the ones displayed in Figure 9.3.4 are much more common than the opposite ones.

In Figure 9.3.4 we also appreciate that some greyscale pixels remain in the inpainted image after applying the modified CH filter. This is due to the fact that the finite-volume scheme has not run for long enough to turn those intensities into the stable white ($\phi = -1$) or black ($\phi = 1$) phase field. This is prone to happening for images with high levels of damage, and one should try to let the simulation run for as long as possible to avoid possible misclassification issues due to those greyscale regions.
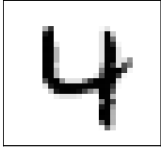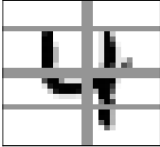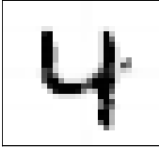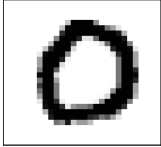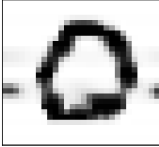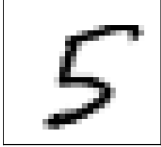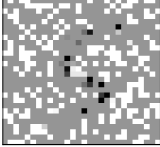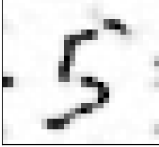
Initial Image	True label	Damaged image	P_D	Inpainted image	P_I
	4		6		4
	0		8		0
	5		4		5

Figure 9.3.4: Particular examples of label predictions for MNIST samples with various types of damage: customized damage C, random damage in 16 rows and random damage in 70% of pixels. The label of the damaged image is wrongly predicted, while the label for the inpainted image is correctly predicted. P_D and P_I represent the label predictions of the damaged and inpainted images, respectively.

Appendix

9.A Sensitivity analysis of the parameters ϵ_1, ϵ_2 and λ

Adequately tuning the two values of ϵ_1, ϵ_2 and λ, is vital to complete a successful image inpainting. These values have to be chosen empirically and depend on the dataset and type of damage. Here we provide further details about finding an adequate combination of these parameters for images of MNIST.

On the one hand, the parameter ϵ is related to the interphase thickness of the solution of the modified CH equation in (9.1). The larger the ϵ, the more diffused the resulting image. However, a large ϵ allows for a large-scale topological reconnection of shapes, which is convenient when the damage breaks a connected colour phase. The strategy proposed in [33] and followed here, consists of employing a two-stage methodology where a large ϵ is employed first (denoted as ϵ_1), followed by a smaller ϵ (denoted as ϵ_2) in order to sharpen the final image. In Figure 9.A.1 we depict some of the image-inpainted images resulting from various combinations of ϵ_1 and ϵ_2. The original image with damage is shown in Figure 9.A.1a, and the optimal combination of ϵ_1 and ϵ_2 following the values in Table 9.3.2 is depicted in Figure 9.A.1b. For this example the optimal parameters are $\epsilon_1 = 1.5$ and $\epsilon_2 = 0.5$, and depend on the mesh size that in this case is $\Delta x = \Delta y = 1$. A larger mesh size requires a larger value of ϵ to produce the same outcome, implying that both parameters are interconnected and have to be of the same order. In order to evaluate the impact of ϵ_1 and ϵ_2 we run two experiments: in the first one we choose both of them with the small value of $\epsilon_1 = \epsilon_2 = 0.5$, and in the second one both parameters take the large value $\epsilon_1 = \epsilon_2 = 1.5$. The outcome with the small value is depicted in Figure 9.A.1c, and clearly there is not an efficient large-scale topological reconnection of the white regions, resulting in some damage left in the image. The outcome with the large value is depicted in Figure 9.A.1d, and the drawback here is that the image is too diffused in comparison to the optimal combination in Figure 9.A.1b.

On the other hand, the parameter λ in the finite-volume scheme in (9.6) ensures that the undamaged pixels are not modified during the image inpainting. The choice of λ has to be sufficiently large to counterbalance the fluxes of the scheme and act as a penalty term that keeps the undamaged pixels invariant. However, λ cannot be too large since otherwise the finite-volume scheme becomes a singularly perturbed problem, and the convergence of the implicit scheme is deteriorated. The value of λ is related to the time step Δt, and greater values of λ are possible if Δt is refined. In our case, with the choice of $\Delta t = 0.1$, our finite-volume scheme does not yield any result and breaks down during the simulation for values of $\lambda \notin [1, 10000]$. To evaluate the impact of λ we run three experiments: in the first one we choose the really small value $\lambda = 0.01$, in the second the small $\lambda = 0.1$ and in the third the large value $\lambda = 10000$. The outcomes with the two small values are depicted in Figure 9.A.1e and Figure 9.A.1f, and we observe that the undamaged pixels have been altered and the overall appearance is not clean. The outcome with the large value is depicted in Figure 9.A.1g, and the image is comparable to the optimal one in Figure 9.A.1b. There is a threshold-λ value above which there is no convergence of the implicit

finite-volume scheme. For the particular image we found this threshold can be approximated by $\lambda = 10000$.

(a) Original image with damage

(b) $\epsilon_1 = 1.5$, $\epsilon_2 = 0.5$ and $\lambda = 1000$

(c) $\epsilon_1 = 0.5$, $\epsilon_2 = 0.5$ and $\lambda = 1000$

(d) $\epsilon_1 = 1.5$, $\epsilon_2 = 1.5$ and $\lambda = 1000$

(e) $\epsilon_1 = 1.5$, $\epsilon_2 = 0.5$ and $\lambda = 0.01$

(f) $\epsilon_1 = 1.5$, $\epsilon_2 = 0.5$ and $\lambda = 0.1$

(g) $\epsilon_1 = 1.5$, $\epsilon_2 = 0.5$ and $\lambda = 10000$
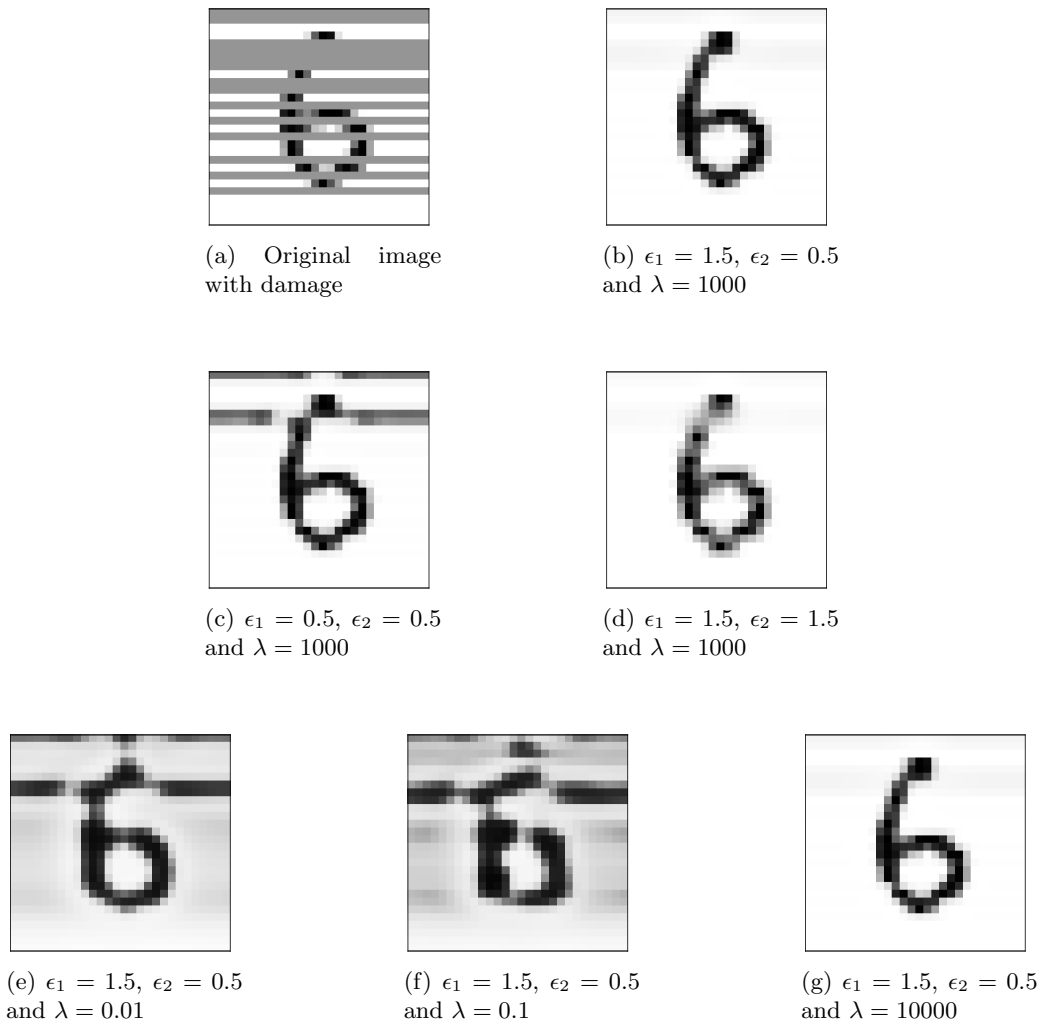
Figure 9.A.1: Sensitivity analysis for different values of ϵ_1, ϵ_2 and λ. (A): original image with damage; (B): optimal choice of parameters; (C): short value of ϵ_1 and ϵ_2; (D): large value of ϵ_1 and ϵ_2; (E): really small value of λ; (F): small value of λ; (G): large value of λ.

Part IV

Finite-volume schemes for fluctuating hydrodynamics

Introduction to fluctuating hydrodynamics

The study of fluid dynamics encounters major challenges due to the inherently multiscale nature of fluids. Not surprisingly, fluid dynamics has been one of the main arenas of activity for numerical analysis and fluids are commonly studied via numerical simulations, either at molecular scale, by using molecular dynamics (MD) or Monte Carlo (MC) simulations; or at macro scale, by utilising deterministic models based on the conservation of fundamental quantities, namely mass, momentum and energy. While atomistic simulations take into account thermal fluctuations, they come with an important drawback, the enormous computational cost of having to resolve at least three degrees of freedom per particle. Despite drastic improvements in computational power over the last few decades, atomistic simulations are only applicable for small fluid volumes. There are also other challenges with such techniques, e.g. the use of a proper thermostat when running non-equilibrium simulations at constant temperature [242]. On the contrary, the convenience of partial differential equations (PDEs), such as continuity and Navier-Stokes, is enormous as they are amenable to both analytical and numerical scrutiny, with numerical simulations being less computationally expensive than MD-MC. However, continuous models based upon PDEs cannot account for the stochastic nature observed in real systems. Fortunately, there is still an approach which lives at the crossroad of mesoscale, namely fluctuating hydrodynamics (FH). Firstly proposed by Landau and Lifshitz [187], FH is formulated in terms of stochastic PDEs which aim at extending Navier-Stokes equations to include thermal fluctuations. FH can then be used to simulate systems undergoing energy-barrier crossing transitions, such as nucleation, which are impossible to describe within the mean-field approximation.

However, the FH formulation by Landau and Lifshitz [187] is phenomenological: they simply included additive stochastic flux terms in the Navier-Stokes equations – we shall refer to these equations as the Landau-Lifshitz-Navier-Stokes (LLNS) equations. A remarkable effort has been made ever since trying to connect FH with MD from first principles [38, 81, 109, 126, 133, 176, 208]. Some of the most widely known attempts to formalise such a connection are the works of Kawasaki [176] and Dean [109]. Theirs provide a formal derivation of the stochastic time-evolution equation for the "density" field of a system of Brownian particles. Nevertheless, their derivation ends up with a time-evolution equation for the microscopic density field, which is nothing but a re-writing of the Brownian equations by using Itô's lemma. For this reason, the Dean-Kawasaki equation has been actively criticised and seen more as a toy model whose derivation does not represent a proper proof of the FH equation. Indeed this model cannot be employed to describe macroscopic quantities, such as density and momentum fields which are obtained by ensemble averaging the corresponding microscopic quantities [7], and thus remains disconnected from the

original Landau-Lifshitz theory. And it is this disconnection that has led to the misconception that the Dean-Kawasaki model describes the evolution of macroscopic observables.

In a recent work [123], a bottom-up derivation of the FH for a system of Brownian particles has been posed. It provided a new first-principle formulation of the governing equations for macroscopic observables in the framework of classical dynamic density functional theory (DDFT). It is worth mentioning that the field of DDFT has gained a lot of traction since the first phenomenological derivations proposed in [115, 127]. Several rigorous derivations have been put forward including effects such as inertia, hydrodynamic interactions and orientation of particles. These derivations have been proposed both for the overdamped and inertial regimes, and we refer the reader to [6, 121, 148, 149, 202] for more details.

In particular, the formulation proposed in [123] allows for a rigorous and systematic derivation of FH but also fluctuating DDFT (FDDFT) which includes the effects of thermal fluctuations on the mean-field DDFT. In that work, it is also shown how the classical DDFT is the most-likely realisation of FDDFT, thus providing closure to a long standing debate in the classical DFT community about the inclusion of fluctuations in DFT. Also, the derivation in [123] stays in tune with the original intuitive treatment of Landau and Lifshitz and at the same time alleviates the misconceptions with the Dean-Kawasaki model. As a remark, it should be noticed that LLNS equations describe a full system of particles, while FDDFT governs the time-evolution of density and momentum fields of subcomponents of a system, e.g. of colloidal particles in a bath. Because of the momentum exchange between colloidal and bath particles, the total momentum in FDDFT for colloidal particles is not conserved, being affected by thermal fluctuations and friction exerted by the bath. Let us also note here that both classical DFT and DDFT, embedded with either exact or approximated models for the density-dependent Helmholtz free-energy functional [207], have been shown to be rather powerful in the study of complex systems at the nano and microscale [119, 147, 148]. Recent advances in DFT and DDFT have extended its applicability to a wide spectrum of applications from nucleation of colloids and macromolecules [124, 201, 204] to fluids in confined geometries [146, 219, 291] and wetting phenomena [220, 290, 295]. But also highly non-uniform systems such as dense liquid droplets and solid clusters [200].

The FDDFT framework in [123], derived for the general case of arbitrarily shaped and thermalized particle, consists of two stochastic PDEs for the number density ρ (known also as particles state probability function) and velocity $\boldsymbol{u}$ fields:

$$
\begin{cases}
\partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^d, \quad t > 0, \\
\partial_t(\rho \boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) = -\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} - \gamma \rho \boldsymbol{u} - \sqrt{k_B T \gamma \rho}\, \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t).
\end{cases}
\tag{10.1}
$$

where $\mathcal{F}[\rho]$ is the density-dependent free-energy functional, γ is a friction parameter describing the interactions between the particles and the bath, k_B is the Boltzmann constant, T is the temperature and $\boldsymbol{\mathcal{W}}$ is a vector of Gaussian stochastic processes delta-correlated in space and time, i.e.

$$
\begin{aligned}
\langle \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t) \rangle &= 0, \\
\langle \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t), \boldsymbol{\mathcal{W}}(\boldsymbol{x}', t') \rangle &= 2\delta(t - t')\delta(\boldsymbol{x} - \boldsymbol{x}').
\end{aligned}
\tag{10.2}
$$

The mass of the particles is taken to be unitary. In the strong damping limit ($\gamma \to \infty$), the high friction between the particles and the bath causes the characteristic time scale of the momentum dynamics to be much shorter than the density one [123, 149]. Thus, as a first approximation, the contributions of the terms $\nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u})$ and $\partial_t(\rho \boldsymbol{u})$ can be neglected. As a result, one obtains the stochastic time-evolution equation for the density field, referred to as overdamped FDDFT [123, 185]:

$$
\partial_t \rho = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \right) + \nabla \cdot \left(\sqrt{k_B T \rho}\, \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t) \right).
\tag{10.3}
$$

Equation (10.3) may be seen as a stochastic version of the gradient-flow equation previously studied, for instance, in [68, 273]. As we later discuss in further detail, (10.3) reduces to the stochastic diffusion equation [178] when considering a system of non-interacting particles (ideal gas), whose free energy would be $\mathcal{F}[\rho] = \int \rho \left(\log \rho - 1 \right) d\mathbf{r}$. However, the presence of a more general functional $\mathcal{F}[\rho]$ allows in principle to introduce non-linear diffusion, external force fields and interparticle interactions. It is also worth mentioning that (10.3) is not well-posed due to the high irregularity originated in the stochastic fluxes and the multiplicity of the noise. This difficulty is typically overcome by introducing some sort of regularization, such as a finite-volume interpretation as employed here and in previous works [120]. One also needs to be careful about the cell size choice and possible nonphysical effects such as negative densities, which may arise from the Gaussian processes.

Previous numerical methodologies for FH have been focused on the LLNS equations for the density and momentum, and the energy equation for the temperature if the systems are non-isothermal. In comparison, the overdamped FDDFT allows us to obtain the density field solving a single equation with stochastic fluxes for isothermal systems. One of the first works on this regard is by Garcia *et al.* [139], where a simple finite-difference scheme to treat the numerical fluxes of the SPDE is constructed. Further works by Bell *et al.* [26, 27] provide an explicit Eulerian discretization of the LLNS equations combined with a third-order Runge-Kutta method with the objective of adequately reproducing the fluctuations in density, energy and momentum. Donev and co-workers [120] exploited the structure factor (equilibrium fluctuation spectrum) to construct finite-volume schemes to solve the LLNS which then allows one to study the accuracy for a given discretization at long wavelengths.

They also proposed a Petrov-Galerkin finite-element discretization of non-linear stochastic diffusion equations embedded with prototypical free-energy functionals, such as the Ginzburg-Landau free energy [270]. However, in order to obtain analytical forms of the structure factors used to assess the performance of the scheme, the study focused on systems at equilibrium (i.e. without density discontinuities), at supercritical temperatures (to avoid phase transition phenomena), and without any external potential. Similarly, methods to solve FH via staggered grids have been constructed [19]. Other works have proposed numerical schemes based on temporal integrators that are implicit-explicit predictor-corrector [111] or two-level leapfrog [145]. Additionally, hybrid schemes have been developed to couple LLNS with MD [107, 108, 110] or with MC [118, 277] simulations of complex fluid systems. Moreover, the LLNS have also been solved to tackle reactive multi-species fluid mixtures [35]. Further works have developed numerical schemes for particular applications of the overdamped FDDFT in (10.3). Specifically, [9, 178] developed numerical methods for reaction-diffusion equations obtained by adding appropriate reaction terms to (10.3) equipped with the ideal-gas free-energy functional.

The works just mentioned have contributed to a better understanding of the effects of thermal fluctuations in complex fluid systems. Nevertheless, an efficient and systematic numerical methodology to solve (10.1) and (10.3) equipped with a general free-energy functional has not yet been developed. Such a methodology would allow for the simulation and scrutiny of a wide range of non-equilibrium phenomena which can be studied within the framework of FDDFT. Relevant examples of these physical phenomena include dynamic evolution of confined systems and energy-barrier crossing transitions, such as nucleation.

10.1 Contribution to numerical schemes for overdamped fluctuating hydrodynamics

In this work we introduce a finite-volume method to solve general stochastic gradient-flow equations with the structure of (10.3) for FDDFT. The main advantages of finite-volume schemes are the

conservation of the total mass of the system and the flexibility to simulate complex geometries. The main contributions of this work can be summarized as follows:

- To provide a space discretization scheme able to deal with fluctuations at discontinuous density profiles. We discretize the deterministic fluxes based on a hybrid approach which takes advantage of both central and upwind schemes.

- To overcome the commonplace challenge of preserving non-negative densities in the presence of noise, by adopting a Brownian bridge technique. Despite previous approaches employing artificial limiters [178], our technique ensures density positivity without altering the Gaussian distribution of the stochastic field.

- To develop a methodology to simulate a family of free-energy functionals, modelling different physical systems. First, we study temporal and spatial correlations, and the structure factor of ideal gas at equilibrium, comparing the results of our finite-volume solver with both MD and theoretical results. Then, we examine the out-of-equilibrium evolution of an ideal gas in a double-well external potential. Subsequently, we simulate homogeneous nucleation kinetics of a fluid consisting of particles interacting through a Lennard-Jones (LJ)-like potential. Providing initial uniform densities corresponding to metastable vapour conditions, we study the phase-transition of the system and compare the results with the mean-field phase diagram.

- To implement and test families of implicit-explicit Euler and Milsten time integrators, together with a weak second-order Runge-Kutta scheme.

- To gain insights into the free-energy decay for stochastic gradient-flow equations (see for instance Figs. 8(d) and 10(b)). The decay of free energy is an important feature of deterministic gradient-flow equations. However, in stochastic gradient-flow equations, the free-energy decay is guaranteed only in the weak noise limit.

Chapter 11 is devoted to this contribution. In its introduction we present the model equation to simulate and outline its main properties. In Section 11.1, we discuss the numerical methodology of our finite-volume scheme, including the discretization of deterministic and stochastic fluxes, time integrators, adaptive time step to preserve density positivity and boundary conditions. Subsequently, several applications to illustrate the validity of our methodology are presented in Section 11.2. Firstly, in Subsection 11.2.1 we begin by validating the numerical scheme for the case of an ideal-gas free energy, due to the extensive analytical results available for this particular choice of free energy. Then, in Subsection 11.2.2 we simulate an ideal gas under a double-well external potential. Finally, in Subsection 11.2.3 we apply our numerical scheme to a homogeneous vapour-liquid transition of a Lennard-Jones fluid.

Finite-volume schemes for overdamped fluctuating hydrodynamics

The contents of this chapter are based on the publication [243], done in collaboration with Antonio Russo, Miguel A. Durán-Olivencia, Peter Yatsyshin, José A. Carrillo and Serafim Kalliadasis:

[243]   A. Russo, S. P. Perez, M. A. Durán-Olivencia, P. Yatsyshin, J. A. Carrillo, and S. Kalliadasis. "A finite-volume method for fluctuating dynamical density functional theory". In: *Journal of Computational Physics* (2020), p. 109796.

Our starting point is the following general SPDE based on the overdamped FDDFT in (10.3),

$$\partial_t \rho = \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \right) + \nabla \cdot \left(\sqrt{\rho/\beta} \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t) \right). \tag{11.1}$$

where $\mathcal{F}[\rho]$ denotes the free energy of the system given by

$$\mathcal{F}[\rho] = \int_{\mathbb{R}^d} \Pi(\rho) d\boldsymbol{x} + \int_{\mathbb{R}^d} V(\boldsymbol{x}) \rho(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{2} \int_{\mathbb{R}^d} K \left(W(\boldsymbol{x}) \star \rho(\boldsymbol{x}) \right) \rho(\boldsymbol{x}) d\boldsymbol{x}, \tag{11.2}$$

with $\Pi(\rho)$ describing the dependency of the free energy $\mathcal{F}[\rho]$ on the local density field $\rho(\boldsymbol{x})$, $V(\boldsymbol{x})$ accounting for the effects of external potentials, K denoting a function depending on the convolution of $\rho(\boldsymbol{x})$ with the symmetric kernel $W(\boldsymbol{x})$ accounting for the interparticle potential. For simplicity, we introduce the constant β, defined as $\beta = (k_B T)^{-1}$.

The mean-field limit of (11.1) in which no stochastic flux is present has received a great deal of attention in the context of gradient flows. As discussed in [123], the most likely path, in the weak noise limit, minimizes the Lagrangian defined as $\mathcal{L} = \| \partial_t \rho - \nabla \cdot \left(\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} \right) \|_{(\sigma \sigma^*)^{-1}}$, where σ is the operator acting on the noise $\boldsymbol{\mathcal{W}}(\boldsymbol{x}, t)$. Thus, the most-likely solution $\langle \rho \rangle$ satisfies

$$\partial_t \langle \rho \rangle = \nabla \cdot \left(\langle \rho \rangle \nabla \frac{\delta \mathcal{F}[\langle \rho \rangle]}{\delta \langle \rho \rangle} \right). \tag{11.3}$$

Equation (11.3) is a generalized diffusion equation, which results in the heat equation if an ideal gas free energy is selected. It has been widely employed not only in the framework of DDFT [148,

149, 291], but also to model thin-liquid films stochastic dynamics[122]. It has the structure of a gradient flow in the Wasserstein metric [224, 273] with applications in a variety of contexts such as granular media [68], materials science and biological swarming [21, 68, 240]. The fundamental property of (11.3) is that the free energy (11.2) is minimized following the decay rate [67, 68, 70]

$$\frac{d}{dt}\mathcal{F}[\langle\rho\rangle] = -\int_{\mathbb{R}^d}\langle\rho\rangle \left|\frac{\delta\mathcal{F}[\langle\rho\rangle]}{\delta\langle\rho\rangle}\right|^2 d\boldsymbol{x}, \tag{11.4}$$

where the variation of the free energy $\mathcal{F}[\rho]$ with respect to the density ρ in the case of (11.2) satisfies

$$\frac{\delta\mathcal{F}[\rho]}{\delta\rho} = f'(\rho) + V(\boldsymbol{x}) + K * (g'(K * \rho)\rho) + g(K * \rho). \tag{11.5}$$

The decay rate in (11.4) is not satisfied by the stochastic gradient flow in (11.1), where occasional free-energy increase can take place during the dynamical evolution. It is precisely these jumps that allow the system to overcome energy barriers leading to phenomena such as phase transitions.

Another quantity of interest in SPDEs of the form (11.1) is the structure factor [120, 178], which represents an important measure of the stochastic properties of the system and it can be experimentally obtained. It is formally defined as the variance of the Fourier transform of the density fluctuations, and we provide its derivation in Appendix 11.A. It is a valuable quantity not only to study the stability of the numerical integrator, but also to compare different schemes.

11.1 Numerical scheme

The one-dimensional (1D) version of (11.1) can be written as

$$\partial_t\rho = \partial_x F_d(\rho) + \partial_x F_s(\rho, \mathcal{W}), \tag{11.6}$$

where F_d and F_s denote the deterministic and stochastic fluxes, respectively,

$$F_d = \rho\partial_x\frac{\delta\mathcal{F}[\rho]}{\delta\rho}, \quad F_s = \sqrt{\rho/\beta}\mathcal{W}. \tag{11.7}$$

The finite-volume formulation of (11.6) is obtained by dividing the domain into grid cells $C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, each one assumed to have the same length $\Delta x = x_{i+1/2} - x_{i-1/2}$, and then approximating in each of them the cell average of ρ defined as

$$\rho_i(t) = \frac{1}{\Delta x}\int_{x_{i-1/2}}^{x_{i+1/2}}\rho(x,t)dx. \tag{11.8}$$

Subsequently, one has to integrate (11.6) spatially over each cell and apply the Gauss divergence theorem, leading eventually to the semi-discrete equation for the temporal evolution of the cell average density,

$$\frac{d\rho_i}{dt} = \frac{F_{d,i+1/2} - F_{d,i-1/2}}{\Delta x} + \frac{F_{s,i+1/2} - F_{s,i-1/2}}{\Delta x}, \tag{11.9}$$

where $F_{d,i+1/2}$ and $F_{s,i+1/2}$ denote the deterministic and stochastic fluxes (11.7) evaluated at the boundary $x_{i+1/2}$. The separation of the physical flux into deterministic and stochastic parts has been effectively applied in previous studies [27, 120], noting though that some of them consider a single flux combining the deterministic and stochastic terms [213]. Here we treat them separately. We now proceed to develop in detail the methodology of our finite-volume scheme.

11.1.1 Deterministic flux

The deterministic flux is evaluated by employing a hybrid method which adopts a central or upwind approximation depending on the relative local total variation of the density. This is a classical technique in deterministic fluid dynamics to construct high-resolution and oscillation-free schemes [265]. On the one hand, central high-order and non-diffusive schemes are applied wherever smooth gradients of the density are found. On the other hand, a diffusive upwind scheme is employed in those regions of the domain with density gradients, in order to prevent the spurious oscillations from central high-order schemes.

Previous works in the field of FH [27, 120, 178] approximate the deterministic flux with a simple second-order central difference approach, even though high-order differences are also proposed but not implemented [120]. Our motivation to adopt a hybrid approach is precisely aimed to avoid possible spurious oscillations. The previous literature is mainly focused on FH with $f'(\rho) = \log \rho$ in (11.5), resulting in a deterministic flux of the form $\partial_x F_d(\rho) = \partial_{xx} \rho$. The treatment of this Laplacian with a central approximation works well for the cases presented in the literature, but as it is shown later in Figure 11.1.1, it can cause spurious oscillations for some solutions.

In the case of (11.6), the stochastic flux leads to non-smooth density. The proposed gradient scheme then compares the local gradient in the density with the neighbouring gradients. When the local gradient is large when compared to the neighbours', an upwind approximation is chosen. If not, the central approximation prevails. As a result, our proposed hybrid scheme for the deterministic flux satisfies

$$F_{d,i+1/2} = \left(1 - \phi(r_{i+1/2})\right) F_{d,i+1/2}^c + \phi(r_{i+1/2}) F_{d,i+1/2}^u, \tag{11.10}$$

where $\phi(r_{i+1/2})$ is a flux limiter with a threshold parameter k, defined as

$$\phi(r_{i+1/2}) = \begin{cases} 0, & \text{if } r_{i+1/2} \leq k, \\ 1, & \text{if } r_{i+1/2} > k, \end{cases}$$

and $r_{i+1/2}$ is a quotient measuring the relative local variation of the density,

$$r_{i+1/2} = \frac{|\rho_{i+1} - \rho_i|}{\sum_{l=i-w}^{i+w} |\rho_{l+1} - \rho_l|}, \tag{11.11}$$

with w indicating the number of neighbouring cell used to compute the total variation. A value $w = 5$ is employed in the numerical experiments of this work, since it gives a good compromise between conservation of local information and effects of the fluctuations.

The threshold parameter k plays a key role and has to be carefully selected. When k is small, the diffusive upwind scheme is chosen more frequently, leading to diffusive behaviour which affects the structure factor and the correlations. On the contrary, when k is large, the central scheme will be predominant, and spurious oscillations may be created. Figure 11.1.1 provides a numerical example to choose an adequate value for k.

Firstly, Figures 11.1.1a and 11.1.1b are obtained by simulating (11.6) with a free energy satisfying $\delta \mathcal{F}/\delta \rho = \log \rho + 0.1x$. The initial density profile has two discontinuities as shown in Figure 11.1.1a. Under these conditions, the numerical solution evolves as a diffusive travelling wave, but the two discontinuities in the initial density trigger spurious oscillations. The oscillations diminish by reducing k (for $k = 0$, which corresponds to only upwind flux, the diffusion eliminates the oscillations). However, a low value of k critically dampens the variance, due to the diffusive nature of the upwind flux, as it is noticed from Figure 11.1.1b.

Secondly, Figure 11.1.1c is obtained from simulating (11.6) with a free energy satisfying $\delta \mathcal{F}/\delta \rho = \log \rho$ and starting from an equilibrium density profile. For this case, the theoretical value of the

structure factor is known and is given by (11.78), meaning that the dampening behaviour of the upwind scheme could be directly evaluated from Figure 11.1.1c. We notice again how the upwind scheme dampens the statistical properties of the system due to the numerical diffusion. As a result, an intermediate value of k needs to be taken in order to find a balance between both numerical flaws. The compromising value is chosen to be $k = 3$.
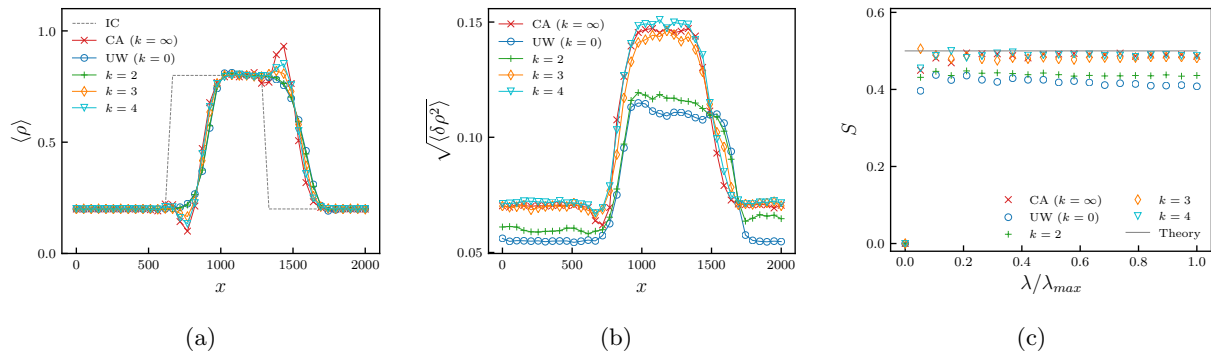

Figure 11.1.1: (a) Mean density and (b) standard deviation for a moving and diffusing initial step function evolving according to (11.6), with $\delta\mathcal{F}/\delta\rho = \log\rho + u_0 x$ and $u_0 = 0.1$. Additionally, we report the structure factor of a uniform system in (c). IC: initial condition, CA: central approximation ($k = \infty$), UW: upwind approximation ($k = 0$). Further values of k are depicted to evaluate the spurious oscillations in the density and the artificial fluctuation dampening in case of both inhomogeneous (a-b) and homogeneous (c) systems. In what follows, we adopt a scheme with $k = 3$, since it gives the compromise between accuracy in the sharp density profile and the fluctuations amplitude. We note that the stochastic term $\partial_x F_s(\rho, \mathcal{W})$ is included in these simulations and is discretized according to (11.24).

After selecting the adequate value of k, we proceed to the detailed construction of the central and upwind deterministic fluxes in (11.10):

a) Upwind approximation of the deterministic flux: it is constructed as proposed in [55], where a first- and second-order finite-volume method for nonlinear equations with gradient-flow structure is constructed. The equations treated in [55] have the form (11.6) without the white noise $\mathcal{W}$. The authors propose to firstly reconstruct the density profile in each cell C_i as a constant profile for the first-order scheme, or as a linear profile for the second-order scheme,

$$\widetilde{\rho}_i(x) = \begin{cases} \rho_i, & x \in C_i, \quad \text{for the first-order scheme,} \\ \rho_i + (\rho_x)_i (x - x_i), & x \in C_i, \quad \text{for the second-order scheme,} \end{cases} \tag{11.12}$$

so that the east and the west density values ρ_i^E and ρ_i^W at the cell interfaces $x_{i+\frac{1}{2}}$ and $x_{i-\frac{1}{2}}$, respectively, are approximated as

$$\begin{aligned} \rho_i^E &= \rho_i + \frac{\Delta x}{2} (\rho_x)_i, \\ \rho_i^W &= \rho_i - \frac{\Delta x}{2} (\rho_x)_i. \end{aligned} \tag{11.13}$$

The numerical derivatives $(\rho_x)_i$ at every cell C_i are computed by means of an adaptive procedure which ensures that the point values (11.13) are second-order and non-negative. This

procedure initially takes centred approximations of the form $(\rho_x)_i = (\rho_{i+1} - \rho_{i-1})/(2\Delta x)$. If it then happens that $\rho_i^E < 0$ or $\rho_i^W < 0$, the scheme employs a *minmod* limiter which ensures that the reconstructed values are non-negative as long as the cell averages ρ_i are non-negative,

$$(\rho_x)_i = \text{minmod}\left(\theta\,\frac{\overline{\rho}_{i+1} - \rho_i}{\Delta x},\, \frac{\overline{\rho}_{i+1} - \overline{\rho}_{i-1}}{2\Delta x},\, \theta\,\frac{\rho_i - \overline{\rho}_{i-1}}{\Delta x}\right), \tag{11.14}$$

where

$$\text{minmod}\,(z_1, z_2, \ldots) = \begin{cases} \min\,(z_1, z_2, \ldots), & \text{if } z_i > 0 \quad \forall i, \\ \max\,(z_1, z_2, \ldots), & \text{if } z_i < 0 \quad \forall i, \\ 0, & \text{otherwise.} \end{cases}$$

The parameter θ controls the numerical viscosity and it is taken to be $\theta = 2$, as in [55].

After completing the density reconstruction, the deterministic flux $F_{d,i+1/2}^u$ is evaluated with an upwind scheme as

$$F_{d,i+1/2}^u = u_{i+1/2}^+\,\rho_i^E + u_{i+1/2}^-\,\rho_{i+1}^W, \tag{11.15}$$

where $u_{i+1/2}$ are discrete values computed from the central difference

$$u_{i+1/2} = -\frac{\left(\dfrac{\delta\mathcal{F}}{\delta\rho}\right)_{i+1} - \left(\dfrac{\delta\mathcal{F}}{\delta\rho}\right)_i}{\Delta x}. \tag{11.16}$$

The upwind formulation of the deterministic flux (11.15) is then accomplished by taking

$$u_{i+1/2}^+ = \max\,(u_{i+1/2}, 0) \quad \text{and} \quad u_{i+1/2}^- = \min\,(u_{i+1/2}, 0). \tag{11.17}$$

Finally, the discrete variation of the free energy with respect to the density $\left(\dfrac{\delta\mathcal{F}}{\delta\rho}\right)_i$ is computed from (11.5), in the case $g(s) = s$, as

$$\left(\frac{\delta\mathcal{F}}{\delta\rho}\right)_i = \Delta x \sum_i K(x_i - x_i)\rho_i + F(\rho_i) + V(x_i). \tag{11.18}$$

For general nonlinearities $g(s)$ a similar treatment is performed.

b) Central approximation for the deterministic flux: this is the main strategy to treat the FH deterministic flux in the literature [27, 120, 178]. In our case, given the generality of the free energy in (11.2), we propose to evaluate the central deterministic flux as

$$F_{d,i+1/2}^c = u_{i+1/2}\,\rho_{i+1/2}, \tag{11.19}$$

where $u_{i+1/2}$ is computed as in (11.16), with the discrete variation of the free energy satisfying (11.18), and $\rho_{i+1/2}$ is taken as the averaged from the adjacent cells,

$$\rho_{i+1/2} = \frac{\overline{\rho}_i + \overline{\rho}_{i+1}}{2}. \tag{11.20}$$

Classical hybrid schemes employ a high-order approximation for the central approximation of the deterministic flux. For this work, however, we just consider the low-order differences (11.16) and (11.20), given that the presence of the stochastic flux limits the spatial order of accuracy. Previous works in the literature also propose this low-order central differences [27, 120, 178].

11.1.2 Stochastic flux

The evaluation of the stochastic flux (11.7) must be done carefully since the divergence of the white noise $\mathcal{W}$ cannot be evaluated pointwise in time and space. This problem is typically overcome by evaluating the noise in the cell by means of a spatiotemporal average, following [120] and subsequently employed by Donev and collaborators in [18, 35, 178] ,

$$\mathcal{W}_i = \frac{1}{\Delta x \Delta t} \int_t^{t+\Delta t} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathcal{W}(x,t) dx \, dt, \qquad (11.21)$$

which, by the definition of the white noise, is equal to a normal distribution with zero mean and variance $(\Delta x \Delta t)^{-1}$, so that

$$\mathcal{W}_i = \mathcal{N}(0,1)/\sqrt{\Delta x \Delta t}. \qquad (11.22)$$

Several approximations for the stochastic flux have been put forward in the literature [27, 178]. They rely on computing the stochastic flux directly at the interfaces using a random number generator, and we refer the reader to [120] for more details about this approach. In this work, however, we aim to employ the spatiotemporal cell average in (11.21) to compute the stochastic fluxes at the interfaces. We are inspired by the literature on numerical methods for hyperbolic problems where it is common to evaluate fluxes in a central or upwind fashion. Of course, here we are not aiming to achieve a higher accuracy at the interface, given that the cell averages are sampled from a distribution. We test the following four different approximations for the stochastic flux, which are compared in Section 11.2:

(a) Forward approximation of the form

$$F_{s,i+1/2} = \left(\sqrt{\frac{\rho}{\beta}} \mathcal{W} \right)_{i+1/2} = \sqrt{\frac{\rho_i}{\beta}} \mathcal{W}_i. \qquad (11.23)$$

(b) Linear approximation of the form

$$F_{s,i+1/2} = \left(\sqrt{\frac{\rho}{\beta}} \mathcal{W} \right)_{i+1/2} = \sqrt{\frac{\rho_{i+1/2}}{\beta}} \mathcal{W}_{i+1/2}, \qquad (11.24)$$

where

$$\rho_{i+1/2} = \frac{\rho_i + \rho_{i+1}}{2}, \quad \mathcal{W}_{i+1/2} = \frac{\mathcal{W}_i + \mathcal{W}_{i+1}}{2}. \qquad (11.25)$$

(c) Parabolic approximation of the form

$$F_{s,i+1/2} = \left(\sqrt{\frac{\rho}{\beta}} \mathcal{W} \right)_{i+1/2} = \sqrt{\frac{\rho_{i+1/2}}{\beta}} \mathcal{W}_{i+1/2}, \qquad (11.26)$$

where

$$\begin{aligned} \rho_{i+1/2} &= \alpha_1 \left(\rho_{i-1} + \rho_{i+2} \right) + \alpha_2 \left(\rho_i + \rho_{i+1} \right), \\ \mathcal{W}_{i+1/2} &= \alpha_1 \left(\mathcal{W}_{i-1} + \mathcal{W}_{i+2} \right) + \alpha_2 \left(\mathcal{W}_i + \mathcal{W}_{i+1} \right), \\ \alpha_1 &= (1 - \sqrt{3})/4, \quad \alpha_2 = (1 + \sqrt{3})/4. \end{aligned} \qquad (11.27)$$

The coefficients α_1 and α_2 are selected as in [27], with the objective of preserving both the average and the variance in each time step.

(d) Upwind approximation, where $\mathcal{W}_i$ is taken as the stochastic velocity, so that a similar expression to the deterministic flux in (11.15) is taken,

$$F_{s,i+1/2} = \left(\sqrt{\frac{\rho}{\beta}} \mathcal{W} \right)_{i+1/2} = \sqrt{\frac{\rho_i^E}{\beta}} \mathcal{W}_{i+1/2}^+ + \sqrt{\frac{\rho_{i+1}^W}{\beta}} \mathcal{W}_{i+1/2}^-, \tag{11.28}$$

where

$$\mathcal{W}_{i+1/2}^+ = \max \left(\mathcal{W}_{i+1/2}, 0 \right), \quad \mathcal{W}_{i+1/2}^- = \min \left(\mathcal{W}_{i+1/2}, 0 \right), \tag{11.29}$$

and $\mathcal{W}_{i+1/2} = (\mathcal{W}_i + \mathcal{W}_{i+1})/2$. The east and west density values ρ_i^E and ρ_i^W are computed as in the deterministic flux, either with a first- or second-order reconstruction (11.13).

11.1.3 Stochastic time integrators

The derivation of the temporal integrators to advance in time the semidiscrete equation (11.9) is accomplished by the equation

$$d\overline{\boldsymbol{\rho}}(t) = \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t)) \, dt + \boldsymbol{\sigma}(\overline{\boldsymbol{\rho}}(t)) \, \overline{\boldsymbol{\mathcal{W}}} \, dt, \tag{11.30}$$

where the vectors $\overline{\boldsymbol{\rho}}(t)$ and $\overline{\boldsymbol{\mathcal{W}}}$ contain the cell averages defined in (11.8) and (11.21), respectively, so that $\overline{\boldsymbol{\rho}}(t) = (\rho_1(t), \rho_2(t), \dots, \rho_n(t))$ and $\overline{\boldsymbol{\mathcal{W}}}(t) = (\mathcal{W}_1(t), \mathcal{W}_2(t), \dots, \mathcal{W}_n(t))$. The vector $\boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))$ and the matrix $\boldsymbol{\sigma}(\overline{\boldsymbol{\rho}}(t))$ depend on the density cell averages $\overline{\boldsymbol{\rho}}(t)$ and their structures vary depending on the choice of the deterministic and stochastic fluxes, respectively.

From (11.30) we employ Itô's lemma to approximate the two functions $\boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))$ and $\boldsymbol{\sigma}(\overline{\boldsymbol{\rho}}(t))$. After integrating in time then we obtain the Taylor expansion of the stochastic process. Truncating this expansion with an error $\mathcal{O}(\Delta t^{1/2})$ and integrating between t and $t + \Delta t$, one can derive the following family of implicit-explicit Euler-Maruyama integrators [182], whose component-wise form satisfies

$$\rho_i(t + \Delta t) = \rho_i(t) + [(1 - \theta) \, \mu_i(\overline{\boldsymbol{\rho}}(t)) + \theta \, \mu_i(\overline{\boldsymbol{\rho}}(t + \Delta t))] \, \Delta t + \sum_{k=1}^n \sigma_{ik}(\overline{\boldsymbol{\rho}}(t)) \overline{W}_k(t) \Delta t. \tag{11.31}$$

The parameter θ allows us to have an explicit ($\theta = 0$), implicit ($\theta = 1$) or semi-implicit ($\theta = 0.5$) temporal integrator. Euler-Maruyama is the highest order integrator for which no multiple stochastic integrals have to be computed, but it has only 0.5 strong order of convergence.

Keeping in the expansion all the terms up to $\mathcal{O}(\Delta t)$, one obtains a derivative-free family of implicit-explicit Milstein integrators with strong order 1.0 and weak order 0.5 [182]. The component-wise version of this scheme is

$$\begin{aligned} \rho_i(t + \Delta t) =& \rho_i(t) + [(1 - \theta) \, \mu_i(\overline{\boldsymbol{\rho}}(t)) + \theta \, \mu_i(\overline{\boldsymbol{\rho}}(t + \Delta t))] \, \Delta t + \sum_{k=1}^n \sigma_{ik}(\overline{\boldsymbol{\rho}}(t)) \overline{W}_k(t) \Delta t \\ &+ \frac{1}{\sqrt{\Delta t}} \sum_{l,m=1}^n [\sigma_{im}(\boldsymbol{\Upsilon}_l(t)) - \sigma_{il}(\overline{\boldsymbol{\rho}}(t))] \, I_{l,m}(t), \end{aligned} \tag{11.32}$$

where the l-th row of the matrix $\boldsymbol{\Upsilon}$ is defined as

$$\boldsymbol{\Upsilon}_l(t) = \overline{\boldsymbol{\rho}}(t) + \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t)) \Delta t + \boldsymbol{\sigma}_l(\overline{\boldsymbol{\rho}}(t)) \sqrt{\Delta t}, \tag{11.33}$$

and multiple stochastic integrals $I_{l,m}(t) = \int_t^{t+\Delta t} \mathcal{W}^l \, \mathcal{W}^m dt$, where $\mathcal{W}_l$ and $\mathcal{W}_m$ are two white noises. These integrals do not have a simple analytical solution, thus are approximated as a function of the white noise cells average in (11.21) as [203]:

$$
I_{l,m}(t) = \begin{cases} \frac{1}{2}\left[\left(\overline{W}^l\right)^2 - 1\right]\Delta t & \text{if } l = m, \\[2ex] \frac{\Delta t}{2}\overline{W}^l\overline{W}^m + \sqrt{k_p\Delta t}(\varphi_l\overline{W}^m - \varphi_m\overline{W}^l) \\[1ex] \qquad + \sum_{r=1}^p \frac{1}{2\pi r}\left[\zeta_{lr}(\sqrt{2}\,\overline{W}^m\sqrt{\Delta t} + \eta_m) - \zeta_{mr}(\sqrt{2}\,\overline{W}^l\sqrt{\Delta t} + \eta_l)\right] & \text{otherwise,} \end{cases}
$$
(11.34)

where φ_l, ζ_{lr} and η_m are pairwise independent variables with distribution $\mathcal{N}(0,\Delta t)$ and k_p is given by

$$
k_p = \frac{1}{12} - \frac{1}{2\pi^2}\sum_1^p \frac{1}{r^2}.
$$
(11.35)

The value p determines the accuracy of the multiple stochastic integral approximation and subsequently the accuracy of the scheme. A value of $p = k/\Delta t$ for some constant k is enough to preserve the accuracy of the scheme [182].

Stochastic time integration schemes of higher strong order have also been proposed in the literature [182]. However, these schemes are very computationally expensive due to the presence of high-order multiple stochastic integrals to be solved. Moreover, in many physical applications, the convergence in probability, also called weak convergence, is more relevant than the strong convergence. For this reason, a last time integration scheme we will study the following explicit weak order 2.0 Runge-Kutta scheme:

$$
\begin{aligned}
\widetilde{\rho}_i(t+\Delta t) &= \rho_i(t) + \frac{1}{2}\left[\mu_i(\mathbf{\Upsilon}(t)) + \mu_i(\overline{\boldsymbol{\rho}}(t))\right]\Delta t + \Phi(t), \\
\rho_i(t+\Delta t) &= \rho_i(t) + \frac{1}{2}\left[\mu_i(\widetilde{\boldsymbol{\rho}}(t+\Delta t)) + \mu_i(\overline{\boldsymbol{\rho}}(t))\right]\Delta t + \Phi(t),
\end{aligned}
$$
(11.36)

where the vector $\Phi(t)$ has components:

$$
\begin{aligned}
\Phi_i(t) = &\frac{1}{4}\sum_{l=1}^n \left[\boldsymbol{\sigma}_{li}(\mathbf{\Lambda}_{l+}(t)) + \boldsymbol{\sigma}_{li}(\mathbf{\Lambda}_{l-}(t)) + 2\boldsymbol{\sigma}_{li}(\overline{\boldsymbol{\rho}}(t))\right]\overline{W}^l(t)\sqrt{\Delta t} \\
&+ \frac{1}{4}\sum_{l=1}^n\sum_{r=1,r\neq l}^n \left[\boldsymbol{\sigma}_{li}(\mathbf{\Xi}_{r+}(t)) + \boldsymbol{\sigma}_{li}(\mathbf{\Lambda}_{r-}(t)) - 2\boldsymbol{\sigma}_{li}(\overline{\boldsymbol{\rho}}(t))\right]\overline{W}^l(t) \\
&+ \frac{1}{4}\sum_{l=1}^n \left[\boldsymbol{\sigma}_{li}(\mathbf{\Lambda}_{l+}(t)) - \boldsymbol{\sigma}_{li}(\mathbf{\Lambda}_{l-}(t))\right]\left[\left(\overline{W}^l(t)\right)^2 - 1\right]\sqrt{\Delta t} \\
&+ \frac{1}{4}\sum_{l=1}^n\sum_{r=1,r\neq l}^n \left[\boldsymbol{\sigma}_{li}(\mathbf{\Xi}_{r+}(t)) - \boldsymbol{\sigma}_{li}(\mathbf{\Xi}_{r-}(t))\right]\left[\overline{W}^l(t)\overline{W}^r(t) + V_{r,i}\right]\sqrt{\Delta t},
\end{aligned}
$$
(11.37)

and the supporting values:

$$
\mathbf{\Upsilon} = \overline{\boldsymbol{\rho}}(t) + \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))\Delta t + \sum_{i=1}^n \boldsymbol{\sigma}_i(\boldsymbol{\rho}(t))\Delta W_i(t),
$$
(11.38)

$$
\mathbf{\Lambda}_{l\pm} = \overline{\boldsymbol{\rho}}(t) + \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))\Delta t \pm \boldsymbol{\sigma}_l(\overline{\boldsymbol{\rho}}(t))\sqrt{\Delta t},
$$
(11.39)

$$
\mathbf{\Xi}_{l\pm} = \overline{\boldsymbol{\rho}}(t) \pm \boldsymbol{\sigma}_l(\overline{\boldsymbol{\rho}}(t))\sqrt{\Delta t}.
$$
(11.40)

The random matrix $\mathbf{V}$ is defined as:

$$V_{r,i}(t) = \begin{cases} \pm 1 & \text{with} \quad p = \frac{1}{2} \quad \text{if} \quad r < i, \\ -1 & \text{if} \quad r = i, \\ -V_{i,r}(t) & \text{if} \quad r > i, \end{cases} \tag{11.41}$$

where p indicates the probability. It has to be emphasised that such a scheme does not involve the computation of multiple stochastic integrals, thus its strong order of convergence is expected to be at most 1.0.

11.1.3.1 Weak and strong order of convergence for temporal integrators

The order of convergence can be measured in the strong and weak sense, for which the strong and weak errors are respectively defined for a particular time τ and a group of trajectories $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_m\}$ as

$$\epsilon_s = \langle |\overline{\boldsymbol{\rho}}^\gamma(\tau) - \overline{\boldsymbol{\rho}}^\gamma_{exact}(\tau)| \rangle_{\gamma \in \Gamma} \quad \text{and} \quad \epsilon_w = \left| \langle \overline{\boldsymbol{\rho}}^\gamma(\tau) \rangle_{\gamma \in \Gamma} - \langle \overline{\boldsymbol{\rho}}^\gamma_{exact}(\tau) \rangle_{\gamma \in \Gamma} \right|, \tag{11.42}$$

where $\overline{\boldsymbol{\rho}}^\gamma(\tau)$ refers to the numerical density cell averages at time τ following trajectory γ, $\overline{\boldsymbol{\rho}}^\gamma_{exact}(\tau)$ denotes the exact or reference solution which is considered to be the true solution of the stochastic equation, the ensemble average $\langle \cdot \rangle$ is taken over the trajectories $\gamma \in \Gamma$, and the norm $|\cdot|$ is taken to be the standard L^1-norm.

In Figure 11.1.2 we evaluate the strong and weak errors for the described stochastic integrators. They are obtained by simulating equation (11.30) in the simplified case of geometric Brownian motion, for which $\boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t)) = -\overline{\boldsymbol{\rho}}(t)$ and $\boldsymbol{\sigma}(\overline{\boldsymbol{\rho}}(t)) = 0.5\overline{\boldsymbol{\rho}}(t)$, thus eliminating the spatial derivatives. As a result, the temporal evolution of the density for a cell j, which is independent from the rest of cells, follows

$$d\rho_i(t) = -\rho_i dt + 0.5\rho_i \mathcal{W}_i dt, \tag{11.43}$$

with the cell averaged white noise $\mathcal{W}_i$ defined as in (11.21). For the simulation we selected $\rho_i(0) = 1$. Geometric Brownian motion is useful to compute the strong and weak errors since the exact solution in analytically known [222].

The results in Figures 11.1.2a and 11.1.2b depict the strong and weak order of convergence for the temporal integrators. Concerning the former, as expected the Euler-Maruyama presents an order of 0.5, while Milstein an order of 1.0. Runge-Kutta is expected to have a strong order of at least 0.5, and in the plot it approaches a value of 1.0.

With respect to the weak order, the whole families of Euler-Maruyama and Milstein solvers are expected to have an order of 1.0, while the Runge-Kutta an order of 2.0. Such theoretical predictions are respected for all schemes, with the exception of the semi-implicit methods which outperforms, giving an order between 1.0 and 2.0.

In Figure 11.1.2c we plot the CPU time against the total number of cells n for each of the temporal integrators. The Euler-Maruyama accounts for $\mathcal{O}(n)$ computations, the Milstein for $\mathcal{O}(n^2)$, and the Runge-Kutta for $\mathcal{O}(n^3)$. However, for $n < 100$ we get a lower CPU time for Runge-Kutta, if compared with all the other integrators except for the explicit Euler-Maruyama.

11.1.4 Positivity of the density through an adaptive time step

A natural constraint for physical systems is the positivity of the density field, and the numerical solution is expected to satisfy such a requirement. Numerical schemes with the property of preserving the positivity of the density have been developed in the literature, specially in the context
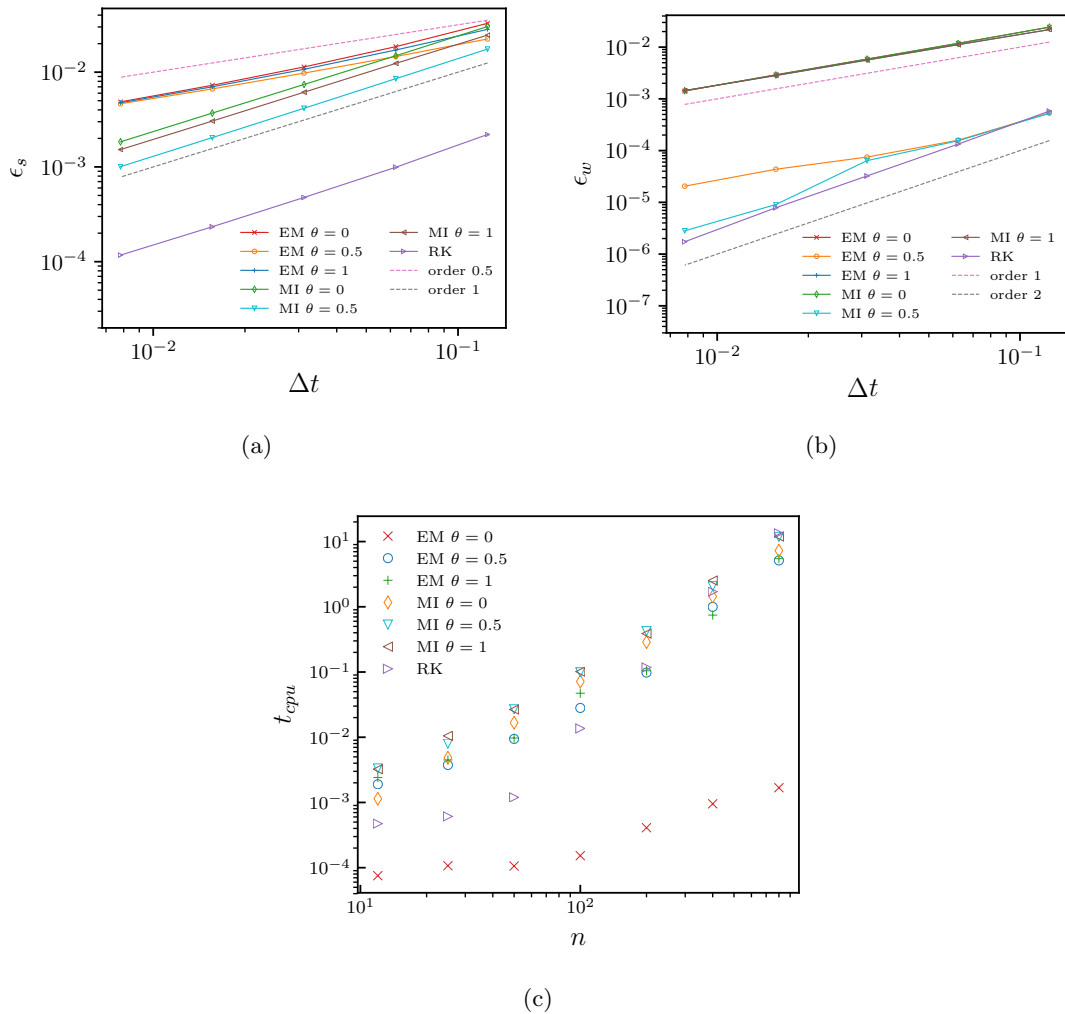
(a)

(b)

(c)

Figure 11.1.2: Strong (a) and weak (b) errors convergence for geometric Brownian motion. In (c) we report the CPU time for each time integration schemes as a function of the number of cells n. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$).

of deterministic conservation law PDEs [34, 55]. The strategy is usually to derive a Courant-Friedrichs-Lewy (CFL) condition which imposes a constraint for the maximum Δt so that the density always remains non-negative.

For particular discretizations of the numerical fluxes (11.7) it is also possible to derive a CFL condition for the SPDE in (11.1). This CFL condition depends on the Gaussian distributions from the white noise (11.22), as well as on the density profile. Following the derivation provided in [55] for a deterministic gradient-flow equation, we proceed to provide an example of the CFL derivation when the upwind discretizations (11.15) and (11.28) for the deterministic and stochastic fluxes, respectively, are employed.

Lemma 11.1.1. *Consider the SPDE* (11.1) *with initial data* $\rho_0(x) > 0$, *together with the semi-discrete finite-volume scheme* (11.9) *with the upwind discretizations for the deterministic* (11.15)-(11.18) *and stochastic* (11.22), (11.28) *fluxes. Assume that the SPDE is temporally discretized with a deterministic Euler forward method. Then, the computed cell averages satisfy* $\rho_i \geq 0$, $\forall j$,

provided that the following two CFL conditions for Δt hold:

$$\frac{1}{2} - \lambda_1 u^+_{i+\frac{1}{2}} - \lambda_2 \mathcal{G}^+_{i+\frac{1}{2}} / \sqrt{\rho_i^E \beta} \geq 0, \quad \frac{1}{2} - \lambda_1 u^-_{i-\frac{1}{2}} - \lambda_2 \mathcal{G}^-_{i-\frac{1}{2}} / \sqrt{\rho_i^W \beta} \geq 0, \qquad (11.44)$$

where

$$\lambda_1 := \frac{\Delta t}{\Delta x}, \quad \lambda_2 := \sqrt{\frac{\Delta t}{\Delta x}}, \quad \overline{\mathcal{G}}_i = \overline{\mathcal{W}}_i \sqrt{\Delta x \Delta t} = \mathcal{N}(0,1), \qquad (11.45)$$

and $\mathcal{G}^+_{i+1/2}$, $\mathcal{G}^-_{i+1/2}$ are constructed as in (11.29), so that

$$\mathcal{G}^+_{i+1/2} = \max\left(\mathcal{G}_{i+1/2}, 0\right), \quad \mathcal{G}^-_{i+1/2} = \min\left(\mathcal{G}_{i+1/2}, 0\right). \qquad (11.46)$$

Proof. Assume that for a given time t the computed solution for the density is known and positive: $\rho_i(t) \geq 0, \forall j$. The new cell averages following a forward Euler temporal scheme in the finite-volume formulation (11.9) satisfy

$$\rho_i(t + \Delta t) = \rho_i(t) - \Delta t \left[\frac{F_{d,i+1/2} - F_{d,i-1/2}}{\Delta x} + \frac{F_{s,i+1/2} - F_{s,i-1/2}}{\Delta x}\right]. \qquad (11.47)$$

Substituting the deterministic and stochastic fluxes for their upwinded discretizations (11.15) and (11.28), respectively, and by employing the notation specified in (11.45), it follows

$$
\begin{aligned}
\rho_i(t + \Delta t) =& \frac{1}{2}(\rho_i^E + \rho_i^W) - \lambda_1 \left[u^+_{i+\frac{1}{2}}\rho_i^E + u^-_{i+\frac{1}{2}}\rho_{i+1}^W - u^+_{i-\frac{1}{2}}\rho_{i-1}^E - u^-_{i-\frac{1}{2}}\rho_i^W\right] \\
& - \lambda_2 \left[\mathcal{G}^+_{i+\frac{1}{2}}\sqrt{\rho_i^E/\beta} + \mathcal{G}^-_{i+\frac{1}{2}}\sqrt{\rho_{i+1}^W/\beta} - \mathcal{G}^+_{i-\frac{1}{2}}\sqrt{\rho_{i-1}^E/\beta} - \mathcal{G}^-_{i-\frac{1}{2}}\sqrt{\rho_i^W/\beta}\right] \\
=& \lambda_1 \left[-u^-_{i+\frac{1}{2}}\rho_{i+1}^W + u^+_{i-\frac{1}{2}}\rho_{i-1}^E\right] + \lambda_2 \left[-\mathcal{G}^-_{i+\frac{1}{2}}\sqrt{\rho_{i+1}^W/\beta} + \mathcal{G}^+_{i-\frac{1}{2}}\sqrt{\rho_{i-1}^E/\beta}\right] \\
& + \left[\frac{1}{2} - \lambda_1 u^+_{i+\frac{1}{2}} - \lambda_2 \mathcal{G}^+_{i+\frac{1}{2}}/\sqrt{\rho_i^E\beta}\right]\rho_i^E + \left[\frac{1}{2} - \lambda_1 u^-_{i-\frac{1}{2}} - \lambda_2 \mathcal{G}^-_{i-\frac{1}{2}}/\sqrt{\rho_i^W\beta}\right]\rho_i^W.
\end{aligned} \qquad (11.48)
$$

Due to the fact that the reconstructed point values for the density ρ_{i-1}^E, ρ_{i+1}^E, ρ_i^W and ρ_{i+1}^W are non-negative, and bearing in mind that $u^+_{i-\frac{1}{2}}$, $\mathcal{G}^+_{i-\frac{1}{2}} \leq 0$ and $u^-_{i+\frac{1}{2}}$, $\mathcal{G}^-_{i+\frac{1}{2}} \geq 0$ due to (11.17) and (11.46), it follows $\rho_i(t + \Delta t) \geq 0$, $\forall j$, provided that the CFL conditions (11.44) hold. $\qquad \square$

The CFL conditions in (11.44) ensure that the density remains non-negative at all times, independently of the values produced by the normal distributions of the white noise spatio-temporal average (11.22). In the case of a rare event in which the Gaussian distribution produces low-probability values located at the tails of the distribution, Δt would be adapted accordingly to ensure the positivity. However, this adaptive time step strategy entails two main disadvantages. First, it requires at each time step the solution of a second-order equation (in 1D) or a two-parameter equation in multi-dimensional problems. Second, since the time-step size is dependent on the random number at each step, higher (or lower) Δt may be favoured by some random numbers, thus not guaranteeing that the correct Brownian path is followed [136].

Previous works in the literature have already addressed the issue of positivity by means of varied approaches. In the context of FH, the authors of [178] have effectively opted for introducing cutting functions based on smoothed Heavisides which prevent the density from becoming negative. The main drawbacks of this strategy are 1) despite reducing the chances of having negative density values, positive densities are not guaranteed, and 2) it affects the density distribution.

A further alternative to preserve positivity lays in the concept of Brownian trees, which were firstly introduced in [136] in order to address the numerical resolution of stochastic differential

equations with variable time steps. The key idea here is that it is vital to respect the Brownian path that is formed after evaluating the normal distributions (11.22). This means that upon advancing our simulation from time t a certain Δt_1 and realising that the density in one of the nodes j has become negative we cannot just simply repeat the time step with a shorter $\Delta t_2 < \Delta t_1$ in order to maintain positivity. The values of the normal distributions after the first trial of advancing Δt_1 have to be respected if the Brownian path is to be preserved. In addition, those values of the normal distributions at $t + \Delta t_1$ have to be employed when computing the values at $t + \Delta t_2$, even if the jump from t to $t + \Delta t_1$ has produced negative densities.

The solution to effectively take the statistical information at $t + \Delta t$ into account when repeating the time step is the so-called Brownian bridge [203, 262]. It allows the computation of $\mathcal{W}_i$ in (11.21) at an intermediate time step $t + \Delta t/2$ by means of the formula

$$\mathcal{W}_i\left(t + \frac{\Delta t}{2}\right) - \mathcal{W}_i(t) = \frac{\mathcal{W}_i\left(t + \Delta t\right) - \mathcal{W}_i\left(t\right)}{2} + \mathcal{N}\left(0, \frac{\Delta t}{4}\right). \tag{11.49}$$

As a result, our tactic consists in initially selecting an adequately small Δt. Then, if after some time the density becomes negative, Δt is halved to compute the intermediate time step from the Brownian bridge (11.49). If that intermediate state leads to further negative densities, the Brownian bridge is applied as many times as needed. The information at $t + \Delta t$ is saved to be employed once all the intermediate time steps with non-negative densities are computed. A pseudocode to implement the Brownian bridge is written in Algorithm 1. As a remark, the adequate choice of a small initial Δt for the simulation is essential to reduce the number of Brownian bridges to a minimum. A compromise is of course needed, since an extremely small Δt does not lead to negative densities but requires a high computational cost for the simulation.

Algorithm 1: Algorithm adopted to overcome the issue of negative density. It is based on an adaptive timestep combined with the Brownian bridge technique, that allows to preserve the properties of the probability distribution underlying the stochastic process.

```
    Input: ρ(t)
    Output: ρ(t + Δt)
 1  NegativeDensity=True;
 2  Δt = Δt₀;
 3  partitions = 0;
 4  while (NegativeDensity==True) do
 5      NegativeDensity=False;
 6      ρ_tmp = ρ;
 7      for i ← 0 to 2^partitions do
 8          compute Brownian bridge;
 9          update ρ_tmp;
10          if (any(ρ_tmp) < 0) then
11              NegativeDensity=True;
12          end
13      end
14      Δt ← Δt/2 ;
15      partitions ← partitions + 1;
16  end
17  ρ(t + Δt) ← ρ_tmp;
18  return ρ(t + Δt);
```

11.1.5   Boundary conditions

In this section we analyse the implementation of boundary conditions for the cases of periodic, confined and open systems. For systems with a periodic boundary, it is sufficient to impose

$$\rho_0 = \rho_N. \tag{11.50}$$

For no-flux conditions, the boundary conditions to impose on the fluxes are

$$F_{i\pm 1/2} = 0 \quad \text{for} \quad i = 0,\, N. \tag{11.51}$$

Open systems in thermal and chemical equilibrium with a reservoir can be represented by a μVT ensemble with constant grand potential $\Omega[\rho] = \mathcal{F}[\rho] - \mu \int \rho dx$, where $\mathcal{F}[\rho] = \mathcal{F}[\rho] + \int V(x)\,\rho dx$ with

$\mathcal{F}[\rho]$ being the Helmholtz free-energy functional, $V(x)$ the external potential acting on the system and μ the chemical potential. Using the fact that the functional derivative of Ω with respect to ρ is null in equilibrium, we obtain $\delta\mathcal{F}[\rho]/\delta\rho = \mu$. Since the system is assumed to be in contact with a reservoir at temperature T_{res} and chemical potential μ_{res}, the corresponding boundary condition applied to compute the velocities $u_{i+1/2}$ at the boundaries in (11.16) is

$$\left(\frac{\delta\mathcal{F}}{\delta\rho}\right)_0 = \left(\frac{\delta\mathcal{F}}{\delta\rho}\right)_N = \mu_{\text{res}}, \tag{11.52}$$

with μ_{res} being the chemical potential of the reservoir. From the value of μ_{res} one can compute the density by solving (11.5) for a fixed value of $\delta\mathcal{F}/\delta\rho$. This implies that the values of ρ_0, ρ_N and any additional ghost node are imposed from (11.52) for all times. Depending on the particular choice of free energy in (11.2), it may be possible to converge to different density profiles depending on the initial condition for the iterative algorithm to solve (11.5). This open boundary condition imposes a positive or negative flux of mass through the boundary, and as a result the total mass is not conserved in time.

11.2 Numerical validation and applications

In this section we provide tests of the numerical schemes developed in Section 11.1. Initially, in Subsection 11.2.1 we conduct a simulation with a purely-diffusive ideal-gas free energy with noise and without external fields or interparticle potentials. There are several theoretical results for such systems [27, 178] allowing us to benchmark the statistical correlation and the structure factor from our numerical schemes. Further validation of the schemes will be offered via comparison with our own MD simulations. The results of these tests show that the Runge-Kutta temporal integrator (11.36) and the linear approximation of the stochastic flux (11.24)-(11.25) accomplish the best accuracy and efficiency to simulate our SPDE (11.6). This choice is maintained in the examples that follow.

Secondly, we provide a simulation for an ideal gas with a local confining external potential $V(x)$, in order to test the mean and variance of the density, the spatial correlation and the decay of the discrete free energy in time.

The simulations of ideal gases are also compared with results from MD simulations using the software LAMMPS [237]. Details of such simulations are explained in Appendix 11.B.

11.2.1 Ideal-gas system in equilibrium

Consider the SPDE in (11.1) without any external or interaction potential ($V(x) = W(x) = 0$) and applied to the classical ideal-gas free energy

$$\mathcal{F}[\rho] = \beta^{-1} \int \rho\left(\ln(\rho) - 1\right) dx, \tag{11.53}$$

leading to a diffusive equation with multiplicative noise of the form

$$\partial_t\rho = \Delta\rho/\beta + \boldsymbol{\nabla} \cdot \left[\sqrt{\rho/\beta}\,\mathcal{W}(x, \rho)\right]. \tag{11.54}$$

The initial density profile is taken as the equilibrium one, with a constant value in all cells of $\rho_i = 0.5$ and a total number of particles of $N = 1000$ for the MD simulation. The mean density profile $\overline{\rho}$ at any time, taken as the mean of the density ensemble averages at every cell, is expected to remain as $\overline{\rho} \approx 0.5$ throughout the simulation due to the equilibrium state. The number of cells in the domain is $n = 40$, the cell size is $\Delta x = 50$ and the time step is computed as $\Delta t = 0.1\Delta x^2$

(selected as in [178]), the number of trajectories is 100, and the number of time steps is 2000, unless otherwise stated. The boundary conditions are periodic and the parameter β is fixed at $\beta = 1$.

The objective is to evaluate how the different numerical methods perform by focusing on four different statistical properties at equilibrium: variance, spatial correlation, time correlation and structural factor. These tests are usually employed in the literature for the validation of stochastic numerical schemes for FH [27, 178, 275]. The advantage of testing these properties at equilibrium is that their exact theoretical values are known and can be used for comparison purposes. Density fluctuations of a system with fixed volume V can be computed as $\langle \delta \rho^2 \rangle = \rho^2 \langle \delta N^2 \rangle / N^2$, where N and $\langle \delta N^2 \rangle$ denote average and variance of the number of particles in V, respectively. As shown in [187], the variance is given by:

$$\langle \delta N^2 \rangle = -T \frac{\bar{N}^2}{V^2} \left(\frac{\partial V}{\partial p} \right)_T \tag{11.55}$$

where T and p are the temperature and pressure of the system, respectively. Employing the equation of state (in reduced units) for an ideal gas, $pV = NT$, we obtain $\langle \delta N^2 \rangle = N$. In the case of infinite systems, the fluctuations of an ideal gas are spatially uncorrelated, namely $\langle \delta \rho_i(t) \delta \rho_i(t) \rangle = \langle \delta \rho^2 \rangle \delta_{ij}^K$. However, for finite systems the constraint on conservation of mass introduces correlations [27]. Expressing the space correlations of density fluctuations as $\langle \delta \rho_i(t) \delta \rho_i(t) \rangle = A \delta_{ij}^K + B$, conservation of mass dictates $\sum_i \langle \delta \rho_i(t) \delta \rho_i(t) \rangle = 0$, which corresponds to the constraint $B = -A/n$, with n being the total number of cells. Moreover, in the limit $n \mapsto \infty$ the fluctuations for an infinite system have to be recovered, thus $A = \langle \delta \rho^2 \rangle$. It follows that the spacial correlation for the closed system can be expressed as:

$$\langle \delta \rho_i(t) \delta \rho_i(t) \rangle = \langle \delta \rho^2 \rangle \left(\delta_{ij}^K - \frac{1}{n} \right). \tag{11.56}$$

The expression for the variance allows us also to obtain a quick estimation of the minimum cell size for which, due to thermal fluctuations, negative density values are likely to occur. The expected value of the density fluctuations for an ideal gas can be expressed through its standard deviation $\sqrt{\langle \delta \rho^2 \rangle} = \sqrt{\rho / \Delta V}$. Thus, with a confidence of 99.7%, the maximum values of the density fluctuations will be $\sqrt{\langle \delta \rho^2 \rangle} \mid_{\max} \sim 3 \sqrt{\frac{\rho}{\Delta V}}$. It follows that the noise fluctuations give negative density values with a probability higher than 0.3% when the following condition is verified:

$$3 \sqrt{\frac{\rho}{\Delta V}} \gtrsim \rho \quad \text{or, equivalently} \quad \Delta V \lesssim \frac{3}{\sqrt{\rho}} \tag{11.57}$$

In Subsection 11.2.1.5 we offer a discussion on the accuracy and efficiency of the temporal integrator and spatial discretization, using the results from the four tests and the computational cost. The justified choices, which are the Runge-Kutta temporal integrator (11.36) and linear approximation of the stochastic flux (11.24)-(11.25), are employed during the four tests, in the sense that the Runge-Kutta temporal integrator is employed when evaluating the different spatial discretizations, and the linear approximation of the stochastic flux is employed when evaluating the temporal integrators.

11.2.1.1 Standard deviation

For this test we aim to evaluate how the standard deviation of the density varies depending on the number of particles per cell N_c. We keep the total length and the total number of particles in the domain as constant, and we only vary the number of particles per cell by enlarging or shortening the cell size Δx. Consequently this analysis helps to elucidate how changing the finite-volume lattice size affects the numerical statistical properties . The mean density of the profile is $\overline{\rho} = N/(n\Delta x)$.

As shown above, the theoretical standard deviation of the diffusion SPDE (11.54) applied in finite systems in equilibrium satisfies

$$\sqrt{\langle \delta \rho^2 \rangle}_{\text{theory}} = \frac{\langle \rho \rangle}{\sqrt{N_c}} \sqrt{1 - \frac{1}{n}}. \tag{11.58}$$

As a remark, in spite of the fact that $\sqrt{\langle \delta \rho^2 \rangle}_{\text{theory}}$ holds for all N_c, previous studies [120, 275] have pointed out that there should be a minimum of 5-10 particles per cell to recover the microscopic statistical properties by means of FH. This occurs because with such low number of particles per cell the particle fluctuations are not accurately modelled with the multiplicative noise in (11.54).

The results of this study are displayed in Figure 11.2.1, depicting a comparison of the temporal schemes (a) and spatial discretizations (b) against the theoretical standard deviation (11.58) and the one computed from MD. It is evident from both plots that all schemes approximate correctly the standard deviation for $N_c > 5$. Below this number of particles per cell, the standard deviations deviate from the expected ones. This result chimes in with the minimum number of 5-10 particles per cell necessary to recover the statistical properties in FH.

There are no remarkable differences between the temporal integrators or spatial discretizations for the stochastic flux.
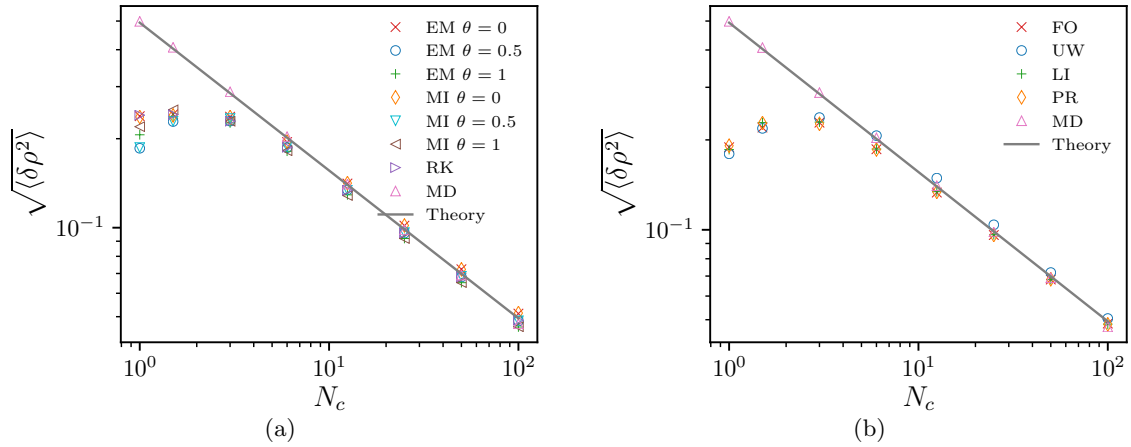

Figure 11.2.1:   Standard deviation $\sqrt{\langle \delta \rho^2 \rangle}$ as a function of the number of particles per cell N_c (or, equivalently, as a function of the cell size given that total volume and density of the system are kept constant throughout the simulations), for an ideal gas in equilibrium. (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward (11.23), UW: Upwind (11.28)-(11.29), LI: Linear (11.24)-(11.25), PR: Parabolic (11.26)-(11.27), Theory: 11.58.

11.2.1.2   Time correlations

The objective of this test is to measure the time correlation of the density in one specific cell of the domain. The normalized time correlation function is defined as

$$C_T(t) = \frac{\langle \delta \rho_i(t) \delta \rho_i(0) \rangle}{\langle \delta \rho_i(0) \delta \rho_i(0) \rangle}, \tag{11.59}$$

where $\delta\bar{\rho}_i(t) = \bar{\rho}_i(t) - \bar{\rho}$. The time correlation function expected to decay in time for any process in equilibrium, including the diffusion equation (11.54). Previous studies [27] have compared the numerical results with the Fourier transform of the time correlation (11.59), which is denoted as the spectral density and for which there are explicit expressions available. In spite of this, these exact expressions for the spectral density do not take into account the finite-size effects in the numerical simulations, leading to a lack of agreement in the results [27].

We have then decided to compare the results obtained from the numerical schemes in Section 11.1 with MD simulations only, which indeed take into account the finite-size effects. The results are displayed in Figure 11.2.2. For all schemes we evidence a clear decay in time of the time correlation. Concerning the temporal integrators, the explicit ones ($\theta = 0$) tend to be closer to the MD simulations initially, while the implicit ones ($\theta = 1$) provide a better approximation in the long-time regimes. With respect to the spatial discretizations for the stochastic flux, the upwind one deviates the most from MD, while the rest of them behave similarly.
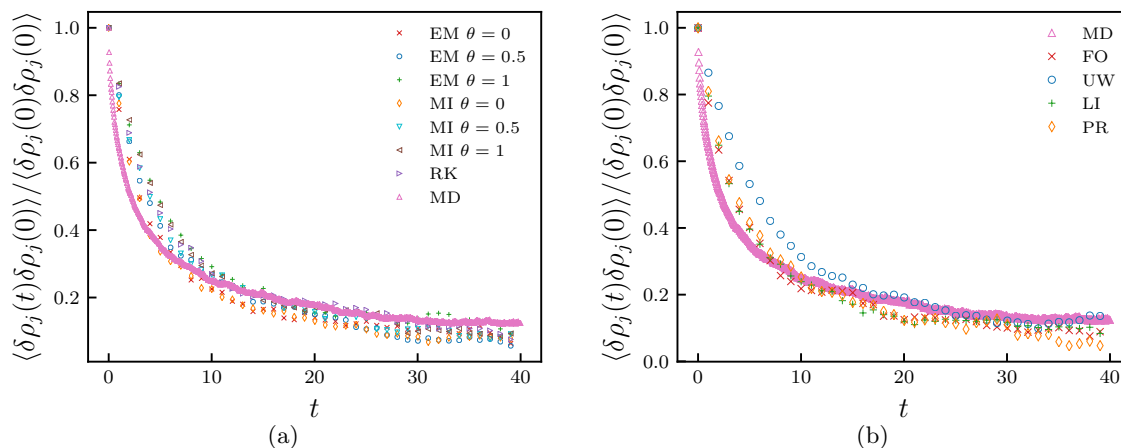

Figure 11.2.2: Temporal decay of the normalized time correlation C_T, defined as in (11.59), for an ideal-gas system in equilibrium (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward (11.23), UW: Upwind (11.28)-(11.29), LI: Linear (11.24)-(11.25), PR: Parabolic (11.26)-(11.27).

11.2.1.3   Spatial correlations

This test seeks to evaluate whether the proposed numerical schemes in Section 11.1 satisfy the exact spatial correlation for finite-size systems derived above:

$$\langle \delta\rho_i(t)\delta\rho_j(t) \rangle = \frac{\langle \rho \rangle}{\Delta x}\left(\delta_{ij} - \frac{1}{n} \right). \tag{11.60}$$

Contrary to the infinite-domain case where there are no spatial correlations between adjacent cells, for the finite-size case there is an extra term $1/n$ which decreases as the number of cells n increases.

The results of this test are depicted in Figure 11.2.3, with the normalized spatial correlation

$$C_S(t) = \frac{\langle \delta\rho_i(t)\delta\rho_i(t) \rangle}{\langle \delta\rho_i(0)\delta\rho_i(0) \rangle} \tag{11.61}$$

with $\delta\overline{\rho}_i(t) = \overline{\rho}_i(t) - \overline{\rho}$, plotted for each of the numerical schemes, the MD simulations and the exact expression (11.60). The main conclusion is that most of the temporal integrators and spatial discretizations approximate adequately the theoretical expression (11.60), as it is depicted in Figure 11.2.3. The fully explicit and implicit Euler-Maruyama and Milstein slightly deviate with respect to the theoretical spatial correlation in the cells adjacent to the central cell, while the semi-implicit schemes perform correctly.
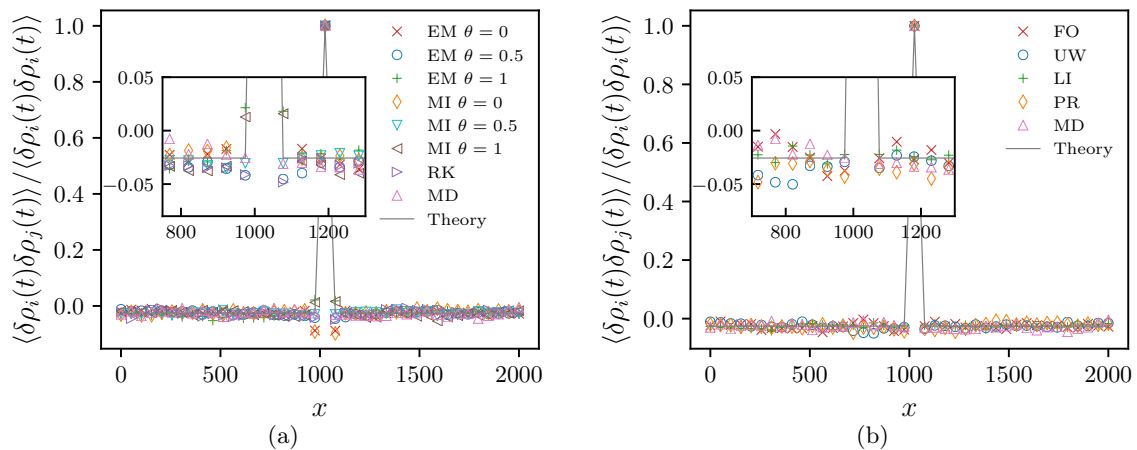

Figure 11.2.3: Normalized spatial correlation (11.61) for an ideal-gas system in equilibrium. (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward (11.23), UW: Upwind (11.28)-(11.29), LI: Linear (11.24)-(11.25), PR: Parabolic (11.26)-(11.27).

11.2.1.4 Structure factor

This test evaluates how the structure factor S at equilibrium is approximated by the temporal and spatial discretizations. Even though its general form satisfies (11.77), its theoretical expression for an ideal gas without external potential is given by (11.78), so that for the current numerical simulation with $\beta = 1$ it follows that $S/\langle\rho\rangle = 1$.

The discrete structure factor is computed from (11.71)-(11.72). First the discrete spatial Fourier transform of the density satisfies:

$$\hat{\rho}(\lambda) = \frac{1}{n}\sum_j \rho_j e^{-i\lambda x_j}. \tag{11.62}$$

Subsequently, the structure factor follows from

$$S(\lambda) = \frac{\langle\delta\hat{\rho}(\lambda)\,\delta\hat{\rho}^*(\lambda)\rangle}{n\Delta x}, \tag{11.63}$$

where $\delta\hat{\rho}(\lambda) = \hat{\rho}(\lambda) - \langle\hat{\rho}(\lambda)\rangle$ and $\hat{\rho}^*$ denotes the complex conjugate of $\hat{\rho}$.

The results of this test for the structure factor at equilibrium are depicted in Figure 11.2.4. The theoretical value of the structure factor, along with the performed MD simulations, allows us to judge whether the temporal integrators and spatial discretizations perform accurately. On the one hand, from Figure 11.2.4a it is evident how the explicit Euler-Maruyama and Milstein temporal integrators overestimate the structure factor for large λ, while their implicit versions underestimate

157

it for large λ too. The semi-implicit schemes and the Runge-Kutta behave correctly, and the small damping in the numerical structure factor for all λ is due to the choice of the hybrid deterministic flux, as it was explained from Figure 11.1.1. On the other hand, from the spatial discretizations of the stochastic flux there is a clear deviation when applying the upwind form. In addition, the forward discretization seems to slightly oscillate for lower λ. The rest of discretizations approximate the theoretical value correctly, with the small damping already mentioned.
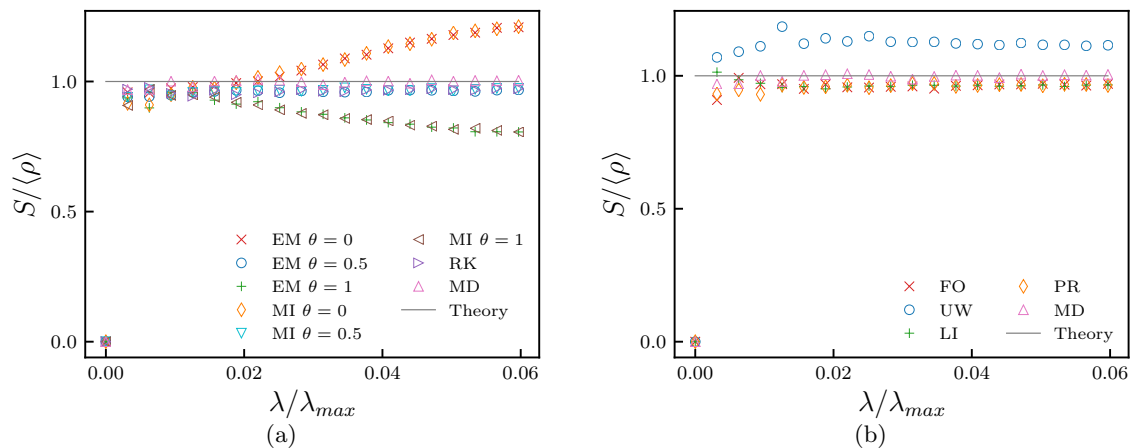

Figure 11.2.4: Structure factor (11.63) for an ideal-gas system in equilibrium. (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward (11.23), UW: Upwind (11.28)-(11.29), LI: Linear (11.24)-(11.25), PR: Parabolic (11.26)-(11.27).

11.2.1.5 Temporal integrators and spatial discretization of the stochastic flux

With respect to the temporal integrators, both the fully explicit and implicit Euler-Maruyama and Milstein present certain deviations in the time correlation (Figure 11.2.2), spatial correlation (Figure 11.2.3) and structure factor (Figure 11.2.4). Their semi-implicit versions and the Runge-Kutta behave similarly in all tests, and approximate adequately the theoretical and MD results. Their relative costs are compared by means of Figure 11.1.2. While the cost of the Runge-Kutta escalates with order $\mathcal{O}(n^3)$, the cost of the semi-implicit Euler-Maruyama and Milstein has an order of $\mathcal{O}(n^2)$. However, due to the different constant coefficient in the cost, the plot clearly shows that for $n < 100$ the Runge-Kutta cost is lower than that of semi-implicit schemes, while for $n > 100$ it is higher.

The Milstein scheme, which guarantees a higher strong order convergence, was tested because in previous works it performed well in conjunction with adaptive time-step algorithms based on Brownian trees [203]. However, from the simulation results, we observed that the higher computational cost of this numerical method did not lead to a increased accuracy compared to the implicit Euler-Maruyama and to the weak Runge-Kutta schemes. Because of these reasons and together with the fact that in the simulations of this work $n < 100$, we select the Runge-Kutta temporal integrator.

Concerning the spatial discretization of the stochastic flux, the upwind choice does not approximate well the time correlation and structure factor, while the forward approximation presents some deviation in the structure factor for small λ. Hence the best choices are the linear and parabolic

approximations, which behave similarly in all test cases. We select the linear approximation due to its lower cost since it only depends on the density and white noise cell averages of two cells and not four.

11.2.2 Ideal-gas system out of equilibrium

For this example we consider a free energy which includes the effects of a double-well external potential, so that

$$\mathcal{F}[\rho] = \int \rho/\beta \left(\ln(\rho) - 1\right) dx + \int V(x)\rho \, dx, \tag{11.64}$$

and the shape of the external potential satisfies

$$V(x) = 5 \left[\left(\frac{x}{n\Delta x/2} \right)^4 - \left(\frac{x}{n\Delta x/2} \right)^2 \right]. \tag{11.65}$$

Numerical simulations for deterministic gradient-flow equations with the free energies of the form (11.64)-(11.65) have already been provided in [55]. Here the objective is to evaluate how the numerical scheme in Section 11.1 for the FDDFT (11.6) with the free energy (11.64)-(11.65) performs by comparing with MD simulations. We also include a comparison with the corresponding deterministic DDFT, which is obtained in the mean-field limit (the most-likely path of FDDFT as noted in the Introduction).

The simulation is performed in a mesh where the number of cells is $n = 40$, each of them with width $\Delta x = 5$. The time steps are $\Delta t = 1$ and the number of time steps is $n_t = 2000$. The ensemble averages are computed from a number of trajectories of $n_{traj} = 1000$. We select $\beta = 1$. The MD simulation is performed by simulating $N = 200$ particles, while the deterministic DDFT simulation applies the numerical scheme in [55] for gradient-flow equations.

The results are depicted in Figure 11.2.5. Figure 11.2.5a displays the ensemble average of the density profile at different times. The three simulations provide similar results and we can conclude that the three approaches are comparable when evaluating the ensemble average profile. Concerning the standard deviation results in Figure 11.2.5b, we find that FDDFT matches with MD and the theoretical results in (11.58), while DDFT, being deterministic, presents zero standard deviation. As already mentioned, the FDDFT values of the standard deviation are slightly lower than the MD and theoretical ones due to the choice of the deterministic flux in a similar fashion to Figures 11.1.1 and 11.2.1. Figure 11.2.5c shows the spatial correlation computed as in (11.60)-(11.61), with the MD and FDDFT results approximating correctly the finite-size theoretical expression in (11.60). DDFT does not have any spatial correlation due to the lack of fluctuations. Finally, in 11.2.5d the temporal evolution of the free-energy functional depending on the ensemble average density is plotted. For the DDFT case one can appreciate that there is decay at all times, while for MD and FDDFT there are short increases of the free energy triggered by the fluctuations, in spite of the fact that during the evolution there is a general decay in the free energy.

11.2.3 Homogeneous nucleation in Lennard-Jones systems

The importance of fluctuations during phase transitions is crucial when considering the homogeneous vapour-liquid transition of a Lennard-Jones (LJ) fluid. Within the framework of DFT, the fluid density profiles of a 1D open system exchanging particles with a reservoir at constant temperature and chemical potential μ, can be obtained from an unconstrained numerical minimization of the grand free-energy functional

$$\Omega[\rho(x)] = \mathcal{F}[\rho(x)] + \int \left(V(x) - \mu\right)\rho(x) \, dx.$$
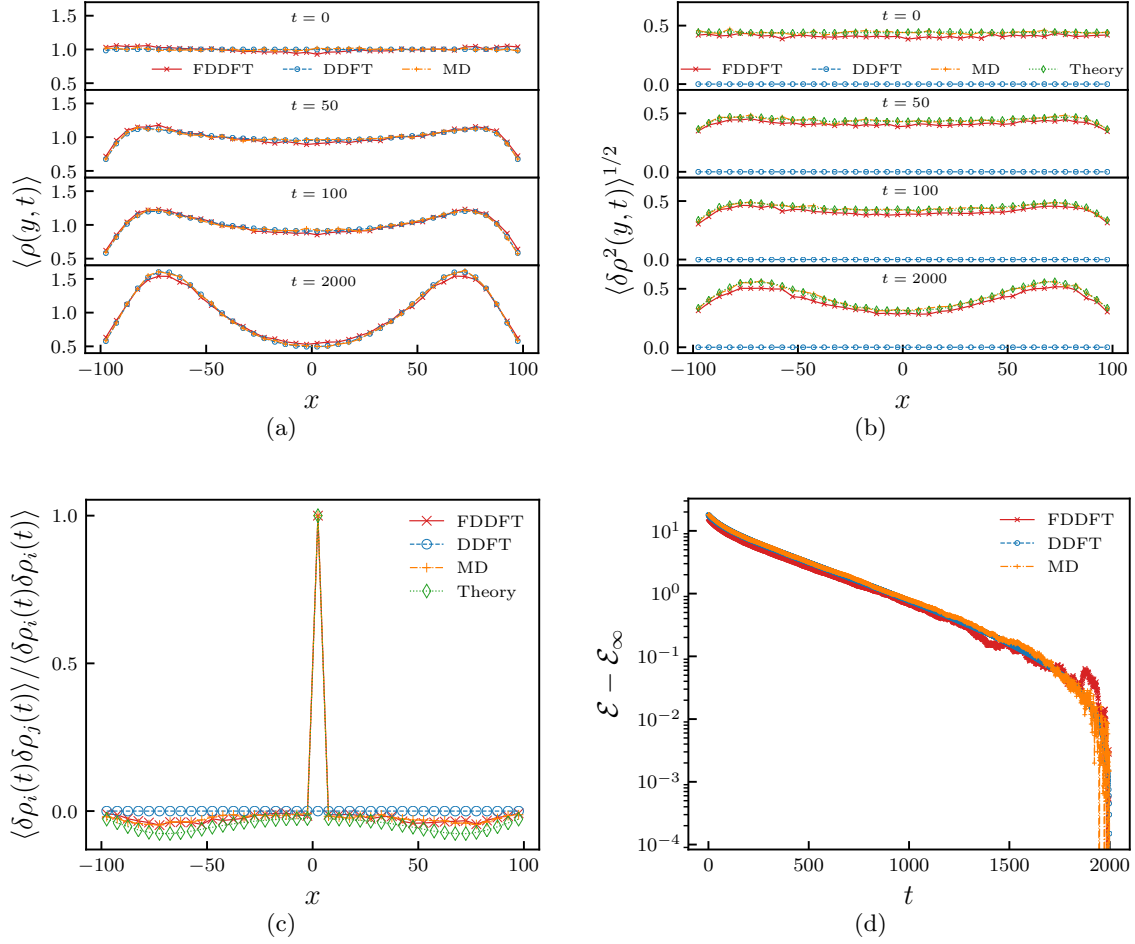
Figure 11.2.5:   Time evolution of mean density (a) and density standard deviation (b) fields computed with FDDFT, DDFT and MD simulations. A comparison in terms of steady state spatial correlations is reported in (c). In (d), we report the decrease in time of the energy functional of the mean density.

In general, $\mathcal{F}[\rho(x)]$ is not analytically obtainable from first principles, except in few cases, i.e. ideal gases and hard-sphere fluids. In the remaining cases, $\mathcal{F}[\rho(x)]$ is either numerically obtained from atomistic simulations or is approximated by means of perturbation expansions around a known free energy [207]. Similarly to previous works on DFT [291, 295], we approximate $\mathcal{F}[\rho(x)]$ of an LJ fluid according to the first-order Barker-Henderson perturbation theory expansion around the hard-sphere fluid free energy [22], namely as

$$\mathcal{F}[\rho(x)] = \int \left\{ f_{\text{ID}}[\rho(x)] + \rho(x) f_{\text{HS}}(\rho(x)) \right\} dx + \frac{1}{2} \int \int \rho(x)\rho(x') W(x, x') \; dx \; dx', \qquad (11.66)$$

where f_{ID}, f_{HS} and $W(x, x')$ denote ideal-gas, hard-sphere repulsive interactions and LJ attractive contributions, respectively. The free energy of an ideal gas is given by

$$f_{\text{ID}}[\rho(x)] = k_{\text{B}}T\rho \left(\ln(\lambda^3 \rho) - 1 \right),$$
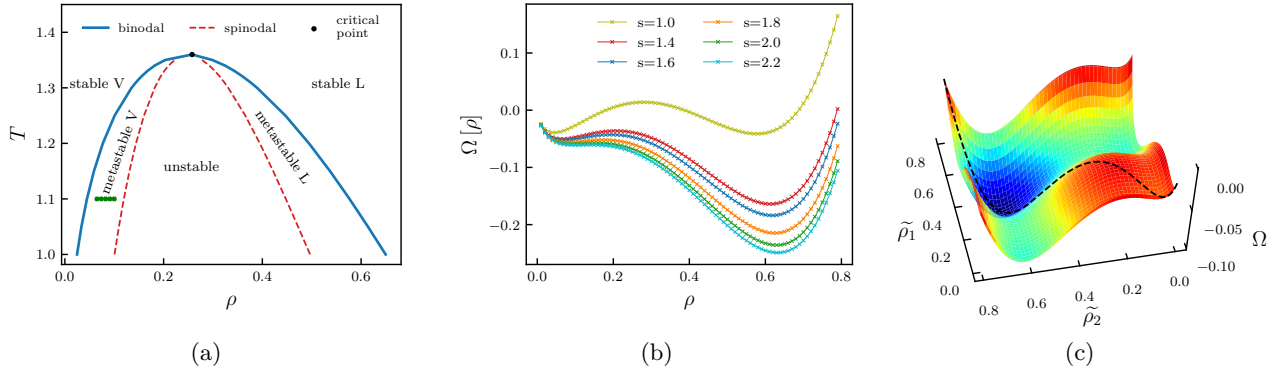
160

(a)

(b)

(c)

Figure 11.2.6:   (a) Bulk phase diagram for the discretized LJ system.   (b) Grand free-energy landscape as function of the system density for some supersaturation ratios adopted in this study. (c) Example of free-energy landscape for systems with a non-uniform density field, with two varying densities $\widetilde{\rho}_1$ and $\widetilde{\rho}_2$. The dotted black line denotes the curve corresponding to bulk uniform systems.

where λ is the thermal de Broglie wavelength. The hard-sphere free-energy density f_{HS} is obtained from the Carnahan-Starling equation of state for the hard sphere fluid, which reads [53]

$$f_{HS}(\rho(x)) = k_{\mathrm{B}}T \left(\frac{4\eta - 3\eta^2}{(1 - \eta)^2} \right), \quad \text{with} \quad \eta = \frac{\pi}{6}\rho\sigma^3$$

where σ is the hard-sphere diameter set to unity in this work. Finally, the LJ (attractive) contributions are taken into account by the following expression:

$$W(x, x') = \begin{cases} -1.2 \; \pi \; \epsilon & \text{if} \quad \mid x - x' \mid \leq 1, \\ \pi \; \epsilon \; \left(0.8 \mid x - x' \mid^{-10} -2 \mid x - x' \mid^{-4} \right) & \text{otherwise,} \end{cases}$$

derived by integrating along y and z the 12-6 LJ potential [295].

In order to analyse the vapour-to-liquid (first-order) phase transitions, we first compute the coexisting density profiles. The coexisting values of vapour and liquid density (binodal line) are denoted as ρ_v and ρ_l respectively, and are obtained by solving the following system of equations:

$$\begin{cases} \left. \frac{\partial \Omega}{\partial \rho} \right|_{\rho_v} = \left. \frac{\partial \Omega}{\partial \rho} \right|_{\rho_l} = 0, \\ \Omega\left[\rho_v\right] - \Omega\left[\rho_l\right] = 0. \end{cases} \tag{11.67}$$

The metastable regions are delimited by the binodal and spinodal lines. The spinodal lines correspond to the inflection points of the grand free energy, hence are evaluated by solving:

$$\left. \frac{\partial^2 \Omega}{\partial \rho^2} \right|_{\rho_v} = \left. \frac{\partial^2 \Omega}{\partial \rho^2} \right|_{\rho_l} = 0. \tag{11.68}$$

Finally, the bulk critical point is given by the intersection between binodal and spinodal lines, and it is thus computed as

$$\left. \frac{\partial \Omega}{\partial \rho} \right|_{\rho_c, T_c} = \left. \frac{\partial^2 \Omega}{\partial \rho^2} \right|_{\rho_c, T_c} = 0. \tag{11.69}$$

In Figure 11.2.6a we report the bulk phase diagram obtained from the discretized grand free energy of LJ fluid. Solid curves depict the binodal, i.e. the locus of liquid-gas coexistence, while dashed
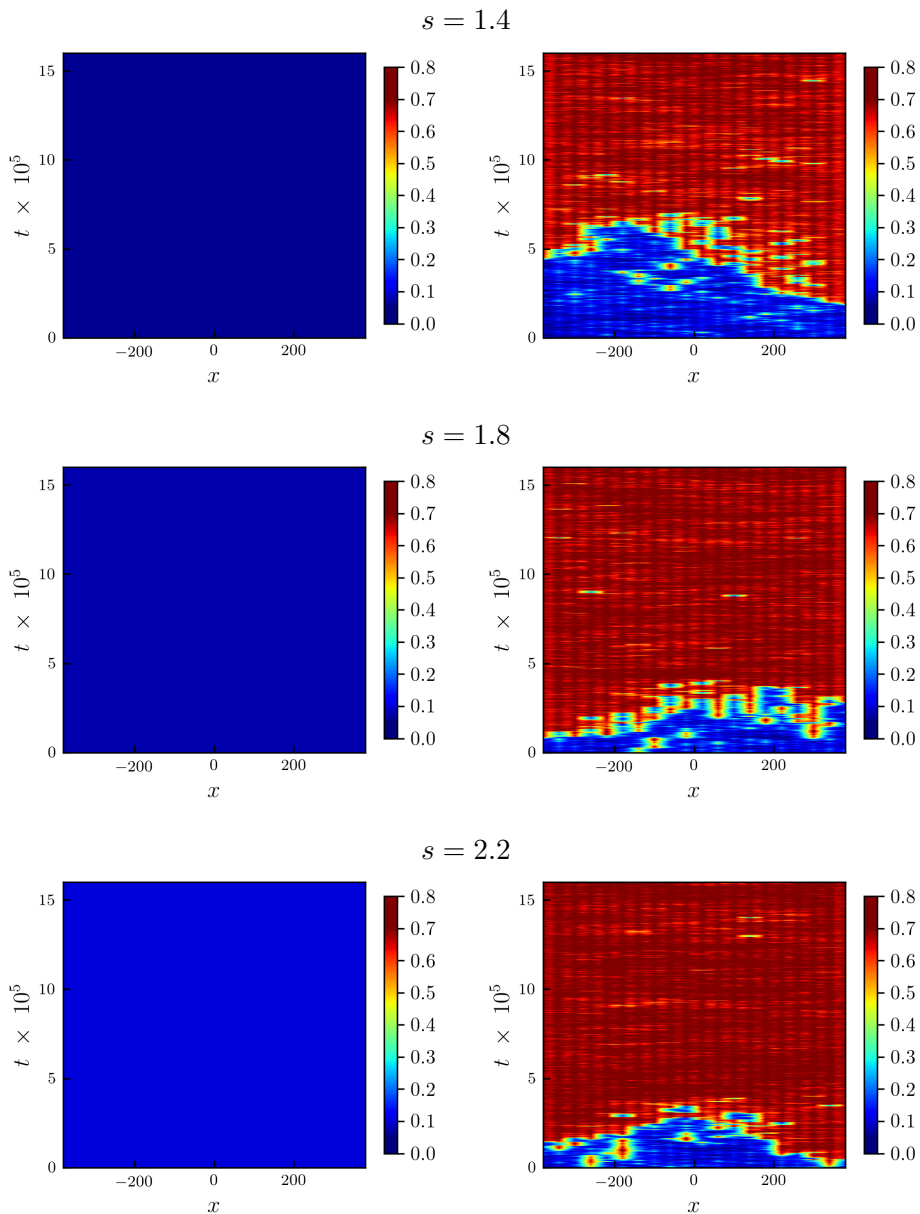
Figure 11.2.7:   Homogeneous nucleation of a vapour LJ system in metastable conditions with supersaturation ratio s. Left column: the mean-field evolution. Right column: a single realization of the stochastic dynamics.

curves depict the spinodal, i.e. the boundary between the metastable and the unstable regions. The black circle designates the bulk critical point at $\rho_c \sim 0.3$ and $T_c \sim 1.35$.

If we denote with ρ_v the vapour coexistence density at a given temperature, the supersaturation ratio is defined as $s = \rho/\rho_v$. We will study the nucleation of vapour systems with identical temperatures, but different initial supersaturation ratios. Figure 11.2.6b depicts the free-energy landscape as a function of the bulk density for such systems. At coexistence $s = 1$, two stable basins are present, which means that the system has equal probability of being in one of the two. Increasing the supersaturation ratio enhances the stability of the liquid phase, thus leaving the vapour density in a metastable condition. Also, the energy barrier that the system has to overcome to pass from the vapour to the liquid phase decreases with s, until it becomes null at

(a)   (b)   (c)

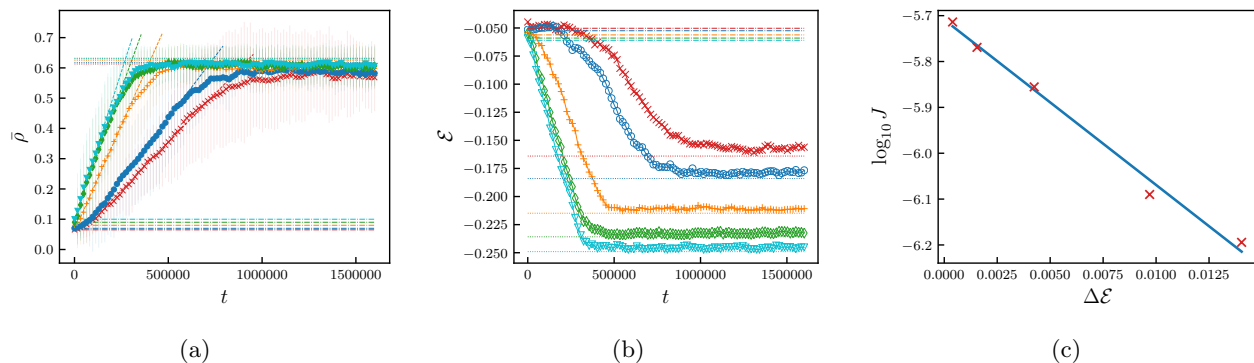Figure 11.2.8: (a) reports the evolution in time of the average system density for the supersaturation ratios adopted in this study. (b) shows the time-evolution of the system free energy for the supersaturation ratios adopted in this study. (c) the nucleation growth rate is plotted against the free-energy barrier.

a supersaturation corresponding to the spinodal line. In such condition only one minimum of the grand free energy exists.

During a phase transition, the system will move from the initial uniform-density state to the final uniform-density state, while during the transition between the two states the density is non-uniform. This means that the bulk grand-free energy in Figure 11.2.6b, being only valid for uniform densities, describes the system in the initial and final stages only, but it does not provide information on the transition path. The grand free energy for non-uniform systems is in general a function of each cell density, i.e. it is an n-dimensional manifold. To give a representative example of this, in Figure 11.2.6c we report our LJ grand-free energy for a non uniform system, constrained to have only two varying densities $\widetilde{\rho}_1 = \left\{\rho_1 = \cdots = \rho_{n/2}\right\}$ and $\widetilde{\rho}_2 = \left\{\rho_{n/2+1} = \cdots = \rho_n\right\}$. The bulk free energy is then recovered for $\widetilde{\rho}_1 = \widetilde{\rho}_2$ (dotted black line).

Single trajectories of the vapour-to-liquid phase transition at different supersaturation ratios, are reported in Figure 11.2.7. For comparison purposes, we perform simulations of the FDDFT and its mean-field (deterministic) counterpart. In order for the transition to occur, the system grand free energy has to overcome an energy barrier. Such passage requires a local injection of energy, thus it is triggered by fluctuations. As a consequence, the mean-field approach fails to describe the transition. Moreover, as expected by looking at the energy barrier in Figure 11.2.6, the transition is favoured by higher supersaturation ratios.

In addition to the presence of fluctuations, the phase transition is allowed due to the open boundary conditions imposed on the system. These boundary conditions are described in Subsection 11.1.5, and model the exchange of particles with a reservoir at constant temperature T_{res} and chemical potential μ_{res}. The mass of the system can then increase (or decrease), thus permitting the transition from the lower-density minima in Figure 11.2.6b to the higher-density ones. However, it is important to remark that these boundary conditions do not simply add (or remove) mass to the system. The imposed chemical potential at the boundary, μ_{res}, can be iteratively solved to obtain the value of the density that satisfies it. We choose μ_{res} so that this iterative algorithm may converge to one of the two minima in Figure 11.2.6b, depending on the initial conditions for the iteration. For the two simulations in Figure 11.2.7 we always select to converge to the lower-density minimum in Figure 11.2.6b. This is why, with identical boundary conditions, the mean-field deterministic simulation in Figure 11.2.7 remains at the lower-density minima in Figure 11.2.6b and conserves the mass. On the contrary, the FDDFT simulation in Figure 11.2.7 is able to increase

the mass thanks to the constant density at the boundary, which allows a continuous exchange of particles.

The trend observed in Figure 11.2.7 is quantitatively analysed in Figure 11.2.8, where we report an ensemble average of 10 nucleation trajectories for each supersaturation ratio. Figure 11.2.8a shows the average density increase as a function of time. The initial and final average system densities are consistent with the vapour and liquid bulk densities predicted by the grand-free energy analysis.

The free energy evaluated at each time as function of the average density is reported in Figure 11.2.8b. The initial free-energy value, corresponding to the vapour metastable basin, evolves in time in order to the reach the more stable liquid basin, as predicted by Figure 11.2.6b. It is interesting that the passage between the two basins implies a slight increase in the free energy due to the energy barrier overcome by the density field fluctuations.

Evidently, the average density kinetics is characterized by three main stages: 1) an initial latency period, 2) a growth period and, 3) an asymptotic relaxation towards a plateau, corresponding to the liquid-phase density. This dynamics is consistent with the multi-stage nucleation pathway experimentally observed and theoretically studied in the phase-transition research community [124]. The growth period exhibits a linear-like trend, with slopes representing the nucleation growth rate J. As reported in the plot in Figure 11.2.8c, an Arrhenius-like relation (as is the case with thermally activated processes) is observed between J and the grand-free energy barrier $\Delta\mathcal{F}$, i.e.

$$J \sim K \exp{-\frac{\Delta\mathcal{F}}{T}}, \tag{11.70}$$

where J is the growth rate K in the limit of a zero-energy barrier. It is worth noticing that the pre-exponential factor K in reality is not a constant, but can be often approximated as constant over limited supersaturation regions [203, 204].

Finally, we remark that the finite-volume scheme is able to accurately simulate processes where the number of particles per cell is greater than 5, as showed in Subsubsection 11.2.1.1. For any process that involves smaller scales one has to rely on MD simulations. This could be relevant for processes, such as nucleation, which may require capturing system features down almost to particle scales at initiation.

Appendix

11.A   Structure factor

The structure factor is a quantity of interest in many fields, including FH [120] as noted earlier and capillary wave theory [227, 228]. As shown in previous works [120, 178], the structure factor represents an important measure of the stochastic properties of the system and it can be experimentally obtained. Thus, the structure factor is a valuable quantity not only to study the stability of the numerical integrator, but also to compare different schemes, as it will be shown in Section 11.1. Here we derive an expression of the structure factor from the linearized FDDFT. If we consider a periodic domain of volume V, the spatial Fourier transform of the density is given by

$$\hat{\rho}_{\boldsymbol{\lambda}} = \frac{1}{V} \int_V \rho(\boldsymbol{x}, t) e^{-i\boldsymbol{\lambda}\cdot\boldsymbol{x}} d\boldsymbol{x}. \tag{11.71}$$

The structure factor is defined as the variance of the Fourier transform of the density fluctuations,

$$S(\boldsymbol{\lambda}) = V \langle \delta\hat{\rho}_{\boldsymbol{\lambda}} \delta\hat{\rho}_{\boldsymbol{\lambda}}^* \rangle, \tag{11.72}$$

where $\delta\hat{\rho}_{\boldsymbol{\lambda}} = \hat{\rho}_{\boldsymbol{\lambda}} - \langle \hat{\rho}_{\boldsymbol{\lambda}} \rangle$, and $\hat{\rho}_{\boldsymbol{\lambda}}^*$ denotes the complex conjugate of $\hat{\rho}_{\boldsymbol{\lambda}}$.

For uniform systems, (11.1) can be formally linearized around its most-likely solution $\langle\rho\rangle$ by means of the Central Limit Theorem, giving

$$\partial_t \rho = \nabla \cdot \left(\rho \nabla \frac{\delta\mathcal{F}[\rho]}{\delta\rho} \right) + \sqrt{\langle\rho\rangle/\beta} \nabla \cdot \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t). \tag{11.73}$$

Taking the Fourier transform of the difference between (11.73) and (11.3), one obtains

$$\partial_t \ \delta\hat{\rho}(\boldsymbol{\lambda}) = i\boldsymbol{\lambda} \cdot \left\{ \mathcal{T}\left(\rho \nabla \frac{\delta\mathcal{F}[\rho]}{\delta\rho} \right) - \mathcal{T}\left(\langle\rho\rangle \nabla \frac{\delta\mathcal{F}[\langle\rho\rangle]}{\delta\langle\rho\rangle} \right) \right\} + i\boldsymbol{\lambda} \cdot \sqrt{\langle\rho\rangle/\beta} \hat{\boldsymbol{\mathcal{W}}}(\boldsymbol{\lambda}). \tag{11.74}$$

where $\mathcal{T}$ denotes the Fourier transform. If the free-energy functional terms in the Fourier space can be expanded at first order around their mean value as

$$\mathcal{T}\left(\rho \nabla \frac{\delta\mathcal{F}[\rho]}{\delta\rho} \right) \sim \mathcal{T}\left(\langle\rho\rangle \nabla \frac{\delta\mathcal{F}[\langle\rho\rangle]}{\delta\langle\rho\rangle} \right) + \frac{\partial \mathcal{T}\left[\rho\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho} \right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}} \delta\hat{\rho}_{\boldsymbol{\lambda}} + \mathcal{O}(\delta\hat{\rho}_{\boldsymbol{\lambda}}), \tag{11.75}$$

then (11.74) yields

$$\partial_t \ \delta\hat{\rho}_{\boldsymbol{\lambda}} = i\boldsymbol{\lambda} \cdot \frac{\partial \mathcal{T}\left[\rho\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho} \right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}} \delta\hat{\rho}_{\boldsymbol{\lambda}} + i\boldsymbol{\lambda} \cdot \sqrt{\langle\rho\rangle/\beta} \hat{\boldsymbol{\mathcal{W}}}(\boldsymbol{\lambda}). \tag{11.76}$$

Since the above equation has the form of an Ornstein-Uhlenbeck process, the structure factor can be computed as its variance:

$$S(\boldsymbol{\lambda}) = \frac{2\left(i\boldsymbol{\lambda}\sqrt{\langle\rho\rangle/\beta}\right)^2}{2i\boldsymbol{\lambda}\frac{\partial\mathcal{T}\left[\rho\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho}\right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial\mathcal{T}\left[\rho\nabla\frac{\delta\mathcal{F}[\rho]}{\delta\rho}\right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} \tag{11.77}$$

For example, in the case of an ideal gas without external potential, $\frac{\delta\mathcal{F}[\rho]}{\delta\rho} = \log\rho$, the structure factor is given by the well-known expression [178]:

$$S(\boldsymbol{\lambda}) = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial\mathcal{T}[\rho\nabla\log\rho]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial\mathcal{T}[\nabla\rho]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial[i\boldsymbol{\lambda}\hat{\rho}]}{\partial\hat{\rho}(\boldsymbol{\lambda})}} = \langle\rho\rangle/\beta. \tag{11.78}$$

11.B   Details of molecular-dynamic simulations

MD simulations are performed using the Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [237]. Particle positions and velocities are integrated in time using the velocity-Verlet algorithm, with a time-step of $dt = 0.001\tau$. The system is simulated at constant temperature and volume, so that particle coordinates are consistent with the canonical ensemble (NVT). Specifically, the temperature $T = 1$ is kept constant during the simulations using a Langevin thermostat. All the physical quantities are expressed in reduced units, i.e. they are nondimensionalized with the fundamental quantities σ, ϵ and m, representing distance, energy and mass, respectively. Further, without loss of generality, σ, ϵ, m and the Boltzmann constant k_B are set equal to unity.

As discussed extensively in [242], a macroscopic field $X(\mathbf{r}, t)$ can be extracted from particle coordinates as $X(\mathbf{r}, t) = \sum_i \chi_i \phi(\boldsymbol{x}_i(t) - \boldsymbol{x})$, where χ_i is the quantity of interest for particle i at position $\mathbf{r}_i$ at time t, and ϕ is a kernel function (commonly a piecewise constant, Gaussian, or polynomial function). In this work, we adopt a piecewise constant function defined as:

$$\phi(y) = \begin{cases} \frac{1}{\Delta x} & \text{for } \|y\| < \Delta x/2, \\ 0 & \text{otherwise}, \end{cases} \tag{11.79}$$

where Δx is the width of each bin. In each comparison, Δx for MD simulations is taken to be the same with that for the discretized FDDFT. Using the above, the instantaneous macroscopic density profile for a single trajectory is computed as:

$$\rho(x, t) = \sum_i m_i \phi(x_i(t) - x), \tag{11.80}$$

where m_i is the mass of the particle i.

11.B.1   Equilibrium simulations

MD simulations of ideal gas fluids in equilibrium are performed using a fixed number of particles $(1,000)$ in a 1D domain of length $2,000$ (in reduced units) with periodic boundary conditions. The system is equilibrated and then a run of 2×10^7 time steps is performed, during which fluid particle positions and velocities are stored every 10^4 time steps for analysis. The process is repeated 10^3 times to generate independent trajectories.
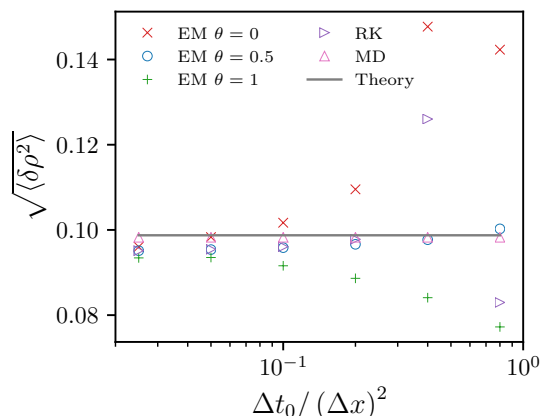
Figure 11.C.1:   Standard deviation $\sqrt{\langle\delta\rho^2\rangle}$ as a function of the normalized initial time step $\Delta t_0/(\Delta x)^2$ for an ideal gas in equilibrium. Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$), Theory: (11.58).

11.B.2   Non-equilibrium simulations

MD simulations of ideal gas fluids in non-equilibrium conditions are performed using a fixed number of particles (200) in a 1D domain of length 200 (in reduced units) with periodic boundary conditions, under an external potential:

$$V(x) = 5\left[\left(\frac{x}{200}\right)^4 - \left(\frac{x}{200}\right)^2\right].\tag{11.81}$$

A run of 2×10^6 time steps is performed, during which fluid particle positions and velocities are stored every 10^3 time steps for analysis. The process is repeated 10^3 times with different (random) initial conditions to generate independent trajectories and gather statistics.

11.C   Time-integrator stability analysis

Both stability and accuracy of the different time-integrators are relevant, given that large time-steps are required in many applications (for instance, for transitions occurring over long time-scales). In the main text, we focused on the accuracy of the schemes comparing finite-volume schemes, MD and theoretical results. Here we analyse the stability of the different time integrators with respect to the time-step size.

   Specifically, in Figure 11.C.1 we report a comparison of the fluctuations' standard deviation obtained from selected time integrators and the MD-theoretical results for varying time step sizes Δt_0. Because of the adaptive time step adopted in the simulations, the actual time step may not be constant throughout the simulations, and in fact it may be lower than Δt_0. The system considered here is the same ideal-gas system (with average density $\bar\rho = 0.5$) used for the analyses in the main text. The cell size adopted is $\Delta x = 50$, corresponding to a number of particles per cell, $N_c = 25$. We do not report the results for the Milstein schemes, as in several tests we did not observe any relevant difference between the Milstein scheme and the Euler-Maruyama one as far as the mean, variance and correlations are concerned. Figure 11.C.1 shows that the semi-implicit scheme outperforms both explicit and implicit schemes at high $\Delta t_0/(\Delta x)^2$, becoming the time-integrator

of choice for computations requiring large time steps. Moreover, the explicit Runge-Kutta scheme shows enhanced stability compared to both implicit and explicit Euler-Maruyama.

Finite-volume schemes for the system of fluctuating hydrodynamics

The fluctuating-hydrodynamic system studied here corresponds to the fluctuating DDFT framework derived in [123], for the general case of arbitrarily shaped and thermalized particles. It consists of two stochastic PDEs for the density ρ and velocity $\boldsymbol{u}$ fields:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^d, \quad t > 0, \\ \partial_t (\rho \boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) = -\rho \nabla \dfrac{\delta \mathcal{F}[\rho]}{\delta \rho} - \gamma \rho \boldsymbol{u} - \sqrt{k_B T \gamma \rho} \, \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t). \end{cases} \tag{12.1}$$

The free energy functional $\mathcal{F}[\rho]$ contains the pressure $P(\rho)$ and generic potential terms $H(\boldsymbol{x}, \rho)$, and can be decomposed as

$$-\rho \nabla \frac{\delta \mathcal{F}}{\delta \rho} = -\nabla P(\rho) - \rho \nabla H(\boldsymbol{x}, \rho).$$

The potential terms $H(\boldsymbol{x}, \rho)$ involve an external field $V(x)$ and an interaction potential $W(x)$ convoluted with the density ρ, satisfying

$$H(\boldsymbol{x}, \rho) = V(\boldsymbol{x}) + W(\boldsymbol{x}) \star \rho.$$

Accordingly, the free-energy functional has the form

$$\mathcal{F}[\rho] = \int_{\mathbb{R}^d} \Pi(\rho) d\boldsymbol{x} + \int_{\mathbb{R}^d} V(\boldsymbol{x}) \rho(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(\boldsymbol{x} - \boldsymbol{y}) \rho(\boldsymbol{x}) \rho(\boldsymbol{y}) d\boldsymbol{x} d\boldsymbol{y}, \tag{12.2}$$

where $\rho \Pi''(\rho) = P'(\rho)$.

With respect to the additive noise term, k_B is the Boltzmann constant, T the temperature, γ the linear damping and $\boldsymbol{\mathcal{W}}$ is a vector of Gaussian stochastic processes delta-correlated in space and time, i.e.

$$\begin{aligned} \langle \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t) \rangle &= 0, \\ \langle \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t), \boldsymbol{\mathcal{W}}(\boldsymbol{x}', t') \rangle &= 2 \delta(t - t') \delta(\boldsymbol{x} - \boldsymbol{x}'). \end{aligned} \tag{12.3}$$

The one-dimensional (1D) version of the fluctuating hydrodynamic system in (12.1) can be written as

$$\partial_t U + \partial_x F(U) = S_H(U, H) + S_D(x, U) + S_{\mathcal{W}}(x, U), \tag{12.4}$$

with

$$U = \begin{pmatrix} \rho \\ \rho u \end{pmatrix}, \quad F(U) = \begin{pmatrix} \rho u \\ \rho u^2 + P(\rho) \end{pmatrix}$$

and

$$S_H(U, H) = \begin{pmatrix} 0 \\ -\rho\partial_x H \end{pmatrix}, \quad S_D(x, U) = \begin{pmatrix} 0 \\ -\gamma\rho u \end{pmatrix}, \quad S_{\mathcal{W}}(x, U) = \begin{pmatrix} 0 \\ -\sqrt{k_B T \gamma \rho}\, \mathcal{W}(x, t) \end{pmatrix},$$

where U are the unknown variables, $F(U)$ the fluxes, and $S_H(U, H)$, $S_D(x, U)$ and $S_{\mathcal{W}}(x, U)$ are the sources related to forces with potential H, damping terms and noise respectively.

The well-balanced first-order finite-volume scheme of the deterministic 1D system (12.4) is detailed in Chapter 5, and its high-order versions in Chapter 6. The objective here is to design a numerical scheme for the system of fluctuating hydrodynamics (12.4), by means of combining the first-order scheme in Chapter 5 with the noise discretization for the stochastic gradient flow equations in Chapter 11.

The finite-volume formulation of (12.4) is obtained by dividing the domain into grid cells $C_i = [x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}]$, centred at x_i and assumed to have the same length $\Delta x = x_{i+1/2} - x_{i-1/2}$, and then approximating in each of them the cell average of ρ and u defined as

$$\rho_i(t) = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} \rho(x, t) dx, \quad u_i(t) = \frac{1}{\Delta x} \int_{x_{i-1/2}}^{x_{i+1/2}} u(x, t) dx. \tag{12.5}$$

These cell averages are approximated with the mid-point rule, which satisfies second-order accuracy,

$$\rho_i(t) = \rho(x_i, t) + \mathcal{O}(\Delta x^2), \quad u_i(t) = u(x_i, t) + \mathcal{O}(\Delta x^2). \tag{12.6}$$

Subsequently, one has to integrate (12.4) spatially over each cell and apply the Gauss divergence theorem, leading to the semi-discrete equation for the temporal evolution of the cell averages,

$$\frac{dU_i}{dt} = -\frac{F_{i+\frac{1}{2}} - F_{i-\frac{1}{2}}}{\Delta x_i} + S_{H,i} + S_{D,i} + S_{\mathcal{W},i}, \tag{12.7}$$

where the cell average of U in the cell $\left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$ is denoted as

$$U_i = \begin{pmatrix} \rho_i \\ \rho_i u_i \end{pmatrix},$$

$F_{i+\frac{1}{2}}$ is an approximation of the flux $F(U)$ at the point $x_{i+\frac{1}{2}}$, and $S_{H,i}$, $S_{D,i}$ and $S_{\mathcal{W},i}$ are cell-average approximations of the source terms $S_H(x, U)$, $S_D(x, U)$ and $S_{\mathcal{W}}(x, U)$ in the cell $\left[x_{i-\frac{1}{2}}, x_{i+\frac{1}{2}}\right]$.

12.1 First-order well-balanced scheme for fluxes and potential sources

The first-order well-balanced scheme proposed in Chapter 5 allows to preserve the discrete version of the steady states for the deterministic system (12.4),

$$\left(\frac{\delta\mathcal{F}}{\delta\rho}\right)_i = \Pi'(\rho_i) + H_i = C_\Gamma \text{ in each } \Lambda_\Gamma, \Gamma \in \mathbb{N}, \tag{12.8}$$

where Λ_Γ, $\Gamma \in \mathbb{N}$, denotes the possible infinite sequence indexed by Γ of subsets Λ_Γ of subsequent indices $i \in \mathbb{Z}$ where $\rho_i > 0$ and $u_i = 0$, and C_Γ the corresponding constant in that connected component of the discrete support.

The flux approximations $F_{i+\frac{1}{2}}$ are evaluated by means of a numerical flux depending on the interface values $U^{\pm}_{i+\frac{1}{2}}$,

$$F_{i+\frac{1}{2}} = \mathcal{F}\left(U^{-}_{i+\frac{1}{2}}, U^{+}_{i+\frac{1}{2}}\right), \quad \text{where } U^{\pm}_{i+\frac{1}{2}} = \begin{pmatrix} \rho^{\pm}_{i+\frac{1}{2}} \\ \rho^{\pm}_{i+\frac{1}{2}} u^{\pm}_{i+\frac{1}{2}} \end{pmatrix}. \tag{12.9}$$

The numerical fluxes employed in Chapter 5 are the Lax-Friedrich flux and the kinetic flux in [236]. For more details about these numerical fluxes we refer the reader to Appendix 5.A.

The interface values $U^{\pm}_{i+\frac{1}{2}}$ are reconstructed from U_i and U_{i+1} by taking into account the steady state relation in (12.8),

$$\Pi'\left(\rho^{-}_{i+\frac{1}{2}}\right) + H_{i+\frac{1}{2}} = \Pi'\left(\rho_i\right) + H_i,$$

$$\Pi'\left(\rho^{+}_{i+\frac{1}{2}}\right) + H_{i+\frac{1}{2}} = \Pi'\left(\rho_{i+1}\right) + H_{i+1},$$

with the discretization of the potential $H(x, \rho)$ satisfying

$$H_i = V_i + \sum_j \Delta x_j W_{ij} \rho_j,$$

where $V_i = V(x_i)$ and $W_{ij} = W(x_i - x_j)$ in case the potential is smooth or choosing W_{ij} as an average value of W on the interval centred at $x_i - x_j$ of length Δx_j in case of general locally integrable potentials W. The term $H_{i+\frac{1}{2}}$ is evaluated to preserve consistency and stability, so that

$$H_{i+\frac{1}{2}} = \max\left(H_i, H_{i+1}\right). \tag{12.10}$$

Then, by denoting as $\xi(s)$ the inverse function of $\Pi'(s)$ for $s > 0$, it follows that the interface values $U^{\pm}_{i+\frac{1}{2}}$ are computed as

$$\begin{aligned} \rho^{-}_{i+\frac{1}{2}} &= \xi\left(\Pi'\left(\rho_i\right) + H_i - H_{i+\frac{1}{2}}\right)_+, \quad u^{-}_{i+\frac{1}{2}} = u_i, \\ \rho^{+}_{i+\frac{1}{2}} &= \xi\left(\Pi'\left(\rho_{i+1}\right) + H_{i+1} - H_{i+\frac{1}{2}}\right)_+, \quad u^{+}_{i+\frac{1}{2}} = u_{i+1}. \end{aligned} \tag{12.11}$$

The source term $S_{H,i}$ is chosen to satisfy the well-balanced property of the deterministic version of system (12.4),

$$S_{H,i} = \frac{1}{\Delta x_i} \begin{pmatrix} 0 \\ P\left(\rho^{-}_{i+\frac{1}{2}}\right) - P\left(\rho^{+}_{i-\frac{1}{2}}\right) \end{pmatrix}. \tag{12.12}$$

The damping source term is simply approximated as

$$S_{D,i} = \begin{pmatrix} 0 \\ -\gamma \rho_i u_i \end{pmatrix}. \tag{12.13}$$

12.2   Spatiotemporal average for the noise source

The cell-average evaluation of the noise term $\mathcal{W}(x, t)$ is done via an spatiotemporal average, as proposed in [120] and applied later in our work of stochastic gradient flow equations in Chapter 11:

$$\mathcal{W}_i = \frac{1}{\Delta x \Delta t} \int_t^{t+\Delta t} \int_{x_{i-\frac{1}{2}}}^{x_{i+\frac{1}{2}}} \mathcal{W}(x, t) dx\, dt, \tag{12.14}$$

which, by definition of the white noise (12.3), is equal to a normal distribution with zero mean and variance $(\Delta x \Delta t)^{-1}$,

$$\mathcal{W}_i = \mathcal{N}(0,1)/\sqrt{\Delta x \Delta t}. \tag{12.15}$$

Note that Δt refers to the time step empoyed in the temporal discretization of the semi-discrete scheme (12.7).

It follows that the souce term $S_{\mathcal{W},i}$ satisfies

$$S_{\mathcal{W},i} = \begin{pmatrix} 0 \\ -\sqrt{k_B T \gamma \rho_i}\,\mathcal{W}_i \end{pmatrix}. \tag{12.16}$$

12.3 Temporal discretization

A myriad of stochastic temporal integrators can be applied to the semi-discrete scheme in (12.7). In our work on stochastic gradient flows in Chapter 11, we tested Euler-Maruyama, Milstein and Runge-Kutta-type integrators with the objective of analyzing their influence in the statistical properties of the numerical solution. Here we simply employ an explicit Euler-Maruyama integrator to the semi-discrete scheme in (12.7). By denoting as U_i^{n+1} and U_i^n as the cell averages of density and velocity at time $t = (n+1)\Delta t$ and $t = n\Delta t$, respectively, and Δt the time step, it follows that the semi-discrete scheme (12.7) becomes

$$U_i^{n+1} = U_i^n - \Delta t \left(\frac{F_{i+\frac{1}{2}}(U_i^n) - F_{i-\frac{1}{2}}(U_i^n)}{\Delta x_i} + S_{H,i}(U_i^n) + S_{D,i}(U_i^n) + S_{\mathcal{W},i}(U_i^n) \right), \tag{12.17}$$

where, for the explicit scheme employed here, all the fluxes and sources at the right-hand side depend on U_i^n and not on U_i^{n+1}.

Remark 12.3.1 (Positivity of the density). It can be shown that for the discrete system (12.17) the density remains positive at all times, as long as the CFL condition of the numerical flux is satisfied. The proof is the same as for the first-order well-balanced scheme in Chapter 5, given that the noise term does not affect the density equation and the interface values $\rho_{i+\frac{1}{2}}^-$ and $\rho_{i-\frac{1}{2}}^+$ are computed identically following (12.11).

12.4 Numerical tests

We test the numerical scheme for two particular choices of the free energy satisfying . Both have a pressure of the form $P(\rho) = \rho^m$, with $m \geq 1$. In Example 12.4.1 we select $m = 1$ so that the pressure satisfies the ideal-gas relation $P(\rho) = \rho$, and the density does not develop vacuum regions during the temporal evolution. For this case the employed numerical flux is the versatile local Lax-Friedrich flux. In Example 12.4.2 we select $m = 2$, and since $m > 1$ then vacuum regions with $\rho = 0$ are generated. This implies that the hyperbolicity of the system (12.1) is lost in those regions, and the local Lax-Friedrich scheme fails. In this case a kinetic solver based on [236] is employed to handle the vacuum regions. Both free energies also include an external potential of the form $V(x) = \frac{x^2}{2}$ and no interaction potential $W(x)$. The time step Δt is chosen to satisfy the CFL conditions of the corresponding numerical fluxes, and more details are available in Appendix 5.A. The constants satisfy $k_B = T = \gamma = 1$. The number of cells per simulation is 50.

The ensemble averages and variances in every cell for density, momentum, variation of the free energy, total energy and free energy are computed from 1000 trajectories. The discrete version of the variation of the free energy satisfies,

$$\left(\frac{\delta \mathcal{F}}{\delta \rho} \right)_i = \Pi'(\rho_i) + V_i + \sum_j \Delta x_j W_{ij}, \quad \text{where} \quad \rho \Pi''(\rho) = P'(\rho), \tag{12.18}$$

while the discrete total energy E^Δ and discrete free energy $\mathcal{F}^\Delta$ are

$$E^\Delta = \sum_i \frac{\Delta x_i}{2} \rho_i u_i^2 + \mathcal{F}^\Delta \quad \text{and} \quad \mathcal{F}^\Delta = \sum_i \Delta x_i \left[\Pi\left(\rho_i\right) + V_i \rho_i \right] + \frac{1}{2} \sum_{i,j} \Delta x_i \Delta x_j W_{ij} \rho_i \rho_j. \quad (12.19)$$

For comparison we also plot the temporal evolution of the deterministic system, solved with the well-balanced first-order finite-volume scheme proposed in Chapter 5.

Example 12.4.1 (Ideal-gas pressure and attractive potential). In this example the pressure satisfies $P(\rho) = \rho$ and there is an external potential of the form $V(x) = \frac{x^2}{2}$. As a result, the relation holding in the deterministic steady state is

$$\frac{\delta \mathcal{F}}{\delta \rho} = \Pi'(\rho) + H = \ln(\rho) + \frac{x^2}{2} = \text{constant on supp}(\rho) \text{ and } u = 0. \quad (12.20)$$

The steady state for the deterministic system, for an initial mass M_0, explicitly satisfies

$$\rho_\infty = M_0 \frac{e^{-x^2/2}}{\int_{\mathbb{R}} e^{-x^2/2} dx}. \quad (12.21)$$

The initial conditions applied in the simulation are

$$\rho(x, t = 0) = M_0 \frac{0.2 + 5 \cos\left(\frac{\pi x}{10}\right)}{\int_{\mathbb{R}} \left(0.2 + 5 \cos\left(\frac{\pi x}{10}\right)\right) dx}, \quad \rho u(x, t = 0) = -0.05 \sin\left(\frac{\pi x}{10}\right), \quad x \in [-5, 5], \quad (12.22)$$
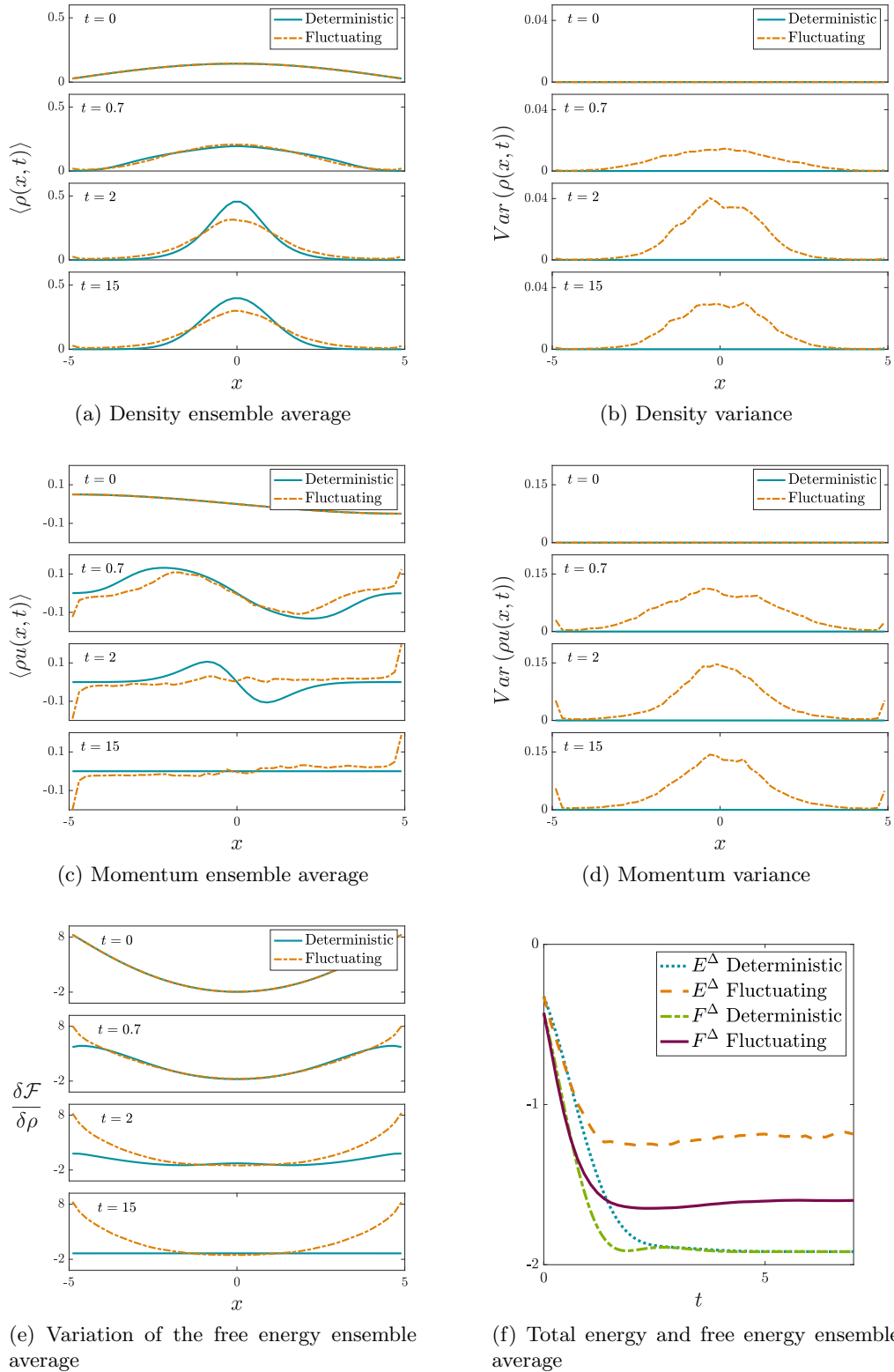
with M_0 equals to 1 so that the total mass is unitary.

Figure 12.4.1 depicts the temporal evolution of the system from $t = 0$ to $t = 15$. Let's compare the solutions of the deterministic and fluctuating simulations. On the one hand, at $t = 15$ the deterministic system has already reached its stationary state (12.21), and we can see that at $t = 15$ there is no momentum in Figure 12.4.1c. There is also a constant variation of the free energy in Figure 12.4.1e, in agreement with (12.20). Furthermore, the total energy and free energy match (due to the abcense of kinetic energy in the steady state) and have reached a plateau, as shown in Figure 12.4.1f.

On the other hand, the evolution of the fluctuating simulation presents some important differences. From Figure 12.4.1f we notice that both free and total energy seem to have reached a plateau. However, there is a gap between them because there is still kinetic energy in the system, and contrary to the deterministic case it cannot be fully dissipated because of the noise. We observe that in Figure 12.4.1c there is indeed momentum at $t = 15$, and the density ensemble average is not the same as the deterministic steady state in Figure 12.4.1a. The variation of the free energy is also not constant at $t = 15$, as shown in Figure 12.4.1e. We also plot the variances of the density and momentum in Figures 12.4.1b and 12.4.1d. In conclusion, the noise prevents the system from reaching a steady state.

Example 12.4.2 (Pressure proportional to square of density and attractive potential). For this example the pressure satisfies $P(\rho) = \rho^2$ and there is an external potential of the form $V(x) = \frac{x^2}{2}$. The choice of $P(\rho) = \rho^2$ implies that regions of vacuum where $\rho = 0$ appear in the evolution and steady solution of the system.

For the fluctuating simulation we find that the time step required by the CFL was extremely low, due to the presence of cells with close to vacuum density. As a result, it is not possible to advance the simulation by taking the exact CFL condition required by the kinetic solver. However, as shown in Chapter 11 and references therein, the fluctuations are not modelled correctly when the number of particles per cell is lower that 5-10 particles. This implies that the noise source

(a) Density ensemble average

(b) Density variance

(c) Momentum ensemble average

(d) Momentum variance

(e) Variation of the free energy ensemble average

(f) Total energy and free energy ensemble average

Figure 12.4.1: Temporal evolution of Example 12.4.1.

term produces numerical artifacts in cells with close to vacuum densities. Due to that we decide to

turn off fluctuations for the cells whose density satisfies $\rho_i < 10^{-4}$. This modification allows us to effectively run the fluctuating simulation, and agrees with the fact that in cells with density close to vacuum there are no enough particles to produce fluctuations.

The steady state for the deterministic system with an initial mass of M_0 satisfies

$$\rho_\infty(x) = \begin{cases} -\dfrac{1}{4}\left(x + \sqrt[3]{3M_0}\right)\left(x - \sqrt[3]{3M_0}\right) & \text{for} \quad x \in \left[-\sqrt[3]{3M_0}, \sqrt[3]{3M_0}\right], \\ 0 & \text{otherwise.} \end{cases} \tag{12.23}$$

The initial conditions applied in the simulation are

$$\rho(x, t = 0) = M_0 \frac{0.1 + e^{-x^2}}{\int_{\mathbb{R}} \left(0.1 + e^{-x^2}\right) dx}, \quad \rho u(x, t = 0) = -0.2 \sin\left(\frac{\pi x}{10}\right), \quad x \in [-5, 5],$$

with M_0 being the mass of the system and equal to 1.

Figure 12.4.2 depicts the temporal evolution of the system from $t = 0$ to $t = 15$. Similar remarks to the ones in Example 12.4.1 apply here. We observe that the deterministic simulation has reached a stady state at $t = 15$, and in Figure 12.4.2f the total and free energies match and have reached a plateau. There is no momentum at $t = 15$ in Figure 12.4.2c and the steady state density profile in Figure 12.4.2a is (12.23).

On the contrary, for the fluctuating simulation there is still kinetic energy at $t = 15$, and in Figure 12.4.2f we see that there is a gap between total and free energy. In Figure 12.4.2c we observe that the momentum is not zero, and in Figure 12.4.2a the density profile is different from the deterministic steady state.

(a) Density ensemble average

(b) Density variance

(c) Momentum ensemble average

(d) Momentum variance

(e) Variation of the free energy ensemble average

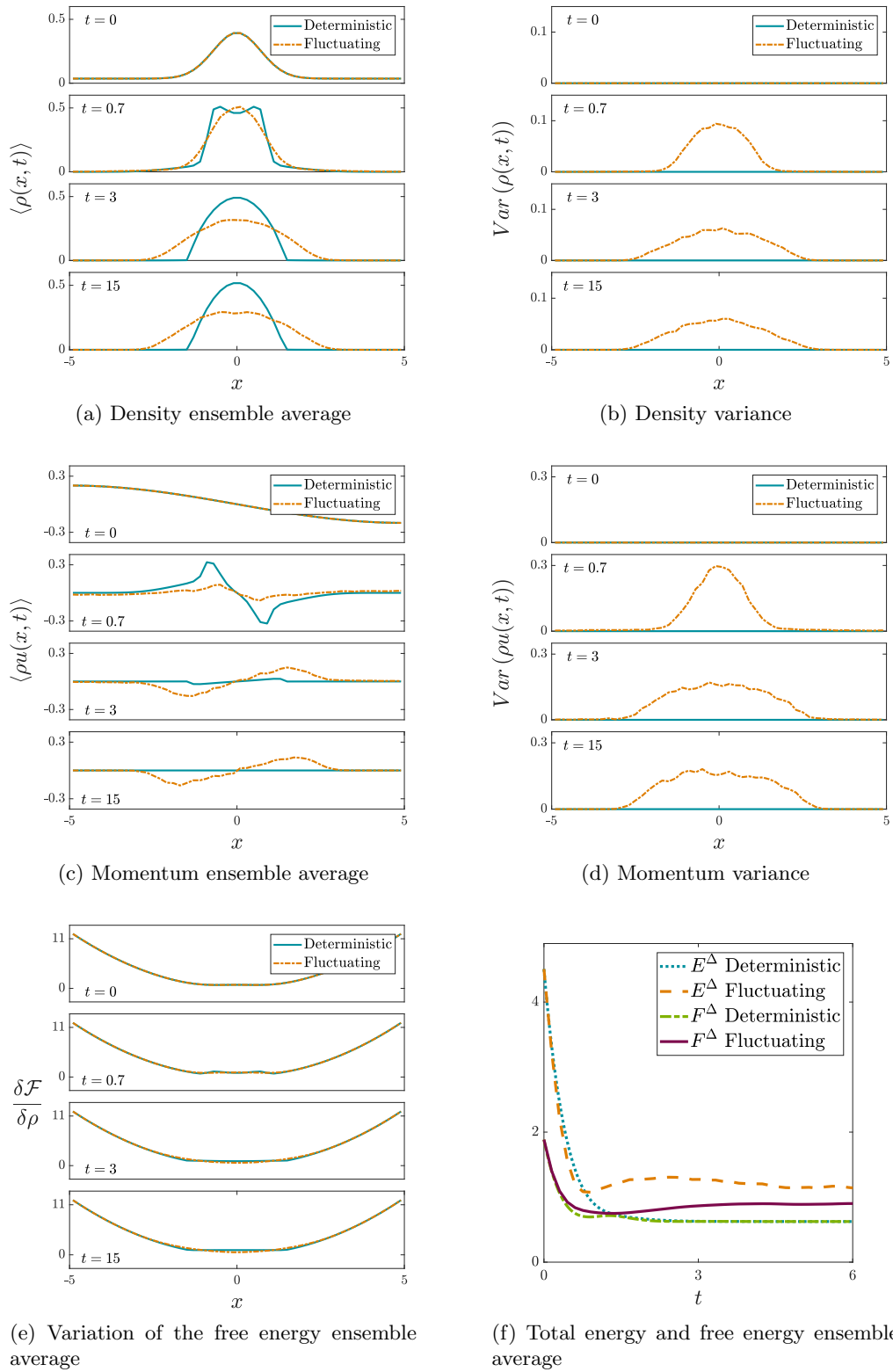(f) Total energy and free energy ensemble average

Figure 12.4.2: Temporal evolution of Example 12.4.2.

Part V

Conclusions

Conclusions and future perspectives

In this chapter we summarize the contributions of this thesis and point out several open questions for future research.

13.1 First- and second-order well-balanced schemes

We have introduced first- and second-order accurate finite volume schemes for a large family of hydrodynamic equations with general free energy, positivity preserving and free energy decaying properties. These hydrodynamic models with damping naturally arise in dynamic density functional theories and the accurate computation of their stable steady states is crucial to understand their phase transitions and stability properties. The models possess a common variational structure based on the physical free energy functional from statistical mechanics. The numerical schemes proposed capture very well steady states and their equilibration dynamics due to the crucial free energy decaying property resulting into well-balanced schemes. The schemes were validated in well-known test cases and the chosen numerical experiments corroborate these conclusions for intricate phase transitions and complicated free energies.

There are also several new avenues of possible future directions. Indeed, we believe the computational framework and associated methodologies presented here can be useful for the study of bifurcations and phase transitions for systems where the free energy is known from experiments only, either physical or in-silico ones, and then our framework can be adopted in a "data-driven" approach. Of particular extension would also be extension to multi-dimensional problems. Two-dimensional problems in particular would be of direct relevance to surface diffusion and therefore to technological processes in materials science and catalysis. We shall examine these and related problems in future studies.

13.2 High-order well-balanced schemes

We propose high-order well-balanced finite-volume schemes for a broad class of hydrodynamic systems with attractive-repulsive interaction forces and linear and nonlinear damping. Our schemes are suitable for free energies containing convolutions of an interaction potential with the density, which are essential for applications such as the Keller-Segel model, more general Euler-Poisson systems, or dynamic-density functional theory. Our schemes are also equipped with a nonnegative-density reconstruction which allows for vacuum regions during the simulation. We provide several prototypical examples from relevant applications, such as shallow-water equations with irregular

topographies, alignment in collective behaviour, and interactive free energies modelling biological or astrophysical systems.

Our work lays the foundations for the construction of well-balanced high-order schemes that may satisfy further fundamental properties of our hyperbolic system, such as the discrete versions of the energy dissipation and entropy identity, or even the well-balanced property for the challenging moving steady states. Developing schemes, amongst the class of positivity-preserving high-order schemes introduced in the present work, satisfying also the entropy stability and energy dissipating properties, is a challenging open question. Our simulations in Chapter 6 satisfy the discrete energy dissipation in practice, even if we are not able to prove it for our finite-volume scheme.

13.3 Finite-volume scheme for the Cahn-Hilliard equation

We have introduced finite-volume schemes for the Cahn-Hilliard equation that unconditionally and discretely satisfy the boundedness of the phase field and the free-energy dissipation. The scheme maintains these two fundamental properties for general potentials including the double-well and logarithmic, wetting conditions as well as more general wall free energies, and degenerate mobilities of the type vanishing when $\phi = \pm 1$. In contrast to previous works, the scheme is not restricted to only particular choices of the free energy potential such as the logarithmic one. In addition, for the case of the double-well potential we do not rely on truncated potentials. Furthermore, the scheme is efficiently extended to higher-dimensional configurations due its flexible and cost-saving dimensional-splitting nature thanks to an upwind and finite-volume formulation. The computational cost can be further reduced with a straightforward parallelization resulting from the dimensional-splitting approach. The numerical schemes are validated and tested in a variety of prototypical configurations with different numbers of dimensions and a rich variety of contact angles between droplets and substrates.

One of the challenges is how to increase the order of accuracy of the scheme. The satisfaction of these unconditional properties imposes a trade-off in the order of our scheme, which in this work is limited to first-order accuracy. Previous works [15, 55] have proposed finite-volume schemes of at least second order that prove similar properties but at the cost of depending on a CFL condition to be satisfied.

13.4 An application of the Cahn-Hilliard equation in image inpainting

We have quantified the prediction improvement of employing a CH image inpainting filter to restore damaged images which are then passed into a neural network. We combined a finite-volume scheme with a neural network for pattern recognition to develop an integrated algorithm summing up the process of adding damage to the images and then predicting their label. Our results for the MNIST dataset suggest that, in general, the accuracy is improved for a wide range of low to moderate damages, while for some particular cases we reach improvements of up to 50%. We also provide the image inpainting outcome of multiple damage scenarios and the benefits of adding the CH filter to predict the label of the image are easily visible.

We believe that our results employing the MNIST dataset lay the foundations towards the application of image inpainting in more complex datasets. Here we have demonstrated the benefit of combining the fields of image inpainting with machine learning, and we believe that many applications can take advantage of it. For instance there are applications such as medical images from MRI or satellite observations where there is typically some inherent noise or damage involved and where there may be potential to employ tools from machine learning as was done here.

MNIST is one of the most well-known and convenient benchmark datasets. It is possible that applying our methodology to increasingly complex datasets might bring about new challenges. At the same time we have relied on two main assumptions about the damage and the images: the first one is that the images are binary, leading to the standard CH potential that has only two wells (i.e. one for each of the two colours); the second one is that the damage is not blind, meaning that the location of the damage is known. Performing image inpainting without these two assumptions becomes substantially more involved as already pointed out in [45, 88, 281]. We will be exploring these and related questions in future works.

13.5 Finite-volume schemes for fluctuating hydrodynamics

We have developed an efficient and robust finite-volume numerical scheme for solving stochastic gradient-flow equations. The scheme was exemplified with FDDFT and allows us to scrutinise the effects of thermal fluctuations on complex phenomena such as phase transitions. Unlike previous numerical methodologies only applicable to a limited range of free energies (e.g. ideal-gas free energies such as in Refs [178, 275]), our proposed scheme deals effectively with general free-energy functionals including external fields or interacting potentials.

Our numerical methodology essentially comprises: a hybrid space discretization based on central and upwind schemes, for both deterministic and stochastic fluxes; a family of implicit-explicit Euler and Milsten time integrators, together with a weak second-order Runge-Kutta scheme; an adaptive time-step scheme, based on the Brownian bridge technique, which ensures the non-negativity of the density; appropriate boundary conditions. What is more, the hybrid approach provides an optimal compromise between statistical properties of the stochastic field and spurious oscillations. Additionally, the adaptive time-step feature of the scheme represents an alternative approach to preserving density positivity without including artificial limiters as in previous schemes.

The scheme is validated by means of several numerical applications. First, we study the variance, temporal and spatial correlations, and structure factor of an ideal gas at equilibrium, comparing the results of the finite-volume solver with theoretical results from the literature and our own MD simulations. In agreement with previous works, we find that a minimum number of $5 - 10$ particles per cell is required in order for FDDFT to match atomistic simulation results. We the examine the out-of-equilibrium evolution of an ideal gas in a double-well external potential. Our stochastic solver accurately reproduces local mean density, local density fluctuations and spatial correlations obtained from MD simulations. It should also be noted that for the deterministic case/DDFT where thermal fluctuations are not included, the results are consistent with both FDDFT and MD. Finally, we simulate homogeneous nucleation kinetics of a fluid consisting of particles interacting through an LJ-like potential. Our results for the phase diagram match the theoretical results and serve so as to illustrate the crucial role of fluctuations to surmount free-energy barriers. As expected, an exponential law is observed for the nucleation growth rate as function of the metastable free-energy barrier.

An open research question is the design of finite-volume schemes for the fluctuating-hydrodynamic system derived in [123], for the general case of arbitrarily shaped and thermalized particles. It consists of two stochastic PDEs for the density ρ and velocity $\boldsymbol{u}$ fields:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \quad \boldsymbol{x} \in \mathbb{R}^d, \quad t > 0, \\ \partial_t(\rho\boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u}) = -\rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho} - \gamma \rho \boldsymbol{u} - \sqrt{k_B T \gamma \rho}\, \boldsymbol{\mathcal{W}}(\boldsymbol{x}, t). \end{cases} \quad (13.1)$$

One of the challenges is how to adapt the numerical flux so that it preserves the statistical properties of the solution. In the case of the overdamped system in Chapter 11, an upwind numerical flux

introduces numerical diffusivity and dampens both the variance and the structural factor of the solution. As a result, we designed a hybrid scheme that combines a non-dissipative central scheme and an upwind scheme. For the case of the fluctuating-hydrodynamic system, it turns out that the versatile Lax-Friedrichs flux employed in Part II dampens again both the variance and the structural factor of the density and momentum. This time, however, relying on central numerical fluxes may not work as well for hyperbolic systems, due to for example their lack of numerical diffusion and the impossibility of proving a CFL condition. A challenging avenue for future work is to study how to modify the Lax-Friedrichs numerical flux so that it preserves the statistical properties of the solution.

Another open research area emerges when the density-momentum system in (13.1) is extended to include the equation for the energy e:

$$\begin{cases} \partial_t \rho + \nabla \cdot (\rho \boldsymbol{u}) = 0, \\ \partial_t(\rho \boldsymbol{u}) + \nabla \cdot (\rho \boldsymbol{u} \otimes \boldsymbol{u} + P) = \nabla \cdot \left(\sqrt{2T}\, \boldsymbol{\eta}(\boldsymbol{x}, t) \right), \\ \partial_t e + \nabla \cdot (\boldsymbol{u}\,(e + P)) = \nabla \cdot (k \nabla T) + \nabla \cdot \left(\sqrt{2T^2}\, \boldsymbol{\zeta}(\boldsymbol{x}, t) \right), \end{cases} \tag{13.2}$$

where P denotes the pressure, T the temperature and k the thermal conductivity. The noise terms $\boldsymbol{\eta}(\boldsymbol{x}, t)$ and $\boldsymbol{\zeta}(\boldsymbol{x}, t)$ are vectors of Gaussian stochastic processes delta-correlated in space and time, i.e.

$$\begin{aligned} \langle \boldsymbol{\eta}(\boldsymbol{x}, t) \rangle &= 0, \\ \langle \boldsymbol{\eta}(\boldsymbol{x}, t), \boldsymbol{\eta}(\boldsymbol{x}', t') \rangle &= 2\delta(t - t')\delta(\boldsymbol{x} - \boldsymbol{x}'). \end{aligned} \tag{13.3}$$

The total energy density satisfies

$$e = \frac{1}{2}\rho |\boldsymbol{u}|^2 + \rho \mathcal{U}(\rho, \theta). \tag{13.4}$$

The specific internal energy $\mathcal{U}(\rho, T)$ is defined through a suitable Equation of State (EoS), which is determined by the Helmholtz free energy $f(\rho, T)$. The internal energy is then derived from $\mathcal{U} = f + Ts$, with the entropy defined as $s = -\partial f / \partial T$ and the pressure as $P = \rho^2 \partial f / \partial \rho$. The Helmholtz free energy plays an analogous role as the free energy $\mathcal{F}[\rho]$ in the overdamped and density-momentum systems of this thesis: by adequately tuning it one can model a wide range of applications. However, this stochastic system is still not well understood even from the analytical point of view, and as a result the design of numerical schemes may be challenging due to the lack of analytical properties to be numerically mimicked.

13.6 Finite-volume schemes for challenging free energies

Most of the numerical simulations in this thesis are performed over free energies that satisfy the general form

$$\mathcal{F}[\rho] = \int_{\mathbb{R}^d} \Pi(\rho) d\boldsymbol{x} + \int_{\mathbb{R}^d} V(\boldsymbol{x})\rho(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} W(\boldsymbol{x} - \boldsymbol{y})\rho(\boldsymbol{x})\rho(\boldsymbol{y}) d\boldsymbol{x} d\boldsymbol{y}, \tag{13.5}$$

where

$$\rho \Pi''(\rho) = P'(\rho). \tag{13.6}$$

There are however more challenging free energies that will be really interesting to explore in future work. For instance, more complex free energies arise in applications related to (D)DFT [121, 147], see [70] for other related free energies and properties. These free energies have the form

$$\mathcal{F}[\rho] = \int_{\mathbb{R}^d} \Pi(\rho) d\boldsymbol{x} + \int_{\mathbb{R}^d} V(\boldsymbol{x})\rho(\boldsymbol{x}) d\boldsymbol{x} + \frac{1}{2} \int_{\mathbb{R}^d} K\left(W(\boldsymbol{x}) \star \rho(\boldsymbol{x})\right) \rho(\boldsymbol{x}) d\boldsymbol{x}, \tag{13.7}$$

where K is a function depending on the convolution of $\rho(\boldsymbol{x})$ with the kernel $W(\boldsymbol{x})$. Its variation with respect to the density satisfies

$$\frac{\delta \mathcal{F}}{\delta \rho} = \Pi'(\rho) + V(\boldsymbol{x}) + \frac{1}{2}K\left(W(\boldsymbol{x}) \star \rho\right) + \frac{1}{2}K'\left(W(\boldsymbol{x}) \star \rho\right)\left(W(\boldsymbol{x}) \star \rho\right). \tag{13.8}$$

Classical (D)DFT is a theoretical framework provided by nonequilibrium statistical mechanics but has increasingly become a widely-employed method for the computational scrutiny of the microscopic structure of both uniform and non-uniform fluids [121, 147, 202, 293, 294]. The starting point in (D)DFT is a functional $\mathcal{F}[\rho]$ for the fluid's free energy which encodes all microscopic information such as the ideal-gas part, short-range repulsive effects induced by molecular packing, attractive interactions and external fields. This functional can be exactly derived only for a limited number of applications, for instance the one-dimensional hard rod system from Percus [230]. However, in general it has to be approximated by making appropriate assumptions, as e.g. in the so-called fundamental-measure theory of Rosenfeld [241]. These assumptions are usually validated by carrying out appropriate test simulations (e.g. of the underlying stochastic dynamics) to compare e.g. the DDFT system with the approximate free-energy functional to the microscopic reference system [148].

In Chapter 5 we show that our well-balanced scheme can be applied to one of the simplest physical free-energy functionals employed in (D)DFT: the hard rods system in one dimension. There are however more sophisticated choices of free energy, such as the fundamental-measure theory of Rosenfeld [241], for which it would be interesting to design robust numerical schemes for the dynamical evolution.

Another area of promising research arises when considering free energies for which $\Pi(\rho)$ is nonconvex. This implies that $\Pi'(\rho)$ may not have an inverse function for positive densities ρ, which was a requirement when designing most of the numerical schemes in this thesis. An example of such free energy is the van der Waals EoS, where the pressure satisfies

$$P = \frac{8\rho T}{3 - \rho} - 3\rho^2. \tag{13.9}$$

For this kind of free energies the eigenvalues are complex, leading to systems that are elliptic instead of hyperbolic. This renders invalid many of the usual strategies in computational fluid dynamics, such as numerical fluxes or characteristic decompositions. The challenge here would be to adapt the schemes presented in this thesis to these nonconvex free energy. An original first strategy is proposed in [255], where the authors put forward a technique that allows to keep the eigenvalues real and distinct, based on modifying the numerical flux of simple overdamped density equations.

13.7 Data-driven approaches to free-energy systems

In the last years machine learning has gained much traction in areas such as autonomous vehicles, computer vision, recommender systems or natural language processing. However, for scientific computing the application of machine learning has been less explored, although recently there has been a surge of interest. Some potential areas of research that build from the tools in this thesis are:

- Data-driven approaches [292] that, given data on collective motion of classical particles, characterise the set of free energies which can describe that particle system.

- Neural ODEs [85] that act as a black-box differential equation solver and are able to trade numerical precision for speed.

- Parameter calibration in agent-based models [205], where the parameters of a certain interaction law or free energy are learned via Gaussian process regression model implemented with Bayesian optimization.

- Physics-informed neural networks (PINNs) [239] are trained to solve supervised learning tasks while respecting any given laws of physics described by general nonlinear PDEs.

Part VI

Bibliography

[1] H. Abels and M. Wilke. "Convergence to equilibrium for the Cahn–Hilliard equation with a logarithmic free energy". In: *Nonlinear Analysis: Theory, Methods & Applications* 67.11 (2007), pp. 3176–3193.

[2] M. Abramowitz, I. A. Stegun, and R. H. Romer. *Handbook of mathematical functions with formulas, graphs, and mathematical tables.* 1988.

[3] G. Albi, L. Pareschi, G. Toscani, and M. Zanella. "Recent advances in opinion modeling: control and social influence". In: *Active Particles, Volume 1* (2017), pp. 49–98.

[4] L. Ambrosio, N. Gigli, and G. Savaré. *Gradient flows: in metric spaces and in the space of probability measures.* Springer Science & Business Media, 2008.

[5] F. Aràndiga, A. Baeza, A. Belda, and P. Mulet. "Analysis of WENO schemes for full and global accuracy". In: *SIAM Journal on Numerical Analysis* 49.2 (2011), pp. 893–915.

[6] A. J. Archer. "Dynamical density functional theory for molecular and colloidal fluids: A microscopic approach to fluid mechanics". In: *J. Chem. Phys* 130.1 (2009), p. 014509.

[7] A. J. Archer and M. Rauscher. "Dynamical density functional theory for interacting Brownian particles: stochastic or deterministic?" In: *J. Phys. A* 37.40 (2004), pp. 9325–9333.

[8] A. C. Aristotelous, O. Karakashian, and S. M. Wise. "A mixed discontinuous Galerkin, convex splitting scheme for a modified Cahn-Hilliard equation and an efficient nonlinear multigrid solver". In: *Discrete & Continuous Dynamical Systems-B* 18.9 (2013), p. 2211.

[9] P. J. Atzberger. "Spatially adaptive stochastic numerical methods for intrinsic fluctuations in reaction–diffusion systems". In: *J. Comput. Phys.* 229.9 (2010), pp. 3474–3501.

[10] E Audusse, C Chalons, and P Ung. "A very simple well-balanced positive and entropy-satisfying scheme for the shallow-water equations". In: *Commun. Math. Sci* 13.5 (2015), pp. 1317–1332.

[11] E. Audusse, F. Bouchut, M.-O. Bristeau, R. Klein, and B. t. Perthame. "A fast and stable well-balanced scheme with hydrostatic reconstruction for shallow water flows". In: *SIAM J. Sci. Comput.* 25.6 (2004), pp. 2050–2065.

[12] E. Audusse and M.-O. Bristeau. "A well-balanced positivity preserving "second-order" scheme for shallow water flows on unstructured meshes". In: *J. Comput. Phys.* 206.1 (2005), pp. 311–333.

[13] B. Aymard, U. Vaes, M. Pradas, and S. Kalliadasis. "A linear, second-order, energy stable, fully adaptive finite element method for phase-field modelling of wetting phenomena". In: *Journal of Computational Physics: X* 2 (2019), p. 100010.

[14] R. Bailo, J. A. Carrillo, and J. Hu. "Bound-Preserving Finite-Volume Schemes for Continuity Equations with Saturation". In: *to be submitted* (2021).

[15] R. Bailo, J. A. Carrillo, and J. Hu. "Fully Discrete Positivity-Preserving and Energy-Dissipating Schemes for Aggregation-Diffusion Equations with a Gradient Flow Structure". In: *Commun. Math. Sci.* 18.5 (2020), pp. 1259–1303.

[16] R. Bailo, J. A. Carrillo, S. Kalliadasis, and S. P. Perez. "Unconditional bound-preserving and energy-dissipating finite-volume schemes for the Cahn-Hilliard equation". In: *arXiv preprint arXiv:2105.05351* (2021).

[17] R. Bailo, J. A. Carrillo, H. Murakawa, and M. Schmidtchen. "Convergence of a Fully Discrete and Energy-Dissipating Finite-Volume Scheme for Aggregation-Diffusion Equations". In: *arXiv: 2002.10821 (M3AS, to appear)* (2020).

[18] K. Balakrishnan, A. L. Garcia, A. Donev, and J. B. Bell. "Fluctuating hydrodynamics of multispecies nonreactive mixtures". In: *Phys. Rev. E* 89.1 (2014), p. 013017.

[19] F. Balboa, J. B. Bell, R. Delgado-Buscalioni, A. Donev, T. G. Fai, B. E. Griffith, and C. S. Peskin. "Staggered schemes for fluctuating hydrodynamics". In: *Multiscale Model. Sim.* 10.4 (2012), pp. 1369–1408.

[20] L. Baňas and R. Nürnberg. "Adaptive finite element methods for Cahn–Hilliard equations". In: *Journal of computational and applied mathematics* 218.1 (2008), pp. 2–11.

[21] A. B. T. Barbaro, J. A. Canizo, J. A. Carrillo, and P. Degond. "Phase Transitions in a Kinetic Flocking Model of Cucker–Smale Type". In: *Multiscale Model. Simul.* 14.3 (2016), pp. 1063–1088.

[22] J. A. Barker and D. Henderson. "Perturbation Theory and Equation of State for Fluids. II. A Successful Theory of Liquids". In: *J. Chem. Phys.* 47.11 (1967), pp. 4714–4721.

[23] J. Barré, P. Degond, and E. Zatorska. "Kinetic theory of particle interactions mediated by dynamical networks". In: *Multiscale Modeling & Simulation* 15.3 (2017), pp. 1294–1323.

[24] J. W. Barrett, J. F. Blowey, and H. Garcke. "Finite Element Approximation of the Cahn–Hilliard Equation with Degenerate Mobility". In: *SIAM Journal on Numerical Analysis* 37.1 (1999), pp. 286–318.

[25] J. W. Barrett, J. F. Blowey, and H. Garcke. "On fully practical finite element approximations of degenerate Cahn-Hilliard systems". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 35.4 (2001), pp. 713–748.

[26] J. B. Bell, A. L. Garcia, and S. A. Williams. "Computational fluctuating fluid dynamics". In: *ESAIM: Math. Model. Numer. Anal.* 44.5 (2010), pp. 1085–1105.

[27] J. B. Bell, A. L. Garcia, and S. A. Williams. "Numerical methods for the stochastic Landau-Lifshitz Navier-Stokes equations". In: *Phys. Rev. E* 76.1 (2007), p. 016708.

[28] N. Bellomo, A. Bellouquid, Y. Tao, and M. Winkler. "Toward a mathematical theory of Keller–Segel models of pattern formation in biological tissues". In: *Math. Model. Methods Appl. Sci.* 25.09 (2015), pp. 1663–1763.

[29] A. Bermudez and M. E. Vázquez. "Upwind methods for hyperbolic conservation laws with source terms". In: *Comput. Fluids* 23.8 (1994), pp. 1049–1071.

[30] M. Bertalmio, A. L. Bertozzi, and G. Sapiro. "Navier-stokes, fluid dynamics, and image and video inpainting". In: *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition. CVPR 2001.* Vol. 1. IEEE. 2001, pp. I–I.

[31] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester. "Image inpainting". In: *Proceedings of the 27th annual conference on Computer graphics and interactive techniques.* 2000, pp. 417–424.

[32] A. Bertozzi, S. Esedoglu, and A. Gillette. "Analysis of a two-scale Cahn–Hilliard model for binary image inpainting". In: *Multiscale Modeling & Simulation* 6.3 (2007), pp. 913–936.

[33] A. L. Bertozzi, S. Esedoglu, and A. Gillette. "Inpainting of binary images using the Cahn–Hilliard equation". In: *IEEE Transactions on image processing* 16.1 (2006), pp. 285–291.

[34] M. Bessemoulin-Chatard and F. Filbet. "A finite volume scheme for nonlinear degenerate parabolic equations". In: *SIAM J. Sci. Comput.* 34.5 (2012), B559–B583.

[35] A. K. Bhattacharjee, K. Balakrishnan, A. L. Garcia, J. B. Bell, and A. Donev. "Fluctuating hydrodynamics of multi-species reactive mixtures". In: *J. Chem. Phys.* 142.22 (2015), p. 224107.

[36] S. Bian and J.-G. Liu. "Dynamic and steady states for multi-dimensional Keller-Segel model with diffusion exponent $m > 0$". In: *Communications in Mathematical Physics* 323.3 (2013), pp. 1017–1070.

[37] J. Binney and S. Tremaine. *Galactic dynamics*. Princeton university press, 2011.

[38] M. Bixon and R. Zwanzig. "Boltzmann-Langevin Equation and Hydrodynamic Fluctuations". In: *Phys. Rev.* 187 (1 1969), pp. 267–272.

[39] F. Bolley, J. A. Canizo, and J. A. Carrillo. "Stochastic mean-field limit: non-Lipschitz forces and swarming". In: *Mathematical Models and Methods in Applied Sciences* 21.11 (2011), pp. 2179–2210.

[40] J. Bosch, D. Kay, M. Stoll, and A. J. Wathen. "Fast solvers for Cahn–Hilliard inpainting". In: *SIAM Journal on Imaging Sciences* 7.1 (2014), pp. 67–97.

[41] J. Bosch and M. Stoll. "A fractional inpainting model based on the vector-valued Cahn–Hilliard equation". In: *SIAM Journal on Imaging Sciences* 8.4 (2015), pp. 2352–2382.

[42] F. Bouchut. *Nonlinear stability of finite Volume Methods for hyperbolic conservation laws: And Well-Balanced schemes for sources*. Springer Science & Business Media, 2004.

[43] F. Boyer, L. Chupin, and P. Fabrie. "Numerical study of viscoelastic mixtures through a Cahn–Hilliard flow model". In: *European Journal of Mechanics - B/Fluids* 23.5 (2004), pp. 759 –780. ISSN: 0997-7546.

[44] C. Buet, S. Dellacherie, et al. "On the Chang and Cooper scheme applied to a linear Fokker-Planck equation". In: *Communications in Mathematical Sciences* 8.4 (2010), pp. 1079–1090.

[45] M. Burger, L. He, and C.-B. Schönlieb. "Cahn–Hilliard inpainting and a generalization for grayvalue images". In: *SIAM Journal on Imaging Sciences* 2.4 (2009), pp. 1129–1167.

[46] L. A. Caffarelli and N. E. Muler. "An L∞ bound for solutions of the Cahn-Hilliard equation". In: *ArRMA* 133.2 (1995), pp. 129–144.

[47] J. W. Cahn. "Critical point wetting". In: *The Journal of Chemical Physics* 66.8 (1977), pp. 3667–3672.

[48] J. W. Cahn and J. E. Hilliard. "Free energy of a nonuniform system. I. Interfacial free energy". In: *The Journal of chemical physics* 28.2 (1958), pp. 258–267.

[49] V. Calvez, J. A. Carrillo, and F. Hoffmann. "Equilibria of homogeneous functionals in the fair-competition regime". In: *Nonlinear Anal.* 159 (2017), pp. 85–128.

[50] V. Calvez, J. A. Carrillo, and F. Hoffmann. "The geometry of diffusing and self-attracting particles in a one-dimensional fair-competition regime". In: *Nonlocal Nonlinear Diffus. Interact. New Methods Dir.* Springer, 2017, pp. 1–71.

[51] A. Canestrelli, A. Siviglia, M. Dumbser, and E. F. Toro. "Well-balanced high-order centred schemes for non-conservative hyperbolic systems. Applications to shallow water equations with fixed and mobile bed". In: *Adv. Water Resour.* 32.6 (2009), pp. 834–844.

[52] G Capdeville. "A central WENO scheme for solving hyperbolic conservation laws on non-uniform meshes". In: *Journal of Computational Physics* 227.5 (2008), pp. 2977–3014.

[53] N. F. Carnahan and K. E. Starling. "Equation of State for Nonattracting Rigid Spheres". In: *J. Chem. Phys.* 51.2 (1969), pp. 635–636.

[54] J. A. Carrillo, M. J. Castro, S. Kalliadasis, and S. P. Perez. "High-Order Well-Balanced Finite-Volume Schemes for Hydrodynamic Equations With Nonlocal Free Energy". In: *SIAM Journal on Scientific Computing* 43.2 (2021), A828–A858.

[55] J. A. Carrillo, A. Chertock, and Y. Huang. "A finite-volume method for nonlinear nonlocal equations with a gradient flow structure". In: *Commun. Comput. Phys.* 17.1 (2015), pp. 233–258.

[56] J. A. Carrillo, Y.-P. Choi, and S. P. Perez. "A review on attractive–repulsive hydrodynamics for consensus in collective behavior". In: *Act. Part. Vol. 1*. Springer, 2017, pp. 259–298.

[57] J. A. Carrillo, Y.-P. Choi, E. Tadmor, and C. Tan. "Critical thresholds in 1D Euler equations with non-local forces". In: *Math. Model. Methods Appl. Sci.* 26.01 (2016), pp. 185–206.

[58] J. A. Carrillo, Y.-P. Choi, and E. Zatorska. "On the pressureless damped Euler–Poisson equations with quadratic confinement: Critical thresholds and large-time behavior". In: *Math. Model. Methods Appl. Sci.* 26.12 (2016), pp. 2311–2340.

[59] J. A. Carrillo, K. Craig, and Y. Yao. "Aggregation-diffusion equations: dynamics, asymptotics, and singular limits". In: *Active Particles, Volume 2*. Springer, 2019, pp. 65–108.

[60] J. A. Carrillo, E. Feireisl, P. Gwiazda, and A. Świerczewska-Gwiazda. "Weak solutions for Euler systems with non-local interactions". In: *J. London Math. Soc.* 95.3 (2017), pp. 705–724.

[61] J. A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. "Asymptotic flocking dynamics for the kinetic Cucker–Smale model". In: *SIAM Journal on Mathematical Analysis* 42.1 (2010), pp. 218–236.

[62] J. A. Carrillo, M. Fornasier, G. Toscani, and F. Vecil. "Particle, kinetic, and hydrodynamic models of swarming". In: *Mathematical modeling of collective behavior in socio-economic and life sciences*. Springer, 2010, pp. 297–336.

[63] J. A. Carrillo, F. Hoffmann, E. Mainini, and B. Volzone. "Ground states in the diffusion-dominated regime". In: *Calculus of variations and partial differential equations* 57.5 (2018), p. 127.

[64] J. A. Carrillo, Y. Huang, and S. Martin. "Explicit flock solutions for quasi-morse potentials". In: *Eur. J. Appl. Math.* 25.5 (2014), pp. 553–578.

[65] J. A. Carrillo, S. Kalliadasis, F. Liang, and S. P. Perez. "Enhancement of damaged-image prediction through Cahn-Hilliard Image Inpainting". In: *Royal Society Open Science* 8 (2021), p. 201294.

[66] J. A. Carrillo, S. Kalliadasis, S. P. Perez, and C.-W. Shu. "Well-balanced finite-volume schemes for hydrodynamic equations with general free energy". In: *Multiscale Modeling & Simulation* 18.1 (2020), pp. 502–541.

[67] J. A. Carrillo, R. J. McCann, and C. Villani. "Contractions in the 2-Wasserstein Length Space and Thermalization of Granular Media". In: *Arch. Ration. Mech. Anal.* 179.2 (2006), pp. 217–263. ISSN: 1432-0673.

[68] J. A. Carrillo, R. J. McCann, C. Villani, et al. "Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates". In: *Rev. Mat. Iberoam.* 19.3 (2003), pp. 971–1018.

[69] J. A. Carrillo, A. Wróblewska-Kamińska, and E. Zatorska. "On long-time asymptotics for viscous hydrodynamic models of collective behavior with damping and nonlocal interactions". In: *arXiv Prepr. arXiv1709.09290* (2018).

[70] J. A. Carrillo, K. Craig, and F. S. Patacchini. "A blob method for diffusion". In: *Calc. Var. Partial Diff.* 58.2 (2019), p. 53.

[71] J. A. Carrillo, S. Kalliadasis, S. P. Perez, and C.-W. Shu. *Well-balanced finite volume schemes for hydrodynamic equations with general free energy*. to appear. 2020.

[72] V. Caselles, J.-M. Morel, and C. Sbert. "An axiomatic approach to image interpolation". In: *IEEE Transactions on image processing* 7.3 (1998), pp. 376–386.

[73] M. J. Castro, J. A. López-García, and C. Parés. "High order exactly well-balanced numerical methods for shallow water systems". In: *Journal of Computational Physics* 246 (2013), pp. 242 –264.

[74] M. Castro, J. Gallardo, and C. Parés. "High order finite volume schemes based on reconstruction of states for solving hyperbolic systems with nonconservative products. Applications to shallow-water systems". In: *Mathematics of computation* 75.255 (2006), pp. 1103–1134.

[75] M. J. Castro, T. M. de Luna, and C. Parés. "Well-balanced schemes and path-conservative numerical methods". In: *Handbook of Numerical Analysis*. Vol. 18. Elsevier, 2017, pp. 131–175.

[76] M. J. Castro, A. Pardo Milanés, and C. Parés. "Well-balanced numerical schemes based on a generalized hydrostatic reconstruction technique". In: *Mathematical Models and Methods in Applied Sciences* 17.12 (2007), pp. 2055–2113.

[77] M. J. Castro and C. Parés. "Well-Balanced High-Order Finite Volume Methods for Systems of Balance Laws". In: *Journal of Scientific Computing* 82 (2020), p. 48.

[78] M. J. Castro and M. Semplice. "Third-and fourth-order well-balanced schemes for the shallow water equations based on the CWENO reconstruction". In: *International Journal for Numerical Methods in Fluids* 89.8 (2019), pp. 304–325.

[79] M. Cates, J.-C. Desplat, P Stansell, A. Wagner, K Stratford, R Adhikari, and I Pagonabarraga. "Physical and computational scaling issues in lattice Boltzmann simulations of binary fluid mixtures". In: *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* 363.1833 (2005), pp. 1917–1935.

[80] C. Chalons, P. Goatin, and L. M. Villada. "High-Order Numerical Schemes for One-Dimensional Nonlocal Conservation Laws". In: *SIAM Journal on Scientific Computing* 40.1 (2018), A288–A305.

[81] P.-H. Chavanis. "Hamiltonian and Brownian systems with long-range interactions: V. Stochastic kinetic equations and theory of fluctuations". In: *Physica A* 387.23 (2008), pp. 5716–5740.

[82] P.-H. Chavanis. "Jeans type instability for a chemotactic model of cellular aggregation". In: *Eur. Phys. J. B-Condensed Matter Complex Syst.* 52.3 (2006), pp. 433–443.

[83] P.-H. Chavanis and C. Sire. "Kinetic and hydrodynamic models of chemotactic aggregation". In: *Phys. A Stat. Mech. its Appl.* 384.2 (2007), pp. 199–222.

[84] C. Chen and X. Yang. "Fast, provably unconditionally energy stable, and second-order accurate algorithms for the anisotropic Cahn–Hilliard model". In: *Computer Methods in Applied Mechanics and Engineering* 351 (2019), pp. 35–59.

[85] R. T. Chen, Y. Rubanova, J. Bettencourt, and D. Duvenaud. "Neural ordinary differential equations". In: *arXiv preprint arXiv:1806.07366* (2018).

[86] W. Chen, C. Wang, X. Wang, and S. M. Wise. "Positivity-preserving, energy stable numerical schemes for the Cahn-Hilliard equation with logarithmic potential". In: *Journal of Computational Physics: X* 3 (2019), p. 100031.

[87] Y. Cheng, A. Chertock, M. Herty, A. Kurganov, and T. Wu. "A new approach for designing moving-water equilibria preserving schemes for the shallow water equations". In: *Journal of Scientific Computing* 80.1 (2019), pp. 538–554.

[88] L. Cherfils, H. Fakih, and A. Miranville. "A Cahn–Hilliard system with a fidelity term for color image inpainting". In: *Journal of Mathematical Imaging and Vision* 54.1 (2016), pp. 117–131.

[89] L. Cherfils, H. Fakih, and A. Miranville. "A complex version of the Cahn–Hilliard equation for grayscale image inpainting". In: *Multiscale Modeling & Simulation* 15.1 (2017), pp. 575–605.

[90] A. Chertock, S. Cui, A. Kurganov, Ş. N. Özcan, and E. Tadmor. "Well-balanced schemes for the Euler equations with gravitation: Conservative formulation using global fluxes". In: *Journal of Computational Physics* 358 (2018), pp. 36–52.

[91] Y.-P. Choi, S.-Y. Ha, and Z. Li. "Emergent dynamics of the Cucker–Smale flocking model and its variants". In: *Active Particles, Volume 1*. Springer, 2017, pp. 299–331.

[92] A. Christlieb, J. Jones, K. Promislow, B. Wetton, and M. Willoughby. "High accuracy solutions to energy gradient flows from material science models". In: *Journal of Computational Physics* 257 (2014), pp. 193–215.

[93] V. Churuksaeva and A. Starchenko. "Mathematical modeling of a river stream based on a shallow water approach". In: *Procedia Computer Science* 66 (2015), pp. 200–209.

[94] N. Condette, C. Melcher, and E. Süli. "Spectral approximation of pattern-forming nonlinear evolution equations with double-well potentials of quadratic growth". In: *Mathematics of computation* 80.273 (2011), pp. 205–223.

[95] M. Copetti and C. M. Elliott. "Numerical analysis of the Cahn-Hilliard equation with a logarithmic free energy". In: *Numerische Mathematik* 63.1 (1992), pp. 39–65.

[96] D. Coppersmith and S. Winograd. "Matrix multiplication via arithmetic progressions". In: *Proceedings of the nineteenth annual ACM symposium on Theory of computing*. 1987, pp. 1–6.

[97] S. Cordier and E. Grenier. "Quasineutral limit of an Euler-Poisson system arising from plasma physics". In: *Communications in Partial Differential Equations* 25.5-6 (2000), pp. 1099–1113.

[98] L. Cozzolino, V. Pepe, L. Cimorelli, A. D'Aniello, R. Della Morte, and D. Pianese. "The solution of the dam-break problem in the Porous Shallow water Equations". In: *Advances in Water Resources* 114 (2018), pp. 83–101.

[99] I. Cravero and M. Semplice. "On the accuracy of WENO and CWENO reconstructions of third order on nonuniform meshes". In: *Journal of Scientific Computing* 67.3 (2016), pp. 1219–1246.

[100] A. Criminisi, P. Perez, and K. Toyama. "Object removal by exemplar-based inpainting". In: *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings*. Vol. 2. IEEE. 2003, pp. II–II.

[101] A. Criminisi, P. Pérez, and K. Toyama. "Region filling and object removal by exemplar-based image inpainting". In: *IEEE Transactions on image processing* 13.9 (2004), pp. 1200–1212.

[102] F. Cucker and S. Smale. "Emergent behavior in flocks". In: *IEEE Trans. Automat. Contr.* 52.5 (2007), pp. 852–862.

[103] F. Cucker and S. Smale. "On the mathematics of emergence". In: *Japanese J. Math.* 2.1 (2007), pp. 197–227.

[104] F. Cucker, S. Smale, and D.-X. Zhou. "Modeling language evolution". In: *Foundations of Computational Mathematics* 4.3 (2004), pp. 315–343.

[105] L. Cueto-Felgueroso and J. Peraire. "A time-adaptive finite volume method for the Cahn–Hilliard and Kuramoto–Sivashinsky equations". In: *Journal of Computational Physics* 227.24 (2008), pp. 9985–10017.

[106] B. A. De Dios, J. A. Carrillo, and C.-W. Shu. "Discontinuous Galerkin methods for the multi-dimensional Vlasov–Poisson problem". In: *Math. Model. Methods Appl. Sci.* 22.12 (2012), p. 1250042.

[107] G. De Fabritiis, R. Delgado-Buscalioni, and P. V. Coveney. "Multiscale modeling of liquids with molecular specificity". In: *Phys. Rev. Lett.* 97.13 (2006), p. 134501.

[108] G. De Fabritiis, M. Serrano, R. Delgado-Buscalioni, and P. V. Coveney. "Fluctuating hydrodynamic modeling of fluids at the nanoscale". In: *Phys. Rev. E* 75.2 (2007), p. 026307.

[109] D. S. Dean. "Langevin equation for the density of a system of interacting Langevin processes". In: *J. Phys. A: Math. Gen.* 29.24 (1996), p. L613.

[110] R. Delgado-Buscalioni and G. De Fabritiis. "Embedding molecular dynamics within fluctuating hydrodynamics in multiscale simulations of liquids". In: *Phys. Rev. E* 76.3 (2007), p. 036709.

[111] S. Delong, B. E. Griffith, E. Vanden-Eijnden, and A. Donev. "Temporal integrators for fluctuating hydrodynamics". In: *Phys. Rev. E* 87.3 (2013), p. 033302.

[112] L. Deng. "The MNIST database of handwritten digit images for machine learning research". In: *IEEE Signal Processing Magazine* 29.6 (2012), pp. 141–142.

[113] Y. Deng, T.-P. Liu, T. Yang, and Z.-a. Yao. "Solutions of Euler-Poisson Equations for Gaseous Stars". In: *Arch. Ration. Mech. Anal.* 164.3 (2002), pp. 261–285.

[114] A. E. Diegel, C. Wang, and S. M. Wise. "Stability and convergence of a second-order mixed finite element method for the Cahn–Hilliard equation". In: *IMA Journal of Numerical Analysis* 36.4 (2016), pp. 1867–1897.

[115] W Dieterich, H. Frisch, and A Majhofer. "Nonlinear diffusion and density functional theory". In: *Z. Phys. B* 78.2 (1990), pp. 317–323.

[116] *Digit Recognizer competition in Kaggle.* https://www.kaggle.com/c/digit-recognizer/overview. Accessed: 2020-03-05.

[117] M. Doi. *Soft matter physics.* Oxford University Press, 2013.

[118] A. Donev, J. B. Bell, A. L. Garcia, and B. J. Alder. "A hybrid particle-continuum method for hydrodynamics of complex fluids". In: *Multiscale Model. Sim.* 8.3 (2010), pp. 871–911.

[119] A. Donev and E. Vanden-Eijnden. "Dynamic density functional theory with hydrodynamic interactions and fluctuations". In: *J. Chem. Phys.* 140.23 (2014), p. 234115.

[120] A. Donev, E. Vanden-Eijnden, A. Garcia, and J. Bell. "On the accuracy of finite-volume schemes for fluctuating hydrodynamics". In: *Commun. Appl. Math. Comput. Sci.* 5.2 (2010), pp. 149–197.

[121] M. A. Durán-Olivencia, B. D. Goddard, and S. Kalliadasis. "Dynamical density functional theory for orientable colloids including inertia and hydrodynamic interactions". In: *J. Stat. Phys.* 164.4 (2016), pp. 785–809.

[122] M. A. Durán-Olivencia, R. S. Gvalani, S. Kalliadasis, and G. A. Pavliotis. "Instability, Rupture and Fluctuations in Thin Liquid Films: Theory and Computations". In: *J. Stat. Phys.* 174.3 (2019), pp. 579–604. ISSN: 1572-9613.

[123] M. A. Durán-Olivencia, P. Yatsyshin, B. D. Goddard, and S. Kalliadasis. "General framework for fluctuating dynamic density functional theory". In: *New J. Phys* 19.12 (2017), p. 123022.

[124] M. A. Durán-Olivencia, P. Yatsyshin, S. Kalliadasis, and J. F. Lutsko. "General framework for nonclassical nucleation". In: *New J. Phys.* 20.8 (2018), p. 083019.

[125] C. M. Elliott and H. Garcke. "On the Cahn–Hilliard equation with degenerate mobility". In: *Siam journal on mathematical analysis* 27.2 (1996), pp. 404–423.

[126] P. Español and A. Donev. "Coupling a nano-particle with isothermal fluctuating hydrodynamics: Coarse-graining from microscopic to mesoscopic dynamics". In: *J. Chem. Phys.* 143.23 (2015), p. 234104.

[127] R. Evans. "The nature of the liquid-vapour interface and other topics in the statistical mechanics of non-uniform, classical fluids". In: *Adv. Phys.* 28.2 (1979), pp. 143–200.

[128] D. J. Eyre. "Unconditionally gradient stable time marching the Cahn-Hilliard equation". In: *MRS Online Proceedings Library Archive* 529 (1998).

[129] F. Filbet. "A finite volume scheme for the Patlak–Keller–Segel chemotaxis model". In: *Numerische Mathematik* 104.4 (2006), pp. 457–488.

[130] F. Filbet, P. Laurençot, and B. Perthame. "Derivation of hyperbolic models for chemosensitive movement". In: *J. Math. Biol.* 50.2 (2005), pp. 189–207.

[131] F. Filbet and C.-W. Shu. "Approximation of hyperbolic models for chemosensitive movement". In: *SIAM J. Sci. Comput.* 27.3 (2005), pp. 850–872.

[132] U. S. Fjordholm, S. Mishra, and E. Tadmor. "Arbitrarily high-order accurate entropy stable essentially nonoscillatory schemes for systems of conservation laws". In: *SIAM Journal on Numerical Analysis* 50.2 (2012), pp. 544–573.

[133] R. F. Fox and G. E. Uhlenbeck. "Contributions to Non-Equilibrium Thermodynamics. I. Theory of Hydrodynamical Fluctuations". In: *Phys. Fluids* 13.8 (1970), pp. 1893–1902.

[134] F. Frank, A. Rupp, and D. Kuzmin. "Bound-preserving flux limiting schemes for DG discretizations of conservation laws with applications to the Cahn–Hilliard equation". In: *Computer Methods in Applied Mechanics and Engineering* 359 (2020), p. 112665.

[135] D. Furihata. "A stable and conservative finite difference scheme for the Cahn-Hilliard equation". In: *Numerische Mathematik* 87.4 (2001), pp. 675–699.

[136] J. G. Gaines and T. J. Lyons. "Variable step size control in the numerical solution of stochastic differential equations". In: *SIAM J. Appl. Math.* 57.5 (1997), pp. 1455–1484.

[137] 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". In: *J. Comput. Phys.* 227.1 (2007), pp. 574–601.

[138] A Gamba, D Ambrosi, A. Coniglio, A De Candia, S Di Talia, E Giraudo, G Serini, L. Preziosi, and F. Bussolino. "Percolation, morphogenesis, and Burgers dynamics in blood vessels formation". In: *Phys. Rev. Lett.* 90.11 (2003), p. 118101.

[139] A. L. Garcia, M. M. Mansour, G. C. Lie, and E. Cementi. "Numerical integration of the fluctuating hydrodynamic equations". In: *J. Stat. Phys.* 47.1-2 (1987), pp. 209–228.

[140] H. Garcke, K. F. Lam, R. Nürnberg, and E. Sitka. "A multiphase Cahn–Hilliard–Darcy model for tumour growth with necrosis". In: *Mathematical Models and Methods in Applied Sciences* 28.03 (2018), pp. 525–577.

[141] J. Garnier, G. Papanicolaou, and T.-W. Yang. "Large deviations for a mean field model of systemic risk". In: *SIAM Journal on Financial Mathematics* 4.1 (2013), pp. 151–184.

[142] L. Gascón and J. Corberán. "Construction of second-order TVD schemes for nonhomogeneous hyperbolic conservation laws". In: *Journal of computational physics* 172.1 (2001), pp. 261–297.

[143]   I. Giardina. "Collective behavior in animal groups: theoretical models and empirical studies". In: *HFSP journal* 2.4 (2008), pp. 205–219.

[144]   J. Giesselmann, C. Lattanzio, and A. E. Tzavaras. "Relative energy for the Korteweg theory and related Hamiltonian flows in gas dynamics". In: *Arch. Ration. Mech. Anal.* 223.3 (2017), pp. 1427–1484.

[145]   V. Y. Glotov, V. M. Goloviznin, S. Karabasov, and A. Markesteijn. "New two-level leapfrog scheme for modeling the stochastic Landau-Lifshitz equations". In: *Comput. Math. & Math. Phys.* 54.2 (2014), pp. 315–334.

[146]   B. D. Goddard, A. Nold, and S. Kalliadasis. "Dynamical density functional theory with hydrodynamic interactions in confined geometries". In: *J. Chem. Phys.* 145.21 (2016), p. 214106.

[147]   B. D. Goddard, A Nold, N Savva, P Yatsyshin, and S Kalliadasis. "Unification of dynamic density functional theory for colloidal fluids to include inertia and hydrodynamic interactions: derivation and numerical experiments". In: *J. Phys. Condens. Matter* 25.3 (2012), p. 35101.

[148]   B. D. Goddard, A. Nold, N. Savva, G. A. Pavliotis, and S. Kalliadasis. "General dynamical density functional theory for classical fluids". In: *Phys. Rev. Lett.* 109.12 (2012), p. 120603.

[149]   B. D. Goddard, G. A. Pavliotis, and S. Kalliadasis. "The overdamped limit of dynamic density functional theory: Rigorous results". In: *Multiscale Model. Simul.* 10.2 (2012), pp. 633–663.

[150]   S. N. Gomes and G. A. Pavliotis. "Mean field limits for interacting diffusions in a two-scale potential". In: *J. nonlinear Sci.* 28.3 (2018), pp. 905–941.

[151]   H. Gómez, V. M. Calo, Y. Bazilevs, and T. J. Hughes. "Isogeometric analysis of the Cahn–Hilliard phase-field model". In: *Computer methods in applied mechanics and engineering* 197.49-50 (2008), pp. 4333–4352.

[152]   L Gosse. "A well-balanced flux-vector splitting scheme designed for hyperbolic systems of conservation laws with source terms". In: *Computers & Mathematics with Applications* 39.9-10 (2000), pp. 135–159.

[153]   L. Gosse. "Asymptotic-preserving and well-balanced schemes for the 1D Cattaneo model of chemotaxis movement in both hyperbolic and diffusive regimes". In: *J. Math. Anal. Appl.* 388.2 (2012), pp. 964–983.

[154]   L. Gosse. *Computing qualitatively correct approximations of balance laws*. Vol. 2. Springer, 2013.

[155]   L. Gosse and A.-Y. Leroux. "A well-balanced scheme designed for inhomogeneous scalar conservation laws". In: *Comptes Rendus L Acad. Des Sci. Ser. I-mathematique* 323.5 (1996), pp. 543–546.

[156]   S. Gottlieb and C.-W. Shu. "Total variation diminishing Runge-Kutta schemes". In: *Math. Comput. Am. Math. Soc.* 67.221 (1998), pp. 73–85.

[157]   J. M. Greenberg and A.-Y. LeRoux. "A well-balanced scheme for the numerical processing of source terms in hyperbolic equations". In: *SIAM J. Numer. Anal.* 33.1 (1996), pp. 1–16.

[158]   J. Gu, L. Zhang, G. Yu, Y. Xing, and Z. Chen. "X-ray CT metal artifacts reduction through curvature based sinogram inpainting". In: *Journal of X-ray Science and Technology* 14.2 (2006), pp. 73–82.

[159] J. Guo, C. Wang, S. M. Wise, and X. Yue. "An H2 convergence of a second-order convex-splitting, finite difference scheme for the three-dimensional Cahn-Hilliard equation". In: *Commun. Math. Sci* 14.2 (2016), pp. 489–515.

[160] Y. Guo and B. Pausader. "Global smooth ion dynamics in the Euler-Poisson system". In: *Communications in Mathematical Physics* 303.1 (2011), pp. 89–125.

[161] S.-Y. Ha and E. Tadmor. "From particle to kinetic and hydrodynamic descriptions of flocking". In: *Kinetic & Related Models* 1.1937-5093-2008-3-415 (2008), p. 415. ISSN: 1937-5093.

[162] M. Hadžić and J. J. Jang. "A class of global solutions to the Euler–Poisson system". In: *Communications in Mathematical Physics* 370.2 (2019), pp. 475–505.

[163] K. A. Hawick and D. P. Playne. "Modelling and Visualizing the Cahn-Hilliard-Cook Equation." In: *MSV*. 2008, pp. 149–155.

[164] Y. He, Y. Liu, and T. Tang. "On large time-stepping methods for the Cahn–Hilliard equation". In: *Applied Numerical Mathematics* 57.5-7 (2007), pp. 616–628.

[165] H. Hildenbrandt, C. Carere, and C. K. Hemelrijk. "Self-organized aerial displays of thousands of starlings: a model". In: *Behavioral Ecology* 21.6 (2010), pp. 1349–1359.

[166] M. Hintermüller, M. Hinze, and M. H. Tber. "An adaptive finite-element Moreau–Yosida-based solver for a non-smooth Cahn–Hilliard problem". In: *Optimization Methods and Software* 26.4-5 (2011), pp. 777–811.

[167] F. Hoffmann. "Keller-Segel-Type Models and Kinetic Equations for Interacting Particles: Long-Time Asymptotic Analysis". PhD thesis. University of Cambridge, 2017.

[168] D. D. Holm and V. Putkaradze. "Formation of clumps and patches in self-aggregation of finite-size particles". In: *Physica D: Nonlinear Phenomena* 220.2 (2006), pp. 183–196.

[169] D. Horstmann and Others. "From 1970 until present: the Keller-Segel model in chemotaxis and its consequences". In: (2003).

[170] R. Jordan, D. Kinderlehrer, and F. Otto. "The variational formulation of the Fokker–Planck equation". In: *SIAM journal on mathematical analysis* 29.1 (1998), pp. 1–17.

[171] S. Kalliadasis and U. Thiele. *Thin films of soft matter*. Vol. 1. Springer, 2007.

[172] T. Katsaounis, B. Perthame, and C. Simeoni. "Upwinding sources at interfaces in conservation laws". In: *Appl. Math. Lett.* 17.3 (2004), 309–316.

[173] T. Katsaounis and C. Simeoni. "First and second order error estimates for the upwind source at interface method". In: *Mathematics of computation* 74.249 (2005), 103–122.

[174] T. Katsaounis and C. Simeoni. "Second order approximation of the viscous Saint-Venant system and comparison with experiments". In: *Hyperbolic Problems: Theory, Numerics, Applications*. Springer, 2003, pp. 633–644.

[175] Y. Katz, K. Tunstrøm, C. C. Ioannou, C. Huepe, and I. D. Couzin. "Inferring the structure and dynamics of interactions in schooling fish". In: *Proceedings of the National Academy of Sciences* 108.46 (2011), pp. 18720–18725.

[176] K. Kawasaki. "Stochastic model of slow dynamics in supercooled liquids and dense colloidal suspensions". In: *Physica A* 208.1 (1994), pp. 35 –64.

[177] E. F. Keller and L. A. Segel. "Initiation of slime mold aggregation viewed as an instability". In: *J. Theor. Biol.* 26.3 (1970), pp. 399–415.

[178] C. Kim, A. Nonaka, J. B. Bell, A. L. Garcia, and A. Donev. "Stochastic simulation of reaction-diffusion systems: A fluctuating-hydrodynamics approach". In: *J. Chem. Phys.* 146.12 (2017), p. 124110.

[179] D. Kim and W. Lu. "Three-dimensional model of electrostatically induced pattern formation in thin polymer films". In: *Physical Review B* 73.3 (2006), p. 035206.

[180] J. Kim, S. Lee, Y. Choi, S.-M. Lee, and D. Jeong. "Basic principles and practical applications of the Cahn–Hilliard equation". In: *Mathematical Problems in Engineering* 2016 (2016).

[181] C. Klingenberg, G. Puppo, and M. Semplice. "Arbitrary order finite volume well-balanced schemes for the Euler equations with gravity". In: *SIAM Journal on Scientific Computing* 41.2 (2019), A695–A721.

[182] P. E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations.* Springer-Verlag Berlin Heidelberg, 1992.

[183] O. Kolb. "On the full and global accuracy of a compact third order WENO scheme". In: *SIAM Journal on Numerical Analysis* 52.5 (2014), pp. 2335–2355.

[184] T. Kolokolnikov, J. A. Carrillo, A. Bertozzi, R. Fetecau, and M. Lewis. "Emergent behaviour in multi-particle systems with non-local interactions". In: *Physica D: Nonlinear Phenomena* 260 (2013), pp. 1–4.

[185] M. Kruger and D. S. Dean. "A Gaussian theory for fluctuations in simple liquids". In: *J. Chem. Phys.* 146.13 (2017), p. 134507.

[186] A. Kurganov, G. Petrova, and Others. "A second-order well-balanced positivity preserving central-upwind scheme for the Saint-Venant system". In: *Commun. Math. Sci.* 5.1 (2007), pp. 133–160.

[187] L. Landau and E. Lifshitz. "Statistical physics, vol. 5". In: *Course of theoretical physics* (1980).

[188] C. B. Laney. *Computational gasdynamics.* Cambridge university press, 1998.

[189] A. Lasota and M. C. Mackey. *Chaos, fractals, and noise: stochastic aspects of dynamics.* Vol. 97. Springer Science & Business Media, 2013.

[190] C. Lattanzio and A. E. Tzavaras. "From gas dynamics with large friction to gradient flows describing diffusion theories". In: *Communications in Partial Differential Equations* 42.2 (2017), pp. 261–290.

[191] D. Lee, J.-Y. Huh, D. Jeong, J. Shin, A. Yun, and J. Kim. "Physical, mathematical, and numerical derivations of the Cahn–Hilliard equation". In: *Computational Materials Science* 81 (2014), pp. 216–225.

[192] H. G. Lee and J. Kim. "Accurate contact angle boundary conditions for the Cahn–Hilliard equations". In: *Computers & fluids* 44.1 (2011), pp. 178–186.

[193] J. Lee, D.-K. Lee, and R.-H. Park. "Robust exemplar-based inpainting algorithm using region segmentation". In: *IEEE Transactions on Consumer Electronics* 58.2 (2012), pp. 553–561.

[194] A. Y. Leroux. "Riemann solvers for some hyperbolic problems with a source term". In: *ESAIM Proc.* Vol. 6. EDP Sciences. 1999, pp. 75–90.

[195] R. J. LeVeque. "Balancing source terms and flux gradients in high-resolution Godunov methods: the quasi-steady wave-propagation algorithm". In: *J. Comput. Phys.* 146.1 (1998), pp. 346–365.

[196] R. J. LeVeque. *Finite volume methods for hyperbolic problems.* Vol. 31. Cambridge university press, 2002.

[197] D. Levy, G. Puppo, and G. Russo. "Central WENO schemes for hyperbolic systems of conservation laws". In: *ESAIM: Mathematical Modelling and Numerical Analysis* 33.3 (1999), pp. 547–571.

[198] D. Levy, G. Puppo, and G. Russo. "Compact central WENO schemes for multidimensional conservation laws". In: *SIAM Journal on Scientific Computing* 22.2 (2000), pp. 656–672.

[199] Q. Liang and F. Marche. "Numerical resolution of well-balanced shallow water equations with complex source terms". In: *Adv. Water Resour.* 32.6 (2009), pp. 873–884.

[200] J. F. Lutsko. "How crystals form: A theory of nucleation pathways". In: *Sci. Adv.* 5 (2019), eaav7399.

[201] J. F. Lutsko. "A dynamical theory of nucleation for colloids and macromolecules". In: *J. Chem. Phys.* 136.3 (2012), p. 034509.

[202] J. F. Lutsko. "Recent developments in classical density functional theory". In: *Adv. Chem. Phys.* 144 (2010), p. 1.

[203] J. F. Lutsko and M. A. Durán-Olivencia. "A two-parameter extension of classical nucleation theory". In: *J. Phys. Condens. Matter* 27.23 (2015), p. 235101.

[204] J. F. Lutsko and M. A. Durán-Olivencia. "Classical nucleation theory from a dynamical approach to nucleation". In: *J. Chem. Phys.* 138.24 (2013), p. 244908.

[205] Z. Mao, Z. Li, and G. E. Karniadakis. "Nonlocal flocking dynamics: Learning the fractional order of pdes from particle simulations". In: *Communications on Applied Mathematics and Computation* 1.4 (2019), pp. 597–619.

[206] F Marche, P. Bonneton, P Fabrie, and N. Seguin. "Evaluation of well-balanced bore-capturing schemes for 2D wetting and drying processes". In: *Int. J. Numer. Methods Fluids* 53.5 (2007), pp. 867–894.

[207] U. M. B. Marconi and P. Tarazona. "Dynamic density functional theory of fluids". In: *J. Chem. Phys.* 110.16 (1999), pp. 8032–8044.

[208] K. T. Mashiyama and H. Mori. "Origin of the Landau-Lifshitz hydrodynamic fluctuations in nonequilibrium systems and a new method for reducing the Boltzmann equation". In: *J. Stat. Phys.* 18.4 (1978), pp. 385–407. ISSN: 1572-9613.

[209] S. Masnou and J.-M. Morel. "Level lines based disocclusion". In: *Proceedings 1998 International Conference on Image Processing. ICIP98 (Cat. No. 98CB36269)*. IEEE. 1998, pp. 259–263.

[210] J. Mendes, A. Russo, S. P. Perez, and S. Kalliadasis. "A finite-volume scheme for gradient-flow equations with non-homogeneous diffusion". In: *Computers & Mathematics with Applications* 89 (2021), pp. 150–162.

[211] P. Minakowski, P. B. Mucha, J. Peszek, and E. Zatorska. "Singular Cucker–Smale Dynamics". In: *Active Particles, Volume 2*. Springer, 2019, pp. 201–243.

[212] A. Miranville and S. Zelik. "Robust exponential attractors for Cahn-Hilliard type equations with singular potentials". In: *Mathematical methods in the applied sciences* 27.5 (2004), pp. 545–582.

[213] K. Mohamed, M. Seaid, and M. Zahri. "A finite volume method for scalar conservation laws with stochastic time–space dependent flux functions". In: *J. Comput. Appl. Math.* 237.1 (2013), pp. 614–632.

[214] S. Motsch and E. Tadmor. "A new model for self-organized dynamics and its flocking behavior". In: *Journal of Statistical Physics* 144.5 (2011), p. 923.

[215] R. Natalini and M. Ribot. "Asymptotic high order mass-preserving schemes for a hyperbolic model of chemotaxis". In: *SIAM Journal on Numerical Analysis* 50.2 (2012), pp. 883–905.

[216] R. Natalini, M. Ribot, and M. Twarogowska. "A well-balanced numerical scheme for a one dimensional quasilinear hyperbolic model of chemotaxis". In: *arXiv preprint arXiv:1211.4010* (2012).

[217] S. Noelle, N. Pankratz, G. Puppo, and J. R. Natvig. "Well-balanced finite volume schemes of arbitrary order of accuracy for shallow water flows". In: *J. Comput. Phys.* 213.2 (2006), pp. 474–499.

[218] S. Noelle, Y. Xing, and C.-W. Shu. "High-order well-balanced finite volume WENO schemes for shallow water equation with moving water". In: *J. Comput. Phys.* 226.1 (2007), pp. 29–58.

[219] A. Nold, B. D. Goddard, P. Yatsyshin, N. Savva, and S. Kalliadasis. "Pseudospectral methods for density functional theory in bounded and unbounded domains". In: *Journal of Computational Physics* 334 (2017), pp. 639–664.

[220] A. Nold, D. N. Sibley, B. D. Goddard, and S. Kalliadasis. "Fluid structure in the immediate vicinity of an equilibrium three-phase contact line and assessment of disjoining pressure models using density functional theory". In: *Phys. Fluids* 26.7 (2014), p. 072001.

[221] *Official website of TensorFlow.* https://www.tensorflow.org/. Accessed: 2020-03-05.

[222] B. Øksendal. "Stochastic differential equations". In: *Stochastic differential equations.* Springer, 2003, pp. 65–84.

[223] S. Osher. "Convergence of generalized MUSCL schemes". In: *SIAM J. Numer. Anal.* 22.5 (1985), pp. 947–961.

[224] F. Otto. *Double degenerate diffusion equations as steepest descent.* Sonderforschungsbereich 256, 1996.

[225] F. Otto. "The geometry of dissipative evolution equations: the porous medium equation". In: (2001).

[226] L. Pareschi and M. Zanella. "Structure preserving schemes for nonlinear Fokker–Planck equations and applications". In: *J. Sci. Comput.* 74.3 (2018), pp. 1575–1600.

[227] A. O. Parry, C Rascón, and R Evans. "The local structure factor near an interface; beyond extended capillary-wave models". In: *J. Phys. Condens. Matter* 28.24 (2016), p. 244013.

[228] M. Paulus, C. Gutt, and M. Tolan. "Static structure factor of capillary waves at large momentum transfer". In: *Phys. Rev. B* 78 (23 2008), p. 235419.

[229] M. A. Peletier. *Energies, gradient flows, and large deviations: a modelling point of view.* 2012.

[230] J. K. Percus. "Equilibrium state of a classical fluid of hard rods in an external field". In: *J. Stat. Phys.* 15.6 (1976), pp. 505–511.

[231] L. Perea, G. Gómez, and P. Elosegui. "Extension of the Cucker-Smale control law to space flight formations". In: *Journal of guidance, control, and dynamics* 32.2 (2009), pp. 527–537.

[232] S. P. Perez. *Videos of the simulations from the work "Well-balanced finite volume schemes for hydrodynamic equations with general free energy".* https://figshare.com/projects/Well-balanced_finite_volume_schemes_for_hydrodynamic_equations_with_general_free_energy/60122. 2018.

[233] S. P. Perez. *Code to reproduce the one- and two-dimensional Cahn-Hilliard simulations.* https://github.com/sergiopperez/Cahn_Hilliard_Finite_Volume. 2020.

[234] S. P. Perez. *Films of the two-dimensional simulations.* https://figshare.com/projects/Unconditional_bound-preserving_and_energy-dissipating_finite-volume_schemes_for_the_Cahn-Hilliard_equation/101126. 2021.

[235] B. Perthame. *Kinetic formulation of conservation laws*. Vol. 21. Oxford University Press, 2002.

[236] B. Perthame and C. Simeoni. "A kinetic scheme for the Saint-Venant system with a source term". In: *Calcolo* 38.4 (2001), pp. 201–231.

[237] S. Plimpton. "Fast parallel algorithms for short-range molecular dynamics". In: *J. Comput. Phys.* 117.1 (1995), pp. 1–19.

[238] M. Queralt-Martín, M Pradas, R. Rodríguez-Trujillo, M Arundell, E. C. Poiré, and A. Hernández-Machado. "Pinning and avalanches in hydrophobic microchannels". In: *Physical review letters* 106.19 (2011), p. 194501.

[239] M. Raissi, P. Perdikaris, and G. E. Karniadakis. "Physics-informed neural networks: A deep learning framework for solving forward and inverse problems involving nonlinear partial differential equations". In: *Journal of Computational Physics* 378 (2019), pp. 686–707.

[240] C. Reina and J. Zimmer. "Entropy production and the geometry of dissipative evolution equations". In: *Phys. Rev. E* 92 (2015), p. 052117.

[241] Y. Rosenfeld. "Free-energy model for the inhomogeneous hard-sphere fluid mixture and density-functional theory of freezing". In: *Phys. Rev. Lett.* 63.9 (1989), p. 980.

[242] A. Russo, M. A. Durán-Olivencia, S. Kalliadasis, and R. Hartkamp. "Macroscopic relations for microscopic properties at the interface between solid substrates and dense fluids". In: *The Journal of chemical physics* 150.21 (2019), p. 214705.

[243] A. Russo, S. P. Perez, M. A. Durán-Olivencia, P. Yatsyshin, J. A. Carrillo, and S. Kalliadasis. "A finite-volume method for fluctuating dynamical density functional theory". In: *Journal of Computational Physics* (2020), p. 109796.

[244] F. Santambrogio. "Euclidean, metric, and Wasserstein gradient flows: an overview". In: *Bulletin of Mathematical Sciences* 7.1 (2017), pp. 87–154.

[245] C.-B. Schönlieb. *Partial differential equation methods for image inpainting*. Vol. 29. Cambridge University Press, 2015.

[246] M. Semplice, A. Coco, and G. Russo. "Adaptive mesh refinement for hyperbolic systems based on third-order compact WENO reconstruction". In: *Journal of Scientific Computing* 66.2 (2016), pp. 692–724.

[247] P. Seppecher. "Moving contact lines in the Cahn-Hilliard theory". In: *International journal of engineering science* 34.9 (1996), pp. 977–992.

[248] G. Serini, D. Ambrosi, E. Giraudo, A. Gamba, L. Preziosi, and F. Bussolino. "Modeling the early stages of vascular network assembly". In: *The EMBO journal* 22.8 (2003), pp. 1771–1779.

[249] J. Shen, J. Xu, and J. Yang. "A new class of efficient and robust energy stable schemes for gradient flows". In: *SIAM Review* 61.3 (2019), pp. 474–506.

[250] J. Shen, J. Xu, and J. Yang. "The scalar auxiliary variable (SAV) approach for gradient flows". In: *Journal of Computational Physics* 353 (2018), pp. 407–416.

[251] J. Shen and X. Yang. "Numerical approximations of allen-cahn and cahn-hilliard equations". In: *Discrete & Continuous Dynamical Systems-A* 28.4 (2010), p. 1669.

[252] M. B. Short, M. R. D'orsogna, V. B. Pasour, G. E. Tita, P. J. Brantingham, A. L. Bertozzi, and L. B. Chayes. "A statistical model of criminal behavior". In: *Mathematical Models and Methods in Applied Sciences* 18.supp01 (2008), pp. 1249–1267.

[253] C.-W. Shu. *Essentially Non-Oscillatory and Weighted Essentially Non-Oscillatory Schemes for Hyperbolic Conservation Laws*. Tech. rep. 1997.

[254] C.-W. Shu. "Essentially non-oscillatory and weighted essentially non-oscillatory schemes for hyperbolic conservation laws". In: *Adv. Numer. Approx. nonlinear hyperbolic equations.* Springer, 1998, pp. 325–432.

[255] C.-W. Shu and S. Adrminimration. "A numerical method for systems of conservation laws of mixed type admitting hyperbolic flux splitting". In: *J. Comput. Phys.* Citeseer. 1992.

[256] C.-W. Shu and S. Osher. "Efficient implementation of essentially non-oscillatory shock-capturing schemes". In: *Journal of Computational Physics* 77.2 (1988), pp. 439–471.

[257] C.-W. Shu and S. Osher. "Efficient Implementation of Essentially Non-oscillatory Shock-capturing Schemes, II". In: *Journal of Computational Physics* 83.1 (1989), pp. 32–78.

[258] D. N. Sibley, A. Nold, and S. Kalliadasis. "Unifying binary fluid diffuse-interface models in the sharp-interface limit". In: *Journal of fluid mechanics* 736 (2013), p. 5.

[259] D. N. Sibley, A. Nold, N. Savva, and S. Kalliadasis. "On the moving contact line singularity: Asymptotics of a diffuse-interface model". In: *The European Physical Journal E* 36.3 (2013), p. 26.

[260] D. N. Sibley, A. Nold, N. Savva, and S. Kalliadasis. "The contact line behaviour of solid-liquid-gas diffuse-interface models". In: *Physics of Fluids* 25.9 (2013), p. 092111.

[261] C. Sire and P.-H. Chavanis. "Thermodynamics and collapse of self-gravitating Brownian particles in D dimensions". In: *Phys. Rev. E* 66.4 (2002), p. 46133.

[262] V. Sotiropoulos and Y. N. Kaznessis. "An adaptive time step scheme for a system of stochastic differential equations with multiple multiplicative noise: chemical Langevin equation, a proof of concept". In: *J. Chem. Phys.* 128.1 (2008), p. 014103.

[263] A. F. Stalder, G. Kulik, D. Sage, L Barbieri, and P. Hoffmann. "A snake-based approach to accurate determination of both contact points and contact angles". In: *Colloids and surfaces A: physicochemical and engineering aspects* 286.1-3 (2006), pp. 92–103.

[264] Z. Sun, J. A. Carrillo, and C.-W. Shu. "A discontinuous Galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials". In: *J. Comput. Phys.* 352 (2018), pp. 76–104.

[265] P. K. Sweby. "High resolution schemes using flux limiters for hyperbolic conservation laws". In: *SIAM J. Numer. Anal.* 21.5 (1984), pp. 995–1011.

[266] E. Tadmor. "Entropy stable schemes". In: *Handbook of Numerical Analysis.* Vol. 17. Elsevier, 2016, pp. 467–493.

[267] A. Thomann, M. Zenk, and C. Klingenberg. "A second-order positivity-preserving well-balanced finite volume scheme for Euler equations with gravity for arbitrary hydrostatic equilibria". In: *International Journal for Numerical Methods in Fluids* 89.11 (2019), pp. 465–482.

[268] G Tierra and F Guillén-González. "Numerical methods for solving the Cahn–Hilliard equation and its applicability to related energy-based models". In: *Archives of Computational Methods in Engineering* 22.2 (2015), pp. 269–289.

[269] E. F. Toro. *Riemann solvers and numerical methods for fluid dynamics: a practical introduction.* Springer Science & Business Media, 2013.

[270] J. A. de la Torre, P. Español, and A. Donev. "Finite element discretization of non-linear diffusion equations with thermal fluctuations". In: *J. Chem. Phys.* 142.9 (2015), p. 094115.

[271] J. Tugaut. "Phase transitions of McKean–Vlasov processes in double-wells landscape". In: *Stochastics An Int. J. Probab. Stoch. Process.* 86.2 (2014), pp. 257–284.

[272] B. Van Leer. "Towards the ultimate conservative difference scheme. V. A second-order sequel to Godunov's method". In: *Journal of computational Physics* 32.1 (1979), pp. 101–136.

[273] C. Villani. *Topics in optimal transportation.* 58. American Mathematical Soc., 2003.

[274] B. P. Vollmayr-Lee and A. D. Rutenberg. "Fast and accurate coarsening simulation with an unconditionally stable time step". In: *Physical Review E* 68.6 (2003), p. 066703.

[275] N. K. Voulgarakis and J.-W. Chu. "Bridging fluctuating hydrodynamics and molecular dynamics simulations of fluids". In: *J. Chem. Phys.* 130.13 (2009), 04B605.

[276] S. Vukovic, N. Crnjaric-Zic, and L. Sopta. "WENO schemes for balance laws with spatially varying flux". In: *J. Comput. Phys.* 199.1 (2004), pp. 87–109.

[277] S. A. Williams, J. B. Bell, and A. L. Garcia. "Algorithm refinement for fluctuating hydrodynamics". In: *Multiscale Model. Sim.* 6.4 (2008), pp. 1256–1280.

[278] X Wu, G. Van Zwieten, and K. Van der Zee. "Stabilized second-order convex splitting schemes for Cahn–Hilliard models with application to diffuse-interface tumor-growth models". In: *International journal for numerical methods in biomedical engineering* 30.2 (2014), pp. 180–203.

[279] C. Wylock, M. Pradas, B. Haut, P. Colinet, and S. Kalliadasis. "Disorder-induced hysteresis and nonlocality of contact line motion in chemically heterogeneous microchannels". In: *Physics of fluids* 24.3 (2012), p. 032108.

[280] Y. Xia, Y. Xu, and C.-W. Shu. "Local discontinuous Galerkin methods for the Cahn–Hilliard type equations". In: *Journal of Computational Physics* 227.1 (2007), pp. 472–491.

[281] J. Xie, L. Xu, and E. Chen. "Image denoising and inpainting with deep neural networks". In: *Advances in neural information processing systems.* 2012, pp. 341–349.

[282] Y. Xing and C.-W. Shu. "A survey of high order schemes for the shallow water equations". In: *J. Math. Study* 47.3 (2014), pp. 221–249.

[283] Y. Xing and C.-W. Shu. "High order finite difference WENO schemes with the exact conservation property for the shallow water equations". In: *J. Comput. Phys.* 208.1 (2005), pp. 206–227.

[284] Y. Xing and C.-W. Shu. "High-order well-balanced finite difference WENO schemes for a class of hyperbolic systems with source terms". In: *J. Sci. Comput.* 27.1-3 (2006), pp. 477–494.

[285] 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". In: *J. Comput. Phys.* 214.2 (2006), pp. 567–598.

[286] Y. Xing and C.-W. Shu. "High order well-balanced WENO scheme for the gas dynamics equations under gravitational fields". In: *Journal of Scientific Computing* 54.2-3 (2013), pp. 645–662.

[287] Y. Xing, C.-W. Shu, and S. Noelle. "On the advantage of well-balanced schemes for moving-water equilibria of the shallow water equations". In: *J. Sci. Comput.* 48.1-3 (2011), pp. 339–349.

[288] J. Xu, Y. Li, S. Wu, and A. Bousquet. "On the stability and accuracy of partially and fully implicit schemes for phase field modeling". In: *Computer Methods in Applied Mechanics and Engineering* 345 (2019), pp. 826–853.

[289] K. Xu. "A well-balanced gas-kinetic scheme for the shallow-water equations with source terms". In: *J. Comput. Phys.* 178.2 (2002), pp. 533–562.

[290] P. Yatsyshin, A. O. Parry, C. Rascón, and S. Kalliadasis. "Wetting of a plane with a narrow solvophobic stripe". In: *Mol. Phys.* 116.15-16 (2018), pp. 1990–1997.

[291] P Yatsyshin, N Savva, and S Kalliadasis. "Density functional study of condensation in capped capillaries". In: *J. Phys. Condens. Matter* 27.27 (2015), p. 275104.

[292] P. Yatsyshin, S. Kalliadasis, and A. B. Duncan. "Data Driven Density Functional Theory: A case for Physics Informed Learning". In: *arXiv preprint arXiv:2010.03374* (2020).

[293] P. Yatsyshin, N. Savva, and S. Kalliadasis. "Geometry-induced phase transition in fluids: Capillary prewetting". In: *Physical Review E* 87.2 (2013), p. 020402.

[294] P. Yatsyshin, N. Savva, and S. Kalliadasis. "Spectral methods for the equations of classical density-functional theory: Relaxation dynamics of microscopic films". In: *The Journal of chemical physics* 136.12 (2012), p. 124113.

[295] P. Yatsyshin, N. Savva, and S. Kalliadasis. "Wetting of prototypical one- and two-dimensional systems: Thermodynamics and density functional theory". In: *J. Chem. Phys.* 142.3 (2015), p. 034708.

[296] R. A. Yeh, C. Chen, T. Yian Lim, A. G. Schwing, M. Hasegawa-Johnson, and M. N. Do. "Semantic image inpainting with deep generative models". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2017, pp. 5485–5493.

[297] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang. "Generative image inpainting with contextual attention". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2018, pp. 5505–5514.

[298] J. Zhang, C. Chen, and X. Yang. "Efficient and energy stable method for the Cahn-Hilliard phase-field model for diblock copolymers". In: *Applied Numerical Mathematics* 151 (2020), pp. 263–281.

[299] X. Zhang and C.-W. Shu. "On maximum-principle-satisfying high order schemes for scalar conservation laws". In: *Journal of Computational Physics* 229.9 (2010), pp. 3091–3120.

[300] S. Zhou and M. Y. Wang. "Multimaterial structural topology optimization with a generalized Cahn–Hilliard model of multiphase transition". In: *Structural and Multidisciplinary Optimization* 33.2 (2007), p. 89.