Imperial College London

Analysis, modelling and prediction of deterministic and stochastic complex systems

Submitted in part fulfilment of the
requirements for the degree of
Doctor of Philosophy of Imperial College London
and the
Diploma of Imperial College London
by

Antonio Russo

Supervisor: Prof. Serafim Kalliadasis
Co-supervisor: Dr. Miguel A. Durán-Olivencia

Department of Chemical engineering
Imperial College London

01/10/2021

Declaration of Originality

I certify that this thesis, 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.

Antonio Russo
01/10/2021

Copyright Declaration

Acknowledgements

I would like to express my sincere gratitude to Imperial College London for giving me the opportunity to live this amazing experience which helped me grow professionally and personally. I would like to particularly thank my supervisor Prof. Serafim Kalliadasis for his dedicated support and guidance throughout my research projects. A special thanks goes to my co-supervisor Dr. Miguel Duran-Olivencia for his enthusiasm, his consistent support and the tireless encouragement. I would like to thank my examiners, Professors Amparo Galindo and Panos Stinis for the engaging discussion and constructive comments. I am extremely grateful to Sergio, Peter and all the other members of the complex multiscale systems research group for all the great collaborations, the interesting discussions and the helpful advice. I would like to thank all the collaborators, including Pietro Asinari, Remco Hartkamp, David Sibley, Yannis G. Kevrekidis and José Carrillo. I am extremely grateful to Virginia, who accompanied me in this journey and encouraged me in the most difficult moments. Finally, a huge thanks goes to my parents, my brother and all my family for the unconditional and tireless support throughout my academic studies.

Abstract

The analysis of complex systems at nano- and micro-scales often requires their numerical simulation. Atomistic simulations, that rely on solving Newton's equation for each component of the system, despite being exact, are often too computationally expensive. In this work, firstly we analyse the properties of confined systems by extracting mesoscopic information directly from particles coordinate.

Then, taking advantage of Mori-Zwanzig projector operator techniques and advanced data-analysis tools, we present a novel approach to parametrize non-Markovian coarse-graining models of molecular system. We focus on the parametrization of the memory terms in the stochastic Generalized Langevin Equation through a deep-learning approach.

Moreover, in the framework of Dynamical Density Functional Theory (DDFT) we derive a continuum non-Markovian formulation, able to describe, given the proper free-energy, the physical properties of an atomistic system. Comparisons between molecular dynamics, fluctuating dynamical density functional theory and fluctuating hydrodynamics simulations validate our approach.

Finally, we propose some numerical schemes for the simulation of DDFT with additional complexities, i.e. with stochastic terms and non-homogeneous non-constant diffusion.

Publications, presentations, scholarships and extra-curricular activities

Publications

Contributions of this work to the literature include:

- A. Russo, M. A. Durán-Olivencia, S. Kalliadasis, & R. Hartkamp. Macroscopic relations for microscopic properties at the interface between solid substrates and dense fluids, *J. Chem. Phys.*, 150, 214705 (2019). It includes materials covered in Section 2.2 of this thesis.

- A. Russo, M. A. Durán-Olivencia, P. Yatsyshin & S. Kalliadasis. Non-Markovian dynamical density functional theory for reacting multi-component systems: theoretical framework and finite volume simulations, *J. Phys. A: Math. Theor.*, 53, 445007 (2020). It includes materials covered in Section 3.2 of this thesis.

- A. Russo, S. P. Perez, J. Carrillo, M. A. Durán-Olivencia, P. Yatsyshin & S. Kalliadasis. A finite-volume method for fuctuating dynamical density functional theory, *J. Comp. Phys.*, 428, 109796 (2020). It includes materials covered in Section 4.1 of this thesis.

- J. Mendes, A. Russo, S. P. Perez & S. Kalliadasis. A finite-volume scheme for gradient flow equations with non-homogeneous diffusion, *Comput. Math. Appl.*, 89, 150–162 (2020). It includes materials covered in Section 4.2 of this thesis.

- A. Russo, M. A. Durán-Olivencia, I. G. Kevrekidis & S. Kalliadasis. Machine learning memory kernels as closure for non-Markovian stochastic processes, *arXiv*:1903.09562 (under-review). It includes materials covered in Section 3.1 of this thesis.

- A. Russo, D. Sibley, P. Asinari & S. Kalliadasis. Liquid-vapour and immiscible fluids interfaces: a non-equilibrium molecular dynamic study (in preparation). It includes materials covered in Section 2.3 of this thesis.

Other contributions to the literature not included in this work, include:

- A. Russo, M. Icardi, M. Elsharkawy, D. Ceglia, P. Asinari & C. M. Megaridis. Numerical simulation of droplet impact on wettability-patterned surfaces, *Phys. Rev. Fluids* (2020).

Presentations

The contents of this work have been presented at the following conferences:

- Mini-symposia "Transport properties at fluids interfaces: macroscopic relations for microscopic phenomena" at the "British applied mathematics colloquium 2017", University of Surrey (10th-12th April 2017);

- Talk "Transport properties at fluids interfaces: a molecular study for a macroscopic modelling" at the "70th Annual Meeting of the APS Division of Fluid Dynamics", Denver, Colorado (19th-21st November 2017);

- Mini-symposia "Computational challenges in fluctuating hydrodynamics" at the "British applied mathematics colloquium 2018", University of St. Andrews (26th-29th March 2018);

- Mini-symposia "A data-driven approach to model reduction: Deep-learning memory effects" at the "British applied mathematics colloquium 2019", University of Bath (24th-26th April 2019);

- Mini-symposia "A finite volume scheme for stochastic PDEs" at the "British applied mathematics colloquium 2019", University of Bath (24th-26th April 2019);

- Talk "Memory effects in dynamic density functional theory with fluctuations" at the "8th DFT days 2020", University of Tuebingen (21st-23rd September 2020);

Scholarships and awards

This work was supported by the following scholarship and awards:

- PhD scholarship of £50k+ for the period Nov 2016-Sep 2020 by the Department of Chemical Engineering at Imperial College London;
- Student Scholarship Award of £100 at the "British applied mathematics colloquium 2017" (University of Surrey);
- Student Scholarship Award of £200 at "British applied mathematics colloquium 2018" (University of St. Andrews).

Graduate school courses

During my PhD, I attended and passed the following modules offered by the Graduate School of Imperial College London:

- Entrepreneurship 3: Idea Generation: Why Didn't I think of That? (2016-2017)
- Writing for Success 4: Publication (2016-2017)
- Careers: Job Search With A Difference (2017-2018)
- Entrepreneurship 5: Generating a Business Model and Plan (2017-2018)
- Entrepreneurship 7b: Strategically Marketing Your Research (2017-2018)
- Global Postgraduate Retreats: Thesis Writing (2018-2019)
- Industrial xChange Retreat: Materials and Technology for the Future (2018-2019)
- Webinar: Teaching: Applying for Associate Fellowship AFHEA (2018-2019)

Extra-curricular activities

During the development of this work, the following academic and non-academic roles were covered during the PhD:

- Graduate teaching assistant in Mathematics 1 during the academic years 2017/2018 and 2018/2019.
- Co-supervisor of undergraduate and master (UROP) projects.

- Accreditation as associate fellow of the higher education academy (fellowship reference number: PR166992).
- Internship in advanced analytics at eBay with responsibilities in business and product analytics, data engineering and modelling, data science and market insights.

Contents

List of Figures

List of Tables

Nomenclature

List of Abbreviations

DDFT	dynamical density functional theory
DFT	density functional theory
FDDFT	fluctuating dynamical density functional theory
FH	fluctuating hydrodynamics
GLE	generalized Langevin equation
HS	hard-spheres
LE	Langevin equation
LJ	Lennard-Jones
MD	molecular dynamics
NVT	canonical ensemble

List of Lower-Case Greek

$\boldsymbol{\theta}$	memory kernel function in GLE
$\boldsymbol{\theta}_0$	friction factor in LE
δ	Dirac delta function
$\dot{\gamma}$	shear rate

ϵ	reduced energy
η^k	kinetic viscosity
η^v	potential viscosity
σ^k	kinetic stress tensor
σ^v	potential stress tensor
ϕ	space kernel function
ρ	density
σ	reduced distance

List of Lower-Case Roman

$\mathbf{p}$	particles momenta vector
$\mathbf{r}$	particles position vector
$\mathbf{u}$	particles velocity vector
k_B	Boltzmann constant
m	particle mass
n	number density

List of Upper-Case Roman

$\mathcal{E}$	functional
$\mathcal{F}$	Helmoltz free-energy functional
$\mathcal{F}^{ex}$	excess free-energy functional
$\mathcal{F}^{id}$	Ideal gas free-energy functional
$\mathcal{H}$	Hamiltonian
$\mathcal{P}$	projection operator
$\mathcal{Q}$	inverse projection operator
Wa	Wall number

E	total energy
K	kinetic energy
N	number of particles in the system
T	temperature
U	potential energy

List of mathematical symbols

$\frac{D}{Dt} = \frac{\partial}{\partial t} + \mathbf{u} \cdot \nabla$	material derivative
$\langle \cdots \rangle$	ensemble average
$\mathbf{A} * \mathbf{f} = \int \mathbf{A}(t-\tau)\mathbf{f}(\tau)d\tau$	convolution
$\mathbf{a} \otimes \mathbf{b} = \mathbf{a}\mathbf{b}^T$	outer product
$\nabla \cdot$	divergence operator
∇	gradient operator
∂_t	partial derivative with respect to t
$\mathrm{Tr}(\mathbf{A}) = \sum_i A_{ii}$	trace of a matrix

Chapter 1

Introduction

1.1. Motivations and historical notes

During the last few decades, the study of physical properties of the matter at a molecular level has become more and more important because of its multiple applications. We have witnessed exceptional technological advances in the fields of micro- and nanofluidics and their applications in engineering [1, 2, 3], colloidal science [4, 5, 6, 7, 8], drug delivery [9] and bio-engineering [10, 11, 12]. For instance some researchers have shown that controllable acoustic mixing of different fluid in micro-channels can be used for the fabrication of therapeutic nanoparticles [10]. Other research groups have developed procedures for the micro-fabrication of organs-on-a-chip, multi-channel microfluidic devices used to simulate entire organs, which could boost pharmaceutical research while reducing hazardous drugs testing on humans and animals [11]. Additionally, numerous research studies have focused on nanofluidics chips for detection and manipulation of DNA molecules, which are fundamental in the cure of genetic diseases [12].

The development of products and processes involving micro-/nanofluidics requires a deep understanding of fluid properties at the nanoscale. This becomes particularly important when dealing with single or multi-phase fluids in micro-/nano-channels, where the atomistic interactions between different fluids and between fluids and solid substrates give rise to a complex landscape of heterogeneous fluid properties. Experimental observations of fluids at the nanoscale, obtained by techniques such as scanning tunnelling microscopy, are extremely complex and mostly limited to static properties. At the

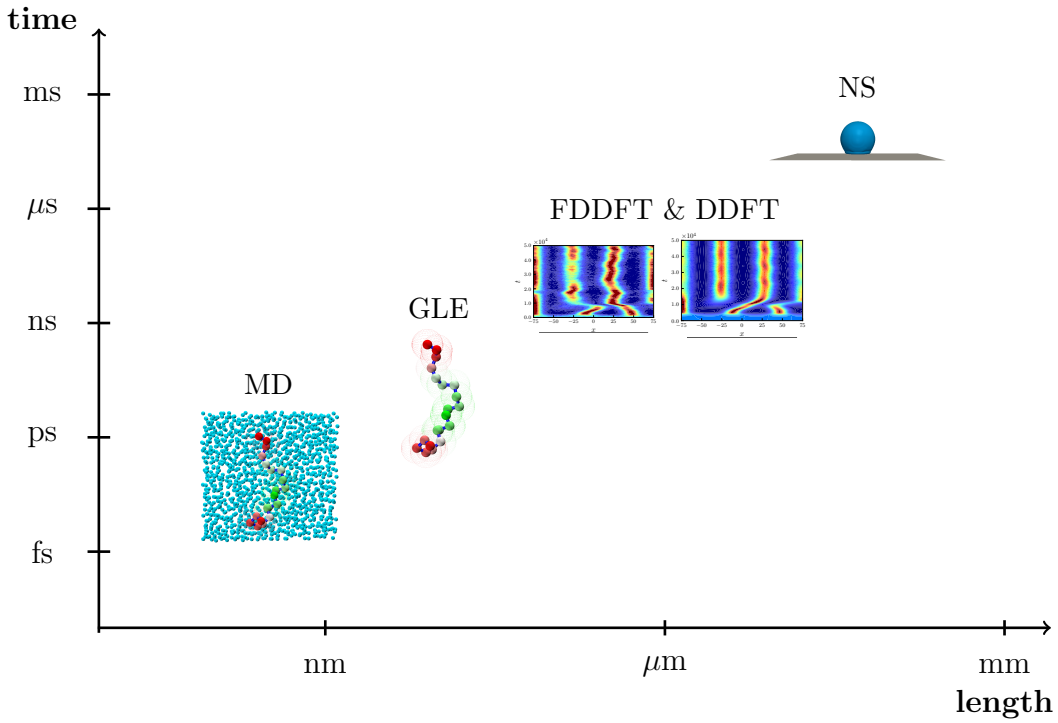

Figure 1.1. Some numerical techniques adopted for the simulations of complex systems across different time- and length-scales. MD: molecular dynamics, GLE: generalized Langevin equation, DDFT: dynamical density functional theory, FDDFT: fluctuating DDFT, NS: Navier-Stokes equations.

same time, continuum models based on Navier-Stokes equations are not able to predict molecular-scale phenomena, where the discrete nature of matter comes to the fore.

For these reasons, extensive studies of fluids behaviour at the nanoscale should be performed using consolidated approaches able to describe inhomogeneous fluids with an atomistic resolution. Moreover, novel frameworks and numerical schemes are needed to overcome several limitations of the existing techniques currently employed in the simulations of complex systems.

1.1.1. Molecular dynamics

Molecular dynamics (MD) is a numerical technique based on the solution of particles equations of motion. It is employed to simulate the time evolution of particles coordinates and, thus, to extract the characteristic properties of the system as a whole. In contrast to the top-down approach of continuum methods, MD simulations require no

prior assumption on the local fluid properties. As such, MD is a suitable technique to gain knowledge on the micro- and macroscopic relationships which intertwine the different physical quantities of a system, with local viscosity being one of the most elusive.

MD was first introduced by Alder and Wainwright in the 1950's in their preliminary work on phase transitions of hard sphere systems [13, 14]. These studies were based on MD simulations of only few hundreds of particles due to the scarce machines computational power available at the time. A few years later, in 1964 a system of 864 particles interacting with a Lennard-Jones potential, which is considered a realistic potential for liquid argon, was simulated for the first time by Rahman [15]. Stillinger and Rahman pushed notably the capabilities of MD, performing the first simulations of liquid water in the 1970' [16, 17]. In 1977, MD was for the first time employed to simulate a protein, the bovine pancreatic trypsin inhibitor, in a work by McCammon *et al.* [18]. Since then, researchers have adopted extensively MD to study the dynamics of complex systems such as homogeneous and heterogeneous fluids [19], lipids [20], protein folding [21], enzymes [22] and DNA [23].

Numerous works have employed MD to describe and eventually model microscopic fluid properties to integrate them into continuum approaches [24, 25, 26, 27, 28, 29, 30]. Nevertheless, such macroscopic relationships are yet to be derived for general conditions and studies on interfacial phenomena often rely on quite simplistic hypotheses. For instance, in widely adopted methods, such as level-set [31] and volume of fluid [32], the variations of density and viscosity at the interface between two fluids characterized by bulk densities ρ_1 and ρ_2, and bulk viscosities η_1 and η_2 take the form:

$$\begin{cases} \rho(\mathbf{r}) = & \rho_1 + (\rho_2 - \rho_1)\, f(\mathbf{r}) \\ \eta(\mathbf{r}) = & \eta_1 + (\eta_2 - \eta_1)\, f(\mathbf{r}) \end{cases} \tag{1.1}$$

where $f(\mathbf{r})$ is a space dependent function. In level-set method $f(\mathbf{r}) = H_\epsilon(\phi(\mathbf{r}))$ with $\phi(\mathbf{r})$ being a level set function and H_ϵ a smoothed step function, while in volume of fluid method $f(\mathbf{r}) = V_2(\mathbf{r})/(V_1(\mathbf{r}) + V_2(\mathbf{r}))$, with $V_1(\mathbf{r})$ with $V_2(\mathbf{r})$ being the local volumes of the two fluids. In both cases, it is assumed that the local viscosity has a linear functional dependence on the local density at the interface, namely

$$\frac{\eta(\mathbf{r}) - \eta_1}{\eta_2 - \eta_1} = \frac{\rho(\mathbf{r}) - \rho_1}{\rho_2 - \rho_1}, \tag{1.2}$$

but this assumption may not always be valid. Similarly, the local viscosity behaviour at the contact point between a fluid and a solid substrate still remains elusive. For

these reasons, more detailed investigations are necessary. In the present work, we will use MD simulation to investigate fluid properties at fluids interface. Specifically, in Chapter 2 we discuss MD technique in more details and employ non-equilibrium MD simulations of Lennard-Jones fluids sheared between two walls to scrutinise inhomogeneous fluid properties at the liquid-solid, liquid-liquid and liquid-vapour interfaces for different temperatures and particles interaction parameters.

1.1.2. Generalized Langevin equation

MD simulations have mainly two limitations. Firstly, typical MD simulation time-steps (function of the particles collision time) is of the order of femtoseconds, while many real fluid phenomena occur on a time scale larger than milliseconds. Secondly, MD cannot deal with systems containing a number of particles of the order of the Avogadro's number($N_{\mathrm{A}} \sim 10^{23}$). This is due to the fact that the computational power required scales at least linearly with the number of particles (more often quadratically due to the two-body atomistic interactions). To overcome this limitations, researchers have developed some dimensionality reduction techniques, which make use of projection operators to reduce the number of observables to simulate [33, 34, 35, 36]. The first application of the projection operators goes back to the works by Mori and Zwanzig who formally derive the dynamics of a Brownian particle, previously described only phenomenologically by Langevin. Mori-Zwanzig formalism allows to derive the dynamic evolution of the observables in the form of generalized Langevin equations (GLEs). GLEs have a stochastic form that includes a non-Markovian term in the form of a time-convolution which is, in general, a complex function of the original high-dimensional system. However, in many relevant cases [37, 38, 39, 40], the memory term can be simplified and expressed as a time-convolution between the set of observables of interest $\boldsymbol{\mathcal{O}}$ and a tensor function $\boldsymbol{\theta}(t)$, known as memory kernel. In these cases, GLEs take the following form:

$$\partial_t \boldsymbol{\mathcal{O}}(t) = \mathbf{F}(\boldsymbol{\mathcal{O}}) - \int_0^t \boldsymbol{\theta}(\tau)\boldsymbol{\mathcal{O}}(t-\tau)d\tau + \mathbf{R}(t). \tag{1.3}$$

The term $\mathbf{F}(\boldsymbol{\mathcal{O}})$ depends only on the current system configuration and corresponds to the mean force term [37, 38, 39, 40]. Several approaches have been developed to compute the potential of mean force of a system, including adaptive biasing forces [41] and umbrella sampling [42].

The vector $\mathbf{R}(t)$, being orthogonal to $\boldsymbol{\mathcal{O}}$, is interpreted as a stochastic term, with correlation given by the fluctuation dissipation theorem $\langle \mathbf{R}(t), \mathbf{R}(t') \rangle = \boldsymbol{\theta}(t-t')\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}} \rangle$.

The non-Markovian term depends on the previous states of the system and is char-

acterized by the memory kernel function $\boldsymbol{\theta}(t)$, which also unequivocally determines the characteristics of the noise term $\mathbf{R}(t)$ through the fluctuation dissipation theorem. The parametrization and numerical simulation of the non-Markovian term often lead to computational challenges because of its convolution structure. To remove such complexity, the memory kernel $\boldsymbol{\theta}(t)$ is commonly approximated by a Dirac-delta function $\boldsymbol{\theta}(t) = 2\boldsymbol{\theta}_0\delta(t)$ with $\boldsymbol{\theta}_0$ being a constant tensor. This procedure leads to the Markovian approximation of the GLE, known as Langevin equation(LE):

$$\partial_t \boldsymbol{\mathcal{O}}(t) = \mathbf{F}(\boldsymbol{\mathcal{O}}) - \boldsymbol{\theta}_0\boldsymbol{\mathcal{O}}(t) + \mathbf{R}(t). \tag{1.4}$$

It is worth highlighting that, because of the properties of Dirac-delta functions, in LE (Eq. (1.4)) the memory term loses its convolution structure, i.e. $\int_0^t \boldsymbol{\theta}(\tau)\boldsymbol{\mathcal{O}}(t-\tau)d\tau = \int_0^t 2\boldsymbol{\theta}_0\delta(\tau)\boldsymbol{\mathcal{O}}(t-\tau)d\tau = \boldsymbol{\theta}_0\boldsymbol{\mathcal{O}}(t)$, and the stochastic term becomes delta-correlated in time, i.e. $\langle \mathbf{R}(t), \mathbf{R}(t')\rangle = 2\boldsymbol{\theta}_0\langle \boldsymbol{\mathcal{O}}_0, \boldsymbol{\mathcal{O}}_0\rangle\delta(t-t')$. Despite being widely adopted for its simplicity, the Markovian approximation above often introduces significant inaccuracies in the coarse-grained system. As a result, in general, GLE (Eq. 1.3) should be employed to preserve the main features of the original high dimensional system into the reduced one and, thus, a robust methodology to extract $\boldsymbol{\theta}(t)$ from the original high-dimensional system is required.

In previous studies, several approaches have been proposed to parametrize the memory kernel. Analytical forms can be only obtained for simple systems, such as a particle in a harmonic oscillator heat bath [43], while numerical parametrization techniques are necessary for more complex systems characterized by non-linear interactions. For instance, in Ref. [44] the authors adopt a perturbation scheme, which is yet "too complex for general use". Despite its accuracy, the algorithm developed in Ref. [45] to parametrize GLEs involves sampling of the full original system, and thus becomes computationally prohibitive for large systems. Another procedure involving large matrix computations and Krylov sub-space approximations has been proposed in Ref. [40]. In Ref. [46], an iterative approach is used to compute a discrete approximation of $\boldsymbol{\theta}(t)$ from the system autocorrelation functions. In both Refs [47, 48], the researchers have proposed to extract the memory kernel by Laplace transforming the correlation functions computed from some historical data of the observables. However, these strategies exhibits serious limitations when the available data on the observables are affected by fluctuations, as shown in Section 3.1. It follows that a robust methodology for an optimal and practical parametrization of the memory kernel is still missing. In Section 3.1, we propose a novel data-driven approach which makes use of machine learning tools to parametrise

the memory kernel in a general GLE.

1.1.3. Fluctuating dynamical density functional theory

The Langevin description discussed above is commonly adopted to simulate only a subset of an high dimensional system. However, in the field of particle fluids, the coarse-grained system obtained is usually still atomistic, and consequently its numerical simulation is characterized by an high computational complexity. A continuum framework capable of overcoming this limitation, while preserving small-scales system behaviour, is offered by density functional theory (DFT). Unlike the atomistic GLEs, DFT and its dynamical counterpart known as dynamical DFT (DDFT) operate with local averaged quantities and are thus computationally more tractable. DDFT formalism, embedded with either exact or approximated models for a density-dependent free-energy functional [49], has already shown its enhanced capabilities in the study of complex system at the nano- and micro-scale [50, 51, 52]. Examples of such applications include the study of nucleation for colloids and macro-molecules [53, 54], the simulation of fluids in confined geometries [55, 56, 57] and the analysis of wetting phenomena [58, 59, 60]. DDFT is usually obtained by deriving the Fokker-Planck equation governing the evolution of the system probability density function, which is then averaged over all but one degrees of freedom [56, 61, 62, 63]. Several extensions of DDFT have been proposed to account for multi-component systems [62], to include hydrodynamic interactions [56, 63, 64], and to embed the effects of particle orientability [65].

The inclusion of thermal fluctuations in the DDFT framework, which is derived through a mean field approach, has posed many questions [66]. In some studies, such as the work by Elder *et al.* [67], an additional noise term had been artificially included. Nevertheless, the physical meaning of this term is dubious. This long standing debate was clarified in the recent work by Durán-Olivencia *et al.* [68], who proposed a derivation of fluctuating DDFT (FDDFT) starting from a Langevin description of the time-evolution of colloidal particles immersed in a thermal bath. Using a proper ensemble average under local equilibrium approximation, they derived the governing equations of the stochastic density ρ and velocity $\mathbf{v}$ fields for a colloidal fluid in a thermal bath, i.e.

$$\partial_t \rho(\mathbf{r},t) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(m^{-1} \rho(\mathbf{r},t) \mathbf{u}(\mathbf{r},t) \right) = 0 \tag{1.5}$$

$$\begin{aligned} \partial_t \left(\rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t) \right) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\rho(\mathbf{r},t)\mathbf{u}(\mathbf{r},t) \otimes \mathbf{u}(\mathbf{r},t) \right) + \rho(\mathbf{r},t) \boldsymbol{\nabla}_{\mathbf{r}} \frac{\delta \mathcal{E}[\rho]}{\delta \rho(\mathbf{r},t)} \\ + \boldsymbol{\theta}_0 \rho(\mathbf{r},t) \mathbf{v}(\mathbf{r},t) + \sqrt{k_B T m \rho(\mathbf{r},t) \boldsymbol{\theta}_0} \boldsymbol{\mathcal{W}}(\mathbf{r},t) = 0 \end{aligned} \tag{1.6}$$

where m is the atomic mass of the colloidal particles, $\mathcal{E}[\rho]$ is a density-dependent energy functional, $\boldsymbol{\theta}_0$ is a friction tensor depending on the interactions between colloidal and bath particles, k_B is the Boltzmann constant, T is the systems temperature and $\boldsymbol{\mathcal{W}}$ is a vector of Gaussian stochastic processes with zero mean and delta-correlated in space and time, i.e.

$$\langle \boldsymbol{\mathcal{W}}(\mathbf{r},t)\rangle = 0, \tag{1.7}$$

$$\langle \boldsymbol{\mathcal{W}}(\mathbf{r},t), \boldsymbol{\mathcal{W}}(\mathbf{r}',t')\rangle = 2\delta(t-t')\delta(\mathbf{r}-\mathbf{r}'). \tag{1.8}$$

Despite the exceptional insights in their derivation, the work by Durán-Olivencia *et al.* [68] relies on some simplifications:

- The colloidal particles are assumed to belong the same single species. However, many real physical systems of interest include multiple components.
- The system particles are considered to be non-reactive. However, many industrial application, such as the analysis of chemical reactors [69], require a framework able to deal with reactive systems.
- The stochastic time-evolution of the Langevin particles system is assumed to follow a Markovian stochastic dynamics (Eq. 1.4). This assumption is satisfied when the system is characterized by a clear scale separation between colloidal particles and bath. However, in many applications, a non-Markovian approach (Eq. 1.3) is required to preserve the atomistic correlations in the continuum description.

To overcome these limitations, in Section 3.2 we propose a non-Markovian FDDFT framework for reacting multi-species particles. Moreover, we develop a finite-volume discretization scheme which enables us to employ FDDFT in the study of relevant physical phenomena, such as Turing patterns formation in binary reacting systems.

1.1.4. Overdamped FDDFT & DDFT

As discussed above, the FDDFT framework derived in Ref. [68] provides the governing equations for the time-evolution of density and momentum fields of a colloidal system. In the strong damping limit (i.e. when $m^{-1}\boldsymbol{\theta}_0 \to \infty$), the characteristic time scale of the momentum dynamics is much shorter than the density one [52, 68] and, thus, the contributions of the terms $\boldsymbol{\nabla}_{\mathbf{r}}\cdot(\rho\mathbf{v}\otimes\mathbf{v})$ and $\frac{\partial \rho\mathbf{v}}{\partial t}$ can be neglected. As a result one obtains

the so called overdamped FDDFT which does not involve the velocity field [68, 70]:

$$\partial_t \rho(\mathbf{r},t) = \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\boldsymbol{\theta}_0^{-1} \rho(\mathbf{r},t) \boldsymbol{\nabla}_{\mathbf{r}} \frac{\delta \mathcal{E}[\rho]}{\delta \rho(\mathbf{r},t)} \right) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\sqrt{k_B T \rho(\mathbf{r},t) \boldsymbol{\theta}_0^{-1}} \boldsymbol{\mathcal{W}}(\mathbf{r},t) \right). \tag{1.9}$$

For the simple ideal-gas free-energy functional $\mathcal{E}[\rho] = \rho\,(\log \rho - 1)$, Eq. (1.9) reduces to the stochastic diffusion equation [71]. However, the presence of the functional $\mathcal{E}[\rho]$ allows for the introduction of non-linear diffusion, external force fields and non-local interactions. Equation (1.9) may be seen as a stochastic version of the gradient flow equation previously studied, for instance, in Refs [72, 73].

Different numerical methods have been proposed in the past to simulate fluctuating equations similar to Eq. (1.9). Many numerical methodologies focused on the Landau-Lifshitz-Navier-Stokes (LLNS) equations [74]. For instance, Garcia *et al.* [75] proposed a finite difference scheme to discretize the LLNS equations for dilute gases. Finite different schemes, embedded with a third-order Runge-Kutta time integrator, were also developed by Bell *et al.* [76, 77] for the discretization of the full LLNS equations. Their scheme was able to reproduce the proper the fluctuation correlations for density, momentum and energy fields. They also introduced the analysis of the structure factor (equilibrium fluctuation spectrum) to assess the accuracy of the scheme in reproducing the proper fluctuations for the different fields. The same group also constructed alternative methods to solve FH via staggered grids [78]. Other works have proposed numerical schemes based on temporal integrators that are implicit-explicit predictor-corrector [79] or two-level leapfrog [80]. Additionally, hybrid schemes coupling LLNS with MD [81, 82, 83] or with Monte Carlo [84, 85] simulations of complex fluid systems have been developed. Finally, other works, such as Refs [71, 86], developed numerical methods for reaction-diffusion equations, obtained by adding reaction terms to the overdamped FDDFT (Eq. (1.9)) equipped with the ideal-gas free-energy functional. These works provided a better understanding of the role of thermal fluctuations in fluids. However, a numerical methodology able to simulate stochastic gradient flow equations such as Eq. (1.9) equipped with a general free-energy functional is not readily available yet. This methodology could be useful in the simulation of noise-driven non-equilibrium phenomena, such as energy-barrier crossing transitions, in the framework of FDDFT. Also, not many work in the literature tackled the numerical simulation of gradient flow equations such as Eq. (1.9) with a non-constant diffusion ($\boldsymbol{\theta}_0$ parameter).

In Section 4.1 we propose a finite volume numerical scheme for solving general stochastic gradient flow equations, which resemble Eq. (1.9) of FDDFT. Additionally, in Section 4.2 we develop a finite volume scheme to simulate the deterministic counterpart

of Eq. (1.9), known as overdamped DDFT, equipped with space- and time-dependent friction parameters $\boldsymbol{\theta}_0$.

1.2. Main aims and outline of the thesis

The general aim of this work is twofold. On the one side, it aims to contribute to the understanding of key properties of the matter at the nanoscale, extending the current knowledge on the topic. On the other side, it aims to extend the current state of the art in numerical simulations of complex systems by developing new theoretical frameworks and numerical methods. More specifically, this thesis focuses on:

- The study of fluid properties in interfacial regions to develop ready-to-use relations for the design and analysis of microfluidic devices. By means of atomistic simulations, we obtain relations between viscosity and density profiles for confined fluids in a variety of conditions. These relations can be employed, for instance, to computed the friction factor of fluid flows in nano-pipes or as a closures for non-equilibrium continuum frameworks (such as DDFT) applied at nanoscopic level.

- The development of mathematical tools for a data-driven stochastic modelling of complex systems. We couple elements of machine learning with the formal structure of the generalized Langevin equation to model complex systems dynamics under noise-governed conditions based on historical data. We shown that our methodology has applications also in climatology and finance.

- The derivation of a new theoretical framework able to describe with a continuum approach the dynamics of colloidal particles in a thermal bath retaining detailed information on the colloids-bath coupling. The developed framework is also extended to multi-species reacting particles systems, and has relevant applications in the study of Turing patterns.

- the development of novel numerical methods for (both deterministic and stochastic) continuum approaches to simulate fluids at the nanoscopic scale. We extend and generalize existing finite-volume methods to include stochastic flux, external fields and inter-particles potentials.

Because of the wide range of topics and fields involved, this work has many direct applications in the fields of technology and science. As an example, a deep knowledge of fluid properties at the nano-scale is crucial in the design and analysis of nanofluidic devices, which are widely adopted in environmental engineering and medicine. Moreover,

our methodology for parametrization of stochastic modelling is successfully employed to model not only chemical and physical systems, but also data series in finance and climatology. Finally, multiple parts of this work were published in recognized open-access journals and presented in international conferences to help advance knowledge and progresses of the whole scientific community. In what follows, we provide a detailed overview of the main topics covered in each chapter.

1.2.1. Chapter 2

In Chapter 2, we employ MD technique to study and model physical properties of fluids at interfaces with an atomistic resolution. Specifically, in Section 2.1, molecular dynamics technique is introduced. We discuss governing equations, common time integration schemes and thermostat algorithms. Next, we include the derivations of formulas relating atomistic information, i.e. particles positions and momenta, with macroscopic observables, e.g. density, stress tensor, ... MD is then employed in Section 2.2 to study the local properties of a fluid system confined between solid walls. Based on MD simulations results, local properties, such as density and stress tensor, are analysed. Moreover, using a semi-empirical approach, we propose local and non-local models relating shear viscosity and density profiles. Finally, in Section 2.3, MD is used to simulate multi-phase Lennard-Jones fluid and gather insights on the fluid properties at fluid interfaces. Following a data-driven approach, we propose a macroscopic relation between shear viscosity and density profiles at interfaces between two immiscible fluids and at liquid-vapour interfaces.

1.2.2. Chapter 3

Chapter 3 deals with dimensionality reduction techniques applied mainly to atomistic physical systems. Section 3.1 provides a discussion of GLEs. They involve a non-trivial parametrization of the non-Markovian term known as memory kernel. In this work, we present a novel data-driven approach, which makes use of neural networks (multilayer perceptron) to obtain an optimal approximation of the memory kernel from historical data. The multi-exponential form of the activation function employed in such approximation allows us to derive an extended dynamics framework to integrate the GLE. The proposed methodology is tested in various application: from coarse-graining of thermal baths to modelling global temperature dynamics in climatology and a stock price index in finance.

In Section 3.2 we propose a theoretical derivation of non-Markovian FDDFT for

multi-species reacting species. The microscopic GLEs governing the particles dynamics is our starting point. After defining density and momentum fields, we derive the non-Markovian FDDFT and its overdamped limit. We then include an extended field variables dynamics framework to overcome the computational challenges due to the convolution structure of the memory term and the time-correlated noise. A finite volume scheme is adopted for the spatio-temporal discretization of the extended variable dynamics. Finally, we include a numerical validation of the non-Markovian FDDFT against MD simulations and an application of FDDFT to the study of Turing patterns in binary systems.

1.2.3. Chapter 4

In Chapter 4 we present numerical methods developed to solve overdamped DDFT in the presence of additional complexities, such as fluctuations or heterogeneous thermal baths. In Section 4.1 we develop a finite volume numerical scheme for solving general stochastic gradient flow equations. We present a space discretization for the deterministic flux based on a hybrid approach which takes advantage of both central and upwind schemes. Families of implicit-explicit Euler-Maruyama and Milsten schemes, with a weak second order Runge-Kutta scheme are employed as time integrators. To overcome the commonplace challenge of preserving non-negative densities in the presence of a stochastic flux, an adaptive time steps based on the Brownian bridge technique is adopted. We use the proposed scheme to simulate the time-evolution of physical systems characterized by different free-energy functionals, validating the simulations results against MD and theory. Additional numerical applications in this section include the time-evolution of an ideal gas in a double well potential and the homogeneous nucleation kinetics of a Lennard-Jones like fluid.

In Section 4.2, we propose a finite-volume schemes to solve gradient flow equations (obtained in the framework of DDFT) with a non-homogeneous diffusion (or friction) coefficient. First- and second-order upwind schemes for the space discretization of one- and two-dimensional systems are derived. We then prove that these schemes satisfy a decay of the free-energy functional in time. Finally, we employ the proposed schemes to simulate colloidal systems immersed in thermal baths characterized by anisotropic, space-dependent and time-dependent properties.

1.2.4. Chapter 5

In Chapter 5, a summary of the results of each chapter is reported together with the main conclusions. We also discuss open questions related to the topics covered in this work and we provide an outlook on future research works.

Chapter 2

Molecular dynamics

2.1. MD: Background and numerical methods

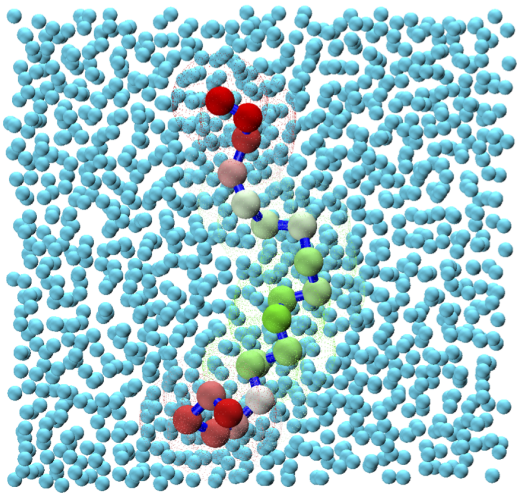

Figure 2.1. Representation of a MD simulation of a chain of particles immersed in a thermal bath.

Molecular dynamics simulations are based on the solution of particles equations of motion. Let us consider a system of N particles with positions $\mathbf{r}(t) = \mathbf{r}_1(t), \ldots, \mathbf{r}_N(t)$ and momenta $\mathbf{p}(t) = \mathbf{p}_1(t), \ldots, \mathbf{p}_N(t)$. The internal energy of an atomistic system is

determined by the Hamiltonian $\mathcal{H}$

$$\mathcal{H} = \sum_{i=1}^{N} \frac{\mathbf{p}_i \cdot \mathbf{p}_i}{m_i} + U\left(\mathbf{r}_1, \ldots, \mathbf{r}_N\right) \tag{2.1}$$

where U is the potential energy of the system which depends in general on the absolute and relative positions of the particles, namely

$$U = \sum_{i=1}^{N} v_1\left(\mathbf{r}_i\right) + \sum_{i=1}^{N}\sum_{j>i}^{N} v_2\left(\mathbf{r}_i, \mathbf{r}_j\right) + \sum_{i=1}^{N}\sum_{j>i}^{N}\sum_{k>j}^{N} v_3\left(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k\right) + \ldots \tag{2.2}$$

with v_1, v_2 and v_3 denote the one-body, two-body and three-body terms, respectively.

The time evolution of position $\mathbf{r}_i(t)$ and momentum $\mathbf{p}_i(t)$ of the particle i is governed by Newton's equations:

$$\frac{d\mathbf{r}_i}{dt} = \frac{d\mathcal{H}}{d\mathbf{p}_i} = \frac{\mathbf{p}_i}{m_i} \tag{2.3}$$

$$\frac{d\mathbf{p}_i}{dt} = -\frac{d\mathcal{H}}{d\mathbf{r}_i} = \mathbf{F}_i \tag{2.4}$$

with $\mathbf{F}_i$ being the total force acting on the particle i.

When system constrains, number of particles and potential interactions do not explicitly dependent on time, the Hamiltonian coincides with the total energy of the system, namely

$$\frac{d\mathcal{H}}{dt} = \sum_{i=1}^{N}\left(\frac{d\mathcal{H}}{\mathbf{r}_i}\cdot\frac{d\mathbf{r}_i}{dt} + \frac{d\mathcal{H}}{\mathbf{p}_i}\cdot\frac{d\mathbf{p}_i}{dt}\right) = \sum_{i=1}^{N}\left(\frac{d\mathcal{H}}{\mathbf{r}_i}\cdot\frac{d\mathcal{H}}{\mathbf{p}_i} - \frac{d\mathcal{H}}{\mathbf{p}_i}\cdot\frac{d\mathcal{H}}{\mathbf{r}_i}\right) = 0 \tag{2.5}$$

In MD, the dynamical evolution of multi-body systems is simulated by using Eq. (2.4). The simulation results are then analysed to extract physical quantities of the system such as density and stress tensor, and transport properties such as shear and volume viscosities.

It is worth highlighting that the integration of Newton's equations for real systems is usually computationally expensive for multiple reasons. First, the time-step size adopted in the time integration must be smaller than the fastest time-scale in the system, typically the bond vibration time which is of the order of femtoseconds. However, typical transitions in real systems occur over a time of the order of seconds. Second, real systems contain a number of particles of the order of Avogadro's number (10^{23}). Thus the computational cost, which scales with the square of the particles number if two-body

interactions are used, could exceed the available machine capabilities.

2.1.1. Time integration schemes

A commonly adopted time-integration algorithm is the velocity Verlet algorithm, which can be expressed in the form:

$$\mathbf{r}_i(t+\Delta t) = \mathbf{r}_i(t) + \frac{\mathbf{p}_i(t)}{m_i}\Delta t - \frac{\boldsymbol{\nabla}_{\mathbf{r}_i}\mathcal{H}(t)}{2m_i}\Delta t^2 \tag{2.6}$$

$$\mathbf{p}_i(t+\Delta t) = \mathbf{p}_i(t) - \frac{\boldsymbol{\nabla}_{\mathbf{r}_i}\mathcal{H}(t) + \boldsymbol{\nabla}_{\mathbf{r}_i}\mathcal{H}(t+\Delta t)}{2}\Delta t \tag{2.7}$$

with Δt being the integration time-step.

Velocity Verlet algorithm is third order accurate in time, i.e. its error is $\mathcal{O}(\Delta t^3)$, and is symplectic, i.e. it conserves the area in phase space delimited by an ensemble of systems. The latter property is particularly important as it ensures the conservation of the average Hamiltonian.

2.1.2. Thermostats

Solving the equation of motion with the classical definition of Hamiltonian generates trajectories of the system with constant energy, consistent with an NVE ensemble. However, many physical scenarios require molecular dynamics simulations to be run at constant temperature, such that the system can be represented by an NVT ensemble. This is achieved by including additional terms in the equations of motion, as if the system was in contact with a thermostat. During the last century, several kinds of thermostats have been proposed. Here, we briefly discuss the Langevin [87] and Nosé-Hoover [88, 89, 90] thermostats.

Langevin thermostat The Langevin thermostat is a stochastic thermostat acting on every particle of the system. Its formulation can be derived through Mori-Zwanzig formalism, assuming the presence of a heat bath in contact with the system. In general a system controlled by a Langevin thermostat has the following equations of motion:

$$\partial_t \mathbf{r}_i = \boldsymbol{\nabla}_{\mathbf{p}_i}\mathcal{H} \tag{2.8}$$

$$\partial_t \mathbf{p}_i = -\boldsymbol{\nabla}_{\mathbf{r}_i}\mathcal{H} - \int_0^t \boldsymbol{\theta}(t-\tau)\cdot\mathbf{p}_i(t)d\tau + \mathbf{R}(t) \tag{2.9}$$

where θ is a time-dependent matrix determining the system-bath coupling and $\mathbf{R}(t)$ a Gaussian process satisfying:

$$\langle \mathbf{R}(t) \rangle = 0 \tag{2.10}$$

$$\langle \mathbf{R}(t)\mathbf{R}^T(s) \rangle = m k_B T \boldsymbol{\theta}(t-s) \tag{2.11}$$

However, such formulation of the thermostat may often be inpractical to adopt in simulations because of its convolution term depending on the whole history of the system and its correlated noise. Thus, it is commonly assumed that the system has no memory of its previous state, i.e. that $\boldsymbol{\theta}(t) \sim 2\boldsymbol{\theta}_0\delta(t)$, with $\boldsymbol{\theta}_0$ being a constant matrix and $\delta(t)$ the Dirac delta function. Moreover, it is also commonly assumed that the thermal bath is isotropic, i.e. that the memory (or friction) term can be rewritten as $\boldsymbol{\theta}_0 = \theta_0\mathbf{1}$. It follows that the commonly adopted Langevin thermostat has then the following form:

$$\partial_t \mathbf{r}_i = \boldsymbol{\nabla}_{\mathbf{p}_i} \mathcal{H} \tag{2.12}$$

$$\partial_t \mathbf{p}_i = -\boldsymbol{\nabla}_{\mathbf{r}_i} \mathcal{H} - \theta_0 \mathbf{p}_i + \mathbf{R}(t) \tag{2.13}$$

where θ_0 determines the coupling between system and thermal bath, and the random contribution is a Gaussian process satisfying

$$\langle \mathbf{R}(t) \rangle = 0 \tag{2.14}$$

$$\langle \mathbf{R}(t)\mathbf{R}^T(s) \rangle = 2 m k_B T \theta_0 \delta(t-s)\mathbf{1} \tag{2.15}$$

Nosé-Hoover thermostats The Nosé-Hoover (NH) thermostat is deterministic and time-reversible. It was derived by modifying the Hamiltonian with an extra degree of freedom for the heat bath s. The NH modified Hamiltonian reads:

$$\mathcal{H}_N(\mathbf{r}, \mathbf{p}, s, p_s) = \sum_{i=1}^{N} \frac{\mathbf{p}_i(t) \cdot \mathbf{p}_i(t)}{2 m_i s} + U(\mathbf{r}) + \frac{p_s^2(t)}{2Q} + (3N+1) k_B T \ln(s). \tag{2.16}$$

where Q determines how fast the system temperature converges to the desired one. The equations of motion resulting from the NH Hamiltonian read:

$$\partial_t \mathbf{r}_i = \boldsymbol{\nabla}_{\mathbf{p}_i} \mathcal{H} \tag{2.17}$$

$$\partial_t \mathbf{p}_i = -\boldsymbol{\nabla}_{\mathbf{r}_i} \mathcal{H} - \xi \mathbf{p}_i \tag{2.18}$$

$$\partial_t \xi = \frac{1}{Q} \left(\sum_{i=1}^{N} \frac{\mathbf{p}_i(t) \cdot \mathbf{p}_i(t)}{2 m_i} - \frac{3}{2} N k_B T \right). \tag{2.19}$$

It has been shown that the NH thermostat performs well for systems in equilibrium and for weakly sheared systems. Unlike the Langevin thermostat, it does not alter the value of the system properties, such as viscosity [91]. However, being a global thermostat, the NH thermostat suffers inhomogeneous conditions of non-equilibrium system where the friction heat is not generated uniformly. To overcome this issue, when necessary, a series of local Nosé-Hoover thermostats is commonly adopted [26].

2.1.3. Boundary conditions

Boundary conditions play a fundamental role in atomistic simulations. Depending on the system under study, particles can be simulated in a box with flexible, periodic or fixed boundaries.

A simulation box with flexible sizes is used when the simulated system can change its size and shape in time. It can be adopted to study, for instance, the properties of a single molecule or cluster.

Periodic boundary conditions are employed to analyse the bulk properties of a large system by simulating only a representative portion of it, i.e. a unit cell. Specifically one simulates only the particles within the unit cell. However the unit cell is surrounded by an infinite number of replicas of itself. It follows that every particle can interact with other particles both in the unit cell and in the neighbouring image cells. In periodic boundary conditions, the minimum-image convention is commonly followed. This prescribe that each particle can interact only with the closest image of every other particle. It follows that the cut-off radius of non-bonded interactions can be at most equal to half the domain length.

Fixed boundary conditions are typically employed when a particles system is confined by an external potential or rigid/semi-rigid atomistic walls placed at the borders of the system. There are at least three ways of modelling an atomistic wall [92]. A first option is to fix particles at their lattice position, forming an atomically smooth surface able to appropriately model very rigid materials. Although this approach is justifiable for the study of various fluid properties that are not notably affected by wall dynamics, rigid walls are not able to absorb heat from the fluid, thus not allowing for a realistic modelling of heat transfer between walls and fluids. Consequently, if the systems generates energy, as in the case of sheared systems, the fluid itself has to be thermostatted in order to dissipate the generated heat. An advantage of the rigid walls is that interactions between wall particles are not computed, thus reducing the computational cost. Alternatively, particles can be connected to their lattice site by means of harmonic springs whose stiffness is related to the 'softness' of wall materials. This kind of walls are able to

absorb energy from the system, (although this may not be sufficient to keep the system temperature constant, e.g. in case of large shear rates). As shown later in this work, a combination of different boundary conditions, one for each direction, can be employed.

2.1.4. Reduced units

When working at the atomistic scales, SI units may not be opportune because often too large compared to the typical values to measure. To overcome this, all the quantities of interest are usually scaled. Lennard-Jones reduced units are obtained by scaling every measure with respect to the fundamental quantities: σ_{LJ}, ϵ_{LJ} and m, corresponding to atomic diameter, energy and atomic mass, respectively. The values of the fundamental quantities depend on the specific fluid and can be used to move from reduced to real units. For instance, LJ parameters for Argon take the following values: $\sigma_{\mathrm{LJ}} = 3.405 \times 10^{-10} m$, $\epsilon_{\mathrm{LJ}} = 1.65 \times 10^{-21} J$ and $m = 6.69 \times 0^{-26} Kg$.

In this work all physical quantities are non-dimensionalized. In Table 2.1 we report the relation between real and reduced units for some relevant physical quantities.

Table 2.1. Physical quantities for a d-dimensional system in reduced units.

Quantity	real units	reduced units
distance	$\mathbf{r}^*$	$\mathbf{r} = \mathbf{r}^*/\sigma$
energy	E^*	$E = E^*/\epsilon$
density	ρ^*	$\rho = \rho^* \sigma^d$
temperature	T^*	$T = k_B T^*/\epsilon$
pressure	P^*	$P = P^* \sigma^d/\epsilon$
time	t^*	$t = t^* \sqrt{\epsilon/m\sigma^2}$

2.1.5. From atomistic coordinates to macroscopic fields

Particles trajectories obtained via MD simulation can be adopted to compute macroscopic fields characterizing the corresponding system, e.g. state variables and transport coefficients. This can be achieved mainly with two methods, which differ in the way they distribute the microscopic information in space.

A first approach consists in dividing the spatial domain into a finite number of bins and assigning the information of each particle to the bin in which the centre of the particle resides. This approach is computationally cheap and easy to implement, but it disregards the finite particle size. Furthermore, when a particle resides anywhere within the confines of a bin, its information is effectively assigned to the central position of a bin, thus potentially introducing errors. The shift of information is averted in the limit

of infinitesimally thin bins, but this also requires extremely long simulations to gather sufficient statistics in each bin. More importantly, it is impossible to analytically calculate gradients of the computed profiles, which instead require an additional numerical approximation.

Some of these limitations can be overcome by locally distributing the atomistic information via smooth differentiable kernels, based on smoothed particle hydrodynamics [93]. In this second approach, a macroscopic field $X(\mathbf{r}, t)$ is expressed as follows,

$$X(\mathbf{r}, t) = \sum_i \chi_i \phi(\mathbf{r}_i(t) - \mathbf{r}), \tag{2.20}$$

where χ_i is the information of interest of the $i-th$ particle located at $\mathbf{r}_i$ at time t, and ϕ is a kernel function with unitary area. The most commonly used kernels in the literature are described by piecewise constant, Gaussian, or polynomial functions. These kernels will be briefly introduced below. For the purpose of this study, profiles are calculated as functions of one spatial coordinate y, and the dependency on time is omitted. Instead, stored information is averaged under the ergodic assumption (i.e., for a given physical macroscopic property X, the ensemble and time averages are considered equivalent: $\langle X \rangle_{\mathrm{NVT}} = \langle X \rangle_{\mathrm{t}}$). A piecewise constant function (mostly adopted in mesh-based approaches) is defined as:

$$\phi(y) = \begin{cases} \frac{1}{2L_x L_z w} & \text{for } \|y\| < w, \\ 0 & \text{otherwise }, \end{cases} \tag{2.21}$$

with L_x and L_z being the total system length along the x and z directions, respectively, and w is the half-width of the function. A Gaussian kernel is given by:

$$\phi(y) = \begin{cases} \frac{1}{L_x L_z \sqrt{2\pi w^2}} e^{-\frac{\|y\|^2}{2w^2}} & \text{for } \|y\| < R_c, \\ 0 & \text{otherwise }, \end{cases} \tag{2.22}$$

where w^2 is the variance. Also linear [94], cubic [95] and quartic [96] splines have been adopted in the literature. For instance, a cubic spline is defined as:

$$\phi(y) = \begin{cases} \frac{1}{L_x L_z w} \left(1 - 3\frac{\|y\|^2}{w^2} + 2\frac{\|y\|^3}{w^3}\right) & \text{for } \|y\| < w, \\ 0 & \text{otherwise }, \end{cases} \tag{2.23}$$

where w controls the width of the kernel.

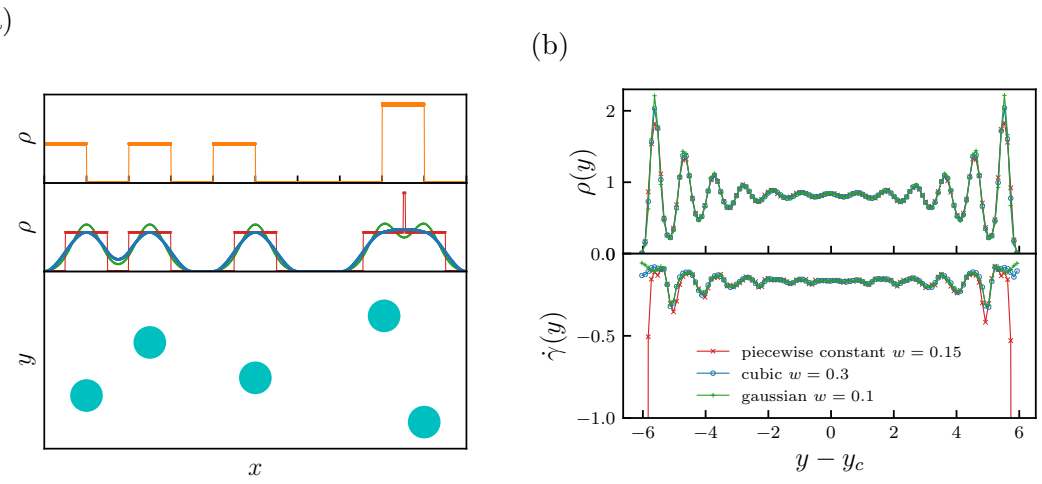

Figure 2.2. (a) Density profiles of a system made of fixed particles (bottom panel). The top panel shows a mesh-based approach with bins of width $w = 0.3$ to compute the density. In the middle panel the density profiles are computed by using three different kernels: piecewise constant with $w = 0.15$ (in red), cubic with $w = 0.3$ (in blue) and Gaussian with $w = 0.1$ (in green). (b) Density and shear rate profile computed for a confined particle system at $T = 1.0$ with three different kernel functions (piecewise constant, Gaussian and cubic). Cubic and Gaussian kernels give similar results, while the piecewise constant kernel introduces some errors at the boundaries. This inaccuracy in the shear rate may directly affect its derived quantities, including shear viscosity.

The Gaussian kernel is more computationally expensive to evaluate than a polynomial function, and provides no compact support since the Gaussian function never reaches zero. However, it can be truncated and reweighed so that the information is conserved. On the other hand, one advantage of the Gaussian kernel is that it allows for exactly calculating the n^{th} derivative of a macroscopic field X as:

$$\frac{dX^n(y)}{dy^n} = \left\langle \sum_i \chi_i \frac{d^n \phi(y_i - y)}{dy^n} \right\rangle_t , \tag{2.24}$$

where the angle brackets with subscript t denote a time average. Conversely, the smoothness and the number of derivatives that can be taken of a polynomial kernel is limited by the order of the polynomial, which also determines the computational cost of the kernel.

In Fig. 2.2(a) we compare, for a system of five particle floating in space, the density profiles computed by using binning and different kernel-based methods. In the binning approach, the domain is divided into ten bins of equal width (top panel). The number density in a bin then equals the number of particles in that bin divided by its volume.

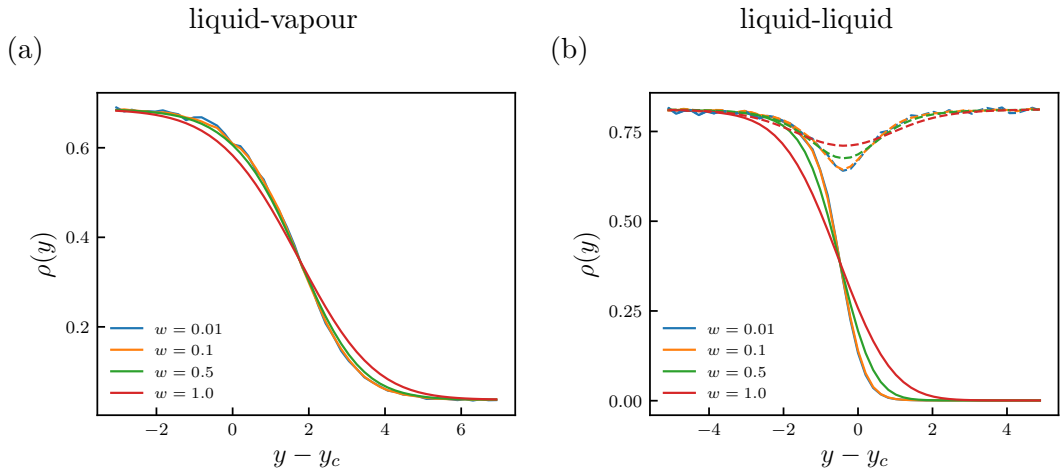

Figure 2.3. Influence of the Gaussian kernel parameter w on the density profiles evaluated at liquid-vapour (a) and liquid-liquid (b) interfaces for a temperature $T = 1.0$. In (b), solid lines represent the single fluid density, while dashed lines refer to the mixture density. The centre of the channel is denoted by y_c.

As shown in the top panel of Fig. 2.2(a), this approach may lead to an unphysical shift of peaks in the the density profile with respect to the particles centres. A second approach consists in adopting a mesh-free discretization employing piecewise constant, cubic and Gaussian kernels. This method allows for an arbitrary number of computational points within the domain and, thus, a finer description of the density. The piecewise constant kernel assumes that the information of a particle is homogeneously distributed across a width of $2w$. Consequently, the piecewise constant profile can still cause some artificial density peaks, as seen on the right in the middle panel of Fig. 2.2(a).

Various recent studies, especially in the areas of granular materials [97, 98] and molecular dynamics [26, 99, 100], have reported the kernel effects on the calculated macroscopic profiles. For example, Weinhart *et al.* [97] found, using granular flow simulations, that the Gaussian smoothing parameter w should be small compared to the smallest length scale to be captured.

For Lennard-Jones fluids confined between two walls, the smallest length scale of interest is approximately $0.8\sigma_{\mathrm{LJ}} - 1.0\sigma_{\mathrm{LJ}}$, corresponding to the period of density oscillations near the walls [101]. For these systems, in this work we adopt a Gaussian kernel with $w = 0.1$, truncated at $6.5\sigma_{\mathrm{LJ}}$ (half domain). The gradients of macroscopic fields can be particularly sensitive to the method used to distribute information. In Fig. 2.2(b) we compare shear rate profile for a sheared Lennard-Jones fluid, calculated with the different kernels. The widths of the kernels are chosen to spread the data approximately

equal. The shear rate profiles are nearly identical far from the walls, where the fluid is more homogeneous. However, close to the walls, the profile calculated with the piecewise constant kernel deviates from the other analytically differentiated profiles.

A Gaussian kernel embedded with a value $w = 0.1$ is also adopted to extract the macroscopic properties in multi-phase Lennard-Jones systems. This choice is tested in Fig 2.3 where we compare the density profiles obtained with several values of w. As shown in Fig 2.3, values of w greater than 0.1 lead to an artificial smoothing in the density profile for both liquid-vapour and liquid-liquid interfaces. On the contrary, the density profiles obtained with $w = 0.1$ and $w = 0.01$ overlap, thus confirming a convergence of the macroscopic field values.

In Sections 2.2 and 2.3 we analyse single and multi-phase fluids with heterogeneous behaviours along the y-direction. Thus, in the following paragraphs we report the formulas adopted to extract the local physical quantities of interest, e.g. density, velocity and stress tensor components, as a function of y.

Density

The steady-state density profile of the α component of a multi-species fluid is evaluated as:

$$\rho_\alpha(y) = \left\langle \sum_{i \in \alpha} m_i \phi(y_i - y) \right\rangle_t \tag{2.25}$$

where m_i is the atomic mass and y_i denotes the y-coordinate of the i-th particle belonging to the species α.

The total mass density is then obtained by summing the density profiles of all the components at every point in space, namely

$$\rho(y) = \sum_\alpha \rho_\alpha(y). \tag{2.26}$$

Velocity and shear rate

The velocity profile of the system component α is evaluated as function of the momentum $\mathbf{j}_\alpha(y)$ and density $\rho_\alpha(y)$ profiles as:

$$\mathbf{u}_\alpha(y) = \frac{\mathbf{j}_\alpha(y)}{\rho_\alpha(y)} = \frac{1}{\rho_\alpha(y)} \left\langle \sum_{i \in \alpha} m_i \mathbf{u}_i \phi(y_i - y) \right\rangle_t, \tag{2.27}$$

with $\mathbf{u}_i$ denoting the velocity vector of the particle i belonging to the species α. Similarly, the velocity profile of the whole system is computed as function of total mass density and total momentum density as $\mathbf{u}(y) = \mathbf{j}(y)/\rho(y)$.

Employing the quotient rule, then we can compute the shear rate profile $\dot{\gamma}_\alpha(y) = \partial u_{\alpha,x}/\partial y$ as function of mass density and x-component of the momentum, namely

$$\dot{\gamma}_\alpha(y) = \frac{1}{\rho_\alpha(y)^2}\left[\rho_\alpha(y)\frac{dJ_{\alpha,x}(y)}{dy} - \frac{d\rho_\alpha(y)}{dy}J_{\alpha,x}(y)\right] \tag{2.28}$$

where the gradients appearing in the numerator can be evaluated through Eq. (2.24). Similarly it is possible to compute the shear rate profile of the total system as $\dot{\gamma}(y) = \partial u_x/\partial y$.

Temperature

The temperature profile of a system is computed along the y axis as function of the number density $\rho(y)$ and the kinetic energy of the particles, namely

$$T(y) = \frac{1}{3\rho(y)}\left\langle \sum_i m_i(\mathbf{u}_i - \mathbf{u})\cdot(\mathbf{u}_i - \mathbf{u})\ \phi(y_i - y)\right\rangle_t . \tag{2.29}$$

Stress tensor

The Irving-Kirkwood-Noll formulation is adopted to compute the stress tensor components from particles coordinates [102, 103]. Following Schofield and Henderson's work, to derive the stress tensor contributions we start differentiating in time the momentum

equation:

$$\frac{\partial}{\partial t}\boldsymbol{j}(\boldsymbol{r};t) = \frac{\partial}{\partial t}\sum_{i=1}^{N}\boldsymbol{p}_i\delta(\boldsymbol{r}-\boldsymbol{r}_i) \tag{2.30}$$

$$= \sum_{i=1}^{N}\dot{\boldsymbol{p}}_i\delta(\boldsymbol{r}-\boldsymbol{r}_i) - \boldsymbol{\nabla}_r\cdot\sum_{i=1}^{N}\frac{\boldsymbol{p}_i\boldsymbol{p}_i}{m_i}\delta(\boldsymbol{r}-\boldsymbol{r}_i) \tag{2.31}$$

$$= \sum_{i=1}^{N}\boldsymbol{F}_i\delta(\boldsymbol{r}-\boldsymbol{r}_i) - \boldsymbol{\nabla}_r\cdot\sum_{i=1}^{N}\frac{\boldsymbol{p}_i\boldsymbol{p}_i}{m_i}\delta(\boldsymbol{r}-\boldsymbol{r}_i) \tag{2.32}$$

$$\begin{aligned}= &\sum_{i=1}^{N}\sum_{j>i}^{N}\boldsymbol{F}_i\left(\delta(\boldsymbol{r}-\boldsymbol{r}_i)-\delta(\boldsymbol{r}-\boldsymbol{r}_j)\right) - \\ &\boldsymbol{\nabla}_r\cdot\left(\rho\mathbf{u}\mathbf{u} + \sum_{i=1}^{N}m_i(\mathbf{u}_i-\mathbf{u})(\mathbf{u}_i-\mathbf{u})\delta(\boldsymbol{r}-\boldsymbol{r}_i)\right)\end{aligned} \tag{2.33}$$

$$= -\boldsymbol{\nabla}_r\cdot(\rho\mathbf{u}\mathbf{u}-\boldsymbol{\sigma}) \tag{2.34}$$

where $\boldsymbol{\sigma}$ is the stress tensor, $\mathbf{u}$ is the streaming velocity, ρ is the density and where we used that $\boldsymbol{\nabla}_{r_i}\delta(\boldsymbol{r}-\boldsymbol{r}_i) = -\boldsymbol{\nabla}_r\delta(\boldsymbol{r}-\boldsymbol{r}_i)$. From the equation above it follows that the stress tensor can be computed using the relation:

$$\boldsymbol{\nabla}_r\cdot\boldsymbol{\sigma} = \sum_{i=1}^{N}\sum_{j>i}^{N}\boldsymbol{F}_i\left(\delta(\boldsymbol{r}-\boldsymbol{r}_i)-\delta(\boldsymbol{r}-\boldsymbol{r}_j)\right) - \boldsymbol{\nabla}_r\cdot\sum_{i=1}^{N}m_i(\mathbf{u}_i-\mathbf{u})(\mathbf{u}_i-\mathbf{u})\delta(\boldsymbol{r}-\boldsymbol{r}_i) \tag{2.35}$$

Moreover, the difference in deltas can be rewritten as:

$$\delta(\boldsymbol{r}-\boldsymbol{r}_i)-\delta(\boldsymbol{r}-\boldsymbol{r}_j) = -\boldsymbol{\nabla}_r\cdot\boldsymbol{r}_{ij}\int_0^1\delta(\boldsymbol{r}-\boldsymbol{r}_i+\lambda\boldsymbol{r}_{ij})d\lambda \tag{2.36}$$

Thus, the stress tensor is defined (although not uniquely) by the following relation:

$$\boldsymbol{\sigma} = -\sum_{i=1}^{N}\sum_{j>i}^{N}\boldsymbol{r}_{ij}\boldsymbol{F}_i\int_0^1\delta(\boldsymbol{r}-\boldsymbol{r}_i+\lambda\boldsymbol{r}_{ij})d\lambda - \sum_{i=1}^{N}m_i(\mathbf{u}_i-\mathbf{u})(\mathbf{u}_i-\mathbf{u})\delta(\boldsymbol{r}-\boldsymbol{r}_i) \tag{2.37}$$

where the first term constitutes the potential contribution and the second term represents the kinetic term.

Eq. (2.37) highlights that the six independent components of the stress tensor are the results of two different microscopic mechanisms. The momentum transfer through particles thermal motion produces the kinetic contribution, which is predominant at

low densities (e.g. in the vapor phase). In a general multi-component system, the kinetic stress tensor acting on the α component of the system is calculated from particles momenta as:

$$\boldsymbol{\sigma}^k_\alpha(y) = -\left\langle \sum_{i\in\alpha} m_i(\mathbf{u}_i - \mathbf{u}_\alpha(y)) \otimes (\mathbf{u}_i - \mathbf{u}_\alpha(y))\ \phi(y_i - y) \right\rangle_t , \tag{2.38}$$

where, $\mathbf{u}_i$ and $\mathbf{u}_\alpha$ are the absolute and the streaming velocity vectors for the species α, respectively. The total kinetic stress is then obtained as function of absolute and streaming velocity vectors of the full system, i.e. $\boldsymbol{\sigma}^k(y) = \sum_\alpha \boldsymbol{\sigma}^k_\alpha(y)$.

The momentum transfer due to interactions between particles give raise to the potential contribution, prevailing in high density components, such as liquids. The potential stress tensor is computed as:

$$\boldsymbol{\sigma}^v_\alpha(y) = -\left\langle \frac{1}{2} \sum_{i\in\alpha} \sum_{j\neq i} \mathbf{F}_{i,j} \otimes (\mathbf{r}_i - \mathbf{r}_j)\ b_\phi(y; y_i, y_j) \right\rangle_t , \tag{2.39}$$

with $\mathbf{r}_i$ being the position vector of the particle i, $\mathbf{F}_{i,j}$ the force acting on the particle i of the species α due to any other particle j, and $b_\phi(y; y_i, y_j) = \int_{s=0}^{1} \phi((1-s)y_i + sy_j - y)ds$ is the bond function [103, 104]. The total potential stress tensor can be then computed by summing the contributions of all system components, i.e. $\boldsymbol{\sigma}^v(y) = \sum_\alpha \boldsymbol{\sigma}^v_\alpha(y)$. It follows that the full stress tensor for a species α is readily evaluated as:

$$\boldsymbol{\sigma}_\alpha(y) = \boldsymbol{\sigma}^k_\alpha(y) + \boldsymbol{\sigma}^v_\alpha(y). \tag{2.40}$$

2.2. Molecular dynamics of single-phase fluids in confined geometries [1]

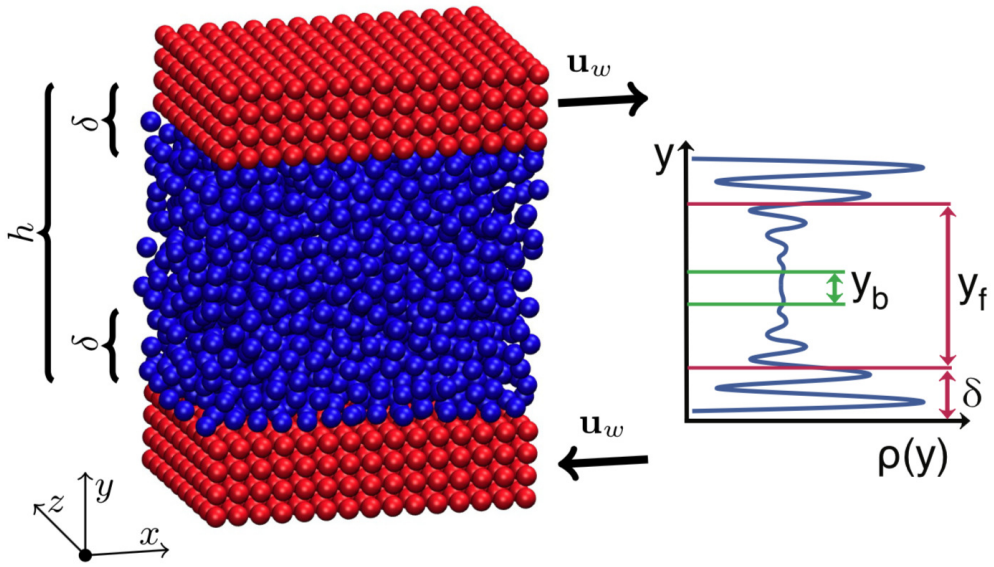

Figure 2.4. Non-equilibrium MD setup for the study of a sheared fluid (blue particles). Walls (red particles) are located at a constant distance $h = 13.0\sigma_{\mathrm{LJ}}$ and move at a constant velocity $\mathbf{u}_w$ in opposite directions . The parameter δ in the figures defines regions at the liquid-solid interfaces where the strong interactions strongly reduce particles mobility, while $y_f = h - \delta$ is its complementary region. The bulk region, which spans $y_b \simeq 4\sigma_{\mathrm{LJ}}$, denotes a part of the domain where no significant layering is observed in the density profile.

In this Section, we present results and analysis of MD simulations used to study the local physical properties of a fluid system confined between two solid substrates. Solid-liquid interactions give rise to non-homogeneous properties close to the wall. In fact, the density profile shows a typical oscillating trend due to the particles layering (Fig. 2.4). Interestingly, also the viscosity profile in the channel exhibits a non-homogeneous behaviour in proximity of the wall. Specifically, we found that this structure of the fluid viscosity is mainly dependent on local density, temperature and wall interaction potential, and thus can be modelled as function of these parameters.

[1] This section includes material from A. Russo, M. A. Durán-Olivencia, S. Kalliadasis, & R. Hartkamp. Macroscopic relations for microscopic properties at the interface between solid substrates and dense fluids, *J. Chem. Phys.*, 150, 214705 (2019).

Modelling the viscosity of fluids is a long standing problem in the fluid-dynamics field. The first successful attempt to model viscosity of a fluid is due to Maxwell in 1860 [105]. Using kinetic theory to study a dilute gas [106], Maxwell concluded that internal friction of a gas is independent of its density. This was later confirmed by Chapman [107] and Enskog [108]. However, shortly after, Enskog showed that, unlike dilute gases, the viscosity of dense gases varies with density and temperature [109]. Since then, researchers have tried to generalize Enskog's theory to inhomogeneous systems by utilizing both theoretical and numerical tools. Din and Michaelides [25] proposed a theoretical framework to model the viscosity of strongly inhomogeneous gases using kinetic theory. They validated such framework against MD simulations, showing that the proposed formulation can only qualitatively predict the viscosity profile of an interacting Lennard-Jones fluid. Later, Zhang *et al.* [110] employed a different approach to study the viscosity profile of inhomogeneous fluids. They parametrized a nonlocal constitutive model for the viscosity by means of an effective space-dependent viscosity kernel, and then applied this approach to confined fluids undergoing a Poiseuille flow. Despite the interesting methodology, Zhang *et al.*'s approach is approximate since it does not take into account any explicit dependence on the density profile and is impractical as it does not provide a general simple formula for the local viscosity. More recently, Hoang *et al.* [27, 111, 112] proposed local average-density models combined with weight functions to include non-local effects on the viscosity profile. In case of fairly adsorbent narrow pores this empirical method leads to modelled viscosity profiles in line with MD results. Finally, Morciano *et al.* [30] adopted a fully data-driven approach to extract an easy-to-use relation between viscosity and density profiles at solid-fluid interfaces. Despite the practical relation proposed in their study, a connection with existing theoretical models was not provided. Moreover the system properties were investigated only for a single temperature.

While considerable efforts have been made to consider effects of temperature and wall-interaction strength on local viscosity, a generic model remains elusive. In this work, we employ non-equilibrium MD simulations of Lennard-Jones fluids sheared between two walls to scrutinise inhomogeneous fluid properties at liquid-solid interfaces. Specifically, we model the viscosity profile of inhomogeneous dense LJ fluids as a functional of density profile, system temperature and wall interactions. The proposed models are validated validated in a range of interfacial conditions and temperatures, and they represent a generalization of the relation proposed by Morciano *et al.* [30]. These relations can be easily employed in the design and analysis of nanofluidic devices, where solid-fluid interfaces play a crucial role.

2.2.1. Simulation details

A Couette flow is simulated by shearing two parallel atomistic walls with a fluid sandwiched between them, as shown in Fig. 2.4. The fluid and wall particles are interacting through a Lennard-Jones potential, typically used to model a charge-neutral and inert material. Wall particles are structured in a square lattice, with a lattice spacing of σ_{LJ}. The interaction between any two particles i and j is described by the two-body Lennard-Jones potential U_{LJ}:

$$U_{\mathrm{LJ}}(r_{ij}) = 4\ \epsilon_{\mathrm{LJ},ij}\ \left[\left(\frac{\sigma_{\mathrm{LJ},ij}}{r_{ij}}\right)^{12} - \left(\frac{\sigma_{\mathrm{LJ},ij}}{r_{ij}}\right)^{6}\right], \tag{2.41}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between the particles, $\epsilon_{\mathrm{LJ},ij}$ is the depth of the potential well and $\sigma_{LJ,ij}$ is the finite atom-atom distance at which the potential is zero. The interaction potential is truncated at a cut-off radius r_c, such that $U_{\mathrm{LJ}}(r_{ij} \geq r_c) = 0$. Walls are modelled as a rigid square particles lattice. Consequently, the fluid itself has to be thermostatted to dissipate the heat generated due to shear. Such direct control over the fluid temperature allows, in principle, for maintaining the desired fluid temperature under arbitrarily high shear rates. Thus, allowing for simulations with a high signal-noise ratio.

The accessible fluid domain (delimited by the centres of the walls particles in direct contact with the fluid) is $[15\sigma_{\mathrm{LJ}}] \times [13\sigma_{\mathrm{LJ}}] \times [10\sigma_{\mathrm{LJ}}]$ and is kept constant throughout the simulations. We impose periodic boundary conditions along the x and z axes, while the fluid is enclosed by the walls along the y direction. A steady shear flow is generated by imposing equal and opposite velocities along the y direction to the two walls [113]. The heating caused by shearing of the fluid is removed from the system by thermostatting the fluid particle velocities in the directions perpendicular to the flow. This is done by applying separate Nosé-Hoover [88, 89, 90] thermostats to slabs of fluid, dividing the fluid region in 13 equally-sized slabs across the channel. The system is simulated at constant temperatures, $1.0 \leq T \leq 4.0$. The influence of the applied strain rate on the system properties with respect to the type of thermostat is discussed in Appendix 2.2.2.

The simulations in this study are performed using the Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS) [114]. The equations of motion are integrated using the velocity-Verlet [115] algorithm, with a time-step of $dt = 0.005\tau$ [116]. Simulations are performed at constant temperature, volume and number of particles.

As regards Lennard-Jones potential parameters, we adopt the following values: $\epsilon_{\mathrm{LJ},f,f} = \epsilon_{\mathrm{LJ}}$, $\epsilon_{\mathrm{LJ},f,w} = 0.6\epsilon_{\mathrm{LJ}}, 1.0\epsilon_{\mathrm{LJ}}, 1.4\epsilon_{\mathrm{LJ}}$, and $\sigma_{\mathrm{LJ},f,f} = \sigma_{\mathrm{LJ},f,w} = \sigma_{\mathrm{LJ}}$. Furthermore, a cutoff

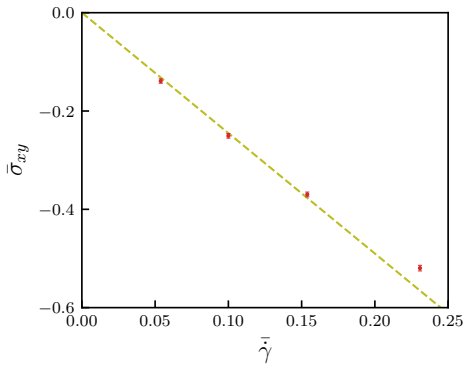

Figure 2.5. Average shear stress against average shear rate is reported for the most critical scenario ($T = 1.0$ and $\epsilon_{f,w} = 0.6$). Non-Newtonian effects are evidenced at average shear rates above 0.2.

radius $r_c = 3.5\sigma_{\mathrm{LJ}}$ is employed. Without loss of generality, σ_{LJ}, ϵ_{LJ}, m and the Boltzmann constant k_B are set equal to unity.

MD simulations are performed as follows. First, the system is equilibrated at equilibrium conditions for 2×10^6 time steps. Then, a constant velocity is imposed to the walls and a nonequilibrium steady state is reached after 2×10^6 time steps, ensured by convergence of the density, velocity and temperature profiles. Finally, a run of 2×10^6 time steps is performed, during which fluid particle positions and velocities are stored every 10^2 time steps for analysis.

2.2.2. Rheological effects and thermostat

Non-Newtonian effects manifest at high shear rates, with the onset shear rate depending on the state point of the fluid. For example, for a dense homogeneous LJ fluid ($\rho = 0.84$, $T = 1.0$) shear-thinning is observed at shear rates above approximately 0.1 [117]. Other rheological phenomena, such as normal stress differences, have been shown to be very small for shear rates up to 0.5 for a homogeneous fluid near the LJ triple point ($\rho = 0.8442$, $T = 0.72$) [118]. In Fig. 2.5, we report the average shear stress in our confined sheared system against the average shear rate for the most critical analysed case ($T = 1.0$ and $\epsilon_{f,w} = 0.6$). No significant rheological effects are evidenced for shear rates below 0.2. Therefore, we adopt $\bar{\dot{\gamma}} \sim 0.15$ in order to maximize the signal-to-noise ratio, while remaining in the Newtonian regime.

The temperature in a MD simulation can be controlled via various thermostats, including Langevin [87], Nosé-Hoover [88, 89, 90] and dissipative particle dynamics [119]

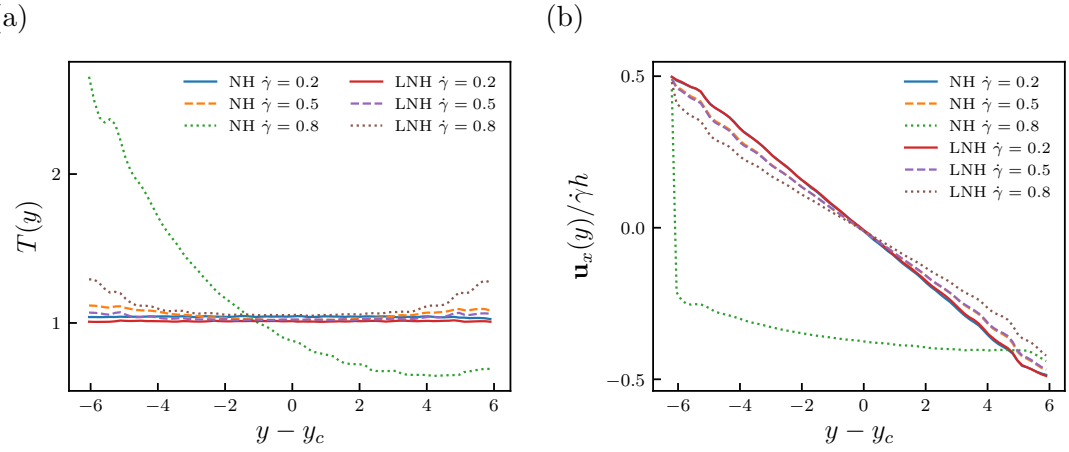

Figure 2.6. (a)Temperature profile for different shear rates, when using a single global Nosé-Hoover thermostat and a series of Nosé-Hoover thermostats, each controlling a region of (non-dimensional) unitary width. (b)Velocity profile induced by the moving walls at several shear rates at the temperature $T = 1.0$.

thermostats. In this work, a Nosé-Hoover (NH) thermostat is adopted since it has been shown to perform well in case of weakly sheared systems and, unlike Langevin and dissipative particle dynamics thermostats, does not alter systems viscosity [91]. However, NH is a global thermostat, and consequently may suffer for any inhomogeneous conditions of the system since the heat is not generated uniformly. To overcome this limitation, we adopted a series of local Nosé-Hoover thermostats (LNH), each of (non-dimensional) unitary width. A similar strategy was already adopted in previous works [26]. Figs 2.6(a-b) compare temperature and velocity profiles obtained with NH and LNH thermostats for different shear rates. The global thermostat is unable to keep a flat temperature in proximity of the walls, where the local shear rate is higher. At the same time, at high shear rates (i.e. 0.8), NH thermostat leads to unphysical velocity and temperature profiles, thus the need for LNH thermostat.

2.2.3. Density

The fluid structure density exhibit a rich behaviour across the channel. The region at the centre of the channel, of width approximately equal to $y_b \simeq 4\sigma_{\mathrm{LJ}}$, does not show significant layering, and thus, it will be denoted as bulk region. The density profiles in Fig. 2.7 reveal that the intensity of the density layering near the wall depends inversely on temperature and directly on $\epsilon_{\mathrm{LJ},f,w}$. On the other hand, the location of the layers remains unaffected, as it depends primarily on the wall structure and on the interaction

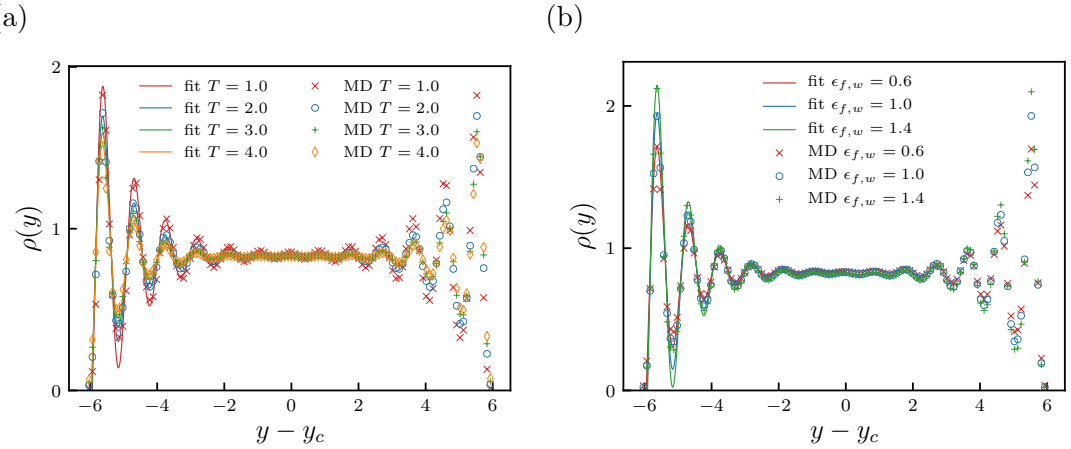

Figure 2.7. (a) Density profile at different temperatures of the domain, $1.0 \leq T \leq 4.0$; (b) Density profile for different interaction potential between walls and fluid, $\epsilon_{\mathrm{LJ},f,w} = 0.6, 1.0, 1.4$. The markers represent MD simulation results, while the fitting relation given by Eq. (2.42) is reported in solid lines.

length scale parameter σ, which are both identical for all our simulations. The dense layers are equidistant, with a mean distance between adjacent peaks of about $l = 0.93$, which is consistent with Refs [19, 26] and in reasonable agreement with the one predicted by kinetic theory: $l \sim \left(\sqrt{2}\sigma_{\mathrm{LJ}}\bar{\rho}\right)^{-1} = 0.88$. The density profiles obtained from MD are well approximated with the following exponential oscillatory relation:

$$\rho(y) = \rho_b \left[1 + Ae^{-B(y-y_w)} \cos\left(\frac{2\pi}{l}(y - y_w)\right)\right], \tag{2.42}$$

where ρ_b is the density in the bulk region of the channel, the fitting parameter $A = A(\epsilon_{\mathrm{LJ},f,w}, T)$ depends on the oscillations amplitude and $B = B(\epsilon_{\mathrm{LJ},f,w}, T)$ indicates the decay away from the interface, located at y_w. It turns out that the amplitude is mainly dependent on the fluid-wall interactions, and can be expressed as $A \sim 3.5\,\epsilon^{0.5}_{\mathrm{LJ},f,w}$, whereas, the decay parameter depends mostly on the temperature as $B \sim 0.85\,T^{0.3}$. The comparison of Eq. (2.42) with MD data (fitting curves shown in Fig. 2.7) leads to relative errors ranging between 2% and 10%.

Wall potential, standoff distance and confinement parameter The mean potential energy acting on the fluid due to its interactions with the walls can be computed with the following two approaches. First, the effective mean potential energy, that takes into account both wall-fluid and fluid-fluid particles interactions, can be calculated from

the density profile at equilibrium according to the Boltzmann distribution [120, 121]:

$$U_{\text{eff}}(y) = -k_B T \log\left(\frac{\rho(y)}{\rho_b}\right), \tag{2.43}$$

where ρ_b is the bulk density. Alternatively, the mean interaction potential acting on a fluid due to an infinite LJ lattice wall can be modelled analytically as [101]:

$$U_{\text{LJ}_{10-4}}(y) = \frac{2\rho_w \epsilon_{\text{LJ},f,w}}{5}\left[2\left(\frac{\sigma_{f,w}}{y - y_w}\right)^{10} - 5\left(\frac{\sigma_{LJ,f,w}}{y - y_w}\right)^{4}\right], \tag{2.44}$$

where ρ_w is the density of the walls. The latter expression has the advantage of providing an analytical form for the wall potential, that is independent from the density profile. However, this approach neglects the influence of the interactions between fluid particles, thus potentially leading to significant inaccuracies in the case of strong fluid-wall interactions.

The density layering at the walls is a result of the competition between potential energy, which favours an ordered structure of the fluid, and particles thermal energy, which enhances the jumps among the different potential wells. Following the work of Chiavazzo *et al.* [122], we introduce the total energy profile as:

$$E(y) = U(y) + \alpha\frac{k_B T}{2}, \tag{2.45}$$

where $U(y)$ is the potential energy acting on the fluid particles due to the walls and the other fluid particles, and α indicates the fraction of the average thermal energy per particle available to escape the wall absorption. Because of the equipartition theorem, each degree of freedom has associated a kinetic energy of $k_B T/2$. Since particles in proximity of a flat wall are able to overcome the potential energy barriers only along a single direction normal to the solid surface (representing half degree of freedom), α is taken equal to 1/2 [122]. $E(y)$ gives access to two important quantities: the standoff distance and the confinement parameter. The standoff distance $y_{\min}$ is defined as the minimum distance from a wall accessible to the fluid:

$$y_{\min} = n_1 - y_w, \tag{2.46}$$

with n_i, $i = 1, \ldots, r$ being the roots of the function $E(y) = 0$, and y_w the position of the first layer of wall particles. It turns out that close to the walls, being $U(y)$ extremely sharp, one can safely assume $U(y) \sim E(y)$. As a result, the standoff distance can be

readily obtained by finding the roots of $U_{\mathrm{LJ}_{10-4}}(y) = 0$, rather than $E(y) = 0$. It follows that $y_{\min} \sim (2/5)^{1/6} \sim 0.86$, as also verified in Wang *et al.* [101]. The comparison with our MD results shows that this approximation is able to accurately model the standoff distance with an error lower than 3%.

The confinement parameter δ denotes the fluid region where wall effects are significant. A way of estimating δ was introduced by Chiavazzo *et al.* [122] and is given by the following expression:

$$\delta = n_r - y_w. \tag{2.47}$$

The confinement parameter δ provides a quantitative means of defining two kinds of region inside the channel: a confined one, where the influence of the wall is significantly affecting particles mobility, and a free region of width y_f where the particles motion is only marginally affected by the presence of the fluid-solid interface. In general the free region does not coincide with the bulk region of width y_b, being the latter completely unaffected in structure by the presence of the wall (i.e. no layering at all is observed in the density profiles). Fig. 2.8(a) shows the total energy landscape $E(y)$ for two limit cases: one with hydrophobic walls at high temperature (weak layering) and one with hydrophilic walls at low temperature (strong layering). The standoff distance $y_{\min}$ and the confinement parameter δ are indicated for the latter case. For the sake of completeness, we also report the potential energy excess at the wall computed directly from inter-atomic interactions. The small discrepancy observed between the energy landscape computed from particle interactions and the one obtained from the Boltzmann distribution does not significantly affect the value of the parameter δ.

Fig 2.8(b) shows the parameter δ, computed with both approximations of $U(y)$, as a function of the Wall number $\mathrm{Wa} = (\rho_{\mathrm{wall}}\sigma_{\mathrm{LJ}}^2\epsilon_{\mathrm{LJ},f,w})/(k_B T)$, which is a measure of the influence of the walls on the fluid [101]. The values of δ obtained for $U_{\mathrm{eff}}(y)$ are, in most cases, greater than the ones obtained for $U_{\mathrm{LJ}_{10-4}}(y)$ and in both cases δ increases with Wa. The greater values for δ obtained for $U_{\mathrm{eff}}(y)$ indicate that intramolecular correlations between fluid particles (not taken into account by $U_{\mathrm{LJ}_{10-4}}(y)$), enlarge the fluid region with strong confinement effects. Furthermore, the values of δ calculated from $U_{\mathrm{LJ}_{10-4}}(y)$ increase smoothly with Wa, whereas those evaluated from $U_{\mathrm{eff}}(y)$ show a stepwise trend. Such behaviour is related to the number of fluid layers intercepted by the horizontal axis in the total energy profile.

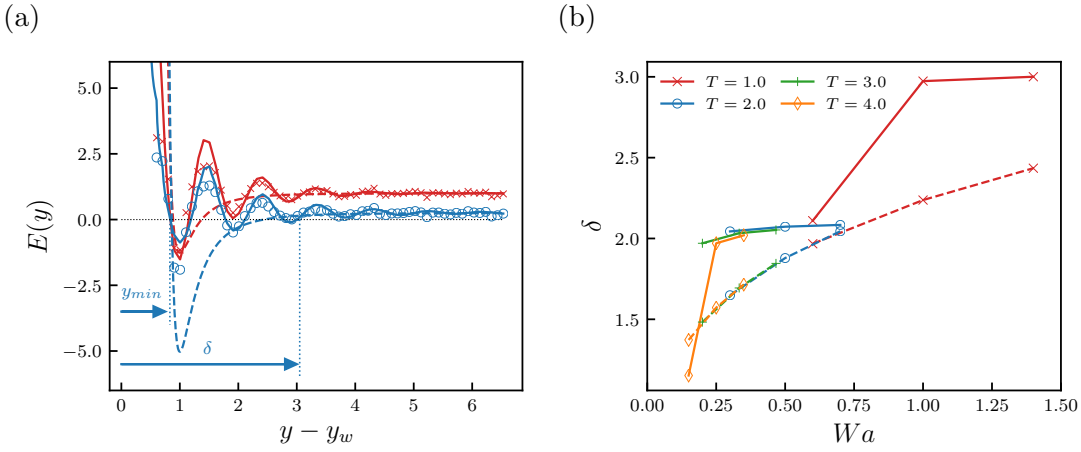

Figure 2.8. (a) Interfacial energy profile computed with potential energy U_{eff} (solid lines) and $U_{\text{LJ}_{10-4}}(y)$ (dashed lines) in two representative cases, one with $\epsilon_{\text{LJ},f,w} = 0.6$ and $T = 4.0$ (in red), the other with $\epsilon_{\text{LJ},f,w} = 1.4$ and $T = 1.0$ (in blue). Markers represent the potential energy excess at the wall computed directly from inter-atomic interactions. The roots of this functions are used to define y_{min} and δ parameters through Eqs (2.46)-(2.47). (b) The parameter δ as function of the Wall number when computed from U_{eff} (solid lines) and from $U_{\text{LJ}_{10-4}}(y)$ (dashed lines).

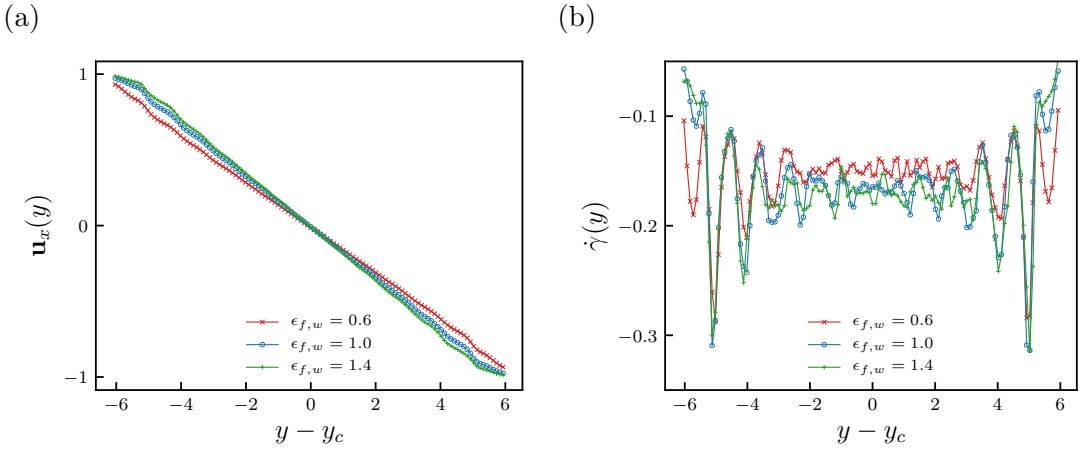

Figure 2.9. (a) Velocity and (b) shear profiles for $\epsilon_{\text{LJ},f,w} = 0.6, 1.0, 1.4$ at $T = 1.0$.

2.2.4. Velocity and shear rate

A Couette flow between two planar walls is expected to produce a linear velocity profile for homogeneous fluids. However the layered fluid structure near the walls causes a non-uniform mobility of the particles, resulting in oscillations of the velocity profile

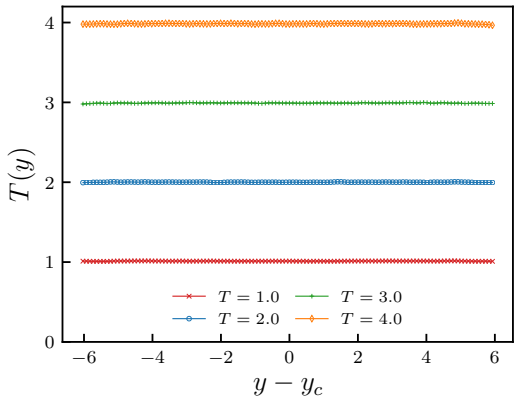

Figure 2.10. Temperature profiles at different temperatures of the domain, namely $1.0 \leq T \leq 4.0$ for $\epsilon_{\mathrm{LJ},f,w} = 1.0$.

(Fig. 2.9(a)). In contact with the solid surface, a layer of fluid atoms is adsorbed to the surface. The thickness of this layer depends on the strength of the fluid-wall interactions. Furthermore, an increase in strength of the walls interaction produces a reduced width of the channel available for the mobile particles, and this results in a higher shear rate in the bulk region of the channel (Fig. 2.9(b)). In addition, the value of the shear rate evaluated at the wall will significantly depend on the wall-fluid interactions, due to slip effects. This phenomenon plays an important role in the study of boundary conditions and has been widely studied in the past, for instance through the analysis of a parameter known as slip length [113]. A comparison between Figs 2.7(b) and 2.9(b) highlights that the lowest peaks in the density correspond to the highest values (in absolute value) of the shear rate. This suggested that a modest slip occurs not only in the contact layer with the wall, but also between fluid layers.

2.2.5. Temperature

The temperature distribution in a nano-channel depends on material and geometry of the walls, type of fluid and shear rate. When a fluid is sheared, energy losses in form of viscous dissipation turn into heat. According to Navier-Stokes energy equation, the heat production is proportional to the fluid viscosity η and to the square of the shear rate $\dot{\gamma}$. Moreover, the amount of heat dissipated through the walls depends on the thermal conductivity according to the Fourier's law and on the specific topology of the walls [123]. Because of this mechanism of heat transfer, the temperature distribution across the channel in real scenarios is strongly dependent on the specific system and its

working conditions.

In this study we restrain our attention to weakly sheared systems with highly conductive walls. In such cases, the generated heat is assumed to be instantaneously dissipated through the walls. As a consequence, the temperature distribution can be assumed to be constant throughout the channel. This assumption, on the one side allows us to use the canonical ensemble to represent the possible states of the fluid system at an atomistic level, on the other side justifies the use of local thermostats, previously adopted in other works [26] and described in Section 2.2.2. The effectiveness of the thermostat in keeping the temperature distribution constant is shown in Fig. 2.10.

2.2.6. Stress tensor

In order to analyse and validate the simulations results for the stress tensor, let us consider the momentum equations and apply it to our system. Firstly, for the system under consideration, the steady-state condition satisfies: $\partial_t \left(\rho(y,t)\mathbf{u}(y,t)\right) \sim 0$. Moreover, due to the periodic conditions along x and z directions, the thin film approximation gives: $\rho(\mathbf{r}) \sim \rho(y)$, $\mathbf{u}(\mathbf{r}) \sim \mathbf{u}(y)$, $\boldsymbol{\sigma}(\mathbf{r}) \sim \boldsymbol{\sigma}(y)$, $\mathbf{s}_b(\mathbf{r}) \sim \mathbf{s}_b(y)$. It follows that the integrated momentum equations integrated in a planar control volume delimited by two general coordinates y_1 and $y_2 > y_1$ can be written as:

$$\sigma_{\alpha y}(y_2) - \sigma_{\alpha y}(y_1) + \int_{y_1}^{y_2} s_{b,\alpha}(y)dy = 0, \tag{2.48}$$

with $\alpha = \{x, y, z\}$. Specifically, if we consider a control volume that contains only fluid particles not subject to any external force, the source term is null, i.e. $s_{b,\alpha}(y) = 0$, and consequently $\sigma_{\alpha y}(y_2) = \sigma_{\alpha y}(y_1)$, meaning that $\sigma_{\alpha y}$ is expected to be constant along the direction normal to the walls. On the contrary, if the control volume includes also the walls particles which are constrained in their positions by means of external forces (or geometric constrains), then the source term is different than zero and $\sigma_{\alpha y}(y)$ is not constant along the y direction.

As reported in Figure 2.11 (a-b), our MD results are consistent with the derivation above. In fact, Figure 2.11(a) shows the diagonal component of the stress tensor $\sigma_{yy}(y)$, including its kinetic and potential components, for a system of temperature $T = 1.0$ and $\epsilon_{\mathrm{LJ},f,w} = 1.0$. The profiles of the kinetic and potential components have both an anisotropic behaviour. However, since they oscillate in antiphase with identical amplitude, the resultant $\sigma^t_{yy}(y)$ is constant (as expected from the momentum equations). Because of the system configuration and its shear, the only non-zero off-diagonal component of the fluid stress is given by $\sigma_{xy}(y)$. Figure 2.11(b) reports kinetic, potential

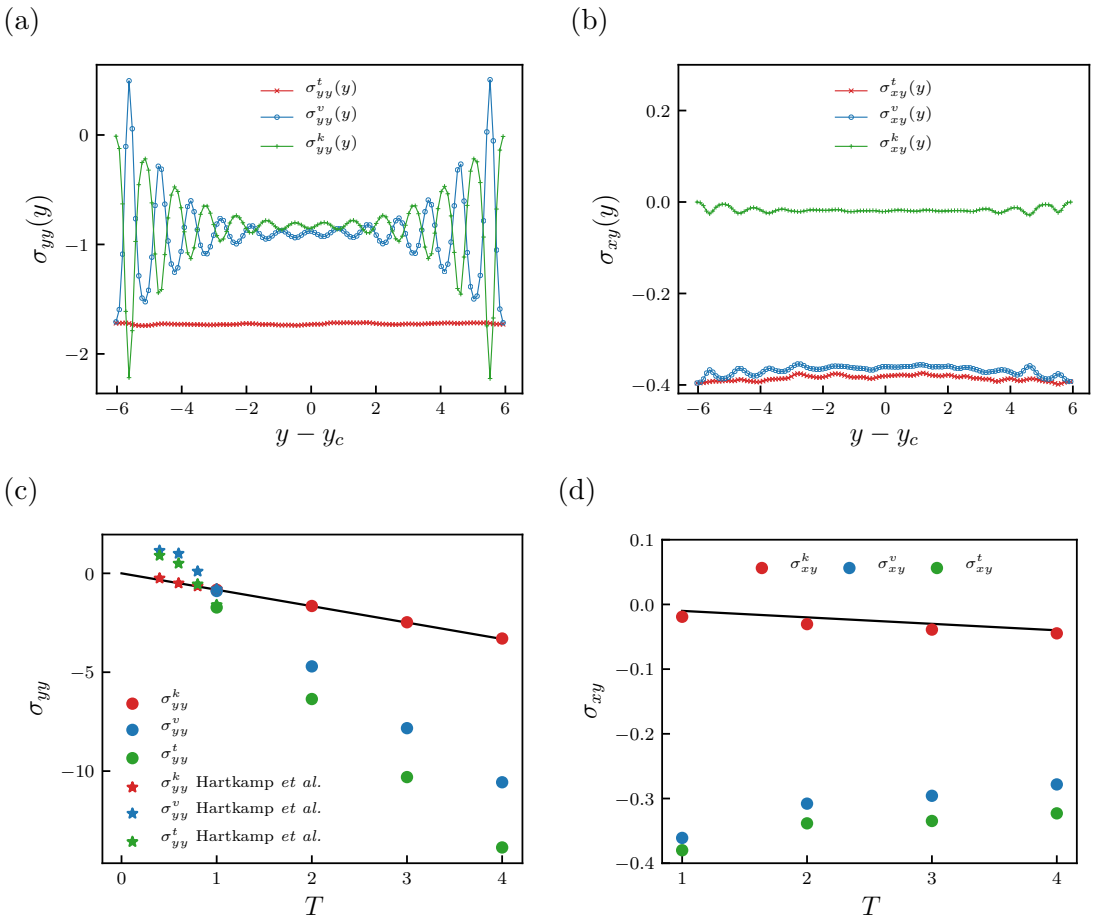

Figure 2.11. (a) Diagonal component σ^t_{yy} and (b) off-diagonal component σ^t_{xy} of the stress tensor evaluated for $T = 1.0$ and $\epsilon_{\mathrm{LJ},f,w} = 1.0$. The bulk values of σ^t_{yy} and σ^t_{xy} evaluated at different temperatures for $\epsilon_{\mathrm{LJ},f,w} = 1.0$ are reported in (c) and (d) respectively. For the sake of completeness and for validation purpose, we also report additional data from Hartkamp *et al.* [26] for $T \leq 1$. The linear dependencies between the kinetic contributions and the temperature are highlighted by the black line.

and total shear stress profiles $\sigma_{xy}(y)$ for $T = 1.0$ and $\epsilon_{\mathrm{LJ},f,w} = 1.0$. Similarly to the normal stress component $\sigma_{yy}(y)$, the total shear stress is approximately constant since its potential and kinetic components have identical oscillations with phase π.

In Fig. 2.11(c) we report the bulk values of σ_{yy} obtained with our MD simulations (at $\epsilon_{\mathrm{LJ},f,w}1.0$) as function of temperature. For the sake of consistency, we also include the bulk values of σ_{yy} obtained by Hartkamp *et al.* [26] in a different range of temperatures ($0.4 \leq T \leq 1.0$) for a system with similar characteristics. First, it is easy to verify that, as expected from equipartition theorem, the kinetic component of σ_{yy} is proportional to

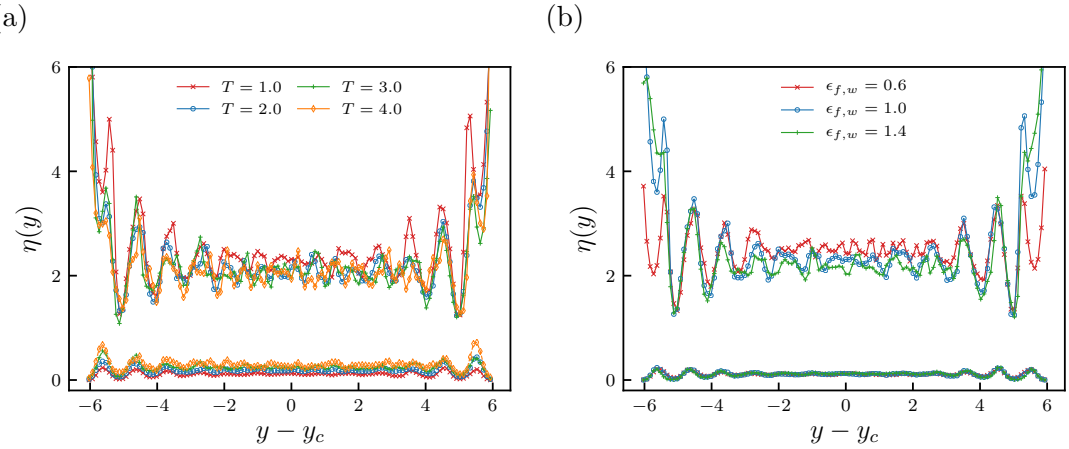

Figure 2.12. Solid lines represent the total shear viscosity profile (a) at different temperatures of the domain, namely $1.0 \leq T \leq 4.0$, for $\epsilon_{\mathrm{LJ},f,w} = 1.0$ and (b) at various wall interaction potentials, i.e. $\epsilon_{\mathrm{LJ},f,w} = 0.6, 1.0, 1.4$ and $T = 1$. The kinetic viscosity profile in (a) and (b) is reported in dashed lines.

the temperature, and more exactly is equal to $\sigma^k_{yy} = -\rho_b T/m$. Physically this means that at higher temperatures particle thermal velocities increase and, thus, more normal momentum is transferred by means of collisions between them. As regards the potential component σ^v_{yy}, its dependency on the temperature is related to the non-linear inter-particles potential. However, it can be noticed that σ^v_{yy} also increases (in absolute value) with the temperature. The reason for this may be found in the fact that higher thermal velocities cause particles to interact repulsively on average at lower distances. Such increase in repulsions lead to more negative normal stress (or, equivalently, to more positive pressure).

Some different behaviours can be detected when analysing the bulk values of the off-diagonal stress σ_{xy} with respect to the temperature (at $\epsilon_{\mathrm{LJ},f,w}1.0$). In this case, the kinetic component still increases in absolute value as a function of the temperature. However, a linear relation seems not to reproduce accurately the MD results at the lowest temperatures and a correction factor is needed, as will be discussed in Section 2.2.7. The potential component σ^v_{xy} decreases in absolute value. A reason for this is that higher thermal velocities reduce the effects of repulsive interactions between fluid layers on the shearing motion of the same layers.

2.2.7. Shear viscosity

Shear viscosity is evaluated as a function of the off-diagonal stress component σ_{xy} and of the shear rate $\dot{\gamma}$ as:

$$\eta(y) = \sigma_{xy}(y)\dot{\gamma}(y)^{-1}. \tag{2.49}$$

In analogy with the stress tensor, one can define the kinetic and potential components of the viscosity, representing the resistance to shearing flows due to particle motion and to inter-particle interactions, respectively. Specifically, the kinetic component ($\propto \rho$ for homogeneous media) is predominant at low fluid densities, while the potential contribution ($\propto \rho^2$ for homogeneous media) prevails at high densities. Figures 2.12(a)-(b) show the kinetic and the total shear viscosity profiles for several temperatures and wall interactions. The oscillatory behavior observed in the vicinity of the walls follows closely (but not exactly) the density profiles. Higher Wa corresponds to more pronounced layering in the viscosity profile, due to fluid adsorption at the solid interface. The viscosity in the bulk region ranges from 2.05 to 2.5, which can be explained by differences in bulk density and temperature of the systems. In the case that $T = 1$ and $\epsilon_{\mathrm{LJ},f,w} = 1.6$, the average density in the bulk region is $\rho_b \sim 0.81$, which is approximately the same value detected in Refs [26, 30]. Thus, under such circumstances, it is possible to compare the value of the viscosity in the bulk region against previous work. From our simulation, we observe a value $\eta_b \sim 2.18$, which is consistent with the values $\eta_b \sim 2.06$ obtained in Ref. [26] and $\eta_b \sim 2.15$ found in Ref. [30].

Analytical expressions to model the viscosity of dense inhomogeneous real fluids are not available up to date. However, for simple hard sphere (HS) systems, Din and Michaelides [25] proposed a theoretical framework that allows to derive the viscosity profile directly from the Boltzmann equation (for more details see Appendix A.1.2). In what follows, such a framework is briefly introduced and, by means of proper empirical correction factors directly obtained from our simulations, it is employed to model the viscosity of LJ fluids. The main aim of the remaining part of this section is to obtain widely applicable functional relations between viscosity and density profiles, which enables us to accurately determine the complex behavior of this transport property in confined fluids.

Analytical model for the viscosity of inhomogeneous hard-sphere fluids The theoretical derivation proposed by Din and Michaelides [25] distinguishes between the kinetic and the potential contributions to the fluid viscosity. The kinetic component is

a local function of the density, that can be written as:

$$\eta^{k}_{\mathrm{HS}}(y) = \frac{16}{5}\eta^{0}\rho(y)b^{*}(y), \tag{2.50}$$

where $\eta^{0} = 5(k_B T m/\pi)^{1/2}/(16\sigma^{2}_{\mathrm{HS}})$ is the zero-density viscosity, and $b^{*}(y)$ is a dimensionless space-dependent function that can be evaluated by solving the following integral equation:

$$\begin{aligned} &4b^{*}(y)\int_{-\sigma_{\mathrm{HS}}}^{\sigma_{\mathrm{HS}}} g(y, y+y')\rho(y+y')dy' \\ &-\frac{25}{3}\int_{-\sigma_{\mathrm{HS}}}^{\sigma_{\mathrm{HS}}} \left[b^{*}(y+y') - b^{*}(y)\right] g(y, y+y')\rho(y+y')dy' \\ &= \frac{5}{2} + \frac{\pi}{3}\int_{-\sigma_{\mathrm{HS}}}^{\sigma_{\mathrm{HS}}} g(y, y+y')\rho(y+y')dy'. \end{aligned} \tag{2.51}$$

The potential contribution to the viscosity for a HS fluid is given by a non-local expression of density:

$$\eta^{v}_{\mathrm{HS}}(y) = \frac{8\pi}{5}\eta^{0}\int_{-\sigma_{\mathrm{HS}}}^{\sigma_{\mathrm{HS}}} y'\ sin^{3}\left(\arccos\left(y'\right)\right)\ I(y, y')\ dy', \tag{2.52}$$

where $I(y, y')$ denotes the following infinite series:

$$\begin{aligned} I(y, y') = y'\sum_{s=1}^{\infty}\frac{(-y'^{s-1})}{s!}\frac{\partial^{s-1}}{\partial^{s-1}y}\Big\{ & g(y, y+y')\rho(y)\rho(y+y') \\ & \times \left[2 + b^{*}(y+y') + b^{*}(y)\right]\Big\}. \end{aligned} \tag{2.53}$$

where g is the pair correlation function.

Among the possible approximations for the pair correlation function of an inhomogeneous fluid $g(y, y+y')$, the one proposed by Fischer *et al.* [124] is adopted here. These authors approximated $g(y, y+y')$ as the radial correlation function of an homogeneous fluid evaluated at an average density $\bar{\rho}(y, y+y')$, namely:

$$g(y, y+y') = g^{hom}[r = \sigma_{\mathrm{HS}}, \bar{\rho}(y, y+y')], \tag{2.54}$$

where $\bar{\rho}(y, y+y')$ is the density average over a sphere centered in $y + \frac{y'}{2}$, with radius

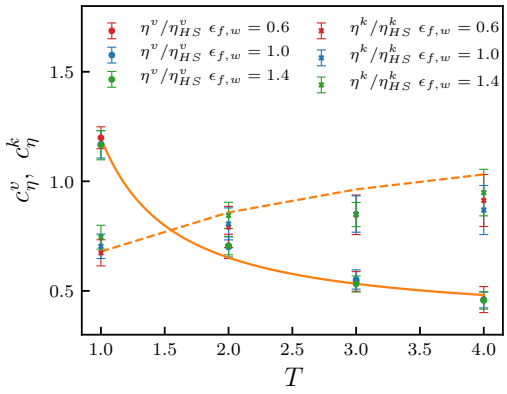

Figure 2.13. Correction factors $c^v_\eta(T)$ (solid line) and $c^k_\eta(T)$(dashed line) for the bulk values of potential and kinetic viscosity respectively. The correction factors are compared against the ratio η/η_{HS} computed directly from our MD results. The error bars indicates the standard errors computed from the distribution of viscosity values inside the bulk region.

$\sigma_{\text{HS}}/2$, i.e.,

$$\bar{\rho}(y, y+y') = \frac{6}{\pi\sigma^3_{\text{HS}}} \int_{\mathbf{r}<\frac{\sigma_{\text{HS}}}{2}} \rho(y + \frac{y'}{2} + \mathbf{r}) d\mathbf{r}. \tag{2.55}$$

Moreover, if the Carnahan-Starling equation of state is adopted [125], then the radial correlation function of an homogeneous fluid can be computed as $g^{hom}[r = \sigma_{\text{HS}},\ \bar{\rho}] = (1-0.5\xi)/(1-\xi)^3$, with $\xi = \pi\sigma^3_{\text{HS}}\bar{\rho}/6$. It is clear from Eqs. (2.52)-(2.53) that, in general, the potential contribution to the viscosity, unlike the kinetic one, cannot be expressed as a local function of the density. However, the effect of the non-locality may be neglected under certain system conditions, as will be shown below.

Shear viscosity of bulk fluid: from Hard-Sphere to Lennard-Jones When using kinetic theory to model a fluid characterized by complex interaction potentials, such as LJ, the temperature dependency of the viscosity is often not accounted properly. To overcome this issue, similarly to Refs [112, 126], we employ some empirical correction factors obtained by comparing the viscosity in the bulk region computed from MD simulations with the HS viscosity given by Eqs (2.50)-(2.52).

The kinetic viscosity in the bulk region can be described by the following expression:

$$\eta^k \sim c^k_\eta(T)\ \eta^k_{\text{HS}}, \tag{2.56}$$

where $c_\eta^k(T)$ is a correction parameter. This parameter can be expressed as $c_\eta^k(T) \sim 1/\Omega^{(2,2)}(T)$, where $\Omega^{(2,2)}$ is the factor introduced by Chapman and Enskog to obtain the viscosity of a LJ dilute gas with the kinetic theory. The value of $\Omega^{(2,2)}(T)$ is difficult to compute directly, and for this reason is typically either reported in tables [127] or provided in form of empirical expressions [128]. Here, $\Omega^{(2,2)}(T)$ is evaluated according to the expression derived by Neufeld [128], that reads:

$$\Omega^{(2,2)}(T) = \frac{A}{T^B} + \frac{C}{e^{DT}} + \frac{D}{e^{FT}} + RT^B \sin\big(ST^W - P\big), \tag{2.57}$$

with A, B ,C ,D ,E, F, R, S, W and P being real numbers.

As regards the potential contribution to the viscosity, it depends on the specific inter-particle interaction potential. In the case of LJ systems, the potential viscosity exhibits an Arrhenius-like dependence on temperature [126], that is not captured by the HS model. Thus, we adjust the HS potential viscosity as:

$$\eta^v \sim c_\eta^v(T)\ \eta_{\mathrm{HS}}^v, \tag{2.58}$$

where $c_\eta^v(T)$ is a correction factor. We model $c_\eta^v(T)$ as function of the temperature by means of our MD simulation results as $c_\eta^v(T) = Ce^{-D/T}$, with $C = 0.35$ and $D = 1.2$. Figure 2.13 shows a comparison between the correction factors $c_\eta^v(T)$ ($c_\eta^k(T)$) and the ratio $\eta^v/\eta_{\mathrm{HS}}^v$ ($\eta^k/\eta_{\mathrm{HS}}^k$) evaluated with MD simulations. The correction factors are able to approximate the values obtained with MD with an accuracy lower than the standard error for almost any value of the temperature considered in this work [1].

Non-local density-viscosity relation If we consider only the zeroth-order terms of Eq. (2.53), the following approximate expression for the viscosity is obtained:

$$\eta(y) = \frac{16\eta^0}{5\Omega}\rho(y)b^*(y) + \frac{8\pi c_\eta^v \eta^0}{5}\rho(y)\int_{-\sigma_{\mathrm{LJ}}}^{\sigma_{\mathrm{LJ}}} K(y,y')\ dy', \tag{2.59}$$

where $K(y,y')$ is a kernel function responsible for the nonlocal contributions to the potential component:

$$\begin{aligned} K(y,y') =& y'^2 \sin^3\left(\arccos(y')\right) g(y, y+y')\rho(y+y') \\ &\times \left[2 + b^*(y+y') + b^*(y)\right]. \end{aligned} \tag{2.60}$$

[1] It is worth underlining that, as implicitly assumed in Ref. [25], the HS diameter σ_{HS} is assumed to be equivalent to the LJ parameter σ_{LJ}.

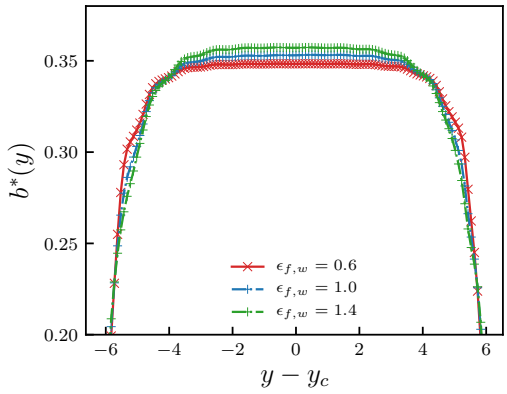

Figure 2.14. Function $b^*(y)$ computed numerically from MD density profiles through Eq. (2.51) for different wall interaction potentials, i.e. $\epsilon_{\mathrm{LJ},f,w} = 0.6$ (red solid line), $\epsilon_{\mathrm{LJ},f,w} = 1.0$ (blue dashed line) and $\epsilon_{\mathrm{LJ},f,w} = 1.4$ (green dashed line), at $T = 1$. It can be noticed that for stronger wall interactions, $b^*(y)$ has higher bulk values and shows a slower decay at the wall interface.

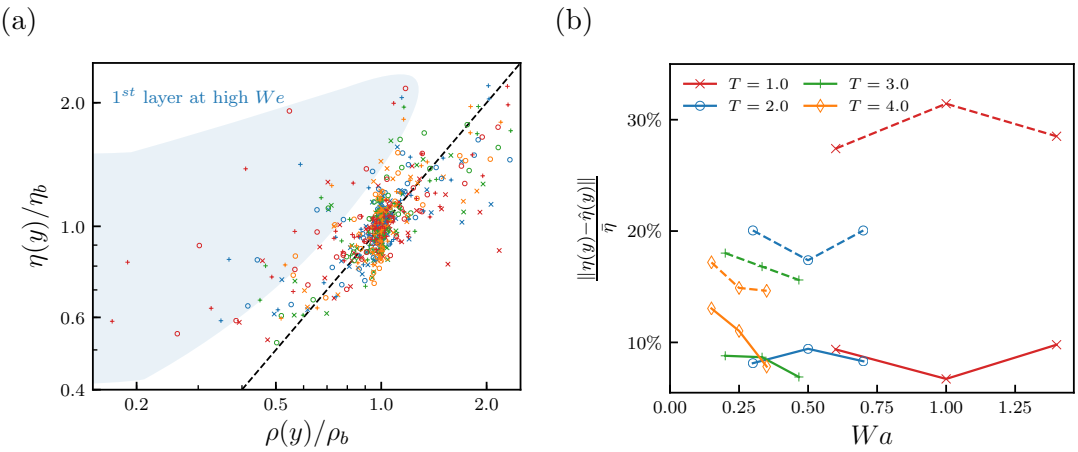

Figure 2.15. (a) Normalized viscosity vs. normalized density profiles for liquid-solid interfaces obtained by MD (symbols) and by Eq. (2.67) (dashed black line). The symbols correspond to different wall potentials (×: $\epsilon_{\mathrm{LJ},f,w} = 0.6$, ○: $\epsilon_{\mathrm{LJ},f,w} = 1.0$ and +: $\epsilon_{\mathrm{LJ},f,w} = 1.4$), whereas distinct colours are used to map the temperatures (red, blue, green and orange indicate $T = 1.0$, $T = 2.0$, $T = 3.0$ and $T = 4.0$, respectively). The points inside the shaded blue area correspond to the viscosity and density of the first fluid layer in contact with the walls for systems at high Wa. (b) Root mean square error normalized with respect to the average channel viscosity when employing Eq. (2.67) to model the viscosity of the whole system (dotted lines) and of the system without the first fluid layer at the walls (solid lines).

Equation (2.59) represents a semi-empirical nonlocal viscosity model, consistent with the kinetic theory and similar in form to local average density models presented in previous works [27, 112, 129]. In this work, we compare our model against the one by Hoang *et al.* [27], showing that the latter is inappropriate for strongly layered fluids. Specifically, by fitting numerous MD simulation data, Hoang *et al.* [27] proposed the following empirical non-local model for the viscosity in a narrow channel:

$$\eta(y) = \frac{\eta^0}{\Omega}\left(\frac{\rho(y)}{\rho_b}\right)^{\gamma} + \eta_{corr}\left[T(y), \rho_{\text{eff}}(y)\right], \tag{2.61}$$

where $\gamma = 0.8e^{0.022T^2}/\rho_b$ and $\rho_{\text{eff}}(y) = \int \omega\left(\mid y - y' \mid\right)\rho(y')dy'$ is a non-local average density with kernel $\omega\left(y\right)$ given by:

$$\omega\left(y\right) = \begin{cases} \frac{6}{\sigma_{\text{LJ}}^3}\left[\left(\frac{\sigma_{\text{LJ}}}{2}\right)^2 - y^2\right] & \text{if} \quad y \leq \frac{\sigma_{\text{LJ}}}{2} \\ 0 \quad \text{otherwise} \end{cases} \tag{2.62}$$

Furthermore, $\eta_{corr}\left[T, \rho_{\text{eff}}(y)\right]$ in Eq. (2.61) takes the following functional expression:

$$\begin{aligned} \eta_{corr}\left[T, \rho_{\text{eff}}(y)\right] =& b_1\left(e^{b_2\ \rho_{\text{eff}}} - 1\right) + b_3\left(e^{b_4\ \rho_{\text{eff}}} - 1\right) \\ &+ \frac{b_5}{T^2}\left(e^{b_6\ \rho_{\text{eff}}} - 1\right), \end{aligned} \tag{2.63}$$

with fitting coefficients b_i given in the study of Galliero *et al.* [130].

Local density-viscosity models The complexity of the previously proposed non-local viscosity model can limit its practicality, for example when limited computational resources are available. Here, we make three assumptions to simplify the model and obtain purely local expressions, which do not require the computation of space convolutions. First, the pair correlation function in the confined region is assumed to not differ significantly from its bulk counterpart, namely $g(y, y + y') \sim g_b(\sigma_{\text{LJ}})$. Second, $b^*(y)$ is assumed to be approximately constant between $y - \sigma_{\text{LJ}}$ and $y + \sigma_{\text{LJ}}$, i.e. $b^*(y + y') \sim b^*(y)$ in Eq. (2.60). Finally, we assume that the value of the integral $\int_{-\sigma_{\text{LJ}}}^{\sigma_{\text{LJ}}} \sin^3\left(\arccos(y')\right)\rho(y + y')\ dy'$ is only weakly dependent on y. Consequently, Eq. (2.59) simplifies to:

$$\eta(y) \simeq \frac{\rho(y)}{\rho_b}\left[\frac{b^*(y)}{b_b^*}\eta_b^k + \frac{1 + b^*(y)}{1 + b_b^*}\eta_b^v\right], \tag{2.64}$$

where η_b, ρ_b and b_b^* are shear viscosity, density and $b^*(y)$ evaluated in the bulk region, respectively. Equation (2.64) provides a local functional expression relating shear viscosity to the density variations at the fluid-solid interface. The analysis of our MD results shows that the ratio $b^*(y)/b_b^*$ mainly depends on the wall interaction, and not on the temperature. The profile of $b^*(y)$ as function of the wall interaction strength is shown in Fig. 2.14. By analysing the MD data, we found that the ratio $b^*(y)/b_b^*$ can be well approximated without using Eq. (2.51), by fitting the following empirical expression with MD simulation results:

$$\frac{b^*(y)}{b_b^*} \simeq 1 - A_b \; e^{B_b(y-y_w)/\epsilon_{\mathrm{LJ},f,w}^{C_b}}, \tag{2.65}$$

where $y - y_w$ is the distance from the wall and the fitting coefficients take the value $A_b = 0.59$, $B_b = 1.61$ and $C_b = 0.15$. Equation (2.65) can be used to have a straightforward evaluation of $b^*(y)$, given that its bulk value b_b^* can be computed from Eq. 2.51 as:

$$b_b^* = \frac{\pi}{12} + \frac{5}{16 \; \sigma_{\mathrm{LJ}} \; \rho_b \; g^{hom}[r = \sigma_{\mathrm{LJ}}, \rho_b]}. \tag{2.66}$$

Assuming that $b^*(y) \sim b_b^*$, Eq. (2.64) reduces to the linear relation:

$$\eta(y) \sim \frac{\eta_b}{\rho_b} \; \rho(y), \tag{2.67}$$

observed also by Morciano *et al.* [30] by analyzing MD data correlations at a fixed temperature $T = 1$. Interestingly, Fig. 2.15(a) shows that the linear relation in Eq. (2.67) can be directly detected from MD results and holds for any analysed temperature. Figure 2.15(b) shows that the corresponding standard error associated with this linear model ranges between 15% and 30% if the first fluid layer in contact with the walls is included and between 5% and 15% if the first fluid layer at the walls is excluded. Eq. (2.67) provides an attractive approximation as it does not require computing the convolution kernel or $b^*(y)$, but its applicability is limited to moderate fluid-wall interaction (low Wa), or to situations in which the first fluid layer in contact with the walls is excluded.

Comparison between the different models In Fig. 2.16, we compare the performance of the proposed viscosity models in predicting MD results for several system conditions, i.e. for varying $\epsilon_{\mathrm{LJ},f,w}$ and T. We also report the viscosity profile predicted by the model recently introduced by Hoang *et al.* [27] and given by Eq. (2.61). As regards the bulk region, the qualitative features of the shear viscosity are well reproduced both by our approach and by the data-driven approach of Refs [27, 130]. On the contrary, for

the inhomogeneous region in proximity of the walls, two regimes may be distinguished. At Wa $\ll 1$, a weak layering is observed in the viscosity profile and non-local effects are negligible. In this regime, both our local models and Hoang *et al.* [27]'s model, which is based on a correlation derived for uniform fluids, are able to accurately reproduce MD results. However, when the Wall number increases, wall effects become more and more important, and models such as Hoang *et al.* [27]'s one fail dramatically to predict the viscosity profile. The models proposed in this work, including the local ones, show a surprisingly good agreement even in case of Wa ~ 1. As expected from the derivation, the nonlocal model outperforms the local ones. Furthermore, when looking at the first fluid layer in contact with the walls, the linear model of Eq. (2.64) is more accurate than the simplified linear model in Eq. (2.67).

Darcy-Weisbach friction factor in nano-pipes The findings of this work can be applied to study, for instance, the Darcy-Weisbach friction factor in nano-geometries. As extensively discussed in the work by Liakopoulos *et al.* [131], it is possible to use a modification of the Darcy-Weisbach friction factor to model flows in nanoscale confined geometries. For an infinitely wide channel with height h, the Darcy-Weisbach friction factor, f, is given by:

$$f = \frac{-2D\dfrac{\partial p}{\partial x}}{\bar{\rho}\bar{V}^2} \tag{2.68}$$

with p being the pressure, $D = 2h$ the hydraulic diameter, and $\bar{\rho} = \frac{1}{h}\int_0^h \rho(y)$ and $\bar{V} = \frac{1}{h}\int_0^h u_x(y)$ the average density and velocity in the cross-section, respectively. From momentum conservation, assuming a Poiseuille flow, the local velocity can be computed by solving the following differential equation:

$$\frac{d}{dy}\left(\eta(y)\frac{du_x(y)}{dy}\right) = \frac{\partial p}{\partial x}. \tag{2.69}$$

For symmetric flows, Eq. (2.69) has the solution:

$$u_x(y) = \frac{\partial p}{\partial x}\int_0^y \frac{y'}{\eta(y')}dy' + u_s \tag{2.70}$$

where $u_s = u_x|_{y=0}$ is the slip velocity at the wall. The integrals in Eq. (2.70) can be solved by using one of our (local or non-local) models to approximate $\eta(y)$. For simplicity, here we will use the local model in Eq. (2.67), equipped with the density

profile approximation in Eq. (2.42). After some algebraic manipulations, we obtain the following expression for the Darcy-Weisbach friction factor in nano-channels:

$$f = \frac{-2D\dfrac{\partial p}{\partial x}}{\dfrac{\rho_b}{h}\int_0^h \left(1 + Ae^{-By}\cos\frac{2\pi y}{l}\right) dy \left[\dfrac{1}{h\eta_b}\dfrac{\partial p}{\partial x}\int_0^h \int_0^y y' \left(1 + Ae^{-By'}\cos\frac{2\pi y'}{l}\right)^{-1} dy'dy + u_s\right]^2}. \tag{2.71}$$

The more accurate non-local viscosity model in Eq. (2.59) can also be employed, but this would increase the computational cost. Moreover, this approach can be generalized to study nano-channel with arbitrary shapes.

The functional relations developed here can be implemented in numerical methods dealing with mesoscopic fluid flows (e.g. Ref [132]) to improve the accuracy of such methods in the description of fluid-solid interfaces. But also models in the framework of dynamic density-functional theory (e.g. Refs [50, 56, 63]), which suffer from unresolved closures, such as the dependence of viscosity on density. Moreover, since the functional relations are derived from a general theoretical framework for HS systems, they can be employed to model the inhomogeneous properties of non-LJ fluids, by appropriately adjusting the empirical factors.

In the next Section, we explore characteristics and properties of liquid-vapor and liquid-liquid interfaces of multi-phase fluids in confined systems, where the treatment of the stress tensor and emergence of surface tension is crucial (e.g. Ref. [133]).

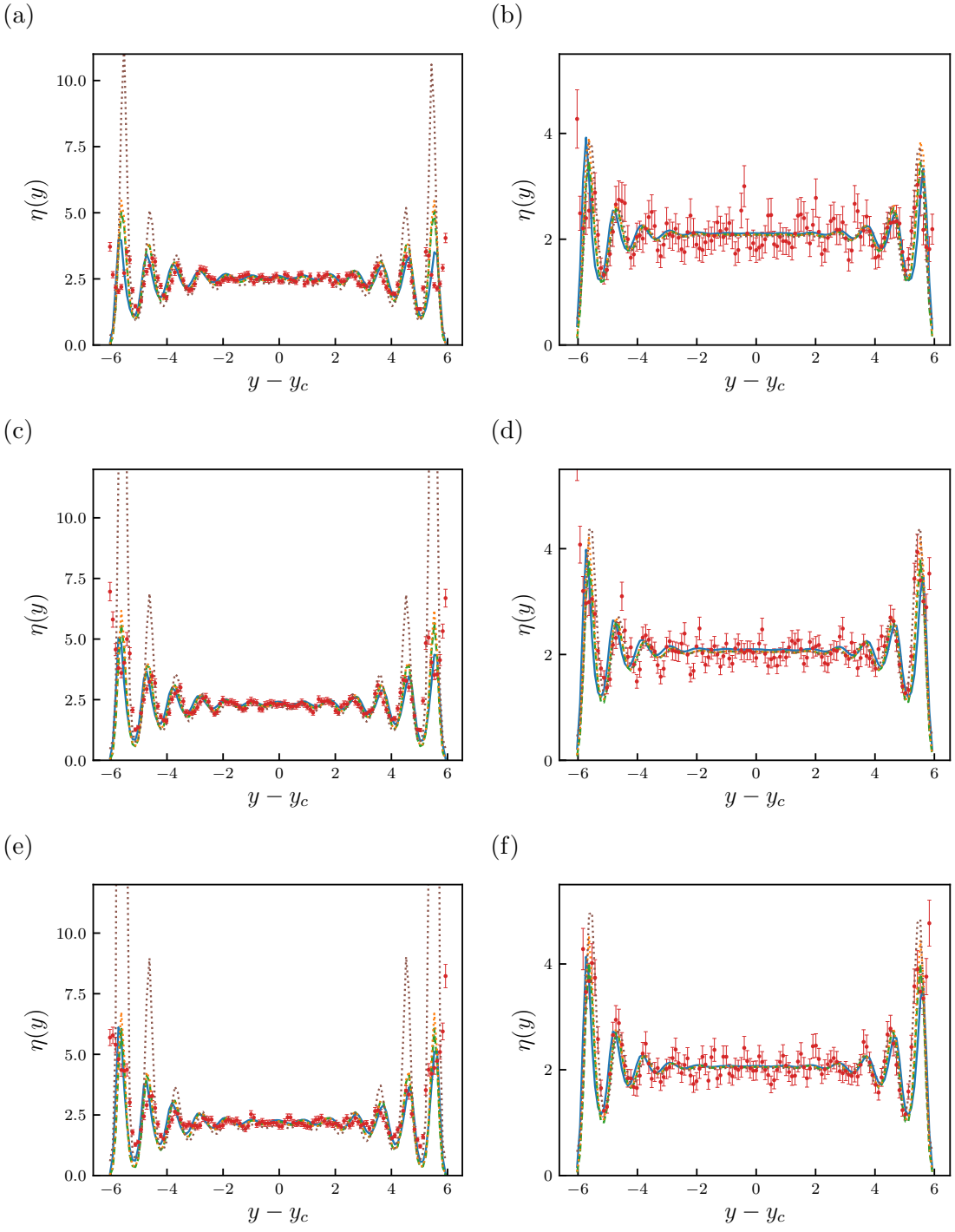

Figure 2.16. Shear viscosity profile for $\epsilon_{\mathrm{LJ},f,w} = 0.6$ (a-b), $\epsilon_{\mathrm{LJ},f,w} = 1.0$ (c-d) and $\epsilon_{\mathrm{LJ},f,w} = 1.4$ (e-f) at temperatures $T = 1.0$ (a-c-e) and $T = 4.0$ (b-d-f). MD results (red dots) are compared against the non-local model given by Eq. (2.59) (solid blue line), the linear model of Eq. (2.64) (dashed green line), the simplified linear model in Eq. (2.67) (dotted orange line), and, finally, the model proposed by Hoang *et al.* [27] (Eq. (2.61)).

2.3. MD simulations of multi-phase fluids in confined geometries

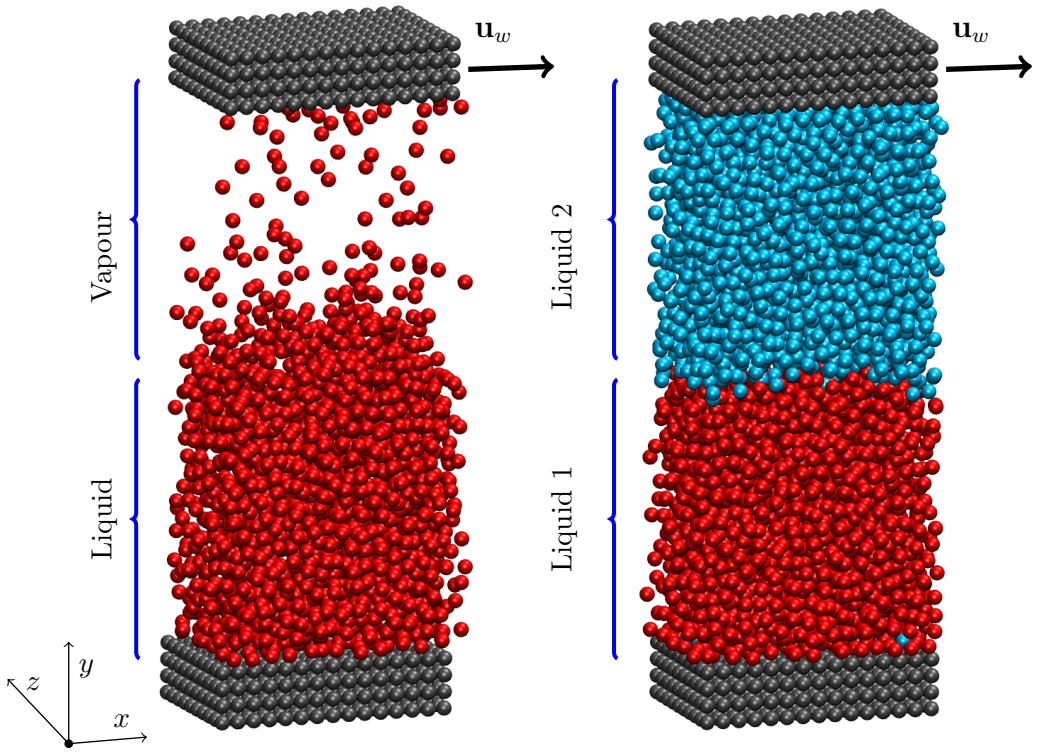

Figure 2.17. MD setup of a liquid-vapour system on the left (a) and of a system with two partially miscible fluids on the right (b).

In widely adopted methods such as level-set [31] and volume of fluid [32], the variations of density and viscosity at the interface between two fluids characterized by bulk densities ρ_1 and ρ_2 and bulk shear viscosities η_1 and η_2, take the form:

$$\begin{cases} \rho(\mathbf{r}) = & \rho_1 + (\rho_2 - \rho_1)\, f(\mathbf{r}) \\ \eta(\mathbf{r}) = & \eta_1 + (\eta_2 - \eta_1)\, f(\mathbf{r}) \end{cases} \tag{2.72}$$

where $f(\mathbf{r})$ is a space dependent function. Specifically, in level-set method $f(\mathbf{r})$ takes the form $f(\mathbf{r}) = H_\epsilon(\phi(\mathbf{r}))$, with $\phi(\mathbf{r})$ being a level set function and H_ϵ a smoothed step function, while in volume of fluid method $f(\mathbf{r}) = V_2(\mathbf{r})/(V_1(\mathbf{r}) + V_2(\mathbf{r}))$, with $V_1(\mathbf{r})$ and $V_2(\mathbf{r})$ being the local volumes of the two fluids. In both cases, it is commonly assumed that the local viscosity has a linear functional dependence on the local density at the

interface, namely

$$\frac{\eta(\mathbf{r}) - \eta_1}{\eta_2 - \eta_1} = \frac{\rho(\mathbf{r}) - \rho_1}{\rho_2 - \rho_1}, \tag{2.73}$$

but this assumption may not always be valid.

Experiments at the nanoscale are extremely difficult and are hampered by errors [4, 134], while continuum models based on Navier-Stokes equations are not suitable to investigate molecular-scale phenomena which are characterized by the discrete behaviour of the matter. Among the different numerical techniques that could be used to investigate fluid interfaces, molecular dynamics (MD) offers an attractive alternative in view of these difficulties and indeed it is often the computational tool of choice in the nanoscale regime. In fact, MD allows us to extract the local properties at the interfaces directly from the coordinates of fluid particles.

Molecular dynamics has already been extensively used for the investigation of liquid-vapour interfaces at equilibrium for Lennard-Jones fluids [135, 136, 137, 138], water [139, 140], as well as for more complex fluids [141, 142]. Some research studies on liquid-vapour interfaces in out-of-equilibrium conditions have also been previously conducted, e.g. analysing fluid evaporation [143] or of water cluster formation [28, 29].

Multiple research studies have also analysed the interfaces of immiscible fluids by means of MD simulations or density functional theory [134, 144, 145, 146, 147]. For instance, Bohn *et al.* [145] analysed the structure and the dynamical properties of a liquid-liquid interface by employing MD simulations. Afterwards, Goel *et al.* [148] proposed a methodology to compute interfacial tension of complex two-phase systems in the framework of Dissipative Particle Dynamics. Later, Braga *et al.* [149] used non-equilibrium MD to investigate the diffusion at liquid-liquid interfaces. Recently, the fundamental role of partially miscible fluids interfaces in the field of ionic liquids have been investigated in multiple works, i.e. Refs [150, 151, 152]. Moreover, a theoretical approach to the study of the density profile at the interface has been proposed by Iatsevitch *et al.* [153]. However, long-standing issues remain to be solved, especially in the investigation of transport properties [145, 149].

In this section, non-equilibrium MD simulations of Lennard-Jones fluids are used to explore and model the variation of the main physical properties at the interfaces between two fluids. Our final aim is to detect existing relations between fluid properties, mainly shear viscosity, and density variations. This approach is motivated by the recent development of mesoscopic numerical frameworks, such as Dynamical Density Functional Theory [51, 154, 155], which is characterized by unresolved closures, such as the

dependence of viscosity on density.

This section is structured as follows. In Sect. 2.3.1 we analyse the local density profiles and model the interface thickness as function of temperature and interaction parameters. Velocity, shear rate and temperature profiles are discussed in Sects 2.3.3-2.3.4. A dedicated analysis of the kinetic and the potential components of the stress tensor is provided in Sect. 2.3.5, where we also analyse surface tension. Hence, functional relations between local shear viscosity and local density at fluid-fluid interfaces are proposed in Sect. 2.3.6. As an example of applications, these models are then employed to compute the Darcy-Weisbach friction factor of multi-phase fluids in nano-geometries (Sect. 2.3.7).

2.3.1. Simulation details

We simulate two different system configurations: one composed of a liquid coexisting with its vapor, and another made of two partially miscible fluids. Both systems are confined between two solid parallel walls, as shown in Fig. 2.17. A steady-state Couette flow is generated by shearing one of the walls [113, 156]. The interaction potential U_i is the given by the sum of the interaction between pairs of fluids particles and between fluid and wall particles, i.e.

$$U(\mathbf{r}_i) = \sum_{j=1}^{N_f} U_{f,f}(\mathbf{r}_{ij}) + \sum_{j=1}^{N_w} U_{f,s}(\mathbf{r}_{ij}), \tag{2.74}$$

where N_f and N_w are the total number of fluid and wall particles, respectively. Both fluid and wall particles are modelled as mono-atomic particles with mass $m = 1$ interacting via a 12-6 Lennard-Jones potential U_{LJ}. In this work we are interested in analyzing the interfaces between liquid-vapor and liquid-liquid which are not directly affected by the fluids behaviour close to the walls. Thus we adopt solid walls, with particles fixed in their lattice positions as in Refs [26, 116, 156]. Regarding U_{LJ} parameters, for all fluid-fluid interactions we adopt the values $\epsilon_{f,f} = \epsilon$ and $\sigma_{f,f} = \sigma$. Moreover, in the LV system we employ $\epsilon_{f,w} = 0.8\epsilon$ and $\sigma_{f,w} = \sigma$ for the wall particles in contact with the liquid, while $\epsilon_{f,w} = 0.4\epsilon$ and $\sigma_{f,w} = \sigma$ for the wall particles in contact with the vapor. This ensures that the liquid phase remains in contact with the same wall throughout the simulations. For the partially miscible liquids we adopt for both fluids $\epsilon_{f,w} = 0.8\epsilon$ and $\sigma_{f,w} = \sigma$.

The simulation box (wall excluded) measures $[15\sigma] \times [33\sigma] \times [10\sigma]$. Periodic boundary conditions are imposed along x and z axes, while the fixed walls wrap the system in the y direction. The periodic boundaries along the directions parallel to the fluid interfaces and

the confinement induced by the limited cross-sectional size of the system act as stabilizer of the planar interface, limiting the formation of capillary waves, phenomena widely discussed in previous works, i.e. Refs [157, 158]. In order to reach an high accuracy in density and surface tension profiles a cut-off radius $r_c = 4.5\sigma$ is adopted (as recommended in Refs [136, 159, 160]).

One of the main purposes of this work is the analysis of shear viscosity profile, which is independent of the shearing velocity of the wall only for limited shear rates (ideally tending to zero). On the other side, in order to achieve an high signal-to-noise ratio, it is convenient to impose a relatively large velocity gradient. Taking into consideration these factors, we adopt an average shear rate of about $\dot{\gamma} = 0.02$ for the liquid-vapour system and $\dot{\gamma} = 0.05$ for the liquid-liquid system.

The motion of the walls causes heating of the flow, which reveals in a continuous energy contribution to the system. However, one of the aims of this study is to analyse the systems in isothermal conditions, for different sub-critical temperatures, i.e. $0.8 \leq T \leq 1.2$. Thus, in order to dissipate the shear energy and maintain a local constant T along the entire domain(consistently with an NVT ensemble), a Nosé-Hoover thermostat is employed [88, 89, 161].

Liquid-vapor coexistence is generated by an isothermal expansion of a uniform system with average density $\rho = 0.67$ along the z direction. In this stage, the liquid phase nucleates at the more attractive wall, and the equilibrium liquid-vapor interface is formed. The interface between the two partially miscible fluids instead is generated by simply allocating the two fluids separately within the domain.

Afterwards, both kinds of system are equilibrated for $2x10^6$ time steps. The shear velocity is then applied to the walls, and an equilibration run of $2x10^6$ time steps is used to reach steady-state conditions, characterized by the stationarity of temperature, density and velocity profiles. The physical properties of interest are computed as a function of particle positions and momenta. Finally, under the ergodic assumption (i.e. for a property 'A': $< A >_{NVT} = < A >_t$), they are averaged over 10 independent trajectories of $t = 4x10^6$ time steps.

Finite size effects

Finite size effects can affect the properties of the system at the liquid-vapour and liquid-liquid interfaces. Since our analysis is focused on understanding interface properties far from the walls, we aim to keep those effects at minimum. Figure 2.18 shows the density profiles of LV (a) and LL (b) systems for varying domain size and same average shear rate. As shown in the insets, for a small size of the system ($L_y = 20$) wall effects cause

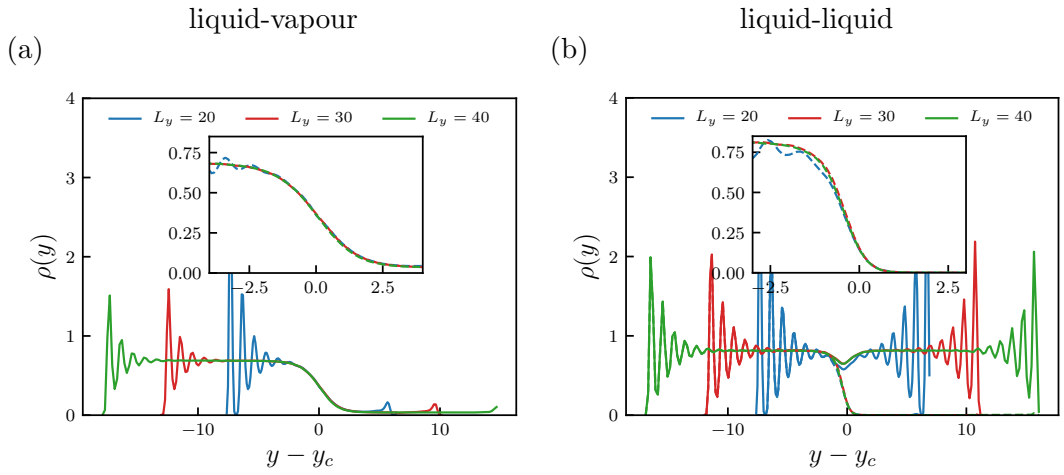

Figure 2.18. Density profiles for LV (a) and LL (b) systems for different sizes of the domain.

oscillating trends in the density profile at both LV and LL interfaces. On the contrary, the profiles obtained with $L_y = 30$ and $L_y = 40$ do not show significant differences. To be on the conservative side, we then adopt a domain length in the y-direction of $L_y = 40$ for both LV and LL systems.

2.3.2. Density

LV systems

In previous studies, i.e. Refs. [28, 140], a typical approach consists in approximating the density profile across liquid-vapour interfaces by means of an hyperbolic tangent function, namely:

$$\rho(y) = \frac{1}{2}\left(\rho^l + \rho^v\right) - \frac{1}{2}\left(\rho^l - \rho^v\right)\tanh\left(\frac{y - y_0}{d}\right), \tag{2.75}$$

where ρ^l and ρ^v are the bulk densities of liquid and vapour respectively, y_0 is the position of the Gibbs dividing interface and d is the interface thickness. In Fig. 2.19(a), we report the density profile obtained with MD simulations and the corresponding fitting curves (Eq. (2.75)) at various temperatures. The fitting above allows us to obtain the values of bulk densities and interface thicknesses at different temperatures. As shown in Fig. 2.19(b), in accordance with the Lennard-Jones fluid phase diagram [162], the difference between the bulk densities of coexisting liquid and vapour reduces for increasing temperatures. The value of the temperature for which such difference becomes zero

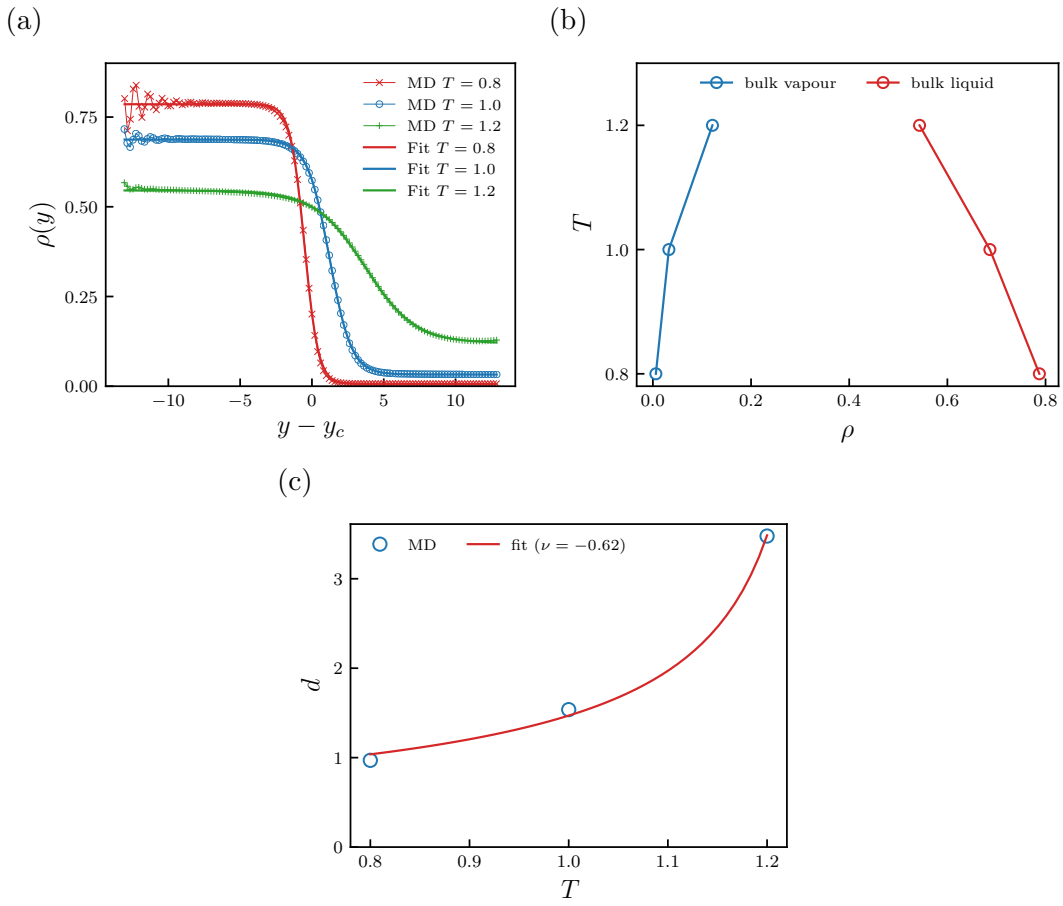

Figure 2.19. (a) LV density profiles at different temperatures, namely $0.8 \leq T \leq 1.2$. (b) Corresponding phase diagram for the finite-size system, obtained from liquid and vapor bulk densities. (c) Interface thickness d as function of the system temperature T obtained by fitting Eq. (2.75) to the MD simulation results.

is known as critical temperature. By following the approach in Ref. [163], the critical temperature T_C is approximated by fitting the expression:

$$\left(\rho^l - \rho^v\right) = A\left(1 - \frac{T}{T_C}\right)^{\beta}, \tag{2.76}$$

where A is a proportionality constant and β is the critical exponent. By fitting Eq. (2.76) with the MD data, we obtained $A \sim 1.073$, $T_C \sim 1.267$ and $\beta = 0.318$ ($R^2 = 0.999$). It is worth noting that our value of β is consistent (only 3% off) with the one ($\beta = 0.328$) reported in Ref. [163].

The values of the interface thickness obtained by fitting Eq. (2.75) with the MD

results, together with the critical temperature T_c estimated with Eq. (2.76), allows us to model the thickness as a function of the temperature with the commonly adopted expression [163]:

$$d(T) = d_0 \left(1 - \frac{T}{T_C}\right)^{-\nu}. \tag{2.77}$$

This fitting (characterized by $R^2 = 0.997$) returns the values $d_0 \sim 0.556$ and $\nu \sim 0.624$. The latter exponent is again very close to the value $\nu = 0.640$ reported in Ref. [163]. In Fig. 2.19(c) we report the interface thickness as a function of the temperature obtained with MD simulations and with Eq. (2.77). From the previous relation, it is clear that the higher the temperature, the wider the interface. Specifically, for a temperature $T = T_c$ the interface width tends to L_y and, thus, the liquid-vapour interface disappears.

LL systems

Figures 2.20(a-b) show the density profiles for one of the fluids (dashed line) and for the mixture of the two identical fluids at different temperatures (a) and different repulsive interaction parameter ϵ_{12} (b). The total density profile exhibits a well at the interface between the two fluids, which is governed by repulsive forces between the two components and temperature. Specifically, the well reduces by increasing ϵ_{12} (which corresponds to weaker repulsive interactions between the two fluids), or by increasing the temperature (which corresponds to higher thermal velocities at the interface).

Similarly to LV systems, we model the density profile for a single fluid across the liquid-vapour interface by means of a hyperbolic tangent function, namely:

$$\rho(y) = \frac{\rho^l}{2}\left(1 - \tanh\left(\frac{y - y_0}{d}\right)\right) \tag{2.78}$$

where ρ^l is the liquid bulk density, y_0 is the position of the dividing interface and d is the interface thickness.

The interface thickness, estimated by fitting Eq. (2.78) to the MD simulations data, is dependent on both ϵ_{12} and T. The relation governing the dependency of the thickness on the temperature for LV systems [163] is applied here to LL systems as well. It reads:

$$d(T) = d_0 \; e^{\epsilon_{12}-1}\left(1 - \frac{T}{T_C}\right)^{-\nu}, \tag{2.79}$$

with fitting parameters $d_0 = 0.647$, $T_c = 1.689$ and $\nu = 0.714$ ($R^2 = 0.947$) which is not far from the value $\nu = 0.640$ obtained in Ref. [163] for LV interfaces. In Fig. 2.20(c)

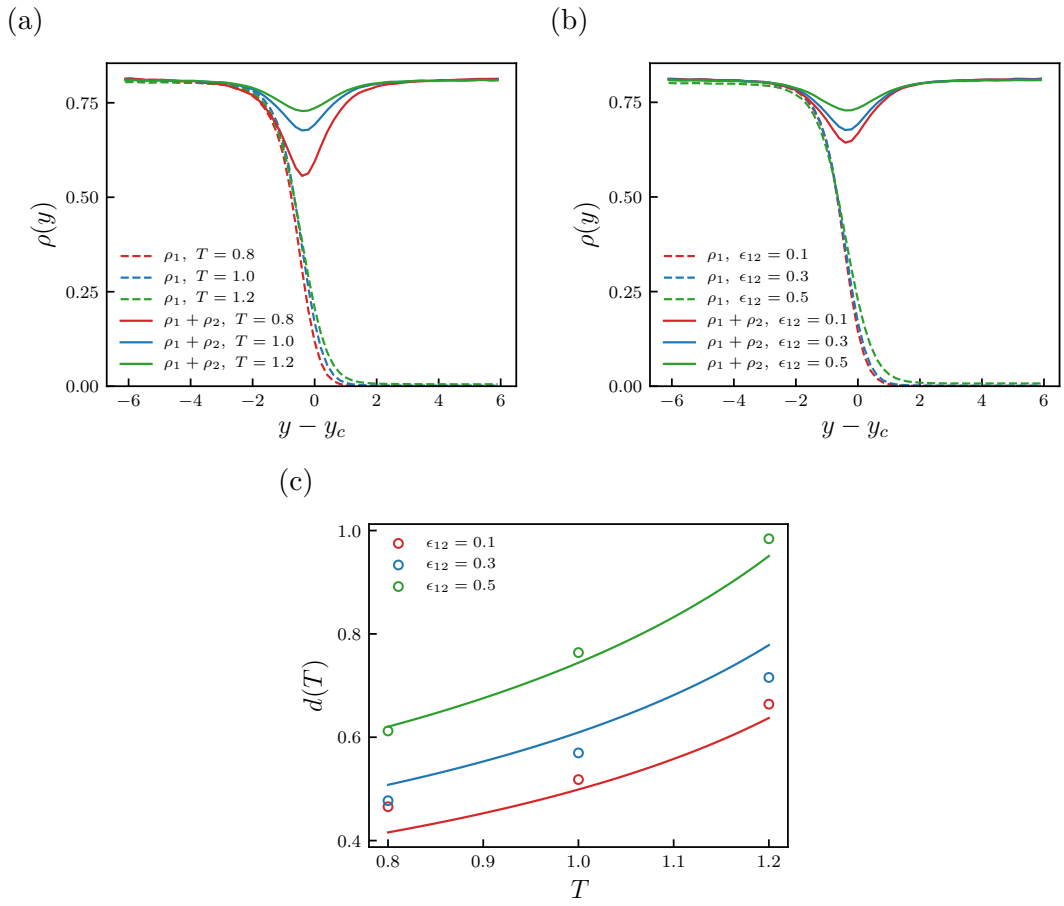

Figure 2.20. Density profiles for one of the fluids (dashed line) and for a mixture of the two fluids (solid line) obtained for different values of temperature $T = 1$ (a) and interaction parameter ϵ_{12} (b). (c) Interface thickness variation with ϵ_{12} and T^*.

we report the interface thickness obtained with MD simulations and the corresponding fitting (Eq. (2.79)).

2.3.3. Velocity and shear rate

In a system with a single homogeneous fluid, the velocity profile generated by a Couette flow is linear. However, when the system is constituted by multiple fluids or phases, the gradient of the velocity profile (i.e. the shear rate) will vary with the density. In Fig. 2.21(a) we report the velocity profile generated by the relative motion of the two walls for the LV system. For such systems the velocity gradient in the vapour is larger than in the liquid, and this gap reduces by increasing the temperature. Moreover, the

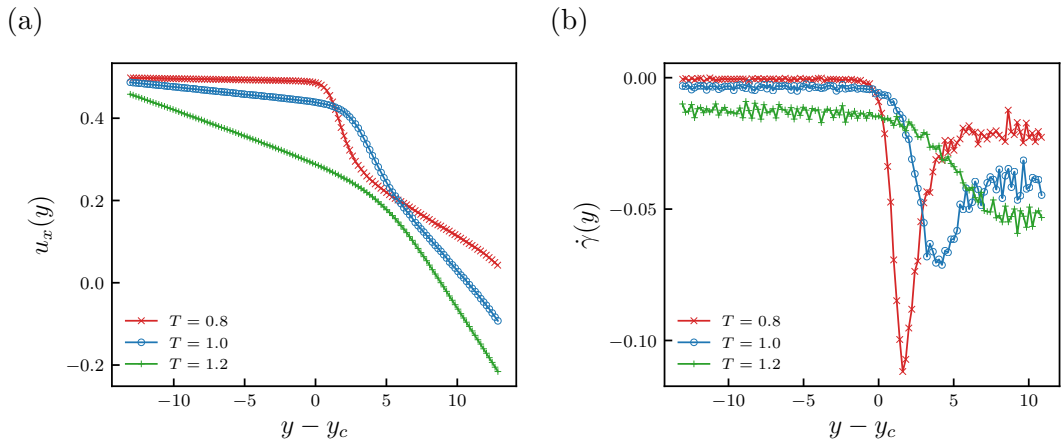

Figure 2.21. (a)Velocity profile and (b) shear rate in LV systems induced by the shearing walls at temperatures $0.8 < T < 1.2$.

sharp variation of the density profile at the interface causes a local increase (in absolute value) of the shear rate (Fig. 2.21(b)).

The velocity and shear rate profiles for the LL system is reported in Fig. 2.22. The velocity profiles exhibit the characteristic linear trend typical of a Couette flow in the bulk regions of the two fluids. However, the interface generated by the repulsive forces between the two fluids generates a zone with particles at higher mobility. This leads to a peak in the shear rate at the interface which decreases (in absolute value) with the temperature. In addition, the value of the shear rate evaluated at the interface depends significantly on the interaction parameter ϵ_{12}, which controls the intensity of the repulsive forces. This suggests that slip occurs among the two fluids, and this phenomenon scales with temperature and interaction potential.

2.3.4. Temperature and thermostat

The shear stress generated by the moving walls induces a viscous friction, which causes an increase in the system temperature, known as viscous heating. In order to ensure an experimentally representative uniform temperature across the domain, we remove the excess of thermal energy by thermostatting the fluid. Among the possible thermostatting techniques, we employ a Nosé-Hoover (NH) thermostat, which, differently from other thermostats such as Langevin and dissipative particle dynamics, does not introduce an artificial contribution to the viscosity [91]. In preliminary tests, we tested a single global NH thermostat and a series of local Nosé-Hoover thermostats similar to the one adopted in other works studying fluid properties close to walls [26, 164]. However, the

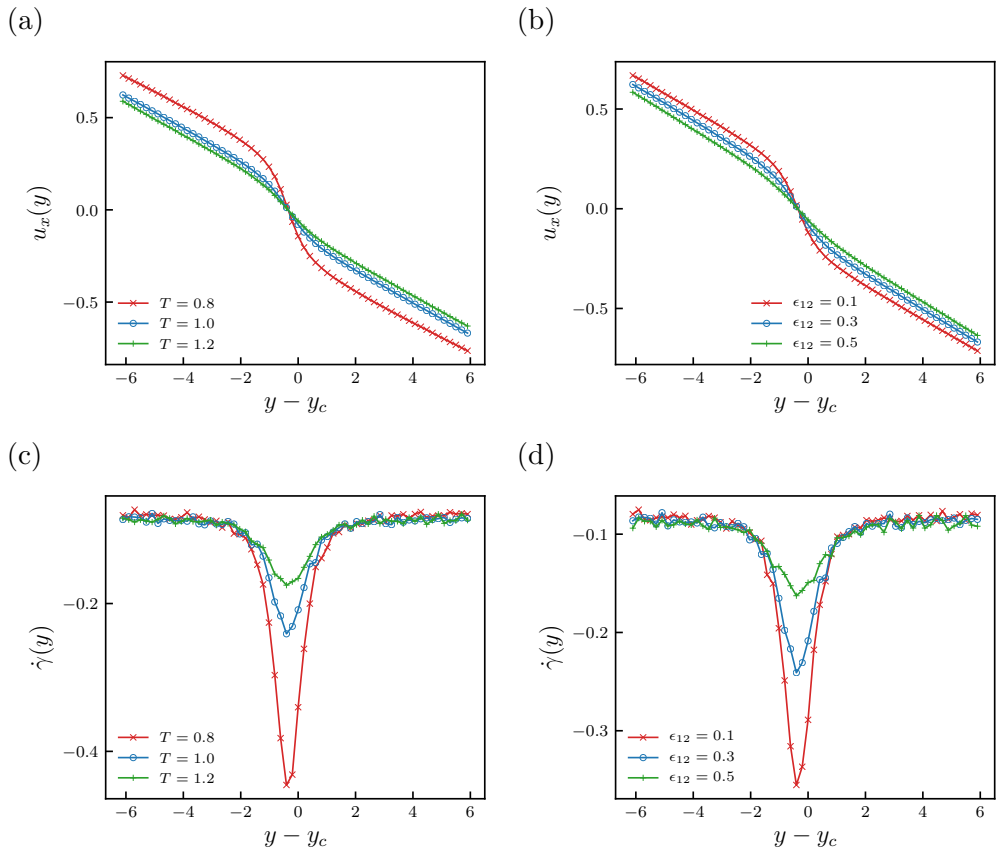

Figure 2.22. (a)Velocity profile and (c) shear rate in LL systems induced by the shearing walls at temperatures $0.8 < T < 1.2$. (b)Velocity profile and (d) shear rate in LL systems induced by the shearing walls for $0.1 \leq \epsilon_{12} \leq 0.5$ at a temperature $T = 1$.

latter strategy was artificially altering the stress tensor profiles, thus a single global NH thermostat is employed. Since such thermostatting strategy suffers strongly non-homogeneous system conditions, we adopted low values of the shear rates. In order to avoid biasing the fluid flow, the thermostat is applied only in the directions perpendicular to the flow, i.e. only in the y and z directions, and the relaxation time adopted is about $\tau = 1.0$, in line with the discussion in Ref. [165].

Figure 2.23 shows the temperature profiles obtained for LV (a) and LL (b) systems. The global thermostat is able to keep a uniform temperature in proximity of liquid interfaces at high temperatures, e.g. $T = 1.2$. At the same time, the temperature profiles exhibit non-constant trends at lower temperatures, e.g. $T = 0.8$, when the shear rate is strongly heterogeneous. However, these local deviations of the temperature from the mean one are less than 10% for LL systems and less than 15% for the LV systems,

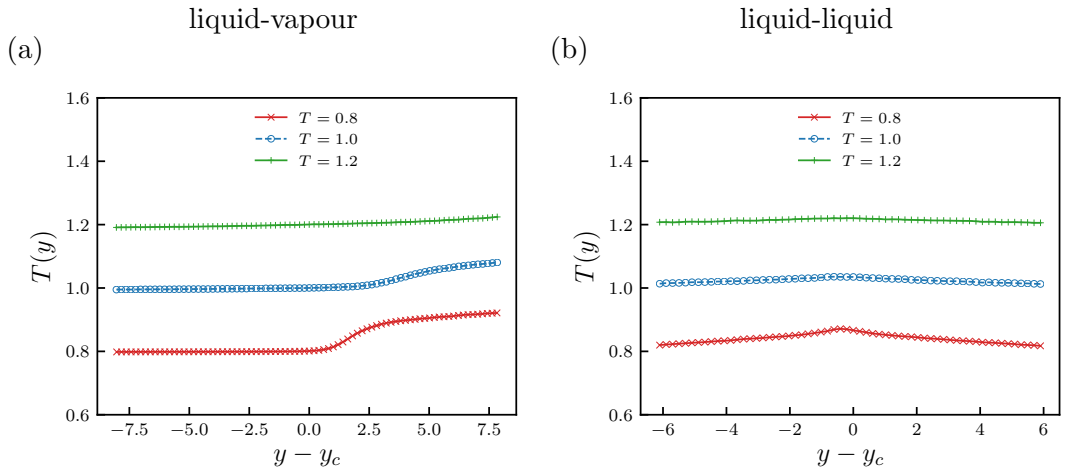

Figure 2.23. Temperature profile at different temperatures $0.8 \leq T \leq 1.2$ for (a) liquid-vapour and (b) liquid-liquid systems.

thus we do not expect significant effects in the following analysis of stress and viscosity profiles.

2.3.5. Stress tensor

As previously discussed in Sect. 2.2.6, the momentum conservation equation states that, in the simulated systems, $\sigma_{\alpha y}(y_2) = \sigma_{\alpha y}(y_1)$ for any two points y_1 and y_2 between the two walls, meaning that the profiles $\sigma_{xy}(y)$, $\sigma_{yy}(y)$ and $\sigma_{zy}(y)$ are constant throughout the phases. This is consistent with the MD results for both LV and LL systems (Fig. 2.24).

As shown in Fig. 2.25, the stress tensor components tangential to the interface $\sigma_{xx}(y)$ and $\sigma_{zz}(y)$ diverge from $\sigma_{yy}(y)$ because of surface tension effects. Based on its derivation from statistical mechanics [134, 145], the interface energy excess E_{if} can be defined as function of the diagonal stress tensor components according to Kirkwood-Buff formula, namely:

$$E_{if}(y) = \frac{\sigma_{xx}(y) + \sigma_{zz}(y)}{2} - \sigma_{yy}(y) \tag{2.80}$$

such that the surface tension γ is equal to:

$$\gamma = \frac{1}{2} \int_{-L}^{L} E_{if}(y)\, dy \tag{2.81}$$

where the interface spans between $-L$ and L. It is worth highlighting that E_{if} is due to the excess of energy introduced by the system heterogeneity into the pressure component

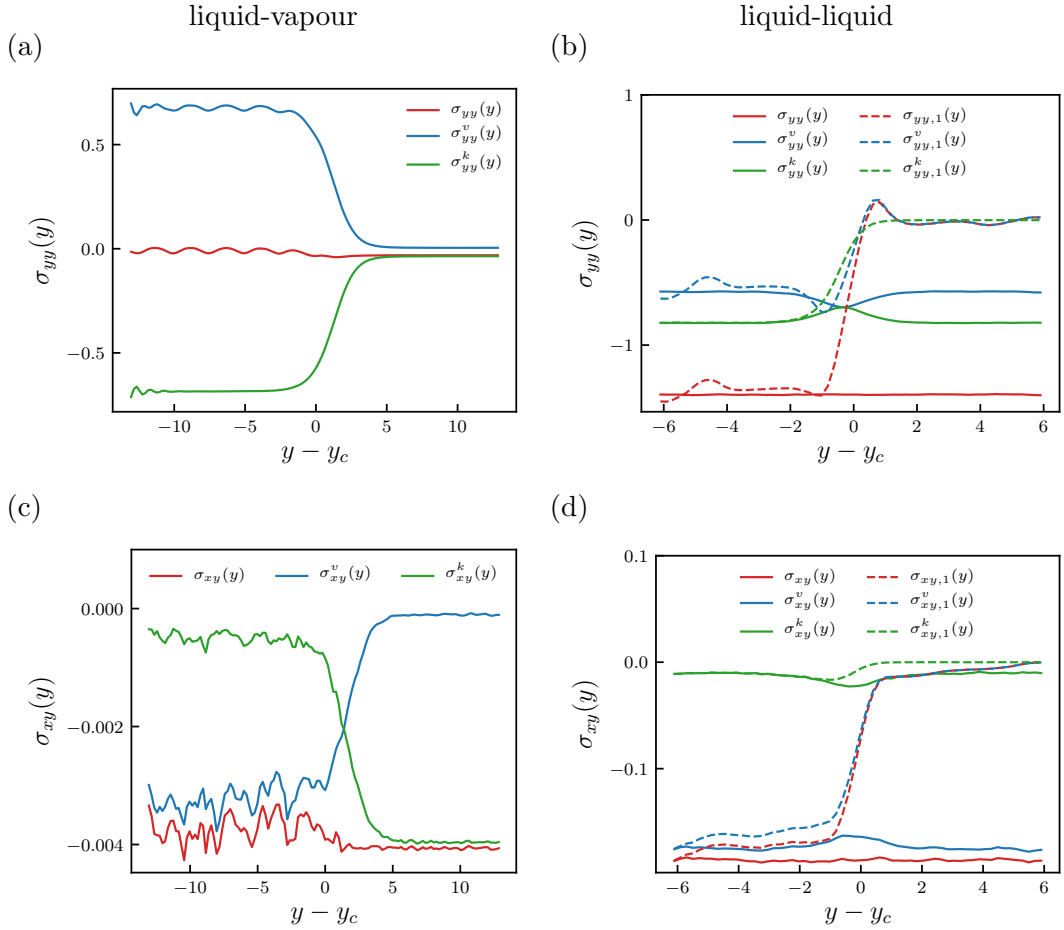

Figure 2.24. Potential and kinetic the stress tensor components σ_{yy} (a-b) and σ_{xy} (c-d) for LL (a-c) and LV (b-d) system at $T = 1$ and $\epsilon_{12} = 0.3$.

tangential to the interface, thus it is equal to zero in the bulk regions where the diagonal components of the stress tensor equal each other (Figs 2.25(a-b)).

LV systems

Even if the kinetic and potential components of $\sigma_{yy}(y)$ (in Fig. 2.24(a)) exhibit a non linear trend in the y-direction and a jump at the interfaces, their sum $\sigma_{yy}(y)$ is constant across the liquid and vapour phases, consistently with the momentum conservation. Similarly, even if the kinetic contribution to the off-diagonal stress tensor component $\sigma_{xy}(y)$ is predominant in the vapour phase, the corresponding potential contribution $\sigma^v_{xy}(y)$ is larger in the liquid phase. Specifically, this leads to a constant profile of the total stress tensor $\sigma_{xy}(y)$ (Fig. 2.24(c)).

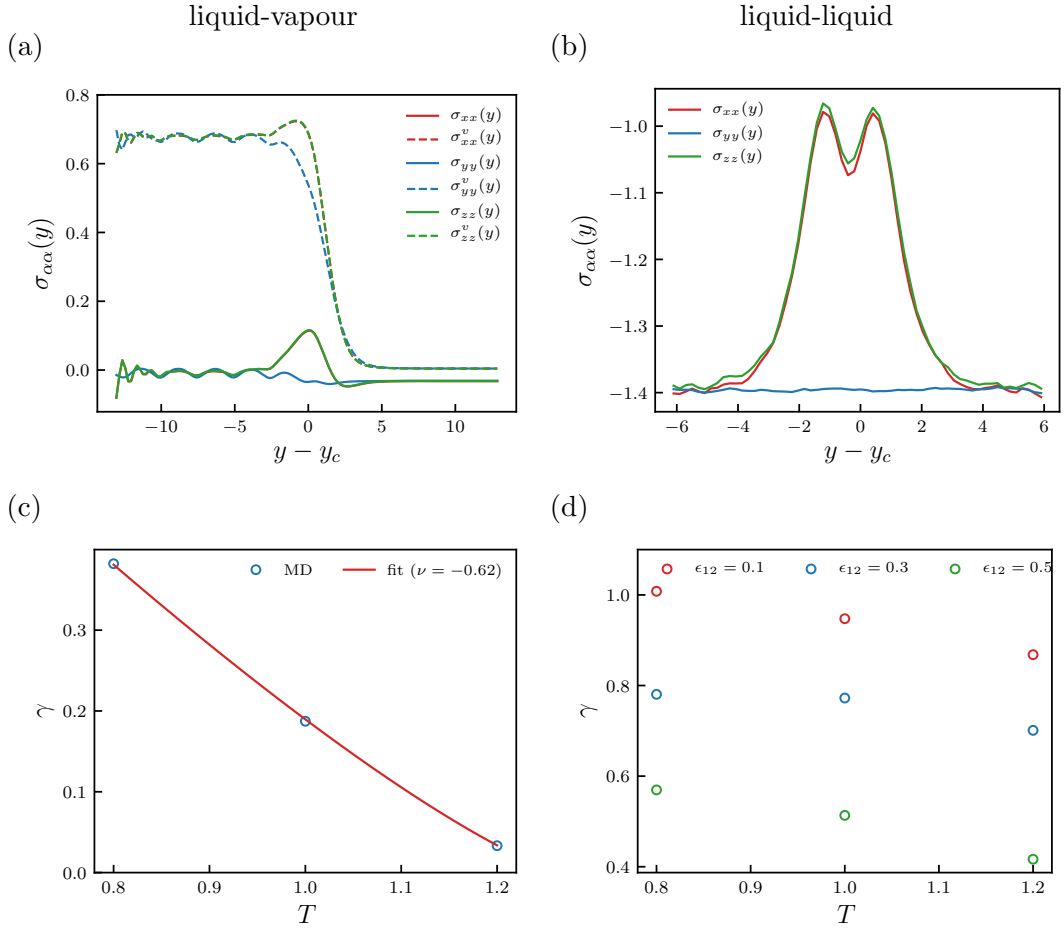

Figure 2.25. In (a-b) we show the diagonal stress tensor components σ_{xx}, σ_{yy} and σ_{zz} for LV (a) and LL (b) systems. We also report the surface tension values as a function of T and ϵ_{12} for LV (c) and LL (d).

The surface tension values for LV systems is shown in Fig. 2.25(a). Given the previously found values $T_C =$ and $\nu \sim 0.624$, we fit the surface tension for LV systems against the temperature employing the expression [163]:

$$\gamma(T) = \gamma_0 \left(1 - \frac{T}{T_C}\right)^{2\nu} . \tag{2.82}$$

The MD data and the fitting curve are shown in Fig. 2.25(a).

Table 2.2. Values of R^2 and relative root mean square error (RRMSE) for the viscosity-density fitting relation in LV systems.

System	Linear		Quadratic	
T	R^2	RRMSE	R^2	RRMSE
0.8	0.553	59.9%	0.802	39.8%
1.0	0.813	32.7%	0.933	19.5%
1.2	0.891	15.7%	0.941	11.5%

LL systems

For LL systems the kinetic component of $\sigma_{yy}(y)$ (Fig. 2.24(b)) shows a non-linear drop at the interface, which is however compensated by an equal and opposite increase in the potential component $\sigma^v_{yy}(y)$. It should be noted also that the total stress tensor components $\sigma_{yy}(y)$ evaluated for a single fluid (dashed line in Fig. 2.24(b)) drops to zero in parts of the domain and, as expected, does not satisfy the momentum conservation equation being only a subset of the total system. Similarly, the off-diagonal stress tensor component $\sigma_{xy}(y)$ (Fig. 2.24(d)), which is the sum of non-constant kinetic and potential contributions, is homogeneous across the system.

The surface tension values for LL systems is shown in Fig. 2.25(b). Similarly to LV systems, surface tension for LL systems decreases with the temperature. Additionally, the parameter ϵ_{12}, which governs the repulsive interactions between the two fluids, has a strong effect on the values of surface tension. Specifically, higher values of ϵ_{12} shift the surface tension-temperature curve down.

2.3.6. Shear viscosity

Local shear viscosity is evaluated as function of the non-diagonal stress component σ_{xy} and of the velocity gradient $\frac{\partial \mathbf{u}_x}{\partial y}$ as:

$$\eta(y) = \sigma_{xy} \left(\frac{\partial \mathbf{u}_x}{\partial y} \right)^{-1}. \tag{2.83}$$

LV systems

Figures 2.26(a-b) show the kinetic and the potential contributions to the shear viscosity profiles, respectively. As expected, the kinetic contribution is approximately one order of magnitude smaller than the potential one in the liquid phase. It is interesting to point out the anomalous behaviour of the kinetic viscosity at the interface, where a well

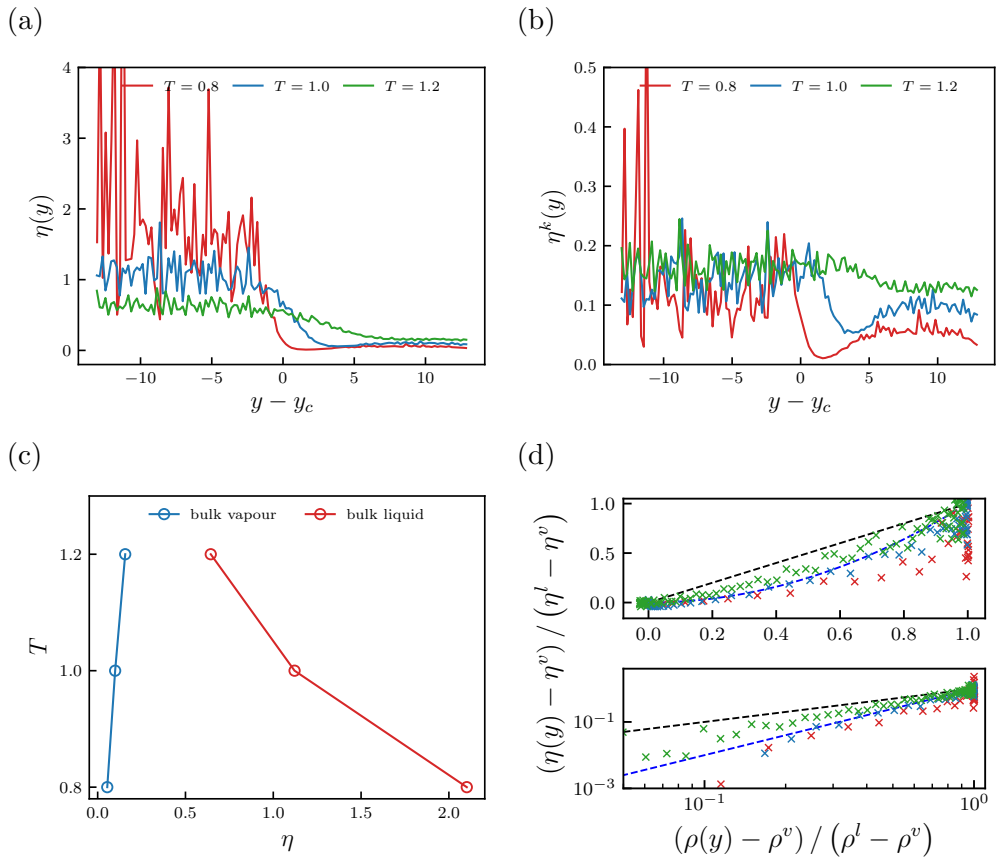

Figure 2.26. Total shear viscosity (a) and kinetic component of the shear viscosity (b) profiles for the LV interface at different temperatures. Bulk liquid and vapour viscosity as a function of the temperature for the LV system (c). Viscosity vs density profile for the LV interface obtained by MD (marks) and by Eq. (2.84) (solid line) (d)

appears at lower temperature, manifestation of a interface-induced confinement. This confirms that the shear viscosity is affected by the confinement induced by the liquid-vapour interface. As a consequence, the commonly adopted approximation of the shear viscosity trend at the interface as a hyperbolic tangent function (e.g. in some Volume of Fluids and Cahn-Hilliard solvers) may introduce significant errors at low temperatures.

Figure 2.26(c) shows the values of the shear viscosity in the bulk liquid and vapour. As expected, the viscosity in the vapour phase (which is dominated by its kinetic component), increases with the temperature. On the contrary, the liquid phase viscosity, mostly determined by its potential component, decreases with T since the thermal motion of particles reduces any effects of interaction between fluid layers.

In Fig. 2.26(d), the shear viscosity variation is plotted as function of the density vari-

Table 2.3. Values of R^2 and RRMSE for the viscosity-density fitting relation in LL systems.

System		Linear		Quadratic	
ϵ_{12}	T	R^2	RRMSE	R^2	RRMSE
0.1	0.8	0.509	54.0%	0.753	38.2%
0.1	1.0	0.862	26.5%	0.934	18.3%
0.1	1.2	0.893	22.7%	0.947	15.9%
0.3	0.8	0.834	29.5%	0.923	20.1%
0.3	1.0	0.905	21.5%	0.952	15.3%
0.3	1.2	0.931	17.9%	0.960	13.7%
0.5	0.8	0.877	24.5%	0.934	17.9%
0.5	1.0	0.938	16.8%	0.959	13.6%
0.5	1.2	0.949	14.5%	0.942	15.5%

ation along the liquid-vapour interface. For the investigated temperatures, the following family of relations between $\eta_s(y)$ and $\rho(y)$ is tested:

$$\eta_s(y) \sim \eta_s^v + \left(\eta_s^l - \eta_s^v\right)\left(\frac{\rho^l - \rho^v}{\rho(y) - \rho^v}\right)^k \tag{2.84}$$

where η_s^v and η_s^l denote shear the bulk viscosity in vapour and liquid respectively, ρ^v and ρ^l denote the bulk density in vapour and liquid respectively, and k is a parameter to be determined. Figure 2.26(d) shows the MD data (markers) against the model in Eq. (2.84) embedded with $k = 1$ (black line) and $k = 2$ (blue line). The model with $k = 2$ exhibit a better agreement with MD, as visible in both the linear and log-log plots. Also the values of R^2 and relative root mean square error reported in Table 2.2 confirm that the quadratic model ($k = 2$) outperform the linear one ($k = 1$) for every temperature T. In the range of temperature analysed, for both $k = 1$ and $k = 2$ an enhanced model accuracy is reached at higher temperature, probably because of reduced confinement effects at the interfaces. The quadratic relation between the two quantities is consistent with the expressions empirically derived in Ref. [166] for bulk fluids in the framework of dissipative particle dynamics. In fact, in Ref. [166] Groot and Warren approximate the viscosity as function of density as $\eta = 2\pi\gamma\rho^2 r_c^5/1575$ where γ and r_c are two constant coefficients.

LL systems

In Figs 2.27(a-b), we report the shear viscosity profiles at the interface between two partially miscible fluids for different values of T and ϵ_{12}. The shear viscosity shows

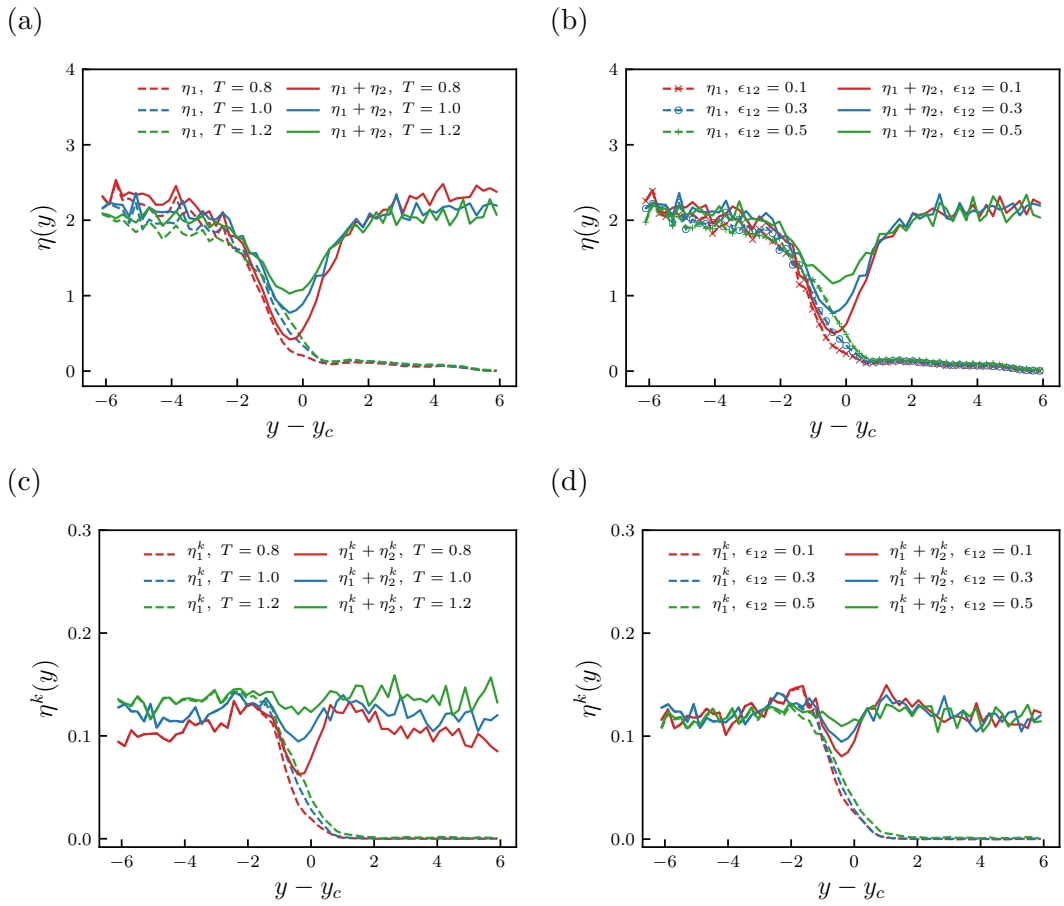

Figure 2.27. Shear viscosity profiles of the mixture of the two fluids (solid lines) and of one of them (dashed lines) for varying temperature T (a) and different values of the interaction parameter ϵ_{12} (b). Kinetic component of the shear viscosity profiles of the mixture of the two fluids (solid lines) and of one of them (dashed lines) for varying temperature T (c) and different values of the interaction parameter ϵ_{12} (d).

a well which reduces at high temperatures and high ϵ_{12} (less repulsive fluids). The corresponding kinetic contribution to the viscosity (shown in Figs 2.27(c-d)) is about an order of magnitude smaller than the total viscosity. Its profile is temperature-dependent and weakly affected by ϵ_{12} (except at the well).

In Fig. 2.28(a), the shear viscosity profile at the liquid-liquid interface is plotted against the corresponding density profile for a single fluid at the interface. Similarly to the LV interfaces, for the investigated T and ϵ_{12}, we tested the following family of relations

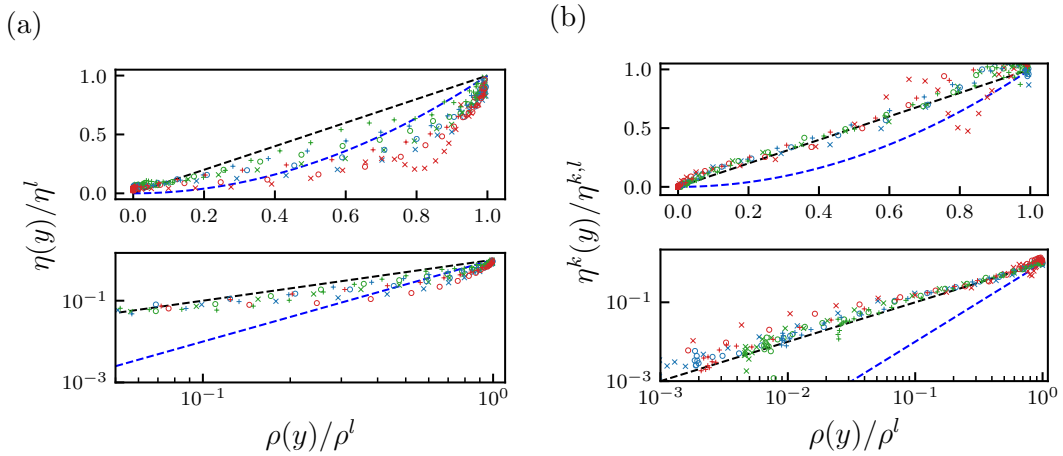

Figure 2.28. Total viscosity (a) and kinetic viscosity (b) profiles against density profiles for one single fluid at the LL interface.

between $\eta_1(y)$ and $\rho_1(y)$:

$$\eta_1(y) \sim \eta^l \left(\frac{\rho_1(y)}{\rho^l} \right)^k \tag{2.85}$$

where η^l and ρ^l denote the values of shear viscosity and density in bulk liquid respectively, and k is an exponent to be determined. In Fig. 2.28(a) we report the MD data (markers) against the model in Eq. (2.84) embedded with $k = 1$ and $k = 2$ (lines). The model with $k = 2$ exhibit a better agreement with MD, as clear from the linear plot. The better fitting accuracy of the quadratic model ($k = 2$) compared to the linear one ($k = q$) is quantitatively confirmed by Table 2.3 which reports the values of R^2 and relative root mean square error. In the range of T and ϵ_{12} analysed, for both $k = 1$ and $k = 2$ a better agreement is reached at higher temperatures and less repulsive interfaces (higher ϵ_{12}).

It is interesting to notice that the quadratic behaviour is mainly driven by the potential component of the viscosity. In fact, as shown in both linear and log-log plots of Fig. 2.28(b), the variation of the (normalized) kinetic viscosity varies linearly with the (normalized) density profile.

The quadratic relation observed for each one of the two (identical) fluids in the binary system allows us to gather some insights on the relation between the viscosity and the density profiles of the total system. The viscosity of the total fluid $\eta(y)$ can be expressed

as:

$$\eta(y) \simeq \eta_1(y) + \eta_2(y) = \frac{\eta_1^l}{{\rho_1^l}^2}\rho_1(y)^2 + \frac{\eta_2^l}{{\rho_2^l}^2}\rho_2(y)^2 \tag{2.86}$$

Since the two fluids in the analysed systems are identical, then $\frac{\eta_1^l}{{\rho_1^l}^2} = \frac{\eta_2^l}{{\rho_2^l}^2} = \frac{\eta^l}{{\rho^l}^2}$. It follows

$$\eta(y) \simeq \frac{\eta^l}{{\rho^l}^2}\left(\rho(y)^2 - 2\rho_1(y)\rho_2(y)\right) \tag{2.87}$$

where $\rho(y) = \rho_1(y) + \rho_2(y)$ is the total density. This simple derivation highlights that the quadratic relation derived for each one of the two (identical) fluids does not holds for the total fluid. Specifically, the multiplicative term proportional to $\rho_1(y)\rho_2(y)$ is different than zero only at the interface and must be subtracted from the quadratic contribution due to the total density.

2.3.7. Friction factor in nano-pipes with multi-phase flow

The proposed relations between density and viscosity profiles in multi-phase fluids can be employed directly to obtain the Darcy-Weisbach friction factor in nanoscale flows. The friction factor for single-component fluids at the nanoscale has been already obtained in previous studies [131, 156]. However, an extension of those approach to multi-phase fluids is still missing. As already reported in Refs [131, 156], for an infinitely wide channel with height h, the Darcy-Weisbach friction factor f is given by:

$$f = -\frac{1}{\bar{\rho}\bar{V}^2}\left(2D\frac{\partial p}{\partial x}\right) \tag{2.88}$$

where $\bar{\rho} = \frac{1}{h}\int_0^h \rho(y)$ is the average density in the channel, $\bar{V} = \frac{1}{h}\int_0^h u_x(y)$ is the average velocity, D is the hydraulic diameter (equal to $2h$ for the studied geometry) and p is the pressure (proportional to the trace of the stress tensor). For a Poiseuille flow, the momentum conservation equation in the x-direction states:

$$\frac{d}{dy}\left(\eta(y)\frac{du_x(y)}{dy}\right) = \frac{\partial p}{\partial x}. \tag{2.89}$$

Integrating Eq. (2.89) twice, one obtains an expression for the x-component of the velocity:

$$u_x(y) = \frac{\partial p}{\partial x}\int_0^y \frac{y'}{\eta(y')}dy' + u_s \tag{2.90}$$

where $u_s = u_x|_{y=0}$ is the slip velocity at one of the walls. Eq. (2.90) can be computed by employing the proposed relations for $\eta(y)$. For LV systems, we adopt the viscosity model in Eq. (2.84) together with the density profile approximation in Eq. (2.75). This allows us to obtain the following expression for the Darcy-Weisbach friction factor in nano-channels:

$$f = \frac{-2D\dfrac{\partial p}{\partial x}}{\frac{1}{2h^3}\int_0^h \left(2\rho^v + (\rho^l - \rho^v)\left(1 - \tanh\left(\frac{y-y_0}{d}\right)\right)\right) dy \left[\dfrac{\partial p}{\partial x}\int_0^h \left(\int_0^y \dfrac{y'}{\eta(y')}dy' + u_s\right) dy\right]^2}. \tag{2.91}$$

with $\eta(y) \sim \eta^v + \left(\eta^l - \eta^v\right)\left[\frac{1-\tanh\left(\frac{y-y_0}{d}\right)}{2}\right]^2$.

Similarly, for LL systems the Darcy-Weisbach friction factor can be obtained using Eqs (2.85) and (2.87), and expressing the total density and viscosity profiles as sum of the corresponding contributions by each fluid:

$$f = \frac{-2D\dfrac{\partial p}{\partial x}}{\frac{1}{h}\int_0^h (\rho_1(y) + \rho_2(y))dy \left[\dfrac{1}{h}\dfrac{\partial p}{\partial x}\int_0^h \int_0^y y' \left(\dfrac{\eta^l}{\rho^l}\left(\rho(y)^2 - 2\rho_1(y)\rho_2(y)\right)\right)^{-1} dy' + u_s\, dy\right]^2}. \tag{2.92}$$

where $\rho_1(y)$ and $\rho_2(y)$ can be approximated by employing Eq. (2.78). It is worth noting that the proposed approach for an infinitely wide channel can be extended to compute the Darcy-Weisbach factor of fluids in nano-pipes with arbitrary geometries.

In this chapter, we have analysed and discussed the microscopic properties of LJ systems at interfaces. Moreover we have extracted macroscopic relations from simulations data. However, MD simulations are typically computationally expensive. In the next chapter we derive and discuss low dimensional models that can be employed to investigate phenomena at atomistic scale.

Chapter 3

Low dimensional models

3.1. Microscopic coarse-graining: From molecular dynamics to generalized Langevin equation [2]

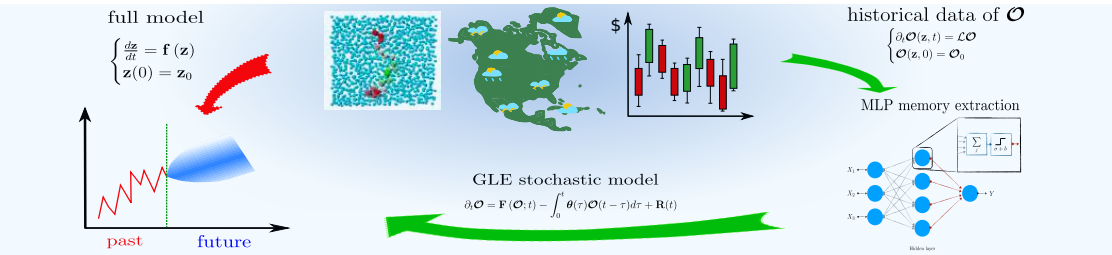

Figure 3.1. Two possible approaches to simulate the time evolution of a dynamical system observable. A first one consists in solving the full deterministic dynamical system (red arrow). Despite the advantage of being exact, this approach is often not suitable either because too computationally expensive or because a model of the full system is not accessible. An alternative approach consists in building a stochastic GLE model for an observable of the dynamical system and parametrizing it given proper historical data (green arrows).

Physical properties at the nano-scale are often investigated in a brute-force manner by solving the Newton's equations of motion for every particle, i.e. by means of molec-

[2] This section includes material from A. Russo, M. A. Durán-Olivencia, I. G. Kevrekidis & S. Kalliadasis. Machine learning memory kernels as closure for non-Markovian stochastic processes, *arXiv*:1903.09562 (2021).

ular dynamics (MD) or Monte Carlo simulations. However, many interesting physical events, such as protein folding [167, 168] and phase nucleation [169, 170], occur at time and length scales several orders of magnitude larger than those accessible by atomistic simulations. Due to the prohibitively high computational cost associated with full simulations, coarse-graining remains a highly relevant instrument in the study of physical systems.

Precursory model reduction techniques go back to the phenomenological description of pollen particles in water by Brown [171]. However, it was not until the work of Langevin [172] that a stochastic model for such phenomena was formulated. Since then, several reduced models have been proposed [33, 34, 35, 36, 40, 43, 44, 45, 61, 173, 174, 175]. A standard way to coarse-grain a physical system of particles is by using projection-operator techniques to decouple the relevant and the irrelevant degrees of freedom in the system [34, 40, 43, 176]. This leads to the following stochastic Generalized Langevin Equations (GLEs) for the time evolution of the resolved variables (or relevant degrees of freedom):

$$\partial_t \boldsymbol{\mathcal{O}}(t) = \mathbf{F}(\boldsymbol{\mathcal{O}}) - \int_0^t \boldsymbol{\theta}(\tau)\boldsymbol{\mathcal{O}}(t-\tau)d\tau + \mathbf{R}(t) \tag{3.1}$$

with $\mathbf{F}(\boldsymbol{\mathcal{O}})$ accounting for the mean force contributions. The term $\mathbf{F}(\boldsymbol{\mathcal{O}})$ depends only on the current system configuration and, in some cases, corresponds to the mean force term [37, 38, 39, 40]. Several approaches have been developed to compute the potential of mean force of a system, including adaptive biasing forces [41] and umbrella sampling [42]. The vector $\mathbf{R}(t)$ is interpreted as a stochastic term, with correlation given by the fluctuation dissipation theorem $\langle \mathbf{R}(t), \mathbf{R}(t') \rangle = \boldsymbol{\theta}(t-t')\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}} \rangle$, where the notation $\langle \mathbf{A}, \mathbf{B} \rangle$ indicates the inner product $\langle \mathbf{A}, \mathbf{B} \rangle = \int \rho(\mathbf{z})\mathbf{A}(\mathbf{z})\mathbf{B}^*(\mathbf{z})\ d\mathbf{z}$, with $\rho(\mathbf{z})$ being a normalized pdf defined in the phase space of the original system, and $\mathbf{B}^*$ the conjugate transpose of $\mathbf{B}$. The convolutional (non-Markovian) term depends on the previous evolution of the system and is characterized by the memory kernel function $\boldsymbol{\theta}(t)$, which also unequivocally determines the characteristics of the noise term $\mathbf{R}(t)$ through the fluctuation dissipation theorem. As a consequence, a proper approximation of $\boldsymbol{\theta}(t)$ is required to preserve the main features of the original high dimensional system into the reduced one. However, the memory kernel depends on both the full set of DoF and the whole history of the complex system, hence making the problem often intractable.

In previous studies, several approaches have been proposed to parametrize GLEs. Analytical forms can be only obtained for specific systems, such as a particle in a harmonic oscillator heat bath [43], while numerical techniques are necessary for more complex

systems characterized by non-linear interactions. For instance, in Ref. [44] the authors adopt a perturbation scheme, which is yet "too complex for general use". Despite its accuracy, the algorithm developed in Ref. [45] to parametrize GLEs involves sampling of the full original system, thus, becoming computationally prohibitive for large systems. Another procedure involving large matrix computations and Krylov sub-space approximations is shown in Ref. [40]. In both Refs [47, 48], the researchers propose to extract the memory kernel by Laplace transforming the correlation functions computed from some historical data of the observables. However, this strategy exhibits serious limitations when the available data on the observables are affected by fluctuations, as shown later in this section.

In this section, we present a novel data-driven approach, which makes use of the GLE structure coupled with a multilayer perceptron (MLP) to achieve an optimal parametrization of the memory kernel. The MLP is provided with proper historical data of the observables of interest obtained either from simulations or existing databases and then executes an optimization procedure to find the optimal approximation of the memory kernel. As we shown later in this section, compared to previous approaches our approximation through MLP shows enhanced robustness, especially when the available data are limited or affected by significant fluctuations. In the this procedure, the memory kernel is extracted in the form of a multi-exponential functions, thus enabling us to derive a tractable stochastic integration algorithm of the non-Markovian process characterized by a time-correlated noise. The universal approximation theorem [177, 178] guarantees a wide applicability of the our methodology which is tested in some relevant case studies from chemistry, biology, climatology and finance.

Correlations equation Let us consider a system in equilibrium condition, such that the historical data of some observables of the system can be considered a realization of a stationary process. If we take the inner product of the GLE with $\boldsymbol{\mathcal{O}}(0)$, one obtains the following correlation equation [48]:

$$\begin{aligned}\mathbf{g}(t) &= -\left\langle \partial_t \boldsymbol{\mathcal{O}}(t) - \mathcal{P}\mathcal{L}\boldsymbol{\mathcal{O}}(t), \boldsymbol{\mathcal{O}}(0)\right\rangle = \left\langle \int_0^t \boldsymbol{\theta}(t-\tau)\boldsymbol{\mathcal{O}}(\tau)d\tau + \mathbf{R}(t), \boldsymbol{\mathcal{O}}(0)\right\rangle \\ &= \int_0^t \boldsymbol{\theta}(t-\tau)\left\langle \boldsymbol{\mathcal{O}}(\tau), \boldsymbol{\mathcal{O}}(0)\right\rangle d\tau + \left\langle \mathbf{R}(t), \boldsymbol{\mathcal{O}}(0)\right\rangle = \int_0^t \boldsymbol{\theta}(t-\tau)\mathbf{h}(\tau)d\tau,\end{aligned} \tag{3.2}$$

where the matrices $\mathbf{g}(t)$ and $\mathbf{h}(t) = \langle \boldsymbol{\mathcal{O}}(t), \boldsymbol{\mathcal{O}}(0)\rangle$ were introduced, and the orthogonality between the random force and the initial value of the observable was employed to set $\langle \mathbf{R}(t), \boldsymbol{\mathcal{O}}(0)\rangle = \mathbf{0}$. For ergodic systems $\mathbf{g}(t)$ and $\mathbf{h}(t)$ are evaluated from the data by

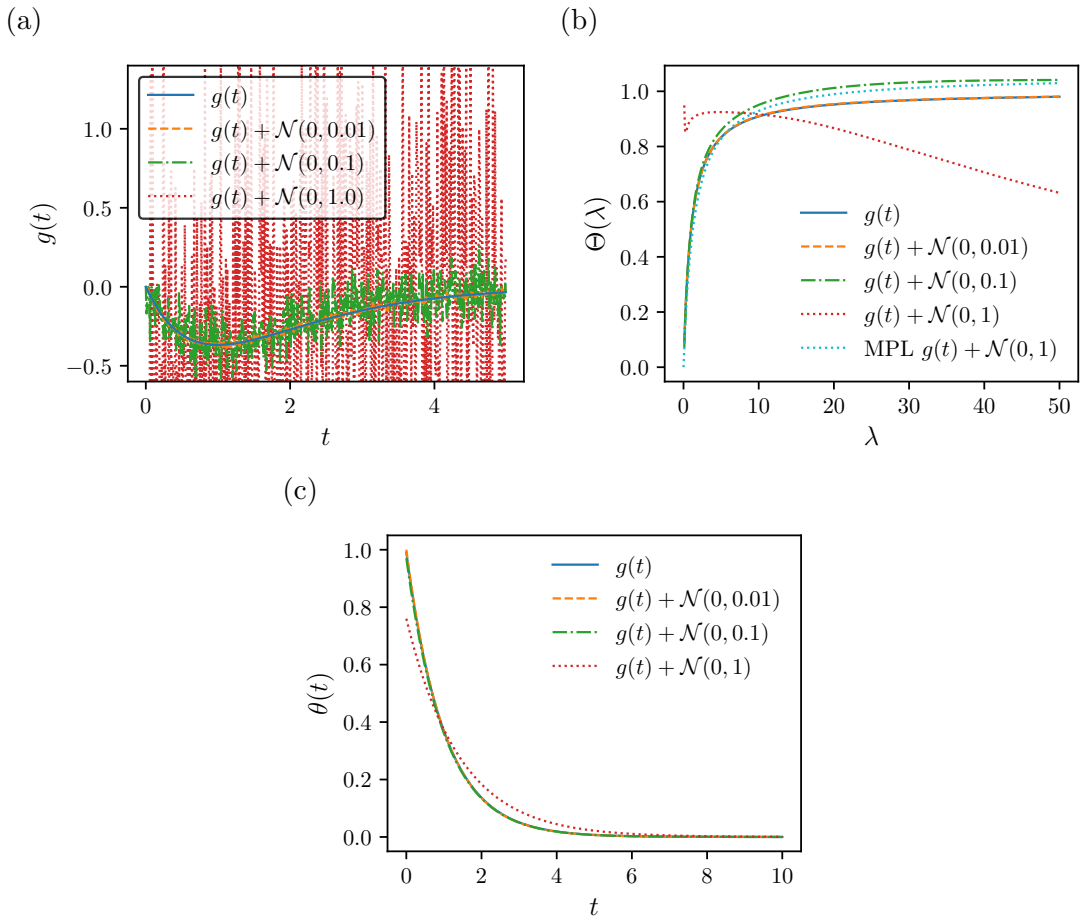

Figure 3.2. Convolution function $g(t)$ affected by random noise with varying amplitudes (a). Comparison between the memory kernel θ computed in the Laplace space (b) and with our MLP-based method (c), for $g(t)$ affected by random noises. For comparison purpose, in (b) we also report the Laplace transform of the memory kernel obtain with our MLP for the strongest noise.

means of time averages. In many scenarios, such as one-dimensional systems or systems of spherical particles, $\boldsymbol{\theta}(t)$, $\mathbf{g}(t)$ and $\mathbf{h}(t)$ are diagonal matrices, i.e. $\boldsymbol{\theta}(t) = \theta(t)\mathbf{1}$, $\mathbf{g}(t) = g(t)\mathbf{1}$ and $\mathbf{h}(t) = h(t)\mathbf{1}$. In such cases, hereinafter, we will denote the scalar functions simply as $\theta(t)$, $g(t)$ and $h(t)$.

3.1.1. Memory kernel in the Laplace space

Recently, in Ref. [48] it was proposed a way to compute the memory kernel using the properties of Laplace transform defined as $\mathcal{L}_p\left(\mathbf{A}(t)\right) = \int_0^\infty \mathbf{A}(t)e^{-t/\lambda}$. In fact, it can be

easily shown that Eq. (3.2) in the Laplace space takes the simple form

$$\mathcal{L}_p\left(\boldsymbol{\theta}(t)\right) = \mathcal{L}_p\left(\mathbf{g}(t)\right)\left[\mathcal{L}_p\left(\mathbf{h}(t)\right)\right]^{-1}. \tag{3.3}$$

Despite its simplicity, this approach is not suitable in case of limited data, which produce correlations affected by random noise. As an example, let us consider a Gaussian error $\boldsymbol{\epsilon}(t)$ affecting only the function $\mathbf{g}(t)$. Then, the error acting on the Laplace transform of the kernel $\boldsymbol{\Theta}(\lambda) = \mathcal{L}_p\left(\boldsymbol{\theta}(t)\right)$, is defined as

$$\Delta\boldsymbol{\Theta}(\lambda) = \tilde{\boldsymbol{\Theta}}(\lambda) - \boldsymbol{\Theta}(\lambda) = \mathcal{L}_p\left(\boldsymbol{\epsilon}(t)\right)\left[\mathcal{L}_p\left(\mathbf{h}(t)\right)\right]^{-1}. \tag{3.4}$$

If we assume $\boldsymbol{\epsilon}(t)$ to be the sum of non-systematic local errors, i.e. $\boldsymbol{\epsilon}(t) = \sum_i \boldsymbol{\epsilon}_i(t) = \sum_i \mathbf{k}_i\delta(t-t_i)$, with $\boldsymbol{\epsilon}_i \sim \mathcal{N}(0,\sigma^2)$, then the total error acting on the memory kernel in the Laplace space becomes

$$\Delta\boldsymbol{\Theta}(\lambda) = \sum_i \mathbf{k}_i e^{-t_i/\lambda}\left[\mathcal{L}_p\left(\mathbf{h}(t)\right)\right]^{-1}. \tag{3.5}$$

This argument shows that local random errors in the real space turn into non-local contributions in the Laplace space. Such a propagation can lead to significant inaccuracies if the memory kernel is approximated in the Laplace space.

As an example, let us consider the simple case, that can be analytically solved, with $h(t) = e^{-t}$ and a noisy $g(t) = -te^{-t}$, as reported in Fig. 3.2(a). Figure 3.2(b) shows that the exact memory kernel computed in the Laplace space (as in Ref. [48]) diverges from the analytical solution when noise becomes significantly large. To overcome this issue, we propose to adopt a MLP-based procedure that gives an optimal approximation of $\boldsymbol{\theta}$ in the real space. Our method is robust and allows to reproduces accurately the expected function even when the strongest noise affects the data, as shown in Fig. 3.2(c).

3.1.2. Memory kernel extraction through a multi-layer perceptron

Artificial neural networks are an interesting substitute of conventional methods in the parametrization of the GLE because of their enhanced capabilities in function approximations. Developed in analogy with the biological processes in the brain, artificial neural networks are essentially made of series of linear and non-linear transformations of some inputs in some output. Among the different possible variants, MLPs have gained popularity because of their versatility and capabilities in the approximation of non-linear functions [179].

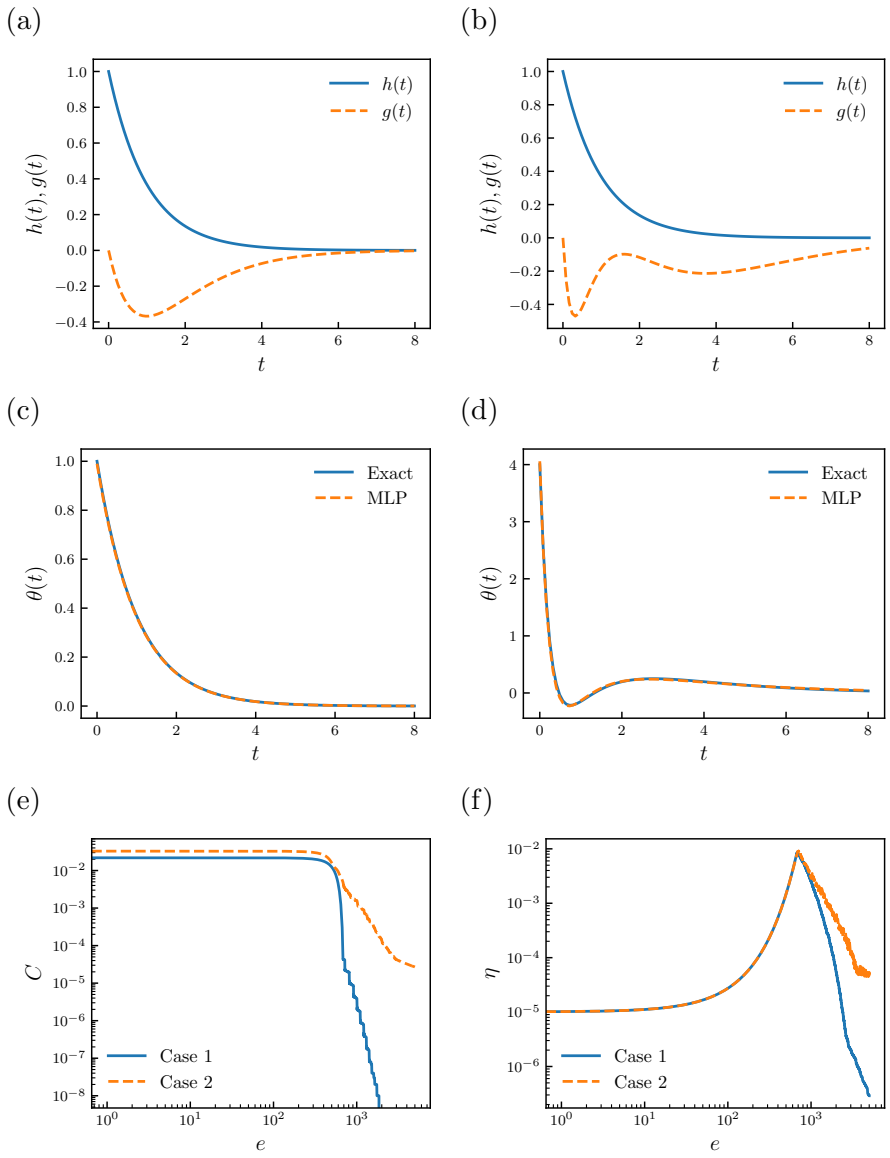

Figure 3.3. In order to test our methodology two representative cases (discussed in the main text) are reported: Case 1 in (a-c) and case 2 in (b-d). The functions $h(t)$ and $g(t)$ (a-b), discretized at 800 points, are provided to the MLP. The comparison between the memory kernel θ computed numerically with our MLP and the exact one is given in (c-d). In (e-f) we show the cost function and learning rate for the two analysed scenarios. In both cases, the numerical approximation is obtained with a MLP trained for 5000 epochs.

MLPs consist of at least three layers, known as input, hidden and output layers, each one including several nodes (Fig. 3.4). Each node i in the layer $l-1$ is connected with

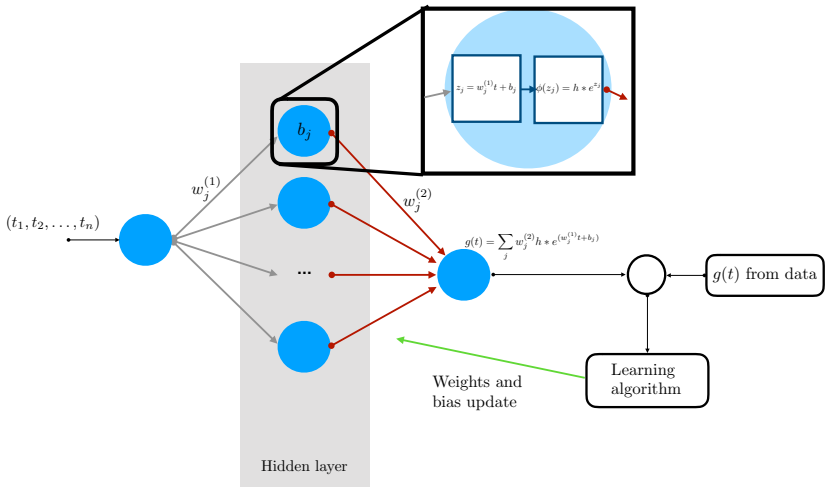

Figure 3.4. Representation of a MLP structure and the learning process

any other node j in the successive layer l and every connections is characterized by a parameter called weight $w^l_{j,i}$. In addition, to every neuron in the network corresponds a parameter called bias b^l_j. The transformation of the dataset at each node is determined by an activation function $\sigma(z^l_j)$. It follows that the output a^l_j of the neuron j of the layer l is computed as $a^l_j = \sigma\left(z^l_j\right)$, with $z^l_j = \sum_i w^l_{j,i} a^{l-1}_i + b_j$. The network learning process consists in an optimization algorithm aiming to find weights $w^l_{j,i}$ and biases b^l_j that minimize a cost (or error) function C computed at the output of the MLP. In this work, we employ a quadratic cost function

$$C = \sum_{t_j}^{N_t} \frac{1}{2N_t} \left(y_j(t_j) - a^L_j(t_j)\right)^2, \tag{3.6}$$

where N_t is the number of data samples. Hence, an algorithm is used to cyclically back-propagate the information about the error evaluated at the output to update weights and bias. Every repetition of this algorithm is called epoch e and the whole procedure is commonly known as learning process.

In the present work, we adopt a three layer MLP with a single input and a single output function. The hidden layer has an arbitrary number of neurons N_n, determining the degree of accuracy of the memory kernel approximation. The universal approximation theorem guarantees that such a structure of the network is able to approximate any continuous function defined on a compact subset of $\mathbb{R}^d$ [177, 178]. Initialization of the MLPs is obtained providing Gaussian distributed random numbers to the weights, and

zeros to the bias. Moreover, no bias is added at the output layer. As regards the activation functions, in the hidden layer we adopt $\sigma(z) = \int_0^t h(t-\tau)e^{z(\tau)}d\tau$, with $h(t)$ being known a priori, while we employ $\sigma(z) = z$ at the output layer. The learning algorithm adopted is the resilient back-propagation algorithm (Rprop) [180], which can be synthesized as follows:

$$\eta(e) = \begin{cases} \eta^+ \cdot \eta(e-1) & \text{if} \quad \frac{\partial C}{\partial \alpha}(e) \cdot \frac{\partial C}{\partial \alpha}(e-1) > 0, \\ \eta^- \cdot \eta(e-1) & \text{if} \quad \frac{\partial C}{\partial \alpha}(e) \cdot \frac{\partial C}{\partial \alpha}(e-1) < 0, \\ \eta(e-1) & \text{otherwise}, \end{cases} \tag{3.7}$$

where $\alpha = \left[w_{j,i}^l; b_j^l\right]$, e indicates the epoch, η is the adaptive learning rate and $0 < \eta^- < 1 < \eta^+$ are fixed parameters. In our experience and according to the literature [180], Rprop algorithm gives an optimal compromise between fastness of the response and solution convergence.

The memory kernel is extracted in the form of an exponential series, namely as:

$$\theta(t) \sim \sum_{k=1}^{N_n} w_k^3 e^{b_k^2} e^{w_k^2 t} = \sum_{k=1}^{N_n} A_k e^{B_k(t)}, \tag{3.8}$$

where N_n is the number of nodes in the hidden layer, $A_k = w_i^3 e^{b_k^2}$ are real numbers and $B_k = w_k^2$ are strictly real negative coefficients. The algorithm presented is adopted to extract the memory kernel in case of diagonal $\boldsymbol{\theta}(t)$, but a generalization of our approach to non diagonal memory kernel matrices is also possible.

As a preliminary test of our approach, we consider triples of simple functions $h(t)$, θ and $g(t)$ satisfying the relation $g(t) = -\int_0^t \theta(t-\tau)h(\tau)d\tau$. Given $h(t)$ and $g(t)$, an approximation of θ is computed with our methodology and is compared with the exact analytical θ. Two tests with different sets of functions are reported here. The functions used for the first test are the following:

$$h(t) = e^{-t}, \quad \theta(t) = e^{-t}, \quad g(t) = -te^{-t} \tag{3.9}$$

In this test, because of the single exponential form of θ, a MLP with a single neuron in the hidden layer is adopted, namely $N_n = 1$.

The functions adopted for the second test are the following:

$$\begin{aligned} h(t) &= e^{-t}, \quad \theta(t) = 6e^{-4t} - 4e^{-t} + 2e^{-t/2} \\ g(t) &= -\left(2e^{-t} - 2e^{-4t} - 4te^{-t} + 4e^{3t/2} - 4e^{-t}\right) \end{aligned} \tag{3.10}$$

For this latter example, we impose $N_n = 3$ neurons in the hidden layer.

Figure 3.3(a-b) reports $h(t)$ and $g(t)$ provided as input to the MLP for both tests. The comparisons between numerical approximations and analytical θ reported in Figs 3.3(c-d) shows that our methodology can accurately learn the exact memory kernel. The behaviours of cost function and learning rate during the learning process for both tests is also shown in Figs 3.3(e-f).

3.1.3. GLE integration

The integration of the GLE dynamics is a not-trivial task for two reasons: first, the convolution integral depends on the full history of the observable, and second, the stochastic term is correlated in time. Several approaches have been proposed to face these issues based on the introduction of a set of auxiliary variables, i.e. Refs [48, 181, 182]. In the present work, we take advantage of the exponential structure of $\boldsymbol{\theta}(t)$ to implement an integration algorithm. The history-dependent convolution term is written as a sum of the additional variables $\mathbf{Z}_k(t)$, each defined as $\mathbf{Z}_k(t) = \int_0^t \mathbf{A}_k e^{\mathbf{B}_k(t-\tau)} \boldsymbol{\mathcal{O}}(\tau) d\tau$, so that their evolution equation can be expressed as $\dot{\mathbf{Z}}_k(t) = \mathbf{B}_k \mathbf{Z}_k(t) - \mathbf{A}_k \boldsymbol{\mathcal{O}}(t)$.

The noise $\mathbf{R}(t)$ has to be generated with proper time correlations in order to satisfy the fluctuation-dissipation theorem. The introduction of an additional set of auxiliary variables $\boldsymbol{\xi}_k(t)$ allows us to express it as function of a standard white noise processes. In details, the noise term is decomposed as $\mathbf{R}(t) = \sum_{k=1}^{Nn} \mathbf{R}_k(t) = \sum_{k=1}^{Nn} \mathbf{b}_k \boldsymbol{\xi}(t)$, so that the corresponding evolution reads $\dot{\mathbf{R}}_k(t) = \mathbf{B}_k \mathbf{R}_k(t) + \mathbf{b}_k \boldsymbol{\xi}(t)$, where $\boldsymbol{\xi}(t)$ is a white noise with zero mean and time correlation $\langle \boldsymbol{\xi}(t) \boldsymbol{\xi}(s) \rangle = 2\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}} \rangle \, \delta(t-s)$, while the coefficients $\mathbf{b}_k$ can be computed numerically (for details see Appendix B.1.2).

As a result, after defining the variables $\mathbf{S}_k(t) = -\mathbf{Z}_k(t) + \mathbf{R}_k(t)$, the GLE can be rewritten in form of extended dynamics as:

$$\begin{cases} \dot{\boldsymbol{\mathcal{O}}} = \mathbf{F}(\boldsymbol{\mathcal{O}}(t)) + \sum_{k=1}^{N_n} \mathbf{S}_k(t) \\ \dot{\mathbf{S}}_k(t) = \mathbf{B}_k \mathbf{S}_k(t) - \mathbf{A}_k \boldsymbol{\mathcal{O}}(t) + \mathbf{b}_k \boldsymbol{\xi}(t). \end{cases} \tag{3.11}$$

The numerical algorithm adopted to solve the system is the following splitting method,

with Euler-Maruyama scheme for $S_k(t)$:

$$\boldsymbol{\mathcal{O}}^{(n+1/2)} = \boldsymbol{\mathcal{O}}^{(n)} + \frac{\Delta t}{2} F^c(\boldsymbol{\mathcal{O}}^{(n)}) + \frac{\Delta t}{2} \sum_{k=1}^{N_n} S_k^{(n)}, \tag{3.12}$$

$$S_k^{(n+1)} = (1 + B_k \Delta t)\, S_k^{(n)} - A_k \boldsymbol{\mathcal{O}}^{(n+1/2)} \Delta t + b_k \xi_k^{(n)}, \tag{3.13}$$

$$\boldsymbol{\mathcal{O}}^{(n+1)} = \boldsymbol{\mathcal{O}}^{(n+1/2)} + \frac{\Delta t}{2} F^c(\boldsymbol{\mathcal{O}}^{(n+1)}) + \frac{\Delta t}{2} \sum_{k=1}^{N_n} S_k^{(n+1)} \tag{3.14}$$

where $\xi_k^{(n)} \sim \mathcal{N}(0, 2\Delta t)$ are independent Gaussian distributed random values.

In order to test the numerical stochastic integrator, similarly to Ref. [181], we consider a one dimensional GLE with single exponential memory kernel and no conservative forces. In this specific case, the time correlation is analytically solvable. Thus, we compare the autocorrelation function computed numerically with the analytical one, which can be expressed as [181]:

$$\frac{\langle \boldsymbol{\mathcal{O}}(t)\boldsymbol{\mathcal{O}}(0)\rangle}{\langle \boldsymbol{\mathcal{O}}(0)\boldsymbol{\mathcal{O}}(0)\rangle} = \begin{cases} e^{\frac{tB}{2}} \left(\cos(\Omega t) - \frac{B}{2\Omega}\sin(\Omega t)\right) & \Omega \neq 0, \\ e^{\frac{tB}{2}} \left(1 - \frac{Bt}{2}\right) & \Omega = 0, \end{cases} \tag{3.15}$$

where it is introduced the complex parameter $\Omega = \sqrt{A - B^2/4}$. Figure 3.5 shows that the numerical integrator is able to accurately reproduce the analytical correlation in the under-damped ($A = 1$ and $B = 1$), in the damped ($A = 1$ and $B = -2$) and in the over-damped cases ($A = 1$ and $B = -4$).

3.1.4. Numerical application I: Single particle in bath

The proposed methodology is tested, first, to model the global effect of a heat bath on a single particle. Data regarding momentum and forces of the target particle, with mass $m = 1$, immersed in a bath of identical particles with masses $m_b = 1.0$ are gathered from equilibrium MD simulations. The interactions between any two particles i and j are modelled via a Lennard-Jones (LJ) potential:

$$v_{\mathrm{LJ}}(\mathbf{r}_{ij}) = \begin{cases} 4\epsilon \left[(\sigma/r_{ij})^{12} - (\sigma/r_{ij})^{6}\right] & \text{if } r_{ij} \leq r_c, \\ 0 & \text{otherwise,} \end{cases} \tag{3.16}$$

where $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between the particles, $\epsilon = 1.0$ is the depth of the potential well, $\sigma = 1.0$ is the finite atom-atom distance at which the potential is zero,

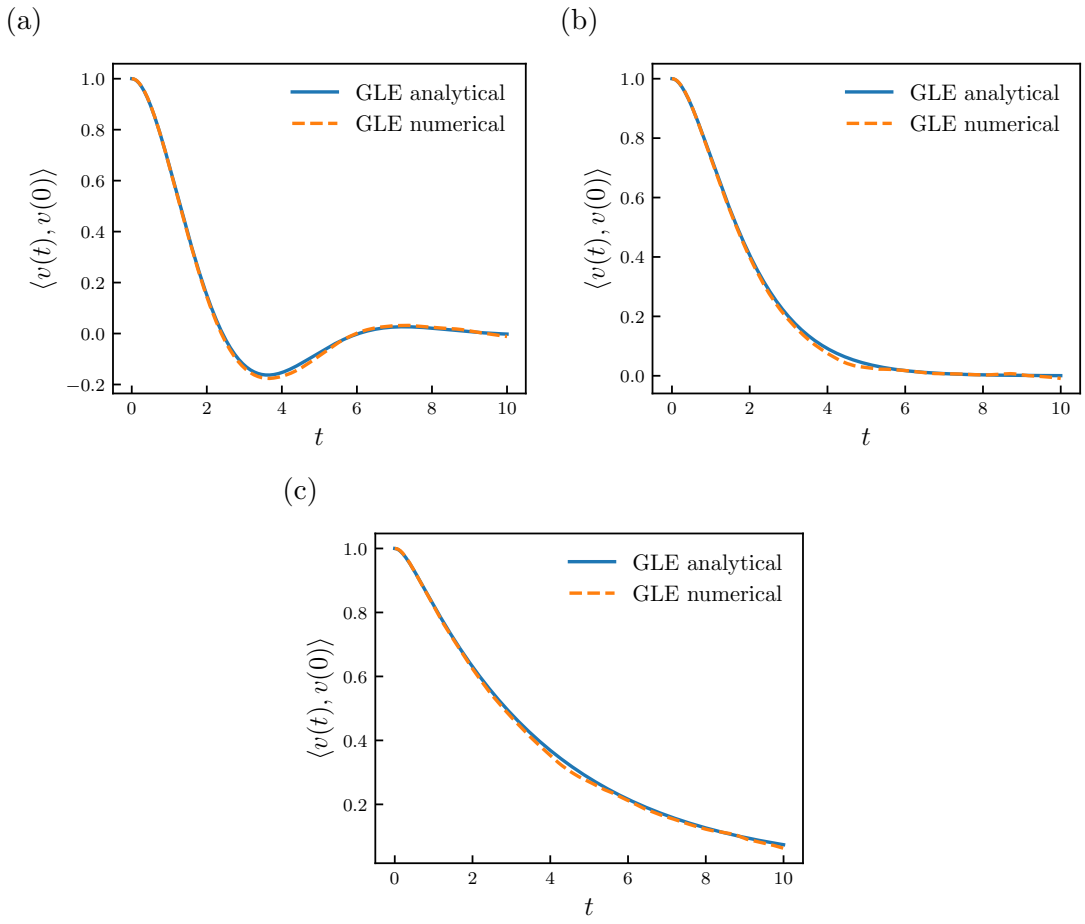

Figure 3.5. Comparison between numerical and analytical time correlation computed over 10^4 independent trajectories for a GLE with memory kernel in the form of a single exponential function [181]. The correlation is computed in the under-damped limit with $A = 1$ and $B = 1$ (a), in the damped case with $A = 1$ and $B = -2$ (b) and in the over-damped limit with $A = 1$ and $B = -4$ (c). In all cases the temperature is set to $T = 1$.

and $r_c = 2.5\sigma$ is the cut-off radius.

The simulation box dimensions are $10\sigma \times 10\sigma \times 10\sigma$, and periodic boundary conditions are imposed along x, y and z axes. A Nosé-Hoover thermostat is used to equilibrate the system at a reduced temperature $T = 1.0$. The following procedure is followed to run the MD simulations. First, the bath particles are randomly generated inside the simulation box. Then, a minimization algorithm is employed to avoid overlaps between particles. Hence, a run of 10^5 time steps is used to equilibrate the system. Finally, data on forces and momenta are gathered over 10^5 time steps. This process is repeated for

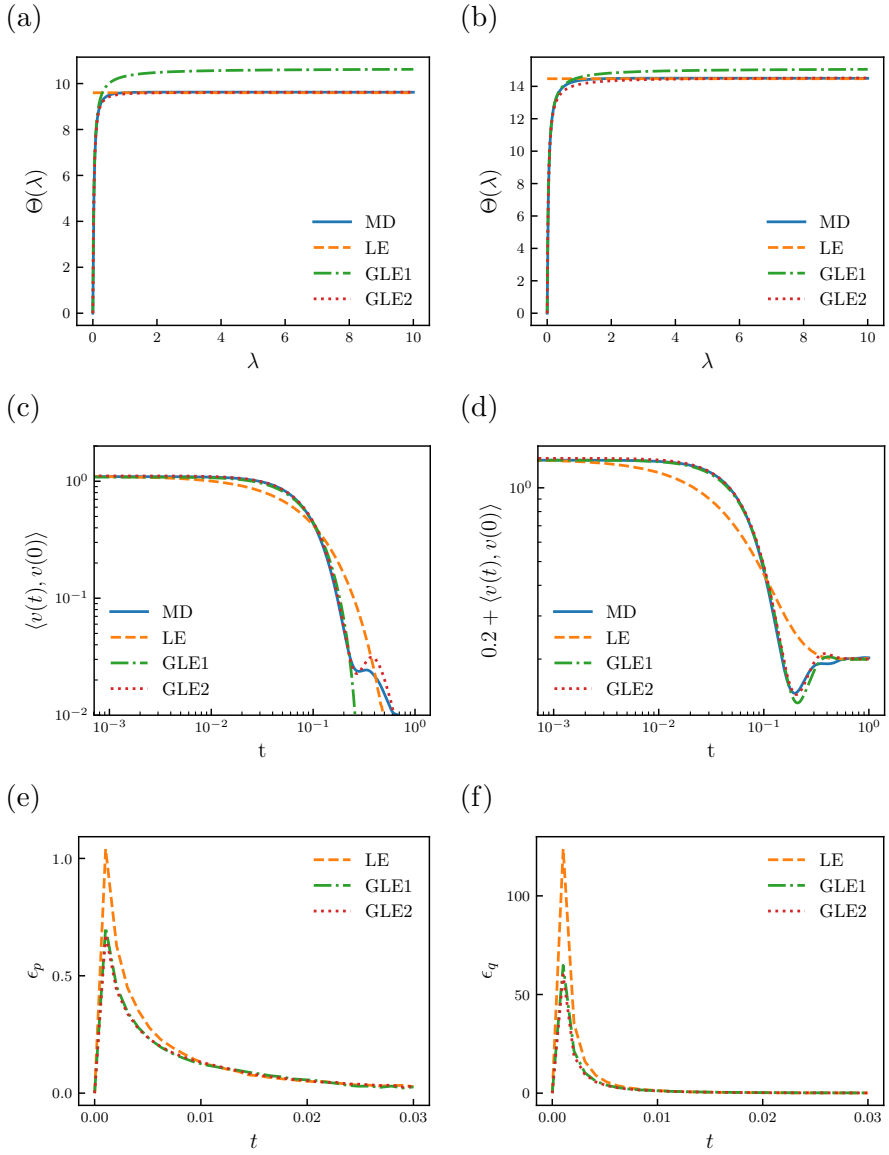

Figure 3.6. Memory kernels computed with the MLP are compared against the one obtained directly from MD simulation in the Laplace space, for (a) LDL and (b) HDL cases. Velocity correlation functions computed from MD, LE and GLE dynamics over 10^4 trajectories for both (c) LDL and (d) HDL cases are also reported. GLE1 and GLE2 refer to the memory kernel approximations obtained respectively with 1 and 2 neurons in the hidden layer. In (e-f) we report the mean square differences $\epsilon_p(t)$ and $\epsilon_q(t)$ between the pdfs of the reduced systems (GLE and LE) and the exact pdf of the full system (MD) as function of the relaxation time.

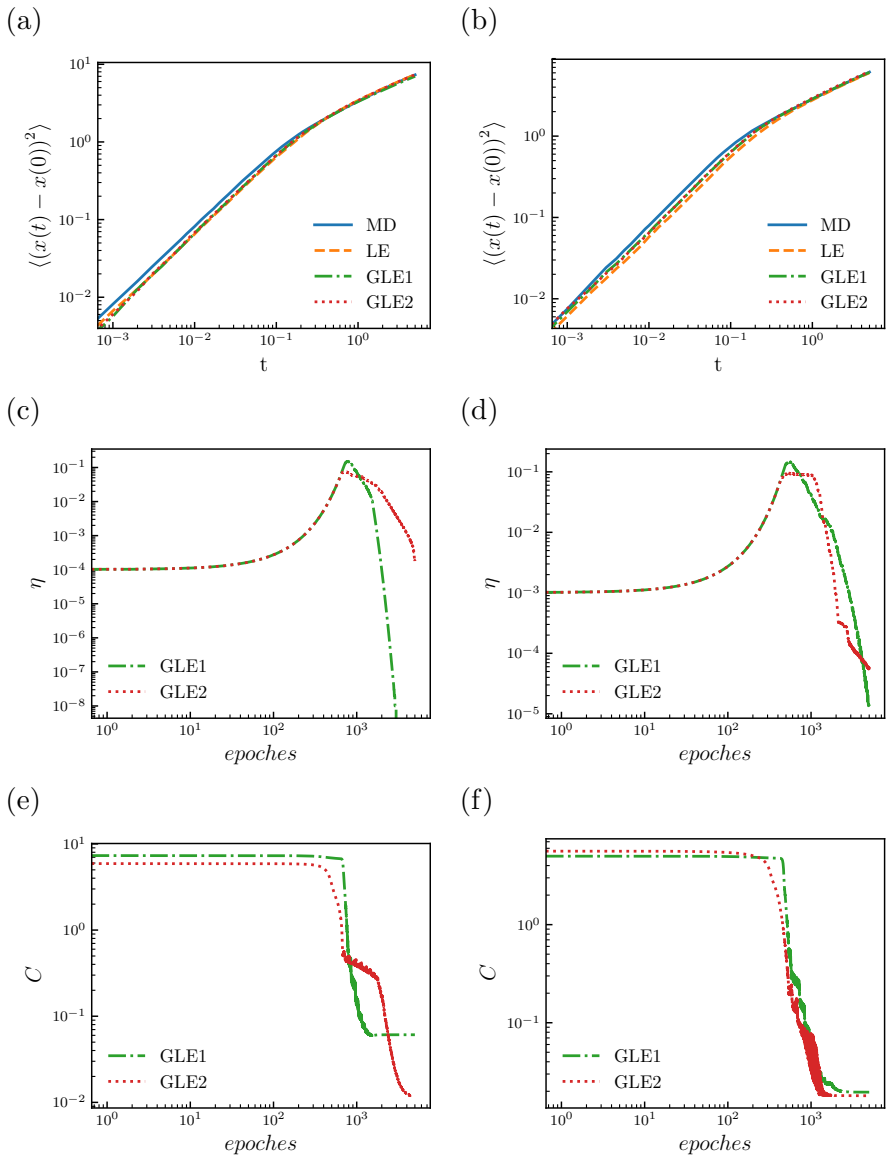

Figure 3.7. Comparison of the mean square displacement (MSD) computed with GLE and MD, for both LDL (a) and HDL (b) cases. GLE1 and GLE2 refer to the memory kernel approximations obtained respectively with 1 and 2 neurons in the hidden layer.

10^2 trajectories in order to enhance the accuracy of the correlations, and consequently, of the memory kernels.

In this study, we consider two bath densities: the low density limit (LDL) with $\rho = 0.699$, and the high density limit (HDL) with $\rho = 0.799$. Figures 3.6(a-b) show the comparison in the Laplace space of the exact memory kernel computed with MD and with

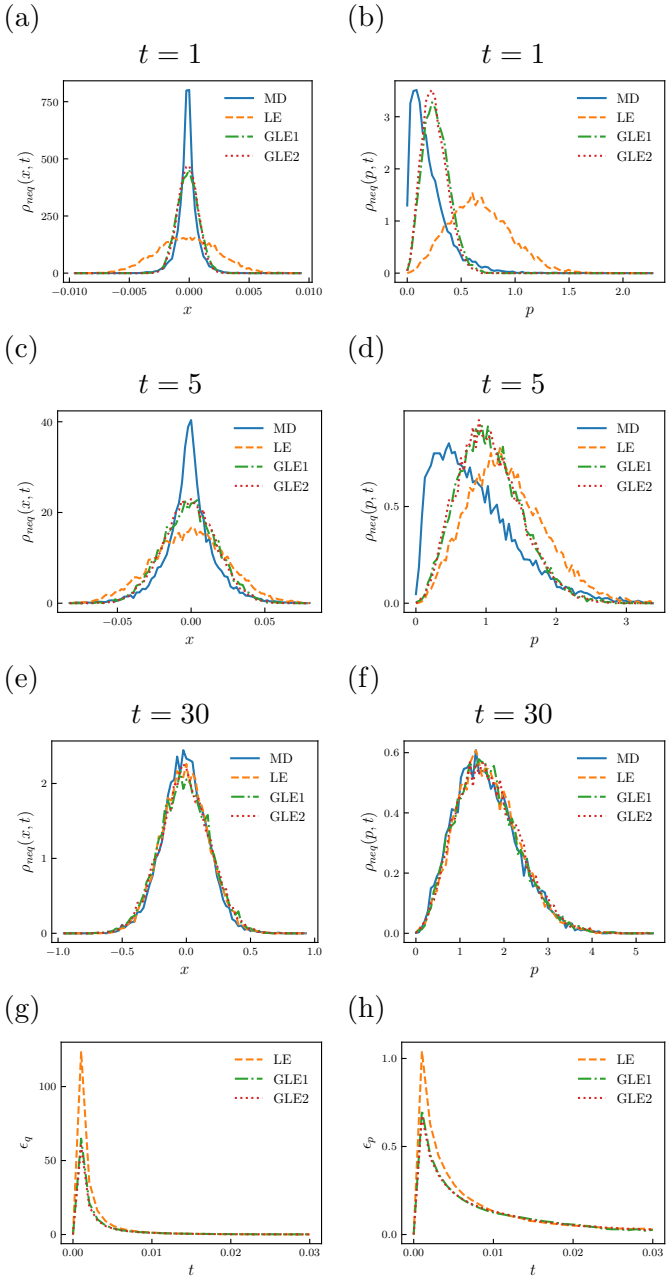

Figure 3.8. Relaxation dynamics of position (a-c-e) and momentum (b-d-f) probability density function from Dirac delta to equilibrium condition computed with MD, LE and GLE over 10^4 trajectories. Corresponding mean square error between LE/GLEs and MD of position (g) and momentum (h) probability density function in time.

our methodology for both LDL and HDL cases. It is worth underlying that the Laplace

space is used only for comparison purpose, since it allows to extract $\boldsymbol{\theta}(t)$ numerically from MD data. From the comparison it emerges that the first order approximation obtained with the MLP outperforms the Markovian approximation at low λ, but is unable to catch the behaviour at high λ in the Laplace space. In contrast, the second order approximations exactly overlay with the MD results for the entire spectrum of λ. The accuracy in the memory kernel approximation directly affects the velocity correlation functions obtained with the different methods, which is shown in Figs 3.6(c-d). In the log-log diagram clearly emerges the limitations of LE, which is unable to accurately replicate the original correlation dynamics obtained with MD simulations. The first order approximation of the GLE, on the contrary, is fairly accurate, but diverges for long times. Finally, the GLE embedded with an higher approximation of the memory kernel follows the exact autocorrelation within a tolerance lower than 1%.

In Figs 3.7(a-b), we report the mean square displacement $\langle (x(t) - x(0))^2 \rangle$ computed with MD, LE and GLE in the LDL and HDL cases. From the comparison, it emerges that both the Markovian and the non-Markovian coarse-graining are able to accurately reproduce the mean square displacement.

Figs 3.7(c-d) show the values of the adaptive learning rate η during the MLP learning process. The log-log plot highlights the wide range of η values, that spans up to 8 orders of magnitude. This variability exemplifies the advantages of an adaptive learning rate over a fixed one. The error (or cost function) evolution during the learning process is reported in Figs 3.7(e-f). The monotonically decreasing trend of C at some point shows a plateau, which corresponds to an end of the learning process.

The effectiveness and the limitations of the coarse-graining out of equilibrium is also tested, by analysing the probability density function ρ. A target particle with zero initial position x and momentum p is immersed in an equilibrated bath of 699 particles identical to the one adopted in the LDL case at equilibrium. Then, 10^5 trajectories of the system relaxation to equilibrium are simulated. This relaxation corresponds to the evolution of a Dirac delta to the equilibrium distribution in the phase space. Similarly, the relaxation of ρ obtained by coarse-graining the bath with GLE and LE is followed. The comparison reported in Figs 3.8(a-f) shows that GLE, even if parametrized with a memory kernel evaluated in equilibrium conditions, significantly outperforms LE. As expected, at equilibrium the distributions obtained with MD, GLE and LE converge. During the relaxation, ρ relaxes faster for LE and GLE with respect to MD. A quantitative estimation of the accuracy of GLE in reproducing the density relaxation is provided by the mean square errors in position ϵ_q and momentum ϵ_p, shown in Figs 3.8(g-h). As expected, both errors are null at the beginning and, asymptotically, when the system

reaches equilibrium. During the first instants of the relaxation, the error reaches a peak, whose value for GLE is lower than LE of about 50% and 35% if considering ϵ_q and ϵ_p, respectively.

3.1.5. Numerical application II: Particle in a bistable potential

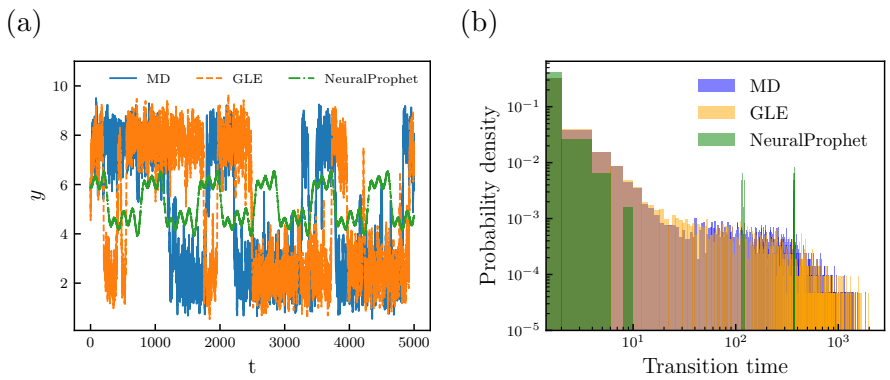

Figure 3.9. (a) Trajectory of a particle in a bistable potential simulated with MD (explicit bath particles), GLE embedded with a memory kernel approximated through our MLP (implicit bath particles) and a standard package using neural network for data series modelling (NeuralProphet). The comparison between transition time probability densities (b) shows that NeuralProphet cannot accurately reproduce the transition dynamics of the particle because of the non-seasonal behaviour of the original MD data series. On the contrary, the transition time distribution obtained with our GLE follows very closely the MD one.

As an additional validation of our GLE approach, we simulated a particle in a bath confined in a double well potential $U^{ext} = 10 \times (y^4 - y^2)$, and compared the GLE dynamics against MD, and against a neural-network forecasting of the dynamics by using NeuralProphet (a standard numerical library for data series modelling). NeuralProphet is based on an open-source software used for time data series forecasts by Facebook's core data science team [183]. It adopts an additive model where non-linear trends are fit with yearly, weekly, and daily seasonality, plus holiday effects. In Figure 3.9(a) we show a particle trajectory in the bistable potential simulated with MD (explicit bath particles), GLE embedded with a memory kernel approximated through our MLP (implicit bath particles) and NeuralProphet. This visualization clearly shows that NeuralProphet is not able to accurately replicate the original non-seasonal dynamics obtained with MD. In Figure 3.9(b), we also report the probability densities of the transition time, defined as the time difference between two consecutive crossings of the saddle point of the po-

Table 3.1. Values of interaction potentials parameters adopted to simulate a particle chain.

Parameters	K_H	K_γ	K_ϕ	r_0	γ_0	$\epsilon_{i,j}$	$\sigma_{i,j}$
Values	100	10	10	1.5	109.5	1	1

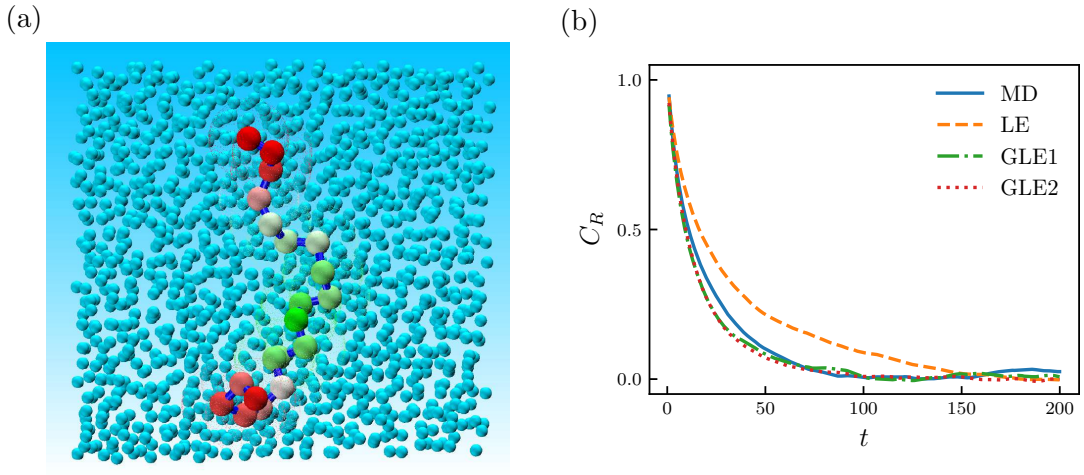

Figure 3.10. (a) Representation of the particle chain in the bath at equilibrium. (b) Time correlation of the gyration radius of the particle chain in a bath at equilibrium computed from LE, GLE and MD simulations.

tential. This comparison between MD, GLE and NeuralProphet confirms that standard packages, such as NeuralProphet, cannot detect and replicate the full kinetics of transition dynamics dominated by non-seasonal events. On the contrary, our GLE approach shows its high capabilities in reproducing the MD transition time.

3.1.6. Numerical application III: Particle chain in bath

As an additional test, we analyse the dynamics of a chain of $N = 20$ particles in a bath. A LJ potential v_{LJ} is used to model pairwise non bonded interactions among chain and bath particles. The chain particle interactions are modelled by the following multi-body Dreiding potential [184], already adopted in Ref. [185] to study polymer chain deformations,

$$v(\mathbf{r}_{i,j,k,l}) = v_{LJ}(\mathbf{r}_{ij}) + v_H(\mathbf{r}_{ij}) + v_\theta(\mathbf{r}_{ijk}) + v_\phi(\mathbf{r}_{ijkl}), \tag{3.17}$$

where $v_H(\mathbf{r}_{ij}) = k_H(\mathbf{r}_{ij} - \mathbf{r}_0)^2$, $v_\theta(r_{ijk}) = k_\theta(\theta_{ijk} - \theta_0)^2$ and $v_\phi(r_{ijkl}) = k_\phi(1 + cos(2\phi_{ijkl}))$ account for linear, angular and dihedral bonds, respectively. Linear covalent bonds are

modelled with the harmonic potential $v_H(\mathbf{r}_{ij}) = k_H(\mathbf{r}_{ij}-\mathbf{r}_0)^2$, where $\mathbf{r}_0$ is the equilibrium position and k_H is a positive constant. Similarly, angular covalent bonds are modelled by $v_\gamma(r_{ijk}) = k_\gamma(\gamma_{ijk}-\gamma_0)^2$, where γ_{ijk} is the angle in i formed by the particles i, j and k, γ_0 is the equilibrium angle and k_γ is a positive constant. Finally, we consider torsional (dihedral) bonds through the potential $v_\phi(r_{ijkl}) = k_\phi(1+cos(2\phi_{ijkl}))$, with ϕ_{ijkl} being the angle between the two planes defined by $\{\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k\}$ and $\{\mathbf{r}_j, \mathbf{r}_k, \mathbf{r}_l\}$ respectively, and k_ϕ being a positive parameter. Table 3.1 reports the values of the interaction parameters adopted in the present work. The bath contains $69,900$ particles interacting with Lennard-Jones potential v_{LJ}. The simulation box measures $50\sigma \times 50\sigma \times 40\sigma$, and periodic boundary conditions are imposed along x, y and z axes. A Nosé-Hoover thermostat is used to equilibrate the system at a reduced temperature $T = 1$ with a time step $\Delta t = 10^{-2}$. The following procedure is followed to run the MD simulations. First, the bath particles are randomly generated inside the simulation box. Then, the chain particle are placed along a straight line, and a minimization algorithm is employed to avoid particle overlaps. Hence, a run of 1.5×10^5 time steps is used to equilibrate the system. Finally, data are gathered over 2×10^7 time steps.

The bath has the same characteristics (density $\rho = 0.699$, temperature $T = 1$ and interaction potential v_{LJ}) of the LDL case for the single particle. This choice, together with the assumption that the potential of mean force acting among the chain particles is approximately equal to $v(\mathbf{r}_{i,j,k,l})$, allows us to use the same memory kernel obtained for the single particle (see Fig. 3.6). Particle chains are usually used to model polymers, whose characteristic dimensions are described by the gyration radius, defined as $R_G^2 = \frac{1}{N}\sum_{k=1}^{N}(\mathbf{r}_k-\mathbf{r}_{\mathrm{CM}})^2$, where $\mathbf{r}_k$ and $\mathbf{r}_{\mathrm{CM}}$ are the position of the particle k and of the centre of mass of the chain respectively. Figure 3.10 shows the radius of gyration autocorrelation, computed as [186, 187] $C_R = \left[\langle R_G(t)^2 R_G(0)^2\rangle - \langle R_G(0)^2\rangle^2\right]\left[\langle R_G(0)^4\rangle - \langle R_G(0)^2\rangle^2\right]^{-1}$, for the particle chain dynamics at equilibrium simulated with LE, GLE and MD. It is interesting to observe that GLE, already with a single neuron, is able to accurately reproduce the bath effects on the chain and outperforms the commonly used Markovian approximation.

3.1.7. Numerical application IV: Modelling global temperature

Several stochastic models have been developed to reproduce and forecast global and local temperature dynamics, for example in Refs [188, 189, 190]. In the present work, we show that GLE, parametrized through our method, is able to accurately model the global daily temperature fluctuations with respect to a properly chosen moving average. It is worth underlying that the same methodology could be also employed to model lo-

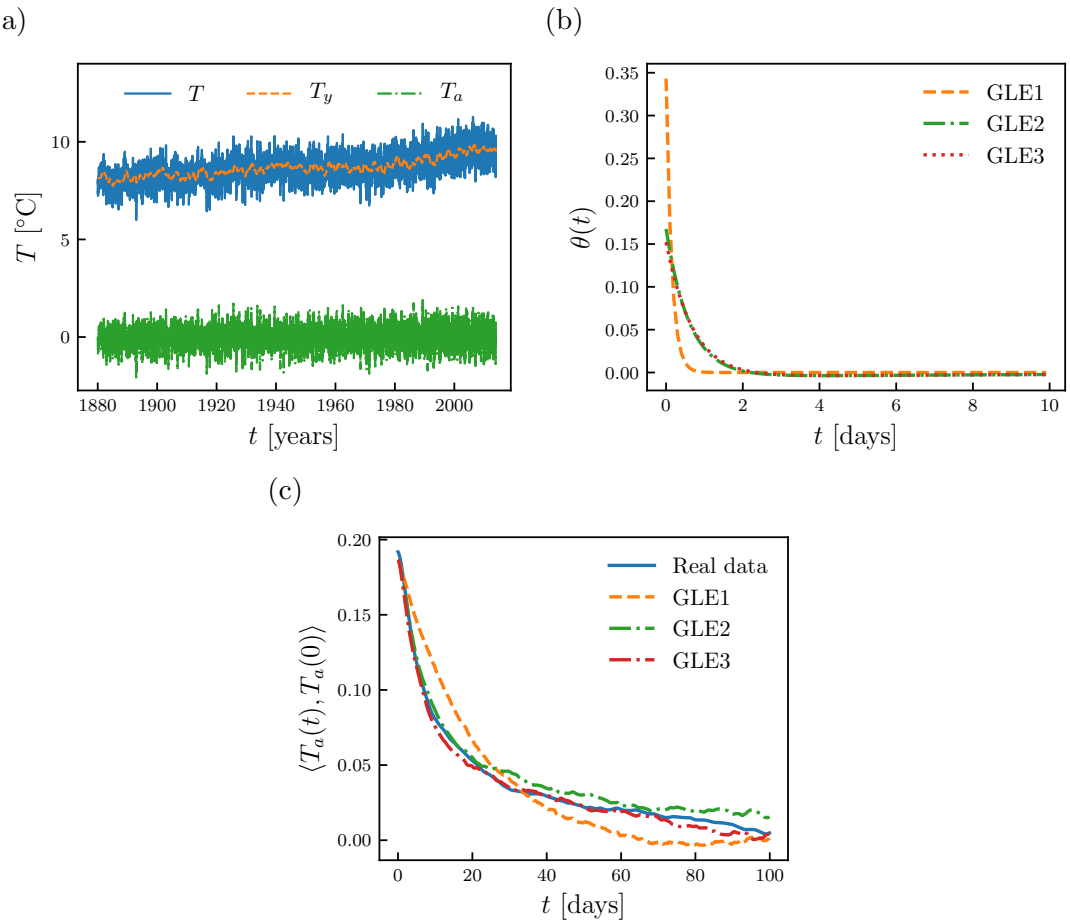

Figure 3.11. (a) Global temperature $T(t)$, annual moving average temperature $T_y(t)$ and the daily anomaly $T_a(t) = T(t) - T_y(t)$ between 1880 and 2014. (b) Memory kernel approximations computed through MLP with 1,2 and 3 neurons in the hidden layer and (c) corresponding time correlations obtained from real data and GLE simulations.

cal temperature dynamics. We consider the daily land-average global temperature $T(t)$ measured during the period 1880-2014, published by the Berkeley Earth [191, 192]. Despite the local temperature showing cyclical trends in the short period due, for instance, to season changes, global temperature does not exhibit a significant seasonal behavior, being a result of the energy balance between solar and earth radiations [193]. Nevertheless, $T(t)$ reveals non-stationarity features due to a long period increasing trend related to global warming, as visible in Fig. 3.11(a). Hence, we first compute the long term

dynamics T_y as an yearly moving average:

$$T_y(t) = \frac{1}{y} \sum_{i=t-y-1}^{t-1} T(i) \tag{3.18}$$

Then, we define the observable of interest as $T_a(t) = T(t) - T_y(t)$, so that the corresponding time series is stationary (see Appendix B.1.3). Consequently, we model $T_a(t)$ with a GLE in the form $\partial_t T_a(t) = -\int_0^t \theta(t-\tau) T_a(\tau) d\tau + R(t)$. In fact, this is a generalization of the Markovian model for weather derivatives proposed in Ref. [188].

In Fig. 3.11(b), we plot various degrees of approximation of the memory kernel extracted with our MLP-based method, while Fig. 3.11(c) shows the corresponding correlations functions. In first place, it emerges an excellent agreement between the correlations obtained with GLE dynamics and the real world data, especially when three neurons are adopted in the hidden layer. Then, matching the relaxation times of memory kernel ($\sim$ days) with the characteristic time of the variable T_y ($\sim$ years), we can obtain the evolution of $T(t)$ as sum of a Markovian yearly (long term) contribution and a non-Markovian daily (short term) term, namely:

$$\partial_t T(t) = (\partial_t + \theta_c)\, T_y(t) - \int_0^t \theta(t-\tau) T(\tau) d\tau + R(t), \tag{3.19}$$

where we introduced the constant $\theta_c = \int_0^t \theta(\tau)$. Eq. (3.19), originating directly from data, reflects the main features of global temperature multi-scale dynamics.

3.1.8. Numerical application V: A stock market model

In more than one study, stochastic models have been employed to model financial instruments, such as bonds and stock prices [194, 195, 196]. In fact, operations such as financial risk management and portfolio optimization require accurate predictions of markets dynamics to maximize profits. However, most of the models used in finance relies on Markovian assumptions, which can potentially introduce inaccuracies. A possible way to overcome such limitations is employing a GLE model. Here we show how the methodology can be applied to model the daily price of the Japanese financial index Nikkei $NI(t)$ between May 1949 and May 2018 [197]. As many other financial instruments, $NI(t)$ exhibits a non-stationary behavior in both mean and variance. Thus, we build an observable defined as $NI_a(t) = [NI(t) - NI_y(t)]\,/\sigma_y(t)$, with $NI_y(t)$ and $\sigma_y(t)$ being respectively a moving average and a moving standard deviation computed over a

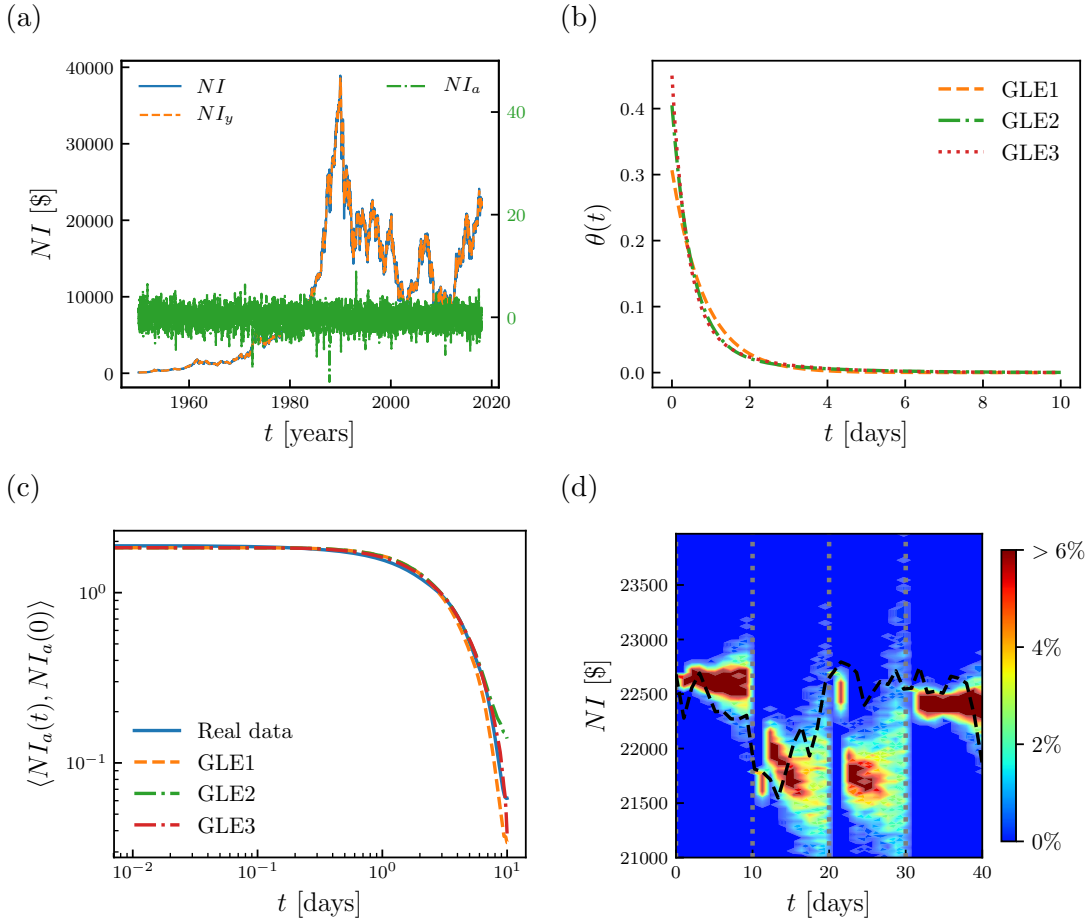

Figure 3.12. (a) Daily close price of Nikkei index $NI(t)$, moving average index $N_y(t)$ computed over a window of 10 days preceding the time t and the normalized index $NI_a(t)$ between 1949 and 2018. (b) Memory kernel approximations computed through MLP with 1,2 and 3 neurons in the hidden layer and (c) the corresponding time correlations obtained from real data and GLE simulation. (d) Comparison between predicted probability distribution (colour-map) and actual market data (dashed black line). Dotted lines in grey delineate the 10 days long investment windows.

period $[t-y, t-1]$ as:

$$NI_y(t) = \frac{1}{y} \sum_{i=t-y-1}^{t-1} NI(i) \tag{3.20}$$

$$\sigma_y(t) = \sqrt{\frac{1}{y} \sum_{i=t-y-1}^{t-1} \left(NI(i) - NI_y(t)\right)^2} \tag{3.21}$$

The parameter y is then properly chosen in order to obtain a stationary $NI_a(t)$; In this work we adopt $y = 10$ days. Hence, we model the normalized stock price $NI_a(t)$ with the following non-Markovian model $\partial_t NI_a(t) = -\int_0^t \theta(t-\tau)\ NI_a(\tau)d\tau + R(t)$.

In Figs 3.12(b-c) we report various degrees of approximations obtained with our method and the corresponding correlations functions. In contrast with the global temperature trend, $NI_a(t)$ do not exhibit a clear time-scale separation between memory kernel and autocorrelation decays. The comparisons between the correlations obtained with GLE dynamics and the real data exhibit a growing accuracy with an increasing number of neurons in the hidden layer. In fact, with the third order approximation we are able to reproduce the correlation decay with a maximum relative errors of order 10^{-2}. The proposed GLE model, parametrized with our MLP equipped with 3 neurons, is employed in a comparison between the predicted probability distribution and actual market data for four time windows, each ten market days long, between 15 Jun 2018 and 10 Aug 2018 (Fig. 3.12(d)). It emerges that our model is able not only to predict most of the actual market trend, but, more important, gives quite accurate information on the local variance of the trend, thus giving the chance of optimizing risk management in short term ($\sim$ weakly) investments.

3.2. Macroscopic coarse-graining: From GLE to non-Markovian FDDFT [3]

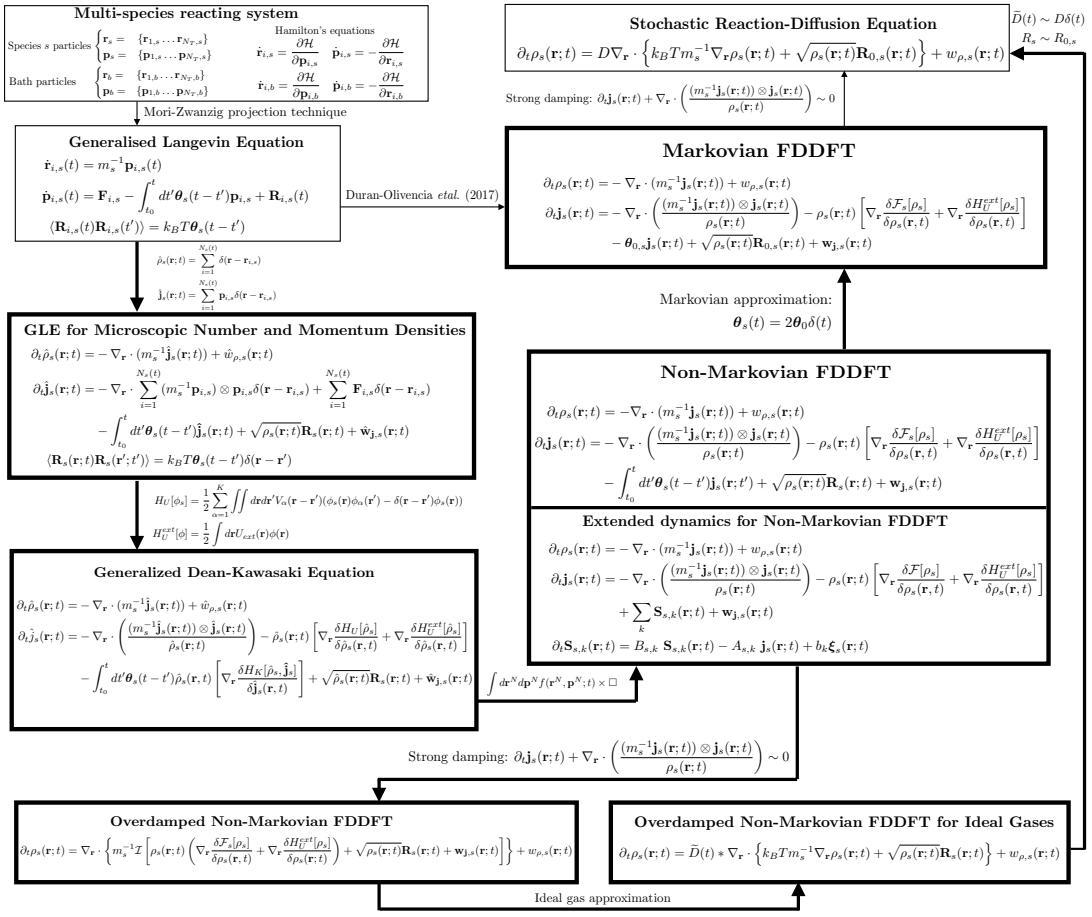

Figure 3.13. Flowchart showing the approach followed to obtain the non-Markovian FDDFT and its over-damped approximation. Connection to previous formulations are also included. Arrows denotes the relations among the different approaches. Thin boxes/arrows: previous approaches. Thick boxes: main results of this work.

The Langevin approach is still atomistic, and thus essentially equivalent to brute-force simulations in computational complexity. A continuum framework capable of describing small-scale systems is offered by dynamical density functional theory (DDFT). Unlike

[3] This section includes material from A. Russo, M. A. Durán-Olivencia, P. Yatsyshin & S. Kalliadasis. Non-Markovian dynamical density functional theory for reacting multi-component systems: theoretical framework and finite volume simulations, *J. Phys. A: Math. Theor.*, 53, 445007 (2020).

the atomistic GLEs, DDFT operates with locally averaged densities and is thus computationally more tractable. DDFT is usually obtained by deriving the Fokker-Planck equation governing the evolution of the system probability density function, which is then averaged over all but one degrees of freedom [56, 61, 62, 63]. Recently, several extensions of DDFT were proposed for multi-component systems [62] to include hydrodynamic interactions [56, 63, 64] and effects of particle orientability [65].

The inclusion of fluctuations in the DDFT framework, which is derived through a mean field approach, has posed many questions [66]. In some studies, such as the work by Elder *et al.* [67], an additional noise term was artificially included. Nevertheless, the physical meaning of this term is dubious. This long standing debate was clarified by the recent derivation of Duran-Olivencia *et al.* [68], who derived fluctuating DDFT (FDDFT) from first principles. In the present section, we extend the derivation by Duran-Olivencia *et al.* [68] to include memory effects in FDDFT for multi-component reacting system.

Our starting point is represented by microscopic GLEs governing the time-evolution of particles belonging to a multi-phase reacting colloidal system in a thermal bath. We then derive the non-Markovian FDDFT governing the dynamics of density and momentum fields. In the overdamped limit, the latter reduces to a single equation for the time-dependent one-body density field. In general, solving the non-Markovian FDDFT is computationally quite challenging, due to the memory term and the time-correlated noise. However, in case of memory terms in the form of decaying exponential series, these issues are overcome by introducing an additional field, whose time evolution accounts for memory term and time-correlations in the stochastic term. Such an approach will be denoted below as extended field dynamics. The resulting non-Markovian DDFT equations are solved using a finite volumes approach. We consider several applications to multi-component reacting systems, including binary configurations leading to the appearance of Turing patterns [198, 199, 200, 201, 202, 203, 204, 205, 206].

The section is organized as follows. In Sect. 3.2.1, we introduce GLEs for an isothermal multi-component system of particles, which provide the starting point for our derivation. Sect. 3.2.2 includes the discussion of microscopic density and momentum fields, and the derivation of the FDDFT equations for a multi-species reacting system. A non-trivial deconvolution operator has to be defined to take the overdamped limit of the previously derived FDDFT in Sect. 3.2.3. The computational challenges due to both the memory term and the correlated noise are overcome by introducing an extended field variables dynamics framework in Sect. 3.2.4. Finite volume method is then used for the spatio-temporal discretization of the extended variable dynamics in Sect. 3.2.5. Sect. 3.2.6

reports several numerical experiments, including a comparison between MD and FH that validates the main results of the work and a study of memory-driven Turing patterns.

3.2.1. Generalized Langevin equations

Consider an isothermal system, which at time t contains a total of $N_{t,s}$ particles, across K different species. We denote the phase-space coordinates of species s as $\mathbf{z}_s = \{\mathbf{r}_s, \mathbf{p}_s\}$, with $\mathbf{r}_s = \{\mathbf{r}_{1,s} \dots \mathbf{r}_{N_{t,s}}\}$ and $\mathbf{p}_s = \{\mathbf{p}_{1,s} \dots \mathbf{p}_{N_{t,s}}\}$ being the particle positions and momenta, respectively. In many cases, such as colloids in suspension, one is interested in the dynamical evolution of only the first N_s of the original $N_{t,s}$ particles. These can be shown to obey the following GLE [34, 40, 43]:

$$\dot{\boldsymbol{r}}_{i,s}(t) = m_s^{-1}\boldsymbol{p}_{i,s}(t)\,, \tag{3.22}$$

$$\dot{\boldsymbol{p}}_{i,s}(t) = \boldsymbol{F}_{i,s} - \int_{t_0}^{t} dt' \boldsymbol{\theta}_s(t-t')\boldsymbol{p}_{i,s} + \boldsymbol{R}_{i,s}(t), \tag{3.23}$$

where $\boldsymbol{F}_{i,s} = -\boldsymbol{\nabla}_{\boldsymbol{r}_{i,s}} U(\boldsymbol{r}_{i,s})$ is the force experienced by the particle, with $U(\boldsymbol{r}_{i,s})$ being the potential of mean force in the form:

$$U(\boldsymbol{r}_{i,s}, \boldsymbol{r}^N) = \sum_{\alpha=1,\alpha\neq s}^{K} \sum_{j=1}^{N_\alpha} V_\alpha(\boldsymbol{r}_{i,s} - \boldsymbol{r}_{j,\alpha}) + \sum_{j=1, j\neq i}^{N_s(t)} V_s(\boldsymbol{r}_{i,s} - \boldsymbol{r}_{j,s}) + U_{ext}(\boldsymbol{r}_{i,s})\,. \tag{3.24}$$

The memory kernel tensor $\boldsymbol{\theta}_s(t-t')$ in the convolution term of Eq. (3.23) accounts for the effects of the thermal bath on the particle dynamics [37, 38, 39, 40]. The stochastic term $\boldsymbol{R}_{i,s}(t)$ in Eq. (3.23) has a Gaussian distribution with mean zero and time correlation function satisfying the fluctuation dissipation theorem, namely

$$\langle \boldsymbol{R}_{i,s}(t) \rangle = 0, \tag{3.25}$$

$$\langle \boldsymbol{R}_{i,s}(t)\boldsymbol{R}_{i,s}(t') \rangle = m_s k_\mathrm{B} T \boldsymbol{\theta}_s(t-t')\,, \tag{3.26}$$

where k_B is the Boltzmann's constant. It is worth underlying that in Eqs (3.22)-(3.23) we have assumed that particles belonging to the same species s all have the same mass m_s, memory kernel tensor $\boldsymbol{\theta}_s$ and interact via a pairwise potential U_s.

Within the Markovian approximation, the characteristic time of the momentum correlation decay is regarded to be much smaller than the typical time scale of the system dynamics. In this case the memory kernel tensor can be approximated as $\boldsymbol{\theta}_s(t) = 2\boldsymbol{\theta}_0\delta(t)$, where $\delta(t)$ is Dirac's delta function and $\boldsymbol{\theta}_0$ is constant. When no clear scale separation between colloidal and bath particles is present, the Markovian hypothesis is inapplicable,

and one must seek approximating the memory kernel [36, 40, 44, 48].

3.2.2. Non-Markovian FDDFT

The microscopic number and momentum density fields of species s, containing a time-dependent number of particles $N_s(t)$, can be defined as follows:

$$\hat{\rho}_s(\boldsymbol{r};t) = \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r} - \boldsymbol{r}_{i,s}) \,, \tag{3.27}$$

$$\hat{\boldsymbol{j}}_s(\boldsymbol{r};t) = \sum_{i=1}^{N_s(t)} \boldsymbol{p}_{i,s}\delta(\boldsymbol{r} - \boldsymbol{r}_{i,s}) \,, \tag{3.28}$$

where N_s is the number of particles in species s.

In order to obtain the number and momentum densities, Eqs (3.27)-(3.28) must be differentiated with respect to time, which can be done by applying the discrete Leibniz integral rule:

$$\partial_t \sum_{i=1}^{N_s(t)} \boldsymbol{f}(i;t)\delta(\boldsymbol{r} - \boldsymbol{r}_{i,s}) = \sum_{i=1}^{N_s(t)} \partial_t \left[\boldsymbol{f}(i;t)\delta(\boldsymbol{r} - \boldsymbol{r}_{i,s})\right] + \boldsymbol{f}(N_s;t)\delta(\boldsymbol{r} - \boldsymbol{r}_{N_s,s})\frac{dN_s(t)}{dt} \,. \tag{3.29}$$

The second term on the right hand side of Eq. (3.29) represent the reaction contribution, which is stochastic in nature and can be determined empirically [71].

Applying the time derivative (Eq. (3.29)) to Eqs (3.27) and (3.28), and using Eqs (3.22) and (3.23), we obtain the time evolution of the number and momentum density fields of each species s:

$$\partial_t \hat{\rho}_s(\boldsymbol{r};t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (m_s^{-1}\hat{\boldsymbol{j}}_s(\boldsymbol{r};t)) + \hat{w}_{\rho,s}(\boldsymbol{r};t) \,, \tag{3.30}$$

$$\begin{aligned}\partial_t \hat{\boldsymbol{j}}_s(\boldsymbol{r};t) = &-\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \sum_{i=1}^{N_s(t)} (m_s^{-1}\boldsymbol{p}_{i,s}) \otimes \boldsymbol{p}_{i,s}\delta(\boldsymbol{r} - \boldsymbol{r}_{i,s}) + \sum_{i=1}^{N_s(t)} \boldsymbol{F}_{i,s}\delta(\boldsymbol{r} - \boldsymbol{r}_{i,s}) \\ &- \int_{t_0}^{t} dt' \boldsymbol{\theta}_s(t-t')\hat{\boldsymbol{j}}_s(\boldsymbol{r};t) + \boldsymbol{\eta}_s(\boldsymbol{r};t) + \hat{\boldsymbol{w}}_{\boldsymbol{j},s}(\boldsymbol{r};t) \,,\end{aligned} \tag{3.31}$$

where $\hat{w}_{\rho,s}$ and $\hat{\boldsymbol{w}}_{\boldsymbol{j},s}$ are the contributions to density and momentum due to chemical

reactions. The local random force $\boldsymbol{\eta}_s$ is given as

$$\boldsymbol{\eta}_s(\boldsymbol{r};t) = \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r}-\boldsymbol{r}_{i,s})\boldsymbol{R}_{i,s}(t)\,. \tag{3.32}$$

If reactions can occur only between resolved particles, additionally the conservation of mass and momentum of the total system must be imposed:

$$\sum_{\alpha=1}^{K} \hat{w}_{\rho,\alpha}(\boldsymbol{r};t) = 0\,, \tag{3.33}$$

$$\sum_{\alpha=1}^{K} \hat{\boldsymbol{w}}_{\boldsymbol{j},s}(\boldsymbol{r};t) = \boldsymbol{0}\,, \tag{3.34}$$

Finally, we can rewrite the stochastic term without changing its statistical properties (see Appendix B.2.1) as

$$\boldsymbol{\eta}_s(\boldsymbol{r};t) = \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t)\,, \tag{3.35}$$

where $\boldsymbol{R}_s$ denotes the spatial-temporal stochastic process satisfying the fluctuation-dissipation theorem

$$\langle \boldsymbol{R}_s(\boldsymbol{r};t)\boldsymbol{R}_s(\boldsymbol{r}';t')\rangle = m_s k_\mathrm{B} T\boldsymbol{\theta}_s(t-t')\delta(\boldsymbol{r}-\boldsymbol{r}'). \tag{3.36}$$

The first and the second terms on the right-hand side of Eq. (3.31) can be rewritten in terms of $\hat{\rho}_s(\boldsymbol{r};t)$ and $\hat{j}_s(\boldsymbol{r};t)$. Specifically, the second term with the force $\boldsymbol{F}_{i,s}$, experienced by the particle can be further rewritten as

$$\begin{aligned}\sum_{i=1}^{N_s(t)} \boldsymbol{\nabla}_{\boldsymbol{r}_{i,s}} U(\boldsymbol{r}_{i,s},\boldsymbol{r}^N)\delta(\boldsymbol{r}-\boldsymbol{r}_{i,s}) = \hat{\rho}_s(\boldsymbol{r};t)\Big[\boldsymbol{\nabla}_{\boldsymbol{r}} U_{ext}(\boldsymbol{r};t) + \\ \sum_{\alpha=1}^{K}\int d\boldsymbol{r}'\boldsymbol{\nabla}_{\boldsymbol{r}} V_\alpha(\boldsymbol{r}-\boldsymbol{r}')\left(\hat{\rho}_\alpha(\boldsymbol{r}';t)-\delta(\boldsymbol{r}-\boldsymbol{r}')\right)\Big]\,.\end{aligned} \tag{3.37}$$

Moreover, the convective flux becomes [68, 207]

$$\sum_{i=1}^{N_s(t)} (m_s^{-1}\boldsymbol{p}_{i,s}) \otimes \boldsymbol{p}_{i,s}\delta(\boldsymbol{r}-\boldsymbol{r}_i) = \frac{\sum_{i=1}^{N_s(t)}\sum_{j=1}^{N_s(t)} (m_s^{-1}\boldsymbol{p}_{i,s}) \otimes \boldsymbol{p}_{i,s}\delta(\boldsymbol{r}-\boldsymbol{r}_{i,s})\delta(\boldsymbol{r}-\boldsymbol{r}_{j,s})}{\hat{\rho}_s(\boldsymbol{r};t)} = \frac{(m_s^{-1}\hat{\boldsymbol{j}}_s(\boldsymbol{r};t)) \otimes \hat{\boldsymbol{j}}_s(\boldsymbol{r};t)}{\hat{\rho}_s(\boldsymbol{r};t)} . \tag{3.38}$$

Thus, the fluctuating momentum equation reads

$$\begin{aligned}\partial_t \hat{j}_s(\boldsymbol{r};t) = &- \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left(\frac{(m_s^{-1}\hat{\boldsymbol{j}}_s(\boldsymbol{r};t)) \otimes \hat{\boldsymbol{j}}_s(\boldsymbol{r};t)}{\hat{\rho}_s(\boldsymbol{r};t)} \right) - \hat{\rho}_s(\boldsymbol{r};t)(\boldsymbol{\nabla}_{\boldsymbol{r}} U_{ext}(\boldsymbol{r};t) \\ &- \hat{\rho}_s(\boldsymbol{r};t) \sum_{\alpha=1}^{K} \int d\boldsymbol{r}' \boldsymbol{\nabla}_{\boldsymbol{r}} V_\alpha(\boldsymbol{r}-\boldsymbol{r}') \left(\hat{\rho}_\alpha(\boldsymbol{r}';t) - \delta(\boldsymbol{r}-\boldsymbol{r}') \right) \\ &- \int_{t_0}^{t} dt' \boldsymbol{\theta}_s(t-t') \hat{\boldsymbol{j}}_s(\boldsymbol{r};t) + \sqrt{\hat{\rho}_s(\boldsymbol{r};t)} \boldsymbol{R}_s(\boldsymbol{r};t) + \hat{\boldsymbol{w}}_{\boldsymbol{j},s}(\boldsymbol{r};t) .\end{aligned} \tag{3.39}$$

The following definition of the internal and external energy functionals

$$H_U[\phi_s] = \frac{1}{2} \sum_{\alpha=1}^{K} \iint d\boldsymbol{r} d\boldsymbol{r}' V_\alpha(\boldsymbol{r}-\boldsymbol{r}')(\phi_s(\boldsymbol{r})\phi_\alpha(\boldsymbol{r}') - \delta(\boldsymbol{r}-\boldsymbol{r}')\phi_s(\boldsymbol{r})) , \tag{3.40}$$

$$H_U^{ext}[\phi] = \frac{1}{2} \int d\boldsymbol{r} U_{ext}(\boldsymbol{r})\phi(\boldsymbol{r}) , \tag{3.41}$$

allows us to rewrite Eq. (3.39) in terms of functional derivative of Eq. (3.40) and Eq. (3.41) as

$$\begin{aligned}\partial_t \hat{j}_s(\boldsymbol{r};t) = &- \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left(\frac{(m_s^{-1}\hat{\boldsymbol{j}}_s(\boldsymbol{r};t)) \otimes \hat{\boldsymbol{j}}_s(\boldsymbol{r};t)}{\hat{\rho}_s(\boldsymbol{r};t)} \right) - \hat{\rho}_s(\boldsymbol{r};t) \left[\boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta H_U[\hat{\rho}_s]}{\delta \hat{\rho}_s(\boldsymbol{r},t)} + \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta H_U^{ext}[\hat{\rho}_s]}{\delta \hat{\rho}_s(\boldsymbol{r},t)} \right] \\ &- \int_{t_0}^{t} dt' \boldsymbol{\theta}_s(t-t') \hat{\hat{\boldsymbol{j}}}_s + \sqrt{\hat{\rho}_s(\boldsymbol{r};t)} \boldsymbol{R}_s(\boldsymbol{r};t) + \hat{\boldsymbol{w}}_{\boldsymbol{j},s}(\boldsymbol{r};t) .\end{aligned} \tag{3.42}$$

Notice that Eq. (3.42) still depends on all particle position and momenta through the terms $\hat{\rho}_s(\boldsymbol{r};t)$ and $\hat{\boldsymbol{j}}_s(\boldsymbol{r};t)$. Thus, an ensemble average is required to obtain a mesoscopic description of the system, as highlighted in [68]. To this end, we introduce a probability density function $f(\boldsymbol{r}^N, \boldsymbol{p}^N; t)$ for the system at local equilibrium:

$$f(\boldsymbol{r}^N, \boldsymbol{p}^N; t) = \mathcal{Z}^{-1} e^{-\beta \mathcal{H}(\boldsymbol{r}^N, \boldsymbol{p}^N; t)} \tag{3.43}$$

where $\mathcal{Z}$ is the partition function, $\beta = (k_{\mathrm{B}}T)^{-1}$ and $\mathcal{H}(\boldsymbol{r}^N, \boldsymbol{p}^N; t)$ is the system Hamiltonian given by:

$$\mathcal{H} = \frac{1}{2}\int d\boldsymbol{r} \sum_{i=1}^{N_s(t)} \Bigg[m^{-1}\boldsymbol{p}_i \cdot \boldsymbol{p}_i + \sum_{\alpha=1,\alpha\neq s}^{K} \sum_{j=1}^{N_\alpha(t)} V_\alpha(\boldsymbol{r}_{i,s} - \boldsymbol{r}_j) + \sum_{j=1, j\neq i}^{N_s(t)} V_s(\boldsymbol{r}_{i,s} - \boldsymbol{r}_j) + U_{ext}(\boldsymbol{r}_{i,s}) \Bigg] \delta(\boldsymbol{r} - \boldsymbol{r}_i). \tag{3.44}$$

Hence, the ensemble average of density, momenta and reaction sources gives:

$$\begin{pmatrix} \hat{\rho}_s(\boldsymbol{r};t) \\ \left\langle \begin{matrix} \hat{w}_{\rho,s}(\boldsymbol{r};t) \\ \hat{\boldsymbol{j}}_s(\boldsymbol{r};t) \end{matrix} \right\rangle \\ \hat{\boldsymbol{w}}_{\boldsymbol{j},s}(\boldsymbol{r};t) \end{pmatrix} = \int f(\boldsymbol{r}^N, \boldsymbol{p}^N; t) \begin{pmatrix} \hat{\rho}_s(\boldsymbol{r};t) \\ \hat{w}_{\rho,s}(\boldsymbol{r};t) \\ \hat{\boldsymbol{j}}_s(\boldsymbol{r};t) \\ \hat{\boldsymbol{w}}_{\boldsymbol{j},s}(\boldsymbol{r};t) \end{pmatrix} d\boldsymbol{r}^N d\boldsymbol{p}^N = \begin{pmatrix} \rho_s(\boldsymbol{r};t) \\ w_{\rho,s}(\boldsymbol{r};t) \\ \boldsymbol{j}_s(\boldsymbol{r};t) \\ \boldsymbol{w}_{\boldsymbol{j},s}(\boldsymbol{r};t) \end{pmatrix} \tag{3.45}$$

Ensemble averaging Eq. (3.42), we obtain the mesoscopic FDDFT for species s:

$$\partial_t \rho_s(\boldsymbol{r};t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (m_s^{-1}\boldsymbol{j}_s(\boldsymbol{r};t)) + w_{\rho,s}(\boldsymbol{r};t) \tag{3.46}$$

$$\partial_t \boldsymbol{j}_s(\boldsymbol{r};t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left(\frac{(m_s^{-1}\boldsymbol{j}_s(\boldsymbol{r};t)) \otimes \boldsymbol{j}_s(\boldsymbol{r};t)}{\rho_s(\boldsymbol{r};t)} \right) - \rho_s(\boldsymbol{r};t) \left[\boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta \mathcal{F}_s[\rho_s]}{\delta \rho_s(\boldsymbol{r},t)} + \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta H_U^{ext}[\rho_s]}{\delta \rho_s(\boldsymbol{r},t)} \right] - \int_{t_0}^{t} dt' \boldsymbol{\theta}_s(t-t') \boldsymbol{j}_s(\boldsymbol{r};t') + \sqrt{\rho_s(\boldsymbol{r};t)} \boldsymbol{R}_s(\boldsymbol{r};t) + \boldsymbol{w}_{\boldsymbol{j},s}(\boldsymbol{r};t)\,. \tag{3.47}$$

Above we introduced the Helmholtz free-energy functional, defined as sum of ideal and excess-over-ideal contributions $\mathcal{F}_s[\rho] = \mathcal{F}_s^{id}[\rho_s] + \mathcal{F}_s[\rho_s]^{ex}$:

$$\mathcal{F}_s^{id}[\rho_s] = k_{\mathrm{B}}T \int d\boldsymbol{r} \rho_s(\boldsymbol{r};t) \left(\log\left(\Lambda^3 \rho_s(\boldsymbol{r};t)\right) - 1 \right) \tag{3.48}$$

where Λ is the thermal de Broglie wavelength. The excess part of the free-energy satisfies the relation:

$$\rho_s(\boldsymbol{r};t) \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta \mathcal{F}_s^{ex}[\rho_s]}{\delta \rho_s(\boldsymbol{r},t)} = \left\langle \hat{\rho}_s(\boldsymbol{r};t) \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta H_U[\hat{\rho}_s]}{\delta \hat{\rho}_s(\boldsymbol{r},t)} \right\rangle = \sum_{\alpha=1}^{K} \int d\boldsymbol{r}' \boldsymbol{\nabla}_{\boldsymbol{r}} V_\alpha(\boldsymbol{r} - \boldsymbol{r}') \rho_{s,\alpha}^{(2)}(\boldsymbol{r}, \boldsymbol{r}';t) \tag{3.49}$$

where $\rho_{s,\alpha}^{(2)}(\boldsymbol{r}, \boldsymbol{r}';t) = \langle \rho_s(\boldsymbol{r};t) \rho_\alpha(\boldsymbol{r}';t) \rangle$ is the inter-species pair-correlation function. Notice that in local equilibrium, the term containing the Helmholtz free-energy is associated with pressure. In fact, the Gibbs–Duhem relation ($dp = \sum_s \rho_s d\mu_s$), and the

Euler–Lagrange equation of equilibrium DFT ($\mu_s = \delta\mathcal{F}_s[\rho_s]/\delta\rho_s$) lead to the following:

$$\boldsymbol{\nabla}_{\boldsymbol{r}} p = \sum_s \rho_s(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta\mathcal{F}_s[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)}, \tag{3.50}$$

which is consistent with the well-known expressions for one-component systems [53, 68].

3.2.3. Markovian and overdamped limits: Connections with DDFT and reaction-diffusion equations

In many cases the memory kernel decays to zero on a timescale much smaller than the characteristic time scale of the density dynamics. Under such circumstances, the density field relaxes much slower than the momentum field, and it is possible to assume that the material derivative of the momentum $D_t\boldsymbol{j}_s \sim 0$. This approximation allows one to reduce the equations of fluctuating hydrodynamics to a single evolution equation for the density. Our system of equations, however, contains a non-Markovian memory term, which has a convolution structure and thus, makes the derivation more complex. Specifically, we define the following "deconvolution" operator

$$\mathcal{I}\left[\int_{t_0}^{t} dt'\phi(t-t')\boldsymbol{\pi}(\boldsymbol{r};t),\phi\right] = \boldsymbol{\pi}(\boldsymbol{r};t), \tag{3.51}$$

which is applied to both sides of the momentum equation (Eq. (3.47)). Assuming that $D_t\boldsymbol{j} \sim 0$, we obtain the non-Markovian time evolution of the density field of the species s in a multi-component system:

$$\begin{aligned}\partial_t\rho_s(\boldsymbol{r};t) =& \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left\{ m_s^{-1}\mathcal{I}\left[\rho_s(\boldsymbol{r};t)\left(\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{F}_s[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)} + \boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta H_U^{ext}[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)}\right)\right.\right. \\ &\left.\left. + \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t) + \boldsymbol{w}_{\boldsymbol{j},s}(\boldsymbol{r};t)\right]\right\} + w_{\rho,s}(\boldsymbol{r};t)\end{aligned} \tag{3.52}$$

In general, the operator $\mathcal{I}$ may be arbitrarily defined, as long as its definition satisfies Eq. (3.51). In this work, we can take advantage of the properties of Laplace transform

and define the operator $\mathcal{I}$ as:

$$\begin{aligned}\mathcal{I}\left[\int_{t_0}^{t} dt'\phi(t-t')\boldsymbol{\pi}(\boldsymbol{r};t),\phi\right] =&\mathcal{L}^{-1}\left\{\frac{\mathcal{L}\left[\int_{t_0}^{t} dt'\phi(t-t')\boldsymbol{\pi}(\boldsymbol{r};t)\right]}{\mathcal{L}\left[\phi(t)\right]}\right\}\\ &= \mathcal{L}^{-1}\left\{\frac{\mathcal{L}\left[\phi(t)\right]\mathcal{L}\left[\boldsymbol{\pi}(\boldsymbol{r};t)\right]}{\mathcal{L}\left[\phi(t)\right]}\right\} = \boldsymbol{\pi}(\boldsymbol{r};t)\end{aligned} \tag{3.53}$$

where $\mathcal{L}[\boldsymbol{\pi}(\boldsymbol{r};t)] = \int_0^\infty \boldsymbol{\pi}(\boldsymbol{r};t)e^{-t/\lambda}dt$ indicates the Laplace transform of a time dependent field.

Applying the deconvolution operator in Eq. (3.51) for a multi-component system of ideal gases, one obtains the following non-Markovian time evolution of the density field of the species s:

$$\begin{aligned}\partial_t\rho_s(\boldsymbol{r};t) = \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\left\{m_s^{-1}\mathcal{I}\Big[\boldsymbol{\nabla}_{\boldsymbol{r}}k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t) + \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t) + \boldsymbol{w}_{\boldsymbol{j},s}(\boldsymbol{r};t)\Big]\right\}\\ + w_{\rho,s}(\boldsymbol{r};t).\end{aligned} \tag{3.54}$$

If we neglect the reaction contributions to the momentum and we make use of the definition of the operator $\mathcal{I}$ in Eq. (3.53), we then obtain

$$\partial_t\rho_s(\boldsymbol{r};t) = \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\left\{m_s^{-1}\mathcal{L}^{-1}\left[\frac{\mathcal{L}\left[k_{\mathrm{B}}T\boldsymbol{\nabla}_{\boldsymbol{r}}\rho_s(\boldsymbol{r};t) + \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t)\right]}{\mathcal{L}\left[\theta(t)\right]}\right]\right\} + w_{\rho,s}(\boldsymbol{r};t). \tag{3.55}$$

By defining the function $\widetilde{D}(t) = \mathcal{L}^{-1}\left[\left(\mathcal{L}\left[\theta(t)\right]\right)^{-1}\right]$, and using the properties of Laplace transform, we get

$$\partial_t\rho_s(\boldsymbol{r};t) = \widetilde{D}(t) * \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\left\{\left(k_{\mathrm{B}}Tm_s^{-1}\boldsymbol{\nabla}_{\boldsymbol{r}}\rho_s(\boldsymbol{r};t) + \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t)\right)\right\} + w_{\rho,s}(\boldsymbol{r};t), \tag{3.56}$$

where $*$ stands for a convolution in time.

For a memory kernel in the form $\theta_s(t) = \sum_k^d A_{s,k}e^{B_{s,k}t}$, the function $\widetilde{D}(t)$ is given by

$$\widetilde{D}(t) = \mathcal{L}^{-1}\left[\left(\sum_k^d \frac{A_{s,k}\lambda}{1-B_{s,k}\lambda}\right)^{-1}\right]. \tag{3.57}$$

In the Laplace transform the highest values of λ correspond to the lowest frequencies,

which are usually dominant. This allows us to simplify the above even further:

$$\begin{aligned}\widetilde{D}(t) \sim \mathcal{L}^{-1}\left[\lim_{\lambda\to\infty}\left(\sum_k^d \frac{A_{s,k}\lambda}{1-B_{s,k}\lambda}\right)^{-1}\right] =& \mathcal{L}^{-1}\left[\left(-\sum_k^d \frac{A_{s,k}}{B_{s,k}}\right)^{-1}\right] \\ =& -\left(\sum_k^d \frac{A_{s,k}}{B_{s,k}}\right)^{-1}\delta(t).\end{aligned} \tag{3.58}$$

Finally, Eq. (3.56) reduces to

$$\partial_t \rho_s(\boldsymbol{r};t) = D_{eff}\boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\left\{k_{\mathrm{B}}T m_s^{-1}\boldsymbol{\nabla}_{\boldsymbol{r}}\rho_s(\boldsymbol{r};t) + \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t)\right\} + w_{\rho,s}(\boldsymbol{r};t), \tag{3.59}$$

where we have introduced the effective diffusion coefficient $D_{eff} = -\left(\sum_k^d \frac{A_{s,k}}{B_{s,k}}\right)^{-1}$.

An alternative derivation of Eq. 3.59 would be to first take the Markovian limit of an ideal gas model, followed by the overdamped limit, assuming that the momentum reaction rate $\boldsymbol{w}_{\boldsymbol{j},s}$ is negligible. In this way Eq. (3.52) would be reduced to the usual reaction-diffusion equation for the density [71]:

$$\partial_t \rho_s(\boldsymbol{r};t) = \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\left[k_{\mathrm{B}}T\left(m_s\boldsymbol{\theta}_{0,s}\right)^{-1}\left(\boldsymbol{\nabla}_{\boldsymbol{r}}\rho_s\right)\right] + \boldsymbol{\nabla}_{\boldsymbol{r}}\cdot\sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_{0,s}(\boldsymbol{r};t) + w_{\rho,s}(\boldsymbol{r};t). \tag{3.60}$$

3.2.4. Extended field dynamics

Numerical solution of the non-Markovian equations Eqs (3.46)-(3.47) is non-trivial. For one thing, it is not obvious how to sample time-correlated noise. Additionally, the direct evaluation of the convolution integral requires one to keep track of the entire history of the momentum field, which drives up the computational cost. Within the so-called extended dynamics approach, additional variables are introduced to capture the memory and noise effects on the observable [48, 181, 182]. In this section we derive extended dynamics for non-Markovian FDDFT, generalizing the work by Kawai [182]. For simplicity we assume that the global effects of the bath are isotropic [48] and the memory kernel can be expressed as an exponential series [181, 182]:

$$\boldsymbol{\theta}_s(t) = \theta_s(t)\mathbf{I} = \sum_k^d A_{s,k}e^{B_{s,k}t}\mathbf{I} \tag{3.61}$$

where $\mathbf{I}$ is the identity matrix, $A_{s,k}$ and $B_{s,k}$ are real-valued, and $B_{s,k} \leq 0$.

Convolution decomposition

The extended field variable associated with the k-th mode of the memory term of species s is defined as:

$$\boldsymbol{Z}_{s,k}(\boldsymbol{r};t) = \int_{t_0}^{t} dt' A_{s,k} e^{B_{s,k}(t-t')} \boldsymbol{j}_s(\boldsymbol{r};t')\,, \tag{3.62}$$

so that the convolution term can be re-written as $\int_{t_0}^{t} dt' \theta_s(t-t')\boldsymbol{j}_s(\boldsymbol{r};t') = \sum_k^d \boldsymbol{Z}_{s,k}(\boldsymbol{r};t)$. Differentiating $\boldsymbol{Z}_{s,k}$ with respect to t, we obtain the time evolution of the field $Z_{s,k}$ in form of a stochastic partial differential equation (SPDE):

$$\partial_t \boldsymbol{Z}_{s,k}(\boldsymbol{r};t) = B_{s,k}\ \boldsymbol{Z}_{s,k}(\boldsymbol{r};t) + A_{s,k}\ \boldsymbol{j}_s(\boldsymbol{r};t)\,. \tag{3.63}$$

Noise decomposition

The noise term is given by

$$\boldsymbol{\eta}_s(\boldsymbol{r};t) = \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t)\,, \tag{3.64}$$

where $\boldsymbol{R}_s$ denotes the spatial-temporal noise with correlations $\langle \boldsymbol{R}_s(\boldsymbol{r};t)\boldsymbol{R}_s(\boldsymbol{r}';t')\rangle = k_{\mathrm{B}}T\theta_s(t-t')\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}$. Because of the symmetry between t and t' in the fluctuation-dissipation theorem, $\theta_s(t)$ must be an even function of time: $\theta_s(t) = \theta_s(-t)$. Additionally, its Fourier transform $\widetilde{\theta}_s(\omega) = \int_{-\infty}^{-\infty} \theta_s(t)e^{-i\omega\tau}dt$ must be real and even too, when ω is real. It follows that in the ω-plane, the roots and singularities of $\widetilde{\theta}_s(\omega)$ must be symmetric with respect to real and imaginary axes. This allows us to introduce the function $\widetilde{\chi}(\omega)$, such that

$$\widetilde{\theta}_s(\omega) = 2\widetilde{\chi}(\omega)\widetilde{\chi}^T(-\omega), \tag{3.65}$$

where

$$\widetilde{\chi}(\omega) = \sum_k -i\left(\omega + iB_{s,k}\right)^{-1} b_{s,k} \tag{3.66}$$

with real coefficients $b_{s,k}$ and $B_{s,k}$. The singular points of $\widetilde{\chi}^{-1}(\omega)$ lie in the lower half of the complex ω-plane. Moreover, we define two function in the Fourier space:

$$\widetilde{\zeta}(\omega) = \widetilde{\chi}^{-1}(\omega), \tag{3.67}$$

and

$$\widetilde{k}_k(\omega) = -i\,(\omega + iB_{s,k})^{-1}\, b_{s,k}\widetilde{\zeta}(\omega), \tag{3.68}$$

and we denote their Fourier inverse transform with $h(t)$ and $k_k(t)$. Combining Eqs (3.66), (3.67) and (3.68), it follows that:

$$\sum_k \widetilde{k}_k(\omega) = 1 \tag{3.69}$$

or, equivalently,

$$\sum_k k_k(t) = \delta(t). \tag{3.70}$$

Moreover, Eq. (3.68) can be rewritten as $(i\omega - B_{s,k})\,\widetilde{k}_k(\omega) = b_{s,k}\widetilde{\zeta}(\omega)$, which in the time domain gives:

$$\frac{d}{dt}k_k(t) - B_{s,k}k_k(t) = b_{s,k}\zeta(t) \tag{3.71}$$

We introduce two more vector fields, corresponding to a delta correlated stochastic field (as proved in Appendix B.2.2) and an auxiliary stochastic field, respectively:

$$\boldsymbol{\xi}_s(\boldsymbol{r};t) = \int_0^{+\infty} \zeta(t-t')\boldsymbol{\eta}_s(\boldsymbol{r};t')dt' \tag{3.72}$$

and

$$\boldsymbol{\eta}_{s,k}(\boldsymbol{r};t) = \int_0^{+\infty} k_k(t-t')\boldsymbol{\eta}_s(\boldsymbol{r};t')dt'. \tag{3.73}$$

From Eqs (3.73) and (3.70) it follows that:

$$\sum_k^d \boldsymbol{\eta}_{s,k}(\boldsymbol{r};t) = \boldsymbol{\eta}_s(\boldsymbol{r};t), \tag{3.74}$$

while, combining Eq. (3.73) and (3.71), we obtain

$$\partial_t\boldsymbol{\eta}_{s,k}(\boldsymbol{r};t) = B_{s,k}\boldsymbol{\eta}_{s,k}(\boldsymbol{r};t) + b_{s,k}\boldsymbol{\xi}_s(\boldsymbol{r};t). \tag{3.75}$$

Eqs. (3.75) and (3.74) allow us to express the original correlated noise as a function of the white noise $\boldsymbol{\xi}(t)$. The discussion of the properties of the process $\boldsymbol{\xi}(\boldsymbol{r};t)$ can be found in Appendix B.2.2.

Non-Markovian FDDFT

After defining an extended field $\boldsymbol{S}_{s,k}(\boldsymbol{r};t) = -\boldsymbol{Z}_{s,k}(\boldsymbol{r};t) + \boldsymbol{\eta}_{s,k}(\boldsymbol{r};t)$, the FDDFT equations can be expressed in the following form:

$$\partial_t \rho_s(\boldsymbol{r};t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (m_s^{-1}\boldsymbol{j}_s(\boldsymbol{r};t)) + w_{\rho,s}(\boldsymbol{r};t) \tag{3.76}$$

$$\begin{aligned}\partial_t \boldsymbol{j}_s(\boldsymbol{r};t) = &-\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left(\frac{(m_s^{-1}\boldsymbol{j}_s(\boldsymbol{r};t)) \otimes \boldsymbol{j}_s(\boldsymbol{r};t)}{\rho_s(\boldsymbol{r};t)}\right) - \rho_s(\boldsymbol{r};t)\left[\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta \mathcal{F}[\rho_s]}{\delta \rho_s(\boldsymbol{r},t)} + \boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta H_U^{ext}[\rho_s]}{\delta \rho_s(\boldsymbol{r},t)}\right] \\ &+ \sum_k^d \boldsymbol{S}_{s,k}(\boldsymbol{r};t) + \boldsymbol{w}_{\boldsymbol{j},s}(\boldsymbol{r};t)\end{aligned} \tag{3.77}$$

$$\partial_t \boldsymbol{S}_{s,k}(\boldsymbol{r};t) = B_{s,k}\ \boldsymbol{S}_{s,k}(\boldsymbol{r};t) - A_{s,k}\ \boldsymbol{j}_s(\boldsymbol{r};t) + b_k \boldsymbol{\xi}_s(\boldsymbol{r};t) \tag{3.78}$$

where $\boldsymbol{\xi}(\boldsymbol{r}_1;t_1)$ is a stochastic process with zero mean and correlations $\langle \boldsymbol{\xi}_s(\boldsymbol{r};t_1)\boldsymbol{\xi}_s(\boldsymbol{r}';t_2)\rangle = 2k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t_1-t_2)\mathbf{I}$.

The system in Eq. (3.78) can be equivalently re-written in the following vector form:

$$\partial_t \mathcal{U}_s(\boldsymbol{r};t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \mathcal{M}_s(\boldsymbol{r};t) + \mathcal{N}_s(\boldsymbol{r};t) + \mathcal{R}_s(\boldsymbol{r};t) + \mathcal{Q}_s. \tag{3.79}$$

In the above equation $\mathcal{U}_s$ is the vector of unknowns, $\mathcal{M}_s$ is the conservative flux, and $\mathcal{Q}_s$ is the stochastic term for species s:

$$\mathcal{U}_s = \begin{pmatrix} \rho_s(\boldsymbol{r};t) \\ \boldsymbol{j}_s(\boldsymbol{r};t) \\ \boldsymbol{S}_{s,1}(\boldsymbol{r};t) \\ \cdots \\ \boldsymbol{S}_{s,k}(\boldsymbol{r};t) \\ \cdots \end{pmatrix}, \quad \mathcal{M}_s = \begin{pmatrix} (m_s^{-1}\boldsymbol{j}_s(\boldsymbol{r};t)) \\ \frac{(m_s^{-1}\boldsymbol{j}_s(\boldsymbol{r};t))\otimes \boldsymbol{j}_s(\boldsymbol{r};t)}{\rho_s(\boldsymbol{r};t)} \\ 0 \\ \cdots \\ 0 \\ \cdots \end{pmatrix}, \quad \mathcal{Q}_s = \begin{pmatrix} 0 \\ \mathbf{0} \\ b_1\ \boldsymbol{\xi}_s(\boldsymbol{r};t) \\ \cdots \\ b_k\ \boldsymbol{\xi}_s(\boldsymbol{r};t) \\ \cdots \end{pmatrix}. \tag{3.80}$$

The source terms are

$$\mathcal{N}_s = \begin{pmatrix} 0 \\ -\rho_s(\boldsymbol{r};t)\left[\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{F}[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)} + \boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta H_U^{ext}[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)}\right] + \sum_k^d \boldsymbol{S}_{s,k}(\boldsymbol{r};t) \\ B_{s,1}\ \boldsymbol{S}_{s,1}(\boldsymbol{r};t) - A_{s,1}\ \boldsymbol{j}_s(\boldsymbol{r};t) \\ \cdots \\ B_{s,k}\ \boldsymbol{S}_{s,k}(\boldsymbol{r};t) - A_{s,k}\ \boldsymbol{j}_s(\boldsymbol{r};t) \\ \cdots \end{pmatrix}, \quad \mathcal{R}_s = \begin{pmatrix} w_{\rho,s}(\boldsymbol{r};t) \\ \boldsymbol{w}_{\boldsymbol{j},s}(\boldsymbol{r};t) \\ 0 \\ \cdots \\ 0 \\ \cdots \end{pmatrix}. \tag{3.81}$$

3.2.5. Finite volume discretization

DFT system of equations and its corresponding deterministic and overdamped counterparts have been solved numerically using different technique, including pseudo-spectral method [57], and finite volume [71, 208, 209]. Because of its enhanced capability in dealing with non-regular fields and complex geometries, in this work we discretize Eq. (3.79) by employing finite volume method. For simplicity, we present the discretization for a one dimensional system, but the generalization to higher dimension is straightforward. A domain of length L is partitioned in n identical cells of width $\Delta x = L/n$. The value of the discretized variable vector for the species s at the cell $1 \leq j \leq n$ is then defined as the following average over the cell volume V_j

$$\mathcal{U}_{s,j}(t) = \frac{1}{\Delta x}\int_{V_j} \mathcal{U}_s(x,t)dx \tag{3.82}$$

Similarly, the time integration is performed by discretizing the total simulation time T in steps of magnitude Δt, such that we can assume $\mathcal{U}_{s,j}^m \equiv \mathcal{U}_{s,j}(m\Delta t)$, with $0 \leq m\Delta t \leq T$. The multiplicative white noise $\boldsymbol{\xi}_s(\boldsymbol{r};t)$ is discretized, using the following spatio-temporal average [71, 208]

$$\boldsymbol{\xi}_{s,j}^m(t) \equiv \frac{1}{\Delta x \Delta t}\int_{m\Delta t}^{(m+1)\Delta t}\int_{V_j} \boldsymbol{\xi}(x,t)dx \sim \sqrt{\frac{\rho_{s,j}(t)}{\Delta x \Delta t}}\ \boldsymbol{G} \tag{3.83}$$

where $\boldsymbol{G}$ represents a vector with component independently drawn from a standard Gaussian distribution $\mathcal{N}(0,1)$. In order to implement time stepping, we define the forward and backward conservative fluxes as:

$$\boldsymbol{\nabla}\cdot\mathcal{M}_s(x;t) = \frac{\mathcal{M}_s(x;t)}{dx} \sim \frac{\mathcal{M}_{s,j\pm1}(t) \mp \mathcal{M}_{s,j}(t)}{\Delta x} \tag{3.84}$$

where $\mathcal{M}_{s,j}(t)$ denotes the average value of the field $\mathcal{M}_s(x;t)$ at the cell j, i.e.

$$\mathcal{M}_s(\mathcal{U}_{s,j}(x;t)) = \mathcal{M}_{s,j}(t). \tag{3.85}$$

Finally, the source terms are discretized simply as:

$$\mathcal{N}_s(\mathcal{U}_{s,j}(x;t)) \equiv \mathcal{N}_{s,j}(t) \tag{3.86}$$

$$\mathcal{R}_s(\mathcal{U}_{s,j}(x;t)) \equiv \mathcal{R}_{s,j}(t) \tag{3.87}$$

Specifically the source term $\mathcal{N}_s(x;t)$ contains the gradient of the variation of the free-energy functional. Within mean-field treatment, the latter takes the following form:

$$\frac{\delta\mathcal{F}[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)} \sim \frac{\delta\mathcal{F}^{id}[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)} + \sum_{\alpha=1}^{K}\int d\boldsymbol{r}' V_\alpha(\boldsymbol{r}-\boldsymbol{r}')\rho_\alpha(\boldsymbol{r}';t). \tag{3.88}$$

The above equation can then be discretized as [209]:

$$\left(\frac{\delta\mathcal{F}[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)}\right)_i \sim \left.\frac{\delta\mathcal{F}^{id}[\rho_s]}{\delta\rho_s(\boldsymbol{r},t)}\right|_{\rho_s=\rho_{s,i}} + \Delta x\sum_{\alpha=1}^{K}\sum_i V_{\alpha,j-i}(t)\rho_{\alpha,i}(t). \tag{3.89}$$

Finally, we adopt a two-stage MacCormack predictor-corrector time-integrator which typically performs well for non-linear systems [210]:

$$\widetilde{\mathcal{U}}^m_{s,j} = \mathcal{U}^m_{s,j} - \frac{\Delta t}{\Delta x}\left(\mathcal{M}^m_{s,j} - \mathcal{M}^m_{s,j-1}\right) + \Delta t\left(\mathcal{N}^m_{s,j} + \mathcal{Q}^m_{s,j} + \mathcal{R}^m_{s,j}\right) \tag{3.90}$$

$$\mathcal{U}^{m+1}_{s,j} = \frac{1}{2}\left[\mathcal{U}^m_{s,j} + \widetilde{\mathcal{U}}^m_{s,j} - \frac{\Delta t}{\Delta x}\left(\widetilde{\mathcal{M}}^m_{s,j+1} - \widetilde{\mathcal{M}}^m_{s,j}\right) + \Delta t\left(\mathcal{N}^m_{s,j} + \mathcal{Q}^m_{s,j} + \mathcal{R}^m_{s,j}\right)\right]. \tag{3.91}$$

3.2.6. Numerical application I: Equilibrium mono-component system

In order to test the numerical stochastic integrator, we consider a non-Markovian FDDFT model of a system of N non-reacting ideal gas particles with a single-exponential memory kernel. The particle momentum time-correlation function of the corresponding GLE can be obtained in a closed analytical form [181, 211]:

$$\hat{C}_{t,i}(t) = \frac{\langle \boldsymbol{p}_i(t)\boldsymbol{p}_i(0)\rangle}{\langle \boldsymbol{p}_i(0)\boldsymbol{p}_i(0)\rangle} = \begin{cases} e^{\frac{tB}{2}}\left(\cos(\Omega t) - \frac{B}{2\Omega}\sin(\Omega t)\right) & \Omega \neq 0, \\ e^{\frac{tB}{2}}\left(1 - \frac{Bt}{2}\right) & \Omega = 0, \end{cases} \tag{3.92}$$

where we introduced a complex-valued parameter $\Omega = \sqrt{A - B^2/4}$.

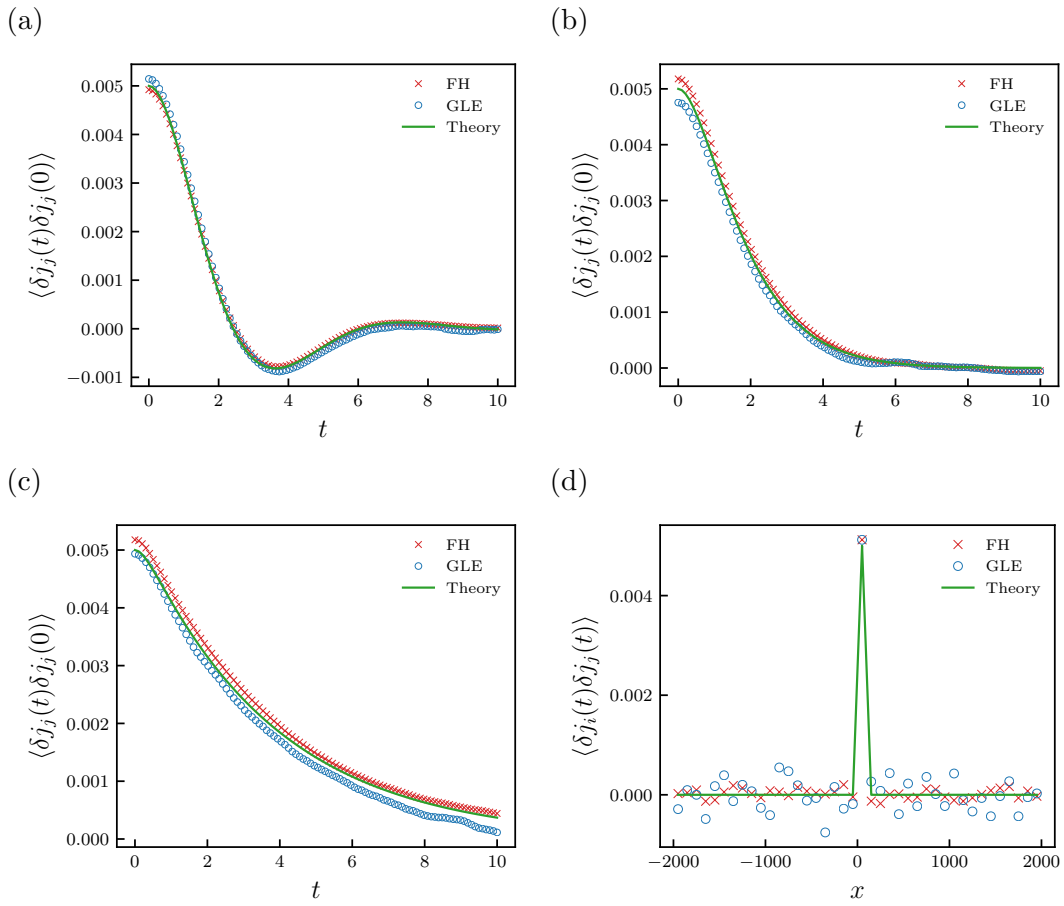

Figure 3.14. Comparison between numerical and theoretical momentum time-correlations for an ideal gas system coupled with a thermal bath at fixed temperature $T = 1$ through a single exponential memory kernel [181]. The time-correlation function is computed in the under-damped limit with $A_1 = 1$ and $B_1 = 1$ (a), in the damped case with $A_1 = 1$ and $B_1 = -2$ (b) and in the over-damped limit with $A_1 = 1$ and $B_1 = -4$ (c). We also report the momentum space correlations between every cell and the middle cell for the overdamped case (d). The system considered is uniform with initial conditions $\bar{\rho} = 0.5$ and $\bar{j} = 0.0$. The domain is discretized in 40 cells of $\Delta x = 100$

To obtain the time-correlation function for the discretized non-Markovian FDDFT model, we proceed in the following manner. First, we define the time-correlation function of the discretized momentum in a cell as:

$$C_{t,j}(t) = \left\langle \frac{1}{\Delta x^2} \int_{V_j} d\boldsymbol{r} \sum_{i=1}^{N_s(t)} \boldsymbol{p}_i(t) \delta(\boldsymbol{r} - \boldsymbol{r}_i(t)) \int_{V_j} d\boldsymbol{r}' \sum_{j=1}^{N_s(t)} \boldsymbol{p}_j(0) \delta(\boldsymbol{r}' - \boldsymbol{r}_j(0)) \right\rangle . \quad (3.93)$$

Next, we employ the independence between $\boldsymbol{p}_i(t)$ and $\boldsymbol{r}_i(t)$, and between $\boldsymbol{p}_i(t)$ and $\boldsymbol{p}_j(t)$, to write

$$\begin{aligned} C_{t,j}(t) &= \frac{1}{\Delta x^2}\int_{V_j} d\boldsymbol{r}\int_{V_j} d\boldsymbol{r}'\sum_{i=1}^{N_s(t)}\sum_{j=1}^{N_s(t)}\langle \boldsymbol{p}_i(t)\boldsymbol{p}_j(0)\rangle\left\langle \delta(\boldsymbol{r}-\boldsymbol{r}_i(t))\delta(\boldsymbol{r}'-\boldsymbol{r}_j(0))\right\rangle \\ &= \frac{1}{\Delta x^2}\int_{V_j} d\boldsymbol{r}\int_{V_j} d\boldsymbol{r}'\sum_{i=1}^{N_s(t)}\langle \boldsymbol{p}_i(t)\boldsymbol{p}_i(0)\rangle\left\langle \delta(\boldsymbol{r}-\boldsymbol{r}_i(t))\delta(\boldsymbol{r}'-\boldsymbol{r}_i(0))\right\rangle \end{aligned} \tag{3.94}$$

Moreover, since all particles are identical and thus have the same momentum correlation, we introduce $\hat{C}_t(t) = \hat{C}_{t,i}(t)$. This gives:

$$\begin{aligned} C_{t,j}(t) &= \frac{\hat{C}_t(t)}{\Delta x^2}\int_{V_j} d\boldsymbol{r}\int_{V_j} d\boldsymbol{r}'\sum_{i=1}^{N_s(t)}\left\langle \delta(\boldsymbol{r}-\boldsymbol{r}_i(t))\delta(\boldsymbol{r}'-\boldsymbol{r}_i(0))\right\rangle = \\ &= \frac{\hat{C}_t(t)}{\Delta x^2}\int_{V_j} d\boldsymbol{r}\left\langle\sum_{i=1}^{N_s(t)}\delta(\boldsymbol{r}-\boldsymbol{r}_i(t))\int_{V_j} d\boldsymbol{r}'\delta(\boldsymbol{r}'-\boldsymbol{r}_i(0))\right\rangle \end{aligned} \tag{3.95}$$

Finally, assuming that both $\boldsymbol{r}_i(0)$ and $\boldsymbol{r}_i(t)$ belong to the same cell V_j, i.e. that the cell size is much larger than the particle mean path, we obtain the time-correlation of the momentum field:

$$C_{t,j}(t) = \frac{\hat{C}_t(t)}{\Delta x^2}\int_{V_j} d\boldsymbol{r}\left\langle\sum_{i=1}^{N_s(t)}\delta(\boldsymbol{r}-\boldsymbol{r}_i(t))\right\rangle = \frac{\hat{C}_t(t)}{\Delta x}\left\langle\frac{1}{\Delta x}\int_{V_j} d\boldsymbol{r}\hat{\rho}(\boldsymbol{r},t)\right\rangle = \hat{C}_t(t)\frac{\langle\rho_j\rangle}{\Delta x} \tag{3.96}$$

We computed the momentum field time-correlation of a uniform ideal gas system with non-Markovian FDDFT, atomistic simulations (GLE), and theory (Eq. 3.96). The ideal gas system is coupled with a thermal bath at fixed temperature $T = 1$ through a single exponential memory kernel. We analysed systems characterized by three different memory kernels, corresponding to under-damped ($A = 1$ and $B = 1$), damped ($A = 1$ and $B = -2$) and over-damped ($A = 1$ and $B = -4$) cases. The domain is discretized in 40 cells of $\Delta x = 100$ (in reduced units).

The GLE simulations are run with the *gld integrator*, which is part of the LAMMPS package [114, 181]. The data are gathered over 500 independent trajectories, each running for 10^4 steps with a time step $\Delta t = 0.001$.

The non-Markovian FDDFT is integrated with the finite volume scheme discussed above. The data are averaged over 1000 independent trajectories, each running for 10^2 steps with a time step $\Delta t = 0.1$.

Figures 3.14(a-c) show that the momentum time-correlation functions obtained using atomistic, continuum and theoretical approaches converge to the same results. It is worth noticing that only exponentially decaying momentum time-correlation functions can be obtained in the Markovian approximation, e.g. the time-correlation function in Fig. 3.14(a) can be recovered only by means of a non-Markovian framework.

We also report the space correlations between every cell and the middle cell for the overdamped case in Fig. 3.14(d). This shows that our numerical scheme does not introduce any spurious space correlations into the ideal gas.

3.2.7. Numerical application II: Non-equilibrium space transition

The main aims of the following section is to demonstrate the significant role of the memory kernel and validate the FDDFT algorithm against GLE simulations in non-equilibrium processes. We study an ideal gas system at temperature $T = 1$ inside a double-well potential with the form:

$$U^{ext} = C\left[\left(\frac{x}{x_0}\right)^4 - \left(\frac{x}{x_0}\right)^2\right] \tag{3.97}$$

with $C = 5$ and $x_0 = 200$. We analyze the time evolution of the system in two scenarios: in the under-damped limit, with $A_1 = 1$ and $B_1 = 1$, and in the over-damped limit, with $A_1 = 1$ and $B_1 = -4$. In order to force mass transport between the two wells, we impose non-uniform initial conditions with $\bar{\rho}^+ = 10$, $\bar{\rho}^- = 5$, where $\bar{\rho}^+$ refers to the density in the left well and $\bar{\rho}^-$ to the density in the right well. As in the previous example, the GLE simulations are run with the *gld integrator*. The data are gathered over 10 independent trajectories, each running for 10^9 steps with a time step $\Delta t = 0.001$.

The non-Markovian FDDFT is integrated with the finite volume scheme discussed above. The data are averaged over 100 independent trajectories, where each was run for 10^6 steps with a time step $\Delta t = 0.1$.

In Figs 3.15(a-b) we compare the time-evolution of the mean density profiles obtained with GLE and FDDFT in the underdamped and overdamped cases. Despite the presence of a sharp gradient, the numerical scheme is able to accurately reproduce the diffusion of the density step function and the mass transfer between the two wells. The time-evolution of the difference between the total mass contained in the left well $m^+(t)$ and the one in the right well $m^-(t)$ is reported in Fig. 3.15(c). The noticeable difference between the over-damped and under-damped cases highlights the importance of the memory effects when modelling out-of-equilibrium systems. This prototypical

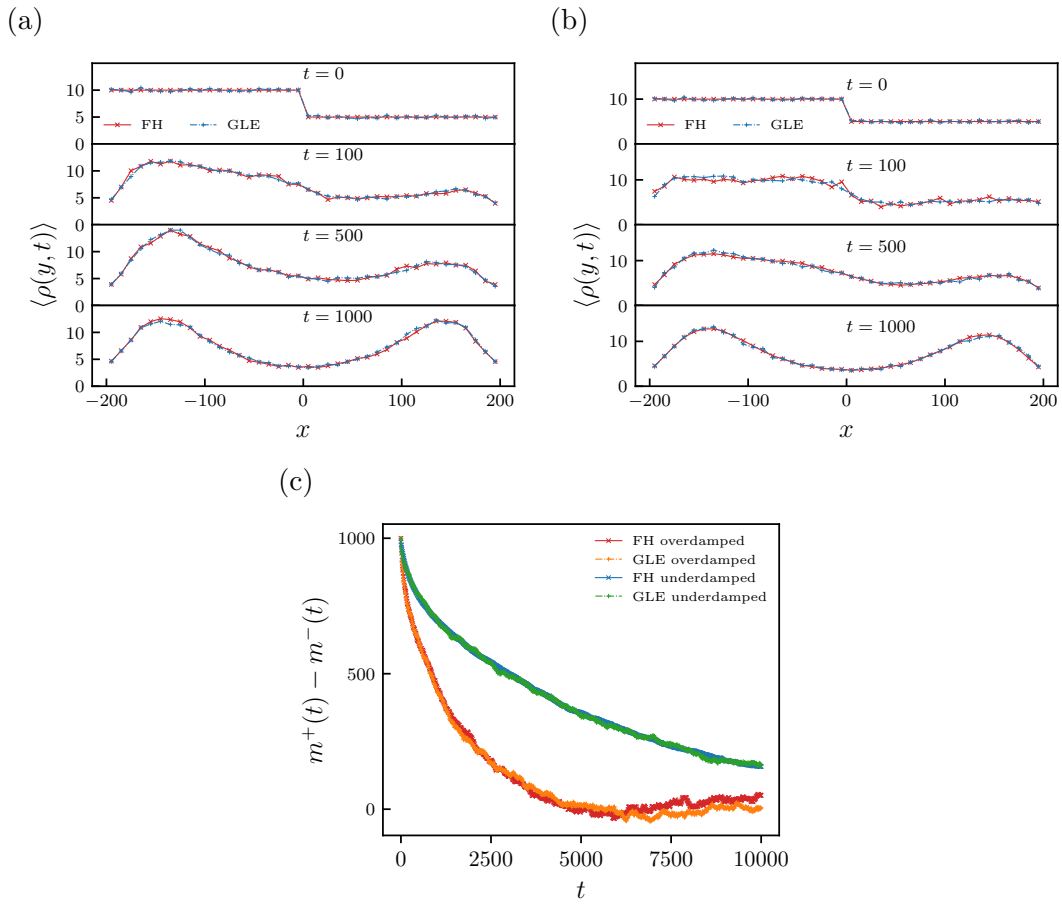

Figure 3.15. Comparison between FDDFT and GLE for a non-uniform ideal gas in a double-well potential at temperature $T = 1$. The system is started off with non-symmetric initial conditions to force mass transport between the two wells. The density time-evolution is computed in the under-damped limit with $A_1 = 1$ and $B_1 = 1$ (a) and in the over-damped limit with $A_1 = 1$ and $B_1 = -4$ (b). In (c) we report the time-evolution of the difference between the total mass contained in the left well $m^+(t)$ and the mass in the right well $m^-(t)$. The domain is discretized in 40 cells of $\Delta x = 10$ and periodic boundary conditions are used.

example suggests that, during energy-barrier transitions, including biological transitions and chemical reactions, the widely adopted Markovian assumption, if not satisfied, could introduce a significant error in the system model.

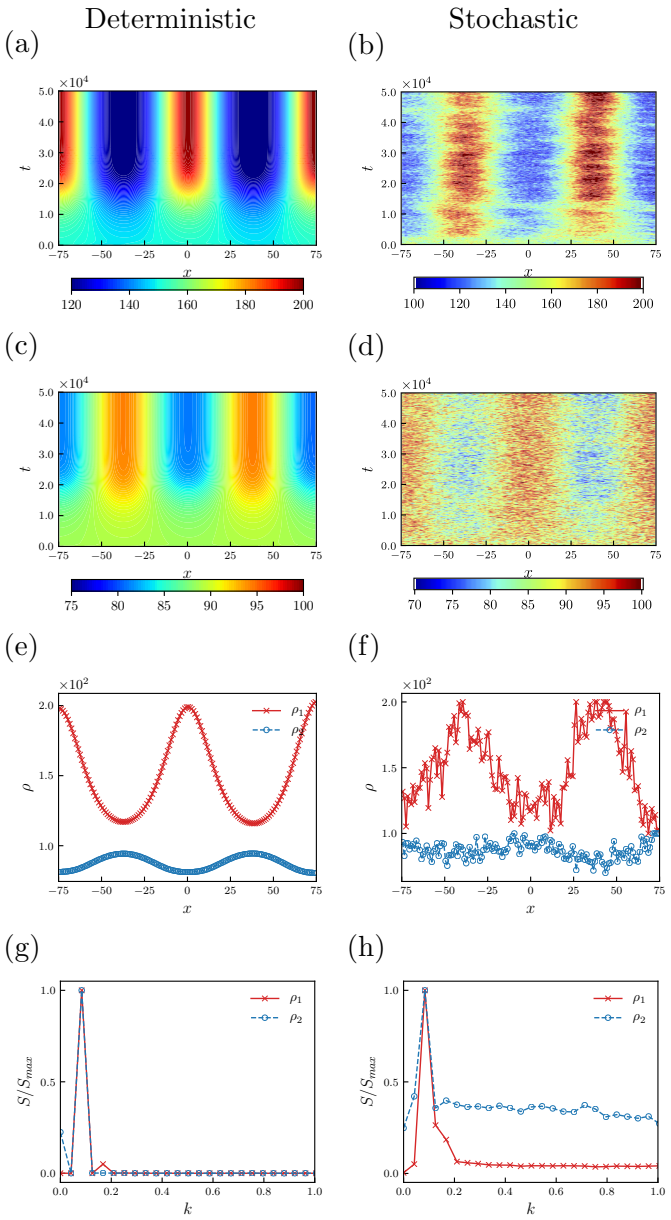

Figure 3.16. Turing patterns for $B_{1,1} = -0.4$ characterized by a wave length $L \sim 70$, which is consistent with the one predicted analytically. The panels (a-c-e-g) and (b-d-f-h) correspond to the deterministic and stochastic cases, respectively. The panels (a-b) and (c-d) show the evolution in time of the densities ρ_1 and ρ_2, respectively. The steady state densities are reported in (c-f). In (g-h) we plot the structure factors of the steady states.

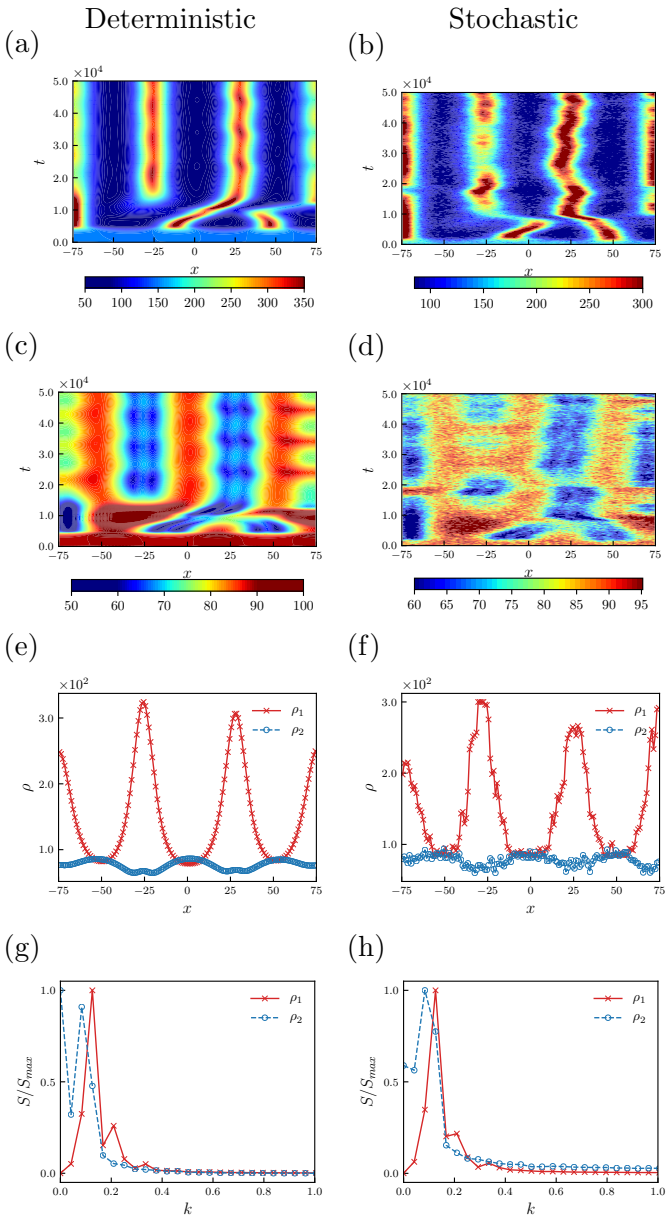

Figure 3.17. Same as Fig. 3.16, but for $B_{1,1} = -0.2$. This configuration is characterized by a wave length $L \sim 70$ $L \sim 50$.

3.2.8. Numerical application III: Memory-driven Turing patterns in binary reacting system

In this section we apply the derived framework to study Turing patterns [198] observed in multi-species chemical reaction-diffusion systems. On of our aim is to study Turing

patterns arising in systems composed of two reacting chemicals characterized by dissimilar memory kernels. Our theoretical and numerical results can be useful in experimental design when studying Turing patterns. In two-component systems, in order to observe Turing patterns, one typically needs to use a third substance, which is fixed in space and reversibly bound to only one of the two component. As a result, the effective diffusion coefficient of the binding component is significantly smaller than that of the other components. Here we analyse a similar scenario, where the binding species consists in a thermal bath with dissimilar memory kernels for the two reacting component.

In this work we consider systems reacting according to the non-linear kinematic chemical model proposed by Schnakenberg [199]. This model consists of the following reactions among the chemicals X_1, A, B and X_2:

$$X \underset{k_{-1}}{\overset{k_1}{\rightleftharpoons}} A, \quad B \xrightarrow{k_2} Y, \quad 2X + Y \xrightarrow{k_3} 3X. \tag{3.98}$$

with k_{-1}, k_1, k_2 and k_3 representing the reaction rate constants. Let us define the densities of X_1, A, B and X_2 as ρ_1, α, β and ρ_2, respectively. According to the law of mass action, the time-evolution of the system components is given by:

$$w_{\rho_1}(\boldsymbol{r};t) = k_{-1}\alpha - k_1\rho_1(\boldsymbol{r};t) + k_3\rho_1^2(\boldsymbol{r};t)\rho_2(\boldsymbol{r};t), \tag{3.99}$$

$$w_{\rho_2}(\boldsymbol{r};t) = k_2\beta - k_3\rho_1^2(\boldsymbol{r};t)\rho_2(\boldsymbol{r};t), \tag{3.100}$$

$$w_{\alpha}(\boldsymbol{r};t) = k_1\rho_1(\boldsymbol{r};t) - k_{-1}\alpha, \tag{3.101}$$

$$w_{\beta}(\boldsymbol{r};t) = -k_2\beta. \tag{3.102}$$

If the reactions are performed in an environment with abundance of A and B, then α and β can be assumed constant. It follows that $w_\alpha(\boldsymbol{r};t) \sim w_\beta(\boldsymbol{r};t) \sim 0$ and the following system of continuity equations governs the time-evolution of the system:

$$\partial_t\rho_1(\boldsymbol{r};t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (m_s^{-1}\boldsymbol{j}_1(\boldsymbol{r};t)) + w_{\rho_1}(\boldsymbol{r};t) \tag{3.103}$$

$$\partial_t\rho_2(\boldsymbol{r};t) = -\boldsymbol{\nabla}_{\boldsymbol{r}} \cdot (m_s^{-1}\boldsymbol{j}_2(\boldsymbol{r};t)) + w_{\rho_2}(\boldsymbol{r};t) \tag{3.104}$$

In our numerical simulations we employ the following parameters: $k_{-1}\alpha = 1.0$, $k_1 = 0.02$, $k_2\beta = 2.0$, $k_3 = 10^{-6}$.

Turing patterns are observed when a uniform dynamical system is stable in the absence of diffusion, but becomes unstable to perturbation in presence of non-uniform diffusion. In our analysis, we use the one-dimensional overdamped density equation Eq. (3.59) in

the weak noise limit:

$$\partial_t \rho_1(x;t) = d_1 \partial_x^2 \rho_1(x;t) + w_{\rho_1}(x;t) \tag{3.105}$$

$$\partial_t \rho_2(x;t) = d_2 \partial_x^2 \rho_1(x;t) + w_{\rho_2}(x;t) \tag{3.106}$$

where we defined $d_1 = k_B T D_{eff,1} m_1^{-1}$ and $d_2 = m_2^{-1} k_B T D_{eff,2}$. First, we compute the stationary states of the system by setting:

$$\begin{cases} w_{\rho_1}(\boldsymbol{r};t) = 0 \\ w_{\rho_2}(\boldsymbol{r};t) = 0 \end{cases} . \tag{3.107}$$

Schnakenberg's model has an unique uniform stationary state given by:

$$\rho_1^b = \frac{k_2\beta + k_{-1}\alpha}{k_1} \sim 150 \tag{3.108}$$

$$\rho_2^b = \frac{k_2\beta k_1^2}{k_3(k_2\beta + k_{-1}\alpha)^2} \sim 89. \tag{3.109}$$

The Jacobian matrix for the reaction-only system evaluated at the stationary state, i.e. with components $J_{i,j} = \frac{\partial w_{\rho_i}}{\partial \rho_j}\Big|_{\rho_i^b}$ is given by:

$$\boldsymbol{J} = \begin{pmatrix} k_1 \dfrac{k_2\beta - k_{-1}\alpha}{k_2\beta + k_{-1}\alpha} & \dfrac{k_3}{k_1^2}(k_2\beta + k_{-1}\alpha)^2 \\ \dfrac{2k_2\beta K_1}{k_2\beta + k_{-1}\alpha} & -\dfrac{k_3}{k_1^2}(k_2\beta + k_{-1}\alpha)^2 \end{pmatrix} \tag{3.110}$$

The reaction parameters chosen in this study satisfy the following conditions:

$$\mathrm{Tr}J = J_{11} + J_{22} \sim -0.0158 < 0 \tag{3.111}$$

$$\mathrm{Det}J = J_{11}J_{22} - J_{12}J_{21} \sim 0.0004 > 0 \tag{3.112}$$

which ensure the stability of the reaction-only system, as required for the appearance of the Turing patterns (see Sect. B.2.3 for more details).

The Jacobian matrix of the reaction-diffusion system is given by $\boldsymbol{J}_k = \boldsymbol{J} - k^2[d_1, d_2]\boldsymbol{I}$. The effective diffusion parameters d_1 and d_2 are functions of the memory kernels, as shown in Sect. 3.2.3. We consider two cases both with single exponential memory kernels, with parameters $A_{1,1} = 1.0$, $A_{2,1} = 1.0$ and $B_{2,1} = -20.0$. The two cases differ due to $B_{1,1}$, which is set to $= -0.4$ in case (a) and to -0.2 in case (b). The respective values of d_1 are 0.4 and 0.2. Note that in both cases the Turing instability conditions are satisfied

because $\mathrm{Det}J_{km} < 0$ (see Sect. B.2.3 for more details). The minimum wave-number for which Turing instability occurs is given by:

$$k_m = \sqrt{\frac{d_1 J_{22} + d_2 J_{11}}{2 d_1 d_2}}. \tag{3.113}$$

Such value corresponds to a wave-length $L = 2\pi/k_m$.

Figure 3.16 shows the time-evolution of the density for case (a) obtained from deterministic and stochastic formulations. The characteristic wave-length for such a system $L \sim 70$, theoretically predicted by Eq. (3.113), is recovered in both deterministic and stochastic solutions. In the stochastic evolution the instability is enhanced by the noise, thus the patterns arise sooner. The structure factor, discussed in Sect. B.2.4, exhibits a clear peak at about the predicted wave number. However, in the stochastic case, higher frequency contributions due to the correlated noise are also visible.

Figure 3.17 shows the deterministic and stochastic density evolution for case (b). Also in this case, the wave-length $L \sim 50$ observed in the simulations is consistent with the one predicted by the theoretical analysis. However, the evolution of this system shows the presence of secondary peaks with low intensity, as is also suggested by the corresponding structure factor.

In this chapter we have derived a theoretical framework (FDDFT) for describing reacting multi-species fluid systems. In the next chapter we propose and test numerical schemes for solving DDFT with additional complexities: (1) DDFT with the stochastic term and (2) DDFT with non-homogeneous diffusion.

Chapter 4

Numerical methods for DDFT with additional complexities

The Markovian limit of the previously derived non-Markovian FDDFT recovers the system of equations for density ρ and velocity $\mathbf{v}$ fields of colloidal fluids in a thermal bath derived by Durán-Olivencia *et al.* [68]:

$$\partial_t \rho(\mathbf{r},t) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(m^{-1} \rho(\mathbf{r},t) \mathbf{v}(\mathbf{r},t) \right) = 0 \tag{4.1}$$

$$\begin{aligned} \partial_t \left(\rho(\mathbf{r},t) \mathbf{v}(\mathbf{r},t) \right) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\rho(\mathbf{r},t) \mathbf{v}(\mathbf{r},t) \otimes \mathbf{v}(\mathbf{r},t) \right) + \rho(\mathbf{r},t) \boldsymbol{\nabla}_{\mathbf{r}} \frac{\delta \mathcal{E}[\rho]}{\delta \rho(\mathbf{r},t)} \\ + \rho(\mathbf{r},t) \boldsymbol{\theta}_0(\mathbf{r},t) \mathbf{v}(\mathbf{r},t) + \sqrt{k_B T m \rho(\mathbf{r},t) \boldsymbol{\theta}_0(\mathbf{r},t)} \boldsymbol{\mathcal{W}}(\mathbf{r},t) = 0 \end{aligned} \tag{4.2}$$

where m is the atomic mass of the colloidal particles, $\mathcal{E}[\rho]$ is the density-dependent free-energy functional, $\boldsymbol{\theta}_0(\mathbf{r},t)$ is the friction tensor accounting for the interactions between colloidal and bath particles, and $\boldsymbol{\mathcal{W}}$ is a vector of Gaussian stochastic processes delta-correlated in space and time, i.e.

$$\langle \boldsymbol{\mathcal{W}}(\mathbf{r},t) \rangle = 0, \tag{4.3}$$

$$\langle \boldsymbol{\mathcal{W}}(\mathbf{r},t), \boldsymbol{\mathcal{W}}(\mathbf{r}',t') \rangle = 2\delta(t-t')\delta(\mathbf{r}-\mathbf{r}'). \tag{4.4}$$

In the over-damped limit ($m^{-1}\boldsymbol{\theta}_0 \to \infty$), the characteristic time scale of the momentum dynamics is much shorter than the density one [52, 68]. Thus, in first approximations, the contributions of the terms $\boldsymbol{\nabla}_{\mathbf{r}} \cdot (\rho \mathbf{v} \otimes \mathbf{v})$ and $\frac{\partial \rho \mathbf{v}}{\partial t}$ can be neglected. As a result, one

obtains the stochastic time-evolution equation for the density field [68, 70]:

$$\partial_t \rho(\mathbf{r},t) = \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left({\boldsymbol{\theta}_0}^{-1} \rho(\mathbf{r},t) \boldsymbol{\nabla}_{\mathbf{r}} \frac{\delta \mathcal{E}[\rho]}{\delta \rho(\mathbf{r},t)} \right) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\sqrt{k_B T \rho(\mathbf{r},t) {\boldsymbol{\theta}_0}^{-1}} \boldsymbol{\mathcal{W}}(\mathbf{r},t) \right). \quad (4.5)$$

where $\mathcal{E}[\rho]$ denotes the free-energy of the system, satisfying

$$\mathcal{E}[\rho] = \int_{\mathbb{R}^d} f(\rho) d\mathbf{r} + \int_{\mathbb{R}^d} V(\mathbf{r}) \rho \, d\mathbf{r} + \frac{1}{2} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} K(\mathbf{r}) \star \rho \, d\mathbf{r}, \quad (4.6)$$

with $f(\rho)$ describing the free-energy part dependent on the local density field ρ, $V(\mathbf{r})$ accounting for external potentials, and $K(\mathbf{r})$ modelling the inter-particle interacting potential. The constant β is defined as $\beta = (k_B T)^{-1}$. As discussed in details later in this work, for the simple ideal-gas free-energy functional $\mathcal{E}[\rho] = \rho\,(\log \rho - 1)$, it reduces to the stochastic diffusion equation [71]. However, the presence of the functional $\mathcal{E}[\rho]$ allows for the introduction of non-linear diffusion, external force fields and non-local interactions.

In Section 4.1 we develop a finite volume scheme to integrate in time the stochastic Eq. (4.5) embedded with the constant parameters $m = 1$ and $\boldsymbol{\theta}_0 = \mathbf{1}$. The presence of the stochastic flux brings an additional layer of complexity to the problem which requires a proper discretization.

In Section 4.2 we propose a numerical scheme to simulate the deterministic version of Eq. (4.5). Our study focuses specifically on the effects of a time- and space-dependent $\boldsymbol{\theta}_0$ on the time evolution of the density field.

4.1. Overdamped FDDFT: Numerical methods and applications [4]

A numerical methodology able to simulate Eq. (4.5) equipped with a general free-energy functional is missing. However, previous works have simulated thermal fluctuations in fluid systems modelled with similar stochastic equations. Specifically, numerical methodologies have been proposed to solve the Landau-Lifshitz-Navier-Stokes (LLNS) equations,

[4] This section includes material from A. Russo, S. P. Perez, J. Carrillo, M. A. Durán-Olivencia, P. Yatsyshin & S. Kalliadasis. A finite-volume method for fluctuating dynamical density functional theory, *J. Comp. Phys.*, 428, 109796 (2021).

which can be expressed as:

$$\partial_t \rho + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(m^{-1} \rho \mathbf{v}\right) = 0 \tag{4.7}$$

$$\partial_t \left(\rho \mathbf{v}\right) + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\rho \mathbf{v} \otimes \mathbf{v} + p\mathbf{I}\right) = \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\boldsymbol{\tau} + \boldsymbol{\mathcal{S}}\right) \tag{4.8}$$

$$\partial_t E + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\mathbf{v}E + p\mathbf{v}\right) = \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left(\boldsymbol{\tau} \cdot \mathbf{v} + \kappa \boldsymbol{\nabla}_{\mathbf{r}} T + \boldsymbol{\mathcal{Q}} + \mathbf{v} \cdot \boldsymbol{\mathcal{S}}\right) \tag{4.9}$$

where p is the pressure, $\boldsymbol{\tau} = \eta\left(\boldsymbol{\nabla}_{\mathbf{r}}\mathbf{v} + \boldsymbol{\nabla}_{\mathbf{r}}\mathbf{v}^T - \frac{2}{3}\mathbf{I}\boldsymbol{\nabla}_{\mathbf{r}} \cdot \mathbf{v}\right)$ is the stress tensor, and η and κ are coefficients of viscosity and thermal conductivity, respectively. Moreover, the stochastic tensors $\boldsymbol{\mathcal{S}}$ and $\boldsymbol{\mathcal{Q}}$ are Gaussian processes delta-correlated in space and time, i.e.

$$\langle \boldsymbol{\mathcal{S}}_{ij}(\mathbf{r}, t), \boldsymbol{\mathcal{S}}_{kl}(\mathbf{r}', t') \rangle = 2k_{\mathrm{B}}\eta T \left(\delta^K_{ik}\delta^K_{jl} + \delta^K_{il}\delta^K_{jk} - \frac{2}{3}\delta^K_{ij}\delta^K_{kl}\right) \delta(t - t')\delta(\mathbf{r} - \mathbf{r}'), \tag{4.10}$$

$$\langle \boldsymbol{\mathcal{Q}}_{i}(\mathbf{r}, t), \boldsymbol{\mathcal{Q}}_{j}(\mathbf{r}', t') \rangle = 2k_{\mathrm{B}}\kappa T^2 \delta^K_{ij} \delta(t - t')\delta(\mathbf{r} - \mathbf{r}') \tag{4.11}$$

with δ^K_{ij} denoting the Kronecker delta. One of the first works on LLNS was developed by Garcia *et al.* [75] and it employs a finite difference scheme to treat the numerical fluxes of the stochastic diffusion equation. Further works by Bell *et al.* [76, 77] propose explicit Eulerian discretizations of the LLNS equations coupled with a third-order Runge-Kutta, with the objective of adequately reproducing the fluctuations in density, energy, and momentum. Donev and co-workers [208] introduce the analysis of the structure factor (equilibrium fluctuation spectrum) to construct optimal finite volume schemes to solve the LLNS and study the accuracy for a given discretization at different wavelengths. The same group also constructed methods to solve LLNS via staggered grids [78]. Other works have proposed implicit-explicit predictor-corrector [79] or two-level leapfrog [80] temporal integrators. Additionally, hybrid schemes have been developed to couple LLNS with MD [81, 82, 83] or with MC [84, 85] simulations. Moreover, LLNS equations have also been employed in numerical studies of reactive multi-species fluid mixtures [212].

Further works have studied numerical methods for particular application of the overdamped FDDFT (Eq. (4.5)). Specifically, in Refs [71, 86] the authors develop numerical methods for reaction-diffusion equations, obtained by adding reaction terms to Eq. (4.5) equipped with the ideal-gas free-energy functional.

All these works have contributed to better understand the effects of thermal fluctuations in fluid systems. In spite of these efforts, a numerical methodology able to simulate Eq. (4.5) equipped with a general free-energy functional is still missing. Such methodology would be needed to study a wide range of non-equilibrium phenomena, which can be studied within the field of FDDFT. Relevant examples of these physical phenomena

include dynamic evolution of confined systems and phase transitions, such as nucleation. In this section, we propose a possible approach to develop such numerical schemes, and we provide exhaustive tests in order to elucidate their strengths and weaknesses. Initially, in Sect. 4.1.3 we conduct simulations of purely-diffusive ideal-gas free-energy with noise and without external potentials, given that the theoretical results for these systems are analytically available. This allows us to compare variance, time-correlations, space-correlations and structure factor from our numerical schemes against the theoretical results. The results of these tests justify the choice of the Runge-Kutta temporal integrator Eq. (4.45) and linear approximation of the stochastic flux Eqs (4.34)-(4.35) as the most accurate and efficient numerical techniques to simulate our SPDE Eq. (4.16). This choice is maintained to simulate the following examples. Secondly, we study ideal gases confined by external potential $V(x)$. Mean, variance, spatial correlation of density and decay of the discrete free-energy in time are compared against MD simulations results. Finally, we study the homogeneous vapor-liquid transition of a Lennard- Jones fluid.

4.1.1. Governing equations and related properties

In the present work, we analyse the following general stochastic partial differential equation, based on the overdamped FDDFT in Eq. (4.5) with $m = 1$ and $\boldsymbol{\theta}_0 = \mathbf{1}$,

$$\begin{cases} \partial_t \rho(\boldsymbol{r};t) = \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left[\rho(\boldsymbol{r};t) \boldsymbol{\nabla}_{\mathbf{r}} \dfrac{\delta \mathcal{E}[\rho]}{\delta_{\mathbf{r}} \rho(\boldsymbol{r};t)} \right] + \boldsymbol{\nabla}_{\mathbf{r}} \cdot \left[\sqrt{\rho(\boldsymbol{r};t)/\beta} \boldsymbol{\mathcal{W}}(\boldsymbol{r};t) \right] \quad \mathbf{r} \in \mathbb{R}^d, \ t > 0, \\ \rho(\mathbf{r};0) = \rho_0(\mathbf{r}), \end{cases} \tag{4.12}$$

Eq. (4.12) may be seen as a stochastic version of the gradient flow equation previously studied in Refs [72, 73]. In fact, the deterministic version of Eq. (4.12) has received a great deal of attention in the context of gradient flows. As discussed in Ref. [68], in the weak noise limit, the most-likely path followed by the system minimizes the Lagrangian defined as $\mathcal{L} = \| \partial_t \rho - \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left(\rho(\boldsymbol{r};t) \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta \mathcal{E}[\rho]}{\delta \rho} \right) \|_{(\sigma\sigma^*)^{-1}}$, where σ is the operator acting on the noise $\boldsymbol{\mathcal{W}}(\boldsymbol{r};t)$. Thus, the most-likely solution $\langle \rho \rangle(\boldsymbol{r};t)$ satisfies

$$\partial_t \langle \rho \rangle(\boldsymbol{r};t) = \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left(\langle \rho \rangle \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta \mathcal{E}[\langle \rho \rangle]}{\delta \langle \rho \rangle} \right). \tag{4.13}$$

Eq. (4.13) has a gradient flows structure [72, 213], and it has applications in a variety of contexts such as granular media [73], material science of biological swarming [214].

The fundamental property of Eq. (4.13) is that the free-energy (Eq. (4.6)) is minimized following the decay rate

$$\frac{d}{dt}\mathcal{E}[\langle\rho\rangle] = -\int_{\mathbb{R}^d}\langle\rho\rangle\left|\frac{\delta\mathcal{E}[\langle\rho\rangle]}{\delta\langle\rho\rangle}\right|^2 d\mathbf{r}, \tag{4.14}$$

where the variation of the free-energy $\mathcal{E}[\rho]$ with respect to the density ρ in the case of Eq. (4.6) satisfies

$$\frac{\delta\mathcal{E}[\rho]}{\delta\rho} = f'(\rho) + V(\mathbf{r}) + K * \rho. \tag{4.15}$$

This property however is not satisfied by the stochastic gradient flow in Eq. (4.12), where punctual increases in the free-energy during the dynamical evolution can take place. Precisely, this jumps allow the system to overcome energy barriers, leading to phenomena such as phase transitions.

4.1.2. Numerical methods

Equation (4.12) (in one-dimension) can be re-written as

$$\partial_t\rho = \partial_x F_d(\rho) + \partial_x F_s(\rho, \mathcal{W}), \tag{4.16}$$

where F_d and F_s denote the deterministic and stochastic fluxes respectively,

$$F_d = \rho\partial_x\frac{\delta\mathcal{E}[\rho]}{\delta\rho}, \quad F_s = \sqrt{\rho/\beta}\mathcal{W}. \tag{4.17}$$

The finite volume discretization of Eq. (4.16) is obtained as follows. First, we divide the domain into grid cells $C_j = [x_{j-\frac{1}{2}}, x_{j+\frac{1}{2}}]$, each of length $\Delta x = x_{j+1/2} - x_{j-1/2}$. Then the cell average of ρ is defined as

$$\overline{\rho}_j(t) = \frac{1}{\Delta x}\int_{x_{j-1/2}}^{x_{j+1/2}}\rho(x,t)dx. \tag{4.18}$$

Subsequently, Eq. (4.16) in integrated in space over each cell. Finally Gauss divergence theorem leads to the semi-discrete equation for the temporal evolution of the cell average density,

$$\frac{d\overline{\rho}_j}{dt} = \frac{F_{d,j+1/2} - F_{d,j-1/2}}{\Delta x} + \frac{F_{s,j+1/2} - F_{s,j-1/2}}{\Delta x}, \tag{4.19}$$

where $F_{d,j+1/2}$ and $F_{s,j+1/2}$ denote the evaluation of the deterministic and stochastic fluxes Eq. (4.17) at the boundary $x_{j+1/2}$, respectively. This separation of the physical flux in two fluxes, deterministic and stochastic, has been effectively applied in previous

studies [76, 208].

Space discretization of the deterministic flux

The deterministic flux at the cell boundaries $F_{d,j+1/2}$ is evaluated following a hybrid method, which employs a central or upwind approximation depending on the relative local variation of the density. Specifically, a central high-order scheme is adopted in the parts of the domain where no sharp gradients of the density are found. An upwind scheme, instead, is employed in the regions of the domain with high density gradients, in order to prevent the spurious oscillations from central high-order schemes. This is a common technique in deterministic fluid dynamics as it allows to construct high-resolution and oscillation-free schemes [215].

Previous works in the field of fluctuating hydrodynamics [71, 76, 208] approximate the deterministic flux with a simple second-order central difference approach. However, the hybrid scheme proposed here aims to avoid possible spurious oscillations. In fact, the previous literature focuses mainly on fluctuating hydrodynamics with a deterministic flux in the form $F_d(\rho) = \partial_x \rho$. The discretization of the flux with central approximation schemes works well for some simple scenarios discussed in the literature, but it can cause spurious oscillations for certain density profiles as shown in Fig. 4.1.

In the case of Eq. (4.16), the inclusion of a stochastic flux prevents the density profile from being smooth. Thus, the proposed hybrid scheme evaluates the local gradient in the density with respect to the neighbouring gradients. When the local gradient is significantly larger than the neighbouring gradients, an upwind approximation is chosen. If not, the central approximation prevails. As a result, our proposed hybrid scheme for the deterministic flux satisfies:

$$F_{d,j+1/2} = \big(1 - \phi(r_{j+1/2})\big)\, F^c_{d,j+1/2} + \phi(r_{j+1/2}) F^u_{d,j+1/2}, \tag{4.20}$$

where $\phi(r_{j+1/2})$ is a flux limiter with a threshold parameter k, defined as

$$\phi(r_{j+1/2}) = \begin{cases} 0, & \text{if } r_{j+1/2} \leq k \\ 1, & \text{if } r_{j+1/2} > k \end{cases}$$

and $r_{j+1/2}$ is a quotient measuring the relative local total variation of the density

$$r_{j+1/2} = \frac{|\rho_{j+1} - \rho_j|}{\sum_{l=-w}^{w} |\rho_{l+1} - \rho_l|}, \tag{4.21}$$

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 selected more frequently, leading to diffusive behaviours which dampens the statistical values of the structure factor and the variance (as will be shown later). On the contrary, for large values of k, the central scheme is predominant, and spurious oscillations may occur. In Figure 4.1 we report the results of a numerical example run for different values of k.

Firstly, Figs 4.1(a-b) are obtained by simulating Eq. (4.16) with a free-energy satisfying $\delta\mathcal{E}/\delta\rho = \log\rho + 0.1x$. The initial density profile presents two discontinuities, as shown in Fig. 4.1(a). Under these conditions, the numerical solution evolves as a diffusive travelling wave, with the two discontinuities in the initial density triggering spurious oscillations. These oscillations diminish by reducing k (e.g. for $k = 0$, which corresponds to a purely upwind flux, the diffusion eliminates any oscillations). However, a low value of k critically dampens the variance, due to the diffuse nature of the upwind flux, as shown in Fig. 4.1(b).

Secondly, Fig. 4.1(c) comes from simulating Eq. (4.16) with a free-energy satisfying $\delta\mathcal{E}/\delta\rho = \log\rho$ and starting from an equilibrium density profile. For this case, the theoretical value of the structure factor is known and is given by Eq. (C.19), meaning that the dampening behaviour of the upwind scheme could be directly evaluated from Fig. 4.1(c). It is clear again that 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. Based on the examples above, such value is chosen to be $k = 3$ in this work.

The central and upwind deterministic fluxes in Eq. (4.20) are built as follows:

1. The upwind discretization of the deterministic flux is constructed as proposed in [209], where a first- and second-order finite-volume method for nonlinear equations with gradient flow structure is constructed. The authors propose to firstly reconstruct the density profile in each cell C_j as a constant profile for the first-order scheme, or as a linear profile for the second-order scheme,

$$\widetilde{\rho}_j(x) = \begin{cases} \overline{\rho}_j, \quad x \in C_j, \quad \text{for the first-order scheme,} \\ \overline{\rho}_j + (\rho_x)_j\,(x - x_j), \quad x \in C_j, \quad \text{for the second-order scheme,} \end{cases} \tag{4.22}$$

   where the east and the west density values ρ_j^E and ρ_j^W at the cell interfaces $x_{j+\frac{1}{2}}$

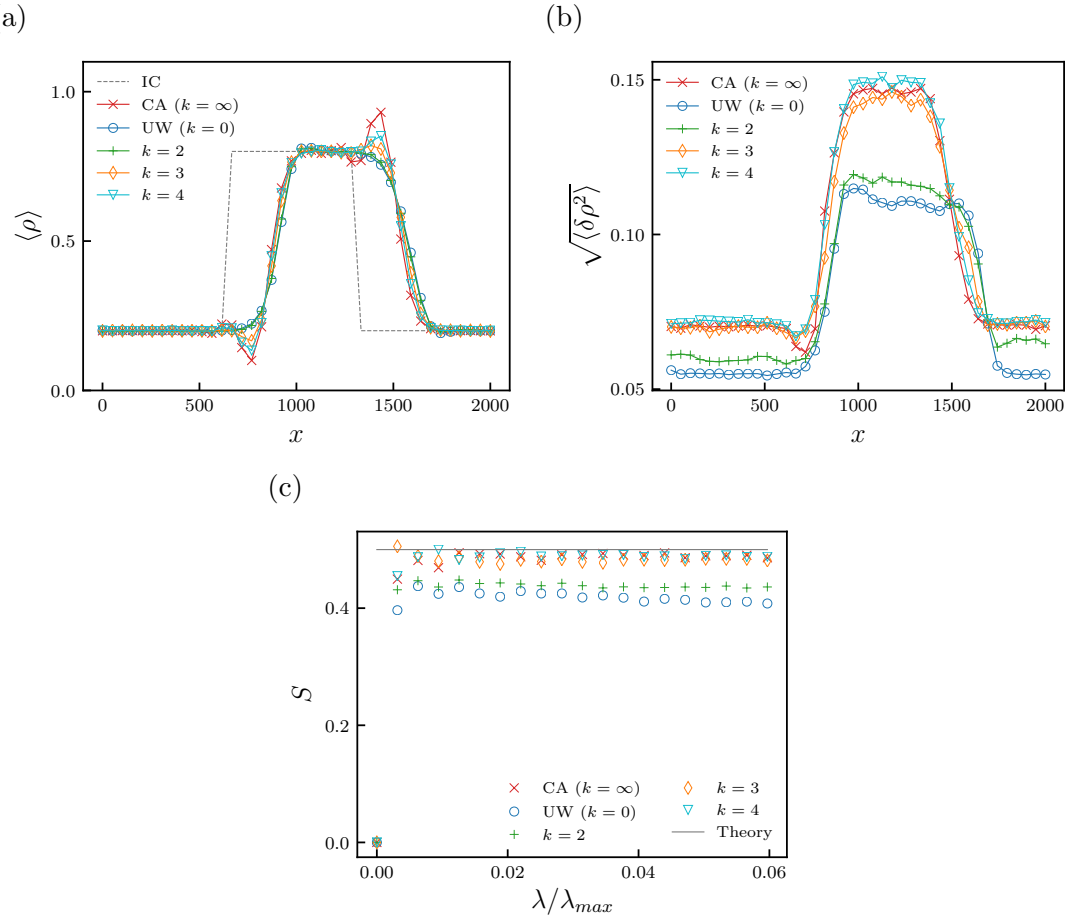

Figure 4.1. (a) Mean density and (b) standard deviation for a moving and diffusing initial step function evolving under Eq. (4.12), with $\delta\mathcal{E}/\delta\rho = \log\rho + u_0 x$ with $u_0 = 0.1$. For clarity, also the structure factor for a uniform system is reported in (c). IC: initial condition, CA: central approximation ($k = \infty$), UW: upwind approximation ($k = 0$).

and $x_{j-\frac{1}{3}}$ are approximated as

$$\begin{aligned} \rho_j^E &= \overline{\rho}_j + \frac{\Delta x}{2} \left(\rho_x\right)_j , \\ \rho_j^W &= \overline{\rho}_j - \frac{\Delta x}{2} \left(\rho_x\right)_j . \end{aligned} \tag{4.23}$$

The numerical derivatives $\left(\rho_x\right)_j$ at every cell C_j are computed by means of an adaptive procedure which ensures that the point values Eq. (4.23) are second-order

and non-negative, i.e.

$$(\rho_x)_j = \text{minmod}\left(\theta\,\frac{\overline{\rho}_{j+1}-\overline{\rho}_j}{\Delta x}, \frac{\overline{\rho}_{j+1}-\overline{\rho}_{j-1}}{2\Delta x}, \theta\,\frac{\overline{\rho}_j-\overline{\rho}_{j-1}}{\Delta x}\right), \tag{4.24}$$

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 $\theta = 2$ in this work. After completing the density discretization, the deterministic flux $F^u_{d,j+1/2}$ is evaluated with an upwind scheme as

$$F^u_{d,j+1/2} = u^+_{j+1/2}\,\rho^E_j + u^-_{j+1/2}\,\rho^W_{j+1}, \tag{4.25}$$

where

$$u^+_{j+1/2} = \max\left(u_{j+1/2}, 0\right) \quad \text{and} \quad u^-_{j+1/2} = \min\left(u_{j+1/2}, 0\right). \tag{4.26}$$

and $u_{j+1/2}$ is computed as

$$u_{j+1/2} = -\frac{\left(\dfrac{\delta\mathcal{E}}{\delta\rho}\right)_{j+1} - \left(\dfrac{\delta\mathcal{E}}{\delta\rho}\right)_j}{\Delta x}. \tag{4.27}$$

In the latter equation, the discrete variation of the free-energy with respect to the density $\left(\dfrac{\delta\mathcal{E}}{\delta\rho}\right)_j$ is evaluated from Eq. (4.15) as

$$\left(\frac{\delta\mathcal{E}}{\delta\rho}\right)_j = \Delta x \sum_i K(x_j - x_i)\rho_i + F(\rho_j) + V(x_j). \tag{4.28}$$

Additional details on the discrete convolution term $\sum_i K(x_j - x_i)\rho_i$ in Eq. (4.28) are given in Ref. [209].

2. The central approximation of the deterministic flux is a common way to treat the deterministic flux in fluctuating hydrodynamic [71, 76, 208]. Specifically the central deterministic flux is evaluated as

$$F^c_{d,j+1/2} = u_{j+1/2}\,\rho_{j+1/2}, \tag{4.29}$$

where $u_{j+1/2}$ is computed as in Eq. (4.27) and $\rho_{j+1/2}$ is the average density of the adjacent cells

$$\rho_{j+1/2} = \frac{\overline{\rho}_j + \overline{\rho}_{j+1}}{2}. \tag{4.30}$$

The 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 Eqs (4.27)-(4.30), given that the presence of the stochastic flux limits the spatial order of accuracy. This is in line with previous works in the literature, e.g. [71, 76, 208].

Space discretization of the stochastic flux

Several approximations for the stochastic flux have been put forward in the literature [71, 76]. Firstly, due to the nature of the white noise, which cannot be evaluated point-wise in time and space, the component $\mathcal{W}$ of the stochastic flux is discretized by means of a spatio-temporal average as in Ref. [208], i.e.

$$\overline{\mathcal{W}}_j = \frac{1}{\Delta x \Delta t} \int_t^{t+\Delta t} \int_{x_{j-\frac{1}{2}}}^{x_{j+\frac{1}{2}}} \mathcal{W}(x,t) dx\, dt. \tag{4.31}$$

The expression above is equal to a normal distribution with zero mean and variance $(\Delta x \Delta t)^{-1}$, so that

$$\overline{\mathcal{W}}_j = \mathcal{N}(0,1)/\sqrt{\Delta x \Delta t}. \tag{4.32}$$

By employing the latter equation, here we propose four different approximations for the stochastic flux:

1. Forward approximation:

$$F_{s,j-1/2} = \left(\sqrt{\frac{\rho}{\beta}}\mathcal{W}\right)_{j-1/2} = \sqrt{\frac{\rho_j}{\beta}}\overline{\mathcal{W}}_j. \tag{4.33}$$

2. Linear approximation:

$$F_{s,j+1/2} = \left(\sqrt{\frac{\rho}{\beta}}\mathcal{W}\right)_{j+1/2} = \sqrt{\frac{\rho_{j+1/2}}{\beta}}\mathcal{W}_{j+1/2}, \tag{4.34}$$

where

$$\rho_{j+1/2} = \frac{\overline{\rho}_j + \overline{\rho}_{j+1}}{2}, \quad \mathcal{W}_{j+1/2} = \frac{\overline{\mathcal{W}}_j + \overline{\mathcal{W}}_{j+1}}{2}. \tag{4.35}$$

3. Parabolic approximation:

$$F_{s,j+1/2} = \left(\sqrt{\frac{\rho}{\beta}}\mathcal{W}\right)_{j+1/2} = \sqrt{\frac{\rho_{j+1/2}}{\beta}}\mathcal{W}_{j+1/2}, \tag{4.36}$$

where

$$\begin{aligned} \rho_{j+1/2} &= \alpha_1\left(\overline{\rho}_{j-1} + \overline{\rho}_{j+2}\right) + \alpha_2\left(\overline{\rho}_j + \overline{\rho}_{j+1}\right), \\ \mathcal{W}_{j+1/2} &= \alpha_1\left(\overline{\mathcal{W}}_{j-1} + \overline{\mathcal{W}}_{j+2}\right) + \alpha_2\left(\overline{\mathcal{W}}_j + \overline{\mathcal{W}}_{j+1}\right), \\ \alpha_1 &= (1-\sqrt{3})/4, \quad \alpha_2 = (1+\sqrt{3})/4. \end{aligned} \tag{4.37}$$

The coefficients α_1 and α_2 are selected in order to preserve both average and variance in each time step as in Ref. [76].

4. Upwind approximation where $\mathcal{W}_j$ is taken as the stochastic velocity, so that a similar expression to the deterministic flux in Eq. (4.25) is obtained,

$$F_{s,j+1/2} = \left(\sqrt{\frac{\rho}{\beta}}\mathcal{W}\right)_{j+1/2} = \sqrt{\frac{\rho_j^E}{\beta}}\mathcal{W}^+_{j+1/2} + \sqrt{\frac{\rho_{j+1}^W}{\beta}}\mathcal{W}^-_{j+1/2}, \tag{4.38}$$

where

$$\mathcal{W}^+_{j+1/2} = \max\left(\mathcal{W}_{j+1/2}, 0\right), \quad \mathcal{W}^-_{j+1/2} = \min\left(\mathcal{W}_{j+1/2}, 0\right), \tag{4.39}$$

and $\mathcal{W}_{j+1/2} = (\overline{\mathcal{W}}_j + \overline{\mathcal{W}}_{j+1})/2$. The east and west density values ρ_j^E and ρ_j^W are computed as in the deterministic flux.

Stochastic time integrator and adaptive time steps

The derivation of the temporal integrator for the semi-discrete equation Eq. (4.19) is achieved by first re-writing the time-evolution equation of the density in the following vectorial form

$$d\overline{\boldsymbol{\rho}}(t) = \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))\, dt + \boldsymbol{\sigma}(\overline{\boldsymbol{\rho}}(t))\, \overline{\boldsymbol{\mathcal{W}}}\, dt, \tag{4.40}$$

where the vectors $\overline{\boldsymbol{\rho}}(t)$ and $\overline{\boldsymbol{\mathcal{W}}}$ contain the cell averages defined in Eqs (4.18)-(4.31) respectively, i.e. $\overline{\boldsymbol{\rho}}(t) = (\overline{\rho}_1(t), \overline{\rho}_2(t), \ldots, \overline{\rho}_n(t))$ and $\overline{\boldsymbol{\mathcal{W}}}(t) = (\overline{\mathcal{W}}_1(t), \overline{\mathcal{W}}_2(t), \ldots, \overline{\mathcal{W}}_n(t))$. 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 structure depends on the discretization of deterministic and stochastic fluxes.

From the vectorial equation Eq. (4.40) one can employ Itô's lemma to expand the two functions $\boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))$ and $\boldsymbol{\sigma}(\overline{\boldsymbol{\rho}}(t))$. Then, after integrating in time, one obtains the

Taylor expansion of the stochastic process. Truncating that expansion with an error $\mathcal{O}(dt^{1/2})$ and integrating between t and $t+\Delta t$, one can derive the following family of implicit-explicit Euler-Maruyama integrator, which component-wise form satisfies [216]

$$\overline{\rho}_j(t+\Delta t) = \overline{\rho}_j(t) + \left[(1-\theta)\,\mu_j(\overline{\boldsymbol{\rho}}(t)) + \theta\,\mu_j(\overline{\boldsymbol{\rho}}(t+\Delta t))\right]\Delta t + \sum_{k=1}^{n}\sigma_{jk}(\overline{\boldsymbol{\rho}}(t))\overline{W}_k(t)\Delta t. \tag{4.41}$$

The parameter θ allows to select 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 and 1.0 weak order of convergence.

Keeping in the expansion all the terms up to the order $\mathcal{O}(dt)$, and using a Runge-Kutta approach to approximate the derivative of $\boldsymbol{\sigma}$, one obtains a derivative-free family of implicit-explicit Milstein integrators with strong order 1.0 [216]. The component-wise version of this scheme reads

$$\overline{\rho}_j(t+\Delta t) = \overline{\rho}_j(t) + \left[(1-\theta)\,\mu_j(\overline{\boldsymbol{\rho}}(t)) + \theta\,\mu_j(\overline{\boldsymbol{\rho}}(t+\Delta t))\right]\Delta t + \sum_{k=1}^{n}\sigma_{jk}(\overline{\boldsymbol{\rho}}(t))\overline{W}_k(t)\Delta t \tag{4.42}$$

$$+ \frac{1}{\sqrt{\Delta t}}\sum_{l,m=1}^{n}\left[\sigma_{jm}(\boldsymbol{\Upsilon}_l(t)) - \sigma_{jl}(\overline{\boldsymbol{\rho}}(t))\right] I_{l,m}(t) \tag{4.43}$$

where the l-th row of the matrix $\boldsymbol{\Upsilon}$ is defined as

$$\boldsymbol{\Upsilon}_l(t) = \overline{\boldsymbol{\rho}}(t) + \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))\Delta t + \boldsymbol{\sigma}_l(\overline{\boldsymbol{\rho}}(t))\sqrt{\Delta t}, \tag{4.44}$$

and the multiple stochastic integrals are defined as $I_{l,m}(t) = \int_t^{t+\Delta t}\mathcal{W}^l\,\mathcal{W}^m dt$, where $\mathcal{W}_l$ and $\mathcal{W}_m$ are two white noises (for more details see Appendix C.1.1).

Stochastic time integration schemes with higher strong order have also been proposed in the literature [216]. 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, as last time integration

scheme we will employ here is the following explicit weak order 2.0 Runge-Kutta scheme:

$$\begin{aligned}\widetilde{\rho}_j(t+\Delta t) &= \overline{\rho}_j(t) + \frac{1}{2}\left[\mu_j(\boldsymbol{\Upsilon}(t)) + \mu_j(\overline{\boldsymbol{\rho}}(t))\right]\Delta t + \Phi(t)\\ \overline{\rho}_j(t+\Delta t) &= \overline{\rho}_j(t) + \frac{1}{2}\left[\mu_j(\widetilde{\boldsymbol{\rho}}(t+\Delta t)) + \mu_j(\overline{\boldsymbol{\rho}}(t))\right]\Delta t + \Phi(t)\end{aligned} \tag{4.45}$$

where the vector s$\Phi(t)$ and $\boldsymbol{\Upsilon}$ are reported in Appendix C.1.1)

The order of that convergence for the temporal integrators above can be measured in strong and weak sense. Specifically, strong and weak errors at a particular time τ and for a group of trajectories $\Gamma = \{\gamma_1, \gamma_2, \ldots, \gamma_m\}$ are defined as

$$\epsilon_s = \langle|\overline{\boldsymbol{\rho}}^{\gamma}(\tau) - \overline{\boldsymbol{\rho}}^{\gamma}_{exact}(\tau)|\rangle_{\gamma\in\Gamma} \quad \text{and} \quad \epsilon_w = \left|\langle\overline{\boldsymbol{\rho}}^{\gamma}(\tau)\rangle_{\gamma\in\Gamma} - \langle\overline{\boldsymbol{\rho}}^{\gamma}_{exact}(\tau)\rangle_{\gamma\in\Gamma}\right|, \tag{4.46}$$

where $\overline{\boldsymbol{\rho}}^{\gamma}(\tau)$ refers to the numerical density cell averages at time τ following the γ trajectory, $\overline{\boldsymbol{\rho}}^{\gamma}_{exact}(\tau)$ denotes the exact (or reference) solution of the stochastic equation, the ensemble average $\langle\cdot\rangle$ is taken over the trajectories $\gamma \in \Gamma$, and $|\cdot|$ denotes the L^1-norm.

In Fig. 4.2 we test strong and weak convergence for the described stochastic integrators for a simple scenario. Specifically, we simulate the system in Eq. (4.40) in the case of a geometric Brownian motion, for which $\boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t)) = -\overline{\boldsymbol{\rho}}(t)$ and $\boldsymbol{\sigma}(\overline{\boldsymbol{\rho}}(t)) = 0.5\overline{\boldsymbol{\rho}}(t)$. In this scenario, the temporal evolution of the density in a cell j, which is independent from the rest of cells, follows

$$d\overline{\rho}_j(t) = -\overline{\rho}_j dt + 0.5\overline{\rho}_j\overline{\mathcal{W}}_j dt, \tag{4.47}$$

with the cell averaged white noise $\overline{\mathcal{W}}_j$ defined as in Eq. (4.31). For the simulation we select $\overline{\rho}_j(0) = 1$. The geometric Brownian motion is chosen to compute strong and weak errors since its exact solution is known [217]. The results in Figs 4.2(a-b) depict the strong and weak order of convergence for the previous temporal integrators. Regarding the strong order of convergence, 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.

On the weak order, the whole families of Euler-Maruyama and Milstein integrators are expected to have an order of 1.0, while the Runge-Kutta an order of 2.0. This is validate numerically for all the schemes, with the semi-implicit methods outperforming and showing an order between 1.0 and 2.0.

Fig. 4.2(c) shows the relation between CPU time and number of cells n for 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 can observe a lower

CPU time for Runge-Kutta, if compared with all the other integrators except for the explicit Euler-Maruyama.

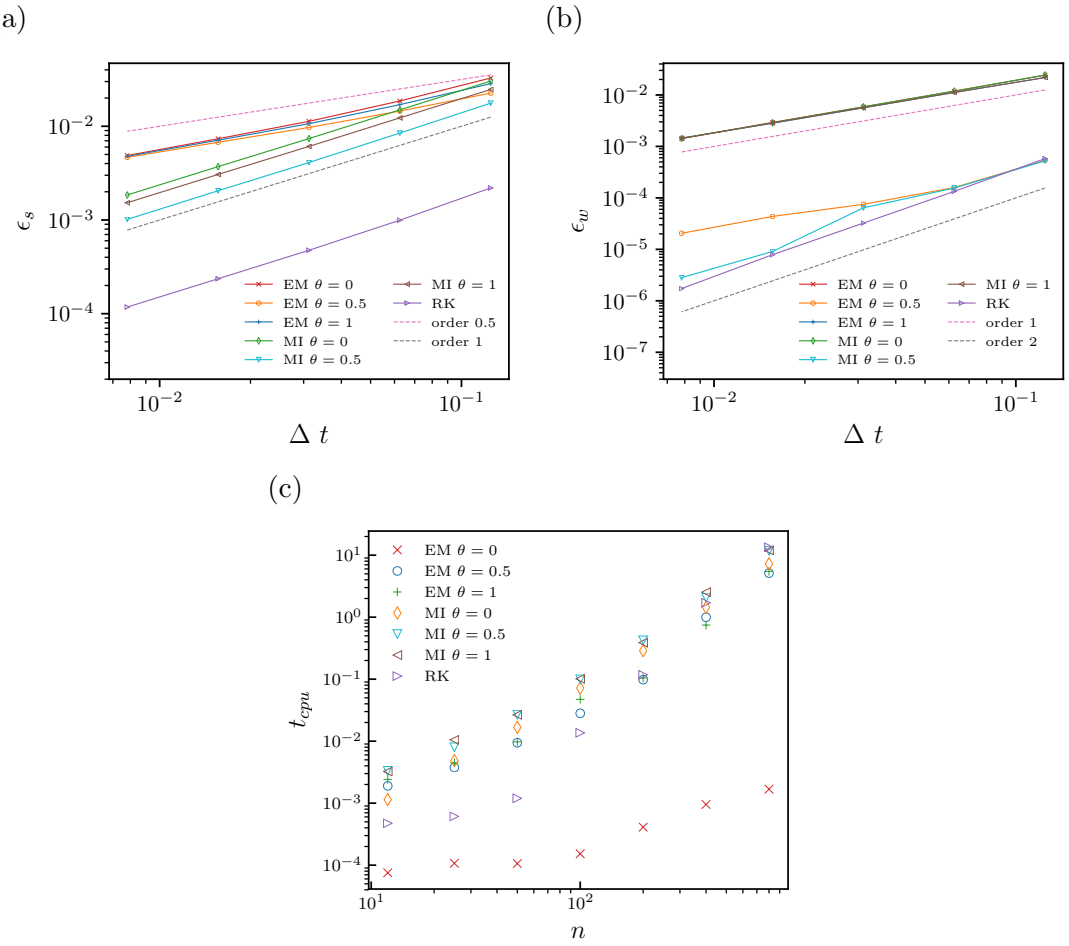

Figure 4.2. Strong (a) and weak (b) errors convergence for geometric Brownian motion. In (c) we report the CPU time for each time-integration scheme as a function of the number of cells n. In the legend EM, MI and RK stand for Euler-Maruyama, Milstein and Runge-Kutta, respectively.

Positivity of the density through an adaptive time step

Previous works in the literature have developed different techniques to addressed the issue of density positivity. In the context of fluctuating hydrodynamics, Kim *et al.* [71] have proposed a smoothed Heaviside cutting function which reduces the chances of density to become negative. However, despite reducing the chances of having negative

density values, this approach does not guarantee positive densities.

A different way to preserve density positivity relies on the concept of Brownian bridge, introduced to solve stochastic differential equations with variable time steps [218]. Unlike deterministic differential equations, stochastic equations requires to follow the Brownian path that is formed after evaluating the normal distributions Eq. (4.32). Specifically, when advancing from a time t by Δt_1 in our simulation, if the numerical density in one of the nodes j becomes negative, one cannot just simply repeat the time step with a shorter $\Delta t_2 < \Delta t_1$ in order to preserve its positivity. The Brownian path generated after the first trial of advancing Δt_1 is to be preserved. Brownian bridge allows to take the statistical information at $t+\Delta t$ into account when repeating the time step [219, 220]. In fact, it provides a way to compute $\overline{\mathcal{W}}_j$ (Eq. (4.31)) at an intermediate time step $t+\Delta t/2$, by means of the formula:

$$\overline{\mathcal{W}}_j\left(t+\frac{\Delta t}{2}\right)-\overline{\mathcal{W}}_j(t)=\frac{\overline{\mathcal{W}}_j\left(t+\Delta t\right)-\overline{\mathcal{W}}_j\left(t\right)}{2}+\mathcal{N}\left(0,\frac{\Delta t}{4}\right). \tag{4.48}$$

Consequently, in our simulations we employ adaptive time steps. If the density becomes negative after a time step, Δt is halved to compute the intermediate time step with the Brownian bridge in Eq. (4.48). If that intermediate state leads to further negative densities, the Brownian bridge is applied again, and so on. A pseudo code to implement the Brownian bridge is written in Algorithm 1.

Algorithm 1: Adaptive time step algorithm adopted to preserve density positivity.

```
   Input: ρ(t)
   Output: ρ(t + Δt)
 1 NegativeDensity=True;
 2 Δt = Δt_0;
 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);
```

Boundary conditions

Here we discuss the numerical implementation of boundary conditions in periodic systems, confined systems and open systems. Periodic boundary conditions along the direction α satisfy

$$\rho_0^\alpha = \rho_N^\alpha \tag{4.49}$$

An hard wall locate at the cell i, with normal direction α, is modelled with the following no-flux conditions:

$$F_i^\alpha \pm 1/2 = S_i^\alpha \pm 1/2 = 0. \tag{4.50}$$

Open systems in thermal and chemical equilibrium with an external reservoir are consistent with a μVT ensemble, with constant grand potential $\Omega[\rho(\mathbf{r})] = \mathcal{F}[\rho(\mathbf{r})] + \int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} - \mu \int \rho(\mathbf{r})d\mathbf{r}$, where $\mathcal{F}[\rho(\mathbf{r})]$ is the Helmholtz free-energy functional, $V(\mathbf{r})$ is the external potential acting on the system and μ is the chemical potential. Using the fact that the functional derivative of Ω with respect to $\rho(\mathbf{r}$ is null at equilibrium, we obtain $\frac{\delta\mathcal{F}[\rho(\mathbf{r})]}{\delta\rho(\mathbf{r})} + V(\mathbf{r}) = U = \mu$. Since the system is assumed to be in contact (at equilibrium) with a reservoir at temperature T and chemical potential U, the corresponding boundary condition is then translated in the following constrains:

$$U_i^\alpha = \mu_{res}, \tag{4.51}$$

$$\rho_i^\alpha = \rho_{res}, \tag{4.52}$$

with μ_{res} and ρ_{res} being chemical potential and density of the reservoir, respectively. Some of the boundary conditions discussed are employed in the numerical applications below.

4.1.3. Numerical application I: Ideal-gas system in equilibrium

As a first system, we solve Eq. (4.12) embedded with a classical ideal-gas free-energy

$$\mathcal{E}[\rho] = \int \rho/\beta \left(\ln(\rho) - 1\right) dx. \tag{4.53}$$

This leads to a diffusive equation with multiplicative noise of the form

$$\partial_t \rho = \Delta\rho/\beta + \boldsymbol{\nabla} \cdot \left[\sqrt{\rho/\beta}\mathcal{W}(x,\rho)\right]. \tag{4.54}$$

In this system configuration we compare the different spacial and temporal numerical methods by focusing on the following statistical properties at equilibrium: variance, space-correlation, time-correlation and structural factor. These metrics are commonly employed in the literature to validate stochastic numerical schemes for fluctuating hydrodynamics [71, 76, 221].

The density fluctuations of a system with a 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 Ref. [74], the variance is given by:

$$\langle \delta N^2 \rangle = -T \frac{\bar{N}^2}{V^2} \left(\frac{\partial V}{\partial p} \right)_T \tag{4.55}$$

with T and p being temperature and pressure of the system, respectively. For an ideal gas, employing the equation of state (in reduced units) $pV = NT$, we obtain $\langle \delta N^2 \rangle = N$.

In case of infinite systems, the fluctuations of an ideal gas are spatially uncorrelated, namely $\langle \delta\rho_i(t)\delta\rho_j(t) \rangle = \langle \delta\rho^2 \rangle \delta^K_{ij}$. However, for finite systems the conservation of mass introduces some correlations [76]. Expressing the space correlations of density fluctuations as $\langle \delta\rho_i(t)\delta\rho_j(t) \rangle = A\delta^K_{ij} + B$, then the conservation of mass gives $\sum_i \langle \delta\rho_i(t)\delta\rho_j(t) \rangle = 0$, which corresponds to the constrain $B = -A/n$, 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:

$$\langle \delta\rho_i(t)\delta\rho_j(t) \rangle = \langle \delta\rho^2 \rangle \left(\delta^K_{ij} - \frac{1}{n} \right). \tag{4.56}$$

Standard deviation As shown above, the theoretical standard deviation of the diffusion SPDE Eq. (4.54) applied in finite systems in equilibrium satisfies

$$\sqrt{\langle \delta\rho^2 \rangle}_{\text{theory}} = \frac{\langle \rho \rangle}{\sqrt{N_c}} \sqrt{1 - \frac{1}{n}}. \tag{4.57}$$

Despite the fact that $\sqrt{\langle \delta\rho^2 \rangle}_{\text{theory}}$ holds for all N_c, previous studies [208, 221] have pointed out that there should be a minimum of 5-10 particles per cell to recover the microscopic statistical properties by means of a continuum approach. This is due to the fact that, with a such low number of particles per cell, particle fluctuations are not accurately modelled with the multiplicative noise in Eq. (4.54).

In Fig. 4.3 we report the comparison between different temporal (a) and spatial (b) discretizations against the theoretical standard deviation in Eq. (4.57) and the one com-

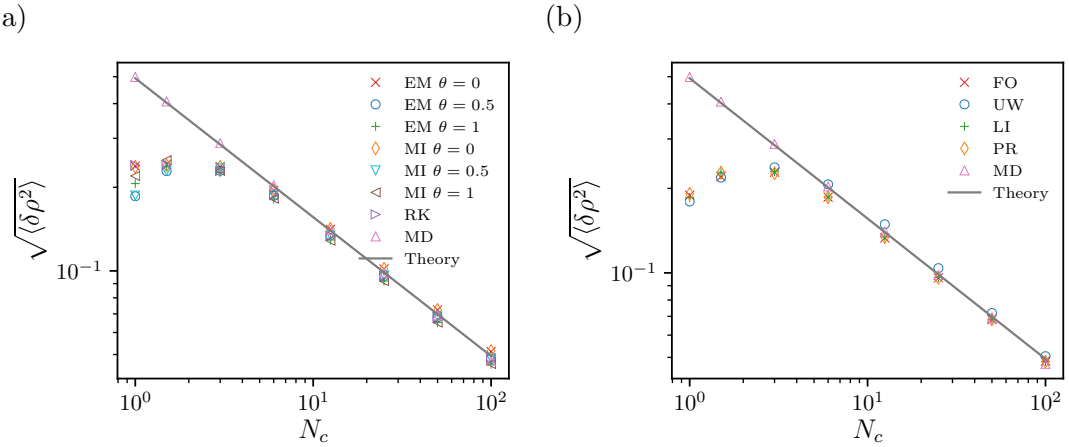

Figure 4.3. The standard deviation $\sqrt{\langle\delta\rho^2\rangle}$ is plotted against the number of particles per cell N_c, for an ideal-gas system at equilibrium. (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward Eq. (4.33), UW: Upwind Eqs (4.38)-(4.39), LI: Linear Eqs (4.34)-(4.35), PR: Parabolic Eqs (4.36)-(4.37), MD: Molecular dynamics.

puted from molecular dynamics. In both plots we can observe that all the schemes approximate well the standard deviation for $N_c > 5$. However, below that number of particles per cell the numerical standard deviation deviate from the expected ones. This result chimes with the minimum number of $5 - 10$ particles per cell necessary to recover the statistical properties in fluctuating hydrodynamics. There are no remarkable differences between the temporal integrators or spatial discretizations for the stochastic flux.

Time correlations The objective of this test is to measure the time-correlation of the density in a cell of the domain. The normalized time correlation function is defined as

$$C_T(t) = \frac{\langle\delta\rho_i(t)\delta\rho_i(0)\rangle}{\langle\delta\rho_i(0)\delta\rho_i(0)\rangle}, \tag{4.58}$$

where $\delta\overline{\rho}_i(t) = \overline{\rho}_i(t) - \overline{\rho}$. This is expected to decay in time for the diffusion equation Eq. (4.54) in equilibrium, although no explicit theoretical expression is known in the literature. Previous studies [76] have compared the numerical results with the Fourier transform of the time correlation Eq. (4.58), 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 from the numerical simulations, leading to a lack of agreement in the results in Ref. [76]. For this reasons we

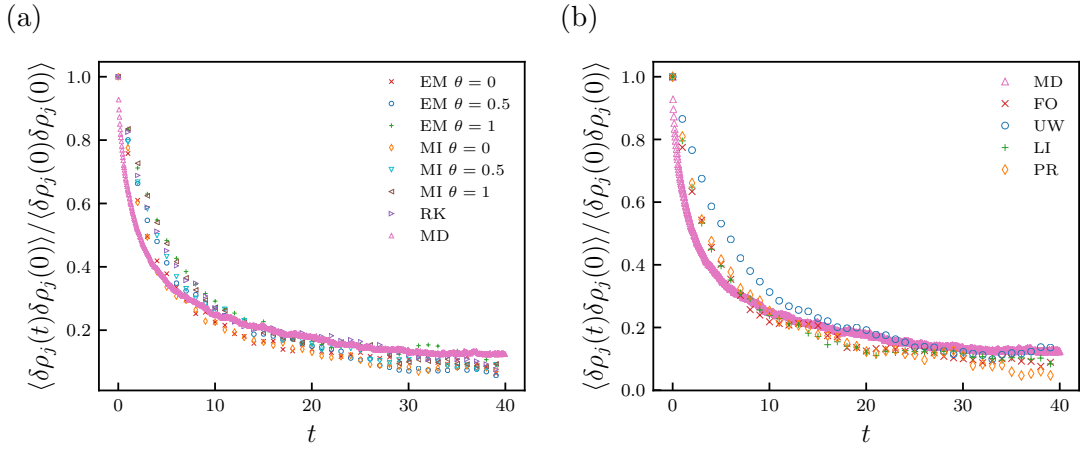

Figure 4.4. Decay of the normalized time-correlation C_T (Eq. (4.58)) for an ideal-gas system in equilibrium (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward Eq. (4.33), UW: Upwind Eqs (4.38)-(4.39), LI: Linear Eqs (4.34)-(4.35), PR: Parabolic Eqs (4.36)-(4.37), MD: Molecular dynamics.

compare the results obtained from the numerical schemes against MD simulations, which indeed take into account the finite-size effects of the numerical simulation. The results are displayed in Fig. 4.4. Regarding the temporal integrators, the explicit ones ($\theta = 0$) tend to behave closer to the MD simulation for the initial short times, while the implicit ones ($\theta = 1$) provide a better approximation in the long-time regimes. Moreover, for different spatial discretizations of the stochastic flux the upwind one deviates the most from the MD simulation, while the rest of them behaves similarly.

Spatial correlations This test aims to understand whether the proposed numerical schemes satisfy the exact spatial correlation for finite-size systems

$$\langle \delta\rho_i(t)\delta\rho_j(t)\rangle = \frac{\langle\rho\rangle}{\Delta x}\left(\delta_{ij} - \frac{1}{n}\right). \tag{4.59}$$

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. Fig. 4.5 shows the normalized spatial correlation

$$C_S(t) = \frac{\langle \delta\rho_i(t)\delta\rho_j(t)\rangle}{\langle \delta\rho_i(0)\delta\rho_i(0)\rangle} \tag{4.60}$$

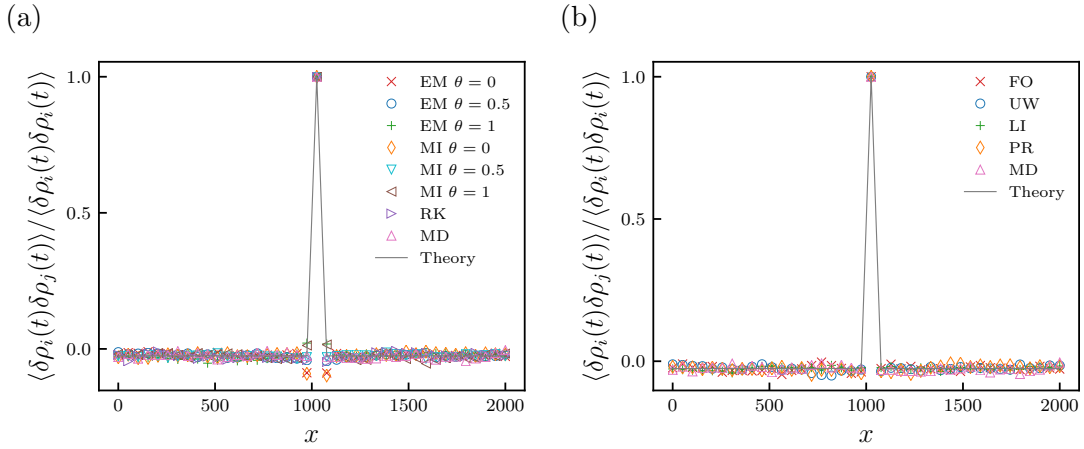

Figure 4.5. Normalized spatial correlation Eq. (4.60) for an ideal-gas system in equilibrium. (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward Eq. (4.33), UW: Upwind Eqs (4.38)-(4.39), LI: Linear Eqs (4.34)-(4.35), PR: Parabolic Eqs (4.36)-(4.37), MD: Molecular dynamics.

with $\delta\overline{\rho}_i(t) = \overline{\rho}_i(t) - \overline{\rho}$, plotted for each of the numerical schemes, the MD simulation and the exact expression (Eq. (4.59)). Most of the temporal integrators and spatial discretizations approximate adequately the theoretical expression Eq. (4.59), as it is depicted in Fig. 4.5. 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 corresponding semi-implicit schemes perform well.

Structure factor The structure factor S at equilibrium fo an ideal gas at equilibrium is known and can be used to evaluate the accuracy of temporal integrators and spatial discretizations. Its theoretical expression for an ideal gas without external potential is given by Eq. (C.19), i.e. $S/\langle\rho\rangle = 1$. The discrete spatial Fourier transform of the density satisfies

$$\hat{\rho}(\lambda) = \frac{1}{n}\sum_j \overline{\rho}_j e^{-i\lambda x_j}. \tag{4.61}$$

Subsequently, the structure factor is computed as

$$S(\lambda) = \frac{\langle \delta\hat{\rho}(\lambda)\, \delta\hat{\rho}^*(\lambda)\rangle}{n\Delta x}, \tag{4.62}$$

where $\delta\hat{\rho}(\lambda) = \hat{\rho}(\lambda) - \langle\hat{\rho}(\lambda)\rangle$, and $\hat{\rho}^*$ denotes the complex conjugate of $\hat{\rho}$ (more details at Appendix C.1.4).

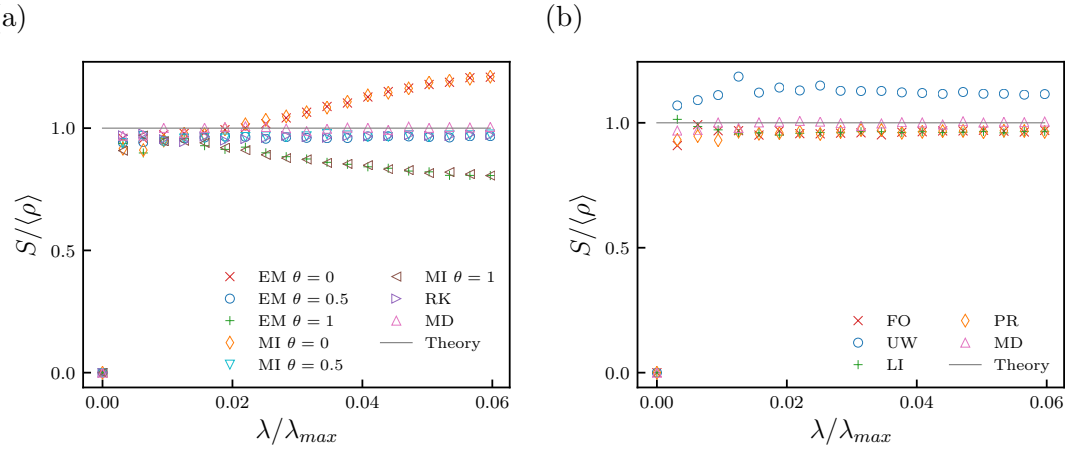

Figure 4.6. Structure factor Eq. (4.62) for an ideal-gas system in equilibrium. (a) Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$). (b) Spatial discretizations of the stochastic flux. FO: Forward Eq. (4.33), UW: Upwind Eqs (4.38)-(4.39), LI: Linear Eqs (4.34)-(4.35), PR: Parabolic Eqs (4.36)-(4.37), MD: Molecular dynamics.

We report the structure factor at equilibrium computed numerically in Fig. 4.6. The theoretical value of the structure factor, along with the performed MD simulation, allows to evaluate the accuracy of the temporal integrators and spatial discretizations. Fig. 4.6(a) shows that the explicit Euler-Maruyama and Milstein temporal integrators overestimate the structure factor for large λ, while their implicit versions underestimate it for large λ. However, the semi-implicit and Runge-Kutta schemes are able to accurately approximate the expected structure factor, 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 Fig. 4.1. Regarding the spatial discretizations of the stochastic flux, a significant error is introduced by the upwind scheme. In addition, the forward discretization introduces some error at lower λ. The rest of discretizations approximate the theoretical value correctly, with the small damping already mentioned.

Discussion on temporal integrators and spatial discretization of the stochastic flux Based on the previous tests, here we discuss the temporal integrator and spatial discretization of choice for the next numerical applications. Regarding the temporal integrators, both the fully explicit and implicit Euler-Maruyama and Milstein show significant deviations in time correlation (Fig. 4.4), spatial correlation (Fig. 4.5) or structure factor (Fig. 4.6). Their semi-implicit versions and the Runge-Kutta behave similarly in all the tests, and approximate adequately the theoretical and MD results.

Their computational costs is compared in Fig. 4.2. Runge-Kutta scales as $\mathcal{O}(n^3)$, while the semi-implicit Euler-Maruyama and Milstein as $\mathcal{O}(n^2)$. However, the comparison clearly shows how for $n < 100$ the Runge-Kutta cost is lower than the semi-implicit schemes, while for $n > 100$ is higher. As a result, since in the simulations of this work $n < 100$, we select the Runge-Kutta temporal integrator.

On the spatial discretization of the stochastic flux, the upwind choice does not approximate well time correlation and structure factor, while the forward approximation shows a significant error in the structure factor for short λ. Thus the best performing choices are the linear and parabolic approximations, which behave similarly in all the test cases. We select the linear approximation due to its lower computational cost.

4.1.4. Numerical application II: Ideal-gas system out of equilibrium

Here we consider a system in a double well potential, i.e. with free-energy:

$$\mathcal{E}[\rho] = \int \rho/\beta \left(\ln(\rho) - 1\right) dx + \int V(x)\rho \, dx, \tag{4.63}$$

with external potential satisfying

$$V(x) = 5\left[\left(\frac{x}{n\Delta x/2}\right)^4 - \left(\frac{x}{n\Delta x/2}\right)^2\right]. \tag{4.64}$$

This is similar to the numerical applications for deterministic gradient flow equations in Ref. [209]. We compare the numerical scheme for the FDDFT Eq. (4.16) against MD simulations. For the sake of completeness, we also include a comparison with the corresponding deterministic equation obtained in the mean field limit [68].

Figure 4.7(a) shows the average density profile at different times. Results for FDDFT, DDFT and MD are consistent, confirming that both FDDFT and DDFT are able to accurately model the average density profile. Regarding the standard deviation (Fig. 4.7(b)), FDDFT results matches with MD and theory (Eq. (4.57)), while the DDFT, being deterministic, does not provide any information on the standard deviation. Figure 4.7(c) shows the spatial correlation computed as in Eqs (4.59)-(4.60). Both MD and FDDFT results approximate correctly the finite-size theoretical expression. As expected, DDFT does not have any spatial correlation due to the lack of fluctuations. Finally, Fig. 4.7(d) exhibits the time-evolution of the free-energy functional, which depends on the ensemble average density. For DDFT a monotonous decay in time is visible, while for MD and FDDFT there are short increases of the free-energy triggered by the fluctuations, in spite

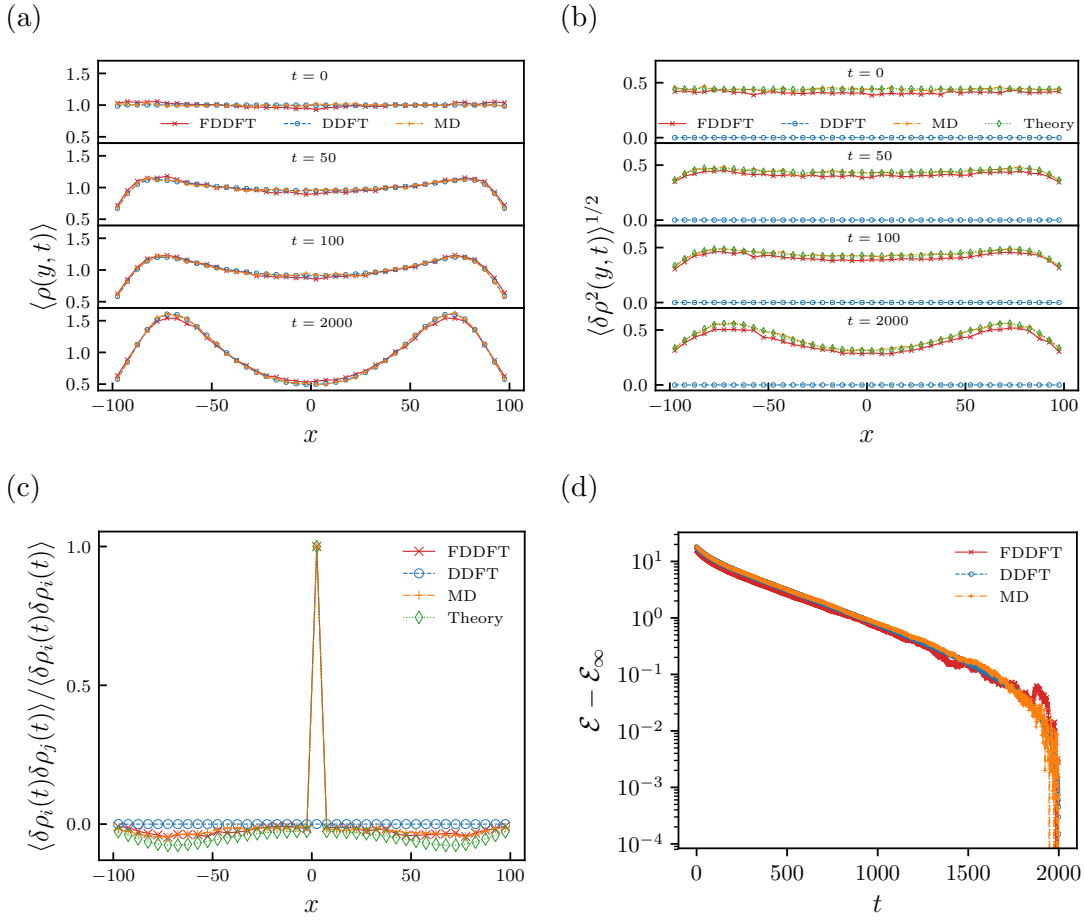

Figure 4.7. Time evolution of mean density (a) and density standard deviation (b) computed with FDDFT, DDFT and MD simulations. A comparison in terms of steady-state spatial correlations is reported in (c). In (d), we report the decrease in time of the energy functional of the mean density.

of the fact that during the evolution there is a general decay in the free-energy.

4.1.5. Numerical application III: Homogeneous nucleation of Lennard-Jones systems

In this section we employ FDDFT to study the homogeneous vapor-liquid transition of a Lennard- Jones fluid. In the framework of DFT, the density profiles of a one-dimensional open system that can exchange particles with a reservoir at constant temperature and chemical potential μ, can be obtained from a numerical minimization of the grand free-

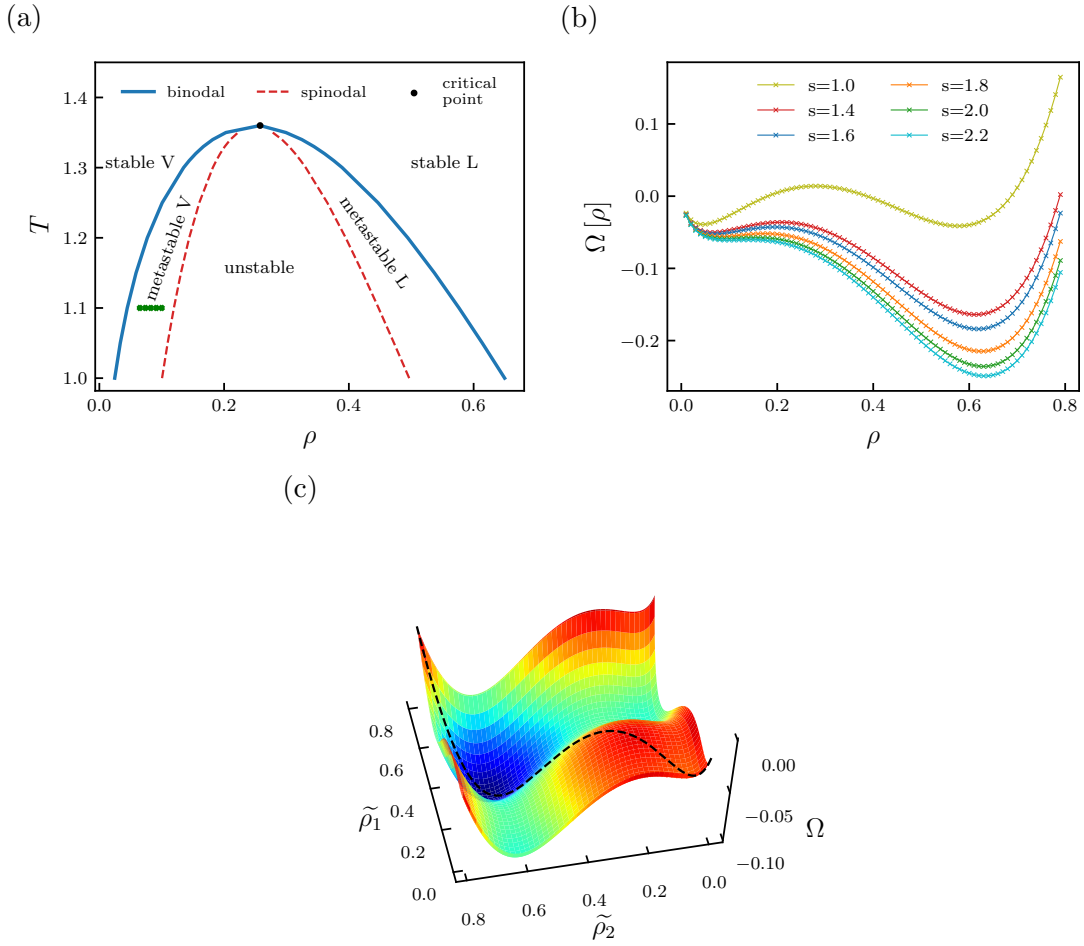

Figure 4.8. Panel (a) shows the bulk phase diagram for the discretized LJ system. Panel (b) shows the grand free-energy landscape as function of the system density for some supersaturation ratios adopted in this study. In (c) we provide an example of free-energy landscape for systems with a non-uniform 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.

energy functional

$$\Omega[\rho(x)] = \mathcal{F}[\rho(x)] + \int (V(x) - \mu)\, \rho(x)\, dx. \tag{4.65}$$

For an system of interacting particles, $\mathcal{F}[\rho(x)]$ is not analytically derivable from first principle, except in few cases, i.e. hard spheres fluids. In the remaining cases, $\mathcal{F}[\rho(x)]$ is either numerically obtained from atomistic simulations or is approximated by means of

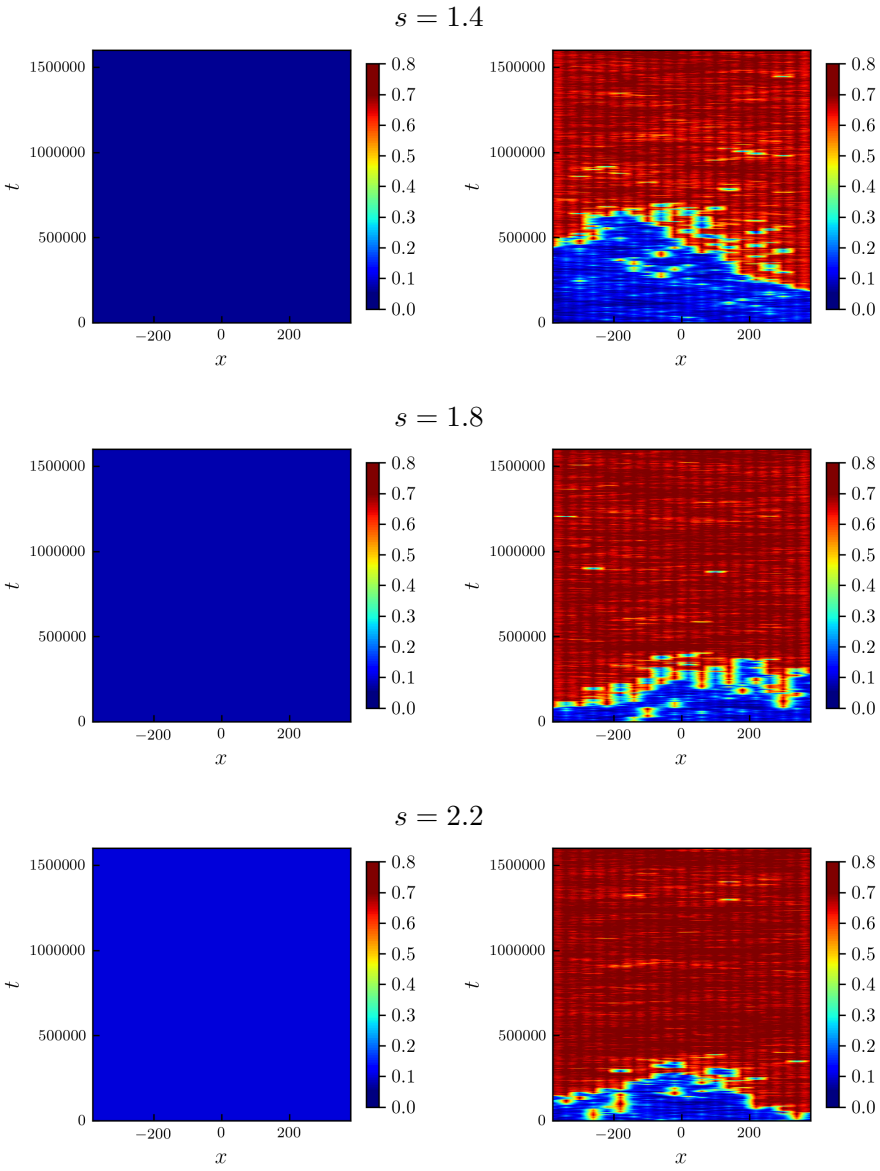

Figure 4.9. Homogeneous nucleation of a vapour LJ system in metastable conditions with supersaturation ratio s. We report on the left column the mean field evolution of density, while on the right a single realization of the stochastic dynamics. The colour map refers to density values.

perturbation expansions around a known free-energy [49]. Similarly to previous works on DDFT [55, 59], we approximate $\mathcal{F}[\rho(x)]$ of a Lennard-Jones fluid according to the first order Barker–Henderson perturbation theory expansion around the hard-sphere fluid

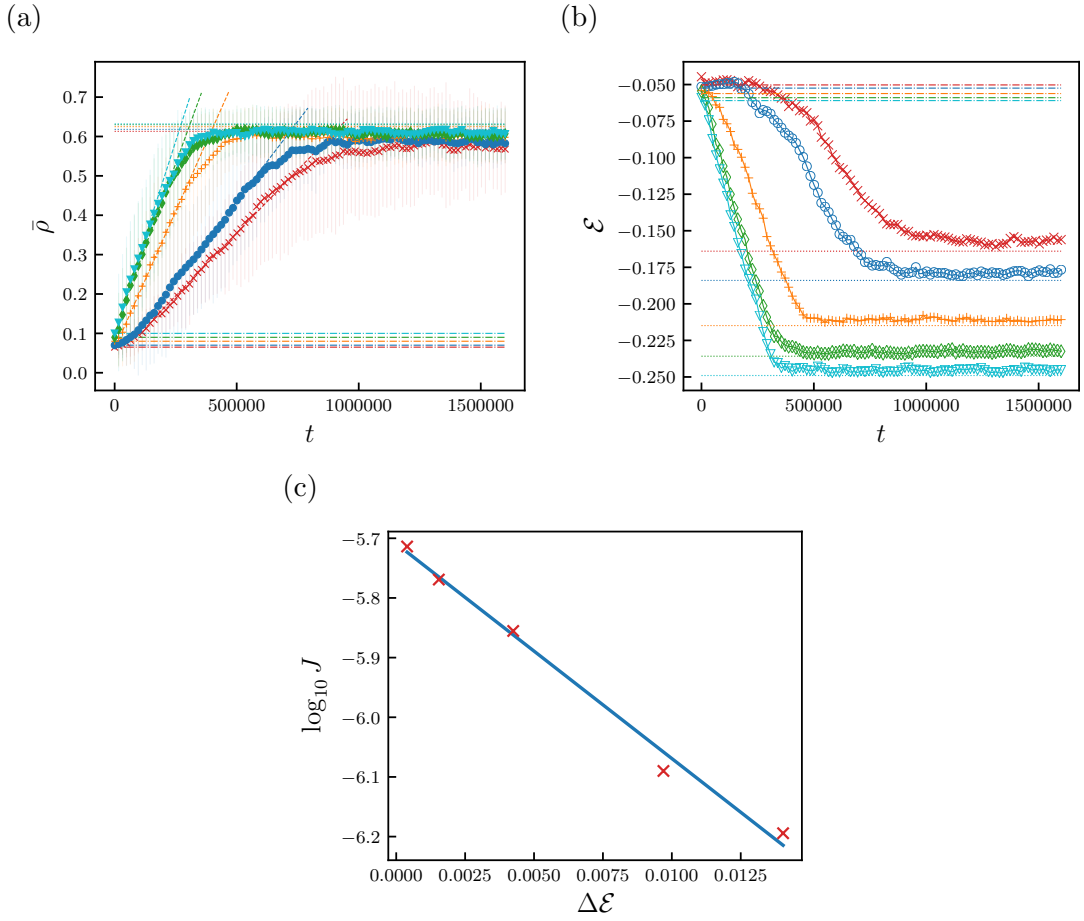

Figure 4.10. On panel (a), we report the evolution in time of the average system density for the supersaturation ratios adopted in this study. Panel (b) shows the time-evolution of the system free-energy for the supersaturation ratios adopted in this study. In (c), the nucleation growth rate J is plotted against the free-energy barrier $\Delta\mathcal{E}$.

free-energy [222], namely as

$$\mathcal{F}[\rho(x)] = \int \{f_{\text{ID}}[\rho(x)] + f_{\text{HS}}[\rho(x)]\}\, dx + \frac{1}{2}\int\int \rho(x)\rho(x')W(x,x')\; dx\; dx' \tag{4.66}$$

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

$$f_{\text{ID}}[\rho(x)] = k_{\text{B}}T\rho\left(\ln\left(\lambda^3\rho\right) - 1\right), \tag{4.67}$$

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 [125]

$$f_{HS}[\rho(x)] = k_{\mathrm{B}}T\left(\frac{4\eta - 3\eta^2}{(1-\eta)^2}\right), \quad \text{with} \quad \eta = \frac{\pi}{6}\rho\sigma^3 \tag{4.68}$$

with σ being the hard sphere diameter (set to 1 in this work). Finally the LJ (attractive) contributions are taken into account by the following expression:

$$W(x,x') = \begin{cases} -1.2\ \pi\ \epsilon & \text{if} \quad \mid x - x' \mid \leq 1 \\ \pi\ \epsilon\ \left(0.8\ \mid x - x' \mid^{-10} - 2\ \mid x - x' \mid^{-4}\right) & \text{otherwise} \end{cases} \tag{4.69}$$

which is derived by integrating along y and z the 12-6 LJ potential [59].

In order to analyse the vapor-liquid phase transitions, first we compute the coexisting density profiles. The coexisting values of vapor and liquid density (on the binodal line) are denoted as ρ_v and ρ_l respectively, and are obtained by solving the following system of equations:

$$\begin{cases} \frac{\partial\Omega}{\partial\rho}\Big|_{\rho_v} = \frac{\partial\Omega}{\partial\rho}\Big|_{\rho_l} = 0 \\ \Omega\left[\rho_v\right] - \Omega\left[\rho_l\right] = 0 \end{cases} \tag{4.70}$$

The meta-stable 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:

$$\frac{\partial^2\Omega}{\partial\rho^2}\Big|_{\rho_v} = \frac{\partial^2\Omega}{\partial\rho^2}\Big|_{\rho_l} = 0 \tag{4.71}$$

Finally, the bulk critical point is given by the intersection between binodal and spinodal lines, and thus is computed as

$$\frac{\partial\Omega}{\partial\rho}\Big|_{\rho_c,T_c} = \frac{\partial^2\Omega}{\partial\rho^2}\Big|_{\rho_c,T_c} = 0 \tag{4.72}$$

In Fig. 4.8(a), we report the bulk phase diagram obtained from the discretized grand free-energy of the LJ fluid. Solid curves depicts the binodal, i.e. the locus of liquid–gas coexistence, while dashed curves depicts the spinodal, i.e. the boundary between metastable and unstable regions. The black circle designates the bulk critical point at $\rho_c \sim 0.3$ and $T_c \sim 1.35$.

The supersaturation ratio is defined as $s = \rho/\rho_v$ with ρ_v denoting the vapor coexistence density at a certain temperature. Here we simulate the nucleation of vapor systems different initial supersaturation ratios. Figure 4.8(b) shows the free-energy landscape as a function of the bulk density for such systems. Coexistence ($s = 1$) corresponds to two stable basins, which means that the system has equal probability of being either in the vapor or liquid phase. Increasing the supersaturation ratio, the high-density basin (corresponding to the liquid phase) enhances its stability, thus leaving the vapor density in a metastable condition. Moreover the energy barrier that the system has to overcome to pass from vapor to liquid phase decreases with s, until it becomes null at a supersaturation corresponding to the spinodal line. In such condition only one minimum of the grand free-energy exists.

A physical description of the phase transition would consist of an initial and final uniform system densities, but also of a non-uniform density field during the transition. This means that the bulk grand free-energy in Fig. 4.8(b), being only valid for uniform densities, describes the system only in the initial and final stages, 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. As an example, in Fig. 4.8(c) we report our LJ grand free-energy for a non uniform system, constrained to have only two varying densities $\widetilde{\rho}_1 = \{\rho_1 = \cdots = \rho_{n/2}\}$ and $\widetilde{\rho}_2 = \{\rho_{n/2+1} = \cdots = \rho_n\}$. The bulk free-energy is then recovered for $\widetilde{\rho}_1 = \widetilde{\rho}_2$ (dotted black line).

A single trajectory of the vapour to liquid transition, at different supersaturation ratios, is shown in Fig. 4.9. For comparison purpose, we perform simulations of the FDDFT and its mean field (deterministic) counterpart. In order for the transition to occur, the system free-energy has to overcome an energy barrier. Such passage requires a local injection of energy that is triggered by fluctuations. As a consequence, the mean field approach fails to describe this transition. As expected, by looking at the energy barrier in Fig. 4.8, the transition is favoured by higher supersaturation ratios.

This trend is quantitatively analysed in Fig. 4.10, where we report the results obtained performing the average of 10 nucleation trajectories for each supersaturation ratio. Figure 4.10(a) 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. Moreover, the values of the free-energy evaluated at each time as function of the average density is reported in Fig. 4.10(b). The initial free-energy value, corresponding to the vapour metastable basin, evolves in time until it reaches the more stable liquid basin, as predicted by Fig. 4.8(b). It is

interesting to notice that the passage between the two basin implies a slight increase in the free-energy due to the energy barrier overcome by the density field fluctuations. The average density kinetic 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. The growth period exhibits a linear-like trend, with slopes representing the nucleation growth rate J. As reported in the plot in Fig. 4.10(c), an Arrhenius-like relation is observed between J and the grand free-energy barrier $\Delta\mathcal{E}$, i.e. $J \propto \exp\{-\frac{\Delta\mathcal{E}}{T}\}$.

The numerical applications above highlight the importance of fluctuations in the study of key phenomena, such as nucleation. In the next section we will propose numerical methods for the deterministic variation of FDDFT embedded with additional complexities, mainly non-homogeneous diffusion.

4.2. Non-homogeneous DDFT: A finite volume approach [5]

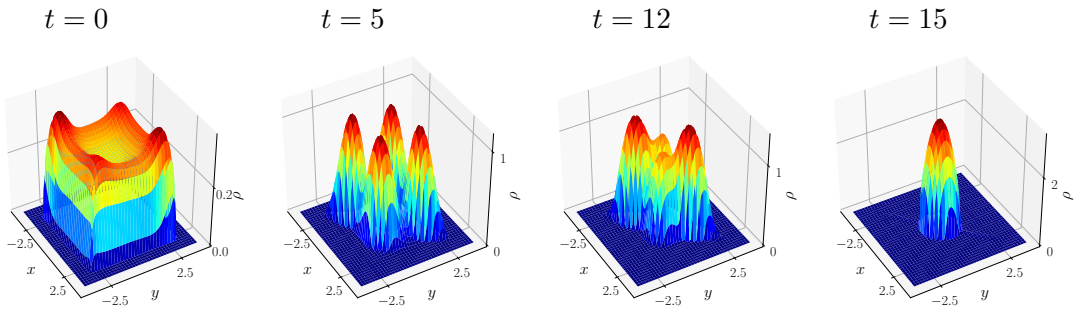

Figure 4.11. Time evolution of the density of a 2-D system characterized by a nonlocal attraction kernel and non uniform diffusion simulated using our finite-volume scheme.

The deterministic version of the time-evolution equation for the density field in Eq. 4.5 [68, 70] reads:

$$\begin{cases} \partial_t \rho(\boldsymbol{r};t) = \boldsymbol{\nabla}_{\boldsymbol{r}} \cdot \left[\boldsymbol{\theta}_0^{-1} \rho(\boldsymbol{r};t) \boldsymbol{\nabla}_{\mathbf{r}} \dfrac{\delta \mathcal{E}[\rho]}{\delta_{\mathbf{r}} \rho(\boldsymbol{r};t)} \right] & \mathbf{r} \in \mathbb{R}^d, \ t > 0, \\ \rho(\mathbf{r};0) = \rho_0(\mathbf{r}), \end{cases} \tag{4.73}$$

This gradient-flow equation (Eq. (4.73)) has an important property, that the value free-energy $\mathcal{E}(\rho)$ decreases monotonically over time [72, 73]. For Eq. (4.73) this can be shown by taking the time-derivative of the free-energy:

$$\begin{aligned} \frac{d}{dt} E(\rho) &= \int_{\mathbb{R}^d} (V + f'(\rho) + K \star \rho) \frac{\partial \rho}{\partial t} d\mathbf{r} \\ &= \int_{\mathbb{R}^d} (V + f'(\rho) + K \star \rho) \boldsymbol{\nabla} \cdot (\boldsymbol{\theta}_0^{-1} \rho \boldsymbol{\nabla} (f'(\rho) + V + K \star \rho)) d\mathbf{r} \\ &= - \int_{\mathbb{R}^d} \boldsymbol{\nabla} (V + f'(\rho) + K * \rho)^T \boldsymbol{\theta}_0^{-1} \rho \boldsymbol{\nabla} (V + f'(\rho) + K \star \rho) d\mathbf{r} \\ &= - \int_{\mathbb{R}^d} \rho \mathbf{v}^T \boldsymbol{\theta}_0^{-1} \mathbf{v} d\mathbf{r} < 0, \end{aligned} \tag{4.74}$$

where $\mathbf{v}$ can be interpreted as the velocity in the gradient flow Eq. (4.73) and satisfies

[5] This section is based on a joint work with Julien Mendes, who materially produced the figures of this section. It includes material from J. Mendes, A. Russo, S. P. Perez & S. Kalliadasis. A finite-volume scheme for gradient flow equations with non-homogeneous diffusion, *Comput. Math. Appl.*, 89, 150–162 (2021). The specific contributions of the author to this thesis include conceptualization, methodology development, supervision and validation.

$\mathbf{v} = -\boldsymbol{\nabla}\left(f'(\rho) + V + K \star \rho\right)$.

The structure of Eq. (4.73) is similar to the gradient-flow equations considered in previous works [209, 223, 224, 225, 226, 227, 228]. However in this work the diffusion parameter $\boldsymbol{\theta}_0^{-1}$ is considered, in general, non-constant in space and time. Here we propose a finite-volume scheme to solve gradient flow equations with non-homogeneous diffusion of the form Eq. (4.73), obtained in the framework of DDFT. The scheme allows us to study systems in non-homogeneous thermal baths and with a general free-energy functional accounting for external fields and inter-particle potentials. Specifically, we adopt a first- and second-order upwind approach for the space discretization in both one- and two-dimensional cases. Moreover, we prove that our scheme satisfies the discrete analogy of the free-energy decay in Eq. (4.74), thanks to the upwind formulation. We exemplify our scheme through several case studies of systems with underlying thermal baths characterized by anisotropic, space-dependent and time-dependent properties.

4.2.1. Numerical methods

The 1D version of Eq. (4.73) can be written as

$$\partial_t \rho = \partial_x \left(\boldsymbol{\theta}_0^{-1} \rho \partial_x \left(f'(\rho) + V + K \star \rho \right) \right), \tag{4.75}$$

where $\boldsymbol{\theta}_0^{-1}$ is a positive function that in general depends on time t, space x and density ρ. The finite-volume formulation of Eq. (4.75) is obtained by dividing the computational domain $[0, L]$ in N cells $C_i = [x_{j-1/2}, x_{j+1/2}]$, all with uniform size $\Delta x = L/N$, so that the centre of the cell satisfies $x_j = (j-1)\Delta x/L + \Delta x/2$, $j \in \{1, \ldots, N\}$. In each of the cells we define the density cell average ρ_j as $\rho_j(t) = \frac{1}{\Delta x}\int_{C_j} \rho(x,t)dx$. Thus, after integrating Eq. (4.75) over a cell C_j, one obtains the usual finite-volume discretization

$$\frac{d\rho_j}{dt} = -\frac{1}{\Delta x}(F_{j+1/2} - F_{j-1/2}), \tag{4.76}$$

where $F_{i\pm1/2}$ denote the numerical fluxes. Their evaluation follows an upwind approach similar to the one proposed by Carrillo and coworkers [209], i.e.

$$F_{j+1/2} = u^+_{j+1/2}\,\rho_j + u^-_{j+1/2}\,\rho_{j+1}, \tag{4.77}$$

where the velocity $u_{i+1/2}$ satisfies

$$u_{j+1/2} = -\left(\boldsymbol{\theta}_0^{-1}\right)_{j+1/2} \frac{\xi_{j+1} - \xi_j}{\Delta x}, \tag{4.78}$$

and

$$u^{+}_{j+1/2} = \max(u_{j+1/2}, 0), \quad u^{-}_{j+1/2} = \min(u_{j+1/2}, 0). \tag{4.79}$$

If $\boldsymbol{\theta}_0^{-1}$ depends only on x or t, than $\left(\boldsymbol{\theta}_0^{-1}\right)_{j+1/2}$ is evaluated explicitly at the interface $x_{j+1/2}$. Otherwise, if it depends on ρ, than it is approximated as $\left(\boldsymbol{\theta}_0^{-1}\right)_{j+1/2} = \left(\boldsymbol{\theta}_0^{-1}(\rho_j) + \boldsymbol{\theta}_0^{-1}(\rho_{j+1})\right)/2$.

The profile in each cell C_j is reconstructed as a constant for the first-order scheme, or as a linear function for the second-order scheme,

$$\widetilde{\rho}_j(x) = \begin{cases} \rho_j, \quad x \in C_i, & \text{for the first-order scheme,} \\ \rho_j + (\rho_x)_j (x - x_j), \quad x \in C_j, & \text{for the second-order scheme,} \end{cases} \tag{4.80}$$

so that the east and the west values ρ_j^E and ρ_j^W at the cell interfaces $x_{j+1/2}$ and $x_{j-1/2}$, respectively, are approximated as

$$\begin{cases} \rho_j^E = \rho_j, \quad \rho_j^W = \rho_j & \text{for the first-order scheme,} \\ \rho_j^E = \rho_j + \frac{\Delta x}{2}(\rho_x)_j, \quad \rho_j^W = \rho_j - \frac{\Delta x}{2}(\rho_x)_j & \text{for the second-order scheme.} \end{cases} \tag{4.81}$$

The numerical derivatives $(\rho_x)_j$ at every cell C_j are computed by means of an adaptive procedure which ensures that the point values Eq. (4.81) are second-order and non-negative, i.e.

$$(\rho_x)_j = \text{minmod}\left(\theta \frac{\rho_{j+1} - \rho_j}{\Delta x}, \frac{\rho_{j+1} - \rho_{j-1}}{2\Delta x}, \theta \frac{\rho_j - \rho_{j-1}}{\Delta x}\right),$$

where

$$\text{minmod}(z_1, z_2, \ldots) = \begin{cases} \min(z_1, z_2, \ldots), & \text{if } z_j > 0 \quad \forall j, \\ \max(z_1, z_2, \ldots), & \text{if } z_j < 0 \quad \forall j, \\ 0, & \text{otherwise.} \end{cases}$$

The parameter θ controls the numerical viscosity and it is taken to be $\theta = 2$. The no-flux conditions are numerically implemented by taking the numerical flux Eq. (4.77) to be zero at the boundaries,

$$\begin{aligned} F^{n+1}_{j-1/2} &= 0 \text{ for } j = 1, \\ F^{n+1}_{j+1/2} &= 0 \text{ for } j = N, \end{aligned} \tag{4.82}$$

As discussed in details in Ref [229], the above scheme satisfies density positivity and

decay of the discrete free-energy defined as

$$E_\Delta(\rho) = \Delta x \sum_j \left(V_j \rho_j + H(\rho_j) + \frac{1}{2} \Delta x \sum_i K_{j-i} \rho_i \rho_j \right). \tag{4.83}$$

Similarly to 1D, the 2D version of the scheme is discussed in Appedix C.2.1 In the next sections, we present various numerical experiments for different test cases. The aim is to show the impact of a variable coefficient $\boldsymbol{\theta}_0^{-1}$ on the dynamical evolution of the solutions, as opposed to a constant $\boldsymbol{\theta}_0^{-1}$. The boundary conditions in all simulations are taken as no-flux and the temporal integrator scheme is the third order TDV Runge-Kutta method [230].

4.2.2. 1D case studies

1D case studies: Order of convergence

First, we numerically test the convergence of the first- and second-order schemes by studying a system satisfying the following governing equation:

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x} \left(\boldsymbol{\theta}_0^{-1} \rho \frac{\partial}{\partial x} \left(\log(\rho) + V \right) \right), \tag{4.84}$$

where $V(x) = \frac{x^2}{2}$. Here we test the convergence both for a constant $\boldsymbol{\theta}_0^{-1} = 1$ and for a non-homogeneous diffusion with a spatial dependency of the form $\boldsymbol{\theta}_0^{-1}(x) = \chi_{x<0} + 0.5$. The computational domain is $[-5, 5]$ and we choose a Gaussian initial condition $\rho_0(x) = \frac{M}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{x^2}{2\sigma^2}\right)$ with $M = 0.1$ and $\sigma = 0.3$.

The order-of-convergence analysis is performed by running the numerical simulation for different number of cells, each of them being twice the one of the previous simulation. Hence, the cell size Δx is halved after each run. We then compute the L^1 error of the numerical solution $\rho_{\Delta x}$ as $\epsilon(\Delta x) = \|\rho_{\Delta x}(x, T) - \rho^*(x, T)\|_{L^1}$, where ρ^* is an approximate exact solution computed using a rather refined mesh (with $12,800$ cells). Finally the order of convergence is obtained from the error as:

$$o(\Delta x) = \log_2 \left(\frac{\epsilon(\Delta x)}{\epsilon\left(\frac{\Delta x}{2}\right)} \right). \tag{4.85}$$

Our results for the homogeneous diffusion with $\boldsymbol{\theta}_0^{-1} = 1$ are depicted in Table 4.1, while the ones for $\boldsymbol{\theta}_0^{-1}(x) = \chi_{x<0} + 0.5$ are shown in Table 4.2. They clearly indicate the first- and second-order convergence for the proposed numerical schemes.

Table 4.1. Accuracy test for $\boldsymbol{\theta}_0^{-1} = 1$ with the first and second-order schemes

Number of cells	First-order		Second-order	
	L^1 error	order	L^1 error	order
25	1.4135E-02	-	3.9482E-04	-
50	7.8529E-03	0.85	2.8628E-04	0.46
100	4.1668E-03	0.91	7.5858E-05	1.92
200	2.1453E-03	0.96	2.1434E-05	1.82

Table 4.2. Accuracy test for $\boldsymbol{\theta}_0^{-1}(x) = \chi_{x<0} + 0.5$ with the first and second-order schemes

Number of cells	First-order		Second-order	
	L^1 error	order	L^1 error	order
25	1.47581E-04	-	1.39116E-04	-
50	7.98831E-03	0.89	1.24210E-04	0.16
100	4.25173E-03	0.91	3.61261E-05	1.78
200	2.18779E-03	0.96	1.23742E-05	1.55

Nonlinear diffusion equation in a radially symmetric inhomogeneous medium

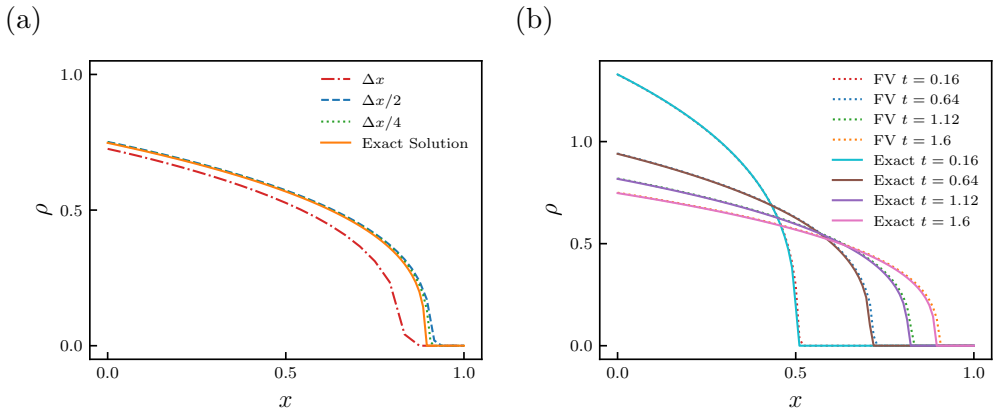

Figure 4.12. (a) Convergence towards the exact solution for Eq. (4.87) at the final computational time $t = 1.6$sec; (b) Comparison between the finite-volume (FV) scheme and the exact solution. Parameter values: $M_\infty = 0.5$, $a = 1$, $b = 3$, $D_0 = 1$, $r_0 = 0.1$, $\Delta x = 0.042$ in (a) and $\Delta x = 0.01$ in (b).

In this numerical application we simulate a system in a radially symmetric inhomogeneous medium [231]. Consider the following ν-dimensional nonlinear diffusion equation

in polar coordinates:

$$\begin{cases} \dfrac{\partial \rho}{\partial t} = \dfrac{1}{r^{\nu-1}} \dfrac{\partial}{\partial r}(r^{\nu-1} D \dfrac{\partial \rho}{\partial r}), \\ \int_0^{r_0} \omega_\nu r^{\nu-1} \rho(r,0) dr = M_\infty, \\ J_\infty(0,t) = 0, \quad t > 0, \\ \rho(r,t) \to 0, \quad r \to \infty, \quad t > 0, \end{cases} \tag{4.86}$$

where r is the radius $r = \|\mathbf{x}\|$, $D(r,\rho) = D_0 r^a \rho^b$ is the diffusion coefficient ($0 < a < 2$, $b > 0$), $J_\infty(r,t) = -\omega_\nu D r^{\nu-1} \frac{\partial \rho}{\partial r}$ is the diffusive flux. This problem describes the evolution of an initial distribution of a diffusing substance. In [231], an analytical solution for this system is derived,

$$\rho(r,t) = \begin{cases} C_0(t) \left[1 - \left(\frac{r}{r_f(t)} \right)^{2-a} \right]^{\frac{1}{b}} & \text{if} \quad r < r_f(t), \\ 0 & \text{otherwise}, \end{cases} \tag{4.87}$$

where $C_0(t)$ is the solution at the origin $r = 0$ and $r_f(t)$ denotes the front position.

In the 1D scenario we have that $r = |x|$, $\nu = 1$ and $\omega_\nu = 1$. We can then re-formulate the problem as:

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial r} \left(D_0 r^a \rho \frac{\partial}{\partial r} \left(\frac{\rho^b}{b} \right) \right). \tag{4.88}$$

Eq. (4.88) is a particular case of Eq. (4.75) with $\boldsymbol{\theta}_0^{-1}(r) = D_0 r^a$, $f'(\rho) = \dfrac{\rho^b}{b}$, $V = 0$ and $K = 0$. The conditions in Eq. (4.86) can be considered as Dirichlet ($\rho(r,t) \to 0$ when $r \to \infty$) and Neumann ($J_\infty(0,t) = 0$). The mass conservation is ensured due to the nature of the finite-volume schemes.

The results for this example are displayed in Fig. 4.12. Figure 4.12(a) depicts the convergence towards the exact solution Eq. (4.87). The discontinuous derivative at the front prevents the numerical solution for larger Δx to approach the exact solution. However, as the mesh is refined the numerical solution converges to the exact one. On the other hand, Fig. 4.12(b) plots the numerical solution at different times comparing to the exact solution Eq. (4.87). Evidently the radial diffusion enlarges the compact support of the solution in time making the front advance.

System with a space-dependent diffusion coefficient in an external field

In this numerical application, we compare the time-evolution of a system embedded with constant and space-dependent $\boldsymbol{\theta}_0^{-1}$. The governing equation for the analysed system

satisfies:

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\boldsymbol{\theta}_0^{-1}\rho\frac{\partial}{\partial x}\left(\nu\rho^{m-1} + V(x)\right)\right), \tag{4.89}$$

with an external potential satisfying $V(x) = \frac{x^4}{4} - \frac{x^2}{2}$ and $m = 3$. Simulations are performed on a domain with boundaries $[-3, 3]$, $\Delta x = 0.06$ and $\nu = 1.48$. Moreover we employ $\boldsymbol{\theta}_0^{-1} = 1$ in a first system and $\boldsymbol{\theta}_0^{-1}(x) = \chi_{x<0} + 0.5$ in a second one. The first case corresponds to the one reported also in [209]. Initial conditions for the system density satisfy $\rho_0(x) = \frac{M}{\sqrt{2\pi\sigma^2}} e^{-\frac{x^2}{2\sigma^2}}$ with $M = 0.1$ and $\sigma^2 = 0.2$.

The time-evolution of the density for systems with constant and space-dependent $\boldsymbol{\theta}_0^{-1}$ are reported in Fig. 4.13. First, Fig. 4.13(a) shows that for a constant $\boldsymbol{\theta}_0 v^{-1}$ the density evolution in time remains spatially symmetric, and, due to the shape of the external potential $V(x)$, the steady state is characterized by two compactly-supported peaks. This is in agreement with the steady-state solution of Eq. (4.89) for $\boldsymbol{\theta}_0^{-1} = 1$, satisfying [209]

$$\rho_\infty(x) = \left(\frac{C(x) - V(x)}{\nu}\right)_+^{\frac{1}{m-1}},$$

where $C(x)$ is a piecewise constant.

Figure 4.13(b) shows the time evolution of the density for $\boldsymbol{\theta}_0^{-1}(x) = \chi_{x<0} + 0.5$. An asymmetric solution characterize the temporal evolution of the density. However the steady state eventually approaches the same solution of the case with $\boldsymbol{\theta}_0^{-1} = 1$. It is worth noticing that the evolution of the density in the left region of the domain is faster due to the higher value of the local diffusion $\boldsymbol{\theta}_0^{-1}(x)$.

Finally, in Fig. 4.13(c) we report the discrete free-energy in Eq. (C.25). It monotonically decreases in time for both choices of $\boldsymbol{\theta}_0^{-1}$. Interestingly the space-dependent $\boldsymbol{\theta}_0^{-1}$ has a faster time convergence towards the steady state value compared to the constant $\boldsymbol{\theta}_0^{-1}$.

Nonlinear diffusion with non local attraction kernel

In this section we analyse systems with constant and space-dependent $\boldsymbol{\theta}_0^{-1}$, as in the previous example. However here we include non-local interactions, i.e.

$$\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\boldsymbol{\theta}_0^{-1}\rho\frac{\partial}{\partial x}\left(\nu\rho^{m-1} + K(x) \star \rho\right)\right). \tag{4.90}$$

This models is characterized by the competition between non-uniform diffusion and non-local attraction. As in the previous section, a first system has $\boldsymbol{\theta}_0^{-1} = 1$ (as in Ref [209].), while a second one has $\boldsymbol{\theta}_0^{-1}(x) = \chi_{x<0} + 1$. Density initial conditions satisfy

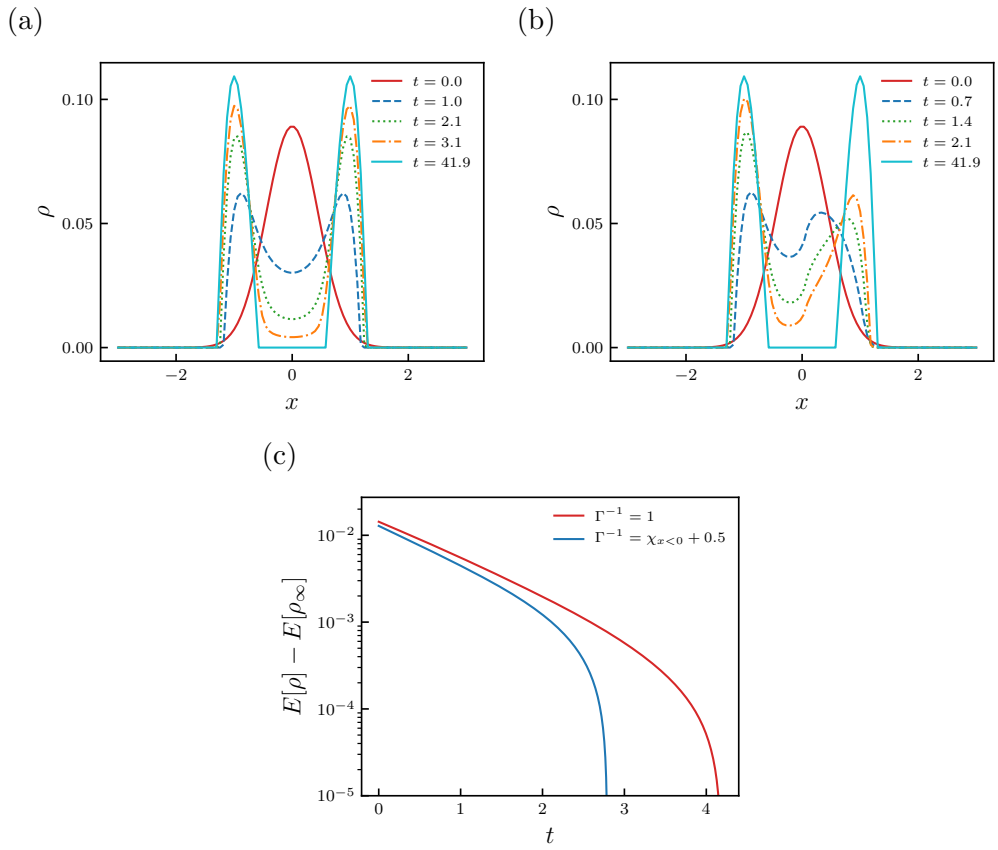

Figure 4.13. Numerical solution with $\boldsymbol{\theta}_0^{-1} = 1$ (a) and $\boldsymbol{\theta}_0^{-1}(x) = \chi_{x<0} + 0.5$ (b). Decay of the discrete free-energy in (c).

$\rho_0(x) = \chi_{[-1.5,1.5]}/4$.

The time evolution of the density for these examples are displayed in Fig. 4.14. Figure 4.14(a) shows results for a constant $\boldsymbol{\theta}_0^{-1}$. It dan be observes that initially two identical bumps are formed but they eventually merge into a single one when reaching the steady state. in this case the solution remains symmetric for all times. Figure 4.14(b) shows the time evolution for the space-dependent $\boldsymbol{\theta}_0^{-1}$. We observe that the dynamics are similar to Fig. 4.14(a), but here the two bumps are no longer symmetric. Specifically the density evolution in the left region of the domain occurs faster due to the larger value of the diffusion coefficient. Eventually the solution recovers the steady state of 4.14(a).

In Fig. 4.14(c) we report the discrete free-energy decay (Eq. (C.25)). The solution of in case of the space-dependent $\boldsymbol{\theta}_0^{-1}$ converges faster to the steady-state. Moreover, for $\boldsymbol{\theta}_0^{-1} = 1$ two characteristic intermediate asymptotic behaviours are observed corresponding to two straight lines with different slopes: the first one corresponds to the merging

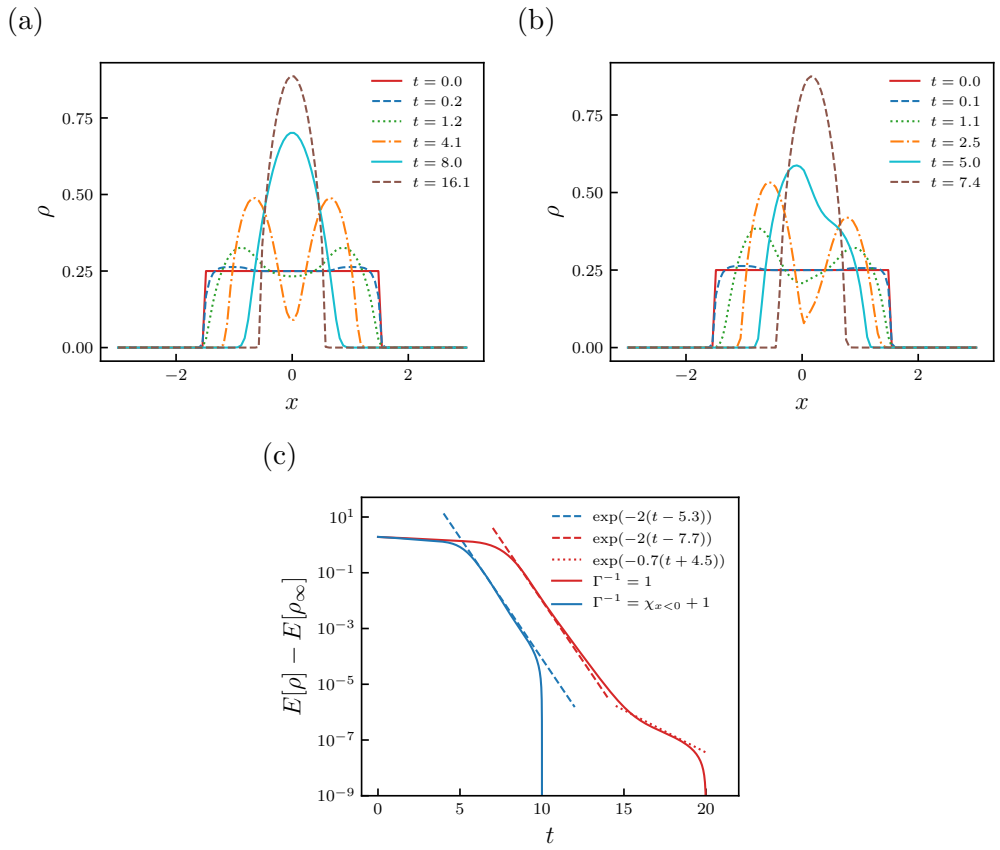

Figure 4.14. (a) Numerical Solution with $\boldsymbol{\theta}_0^{-1} = 1$, (b) Numerical Solution with $\boldsymbol{\theta}_0^{-1}(x) = 1 + \chi_{x<0}$, (c) Decay of the discrete free-energy. Parameter values: $\nu = 0.5$, $m = 3$, $K(x) = -(1 - |x|)\chi_{|x|<1}$,$\Delta x = 0.08$.

of the two bumps, while the second one to the formation of a single mass cluster. In the case of space-dependent $\boldsymbol{\theta}_0^{-1}$, the free-energy has a single exponential decays in time, and the overall evolution is faster.

Time-dependent diffusion coefficient $\boldsymbol{\theta}_0^{-1}(t)$

Here we study the heat equation $\frac{\partial \rho}{\partial t} = \frac{\partial}{\partial x}\left(\boldsymbol{\theta}_0^{-1}\frac{\partial}{\partial x}(\rho)\right)$, embedded with the following time-dependent diffusion coefficient $\boldsymbol{\theta}_0^{-1}$:

$$\boldsymbol{\theta}_0^{-1}(t) = \begin{cases} 1 & \text{if} \quad \cos(2\pi t) > 0.5, \\ 0 & \text{otherwise.} \end{cases}$$

The initial density for the simulations is $\rho_0(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, the spatial domain is $[-3, 3]$ and $\Delta x = 0.06$.

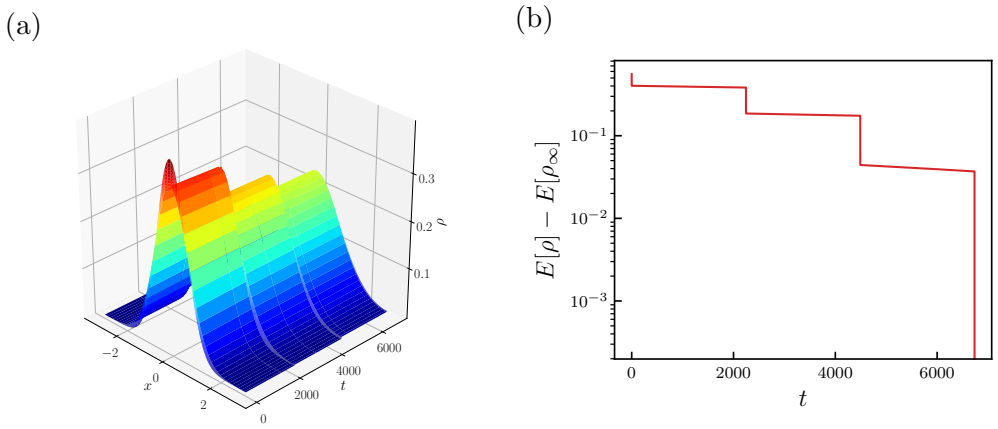

Figure 4.15. Heat equation with a time-dependent diffusion $\boldsymbol{\theta}_0^{-1}$. (a) Three-dimensional plot of the numerical solution and (b) Ddcay of the discrete free-energy.

The time-evolution of the density profile is shown in Fig. 4.15. This is characterized by a discontinuous time evolution which can be analysed as follows. Let us define I as a union of intervals I_p such that $\forall t \in I : \boldsymbol{\theta}_0^{-1}(t) = 1$. The density evolution varies according to the time interval. For $t \in I_p =: [t_{2p}, t_{2p+1}]$, the time-evolution of the density follows the usual diffusion evolution. For $t \notin I$, we have a stationary problem: $\frac{\partial \rho}{\partial t} = 0 \quad \forall t \in [t_{2p+1}, t_{2p+2}]$, and thus $\rho(x,t) = \rho(x, t_{2p+1}) \quad \forall t \in [t_{2p+1}, t_{2p+2}]$. Fig. 4.15(b) shows that this discontinuous time evolution is observed also in the free-energy decay.

4.2.3. 2D case studies

$\boldsymbol{\theta}_0{}^{-1}$ constant matrix

In 2D applications, $\boldsymbol{\theta}_0^{-1}$ in Eq. (4.73) becomes a 2×2 matrix. Let us consider the following anisotropic diffusion problem embedded with a constant $\boldsymbol{\theta}_0^{-1}$ matrix:

$$\frac{\partial \rho}{\partial t} = \boldsymbol{\nabla} \cdot (\boldsymbol{\theta}_0^{-1} \boldsymbol{\nabla} \rho) \quad on \quad \Omega = [-A, A]^2. \tag{4.91}$$

An analytical solution of Eq. (4.91) can be obtained by employing 2D Fourier series. Specifically if the time-evolution of the system density is periodic in space at all times,

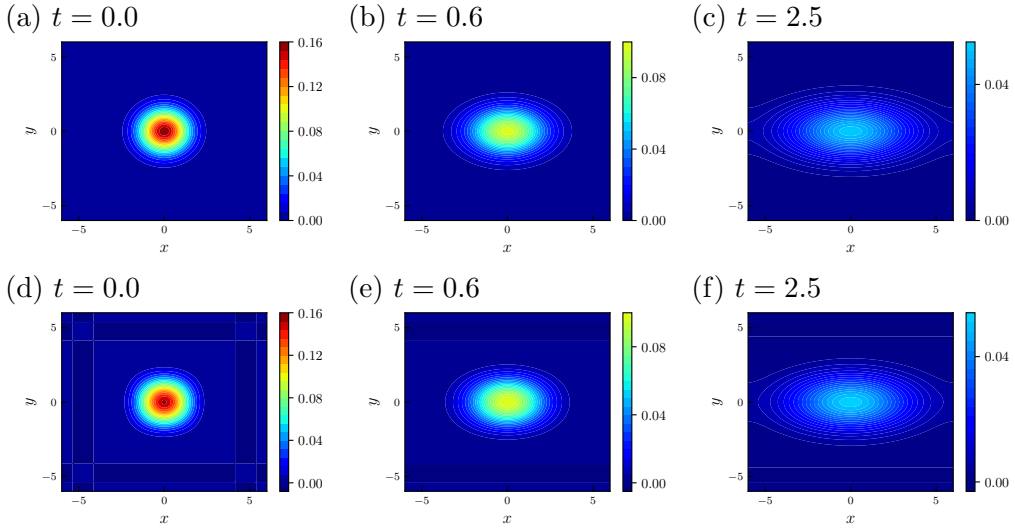

Figure 4.16. Diffusion in $(1, 0)$ direction. (a) Numerical solution, (b) approximate (Fourier series) solution.

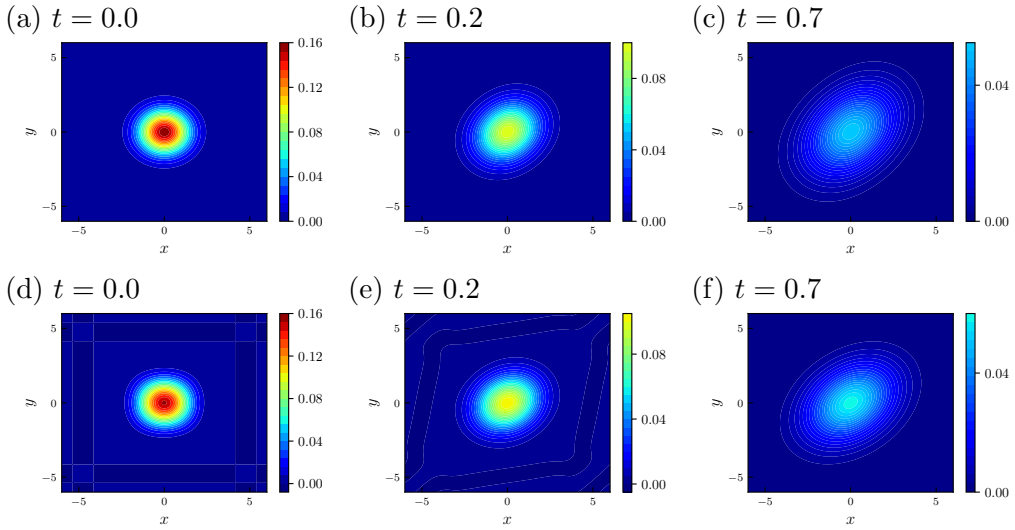

Figure 4.17. Diffusion in (1,1) direction. (a) Numerical solution, (b) approximate (Fourier series) solution.

than its density Fourier expansion can be written as:

$$\rho(x, y, t) = \sum_{k,l} C_{k,l}(t) e^{i\frac{\pi}{A}(kx+ly)}. \tag{4.92}$$

where the Fourier coefficients satisfy:

$$\frac{dC_{k,l}}{dt}(t) = -\left(\frac{\pi}{A}\right)^2 \alpha_{k,l} C_{k,l}(t), \tag{4.93}$$

with $\alpha_{k,l} = ((\boldsymbol{\theta}_0^{-1})^T(k,l)^T).((k,l)^T)$. It follow that $C_{k,l}(t) = C_{k,l}(0)\exp\left(-\left(\frac{\pi}{A}\right)^2 \alpha_{k,l} t\right)$, with the coefficients $C_{k,l}(0)$ being the Fourier coefficients of the initial conditions. In this case, we can approximate ρ with a truncated Fourier series:

$$\rho(x,y,t) \approx \sum_{|k|+|l|\leq N} C_{k,l}(t) e^{i\frac{\pi}{A}(kx+ly)}. \tag{4.94}$$

Here our objective is to compare the numerical simulation of Eq. (4.91) against the approximated exact solution obtained with the truncated Fourier series. In order to show the effect of $\boldsymbol{\theta}_0$ on the diffusion directions, we consider systems characterized by the following two $\boldsymbol{\theta}_0$ matrices:

$$\boldsymbol{\theta}_0^{(1)} = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\theta}_0^{(2)} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}. \tag{4.95}$$

If $\boldsymbol{\theta}_0$ is diagonal, the main diffusion direction corresponds to the highest coefficient of $\boldsymbol{\theta}_0^{-1}$. However, in general the diffusion direction is given by the eigenvector associated with the highest eigenvalue. If the two eigenvalues are equal, then heat is diffused along both eigenvectors directions with the same speed.

The initial density field is a normalized Gaussian distribution, and the numerical domain is $[6 \times 6]$ with $\Delta x = \Delta y = 0.2$. Simulations result for $\boldsymbol{\theta}_0^{(1)}$ and $\boldsymbol{\theta}_0^{(2)}$ are reported in Fig. 4.16 and Fig. 4.17, respectively. As expected, in Fig. 4.16 the main diffusion direction is parallel to the horizontal axis, while in Fig. 4.17 it is diagonal and aligned with the direction of the highest eigenvalue. Finally, the numerical and Fourier time-evolution of the density are in good agreement.

Space-dependent diffusion

Here we simulate the non-homogenous diffusion problem $\boldsymbol{\theta}_0^{-1}$

$$\frac{\partial \rho}{\partial t} = \boldsymbol{\nabla} \cdot (\boldsymbol{\theta}_0^{-1} \boldsymbol{\nabla} \rho). \tag{4.96}$$

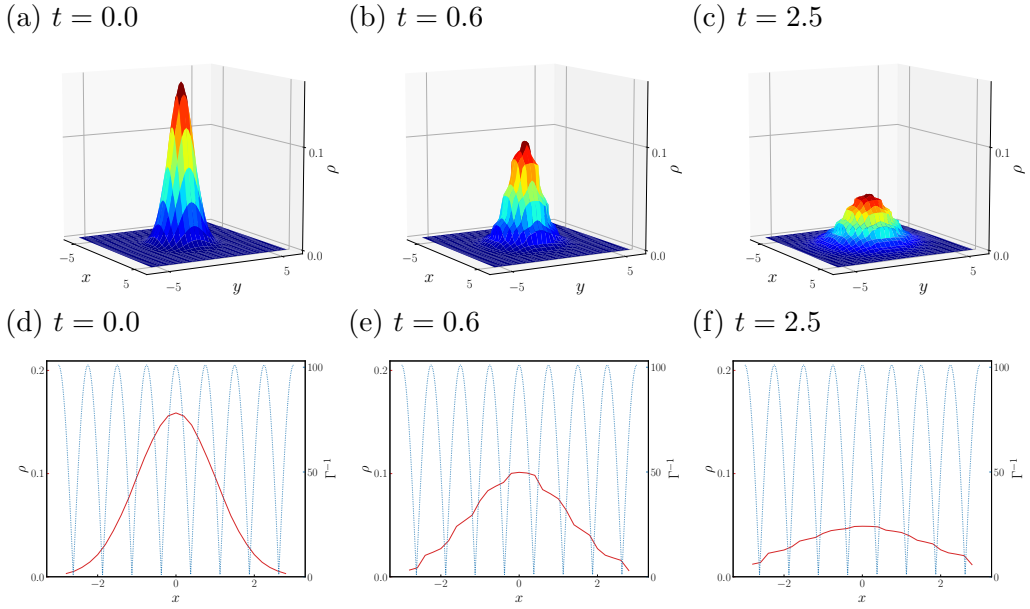

Figure 4.18. (a)-(c) 3D plots of the numerical solution and (d)-(f) slices of the numerical solution for $y = 0$. Parameter values: $\boldsymbol{\theta}_0^{-1}(x, y) = 1 + 100|\cos\left(\frac{2\pi}{T}\sqrt{x^2 + y^2}\right)|$, $(T = 1.5)$, $\Delta x = \Delta y = 0.2$.

embedded with the following space-dependent $\boldsymbol{\theta}_0^{-1}$:

$$\boldsymbol{\theta}_0^{-1}(x, y) \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1 + 100 \left|\cos\left(\frac{2\pi}{T}\sqrt{x^2 + y^2}\right)\right| \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{4.97}$$

which varies sinusoidally in space. The initial density is a normalized Gaussian distribution, while the numerical domain is $[6 \times 6]$ with $\Delta x = \Delta y = 0.2$.

In Figs 4.18(a-b-c) we plot the 2D numerical solution, while in Figs 4.18(d-e-f) we depict specific slices of the numerical solution as well as of $\boldsymbol{\theta}_0^{-1}$ at $y = 0$. Figures 4.18(e-f) shows that the slopes of the solution vary following the space-dependency of $\boldsymbol{\theta}_0^{-1}$. Eventually ρ spreads out due to the diffusion.

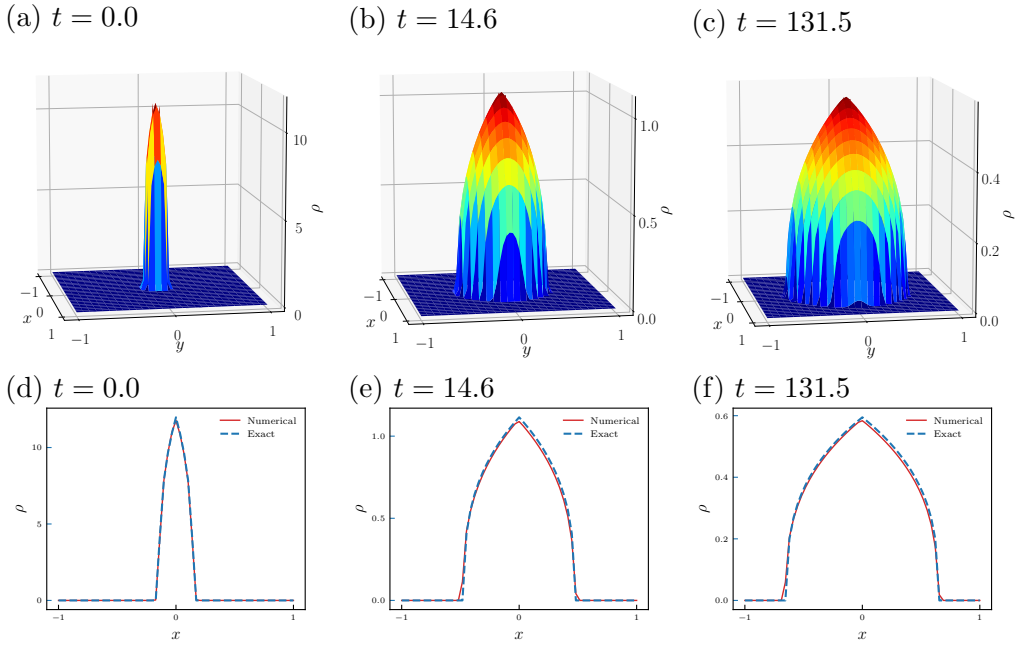

Figure 4.19. (a): 3-D plot of the numerical solution, (b): Slices of numerical solution and exact solution for $y = 0$. Parameter values: $a = 1$, $b = 3$, $D_0 = 0.01$, $\Delta x = \Delta y = 0.08$, $r_0 = 0.14$, $M_\infty = 0.5$.

Nonlinear diffusion equation in a radially-symmetric inhomogeneous medium in 2D

Here we analyse the 2D extension of the 1D problem in Eq. (4.86). Here we choose:

$$\boldsymbol{\theta}_0^{-1}(x, y) = D_0 r^a \rho^b \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \tag{4.98}$$

and we select the rest of parameters as $\nu = 2$, $r = \sqrt{x^2 + y^2}$ and $w_\nu = 2\pi$. The initial condition for the simulations is $\rho_0(r) = C_0(0)\left(1 - (r/r_0)^{2-a}\right)^{\frac{1}{b}}$ where $a = 1$, $b = 3$, $r_0 = 0.14$, and $C_0(0)$ satisfies the conservation of mass, with $M_\infty = 0.5$. The computational domain is $[-1, 1]^2$ with $\Delta x = \Delta y = 0.08$.

In Figs 4.19(a-b-c) we plot the numerical solution obtained from our finite-volume scheme. We can observe how the compactly-supported density spreads out in time due to diffusion. In Figs 4.19(d-e-f) we compare slices at $y = 0$ of the numerical and the exact solutions Eq. (4.87) derived in Ref. [231]. Evidently the two methods are in good agreement at all times.

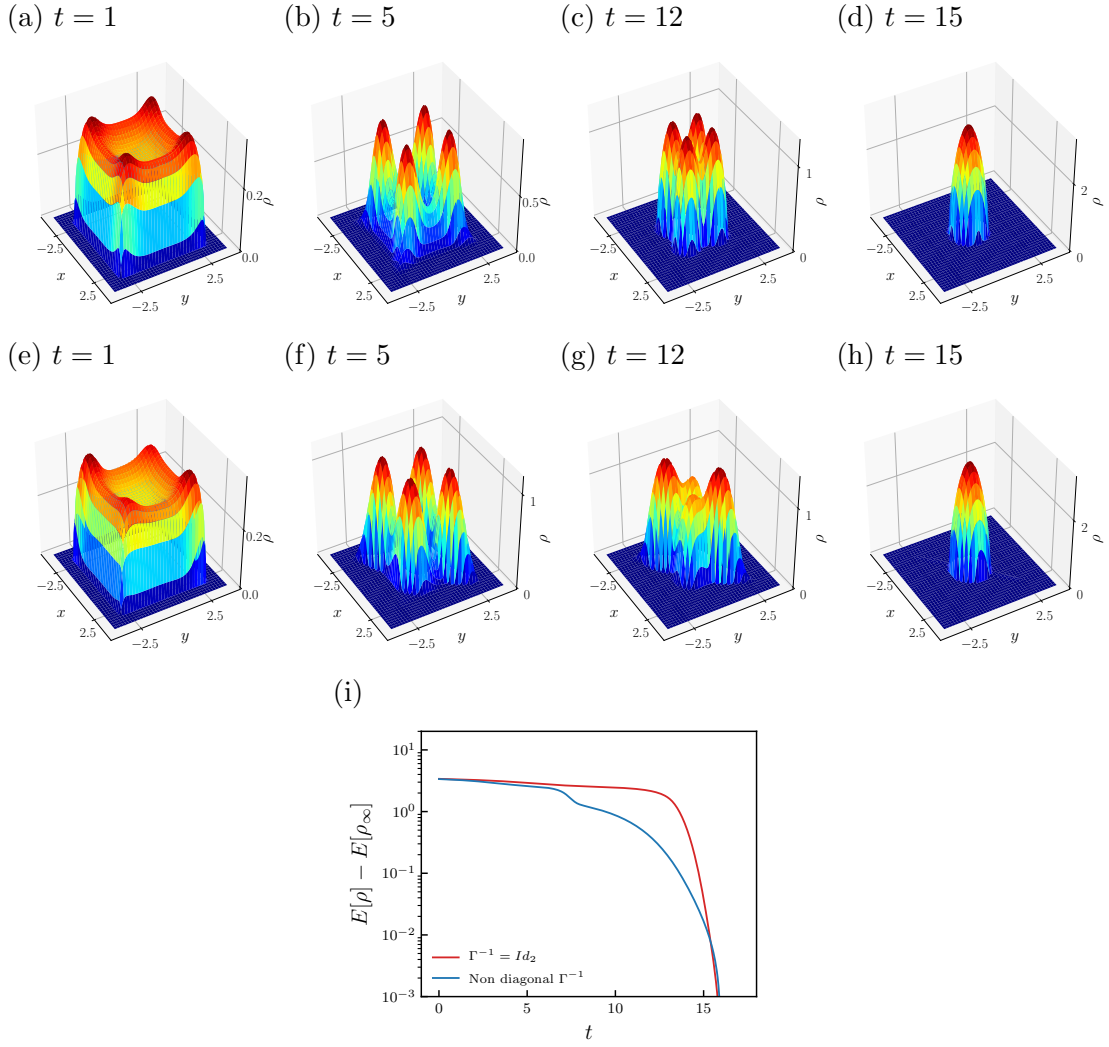

Figure 4.20. (a)-(d): 3-D plot of the numerical solution with $\boldsymbol{\theta}_0^{(1)}$, (e)-(h): 3-D plot of the numerical solution with $\boldsymbol{\theta}_0^{(2)}$, (i): Decay of the discrete free-energy. Parameter values: $m = 3$, $\Delta x = \Delta y = 0.05$, $\rho_0(x, y) = \frac{1}{4}\chi_{[-3,3]^2}(x, y)$.

Nonlinear diffusion with nonlocal attraction in 2D

In this final example we consider the gradient flow in Eq. (4.73) with a non zero attraction kernel $K(x, y) = -\frac{1}{\pi}e^{-x^2-y^2}$:

$$\frac{\partial \rho}{\partial t} = \boldsymbol{\nabla} \cdot \left(\boldsymbol{\theta}_0^{-1}\rho \boldsymbol{\nabla} \cdot \left(\nu\rho^{m-1} + K(x, y) \star \rho\right)\right). \tag{4.99}$$

The aim of this example is to compare the dynamical evolution with two different $\boldsymbol{\theta}_0^{-1}$ matrices:

$$\boldsymbol{\theta}_0^{(1)} = \begin{bmatrix} 10 & 0 \\ 0 & 1 \end{bmatrix}, \quad \boldsymbol{\theta}_0^{(2)} = \begin{bmatrix} 1 & 0.5 \\ 0.5 & 1 \end{bmatrix}.$$

The eigenvectors of the matrices above determine different directions of diffusion during the time-evolution of the density. The space domain taken in the simulations is $[-3,3]^2$ meshed with $\Delta x = \Delta y = 0.03$, and the initial conditions satisfy $\frac{1}{4}\chi_{[-3,3]^2}(x,y)$.

Figures 4.20(a-b-c-d) depict the evolution with $\boldsymbol{\theta}_0^{(1)}$. We observe a similar dynamics to the ones in the 1D example Fig. 4.2.2(a). In this case four identical bumps are formed, eventually merging into a single one. The non-local attraction (convolution) dominates the diffusion during the dynamical evolution, until the steady state is approached, and the dynamics is symmetric with respect to $(0,0)$.

Figures 4.20(e-h) show the results with $\boldsymbol{\theta}_0^{(2)}$. The steady state is identical to the case $\boldsymbol{\theta}_0^{(1)}$, however the merging of the peaks is accomplished at two different timescales. This is due to the different eigenvectors of $\boldsymbol{\theta}_0^{(1)}$ and $\boldsymbol{\theta}_0^{(2)}$. Indeed, the discrete free-energy plot in Fig. 4.20(i) shows that for $\boldsymbol{\theta}_0^{(1)}$ a plateau is reached before a drastic decrease of free-energy. For $\boldsymbol{\theta}_0^{(2)}$ in Figure 4.20(i) that there are two distinct phases during the decrease of the free-energy due to the two merging that take place during the time-evolution of the density.

Chapter 5

Conclusions and Outlook

In this final Chapter, we summarize this work and discuss future investigations on some of the analysed topics. Moreover, we include a dedicated section on additional research topics to be investigated in future projects as an extension of the present work.

5.1. Summary and conclusions

5.1.1. Molecular dynamics simulation of liquid-solid, liquid-vapor and liquid-liquid interfaces

In the present work, we used MD to obtain a microscopic description of confined LJ systems and extract macroscopic relations from simulations data. A detailed analysis of the density profile at the interfaces for several conditions of the system is provided. Furthermore, the non-homogeneous stress tensor and its kinetic an potential contributions are discussed. Hence, the shear viscosity profile of the system is investigated, and modelled as function of density, temperature and wall interactions via local and non-local models. Specifically, also by leveraging a theoretical framework proposed in Ref [25], we propose and test relations between viscosity and density variation at interfaces, which constitute a generalization of the purely data-driven model proposed by Morciano *et al.* [30]. Such relations are tested against MD simulations in a wide range of interfacial conditions and temperatures, showing a good agreement.

Our models can be employed in continuous numerical methods for mesoscopic fluid flows, such as DDFT, to improve their accuracy at fluid-solid interfaces. In future

works on the topic, the proposed relations should be further generalized to additional inter-particle potentials. Also, a detailed analysis of fluid properties for various wall geometries, shapes and roughness would be of great interest for the engineering community and is still missing.

5.1.2. Low dimensional models

GLE

In this work, we have propose a novel methodology to parametrize a GLE dynamics of an observable by means of a multi-layers perceptron. By using machine learning to eliminate the unknowns from GLEs, our methodology outperforms previous approaches in terms of efficiency and robustness. Differently from traditional approaches employing Laplace transform, we have shown that the presented methodology does not suffer random data fluctuations typically present in real system data-sets. The general applicability of our approach, guaranteed by the universal approximation theorem, makes its use appealing in a variety of applications. In fact, our methodology is tested against several prototypical examples, from colloidal systems to particle chains in a bath, to climatology and financial assets. In all cases, we report a good agreement between actual and modelled dynamics of the observables under consideration. Thus, coupling deep learning with the most general equation of statistical physics, namely GLE, opens the doors for a new way of modelling and understanding complex systems. Future developments of our method will involve MLPs equipped with complex exponential functions, since this may lead to enhanced approximations of oscillatory memory kernels.

Non-Markovian FDDFT

Here we have derived a theoretical framework for describing colloidal reacting multi-species fluid systems. From Newton's equations of motion, applying Mori-Zwanzig projection-operator techniques, one can obtain a system of GLEs, describing the motion of a relevant part of the original system only. Unlike most approaches, GLEs have a stochastic non-Markovian form, with a memory term, given in the form of a time-convolution. Defining the local number and momentum density fields, we obtain the corresponding non-Markovian Navier-Stokes like system of microscopic fluctuating hydrodynamics. Introducing the free-energy functional from classical density-functional theory, we obtain a non-Markovian fluctuating dynamical density functional theory for reacting multi-species systems, which is the main result of the present work. By introducing a proper deconvolution operator and taking the overdamped limit, we also

obtained the corresponding limiting case expression for the non-Markovian overdamped evolution of the density field. Additionally, we proposed a finite volume scheme to discretize the non-Markovian FDDFT, and we provided several numerical applications. These include ideal gas in and out of equilibrium, and a two-species reacting system exhibiting memory-driven Turing patterns.

The theoretical and numerical framework developed in this work can be adopted to investigate many physical phenomena in both active and passive matter. In fact, we showed that memory, by inducing time-correlations and variations of the fluid mobility that alter the transition path of the system, have a significant role in the dynamical evolution of systems undergoing state transitions. As a result, in general the treatment of the memory term requires a great attention. Thus, an interesting extension of the present work would include the usage of our framework to the study of phase transitions in biological fluids, where memory effects should play an important role. Such systems are often characterized by space- and time-scales not accessible with atomistic simulations [167, 168]. Another possible line of investigation could include the analysis of geometrical confinement on colloidal systems [55, 232, 233].

5.1.3. DDFT with additional complexities

In this work we also proposed and tested finite volume numerical schemes for solving gradient flow equations derived in the framework of FDDFT. We considered DDFT with two different complexity: (1) DDFT with the stochastic term and (2) DDFT with non-homogeneous diffusion.

Regarding stochastic DDFT, while previous works have developed numerical methodology only applicable to a limited range of free-energies (e.g. ideal gas free-energy in Refs [71, 221]), our proposed scheme effectively deals with general free-energy functionals, including external fields and inter-particle potential components. This contributes to a better understanding the effects of thermal fluctuations in relevant physical phenomena. Our numerical scheme relies on a hybrid space discretization based on central and upwind schemes. This approach provides an optimal compromise between statistical properties of the stochastic field and spurious oscillations. Moreover, an adaptive time-step based on Brownian bridge technique is employed to ensure the non-negativity of the density. We validated the proposed scheme by means of several numerical applications. First, we analysed variance, temporal and spatial correlations, and structure factor of ideal gas at equilibrium, comparing the results of our finite volume solver with both MD and theoretical results. Consistently with previous works, we showed that a minimum number of $5 - 10$ particle per cell is required in order for FDDFT to match atomistic

simulations results. Then, we examined the out of equilibrium evolution of a system in a double well potential. As expected in this case, our stochastic solver accurately reproduced local mean density, local density fluctuations and spatial correlations obtained with MD simulations. We also highlighted that, in this specific case, the deterministic the outcomes from DDFT simulations, which do not include thermal fluctuations, are consistent with the FDDFT and MD results. Finally, we simulated homogeneous nucleation kinetics of a fluid consisting of particles interacting through a Lennard-Jones like potential. The outcome of these simulations matches the theoretical results predicted by the phase diagram, and illustrate the crucial role of fluctuations to surmount free-energy barriers. Moreover, an Arrhenius-like law is observed for the nucleation growth rate as function of the metastable free-energy barrier.

It is worth highlighting that the approach employed here to preserve density positivity may not be applicable to simulations requiring the solution of both density and momentum fields. This could be employed, for instance, in research studies of bubble nucleation and collapse.

Regarding the numerical schemes for DDFT with non-homogeneous diffusion, we developed an accurate, efficient and robust finite-volume scheme to solve gradient-flow equations with non-homogeneous properties. Our numerical methodology extends the state-of-art in the finite-volume literature [209] by accounting for spatio-temporal dependence of the friction tensor $\boldsymbol{\theta}_0$. We propose 1D and 2D first- and second-order finite-volume schemes satisfying density positivity (under a certain CFL condition) as well as the discrete-energy decay. The order of the scheme is tested for both homogeneous and non-homogeneous friction tensors demonstrating first- and second-order spatial convergence. moreover, our approach is applied to simulate several prototypical systems with thermal baths characterized by anisotropic, space-dependent and time-dependent properties, both in 1D and 2D. The dynamical evolution of density profile and the decay of discrete free-energy allow us to scrutinise how the particular spatio-temporal dependency of $\boldsymbol{\theta}_0$ affects the dynamical evolution of the numerical solutions. We also performed several numerical experiments showing non-trivial dynamical evolution arising from the particular choice $\boldsymbol{\theta}_0$ which in turn strongly influences the characteristic timescales of the system and decay rates of the discrete free-energy.

5.2. Additional open directions and future research topics

In this section we discuss some topics to be investigated in future research studies to extend the work in this thesis.

5.2.1. Volume (or bulk) viscosity for inhomogeneous fluids

A general fluid flow can be described by using two transport coefficients, shear and volume (or bulk) viscosities. They are related to the irreversibility of a generic process acting on the fluid. The shear viscosity η accounts for shape deformations, while the bulk viscosity η_b is related to size deformations. Since in many physical systems, such as incompressible fluids or dilute monoatomic gases, changes in shape are predominant, the effects of bulk viscosity are often neglected. However, when dealing with phenomena involving fluids compressions or expansions, such as shockwaves, bulk viscosity plays a key role. Moreover, in confined fluids, bulk viscosity is in general non-uniform throughout the domain, thus a methodology to investigate its local value is required. A generic stress tensor can be expressed as:

$$\sigma = -p_{eq}\mathbf{I} + \left(\eta_b - \frac{2}{3}\eta_s\right)\boldsymbol{\nabla}\cdot\mathbf{u}\mathbf{I} + \eta_s\left(\boldsymbol{\nabla}\mathbf{u} + \boldsymbol{\nabla}\mathbf{u}^T\right) \tag{5.1}$$

where σ is the stress tensor out of equilibrium, $p_{eq} = -(\sigma_{eq,xx} + \sigma_{eq,yy} + \sigma_{eq,zz})/3$ is the pressure, η is the shear viscosity, $\mathbf{u}$ is the velocity and η_b is the bulk (or volume) viscosity, which equal to zero for dilute monatomic gases and strictly positive for the other cases. The bulk viscosity is an empirical coefficient depending on the nature of the fluid and is usually quite difficult to compute locally.

We propose to employ MD simulations in future studies to estimate local bulk viscosity. Specifically, Green-Kubo formulation provides a relation between bulk viscosity and pressure fluctuation $\delta p(t) = p(t) - \bar{p}$ for an NVE ensemble [234], namely

$$\eta_b = \frac{V}{k_B T}\int_0^\infty < \delta p(0)\delta p(t) > dt \tag{5.2}$$

where V is the volume, k_B is the Boltzmann's constant, T is the temperature, t is the time and $< \delta p(0)\delta p(t) >$ indicates the average autocorrelation of $\delta p(t)$. This expression has been widely used for hard-sphere, Lennard-Jones [235, 236, 237] and liquid water [238]. Non-equilibrium approaches to the computations of the bulk viscosity are difficult since, as Hoover *et al.* [237] state, "a steady-state measurement, possible in case of shear (viscosity), is impossible. At best the bulk viscosity can be measured a cyclic compression and expansion". However, cyclic compressions and expansions would be difficult to implement for systems confined between solid interfaces, like the ones we are interested in.

We performed some preliminay MD simulations on this analysis. A fluid is confined between two solid parallel walls. For each fluid particle i^{th}, the potential U_i is computed

as:

$$U(r_i)_{fluid} = \sum_{j=1}^{N_f} U_{LJ}(r_{ij})_{fluid-fluid} + \sum_{j=1}^{N_w} U_{LJ}(r_{ij})_{solid-fluid} \tag{5.3}$$

where N_f and N_w are the total number of respectively fluid and wall particles within a distance lower than its cutting radius r_c. Walls particles are maintained fixed in their lattice positions. The fluid is modelled as mono-atomic particles interacting by a 12-6 Lennard-Jones potential U_{LJ}. The simulation box (wall excluded) measures $[30\sigma] \times [30\sigma] \times [30\sigma]$. Periodic boundary conditions are imposed along x and y axes, while non-periodic,shrink-wrapped conditions contains the system in the z direction. The system is equilibrated with a a Nosé-Hoover thermostat to a temperature $Tk_b/\epsilon = 1.0$, and simulations are carried out in an NVE ensemble. The data are gathered by following the approach in [239]: the pressure autocorrelations to be averaged are computed from subset of a single trajectory, with interval between the time origins of $30n_{steps}$.

Figure 5.1 reports the average autocorrelation function and the bulk viscosity for the whole domain evaluated along the confined direction. It can be noticed that η_b tends to a constant value in the bulk fluid and increases close to the walls (as expected due to the density layering [239]). In this preliminary investigation, we compute the bulk viscosity inside each bin with a relation that is rigorously valid only for NVE ensembles. However, a relation for grand-canonical ensembles is also available and takes the following form [234]:

$$\eta_b = \frac{V}{k_B T} \int_0^\infty < \delta p(0) \delta p(t) > dt \tag{5.4}$$

where $\delta p(t) = p(t) - \bar{p}_{corr}$ and $\bar{p}_{corr}$ can be expressed as:

$$\bar{p}_{corr} = \bar{p} + \frac{\partial \bar{p}}{\partial E(t)} \left(E(t) - \bar{E} \right) + \frac{\partial \bar{p}}{\partial N(t)} \left(N(t) - \bar{N} \right) \tag{5.5}$$

with $E(t)$ and $\bar{E}$ being the instantaneous and average total energies respectively, and $N(t)$ and $\bar{N}$ being the instantaneous and average number of particle in the system respectively. Specifically, an equilibrium NVT system could be simulated, such that we can consider volume, temperature and chemical potential to be constant within each bin (for sufficiently large bins). Then, the Green-Kubo formulation for grand-canonical ensembles could be employed to accurately estimate the local bulk viscosity. It is worth mentioning however that one would need to simulate the system at different energy and density values to evaluate for each bin the two derivatives $\frac{\partial \bar{p}}{\partial E(t)}$ and $\frac{\partial \bar{p}}{\partial N(t)}$ in Eq. 5.5.

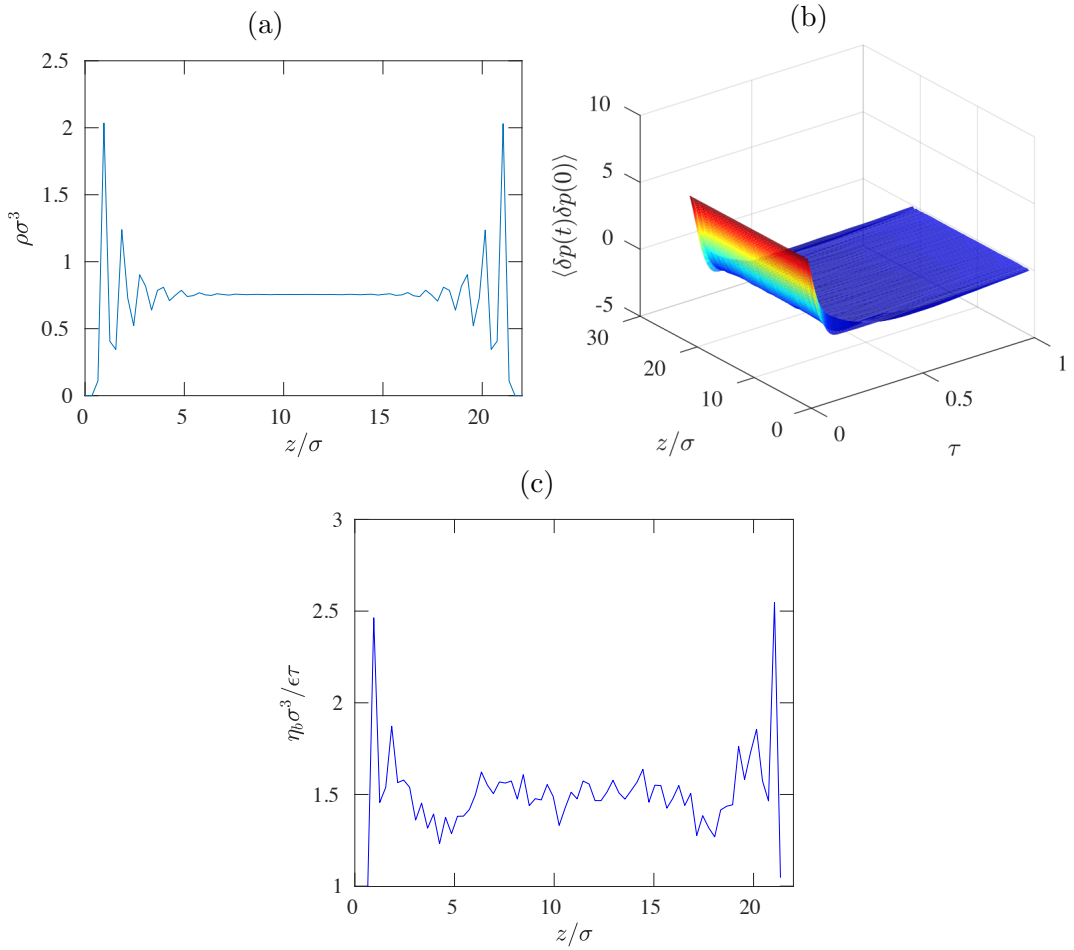

Figure 5.1. (a)Density profile, (b) average autocorrelation of the pressure fluctuation and (c) local bulk viscosity $\eta_b(y)$ evaluated in the simulated domain.

Bibliography

[1] M. A. Budroni, L. A. Riolfo, L. Lemaigre, F. Rossi, M. Rustici, and A. De Wit. Chemical control of hydrodynamic instabilities in partially miscible two-layer systems. *J. Phys. Chem. Lett.*, 5(5):875–881, 2014.

[2] Mohamed H. Nasr, Craig E. Green, Peter A. Kottke, Xuchen Zhang, Thomas E. Sarvey, Yogendra K. Joshi, Muhannad S. Bakir, and Andrei G. Fedorov. Hotspot thermal management with flow boiling of refrigerant in ultrasmall microgaps. *J. Electron. Packag.*, 139:8, 2017.

[3] Ran Peng and Dongqing Li. Electrokinetic motion of single nanoparticles in single pdms nanochannels. *Microfluid Nanofluidics*, 21(1):12, 2017.

[4] Jongho Lee, Tahar Laoui, and Rohit Karnik. Nanofluidic transport governed by the liquid/vapour interface. *Nat. Nanotechnol.*, 9:317–323, 2014.

[5] P. Malgaretti, M. N. Popescu, and S. Dietrich. Active colloids at fluid interfaces. *Soft Matter*, 12:4007–4023, 2016.

[6] Cleber L. Filomeno, Mansour Kouyaté, Véronique Peyre, Gilles Demouchy, Alex F. C. Campos, Régine Perzynski, Francisco A. Tourinho, and Emmanuelle Dubois. Tuning the solid/liquid interface in ionic colloidal dispersions: Influence on their structure and thermodiffusive properties. *J. Phys. Chem. C*, 121(10):5539–5550, 2017.

[7] Gunnar Dunér, Stephen Garoff, Todd M. Przybycien, and Robert D. Tilton. Transient marangoni transport of colloidal particles at the liquid/liquid interface caused

by surfactant convective-diffusion under radial flow. *J. Colloid Interface Sci.*, 462:75 – 87, 2016.

[8] Antonio Russo, Miguel A Durán-Olivencia, Ioannis G Kevrekidis, and Serafim Kalliadasis. Deep learning as closure for irreversible processes: A data-driven generalized langevin equation. *arXiv preprint arXiv:1903.09562*, 2019.

[9] Uday B. Kompella and Kavitha Koushik. Preparation of drug delivery systems using supercritical fluid technology. *Crit. Rev. Ther. Drug*, 18(2), 2001.

[10] Christoph Westerhausen, Lukas G. Schnitzler, Dominik Wendel, Rafal Krzysztoń, Ulrich Lächelt, Ernst Wagner, Joachim O. Rädler, and Achim Wixforth. Controllable acoustic mixing of fluids in microchannels for the fabrication of therapeutic nanoparticles. *Micromachines*, 7(9):150, 2016.

[11] Dongeun Huh, Hyun Jung Kim, Jacob P. Fraser, Daniel E. Shea, Mohammed Khan, Anthony Bahinski, Geraldine A. Hamilton, and Donald E. Ingber. Microfabrication of human organs-on-chips. *Nat. Protoc.*, 8:2135–2157, 2013.

[12] Chao Wang, Sung-Wook Nam, John M. Cotte, Christopher V. Jahnes, Evan G. Colgan, Robert L. Bruce, Markus Brink, Michael F. Lofaro, Jyotica V. Patel, Lynne M. Gignac, Eric A. Joseph, Satyavolu Papa Rao, Gustavo Stolovitzky, Stanislav Polonsky, and Qinghuang Lin. Wafer-scale integration of sacrificial nanofluidic chips for detecting and manipulating single dna molecules. *Nat. Commun.*, 8(14243):9, 2017.

[13] B. J. Alder and T. E. Wainwright. Phase transition for a hard sphere system. *J. Chem. Phys.*, 27(5):1208–1209, 1957.

[14] B. J. Alder and T. E. Wainwright. Studies in molecular dynamics. i. general method. *J. Chem. Phys.*, 31(2):459–466, 1959.

[15] A. Rahman. Correlations in the motion of atoms in liquid argon. *Phys. Rev.*, 136:A405–A411, Oct 1964.

[16] A. Rahman and F. H. Stillinger. Hydrogen-bond patterns in liquid water. *J. Am. Chem. Soc.*, 95(24):7943–7948, 1973.

[17] Frank H. Stillinger and Aneesur Rahman. Improved simulation of liquid water by molecular dynamics. *J. Chem. Phys.*, 60(4):1545–1557, 1974.

[18] J. Andrew McCammon, Bruce R. Gelin, and Martin Karplus. Dynamics of folded proteins. *Nature*, 267(5612):585–590, 1977.

[19] Remco Hartkamp and Stefan Luding. Lennard-jones fluids in a nanochannel. In Nuernberg Messe GmbH, editor, *7th International Conference on Multiphase Flow*, pages 1–7. Nuernberg Messe GmbH, 2010.

[20] Scott E. Feller. Molecular dynamics simulations of lipid bilayers. *Curr. Opin. Colloid Interface Sci.*, 5(3):217 – 223, 2000.

[21] Sabab Hasan Khan, Amresh Prakash, Preeti Pandey, Andrew M. Lynn, Asimul Islam, Md. Imtaiyaz Hassan, and Faizan Ahmad. Protein folding: Molecular dynamics simulations and in vitro studies for probing mechanism of urea- and guanidinium chloride-induced unfolding of horse cytochrome-c. *Int. J. Biol. Macromol.*, 122:695 – 704, 2019.

[22] Rafael C Bernardi, Isaac Cann, and Klaus Schulten. Molecular dynamics study of enhanced man5b enzymatic activity. *Biotechnol. Biofuels*, 7, 2014.

[23] D. L. Beveridge, K. J. McConnell, M. A. Young, S. Vijayakumar, and G. Ravishanker. Molecular dynamics simulations of dna and a protein-dna complex including solvent. *Molecular Engineering*, 5(1):255–269, 1995.

[24] David M. Heyes. Transport coefficients of lennard-jones fluids: A molecular-dynamics and effective-hard-sphere treatment. *Phys. Rev. B*, 37:5677–5696, Apr 1988.

[25] Xue-Dong Din and Efstathios E. Michaelides. Kinetic theory and molecular dynamics simulations of microscopic flows. *Phys. Fluids*, 9(12):3915–3925, 1997.

[26] Remco Hartkamp, A. Ghosh, T. Weinhart, and S. Luding. A study of the anisotropy of stress in a fluid confined in a nanochannel. *J. Chem. Phys.*, 137(4):044711, 2012.

[27] H. Hoang and G. Galliero. Local viscosity of a fluid confined in a narrow pore. *Phys. Rev. E*, 86:021202, Aug 2012.

[28] Liem X. Dang and Tsun-Mei Chang. Molecular dynamics study of water clusters, liquid, and liquid-vapor interface of water with many-body potentials. *J. Chem. Phys.*, 106(19):8149–8159, 1997.

[29] I-Feng W. Kuo and Christopher J. Mundy. An ab initio molecular dynamics study of the aqueous liquid-vapor interface. *Science*, 303(5658):658–660, 2004.

[30] M. Morciano, M. Fasano, A. Nold, C. Braga, P. Yatsyshin, D. N. Sibley, B. D. Goddard, E. Chiavazzo, P. Asinari, and S. Kalliadasis. Nonequilibrium molecular dynamics simulations of nanoconfined fluids at solid-liquid interfaces. *J. Chem. Phys.*, 146(24):244507, 2017.

[31] Y.C. Chang, T.Y. Hou, B. Merriman, and S. Osher. A level set formulation of eulerian interface capturing methods for incompressible fluid flows. *J. Comput. Phys.*, 124(2):449–464, March 1996.

[32] M. J. Nieves-Remacha, L. Yang, and K. F. Jensen. Openfoam computational fluid dynamic simulations of two-phase flow and mass transfer in an advanced-flow reactor. *Ind. Eng. Chem. Res.*, 54(26):6649–6659, 2015.

[33] Robert Zwanzig. Ensemble method in the theory of irreversibility. *J. Chem. Phys.*, 33(5):1338–1341, 1960.

[34] Hazime Mori. Transport, collective motion, and brownian motion. *Prog. Theor. Phys.*, 33(3):423–455, 1965.

[35] Robert Zwanzig. Nonlinear generalized langevin equations. *J. Stat. Phys.*, 9(3):215–220, 1973.

[36] Alexandre Chorin and Panagiotis Stinis. Problem reduction, renormalization, and memory. *Commun. Appl. Math. Comput. Sci.*, 1(1):1–27, 2006.

[37] Tomoyuki Kinjo and Shi-aki Hyodo. Equation of motion for coarse-grained simulation based on microscopic description. *Phys. Rev. E*, 75:051109, May 2007.

[38] Carmen Hijón, Pep Español, Eric Vanden-Eijnden, and Rafael Delgado-Buscalioni. Mori–zwanzig formalism as a practical computational tool. *Faraday Discuss.*, 144:301–322, 2010.

[39] Shi-aki Hyodo. Coarse-grained equation of motion for many particle system containing internal degrees of freedom. *Japan J. Appl. Ind. Math.*, 28(1):69–87, 2011.

[40] Minxin Chen, Xiaotao Li, and Chun Liu. Computation of the memory functions in the generalized langevin models for collective dynamics of macromolecules. *J. Chem. Phys.*, 141(6):064112, 2014.

[41] Jeffrey Comer, James C. Gumbart, Jérôme Hénin, Tony Lelièvre, Andrew Pohorille, and Christophe Chipot. The adaptive biasing force method: Everything you always wanted to know but were afraid to ask. *J. Phys. Chem. B*, 119(3):1129–1151, 2015.

[42] Turgut Baştuğ, Po-Chia Chen, Swarna M. Patra, and Serdar Kuyucak. Potential of mean force calculations of ligand binding to ion channels from jarzynski's equality and umbrella sampling. *J. Chem. Phys.*, 128(15):155104, 2008.

[43] Robert Zwanzig. *Nonequilibrium Statistical Mechanics*. Oxford University Press, 2001.

[44] Alexandre J. Chorin, Ole H. Hald, and Raz Kupferman. Optimal prediction and the mori–zwanzig representation of irreversible processes. *PNAS*, 97(7):2968–2973, 2000.

[45] Eric Darve, Jose Solomon, and Amirali Kia. Computing generalized langevin equations and generalized fokker–planck equations. *PNAS*, 106(27):10884–10889, 2009.

[46] Alexis Torres-Carbajal, Salvador Herrera-Velarde, and Ramón Castañeda Priego. Brownian motion of a nano-colloidal particle: the role of the solvent. *Phys. Chem. Chem. Phys.*, 17:19557–19568, 2015.

[47] Oliver F. Lange and Helmut Grubmüller. Collective langevin dynamics of conformational motions in proteins. *J. Chem. Phys.*, 124(21):214903, 2006.

[48] Huan Lei, Nathan A. Baker, and Xiantao Li. Data-driven parameterization of the generalized langevin equation. *PNAS*, 113(50):14183–14188, 2016.

[49] Umberto Marini Bettolo Marconi and Pedro Tarazona. Dynamic density functional theory of fluids. *J. Chem. Phys.*, 110(16):8032–8044, 1999.

[50] 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. *J. Phys. Condens. Matter*, 25(3):035101, dec 2012.

[51] Benjamin D. Goddard, Andreas Nold, Nikos Savva, Grigorios A. Pavliotis, and Serafim Kalliadasis. General dynamical density functional theory for classical fluids. *Phys. Rev. Lett.*, 109:120603, Sep 2012.

[52] B. Goddard, G. Pavliotis, and S. Kalliadasis. The overdamped limit of dynamic density functional theory: Rigorous results. *Multiscale Modeling & Simulation*, 10(2):633–663, 2012.

[53] James F. Lutsko. A dynamical theory of nucleation for colloids and macromolecules. *J. Chem, Phys.*, 136(3):034509, 2012.

[54] James F Lutsko and Miguel A Durán-Olivencia. Classical nucleation theory from a dynamical approach to nucleation. *J. Chem. Phys*, 138(24):244908, 2013.

[55] P Yatsyshin, N Savva, and S Kalliadasis. Density functional study of condensation in capped capillaries. *J. Phys. Cond. Matter*, 27(27):275104, jun 2015.

[56] B.D. Goddard, A. Nold, and S. Kalliadasis. Dynamical density functional theory with hydrodynamic interactions in confined geometries. *J. Chem. Phys.*, 145:214106, 2016.

[57] Andreas Nold, Benjamin D. Goddard, Peter Yatsyshin, Nikos Savva, and Serafim Kalliadasis. Pseudospectral methods for density functional theory in bounded and unbounded domains. *J. Comput. Phys.*, 334:639–664, 2017.

[58] Andreas Nold, David N. Sibley, Benjamin D. Goddard, and Serafim Kalliadasis. Fluid structure in the immediate vicinity of an equilibrium three-phase contact line and assessment of disjoining pressure models using density functional theory. *Phys. Fluids*, 26(7):072001, 2014.

[59] Petr Yatsyshin, Nikos Savva, and Serafim Kalliadasis. Wetting of prototypical one- and two-dimensional systems: Thermodynamics and density functional theory. *J. Chem. Phys.*, 142(3):034708, 2015.

[60] P. Yatsyshin, A. O. Parry, C. Rascón, and S. Kalliadasis. Wetting of a plane with a narrow solvophobic stripe. *Mol. Phys.*, 116(15-16):1990–1997, 2018.

[61] A J Archer. Dynamical density functional theory for dense atomic liquids. *J. Phys. Condens. Matter*, 18(24):5617, 2006.

[62] B. D. Goddard, A. Nold, and S. Kalliadasis. Multi-species dynamical density functional theory. *J. Chem. Phys.*, 138(14):144904, 2013.

[63] 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 hydro-

dynamic interactions: derivation and numerical experiments. *J. Phys. Condens. Matter*, 25(3):035101, 2013.

[64] E. S. Kikkinides and P. A. Monson. Dynamic density functional theory with hydrodynamic interactions: Theoretical development and application in the study of phase separation in gas-liquid systems. *J. Chem. Phys.*, 142(9):094706, 2015.

[65] Miguel A. Durán-Olivencia, Benjamin D. Goddard, and Serafim Kalliadasis. Dynamical density functional theory for orientable colloids including inertia and hydrodynamic interactions. *J. Stat. Phys.*, 164(4):785–809, Aug 2016.

[66] Andrew J Archer and Markus Rauscher. Dynamical density functional theory for interacting brownian particles: stochastic or deterministic? *J. Phys. A: Math. Gen.*, 37(40):9325, 2004.

[67] K. R. Elder, Mark Katakowski, Mikko Haataja, and Martin Grant. Modeling elasticity in crystal growth. *Phys. Rev. Lett.*, 88:245701, 2002.

[68] Miguel A Durán-Olivencia, Peter Yatsyshin, Benjamin D Goddard, and Serafim Kalliadasis. General framework for fluctuating dynamic density functional theory. *New J. Phys*, 19(12):123022, 2017.

[69] Chao Yang and Zai-Sha Mao. *Numerical Simulation of Multiphase Reactors with Continuous Liquid Phase*. Academic Press, Oxford, 2014.

[70] Matthias Kruger and David S. Dean. A gaussian theory for fluctuations in simple liquids. *J. Chem. Phys.*, 146(13):134507, 2017.

[71] Changho Kim, Andy Nonaka, John B. Bell, Alejandro L. Garcia, and Aleksandar Donev. Stochastic simulation of reaction-diffusion systems: A fluctuating-hydrodynamics approach. *J. Chem. Phys.*, 146(12):124110, 2017.

[72] Cédric Villani. *Topics in optimal transportation*. American Mathematical Soc., 2003.

[73] José A Carrillo, Robert J McCann, Cédric Villani, et al. Kinetic equilibration rates for granular media and related equations: entropy dissipation and mass transportation estimates. *Revista Matematica Iberoamericana*, 19(3):971–1018, 2003.

[74] L.D. Landau, E.M. Lifshitz, and L.P. Pitaevskii. *Statistical Physics: Course of Theoretical Physics*, volume 5. Pergamon Press, 1980.

[75] Alejandro L Garcia, M Malek Mansour, George C Lie, and Enrico Cementi. Numerical integration of the fluctuating hydrodynamic equations. *J. Stat. Phys.*, 47(1-2):209–228, 1987.

[76] John B Bell, Alejandro L Garcia, and Sarah A Williams. Numerical methods for the stochastic landau-lifshitz navier-stokes equations. *Phys. Rev. E*, 76(1):016708, 2007.

[77] John B Bell, Alejandro L Garcia, and Sarah A Williams. Computational fluctuating fluid dynamics. *ESAIM Math. Model. Numer. Anal.*, 44(5):1085–1105, 2010.

[78] Florencio Balboa, John B Bell, Rafael Delgado-Buscalioni, Aleksandar Donev, Thomas G Fai, Boyce E Griffith, and Charles S Peskin. Staggered schemes for fluctuating hydrodynamics. *Multiscale Modeling & Simulation*, 10(4):1369–1408, 2012.

[79] Steven Delong, Boyce E Griffith, Eric Vanden-Eijnden, and Aleksandar Donev. Temporal integrators for fluctuating hydrodynamics. *Phys. Rev. E*, 87(3):033302, 2013.

[80] V Yu Glotov, Vasilii Mikhailovich Goloviznin, SA Karabasov, and AP Markesteijn. New two-level leapfrog scheme for modeling the stochastic landau-lifshitz equations. *Comput. Math. Math. Phys.*, 54(2):315–334, 2014.

[81] R Delgado-Buscalioni and G De Fabritiis. Embedding molecular dynamics within fluctuating hydrodynamics in multiscale simulations of liquids. *Phys. Rev. E*, 76(3):036709, 2007.

[82] G De Fabritiis, M Serrano, R Delgado-Buscalioni, and PV Coveney. Fluctuating hydrodynamic modeling of fluids at the nanoscale. *Phys. Rev. E*, 75(2):026307, 2007.

[83] G De Fabritiis, R Delgado-Buscalioni, and PV Coveney. Multiscale modeling of liquids with molecular specificity. *Phys. Rev. Lett.*, 97(13):134501, 2006.

[84] Sarah A Williams, John B Bell, and Alejandro L Garcia. Algorithm refinement for fluctuating hydrodynamics. *Multiscale Modeling & Simulation*, 6(4):1256–1280, 2008.

[85] Aleksandar Donev, John B Bell, Alejandro L Garcia, and Berni J Alder. A hybrid particle-continuum method for hydrodynamics of complex fluids. *Multiscale Modeling & Simulation*, 8(3):871–911, 2010.

[86] Paul J Atzberger. Spatially adaptive stochastic numerical methods for intrinsic fluctuations in reaction–diffusion systems. *J. Comp. Phys.*, 229(9):3474–3501, 2010.

[87] T. Schneider and E. Stoll. Molecular-dynamics study of a three-dimensional one-component model for distortive phase transitions. *Phys. Rev. B*, 17:1302–1322, Feb 1978.

[88] Shuichi Nosé. A unified formulation of the constant temperature molecular dynamics methods. *J. Chem. Phys.*, 81(1):511–519, 1984.

[89] Shuichi Nosé. A molecular dynamics method for simulations in the canonical ensemble. *Mol. Phys.*, 52(2):255–268, 1984.

[90] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, Mar 1985.

[91] Xin Yong and Lucy T. Zhang. Thermostats and thermostat strategies for molecular dynamics simulations of nanofluidics. *J. Chem. Phys.*, 138(8):084503, 2013.

[92] Reza Kamali and Ali Kharazmi. Molecular dynamics simulation of surface roughness effects on nanoscale flows. *Int. J. Therm. Sci.*, 50(3):226 – 232, 2011.

[93] L. B. Lucy. A numerical approach to the testing of the fission hypothesis. *Astron. J.*, 82(12):1013–1024, 1977.

[94] Juan M. Vanegas, Alejandro Torres-Sánchez, and Marino Arroyo. Importance of force decomposition for local stress calculations in biomembrane molecular simulations. *J. Chem. Theory Comput.*, 10(2):691–702, 2014.

[95] J A Zimmerman, E B WebbIII, J J Hoyt, R E Jones, P A Klein, and D J Bammann. Calculation of stress in atomistic simulation. *Model. Simul. Mater. Sci. Eng.*, 12(4):S319, 2004.

[96] T. Belytschko, Y. Krongauz, D. Organ, M. Fleming, and P. Krysl. Meshless methods: An overview and recent developments. *Comput. Methods Appl. Mech. Eng.*, 139(1):3 – 47, 1996.

[97] Thomas Weinhart, Remco Hartkamp, Anthony R. Thornton, and Stefan Luding. Coarse-grained local and objective continuum description of three-dimensional granular flows down an inclined surface. *Phys. Fluids*, 25(7):070605, 2013.

[98] Isaac Goldhirsch. Stress, stress asymmetry and couple stress: from discrete particles to continuous fields. *Granul. Matter*, 12(3):239–252, May 2010.

[99] Wm. G. Hoover, Carol G. Hoover, and James F. Lutsko. Microscopic and macroscopic stress with gravitational and rotational forces. *Phys. Rev. E*, 79:036709, Mar 2009.

[100] Satya N. Atluri and Shengping Shen. The basis of meshless domain discretization: the meshless local petrov–galerkin (mlpg) method. *Adv. Comput. Math.*, 23(1):73–93, Jul 2005.

[101] Gerald J. Wang and Nicolas G. Hadjiconstantinou. Molecular mechanics and structure of the fluid-solid interface in simple fluids. *Phys. Rev. Fluids*, 2:094201, Sep 2017.

[102] Walter Noll. Die herleitung der grundgleichungen der thermomechanik der kontinua aus der statistischen mechanik. *J. Rational Mech. Anal.*, 4:627–646, 1955.

[103] Nikhil Chandra Admal and E. B. Tadmor. A unified interpretation of stress in molecular systems. *J. Elast.*, 100(1):63–143, Jun 2010.

[104] P. Schofield and J. R. Henderson. Statistical mechanics of inhomogeneous fluids. *Proc. Royal Soc. A*, 379(1776):231–246, 1982.

[105] J. C. Maxwell. V. illustrations of the dynamical theory of gases. – part i. on the motions and collisions of perfectly elastic spheres. *Philos. Mag.*, 19(124):19–32, 1860.

[106] R.L. Liboff. *Kinetic Theory: Classical, Quantum, and Relativistic Descriptions*. Springer New York, 2003.

[107] S. Chapman. The kinetic theory of simple and composite monatomic gases : viscosity, thermal conduction, and diffusion. *Proc. Royal Soc. A*, 93(646):1–20, 1916.

[108] David Enskog. *Kinetische Theorie der Vorgänge in mässig verdünnten Gasen*. PhD thesis, Uppsala Universitet, 1917.

[109] David Enskog. *Kinetische Theorie der Wärmeleitung: Reibung und Selbst-diffusion in Gewissen verdichteten gasen und flüssigkeiten*. Almqvist & Wiksells boktryckeri-a.-b., 1922.

[110] Junfang Zhang, B. D. Todd, and Karl P. Travis. Viscosity of confined inhomogeneous nonequilibrium fluids. *J. Chem. Phys.*, 121(21):10778–10786, 2004.

[111] H. Hoang and G. Galliero. Shear viscosity of inhomogeneous fluids. *J. Chem. Phys.*, 136(12):124902, 2012.

[112] H. Hoang and G. Galliero. Local shear viscosity of strongly inhomogeneous dense fluids: from the hard-sphere to the lennard-jones fluids. *J. Phys. Condens. Matter*, 25(48):485001, 2013.

[113] Peter A. Thompson and Mark O. Robbins. Shear flow near solids: Epitaxial order and flow boundary conditions. *Phys. Rev. A*, 41:6830–6837, Jun 1990.

[114] Steve Plimpton. Fast parallel algorithms for short-range molecular dynamics. *J. Comput. Phys.*, 117(1):1 – 19, 1995.

[115] Loup Verlet. Computer "experiments" on classical fluids. i. thermodynamical properties of lennard-jones molecules. *Phys. Rev.*, 159:98–103, 1967.

[116] Peter A. Thompson and Sandra M. Troian. A general boundary condition for liquid flow at solid surfaces. *Nature*, 389:360–362, 1997.

[117] S. Hess. Structure and nonlinear flow behavior of simple and complex fluids. *Int. J. Thermophys.*, 23(4):905–920, Jul 2002.

[118] Remco Hartkamp, B. D. Todd, and Stefan Luding. A constitutive framework for the non-newtonian pressure tensor of a simple fluid under planar flows. *J. Chem. Phys.*, 138(24):244508, 2013.

[119] Thomas Soddemann, Burkhard Dünweg, and Kurt Kremer. Dissipative particle dynamics: A useful thermostat for equilibrium and nonequilibrium molecular dynamics simulations. *Phys. Rev. E*, 68:046702, Oct 2003.

[120] Laurent Joly, Christophe Ybert, Emmanuel Trizac, and Lydéric Bocquet. Hydrodynamics within the electric double layer on slipping surfaces. *Phys. Rev. Lett.*, 93:257805, 2004.

[121] Pierre-Andre Cazade, Remco Hartkamp, and Benoit Coasne. Structure and dynamics of an electrolyte confined in charged nanopores. *J. Phys. Chem. C*, 118(10):5061–5072, 2014.

[122] Eliodoro Chiavazzo, Matteo Fasano, Pietro Asinari, and Paolo Decuzzi. Scaling behaviour for the water transport in nanoconfined geometries. *Nat. Commun.*, 5:3565, 2014.

[123] Baotong Li, Jun Hong, and Xiangyang Tian. Generating optimal topologies for heat conduction by heat flow paths identification. *Int. Commun. Heat Mass Transfer*, 75:177 – 182, 2016.

[124] Johann Fischer and Michael Methfessel. Born-green-yvon approach to the local densities of a fluid at interfaces. *Phys. Rev. A*, 22:2836–2843, Dec 1980.

[125] Norman F. Carnahan and Kenneth E. Starling. Equation of state for nonattracting rigid spheres. *J. Chem. Phys.*, 51(2):635–636, 1969.

[126] S Morioka. The dense gas-like model of the viscosity for liquid metals. *J. Non-Cryst. Solids*, 341(1):46 – 52, 2004.

[127] J. O. Hirschfelder, C. F. Curtiss, and R. B. Bird. *Molecular theory of gases and liquids*. Wiley, New York, 1954.

[128] Philip D. Neufeld, A. R. Janzen, and R. A. Aziz. Empirical equations to calculate 16 of the transport collision integrals for the lennard-jones (12-6) potential. *J. Chem. Phys.*, 57(3):1100–1102, 1972.

[129] I. Bitsanis, J. J. Magda, M. Tirrell, and H. T. Davis. Molecular dynamics of flow in micropores. *J. Chem. Phys.*, 87(3):1733–1750, 1987.

[130] Guillaume Galliéro, Christian Boned, and Antoine Baylaucq. Molecular dynamics study of the lennard jones fluid viscosity: Application to real fluids. *Ind. Eng. Chem. Res.*, 44(17):6963–6972, 2005.

[131] A. Liakopoulos, F. Sofos, and T. E. Karakasidis. Darcy-weisbach friction factor at the nanoscale: From atomistic calculations to continuum models. *Phys. Fluids*, 29(5):052003, 2017.

[132] Sathish K. P. Sanjeevi, Ahad Zarghami, and Johan T. Padding. Choice of no-slip curved boundary condition for lattice boltzmann simulations of high-reynolds-number flows. *Phys. Rev. E*, 97:043305, Apr 2018.

[133] C. Braga, E.R. Smith, A. Nold, D.N. Sibley, and S. Kalliadasis. The pressure tensor across a liquid-vapour interface. *J. Chem. Phys.*, 149:044705, 2018.

[134] Ilan Benjamin. Molecular structure and dynamics at liquid-liquid interfaces. *Annu. Rev. Phys. Chem.*, 48(1):407–451, 1997.

[135] J Winkelmann. The liquid-vapour interface of pure fluids and mixtures: application of computer simulation and density functional theory. *J. Phys. Condens. Matter*, 13(21):4739, 2001.

[136] Andrij Trokhymchuk and José Alejandre. Computer simulations of liquid/vapor interface in lennard-jones fluids: Some questions and answers. *J. Chem. Phys.*, 111(18):8510–8523, 1999.

[137] J. López-lemus and J. Alejandre. Thermodynamic and transport properties of simple fluids using lattice sums: bulk phases and liquid-vapour interface. *Mol. Phys.*, 100(18):2983–2992, 2002.

[138] Carlos Braga, Edward R. Smith, Andreas Nold, David N. Sibley, and Serafim Kalliadasis. The pressure tensor across a liquid-vapour interface. *J. Chem. Phys.*, 149(4):044705, 2018.

[139] Michael A. Wilson, Andrew. Pohorille, and Lawrence R. Pratt. Molecular dynamics of the water liquid-vapor interface. *J. Phys. Chem.*, 91(19):4873–4878, 1987.

[140] Ramona S. Taylor, Liem X. Dang, and Bruce C. Garrett. Molecular dynamics simulations of the liquid/vapor interface of spc/e water. *J. Phys. Chem.*, 100(28):11720–11725, 1996.

[141] Jared T. Fern, David J. Keffer, and William V. Steele. Vapor-liquid equilibrium of ethanol by molecular dynamics simulation and voronoi tessellation. *J. Phys. Chem. B*, 111(46):13278–13286, 2007.

[142] Sudhir K. Singh, Ankit Sinha, Goutam Deo, and Jayant K. Singh. Vapor-liquid phase coexistence, critical properties, and surface tension of confined alkanes. *J. Phys. Chem. C*, 113(17):7170–7180, 2009.

[143] Shengfeng Cheng, Jeremy B. Lechman, Steven J. Plimpton, and Gary S. Grest. Evaporation of lennard-jones fluids. *J. Chem. Phys.*, 134(22), 2011.

[144] J. Stecki and S. Toxvaerd. Correlations in the liquid interfaces of simple liquids. *J. Chem. Phys.*, 103(22):9763–9771, 1995.

[145] J. B. Buhn, P. A. Bopp, and M. J. Hampe. A molecular dynamics study of a liquid-liquid interface: structure and dynamics. *Fluid Ph. Equilibria*, 224(2):221 – 230, 2004.

[146] P. Geysermans, N. Elyeznasni, and V. Russier. Layered interfaces between immiscible liquids studied by density-functional theory and molecular-dynamics simulations. *J. Chem. Phys.*, 123(20), 2005.

[147] Sanjib Senapati and Max L. Berkowitz. Computer simulation study of the interface width of the liquid/liquid interface. *Phys. Rev. Lett.*, 87:176101, 2001.

[148] Himanshu Goel, P. Rakesh Chandran, Kishalay Mitra, Saptarshi Majumdar, and Partha Ray. Estimation of interfacial tension for immiscible and partially miscible liquid systems by dissipative particle dynamics. *Chem. Phys. Lett.*, 600:62 – 67, 2014.

[149] Carlos Braga, Amparo Galindo, and Erich A. Müller. Nonequilibrium molecular dynamics simulation of diffusion at the liquid-liquid interface. *J. Chem. Phys.*, 141(15):154101, 2014.

[150] B. Heggen, W. Zhao, F. Leroy, A. J. Dammers, and F. Müller-Plathe. Interfacial properties of an ionic liquid by molecular dynamics. *J. Phys. Chem. B*, 114:6954, 2010.

[151] G. Hantal, I. Voroshylova, M. N. D. S. Cordeiro, and M. Jorge. A systematic molecular simulation study of ionic liquid surfaces using intrinsic analysis methods. *Phys. Chem. Chem. Phys.*, 14:5200, 2012.

[152] G. Hantal, M. Sega, S. Kantorovich, C. Schröder, and M. Jorge. Intrinsic structure of the interface of partially miscible fluids: An application to ionic liquids. *J. Phys. Chem. C*, 119(51):28448–28461, 2015.

[153] Stanislav Iatsevitch and Frank Forstmann. Density profiles at liquid-vapor and liquid-liquid interfaces: An integral equation study. *J. Chem. Phys.*, 107(17):6925–6935, 1997.

[154] Markus Rauscher. Dynamic density functional theory (ddft). In Dongqing Li, editor, *Encyclopedia of Microfluidics and Nanofluidics*, pages 428–433. Springer US, Boston, MA, 2008.

[155] Antonio Russo, Miguel A. Durán-Olivencia, Peter Yatsyshin, and Serafim Kalliadasis. Memory effects in dynamic density functional theory with fluctuation: Theory and simulations. *J. Phys. A*, 2020.

[156] Antonio Russo, Miguel A. Durán-Olivencia, Serafim Kalliadasis, and Remco Hartkamp. Macroscopic relations for microscopic properties at the interface between solid substrates and dense fluids. *J. Chem. Phys.*, 150(21):214705, 2019.

[157] Ahmed E. Ismail, Gary S. Grest, and Mark J. Stevens. Capillary waves at the liquid-vapor interface and the surface tension of water. *J. Chem. Phys.*, 125(1):014702, 2006.

[158] Scott W. Sides, Gary S. Grest, and Martin-D. Lacasse. Capillary waves at liquid-vapor interfaces: A molecular dynamics simulation. *Phys. Rev. E*, 60:6708–6713, 1999.

[159] M. Rao and D. Levesque. Surface structure of a liquid film. *J. Chem. Phys.*, 65(8):3233–3236, 1976.

[160] J.P.R.B. Walton, D.J. Tildesley, J.S. Rowlinson, and J.R. Henderson. The pressure tensor at the planar surface of a liquid. *Mol. Phys.*, 48(6):1357–1368, 1983.

[161] William G. Hoover. Canonical dynamics: Equilibrium phase-space distributions. *Phys. Rev. A*, 31:1695–1697, Mar 1985.

[162] Shiang-Tai Lin, Mario Blanco, and William A. Goddard III. The two-phase model for calculating thermodynamic properties of liquids from molecular dynamics: Validation for the phase diagram of lennard-jones fluids. *J. Chem. Phys.*, 119(22):11792–11805, 2003.

[163] Hiroshi Watanabe, Nobuyasu Ito, and Chin-Kun Hu. Phase diagram and universality of the lennard-jones gas-liquid system. *J. Chem. Phys.*, 136(20):204102, 2012.

[164] Shaofan Li and Ni Sheng. On multiscale non-equilibrium molecular dynamics simulations. *Int. J. Numer. Met. Engng*, 83:998–1038, 2010.

[165] Brad Lee Holian, Arthur F. Voter, and Ramon Ravelo. Thermostatted molecular dynamics: How to avoid the toda demon hidden in nosé-hoover dynamics. *Phys. Rev. E*, 52:2338–2347, Sep 1995.

[166] Robert D. Groot and Patrick B. Warren. Dissipative particle dynamics: Bridging the gap between atomistic and mesoscopic simulation. *J. Chem. Phys.*, 107(11):4423–4435, 1997.

[167] Harold A. Scheraga, Mey Khalili, and Adam Liwo. Protein-folding dynamics: Overview of molecular simulation techniques. *Annu. Rev. Phys. Chem.*, 58(1):57–83, 2007. PMID: 17034338.

[168] Rafael C. Bernardi, Marcelo C.R. Melo, and Klaus Schulten. Enhanced sampling techniques in molecular dynamics simulations of biological systems. *Biochim. Biophys. Acta*, 1850(5):872–877, 2015.

[169] Matteo Salvalaglio, Claudio Perego, Federico Giberti, Marco Mazzotti, and Michele Parrinello. Molecular-dynamics simulations of urea nucleation from aqueous solution. *PNAS*, 112(1):E6–E14, 2015.

[170] Yang Sun, Huajing Song, Feng Zhang, Lin Yang, Zhuo Ye, Mikhail I. Mendelev, Cai-Zhuang Wang, and Kai-Ming Ho. Overcoming the time limitation in molecular dynamics simulation of crystal nucleation: A persistent-embryo approach. *Phys. Rev. Lett.*, 120:085703, 2018.

[171] Robert Brown. Xxvii. a brief account of microscopical observations made in the months of june, july and august 1827, on the particles contained in the pollen of plants; and on the general existence of active molecules in organic and inorganic bodies. *Philos. Mag.*, 4(21):161–173, 1828.

[172] Don S. Lemons and Anthony Gythiel. Paul langevin's 1908 paper "on the theory of brownian motion" ["sur la théorie du mouvement brownien", c. r. acad. sci. (paris) 146, 530-533 (1908)]. *Am. J. Phys.*, 65(11):1079–1081, 1997.

[173] John G. Kirkwood. The statistical mechanical theory of transport processes i. general theory. *J. Chem. Phys.*, 14(3):180–201, 1946.

[174] J. H. Irving and John G. Kirkwood. The statistical mechanical theory of transport processes. iv. the equations of hydrodynamics. *J. Chem. Phys.*, 18(6):817–829, 1950.

[175] Denis J. Evans and Gary P. Morriss. *Statistical Mechanics of Nonequilibrium Liquids*. Academic Press, 1990.

[176] J. Xing and K. S. Kim. Application of the projection operator formalism to non-hamiltonian dynamics. *J. Chem. Phys.*, 134(4):044132, 2011.

[177] Kurt Hornik, Maxwell Stinchcombe, and Halbert White. Multilayer feedforward networks are universal approximators. *Neural Netw.*, 2(5):359–366, 1989.

[178] Kurt Hornik. Approximation capabilities of multilayer feedforward networks. *Neural Netw.*, 4(2):251–257, 1991.

[179] M A Nielsen. *Neural Networks and Deep Learning*. Determination Press, 2015.

[180] Martin Riedmiller. Advanced supervised learning in multi-layer perceptrons– from backpropagation to adaptive learning algorithms. *Comp. Stand. Inter.*, 16(3):265–278, 1994.

[181] Andrew D. Baczewski and Stephen D. Bond. Numerical integration of the extended variable generalized langevin equation with a positive prony representable memory kernel. *J. Chem. Phys.*, 139(4):044107, 2013.

[182] Shinnosuke Kawai. On the environmental modes for the generalized langevin equation. *J. Chem. Phys.*, 143(9):094101, 2015.

[183] Sean J. Taylor and Benjamin Lethamn. Forecasting at scale. *PeerJPreprint*, 2017.

[184] Stephen L. Mayo, Barry D. Olafson, and William A. Goddard. Dreiding: a generic force field for molecular simulations. *J. Phys. Chem.*, 94(26):8897–8909, 1990.

[185] D. Hossain, M.A. Tschopp, D.K. Ward, J.L. Bouvard, P. Wang, and M.F. Horstemeyer. Molecular dynamics simulations of deformation mechanisms of amorphous polyethylene. *Polymer*, 51(25):6071 – 6083, 2010.

[186] Marvin Bishop, M. H. Kalos, and H. L. Frisch. Molecular dynamics of polymeric systems. *J. Chem. Phys.*, 70(3):1299–1304, 1979.

[187] D. I. Dimitrov, A. Milchev, Kurt Binder, Leonid I. Klushin, and Alexander M. Skvortsov. Universal properties of a single polymer chain in slit: Scaling versus molecular dynamics simulations. *J. Chem. Phys.*, 128(23):234902, 2008.

[188] Peter Alaton, Boualem Djehiche, and David Stillberger. On modeling and pricing weather derivatives. *Appl. Math. Fin.*, 9(1):1–20, 02 2002.

[189] F. E. Benth and J. S. Benth. Weather derivatives and stochastic modelling of temperature. *Int. J. Stoch. Anal.*, 2011:1–21, 2011.

[190] E. Moreles and D. Martínez-López. Analysis of the simulated global temperature using a simple energy balance stochastic model. *Atmósfera*, 29(4):279–297, 2016.

[191] Berkeley-Earth. Experimental land-average temperature for period 1880–2014. http://berkeleyearth.lbl.gov/auto/Global/Complete_TAVG_daily.txt, 2018.

[192] Robert Rohde, Richard Muller, Robert Jacobsen, Saul Perlmutter, Arthur Rosenfeld, Jonathan Wurtele, Judith Curry, Charlotte Wickham, and Steven Mosher. Berkeley earth temperature averaging process. *Geoinfor. Geostat.: An Overview*, 1(2), 2013.

[193] E. Friis-Christensen and K. Lasses. Length of the solar cycle: An indicator of solar activity closely associated with climate. *Science*, 254(5032):698–700, 1991.

[194] Masafumi Takahashi. Non-ideal brownian motion, generalized langevin equation and its application to the security market. *Fin. Eng. Japanese Markets*, 3(2):87–119, Jul 1996.

[195] Min G. Lee, Akihiko Oba, and Hideki Takayasu. Parameter estimation of a generalized langevin equation of market price. In Hideki Takayasu, editor, *Empirical Science of Financial Fluctuations: The Advent of Econophysics*, pages 260–270. Springer Japan, Tokyo, 2002.

[196] Robert C. Merton. Lifetime portfolio selection under uncertainty: The continuous-time case. *Rev. Econ. Stat.*, 51(3):247–257, 1969.

[197] www.macrotrends.net. Nikkei 225 index - 67 year historical chart. https://www.macrotrends.net/2593/nikkei-225-index-historical-chart-data, 2018.

[198] Alan Mathison Turing. The chemical basis of morphogenesis. *Philos. Trans. Royal Soc. B*, 237(641):37–72, 1952.

[199] J. Schnakenberg. Simple chemical reaction systems with limit cycle behaviour. *J. Theor. Biol.*, 81(3):389–400, 1979.

[200] P. Arcuri and J. D. Murray. Pattern sensitivity to boundary and initial conditions in reaction-diffusion models. *J. Math. Biol.*, 24(2):141–165, Jul 1986.

[201] V. Dufiet and J. Boissonade. Conventional and unconventional turing patterns. *J. Chem. Phys.*, 96(1):664–673, 1992.

[202] P. K. Maini, D. L. Benson, and J. A. Sherratt. Pattern formation in reaction-diffusion models with spatially inhomogeneous diffusion coefficients. *Math. Med. Biol.*, 9(3):197–213, 1992.

[203] Adolfo Sanz-Anchelergues, Anatol M. Zhabotinsky, Irving R. Epstein, and Alberto P. Muñuzuri. Turing pattern formation induced by spatially correlated noise. *Phys. Rev. E*, 63:056124, Apr 2001.

[204] Tommaso Biancalani, Duccio Fanelli, and Francesca Di Patti. Stochastic turing patterns in the brusselator model. *Phys. Rev. E*, 81:046215, Apr 2010.

[205] Philip K. Maini, Thomas E. Woolley, Ruth E. Baker, Eamonn A. Gaffney, and S. Seirin Lee. Turing's model for biological pattern formation and the robustness problem. *Interface Focus*, 2(4):487–496, 2012.

[206] Kaushik Balakrishnan, John B. Bell, Aleksandar Donev, and Alejandro L. Garcia. Fluctuating hydrodynamics and direct simulation monte carlo. *AIP Conference Proceedings*, 1501(1):695–704, 2012.

[207] Takenobu Nakamura and Akira Yoshimori. Derivation of the nonlinear fluctuating hydrodynamic equation from the underdamped langevin equation. *J. Phys. A.: Math. Theor.*, 42(6):065001, jan 2009.

[208] Aleksandar Donev, Eric Vanden-Eijnden, Alejandro Garcia, and John Bell. On the accuracy of finite-volume schemes for fluctuating hydrodynamics. *Commun. Appl. Math. Comput. Sci.*, 5(2):149–197, 2010.

[209] José A. Carrillo, Alina Chertock, and Yanghong Huang. A finite-volume method for nonlinear nonlocal equations with a gradient flow structure. *Comm. Comput. Phys.*, 17(1):233–258, 2015.

[210] R. W. MacCormack. The effect of viscosity in hypervelocity impact cratering. *AIAA Paper*, 69(354), 1969.

[211] Bruce J. Berne, Jean Pierre Boon, and Stuart A. Rice. On the calculation of autocorrelation functions of dynamical variables. *J. Chem. Phys.*, 45(4):1086–1096, 1966.

[212] Amit Kumar Bhattacharjee, Kaushik Balakrishnan, Alejandro L Garcia, John B Bell, and Aleksandar Donev. Fluctuating hydrodynamics of multi-species reactive mixtures. *J. Chem. Phys*, 142(22):224107, 2015.

[213] Felix Otto. *Double degenerate diffusion equations as steepest descent*. Sonderforschungsbereich 256, 1996.

[214] Alethea BT Barbaro, José A Canizo, José A Carrillo, and Pierre Degond. Phase transitions in a kinetic flocking model of cucker–smale type. *Multiscale Modeling & Simulation*, 14(3):1063–1088, 2016.

[215] Peter K Sweby. High resolution schemes using flux limiters for hyperbolic conservation laws. *SIAM J. Numer. Anal.*, 21(5):995–1011, 1984.

[216] Peter E. Kloeden and Eckhard Platen. *Numerical Solution of Stochastic Differential Equations*. Springer-Verlag Berlin Heidelberg, 1992.

[217] Bernt Øksendal. Stochastic differential equations. In *Stochastic differential equations*, pages 65–84. Springer, 2003.

[218] Jessica G Gaines and Terry J Lyons. Variable step size control in the numerical solution of stochastic differential equations. *SIAM J. Appl. Math*, 57(5):1455–1484, 1997.

[219] Vassilios Sotiropoulos and Yiannis 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. *J. Chem. Phys.*, 128(1):014103, 2008.

[220] James F Lutsko and Miguel A Durán-Olivencia. A two-parameter extension of classical nucleation theory. *J. Phys. Condens. Matter*, 27(23):235101, may 2015.

[221] Nikolaos K Voulgarakis and Jhih-Wei Chu. Bridging fluctuating hydrodynamics and molecular dynamics simulations of fluids. *J. Chem. Phys*, 130(13):04B605, 2009.

[222] J. A. Barker and D. Henderson. Perturbation theory and equation of state for fluids. ii. a successful theory of liquids. *J. Chem. Phys.*, 47(11):4714–4721, 1967.

[223] Martin Burger, José A. Carrillo, and Marie-Therese Wolfram. A mixed finite element method for nonlinear diffusion equations. *Kinet. Relat. Models*, 3:59–83, 2010.

[224] Marianne Bessemoulin-Chatard and Francis Filbet. A finite volume scheme for nonlinear degenerate parabolic equations. *SIAM J. Sci. Comput.*, 34(5):B559–B583, 2012.

[225] Rafael Bailo, Jose A Carrillo, and Jingwei Hu. Fully discrete positivity-preserving and energy-decaying schemes for aggregation-diffusion equations with a gradient flow structure. *arXiv preprint arXiv:1811.11502*, 2018.

[226] Zheng Sun, José A. Carrillo, and Chi-Wang Shu. A discontinuous galerkin method for nonlinear parabolic equations and gradient flow problems with interaction potentials. *J. Comp. Phys.*, 352:76 – 104, 2018.

[227] Luis Neves De Almeida, Federica Bubba, Benoît Perthame, and Camille Pouchol. Energy and implicit discretization of the fokker-planck and keller-segel type equations. *arXiv preprint arXiv:1803.10629*, 2018.

[228] S. Carillo, J.A. amd Kalliadasis, S.P. Perez, and C.-W. Shu. Well-balanced finite-volume schemes for hydrodynamic equations with general free energy. *arXiv preprint arXiv:1812.00980*, 2020.

[229] Julien Mendes, Antonio Russo, Sergio P. Perez, and Serafim Kalliadasis. A finite-volume scheme for gradient-flow equations with non-homogeneous diffusion. *Computers & Mathematics with Applications*, 89:150–162, 2021.

[230] Sigal Gottlieb and Chi-Wang Shu. Total variation diminishing Runge-Kutta schemes. *Math. Comput. Am. Math. Soc.*, 67(221):73–85, 1998.

[231] Mohamed Hayek. An exact solution for a nonlinear diffusion equation in a radially symmetric inhomogeneous medium. *Comput. Math. Appl.*, 68(12):1751–1757, 2014.

[232] Petr Yatsyshin, Nikos Savva, and Serafim Kalliadasis. Geometry-induced phase transition in fluids: Capillary prewetting. *Phys. Rev. E*, 87:020402, Feb 2013.

[233] P Yatsyshin, A O Parry, C Rascón, and S Kalliadasis. Classical density functional study of wetting transitions on nanopatterned surfaces. *J. Phys. Condens. Matter*, 29(9):094001, 2017.

[234] R. Zwanzig. Time-correlation functions and transport coefficients in statistical mechanics. *Annu. Rev. Phys. Chem.*, 16(1):67–102, 1965.

[235] William G. Hoover, Denis J. Evans, Richard B. Hickman, Anthony J. C. Ladd, William T. Ashurst, and Bill Moran. Lennard-jones triple-point bulk and shear viscosities. green-kubo theory, hamiltonian mechanics, and nonequilibrium molecular dynamics. *Phys. Rev. A*, 22:1690–1697, Oct 1980.

[236] C. Hoheisel, R. Vogelsang, and M. Schoen. Bulk viscosity of the lennard-jones fluid for a wide range of states computed by equilibrium molecular dynamics. *J. Chem. Phys.*, 87(12):7195–7198, 1987.

[237] William G. Hoover, Anthony J. C. Ladd, Richard B. Hickman, and Brad Lee Holian. Bulk viscosity via nonequilibrium and equilibrium molecular dynamics. *Phys. Rev. A*, 21:1756–1760, May 1980.

[238] G. J. Guo and Y. Zhang. Equilibrium molecular dynamics calculation of the bulk viscosity of liquid water. *Mol. Phys.*, 99(4):283–289, 2001.

[239] Vladimir G. Baidakov and Sergey P. Protsenko. Metastable lennard-jones fluids. iii. bulk viscosity. *The Journal of Chemical Physics*, 141(11), 2014.

[240] H. Van Beijeren and M.H. Ernst. The modified enskog equation. *Physica*, 68(3):437 – 456, 1973.

[241] Dror Givon, Raz Kupferman, and Andrew Stuart. Extracting macroscopic dynamics: model problems and algorithms. *Nonlinearity*, 17(6):R55, 2004.

[242] Georg A. Gottwald, Daan T. Crommelin, and Christian L. E. Franzke. Stochastic climate theory. In Christian L. E. Franzke and Terence J.Editors O'Kane, editors, *Nonlinear and Stochastic Climate Dynamics*, pages 209–240. Cambridge University Press, 2017.

[243] David A. Dickey and Wayne A. Fuller. Distribution of the estimators for autoregressive time series with a unit root. *J. Am. Stat. Assoc.*, 74(366a):427–431, 1979.

Chapter A

Chapter 2: Additional details and results

A.1. MD simulations of a single fluid in a confined geometry

A.1.1. Global density

Using the partition of the fluid domain into an interface region and a bulk region, we propose an analysis of the averaged fluid property in these regions. Such an approach can be generalized to any confined fluid system and can thus present a practical tool for the design of engineering nano-devices. First, we define the volume of the confined region for a channel with cross surface S as $V_c = \delta(T, \epsilon_{\mathrm{LJ},f,w})S$, and consequently the free volume as $V_f = V_{tot} - V_c$, where V_{tot} is the total volume of the channel. Thus, we compute the average confined density as $\rho_c = \frac{1}{V_c}\int_{V_c} \rho(y)dV$ and the average free density as $\rho_f = \frac{1}{V_f}\int_{V_f} \rho(y)dV$. The study of this quantities normalized with respect to the bulk density ρ_b is reported in Fig. A.1. It emerges that the average free density is approximately equal to the bulk density, regardless of the fluid temperature or the fluid-wall interaction strength. On the contrary, the effect of the layering at the walls is important when considering $\rho_c/\rho_b = \rho_c/\rho_b(T, \epsilon)$. Within the range of conditions examined here, a satisfactory estimation of the dependency is given by the relation $\rho_c/\rho_b(T, \epsilon) \sim A + B\mathrm{Wa}\sqrt{T}$, with A and B being two constant parameters. This relation

allows us to compute the average density of the channel $\bar{\rho}$ as function of the bulk density:

$$\bar{\rho} \sim \rho_b \left[\frac{h - 2\delta}{h} + \frac{2\delta}{h} \left(A + B \ \mathrm{Wa}\sqrt{T} \right) \right], \tag{A.1}$$

with A and B being fitting parameters with values $A \sim 0.8$ and $B \sim 0.1$.

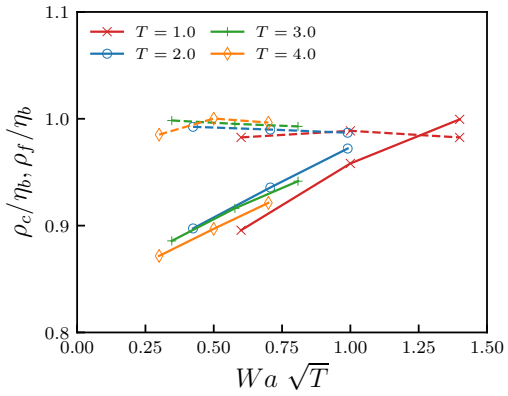

Figure A.1. Normalized average free density ρ_f/ρ_b (dashed lines) and normalized average confined density ρ_c/ρ_b (solid lines) for several temperatures and walls interactions. The ratio ρ_f/ρ_b does not show a significant dependency on the system conditions and is close to unity, while ρ_c/ρ_b exhibits a dependency well approximated by the relation $\rho_c/\rho_b(T, \epsilon) \sim A + B\mathrm{Wa}\sqrt{T}$, with A and B being constant parameters.

A.1.2. Derivation of the viscosity of a dense inhomogeneous hard-sphere fluid from kinetic theory

For dense inhomogeneous hard sphere system, Din and Michaelides [25] derived an expression for the shear viscosity directly from Boltzmann equation. In this section we report the main steps of the procedure, and we suggest the reader to refer to the original work for more details. The starting point is to consider the revised Enskog equation for

the singlet probability density function $W_1(\mathbf{r}_1, \mathbf{u}_1, t)$ of inhomogeneous dense fluid [240]:

$$\begin{aligned}
&\frac{\partial W_1}{\partial t} + \frac{\partial W_1}{\partial \mathbf{r}_1}\mathbf{u}_1 + \frac{\partial W_1}{\partial \mathbf{u}_1}\dot{\mathbf{u}}_1 = \\
&\int\int_{\mathbf{k}\cdot\mathbf{u}_{21}>0} \Big[g(\mathbf{r}_1, \mathbf{r}_1 + \sigma_{\mathrm{HS}}\mathbf{k}) W_1(\mathbf{r}_1, \mathbf{u'}_1, t) \times \\
&\times W_1(\mathbf{r}_1 + \sigma_{\mathrm{HS}}\mathbf{k}, \mathbf{u'}_2, t) - g(\mathbf{r}_1, \mathbf{r}_1 - \sigma_{\mathrm{HS}}\mathbf{k}) W_1(\mathbf{r}_1, \mathbf{u}_1, t) \times \\
&\times W_1(\mathbf{r}_1 - \sigma_{\mathrm{HS}}\mathbf{k}, \mathbf{u}_2, t) \Big] \sigma_{\mathrm{HS}}^2 \, (\mathbf{k}\cdot\mathbf{u}_{21}) \, d\ \mathbf{k}\ d\ \mathbf{u}_{21},
\end{aligned} \tag{A.2}$$

where σ_{HS} is the hard sphere diameter, $\mathbf{u}_{21} = \mathbf{u}_2 - \mathbf{u}_1$ is the relative velocity between two particles, $\mathbf{k} = (\mathbf{r}_2 - \mathbf{r}_1) \,/\, \mid \mathbf{r}_2 - \mathbf{r}_1 \mid$ is a unit vector connecting two particles, and $\mathbf{u'}_1 = \mathbf{u}_1 + (\mathbf{k}\cdot\mathbf{u}_{21})\,\mathbf{k}$ and $\mathbf{u'}_2 = \mathbf{u}_2 - (\mathbf{k}\cdot\mathbf{u}_{21})\,\mathbf{k}$ are the post-collision velocities of the two particle. In order to solve the previous equation for a steady-state Couette flow, the singlet pdf is assumed to have the following form:

$$W_1(\mathbf{r}_1, \mathbf{u}_1) = W_1^{(0)}(\mathbf{r}_1, \mathbf{u}_1 - \mathbf{u}) \left[1 + \Phi(\mathbf{r}_1, \mathbf{u}_1 - \mathbf{u})\right], \tag{A.3}$$

where $W_1^{(0)}$ is the equilibrium distribution given by

$$W_1^{(0)}(\mathbf{r}_1, \mathbf{u}_1 - \mathbf{u}) = n(\mathbf{r}_1) \left(\frac{m}{2\pi k_B T}\right)^{3/2} e^{-\frac{m(\mathbf{u}_1 - \mathbf{u})^2}{2k_B T}}, \tag{A.4}$$

and $\Phi(\mathbf{r}_1, \mathbf{u}_1 - \mathbf{u})$ is a perturbation function due to non-equilibrium effects. After some algebraic manipulation, it turns out that the perturbation function can be expressed as a function directly proportional to the velocity gradient, namely:

$$\begin{aligned}
\Phi(\mathbf{r}_1, \mathbf{u}_1 - \mathbf{u}) &= -b(\mathbf{r}_1) \left(\frac{m}{2k_B T}\right) \times \\
&\times \left[(\mathbf{u}_1 - \mathbf{u})\,(\mathbf{u}_1 - \mathbf{u}) - \mid \mathbf{u}_1 - \mathbf{u} \mid^2 \mathbf{1}/3\right] : \boldsymbol{\nabla}_{\mathbf{r}_1}\mathbf{u}.
\end{aligned} \tag{A.5}$$

By using $W_1(\mathbf{r}_1, \mathbf{u}_1)$, it follows that the kinetic component of the viscosity can be written as:

$$\eta_{\mathrm{HS}}^k(y) = \frac{16}{5}\eta^0 \rho(y) b^*(y), \tag{A.6}$$

where $\eta^0 = 5(k_B T m/\pi)^{1/2}/(16\sigma_{\mathrm{HS}}^2)$ is the zero-density viscosity, and $b^*(y) = \sqrt{\pi k_B T/4m\sigma_{\mathrm{HS}}^2}\, b(y)$ is a dimensionless space-dependent function, that can be evaluated by solving numeri-

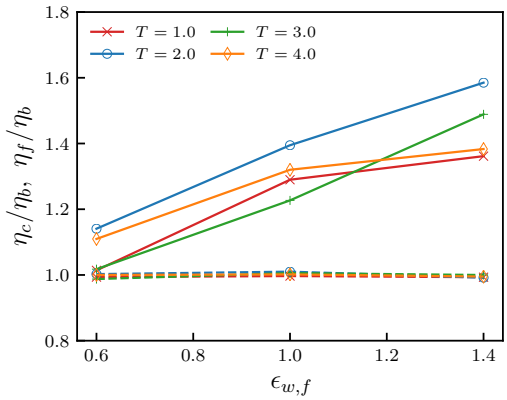

Figure A.2. Free viscosity (dashed lines) and confined viscosity (solid lines) normalized with respect to the viscosity in the bulk region.

cally the following integral equation:

$$\begin{aligned} &4b^*(y)\int_{-\sigma_{\rm HS}}^{\sigma_{\rm HS}} g(y,y+y')\rho(y+y')dy' \\ &-\frac{25}{3}\int_{-\sigma_{\rm HS}}^{\sigma_{\rm HS}}\left[b^*(y+y')-b^*(y)\right]g(y,y+y')\rho(y+y')dy' \\ &=\frac{5}{2}+\frac{\pi}{3}\int_{-\sigma_{\rm HS}}^{\sigma_{\rm HS}} g(y,y+y')\rho(y+y')dy'. \end{aligned} \tag{A.7}$$

Moreover, the potential contribution to the viscosity for a hard sphere fluid is given by:

$$\eta^v_{\rm HS}(y) = \frac{8\pi}{5}\eta^0\int_{-\sigma_{\rm HS}}^{\sigma_{\rm HS}} y'\ sin^3\left(\arccos(y')\right)\ I(y,y')\ dy', \tag{A.8}$$

where $I(y,y')$ is given by the infinite series:

$$\begin{aligned} I(y,y') = y'\sum_1^\infty \frac{(-y'^{s-1})}{s!}\frac{\partial^{s-1}}{\partial^{s-1}y}\Big\{ &g(y,y+y')\rho(y)\rho(y+y') \\ &\times\left[2+b^*(y+y')+b^*(y)\right]\Big\}. \end{aligned} \tag{A.9}$$

A.1.3. Global shear viscosity for a fluid in a confined geometry

Following the same approach of Sect. A.1.1, we define the confined viscosity $\eta_c = \frac{1}{V_c}\int_{V_c}\eta(y)dV$ and the free viscosity $\eta_f = \frac{1}{V_f}\int_{V_f}\eta(y)dV$. With prior knowledge of η_c

and η_f, the average viscosity of the system in the channel $\bar{\eta}$ can then be expressed as:

$$\bar{\eta} = \eta_b \left(\frac{V_f}{V_{tot}} \frac{\eta_f}{\eta_b} + \frac{V_c}{V_{tot}} \frac{\eta_c}{\eta_b} \right). \tag{A.10}$$

where the boundary of the confinement region is defined through δ. This expression can be then parametrized, knowing the variations of η_c and η_f as function of the temperature and the fluid-wall interaction. Figure A.2 shows that the free viscosity can be considered approximately equal to the viscosity in the bulk region, regardless of the fluid temperature and the fluid-wall interaction strength. Conversely, the ratio η_c/η_b varies (approximately linearly) with $\epsilon_{\mathrm{LJ},f,w}$ (and does not show strong dependencies on T for the analysed systems). Based on these observations, Eq. (A.10) can be approximated as:

$$\bar{\eta} \sim \eta_b \left[\frac{h - 2\delta}{h} + \frac{2\delta}{h} \left(A_\eta + B_\eta \, \epsilon_{\mathrm{LJ},f,w} \right) \right], \tag{A.11}$$

with A_η and B_η being some fitting parameters. Specifically, the values $A_\eta \sim 1.3$ and $B_\eta \sim 0.5$ lead to a discrepancy between Eq. (A.11) and MD results lower than 10%.

Eq. (A.11) represents an useful tool in engineering nano-devices design, since it allows to predict the average viscosity in a channel based on its geometry (through h), on the fluid-wall coupling thermodynamic parameter δ, and on the fluid viscosity in the bulk, which can be accurately modelled with formulas developed for homogeneous fluids (i.e. as in Sect. 2.2.7 or as in Ref. [111]).

A.2. MD simulations of multi-phase systems

A.2.1. LV interface vs shear rate

Fluid interfaces in non-equilibrium systems can be affected and its structure modified when the system is sheared. Since this effect is expected to be more relevant at high shear rates, we tested the system under various shear rates. As shown in Fig. A.3, shear rates above 0.1 are already high enough to introduce a significant deviation of the density profile.

A.2.2. Stability of the LL interface in time

In order to measure the effects of the repulsive potential, the LL systems analysed in this work are composed of two partially miscible fluids. Being not in-miscible, the two

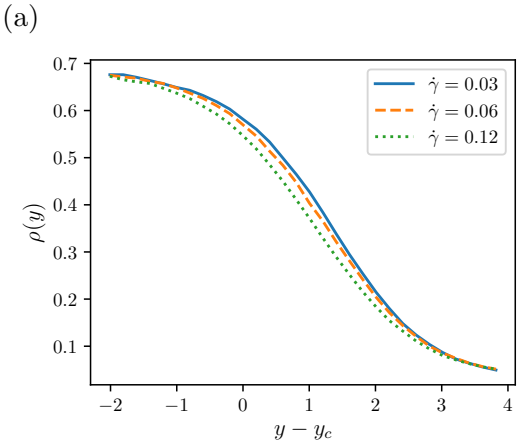

Figure A.3. Density profile at the LV interface for different shear rates.

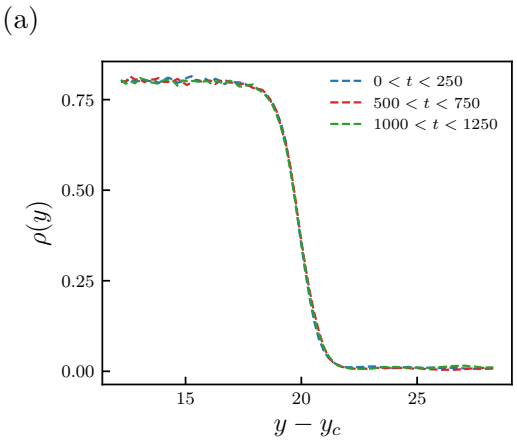

Figure A.4. Density profiles in time at the LV interface.

fluids will start mixing if the system is simulated long enough. For this reason, we gathered data for each trajectory by performing multiple simulations for a limited time. To measure what is the maximum time allowed before mixing becomes important, we observed the average density profile evaluated in different time windows. As shown in Fig. A.3, for time windows within $t < 1250$ no difference can be seen in the density profiles.

Chapter B

Chapter 3: additional details and results

B.1. GLE

B.1.1. Theoretical background

Mori-Zwanzig's formalism Let us consider the following (linear of non-linear) deterministic dynamical system:

$$\begin{cases} \frac{d\mathbf{z}}{dt} = \mathbf{f}(\mathbf{z}) \\ \mathbf{z}(0) = \mathbf{z}_0 \end{cases} \tag{B.1}$$

where $\mathbf{z} \in \mathbb{R}^n$ is a vector of independent variables. For the system in Eq. (B.1), it can be defined a set of observables $\boldsymbol{\mathcal{O}}(\mathbf{z}, t) = \phi(\mathbf{z}(t))$, where ϕ represent the transformation map between $\mathbf{z}$ and $\boldsymbol{\mathcal{O}}$. By using the chain rule, it is easy to show that the evolution equation of $\boldsymbol{\mathcal{O}}(\mathbf{z}, t)$ can be written as:

$$\begin{cases} \frac{\partial \boldsymbol{\mathcal{O}}}{\partial t}(\mathbf{z}, t) = \mathcal{L}\boldsymbol{\mathcal{O}} \\ \boldsymbol{\mathcal{O}}(\mathbf{z}, 0) = \boldsymbol{\mathcal{O}}_0 \end{cases} \tag{B.2}$$

where it was introduced the operator $\mathcal{L} = \mathbf{f}(\mathbf{z}) \cdot \boldsymbol{\nabla}_{\mathbf{z}}$. It follows that the solution of Eq. (B.2) can be written as:

$$\boldsymbol{\mathcal{O}}(\mathbf{z}, t) = e^{\mathcal{L}t} \boldsymbol{\mathcal{O}}_0 \tag{B.3}$$

where the exponential has to be intended as the power series that defines the exponential map between matrices.

If we are only interested in the dynamics of some observables $\boldsymbol{\mathcal{O}}$, rather then the whole solution $\mathbf{z}(t)$, we can define a projection operator $\mathcal{P}$, which maps functions of $\mathbf{z}$ into function of $\boldsymbol{\mathcal{O}}$. It is worth underlining that, in general, the set of observables $\boldsymbol{\mathcal{O}}$ may be defined by a linear or nonlinear transformation of $\mathbf{z}$, but in any case the evolution of $\boldsymbol{\mathcal{O}}$ is supposed to be unitary, i.e. $\mid \boldsymbol{\mathcal{O}}(t) \mid^2 = \mid \boldsymbol{\mathcal{O}}(0) \mid^2$. A simple, but still important, scenario is given by $\boldsymbol{\mathcal{O}}$ being a subset of $\mathbf{z}$. As we will see later, this case plays a fundamental role in dimensional reductions of multi-component systems, i.e. colloidal particles in a thermal bath. Given a projection operator $\mathcal{P}$, namely a transformation from a vector space to itself such that $\mathcal{P}^2 = \mathcal{P}$, one can follow Mori-Zwanzig's formalism [33, 34, 35] to obtain a form of Eq. (B.2) suitable for system dimensionality reduction. Note that at this point no constrain is put on the form of the projection operator. After defining the operator $\mathcal{Q} = \mathbf{1} - \mathcal{P}$, orthogonal to $\mathcal{P}$, Eq. (B.2) can be rewritten as:

$$\frac{\partial \boldsymbol{\mathcal{O}}}{\partial t}(\mathbf{z}, t) = \mathcal{L} e^{\mathcal{L}t} \boldsymbol{\mathcal{O}}_0 = e^{\mathcal{L}t} \mathcal{P}\mathcal{L}\boldsymbol{\mathcal{O}}_0 + e^{\mathcal{L}t} \mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0 \tag{B.4}$$

Duhamel-Dyson's formula allows to rewrite the exponential term $e^{\mathcal{L}t}$ as:

$$e^{\mathcal{L}t} = e^{\mathcal{Q}t} + \int_0^t e^{\mathcal{L}(t-\tau)} \ \mathcal{P} \ e^{\mathcal{Q}\tau} \ d\tau \tag{B.5}$$

and, consequently, Eq. (B.4) becomes the so called Mori-Zwanzig's equation:

$$\frac{\partial \boldsymbol{\mathcal{O}}}{\partial t}(\mathbf{z}, t) = e^{\mathcal{L}t} \mathcal{P}\mathcal{L}\boldsymbol{\mathcal{O}}_0 + \int_0^t e^{\mathcal{L}(t-\tau)} \mathcal{P}\mathcal{L} e^{\mathcal{Q}\tau} \mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0 \ d\tau + e^{\mathcal{Q}\mathcal{L}t} \mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0 \tag{B.6}$$

The first term is the Markovian contribution, the second constitutes the memory term and the last one is often interpreted as the noise. It is worth noticing that, at this stage, Eq. (B.6) is exactly equivalent to Eq. (B.1) and is valid independently from the specific choice of the projection operator $\mathcal{P}$. Mori and Zwanzig [34, 35, 43] proposed two different projection operators leading to different forms of GLE, that we will briefly discuss in next sections.

If we name the noise term $\mathbf{R}(t) = e^{\mathcal{Q}\mathcal{L}t} \mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0$, then the following dynamical system

remains determined:

$$\begin{cases} \dfrac{\partial \mathbf{R}}{\partial t}(\boldsymbol{\mathcal{O}}_0, t) = \mathcal{Q}\mathcal{L}\mathbf{R}(\boldsymbol{\mathcal{O}}_0, t), \\ \mathbf{R}(\boldsymbol{\mathcal{O}}_0, t) = \mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0. \end{cases} \tag{B.7}$$

Projecting Eq. (B.7) according to $\mathcal{P}$, it follows:

$$\begin{cases} \mathcal{P}\dfrac{\partial \mathbf{R}}{\partial t}(\boldsymbol{\mathcal{O}}_0, t) = \mathcal{P}\mathcal{Q}\mathcal{L}\mathbf{R}(\boldsymbol{\mathcal{O}}_0, t) = \mathbf{0}, \\ \mathcal{P}\mathbf{R}(\boldsymbol{\mathcal{O}}_0, t) = \mathcal{P}\mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0 = \mathbf{0}, \end{cases} \tag{B.8}$$

where we have used the property of the projection operator $\mathcal{P}\mathcal{Q} = \mathbf{0}$. This shows that $\mathbf{R}(t)$ is orthogonal to the range of $\mathcal{P}$ at any time t. However, in order to express $\mathbf{R}(t)$ as a stochastic process, it is necessary to have either time scale separation or weak coupling between resolved and unresolved variables [241]. When at least one of such conditions occurs, at least asymptotically, the influence of the unresolved variables may be interpreted as sum of many uncorrelated events, and consequently can be treated with Central Limit Theorem [242]. Thus, it is the Central Limit Theorem that determines the Gaussian shape for the distribution of $\mathbf{R}(t)$, while its time correlation follows from the fluctuation dissipation theorem, as shown in what follows.

Mori's projection operator The projection operator introduced by Mori [34], when applied to a general variable $\mathbf{A}(\mathbf{z})$, is given by:

$$\mathcal{P}\mathbf{A}(\mathbf{z}) = \langle \mathbf{A}, \boldsymbol{\mathcal{O}}_0 \rangle \langle \boldsymbol{\mathcal{O}}_0, \boldsymbol{\mathcal{O}}_0 \rangle^{-1} \boldsymbol{\mathcal{O}}_0 \tag{B.9}$$

where the inner product $\langle \mathbf{A}, \mathbf{B} \rangle$ is defined as

$$\langle \mathbf{A}, \mathbf{B} \rangle = \int \rho(\mathbf{z}) \mathbf{A}(\mathbf{z}) \mathbf{B}^*(\mathbf{z}) \, d\mathbf{z} \tag{B.10}$$

with $\rho(\mathbf{z})$ being a normalized probability density function defined in the phase space of the original system and $\mathbf{B}^*$ the conjugate transpose of $\mathbf{B}$. In case of systems with Hamiltonian $\mathcal{H}$ in a canonical ensemble, the probability density function is $\rho(\mathbf{z}) = Z^{-1} e^{-\beta \mathcal{H}(\mathbf{z})}$, where Z is the partition function and $\beta = k_B T$. Employing Mori's operator in Eq. (B.6), we obtain the Markovian term:

$$e^{\mathcal{L}t}\mathcal{P}\mathcal{L}\boldsymbol{\mathcal{O}}_0 = \langle \mathcal{L}\boldsymbol{\mathcal{O}}_0, \boldsymbol{\mathcal{O}}_0 \rangle \langle \boldsymbol{\mathcal{O}}_0, \boldsymbol{\mathcal{O}}_0 \rangle^{-1} \boldsymbol{\mathcal{O}}(t). \tag{B.11}$$

Moreover, from the definition of $\mathbf{R}(t)$, we obtain the memory term:

$$\int_0^t e^{\mathcal{L}(t-\tau)}\mathcal{P}\mathcal{L}e^{\mathcal{Q}\tau}\mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0\ d\tau = -\int_0^t \boldsymbol{\theta}(\tau)\boldsymbol{\mathcal{O}}(t-\tau)\ d\tau \tag{B.12}$$

where the memory kernel is defined as $\boldsymbol{\theta}(t) = -\langle\mathcal{L}\mathbf{R}(t), \boldsymbol{\mathcal{O}}_0\rangle\langle\boldsymbol{\mathcal{O}}_0, \boldsymbol{\mathcal{O}}_0\rangle^{-1}$. Since $\mathcal{Q}\mathbf{R}(t) = \mathbf{R}(t)$, and $\mathcal{L}$ is an anti-Hermitian operator [43], it follows that $\langle\mathcal{L}\mathbf{R}(t), \boldsymbol{\mathcal{O}}_0\rangle = -\langle\mathbf{R}(t), \mathcal{L}\boldsymbol{\mathcal{O}}_0\rangle = -\langle\mathbf{R}(t), \mathcal{Q}\mathcal{L}\boldsymbol{\mathcal{O}}_0\rangle = -\langle\mathbf{R}(t), \mathbf{R}(0)\rangle$. Hence, we obtain the following relation:

$$\boldsymbol{\theta}(t) = \langle\mathbf{R}(t), \mathbf{R}(0)\rangle\langle\boldsymbol{\mathcal{O}}_0, \boldsymbol{\mathcal{O}}_0\rangle^{-1}, \tag{B.13}$$

which constitutes the fluctuation dissipation theorem.

Zwanzig projection operator As Zwanzig pointed out, Mori's projection operator leads to a linearised GLE [43]. Zwanzig [33, 43] defined the projection operator applied to the variable $\mathbf{A}(\mathbf{z})$ through the following conditional expectation:

$$\mathcal{P}\mathbf{A}(\mathbf{z}) = \frac{\int \rho(\mathbf{z})\ \mathbf{A}(\mathbf{z})\ \delta(\boldsymbol{\mathcal{O}} - \phi(\mathbf{z}))\ d\mathbf{z}}{\int \rho(\mathbf{z})\ \delta(\boldsymbol{\mathcal{O}} - \phi(\mathbf{z}))\ d\mathbf{z}}, \tag{B.14}$$

where $\delta(\boldsymbol{\mathcal{O}} - \phi(\mathbf{z})) = \prod_j \delta(\boldsymbol{\mathcal{O}}_j - \phi_j)$.

In molecular dynamics, the set of observables is often defined as a subset of the original coordinates, namely $\boldsymbol{\mathcal{O}} \subseteq \mathbf{z}$. In this case, Zwanzig's projection operator allows to express the Markovian term in Eq. (B.6) as function of the potential of mean force. To show this, let us consider an isothermal Hamiltonian system of N particles with coordinates $\mathbf{z} = \{\mathbf{r}, \mathbf{p}\}$, where $\mathbf{r} = \{\mathbf{r}_1 \dots \mathbf{r}_N\}$ and $\mathbf{p} = \mathbf{p}_1 \dots \mathbf{p}_N$ are position and momenta, respectively. With $\mathbf{f}(\mathbf{z}) = -\nabla_{\mathbf{z}} V(\mathbf{z})$, Eq. (B.1) gives the Newton's equations of motion for a system of interacting particles. Suppose one is interested in the dynamical evolution of only n of the original N particles, whose coordinates (called relevant variables) and are indicated as $\tilde{\mathbf{z}} = \{\mathbf{r}_1 \dots \mathbf{r}_n, \mathbf{p}_1 \dots \mathbf{p}_n\}$. The remaining variables, called unresolved variables, are denoted by $\hat{\mathbf{z}} = \{\mathbf{r}_{n+1} \dots \mathbf{r}_N, \mathbf{p}_{n+1} \dots \mathbf{p}_N\}$. Hence, inserting Zwanzig's operator in Eq. (B.6), we obtain the Markovian term in the form:

$$\mathcal{P}\mathcal{L}\tilde{\mathbf{z}} = \frac{\int\ -\nabla_{\mathbf{z}} V(\mathbf{z}) e^{-\beta\mathcal{H}(\mathbf{z})}\ \delta(\mathbf{z} - \tilde{\mathbf{z}})\ d\mathbf{z}}{\int e^{-\beta\mathcal{H}(\mathbf{z})}\ \delta(\mathbf{z} - \tilde{\mathbf{z}})\ d\mathbf{z}} = -\nabla_{\tilde{\mathbf{z}}} V^{\mathrm{PMF}}(\tilde{\mathbf{z}}) \tag{B.15}$$

where V^{PMF} is known as potential of mean force. Moreover, the memory term can be

written in terms of the noise term as:

$$\int_0^t e^{\mathcal{L}(t-\tau)}\mathcal{P}\mathcal{L}\mathbf{R}(\tau)\ d\tau. \tag{B.16}$$

Ref. [40] has shown that the term in Eq. (B.16) is null for the position coordinates $\mathbf{r}$, while can expressed for the momentum coordinates $\mathbf{p}$ as:

$$-\int_0^t \boldsymbol{\theta}(\tau)\mathbf{p}(t-\tau)d\tau \tag{B.17}$$

B.1.2. GLE integration

The integration of the GLE dynamics is a not-trivial task for two reasons: first, the convolution integral depends on the full history of the observable, and second, the noise term correlations have to be consistent with the fluctuation dissipation theorem. Several approaches based on the introduction of a set of auxiliary variable have been proposed to face these issues, i.e. Refs [48, 181, 182]. In the present work, we take advantage of the exponential structure of $\boldsymbol{\theta}(t)$ to implement an integration algorithm, which, for a scalar memory kernel, reduces to the one proposed in Ref. [182].

Convolution decomposition The history-dependent convolution term is written as a sum of the additional variable vectors $\mathbf{Z}_k(t)$, whose components are $Z_{k,i}(t) = \int_0^t A_{k,ij}e^{B_{k,ij}(t-\tau)}\boldsymbol{\mathcal{O}}_j(\tau)d\tau$. Applying Leibniz's integral rule, and taking advantage of the symmetry of the matrices $\mathbf{B}_k$ it follows that the time evolution of $\mathbf{Z}_k(t)$ can be expressed as:

$$\dot{\mathbf{Z}}_k(t) = \mathbf{B}_k\mathbf{Z}_k(t) - \mathbf{A}_k\boldsymbol{\mathcal{O}}(t) \tag{B.18}$$

Hence, the original GLE is rewritten in form of the equivalent system:

$$\begin{cases} \partial_t\boldsymbol{\mathcal{O}}(t) = \mathcal{P}\mathcal{L}\boldsymbol{\mathcal{O}} - \sum_k \mathbf{Z}_k(t) + \mathbf{R}(t) \\ \partial_t\mathbf{Z}_k(t) = \mathbf{B}_k\mathbf{Z}_k(t) + \mathbf{A}_k\boldsymbol{\mathcal{O}}(t) \end{cases} \tag{B.19}$$

Random force decomposition In this section we provide the theoretical derivation of the random force decomposition for a general real tensor function $\boldsymbol{\theta}(t)$. It is worth noticing that such formulation is valid for any form of the memory kernel, not just exponential ones. First, let us notice that, because of the symmetry between t and t' in the fluctuation dissipation theorem, $\boldsymbol{\theta}(t)$ is an even function of time, i.e. $\boldsymbol{\theta}(t) = \boldsymbol{\theta}(-t)$. We define the Fourier transform of $\boldsymbol{\theta}(t)$ as $\widetilde{\boldsymbol{\theta}}(\omega) = \int_{-\infty}^{-\infty}\boldsymbol{\theta}(t)e^{-i\omega\tau}dt$. Since $\boldsymbol{\theta}(t)$ is real

and even in time, also $\widetilde{\boldsymbol{\theta}}(\omega)$ is real and even for real ω. It follows that both zeros and singular points of $\widetilde{\boldsymbol{\theta}}(\omega)$ are symmetric with respect to both real and imaginary axis in the ω-plane. Then we introduce the function $\widetilde{\boldsymbol{\chi}}(\omega)$ given by

$$\widetilde{\boldsymbol{\chi}}(\omega) = \sum_k -i \left(\omega \mathbf{1} + i\mathbf{B'}_k\right)^{-1} \mathbf{b}_k \tag{B.20}$$

where the real matrices $\mathbf{b}_k$ and $\mathbf{B'}_k$ are such that:

$$\widetilde{\boldsymbol{\theta}}(\omega)\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}} \rangle = 2\widetilde{\boldsymbol{\chi}}(\omega)\widetilde{\boldsymbol{\chi}}^T(-\omega), \tag{B.21}$$

and the singular points of $\widetilde{\boldsymbol{\chi}}^{-1}(\omega)$ lie in the lower-half complex ω-plane. Moreover, we define the two matrices:

$$\widetilde{\boldsymbol{\zeta}}(\omega) = \widetilde{\boldsymbol{\chi}}^{-1}(\omega), \tag{B.22}$$

and

$$\widetilde{\mathbf{k}}_k(\omega) = -i \left(\omega \mathbf{1} + i\mathbf{B'}_k\right)^{-1} \mathbf{b}_k \widetilde{\boldsymbol{\zeta}}(\omega), \tag{B.23}$$

and we denote their Fourier inverse transform with $\mathbf{h(t)}$ and $\mathbf{k}_k(t)$. Combining Eqs (B.20), (B.22) and (B.23), it follows that:

$$\sum_k \widetilde{\mathbf{k}}_k(\omega) = \mathbf{1} \tag{B.24}$$

or, equivalently,

$$\sum_k \mathbf{k}_k(t) = \mathbf{1}\delta(t). \tag{B.25}$$

Moreover Eq. (B.23) can be rewritten as $(i\omega \mathbf{1} - \mathbf{B'}_k)\, \widetilde{\mathbf{k}}_k(\omega) = \mathbf{b}_k \widetilde{\boldsymbol{\zeta}}(\omega)$, that in the time domain gives:

$$\frac{d}{dt}\mathbf{k}_k(t) - \mathbf{B'}_k \mathbf{k}_k(t) = \mathbf{b}_k \boldsymbol{\zeta}(t) \tag{B.26}$$

Finally, the following vector variables are introduced:

$$\boldsymbol{\xi}(t) = \int_0^{+\infty} \boldsymbol{\zeta}(t - t')\mathbf{R}(t')dt' \tag{B.27}$$

and

$$\mathbf{R}_k(t) = \int_0^{+\infty} \mathbf{k}_k(t-t')\mathbf{R}(t')dt'. \tag{B.28}$$

From Eqs (B.28)-(B.25) it follows that:

$$\sum_k \mathbf{R}_k(t) = \mathbf{R}(t), \tag{B.29}$$

while, combining Eqs (B.28)-(B.26)

$$\frac{d}{dt}\mathbf{R}_k(t) = \mathbf{B'}_k\mathbf{R}_k(t) + \mathbf{b}_k\boldsymbol{\xi}(t). \tag{B.30}$$

Eqs (B.30)-(B.29) are the main result of the section since they allow to express the correlated noise of the original GLE as a function of white noises $\boldsymbol{\xi}(t)$.

In what follows, we discuss the properties of the stochastic process $\boldsymbol{\xi}(t)$. First, since all the singularities of $\widetilde{\boldsymbol{\zeta}}(\omega) = \widetilde{\boldsymbol{\chi}}^{-1}(\omega)$ lie in the lower-half complex ω-plane, then for $\tau > 0$:

$$\begin{aligned}\boldsymbol{\zeta}(\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\widetilde{\boldsymbol{\zeta}}(\omega)e^{i\omega\tau} = \lim_{a\to\infty}\frac{1}{2\pi}\int_{-a}^{a} d\omega\widetilde{\boldsymbol{\zeta}}(\omega)e^{i\omega\tau} = \\ &= \frac{1}{2\pi}\oint_{C^+} d\omega\widetilde{\boldsymbol{\zeta}}(\omega)e^{i\omega\tau} - \lim_{a\to\infty}\frac{1}{2\pi}\int_{\mathrm{arc}(a\to-a)} d\omega\widetilde{\boldsymbol{\zeta}}(\omega)e^{i\omega\tau} = \mathbf{0}\end{aligned} \tag{B.31}$$

where $\oint_{C^+} d\omega$ indicates the integral over a closed contour C^+ that goes along the real line from $-a$ to a and then along a semicircle centred at 0 from a to $-a$, while $\int_{\mathrm{arc}(a\to-a)} d\omega$ is the integral along an arc centred at 0 from a to $-a$. Hence, for $t > 0$ we can write

$$\boldsymbol{\xi}(t) = \int_0^{+\infty} \boldsymbol{\zeta}(t-t')\mathbf{R}(t')dt' = \int_{-\infty}^{+\infty} \boldsymbol{\zeta}(t-t')\mathbf{R}(t')dt' \tag{B.32}$$

Thus, the correlation function of $\boldsymbol{\xi}(t)$ at t_1 and t_2 is given by:

$$\begin{aligned}
&\left\langle \boldsymbol{\xi}(t_1)\boldsymbol{\xi}^T(t_2)\right\rangle = \\
&= \left\langle \int_{-\infty}^{+\infty} dt_1' \int_{-\infty}^{+\infty} dt_2' \boldsymbol{\zeta}(t_1 - t_1')\mathbf{R}(t_1')\left[\boldsymbol{\zeta}(t_2 - t_2')\mathbf{R}(t_2')\right]^T \right\rangle = \\
&= \int_{-\infty}^{+\infty} dt_1' \int_{-\infty}^{+\infty} dt_2' \boldsymbol{\zeta}(t_1 - t_1')\left\langle \mathbf{R}(t_1')\mathbf{R}^T(t_2')\right\rangle \boldsymbol{\zeta}^T(t_2 - t_2') = \\
&= \int_{-\infty}^{+\infty} dt_1' \int_{-\infty}^{+\infty} dt_2' \boldsymbol{\zeta}(t_1 - t_1')\boldsymbol{\theta}(t_1' - t_2')\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}}\rangle \boldsymbol{\zeta}^T(t_2 - t_2')
\end{aligned} \tag{B.33}$$

where we used the fluctuation dissipation theorem. Using the definition of Fourier transform of $\boldsymbol{\theta}$, it follows

$$\begin{aligned}
&\left\langle \boldsymbol{\xi}(t_1)\boldsymbol{\xi}^T(t_2)\right\rangle = \\
&= \int_{-\infty}^{+\infty} dt_1' \int_{-\infty}^{+\infty} dt_2' \boldsymbol{\zeta}(t_1 - t_1')\frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \widetilde{\boldsymbol{\theta}}(\omega) e^{i\omega(t_1' - t_2')} \\
&\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}}\rangle \boldsymbol{\zeta}^T(t_2 - t_2') = \\
&= \int_{-\infty}^{+\infty} dt_1' \int_{-\infty}^{+\infty} dt_2' \boldsymbol{\zeta}(t_1 - t_1')\frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \widetilde{\boldsymbol{\theta}}(\omega) e^{i\omega(t_1' - t_2')} \\
&\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}}\rangle \boldsymbol{\zeta}^T(t_2 - t_2') e^{-i\omega(t_1 - t_2)} e^{i\omega(t_1 - t_2)} = \\
&= \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \int_{-\infty}^{+\infty} dt_1' \boldsymbol{\zeta}(t_1 - t_1') e^{-i\omega(t_1 - t_1')}\widetilde{\boldsymbol{\theta}}(\omega)\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}}\rangle \\
&\int_{-\infty}^{+\infty} dt_2' \boldsymbol{\zeta}^T(t_2 - t_2') e^{i\omega(t_2' - t_2')} e^{i\omega(t_1 - t_2)}
\end{aligned} \tag{B.34}$$

Applying the definition of Fourier transform of $\boldsymbol{\zeta}(t)$, and taking advantage of Eqs (B.21)-(B.22), we finally obtain:

$$\left\langle \boldsymbol{\xi}(t_1)\boldsymbol{\xi}(t_2)^T\right\rangle = \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \widetilde{\boldsymbol{\zeta}}(\omega)\widetilde{\boldsymbol{\theta}}(\omega)\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}}\rangle \widetilde{\boldsymbol{\zeta}}^T(-\omega) e^{i\omega(t_1 - t_2)} = \tag{B.35}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega \widetilde{\boldsymbol{\zeta}}(\omega) 2\widetilde{\boldsymbol{\chi}}(\omega)\widetilde{\boldsymbol{\chi}}(-\omega)^T \widetilde{\boldsymbol{\zeta}}^T(-\omega) e^{i\omega(t_1 - t_2)} = \tag{B.36}$$

$$= \frac{1}{2\pi}\int_{-\infty}^{+\infty} d\omega 2\mathbf{1} e^{i\omega(t_1 - t_2)} = 2\delta(t_1 - t_2). \tag{B.37}$$

It follows that $\boldsymbol{\xi}(t)$ is a delta correlated stochastic process.

Since we adopted a approximation of $\boldsymbol{\theta}(t)$ whose components are in the exponential

form $\theta_{i,j}(t) = \sum_k A_{k,ij} e^{B_{k,ij}(t)}$, its Fourier transform is given by:

$$\widetilde{\boldsymbol{\theta}}(\omega) = \sum_k \left[-i\mathbf{A}_k \oslash (\omega\mathbf{J} + i\mathbf{B}_k) + i\mathbf{A}_k \oslash (\omega\mathbf{J} - i\mathbf{B}_k)\right] \tag{B.38}$$

where $\oslash$ indicates the Hadamard division and $\mathbf{J}$ is the $n \times n$ matrix of ones. Now $\boldsymbol{\theta}(t)$ is a real and even function of t, then $\widetilde{\boldsymbol{\theta}}(\omega)$ has to be real and even for real values of ω. As a consequence, the singular points of $\widetilde{\boldsymbol{\theta}}(\omega)$ has to be symmetrical with respect to the real and imaginary axes, namely in the form of pairs $\pm i\mathbf{B}_k$. For the same reason, the roots of $\widetilde{\boldsymbol{\theta}}(\omega)$ have to be symmetric with respect to the real and imaginary axes. Thus, putting Eq. (B.38) into a common denominator, factorizing, and calling $\boldsymbol{\beta}_n$ and $\boldsymbol{\beta}_n^*$ the conjugate matrices containing the zeros of the numerator, we obtain:

$$\widetilde{\boldsymbol{\theta}}(\omega) = \mathbf{K} \odot \left(\prod_n (\omega\mathbf{J} - \boldsymbol{\beta}_n) \odot (\omega\mathbf{J} - \boldsymbol{\beta}_n^*)\right) \oslash \tag{B.39}$$

$$\left(\prod_k (\omega\mathbf{J} + i\mathbf{B}_k) \odot (\omega\mathbf{J} - i\mathbf{B}_k)\right) \tag{B.40}$$

where $\odot$ is the Hadamard product, $\mathbf{K}$ is a matrix of positive real numbers and it is assumed that $Im(\beta_n) > 0$ and $Im(\beta_n^*) < 0$. It is worth noticing that, since $\widetilde{\boldsymbol{\theta}}(\omega)$ is non-negative, then $\mathbf{K}$ contains only positive values [182]. Now we define the function $\widetilde{\boldsymbol{\chi}}(\omega)$ as:

$$\begin{aligned}\widetilde{\boldsymbol{\chi}}(\omega) = \frac{\mathbf{K}^{1/2}}{\sqrt{2}} \odot \langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}} \rangle^{\odot 1/2} \odot \prod_n i\left(\omega\mathbf{J} - \boldsymbol{\beta}_n^*\right) \oslash \prod_k i\left(\omega\mathbf{J} + i\mathbf{B}_k\right) = \\ = \sum_k -i\left(\omega\mathbf{1} + i\mathbf{B}_k\right)^{-1} \mathbf{b}_k\end{aligned} \tag{B.41}$$

Eq. (B.41) has to be solved to find the matrices $\mathbf{b}_k$.

In case of diagonal memory kernel matrix, $\mathbf{b}_k$ can be easily found by solving the easier relation:

$$\begin{aligned}\frac{\mathbf{K}^{1/2}}{\sqrt{2}} \langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}} \rangle^{1/2} \prod_n i\left(\omega\mathbf{1} - \boldsymbol{\beta}_n^*\right) \prod_k i\left(\omega\mathbf{1} + i\mathbf{B}_k\right)^{-1} = \\ = \sum_k -i\left(\omega\mathbf{1} + i\mathbf{B}_k\right)^{-1} \mathbf{b}_k\end{aligned} \tag{B.42}$$

Finally, for a one dimensional GLE the presented formulation reduces to the one derived by Kawai [182], therefore the coefficients b_k can be evaluated from the following

Table B.1. Results of the augmented Dickey–Fuller(ADF) test for modified global temperature.

ADF Statistic:	-21.377945
p-value:	$< 10^{-16}$
lags:	54
Critical Values:	
1%:	-2.566
5%:	-1.941
10%:	-1.617

relation:

$$\langle \boldsymbol{\mathcal{O}}, \boldsymbol{\mathcal{O}} \rangle A_k = -2b_k \sum_n \frac{b_n}{B_k + B_n} \tag{B.43}$$

obtained by Eqs (B.20), (B.21) and (B.38).

Extended dynamics and integration algorithm For a general $\boldsymbol{\theta}(t)$, the extended dynamics is then expressed as:

$$\begin{cases} \partial_t \boldsymbol{\mathcal{O}}(t) = \mathcal{P}\mathcal{L}\boldsymbol{\mathcal{O}} - \int_0^t \boldsymbol{\theta}(\tau)\boldsymbol{\mathcal{O}}(t-\tau)d\tau + \sum_k \mathbf{R}_k(t) \\ \partial_t \mathbf{R}_k(t) = \mathbf{B'}_k \mathbf{R}_k(t) + \mathbf{b}_k \boldsymbol{\xi}(t) \end{cases} \tag{B.44}$$

where the convolution can be decomposed in different ways depending on the structure of $\boldsymbol{\theta}(t)$. In our case $\boldsymbol{\theta}(t)$ has an exponential form, thus $\mathbf{B'}_k = \mathbf{B}_k$ and the variables $\mathbf{S}_k(t) = -\mathbf{Z}_k(t) + \mathbf{R}_k(t)$ can be defined, so that the GLE can is rewritten in following form:

$$\begin{cases} \partial_t \boldsymbol{\mathcal{O}}(t)\boldsymbol{\mathcal{O}} = \mathbf{F}(\boldsymbol{\mathcal{O}}(t)) + \sum_{k=1}^{N_n} \mathbf{S}_k(t) \\ \partial_t \mathbf{S}_k(t) = \mathbf{B}_k \mathbf{S}_k(t) - \mathbf{A}_k \boldsymbol{\mathcal{O}}(t) + \mathbf{b}_k \boldsymbol{\xi}(t), \end{cases} \tag{B.45}$$

with $\mathbf{F}(\boldsymbol{\mathcal{O}}(t)) = \mathcal{P}\mathcal{L}\boldsymbol{\mathcal{O}}$ accounting for the conservative mean force contributions.

B.1.3. Modelling global temperature: Additional results

In order to test the statistical properties ot T_a, we adopted some qualitative and quantitative tests. Figure B.1(a) shows the QQ (quantile-quantile) plot, which compares the data distribution against the normal one for each quantile. It emerges that the time series data are well approximated by a normal distribution, especially in the theoretical

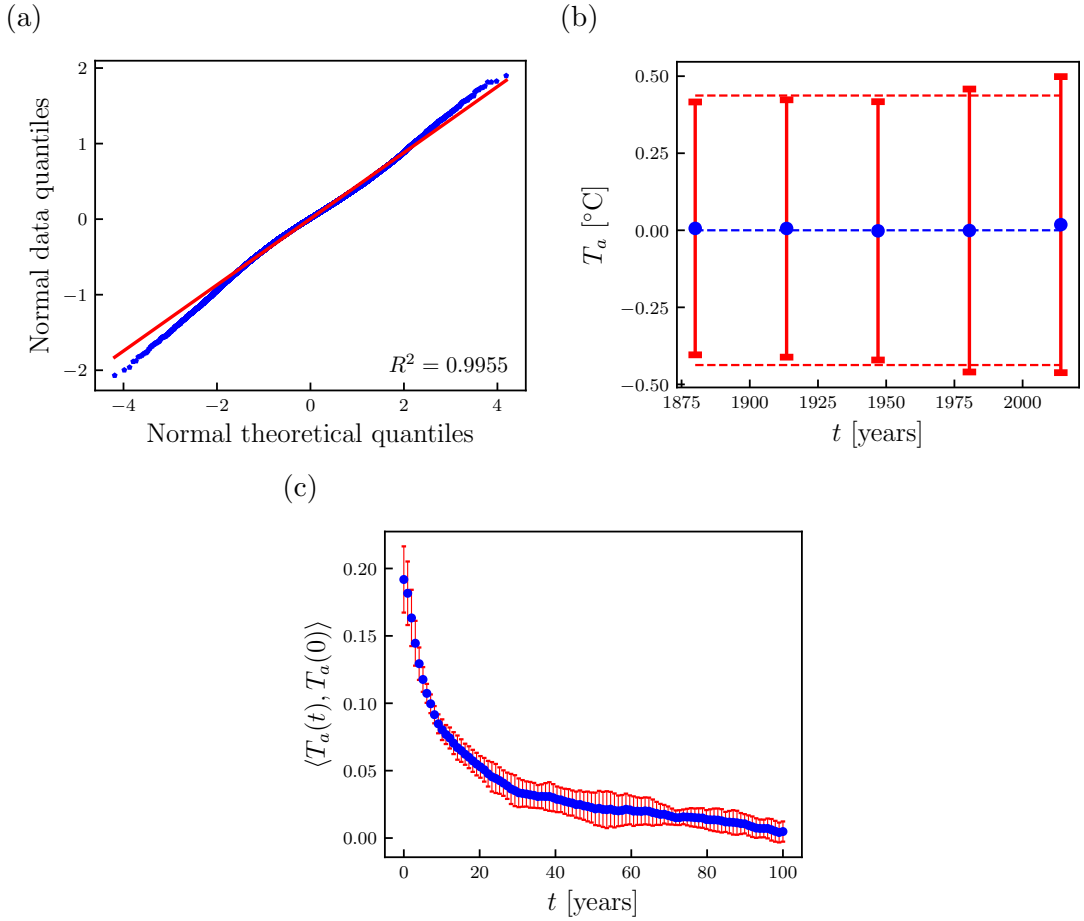

Figure B.1. (a) QQ-plot for $T_a(t) = T(t) - T_y(t)$. (b) Mean evaluated for 5 different data windows in time (blue dots) and corresponding standard deviations represented as red error bars. (c) Average time correlation function (blue dots) and standard error evaluated at each time from the time correlations of 5 different data windows.

quantile range $-3 < Q < 3$. Some tail effects are visible, but the overall agreement is quantitatively verified by the R-squared test, which gives a value $R^2 = 0.9955$.

To test the stationarity of mean variance, and time correlations, we split the data in 5 windows. Figure B.1(b) shows that, assuming the stationarity of the series, we commit maximum errors for mean and standard deviation of 0.0183 and 0.0430, respectively. Moreover, as reported in Fig. B.1(c), the maximum standard error between the windows time correlation and their mean is 0.0246.

Additionally, to test the stationarity of the modified time series, also the augmented Dickey–Fuller (ADF) test is adopted [243]. ADF test is useful to establish if a unit root

Table B.2. Results of the augmented Dickey–Fuller(ADF) test for modified Nikkei index.

ADF Statistic:	-29.805726
p-value:	$< 10^{-16}$
lags:	10
Critical Values:	
1%:	-2.566
5%:	-1.941
10%:	-1.617

is present in the stochastic data series. Specifically, the null hypothesis of a unit root is rejected in favour of the stationary alternative if the test statistic is more negative than some critical values. The results of the ADF test reported in Table B.1 allows us to reject the unit root hypothesis with a probability higher than 99%.

B.1.4. Modelling Nikkei index: Additional results

Figure B.2(a) shows the QQ (quantile-quantile) plot. The time series distribution is well approximated with a normal distribution in the theoretical quantile range $-2.5 < Q < 2.5$. Heavy tails effects are present. This means that the Gaussian approximation, and consequently the GLE for $NI(t)$, remains valid as long as extreme market events, such as market crashes or crisis, are not considered. The overall agreement is quantitatively verified by the R-squared test, which gives $R^2 = 0.9894$.

In order to test the stationarity of mean variance, and time correlations, we split the data in 5 equally sized sets and, for each one, we analyse the statistical properties. Figure B.2(b) shows that, assuming the stationarity of the series, we commit maximum errors for mean and standard deviation of 0.2787 and 0.0234, respectively. Moreover, as reported in Fig. B.2(c), the maximum standard error between the windows time correlation and their mean is 0.1082. The results of the ADF test reported in Table B.2 allows us to reject the unit root hypothesis with a probability higher than 99%.

B.2. Non-Markovian FDDFT

B.2.1. Derivations of noise in FDDFT

The noise term in Eq. (3.31) is $\boldsymbol{\eta}_s(\boldsymbol{r};t) = \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r} - \boldsymbol{r}_{i,s})\boldsymbol{R}_{i,s}(t)$, where $\boldsymbol{R}_{i,s}(t)$ has zero mean and correlation function $\langle \boldsymbol{R}_{i,s}(t)\boldsymbol{R}_{i,s}(t')\rangle = k_{\mathrm{B}}T\boldsymbol{\theta}_s(t-t')$. It follows that the

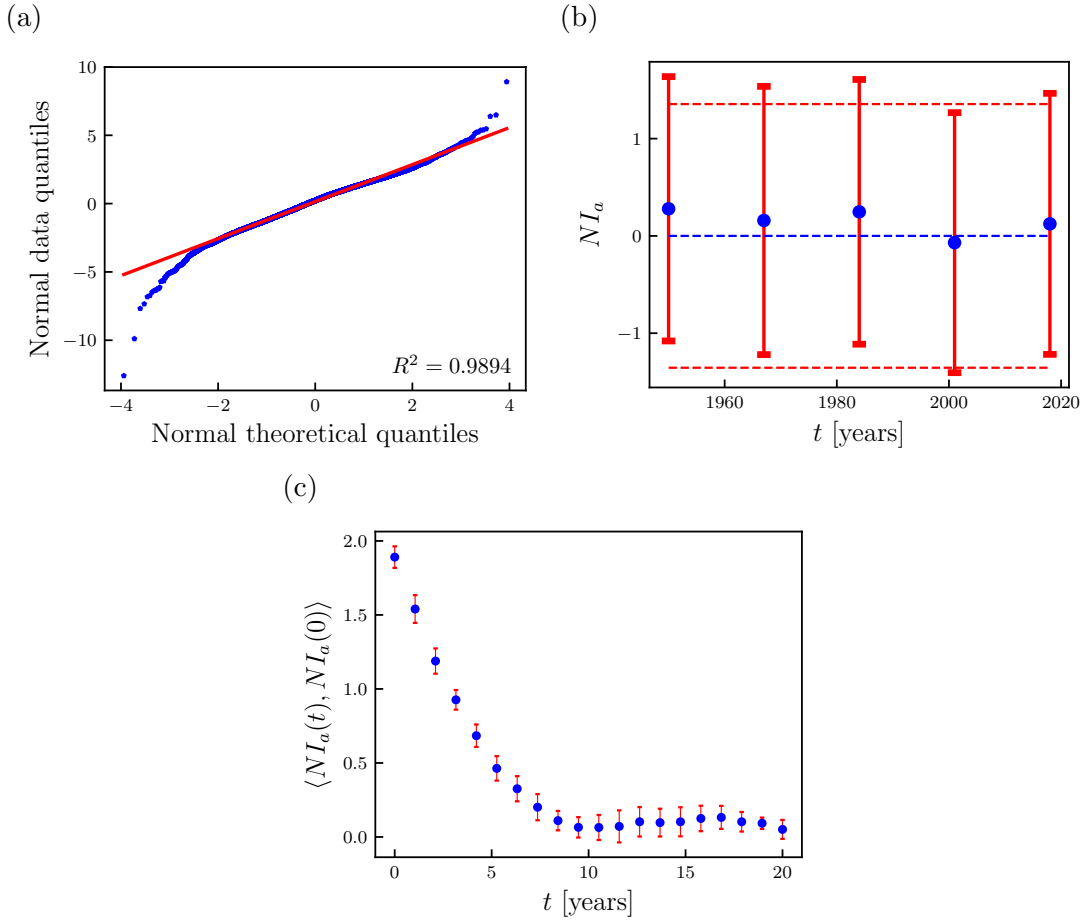

Figure B.2. (a) QQ-plot for $NI_a(t) = NI(t) - NI_y(t)$. (b) Mean evaluated for 5 different data windows in time (blue dots) and corresponding standard deviations represented as red error bars. (c) Average time correlation function (blue dots) and standard error evaluated at each time from the time correlations of 5 different data windows.

spatio-temporal correlations of $\boldsymbol{\eta}_s$ can be expressed as:

$$\langle \boldsymbol{\eta}_s(\boldsymbol{r};t)\boldsymbol{\eta}_s(\boldsymbol{r}';t')\rangle = \left\langle \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r}-\boldsymbol{r}_{i,s}) \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r}'-\boldsymbol{r}_{i,s})\boldsymbol{R}_{i,s}(t)\boldsymbol{R}_{i,s}(t') \right\rangle . \tag{B.46}$$

Using the property of the Dirac delta function

$$\sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r}-\boldsymbol{r}_{i,s}) \sum_{i=1}^{N_s(t')} \delta(\boldsymbol{r}'-\boldsymbol{r}_{i,s}) = \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r}-\boldsymbol{r}_{i,s})\delta(\boldsymbol{r}-\boldsymbol{r}') = \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r}'-\boldsymbol{r}_{i,s})\delta(\boldsymbol{r}-\boldsymbol{r}'), \tag{B.47}$$

the spatio-temporal correlations of η_s can be rewritten as:

$$\begin{aligned}\langle \boldsymbol{\eta}_s(\boldsymbol{r};t)\boldsymbol{\eta}_s(\boldsymbol{r}';t')\rangle &= \left\langle \sum_{i=1}^{N_s(t)} \delta(\boldsymbol{r}-\boldsymbol{r}_{i,s})\delta(\boldsymbol{r}-\boldsymbol{r}')\boldsymbol{R}_{i,s}(t)\boldsymbol{R}_{i,s}(t') \right\rangle \\ &= \left\langle \rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')k_{\mathrm{B}}T\boldsymbol{\theta}_s(t-t') \right\rangle .\end{aligned} \tag{B.48}$$

Thus, introducing a spatio-temporal noise satisfying $\langle \boldsymbol{R}_s(\boldsymbol{r};t)\boldsymbol{R}_s(\boldsymbol{r}';t')\rangle = k_{\mathrm{B}}T\boldsymbol{\theta}_s(t-t')\delta(\boldsymbol{r}-\boldsymbol{r}')$, the local fluctuation term can be rewritten as

$$\boldsymbol{\eta}_s(\boldsymbol{r};t) = \sqrt{\rho_s(\boldsymbol{r};t)}\boldsymbol{R}_s(\boldsymbol{r};t)\,, \tag{B.49}$$

B.2.2. Extended field dynamics in Non-Markovian FDDFT

In this section, we discuss the properties of the stochastic process $\boldsymbol{\xi}(\boldsymbol{r};t)$. First, since all the singularities of $\widetilde{\zeta}(\omega) = \widetilde{\chi}^{-1}(\omega)$ lie in the lower-half complex ω-plane, then for $\tau > 0$:

$$\begin{aligned}\zeta(\tau) &= \frac{1}{2\pi}\int_{-\infty}^{\infty} d\omega\widetilde{\zeta}(\omega)e^{i\omega\tau} = \lim_{a\to\infty}\frac{1}{2\pi}\int_{-a}^{a} d\omega\widetilde{\zeta}(\omega)e^{i\omega\tau} = \\ &= \frac{1}{2\pi}\oint_{C^+} d\omega\widetilde{\zeta}(\omega)e^{i\omega\tau} - \lim_{a\to\infty}\frac{1}{2\pi}\int_{\mathrm{arc}(a\to-a)} d\omega\widetilde{\zeta}(\omega)e^{i\omega\tau} = 0\end{aligned} \tag{B.50}$$

where $\oint_{C^+} d\omega$ indicates the integral over a closed contour C^+ that goes along the real line from $-a$ to a and then along a semicircle centred at 0 from a to $-a$, while $\int_{\mathrm{arc}(a\to-a)} d\omega$ is the integral along an arc centred at 0 from a to $-a$. Hence, for $t > 0$ we can write

$$\boldsymbol{\xi}_s(\boldsymbol{r};t) = \int_0^{+\infty} \zeta(t-t')\boldsymbol{\eta}_s(t')dt' = \int_{-\infty}^{+\infty} \zeta(t-t')\boldsymbol{\eta}_s(t')dt' \tag{B.51}$$

Thus, the spatio-temporal correlation function of $\boldsymbol{\xi}_s(\boldsymbol{r};t)$ is given by:

$$\begin{aligned}&\left\langle \boldsymbol{\xi}_s(\boldsymbol{r};t_1)\boldsymbol{\xi}_s(\boldsymbol{r}';t_2)\right\rangle = \\ &= \left\langle \int_{-\infty}^{+\infty} dt_1' \int_{-\infty}^{+\infty} dt_2'\zeta(t_1-t_1')\boldsymbol{\eta}_s(\boldsymbol{r};t_1')\left[\zeta(t_2-t_2')\boldsymbol{\eta}_s(\boldsymbol{r}';t_2')\right]^T \right\rangle = \\ &= \int_{-\infty}^{+\infty} dt_1' \int_{-\infty}^{+\infty} dt_2'\zeta(t_1-t_1')\left\langle \boldsymbol{\eta}_s(\boldsymbol{r};t_1')\boldsymbol{\eta}_s^T(\boldsymbol{r}';t_2')\right\rangle \zeta(t_2-t_2')\end{aligned} \tag{B.52}$$

where we used the fluctuation dissipation theorem.

Employing the definition of Fourier transform of θ_s and $\zeta(t)$, together with Eqs (3.65)-

(3.67), we obtain:

$$
\begin{aligned}
&\left\langle \boldsymbol{\xi}_s(\boldsymbol{r};t_1)\boldsymbol{\xi}_s(\boldsymbol{r}';t_2)\right\rangle = \\
&= k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}\int_{-\infty}^{+\infty}dt_1'\int_{-\infty}^{+\infty}dt_2'\zeta(t_1-t_1')\zeta(t_2-t_2')\theta_s(t_2'-t_1') = \\
&= k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}\int_{-\infty}^{+\infty}dt_1'\int_{-\infty}^{+\infty}dt_2'\zeta(t_1-t_1')\frac{1}{2\pi}\times \\
&\times\int_{-\infty}^{+\infty}d\omega\widetilde{\theta}_s(\omega)e^{i\omega(t_1'-t_2')}\zeta(t_2-t_2') = \\
&= k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}\int_{-\infty}^{+\infty}dt_1'\int_{-\infty}^{+\infty}dt_2'\zeta(t_1-t_1')\frac{1}{2\pi}\times \\
&\times\int_{-\infty}^{+\infty}d\omega\widetilde{\theta}_s(\omega)e^{i\omega(t_1'-t_2')}\zeta(t_2-t_2')e^{-i\omega(t_1-t_2)}e^{i\omega(t_1-t_2)} = \\
&= k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\omega\int_{-\infty}^{+\infty}dt_1'\zeta(t_1-t_1')e^{-i\omega(t_1-t_1')}\widetilde{\theta}_s(\omega)\times \\
&\times\int_{-\infty}^{+\infty}dt_2'\zeta(t_2-t_2')e^{i\omega(t_2'-t_2')}e^{i\omega(t_1-t_2)} = \\
&= k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\omega\widetilde{\zeta}(\omega)\widetilde{\theta}_s(\omega)\widetilde{\zeta}(-\omega)e^{i\omega(t_1-t_2)} = \\
&= k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\omega\widetilde{\zeta}(\omega)2\widetilde{\chi}(\omega)\widetilde{\chi}(-\omega)\widetilde{\zeta}(-\omega)e^{i\omega(t_1-t_2)} = \\
&= k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\mathbf{I}\frac{1}{2\pi}\int_{-\infty}^{+\infty}d\omega 2e^{i\omega(t_1-t_2)} \\
&= 2k_{\mathrm{B}}T\rho_s(\boldsymbol{r};t)\delta(\boldsymbol{r}-\boldsymbol{r}')\delta(t_1-t_2)\mathbf{I}.
\end{aligned}
\tag{B.53}
$$

It follows that $\boldsymbol{\xi}_s(\boldsymbol{r};t)$ is delta correlated in space and time, and consequently can be easily generated, i.e. as in Ref. [208].

B.2.3. Linear stability analysis for reaction-diffusion systems

Turing patterns are observed when an uniform system is stable in absence of diffusion, but become unstable to perturbation with diffusion. For the analysis, we consider the 1D density equation:

$$\partial_t\rho_1(x;t) = d_1\partial_x^2\rho_1(x;t) + w_{\rho_1}(x;t) \tag{B.54}$$

$$\partial_t\rho_2(x;t) = d_2\partial_x^2\rho_1(x;t) + w_{\rho_2}(x;t) \tag{B.55}$$

with stationary states given by:

$$w_{\rho_1}(\boldsymbol{r};t) = 0 \tag{B.56}$$
$$w_{\rho_2}(\boldsymbol{r};t) = 0. \tag{B.57}$$

Hence, we study the linear stability of the system around the stationary state for both the corresponding reacting-only system and for the full system. The time evolution of the infinitesimal perturbations $\delta\rho_1$ and $\delta\rho_2$ is given by:

$$\partial_t \delta\rho_1 = d_1 \partial_x^2 \delta\rho_1 + J_{11}\delta\rho_1 + J_{12}\delta\rho_2 \tag{B.58}$$
$$\partial_t \delta\rho_2 = d_2 \partial_x^2 \delta\rho_1 + J_{21}\delta\rho_1 + J_{22}\delta\rho_2 \tag{B.59}$$

where the coefficients of the Jacobian matrix evaluated at the stationary state $J_{i,j} = \frac{\partial w_{\rho_i}}{\partial \rho_j}\Big|_{\rho_i^b}$ are introduced. A particular solution is given by

$$\begin{pmatrix} \delta\rho_1 \\ \delta\rho_2 \end{pmatrix} = \begin{pmatrix} \delta\rho_1^0 \\ \delta\rho_2^0 \end{pmatrix} e^{\lambda_k t} e^{ikx} \tag{B.60}$$

Combining Eq. (B.59) and Eq. (B.60), we obtain the eigenvalue problem:

$$(\boldsymbol{J}_k - k\boldsymbol{I}) \begin{pmatrix} \delta\rho_1 \\ \delta\rho_2 \end{pmatrix} = \boldsymbol{0} \tag{B.61}$$

where

$$\boldsymbol{J}_k = \begin{pmatrix} J_{1,1} - d_1 k^2 & J_{1,2j} \\ J_{2,1} & J_{2,2} - d_2 k^2 \end{pmatrix} \tag{B.62}$$

The uniform stationary solution is stable if both eigenvalues $\lambda_{k,i}$ at a wave number k have negative real part. Such a condition is verified if:

$$\mathrm{Tr} J_k = J_{11} + J_{22} - (d_1 + d_2)k^2 < 0 \tag{B.63}$$
$$\mathrm{Det} J_k = (J_{11} - d_1 k^2)(J_{22} - d_2 k^2) - J_{12}J_{21} > 0 \tag{B.64}$$

Turing patterns arises when a reaction-only system, corresponding to $d_1 = d_2 = 0$, is stable, but the corresponding reaction-diffusion system is unstable. The reaction-only

system is stable when the following two conditions are satisfied:

$$\mathrm{Tr}J = J_{11} + J_{22} < 0 \tag{B.65}$$

$$\mathrm{Det}J = J_{11}J_{22} - J_{12}J_{21} > 0 \tag{B.66}$$

The reaction-diffusion system is unstable when at least one of Eqs (B.63)-(B.64) is not verified. Because of Eq. (B.65), $\mathrm{Tr}J_k < \mathrm{Tr}J < 0$, thus Eq. (B.63) is always verified. Thus, instability occurs if and only if $\mathrm{Det}J_k < 0$ for at least a value of k. The minimum value of k for which such a condition occurs is given by:

$$k_m = \sqrt{\frac{d_1 J_{22} + d_2 J_{11}}{2 d_1 d_2}} \tag{B.67}$$

which corresponds to a wave number $L_k = 2\pi/k$. It directly follows that the corresponding $\mathrm{Det}J_{km} < 0$ when:

$$\mathrm{Det}J_{km} = J_{11}J_{22} - J_{12}J_{21} - \frac{d_1 J_{22} + d_2 J_{11}}{4 d_1 d_2} < 0 \tag{B.68}$$

B.2.4. Structure factor

As shown in previous works [71, 208], the structure factor is useful not only to study the stability of the numerical integrator, but also in multi-phase systems to obtain the frequency components characterizing the system. If we consider a periodic domain of volume V, the spatial Fourier transform of the number density for the species s is given by

$$\hat{\rho}_{s,\mathbf{k}}(t) = \frac{1}{V}\int_V \rho_s(\mathbf{r}, t) e^{-i\mathbf{k}\cdot\mathbf{r}} d\mathbf{r}. \tag{B.69}$$

Thus, the structure factor is defined as the variance of the Fourier transform of the density fluctuations of species s,

$$S(\mathbf{k}) = V\langle \delta\hat{\rho}_{s,\mathbf{k}} \delta\hat{\rho}^*_{s,\mathbf{k}} \rangle \tag{B.70}$$

with the Fourier transform of the density fluctuations being $\delta\hat{\rho}_{s,\mathbf{k}} = \hat{\rho}_{s,\mathbf{k}} - \langle\hat{\rho}_{s,\mathbf{k}}\rangle$. In case of uniform ideal gases, the structure factor is computed as [71]:

$$S(\mathbf{k}) = \bar{\rho}_s \tag{B.71}$$

Chapter C

Chapter 4: Additional details and results

C.1. Overdamped FDDFT

C.1.1. Time integrators

Milstein The multiple integrals in Milstein's integrator do not have a simple analytical solutions, thus are approximated as function of the white noise cells average in Eq. (4.31) as in Ref. [220]:

$$I_{l,m}(t) = \begin{cases} \frac{1}{2}\left[\left(\overline{W}^l\right)^2 - 1\right]\Delta t & \text{if} \quad l = m \\ \frac{1}{2}\overline{W}^l\overline{W}^m\Delta t + \sqrt{k_p\ \Delta t}(\varphi_l\overline{W}^m - \varphi_m\overline{W}^l) & \\ +\frac{1}{2\pi}\sum_{r=1}^{p}\frac{1}{r}\left[\zeta_{lr}(\sqrt{2}\ \overline{W}^m\sqrt{\Delta t} + \eta_m) - \zeta_{mr}(\sqrt{2}\ \overline{W}^l\sqrt{\Delta t} + \eta_l)\right] & \text{otherwise} \end{cases} \tag{C.1}$$

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

$$k_p = \frac{1}{12} - \frac{1}{2\pi^2}\sum_{1}^{p}\frac{1}{r^2}. \tag{C.2}$$

The value p determines the accuracy of the multiple stochastic integral approximation, and then of the scheme. A value of $p = k/\Delta t$ for some constant k is enough to preserve the accuracy of the scheme [216].

Runge-Kutta The explicit weak order 2.0 Runge-Kutta scheme reads

$$\begin{aligned}\widetilde{\rho}_j(t+\Delta t) &= \overline{\rho}_j(t) + \frac{1}{2}\left[\mu_j(\boldsymbol{\Upsilon}(t)) + \mu_j(\overline{\boldsymbol{\rho}}(t))\right]\Delta t + \Phi(t)\\ \overline{\rho}_j(t+\Delta t) &= \overline{\rho}_j(t) + \frac{1}{2}\left[\mu_j(\widetilde{\boldsymbol{\rho}}(t+\Delta t)) + \mu_j(\overline{\boldsymbol{\rho}}(t))\right]\Delta t + \Phi(t)\end{aligned} \tag{C.3}$$

where the vector $\Phi(t)$ has components

$$\begin{aligned}\Phi_j(t) &= \frac{1}{4}\sum_{l=1}^{n}\left[\boldsymbol{\sigma}_{lj}(\boldsymbol{\Lambda}_{l+}(t)) + \boldsymbol{\sigma}_{lj}(\boldsymbol{\Lambda}_{l-}(t)) + 2\boldsymbol{\sigma}_{lj}(\overline{\boldsymbol{\rho}}(t))\right]\overline{W}^{l}(t)\sqrt{\Delta t}\\ &+ \frac{1}{4}\sum_{l=1}^{n}\sum_{r=1,r\neq l}^{n}\left[\boldsymbol{\sigma}_{lj}(\boldsymbol{\Xi}_{r+}(t)) + \boldsymbol{\sigma}_{lj}(\boldsymbol{\Lambda}_{r-}(t)) - 2\boldsymbol{\sigma}_{lj}(\overline{\boldsymbol{\rho}}(t))\right]\overline{W}^{l}(t)\\ &+ \frac{1}{4}\sum_{l=1}^{n}\left[\boldsymbol{\sigma}_{lj}(\boldsymbol{\Lambda}_{l+}(t)) - \boldsymbol{\sigma}_{lj}(\boldsymbol{\Lambda}_{l-}(t))\right]\left[\left(\overline{W}^{l}(t)\right)^2 - 1\right]\sqrt{\Delta t}\\ &+ \frac{1}{4}\sum_{l=1}^{n}\sum_{r=1,r\neq l}^{n}\left[\boldsymbol{\sigma}_{lj}(\boldsymbol{\Xi}_{r+}(t)) - \boldsymbol{\sigma}_{lj}(\boldsymbol{\Xi}_{r-}(t))\right]\left[\overline{W}^{l}(t)\overline{W}^{r}(t) + V_{r,j}\right]\sqrt{\Delta t}\end{aligned} \tag{C.4}$$

and supporting values

$$\boldsymbol{\Upsilon} = \overline{\boldsymbol{\rho}}(t) + \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))\Delta t + \sum_{j=1}^{n}\boldsymbol{\sigma}_j(\boldsymbol{\rho}(t))\Delta W_j(t) \tag{C.5}$$

$$\boldsymbol{\Lambda}_{l\pm} = \overline{\boldsymbol{\rho}}(t) + \boldsymbol{\mu}(\overline{\boldsymbol{\rho}}(t))\Delta t \pm \boldsymbol{\sigma}_l(\overline{\boldsymbol{\rho}}(t))\sqrt{\Delta t} \tag{C.6}$$

$$\boldsymbol{\Xi}_{l\pm} = \overline{\boldsymbol{\rho}}(t) \pm \boldsymbol{\sigma}_l(\overline{\boldsymbol{\rho}}(t))\sqrt{\Delta t} \tag{C.7}$$

$$\tag{C.8}$$

The random matrix $\mathbf{V}$ is defined as:

$$V_{r,j}(t) = \begin{cases} \pm 1 \quad \text{with} \quad p = \frac{1}{2} \quad \text{if} \quad r < j \\ -1 \quad \text{if} \quad r = j \\ -V_{j,r}(t) \quad \text{if} \quad r > j \end{cases} \tag{C.9}$$

where p indicates the probability.

C.1.2. MD simulations details

MD simulations are performed using the Large-Scale Atomic/Molecular Massively Parallel Simulator (LAMMPS). 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 [156], a macroscopic field $X(\mathbf{r}, t)$ can be extracted from particle coordinates as $X(\mathbf{r}, t) = \sum_i \chi_i \phi(\mathbf{r}_i(t) - \mathbf{r})$, where χ_i is the information of interest of 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:

$$\phi(y) = \begin{cases} \frac{1}{\Delta x} & \text{for } \|y\| < \Delta x/2, \\ 0 & \text{otherwise }, \end{cases} \tag{C.10}$$

with Δx being the width of each bin. In each comparison, we match the Δx for MD simulations with the one for the discretized FDDFT. Using the above, the instantaneous macroscopic density profile for a single trajectory is computed as:

$$\rho(x, t) = \sum_i m_i \phi(x_i(t) - x), \tag{C.11}$$

where m_i is the mass of the particle i.

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.

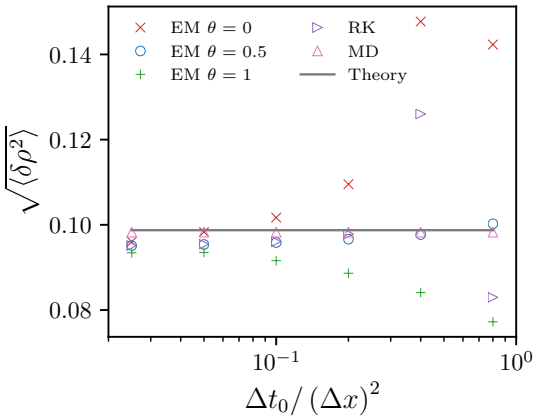

Figure C.1. Standard deviation $\sqrt{\langle \delta\rho^2 \rangle}$ as a function of the initial time-step Δt_0 for an ideal gas in equilibrium. Temporal integrators. EM: Euler-Maruyama, MI: Milstein, RK: Runge-Kutta, MD: Molecular dynamics. Explicit ($\theta = 0$), semi-implicit ($\theta = 0.5$) and implicit ($\theta = 1$), MD: Molecular dynamics, Theory: Eq. (4.57).

Non-equilibrium simulations MD simulations of ideal gas fluids in non-equilibrium conditions are performed using a fixed number of particles (200) in a one-dimensional (1D) domain of length 200 (in reduced units) with periodic boundary conditions, under an external potential:

$$V(x) = 5\left[\left(\frac{x}{200}\right)^4 - \left(\frac{x}{200}\right)^2\right]. \tag{C.12}$$

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.

C.1.3. Time integrators 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 analyze the stability of the different time integrators with respect to the time-step size. Specifically, in Fig. C.1 we report a comparison of the fluctuations standard deviation obtained with some selected

time integrators and the MD-theoretical results for varying time step sizes Δt_0. It is worth underlying that, because of the adaptive time step adopted in the simulations, Δt_0 represents only the initial time-step and the effective time-step may not be constant throughout the simulations, i.e. 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 $N_c = 25$. Moreover, we do not report the results for the Milstein schemes, since in previous 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 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 an enhanced stability compared to both implicit and explicit Euler-Maruyama.

C.1.4. Structure factor

As shown in previous works [71, 208], the structure factor represents an important measure of the stochastic properties of the system. It can be not only to study the stability of the numerical integrator, but also to compare different schemes. In this work, we derive an expression of the structure factor from the linearised FDDFT. If we consider a periodic domain of volume V, the spatial Fourier transform of the density is given by

$$\hat{\rho}_{\boldsymbol{\lambda}} = \frac{1}{V} \int_V \rho(\mathbf{r}, t) e^{-i\boldsymbol{\lambda}\cdot\mathbf{r}} d\mathbf{r}. \tag{C.13}$$

The structure factor is defined as the variance of the Fourier transform of the density fluctuations,

$$S(\boldsymbol{\lambda}) = V \langle \delta\hat{\rho}_{\boldsymbol{\lambda}} \delta\hat{\rho}^*_{\boldsymbol{\lambda}} \rangle \tag{C.14}$$

with the Fourier transform of the density fluctuations being $\delta\hat{\rho}_{\boldsymbol{\lambda}} = \hat{\rho}_{\boldsymbol{\lambda}} - \langle \hat{\rho}_{\boldsymbol{\lambda}} \rangle$, and $\hat{\rho}^*_{\boldsymbol{\lambda}}$ defined as the complex conjugate of $\hat{\rho}_{\boldsymbol{\lambda}}$. Taking the Fourier transform of the difference between the linearised version of Eq.(4.5) and Eq.(4.13), one obtains

$$\partial_t \, \delta\hat{\rho}(\boldsymbol{\lambda}) = i\boldsymbol{\lambda} \cdot \left\{ \mathcal{T} \left[\rho(\boldsymbol{r}; t) \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta \mathcal{E}[\rho]}{\delta \rho} \right] - \mathcal{T} \left\langle \rho(\boldsymbol{r}; t) \boldsymbol{\nabla}_{\boldsymbol{r}} \frac{\delta \mathcal{E}[\rho]}{\delta \rho} \right\rangle \right\} + i\boldsymbol{\lambda} \cdot \sqrt{\langle \rho(\boldsymbol{r}) \rangle / \beta} \hat{\boldsymbol{\mathcal{W}}}(\boldsymbol{\lambda}). \tag{C.15}$$

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

$$\mathcal{T}\left[\rho(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{E}[\rho]}{\delta\rho}\right] \sim \mathcal{T}\left\langle\rho(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{E}[\rho]}{\delta\rho}\right\rangle + \frac{\partial\mathcal{T}\left[\rho(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{E}[\rho]}{\delta\rho}\right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}\delta\hat{\rho}_{\boldsymbol{\lambda}} + \mathcal{O}(\delta\hat{\rho}_{\boldsymbol{\lambda}}), \tag{C.16}$$

then Eq. C.15 yields

$$\partial_t\ \delta\hat{\rho}_{\boldsymbol{\lambda}} = i\boldsymbol{\lambda}\cdot\frac{\partial\mathcal{T}\left[\rho(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{E}[\rho]}{\delta\rho}\right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}\delta\hat{\rho}_{\boldsymbol{\lambda}} + i\boldsymbol{\lambda}\cdot\sqrt{\langle\rho\rangle/\beta}\hat{\boldsymbol{\mathcal{W}}}(\boldsymbol{\lambda}). \tag{C.17}$$

Since the last equation has the form of a Ornstein–Uhlenbeck process, the structure factor can be computed as its variance:

$$S(\boldsymbol{\lambda}) = \frac{2\left(i\boldsymbol{\lambda}\sqrt{\langle\rho\rangle/\beta}\right)^2}{2i\boldsymbol{\lambda}\frac{\partial\mathcal{T}\left[\rho(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{E}[\rho]}{\delta\rho}\right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial\mathcal{T}\left[\rho(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}}\frac{\delta\mathcal{E}[\rho]}{\delta\rho}\right]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} \tag{C.18}$$

For example, in case of ideal gas without external potential, $\frac{\delta\mathcal{E}[\rho]}{\delta\rho} = \log\rho(\boldsymbol{r},t)$, the structure factor is given by:

$$S(\boldsymbol{\lambda}) = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial\mathcal{T}[\rho(\boldsymbol{r};t)\boldsymbol{\nabla}_{\boldsymbol{r}}\log\rho(\boldsymbol{r},t)]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial\mathcal{T}[\boldsymbol{\nabla}_{\boldsymbol{r}}\rho(\boldsymbol{r},t)]}{\partial\hat{\rho}_{\boldsymbol{\lambda}}}} = \frac{i\boldsymbol{\lambda}\langle\rho\rangle/\beta}{\frac{\partial[i\boldsymbol{\lambda}\hat{\rho}]}{\partial\hat{\rho}(\boldsymbol{\lambda})}} = \langle\rho\rangle/\beta \tag{C.19}$$

which is consistent with the expression of the structure factor derived in Ref. [71].

C.2. Non-homogeneous DDFT

C.2.1. Numerical methods: 2D scheme

In this section, we describe the semi-discrete finite-volume formulation for the 2D version of Eq. (4.73). In 2D the non-homogeneous diffusion $\boldsymbol{\theta}_0^{-1}$ is a 2×2 matrix that in general depends on x, y, t and ρ. $\boldsymbol{\theta}_0^{-1}$ is symmetric and positive definite, and we denote its components as

$$\boldsymbol{\theta}_0^{-1} = \begin{pmatrix} a & b \\ b & c \end{pmatrix}. \tag{C.20}$$

We divide the computational domain in rectangular cells $C_{j,k} = [x_{j-1/2}, x_{j+1/2}]\times[y_{k-1/2}, y_{k+1/2}]$, which are assumed to have uniform size $\Delta x\,\Delta y$ so that $x_{j+1/2} - x_{j-1/2} =$

$\Delta x,\ \forall j$, and $y_{k+1/2}-y_{k-1/2}=\Delta y,\ \forall k$. In each cell $C_{j,k}$ we define the density cell average as $\rho_{j,k}(t)=\frac{1}{\Delta x\Delta y}\int_{C_{j,k}}\rho(x,y,t)\,dx\,dy$. Integrating Eq. (4.73) over $C_{i,j}$, and applying the divergence theorem, it follows

$$\frac{d\rho_{j,k}}{dt}=-\frac{F^x_{j+1/2,k}-F^x_{j-1/2,k}}{\Delta x}-\frac{F^y_{j,k+1/2}-F^y_{j,k-1/2}}{\Delta y}. \tag{C.21}$$

with numerical fluxes given by

$$F^x_{j+1/2,k}=u^+_{j+1/2,k}\,\rho^E_{j,k}+u^-_{j+1/2,k}\,\rho^W_{j+1,k},\quad F^y_{j,k+1/2}=v^+_{j,k+1/2}\,\rho^N_{j,k}+v^-_{j,k+1/2}\,\rho^S_{j,k+1}. \tag{C.22}$$

The upwind approach is accomplished by

$$\begin{aligned}\left(u_{j+1/2,k}\right)^+&=\max\left(u_{j+1/2,k},0\right),\quad \left(u_{j+1/2,k}\right)^-=\min\left(u_{j+1/2,k},0\right),\\ \left(v_{j,k+1/2}\right)^+&=\max\left(v_{j,k+1/2},0\right),\quad \left(v_{j,k+1/2}\right)^-=\min\left(v_{j,k+1/2},0\right),\end{aligned} \tag{C.23}$$

with velocity satisfying

$$\begin{aligned}u_{j+1/2,k}&=-a_{j+1/2,k}\left(\frac{\xi_{j+1,k}-\xi_{j,k}}{\Delta x}\right)-b_{j+1/2,k}\left(\frac{\xi_{j,k+1}-\xi_{j,k}}{\Delta y}\right),\\ v_{j,k+1/2}&=-b_{j,k+1/2}\left(\frac{\xi_{j+1,k}-\xi_{j,k}}{\Delta x}\right)-c_{j,k+1/2}\left(\frac{\xi_{j,k+1}-\xi_{j,k}}{\Delta y}\right).\end{aligned} \tag{C.24}$$

Finally the discretized variation of the free-energy $\xi_{j,k}$ is computed as

$$\xi_{j,k}=\Delta x\Delta y\sum_{l,m}K_{j-l,k-m}\,\rho_{l,m}+f'(\rho_{j,k})+V_{j,k}. \tag{C.25}$$

Similarly to the one-dimensional case, the reconstructions of the density $\rho^E_{j,k}$, $\rho^W_{j,k}$, $\rho^N_{j,k}$ and $\rho^S_{j,k}$ are computed by means of a first- or second-order reconstruction of the form

$$\begin{cases}\rho^E_{j,k}=\rho_{j,k},\quad \rho^W_{j,k}=\rho_{j,k}, & \text{for the first-order scheme,}\\ \rho^E_{j,k}=\rho_{j,k}+\frac{\Delta x}{2}\left(\rho_x\right)_{j,k},\quad \rho^W_{j,k}=\rho_{j,k}-\frac{\Delta x}{2}\left(\rho_x\right)_{j,k}, & \text{for the second-order scheme,}\end{cases} \tag{C.26}$$

for $\rho^E_{j,k}$ and for $\rho^W_{j,k}$, and of the form

$$\begin{cases} \rho^N_{j,k} = \rho_{j,k}, \quad \rho^S_{j,k} = \rho_{j,k}, & \text{for the first-order scheme,} \\ \rho^N_{j,k} = \rho_{j,k} + \frac{\Delta y}{2}\left(\rho_y\right)_{j,k}, \quad \rho^S_{j,k} = \rho_{j,k} - \frac{\Delta y}{2}\left(\rho_y\right)_{j,k}, & \text{for the second-order scheme,} \end{cases} \tag{C.27}$$

for $\rho^N_{j,k}$ and for $\rho^S_{j,k}$. As with the 1D set up, the slopes are initially computed by means of central-difference approximations satisfying

$$\left(\rho_x\right)_{j,k} = \frac{\rho_{j+1,k} - \rho_{j,k}}{2\Delta x} \quad \text{and} \quad \left(\rho_y\right)_{j,k} = \frac{\rho_{j,k+1} - \rho_{j,k}}{2\Delta y}. \tag{C.28}$$

If at any point in the time evolution the reconstructions become negative, we recompute the slopes $\left(\rho_x\right)_{j,k}$ and $\left(\rho_y\right)_{j,k}$ by means of a monotone nonlinear limiter which ensures non-negativity. Following [209], we apply the generalized minmod limiters with $\theta \in [1,2]$:

$$\begin{aligned} \left(\rho_x\right)_{j,k} &= \text{minmod}\left(\theta\,\frac{\rho_{j+1,k} - \rho_{j,k}}{\Delta x}, \frac{\rho_{j+1,k} - \rho_{j-1,k}}{2\Delta x}, \theta\,\frac{\rho_{j,k} - \rho_{j-1,k}}{\Delta x}\right), \\ \left(\rho_y\right)_{j,k} &= \text{minmod}\left(\theta\,\frac{\rho_{j,k+1} - \rho_{j,k}}{\Delta y}, \frac{\rho_{j,k+1} - \rho_{j,k-1}}{2\Delta y}, \theta\,\frac{\rho_{j,k} - \rho_{j,k-1}}{\Delta y}\right). \end{aligned} \tag{C.29}$$

The numerical no-flux conditions are implemented by taking the numerical flux Eq. (C.22) to be zero at the boundaries:

$$\begin{aligned} F^x_{j-\frac{1}{2},k} &= 0 \text{ for } j = 1, \forall k, \quad F^x_{j+\frac{1}{2},k} = 0 \text{ for } j = N, \forall k, \\ F^y_{j,k-\frac{1}{2}} &= 0 \text{ for } k = 1, \forall j; \quad F^y_{j,k+\frac{1}{2}} = 0 \text{ for } k = N, \forall j. \end{aligned} \tag{C.30}$$

We conclude this subsection by proving the 2D version of the positivity of the density and the decay of the discrete free-energy.

Theorem C.2.1 (Positivity preserving). The 2D semi-discrete finite-volume scheme Eqs (C.21)-(C.25), coupled with a positivity preserving reconstruction Eqs (C.26)-(C.27) for ρ, preserves the positivity of the cell-averages $\rho_{j,k} \geq 0\,\forall\{j,k\}$ if the following CFL condition is satisfied:

- For the first-order reconstruction in Eqs (C.26)-(C.27):

$$\Delta t \leq \min_{i,j}\left\{\frac{u^+_{j+1/2,k} - u^-_{j-1/2,k}}{\Delta x} + \frac{v^+_{j,k+1/2} - v^-_{j,k-1/2}}{\Delta y}\right\}$$

- For the second-order reconstruction in Eqs (C.26)-(C.27):

$$\Delta t \leq \min\left\{ \frac{\Delta x}{4\max\left\{u^{+}_{j+1/2,k}, u^{-}_{j+1/2,k}\right\}}, \frac{\Delta y}{4\max\left\{v^{+}_{j,k+1/2}, v^{-}_{j,k+1/2}\right\}} \right\}$$

Proof. The proof follows the same lines as in [209].

□

For the decay of the discrete free-energy it is important to note that a full $\boldsymbol{\theta}_0^{-1}$ matrix Eq. (C.20) where $b \neq 0$ produces a dimensional coupling in the gradient flow Eq. (4.73), since the velocities Eq. (C.24) now contain terms from orthogonal directions. This can be avoided by performing a diagonalization of the matrix $\boldsymbol{\theta}_0^{-1}$, which is ensured due to the conditions imposed over it (namely that it is symmetric and positive definite). To show this we can first diagonalize $\boldsymbol{\theta}_0^{-1}$ as

$$\boldsymbol{\theta}_0^{-1} = QDQ^T,$$

where $QQ^T = Id_2$ and D is a positive diagonal matrix

$$D = \begin{bmatrix} \alpha & 0 \\ 0 & \beta \end{bmatrix}.$$

Formally, we can project the gradient operator onto the eigenvectors basis: $\boldsymbol{\nabla} = Q\boldsymbol{\nabla}_e$, and $\boldsymbol{\nabla} \cdot F = \boldsymbol{\nabla}^T F = \boldsymbol{\nabla}_e^T (Q^T F) = \boldsymbol{\nabla}_e \cdot (Q^T F)$. Therefore, the gradient-flow equation in Eq. (4.73) becomes

$$\frac{\partial \rho}{\partial t} = (Q\boldsymbol{\nabla}_e)^T \left(QDQ^T Q\boldsymbol{\nabla}_e \frac{\delta \mathcal{E}[\rho]}{\delta \rho} \right) = (Q\boldsymbol{\nabla}_e)^T \left(QD\boldsymbol{\nabla}_e \frac{\delta \mathcal{E}[\rho]}{\delta \rho} \right) = \boldsymbol{\nabla}^T \left(D\boldsymbol{\nabla}_e \frac{\delta \mathcal{E}[\rho]}{\delta \rho} \right),$$

so that Eq. (4.73) can now be written in the eigenvector basis as,

$$\frac{\partial \rho}{\partial t} = \boldsymbol{\nabla}_e \cdot (D\boldsymbol{\nabla}_e \rho). \tag{C.31}$$

The advantage of this formulation is that for Eq. (C.31) the diffusion matrix D is diagonal which then simplifies the velocities in Eq. (C.24) since $b = 0$. Following this transformation one can then show the decay of the discrete free-energy for Eq. (C.31).

Such discrete energy in 2D satisfies:

$$E_\Delta = \Delta x \Delta y \sum_{j,k} V_{j,k}\, \rho_{j,k} + H(\rho_{j,k}) + \frac{1}{2}\Delta x \Delta y \sum_{l,m} K_{j-l,k-m}\, \rho_{l,m}\, \rho_{j,k}. \tag{C.32}$$

Theorem C.2.2 (Discrete-energy decay). The 2D semi-discrete finite-volume scheme Eqs (C.21)-(C.25), coupled with a positivity preserving reconstruction Eqs (C.26)-(C.27) for ρ, satisfies the following decay of the discrete free-energy Eqs (C.32) for gradient flows where $b = 0$ in $\boldsymbol{\theta}_0^{-1}$ Eq. (C.20) or previously diagonalized as in Eq. (C.31):

$$\frac{d}{dt} E_\Delta(t) \leq -I_\Delta(t), \tag{C.33}$$

where the discrete version of the entropy dissipation results in

$$I_\Delta(t) = -\Delta x \Delta y \sum_{j,k} \left[\frac{\left(u_{j+1/2,k}\right)^2}{a_{j+1/2,k}} + \frac{\left(v_{j,k+1/2}\right)^2}{c_{j,k+1/2}} \right] \min_{j,k} \left(\rho^E_{j,k}, \rho^W_{j+1,k}, \rho^N_{j,k}, \rho^S_{j,k+1}\right). \tag{C.34}$$

Proof. Following the 1D proof in Sect. 4.2.1, the temporal decay of the discrete free-energy in Eq. (C.32) yields

$$\frac{d}{dt} E_\Delta(t) = -\Delta x \Delta y \sum_{j,k} \xi_{j,k} \left(\frac{F^x_{j+1/2,k} - F^x_{j-1/2,k}}{\Delta x} + \frac{F^y_{j,k+1/2} - F^y_{j,k-1/2}}{\Delta y} \right).$$

Then, by applying integration by parts with respect to both x and y yields:

$$\frac{d}{dt} E_\Delta(t) = -\Delta x \Delta y \sum_{j,k} \left[\left(\frac{\xi_{j,k} - \xi_{j+1,k}}{\Delta x} \right) F^x_{j+1/2,k} + \left(\frac{\xi_{j,k} - \xi_{j,k+1}}{\Delta y} \right) F^y_{j,k+1/2} \right).$$

With $b = 0$ in $\boldsymbol{\theta}_0^{-1}$ Eq. (C.20) and the expressions for the velocities in Eq. (C.24) and the fluxes in Eq. (C.22), it follows that:

$$\begin{aligned}
\frac{d}{dt} E_\Delta(t) =& -\Delta x \Delta y \sum_{j,k} \left[\frac{u_{j+1/2,k}}{a_{j+1/2,k}} \left(u^+_{j+1/2,k}\, \rho^E_{j,k} + u^-_{j+1/2,k}\, \rho^W_{j+1,k} \right) \right] \\
& - \Delta x \Delta y \sum_{j,k} \left[\frac{v_{j,k+1/2}}{c_{j,k+1/2}} \left(v^+_{j,k+1/2}\, \rho^N_{j,k} + v^-_{j,k+1/2}\, \rho^S_{j,k+1} \right) \right] \\
& \leq -\Delta x \Delta y \sum_{j,k} \left[\frac{\left(u_{j+1/2,k}\right)^2}{a_{j+1/2,k}} + \frac{\left(v_{j,k+1/2}\right)^2}{c_{j,k+1/2}} \right] \min_{j,k} \left(\rho^E_{j,k}, \rho^W_{j+1,k}, \rho^N_{j,k}, \rho^S_{j,k+1}\right).
\end{aligned}$$

Thus, the decay of the discrete free-energy in Eq. (C.33) is satisfied with a discrete

entropy dissipation of the form Eq. (C.34). □