THE UNIVERSITY *of* EDINBURGH

Granular media at multiple scales: mathematical analysis, modelling and computation

Tim. D. Hurst

Doctor of Philosophy
University of Edinburgh
June 17, 2020

Declaration

I declare that this thesis was composed by myself and that the work contained therein is my own, except where explicitly stated otherwise in the text.

(*Tim. D. Hurst*)
June 17, 2020

Dedicated to my family, they're an alright bunch.

Lay Summary

A granular medium is defined as a system of particles that are unaffected by quantum or thermal fluctuations, where the particles lose energy through collisions. The classic example to take here is sand, but the uses and applications of granular media are much more numerous and affect everyone in their day-to-day life, not just beach-goers. The cup of freshly ground coffee you enjoy in the morning, your walk or drive to work, and the aspirin you have to take after reading this thesis all involve granular media at different scales. The natural and industrial processes that rely on such systems of particles cannot be overstated: geology, astronomy, pharmacology and agricultural sciences are just some of the areas that utilise or depend on granular media.

The properties that make granular media so beneficial are the same properties that make it challenging to model mathematically: hard particles that interact in complicated ways at small timescales can change the behaviour of the entire system over long timescales. We can use modern numerical methods to model every particle in the system, and track how it interacts with others, but there's only so many particles a computer can handle. The alternative is to approximate the collection of particles as a fluid, then simulate the dynamics of the fluid using a computer. But how do we ensure that the fluid correctly models the interactions that the particles undergo, at a macroscopic level?

In this thesis we consider the fundamental microscopic and macroscopic properties of granular media, and construct a new mathematical model that describes a system of hard particles as a fluid. In doing so we will touch upon some important theoretic properties that must be accounted for when deriving a continuum model from the microscopic dynamics. We also look at efficient and accurate microscopic simulations, and show how we can use the results of these simulations to improve the validity of the continuum model.

Abstract

There are many challenges in modelling granular media, in particular due to hard particle interactions such as collisions. Modelling and simulating at a microscopic level produces very accurate results, but simulations are generally restricted to relatively small systems of particles. It is also difficult to construct a simple continuum model which accurately describes all the properties of granular media.

In this thesis, we consider a number of the problems associated with modelling granular media. We first look at the microscopic dynamics of individual particles and how to derive physically appropriate interactions between them, and discuss Event-Driven Particle Dynamics (EDPD) as an accurate and efficient way to model a system of hard, spherical particles. We then present a novel derivation of the weak form of the Liouville equation which can model systems where particles interact instantaneously (*e.g.* via inelastic collisions). From here we construct the BBGKY hierarchy and use moment closure methods to construct a new, accurate continuum model for granular media, based on Dynamical Density Functional Theory (DDFT). We then use EDPD to construct approximations for the radial correlation function which accounts for friction, packing fraction and inelasticity. This is then included in the DDFT in simulated examples.

Acknowledgements

I've been lucky enough during my PhD to meet and work with lots of people. It's impossible to thank everyone here who has contributed to my experience, but hopefully I can make a bit of a dent here.

Firstly, the PhD was supported by The Maxwell Institute Graduate School in Analysis and its Applications (MIGSAA), a Centre for Doctoral Training funded by the UK Engineering and Physical Sciences Research Council (EPSRC) (grant EP/L016508/01), the Scottish Funding Council, Heriot-Watt University and the University of Edinburgh. The majority of the knowledge and skills I've gained over the last four or five years were all made possible because of the funding and opportunities provided by MIGSAA.

At the end of the third year of my PhD I took part in a research internship at the Alan Turing Institute (supported by EPSRC grant EP/N510129/1). Although unrelated to my main research, the experience was very enriching, and led to collaborations that continued for the rest of my PhD. In particular I'd like to thank Matt Thorpe, Kostas Zygalakis and Oliver Crook, as well as Mihai Cucuringu, Carola-Bibaine Schönlieb and Elizabeth Soilleux for their participation in the project.

I'd like to thank Mark Wilkinson, for a number of energising discussions that led to an important and fulfilling chunk of my research. I would like to show my gratitude to Marcus Bannerman, who quickly became a rewarding collaborator during the last year of my PhD. I also appreciate the meetings I had with the powder group in the Department of Chemical Engineering at Heriot-Watt university, that made a mathematician feel very welcome.

An honourable mention also goes to the guys at Brainnwave, a data analytics company based in Edinburgh, where I worked during a placement in the last year of my PhD. They provided me with the opportunity to apply mathematical and statistical research to modern real-world problems and could not have been more welcoming or supportive. This placement was funded by an innovation grant (EP/S51505X/1) from the EPSRC.

And finally (and perhaps most importantly) I would like to thank my supervisors Ben Goddard and Raffaella Ocone, who provided me with the freedom and guidance that I needed to get the work done, and continually showed me how rewarding and enjoyable academic research can be.

Contents

Chapter 1

Introduction

Systems of granular media play an important role in many industrial processes and natural phenomena. When you consider the generality of its definition, as a conglomeration of solid 'macroscopic' particles (*i.e.* that are unaffected by quantum or thermal fluctuations) that lose energy through particle interactions, it is unsurprising that granular media are crucial in engineering, chemical, material, agricultural, pharmaceutical, geological and astronomical sciences [1], [2]. In nature we witness systems of granular media in pattern formation on sandy beaches [3]–[5], in the flow of solid particles down an inclined slope during an avalanche or landslide [6]–[9], or in the formation and movement of asteroid belts in space [10]–[12]. In industry the properties of a granular medium are utilised for various purposes, for example in the storage or processing of grains in a silo [13], [14], in mixing via the movement of material in a rotating drum [6], [15], [16], or the use of hard particles in various fluidisation processes, such as in catalysis [17]–[20].

Given its importance in nature and industry, it is very advantageous to have well-formed mathematical models that can be used to accurately predict how systems of granular media behave. There are many different models available for this purpose in the literature [21]–[23]. At a microscopic level the dynamic of interacting particles is well-studied and in particular are popularly modelled using Discrete Element Methods (DEM) (introduced in [24]), or Event-Driven Particle Dynamics (EDPD) (one of the earliest particle simulation methods, first discussed in [25]).

The discrete element method has become particularly popular in modern simulations of granular media [23], [26], [27]. By checking for particle-particle interactions at each timestep during a simulation, DEM can use contact-force models to include complex, sustained interactions between particles in a system [28], [29]. With the use of modern computational power we can then simulate systems of hundreds of millions (or even billions) of particles. Many applications have been considered using DEM, and the results have been validated with experiments [30]–[34].

However, even with state-of-the-art supercomputers, DEM simulations are generally restricted to short times, due to the small timescales and the complexity of particle interactions. To combat this, various coarse-graining techniques have been constructed [35], [36], where a number of particles are simulated together as a sort of 'super-particle'. Inevitably these approximations remove small-scale physical phenomena from the simulation to speed up computation, which may reduce the accuracy at the microscopic level.

Furthermore, general DEM simulations are inefficient in systems where particle interactions

can be considered instantaneous. In this case identifying and resolving contacts at every timestep becomes overly expensive. As an alternative one can utilise EDPD , where collisions are predicted from analytic particle trajectories, then resolved until the next collision takes place, under the assumption that collisions are approximately binary and instantaneous. There are then various algorithms that have been included in EDPD algorithms to render the cost of simulation to $\mathcal{O}(1)$ per collision [37] (which we will discuss in chapter 3). Thus, for low-density or fast-flowing systems, EDPD can have a competitive advantage, although to ensure it is efficient the free particle dynamics must be relatively simple, and industrial-scale systems where the number of particles can exceed 10^{20} (for example in fluidisation processes) are computationally intractable. In this case any microscopic simulation method will suffer, either in its efficiency or its accuracy. For DEM simulations, the search for interacting particles at each timestep will become prohibitively expensive, and for EDPD simulations the number of collisions between particles will render the system computationally intractable. It is then appropriate to instead consider a macroscopic continuum model.

The kinetic theory of granular flow is one of the popular and well-studied models for continuum simulation of granular flows [38]–[40]. Originally designed for the study of hard-sphere gases, kinetic theory was modified to account for dissipative particle interactions, so that it could be suitably applied to systems of granular gases. Since then, models have been adapted include various particle properties, such as particle shape [41] and higher densities [42]–[44]. However, although there is some work on linking the microscopic dynamics to the continuum models, many of the derivations for kinetic theories applied to granular flow are not derived directly from the microscopic dynamics, instead relying upon the Boltzmann equation, or are constructed by somewhat ad-hoc modifications of known results. There are some results available for constructing the Boltzmann equation from microscopic dynamics [45]–[47], but generally they avoid hard-particle collisions or do not account for other interaction forces.

Alternatively, recent developments in $\mu(I)$-rheology have shown promise in the development of a model for the rheology of a dense granular flow [48]–[50]. There are many interesting results in this area, as well as some preliminary results in ensuring that the system of equations are well-posed [51]. However instead of determining the model by considering microscopic results, Navier Stokes-like equations are developed and modified by considering the physical properties of a macroscopic granular flow. Furthermore,$\mu(I)$-rheology models are designed to model high-density systems, so we cannot expect them to perform as well in the low-density regime.

There are also models that combine microscopic and macroscopic methods to accurately describe the dynamics of granular media [52], [53]. These methods are particularly relevant for systems where granular particles are agitated and interact with another fluid. This complex fluid-particle relationship is utilised in fluidisation processes such as catalysis. By coupling a DEM model with computational fluid dynamics, fluidisation can be modelled directly using CFD-DEM [54]–[56]. The resulting simulations used to model fluidisation processes [57], [58], however the coupling introduces additional computational cost to the simulation. Alternatively we can again approximate the solid phase with a continuum model (usually by using the kinetic theory of granular flow to simulate the granular medium as a non-Newtonian fluid), producing a two-fluid model [59]–[63]. Two-fluid models are also a thriving topic of modern research, but their accuracy will still rely on transporting careful, valid assumptions from the microscopic dynamics into the macroscopic

model.

There are so many different models for granular media because there are several physical phenomena that a particular system of interest can experience. The particles in a granular material are defined to have a finite volume and cannot overlap, which can cause interesting packing properties. The energy lost through particle collisions can cause non-homogeneous aggregation at boundaries or in the bulk of a system. Additional microscopic properties such as cohesion [64] or charge [65], [66] (for systems of wet or dry particles respectively) can lead to further macroscopic changes to the system. It is important to consider relevant physical phenomena in the construction of a mathematical model. Doing so in an accurate way leads to many difficulties which make microscopically simulating granular media computationally expensive. For example, energy loss due to particle interactions can cause *inelastic collapse* [67], where an infinite number of collisions can occur in finite time, which can effectively jam numerical simulations. Phenomena such as this must be included in the continuum model in a physically-relevant way, to ensure that the result is accurate.

Moving from a microscopic model to a continuum can drastically reduce computational cost, as simulation time no longer depends explicitly on the number of particles (although fine computational grids may have to be implemented to capture behaviour on the scale of particle size). Furthermore, complicated interparticle interactions that have to be resolved at the microscopic level can be incorporated in the governing continuum equations without incurring a similar penalty in the computational cost of the system. However, care has to be taken to ensure that the relevant physical phenomena modelled at the microscopic level are correctly represented in the continuum model. Approximations will inevitably have to be made, for example to close the system of equations and remove a reliance on hihger-order moments or particle distributions (as discussed in chapter 6), to construct a system of equations for a continuum model which is computationally tractable,

In this thesis, we consider the mathematical derivation of a continuum model for granular media. Our derivation considers fast-flowing or low-density systems, where interactions between individual particles can be approximated by binary, instantaneous interactions. At the particle level, we consider what constitutes a instantaneous interaction between particles, and how to ensure it is physically valid. We then use novel approaches to move from the microscopic (where the movement of particles are described the Newton equations) to the continuum (where in the first instance the Liouville equation models the evolution of the distribution of all particles in a system), accounting for volume exclusion and instantaneous particle interactions such as collisions. This new derivation leads to additional terms in the Liouville equation that account for instantaneous interactions in the evolution of the distribution of particles.

Following this, we produce a new state-of-the-art Dynamical Density Functional Theory (DDFT), that can be used to model granular media, and accounts for inelastic collisions between particles, as well as other external or interparticle effects. The application of DDFT to granular media is motivated by successful models in other complex fluids such as colloidal flows, where complex particle interactions can be modelled using the equilibrium approximation of the Helmholtz free energy functional. This new application explores how DDFT methods can be modified to model granular flows. To derive this model we first pass to the Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy, then perform a moment closure on the first equation of the hierarchy. The resulting DDFT involves additional terms when compared to previously derived equations for col-

loids [68]–[73] or one-dimensional thermostatted granular systems [74], as well as an additional equation to model granular temperature in the DDFT.

There are quantities in the DDFT model that need to be reasonably approximated to close the system of governing equations. To do this, we simulate systems of a small number of particles, extract any required quantities by repeating the numerical experiments and taking the average result, and including these value in the DDFT simulation. As our derivation assumes that particle collisions are binary and instantaneous, EDPD is chosen as a suitable candidate for the microscopic simulations, after we adapt known algorithms to account for viscous drag in the system. In particular we study the effect of various microscopic parameters on the radial correlation function. We will show that viscous and dissipative collisional terms in the microscopic dynamics have a strong effect on the behaviour of the function, unlike in the well-studied steady-state approximation which is popular in continuum models [62], [75]. We then finally present some examples of the DDFT which test the theory, and suggest that the newly-derived model and the parametrised form of the radial correlation function have an important effect on the dynamics of a granular medium.

We note that, although the results presented show promise, the constructed DDFT is somewhat limited in its scope. Modern industrial applications that can be treated in this way are generally multi-phase, and in our derivation we have simplified the interaction between the external fluid and granular particles. However, the methodology laid out in this thesis could also be applied to multiphase systems (in a similar way to the colloidal case [76]), to produce a system of equations based on DDFT results that could be used to model industrial processes such as fluidisation, as a competitor to other two-fluid models. In this thesis we instead focus on theoretical aspects of the derivation, and the numerical experiments that are included are designed to test the theories presented.

1.1 Structure of the thesis

We here provide an overview of the results presented in each chapter of the thesis.

Chapter 2. Microscopic dynamics: modelling

In this chapter we outline the microscopic model considered in this thesis. This involves describing the free dynamics undertaken by particles unaffected by instantaneous interactions, then constructing an operator that describes all physically valid instantaneous interactions between hard spheres, including inelastic collisions and refractive events caused by discrete potential interactions. This original work resulted from collaborations with Dr. Mark Wilkinson[1,2] and one of my academic supervisors, Dr. Ben Goddard[3,4], as discussed in [77].

Chapter 3. Microscopic dynamics: simulation

Following chapter 2 we describe our preferred method for simulating microscopic models: Event-Driven Particle Dynamics (EDPD). There are several computational methods that make EDPD an

[1] mark.wilkinson@hw.ac.uk
[2] Department of Mathematics, Heriot-Watt University, Edinburgh, UK, EH14 4AS
[3] bgoddard@ed.ac.uk
[4] School of Mathematics and the Maxwell Institute for Mathematical Sciences, University of Edinburgh, Edinburgh, UK, EH9 3FD

efficient and accurate method for dynamics with particles that experience instantaneous interactions. We implement EDPD in `MATLAB` in one dimension, and use open source software `DynamO` to perform microsopic simulations in two and three dimensions for this work. Collaborating with Dr. Marcus Bannerman[5,6], we have also included frictional terms in `DynamO` computations. This work was also undertaken with both of my academic supervisors, Dr. Ben Goddard and Prof. Raffaella Ocone[7,8].

Chapter 4. the Liouville equation for hard particles

In this chapter, we consider the weak formulation of the Liouville equation for particles that undergo instantaneous interactions. This is an important step to constructing a valid physical model for granular media; the classical result for the Liouville equation is not valid for hard spheres as volume exclusion and collisional effects must be included in the derivation. The investigation produces additional terms in the weak Liouville equation as a consequence of volume exclusion, that are not present in the standard derivation of the Liouville equation, and lead to well-known collision operators in the BBGKY hierarchy. This work is also discussed in [77].

Chapter 5. The BBGKY hierachy for granular media

The Liouville equation constructed in chapter 4 is the first continuum model we consider for granular media. To derive a continuum model that can be efficiently simulated, we need to reduce the dimensionality of the PDEs under consideration. We therefore consider the BBGKY hierarchy for hard spheres. Many of the results here are well-known, but additional care must be taken for hard sphere terms resulting from the weak formulation of the Liouville equation. We also construct moments of the first two equations in the BBGKY hierarchy, which will later be used to create a low dimensional system of PDEs for modelling granular media.

Chapter 6. Moment closure methods

In this chapter we consider two moment closure methods; maximum entropy closure and equilibrium approximations leading to a Dynamical Density Functional Theory (DDFT). We motivate the moment closure method by first considering the maximum entropy closure scheme, based on work in [78]. This scheme shows a promising methodology for constructing a reasonable computational model, but is limited in scope due to its dimensionality. Following these results we construct a novel DDFT for granular media, under assumptions on the local equilibrium and correlations in the system. This DDFT model overcomes dimensionality issues of the model derived using maximum entropy closure. This is based on work with Dr. Ben Goddard and Prof. Raffaella Ocone [79]. At the end of this section we provide additional detail on the maximum entropy scheme.

Chapter 7. The radial correlation function: Numerical investigation

In this section we use the numerical methods discussed in chapter 3 to investigate long-term dynamics of systems of inelastic hard particles. In many continuum models, the radial correlation

[5] m.campbellbannerman@abdn.ac.uk
[6] School of Engineering, University of Aberdeen, Aberdeen AB24 3FX
[7] r.ocone@hw.ac.uk
[8] School of Engineering and Physical Sciences, Heriot-Watt University, Edinburgh, UK, EH14 4AS

function is approximated by its value at steady (or homogeneous cooling) state. For our purposes, this is not the optimal choice for the correlation function, as our continuum models are not simulated on steady-state timescales. We consider the time taken for systems to reach a steady state, and construct the radial correlation function for hard rods, discs and spheres, with and without a viscous term in the microscopic equations of motion, by using EDPD. We discuss the behaviour of the radial correlation function for different coefficients of restitutions and packing fractions, as well as other properties of the microscopic model.

Chapter 8. Granular DDFT examples

In this chapter we present one and two-dimensional proof-of-concept examples of the constructed DDFT. The examples presented show the importance of each term in the DDFT, and we use parametrised data from the empirical radial correlation functions constructed in chapter 7 in the collision operator in the models considered. These results are also presented in [79].

Chapter 9. Conclusions and future work

In this section we briefly conclude the results of work discussed in the main body of the thesis and discuss open problems.

Appendices

The five appendices in this thesis provide additional information related to, and work unrelated to, the main content. In appendix A, we construct simplifications of the collision operator in one and two dimensions, using the Maxwellian local equilibrium assumption discussed in chapter 6, which are then used for faster computations in chapter 8.

In appendix B we discuss the code used to construct the results presented in this thesis, including the `MATLAB` library for EDPD, uses of the `DynamO` library for two and three dimensional EDPD simulations, and extensions of the `2dchebclass` [80] for granular DDFT simulations. All code for simulations in this thesis is available upon request: please email `t.hurst@sms.ed.ac.uk` for more information.

In appendices C and D we discuss some work undertaken in the initial stages of the PhD, investigating efficient methods to accurately approximate non-adiabatic transitions in quantum mechanics. This is joint work with Prof. Volker Betz[9,10] and Dr. Ben Goddard, the results of which are also available in [81] and [82]. A brief description of the code used is also provided.

Finally in appendices E and F work undertaken during a research internship at the Alan Turing Institute is presented. The work focuses on different problems in semi-supervised learning algorithms, firstly in novel and efficient algorithms that utilise continuum results of the discrete p-Dirichlet energy minimisation problems [83], and secondly the computation of TL^p transport distances using a linearisation approximation. This work is collaborative with Dr. Matthew

[9] `betz@mathematik.tu-darmstadt.de`
[10] Fachbereich Mathematik, Technische Universität Darmstadt, 64289 Darmstadt, Germany

Thorpe[11,12] , Oliver Crook [13,14] , Dr. Kostas Zygalakis[15,4], Prof. Mihai Cucuringu[16,17], and Prof. Carola-Bibaine Schonlieb[18,12].

The work presented in appendices C to F are used as a personal record of research undertaken during the PhD, and are not intended to be considered as part of the main content of this thesis.

[11] m.thorpe@maths.cam.ed.ac.uk
[12] Department of Applied Mathematics and Theoretical Physics, University of Cambridge, Cambridge, CB3 0WA, UK
[13] omc25@cam.ed.ac.uk
[14] MRC Biostatistics Unit, School of Clinical Medicine, University of Cambridge, Cambridge CB2 0SR, UK
[15] K.Zygalakis@ed.ac.uk
[16] mihai.cucuringu@stats.ox.ac.uk
[17] Department of Statistics University of Oxford, Oxford 24-29 St Giles' Oxford OX1 3LB, United Kingdom
[18] cbs31@cam.ac.uk

Chapter 2

Microscopic dynamics - modelling

At the microscopic level, granular media can be seen as a collection of interacting particles. These particles can be affected by external forces such as gravity or friction, but also by interaction forces between particles, such as repulsion or attraction.

Throughout, we consider a system of N particles in d dimensions. For $i = 1, \dots, N$, we let $\boldsymbol{r}_i(t) \in \mathbb{R}^d, \boldsymbol{p}_i(t) \in \mathbb{R}^d$ be the position and momentum of the i^{th} particle at time $t \in \mathbb{R}$. We define $\boldsymbol{r}^N(t) \in \mathbb{R}^{dN}, \boldsymbol{p}^N(t) \in \mathbb{R}^{dN}$ to be the concatenation of all particle positions and momenta respectively, *i.e.* $\boldsymbol{r}^N(t) = (\boldsymbol{r}_1(t), \boldsymbol{r}_2(t), \dots, \boldsymbol{r}_N(t))$ and similarly for $\boldsymbol{p}^N(t)$. We assign each particle a mass $m_i > 0$, and define $\boldsymbol{M} \in \mathbb{R}^{N \times N}$ to be the diagonal matrix with particle masses in its entries. The velocity of a particle is then given by $\boldsymbol{v}_i = \boldsymbol{p}_i / m_i$, and we write $\boldsymbol{v}^N = \boldsymbol{M}^{-1}\boldsymbol{p}^N$, noting that when $\boldsymbol{M} = \boldsymbol{I}$ (the identity matrix), velocity and momentum are the same. We provide the initial condition $\boldsymbol{r}^N = \boldsymbol{r}_0^N \in \mathbb{R}^{dN}, \boldsymbol{p}^N = \boldsymbol{p}_0^N \in \mathbb{R}^{dN}$. Initial conditions for the i^{th} particle in the system are labelled by $\boldsymbol{r}_{i,0}, \boldsymbol{p}_{i,0} \in \mathbb{R}^d$.

2.1 Particle dynamics

Given initial conditions $\boldsymbol{r}^N(0) = \boldsymbol{r}_0^N \in \mathbb{R}^{dN}, \boldsymbol{p}^N(0) = \boldsymbol{p}_0^N \in \mathbb{R}^{dN}$, the movement of point-like particles can be modelled by Newton's equations, a pair of coupled ordinary differential equations:

$$\frac{\mathrm{d}\boldsymbol{r}^N}{\mathrm{d}t} = \boldsymbol{M}^{-1}\boldsymbol{p}^N, \tag{2.1}$$

$$\frac{\mathrm{d}\boldsymbol{p}^N}{\mathrm{d}t} = -\nabla_{\boldsymbol{r}^N} V(\boldsymbol{r}^N, t) - \boldsymbol{\Gamma}(\boldsymbol{r}^N, \boldsymbol{p}^N, t)\boldsymbol{p}^N. \tag{2.2}$$

Equation (2.1) tells us that the position of all particles changes due to their momentum. The change in momentum is then determined by the second equation, eq. (2.2). The first term on the right-hand side of eq. (2.2) is the gradient of a smooth potential term which incorporates all external and interaction potentials that affect the particles, *e.g.* gravity and electrostatic forces. We can separate $V(\boldsymbol{r}^N, t)$ into individual interaction potentials:

$$V(\boldsymbol{r}^N, t) = \sum_i V^{\text{ext}}(\boldsymbol{r}_i, t) + \frac{1}{2}\sum_{i,j} V_2(\boldsymbol{r}_i, \boldsymbol{r}_j, t)$$

$$+\frac{1}{3!}\sum_{i,j,k} V_3(\boldsymbol{r}_i,\boldsymbol{r}_j,\boldsymbol{r}_k,t)+\dots \tag{2.3}$$

Here $V^{\text{ext}}(\boldsymbol{r}_i,t)$ is the effect of external potentials on the i^{th} particle. The potentials V_l where $2\leq l\leq n$ provide smooth interactions between particles, and the inverse factorial prefactors are a modelling choice, that will cancel later in the derivation of a continuum model (see chapter 5). The final term on the right-hand side of eq. (2.2) describes viscous effects with the external bath, which in full generality can depend on the position and momenta of every particle. The effects are determined by the values of the tensor $\boldsymbol{\Gamma}(\boldsymbol{r}^N,\boldsymbol{p}^N,t)$, which is sometimes replaced with a scalar friction coefficient $\gamma>0$ as an approximation (*e.g.* in colloidal systems [68]). In a system with $V(\boldsymbol{r}^N(t),t)=0$ this causes particle velocities to slow at an exponential speed. Alternatively in systems used to model fluidisation of a granular medium, this viscous term may depend on the relative velocities between particles and the external fluid [61].

Point-like particles can also be modelled stochastically using the Langevin equations [84], a coupled set of stochastic differential equations (where for ease of notation we have assumed that all particles have mass m):

$$\frac{\mathrm{d}\boldsymbol{r}^N}{\mathrm{d}t}=\frac{\boldsymbol{p}^N}{m}, \tag{2.4}$$

$$\frac{\mathrm{d}\boldsymbol{p}^N}{\mathrm{d}t}=-\nabla_{\boldsymbol{r}^N}V(\boldsymbol{r}^N,t)-\boldsymbol{\Gamma}(\boldsymbol{r}^N,\boldsymbol{p}^N,t)\boldsymbol{p}^N+\sqrt{mk_bT\gamma}\boldsymbol{\xi}^N(t). \tag{2.5}$$

The additional term $\boldsymbol{\xi}^N(t)=(\boldsymbol{\xi}_1(t),\boldsymbol{\xi}_2(t),\dots,\boldsymbol{\xi}_N(t))$ in eq. (2.5) is a concatenation of N d-dimensional Brownian motion terms [85], *i.e.* $\xi=(\xi_1(t),\xi_2(t),\dots,\xi_d(t))$ where

$$\langle\xi_i(t)\rangle=0,\quad \langle\xi_i(t)\xi_j(t)\rangle=\delta_{ij}\delta(t-t'),\quad i,j\in\{1,\dots,d\}, \tag{2.6}$$

where δ_{ij} is the Kronecker delta:

$$\delta_{ij}=\begin{cases}1, & i=j,\\ 0, & i\neq j.\end{cases} \tag{2.7}$$

and $\delta(x)$ is the Dirac function, which we heuristically define by the identities:

$$\delta(x)=\begin{cases}+\infty, & x=0,\\ 0, & x\neq 0,\end{cases} \tag{2.8}$$

$$\int_{-\infty}^{\infty}\delta(x)\,\mathrm{d}x=1, \tag{2.9}$$

but can be rigorously defined as a distribution or measure [86]. Informally, stochasticity is included by adding white noise in the Newton equations. The prefactor of the Brownian term is determined by a fluctuation-dissipation theorem [87]. The Langevin equations are popular in molecular dynamics [21] and are also used in colloidal systems [68], where particles are small enough to be affected by thermal fluctuations in the external fluid. Although granular media is generally unaffected by thermal fluctuations by definition, the Brownian term is sometimes included in models for granular media [74], as a thermostat for the dynamics. We choose to consider the deterministic

dynamics in our microscopic derivations, although the addition of white noise is well studied for other systems such as colloids, and can be informally included in the DDFT derivation discussed in chapter 6.

As we have discussed, it is well-known that volume exclusion and collisional effects play an important role in a granular medium. We now modify the dynamics discussed here to include these physical phenomena.

2.2 Effects of hard spheres

One of the defining properties of particles in granular media is that they are *hard*; they take up a finite, non-zero volume. For our purposes, we will approximate the shape of all particles as $(d-1)$-spheres with diameter ε. There are some important consequences to the finite volume assumption. Firstly, as particles cannot overlap there is a maximum number of particles that can fit in a finite volume. Many models incorporate these volume exclusion effects by including an infinite step potential:

$$V_\infty(\boldsymbol{r}_i(t), \boldsymbol{r}_j(t)) = \begin{cases} \infty, & \|\boldsymbol{r}_i - \boldsymbol{r}_j\| < \varepsilon, \\ 0, & \|\boldsymbol{r}_i - \boldsymbol{r}_j\| \geq \varepsilon. \end{cases} \tag{2.10}$$

when simulating microscopic dynamics numerically this potential is generally approximated by a smooth, steep potential. To avoid massive velocity changes when particles come close to one another, a small timestep is also generally required. By studying this infinite step potential one can derive continuum models with volume exclusion in one dimension (by considering the Percus free energy [88]), or in higher dimensions (using fundamental measure theory [89]).

The infinite step potential helps include the effects of finite volume, but does not directly model collisional dynamics that arise as a consequence of hard particles.

As an alternative, particles in a granular medium can be simulated using non-instantaneous contact force models. For example, using the Discrete Element Method, where particles are considered 'soft', interactions between particles are resolved using various contact force models that depend on the particle overlap, for example using a linear spring-dashpot model [24], or more complicated models [28], [90] that account for tangential forces and torque. For a more detailed review of such contact models we refer the reader to [26]. As a benefit, these schemes allow for many-body interactions and additional collisional effects, and are therefore more appropriate for high-density systems where sustained contacts between particles are more common. However, in systems of low density (or where interactions can be reasonably approximated as instantaneous), the interactions that we consider here as a consequence of momentum conservation will be more appropriate.

We include collisional effects by restricting the dynamics to the *hard-sphere domain*; for all times $t \in \mathbb{R}$ we require that

$$\boldsymbol{r}^N(t) \in \mathcal{P}_\varepsilon := \{\boldsymbol{x}^N \in \mathbb{R}^{dN} : \forall i, j \in \{1, \dots, N\}, i \neq j, \|\boldsymbol{x}_i - \boldsymbol{x}_j\| \geq \varepsilon\}. \tag{2.11}$$

We note that the hard-sphere domain is a real analytic manifold and define its boundary as $\partial\mathcal{P}_\varepsilon$. This supplies an additional restriction to the initial conditions provided in the dynamics; initial

data is only admissible if it satisfies eq. (2.11). We note that for given initial position data $\boldsymbol{r}^N \in \mathcal{P}_\varepsilon$, the set of admissible initial velocity data may not encompass all of $\mathbb{R}^{dN}$. We will assume that the admissible velocity data $\mathcal{V}(\boldsymbol{r}_0^N)$ is an analytic submanifold of $\mathbb{R}^{dN}$ which depends on the initial position data provided. To stop particles from overlapping, we must instantaneously change the velocities of particles when they come into contact; we call this transformation a *collision rule*. We will show that any valid collision rule must take a particular form, which is satisfied by standard collision rules available in the literature. For these arguments we will set $m_i = 1$ for all $i = 1, \dots, N$, which simplifies notation by setting momentum equal to velocity, the result is similar for particles with non-unit mass.

We will assume that all contacts are binary, *i.e.* they only involve two particles, which is a standard assumption for fast-flowing or low-density systems, where contacts between particles are very short and can be approximated as instantaneous, so that the probability of many-body interactions becomes vanishingly small. To ensure that the collision rule is physically valid, it must conserve linear and angular momentum, as described in [91]. Let particles i and j be such that at time t, $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon$, and define $\boldsymbol{v}_i^{\text{in}}, \boldsymbol{v}_j^{\text{in}}$ as the incoming velocities, and $\boldsymbol{v}_i^{\text{out}}, \boldsymbol{v}_j^{\text{out}}$ as the outgoing velocities after a collision has occurred. Then the incoming and outgoing velocities must satisfy the conservation of linear momentum:

$$\boldsymbol{v}_i^{\text{out}} + \boldsymbol{v}_j^{\text{out}} = \boldsymbol{v}_i^{\text{in}} + \boldsymbol{v}_j^{\text{in}}, \tag{2.12}$$

and the conservation of angular momentum:

$$(\boldsymbol{r}_i - \boldsymbol{a}) \times \boldsymbol{v}_i^{\text{out}} + (\boldsymbol{r}_j - \boldsymbol{a}) \times \boldsymbol{v}_j^{\text{out}} = (\boldsymbol{r}_i - \boldsymbol{a}) \times \boldsymbol{v}_i^{\text{in}} + (\boldsymbol{r}_j - \boldsymbol{a}) \times \boldsymbol{v}_j^{\text{in}}, \quad \forall \boldsymbol{a} \in \mathbb{R}^3, \tag{2.13}$$

where $\boldsymbol{a} \times \boldsymbol{b}$ is the standard cross product. Using these restrictions, we can show that any instantaneous change must have a particular form. Firstly, we have the following geometric result concerning the conservation laws in three dimensions, which can also be simplified for $d = 1$ or $d = 2$.

Proposition 1. For $d = 3$, eq. (2.12) and eq. (2.13) are true for all $\boldsymbol{a} \in \mathbb{R}^3$ if and only if

$$(\boldsymbol{r}_i - \boldsymbol{q}_k) \times \boldsymbol{v}_i^{\text{out}} + (\boldsymbol{r}_j - \boldsymbol{q}_k) \times \boldsymbol{v}_j^{\text{out}} = (\boldsymbol{r}_i - \boldsymbol{q}_k) \times \boldsymbol{v}_i^{\text{in}} + (\boldsymbol{r}_j - \boldsymbol{q}_k) \times \boldsymbol{v}_j^{\text{in}}, \tag{2.14}$$

for $k = 1, \dots, 4$, where $\{\boldsymbol{q}_k\}_{k=1}^4 \subset \mathbb{R}^3$ are the vertices of a (non-degenerate) polytope in $\mathbb{R}^3$, *i.e* for example $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q_3}$ form a linearly independent set, and $\boldsymbol{q}_4$ can be written as a linear combination of $\boldsymbol{q}_1, \boldsymbol{q}_2, \boldsymbol{q_3}$ with non-zero coefficients.

Proof. By the conservation of angular momentum, necessity is trivial. For sufficiency, by eq. (2.12) we may only consider $\boldsymbol{r}_i = 0$ without loss of generality, and so for each $\boldsymbol{q}_k$, for all constants $c_k \in \mathbb{R}$

$$\begin{aligned} &- c_k \boldsymbol{q}_j \times \boldsymbol{v}_i^{\text{out}} + c_k(\boldsymbol{r}_j - \boldsymbol{q}_k) \times \boldsymbol{v}_j^{\text{out}} = -c_k \boldsymbol{q}_j \times \boldsymbol{v}_i^{\text{in}} + c_k(\boldsymbol{r}_j - \boldsymbol{q}_k) \times \boldsymbol{v}_j^{\text{in}}, \\ \implies &- \sum_{k=1}^4 c_k \boldsymbol{q}_j \times \boldsymbol{v}_i^{\text{out}} + \sum_{k=1}^4 c_k(\boldsymbol{r}_j - \boldsymbol{q}_k) \times \boldsymbol{v}_j^{\text{out}} = - \sum_{k=1}^4 c_k \boldsymbol{q}_j \times \boldsymbol{v}_i^{\text{in}} + \sum_{k=1}^4 c_k(\boldsymbol{r}_j - \boldsymbol{q}_k) \times \boldsymbol{v}_j^{\text{in}}. \end{aligned}$$

If we now suppose that $\sum_{j=1}^{4} c_j = 1$, then we have that eq. (2.13) is satisfied for all q in the set

$$\mathcal{C} = \left\{ \sum_{i=1}^{4} c_j \boldsymbol{q}_j : \sum_{j=1}^{4} c_j = 1 \right\}. \tag{2.15}$$

As $\mathcal{C}$ is a convex set in $\mathbb{R}^3$, we infer that $\mathcal{C} = \mathbb{R}^3$, as required. □

We then define a map $\boldsymbol{\sigma}^+(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{v}_i, \boldsymbol{v}_j) : \mathbb{R}^6 \to \mathbb{R}^6$ from pre-collisional to post-collisional velocities (which we call the *forward-time event map*):

$$\begin{bmatrix} \boldsymbol{v}_i^{\text{out}} \\ \boldsymbol{v}_j^{\text{out}} \end{bmatrix} := \boldsymbol{\sigma}^+(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{v}_i, \boldsymbol{v}_j) \begin{bmatrix} \boldsymbol{v}_i^{\text{in}} \\ \boldsymbol{v}_j^{\text{in}} \end{bmatrix}. \tag{2.16}$$

Analogously we define the *backward-time event map* $\boldsymbol{\sigma}^-(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{v}_i, \boldsymbol{v}_j) : \mathbb{R}^6 \to \mathbb{R}^6$ to take post-collisional velocities to pre-collisional velocities:

$$\begin{bmatrix} \boldsymbol{v}_i^{\text{in}} \\ \boldsymbol{v}_j^{\text{in}} \end{bmatrix} := \boldsymbol{\sigma}^-(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{v}_i, \boldsymbol{v}_j) \begin{bmatrix} \boldsymbol{v}_i^{\text{out}} \\ \boldsymbol{v}_j^{\text{out}} \end{bmatrix}. \tag{2.17}$$

We can then define the structure of event maps in $d = 3$ by considering conservation laws and the result above. For this theorem we will use the *tensor product*, which will now define in full.

Let

$$\boldsymbol{A} = (a_{ij})_{i=1,\dots,n, j=1,\dots,m} \in \mathbb{R}^{n\times m} \quad \boldsymbol{B} = (b_{ij})_{i=1,\dots,k, j=1,\dots,l} \in \mathbb{R}^{k\times l}.$$

The tensor product $\boldsymbol{A} \otimes \boldsymbol{B} \in \mathbb{R}^{nk\times ml}$ is given by the following element-wise multiplication:

$$\boldsymbol{A} \otimes \boldsymbol{B} = \begin{pmatrix} a_{11}b_{11} & a_{11}b_{12} & \dots & a_{11}b_{1l} & \dots & \dots & a_{1m}b_{11} & a_{1m}b_{12} & \dots & a_{1m}b_{1l} \\ a_{11}b_{21} & a_{11}b_{22} & \dots & a_{11}b_{2l} & \dots & \dots & a_{1m}b_{21} & a_{1m}b_{22} & \dots & a_{1m}b_{2l} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{11}b_{k1} & a_{11}b_{k2} & \dots & a_{11}b_{kl} & \dots & \dots & a_{1m}b_{k1} & a_{1m}b_{k2} & \dots & a_{1m}b_{kl} \\ \vdots & \vdots & & \vdots & \ddots & & \vdots & \vdots & & \vdots \\ \vdots & \vdots & & \vdots & & \ddots & \vdots & \vdots & & \vdots \\ a_{n1}b_{11} & a_{n1}b_{12} & \dots & a_{n1}b_{1l} & \dots & \dots & a_{nm}b_{11} & a_{nm}b_{12} & \dots & a_{nm}b_{1l} \\ a_{n1}b_{21} & a_{n1}b_{22} & \dots & a_{n1}b_{2l} & \dots & \dots & a_{nm}b_{21} & a_{nm}b_{22} & \dots & a_{nm}b_{2l} \\ \vdots & \vdots & \ddots & \vdots & & & \vdots & \vdots & \ddots & \vdots \\ a_{n1}b_{k1} & a_{n1}b_{k2} & \dots & a_{n1}b_{kl} & \dots & \dots & a_{nm}b_{k1} & a_{nm}b_{k2} & \dots & a_{nm}b_{kl} \end{pmatrix}.$$

We extend the tensor products to vectors $\boldsymbol{a} = (a_i)_{i=1,\dots,n,} \in \mathbb{R}^n$ and $\boldsymbol{b} = (b_i)_{i=1,\dots,m} \in \mathbb{R}^m$ as $\boldsymbol{a} \otimes \boldsymbol{b} \in \mathbb{R}^{n\times m}$ where

$$\boldsymbol{a} \otimes \boldsymbol{b} = \begin{pmatrix} a_1b_1 & a_1b_2 & \dots & a_1b_m \\ a_2b_1 & a_2b_2 & \dots & a_2b_m \\ \vdots & \vdots & \ddots & \vdots \\ a_nb_1 & a_nb_2 & \dots & a_nb_m \end{pmatrix}.$$

Tensor product notation is particularly useful when considering multi-dimensional derivatives of vectors; for $\boldsymbol{a} \in \mathbb{R}^n$ and a differentiable function $\boldsymbol{f} : \boldsymbol{a} \to \mathbb{R}^m$ we define $\nabla_{\boldsymbol{a}} \boldsymbol{F}(\boldsymbol{a}) \in \mathbb{R}^{n \times m}$ as

$$\nabla_{\boldsymbol{a}} \boldsymbol{f}(\boldsymbol{a}) = \begin{pmatrix} \partial_{a_1} & \partial_{a_2} & \dots & \partial_{a_n} \end{pmatrix} \otimes \boldsymbol{f}(\boldsymbol{a}).$$

For example,

$$\nabla_{\boldsymbol{r_2}} (\boldsymbol{r}_1 - \boldsymbol{r}_2) = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix} = \begin{pmatrix} 1 \\ -1 \end{pmatrix} \otimes \boldsymbol{I}_3$$

where $\boldsymbol{I}_3 \in \mathbb{R}^{3 \times 3}$ is the identity matrix. Derivations also involve matrix-vector products, and to avoid ambiguity, for a vector $\boldsymbol{a} \in \mathbb{R}^n$ and matrix $\boldsymbol{B} \in \mathbb{R}^{n \times n}$, we define $\boldsymbol{a} \cdot \boldsymbol{B}$ and $\boldsymbol{B} \cdot \boldsymbol{a}$ elementwise by

$$(\boldsymbol{a} \cdot \boldsymbol{B})_i = \sum_{j=1}^{n} a_j B_{ji},$$
$$(\boldsymbol{B} \cdot \boldsymbol{a})_i = \sum_{j=1}^{n} a_j B_{ij},$$

for $i = 1, \dots, n$.

Theorem 2. Let $\boldsymbol{R} \in \mathbb{R}^6$ and $\boldsymbol{V} \in \mathbb{R}^6$, and set $\boldsymbol{R} = [\boldsymbol{r}, \bar{\boldsymbol{r}}], \boldsymbol{V} = [\boldsymbol{v}, \bar{\boldsymbol{v}}]$. Assume that $\bar{\boldsymbol{v}} \neq \boldsymbol{v}$, and set

$$\boldsymbol{N}(\boldsymbol{R}) = \frac{1}{\|\boldsymbol{r} - \bar{\boldsymbol{r}}\|} [\boldsymbol{r} - \bar{\boldsymbol{r}}, \bar{\boldsymbol{r}} - \boldsymbol{r}]. \tag{2.18}$$

Then the following are equivalent:

1. $$\boldsymbol{\sigma}(\boldsymbol{R}, \boldsymbol{V}) = I - \eta(\boldsymbol{R}, \boldsymbol{V}) \boldsymbol{N}(\boldsymbol{R}) \otimes \boldsymbol{N}(\boldsymbol{R}), \tag{2.19}$$

   where $\boldsymbol{A} \otimes \boldsymbol{B}$ is the tensor product, and $\eta(\boldsymbol{R}, \boldsymbol{V}) : \mathbb{R}^{12} \to \mathbb{R}$.

2. The velocities $\bar{\boldsymbol{v}}, \boldsymbol{v}$ and $\bar{\boldsymbol{v}}', \boldsymbol{v}'$ defined by

   $$\begin{bmatrix} \boldsymbol{v}' \\ \bar{\boldsymbol{v}}' \end{bmatrix} = \boldsymbol{\sigma}(\boldsymbol{R}, \boldsymbol{V}) \begin{bmatrix} \boldsymbol{v} \\ \bar{\boldsymbol{v}} \end{bmatrix} \tag{2.20}$$

   satisfy the conservation of linear eq. (2.12) and angular eq. (2.13) momentum.

Proof. We note that eq. (2.13) can be written as $\boldsymbol{A_a}\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V} = \boldsymbol{A_a}\boldsymbol{V}$, where

$$\boldsymbol{A_a} = \begin{pmatrix} 0 & -r_3^a & r_2^a & 0 & -\bar{r}_3^a & \bar{r}_2 \\ r_3^a & 0 & -r_1^a & \bar{r}_3^a & 0 & -\bar{r}_1^a \\ -r_2^a & r_1^a & 0 & -\bar{r}_2^a & \bar{r}_1^a & 0 \end{pmatrix}, \tag{2.21}$$

where $\boldsymbol{r} = (r_1, r_2, r_3)$ and we have written $r_i^a = r_i - a_i$.

- (1 $\implies$ 2). This can be shown by a direct calculation:

$$\begin{aligned} \boldsymbol{A_a}\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V}) &= \boldsymbol{A_a}(\boldsymbol{I} - \eta(\boldsymbol{R},\boldsymbol{V})\boldsymbol{N}(\boldsymbol{R}) \otimes \boldsymbol{N}(\boldsymbol{R})), \\ &= \boldsymbol{A_a} - \eta(\boldsymbol{R},\boldsymbol{V})\boldsymbol{A_a}\boldsymbol{N}(\boldsymbol{R}) \otimes \boldsymbol{N}(\boldsymbol{R}), \\ &= \boldsymbol{A_a}, \end{aligned}$$

  where we have used that $\boldsymbol{A_a}\boldsymbol{N}(\boldsymbol{R}) = (0,0,0)$.

- (2 $\implies$ 1). Before starting the calculation, we note that the matrix $\boldsymbol{A_a}$ is rank 3 (by considering its row-echelon form), and by proposition 1 it is enough to show that eq. (2.13) holds for $\boldsymbol{a}$ determined by the vertices of a polytope, if eq. (2.12) holds.

  We note that $\boldsymbol{A_a}\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V} = \boldsymbol{A_a}\boldsymbol{V}$ for all $\boldsymbol{V} \in \mathbb{R}^6$ implies that $\boldsymbol{A_a}(\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V} - \boldsymbol{V}) = 0$. Therefore $\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V} - \boldsymbol{V} \in \ker(\boldsymbol{A_a})$ for all $\boldsymbol{a} \in \mathbb{R}^6$. Let $\boldsymbol{Y} = (\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V} - \boldsymbol{V})$, then

$$\begin{aligned} -Y_2 r_3^a + Y_3 r_2^a - Y_5 \bar{r}_3^a + Y_6 \bar{r}_2^a = 0, \\ Y_1 r_3^a - Y_3 r_1^a + Y_4 \bar{r}_3^a - Y_6 \bar{r}_1^a = 0, \\ -Y_1 r_2^a + Y_2 r_1^a - Y_4 \bar{r}_2^a + Y_5 \bar{r}_1^a = 0. \end{aligned}$$

  We consider the values

$$\begin{aligned} \boldsymbol{a}^1 &= (r_1, r_2, r_3), \\ \boldsymbol{a}^2 &= (r_1, r_2, \bar{r}_3), \\ \boldsymbol{a}^3 &= (r_1, \bar{r}_2, \bar{r}_3), \\ \boldsymbol{a}^4 &= (\bar{r}_1, \bar{r}_2, \bar{r}_3), \end{aligned}$$

  which form the four vertices of a tetrahedron, and result in

$$\boldsymbol{Y} = \tilde{\eta}(\boldsymbol{R},\boldsymbol{V}) \begin{pmatrix} r_1 - \bar{r}_1 \\ r_2 - \bar{r}_2 \\ r_3 - \bar{r}_3 \\ \bar{r}_1 - r_1 \\ \bar{r}_2 - r_2 \\ \bar{r}_3 - r_3 \end{pmatrix} = \eta(\boldsymbol{R},\boldsymbol{V})\boldsymbol{N}(\boldsymbol{R}), \tag{2.22}$$

  where $\eta(\boldsymbol{R},\boldsymbol{V}) = \|\boldsymbol{r} - \bar{\boldsymbol{r}}\|\tilde{\eta}(\boldsymbol{R},\boldsymbol{V})$, without loss of generality. If we take the dot product on

both sides of eq. (2.22) with $N(\boldsymbol{R})$ and rearrange, we find that

$$\begin{aligned}&\boldsymbol{N}(\boldsymbol{R})\cdot(\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})-\boldsymbol{V})=\eta(\boldsymbol{R},\boldsymbol{V}),\\ \implies&(\boldsymbol{I}-\boldsymbol{N}(\boldsymbol{R})\otimes\boldsymbol{N}(\boldsymbol{R}))\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V}=(\boldsymbol{I}-\boldsymbol{N}(\boldsymbol{R})\otimes\boldsymbol{N}(\boldsymbol{R}))\boldsymbol{V},\end{aligned}$$

i.e. $\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})$ can only change the component of $\boldsymbol{V}$ in the direction of $\boldsymbol{N}(\boldsymbol{R})$. Thus, if $\boldsymbol{N}(\boldsymbol{R})\cdot\boldsymbol{V}=0$, then $\boldsymbol{V}$ is contained in the hyperplane orthogonal to $\boldsymbol{N}(\boldsymbol{R})$, we must have that $\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V}=\boldsymbol{V}$. Therefore, without loss of generality for any $\boldsymbol{V}\in\mathbb{R}^6$ we can take $\eta(\boldsymbol{R},\boldsymbol{V})=\boldsymbol{N}(\boldsymbol{R})\cdot\boldsymbol{V}\tilde{\eta}(\boldsymbol{R},\boldsymbol{V})$, and so

$$\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V}=(I-\eta(\boldsymbol{R},\boldsymbol{V})\boldsymbol{N}(\boldsymbol{R})\otimes\boldsymbol{N}(\boldsymbol{R}))\boldsymbol{V},\tag{2.23}$$

which is as claimed.

□

To determine what form $\eta(\boldsymbol{r}_i,\boldsymbol{r}_j,\boldsymbol{v}_i,\boldsymbol{v}_j)$ takes, we require an additional constraint on the collision rule, determined by the properties of the collision. In addition we can assume that the event map is time-reversible:

$$\boldsymbol{\sigma}^+(\boldsymbol{R},\boldsymbol{V})\boldsymbol{\sigma}^-(\boldsymbol{R},\boldsymbol{V})=\boldsymbol{I}.\tag{2.24}$$

Then we have that

$$\eta^+(\boldsymbol{R},\boldsymbol{V})\eta^-(\boldsymbol{R},\boldsymbol{V})+\eta^+(\boldsymbol{R},\boldsymbol{V})+\eta^-(\boldsymbol{R},\boldsymbol{V})=0.\tag{2.25}$$

We provide three example constraints that produce different collision rules.

Example 3 (Elastic collisions)**.** If in addition we assume that collisions are symmetric in time we have that $\eta^+(\boldsymbol{R},\boldsymbol{V})=\eta^-(\boldsymbol{R},\boldsymbol{V})=:\eta(\boldsymbol{R},\boldsymbol{V})$, we find that

$$\eta(\boldsymbol{R},\boldsymbol{V})=0\text{ or }-2.\tag{2.26}$$

The former result produces the identity matrix; meaning the two particles pass through each other. The latter result exchanges and reflects the part of velocity in the direction of the outward pointing normal. In terms of incoming and outgoing velocities, we have that

$$\begin{aligned}\boldsymbol{v}_i^{\text{out}}&=\boldsymbol{v}_i^{\text{in}}-(\boldsymbol{r}_i-\boldsymbol{r}_j)\cdot(\boldsymbol{v}_i^{\text{in}}-\boldsymbol{v}_j^{\text{in}})\frac{(\boldsymbol{r}_i-\boldsymbol{r}_j)}{\|\boldsymbol{r}_i-\boldsymbol{r}_j\|^2},\\ \boldsymbol{v}_j^{\text{out}}&=\boldsymbol{v}_j^{\text{in}}+(\boldsymbol{r}_i-\boldsymbol{r}_j)\cdot(\boldsymbol{v}_i^{\text{in}}-\boldsymbol{v}_j^{\text{in}})\frac{(\boldsymbol{r}_i-\boldsymbol{r}_j)}{\|\boldsymbol{r}_i-\boldsymbol{r}_j\|^2}.\end{aligned}\tag{2.27}$$

This is the elastic Boltzmann collision rule [92]. We note that as a consequence of making the event map symmetric in time, the event map derived here also conserves kinetic energy:

$$\frac{1}{2}\|\boldsymbol{\sigma}(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V}\|^2=\frac{1}{2}\|\boldsymbol{V}\|^2.\tag{2.28}$$

Example 4 (Inelastic collisions)**.** An important trait of collisions in granular media is that they are *not* energy preserving. As velocity can only be exhanged in the direction of the collision, we must have that the forward-time scalar value $\eta^+(\boldsymbol{R},\boldsymbol{V})$ is a scalar multiple of the backward-time scalar value $\eta^-(\boldsymbol{R},\boldsymbol{V})$, *i.e.*

$$\eta^+(\boldsymbol{R},\boldsymbol{V}) = \alpha\eta^-(\boldsymbol{R},\boldsymbol{V}) \tag{2.29}$$

for $\alpha \in (0,1]$, so that (ignoring the trivial result)

$$\eta^- = -\frac{1+\alpha}{\alpha}, \quad \eta^+ = -(1+\alpha). \tag{2.30}$$

In terms of incoming and outgoing velocities, this is written as the inelastic Boltzmann collision operator:

$$\begin{aligned}
\boldsymbol{v}_i^{\text{out}} &= \boldsymbol{v}_i^{\text{in}} - \frac{1+\alpha}{2}(\boldsymbol{r}_i-\boldsymbol{r}_j)\cdot(\boldsymbol{v}_i^{\text{in}}-\boldsymbol{v}_j^{\text{in}})\frac{(\boldsymbol{r}_i-\boldsymbol{r}_j)}{\|\boldsymbol{r}_i-\boldsymbol{r}_j\|^2},\\
\boldsymbol{v}_j^{\text{out}} &= \boldsymbol{v}_j^{\text{in}} + \frac{1+\alpha}{2}(\boldsymbol{r}_i-\boldsymbol{r}_j)\cdot(\boldsymbol{v}_i^{\text{in}}-\boldsymbol{v}_j^{\text{in}})\frac{(\boldsymbol{r}_i-\boldsymbol{r}_j)}{\|\boldsymbol{r}_i-\boldsymbol{r}_j\|^2}.
\end{aligned} \tag{2.31}$$

It is clear in this case that energy is not preserved, except when taking $\alpha = 1$, when we return to the elastic case. In fig. 2.1 we provide trajectories of two particles in 2 dimensions that experience inelastic collisions for different coefficients of restitution, as well as the dynamics in the reduced-difference space $\boldsymbol{r}_1 - \boldsymbol{r}_2$. The free dynamics of the trajectories are given by the dynamics in example 9.

We note that the average velocity of two particles remains unchanged by an inelastic collision:

$$\frac{1}{2}\left(\boldsymbol{v}_i^{\text{out}} + \boldsymbol{v}_j^{\text{out}}\right) = \frac{1}{2}\left(\boldsymbol{v}_i^{\text{in}} + \boldsymbol{v}_j^{\text{in}}\right). \tag{2.32}$$

Thus the effect of inelasticity is only present in higher order moments.

Example 5 (Discrete step potential systems)**.** The event maps $\boldsymbol{\sigma}^+, \boldsymbol{\sigma}^-$ are not necessarily limited to collisional events. In general, particles can be given multiple interaction diameters, where velocity can be exchanged for potential energy. These interactions can be written in terms of discrete potential forces; we define interaction diameters $\varepsilon_k, k = 0,\dots K \in \mathbb{N}$ such that $0 < \varepsilon_1 < \varepsilon_2 < \cdots < \varepsilon_K$. The pairwise interaction can then be written as a step potential, with an infinite 'core'.

$$V(\boldsymbol{r}_i,\boldsymbol{r}_j) = \begin{cases} 0, & \|\boldsymbol{r}_i-\boldsymbol{r}_j\| > \varepsilon_K,\\ a_{K-1}, & \varepsilon_K > \|\boldsymbol{r}_i-\boldsymbol{r}_j\| > \varepsilon_{K-1},\\ \vdots & \\ a_1, & \varepsilon_1 > \|\boldsymbol{r}_i-\boldsymbol{r}_j\| > \varepsilon_0,\\ \infty, & \|\boldsymbol{r}_i-\boldsymbol{r}_j\| < \varepsilon_0, \end{cases} \tag{2.33}$$

for some $a_1, a_2, \dots, a_{K_1} \in \mathbb{R}$. The simplest example of a discrete step potential used in modelling and simulation is the *square-well* or *square-shoulder* potential, where $K = 1$ and $a_1 < 0$ or $a_1 > 0$ respectively. Step potentials can also be used to approximate more complicated interaction

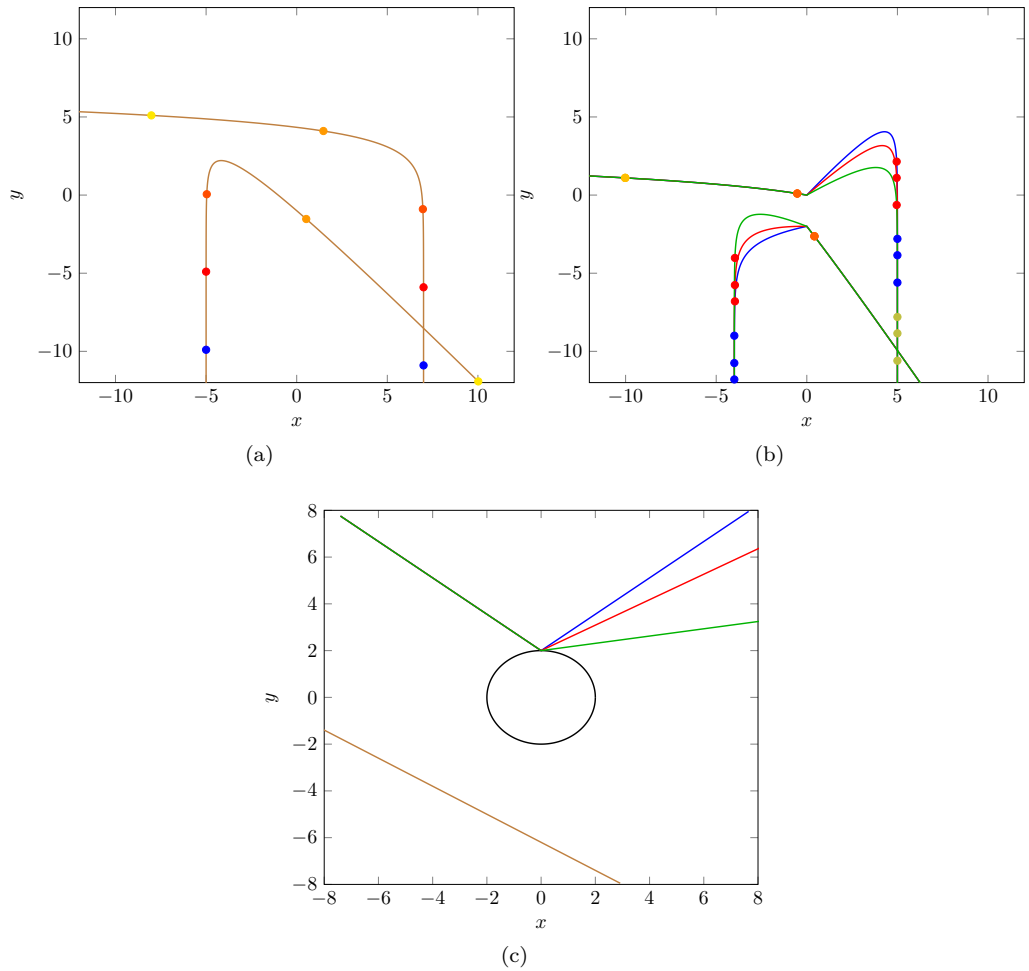


Figure 2.1: Example trajectories of particles that are affected by gravity and friction (see example 9) and collide with different coefficients of restitution α. Figure 2.1a: trajectories of two particles that never collide. Figure 2.1b: trajectories where two particles collide with $\alpha = 1$ (blue), $\alpha = 0.7$ (red) and $\alpha = 0.2$ (green). The coloured dots on each trajectory represent the position at a particular time t. Figure 2.1c shows the same trajectories (using the same colours) considered in the reduced-difference space $(\boldsymbol{r}_1 - \boldsymbol{r}_2)$, where the black circle is the boundary $\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| = \varepsilon$. The reduced-difference space diagram shows the effect of different α on the trajectories of colliding particles.

potentials, for example the *Lennard-Jones* potential [93] which is popular in molecular dynamics:

$$V_{\mathrm{LJ}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = 4\mu \left(\left(\frac{\varepsilon}{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|} \right)^{12} - \left(\frac{\varepsilon}{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|} \right)^{6} \right), \tag{2.34}$$

where μ is the depth of the well present in the potential. We can approximate this with (for example) [94]:

$$\tilde{V}_{\mathrm{LJ}}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \begin{cases} \infty, & \|\boldsymbol{r}_1 - \boldsymbol{r}_2\| < \varepsilon_0, \\ \min_{\varepsilon_i < \|\boldsymbol{r}_1 - \boldsymbol{r}_2\| < \varepsilon_{i+1}} \left(V_{\mathrm{LJ}}(\boldsymbol{r}_1, \boldsymbol{r}_2) \right), & \varepsilon_i < \|\boldsymbol{r}_1 - \boldsymbol{r}_2\| < \varepsilon_{i+1}, i = 1, \ldots, K-1, \end{cases} \tag{2.35}$$

for diameters $0 < \varepsilon_0 < \varepsilon_1 < \cdots < \varepsilon_K$, $K \in \mathbb{N}$. In fig. 2.2 we provide diagrammatic examples of discrete-interaction potentials.

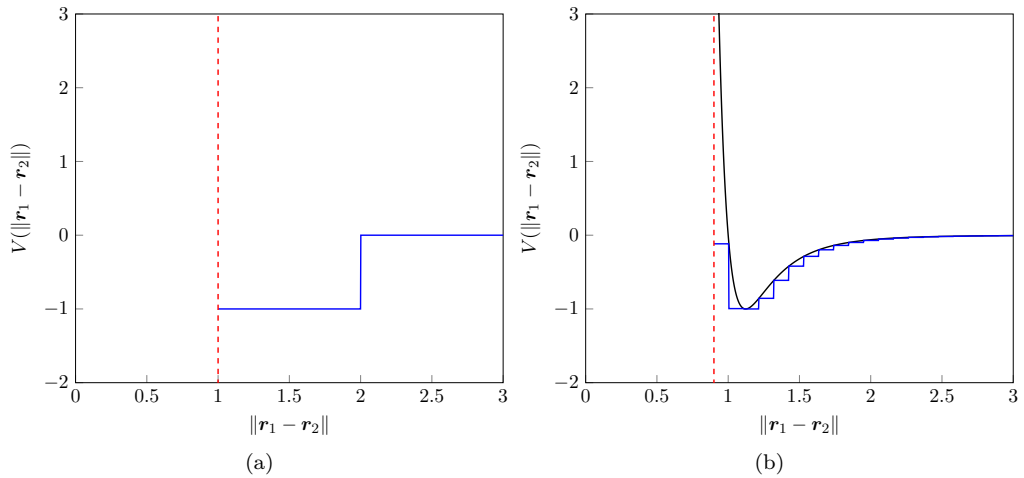


Figure 2.2: Two examples of discrete-interaction potentials. Left: a square-well potential with $\varepsilon_0 = 1, \varepsilon_1 = 1$ and $a_1 = -1$, right: a discrete potential (blue) used to approximate the Lennard-Jones potential (black). In both cases we use a red dashed line to display the hard-particle diameter. (red).

We remark that the direct simulation of systems with many interaction potentials will become prohibitively expensive using the methods discussed in chapter 3. However, they still define a reasonable microscopic model that can be studied, and in chapter 4 we derive a continuum model that accounts for any discrete interactions between particles.

To determine the event maps for this interaction, we require an understanding of what happens when particles reach interaction diameters. When $\|\boldsymbol{r}_i - \boldsymbol{r}_j\|$ reaches an interaction boundary, the effect on the velocities of the particles depends on the energies in the direction of the collision. If the energy is too low to overcome the potential energy barrier, the particles will be reflected, otherwise the velocities of the particles will be changed by the potential energy. These are examples of single interactions, when multiple interactions can occur complications can arise, for example two particles can get stuck in a potential well.

To construct the collision rule for discrete barriers we summarise the events in different cases. Assume that particles i and j are distance $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_i$ from one another. The discrete potential function takes values a_{i+1} for $\boldsymbol{r} > \varepsilon_i$ and a_i for $\boldsymbol{r} < \varepsilon_i$. Define $\Delta a = a_i - a_{i+1}$. Then there are four possible cases:

1. $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j) < 0$ and $\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j))^2$. In this case the two particles are moving towards each other at contact and have enough energy to overcome the potential barrier. Their velocities are *refracted* in the direction of the collision, by an amount proportional to their incoming velocities.

2. $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j) < 0$ and $\Delta a > ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j))^2$. In this case the particles are moving toward one another but do not have the energy to overcome the potential barrier. In this case the velocities of the particles are reflected.

3. $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j) < 0$ and $-\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j))^2$. In this case the two particles are moving away from one other at contact and have enough energy to overcome the potential barrier. Their velocities are *refracted* in the direction of the collision, by an amount proportional to their incoming velocities. We note the addition of a minus sign in the second condition; this is introduced because the particles are moving away from one another.

4. $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j) < 0$ and $-\Delta a > ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j))^2$. In this case the particles are moving away from one another but do not have the energy to overcome the potential barrier. In this case the velocities of the particles are reflected. Again we note the addition of a minus sign in the second condition.

In cases 2 and 4, we use the condition already described in example 3, derived in an analogous way. In cases 1 and 3 we require a new collision rule which provides the correct dynamics; η^+ and η^- should change the speed of colliding particles in a collision, proportional to the energy barrier and the speed of the incoming particles. For these reasons we propose that

$$\eta^+ = -\frac{\|(\boldsymbol{N} \otimes \boldsymbol{N})\boldsymbol{V}\| + \varepsilon_i^2 \Delta a}{\|(\boldsymbol{N} \otimes \boldsymbol{N})V\|}\eta^-. \tag{2.36}$$

Then by using the identity eq. (2.25), we have that:

$$\eta^+ = -\left(1 - \frac{\|(\boldsymbol{N} \otimes \boldsymbol{N})\boldsymbol{V}\| + \varepsilon_i^2 \Delta a}{\|(\boldsymbol{N} \otimes \boldsymbol{N})V\|}\right), \quad \eta^- = -\left(1 - \frac{\|(\boldsymbol{N} \otimes \boldsymbol{N})\boldsymbol{V}\|}{\|(\boldsymbol{N} \otimes \boldsymbol{N})\boldsymbol{V}\| + \varepsilon_i^2 \Delta a}\right).$$

In terms of incoming and outgoing velocities this can be written as

$$\begin{aligned}
\boldsymbol{v}_i^{\text{out}} =& \boldsymbol{v}_i^{\text{in}} - \frac{1}{2}\Bigg[(\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|} \\
&- \text{sign}\left((\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|}\right)\left(((\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|})^2 + \Delta a\right)^{\frac{1}{2}}\Bigg]\frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|}, \\
\boldsymbol{v}_j^{\text{out}} =& \boldsymbol{v}_j^{\text{in}} - \frac{1}{2}\Bigg[(\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|} \\
&- \text{sign}\left((\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|}\right)\left(((\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|})^2 + \Delta a\right)^{\frac{1}{2}}\Bigg]\frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|}.
\end{aligned} \tag{2.37}$$

This fully describes all event types for discrete potential systems. For square-well and square-shoulder systems (where $K = 2$ and $a_2 = 0$) all events can be fully described from their pre-collisional velocities:

1. **Core.** When particles are in contact *i.e.* $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_0$, the particles collide using rule eq. (2.31).

2. **Capture** If $a_1 < 0$, $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_1$ and $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) < 0$, then $\Delta a < 0$ and $\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}))^2$ is immediate, and so particles continue moving towards each other and their velocities are refracted with eq. (2.37).

3. **Release** If $a_1 > 0$, $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_1$ and $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) > 0$, then $\Delta a > 0$ and $-\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i - \boldsymbol{v}_j))^2$ is again immediate, and particles move away from one another after their velocities are updated with eq. (2.37).

4. **Disassociation** If $a_1 < 0$, $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_1$ and $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) > 0$, when $\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}))^2$ then particles are moving away from each other, and continue to do so after their velocities have been updated with eq. (2.37).

5. **Association** If $a_1 > 0$, $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_1$ and $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) < 0$, when $-\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}))^2$ the two particles are moving towards each other and continue to do so, once their velocities have been updated with eq. (2.37).

6. **Internal Bounce** If $a_1 < 0$, $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_1$ and $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) < 0$, when $\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}))^2$ the two particles are moving away from each other but do not have the energy to overcome the energy barrier, so their velocities are reflected using eq. (2.27).

7. **External Bounce** If $a_1 > 0$, $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| = \varepsilon_1$ and $(\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}) < 0$, when $-\Delta a \leq ((\boldsymbol{r}_i - \boldsymbol{r}_j) \cdot (\boldsymbol{v}_i^{\text{in}} - \boldsymbol{v}_j^{\text{in}}))^2$ the two particles are moving toward each other but do not have the energy to overcome the energy barrier, so their velocities are reflected using eq. (2.27).

These interactions are of interest because they are the first approximation to interaction potentials which model attractive forces between particles, and can be used as approximations of general potentials. In fig. 2.3 we provide trajectories of two particles in 2 dimensions that have square-shoulder potential interactions (*i.e.* $a_1 > 0$), and in fig. 2.4a we show how those interactions appear in the reduced-difference space $\boldsymbol{r}_1 - \boldsymbol{r}_2$.

In some simulations it may be appropriate to include collisions between particles and boundaries. An inelastic collision between a particle and a point on the boundary located at $\boldsymbol{r} \in \mathbb{R}^d$ occurs when $\|\boldsymbol{r}_i - \boldsymbol{r}\| = \varepsilon/2$ and is given by

$$\boldsymbol{v}_i^{\text{out}} = \boldsymbol{v}_i^{\text{in}} - \frac{1+\beta}{2}\left(\boldsymbol{v}_i^{\text{in}} \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r})}{\|\boldsymbol{r}_i - \boldsymbol{r}\|}\right)\frac{(\boldsymbol{r}_i - \boldsymbol{r})}{\|\boldsymbol{r}_i - \boldsymbol{r}\|} \tag{2.38}$$

where $\beta \in (0, 1]$ is the coefficient of restitution between particles and walls.

As mentioned earlier, these results can be modified in a trivial manner to account for systems where particles have different masses; the result is equivalent to substituting in the equality $\boldsymbol{v}_i = \boldsymbol{p}_i/m_i$. We can also produce collision rules where the coefficient of restitution depends explicitly on the pair of particles colliding $\alpha_{ij} \in (0, 1]$, or particles of different diameters $\varepsilon_i > 0$, by a straightforward extension of the results above.

2.3 Discussion

In this chapter we have fully described the microscopic (particle) model for the system of interest. The main challenge here was to introduce instantaneous interactions to the Newtonian dynamics in a physically reasonable way. We overcame this challenge by introducing the event map $\boldsymbol{\sigma}(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{v}_i, \boldsymbol{v}_j)$ which maps pre-collisional to post-collisional velocities, and showed that the event map only changes the velocities of the particles in the direction of the collision. Various examples in this section have shown the generality of this result, and we expect that a similar methodology could produce reasonable instantaneous interactions in systems with more degrees of freedom, *e.g.* rotation or charge. As discussed, if instead we were to consider non-instantaneous interactions,

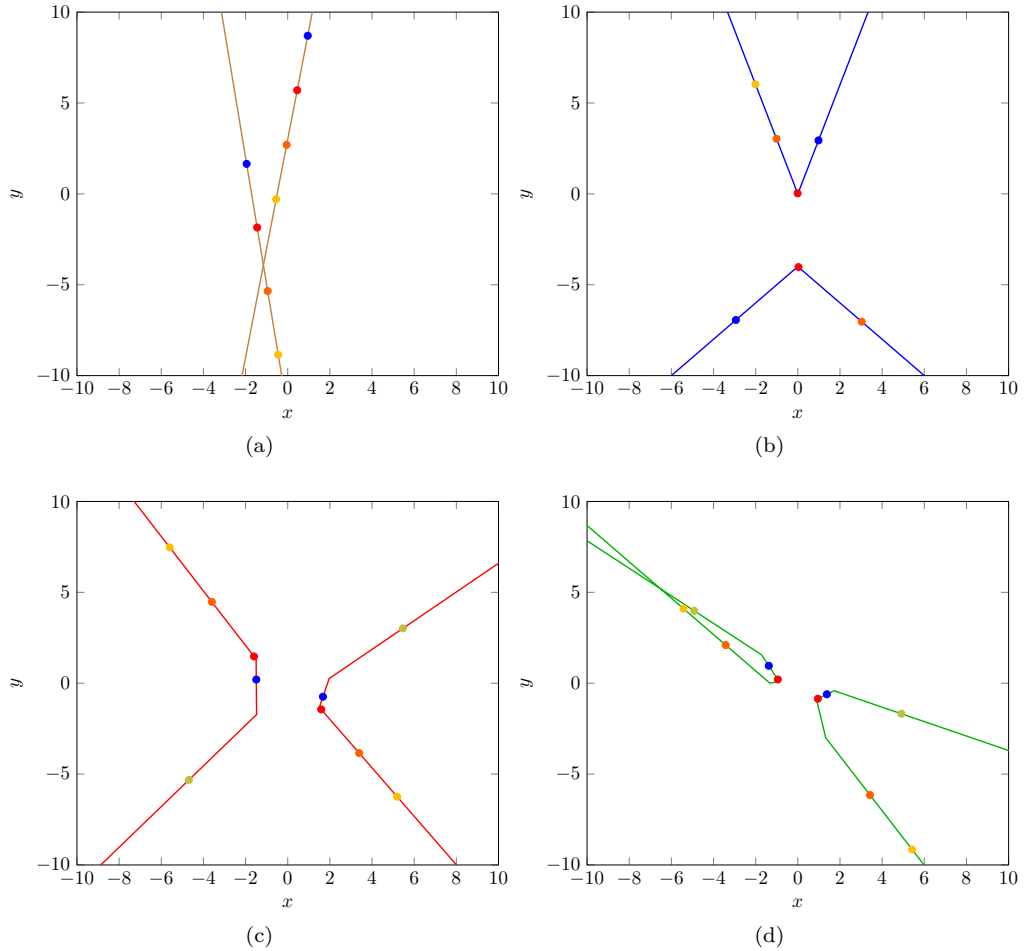
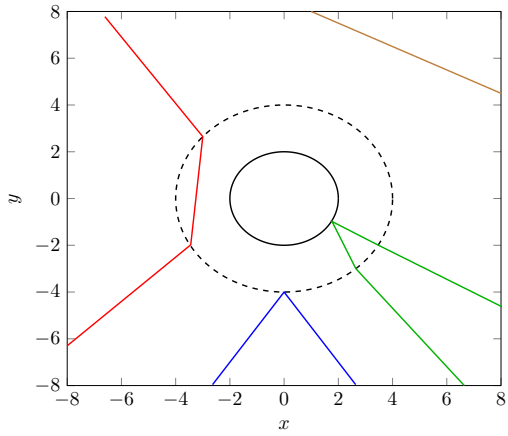


Figure 2.3: Example trajectory pairs between two square-shoulder particles. Figure 2.3a: trajectories of two particles that do not collide. Figure 2.3b: trajectories where two particles experience an elastic collision at the external interaction diameter. Figure 2.3c: the two particles experience two refraction events. Figure 2.3d: the two particles experience a refraction, an elastic colliison with the interior interaction diameter and a further refraction event. Each trajectory has a number of coloured points to display the location of particles at particular times t.



(a)

Figure 2.4: The corresponding reduced-difference representations of the trajectories considered in fig. 2.3. The black, solid line represents the surface given by $\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| = \varepsilon_0$ where elastic collisions take place, an the black, dashed line is the surface $\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| = \varepsilon_1$, where particles are either refracted or reflected.

we would instead consider one of the contact-force models which are commonly used in DEM simulations [95].

The resulting microscopic dynamics are popular in systems where collisions are assumed to be binary and instantaneous. In the next chapter we will see that the instantaneous interaction model discussed here can be efficiently simulated using event-driven particle dynamics, which is a competitive model for simulating systems of inelastic hard spheres.

Chapter 3

Microscopic dynamics - simulation

3.1 Motivation

To simulate molecular dynamics it is popular to use time-stepping methods, where (stochastic) differential equations are discretized in time, and positions and velocities are updated after a discrete timestep $\Delta t > 0$. The fundamental method for deterministic dynamics is the Euler method [96]; given an ODE

$$\frac{\mathrm{d}x(t)}{\mathrm{d}t} = f(x) \tag{3.1}$$

for $t \in [0, T]$, we discretize time into n parts by setting $\Delta t = T/n$, $t_0 = 0$ and $t_i = t_{i-1} + \Delta t$ for $i = 1, \ldots, n$. Then assuming that f is differentiable in x and that Δt is sufficiently small, we have that

$$x(t_i) \approx x_i := x_{i-1} + \Delta t f(x_{i-1}). \tag{3.2}$$

From this simple method a plethora of numerical schemes have been constructed, in an area that is still a thriving topic of modern research. Various numerical schemes have been constructed that utilise state-of-the-art computational power to accurately simulate hundreds of millions of particles. There are higher order schemes that increase the rate of convergence (linked to the finite difference methods discussed in chapter 8). Verlet integration is an example of a second-order method that is commonly used in molecular dynamics due to its numerical stability [21]. Implicit and explicit methods produce better stability or convergence rates for simulating molecular or particle dynamics [97], but do not account for hard-sphere interactions. Further to this, various splitting schemes are available and allow simulation of complex particle or molecular dynamics, that are determined for stochastic differential equations [95], [98] can provide further computation speed up.

Many of these methods are also available in various molecular dynamics software, that are optimised to accurately simulate millions (or even billions) of particles [99]–[101]. Research in numerical methods for molecular dynamics is vast and ever-growing, For more in-depth discussion of popular methods and their mathematical properties, we refer the reader to [21], [102].

Numerical schemes for molecular dynamics must be adapted to account for hard-sphere inter-

actions for our purposes. A direct method that is sometimes considered is to introduce an infinite potential such as eq. (2.10) [103], but this can lead to first order errors in the energy, which renders the simulation unphysical over long times. Furthermore, it is unclear how to correctly include inelastic effects in these simulations.

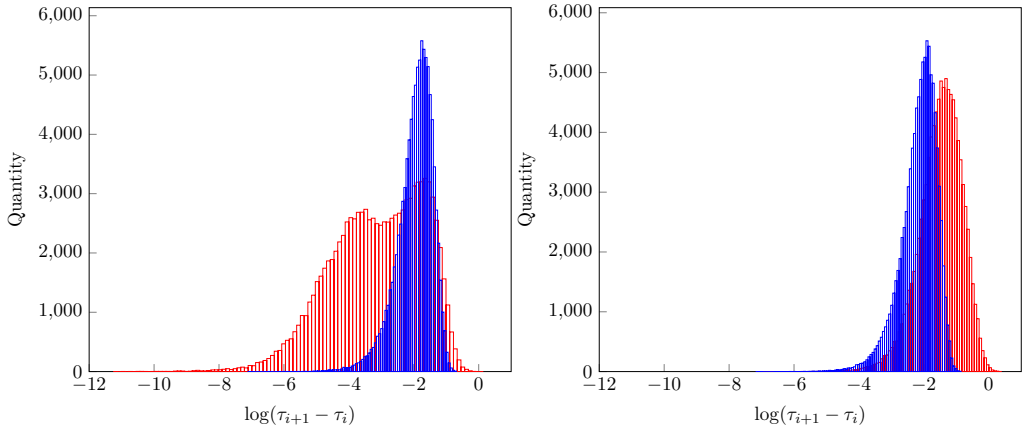
The Verlet method has also been adapted to account for collisional dynamics in [103], which is known as the collision Verlet. This scheme aims to predict future collisions by solving the trajectory approximated by a single Verlet timestep (equivalent to solving a quartic equation), collisions between particles can then be resolved before future collisions are predicted. If no collisions occur then the timestep is taken, and new collisions are predicted at the updated time. This method is in a similar vein to event-driven systems discussed in this chapter, but by generalising the scheme to consider more complicated dynamics it risks a larger computational overhead. We argue that for our purposes we only need to consider simple free dynamics, as we will use these microscopic simulations to approximate local parameters in a continuum equation. We note that, for systems where the mean-free path of individual particles can be accurately approximated using a Verlet timestep, there will be little additional computational cost in using the collision Verlet scheme compared to EDPD.

The most popular modern simulation method for systems of dissipative particles is the Discrete Element Method (DEM) [104], where interactions between particles are modelled as non-instantaneous, and are resolved for each timestep using contact force models. By searching for contacts between particles at each timestep and resolving the contact using a force model, they can incorporate sustained, multi-body interactions between particles. Their ability to model a variety of granular systems [105]–[107] where they can simulate very large collections of particles has made DEM a modern staple in research communities. There are also various open-source and proprietary code libraries that can efficiently simulate hard-particle systems using DEM [108]. Computational fluid dynamics coupled with DEM has also become popular for simulation of fluidised beds [56], and there are also various coarse-graining schemes that reduce computational cost of DEM by modelling agglomerates rather than individual particles [35], at the expense of a loss of some fine-grain accuracy (although approaches have been developed to improve accuracy of coarse-grain systems by for example considering different coarse-grain scales in simulation [36]).

However, it is still challenging to simulate dissipative particles over the timescales of industrial processes using DEM. Furthermore, they are less suited to systems of low density, where interactions between particles can become well-separated in time. The comprehensive neighbour search that DEM models require then becomes computationally restrictive. In our model, as collisions are treated as binary and instantaneous, we do not need to rely on contact force models to determine the effect of particle interactions. Furthermore, as we are interested in parameters that are dependent on the diameter of a single particle, coarse-grain DEM simulations are not appropriate for our systems.

To include collisions in a system of hard inelastic particles under the model discussed in chapter 2, we must use a small timestep to capture collision events. Discrete-time numerical methods for collisional dynamics generally require a check after every timestep to see if particles are overlapping, *i.e.* if there are any $i, j \in \{1, \dots, N\}$ where $i \neq j$ and $\|\boldsymbol{r}_i - \boldsymbol{r}_j\| < \varepsilon$. If an overlap is detected, particle velocities need to be appropriately adjusted to ensure the simulation accurately approximates the collision.

For systems where the density of the particles is uniformly low across the domain and collisions are elastic, timestepping methods can be used effectively to simulate a system of particles. However, when there are areas of high density in a system of hard particles, collisions can occur very quickly after one another, so a very small timestep would be required to capture all events and ensure that events are processed sequentially in time. In fig. 3.1 we show a histogram of times between consecutive collisions in a one and two-dimensional system. The results show that even for relatively simple systems of particles, *e.g.* $d = 1$ and $\alpha = 1$, the time between collisions can cover 5 orders of magnitude, even when excluding rare events (which cover 10 orders of magnitude in the example presented). When $\alpha < 1$ the range of times between collisions increases, and the distribution becomes less trivial. We note that in fig. 3.1a, the case $\alpha = 0.5$ produces a distribution that looks bimodal; this is due to the introduction of the TC method discussed in section 3.5.2. This is not evident in the two-dimensional example, as the number of collisions considered is not sufficient to cause particle streaming in two dimensions, where the TC method will take effect.



(a) $d = 1$ (b) $d = 2$

Figure 3.1: Histograms of times between collisions in 100 samples of system of $N = 100$ obeying linear dynamics with average velocity $\bar{\boldsymbol{v}} = 0$, with periodic boundary conditions, in a system with density $\rho = 0.5$. Left: one dimension, right: two dimensions, red: $\alpha = 0.5$, blue: $\alpha = 1$. In each sample we perform 1000 collisions. The results show that for elastic particles, collisions are separated by a variety of orders of magnitude when accounting for rare events, and when particles are inelastic this distribution widens.

It is therefore preferable to apply a method which does not discretize time independent of collisions, but instead determines the timestep by predicting future collisions between particles. In Event Driven Particle Dynamics (EDPD), dynamics are solved analytically up to a collision, at which point the collision is processed, new collision predictions are made and the dynamics are then advanced to the next collision. There are two requirements that the system dynamics must satisfy to allow for event-driven simulation, or at least an efficient implementation. Firstly, if collisions are neglected the dynamics of the system must be analytic. We will call the dynamics of the system of particles without collisions as the *free dynamics* of the system. Secondly, given the free dynamics of the system, the time at which any two particles collide must be analytically solvable. We define the collision time $\tau : \mathbb{R}^{4dN} \to \mathbb{R}$ as the first time at which two particles collide:

$$\tau(\boldsymbol{r}_i^0, \boldsymbol{r}_j^0, \boldsymbol{v}_i^0, \boldsymbol{v}_j^0) = \arg\min\{t > 0 : \|\boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)\| = \varepsilon\}. \tag{3.3}$$

When particles are different sizes, we replace ε in eq. (3.3) with $\frac{\varepsilon_i+\varepsilon_j}{2}$, where $\varepsilon_i, \varepsilon_j$ are the diameters of particles i and j respectively.

For domains which are not periodic, it is also necessary to have an analytic form for when a particle meets the boundary. If the dynamics occur on the domain $\Omega \subset \mathbb{R}^d$ with boundary $\partial\Omega$:

$$\tau_{\text{wall}}(\boldsymbol{r}_{i,0}, \boldsymbol{v}_{i,0}) = \arg\min\{t > 0 : \|\boldsymbol{r}_i(t) - \boldsymbol{r}\| = \varepsilon/2, \boldsymbol{r} \in \partial\Omega\}. \tag{3.4}$$

In most simulations considered in this work we use periodic boundary conditions, in simulations where this is not true we only consider flat walls, which generally simplifies the computation. A popular method for simulating more complicated domains is to construct the boundary of the domain using particles with fixed positions.

In principle these restrictions can be relaxed; given a numerical solution of the free dynamics intersections of individual trajectories can be computed. However, implementations of event-driven particle dynamics are much more efficient when analytic forms of the free dynamics and event times are available.

3.2 Dynamics with analytic solutions

The requirements that the dynamics must satisfy to use EDPD are quite restrictive, however we shall see that for our purposes they are sufficient for accurate approximation. Here we provide examples of microscopic dynamics which satisfy the requirements. We introduce some notation to reduce complexity: for particle quantities $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$, we write $\boldsymbol{x}_{ij} = \boldsymbol{x}_i - \boldsymbol{x}_j$. For ease of notation we will set $m_i = 1$ for all particles.

Example 6 (Linear dynamics)**.** In a system where

$$\frac{\mathrm{d}\boldsymbol{r}^N(t)}{\mathrm{d}t} = \boldsymbol{v}^N(t), \quad \frac{\mathrm{d}\boldsymbol{v}^N(t)}{\mathrm{d}t} = 0, \tag{3.5}$$

the free dynamics of an individual particle is then given by

$$\boldsymbol{r}_i(t) = \boldsymbol{v}_{i,0}t + \boldsymbol{r}_{i,0}, \quad \boldsymbol{v}_i(t) = \boldsymbol{v}_{i,0}. \tag{3.6}$$

We can then determine the collision time of any two particles by solving a quadratic polynomial in t:

$$\begin{aligned}\tau(\boldsymbol{r}_{i,0}, \boldsymbol{r}_{j,0}, \boldsymbol{v}_{i,0}, \boldsymbol{v}_{j,0}) = \min\Bigg\{t = &\frac{1}{\|\boldsymbol{v}_{ij,0}\|}\Bigg[-\widehat{\boldsymbol{v}_{ij,0}} \cdot \boldsymbol{r}_{ij,0} \\ &\pm \left[(\widehat{\boldsymbol{v}_{ij,0}} \cdot \boldsymbol{r}_{ij,0})^2 - (\|\boldsymbol{r}_{ij,0}\|^2 - \varepsilon^2)\right]^{\frac{1}{2}}\Bigg] : t > 0\Bigg\}\end{aligned} \tag{3.7}$$

and the collision time of particles with the boundary of the domain is given by

$$\begin{aligned}\tau_{\text{wall}}(\boldsymbol{r}_{i,0}, \boldsymbol{v}_{i,0}) = \min\Big\{t = &-\widehat{\boldsymbol{v}_{i,0}} \cdot (\boldsymbol{r}_{i,0} - \boldsymbol{r}) \pm \\ &\left[(\widehat{\boldsymbol{v}_{i,0}} \cdot (\boldsymbol{r}_{i,0} - \boldsymbol{r}))^2 - (\|\boldsymbol{r}_{i,0} - \boldsymbol{r}\|^2 - \varepsilon^2/4)\right]^{\frac{1}{2}}, t > 0, \boldsymbol{r} \in \partial\Omega\Big\}\end{aligned} \tag{3.8}$$

where $\widehat{\boldsymbol{x}} = \boldsymbol{x}/\|\boldsymbol{x}\|$. We note that these dynamics are independent of particle mass. An example of hard-particle dynamics which have free dynamics governed by eq. (3.5) was considered in fig. 2.3.

Example 7 (Constant external potential)**.** We can include simple external effects such as gravity by including a constant external force:

$$\frac{\mathrm{d}\boldsymbol{r}^N(t)}{\mathrm{d}t} = \boldsymbol{v}^N(t), \quad \frac{\mathrm{d}\boldsymbol{v}^N(t)}{\mathrm{d}t} = -\boldsymbol{G}^N, \tag{3.9}$$

where $\boldsymbol{G}^N = [\boldsymbol{G}, \boldsymbol{G}, \ldots, \boldsymbol{G}] \in \mathbb{R}^{dN}$ $\boldsymbol{G} \in \mathbb{R}^d$. In this case the free dynamics are given by

$$\boldsymbol{r}^N(t) = -\frac{\boldsymbol{G}^N t^2}{2} + \boldsymbol{v}_0^N t + \boldsymbol{r}_0^N, \quad \boldsymbol{v}^N(t) = -\boldsymbol{G}^N t + \boldsymbol{v}_0^N. \tag{3.10}$$

In this case for particles of the same mass the collision times between particles are given by eq. (3.7), as it only depends on the relative positions of two particles. For interactions with the boundary, the collision time is given by the smallest positive value of a quartic:

$$\begin{aligned}\tau_{\text{wall}}(\boldsymbol{r}_{i,0}, \boldsymbol{v}_{i,0}) = &\min\{t : t > 0 \text{ and } \|\boldsymbol{G}\|t^4 + \boldsymbol{G} \cdot \boldsymbol{r}_{i,0} t^3 \\ &+ (g \cdot (\boldsymbol{r}_{i,0} - \boldsymbol{r}) + \|\boldsymbol{v}_{i,0}\|^2)t^2 + 2\boldsymbol{v}_{i,0} \cdot (\boldsymbol{r}_{i,0} - \boldsymbol{r})t + \|\boldsymbol{r}_{i,0} - \boldsymbol{r}\|^2 = \varepsilon^2, \boldsymbol{r} \in \partial\Omega\}\end{aligned} \tag{3.11}$$

More generally, we can include a discrete external potential, introducing new events at each interaction diameter. We note that the efficiency of EDPD algorithms is directly related to the number of events that must be predicted, so systems of particles which have multiple interaction diameters may become computationally intractable.

Example 8 (Viscous dynamics)**.** We may also include a viscous term:

$$\frac{\mathrm{d}\boldsymbol{r}^N(t)}{\mathrm{d}t} = \boldsymbol{v}^N(t), \quad \frac{\mathrm{d}\boldsymbol{v}^N(t)}{\mathrm{d}t} = -\gamma\boldsymbol{v}^N(t), \tag{3.12}$$

for $\gamma \in \mathbb{R}$. In this case the free dynamics are given by

$$\boldsymbol{r}^N(t) = \frac{\boldsymbol{v}_0^N}{\gamma}(1 - e^{-\gamma t}) + \boldsymbol{r}_0^N, \quad \boldsymbol{v}^N(t) = \boldsymbol{v}_0^N e^{-\gamma t}, \tag{3.13}$$

and the collision times are given by

$$\begin{aligned}\tau(\boldsymbol{r}_{i,0}, \boldsymbol{r}_{j,0}, \boldsymbol{v}_{i,0}, \boldsymbol{v}_{j,0}) = \Bigg\{ t = -\frac{1}{\gamma}\log\Bigg(1 - \frac{\gamma}{\|\boldsymbol{v}_{ij,0}\|}\Bigg[&- \boldsymbol{r}_{ij,0} \cdot \widehat{\boldsymbol{v}_{ij,0}} \\ &\pm \left[(\boldsymbol{r}_{ij,0} \cdot \widehat{\boldsymbol{v}_{ij,0}})^2 - (\|\boldsymbol{r}_{ij,0}\|^2 - \varepsilon^2)\right]^{\frac{1}{2}}\Bigg]\Bigg) : t > 0\Bigg\},\end{aligned} \tag{3.14}$$

$$\begin{aligned}\tau_{\text{wall}}(\boldsymbol{r}_{i,0}, \boldsymbol{v}_{i,0}) = \min\Bigg\{ t = -\frac{1}{\gamma}\log\Bigg(1 - \frac{\gamma}{\|\boldsymbol{v}_{i,0}\|}\Bigg[&- (\boldsymbol{r}_{i,0} - \boldsymbol{r}) \cdot \widehat{\boldsymbol{v}_{i,0}} \\ &\pm \left[((\boldsymbol{r}_{i,0} - \boldsymbol{r}) \cdot \widehat{\boldsymbol{v}_{i,0}})^2 - (\|\boldsymbol{r}_{i,0} - \boldsymbol{r}\|^2 - \varepsilon^2/4)\right]^{\frac{1}{2}}\Bigg]\Bigg) : t > 0, \boldsymbol{r} \in \partial\Omega\Bigg\}.\end{aligned} \tag{3.15}$$

We note that in this case the viscous term is constant, but one can extend the results here to approximate non-constant γ, by modifying the value of γ after every collision. This should produce

reasonably accurate results, provided that γ is approximately constant in the timescale of the mean-free path of a particle.

Example 9 (Viscous dynamics with constant external force)**.** We may also combine eq. (3.9) and eq. (3.12) to include viscous and gravitational effects:

$$\frac{\mathrm{d}\boldsymbol{r}^N(t)}{\mathrm{d}t} = \boldsymbol{v}^N(t), \quad \frac{\mathrm{d}\boldsymbol{v}^N(t)}{\mathrm{d}t} = -\boldsymbol{G}^N - \gamma \boldsymbol{v}^N(t), \tag{3.16}$$

in which case the trajectories of the particles under free dynamics are given by

$$\boldsymbol{r}^N(t) = \frac{1}{\gamma}\left(\boldsymbol{v}_0^N + \frac{\boldsymbol{G}^N}{\gamma}\right)(1 - e^{-\gamma t}) - \frac{\boldsymbol{G}^N t}{\gamma} + \boldsymbol{r}_0^N, \quad \boldsymbol{v}^N(t) = \left(\boldsymbol{v}_0^N + \frac{\boldsymbol{G}^N}{\gamma}\right)e^{-\gamma t}. \tag{3.17}$$

These are the free dynamics that were considered in fig. 2.1a. Again we see that, as collision times only depend on the relative position of particles, the introduction of a constant external potential has no effect on the particle collision time and is given by eq. (3.14). However, it is no longer possible to write out an analytic form for the collision time with a point on the boundary; the equation to solve has a mixture of exponentials and polynomials in time. In one dimension, the solution can be written in terms of the product-log function (also known as the Lambert-$\mathcal{W}$ function [109]). The collision time with the boundary is given by

$$\begin{aligned}\tau_{\mathrm{wall}}(r_{i,0}, v_{i,0}) = \min\Bigg\{ & t = -\frac{1}{G}\left(r_{i,0} - r + \frac{1}{\gamma}\left(v_{i,0} + \frac{G}{\gamma}\right)\right) \\ & + \frac{1}{\gamma}\mathcal{W}_0\left[-\left(\frac{v_{i,0}}{G} + \frac{1}{\gamma}\right)\exp\left(-\frac{\gamma}{G}\left(r_{i,0} + \frac{1}{\gamma}\left(v_{i,0} + \frac{G}{\gamma}\right) - r\right)\right)\right] : t > 0, r \in \partial\Omega\Bigg\},\end{aligned} \tag{3.18}$$

where $\mathcal{W}_0(x)$ is the principal branch of the Lambert-$\mathcal{W}$ function, *i.e.* $\mathcal{W}_0(x)$ is the real solution to the equation

$$ye^y = x. \tag{3.19}$$

There is no analytic form for the Lambert-$\mathcal{W}$ function, but numerical approximations to machine precision are available, and are implemented in many coding languages.

We now provide a review of algorithms in event-driven particle dynamics, and methods considered to improve efficiency of simulation.

3.3 A naive algorithm

Given the analytic form of the free particle dynamics and collision times for particles, one can construct a straightforward algorithm to simulate a collection of hard spherical particles. We outline the algorithm as follows:

Algorithm 10.

1. Initialise a system of N particles with positions $\boldsymbol{r}_0^N \in \mathcal{P}_\varepsilon$ and velocities $\boldsymbol{v}_0^N \in \mathbb{R}^{dN}$.
2. Using the analytic expression of the free dynamics, determine the collision times τ_{ij} and τ_i^{wall} for particles $i, j \in \{1, \ldots, N\}$ (*event prediction*).

3. Find the event $\tilde{\tau} = \min\{\tau_{ij}, \tau_{ij}^{\text{wall}}, i, j = 1, \dots, N\}$ which happens first (*event scheduling*).

4. Advance the simulation up to time $\tilde{\tau}$, using the analytic expression of the dynamics (*advancing free dynamics*).

5. Apply the collision rule to the particle(s) involved in the event (*event processing*).

6. Repeat steps (2-5) until a stopping criterion (*e.g.* a time $T > 0$) is reached.

This algorithmic structure is fundamental to all modern EDPD algorithms, however as a stand-alone algorithm it is very expensive; each of steps (2-4) is an $\mathcal{O}(N^2)$ calculation. This becomes problematic when the number of particles is large, or when the number of collisions is large, which could occur in areas of a system with high density.

3.4 Improvements in efficiency

We now introduce various methods which can improve the efficiency of an EDPD algorithm. The methods can be separated into three types: advancing free dynamics, event prediction and event scheduling. We shall see that the cost of computation can be reduced from $\mathcal{O}(N^2)$ to an effective cost of $\mathcal{O}(1)$ per collision.

3.4.1 Advancing free dynamics

To improve the efficiency of step 2 in algorithm 10, we can make the dynamics *asynchronous* [110]. This involves giving each particle a *local time* $t_i \in \mathbb{R}$, so that at any given time in the simulation the initial condition of all particles will be given by $\boldsymbol{r}_i(t_i), \boldsymbol{v}_i(t_i)$, where t_i are the times of the last event the particle was involved in.

By making the dynamics asynchronous, we only need to update the positions and velocities of particles involved in the next scheduled event, thus reducing the cost to $\mathcal{O}(1)$. The caveat is that the t_i must be accounted for in the collision time calculations. Revisiting the examples of analytical dynamics we find that (where quantities $\boldsymbol{r}_i\boldsymbol{r}_j, \boldsymbol{v}_i, \boldsymbol{v}_j$ are given at times t_i, t_j):

$$\begin{aligned}\tau_{ij} = \min\Bigg\{ t = \frac{1}{\|\boldsymbol{v}_{ij}\|}\Bigg[&- (\boldsymbol{r}_{ij} - (\boldsymbol{v}_i t_i - \boldsymbol{v}_j t_j)) \cdot \boldsymbol{v}_{ij} \\ &\pm \Big[((\boldsymbol{r}_{ij} - (\boldsymbol{v}_i t_i - \boldsymbol{v}_j t_j)) \cdot \boldsymbol{v}_{ij})^2 - (\|\boldsymbol{r}_{ij} - (\boldsymbol{v}_i t_i - \boldsymbol{v}_j t_j)\| - \varepsilon^2)\Big]^{\frac{1}{2}}\Bigg] : t > 0\Bigg\},\end{aligned} \tag{3.20}$$

$$\begin{aligned}\tau_{ij} = \min\Bigg\{ t = \frac{1}{\|\boldsymbol{v}_{ij} + \boldsymbol{G}t_{ij}\|} \times \\ \Bigg[&- (-\boldsymbol{G}t_i^2/2 - \boldsymbol{v}_i t_i + \boldsymbol{r}_i + \boldsymbol{G}t_j^2/2 + \boldsymbol{v}_j t_j - \boldsymbol{r}_j) \cdot \widehat{(\boldsymbol{v}_{ij} + \boldsymbol{G}t_{ij})} \\ &\pm \big\{((-\boldsymbol{G}t_i^2/2 - \boldsymbol{v}_i t_i + \boldsymbol{r}_i + \boldsymbol{G}t_j^2/2 + \boldsymbol{v}_j t_j - \boldsymbol{r}_j) \cdot \widehat{(\boldsymbol{v}_{ij} + \boldsymbol{G}t_{ij})})^2 \\ &\qquad - (\| - \boldsymbol{G}t_i^2/2 - \boldsymbol{v}_i t_i + \boldsymbol{r}_i + \boldsymbol{G}t_j^2/2 + \boldsymbol{v}_j t_j - \boldsymbol{r}_j\|^2 - \varepsilon^2)\big\}\Bigg], t > 0\Bigg\},\end{aligned} \tag{3.21}$$

$$\tau_{ij} = \min\left\{ t = -\frac{1}{\gamma}\left[\frac{1}{\|\boldsymbol{v}_i e^{\gamma t_i} - \boldsymbol{v}_j e^{\gamma t_j}\|} \log\left[- (\boldsymbol{r}_{ij} + \boldsymbol{v}_{ij}/\gamma) \cdot \widehat{(\boldsymbol{v}_i e^{\gamma t_i} - \boldsymbol{v}_j e^{\gamma t_j})}\right.\right.\right. \tag{3.22}$$

$$\pm\left[((\boldsymbol{r}_{ij}+\boldsymbol{v}_{ij}/\gamma)\cdot(\boldsymbol{v}_i e^{\gamma t_i}-\widehat{\boldsymbol{v}_j e^{\gamma t_j}}))^2-(\|\boldsymbol{r}_{ij}+\boldsymbol{v}_{ij}/\gamma\|-\varepsilon^2)\right]\Bigg],t>0\Bigg\},$$

$$\tau_{ij}=\min\Bigg\{t=-\frac{1}{\gamma}\log\Bigg[\frac{\gamma}{\|\boldsymbol{v}_{ij}\|}\Bigg(\frac{1}{\gamma}\left(\boldsymbol{r}_{ij}+\frac{\boldsymbol{G}}{\gamma}(t_i-t_j)+\frac{1}{\gamma}\boldsymbol{v}_{ij}\right)\cdot(\boldsymbol{v}_{ij}) \tag{3.23}$$

$$\pm\Bigg[\frac{1}{\gamma^2}\left(\left(\boldsymbol{r}_{ij}+\frac{\boldsymbol{G}}{\gamma}(t_i-t_j)+\frac{1}{\gamma}\boldsymbol{v}_{ij}\right)\cdot(\boldsymbol{v}_{ij})\right)^2$$

$$-\left(\left\|\boldsymbol{r}_{ij}+\frac{\boldsymbol{G}}{\gamma}(t_i-t_j)+\frac{1}{\gamma}\boldsymbol{v}_{ij}\right\|-\varepsilon^2\right)\Bigg]\Bigg)\Bigg],t>0\Bigg\},$$

for examples 6 to 9 respectively. As solutions are still analytic, these calculations can still be computed quickly.

3.4.2 Event prediction

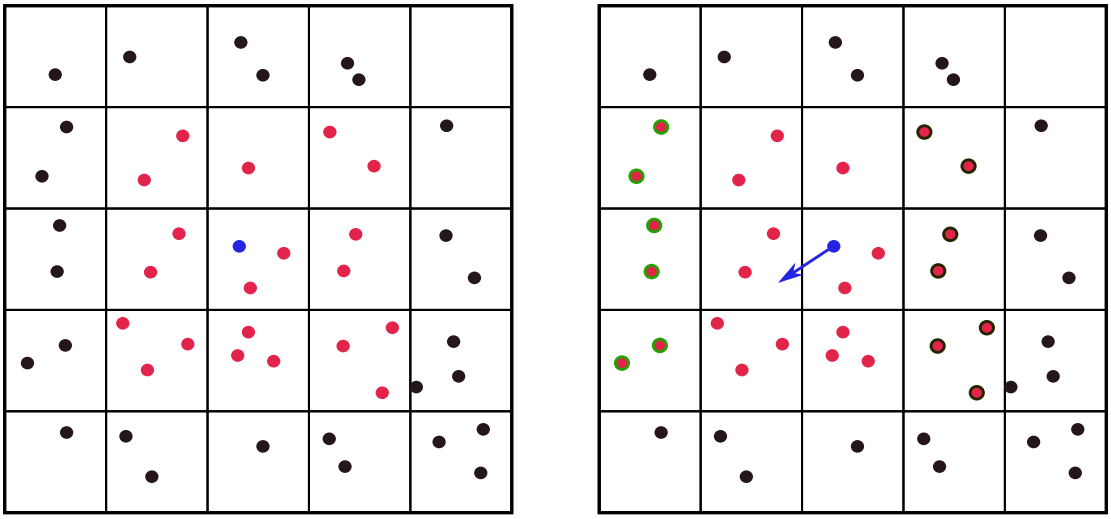
To improve the efficiency of the event prediction step, we want to reduce the number of predictions made after each collision. This should be possible; we would expect that most collision times are unaffected by the occurrence of a single binary collision.

However, due to the chaotic nature of particle collisions, it can be very hard to determine which events will still be valid after a collision has occurred. In the first instance we therefore need to recalculate all possible future events involving the particles involved in the current event, reducing the number of calculations to $\mathcal{O}(N)$. In fact, the number of events to be predicted will be $N-1+C$, where C is the number of boundary events that must be predicted.

To reduce the computation cost of event prediction we split the domain into *cells* [25]. We consider a simple square domain split in to n^d cells of equal size, for $n\in\mathbb{N}$. Consider the prediction of events for a particle i. Instead of predicting future collisions with all other particles, we only make collision predictions for particles in *neighbouring cells*, including the particle cell itself (fig. 3.2a). For d dimensions there are then 3^d neighbouring cells, and the number of predictions per particle then depends on the number of particles in neighbouring cells. As particle size and cell volume is fixed, a finite number of particles can fit in neighbouring cells, independent of N. By making the neighbouring cells sufficiently small we can minimise the number of collision predictions for a particle. However, there is also a chance that before any collision with neighbouring particles occur, particle i will move to a different cell. We must therefore introduce a new event for prediction, a *cell transfer*. After a cell transfer new predictions for collisions with new neighbours have to be made, and predictions with old neighbours may be discarded (see figure 3.2b).

The introduction of the cell method requires us to keep track of an additional parameter per particle: the cell the particle is located in. From there we can determine the neighbouring cells and therefore the neighbouring particles.

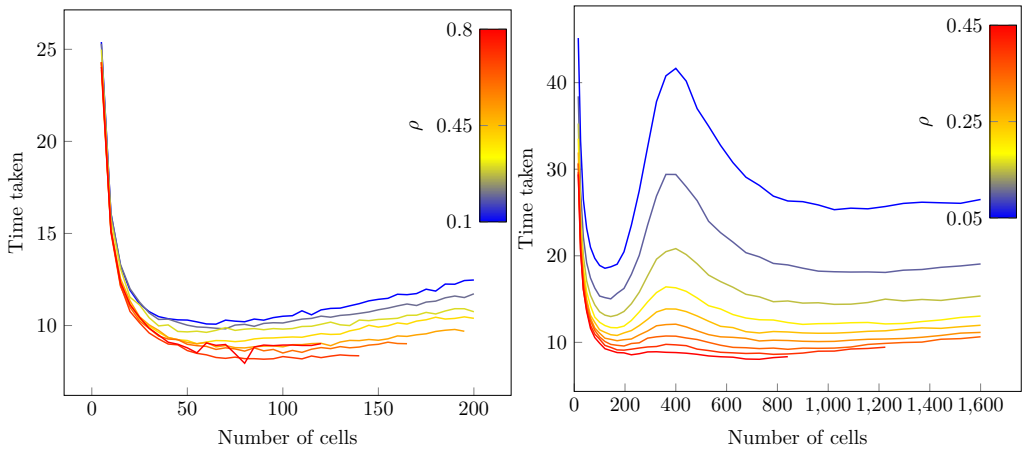
We also need to consider the ideal cell size $C>0$ for a given simulation. Depending on the density of the system, there are different optimal choices of C. The minimum size of a cell is determined by the size of the largest particle in the system. In [111] it is stated that for high densities, the number of cells should be proportional to the number of particles, so that each cell contains at most one particle. For systems of lower density fewer cells are more appropriate to reduce the number of cell transfer events. In fig. 3.3 we show how computational time changes depending on the particle density and the number of cells. In fig. 3.3b we see that the optimal



(a) (b)

Figure 3.2: Definition and recycling of neighbours in the cell method. Left: We have split the domain into cells. Cell neighbours for a given particle (blue) are coloured red, right: new cell neighbours (circled green) and old cell neighbours (circled black) after the blue particle has transferred cell.

choice of the number of cells is non-trivial for low densities, but by choosing a system with a large number of cells we can generally improve the efficiency of the simulation.



(a) $d = 1$ (b) $d = 2$

Figure 3.3: Computation times for different densities when using a different number of cells in the simulation. In these simulations we consider systems of $N = 100$ particles, and simulate 5000 collisions. The simulations were repeated 100 times with different initial conditions, and the results were averaged. We see that there is an optimal choice for the number of cells in each case, but in general by choosing a large number of cells we increase computational efficiency.

We note that there are various cell methods for systems where the standard cell method is not efficient; for example systems with high polydispersity. In this case many authors make the cell method adaptive with a quadtree or octree structure, as discussed in [111]. The alternative approach also discussed in [111] for polydisperse (or non-spherical particle) systems is to add an

additional diameter to each particle; when particles are within this perimeter of the particle, their collision times are predicted. The additional overhead for this method is then the storage and maintenance of *neighbour lists* (also known as *Verlet lists* in molecular dynamics [112]) for each particle.

By updating the positions of particles after a cell transfer, the resulting dynamics will depend on the number of cells, due to machine-precision errors in the summation of distances after cell transfers occur. This effect is amplified by the chaotic nature of the collision rule. A diagram showing this is given in fig. 3.4, where two particles are moving towards each other and colliding. The trajectory which is updated after every cell transfer picks up floating point error differences that change the position at which the particles collide. In the diagram the machine-precision errors have been exaggerated to show the effect. Although the difference in the trajectories is small, after many collisions have occurred the resulting particle positions can be completely different.

Therefore to ensure that the result of the dynamics is independent of the number of cells, we choose not to update the local time of a particle after a cell transfer, instead we label each particle with its cell and a list of neighbouring particles, and update the label after a cell transfer, then update its list of neighbouring particles.

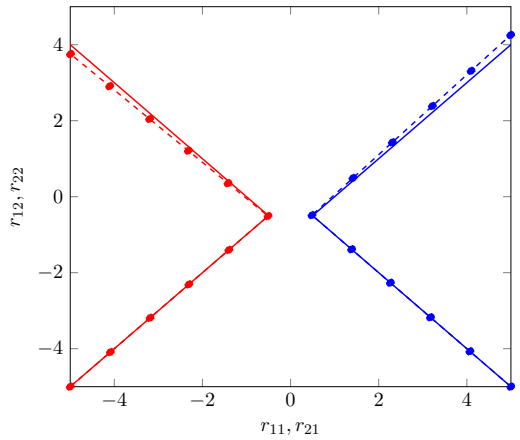


Figure 3.4: A trajectory of two particles in two dimensions, where the cell method is (dashed) implemented, and particle trajectories are updated for every cell transfer, given by marked points on the trajectory. Machine precision differences have been exaggerated to make the result clearer. The solid line is the trajectory for when the cell method is not implemented. We can remove these small changes in particle trajectories by only updating position of particles when they experience collisions.

3.4.3 Event scheduling

Once all events are predicted, we must determine which event comes first. After applying the cell method, for each particle we have a constant number of predictions, which depends on the cell size. Therefore, the list of events is of size N, consisting of the first event for each particle. To reduce the cost of finding the minimum value in this list we can implement a data structure. The data structure must be able to support the following operations efficiently:

1. Identify the first event in the list,

2. Remove any events that are no longer valid after a collision has occurred,

3. Add any new events that have been calculated after a collision.

These three qualities define a *priority queue*. There are various implementations of priority queues [72], we focus on binary search trees and bounded increasing priority queues.

Binary search trees

We define *binary trees* [113] as N *nodes* connected by $N-1$ *left and right branches* at different *levels*. We define the nodes which have no lower branches to be the *leaves*. The *children* of a node i are the nodes connected to it on the next level down, the *parent* of i is the node connected on the next level up. These definitions are depicted in figure 3.5.

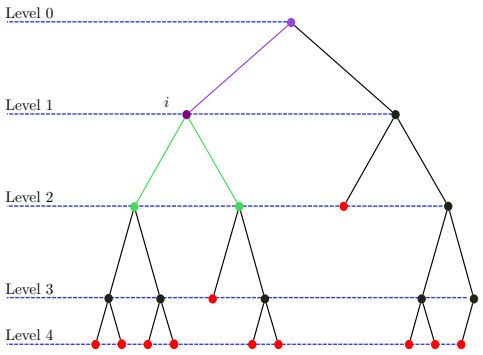


Figure 3.5: An example of a binary tree. Nodes are represented by points, branches are lines connected by points, and leaves are coloured red. The parent of node i is coloured purple, and the children of i are green.

To construct a binary tree, given a list of events L constructed at the initialization stage, we use the following algorithm. An instance of the algorithm is given in figure 3.6.

1. Remove an event $(t_i, \boldsymbol{b}_i)$ from L, place it at node 0,
2. Remove an event $(t_j, \boldsymbol{b}_j)$ from L. Start at position 0, set $n = 0$.
3. We consider 2 cases
   (a) $t_j < t_i$. If there is no left branch from node n, construct a left branch connecting a new node $n+1$ and place event $(t_j, \boldsymbol{b}_j)$ there. If there is a node at position $n+1$, set $n \to n+1$, and repeat step (2).
   (b) $t_j > t_i$. If there is no left branch from node n, construct a left branch connecting a new node $n+2$ and place event $(t_j, \boldsymbol{b}_j)$ there. If there is a node at position $n+1$, set $n \to n+2$, and repeat step (2).
4. Repeat (2)-(3) as necessary.

The first operation we use on the binary tree is a *search* for a particular element, shown in figure 3.7. To locate the event t_i, we start at node $n = 0$ and travel left or right by comparing the event time at node n to t_i, until we reach the event. This will require on average $\mathcal{O}(\log_2(N))$ checks.

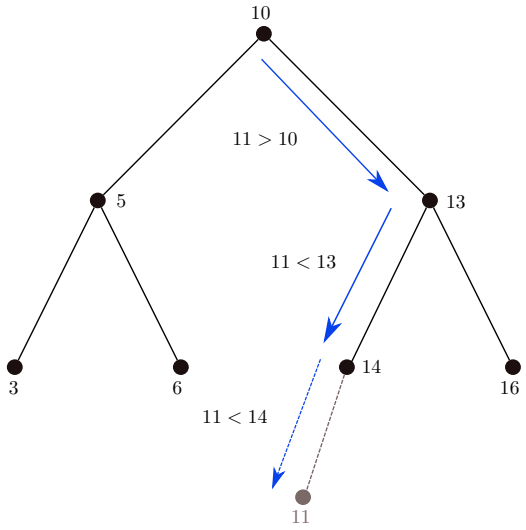


Figure 3.6: An example of an insert in a binary tree, inserting a node with value 11, by following the blue arrows in the diagram. This algorithm can also be used for construction.

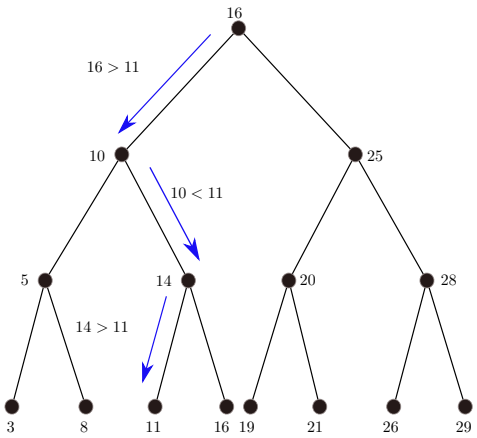


Figure 3.7: An example of a search in a binary tree. We search for a node with value 11.

In the worst case, the element will be one of the leaves. If construction was optimal, each node has two branches so there are m leaves, such that

$$N \approx \sum_{i=1}^{m} 2^i, \implies N \approx 2^{m-1}$$
$$\implies m = \mathcal{O}(\log_2(N)).$$

Therefore in the optimal construction for the worst case a search costs $\mathcal{O}(\log_2(N))$. Of course, the tree may not be in this optimal shape; in the most pathological case there can be N leaves. There are restructuring algorithms to avoid these cases, however in EDPD computations the structure of the tree tends to be maintained naturally.

For *addition* of an event to the binary tree, we simply use the construction algorithm defined above. In the optimal case, this would involve traversing $m/2$ levels, so therefore the operation is $\mathcal{O}(\log_2(N))$ on average.

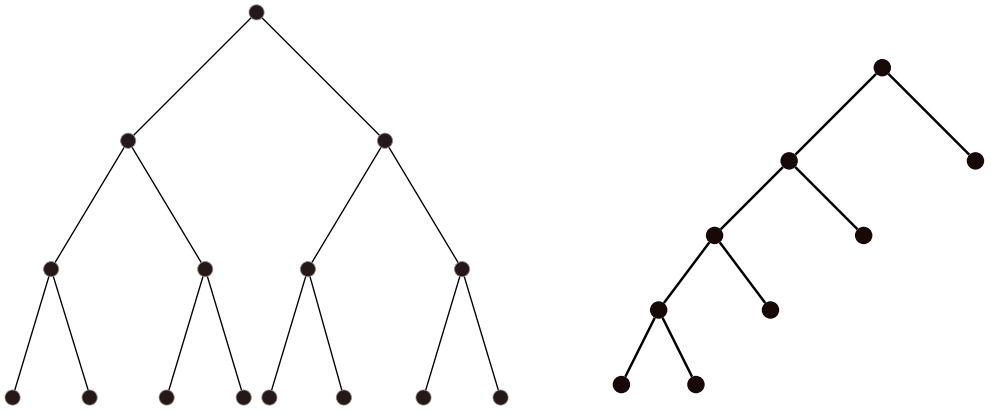


Figure 3.8: Examples of binary trees. Left: An example of a well structured binary tree. Right: An example of a poorly structured binary tree; by restructuring the list of events we can avoid such pathological examples.

For *removal* of an event, we need to locate the event in the tree, then restructure the tree. Restructuring depends on the number of children the node for t_i has. There are three cases. If the event is a leaf, removal is straightforward. For one child, the single child event takes the place of t_i. If there are two children, the minimum $t_{\min}$ of the right child subtree replaces t_i in the node, and the leaf of $t_{\min}$ is removed. Each of these removal actions are depicted in fig. 3.9. All operations performed to construct, search, add, and remove nodes in a binary tree are therefore $\mathcal{O}(\log_2(N))$.

Bounded increasing priority queues

To reduce the computational cost to $\mathcal{O}(1)$, we may implement a *bounded increasing priority queue* (BIPQ). The BIPQ was introduced in [37], and relies on two assumptions:

1. A new event t to be added to the queue is always larger than the minimum event removed due to a search t_{last}.

2. There exists a constant $t_{\max}$ which bounds the difference between a new predicted event and the global simulation time.

Therefore

$$t_{\max} > t - t_{\text{last}} \geq 0.$$

The first assumption is fair as all predicted events occur after the global simulation time, or they are invalid. The second assumption is more of a computational construct; for example if particle i has very low velocity it could take a long time for an event to occur. Instead we may choose a value $t_{\max}$ which bounds all *reasonable* event times (for example, $t_{\max} = t_{\text{end}}$, the required simulation time). Any events with $t > t_{\max}$ can be put into an *overflow list*.
The idea behind a BIPQ is to introduce some order to L, without fully ordering all events. We initialize the BIPQ with creation of an array of n empty lists l_i, $1 \leq i \leq n$, for a suitable choice of n (to be determined), and an empty binary search tree structure. We also set $i^* = 0$ and $i_0 = 0$, the *current index* and *base set* of the BIPQ. The operations of a priority queue are applied to a

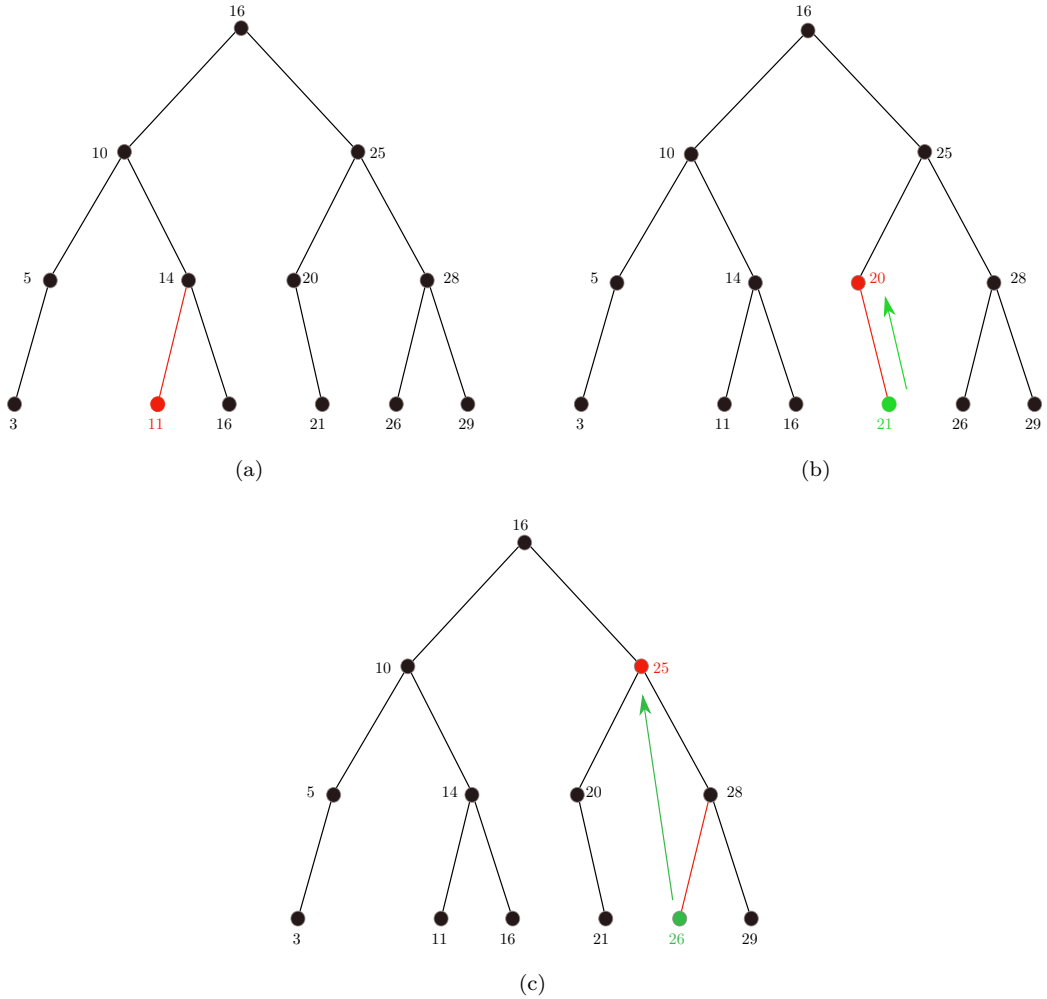


Figure 3.9: Removing nodes from a binary tree. Red branches and nodes are deleted, while green nodes are moved. Top left: Removal of a node with no children. Top right: Removal of a node with one child. Bottom: Removal of a node with two children.

BIPQ as follows:

- Addition of an event: given an event with time t, we determine the *index* of the event with the calculation

$$i = \lfloor st - i_0 \rfloor,$$

  for some scalar factor s (to be determined). If $i \neq i^*$, we add the event to list i. If $i = i^*$, the event is added to the binary tree.

- Identification of the shortest event. This is given by the root of the binary tree. If the binary tree is empty, then the next list l_{i^*+1} is used to fill the binary tree, and we increment i^* by 1. l_{i^*+1} is also empty, we increment i^* until we find a non-empty list. If all l_i are empty for $i^* < i < n$, we reset $i^* = 0$ and increment i_0 by n. If we are using an overflow list l_{over}, when i^* reaches n, we recalculate i for events in l_{over} and redistribute the events into the lists.

- Deletion of an event: given an event with time t we use its index to locate the list or binary tree it is in, and delete it from the list.

To ensure the BIPQ is efficient, the correct choice of n and s needs to be made. In [37], the choice of s is made to ensure that the number of events in the binary tree is consistently small, around 10-20. The number of lists n is then given by

$$n = st_{\max}.$$

However, as mentioned the value for $t_{\max}$ can be very large, as can s. This leads to a significant number of lists in the BIPQ. In [114], a smaller value for n is used in a trade-off with computation time, leading to larger (but still small) number of events in the binary tree.

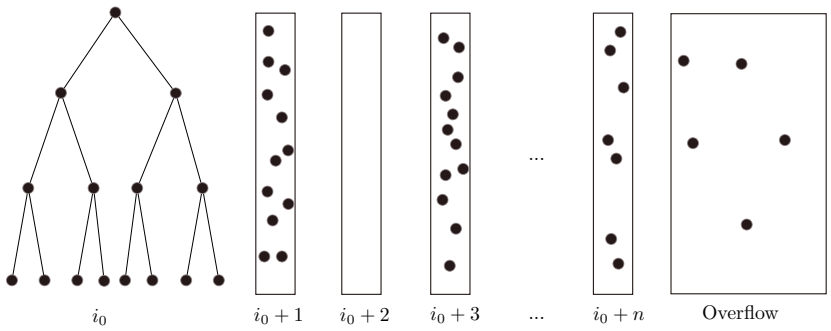


Figure 3.10: Diagram for the BIPQ. at position i_0 there is a binary tree ordering events. Lists $i_0+1,\ldots,i_0+n$ contain unordered events, but introduce a *semi-ordering*. We use an overflow list for events which occur much later on. by introducing a semi ordering to the event times we reduce search time by considering a smaller binary tree.

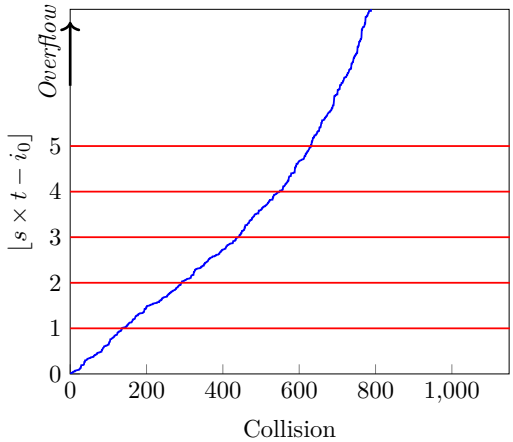


Figure 3.11: How events are distributed into lists, for an example set of events that has been ordered (blue). Events (which here are sorted) between $\lfloor st - i_0 \rfloor \in [0, 1)$ are put into the binary tree, while others have been put into corresponding lists. Events with are large event time are put into the overflow list.

As the size of the binary tree is set to be constant, the structure of the BIPQ reduces computational cost of first event prediction to $\mathcal{O}(1)$. For systems with a small number of particles, the additional structure may not be beneficial, but for larger systems the introduction of a BIPQ will

significantly reduce the computation time. By combining the cell method and the BIPQ we reduce the cost of event-driven particle dynamics simulations to $\mathcal{O}(1)$ per event.

3.5 Improving accuracy

3.5.1 Stability of algorithms

In [115] it is noted that EDPD simulations can become numerically jammed due to machine-precision differences in distance calculations. When two particles are advanced up to a point of contact, it may result in a small, unphysical overlap. After modifying the velocity of the particles to account for the collision, the next collision scheduled between the particles will be when the overlap between the particles is removed. This numerical error can then lead to the simluation jamming.

To resolve this [115] recommends using an *overlap function* and its derivative to determine whether the collision is stable. An overlap function between two particles is given by

$$f_{ij}^{\mathrm{HS}}(t) = \|\boldsymbol{r}_i(t) - \boldsymbol{r}_j(t)\|^2 - \varepsilon^2. \tag{3.24}$$

To ensure that the collision is valid the collision time τ_{ij} must satisfy

$$f_{ij}^{\mathrm{HS}}(\tau_{ij}) \leq 0, \text{ and } \frac{\mathrm{d}}{\mathrm{d}t} f_{ij}^{\mathrm{HS}}(t)|_{t=\tau_{ij}} < 0. \tag{3.25}$$

By implementing this check we can ensure that machine-precision differences do not restrict simulation time.

3.5.2 Inelastic collapse

The assumption that collisions are binary and instantaneous can lead to *inelastic collapse* when using inelastic collisions; an infinite number of collisions can occur in finite time. A fundamental example of this is a single inelastic particle under gravity bouncing on a solid surface (see fig. 3.12a). In [67] it is noted that for small enough coefficient of restitution ($\alpha < 7 - 4\sqrt{3}$) only three particles in one dimension are needed to achieve inelastic collapse, and also provides investigation into the behaviour of inelastic collapse in two dimensions. Numerical experiments generally show inelastic collapse can occur for much larger values of α than those predicted theoretically. In fig. 3.12b We provide an example of inelastic collapse between particles in one dimension. We note that inelastic collapse is a physical phenomenon, and not an artefact of the numerical scheme.

For event-driven simulation inelastic collapse poses a problem; if an infinite number of collisions can occur in a finite time the simulation will not progress past this time without producing non-physical effects. There are several methods to rectify this, as listed in [116]. We employ the *TC method*, which is preferred in [116], where it is argued that when two events for a particular particle are predicted to occur after a short time interval $t_c > 0$, they should not be treated as separate events, and (non-instantaneous) collisions which overlap in time lead to a smaller amount of dissipation. Therefore if a collision happens between particles i and j within t_c of their previous collision, we set coefficient of restitution to $\alpha = 1$, rendering the collision elastic. The resulting dynamics are also displayed in fig. 3.12. We can see that the implementation of

the TC method allows the EDPD simulation to continue past inelastic collapse without producing unrealistic results. This relies on an appropriate choice for t_c for the system considered.

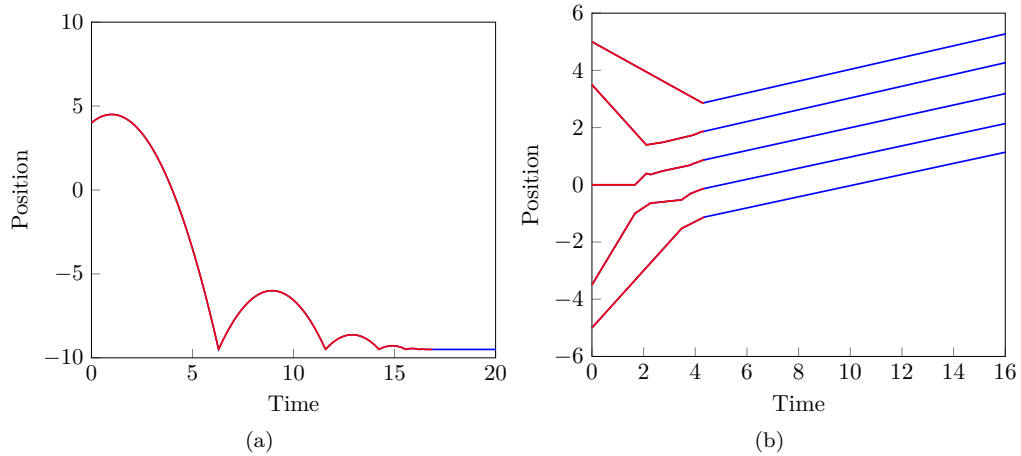


Figure 3.12: Left: inelastic collapse of a single particle with gravity without the TC method (red), the TC method is then implemented using $t_c = 10^{-5}$ to continue the dynamics (blue), right: inelastic collapse between particles in one dimension without (red) the TC method, at which point the TC method is implemented using $t_c = 10^{-5}$ (blue). We note that the dynamics are identical up to the point of inelastic collapse, at which point the TC method allows the simulation to advance in such a way that the trajectories are still physically appropriate.

3.6 `MATLAB` implementation

We have implemented a state-of-the-art (stable) EDPD algorithm in `MATLAB` which utilises asynchronous dynamics and the cell method. The code can support linear dynamics, dynamics with friction, and dynamics with constant external potential in one, two, or three dimensions, and supports elastic and inelastic collisions by utilising the TC method described above. The domain can be periodic or bounded in each dimension. We also include the capability to run multiple samples and record statistical quantities from them, such as the radial correlation function discussed in future chapters.

3.6.1 Chaotic dynamics

The implementation of the cell method and a bounded increasing priority queue produce an efficient algorithm for event-driven particle dynamics, which is independent of the number of particles in the system. However the system has a chaotic nature, as small changes in the initial conditions of the particles in the system produce completely different results after a small number of collisions, due to the sensitivity of the collision rule. Using the `MATLAB` implementation, we consider a system of 25 particles with $\varepsilon = 1$ in two dimensions with no external potentials or friction. Figure 3.13 shows the state of the system after 1000 collisions, given initial conditions $\boldsymbol{x}_0^N, \boldsymbol{v}_0^N$, and initial conditions $\boldsymbol{x}_0^N, \tilde{\boldsymbol{v}}_0^N$, where $\tilde{\boldsymbol{v}}_0^N = \boldsymbol{v}_0^N + (\epsilon, 0, 0, \ldots, 0)$ where ϵ is machine-precision small. The result is an example of the chaotic nature of the particle system. Our uses of EDPD are not dependent on a particular realisation of the system, but statistics on the system over many samples, where

chaotic properties of the dynamics will average out, so we do not need to ensure that chaotic properties are mitigated in our simulations.

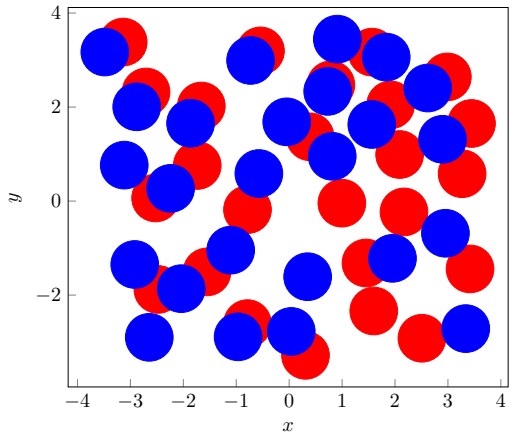


Figure 3.13: Chaotic nature of EDPD. A collection of $N = 25$ particles were simulated for 1000 collisions with random initial positions and velocities, resulting in the red particle positions. One of the initial velocities of the particles was then changed by a machine-precision amount to produce the blue particles after the same number of collisions: small changes in collision outcomes propagate to change the long term dynamics of the system.

3.7 DynamO

DynamO [117] is an open-source event-driven particle simulator which utilises state-of-the-art algorithms and code to run EDPD simulations. It can simulate constant external and discrete interaction potentials in two and three dimensions. It also supports more complex domains, non-spherical particle shapes, and polydispersity, and can handle millions of particles and processes tens of thousands of events per second.

We have implemented viscous drag dynamics given by eq. (3.13) in DynamO. We then use DynamO for larger systems of particles in two and three dimensions and modify the simulations to include viscous drag with the dynamics given by eq. (3.13).

3.8 Discussion

In this section we have discussed how EDPD algorithms can efficiently simulate the dynamics of hard spheres, with machine-precision accuracy. The algorithms presented show that EDPD can be $\mathcal{O}(1)$ per collision and the code libraries discussed implement these methods. We have also presented novel implementations of EDPD in MATLAB and using open-source software DynamO, which can simulate dynamics that include viscous drag terms. The resulting schemes are state-of-the-art, and can compete against other microscopic numerical methods, with the potential to outperform competitors when the system is suited for EDPD (*i.e.* low-density systems where collisions and instantaneous interactions play an important role).

Without using multiple cores, event-driven code libraries such as DynamO can simulate systems of tens of millions of particles, processing nearly 100,000 events per second, on a standard home computer. Furthermore, by determining the step size of the simulation via collisions, we can

reach long simulation times that are required for construction of continuum model parameters. These methods also exactly solve (up to machine precision) the particle dynamics that we used to derive continuum models in this work. Thus for systems of interest to us, these results give EDPD a competitive advantage when compared to collision Verlet schemes or DEM. The main restriction of many implementations is that they can only model systems where particles obey simple differential equations such as the examples presented above. Event-driven methods can also be used to simulate particles that interact via step potentials such as eq. (2.33), however as the number of potential events increases, these methods will lose their competitive edge.

If the free dynamics of interest can be approximated to high precision by an analytic solution, in the timescale of the mean-free path, the same method could also be applied to more complicated systems of dynamics without drastically reducing efficiency of the algorithm, as in the collision Verlet scheme [103]. This may then allow for systems where there is an external flow, suggesting that EDPD could also be used to model fluidisation processes.

We could increase efficiency further by parallelising the dynamics over several cores. Due to causality issues and the chaotic natural or hard-particle interactions EDPD algorithms are naturally very hard to parallelize. In [114] the conditions which restrict parallelization are examined, and an *event-based decomposition* is constructed which produces significant speed-up of EDPD algorithms through parallelization, compared to previous methods which tend to apply a spatial decomposition. We have not implemented this parallelisation scheme for our simulations, as we consider many samples of systems with a small number of particles, so we can simply parallelise over the number of samples instead.

When considering a high density system, where the number of collisions in a short amount of time can be very large, any method to model the system as a collection of particles can still be too computationally expensive for practical uses, especially if we want to model the dynamics over macroscopic timescales. Even with state-of-the-art algorithmic approaches applied, the number of particles in the system will always eventually be a computational bottleneck. For this reason it is inevitably necessary to construct a continuum model to approximate the dynamics in large systems of granular media. In the following chapter we will progress from a microscopic regime to the continuum, by constructing the Liouville equation for hard particles.

Chapter 4

The Liouville equation for hard particles

We wish to construct a continuum model directly from the microscopic dynamics. We therefore require a way to get from the set of differential equations describing the microscopic particle dynamics to a set of equations describing the macroscopic continuum dynamics. Under the assumption that the particle dynamics are smooth, associated with the microscopic dynamics is the *Liouville equation* [118], a partial differential equation which determines the dynamics of the N-body distribution function $f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t)$. For example, when $m = 1$ we have that:

$$\mathcal{L}[f^{(N)}] := \left[\frac{\partial}{\partial t} + \boldsymbol{v}^N \cdot \nabla_{\boldsymbol{r}^N} + \nabla_{\boldsymbol{v}^N} \cdot (\nabla_{\boldsymbol{r}^N} V(\boldsymbol{r}^N, t) + \boldsymbol{\Gamma}(\boldsymbol{r}^N, \boldsymbol{v}^N, t)\boldsymbol{v}^N)\right] f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) = 0. \tag{4.1}$$

For point-like particles and sufficiently smooth $V(\boldsymbol{r}^N, t)$, eq. (4.1) fully describes the evolution of an initial configuration of particles. The result is derived by assuming that particle positions and velocities can be determined to arbitrary precision, and that the microscopic equations of motion fully describe the motion of particles, by using statistical-mechanical methods that determine the Liouville equation by exploring volume elements of the phase space.

However, this derivation does not account for collisional effects; when particles are of finite volume adjustments have to be made to the microscopic dynamics to avoid non-physical particle overlap. At an informal level the collisional effect is not recognised at the level of the Liouville equation, but is derived in the Bogoliubov-Born-Green-Kirkwood-Yvon (BBGKY) hierarchy [46], [119], for example as a consequence of additional assumptions on an interaction force [120]. However, by instantaneously changing the velocities of two particles that undergo a collision a fundamental assumption in the Liouville derivation is no longer valid; the dynamics of an individual particle are no longer smooth. We should therefore model collisional effects at the level of the Liouville equation, rather than when considering the derivation of the BBGKY hierarchy from the Liouville equation. We cannot rely on the classical form of the Liouville equation, and must resort to an alternative formulation to derive an equation for $f^{(N)}(\boldsymbol{R}, \boldsymbol{V}, t)$.

One method that is used to include hard-sphere effects within the Liouville equation is to define

so-called *mild* solutions to eq. (4.88) [121]. Once such solutions have been defined the associated BBGKY hierarchy can be constructed, at which point the Boltzmann collision operator is included when mild solutions are considered. As an alternative, we will consider the *weak formulation* of the Liouville equation [91]. The resulting derivation of the BBGKY hierarchy from the weak form is much more straightforward, when compared to the mild-solution counterpart.

In [91] a system of $N = 2$ spherical particles with diameter $\varepsilon > 0$, $V(\boldsymbol{r}^N, t) = 0$, and $\gamma = 0$ is fully characterised, and under no additional assumptions it is shown that for all smooth, compactly-supported test functions Φ, given initial data $f_0^{(2)} \in C^0(\mathcal{D}) \cap L^1(\mathcal{D})$ (*i.e.* the space of compactly-supported continuous functions), such that $f_0^{(2)}$ integrates to 1 on the phase space, and is always positive, there exists a unique $f^{(2)} \in C^0((-\infty, \infty), L^1(\mathcal{D}))$ which satisfies

$$\begin{aligned}\int_{\mathcal{P}_\varepsilon}\int_{\mathbb{R}^6}\int_{-\infty}^{\infty}\left[\frac{\partial\Phi(\boldsymbol{R},\boldsymbol{V},t)}{\partial t}+(\boldsymbol{V}\cdot\nabla_{\boldsymbol{R}})\Phi(\boldsymbol{R},\boldsymbol{V},t)\right]f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R}\\ =\\ -\int_{\partial\mathcal{P}_\varepsilon}\int_{\mathbb{R}^6}\int_{-\infty}^{\infty}\Phi(\boldsymbol{R},\boldsymbol{V},t)f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{V}\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{R}}(\boldsymbol{R})\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R}),\end{aligned} \tag{4.2}$$

and conserves linear and angular momentum, and kinetic energy. Here

- $\mathcal{P}_\varepsilon = \{\boldsymbol{R} \in \mathbb{R}^6 : \|\boldsymbol{r}_1 - \boldsymbol{r}_2\| \geq \varepsilon\}$ is the hard-sphere domain which restricts particles from overlapping,
- $\mathcal{D} \subseteq \mathbb{R}^{12}$ represents all possible configurations of positions and velocities, given that two particles cannot overlap ($\boldsymbol{R} \in \mathcal{P}_\varepsilon$).
- The vector $\hat{\boldsymbol{\nu}}_{\boldsymbol{R}} \in \mathbb{R}^6$ is the outward unit normal to the surface $\partial\mathcal{P}_\varepsilon$,
- $\mathcal{H}$ is the Hausdorff measure [86] on $\partial\mathcal{P}_\varepsilon$.

In this case, we say that $f^{(2)}$ is a global-in-time weak solution of the Liouville equation:

$$\frac{\partial f^{(2)}}{\partial t} + (\boldsymbol{V}\cdot\nabla_{\boldsymbol{R}})f^{(2)} = C[f^{(2)}], \tag{4.3}$$

where $C[f^{(2)}]$ is determined in the weak sense against test functions $\Phi \in C_1^1((-\infty, \infty), \mathcal{D})$:

$$\langle C[f^{(2)}], \Phi\rangle = \int_{\partial\mathcal{P}_\varepsilon}\int_{\mathbb{R}^6} f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{V}\cdot\hat{\nu}_{\boldsymbol{R}}\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R}). \tag{4.4}$$

In contrast to the classical formulation, a collisional term $C[f^{(2)}]$ is derived at the level of the Liouville equation. Furthermore, under the assumptions of molecular chaos, $C[f^{(2)}]$ admits the elastic Boltzmann collision operator [47], [122] in the first equation of the weak formulation of the BBGKY hierarchy, agreeing with classical results in the strong form [123], [124], where the collision operator is derived by considering macroscopic quantities using statistical mechanics, rather than constructing the macroscopic result as a consequence of microscopic quantities. Under restrictions on initial data, for example on the particle density of the system, an analogue to eq. (4.3) should also hold in systems with more than 2 particles.

From this point a clear question to consider is how more complicated dynamics, or other instantaneous interactions between particles, affect the derivation. In this section we will answer

this question and create a weak form of the Liouville equation that accurately models systems of particles with general free dynamics and any physically appropriate instantaneous interactions, and prove existence of solutions. This novel derivation will differ from the classical result of the strong form eq. (4.88) by accounting for collisions, and fully generalises the weak form eq. (4.3), which is restricted to linear, elastic dynamics. By the end of this chapter we will be in a strong theoretical position to derive a closed set of equations to model a granular medium as a continuum, which has been derived directly from the microscopic dynamics.

4.1 Set-up

We will consider a system with $N = 2$ particles with initial positions $\boldsymbol{r}_1, \boldsymbol{r}_2 \in \mathcal{P}_\varepsilon$ and initial velocities $\boldsymbol{v}_1^0, \boldsymbol{v}_2^0 \in \mathcal{V}^f(\boldsymbol{r}_1, \boldsymbol{r}_2) \subset \mathbb{R}^6$, where $\mathcal{V}^f(\boldsymbol{r}_1, \boldsymbol{r}_2)$ is the set of all admissible initial velocities, given valid initial position data. We define $\boldsymbol{R} = [\boldsymbol{r}_1^0, \boldsymbol{r}_2^0]$ and $\boldsymbol{V} = [\boldsymbol{v}_i^0, \boldsymbol{v}_j^0]$ (in an analogous way to the approach of theorem 2), we define the *free-particle flow maps* for position and velocity by $\boldsymbol{\Phi}_t^x(\boldsymbol{R}, \boldsymbol{V})$ and $\boldsymbol{\Phi}_t^v(\boldsymbol{R}, \boldsymbol{V})$ respectively, and assume that they satisfy the Newton equations

$$\partial_t \boldsymbol{\Phi}_t^x(\boldsymbol{R}, \boldsymbol{V}, t) = \boldsymbol{\Phi}_t^v(\boldsymbol{R}, \boldsymbol{V}, t), \quad \partial_t \boldsymbol{\Phi}_t^v(\boldsymbol{R}, \boldsymbol{V}, t) = \boldsymbol{H}(\boldsymbol{\Phi}_t^x(\boldsymbol{R}, \boldsymbol{V}, t), \boldsymbol{\Phi}_t^v(\boldsymbol{R}, \boldsymbol{V}, t), t). \tag{4.5}$$

Here $\boldsymbol{H} : \mathbb{R}^6 \times \mathbb{R}^6 \times \mathbb{R} \to \mathbb{R}^6$ is differentiable in all its arguments.

We stress that this is a very general set of equations to consider; it also reduces to eqs. (2.1) and (2.2) in chapter 2 by setting:

$$\boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t) = -\nabla_{\boldsymbol{R}} V(\boldsymbol{R}, t) + \boldsymbol{\Gamma}(\boldsymbol{R}, \boldsymbol{V}, t)\boldsymbol{V}. \tag{4.6}$$

Thus the result also accounts for any interaction potentials (*e.g.* cohesive forces) and viscous properties.

We write $\boldsymbol{\Phi}_t = [\boldsymbol{\Phi}_t^r, \boldsymbol{\Phi}_t^v]$ as the flow map for initial data $\boldsymbol{Z} = [\boldsymbol{R}, \boldsymbol{V}] \in \mathcal{D}$. We also assume that the flow maps produce unique trajectories for any given initial data. We also introduce the reduced-difference notation $\tilde{\boldsymbol{x}} = \boldsymbol{x}_1 - \boldsymbol{x}_2$ for $\boldsymbol{X} \in \mathbb{R}^6$, which is useful when discussing relative distances and collision times of particles.

4.1.1 Event times

We will characterise each instantaneous event by an event time, an interaction diameter and an event map that changes the particle velocities. We define the set of interaction diameters by $\Theta = \{\varepsilon_i \in \mathbb{R}^+, i = 1, \dots, m : \varepsilon_m > \varepsilon_{m-1} > \cdots > \varepsilon_0 > 0\}$ where m is the number of interaction diameters of an individual particle. As interactions are instantaneous, the event times can be enumerated as a discrete set, which can be finite or infinite. We write the event times as $\tau_i \in \mathbb{R}$ for $i = -M, -M+1, \dots, -1, 0, 1, \dots, N-1, N$, where $M = M(\boldsymbol{R}, \boldsymbol{V}), N = N(\boldsymbol{R}, \boldsymbol{V}) \in \mathbb{N} \cup \{\infty\}$, and

$$-\infty = \tau_{-M}(\boldsymbol{R}, \boldsymbol{V}) < \tau_{-M+1}(\boldsymbol{R}, \boldsymbol{V}) < \cdots < \tau_{N-1}(\boldsymbol{R}, \boldsymbol{V}) < \tau_N(\boldsymbol{R}, \boldsymbol{V}) = \infty, \tag{4.7}$$

where we choose τ_0 to be the closest event time to time $t = 0$:

$$\tau_0 = \arg\min_{|s|}\{s \in \mathbb{R} : \tilde{\boldsymbol{\Phi}}_s(\boldsymbol{R}, \boldsymbol{V}) \in \partial\mathcal{P}_{\bar{\varepsilon}}, \bar{\varepsilon} \in \Theta\}. \tag{4.8}$$

In turn each τ_i can be defined in terms of the previous or next event time. When the previous event time has been defined:

$$\tau_{i+1}(\boldsymbol{R}, \boldsymbol{V}) := \arg\min_{|s|}\{s \in (\tau_i, \infty) : \boldsymbol{\Phi}_{s-\tau_i}(\boldsymbol{\Psi}^x_{\tau_i}, \boldsymbol{\Psi}^v_{\tau_i}) \in \partial\mathcal{P}_{\bar{\varepsilon}}, \bar{\varepsilon} \in \Theta\}, \tag{4.9}$$

and when the next event has been defined:

$$\tau_{i-1}(\boldsymbol{R}, \boldsymbol{V}) := \arg\min_{|s|}\{s \in (-\infty, \tau_i) : \boldsymbol{\Phi}_{s-\tau_i}(\boldsymbol{\Psi}^x_{\tau_i}, \boldsymbol{\Psi}^v_{\tau_i}) \in \partial\mathcal{P}_{\bar{\varepsilon}}, \bar{\varepsilon} \in \Theta\}. \tag{4.10}$$

These events occur when particles reach an interaction diameter. We write $\boldsymbol{\Psi} = [\boldsymbol{\Psi}^r, \boldsymbol{\Psi}^v]$ as the flow map for the dynamics with instantaneous interactions. We will assume that there exists $\delta > 0$ such that for all $i = -M, -M+1, \ldots, N-1, N$, $\tau_i - \tau_{i-1} > \delta$, and we write $\tau_{-M}(\boldsymbol{R}, \boldsymbol{V}) = -\infty$, $\tau_N(\boldsymbol{R}, \boldsymbol{V}) = \infty$ for the behaviour of the dynamics in the limit as $t \to \pm\infty$. Between each pair of event times the particle dynamics are determined by the free dynamics in eqs. (2.1) and (2.2), where initial conditions can be given by the positions and velocities after the previous event time.

4.1.2 Flow maps

We can fully define the flow maps for dynamics with instantaneous interactions using $\boldsymbol{\Phi}^x_t, \boldsymbol{\Phi}^v_t, \tau_i$ and $\boldsymbol{\sigma}^{\pm}_i$. There are two possible cases for the dynamics when including instantaneous interactions.

No instantaneous events. If $\boldsymbol{R}, \boldsymbol{V}$ are such that no instantaneous events happen, then $M+N = 1$ and $\boldsymbol{\Psi}^r_t, \boldsymbol{\Psi}^v_t$ obey

$$\partial_t \boldsymbol{\Psi}^x_t(\boldsymbol{R}, \boldsymbol{V}) = \boldsymbol{\Phi}^x_t(\boldsymbol{R}, \boldsymbol{V}), \tag{4.11}$$

$$\partial_t \boldsymbol{\Psi}^v_t(\boldsymbol{R}, \boldsymbol{V}) = \boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t). \tag{4.12}$$

Instantaneous events. When $\boldsymbol{R}, \boldsymbol{V}$ are such that instantaneous events occur in finite time, then

$$\partial_t \boldsymbol{\Psi}^x_t(\boldsymbol{R}, \boldsymbol{V}) = \begin{cases} \boldsymbol{\Phi}^v_t(\boldsymbol{R}, \boldsymbol{V}), & \tau_{i_0-1} \le t \le \tau_{i_0}, \\ \boldsymbol{\Phi}^v_{t-\tau_i}(\boldsymbol{\Psi}^x_{\tau_i}(\boldsymbol{R}, \boldsymbol{V}), \boldsymbol{\sigma}^+_i(\boldsymbol{R}, \boldsymbol{V})\boldsymbol{\Psi}^v_{\tau_i}(\boldsymbol{R}, \boldsymbol{V})), & \tau_i < t \le \tau_{i+1}, \\ & i = i_0, \ldots, N-1, \\ & \tau_{i-1} \le t < \tau_i, \\ \boldsymbol{\Phi}^v_{t-\tau_i}(\boldsymbol{\Psi}^x_{\tau_i}(\boldsymbol{R}, \boldsymbol{V}), \boldsymbol{\sigma}^-_i(\boldsymbol{R}, \boldsymbol{V})\boldsymbol{\Psi}^v_{\tau_i}(\boldsymbol{R}, \boldsymbol{V})), & i = -(M+1), \ldots, \\ & \qquad (i_0 - 1). \end{cases} \tag{4.13}$$

and

$$\partial_t \Psi_t^v(\boldsymbol{R},\boldsymbol{V}) = \begin{cases} \boldsymbol{H}(\boldsymbol{R},\boldsymbol{V},t), & \tau_{i_0-1} \le t \le \tau_{i_0}, \\ \boldsymbol{H}(\boldsymbol{\Psi}^x_{\tau_i}(\boldsymbol{R},\boldsymbol{V}), \boldsymbol{\sigma}_i^+(\boldsymbol{R},\boldsymbol{V})\boldsymbol{\Psi}^v_{\tau_i}(\boldsymbol{R},\boldsymbol{V}), t-\tau_i), & \tau_i < t \le \tau_{i+1}, \\ & i = i_0,\dots,N-1, \\ & \tau_{i-1} \le t < \tau_i, \\ \boldsymbol{H}(\boldsymbol{\Psi}^x_{\tau_i}(\boldsymbol{R},\boldsymbol{V}), \boldsymbol{\sigma}_i^-(\boldsymbol{R},\boldsymbol{V})\boldsymbol{\Psi}^v_{\tau_i}(\boldsymbol{R},\boldsymbol{V}), t-\tau_i), & i = -(M+1),\dots, \\ & \qquad (i_0-1). \end{cases} \tag{4.14}$$

where $i_0 = 1$ if $\tau_0 < 0$ and $i_0 = 0$ if $\tau_0 > 0$.

Note that free-particle flow maps need not be defined globally in time to be used in these flow maps if the instantaneous interaction renders the trajectory admissible. We then define $\mathcal{V}(\boldsymbol{R})$ to be the set of admissible initial velocity data, given $\boldsymbol{R} \in \mathcal{P}_\varepsilon$, which accounts for instantaneous interactions. We now provide some results on instantaneous event times, which will prove useful in the derivation. The first result is a generalisation of a result in [91]: where it is shown that for linear dynamics there is at most one non-infinite event time τ, and it satisfies

$$\boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \tau(\boldsymbol{R},\boldsymbol{V}) = -1, \tag{4.15}$$

It is difficult to determine a physical meaning from this result. We shall see that the generalisation provides more insight.

Proposition 11 (Transport Identity I)**.** Given $\boldsymbol{R} \in \mathbb{R}^6$ and $\boldsymbol{V} \in \mathcal{V}(\boldsymbol{R})$, let $\tau(\boldsymbol{R},\boldsymbol{V})$ be a particular event time in the dynamics determined by $\boldsymbol{\Psi}^x_t(\boldsymbol{R},\boldsymbol{V})$ and $\boldsymbol{\Psi}^v_t(\boldsymbol{R},\boldsymbol{V})$. Then for all $t \in (\tau_{i-1},\tau_{i+1})$, (omitting flow map arguments for ease of notation)

$$[\boldsymbol{\Phi}^v_t \cdot \nabla_{\boldsymbol{Y}} + \boldsymbol{H}(\boldsymbol{\Phi}^x_t, \boldsymbol{\Phi}^v_t) \cdot \nabla_{\boldsymbol{W}}] \tau(\boldsymbol{Y},\boldsymbol{W})|_{\boldsymbol{Y}=\boldsymbol{\Psi}^v_t, \boldsymbol{W}=\boldsymbol{\Psi}^v_t} = -1. \tag{4.16}$$

Proof. For ease of notation, we omit the arguments of $\boldsymbol{\Phi}^x_t$ and $\boldsymbol{\Phi}^v_t$ where possible. Firstly, we note that the event time $\tau(\boldsymbol{R},\boldsymbol{V})$ can be written as [91]

$$\tau(\boldsymbol{R},\boldsymbol{V}) = \arg\min_{|s|} \{s : \boldsymbol{\Phi}^x_s(\boldsymbol{R},\boldsymbol{V}) \in \partial\mathcal{P}_\varepsilon\}.$$

We first consider τ as a function of the data at time t, *i.e.*

$$\tilde{\tau}(\boldsymbol{\Psi}_t(\boldsymbol{R},\boldsymbol{V}), \boldsymbol{\Psi}_t(\boldsymbol{R},\boldsymbol{V})) = \tau(\boldsymbol{R},\boldsymbol{V}).$$

To construct the time derivative of $\tilde{\tau}$ the classical definition of the derivative. Let $h > 0$ and assume that $h \ll \delta$. Then

$$\begin{aligned} \tau(\boldsymbol{\Psi}^r_{t+h}, \boldsymbol{\Psi}^v_{t+h}) &= \arg\min\{s \in (\tau_{i-1},\tau_{i+1}) : \|\tilde{\boldsymbol{\Phi}}^x_s(\boldsymbol{\Psi}^x_{t+h}, \boldsymbol{\Psi}^v_{t+h})\| = \varepsilon\} \\ &= \arg\min\{s \in (\tau_{i-1},\tau_{i+1}) : \|\tilde{\boldsymbol{\Phi}}^r_{s-h}(\boldsymbol{\Psi}^r_t, \boldsymbol{\Psi}^v_t)\| = \varepsilon\} \\ &= \arg\min\{s+h \in (\tau_{i-1},\tau_{i+1}) : \|\tilde{\boldsymbol{\Phi}}^r_s(\boldsymbol{\Phi}^r_t, \boldsymbol{\Phi}^v_t)\| = \varepsilon\}. \end{aligned}$$

We have assumed that there exists a unique value $\bar{s}$ for which this is true. It follows that

$$\frac{1}{h}(\tau(\boldsymbol{\Psi}_t^r, \boldsymbol{\Psi}_t^v) - \tau(\boldsymbol{\Psi}_{t+h}^r, \boldsymbol{\Psi}_{t+h}^v)) = \frac{1}{h}(\bar{s} - (\bar{s} + h)) = -1.$$

Thus in the limit as $h \to 0$ we find that the time derivative of τ is -1. To relate the time derivative of τ to the result on the left-hand side of eq. (4.16), we use results from generator theory [125]. Defining the generator

$$\mathcal{L}\boldsymbol{v} = \begin{pmatrix} \boldsymbol{V} \\ \boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}) \end{pmatrix} \cdot \nabla_{\boldsymbol{Z}} \boldsymbol{v}, \tag{4.17}$$

we have that

$$\begin{aligned} \partial_t \tau(\boldsymbol{Z}(t)) =& \nabla_{\boldsymbol{Z}} \tau(\boldsymbol{Z}(t)) \cdot \frac{\mathrm{d}\boldsymbol{Z}(t)}{\mathrm{d}t} \\ =& \nabla_{\boldsymbol{Z}} \tau(\boldsymbol{Z}(t)) \cdot \begin{pmatrix} \Phi_t^v \\ \boldsymbol{H}(\boldsymbol{\Phi}_t^r, \boldsymbol{\Phi}_t^v) \end{pmatrix} = \mathcal{L}\tau(\boldsymbol{Z}(t)). \end{aligned}$$

Thus we find that, for $t \in (\tau_{i-1}, \tau_{i+1})$:

$$\partial_t \tau(\boldsymbol{\Psi}_t^r, \boldsymbol{\Psi}_t^v) = [\partial_t \boldsymbol{\Phi}_t^r \cdot \nabla_{\boldsymbol{Y}} + \partial_t \boldsymbol{\Phi}_t^v \cdot \nabla_{\boldsymbol{W}}] \, \tau(\boldsymbol{Y}, \boldsymbol{W})|_{\boldsymbol{Y}=\boldsymbol{\Phi}_t^r, \boldsymbol{W}=\boldsymbol{\Phi}_t^v}.$$

By using the Newton equations for the free-particle dynamics we arrive at the required result. □

The physical meaning of this result is straightforward; advancing forward in time towards an event in the future decreases the time until the event proportionally. The result also simplifies to eq. (4.15) when considering linear dynamics. The second transport identity of interest involves the event maps $\boldsymbol{\sigma}^{\pm}$.

We can also provide an analogous result to the second transport identity in [91].

Proposition 12 (Transport Identity II)**.** Given $\boldsymbol{R} \in \mathbb{R}^6$ and $\boldsymbol{V} \in \mathcal{V}(\boldsymbol{R})$, let $\boldsymbol{\sigma}_i(\boldsymbol{R}, \boldsymbol{V})$ be a particular event map in the dynamics determined by $\boldsymbol{\Psi}_t^r(\boldsymbol{R}, \boldsymbol{V})$ and $\boldsymbol{\Psi}_t^v(\boldsymbol{R}, \boldsymbol{V})$. Then for all $t \in (\tau_{i-1}, \tau_{i+1})$, (omitting flow map arguments for ease of notation)

$$[\boldsymbol{\Phi}_t^v \cdot \nabla_{\boldsymbol{Y}} + \boldsymbol{H}(\boldsymbol{\Phi}_t^r, \boldsymbol{\Phi}_t^v) \cdot \nabla_{\boldsymbol{W}}] \boldsymbol{\sigma}(\boldsymbol{Y}, \boldsymbol{W})|_{\boldsymbol{Y}=\boldsymbol{\Psi}_t^r, \boldsymbol{W}=\boldsymbol{\Psi}_t^v} = 0. \tag{4.18}$$

Proof. By using eq. (4.17), we see that the time derivative of $\boldsymbol{\sigma}_i(\boldsymbol{R}, \boldsymbol{V})$ and the left-hand side of eq. (4.18) are equal. We note that, by shifting in time,

$$\begin{aligned} \partial_t \boldsymbol{\sigma}(\boldsymbol{\Phi}_t^x, \boldsymbol{\Phi}_t^v) =& \partial_t \big(\eta[\boldsymbol{\Phi}_{\tau_i - t}^r(\boldsymbol{\Psi}_t^x, \boldsymbol{\Psi}_t^v), \boldsymbol{\Phi}_{\tau_i - t}^v(\boldsymbol{\Psi}_t^r, \boldsymbol{\Psi}_t^v)] \\ & \times \big(\boldsymbol{N}[\boldsymbol{\Phi}_{\tau_i - t}^r(\boldsymbol{\Psi}_t^r, \boldsymbol{\Psi}_t^v)] \otimes \boldsymbol{N}[\boldsymbol{\Phi}_{\tau_i - t}^r(\boldsymbol{\Psi}_t^r, \boldsymbol{\Psi}_t^v)] \big) \big), \\ =& \partial_t \left(\eta \left[\boldsymbol{\Phi}_{\tau_i}^r, \boldsymbol{\Phi}_{\tau_i}^v \right] \boldsymbol{N}[\boldsymbol{\Phi}_{\tau_i}^x] \otimes \boldsymbol{N}[\boldsymbol{\Phi}_{\tau_i}^r] \right) = 0, \end{aligned}$$

which completes the proof. □

This result again has a straightforward physical meaning; event maps do not depend on time; only on the instantaneous positions and velocities at the time of the interaction.

4.2 Formulation

We state the following definition for solutions of the Liouville equation for dynamics with instantaneous interactions.

Definition 13. (Global in time weak solutions of the Liouville equation) Suppose we are given an initial condition $f_0 \in C^0(\mathcal{D}) \cap L^1(\mathcal{D})$ (*i.e.* compactly supported and continuous), such that

$$\int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} f_0(\boldsymbol{R}, \boldsymbol{V}) \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} = 1, \quad f_0(\boldsymbol{R}, \boldsymbol{V}) \geq 0. \tag{4.19}$$

Then $f \in C^0((-\infty, \infty), L^1(\mathcal{D}))$ is a physical global-in-time solution of the Liouville equation

$$\partial_t f + \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} f + \nabla_{\boldsymbol{V}} \cdot (\boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t) f) = C_{\boldsymbol{R}}[f^{(2)}] + C_{\boldsymbol{V}}[f^{(2)}] \tag{4.20}$$

if and only if for all test functions $\Phi \in C_c^\infty(\mathbb{R}^6 \times (-\infty, \infty))$ (*i.e.* functions that are compactly supported and smooth on the space $\mathbb{R}^6 \times (-\infty, \infty)$):

$$\begin{aligned}
&\int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}, t) [\partial_t \Phi(\boldsymbol{R}, \boldsymbol{V}, t) + \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \\
&\qquad\qquad + \nabla_{\boldsymbol{V}} \cdot (\boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t) \Phi(\boldsymbol{R}, \boldsymbol{V}), t))] \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} \\
&= - \int_{\partial \mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}, t) \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{V} \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{R}} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\mathcal{H}(\boldsymbol{R}) \\
&- \int_{\mathcal{P}_\varepsilon} \int_{\partial \mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}), t) \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t) \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{V}} \, \mathrm{d}t \, \mathrm{d}\mathcal{H}(\boldsymbol{R}, \boldsymbol{V}) \, \mathrm{d}\boldsymbol{R},
\end{aligned} \tag{4.21}$$

where $\hat{\boldsymbol{\nu}}_{\boldsymbol{R}}$ and $\hat{\boldsymbol{\nu}}_{\boldsymbol{V}}$ are the outward unit normals of the surfaces $\partial \mathcal{P}_\varepsilon$ and $\partial \mathcal{V}(\boldsymbol{R})$ respectively, and f obeys the conservation of linear and angular momentum for all $t \in (-\infty, \infty)$:

$$\int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} (\boldsymbol{v}_1 + \boldsymbol{v}_2) f(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} (\boldsymbol{v}_1 + \boldsymbol{v}_2) f_0(\boldsymbol{R}, \boldsymbol{V}) \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}, \tag{4.22}$$

$$\begin{aligned}
&\int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} (\boldsymbol{r}_1 \times \boldsymbol{v}_1 + \boldsymbol{r}_2 \times \boldsymbol{v}_2) f(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} = \\
&\qquad \int_{\mathcal{P}} \int_{\mathcal{V}(\boldsymbol{R})} (\boldsymbol{r}_1 \times \boldsymbol{v}_1 + \boldsymbol{r}_2 \times \boldsymbol{v}_2) f_0(\boldsymbol{R}, \boldsymbol{V}) \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R},
\end{aligned} \tag{4.23}$$

and the microscopic dynamics satisfy the associated event map constraints.

Now that we have given a definition we can state an existence theorem.

Theorem 14. For any $f_0 \in C^0(\mathcal{D}) \cap L^1(\mathcal{D})$, there exists a physical global-in-time weak solution to the Liouville equation eq. (4.21).

The Liouville equation here is derived for a system of two particles. In principle, systems of many particles may involve many-body interactions. However, given the correct subset of initial data, which ensures that all instantaneous interactions are pairwise, the equation eq. (4.21) should also be accurate for systems of many particles, and may be used to approximate systems where the initial data is not so carefully constructed.

We can also state a general form of the BBGKY hierarchy. We start with a definition of global-in-time weak solutions.

Definition 15. Let $f_0 \in C^1(\mathcal{D}) \cap L^1(\mathcal{D})$ be symmetric in its particle arguments (*i.e.* $[\boldsymbol{r}_1, \boldsymbol{v}_1]$ and $[\boldsymbol{r}_2, \boldsymbol{v}_2]$ can be interchanged). We say that a pair of maps $(f_0^{(1)}, f_0^{(2)})$ with membership

$$f^{(1)} \in C^0((-\infty, \infty), L^1(T\mathbb{R}^3)), \quad f^{(2)} \in C^0((-\infty, \infty), L^2(\mathcal{D})) \tag{4.24}$$

is a global-in-time weak solution of the BBGKY hierarchy associated to the initial data

$$f_0^{(1)} = \int_{\mathbb{R}^3 \setminus \mathcal{B}_\varepsilon(\boldsymbol{r}_1)} \int_{\mathcal{V}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{v}_1)} f_0(\boldsymbol{R}, \boldsymbol{V}) \, \mathrm{d}\boldsymbol{v}_2 \, \mathrm{d}\boldsymbol{r}_2, \quad \text{for all } [\boldsymbol{r}_1, \boldsymbol{v}_1] \in T\mathbb{R}^3, \tag{4.25}$$

and

$$f_0^{(2)}(\boldsymbol{R}, \boldsymbol{V}) = f_0(\boldsymbol{R}, \boldsymbol{V}), \quad \text{for all } [\boldsymbol{R}, \boldsymbol{V}] \in \mathcal{D}, \tag{4.26}$$

if and only if, for all test functions $\phi \in C_c^\infty(T\mathbb{R}^3 \times (-\infty, \infty))$,

$$\begin{aligned}
&\int_{\mathbb{R}^3} \int_{\mathbb{R}^3} \int_{-\infty}^{\infty} (\partial_t + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{r}_1}) \phi(\boldsymbol{r}_1, \boldsymbol{v}_1, t) f^{(1)}(\boldsymbol{r}_1, \boldsymbol{v}_1, t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}_1 \, \mathrm{d}\boldsymbol{r}_1 \\
&+ \int_{\mathbb{R}^3} \int_{\mathbb{R}^3 \setminus \mathcal{B}_\varepsilon(\boldsymbol{r}_1)} \int_{\mathbb{R}^3} \int_{\mathcal{V}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{v}_1)} \int_{-\infty}^{\infty} \nabla_V \cdot (\boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t) \phi(\boldsymbol{r}_1, \boldsymbol{v}_1, t)) \\
&\qquad \times f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}_2 \, \mathrm{d}\boldsymbol{v}_1 \, \mathrm{d}\boldsymbol{r}_2 \, \mathrm{d}\boldsymbol{r}_1 \\
&= \\
&-\frac{1}{\sqrt{2}} \int_{\mathbb{R}^3} \int_{\mathbb{S}^2} \int_{\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} \phi(\boldsymbol{r}_1, \boldsymbol{v}_1, t) f^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_2 + \varepsilon \boldsymbol{n}], [\boldsymbol{v}_1, \boldsymbol{v}_2], t) (\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot \boldsymbol{n} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{n} \, \mathrm{d}\boldsymbol{r}_1 \\
&\quad - \int_{\mathcal{P}_\varepsilon} \int_{\partial \mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} \phi(\boldsymbol{r}_1, \boldsymbol{v}_1, t) f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t) \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{V}} \, \mathrm{d}t \, \mathrm{d}\mathcal{H}(\boldsymbol{V}, \boldsymbol{R}) \, \mathrm{d}\boldsymbol{R},
\end{aligned} \tag{4.27}$$

and $f^{(2)}$ satisfies eq. (4.21).

After a partition of the phase space the BBGKY hierarchy will follow as a corollary of theorem 14.

Corollary 16. For any $f_0 \in C^0(\mathcal{D}) \cap L^1(\mathcal{D})$ there exists a global-in-time solution to the BBGKY hierarchy given by definition 15.

We shall see in later examples that for more intricate results one must have a good understanding of the admissible data. For example, in linear elastic hard-sphere dynamics the derivation of the BBGKY hierarchy result in the Boltzmann collision operator on the right-hand side of eq. (4.27) [91].

4.3 Derivation

The method to derive the Liouville equation in both of our examples is similar to the method used in [91]: we wish to find weak solutions $f^{(2)}$ to the equation

$$\mathcal{L}[f^{(2)}] = C[f^{(2)}],$$

where the operator $\mathcal{L}$ is the Liouville equation associated to our choice of dynamics, for point-like particles. To derive the operator on the right-hand side we consider the standard Liouville equations on test functions $\Phi \in C_c^\infty(\mathcal{D} \times (-\infty, \infty))$, multiply by $f^{(2)}$, and integrate to find

$$\int_{\mathcal{P}_\varepsilon} \int_{\mathbb{R}^6} \int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \mathcal{L}[\Phi(\boldsymbol{R}, \boldsymbol{V}, t)] \, \mathrm{d}\boldsymbol{R} \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}t = 0. \tag{4.28}$$

By using standard integration identities we will transfer the operator $\mathcal{L}$ from the function Φ to $f^{(2)}$, then any remainders will contribute to the surface terms on the right-hand side of eq. (4.28).

We partition the derivation into three parts:

$$I(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \partial_t \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}, \tag{4.29}$$

$$J(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}, \tag{4.30}$$

$$K(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \nabla_{\boldsymbol{V}} \cdot [\boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, t) \Phi(\boldsymbol{R}, \boldsymbol{V}, t)] \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}, \tag{4.31}$$

which we call the time, space and velocity derivative terms respectively. For each of these integrals we separate the phase space into parts where $f^{(2)}$ is smooth, use standard integral identities on these parts (taking care of any surface terms that arise), and collect like terms. We will then sum $I(\Phi), J(\Phi)$ and $K(\Phi)$ together to achieve the result.

4.3.1 The time derivative term

Using the flow maps $\boldsymbol{\Psi}_t^x, \boldsymbol{\Psi}_t^v$ we can write the time integral in $I(\Phi)$ as a sum of integrals:

$$I(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \sum_{i=-M(\boldsymbol{R},\boldsymbol{V})+1}^{N(\boldsymbol{R},\boldsymbol{V})} \int_{-\tau_i}^{-\tau_{i-1}} f_0^{(2)}(\boldsymbol{\Psi}_{-t}^x, \boldsymbol{\Psi}_{-t}^v) \partial_t \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}. \tag{4.32}$$

On each interval (τ_{i-1}, τ_i) the flow is described by the free dynamics, so we may apply integration by parts, which we state here for completion.

Theorem 17. (Integration by parts [126]) For a vector $\boldsymbol{v}$, scalar function ϕ over a volume Ω with surface $\mathrm{d}\Omega$,

$$\int_\Omega \phi \nabla \cdot \boldsymbol{v} \, \mathrm{d}\boldsymbol{V} = \int_{\partial\Omega} \phi \boldsymbol{v} \cdot \mathrm{d}\boldsymbol{S} - \int_\Omega \boldsymbol{v} \cdot \nabla \phi \, \mathrm{d}\boldsymbol{V}. \tag{4.33}$$

Keeping in mind that evaluating $\boldsymbol{\Psi}_t(\boldsymbol{R}, \boldsymbol{V})$ at event times from the left or right provides a different result (due the the application of $\boldsymbol{\sigma}_i(\boldsymbol{R}, \boldsymbol{V})$ at each event time τ_i), and that compactly-supported functions yield zero at $\tau_{-M} = -\infty, \tau_N = \infty$, we obtain

$$\begin{aligned} I(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \sum_{i=-M+1}^{N-1} \left\{ \Phi(\boldsymbol{R}, \boldsymbol{V}, \tau_i) (f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^-}) - f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^+})) \right\} \\ - \sum_{i=-M+1}^{N} \left\{ \int_{-\tau_i}^{-\tau_{i-1}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \partial_t f_0^{(2)}(\boldsymbol{\Psi}_{-t}^x, \boldsymbol{\Psi}_{-t}^v) \, \mathrm{d}t \right\} \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}, \end{aligned}$$

where

$$\boldsymbol{\Psi}_{\tau^-} = \lim_{t\to\tau^-} \boldsymbol{\Psi}_t, \quad \boldsymbol{\Psi}_{\tau^+} = \lim_{t\to\tau^+} \boldsymbol{\Psi}_t.$$

For each interval (τ_{i-1}, τ_i), the flow maps $\boldsymbol{\Psi}^x_t, \boldsymbol{\Psi}^v_t$ are determined by the free dynamics $\boldsymbol{\Phi}^x_t, \boldsymbol{\Phi}^v_t$ with particular initial conditions. We apply the chain rule on each of these intervals. For the second term in this result we consider two separate cases.

If there are no particle-particle interactions for initial data $\boldsymbol{R}, \boldsymbol{V}$, then $M + N = 1$, and

$$\begin{aligned}
\partial_t f_0^{(2)}(\boldsymbol{\Psi}^x_{-t}, \boldsymbol{\Psi}^v_{-t}) =& [\partial_t \boldsymbol{\Phi}^x_{-t} \cdot \nabla_{\boldsymbol{Y}} + \partial_t \boldsymbol{\Phi}^v_{-t} \cdot \nabla_{\boldsymbol{W}}] f_0^{(2)}|_{\boldsymbol{Y}=\boldsymbol{\Phi}^x_{-t}, \boldsymbol{W}=\boldsymbol{\Phi}^v_{-t}}, \\
=& [\boldsymbol{\Phi}^v_{-t} \cdot \nabla_{\boldsymbol{Y}} + \boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, -t) \cdot \nabla_{\boldsymbol{W}}] f_0^{(2)}|_{\boldsymbol{Y}=\boldsymbol{\Phi}^x_{-t}, \boldsymbol{W}=\boldsymbol{\Phi}^v_{-t}}, \\
=& -[\boldsymbol{\Psi}^v_{-t} \cdot \nabla_{\boldsymbol{R}} + \boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, -t) \cdot \nabla_{\boldsymbol{V}}] f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t).
\end{aligned}$$

For initial data which causes particle-particle collisions,

$$\partial_t f_0^{(2)} = \begin{cases}
-[\boldsymbol{\Psi}^v_{-t} \cdot \nabla_{\boldsymbol{R}} + \boldsymbol{H}(\boldsymbol{R}, \boldsymbol{V}, -t) \cdot \nabla_{\boldsymbol{V}}] f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t), & \tau_{i_0-1} \le t \le \tau_{i_0}, \\
-[\boldsymbol{\Psi}^v_{-(t-\tau_i)} \cdot \nabla_{\boldsymbol{Y}} + \boldsymbol{H}(\boldsymbol{\Psi}^x_{-\tau_i}, \boldsymbol{\sigma}_i^+ \boldsymbol{\Psi}^v_{-\tau_i}, -(t-\tau_i)) \cdot \nabla_{\boldsymbol{W}}] & \tau_i < t < \tau_{i+1}, \\
\qquad \circ f^{(2)}(\boldsymbol{Y}, \boldsymbol{W}, t)|_{\boldsymbol{Y}=\boldsymbol{\Phi}^x_{-t}, \boldsymbol{W}=\boldsymbol{\Phi}^v_{-t}} & i = i_0, \dots, N-1, \\
-[\boldsymbol{\Psi}^v_{-(t-\tau_i)} \cdot \nabla_{\boldsymbol{Y}} + \boldsymbol{H}(\boldsymbol{\Psi}^x_{-\tau_i}, \boldsymbol{\sigma}_i^- \boldsymbol{\Psi}^v_{-\tau_i}, -(t-\tau_i)) \cdot \nabla_{\boldsymbol{W}}] & \tau_i < t < \tau_{i+1}, \\
\qquad \circ f^{(2)}(\boldsymbol{Y}, \boldsymbol{W}, t)|_{\boldsymbol{Y}=\boldsymbol{\Phi}^x_{-t}, \boldsymbol{W}=\boldsymbol{\Phi}^v_{-t}} & i = -M+1, \dots, (i_0-1),
\end{cases}$$

where $i_0 = 0$ if $\tau_0 < 0$ (*i.e.* the closest event to $t = 0$ is in the past) and $i_0 = 1$ if $\tau_0 > 0$ (the closest event is in the future).

We partition the result into the surface terms and the new integrals:

$$I_1(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \sum_{i=-M+1}^{N-1} \Phi(\boldsymbol{R}, \boldsymbol{V}, \tau_i)(f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^-}) - f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^+}))\, \mathrm{d}\boldsymbol{V}\, \mathrm{d}\boldsymbol{R}, \tag{4.34}$$

$$I_2(\Phi) = -\int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \sum_{i=-M+1}^{N} \int_{-\tau_i}^{-\tau_{i-1}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \partial_t f_0^{(2)}(\boldsymbol{\Psi}^x_{-t}, \boldsymbol{\Psi}^v_{-t})\, \mathrm{d}t\, \mathrm{d}\boldsymbol{V}\, \mathrm{d}\boldsymbol{R}. \tag{4.35}$$

4.3.2 The space derivative term

We partition the time integral in $J(\Phi)$ into terms where the argument is smooth:

$$J(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \sum_{i=-M+1}^{N} \int_{-\tau_i}^{-\tau_{i-1}} f_0^{(2)}(\boldsymbol{\Psi}^x_{-t}, \boldsymbol{\Psi}^v_{-t}) \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \Phi(\boldsymbol{Z}, t)\, \mathrm{d}t\, \mathrm{d}\boldsymbol{V}\, \mathrm{d}\boldsymbol{R}, \tag{4.36}$$

We apply the following rule on each term:

$$f_0^{(2)} \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \Phi = \nabla_{\boldsymbol{R}} \cdot (\boldsymbol{V} f_0^{(2)} \Phi) - \Phi \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} f_0^{(2)}, \tag{4.37}$$

which results in:

$$J(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \sum_{i=-M+1}^{N} \left\{ \int_{-\tau_i}^{-\tau_{i-1}} \nabla_{\boldsymbol{R}} \cdot (\boldsymbol{V} f_0^{(2)} \Phi)\, \mathrm{d}t \right. \tag{4.38}$$

$$- \int_{-\tau_i}^{-\tau_{i-1}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}t \Bigg\} \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}.$$

It remains to consider the first (divergence) term in $J(\Phi)$. We split this integral by the velocities $\boldsymbol{V} \in \mathcal{C}(\boldsymbol{R})$ that cause interactions and $\boldsymbol{V} \in \mathcal{V}(\boldsymbol{R}) \backslash \mathcal{C}(\boldsymbol{R})$ that do not. For the latter case we have

$$M_1(\Phi) := \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R}) \backslash \mathcal{C}(\boldsymbol{R})} \int_{-\infty}^{\infty} \nabla_{\boldsymbol{R}} \cdot (\boldsymbol{V} f_0^{(2)}(\boldsymbol{\Phi}_{-t}^x, \boldsymbol{\Phi}_{-t}^v) \Phi(\boldsymbol{R}, \boldsymbol{V}, t)) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}. \tag{4.39}$$

By the divergence theorem we can write:

$$M_1(\Phi) = \int_{\partial \mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R}) \backslash \mathcal{C}(\boldsymbol{R})} \int_{-\infty}^{\infty} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{V} \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{R}}(\boldsymbol{R}, \boldsymbol{V}) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\mathcal{H}(\boldsymbol{R}), \tag{4.40}$$

where $\hat{\nu}(\boldsymbol{R}, \boldsymbol{V})$ is the outward unit normal of $\partial \mathcal{P}_\varepsilon$.

For the collisional integrals we use Reynold's transport theorem [127] on each term in the sum to find

$$\begin{aligned}
&\int_{\mathcal{P}_\varepsilon} \int_{\mathcal{C}(\boldsymbol{R})} \sum_{i=-M+1}^{N} \int_{-\tau_i}^{-\tau_{i-1}} \nabla_{\boldsymbol{R}} \cdot (\boldsymbol{V} f_0^{(2)} \Phi) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} \\
&\quad = \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{C}(\boldsymbol{R})} \sum_{i=-M+1}^{N} \nabla_{\boldsymbol{R}} \cdot \int_{-\tau_i}^{-\tau_{i-1}} \boldsymbol{V} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) f_0^{(2)}(\boldsymbol{\Psi}_{-t}) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} \\
&\qquad + \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{C}(\boldsymbol{R})} \sum_{i=-M+1}^{N-1} (\boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \tau_i) \Phi(\boldsymbol{R}, \boldsymbol{V}, \tau_i) [f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^-}) - f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^+})] \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}.
\end{aligned} \tag{4.41}$$

Once again we split the result $J(\Phi)$ into several parts:

$$J_1(\Phi) := \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{C}(\boldsymbol{R})} \sum_{i=-M+1}^{N-1} (\boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \tau_i) \Phi(\boldsymbol{R}, \boldsymbol{V}, \tau_i) [f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^-}) - f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^+})] \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} \tag{4.42}$$

$$J_2(\Phi) := - \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}, \tag{4.43}$$

$$\begin{aligned}
J_3(\Phi) := &\int_{\partial \mathcal{P}_\varepsilon} \int_{\mathcal{V}(\boldsymbol{R}) \backslash \mathcal{C}(\boldsymbol{R})} \int_{-\infty}^{\infty} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{V} \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{R}}(\boldsymbol{R}, \boldsymbol{V}) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\mathcal{H}(\boldsymbol{R}) \\
&+ \int_{\mathcal{P}_\varepsilon} \int_{\mathcal{C}(\boldsymbol{R})} \sum_{i=-M+1}^{N} \nabla_{\boldsymbol{R}} \cdot \int_{-\tau_i}^{-\tau_{i-1}} \boldsymbol{V} \Phi(\boldsymbol{R}, \boldsymbol{V}, t) f_0^{(2)}(\boldsymbol{\Psi}_{-t}) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R}.
\end{aligned} \tag{4.44}$$

4.3.3 The velocity derivative term

The velocity derivative follows a similar argument to the one above; we use the following calculus identity between each two event times:

$$f^{(2)} \nabla_{\boldsymbol{V}} \cdot (\boldsymbol{H} \Phi) = \nabla_{\boldsymbol{V}} \cdot (\boldsymbol{H} f^{(2)} \Phi) - \Phi \boldsymbol{H} \cdot \nabla_{\boldsymbol{V}} f^{(2)}. \tag{4.45}$$

It remains to consider the first term in eq. (4.45). First we consider the case of no interactions. For each $X \in \mathcal{P}_\varepsilon$, $\mathcal{V}(\boldsymbol{R})\backslash\mathcal{C}(\boldsymbol{R})$ the divergence theorem provides us with the following result:

$$\begin{aligned} M_2(\Phi) &:= \int_{\mathcal{P}_\varepsilon}\int_{\mathcal{V}(\boldsymbol{R})\backslash\mathcal{C}(\boldsymbol{R})}\int_{-\infty}^{\infty} \nabla_{\boldsymbol{V}}\cdot(G(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)f_0^{(2)}(\boldsymbol{\Phi}_{-t}))\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R}, \\ &= \int_{\mathcal{P}_\varepsilon}\int_{\partial(\mathcal{V}(\boldsymbol{R})\backslash\mathcal{C}(\boldsymbol{R}))}\int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{H}\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{V}}(\boldsymbol{V})\,\mathrm{d}t\,\mathrm{d}\mathcal{H}(\boldsymbol{V})\,\mathrm{d}\boldsymbol{R}. \end{aligned} \tag{4.46}$$

For the collisional part we can use Reynolds' transport theorem in an analogous fashion to before, as we have assumed that $\mathcal{C}(\boldsymbol{R})$ is a subset of $\mathcal{V}(\boldsymbol{R})$ with a piecewise-smooth boundary:

$$\begin{aligned} &\int_{\mathcal{P}_\varepsilon}\int_{\mathcal{C}(\boldsymbol{R})}\sum_{i=-M+1}^{N}\int_{-\tau_i}^{\tau_{i-1}}\nabla_{\boldsymbol{V}}\cdot(\boldsymbol{H}\Phi f_0^{(2)}(\boldsymbol{\Psi}_{-t}))\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R} \\ &\quad= \int_{\mathcal{P}_\varepsilon}\int_{C(\boldsymbol{R})}\sum_{i=-M+1}^{N}\nabla_{\boldsymbol{V}}\cdot\int_{-\tau_i}^{-\tau_{i-1}}\boldsymbol{H}f^{(2)}(\boldsymbol{Z},t)\Phi(\boldsymbol{Z},t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R} \\ &\quad+ \int_{\mathcal{P}_\varepsilon}\int_{C(\boldsymbol{R})}\sum_{i=-M+1}^{N-1}[\boldsymbol{H}\cdot\nabla_{\boldsymbol{V}}\tau_i]\Phi(\boldsymbol{Z},\tau_i)[f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^-})-f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^+})]\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R}. \end{aligned} \tag{4.47}$$

We partition the result as follows:

$$K_1(\Phi) = \int_{\mathcal{P}_\varepsilon}\int_{C(\boldsymbol{R})}\sum_{i=-M+1}^{N-1}[\boldsymbol{H}\cdot\nabla_{\boldsymbol{V}}\tau_i]\Phi(\boldsymbol{Z},\tau_i)[f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^-})-f_0^{(2)}(\boldsymbol{\Psi}_{\tau_i^+})]\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R} \tag{4.48}$$

$$K_2(\Phi) = -\int_{\mathcal{P}_\varepsilon}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty}\Phi(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{H}\cdot\nabla_{\boldsymbol{V}}f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R}, \tag{4.49}$$

$$\begin{aligned} K_3(\Phi) &= \int_{\mathcal{P}_\varepsilon}\int_{\partial(\mathcal{V}(\boldsymbol{R})\backslash\mathcal{C}(\boldsymbol{R}))}\int_{-\infty}^{\infty}f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{H}\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{V}}(\boldsymbol{V})\,\mathrm{d}t\,\mathrm{d}\mathcal{H}(\boldsymbol{V})\,\mathrm{d}\boldsymbol{R} \\ &\quad+\int_{\mathcal{P}_\varepsilon}\int_{C(\boldsymbol{R})}\sum_{i=-M+1}^{N}\nabla_{\boldsymbol{V}}\cdot\int_{-\tau_i}^{-\tau_{i-1}}\boldsymbol{H}f^{(2)}(\boldsymbol{Z},t)\Phi(\boldsymbol{Z},t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R}. \end{aligned} \tag{4.50}$$

4.3.4 Combining all terms

We now combine all contributions into one equation. We note that by proposition 11,

$$I_1(\Phi) = -(J_1(\Phi)+K_1(\Phi)), \tag{4.51}$$

and so when combining all contributions, these terms disappear. By the results of propositions 11 and 12 we see that

$$I_2(\Phi) = -(J_2(\Phi)+K_2(\Phi)). \tag{4.52}$$

The remaining terms $J_3(\Phi)$ and $K_3(\Phi)$ are surface terms in position and velocity phase space respectively. By applying the divergence theorem to the second integral in $J_3(\Phi)$ and $K_3(\Phi)$, we can combine terms to find

$$J_3(\Phi) = \int_{\partial\mathcal{P}_\varepsilon}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty}f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)V\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{R}}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R}), \tag{4.53}$$

$$K_3(\Phi) = \int_{\mathcal{P}_\varepsilon} \int_{\partial\mathcal{V}(\boldsymbol{R})} \int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{H}(\boldsymbol{R},\boldsymbol{V},t)\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{V}}(\boldsymbol{R},\boldsymbol{V})\,\mathrm{d}t\,\mathrm{d}\mathcal{H}(\boldsymbol{R},\boldsymbol{V})\,\mathrm{d}\boldsymbol{R}. \quad (4.54)$$

This concludes the proof of theorem 14.

4.3.5 The BBGKY hierarchy

After the formulation of the Liouville equation the construction of the left-hand side of eq. (4.27) is straightforward. The second term on the right-hand side of the equation remains unchanged, as to progress further we require more information on the outward normal on the space of admissible velocity data $\hat{\boldsymbol{\nu}}_{\boldsymbol{V}}$. The first term on the right-hand side of eq. (4.27) can be derived by first noting that

$$\begin{aligned} &-\int_{\partial\mathcal{P}_\varepsilon}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\boldsymbol{V}\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{R}}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R}) \quad (4.55)\\ &=\frac{1}{\sqrt{2}\varepsilon^d}\int_{\mathbb{R}^6}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty}\delta_{\|\boldsymbol{r}-\bar{\boldsymbol{r}}\|-\varepsilon}\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_2],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\\ &\qquad\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}_1\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{r}_1\,\mathrm{d}\boldsymbol{r}_2, \end{aligned}$$

where δ_x is the Dirac measure for $x\in\mathbb{R}$. We perform a substitution on $\boldsymbol{r}_2$ using $\boldsymbol{n}=\boldsymbol{r}_2-\boldsymbol{r}_1$, which has Jacobian 1, to get

$$\begin{aligned} &-\int_{\partial\mathcal{P}_\varepsilon}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{V}\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{R}}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R}) \quad (4.56)\\ &=-\frac{1}{\sqrt{2}\varepsilon^d}\int_{\mathbb{R}^6}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty}\delta_{\|\boldsymbol{n}\|-\varepsilon}\phi(\boldsymbol{R},\boldsymbol{V},t)f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_2+\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\\ &\qquad\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}_1\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r}. \end{aligned}$$

Then by the definition of the Dirac measure we have that for any function $f:\mathbb{R}^d\to\mathbb{R}, a,b\in\mathbb{R}, b\neq 0$,

$$b^{d-1}\int_{\mathbb{R}^d} f(\boldsymbol{x})\delta_{\|\boldsymbol{x}\|-a}\,\mathrm{d}\boldsymbol{x}=\int_{\mathbb{R}^d} f(b\boldsymbol{x})\delta_{\|\boldsymbol{x}\|-\frac{a}{b}}\,\mathrm{d}\boldsymbol{x} \quad (4.57)$$

Thus we arrive at the result

$$\begin{aligned} &-\int_{\partial\mathcal{P}_\varepsilon}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty} f(\boldsymbol{R},\boldsymbol{V},t)\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\boldsymbol{V}\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{R}}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R}) \quad (4.58)\\ &=-\frac{1}{\sqrt{2}}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\int_{\mathcal{V}(\boldsymbol{R})}\int_{-\infty}^{\infty}\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)f([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\\ &\qquad\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}_1\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r}_1. \end{aligned}$$

This concludes the proof of corollary 16. For a more detailed set of equations we require more specific dynamics and interactions, so that collisional terms can be derived using the event rule and the admissible data for the system. We now construct the weak formulation of the Liouville equation and the BBGKY hierarchy for dynamics with constant external force and friction.

4.4 Example: Inelastic particles with gravity and friction

We consider the dynamics where

$$\boldsymbol{H}(\boldsymbol{R},\boldsymbol{V},t) = -\boldsymbol{G} - \gamma \boldsymbol{V}, \tag{4.59}$$

where $\boldsymbol{G} = [\boldsymbol{g},\boldsymbol{g}]$ for $\boldsymbol{g} \in \mathbb{R}^3$ and $\gamma > 0$. These are the same dynamics as in example 9, so we know the analytic forms of the free dynamics, and can fully determine the behaviour of the associated Liouville equation in the two-particle case.

4.4.1 Velocity cones and collision times

The reduced-difference dynamics $\tilde{\boldsymbol{\Phi}}^x_t(\boldsymbol{R},\boldsymbol{V})$ provide the dynamics of the relative position of the two particles:

$$\tilde{\boldsymbol{\Phi}}^x_t(\boldsymbol{R},\boldsymbol{V}) = \tilde{\boldsymbol{R}} + \frac{1}{\gamma}\tilde{\boldsymbol{V}}(1 - e^{-\gamma t}) \tag{4.60}$$

As the reduced-difference dynamics $\tilde{\boldsymbol{\Phi}}^x_t(\boldsymbol{R},\boldsymbol{V})$ follow straight lines (parametrised exponentially in $-\gamma t$) we see that two particles satisfying eq. (3.16) can experience at most one collision. The initial data can then be partitioned into non-interacting, pre-collisional and post-collisional situations.

For $\boldsymbol{R} \in \mathcal{P}_\varepsilon$ and $\boldsymbol{V} \in \mathbb{R}^6$, we write $L(\boldsymbol{R},\boldsymbol{V}) \subset \mathbb{R}^6$ to denote the line

$$L(\boldsymbol{R},\boldsymbol{V}) = \left\{ \boldsymbol{R} - \frac{\boldsymbol{g}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{V} + \frac{\boldsymbol{g}}{\gamma}\right)(1 - e^{-\gamma t}) : t \in \mathbb{R} \right\} \tag{4.61}$$

and define the two infinite half lines

$$L^-(\boldsymbol{R},\boldsymbol{V}) = \left\{ \boldsymbol{R} - \frac{\boldsymbol{g}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{V} + \frac{\boldsymbol{g}}{\gamma}\right)(1 - e^{-\gamma t}) : t \leq 0 \right\}, \tag{4.62}$$

$$L^+(\boldsymbol{R},\boldsymbol{V}) = \left\{ \boldsymbol{R} - \frac{\boldsymbol{g}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{R} + \frac{\boldsymbol{g}}{\gamma}\right)(1 - e^{-\gamma t}) : t \geq 0 \right\} \tag{4.63}$$

The velocity collision cone is then defined analogously

$$C(\boldsymbol{R}) = \left\{ \boldsymbol{V} \in \mathbb{R}^6 : L(\boldsymbol{R},\boldsymbol{V}) \cap \partial\mathcal{P}_\varepsilon \neq \emptyset \right\} \tag{4.64}$$

And we can split this set into pre-collisional and post-collisional velocities respectively:

$$C^-(\boldsymbol{R}) = \left\{ \boldsymbol{V} \in C(\boldsymbol{R}) : L^+(\boldsymbol{R},\boldsymbol{V}) \cap \partial\mathcal{P}_\varepsilon \neq \emptyset \right\}, \tag{4.65}$$

$$C^+(\boldsymbol{R}) = \left\{ \boldsymbol{V} \in C(\boldsymbol{R}) : L^-(\boldsymbol{R},\boldsymbol{V}) \cap \partial\mathcal{P}_\varepsilon \neq \emptyset \right\}. \tag{4.66}$$

We note that a constant external potential $\boldsymbol{G}$ does not have an affect on the shape of the collision cones, as it does not effect the dynamics determined by the relative distance of the particles. The frictional constant γ truncates the pre-collisional velocity cone (when compared to linear dynamics). However, for these dynamics, for any given $\boldsymbol{R} \in \mathcal{P}_\varepsilon$, all initial velocities $\boldsymbol{V} \in \mathbb{R}^6$ are admissible (*i.e.* $\mathcal{V}(\boldsymbol{R}) = \mathbb{R}^6$), and so in particular the second surface term in the weak formulation of the Liouville equation disappears as the surface of the space $\mathcal{V}(\boldsymbol{R}) = \mathbb{R}^6$ is the empty set. We also have an

analytic form of the unique event time $\tau(\boldsymbol{R}, \boldsymbol{V})$, given by eq. (3.14). We can also determine which root to take, depending on if the initial data is pre-collisional or post-collisional.

Lemma 18 (Characterisation of the Collision Time Map for dynamics with gravity and friction)**.** For any $\boldsymbol{R} \in \mathcal{P}_\varepsilon$,

1. If $V \in \mathcal{C}^+(\boldsymbol{R})$ then $\tilde{\boldsymbol{V}} \cdot \tilde{\boldsymbol{R}} > 0$ and

$$\tau(\boldsymbol{R}, \boldsymbol{V}) = -\frac{1}{\gamma} \log\left(1 + \frac{\gamma}{\|\tilde{\boldsymbol{V}}\|}\left\{\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}} + \left[(\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}})^2 - (\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2)\right]^{\frac{1}{2}}\right\}\right). \tag{4.67}$$

2. If $V \in C^-(\boldsymbol{R})$ then $-\frac{1}{2}\left(\gamma(\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2) + \frac{\|\tilde{\boldsymbol{V}}\|^2}{\gamma}\right) < (\tilde{\boldsymbol{R}}) \cdot (\tilde{\boldsymbol{V}}) < 0$ and

$$\tau(\boldsymbol{R}, \boldsymbol{V}) = -\frac{1}{\gamma} \log\left(1 + \frac{\gamma}{\|\tilde{\boldsymbol{V}}\|}\left\{\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}} - \left[((\tilde{\boldsymbol{R}}) \cdot \hat{\tilde{\boldsymbol{V}}})^2 - (\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2)\right]^{\frac{1}{2}}\right\}\right). \tag{4.68}$$

Proof. The collision occurs when

$$\|\tilde{\Phi}_t^x(\boldsymbol{R}, \boldsymbol{V})\|^2 = \varepsilon^2 \implies \left\|\tilde{\boldsymbol{R}} + \frac{1}{\gamma}\tilde{\boldsymbol{V}}(1 - e^{-\gamma\tau})\right\|^2 = \varepsilon^2.$$

If the particles collide then $\tau(\boldsymbol{R}, \boldsymbol{V})$ must take one of the two following values

$$\tau^-(\boldsymbol{R}, \boldsymbol{V}) := -\frac{1}{\gamma} \log\left(1 + \frac{\gamma}{\|\tilde{\boldsymbol{V}}\|}\left\{\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}} - \left[(\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}})^2 - (\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2)\right]^{\frac{1}{2}}\right\}\right),$$
$$\tau^+(\boldsymbol{R}, \boldsymbol{V}) := -\frac{1}{\gamma} \log\left(1 + \frac{\gamma}{\|\tilde{\boldsymbol{V}}\|}\left\{\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}} + \left[(\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}})^2 - (\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2)\right]^{\frac{1}{2}}\right\}\right).$$

We note that the argument in the square root requires

$$\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}} < -(\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2)^{\frac{1}{2}} \text{ or } \tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}} > (\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2)^{\frac{1}{2}},$$

and if both $\tau^\pm(\boldsymbol{R}, \boldsymbol{V})$ exist, then $\tau^-(\boldsymbol{R}, \boldsymbol{V}) < \tau^+(\boldsymbol{R}, \boldsymbol{V})$. Assume now that $\boldsymbol{V} \in C^+(\boldsymbol{R})$, then $\tau^\pm(\boldsymbol{R}, \boldsymbol{V}) < 0$, and so $\tau^+(\boldsymbol{R}, \boldsymbol{V}) > \tau^-(\boldsymbol{R}, \boldsymbol{V})$, so $\tau(\boldsymbol{R}, \boldsymbol{V}) = \tau^+(\boldsymbol{R}, \boldsymbol{V})$ is required.

Furthermore,

$$\tilde{\boldsymbol{R}} \cdot \tilde{\boldsymbol{V}} > \frac{\|\tilde{\boldsymbol{V}}\|}{\gamma} + ((\tilde{\boldsymbol{R}} \cdot \tilde{\boldsymbol{V}})^2 - (\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2))^{\frac{1}{2}} > 0.$$

and so $\tilde{\boldsymbol{R}} \cdot \tilde{\boldsymbol{V}} > 0$. Alternatively, if $\boldsymbol{V} \in C^-(\boldsymbol{R})$, $\tau(\boldsymbol{R}, \boldsymbol{V}) > 0$ and so we take $\tau(\boldsymbol{R}, \boldsymbol{V}) = \tau^-(\boldsymbol{R}, \boldsymbol{V})$. Thus

$$\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}} + \frac{\|\tilde{\boldsymbol{V}}\|}{\gamma} > ((\tilde{\boldsymbol{R}} \cdot \hat{\tilde{\boldsymbol{V}}})^2 - (\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2))^{\frac{1}{2}},$$

Squaring both sides and rearranging, we find

$$\tilde{\boldsymbol{R}} \cdot \tilde{\boldsymbol{V}} > -\frac{1}{2}\left(\gamma(\|\tilde{\boldsymbol{R}}\|^2 - \varepsilon^2) + \frac{\|\tilde{\boldsymbol{V}}\|}{\gamma}\right).$$

□

The inequality in the pre-collisional case relates to the presence of friction in the dynamics; if particles do not have enough energy in the direction $\tilde{\boldsymbol{R}}$ then the particles will never meet.

4.4.2 Flow maps

Given the free-particle dynamics, the scattering maps, and a full characterisation of the admissible data, we are now in a position to fully define the hard-sphere flow maps $\boldsymbol{T}_t : \mathbb{R}^6 \times \mathbb{R}^6 \times \mathbb{R} \to \mathbb{R}^6 \times \mathbb{R}^6$ for this example. We define $\boldsymbol{\Pi}_1 : \mathbb{R}^6 \times \mathbb{R}^6 \to \mathbb{R}^6, \boldsymbol{\Pi}_2 : \mathbb{R}^6 \times \mathbb{R}^6 \to \mathbb{R}^6$ to be functions that map the result of $\boldsymbol{T}_t$ to the position and velocity trajectories respectively. To write down the flow map in full, we consider collision-free and collisional dynamics separately.

Collision-free Dynamics

If $\boldsymbol{R} \in \mathcal{P}_\varepsilon$ and $\boldsymbol{V} \in \mathbb{R}^6 \backslash \mathcal{C}(\boldsymbol{R})$, then

$$(\boldsymbol{\Pi}_1 \circ \boldsymbol{T}_t)\boldsymbol{Z} = \boldsymbol{R} - \frac{\boldsymbol{G}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{V} + \frac{\boldsymbol{G}}{\gamma}\right)(1 - e^{-\gamma t}), \tag{4.69}$$

$$(\boldsymbol{\Pi}_2 \circ T_t)\boldsymbol{Z} = -\frac{\boldsymbol{G}}{\gamma} + \left(\boldsymbol{V} + \frac{\boldsymbol{G}}{\gamma}\right)e^{-\gamma t}. \tag{4.70}$$

Collisional Dynamics

Firstly, if $\boldsymbol{R} \in \mathcal{P}_\varepsilon$ and $\boldsymbol{V} \in \mathcal{C}^-(\boldsymbol{R})$, then

$$(\boldsymbol{\Pi}_1 \circ \boldsymbol{T}_t)\boldsymbol{Z} = \begin{cases} \boldsymbol{R} - \frac{\boldsymbol{G}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{V} + \frac{\boldsymbol{G}}{\gamma}\right)(1 - e^{-\gamma t}), & \text{if } t < \tau(\boldsymbol{R}, \boldsymbol{V}) \\ \left[\boldsymbol{R} - \dfrac{\boldsymbol{G}\tau}{\gamma} + \dfrac{1}{\gamma}\left(\boldsymbol{V} + \dfrac{\boldsymbol{G}}{\gamma}\right)(1 - e^{-\gamma\tau})\right] - \dfrac{\boldsymbol{G}(t-\tau)}{\gamma} \\ + \dfrac{1 - e^{-\gamma(t-\tau)}}{\gamma}\left[\boldsymbol{\sigma}_-(\boldsymbol{R}, \boldsymbol{V})\left(\left(\boldsymbol{V} + \dfrac{\boldsymbol{G}}{\gamma}\right)e^{-\gamma\tau} - \dfrac{\boldsymbol{G}}{\gamma}\right) + \dfrac{\boldsymbol{G}}{\gamma}\right], & \text{if } \tau(\boldsymbol{R}, \boldsymbol{V}) < t \end{cases} \tag{4.71}$$

and

$$(\boldsymbol{\Pi}_2 \circ \boldsymbol{T}_t)\boldsymbol{Z} = \begin{cases} -\frac{\boldsymbol{G}}{\gamma} + \left(\boldsymbol{V} + \frac{\boldsymbol{G}}{\gamma}\right)e^{-\gamma t}, & \text{if } t < \tau(\boldsymbol{R}, \boldsymbol{V}) \\ -\frac{\boldsymbol{G}}{\gamma} + \left[\boldsymbol{\sigma}_-(\boldsymbol{R}, \boldsymbol{V})\left((\boldsymbol{V} + \frac{\boldsymbol{G}}{\gamma})e^{-\gamma\tau} - \frac{\boldsymbol{G}}{\gamma}\right) + \frac{\boldsymbol{G}}{\gamma}\right]e^{-\gamma(t-\tau)}, & \text{if } \tau(\boldsymbol{R}, \boldsymbol{V}) < t \end{cases} \tag{4.72}$$

If $\boldsymbol{R} \in \mathcal{P}_\varepsilon$ and $\boldsymbol{V} \in \mathcal{C}^+(\boldsymbol{R})$, then

$$(\boldsymbol{\Pi}_1 \circ \boldsymbol{T}_t)\boldsymbol{Z} = \begin{cases} \left[\boldsymbol{R} - \dfrac{\boldsymbol{G}\tau}{\gamma} + \dfrac{(1 - e^{-\gamma\tau})}{\gamma}\left(\boldsymbol{V} + \dfrac{\boldsymbol{G}}{\gamma}\right)\right] - \dfrac{\boldsymbol{G}(t-\tau)}{\gamma} \\ + \dfrac{1 - e^{-\gamma(t-\tau)}}{\gamma}\left[\boldsymbol{\sigma}_+(\boldsymbol{R}, \boldsymbol{V})\left(\left(\boldsymbol{V} + \dfrac{\boldsymbol{G}}{\gamma}\right)e^{-\gamma\tau} - \dfrac{\boldsymbol{G}}{\gamma}\right) + \dfrac{\boldsymbol{G}}{\gamma}\right], & \text{if } t < \tau(\boldsymbol{R}, \boldsymbol{V}) \\ \boldsymbol{R} - \frac{\boldsymbol{G}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{V} + \frac{\boldsymbol{G}}{\gamma}\right)(1 - e^{-\gamma t}), & \text{if } \tau(\boldsymbol{R}, \boldsymbol{V}) < t \end{cases} \tag{4.73}$$

and

$$(\mathbf{\Pi}_2 \circ \boldsymbol{T}_t)\boldsymbol{Z} = \begin{cases} -\frac{\boldsymbol{G}}{\gamma} + \left[\boldsymbol{\sigma}_+(\boldsymbol{R},\boldsymbol{V})\left((\boldsymbol{V}+\frac{\boldsymbol{G}}{\gamma})e^{-\gamma\tau} - \frac{\boldsymbol{G}}{\gamma}\right) + \frac{\boldsymbol{G}}{\gamma}\right] e^{-\gamma(t-\tau)}, & \text{if } t < \tau(\boldsymbol{R},\boldsymbol{V}) \\ -\frac{\boldsymbol{G}}{\gamma} + \left(\boldsymbol{V}+\frac{\boldsymbol{G}}{\gamma}\right) e^{-\gamma t}, & \text{if } \tau(\boldsymbol{R},\boldsymbol{V}) < t. \end{cases} \tag{4.74}$$

Finally, if $\boldsymbol{R} \in \partial\mathcal{P}_\varepsilon$

$$(\mathbf{\Pi}_1 \circ \boldsymbol{T}_t)\boldsymbol{Z} = \begin{cases} \boldsymbol{R} - \frac{\boldsymbol{G}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{V}+\frac{\boldsymbol{G}}{\gamma}\right)(1-e^{-\gamma t}), & \text{if } t < 0 \\ \boldsymbol{R} - \frac{\boldsymbol{G}t}{\gamma} + \frac{1}{\gamma}\left(\boldsymbol{\sigma}_-(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V}+\frac{\boldsymbol{G}}{\gamma}\right)(1-e^{-\gamma t}), & \text{if } 0 < t \end{cases} \tag{4.75}$$

and

$$(\mathbf{\Pi}_2 \circ \boldsymbol{T}_t)\boldsymbol{Z} = \begin{cases} -\frac{\boldsymbol{G}}{\gamma} + \left(\boldsymbol{V}+\frac{\boldsymbol{G}}{\gamma}\right) e^{-\gamma t}, & \text{if } -\infty < t < 0 \\ -\frac{\boldsymbol{G}}{\gamma} + \left(\boldsymbol{\sigma}_-(\boldsymbol{R},\boldsymbol{V})\boldsymbol{V}+\frac{\boldsymbol{G}}{\gamma}\right) e^{-\gamma t}, & \text{if } 0 < t < \infty. \end{cases} \tag{4.76}$$

Two-dimensional diagrams of the trajectories that can occur in this system of dynamics were provided in fig. 2.1, the result in three dimensions is similar. We note that in fig. 2.1a, although the trajectories overlap, the particles do not collide as they reach the only potential intersection point at different times.

4.4.3 The Liouville equation

Using theorem 14, as we understand the admissible data $\boldsymbol{R},\boldsymbol{V}$ for the particle dynamics, we can write down the Liouville equation for these particular dynamics. For any $f_0^{(2)} \in C^0(\mathcal{D}) \cap L^1(\mathcal{D})$, there exists a physical global-in-time weak solution of

$$\left[\frac{\partial}{\partial t} + \boldsymbol{V}\cdot\nabla_{\boldsymbol{R}} - \boldsymbol{G}\cdot\nabla_{\boldsymbol{V}} - \nabla_{\boldsymbol{V}}\cdot(\gamma\boldsymbol{V})\right] f^{(2)}(\boldsymbol{R},\boldsymbol{V},t) = C_{\boldsymbol{R}}[f^{(2)}], \tag{4.77}$$

where the scattering map defining collisions satisfies eq. (2.29).

4.4.4 The BBGKY hierarchy

We define $f^{(1)}(\boldsymbol{r}_1,\boldsymbol{v}_1,t)$ as in corollary 16, which then satisfies (by using $\boldsymbol{G} = [\boldsymbol{g},\boldsymbol{g}]$),

$$\begin{aligned} &\int_{\mathcal{P}_\varepsilon}\int_{\mathbb{R}^6}\int_{-\infty}^{\infty} (\partial_t + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{r}_1} + \boldsymbol{g}\cdot\nabla_{\boldsymbol{v}_1})\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t) f^{(1)}(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R} \\ &= -\frac{1}{\sqrt{2}}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\int_{\mathbb{R}^6}\int_{-\infty}^{\infty} \phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t) f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r}_1. \end{aligned} \tag{4.78}$$

The collisional term $C(f^{(2)})$ on the right-hand side of the above equation can then be separated into a pre-collisional and a post-collisional term:

$$\begin{aligned} &C[f^{(2)}] = \\ &-\frac{1}{\sqrt{2}}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\int_{C^-(\boldsymbol{R})}\int_{-\infty}^{\infty} \phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t) f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r}_1 \end{aligned}$$

$$-\frac{1}{\sqrt{2}}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\int_{C^+(\boldsymbol{R})}\int_{-\infty}^{\infty}\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_2+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r}_1,$$

where

$$C^-(\boldsymbol{n})=\{\boldsymbol{V}\in\mathbb{R}^6:(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0\}$$

and

$$\begin{aligned}C^+(\boldsymbol{n})=&\{\boldsymbol{V}=[\boldsymbol{v}_1,\boldsymbol{v}_2]\in\mathbb{R}^6:-\frac{\|\boldsymbol{v}_1-\boldsymbol{v}_2\|^2}{2\gamma}<(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0\}\\=&\{\boldsymbol{V}=[\boldsymbol{v}_1,\boldsymbol{v}_2]\in\mathbb{R}^6:(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0\},\end{aligned}$$

where the lower bound has disappeared because the integral is over $\boldsymbol{R}\in\partial\mathcal{P}_\varepsilon$. We introduce the change of variables for the pre-collisional integral that is provided by the forward-time scattering map:

$$\boldsymbol{V}\mapsto\left(I-\frac{1+\alpha}{2}\begin{Bmatrix}\boldsymbol{n}\\-\boldsymbol{n}\end{Bmatrix}\otimes\begin{Bmatrix}\boldsymbol{n}\\-\boldsymbol{n}\end{Bmatrix}\right)\boldsymbol{V}.$$

This transform has Jacobian $1/\alpha$, which can be calculated directly or by computing the eigenvalues of $\boldsymbol{\sigma}^+(\boldsymbol{R},\boldsymbol{V})$, taking the product and the absolute value of the result. We note that

$$(\boldsymbol{v}_1^{\mathrm{in}}-\boldsymbol{v}_2^{\mathrm{in}})\cdot\boldsymbol{n}=-\frac{1}{\alpha}(\boldsymbol{v}_1^{\mathrm{out}}-\boldsymbol{v}_2^{\mathrm{out}})\cdot\boldsymbol{n}.$$

Using this change of variables, we have the weak inelastic Boltzmann collision operator appearing in the BBGKY hierarchy, an analogous result to the elastic case in [91]:

$$\begin{aligned}&\int_{\partial\mathcal{P}_\varepsilon}\int_{\mathbb{R}^6}\int_{-\infty}^{\infty}\Phi(\boldsymbol{R},\boldsymbol{V},t)f^{(2)}(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{V}\cdot\tilde{\boldsymbol{\nu}}_{\boldsymbol{R}}(\boldsymbol{R})\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R},\\=&\frac{1}{\sqrt{2}}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\int_{C^+(\boldsymbol{n})}\int_{-\infty}^{\infty}\Bigg[\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\\&\qquad-\frac{1}{\alpha^2}\phi(\boldsymbol{r}_1,\boldsymbol{v}_1^{\mathrm{out}},t)f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_2+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\mathrm{out}},\boldsymbol{v}_2^{\mathrm{out}}],t)\Bigg](\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r}_1.\end{aligned}$$

By a suitable choice of the test functions ϕ (*i.e.* Dirac functions in their arguments) we have a suitable candidate for the strong form of the collision operator for spherical particles of diameter ε, which we will use in derivations for a continuum model:

$$\begin{aligned}\mathcal{L}_{\mathrm{coll}}[f^{(2)}]=&-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\\&\times\left[\frac{1}{\alpha^2}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\mathrm{out}},\boldsymbol{v}_2^{\mathrm{out}}],t)-f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\right]\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}.\end{aligned}\tag{4.79}$$

The result eq. (4.79) is similar, but not equivalent to, the Revised-Enskog-Theory (RET) collision operator [128], which is popular in the literature:

$$\mathcal{L}^{\mathrm{RET}}_{\mathrm{coll}}[f^{(2)}] = -\varepsilon^{d-1}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n} \tag{4.80}$$
$$\times\left[\frac{1}{\alpha^2}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1-\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\mathrm{out}},\boldsymbol{v}_2^{\mathrm{out}}],t)-f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\right]\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}.$$

The RET collision operator is popular in the modelling of dense granular flow [74], [129] and it is clear that in one dimension with $\varepsilon\to 0$ the two collision operators are equivalent. In comparison to eq. (4.79) the normalisation prefactor scales with the size of the particle, and there is a negation in one of the terms. Neglecting the normalisation term at first it appears that by using suitable variable changes, eq. (4.79) and eq. (4.80) are equivalent. However, as we see in one dimension:

$$\begin{aligned}\mathcal{L}_{\mathrm{coll}}[f^{(2)}] = \frac{1}{\sqrt{2}}\int_{\mathbb{R}}&\chi_{(v_1-v_2)>0}(v_1-v_2)\\ &\times\left[\frac{1}{\alpha^2}f^{(2)}([r_1,r_1+\varepsilon],[v_1^{\mathrm{out}},v_2^{\mathrm{out}}],t)-f^{(2)}([r_1,r_1+\varepsilon],[v_1,v_2],t)\right]\\ &-\chi_{(v_1-v_2)<0}(v_1-v_2)\\ &\times\left[\frac{1}{\alpha^2}f^{(2)}([r_1,r_1-\varepsilon],[v_1^{\mathrm{out}},v_2^{\mathrm{out}}],t)-f^{(2)}([r_1,r_1-\varepsilon],[v_1,v_2],t)\right]\mathrm{d}v_2,\end{aligned}$$

$$\begin{aligned}\mathcal{L}^{\mathrm{RET}}_{\mathrm{coll}}[f^{(2)}] = \int_{\mathbb{R}}(v_1-v_2)&\left[\frac{1}{\alpha^2}\chi_{(v_1-v_2)>0}f^{(2)}([r_1,r_1+\varepsilon],[v_1^{\mathrm{out}},v_2^{\mathrm{out}}],t)\right.\\ &\left.+\chi_{(v_1-v_2)<0}f^{(2)}([r_1,r_1+\varepsilon],[v_1,v_2],t)\right]\\ -(v_1-v_2)&\left[\frac{1}{\alpha^2}\chi_{(v_1-v_2)<0}f^{(2)}([r_1,r_1-\varepsilon],[v_1^{\mathrm{out}},v_2^{\mathrm{out}}],t)\right.\\ &\left.+\chi_{(v_1-v_2)>0}f^{(2)}([r_1,r_1-\varepsilon],[v_1,v_2],t)\right]\mathrm{d}v_2.\end{aligned}$$

The placement of characteristic functions in the one-dimensional RET operator means that no valid variable change will make the two integrals identical. As collisions adjust velocities in one dimension, from this result (or from equivalent calculations in higher dimensions), we can see that the two collision operators are not proportional.

To justify the derived collision operator we consider the classical derivation of the Boltzmann equation in one dimension, as discussed in [130]. Here the Boltzmann equation and the associated collision operator is derived by considering gains and losses caused by a collision in a range of velocities $\mathrm{d}\boldsymbol{v}_1$. The result is the Boltzmann collision integral

$$\mathcal{L}^{\mathrm{Boltz}}_{\mathrm{coll}}[f^{(1)}] = \int|\boldsymbol{v}_1-\boldsymbol{v}_2|[f^{(1)}(\boldsymbol{r}_1,\boldsymbol{v}_1^{\mathrm{out}})f^{(1)}(\boldsymbol{r}_2,\boldsymbol{v}_2^{\mathrm{out}})-f^{(1)}(\boldsymbol{r}_1,\boldsymbol{v}_1)f^{(1)}(\boldsymbol{r}_2,\boldsymbol{v}_2)]\,\mathrm{d}\sigma\,\mathrm{d}\boldsymbol{v}_1 \tag{4.81}$$

where $\mathrm{d}\sigma$ integrates over all positions $\boldsymbol{r}_2$ that omit a collision, and $|\boldsymbol{v}_1-\boldsymbol{v}_2|$ occurs due to a relation on the collision cross section. In our case we consider the relative velocity in the direction of the collision $|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}|$ (which can be treated similarly to the one-dimensional case, in principle)

and the positions that omit a collision are restricted to $\boldsymbol{r}_2 = \boldsymbol{r}_1 + \varepsilon\boldsymbol{n}$ where $(\boldsymbol{v}_1 - \boldsymbol{v}_2)\cdot(\boldsymbol{r}_1 - \boldsymbol{r}_2) > 0$. Thus the Boltzmann collision integral becomes

$$\mathcal{L}_{\text{coll}}^{\text{Boltz}}[f^{(1)}] = \int \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}|(\boldsymbol{v}_1 - \boldsymbol{v}_2)\cdot\boldsymbol{n}| \times [f^{(1)}(\boldsymbol{r}_1, \boldsymbol{v}_1^{\text{out}})f^{(1)}(\boldsymbol{r}_1 + \varepsilon\boldsymbol{n}, \boldsymbol{v}_2^{\text{out}}) - f^{(1)}(\boldsymbol{r}_1, \boldsymbol{v}_1)f^{(1)}(\boldsymbol{r}_1 + \varepsilon\boldsymbol{n}, \boldsymbol{v}_2)]\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{v}_1. \tag{4.82}$$

The Boltzmann equation also assumes *molecular chaos*, *i.e.* that one-body distributions are uncorrelated:

$$f^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{v}_1, \boldsymbol{v}_2, t) = f^{(1)}(\boldsymbol{r}_1, \boldsymbol{v}_1, t)f^{(1)}(\boldsymbol{r}_1, \boldsymbol{v}_1, t). \tag{4.83}$$

If we 'reverse' the molecular chaos assumption in eq. (4.82), we arrive at eq. (4.79). This suggests that eq. (4.79) is a reasonable choice.

Furthermore, in contrast to the RET collision operator there is no prefactor involving ε in the Boltzmann collision operator. To argue for the exclusion of a diameter prefactor in the collision operator we consider a dimensional argument. We define scales

$$\boldsymbol{v} \sim V, \boldsymbol{r} \sim L, t \sim T. \tag{4.84}$$

Then

$$\int_{\mathcal{P}_\varepsilon}\int_{\mathbb{R}^3} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}t)\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{r}_2 \sim VLf^{(1)}(\boldsymbol{r}, \boldsymbol{v}, t), \tag{4.85}$$

$$(\partial_t + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{r}_1}) \sim \frac{1}{T}, \quad \mathcal{L}_{\text{coll}}[f^{(2)}] \sim \frac{1}{T} \quad \mathcal{L}_{\text{coll}}^{\text{RET}}[f^{(2)}] \sim \frac{[\varepsilon]^{d-1}}{T} \tag{4.86}$$

so that the left and right-hand side of eq. (4.78) balance. When using the RET collision operator an additional dimensional parameter ε is introduced on the right-hand side of eq. (4.78), and the right-hand side no longer balances the left-hand side.

4.4.5 Extension to general external and interaction potentials

By considering this example we have constructed a valid form of collision operator for inelastic hard-sphere systems. In this example, as the relative trajectory of the particles is linear, two particles can only experience one event, namely the inelastic collision. By including more complicated smooth interaction potentials (*e.g.* a quadratic interaction potential which attracts particles to one another) multiple collisions can occur between two particles. The resulting Lioville equation will still be described by eq. (4.21), under the condition that the collisions that particles experience are well separated.

Alternatively, multiple events can occur between particles by including additional interaction diameters where non-collisional events can occur. We now present an example where multiple instantaneous interactions can occur between two particles.

4.5 Example: Square shoulder particles

We now include an additional interaction diameter $\varepsilon_1 > \varepsilon > 0$ where particle trajectories are refracted away from one another, provided they can overcome an energy barrier. If the particles do not have enough energy to overcome the barrier then they are instead reflected away.

For this example we therefore consider the free dynamics given by eq. (3.5), so that event times between the particles at ε and ε_1 can be constructed using the collision time in eq. (3.7). The event at the diameter ε_1 is given by example 5 with $a_1 > 0$, and at the core diameter ε we consider a pure elastic collision given by example 3.

4.5.1 Flow maps

Given the collision times $\tau(\boldsymbol{R}, \boldsymbol{V})$ for all initial data $\boldsymbol{R} \in \mathcal{P}_\varepsilon$, $\boldsymbol{V} \in \mathbb{R}^6$, it is then possible to construct flow maps for every possible pair of trajectories for two particles. We reduce notational burden by omitting the precise details of the flow maps, and including an overview of the types of trajectories that can occur. A diagram of the possible trajectories were given in fig. 2.3, along with the corresponding trajectories in the reduced-difference phase space in fig. 2.4.

We define the subset of initial data that has an interaction with the surface $\partial\mathcal{P}_{\varepsilon_1}$ as

$$\mathcal{C}_1(\boldsymbol{R}) = \{\boldsymbol{V} \in \mathbb{R}^6 : \exists t \in \mathbb{R} : \|\tilde{\Psi}_t^{\boldsymbol{r}}\| = \varepsilon_1\}. \tag{4.87}$$

Collision-free dynamics

For $\boldsymbol{R} \in \mathcal{P}_{\varepsilon_1}$ and $\boldsymbol{V} \in \mathbb{R}^6 \backslash \mathcal{C}_1(\boldsymbol{R})$, the two particles never reach the interaction perimeter $\partial\mathcal{P}_{\varepsilon_1}$, so their dynamics are described by the free dynamics eq. (3.5).

External bounce

When $\boldsymbol{R} \in \mathcal{P}_{\varepsilon_1}$ and $\boldsymbol{V} \in \mathcal{C}_1(\boldsymbol{R})$, such that $((\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2))^2 < a$, the particles experience a single elastic collision at the boundary $\partial\mathcal{P}_{\varepsilon_1}$ and are reflected away from one another.

Refractive dynamics

For all other initial data, *i.e.* either $\boldsymbol{R} \in \mathcal{P}_\varepsilon \backslash \mathcal{P}_{\varepsilon_1}$ and $\boldsymbol{V} \in \mathbb{R}^6$, or $\boldsymbol{R} \in \mathcal{P}_{\varepsilon_1}$ and $\boldsymbol{V} \in \mathcal{C}_1(\boldsymbol{R})$ such that $((\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2))^2 > a$, there are two possible trajectory types:

1. The particles are refracted at the boundary $\partial\mathcal{P}_{\varepsilon_1}$ at time τ_0 with outgoing velocity $\Psi^{\boldsymbol{v}}_{\tau_0}(\boldsymbol{R}, \boldsymbol{V})$ such that $\Psi^{\boldsymbol{v}}_{\tau_0}(\boldsymbol{R}, \boldsymbol{V}) \in \mathbb{R}^6 \backslash \mathcal{C}(\Psi^{\boldsymbol{r}}_{\tau_0}(\boldsymbol{R}, \boldsymbol{V}))$. In this case the particles will experience a further refraction event at time τ_1 when leaving the space $\mathcal{P}_\varepsilon \backslash \mathcal{P}_{\varepsilon_1}$

2. The particles are refracted at the boundary $\partial\mathcal{P}_{\varepsilon_1}$ at time τ_0, with outgoing velocity $\Psi^{\boldsymbol{v}}_{\tau_0}(\boldsymbol{R}, \boldsymbol{V})$ such that $\Psi^{\boldsymbol{v}}_{\tau_0}(\boldsymbol{R}, \boldsymbol{V}) \in \mathcal{C}(\Psi^{\boldsymbol{r}}_{\tau_0}(\boldsymbol{R}, \boldsymbol{V}))$. The particles will then experience a core collision at time τ_1 before leaving the space $\mathcal{P}_\varepsilon \backslash \mathcal{P}_{\varepsilon_1}$ at time τ_2 and experiencing a final refraction at the perimeter $\partial\mathcal{P}_{\varepsilon_1}$.

We consider these two interactions together as we shall see that they provide the same contribution to the Liouville equation.

4.5.2 The Liouville equation

To construct the Liouville equation we partition the space of admissible initial data into three parts, then sum the corresponding Liouville equations to construct the Liouville equation for square-shoulder particles.

Collision-free dynamics

We consider trajectories where $\boldsymbol{R} \in \mathcal{P}_{\varepsilon_1}$ and $\boldsymbol{V} \in \mathbb{R}^6 \backslash \mathcal{C}_1(\boldsymbol{R})$. Then eq. (4.21) provides us with the Liouville equation for this system of admissible initial data

$$\begin{aligned}
&\int_{\mathcal{P}_{\varepsilon_1}} \int_{\mathbb{R}^6 \backslash \mathcal{C}(\boldsymbol{R})} \int_{-\infty}^{\infty} f^{(2)}(\boldsymbol{R}, \boldsymbol{V}, t)[\partial_t \Phi(\boldsymbol{R}, \boldsymbol{V}, t) + \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t)] \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} = \\
&- \int_{\partial \mathcal{P}_{\varepsilon_1}} \int_{\mathbb{R}^6 \backslash \mathcal{C}_1(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}, t) \Phi(\boldsymbol{R}, \boldsymbol{V}, t) \boldsymbol{V} \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{R}} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\mathcal{H}(\boldsymbol{R}).
\end{aligned} \tag{4.88}$$

In this case the collision term is zero, as for $\boldsymbol{R} \in \partial \mathcal{P}_{\varepsilon_1}$, $\mathbb{R}^6 \backslash \mathcal{C}_1(\boldsymbol{R}) = \emptyset$.

External bounce

For the set $\boldsymbol{R} \in \mathcal{P}_{\varepsilon_1}$ and $\boldsymbol{V} \in \{[\boldsymbol{v}_1, \boldsymbol{v}_2] \in \mathbb{R}^6 : |(\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2)|^2 < a\} =: \mathcal{V}_1(\boldsymbol{R})$ we have

$$\begin{aligned}
&\int_{\mathcal{P}_{\varepsilon_1}} \int_{\mathcal{V}_1(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}, t)[\partial_t \Phi(\boldsymbol{R}, \boldsymbol{V}, t) + \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t)] \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} = \\
&- \int_{\partial \mathcal{P}_{\varepsilon_1}} \int_{V_1(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}, t) \Phi(\boldsymbol{R}, \boldsymbol{V}, t) V \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{R}} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\mathcal{H}(\boldsymbol{R}).
\end{aligned} \tag{4.89}$$

When $\boldsymbol{R} \in \partial \mathcal{P}_{\varepsilon_1}$, $\mathcal{V}(\boldsymbol{R}) = \{\boldsymbol{V} \in \mathcal{C}_1(\boldsymbol{R}) : |(\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2)|^2 < a\}$.

Refractive dynamics

For dynamics where the particle has sufficient energy to overcome the barrier at ε_1 we must consider $\boldsymbol{R} \in \mathcal{P}_{\varepsilon_1}$ and $\boldsymbol{V}$ in the set

$$\mathcal{V}_2(\boldsymbol{R}) := \{\boldsymbol{V} = [\boldsymbol{v}_1, \boldsymbol{v}_2] \in \mathbb{R}^6 : \boldsymbol{V} \in \mathcal{C}_1(\boldsymbol{R}) \text{ and } |(\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot (\boldsymbol{r}_1 - \boldsymbol{r}_2)|^2 > a \text{ if } \boldsymbol{R} \in \mathcal{P}_{\varepsilon_1}\}. \tag{4.90}$$

In this case there is more than one instantaneous interaction in the system, but by the results of theorem 14 we know that these interactions will not directly contribute to the Liouville equation but will remain embedded in $f^{(2)}$. Thus we have

$$\begin{aligned}
&\int_{\mathcal{P}_{\varepsilon}} \int_{\mathcal{V}_2(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}, t)[\partial_t \Phi(\boldsymbol{R}, \boldsymbol{V}, t) + \boldsymbol{V} \cdot \nabla_{\boldsymbol{R}} \Phi(\boldsymbol{R}, \boldsymbol{V}, t)] \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\boldsymbol{R} = \\
&- \int_{\partial \mathcal{P}_{\varepsilon}} \int_{\mathcal{V}_2(\boldsymbol{R})} \int_{-\infty}^{\infty} f(\boldsymbol{R}, \boldsymbol{V}, t) \Phi(\boldsymbol{R}, \boldsymbol{V}, t) V \cdot \hat{\boldsymbol{\nu}}_{\boldsymbol{R}} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{V} \, \mathrm{d}\mathcal{H}(\boldsymbol{R}).
\end{aligned} \tag{4.91}$$

For the surface integral term, when $\boldsymbol{R} \in \partial \mathcal{P}_{\varepsilon}$ we have $\boldsymbol{V} \in \mathbb{R}^6$.

Combining terms

Combining the results of eqs. (4.88) to (4.89) we return the Liouville equation for square-shoulder particle dynamics:

$$\begin{aligned}
&\int_{\mathcal{P}_\varepsilon}\int_{\mathbb{R}^6}\int_{-\infty}^{\infty} f(\boldsymbol{R},\boldsymbol{V},t)[\partial_t\Phi(\boldsymbol{R},\boldsymbol{V},t)+\boldsymbol{V}\cdot\nabla_{\boldsymbol{R}}\Phi(\boldsymbol{R},\boldsymbol{V},t)]\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\boldsymbol{R}=\\
&-\int_{\partial\mathcal{P}_\varepsilon}\int_{\mathbb{R}^6}\int_{-\infty}^{\infty} f(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)\boldsymbol{V}\cdot\hat{\boldsymbol{\nu}}_{\boldsymbol{R}}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R})\\
&-\int_{\partial\mathcal{P}_{\varepsilon_1}}\int_{\mathcal{V}_3(\boldsymbol{R})}\int_{-\infty}^{\infty} f(\boldsymbol{R},\boldsymbol{V},t)\Phi(\boldsymbol{R},\boldsymbol{V},t)V\cdot\hat{\nu}_{\boldsymbol{R}}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{V}\,\mathrm{d}\mathcal{H}(\boldsymbol{R})
\end{aligned}\tag{4.92}$$

where

$$\mathcal{V}_3(\boldsymbol{R}):=\{[\boldsymbol{v}_1,\boldsymbol{v}_2]\in\mathbb{R}^6:|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)|^2<a\}.\tag{4.93}$$

Thus the inclusion of a square-shoulder potential term produces a partial collision operator term at the Liouville equation, and information on refraction events between particles is absorbed into the left-hand side of the equation.

We note that upon construction of the BBGKY hierarchy for square-shoulder potential systems, in addition to the standard Boltzmann collision operator given by eq. (4.79), another collision term is included of the form

$$\begin{aligned}
\mathcal{L}^a_{\mathrm{coll}}[f^{(2)}]=&\frac{1}{\sqrt{2}}\int_{\mathbb{R}^3}\int_{\mathbb{S}^2}\int_{C_a^+(\boldsymbol{n})}\int_{-\infty}^{\infty}\Phi(\boldsymbol{r},\boldsymbol{v}_1,y)\Big[f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon_1\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\\
&-f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon_1\boldsymbol{n}],[\boldsymbol{v}_1^{\mathrm{out}},\boldsymbol{v}_2^{\mathrm{out}}],t)\Big](\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}_1\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r},
\end{aligned}\tag{4.94}$$

where

$$C_a^+(n)=\{[\boldsymbol{v}_1,\boldsymbol{v}_2]\in\mathbb{R}^6:(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)>0\text{ and }|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot(\boldsymbol{r}_1-\boldsymbol{r}_2)|^2<a\}.\tag{4.95}$$

The additional contribution is due to the collisional effects at the external interaction diameter. In addition the function $f^{(2)}$ (and in turn $f^{(1)}$) must obey a kinetic energy law on $\partial\mathcal{P}_{\varepsilon_1}$, so that it obeys eq. (2.36). We expect that this will result in a scaling prefactor a in the Liouville equation, within the exterior interaction perimeter.

4.6 Modelling granular media using the Liouville equation

In this chapter we have constructed the weak formulation of the Liouville equation using minimal assumptions on the microscopic dynamics. Our derivation only considers two particles but it is reasonable to assume that, under the assumptions that collisions are binary, results on the collision operator can be used in systems of N particles. As a result of a rigorous argument we have derived additional terms (that are not present in the standard derivation of the Liouville equation [118]) in the weak form of the Liouville equation, by carefully identifying the set of admissible data that accounts for possible instantaneous interactions between particles. As a result of this careful

derivation, the construction of the BBGKY hierarchy, including associated collision operators, is straightforward. The results are equivalent to other rigorous derivations of the BBGKY hierarchy for hard spheres, *e.g.* [46], [120], but our method allows us to easily include additional interaction potentials and instantaneous interactions. We have provided motivating examples to show the breadth of the results presented here.

However, due to the high dimensionality of the Liouville equation for N particles it is not usually possible to simulate a system of granular media directly using the PDE described; the Liouville equation is an dN-dimensional partial differential equation, which is far too large to discretize and simulate numerically using standard schemes. We therefore need to derive a continuum model from the Liouville equation, which reduces the dimensionality of the problem with minimal effect on the accuracy of the model. In the next chapter we will start to reduce the dimensionality of the continuum model by constructing the BBGKY hierarchy for a system of N particles with dynamical properties of interest.

Inspired by the weak formulation of the Liouville equation we shall assume that collisional effects can be included in our continuum model for N particles by using the collision operator in eq. (4.79). By doing so, we are assuming that the only instantaneous interactions in the system are due to inelastic collisions, and that all $\boldsymbol{v}_0^N \in \mathbb{R}^{dN}$ are admissible as initial conditions. We also assume that all collisions are binary and that there is a minimum time $\delta > 0$ that separates all collisions. Although this will remove physical phenomena from the system (*i.e.* inelastic collapse), this is valid for the microscopic collisions we consider, where inelastic collapse is removed from the system using the TC method.

Chapter 5

The BBGKY hierarchy for granular media

The Liouville (or Kramers) equation derived in chapter 4 provides us with an equation for the evolution of a distribution function $f^N(\boldsymbol{r}^N, \boldsymbol{v}^N, t)$, which tells us the probability of finding N particles with positions $\boldsymbol{r}^N$ and velocities $\boldsymbol{v}^N$ at time t, given initial conditions $\boldsymbol{r}_0^N, \boldsymbol{v}_0^N$ at time $t = 0$ (neglecting the collision operator):

$$\left[\frac{\partial}{\partial t} + \boldsymbol{v}^N \cdot \boldsymbol{\nabla}_{\boldsymbol{r}^N} - \boldsymbol{\nabla}_{\boldsymbol{r}^N} V(\boldsymbol{r}^N, t) \cdot \boldsymbol{\nabla}_{\boldsymbol{v}^N}\right] f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \\ - \boldsymbol{\nabla}_{\boldsymbol{v}^N} \cdot \left[\gamma(\boldsymbol{v}^N + mk_BT\boldsymbol{\nabla}_{\boldsymbol{v}^N}) f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t)\right] = 0. \tag{5.1}$$

From this point onwards we will set $\boldsymbol{M} = I$ and $\boldsymbol{\Gamma}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) = \gamma > 0$ to be constant. When $\boldsymbol{\Gamma}$ depends on (binary) relative positions of particles, it is expected that a similar result will hold as in the colloidal case [69]. When $\boldsymbol{\Gamma}$ depends on the relative velocities of particles, some of the integral identities we use are no longer applicable, so further investigation will be required. For systems where $\boldsymbol{M} \neq I$, it may be more appropriate to consider a multi-species derivation of the BBGKY hierarchy and associated continuum models.

We introduce the *n-reduced-phase-space particle distribution function*, defined by

$$f^{(n)}(\boldsymbol{r}^n, \boldsymbol{v}^n, t) = \frac{N!}{(N-n)!} \int_{\mathbb{R}^{(N-n)d}} \int_{\mathbb{R}^{(N-n)d}} \mathrm{d}\boldsymbol{r}^{(N-n)}\, \mathrm{d}\boldsymbol{v}^{(N-n)} f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \chi_{\mathcal{D}_N(\boldsymbol{r}^n, t)}, \tag{5.2}$$

where $\boldsymbol{r}^{N-n} = (\boldsymbol{r}_{n+1}, \ldots, \boldsymbol{r}_N)$, $\boldsymbol{r}^n = (\boldsymbol{r}_1, \boldsymbol{r}_2, \ldots, \boldsymbol{r}_n)$, and similar for $\boldsymbol{v}^{N-n}, \boldsymbol{v}^n$, and for $n \leq N$,

$$\mathcal{D}_N(\boldsymbol{r}^n, t) = \left\{\boldsymbol{r}^{N-n} \in \mathbb{R}^{d(N-n)}, \boldsymbol{v}^N \in \mathbb{R}^{dN} : \forall i, j \in \{1, \ldots, N\}, \|\boldsymbol{r}_i - \boldsymbol{r}_j\| \geq \varepsilon\right\}, \tag{5.3}$$

and $\chi_{\mathcal{F}}$ is the characteristic function over the set $\mathcal{F}$:

$$\chi_{\mathcal{F}}(\boldsymbol{x}) = \begin{cases} 1 & \boldsymbol{x} \in \mathcal{F}, \\ 0 & \text{otherwise.} \end{cases} \tag{5.4}$$

The characteristic function $\chi_{\mathcal{D}_n(\boldsymbol{r}^k, t)}$ then ensures that we only consider data where particles do

not overlap. The data in the reduced-phase-space particle distribution functions must also ensure that particles do not overlap; the functions $f^{(n)}$ are defined on

$$\tilde{\mathcal{D}}_N^n = \{\boldsymbol{r}^n \in \mathbb{R}^{dn} : \forall i, j \in 1 \dots n, \|\boldsymbol{r}_i - \boldsymbol{r}_j\| \geq \varepsilon\}. \tag{5.5}$$

To ease notation in the derivation we will write $\boldsymbol{r} = \boldsymbol{r}_1$, $\boldsymbol{v} = \boldsymbol{v}_1$. These distribution functions represent the probability of finding n particles in a system of N particles with positions $\boldsymbol{r}^n$ and velocities $\boldsymbol{v}^n$. We can use this definition to construct a hierarchy of partial differential equations for all reduced-phase-space particle distribution functions, *i.e.* the BBGKY hierarchy.

The hierarchy of equations discussed here has a rich history in the literature, from its original consideration in the literature [131]–[134], to its application to modern theory and applications [45], [46], [135], [136], where it has been extended to include hard-sphere systems, and is commonly used as a starting point to derive generalised kinetic equations for various particle systems [78], [137], [138]. The hierarchy is linked to the Boltzmann equation: under molecular chaos assumptions and in the absence of any particle-particle interactions one can obtain the Boltzmann equation for hard spheres from the first equation in the BBGKY hierarchy [46]. The identifying trait of the hierarchy is the relationship between low and high-order distributions: the equation for the i^{th} reduced-phase-space distribution function depends on the $(i+1)^{\text{th}}$ distribution function. Although this means that finding solutions to the BBGKY hierarchy is as complex as finding solutions to the Liouville equation, by constructing an equation of motion for $f^{(1)}$, we can then truncate the hierarchy so that we have a closed equation for the one-body distribution, which removes the explicit dependence on N.

5.1 The BBGKY Hierarchy

We arrive at the n^{th} equation in the BBGKY hierarchy by integrating out the last $N-n$ positions and velocities in the Liouville equation. The first equation of the BBGKY hierarchy is obtained by integrating all but one position and velocity in the Liouville equation. We provide a term-by-term derivation of the first equation in the BBGKY hierarchy.

Firstly, the time derivative term is straightforward as the time derivative can be moved outside the integral:

$$\begin{aligned}
&\frac{N!}{(N-1)!}\int_{\mathbb{R}^{(N-1)d}}\int_{\mathbb{R}^{(N-1)d}} \mathrm{d}\boldsymbol{r}^{(N-1)}\,\mathrm{d}\boldsymbol{v}^{(N-1)}\,\frac{\partial}{\partial t}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)} \\
&= \frac{\partial}{\partial t}\frac{N!}{(N-1)!}\int_{\mathbb{R}^{(N-1)d}}\int_{\mathbb{R}^{(N-1)d}} \mathrm{d}\boldsymbol{r}^{(N-1)}\,\mathrm{d}\boldsymbol{v}^{(N-1)}\,f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}, \\
&= \frac{\partial}{\partial t}f^{(1)}(\boldsymbol{r}_1,\boldsymbol{v}_1,t).
\end{aligned} \tag{5.6}$$

For other terms integration by parts will be applied. For the next term in the Liouville equation we split the gradient term into two parts

$$\boldsymbol{v}^N \cdot \boldsymbol{\nabla}_{\boldsymbol{r}^N} f^{(N)} = \boldsymbol{v}^{N-1} \cdot \boldsymbol{\nabla}_{\boldsymbol{r}^{N-1}} f^{(N)} + \boldsymbol{v}_1 \cdot \boldsymbol{\nabla}_{\boldsymbol{r}_1} f^{(N)}. \tag{5.7}$$

The integral of the first term then vanishes using integration by parts on the position integral, and

the second term can be taken outside the integral to find

$$\frac{N!}{(N-1)!}\int_{\mathbb{R}^{(N-1)d}}\int_{\mathbb{R}^{(N-1)d}} \mathrm{d}\boldsymbol{r}^{(N-1)}\,\mathrm{d}\boldsymbol{v}^{(N-1)}\frac{1}{m}\boldsymbol{v}^N\cdot\boldsymbol{\nabla}_{\boldsymbol{r}^N} f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)} = \frac{1}{m}\boldsymbol{v}_1\cdot\boldsymbol{\nabla}_{\boldsymbol{r}_1} f^{(1)}. \tag{5.8}$$

For the term involving the potential $V(\boldsymbol{r}^N,t)$ we separate out the dot product again and use the expanded version of $V(\boldsymbol{r}^N,t)$ eq. (2.3), which gives us

$$\begin{aligned}
&\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}^N}V(\boldsymbol{r}^N)\cdot\nabla_{\boldsymbol{v}^N}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad=\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}^{N-1}}\left(V(\boldsymbol{r}^N)\right)\cdot\nabla_{\boldsymbol{v}^{N-1}}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad+\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}\left(\sum_{i=1}^N V^{ext}(\boldsymbol{r}_i)\right)\cdot\nabla_{\boldsymbol{v}_1}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad+\frac{1}{2}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}\left(\sum_{i,j=1}^N V_2(\boldsymbol{r}_i,\boldsymbol{r}_j)\right)\cdot\nabla_{\boldsymbol{v}_1}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad+\frac{1}{6}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}\left(\sum_{i,j,k=1}^N V_3(\boldsymbol{r}_i,\boldsymbol{r}_j,\boldsymbol{r}_k)\right)\cdot\nabla_{\boldsymbol{v}_1}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad+\ldots.
\end{aligned}\tag{5.9}$$

Using integration by parts on the velocity integral, the first term on the right-hand side is zero. For the second term of (5.9) all but one entry of the summand is independent of $\boldsymbol{r}_1$, so disappear, and we find

$$\begin{aligned}
&\frac{N!}{(N-1)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}\left(\sum_{i=1}^N V^{ext}(\boldsymbol{r}_i)\right)\cdot\nabla_{\boldsymbol{v}_1}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad=\frac{N!}{(N-1)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}V^{ext}(\boldsymbol{r}_1)\cdot\nabla_{\boldsymbol{v}_1}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad=\nabla_{\boldsymbol{r}_1}V^{ext}(\boldsymbol{r}_1)\cdot\nabla_{\boldsymbol{v}_1}\left(\frac{N!}{(N-1)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\right)\\
&\qquad=\nabla_{\boldsymbol{r}_1}V^{ext}(\boldsymbol{r}_1)\cdot\nabla_{\boldsymbol{v}_1}f^{(1)}.
\end{aligned}\tag{5.10}$$

For the third term of (5.9) only terms that have $\boldsymbol{r}_1$ in their argument survive the gradient, so

$$\nabla_{\boldsymbol{r}_1}\sum_{i,j=1}^N V_2(\boldsymbol{r}_i,\boldsymbol{r}_j) = \nabla_{\boldsymbol{r}_1}\sum_{i,j=1}^N\left[V_2(\boldsymbol{r}_1,\boldsymbol{r}_j)+V_2(\boldsymbol{r}_j,\boldsymbol{r}_1)\right]. \tag{5.11}$$

Then within the integral we can argue that the symmetry of $f^{(N)}$ in its arguments means that the integral

$$\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}V_2(\boldsymbol{r}_1,\boldsymbol{r}_j)\cdot\nabla_{\boldsymbol{v}_1}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)} \tag{5.12}$$

is equivalent for each $j = 2, \dots n$. So we find

$$\begin{aligned}
&\frac{N!}{2(N-1)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}\left(\sum_{i,j}^{N} V_2(\boldsymbol{r}_i,\boldsymbol{r}_j)\right)\cdot\nabla_{\boldsymbol{v}_1} f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
=&\frac{N!}{2(N-1)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}(N-1)(V_2(\boldsymbol{r}_1,\boldsymbol{r}_2)+V_2(\boldsymbol{r}_2,\boldsymbol{r}_1))\cdot\nabla_{\boldsymbol{v}_1} f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
=&\frac{N!(N-1)}{2(N-1)!}\int\int \mathrm{d}\boldsymbol{r}_2\,\mathrm{d}\boldsymbol{v}_2\nabla_{\boldsymbol{r}_1}[V_2(\boldsymbol{r}_1,\boldsymbol{r}_2)+V_2(\boldsymbol{r}_2,\boldsymbol{r}_1)]\\
&\qquad\cdot\nabla_{\boldsymbol{r}_1}\left(\int\int \mathrm{d}\boldsymbol{r}^{N-2}\,\mathrm{d}\boldsymbol{v}^{N-2} f^{(N)}\chi_{\mathcal{D}_N([\boldsymbol{r}_1,\boldsymbol{r}_2],t)}\right)\\
=&\int\int \mathrm{d}\boldsymbol{r}_2\,\mathrm{d}\boldsymbol{v}_2\nabla_{\boldsymbol{r}_1}\left[\frac{1}{2}(V_2(\boldsymbol{r}_1,\boldsymbol{r}_2)+V_2(\boldsymbol{r}_2,\boldsymbol{r}_1))\right]\cdot\nabla_{\boldsymbol{v}_1} f^{(2)}.
\end{aligned}\tag{5.13}$$

Similar methods are used on higher-order interparticle potentials, for example

$$\begin{aligned}
&\frac{1}{6}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}_1}\left(\sum_{i,j,k=1}^{N} V_3(\boldsymbol{r}_i,\boldsymbol{r}_j,\boldsymbol{r}_k)\right)\cdot\nabla_{\boldsymbol{v}_1} f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad\qquad =\\
&\int\int\int\int \mathrm{d}\boldsymbol{r}_2\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{r}_3\,\mathrm{d}\boldsymbol{v}_3\nabla_{\boldsymbol{r}_1}\left[\frac{1}{6}\sum_{i,j,k=1,2,3} V_3(\boldsymbol{r}_i,\boldsymbol{r}_j,\boldsymbol{r}_k)\right]\cdot\nabla_{\boldsymbol{v}_1} f^{(3)}.
\end{aligned}\tag{5.14}$$

We can then define:

$$v_2(\boldsymbol{r}_1,\boldsymbol{r}_2)=\frac{1}{2}\left(V_2(\boldsymbol{r}_1,\boldsymbol{r}_2)+V_2(\boldsymbol{r}_2,\boldsymbol{r}_1)\right),\quad v_3(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3)=\frac{1}{6}\sum_{i,j,k=1,2,3} V_3(\boldsymbol{r}_i,\boldsymbol{r}_j,\boldsymbol{r}_k),\dots.\tag{5.15}$$

The final term on the right-hand side of eq. (5.1) can be treated again by splitting the dot product into two parts and using integration by parts to find

$$\begin{aligned}
&\frac{N!}{(N-1)!}\int_{\mathbb{R}^{(N-n)d}}\int_{\mathbb{R}^{(N-n)d}} \mathrm{d}\boldsymbol{r}^{(N-n)}\,\mathrm{d}\boldsymbol{v}^{(N-n)}\nabla_{\boldsymbol{r}^N}\cdot\left[\gamma(\boldsymbol{v}^N+mk_BT\nabla_{\boldsymbol{v}^N})f^{(N)}\right]\chi_{\mathcal{D}_N(\boldsymbol{r}_1,t)}\\
&\qquad\qquad =\\
&\nabla_{\boldsymbol{v}_1}\cdot\left[\gamma(\boldsymbol{v}_1+mk_BT\nabla_{\boldsymbol{v}_1})f^{(1)}\right].
\end{aligned}\tag{5.16}$$

We note that this final result will have to be modified in the case where γ depends on the relative velocity of particles, which will lead to further dependence on higher-order particle distributions.

5.1.1 The collision operator

As discussed in the chapter 4, the Liouville equation for hard spheres can only be considered in the weak sense, and must include an additional term on the right-hand side of the equation. The derivation of all other terms in the first equation of the BBGKY is equivalent; only the collision term must be considered in the weak sense.

We infer the collision operator in the N particle system from the result derived from the two-

particle Liouville equation in theorem 14 (using test functions $\Phi \in C_c^1(\mathbb{R}^{2dN} \times (-\infty, \infty))$:

$$C_{\boldsymbol{r}^N}[f^{(N)}] = \int_{\partial \mathcal{P}_\varepsilon^N} \int_{\mathbb{R}^{dN}} \int_{-\infty}^{\infty} f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \Phi(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \boldsymbol{v}^N \cdot \hat{\boldsymbol{\nu}}(\boldsymbol{r}^N) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^N \, \mathrm{d}\mathcal{H}(\boldsymbol{r}^N), \tag{5.17}$$

where the surface $\partial \mathcal{P}_\varepsilon^N$ is the space where exactly one pair of particles is in contact:

$$\begin{aligned} \partial \mathcal{P}_\varepsilon^N = \{ \boldsymbol{r}^N \in \mathbb{R}^{dN} : \forall i, j \in \{1, \dots, N\}, \\ i \neq j, \|\boldsymbol{r}_i - \boldsymbol{r}_j\| \geq \varepsilon, \exists! i', j' \in \{1, \dots, N\} : \|\boldsymbol{r}_{i'} - \boldsymbol{r}_{j'}\| = \varepsilon \}, \end{aligned} \tag{5.18}$$

and we have assumed that all $\boldsymbol{v}^N \in \mathbb{R}^{dN}$ are admissible. We will assume that at any time $t \in \mathbb{R}$ at most two particles are in contact. By partitioning the space into sections where only one collision is occurring for a given period of time we should be able to relax this assumption. Then $\partial \mathcal{P}_\varepsilon^N$ can be partitioned into non-intersecting sets:

$$\begin{aligned} \partial \mathcal{P}_\varepsilon^N &= \bigcup_{k,l=1,l>k}^{N} \{ \boldsymbol{r}^N \in \mathbb{R}^{dN} : \forall i, j \in \{1, \dots, N\}, i \neq j, \|\boldsymbol{r}_i - \boldsymbol{r}_j\| > \varepsilon, \|\boldsymbol{r}_k - \boldsymbol{r}_l\| = \varepsilon \}, \\ &=: \bigcup_{k,l=1,l>k}^{N} \partial \mathcal{P}_{kl}^N. \end{aligned}$$

We note that in the union we have taken $l > k$ to avoid double counting sets. Thus we have that

$$C_{\boldsymbol{r}^N}[f^{(N)}] = \sum_{i,j=1,ij>i}^{N} \int_{\partial \mathcal{P}_{ij}^N} \int_{\mathbb{R}^{dN}} \int_{-\infty}^{\infty} f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \Phi(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \boldsymbol{v}^N \cdot \hat{\boldsymbol{\nu}}_{ij}(\boldsymbol{r}^N) \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^N \, \mathrm{d}\mathcal{H}(\boldsymbol{r}^N).$$

The outward unit normals $\hat{\boldsymbol{\nu}}_{ij}$ are then given by

$$\hat{\boldsymbol{\nu}}_{ij}(\boldsymbol{r}^N) = \begin{cases} 0, & k = \{1, \dots, N\} \backslash \{i, j\}, \\ \frac{\boldsymbol{r}_i - \boldsymbol{r}_j}{\sqrt{2}\|\boldsymbol{r}_i - \boldsymbol{r}_j\|}, & k = i, \\ \frac{\boldsymbol{r}_j - \boldsymbol{r}_i}{\sqrt{2}\|\boldsymbol{r}_i - \boldsymbol{r}_j\|}, & k = j. \end{cases} \tag{5.19}$$

We find that

$$\begin{aligned} C_{\boldsymbol{r}^N}[f^{(N)}] = \frac{1}{\sqrt{2}} \sum_{i,j=1,i>j}^{N} \int_{\partial \mathcal{P}_{ij}^N} \int_{\mathbb{R}^{dN}} \int_{-\infty}^{\infty} f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \Phi(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \\ \times (\boldsymbol{v}_i - \boldsymbol{v}_j) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^N \, \mathrm{d}\mathcal{H}(\boldsymbol{r}^N). \end{aligned} \tag{5.20}$$

For the first moment of the hierarchy we want to consider the result when all positions and velocities except for $\boldsymbol{r}_1$ and $\boldsymbol{v}_1$ are integrated out of the collision operator, and the test function only depends on $\boldsymbol{r}_1$ and $\boldsymbol{v}_1$. We note that $j \neq 1$, and so there are two cases to consider.

Firstly, when $i \neq 1$ contributions in the weak formulations are of the form

$$\frac{1}{\sqrt{2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{-\infty}^{\infty} \phi(\boldsymbol{r}_1, \boldsymbol{v}_1, t) \mathcal{B}_{ij}[f^{(N)}] \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}_1 \, \mathrm{d}\boldsymbol{r}_1$$

where

$$\mathcal{B}_{ij}[f^{(N)}] := \int_{\partial\mathcal{P}_{ij}^N(\boldsymbol{r}_1)} \int_{\mathbb{R}^{d(N-1)}} f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t)(\boldsymbol{v}_i - \boldsymbol{v}_j) \cdot \frac{(\boldsymbol{r}_i - \boldsymbol{r}_j)}{\|\boldsymbol{r}_i - \boldsymbol{r}_j\|} \,\mathrm{d}\boldsymbol{v}^{N-1} \,\mathrm{d}\mathcal{H}(\boldsymbol{r}^{N-1}) \tag{5.21}$$

and

$$\begin{aligned}\partial\mathcal{P}_{ij}^N(\boldsymbol{r}_1) :=& \{\boldsymbol{r}^{N-1} \in \mathbb{R}^{d(N-1)} : \forall i,j \in \{1,\dots,N\}, i \neq j, \|\boldsymbol{r}_i - \boldsymbol{r}_j\| > \varepsilon, \|\boldsymbol{r}_k - \boldsymbol{r}_l\| = \varepsilon\},\\ =& \bigcup_{\boldsymbol{r}_j = \boldsymbol{r}_i + \varepsilon\boldsymbol{n}, \boldsymbol{n} \in \mathbb{S}^{d-1}} \mathcal{D}_N(\boldsymbol{r}_i, \boldsymbol{r}_j).\end{aligned}$$

By using the symmetry in the arguments of $f^{(N)}$ we only need to consider $i = 2, j = 3$, the contribution from other values for i, j will be identical. We can then repeat the result from this case $(N^2 - 3N + 2)/2$ times. We can write

$$\begin{aligned}\mathcal{B}_{23}(f^{(N)}) = \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \chi_{\substack{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| > \varepsilon \\ \|\boldsymbol{r}_1 - \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}\| > \varepsilon}} \int_{\mathcal{D}_N(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n})} \\ \left[\int_{\mathbb{R}^{d(N-1)}} f^{(N)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}, \boldsymbol{r}^{N-3}], \boldsymbol{v}^N, t)(\boldsymbol{v}_2 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \,\mathrm{d}\boldsymbol{v}^{N-1}\right] \\ \mathrm{d}\boldsymbol{r}^{N-3} \,\mathrm{d}\boldsymbol{n} \,\mathrm{d}\boldsymbol{r}_2.\end{aligned}$$

Note that we have included a characteristic function to ensure that the remaining particles can not overlap; it is equivalent to considering the space $\tilde{\mathcal{D}}_N^2$. By separating the integral in the velocity difference $\boldsymbol{v}_2 - \boldsymbol{v}_3$, changing labelling $\boldsymbol{v}_2 \leftrightarrow \boldsymbol{v}_3$ in the second equation and using the symmetry of $f^{(N)}$, we have that (after including a prefactor of $N!/(N-1)!$)

$$\begin{aligned}\mathcal{B}_{23}(f^{(N)}) =& \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \chi_{\substack{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| > \varepsilon \\ \|\boldsymbol{r}_1 - \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}\| > \varepsilon}} \int_{\mathcal{D}_N(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n})} \int_{\mathbb{R}^{d(N-1)}} \boldsymbol{v}_2 \cdot \boldsymbol{n} \\ & \left(f^{(N)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}, \boldsymbol{r}^{N-3}], \boldsymbol{v}^N, t) - f^{(N)}([\boldsymbol{r}_1, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}, \boldsymbol{r}_2, \boldsymbol{r}^{N-3}], \boldsymbol{v}^N, t)\right) \\ & \mathrm{d}\boldsymbol{v}^{N-1} \,\mathrm{d}\boldsymbol{r}^{N-3} \,\mathrm{d}\boldsymbol{n} \,\mathrm{d}\boldsymbol{r}_2, \\ =& \frac{1}{(N-2)(N-1)} \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^{2d}} \chi_{\substack{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| > \varepsilon \\ \|\boldsymbol{r}_1 - \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}\| > \varepsilon}} \boldsymbol{v}_2 \cdot \boldsymbol{n} \\ & \left(f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}, \boldsymbol{r}^{N-3}], \boldsymbol{v}^3, t) - f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}, \boldsymbol{r}_2], \boldsymbol{v}^3, t)\right) \\ & \mathrm{d}\boldsymbol{v}_2 \,\mathrm{d}\boldsymbol{v}_3 \,\mathrm{d}\boldsymbol{n} \,\mathrm{d}\boldsymbol{r}_2.\end{aligned}$$

When considering all $(N^2 - 3N + 2)/2$ terms the prefactor $\frac{1}{(N-1)(N-2)}$ will cancel. The total contribution from terms where $i \neq 1$ is then given by

$$\begin{aligned}\frac{1}{2\sqrt{2}} \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \int_{-\infty}^{\infty} \phi(\boldsymbol{r}_1, \boldsymbol{v}_1, t) \Big[\int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^{2d}} \chi_{\substack{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| > \varepsilon \\ \|\boldsymbol{r}_1 - \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}\| > \varepsilon}} \boldsymbol{v}_2 \cdot \boldsymbol{n} \\ \left(f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}], \boldsymbol{v}^3, t) - f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}, \boldsymbol{r}_2], \boldsymbol{v}^3, t)\right) \\ \mathrm{d}\boldsymbol{v}_2 \,\mathrm{d}\boldsymbol{v}_3 \,\mathrm{d}\boldsymbol{n} \,\mathrm{d}\boldsymbol{r}_2 \Big] \,\mathrm{d}t \,\mathrm{d}\boldsymbol{v}_1 \,\mathrm{d}\boldsymbol{r}_1.\end{aligned}$$

These contributions from the collision operator are therefore small, and when $\varepsilon \to 0$ the contribution disappears. We therefore choose to neglect these terms from this point on in the derivation.

For $i=1$ there are $N-1$ choices of j, by using the symmetry of $f^{(2)}$ we have that

$$\begin{aligned}&\sum_{j=2}^{N}\int_{\partial\mathcal{P}_{1j}}\int_{\mathbb{R}^{dN}}\int_{-\infty}^{\infty}\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)f^{(N)}(\boldsymbol{r}^N,\boldsymbol{v}^N,t)(\boldsymbol{v}_1-\boldsymbol{v}_j)\cdot\frac{(\boldsymbol{r}_1-\boldsymbol{r}_j)}{\|\boldsymbol{r}_1-\boldsymbol{r}_j\|}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}^N\,\mathrm{d}\mathcal{H}(\boldsymbol{r}^N)\\&=\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{-\infty}^{\infty}(N-1)\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\mathcal{B}_{12}(f^{(N)})\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}_1\,\mathrm{d}\boldsymbol{r}_1\end{aligned}$$

where

$$\begin{aligned}\mathcal{B}_{12}(f^{(N)})=\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d}\left[\int_{\mathcal{D}_N(\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n})}\int_{\mathbb{R}^{d(N-2)}}f^{(N)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}^{N-2}],\boldsymbol{v}^N,t)\,\mathrm{d}\boldsymbol{v}^{N-2}\,\mathrm{d}\boldsymbol{r}^{N-2}\right]\\\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}.\end{aligned}$$

Thus when we consider the first equation of the BBGKY hierarchy this term contributes (premultiplying by $N!/(N-1)!$ and combining all $N-1$ terms):

$$\begin{aligned}&\frac{N!}{\sqrt{2}(N-1)!}\int_{\mathbb{R}^{2d}}\int_{-\infty}^{\infty}(N-1)\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\mathcal{B}_{12}(f^{(N)})\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}^2\\&=\int_{\mathbb{R}^{2d}}\int_{-\infty}^{\infty}\phi(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\left[\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}\right]\mathrm{d}t\,\mathrm{d}\boldsymbol{v}_1\,\mathrm{d}\boldsymbol{r}_1.\end{aligned}$$

Note that in this case we do not need to include a characteristic function as after the bracketed integrals only one particle remains in the system. By separating into pre and post-collisional initial data as in section 4.4.4, we arrive at the inelastic Boltzmann collision operator eq. (4.79) in the first equation of the BBGKY hierarchy in the weak form. We use this result to argue eq. (4.79) is a suitable collision operator to include in the N-particle system of equations.

In all, we conclude that the first equation of the BBGKY hierarchy has the form (where we choose to include only two and three-body interparticle potentials to reduce notation)

$$\begin{aligned}&\frac{\partial}{\partial t}f^{(1)}+\frac{1}{m}\boldsymbol{v}_1\cdot\nabla_{\boldsymbol{r}_1}f^{(1)}-\nabla_{\boldsymbol{r}_1}V^{ext}(\boldsymbol{r}_1)\cdot\nabla_{\boldsymbol{v}_1}f^{(1)}\\&\quad-\int\int\mathrm{d}\boldsymbol{r}_2\,\mathrm{d}\boldsymbol{v}_2\nabla_{\boldsymbol{r}_1}v_2(\boldsymbol{r}_1,\boldsymbol{r}_2)\cdot\nabla_{\boldsymbol{v}_1}f^{(2)}\\&\quad-\int\int\int\int\mathrm{d}\boldsymbol{r}_2\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{r}_3\,\mathrm{d}\boldsymbol{v}_3\nabla_{\boldsymbol{r}_1}v_3(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3)\cdot\nabla_{\boldsymbol{v}_1}f^{(3)}\\&=\nabla_{\boldsymbol{v}_1}\cdot\left[\gamma(\boldsymbol{v}_1+mk_BT\nabla_{\boldsymbol{v}_1})f^{(1)}\right]+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^2}\int_{\mathbb{R}^3}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}\\&\quad\left[\frac{1}{\alpha^2}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\mathrm{out}},\boldsymbol{v}_2^{\mathrm{out}}],t)-f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2],t)\right]\\&\qquad\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n},\end{aligned}\tag{5.22}$$

where we have included the characteristic function in the collision operator as an alternative to the pre-collisional argument in the velocity integral.

5.1.2 The second equation in the BBGKY hierarchy

We apply similar methods to derive the second equation in the BBGKY hierarchy as in the first. For the time derivative term the result is straightforward:

$$\begin{aligned}
&\frac{N!}{(N-2)!}\int_{\mathbb{R}^{(N-2)d}}\int_{\mathbb{R}^{(N-2)d}} \mathrm{d}\boldsymbol{r}^{(N-2)}\,\mathrm{d}\boldsymbol{v}^{(N-2)}\frac{\partial}{\partial t}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&=\frac{\partial}{\partial t}\frac{N!}{(N-2)!}\int_{\mathbb{R}^{(N-2)d}}\int_{\mathbb{R}^{(N-2)d}} \mathrm{d}\boldsymbol{r}^{(N-2)}\,\mathrm{d}\boldsymbol{v}^{(N-2)}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)},\\
&=\frac{\partial}{\partial t}f^{(2)}(\boldsymbol{r}^2,\boldsymbol{v}^2,t).
\end{aligned} \tag{5.23}$$

And in a similar fashion to the derivation of the first BBGKY equation, the next term in the Liouville equation can be split into two parts

$$\boldsymbol{v}^N\cdot\boldsymbol{\nabla}_{\boldsymbol{r}^N}f^{(N)}=\boldsymbol{v}^{N-2}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}^{N-2}}f^{(N)}+\boldsymbol{v}^2\cdot\boldsymbol{\nabla}_{\boldsymbol{r}^2}f^{(N)}. \tag{5.24}$$

The integral of the first term then vanishes using integration by parts, so that we are left with

$$\frac{N!}{(N-2)!}\int_{\mathbb{R}^{(N-2)d}}\int_{\mathbb{R}^{(N-2)d}}\frac{1}{m}\boldsymbol{v}^N\cdot\boldsymbol{\nabla}_{\boldsymbol{r^N}}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\,\mathrm{d}\boldsymbol{r}^{(N-2)}\,\mathrm{d}\boldsymbol{v}^{(N-2)}=\frac{1}{m}\boldsymbol{v}^2\cdot\boldsymbol{\nabla}_{\boldsymbol{r^2}}f^{(2)}. \tag{5.25}$$

More care has to be taken when considering the potential term;

$$\begin{aligned}
\int\int \mathrm{d}\boldsymbol{r}^{N-2}\,\mathrm{d}\boldsymbol{v}^{N-2}&\nabla_{\boldsymbol{r}^N}V(\boldsymbol{r}^N)\cdot\nabla_{\boldsymbol{v}^N}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&=\int\int \mathrm{d}\boldsymbol{r}^{N-2}\,\mathrm{d}\boldsymbol{v}^{N-2}\nabla_{\boldsymbol{r}^{N-2}}\left(V(\boldsymbol{r}^N)\right)\cdot\nabla_{\boldsymbol{v}^{N-2}}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&+\int\int \mathrm{d}\boldsymbol{r}^{N-2}\,\mathrm{d}\boldsymbol{v}^{N-2}\nabla_{\boldsymbol{r}^2}\left(\sum_{i=1}^N V^{ext}(\boldsymbol{r}_i)\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&+\frac{1}{2}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}^2}\left(\sum_{i,j=1}^N V_2(\boldsymbol{r}_i,\boldsymbol{r}_j)\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&+\frac{1}{6}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}^2}\left(\sum_{i,j,k=1}^N V_3(\boldsymbol{r}_i,\boldsymbol{r}_j,\boldsymbol{r}_k)\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&+\ldots.
\end{aligned} \tag{5.26}$$

The first term on the right-hand side of the equality is zero by the standard integration by parts argument used previously. For the second term, the only potentials terms that remain due to the spatial gradient are

$$\nabla_{\boldsymbol{r}^2}V^{\mathrm{ext}}(\boldsymbol{r}_i)=\nabla_{\boldsymbol{r}^2}(V^{\mathrm{ext}}(\boldsymbol{r}_1)+V^{\mathrm{ext}}(\boldsymbol{r}_2)). \tag{5.27}$$

Thus the second term contributes

$$\begin{aligned}
&\int\int \mathrm{d}\boldsymbol{r}^{N-2}\,\mathrm{d}\boldsymbol{v}^{N-2}\nabla_{\boldsymbol{r}^2}\left(\sum_{i=1}^N V^{ext}(\boldsymbol{r}_i)\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&=
\end{aligned} \tag{5.28}$$

$$\nabla_{\boldsymbol{r}^2}\left(V^{ext}(\boldsymbol{r}_1)+V^{ext}(\boldsymbol{r}_2)\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(2)}.$$

For the next term, we have that

$$\nabla_{\boldsymbol{r}^2}\sum_{i,j=1}^{N}V_2(\boldsymbol{r}_i,\boldsymbol{r}_j)=\nabla_{\boldsymbol{r}^2}\left([V_2(\boldsymbol{r}_1,\boldsymbol{r}_2)+V_2(\boldsymbol{r}_2,\boldsymbol{r}_1)]+\sum_{i=1}^{2}\sum_{j=3}^{N}(V_2(\boldsymbol{r}_i,\boldsymbol{r}_j)+V_2(\boldsymbol{r}_j,\boldsymbol{r}_i))\right). \quad (5.29)$$

By using the symmetry of the distribution function we then have that

$$\begin{aligned}
&\frac{N!}{2(N-2)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}^2}\left(\sum_{i,j=1}^{N}V_2(\boldsymbol{r}_i,\boldsymbol{r}_j)\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
=&\frac{N!}{2(N-2)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}^2}[V_2(\boldsymbol{r}_1,\boldsymbol{r}_2)+V_2(\boldsymbol{r}_2,\boldsymbol{r}_1)]\cdot\nabla_{\boldsymbol{v}^2}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
&+\frac{N!(N-2)}{2(N-2)!}\int\int \mathrm{d}\boldsymbol{r}^{N-1}\,\mathrm{d}\boldsymbol{v}^{N-1}\nabla_{\boldsymbol{r}^2}\left(\sum_{i=1,2}(V_2(\boldsymbol{r}_i,\boldsymbol{r}_3)+V_2(\boldsymbol{r}_3,\boldsymbol{r}_i))\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(N)}\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\\
=&\nabla_{\boldsymbol{r}^2}\left[\frac{1}{2}(V_2(\boldsymbol{r}_1,\boldsymbol{r}_2)+V_2(\boldsymbol{r}_2,\boldsymbol{r}_1))\right]\cdot\nabla_{\boldsymbol{v}^2}f^{(2)}\\
&\qquad+\nabla_{\boldsymbol{r}^2}\left(\frac{1}{2}\sum_{i=1,2}(V_2(\boldsymbol{r}_i,\boldsymbol{r}_3)+V_2(\boldsymbol{r}_3,\boldsymbol{r}_i))\right)\cdot\nabla_{\boldsymbol{v}^2}f^{(3)}.
\end{aligned}$$

Similar derivations can be considered for higher-order interaction potentials. For the final term on the left-hand side of the Liouville equation the derivation is analogous to previous methods, to get

$$\begin{aligned}
&\frac{N!}{(N-2)!}\int\int\nabla_{\boldsymbol{v}^N}[\gamma(\boldsymbol{v}^N+mk_BT\nabla_{\boldsymbol{v}^N})f^{(N)}]\chi_{\mathcal{D}_N(\boldsymbol{r}^2,t)}\,\mathrm{d}\boldsymbol{v}^{N-2}\,\mathrm{d}\boldsymbol{r}^{N-2}\\
&\qquad=\nabla_{\boldsymbol{v}^2}[\gamma(\boldsymbol{v}^2+mk_BT\nabla_{\boldsymbol{v}^2})f^{(2)}]. \quad (5.30)
\end{aligned}$$

Finally we consider the collision operator starting from eq. (5.21):

$$\begin{aligned}
C_{\boldsymbol{r}^N}[f^{(N)}]=\frac{1}{\sqrt{2}}\sum_{i,j=1,i>j}^{N}\int_{\partial\mathcal{P}_{ij}^N}\int_{\mathbb{R}^{dN}}\int_{-\infty}^{\infty}f^{(N)}(\boldsymbol{r}^N,\boldsymbol{v}^N,t)\Phi(\boldsymbol{r}^N,\boldsymbol{v}^N,t) \quad (5.31)\\
\times(\boldsymbol{v}_i-\boldsymbol{v}_j)\cdot\frac{(\boldsymbol{r}_i-\boldsymbol{r}_j)}{\|\boldsymbol{r}_i-\boldsymbol{r}_j\|}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}^N\,\mathrm{d}\mathcal{H}(\boldsymbol{r}^N).
\end{aligned}$$

For the second equation in the hierarchy, we want to integrate out all positions and velocities except for $\boldsymbol{r}_1,\boldsymbol{r}_2$ and $\boldsymbol{v}_1,\boldsymbol{v}_2$, and the test function depends only on these quantities. There are four cases to consider.

Firstly, the case $i=1$ and $j=2$ contributes one term to the summation in eq. (5.21):

$$\begin{aligned}
&\frac{1}{\sqrt{2}}\int_{\partial\mathcal{P}_{12}^N}\int_{\mathbb{R}^{dN}}\int_{-\infty}^{\infty}f^{(N)}(\boldsymbol{r}^N,\boldsymbol{v}^N,t)\phi(\boldsymbol{r}^2,\boldsymbol{v}^2,t)(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\frac{(\boldsymbol{r}_1-\boldsymbol{r}_2)}{\|\boldsymbol{r}_1-\boldsymbol{r}_2\|}\,\mathrm{d}t\,\mathrm{d}\boldsymbol{v}^N\,\mathrm{d}\mathcal{H}(\boldsymbol{r}^N)\\
&=\frac{1}{\sqrt{2}}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\int_{-\infty}^{\infty}\phi([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\\
&\qquad\times\left[\int_{\mathcal{D}_N(\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n})}\int_{\mathbb{R}^{d(N-2)}}f^{(N)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}^{N-2}],\boldsymbol{v}^N,t)\,\mathrm{d}\boldsymbol{v}^{N-2}\,\mathrm{d}\boldsymbol{r}^{N-2}\right]\mathrm{d}t\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\,\mathrm{d}\boldsymbol{r}_1,
\end{aligned}$$

$$= \frac{1}{\sqrt{2}} \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^{2d}} \int_{-\infty}^{\infty} \phi([\boldsymbol{r}_1, \boldsymbol{r}_1 + \varepsilon\boldsymbol{n}], \boldsymbol{v}^2, t) f^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1 + \varepsilon\boldsymbol{n}], \boldsymbol{v}^2, t) \\ \times (\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot \boldsymbol{n} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^2 \, \mathrm{d}\boldsymbol{n} \, \mathrm{d}\boldsymbol{r}_1.$$

This term only contributes to the system at the interaction diameter ε. Next, we consider the case where $i = 1$ and $j \geq 3$, each of which contributes $(N-2)$ terms to $C_{\boldsymbol{r}^N}[f^{(N)}]$. Using standard symmetry arguments in $f^{(N)}$ we only need to consider the case $j = 3$. Then

$$\frac{N!}{\sqrt{2}(N-2)!} \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \int_{-\infty}^{\infty} \phi(\boldsymbol{r}^2, \boldsymbol{v}^2, t) \\ \times \left[\int_{\mathbb{S}^{d-1}} \chi_{\substack{\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| > \varepsilon \\ \|\boldsymbol{r}_1 - (\boldsymbol{r}_2 + \varepsilon\boldsymbol{n})\|}} \int_{\mathcal{D}_N(\boldsymbol{r}^2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n})} \int_{\mathbb{R}^{d(N-2)}} f^{(N)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1 + \varepsilon\boldsymbol{n}, \boldsymbol{r}^{N-3}], \boldsymbol{v}^N, t) \right. \\ \left. \times (\boldsymbol{v}_1 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \, \mathrm{d}\boldsymbol{v}^{N-2} \, \mathrm{d}\boldsymbol{r}^{N-2} \, \mathrm{d}\boldsymbol{n} \right] \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^2 \, \mathrm{d}\boldsymbol{r}^2, \\ = \frac{1}{\sqrt{2}} \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \int_{-\infty}^{\infty} \frac{\phi(\boldsymbol{r}^2, \boldsymbol{v}^2, t)}{(N-2)} \\ \times \left[\int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^d} f^{(3)}([\boldsymbol{r}^2, \boldsymbol{r}_1 + \varepsilon\boldsymbol{n}], \boldsymbol{v}^3, t)(\boldsymbol{v}_1 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \, \mathrm{d}\boldsymbol{v}_3 \, \mathrm{d}\boldsymbol{n} \right] \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^2 \, \mathrm{d}\boldsymbol{r}^2.$$

Then using the results of section 4.4.4, we see that the contribution from these terms to the second BBGKY takes the form of the collision operator on $f^{(3)}$ between particles 1 and 3:

$$\frac{1}{\sqrt{2}} \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^d} \chi_{(\boldsymbol{v}_1 - \boldsymbol{v}_3) \cdot \boldsymbol{n} > 0} (\boldsymbol{v}_1 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \Big[f^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1 + \varepsilon\boldsymbol{n}, \boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{v}_3, t) \\ - \frac{1}{\alpha^2} f^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1 + \varepsilon\boldsymbol{n}, \boldsymbol{v}_1', \boldsymbol{v}_2, \boldsymbol{v}_3', t) \Big] \mathrm{d}\boldsymbol{v}_3 \, \mathrm{d}\boldsymbol{n}.$$

The result is analogous for the next case; $i = 2$ and $j = 3$. The final case to consider is then $i > 3$ and $j > i$. As usual we appeal to symmetry so that we only have to consider $i = 3$ and $j = 4$, and note that there are $(N-2)(N-3)/2$ terms of this form. We can use an identical argument to the equivalent case in the first BBGKY equation (when $i > 2$ and $j > i$), to find

$$\frac{N!}{(N-2)!} \mathcal{B}_{34}[f^{(N)}] \\ = \int_{\partial \mathcal{P}_{34}^{(N)}} \int_{\mathbb{R}^{dN}} \int_{-\infty}^{\infty} f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \phi(\boldsymbol{r}^2, \boldsymbol{v}^2, t)(\boldsymbol{v}_3 - \boldsymbol{v}_4) \cdot \frac{(\boldsymbol{r}_3 - \boldsymbol{r}_4)}{\|\boldsymbol{r}_3 - \boldsymbol{r}_4\|} \, \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^N \, \mathrm{d}\mathcal{H}(\boldsymbol{r}^N), \\ = \frac{1}{2\sqrt{2}} \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^{2d}} \int_{-\infty}^{\infty} \Big\{ \int_{\mathbb{R}^{2d}} \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \boldsymbol{v}_3 \cdot \boldsymbol{n} \big[f^{(4)}([\boldsymbol{r}^2, \boldsymbol{r}_3, \boldsymbol{r}_3 + \varepsilon\boldsymbol{n}], \boldsymbol{v}^4, t) \\ - f^{(4)}([\boldsymbol{r}^2, \boldsymbol{r}_3 + \varepsilon\boldsymbol{n}, \boldsymbol{r}_3], \boldsymbol{v}^4, t) \big] \mathrm{d}\boldsymbol{n} \, \mathrm{d}\boldsymbol{r}_3 \, \mathrm{d}\boldsymbol{v}_3 \, \mathrm{d}\boldsymbol{v}_4 \Big\} \mathrm{d}t \, \mathrm{d}\boldsymbol{v}^2 \, \mathrm{d}\boldsymbol{r}^2.$$

The term in the square brackets is again small and vanishes when $\varepsilon \to 0$, so we choose to neglect this term.

This concludes the derivation of the second equation in the BBGKY hierarchy; by combining all terms considered we arrive the second equation in the BBGKY hierarchy (where for ease of notation we have restricted to two and three-body interparticle interaction terms):

$$\frac{\partial}{\partial t} f^{(2)} + \frac{1}{m} \boldsymbol{v}^2 \cdot \boldsymbol{\nabla}_{\boldsymbol{r}^2} f^{(2)} - \nabla_{\boldsymbol{r}^2} [V^{ext}(\boldsymbol{r}_1) + V^{ext}(\boldsymbol{r}_2)] \cdot \nabla_{\boldsymbol{v}^2} f^{(2)}$$

$$
\begin{aligned}
&- \nabla_{\boldsymbol{r}^2} v_2(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot \nabla_{\boldsymbol{v}^2} f^{(2)} \\
&- \int\int \mathrm{d}\boldsymbol{r}_3 \, \mathrm{d}\boldsymbol{v}_3 \nabla_{\boldsymbol{r}^2} \left[v_2(\boldsymbol{r}_2, \boldsymbol{r}_3) + v_2(\boldsymbol{r}_1, \boldsymbol{r}_3) \right] \cdot \nabla_{\boldsymbol{v}^2} f^{(3)} \\
&- \int\int \mathrm{d}\boldsymbol{r}_3 \, \mathrm{d}\boldsymbol{v}_3 \nabla_{\boldsymbol{r}^2} v_3(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) \cdot \nabla_{\boldsymbol{v}^2} f^{(3)} \\
=&\nabla_{\boldsymbol{v}^2} \cdot \left[\gamma(\boldsymbol{v}^2 + m k_B T \nabla_{\boldsymbol{v}^2}) f^{(2)} \right] + \frac{1}{\sqrt{2}} \chi_{\boldsymbol{r}_2 = \boldsymbol{r} + 1 + \varepsilon \boldsymbol{n}} (\boldsymbol{r}_1 - \boldsymbol{r}_2) \cdot \boldsymbol{n} f^{(2)} \\
&+ \frac{1}{\sqrt{2}} \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} \chi_{(\boldsymbol{v}_1 - \boldsymbol{v}_3) \cdot \boldsymbol{n} > 0} \left[\frac{1}{\alpha^2} f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1 + \varepsilon \boldsymbol{n}], [\boldsymbol{v}_1^{\mathrm{out}}, \boldsymbol{v}_2, \boldsymbol{v}_3^{\mathrm{out}}], t) \right. \\
&\qquad \left. - f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1 + \varepsilon \boldsymbol{n}], \boldsymbol{v}^3, t) \right] (\boldsymbol{v}_1 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \, \mathrm{d}\boldsymbol{v}_3 \, \mathrm{d}\boldsymbol{n} \\
&+ \frac{1}{\sqrt{2}} \int_{\mathbb{S}^2} \int_{\mathbb{R}^3} \chi_{(\boldsymbol{v}_2 - \boldsymbol{v}_3) \cdot \boldsymbol{n} > 0} \left[\frac{1}{\alpha^2} f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon \boldsymbol{n}], [\boldsymbol{v}_1, \boldsymbol{v}_2^{\mathrm{out}}, \boldsymbol{v}_3^{\mathrm{out}}], t) \right. \\
&\qquad \left. - f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon \boldsymbol{n}], \boldsymbol{v}^3, t) \right] (\boldsymbol{v}_2 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \, \mathrm{d}\boldsymbol{v}_3 \, \mathrm{d}\boldsymbol{n}. \qquad (5.32)
\end{aligned}
$$

Note that the collision operator still affects the second equation in the BBGKY hierarchy, and will also appear in higher-order equations.

5.2 Uses of the BBGKY hierarchy

We have now constructed an equation which shows how the one-body reduced-phase-space probability distribution evolves over time. However, the one-body distribution function also relies on higher-order distribution functions which are also determined by other PDEs in the BBGKY hierarchy. Furthermore, even if terms involving higher-order distributions are neglected and we only consider the first equation in the BBGKY hierarchy, we are still left with a $2d$-dimensional PDE to solve, which for $d > 2$ will become computationally intractable using standard methods. We therefore cannot use the hierarchy directly for simulation, we must make some approximations to *truncate* the hierarchy.

5.3 Moments of the hierarchy

It is common to consider moments of the BBGKY hierarchy to construct closed continuum models for particle systems. By taking moments in velocity we lose precise velocity characteristics of particles in the system, but are able to half the dimension of the system of PDEs. In many cases the quantities of interest for a particle system are provided by these moments, for example local densities.

A general k-particle n^{th} order velocity moment is given by

$$
\langle \mathcal{Q}^{n_1, \ldots, n_k} f^{(k)} \rangle (\boldsymbol{r}^k, t) := \int \mathrm{d}\boldsymbol{v}_1 \ldots \mathrm{d}\boldsymbol{v}_k \boldsymbol{v}_1^{\otimes n_1} \otimes \cdots \otimes \boldsymbol{v}_k^{\otimes n_k} f^{(k)}(\boldsymbol{r}^k, \boldsymbol{v}^k, t), \qquad (5.33)
$$

where $\otimes$ is the pre-defined tensor product and we define as the tensor product of a vector with

itself n times:

$$\boldsymbol{x}^{\otimes n} = \boldsymbol{x} \otimes \boldsymbol{x} \otimes \cdots \otimes \boldsymbol{x}, \tag{5.34}$$

and $\sum_{i=1}^{k} n_k = n$. We consider moments in $\boldsymbol{v}_1$ but moments in $\boldsymbol{v}_2$ follow symmetrically. For simplicity, we will not include any interaction terms that are more than pairwise, but their inclusion is straightforward.

We use the multi-dimensional product rule and integration by parts to determine that

$$\int_{\mathbb{R}^d} \boldsymbol{v}_1^{\otimes n} (\boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{r}_1} f^{(1)}) \,\mathrm{d}\boldsymbol{v}_1 = \int_{\mathbb{R}^d} \nabla_{\boldsymbol{r}_1} \cdot (\boldsymbol{v}^{\otimes(n+1)} f^{(1)}) \,\mathrm{d}\boldsymbol{v}_1, \tag{5.35}$$

$$\int_{\mathbb{R}^d} \boldsymbol{v}_1^{\otimes n} \nabla_{\boldsymbol{r}_1} V^{\mathrm{ext}} \cdot \nabla_{\boldsymbol{v}_1} f^{(1)} \,\mathrm{d}\boldsymbol{v}_1 = -n \int_{\mathbb{R}^d} f^{(1)} \boldsymbol{v}_1^{\otimes(n-1)} \otimes \nabla_{\boldsymbol{r}_1} V^{\mathrm{ext}} \,\mathrm{d}\boldsymbol{v}_1, \tag{5.36}$$

$$\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d} \boldsymbol{v}_1^{\otimes n} \nabla_{\boldsymbol{r}_1} v_2 \cdot \nabla_{\boldsymbol{v}_1} f^{(1)} \,\mathrm{d}\boldsymbol{r}_2 \,\mathrm{d}\boldsymbol{v}_2 \,\mathrm{d}\boldsymbol{v}_1 = -n \int_{\mathbb{R}^d}\int_{\mathbb{R}^{2d}} f^{(1)} \boldsymbol{v}_1^{\otimes(n-1)} \otimes \nabla_{\boldsymbol{r}_1} v_2 \,\mathrm{d}\boldsymbol{v}^2 \,\mathrm{d}\boldsymbol{r}_2. \tag{5.37}$$

The moments for the first BBGKY equation are then as follows:

$$\begin{aligned}
\frac{\partial}{\partial t}\langle \mathcal{Q}^n f^{(1)} \rangle = &-\frac{1}{m} \nabla_{\boldsymbol{r}_1} \cdot \langle \mathcal{Q}^{n+1} f^{(1)} \rangle - n(\nabla_{\boldsymbol{r}_1} V^{\mathrm{ext}}) \langle \mathcal{Q}^{n-1} f^{(1)} \rangle \\
&- n \int \mathrm{d}\boldsymbol{r}_2 (\nabla_{\boldsymbol{r}_1} v_2(\boldsymbol{r}_1, \boldsymbol{r}_2)) \langle \mathcal{Q}^{n-1,0} f^{(2)} \rangle + \langle \mathcal{Q}^n \mathcal{L}_{\mathrm{diss}}^{(1)} f^{(1)} \rangle \\
&- \frac{1}{\sqrt{2}} \int_{\mathbb{S}^{d-1}} \left\{ \int_{\mathbb{R}^{2d}} \boldsymbol{v}_1^{\otimes n} (\boldsymbol{v}_1 - \boldsymbol{v}_2) \cdot \boldsymbol{n} \chi_{(\boldsymbol{v}_1 - \boldsymbol{v}_2)\cdot \boldsymbol{n} > 0} \tilde{f}_{\alpha,\varepsilon}^{(2)}(\boldsymbol{r}_1, \boldsymbol{n}, \boldsymbol{v}^2) \,\mathrm{d}\boldsymbol{v}_1 \,\mathrm{d}\boldsymbol{v}_2 \right\} \mathrm{d}\boldsymbol{n},
\end{aligned} \tag{5.38}$$

where

$$\tilde{f}_{\alpha,\varepsilon}^{(2)}(\boldsymbol{r}_1, \boldsymbol{n}, \boldsymbol{v}^2) = \left[f^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1 + \varepsilon \boldsymbol{n}], [\boldsymbol{v}_1, \boldsymbol{v}_2], t) - \frac{1}{\alpha^2} f^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1 + \varepsilon \boldsymbol{n}], [\boldsymbol{v}_1^{\mathrm{out}}, \boldsymbol{v}_2^{\mathrm{out}}], t) \right] \tag{5.39}$$

and

$$\langle \mathcal{Q}^n \mathcal{L}_{\mathrm{diss}}^{(1)} f^{(1)} \rangle = -n\gamma \langle \mathcal{Q}^n f^{(1)} \rangle + n(n-1)\gamma m k_B T_{\mathrm{res}} \langle \mathcal{Q}^{n-2} f^{(1)} \rangle \mathbf{1}, \tag{5.40}$$

where $\mathbf{1}$ is the $d \times d$ identity matrix and we define the dot product in the collision operator as the dot product in the final dimension of the tensor product.

The $\boldsymbol{v}_1$-moments of the second equation of the BBGKY hierarchy are then given by

$$\begin{aligned}
\frac{\partial}{\partial t}\langle \mathcal{Q}^{n,0} f^{(2)} \rangle = &-\frac{1}{m} \nabla_{\boldsymbol{r}_1} \cdot \langle \mathcal{Q}^{n+1,0} f^{(2)} \rangle - \frac{1}{m} \nabla_{\boldsymbol{r}_2} \cdot \langle \mathcal{Q}^{n,1} f^{(2)} \rangle \\
&- n(\nabla_{\boldsymbol{r}_1} V^{\mathrm{ext}} + \nabla_{\boldsymbol{r}_1} v_2(\boldsymbol{r}_1, \boldsymbol{r}_2)) \langle \mathcal{Q}^{n-1,0} f^{(2)} \rangle \\
&- n \int \mathrm{d}\boldsymbol{r}_3 (\nabla_1 v_2(\boldsymbol{r}_1, \boldsymbol{r}_3)) \langle \mathcal{Q}^{n-1,0,0} f^{(3)} \rangle + \langle \mathcal{Q}^{n,0} \mathcal{L}_{\mathrm{diss}}^{(2)} f^{(2)} \rangle \\
&- \frac{1}{\sqrt{2}} \chi_{\substack{\boldsymbol{r}_2 = \boldsymbol{r}_1 + \varepsilon \boldsymbol{n} \\ \boldsymbol{n} \in \mathbb{S}^{d-1}}} \left(\langle \mathcal{Q}^{n+1,0} f^{(2)} \rangle + \langle \mathcal{Q}^{n,1} f^{(2)} \rangle \right) \cdot \boldsymbol{n} \\
&- \frac{1}{\sqrt{2}} \int_{\mathbb{S}^{d-1}} \left\{ \int_{\mathbb{R}^{3d}} \boldsymbol{v}_1^{\otimes n} (\boldsymbol{v}_1 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \chi_{(\boldsymbol{v}_1 - \boldsymbol{v}_3)\cdot \boldsymbol{n} > 0} \tilde{f}_{\alpha,\varepsilon}^{(3)}(\boldsymbol{r}^2, \boldsymbol{n}, \boldsymbol{v}^3) \,\mathrm{d}\boldsymbol{v}^3 \right\} \mathrm{d}\boldsymbol{n} \\
&- \frac{1}{\sqrt{2}} \int_{\mathbb{S}^{d-1}} \left\{ \int_{\mathbb{R}^{3d}} \boldsymbol{v}_1^{\otimes n} (\boldsymbol{v}_2 - \boldsymbol{v}_3) \cdot \boldsymbol{n} \chi_{(\boldsymbol{v}_2 - \boldsymbol{v}_3)\cdot \boldsymbol{n} > 0} \tilde{f}_{\alpha,\varepsilon}^{(3)}(\boldsymbol{r}^2, \boldsymbol{n}, \boldsymbol{v}^3) \,\mathrm{d}\boldsymbol{v}^3 \right\} \mathrm{d}\boldsymbol{n},
\end{aligned} \tag{5.41}$$

where

$$\tilde{f}^{(3)}_{\alpha,\varepsilon}(\boldsymbol{r}^2,\boldsymbol{n},\boldsymbol{v}^3) = \left[f^{(3)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3],t) - \frac{1}{\alpha^2}f^{(3)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\text{out}},\boldsymbol{v}_2,\boldsymbol{v}_3^{\text{out}}],t)\right],$$

$$\bar{f}^{(3)}_{\alpha,\varepsilon}(\boldsymbol{r}^2,\boldsymbol{n},\boldsymbol{v}^3) = \left[f^{(3)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_2+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3],t) - \frac{1}{\alpha^2}f^{(3)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_2+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2^{\text{out}},\boldsymbol{v}_3^{\text{out}}],t)\right]$$

and

$$\langle\mathcal{Q}^{n,0}\mathcal{L}^{(2)}_{\text{diss}}\rangle = -n\gamma\langle\mathcal{Q}^{n,0}f^{(2)}\rangle + n(n-1)\gamma m k_B T_{\text{res}}\langle\mathcal{Q}^{n-2,0}f^{(2)}\rangle\mathbf{1}. \tag{5.42}$$

In a similar approach we can then determine moments for any equation in the BBGKY hierarchy.

5.4 Discussion

In this chapter we have constructed the first and second equations in the BBGKY hierarchy for systems of hard, spherical particles that collide inelastically. The resulting equations are well-known in the literature, for point-like particles [139] and in the hard-sphere case [46], as collision operators are generally included at this point in the derivation. However, in our case we have shown that the collision operators also occur as a result of the rigorously-derived Liouville equation in chapter 4. This involved additional integral calculations, but as we knew that the collisional effects were contained in the Liouville equation, little additional theory was required. The resulting equations have an explicit dependence on higher-order distributions, due to particle-particle interactions governed by interaction potentials and instantaneous events.

Following the construction of the BBGKY hierarchy, we then determined the general form of velocity moments of the first and second equation. This halves the dimensionality of the equations under consideration, however in addition to the dependence on higher-order distributions, we saw that the n^{th} moments of the k^{th} BBGKY hierarchy also depend on the $(n+1)^{\text{th}}$ moment of the same equation. As a result we cannot use the BBGKY equations to directly simulate granular media. We therefore require approximations that allow us to write the $(n+1)^{\text{th}}$ moment of distributions in terms of the first n moments, so that we can create a closed system of equations from the results of the BBGKY hierarchy. Thus, in the next chapter we investigate suitable moment closure schemes that help us construct a computationally-tractable continuum model.

Chapter 6

Moment closure methods

6.1 Motivation

Moment closure is a common method that is used to estimate moments of a stochastic process [140]. In systems that are approximated by a system of particles at the microscopic level, the moments of the stochastic or deterministic process of the microscopic dynamics may be of interest. It is common that the evolution of these moments will be determined by a set of differential equations, where the equation for the i^{th} moment will depend on the $(i+1)^{\text{th}}$ moment. Such an example is the BBGKY hierarchy, as derived in chapter 5. In this case the moments are of interest because they reduce the dimensionality of the system, and represent understandable physical quantities.

There are many different moment closure methods available in the literature, and the suitable moment closure scheme will depend on the system of interest. The simplest method is simply to set all moments which are a higher order than a selected value j to zero [140], which directly truncates the hierarchy. More relevant to kinetic theory is Grad's 13 moment equations [141], that utilize a moment closure method, where the phase density is approximated by a Hermite polynomial expansion about a local Maxwellian equilibrium distribution [142]. For more information on various moment closure methods and their applications, we refer the reader to [143].

One moment-closure method that produces a closed set of equations from the BBGKY hierachy is the maximum entropy method [78]. By asserting that the particle distributions must maximise statistical entropy, one can write any unknown quantities in a chosen truncated set of equations in the hierarchy in terms of known quantities. We provide more information on maximum entropy moment closure in section 6.5. As an example, the simplest maximum entropy moment closure case leads to a system of 5 equations of motion (in the three-dimensional case): one for the density $\rho(\boldsymbol{r}, t)$, three for the local average velocity $\boldsymbol{v}(\boldsymbol{r}, t)$, and one for the correlation function $g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, t)$. These results can be modified to include collisional terms in a straightforward manner. Unfortunately the resulting equations involve six-dimensional quantities that will be computationally intractable using standard numerical schemes. We therefore require a closure scheme that results in only three-dimensional quantities.

We wish to close the moments of the BBGKY hierarchy with the aim to construct a set of equations we can use for efficient simulations. Before we apply any moment closure procedure, we

introduce some new notation. Let

$$\rho^{(k)}(\boldsymbol{r}^k,t)=\int \mathrm{d}\boldsymbol{v}^k f^{(k)}(\boldsymbol{r}^k,\boldsymbol{v}^k,t), \tag{6.1}$$

$$\boldsymbol{u}_j^{(k)}(\boldsymbol{r}^k,t)=\int \mathrm{d}\boldsymbol{v}^k \boldsymbol{v}_j f^{(k)}(\boldsymbol{r}^k,\boldsymbol{v}^k,t), \tag{6.2}$$

$$\boldsymbol{\Pi}_{ij}^{(k)}(\boldsymbol{r}^k,t)=\int \mathrm{d}\boldsymbol{v}^k \boldsymbol{v}_i\boldsymbol{v}_j f^{(k)}(\boldsymbol{r}^k,\boldsymbol{v}^k,t), \tag{6.3}$$

$$\boldsymbol{Q}_{ijl}^{(k)}(\boldsymbol{r}^k,t)=\int \mathrm{d}\boldsymbol{v}^k \boldsymbol{v}_i\boldsymbol{v}_j\boldsymbol{v}_l f^{(k)}(\boldsymbol{r}^k,\boldsymbol{v}^k,t). \tag{6.4}$$

The quantities $\rho^{(k)},\boldsymbol{u}_j^{(k)},\boldsymbol{\Pi}_{ij}^{(k)}$ and $\boldsymbol{Q}_{ijl}^{k}$ are then scalars, vectors, matrices and 3-tensors respectively, where $\rho^{(k)}$ is the k-body density, $\boldsymbol{u}_j^{(k)}$ is the k-body local average velocity of particle j, $\boldsymbol{\Pi}_{ij}^{(k)}$ is the un-centred covariance in the local velocity between particles i and j, and $\boldsymbol{Q}_{ijl}^{(k)}$ provides higher order covariances between particles i,j and l.

We will also use the collision operator $\mathcal{L}_{\text{coll}}$ considered in *eq.* (4.79). It is difficult to construct analytic forms for the collision operator due to the presence of the characteristic function. Before considering moment closure methods we therefore discuss calculations necessary to construct moments of the collision operator.

6.2 Moments of the collision operator

We are interested in the first three moments of the derived collision operator in the first equation of the BBGKY hierarchy:

$$\begin{aligned}\mathcal{L}_{\text{coll}}[f^{(2)}] = &-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n} \\ &\times\left[\frac{1}{\alpha^2}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\text{out}},\boldsymbol{v}_2^{\text{out}}],t)-f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\right]\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.5}$$

The i^{th} moment of the collision operator is then given by

$$\begin{aligned}\boldsymbol{\mathcal{M}}_i\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right) = &-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}\boldsymbol{v}_1^{\otimes i}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n} \\ &\times\left[\frac{1}{\alpha^2}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\text{out}},\boldsymbol{v}_2^{\text{out}}],t)-f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\right]\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.6}$$

We then use the reverse collision rule as a substitution on the first term in the square brackets:

$$\boldsymbol{v}_1'=\boldsymbol{v}_1-\frac{1+\alpha}{2\alpha}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]\boldsymbol{n},\qquad \boldsymbol{v}_1'=\boldsymbol{v}_1-\frac{1+\alpha}{2\alpha}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]\boldsymbol{n}.$$

This results in the following equation:

$$\begin{aligned}\boldsymbol{\mathcal{M}}_i\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right) = &\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}\boldsymbol{v}_1^{\otimes i}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n} \\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\left(\boldsymbol{v}_1-\frac{1+\alpha}{2}((\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n})\boldsymbol{n}\right)^{\otimes i} \\ &\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.7}$$

When $i = 0$ the two integrals can be combined and so we find

$$\begin{aligned}\mathcal{M}_0\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right) =& \frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}|f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}|f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\ =&\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\ =&\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{u}_1^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)-\boldsymbol{u}_2^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}. \end{aligned} \tag{6.8}$$

We will see that that as $\varepsilon \to 0$ the values of $\boldsymbol{u}_i^{(2)}$ in $\mathcal{M}_0(\mathcal{L}_{\text{coll}}(f^{(2)})$ are the same, so the terms involved disappear. We are then left with the integral of an odd function over a unit sphere, which is zero. When $i = 1$ we can write

$$\begin{aligned}\boldsymbol{\mathcal{M}}_1\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right) =& \frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\boldsymbol{v}_1(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\ &-\frac{1+\alpha}{2\sqrt{2}}\int_{\mathbb{S}^{d-1}}\boldsymbol{n}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n},\\ =&\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{\Pi}_{11}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)-\boldsymbol{\Pi}_{12}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}\\ &-\frac{1+\alpha}{2\sqrt{2}}\int_{\mathbb{S}^{d-1}}\boldsymbol{n}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.\end{aligned}$$

Again we see that when $\varepsilon \to 0$ the first integral in $\mathcal{M}_1(\mathcal{L}_{\text{coll}}f^{(2)})$ will disappear as the argument of the integral becomes odd in n.

When $i = 2$ we expand

$$\begin{aligned}\left(\boldsymbol{v}_1-\frac{1+\alpha}{2}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]\boldsymbol{n}\right)^{\otimes 2} =&\boldsymbol{v}_1\otimes\boldsymbol{v}_1-\frac{(1+\alpha)}{2}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}(\boldsymbol{v}_1\otimes\boldsymbol{n}+\boldsymbol{n}\otimes\boldsymbol{v}_1)\\ &+\frac{(1+\alpha)^2}{4}((\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n})^2\boldsymbol{n}\otimes\boldsymbol{n}\end{aligned}$$

to rewrite the next moment as

$$\begin{aligned}\boldsymbol{\mathcal{M}}_2\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right) =& \frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\boldsymbol{v}_1\otimes\boldsymbol{v}_1(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\ -\frac{1+\alpha}{2\sqrt{2}}\int_{\mathbb{S}^{d-1}}&\Big[\left(\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\boldsymbol{v}_1[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\right)\otimes\boldsymbol{n}\\ &+\boldsymbol{n}\otimes\left(\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\boldsymbol{v}_1[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\right)\Big]\,\mathrm{d}\boldsymbol{n}\\ +\frac{(1+\alpha)^2}{4\sqrt{2}}\int_{\mathbb{S}^{d-1}}&(\boldsymbol{n}\otimes\boldsymbol{n})\left[\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}|^3 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\right]\mathrm{d}\boldsymbol{n}.\end{aligned}$$

The first term is given by

$$\begin{aligned}&\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\boldsymbol{v}_1\otimes\boldsymbol{v}_1(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\ &\qquad=\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{Q}_{111}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)-\boldsymbol{Q}_{112}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}.\end{aligned}$$

The remaining integral terms in the first three moments are:

$$\boldsymbol{A}_1 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}, \tag{6.9}$$

$$\boldsymbol{A}_2 = \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[\boldsymbol{n}\otimes\boldsymbol{v}_1+\boldsymbol{v}_1\otimes\boldsymbol{n}][(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}, \tag{6.10}$$

$$\boldsymbol{A}_3 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n}\otimes\boldsymbol{n} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^3 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}. \tag{6.11}$$

For $\boldsymbol{A}_1$ if we expand the scalar product in the integral we have

$$\begin{aligned}\boldsymbol{A}_1 = &\int_{\mathbb{S}^{d-1}} \boldsymbol{n} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1\cdot\boldsymbol{n})^2-(\boldsymbol{v}_1\cdot\boldsymbol{n})(\boldsymbol{v}_2\cdot\boldsymbol{n})] f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\ &+\int_{\mathbb{S}^{d-1}} \boldsymbol{n} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_2\cdot\boldsymbol{n})^2-(\boldsymbol{v}_1\cdot\boldsymbol{n})(\boldsymbol{v}_2\cdot\boldsymbol{n})] f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.\end{aligned}$$

In the second integral we then switch the labelling for $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ and use the symmetry of $f^{(2)}$ to find:

$$\begin{aligned}\boldsymbol{A}_1 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n} \Bigg[&\int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1\cdot\boldsymbol{n})^2-(\boldsymbol{v}_1\cdot\boldsymbol{n})(\boldsymbol{v}_2\cdot\boldsymbol{n})] f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\\ &+\int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}[(\boldsymbol{v}_2\cdot\boldsymbol{n})^2-(\boldsymbol{v}_1\cdot\boldsymbol{n})(\boldsymbol{v}_2\cdot\boldsymbol{n})] f^{(2)}([\boldsymbol{r}_1+\varepsilon\boldsymbol{n},\boldsymbol{r}_1],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2 \Bigg]\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.12}$$

Thus when $\varepsilon \to 0$ we are left with the integral of an odd function over the unit sphere, which is zero. This shows that the effect of collisions on the local average velocity is negligible when particles are small.

The integrals A_2 and A_3 cannot be reduced to odd functions over the unit sphere in the limit as $\varepsilon \to 0$, and so are non-zero and contribute to the dynamics. It is difficult to simplify these equations any further independently of dimension due to the inclusion of the characteristic function. We will consider the calculations in $d = 1$ and $d = 2$ in appendix A.

6.2.1 The RET collision operator

We also provide a comparison with the RET collision operator

$$\begin{aligned}\mathcal{L}^{\mathrm{RET}}_{\mathrm{coll}}[f^{(2)}] = &-\sigma^{d-1}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^d} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\\ &\times\left[\frac{1}{\alpha^2}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1-\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\mathrm{out}},\boldsymbol{v}_2^{\mathrm{out}}],t)-f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\right]\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.13}$$

We consider the moments of the RET collision operator and apply the collisional substitution so that

$$\begin{aligned}\boldsymbol{\mathcal{M}}_i\left(\mathcal{L}^{\mathrm{RET}}_{\mathrm{coll}}[f^{(2)}]\right) = &\sigma^{d-1}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}\boldsymbol{v}^{\otimes i}\\ &\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n} f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\end{aligned} \tag{6.14}$$

$$+\sigma^{d-1}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\left(\boldsymbol{v}_1-\frac{1+\alpha}{2}((\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n})\boldsymbol{n}\right)^{\otimes i} \tag{6.15}$$
$$\times(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1-\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.$$

For the first moment by switching the notation in the second integral and using the symmetry, we have that

$$\mathcal{M}_0\left(\mathcal{L}^{\mathrm{RET}}_{\mathrm{coll}}[f^{(2)}]\right)=\sigma^{d-1}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n} \tag{6.16}$$
$$\times[f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)-f^{(2)}([\boldsymbol{r}_1-\varepsilon\boldsymbol{n},\boldsymbol{r}_1],\boldsymbol{v}^2,t)]\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.$$

For the second moment,

$$\begin{aligned}\boldsymbol{\mathcal{M}}_1\left(\mathcal{L}^{\mathrm{RET}}_{\mathrm{coll}}[f^{(2)}]\right)=&\sigma^{d-1}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}\boldsymbol{v}_1(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\\&\times[f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)-f^{(2)}([\boldsymbol{r}_1-\varepsilon\boldsymbol{n},\boldsymbol{r}_1],\boldsymbol{v}^2,t)]\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\&-\frac{\sigma^{d-1}(1+\alpha)}{2}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\boldsymbol{n}((\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n})^2f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1-\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n},\end{aligned}$$

and for the third

$$\begin{aligned}\boldsymbol{\mathcal{M}}_2\left(\mathcal{L}^{\mathrm{RET}}_{\mathrm{coll}}[f^{(2)}]\right)=&\sigma^{d-1}\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}>0}\boldsymbol{v}_1^{\otimes2}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}\\&\times[f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)-f^{(2)}([\boldsymbol{r}_1-\varepsilon\boldsymbol{n},\boldsymbol{r}_1],\boldsymbol{v}^2,t)]\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\-\frac{\sigma^{d-1}(1+\alpha)}{2}&\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}((\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n})^2\\&\times[\boldsymbol{n}\otimes\boldsymbol{v}_1+\boldsymbol{v}_1\otimes\boldsymbol{n}]f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1-\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}\\+\frac{\sigma^{d-1}(1+\alpha)^2}{4}&\int_{\mathbb{S}^{d-1}}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\boldsymbol{n}\otimes\boldsymbol{n}((\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n})^2f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1-\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.\end{aligned}$$

Thus aside from the first terms in each of the moments the results from the Boltzmann collision operator can also be applied to the RET collision operator, after a change of sign.

6.2.2 Collision terms in the second BBGKY equation

Collision terms also contribute to the first three moments of the second equation of the BBGKY hierarchy. The terms that arise in this case are

$$\begin{aligned}\mathcal{L}^2_{\mathrm{coll}}[f^{(2)},f^{(3)}]:=&-\frac{1}{\sqrt{2}}\chi_{\substack{\boldsymbol{r}_2=\boldsymbol{r}+1+\varepsilon\boldsymbol{n}\\\boldsymbol{n}\in\mathbb{S}^{d-1}}}(\boldsymbol{r}_1-\boldsymbol{r}_2)\cdot\boldsymbol{n}f^{(2)}(\boldsymbol{r}^2,\boldsymbol{v}^2,t)\\&-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^2}\int_{\mathbb{R}^d}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}>0}\left[\frac{1}{\alpha^2}f^{(3)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1^{\mathrm{out}},\boldsymbol{v}_2,\boldsymbol{v}_3^{\mathrm{out}}],t)\right.\\&\qquad\left.-f^{(3)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^3,t)\right](\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{v}_3\,\mathrm{d}\boldsymbol{n}\\&-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^2}\int_{\mathbb{R}^d}\chi_{(\boldsymbol{v}_2-\boldsymbol{v}_3)\cdot\boldsymbol{n}>0}\left[\frac{1}{\alpha^2}f^{(3)}([\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_2+\varepsilon\boldsymbol{n}],[\boldsymbol{v}_1,\boldsymbol{v}_2^{\mathrm{out}},\boldsymbol{v}_3^{\mathrm{out}}],t)\right.\end{aligned}$$

$$- f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2 + \varepsilon\boldsymbol{n}], \boldsymbol{v}^3, t)\Big](\boldsymbol{v}_2 - \boldsymbol{v}_3)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{v}_3\,\mathrm{d}\boldsymbol{n}. \tag{6.17}$$

Moments of the first term in $\mathcal{L}^2_{\mathrm{coll}}[f^{(2)}, f^{(3)}]$ can be constructed easily as the characteristic function does not involve velocities $\boldsymbol{v}^2$. The other terms can be considered in the same way as in the first moment, we are left with

$$\begin{aligned}\mathcal{M}_0\left(\mathcal{L}^2_{\mathrm{coll}}[f^{(2)}, f^{(3)}]\right) = &-\frac{1}{\sqrt{2}}\chi_{\boldsymbol{r}_2=\boldsymbol{r}+1+\varepsilon\boldsymbol{n}}\underset{\boldsymbol{n}\in\mathbb{S}^{d-1}}{}\left(\boldsymbol{u}_1^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t) - \boldsymbol{u}_2^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t)\right)\cdot\boldsymbol{n}\\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left(\boldsymbol{u}_1^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t) - \boldsymbol{u}_3^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t)\right)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}\\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left(\boldsymbol{u}_2^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2+\varepsilon\boldsymbol{n}], t) - \boldsymbol{u}_3^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2+\varepsilon\boldsymbol{n}], t)\right)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n},\end{aligned} \tag{6.18}$$

$$\begin{aligned}\boldsymbol{\mathcal{M}}_1\left(\mathcal{L}^2_{\mathrm{coll}}[f^{(2)}, f^{(3)}]\right) = &-\frac{1}{\sqrt{2}}\chi_{\boldsymbol{r}_2=\boldsymbol{r}+1+\varepsilon\boldsymbol{n}}\underset{\boldsymbol{n}\in\mathbb{S}^{d-1}}{}\left(\boldsymbol{\Pi}_{11}^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t) - \boldsymbol{\Pi}_{12}^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t)\right)\cdot\boldsymbol{n}\\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{\Pi}_{11}^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t) - \boldsymbol{\Pi}_{13}^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}\\ &-\frac{1+\alpha}{2\sqrt{2}}\int_{\mathbb{S}^{d-1}}\boldsymbol{n}\int_{\mathbb{R}^{3d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}]^2 f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], \boldsymbol{v}^3, t)\,\mathrm{d}\boldsymbol{v}^3\,\mathrm{d}\boldsymbol{n}\\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left(\boldsymbol{\Pi}_{12}^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2+\varepsilon\boldsymbol{n}, t) - \boldsymbol{\Pi}_{13}^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2+\varepsilon\boldsymbol{n}, t)\right)\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n},\end{aligned} \tag{6.19}$$

$$\begin{aligned}\boldsymbol{\mathcal{M}}_2\left(\mathcal{L}^2_{\mathrm{coll}}[f^{(2)}, f^{(3)}]\right) = &-\frac{1}{\sqrt{2}}\chi_{\boldsymbol{r}_2=\boldsymbol{r}+1+\varepsilon\boldsymbol{n}}\underset{\boldsymbol{n}\in\mathbb{S}^{d-1}}{}\left(\boldsymbol{Q}_{111}^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t) - \boldsymbol{Q}_{112}^{(2)}([\boldsymbol{r}_1, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t)\right)\cdot\boldsymbol{n}\\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{Q}_{111}^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t) - \boldsymbol{Q}_{113}^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}\\ &-\frac{1+\alpha}{2\sqrt{2}}\int_{\mathbb{S}^{d-1}}\Big[\left(\int_{\mathbb{R}^{3d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}<0}\boldsymbol{v}_1[(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}]^2 f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], \boldsymbol{v}^3, t)\,\mathrm{d}\boldsymbol{v}^3\right)\otimes\boldsymbol{n}\\ &\qquad+\boldsymbol{n}\otimes\left(\int_{\mathbb{R}^{3d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}<0}\boldsymbol{v}_1[(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}]^2 f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], \boldsymbol{v}^3, t)\,\mathrm{d}\boldsymbol{v}^3\right)\Big]\,\mathrm{d}\boldsymbol{n}\\ &+\frac{(1+\alpha)^2}{4\sqrt{2}}\int_{\mathbb{S}^{d-1}}(\boldsymbol{n}\otimes\boldsymbol{n})\left[\int_{\mathbb{R}^{3d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_3)\cdot\boldsymbol{n}<0}|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}|^3 f^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_1+\varepsilon\boldsymbol{n}], \boldsymbol{v}^3, t)\,\mathrm{d}\boldsymbol{v}^2\right]\mathrm{d}\boldsymbol{n}\\ &+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{Q}_{112}^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2+\varepsilon\boldsymbol{n}], t) - \boldsymbol{Q}_{113}^{(3)}([\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_2+\varepsilon\boldsymbol{n}], t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.20}$$

The integrals here that involve characteristic functions are analogous to $\boldsymbol{A}_1$, $\boldsymbol{A}_2$ and $\boldsymbol{A}_3$ considered in the first case, which implies that the derivations considered in Appendix A can be modified and used to construct computationally efficient forms for the terms here.

We can now rewrite the first three moments of the first two equations of the BBGKY hierarchy.

6.3 The BBGKY equations

The first three moments of the first BBGKY hierarchy can then be written as:

$$\frac{\partial}{\partial t}\rho^{(1)}(\boldsymbol{r}_1,t) = -\frac{1}{m}\nabla_{\boldsymbol{r}_1}\cdot\boldsymbol{u}^{(1)}(\boldsymbol{r}_1,t) + \frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{u}_1^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t) - \boldsymbol{u}_2^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}, \tag{6.21}$$

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{u}_1^{(1)}(\boldsymbol{r}_1,t) = &-\frac{1}{m}(\nabla_{\boldsymbol{r}_1}\cdot\boldsymbol{\Pi}_{11}^{(1)}(\boldsymbol{r}_1,t)) - (\nabla_{\boldsymbol{r}_1}V^{\mathrm{ext}})\rho^{(1)}(\boldsymbol{r}_1,t)\\ &-\int d\boldsymbol{r}_2(\nabla_{\boldsymbol{r}_1}v_2(\boldsymbol{r}_1,\boldsymbol{r}_2))\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t)\\ &-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{\Pi}_{11}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t) - \boldsymbol{\Pi}_{12}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}\\ &+\frac{1+\alpha}{2\sqrt{2}}\int_{\mathbb{S}^{d-1}}\boldsymbol{n}\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n},\end{aligned} \tag{6.22}$$

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{\Pi}_{11}^{(1)}(\boldsymbol{r}_1,t) = &-\frac{1}{m}(\nabla_{\boldsymbol{r}_1}\cdot\boldsymbol{Q}_{111}^{(1)}(\boldsymbol{r}_1,t)) - (\nabla_{\boldsymbol{r}_1}V^{\mathrm{ext}})\boldsymbol{u}_1^{(1)}(\boldsymbol{r}_1,t)\\ &-\int d\boldsymbol{r}_2(\nabla_{\boldsymbol{r}_1}v_2(\boldsymbol{r}_1,\boldsymbol{r}_2))\boldsymbol{u}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\\ &-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{Q}_{111}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}]) - \boldsymbol{Q}_{112}^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}])\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}\\ &+\frac{1+\alpha}{2\sqrt{2}}\int_{\mathbb{S}^{d-1}}\Big[\left(\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\boldsymbol{v}_1[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\right)\otimes\boldsymbol{n}\\ &\qquad+\boldsymbol{n}\otimes\left(\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}\boldsymbol{v}_1[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\right)\Big]\,\mathrm{d}\boldsymbol{n}\\ &-\frac{(1+\alpha)^2}{4\sqrt{2}}\int_{\mathbb{S}^{d-1}}(\boldsymbol{n}\otimes\boldsymbol{n})\left[\int_{\mathbb{R}^{2d}}\chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}|(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}|^3 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\right]\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.23}$$

The moments for the second equation read:

$$\frac{\partial}{\partial t}\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t) = -\frac{1}{m}\nabla_{\boldsymbol{r}_1}\cdot\boldsymbol{u}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t) - \frac{1}{m}\nabla_{\boldsymbol{r}_2}\cdot\boldsymbol{u}_2^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t) + \mathcal{M}_0\left(\mathcal{L}_{\mathrm{coll}}^2[f^{(2)},f^{(3)}]\right), \tag{6.24}$$

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{u}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t) = &-\frac{1}{m}\nabla_{\boldsymbol{r}_1}\cdot\boldsymbol{\Pi}_{11}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t) - \frac{1}{m}\nabla_{\boldsymbol{r}_1}\cdot\boldsymbol{\Pi}_{12}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t)\\ &-(\nabla_{\boldsymbol{r}_1}V^{\mathrm{ext}}+\nabla_{\boldsymbol{r}_1}v_2(\boldsymbol{r}_1,\boldsymbol{r}_2))\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t)\\ &-\int\mathrm{d}\boldsymbol{r}_3(\nabla_{\boldsymbol{r}_1}V_2(\boldsymbol{r}_1,\boldsymbol{r}_3))\rho^{(3)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3,t)\\ &-\boldsymbol{\mathcal{M}}_1\left(\mathcal{L}_{\mathrm{coll}}^2[f^{(2)},f^{(3)}]\right),\end{aligned} \tag{6.25}$$

$$\begin{aligned}\frac{\partial}{\partial t}\boldsymbol{\Pi}_{11}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t) = &-\frac{1}{m}\nabla_1\cdot\boldsymbol{Q}_{111}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t) - \frac{1}{m}\nabla_2\cdot\boldsymbol{Q}_{112}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t)\\ &-2(\nabla_{\boldsymbol{r}_1}V_1+\nabla_{\boldsymbol{r}_1}V_{12})\boldsymbol{u}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,t)\end{aligned} \tag{6.26}$$

$$- \int \mathrm{d}\boldsymbol{r}_3 (\nabla_{\boldsymbol{r}_1} V_{13}) \boldsymbol{u}_1^{(3)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3, t)$$
$$- \boldsymbol{\mathcal{M}}_2 \left(\mathcal{L}^2_{\mathrm{coll}}[f^{(2)}, f^{(3)}] \right),$$

where $\mathcal{M}_i \left(\mathcal{L}^2_{\mathrm{coll}}[f^{(2)}, f^{(3)}] \right)$ are the moments of the collisional terms in the second BBGKY hierarchy.

6.4 Dynamical density functional theory

After constructing the BBGKY hierarchy, we must make some assumptions that allow us to truncate the chain of PDEs in a way that reduces the dimensionality of the system to something which is accessible to fast numerical methods.

Recently, dynamical density functional theory (DDFT) has provided useful results for modelling a variety of particle systems. Derivations of a DDFT are motivated by the fact that the following rigorously known relationship between the N-particle distribution function and the one-body position density [144]:

$$\rho(\boldsymbol{r}_1, t) = N \int_{\mathbb{R}^{(N-1)d}} \int_{\mathbb{R}^{Nd}} f^{(N)}(\boldsymbol{r}^N, \boldsymbol{v}^N, t) \chi_{\mathcal{D}_N(\boldsymbol{r}_1, t)} \, \mathrm{d}\boldsymbol{v}^N \, \mathrm{d}\boldsymbol{r}^{N-1}. \tag{6.27}$$

DDFTs have been constructed for various particle systems, usually focussing on colloids (soft particles suspended in a bath). One of the first DDFTs constructed was in the *over-damped* case, where the friction coefficient γ is large. In this case, [71] constructed a PDE determining how the density of the system of particles changes over time, relying on assumptions on the correlations between higher and lower-order distributions:

$$\frac{\partial \rho}{\partial t} = \frac{1}{\gamma m} \nabla \cdot \left[\rho(\boldsymbol{r}, t) \nabla \frac{\delta \mathcal{F}[\rho(\boldsymbol{r}, t)]}{\delta \rho(\boldsymbol{r}, t)} \right], \tag{6.28}$$

where $\mathcal{F}[\rho]$ is the Helmholtz free energy functional (defined at equilibrium) and $\delta \mathcal{F}[\rho(\boldsymbol{r}, t)]/\delta \rho(\boldsymbol{r}, t)$ is its functional derivative. Thus given a good understanding of $\mathcal{F}[\rho]$, a well-studied energy functional from classical density functional theory, non-equilibrium dynamics can be simulated, with the guarantee that the equilibrium profile will be correct.

Following from this, DDFTs that incorporated inertial effects were constructed [68], [73], for example in [68] they provide

$$\frac{\partial \rho}{\partial t} + \nabla \cdot (\rho \boldsymbol{v}) = 0, \tag{6.29}$$
$$\rho \left(\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v} \cdot \nabla \boldsymbol{v}) \right) + \gamma \rho \boldsymbol{v} = -\frac{1}{m} \rho \nabla \frac{\delta \mathcal{F}[\rho]}{\delta \rho}. \tag{6.30}$$

Both equations can include volume exclusion effects by including terms in the free energy functional. Results are also available for volume exclusion ([88], [89]), as well as hydrodynamic interactions [69], non-spherical particles [145] and multiple species systems [76]. However, these results cannot be appropriate for granular media because they do not involve (inelastic) collisional effects. In [74], the RET collision operator is used in a one-dimensional derivation to construct a DDFT for the density of a system of particles that undergo inelastic effects. This paper provides important insight

into the thermodynamic properties of hard, inelastic rods. However, the results are restricted to one dimension, and a single equation is constructed that is used to model the density of the system of particles incorporates effects that occur due to collisions, when we have seen that the effects are most prevalent in the kinetic energy of the system. Furthermore, an analytic approximation on form of the the correlation function $g^{(2)}$ (which we introduce in eq. (6.32)) is used, which we will see in chapter 7 is not appropriate for inelastic systems.

In this case, we prefer to construct a set of equations that model the density, local average velocity, and fluctuations in the kinetic temperature of a system of inelastic hard spheres in d dimensions, using the results we have obtained so far, by making three assumptions on the system under consideration. We make a weaker assumption on the form of the correlation function, instead choosing to construct $g^{(2)}$ numerically in the examples we consider.

We begin the construction of a DDFT for granular media by considering the first three moments of the first equation in the BBGKY hierarchy, eqs. (6.21) to (6.23).

6.4.1 Assumptions for the DDFT

There are three assumptions that are made to close the system of equations.

1. We approximate many-body interactions in the non-equilibrium system using those of a fluid at equilibrium with the same one-body density profile [68]:

$$-\int \mathrm{d}\boldsymbol{v}_1 \Bigg[\mathrm{d}\boldsymbol{r}_2\, \mathrm{d}\boldsymbol{v}_2 \nabla_{\boldsymbol{r}_1} v_2(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot \nabla_{\boldsymbol{v}_1} f^{(2)} \tag{6.31}$$
$$+ \int\int\int\int d\boldsymbol{r}_2\, \mathrm{d}\boldsymbol{v}_2\, \mathrm{d}\boldsymbol{r}_3\, \mathrm{d}\boldsymbol{v}_3 \nabla_{\boldsymbol{r}_1} v_3(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{r}_3) \cdot \nabla_{\boldsymbol{v}_1} f^{(3)} + \ldots \Bigg] = \rho(\boldsymbol{r}_1) \nabla_{\boldsymbol{r}_1} \frac{\delta \mathcal{F}_{\mathrm{ex}}[\rho(\boldsymbol{r}_1)]}{\delta \rho(\boldsymbol{r}_1)}$$

   where $\frac{\delta \mathcal{F}_{\mathrm{ex}}[\rho(\boldsymbol{r})]}{\delta \rho(\boldsymbol{r})}$ is the functional derivative of the excess part of the Helmholtz free energy functional [146]. This term is well studied and numerical schemes to approximate various interparticle potentials are available. The free energy functional is also used to incorporate volume exclusion in the system using the (exact) Percus free energy functional in one dimension [88], or the (accurate) fundamental measure theory in multiple dimensions [89]. In our examples we will use the free energy functional to approximate volume exclusion effects.

2. We can write higher-order distributions as follows:

$$f^{(k)}(\boldsymbol{r}^k, \boldsymbol{v}^k, t) = g^{(k)}(\boldsymbol{r}^k, t) \prod_{i=1}^{k} f^{(1)}(\boldsymbol{r}_i, \boldsymbol{v}_i), \tag{6.32}$$

   i.e. higher-order distributions are uncorrelated in velocity. Although the effect of inelastic collisions between particles produces correlations in particle velocities, we assume that the effects are short-range so are dominated by correlations in position. This assumption is also implemented in the maximum entropy moment closure method when a Maxwellian reference distribution is considered.

3. The one-particle distribution function can be approximated by a local equilibrium Maxwell-

Boltzmann distribution [139]:

$$f_{\text{le}}^{(1)}(\boldsymbol{r},\boldsymbol{v},t)=\frac{\rho(\boldsymbol{r},t)}{|2\pi m k_B T\boldsymbol{E}(\boldsymbol{r},t)|}\exp\left(-\frac{1}{2mk_BT}(\boldsymbol{v}-\bar{\boldsymbol{p}}(\boldsymbol{r},t))^T\boldsymbol{E}(\boldsymbol{r},t)^{-1}(\boldsymbol{v}-\bar{\boldsymbol{p}}(\boldsymbol{r},t))\right). \tag{6.33}$$

We note that the local equilibrium of a granular fluid is known not to be Maxwellian [147]. However, by assuming a Maxwellian distribution we are able to write the second moment of $f^{(1)}$ in terms of known quantities. In fact, under these assumptions

$$\int_{\mathbb{R}^d} f_{\text{le}}^{(1)}(\boldsymbol{r},\boldsymbol{v},t)\,\mathrm{d}\boldsymbol{v}=\rho(\boldsymbol{r},t), \tag{6.34}$$

$$\int_{\mathbb{R}^d} \boldsymbol{v} f_{\text{le}}^{(1)}(\boldsymbol{r},\boldsymbol{v},t)\,\mathrm{d}\boldsymbol{v}=\rho(\boldsymbol{r},t)\bar{\boldsymbol{p}}(\boldsymbol{r},t), \tag{6.35}$$

$$\int_{\mathbb{R}^d} \boldsymbol{v}\otimes\boldsymbol{v} f_{\text{le}}^{(1)}(\boldsymbol{r},\boldsymbol{v},t)\,\mathrm{d}\boldsymbol{v}=\frac{k_BT\rho(\boldsymbol{r},t)}{m}\boldsymbol{E}(\boldsymbol{r},t)+\frac{\rho(\boldsymbol{r},t)}{m^2}\bar{\boldsymbol{p}}(\boldsymbol{r},t)\otimes\bar{\boldsymbol{p}}(\boldsymbol{r},t). \tag{6.36}$$

There have been other approximations for the local equilibrium, such as the Sonine polynomial approximation [129], that could also be implemented here (and may be more appropriate). Alternatively additional correctional terms can be included, which have been considered in the colloidal case [70], [148].

Under these assumptions we can use the first three moments of the first equation in the BBGKY hierarchy to create a DDFT for granular media.

6.4.2 A DDFT for granular media

The first equation in the DDFT is determined by the first moment of eq. (5.22). The derivation is straightforward, as most terms in the first moment of the first BBGKY hierarchy cancel and we are left with:

$$\begin{aligned}\frac{\partial}{\partial t}\rho(\boldsymbol{r},t)&=-\frac{1}{m}\nabla_{\boldsymbol{r}}\cdot\boldsymbol{u}_1^{(1)}(\boldsymbol{r},t)-\mathcal{M}_0\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right)\\&=-\frac{1}{m}\nabla_{\boldsymbol{r}}\cdot\boldsymbol{u}_1^{(1)}(\boldsymbol{r},t)-\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[\boldsymbol{u}_1^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)-\boldsymbol{u}_2^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],t)\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}.\end{aligned} \tag{6.37}$$

By writing

$$\begin{aligned}\boldsymbol{u}_1^{(1)}&=\rho\bar{\boldsymbol{p}},\\\boldsymbol{u}_1^{(2)}(\boldsymbol{r},\boldsymbol{r}+\varepsilon\boldsymbol{n})&=g(\boldsymbol{r},\boldsymbol{r}+\varepsilon\boldsymbol{n})\rho(\boldsymbol{r})\rho(\boldsymbol{r}+\varepsilon\boldsymbol{n})\bar{\boldsymbol{p}}(\boldsymbol{r}),\\\boldsymbol{u}_2^{(2)}(\boldsymbol{r},\boldsymbol{r}+\varepsilon\boldsymbol{n})&=g(\boldsymbol{r},\boldsymbol{r}+\varepsilon\boldsymbol{n})\rho(\boldsymbol{r})\rho(\boldsymbol{r}+\varepsilon\boldsymbol{n})\bar{\boldsymbol{p}}(\boldsymbol{r}+\varepsilon\boldsymbol{n})\end{aligned}$$

we arrive at the continuity equation with collisional corrections:

$$\boxed{\frac{\partial}{\partial t}\rho(\boldsymbol{r},t)=-\frac{1}{m}\nabla_{\boldsymbol{r}}\cdot(\rho\bar{\boldsymbol{p}})+\frac{1}{\sqrt{2}}\int_{\mathbb{S}^{d-1}}\left[g^{(2)}(\boldsymbol{r},\boldsymbol{r}+\varepsilon\boldsymbol{n})\rho\rho_+\bar{\boldsymbol{p}}-g^{(2)}(\boldsymbol{r},\boldsymbol{r}+\varepsilon\boldsymbol{n})\rho\rho_+\bar{\boldsymbol{p}}_+\right]\cdot\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}.} \tag{6.38}$$

where we write $\bar{\boldsymbol{p}}_+ = \bar{\boldsymbol{p}}(\boldsymbol{r} + \varepsilon \boldsymbol{n})$ and similar for ρ_+. We have already noted that collisional corrections in the density are small when ε is small as they consider the difference in local momentum at the diameter of a single particle, if they are neglected we arrive at the continuity equation given in previous DDFTs:

$$\frac{\partial}{\partial t}\rho(\boldsymbol{r},t) = -\frac{1}{m}\nabla_{\boldsymbol{r}} \cdot (\rho(\boldsymbol{r},t)\bar{\boldsymbol{p}}(\boldsymbol{r},t)). \tag{6.39}$$

The effect of this equation is intuitive; local density is transported due to changes in local velocity, weighted by the density (*i.e.* the flux).

For the second equation we have that

$$\begin{aligned}\frac{\partial}{\partial t}(\rho\bar{\boldsymbol{p}}) =& -\frac{1}{m}\nabla_{\boldsymbol{r}} \cdot \left(\frac{K_B T\rho(\boldsymbol{r},t)}{m}\boldsymbol{E}(\boldsymbol{r},t) + \frac{\rho(\boldsymbol{r},t)}{m^2}\bar{\boldsymbol{p}}(\boldsymbol{r},t)\otimes\bar{\boldsymbol{p}}(\boldsymbol{r},t)\right) \\ &- \rho\nabla_{\boldsymbol{r}} V^{\text{ext}} - \rho\nabla_{\boldsymbol{r}}\frac{\delta\mathcal{F}_{\text{ex}}[\rho]}{\delta\rho} - \gamma\rho\bar{\boldsymbol{p}} - \boldsymbol{\mathcal{M}}_1\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right).\end{aligned} \tag{6.40}$$

Using $\nabla_{\boldsymbol{r}}\rho = \rho\nabla_{\boldsymbol{r}}\ln(\rho)$ we can include and remove the ideal gas contribution to the equation

$$\begin{aligned}\frac{\partial}{\partial t}(\rho\bar{\boldsymbol{p}}) =& -\frac{1}{m}\nabla_{\boldsymbol{r}} \cdot \left(\frac{K_B T\rho(\boldsymbol{r},t)}{m}\boldsymbol{E}(\boldsymbol{r},t) + \frac{\rho(\boldsymbol{r},t)}{m^2}\bar{\boldsymbol{p}}(\boldsymbol{r},t)\otimes\bar{\boldsymbol{p}}(\boldsymbol{r},t)\right) + \frac{k_B T}{m}\nabla_{\boldsymbol{r}}\rho \\ &- \frac{1}{m}\rho\nabla_{\boldsymbol{r}}\frac{\delta\mathcal{F}[\rho]}{\delta\rho} - \gamma\rho\bar{\boldsymbol{p}} - \boldsymbol{\mathcal{M}}_1\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right) - \rho\nabla_{\boldsymbol{r}} V^{\text{ext}} - \rho\nabla_{\boldsymbol{r}}\frac{\delta\mathcal{F}_{\text{ex}}[\rho]}{\delta\rho}.\end{aligned}$$

Here $\mathcal{F}[\rho]$ is the full Helmholtz free energy functional

$$\mathcal{F}[\rho] = k_B T\int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{r}\rho(\boldsymbol{r})[\ln(\Lambda^3\rho(\boldsymbol{r},t)) - 1] + \mathcal{F}_{\text{ex}}[\rho(\boldsymbol{r},t)] + \int_{\mathbb{R}^d} V^{\text{ext}}(\boldsymbol{r},t)\rho(\boldsymbol{r},t), \tag{6.41}$$

where Λ is the (irrelevant) De Broglie wavelength. Note that in the DDFT derivation with two moments, the term $\nabla_{\boldsymbol{r}}\rho(\boldsymbol{r},t)$ is not present outside of the free energy functional, as $\boldsymbol{E} = \boldsymbol{I}$.

To advance we use the following calculus results [68]:

$$\begin{aligned}\nabla_{\boldsymbol{r}} \cdot (\rho\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}) =& \rho[\nabla_{\boldsymbol{r}} \cdot (\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}})] + (\nabla_{\boldsymbol{r}}\rho) \cdot (\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}) \\ =& \rho[\bar{\boldsymbol{p}}(\nabla \cdot \bar{\boldsymbol{p}}) + \bar{\boldsymbol{p}} \cdot \nabla\otimes\bar{\boldsymbol{p}}] + \bar{\boldsymbol{p}}(\bar{\boldsymbol{p}} \cdot \nabla_{\boldsymbol{r}}\rho),\end{aligned}$$

where $\nabla_{\boldsymbol{r}}\otimes\bar{\boldsymbol{p}} = \left(\frac{\partial}{\partial\boldsymbol{r}_i}\bar{p}_j\right)_{i,j=1,\ldots,d}$ is a tensor product derivative. Furthermore,

$$\begin{aligned}\frac{\partial}{\partial t}(\rho\bar{\boldsymbol{p}}) =& \rho\frac{\partial\bar{\boldsymbol{p}}}{\partial t} + \bar{\boldsymbol{p}}\frac{\partial\rho}{\partial t}, \\ =& \rho\frac{\partial\bar{\boldsymbol{p}}}{\partial t} - \bar{\boldsymbol{p}}[\nabla_{\boldsymbol{r}} \cdot (\rho\bar{\boldsymbol{p}})] - \bar{\boldsymbol{p}}\mathcal{M}_0\left(\mathcal{L}_{\text{coll}} f^{(2)}\right), \\ =& \rho\frac{\partial\bar{\boldsymbol{p}}}{\partial t} - \rho\bar{\boldsymbol{p}}(\nabla_{\boldsymbol{r}} \cdot \bar{\boldsymbol{p}}) - \bar{\boldsymbol{p}}(\bar{\boldsymbol{p}} \cdot \nabla_{\boldsymbol{r}}\rho) - \bar{\boldsymbol{p}}\mathcal{M}_0\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right).\end{aligned}$$

After inserting these results into the equation and dividing by ρ, and using $\nabla_{\boldsymbol{r}}\rho = \nabla_{\boldsymbol{r}} \cdot (\rho\boldsymbol{I})$, we

arrive at the momentum equation

$$\boxed{\frac{\partial \boldsymbol{v}}{\partial t} + \boldsymbol{v}\cdot\boldsymbol{\nabla}_{\boldsymbol{r}}\boldsymbol{v} + \gamma\boldsymbol{v} + \frac{k_BT}{m\rho}\nabla_{\boldsymbol{r}}\cdot(\rho(\boldsymbol{E}-\boldsymbol{I})) + \frac{1}{m}\nabla_{\boldsymbol{r}}\frac{\delta\mathcal{F}[\rho]}{\delta\rho} + \frac{1}{m\rho}\boldsymbol{\mathcal{M}}_1^c(\mathcal{L}_{\text{coll}}[f^{(1)},g^{(2)}]) = 0.} \tag{6.42}$$

where $\boldsymbol{\mathcal{M}}_1^c(\mathcal{L}_{\text{coll}}[f^{(2)}])$ is the first centred moment of the collision operator:

$$\begin{aligned}\boldsymbol{\mathcal{M}}_1^c(\mathcal{L}_{\text{coll}}[f^{(1)},g^{(2)}]) :=& \int_{\mathbb{R}^d}\mathrm{d}\boldsymbol{v}\frac{(\boldsymbol{v}-\bar{\boldsymbol{p}})}{m}\mathcal{L}_{\text{coll}}[f^{(1)},g^{(2)}], &(6.43)\\ =&\boldsymbol{\mathcal{M}}_1(\mathcal{L}_{\text{coll}}[f^{(1)},g^{(2)}]) - \bar{\boldsymbol{p}}\mathcal{M}_0(\mathcal{L}_{\text{coll}}[f^{(1)},g^{(2)}]). &(6.44)\end{aligned}$$

We note that (as it will be useful in the derivation of the third equation) using Einstein summation notation the momentum equation can be written as:

$$\frac{\partial}{\partial t}\bar{p}_j = -\bar{p}_i\partial_i\bar{p}_j - \gamma\bar{p}_j - \frac{k_BT}{m\rho}\partial_i(\rho(E_{ij}-I_{ij})) - \frac{1}{m}\partial_j\frac{\delta\mathcal{F}}{\delta\rho} - \frac{1}{m}(\boldsymbol{\mathcal{M}}_1^c\mathcal{L}_{\text{coll}}(f^{(1)},g^{(2)}))_j. \tag{6.45}$$

It remains to consider the third moment of the first equation of the BBGKY hierarchy. We have that

$$\begin{aligned}\frac{\partial}{\partial t}\left(\frac{k_BT}{m}\rho\boldsymbol{E} + \frac{\rho}{m^2}\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}\right) =& -\frac{1}{m}\nabla\cdot\boldsymbol{Q}_{111}^{(1)} - 2(\nabla_{\boldsymbol{r}}V^{\text{ext}})\otimes(\rho\bar{\boldsymbol{p}}) &(6.46)\\ &- 2\int_{\mathbb{R}^d}\nabla_{\boldsymbol{r}}v_2\otimes\boldsymbol{u}_1^{(2)}\,\mathrm{d}\boldsymbol{r}_2 - 2\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\nabla_{\boldsymbol{r}}v_3\otimes\boldsymbol{u}_1^{(3)}\,\mathrm{d}\boldsymbol{r}_3\,\mathrm{d}\boldsymbol{r}_2\\ &- 2\gamma\left(\frac{k_BT}{m}\rho\boldsymbol{E} + \frac{\rho}{m^2}\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}\right) + 2\gamma mk_BT\rho\mathbf{1} - \boldsymbol{\mathcal{M}}_2\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right).\end{aligned}$$

Using the correlation approximation we have that for $n\in\mathbb{N}$:

$$\begin{aligned}\boldsymbol{u}_1^{(n)}(\boldsymbol{r},t) =& \int_{\mathbb{R}^{dn}}\boldsymbol{v}_1 f^{(n)}(\boldsymbol{r}^n,\boldsymbol{v}^n,t)\,\mathrm{d}\boldsymbol{v}^n\\ =& \int_{\mathbb{R}^{dn}}\boldsymbol{v}_1 g^{(n)}(\boldsymbol{r}^n,t)\prod_{i=1}^n f^{(1)}(\boldsymbol{r}_i,\boldsymbol{v}_i,t)\,\mathrm{d}\boldsymbol{v}^n\\ =& g^{(n)}(\boldsymbol{r}^n,t)\left[\int_{\mathbb{R}^d}\boldsymbol{v}_1 f^{(1)}(\boldsymbol{r}_1,\boldsymbol{v}_1,t)\,\mathrm{d}\boldsymbol{v}_1\right]\prod_{i=2}^n\left[\int_{\mathbb{R}^d}f^{(1)}(\boldsymbol{r}_i,\boldsymbol{v}_i,t)\,\mathrm{d}\boldsymbol{v}_i\right]\\ =& g^{(n)}(\boldsymbol{r}^n,t)\bar{\boldsymbol{p}}(\boldsymbol{r}_1,t)\prod_{i=1}^n\rho(\boldsymbol{r}_i,t)\\ =& \bar{\boldsymbol{p}}(\boldsymbol{r}_1,t)\rho^{(n)}(\boldsymbol{r}^n,t).\end{aligned}$$

Thus terms in the third moment with interaction potentials can be written as

$$\int_{\mathbb{R}^d}(\nabla_{\boldsymbol{r}_1}v_2(\boldsymbol{r}_1,\boldsymbol{r}_2))\boldsymbol{u}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\,\mathrm{d}\boldsymbol{r}_2 + \int_{\mathbb{R}^d}\int_{\mathbb{R}^d}(\nabla_{\boldsymbol{r}_1}v_3(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3))\boldsymbol{u}_1^{(3)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3)\,\mathrm{d}\boldsymbol{r}_3\,\mathrm{d}\boldsymbol{r}_2 + \ldots$$

$$=$$

$$-\rho(\boldsymbol{r}_1)\bar{\boldsymbol{p}}(\boldsymbol{r}_1,t)\otimes\nabla_{\boldsymbol{r}_1}\frac{\delta\mathcal{F}_{\text{ex}}[\rho(\boldsymbol{r}_1)]}{\delta\rho(\boldsymbol{r}_1)}.$$

By introducing and removing the term $\bar{\boldsymbol{p}}\otimes\nabla_{\boldsymbol{r}}\rho$ we can then include the Helmholtz free energy in

the third moment:

$$\frac{\partial}{\partial t}\left(\frac{k_BT}{m}\rho\boldsymbol{E}+\frac{\rho}{m^2}\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}\right)=-\frac{1}{m}\nabla_{\boldsymbol{r}}\cdot\boldsymbol{Q}^{(1)}_{111}+\frac{k_BT}{m}\nabla_{\boldsymbol{r}}\rho-\frac{2}{m}\rho\bar{\boldsymbol{p}}\otimes\nabla_{\boldsymbol{r}}\frac{\delta\mathcal{F}}{\delta\rho} \tag{6.47}$$
$$-2\gamma\left(\frac{k_BT}{m}\rho\boldsymbol{E}+\frac{\rho}{m^2}\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}\right)+2\gamma mk_BT\rho\mathbf{1}-\boldsymbol{\mathcal{M}}_2\left(\mathcal{L}_{\text{coll}}[f^{(2)}]\right).$$

Furthermore, using the local equilibrium approximation we can show that:

$$\boldsymbol{Q}^{(1)}_{111}(\boldsymbol{r},t)=\frac{k_BT}{m^2}\rho(\boldsymbol{r},t)\left[\bar{\boldsymbol{p}}\otimes\boldsymbol{E}+\boldsymbol{E}\otimes\bar{\boldsymbol{p}}+(\bar{p}_jE_{ik})_{i,j,k=1,\dots d}\right]+\frac{\rho}{m^2}\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}. \tag{6.48}$$

Using Einstein summation notation the divergence of this moment can then be written as

$$\begin{aligned}\partial_i(\boldsymbol{Q}^{(1)}_{111})_{i,j,k}=\frac{k_BT}{m}\Big\{&(\partial_i\rho)[\bar{p}_iE_{jk}+\bar{p}_kE_{ij}+\bar{p}_jE_{ik}]+\rho[(\partial_i\bar{p}_i)E_{jk}+(\partial_i\bar{p}_j)E_{ik}+(\partial_i\bar{p}_jE_{ik})\\&+\bar{p}_i(\partial_iE_{jk})+\bar{p}_k(\partial_iE_{ij})+\bar{p}_j(\partial_iE_{ik})]\Big\}\\+\frac{1}{m^2}\Big\{&(\partial_i\rho)\bar{p}_i\bar{p}_j\bar{p}_k+\rho[(\partial_i\bar{p}_i)\bar{p}_j\bar{p}_k+\bar{p}_i(\partial_i\bar{p}_j)\bar{p}_k+\bar{p}_j\bar{p}_j(\partial_i\bar{p}_k)]\Big\}.\end{aligned}$$

When we consider the time derivative we have that

$$\frac{\partial}{\partial t}\boldsymbol{\Pi}^{(1)}_{11}(\boldsymbol{r},t)=\frac{\partial}{\partial t}\left(\frac{k_BT}{m}\rho\boldsymbol{E}+\frac{\rho}{m^2}\bar{\boldsymbol{p}}\otimes\bar{\boldsymbol{p}}\right).$$

In Einstein summation notation and using eq. (6.39), we find

$$\begin{aligned}\left(\frac{\partial}{\partial t}\boldsymbol{\Pi}^{(1)}_{11}\right)_{j,k}=&\frac{k_BT}{m}\left[\rho\partial_tE_{jk}-\frac{1}{m}E_{jk}[(\partial_i\rho)\bar{p}_i+\rho(\partial_i\bar{p}_i)]\right]\\&+\frac{1}{m^2}\left[\rho[(\partial_t\bar{p}_j)\otimes\bar{p}_k+\bar{p}_j\otimes(\partial_t\bar{p}_k)]-\bar{p}_j\bar{p}_k\frac{1}{m}[(\partial_i\rho)\bar{p}_i+\rho(\partial_i\bar{p}_i)]\right].\end{aligned}$$

Considering the contributions from the time derivative term and the divergence term together we see that some terms cancel:

$$\begin{aligned}\left(\frac{\partial}{\partial t}\boldsymbol{\Pi}^{(1)}_{11}\right)_{j,k}+\frac{1}{m}\partial_i(\boldsymbol{Q}^{(1)}_{111})_{i,j,k}=&\frac{k_bT\rho}{m}\partial_tE_{jk}+\frac{\rho}{m^2}[(\partial_t\bar{p}_j)\bar{p}_k+\bar{p}_j(\partial_t\bar{p}_k)]\\&+\frac{k_BT}{m^2}\Big\{(\partial_i\rho)[\bar{p}_kE_{ij}+\bar{p}_jE_{ik}]+\rho[(\partial_i\bar{p}_j)E_{ik}+(\partial_i\bar{p}_jE_{ik})\\&+\bar{p}_i(\partial_iE_{jk})+\bar{p}_k(\partial_iE_{ij})+\bar{p}_j(\partial_iE_{ik})]\Big\}\\&+\frac{\rho}{m^2}\left\{\bar{p}_i(\partial_i\bar{p}_j)\bar{p}_k+\bar{p}_j\bar{p}_j(\partial_i\bar{p}_k)\right\}.\end{aligned}$$

We then substitute this result into eq. (6.47). By inserting the momentum equation further terms disappear and we are left with an equation explaining the dynamics of the granular temperature

term.

$$\boxed{\partial_t \boldsymbol{E} + \boldsymbol{v}\cdot\nabla_{\boldsymbol{r}} E + (\boldsymbol{E}\nabla_{\boldsymbol{r}}\boldsymbol{v}) + (\boldsymbol{E}\nabla_{\boldsymbol{r}}\boldsymbol{v})^T + 2\gamma(\boldsymbol{E}-\boldsymbol{I}) + \frac{1}{k_B T\rho}\boldsymbol{\mathcal{M}}_2^c(\mathcal{L}_{\text{coll}}[f^{(1)}, g^{(2)}]) = 0,} \tag{6.49}$$

where $\boldsymbol{\mathcal{M}}_2(\mathcal{L}_{\text{coll}}[f^{(1)}, g^{(2)}])$ is the second centred moment of the collision operator:

$$\boldsymbol{\mathcal{M}}_2^c(\mathcal{L}_{\text{coll}}[f^{(1)}, g^{(2)}]) = \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{p} \frac{(\boldsymbol{p}-\bar{\boldsymbol{p}})\otimes(\boldsymbol{p}-\bar{\boldsymbol{p}})}{m^2}\mathcal{L}_{\text{coll}}[f^{(1)}, g^{(2)}]. \tag{6.50}$$

Given expressions or approximations of the two centred moments of the collision operator, and an approximation for the correlation function $g^{(2)}$, eqs. (6.39), (6.42) and (6.49) constitute a closed system of equations to model granular media. Furthermore, the set of equations that have been constructed are all d-dimensional, which reduces the dimension down to a level which is suitable for simulation using standard methods. In the next chapter we will construct $g^{(2)}$ for inelastic hard particles in the presence of external friction to use in computed examples of the DDFT considered here.

6.4.3 Discussion

The system of equations derived here is a new approach to modelling granular media, based on results in density functional theory that are used in colloidal systems. The continuity equation eq. (6.38) states that changes in density over time is determined by variations in the local average velocity. The additional integral term here accounts for changes in the density due to collisional effects. We have already noted that this term is small, and vanishes when $\varepsilon \to 0$.

The momentum equation eq. (6.42) contains several terms of note. Firstly, there is a viscous term $\gamma\boldsymbol{v}$ that shows how frictional effects reduce the local velocity over time. The term involving granular temperature is a new addition to previous DDFTs, *e.g.* eqs. (6.29) and (6.30) [68], as the granular temperature is usually neglected in colloidal flows. This term dictates that local average velocity fluctuations also occur because of changes in the local granular temperature. We note the present of the identity matrix in this term: this is due to dissipative effects as a result of thermal fluctuations in the medium. In the case where particle dynamics are deterministic the identity matrix can be removed.

The next term involves the Helmholtz free energy functional. This term provides the DDFT with the potential to model many-body interactions and volume exclusion effects. By approximating these effects in this term, we are able to draw on any results that are already provided for the free energy functional [149].

Finally, the first moment of the inelastic Boltzmann collision operator is included in the equation. The inclusion of this term ensures that fluctuations in the local average velocity as a result of collisional effects are included. However, we expect the effect of this term to be small, as we know that binary collisions do not change the average velocity between particles. Indeed, in appendix A, this term vanishes as the size of the particle tends to zero, $\varepsilon \to 0$.

Coupled with the momentum equation is the equation that determines the evolution of granular temperature, eq. (6.49). The terms involving spatial derivatives here explain how the granular temperature is affected by itself and the local average velocity. These terms arise as a consequence of the moment closure scheme. The frictional term shows how viscous effects dampen the granular

temperature over time. We note the inclusion of the identity matrix here insures that the granular temperature never dissipates below the level provided by thermal fluctuations. Again, the identity matrix should be removed when considering systems unaffected by thermal fluctuations, but we leave it in here to present the more general result.

The final term in eq. (6.49) introduces changes in the granular temperature, due to collisional effects. We expect this term to have a strong effect on the dynamics of the granular temperature when $\alpha < 1$, as at the microscopic level inelastic collisions reduce the variance in velocity between particles. In appendix A, we show that the second moment of the collision operator is non-zero in the limit as $\varepsilon \to 0$ when $\alpha < 1$, which is as we expected. Rather than approximate this term using arguments from kinetic theory (where by approximating granular media as a non-Newtonian fluid collisional effects are sometimes included by considering the viscosity and strain tensor of the fluid), we can compute a simplified form of the collision operator directly from the Maxwellian assumption. The result is a spherical integral which can be efficiently calculated numerically; the details of this calculation are in appendix A for one and two dimensions.

We note that, under the assumption that $\boldsymbol{E}$ is constant, and collisional effects are ignored, the granular temperature equation disappears, and the momentum equation is identical to the result in [68]. We expect that further generalisations (*i.e.* involving hydrodynamic interactions [69], multiple species [76], or orientable particles [145]) will mimic results already known for colloidal systems, but with the addition of valid collisional terms and the equation for granular temperature.

There are some similarities between the DDFT and results where the kinetic theory of granular flow used is as a stand-alone model for granular media or in two-fluid models, for example in [40], [61], [150], [151]. Firstly the solid phase is described by equations of motion which determine the density, local average velocity, and granular temperature. The density equation is identical in each case (with some additional complexity introduced in multiphase flows), which is to be expected, as most of the dynamics in the system are considered in the momentum equation. Furthermore, in most models there are terms that are used to model collisions effects in the momentum and granular-temperature equations, and in general in the literature there are additional assumptions implemented to simplify the collisional term (for example by taking the limit as particle size vanishes). In our case we construct the analytic form of the collision operator, under the same assumptions that are used to derive the rest of the model.

We note that eqs. (6.39), (6.42) and (6.49) do not include interaction terms between the fluid and solid phase, as in the microscopic dynamics used in the DDFT only one (solid) phase was considered. A multi-species granular DDFT could also be constructed under the same methodology, so that we can include complex interparticle interactions in a two-fluid model for granular media, in a well-studied manner, as in the single-species case.

6.5 Details on maximum entropy closure

One method which we describe in more detail is a closure method in [78], which constructs a closed system of equations for point particles with correlations by using the *maximum entropy principle*. We present the results in detail here. Additional terms will also occur due to collisional effects; we neglect the terms in this discussion, but after considering the half-moment terms in the moments of the BBGKY equations, collisional terms can be included in the resulting equations.

The maximum entropy principle states that the best unbiased distribution function based on the available information is the distribution which maximises statistical entropy:

$$\mathcal{S}^{(k)}[f^{(k)}] = -\frac{1}{k!}\int \mathrm{d}\boldsymbol{r}^k\,\mathrm{d}\boldsymbol{v}^k f^{(k)} \ln\left(\frac{f^{(k)}}{f_0^{(k)}}\right), \tag{6.51}$$

where $f_0^{(k)}$ is some reference distribution. We wish to maximise this with the constraint that the first k moments are specified and are consistent with the maximum entropy constraint. This is then written as a Lagrange functional. The k^{th} maximum entropy distribution then takes the form [78]

$$f_{ME}^{(k)} = f_0^{(k)} \exp\left(\sum_{i=1}^{k}\sum_{n_1=0}^{N_1} \lambda_{n_1}^{(1)}(\boldsymbol{r}_i)\cdot\boldsymbol{v}_i^{n_1}\right) \exp\left(\sum_{i=1}^{k}\sum_{j=i+1}^{K}\sum_{n_1+n_2=0}^{N_2} \lambda_{n_1,n_2}^{(2)}(\boldsymbol{r}_i,\boldsymbol{r}_j)\cdot\boldsymbol{v}_i^{n_1}\boldsymbol{v}_j^{n_2}\right)\times\ldots. \tag{6.52}$$

We now consider different moment closures using maximum entropy. To calculate the Lagrange multipliers, we match the moments of the ME distribution function with the moments we know by evolution equations. We can write the moments in terms of hydrodynamic quantities:

$$\langle f^{(k)}\rangle(\boldsymbol{r}^k,t) = \rho^{(k)}(\boldsymbol{r}^k,t), \tag{6.53}$$

$$\langle \mathcal{P}^{(0,\ldots,0,1,0,\ldots)} f^{(k)}\rangle(\boldsymbol{r}^k,t) = \boldsymbol{u}_i^{(k)}(\boldsymbol{r}^k,t) = \bar{\boldsymbol{p}}_i^{(k)}\rho^{(k)}, \tag{6.54}$$

$$\langle \mathcal{P}^{(0,\ldots,0,1,0,\ldots,0,1,0,\ldots)} f^{(k)}\rangle(\boldsymbol{r}^k,t) = \boldsymbol{\Pi}_{ij}^{(k)}(\boldsymbol{r}^k,t) = \rho^{(k)}\bar{\boldsymbol{p}}_i^{(k)}\otimes\bar{\boldsymbol{p}}_j^{(k)} + \rho^{(k)}\boldsymbol{E}_{ij}^{(k)}, \tag{6.55}$$

where $\rho^{(1)}$ is the local density, $\bar{\boldsymbol{p}}^{(1)}$ is the momentum density (a vector in $\mathbb{R}^d$), and $\boldsymbol{E}^{(1)}$ is the kinetic energy density (a symmetric second rank tensor with arguments in $\mathbb{R}^d$), and $\rho^{(k)}, \bar{\boldsymbol{p}}^{(k)}$ and $\boldsymbol{E}^{(k)}$ are the k^{th}-order equivalents.

6.5.1 4+1 moment closure with Maxwellian reference distribution

For the 4+1 moment closure we consider eqs. (6.21), (6.22) and (6.24), *i.e.* the first four equations of the moments of the first BBGKY hierarchy and the first equation of the moments in the second BBGKY hierarchy. The set of variables that we can determine by the PDEs considered are then $\{\rho^{(1)}, \boldsymbol{u}_1^{(1)}, \rho^{(2)}\}$. We have to approximate $\{\boldsymbol{\Pi}_{11}^{(1)}, \boldsymbol{u}_1^{(2)}, \boldsymbol{u}_2^{(2)}\}$, as well as the partial collision moments $\{\tilde{\boldsymbol{\Pi}}_{11}^{(2)}, \tilde{\boldsymbol{\Pi}}_{12}^{(2)}\}$, by using maximum entropy methods, so that the set of equations can be closed.

Following the result of eq. (6.52), the maximum entropy estimate of the two-particle distribution is given by

$$f_{ME}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{v}_1,\boldsymbol{v}_2,t) = f_0^{(2)}(\boldsymbol{v}_1,\boldsymbol{v}_2)\exp\left(\sum_{n=0}^{1}\lambda_n^{(1)}(\boldsymbol{r}_1)\cdot\boldsymbol{v}_1^{\otimes n} + \sum_{n=0}^{1}\lambda_n^{(1)}(\boldsymbol{r}_2)\cdot\boldsymbol{v}_2^{\otimes n} + \lambda_{00}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\right), \tag{6.56}$$

where the functions denoted by λ are Langrange multipliers that must be determined. We use the

reference distribution

$$f_0^{(2)}(\boldsymbol{v}_1, \boldsymbol{v}_2) = \prod_{i=1}^{2} f_0(\boldsymbol{v}_i), \tag{6.57}$$

$$f_0(\boldsymbol{v}_i) = \frac{1}{(2\pi m k_B T)^{3/2}} \exp\left(-\frac{\boldsymbol{v}_i^2}{2mk_BT}\right). \tag{6.58}$$

This is the *Maxwellian* reference distribution, a well-studied distribution considered in the kinetic theory of gases, and known to be the equilibrium distribution for the one-particle distribution function [139].

Then, by matching the available variables $\{\rho^{(1)}, \boldsymbol{u}^{(1)}, \rho^{(2)}\}$ and considering moments of the maximum entropy estimate, [78] obtains:

$$\lambda_0^{(1)}(\boldsymbol{r}_1) = \ln(\rho^{(1)}(\boldsymbol{r}_1)) - \frac{\bar{\boldsymbol{p}}_1 \cdot \bar{\boldsymbol{p}}_1}{2mk_BT}, \quad \boldsymbol{\lambda}_1^{(1)}(\boldsymbol{r}_1) = \frac{\bar{\boldsymbol{p}}^{(1)}}{mk_BT}, \quad \lambda_{00}^{(2)} = \ln(g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)), \tag{6.59}$$

where $g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = \rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)/(\rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2))$ is the pair correlation function. Then eq. (6.56) becomes

$$f_{ME}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{v}_1, \boldsymbol{v}_2) = \left(\prod_{i=1}^{2} f_{ME}^{(1)}(\boldsymbol{r}_i, \boldsymbol{v}_i)\right) g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2), \tag{6.60}$$

where

$$f_{ME}^{(1)}(\boldsymbol{r}_i, \boldsymbol{v}_i) = \frac{\rho^{(1)}(\boldsymbol{r}_i)}{(2\pi m k_B T)^{3/2}} \exp\left(-\frac{(\boldsymbol{v}_i - \bar{\boldsymbol{p}}(\boldsymbol{r}_i))^2}{2mk_BT}\right). \tag{6.61}$$

Then

$$\boldsymbol{\Pi}_{11,ME}^{(1)}(\boldsymbol{r}_1) = \rho^{(1)}\bar{\boldsymbol{p}}^{(1)}\bar{\boldsymbol{p}}^{(1)} + \rho^{(1)}mk_BT\mathbf{1}, \tag{6.62}$$

$$\begin{aligned}\boldsymbol{u}_{i,ME}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_i)\rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2)g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2), \\ &= \bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_i)\rho^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2).\end{aligned} \tag{6.63}$$

Neglecting the collision term, we arrive at a set of equations for $\rho^{(1)}, \bar{\boldsymbol{p}}^{(1)}$ and the pair correlation function $g^{(2)}$ [78]:

$$\frac{\partial}{\partial t}\rho^{(1)}(\boldsymbol{r}_1, t) = -\frac{1}{m}\nabla_{\boldsymbol{r}_1}(\rho^{(1)}(\boldsymbol{r}_1, t)\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1, t)), \tag{6.64}$$

$$\begin{aligned}\frac{\partial}{\partial t}(\rho^{(1)}(\boldsymbol{r}_1, t)\bar{\boldsymbol{p}}(\boldsymbol{r}_1, t)) &= -\frac{1}{m}\nabla_{\boldsymbol{r}_1} \cdot (\rho^{(1)}(\boldsymbol{r}_1, t)\bar{\boldsymbol{p}}(\boldsymbol{r}_1, t)\bar{\boldsymbol{p}}(\boldsymbol{r}_1, t)) \\ &\quad - k_BT\nabla_{\boldsymbol{r}_1}\rho^{(1)}(\boldsymbol{r}_1, t) - (\nabla_{\boldsymbol{r}_1}V_1)\rho^{(1)}(\boldsymbol{r}_1, t) \\ &\quad - \int d\boldsymbol{r}_2(\nabla_{\boldsymbol{r}_1}V_2(\boldsymbol{r}_1, \boldsymbol{r}_2))\rho^{(1)}(\boldsymbol{r}_1, t)\rho^{(1)}(\boldsymbol{r}_2, t)g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, t),\end{aligned} \tag{6.65}$$

$$\frac{\partial}{\partial t}g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, t) = -\frac{1}{m}\left(\bar{\boldsymbol{p}}(\boldsymbol{r}_1, t) \cdot \nabla_{\boldsymbol{r}_1}g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, t) + \bar{\boldsymbol{p}}(\boldsymbol{r}_2, t) \cdot \nabla_{\boldsymbol{r}_2}g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, t)\right). \tag{6.66}$$

To incorporate collisional effects we must also include the partial collision moments in the equation for $\bar{\boldsymbol{p}}$. These equations can then be used to model granular media. However, we suspect that important collisional effects will be neglected in these equations; as we have noted in example 4,

the effect of inelastic collisions does not change the average velocity in a system of particles, an inelastic collision reduces the variance of particle velocities. Furthermore in the construction of the moments of the BBGKY hierarchy in section 6.2, we saw that terms in the first and second moment of the collision operator were negligible, and it was only in the third moment that terms occurred that could have a significant effect on the dynamics. Thus equations that only model the local average velocity and local density will not represent these effects. We therefore must consider higher-order moment closures.

6.5.2 10+4 moment closure

For a higher-order moment closure we consider eqs. (6.21) to (6.25). Then the variables which we determine by partial differential equations are $\{\rho^{(1)}, \boldsymbol{u}^{(1)}, \boldsymbol{\Pi}_{11}^{(1)}, \rho^{(2)}, \boldsymbol{u}_1^{(2)}\}$ and we must approximate $\{\boldsymbol{\Pi}_{11}^{(2)}, \boldsymbol{\Pi}_{12}^{(2)}, \boldsymbol{Q}_{111}^{(1)}, \rho^{(3)}\}$, as well as the partial collision moments $\{\tilde{\boldsymbol{Q}}_{111}^{(2)}, \tilde{\boldsymbol{Q}}_{112}^{(2)}, \tilde{\boldsymbol{\Pi}}_{11}^{(3)}, \tilde{\boldsymbol{\Pi}}_{12}^{(3)}\}$ to close the system of equations. The derivation of the Lagrange multipliers for $f_{ME}^{(3)}$ is described in the appendix in [78].

The three-body distribution function is approximated using

$$f_{ME}^{(3)}(\boldsymbol{r}^3, \boldsymbol{v}^3, t) = \exp\left(\sum_{n=0}^{2} \boldsymbol{\lambda}_n^{(1)}(\boldsymbol{r}_1) \cdot \boldsymbol{v}_1 + \sum_{n=0}^{2} \boldsymbol{\lambda}_n^{(1)}(\boldsymbol{r}_2) \cdot \boldsymbol{v}_2^{\otimes n} + \sum_{n=0}^{2} \boldsymbol{\lambda}_n^{(1)}(\boldsymbol{r}_3) \cdot \boldsymbol{v}_3^{\otimes n} \right. \tag{6.67}$$
$$\left. + \sum_{n+n'=0}^{1} (\boldsymbol{\lambda}_{nn'}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot \boldsymbol{v}_1^{\otimes n} \otimes \boldsymbol{v}_2^{\otimes n'} + \boldsymbol{\lambda}_{nn'}^{(2)}(\boldsymbol{r}_2, \boldsymbol{r}_3) \cdot \boldsymbol{v}_2^{\otimes n} \otimes \boldsymbol{v}_3^{\otimes n'} + \boldsymbol{\lambda}_{nn'}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_3) \cdot \boldsymbol{v}_1^{\otimes n} \otimes \boldsymbol{v}_3^{\otimes n'}) \right).$$

Also

$$f_{ME}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{v}_1, \boldsymbol{v}_2) = \left(\prod_{i=1}^{2} f_{ME}^{(1)}(\boldsymbol{r}_i, \boldsymbol{v}_i) \right) \exp\left(\lambda_{00}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) + \lambda_{10}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot \boldsymbol{v}_1 + \lambda_{01}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot \boldsymbol{v}_2 \right). \tag{6.68}$$

Then we can rewrite the three-body distribution as

$$f_{ME}^{(3)}(\boldsymbol{r}^3, \boldsymbol{v}^3, t) = \frac{f_{ME}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{v}_1, \boldsymbol{v}_2) f_{ME}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_3, \boldsymbol{v}_1, \boldsymbol{v}_3) f_{ME}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2, \boldsymbol{v}_1, \boldsymbol{v}_2)}{f_{ME}^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_1) f_{ME}^{(1)}(\boldsymbol{r}_2, \boldsymbol{v}_2) f_{ME}^{(1)}(\boldsymbol{r}_3, \boldsymbol{v}_3)}. \tag{6.69}$$

By comparing moments of the maximum entropy function and various density identities in [78], the Lagrange multipliers are then calculated as

$$\begin{aligned}
\lambda_0^{(1)}(\boldsymbol{r}_1) &= \ln\left(\frac{\rho^{(1)}(\boldsymbol{r}_1)}{\det(2\pi \boldsymbol{E}^{(1)}(\boldsymbol{r}_1))^{1/2}} \right) - \frac{1}{2} \bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1) \cdot [\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_1) \cdot \bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1), \\
(\boldsymbol{\lambda}_1^{(1)})_k(\boldsymbol{r}_1) &= ([\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_1) \cdot \bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1))_k, \\
(\boldsymbol{\lambda}_2^{(1)})_{kl}(\boldsymbol{r}_1) &= -\frac{1}{2} [\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_1), \\
\lambda_{00}^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) &= \ln(g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)) - \frac{1}{2} \sum_{i=1}^{2} \left(\bar{\boldsymbol{p}}_i^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot [\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_i) \cdot \bar{\boldsymbol{p}}_i^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \right. \\
&\qquad \left. - \bar{\boldsymbol{p}}_i^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot [\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_i) \cdot \bar{\boldsymbol{p}}_i^{(1)}(\boldsymbol{r}_1, \boldsymbol{r}_2) \right), \\
(\lambda_{10}^{(2)})_k(\boldsymbol{r}_1, \boldsymbol{r}_2) &= ([\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_1) \cdot \Delta \bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2)),
\end{aligned} \tag{6.70}$$

$$(\lambda_{01}^{(2)})_k(\boldsymbol{r}_1,\boldsymbol{r}_2) = ([\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_2)\cdot\Delta\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)), \tag{6.71}$$

where

$$\Delta\bar{\boldsymbol{p}}_n^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \bar{\boldsymbol{p}}_n^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) - \bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_n), \tag{6.72}$$

for $n = 1, 2$. The maximum entropy distributions can then be written as

$$f_{ME}^{(1)}(\boldsymbol{r}_i,\boldsymbol{v}_i) = \frac{\rho^{(1)}(\boldsymbol{r}_i)}{\det(2\pi\boldsymbol{E}^{(1)}(\boldsymbol{r}_i))^{1/2}}\exp\left(-\frac{1}{2}(\boldsymbol{v}^{(1)}-\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1))[\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_1)(\boldsymbol{v}^{(1)}-\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1))\right). \tag{6.73}$$

and

$$f_{ME}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{v}_1,\boldsymbol{v}_2) = \left(\prod_{i=1}^{2}\tilde{f}_{i,ME}^{(1)}(\boldsymbol{r}_i,\boldsymbol{v}_i;\boldsymbol{r}_{j\neq i})\right)g^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2), \tag{6.74}$$

where

$$\tilde{f}_{i,ME}^{(1)}(\boldsymbol{r}_i,\boldsymbol{v}_i;\boldsymbol{r}_{j\neq i}) = \quad = \frac{\rho^{(1)}(\boldsymbol{r}_1)}{\det(2\pi\boldsymbol{E}^{(1)}(\boldsymbol{r}_1))^{1/2}}\exp\left[-\frac{1}{2}(\boldsymbol{v}_1-\bar{\boldsymbol{p}}_1^{(2)})\cdot[\boldsymbol{E}^{(1)}]^{-1}\cdot(\boldsymbol{v}_1-\bar{\boldsymbol{p}}_1^{(2)})\right], \tag{6.75}$$

and $f_{ME}^{(3)}$ is determined by eq. (6.69). Then non-collisional higher-order moments are determined by

$$\boldsymbol{\Pi}_{11,ME}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) + \rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\boldsymbol{E}^{(1)}(\boldsymbol{r}_1), \tag{6.76}$$

$$\boldsymbol{\Pi}_{12,ME}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2) = \rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2), \tag{6.77}$$

$$\boldsymbol{Q}_{111,ME}^{(1)}(\boldsymbol{r}_1) = \rho^{(1)}(\boldsymbol{r}_1)\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1)\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1)\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1) + 3\rho^{(1)}\boldsymbol{E}^{(1)}(\boldsymbol{r}_1)\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1). \tag{6.78}$$

Further,

$$\rho_{ME}^{(3)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3) = \frac{\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\rho^{(2)}(\boldsymbol{r}_2,\boldsymbol{r}_3)\rho^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_3)}{\rho^{(1)}(\boldsymbol{r}_1)\rho^{(1)}(\boldsymbol{r}_2)\rho^{(1)}(\boldsymbol{r}_3)}\delta g^{(1)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3), \tag{6.79}$$

where we write the overall correlation term $\delta g^{(1)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3)$ with the Lagrange multipliers:

$$\begin{aligned}
\delta g^{(3)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3) =& \exp\left(-\frac{1}{2}\boldsymbol{\lambda}_{10}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\cdot(\boldsymbol{\lambda}_2^{(1)}(\boldsymbol{r}_1))^{-1}\cdot\lambda_{10}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_3)\right)\\
&\times\exp\left(-\frac{1}{2}\boldsymbol{\lambda}_{10}^{(2)}(\boldsymbol{r}_2,\boldsymbol{r}_3)\cdot(\boldsymbol{\lambda}_2^{(1)}(\boldsymbol{r}_2))^{-1}\cdot\lambda_{01}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\right)\\
&\times\exp\left(-\frac{1}{2}\boldsymbol{\lambda}_{01}^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_3)\cdot(\boldsymbol{\lambda}_2^{(1)}(\boldsymbol{r}_3))^{-1}\cdot\lambda_{01}^{(2)}(\boldsymbol{r}_2,\boldsymbol{r}_3)\right),\\
=&\exp\left(\Delta\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\cdot[\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_1)\cdot\Delta\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_3)\right)\\
&\times\exp\left(\Delta\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_2,\boldsymbol{r}_3)\cdot[\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_2)\cdot\Delta\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\right)\\
&\times\exp\left(\Delta\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_3)\cdot[\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_3)\cdot\Delta\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_2,\boldsymbol{r}_3)\right).
\end{aligned}$$

For the three-body identities we consider integrals of the form

$$\begin{aligned}\rho^{(3)}_{ME}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3) =& \exp(\lambda^{(2)}_{00}(\boldsymbol{r}_1,\boldsymbol{r}_2)+\lambda^{(2)}_{00}(\boldsymbol{r}_1,\boldsymbol{r}_3)+\lambda^{(2)}_{00}(\boldsymbol{r}_2,\boldsymbol{r}_3)),\\ &\times\int \mathrm{d}\boldsymbol{v}_1 f^{(1)}_{ME}(\boldsymbol{r}_1)\exp(\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_1,\boldsymbol{r}_2)+\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_1,\boldsymbol{r}_2)\cdot\boldsymbol{v}_1)\\ &\times\int \mathrm{d}\boldsymbol{v}_2 f^{(1)}_{ME}(\boldsymbol{r}_2)\exp(\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_2,\boldsymbol{r}_3)+\boldsymbol{\lambda}^{(2)}_{01}(\boldsymbol{r}_1,\boldsymbol{r}_2)\cdot\boldsymbol{v}_2)\\ &\times\int \mathrm{d}\boldsymbol{p}_3 f^{(1)}_{ME}(\boldsymbol{r}_3)\exp(\boldsymbol{\lambda}^{(2)}_{01}(\boldsymbol{r}_1,\boldsymbol{r}_3)+\boldsymbol{\lambda}^{(2)}_{01}(\boldsymbol{r}_2,\boldsymbol{r}_3)\cdot\boldsymbol{p}_3),\end{aligned}$$

(with additional multiples of $\boldsymbol{v}_i$ for higher moments) we use the substitutions

$$\begin{aligned}\tilde{\boldsymbol{v}}_1 &\mapsto \boldsymbol{v}_1+\frac{1}{2}(\boldsymbol{\lambda}^{(1)}_2)^{-1}(\boldsymbol{r}_1)\cdot\left(\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_1,\boldsymbol{r}_2)+\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_1,\boldsymbol{r}_3)\right),\\ \tilde{\boldsymbol{v}}_2 &\mapsto \boldsymbol{v}_2+\frac{1}{2}(\boldsymbol{\lambda}^{(1)}_2)^{-1}(\boldsymbol{r}_2)\cdot\left(\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_2,\boldsymbol{r}_3)+\boldsymbol{\lambda}^{(2)}_{01}(\boldsymbol{r}_1,\boldsymbol{r}_2)\right),\\ \tilde{\boldsymbol{v}}_3 &\mapsto \boldsymbol{v}_3+\frac{1}{2}(\boldsymbol{\lambda}^{(1)}_2)^{-1}(\boldsymbol{r}_3)\cdot\left(\boldsymbol{\lambda}^{(2)}_{01}(\boldsymbol{r}_1,\boldsymbol{r}_3)+\boldsymbol{\lambda}^{(2)}_{01}(\boldsymbol{r}_2,\boldsymbol{r}_3)\right),\end{aligned}$$

where, for example

$$\frac{1}{2}(\boldsymbol{\lambda}^{(1)}_2)^{-1}(\boldsymbol{r}_1)\cdot\left(\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_1,\boldsymbol{r}_2)+\boldsymbol{\lambda}^{(2)}_{10}(\boldsymbol{r}_1,\boldsymbol{r}_3)\right)=\Delta\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_2)+\Delta\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_3).$$

and so

$$\boldsymbol{\Pi}^{(3)}_{11}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3)=((\bar{\boldsymbol{p}}^{(1)})^2(\boldsymbol{r}_1)+\boldsymbol{\sigma}^{(1)}(\boldsymbol{r}_1)+(\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_2)+\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_3))^2)\rho^{(3)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3), \tag{6.80}$$

$$\begin{aligned}\boldsymbol{\Pi}^{(3)}_{12}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3) =& (\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_2)+\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_3)-\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1))\\ &\times(\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_2,\boldsymbol{r}_3)+\bar{\boldsymbol{p}}^{(2)}_2(\boldsymbol{r}_1,\boldsymbol{r}_2)-\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_2))\rho^{(3)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3),\end{aligned} \tag{6.81}$$

$$\begin{aligned}\boldsymbol{\Pi}^{(3)}_{13}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3) =& (\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_2)+\bar{\boldsymbol{p}}^{(2)}_1(\boldsymbol{r}_1,\boldsymbol{r}_3)-\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_1))\\ &\times(\bar{\boldsymbol{p}}^{(2)}_2(\boldsymbol{r}_1,\boldsymbol{r}_3)+\bar{\boldsymbol{p}}^{(2)}_2(\boldsymbol{r}_2,\boldsymbol{r}_3)-\bar{\boldsymbol{p}}^{(1)}(\boldsymbol{r}_3))\rho^{(3)}(\boldsymbol{r}_1,\boldsymbol{r}_2,\boldsymbol{r}_3).\end{aligned} \tag{6.82}$$

We note that, using the product rule and eq. (6.21),

$$\begin{aligned}\frac{\partial}{\partial t}(\rho^{(1)}_1\bar{\boldsymbol{p}}^{(1)}_1) &= \rho^{(1)}\frac{\partial}{\partial t}\bar{\boldsymbol{p}}^{(1)}_1+\bar{\boldsymbol{p}}^{(1)}_1\frac{\partial}{\partial t}\rho^{(1)}_1,\\ &= \rho^{(1)}\frac{\partial}{\partial t}\bar{\boldsymbol{p}}^{(1)}_1-\bar{\boldsymbol{p}}^{(1)}_1[\nabla\cdot(\rho^{(1)}_1\bar{\boldsymbol{p}}^{(1)}_1)],\\ &= \rho^{(1)}\frac{\partial}{\partial t}\bar{\boldsymbol{p}}^{(1)}_1-\rho^{(1)}_1\bar{\boldsymbol{p}}^{(1)}_1[\nabla\cdot\bar{\boldsymbol{p}}^{(1)}_1]-\bar{\boldsymbol{p}}^{(1)}_1(\bar{\boldsymbol{p}}^{(1)}_1\cdot\nabla\rho^{(1)}_1).\end{aligned} \tag{6.83}$$

Also

$$\begin{aligned}\nabla\cdot(\rho^{(1)}_1\bar{\boldsymbol{p}}^{(1)}_1\bar{\boldsymbol{p}}^{(1)}_1) &= \rho[\nabla\cdot(\bar{\boldsymbol{p}}^{(1)}_1\bar{\boldsymbol{p}}^{(1)}_1)]+(\nabla\rho^{(1)}_1)\cdot(\bar{\boldsymbol{p}}^{(1)}_1\bar{\boldsymbol{p}}^{(1)}_1),\\ &= \rho^{(1)}_1[\bar{\boldsymbol{p}}^{(1)}_1(\nabla\bar{\boldsymbol{p}}^{(1)}_1\cdot\bar{\boldsymbol{p}}^{(1)}_1)+\bar{\boldsymbol{p}}^{(1)}_1\cdot\nabla\bar{\boldsymbol{p}}^{(1)}_1]+\bar{\boldsymbol{p}}^{(1)}_1(\bar{\boldsymbol{p}}^{(1)}_1\cdot\nabla\rho^{(1)}_1).\end{aligned} \tag{6.84}$$

Therefore the first two terms in eq. (6.22) introduce a material derivative D/Dt. Further, a similar approach with eq. (6.23) produces a material derivative in $\bar{\boldsymbol{p}}^{(1)}_1\otimes\bar{\boldsymbol{p}}^{(1)}_1+\boldsymbol{E}^{(1)}_1$, and a material derivative with respect to the two-body velocity occurs in eq. (6.25). Finally by writing $\rho^{(2)}(\boldsymbol{r}_i,\boldsymbol{r}_j)=\rho^{(2)}_{ij}=\rho^{(1)}_i\rho^{(1)}_j g_{ij}$ for $i,j=1,2,3$, using the product rule and eq. (6.21), eq. (6.24)

produces an evolution equation for the correlation function g. In all, we obtain the following closed set of equations

$$\frac{\partial}{\partial t}\rho_1^{(1)} = -\nabla_1 \cdot (\rho_1^{(1)}\bar{\boldsymbol{p}}_1^{(1)}), \tag{6.85}$$

$$\begin{aligned}\frac{D}{Dt}\bar{\boldsymbol{p}}_1^{(1)} =& -\nabla_{\boldsymbol{r}-1}\cdot \boldsymbol{E}_1^{(1)} - \boldsymbol{E}_1^{(1)}\cdot\nabla_{\boldsymbol{r}_1}(\ln\rho_1^{(1)}) - \frac{1}{m}(\nabla_{\boldsymbol{r}_1}V^{\text{ext}})\rho_1^{(1)} \\ & - \frac{1}{m}\int \mathrm{d}\boldsymbol{r}_2[\rho_2^{(1)} g_{12}(\nabla_{\boldsymbol{r}_1} v_2(\boldsymbol{r}_1,\boldsymbol{r}_2))],\end{aligned} \tag{6.86}$$

$$\begin{aligned}\frac{D}{Dt}(\bar{\boldsymbol{p}}_1^{(1)}\otimes\bar{\boldsymbol{p}}_1^{(1)} + \boldsymbol{E}_1^{(1)}) =& -\nabla_1\cdot(\boldsymbol{E}_1^{(1)}\bar{\boldsymbol{p}}_1^{(1)}) - 2\boldsymbol{E}_1^{(1)}\bar{\boldsymbol{p}}_1^{(1)}\cdot\nabla_{\boldsymbol{r}_1}(\ln(\rho_1^{(1)})) \\ & - \frac{1}{m}(\nabla_{\boldsymbol{r}_1}V^{\text{ext}})\otimes\bar{\boldsymbol{p}}_1^{(1)} - \frac{1}{m}\int \mathrm{d}\boldsymbol{r}_2[(\nabla_{\boldsymbol{r}_1}v_2(\boldsymbol{r}_1,\boldsymbol{r}_2))\rho_2^{(1)} g_{12}\bar{\boldsymbol{p}}_{1,12}^{(2)}],\end{aligned} \tag{6.87}$$

$$\begin{aligned}\frac{\partial}{\partial t}g_{12} =& -g_{12}\big[\nabla_1\cdot(\Delta\bar{\boldsymbol{p}}_{1,12}) + \nabla_{\boldsymbol{r}_2}\cdot(\Delta\bar{\boldsymbol{p}}_{2,12}) \\ & \qquad + \Delta\bar{\boldsymbol{p}}_{1,12}\cdot\nabla_{\boldsymbol{r}_1}(\ln(\rho_1^{(1)})) + \Delta\bar{\boldsymbol{p}}_{2,12}\cdot\nabla_{\boldsymbol{r}_2}(\ln(\rho_2^{(1)}))\big] \\ & \qquad - \left[\bar{\boldsymbol{p}}_{1,12}\cdot\nabla_{\boldsymbol{r}_1}g_{12} + \bar{\boldsymbol{p}}_{2,12}\cdot\nabla_{\boldsymbol{r}_2}g_{12}\right],\end{aligned} \tag{6.88}$$

$$\begin{aligned}\frac{D_1^{(2)}}{Dt}\bar{\boldsymbol{p}}_{1,12}^{(2)} =& -\nabla_{\boldsymbol{r}_1}\cdot\boldsymbol{E}_1^{(1)} - \boldsymbol{E}_1^{(1)}\cdot\nabla_{\boldsymbol{r}_1}(\ln(\rho_1^{(1)}\rho_2^{(1)}g_{12})) \\ & - \frac{1}{m}(\nabla_{\boldsymbol{r}_1}V^{\text{ext}} + \nabla_{\boldsymbol{r}_2}v_2(\boldsymbol{r}_1,\boldsymbol{r}_2)) \\ & - \frac{1}{m}\int \mathrm{d}\boldsymbol{r}_3(\nabla_1 V_2(\boldsymbol{r}_1,\boldsymbol{r}_3))\rho_3^{(1)}g_{13}g_{23}\delta g_{123},\end{aligned} \tag{6.89}$$

where

$$\begin{aligned}\delta g_{123} =& \exp\left(\Delta\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\cdot[\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_1)\cdot\Delta\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_3)\right) \\ & \times\exp\left(\Delta\bar{\boldsymbol{p}}_1^{(2)}(\boldsymbol{r}_2,\boldsymbol{r}_3)\cdot[\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_2)\cdot\Delta\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_2)\right) \\ & \times\exp\left(\Delta\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_1,\boldsymbol{r}_3)[\boldsymbol{E}^{(1)}]^{-1}(\boldsymbol{r}_3)\cdot\Delta\bar{\boldsymbol{p}}_2^{(2)}(\boldsymbol{r}_2,\boldsymbol{r}_3)\right), \\ \Delta\bar{\boldsymbol{p}}_{i,jk}^{(2)} =& \bar{\boldsymbol{p}}_{i,jk}^{(2)} - \bar{\boldsymbol{p}}_i^{(1)}, \\ \frac{D_1^{(2)}}{Dt}\bar{\boldsymbol{p}}_{1,12}^{(2)} =& \left(\frac{\partial}{\partial t} + \bar{\boldsymbol{p}}_{1,12}^{(2)}\cdot\nabla\right)\bar{\boldsymbol{p}}_{1,12}^{(2)}.\end{aligned}$$

Note that eq. (6.89) has an analogous equation for $\bar{\boldsymbol{p}}_{2,12}^{(2)}$. With the addition of the collisional moments under the maximum entropy results the set of equations has been closed and should be more appropriate as a model for granular media.

The methodology behind this derivation is appealing; by using some physically appropriate assumptions on the entropy of the system, we can derive a closed set of equations to model a system of particles. However, the equations constructed are still of dimension $2d$, and also involve d-dimensional integrals of $3d$-dimensional distributions to resolve collisional terms. When $d \geq 2$ this becomes expensive to simulate accurately. The moment closure scheme considered has provided more insight into the equations of interest (for example by deriving the Maxwellian local equilibrium approximation as a consequence of the closure scheme), but to create a model that can be efficiently simulated we have to include assumptions that reduce the dimensionality further.

Chapter 7

The radial correlation function: Numerical investigation

To simulate granular media using the derived DDFT we need an appropriate form for the correlation function $g^{(2)}$. We will assume that the correlation function is *radial*, *i.e.* that $g^{(2)}(\boldsymbol{r}_1, \boldsymbol{r}_2) = g^{(2)}(\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|)$. For systems of hard spheres with interaction potentials that are also radial, this is a valid assumption in the bulk of the system, *i.e.* away from any boundary. In the collision operator the correlation function is only evaluated at contact where $\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| = \varepsilon$, so we only need to construct $g^{(2)}(\varepsilon)$. However, for more complicated dynamics that involve smooth interaction potentials, we require the correlation function at all points $\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| > \varepsilon$.

Analytic forms for the radial correlation function are difficult to construct in general, results in the literature are limited to particular particle dynamics and domains, for example hard rods in a non-periodic domain with elastic collisions with a constant external field [152], and it is initially unclear how inelasticity will affect the dynamics.

In [75] the value of $g^{(2)}(\varepsilon)$ is considered for hard spheres that collide elastically, at steady state. It is noted that in [153] computer simulations are used to fit parameters for the form

$$g^{(2)}(\varepsilon) = \sum_{i=1}^{d} \frac{c_i \phi^i}{1 - \phi/\phi_c}, \tag{7.1}$$

where $\phi(\boldsymbol{r}, t)$ is the *local packing fraction* of the system:

$$\phi(\boldsymbol{r}, t) = \int_{\mathbb{S}^{d-1}(\boldsymbol{r})} \rho(\boldsymbol{r} + \varepsilon \boldsymbol{n}/2, t)\, \mathrm{d}\boldsymbol{n}, \tag{7.2}$$

and ϕ_c is the *random close packing fraction*; for $d = 1$ $\phi_c = 1$, while for $d = 2$ and $d = 3$, $\phi_c \approx 0.824$ and $\phi_c \approx 0.639$ respectively. We note that for small ε the relationship between ϕ and ρ is well approximated by a linear relationship. In [75] the preferred analytic form of $g^{(2)}$ at contact is given by a modified version of the Carnahan-Starling approximation:

$$g^{(2)}_{\mathrm{CH}'}(\varepsilon, [\phi(\boldsymbol{r}, t)]) = \begin{cases} \frac{1-\phi/2}{(1-\phi)^d}, & 0 \leq \phi \leq \phi_f, \\ \frac{(1-\phi_f/2)}{(1-\phi_f)^d} \frac{\phi_c - \phi_f}{\phi_c - \phi}, & \phi_f \leq \phi \leq \phi_c \end{cases} \tag{7.3}$$

where ϕ_f is called the *freezing packing fraction*, and is defined as a point where the hard sphere fluid experiences the liquid-solid 'phase transition' [154]. The values for the freezing packing fraction are given by $\phi_f \approx 1, 0.69$ and 0.49 for $d = 1, 2, 3$ respectively [75]. The result is an extension of the equilibrium radial correlation function considered in [74], [88] and is a well-founded reliable result for the equilibrium form of $g^{(2)}$. We note that as $\phi \to 0$, $g^{(2)}(\varepsilon) \to 1$, so that for small densities particles are uncorrelated (particles are unlikely to come into contact), and when $\phi \to \phi_c$, $g^{(2)}(\varepsilon) \to \infty$ (particles are all in direct contact). The change in behaviour at the freezing packing fraction is included to account for non-instantaneous interactions that occur at high packing fractions. In one dimension the Carnahan-Starling approximation is equivalent to the Percus approximation discussed in [74]:

$$g_{\mathrm{P}}^{(2)}(\varepsilon, [\phi(\boldsymbol{r}, t)]) = \frac{1}{1 - \phi(\boldsymbol{r}, t)}. \tag{7.4}$$

However the equilibrium form of the radial correlation function for elastic hard spheres is not necessarily appropriate for inelastic hard spheres, as the equilibrium dynamics are strongly affected the energy loss due to collisions. As analytic forms are difficult to construct for inelastic dynamics, we will construct numerical approximations of $g^{(2)}$ at long times on small systems of particles, and use the resulting numerical form of $g^{(2)}$ as a local approximation in DDFT simulations. We therefore consider constructing the radial correlation function using accurate EDPD simulations. In contrast to the methods in [153], where $g^{(2)}(\varepsilon)$ is constructed using equilibrium ensembles of particles, we use dynamic simulations to parametrise the correlation function at long times.

7.1 Constructing the correlation function using EDPD

We will consider the correlation function after long times in one, two, and three dimensions, and construct $g^{(2)}$ for small systems of particles. We are aiming to use our results as a local approximation to the radial correlation function in the bulk of the system. We will therefore construct $g^{(2)}$ on a periodic domain.

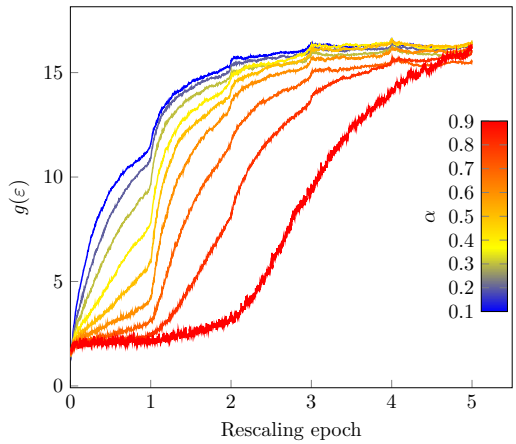
We will use the results of this chapter to approximate the radial correlation function locally, by parametrising in terms of the local packing fraction. We assume that the local approximation is fine enough that the external potential is approximately constant in this area. As relative positions of particles are unaffected by external potentials, we do not need to include an external potential in the microscopic simulations to get an accurate parametrisation.

7.1.1 The homogeneous cooling state

The Carnahan-Starling approximation discussed above approximates the radial correlation function at contact using its steady-state form. The steady state can be reached in systems of elastic particles by running event-driven simulations until the mean-free path of particles stabilises. As we have previously noted, this steady state is not reached in systems of inelastic particles. Instead it is popular to approximate the value of the radial correlation function at contact by the *homogeneous cooling state*, when inelastic collisions between particles have removed all variance in particle velocities, and the value of the radial correlation function stabilizes.

It is known [39] that in one dimension for most values of α (approximately for $\alpha < 0.9$), the value of the radial correlation function at the homogeneous cooling state is independent of the coefficient of restitution: all particles will eventually be moving aside one another at identical velocities. This state is reached in accordance with Haff's law [39], which describes the exponential decay of the granular temperature of the system, due to inelastic collisions.

The homogeneous cooling state is only reached after very long times. We can use long-time dynamics of event-driven simulations to construct $g^{(2)}$ at the homogeneous cooling state. In the case where $\gamma = 0$, to reach this state we need to rescale the velocity variance in the system after it has been reduced by particle collisions. This effectively rescales time in the system, so that we can approach the homogeneous cooling state. In fig. 7.1, we present the results of various one-dimensional simulations, where we rescale the velocity variance in the system after 99% of the variance has been removed due to inelastic collisions. We repeat the rescaling 5 times, and as a result reach the same value for $g^{(2)}(\varepsilon)$ for $\alpha < 0.75$, as expected. We also see that the rate at which systems reach the homogeneous cooling state depends on the choice of α: if α is small then the value of $g^{(2)}$ converges quickly.



(a)

Figure 7.1: The value of $g^{(2)}(\varepsilon)$ over simulation time, in a system of 100 particles, where the total packing fraction is $\phi = 0.5$. We parametrise time by rescaling epochs, when particle velocities are rescaled so that the velocity variance of the system is 1. Between each of these points we present the value of $g(\varepsilon)$ interpolated uniformly over the epoch. We see that over time the value of $g^{(2)}(\varepsilon)$ becomes independent of the value of α.

However, for our purposes the homogeneous cooling state may note be valid. Firstly, when $\gamma > 0$ the system of particles will not reach the homogeneous cooling state, as viscous drag stops all movement before the steady state is reached. Furthermore, the timescales of interest in the continuum model are much shorter than the timescales taken to reach the steady state, so using the value of $g^{(2)}(\varepsilon)$ at this time seems physically inappropriate. Instead, we will construct $g^{(2)}$ at timescales determined by other properties of the system. In short, simulations will run up to a time which is much shorter than the timescale of the continuum model, but much longer than the timescale of collisional dynamics (the mean-free path).

We must then choose what criteria determines we have reached long times in the simulation. There are at least three choices we could consider:

1. We could stop the simulation when we reach some large time T. In general this may not

be possible; for high density systems where many collisions occur, simulations reaching some large time T may take too long to gain accurate statistical information. It is also unclear how large T needs to be for a given system of particles. However, this approach is simple to implement and if simulations are efficient enough very large times should be achievable. Furthermore, for simulations with a non-zero friction coefficient γ long time simulations should be achievable as particles lose energy at an exponential rate. When $\gamma = 0$ it may be appropriate to rescale the system to reach the preferred time T.

2. We can wait until the total variance in particle velocities

$$\mathrm{Var}(\boldsymbol{v}^N) = \frac{1}{N}\sum_{i=1}^{N} \|\boldsymbol{v}_i - \bar{\boldsymbol{v}}\|^2$$

has significantly reduced. This method is valid for when $\alpha < 1$ where velocity variances are reduced due to inelastic collisions. For simulations with friction or gravity, this method will not be valid as the variance in velocity will also scale with the friction coefficient γ. However for systems with linear dynamics this criterion should capture when the system has reached a settled state. We note that if the TC method (discussed in section 3.5.2) has been implemented to mitigate inelastic collapse there will be a small portion of the velocity variance that will not dissipate.

3. We can wait until the energy of the system

$$\mathrm{En}(\boldsymbol{v}^N) = \frac{1}{2}\sum_{i=1}^{N} \|\boldsymbol{v}_i\|^2$$

has significantly reduced. This is valid for when the friction coefficient $\gamma > 0$. However when $\gamma = 0$ the system can equilibrate without losing all its energy; the average velocity of the system will be conserved by inelastic collisions, the total energy of the system will not become arbitrarily small. In this situation we can measure when the total energy of the system stabilises instead. For elastic systems the kinetic energy is fully conserved, so this approach cannot be implemented.

The choice of criterion therefore depends on the dynamics considered. In our results, we use criterion 2 for dynamics without friction and criterion 3 for dynamics with friction. For two and three-dimensional systems, where the use of `DynamO` allows for very efficient simulations, we use criterion 1.

Once the criterion has been met we need to construct the correlation function from the resulting particle positions. To do this we histogram the relative positions of the particles in each sample and average the result.

7.2 Results

7.2.1 Hard rods

To construct the radial correlation function for systems of one-dimensional particles, we use the code discussed in section 3.6. For all the simulations considered, we use $N = 100$ particles and

run at least 5000 samples for each choice of parameters, where particle velocities are sampled from a Gaussian distribution with mean 0 and variance 1, the results are largely independent of the initial scale. This provides enough information to construct the radial correlation function empirically. For initial conditions, we place particles in the domain by sampling their positions from a uniform distribution until the required density is reached, rejecting any samples that produce particle overlap. If the required density is high this method may not be sufficient to stop particles from overlapping; in this situation we place particles uniformly and perform a large number of collisions, with velocities sampled from a Maxwellian distribution, before resetting the velocities and performing the dynamics for the given sample. This insures that the particles are randomly placed in the domain. Alternatively we could sample from a probability distribution of non-overlapping particles.

When considering systems with $\gamma = 0$, we run simulations until the variance in the particle velocities is 0.1% of the initial variance; this is the criterion 2 defined in section 7.1. For $\gamma > 0$ we run the simulation until 0.1% of the energy in the system remains. When $\gamma = 0$ if the dynamics are run for much longer times the distribution would converge to the homogeneous cooling state, but would not effectively characterise the difference between different parameters, would not be suitable for the timescales of interest, and would drastically increase computation times.

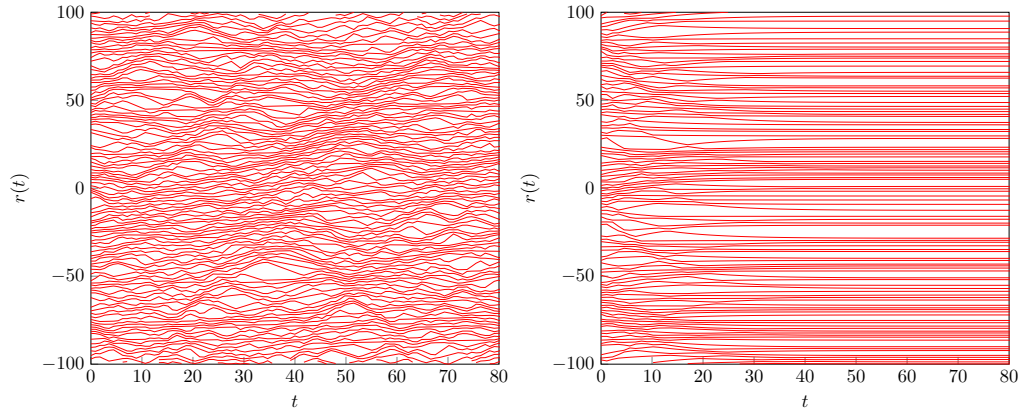
We present some examples of samples constructed in the one-dimensional case with different choices of the parameters γ and α, in fig. 7.2, where we limit the dynamics to a total of 5000 collisions. The results show clear differences in the behaviour of the system, which will be reflected in the correlation function. Firstly, we see that when $\alpha = 1$ and $\gamma = 0$ in fig. 7.2a, as energy is fully conserved, the distribution of positions at any given time is similar. When $\gamma \neq 0$ there is a stark difference in the long-time dynamics, as particles become stationary over time. Furthermore, when we set $\alpha < 1$, particles start to stream together as the local velocity in the system becomes more uniform. We can see in fig. 7.2c that, although small, velocities are still non-zero after a long time, while in fig. 7.2d frictional effects render the particles stationary over time, although particle streaming is still observed.

Furthermore, we note that when $\alpha < 1$ it takes less time for the system of particles to experience 5000 collisions. This is because of particle streaming; after particles collide they remain closer to one another, allowing collisions to take place in shorter timescales.

Polydispersity in hard rods

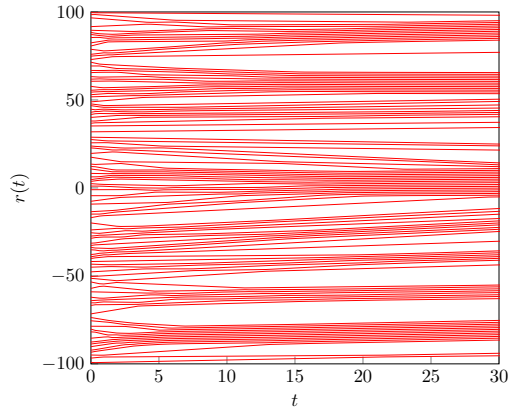
We first note a property of monodisperse inelastic hard rods that makes an empirical construction of the correlation function difficult. When collisions are inelastic in one dimension and the free dynamics are linear, after long times the probability that particles are in contact is very high, leading to a very large value in the correlation function at contact, making it difficult to parametrise the correlation function. This is to be expected; if all particles are in contact with at least one other particle the correlation function will diverge. This is a microscopic analogue to the packing fraction limit of the Carnahan-Starling approximation eq. (7.3).

We see this in fig. 7.3; the plot is on a log-log scale and the value $g^{(2)}(\varepsilon)$ is linearly increasing as $\log(\varepsilon_V)$, the log of the variance in the particle diameter (where particle diameters are sampled from a uniform distribution) decreases. The example considered here uses $\phi = 0.5$, similar results are available for other densities. Thus we expect that the value of $g^{(2)}(\varepsilon)$ diverges in the limit of

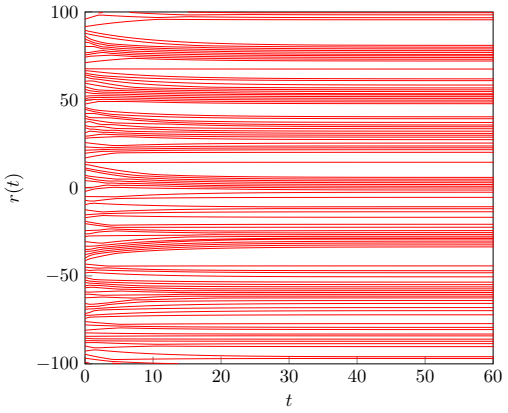


(a) $\gamma = 0, \alpha = 1$

(b) $\gamma = 1, \alpha = 1$



(c) $\gamma = 0, \alpha = 0.5$



(d) $\gamma = 1, \alpha = 0.5$

Figure 7.2: Examples of samples used in the one-dimensional case for different parameters. In each of these examples the packing fraction is set to $\phi = 0.5$, and select different choices of α and γ. For each sample we run the dynamics for a total of 5000 collisions. The results show different behaviour depending on the choice of parameters.

monodispersity, and becomes a Dirac function at the contact point.

This poses a new issue: we cannot produce a parametrised approximation of $g^{(2)}(\varepsilon)$ for monodisperse particles, without involving discontinuities that will lead to precision errors in any continuum models. To construct an approximation for the one-dimensional radial correlation function we therefore include some polydispersity in the system; from now on in one dimension we consider polydisperse particles with variance in diameter given by $\varepsilon_V = \varepsilon/10$. One could then modify the DDFT to account for polydispersity in the collision operator, but we suggest that the monodisperse DDFT in one dimension will be a good approximation for this: it is informally equivalent to using the average result of the collision operator in the polydisperse system.

Equilibrium times for hard rods

When we consider linear dynamics between particles, and inelastic collisions, an interesting question to ask is how long it takes for the stopping criterion provided by criterion 2 to be reached. In fig. 7.4 we present the average time taken for samples to reach this point, given choices of α and ϕ. The results show a complicated and interesting relationship between the coefficient of restitution and

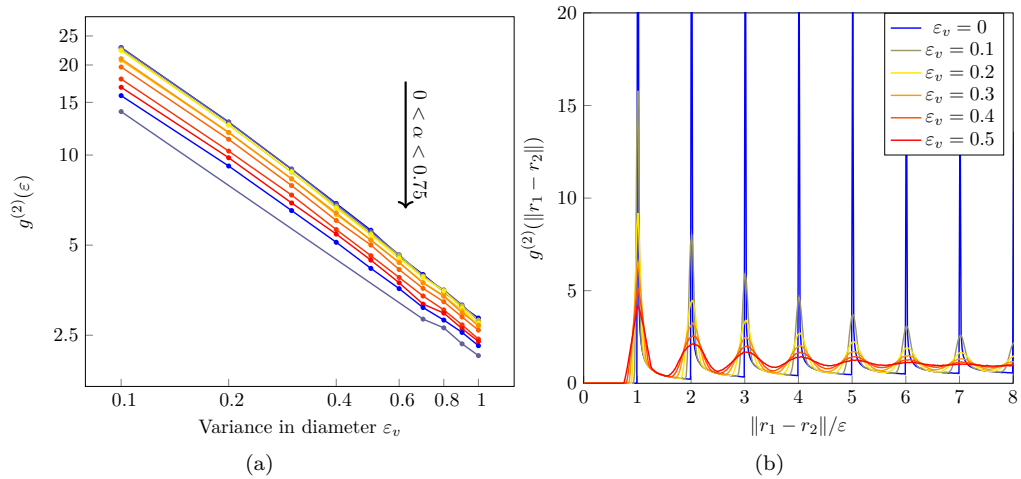


Figure 7.3: The effect of particle diameter on the radial correlation function over long times. Left: $g^{(2)}(\varepsilon)$ for different choices of α when the friction coefficient $\gamma = 0$ and $\phi = 0.5$, when the variance ε_V in the diameter of particles in the system is varied. We see that as $\varepsilon_V \to 0$, the results predict that the value of $g^{(2)}(\varepsilon)$ will diverge. Right: The radial correlation function for $\alpha = 0.5, \phi = 0.5$ and different values of ε_V. Peaks are more pronounced here when ε_V is small.

the time taken. They suggest that there are two regimes where the coefficient of restitution α has a different effect on the system. In fig. 7.5 we compare the time taken for particles to lose 90% of their

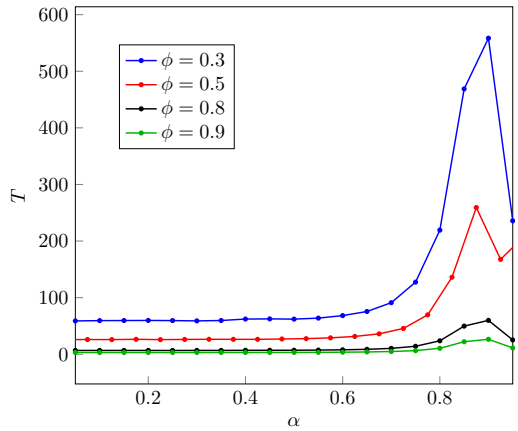


Figure 7.4: The average time taken T for each sample to reach equilibrium, for different values of α, for different packing fractions ϕ. We see two regions of behaviour in the results.

velocity variance, so that we can efficiently simulate systems with coefficients of restitution close to 1. The results further stress that systems close to $\alpha = 1$ display different properties to when α is small. In particular, for $\alpha < 0.7$, the trend of equilibrium slowly increases and accelerates as α increases, but when $\alpha > 0.7$ the average equilibrium time oscillates as α increases. This represents a complicated trade-off between energy loss due to inelastic interactions and increased collisions occurring due to variance in particle velocities.

In fig. 7.6 we construct kernel density estimates of equilibrium times of all samples for choices of α (normalised by the total number of samples), in a system with $\phi = 0.3$. For the case $\alpha = 1$ we consider the result after 10000 collisions has taken place, for reference. When $\alpha < 0.8$ all densities

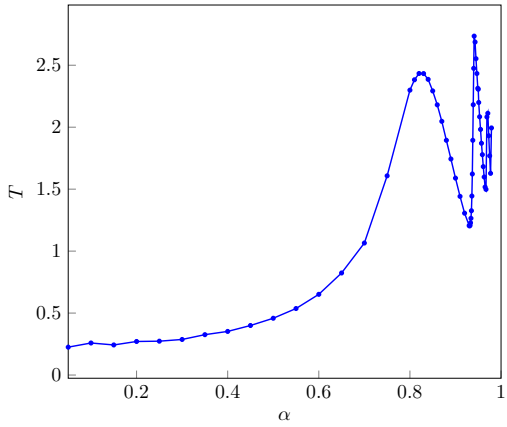


Figure 7.5: The average time taken T for each sample to lose 90% of the particle velocity variance, for different values of α, with $\phi = 0.5$. Considered values of alpha are marked by points. Faster oscillations near $\alpha = 1$ show interesting behaviour due to a trade-off between α and the number of collisions.

are similar to the ones presented, while when $\alpha > 0.8$ the resulting distributions are much wider; the equilibrium times for simulations has very large variance and the number of outliers is small. When $\alpha = 1$ the time taken for 10000 collisions to take place is much narrower with fewer outliers.

Another measurement to consider would be the number of collisions that occur before equilibrium. However, due to the possibility of inelastic collapse and the implementation of the TC method (see section 3.5.2) in our simulations, this is a difficult value to accurately measure; when inelastic collapse occurs an infinite number of collisions is approximated by a large finite number of collisions using the T C method.

In figs. 7.7 and 7.8 we show how the radial correlation function changes over time. We can see that the evolution of the radial correlation function also depends on the dynamics considered. In the case $\gamma = 0, \alpha = 1$ the known equilibrium for elastic hard spheres is reached. For other systems, the presence of the non-zero friction coefficient γ slows the dynamics at an exponential rate, so that the rate at which the correlation function changes is exponential rather than linear. When $\alpha \neq 1$ the radial correlation function develops peaks over time at $n\varepsilon$ for $n \in \mathbb{N}$, showing that particles are clustering due to inelastic collisions, and in this case when $\gamma \neq 0$ the convergence to a stationary function is exponential. We will see that the height of the peaks depends on the coefficient of restitution, the local packing fraction and the friction coefficient.

The correlation function at contact

As previously mentioned the radial correlation function occurs in eqs. (6.39), (6.42) and (6.49) in the collision operator only at the value $\|\boldsymbol{r}_1 - \boldsymbol{r}_2\| = \varepsilon$. We provide the value $g^{(2)}(\varepsilon)$ for different choices of α, γ, and ϕ, when particles have polydispersity with variance $\varepsilon_V = \varepsilon/10$, in fig. 7.9.

The results for hard rods show a lot of interesting dynamical properties, which are also evident in the radial correlation functions in figs. 7.7 and 7.8. The two regimes of high and low coefficient of restitution seen in fig. 7.4 are also evident in figs. 7.9a and 7.10a, where for small values of α, $g^{(2)}(\varepsilon)$ is approximately constant for a given density, while when $\alpha > 0.75$ $g^{(2)}(\varepsilon)$ experiences a rapid change.

We see in figs. 7.9b and 7.10b that for low densities when the coefficient of restitution is low,

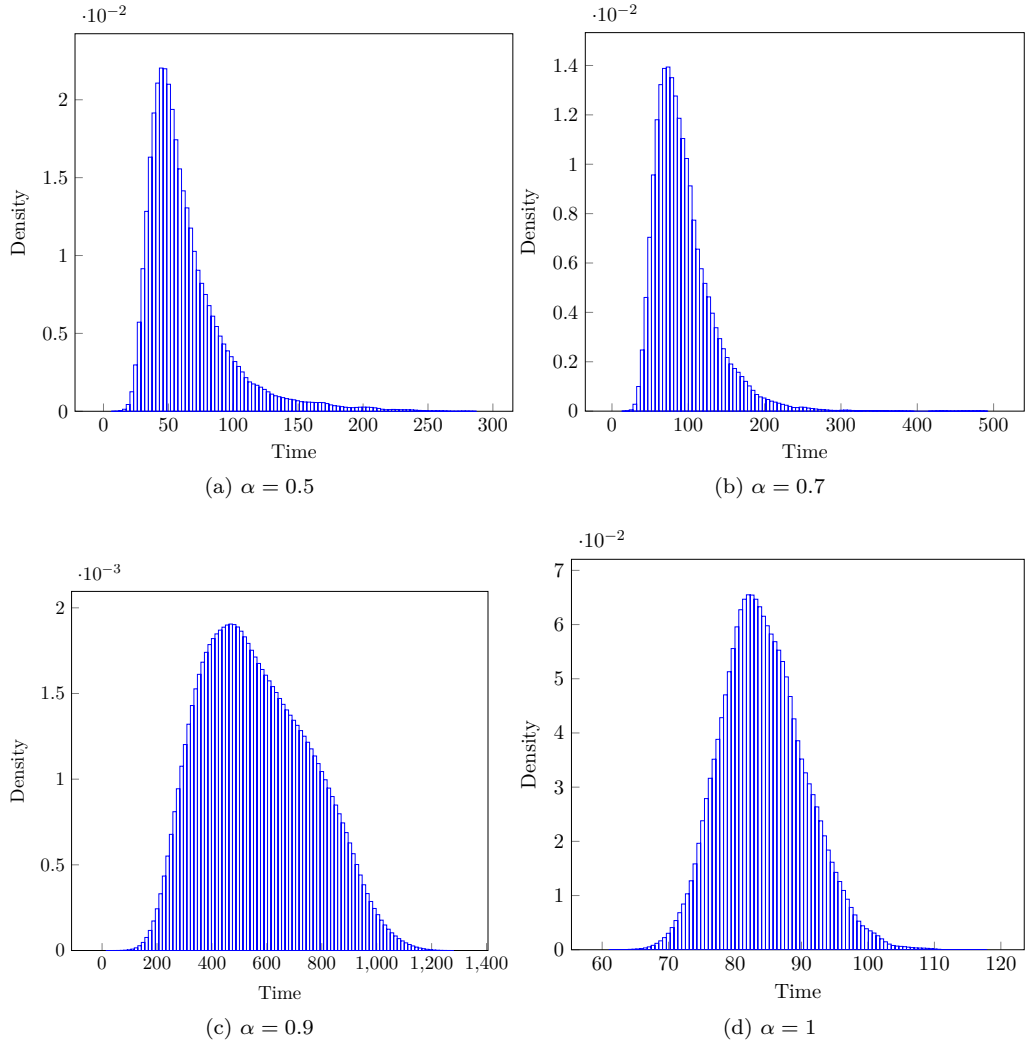
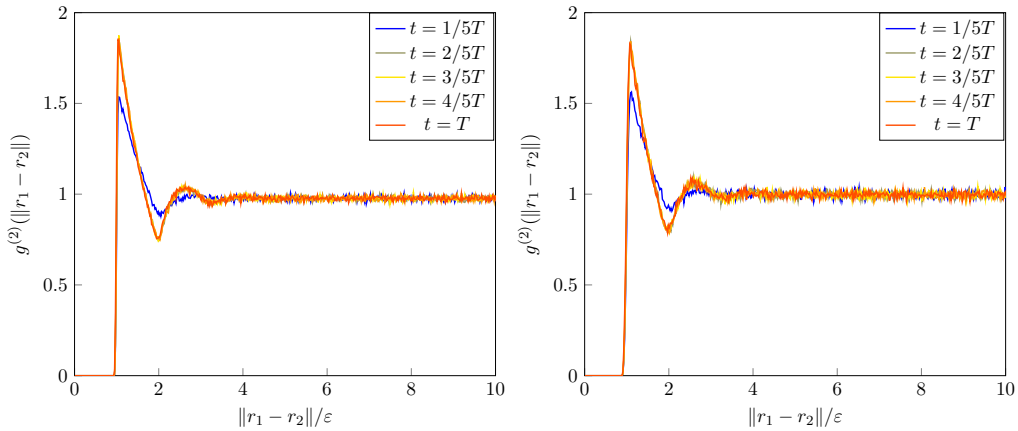


Figure 7.6: Densities of time taken for 99% of the variance in particle velocities to be lost due to inelasticity. Each system has a total packing fraction of $\phi = 0.3$, and we consider different coefficients of restitution α. For small α the result is similar, when α is close to 1, the distribution of times is much wider.

the value of $g^{(2)}$ is also large, which is evidence of clustering or streaming occurring between the particles. We note that the Carnahan-Starling approximation is accurate for low densities in the elastic case, but the result diverges for high densities, where polydispersity causes the close packing fraction limit of $g^{(2)}$ to be finite. We note that the behaviour of $g^{(2)}$ is similar for values of α close to 1 as the elastic case, but quickly changes for $\alpha < 0.9$.

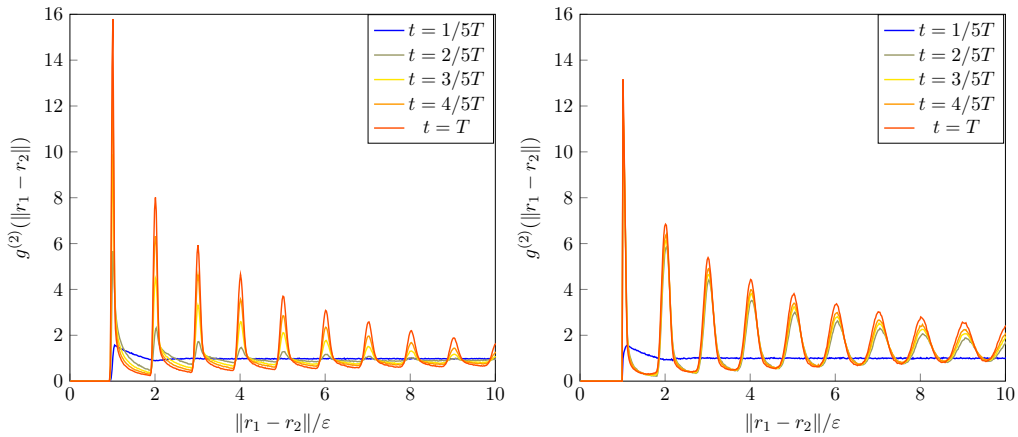
7.2.2 Hard discs

To construct $g^{(2)}(\varepsilon)$ empirically for hard discs (*i.e.* $d = 2$) we use the `DynamO` code library. As previously discussed in section 3.7 the highly optimised algorithms used in `Dynamo` allow for a large number of collisions to be calculated in a small computation time. We can therefore simulate particle dynamics for very long times, so we choose criterion 1 in section 7.1 to determine when the system of particles has stabilised. Furthermore, the code base can also handle much larger



(a) $\gamma = 0$ (b) $\gamma = 1$

Figure 7.7: The evolution of the radial correlation function at different times t for $\alpha = 1$, where total times T are averages over the samples considered. The results show minor changes in $g^{(2)}$, independent of the friction coefficient γ.



(a) $\gamma = 0$ (b) $\gamma = 1$

Figure 7.8: The evolution of the radial correlation function over time for $\alpha = 0.5$, where times presented are averages over the samples considered. In this case large peaks form at contact diameters, when $\gamma > 0$ the height of the peaks is reduced.

systems of particles. In simulations involving hard discs using `DynamO` we use $N = 900$ particles and consider at least 1000 samples for each choice of parameters, where initial particle velocities are taken from a Gaussian distribution with mean 0 and variance 1. To ensure that the initial positions of particles are randomly spaced, we first equally space the particles in the domain, and elastic linear dynamics are run until the mean-free path between particles converges, which implies that the dynamics have reached an equilibrium state for elastic particles. Furthermore, by resampling the velocities of particles after the inelastic dynamics are completed, and rerunning the simulation for the same amount of time, we have seen that the initial positions of particles has little effect on the resulting radial correlation functions.

We note that `DynamO` also has the ability to construct the radial correlation function by histogramming relative positions or particles, accounting for the volume element of the domain. As a

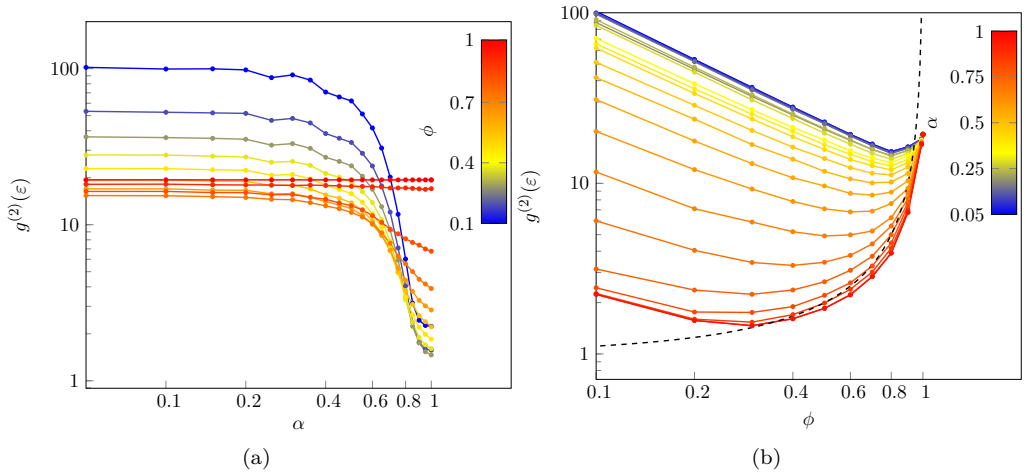


Figure 7.9: Values of the radial correlation function for hard rods at contact, for dynamics without friction ($\gamma = 0$). In fig. 7.9b we include the approximation $g^{(2)}(\varepsilon) = g_{\mathrm{P}}^{(2)}(\varepsilon)$ with a black, dashed line. We see non-trivial behaviour across the parameter values.

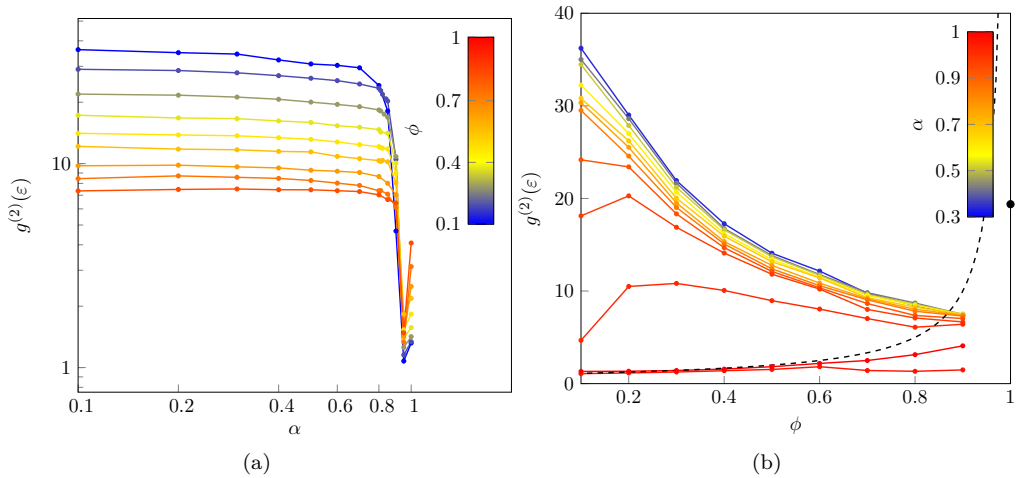
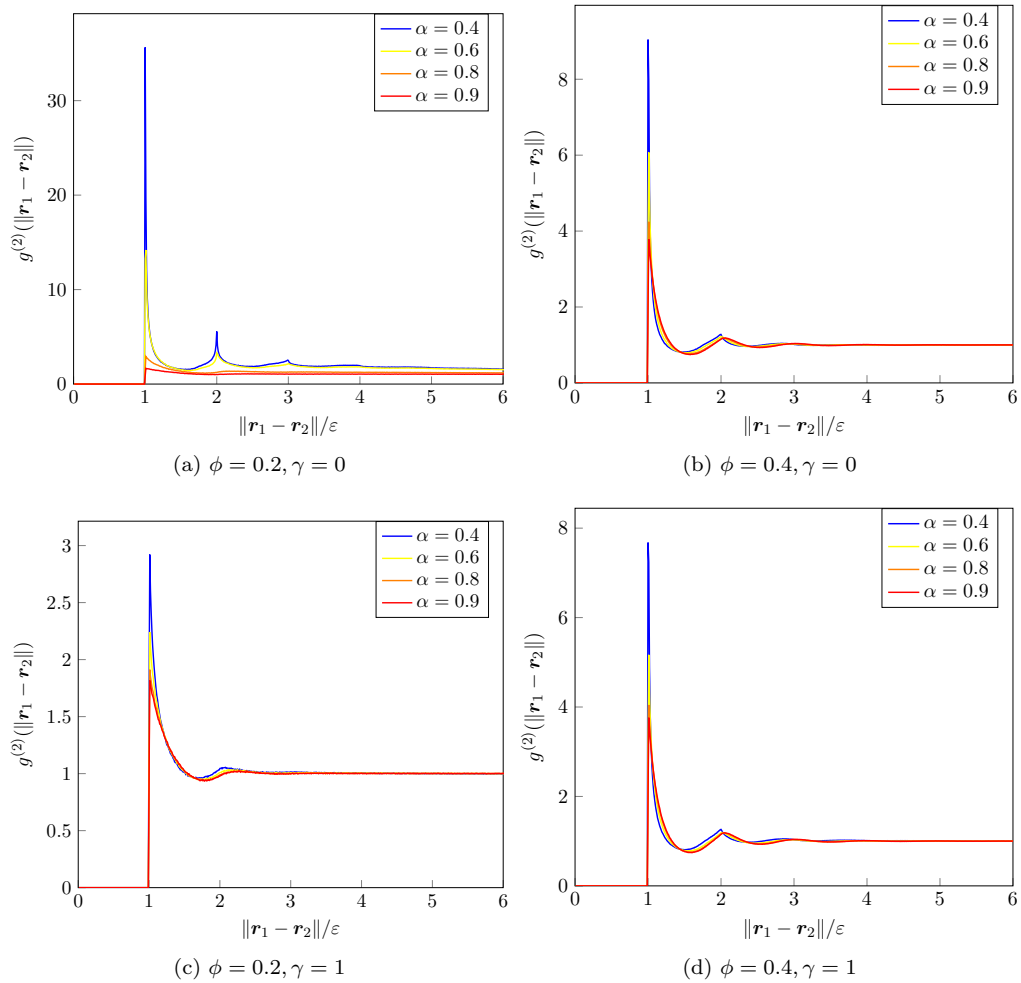


Figure 7.10: Values of the radial correlation function for hard rods at contact, for dynamics with friction ($\gamma = 1$). In fig. 7.9b we include the approximation $g^{(2)}(\varepsilon) = g_{\mathrm{P}}^{(2)}(\varepsilon)$ with a black, dashed line. The value of $g^{(2)}(\varepsilon)$ at $\phi = \phi_c$ is marked, but left unconnected from the other points to improve clarity. When compared to fig. 7.9, results are qualitatively similar, but the presence of friction has reduced the magnitude of results, and changed the behaviour for low densities.

standard, `DynamO` constructs the correlation function for particles in three dimensions, so we must correct the surface element to construct the two-dimensional result.

In fig. 7.11 we present the radial correlation function for different parameters. In comparison to the results in one dimension we see that the peaks in the correlation function are less pronounced and there are fewer peaks; this is because more than one particle can be in contact with an individual particle, so the full diameter of a cluster involving N hard discs will be significantly smaller than the diameter of a cluster involving N hard discs. There is still a difference in results for different choices of α; smaller values for the coefficient of restitution lead to larger peaks. Furthermore, when we include a friction coefficient γ the height of the peaks is reduced as viscous

effects with the external bath stop particles from colliding in long times.



(a) $\phi = 0.2, \gamma = 0$ (b) $\phi = 0.4, \gamma = 0$

(c) $\phi = 0.2, \gamma = 1$ (d) $\phi = 0.4, \gamma = 1$

Figure 7.11: Examples of the radial correlation function for different values of α. As in one dimension, peaks form when low coefficients of restitution are considered, and the height of the peaks is reduced when $\gamma > 0$. The qualitative behaviour of $g^{(2)}$ is different due to the additional degree of freedom.

We can also consider the radial correlation function at contact for hard discs. In fig. 7.12 we look at different values of $g^{(2)}(\varepsilon)$ for choices of α and ϕ. When the coefficient of restitution α is small, and $\gamma > 0$, collisions can occur between particles with very low velocities. In this situation the velocity transfer between particles can be very small, and floating point errors can then lead to the simulation jamming. To reduce the chance of this happening, we can rescale the incoming velocities, process the collisions, then scale back. However for very low densities and small α the issue cannot be resolved by a constant rescaling. We therefore only consider values of α, ϕ which are not affected by machine precision errors.

We see some similarities in the two-dimensional case to the one-dimensional case; the presence of friction reduces the value of $g^{(2)}(\varepsilon)$. However there are some differences to the one-dimensional case; in particular when ϕ is small the value of $g^{(2)}$ is small for all values of α, whereas in the one-dimensional case we witness larger values of $g^{(2)}$ as ϕ decreases. We note that for very low densities when $\gamma = 0$ the value of $g^{(2)}$ reduces, as collisions between particles become very uncommon. For

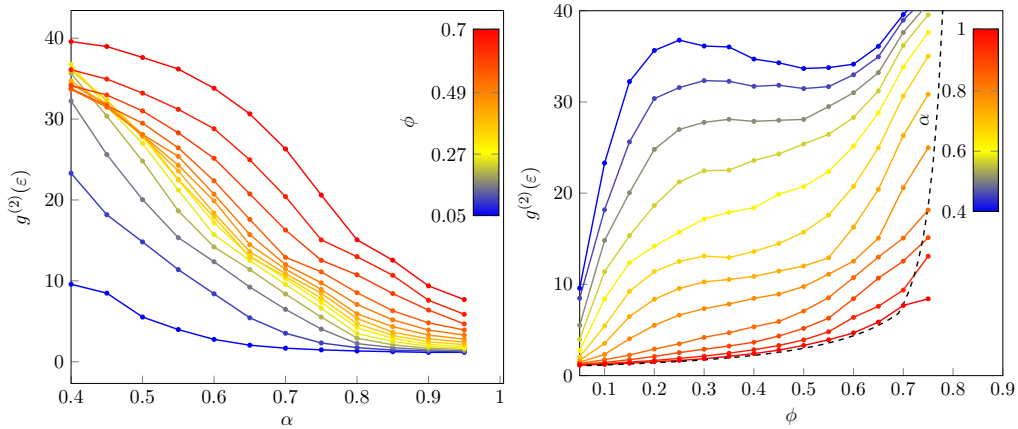


Figure 7.12: Values of the radial correlation function for hard discs at contact, for dynamics without friction ($\gamma = 0$). We include the result for $g^{(2)} = g^{(2)}_{\mathrm{CH}'}$ for hard discs with a black, dashed line. The results show non-trivial relationships between ϕ, α, and $g^{(2)}(\varepsilon)$.

high densities the value of $g^{(2)}$ from simulations differs from $g^{(2)}_{\mathrm{CH}'}$ for all values of α, as EDPD simulations do not account for non-instantaneous interactions that occur above the freezing packing fraction.

We note that the behaviour of the correlation function at contact is similar for $\alpha > 0.75$, but the behaviour changes significantly for lower coefficients of restitution. This is a lower value than in the hard rod case, where the qualitative change in $g^{(2)}$ occurs for around $\alpha < 0.9$. We expect this, as the larger number of degrees of freedom reduce the probability of particle interactions.

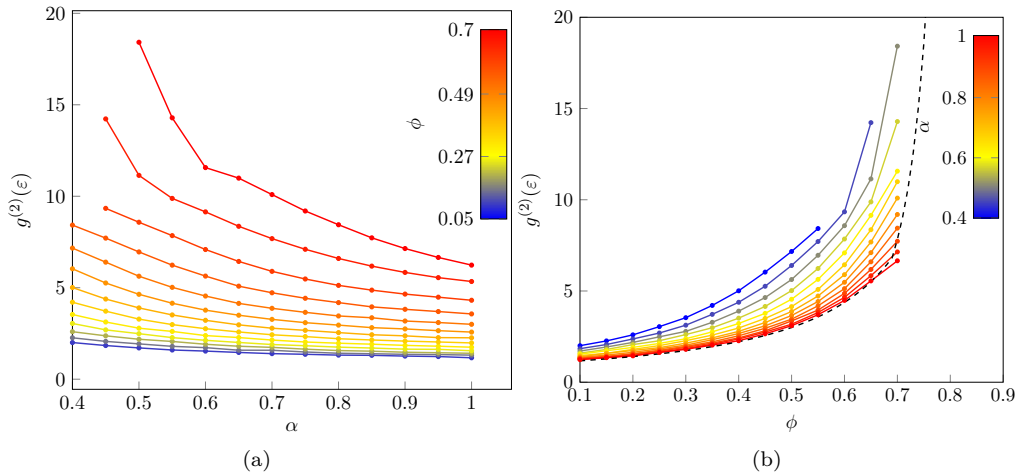
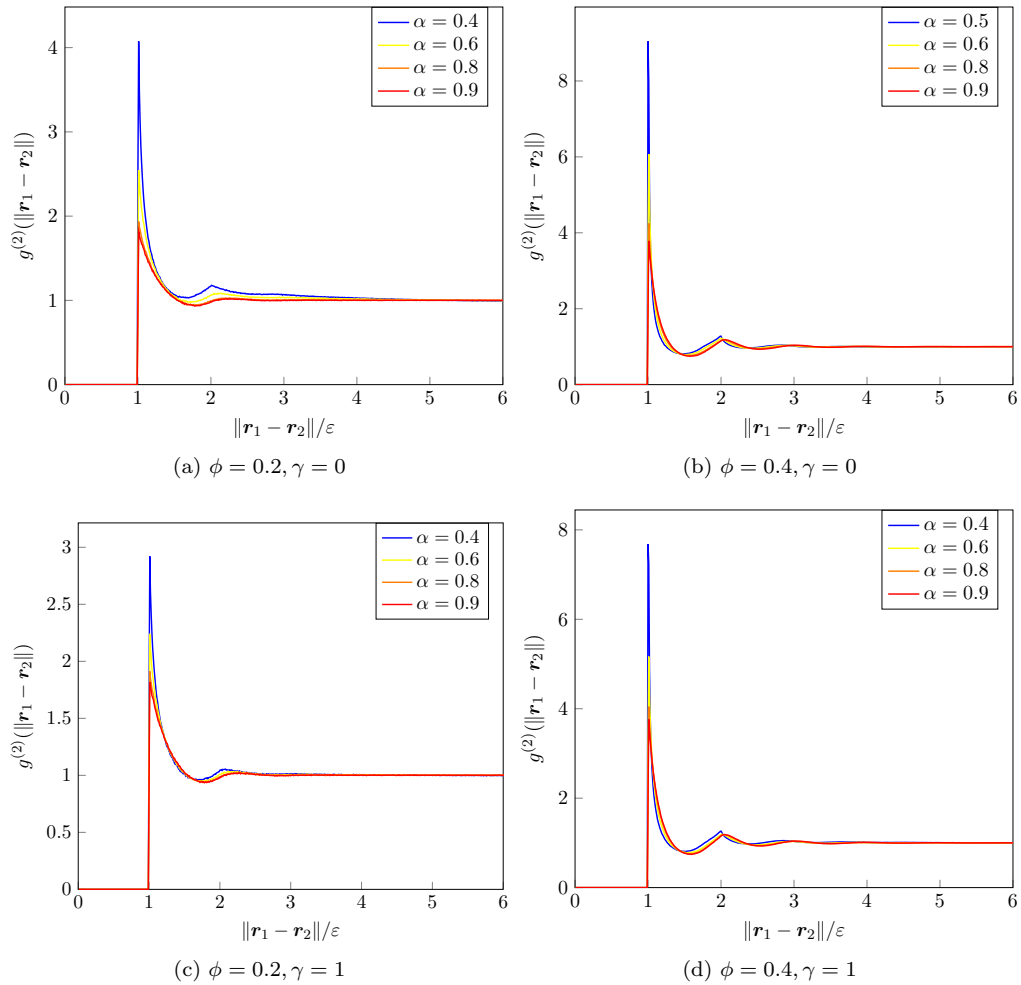


Figure 7.13: Values of the radial correlation function for hard discs at contact, for dynamics with friction ($\gamma = 1$). We include the result for $g^{(2)} = g^{(2)}_{\mathrm{CH}'}$ for hard discs with a black, dashed line. We see that the dependence on parameters ϕ and α is not as strong as when $\gamma = 0$, as expected.

7.2.3 Hard spheres

In three dimensions we also use `DynamO` to construct the radial correlation function, using criterion 1 in section 7.1 as a stopping criterion. We use $N = 1372$ particles and consider at least 1000 samples for each choice of parameters, where velocities are sampled from a Gaussian distribution with mean 0 and variance 1. In fig. 7.14 we present the histograms constructed from samples for different parameters and compare against the known analytic form. The results are qualitatively similar to the two-dimensional case, but there are some small differences in the peak heights. Again, we see that the value of α and ϕ have effects on the heights of peaks in $g^{(2)}(\|\boldsymbol{r}_1 - \boldsymbol{r}_2\|)$.

We note that in this case, when checking whether the initial positions changes the value of the correlation function after long times, we found a mild dependence when $\alpha = 0.4, \phi \geq 0.3$, and $\gamma = 0.5$. Further consideration of the results in these parameters is required to ensure that the parametrisation of $g^{(2)}$ is valid.



(a) $\phi = 0.2, \gamma = 0$ (b) $\phi = 0.4, \gamma = 0$

(c) $\phi = 0.2, \gamma = 1$ (d) $\phi = 0.4, \gamma = 1$

Figure 7.14: Examples of the radial correlation function for different values of α. The overall behaviour is similar to the two-dimensional case.

The radial correlation function at contact for hard spheres is presented in figs. 7.15 and 7.16, for choices of α and ϕ with $\gamma = 0$ and $\gamma = 1$ respectively. The results here are similar to the other cases, for values of α near 1. We note that in the frictionless case there is less variation in

the behaviour of $g^{(2)}(\varepsilon)$ for different α. This is to be expected as the additional spatial dimension reduces the probability of collisions, but we also expect that for lower values of α there would be a change in behaviour in the correlation function at contact. When friction is included the variation in values of $g^{(2)}(\varepsilon)$ is reduced, but we still see different values for choices of α and ϕ. Again, the values of $g^{(2)}(\varepsilon)$ for any α do not match $g^{(2)}_{\mathrm{CH}'}(\varepsilon)$ as EDPD does not account for non-instantaneous interactions at high packing fractions.

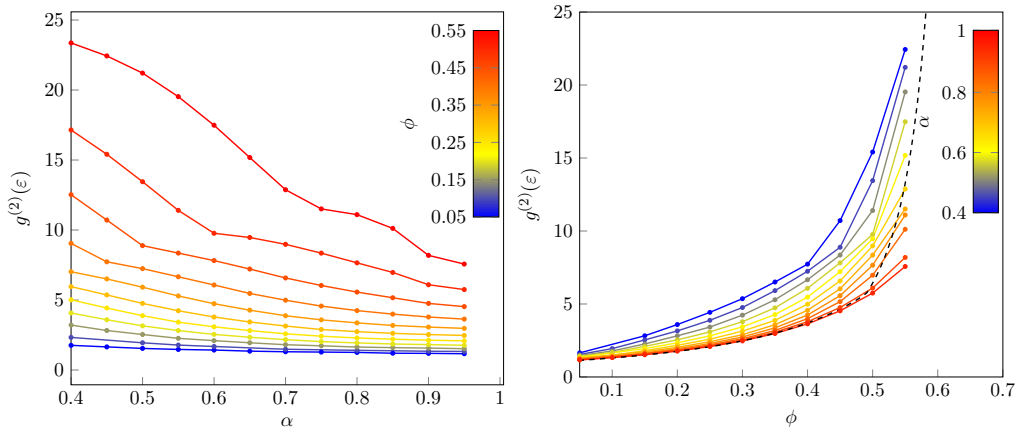


Figure 7.15: Values of the radial correlation function for hard spheres at contact, for dynamics without friction ($\gamma = 0$). We include the result for $g^{(2)} = g^{(2)}_{\mathrm{CH}'}$ for hard spheres with a black, dashed line. We see that α and ϕ both affect the value of $g^{(2)}(\varepsilon)$.

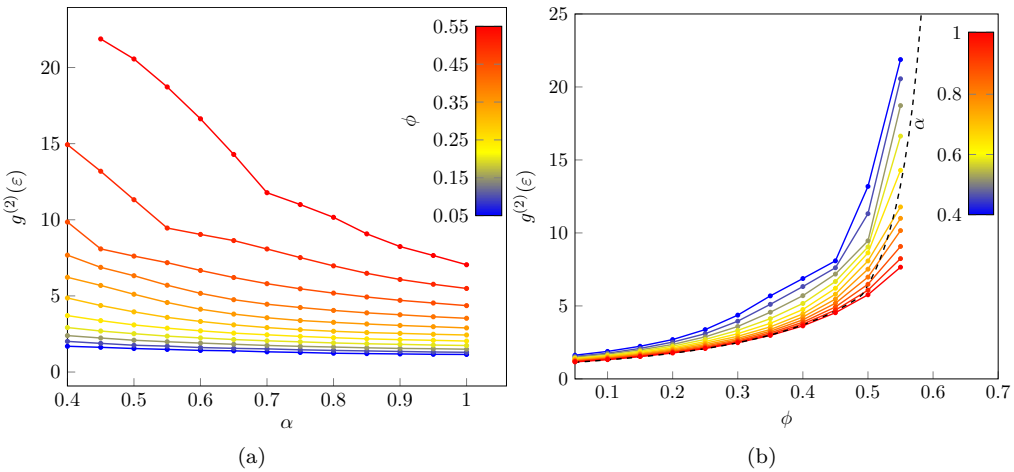


Figure 7.16: Values of the radial correlation function for hard spheres at contact, for dynamics with friction ($\gamma = 1$). We include the result for $g^{(2)} = g^{(2)}_{\mathrm{CH}'}$ for hard spheres with a black, dashed line. Behaviour is similar to the frictionless case, but not as pronounced.

7.3 Discussion

For hard rods, discs, and spheres we have presented a complicated behaviour between the parameters α, γ, ϕ and $g^{(2)}(\varepsilon)$, which encourages us to use a more intricate approximation of the radial correlation function than the Carnahan-Starling approximations, or any other approximations discussed in the start of the chapter. This new insight into the properties of the radial correlation function at relevant timescales for continuum simulation, rather than at the steady state, suggest that these parameters should also have an effect on the macroscopic dynamics. These new insights were achieved by considering the properties of the radial correlation function away from the homogeneous cooling state, but still within a reasonable timescale, unlike most examples considered in the literature, for example in [75].

These results also motivate further investigation of the radial correlation function for different systems of particles, so that we no longer have to rely on steady-state versions of $g^{(2)}$ in systems where this approximation is not relevant. We have restricted our investigation to systems of particles that are only affected by (inelastic) collisions and viscous drag; but `DynamO` allows us to investigate discrete potential systems, and include non-periodic boundaries. In general, we suggest that for the system of interest, quantities such as the radial correlation function should be locally approximated using statistical results from microscopic systems involving a small number of particles over reasonable timescales, by utilising modern computational power and state-of-the-art particle dynamics algorithms. These quantities can then be included in the chosen continuum model for the entire system of particles over longer simulation times.

From the numerical experiments considered, in the next chapter we will construct a parametrisation of the results (using cubic spline interpolation [155]) and use them as an approximation in the continuum model.

Chapter 8

Granular DDFT examples

Now that we have a grasp of the radial correlation function for different densities, friction, and restitution coefficients, we want to simulate eqs. (6.39), (6.42) and (6.49) using parametrised versions of the radial correlation function.

The general form of eqs. (6.39), (6.42) and (6.49) are non-linear integro-differential PDEs, and are non-trivial to implement efficiently. In our experiments we use pseudospectral numerical methods, to accurately simulate the required dynamics.

There are various numerical schemes considered in the literature to test continuum models for granular flows. In particular, (annular) shear flows are used in both theory and experiment to help compare results [156]–[159]. Alternatively, to test models designed to simulate fluidisation, a simplified geometry of a fluidisation chamber is usually considered [61], [62], [160], [161] in fluid-fluid and particle-fluid simulations, to help compare the results to real-world industrial problems.

Due to the infancy of the model considered here, we cannot consider many of the common example systems in the literature. For example, as the system does not account for the interaction between the external fluid and the granular media, fluidisation simulations are not appropriate for the model. Furthermore, as we have not considered how boundaries will affect the radial correlation function, it would be inappropriate to consider numerical experiments involving boundaries, *e.g.* sheared systems of particles, without further investigation in the model derivation.

Therefore in this chapter we will consider some fundamental cases to test the derived continuum model. We use these simple examples to examine the effect of different terms in the DDFT and test the theory we have considered. The resulting qualitative features show that the system of equations captures expected effects of inelastic collisional dynamics.

8.1 Implementation

Pseudospectral methods encompass spectral and collocation methods, which can be used to efficiently calculate derivatives and integrals with high precision, under the assumption that the system satisfies smoothness criteria. We choose to use pseudospectral methods because of their potential to simulate integro-differential partial differential equations, and because there is open-source `MATLAB` code that is already available for DDFT simulation. We provide a brief introduction to methods used for the examples considered here, primarily based on work in [162]. There are three

main calculations that are required to solve the equations of interest; differentiation, integration and interpolation.

8.1.1 Differentiation

To simulate the DDFT equations we need to be able to accurately approximate spatial derivatives. We will consider the construction of a spectral differentiation matrix on a periodic domain. Spectral differentiation methods can be motivated by finite difference schemes. We discretize the domain $[0, 1]$ into N points, and define $t_0 = 0$, $t_{i+1} = t_i + \Delta t$ points in the domain, where $\Delta t = 1/N$, so that $t_0 = r_N$. Given a function u which has values $u(t_i)$ at points t_i, we can approximate the derivative $\mathrm{d}u(t)/\,\mathrm{d}t$ at points t_i using a finite difference approach:

$$\left.\frac{\mathrm{d}u(t)}{\mathrm{d}t}\right|_{t=t_i} \approx \frac{u(t_{i+1}) - u(t_{i-1})}{2\Delta t} =: w_i. \tag{8.1}$$

In matrix form this can be written as

$$\begin{pmatrix} w_1 \\ w_2 \\ \\ \vdots \\ \\ w_N \end{pmatrix} = \frac{1}{\Delta t} \begin{pmatrix} 0 & \frac{1}{2} & & & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \ddots & & \\ & & \ddots & & \\ & & \ddots & 0 & \frac{1}{2} \\ \frac{1}{2} & & & -\frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} u(t_1) \\ u(t_2) \\ \\ \vdots \\ \\ u(t_N) \end{pmatrix}. \tag{8.2}$$

The resulting approximation is $\mathcal{O}(\Delta t^2)$, and by including more points in the finite difference approximation higher-order finite difference schemes can be constructed, which can also be represented as matrix products. However, higher-order matrices are also more dense, making computation more expensive.

Intuitively spectral differentiation matrices can be seen as a sort of 'limit' of finite difference matrices; they are fully dense and have fast convergence rates. Due to their high accuracy fewer discretization points are necessary, so although the matrices considered are dense the calculations can be very efficient.

As described in [162], to construct a spectral differentiation matrix, one considers an analytic function p which is identical to u at the discrete points t_i, then you approximate u' with the derivative of p. By considering appropriate functions p for the periodic domain (trigonometric polynomials, inspired by applications of the discrete Fourier transform), the following spectral

differentiation matrix is constructed for the periodic domain (where N is chosen to be even):

$$D_N = \begin{pmatrix} & \vdots & \\ \ddots & \frac{1}{2}\cot\left(\frac{3\Delta t}{2}\right) & \\ \ddots & -\frac{1}{2}\cot\left(\frac{2\Delta t}{2}\right) & \\ \ddots & \frac{1}{2}\cot\left(\frac{\Delta t}{2}\right) & \\ & 0 & \\ & -\frac{1}{2}\cot\left(\frac{\Delta t}{2}\right) & \ddots \\ & \frac{1}{2}\cot\left(\frac{2\Delta t}{2}\right) & \ddots \\ & -\frac{1}{2}\cot\left(\frac{3\Delta t}{2}\right) & \ddots \\ & \vdots & \end{pmatrix}. \tag{8.3}$$

Using this differentiation matrix results in an approximation which is $\mathcal{O}(N^{-m})$, where m depends on the smoothness of the solution. We therefore need very few points to accurately calculate spatial derivatives in one dimension, provided the solution is suitably smooth. Generally to display results one can then make use of polynomial interpolation. Higher-order differentiation matrices can also be constructed by considering higher-order derivatives of the function p.

8.1.2 Interpolation

For the free energy term and collisional terms in eqs. (6.39), (6.42) and (6.49) we are interested in quantities that have been shifted *e.g.* $\rho(r+\varepsilon)$ in one dimension. Numerically we want to interpolate ρ onto points $r+\varepsilon$, given the points $\rho(r_i), r_i = 2\pi\frac{i}{N}$. We can do this by constructing an interpolation matrix. In general given a polynomial

$$p(r) = \sum_{j=0}^{N} a_j r^j \tag{8.4}$$

to construct the values of p at points $x_i, i = 1, \ldots, M$, we can apply the Vandermonde matrix [163] to the polynomial coefficients:

$$\begin{pmatrix} p(x_0) \\ p(x_1) \\ \vdots \\ p(x_N) \end{pmatrix} = \begin{pmatrix} x_0^N & x_0^{N-1} & x_0^{N-2} & \ldots & x_0 & 1 \\ x_1^N & x_1^{N-1} & x_1^{N-2} & \ldots & x_1 & 1 \\ \vdots & \vdots & \vdots & & \vdots & \vdots \\ x_M^N & x_M^{N-1} & x_M^{N-2} & \ldots & x_M & 1 \end{pmatrix} \begin{pmatrix} a_N \\ a_{N-1} \\ \vdots \\ a_0 \end{pmatrix}. \tag{8.5}$$

In the case of a Fourier domain the appropriate approximation [162] is a linear combination of periodic sinc functions:

$$p(x) = \sum_{j=1}^{N} a_j S_N(x - r_j), \tag{8.6}$$

where

$$S_N(x) = \frac{\sin(\pi x/\Delta t)}{(2\pi/\Delta t)\tan(x/2)} \tag{8.7}$$

and r_j are the equispaced points in the periodic domain, where the value of the function is known and given by a_j. The Vandermonde matrix is then given by

$$P_N(\boldsymbol{x}) = \begin{pmatrix} S_N(x_0 - r_N) & S_N(x_0 - r_{N-1}) & \dots & S_N(x_0 - r_1) & S_N(x_0 - r_0) \\ S_N(x_1 - r_N) & S_N(x_1 - r_{N-1}) & \dots & S_N(x_1 - r_1) & S_N(x_1 - r_0) \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ S_N(x_M - r_N) & S_N(x_M - r_{N-1}) & \dots & S_N(x_M - r_1) & S_N(x_M - r_0) \end{pmatrix}. \tag{8.8}$$

We can then use this matrix to translate functions, as is required in the collisional terms in eq. (6.49); $P_N(\boldsymbol{x}) \cdot \boldsymbol{a}$ maps points $\boldsymbol{x}$ to points on the trigonometric polynomial p. Alternatively to evaluating $S_N(x_i - r_j)$ for each i, j, by using the fast Fourier transform we can construct an interpolation matrix to apply to the vector of computation points [162].

8.1.3 Integration

Finally, in free energy terms in one dimension, and free energy and collisional terms in higher dimensions, we require accurate and efficient integration techniques. Following the same methodology as for differentiation, integration can be constructed on a periodic domain by interpolating using trigonometric polynomials, and integrating the interpolant analytically. This results in a vector of weights, which approximate the integral using the inner product:

$$\int_0^{2\pi} u(x)\,\mathrm{d}x \approx \frac{2\pi}{N}\begin{pmatrix} 1 & 1 & \dots & 1 \end{pmatrix}\begin{pmatrix} u(x_0) \\ u(x_1) \\ \vdots \\ u(x_N) \end{pmatrix}, \tag{8.9}$$

i.e. for periodic domains the weights are all 1.

In general we may want to construct local integrals for the DDFT, for example the functional derivative of the Percus free energy functional is given by:

$$\frac{\delta \mathcal{F}_{\mathrm{ex}}[\rho]}{\delta \rho}(r) = \log\left(1 - \int_r^{r+\varepsilon} \rho(r')\,\mathrm{d}r'\right) + \int_{r-\varepsilon}^{r} \frac{\rho(r'')}{1 - \int_{r''}^{r''+\varepsilon} \rho(r')\,\mathrm{d}r'}\,\mathrm{d}r''. \tag{8.10}$$

Thus we must consider integrals over non-periodic intervals. For this instead of Fourier points we must use *Chebyshev* points to construct the integration interval; these are non-uniform points in the domain that are suitable for differentiation and integration. Appropriate weights can be constructed using the same methodology as the periodic case; we refer the reader to [162] for details.

8.1.4 Timestepping

Finally, once we have accurate methods to approximate the integrals and spatial derivatives in the DDFT, we need a numerical method to perform a timestepping scheme. For this we use `MATLAB's` built in ODE solvers; in particular we use `ode15s`, as we expect the ODEs considered to be stiff. As an alternative we could also apply `ode45s`; the additional benefit of `ode15s` is that it can solve algebraic-differential equations to solve dynamics with boundaries, which is not required on a periodic domain.

8.1.5 Multiple dimensions

For simulations with more than one spatial dimension, we may use Kronecker products of differentiation matrices in each dimension. For example; to compute the Laplacian of a function $u(x, y)$, we construct the differentiation matrices D_x^2, D_y^2 on the correct discretization points x_i, y_i using the one-dimensional methods above, then construct the matrix

$$L_N = D_x^2 \otimes I + I \otimes D_y^2 \tag{8.11}$$

and perform the matrix product on the vectorised matrix $u(x_i, y_j), i, j = 1, \dots N$. Using the tensor product we can then construct domains that are periodic in both dimensions (which we consider for simulation in section 8.3.1), as well as Fourier-spectral domains, which are periodic in one dimension and bounded in the other. These domains are used in our simulations for volume exclusion terms via FMT, where integrals over surfaces (circles) and volumes (discs) are required in two dimensions. We note that the derivation of fundamental measure theory is motivated by ideas of the Percus free energy functional extended to multiple dimensions, we refer to [149] for more details.

8.1.6 Extensions

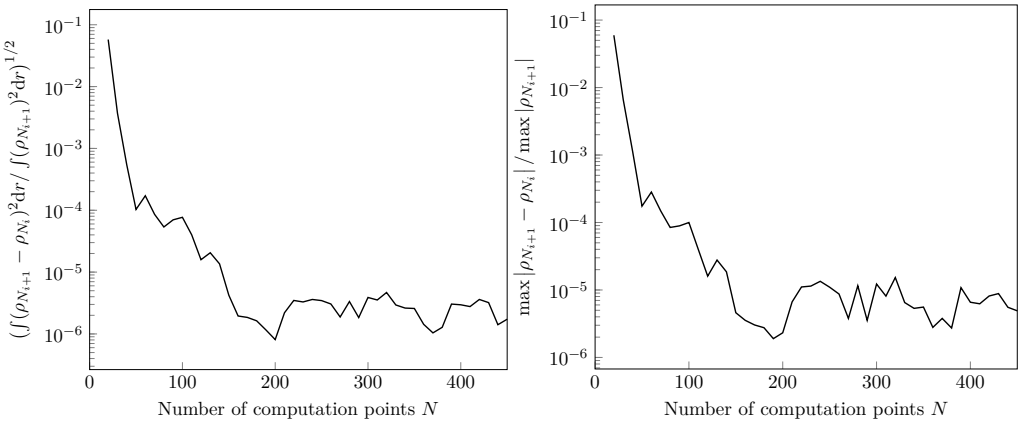
We have mentioned that for non-periodic domains differentiation matrices and integration vectors can be constructed using the same methodology. Chebychev points are then the suitable discretization for bounded intervals, but pseudospectral methods can also be used to discretize infinite and half-infinite lines, and polar domains. For more complicated domains in two and three dimensions boundary patching methods can be used to patch periodic and bounded domains together to create non-trivial shapes. For more information on these methods, we refer the reader to [164] and [162].

All of the methods discussed here are implemented [165] and available in the `2dChebClass` code library [80]. For the examples we have considered we have augmented this library to consider periodic domains, and have include collisional operators and volume exclusion energy functionals for the domains of interest.

8.2 One dimension

In one dimension we consider two examples. Unless otherwise stated, we perform the dynamics on a grid of 150 computational points. The results are then interpolated onto a finer grid of 2000 interpolation points before being presented. Performing the same dynamics with more computational

points produced similar results. In fig. 8.1 we provide estimates of the relative L^2 and L^∞ errors in the density ρ for using more computational points for the dynamics considered in section 8.2.1. We see that for more that 150 computational points the error is at of order $\mathcal{O}(10^{-6})$, which is the chosen tolerance for the ODE solver. Thus 150 computational points is enough to ensure that the results are accurate.



(a) (b)

Figure 8.1: Relative L^2 (left) and L^∞ (right) errors in the density ρ for the dynamics considered in section 8.2.1. We see that choosing 150 computational points is sufficient to guarantee reasonable accuracy.

8.2.1 Example: Gaussian bump

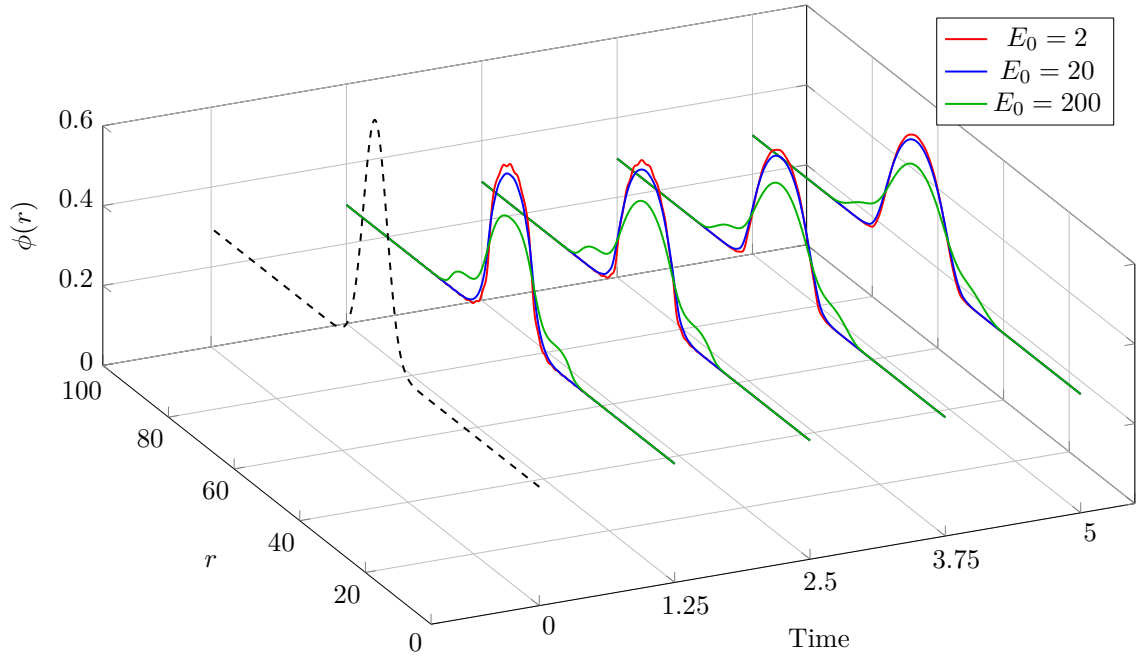
For our first example we consider a periodic domain $[0, 100]$ with the initial conditions:

$$\rho_0(r) = \frac{\phi_v}{N_v} \exp\left(-\frac{(r-50)^2}{25}\right) + 0.5, \quad v_0(r) = 0, \quad E_0(r) = E_0 \tag{8.12}$$

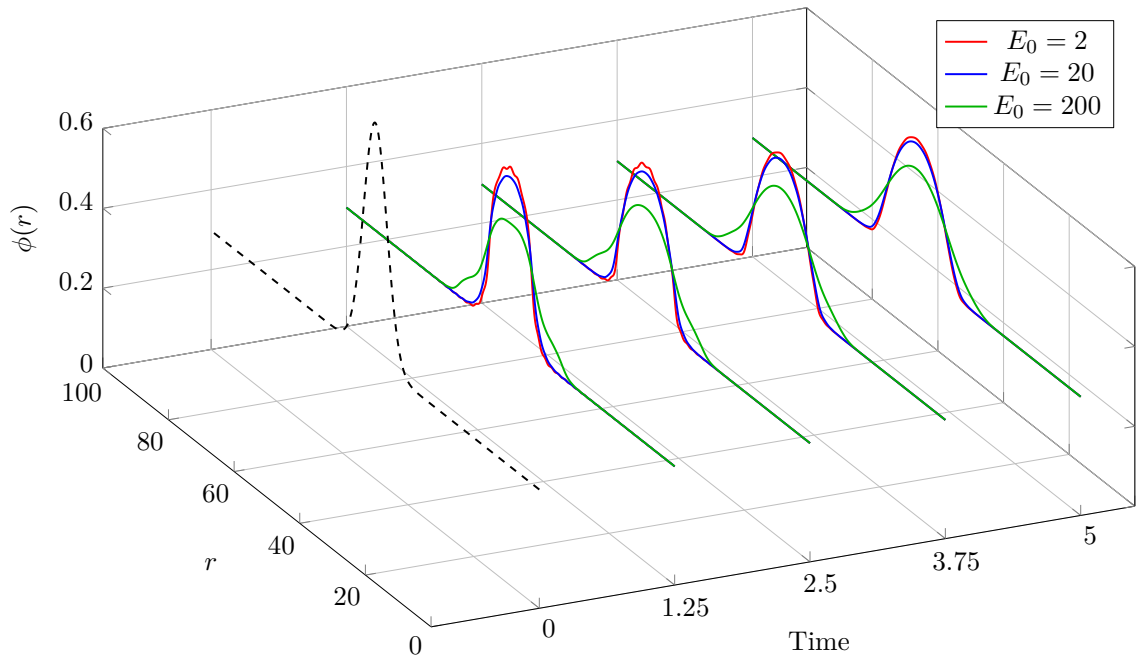
where $E_0 > 0$, $\phi_v = 0.35$ is the total packing fraction of the system and N_v is a normalisation constant that insures that the density integrates to N. We consider a system of $N = 140000$ particles, so that the fully packed system would contain 400000 particles of diameter $\varepsilon = 2.5 \times 10^{-4}$. We use this example to consider the effect of the quantity $E(r, t)$ in the set of equations eqs. (6.39), (6.42) and (6.49). Initially we consider no external potential, and we include the Percus free energy functional to ensure particles cannot overlap. We also include the collision operator, with $g^{(2)}$ parametrised by numerical results presented in chapter 7.

In fig. 8.2a we consider different choices of E_0 for $\alpha = 1$, and in fig. 8.2b we consider different choices of E_0 for $\alpha = 0.5$. We see that there is a considerable difference in the dynamics of the system when large values of E_0 are considered, independent of the choice of α. The effect of inelasticity is more obvious when $E_0 = 200$, where bumps that occur in the density where large changes were present in the initial condition are smoother.

We can also introduce an external potential V^{ext} to the system. In fig. 8.3 we include a constant external force $V^{\text{ext}} = V_0 r$, for the two cases $V_0 = 0$ and $V_0 = -5$. We note that introducing a constant external potential has a translational effect but does not change the relative positions of particles, for any choice of α presented, as expected.

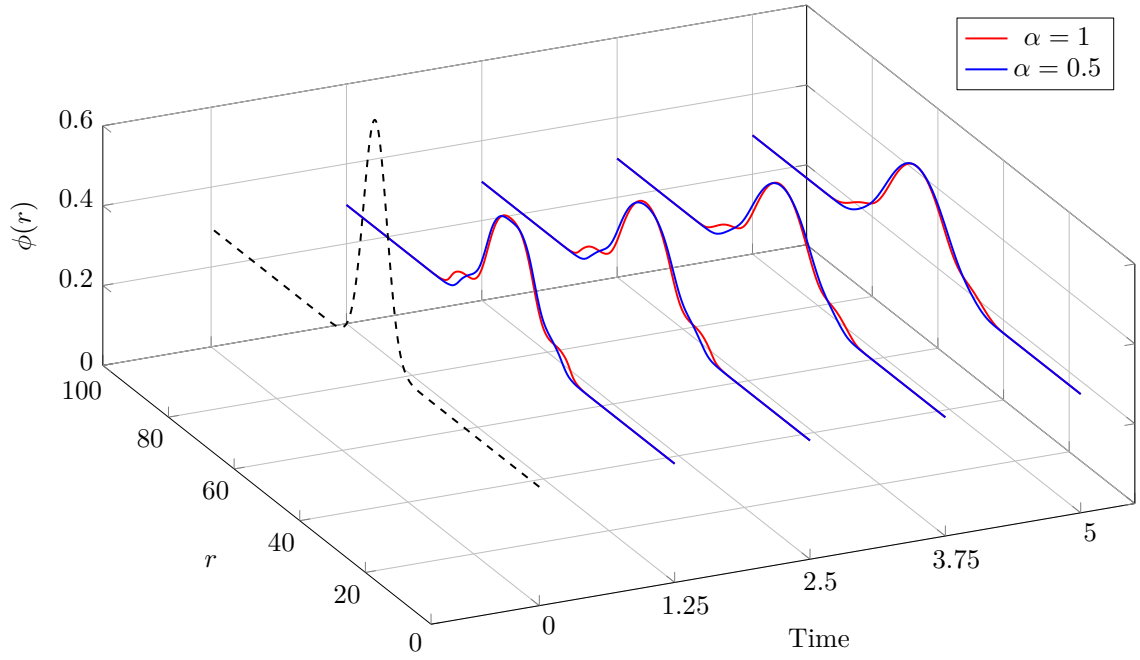


(a) $\alpha = 1$

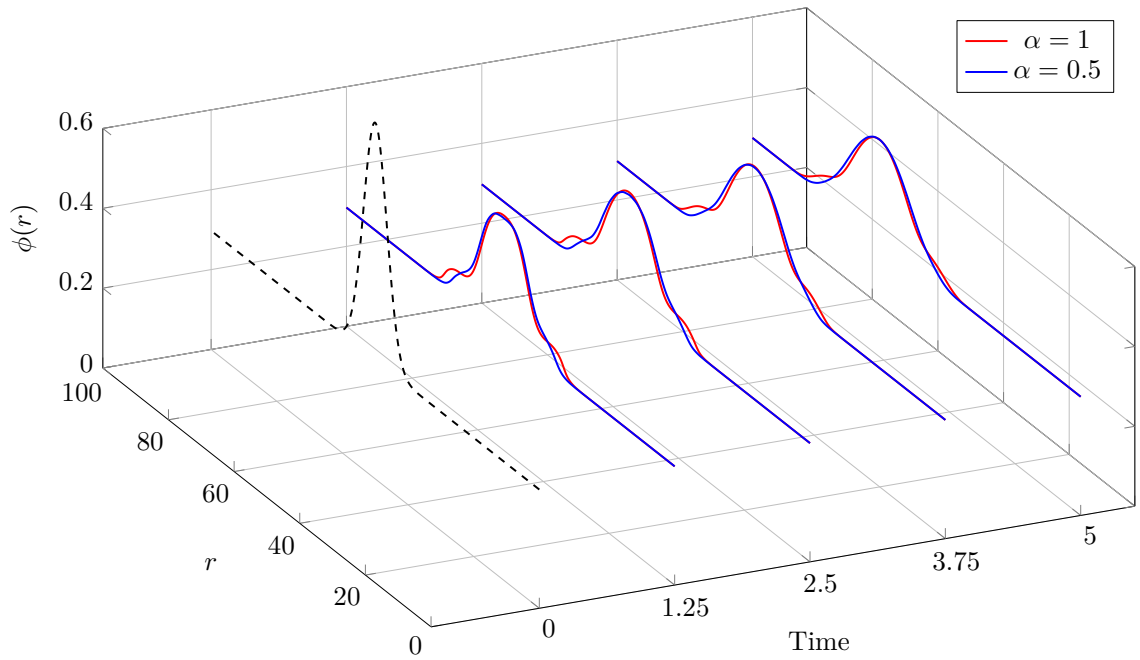


(b) $\alpha = 0.5$

Figure 8.2: The density of a system of particles obeying eqs. (6.39), (6.42) and (6.49), with initial conditions given by eq. (8.12), where we take $v_0 = 0$ and various values for E_0. We see a similar effect from the choice of E_0 in both cases, and some changes when collisions are inelastic.



(a) $V^{\text{ext}} = 0$



(b) $V^{\text{ext}} = -5r$

Figure 8.3: The density of a system of particles obeying eqs. (6.39), (6.42) and (6.49) with initial conditions given by eq. (8.12), where $v_0 = 0$ and $E_0 = 200$. The inclusion of an external potential translates the density, but does not affect it otherwise.

8.2.2 Example: Merging masses

In this example we will show that inelasticity in the collision operator can cause clustering effects in simulations. On the periodic domain $[0, 100]$, we consider the initial conditions:

$$\rho_0(r) = \frac{\phi_v}{N_v}\left(e^{\frac{(r-25)^2}{25}} + e^{\frac{(r-75)^2}{25}} + 0.5\right), \quad v_0(r) = 20\sin\left(\frac{2\pi r}{100}\right), \quad E_0(r) = 250, \tag{8.13}$$

where $\phi_v = 0.3$ is the total packing fraction of the system and N_v is a normalisation constant to ensure that the density integrates to N. In fig. 8.4 we provide graphs of the initial conditions ϕ_0 and v_0. For this example we take $\gamma = 2$ and $V^{\text{ext}} = 0$ with $\varepsilon = 2.5 \times 10^{-4}$, and include the Percus

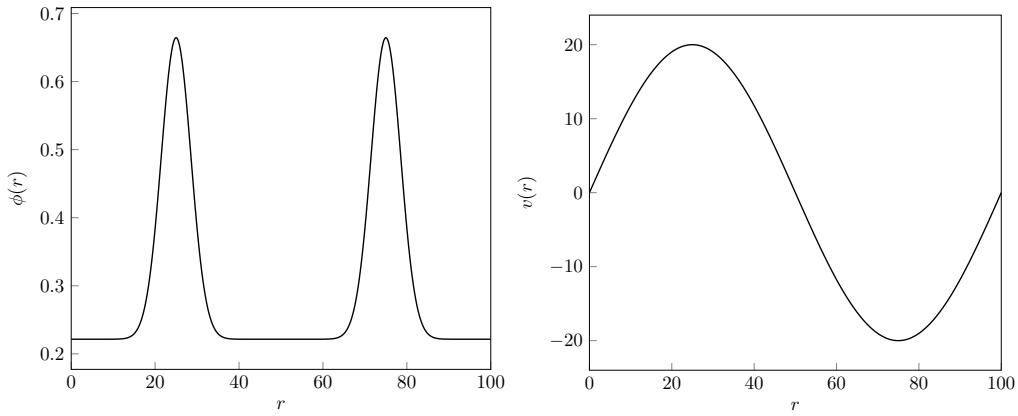


Figure 8.4: The initial conditions given by eq. (8.13). We have chosen these conditions so that two masses of particles will 'collide', which we anticipate will cause different behaviour, depending on the choice of α.

free energy functional to ensure that the effects of volume exclusion are included in the system. We use the collisional term to include inelastic affects for varying choices of $\alpha \in [0, 1]$. As $\rho = 0.3$, for the particle diameter considered there are $N = 125000$ particles in the system and the total packing fraction of this domain is $N = 400000$.

In fig. 8.5 we show the effect of each term in the DDFT. We can see that without the collision term and the volume exclusion term, the two masses of particles are partially reflected away from one another. This still occurs when volume exclusion and collisional terms are included, but when we use the values for $g^{(2)}(\varepsilon)$ that are provided by our numerical experiments the masses coalesce instead of being reflected. We note that when the collision term is included and the one-body distributions are assumed to be uncorrelated (*i.e.* $g^{(2)}(\varepsilon) = 1$) the collisional terms have less of an effect.

Figure 8.6 presents the results of the DDFT for different choices of α, where all terms are included. We see that when α is small the two masses coalesce, while when α is close to 1, a larger portion of the masses are reflected.

In fig. 8.7, we fix the packing fraction of the system but consider different particle numbers in the simulation. We note that in this case the size of the domain is modified so that the particle diameter is constant. We see that for small values of N the results differ; the size of the reflected masses is larger when N is small, but for $N = 1200$ or $N = 120000$ the results are roughly the same.

Alternatively we can reduce the number of particles in the system while keeping the domain size constant, *i.e.* reduce the total packing fraction. In fig. 8.8, we consider different values of ϕ_v, and see that when ϕ increases the two masses coalesce for lower values of the coefficient of restitution, as the volume exclusion effects become more prominent.

Furthermore, in fig. 8.9 we consider the same dynamics for different number of particles and compare the difference between using the limit of the collision operator eq. (4.79) as the particle

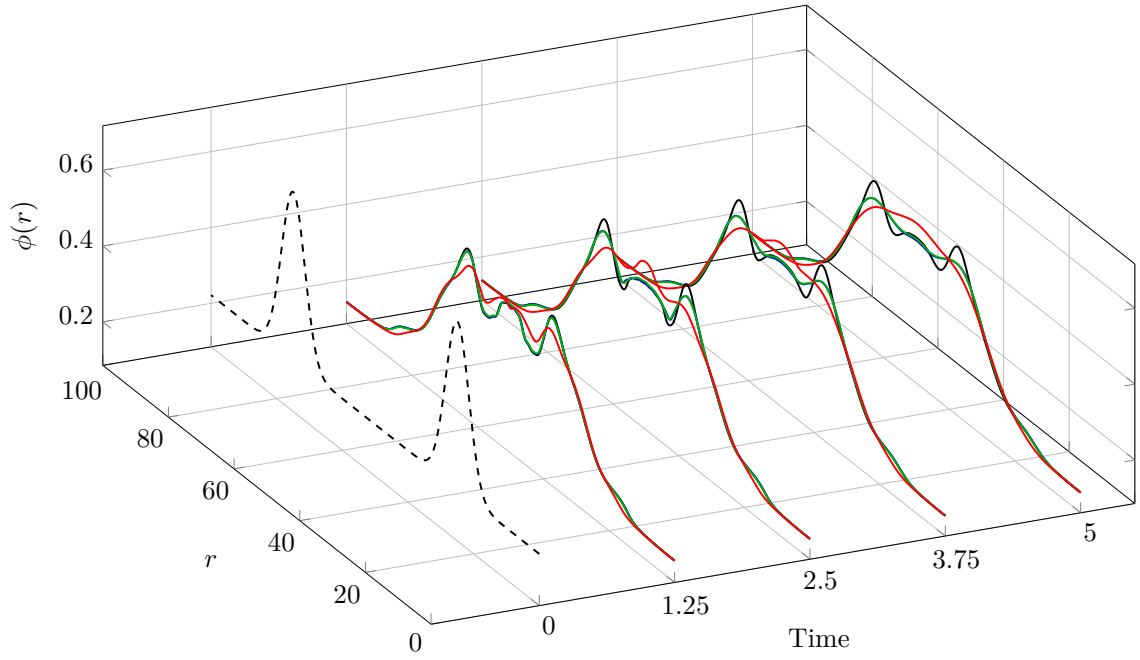


Figure 8.5: Results from the DDFT simulation for $\alpha = 0.5$. Each simulation has the same initial condition (black, dashed) at time 0. The black line neglects the collision operator and the free energy term. The blue line includes the free energy term but not the collision operator. The green line includes both terms, with $g^{(2)}(\sigma) = 1$, and the red line includes both terms and uses $g^{(2)}(\sigma)$ determined by particle simulations. The results show the importance of including each term in the model.

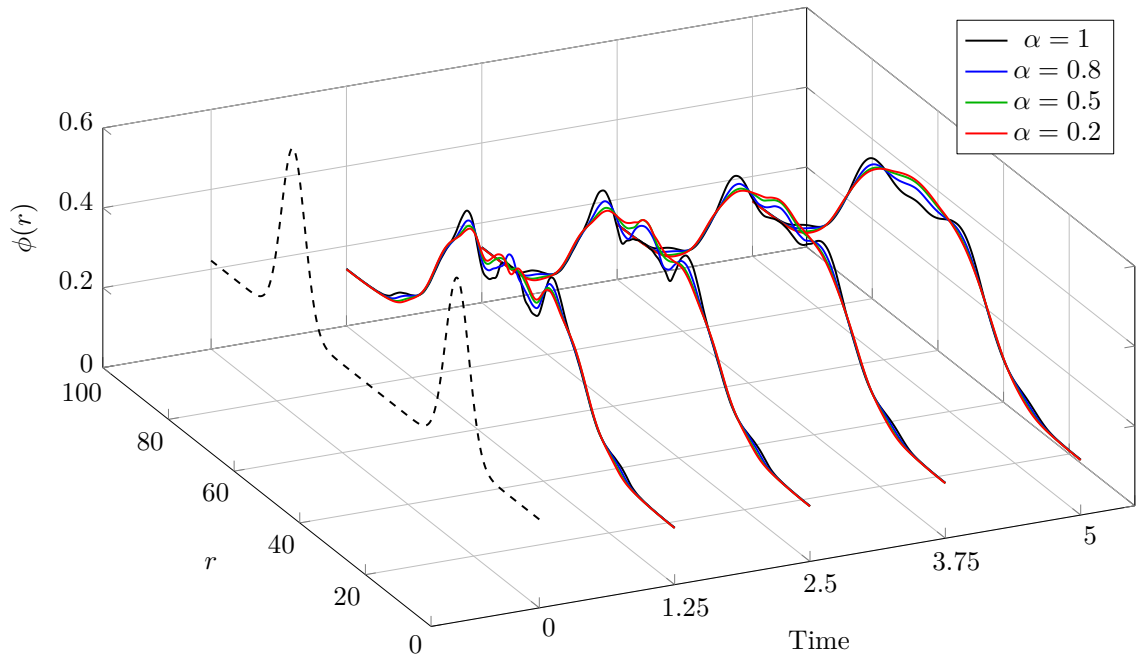
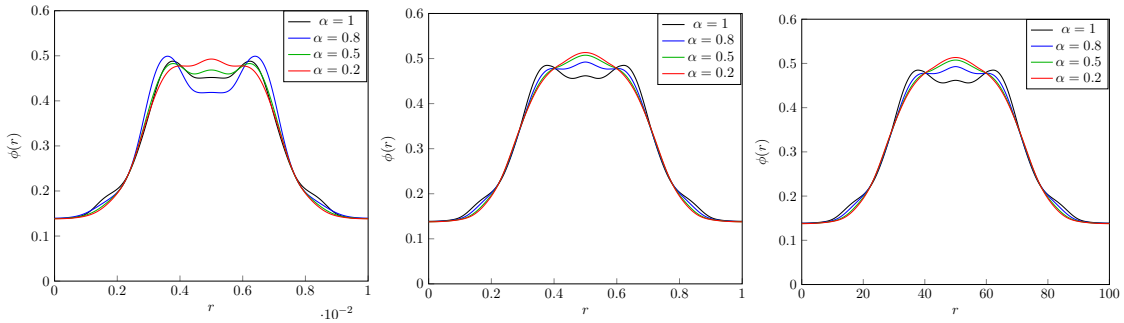
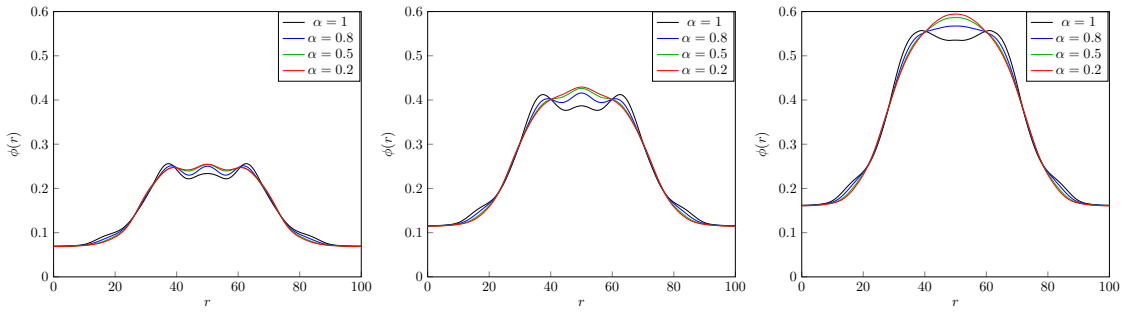


Figure 8.6: Results from the DDFT simulation for different values of α. Each simulation has the same initial condition (black, dashed) at time 0. When $\alpha = 1$ the two masses are reflected, while for $\alpha < 1$ the masses tend to coalesce.



(a) $N = 12$ (b) $N = 1200$ (c) $N = 120000$

Figure 8.7: Simulations involving different numbers of particles, given initial condition eq. (8.13), and dynamics are ran up to time $t = 5$ for different values of α.The results differ when considering a smaller number of particles, but are more similar when N is large.



(a) $\phi_v = 0.15$ (b) $\phi_v = 0.25$ (c) $\phi_v = 0.35$

Figure 8.8: Simulations involving different packing fractions, given initial condition eq. (8.13), and dynamics are ran up to time $t = 5$ for different values of α. For large total packing fractions the resulting shape of the density is smoother, as volume exclusion properties stop the formation of a high peak at the centre of the domain.

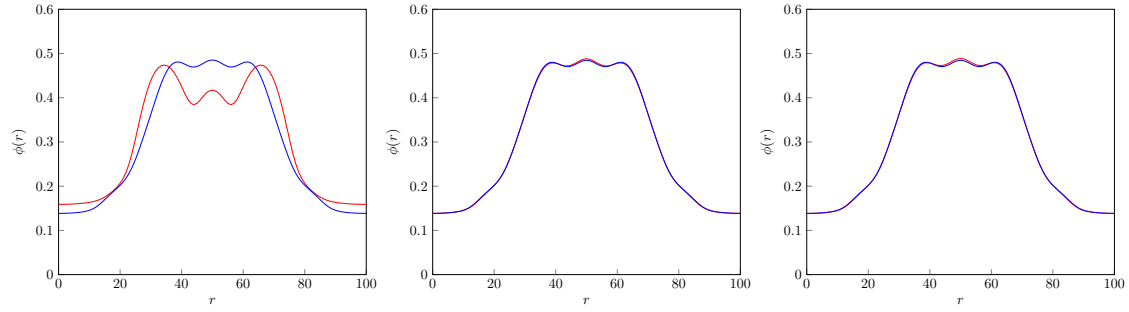
diameter vanishes. In this case we rescale the particle diameter to include more particles in the domain, rather than rescaling the domain. We see that there is a large difference when the number of particles is small, and as expected as the number of particles increases (and the particle diameter becomes small) the difference in the dynamics is small. We expect this when considering the collision operator for inelastic hard rods; as $\varepsilon \to 0$ the second centred moment of the collision operator is $\mathcal{O}(E^{3/2})$.

In fig. 8.10 we consider a system with $N = 175,000$ particles, which has total packing fraction $\phi_v = 0.4375$, and increase the magnitude of the initial velocity; we set:

$$v_0(r) = 26 \sin\left(\frac{2\pi r}{100}\right) \tag{8.14}$$

We show the importance of including the volume exclusion term independently of the collision operator, when using a parametrisation of the radial correlation function. In the example where the Percus free energy term is not included, the dynamics lead to non-physical densities; at time $t = 1.5$ we see that the density exceeds 1.

If instead we use the analytic form for $g^{(2)}(\varepsilon)$ discussed in chapter 7, the dynamics do not have a



(a) $N = 24$ (b) $N = 2400$ (c) $N = 240000$

Figure 8.9: Simulations involving different numbers of particles, given initial condition eq. (8.13), and dynamics are ran up to time $t = 5$ for $\alpha = 0.5$ and fixed total packing fraction $\rho_v = 0.3$, using the collision operator eq. (4.79) (red) and its limit as $\varepsilon \to 0$ (blue). For a small number of particles the results differ greatly, but as the number of particles increase the result is nearly identical, as expected.

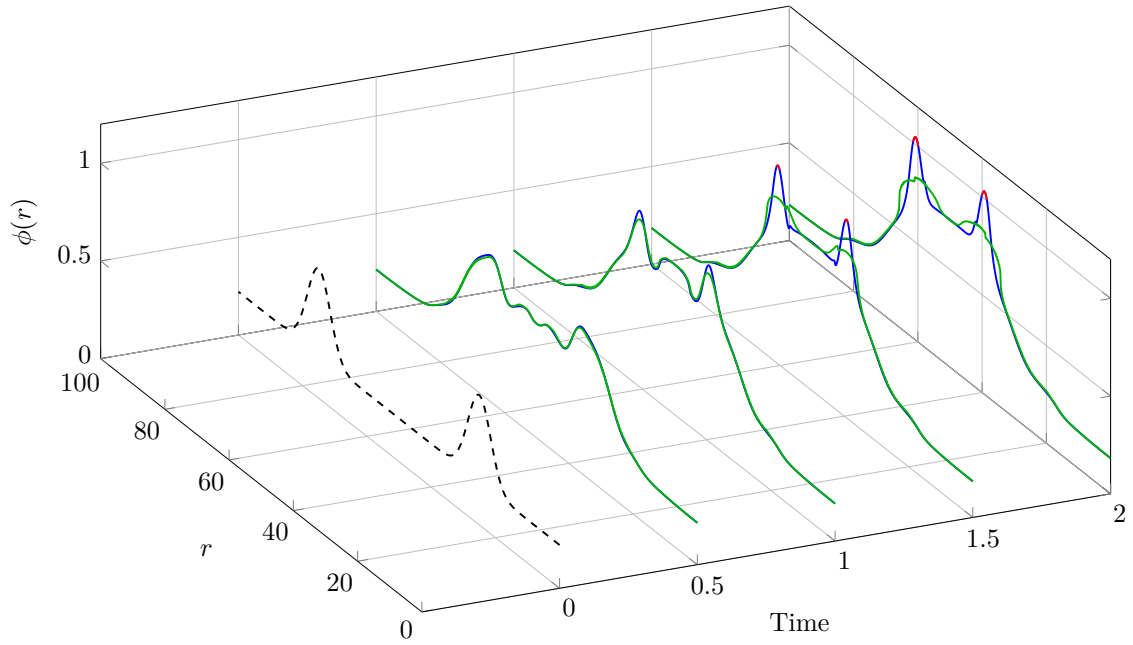


Figure 8.10: Dynamics where the volume exclusion term is included (green) and neglected (blue), with non-physical densities highlighted in red. For this example to ensure accuracy we preform the dynamics on 600 computational points. When volume exclusion is neglected the system reaches non-physical packing fractions.

strong relationship with the coefficient of restitution α, as seen in fig. 8.11. In this case the analytic form of $g^{(2)}(\varepsilon) = g^{(2)}_{\mathcal{P}}(\varepsilon)$ ensures that the volume does not approach non-physical values, but the effect of the collision operator is minimal when we compare to the results in fig. 8.6. Furthermore, there is less evidence of volume exclusion effects at lower densities.

Finally, we present the dynamics of the system for different values of α, at time $t = 50$. In this case the frictional effects present dominate the dynamics, so the results of different choices of α are more similar.

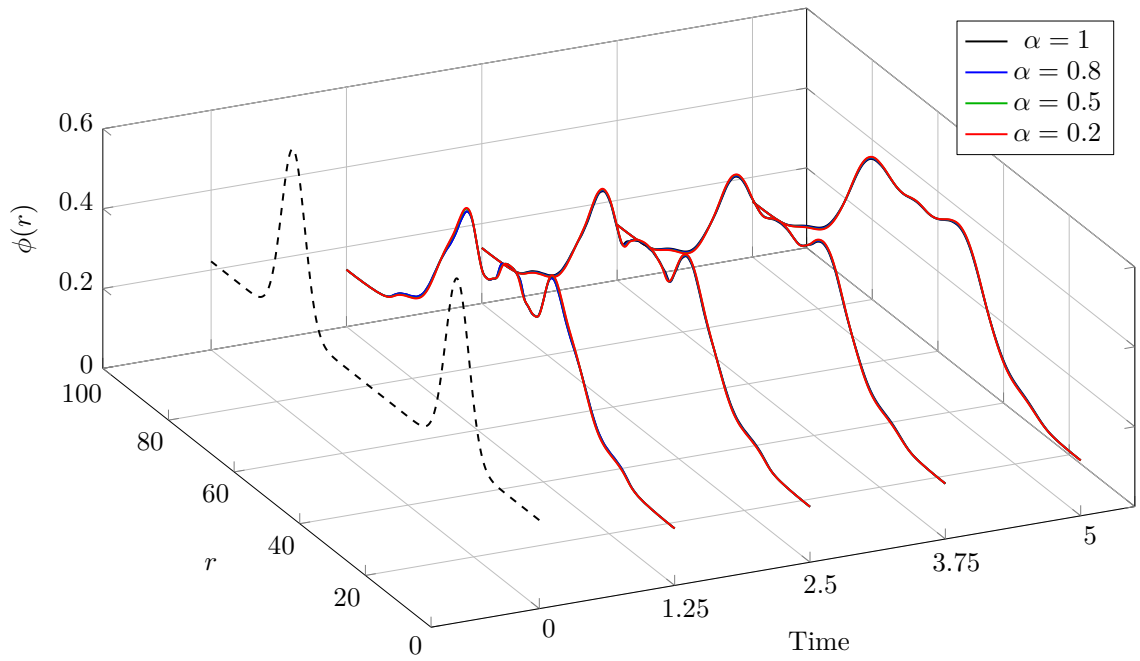


Figure 8.11: The dynamics when eq. (7.3) is used with initial conditions eq. (8.13), and the Percus free energy functional is neglected. In this case the effect of the collision operator has significantly reduced.

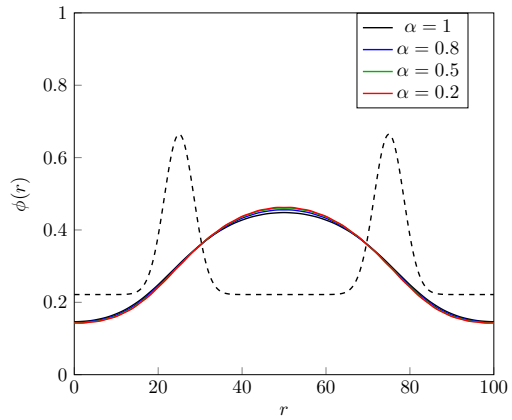


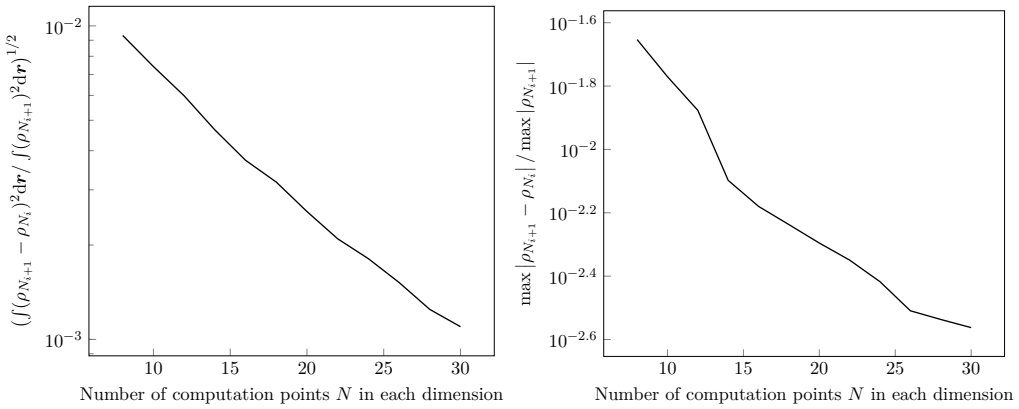
Figure 8.12: Long time dynamics of the system with initial conditions eq. (8.13), for different values of α. After long times the viscous effects in the system dominate the dynamics, and inelastic effects are less evident.

8.3 Two dimensions

Using the methods discussed at the start of the chapter, and the results in appendix A, we can perform simulations of eqs. (6.39), (6.42) and (6.49) in two dimensions. As we have not considered boundary conditions in the numerical approximation of the correlation function in chapter 7 we will restrict our simulations to periodic domains. In all simulations we will use fundamental measure theory to include volume exclusion in the dynamics; for this we use the results discussed in [89].

In fig. 8.13 we provide the relative L^2 and L^∞ errors for the dynamics considered in section 8.3.1, for a variety of computational points. We see that the relative error is approximately 0.1% for a computational grid of 30 points in each dimension, but for a grid of 20 computational points

in each dimension the error is also small. In the following examples we perform all dynamics on 20×20 computational grids.



(a) (b)

Figure 8.13: Relative L^2 (left) and L^∞ (right) errors in the density ρ for the dynamics considered in section 8.3.1. The error is small for $N = 20$, although it is clear that increasing the number of computational points will further increase accuracy, but will increase computation time.

8.3.1 Example: Gaussian bump

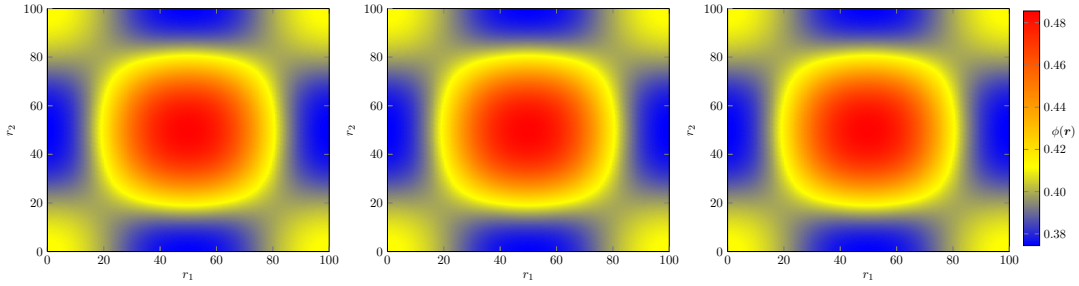
For our first example we consider the analogue to the one-dimensional case considered in section 8.2.1, to consider the effect of the matrix $\boldsymbol{E}(\boldsymbol{r}, t)$ and the coefficient of restitution α. We use initial conditions:

$$\rho_0(\boldsymbol{r}) = \frac{\rho_v}{N_c}\left(\exp\left(-\frac{(r_1-50)^2}{25}\right)\exp\left(-\frac{(r_2-50)^2}{25}\right)+1\right), \quad \boldsymbol{v}_0(\boldsymbol{r}) = 0, \quad \boldsymbol{E}_0(\boldsymbol{r}) = E_0\boldsymbol{I}, \tag{8.15}$$

where $E_0 > 0$, ρ_c is the total packing fraction of the system, and N_c is a normalisation constant for the Gaussian term. We perform the dynamics up to time $t = 5$ and consider the results. Unless otherwise stated we take $\rho_v = 0.4, N = 500000$, and $E_0 = 150$. In fig. 8.14 we display the results for different coefficients of restitution using the values of $g^{(2)}(\varepsilon)$ constructed from microscopic dynamics in chapter 7, where we have set $E_0 = 200$. Compared to the results in one dimension, the effect of the collision operator for different α is much smaller, in fig. 8.15 we present the difference in the result at time $t = 5$ for $\alpha = 0.5$ and $\alpha = 1$, the scale of the difference is above the scale of the error but is still small compared to the overall density; for the duration of the dynamics the largest difference between the two cases is around 2%. We note that the effect of the collision operator is strongest at areas where there are higher gradients in the density; where inelasticity in collisions is most likely to cause a difference in the dynamics.

The effect of inelasticity would be more significant in systems where $\gamma = 0$, where the effect of the collision operator in the granular temperature equation does not have to compete with the viscous drag in the system. However, we require viscous drag in the system to validate the Maxwellian assumption in the derivation of the DDFT, so cannot consider non-viscous cases with the current theory.

In fig. 8.16 we present the diagonal terms in the second moment for the initial conditions



(a) $\alpha = 0.5$ (b) $\alpha = 0.7$ (c) $\alpha = 1$

Figure 8.14: The density at time $t = 5$ using initial conditions eq. (8.15), for different coefficients of restitution α. The small differences in the resulting densities are hard to capture with the human eye.

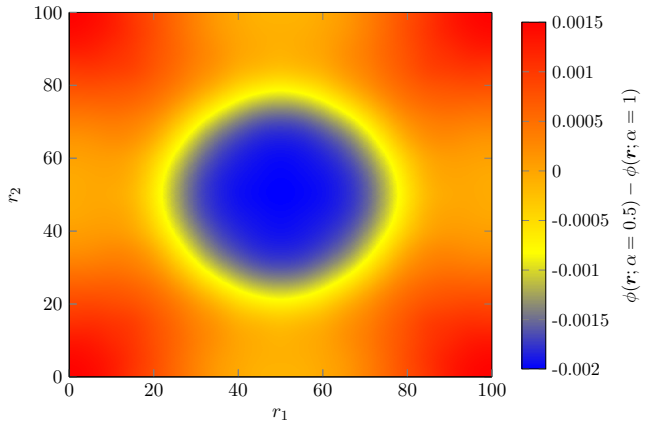


Figure 8.15: The difference in density at time $t = 5$ using initial conditions eq. (8.15), between dynamics using $\alpha = 0.5$ and $\alpha = 1$. This result shows that inelasticity does play a small role in the dynamics of the system.

eq. (8.15), for $\alpha = 0.5$ and $\alpha = 1$. We see that in the inelastic case the second moment takes larger values at higher densities. We note that for $\alpha = 1$ the second moment will vanish as $\varepsilon \to 0$.

We can also consider dynamics where the off-diagonal terms of $\boldsymbol{E}_0$ are non-zero and consider a larger number of particles; The results are presented in fig. 8.17. In this case we see some changes in the symmetry of the density.

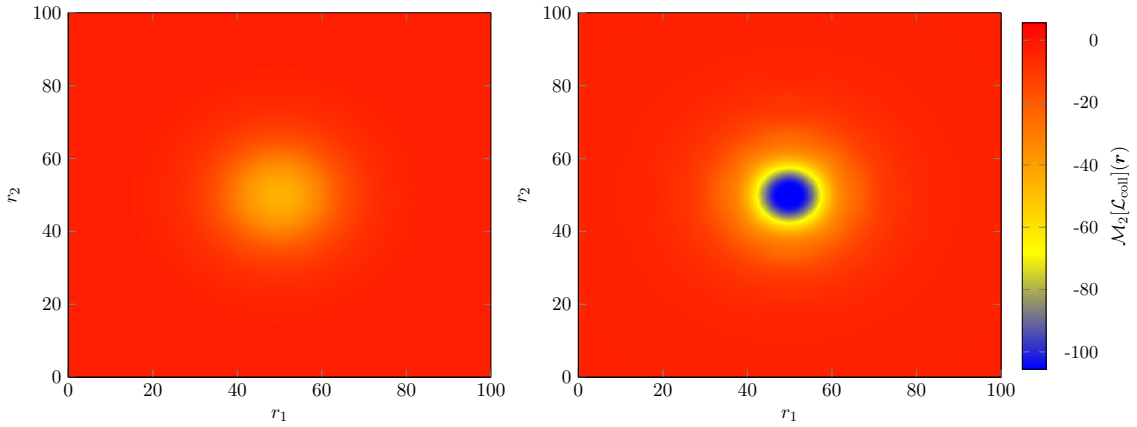
8.3.2 Example: Moving Masses

In this example we use initial conditions that are analogous to the one-dimensional moving masses case, where we expect the value of α will determine how much particles in the system cluster:

$$\rho_0(\boldsymbol{r}) = \frac{\phi_v}{N_c}\left(\exp\left(-\frac{(r_1-25)^2}{25}\right)\exp\left(-\frac{(r_2-25)^2}{25}\right) + \exp\left(-\frac{(r_1-75)^2}{25}\right)\exp\left(-\frac{(r_2-75)^2}{25}\right) + 1\right),$$

$$\boldsymbol{v}_0(\boldsymbol{r}) = \left(v_1\cos\left(2\pi r_1/L_1\right), v_2\cos\left(2\pi r_2/L_2\right)\right),$$

$$\boldsymbol{E}_0(\boldsymbol{r}) = E_0\boldsymbol{I}. \tag{8.16}$$



(a) $\alpha = 1$ (b) $\alpha = 0.5$

Figure 8.16: The diagonal entries of the second moment of the collision operator in eq. (6.49) for different values of α, with $\rho, \boldsymbol{v}, \boldsymbol{E}$ given by eq. (8.15). Again, we see that α will affect the local granular temperature.

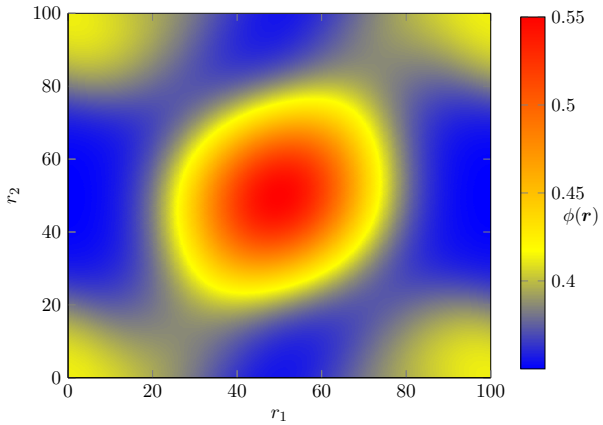
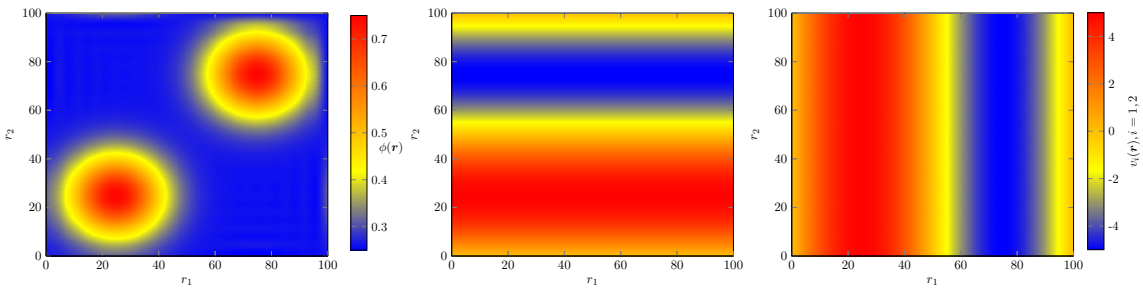


Figure 8.17: The density at time $t = 5$, with initial conditions given by eq. (8.15), using $\alpha = 0.5, \rho_v = 0.5$ and $N = 500000$, but with $\boldsymbol{E}_0 = (200, 100; 100, 200)$. There is a noticeable change in the resulting density when compared to fig. 8.14.

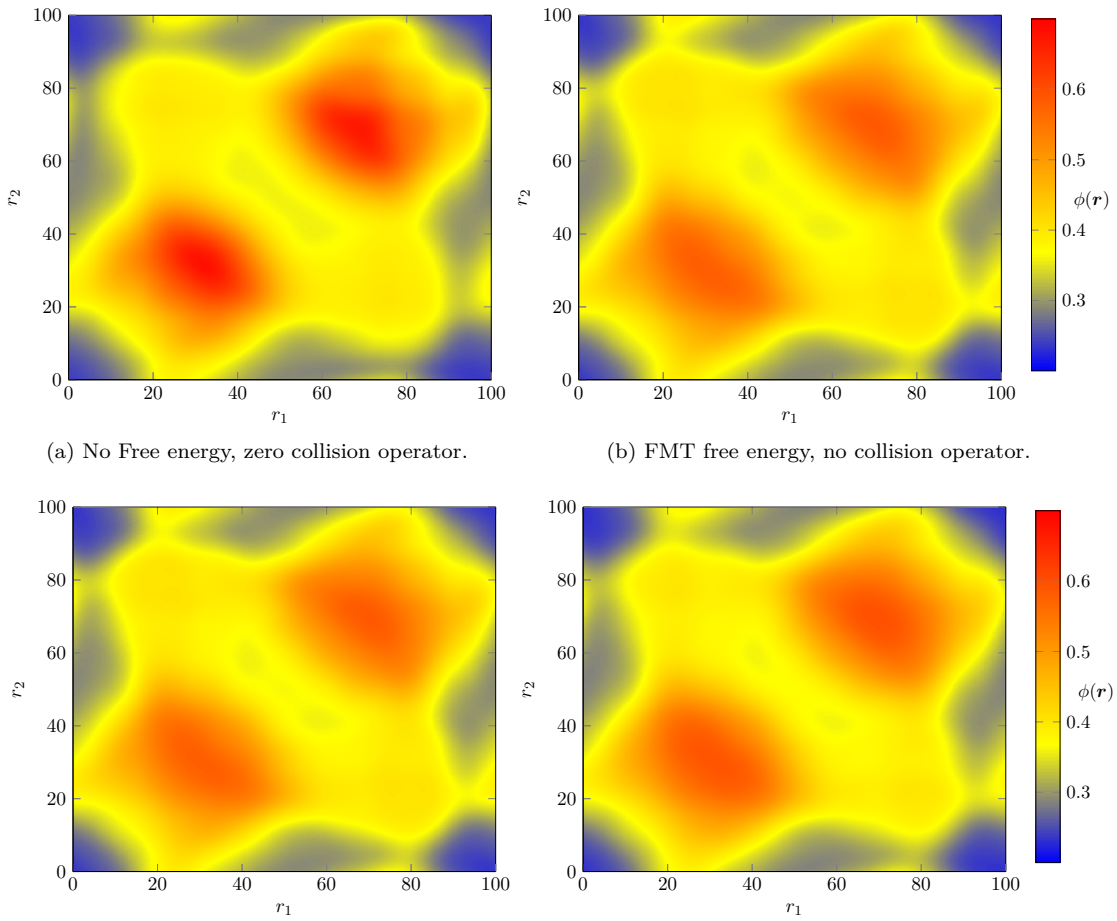
For $v_1, v_2 \in \mathbb{R}$ and $E_0 > 0$. Unless otherwise stated, we will use $\phi_v = 0.5, \alpha = 0.5, E_0 = 200$ and $N = 500000$. The initial conditions for ϕ_0 and $\boldsymbol{v}_0$ are then presented in fig. 8.18.



(a) ρ_0 (b) $v_1(\boldsymbol{r}, 0)$ (c) $v_2(\boldsymbol{r}, 0)$

Figure 8.18: Initial conditions described in eq. (8.16). In a similar way to the one-dimensional case, we have chosen initial conditions that cause two masses of particles to move towards one another.

We use this example to consider the effect of each term in the DDFT, and the effect of the magnitude of $\boldsymbol{E}$ in the equations. In fig. 8.19 we present the solution of the equations eqs. (6.39), (6.42) and (6.49) at time $t = 5$, where various terms are neglected and we set $E_0 = 150, v_1 = v_2 = 2.5$. We see that when the FMT free energy term is included, the peaks in the density at $t = 5$ are broader and lower, as the effects of volume exclusion stop the density from reaching higher values. Again we see that the effect of the collision operator is not as strong as expected. In fig. 8.20 we present the difference between the densities given in fig. 8.19. The largest difference occurs when FMT is included, however we note that the effect of the collision terms is non-zero and higher than the error present in the simulations. Furthermore when using the data from numerical simulations there is a larger change in the resulting density at time $t = 5$.



(a) No Free energy, zero collision operator.

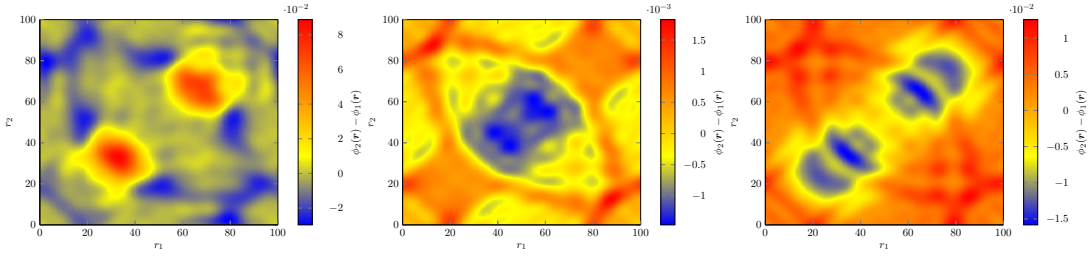
(b) FMT free energy, no collision operator.

(c) FMT free energy, collision operator with $g = 1$.

(d) FMT free energy, collision operator with g derived from numerical results.

Figure 8.19: The density at time $t = 5$, where at time $t = 0$ the initial conditions eq. (8.16) are used, and dynamics are given by the granular DDFT, applying different terms and using $\alpha = 0.5$. This displays the role that each term in the DDFT plays, although some of the effects are too small to capture visually.

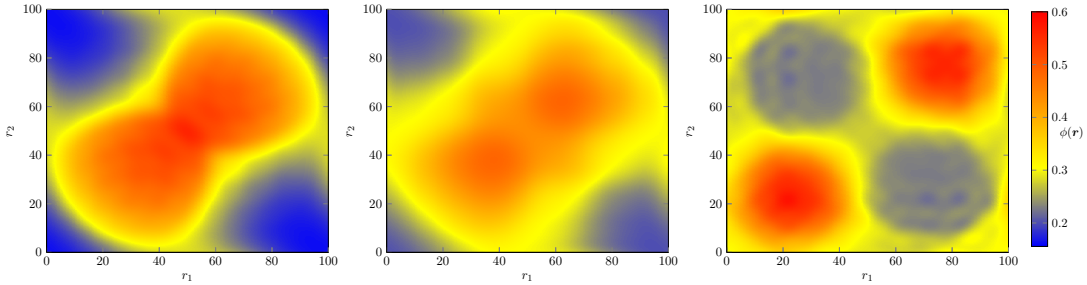
In fig. 8.21 we investigate the effect of the magnitude of E_0 on the dynamics. When E_0 is small, the two masses coalesce, while for E_0 large the masses are reflected, which is similar to the behaviour in the one-dimensional case. We also note evidence of potential numerical inaccuracies in the result in fig. 8.21c (in the form of local oscillations in the density); these are likely due to fast



(a) (b) (c)

Figure 8.20: The difference in results presented in fig. 8.19. In fig. 8.20a we show the difference between fig. 8.19a and fig. 8.19b, in fig. 8.20b we consider fig. 8.19b versus fig. 8.19c, and in fig. 8.20c we consider fig. 8.19c against fig. 8.19d. In each case we see a non-zero change in the density.

changes in the local velocity, which are difficult to simulate using pseudospectral methods where smoothness of solutions is important, and hard to model accurately under the Maxwellian local equilibrium assumption, as discussed in the derivation of the DDFT.



(a) $E_0 = 5$ (b) $E_0 = 50$ (c) $E_0 = 500$

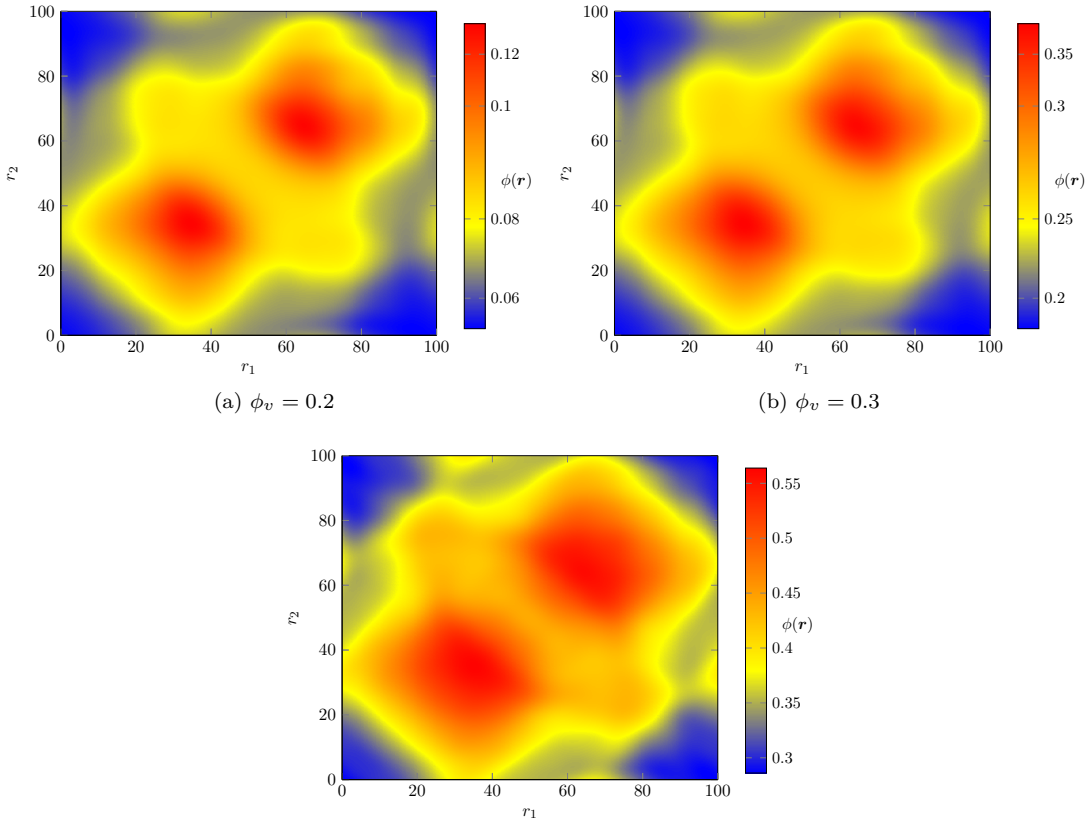
Figure 8.21: The density at time $t = 5$, where at time $t = 0$ the initial conditions eq. (8.16) are used, and dynamics are given by the granular DDFT, using $\alpha = 0.5$ and different values for E_0. We see a noticeable change in the density as E_0 increases, but when E_0 is too large the numerical scheme struggles to simulate the fast changes in density.

We can consider the effect of different packing fractions on the system in fig. 8.22. As expected, when the packing fraction is increased the areas of highest density become more spread as volume exclusion reduces the local density. Finally in fig. 8.23 we present the same example where instead of changing densities we consider different numbers of particles. The effect is much smaller than witnessed in the one-dimensional case.

8.3.3 Example: Periodic waves

For our final example we want to consider initial conditions that are less conveniently structured, to show how the system behaves for more arbitrary initial conditions. We set:

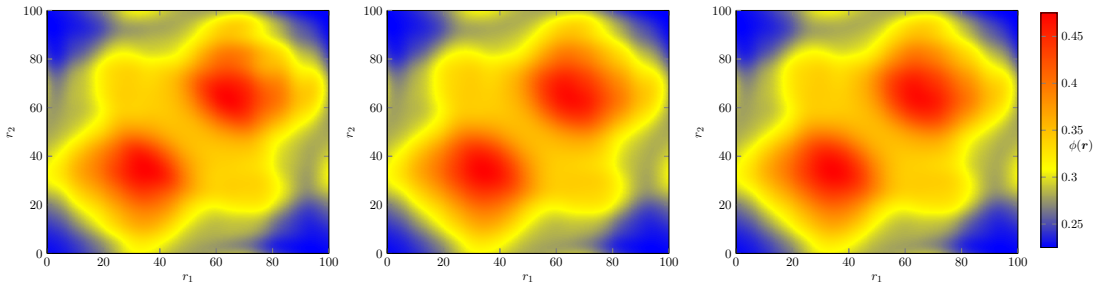
$$\rho_0(\boldsymbol{r}) = \frac{0.5}{N_c}\Big\{3\cos\left(2\pi\left(\frac{r_1}{100} - 0.1\right)\right) + \sin\left(2\pi\left(2\frac{r_1}{100} + \frac{r_2}{100} + 0.3\right)\right)$$
$$+1.2\cos\left(\frac{2\pi r_1}{100}\right)\sin\left(\frac{2\pi r_2}{100}\right) + \sin\left(\frac{4\pi r_1 r_2}{100^2}\right)^2 + 5.7\Big\},$$
$$v_{1,0}(\boldsymbol{r}) = \frac{3}{2}\sin\left(\frac{2\pi r_1}{100}\right)\cos\left(\frac{2\pi r_2}{100}\right),$$



(a) $\phi_v = 0.2$ (b) $\phi_v = 0.3$

(c) $\phi_v = 0.5$

Figure 8.22: The density at time $t = 5$, where at time $t = 0$ the initial conditions eq. (8.16) are used, and dynamics are given by the granular DDFT, using $\alpha = 0.5$, $E_0 = 250$ and different values of ϕ_V. Note that the colour scale varies, depending on the plot. We see that high packing fractions dramatically change the result of the system, as volume exclusion effects become more prevalent.



(a) $N = 40$ (b) $N = 4000$ (c) $N = 400000$

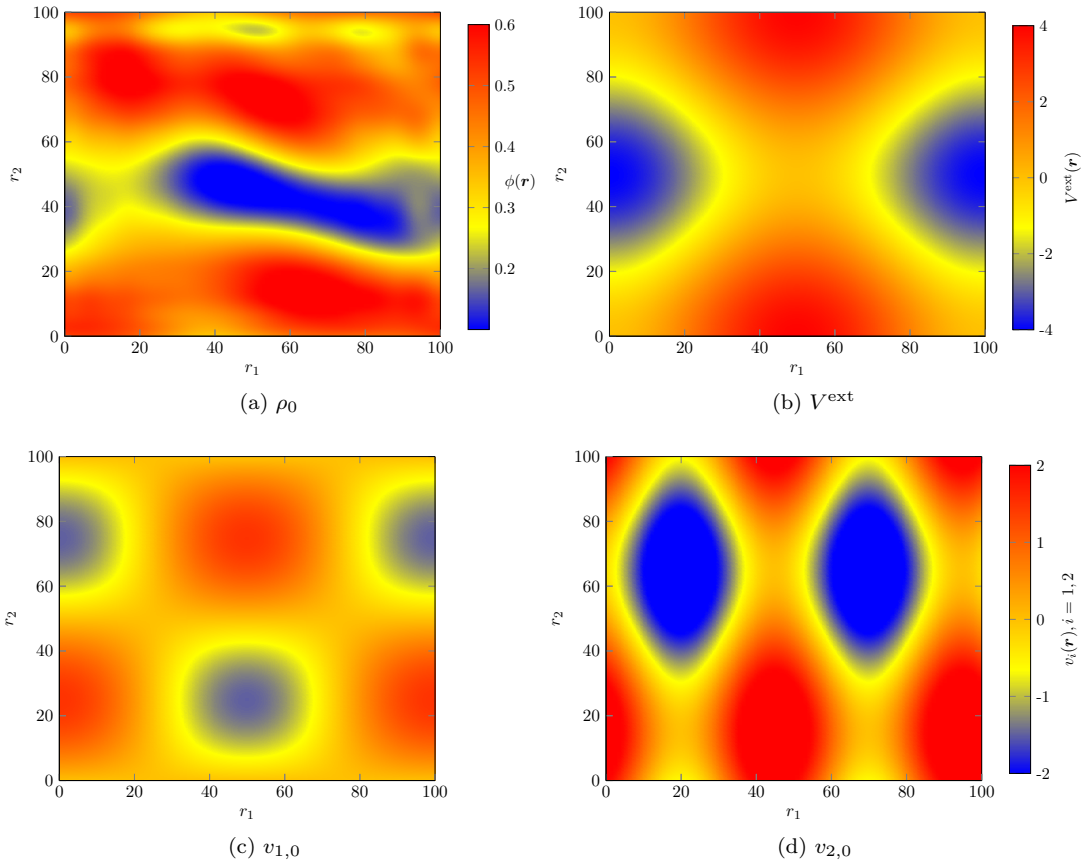
Figure 8.23: The density at time $t = 5$, where at time $t = 0$ the initial conditions eq. (8.16) are used, and dynamics are given by the granular DDFT, using $\alpha = 0.5$, $E_0 = 250$ and different values of N, keeping ρ_V constant. When N is small there is a change in behaviour in the local density, but for large values of N the behaviour is more similar.

$$v_{2,0}(\boldsymbol{r}) = -\frac{3}{2}\left\{\cos\left(4\pi\left(\frac{r_2}{100} - 0.2\right)\right) - \sin\left(2\pi\left(\frac{r_1}{100} - 0.9\right)\right)\right\},$$
$$\boldsymbol{E}_0 = \begin{pmatrix} 250 & 0 \\ 0 & 200 \end{pmatrix}. \tag{8.17}$$

We also consider a trigonometric external potential:

$$V^{\text{ext}}(\boldsymbol{r}) = 2\cos\left(\frac{2\pi r_1}{100}\right) - 2\cos\left(\frac{2\pi r_2}{100}\right). \tag{8.18}$$

The initial conditions ρ_0 and $\boldsymbol{v}_0$ are presented in fig. 8.24, as well as the external potential V^{ext}. We consider a system with $N = 500000$ particles.



(a) ρ_0 (b) V^{ext}

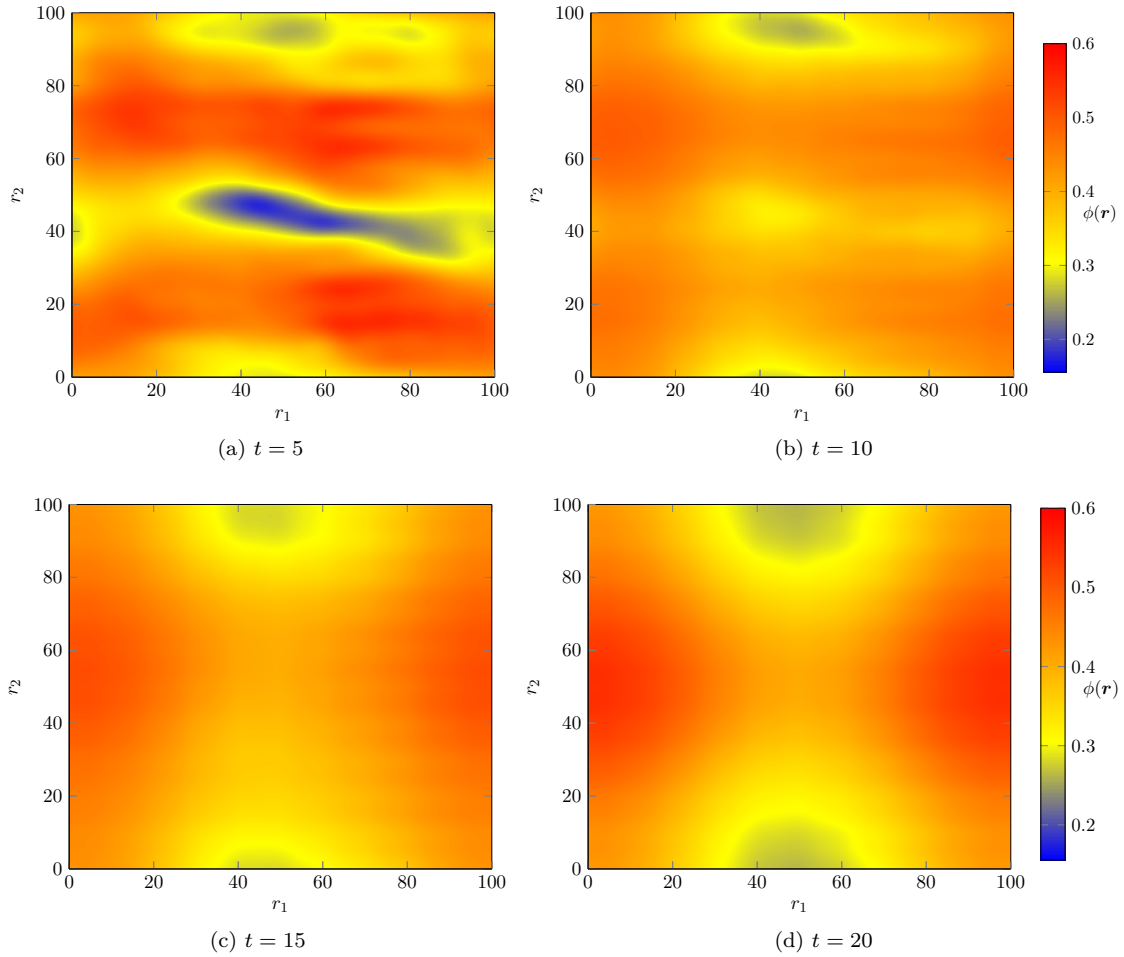
(c) $v_{1,0}$ (d) $v_{2,0}$

Figure 8.24: The initial conditions given by eq. (8.17), which are constructed using sums and products of trigonometric functions that are periodic in the domain.

In fig. 8.25 We present the results at different times t. We see that the external potential transports the density, while the peaks and troughs of the density also change due to volume exclusion and collisional effects. The long term dynamics lead to accumulations of particles in the areas of low external potential.

8.4 Discussion

In this chapter we have provided basic numerical examples in one and two dimensions that present the effect of each term in the DDFT. We have shown that including the parametrised versions of the radial correlation function in the continuum model, rather than their steady-state counterparts, produces changes in the dynamics of the system that are qualitatively expected. Furthermore, we produced expected results when increasing the number of particles or local packing fraction of the



(a) $t = 5$ (b) $t = 10$

(c) $t = 15$ (d) $t = 20$

Figure 8.25: The dynamics of the granular DDFT with initial conditions given by eq. (8.17). We see that over time the external potential field determines the shape of the density. Further investigation using alternative numerical schemes is requires to determine whether the horizontal bands in the system are an artefact of the pseudospectral scheme.

system. The inclusion of the free energy functional also allows us to implement a volume exclusion term, which further changes the overall behaviour of the density.

These promising initial results imply that the DDFT could be an appropriate model for systems of granular media. Before applying the DDFT to more realistic problems and real-world applications, we need the theory to account for multiple species (*i.e.* construct a suitable DDFT alternative for the two-fluid model), and remove the necessity for a viscous drag term, so that we can use the results to simulate fluidisation. This involves removing or replacing the Maxwellian local equilibrium assumption that we used to close the system of equations and derive the DDFT in chapter 6, which requires a sufficiently large value for γ to ensure the approximation is accurate. Furthermore, we need to investigate the effects of boundaries on the derivation, and on the radial correlation function. Once these issues have been investigated, we will be in a suitable position to simulate fluidisation examples that are popular in the literature, and can be compared with experimental results.

For these more complicated systems, pseudospectral methods may no longer be appropriate, as

they may not capture small scale variation accurately without drastically increasing computational cost; if the numerics are not sufficiently smooth then artefacts of the pseudospectral method will arise, as witnessed in some of the two-dimensional examples already considered. Therefore to accurately identify the behaviour of the DDFT in future numerical experiments on physically relevant systems of granular media, we must investigate other valid numerical approaches. There are various code libraries, including the open-source library `OpenFOAM` [166], or the proprietary software `COMSOL` [167], that may be suitable for DDFT simulation.

Chapter 9

Conclusions and future work

In this work we have considered different problems in aspects of granular media. Firstly we have considered granular media at a microscopic scale; constructing a general form for instantaneous interactions between particles, and discussing efficient methods to simulate the dynamics of systems of interacting hard particles. This new approach fully characterises all possible instantaneous interactions under the assumption that linear and angular momentum are conserved. We have also implemented event-driven for particles that are affected by external friction in one, two, and three dimensions.

Following this, our careful construction of the Liouville equation for hard particles produced additional terms relating the the admissible position and velocity data, that lead to the Boltzmann collision operator for hard inelastic spheres of diameter ε. This novel derivation allows any instantaneous interactions to be included in the Liouville equation. From here we have derived an original DDFT which effectively models systems of hard particles that experience inelastic collisions as a continuum. We used event-driven simulations to construct the radial correlation function from microscopic systems of particles. The results show that the radial correlation function displays a rich behaviour which depends on many parameters of the system that are not usually considered, *e.g.* the coefficient of restitution. Simple numerical examples of the DDFT then showed how the properties of the correlation function carry over to the macroscopic regime.

A common characteristic of research is that by finding answers, we tend to end up asking more questions. We now describe some open problems that could be considered from this work.

9.1 Open problems

9.1.1 Microscopic modelling

In Theorem 2 by considering the conservation of linear and angular momentum we reduced the six-dimensional set of equations that determine the effect of an instantaneous interaction between hard spheres to a one-dimensional equation. The resulting equation then allowed us to construct the inelastic collision rule eq. (2.31), as well as refraction rules eq. (2.37) used in systems with discrete step potentials.

It is also of interest to construct interaction rules for dynamics with additional degrees of freedom, for example by including angular rotation or charge in the properties of the system.

Identifying these interaction rules will then allow for a derivation of the Liouville equation for more complex dynamics.

9.1.2 Microscopic simulation

We have modified event driven methods to simulate dynamics where particles are affected by external friction. The resulting dynamics were then used to construct correlation functions for dynamics after long times. Aside from improvements in the efficiency of algorithms considered, there are some improvements that can be made to the simulation.

Firstly, for low coefficients of restitution ($\alpha < 0.4$) and for higher densities, microscopic simulations are jammed. This could be due to stability issues in the algorithm, as discussed in section 3.5.1, or due to cases of inelastic collapse occurring, as discussed in section 3.5.2. An in-depth investigation into systems that experience jamming is required to understand the issue. In either case, we require a method to stop the simulation from jamming, while preserving the properties of the system as well as possible.

Furthermore, we could also consider systems with external friction and where particles interact with one another via discrete step potentials, or interact with an external boundary. In each case we need to construct analytic forms of the time where the interaction takes place, so that we can use EDPD efficiently. It would also be useful to include more complicated dynamics in EDPD simulations where possible, for example by including additional degrees of freedom (after a physically valid event rule has been defined for such interactions).

9.1.3 The weak formulation of the Liouville equation

After careful construction of the Liouville equation for hard spheres, we showed that the Boltzmann collision operator present in the first equation of the BBGKY hierachy is a consequence of an additional term in the Liouville equation. Our result is accurate for systems with any external or pairwise interaction potentials, provided that the instantaneous events occurring between particles are well-separated. For our purposes, we considered a system where $N = 2$, then heuristically argued that through a partition of the phase space the same result could approximate the N particle system. It would be beneficial to have a precise derivation of this result.

We also considered inelastic particles that are unaffected by additional degrees of freedom such as angular rotation or charge. Following a construction of the correct possible interactions between particles given additional degrees of freedom, we could analogously formulate the weak Liouville equation.

There are also particular examples that would be interesting to study using the weak formulation of the Liouville equation. In section 4.5 we considered a system of particles which are affected by square shoulder interactions. This resulted in an additional surface term in the BBGKY hierarchy at the second interaction diameter. If instead we consider $a < 0$ in this example (*i.e.* square-well particles), the dynamics are more complicated; for example particles can be trapped between the two interaction diameters, and there are initial velocities where dynamics are not defined globally in time, leading to additional terms in the Liouville equation. A careful study of this system, followed by homogenisation methods on the additional terms in the Liouville equation, could lead to effective potential terms in the Liouville equation for square well particles, or even

arbitrary stepped potentials.

As a final note on the results of Chapter 4, it would be very beneficial to study the properties of a system where events are not well separated; for example in systems where inelastic collapse can occur. In such systems we can no longer assume that dynamics are unique for all times, or event that they are defined. A detailed study of this problem could lead to a better understanding of inelastic collapse in continuum systems.

9.1.4 Dynamical density functional theory for granular media

By considering the first three moments of the BBGKY hierarchy for inelastic hard spheres, we successfully constructed a new dynamical density functional theory for modelling granular media. Importantly, the model includes terms that determine the dynamics of $\boldsymbol{E}(\boldsymbol{r}, t)$, the variance in the local average velocity in the system. The model produced some expected properties of granular media.

When constructing eqs. (6.39), (6.42) and (6.49), we made some simplifying assumptions on the system. In particular, we assumed that the local equilibrium of the system was Maxwellian, an approximation that is known to be inaccurate for granular media. For systems where the external friction is small, this leads to ill-posed equations of motion. The effects of this approximation are then present in our simulations, when γ is small, or equivalently $\boldsymbol{v}$ is large.

To improve this result, we could instead consider a sonine polynomial approximation, as is discussed in [129]. Alternatively, an investigation of corrective terms to the Maxwellian assumption as in [148], where in addition we consider collision terms and the third moment of the BBGKY hierarchy, could lead to a more accurate DDFT for modelling granular media.

A careful consideration of the granular DDFT under various limits (for example taking the limit as the number of particles increases while scaling the particle diameter) and approximations (in alignment to the results in [70]) may also provide insight into the properties of the three-moment DDFT.

Most importantly, by considering a particle system with multiple species (one of which models hard, inelastic particles), the equations derived in chapter 6 could be extended to a coupled system of equations that model fluidisation processes, by accounting for interactions between the fluid and the granular medium. This could produce a two-fluid model that rivals current state-of-the-art models that are used to simulate modern industrial processes.

9.1.5 Parametrisation of the radial correlation function

The granular DDFT that we construct requires an understanding of the radial correlation function for the system of particles considered. It is popular to use the steady-state form of $g^{(2)}$ for elastic hard spheres, however we have shown that the long-time dynamics of the radial correlation function for inelastic hard spheres behaves very differently to the steady-state form. We therefore used an approximation of the radial correlation function derived from long-time dynamics of particle systems, and parametrised $g^{(2)}$ in terms of the local density.

We could extend this approach further by constructing the radial correlation function at provided points in time in the continuum simulation: given times $t_0 = 0, t_1, \ldots, t_N = T$, at time t_i we sample from the density of the system, and use the highly efficient EDPD simulations to run

a microscopic simulation for time $t_{\Updownarrow\rangle\rfloor\nabla\wr}$, where $t_{\Updownarrow\rangle\rfloor\nabla\wr} \ll t_{i+1} - t_i$. From the result we extract the radial correlation function, which we use in the continuum simulation until time t_{i+1}, where we repeat the process. This method may construct more physically valid results by reactively updating the radial correlation function.

As an alternative it would be more intuitive to parametrise $g^{(2)}$ in terms of the variance in the local average velocity, $\boldsymbol{E}(\boldsymbol{r}, t)$, as we know that inelastic collisions affect the local variance more than the local density and velocity. An accurate construction of the radial correlation function may help produce the expected effect of inelasticity in the DDFT simulations. Furthermore, it would also be useful to parametrise $g^{(2)}$ in a form similar to eq. (7.3), to align with known results in the literature.

Bibliography

[1] R. A. Bagnold, *The physics of blown sand and desert dunes*. Courier Corporation, 2012.

[2] P. Richard, M. Nicodemi, R. Delannay, P. Ribiere, and D. Bideau, "Slow relaxation and compaction of granular systems," *Nature Materials*, vol. 4, no. 2, pp. 121–128, 2005.

[3] R. Hoyle and R. B. Hoyle, *Pattern formation: an introduction to methods*. Cambridge University Press, 2006.

[4] G. H. Ristow, *Pattern formation in granular materials*. Springer Science & Business Media, 2000.

[5] E. Clément, L. Vanel, J. Rajchenbach, and J. Duran, "Pattern formation in a vibrated two-dimensional granular layer," *Physical Review E*, vol. 53, no. 3, p. 2972, 1996.

[6] P. Tegzes, T. Vicsek, and P. Schiffer, "Avalanche dynamics in wet granular materials," *Physical review letters*, vol. 89, no. 9, p. 094 301, 2002.

[7] A. Daerr and S. Douady, "Two types of avalanche behaviour in granular media," *Nature*, vol. 399, no. 6733, pp. 241–243, 1999.

[8] S. P. Pudasaini and K. Hutter, *Avalanche dynamics: dynamics of rapid flows of dense granular avalanches*. Springer Science & Business Media, 2007.

[9] A. N. Edwards, A. S. Russell, C. G. Johnson, and J. M. N. T. Gray, "Frictional hysteresis and particle deposition in granular free-surface flows," *Journal of Fluid Mechanics*, vol. 875, pp. 1058–1095, 2019.

[10] K. J. Walsh, D. C. Richardson, and P. Michel, "Spin-up of rubble-pile asteroids: Disruption, satellite formation, and equilibrium shapes," *Icarus*, vol. 220, no. 2, pp. 514–529, 2012.

[11] D. J. Scheeres, C. M. Hartzell, P. Sánchez, and M. Swift, "Scaling forces to asteroid surfaces: The role of cohesion," *Icarus*, vol. 210, no. 2, pp. 968–984, 2010.

[12] E. Asphaug, S. J. Ostro, R. Hudson, D. J. Scheeres, and W. Benz, "Disruption of kilometre-sized asteroids by energetic collisions," *Nature*, vol. 393, no. 6684, pp. 437–440, 1998.

[13] P. Lawton, "Coefficients of friction between cereal grain and various silo wall materials," *Journal of Agricultural Engineering Research*, vol. 25, no. 1, pp. 75–86, 1980.

[14] R. Kobyłka, M. Molenda, and J. Horabik, "Loads on grain silo insert discs, cones, and cylinders: Experiment and DEM analysis," *Powder technology*, vol. 343, pp. 521–532, 2019.

[15] K. M. Hill and J. Kakalios, "Reversible axial segregation of rotating granular media," *Physical Review E*, vol. 52, no. 4, p. 4393, 1995.

[16] I. S. Aranson and L. S. Tsimring, "Patterns and collective behavior in granular media: Theoretical concepts," *Reviews of modern physics*, vol. 78, no. 2, p. 641, 2006.

[17] S. Rapagnà, N. Jand, A. Kiennemann, and P. Foscolo, "Steam-gasification of biomass in a fluidised-bed of olivine particles," *Biomass and bioenergy*, vol. 19, no. 3, pp. 187–197, 2000.

[18] T. Nordgreen, T. Liliedahl, and K. Sjöström, "Metallic iron as a tar breakdown catalyst related to atmospheric, fluidised bed gasification of biomass," *Fuel*, vol. 85, no. 5-6, pp. 689–694, 2006.

[19] J. Richardson and W. Zaki, "Sedimentation and fluidisation: Part I," *Chemical Engineering Research and Design*, vol. 75, S82–S100, 1997.

[20] H. Mickley and D. F. Fairbanks, "Mechanism of heat transfer to fluidized beds," *AIChE Journal*, vol. 1, no. 3, pp. 374–384, 1955.

[21] B. Leimkuhler and C. Matthews, *Molecular Dynamics*. Springer, 2016.

[22] R. Hockney and J. Eastwood, *Computer Simulation Using Particles*. CRC Press, 1988, ISBN: 9781439822050.

[23] F. Radjaï and F. Dubois, *Discrete-element modeling of granular materials*. Wiley-Iste, 2011.

[24] P. A. Cundall and O. D. Strack, "A discrete numerical model for granular assemblies," *Géotechnique*, vol. 29, no. 1, pp. 47–65, 1979.

[25] B. J. Alder and T. E. Wainwright, "Studies in molecular dynamics. I. General method," *Journal of Chemical Physics*, vol. 31, no. 2, pp. 459–466, 1959.

[26] H. Zhu, Z. Zhou, R. Yang, and A. Yu, "Discrete particle simulation of particulate systems: Theoretical developments," *Chemical Engineering Science*, vol. 62, no. 13, pp. 3378–3396, 2007.

[27] ——, "Discrete particle simulation of particulate systems: A review of major applications and findings," *Chemical Engineering Science*, vol. 63, no. 23, pp. 5728–5770, 2008.

[28] A. Di Renzo and F. P. Di Maio, "Comparison of contact-force models for the simulation of collisions in DEM-based granular flow codes," *Chemical engineering science*, vol. 59, no. 3, pp. 525–541, 2004.

[29] F. Radjai and V. Richefeu, "Contact dynamics as a nonsmooth discrete element method," *Mechanics of Materials*, vol. 41, no. 6, pp. 715–728, 2009.

[30] R. Kobyłka and M. Molenda, "DEM simulations of loads on obstruction attached to the wall of a model grain silo and of flow disturbance around the obstruction," *Powder Technology*, vol. 256, pp. 210–216, 2014.

[31] C. González-Montellano, F. Ayuga, and J. Ooi, "Discrete element modelling of grain flow in a planar silo: Influence of simulation parameters," *Granular Matter*, vol. 13, no. 2, pp. 149–158, 2011.

[32] B. Mishra and R. K. Rajamani, "The discrete element method for the simulation of ball mills," *Applied Mathematical Modelling*, vol. 16, no. 11, pp. 598–604, 1992.

[33] J. Ghaboussi and R. Barbosa, "Three-dimensional discrete element method for granular materials," *International Journal for Numerical and Analytical Methods in Geomechanics*, vol. 14, no. 7, pp. 451–472, 1990.

[34] G. Lu, J. Third, and C. Müller, "Discrete element models for non-spherical particle systems: From theoretical developments to applications," *Chemical Engineering Science*, vol. 127, pp. 425–465, 2015.

[35] M. Sakai, M. Abe, Y. Shigeto, S. Mizutani, H. Takahashi, A. Viré, J. R. Percival, J. Xiang, and C. C. Pain, "Verification and validation of a coarse grain model of the DEM in a bubbling fluidized bed," *Chemical Engineering Journal*, vol. 244, pp. 33–43, 2014.

[36] D. Queteschiner, T. Lichtenegger, S. Pirker, and S. Schneiderbauer, "Multi-level coarse-grain model of the DEM," *Powder Technology*, vol. 338, pp. 614–624, 2018.

[37] G. Paul, "A complexity $\mathcal{O}(1)$ priority queue for event driven molecular dynamics simulations," *Journal of Computational Physics*, vol. 221, no. 2, pp. 615–625, 2007, ISSN: 00219991. DOI: 10.1016/j.jcp.2006.06.042.

[38] J. Jenkins, "Particle segregation in collisional flows of inelastic spheres," in *Physics of dry granular media*, Springer, 1998, pp. 645–658.

[39] N. V. Brilliantov and T. Pöschel, *Kinetic theory of granular gases*. Oxford University Press, 2010.

[40] J. Ding and D. Gidaspow, "A bubbling fluidization model using kinetic theory of granular flow," *AIChE journal*, vol. 36, no. 4, pp. 523–538, 1990.

[41] G. S. Singh and B. Kumar, "Kinetic theory of dense fluids of rigid biaxial ellipsoids," *Phys. Rev. E*, vol. 62, pp. 7927–7940, 6 Dec. 2000. DOI: 10.1103/PhysRevE.62.7927. [Online]. Available: https://link.aps.org/doi/10.1103/PhysRevE.62.7927.

[42] R. G. González and V. Garzó, "Transport coefficients for granular suspensions at moderate densities," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2019, no. 9, p. 093 204, 2019.

[43] J. T. Jenkins and D. Berzi, "Dense inclined flows of inelastic spheres: Tests of an extension of kinetic theory," *Granular Matter*, vol. 12, no. 2, pp. 151–158, 2010.

[44] J. T. Jenkins, "Dense inclined flows of inelastic spheres," *Granular matter*, vol. 10, no. 1, pp. 47–52, 2007.

[45] M. Pulvirenti and S. Simonella, "The Boltzmann–Grad limit of a hard sphere system: Analysis of the correlation error," *Inventiones mathematicae*, vol. 207, no. 3, pp. 1135–1237, 2017.

[46] S. Simonella, T. A. G. Genovese, G. Gentile, A. Giuliani, A. Pellegrinotti, and M. Pulvirenti, "BBGKY hierarchy for hard sphere systems," Sapienza University of Rome, Tech. Rep., 2011.

[47] I. Gallagher, L. Saint-Raymond, and B. Texier, *From Newton to Boltzmann: hard spheres and short range potentials*, 2012. arXiv: 1208.5753.

[48] E. R. Parra and K. Kamrin, "Capturing transient granular rheology with extended fabric tensor relations," *Granular Matter*, vol. 21, no. 4, p. 89, 2019.

[49] D. Schaeffer, T. Barker, D. Tsuji, P. Gremaud, M. Shearer, and J. Gray, "Constitutive relations for compressible granular flow in the inertial regime," *Journal of Fluid Mechanics*, vol. 874, pp. 926–951, 2019.

[50] P. Jop, Y. Forterre, and O. Pouliquen, "A constitutive law for dense granular flows," *Nature*, vol. 441, no. 7094, pp. 727–730, 2006.

[51] T. Barker, D. Schaeffer, M. Shearer, and J. Gray, "Well-posed continuum equations for granular flow with compressibility and $\mu(I)$-rheology," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 473, no. 2201, p. 20 160 846, 2017.

[52] B. G. M. Van Wachem and A. E. Almstedt, "Methods for multiphase computational fluid dynamics," *Chemical Engineering Journal*, vol. 96, no. 1-3, pp. 81–98, 2003.

[53] J. Zhao and T. Shan, "Coupled CFD–DEM simulation of fluid–particle interaction in geomechanics," *Powder Technology*, vol. 239, pp. 248–258, 2013.

[54] C. Kloss, C. Goniva, A. Hager, S. Amberger, and S. Pirker, "Models, algorithms and validation for opensource DEM and CFD–DEM," *Progress in Computational Fluid Dynamics, an International Journal*, vol. 12, no. 2-3, pp. 140–152, 2012.

[55] K. Chu, J. Chen, and A. Yu, "Applicability of a coarse-grained CFD–DEM model on dense medium cyclone," *Minerals Engineering*, vol. 90, pp. 43–54, 2016.

[56] W. Zhong, A. Yu, X. Liu, Z. Tong, and H. Zhang, "DEM/CFD–DEM modelling of non-spherical particulate systems: Theoretical developments and applications," *Powder technology*, vol. 302, pp. 108–152, 2016.

[57] F. Alobaid, J. Ströhle, and B. Epple, "Extended CFD/DEM model for the simulation of circulating fluidized bed," *Advanced Powder Technology*, vol. 24, no. 1, pp. 403–415, 2013.

[58] D. Jajcevic, E. Siegmann, C. Radeke, and J. G. Khinast, "Large-scale CFD–DEM simulations of fluidized granular systems," *Chemical Engineering Science*, vol. 98, pp. 298–310, 2013.

[59] M. Ishii and K. Mishima, "Two-fluid model and hydrodynamic constitutive relations," *Nuclear Engineering and design*, vol. 82, no. 2-3, pp. 107–126, 1984.

[60] S. Schneiderbauer and S. Pirker, "Filtered and heterogeneity-based subgrid modifications for gas–solid drag and solid stresses in bubbling fluidized beds," *AIChE journal*, vol. 60, no. 3, pp. 839–854, 2014.

[61] S. Schneiderbauer, "A spatially-averaged two-fluid model for dense large-scale gas-solid flows," *AIChE Journal*, vol. 63, no. 8, pp. 3544–3562, 2017.

[62] ——, "Validation study on spatially averaged two-fluid model for gas-solid flows: II. Application to risers and fluidized beds," *AIChE Journal*, vol. 64, no. 5, pp. 1606–1617, 2018.

[63] E. B. Pitman and L. Le, "A two-fluid model for avalanche and debris flows," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 363, no. 1832, pp. 1573–1601, 2005.

[64] Y. T. Makkawi, P. C. Wright, and R. Ocone, "The effect of friction and inter-particle cohesive forces on the hydrodynamics of gas–solid flow: A comparative analysis of theoretical predictions and experiments," *Powder Technology*, vol. 163, no. 1-2, pp. 69–79, 2006.

[65] J. Kolehmainen, A. Ozel, and S. Sundaresan, "Eulerian modelling of gas–solid flows with triboelectric charging," *Journal of Fluid Mechanics*, vol. 848, pp. 340–369, 2018.

[66] V. Lee, S. R. Waitukaitis, M. Z. Miskin, and H. M. Jaeger, "Direct observation of particle interactions and clustering in charged granular streams," *Nature Physics*, vol. 11, no. 9, p. 733, 2015.

[67] S. McNamara and W. R. Young, "Inelastic collapse in two dimensions," vol. 50, no. 1, 1994, ISSN: 1063651X.

[68] A. J. Archer, "Dynamical density functional theory for molecular and colloidal fluids: a microscopic approach to fluid mechanics," *Journal of Chemical Physics*, vol. 130, p. 014 509, 2009.

[69] B. D. Goddard, A. Nold, N. Savva, G. A. Pavliotis, and S. Kalliadasis, "General dynamical density functional theory for classical fluids," *Physical Review Letters*, vol. 109, no. 12, pp. 1–5, 2012, ISSN: 00319007.

[70] B. D. Goddard, G. A. Pavliotis, and S. Kalliadasis, "The overdamped limit of dynamic density functional theory: Rigorous results," *Multiscale Modeling & Simulation*, vol. 10, no. 2, pp. 633–663, 2012.

[71] U. M. B. Marconi and P. Tarazona, "Dynamic density functional theory of fluids," *Journal of Chemical Physics*, vol. 110, no. 16, pp. 8032–8044, 1999.

[72] M. Marín and P. Cordero, "An empirical assessment of priority queues in event-driven molecular dynamics simulation," *Computer Physics Communications*, vol. 92, no. 2, pp. 214–224, 1995, ISSN: 0010-4655.

[73] U. Marini Bettolo Marconi and P. Tarazona, "Nonequilibrium inertial dynamics of colloidal systems," *Journal of Chemical Physics*, vol. 124, no. 16, p. 164 901, 2006.

[74] U. Marconi, P. Tarazona, and F. Cecconi, "Theory of thermostatted inhomogeneous granular fluids: A self-consistent density functional description," *Journal of Chemical Physics*, vol. 126, no. 16, pp. 1–13, 2007.

[75] S. Torquato, "Nearest-neighbor statistics for packings of hard spheres and disks," *Physical Review E*, vol. 51, no. 4, p. 3170, 1995.

[76] B. Goddard, A. Nold, and S. Kalliadasis, "Multi-species dynamical density functional theory," *Journal of Chemical Physics*, vol. 138, no. 14, p. 144 904, 2013.

[77] T. D. Hurst, B. D. Goddard, and M. Wilkinson, *A derivation of the Liouville equation for hard particle dynamics with non-conservative interactions*, 2019. arXiv: 1910.06656 [cond-mat.stat-mech].

[78] K. H. Hughes and I. Burghardt, "Maximum-entropy closure of hydrodynamic moment hierarchies including correlations," *Journal of Chemical Physics*, vol. 136, no. 21, p. 214 109, 2012.

[79] B. D. Goddard, T. D. Hurst, and R. Ocone, *Modelling inelastic granular media using dynamical density functional theory*, 2020. arXiv: 2003.07327 [cond-mat.stat-mech].

[80] B. D. Goddard, A. Nold, and S. Kalliadasis, *2DChebClass [Software]*, 2017. [Online]. Available: dx.doi.org/10.7488/ds/1991.

[81] B. D. Goddard and T. Hurst, *Multiple superadiabatic transitions and Landau-Zener formulas*, 2018. arXiv: 1804.04660 [physics.chem-ph].

[82] V. Betz, B. D. Goddard, and T. D. Hurst, "Nonadiabatic transitions in multiple dimensions," *SIAM Journal on Scientific Computing*, vol. 41, no. 5, B1011–B1033, 2019.

[83] O. M. Crook, T. Hurst, C.-B. Schönlieb, M. Thorpe, and K. C. Zygalakis, *PDE-Inspired algorithms for semi-supervised learning on point clouds*, 2019. arXiv: 1909.10221 [math.NA].

[84] A. Einstein, "Über die von der molekularkinetischen theorie der wärme geforderte bewegung von in ruhenden flüssigkeiten suspendierten teilchen," *Annalen der Physik (Berlin, Germany)*, vol. 322, no. 8, pp. 549–560, 1905, ISSN: 1521-3889.

[85] R. Brown, "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.," *Philosophical Magazine*, vol. 4, no. 21, pp. 161–173, 1828.

[86] L. C. Evans and R. F. Gariepy, *Measure theory and fine properties of functions*. Chapman and Hall/CRC, 2015.

[87] L. E. Ermak and J. A. McCammon, "Brownian dynamics with hydrodynamic interactions," *Journal of Chemical Physics*, vol. 69, no. 4, pp. 1352–1360, 1978.

[88] J. K. Percus, "Equilibrium state of a classical fluid of hard rods in an external field," *Journal of Statistical Physics*, vol. 15, no. 6, pp. 505–511, Dec. 1976, ISSN: 1572-9613.

[89] R. Roth, "Fundamental measure theory for hard-sphere mixtures: A review," *Journal of Physics: Condensed Matter*, vol. 22, no. 6, p. 063 102, 2010.

[90] O. R. Walton, "Numerical simulation of inclined chute flows of monodisperse, inelastic, frictional spheres," *Mechanics of materials*, vol. 16, no. 1-2, pp. 239–247, 1993.

[91] M. Wilkinson, *On global-in-time chaotic weak solutions of the Liouville equation for hard spheres*, 2018. arXiv: 1807.01017.

[92] M. N. Bannerman, T. E. Green, P. Grassia, and L. Lue, "Collision statistics in sheared inelastic hard spheres," *Physical Review E*, vol. 79, no. 4, p. 041 308, 2009.

[93] J. E. Lennard-Jones, "The determination of molecular orbitals," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 198, no. 1052, pp. 1–13, 1949.

[94] C. Thomson, L. Lue, and M. N. Bannerman, "Mapping continuous potentials to discrete forms," *Journal of Chemical Physics*, vol. 140, no. 3, p. 034 105, 2014.

[95] Z. Zhang, X. Liu, Z. Chen, H. Zheng, K. Yan, and J. Liu, "A unified thermostat scheme for efficient configurational sampling for classical/quantum canonical ensembles via molecular dynamics," *The Journal of chemical physics*, vol. 147, no. 3, p. 034 109, 2017.

[96] L. Euler, *Institutionum calculi integralis*. impensis Academiae imperialis scientiarum, 1824, vol. 1.

[97] D. F. Griffiths and D. J. Higham, *Numerical methods for ordinary differential equations: initial value problems*. Springer Science & Business Media, 2010.

[98] B. Leimkuhler and C. Matthews, "Efficient molecular dynamics using geodesic integration and solvent–solute splitting," *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 472, no. 2189, p. 20 160 138, 2016.

[99] C. Grindon, S. Harris, T. Evans, K. Novik, P. Coveney, and C. Laughton, "Large-scale molecular dynamics simulation of DNA: Implementation and validation of the AMBER98 force field in LAMMPS," *Philosophical Transactions of the Royal Society of London. Series A: Mathematical, Physical and Engineering Sciences*, vol. 362, no. 1820, pp. 1373–1386, 2004.

[100] P. Eastman, M. S. Friedrichs, J. D. Chodera, R. J. Radmer, C. M. Bruns, J. P. Ku, K. A. Beauchamp, T. J. Lane, L.-P. Wang, D. Shukla, *et al.*, "OpenMM 4: A reusable, extensible, hardware independent library for high performance molecular simulation," *Journal of chemical theory and computation*, vol. 9, no. 1, pp. 461–469, 2013.

[101] D. Van Der Spoel, E. Lindahl, B. Hess, G. Groenhof, A. E. Mark, and H. J. Berendsen, "GROMACS: Fast, flexible, and free," *Journal of computational chemistry*, vol. 26, no. 16, pp. 1701–1718, 2005.

[102] D. C. Rapaport and D. C. R. Rapaport, *The art of molecular dynamics simulation*. Cambridge university press, 2004.

[103] Y. A. Houndonougbo, B. B. Laird, and B. J. Leimkuhler, "A molecular dynamics algorithm for mixed hard-core/continuous potentials," *Molecular Physics*, vol. 98, no. 5, pp. 309–316, 2000.

[104] C. Thornton, *Granular dynamics, contact mechanics and particle system simulations*. Springer, 2015.

[105] S. Yuu, T. Umekage, and Y. Johno, "Numerical simulation of air and particle motions in bubbling fluidized bed of small particles," *Powder Technology*, vol. 110, no. 1-2, pp. 158–168, 2000.

[106] M. Sakai, K. Shibata, and S. Koshizuka, "Development of a criticality evaluation method involving the granular flow of the nuclear fuel in a rotating drum," *Nuclear science and engineering*, vol. 154, no. 1, pp. 63–73, 2006.

[107] Y. Tsuji, T. Tanaka, and T. Ishida, "Lagrangian numerical simulation of plug flow of cohesionless particles in a horizontal pipe," *Powder technology*, vol. 71, no. 3, pp. 239–250, 1992.

[108] S. Hentz, L. Daudeville, and F. V. Donzé, "Identification and validation of a discrete element model for concrete," *Journal of engineering mechanics*, vol. 130, no. 6, pp. 709–719, 2004.

[109] R. M. Corless, G. H. Gonnet, D. E. Hare, D. J. Jeffrey, and D. E. Knuth, "On the Lambert-W function," *Advances in Computational Mathematics*, vol. 5, no. 1, pp. 329–359, 1996.

[110] A. Donev, "Asynchronous event-driven particle algorithms," *Simulation*, vol. 85, no. 4, pp. 229–242, 2009.

[111] A. Donev, S. Torquato, and F. H. Stillinger, "Neighbor list collision-driven molecular dynamics simulation for nonspherical hard particles," *Journal of Computational Physics*, vol. 202, no. 2, pp. 765–793, 2005.

[112] A. A. Chialvo and P. G. Debenedetti, "On the use of the Verlet neighbor list in molecular dynamics," *Computer physics communications*, vol. 60, no. 2, pp. 215–224, 1990.

[113] *Binary trees*, Accessed: 19/07/2017. [Online]. Available: http://cslibrary.stanford.edu/110/BinaryTrees.html.

[114] M. A. Khan and M. C. Herbordt, "Parallel discrete molecular dynamics simulation with speculation and in-order commitment," *Journal of Computational Physics*, vol. 230, no. 17, pp. 6563–6582, 2011, ISSN: 0021-9991.

[115] M. N. Bannerman, S. Strobl, A. Formella, and T. Pöschel, "Stable algorithm for event detection in event-driven particle dynamics," *Computational Particle Mechanics*, vol. 1, no. 2, pp. 191–198, 2014.

[116] S. Luding and S. McNamara, "How to handle the inelastic collapse of a dissipative hard-sphere gas with the TC model," *Granular Matter*, vol. 1, no. 3, pp. 113–128, 1998.

[117] M. N. Bannerman, R. Sargant, and L. Lue, "DynamO: A free $\mathcal{O}(N)$ general event-driven molecular dynamics simulator," *Journal of Computational Chemistry*, vol. 32, no. 15, pp. 3329–3338, 2011.

[118] D. J Evans and G. P Morriss, *Statistical mechanics of nonequilbrium liquids*. ANU Press, 2007.

[119] R. Balescu, "Equilibrium and nonequilibrium statistical mechanics," *NASA STI/Recon Technical Report A*, vol. 76, p. 756, May 1975.

[120] K. Huang, *Statistical mechanics*. Wiley, 1987.

[121] C. Cercignani, R. Illner, and M. Pulvirenti, *The mathematical theory of dilute gases*. Springer Science & Business Media, 2013, vol. 106.

[122] C. Cercignani, "The Boltzmann equation," in *The Boltzmann equation and its applications*, Springer, 1988, pp. 40–103.

[123] J. J. Brey, J. W. Dufty, and A. Santos, "Dissipative dynamics for hard spheres," *Journal of statistical physics*, vol. 87, no. 5-6, pp. 1051–1066, 1997.

[124] M. Ernst, J. Dorfman, W. Hoegy, and J. Van Leeuwen, "Hard-sphere dynamics and binary-collision operators," *Physica*, vol. 45, no. 1, pp. 127–146, 1969.

[125] G. Pavliotis and A. Stuart, *Multiscale methods: averaging and homogenization*. Springer Science & Business Media, 2008.

[126] M. Protter and C. Morrey, *Intermediate Calculus*, ser. Undergraduate Texts in Mathematics. Springer New York, 2012.

[127] J. E. Marsden and A. Tromba, *Vector calculus*. Macmillan, 2003.

[128] H. v. Beijeren and M. H. Ernst, "The modified Enskog equation," *Physica*, vol. 68, no. 3, pp. 437–456, 1973.

[129] V. Garzó, C. M. Hrenya, and J. W. Dufty, "Enskog theory for polydisperse granular mixtures. II. Sonine polynomial approximation," *Physical Review E.*, vol. 76, no. 3, p. 031 304, 2007.

[130] L. Pitaevskii and E. Lifshitz, *Physical kinetics*. Butterworth-Heinemann, 2012, vol. 10.

[131] J. Yvon, *La théorie statistique des fluides et l'équation d'état*. Hermann & cie, 1935, vol. 203.

[132] N. N. Bogoliubov, "Kinetic equations," *Journal of Physics USSR*, vol. 10, no. 3, pp. 265–274, 1946.

[133] J. G. Kirkwood, "The statistical mechanical theory of transport processes I. General theory," *The Journal of Chemical Physics*, vol. 14, no. 3, pp. 180–201, 1946.

[134] M. Born and H. Green, "A general kinetic theory of liquids I. The molecular distribution functions," *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, vol. 188, no. 1012, pp. 10–18, 1946.

[135] C. Cercignani, V. Gerasimenko, and D. Y. Petrina, "The BBGKY hierarchy," in *Many-Particle Dynamics and Kinetic Equations*, Springer, 1997, pp. 7–65.

[136] M. Tessarotto and C. Cremaschini, "Modified BBGKY hierarchy for the hard-sphere system," *The European Physical Journal Plus*, vol. 129, no. 11, p. 243, 2014.

[137] V. Gerasimenko and D. Y. Petrina, "The generalized kinetic equation generated by the BBGKY hierarchy," *Ukrayins' kij Fyizichnij Zhurnal*, vol. 43, no. 6-7, pp. 697–702, 1998.

[138] W. Cassing and A. Pfitzner, "Self-consistent truncation of the BBGKY hierarchy on the two-body level," *Zeitschrift für Physik A Hadrons and Nuclei*, vol. 342, no. 2, pp. 161–167, 1992.

[139] J. Hansen and I. R. McDonald, *Theory of simple liquids*. Elsevier, 1990.

[140] L. Socha, *Linearization Methods for Stochastic Dynamic Systems*. Springer, 207.

[141] H. Grad, "On the kinetic theory of rarefied gases," *Communications on pure and applied mathematics*, vol. 2, no. 4, pp. 331–407, 1949.

[142] H. Struchtrup and M. Torrilhon, "Regularization of Grad's 13 moment equations: Derivation and linear analysis," *Physics of Fluids*, vol. 15, no. 9, pp. 2668–2680, 2003.

[143] C. Kuehn, "Moment closure-—a brief review," in *Control of self-organizing nonlinear systems*, Springer, 2016, pp. 253–271.

[144] G. K. Chan and R. Finken, "Time-dependent density functional theory of classical fluids," *Physical Review Letters*, vol. 94, p. 183 001, 18 May 2005.

[145] M. A. Durán-Olivencia, B. D. Goddard, and S. Kalliadasis, "Dynamical density functional theory for orientable colloids including inertia and hydrodynamic interactions," *The Journal of Statistical Physics*, vol. 164, no. 4, pp. 785–809, 2016.

[146] D. Henderson, *Fundamentals of inhomogeneous fluids*. CRC Press, 1992.

[147] D. Benedetto, E. Caglioti, J. A. Carrillo, and M. Pulvirenti, "A non-Maxwellian steady distribution for one-dimensional granular media," *Journal of Statistical Physics*, vol. 91, no. 5-6, pp. 979–990, 1998.

[148] B. 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," *Journal of Physics: Condensed Matter*, vol. 25, no. 3, p. 035 101, 2012.

[149] J. F. Lutsko, "Recent developments in classical density functional theory," *Advances in chemical physics*, vol. 144, p. 1, 2010.

[150] L. Huilin, D. Gidaspow, J. Bouillard, and L. Wentie, "Hydrodynamic simulation of gas–solid flow in a riser using kinetic theory of granular flow," *Chemical Engineering Journal*, vol. 95, no. 1-3, pp. 1–13, 2003.

[151] J. W. Dufty, *Kinetic theory and hydrodynamics for rapid granular flow - a perspective*, 2001. arXiv: cond-mat/0108444 [cond-mat.stat-mech].

[152] J. Ibsen, P. Cordero, and R. Tabensky, "Hard rods in the presence of a uniform external field," *Journal of Chemical Physics*, vol. 107, no. 14, pp. 5515–5523, 1997.

[153] Y. Song, R. M. Stratt, and E. Mason, "The equation of state of hard spheres and the approach to random closest packing," *Journal of Chemical Physics*, vol. 88, no. 2, pp. 1126–1133, 1988.

[154] J. Tobochnik and P. M. Chapin, "Monte Carlo simulation of hard spheres near random closest packing using spherical boundary conditions," *Journal of Chemical Physics*, vol. 88, no. 9, pp. 5824–5830, 1988.

[155] C. H. Reinsch, "Smoothing by spline functions. II," *Numerische Mathematik*, vol. 16, no. 5, pp. 451–454, Feb. 1971, ISSN: 0945-3245. DOI: 10.1007/BF02169154. [Online]. Available: https://doi.org/10.1007/BF02169154.

[156] S. Savage and M. Sayed, "Stresses developed by dry cohesionless granular materials sheared in an annular shear cell," *Journal of Fluid Mechanics*, vol. 142, pp. 391–430, 1984.

[157] R. Wildman, T. Martin, J. Huntley, J. Jenkins, H. Viswanathan, X. Fen, and D. Parker, "Experimental investigation and kinetic-theory-based model of a rapid granular shear flow," *Journal of Fluid Mechanics*, vol. 602, pp. 63–79, 2008.

[158] C. Lun and A. Bent, "Numerical simulation of inelastic frictional spheres in simple shear flow," *Journal of Fluid Mechanics*, vol. 258, pp. 335–353, 1994.

[159] C. S. Campbell, "The stress tensor for simple shear flows of a granular material," *Journal of Fluid Mechanics*, vol. 203, pp. 449–473, 1989.

[160] B. Hoomans, J. Kuipers, W. J. Briels, and W. P. M. van Swaaij, "Discrete particle simulation of bubble and slug formation in a two-dimensional gas-fluidised bed: A hard-sphere approach," *Chemical Engineering Science*, vol. 51, no. 1, pp. 99–118, 1996.

[161] B. Hoomans, J. Kuipers, and W. P. M. van Swaaij, "Granular dynamics simulation of segregation phenomena in bubbling gas-fluidised beds," *Powder Technology*, vol. 109, no. 1-3, pp. 41–48, 2000.

[162] L. Trefethen, *Spectral Methods in MATLAB*. Society for Industrial and Applied Mathematics, 2000.

[163] K. Hoffman and R. Kunze, *Linear Algebra*, ser. Featured Titles for Linear Algebra (Advanced) Series. Prentice-Hall, 1971.

[164] J. Boyd, *Chebyshev and Fourier Spectral Methods: Second Revised Edition*, ser. Dover Books on Mathematics. Dover Publications, 2001.

[165] A. Nold, B. D. Goddard, P. Yatsyshin, N. Savva, and S. Kalliadasis, "Pseudospectral methods for density functional theory in bounded and unbounded domains," *Journal of Computational Physics*, vol. 334, pp. 639–664, 2017.

[166] H. Jasak, A. Jemcov, Z. Tukovic, *et al.*, "OpenFOAM: A C++ library for complex physics simulations," in *International workshop on coupled methods in numerical dynamics*, IUC Dubrovnik Croatia, vol. 1000, 2007, pp. 1–20.

[167] R. W. Pryor, *Multiphysics modeling using COMSOL®: a first principles approach*. Jones & Bartlett Publishers, 2009.

Appendix A

Computationally efficient forms of the collision operator

Considering the i^{th} moments of the collision operator in eq. (6.6) resulted in three integrals eqs. (6.9) to (6.11) that involve characteristic functions, which make them difficult to solve analytically:

$$\boldsymbol{A}_1 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n},$$
$$\boldsymbol{A}_2 = \int_{\mathbb{S}^{d-1}} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[\boldsymbol{n}\otimes\boldsymbol{v}_1+\boldsymbol{v}_1\otimes\boldsymbol{n}][(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^2 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n},$$
$$\boldsymbol{A}_3 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n}\otimes\boldsymbol{n} \int_{\mathbb{R}^{2d}} \chi_{(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}<0}[(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\boldsymbol{n}]^3 f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t)\,\mathrm{d}\boldsymbol{v}^2\,\mathrm{d}\boldsymbol{n}.$$

Under the assumption on the correlation function, and the Maxwellian assumption on the local equilibrium, we can write $f^{(2)}$ as a product of Gaussians in velocity:

$$f^{(2)}([\boldsymbol{r}_1,\boldsymbol{r}_1+\varepsilon\boldsymbol{n}],\boldsymbol{v}^2,t) = \frac{\rho\rho_+ g}{|2\pi m k_B T\boldsymbol{E}|^{\frac{1}{2}}|2\pi m k_B T\boldsymbol{E}_+|^{\frac{1}{2}}} \tag{A.1}$$
$$\times \exp\left(-\frac{1}{2mk_BT}(\boldsymbol{v}-m\bar{\boldsymbol{p}})^T\boldsymbol{E}^{-1}(\boldsymbol{v}-m\bar{\boldsymbol{p}})\right)\exp\left(-\frac{1}{2mk_BT}(\boldsymbol{v}-m\bar{\boldsymbol{p}}_+)^T\boldsymbol{E}_+^{-1}(\boldsymbol{v}-m\bar{\boldsymbol{p}}_+)\right)$$

For the following arguments we will set $m=1, k_BT=1$ to simplify notation.

A.0.1 Useful integral identities

For the derivations in one and two dimensions, the following integrals will be of use. Firstly we recall the moments of Gaussian functions:

$$\int_{-\infty}^{\infty}\exp\left(-\frac{(x-\mu)^2}{\xi}\right)\mathrm{d}x = \sqrt{\pi\xi},\quad \int_{-\infty}^{\infty}x\exp\left(-\frac{(x-\mu)^2}{\xi}\right)\mathrm{d}x = \mu\sqrt{\pi\xi},$$
$$\int_{-\infty}^{\infty}x^2\exp\left(-\frac{(x-\mu)^2}{\xi}\right)\mathrm{d}x = \frac{\sqrt{\pi\xi}}{2}(\xi+2\mu^2).$$
$$\int_{-\infty}^{\infty}x^3\exp\left(-\frac{(x-\mu)^2}{\xi}\right)\mathrm{d}x = \frac{\sqrt{\pi\xi}}{2}\mu(2\mu^2+3\xi). \tag{A.2}$$

To deal with the characteristic function in the integrals we have the following identities involving error functions:

$$\begin{aligned}
\int_{-\infty}^{\infty} \chi_{\pm x>0} \exp\left(-\frac{(x-\mu)^2}{\xi}\right) \mathrm{d}x &= \frac{\sqrt{\pi\xi}}{2}\left(1 \pm \operatorname{erf}\left(\frac{\mu}{\sqrt{\xi}}\right)\right), \\
\int_{-\infty}^{\infty} \chi_{\pm x>0} x \exp\left(-\frac{(x-\mu)^2}{\xi}\right) \mathrm{d}x &= \frac{\mu\sqrt{\pi\xi}}{2}\left(1 \pm \operatorname{erf}\left(\frac{\mu}{\sqrt{\xi}}\right)\right) \pm \frac{\xi}{2}\exp\left(-\frac{\mu^2}{\xi}\right), \\
\int_{-\infty}^{\infty} \chi_{\pm x>0} x^2 \exp\left(-\frac{(x-\mu)^2}{\xi}\right) \mathrm{d}x &= \frac{\sqrt{\pi\xi}}{2}\left(1 \pm \operatorname{erf}\left(\frac{\mu}{\sqrt{\xi}}\right)\right)\left(\frac{\xi}{2}+\mu^2\right) \pm \frac{\mu\xi}{2}\exp\left(-\frac{\mu^2}{\xi}\right), \\
\int_{-\infty}^{\infty} \chi_{\pm x>0} x^3 \exp\left(-\frac{(x-\mu)^2}{\xi}\right) \mathrm{d}x &= \frac{\mu\sqrt{\pi\xi}}{2}\left(\frac{3\xi}{2}+\mu^2\right)\left(1 \pm \operatorname{erf}\left(\frac{\mu}{\sqrt{\xi}}\right)\right) \pm \frac{\xi}{2}\exp\left(-\frac{\mu^2}{\xi}\right)(\mu^2+\xi).
\end{aligned} \tag{A.3}$$

Where we have defined the error function as

$$\operatorname{erf}(x) := \frac{2}{\sqrt{\pi}}\int_0^x \exp(-t^2)\,\mathrm{d}t. \tag{A.4}$$

Finally we consider moments of products of Gaussians and error functions. We have that for $a \neq 0, b \in \mathbb{R}$ [1]:

$$\int_{-\infty}^{\infty} \operatorname{erf}(x)\exp(-(ax+b)^2)\,\mathrm{d}x = -\frac{\sqrt{\pi}}{a}\operatorname{erf}\left(\frac{b}{\sqrt{a^2+1}}\right) \tag{A.5}$$

Then by using integration by parts with

$$\begin{aligned}
u = \operatorname{erf}(x), \quad \mathrm{d}u = \frac{2}{\sqrt{pi}}\exp(-x^2), \\
\mathrm{d}v = x\exp(-(ax+b)^2), \quad v = -\frac{1}{2a^2}[\sqrt{\pi}b\operatorname{erf}(ax+b) + \exp(-(ax+b)^2)],
\end{aligned}$$

we have

$$\int_{-\infty}^{\infty} x\operatorname{erf}(x)\exp(-(ax+b)^2)\,\mathrm{d}x = \frac{\sqrt{\pi}b}{a^2}\operatorname{erf}\left(\frac{b}{\sqrt{a^2+1}}\right) + \frac{1}{a^2\sqrt{a^2+1}}\exp\left(-\frac{b^2}{a^2+1}\right) \tag{A.6}$$

With a similar choice of u and $\mathrm{d}v$ we have that

$$\begin{aligned}
\int_{-\infty}^{\infty} x^2\operatorname{erf}(x)\exp(-(ax+b)^2)\,\mathrm{d}x = &-\frac{\sqrt{\pi}(2b^2+1)}{2a^3}\operatorname{erf}\left(\frac{b}{\sqrt{a^2+1}}\right) \\
&- \frac{b(2a^2+1)}{a^3(a^2+1)^{3/2}}\exp\left(-\frac{b^2}{a^2+1}\right),
\end{aligned} \tag{A.7}$$

$$\begin{aligned}
\int_{-\infty}^{\infty} x^3\operatorname{erf}(x)\exp(-(ax+b)^2)\,\mathrm{d}x = &\frac{\sqrt{\pi}b(2b^2+3)}{2a^4}\operatorname{erf}\left(\frac{b}{\sqrt{a^2+1}}\right) \\
&+ \frac{1}{2a^4(a^2+1)^{5/2}}\left[b^2(6a^4+6a^2+2)+(3a^4+5a^2+2)\right]\exp\left(-\frac{b^2}{a^2+1}\right)
\end{aligned} \tag{A.8}$$

A.0.2 One dimension

In one dimension the integral over $\mathbb{S}^{d-1}$ reduces to a sum over ± 1. We have that, after a change of variables $\boldsymbol{v}_1 \to (\boldsymbol{v}_1 - \boldsymbol{v}_2)$,

$$\begin{aligned} A_1 =& \frac{\rho\rho_+ g_+}{|\pi E|^{1/2}|\pi E_+|^{1/2}} \int_{\mathbb{R}^2} \chi_{v_1<0} v_1^2 \exp\left(-\frac{(v_1+v_2-\bar{v})^2}{E}\right) \exp\left(-\frac{(v_2-\bar{v}_+)^2}{E_+}\right) \mathrm{d}v^2 \\ & - \frac{\rho\rho_- g_-}{|\pi E|^{1/2}|\pi E_-|^{1/2}} \int_{\mathbb{R}^2} \chi_{v_1>0} v_1^2 \exp\left(-\frac{(v_1+v_2-\bar{v})^2}{E}\right) \exp\left(-\frac{(v_2-\bar{v}_-)^2}{E_-}\right) \mathrm{d}v^2 \end{aligned}$$

We then rearrange the exponential terms to get v_2 into a single Gaussian:

$$\begin{aligned} &\exp\left(-\frac{(v_1+v_2-\bar{v})^2}{E}\right) \exp\left(-\frac{(v_2-\bar{v}_\pm)^2}{E_\pm}\right) \\ &= \\ &\exp\left[-\frac{1}{2(E+E_\pm)}\left(v_2 + \frac{E_\pm}{E+E_\pm}\left(v_1 - \left(\bar{v} - \frac{E}{E_\pm}\bar{v}_\pm\right)\right)\right)^2\right] \exp\left[-\frac{1}{2(E+E_\pm)}\left(v_1 - (\bar{v}-\bar{v}_\pm)\right)^2\right] \end{aligned} \tag{A.9}$$

We then use eq. (A.2) to solve the integral in v_2 in each term in A_1:

$$\begin{aligned} A_1 =& \frac{\rho\rho_+ g_+}{\sqrt{2\pi(E+E_+)}} \int_{\mathbb{R}} \chi_{v_1>0} v_1^2 \exp\left[-\frac{(v_1-(\bar{v}-\bar{v}_+))^2}{2(E+E_+)}\right] \mathrm{d}v_1 \\ & - \frac{\rho\rho_- g_-}{\sqrt{2\pi(E+E_-)}} \int_{\mathbb{R}} \chi_{v_1<0} v_1^2 \exp\left[-\frac{(v_1-(\bar{v}-\bar{v}_-))^2}{2(E+E_-)}\right] \mathrm{d}v_1 \end{aligned}$$

We then use eq. (A.3) to solve the remaining integrals in v_1:

$$\begin{aligned} A_1 =& \frac{\rho\rho_+ g_+}{2}\left[((E+E_+)+(\bar{v}-\bar{v}_+)^2)\left(1-\operatorname{erf}\left(\frac{\bar{v}-\bar{v}_+}{\sqrt{2(E+E_+)}}\right)\right)\right. \\ & \qquad \left. - (\bar{v}-\bar{v}_+)\sqrt{\frac{(E+E_+)}{2\pi}} \exp\left(-\frac{(\bar{v}-\bar{v}_+)^2}{2(E+E_+)}\right)\right] \\ & - \frac{\rho\rho_- g_-}{2}\left[((E+E_-)+(\bar{v}-\bar{v}_-)^2)\left(1+\operatorname{erf}\left(\frac{\bar{v}-\bar{v}_-}{\sqrt{2(E+E_-)}}\right)\right)\right. \\ & \qquad \left. + (\bar{v}-\bar{v}_-)\sqrt{\frac{(E+E_-)}{2\pi}} \exp\left(-\frac{(\bar{v}-\bar{v}_-)^2}{2(E+E_-)}\right)\right] \end{aligned} \tag{A.10}$$

Which fully defines A_1 in terms of error functions and exponentials. We note that the terms with $\bar{v} - \bar{v}_\pm$ disappear as $\varepsilon \to 0$, and the remaining terms in E cancel. Thus the we are left with zero in the second moment of the collision operator in the limit as the particle size vanishes, as expected.

In one dimension we have that

$$\begin{aligned} A_2 =& 2\int_{\mathbb{R}^2} \chi_{(v_1-v_2)<0} v_1 (v_1-v_2)^2 f^{(2)}(r, r+\varepsilon, v_1, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2 \\ & - 2\int_{\mathbb{R}^2} \chi_{(v_1-v_2)>0} v_1 (v_1-v_2)^2 f^{(2)}(r, r-\varepsilon, v_1, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2, \end{aligned} \tag{A.11}$$

$$A_3 = \int_{\mathbb{R}^2} \chi_{(v_1-v_2)<0} (v_1-v_2)^3 f^{(2)}(r, r+\varepsilon, v_1, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2 \tag{A.12}$$

$$- \int_{\mathbb{R}^2} \chi_{(v_1-v_2)<0}(v_1-v_2)^3 f^{(2)}(r, r-\varepsilon, v_1, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2.$$

Using the substitution $\tilde{v} = v_1 - v_2$ we see that

$$A_2 = 2\tilde{A}_1 + 2\tilde{A}_2, \quad A_3 = \tilde{A}_1 \tag{A.13}$$

where

$$\begin{aligned}\tilde{A}_1 = &\int_{\mathbb{R}^2} \chi_{(v_1-v_2)<0} v_1^3 f^{(2)}(r, r+\varepsilon, v_1+v_2, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2 \\ &- \int_{\mathbb{R}^2} \chi_{(v_1-v_2)>0} v_1^3 f^{(2)}(r, r-\varepsilon, v_1+v_2, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2,\end{aligned} \tag{A.14}$$

$$\begin{aligned}\tilde{A}_2 = &\int_{\mathbb{R}^2} \chi_{(v_1-v_2)<0} v_2 v_1^2 f^{(2)}(r, r+\varepsilon, v_1+v_2, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2 \\ &- \int_{\mathbb{R}^2} \chi_{(v_1-v_2)>0} v_2 v_1^2 f^{(2)}(r, r-\varepsilon, v_1+v_2, v_2, t)\,\mathrm{d}v_1\,\mathrm{d}v_2.\end{aligned} \tag{A.15}$$

For $\tilde{A}_1$ we use eq. (A.9) to solve the integral in v_2, so that we are left with

$$\begin{aligned}\tilde{A}_1 = &\frac{\rho\rho_+ g_+}{\sqrt{2\pi(E+E_+)}} \int_{\mathbb{R}} \chi_{v_1<0} v_1^3 \exp\left[-\frac{(v_1-(\bar{v}-\bar{v}_+))^2}{2(E+E_+)}\right]\mathrm{d}v_1 \\ &- \frac{\rho\rho_- g_-}{\sqrt{2\pi(E+E_-)}} \int_{\mathbb{R}} \chi_{v_1>0} v_1^3 \exp\left[-\frac{(v_1-(\bar{v}-\bar{v}_-))^2}{2(E+E_-)}\right]\mathrm{d}v_1.\end{aligned} \tag{A.16}$$

Then, using the third partial moment in eq. (A.3), we find

$$\begin{aligned}\tilde{A}_1 = &\rho\rho_+ g_+ \left[\frac{(\bar{v}-\bar{v}_+)}{2}(3(E+E_+)+(\bar{v}-\bar{v}_+)^2)\left(1-\operatorname{erf}\left(\frac{(\bar{v}-\bar{v}_+)}{\sqrt{2(E+E_+)}}\right)\right)\right. \\ &\left. - \sqrt{\frac{E+E_+}{2\pi}}((\bar{v}-\bar{v}_+)^2+2(E+E_+))\exp\left(-\frac{(\bar{v}-\bar{v}_+)^2}{2(E+E_+)}\right)\right] \\ &- \rho\rho_- g_- \left[\frac{(\bar{v}-\bar{v}_-)}{2}(3(E+E_-)+(\bar{v}-\bar{v}_-)^2)\left(1+\operatorname{erf}\left(\frac{(\bar{v}-\bar{v}_-)}{\sqrt{2(E+E_-)}}\right)\right)\right. \\ &\left. + \sqrt{\frac{E+E_+}{2\pi}}((\bar{v}-\bar{v}_-)^2+2(E+E_-))\exp\left(-\frac{(\bar{v}-\bar{v}_-)^2}{2(E+E_-)}\right)\right].\end{aligned} \tag{A.17}$$

For $\tilde{A}_2$, we use eq. (A.9) and solve the integral in v_2 using the second moment in eq. (A.2) to find:

$$\begin{aligned}\tilde{A}_2 = &-\frac{\rho\rho_+}{\sqrt{2\pi(E+E_+)}} \int_{\mathbb{R}} \chi_{v_1<0} \frac{E_+}{E+E_+}\left[v_1-\left(\bar{v}-\frac{E\bar{v}_+}{E_+}\right)\right] v_1^2 \exp\left[-\frac{(v_1-(\bar{v}-\bar{v}_+))^2}{2(E+E_+)}\right]\mathrm{d}v_1 \\ &+ \frac{\rho\rho_-}{\sqrt{2\pi(E+E_-)}} \int_{\mathbb{R}} \chi_{v_1>0} \frac{E_-}{E+E_-}\left[v_1-\left(\bar{v}-\frac{E\bar{v}_-}{E_-}\right)\right] v_1^2 \exp\left[-\frac{(v_1-(\bar{v}-\bar{v}_-))^2}{2(E+E_-)}\right]\mathrm{d}v_1 \\ = &-\frac{E_+}{E+E_+}\tilde{A}_1^+ + \frac{E_+\bar{v}-E\bar{v}_+}{E+E_+}A_1^+ + \frac{E_-}{E+E_-}\tilde{A}_1^- + \frac{E_-\bar{v}-E\bar{v}_-}{E+E_-}A_1^-,\end{aligned}$$

where $\tilde{A}_1^{\pm}$ are the positive and negative diameter contributions of $\tilde{A}_1$, and similar for A_1. This fully determines all contributions to the first three moments of the collision operator in terms of

error functions. We note that when $\varepsilon \to 0$ we have that

$$\lim_{\varepsilon\to 0} \tilde{A}_1 = -\frac{8\rho^2 g E^{3/2}}{\sqrt{\pi}},$$
$$\lim_{\varepsilon\to 0} \tilde{A}_2 = -\frac{1}{2}\lim_{\varepsilon\to 0}\tilde{A}_1 + \bar{v}\lim_{\varepsilon\to 0} A_1 = \frac{4\rho^2 g E^{3/2}}{\sqrt{\pi}}.$$

In the third moment the terms of interest are of the form

$$\begin{aligned}-\frac{1+\alpha}{2\sqrt{2}}A_2 + \frac{(1+\alpha)^2}{4\sqrt{2}}A_3 =& -\frac{1+\alpha}{2\sqrt{2}}(2\tilde{A}_1 + 2\tilde{A}_2) + \frac{(1+\alpha)^2}{4\sqrt{2}}\tilde{A}_1 \\ \to& -\frac{1+\alpha}{2\sqrt{2}}\lim_{\varepsilon\to 0}\tilde{A}_1 + \frac{(1+\alpha)^2}{4\sqrt{2}}\lim_{\varepsilon\to 0}\tilde{A}_1 \\ =& \frac{2\rho^2 g E^{3/2}}{\sqrt{2\pi}}(1-\alpha^2),\end{aligned}$$

which shows that the affect of inelastic collisions on the kinetic energy does not vanish in the limit as the particle size vanishes, as expected.

A.0.3 Two dimensions

When $d = 2$, we note that

$$(\boldsymbol{v}_1 \cdot \boldsymbol{n})^i = (\boldsymbol{v}_1^{\otimes i}) \cdot^i (\boldsymbol{n}^{\otimes i}),$$

where $\cdot^i$ represents the scalar product between all i dimensions of the matrix formed by the tensor products. Thus the integrals of interest are then

$$A_1 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n}\tilde{A}_1 \cdot^2 \boldsymbol{n}\otimes\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}, \tag{A.18}$$
$$A_2 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n}\otimes[(\tilde{A}_2 + \tilde{A}_3)\cdot^2(\boldsymbol{n}\otimes\boldsymbol{n})] + [(\tilde{A}_2+\tilde{A}_3)\cdot^2(\boldsymbol{n}\otimes\boldsymbol{n})]\otimes\boldsymbol{n}\,\mathrm{d}\boldsymbol{n}, \tag{A.19}$$
$$A_3 = \int_{\mathbb{S}^{d-1}} \boldsymbol{n}\otimes\boldsymbol{n}\tilde{A}_2 \cdot^3 (\boldsymbol{n}\otimes\boldsymbol{n}\otimes\boldsymbol{n})\,\mathrm{d}\boldsymbol{n}, \tag{A.20}$$

where

$$\begin{aligned}\tilde{A}_1 =& \int_{\mathbb{R}^d} \exp\left[-\frac{1}{2}(\boldsymbol{v}_2-\bar{\boldsymbol{v}}_+)^T E_+^{-1}(\boldsymbol{v}_2-\bar{\boldsymbol{v}}_+)\right] \\ &\times\left\{\int_{\mathbb{R}^d}\chi_{\boldsymbol{v}_1\cdot n<0}(\boldsymbol{v}_1\otimes\boldsymbol{v}_1)\exp\left[-\frac{1}{2}(\boldsymbol{v}_1+\boldsymbol{v}_2-\bar{\boldsymbol{v}})^T E^{-1}(\boldsymbol{v}_1+\boldsymbol{v}_2-\bar{\boldsymbol{v}})\right]\mathrm{d}\boldsymbol{v}_1\right\}\mathrm{d}\boldsymbol{v}_2 \\ \tilde{A}_2 =& \int_{\mathbb{R}^d} \exp\left[-\frac{1}{2}(\boldsymbol{v}_2-\bar{\boldsymbol{v}}_+)^T E_+^{-1}(\boldsymbol{v}_2-\bar{\boldsymbol{v}}_+)\right] \\ &\times\left\{\int_{\mathbb{R}^d}\chi_{\boldsymbol{v}_1\cdot n<0}\boldsymbol{v}_1\otimes\boldsymbol{v}_1\otimes\boldsymbol{v}_1\exp\left[-\frac{1}{2}(\boldsymbol{v}_1+\boldsymbol{v}_2-\bar{\boldsymbol{v}})^T E^{-1}(\boldsymbol{v}_1+\boldsymbol{v}_2-\bar{\boldsymbol{v}})\right]\mathrm{d}\boldsymbol{v}_1\right\}\mathrm{d}\boldsymbol{v}_2, \\ \tilde{A}_3 =& \int_{\mathbb{R}^d} \exp\left[-\frac{1}{2}(\boldsymbol{v}_2-\bar{\boldsymbol{v}}_+)^T E_+^{-1}(\boldsymbol{v}_2-\bar{\boldsymbol{v}}_+)\right] \\ &\times\left\{\int_{\mathbb{R}^d}\chi_{\boldsymbol{v}_1\cdot n<0}\boldsymbol{v}_2\otimes\boldsymbol{v}_1\otimes\boldsymbol{v}_1\exp\left[-\frac{1}{2}(\boldsymbol{v}_1+\boldsymbol{v}_2-\bar{\boldsymbol{v}})^T E^{-1}(\boldsymbol{v}_1+\boldsymbol{v}_2-\bar{\boldsymbol{v}})\right]\mathrm{d}\boldsymbol{v}_1\right\}\mathrm{d}\boldsymbol{v}_2.\end{aligned}$$

We then note that the product of two Gaussian densities is also Gaussian: given two multivariate Gaussians with means μ_1, μ_2 and covariance matrices σ_1, σ_2, their product can be written as:

$$\begin{aligned}
&\exp\left(-\frac{1}{2}(x-\mu_1)^T\sigma_1^{-1}(x-\mu_1)\right)\exp\left(-\frac{1}{2}(x-\mu_2)^T\sigma_2^{-1}(x-\mu_2)\right)\\
=&\exp\left(-\frac{1}{2}(\mu_1-\mu_2)^T(\sigma_1+\sigma_2)^{-1}(\mu_1-\mu_2)\right)\\
&\times\exp\Bigg[-\frac{1}{2}(x-(\sigma_1^{-1}+\sigma_2^{-1})^{-1}(\sigma_1^{-1}\mu_1+\sigma_2^{-1}\mu_2))^T\\
&\qquad\times(\sigma_1^{-1}+\sigma_2^{-1})(x-(\sigma_1^{-1}+\sigma_2^{-1})^{-1}(\sigma_1^{-1}\mu_1+\sigma_2^{-1}\mu_2))\Bigg].
\end{aligned}$$

We can then take $\mu_1=\bar{\boldsymbol{v}}_+, \mu_2=\bar{\boldsymbol{v}}-\boldsymbol{v}_1$ and $\sigma_1=E_+, \sigma_2=E$, which allows us to solve the integral with respect to $\boldsymbol{v}_2$ in A_1 and A_2 to find

$$\tilde{A}_1=\frac{1}{2\pi|E+E_+|^{1/2}}\int_{\mathbb{R}^2}\chi_{\boldsymbol{v}_1\cdot n<0}\boldsymbol{v}_1^{\otimes 2}\exp\left(-\frac{1}{2}(\boldsymbol{v}_1-(\bar{\boldsymbol{v}}-\bar{\boldsymbol{v}}_+))^T(E+E_+)^{-1}(\boldsymbol{v}_1-(\bar{\boldsymbol{v}}-\bar{\boldsymbol{v}}_+))\right)\mathrm{d}\boldsymbol{v}_1, \tag{A.21}$$

$$\tilde{A}_2=\frac{1}{2\pi|E+E_+|^{1/2}}\int_{\mathbb{R}^2}\chi_{\boldsymbol{v}_1\cdot n<0}\boldsymbol{v}_1^{\otimes 3}\exp\left(-\frac{1}{2}(\boldsymbol{v}_1-(\bar{\boldsymbol{v}}-\bar{\boldsymbol{v}}_+))^T(E+E_+)^{-1}(\boldsymbol{v}_1-(\bar{\boldsymbol{v}}-\bar{\boldsymbol{v}}_+))\right)\mathrm{d}\boldsymbol{v}_1, \tag{A.22}$$

$$\begin{aligned}
\tilde{A}_3=&\frac{1}{2\pi|E+E_+|^{1/2}}\int_{\mathbb{R}^2}\chi_{\boldsymbol{v}_1\cdot n<0}\left[(E^{-1}+E_+^{-1})^{-1}(E^{-1}(\bar{\boldsymbol{v}}-\boldsymbol{v}_1)+E_+^{-1}\bar{\boldsymbol{v}}_+)\right]\otimes\boldsymbol{v}_1\otimes\boldsymbol{v}_1 \qquad (A.23)\\
&\qquad\times\exp\left(-\frac{1}{2}(\boldsymbol{v}_1-(\bar{\boldsymbol{v}}-\bar{\boldsymbol{v}}_+))^T(E+E_+)^{-1}(\boldsymbol{v}_1-(\bar{\boldsymbol{v}}-\bar{\boldsymbol{v}}_+))\right)\mathrm{d}\boldsymbol{v}_1,\\
=&-(E^{-1}+E_+^{-1})^{-1}E^{-1}\tilde{A}_2+(E^{-1}+E_+^{-1})^{-1}(E^{-1}\bar{\boldsymbol{v}}+E_+^{-1}\bar{\boldsymbol{v}}_+)\otimes\tilde{A}_1.
\end{aligned}$$

We see again that by constructing $\tilde{A}_1$ and $\tilde{A}_2$, we can then construct $\tilde{A}_3$ from the results. To work with the characteristic function we consider $\boldsymbol{v}_1=(v_{11},v_{12})$ in two dimensions, by considering the cases $n_1>0$ and $n_1<0$, we partition the integral so that we can consider the one-dimensional integrals in v_{11}. We define $b=(E+E_+)^{-1}$ and $\boldsymbol{a}=\bar{\boldsymbol{v}}-\bar{\boldsymbol{v}}_+$ and use the transform $\tilde{v}_{11}=v_{11}-\frac{n_2}{n_1}v_{12}$, then for example the integrals with respect to $\boldsymbol{v}_1$ in $\tilde{A}_1$ can be written as

$$\begin{aligned}
&\chi_{n_1>0}\int_{\mathbb{R}}v_{12}^j\exp\left[-\frac{|b|}{2b_1}(v_{12}-a_2)^2\right]\int_{\mathbb{R}}\chi_{v_{11}<0}\left(v_{11}-v_{12}\frac{n_2}{n_1}\right)^i\exp\left[-\frac{b_1}{2}F(\boldsymbol{v}_1)^2\right]\mathrm{d}v_{11}\,\mathrm{d}v_{12}\\
&+\chi_{n_1<0}\int_{\mathbb{R}}v_{12}^j\exp\left[-\frac{|b|}{2b_1}(v_{12}-a_2)^2\right]\int_{\mathbb{R}}\chi_{v_{11}>0}\left(v_{11}-v_{12}\frac{n_2}{n_1}\right)^i\exp\left[-\frac{b_1}{2}F(\boldsymbol{v}_1)^2\right]\mathrm{d}v_{11}\,\mathrm{d}v_{12}
\end{aligned}$$

where

$$F(\boldsymbol{v}_1)=\left(v_{11}+v_{12}\left(\frac{b_2}{b_1}-\frac{n_2}{n_1}\right)-\left(a_1+\frac{b_2}{b_1}a_2\right)\right)$$

and $i,j=0,1,2$ such that $i+j=2$. Using the same methods, the integral with respect to $\boldsymbol{v}_1$ in $\tilde{A}_2$ follows similarly with $i,j=0,1,2,3$ such that $i+j=3$. We consider the first term in eq. (A.24),

the second one will follow by symmetric arguments. Integrals to consider are of the form:

$$M_{ij} = \int_{\mathbb{R}} v_{12}^j \exp\left[-\frac{|b|}{2b_1}(v_{12}-a_2)^2\right] \int_{\mathbb{R}} \chi_{v_{11}<0} \left(v_{11} - v_{12}\frac{n_2}{n_1}\right)^i \exp\left[-\frac{b_1}{2}F(\boldsymbol{v}_1)^2\right] \mathrm{d}v_{11}\,\mathrm{d}v_{12}. \tag{A.24}$$

Then let us define

$$N_j = \sqrt{\frac{\pi}{2b_1}} \int_{\mathbb{R}} v_{12}^j \exp\left[-\frac{|b|}{2b_1}(v_{12}-a_2)^2\right] \tag{A.25}$$

$$\times \left[1 - \mathrm{erf}\left(\sqrt{\frac{b_1}{2}}\left(v_{12}\left(\frac{n_2}{n_1}-\frac{b_2}{b_1}\right) + \left(a_1 + \frac{b_2}{b_1}a_2\right)\right)\right)\right] \mathrm{d}v_{12}, \tag{A.26}$$

$$\begin{aligned} P_j =& \frac{1}{b_1}\int_{\mathbb{R}} v_{12}^j \exp\left[-\frac{|b|}{2b_1}(v_{12}-a_2)^2\right] \exp\left[-\frac{b_1}{2}\left(v_{12}\left(\frac{n_2}{n_1}-\frac{b_2}{b_1}\right) + \left(a_1+\frac{b_2}{b_1}a_2\right)\right)^2\right] \mathrm{d}v_{12}, \\ =& \frac{1}{b_1}\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right] \int_{\mathbb{R}} v_{12}^j \exp\left[-\frac{\boldsymbol{n}^\perp\boldsymbol{b}\boldsymbol{n}}{2n_1^2}\left(v_{12} - n_1\frac{(\boldsymbol{n})^\perp\boldsymbol{b}\boldsymbol{a}}{(\boldsymbol{n})^\perp\boldsymbol{b}\boldsymbol{n}^\perp}\right)^2\right] \mathrm{d}v_{12}. \end{aligned} \tag{A.27}$$

Then using the exponential moments eqs. (A.2) and (A.5) to (A.8), we find

$$N_0 = \frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left(1 - \mathrm{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right), \tag{A.28}$$

$$N_1 = \frac{\pi}{\sqrt{|\boldsymbol{b}|}}a_2\left(1 - \mathrm{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right) - \frac{\sqrt{2\pi}(n_2b_1 - n_1b_2)}{|\boldsymbol{b}|\sqrt{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right], \tag{A.29}$$

$$\begin{aligned} N_2 =& \frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[\frac{b_1}{|\boldsymbol{b}|} + a_2^2\right]\left(1 - \mathrm{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right) \\ & - \frac{\sqrt{2\pi}(n_2b_1 - n_1b_2)}{|\boldsymbol{b}|((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)^{3/2}}\left[2a_2((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp) - (n_2b_1 - n_1b_2)\boldsymbol{a}^T\boldsymbol{n}\right] \\ & \times \exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right], \end{aligned} \tag{A.30}$$

$$\begin{aligned} N_3 =& \frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[a_2^3 + 3\frac{b_1}{|\boldsymbol{b}|}a_2\right]\left(1 - \mathrm{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right) \\ & - \frac{\sqrt{2\pi}(n_2b_1 - n_1b_2)}{|\boldsymbol{b}((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)^{5/2}}\left[((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)^2\left(3a_2^2 + 3\frac{b_1}{|\boldsymbol{b}|}\right)\right. \\ & \left. - ((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)\left(3a_2(\boldsymbol{a}^T\boldsymbol{n})(n_2b_1 - n_1b_2) + \frac{(n_2b_1 - n_1b_2)^2}{|\boldsymbol{b}|}\right) + (\boldsymbol{a}^T\boldsymbol{n})^2(n_2b_1 - n_1b_2)^2\right] \\ & \times \exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right] \end{aligned} \tag{A.31}$$

and

$$\mathcal{P}_0 = \frac{\sqrt{2\pi} n_1}{b_1((\boldsymbol{n}^\perp)^T \boldsymbol{b} \boldsymbol{n}^\perp)^{1/2}} \exp\left(-\frac{1}{2}\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right), \tag{A.32}$$

$$\mathcal{P}_1 = \frac{\sqrt{2\pi} n_1^2 (\boldsymbol{n}^\perp)^T \boldsymbol{b}\boldsymbol{a}}{b_1((\boldsymbol{n}^\perp)^T \boldsymbol{b} \boldsymbol{n}^\perp)^{3/2}} \exp\left(-\frac{1}{2}\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right), \tag{A.33}$$

$$\mathcal{P}_2 = \frac{\sqrt{2\pi} n_1^3}{b_1((\boldsymbol{n}^\perp)^T \boldsymbol{b} \boldsymbol{n}^\perp)^{5/2}} [(\boldsymbol{n}^\perp)^T \boldsymbol{b}\boldsymbol{n}^\perp + ((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{a})^2] \exp\left(-\frac{1}{2}\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right), \tag{A.34}$$

$$\mathcal{P}_3 = \frac{\sqrt{2\pi} n_1^4 (\boldsymbol{n}^\perp)^T \boldsymbol{b}\boldsymbol{a}}{b_1((\boldsymbol{n}^\perp)^T \boldsymbol{b} \boldsymbol{n}^\perp)^{7/2}} [3(\boldsymbol{n}^\perp)^T \boldsymbol{b}\boldsymbol{n}^\perp + ((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{a})^2] \exp\left(-\frac{1}{2}\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right). \tag{A.35}$$

We are interested in the integrals with $\boldsymbol{v}_1 \otimes \boldsymbol{v}_1$ and $\boldsymbol{v}_1 \otimes \boldsymbol{v}_1 \otimes \boldsymbol{v}_1$. Thus we want to find

$$\begin{pmatrix} \mathcal{M}_{20} & \mathcal{M}_{11} \\ \mathcal{M}_{11} & \mathcal{M}_{02} \end{pmatrix}, \quad \begin{pmatrix} \mathcal{M}_{30} & \mathcal{M}_{21} \\ \mathcal{M}_{21} & \mathcal{M}_{12} \\ \mathcal{M}_{21} & \mathcal{M}_{12} \\ \mathcal{M}_{12} & \mathcal{M}_{03} \end{pmatrix}. \tag{A.36}$$

Then

$$\mathcal{M}_{0j} = \mathcal{N}_j, \tag{A.37}$$

$$\mathcal{M}_{1j} = -\frac{b_2}{b_1}\mathcal{N}_{j+1} + \left(a_1 + \frac{b_2}{b_1}a_2\right)\mathcal{N}_j - \mathcal{P}_j, \tag{A.38}$$

$$\begin{aligned}\mathcal{M}_{2j} = &\left(\frac{b_2^2}{b_1^2}\right)\mathcal{N}_{j+2} - 2\frac{b_2}{b_1}\left(a_1 + \frac{b_2}{b_1}a_2\right)\mathcal{N}_{j+1} + \left(\frac{1}{b_1} + \left(a_1 + \frac{b_2}{b_1}a_2\right)^2\right)\mathcal{N}_j \\ &+ \left(\frac{n_2}{n_1} + \frac{b_2}{b_1}\right)\mathcal{P}_{j+1} - \left(a_1 + \frac{b_2}{b_1}a_2\right)\mathcal{P}_j,\end{aligned} \tag{A.39}$$

$$\begin{aligned}M_{3j} = &-\left(\frac{b_2^3}{b_1^3}\right)\mathcal{N}_{j+3} + 3\frac{b_2^2}{b_1^2}\left(a_1 + \frac{b_2}{b_1}a_2\right)\mathcal{N}_{j+2} \\ &- \left[3\frac{b_2}{b_1}\left(a_1 + \frac{b_2}{b_1}a_2\right)^2 + 3\frac{b_2}{b_1^2}\right]\mathcal{N}_{j+1} + \left[\left(a_1 + \frac{b_2}{b_1}a_2\right)^3 + \frac{3}{b_1}\left(a_1 + \frac{b_2}{b_1}a_2\right)\right]\mathcal{N}_j \\ &- \left[\left(\frac{n_2}{n_1} + \frac{b_2}{b_1}\right)^2 - \frac{n_2 b_2}{n_1 b_1}\right]\mathcal{P}_{j+2} + \left(a_1 + \frac{b_2}{b_1}a_2\right)\left[\frac{n_2}{n_1} + 2\frac{b_2}{b_1}\right]\mathcal{P}_{j+1} - \left[\left(a_1 + \frac{b_2}{b_1}a_2\right)^2 + \frac{2}{b_1}\right]\mathcal{P}_j.\end{aligned} \tag{A.40}$$

This provides us with everything we need to construct the first three moments of the collision operator in two dimensions, as a one dimensional integral over a circle. If we investigate further,

we see some similarities between the one and two dimensional cases; in particular in the prefactor to the error function:

$$\begin{aligned}\mathcal{M}_{02} =&\frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[\frac{b_1}{|\boldsymbol{b}|}+a_2^2\right]\left(1-\operatorname{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right)\\ &-\frac{\sqrt{2\pi}(b_1n_2-b_2n_1)}{|\boldsymbol{b}|((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)^{3/2}}\left[2a_2((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)-(n_2b_1-n_1b_2)\boldsymbol{a}^T\boldsymbol{n}\right]\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right],\end{aligned} \tag{A.41}$$

$$\begin{aligned}\mathcal{M}_{20} =&\frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[\frac{b_3}{|\boldsymbol{b}|}+a_1^2\right]\left[1-\operatorname{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right],\\ &-\frac{\sqrt{2\pi}(b_2n_2-b_3n_1)}{|\boldsymbol{b}|((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)^{3/2}}\left[2a_1((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)-(n_2b_2-n_1b_3)\boldsymbol{a}^T\boldsymbol{n}\right]\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right]\\ &-\frac{\sqrt{2\pi}}{b_1|\boldsymbol{b}|((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)^{3/2}}\left[n_1|\boldsymbol{b}|\boldsymbol{n}^\perp\boldsymbol{b}\boldsymbol{n}^\perp\right]\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right],\end{aligned} \tag{A.42}$$

$$\begin{aligned}\mathcal{M}_{11} =&\frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[-\frac{b_2}{|\boldsymbol{b}|}+a_1a_2\right]\left[1-\operatorname{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right]\\ &-\frac{\sqrt{2\pi}}{|\boldsymbol{b}|((\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp)^{3/2}}\boldsymbol{a}^T\begin{pmatrix}-n_2(n_2b_1-n_1b_2)^2\\ n_1(n_2b_2-n_1b_3)^2\end{pmatrix}\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right].\end{aligned} \tag{A.43}$$

The additional terms in $\mathcal{M}_{20}$ when compared to $\mathcal{M}_{02}$ are odd in $\boldsymbol{n}$ so integrate to zero in the full integral. The terms which contribute to the third moment are then

$$\mathcal{M}_{03} =\frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[a_2^3+3\frac{b_1}{|\boldsymbol{b}|}a_2\right]\left(1-\operatorname{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right)+Q_{30}(a_1,a_2)\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right], \tag{A.44}$$

$$\begin{aligned}\mathcal{M}_{21} =&\frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[a_1^2a_2-2\frac{a_1b_2}{|\boldsymbol{b}|}+\frac{b_3a_2}{|\boldsymbol{b}|}\right]\left(1-\operatorname{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right)\\ &+Q_{21}(a_1,a_2)\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right],\end{aligned} \tag{A.45}$$

$$\begin{aligned}\mathcal{M}_{12} =&\frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[a_2^2a_1-2\frac{a_2b_2}{|\boldsymbol{b}|}+\frac{b_1a_1}{|\boldsymbol{b}|}\right]\left(1-\operatorname{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right)\\ &+Q_{12}(a_1,a_2)\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right],\end{aligned} \tag{A.46}$$

$$\mathcal{M}_{30} =\frac{\pi}{\sqrt{|\boldsymbol{b}|}}\left[a_1^3+3\frac{b_3}{|\boldsymbol{b}|}a_1\right]\left(1-\operatorname{erf}\left(\frac{\sqrt{|\boldsymbol{b}|}(\boldsymbol{a}^T\boldsymbol{n})}{\sqrt{2(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}}\right)\right)+Q_{03}(a_1,a_2)\exp\left[-\frac{1}{2}\left(\frac{|\boldsymbol{b}|(\boldsymbol{a}^T\boldsymbol{n})^2}{(\boldsymbol{n}^\perp)^T\boldsymbol{b}\boldsymbol{n}^\perp}\right)\right], \tag{A.47}$$

where $Q_{ij}(a_1,a_2)$ are quadratic polynomials in a_1 and a_2,with prefactors in $\boldsymbol{b}$ and $\boldsymbol{n}$.

References

[1] M. Geller, "A table of integrals of the error functions," *J. Res. Nat. Bureau Stand., B*, vol. 73, pp. 1–20, 1969.

Appendix B

Description of code

We can provide code used to create all the figures and results in this thesis upon request: please email t.hurst@sms.ed.ac.uk for more details.

By running the script `addPaths`, all directories needed for computations will be included in the `MATLAB` path.

B.1 EDPD - MATLAB

For results presented in chapter 3 and one dimensional results in chapter 7, we constructed a `MATLAB` code library for event driven particle dynamics. The folder `EDPD` contains all the files necessary to run hard rod samples with and without friction, histogram the results, and plot any results that are presented in the thesis as MATLAB figures.

B.1.1 Parameters for simulation

To construct a sample, the code needs to be provided with a struct `HSParams` which includes all the parameters for the simulation. Any parameters that are not provided by the user are given by defaults. Invalid parameters (*e.g.* including a number of particles that leads to a non-physical density) may be replaced. Here we describe the parameters and outputs of the simulation; the list may not be exhaustive but any remaining parameters should be intuitive.

`HSParams` is separated into individual structs:

- `Particles` contains all the particle properties:
  - `N`: the number of particles in the system.
  - `CoRPart`: the coefficient of restitution for particle-particle collisions.
  - `CoRWall`: the coefficient of restitution for particle-wall collisions.
  - `Species`: the species of each particles.
  - `nSpecies`: the total number of unique species.
  - `mass`: the mass of each particle: currently set to mass=1.

  - `pSize`: information on the diameter of the particles: `sigmaA` gives the average diameter, `sigmaV` gives the variance of the particles (assuming a uniform distribution) and `sigma` is a vector of all particle diameters.

- `Domain` contains information on the domain:
  - `d`: the physical dimension of the simulation. The code can run 2 and 3 dimensional examples (including interpolation and plotting), but for much faster dynamics we refer the reader to the `Dynamo` software [1].
  - `density`: the density of the system.
  - `BCs`: the boundary condition: 1 refers to reflective surfaces, 0 refers to periodic boundarys (rendering `Particles.CoRWall` unnecessary).
  - `L`: half the length of the domain.

- `Dynamics`: contains information on the dynamics:
  - `g`: gravitational constant.
  - `gamma`: friction coefficient *warning!* Setting non-zero `g` *and* `gamma` will lead to very slow/inaccurate dynamics!
  - `TC`: the value used in the TC method to stop inelastic collapse.
  - `Functions`: functions used in the dynamics. Will be set automatically given choice of `g` and `gamma`.

- ICs supplies the initial conditions:
  - `xIC`: the initial positions of particles. If this is not given the particles will be randomly distributed in the domain, and not overlapping (by placing one particle in the domain at a time and throwing away overlapping particles). If the density is high the particles will be positioned uniformly, then elastic linear dynamics will be used to 'randomise' the positions.
  - `vIC`: the initial conditions for velocity.
  - `velAv`: the average velocity of the particles.
  - `velVar`: the variance in velocity of the particles. `AvVel` and `VarVel` can be used to construct normal random initial velocities.

- `OutputOps` provides the output options for the dynamics:
  - `NColl`: number of collisions per simulation (epoch).
  - `EDPD`: a struct containing the parameters for the EDPD algorithm:
    - `PriQ`: flag for priority queues. Currently not available so set to 0.
    - `NumCells`: number of cells used in the cell method. Note: this needs to be at least 4 or the code will produce non-physical dynamics.
  - `makexOut`: whether to make vectors containing all positions and velocities at each collision time.

- `Checks`: check flags and information on the dynamics. Flags are currently used when multiple samples are constructed in `EDPDnSamples` script.
  - `doChecks`: overall flag on whether to do checks during the simulation.
  - `allTimes`: whether to do the checks at all times or just at the end of each epoch.
  - `TotEnergy`: calculate the total kinetic energy of the particles.
  - `EventPart`: list the particles evolved in each event at collision times. If the collision is between a particle and a wall, the second entry will be `NaN`.
  - `CoM`: calculate the centre of mass of the collision
  - `AvVel`: calculate the average velocity of the system
  - `VarVel`: calculate the variance in velocity of the system.
- Interp: options for interpolation. For the use of `InterpolateEDPD` script:
  - `doInterp`: whether to perform interpolation or not. Used for histogramming over multiple samples.
  - `NumPts`: the number of points in time to interpolate over.
  - `GiveTime`: what time to interpolate to. If this is not given then the simulation with be interpolated to the final recorded collision.
- Plot: Plotting options for the simulation. For the use of `doPlotting` script.
  - `doPlots`: flag for automatic plotting.
  - `PlottingType`: there are three types of plotting available: `Simple`, `Full` and `Video`. For one dimensional simulations, `Simple` should be enough, and accounts for periodic boundary conditions. `Full` and `Video` plot circles with diameters given by the particle sizes, and accounts for periodic boundary conditions. This is a more time intensive plotting method.
  - `WhatToPlot`: whether to plot `Start`, `End`, Exact, `Interp`, or `Choosex` for a particular `EDPD.x` and `EDPD.T`.
  - `PlotQuiver`: flag on whether to plot the quiver for each particle (N vector).
  - `PlotParticles`: flag on whether to plot each particle
  - `ParticleColours`: the colours of particles. Each species is assigned a different colour, up to a total of 6 different species.
  - `QuiverColours`: colours of quivers for particles.
  - `FileName`: name of the file for save the video to.
  - ASF auto-scale factor for quivers.

- `EDPD`: this struct is populated by the results of the simulation:
  - `xOut`: position of particles at each collision time.
  - `vOut`: velocity of particles at each collision time.
  - `xEnd`: position of all particles at the end of the simulation.
  - `vEnd`: position and velocity of all particles at the end of the simulation.
  - `CollTs`: vector of collision times.

- `Checks`: results of the dynamics checks flagged in `OutputOps.Checks`.
- `Cells`: information on the cell method parameters:
  - `nCell`: total number of cells in each dimension'
  - `Part`: cell array providing the cell location of each particle.
  - `NeighFinder`: an array to determine which cells are adjacent to a cell of interest.
  - `CoordToLabel`: match cell coordinates to cell labels.
  - `Labels`: integer labels of each cell.
  - `Coords`: coordinates of each cell.
- `NoMoreCollisions`: flag to show when there are no more future collisions. Useful for when viscous dynamics are considered *i.e.* `gamma`> 0.
- `xInterp`: positions of particles at interpolated times.
- `vInterp`: velocities of particles at interpolated times.
- `InterpTimes`: interpolation times.

- `nSamples`: when making multiple samples, this struct is used:
  - `n`: the number of samples used
  - `stopCriterion`: the criterion used for stopping *e.g.* `Energy`. Can also use the variance in velocity or time, depending on the dynamics different criteria can be appropriate.
  - `tol`: the tolerance of the stopping criterion *e.g.* `tol`$= 10^{-3}$ and `stopCriterion=Energy` means the simulation will run until 99.9% of the energy is dissipated.
  - `Interp`: whether to interpolate the result for each sample.
  - `NumPts`: the number of interpolation points.
  - `repLim`: a limit on the number of epochs.
  - `L`: the length of each domain in the sample set.
  - `Time`: the length of each simulation.
  - `endTol`: the tolerance reached by the simulation.
  - `x`: output positions of each sample.
  - `v`: output velocities of the sample.

- `Hist`: information from histogramming the samples:
  1. `nBins`: number of bins for the histogramming.
  2. `nR`: or `g`: the value of the correlation function.
  3. `xR`: positions of the correlation function.

For a particular simulation, not all parameters need to be defined; you may find they are defined correctly for you (although you may want to check the results!).

For the default example, you can run the following code:

```
HS = EventDrivenDynamics;
HS = InterpolateEDPD(HS);
doPlotting(HS)
```

To construct any radial correlation functions in one dimension (*i.e.* results seen in figs. 7.3b, 7.7 and 7.8) we use the script `Examples/makeRCFHardRods.m` with a choice of parameters. For results presenting $g^{(2)}(\varepsilon)$, we use the script `Examples/makegRadial1D` to load all computed correlation functions and extract the value of $g^{(2)}$ at contact, which is then saved in `EDPDData/gRadial1D.mat`.

We also provide scripts for all the results presented in the thesis. These scripts are available in the subdirectory `EDPD/Examples`. Some of these scripts will take a very long time to run, and were generally computed by parallelising samples over many cores. For these results, we also provide `.mat` files with the associated data and scripts to plot the results.

B.2 EDPD - DynamO

For information on how to download, install and use `Dynamo`, we refer the reader to the website `http://dynamomd.org/`, which has plenty of instructions and examples to follow.

We supply some `bash` scripts that were used to create correlation data for hard discs and spheres using `DynamO`, in the `DynamO` subdirectory of the code. The script `allSamples.sh` can be used to run all samples considered in this thesis for two and three dimensional radial correlation functions, by calling `Correlations2D.sh`, `FrictionCorrelations2D.sh`, `Correlations3D.sh`, and the script `FrictionCorrelations3D.sh`, wherein parameters such as α and ϕ can be adjusted.

After exporting the data as `xml` files, we then use `MATLAB` scripts to accumulate the results for figs. 7.11 to 7.16, where we consider the radial correlation function at contact for hard discs and spheres. For the conversion of `xml` files in `MATLAB` we use the `xml2struct` function, available on the `MATLAB` file exchange. The value of $g^{(2)}$ at contact for $d = 2, 3$ is constructed using `makegRadial2d.m, makegRadial3d.m` and results are provided in `EDPDData/gRadial2D.mat` and `EDPDData/gRadial3D.mat` respectively.

Figures 7.11 and 7.14 are produced in `MATLAB` using the scripts `makeg2DExamples.m` and `makeg3DExamples.m` respectively. The script `makegContact2D.m` constructs figs. 7.12 and 7.16, and `makegContact3D.m` constructs figs. 7.12 and 7.16.

B.3 DDFT for granular media

Code for DDFT simulations is based on the `2DChebClass` code library [2]. We supply one dimensional code in `ChebClassGranular` and two dimensional code in `2DChebClassGranular`.

B.3.1 One dimension

To run simulations using one dimensional DDFT, use the function `HardRodsIntervalGM`. this function takes a struct `Input`. Some of the important fields that are included in `Input` are:

- `nParticles`: the number of particles in the domain.
- `FullPacking`: the number of particles in a full packing.

Figure	Script in `Examples`	Data in `EDPDData`	Plot data in `Examples`
2.1	`InelasticTrajectories.m`	N/A	N/A
2.2	`makePotentialExamples.m`	N/A	N/A
2.3 & 2.4	`SquareShoulderTrajectories.m`	N/A	N/A
3.1a	`collisionTimes1D.m`	`collisionHistograms1D.mat`	`plotCollisionTimes1D.m`
3.1b	`collisionTimes2D.m`	`collisionHistograms2D.mat`	`plotCollisionTimes2D.m`
3.3a	`cellTimes1d.m`	`cellTimes1D.mat`	`plotCellTimes1D.m`
3.3b	`cellTimes2d.m`	`cellTimes2D.mat`	`plotCellTimes2D.m`
3.4	`machinePrecision.m`	N/A	N/A
3.12a	`InelasticCollapseGravity.m`	N/A	N/A
3.12b	`InelasticCollapseParticles.m`	N/A	N/A
3.13	`chaoticNature.m`	N/A	N/A
7.2	`sampleTrajectories.m`	N/A	N/A
7.3a	N/A	`gRadial1D.mat`	`particleVariance.m`
7.3b	N/A	`PolyDispersityExamples`	`plotParticleDiameterRCFs.m`
7.4 & 7.5	N/A	`eqTimes.mat`	`plotEqTimes.m`
7.6	N/A	`EqTimes`	`plotEqTimeHist.m`
7.7 & 7.8	N/A	`OneDExamples`	`plot1DRCFExamples.m`
7.9 & 7.10	N/A	`gRadial1D.mat`	`plotgContact1D.m`

Table B.1: `MATLAB` scripts used to construct EDPD figures presented in the thesis.

- `timeRatio`: a coefficient determining the length of the simulation.
- `yMin` and `yMax`: the start and end points of the domain.
- `CoR`: the particle-particle coefficient of restitution.
- `FE`: the free energy functional, either `'Zero'` or `'Percus'` for the examples considered.
- `gData`: the value of the radial correlation function at particular packing fractions (*i.e.* an $K \times 2$ vector with packing fraction in the first entry and RCF in the second). We note that the function `makegCH1D.m` can be used to calculate the 1D Carnahan-Starling approximation of the collision operator.
- `gamma`: the friction coefficient.
- `NumPts` and `InterpPts`: the number of computational and interpolation points for the simulation.
- `FexNum`: a struct for the free energy functional, which has entries:
  - `Fex` the free energy functional, either `'Zero'` or `'Percus'`.
  - `N` the number of computation points for free energy integral calculations.
- `CollisionOperators` a struct which contains strings for collision operator functions `M0`, `M1C` and `M2C`. The options are:
  - `makeM0CLimit1D`, `makeM1CLimit1D` and `makeM2CLimit1D` for the vanishing particle limit approximation of the collision operator.
  - `makeM0C1D`, `makeM1C1D` and `makeM2C1D` for the analytic form of the (inelastic) Boltzmann collision operator.
  - `makeM1CRET1D` and `makeM2CRET1D` for the analytic form of the RET collision operator.
- `InitialCondition`: a struct containing the parameters for the initial conditions
  - `vCoefficient`: the coefficient in the initial velocity.
  - `ECoefficient`: the coefficient in the initial granular temperature.
  - `function`: the choice of initial condition, a string which provides the initial condition function. The functions `makeGranularIC.m`, `makeGranularIC2.m`, `makeGranularIC3.m` are some of the possible functions for use here.
- `V1`: a struct that provides all the information for the external potential, with entries:
  - `V1DV1`: the name of the function that constructs `V1` and `DV1`. Functions available for the examples considered here are `zeroPotential1D.m` and `constant1D.m`.
- `DDFTCode`: the code used to calculate the dynamics, either `GranularDDFT1D` (which uses that $E(r,t)$ is positive), or `GranularDDFT1Dv2`.

The results and figures presented in section 8.2 can also be constructed as `MATLAB` figures, scripts for each figure are given in fig. B.1.

Figure	Script
8.1 8.2a	ConvergenceTest1D.m Example1CoR1DifferentE.m
8.2b	Example05CoR1DifferentE.m
8.3a	Example1NoExternalPotential.m
8.3b	Example1ExternalPotential.m
8.4	Example2Initial.m
8.5	Example2DifferentTerms.m
8.6	movingMassDifferentCoR.m
8.7	Example2DifferentNumParticles.m
8.8	Example2DifferentPackingFraction.m
8.9	Example2DifferentCollOperators.m
8.10	Example2NonPhysical.m
8.11	Example2Analyticg.m
8.12	Example2LongTime

Figure B.1: MATLAB scripts used to construct 1D DDFT figures presented in the thesis.

B.3.2 Two dimensions

For examples in two dimensions we provide a separate directory that contains all the scripts required. Many of the parameters in these examples are the same as in one dimension, but for clarity we describe some of the important parameters for two dimensional examples separately. This set is incomplete but is included as a guide to parameters in the code library.

- `nParticles`: the number of particles in the domain.
- `FullPacking`: the number of particles in a full packing.
- `timeRatio`: a coefficient determining the length of the simulation
- `y1Min`, `y1Max` and `y2Min`, `y2Max`: the start and end points of the domain in each dimension.
- `CoR`: the particle-particle coefficient of restitution.
- `gData`: the value of the radial correlation function at particular packing fractions (*i.e.* an $K \times 2$ vector with packing fraction in the first entry and RCF in the second). We note that the function `makegCH2D.m` can be used to calculate the 2D Carnahan-Starling approximation of the collision operator.
- `gamma`: the friction coefficient.
- `NumPts` and `InterpPts`: the number of computational and interpolation points for the simulation.
- `FexNum`: a struct for the free energy functional, which has entries:
  - `Fex`: the free energy functional, either `'Zero'`, `'FMTRoth'` or `FMTRosenfeld`.
  - `N1disc` and `N2disc`: the number of computation points for free energy integral calculations over a disc.
  - `Ncircle`: the number of computation points for free energy integral calculations over a circle.

- `CollTerms`: an entry in the `optsNum` struct that determines the collision operator considered. Available options are `getCentredMoments`, `getCentredMomentsLimits` and the function `getCentredMomentsZeros`, which sets all collision operator moments to zero.
- `InitialCondition`: a struct containing the parameters for the initial conditions:
  - `vCoefficient`: the coefficients in the initial velocity.
  - `ECoefficient`: the coefficients in the initial granular temperature.
  - `function`: the choice of initial condition, a string which provides the initial condition function. The functions with prefix `makeGranularInitialCondition2D` can be used here.
- `V1`: a struct that provides all the information for the external potential, with entries:
  - `V1DV1`: the name of the function that constructs `V1` and the derivatives `DV1`. Functions available are `zeroPotential2D.m` and `constantPotential2D.m` and `trigPotential2D`.
- `DDFTCode`: the code used to calculate the dynamics, either `GranularDDF2D` (which uses that the diagonals of $\boldsymbol{E}(r,t)$ are positive), or `GranularDDFT2Dv2`.

There are other parameters in the 2D DDFT code, but their purpose should be understandable. The results and figures presented in section 8.3 can also be constructed as `MATLAB` figures, scripts for each figure are given in fig. B.2.

Figure	Script
8.13	`ConvergenceTest2D.m`
8.14	`GaussianBump2DDifferentCoR.m`
8.15	`GaussianBump2DCoRDiff.m`
8.16	`GaussianBumpCollisionMoments.m`
8.17	`GaussianBump2DOffDiagonalEnergies.m`
8.18	`MovingMass2DIC.m`
8.19	`MovingMassDifferentTerms.m`
8.20	`MovingMassTermsDiff.m`
8.21	`MovingMass2DDifferentEnergies.m`
8.22	`MovingMass2DDensity.m`
8.23	`MovingMass2DnParticles.m`
8.24	`PeriodicMess2IC.m`
8.25	`PeriodicMessExample.m`

Figure B.2: `MATLAB` scripts used to construct 1D DDFT figures presented in the thesis.

B.4 Other figures

Figures where code has not been provided are diagrammatic and were drawn using using `Inkscape`.

References

[1] M. N. Bannerman, R. Sargant, and L. Lue, "DynamO: A free $\mathcal{O}(N)$ general event-driven molecular dynamics simulator," *Journal of Computational Chemistry*, vol. 32, no. 15, pp. 3329–3338, 2011.

[2] B. D. Goddard, A. Nold, and S. Kalliadasis, *2DChebClass [Software]*, 2017. [Online]. Available: dx.doi.org/10.7488/ds/1991.

Appendix C

Multiple superadiabatic transitions and the Landau-Zener formula

We consider nonadiabatic systems in which the classical Born-Oppenheimer approximation breaks down. We present a general theory that accurately captures the full transmitted wavepacket after multiple transitions through either a single or distinct avoided crossings, including interference effects. In suitable limits we recover both the celebrated Landau-Zener formula and standard surface-hopping algorithms. In contrast to these methods, we compute the full transmitted quantum wavepacket, including phase information. The derived analytic transition probability can be incorporated into surface-hopping algorithms in place of the Landau-Zener factor. Our algorithm shows excellent agreement with the full quantum dynamics for a range of avoided crossing systems, and can also be applied to single full crossings with similar accuracy.

C.1 Introduction

The Born-Oppenheimer approximation (BOA) [1] is one of the most widely used methods used to study the quantum dynamics of molecules. Intuitively, it is motivated by the fact that electrons are much lighter, and therefore much faster, than the nuclei, and hence rapidly adjust their positions with respect to the nuclei. This scale separation allows, in many cases, for the electronic and nuclear dynamics to be decoupled. In particular, if the electrons start in a particular bound state, for a fixed set of nuclei positions, then they should remain in this bound state even though the nuclei are slowly moving. Hence the nuclear dynamics can be determined by considering their motion on only one (electronic) potential energy surface.

However, there are interesting situations in which the BOA breaks down [2]–[5]. For example, in many photochemical processes the nuclear motion cannot be restricted to a single potential energy surface because, for some nuclear configurations, two such surfaces become close, or even cross. In the former case, known as avoided crossing, the BOA is still valid to leading order (in the small parameter ϵ, which is the square root of the ratio of the electronic and nuclear masses), but the

remaining corrections are of fundamental interest and, in fact, determine the associated chemistry. In the latter case, which generally take the form of conical intersections, the BOA breaks down completely.

Here we are predominantly interested in cases where the transmitted wavepacket is (exponentially) small [6]–[8], for example when there is an avoided crossing, or when the wavepacket does not pass directly over the conical intersection. Such regimes are, in some sense, generic, as avoided crossings are generic in 1D [9], and in higher dimensions the probability of an arbitrary wavepacket exactly hitting a conical intersection is vanishingly small [5]. In particular, we consider cases where the wavepacket passes through multiple avoided crossings, or repeatedly through the same crossing. In such cases the transmitted wavepackets can interfere, and thus it is necessary to understand their phases. This suggests that a full quantum mechanical treatment of the problem is required. However, in even moderate dimensions, such treatments are numerically intractable, especially for multiple, coupled electronic potential surfaces.

In order to overcome this, a range of coupled quantum-classical and semiclassical methods have been developed. These include the multiple-spawning wavepacket method [10]–[12], the frozen Gaussian wavepacket method [13], Ehrenfest dynamics [14]–[16], and the semiclassical initial value representation [17]–[19]. The main advantage of such schemes is the significantly reduced computational cost. The main disadvantage, at least with respect to the problem at hand, is the lack of phase information from almost all such schemes. Along with those mentioned above, one of the most widely-used quantum-classical approaches is surface hopping [20]–[31], in which particles are evolved under classical dynamics on a single surface and can 'hop' to other surfaces with a specified probability. Perhaps the most common approach is to only allow hops at points in the trajectory when the gap between energy surfaces has a local minimum (*i.e.* at an avoided crossing), and the probability of the hop is given by a Landau-Zener (LZ) formula [32], [33]. Such methods give good results for a single transition, especially when the transmitted wavepacket is reasonably large, but fail completely when multiple transitions are involved, due to the complete lack of phase information [34]. We note here that there is at least one such scheme [35] that does aim to retain the phase information, but this is limited to small gaps between the potential energy surfaces, which in turn leads to large transmitted wavepackets. The same restriction is true for other mathematical approaches that lead to explicit formulae for the transmitted wavepacket [7]. It has been shown that, if the gap scales with ϵ, then the transitions are of order one and dominated by the Landau-Zener factor [36], [37].

An alternative approach, inspired by the work of Berry on superadiabatic representations [38], [39] considers the whole quantum mechanical wavepacket. These results, which are restricted to the semiclassical regime where the nuclei move classically, were later made rigorous [40], [41]. It was later shown that, through the use of such superadiabatic representation (which are generalisations of the well-known adiabatic representation), it is possible to derive a formula for the transmitted wavepacket, including phase, at an avoided crossing [42]–[45]. The associated algorithm requires only the quantum evolution on a single energy surface. Whilst this is still computationally demanding if one wants to solve the full Schrödinger equation, there are approximate methods, such as Hagedorn wavepackets [36], [46], [47] or standard quantum chemistry techniques such as MCTDH [48], which make small relative errors and are computationally much more tractable. Such methods have so far been applied to single transitions through avoided crossings [42]–[44],

and to multiple transitions of a single crossing in the case of the photodissociation of NaI [45]. The main goals here are to extend the methodology to multiple transitions through different avoided crossings and to systematically study the effects of making various approximations that lead to a LZ-like transition probability. We also demonstrate that, although not designed to tackle such problems, the methodology can be successfully applied to single transitions of full crossings.

We present an algorithm that has a number of advantages. We have already mentioned: (i) Preservation of phase information, which allows the accurate study of interference effects; (ii) Only a single surface is required, which significantly reduces the computational cost when compared to a fully-coupled system, whilst also allowing the use of state-of-the art numerical schemes. The main other benefits are: (iii) Only the adiabatic surfaces (which are the most commonly obtained surfaces from quantum chemistry calculations) are required, in particular there is no need for a diabatization scheme, or the determination of the adiabatic coupling elements; (iv) Such surfaces are only required locally, and thus can be computed on-the-fly; (v) The transmitted wavepacket is created instantaneously, and hence there is no reliance on complicated numerical cancellations of highly-oscillatory wavepackets, which are generally present in the adiabatic representation; (vi) The methodology is easily extended to multiple adiabatic surfaces; (vii) The derived formula is accurate for a wide range of potential energy gaps and small parameters ϵ, and for any semiclassical wavepacket, *i.e.* one of typical width or order $\sqrt{\epsilon}$.

There are, of course, also some disadvantages when compared to the more widely-used schemes: (i) In order to capture the phase information, the one-level dynamics must retain at least some of their quantum nature, and this are inherently more computationally demanding than the analogous classical dynamics; (ii) In the full formalism, it is necessary to be able to extend the potential surface into the complex plane, at least in the region of an avoided crossing. This is essential to be able to accurately compute the transition probabilities. However, in some regimes, for example when the LZ formula is accurate, we can bypass this requirement; (iii) The scheme is, in principle, restricted to wavepackets that are semiclassical near the avoided crossing. However, due to the linearity of the Schrödinger equation, and as demonstrated in [45], it is possible to 'slice' the wavepacket at the crossing. However, this may be more problematic in higher dimensions; (iv) As it stands, the method is restricted to 1D. However, we have successfully extended it to higher dimensions through a slicing procedure [49].

To outline our approach, we will first review the standard model for nonadiabatic transitions (Section C.2) and avoided crossings (Section C.3). We will then, in Section C.4, give a brief overview of existing surface hopping models and the LZ formula. We then outline the superadiabatic approach and give the resulting formula in Section C.5, before describing the associated algorithm in Section C.6. In Section C.7 we systematically investigating its accuracy, and the effects of replacing the true transition probability by a LZ-like one. Finally, in Section C.8, we summarise our results and discuss some open problems.

C.2 The model

The Schrödinger equation governing the quantum dynamics of a molecular system can be written as

$$\mathrm{i}\hbar\partial_t\psi(\boldsymbol{x}_\mathrm{n},\boldsymbol{x}_\mathrm{e},t) = H_\mathrm{mol}\psi(\boldsymbol{x}_\mathrm{n},\boldsymbol{x}_\mathrm{e},t).$$

where $\boldsymbol{x}_\mathrm{n}$ and $\boldsymbol{x}_\mathrm{e}$ are the nuclear and electronic positions, respectively, and the Hamiltonian is given by

$$H_\mathrm{mol} = -\frac{\hbar^2}{2m_\mathrm{n}}\Delta_{\boldsymbol{x}_\mathrm{n}} - \frac{\hbar^2}{2m_\mathrm{e}}\Delta_{\boldsymbol{x}_\mathrm{e}} + V_\mathrm{n}(\boldsymbol{x}_\mathrm{n}) + V_\mathrm{e}(\boldsymbol{x}_\mathrm{e}) + V_\mathrm{n,e}(\boldsymbol{x}_\mathrm{n}, \boldsymbol{x}_\mathrm{e}).$$

Here the first two terms are the kinetic energies of the nuclei and electrons with masses m_n and m_e, respectively. Note that the masses of the nuclei may all be chosen to be the same by a rescaling of the nuclear coordinates. The potentials V_n and V_e denote the nuclear and electronic Coulomb repulsions, respectively, whilst $V_\mathrm{n,e}$ is the attraction between the nuclei and electrons.

We now change to atomic units ($\hbar = m_\mathrm{e} = e = 1$) and definine $\epsilon = 1/\sqrt{m_\mathrm{e}}$ and the electronic Hamiltonian for fixed nuclear positions $\boldsymbol{x}_\mathrm{n} = \boldsymbol{x}$

$$H_\mathrm{e}(\boldsymbol{x}) = -\frac{1}{2}\Delta_{\boldsymbol{x}_\mathrm{e}} + V_\mathrm{n}(\boldsymbol{x}) + V_\mathrm{e} + V_\mathrm{n,e}(\boldsymbol{x}, \cdot).$$

Suppose that $U^\pm(\boldsymbol{x})$ are two eigenvalues of the electronic Hamiltonian (*i.e.* two adiabatic potential energy surfaces) of multiplicity one and well-separated from the rest of the electronic spectrum. Then, from Born-Oppenheimer theory [50], [51], the effective nuclear Schrödinger equation is

$$\mathrm{i}\epsilon\partial_t\psi(\boldsymbol{x}, t) = \Big(-\frac{\epsilon^2}{2}\Delta_{\boldsymbol{x}} + V(\boldsymbol{x})\Big)\psi(\boldsymbol{x}, t), \tag{C.1}$$

where V is a 2×2 matrix with eigenvalues $U^\pm$, *i.e.* a diabatic matrix. In general, V is symmetric and has the form

$$V(\boldsymbol{x}) = \begin{pmatrix} V_1(\boldsymbol{x}) & V_{12}(\boldsymbol{x}) \\ V_{12}(\boldsymbol{x}) & V_2(\boldsymbol{x}) \end{pmatrix}.$$

For notational convenience, and to connect back to previous work [42]–[45], we find it useful to define

$$Z = (V_1 - V_2)/2, \quad X = V_{12}, \quad d = (V_1 + V_2)/2, \quad \rho = \sqrt{X^2 + Z^2}$$

and so

$$V(\boldsymbol{x}) = d(\boldsymbol{x}) + \begin{pmatrix} Z(\boldsymbol{x}) & X(\boldsymbol{x}) \\ X(\boldsymbol{x}) & -Z(\boldsymbol{x}) \end{pmatrix}.$$

It is easy to see that the adiabatic surfaces are then given by $U^\pm(\boldsymbol{x}) = d(\boldsymbol{x}) \pm \rho(\boldsymbol{x})$ and so ρ is half the energy gap between the two surfaces.

C.3 Avoided crossings

In the adiabatic representation, the electronic Hamiltonian is diagonal at each choice of $\boldsymbol{x}$, and transitions between the adiabatic surfaces are governed by the kinetic energy term, which introduces off-diagonal coupling elements. Typically, when the adiabatic potentials are well-separated, the coupling elements are small and then two levels may be treated separately via the Born-Oppenheimer approximation. However, if the adiabatic surfaces become close, but do not cross, the coupling terms typically become large (but do not diverge). Such nuclear configurations are known as avoided crossings As a result of the large coupling elements, a small, but not negligible, part of the nuclear wavepacket is transferred between the adiabatic surfaces.

Suppose, for clarity of exposition, that the wavepacket initially occupies the upper adiabatic

level. The aim of this work is to determine the transmitted wavepacket (on the lower adiabatic level) well away from the crossing (in the scattering regime). Whilst one can, in principle, compute this by a standard numerical solution of the Schrödinger equatiion, there are a number of challenges that prevent this from being a realistic option for most systems of interest:

1. In order to compute the dynamics, one needs an accurate representation of the potential energy surfaces. Typically the adiabatic surfaces are calculated using quantum chemistry methods, such as Density Functional Theory, but it is computationally expensive to determine such surfaces, especially when the number of degrees of freedom (dimension of $\boldsymbol{x}$) is large. In such cases, it is desirable to design methods which can utilise on-the-fly surfaces, determined only locally. Additionally, practical methods for determining surfaces for excited states are still in their infancy, and one also needs to determine the off-diagonal coupling elements. Finally, we note that diabatic representations are not unique, and those obtained in two- and multiple-state cases may differ significantly [52].

2. The wavepackets we wish to compute are highly oscillatory, typically oscillating with frequency of order ϵ^{-1} in space. This can be seen by comparing the kinetic and potential terms in (C.1). When using a standard numerical scheme, such as Strang splitting, correctly resolving such oscillations requires very fine grids in both position and momentum space. The curse of dimensionality (for N points in d dimensions, one requires N^d points) results in such approaches being impractical for all but very small dimensional systems.

3. Away from avoided crossings, the transmitted wavepacket is typically exponentially small in both the gap size δ and $1/\epsilon$. This can be seen from the formula (C.4) in Section C.5.2 or the standard LZ transition probabilities (C.2) and (C.3), where $\rho_{x_c} = \delta$. In contrast, globally in time, the transitions in the adiabatic representation are of order ϵ. The necessary cancellations in the transmitted wavepacket occur through Stückelberg oscillations. See Figure C.1 for an example. There are two challenges here. The first is to correctly resolve these cancellations, which can require very small time steps. The second is the more general challenge of computing an exponentially small quantity; any absolute errors in the numerical scheme must also be exponentially small or they will overwhelm the desired results.

C.4 Existing approaches and Landau-Zener

In this section we discuss some existing approaches to calculate the transition probability or the transmitted wavepacket.

C.4.1 Surface-hopping algorithms

Here we give a brief overview of surface hopping methods, which are one of the most successful approaches for simulating nonadiabatic dynamics. This is a mixed quantum-classical approach, where particles are transported classically on the adiabatic surfaces and hop between them under certain conditions, which simulates the quantum effects. A general surface hopping algorithm consists of four steps:

1. Sampling of the initial condition.

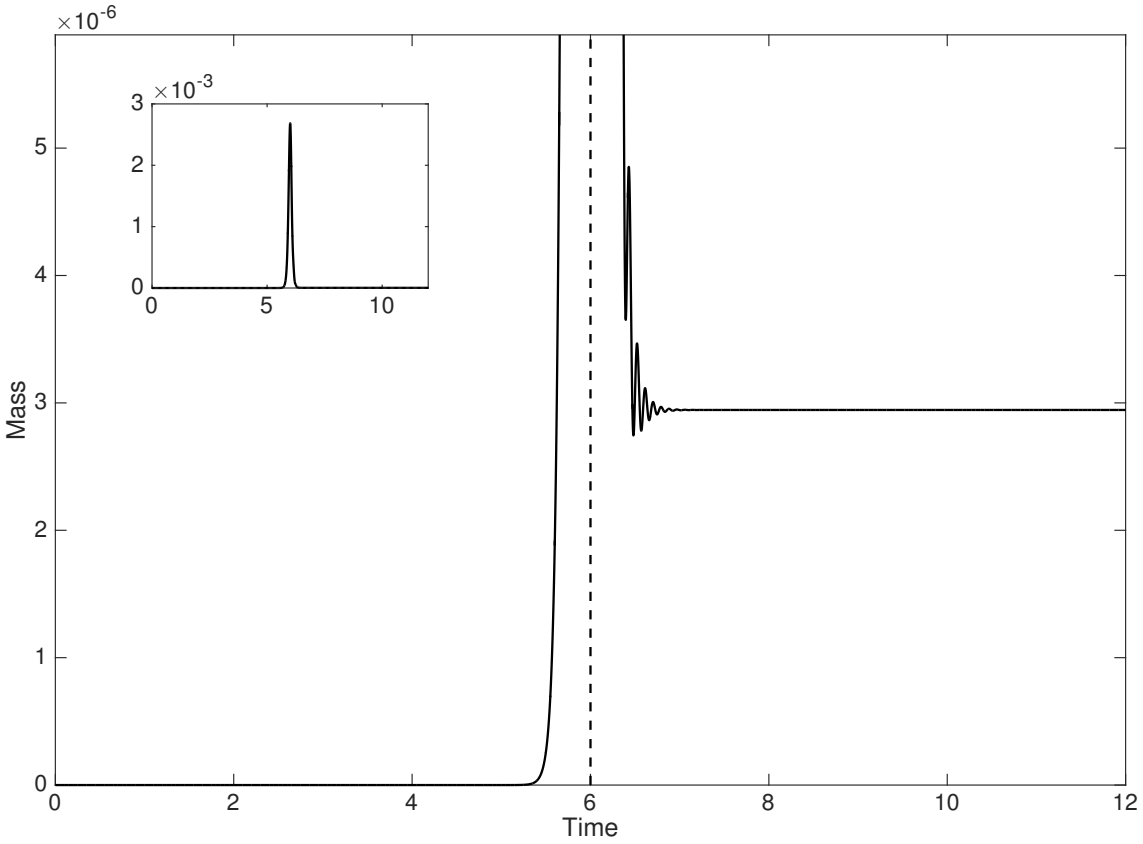


Figure C.1: Inset: The mass of the transmitted wavepacket against time as the wavepacket on the original adiabatic surface moves through an avoided crossing. Main figure: Zoom for clarity of the final transmitted mass and Stückelberg oscillations. The time at which the centre of mass of the original wavepacket reaches the avoided crossing is marked with a dashed vertical line and coincides with the maximum transmitted mass.

2. Classical evolution via $\dot{x} = p$, $\dot{p} = -\nabla U^{\pm}(x)$.

3. Surface hopping.

4. Computation of observables.

There are many such schemes, both deterministic and probabilistic and we refer to [20]–[31] for further details.

Of particular interest here is the surface hopping step. Typically this is performed when the gap between the two adiabatic surfaces is minimal along a classical trajectory. Whenever such a trajectory reaches a local minimum, a transition to the other surface is performed with a certain probability, usually derived from a simplified quantum mechanical model. The standard approach is to use a LZ formula, which we describe in the next Section. The choice of this hopping probability is the main distinguishing feature of different surface hopping models.

The principal advantage of surface hopping algorithms is their simplicity. Due to their use of classical dynamics, which only require local properties of the potential energy surfaces, the methods can be applied in relatively high dimensions, using on-the-fly surfaces. As mentioned previously, such high dimensional systems are beyond the reach of full quantum mechanical methods.

The principal disadvantage is that they lose all phase information, and so cannot reliably treat

systems in which interference effects are important, or determine observables in which the relative phase of the wavepackets on the adiabatic surfaces is required [30]. Additionally, they are accurate only when the specified hopping probability is accurate; we will investigate this in Section C.7.

C.4.2 The Landau-Zener formula

In order to compute the transition probability, it is common to use a Landau-Zener formula. Whilst the LZ model provides a simple formula for the transition probability, it is generally formulated as a one-dimensional, two-level model in the diabatic representation. However, practical applications occur in multiple dimensions and, as discussed above, the potential energy surfaces are normally calculated in the adiabatic representation. There are a number of formulations of the LZ probability, including the extension to multiple dimensions in the diabatic formalism [28], and versions which only require knowledge of the adiabatic potentials [53], [54]. Here we restrict ourselves to two formalisms, the first is a diabatic representation. which requires knowledge of the diabatic matrix elements, whilst the second is an adiabatic representation, which only requires the gap between the adiabatic potentials.

Consider a classical particle with trajectory $\big(x(t), p(t)\big)$ in phase space. Denote the position where ρ attains a minimum by x_c, and the momentum of the particle at the corresponding time t_c by p_c. Then the diabatic LZ transition probability is given by

$$P_{\mathrm{d}} = \exp\Big(-\frac{\pi}{\epsilon}\frac{\rho(x_c)^2}{|p_c|\sqrt{X'(x_c)^2 + Z'(x_c)^2}}\Big). \tag{C.2}$$

The corresponding adiabatic transition probability is given by

$$P_{\mathrm{a}} = \exp\Big(-\frac{\pi}{\epsilon}\sqrt{\frac{\rho(x_c)^3}{\frac{\mathrm{d}^2}{\mathrm{d}t^2}\rho(x(t))|_{t=t_c}}}\Big).$$

If one has knowledge of the diabatic matrix elements, and hence X and Z, this can be rewritten as

$$P_{\mathrm{a}} = \exp\Big(-\frac{\pi}{\epsilon}\frac{\rho(x_c)^2}{|p_c|\sqrt{X'(x_c)^2 + Z'(x_c)^2 + X(x_c)X''(x_c) + Z(x_c)Z''(x_c)}}\Big). \tag{C.3}$$

Note that in the corresponding multidimensional formula [30] there is an additional term, which in 1D would be $[X(x_c)X'(x_c) + Z(x_c)Z'(x_c)](U^{\pm})'(x_c)$. However, since an avoided crossing is defined as a minimum of ρ, and $\rho' = (XX' + ZZ')/\rho$, this term is zero in 1D.

C.5 Superadiabatic representations and the formula

In this section we will briefly review the ideas behind the use of superadiabatic representations to compute the transmitted wavepacket and refer the reader to the cited works for more details. We will then present a generalisation of a previously-derived formula, which is applicable to 1D avoided crossings not centred at the origin.

C.5.1 Superadiabatic representations

Superadiabatic representations were first introduced by Berry [38], [39], under the additional approximation that the nuclei move classically. More recently this has been extended to the full BOA [42]–[45].

As suggested by the name, superadiabatic representations are refinements of the adiabatic representation. The adiabatic representation introduces a unitary matrix such that the electronic Hamiltonian is diagonal at each $\boldsymbol{x}$ and, as such, introduces kinetic coupling elements. In 1D, these coupling elements are of the form $\pm\epsilon\kappa(x)(\epsilon\partial_x)$, where $\kappa = (X'Z - Z'X)/\rho^2$. Since, as described in Section C.3, the wavepackets are rapidly oscillating, the term $\epsilon\partial_x$ is actually of order 1, and hence we see that, naïvely, the transitions are of order ϵ. However, as discussed previously, the transitions are exponentially small in $1/\epsilon$ away from the avoided crossings. Hence, under the adiabatic approximation, the dynamics are complicated, as demonstrated by the population on the lower level during a typical transition, see Figure C.1. This reliance on large cancellations to leave an exponentially small wavepacket suggests that the adiabatic representation may not be the ideal frame of reference in which to study transitions at avoided crossings.

Superadiabatic representations improve on the adiabatic one by simplifying the dynamics near an avoided crossing. The superadiabatic representations can be enumerated, and, initially, moving to successively higher superadiabatic representations reduces the spurious oscillations in the dynamics until the transmitted population builds up monotonically as the wavepacket travels through the avoided crossing. This is known as the optimal superadiabatic representation. However, moving to even higher representations results in the spurious oscillations returning. Previous results give a reliable method to determine the optimal superadiabatic representation [42], [44]. However, computing the unitary operators for this representation is highly challenging, and performing the numerical computations in such a representation is similarly difficult.

The main benefit of superadiabatic representations for our purposes is that they allow the derivation of an explicit formula for the transmitted wavepacket in the optimal superadiabatic representation, without requiring the associated unitary matrix. By general theory [55], all of the superadiabatic representations agree with the adiabatic one away from any avoided crossing. This leads to a simple algorithm to compute the transition through an avoided crossing in the adiabatic representation, as described in Section C.6.

C.5.2 The formula

Following [39], it is useful to introduce a nonlinear rescaling in which the adiabatic coupling elements obtain a universal form, known as the natural scale:

$$\tau(x) = 2\int_{x_c}^{x} \rho(s)\mathrm{d}s,$$

where x_c is the position of the avoided crossing. We now extend ρ and τ into the complex plane and, by the theory of Stokes lines [56], the analytic continuation of ρ has a pair of complex conjugate zeros, close to x_c, at x_{cz} and x_{cz}^*. We define

$$\tau_{x_c} = \tau(x_{cz}) = \tau_r + \mathrm{i}\tau_c.$$

Let $\phi^{\pm}(x, t_c)$ be the incoming wavepacket on the corresponding adiabatic surface $U^{\pm}$ at time t_c when the centre of mass coincides with an avoided crossing at x_c. Then, for $t > t_c$, the transmitted wavepacket on the other adiabatic surface $U^{\mp}$ can be approximated by

$$\psi(x,t) = \mathrm{e}^{-(\mathrm{i}/\epsilon)(t-t_c)H^{\mp}}\,\psi^{\mp}(x)$$

where $\psi^{\mp}(x)$ is a wavepacket instantaneously created at time t_c, which is more easily expressed in Fourier space via

$$\begin{aligned}\hat{\psi}^{\mp}(p) = \Theta(p^2 \mp 4\delta)\frac{p+\eta^{\mp}}{2|\eta^{\mp}|}\exp\Big(-\frac{\tau_c}{2\delta\epsilon}|p-\eta^{\mp}|\Big)\exp\Big(-\mathrm{i}\frac{\tau_r}{2\delta\epsilon}(p-\eta^{\mp})\Big)\\ \times\exp\Big(-\mathrm{i}x_c(p-\eta^{\mp})\Big)\hat{\phi}^{\pm}(\eta^{\mp}).\end{aligned} \tag{C.4}$$

Here

$$\delta = \rho(x_c), \quad \eta^{\mp} = \mathrm{sign}(p)\sqrt{p^2 \mp 4\delta},$$

Θ is the Heaviside function and the Fourier transform needs to be performed under the correct scaling:

$$\hat{\psi}(p) = \frac{1}{2\pi\epsilon}\int \mathrm{e}^{-(\mathrm{i}/\epsilon)px}\,\psi(x)\mathrm{d}x.$$

We note that the principle difference from previous presentations of the formula is the final exponential factor involving x_c, the position of the avoided crossing. In previous work, this position has been taken to be zero, in which case the factor is simply 1. It arises from the Fourier transform of the nth superadiabatic coupling function $\hat{\kappa_n}(p-\eta)$ (see *e.g.* [44, p. 2257]), where the shift of the crossing in space leads to a phase shift in the Fourier transform.

C.5.3 Analysis of the formula

We now present a brief analysis of the formula, which allows us to connect to the surface hopping approaches, as well as the LZ formula.

Firstly we note that the formula involves the same momentum adjustment that is phenomenologically introduced in surface hopping algorithms. We note that $\eta\mp$ is precisely the classical incoming momentum required to give outgoing momentum p when moving down/up, respectively, a potential energy gap of 2δ and requiring (classical) energy conservation. Relatedly, when passing from the upper to the lower level, the Heaviside function ensures that the transmitted wavepacket has (absolute) momentum at least 2δ, whereas when passing from the lower to upper level it is trivially 1, indicating no restriction on the transmitted momentum. The analogous restriction that a classical particle can only be transmitted to the upper level if it has sufficient kinetic energy is accounted for by the $\hat{\psi}^{\pm}(\eta^{\mp})$ term.

We now discuss how, in appropriate limits, the formula essentially reduces to a LZ transition for each point in momentum space. We make a number of independent approximations:

1. $x_c = 0$.

   For a single avoided crossing we may do this without loss of generality by shifting the space variable.

2. $\tau_r = 0$.

This is the case, for example, when the potential is symmetric around the avoided crossing.

3. δ is small.
   This produces two simplifications to the formula using that $\eta^{\mp} \approx p \mp 2\delta/p$:

   - The prefactor simplifies to $\frac{\eta^{\mp}+p}{2|\eta^{\mp}|} \approx p/|p| = \text{sign}(p)$;
   - The factor in the exponential simplifies to $|p - \eta^{\mp}| \approx 2\delta/|p|$.

   Note that the small parameter in these expansions is actually δ/p_0, and so we expect these approximations to be more accurate for either small δ or large incoming momentum.

4. Second order expansion of ρ.
   It is well known [39] that a naïve second order expansion of ρ is incorrect as the analytic continuation of ρ must vanish like a square root at its complex zeros. We therefore approximate ρ via

   $$\rho(x) \approx \sqrt{\delta^2 + g(x - x_c)}$$

   with g a smooth function with $g(0) = g'(0) = 0$. Performing a second order expansion of g then gives

   $$\rho(x) \approx \sqrt{\delta^2 + \tfrac{1}{2}g''(0)(x - x_c)^2}.$$

   In this case both x_{cz} and τ_c can be computed analytically to give

   $$\tau_c \approx \mathrm{i}\frac{\pi\delta^2}{2\alpha}$$

   where $\alpha^2 = \frac{1}{2}g''(0)$. To connect purely to ρ, we note that $\frac{1}{2}g''(0) = \delta\rho''(x_c)$ and hence

   $$\tau_c \approx \mathrm{i}\frac{\pi\delta^{3/2}}{2\sqrt{\rho''(x_c)}}.$$

   Finally, in order to connect to the LZ formulas, an explicit computation gives that

   $$\rho''(x) = \frac{X'(x)^2 + Z'(x)^2 + X(x)X''(x) + Z(x)Z''(x)}{\rho(x)} - \frac{\big(X(x)X'(x) + Z(x)Z'(x)\big)^2}{\rho(x)^3}$$

   As before $\rho'(x_c) = (X(x_c)X'(x_c) + Z(x_c)Z'(x_c))/\rho(x_c) = 0$ and $\rho(x_c) = \delta$. Hence, we have that

   $$\rho''(x_c) = \frac{X'(x_c)^2 + Z'(x_c)^2 + X(x_c)X''(x_c) + Z(x_c)Z''(x_c)}{\delta}$$

   and so

   $$\tau_c \approx \mathrm{i}\frac{\pi\delta^2}{2\sqrt{X'(x_c)^2 + Z'(x_c)^2 + X(x_c)X''(x_c) + Z(x_c)Z''(x_c)}}.$$

Suppose now that we make all four approximations. Then (C.4) becomes

$$\begin{aligned}\hat{\psi}^{\mp}(p) =& \text{sign}(p)\Theta(p^2 \mp 4\delta) \\ & \exp\Big(-\frac{\pi\delta^2}{2\epsilon|p|\sqrt{X'(x_c)^2 + Z'(x_c)^2 + X(x_c)X''(x_c) + Z(x_c)Z''(x_c)}}\Big)\hat{\psi}^{\pm}(\eta^{\mp}).\end{aligned} \tag{C.5}$$

It is now clear that the exponential factor corresponds precisely to the adiabatic LZ transition

probability in (C.3), with the additional factor of 1/2 accounting for the fact that we are determining the size of the transmitted wavepacket rather than the transition probability, which is proportional to the square of the wavepacket. The Heaviside function is also included indirectly in surface hopping models, which explicitly exclude classically-forbidden transitions.

Note that if we are only interested in the transition probability then the first two approximations are irrelevant as they only affect the phase. However, when dealing with multiple transitions these terms are crucial in understanding interference effects. In Section C.7 we will investigate the effects of these approximations in some example systems.

After approximations (1)–(4) have been made, the resulting formula (C.5) can be thought of as a generalisation of surface hopping that retains phase information. This can be seen by noting that the formula decouples in momentum space. Thus, if we replace the classical transport of individual particles, the ensemble of which represents the initial wavepacket, with quantum evolution of the initial wavepacket, and then replace particle hopping with hopping of momentum components of the wavepacket, then we have a clear analogue of the surface hopping methods. One promising avenue of further work is to investigate the use of the formula transmission probability (instead of the LZ one) in surface hopping algorithms. In contrast, we can recover the surface hopping methodology (but retaining phase information) by dividing the wavepacket into small pieces (the surface hopping particles), evolving them classically on the initial level (*e.g.* using Hagedorn's wavepacket approach[36], [46], [47]) until they reach an avoided crossing, and then applying the formula either with the full transition probability, or the LZ approximation, and reconstructing the wavepacket on the other level.

C.6 The algorithm

The general algorithm is similar to that presented in previous work, but here we describe its extension to multiple transitions and different levels of approximation, which ultimately lead to an analogue of the Landau-Zener formula, but applied to wavepackets, rather than simply as a transition probability. The transmitted wavepacket is computed via the following algorithm:

1. **Initial Condition:** The initial wavepacket should be specified on either the upper or lower adiabatic level, well away from any of the avoided crossings. Note that, in such regions, the adiabatic, superadiabatic, and diabatic levels are very close, so one may instead specify the wavepacket on a single diabatic level. If the initial wavepacket is given close to an avoided crossing, for example as the result of a laser excitation, then it must be evolved away (into the scattering regime) on the corresponding adiabatic level under the BO approximation.

2. **One-Level Dynamics:** The initial wavepacket is now evolved under the BO approximation until the final, specified time, or until another termination condition is satisfied (such as the wavepacket reaching a minimum distance from an avoided crossing). This can be done using any one-level scheme that provides sufficient accuracy, such as Strang splitting, Hagedorn wavepackets [36], [46], [47], or MCTDH [48]. An additional one-level computation is required for the other adiabatic potential surface, but that level is initially unoccupied.

3. **Detection of Avoided Crossings:** Here an avoided crossing is defined as a (local) minimum of the gap ρ. Whenever the centre of mass of the wavepacket reaches such a minimum, apply

the formula as described in the next step. Such local minima may be determined *a priori*, for example when the potentials are given analytically, or on-the-fly by monitoring ρ.

4. **Application of the Formula:** Apply the formula (C.4) to the wavepacket at the avoided crossing and add the resulting wavepacket to the lower level. Note that the formula implicitly requires the potentials to be extended into the complex plane in order to compute τ. However, as described in the following Section, this requirement may be bypassed by using an analogue of the Landau-Zener formula, at a cost to accuracy which is investigated for some examples in Section C.7.

5. **Computation of Observables:** At any time step the wavepackets on the two levels may be used to compute observables, such as mean position, momentum and the level populations, including those which require phase information such as inter-level observables. Note, however (as discussed in Section C.5.1), that these will only agree with the corresponding quantities computed for the adiabatic populations well away from any avoided crossings. An extreme example of this is that, before the wavepacket on the initial level reaches the avoided crossing, the other level us completely unoccupied; see Figure C.8.

In the following Section we will investigate the accuracy of this algorithm. One restriction for its application to multiple crossings is that the transmitted wavepacket must be small, or, more precisely, the wavepacket remaining on the original surface must not change significantly when compared to its evolution on a single adiabatic surface. This is due to the perturbative nature of the derivation, which assumes that the original wavepacket is unchanged during a transition.

C.7 Numerics

C.7.1 Jahn-Teller

We consider first a simple example in order to demonstrate the effects of the approximations in Section C.5.3. We choose

$$V(x) = \begin{pmatrix} x & \delta \\ \delta & -x \end{pmatrix}$$

where we have $X = \delta$, $Z = x$, $\rho = \sqrt{x^2 + \delta^2}$. There is a single avoided crossing at $x_c = 0$, with gap 2δ. It is clear that $x_{cz} = \mathrm{i}\delta$ and a straightforward calculation shows that $\tau_{x_c} = \mathrm{i}\delta^2\pi/2$. Note, therefore, that assumptions (1), (2) and (4) of Section C.5.3 hold exactly. Furthermore, since $X(x_c)X''(x_c) + Z(x_c)Z''(x_c) = 0$, the diabatic and adiabatic LZ transition probabilities given in (C.2) and (C.3) are identical in this case. This simple model allows us to investigate the effects of approximation (3), *i.e.* the difference between the full formula (C.4) and the LZ approximation for a range of values of δ. From the arguments in Section C.5.3, we expect the two results to agree to high accuracy when δ is small, and hence the transition is large, but we expect the full formula result to be more accurate in the regime we are interested (relatively large δ and small transitions).

We choose to specify the wavepacket at the avoided crossing, and determine the initial condition by evolving it backwards in time away from the crossing on a single adiabatic surface. This ensures that the wavepacket is semiclassical (*i.e.* of width order $\sqrt{\epsilon}$) when it reaches the avoided crossing.

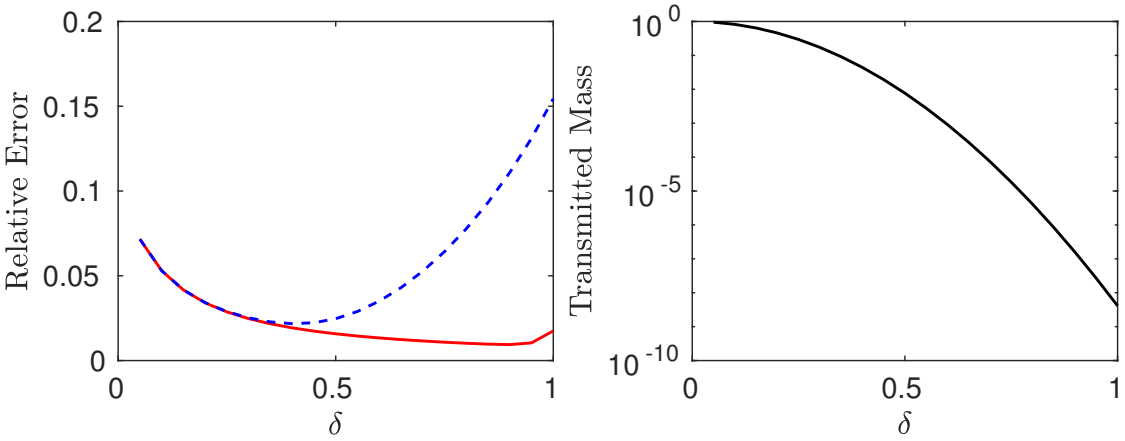


Figure C.2: Left: The relative error between the 'exact' numerical solution and the application of the algorithm using formula (C.4) [red, solid] and (C.5) [blue, dashed]. Right: The 'exact' transmitted mass, which is in excellent agreement with that computed using (C.4) for all values of δ.

As noted above, due to the linearity of the Schrödinger equation, if this were not the case then we could apply a slicing procedure to obtain similarly accurate results. In particular, we choose

$$\hat{\psi}(p) = \frac{1}{(\pi\epsilon)^{1/4}} \exp\Big(-\frac{\mathrm{i}}{\epsilon} p_0 x_0 - \frac{1}{2\epsilon}(p-p_0)^2 - \frac{\mathrm{i}}{\epsilon} x_0 (p-p_0) \Big) \tag{C.6}$$

where, along with δ, the free parameters are ϵ and p_0. For this example we fix $\epsilon = 1/50$, which is similar to the value chosen in surface hopping works *e.g.* [28], [30], [34] (and approximately correct for real-world systems e.g. [45]) and $p_0 = 8$. We could, in principle, vary these parameters, and we will do so in later examples. We note that due to the nature of the potential, in order to start sufficiently far away from the avoided crossing (such that the adiabatic and superadiabatic representations agree) the initial potential energy must be reasonably large, leading to a minimum value of p_0 at the avoiding crossing. See Figure C.3 for the adiabatic potentials.

Here we evolve backwards to a start time of $40/p_0$ with timestep $1/(1000p_0)$ and then forwards through the avoided crossing for time $80/p_0$ with the same timestep. We perform the numerics with a spatial grid with 2^{15} points and endpoints ± 60.

In Figure C.2 we show the relative error in the transmitted wavepacket and the transmitted mass. This clearly demonstrates that, for small δ (and large transmitted mass), both our formula (C.4) and the LZ-like version (C.5) give very good results. However, as δ increases, the simplified version becomes increasingly inaccurate.

C.7.2 Simple avoided crossing

We now consider a simple example, which will both allow us to systematically investigate the accuracy of our method for different parameter regimes, as well as providing a benchmark for the accuracy of a single transition; this is, at least heuristically, a lower bound for the accuracy for multiple transitions. We choose

$$V(x) = \begin{pmatrix} \frac{1}{2}\tanh(x) & \delta \\ \delta & -\frac{1}{2}\tanh(x) \end{pmatrix}$$

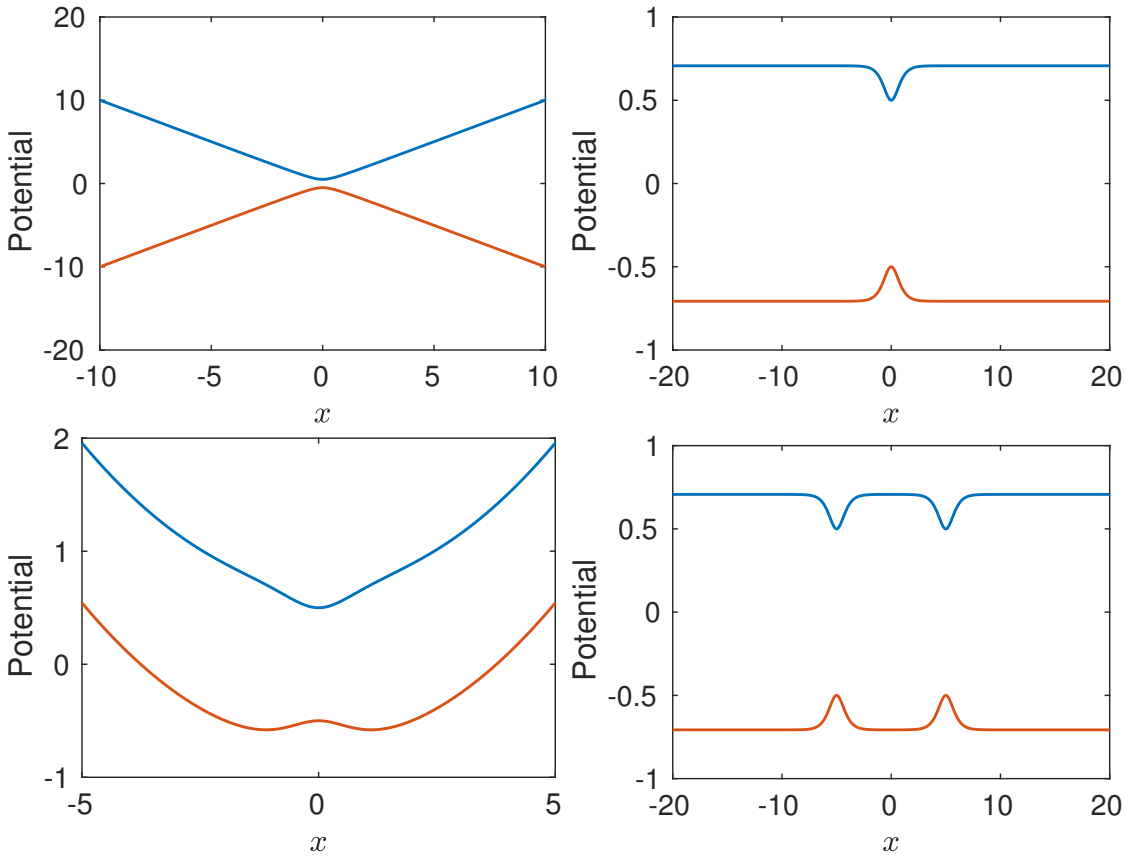


Figure C.3: The adiabatic potentials for [clockwise from top left] the Jahn-Teller, Simple, Quadratic and Dual Potentials

.

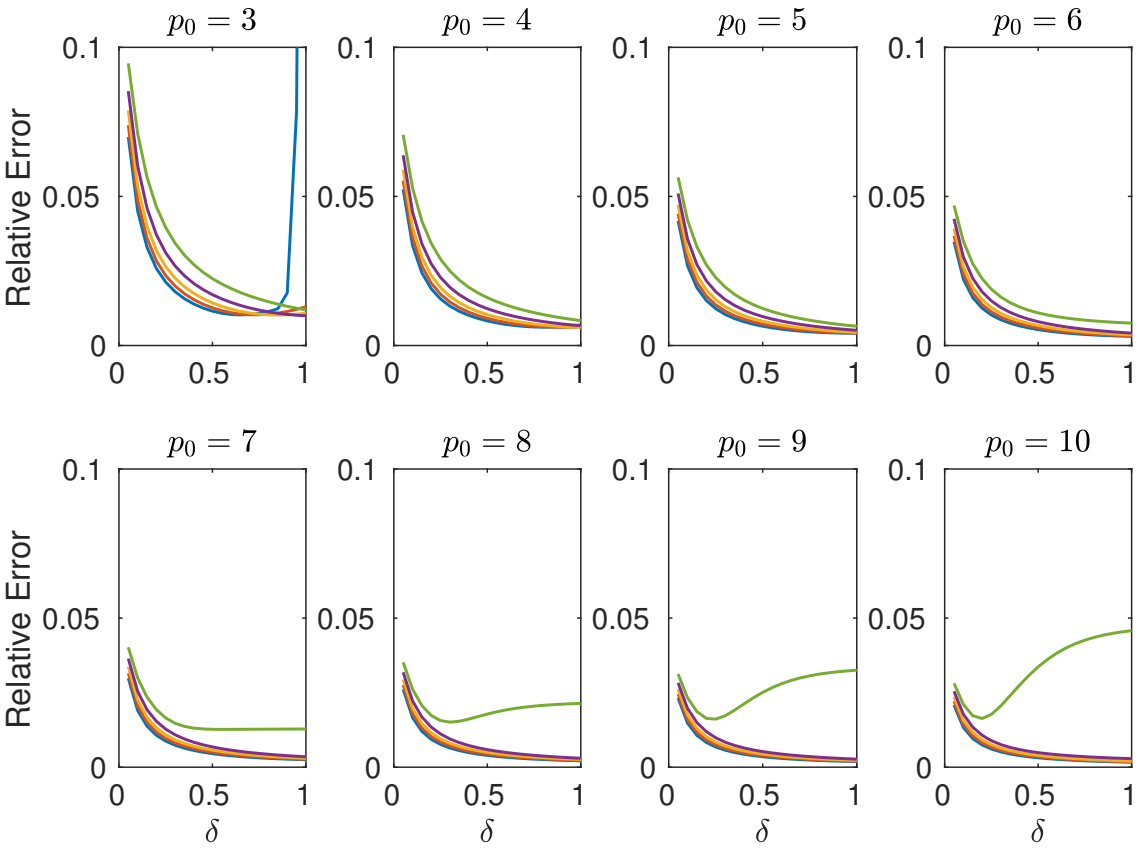


Figure C.4: The relative error between the 'exact' numerical solution and the application of the algorithm using formula (C.4). Each subplot shows the result for a different value of p_0 for a range of δ values. Different colour curves {green, purple, yellow, red, blue} correspond to $\epsilon = \{1/10, 1/20, 1/30, 1/40, 1/50\}$, respectively. Note that, apart from the largest value $\epsilon = 1/10$, the errors are very similar.

where we have $X = \delta$, $Z = \frac{1}{2}\tanh(x)$, $\rho = \sqrt{\frac{1}{4}\tanh(x)^2 + \delta^2}$. See Figure C.3 for the adiabatic potentials with $\delta = 1/2$.

As in the previous example, in order to control the (mean) momentum of the wavepacket when it reaches the crossing, we specify the wavepacket in momentum space at the avoided crossing and then evolve it backwards in time on a single adiabatic surface to obtain an initial wavepacket for the computations. In particular, we take a Gaussian wavepacket as given in (C.6) for a range of values of ϵ and p_0. We compute the results for a single transition of the avoided crossing, both using the full formula (C.4) and the approximate one (C.5), which corresponds to using the LZ transition probability. As can be seen from Figure C.4, the relative error is typically of the order of a few percent, with increasing accuracy as δ and/or p_0 increase. The deviation of the green curve, which corresponds to $\epsilon = 1/10$ is a result of the asymptotic nature of the formula. The odd behaviour of the blue curve for $p_0 = 3$, $\epsilon = 1/50$ and $\delta \approx 1$ is a result of parts of the wavepacket becoming 'trapped' near the avoided crossing, which violates the assumption of a single transition.

Figures C.5 and C.6 demonstrate the effects of using the algorithm with the approximate formula (C.5). As can be seen from Figure C.5, for moderate values of δ, the results become very poor. However, as expected, Figure C.6 shows that, for small δ, the results are very similar to using the full formula (C.4).

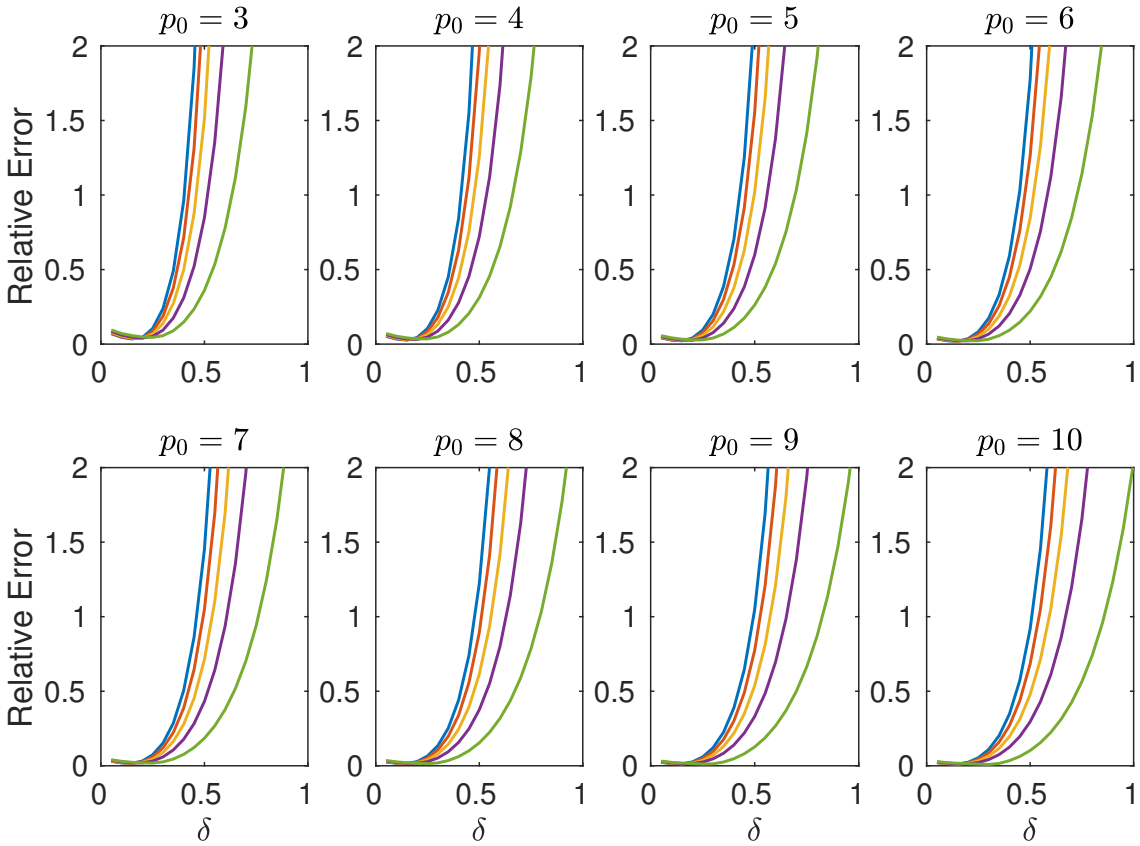


Figure C.5: As Figure C.4 but using formula (C.5). Note that the results for all but the smallest values of δ are significantly worse than those in Figure C.4.

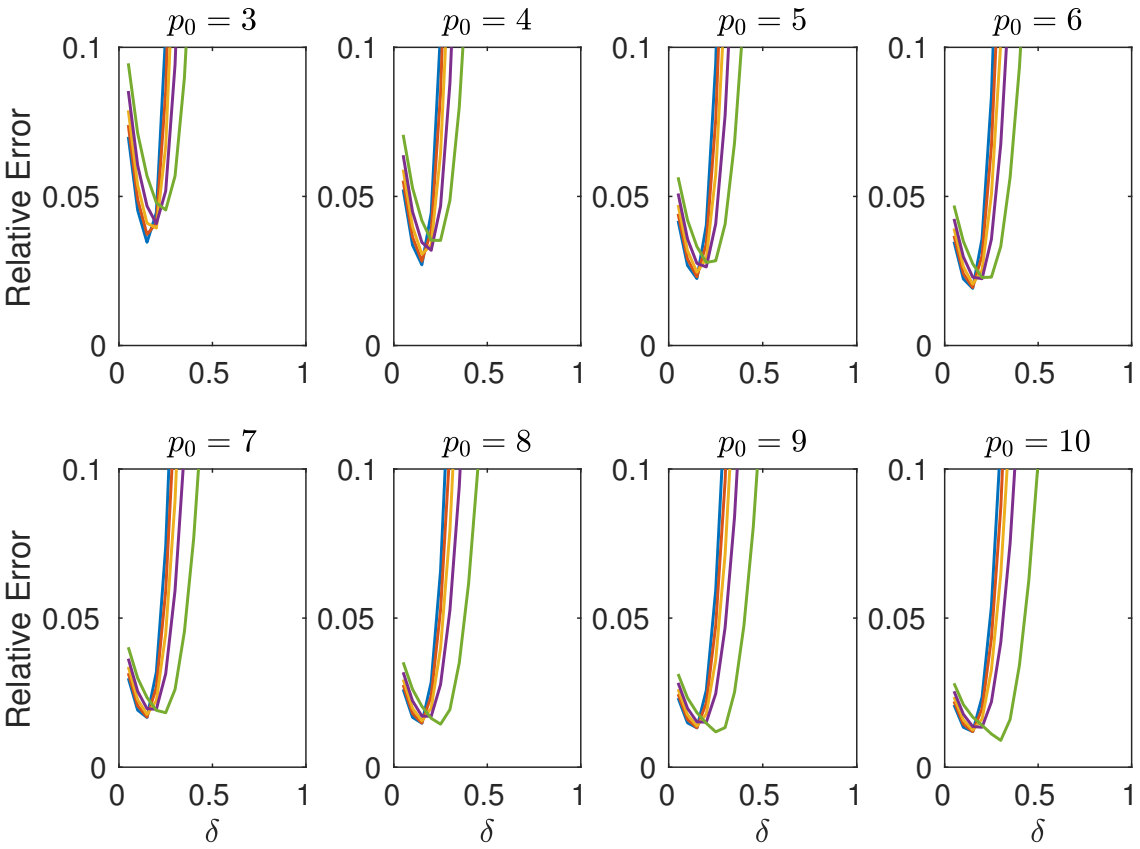


Figure C.6: Zoom of Figure C.5.

For completeness, we give the numerical details: The spatial grid uses 2^{14} points with limits ∓ 40. We use a time step of $1/(100p_0)$ and obtain the initial wavepacket by evolving the wavepacket backwards from the crossing for time $20/p_0$. The system is then evolved forwards for time $40/p_0$. We note that halving the time step and doubling the number of grid points does not significantly affect the results.

As a further test of the accuracy of the algorithm we perform the same calculation as for the Gaussian wavepackets in the previous example, but with a wavepacket on the upper level at the crossing given by

$$\hat{\psi}(p) = \sum_{j=1}^{3} w_i \hat{\psi}(x_{0,i}, p_{0,i}, p), \tag{C.7}$$

where $\hat{\psi}$ is a Gaussian as given by (C.6). We choose $\epsilon = 1/50$, $w = [0.7, 1, 0.9]$, $p_0 = [4.6, 5, 5.3]$ and $x_0 = [0.1, 0, -0.05]$. However, we note that the results are robust under these choices for a wide range of values. We show the resulting transmitted wavepacket in Figure C.7 which for convenience of displaying the phase, we have evolved backwards to the avoided crossing on the lower level. Note that the relative error in this case is 0.0057. In particular, Figure C.7 demonstrates that higher-momentum wavepackets are more likely to make the transition.

We note here that the results for wavepackets starting on the lower level are very similar, and we will investigate such a situation in the following Section.

Finally, we consider if the algorithm is applicable to full crossings (with $\delta = 0$). In such a case, the approximations made in Section C.5.3 lead to the conclusion that the transmitted wavepacket

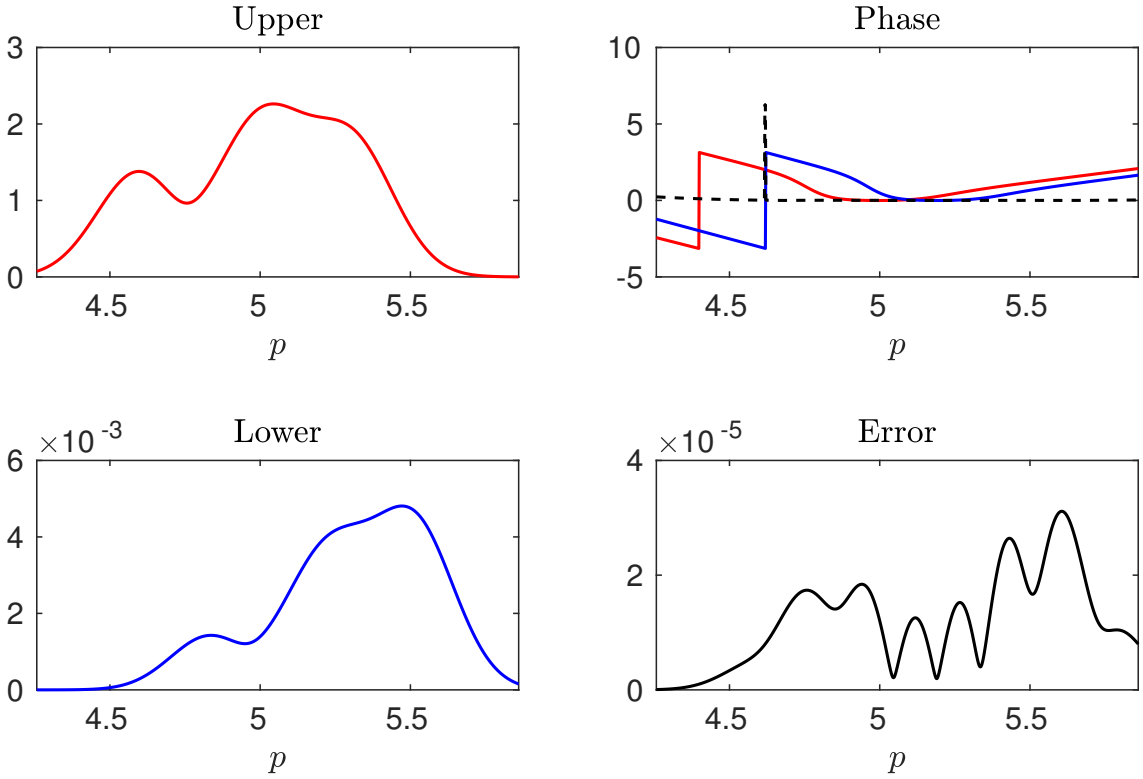


Figure C.7: The wavepacket in momentum space. 'Upper' denotes the wavepacket on the upper level at the avoided crossing, as given by (C.7). 'Lower' denotes the transmitted wavepacket, computed using the algorithm, evolved backwards on the lower level to the avoided crossing. 'Phase' shows the phase of the upper (red), lower (blue) and error(black, dashed). 'Relative Error' displays the relative error between the transmitted wavepackets given by the 'exact' solution and the result of the algorithm.

is approximately equal to the incoming wavepacket. Applying this in the case $p_0 = 5$, $\epsilon = 1/50$ and $\delta = 0$ gives a relative of 0.0856 for both the formula (C.4) and LZ approximation (C.5). This is comparable to the relative error for small, but non-zero δ (see Figure C.4). This indicates that the methodology can also be used for full crossings. This is important in higher dimensions, where part of the wavepacket may travel across a full crossing (conical intersection), whilst other parts experience an effective avoided crossing, in which case we need only one algorithm to accurately treat the whole wavepacket.

C.7.3 Multiple transitions of a single crossing

We now demonstrate the algorithm when the wavepacket makes multiple transitions of a single avoided crossing. Here we add a quadratic confining potential, which causes the wavepacket to oscillate backwards and forwards through the avoided crossing:

$$V(x) = \alpha x^2 + \begin{pmatrix} \frac{1}{2}\tanh(x) & \delta \\ \delta & -\frac{1}{2}\tanh(x) \end{pmatrix},$$

where we have $X = \delta$, $Z = \frac{1}{2}\tanh(x)$, $\rho = \sqrt{\frac{1}{4}\tanh(x)^2 + \delta^2}$. and $d(x) = \alpha x^2$. We choose $\alpha = 0.05$, which gives a relatively weak confining quadratic potential. We use the same grid and

time step as for the simple case in Section C.7.2 but here evolve back to $t = -5$ and forwards to $t = 30$, which gives 3 complete transitions of the avoided crossing. Here we start with a wavepacket of the form (C.6) on the lower level with $x_0 = 0$ and $p_0 = 5$. Again we choose $\epsilon = 1/50$. See Figure C.3 for the adiabatic potentials.

As can be seen in Figure C.8, the 'exact' dynamics require extreme numerical cancellations at each transition in order to produce the true wavepacket. Although not shown in the Figure, the maximum transmitted mass is 0.0028, which is around 200 times larger than the final mass. Note that the results of both formulas are of a similar accuracy to the results for a single crossing, with relative errors 0.0123 and 3.637 for (C.4) and (C.5), respectively. In particular, the agreement between the 'exact' and formula (C.4) results is excellent whilst, in this case, (C.5) significantly underestimates the size of the transmitted wavepackets.

Whist, in principle, we would expect the results of using (C.5) to improve when δ decreases (*i.e.* when the transmitted wavepacket is larger) this adds a complication to the algorithm: When the transmitted wavepacket is large, this significantly affects the wavepacket on the original level, which is used explicitly in the formula for the next transmitted wavepacket. Due to the perturbative nature of the derivation of the formula (C.4) (see *e.g.* [42], [43]), the wavepacket on the original level is not treated explicitly, and so we do not have access to this unless it can be assumed that it is largely unaffected by the transition. A necessary requirement for this, due to mass conservation, is that the transmitted wavepacket is small.

C.7.4 Dual avoided crossings

As we have seen, for multiple transitions at avoided crossings, the algorithm described in Section C.6 works as expected, determining the correct phase of the wavepackets, and therefore also the correct interference effects. However, for transitions at separate avoided crossings there is an extra difficulty that arises from the definition of the diabatic and adiabatic potentials. As can be seen from Figure C.9, in an example with two identical avoided crossings at $\pm x_c$, the diabatic eigenfunctions are even. Hence, treating the two crossings independently, the dynamics through the second crossing could be computed by flipping the surfaces in space (which gives the diabatic surfaces associated with the first crossing) and reversing the momentum of the wavepacket. From (C.4), we see that reversing the momentum introduces a sign change in the transmitted wavepacket, which must be taken into account when computing the total transmitted wavepacket. Note that this argument generalises to the case where there are multiple non-identical crossings; the case here was chosen for clarity.

Here we choose potentials

$$V(x) = \begin{pmatrix} \frac{1}{2}\big(\tanh(x-5) + \tanh(x+5) + 1\big) & \delta \\ \delta & -\frac{1}{2}\big(\tanh(x-5) + \tanh(x+5) + 1\big) \end{pmatrix}$$

where we have $X = \delta$ and $Z = \frac{1}{2}\big(\tanh(x-5) + \tanh(x+5) + 1\big)$. We choose $\epsilon = 1/50$, $\delta = 1/2$ and an initial Gaussian condition using (C.6) with $x_0 = 0$ and $p_0 = 0$. We use the same numerical scheme as in Section C.7.2], first evolving backwards for $t = 5$ and then forwards for $t = 10$. See Figure C.3 for the adiabatic potentials.

As can be seen from Figure C.10, the results using (C.4) are once again very good (with a relative

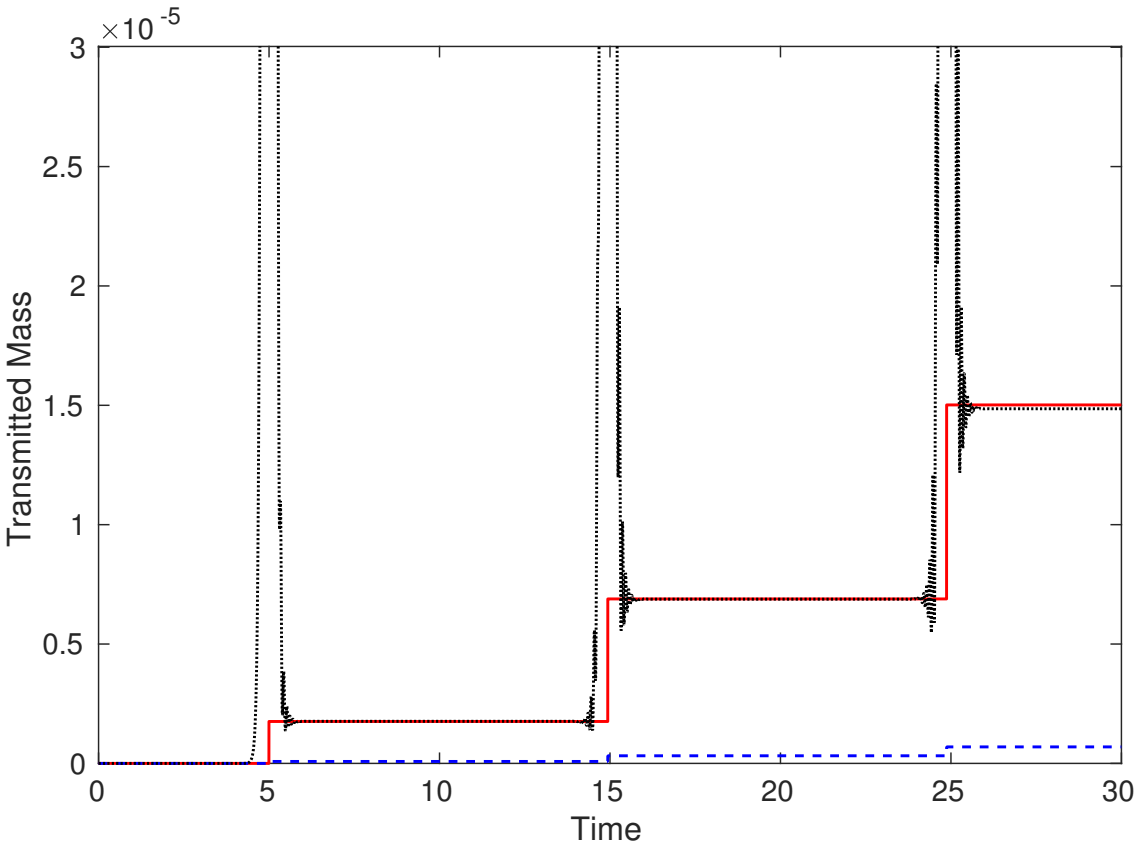


Figure C.8: Mass of transmitted wavepacket on the upper adiabatic level over time for the 'exact' dynamics [black, dotted] and using the algorithm with formulas (C.4) [red, solid] and (C.5) [blue, dashed]. The centre of mass of the wavepacket on the lower level reaches the avoided crossing three times, at approximately $t = 5, 15, 25$, as indicated by the jumps in the formula masses.

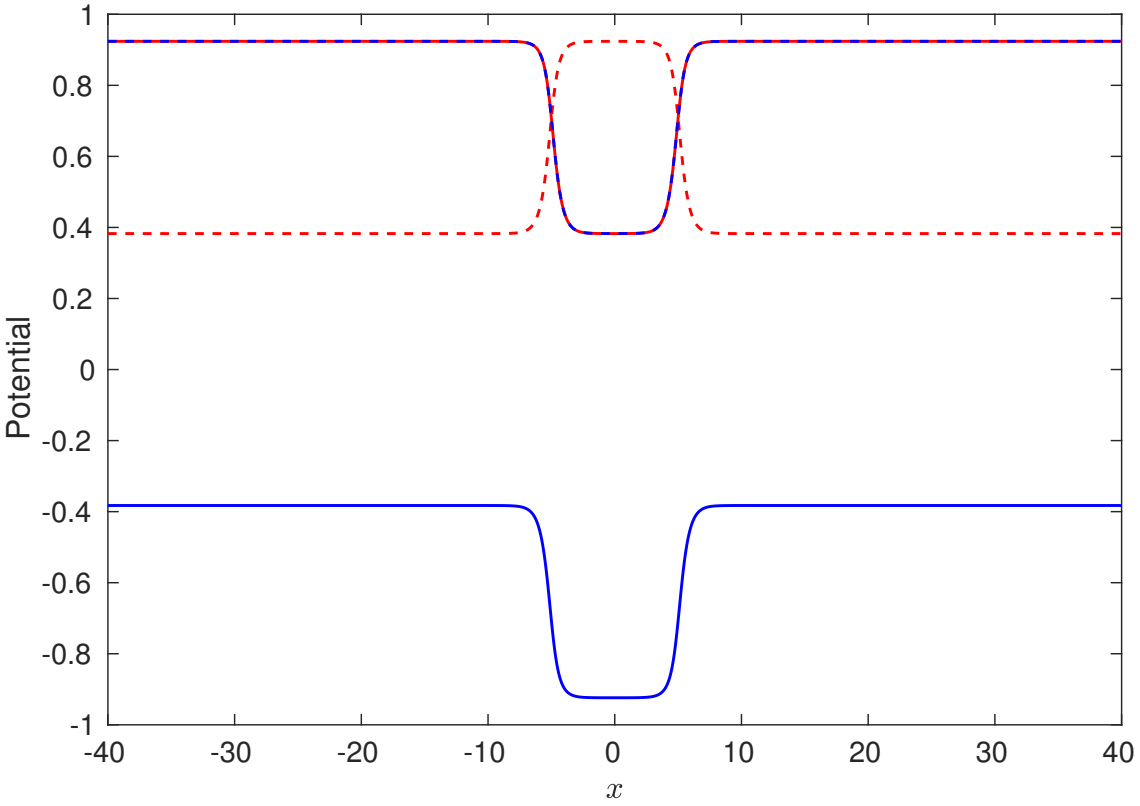


Figure C.9: Components of the two diabatic eigenvectors Φ_1 and Φ_2 in red and blue, respectively. Solid and dashed lines show the two components, denoted by $\Phi_j^{\pm}$. Note that $\Phi_1^+ = \Phi_2^-$ and $\Phi_2^+ = -\Phi_1^-$. The avoided crossings are at $x \approx \pm 5$.

error of 0.0295), whilst those using the approximate formula (C.5) are much poorer (relative error 2.193). Note that if we do not include the additional phase correction described above then the results using (C.4) are also very poor.

C.8 Conclusions and open problems

We have presented a general scheme for the computation of wavepackets transmitted during multiple transitions through avoided crossings (at least when the transmitted wavepacket is small), which is also applicable to single transitions through full crossings. In fact, since, in the latter case, almost the entire wavepacket is transmitted, the scheme should also give accurate results for multiple transitions of full crossings.

The principal advantage of our algorithm is that it produces the full quantum wavepacket, including its phase, in particular allowing the investigation interference effects during multiple transitions. This is in contrast to standard surface-hopping algorithms that lose all phase information.

Open problems, which will be the subject of future works, are (i) Approximation of the wavepacket that remains on the original level when the transmitted wavepacket is not small, which would allow the study of multiple transitions of general crossings; (ii) Extension to higher dimensions. This can be done via a slicing algorithm; preliminary results for model systems are

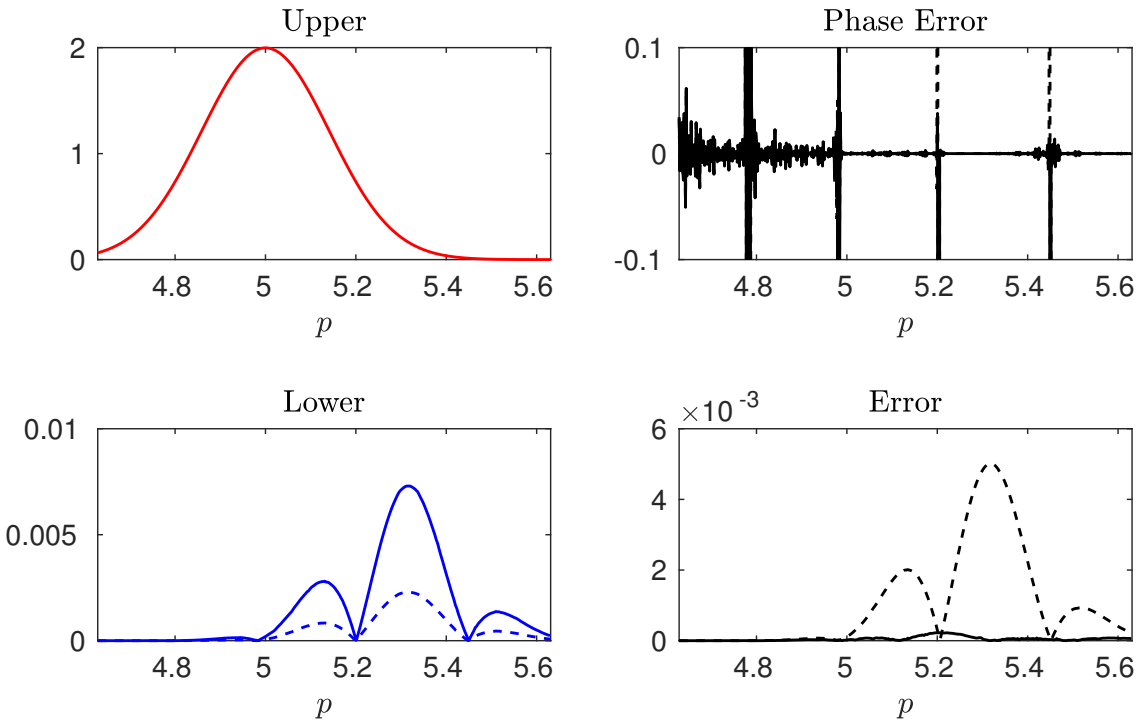


Figure C.10: Top left: The final wavepacket on the upper level. Bottom left: The final transmitted wavepackets on the lower level using (C.4) [solid] and (C.5) [dashed]. Bottom right: The associated errors when compared to the 'exact' numerical solution. Top right: The phase error, which is very small in both cases apart from when the amplitude of the wavepacket is very small.

presented in [49]; (iii) Implementation of our more accurate transition rate in surface hopping models, which should extend their range of validity to systems when the transmitted wavepacket is significantly smaller than those which can be accurately captured by existing schemes.

References

[1] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Annalen der Physik (Berlin, Germany)*, vol. 84, pp. 457–484, 1927.

[2] W. Domcke, D. Yarkony, and K. Horst, *Conical intersections: electronic structure, dynamics and spectroscopy*. World Scientific, 2004, vol. 15.

[3] W. Domcke, D. Yarkony, and K. Köppel, *Conical intersections: theory, computation and experiment*. World Scientific, 2011, vol. 17.

[4] H. Nakamura, *Nonadiabatic transition: concepts, basic theories and applications*. World Scientific, 2012.

[5] J. Tully, "Perspective: Nonadiabatic dynamics theory," *Journal of Chemical Physics*, vol. 137, no. 22, 22A301, 2012.

[6] G. Hagedorn and A. Joye, "A time-dependent Born-Oppenheimer approximation with exponentially small error estimates," *Communications in Mathematical Physics*, vol. 223, no. 3, pp. 583–626, 2001.

[7] ——, "Determination of non-adiabatic scattering wave functions in a born-oppenheimer model," *Annales Henri Poincaré*, vol. 6, no. 5, pp. 937–990, 2005.

[8] A. Martinez and V. Sordoni, "A general reduction scheme for the time-dependent Born-Oppenheimer approximation," *Comptes Rendus Mathematique*, vol. 334, no. 3, pp. 185–188, 2002.

[9] J. Von Neumann and E. Wigner, "Über das Verhalten von Eigenwerten bei Adiabatischen Prozessen," *Physikalische Zeitschrift*, vol. 30, p. 467, 1929.

[10] M. Ben-Nun and T. Martınez, "Nonadiabatic molecular dynamics: Validation of the multiple spawning method for a multidimensional problem," *Journal of Chemical Physics*, vol. 108, no. 17, pp. 7244–7257, 1998.

[11] M. Ben-Nun, J. Quenneville, and T. Martínez, "Ab initio multiple spawning: Photochemistry from first principles quantum molecular dynamics," *Journal of Physical Chemistry A*, vol. 104, no. 22, pp. 5161–5175, 2000.

[12] A. Virshup, J. Chen, and T. Martínez, "Nonlinear dimensionality reduction for nonadiabatic dynamics: The influence of conical intersection topography on population transfer rates," *Journal of Chemical Physics*, vol. 137, no. 22, 22A519, 2012.

[13] E. Heller, "Cellular dynamics: A new semiclassical approach to time-dependent quantum mechanics," *Journal of Chemical Physics*, vol. 94, no. 4, pp. 2723–2729, 1991.

[14] A. McLachlan, "A variational solution of the time-dependent schrodinger equation," *Molecular Physics*, vol. 8, no. 1, pp. 39–44, 1964.

[15] H.-D. Meyer and W. Miller, "A classical analog for electronic degrees of freedom in nonadiabatic collision processes," *Journal of Chemical Physics*, vol. 70, no. 7, pp. 3214–3223, 1979.

[16] S.-I. Sawada, A. Nitzan, and H. Metiu, "Mean-trajectory approximation for charge-and energy-transfer processes at surfaces," *Physical Review B*, vol. 32, no. 2, p. 851, 1985.

[17] W. Miller, "Classical s matrix: Numerical application to inelastic collisions," *Journal of Chemical Physics*, no. 9, pp. 3578–3587, 1970.

[18] H. Kreek and R. Marcus, "Semiclassical collision theory. multidimensional integral method," *Journal of Chemical Physics*, vol. 61, no. 8, pp. 3308–3312, 1974.

[19] W. Miller, "The semiclassical initial value representation: A potentially practical way for adding quantum effects to classical molecular dynamics simulations," *Journal of Physical Chemistry A*, vol. 105, no. 13, pp. 2942–2955, 2001.

[20] J. Tully and R. Preston, "Trajectory surface hopping approach to nonadiabatic molecular collisions: The reaction of h+ with d2," *Journal of Chemical Physics*, vol. 55, no. 2, pp. 562–572, 1971.

[21] W. Miller and T. George, "Semiclassical theory of electronic transitions in low energy atomic and molecular collisions involving several nuclear degrees of freedom," *Journal of Chemical Physics*, vol. 56, no. 11, pp. 5637–5652, 1972.

[22] J. Stine and J. Muckerman, "On the multidimensional surface intersection problem and classical trajectory 'surface hopping'," *Journal of Chemical Physics*, vol. 65, no. 10, pp. 3975–3984, 1976.

[23] P. Kuntz, J. Kendrick, and W. Whitton, "Surface-hopping trajectory calculations of collision-induced dissociation processes with and without charge transfer," *Chemical Physics*, vol. 38, no. 2, pp. 147–160, 1979.

[24] J. Tully, "Molecular dynamics with electronic transitions," *Journal of Chemical Physics*, vol. 93, no. 2, pp. 1061–1071, 1990.

[25] S. Hammes-Schiffer and J. Tully, "Proton transfer in solution: Molecular dynamics with quantum transitions," *Journal of Chemical Physics*, vol. 101, no. 6, pp. 4657–4667, 1994.

[26] U. Müller and G. Stock, "Surface-hopping modeling of photoinduced relaxation dynamics on coupled potential-energy surfaces," *Journal of Chemical Physics*, vol. 107, no. 16, pp. 6230–6245, 1997.

[27] E. Fabiano, G. Groenhof, and W. Thiel, "Approximate switching algorithms for trajectory surface hopping," *Chemical Physics*, vol. 351, no. 1, pp. 111–116, 2008.

[28] C. Fermanian Kammerer and C. Lasser, "Single switch surface hopping for molecular dynamics with transitions," *Journal of Chemical Physics*, vol. 128, no. 14, p. 144 102, 2008.

[29] C. Lasser and T. Swart, "Single switch surface hopping for a model of pyrazine," *Journal of Chemical Physics*, vol. 129, no. 3, p. 034 302, 2008.

[30] A. Belyaev, C. Lasser, and G. Trigila, "Landau–zener type surface hopping algorithms," *Journal of Chemical Physics*, vol. 140, no. 22, p. 224 108, 2014.

[31] A. Belyaev, W. Domcke, C. Lasser, and G. Trigila, "Nonadiabatic nuclear dynamics of the ammonia cation studied by surface hopping classical trajectory calculations," *Journal of Chemical Physics*, vol. 142, no. 10, p. 104 307, 2015.

[32] D. Zener, "Non-adiabatic crossings of energy levels," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 137, pp. 696–702, 1932.

[33] L. Landau, *Collected Papers of L.D. Landau*. Oxford: Pergamon Press, 1965.

[34] C. Fermanian Kammerer and C. Lasser, "An egorov theorem for avoided crossings of eigenvalue surfaces," *Communications in Mathematical Physics*, pp. 1–47, 2017.

[35] L. Chai, S. Jin, Q. Li, and O. Morandi, "A multiband semiclassical model for surface hopping quantum dynamics," *Multiscale Modellling and Simulation*, vol. 13, no. 1, pp. 205–230, 2015.

[36] G. Hagedorn, *Molecular propagation through electron energy level crossings*. American Mathematical Soc., 1994, vol. 536.

[37] G. Hagedorn and A. Joye, "Landu-zener transitions through small electronic eigenvalue gaps in the born-oppenheimer approximation," *Annales Henri Poincaré*, vol. 68, p. 85, 1998.

[38] M. Berry, "Histories of adiabatic quantum transitions," *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, vol. 429, no. 1876, pp. 61–72, 1990.

[39] M. Berry and R. Lim, "Universal transition prefactors derived by superadiabatic renormalization," *Journal of Physics A: Mathematical and General*, vol. 26, no. 18, pp. 4737–4747, Jan. 1993.

[40] G. Hagedorn and A. Joye, "Time development of exponentially small non-adiabatic transitions," *Communications in Mathematical Physics*, vol. 250, no. 2, pp. 393–413, 2004.

[41] V. Betz and S. Teufel, "Precise coupling terms in adiabatic quantum evolution," *Annales Henri Poincaré*, vol. 6, no. 2, pp. 217–246, 2005.

[42] V. Betz, B. Goddard, and S. Teufel, "Superadiabatic transitions in quantum molecular dynamics," *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, vol. 465, no. 2111, pp. 3553–3580, Nov. 2009.

[43] V. Betz and B. Goddard, "Accurate prediction of nonadiabatic transitions through avoided crossings," *Physical Review Letters*, vol. 103, p. 213 001, 2009.

[44] ——, "Nonadiabatic transitions through tilted avoided crossings," *SIAM Journal of Scientific Computing*, vol. 33, no. 5, pp. 2247–2276, 2011.

[45] V. Betz, B. Goddard, and U. Manthe, "Wave packet dynamics in the optimal superadiabatic approximation," *Journal of Chemical Physics*, vol. 144, no. 22, p. 224 109, 2016.

[46] G. Hagedorn, "Semiclassical quantum mechanics. III. the large order asymptotics and more general states," *Annals of Physics*, vol. 135, no. 1, pp. 58–70, 1981.

[47] C. Lubich, *From quantum to classical molecular dynamics: reduced models and numerical analysis*. European Mathematical Society, 2008.

[48] F. Huarte-Larrañaga and U. Manthe, *Multidimensional Quantum Dynamics: MCTDH Theory and Applications*. Wiley, 2009.

[49] V. Betz, B. D. Goddard, and T. D. Hurst, "Nonadiabatic transitions in multiple dimensions," *SIAM Journal on Scientific Computing*, vol. 41, no. 5, B1011–B1033, 2019.

[50] G. A. Hagedorn, "A time dependent born-oppenheimer approximation," *Communincations in Mathematical Physics*, vol. 77, no. 1, pp. 1–19, 1980.

[51] H. Spohn and S. Teufel, "Adiabatic decoupling and time-dependent born–oppenheimer theory," *Communications in Mathematical Physics*, vol. 224, no. 1, pp. 113–132, 2001.

[52] A. Belyaev, P. Barklem, A. Dickinson, and F. Gadéa, "Cross sections for low-energy inelastic h+ na collisions," *Physical Review A*, vol. 81, no. 3, p. 032 706, 2010.

[53] C. Zhu, Y. Teranishi, and H. Nakamura, "Nonadiabatic transitions due to curve crossings: Complete solutions of the landau-zener-stueckelberg problems and their applications," *Advances in Chemical Physics*, vol. 117, pp. 127–233, 2001.

[54] A. Belyaev and O. Lebedev, "Nonadiabatic nuclear dynamics of atomic collisions based on branching classical trajectories," *Physical Review A*, vol. 84, no. 1, p. 014 701, 2011.

[55] S. Teufel, *Adiabatic perturbation theory in quantum dynamics*. Springer, 2003.

[56] A. Joye, G. Mileti, and C. Pfister, "Interferences in adiabatic transition probabilities mediated by stokes lines," *Physical Review A*, vol. 44, no. 7, p. 4280, 1991.

Appendix D

Non-adiabatic transitions in multiple dimensions

We consider non-adiabatic transitions in multiple dimensions, which occur when the Born-Oppenheimer approximation breaks down. We present a general, multi-dimensional algorithm which can be used to accurately and efficiently compute the transmitted wavepacket at an avoided crossing. The algorithm requires only one-level Born-Oppenheimer dynamics and local knowledge of the potential surfaces. Crucially, in contrast to many standard methods in the literature, we compute the whole wavepacket, including its phase, rather than simply the transition probability. We demonstrate the excellent agreement with full quantum dynamics for a a range of examples in two dimensions. We also demonstrate surprisingly good agreement for a system with a full conical intersection.

D.1 Introduction

Many computations in quantum molecular dynamics rely on the Born-Oppenheimer Approximation (BOA) [1], which utilises the small ratio ε^2 of electronic and reduced nuclear masses to replace the electronic degrees of freedom with *Born-Oppenheimer* potential surfaces. When these surfaces are well separated, the BOA further reduces computational complexity by decoupling the dynamics to individual surfaces.

However, there are many physical examples (see *e.g.* [2],[3],[4] and [5]) where the surfaces are not well separated (known as an *avoided crossing*) or even have a full intersection. In these regions the BOA breaks down, and the coupled dynamics must be considered; when a wavepacket travels over a region where the surfaces are separated by a small but non-vanishing amount, a chemically crucial portion of the wavepacket can move to a different energy level via a *non-adiabatic transition*. The existence of the small parameter ε introduces several challenges when attempting to numerically approximate the dynamics. First, and independently of the existence of an avoided or full crossing, the wavepacket oscillates with frequency $1/\varepsilon$ and hence a very fine computational grid is required. Furthermore, in the region of an avoided crossing, the dynamics produce rapid oscillations and, in turn, cancellations in the wavepacket; the transmitted wavepacket very close to the crossing is $\mathcal{O}(\varepsilon)$, but in the scattering regime the transmission is exponentially small. It is therefore necessary to travel far from the avoided crossing (in position space) with a small time-

step to accurately calculate the phase, size and shape of the transmitted wavepacket. In order to calculate the exponentially small wavepacket, one must ensure that the absolute errors in a given numerical scheme are also exponentially small, or they will swamp the true result. Finally, the number of gridpoints in the domain increases exponentially as the dimension of the system increases. Thus standard numerical algorithms quickly become computationally intractable.

Many efforts have been made to avoid computational expense by approximating the transmitted wavepacket while avoiding the coupled dynamics. *Surface hopping algorithms* discussed in [6]–[18] approximate the transition using classical dynamics, where the *Landau-Zener transition rate* [19], [20] is sometimes used to determine the size of the transmitted wavepacket. This method has enjoyed some success, and has been applied to higher dimensional systems (in particular see [16], [17]). However, the full transmitted quantum wavepacket is not always calculated; phase information is lost, although surface hopping approaches have been considered which try to incorporate phase information [21]–[26]. Such information is crucial when considering systems with interference effects, *e.g.* ones in which the initial wavepacket makes multiple transitions through an avoided crossing. In contrast, in [27] and [28], a formula is derived to accurately approximate the full transmitted wavepacket, in one dimension, using only decoupled dynamics. The formula has been applied to a variety of examples with accurate results, including the transmitted wavepacket due to photo-dissociation of sodium iodide [29].

In this paper we construct a method to apply the formula derived in [27] and [28] to higher dimensional problems. We set up the problem, state assumptions, and the main result and algorithm in Appendix D.2. Our derivation is motivated by the derivation of the formula in one dimension [27], which we outline in Appendix D.3 and extend to d dimensions in Appendix D.4. In Appendix D.5 we create a d-dimensional formula for systems in which near the avoided crossing, when the derivatives of the adiabatic potential surfaces are slowly varying in all but the direction in which the wavepacket is travelling. We then extend this result via a simple algorithm to obtain a general d-dimensional formula. We provide some examples and results in Appendix D.6 and note conclusions and future work in Appendix D.7. Additional information on code and some instructive diagrams are available in appendix D.8 and appendix D.9 respectively.

D.2 Set-up and Main Results

We consider the evolution of a semiclassical wavepacket $\psi : \mathbb{R}^d \to \mathbb{C}^2$ at time t, $\psi = \begin{pmatrix} \psi_1(\boldsymbol{x},t) \\ \psi_2(\boldsymbol{x},t) \end{pmatrix}$, governed by the equation:

$$i\varepsilon\partial_t\psi(\boldsymbol{x},t) = H\psi(\boldsymbol{x},t), \tag{D.1}$$

where ε^2 is the ratio between an electron and the reduced nuclear mass of the molecule, *i.e.* $\varepsilon \ll 1$ and H is a Hamiltonian operator. This system is derived after a standard rescaling of a full two level Schrödinger equation involving the kinetic and potential terms between electrons and nuclei. We use the ε-scaled Fourier transform to transform the wavepackets ψ_1, ψ_2 and operators such as H into momentum space:

Definition 19. In d dimensions the wavepacket $f : \mathbb{R}^d \to \mathbb{C}$ in scaled momentum space is given

using the *ε-scaled Fourier transform*

$$\widehat{f^{\varepsilon}}(\boldsymbol{k}) = \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} f(\boldsymbol{x}) \exp\left(-\frac{i}{\varepsilon}\boldsymbol{k}\cdot\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x}. \tag{D.2}$$

For any (sufficiently nice) function $f : \mathbb{R}^d \to \mathbb{C} \in L^2(\mathbb{R}^d)$, the *$\varepsilon$-scaled Fourier transform* $\widehat{A}^{\varepsilon}$ of an operator A is given by

$$\widehat{A}^{\varepsilon}\widehat{f^{\varepsilon}}(\boldsymbol{k},t) \coloneqq \widehat{Af}^{\varepsilon}(\boldsymbol{k},t) = \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} Af(\boldsymbol{x},t) \exp\left(-\frac{i}{\varepsilon}\boldsymbol{k}\cdot\boldsymbol{x}\right) \mathrm{d}\boldsymbol{x}. \tag{D.3}$$

We also define the Weyl quantization [30] in multiple dimensions, which is used throughout this work.

Definition 20. For a symbol $H(\varepsilon, \boldsymbol{p}, \boldsymbol{q})$, given a test function ψ, we define the *Weyl quantization* of H by

$$(\mathcal{W}_{\varepsilon}H\psi)(\boldsymbol{x}) = \frac{1}{(2\pi\varepsilon)^d} \int_{\mathbb{R}^{2d}} \mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{y} H(\varepsilon, \boldsymbol{\xi}, \frac{1}{2}(\boldsymbol{x}+\boldsymbol{y}))e^{\frac{i}{\varepsilon}(\boldsymbol{\xi}\cdot(\boldsymbol{x}-\boldsymbol{y}))}\psi(\boldsymbol{y}). \tag{D.4}$$

The Hamiltonian in eq. (D.1) is given by [28]

$$H(\boldsymbol{x}) = -\frac{\varepsilon^2}{2}\nabla_{\boldsymbol{x}}^2 I + V(\boldsymbol{x}) + d(\boldsymbol{x})I, \tag{D.5}$$

where

$$V(\boldsymbol{x}) = \begin{pmatrix} Z(\boldsymbol{x}) & X(\boldsymbol{x}) \\ X(\boldsymbol{x}) & -Z(\boldsymbol{x}) \end{pmatrix} \tag{D.6}$$

and $d(\boldsymbol{x})$ is the part of the potential operator with non-zero trace. In general $V(\boldsymbol{x})$ can be given by a Hermitian matrix, but as noted in [31], any Hermitian $V(\boldsymbol{x})$ can be transformed into real symmetric form. This is known as the *diabatic representation* of the system. We define $V_1 = Z(\boldsymbol{x}) + d(\boldsymbol{x})$ and $V_2 = -Z(\boldsymbol{x}) + d(\boldsymbol{x})$ as the two *diabatic potentials*, with the *diabatic coupling element* as the off-diagonal element $V_{12} = X(\boldsymbol{x})$. It is useful to define $\theta(\boldsymbol{x}) = \tan^{-1}\left(\frac{X(\boldsymbol{x})}{Z(\boldsymbol{x})}\right)$, so that we can write the polar decomposition of eq. (D.5):

$$\cos(\theta(\boldsymbol{x})) = \frac{Z(\boldsymbol{x})}{\sqrt{X(\boldsymbol{x})^2 + Z(\boldsymbol{x})^2}}, \qquad \sin(\theta(\boldsymbol{x})) = \frac{X(\boldsymbol{x})}{\sqrt{X(\boldsymbol{x})^2 + Z(\boldsymbol{x})^2}}. \tag{D.7}$$

Then, defining $\rho(\boldsymbol{x}) = \sqrt{X(\boldsymbol{x})^2 + Z(\boldsymbol{x})^2}$, gives

$$V(\boldsymbol{x}) = \rho(\boldsymbol{x}) \begin{pmatrix} \cos(\theta(\boldsymbol{x})) & \sin(\theta(\boldsymbol{x})) \\ \sin(\theta(\boldsymbol{x})) & -\cos(\theta(\boldsymbol{x})) \end{pmatrix}. \tag{D.8}$$

Consider the unitary matrix U_0 which diagonalises the potential operator $V(x)$:

$$U_0(\boldsymbol{x}) = \begin{pmatrix} \cos\left(\frac{\theta(\boldsymbol{x})}{2}\right) & \sin\left(\frac{\theta(\boldsymbol{x})}{2}\right) \\ \sin\left(\frac{\theta(\boldsymbol{x})}{2}\right) & -\cos\left(\frac{\theta(\boldsymbol{x})}{2}\right) \end{pmatrix}. \tag{D.9}$$

If we define $\psi_0(\boldsymbol{x},t) = \begin{pmatrix}\psi^+(\boldsymbol{x},t)\\ \psi^-(\boldsymbol{x},t)\end{pmatrix} = U_0(\boldsymbol{x})\psi(\boldsymbol{x},t)$, then we arrive at the *adiabatic Schrödinger equation*

$$i\varepsilon\partial_t\psi_0(\boldsymbol{x},t) = H_0\psi_0(\boldsymbol{x},t). \tag{D.10}$$

Here $H_0 = U_0 H U_0^{-1}$ is given by

$$H_0(\boldsymbol{x}) = -\frac{\varepsilon^2}{2}\nabla_{\boldsymbol{x}}^2\boldsymbol{I} + \begin{pmatrix}\rho(\boldsymbol{x}) + d(\boldsymbol{x}) + \varepsilon^2\frac{\|\nabla_{\boldsymbol{x}}\theta(\boldsymbol{x})\|^2}{8} & -\varepsilon\frac{\nabla_{\boldsymbol{x}}\theta(\boldsymbol{x})}{2}\cdot(\varepsilon\nabla_{\boldsymbol{x}}) - \varepsilon^2\frac{\nabla_{\boldsymbol{x}}^2\theta(\boldsymbol{x})}{4}\\ \varepsilon\frac{\nabla_{\boldsymbol{x}}\theta(\boldsymbol{x})}{2}\cdot(\varepsilon\nabla_{\boldsymbol{x}}) + \varepsilon^2\frac{\nabla_{\boldsymbol{x}}^2\theta(\boldsymbol{x})}{4} & -\rho(\boldsymbol{x}) + d(\boldsymbol{x}) + \varepsilon^2\frac{\|\nabla_{\boldsymbol{x}}\theta(\boldsymbol{x})\|^2}{8}\end{pmatrix}. \tag{D.11}$$

The *adiabatic potential surfaces* are given by the diagonal entries of the adiabatic potential matrix to leading order,

$$V_U(\boldsymbol{x}) = \rho(\boldsymbol{x}) + d(\boldsymbol{x}), \quad V_L(\boldsymbol{x}) = -\rho(\boldsymbol{x}) + d(\boldsymbol{x}), \tag{D.12}$$

where V_U is the upper adiabatic potential surface, and V_L is the lower adiabatic potential surface. The off-diagonal entries of eq. (D.12) are coupling terms, which are negligible when the two adiabatic surfaces are well separated. An avoided crossing occurs when two adiabatic surfaces become close to one another, and the coupling terms have a non-negligible effect. Note that, as we are considering semiclassical wavepackets, derivatives are of order $1/\varepsilon$ and hence the leading order off-diagonal elements are of order ε.

For a more precise definition of an avoided crossing, we direct the reader to [32] (although it should be noted that the precise meaning of avoided crossing does vary in the literature), but for the purposes of this paper we will work with a definition of an avoided crossing with respect to the wavepacket. We define the *centre of mass* of the wavepacket $\psi^\pm$ at time t by

$$\boldsymbol{x}_{\mathrm{COM}}(t) = \frac{\int_{\mathbb{R}^n}\mathrm{d}\boldsymbol{x}\,\boldsymbol{x}|\psi^\pm(\boldsymbol{x},t)|^2}{\int_{\mathbb{R}^n}\mathrm{d}\boldsymbol{x}|\psi^\pm(\boldsymbol{x},\boldsymbol{t})|^2}, \tag{D.13}$$

and the *centre of momentum* of $\psi^\pm$ as

$$\boldsymbol{p}_{\mathrm{COM}}(t) = \frac{\int_{\mathbb{R}^n}\mathrm{d}\boldsymbol{p}\,\boldsymbol{p}|\widehat{\psi^\pm}^\varepsilon(\boldsymbol{p},t)|^2}{\int_{\mathbb{R}^n}\mathrm{d}\boldsymbol{p}|\widehat{\psi^\pm}^\varepsilon(\boldsymbol{p},t)|^2}. \tag{D.14}$$

Definition 21. Let V_U and V_L be the adiabatic surfaces defined in eq. (D.12) such that $V_U(\boldsymbol{x}) - V_L(\boldsymbol{x}) = 2\rho(\boldsymbol{x})$. A wavepacket $\psi^\pm$ on the upper/lower level is said to reach an *avoided crossing* at time t when $\rho\big(\boldsymbol{x}_{\mathrm{COM}}(t)\big)$ reaches a local minimum of ρ along its trajectory. Furthermore, we say that the avoided crossing is *tilted* when, near the avoided crossing, the non-symmetric part $d(\boldsymbol{x})$ of V_U and V_L can be written as $d(\boldsymbol{x}) = \boldsymbol{\lambda}\cdot\boldsymbol{x} + \mathcal{O}(\|\boldsymbol{x}\|^2)$, where $\boldsymbol{\lambda}$ is non-zero in the direction $\boldsymbol{p}_{\mathrm{COM}}(t)$.

We note that, at an avoided crossing, the derivative couplings in eq. (D.11) are non-negligible, and it is in such regions that we expect the transitions between the adiabatic states to occur. In the following we consider only cases in which the avoided crossing is of dimension zero, either due to the nature of the potential energy surfaces, or the path of the wavepacket. In cases where the dimension is higher, for example, when the wavepacket travels along a 'seam' of avoided crossings, we expect the method to break down. For the case of 'tilted' crossings in 1D, we refer the reader

to [33] and note that we will soon make the assumption that $\|\boldsymbol{\lambda}\|$ is small in the direction of $\boldsymbol{p}_{\mathrm{COM}}$, and thus not treat the 'tilted' case here.

We will assume that the initial wavepacket is purely on the upper level, $\psi^0(\boldsymbol{x}) = \begin{pmatrix} \psi^{0,+}(\boldsymbol{x}) \\ 0 \end{pmatrix}$ and, without loss of generality, that the centre of mass of the wavepacket in position space reaches an avoided crossing of height 2δ at position $\boldsymbol{x}_0$ at time t_{ac}, and is moving in the direction of $\boldsymbol{q}_1$. The adiabatic representation approximates the wavepacket transmitted through an avoided crossing to leading order by the perturbative solution [34]

$$\psi_0^-(\boldsymbol{x},t) = -i\varepsilon \int_{-\infty}^{t} e^{-\frac{i}{\varepsilon}(t-s)H^-(\boldsymbol{x})} \kappa_1^-(\boldsymbol{x}) \cdot (\varepsilon\partial_{\boldsymbol{x}}) e^{-\frac{i}{\varepsilon}sH^+(\boldsymbol{x})} \psi^{0,+}(\boldsymbol{x})\, \mathrm{d}s, \tag{D.15}$$

where

$$H^{\pm}(\boldsymbol{x}) = -\frac{\varepsilon^2}{2}\nabla_{\boldsymbol{x}}^2 \pm \rho(\boldsymbol{x}) + d(\boldsymbol{x}), \quad \kappa_1^{\pm}(\boldsymbol{x}) = \pm\frac{\partial_{\boldsymbol{x}}\theta(\boldsymbol{x})}{2}. \tag{D.16}$$

The perturbative solution in the adiabatic representation does not offer much explanation as to the properties of the transmitted wavepacket. For instance, the constructed wavepacket at first looks to be $\mathcal{O}(\varepsilon)$. However due to the adiabatic coupling operator $\kappa_1^{\pm}$, fast oscillations and cancellations between upper and lower transmissions occur near the avoided crossing, so that far from in position space the crossing the transmitted wavepacket is much smaller than the transition at the crossing point (fig. D.1). For this reason, the transmitted wavepacket is better approximated using the

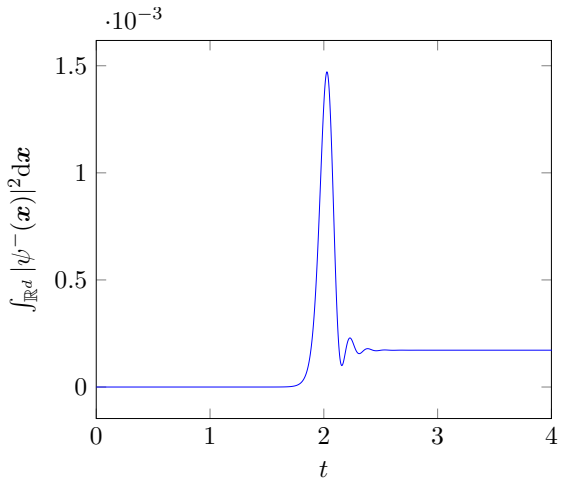


Figure D.1: The total mass of wavepacket $\psi^-(\boldsymbol{x})$ on the lower potential surface against time t, for the system described in example 27 with parameters in eq. (D.64). The centre of mass of the wavepacket reaches the avoided crossing at $t = 2$.

perturbative solution from the n^{th} *superadiabatic* representation [27], for some optimal choice of n. The n^{th} superadiabatic representation is produced by creating and applying unitary *pseudodifferential* operators U_n, such that the off-diagonal elements of the potential operator have prefactor ε^{n+1}, and the diagonal elements are the same to leading order as in the adiabatic representation. Existence of such operators is discussed in [27]. The Hamiltonian H_n in the n^{th} superadiabatic

representation is given by

$$H_n(\boldsymbol{x}) = -\frac{\varepsilon^2}{2}\nabla_{\boldsymbol{x}}^2 \boldsymbol{I} + \begin{pmatrix} \rho(\boldsymbol{x}) + d(\boldsymbol{x}) + \mathcal{O}(\varepsilon^2) & \varepsilon^{n+1} K_{n+1}^+ \\ \varepsilon^{n+1} K_{n+1}^- & -\rho(\boldsymbol{x}) + d(\boldsymbol{x}) + \mathcal{O}(\varepsilon^2) \end{pmatrix}, \tag{D.17}$$

for some pseudodifferential coupling operators $K_{n+1}^{\pm}$, which are of order one. The perturbative solution in the n^{th} superadiabatic representation is then given by

$$\psi_n^-(\boldsymbol{x}, t) = -i\varepsilon^n \int_{-\infty}^{t} e^{-\frac{i}{\varepsilon}(t-s)H^-(\boldsymbol{x})} K_{n+1}^-(\boldsymbol{x}) e^{-\frac{i}{\varepsilon}sH^+(\boldsymbol{x})} \psi^{0,+}(\boldsymbol{x})\, \mathrm{d}s, \tag{D.18}$$

Direct computation of the pseudodifferential operators K_{n+1} and U_n is recursive in n (see Appendix D.4), and leads to very complex operators, so we cannot produce a practical numerical scheme directly using superadiabatic representations. However we will use superadiabatic representations to construct a simple and accurate algorithm.
In [28], where a formula approximating the transmitted wavepacket in one dimension is constructed, five assumptions are made:

(A1) The avoided crossing is 'flat', *i.e.* $\|\boldsymbol{\lambda}\|$ in Definition 21 is small (in the direction of $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$) compared to the energy gap, 2δ. This approximation can be removed in 1D [33], but the resulting algorithm is more complicated; we will pursue the multidimensional version of this in future work.

(A2) The momentum of the wavepacket near the avoided crossing is sufficiently large. Furthermore, by a coordinate rotation we can assume without loss of generality that the momentum is concentrated in the first dimension. This allows the quantum symbol of the coupling operator K_{n+1} to be approximated by its highest order polynomial term, as discussed in Appendix D.4.

(A3) The first order Taylor approximation of the adiabatic (Born-Oppenheimer) energy surfaces near $\boldsymbol{x}_0$ leads to a dynamics that is a good approximation of the true dynamics near $\boldsymbol{x}_0$, *i.e.* we can write the adiabatic propagators near the avoided crossing as

$$H^{\pm} \approx -\frac{\varepsilon^2}{2}\nabla_{\boldsymbol{x}}^2 \pm \delta + \boldsymbol{\lambda} \cdot \boldsymbol{x}, \tag{D.19}$$

(A4) The width of the wavepacket is $\mathcal{O}(\varepsilon)$. For the 1D case, it has been shown[29] that, by the linearity of the Schrödinger equation, we can consider wider wavepackets through a slicing method. We expect this to also hold in higher dimensions.

(A5) The functions ρ and θ are analytic in a strip containing the real axis.

In the multidimensional derivation we will make one additional assumption:

6. The adiabatic potential surfaces near the avoided crossing point vary slowly in all but the direction of $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$.

We are now ready to state the main result of this paper. Under the assumptions (A2) to 6, we approximate the transmitted wavepacket at the avoided crossing point using the formula:

$$\widehat{\psi^-}^{\varepsilon}(\boldsymbol{k},t) = e^{-\frac{i}{\varepsilon}t\widehat{H^-}^{\varepsilon}}\frac{\nu(k_1)+k_1}{2|\nu(k_1)|}e^{-\frac{i}{\varepsilon}(k_1-\nu(k_1))(x_0+\frac{\tau_r}{2\delta})}e^{-\frac{\tau_c}{2\delta\varepsilon}|k_1-\nu(k_1)|}$$
$$\times\, \chi_{k_1^2>4\delta}\widehat{\phi^+}^{\varepsilon}(\nu(k_1),k_2,\ldots,k_d), \quad \text{(D.20)}$$

where ξ, ν, τ_c and τ_r are the d-dimensional analogues of those quantities defined in one dimension in (D1) to (D4), and are discussed in Appendix D.4 and Appendix D.5. Here, as described precisely in Algorithm 22 below, ϕ^+ is the wavepacket on the upper level at the avoided crossing.

We outline the method through which eq. (D.20) may be used to compute the transmitted wavepacket using only one-level dynamics via the following algorithm and 2D diagrams available in fig. D.10:

Algorithm 22.

(B1) Begin with an initial wave packet $\psi^{0,+}(\boldsymbol{x})$ on the upper adiabatic energy surface, far from the crossing in position space, with momentum such that $\rho\big(\boldsymbol{x}_{\mathrm{COM}}(t)\big)$ will attain a minimum value (fig. D.10a).

(B2) Evolve $\psi^{0,+}$ on the upper level, *i.e.* under the BOA, until its centre of mass reaches a local minimum at time t_{ac}. Define

$$\phi^+(\boldsymbol{x}) \coloneqq e^{-\frac{i}{\varepsilon}t_{\mathrm{ac}}H^+(\boldsymbol{x})}\psi^{0,+}(\boldsymbol{x}). \quad \text{(D.21)}$$

(B3) Divide up the full d-dimensional space into d-dimensional strips parallel to $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$. The width of the strips in all directions perpendicular to $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$ should be of the order of the width of the transition region (along $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$) in the optimal superadiabatic basis. In practice we restrict these strips to the region of space where the wavepacket has significant mass.

(B4) On each strip, replace the true potential energy matrix by an approximation that is flat perpendicular to the direction of $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$. In practice, we take the potential along $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$ in the middle of the strip and replicate it in the directions perpendicular to $\boldsymbol{p}_{\mathrm{COM}}(t_{\mathrm{ac}})$. Note in particular that the new potential may be different for each strip.

(B5) Compute the transmitted wavepacket on the lower level for each strip by applying the formula eq. (D.20) along $\boldsymbol{p}_{\mathrm{COM}}$ (fig. D.10c) and sum them together: $\widehat{\psi^-}^{\varepsilon}(\boldsymbol{k},t_{\mathrm{ac}}) = \sum_{j=1}^{n}\widehat{\psi_j^-}^{\varepsilon}(\boldsymbol{k},t_{\mathrm{ac}})$.

(B6) Evolve the transmitted wavepacket away from the avoided crossing on the lower level, say to time $t_{\mathrm{ac}}+s$, using the BOA (fig. D.10e): $\widehat{\psi^-}^{\varepsilon}(\boldsymbol{k},t_{\mathrm{ac}}+s) = e^{-\frac{i}{\varepsilon}s\widehat{H^-}^{\varepsilon}}\widehat{\psi^-}^{\varepsilon}(\boldsymbol{k},t_{\mathrm{ac}})$.

To summarise, we have derived an algorithm for approximating the transmitted wavepacket for an avoided crossing in any dimension, which only requires one-level dynamics, and local information about the adiabatic electronic surfaces, *i.e.* δ and τ^{cz}. The dependence on the n^{th} superadiabatic representation is also removed due to cancellations in the derivation. This seems peculiar to the case where (A1) applies and is not expected to be true in general. A similar method can be used to determine transmitted wavepackets from lower to upper levels. While we note that when the dimension of the system is large, we still require a high dimensional discretization for simulation of the one-level dynamics. However, methods (*e.g.* [35]) which improve performance of one-level

dynamics can be applied to significantly reduce computational cost. In the following section, we derive algorithm 22 and provide numerical examples. We note that for a particular asymptotic limit in one dimension, error bounds have been constructed for this approximation [27], but for general $\boldsymbol{p}_{\mathrm{COM}}, \varepsilon$ only empirical estimates are available.

D.3 Motivation: Approximating the transmitted wavepacket in one dimension

The formula is derived in one dimension using the superadiabatic perturbative solution eq. (D.18) by

(C1) Finding algebraic recursive differential equations to calculate the *quantum symbol* $\kappa_{n+1}^{\pm}$, where $K_{n+1}^{\pm}$ is the Weyl quantisation of $\kappa_{n+1}^{\pm}$.

(C2) Introducing by a change of variables $\tilde{\kappa}^{\pm}(\tau(q)) = \kappa_{n+1}^{\pm}(q)$, where

$$\tau(q) = 2\int_0^q \rho(r)\,\mathrm{d}r, \tag{D.22}$$

(which is the *natural scale* discussed in [31]) then approximating $\tilde{\kappa}_{n+1}^{\pm}$ in an analogous way to the time-adiabatic case in [36].

(C3) Applying the Avron-Herbst formula [37] to $H^{\pm} \approx \frac{\varepsilon^2}{2}\partial_x^2 \pm \delta + \lambda x$ by using (A3).

(C4) Applying a stationary phase argument (with small λ) to evaluate the remaining integral.

Following this derivation leads to an approximation of the transmitted wavepacket in scaled momentum space, far from the avoided crossing in momentum space:

$$\widehat{\psi^-}^{\varepsilon}(k,t) = e^{-\frac{i}{\varepsilon}t\widehat{H^-}^{\varepsilon}} \frac{\nu(k)+k}{2|\nu(k)|} e^{-\frac{i}{\varepsilon}(k-\nu(k))\left(x_0+\frac{\tau_r}{2\delta}\right)} e^{-\frac{\tau_c}{2\delta\varepsilon}|k-\nu(k)|} \chi_{k^2>4\delta} \widehat{\phi^+}^{\varepsilon}(\nu(k)), \tag{D.23}$$

where

(D1) The indicator function $\chi_{k^2>4\delta}$ (which is one when $k^2 > 4\delta$ and zero otherwise) relates to (classical) energy conservation: kinetic energy from the potential energy difference between two levels must be gained by the wavepacket.

(D2) The dependence on the n^{th} superadiabatic representation is removed during the formula derivation.

(D3) $\nu(k) = \mathrm{sgn}(k)(\sqrt{k^2 - 4\delta})$, the initial momentum a classical particle would need to have momentum k after falling down a potential energy difference of 2δ, *i.e.* the distance between the potential surfaces at the avoided crossing, which shifts the wavepacket in momentum space. This arises naturally; it is often enforced in surface hopping algorithms.

(D4) $\tau^{cz} := \tau_r + i\tau_c = 2\int_0^{q^{cz}} \rho(q)\,\mathrm{d}q$, where $q^{cz} \in \mathbb{C}$ is the closest value to the local minimum of ρ such that $\rho(q^{cz}) = 0$, when ρ is extended to the complex plane. The prefactor $e^{-\frac{\tau_c}{2\delta\varepsilon}|\nu(k)-k|}$ determines the size of the transmitted wavepacket. In appendix C, we saw that under appropriate approximations of the momentum and potential surfaces, this prefactor is comparable

to the Landau-Zener transition prefactor used in surface hopping algorithms such as in [17]. An additional change in phase occurs due to τ_r, which is present when the potential is not symmetric about the avoided crossing.

The constructed formula eq. (D.23) allows us to approximate the size and shape of the transmitted wave packet due to an avoided crossing, and avoid computing expensive coupled dynamics. The method for applying the algorithm is as follows:

Algorithm 23 (1D version of Algorithm 22)**.** (E1) Begin with an initial wave packet ψ_0^+ on the upper adiabatic energy surface, far from the crossing in position space, with momentum such that the wave packet will cross the minimum of ρ (fig. D.2a).

(E2) Evolve ψ_0^+ according to the BOA on the upper adiabatic level until the centre of mass is at the avoided crossing, at time t_{ac} (fig. D.2b),

$$\phi^+(x) \coloneqq e^{-\frac{i}{\varepsilon}t_{\text{ac}}H^+(x)}\psi^{0,+}(x), \tag{D.24}$$

(E3) Apply the one dimensional formula to the ε-Fourier transform of the wave packet at the crossing (fig. D.2c):

$$\widehat{\psi^-}^\varepsilon(k,t_{\text{ac}}) = \frac{\nu(k)+k}{2|\nu(k)|}e^{-\frac{i}{\varepsilon}(k-\nu(k))(x_0+\frac{\tau_r}{2\delta})}e^{-\frac{\tau_c}{2\delta\varepsilon}|k-\nu(k)|}\chi_{k^2>4\delta}\widehat{\phi^+}^\varepsilon(\nu(k)), \tag{D.25}$$

(E4) Evolve the transmitted wave packet far away enough from the crossing in position space, say to time $t_{\text{ac}}+s$, using the BOA (fig. D.2d): $\widehat{\psi^-}^\varepsilon(x,t_{\text{ac}}+s) = e^{-\frac{i}{\varepsilon}s\widehat{H^-}^\varepsilon}\widehat{\psi^-}^\varepsilon(x,t_{\text{ac}})$.

Applications of the one dimensional formula have been widely successful on a variety of examples. In addition to the sodium iodide example [29] already mentioned, tilted avoided crossings have been examined, and a formula developed which in contrast is dependent on n.

Finally, the above derivation can also be modified for reverse transitions (from lower to upper surface). If we consider an initial wavepacket ψ_0^- far from the avoided crossing in position space on the lower energy level, the above algorithm can be applied analogously, where to approximate the wavepacket transmitted to the upper level, eq. (D.25) is replaced by

$$\widehat{\psi^+}^\varepsilon(k,t_{\text{ac}}) = \frac{\tilde{\nu}(k)+k}{2|\tilde{\nu}(k)|}e^{-\frac{i}{\varepsilon}(k-\tilde{\nu}(k))(x_0+\frac{\tau_r}{2\delta})}e^{-\frac{\tau_c}{2\delta\varepsilon}|k-\tilde{\nu}(k)|}\widehat{\phi^-}^\varepsilon(\tilde{\nu}(k),t_{\text{ac}}), \tag{D.26}$$

where $\tilde{\nu}(k) = \operatorname{sgn}(k)\sqrt{k^2+4\delta}$ contributes a loss of momentum due to the potential energy difference between the two surfaces.

D.4 Coupling operators in higher dimensions

The first step in deriving eq. (D.23) in [27] was to approximate the superadiabatic coupling operators $K_{n+1}^\pm$. We now consider these operators in higher dimensions. We restrict the calculations here to two dimensions for clarity, but they can easily be adapted to d dimensions.

Lemma 24. In two dimensions, $\kappa_{n+1}^\pm$ is given by

$$\kappa_{n+1}^\pm(\boldsymbol{p},\boldsymbol{q}) = -2\rho(\boldsymbol{q})(x_{n+1}(\boldsymbol{p},\boldsymbol{q}) \pm y_{n+1}(\boldsymbol{p},\boldsymbol{q})). \tag{D.27}$$

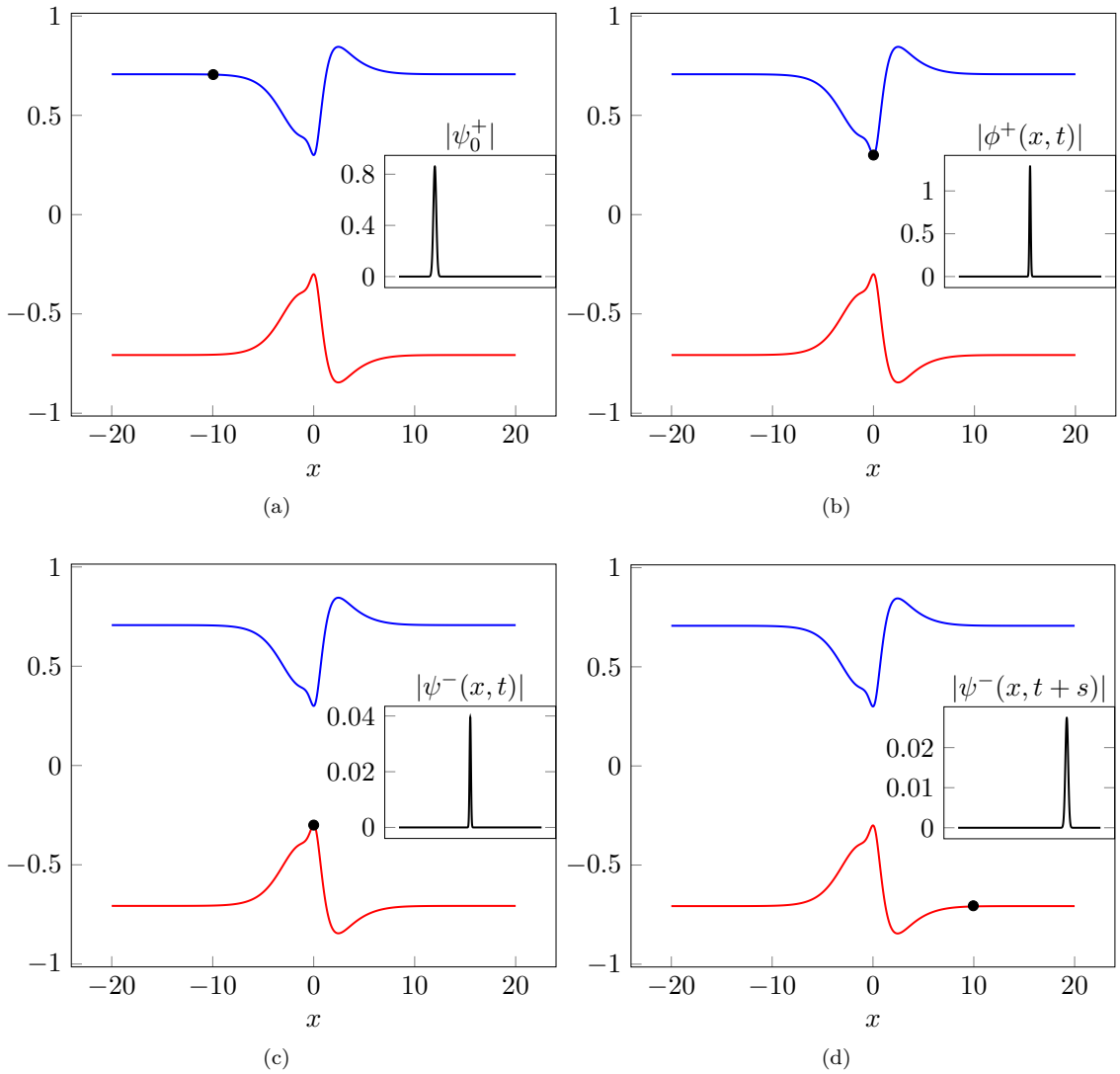


Figure D.2: Application of the 1D formula for a particular system discussed in [28]. The centre of mass of the associated wavepacket (inset) is represented by a black point on either the upper (blue) and lower (red) adiabatic potential surfaces.

where $x_{n+1}(\boldsymbol{p},\boldsymbol{q}), y_{n+1}(\boldsymbol{p},\boldsymbol{q})$ are given by the following algebraic recursive differential equations (where we omit the arguments of symbols to ease notation):

$$x_1 = z_1 = w_1 = 0, \quad y_1 = -\frac{i}{4\rho}(\boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}\theta). \tag{D.28}$$

and

$$y_n = 0,\ n \text{ even}, \quad x_n = z_n = w_n = 0,\ n \text{ odd}, \tag{D.29}$$

where $\rho = \rho(\boldsymbol{q})$. For n odd, we have

$$x_{n+1} = -\frac{1}{2\rho}\left[\frac{1}{i}(\boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}y_n) - 2\sum_{j=1}^{n}\frac{1}{(2i)^j j!}\sum_{|\boldsymbol{\alpha}|=j}\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}}(b_{\boldsymbol{\alpha}}z_{n+1-j} - a_{\boldsymbol{\alpha}}x_{n+1-j})\right], \tag{D.30}$$

and for n even

$$y_{n+1} = -\frac{1}{2\rho}\Bigg[\frac{1}{i}((\boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}x_n) - z_n(\boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}\theta)) \\ - 2\sum_{j=1}^{n}\frac{1}{(2i)^j j!}\sum_{|\boldsymbol{\alpha}|=j}\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}}(-a_{\boldsymbol{\alpha}}y_{n+1-j} + b_{\boldsymbol{\alpha}}w_{n+1-j})\Bigg], \quad \text{(D.31)}$$

$$\frac{1}{i}((\boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}z_n) - x_n(\boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}\theta)) = \sum_{j=1}^{n}\frac{1}{(2i)^j j!}\sum_{|\boldsymbol{\alpha}|=j}\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}}(b_{\boldsymbol{\alpha}}y_{n+1-j} + a_{\boldsymbol{\alpha}}w_{n+1-j}), \quad \text{(D.32)}$$

$$\frac{1}{i}(\boldsymbol{p}\cdot\nabla_{\boldsymbol{q}}w_n) = 2\sum_{j=1}^{n}\frac{1}{(2i)^j j!}\sum_{|\boldsymbol{\alpha}|=j}\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}}(a_{\boldsymbol{\alpha}}z_{n+1-j} + b_{\boldsymbol{\alpha}}x_{n+1-j}), \quad \text{(D.33)}$$

where $\boldsymbol{\alpha} = (\alpha_1, \alpha_2)$, $\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}} = \partial_{p_1}^{\alpha_1}\partial_{p_2}^{\alpha_2}$, and $a_{\boldsymbol{\alpha}} = a_{\boldsymbol{\alpha}}(\boldsymbol{q}), b_{\boldsymbol{\alpha}} = b_{\boldsymbol{\alpha}}(\boldsymbol{q})$ depend only on $\boldsymbol{q}$, and are given by the recursions

$$a_0 = \rho(\boldsymbol{q}), \quad b_0 = 0,$$
$$a_{(\alpha_1+1,\alpha_2)} = \partial_{q_1}a_{(\alpha_1,\alpha_2)} + (\partial_{q_1}\theta)b_{(\alpha_1,\alpha_2)}, \quad b_{(\alpha_1+1,\alpha_2)} = \partial_{q_1}b_{(\alpha_1,\alpha_2)} - (\partial_{q_1}\theta)a_{(\alpha_1,\alpha_2)},$$
$$a_{(\alpha_1,\alpha_2+1)} = \partial_{q_2}a_{(\alpha_1,\alpha_2)} + (\partial_{q_2}\theta)b_{(\alpha_1,\alpha_2)}, \quad b_{(\alpha_1,\alpha_2+1)} = \partial_{q_2}b_{(\alpha_1,\alpha_2)} - (\partial_{q_2}\theta)a_{(\alpha_1,\alpha_2)}.$$

Proof. The method is a straightforward extension of [27, Sections 2 and 3], in particular we direct the reader to Proposition 3.3 (page 3654). □

The result of Lemma 24 shows that x_n, y_n, z_n, w_n can be written as polynomials in $\boldsymbol{p}$ of order n, as the recursive definitions involve finite products, derivatives and sums of the initial x_0, y_0, z_0, w_0, which are polynomials in $\boldsymbol{p}$. We therefore write

$$x_n(\boldsymbol{p},\boldsymbol{q}) = \sum_{m=0}^{n}\sum_{k=0}^{m}p_1^k p_2^{m-k}x_n^{k,m-k}(\boldsymbol{q}), \quad \text{(D.34)}$$

for some $x_n^{k,m-k}(\boldsymbol{q})$, and similarly for y_n, z_n, w_n. For a given j, we write $\boldsymbol{\alpha}_j = (\alpha, j-\alpha)$ for each $\alpha \le j$.

Consider for example

$$\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}_j}x_{n+1-j} = \sum_{m=0}^{n+1-j}\sum_{k=0}^{m}(\partial_{p_1}^{\alpha}p_1^k)\big(\partial_{p_2}^{j-\alpha}p_2^{m-k}\big)x_{n+1-j}^{k,m-k}(\boldsymbol{q}),$$

where by a direct calculation

$$\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}_j}p_1^k p_2^{m-k} = \begin{cases}\frac{k!}{(m-k)!}\frac{(m-k)!}{(m-k-j+\alpha)!}p_1^{k-\alpha}p_2^{m-k-j+\alpha}, & k \ge \alpha \text{ and } m \ge j,\\ 0, & \text{otherwise.}\end{cases} \quad \text{(D.35)}$$

Therefore

$$\partial_{\boldsymbol{p}}^{\boldsymbol{\alpha}_j}x_{n+1-j} = \sum_{m=j}^{n+1-j}\sum_{k=\alpha}^{m-\alpha+j}\frac{k!}{(k-\alpha)!}\frac{(m-k)!}{(m-k-j+\alpha)!}p_1^{k-\alpha}p_2^{m-k-j+\alpha}x_{n+1-j}^{k,m-k}(\boldsymbol{q}),$$

so that

$$\mathcal{A}(\boldsymbol{p},\boldsymbol{q}) \coloneqq \sum_{j=1}^{n} \frac{1}{(2i)^j j!} \sum_{\alpha=0}^{j} a_{\boldsymbol{\alpha}_j} \partial_{p_1}^{\alpha} \partial_{p_2}^{j-\alpha} x_{n+1-j}(\boldsymbol{p},\boldsymbol{q}) \tag{D.36}$$

can be rewritten as

$$\sum_{j=1}^{n} \frac{1}{(2i)^j j!} \sum_{\alpha=0}^{j} a_{\boldsymbol{\alpha}_j} \sum_{m=j}^{n+1-j} \sum_{k=\alpha}^{m+\alpha-j} \frac{k!}{(k-\alpha)!} \frac{(m-k)!}{(m-k-j+\alpha)!} p_1^{k-\alpha} p_2^{m-k-j+\alpha} x_{n+1-j}^{k,m-k}(\boldsymbol{q}).$$

We now want to extract p_1 and p_2 from the final two summations, so that we can compare coefficients on either side of the results of lemma 24 to construct recursive equations for $x_n^{A,B}$ for $A+B<n$. Consider terms where $j > \frac{n+1}{2}$. By the limits of the third summand, we find that $m > \frac{n+1}{2}$, and that $m < \frac{n+1}{2}$, a contradiction. Therefore we restrict the limits of first summand. Defining $b = k-\alpha$, and $c = m-j$, we find

$$\mathcal{A} = \sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \sum_{\alpha=0}^{j} \sum_{c=0}^{n+1-2j} \sum_{b=0}^{c} \frac{a_{\boldsymbol{\alpha}_j} (b+\alpha)!((c+j)-(b+\alpha))!}{(2i)^j j! b! (c-b)!} p_1^b p_2^{c-b} x_{n+1-j}^{b+\alpha,(c+j)-(b+\alpha)}(\boldsymbol{q}).$$

We now want to switch the order of summation. We note that, for an arbitrary $\mathcal{B}$,

$$\sum_{j=1}^{\lfloor \frac{n+1}{2} \rfloor} \sum_{c=0}^{n+1-2j} \mathcal{B}_{c,j} = \sum_{c=0}^{n+1} \sum_{j=1}^{\lfloor \frac{c}{2} \rfloor} \mathcal{B}_{n+1-c,j},$$

which can be shown directly (note that the terms where $c=0$, $c=1$ are zero). Using this, we finally have that

$$\begin{aligned} \mathcal{A} = \sum_{c=0}^{n+1} \sum_{b=0}^{n+1-c} & p_1^b p_2^{n+1-c-b} \\ & \times \sum_{j=1}^{\lfloor \frac{c}{2} \rfloor} \sum_{\alpha=0}^{j} \frac{a_{\boldsymbol{\alpha}_j} (b+\alpha)!(n+1-c+j-b-\alpha)!}{(2i)^j j! b! (n+1-c-b)!} x_{n+1-j}^{b+\alpha,(n+1-c+j)-(b+\alpha)}(\boldsymbol{q}). \end{aligned} \tag{D.37}$$

Importantly, p_1 and p_2 have been extracted from two of the summations. Note that eq. (D.37) reduces to the 1D result in [27] for p_2 and p_1, by taking $b=0$ and $\alpha=0$, or $j-\alpha=0$ and $n+1-c-b=0$ respectively. We then obtain the following result.

Proposition 25. The coefficients $x_n^{A,B}(\boldsymbol{q})$ to $w_n^{A,B}(\boldsymbol{q})$ are determined by the following algebraic-differential recursive equations. We have (omitting arguments of symbols for ease of notation):

$$x_1^{A,B} = z_1^{A,B} = w_1^{A,B} = 0, \quad A+B \in \{0,1\}, \tag{D.38}$$

$$y_1^{0,0} = y_1^{1,1} = 0, \quad y_1^{1,0} = -\frac{i}{4\rho}\partial_{q_1}\theta, \quad y_1^{0,1} = -\frac{i}{4\rho}\partial_{q_2}\theta. \tag{D.39}$$

Further, when n is odd,

$$x_{n+1}^{A,B} = -\frac{1}{2\rho}\Bigg[\frac{1}{i}(\partial_{q_1} y_n^{A-1,B} + \partial_{q_2} y_n^{A,B-1}) - 2\sum_{j=1}^{\lfloor\frac{n+1-(A+B)}{2}\rfloor}\sum_{\alpha=0}^{j}\frac{1}{(2i)^j j!} \times \frac{(A+\alpha)!}{A!}\frac{(B+j-\alpha)!}{B!}\left(b_{\boldsymbol{\alpha}_j} z_{n+1-j}^{A+\alpha,B+j-\alpha} - a_{\boldsymbol{\alpha}_j} x_{n+1-j}^{A+\alpha,B+j-\alpha}\right)\Bigg], \quad \text{(D.40)}$$

When n is even, we have

$$y_{n+1}^{A,B} = -\frac{1}{2\rho}\Bigg[\frac{1}{i}((\partial_{q_1} x_n^{A-1,B} + \partial_{q_2} x_n^{A,B-1}) - (z_n^{A-1,B}\partial_{q_1}\theta + z_n^{A,B-1}\partial_{q_2}\theta)) - 2\sum_{j=1}^{\lfloor\frac{n+1-(A+B)}{2}\rfloor}\sum_{\alpha=0}^{j}\frac{1}{(2i)^j j!}\frac{(A+\alpha)!}{A!}\frac{(B+j-\alpha)!}{B!} \times \left(-a_{\boldsymbol{\alpha}_j} y_{n+1-j}^{A+\alpha,B+j-\alpha} + b_{\boldsymbol{\alpha}_j} w_{n+1-j}^{A+\alpha,B+j-\alpha}\right)\Bigg], \quad \text{(D.41)}$$

$$0 = \frac{1}{i}((\partial_{q_1} z_n^{A-1,B} + \partial_{q_2} z_n^{A,B-1}) + (x_n^{A-1,B}\partial_{q_1}\theta + x_n^{A,B-1}\partial_{q_2}\theta)) - 2\sum_{j=1}^{\lfloor\frac{n+1-(A+B)}{2}\rfloor}\sum_{\alpha=0}^{j}\frac{1}{(2i)^j j!}\frac{(A+\alpha)!}{A!}\frac{(B+j-\alpha)!}{B!} \times \left(b_{\boldsymbol{\alpha}_j} y_{n+1-j}^{A+\alpha,B+j-\alpha} + a_{\boldsymbol{\alpha}_j} w_{n+1-j}^{A+\alpha,B+j-\alpha}\right), \quad \text{(D.42)}$$

$$0 = \frac{1}{i}((\partial_{q_1} w_n^{A-1,B} + \partial_{q_2} w_n^{A,B-1}) - 2\sum_{j=1}^{\lfloor\frac{n+1-(A+B)}{2}\rfloor}\sum_{\alpha=0}^{j}\frac{1}{(2i)^j j!}\frac{(A+\alpha)!}{A!}\frac{(B+j-\alpha)!}{B!} \times \left(a_{\boldsymbol{\alpha}_j} z_{n+1-j}^{A+\alpha,B+j-\alpha} + b_{\boldsymbol{\alpha}_j} x_{n+1-j}^{A+\alpha,B+j-\alpha}\right). \quad \text{(D.43)}$$

Proof. We substitute eq. (D.34) into the results of lemma 24 and compare coefficients in powers of p_1, p_2 on either side, using eq. (D.37). □

As with the coefficients x_n and y_n in eq. (D.27), $\kappa_{n+1}^{\pm}$ has polynomial form:

$$\kappa_{n+1}^{\pm}(\boldsymbol{p},\boldsymbol{q}) = \sum_{m=0}^{n}\sum_{j=0}^{m} p_1^j p_2^{m-j}\kappa_{n+1}^{(j,m-j)\pm}(\boldsymbol{q}). \quad \text{(D.44)}$$

Here we apply assumption (A2): $\kappa_{n+1}^{\pm} \approx p_1^n \kappa_{n+1}^{(n,0)\pm}(\boldsymbol{q})$. In the one dimensional case this has been shown to be accurate for sufficiently large p, but in practice holds for much smaller values. By directly constructing the Weyl quantisation of $p_1^n \kappa_{n+1}^{(n,0)\pm}(\boldsymbol{q})$ as in [27, pg. 3570], we see that the effect of the coupling operator is negligible outside a small region near the avoided crossing, determined by the small parameter ε which shows that it is reasonable to take the leading term in $\kappa_{n+1}^{\pm}$. The 2D algebraic differential recursive equations then reduce to the one dimensional case in

[27]:

$$x_{n+1}^{n+1,0} \approx \frac{i}{2\rho}(\partial_{q_1} y_n^{n,0}),$$
$$y_{n+1}^{n+1,0} \approx \frac{i}{2\rho}((\partial_{q_1} x_n^{n,0})' - (\partial_{q_1}\theta) z_n^{n,0}), \quad 0 \approx \partial_{q_1} z_n^{n,0} + (\partial_{q_1}\theta) x_n^{n,0}. \tag{D.45}$$

To ease notation, redefine $x_{n+1} = x_{n+1}^{n+1,0}$, and similar for y_{n+1}, z_{n+1}. It is unclear what the analogue of eq. (D.22), introduced initially in [38] for the time-adiabatic case, would be for multidimensional systems. We introduce the natural scaling in the first dimension

$$\tau(\boldsymbol{q}) = 2\int_0^{q_1} \rho(r, q_2)\,\mathrm{d}r. \tag{D.46}$$

Defining $\tilde{f}(\tau(\boldsymbol{q})) = f(\boldsymbol{q})$ the recursive relations eq. (D.45) then become

$$\tilde{x}_{n+1}^0 = i\tilde{y}_{n+1}^0, \quad \tilde{y}_{n+1}^0 = i((\tilde{x}_n^0)' + \tilde{\theta}'\tilde{z}_n^0), \quad 0 = (\tilde{z}_n^0)' + \tilde{\theta}'\tilde{x}_n^0, \tag{D.47}$$

where $\tilde{\theta}' = \frac{d}{d\tau(q_1,q_2)}\tilde{\theta}$. These recursive equations also occur in [36], where they are solved in one dimension, under the assumption that

$$\frac{d}{d\tau}\tilde{\theta}(\tau) = \frac{i\gamma}{\tau - \bar{\tau}^{cz}} - \frac{i\gamma}{\tau - \tau^{cz}} + \tilde{\theta}_r'(\tau), \tag{D.48}$$

where τ^{cz} is a first order complex singularity of $\tilde{\theta}$, and $\tilde{\theta}_r$ has no singularities closer to the real axis than τ^{cz}. If the avoided crossing occurs at 0, we can write $\rho^2(q) = \delta^2 + g(q)^2$, for some analytic function g such that $g(0) \approx 0$, and g^2 is quadratic in the neighbourhood of $q = 0$. Therefore a Stokes line (*i.e.* a curve with $\mathrm{Im}(\rho) = 0$) crosses the real axis perpendicularly [39], and following this line leads to a pair of complex conjugate points $q^{cz}, \bar{q}^{cz}$ which are complex zeros of ρ. Defining $\tau^{cz} = \tau(q^{cz})$, it is shown in [38] that first order complex singularities of the adiabatic coupling function arise at these complex zeros. This derivation is still valid in our case, for each q_2. The recursive algebraic differential equations solved in [36] then give us κ_n^- to leading order:

$$\kappa_n^-(\boldsymbol{q}) \approx \kappa_{n,0}^-(\boldsymbol{q}) := \frac{i^n}{\pi}\rho(\boldsymbol{q})(n-1)!\left(\frac{i}{(\tau(\boldsymbol{q}) - \bar{\tau}^{cz}(q_2))^n} - \frac{i}{(\tau(\boldsymbol{q}) - \tau^{cz}(q_2))^n}\right). \tag{D.49}$$

It is clear that the results of this section can be extended to higher dimensions, by assuming the direction of travel of the wavepacket is in the first dimension. We will now use this observation to design an algorithm for multi-dimensional transitions using only the 1D transition formula.

D.5 Multi-dimensional formula derivation

The derivation of a multidimensional formula, under the assumptions above, follows similarly to the one dimensional case. We want to approximate the pseudodifferential operator K_n, which is given by the Weyl quantisation of κ_n. The polynomial form of κ_n allows us to simplify the Weyl quantisation as follows.

Proposition 26. Let $\kappa(\boldsymbol{p},\boldsymbol{q}) = g(\boldsymbol{q})\prod_{i=1}^{d} p_i^{A_i}$, for $A_i \in \mathbb{N}$. Then

$$\widehat{(\mathcal{W}_\varepsilon \kappa\psi)}^{\varepsilon}(\boldsymbol{k}) = \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} \widehat{g}^{\varepsilon}(\boldsymbol{k}-\boldsymbol{\eta}) \prod_{i=1}^{d} \left(\frac{k_i+\eta_i}{2}\right)^{A_i} \widehat{\psi}^{\varepsilon}(\boldsymbol{\eta})\,\mathrm{d}\boldsymbol{\eta}. \tag{D.50}$$

Proof. The proof is a multi-dimensional extension of [28, Lemma 4.1]. Firstly, using that $\psi(\boldsymbol{y}) = (2\pi\varepsilon)^{-d/2}\int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\eta}\widehat{\psi}^{\varepsilon}(\boldsymbol{\eta})\exp(i(\boldsymbol{\eta}\cdot\boldsymbol{y})/\varepsilon)$,

$$\begin{aligned}(\mathcal{W}_\varepsilon\kappa\psi)(\boldsymbol{x}) =& \frac{1}{(2\pi\varepsilon)^{d}} \int_{\mathbb{R}^{2d}} \mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{y} \left(\prod_{i=1}^{d}\xi_i^{A_i}\right) g\left(\frac{\boldsymbol{x}+\boldsymbol{y}}{2}\right) e^{\frac{i}{\varepsilon}(\boldsymbol{\xi}\cdot(\boldsymbol{x}-\boldsymbol{y}))}\psi(\boldsymbol{y}),\\ =& \frac{1}{(2\pi\varepsilon)^{\frac{3d}{2}}} \int_{\mathbb{R}^{3d}} d\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{y}\,\mathrm{d}\boldsymbol{\eta} \left(\prod_{i=1}^{d}\xi_i^{A_i}\right) g\left(\frac{\boldsymbol{x}+\boldsymbol{y}}{2}\right) e^{\frac{i}{\varepsilon}(\boldsymbol{\xi}\cdot(\boldsymbol{x}-\boldsymbol{y})+\boldsymbol{\eta}\cdot\boldsymbol{y})}\widehat{\psi}^{\varepsilon}(\boldsymbol{\eta}).\end{aligned}$$

Now define $\tilde{y}_i = (x_i+y_i)/2$, $i=1,\dots,d$. Then

$$\begin{aligned}(\mathcal{W}_\varepsilon\kappa\psi)(\boldsymbol{x}) &= \frac{2^d}{(2\pi\varepsilon)^{\frac{3d}{2}}} \int_{\mathbb{R}^{3d}} \mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\tilde{\boldsymbol{y}}\,\mathrm{d}\boldsymbol{\eta} \left(\prod_{i=1}^{d}\xi_i^{A_i}\right) g(\tilde{\boldsymbol{y}}) e^{\frac{i}{\varepsilon}(\boldsymbol{\xi}\cdot\boldsymbol{x}+(2\tilde{\boldsymbol{y}}-\boldsymbol{x})\cdot(\boldsymbol{\eta}-\boldsymbol{\xi}))}\psi(\boldsymbol{\eta}),\\ &= \frac{2^d}{(2\pi\varepsilon)^{d}} \int_{\mathbb{R}^{2d}} \mathrm{d}\boldsymbol{\xi}\,\mathrm{d}\boldsymbol{\eta} \left(\prod_{i=1}^{d}\xi_i^{A_i}\right) e^{\frac{i}{\varepsilon}(\boldsymbol{x}\cdot(2\boldsymbol{\xi}-\boldsymbol{\eta}))}\psi(\boldsymbol{\eta})\hat{g}^{\varepsilon}(2(\boldsymbol{\xi}-\boldsymbol{\eta})).\end{aligned}$$

We perform a second change of variables $\tilde{\xi}_i = 2\xi_i$ and find

$$(\mathcal{W}_\varepsilon\kappa\psi)(\boldsymbol{x}) = \frac{1}{(2\pi\varepsilon)^{d}} \int_{\mathbb{R}^{2d}} \mathrm{d}\tilde{\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{\eta} \left(\prod_{i=1}^{d}\left(\frac{\tilde{\xi}_i}{2}\right)^{A_i}\right) e^{\frac{i}{\varepsilon}(\boldsymbol{x}\cdot(\tilde{\boldsymbol{\xi}}-\boldsymbol{\eta}))}\hat{\psi}^{\varepsilon}(\boldsymbol{\eta})\hat{g}^{\varepsilon}(\tilde{\boldsymbol{\xi}}-2\boldsymbol{\eta}).$$

We apply the scaled Fourier transform to both sides of this equation:

$$\widehat{(\mathcal{W}_\varepsilon\kappa\psi)}^{\varepsilon}(\boldsymbol{k}) = \frac{1}{(2\pi\varepsilon)^{\frac{3d}{2}}} \int_{\mathbb{R}^{3d}} \mathrm{d}\tilde{\boldsymbol{\xi}}\,\mathrm{d}\boldsymbol{\eta}\,\mathrm{d}\boldsymbol{x} \left(\prod_{i=1}^{d}\left(\frac{\tilde{\xi}_i}{2}\right)^{A_i}\right) e^{\frac{i}{\varepsilon}(\boldsymbol{x}\cdot(\tilde{\boldsymbol{\xi}}-\boldsymbol{\eta}-\boldsymbol{k}))}\hat{\psi}^{\varepsilon}(\boldsymbol{\eta})\hat{g}^{\varepsilon}(\tilde{\boldsymbol{\xi}}-2\boldsymbol{\eta}).$$

Using that $(2\pi\varepsilon)^{-d}\int \mathrm{d}\boldsymbol{x}\exp(i(\boldsymbol{a}\cdot\boldsymbol{x})/\varepsilon) = \delta(\boldsymbol{a})$ allows us to directly compute the $\boldsymbol{x}$ integral, giving eq. (D.50). □

Next we linearise the dynamics near the avoided crossing. By (A3), to leading order the uncoupled propagators in eq. (D.16) can be approximated by

$$H_1^{\pm} = -\frac{\varepsilon^2}{2}\nabla_{\boldsymbol{x}}^2 \pm \delta + \boldsymbol{\lambda}\cdot\boldsymbol{x}. \tag{D.51}$$

Then, by the fundamental theorem of calculus,

$$e^{\frac{i}{\varepsilon}sH_1^{\pm}} - e^{\frac{i}{\varepsilon}sH^{\pm}} = \left\{\int_0^s e^{\frac{i}{\varepsilon}rH_1^{\pm}}\left[\frac{i}{\varepsilon}(H_1^{\pm}-H^{\pm})\right]e^{-\frac{i}{\varepsilon}rH^{\pm}}\,\mathrm{d}r\right\}e^{\frac{i}{\varepsilon}sH^{\pm}}. \tag{D.52}$$

Since $H_1^{\pm} - H^{\pm}$ is quadratic near zero, the integrand in eq. (D.52) is of order 1 in an $\sqrt{\varepsilon}$-neighbourhood of zero. Outside of this region the coupling function provides a negligible result, as seen in the one dimensional case [27]. We also use the d dimensional Avron-Herbst formula [37],

which shows that

$$e^{-\frac{i}{\varepsilon}s\widehat{H_1^{\pm}}^{\varepsilon}} = e^{-\frac{i\|\boldsymbol{\lambda}\|^2 s^3}{6\varepsilon}} e^{s(\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{k}})} e^{-\frac{i}{2\varepsilon}((\|\boldsymbol{k}\|^2\pm 2\delta)s-(\boldsymbol{\lambda}\cdot\boldsymbol{k})s^2)}. \tag{D.53}$$

Then

$$\begin{aligned}\widehat{\psi_n^-}^{\varepsilon}(\boldsymbol{k},t) \approx -i\varepsilon^n e^{-\frac{i}{\varepsilon}t\widehat{H^-}^{\varepsilon}} \int_{-\infty}^{t} e^{-\frac{i\|\boldsymbol{\lambda}\|^2 s^3}{6\varepsilon}} e^{s(\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{k}})} e^{-\frac{i}{2\varepsilon}((\|\boldsymbol{k}\|^2-2\delta)s-(\boldsymbol{\lambda}\cdot\boldsymbol{k})s^2)} \widehat{K_{n+1}^-}^{\varepsilon} \\ \times e^{-\frac{i\|\boldsymbol{\lambda}\|^2 s^3}{6\varepsilon}} e^{s(\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{k}})} e^{-\frac{i}{2\varepsilon}((\|\boldsymbol{k}\|^2+2\delta)s-(\boldsymbol{\lambda}\cdot\boldsymbol{k})s^2)} \widehat{\phi^+}^{\varepsilon}(\boldsymbol{k})\,\mathrm{d}s.\end{aligned} \tag{D.54}$$

Using proposition 26 for the coupling function shows that

$$\begin{aligned}\widehat{\psi_n^-}^{\varepsilon}(\boldsymbol{k},t) \approx -i\frac{\varepsilon^n}{(2\pi\varepsilon)^{d/2}} e^{-\frac{i}{\varepsilon}t\widehat{H^-}^{\varepsilon}} \int_{-\infty}^{t} \mathrm{d}s e^{-\frac{i\|\boldsymbol{\lambda}\|^2 s^3}{6\varepsilon}} e^{s(\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{k}})} e^{-\frac{i}{2\varepsilon}((\|\boldsymbol{k}\|^2-2\delta)s-(\boldsymbol{\lambda}\cdot\boldsymbol{k})s^2)} \\ \times \int_{\mathbb{R}^d} \mathrm{d}\boldsymbol{\eta} \left\{ \sum_{A_i=1,i=1,..,d}^{n+1} \widehat{\kappa_{n+1}^{\boldsymbol{A},-}}^{\varepsilon}(\boldsymbol{k}-\boldsymbol{\eta}) \left(\prod_{i=1}^{d} \left(\frac{k_i+\eta_i}{2} \right)^{A_i} \right) \right\} \\ \times e^{-\frac{i(\|\boldsymbol{\lambda}\|^2 s^3}{6\varepsilon}} e^{s(\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{\eta}})} e^{-\frac{i}{2\varepsilon}((\|\boldsymbol{\eta}\|^2+2\delta)s-(\boldsymbol{\lambda}\cdot\boldsymbol{\eta})s^2)} \widehat{\phi^+}^{\varepsilon}(\boldsymbol{\eta_1}),\end{aligned}$$

where $\boldsymbol{A} = (A_1 \dots A_d)$. The operator $e^{s\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{k}}}$ is a *shift operator*, so $e^{s\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{k}}} f(\boldsymbol{k}) = f(\boldsymbol{k}+\boldsymbol{\lambda}s)$. Instead of applying the shift operator to the right, we use the fact that the integral is invariant under the transform $\boldsymbol{\eta} \mapsto \boldsymbol{\eta} - \boldsymbol{\lambda}s$ to apply it to the left: in this case $f(\boldsymbol{\eta})e^{-s\boldsymbol{\lambda}\cdot\partial_{\boldsymbol{\eta}}} = f(\boldsymbol{\eta}-\boldsymbol{\lambda}s)$. The following transformations take place in the integrand:

$$\begin{aligned}&\widehat{\kappa_{n+1}^{\boldsymbol{A},-}}^{\varepsilon}(\boldsymbol{k}-\boldsymbol{\eta}) \mapsto \widehat{\kappa_{n+1}^{\boldsymbol{A},-}}^{\varepsilon}(\boldsymbol{k}-\boldsymbol{\eta}), \quad \boldsymbol{k}+\boldsymbol{\eta} \mapsto \boldsymbol{k}+\boldsymbol{\eta}-2\boldsymbol{\lambda}s, \\ &e^{\frac{i}{2\varepsilon}((\|\boldsymbol{k}\|^2\pm 2\delta)s-(\boldsymbol{\lambda}\cdot\boldsymbol{k})s^2)} \mapsto e^{\frac{i}{2\varepsilon}((\|\boldsymbol{k}-\boldsymbol{\lambda}s\|^2\pm 2\delta)s-(\boldsymbol{\lambda}\cdot(\boldsymbol{k}-\boldsymbol{\lambda}s))s^2)}.\end{aligned}$$

Rearranging gives

$$\begin{aligned}\widehat{\psi_n^-}^{\varepsilon}(\boldsymbol{k},t) \approx -i\frac{\varepsilon^n}{(2\pi\varepsilon)^{d/2}} e^{-\frac{i}{\varepsilon}t\widehat{H^-}^{\varepsilon}} \\ \times \int_{-\infty}^{t}\int_{\mathbb{R}^d} \mathrm{d}s\,\mathrm{d}\boldsymbol{\eta} \left\{ \sum_{A,B=1}^{n+1} \widehat{\kappa_{n+1}^{\boldsymbol{A},-}}^{\varepsilon}(\boldsymbol{k}-\boldsymbol{\eta}) \left(\prod_{i=1}^{d} \left(\frac{k_i+\eta_i-2\lambda_i s}{2} \right)^{A_i} \right) \right\} \\ \times \widehat{\phi^+}^{\varepsilon}(\boldsymbol{\eta}) \exp\left\{ \frac{i}{2\varepsilon} \left[(\|\boldsymbol{k}\|^2 - \|\boldsymbol{\eta}\|^2 - 4\delta)s - (\boldsymbol{\lambda}\cdot(\boldsymbol{k}-\boldsymbol{\eta}))s^2 \right] \right\}.\end{aligned} \tag{D.55}$$

We approximate κ_{n+1}^- with eq. (D.49), then calculate the scaled Fourier transform:

$$\begin{aligned}&\widehat{\kappa_n^-}^{\varepsilon}(\boldsymbol{k}) = \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} e^{-(i/\varepsilon)\boldsymbol{k}\cdot\boldsymbol{q}} \kappa_n^-(\boldsymbol{q})\,\mathrm{d}\boldsymbol{q}, \\ &\approx \frac{(n-1)!}{(2\pi\varepsilon)^{d/2}} \frac{i^n}{\pi} \int_{\mathbb{R}^d} \rho(\boldsymbol{q}) \left[\frac{i}{(\tau(\boldsymbol{q}) - \bar{\tau}^{cz}(\boldsymbol{q}^{d-1}))^n} - \frac{i}{(\tau(\boldsymbol{q}) - \tau^{cz}(\boldsymbol{q}^{d-1}))^n} \right] e^{-(i/\varepsilon)\boldsymbol{k}\cdot\boldsymbol{q}}\,\mathrm{d}\boldsymbol{q},\end{aligned}$$

where $\boldsymbol{q}^{d-1} = (q_2, \dots, q_d)$. Using 6 $\rho(\boldsymbol{q}) \approx \rho(q_1)$ and consequently $\tau(\boldsymbol{q}) = \tau(q_1)$, $\tau^{cz}(\boldsymbol{q}^{d-1}) = \tau^{cz}$. Therefore the Fourier transform in other dimensions produces a Dirac function, $\frac{1}{\sqrt{2\pi\varepsilon}} \int_{-\infty}^{\infty} e^{-\frac{ikx}{\varepsilon}}\,\mathrm{d}x = \sqrt{2\pi\varepsilon}\delta(k)$. As $\tau(\boldsymbol{q}) \approx \tau(q_1)$, we only need to consider the one dimensional case. This is discussed

in [27]. A simple extension to d dimensions therefore shows that

$$\widehat{\kappa_{n,0}^{-}}^{\varepsilon}(\boldsymbol{k}) = \frac{i}{\sqrt{2\pi\varepsilon}}\left(\frac{k_1}{2\delta\varepsilon}\right)^{n-1} e^{-i\tau_r\frac{k_1}{2\delta\varepsilon}} e^{-\tau_c\frac{|k_1|}{2\delta\varepsilon}} \sqrt{2\pi\varepsilon}^{(d-1)}\delta(k_2,\ldots,k_d). \tag{D.56}$$

We insert eq. (D.56) into eq. (D.55), and rearrange to find

$$\begin{aligned}\widehat{\psi_n^{-}}^{\varepsilon}(\boldsymbol{k},t) = \frac{1}{4\pi\varepsilon}e^{-\frac{i}{\varepsilon}t\widehat{H^{-}}^{\varepsilon}}\int_0^\infty \mathrm{d}s\int_{\mathbb{R}}\mathrm{d}\eta_1\left(\frac{k_1^2-\eta_1^2}{4\delta}\right)^n\left(1-\frac{2\lambda_1 s}{k_1+\eta_1}\right)^{n+1}\\ \times e^{-\frac{i\tau_r(k_1-\eta_1)}{2\delta\varepsilon}}e^{-\frac{\tau_c(|k_1-\eta_1|)}{2\delta\varepsilon}}\\ \times\left\{\int_{\mathbb{R}^{d-1}}\mathrm{d}\eta_2\ldots\mathrm{d}\eta_d\widehat{\phi^{+}}^{\varepsilon}(\boldsymbol{\eta})e^{\frac{i}{2\varepsilon}[(\|\boldsymbol{k}\|^2-\|\boldsymbol{\eta}\|^2)s-\boldsymbol{\lambda}\cdot(\boldsymbol{k}-\boldsymbol{\eta})s^2]}\delta(k_2-\eta_2,\ldots,k_d-\eta_d)\right\}.\end{aligned}$$

By the identity $f(x)=\int_{-\infty}^{\infty}\delta(x-a)f(a)\,\mathrm{d}a$, the integral in the dimensions $2,\ldots,d$ can be evaluated to find

$$\begin{aligned}\widehat{\psi_n^{-}}^{\varepsilon}(\boldsymbol{k},t) = \frac{1}{4\pi\varepsilon}e^{-\frac{i}{\varepsilon}t\widehat{H^{-}}^{\varepsilon}}\int_0^\infty \mathrm{d}s\int_{\mathbb{R}}\mathrm{d}\eta_1\left(\frac{k_1^2-\eta_1^2}{4\delta}\right)^n\left(1-\frac{2\lambda_1 s}{k_1+\eta_1}\right)^{n+1}\\ \times e^{-\frac{i\tau_r(k_1-\eta_1)}{2\delta\varepsilon}}e^{-\frac{\tau_c(|k_1-\eta_1|)}{2\delta\varepsilon}}\\ \times\widehat{\phi^{+}}^{\varepsilon}(\eta_1,k_2,\ldots,k_d)e^{\frac{i}{2\varepsilon}[(|k_1|^2-|\eta_1|^2-4\delta)s-\lambda_1(k_1-\eta_1)s^2]}.\end{aligned}$$

By (A1), λ_1 is small and so can be neglected, so that

$$\begin{aligned}\widehat{\psi_n^{-}}^{\varepsilon}(\boldsymbol{k},t) =&\frac{1}{4\pi\varepsilon}e^{-\frac{i}{\varepsilon}t\widehat{H^{-}}^{\varepsilon}}\int_0^\infty \mathrm{d}s\int_{\mathbb{R}}\mathrm{d}\eta\left(\frac{k_1^2-\eta_1^2}{4\delta}\right)^n e^{-\frac{i\tau_r(k_1-\eta_1)}{2\delta\varepsilon}}e^{-\frac{\tau_c(|k_1-\eta_1|)}{2\delta\varepsilon}}\\ &\times\widehat{\phi^{+}}^{\varepsilon}(\eta_1,k_2,\ldots,k_d)e^{\frac{i}{2\varepsilon}(|k_1|^2-|\eta_1|^2-4\delta)s}.\end{aligned} \tag{D.57}$$

From here we can follow the derivation in [27] and obtain an extension of its main result to d dimensions, given by eq. (D.20). In this derivation, cancellations in the integral remove all dependence on n. Therefore for implementation of eq. (D.20) we do not need to calculate the pseudodifferential operators $K_{n+1}^{\pm}$, or in fact find the optimal choice for n, but have utilised superadiabatic representations in its construction.

As justification for the proposed algorithm we note that we evolve the wavepacket on the new potential energy surface, restricted to each strip. As such, we discard any part of the wavepacket that leaves the strip and ignore any additional parts entering from other strips. Since the Schrödinger equation is linear, this introduces two types of error, due to: (i) the modification of the potential in each strip, and (ii) the wavepacket broadening out of the selected strip, or into it from the outside. Both errors are small, the first because the strip is quite narrow (so the potential is approximately constant), the second because the time that we actually evolve for is small (of the order of the crossing region in the optimal superadiabatic basis).

In practice, for the examples in Appendix D.6, we compute the BOA dynamics on a uniform 2-dimensional grid. Once the centre of mass of the wavepacket reaches the avoided crossing, we interpolate the wavepacket onto a grid with the new p_1 direction parallel to that of $\boldsymbol{p}_{\mathrm{COM}}$. Instead of treating strips of the appropriate width, we simply apply the formula eq. (D.20) along each of the 1D lines parallel to p_1 (or $\boldsymbol{p}_{\mathrm{COM}}$); this reduces to applying the 1D formula. For small ε, this is

essentially equivalent to the algorithm above as the approximate potentials of neighbouring lines are very similar and the evolution time in the optimal superadiabatic basis is very short.

D.6 Numerical results

We perform the algorithm on a selection of examples, and compare it to the two level 'exact' computation, where the Strang splitting method is used. For all examples we consider two wavepackets given in momentum space by:

$$\widehat{\psi_0}^{\varepsilon}(\boldsymbol{p}) = \frac{1}{N_\psi}\exp\left(-\frac{\|\boldsymbol{p}-\boldsymbol{p}_0\|^2}{2\varepsilon}\right)\exp\left(-i\frac{(\boldsymbol{p}-\boldsymbol{p}_0)\cdot\boldsymbol{x}_0}{\varepsilon}\right), \tag{D.58}$$

$$\widehat{\phi}^{\varepsilon}(\boldsymbol{p}) = \frac{1}{N_\phi}\exp\left(-\frac{\|\boldsymbol{p}-\boldsymbol{p}_0\|^6}{2\varepsilon}\right)\exp\left(-i\frac{(\boldsymbol{p}-\boldsymbol{p}_0)\cdot\boldsymbol{x}_0}{\varepsilon}\right), \tag{D.59}$$

where N_α are normalisation constants. To ensure that the wavepacket has sufficient momentum to travel through the avoided crossing, we choose to define the wavepackets at the avoided crossing point, then evolve backwards in time away from the avoided crossing using one level dynamics, before evolving forwards and applying the formula. In practice the initial wavepacket can be given in any initial location, provided it is far enough from the avoided crossing to be unaffected by coupling effects.

To compare the formula results to exact calculations we use the L^2-relative error:

$$Er_{\text{rel}}(\psi_1,\psi_2) = \max\left(\frac{\|\psi_1 \pm \psi_2\|}{\|\psi_1\|}, \frac{\|\psi_1 \pm \psi_2\|}{\|\psi_2\|}\right), \tag{D.60}$$

Where $\|\cdot\|$ is the standard L^2-norm. For comparison to other algorithms which do not calculate phase, it is also beneficial to consider the relative absolute error

$$Er_{\text{abs}}(\psi_1,\psi_2) = \max\left(\frac{\||\psi_1|-|\psi_2|\|}{\|\psi_1\|}, \frac{\||\psi_1|-|\psi_2|\|}{\|\psi_2\|}\right). \tag{D.61}$$

or the relative mass error

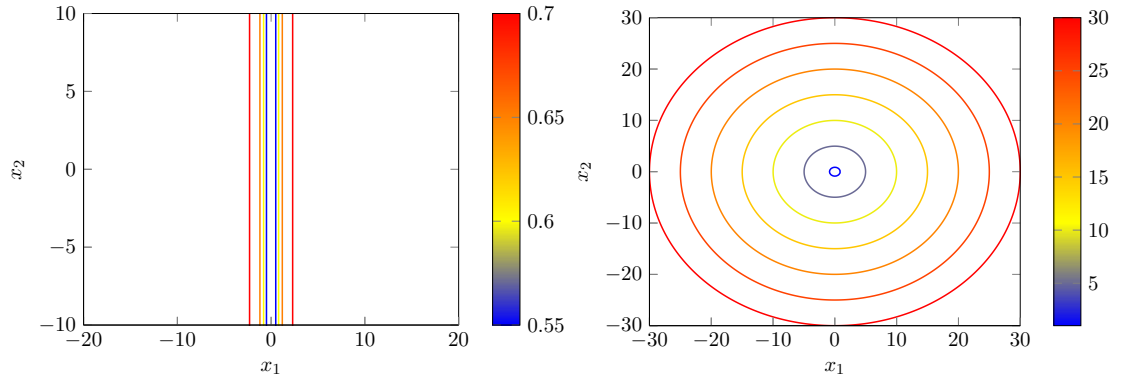
$$Er_{\text{mass}}(\psi_1,\psi_2) = \max\left(\frac{\|\psi_1\|}{\|\psi_2\|}, \frac{\|\psi_2\|}{\|\psi_1\|}\right) - 1. \tag{D.62}$$

Example 27. Consider the diabatic potential matrix

$$V(\boldsymbol{x}) = \begin{pmatrix} \tanh(x_1) & \delta \\ \delta & -\tanh(x_1) \end{pmatrix}. \tag{D.63}$$

This is a direct extension of a one dimensional problem, and as there is no dependence in x_2, the assumptions made in the derivation in Appendix D.5 are exactly valid, if the direction of the wavepacket is independent of p_2. The lower surface is given by $V_L = -V_U$. The upper adiabatic surface is shown in fig. D.3a. We take parameters

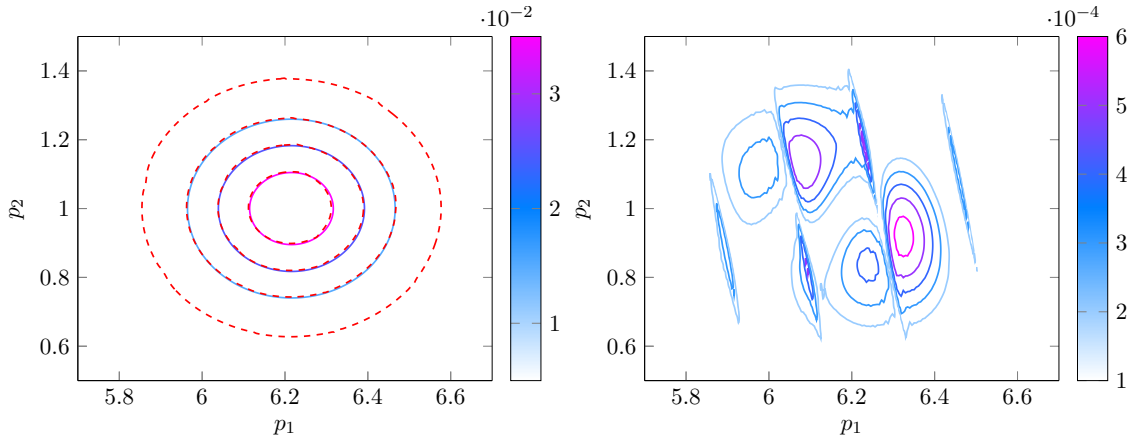
$$\{\varepsilon, \delta, \boldsymbol{p}_0, \boldsymbol{x_0}\} = \left\{\frac{1}{30}, \frac{1}{2}, (6,1), (0,0)\right\}. \tag{D.64}$$



(a) (b)

Figure D.3: Contour plot of the upper adiabatic potential surfaces for example 27 (left) and example 28 (right). In these examples, $V_U = -V_L$.

Using a mesh of $2^{13} \times 2^{13}$ points on the domain $[-20, 20]^2$, starting at time 0, we evolve the wavepacket back to time -2 with time-step $1/(50\|\boldsymbol{p}_0\|)$, then evolve forwards to time 2, applying the algorithm, and compare to the exact calculation. For the Gaussian wavepacket ψ, $Er_{\text{rel}} = 0.0151$, $Er_{\text{abs}} = 0.0151$, and $Er_{\text{mass}} = 0.0016$. For non-Gaussian ϕ $Er_{\text{rel}} = 0.0389$, $Er_{\text{abs}} = 0.0387$, and $Er_{\text{mass}} = 0.0023$. The result of the formula and corresponding error are shown in figs. D.4 and D.5.



(a) (b)

Figure D.4: Results for example 27, when using parameters in eq. (D.64) with initial wavepacket of form eq. (D.58). Left: exact calculation (solid line) versus formula result (dashed line). Contours for the formula result are at the same values as the neighbouring exact contours. Right: relative error.

Example 28. We consider the diabatic potential matrix described in [23]

$$V(\boldsymbol{x}) = \begin{pmatrix} x_1 & \sqrt{x_2^2 + \delta^2} \\ \sqrt{x_2^2 + \delta^2} & -x_1 \end{pmatrix}, \tag{D.65}$$

which is a modified Jahn-Teller diabatic potential, where the conical intersection is replaced with

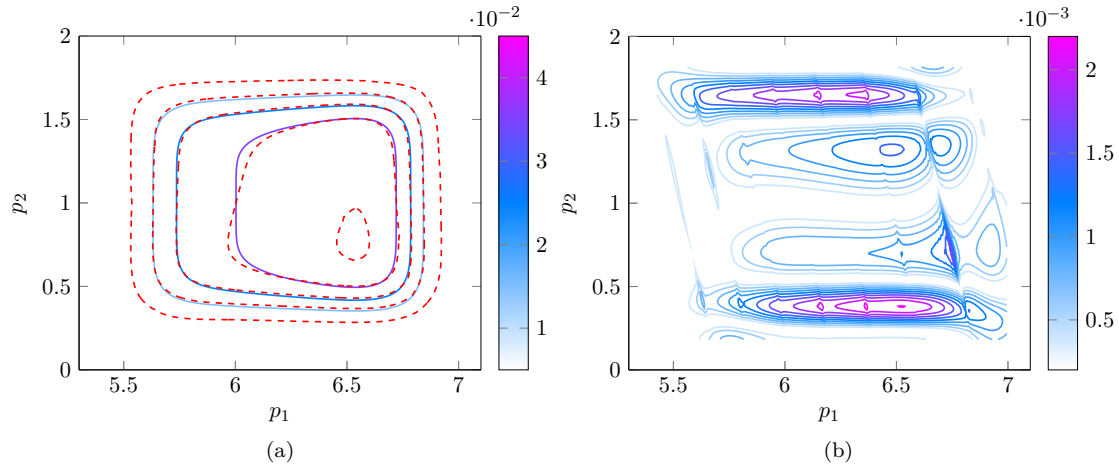


Figure D.5: As in fig. D.4, but with initial wavepacket eq. (D.59).

an avoided crossing with gap 2δ. The upper adiabatic surface is shown in fig. D.3b. We use parameters

$$\{\varepsilon, \delta, \boldsymbol{p}_0, \boldsymbol{x_0}\} = \left\{\frac{1}{30}, 0.5, (5,2), (0,0)\right\}, \tag{D.66}$$

a mesh of $2^{13} \times 2^{13}$ points on the domain $[-40, 40]^2$, we start at time 0, and evolve backwards with time-step $1/(50\|\boldsymbol{p}_0\|)$ to time $-20/\|\boldsymbol{p}_0\|^2$, then forwards to $20/\|\boldsymbol{p}_0\|^2$, we find $Er_{\text{rel}} = 0.0351, Er_{\text{abs}} = 0.0304$, and $Er_{\text{mass}} = 0.0029$ using Gaussian initial wavepacket ψ_0, and $Er_{\text{rel}} = 0.0679, Er_{\text{abs}} = 0.0616$, and $Er_{\text{mass}} = 0.0033$ for non-Gaussian initial wavepacket ϕ. Figures D.6 and D.7 display the result of the formula compared to the exact calculation. We now use the

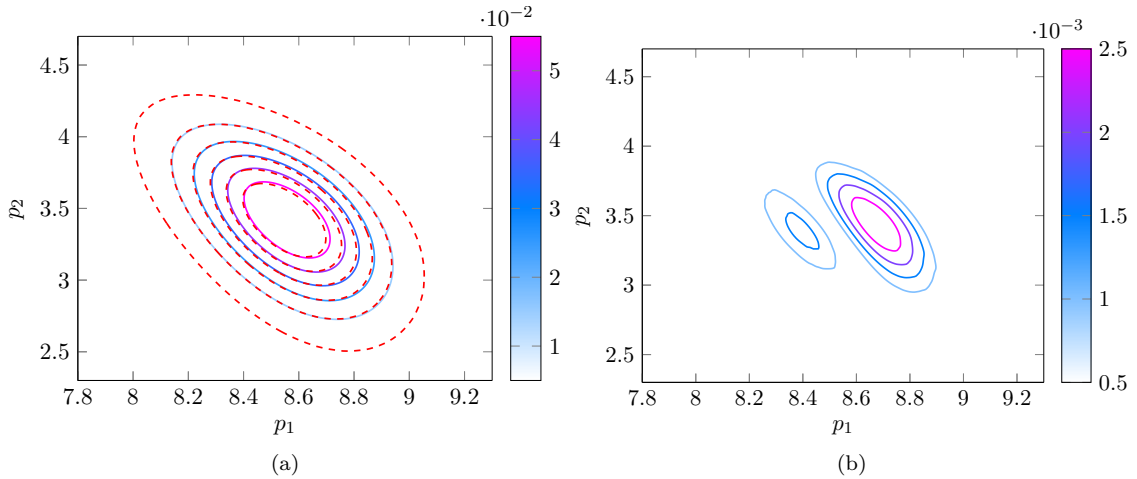


Figure D.6: Results for example 28, when using parameters in eq. (D.66) with initial wavepackets of form eq. (D.58). Results are presented as in fig. D.4.

parameters

$$\{\varepsilon, \delta, \boldsymbol{p}_0, \boldsymbol{x_0}\} = \left\{\frac{1}{30}, 0, (5,0), (0,0.5)\right\}. \tag{D.67}$$

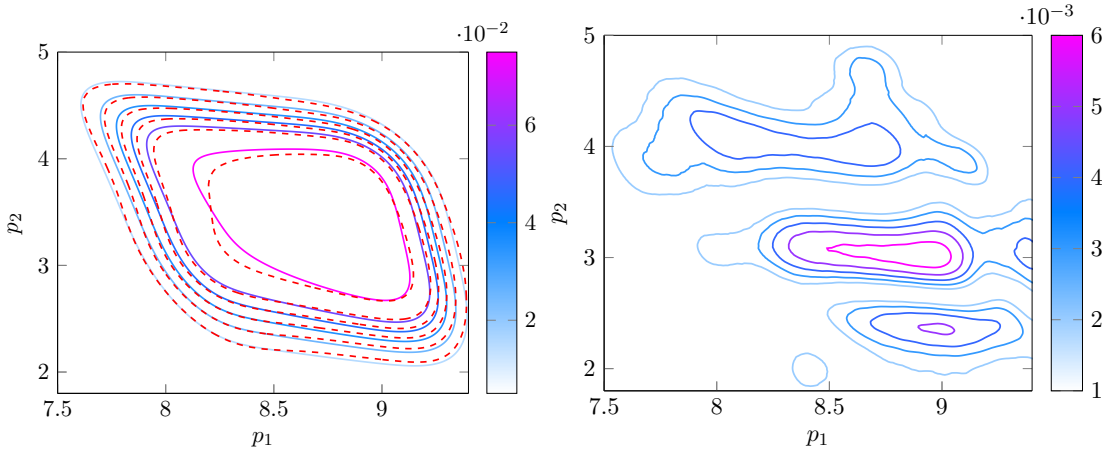


Figure D.7: As in fig. D.6, but with initial wavepacket eq. (D.59).

In addition, we included the sign of x_2 in the off-diagonal elements of $V(\mathbf{x})$, which then gives the standard Jahn-Teller Hamiltonian. However, let us stress that non-adiabatic transitions must be exactly the same for the Hamiltonian with and without the sign included. The reason is that by that choice, we have just chosen a different diabatic representation, but the (unique) adiabatic representation remains the same. It is an advantage of our method, which only uses the adiabatic energy surfaces, that it is insensitive to such a change. The Jahn-Teller Hamiltonian has a conical intersection. We have chosen momentum such that the centre of mass of the wavepacket does not cross the intersection. We evolve back to $-25/\|\boldsymbol{p}_0\|^2$ with a time-step of $1/(50\|\boldsymbol{p}_0\|)$, then evolve forwards to $25/\|\boldsymbol{p}_0\|^2$ using the algorithm, and compare with the exact calculation. Then $Er_{\text{rel}} = 0.0638, Er_{\text{abs}} = 0.0550$, and $Er_{\text{mass}} = 0.0309$ for initial wavepacket of form ψ_0 and $Er_{\text{rel}} = 0.1511, Er_{\text{abs}} = 0.0850$, and $Er_{\text{mass}} = 0.0604$ for ϕ, the transmitted wavepacket and error is given in fig. D.6. Although the relative error is large in this final calculation, the absolute error and mass error shows that the algorithm has performed well, given that it is not designed for systems where δ is small or vanishing. fig. D.9 also shows that the shape of the wavepacket is still well approximated qualitatively.

We note that the relative and absolute error in example 28 differ, while in example 27 they are the same. We believe this is due to a change in phase when ρ is not flat in q_2, so the error due to the modification of the potential surface for each strip is larger.

D.7 Conclusions and future work

In this paper we have constructed an algorithm which can be used to approximate the transmitted wavepacket in non-adiabatic transitions in multiple dimensions, by constructing a formula based on the one dimensional result in [28], and appealing to the linearity of the Schrödinger equation to decompose the dynamics onto strips with potentials that are constant in all but one direction. Presented examples in two dimensions show similar accuracy to one dimensional analogues, and are accurate in the phase, which is beyond the capability of standard surface hopping models.

Correctly approximating the phase of the wavepacket becomes important when more than one

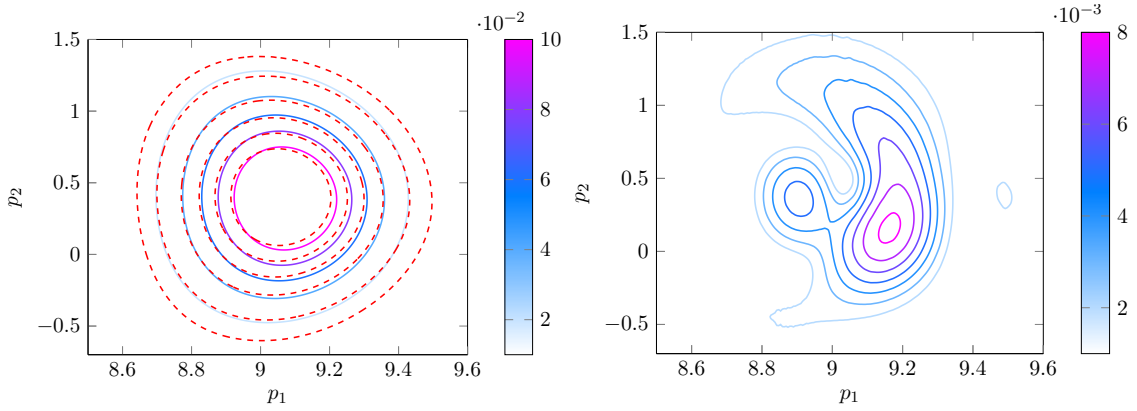


Figure D.8: As in fig. D.6, but with parameters eq. (D.67).

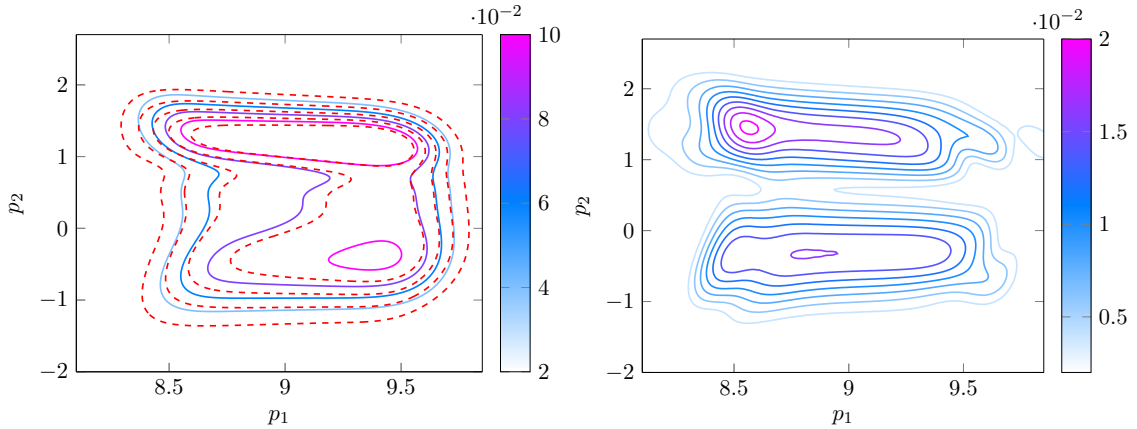


Figure D.9: As in fig. D.8, but with initial wavepacket eq. (D.59).

transition takes place. In future work we will consider multiple transitions in two dimensions using the algorithm. This will involve taking into account the effect of geometric phase [40] due to multiple avoided crossings, as well as constructing an approximation of the wavepacket which remains on the upper level after a transition has taken place. We also will compare the results of the algorithm considered in this paper with other algorithms designed to approximate non-adiabatic transitions, *e.g.* [41].

D.8 Details on code

Here we put some information about the code provided as supplementary material, available in the QMD section of the thesis code, available upon request: please email `t.hurst@sms.ed.ac.uk` for more information. The code provided compares exact calculations (using a Strang splitting method) with results produced by the algorithm algorithm 22, in one and two dimensions, for one avoided crossing. The code is separated into seven primary folders: [`Wavepackets.`] This folder contains functions to construct eqs. (D.58) and (D.59), for given parameters.

[`Potentials.`] This folder contains the potentials used in examples 27 and 28.

[`FormulaCalculations.`] All calculations needed to construct the formula result in 1D and 2D are in this folder. For 2D calculations, if the direction of momentum of the wavepacket depends on the second dimension, it is necessary to use the `Rotated` functions, which slightly increase computation time due to interpolations.

[`Examples.`] This folder contains scripts to produce the results of examples in this paper.

[`ExactCalculations.`] This folder includes all calculations needed for one and two-level dynamics with the Schrödinger equation using a Strang Splitting method, which is detailed below. The function `DetectAvoidedCrossing` finds the first local minimum of the potential surface and applies the formula at this point. The function `ExactTwoLevelWithFormula` can be used to evolve the exact and formula results simultaneously, for comparison.

[`Computations`.] These functions can be used to compute the formula result using `makeFormula`, the exact result using `makeExact`, or both using `ExactvsFormula`, for a given wavepacket.

[`Auxiliary`.] This folder contains auxiliary functions used for the handling of structs and variables in the code, for plotting and creating `.avi` animations, for saving results of computations and ε-Fourier transforms.

The code allows the user to choose different parameters for the examples, where to compare the exact and formula result, and whether to plot the dynamics. If parameters are not given, defaults are selected. After computations are calculated once the results are saved in `..\Data`. If the same parameters are used again the data can then be loaded instead of recomputed.

We now detail some of the numerical methods used in the computations.

D.8.1 The scaled Fourier transform

We make use to the `MATLAB` fast Fourier transform `fftn` in d dimensions, to apply the scaled d-dimensional Fourier transform. The d dimensional ε-Fourier transform is given by

$$\widehat{f^{\varepsilon}}(\boldsymbol{p}) = \frac{1}{(2\pi\varepsilon)^{d/2}} \int_{\mathbb{R}^d} f(\boldsymbol{x}) \exp\left(-\frac{i}{\varepsilon}\boldsymbol{p}\cdot\boldsymbol{x}\right) d\boldsymbol{x}. \tag{D.68}$$

The Matlab function `fftn` for a function f in d dimensions at $\boldsymbol{N} = (N_1, \ldots, N_d) \in \mathbb{N}^d$ spacial points, with values $f(\boldsymbol{j})$, at point $\boldsymbol{j}$, is defined for momentum point $\boldsymbol{k}$ with values $F(\boldsymbol{k})$, where

$$\begin{aligned} F(\boldsymbol{k}) = \sum_{j_1=1}^{N} \cdots \sum_{j_d=1}^{N} f(j_1, \ldots, j_d) \exp\left(-\frac{2\pi i}{N_1}(j_1-1)(k_1-1)\right) \ldots \\ \cdots \times \exp\left(-\frac{2\pi i}{N_d}(j_d-1)(k_d-1)\right), \end{aligned} \tag{D.69}$$

(where we have used $j-1, k-1$, as counting in `MATLAB` starts at 1) which is the discrete approximation to the integral

$$\widehat{f}(\boldsymbol{p}) = \int_{\mathbb{R}^d} f(\boldsymbol{x}) \exp(-2\pi i \boldsymbol{p}\cdot\boldsymbol{x}) d\boldsymbol{x}. \tag{D.70}$$

We compute on a bounded domain $\mathcal{D} \in \mathbb{R}^d$. In each dimension m we are therefore considering an interval $[x_m^S, x_m^E]$. We discretize each interval using N_m points, where

$$\begin{aligned} x_m(1) = x_m^S, \quad x_m(N_m) = x_m^E, \\ dx_m \frac{x_m(N_m) - x_m(1)}{N_m - 1}, \\ x_m(j_m) = x_m(1) + (j_m - 1)dx_m, \quad j_m = 1, \ldots, N_m \end{aligned}$$

The corresponding momentum dimension is discretized with

$$p_m(k_m + 1) = p_m(k_i)(k_m - 1 - N_m/2)dp_m, \quad k_m = 1, \ldots, N_m, \tag{D.71}$$

$$dp_m = \frac{2\pi\varepsilon}{N_m dx_m}. \tag{D.72}$$

So that $\boldsymbol{p}(\boldsymbol{k}) = (p_1(k_1), \ldots, p_d(k_d))$, $\boldsymbol{x}(\boldsymbol{j}) = (x_1(k_1), \ldots, x_d(k_d))$. We approximate the integral in eq. (D.68) by

$$\begin{aligned} \widehat{f^\varepsilon}(\boldsymbol{p}(\boldsymbol{k})) &= \int_{\mathbb{R}^n} f(\boldsymbol{x}) \exp\left(-\frac{i}{\varepsilon}\boldsymbol{p}(\boldsymbol{k}) \cdot \boldsymbol{x}\right) d\boldsymbol{x}, \\ &\approx dx_1 \ldots dx_n \sum_{j_1=1}^{N_1} \cdots \sum_{j_d=1}^{N_d} f(\boldsymbol{x}(\boldsymbol{j})) \exp\left(-\frac{i}{\varepsilon}\boldsymbol{p}(\boldsymbol{k}) \cdot \boldsymbol{x}(\boldsymbol{j})\right). \end{aligned}$$

Now consider the exponent

$$-\frac{i}{\varepsilon} p_m(k_m) x_m(j_m). \tag{D.73}$$

For ease of notation, we suppress the dimension. Inserting identities

$$\begin{aligned} -\frac{i}{\varepsilon} p(k)x(j) &= -\frac{i}{\varepsilon}(k - 1 - N/2)dp(x(1) + (j-1)dx), \\ &= -\frac{2\pi i}{N}(k-1)(j-1) + \frac{i\pi}{dx}(x(j) - x(1)) - \frac{i}{\varepsilon}x(1)p(k). \end{aligned}$$

The first term is the exponent in the `fftn`, and the third term is independent of j, so the exponent can be taken out of the summand. In conclusion we find

$$\widehat{f^\varepsilon}(\boldsymbol{p}(\boldsymbol{k})) \approx \frac{dx_1 \ldots dx_n}{(2\pi\varepsilon)^{d/2}} \exp\left(-\frac{i}{\varepsilon}\boldsymbol{x}(\boldsymbol{1}) \cdot \boldsymbol{p}(\boldsymbol{k})\right) \mathcal{F}\left[\exp\left(i\pi \sum_{m=1}^{d} \frac{x_m(\cdot) - x_m(1)}{dx_m}\right) f(\boldsymbol{x}(\cdot))\right](\boldsymbol{k}), \tag{D.74}$$

where $\boldsymbol{1} = (1, \ldots, 1)$. Similarly, we can used the inverse fast Fourier transform `ifftn` to apply the inverse ε-Fourier transform. The inverse Fourier transform for a function F defined at $\boldsymbol{N}$ momentum points with values $F(\boldsymbol{k})$, is defined by a function f at $\boldsymbol{N}$ spatial points with values $f(\boldsymbol{j})$ given by

$$f(\boldsymbol{j}) = \tilde{\mathcal{F}}(F)(\boldsymbol{j}) = \frac{1}{N_1 N_2 \ldots N_d} \sum_{k_1=1}^{N_1} \cdots \sum_{k_d=1}^{N_d} F(k_1 \ldots, k_d) \exp\left(\frac{2\pi i}{N_1}(j_1 - 1)(k_1 - 1)\right) \ldots$$

$$\cdots \times \exp\left(-\frac{2\pi i}{N_d}(j_d-1)(k_d-1)\right). \tag{D.75}$$

Then by a similar calculation, the discretization of the inverse ε-Fourier transform can be written using the inverse Fourier transform

$$f(\boldsymbol{x}(\boldsymbol{j})) = \frac{N_1 \ldots N_d dp_1 \ldots dp_d}{(2\pi\varepsilon)^{\frac{d}{2}}} \exp\left(\sum_{m=1}^{d} \frac{x_m(j_m) - x_m(1)}{dx_m}\right) \tilde{\mathcal{F}}\left[\exp\left(\frac{i}{\varepsilon}\boldsymbol{x}(\mathbf{1})\cdot\boldsymbol{p}(\cdot)\right) F(\cdot)\right](\boldsymbol{j}). \tag{D.76}$$

D.8.2 The Strang splitting method

We use the Strang splitting method to compute the exact dynamics of non-adiabatic transitions. The Schrödinger equation in the diabatic representation is given by:

$$\begin{aligned} i\varepsilon\partial_t\psi(\boldsymbol{x},t) &= H\psi(\boldsymbol{x},t), \\ &= \left\{-\frac{\varepsilon^2}{2}\mathbf{1}\nabla^2 + V(\boldsymbol{x})\right\}\psi(\boldsymbol{x},t). \end{aligned} \tag{D.77}$$

The solution for eq. (D.77) can be formally represented as

$$\psi(\boldsymbol{x},t) = \exp\left(-\frac{i}{\varepsilon}\left(-\frac{\varepsilon^2}{2}\mathbf{1}\nabla^2 + V(\boldsymbol{x})\right)t\right)\psi_0(\boldsymbol{x}), \tag{D.78}$$

for some initial condition $\psi_0(\boldsymbol{x})$. The position and momentum operators in the exponant satisfy the Campbell-Baker-Hausdorff theorem [42]: for operators A, B, C:

$$\begin{aligned} \exp(A)\exp(B)\exp(C) = \exp\Bigg(A + B + C + \frac{1}{2}[A,B] + \frac{1}{2}[A,C] + \frac{1}{2}[B,C] \\ + \frac{1}{2}[[A,B],A+B+C] + \ldots\Bigg). \end{aligned} \tag{D.79}$$

If we take $A = C = \frac{i\varepsilon\Delta t}{4}\nabla^2$, $B = -\frac{i\Delta t}{\varepsilon}V(\boldsymbol{x})$, then by eq. (D.79),

$$\exp\left(\frac{i\varepsilon\Delta t}{4}I\nabla^2\right)\exp\left(-\frac{i\Delta t}{\varepsilon}V(x)\right)\exp\left(i\frac{\varepsilon}{4}I\nabla^2\Delta t\right) = \exp\left(-\frac{i}{\varepsilon}\Delta t H + \mathcal{O}(\Delta t^3)\right). \tag{D.80}$$

This approximation is second order. We use eq. (D.80) via a *Strang splitting method* for exact one level and two level calculations. We start with initial condition $\psi(\boldsymbol{x},0)$ in the diabatic representation, and perform the dynamics on the interval $[0,t]$, which is discretized into n steps of length $\Delta t = t/n$. We then:

1. Perform the ε-Fourier transform on the initial condition, $\psi(x,0) \to \widehat{\psi}^{\varepsilon}(k,0)$,
2. Apply the kinetic energy operator with a half time step, in momentum space

$$\widehat{\psi}^{\varepsilon}(k,0) \to \exp\left(\frac{i\varepsilon\Delta t}{4}I\|\boldsymbol{k}\|^2\right)\widehat{\psi}^{\varepsilon}(\boldsymbol{k},0) := \widehat{\psi}_1^{\varepsilon}(\boldsymbol{k}),$$

3. Perform the inverse scaled Fourier transform $\widehat{\psi}_1^{\varepsilon}(\boldsymbol{k}) \to \psi_1(\boldsymbol{x})$,

4. Apply the potential operator with a full time step $\psi_1(\boldsymbol{x}) \to \exp\left(-\frac{i\Delta t}{\varepsilon}V(\boldsymbol{x})\right)\psi_1(\boldsymbol{x}) := \psi_2(\boldsymbol{x})$,
5. Perform the scaled Fourier transform, $\psi_2(\boldsymbol{x}) \to \widehat{\psi}_2^{\varepsilon}(\boldsymbol{k})$,
6. Apply the kinetic energy operator with a *full* time step, in momentum space $\widehat{\psi}_2^{\varepsilon}(x, \Delta t) \to \exp\left(-i\frac{\varepsilon}{2}Ik^2\Delta t\right)\widehat{\psi}_2^{\varepsilon}(x, \Delta t) := \widehat{\psi}_1^{\varepsilon}(x, \Delta t)$.
7. Perform the inverse scaled Fourier transform $\widehat{\psi}_1^{\varepsilon}(k) \to \psi_1(x)$,
8. Repeat steps (4-7) $n-1$ times,
9. Repeat steps (1-3), to return $\psi(x, n\Delta t) = \psi(x, t)$.

By simply applying a full kinetic timestep instead of half twice, we reduce the number of computations needed.

D.9 Diagrams for the 2D algorithm

figs. D.10a to D.10e are to help explain the algorithm discussed in appendix D.5, for approximating the transmitted wavepacket using the derived formula eq. (D.20), using example 28 with wavepacket eq. (D.59) and parameters eq. (D.66).

References

[1] M. Born and R. Oppenheimer, "Zur Quantentheorie der Molekeln," *Annalen der Physik (Berlin, Germany)*, vol. 84, pp. 457–484, 1927.

[2] W. Domcke, D. Yarkony, and K. Horst, *Conical intersections: electronic structure, dynamics and spectroscopy*. World Scientific, 2004, vol. 15.

[3] W. Domcke, D. Yarkony, and K. Köppel, *Conical intersections: theory, computation and experiment*. World Scientific, 2011, vol. 17.

[4] H. Nakamura, *Nonadiabatic transition: concepts, basic theories and applications*. World Scientific, 2012.

[5] J. Tully, "Perspective: Nonadiabatic dynamics theory," *Journal of Chemical Physics*, vol. 137, no. 22, 22A301, 2012.

[6] J. Tully and R. Preston, "Trajectory surface hopping approach to nonadiabatic molecular collisions: The reaction of h+ with d2," *Journal of Chemical Physics*, vol. 55, no. 2, pp. 562–572, 1971.

[7] W. Miller and T. George, "Semiclassical theory of electronic transitions in low energy atomic and molecular collisions involving several nuclear degrees of freedom," *Journal of Chemical Physics*, vol. 56, no. 11, pp. 5637–5652, 1972.

[8] J. Stine and J. Muckerman, "On the multidimensional surface intersection problem and classical trajectory 'surface hopping'," *Journal of Chemical Physics*, vol. 65, no. 10, pp. 3975–3984, 1976.

[9] P. Kuntz, J. Kendrick, and W. Whitton, "Surface-hopping trajectory calculations of collision-induced dissociation processes with and without charge transfer," *Chemical Physics*, vol. 38, no. 2, pp. 147–160, 1979.

[10] J. Tully, "Molecular dynamics with electronic transitions," *Journal of Chemical Physics*, vol. 93, no. 2, pp. 1061–1071, 1990.

[11] S. Hammes-Schiffer and J. Tully, "Proton transfer in solution: Molecular dynamics with quantum transitions," *Journal of Chemical Physics*, vol. 101, no. 6, pp. 4657–4667, 1994.

[12] U. Müller and G. Stock, "Surface-hopping modeling of photoinduced relaxation dynamics on coupled potential-energy surfaces," *Journal of Chemical Physics*, vol. 107, no. 16, pp. 6230–6245, 1997.

[13] V. Rousse, "Landau–Zener transitions for eigenvalue avoided crossings in the adiabatic and Born–Oppenheimer approximations," *Asymptotic Analysis*, vol. 37, no. 3, 4, pp. 293–328, 2004.

[14] E. Fabiano, G. Groenhof, and W. Thiel, "Approximate switching algorithms for trajectory surface hopping," *Chemical Physics*, vol. 351, no. 1, pp. 111–116, 2008.

[15] C. Fermanian Kammerer and C. Lasser, "Single switch surface hopping for molecular dynamics with transitions," *Journal of Chemical Physics*, vol. 128, no. 14, p. 144 102, 2008.

[16] C. Lasser and T. Swart, "Single switch surface hopping for a model of pyrazine," *Journal of Chemical Physics*, vol. 129, no. 3, p. 034 302, 2008.

[17] A. Belyaev, C. Lasser, and G. Trigila, "Landau–zener type surface hopping algorithms," *Journal of Chemical Physics*, vol. 140, no. 22, p. 224 108, 2014.

[18] A. Belyaev, W. Domcke, C. Lasser, and G. Trigila, "Nonadiabatic nuclear dynamics of the ammonia cation studied by surface hopping classical trajectory calculations," *Journal of Chemical Physics*, vol. 142, no. 10, p. 104 307, 2015.

[19] D. Zener, "Non-adiabatic crossings of energy levels," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 137, pp. 696–702, 1932.

[20] L. Landau, *Collected Papers of L.D. Landau*. Oxford: Pergamon Press, 1965.

[21] V. N. Gorshkov, S. Tretiak, and D. Mozyrsky, "Semiclassical Monte-Carlo approach for modelling non-adiabatic dynamics in extended molecules," *Nature communications*, vol. 4, p. 2144, 2013.

[22] J. Lu and Z. Zhou, "Frozen gaussian approximation with surface hopping for mixed quantum-classical dynamics: A mathematical justification of fewest switches surface hopping algorithms," *Mathematics of Computation*, vol. 87, no. 313, pp. 2189–2232, 2018.

[23] L. Chai, S. Jin, Q. Li, and O. Morandi, "A multiband semiclassical model for surface hopping quantum dynamics," *Multiscale Modellling and Simulation*, vol. 13, no. 1, pp. 205–230, 2015.

[24] R. Kapral and G. Ciccotti, "Mixed quantum-classical dynamics," *Journal of Chemical Physics*, vol. 110, no. 18, pp. 8919–8929, 1999.

[25] I. Horenko, C. Salzmann, B. Schmidt, and C. Schütte, "Quantum-classical liouville approach to molecular dynamics: Surface hopping gaussian phase-space packets," *Journal of Chemical Physics*, vol. 117, no. 24, pp. 11 075–11 088, 2002.

[26] R. Kapral, "Surface hopping from the perspective of quantum–classical liouville dynamics," *Chemical Physics*, vol. 481, pp. 77–83, 2016.

[27] V. Betz, B. Goddard, and S. Teufel, "Superadiabatic transitions in quantum molecular dynamics," *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, vol. 465, no. 2111, pp. 3553–3580, Nov. 2009.

[28] V. Betz and B. Goddard, "Accurate prediction of nonadiabatic transitions through avoided crossings," *Physical Review Letters*, vol. 103, p. 213 001, 2009.

[29] V. Betz, B. Goddard, and U. Manthe, "Wave packet dynamics in the optimal superadiabatic approximation," *Journal of Chemical Physics*, vol. 144, no. 22, p. 224 109, 2016.

[30] A. Bach, *An Introduction to Semiclassical and Microlocal Analysis*, 1st ed., ser. Universitext. Springer, 2002, ISBN: 978-1-4757-4495-8.

[31] M. Berry, "Histories of adiabatic quantum transitions," *Proceedings of the Royal Society of London, Series A: Mathematical, Physical and Engineering Sciences*, vol. 429, no. 1876, pp. 61–72, 1990.

[32] G. A. Hagedorn and A. Joye, "Molecular propagation through small avoided crossings of electron energy levels," *Reviews in Mathematical Physics*, vol. 11, no. 01, pp. 41–101, 1999.

[33] V. Betz and B. Goddard, "Nonadiabatic transitions through tilted avoided crossings," *SIAM Journal of Scientific Computing*, vol. 33, no. 5, pp. 2247–2276, 2011.

[34] S. Teufel, *Adiabatic perturbation theory in quantum dynamics*. Springer, 2003.

[35] T. Katsaounis and I. Kyza, "A posteriori error control and adaptivity for crank–nicolson finite element approximations for the linear schrödinger equation," *Numerische Mathematik*, vol. 129, no. 1, pp. 55–90, 2015.

[36] V. Betz and S. Teufel, "Precise coupling terms in adiabatic quantum evolution," *Annales Henri Poincaré*, vol. 6, no. 2, pp. 217–246, 2005.

[37] J. Avron and I. Herbst, "Spectral and scattering theory of Schrödinger operators related to the Stark effect," *Communications in Mathematical Physics*, vol. 52, no. 3, pp. 239–254, 1977.

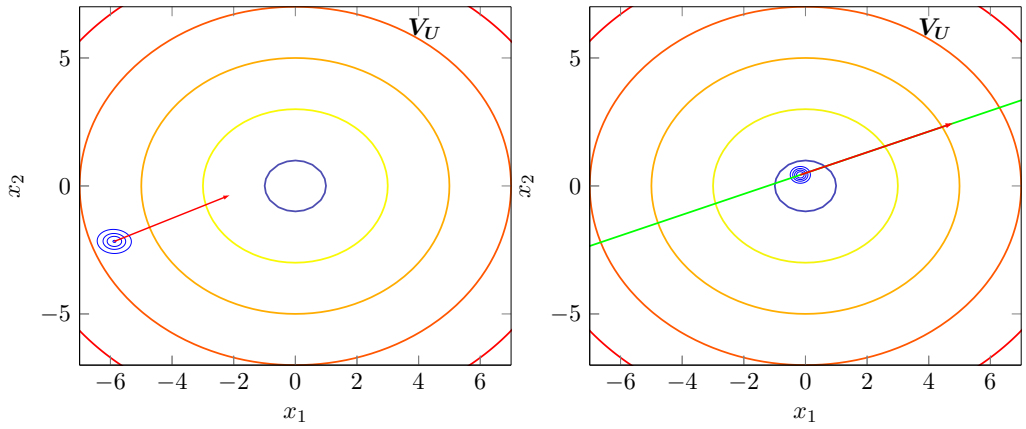
[38] M. Berry and R. Lim, "Universal transition prefactors derived by superadiabatic renormalization," *Journal of Physics A: Mathematical and General*, vol. 26, no. 18, pp. 4737–4747, Jan. 1993.

[39] A. Joye, G. Mileti, and C. Pfister, "Interferences in adiabatic transition probabilities mediated by stokes lines," *Physical Review A*, vol. 44, no. 7, p. 4280, 1991.

[40] A. Bohm, A. Mostafazedeh, H. Koizumi, Q. Niu, and J. Zwanziger, *The geometric phase in quantum systems*. Texts and Monographs in Physics, Springer, Heidelberg, 2003.

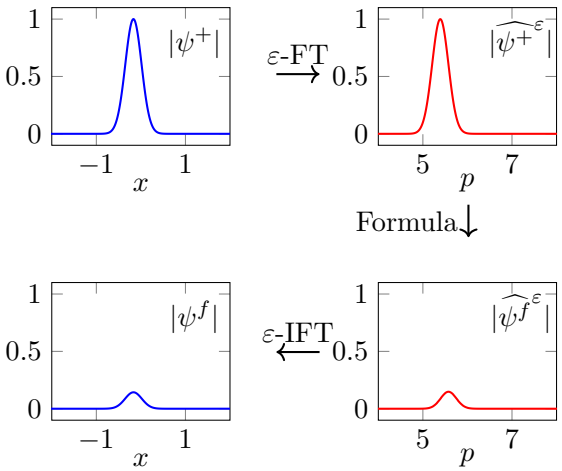
[41] C. Fermanian Kammerer and C. Lasser, "An egorov theorem for avoided crossings of eigenvalue surfaces," *Communications in Mathematical Physics*, pp. 1–47, 2017.

[42] B. Leimkuhler and C. Matthews, *Molecular Dynamics*. Springer, 2016.

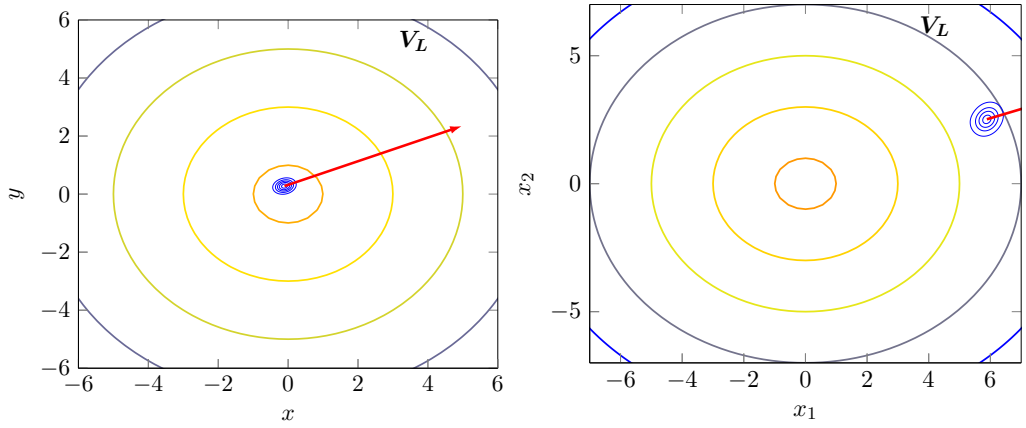


(a) ψ^+ far from avoided crossing.

(b) ψ^+ at the avoided crossing.



(c) Application of the formula to the wavepacket on a 1D strip.



(d) ψ^- approximated by the formula at the crossing.

(e) ψ^- approximated by the formula, far from the crossing.

Figure D.10: Application of the algorithm, with the system described in example 28. Large coloured ovals are contours of the potential surface being considered, dense blue contours are the wavepackets. The strip which is considered in fig. D.10c is that along the green line in fig. D.10b. in fig. D.10c, blue lines are the wavepacket in position space, and red lines are in momentum space. In each contour plot, the direction and magnitude of the wavepacket's momentum is represented by a red arrow.

Appendix E

Algorithms for semi-supervised learning on point clouds

Given a data set and a subset of labels the problem of semi-supervised learning on point clouds is to extend the labels to the entire data set. In this paper we extend the labels by minimising the constrained discrete p-Dirichlet energy. Under suitable conditions the discrete problem can be connected, in the large data limit, with the minimiser of a weighted continuum p-Dirichlet energy with the same constraints. We take advantage of this connection by designing numerical schemes that first estimate the density of the data and then apply PDE methods, such as pseudo-spectral methods, to solve the corresponding Euler-Lagrange equation. We prove that our scheme is consistent in the large data limit for two methods of density estimation: kernel density estimation and spline kernel density estimation.

E.1 Introduction

In many machine learning problems, such as classification or labelling, one often aims to exploit the usually large quantities of data in order to capture its geometry. Frequently in applications labels for some of the data are available but often in low quantities because of the cost of labelling points. In the semi-supervised learning setting we are given a data set $\boldsymbol{x}_i \in \mathbb{R}^d$, $i = 1, \dots, n$ sampled from an unknown probability measure μ, and a small subset of labelled pairs $(\boldsymbol{x}_i, y_i)$, $i = 1, \dots, N$ where $y_i \in \mathbb{R}$ and we work in the regime $N \ll n$. Here, the labels y_i for the first N data points are known and we aim to estimate the labels $\{y_i\}_{i=N+1}^n$ for the remaining data points $\{\boldsymbol{x}_i\}_{i=N+1}^n$.

One method to assign these labels is to minimise an objective function, which penalises smoothness of assigned labels, under the constraint that known labels are preserved. A common choice of such an objective function is the graph p-Dirichlet energy [1], [2], which allows one to define a discrete version of a Dirichlet energy. More precisely, we consider the graph $(\Omega_n, \boldsymbol{W})$ of nodes $\Omega_n = \{\boldsymbol{x}_i\}_{i=1}^n$ and edge weights $\boldsymbol{W} = (W_{ij})_{i,j=1}^n$ where W_{ij} is the edge weight between data points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ (by convention we say there is no edge between $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ if $W_{ij} = 0$). We use the random geometric graph model with length scale ε_n for defining the edge weights. Approximately, the parameter ε_n determines the range at which two nodes become connected; the explicit construction is given in the following section. The objective functional is defined as the difference

between labels weighted by edge weights:

$$\mathcal{E}^{(p)}_{n,\mathrm{con}}(f) = \begin{cases} \frac{1}{\varepsilon_n^p n^2} \sum_{i,j=1}^n W_{ij} |f(\boldsymbol{x}_i) - f(\boldsymbol{x}_j)|^p, & \text{if } f(\boldsymbol{x}_i) = y_i \text{ for all } i = 1, \dots, N, \\ +\infty & \text{else.} \end{cases} \tag{E.1}$$

Computing the minimiser of the graph p-Dirichlet energy becomes computationally expensive when considering a large number of data points. However, it has been shown [3] that under an admissible scaling regime in ε_n (and when $p > d$), minimisers of (E.1) converge to minimisers of the continuum p-Dirichlet energy:

$$\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho) = \begin{cases} \sigma_\eta \int_\Omega |\nabla f(\boldsymbol{x})|^p \rho^2(\boldsymbol{x}) \, \mathrm{d}\boldsymbol{x} & \text{if } f \in W^{1,p}(\Omega) \text{ and } f(\boldsymbol{x}_i) = y_i \text{ for } i = 1, \dots, N, \\ +\infty & \text{else} \end{cases} \tag{E.2}$$

where ρ is the density of the data points, and σ_η is some constant depending only on a weight function η (satisfying Definition 29). This result shows that the minimiser of the continuum p-Dirichlet energy is an accurate estimate of the minimiser of the graph p-Dirichlet energy when considering a large amount of data. However, the dependency of the continuum p-Dirichlet energy on the underlying data density poses a new problem as it is unrealistic to assume that we know the density of the data. The objective of this paper is to develop the framework for a new numerical method for finding minimisers of $\mathcal{E}^{(p)}_n$ based on its connection to the continuum variational problem of minimising $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}$. In particular, it is our aim to develop a numerical scheme that is efficient for large ($n \gg 1$) datasets. We refer to the work by Flores Rios, Calder and Lerman [4] for algorithms based on the discrete problem.

We note that an associated *non-local* continuum p-Dirichlet energy is also of interest;

$$\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f;\rho) = \begin{cases} \dfrac{1}{\varepsilon^p} \displaystyle\int_\Omega \int_\Omega \eta_{\varepsilon_n}(|\boldsymbol{x} - \boldsymbol{z}|) |f(\boldsymbol{x}) - f(\boldsymbol{z})|^p & \text{if } f \in W^{1,p}(\Omega) \text{ and } f(\boldsymbol{x}_i) = y_i, i = 1, \dots, N, \\ \qquad \times \rho(\boldsymbol{x}) \rho(\boldsymbol{z}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{z}. & \\ +\infty & \text{else} \end{cases}$$

for example, $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f;\rho)$ was considered in [3] as an intermediary functional to provide convergence between $\mathcal{E}^{(p)}_{n,\mathrm{con}}(f)$ and $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}$. We also develop a numerical method for computing minimisers of the above functional that are efficient for large n.

In our approach we, rather than minimise (E.1), aim instead to minimise (E.2). Since the density ρ is unknown we are required to estimate it from the data Ω_n. Density estimation is a well-studied problem in statistics dating back, at least, to Fix and Hodges in 1951 [5] and Akaike in 1954 [6]. The first method we consider here to estimate the density is *kernel density estimation* [7], [8]. The idea behind kernel density estimation is to replace the empirical measure $\mu_n = \sum_{i=1}^n \delta_{x_i}$, where $x_i \overset{\text{iid}}{\sim} \mu$, with a smooth approximation; in particular, $\mu_n(A) = \frac{1}{n} \sum_{i=1}^n \delta_{\boldsymbol{x}_i}(A) \approx \int_A \frac{1}{n} \sum_{i=1}^n K_h(\boldsymbol{x} - \boldsymbol{x}_i) \, \mathrm{d}\boldsymbol{x}$. Formally, for n large enough $\mu_n \approx \mu$ and then one expects $\frac{1}{n} \sum_{i=1}^n K_h(\cdot - \boldsymbol{x}_i)$ to approximate the density ρ of μ. Analysing this approximation has been the interest of many statisticians, see for example [9]–[15]. Our methods are an adaptation of the results by Giné and Guillon [11]. We refer to [16] for an overview on kernel density estimation.

The second type of density estimation we consider is a regularised version of the kernel density estimate. Our method is to use the kernel density estimate to estimate the value of the density at knot points (that we are free to choose), we then use smoothing splines to produce an estimate of the density. We call this method the *spline kernel density estimate*. Although we are unaware of previous work using splines to estimate the density of distributions the idea of introducing regularisation in density estimation is not new, see for example [17].

The advantage of using splines is that they introduce additional smoothness into our estimate of the density which allows for better approximations of the density and, due to the volume of results in the literature, are theoretically well understood. Furthermore, we can take advantage of fast computational methods for solving the spline smoothing problem. We refer to [18] for an overview and mention a few select references here. Convergence in norm of special splines under various settings have been studied in [19]–[27], and general splines in [28]–[33]. Similarly weak (pointwise) convergence of special splines has been studied in [34]–[38] and general splines in [39].

Due to its interest in machine learning the convergence of variational problem (E.1) to (E.2) has attracted much interest. For example, pointwise convergence results have been used to motivate the choice of $p > d$ [40]–[49]; however, pointwise convergence is not enough, in general, to imply variational convergence (convergence of minimisers). Spectral convergence [43], [50]–[54] (and error bounds [55], [56]) shows convergence of minimisers only when $p = 2$. The framework to analyse the discrete-to-variational was developed by García-Trillos and Slepčev [57] and later applied to the constrained problem to show variational convergence when $p > d$ and ε_n satisfies an upper bound [3]. Using PDE methods Calder [4] studies the large data limits of two closely related problems. The first is Lipschitz learning (which corresponds to choosing $p = \infty$) [58], and the second is the game theoretic p-Dirichlet energy [59].

In this paper we show that minimisers of the continuum p-Dirichlet energy, in which the density is estimated from the data, converge to a minimiser of the continuum p-Dirichlet energy in the large data limit. After setting up notation and listing the main results in Section E.2, we do this in two parts. The first part, in Section E.3, gives sufficient conditions for the convergence of $\rho_n \to \rho$ to imply the convergence of minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ to minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$, and for minimisers of $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho_n)$ to converge to minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$. This complements the results of [3] which prove convergence of minimisers of $\mathcal{E}^{(p)}_{n,\mathrm{con}}$ to minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$ via the intermediary functional $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho)$ (see Figure E.1 for a summary). Then, in the second part, we provide two examples of density estimation schemes, the kernel density estimate and the spline kernel density estimate, that satisfy the conditions in the previous section (see Section E.4). Numerical illustration of the results in two dimensions are provided in Section E.5 and we conclude in Section E.6.

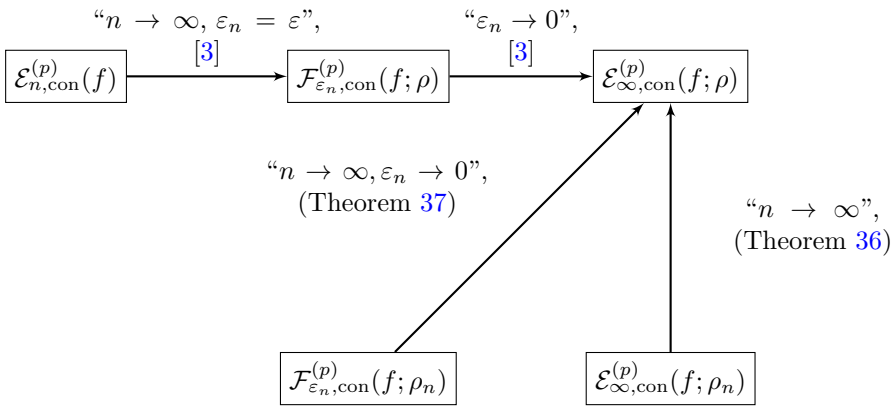


Figure E.1: Diagram depicting how various p-Dirichlet energies discussed in this paper are related, and where their convergence results can be found.

E.2 Setting and main results

E.2.1 Notation

We consider functions on an open, bounded and connected domain $\Omega \subseteq \mathbb{R}^d$ with Lipschitz boundary. Given a positive Radon measure $\mu \in \mathcal{M}_+(\Omega)$ (where usually μ is a probability measure) we let $L^p(\mu)$ denote the space of functions for which the p^{th} power of the absolute value is integrable with respect to μ and the usual norm $\|f\|_{L^p(\mu)}$. When $\mu = \mathcal{L}\lfloor_\Omega$, the Lebesgue measure on Ω we write, with a small abuse of notation, $L^p(\Omega)$ instead of $L^p(\mathcal{L}\lfloor_\Omega)$. Sobolev spaces, denoted by $W^{m,p}(\Omega)$, are the space of functions where the p^{th} power of the absolute value of the first m (weak) derivatives are integrable with respect to the Lebesgue measure. When $p = 2$ we also write $W^{m,2} = H^m$. The norm $\|f\|_{W^{m,p}(\Omega)}$ on $W^{m,p}$ is defined in the usual way. We use $C^{0,\alpha}$ to define the Hölder space with norm $\|f\|_{C^{0,\alpha}(\Omega)}$.

Throughout we assume we have a data set $\Omega_n = \{\boldsymbol{x}_i\}_{i=1}^n$ where $\boldsymbol{x}_i \overset{\text{iid}}{\sim} \mu$ and $\mu \in \mathcal{P}(\Omega)$. Given such a data set we define the *empirical measure* by

$$\mu_n(\boldsymbol{x}) = \frac{1}{n}\sum_{i=1}^n \delta(\boldsymbol{x} - \boldsymbol{x}_i),$$

where $\delta(\boldsymbol{x})$ is the Dirac function.

We write $\Omega' \subset\subset \Omega$ to mean that Ω' is a compact subset of Ω and d_{H} is the Hausdorff distance between sets in $\mathbb{R}^d$.

E.2.2 Dirichlet energies: setup

The following definition is used to construct the weights of the graph given data $\Omega_n = \{\boldsymbol{x}_i\}_{i=1}^n$.

Definition 29. A function $\eta : [0, +\infty) \to [0, +\infty)$ is a **weight function** if it is a decreasing function with $\lim_{r\to+\infty} \eta(r) = 0$, and is positive and continuous at $r = 0$.

We prescribe weights between the points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ using a weight function η: for a fixed $\varepsilon > 0$, the weight between two points $\boldsymbol{x}_i$ and $\boldsymbol{x}_j$ is defined by

$$W_{ij} = \eta_\varepsilon(|\boldsymbol{x}_i - \boldsymbol{x}_j|), \tag{E.3}$$

where $\eta_\varepsilon(\cdot) = \frac{1}{\varepsilon^d}\eta\left(\frac{\cdot}{\varepsilon}\right)$. This can be used to define a maximum distance ε for which particles have a non-zero weight; for example $\eta(t) = 1$ for $t < 1$ and $\eta(t) = 0$ otherwise. Then then W_{ij} is positive only when $|\boldsymbol{x}_i - \boldsymbol{x}_j| < \varepsilon$. In the remainder of this section we define the various Dirichlet energies which are considered in the sequel.

Definition 30. Let $p \in (1, +\infty)$, $\Omega_n = \{\boldsymbol{x}_i\}_{i=1}^n \subset \mathbb{R}^d$ and define μ_n to be the empirical measure and W_{ij} as in (E.3) where η is a weight function. We define the **discrete p-Dirichlet energy** by

$$\begin{aligned}\mathcal{E}_n^{(p)} &: L^p(\mu_n) \to [0, +\infty)\\ \mathcal{E}_n^{(p)}(f) &= \frac{1}{\varepsilon^p n^2}\sum_{i,j=1}^n W_{ij}|f(\boldsymbol{x}_i) - f(\boldsymbol{x}_j)|^p.\end{aligned}$$

When the data generating distribution μ has density ρ (with respect to the Lebesgue measure) then the large data limit, in the sense of Γ-convergence, is given by the continuum p-Dirichlet energy $\mathcal{E}_\infty^{(p)}$ defined below. Indeed, when $p = 1$ it was shown in [57] that $\Gamma\text{-}\lim_{n\to+\infty} \mathcal{E}_n^{(1)} = \mathcal{E}_\infty^{(1)}(\cdot;\rho)$ where $\mathcal{E}_\infty^{(1)}(\cdot;\rho)$ is a weighted total variation (and takes a slightly different form to the class of energies given below). The proof, however, generalises to any $p \in [1, +\infty)$.

Definition 31. Let $p \in (1, +\infty)$, $\Omega \subset \mathbb{R}^d$ be an open, bounded and connected domain with Lipschitz boundary, and $\rho \in L^\infty(\Omega)$ be a non-negative function. Let η be a weight function and assume that

$$\sigma_\eta := \int_{\mathbb{R}^d} \eta(|\boldsymbol{x}|)|\boldsymbol{x}\cdot\boldsymbol{e}_1|^p \,\mathrm{d}\boldsymbol{x} < +\infty, \tag{E.4}$$

where $\boldsymbol{e}_1 = (1, 0, \ldots, 0)$. We define the **continuum p-Dirichlet energy with respect to ρ** by

$$\begin{aligned}&\mathcal{E}_\infty^{(p)}(\cdot;\rho) : L^p(\Omega) \to [0, +\infty],\\ &\mathcal{E}_\infty^{(p)}(f;\rho) = \begin{cases}\sigma_\eta \int_\Omega |\nabla f(\boldsymbol{x})|^p \rho^2(\boldsymbol{x}) \,\mathrm{d}\boldsymbol{x} & \text{if } f \in W^{1,p}(\Omega),\\ +\infty & \text{else}\end{cases}\end{aligned} \tag{E.5}$$

The parameter p controls the amount of regularity. The results of this paper concern a finite choice of p since for $p = +\infty$ the Dirichlet energy loses sensitivity to the density of the data. In fact, it is easy to show that the variational limit as $p \to +\infty$ (with n fixed and after renormalising with respect to p) is the Lipschitz learning problem:

$$\mathcal{E}_n^{(\infty)}(f) = \max_{i,j\in\{1,\ldots,n\}} W_{ij}|f(\boldsymbol{x}_i) - f(\boldsymbol{x}_j)|,$$

see for example [60] for the computation with a similar objective. The objective of this paper is to build numerical methods by estimating the density of data. For Lipschitz learning the data distribution appears in the continuum limit only through it's support; in particular, the intensity is irrelevant. More precisely, it is known (see [42] for pointwise limits and [58] for variational limits in the semi-supervised setting) that the large data limit of $\mathcal{E}_n^{(\infty)}$ is

$$\mathcal{E}_\infty^{(\infty)}(f) = \sup_{\boldsymbol{x}\in\Omega} |\nabla f(\boldsymbol{x})|.$$

Hence, one needs only to estimate the support of the data, not the density and so Lipschitz learning

falls outside the scope of our method.

We work in the semi-supervised setting; that is, we assume that we have labels y_i, $i = 1, \dots, N$ for the first N data points where N is fixed. To estimate labels at the remaining $n - N$ data points we use the Dirichlet energies to define a notion of regularity. More precisely, we minimise the Dirichlet energies subject to agreeing with the training data:

$$\text{minimise } \mathcal{E}_n^{(p)}(f) \qquad \text{subject to} \qquad f(\boldsymbol{x}_i) = y_i \quad \forall i = 1, \dots, N. \tag{E.6}$$

Analogously for the continuum Dirichlet energy. It will be convenient to define the constrained energies as follows.

Definition 32. Under the setting and notation of Definition 30, and given labels y_i for $i = 1, \dots, N$ we define the **constrained discrete p-Dirichlet energy** $\mathcal{E}_{n,\text{con}}^{(p)} : L^p(\mu_n) \to [0, +\infty]$ by eq. (E.1).

Definition 33. In addition to the setting and notation of Definition 31 assume that $p > d$. Then given labels y_i for $i = 1, \dots, N$ we define the **constrained continuum p-Dirichlet energy with respect to** ρ $\mathcal{E}_{\infty,\text{con}}^{(p)} : L^p(\Omega) \to [0, +\infty]$ by eq. (E.2)

Note that we require $p > d$ in order for the constrained continuum p-Dirichlet energy to be well defined. More precisely, for any f with $\mathcal{E}_\infty^{(p)}(f; \rho) < +\infty$ we necessarily have that $f \in W^{1,p}(\Omega)$ and hence by Sobolev embedding (Morrey's inequality) f can be identified with a continuous function, and therefore pointwise evaluation $f(\boldsymbol{x}_i)$ can be defined. For $p \leq d$ the constrained Dirichlet energy can no longer be defined in the continuum setting.

We also define the non-local continuum approximation $\mathcal{F}_{\varepsilon_n}^{(p)}$ of $\mathcal{E}_n^{(p)}$. This has been used as an intermediary functional in the discrete-to-continuum analysis of the p-Dirichlet energies, for example [3], [57] (as mentioned in Section E.1), but is also of interest in its own right.

Definition 34. Let $p \in (1, +\infty)$, $\Omega \subset \mathbb{R}^d$ be an open, bounded and connected domain with Lipschitz boundary, and $\rho \in L^\infty(\Omega)$ be a non-negative function. Let η be a weight function and then, we define the **non-local continuum p-Dirichlet energy with respect to** ρ by

$$\begin{aligned}
&\mathcal{F}_{\varepsilon_n}^{(p)}(\cdot, \rho) : L^p(\Omega) \to [0, +\infty), \\
&\mathcal{F}_{\varepsilon_n}^{(p)}(f; \rho) = \frac{1}{\varepsilon^p} \int_\Omega \int_\Omega \eta_{\varepsilon_n}(|\boldsymbol{x} - \boldsymbol{z}|) |f(\boldsymbol{x}) - f(\boldsymbol{z})|^p \rho(\boldsymbol{x}) \rho(\boldsymbol{z}) \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{z}.
\end{aligned}$$

We note that we are no longer able to impose pointwise constraints on $\mathcal{F}_{\varepsilon_n}^{(p)}$. Although $\mathcal{F}_{\varepsilon_n}^{(p)}$ is approximating a Sobolev semi-norm (when ε is small) we are still working on an L^p space with no continuity implied, and therefore one cannot impose pointwise constraints We overcome this by instead imposing the constraints on small balls around $\boldsymbol{x}_i$, $i = 1, \dots, N$. For our analysis we require that the balls have radius at least ε which leads us to define the constrained non-local continuum model as follows.

Definition 35. Under the setting and notation of Definition 34, and given labels y_i for $i = 1, \dots, N$ we define the **constrained non-local continuum p-Dirichlet energy with respect to** ρ by

$$\mathcal{F}_{\varepsilon,\text{con}}^{(p)} : L^p(\Omega) \to [0, +\infty],$$

$$\mathcal{F}^{(p)}_{\varepsilon,\mathrm{con}}(f;\rho) = \begin{cases} \mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho), & \text{if } f(\boldsymbol{x}) = y_i \text{ for } i = 1,\dots,N, \text{ and } \boldsymbol{x} \in B(\boldsymbol{x}_i,\varepsilon), \\ +\infty, & \text{else.} \end{cases}$$

In the above definition we make the assumption that $B(\boldsymbol{x}_i,\varepsilon) \cap B(\boldsymbol{x}_j,\varepsilon) = \emptyset$ for all $i,j = 1,\dots,N$. This is clearly satisfied for ε sufficiently small.

E.2.3 Large data asymptotics for Dirichlet energies

For the results given in this section we make the following assumptions:

(A1) $\Omega \subset \mathbb{R}^d$ is an open, connected, bounded domain with Lipschitz boundary, with associated probability measure $\mu \in \mathcal{P}(\Omega)$.

(A2) The probability measure μ has continuous density $\rho \in L^\infty(\Omega)$.

(A3) The density ρ is bounded above and below by strictly positive constants.

(A4) For $i = 1,\dots,N$, the points $\boldsymbol{x}_i \in \Omega$ are labelled with values $y_i \in \mathbb{R}$.

(A5) For $n \geq i > N$, the points $\boldsymbol{x}_i \in \Omega$ are i.i.d. samples of μ.

(A6) $\eta : [0,\infty) \to [0,\infty)$ is a weight function, and weights W_{ij} are defined by (E.3) for $i,j = 1,\dots,n$ and $\varepsilon = \varepsilon_n$.

(A7) The integral σ_η as defined in (E.4) is finite.

(A8) The smoothing parameter takes a value larger than the dimension of the data, $p > d$.

Under assumptions (A1) to (A8), and the following scaling on $\varepsilon_n = \varepsilon$,

$$\left(\frac{1}{n}\right)^{\frac{1}{p}} \gg \varepsilon_n \gg \begin{cases} \left(\frac{\log(n)}{n}\right)^{\frac{1}{d}} & \text{if } d \geq 3 \\ \frac{(\log(n))^{\frac{3}{4}}}{\sqrt{n}} & \text{if } d = 2 \end{cases} \tag{E.7}$$

minimisers of (E.1) converge to minimisers of (E.2) [3]. Furthermore, when $p < d$, minimisers of (E.1) converge to minimisers of (E.5) (*i.e.* constants), and so the constraints are lost as $n \to \infty$. This result allows us to approximate minimisers of (E.1) by its continuum analogue, (E.2). However, in general we may not know the density ρ. To make use of the continuum formulation for finite data, it is therefore necessary to estimate the density ρ using the information available; the data points $\boldsymbol{x}_i$, $i = 1,\dots,n$. This is the focus of the first main result of this paper.

For this result, we include the following assumption on the estimate of ρ:

(A9) The density estimate $\rho_n : \Omega \to \mathbb{R} \in L^\infty(\Omega)$, satisfies $\sup_{n\in\mathbb{N}} \|\rho_n\|_{L^\infty(\mathbb{R}^d)} < +\infty$ and $\rho_n \to \rho$ in $L^\infty_{\mathrm{loc}}(\Omega)$, *i.e.*, for all $\Omega' \subset\subset \Omega$

$$\sup_{\boldsymbol{x}\in\Omega'} |\rho_n(\boldsymbol{x}) - \rho(\boldsymbol{x})| \to 0.$$

Under the above assumptions we can prove the convergence of minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ to a minimiser of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$. A sequence f_n is a sequence of almost minimisers if there exists $\delta_n \to 0^+$

such that

$$\min_{f\in W^{1,p}} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho_n) \geq \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho) - \delta_n.$$

Theorem 36 (Convergence of minimisers of the local model)**.** Assume $\Omega, \mu, \eta, p, \rho, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, (i) minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ are precompact in $L^p_{\mathrm{loc}}(\Omega)$,

$$\text{(ii)} \min_{f\in W^{1,p}} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho) = \lim_{n\to\infty} \min_{f\in W^{1,p}} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n),$$

and (iii) any converging sequence of almost minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ converges in $L^\infty_{\mathrm{loc}}(\Omega)$ to a minimiser of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot,\rho)$.

Furthermore, we show an analogous non-local result.

Theorem 37 (Convergence of minimisers of the non-local model)**.** Assume $\Omega, \mu, \eta, p, \rho, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, (i) minimisers of $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho_n)$ are precompact in $L^p_{\mathrm{loc}}(\Omega)$,

$$\text{(ii)} \min_{f\in W^{1,p}} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho) = \lim_{n\to\infty} \inf_{f\in W^{1,p}} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho_n),$$

and (iii) any converging sequence of minimisers of $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho_n)$ converges in $L^\infty_{\mathrm{loc}}(\Omega)$ to a minimiser of $\mathcal{E}^{(p)}_{\infty}(\cdot;\rho_n)$.

The proof of both theorems is given in Section E.3. In Sections E.2.4 and E.2.5 we give two examples on how to construct density estimates that satisfy Assumption (A9) (with probability one).

E.2.4 Density estimates: set up

Both of the convergence results from the previous section rely on having a density estimate which converge locally uniformly with probability one. We consider two examples of density estimates with this convergence property: the kernel density estimate (KDE) and the closely related spline kernel density estimate (SKDE).

Kernel Density Estimate Recall the empirical measure $\mu_n = \frac{1}{n}\sum_{i=1}^n \delta(\cdot - \boldsymbol{x}_i)$. The kernel density estimate (KDE) can be viewed as a continuous approximation to the empirical measure, where each Dirac function is approximated by a function with particular properties, known as a *kernel function*; which is a function integrating to unity, *i.e.* $\int_{\mathbb{R}^d} K(\boldsymbol{x}) = 1$. A popular choice is the *Gaussian kernel function*

$$K(x) = \frac{1}{(2\pi)^{d/2}} \exp\left(-\frac{\|\boldsymbol{x}\|^2}{2}\right).$$

Other popular choices include the uniform and Epanechnikov kernels. In general kernel functions do not have to be symmetric or positive, we refer to [61], [62] for more examples of kernel functions. In our numerical experiments in Section E.5 we choose the Gaussian kernel.

We define the kernel density estimate as follows.

Definition 38. Given $\boldsymbol{x}_i \subset \Omega$ for $i = 1, \dots, n$ and a bandwidth $h > 0$, the *Kernel density estimate* $\rho_{n,h} : \Omega \to \mathbb{R}$ is defined by

$$\rho_{n,h}(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} K_h \left(\boldsymbol{x} - \boldsymbol{x}_i \right). \tag{E.8}$$

where $K_h(\boldsymbol{x}) = \frac{1}{h^d} K\left(\boldsymbol{x}/h\right)$ and $K : \mathbb{R}^d \to \mathbb{R}$ integrates to unity.

We note that, with additional notational complexity one can generalise the bandwidth to a positive semidefinite matrix $\boldsymbol{H}$, *i.e.*

$$\rho_{n,\boldsymbol{H}}(\boldsymbol{x}) = \frac{1}{n} \sum_{i=1}^{n} K_{\boldsymbol{H}} \left(\boldsymbol{x} - \boldsymbol{x}_i \right)$$

where $K_{\boldsymbol{H}}(\boldsymbol{x}) = \frac{1}{|\boldsymbol{H}|} K\left(H^{-1}\boldsymbol{x}\right)$. In the sequel we treat the special case where $\boldsymbol{H} = h\mathrm{Id}$.

When $h \to 0$ we regain the empirical measure μ_n, *i.e.* $\lim_{h \to 0^+} \int_A \mathrm{d}\rho_{n,h}(\boldsymbol{x}) = \mu_n(A)$ for all open sets A. We shall see that to guarantee convergence to the continuous density ρ, we will require a lower bound on the rate at which $h \to 0$. We state the convergence result in the next subsection.

Spline kernel density estimate To find the minimiser of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot; \rho_n)$ we use a gradient flow which involves the density estimate ρ_n and its derivative. It is therefore of interest to have a smooth approximation of ρ. Our strategy is to regularise the kernel density estimate.

One way to do this is to solve the variational problem:

$$\text{minimise } \|u - \rho_{n,h}\|^2_{L^2(\Omega)} + \lambda \|\nabla^m u\|^2_{L^2(\Omega)} \quad \text{over } u \in H^m(\Omega).$$

Drawing inspiration from the spline smoothing community we approximate the first term by

$$\|u - \rho_{n,h}\|^2_{L^2(\Omega)} \approx \frac{1}{T} \sum_{i=1}^{T} |u(\boldsymbol{t}_i) - \rho_{n,h}(\boldsymbol{t}_i)|^2$$

where $\{\boldsymbol{t}_i\}_{i=1}^T$ are called *knot points* (which we are free to choose). We therefore consider the variational problem:

$$S_{\lambda,T}(f) := \operatorname*{argmin}_{u \in H^m(\Omega)} \left\{ \frac{1}{T} \sum_{i=1}^{T} (u(\boldsymbol{t}_i) - f_i)^2 + \lambda \|\nabla^m u\|^2_{L^2(\Omega)} \right\}. \tag{E.9}$$

We define the projection operator

$$P_T : H^m(\Omega) \to \mathbb{R}^T, \qquad P_T(f) = f(\boldsymbol{t}_i) \tag{E.10}$$

which is well-defined whenever $m > d/2$.

Definition 39. Given a set of knot points $\{\boldsymbol{t}_i\}_{i=1}^T \subset \Omega$ and the kernel density estimate $\rho_{n,h}$ we define the *spline kernel density estimate* (SKDE) by

$$\rho_{n,h,\lambda,T} = S_{\lambda,T}(P_T(\rho_{n,h})) \tag{E.11}$$

where $S_{\lambda,T}$ is defined by (E.9) and P_T is defined by (E.10).

In the following subsection we give L^∞ convergence rates with probability 1 for the SKDE (in fact these results are a corollary of almost sure convergence results in H^m which when $m > d/2$ imply uniform convergence via Sobolev embeddings - we refer to Section E.4 for details).

We give a numerical comparison of the KDE and the SKDE in Section E.5. Moreover, we use the KDE and SKDE estimates to construct numerical methods to calculate the minimisers of (E.2) with $\rho_n = \rho_{n,h}$ and $\rho_n = \rho_{n,h,\lambda,T}$, and provide examples showing rates of convergence and computational efficiency.

E.2.5 Large data asymptotics for density estimation

We first discuss the almost sure locally uniform convergence of the kernel density estimate. This result follows almost immediately from known results in the literature. In particular, it is known from [11] that the L^∞ norm between the KDE and its expected value converges with a certain rate to 0. From here it is not difficult to show that the KDE converges locally uniformly to its true value.

The conditions we use to prove the convergence of the KDE estimate are the following.

(B1) $\Omega \subset \mathbb{R}^d$ is an open and bounded domain with Lipschitz boundary, with associated probability measure $\mu \in \mathcal{P}(\Omega)$.

(B2) The probability measure μ has a bounded density ρ.

(B3) For $i = 1, \dots,$ the points $\boldsymbol{x}_i \in \Omega$ are i.i.d. samples of μ.

(B4) $K : \mathbb{R}^d \to \mathbb{R}$ satisfies $\int_{\mathbb{R}^d} K(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} = 1$, has compact support in $B(0, M)$ for some $M > 0$ and can be written $K = \phi \circ \xi$ where ϕ is a bounded function of bounded variation and ξ is a polynomial.

(B5) $h = h_n$ satisfies

$$h_n \to 0^+, \quad \frac{nh_n^d}{|\log(h_n)|} \to \infty, \quad \frac{|\log(h_n)|}{\log\log(n)} \to \infty, \quad \text{and} \quad h_n \leq ch_{2n}$$

for some $c > 0$.

(B6) K satisfies the integrability condition $\int_{\mathbb{R}^d} |K(\boldsymbol{x})|\|\boldsymbol{x}\|\,\mathrm{d}\boldsymbol{x} < +\infty$.

(B7) ρ is continuous on Ω.

Assumptions (B1)-(B5) are the same as those used in [11] to prove the almost sure convergence of the bias:

$$\mathrm{Bias}(\rho_{n,h}) = \sup_{\boldsymbol{x}\in\Omega} |\mathbb{E}[\rho_{n,h}(x)] - \rho_{n,h}(x)|$$

to zero. We use Assumptions (B6)-(B7) to show that $\mathbb{E}[\rho_{n,h}] \to \rho$ locally uniformly with probability one.

In fact the assumption that $K = \phi \circ \xi$ where ϕ is bounded and of bounded variation and ξ is a polynomial can be relaxed. Our result is a simple application of [11, Theorem 2.3] which uses a

class of kernels that satisfy a technical condition that is sufficient to bound the Vapnik-Červonenkis (VC) dimension of functions of the form $K\left(\frac{\boldsymbol{x}-\cdot}{h}\right)$ for $\boldsymbol{x} \in \mathbb{R}^d$ and $h > 0$. As the authors remark the technical assumption is satisfied for functions of the form $K = \phi \circ \xi$ and since this includes the kernels we are interested in, *e.g.* Gaussian kernels, we satisfy ourselves with this less general case that can be stated more easily.

We state the converge result for the KDE here, the proof is given in Section E.4.1.

Theorem 40. Assume $\Omega, \mu, \rho, K, h_n$ and $\{\boldsymbol{x}\}_{i=1}^n$ satisfy Assumptions (B1)-(B7). Define $\rho_{n,h}$ as in Definition 38. Then, with probability one, for all $\Omega' \subset\subset \Omega$ we have

$$\lim_{n\to\infty} \|\rho_{n,h_n} - \rho\|_{L^\infty(\Omega')} \to 0.$$

We now turn our attention to the spline kernel density estimate. To prove convergence we will need some additional assumptions which we state now.

(B8) The number of derivatives penalised is greater than half the dimension of the data, $m > d/2$.

(B9) Let $d_{\mathrm{H}}(T) = d_{\mathrm{H}}(\{\boldsymbol{t}_i\}_{i=1}^T, \Omega)$, where d_{H} is the Hausdorff distance, and

$$\mathrm{Sep}(T) = \min\{|\boldsymbol{t}_i - \boldsymbol{t}_j| \,:\, i \neq j, i, j \in \{1, \ldots, T\}\}$$

then T_n and h_n satisfy

$$d_{\mathrm{H}}(T_n) \to 0 \qquad \text{and} \qquad d_{\mathrm{H}}(T_n) = O(\mathrm{Sep}(T_n))$$

(B10) $\lambda_n \to 0^+$ satisfies

$$\lambda_n T_n d_{\mathrm{H}}(T_n)^d = O(1), \quad T_n \lambda_n^{\frac{2m+d}{2m}} \gg n^\theta \quad \text{and} \quad T_n \lambda_n^{\frac{d}{2m}} \geq 1$$

for some $\theta > 0$.

(B11) Ω satisfies the uniform cone condition: *i.e.* there exists $r > 0$ and $\tau > 0$ such that for any $\boldsymbol{t} \in \Omega$, there exists a unit vector $\boldsymbol{\xi}(\boldsymbol{t}) \in \mathbb{R}^d$ such that the cone

$$C(\boldsymbol{t}, \boldsymbol{\xi}(t), \tau, r) = \{\boldsymbol{t} + \lambda\boldsymbol{\eta} \,:\, |\boldsymbol{\eta}| = 1, \boldsymbol{\eta} \cdot \boldsymbol{\xi}(\boldsymbol{t}) \geq \cos(\tau), 0 \leq \lambda \leq r\}$$

is fully contained in Ω.

(B12) T_n and h_n satisfy $\mathrm{Sep}(T_n) \geq 2Mh_n$ where M is given in Assumption (B4).

(B13) ρ is Lipschitz continuous on Ω and $\rho \in H^m(\Omega)$.

These conditions are necessary to apply the spline smoothing results of [63] and [27]; in particular Assumptions (B1), (B8)-(B11) are needed for the H^k convergence of splines in [27] and Assumptions (B12)-(B13) are used to match our setting here to their setting. We note that if we distribute the knot points uniformly over Ω then $d_{\mathrm{H}}(T_n) \sim \frac{1}{T_n^{\frac{1}{d}}}$ and $\mathrm{Sep}(T_n) \sim \frac{1}{T_n^{\frac{1}{d}}}$ which satisfy Assumption (B9). Assumption (B10) is satisfied for any $\lambda_n \to 0^+$ with $\lambda_n T_n^{\frac{2m}{d+2m}} \gg n^\theta$ for some

$\theta > 0$. The result we are interested in is stated below, the proof is a simple corollary of Theorem 56 in Section E.4.2 which gives convergence in $H^k_{\mathrm{loc}}(\Omega)$.

Theorem 41. Assume $\Omega, \mu, \rho, K, h_n, T_n, \lambda_n, m, \{\boldsymbol{t}_i\}_{i=1}^{T_n}$ and $\{\boldsymbol{x}\}_{i=1}^{n}$ satisfy Assumptions (B1)-(B13). We define $\rho_{n,h_n,\lambda_n,T_n}$ as in Definition 39. Then, with probability one, for all $\Omega' \subset\subset \Omega$ we have

$$\|\rho_{n,h_n,\lambda_n,T_n} - \rho\|_{L^\infty(\Omega')} \to 0.$$

E.3 Convergence of minimisers

To show that minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ and $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho_n)$, where $\rho_n \to \rho$, converge to minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ we require a notion of convergence for functionals. For variational convergence the correct notion is Γ-convergence that we recall now.

Definition 42. Let $(\mathcal{X}, d)$ be a metric space. Let $F_n : \mathcal{X} \to \mathbb{R}$ for each $n \in \mathbb{N}$. We say that $(F_n)_{n\in\mathbb{N}}$ Γ-*converges* to $F : \mathcal{X} \to \mathbb{R}$ and write $\Gamma\text{-}\lim_{n\to+\infty} F_n = F$ if

1. (liminf inequality) for every $x \in \mathcal{X}$ and every $(x_n)_{n\in\mathbb{N}}$ such that $x_n \to x$ in $\mathcal{X}$,
$$F(x) \le \liminf_{n\to+\infty} F_n(x_n);$$
2. (existence of recovery sequences) for every $x \in \mathcal{X}$, there exists some sequence $(x_n)_{n\in\mathbb{N}}$ such that $x_n \to x$ in $\mathcal{X}$ and
$$F(x) \ge \limsup_{n\to+\infty} F_n(x_n).$$

The following result then provides conditions for convergence of minimisers, the proof can be found in, for example, [64], [65].

Theorem 43. Let $(\mathcal{X}, d)$ be a metric space and $F_n : \mathcal{X} \to [0, +\infty]$ be a sequence of functionals. Let x_n be a minimising sequence for F_n. If the set $\{x_n\}_{n=1}^\infty$ is precompact and $F_\infty = \Gamma\text{-}\lim_{n\to+\infty} F_n$ where $F_\infty : \mathcal{X} \to [0, +\infty]$ is not identically $+\infty$ then

$$\min_{\mathcal{X}} F_\infty = \lim_{n\to+\infty} \inf_{\mathcal{X}} F_n.$$

Furthermore any cluster point of $\{x_n\}_{n=1}^\infty$ is a minimiser of F_∞.

The following lemma will be useful when considering our cases of Γ-convergence, and compactness of minimisers.

Lemma 44 (Morrey's Theorem)**.** Let $\Omega \subset \mathbb{R}^d$ be an extension domain for $W^{1,p}(\Omega)$ with finite measure (*i.e.* there exist a bounded linear operator $E : W^{1,p}(\Omega) \to W^{1,p}(\mathbb{R}^d)$ such that $Ef|_\Omega = f$ on Ω for every $f \in W^{1,p}(\Omega)$). Let $\{f_n\}_{n\in\mathbb{N}} \subset W^{1,p}(\Omega)$ be a uniformly bounded sequence. Then, if $p > d$ there exist a subsequence $\{f_{n_k}\}_{k\in\mathbb{N}}$ of $\{f_n\}_{n\in\mathbb{N}}$ and a function $f \in C^{0,\alpha}(\Omega)$ such that $f_{n_k} \to f$ as $k \to \infty$ in $C^{0,\alpha}(\Omega)$, for any $0 < \alpha < 1 - \frac{d}{p}$.

Remark 45. We note that as $C^{0,\alpha}(\Omega) \subset L^\infty(\Omega)$, compactness in $L^\infty(\Omega)$ follows.

To show compactness of minimisers, we require a slightly modified Poincaré inequality.

Lemma 46 (Poincaré Inequality)**.** Let $p > d$ and $\Omega \subset \mathbb{R}^d$ be a connected extension domain for $W^{1,p}(\Omega)$ with finite measure. For $f \in W^{1,p}(\Omega)$, let

$$\bar{f} = \frac{1}{N} \sum_{i=1}^{N} f(\boldsymbol{x}_i)$$

for a fixed set $\{\boldsymbol{x}_i\}_{i=1}^N \subset \Omega$. Then, there exists a constant C such that for all $f \in W^{1,p}(\Omega)$,

$$\|f - \bar{f}\|_{L^p(\Omega)} \leq C \|\nabla f\|_{L^p(\Omega)}$$

Proof. The proof is very similar to the proof of the Poincaré inequality found in [66, Theorem 12.23], we just check that one can take the average value of f over finitely many points. Assume for a contradiction that there exists a sequence $f_n \in W^{1,p}(\Omega)$ such that

$$\|f_n - \bar{f}_n\|_{L^p(\Omega)} \geq n \|\nabla f_n\|_{L^p(\Omega)} > 0.$$

Define a centralised, normalised sequence

$$v_n = \frac{f_n - \bar{f}_n}{\|f_n - \bar{f}_n\|_{L^p(\Omega)}},$$

then $v_n \in W^{1,p}(\Omega)$, $\|v_n\|_{L^p(\Omega)} = 1, \bar{v}_n = 0$, and

$$\|\nabla v_n\|_{L^p(\Omega)} = \left\| \nabla \left(\frac{f_n - \bar{f}_n}{\|f_n - \bar{f}_n\|_{L^p(\Omega)}} \right) \right\|_{L^p(\Omega)} = \frac{\|\nabla f_n\|_{L^p(\Omega)}}{\|f_n - \bar{f}_n\|_{L^p(\Omega)}} \leq \frac{1}{n}.$$

Thus by Lemma 44, there exists a subsequence $\{v_{n_k}\}_{n \in \mathbb{N}}$ such that $v_{n_k} \to v$ in $L^\infty(\Omega)$. Further, we must have $\|v\|_{L^p(\Omega)} = 1$ and $\bar{v} = 0$.

Now consider a differentiable, compactly supported function $\varphi : \Omega \to \mathbb{R}$. Then for each derivative of φ (using the Lebesgue dominated convergence theorem, integration by parts and Hölder's inequality)

$$\begin{aligned} \left| \int_\Omega v \frac{\partial \varphi}{\partial x_i} \, \mathrm{d}x \right| &= \lim_{k \to \infty} \left| \int_\Omega v_{n_k} \frac{\partial \varphi}{\partial x_i} \, \mathrm{d}x \right| \\ &= \lim_{k \to \infty} \left| \int_\Omega \varphi \frac{\partial v_{n_k}}{\partial x_i} \, \mathrm{d}x \right| \\ &\leq \|\varphi\|_{L^{p'}(\Omega)} \left\| \frac{\partial v_{n_k}}{\partial x_i} \right\|_{L^p(\Omega)} = 0. \end{aligned}$$

Then $\|\nabla v\|_{L^p(\Omega)} = 0$ so v is constant, and as $\bar{v} = 0$ we must have that $v = 0$ which contradicts $\|v\|_{L^p(\Omega)} = 1$. Hence the required result holds. □

E.3.1 Convergence of the local model

We now state the compactness property for $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot; \rho_n)$. Compactness of minimisers is a corollary.

Proposition 47. Assume that $\Omega, \mu, \eta, p, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, any sequence $\{f_n\}_{n \in \mathbb{N}}$ satisfying $\sup_{n \in \mathbb{N}} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n; \rho_n) < +\infty$ is bounded in $W^{1,p}(\Omega')$ and precom-

pact in $C^{0,\alpha}(\Omega')$ for any $\Omega' \subset\subset \Omega$ and any $0 < \alpha < 1 - \frac{d}{p}$.

Proof. We show that $\{f_n\}_{n\in\mathbb{N}}$ is uniformly bounded in $W^{1,p}(\Omega')$. Compactness in $C^{0,\alpha}(\Omega')$ then follows from Lemma 44.

First, consider $\|\nabla f_n\|_{L^p(\Omega')}$. We note for sufficiently large n that ρ_n is strictly positive for a.e. $\boldsymbol{x} \in \Omega'$, moreover by (A3) and (A9),

$$\rho_n(\boldsymbol{x}) \geq \frac{\rho(\boldsymbol{x})}{\sqrt{2}} \geq \frac{\min_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x})}{\sqrt{2}},$$

for a.e. $\boldsymbol{x} \in \Omega'$ and n sufficiently large. Hence,

$$\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n) \geq \mathcal{E}^{(p)}_{\infty}(f_n;\rho_n) = \sigma_\eta \int_\Omega |\nabla f_n(\boldsymbol{x})|^p \rho_n^2(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \geq \frac{\sigma_\eta}{2} \min_{\boldsymbol{x}\in\Omega}(\rho^2(\boldsymbol{x})) \|\nabla f_n\|^p_{L^p(\Omega')}.$$

Therefore $\sup_{n\in\mathbb{N}} \|\nabla f_n\|_{L^p(\Omega')} < +\infty$.

We are left to show $\sup_{n\in\mathbb{N}} \|f_n\|_{L^p(\Omega')} < +\infty$. By Minkowski's inequality and Lemma 46:

$$\|f_n\|_{L^p(\Omega')} \leq \|f_n - \bar{f}_n\|_{L^p(\Omega')} + |\bar{f}_n| \sqrt[p]{\mathrm{Vol}(\Omega')} \leq C\left(\|\nabla f_n\|_{L^p(\Omega')} + \frac{1}{N}\left|\sum_{i=1}^N y_i\right|\right).$$

Since, by the previous argument, we can bound $\|\nabla f_n\|_{L^p(\Omega')}$ independently of n we have that $\{f_n\}_{n\in\mathbb{N}}$ is bounded in $L^p(\Omega')$ as required. □

An immediate corollary of the previous result is that minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ are bounded.

Corollary 48. Assume that $\Omega, \mu, \eta, p, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ are bounded in $W^{1,p}(\Omega')$ and precompact in $C^{0,\alpha}(\Omega')$ for any $\Omega' \subset\subset \Omega$ and any $\alpha \in (0, 1 - \frac{d}{p})$.

Proof. Choose any $f^\dagger$ that smoothly interpolates between constraints, *i.e.* $\|f^\dagger\|_{W^{1,p}(\Omega)} < +\infty$ and $f^\dagger(\boldsymbol{x}_i) = y_i$ for all $i = 1, \ldots, N$. Let f_n be a minimiser of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$. Clearly $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n) \leq \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f^\dagger;\rho_n) \leq \sigma_\eta \|\rho_n\|^2_{L^\infty(\Omega)} \|\nabla f^\dagger\|^p_{L^p(\Omega)}$. In particular $\sup_{n\in\mathbb{N}} \mathcal{E}^{(p)}_{\infty}(f_n;\rho_n) < +\infty$, hence the result follows from Proposition 47. □

We now consider Γ-convergence for $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$.

Lemma 49. Assume that $\Omega, \mu, \eta, p, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_n)$ Γ-converges to $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$.

Proof. Let $\Omega' \subset\subset \Omega$. We note that for each n there exists δ_n such that for a.e. $\boldsymbol{x} \in \Omega'$,

$$\rho(\boldsymbol{x})\left(1 - \frac{\delta_n}{\inf_{\boldsymbol{x}\in\Omega}(\rho(\boldsymbol{x}))}\right) \leq \rho_n(\boldsymbol{x}) \leq \rho(\boldsymbol{x})\left(1 + \frac{\delta_n}{\inf_{\boldsymbol{x}\in\Omega}(\rho(\boldsymbol{x}))}\right) \tag{E.12}$$

where $\delta_n \to 0$ as $n \to +\infty$. This implies that

$$\left(1 - \frac{\delta_n}{\inf_{\boldsymbol{x}}(\rho(\boldsymbol{x}))}\right)^2 \mathcal{E}^{(p)}_{\infty}(f;\rho\lfloor_{\Omega'}) \leq \mathcal{E}^{(p)}_{\infty}(f;\rho_n\lfloor_{\Omega'}) \leq \left(1 + \frac{\delta_n}{\inf_{\boldsymbol{x}}(\rho(\boldsymbol{x}))}\right)^2 \mathcal{E}^{(p)}_{\infty}(f;\rho\lfloor_{\Omega'}). \tag{E.13}$$

Note that $\mathcal{E}^{(p)}_{\infty}(f;\rho) = \sigma_\eta \|\rho^{\frac{2}{p}} \nabla f\|^p_{L^p(\Omega)}$. This identity and (E.13) will be used to prove the two conditions for Γ-convergence.

(Liminf inequality.) Assume that $f_n \to f$ in $L^p(\Omega)$ and $\liminf_{n\to+\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n) < +\infty$ (else the result is trivial). By recourse to a subsequence (not relabelled) we may assume that

$$\liminf_{n\to+\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n) = \lim_{n\to+\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n)$$

and therefore by the compactness property, Proposition 47, we have that $\{f_n\}_{n\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega')$ and hence there exists a further subsequence (not relabelled) weakly converging in $W^{1,p}(\Omega')$ and, by Lemma 44, strongly in $L^\infty(\Omega')$. Strong convergence in $L^\infty(\Omega')$ implies that f must also satisfy the constraints $f(\boldsymbol{x}_i) = y_i$ for $i = 1,\dots,N$ (where we assume $\boldsymbol{x}_i \in \Omega'$ for all $i = 1,\dots,N$). We note also that $f_n\rho^{\frac{2}{p}}$ is weakly convergent in $W^{1,p}(\Omega')$. Hence, by (E.13) and weak lower semi-continuity of norms

$$\begin{aligned}\liminf_{n\to+\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n) &\geq \liminf_{n\to\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n\lfloor_{\Omega'}) \geq \liminf_{n\to+\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho\lfloor_{\Omega'}) \\ &= \liminf_{n\to+\infty} \sigma_\eta \|\nabla f_n \rho^{\frac{2}{p}}\|^p_{L^p(\Omega')} \geq \sigma_\eta \|\nabla f \rho^{\frac{2}{p}}\|^p_{L^p(\Omega')}.\end{aligned}$$

Taking $\Omega' \to \Omega$ and applying Fatou's lemma we have

$$\liminf_{n\to+\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f_n;\rho_n) \geq \sigma_\eta \|\nabla f \rho^{\frac{2}{p}}\|^p_{L^p(\Omega)} = \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho)$$

as required.

(Recovery sequence.) For a given $f \in L^p(\Omega)$ we choose $f_n = f$ and applying the upper bound in (E.13):

$$\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho_n) \leq \left(1 + \frac{\delta_n}{\inf_{\boldsymbol{x}\in\Omega} \rho(\boldsymbol{x})}\right)^2 \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho) + \sigma_\eta \int_{\Omega\setminus\Omega'} |\nabla f(x)|^p \rho_n^2(x)\,\mathrm{d}x. \tag{E.14}$$

Hence,

$$\limsup_{n\to\infty} \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho_n) \leq \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho) + C\sigma_\eta \left\|\mathbb{1}_{\Omega\setminus\Omega'}\nabla f\right\|^p_{L^p(\Omega)}. \tag{E.15}$$

Taking $\Omega' \to \Omega$ and applying the dominated convergence theorem we get

$$\lim_{\Omega'\to\Omega} \left\|\mathbb{1}_{\Omega\setminus\Omega'}\nabla f\right\|^p_{L^p(\Omega)} \to 0 \tag{E.16}$$

as required. □

The proof of Theorem 36 is then a simple application of Theorem 43 to Corollary 48 and Lemma 49.

E.3.2 Convergence of the non-local model

We start with the compactness result.

Proposition 50. Assume that $\Omega, \mu, \eta, p, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, any sequence $\{f_n\}_{n\in\mathbb{N}}$ satisfying $\sup_{n\in\mathbb{N}} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) < +\infty$ is precompact in $L^p(\Omega')$ for any

$\Omega' \subset\subset \Omega$. Furthermore, if f is a cluster point of $\{f_n\}_{n\in\mathbb{N}}$ in $L^p(\Omega)$ then $f \in C^{0,\alpha}(\Omega')$ for any $0 < \alpha < 1 - \frac{d}{p}$ and $f(\boldsymbol{x}_i) = y_i$ for all $i = 1, \dots, N$.

Proof. By [3, Lemma 4.3] there exists a mollifier J such that J has compact support in $\overline{B(0,1)}$, $J \le C\eta$ for some C, and

$$\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) \ge C\mathcal{E}^{(p)}_{\infty}(J_{\varepsilon_n} * f_n; \mathbb{1}_{\Omega'}\rho_n) \tag{E.17}$$

for any $\Omega' \subset\subset \Omega$ with $\mathrm{dist}(\Omega', \partial\Omega) > \varepsilon_n$, where $\mathbb{1}_{\Omega'}$ is an indicator function over Ω'. We have that $\sup_{n\in\mathbb{N}} \mathcal{E}^{(p)}_{\infty}(\tilde{f}_n; \mathbb{1}_{\Omega'}\rho_n) < +\infty$ where $\tilde{f}_n = J_{\varepsilon_n} * f_n$. Since, for all $i = 1, \dots, N$,

$$\tilde{f}_n(\boldsymbol{x}_i) = \int_{B(0,\varepsilon_n)} J_{\varepsilon_n}(\boldsymbol{z}) f_n(\boldsymbol{x}_i - \boldsymbol{z})\,\mathrm{d}\boldsymbol{z} = y_i \int_{B(0,\varepsilon_n)} J_{\varepsilon_n}(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z} = y_i \tag{E.18}$$

then $\mathcal{E}^{(p)}_{\infty}(\tilde{f}_n; \mathbb{1}_{\Omega'}\rho_n) = \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\tilde{f}_n; \mathbb{1}_{\Omega'}\rho_n)$. By Proposition 47, $\{\tilde{f}_n\}_{n\in\mathbb{N}}$ is bounded in $W^{1,p}(\Omega')$ and precompact in $C^{0,\alpha}(\Omega')$. Let $\tilde{f}_{n_k} \to f$ in $C^{0,\alpha}(\Omega')$ as $k \to +\infty$. We claim that $f_{n_k} \to f$ in $L^p(\Omega')$.

It is enough to show that $\|f_{n_k} - \tilde{f}_{n_k}\|_{L^p(\Omega')} \to 0$. By Jensen's inequality we have

$$\begin{aligned}
\|f_{n_k} - \tilde{f}_{n_k}\|^p_{L^p(\Omega')} &= \int_{\Omega'} \left| f_{n_k}(\boldsymbol{x}) - \int_{B(0,\varepsilon_{n_k})} J_{\varepsilon_{n_k}}(\boldsymbol{x}-\boldsymbol{z}) f_{n_k}(\boldsymbol{z})\,\mathrm{d}\boldsymbol{z} \right|^p \mathrm{d}\boldsymbol{x} \\
&\le \int_{\Omega'} \int_{B(0,\varepsilon_{n_k})} J_{\varepsilon_{n_k}}(\boldsymbol{x}-\boldsymbol{z}) \left| f_{n_k}(\boldsymbol{x}) - f_{n_k}(\boldsymbol{z}) \right|^p \mathrm{d}\boldsymbol{z}\,\mathrm{d}\boldsymbol{x} \\
&\le \frac{C\varepsilon^p_{n_k}}{\inf_{\boldsymbol{x}\in\Omega} \rho^2_{n_k}(\boldsymbol{x})} \mathcal{F}^{(p)}_{\varepsilon_{n_k}}(f_{n_k};\rho_{n_k}) \to 0
\end{aligned}$$

It follows that $\{f_n\}_{n\in\mathbb{N}}$ is compact in $L^p(\Omega')$. □

As in the previous section the above compactness property can be applied to minimisers.

Corollary 51. Assume that $\Omega, \mu, \eta, p, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, minimisers of $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho_n)$ are precompact in $L^p(\Omega')$, for any $\Omega' \subset\subset \Omega$. Furthermore if f is a cluster point of $\{f_n\}_{n\in\mathbb{N}}$ in $L^p(\Omega')$ then $f \in C^{0,\alpha}(\Omega')$ for any $0 < \alpha < 1 - \frac{d}{p}$ and $f(\boldsymbol{x}_i) = y_i$ for all $i = 1, \dots, N$.

Proof. Let $R > 0$ satisfy $\min_{i\neq j\in\{1,\dots,N\}} |\boldsymbol{x}_i - \boldsymbol{x}_j| \ge 3R$. Choose any $f^\dagger \in C^\infty(\Omega')$ that smoothly interpolates between constraints on balls of radius R around each $\boldsymbol{x}_i$ (for $i = 1, \dots, N$), *i.e.* $\|f^\dagger\|_{W^{1,p}(\Omega')} < +\infty$ and $f^\dagger(\boldsymbol{x}) = y_i$, for $\boldsymbol{x} \in B(\boldsymbol{x}_i, R)$, $i = 1, \dots, N$. We assume $\varepsilon_n < R$ and let L be the Lipschitz constant for $f^\dagger$.

Let f_n be a sequence of minimisers of $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}$. Then,

$$\begin{aligned}
\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) &\le \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f^\dagger;\rho_n) \\
&\le \frac{L\|\rho_n\|^2_{L^\infty(\Omega)}}{\varepsilon^p_n} \int_\Omega \int_\Omega \eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|\boldsymbol{x}-\boldsymbol{z}|^p\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} \\
&= L\|\rho_n\|^2_{L^\infty(\Omega)} \mathrm{Vol}(\Omega) \int_\Omega \eta(|\boldsymbol{x}|)|\boldsymbol{x}|^p\,\mathrm{d}\boldsymbol{x} \\
&= d\sigma_\eta L\|\rho_n\|^2_{L^\infty(\Omega)} \mathrm{Vol}(\Omega).
\end{aligned}$$

So, $\sup_{n\in\mathbb{N}} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) < +\infty$, hence the result follows from Proposition 50. □

We now prove Γ-convergence.

Lemma 52. Assume that $\Omega, \mu, \eta, p, \rho_n$ and $\{\boldsymbol{x}_i\}_{i=1}^n$ satisfy Assumptions (A1)-(A9). Then, $\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(\cdot;\rho_n)$ Γ-converges to $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$.

Proof. We recall (E.12) and therefore for there exists a sequence $\gamma_n \to 0$ such that

$$(1-\gamma_n)^2 \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f;\rho\lfloor_{\Omega'}) \leq \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f;\rho_n\lfloor_{\Omega'}) \leq (1+\gamma_n)^2 \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f;\rho\lfloor_{\Omega'}) \tag{E.19}$$

for all $f \in L^p(\Omega)$.

(Liminf inequality.) Assume $f_n \to f$ in $L^p(\Omega)$ and $\liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) < \infty$ else the result is trivial. By recourse to a subsequence (relabelled) we assume that

$$\liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) = \lim_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n).$$

By the compactness property (Proposition 50) we have that $f(x_i) = y_i$ for all $i = 1,\dots,N$. Now,

$$\liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) \geq \liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n\lfloor_{\Omega'}) \geq \liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n}(f_n;\rho\lfloor_{\Omega'}) \geq \mathcal{E}^{(p)}_{\infty}(f;\rho\lfloor_{\Omega'})$$

by [3, Lemma 4.6] and (E.19). By Fatou's lemma $\liminf_{\Omega'\to\Omega} \mathcal{E}^{(p)}_{\infty}(f,\rho\lfloor_{\Omega'}) \geq \mathcal{E}^{(p)}_{\infty}(f,\rho)$, hence

$$\liminf_{n\to\infty} \mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) \geq \mathcal{E}^{(p)}_{\infty}(f,\rho).$$

Since the constraints are satisfied then $\mathcal{E}^{(p)}_{\infty}(f;\rho) = \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho)$.

(Recovery sequence.) We prove the recover sequence in three parts.

Part 1: Assume $f \in W^{1,p}(\Omega)$ is Lipschitz continuous with $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho) < \infty$ and η has compact support in $B(0,M)$. We define

$$f_n(x) = \begin{cases} y_i & \text{if } |x - x_i| < \varepsilon_n \text{ for } i = 1,\dots,N \\ f(x) & \text{else.} \end{cases}$$

Then as f_n and f agree away from the constraints,

$$\begin{aligned}
\|f_n - f\|^p_{L^p(\Omega)} &= \int_\Omega |f_n(\boldsymbol{x}) - f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}, \\
&= \int_{\cup_{i=1}^N B(\boldsymbol{x}_i,\varepsilon_n)} |y_i - f(\boldsymbol{x})|^p \,\mathrm{d}\boldsymbol{x}, \\
&\leq \mathrm{Lip}(f)^p \sum_{i=1}^N \int_{B(\boldsymbol{x}_i,\varepsilon_n)} |\boldsymbol{x}_i - \boldsymbol{x}|^p \,\mathrm{d}x \\
&\leq \mathrm{Lip}(f)^p N \varepsilon_n^p \mathrm{Vol}(B(0,\varepsilon_n)) \\
&\leq C\varepsilon_n^{p+d}.
\end{aligned}$$

Hence, $f_n \to f$ in $L^p(\Omega)$.

We recall the following: for all $\xi > 0$ there exists a constant $C_\xi > 0$ such that, for any $a, b \in \mathbb{R}^d$,

$$|a|^p - |b|^p \le \xi |b|^p + C_\xi |a-b|^p.$$

So, for a fixed $\xi > 0$,

$$\begin{aligned}
&\mathcal{F}^{(p)}_{\varepsilon_n}(f_n;\rho) - \mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) \\
&\quad = \frac{1}{\varepsilon_n^p}\int_\Omega\int_\Omega \eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)\Big(|f_n(\boldsymbol{x})-f_n(\boldsymbol{z})|^p - |f(\boldsymbol{x})-f(\boldsymbol{z})|^p\Big)\rho(\boldsymbol{x})\rho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} \\
&\quad \le \frac{\xi}{\varepsilon_n^p}\int_\Omega\int_\Omega \eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|f(\boldsymbol{x})-f(\boldsymbol{z})|^p\rho(\boldsymbol{x})\rho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} \\
&\qquad + \frac{C_\xi}{\varepsilon_n^p}\int_\Omega\int_\Omega \eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|f_n(\boldsymbol{x})-f_n(\boldsymbol{z})-f(\boldsymbol{x})+f(\boldsymbol{z})|^p\rho(\boldsymbol{x})\rho(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} \\
&\quad \le \xi\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) + \frac{2pC_\xi\|\rho\|^2_{L^\infty(\Omega)}}{\varepsilon_n^p}\int_\Omega\int_\Omega \eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|f_n(\boldsymbol{x})-f(\boldsymbol{x})|^p\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} \\
&\quad = \xi\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) + \frac{2pC_\xi\|\rho\|^2_{L^\infty(\Omega)}}{\varepsilon_n^p}\int_{\mathbb{R}^d}\eta(|\boldsymbol{x}|)\,\mathrm{d}\boldsymbol{x}\|f_n-f\|^p_{L^p(\Omega)} \\
&\quad \le \xi\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) + \tilde{C}\varepsilon_n^d.
\end{aligned}$$

Hence, $\mathcal{F}^{(p)}_{\varepsilon_n}(f_n;\rho) \le (1+\xi)\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) + O(\varepsilon_n^d)$.

Applying the above to (E.19) we have,

$$\begin{aligned}
\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) &\le \mathcal{F}^{(p)}_{\varepsilon_n}(f_n;\rho) + \frac{2}{\varepsilon_n^p}\int_\Omega\int_{\Omega\setminus\Omega'}\eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|f(\boldsymbol{x})-f(\boldsymbol{z})|^p\rho_n(\boldsymbol{x})\rho_n(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} \\
&\le (1+\xi)\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) + O(\varepsilon_n^d) \\
&\quad + \frac{2\|\rho_n\|^2_{L^\infty}\mathrm{Lip}(f)}{\varepsilon_n^p}\int_{\mathrm{dist}(z,\partial\Omega)\le d_{\mathrm{H}}(\Omega,\Omega')+M\varepsilon_n}\int_{\Omega\setminus\Omega'}\eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|\boldsymbol{x}-\boldsymbol{z}|^p\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z} \\
&\le (1+\xi)\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) + O(\varepsilon_n^d) \\
&\quad + 2d\sigma_\eta\mathrm{Lip}(f)\mathrm{Vol}\left(\{z : \mathrm{dist}(z,\partial\Omega)\le d_{\mathrm{H}}(\Omega,\Omega')+M\varepsilon_n\}\right)\sup_{n\in\mathbb{N}}\|\rho_n\|^2_{L^\infty}.
\end{aligned}$$

By [3, Lemma 4.6] we have $\limsup_{n\to\infty}\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho) \le \mathcal{E}^{(p)}_\infty(f;\rho) = \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho)$. So taking $n\to\infty$ followed by $\xi\to 0$ and $\Omega'\to\Omega$ we have,

$$\limsup_{n\to\infty}\mathcal{F}^{(p)}_{\varepsilon_n,\mathrm{con}}(f_n;\rho_n) \le \mathcal{E}^{(p)}_{\infty,\mathrm{con}}(f;\rho).$$

Part 2: We still assume that f is Lipschitz continuous but relax the compact support assumption on η. Assume η satisfies the integrability condition in (A7). We define $\mathcal{F}^{(p)}_{\varepsilon_n}(\cdot;\rho,\eta)$ to be the functional $\mathcal{F}^{(p)}_{\varepsilon_n}(\cdot;\rho)$ with weight function η. Then, we let η^M be the truncated weight function $\eta^M(t) = \eta(t)\mathbb{1}_{t\le M}$. Now,

$$\mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho_n,\eta) = \mathcal{F}^{(p)}_{\varepsilon_n}(f;\rho_n,\eta^M) + \frac{1}{\varepsilon_n^p}\int\int_{|\boldsymbol{x}-\boldsymbol{z}|>M\varepsilon_n}\eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|f(\boldsymbol{x})-f(\boldsymbol{z})|^p\rho_n(\boldsymbol{x})\rho_n(\boldsymbol{z})\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{z}.$$

We can apply part 1 to the first term on the right hand side. For the second term, for each $\boldsymbol{z} \in \Omega$,

$$\frac{1}{\varepsilon_n^p} \int_{|\boldsymbol{x}-\boldsymbol{z}|>M\varepsilon_n} \eta_{\varepsilon_n}(|\boldsymbol{x}-\boldsymbol{z}|)|f(\boldsymbol{x})-f(\boldsymbol{z})|^p \rho_n(\boldsymbol{x})\,\mathrm{d}\boldsymbol{x} \leq \mathrm{Lip}(f)\|\rho_n\|_{L^\infty} \int_{|\boldsymbol{w}|\geq M} \eta(|\boldsymbol{w}|)|\boldsymbol{w}|^p\,\mathrm{d}\boldsymbol{w}.$$

Hence, (using $\eta^M \leq \eta$)

$$\limsup_{n\to\infty} \mathcal{F}_{\varepsilon_n}^{(p)}(f;\rho_n,\eta) \leq \mathcal{E}_\infty^{(p)}(f;\rho,\eta) + \mathrm{Lip}(f) \sup_{n\in\mathbb{N}} \|\rho_n\|_{L^\infty}^2 \mathrm{Vol}(\Omega) \int_{|\boldsymbol{w}|\geq M} \eta(|\boldsymbol{w}|)|\boldsymbol{w}|^p\,\mathrm{d}\boldsymbol{w}.$$

By the monotone convergence theorem, taking $M \to \infty$ we have

$$\limsup_{n\to\infty} \mathcal{F}_{\varepsilon_n}^{(p)}(f;\rho_n,\eta) \leq \mathcal{E}_\infty^{(p)}(f;\rho,\eta)$$

as required.

Part 3: Since Lipschitz functions are dense in $W^{1,p}$ we can, as is usual in Γ-convergence arguments, conclude by a diagonalisation argument. □

E.4 Convergence of density estimates

In the following two subsections we prove that the kernel density estimate, and the spline kernel density estimate, satisfy Assumption (A9).

E.4.1 Convergence of the kernel density estimate

Our result is an easy consequence of the following theorem due to [11, Theorem 2.3].

Theorem 53. Let $K = \phi \circ \xi$ where ϕ is a bounded function of bounded variation and ξ is a polynomial. Assume $x_i \overset{\text{iid}}{\sim} \mu$ where $\mu \in \mathcal{P}(\mathbb{R}^d)$ has a bounded density ρ. Assume h_n satisfies

$$h_n \to 0^+, \quad \frac{nh_n^d}{|\log(h_n)|} \to \infty, \quad \frac{|\log(h_n)|}{\log\log(n)} \to \infty, \quad \text{and} \quad h_n \leq ch_{2n}$$

for some $c > 0$. Define $\rho_{n,h}$ as in (E.8) and $\bar{\rho}_h$ by

$$\bar{\rho}_h(\boldsymbol{x}) = \mathbb{E}\rho_{n,h}(x) = \int_{\mathbb{R}^d} K_h(\boldsymbol{x}-\boldsymbol{z})\,\mathrm{d}\mu(\boldsymbol{z}).$$

Then there exists $C > 0$ such that, with probability one,

$$\limsup_{n\to\infty} \sqrt{\frac{nh_n^d}{|\log h_n|}} \|\rho_{n,h_n} - \bar{\rho}_{h_n}\|_{L^\infty(\mathbb{R}^d)} = C.$$

As remarked in Section E.2.5 [11] treats a more general class of kernels K and for example one could also include kernels of the form $K = \mathbb{1}_{[-1,1]^d}$. We now prove Theorem 40.

Proof of Theorem 40. We extend ρ to the whole of $\mathbb{R}^d$ by setting $\rho(\boldsymbol{x}) = 0$ for all $\boldsymbol{x} \in \mathbb{R}^d \setminus \Omega$. Now, for any $\Omega' \subset\subset \Omega$,

$$\|\rho_{n,h_n} - \rho\|_{L^\infty(\Omega')} \leq \|\rho_{n,h_n} - \bar{\rho}_{h_n}\|_{L^\infty(\Omega')} + \|\bar{\rho}_{h_n} - \rho\|_{L^\infty(\Omega')}$$

the first term on the RHS goes to zero by Theorem 53. For the second term we define

$$\tilde{\Omega} = \left\{ \boldsymbol{x} \in \Omega \,:\, \inf_{\boldsymbol{z}\in\Omega'} |\boldsymbol{x}-\boldsymbol{z}| \le \frac{1}{2} d_{\mathrm{H}}(\Omega', \Omega) \right\}$$

where d_{H} is the Hausdorff distance, and choose $\delta > 0$. Since ρ is uniformly continuous on $\tilde{\Omega}$ there exists $R_\delta > 0$ such that for all $\boldsymbol{x}, \boldsymbol{z} \in \tilde{\Omega}$ with $|\boldsymbol{x}-\boldsymbol{z}| < R_\delta$ we have $|\rho(\boldsymbol{x}) - \rho(\boldsymbol{z})| \le \delta$. Let $\mathrm{spt}(K) \subset B(0, M)$, and assume n is large enough so that

$$h_n \le \frac{1}{M} \min\left\{ \frac{1}{2} d_{\mathrm{H}}(\Omega', \Omega), R_\delta \right\}.$$

Then, for any $x \in \Omega'$,

$$\begin{aligned} |\bar{\rho}_{h_n}(\boldsymbol{x}) - \rho(\boldsymbol{x})| &= \left| \int_{\mathbb{R}^d} K_{h_n}(\boldsymbol{x}-\boldsymbol{z}) \left(\rho(\boldsymbol{z}) - \rho(\boldsymbol{x})\right) \mathrm{d}\boldsymbol{z} \right| \\ &\le \int_{\mathbb{R}^d} |K_{h_n}(\boldsymbol{x}-\boldsymbol{z})| \, |\rho(\boldsymbol{z}) - \rho(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{z} \\ &\le \delta \|K\|_{L^1(\mathbb{R}^d)}. \end{aligned}$$

Hence $\lim_{n\to\infty} \|\bar{\rho}_{h_n} - \rho\|_{L^\infty(\Omega')} \le \delta \|K\|_{L^1(\mathbb{R}^d)}$. Since $\delta > 0$ is arbitrary we have shown $\lim_{n\to\infty} \|\bar{\rho}_{h_n} - \rho\|_{L^\infty(\Omega')} = 0$ as required. □

The following result allows us to extend the convergence to sets Ω_h where

$$\Omega_h = \{x \in \Omega \,:\, \mathrm{dist}(x, \partial\Omega) \ge h\}. \tag{E.20}$$

Lemma 54. If in addition to the assumptions in Theorem 40 we assume that ρ is Lipschitz continuous on Ω, then

$$\lim_{n\to\infty} \|\rho_{n,h_n} - \rho\|_{L^\infty(\Omega_{Mh_n})} = 0$$

where Ω_h is defined by (E.20).

Proof. Analogously to the proof of Theorem 40 we have

$$\|\rho_{n,h_n} - \rho\|_{L^\infty(\Omega_{Mh_n})} \le \|\rho_{n,h} - \bar{\rho}_{h_n}\|_{L^\infty(\mathbb{R}^d)} + \|\bar{\rho}_{h_n} - \rho\|_{L^\infty(\Omega_{h_n})}$$

where the first term goes to zero by Theorem 53. The second term goes to zero uniformly by, for all $x \in \Omega_{Mh_n}$,

$$\begin{aligned} |\bar{\rho}_{h_n}(\boldsymbol{x}) - \rho(\boldsymbol{x})| &\le \int_{\mathbb{R}^d} |K_{h_n}(\boldsymbol{x}-\boldsymbol{z})| |\rho(\boldsymbol{z}) - \rho(\boldsymbol{x})| \, \mathrm{d}\boldsymbol{z} \\ &\le \mathrm{Lip}(\rho) h_n \int_{\mathbb{R}^d} |K(\boldsymbol{w})| \|\boldsymbol{w}\| \, \mathrm{d}\boldsymbol{w}. \end{aligned}$$

Hence $\|\bar{\rho}_{h_n} - \rho\|_{L^\infty(\Omega_{h_n})} \to 0$ as required. □

E.4.2 Convergence of the spline kernel density estimate

Our method relies on the result of [27] (given below), where almost sure H^m error estimates were constructed for multivariate spline functions from data with uncorrelated, centred noise with results from [63]. Here we adapt the results to suit SKDE. By the linearity of the smoothing spline functional we can write

$$\rho_{n,h,\lambda,T} = S_{\lambda,T}(\rho + v_{n,h} - \bar{v}_{n,h}) + S_{\lambda,T}(\bar{v}_{n,h})$$

where

$$\begin{aligned} v_{n,h} &= \rho_{n,h} - \rho \\ \bar{v}_{n,h} &= \mathbb{E}v_n = \int_\Omega K(\boldsymbol{x})\left(\rho(\cdot - h\boldsymbol{x}) - \rho(\cdot)\right)\,\mathrm{d}\boldsymbol{x}. \end{aligned}$$

By the triangle inequality, for $\Omega' \subset\subset \Omega$,

$$\|\rho_{n,h,\lambda,T} - \rho\|_{H^m(\Omega')} \leq \|S_{\lambda,T}(\rho + v_{n,h} - \bar{v}_{n,h}) - \rho\|_{H^m(\Omega')} + \|S_{\lambda,T}(\bar{v}_{n,h})\|_{H^m(\Omega')}.$$

The first term on the right hand side can be bounded by following theorem found in [27, Theorem 4.1].

Theorem 55. In addition to Assumptions (B1),(B8)-(B11) on Ω, T_n, m and $\{\boldsymbol{t}_i\}_{i=1}^{T_n}$, assume $\delta_n \in \mathbb{R}^{T_n}$ are random variables satisfying $\mathbb{E}\delta_{n,i} = 0$, for all i, $\delta_{n,i}$ is independent of $\delta_{n,j}$ for all $i \neq j$, and

$$\forall r \in \mathbb{N}, \quad \exists C = C(r) \quad \text{such that} \quad \forall n \in \mathbb{N}, \text{ and } \forall i = 1, \ldots, T_n \quad \text{we have} \quad \mathbb{E}|\delta_{n,i}|^{2r} \leq C.$$

Define $S_{\lambda,T}$ by (E.11) and P_T by (E.10). Then, for $f \in H^m(\Omega)$,

$$\lim_{n\to\infty} \|S_{\lambda_n,T_n}(P_{T_n}(f) + \delta_n) - f\|_{H^m(\Omega)} = 0$$

with probability one.

It is easy to check that for $\delta_n = P_{T_n}(v_{n,h_n} - \bar{v}_{n,h_n})$ that $\delta_{n,i}$ are independent whenever $\mathrm{Sep}(T_n) \geq 2Mh_n$ where $\mathrm{spt}(K) \subset B(0, M)$. Moreover,

$$\mathbb{E}|\delta_{n,i}|^{2r} \leq 2^{2r}\mathbb{E}\|\rho_{n,h_n} - \rho\|^{2r}_{L^\infty(\Omega_{h_n})}$$

for $\{\boldsymbol{t}_i\}_{i=1}^{T} \subset \Omega_{Mh}$ and therefore, by Lemma 54, the RHS converges to zero. Hence, we can apply the above theorem to infer that $\|S_{\lambda_n,T_n}(\rho + v_{n,h_n} - \bar{v}_{n,h_n}) - \rho\|_{H^m(\Omega')} \to 0$ with probability one.

Theorem 56. Under the conditions of Theorem 41 with probability one, for all $\Omega' \subset\subset \Omega$, we have

$$\|\rho_{n,h_n,\lambda_n,T_n} - \rho\|_{H^m(\Omega')} \to 0.$$

Proof. By the preceding argument and Theorem 55 it is enough to show that $\|S_{\lambda_n,T_n}(\bar{v}_{n,h_n})\|_{H^m(\Omega')} \to 0$. Let $\gamma_n = S_{\lambda_n,T_n}(\bar{v}_{n,h_n})$ and Γ_n be the $T_n \times T_n$ matrix satisfying

$$\boldsymbol{z}^\top \Gamma_n \boldsymbol{z} = \min\left\{\|\nabla^m u\|^2_{L^2(\Omega)} \,:\, u \in H^m(\Omega), u(\boldsymbol{t}_i) = z_i\, \forall i = 1, \ldots, T_n\right\}.$$

One has, see [63, Section 5],

$$\frac{1}{T_n}\sum_{i=1}^{T_n}|\gamma_n(\boldsymbol{t}_i)|^2 = \frac{1}{T_n}\bar{v}_{n,h_n}^\top A_n^2 \bar{v}_{n,h} \le \frac{1}{T_n}\|\bar{v}_{n,h_n}\|^2_{L^\infty(\Omega_{Mh_n})}\mathrm{tr}(A_n^2)$$
$$\|\nabla^m\gamma_n\|_{L^2(\Omega)} = \frac{1}{T_n\lambda_n}\bar{v}_{n,h_n}^\top (A_n - A_n^2)\bar{v}_{n,h_n} \le \frac{1}{T_n\lambda_n}\|\bar{v}_{n,h_n}\|^2_{L^\infty(\Omega_{n,h_n})}\mathrm{tr}(A_n)$$

where $A_n = (\mathrm{Id} + \lambda_n T_n \Gamma_n)^{-1}$. By [63, Theorem 5.3] there exists $C_1 > 0$, $C_2 > 0$ and $M > 0$ such that

$$\alpha_i^{(n)} = 0 \quad \forall i = 1,\dots,M \qquad \text{and} \qquad C_1 i^{\frac{2m}{d}} \le \alpha_i^{(n)} \le C_2 i^{\frac{2m}{d}} \quad \forall i = M+1,\dots,T_n$$

where $\alpha_i^{(n)}$ are the ordered eigenvalues of $T_n\Gamma_n$. Hence,

$$\begin{aligned}
\mathrm{tr}(A_n^2) &= \sum_{i=1}^{T_n}\frac{1}{(\alpha_i^{(n)}\lambda_n + 1)^2} \\
&\le M + \sum_{i=M+1}^{T_n}\frac{1}{(C_1 i^{\frac{2m}{d}}\lambda_n + 1)^2} \\
&\le M + \int_M^{T_n}\frac{1}{(C_1 t^{\frac{2m}{d}} + 1)^2}\,\mathrm{d}t \\
&\le M + \frac{1}{(C_1\lambda_n)^{\frac{d}{2m}}}\int_0^\infty \frac{1}{(s^{\frac{2m}{d}}+1)^2}\,\mathrm{d}s
\end{aligned}$$

and similarly,

$$\mathrm{tr}(A_n) \le M + \frac{1}{(C_1\lambda_n)^{\frac{d}{2m}}}\int_0^\infty \frac{1}{s^{\frac{2m}{d}}+1}\,\mathrm{d}s.$$

So,

$$\frac{1}{T_n}\sum_{i=1}^{T_n}|\gamma_n(\boldsymbol{t}_i)|^2 \le \frac{\tilde{C}}{T_n\lambda_n^{\frac{d}{2m}}}\|\bar{v}_{n,h_n}\|^2_{L^\infty(\Omega_{Mh_n})}$$
$$\|\nabla^m\gamma_n\|_{L^2(\Omega)} \le \frac{\tilde{C}}{T_n\lambda_n^{1+\frac{d}{2m}}}\|\bar{v}_{n,h_n}\|^2_{L^\infty(\Omega_{n,h_n})}.$$

By [63, Theorem 3.4] we can bound

$$\|\gamma_n\|^2_{H^m(\Omega')} \le C'\left(\frac{1}{T_n}\sum_{i=1}^{T_n}|\gamma_n(\boldsymbol{t}_i)|^2 + \|\nabla^m\gamma_n\|^2_{L^2(\Omega')}\right) \le \frac{\hat{C}\|\bar{v}_{n,h_n}\|^2_{L^\infty(\Omega_{Mh_n})}}{T_n\lambda_n^{\frac{2m+d}{2m}}}.$$

Hence, $\gamma_n \to 0$ in $H^m(\Omega')$ as required. □

The proof of Theorem 41 is now, due to Sobolev embeddings (in particular Morrey's inequality), just a corollary of the above theorem since $m > d/2$.

E.5 Numerical experiments

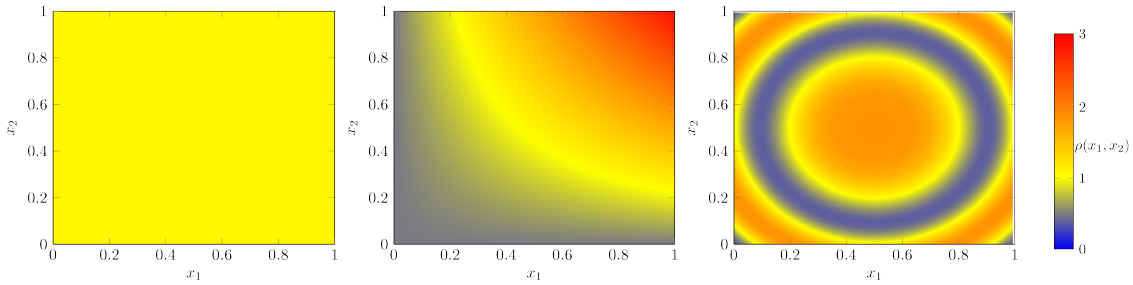
To provide evidence of the convergence results stated in this paper, we consider numerical examples of the KDE and SKDE, construct methods to determine the minimisers of different p-Dirichlet energies, and provide example computations and error estimates. We compare our results in terms of computation time with [4]. By approximating the densities and discretising on a coarser grid we introduce another source of error which could be significant for small data sizes, hence our method is not state-of-the-art in the small data regime. On the other hand, discrete based methods fairly quickly become computationally infeasible whereas the continuum limit based numerical method controls the computational cost allowing one to apply the method to very large datasets; this is the regime where our approach is state-of-the-art.

E.5.1 Setup

We consider the domain $\Omega = [0,1] \times [0,1]$, and sample from three different densities:

$$\begin{aligned}
\rho_1(x,y) &= 1, \\
\rho_2(x,y) &= \frac{1}{\mathcal{N}_2}(xy + 0.2), \\
\rho_3(x,y) &= \frac{1}{\mathcal{N}_3}(\cos\left(6\pi((x-0.5)^2 + (y-0.2)^2)\right)/3 + 0.5),
\end{aligned}$$

where $\mathcal{N}_2, \mathcal{N}_3$ are normalisation constants. The densities are plotted in Figure E.2.



(a) (b) (c)

Figure E.2: Three densities considered in the examples. Left: ρ_1, centre: ρ_2, right: ρ_3.

For simplicity, we position 16 constraints uniformly across the domain, and labels are given using the formula:

$$C(x,y) = 4\left(x - \frac{1}{2}\right)^2 + \left(y - \frac{1}{2}\right)^2.$$

The constraints are presented graphically in Figure E.3.

Code is based on the pseudo-spectral code base `2DChebClass`, [67]. The numerical methods used in this paper also rely on boundary patching methods. A version of `2DChebClass` which includes boundary patching and p-Dirichlet minimisation is available upon request.

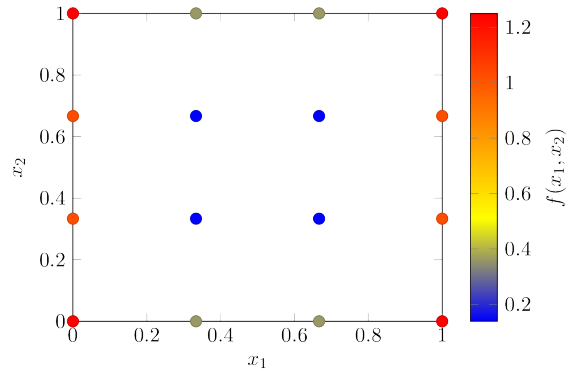


Figure E.3: The position and value of the constraints used in each example.

E.5.2 Density estimation

Numerical method

For the numerical results of the density estimate, we discretise the domain Ω uniformly with $D = 2^{10}$ evenly spaced grid points in each dimension and consider

$$\tilde{\Omega}_D = \{(x_i, y_j), x_i = \frac{i}{D-1}, y_j = \frac{j}{D-1}, i, j = 0, \dots, (D-1)\}.$$

We choose D large, so that discretization errors are small. We sample from non-uniform densities using the `MATLAB` function `pinky` [68]. For the density estimates we consider two measures of error, given a density estimate ρ_n, we consider the L^2 error:

$$\|\rho_n - \rho\|^2_{L^2(\Omega')} = \int_{\Omega'} |\rho_n(\boldsymbol{x}) - \rho(\boldsymbol{x})|^2 \, \mathrm{d}\boldsymbol{x} \approx \frac{1}{D^2} \sum_{i,j=0}^{D-1} \mathbb{1}_{(x_i,y_j)\in\Omega'} |\rho_n(x_i, y_j) - \rho(x_i, y_j)|^2,$$

and the L^∞ error:

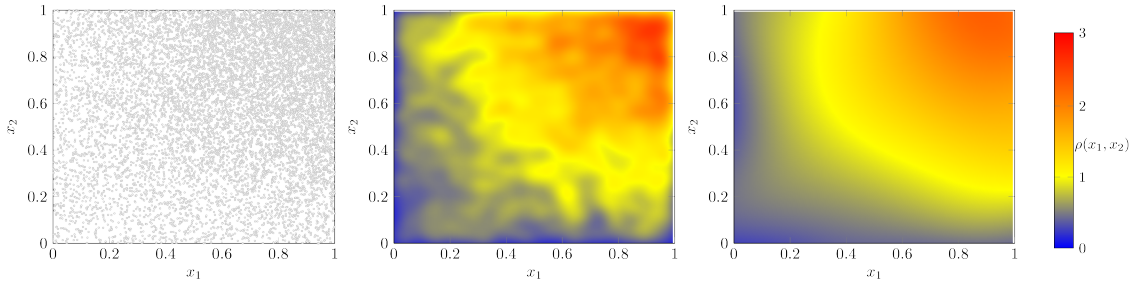
$$\|\rho_n - \rho\|_{L^\infty(\Omega')} = \sup_{\boldsymbol{x}\in\Omega'} |\rho_n(\boldsymbol{x}) - \rho(\boldsymbol{x})| \approx \sup_{(x,y)\in\tilde{\Omega}_D\cap\Omega'} |\rho_n(x, y) - \rho(x, y)|.$$

As we are only interested in the local approximation to the density, we construct the L^∞ and L^2 errors on $\Omega' = [0.01, 0.99] \times [0.01, 0.99]$. Kernel density estimates and smoothing splines are well studied, so we can utilise built-in functions in `MATLAB` in our calculations. We construct the KDE using the built-in function `mvksdensity` using Gaussian kernels (which are kernel functions of order 2). To construct the SKDE we use the built-in function `spaps`. For simplicity, we take the knots $\{t_i\}_{i=1,\dots,T}$ to be evenly spaced across the domain. An example of samples from ρ_2 and the associated KDE and SKDE is given in Figure E.4.

Results and discussion

We present the L^∞ errors for the density and the derivative in Figure E.5 and Figure E.6. We take $T = 2^{12}$ and $\lambda = 10^{-6}$. For simplicity T and λ remain fixed in the experiments.

Figures E.5 and E.6 show that the SKDE does no worse that the KDE, and often performs better, in terms of L^∞ error. Occasionally, we see that the L^∞ error for the SKDE is greater

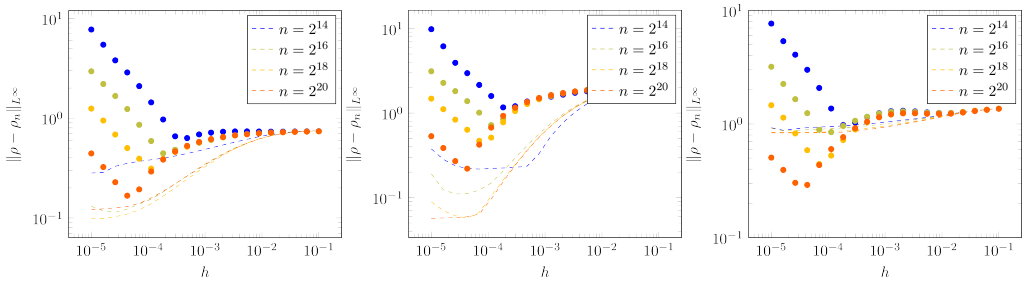


(a) (b) (c)

Figure E.4: Left:10000 samples from ρ_2 using `pinky`, centre: the KDE with $h = 0.03$ using `mvksdensity`, right: the SKDE with $h = 0.03, \lambda = 10^{-6}, T = 2^{12}$ using `spaps`. Density estimates are calculated on a mesh of $2^{10} \times 2^{10}$ points.

than the KDE. This may be because of larger fluctuations in the derivatives, where the SKDE over-smooths the KDE. For all three densities, the KDE density error has an optimal choice of bandwidth h, as predicted by the theory.

Although in some cases the improved approximation due to smoothing splines is small, they also provide additional robustness in the choice of bandwidth h, with very little additional computation cost. We present the computation time for the KDE and SKDE in Figure E.7. The computation time for different densities is almost identical, and the inclusion of a smoothing spline approximation is negligible in cost. We note that the computational cost of kernel density estimation can also be significantly reduced using parallelisation.

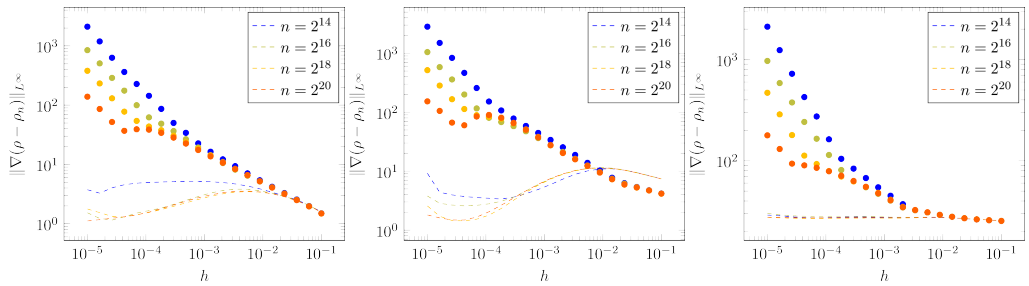


(a) (b) (c)

Figure E.5: L^∞ errors for the three densities using the two density estimates. Dotted lines represent the KDE, while dashed lines represent the SKDE. Left: ρ_1, centre: ρ_2, right: ρ_3.

E.5.3 p-Dirichlet energy minimisation

We wish to compare the accuracy and efficiency of different Dirichlet energies on different densities. In contrast to the discrete p-Dirichlet energies, the continuum p-Dirichlet energies are not prohibitively expensive when n is large. However, when d is large, standard numerical methods become computationally intractable as the number of discretization points increases exponentially in dimension. It is an area of future work to construct a numerical scheme which can find the minimiser of (E.2) when d is large, perhaps under additional assumptions on the underlying probability density of the data. For now, we restrict our numerical investigation to problems with $d = 2$, and



(a) (b) (c)

Figure E.6: L^∞ errors for the derivatives of the three densities using the two density estimates. Dotted lines represent the KDE, while dashed lines represent the SKDE. Left: ρ_1, centre: ρ_2, right: ρ_3.

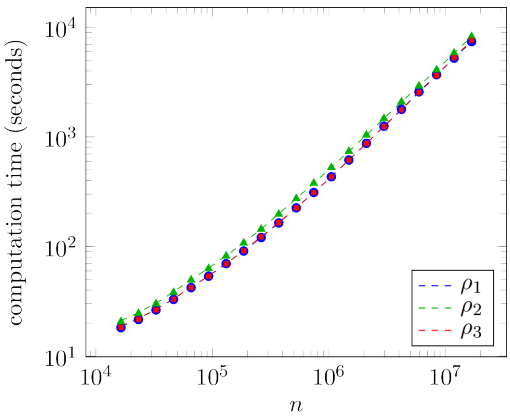


Figure E.7: Computation time for KDE and SKDE density estimates. Dashed lines represent KDE computation times, the larger circles, squares and triangles of the same color are SKDE computation times for each density.

focus on the large data problem, rather than the large dimension problem.

Numerical methods

When $p \neq 2$, the minimisation problem (E.6) we must consider is nonlinear. Therefore, to construct minimisers for the different p-Dirichlet energies we use gradient descent.

Discrete p-Dirichlet energies For (E.1) we consider the gradient flow

$$\begin{aligned}\frac{\partial f(\boldsymbol{x}_i)}{\partial t} &= \begin{cases} 0 & i = 1, \dots, N, \\ -\frac{p}{\varepsilon_n^p n^2} \sum_{j=1}^n W_{ij}(f(\boldsymbol{x}_j) - f(\boldsymbol{x}_i))|f(\boldsymbol{x}_j) - f(\boldsymbol{x}_i)|^{p-2} & \text{else.} \end{cases} \\ &=: \nabla \mathcal{E}^{(p)}_{n,\mathrm{con}}(f)(\boldsymbol{x}_i).\end{aligned}$$

By construction, the solution to $\nabla \mathcal{E}^{(p)}_{n,\mathrm{con}}(f) = 0$ is the minimiser of $\mathcal{E}^{(p)}_{n,\mathrm{con}}$. To find the minimiser we start with an initial guess f_0 (where f_0 agrees with the constraints), discretise time and advance via a timestep τ. Thus, at the k^{th} step:

$$f_{k+1}(\boldsymbol{x}_i) = f_k(\boldsymbol{x}_i) - \tau \nabla \mathcal{E}^{(p)}_{n,\mathrm{con}}(f_k)(\boldsymbol{x}_i) \qquad i = N+1, N+2, \dots, n.$$

Gradient descent is an important and well studied method in optimisation. Nesterov accelarated gradient descent [69] improves convergence to $\mathcal{O}(\frac{1}{k^2})$, compared to the result for standard gradient descent, which has a convergence rate of $\mathcal{O}(\frac{1}{k})$. Adaptive gradient descent methods such as ADAM [70] can further speed up convergence. Proof of convergence of ADAM for convex functions was originally provided in [70], and improvements in the proof were later provided by [71], although there is still some contention in the literature of this result [72].

However, independent of the gradient descent algorithm applied, the computation time of the minimisation problem scales as $\mathcal{O}(n^3)$. Difficulties also arise when trying to choose the correct value for ε_n. The asymptotic bounds (E.7) provide some reference for a good value to take, but it is uncertain what value will provide a good result for a particular number of samples n.

For state-of-the art calculation of the discrete p-Dirichlet minimiser, we use the Newton iteration method and homotopy discussed in [4], which reduces the computational cost to $\mathcal{O}(n^2)$. In this case the graph is connected using k nearest neighbours calculations, to avoid complications in choosing ε_n. We present the computation times for this method on our examples in fig. E.11b.

Continuum p-Dirichlet energies For the continuum p-Dirichlets, we require gradient descent on a continuum rather than on discrete data points. The associated gradient flow is found by calculating the Gateaux derivative of $\mathcal{E}_\infty^{(p)}$. For any $v \in W^{1,p}(\Omega)$,

$$\begin{aligned}\partial\mathcal{E}_\infty^{(p)}(u;v) :=& \lim_{\delta\to 0^+}\frac{1}{\delta}\left(\mathcal{E}_\infty^{(p)}(u+\delta v)-\mathcal{E}_\infty^{(p)}(u)\right),\\ =& p\sigma_\eta\int_{\partial\Omega} v|\nabla u|^{p-2}\rho^2\nabla u\cdot \mathrm{d}S - p\sigma_\eta\int_\Omega v\,\mathrm{div}(\nabla u|\nabla u|^{p-2}\rho^2)\,\mathrm{d}x.\end{aligned}$$

The minimiser will therefore satisfy

$$|\nabla u|^{p-2}\nabla u\cdot \mathrm{n}\rho^2\mathbb{1}_{\partial\Omega} - \mathrm{div}(\nabla u|\nabla u|^{p-2}\rho^2) = 0,$$

where $\mathbb{1}$ is an indicator function and n is the outward unit normal to the surface $\partial\Omega$. To find the minimiser we can perform gradient descent using

$$\frac{\partial u}{\partial t} = \begin{cases}\mathrm{div}(\nabla u|\nabla u|^{p-2}\rho^2) & \text{on } \Omega\setminus\partial\Omega,\\ -|\nabla u|^{p-2}\nabla u\cdot\mathrm{n}\rho^2 + \beta\mathrm{div}(\nabla u|\nabla u|^{p-2}\rho^2) & \text{on } \partial\Omega,\\ 0, & \text{at } \boldsymbol{x}_i, i=1,\dots,N\end{cases} \tag{E.21}$$

where β is a parameter which allows flux through the boundary $\partial\Omega$. In our simulations we take $\beta = 0.01$.

For each gradient step, we need to approximate spacial derivatives of ρ and f. We use pseudo-spectral methods [73] and domain decomposition [74] to accurately and efficiently apply gradient descent. Pseudo-spectral (or collocation) methods are popular methods in the construction of numerical solutions to PDEs. Provided the function in consideration is suitably smooth, these methods produce high precision on coarse meshes.

To gain some intuition of the pseudo-spectral methodology, we consider the one dimensional case on a periodic domain. For a function $f:\mathbb{R}\to\mathbb{R}$ discretised on a uniform grid $\{x_1,\dots,x_D\}$, where $|x_i - x_{i+1}| = h$, the finite difference approximation $g:\mathbb{R}\to\mathbb{R}$ of the derivative f' at points

x_i is determined by

$$f'(x_i) \approx g(x_i) = \frac{f(x_{i+1}) - f(x_{i-1})}{2h}$$

We can write this as a matrix multiplication, involving a sparse matrix,

$$\begin{pmatrix} g(x_1) \\ g(x_2) \\ \\ \vdots \\ \\ g(x_D) \end{pmatrix} = h^{-1} \begin{pmatrix} 0 & \frac{1}{2} & & & -\frac{1}{2} \\ -\frac{1}{2} & 0 & \ddots & & \\ & 0 & \ddots & & \\ & 0 & \ddots & 0 & \frac{1}{2} \\ \frac{1}{2} & & & -\frac{1}{2} & 0 \end{pmatrix} \begin{pmatrix} f(x_1) \\ f(x_2) \\ \\ \vdots \\ \\ f(x_D) \end{pmatrix}.$$

This approximation is then accurate to $O(h^2)$, where h is the distance between two grid points. We may also consider a finite difference approximation which includes the four nearest points,

$$f'(x_i) \approx g(x_i) = \frac{-f(x_{i+2}) + 8f(x_{i+1}) - 8f(x_{i-1}) + f(x_{i-2})}{12h}$$

which can be written as another matrix multiplication with matrix,

$$\begin{pmatrix} g(x_1) \\ g(x_2) \\ \\ \vdots \\ \\ g(x_D) \end{pmatrix} = h^{-1} \begin{pmatrix} & \ddots & & & \frac{1}{12} & -\frac{2}{3} \\ & \ddots & -\frac{1}{12} & & & \frac{1}{12} \\ & \ddots & \frac{2}{3} & \ddots & & \\ & \ddots & 0 & \ddots & & \\ & \ddots & -\frac{2}{3} & \ddots & & \\ -\frac{1}{12} & & \frac{1}{12} & \ddots & & \\ \frac{2}{3} & -\frac{1}{12} & & \ddots & & \end{pmatrix} \begin{pmatrix} f(x_1) \\ f(x_2) \\ \\ \vdots \\ \\ f(x_D) \end{pmatrix}$$

this then increases the order of convergence to $\mathcal{O}(h^4)$, but computation is more expensive, as the matrix representation of differentiation is less sparse. Spectral methods can be thought of as a limiting finite difference approximation. Under the assumption that the solution is periodic and infinitely differentiable, by considering a discretization in the Fourier domain, we can construct dense differentiation matrices in the spatial domain which can provide $\mathcal{O}(h^m)$ convergence for every m. These results can be extended to non-periodic domains by using Chebyshev polynomials instead of a Fourier basis, which for example avoids the *Runge phenomenom* (where interpolation errors increase exponentially as the number of gridpoints increases). For example, on the grid $[-1, 1]$ we use the points

$$x_i = \cos\left(\frac{i\pi}{D}\right), \quad i = 1, \ldots, D$$

Higher order derivatives and boundary conditions are also included in an intuitive manner. In higher dimensions, the differential matrices are constructed by taking Kronecker products of one-dimensional pseudo-spectral matrices. For more details and a concise introduction to pseudo-spectral methods, see [73]. For some convergence results related to non-linear PDEs and spectral methods, see [74]. In the two dimensional problems we are considering, we construct pseudospectral

differential matrices $\mathcal{D}_x$ and $\mathcal{D}_y$ using the methods provided in [73], [74]. Heuristically, once suitable pseudospectral matrices are constructed, the non-linear partial differential equation is discretised in space by replacing gradients ∂_x, ∂_y with the matrices $\mathcal{D}_x, \mathcal{D}_y$ respectively.

For a system where the pointwise constraints $\{\boldsymbol{x}_i, i = 1, \dots N\}$ are located on the boundary $\partial\Omega$ at Chebyshev gridpoints, (E.21) can be discretised using an explicit Euler method with timestep τ. Given a matrix of points defined by a Kronecker product grid of Chebyshev points $(x_i)_{i=1,\dots D}, (y_i)_{i=1,\dots D}$, we define $\boldsymbol{u}^n = (u(x_i, y_j), t^n)_{i,j=1,\dots D}$ as the discretised approximation of $u(\boldsymbol{x}, n\tau)$ at time $n\tau$ for $n \in \mathbb{N}$. Given initial condition $\boldsymbol{u}^0 = \boldsymbol{u}_0$, we then have:

$$\frac{\boldsymbol{u}^{n+1} - \boldsymbol{u}^n}{\tau} = \begin{cases} \begin{aligned} &\mathcal{D}_x(\mathcal{D}_x\boldsymbol{u}((\mathcal{D}_x\boldsymbol{u}^n)^2 + (\mathcal{D}_y\boldsymbol{u}^n)^2)^{\frac{p-2}{2}} \cdot \boldsymbol{\rho}^2) \\ &\quad + \mathcal{D}_y(\mathcal{D}_y\boldsymbol{u}^n((\mathcal{D}_x\boldsymbol{u}^n)^2 + (\mathcal{D}_y\boldsymbol{u}^n)^2)^{\frac{p-2}{2}} \cdot \boldsymbol{\rho}^2), \end{aligned} & \text{on } \Omega \setminus \partial\Omega, \\ \begin{aligned} &-\boldsymbol{\rho}^2 \cdot ((\mathcal{D}_x\boldsymbol{u}^n)^2 + (\mathcal{D}_y\boldsymbol{u}^n)^2)^{\frac{p-2}{2}}((\mathcal{D}_x\boldsymbol{u}^n)\boldsymbol{n}_x + (\mathcal{D}_x\boldsymbol{u}^n)\boldsymbol{n}_y) \\ &\quad + \beta\Big[\mathcal{D}_x(\mathcal{D}_x\boldsymbol{u}^n((\mathcal{D}_x\boldsymbol{u}^n)^2 + (\mathcal{D}_y\boldsymbol{u}^n)^2)^{\frac{p-2}{2}} \cdot \boldsymbol{\rho}^2) \\ &\qquad + \mathcal{D}_y(\mathcal{D}_y\boldsymbol{u}^n((\mathcal{D}_x\boldsymbol{u}^n)^2 + (\mathcal{D}_y\boldsymbol{u}^n)^2)^{\frac{p-2}{2}} \cdot \boldsymbol{\rho}^2)\Big], \end{aligned} & \begin{aligned} &\text{on } (x_i, y_j) \\ &\in \partial\Omega \backslash \{\boldsymbol{x}_i, i = 1, \dots N\} \end{aligned} \\ 0, & \text{on } \boldsymbol{x}_i, i = 1, \dots N. \end{cases} \tag{E.22}$$

where $\boldsymbol{n}_x, \boldsymbol{n}_y$ are matrices which are zero on $\Omega \backslash \partial\Omega$, and are the x and y components of the outward unit normal of the domain on $\partial\Omega$ respectively, $\boldsymbol{\rho} = (\rho(x_i, y_j))_{i,j=1,\dots,D}$, and $\cdot$ represents the pointwise product between two matrices.

However, sharp peaks are generally observed around constrained points in the interior of the domain, which is particularly evident for small values of p. Between each constraint we expect the function to be smooth. We therefore decompose the domain of interest, such that constraints lie on boundaries between patches of the domain, and match boundary conditions between each patch.

Inside each patch, and on the boundary of the entire domain, the system obeys (E.21). If two patches share a boundary, we need to ensure that their values match, and that the flux also matches. Thus for patches i and j which share the boundary $\partial\Omega_{ij}$, at shared points

$$\begin{aligned} u_i &= u_j, \\ \rho^2|\nabla u_i|^{p-2}\nabla u_i \cdot \mathrm{n}_i &= -\rho^2|\nabla u_j|^{p-2}\nabla u_j \cdot \mathrm{n}_j, \end{aligned} \tag{E.23}$$

where n_i is the normal from patch i to $\partial\Omega_{ij}$, and u_i is the value of the function in patch i, and similarly for n_j and patch j. Due to the shape of the grid, for the best accuracy, constraints should be placed between shared corners, however in practice this causes degeneracies due to discretization. To solve this the boundary conditions between patches that share corners can be altered to account for this, but for simplicity we place constraints shared by more than two patches near, not on, shared corners.

The additional matching conditions between patches converts the problem E.21 to a set of

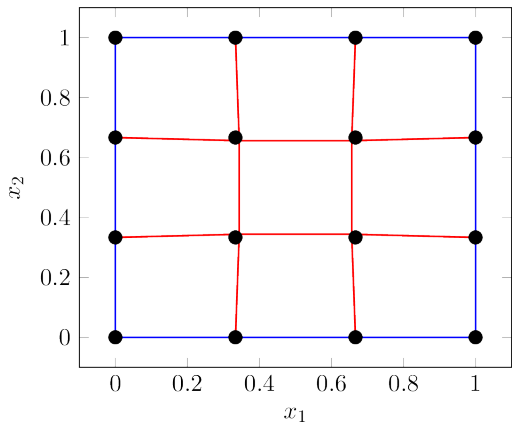


Figure E.8: Boundary patching method. The blue lines indicate the boundary of the domain, where the original boundary condition in (E.21) is used. The interior red lines are where (E.23) is used. Constraints (black points) are placed on the corners of patches, with the exception of interior corners, which are offset.

differential algebraic equations (DAEs) [75]. Our system is semi-explicit, in that defining $\Gamma = (\cup_{i,j} \partial\Omega_{ij})$ to be the boundary between patches, $g = u\lfloor_\Gamma$ as the value of u on Γ, and $f = u\lfloor_{\Omega\setminus\Gamma}$ as the value of u away from the boundaries, we can write

$$\begin{aligned}\frac{\mathrm{d}\boldsymbol{f}}{\mathrm{d}t} &= F(\boldsymbol{f}, \boldsymbol{g}),\\ 0 &= G(\boldsymbol{f}, \boldsymbol{g}).\end{aligned}$$

where F is given by the equations (E.21) restricted to $\Omega \setminus \Gamma$, and G is given by (E.23), and $\boldsymbol{f}, \boldsymbol{g}$ are the spatially discretized f and g. It is necessary to use (semi-)implicit methods on DAEs, so that variables determined by algebraic equations can be updated for each timestep. Many implicit methods are available and well studied in the literature [76], we consider the semi-implicit Euler method [77]. At step $n+1$, we solve the linear system:

$$\begin{aligned}\boldsymbol{f}^{n+1} - \boldsymbol{f}^n &= \tau F(\boldsymbol{f}^{n+1}, \boldsymbol{g}^{n+1}),\\ 0 &= G(\boldsymbol{f}^{n+1}, \boldsymbol{g}^{n+1}).\end{aligned}$$

After discretising F and G using the pseudospectral matrices $\mathcal{D}_x, \mathcal{D}_y$ in an analogous way to (E.22), numerically this involves computing a Jacobian in F and G, and solving a linear system of equations at each timestep.

Results

Figure E.9 shows an example of the minimiser of $\mathcal{E}^{(p)}_{n,\mathrm{con}}$ (Figures E.9a, E.9b and E.9c) with $n = 1500$, the minimiser of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$ (Figures E.9d, E.9e and E.9f), the minimiser of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_{n,h})$ (Figure E.9g, E.9h and E.9i) and the minimiser of the minimiser of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;S_{T,\lambda}(\rho_{n,h}))$ (Figures E.9j, E.9k and E.9l), with $n = 2^{17}$ samples for each density, and $p = 3$, where the minimisers are achieved using gradient descent methods explained above with a tolerance of 10^{-5}. We construct the KDE using $h = 0.01$, and for the SKDE we choose $T = 2^{12}$ and $\lambda = 10^{-6}$. Each patch is discretised with 100 Chebyshev points. For Figures E.9a, E.9b and E.9c we chose ε via the tuning

procedure described above.

In Figure E.10, we give the L^∞ errors between the minimisers of $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho)$ and $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_{n,h})$ or $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;S_{T,\lambda}(\rho_{n,h}))$, for different values of n. The results reflect the conclusions in Section E.5.2, the SKDE is an improvement for ρ_1 and ρ_2, but due to the large fluctuations in gradient in ρ_3, smoothing the KDE increases the associated error. Although using the SKDE in the minimisation problem improves the accuracy of the result, we note that the choices of λ and T were not necessarily optimal in these examples, as the optimal choice of parameters can lead to density estimates which have negative values, causing numerical errors during gradient descent. It is a topic of future work to consider and apply the associated smoothing spline problem for strictly positive functions in the SKDE. Finally, we present the CPU time for each gradient descent computation in Figure E.11. For low dimensional problems with large amounts of data, a continuum approach is shown to be computationally cheaper, and we note that density estimation can be parallelised to reduce computation time of the continuum method.

E.6 Conclusions and future work

We have shown that the appropriate limit is attained when using a density estimate in the constrained continuum p-Dirichlet minimisation problem, provided the density estimate converges uniformly almost surely. In addition, we have shown that the kernel density estimate meets the convergence criterion, and using smoothing splines can improve the approximation without affecting convergence. The non-local p-Dirichlet energy convergence result also provides an insight into the link between the discrete p-Dirichlet energy and density estimation in the continuum analogue.

We have also provided numerical examples using different probability densities, which show that the constrained continuum p-Dirichlet energy can be used effectively in problems with a large amount of data and low dimension.

Future improvements to the scheme include incorporating a positivity constraint in the smoothing spline calculation for more robust density estimation. We would also like to consider different density (parametric or non-parametric) estimation methods in the minimisation problem, and construct numerically stable continuum methods for when p is large.

References

[1] X. Zhu, Z. Ghahramani, and J. D. Lafferty, "Semi-supervised learning using Gaussian fields and harmonic functions," in *Proceedings of the 20th International conference on Machine learning*, 2003, pp. 912–919.

[2] D. Zhou and B. Schölkopf, "Regularization on discrete spaces," in *Joint Pattern Recognition Symposium*, 2005, pp. 361–368.

[3] D. Slepcev and M. Thorpe, "Analysis of p-Laplacian regularization in semi-supervised learning," *SIAM Journal on Mathematical Analysis*, vol. 51, no. 3, pp. 2085–2120, 2019.

[4] M. F. Rios, J. Calder, and G. Lerman, *Algorithms for ℓ_p-based semi-supervised learning on graphs*, 2019. arXiv: 1901.05031 [math.NA].

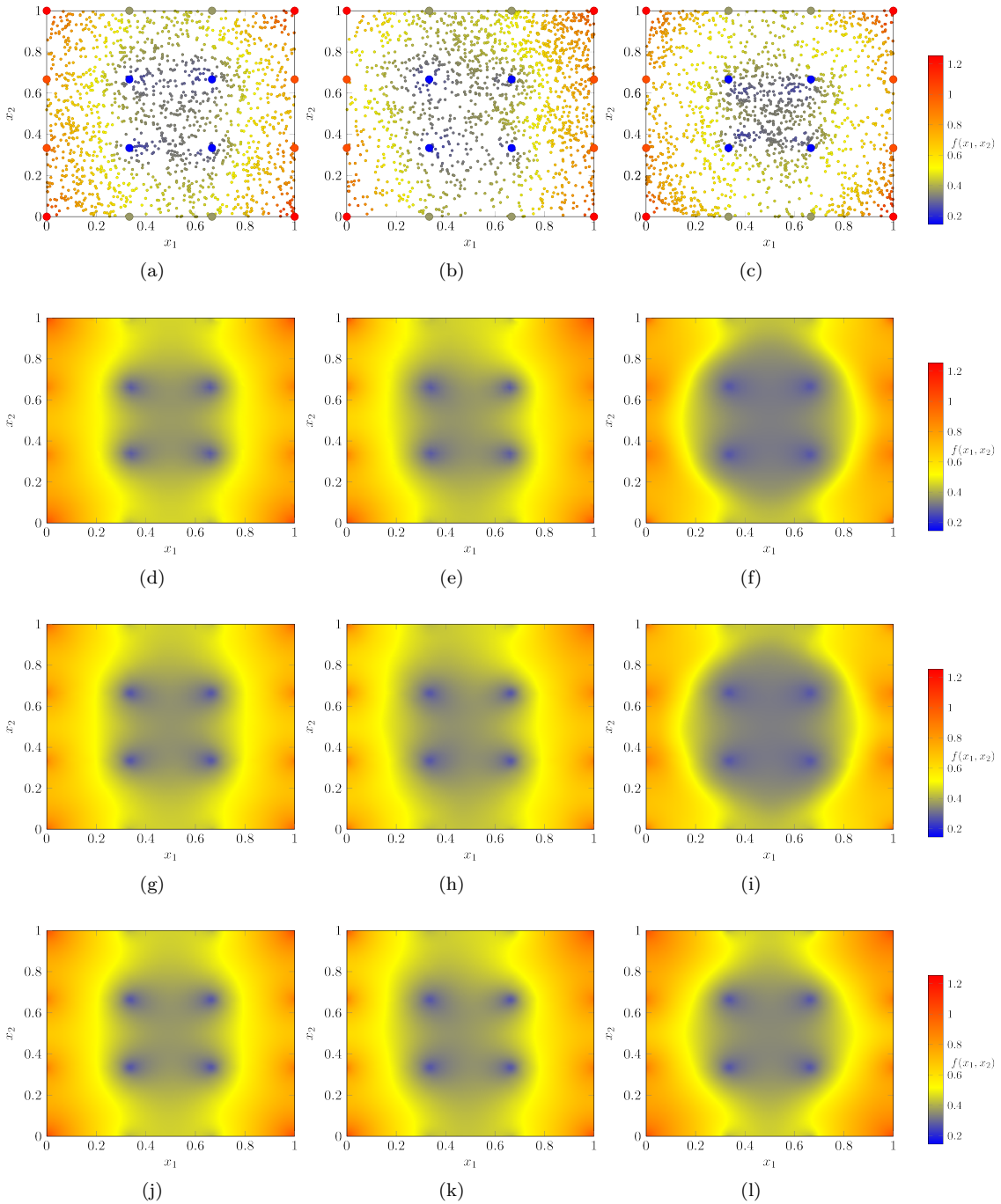


Figure E.9: Numerically constructed minimisers for $\mathcal{E}^{(p)}_{n,\mathrm{con}}$ (first row), ground truth $\mathcal{E}^{(p)}_{\infty}(\cdot;\rho)$ (second row), the KDE $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;\rho_{n,h})$ (third row), and the SKDE $\mathcal{E}^{(p)}_{\infty,\mathrm{con}}(\cdot;S_{T,\lambda}(\rho_{n,h}))$ (fourth row). The first column is for data points sampled from ρ_1, the second column for data points sampled from ρ_2, and the third column is for data points sampled from ρ_3. In these examples we have taken $p = 3$.

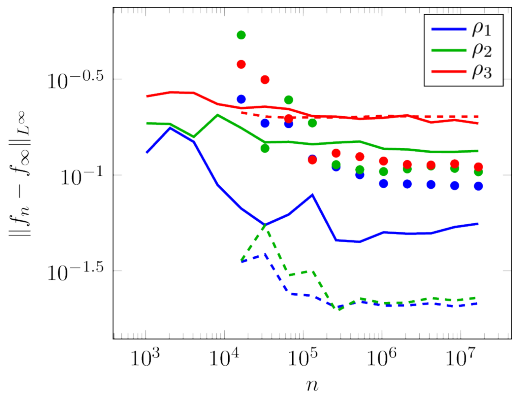


Figure E.10: L^∞ error between the computed minimisers f_n of $\mathcal{E}_\infty^{(p)}(\cdot;\rho_n)$ against the minimiser f_∞ of $\mathcal{E}_\infty^{(p)}(\cdot,\rho)$, for the KDE (dotted) and SKDE (dashed) in the p-Dirichlet energy minimisation problem, where $\lambda = 10^{-6}$ and $T = 2^{12}$ and $p = 3$, and results from the method discussed in [4] (solid).

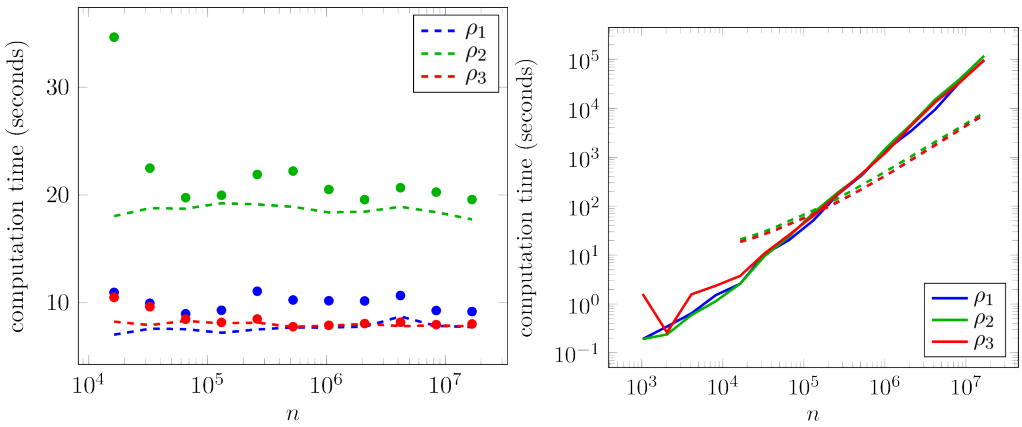


Figure E.11: Computation times for the different methods. Left: computation times for the continuum p Dirichlet energy calculation, once the density estimate is constructed. Dotted lines represent the minimisation problem using the KDE, dashed are using the SKDE, where $\lambda = 10^{-6}$ and $T = 2^{12}$ and $p = 3$. We note that computation time using the exact density is 11.03, 11.32, and 21.77 seconds respectively for ρ_1, ρ_2 and ρ_3. Right: Comparison of the continuum method presented (using SKDE to approximate the density, dashed line) in this paper and the discrete methodology discussed in [4] (solid line).

[5] E. Fix and J. L. Hodges, "Discriminatory analysis. Nonparametric discrimination: Consistency properties," *International Statistical Review / Revue Internationale de Statistique*, vol. 57, no. 3, pp. 238–247, 1989, Originally appeared as Report Number 4, Project Number 21-49-004, USAF School of Aviation Medicine, Randolph Field, Texas, in February 1951.

[6] H. Akaike, "An approximation to the density function," *Annals of the Institute of Statistical Mathematics*, vol. 6, no. 2, pp. 127–132, 1954.

[7] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *Ann. Math. Statist.*, vol. 27, no. 3, pp. 832–837, 1956. DOI: 10.1214/aoms/1177728190. [Online]. Available: https://doi.org/10.1214/aoms/1177728190.

[8] E. Parzen, "On estimation of a probability density function and mode," *The Annals of Mathematical Statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[9] H. Jiang, "Uniform convergence rates for kernel density estimation," in *Proceedings of the 34th International Conference on Machine Learning*, 2017, pp. 1694–1703.

[10] A. Rinaldo and L. Wasserman, "Generalized density clustering," *The Annals of Statistics*, vol. 38, no. 5, pp. 2678–2722, 2010.

[11] E. Giné and A. Guillou, "Rates of strong uniform consistency for multivariate kernel density estimators," *Annales de l'Institut Henri Poincare (B) Probability and Statistics*, vol. 38, no. 6, pp. 907–921, 2002, ISSN: 0246-0203.

[12] B. W. Silverman, "Weak and strong uniform consistency of the kernel estimate of a density and its derivatives," *The Annals of Statistics*, vol. 6, no. 1, pp. 177–184, 1978, ISSN: 00905364.

[13] W. Stute, "The oscillation behavior of empirical processes: The multivariate case," *The Annals of Probability*, pp. 361–379, 1984.

[14] E. A. Nadaraya, "On non-parametric estimates of density functions and regression curves," *Theory of Probability & Its Applications*, vol. 10, no. 1, pp. 186–190, 1965.

[15] A. Tsybakov, *Introduction to Nonparametric Estimation*. Springer, 2009.

[16] L. Devroye and G. Lugosi, *Combinatorial Methods on Density Estimation*. Springer, 2001.

[17] P. Whittle, "On the smoothing of probability density functions," *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 334–343, 1958.

[18] G. Wahba, *Spline models for observational data*. Philadelphia: Society for Industrial and Applied Mathematics (SIAM), 1990.

[19] N. Bissantz, T. Hohage, and A. Munk, "Consistency and rates of convergence of nonlinear Tikhonov regularization with random noise," *Inverse Problems*, vol. 20, no. 6, pp. 1773–1789, 2004.

[20] N. Bissantz, T. Hohage, A. Munk, and F. Ruymgaart, "Convergence rates of general regularization methods for statistical inverse problems and applications," *SIAM Journal on Numerical Analysis*, vol. 45, no. 6, pp. 2610–2636, 2007.

[21] G. Claeskens, T. Krivobokova, and J. D. Opsomer, "Asymptotic properties of penalized spline estimators," *Biometrika*, vol. 96, no. 3, pp. 529–544, 2009.

[22] P. Hall and J. D. Opsomer, "Theory for penalised spline regression," *Biometrika*, vol. 92, no. 1, pp. 105–118, 2005.

[23] G. Kauermann, T. Krivobokova, and L. Fahrmeir, "Some asymptotic results on generalized penalized spline smoothing," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, vol. 71, no. 2, pp. 487–503, 2009.

[24] M.-J. Lai and L. Wang, "Bivariate penalized splines for regression," *Statistica Sinica*, vol. 23, pp. 1399–1417, 2013.

[25] M. A. Lukas, "Robust generalized cross-validation for choosing the regularization parameter," *Inverse Problems*, vol. 22, no. 5, pp. 1883–1902, 2006.

[26] X. Wang, J. Shen, and D. Ruppert, "On the asymptotics of penalized spline smoothing," *Electronic Journal of Statistics*, vol. 5, pp. 1–17, 2011.

[27] R. Arcangeli and B. Ycart, "Almost sure convergence of smoothing D^m-splines for noisy data," *Numerische Mathematik*, vol. 66, no. 1, pp. 281–294, 1993.

[28] G. Wahba, "A comparison of GCV and GML for choosing the smoothing parameter in the generalized spline smoothing problem," *The Annals of Statistics*, vol. 13, no. 4, pp. 1378–1402, 1985.

[29] G. S. Kimeldorf and G. Wahba, "A correspondence between Bayesian estimation on stochastic processes and smoothing by splines," *The Annals of Mathematical Statistics*, vol. 41, no. 2, pp. 495–502, 1970.

[30] D. D. Cox, "Approximation of method of regularization estimators," *The Annals of Statistics*, vol. 16, no. 2, pp. 694–712, 1988.

[31] D. W. Nychka and D. D. Cox, "Convergence rates for regularized solutions of integral equations from discrete noisy data," *The Annals of Statistics*, vol. 17, no. 2, pp. 556–572, 1989.

[32] R. J. Carroll, A. C. M. Van Rooij, and F. H. Ruymgaart, "Theoretical aspects of ill-posed problems in statistics," *Acta Applicandae Mathematica*, vol. 24, no. 2, pp. 113–140, 1991.

[33] B. A. Mair and F. H. Ruymgaart, "Statistical inverse estimation in Hilbert scales," *SIAM Journal on Applied Mathematics*, vol. 56, no. 5, pp. 1424–1444, 1996.

[34] Y. Li and D. Ruppert, "On the asymptotics of penalized splines," *Biometrika*, vol. 95, no. 2, pp. 415–436, 2008.

[35] J. Shen and X. Wang, "Estimation of monotone functions via P-splines: A constrained dynamical optimization approach," *SIAM Journal on Control and Optimization*, vol. 49, no. 2, pp. 646–671, 2011.

[36] L. Xiao, Y. Li, T. V. Apanasovich, and D. Ruppert, *Local asymptotics of p-splines*, 2012. arXiv: 1201.0708 [math.ST].

[37] T. Yoshida and K. Naito, "Asymptotics for penalized additive B-spline regression," *Journal of the Japan Statistical Society*, vol. 42, no. 1, pp. 81–107, 2012.

[38] ——, "Asymptotics for penalised splines in generalised additive models," *Journal of Nonparametric Statistics*, vol. 26, no. 2, pp. 269–289, 2014.

[39] M. Thorpe and A. M. Johansen, "Pointwise convergence in probability of general smoothing splines," *Annals of the Institute of Statistical Mathematics*, Apr. 2017, ISSN: 1572-9052.

[40] B. Nadler, N. Srebro, and X. Zhou, "Statistical analysis of semi-supervised learning: The limit of infinite unlabelled data," in *Advances in Neural Information Processing Systems (NIPS)*, 2009, pp. 1330–1338.

[41] M. Alamgir and U. Von Luxburg, "Phase transition in the family of p-resistances," in *Advances in Neural Information Processing Systems (NIPS)*, 2011, pp. 379–387.

[42] A. El Alaoui, X. Cheng, A. Ramdas, M. J. Wainwright, and M. I. Jordan, "Asymptotic behavior of ℓ_p-based laplacian regularization in semi-supervised learning," in *Conference on Learning Theory*, 2016, pp. 879–906.

[43] M. Belkin and P. Niyogi, "Convergence of Laplacian eigenmaps," in *Advances in Neural Information Processing Systems (NIPS)*, 2007, pp. 129–136.

[44] R. R. Coifman and S. Lafon, "Diffusion maps," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 5–30, 2006, ISSN: 1063-5203. DOI: 10.1016/j.acha.2006.04.006. [Online]. Available: http://dx.doi.org/10.1016/j.acha.2006.04.006.

[45] E. Giné and V. Koltchinskii, "Empirical graph Laplacian approximation of Laplace-Beltrami operators: Large sample results," in *High dimensional probability*, ser. IMS Lecture Notes Monograph Series, vol. 51, Beachwood, OH: Institute of Mathematical Statistics, 2006, pp. 238–259.

[46] M. Hein, J.-Y. Audibert, and U. von Luxburg, "From graphs to manifolds–weak and strong pointwise consistency of graph Laplacians," in *Learning theory*, Springer, 2005, pp. 470–485.

[47] M. Hein, "Uniform convergence of adaptive graph-based regularization," in *International Conference on Computational Learning Theory*, 2006, pp. 50–64.

[48] A. Singer, "From graph to manifold Laplacian: The convergence rate," *Applied and Computational Harmonic Analysis*, vol. 21, no. 1, pp. 128–134, 2006.

[49] D. Ting, L. Huang, and M. I. Jordan, "An analysis of the convergence of graph Laplacians," in *Proceedings of the 27th International Conference on Machine Learning*, 2010.

[50] U. von Luxburg, M. Belkin, and O. Bousquet, "Consistency of spectral clustering," *The Annals of Statistics*, vol. 36, no. 2, pp. 555–586, 2008.

[51] B. Pelletier and P. Pudlo, "Operator norm convergence of spectral clustering on level sets," *Journal of Machine Learning Research*, vol. 12, pp. 385–416, 2011, ISSN: 1532-4435.

[52] A. Singer and H.-T. Wu, "Spectral convergence of the connection Laplacian from random samples," *Information and Inference: A Journal of the IMA*, vol. 6, no. 1, pp. 58–123, 2017.

[53] N. García Trillos and D. Slepčev, "A variational approach to the consistency of spectral clustering," *Applied and Computational Harmonic Analysis*, 2016.

[54] D. Burago, S. Ivanov, and Y. Kurylev, "A graph discretization of the Laplace-Beltrami operator," *Journal of Spectral Theory*, vol. 4, no. 4, pp. 675–714, 2014, ISSN: 1664-039X.

[55] X. Wang, *Spectral convergence rate of graph laplacian*, 2015. arXiv: 1510.08110 [stat.ML].

[56] N. G. Trillos, M. Gerlach, M. Hein, and D. Slepčev, "Error estimates for spectral convergence of the graph laplacian on random geometric graphs toward the laplace–beltrami operator," *Foundations of Computational Mathematics*, pp. 1–61, 2019.

[57] N. García Trillos and D. Slepčev, "Continuum limit of total variation on point clouds," *Archive for Rational Mechanics and Analysis*, vol. 220, no. 1, pp. 193–241, 2016.

[58] J. Calder, "Consistency of lipschitz learning with infinite unlabeled data and finite labeled data," *SIAM Journal on Mathematics of Data Science*, vol. 1, no. 4, pp. 780–812, 2019.

[59] ——, "The game theoretic p-Laplacian and semi-supervised learning with few labels," *Nonlinearity*, vol. 32, no. 1, p. 301, 2018.

[60] R. Cristoferi and M. Thorpe, "Large data limit for a phase transition model with the p-laplacian on point clouds," *European Journal of Applied Mathematics*, pp. 1–47,

[61] Q. Li and J. Racine, *Nonparametric Econometrics: Theory and Practice*. Princeton University Press, Jan. 2007.

[62] B. E. Hansen, "Lecture notes on nonparametrics," University of Wisconsin-Madison, 2009.

[63] F. I. Utreras, "Convergence rates for multivariate smoothing spline functions," *Journal of Approximation Theory*, vol. 52, no. 1, pp. 1–27, 1988, ISSN: 0021-9045.

[64] A. Braides, *Gamma-convergence for Beginners*, ser. Oxford Lecture Series in Mathematics. Oxford University Press, 2002.

[65] G. Dal Maso, *An Introduction to Γ-Convergence*. Springer, 1993.

[66] G. Leoni, *A First Course in Sobolev Spaces*. American Mathematical Society, 2009, vol. 105.

[67] B. D. Goddard, A. Nold, and S. Kalliadasis, *2DChebClass [Software]*, 2017. [Online]. Available: dx.doi.org/10.7488/ds/1991.

[68] T. Ursell, *pinky [Software]*, MathWorks File Exchange, https://uk.mathworks.com/matlabcentral/fileexchange/35797-generate-random-numbers-from-a-2d-discrete-distribution, 2016.

[69] W. Su, S. Boyd, and E. J. Candès, "A differential equation for modeling Nesterov's accelerated gradient method: Theory and insights," *Journal of Machine Learning Research*, vol. 17, no. 153, pp. 1–43, 2016.

[70] D. P. Kingma and J. Ba, *Adam: A method for stochastic optimization*, 2014. arXiv: 1412.6980 [cs.LG].

[71] S. Bock, J. Goppold, and M. Weiß, "An improvement of the convergence proof of the ADAM-Optimizer," *arXiv e-prints*, Apr. 2018. arXiv: 1804.10587.

[72] S. J. Reddi, S. Kale, and S. Kumar, "On the convergence of adam and beyond," in *International Conference on Learning Representations*, 2018.

[73] L. Trefethen, *Spectral Methods in MATLAB*. Society for Industrial and Applied Mathematics, 2000.

[74] J. Boyd, *Chebyshev and Fourier Spectral Methods: Second Revised Edition*, ser. Dover Books on Mathematics. Dover Publications, 2001.

[75] P. Kunkel, V. Mehrmann, and V. L. Mehrmann, *Differential-algebraic Equations: Analysis and Numerical Solution*, ser. EMS textbooks in mathematics. European Mathematical Society, 2006.

[76] E. Hairer and G. Wanner, *Solving Ordinary Differential Equations II: Stiff and Differential-Algebraic Problems*, ser. Springer series in computational mathematics. Springer-Verlag, 1991.

[77] P. Deuflhard, E. Hairer, and J. Zugck, "One-step and extrapolation methods for differential-algebraic systems," *Numerische Mathematik*, vol. 51, no. 5, pp. 501–516, Sep. 1987, ISSN: 0945-3245.

Appendix F

A linear transportation L^p distance for pattern recognition

The transportation L^p distance, denoted TL^p, has been proposed as a generalisation of optimal transport motivated by the property that it can be applied directly to colour or multi-channelled images, as well as multivariate time-series without normalisation or mass constraints. These distances, as with optimal transport, are powerful tools in modelling data with spatial or temporal perturbations. However, their computational cost can make them infeasible to apply to even moderate pattern recognition tasks. We propose linear versions of these distances and show that the linear TL^p distance significantly improves over linear optimal transport on signal processing tasks. Applications to real-world datasets shows vastly improved performance over linear optimal transport. We also demonstrate that the linear TL^p distance is several orders of magnitude faster to compute than the TL^p distance, but with competitive performance on the classification task.

F.1 Introduction

Optimal transport has gained in recent popularity because of its ability to model diverse data distributions in the signal and image processing fields [1]. Transportation-based methods have been successfully applied to image analysis, including medical images [2] and facial recognition [3], as well as cosmology [4], [5] and voice recognition. Machine learning and Bayesian statistics have also benefited from transport-based approaches [6]–[11].

The popularity of optimal transport is, in part, due to the rise in the number of problems in the experimental and social sciences in which techniques are required to compare signal perturbations across spatial or temporal domains. For example, optimal transport based methods for image registration and warping [12] and image morphing [13] have existed for many years. Transportation techniques provide non-linear methods that jointly model locations and intensities, making transportation based approaches a powerful tool in many problems.

Optimal transport methods, in particular Wasserstein distances, are grounded in a wealth of mathematical theory and excellent introductions to the advanced mathematical theory of optimal transport are presented in [14], [15], whilst [16] presents the theory with a more applied perspective. Many technical aspects of optimal transport have been explored, including its geometric properties

[17] and its links to evolutionary PDEs [18].

There is much interest in developing efficient methods to compute optimal transport distances and maps. For discrete measures the optimal transport problem can be solved using linear programming approaches [16]. Since solving a linear programme can be costly, [19] proposed a multi-scale linear programme for efficient computations. Other approaches include flow minimisation techniques [12], [18], [20], [21] and gradient descent approaches [22], as well as multi-scale methods [23], [24]. More recently, Cuturi proposed entropy-regularisation based approaches to compute approximations of the optimal transport problem [25]. These methods have been explored in-depth with many extensions [8], [26]–[28]. In order to efficiently compute pairwise optimal transport distances on a large data set a framework called linear optimal transport was proposed in [29], [30].

Higher-order transportation methods have been proposed in the mathematical analysis literature [31], [32]. In [31] García Trillos and Slepčev proposed the transportation L^p (TL^p) distance to define discrete-to-continuum convergence of variational problems on point clouds. This was further extended in [32] to a transportation $\mathrm{W}^{k,p}$ distance (where the notation relates to Sobolev spaces). These transportation distances have several key advantages over the optimal transport distance; these include [33]:

1. They have no need for mass normalisation.
2. They are not restricted to non-negative measures.
3. They are sensitive to high frequency perturbations.
4. They can track translations further than L^p.
5. They can compare signals with different discretizations.

Additionally the transportation $\mathrm{W}^{k,p}$ distance can include information on the derivative of the signal. The transportation distances were proposed in [33] to tackle problems in signal analysis; for example, they were able to apply transportation methods to colour images. They experimentally showed that the TL^p distance out performed both L^p distances and optimal transport distances in classification tasks. However, these higher-order transportation distances require the computation of an optimal transport map (on a higher dimensional space) and this thwarts its application to larger scale pattern recognition tasks, where pairwise distances are needed. We propose linear transportation distances to take advantage of the flexibility of the TL^p (and transportation $\mathrm{W}^{k,p}$) distances, whilst alleviating computation costs.

Our proposed method can be seen as an extension of the linear optimal transport framework [29] to higher order transportation distances. Suppose we wish to compare N signals/images, then application of optimal transport methods would require computation of $N(N-1)/2$ distances. In the linear optimal transport framework only N distances need to be computed. From here signals/images are then embedded into Euclidean space allowing linear statistical methods to be applied, whilst preserving much of the geometry of the original optimal transport space.

This manuscript begins by reviewing optimal transport and the TL^p distance. The linear optimal transport framework is reviewed in appendix F.2.3 and we propose our extension to TL^p in appendix F.2.4. We include further sections (see appendices F.2.5 and F.2.7) on computing the linear TL^p distance, including derivatives in the distance, and geodesics and interpolation

in the linear TL^p space. In appendices F.3.1 and F.3.2 we apply our method to two synthetic datasets and demonstrate that linear TL^p improves over the linear Wasserstein distance. We then apply our method to classification problems in cell morphometry (appendix F.3.3), Australian sign language (appendix F.3.4), breast cancer histopathology (appendix F.3.5) and financial time series (appendix F.3.6), and show that linear TL^p outperforms the linear Wasserstein distance and that it has similar performance to TL^p, but is several orders of magnitude faster.

F.2 Methods

F.2.1 Optimal transport

To fix notation, we review the modern Monge-Kantorovich formulation of optimal transport and we refer to the excellent monographs [14], [15] for a thorough exposition. Let μ and ν be probability measures on measure spaces X and Y, respectively. Further, let $\mathcal{B}(X)$ denote the Borel σ-algebra on X. We define the pushforward of a measure $\mu \in \mathcal{P}(X)$ by a function $h : X \to Z$ by $h_*\mu(A) := \mu(h^{-1}(A))$ for all $A \in \mathcal{B}(Z)$. The inverse of h is understood as being in the set theoretic sense, i.e. $h^{-1}(A) := \{x : h(x) \in A\}$. We denote by $\Pi(\mu, \nu)$ the set of all measures on $X \times Y$ such that the first marginal is μ and the second marginal is ν. To be precise, if $P^X : X \times Y \to X$ and $P^Y : X \times Y \to Y$ are the canonical projections then $P^X_*\pi = \mu$ and $P^Y_*\pi = \nu$. We call any $\pi \in \Pi(\mu, \nu)$ a transportation plan between μ and ν (also called a coupling between μ and ν).

The Kantorovich optimal transport problem is the following variational problem

$$K(\mu, \nu) = \inf_{\pi \in \Pi(\mu,\nu)} \int_{X \times Y} c(x, y) \, \mathrm{d}\pi(x, y), \tag{F.1}$$

where $c(x, y)$ is a cost function. The minimiser of this problem is called the optimal transport plan $\pi^\dagger$ and such a minimiser exists when c is lower semi-continuous (see, for example [16]). The prototypical example for c (when $X = Y = \mathbb{R}^d$) is $c(x, y) = |x - y|_p^p := \sum_{i=1}^d |x_i - y_i|^p$, for this choice of c one can define the Wasserstein distance by $d_{\mathrm{W}^p}(\mu, \nu) = \sqrt[p]{K(\mu, \nu)}$ (see also (F.3) below). When c is a metric then (F.1) is also known as the earth mover's distance.

Now, considering a different formulation, let $T : X \to Y$ be a Borel measurable function such that $T_*\mu = \nu$. The Monge optimal transport problem is to solve

$$M(\mu, \nu) = \inf_{T \,:\, T_*\mu = \nu} \int_{X \times Y} c(x, T(x)) \, \mathrm{d}\mu(x) \tag{F.2}$$

We call any T that satisfies $T_*\mu = \nu$ a transport map between μ and ν, and the solution to the optimisation problem $T^\dagger$ is called the optimal transport map.

It is worthwhile noting that the formulation of the optimal transport problems in equations (F.1) and (F.2) are not, in general, equivalent. However, if the optimal transport plan $\pi^\dagger$ can be written in the form $\pi^\dagger = (\mathrm{Id} \times T^\dagger)_*\mu$ then it follows that $T^\dagger$ is an optimal transport map and the two formulations are equivalent, i.e. $K(\mu, \nu) = M(\mu, \nu)$. A sufficient condition to show that such an optimal transport plan exists is to require that μ is absolutely continuous with respect to the Lebesgue measure on a compact domain $\Omega \subset \mathbb{R}^d$ and, in addition, $c(x, y) = h(x - y)$ where

$h : \Omega \to [0, \infty)$ is strictly convex and superlinear, see [14] Theorem 2.44. Note that it is easy to find examples where there do not exist transport maps at all. For example, if $\mu = \frac{1}{3}\delta_{x_1} + \frac{1}{3}\delta_{x_2} + \frac{1}{3}\delta_{x_3}$ and $\nu = \frac{1}{2}\delta_{y_1} + \frac{1}{2}\delta_{y_2}$.

Let $\Omega \subset \mathbb{R}^n$ and $\mathcal{P}_p(\Omega)$ be the set of Radon measures on Ω with finite p^{th} moment. For $p \in [1, \infty)$, we define the Wasserstein distance between μ and ν in $\mathcal{P}_p(\Omega)$ by

$$d_{\mathrm{W}^p}(\mu, \nu) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\Omega \times \Omega} |x - y|_p^p \, \mathrm{d}\pi(x, y) \right)^{1/p}. \tag{F.3}$$

For the case $p = \infty$, we can define a distance on $\mathcal{P}_\infty(\Omega)$ by

$$d_{\mathrm{W}^\infty}(\mu, \nu) := \inf_{\pi \in \Pi(\mu,\nu)} \operatorname{esssup}_\pi \{|x - y| : (x, y) \in \Omega \times \Omega\}.$$

We briefly review the features that make optimal transport particularly suited to signal and image processing. For an extended survey of these ideas see [1]. Optimal transport is able to provide generative models which can represent diverse data distributions and can capture signal variations as a result of spatial perturbations. Furthermore, there is a well formulated theoretical basis with interesting geometrical properties, such as existence of minimisers [15], [34], the Riemannian structure of Wasserstein spaces when $p = 2$ [18], [35] and characterisation as the weak* convergence when Ω is compact [16]. The Riemannian structure allows the characterisation of geodesics (shortest curves) on the space $\mathcal{P}_2(\Omega)$. In addition, there are many methods to compute the optimal transport distance, we review a selection in appendix F.2.5 in the context of computing the linear TL^p distance.

F.2.2 The transportation L^p distance

The TL^p distance was first introduced in [31] to define a discrete-to-continuum convergence on point clouds. This tool has been been extensively used to study similar statistical problems, e.g. [36]–[48], and recently has been shown to be a valuable tool in signal analysis [33], [49]. Further properties of this metric have been established in [32]. In this section, we review the definitions and properties of the TL^p distance and space.

Given an open and bounded domain $\Omega \subset \mathbb{R}^d$ we define the TL^p space as the set of pairs (μ, f) such that $f \in \mathrm{L}^p(\mu; \mathbb{R}^m)$ and $\mu \in \mathcal{P}_p(\Omega)$. We do not make any assumption on the dimension m of the range of f. Working in this formal and abstract framework of measure theory allows us to formulate our methods for both discrete and continuous signals, simultaneously. Importantly, this framework allows us to compare signals with different discretizations. Thus, we define the TL^p space as

$$\mathrm{TL}^p := \{(\mu, f) : \mu \in \mathcal{P}_p(\Omega), f \in \mathrm{L}^p(\mu; \mathbb{R}^m)\}.$$

We construct the TL^p distance between pairs $(\mu, f) \in \mathrm{TL}^p$ and $(\nu, g) \in \mathrm{TL}^p$ as follows:

$$d_{\mathrm{TL}^p}((\mu, f), (\nu, g)) = \left(\inf_{\pi \in \Pi(\mu,\nu)} \int_{\Omega \times \Omega} |x - y|_p^p + |f(x) - g(y)|_p^p \, \mathrm{d}\pi(x, y) \right)^{1/p}. \tag{F.4}$$

Intuitively, we see that TL^p optimal transport plans (that is plans in $\Pi(\mu,\nu)$ that achieve the minimum in the above variational problem) strike a balance between matching spatially, i.e. minimising $\int_{\Omega\times\Omega}|x-y|_p^p\,\mathrm{d}\pi(x,y)$, and matching signal features, i.e. minimising $\int_{\Omega\times\Omega}|f(x)-g(y)|_p^p\,\mathrm{d}\pi(x,y)$.

Example 57. Later we consider an application to synthetic images (see appendix F.3.2), let us explain here how images can be represented in TL^p. Let $\{x_i\}_{i=1}^n$ be the location of pixels (which usually form a grid over $[0,1]\times[0,1]$). We apply TL^p by choosing a base measure μ, this is commonly the uniform measure over $\{x_i\}_{i=1}^n$, i.e. $\mu=\frac{1}{n}\sum_{i=1}^n\delta_{x_i}$. An image is then represented by the pair (μ,f) where $f:\{x_i\}_{i=1}^n\to\mathbb{R}^3$ for RGB images and $f(x_i)$ are the RGB values for the pixel at location x_i. Similarly, for greyscale images one would have $f:\{x_i\}_{i=1}^n\to\mathbb{R}$ where $f(x_i)$ is now the greyscale value for the pixel at location x_i. Of course, one can make different choices for μ in order to emphasise regions/features of the images.

To understand the TL^p distance, we reformulate it as a Wasserstein distance supported on the graphs of functions. Recall that the graph of a function is defined as:

$$\mathrm{Gra}(f):=\{(x,f(x)):x\in\Omega\}\,.$$

Note that the $\mathrm{Gra}(f)\subset\Lambda:=\Omega\times\mathbb{R}^m$. We define the following *lifted measure* on $\mathrm{Gra}(f)$:

$$\tilde{\mu}(A\times B)=(\mathrm{Id}\times f)_*\mu(A\times B)=\mu(\{x:x\in A,f(x)\in B\}),$$

where $A\times B\subset\Lambda$. It is clear that $\tilde{\mu}$ is a well-defined measure on $\mathrm{Gra}(f)$. We can characterise the TL^p distance as a Wasserstein distance in the following way [31]:

$$\begin{aligned}
d^p_{\mathrm{TL}^p}((\mu,f),(\nu,g)) &= \inf_{\pi\in\Pi(\mu,\nu)}\int_{\Omega\times\Omega}|x-y|_p^p+|f(x)-g(y)|_p^p\,\mathrm{d}\pi(x,y)\\
&= \inf_{\tilde{\pi}\in\tilde{\Pi}(\tilde{\mu},\tilde{\nu})}\int_{\Lambda\times\Lambda}|\boldsymbol{x}-\boldsymbol{y}|_p^p\,\mathrm{d}\tilde{\pi}(\boldsymbol{x},\boldsymbol{y})\\
&= d^p_{\mathrm{W}^p}(\tilde{\mu},\tilde{\nu}).
\end{aligned}$$

Thus, we can see that the TL^p distance is the Wasserstein distance between the appropriate measures on the graphs of function. This allows us to make the following identification between Wasserstein spaces and TL^p through the mapping.

$$\mathrm{TL}^p\to\mathcal{P}_p(\Omega\times\mathbb{R}^m) \tag{F.5}$$

$$(\mu,f)\mapsto\tilde{\mu}=(\mathrm{Id}\times f)_*\mu. \tag{F.6}$$

This connection of the TL^p distance and the OT distance facilitates the transfer of certain Wasserstein properties to the TL^p setting; for example, metric properties and existence of minimisers.

It is easy to see that, for any $\mu,\nu\in\mathcal{P}_p(\Omega)$ and $f,g\in L^p(\mu)$:

$$\begin{aligned}
d_{\mathrm{TL}^p}((\mu,f),(\mu,g))&\le\|f-g\|_{L^p(\mu)}\\
d_{\mathrm{W}^p}(\mu,\nu)&\le d_{\mathrm{TL}^p}((\mu,\mathbb{1}),(\nu,\mathbb{1})).
\end{aligned}$$

In fact, one can also prove the converse inequalities (up to a constant) and hence TL^p can be seen to generalise both weak* convergence of measures and L^p convergence of functions (see [31] or proposition 58 below).

We can further connect TL^p and optimal transport by observing that the Kantorovich optimal transport problem between two measures μ and ν with cost function $c(x, y; f, g) = |x-y|_p^p + |f(x) - g(y)|_p^p$ is equal to the TL^p distance between the two signals (μ, f) and (ν, g).

For reference, we also state a Monge-type formulation of the TL^p distance as follows:

$$d^p_{\mathrm{TL}^p}((\mu, f), (\nu, g)) = \inf_{T:T_*\mu=\nu} \int_\Omega |x - T(x)|_p^p + |f(x) - g(T(x))|_p^p \,\mathrm{d}\mu(x), \tag{F.7}$$

where T is a transportation map. When we write the Monge formulation of optimal transport we are assuming that there is an equivalence between (F.4) and (F.7). This is in general difficult to verify since the the application of Brenier's theorem does not lead to natural conditions. (Assuming that μ does not give mass to small sets and both μ and ν have a sufficient number of bounded moments then one can apply Brenier's theorem to $K(\mu, \nu)$ where $c(x, y) = |x - y|_p^p + |f(x) - g(y)|_p^p$ and K is defined by (F.1), if c is strictly convex; practically this is not reasonable.) However, when μ and ν are discrete uniform measures with supports of equal size the Monge formulation (F.7) coincides with the Kantorovich formulation (F.4) (see the proposition below).

Let us recall the identification (eqs. (F.5) and (F.6)) and the identity $d_{\mathrm{W}^p}(\tilde{\mu}, \tilde{\nu}) = d_{\mathrm{TL}^p}((\mu, f), (\nu, g))$. Then, there is a corresponding equivalence between transport maps. That is (assuming all transport maps exist and are unique) let $T^\dagger$ achieve the minimum in eq. (F.7) and $\tilde{T}^\dagger$ achieve the minimum in $M(\tilde{\mu}, \tilde{\nu})$ where M is given by eq. (F.2) with $c(x, y) = |x - y|_p^p$. It follows that $\tilde{T}^\dagger(\boldsymbol{x}) = (T^\dagger(x), g(T^\dagger(x)))$ for μ-almost every $x \in \Omega$.

Proposition 58. Let $\Omega \subset \mathbb{R}^d$ be open and bounded and $p \in (1, \infty)$. Then the following holds

1. [31, Remark 3.4] $(\mathrm{TL}^p, d_{\mathrm{TL}^p})$ is a metric space;
2. [16, Theorem 5.10] $\mu_n \overset{*}{\rightharpoonup} \mu$ if and only if $(\mu_n, \mathbb{1}) \overset{\mathrm{TL}^p}{\to} (\mu, \mathbb{1})$;
3. [31, Proposition 3.12] $f_n \to f$ in $\mathrm{L}^p(\mu)$ if and only if $(\mu, f_n) \overset{\mathrm{TL}^p}{\to} (\mu, f)$;
4. [32, Theorem 2.2] for any $(\mu, f), (\nu, g) \in \mathrm{TL}^p$ there exists a transport plan $\pi^\dagger \in \Pi(\mu, \nu)$ realising the minimum in $d_{\mathrm{TL}^p}((\mu, f), (\nu, g))$, i.e.

$$d^p_{\mathrm{TL}^p}((\mu, f), (\nu, g)) = \int_{\Omega\times\Omega} |x - y|_p^p + |f(x) - g(y)|_p^p \,\mathrm{d}\pi^\dagger(x, y);$$

5. [33, Proposition 3.5] if $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $\nu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$ then for any $f \in L^p(\mu)$ and $g \in L^p(\nu)$ there exists $T^\dagger : \{x_i\}_{i=1}^n \to \{y_j\}_{j=1}^n$ such that $T^\dagger_*\mu = \nu$ and

$$d^p_{\mathrm{TL}^p}((\mu, f), (\nu, g)) = \int_\Omega |x - T^\dagger(x)|_p^p + |f(x) - g(T^\dagger(x))|_p^p \,\mathrm{d}\mu(x),$$

   i.e. the Monge and Kantorovich formulations of TL^p (given by (F.7) and (F.4) respectively) are equivalent.

Note that not all properties of Wasserstein spaces carry through to TL^p. For example $(\mathrm{TL}^p, d_{\mathrm{TL}^p})$ is not complete. Indeed, following [31], let $\Omega = (0, 1)$ and note that $f_{n+1}(x) = \operatorname{sign}\sin(2^n \pi x)$,

$\mu_n = \mathcal{L}\lfloor_{(0,1)}$ (the Lebesgue measure on $(0,1)$) is a Cauchy sequence in $(\mathrm{TL}^p, d_{\mathrm{TL}^p})$. However, $\{f_n\}$ does not converge in L^p and therefore $\{(\mu_n, f_n)\}$ cannot converge in TL^p, by part 3 of the above proposition. The completion of TL^p can be identified with the set of Young measures, and therefore the space $\mathcal{P}(\Omega \times \mathbb{R}^m)$, see [31, Remark 3.6] or [32, Theorem 5.6]. We note also that there do not exist geodesics in TL^p.

Now that we have established the theory required, we compare properties relevant to pattern recognition of the TL^p distance with the Wasserstein and L^p distances.

Properties of the TL^p distance in pattern recognition

Mass re-normalisation. Wasserstein distances require measures to be non-negative and have equal mass integrating to unity. If a signal is negative then re-normalisation must be applied to make it non-negative. In the Wasserstein framework this results in signal compression and reduces sensitivity to translations. On the other hand, the L^p and TL^p distances are invariant to addition by a positive constant and thus signals can be made positive without fear of reduced sensitivity of the distance.

High frequency perturbations. Consider $f = g + A\chi$, where $A\chi$ is a high frequency perturbation with length scale ω and amplitude A. Each x, in the Monge formulation of the Wasserstein distance, is moved a distance which is on the order of the wavelength ω of χ. This distance is small and independent of the amplitude A. The TL^p distance inherits sensitivity to high frequency perturbations through the L^p norm and the distance scales linearly with the amplitude A.

Tracking translations. Let $f = c\mathbb{1}_{[0,1]}$ be a constant multiple of the indicator function on $[0,1]$. Let us perturb f by distance ℓ; that is, let $g = f(x - \ell)$. For any $\ell > 1$, the L^p distance cannot detect further translations of f. However, the Wasserstein and TL^p distance can continue to track translations of f for all ℓ.

The TL^p appears an excellent tool to exploit in pattern recognition problems such as images or times series, however it is as computational demanding as the Wasserstein distance and despite recent advances in computation of optimal transport, e.g. [25], it is still challenging to apply it to large scale problems. In the next section, we review the linear optimal transport framework, which was introduced to allow application of optimal transport methods to classification problems [29]. In later sections, we combine the ideas of LOT and TL^p.

F.2.3 Linear optimal transport

The linear optimal transport (LOT) framework was proposed in [29], as a way to apply optimal transport techniques (in particular Wasserstein distances) to large scale classification problems for image analysis. Given a set of N images, one would need to compute all pairwise Wasserstein distances in order to use methods such as k-nearest neighbour classifiers. The LOT framework was developed so that only N Wasserstein distances need to be computed. In particular, it is the optimal transport maps between signals that are computed. From here, the images are embedded in a Euclidean space therefore allowing linear statistical techniques to be applied [50]. This technique was successfully applied in [2] to detect morphological difference in cancer cells. The technique has

been further refined and extended to super-resolution images [3], [30]. In this section, we briefly review the ideas of linear optimal transport.

The idea behind the LOT framework is to find the optimal transport maps with respect to one (reference) measure. In this section the cost c will be assumed to be given by $c(x,y) = |x-y|_p^p$ and therefore the Wasserstein distance relates to the optimal transport problems through $d_{\mathrm{W}^p}(\mu,\nu) = \sqrt[p]{M(\mu,\nu)} = \sqrt[p]{K(\mu,\nu)}$ (for simplicity we assume that the Monge problem is equavalent to the Kantorovich problem and, in particular, there exists optimal transport maps). Via an embedding of the transport map into Euclidean space the Wasserstein distance between any two pairs is estimated. More precisely, the LOT framework provides a linear embedding for $\mathcal{P}_p(\Omega)$ with respect to a fixed measure $\sigma \in \mathcal{P}_p(\Omega)$ [1]. This means the Euclidean distance of the embedded measure and the fixed measure σ is equal to the Wasserstein distance of the measure and the fixed measure. The Euclidean distance between any two measures is then an approximation to the Wasserstein distance between these measures. These linear embeddings then facilitate the application of standard statistical techniques such as PCA, LDA and K-means. The LOT framework is also invertible and so synthetic, but physically possible signals, can be realised [51].

Let $\mu_1, \mu_2 \in \mathcal{P}(\Omega)$ and $\sigma \in \mathcal{P}(\Omega)$ is our reference measure. Throughout this section, we assume that optimal transport maps $T^{\mu_i} : \Omega \to \Omega$ exist between σ and μ_i, i.e.

$$d_{\mathrm{W}^p}(\sigma,\mu_i) = \sqrt[p]{\int_\Omega |x - T^{\mu_i}(x)|_p^p \,\mathrm{d}\sigma(x)},$$

and $T^{\mu_i}_* \sigma = \mu_i$. If optimal transport maps do not exist then one can still define the LOT distance but there is not a natural way to embed this distance into a Euclidean space. We refer to [29, Section 2.3] on how to define the LOT distance using generalised geodesics.

In the setting considered here the *Linear Optimal Transport Distance* is defined by:

$$d_{\mathrm{LOT},\sigma}(\mu_1,\mu_2) := \sqrt[p]{\int_\Omega |T^{\mu_1}(x) - T^{\mu_2}(x)|_p^p \,\mathrm{d}\sigma(x)}.$$

We observe that $d_{\mathrm{LOT},\sigma}$ is a metric and

$$d_{\mathrm{LOT},\sigma}(\mu_1,\mu_2) = \|T^{\mu_1} - T^{\mu_2}\|_{L^p(\sigma)}.$$

Let us assume that σ has a density ρ with respect to the Lebesgue measure and define

$$P_c(\mu) = (T^\mu - \mathrm{Id})\rho^{\frac{1}{p}}. \tag{F.8}$$

Then

$$d_{\mathrm{LOT},\sigma}(\mu_1,\mu_2) = \|P_c(\mu_1) - P_c(\mu_2)\|_{L^p(\Omega)}. \tag{F.9}$$

The map P_c is our linear embedding from the Wasserstein space to a Euclidean space. We make the following claims on the embedding.

Proposition 59. Assume $\Omega \subset \mathbb{R}^d$ is bounded and $\sigma \in \mathcal{P}(\Omega)$ has a density ρ with respect to the

Lebesgue measure. Define $P = P_c$ where P_c is given by (F.8). Then, the following holds:

1. $P(\mu) \in L^p(\Omega)$ for any $\mu \in \mathcal{P}(\Omega)$,
2. $P(\sigma) = 0$,
3. $d_{\mathrm{LOT},\sigma}(\mu_1, \mu_2) = \|P(\mu_1) - P(\mu_2)\|_{L^p(\Omega)}$ for any $\mu_1, \mu_2 \in \mathcal{P}(\Omega)$,
4. $d_{\mathrm{LOT},\sigma}(\sigma, \mu) = d_{\mathrm{W}^p}(\sigma, \mu)$ for any $\mu \in \mathcal{P}(\Omega)$.

Proof. Since σ has a density with respect to the Lebesgue measure then transport maps T^{μ_i}, T^{μ} exist. Since the Wasserstein distance is finite (as Ω is bounded), it follows that $T^{\mu_i} - \mathrm{Id} \in L^p(\sigma)$ which proves (1). (2) follows directly from $T^{\sigma} = \mathrm{Id}$. (3) was shown already in (F.9). Finally, (4) follows

$$d_{\mathrm{LOT},\sigma}(\sigma, \mu) = \|(T^{\mu} - \mathrm{Id})\rho^{\frac{1}{p}}\|_{L^p(\Omega)} = \|T^{\mu} - \mathrm{Id}\|_{L^p(\sigma)} = d_{\mathrm{W}^p}(\sigma, \mu)$$

where we use $P(\sigma) = 0$. □

We make a similar definition for discrete measures. If $\sigma = \sum_{j=1}^n \rho_j \delta_{x_j}$ for some $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ then we define

$$[P_d(\mu_i)]_j = (T^{\mu_i}(x_j) - x_j)\rho_j^{\frac{1}{p}}. \tag{F.10}$$

Analogously to the Lebesgue density case we have

$$d_{\mathrm{LOT},\sigma}(\mu_1, \mu_2) = |P_d(\mu_1) - P_d(\mu_2)|_p$$

where we recall that $|\cdot|_p$ is the Euclidean p-norm: $|x|_p := \sqrt[p]{\sum_{j=1}^n |x_j|^p}$. For discrete σ the map P_d is our linear embedding from the Wasserstein space to a Euclidean space.

Proposition 60. Let $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ and assume $\sigma = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. Define $P = P_d$ where P_d is given by (F.10) with $\rho_j = \frac{1}{n}$. Then, the following holds:

1. $P(\mu) \in L^p(\Omega)$ for any $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$,
2. $P(\sigma) = 0$,
3. $d_{\mathrm{LOT},\sigma}(\mu_1, \mu_2) = |P(\mu_1) - P(\mu_2)|_p$ for any $\mu_1 = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, $\mu_2 = \frac{1}{n}\sum_{j=1}^n \delta_{z_j}$,
4. $d_{\mathrm{LOT},\sigma}(\sigma, \mu) = d_{\mathrm{W}^p}(\sigma, \mu)$ for any $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$.

Proof. The particular forms of all the measures $\sigma, \mu, \mu_1, \mu_2$ is enough to guarantee that transport maps $T^{\mu}, T^{\mu_1}, T^{\mu_2}$ all exist. The proof is then analogous to the proof of proposition 59. □

Example 61. Let us consider how to generate a new image using the linear embedding. Suppose we have a reference measure $\sigma \in \mathcal{P}(\mathbb{R}^d)$ with density ρ and a set of measures $\{\mu_i\}_{i=1}^N \subset \mathcal{P}(\mathbb{R}^d)$ with optimal transport maps T^{μ_i} which form the linear embedding through $\alpha_i = P(\mu_i) = (T^{\mu_i} - \mathrm{Id})\rho^{\frac{1}{p}}$. Given a new point α in the linear space we can define a transport map by $T = \alpha\rho^{-\frac{1}{p}} + \mathrm{Id}$. We generate a new image by $\mu = T_*\sigma$. Note that to generate the new image we only required the reference measure σ and a new point α in the linear space. However, in order to generate the new point α it will often be sensible to use the statistics of $\{\alpha_i\}_{i=1}^N$, for example see [51].

In both the Lebesgue density and uniform discrete case P preserves the Wasserstein distance between the reference measure and any given μ (where for discrete measures μ is also uniform discrete). Between μ_1, μ_2 one approximates $d_{\mathrm{W}^p}(\mu_1, \mu_2) \approx d_{\mathrm{LOT},\sigma}(\mu_1, \mu_2)$. The next section proposes our extension of the LOT framework to the TL^p distance.

F.2.4 A linear TL^p framework

In this section, we propose a linear TL^p framework. Recall that the TL^p distance can be defined as an optimal transport distance between measures supported on the graph of a function. Let $(\sigma, h) \in \mathrm{TL}^p$ be the TL^p *reference signal* and $\tilde{\sigma} = (\mathrm{Id} \times h)_* \sigma \in \mathcal{P}_p(\Lambda)$ the measure in $\Lambda = \Omega \times \mathbb{R}^m$ with support on the graph of h. Let $(\mu_i, f_i) \in \mathrm{TL}^p$, $i = 1, 2$, and define $\tilde{\mu}_i = (\mathrm{Id} \times f_i)_* \mu$. As in the previous section we will assume that optimal transport maps $\tilde{T}^{\tilde{\mu}_i} : \Lambda \to \Lambda$ exist between $\tilde{\sigma}$ and $\tilde{\mu}_i$, i.e.

$$d_{\mathrm{W}^p}(\tilde{\sigma}, \tilde{\mu}_i) = \sqrt[p]{\int_\Lambda |\boldsymbol{x} - \tilde{T}^{\tilde{\mu}_i}(\boldsymbol{x})|_p^p \, \mathrm{d}\tilde{\sigma}(\boldsymbol{x})}.$$

Recall that we can write $\tilde{T}^{\tilde{\mu}_i}$ in the form $\tilde{T}^{\tilde{\mu}_i}(\boldsymbol{x}) = (T^{\mu_i}(x), f_i(T^{\mu_i}(x)))$ where $\boldsymbol{x} = (x, y) \in \mathbb{R}^d \times \mathbb{R}^m$ and T^{μ_i} is the optimal plan for the Monge problem (F.2) between μ_i and σ with cost $c(x, y) = |x - y|_p^p + |f(x) - g(y)|_p^p$. The *Linear Transportation* L^p *Distance* is defined as

$$d_{\mathrm{LTL}^p,(\sigma,h)}((\mu_1, f_1), (\mu_2, f_2)) := \sqrt[p]{\int_\Lambda |\tilde{T}^{\tilde{\mu}_1}(\boldsymbol{x}) - \tilde{T}^{\tilde{\mu}_2}(\boldsymbol{x})|_p^p \, \mathrm{d}\tilde{\sigma}(\boldsymbol{x})}$$

Simple manipulations of the linear TL^p distance imply

$$\begin{aligned} d_{\mathrm{LTL}^p,(\sigma,h)}((\mu_1, f_1), (\mu_2, f_2)) &= \|\tilde{T}^{\tilde{\mu}_1} - \tilde{T}^{\tilde{\mu}_2}\|_{L^p(\tilde{\sigma})} \\ &= \sqrt[p]{\int_\Omega |T^{\mu_1}(x) - T^{\mu_2}(x)|_p^p + |f_1(T^{\mu_1}(x)) - f_2(T^{\mu_2}(x))|_p^p \, \mathrm{d}\sigma(x)}. \end{aligned}$$

Following the construction of the embedding in the previous section we go directly to the discrete case (since $\tilde{\sigma}$ has support on the graph it cannot have a density with respect to the Lebesgue measure). We assume that $\sigma = \sum_{i=1}^n \rho_j \delta_{x_j}$ for some $\{x_j\}_{j=1}^n \subset \mathbb{R}^d$ and we define

$$[P_d((\mu_i, f_i))]_j = (T^{\mu_i}(x_j) - x_j)\rho_j^{\frac{1}{p}} \tag{F.11}$$

$$[Q_d((\mu_i, f_i))]_j = (f_i(T^{\mu_i}(x_j)) - h(x_j))\rho_j^{\frac{1}{p}} \tag{F.12}$$

$$\tilde{P}_d((\mu_i, f_i)) = (P_d((\mu_i, f_i)), Q_d((\mu_i, f_i))). \tag{F.13}$$

Given this definition we can write

$$d_{\mathrm{LTL}^p,(\sigma,h)}((\mu_1, f_1), (\mu_2, f_2)) = |\tilde{P}_d((\mu_1, f_1)) - \tilde{P}_d((\mu_2, f_2))|_p.$$

The map $\tilde{P}_d$ embeds our signals into a Euclidean space. We have the following properties of the embedding (analogous to proposition 59 and proposition 60).

Proposition 62. Let $\{x_i\}_{i=1}^n \subset \mathbb{R}^d$ and assume $\sigma = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$. Define $\tilde{P} = \tilde{P}_d$ where $\tilde{P}_d$ is given by (eq. (F.11)-eq. (F.13)) with $\rho_j = \frac{1}{n}$. Then, the following holds:

1. $\tilde{P}((\mu, f)) \in \ell^p$ for any $(\mu, f) \in \mathrm{TL}^p$ with $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$,

2. $P((\sigma, h)) = 0$,

3. $d_{\mathrm{TL}^p}((\mu_1, f_1), (\mu_2, f_2)) = |\tilde{P}((\mu_1, f_1)) - \tilde{P}((\mu_2, f_2))|_p$ for any $(\mu_i, f_i) \in \mathrm{TL}^p$ with $\mu_1 = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$, $\mu_2 = \frac{1}{n}\sum_{j=1}^n \delta_{z_j}$,

4. $d_{\mathrm{TL}^p}((\sigma, h), (\mu, f)) = d_{\mathrm{TL}^p}((\sigma, h), (\mu, f))$ for any $(\mu, f) \in \mathrm{TL}^p$ with $\mu = \frac{1}{n}\sum_{j=1}^n \delta_{y_j}$.

Proof. By proposition 58(5) the transport maps $T^{\mu}, T^{\mu_1}, T^{\mu_2}$ exist. The rest of the proof follows as in the proof of proposition 59. □

As for LOT we have that LTL^p is exactly TL^p when comparing with the reference measure, i.e. $d_{\mathrm{LTL}^p,(\sigma,h)}((\sigma, h), (\mu, f)) = d_{\mathrm{TL}^p}((\sigma, h), (\mu, f))$. When we are comparing two measures, neither of which are the base measure, then we make the approximation

$$d_{\mathrm{LTL}^p,(\sigma,h)}((\mu_1, f_1), (\mu_2, f_2)) \approx d_{\mathrm{TL}^p}((\mu_1, f_1), (\mu_2, f_2)).$$

Example 63. Let us consider how to generate a new TL^p image from the linear space. We recall that colour images can be represented by (μ, f) where $\{x_i\}_{i=1}^n$ are the locations of pixels (which are uniform across $[0, 1] \times [0, 1]$), $\mu = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$ and $f : \{x_i\}_{i=1}^n \to \mathbb{R}^3$ represents the RGB values for each pixel. We take a reference image (σ, h) of the same form (in particular $\sigma = \frac{1}{n}\sum_{i=1}^n \delta_{x_i}$), and note that the embedding is given by $\alpha^{\tilde{\mu}} = (\alpha_1^{\tilde{\mu}}, \ldots, \alpha_n^{\tilde{\mu}}) \in \mathbb{R}^{5n}$ where

$$\alpha_i^{\tilde{\mu}} = \frac{1}{n^{\frac{1}{p}}}\left(\tilde{T}^{\tilde{\mu}}(x_i) - (x_i, h(x_i))\right) \in \mathbb{R}^5.$$

To generate a new image we need to invert this mapping. Let $\alpha = (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^{5n}$. We define $\tilde{T}_i = n^{\frac{1}{p}}\alpha_i + (x_i, h(x_i)) \in \mathbb{R}^5$. Then $y_i = (\tilde{T}_i)_{1:2} \in \mathbb{R}^2$ are the location of the pixels and $c_i = (\tilde{T}_{3:5})$ are the RGB values in the new image. In the TL^p space the new image is represented by (ν, g) where $\nu = \frac{1}{n}\sum_{i=1}^n \delta_{y_i}$ and $g(y_i) = c_i$. This is only well defined if y_i are all unique. If not, then we use Barycentric projection (see also appendix F.2.7), for example if $y_i = y$ for all $i \in \mathcal{I}$ then we define $g(y) = \frac{1}{|\mathcal{I}|}\sum_{\mathcal{I}} c_i$ to be the empirical average.

F.2.5 Computing the linear transportation L^p embedding

In this subsection we review some methods for computing the linear TL^p embedding. In principle, any algorithm that can compute optimal transport distances can be adapted to compute TL^p by either interpreting TL^p as a Wasserstein distance on the graphs of functions, or as an optimal transport problem with cost $c(x, y; f, g) = |x - y|_p^p + |f(x) - g(y)|_p^p$. We refer to [52] for a thorough review of computational methods for optimal transport. Here, we review the following methods in the setting of TL^p: linear programming, entropy regularised optimal transport, and flow minimisation.

Once we have obtained optimal TL^p maps $\tilde{T}^{\tilde{\mu}_i} : \Omega \times \mathbb{R}^m \to \Omega \times \mathbb{R}^m$ for each of the transportation problems between $\tilde{\sigma} \in \mathcal{P}(\Omega \times \mathbb{R}^m)$ and $\tilde{\mu}_i \in \mathcal{P}(\Omega \times \mathbb{R}^m)$ for $i = 1, ..., N$ we can embed into Euclidean space by (eq. (F.11)-eq. (F.13)) where $\tilde{T}^{\tilde{\mu}_i} = (T^{\mu_i}, f_i \circ T^{\mu_i})$. Hence, linear statistical methods can be applied. In some cases we maybe interested in using derivative information of the signals,

in appendix F.2.6 we further extend the theory to the case of the linear transportation Sobelev distance.

Linear programming

Consider the case when

$$\mu = \sum_{i=1}^{m} p_i \delta_{x_i}, \quad \nu = \sum_{j=1}^{n} q_j \delta_{y_j}, \quad f_i = f(x_i) \quad \text{and} \quad g_j = g(y_j). \tag{F.14}$$

Then the TL^p distance can be written as

$$d^p_{\mathrm{TL}^p}((\mu, f), (\nu, g)) = \min_{\pi} \sum_{i=1}^{m} \sum_{j=1}^{n} \left(|x_i - y_j|_p^p + |f_i - g_j|_p^p \right) \pi_{ij}$$

where the minimum is taken over matrices $\pi \in \mathbb{R}_+^{n \times m}$ such that the row sums are $(p_1, \dots, p_m)$ and the column sums are $(q_1, \dots, q_n)$. This is a linear programme and can be solved, for example, by using the simplex or interior point methods. Although this works well for small problems (i.e. when n and m are small) linear programming algorithms quickly become infeasible for large problems. There are ad-hoc multiscale methods such as [19] that allow for the application of linear programming methods to larger problems but these do not have theoretical guarantees (for example, there are no guarantees regarding finding the global minimiser).

An entropy regularisation approach

We assume two pairs $(\mu, f), (\nu, g) \in \mathrm{TL}^p$ can be written in the form (F.14). It was proposed in [25] to consider the entropy regularised problem

$$S_\varepsilon((\mu, f), (\nu, g)) = \min_{\pi} \left(\sum_{i=1}^{m} \sum_{j=1}^{n} \left(|x_i - y_j|_p^p + |f_i - g_j| \right) \pi_{ij} - \varepsilon H(\pi) \right) \tag{F.15}$$

where $\varepsilon > 0$ is a positive parameter that controls the amount of regularisation, H is entropy and defined by

$$H(\pi) = -\sum_{i=1}^{n} \sum_{j=1}^{m} \pi_{ij} \log \pi_{ij}$$

and the minimum in (F.15) is taken over matrices $\pi \in \mathbb{R}_+^{n \times m}$ such that the row sums are $\boldsymbol{p} = (p_1, \dots, p_m)$ and the column sums are $\boldsymbol{q} = (q_1, \dots, q_n)$. When $\varepsilon \to 0$, the results of [53] imply that $S_\varepsilon((\mu, f), (\nu, g)) \to d^p_{\mathrm{TL}^p}((\mu, f), (\nu, g))$. Subsequent developments of the entropy regularised approach have appeared in [26], [27]. The measure S_ε is referred to as the Sinkhorn distance. It is easy to see that

$$S_\varepsilon((\mu, f), (\nu, g)) = \varepsilon \inf_{\pi} \{\mathrm{KL}(\pi | K)\},$$

where $K_{ij} = \exp\left(-\frac{C_{ij}}{\varepsilon}\right)$ is the Gibbs distribution, $C_{ij} = |x_i - y_j|_p^p + |f_i - g_j|_p^p$, and KL denotes the *Kullback-Leibler* divergence. The minimisation is taken over the same set as in (F.15) The

optimal choice for π can be written in the following form:

$$\pi^\dagger = \mathrm{diag}(u) K \mathrm{diag}(v),$$

where u, v are the limits, as $r \to \infty$, of the sequence

$$v^{(0)} = \mathbf{1},\ u^{(r)} = \frac{\boldsymbol{p}}{Kv^{(r)}},\ v^{(r+1)} = \frac{\boldsymbol{q}}{K^T u^{(r)}},$$

see [27]. The entropy regularisation means the optimal π for S_ε cannot be written as a transport map. To obtain an approximation to the optimal transport map one can use Barycentric projections as in [19, Section 2.3].

A flow minimisation approach

Following [12] we derive a flow minimization method for finding the transportation map in TL^2. Let $\Omega \subset \mathbb{R}^d$ be a compact domain with smooth boundary and let (μ, f) and (ν, g) be signals in $\mathrm{TL}^2(\Omega, \mathbb{R}^m)$, where $f, g : \Omega \to \mathbb{R}^m$ are square-integrable functions. Furthermore, we assume that the measures μ and ν admit densities with respect to the Lebesgue measure. Abusing notation we write $\mathrm{d}\mu(x) = \mu(x)\mathrm{d}x$. The variational TL^2 problem is finding the diffeomorphic map $T : \Omega \to \Omega$, which minimises the following energy

$$\varepsilon(T) = \int_\Omega \left(|T(x) - x|_2^2 + |g(T(x)) - f(x)|_2^2 \right) \mu(x)\, \mathrm{d}x \tag{F.16}$$

$$\text{subject to } T_*\mu = \nu. \tag{F.17}$$

We assume the following polar factorization of T. Let $s : \Omega \times [0, \infty) \to \Omega$ and assume the second coordinate is time. We further assume for any fixed t, $[s(\cdot, t)]_*\mu = \mu$. That is, $s(\cdot, t) : \Omega \to \Omega$ is a mass preserving rearrangement of μ. Let $T^0 : \Omega \to \Omega$ be an initial mass preserving map between μ and ν, i.e. $T^0_*\mu = \nu$, for example the Knothe-Rosenblatt coupling [15]. We assume that $s(\cdot, t)$ is invertible in x for every t and with an abuse of notation we write s^{-1} for this inverse, i.e.

$$s^{-1}(s(x,t),t) = x = s(s^{-1}(x,t),t) \quad \text{for all } x \in \Omega \text{ and for all } t \in [0,\infty). \tag{F.18}$$

We require that $T = T^0 \circ s^{-1}$. The strategy in [12] is to evolve $s(\cdot, t)$ using a gradient descent step such that it converges to a minimiser of (F.16) satisfying the constraint (F.17) as $t \to \infty$. We first consider sufficient conditions on s in order to guarantee that (F.17) holds for all $t > 0$. The proof of the proposition can be found in [12, Section A.2].

Proposition 64. Let $\Omega \subset \mathbb{R}^d$ be a compact domain with a smooth boundary and $\mu, \nu \in \mathcal{P}(\Omega)$. Assume that μ and ν have C^1 densities with respect to the Lebesgue measure on Ω and with an abuse of notation write $\mathrm{d}\mu(x) = \mu(x)\,\mathrm{d}x$ and $\mathrm{d}\nu(x) = \nu(x)\,\mathrm{d}x$ Let χ be a C^1 vector field on Ω satisfying $\mathrm{div}(\chi) = 0$ on Ω and $\chi \cdot \mathbf{n} = 0$ on $\partial\Omega$ where $\mathbf{n}$ is the normal to the boundary of Ω. Assume $s : \Omega \times [0, \infty) \to \Omega$ is differentiable and invertible in the sense of (F.18), and T^0 satisfies $T^0_*\mu = \nu$. If, $s(\cdot, 0) = \mathrm{Id}$ and for all $t > 0$

$$\frac{\partial s}{\partial t}(x,t) = \frac{1}{\mu(s(x,t))} \chi(s(x,t)) \tag{F.19}$$

then $[s(\cdot,t)]_*\mu = \mu$. Furthermore $\frac{\partial T^t}{\partial t} = -\frac{1}{\mu}\nabla T^t \chi$ and $T^t_*\mu = \nu$ where

$$T^t = T^0(s^{-1}(\cdot,t)). \tag{F.20}$$

If we restrict ourselves to look for transport maps of the form (eq. (F.19)-eq. (F.20)) then we must decide how to choose χ. Let us define $s_\chi : \Omega \times [0,\infty) \to \Omega$ by (F.19) with $s_\chi(\cdot,0) = \mathrm{Id}$ and $T^t_\chi : \Omega \to \Omega$ by (F.20) with $s = s_\chi$. An obvious criterion is to choose χ so that $\varepsilon(T^t_\chi)$ decreases quickest over all choices of χ. To this end we compute the derivative of $\varepsilon(T^t_\chi)$ with respect to t.

Lemma 65. In addition to the assumptions and notation of proposition 64 let $f \in C^1(\Omega;\mathbb{R}^m)$ and $g \in L^2(\nu)$ and define ε by (F.16). Define $s_\chi : \Omega \times [0,\infty) \to \Omega$ by (F.19) with $s_\chi(\cdot,0) = \mathrm{Id}$ and $T^t_\chi : \Omega \to \Omega$ by (F.20) with $s = s_\chi$. Then we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(T^t_\chi) = -\int_\Omega Q(x,t)\cdot\chi(x)\,\mathrm{d}x$$

where

$$Q(t,x) = 2T^t_\chi(x) + 2\sum_{i=1}^m g_i(T^t_\chi(x))\nabla f_i(x). \tag{F.21}$$

Proof. We define

$$\tilde{\varepsilon}(T;f,g) = \int_\Omega |g(T(x)) - f(x)|_2^2\,\mathrm{d}\mu(x)$$

which we can also write as

$$\tilde{\varepsilon}(T;f,g) = \int_\Omega |g(y)|_2^2\,\mathrm{d}\nu(y) + \int_\Omega |f(x)|_2^2\,\mathrm{d}\mu(x) - 2\int_\Omega g(T(x))\cdot f(x)\,\mathrm{d}\mu(x).$$

By a change of variables $y = s_\chi^{-1}(x,t)$, and since $[s_\chi(\cdot,t)]_*\mu = \mu$ we have

$$\begin{aligned}\tilde{\varepsilon}(T^t_\chi;f,g) &= \int_\Omega |g(y)|_2^2\,\mathrm{d}\nu(y) + \int_\Omega |f(x)|_2^2\,\mathrm{d}\mu(x) - 2\int_\Omega g(T^0(s_\chi^{-1}(x,t)))\cdot f(x)\,\mathrm{d}\mu(x)\\ &= \int_\Omega |g(y)|_2^2\,\mathrm{d}\nu(y) + \int_\Omega |f(x)|_2^2\,\mathrm{d}\mu(x) - 2\int_\Omega g(T^0(y))\cdot f(s_\chi(y,t))\,\mathrm{d}\mu(y).\end{aligned}$$

Differentiating the above we obtain,

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\tilde{\varepsilon}(T^t_\chi;f,g) &= -2\sum_{i=1}^m \int_\Omega g_i(T^0(y))\nabla f_i(s_\chi(y,t))\cdot\frac{\partial s_\chi}{\partial t}(y,t)\,\mathrm{d}\mu(y)\\ &= -2\sum_{i=1}^m \int_\Omega g_i(T^t_\chi(x))\nabla f_i(x)\cdot\chi(x)\,\mathrm{d}x.\end{aligned}$$

We note that $\varepsilon(T^t_\chi) = \tilde{\varepsilon}(T^t_\chi;\mathrm{Id},\mathrm{Id}) + \tilde{\varepsilon}(T^t_\chi;f,g)$ hence $\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(T^t_\chi) = -\int_\Omega Q(x,t)\cdot\chi(x)\,\mathrm{d}x$. □

When $d = 2$ by the Helmholtz decomposition (in 2D) we can find, for each $t > 0$, two scalar fields $w : \Omega \to \mathbb{R}$ and $\alpha : \Omega \to \mathbb{R}$ such that $Q(\cdot,t) = \nabla w + \nabla^\perp\alpha$ (where the t dependence on α and w is supressed) and $\alpha = 0$ on $\partial\Omega$ where $\nabla^\perp f = \left(-\frac{\partial f}{\partial x_2}, \frac{\partial f}{\partial x_1}\right)$ for a function $f(x) = f(x_1,x_2)$. To find the direction of steepest descent we let $\psi = \nabla^\perp\alpha$ and $\chi = \nabla^\perp\beta$ and compute

$$\frac{\mathrm{d}}{\mathrm{d}t}\varepsilon(T^t_\chi) = -\int_\Omega (\nabla w(x) + \psi(x))\cdot\chi(x)\,\mathrm{d}x$$

$$
\begin{aligned}
&= -\int_\Omega \left(\mathrm{div}(w\chi)(x) - w(x)\mathrm{div}(\chi)(x)\right)\,\mathrm{d}x - \int_\Omega \psi(x)\cdot\chi(x)\,\mathrm{d}x \\
&= -\int_{\partial\Omega} w(x)\chi(x)\cdot\mathbf{n}(x)\,\mathrm{d}S(x) - \int_\Omega \psi(x)\cdot\chi(x)\,\mathrm{d}x \\
&= -\int_\Omega \psi(x)\cdot\chi(x)\,\mathrm{d}x \\
&= -\int_\Omega \nabla\alpha(x)\cdot\nabla\beta(x)\,\mathrm{d}x
\end{aligned} \tag{F.22}
$$

where the third line follows from the divergence theorem and since $\mathrm{div}(\chi) = 0$ on Ω, and the fourth line follows from $\chi(x)\cdot\mathbf{n}(x) = 0$ on $\partial\Omega$. It follows that the direction of steepest descent is $\alpha = \beta$.

To find α, we need to observe that $\nabla\alpha = -Q^\perp - \nabla^\perp w$ where $\perp$ is rotation clockwise by $\pi/2$, i.e. $Q^\perp = (-Q_2, Q_1)$. Taking the divergence we have

$$
\Delta\alpha = \mathrm{div}(\nabla\alpha) = \mathrm{div}(-Q^\perp - \nabla^\perp w) = -\mathrm{div}(Q^\perp).
$$

Hence, α solves the Poisson equation with Dirichlet boundary conditions:

$$
\Delta\alpha = -\mathrm{div}(Q^\perp) \qquad \text{in } \Omega \tag{F.23}
$$

$$
\alpha = 0 \qquad \text{on } \partial\Omega. \tag{F.24}
$$

To summarise, the flow minimization scheme for TL^p, given a step size τ is as follows.

1. Construct T^0 and set $t = 0$.
2. Compute $Q(\cdot, t)$ defined by (F.21).
3. Find α by solving (eq. (F.23)-eq. (F.24)).
4. Update $T^{t+\tau} = T^t - \frac{\tau}{\mu}\nabla T^t \nabla^\perp \alpha$.
5. Set $t \mapsto t + \tau$.
6. Repeat 2-5 until convergence.

F.2.6 Linear transportation Sobelev distance

A natural way to include information about the derivative of a signal is to augment the signal. Briefly, assume $f : \mathbb{R} \to \mathbb{R}$ is in the Sobelev space $\mathrm{W}^{1,p}$, this means that the first weak derivative of f is integrable in L^p, with the base measure being $\mathcal{L}^d$, the Lebesgue measure on $\mathbb{R}^d$, for simplicity. We can then define the augmented signal $\boldsymbol{f} = (f, \frac{df}{dx})$ and we note that $\boldsymbol{f} : \mathbb{R} \to \mathbb{R}^2$ is simply a function. Thus we can define a transportation Sobelev distance:

$$
d_{\mathrm{TW}^{1,p}}((\mathcal{L}^d, f), (\mathcal{L}^d, g)) = d_{\mathrm{TL}^p}((\mathcal{L}^d, \boldsymbol{f}), (\mathcal{L}^d, \boldsymbol{g})).
$$

This was first proposed in [32] and we refer the readers there for the theoretical background. Furthermore, our construction from the previous section carry through to the Sobolev setting and so we can propose a linear transportation $W^{1,p}$ distance. There is no need to restrict ourselves to only one derivative. Assuming that $f \in \mathrm{W}^{k,p}$; that is, f has k weak derivatives integrable in L^p,

then we can define $\boldsymbol{f} = (f, \frac{df}{dx}, \frac{d^2f}{dx^2}, ..., \frac{d^kf}{dx^k})$. Thus a transportation Sobelev distance, which we denote $\mathrm{TW}^{k,p}$ is given by

$$d_{\mathrm{TW}^{k,p}}(f,g) = d_{\mathrm{TL}^p}((\mathcal{L}^d, \boldsymbol{f}), (\mathcal{L}^d, \boldsymbol{g})).$$

The linear theory we have developed in previous section then applies. In addition, multi-variable functions $f : \mathbb{R}^m \to \mathbb{R}^m$ with k weak derivative integrable in L^p can be handled in a simple manner by defining

$$J_k f = \left(f_1, ..., f_m, \frac{\partial f_1}{\partial x_1}, ..., \frac{\partial f_1}{\partial x_m}, \frac{\partial f_2}{\partial x_1}, ..., \frac{\partial f_2}{\partial x_m}, ..., \frac{\partial^k f_m}{\partial x_m^k} \right).$$

A full characterization of the space $(\mathrm{TW}^{k,p}, d_{\mathrm{TW}^{k,p}})$ is given in [32]; in particular, it is a metric space.

F.2.7 Geodesics and interpolation

The space $(\mathcal{P}_p(\Omega), d_{\mathrm{W}^p})$ is a geodesic space, with easily characterisable geodesics. Letting $\pi^\dagger \in \Pi(\sigma, \mu)$ be the optimal transport plan that minimises the transport problem given by (F.1), we define $I_t : \Omega \times \Omega \to \Omega$, where $t \in [0, 1]$, to be a linear interpolation, as follows:

$$I_t(x, y) = (1 - t)x + ty.$$

Then the geodesic in Wasserstein space is given by $\mu(t) = [I_t]_* \pi^\dagger$. When there exists transport maps, i.e. $\pi^\dagger = (\mathrm{Id} \times T^\mu)_* \mu$ then the geodesic can be written $\mu_t = [T_t^\mu]_* \mu$ where $T_t^\mu(x) = I_t(x, T^\mu(x)) = (1 - t)x + tT^\mu(x)$. Let P be defined by $P = P_c$ in (F.8) then since

$$P(\mu_t) = (T_t^\mu - \mathrm{Id})\rho^{\frac{1}{p}} = ((1 - t)\mathrm{Id} + tT^\mu - \mathrm{Id})\,\rho^{\frac{1}{p}} = t\,(T^\mu - \mathrm{Id})\,\rho^{\frac{1}{p}} = tP(\mu)$$

we see that the projection of the geodesic onto the Euclidean space is the geodesic between the projections. In particular, the geodesic between $P(\sigma) = 0$ and $P(\mu)$ in Euclidean space is simply $tP(\mu)$. The same argument holds in the discrete case where $P = P_d$ is defined by (F.10). Since the projection is invertible we can map any point in the Euclidean embedding back to the optimal transport space. Notably, this allows one to translate principal eigenvectors in PCA space (of the linear embedding) into modes of variation in optimal transport space, see [29] for more details.

This argument does not directly apply to the TL^p space since, by the following remark, the TL^p space does not permit geodesics.

Remark 66. Consider the measure $\mu = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ and the functions $f(0) = 0, f(1) = 10, g(0) = 10, g(1) = 0$. Then the transport between (μ, f) and (μ, g) is from $(0, 0)$ to $(1, 0)$ and from $(1, 10)$ to $(0, 10)$. The "half way" point would be the measure $\mu_{\frac{1}{2}} = \delta_{\frac{1}{2}}$ and the function that takes the value 10 and 0 at $x = \frac{1}{2}$, which is not a function.

However, this does not prevent us from interpolating and visualising modes of variation. Indeed, let $\tilde{T}^{\tilde{\mu}}(\boldsymbol{x}) = (T^\mu(x), f(T^\mu(x)))$ be the optimal TL^p map pushing (σ, h) to (μ, f), then the map

$$\tilde{T}_t^{\tilde{\mu}}(\boldsymbol{x}) = ((1 - t)x + tT^\mu(x), (1 - t)h(x) + t(f(T^\mu(x))))$$

interpolates between the signals (σ, h) and (μ, g). In fact, this is the geodesic in $(\mathcal{P}_p(\Omega \times \mathbb{R}^m), d_{\mathrm{W}^p})$; that is $\tilde{\mu}_t = \tilde{T}_t^{\tilde{\mu}}$ is the geodesic in $(\mathcal{P}_p(\Omega \times \mathbb{R}^m), d_{\mathrm{W}^p})$ between $\tilde{\sigma}$ and $\tilde{\mu}$. Although we can invert

$\tilde{P} = \tilde{P}_d$ (defined in (F.11)) in the Wasserstein space, i.e. for all $p \in \mathbb{R}^{nm}$ there exists $\tilde{\nu} \in \mathcal{P}(\Omega \times \mathbb{R}^m)$ such that $P(\tilde{\nu}) = p$ (note that this is P and not $\tilde{P}$ since we are inverting with respect to the Wasserstein embedding) we cannot guarantee that $\tilde{\nu}$ can be written in the form $\tilde{\nu} = (\mathrm{Id} \times g)_* \nu$. Hence, we cannot in general invert the linear embeddings from TL^p back into TL^p. Instead we use an approximate inversion. We define $\tilde{P}^{-1}(p) = (\nu, \bar{g})$ where $\tilde{\nu}$ satisfies $P(\tilde{\nu}) = p$, $\bar{g}(x) = \mathbb{E}\tilde{\nu}_x$ and (where we use disintegration of measures) $\tilde{\nu} = \tilde{\nu}_x \otimes \nu$ with the latter meaning

$$\tilde{\nu}(A \times B) = \int_A \tilde{\nu}_x(B)\, \mathrm{d}\nu(x) \qquad \text{for all measureable } A \subset \Omega, B \subset \mathbb{R}^m.$$

In other words, we define the "inverse" map from the linear embedding of TL^p back into TL^p as the inverse map in Wasserstein distance and projected onto the TL^p space:

1: $p \in \{\text{linear } \mathrm{TL}^p \text{ space}\} \mapsto \tilde{\nu} \in \mathcal{P}(\Omega \times \mathbb{R}^m)$ using inverse of LOT in the space $\Omega \times \mathbb{R}^m$

2: $\tilde{\nu} \mapsto (\nu, g) \in \mathrm{TL}^p$ by projecting the Wasserstein space on $\Omega \times \mathbb{R}^m$ onto TL^p.

The projection onto TL^p is done by taking the mean across each fibre in $x \in \Omega$. With this definition we are also able to visualise any point in the linear embedding in TL^p space.

F.3 Results

In this section we apply the linear TL^p framework to synthetic and real world examples. Our first synthetic example is a 1D class of functions that is based on the synthetic example in [33]. The second synthetic example is in 2D. We consider real world applications to cell morphometry, auslan (Australian sign language), breast cancer histopathology and financial time series.

F.3.1 One-dimensional synthetic signal processing

We first consider a one dimensional signal processing problem, to test the ability of OT and TL^p to discriminate between different signals. Throughout, we take the Euclidean distance as the cost function and $p = 2$ for TL^p distance. We consider the task of discriminating between double hump and a high-frequency perturbation of the hump function: a chirp function. A double hump function is of the form:

$$f = K_1 \cdot (\mathbb{1}_{[l,l+r]} + \mathbb{1}_{[l+b+r,l+b+2r]}), \tag{F.25}$$

where $\mathbb{1}_{[\alpha,\beta]}$ denotes the indicator function on the interval $[\alpha, \beta]$ and $l \in [0, 1-b-2r]$. The constant K_1 is chosen such that f integrates to unity. A chirp-hump function is given as:

$$f = K_2 \cdot \left(\sum_{j=0}^{\frac{r}{\gamma}-1} \mathbb{1}_{[l+j\gamma, l+\frac{(2j+1)\gamma}{2}]} + \frac{1}{4}\mathbb{1}_{[l+b+r,l+b+2r]} \right), \tag{F.26}$$

where γ controls the high-frequency perturbation and K_2 is chosen so that f integrates to unity. To generate our synthetic dataset we proceed as follows, fixing l, r and b, we generate $f_1, ..., f_{30}$ from (F.25). We corrupt each signal with standard Gaussian noise to obtain 30 noisy double hump functions. We then obtain two separate classes from the chirp-hump functions by first randomly sampling $\gamma \in \{\gamma_1, \gamma_2\}$ with equal probability. Each chirp-hump function is then corrupted

with standard Gaussian noise. We then obtain functions $f_{31}, ..., f_{60}$ as chirp-hump functions with proportion R_1 having perturbation parameter γ_1 and proportion R_2 having perturbation parameter γ_2. All functions are defined on $[0, 1]$ discretized on a uniform gird of length $N = 150$.

To apply LOT to discriminate between these signals we first need to satisfy positivity and mass constraints. Thus, each function f is normalised as follows $g = \frac{f+\chi}{\int(f+\chi)}$, for a small number χ. Discrete measures $\mu_1, ..., \mu_{60}$ are then defined to be the probability measures with density $g_1, \ldots, g_{60}$ with respect to the uniform grid on $[0, 1]$. A reference measure σ is constructed as an empirical average of all these measures. We then use entropy regularised methods, see [25], [27] or appendix F.2.5 to compute optimal transport plans between σ and μ_i, where $i = 1, ..., 60$; after which an optimal transport map is computed using Barycentric projection. We then embed the measures into Euclidean space as described in appendix F.2.3. Note that for this linear embedding $E \in \mathbb{R}^{150\times 60}$. This method requires only the computation of 60 transport plans, rather than 59×30 if all pairwise OT distance were computed.

The linear TL^p framework can be applied directly without ad-hoc pre-processing and normalisation. The base measure is taken as the uniform measure on $[0, 1]$. The reference measure σ is taken to be the base measure and the reference signal h is the empirical average of all signals. Optimal TL^p plans are computed again using entropy regularised methods from (σ, h) to (μ_i, f_i) for $i = 1, ..., 60$. Recall that this requires the computation of the optimal transport plan from $\tilde{\sigma}$ to $\tilde{\mu_i}$ for $i = 1, ..., 60$. The map is then obtained from the plan via Barycentric projection. A linear embedding U is obtained as detailed in appendix F.2.4. Note this linear embedding is higher dimensional and $U \in \mathbb{R}^{300\times 60}$.

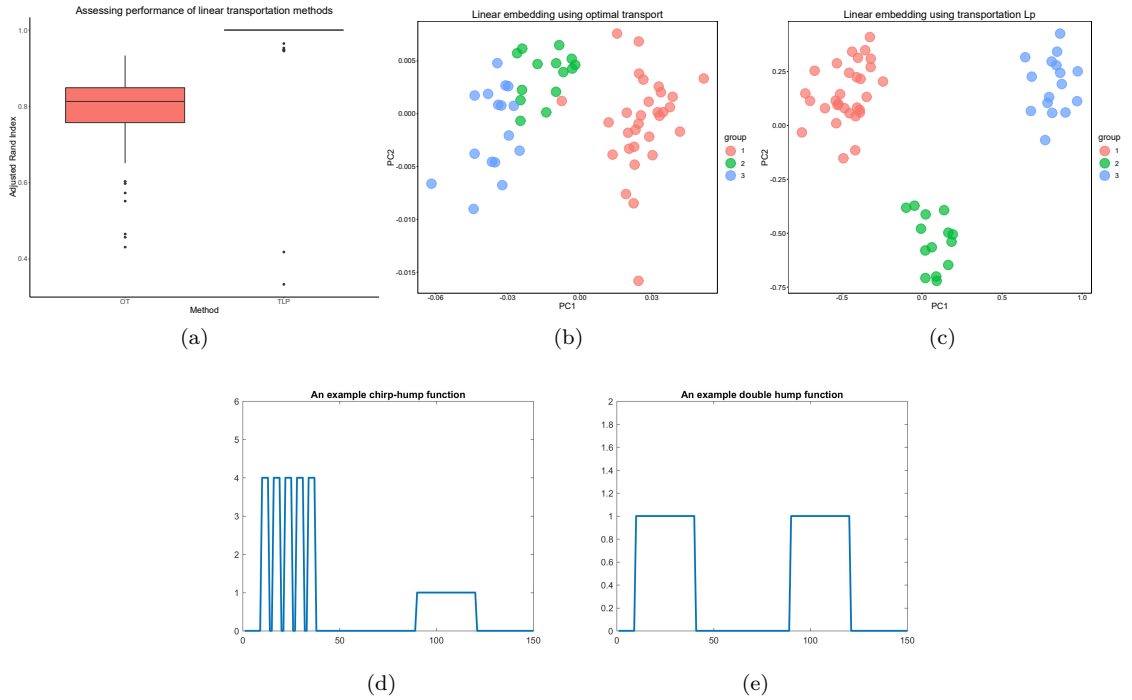
To assess the discriminating ability of OT and TL^p we apply K-means clustering with $K = 3$ to the linear embedding to see if we can recover the true underlying classes. Since K-means attempts to minimise the within class distance and maximise between class distance, we expect a distance which is able to detect the differences between the classes to have the best performance. We take the clustering returned from the K-means algorithm and compare it to the true clustering using the adjusted Rand index (ARI) [54], [55]. An ARI is a score with 1 indicating perfect agreement, 0 indicating the method performs as well as one would expect if random assignment where made and the ARI can be negative if the method is worse than random. We repeat our method 100 times to produce a distribution of scores.

Figure F.1 demonstrate the improved performance of using the TL^p distance to form a linear embedding of the data. The median ARI using the TL^p distance was 1, whilst the median for using the linear OT approach was 0.8129. The TL^p approach outperforms the OT approach significantly (Kolmogorov-Smirnov (KS) Test, $p < 10^{-4}$). Furthermore, fig. F.1 panels (b) and (c) demonstrates that the linear embedding produce much tighter and therefore more interpretable clusters when using the TL^p approach as compared to the OT approach.

F.3.2 Two-dimensional synthetic signal processing

We consider a more challenging synthetic two dimensional signal processing problem. As in the previous section, we take $p = 2$ as the exponent of the cost function in both settings. Consider the following class of functions, defined on a grid on $[0, 1]^2$:

$$\mathcal{M}_1 = \left\{ f : [0, 1]^2 \to \mathbb{R} \,|\, f(x_i, y_j) = \alpha x_i \mathrm{e}^{-x_i^2 - y_j^2} + \sigma_{ij}, \alpha \sim \mathcal{N}(0, 1), \sigma_{ij} \sim \mathcal{N}(0, 1) \right\},$$



(a) (b) (c)

(d) (e)

Figure F.1: (a) A boxplot showing the distribution of adjusted Rand index over different runs of the algorithm. The TL^p based method clearly outperform the OT approach. (b) An example of a linear OT embedding showing classes that appear dispersed, and between class distance is smaller than within class distance. (c) An example of a linear TL^p embedding, between class distance is clearly greater than within class distance. This results in tighter clusters. (d) An example chirp-hump function (without noise). (e) An example double hump function (without noise).

and

$$\mathcal{M}_2 = \left\{ f : [0,1]^2 \to \mathbb{R} \,|\, f(x_i, y_j) = \alpha x_i \mathrm{e}^{-x_i^2 - y_j^2} + \sigma_{ij}, \alpha \sim \mathcal{N}(-4, 1.5), \sigma_{ij} \sim \mathcal{N}(0,1) \right\}.$$

Furthermore, we introduce a perturbation to functions in $\mathcal{M}_1$, by first sampling an integer $n \in \{10, ..., 20\}$ each with equal probability and then setting $f = -2$ for n randomly chosen coordinates on the grid, with each coordinate having equal probability of being chosen.

We note that these signals take positive and negative values and thus to apply linear OT to this problem we perform normalisation. We add a constant to each (random) function and then ensure that mass still integrates to unity. This ad-hoc normalisation procedure introduces signal compression into the problem causing important features to become suppressed. The TL^p distance can be applied without normalisation or pre-processing.

We generate 25 random functions from each class. We then apply both the linear OT and linear TL^p methods to the resulting dataset. We performed PCA and K-means clustering, with $K = 2$, on the linear embedding to see if we can discriminate between the two classes. As in the previous section, we compare the resultant clustering with ground truth using the ARI. We repeat this process 100 times to obtain a distribution of scores.

Figure F.2 show that a linear embedding based on a TL^p distance outperforms the LOT embedding. The median ARI for the TL^p approach is 0.92 and the median ARI for OT is 0.5689. The linear TL^p distance produced significantly better results (KS-test, $p < 10^{-4}$) and the PCA

plots in fig. F.2 demonstrate that the two classes overlap when using OT but separate when using TL^p.

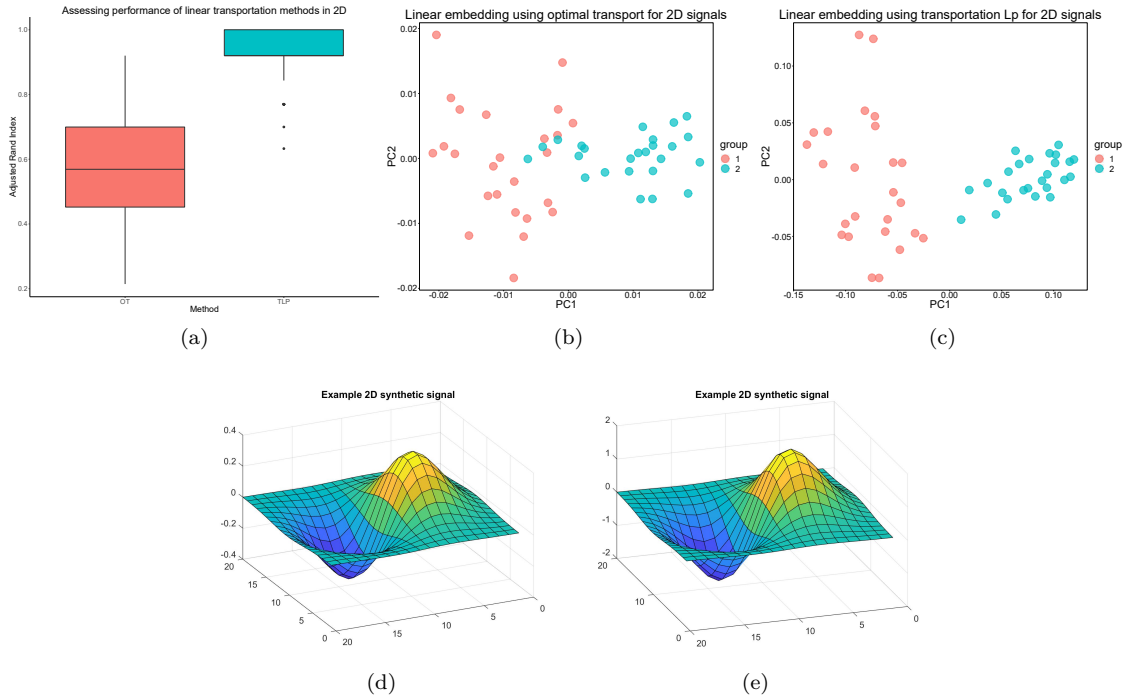


Figure F.2: (a) A boxplot showing the distribution of adjusted Rand index over different runs of the algorithm. The TL^p based method clearly outperforms the OT approach. (b) An example of a linear OT embedding where we see overlap between the classes as a result of signal compression. (c) An example of a linear TL^p embedding where there is a clear separation between classes. (d) An example signal (without noise) from $\mathcal{M}_1$. (e) An example signal (without noise) from $\mathcal{M}_2$.

F.3.3 Application: Cell morphometry

In this section, we analyse the liver dataset of [2] containing 250 normal and 250 cancerous liver cells. [2] proposed a transportation based morphometry analysis and this facilitated high accuracy classification. Furthermore, the generative nature of optimal transport allowed them to visualise the modes of variation in the dataset, allowing for superior interpretation of the data. Each cell image is defined on a 192×192 pixel grid with a single intensity channel. Thus both linear OT and linear TL^p are applicable.

For consistency we apply flow minimisation techniques to compute the transport maps in each setting [30]. To alleviate numerical issues, as in [30], we apply a Gaussian low-pass filter with standard deviation 2 to smooth the data. We perform mass normalisation so that optimal transport can be applied and these signals are also used for TL^p, so that the spatial and intensity features are on the same scale. A linear embedding is obtained from the transport maps.

Once this linear embedding is obtained we use the 1 nearest neighbour (1NN) algorithm to predict normal or cancerous from the linear embedding of the signals. We assess performance with a 5-fold cross-validation framework; that is, an 80/20 split between training and testing partitions. As an assessment of performance we use the macro-F1 score (the harmonic mean of the precision

and recall) [56]. This score was then averaged across each of the 5-folds and the whole process was repeated 100 times to produce a distribution of scores.

Figure F.3 shows that on this particular task the linear OT and linear TL^p framework have similar classification performance when differentiating between cancerous and normal cells. This is unsurprising as these images only have a single intensity channel. We also visualise, in a PCA plot, the linear TL^p embedding showing the variability across the dataset. Figure F.3 panel (c) visualises the modes of variation in image space for the first 5 principal components of the data.

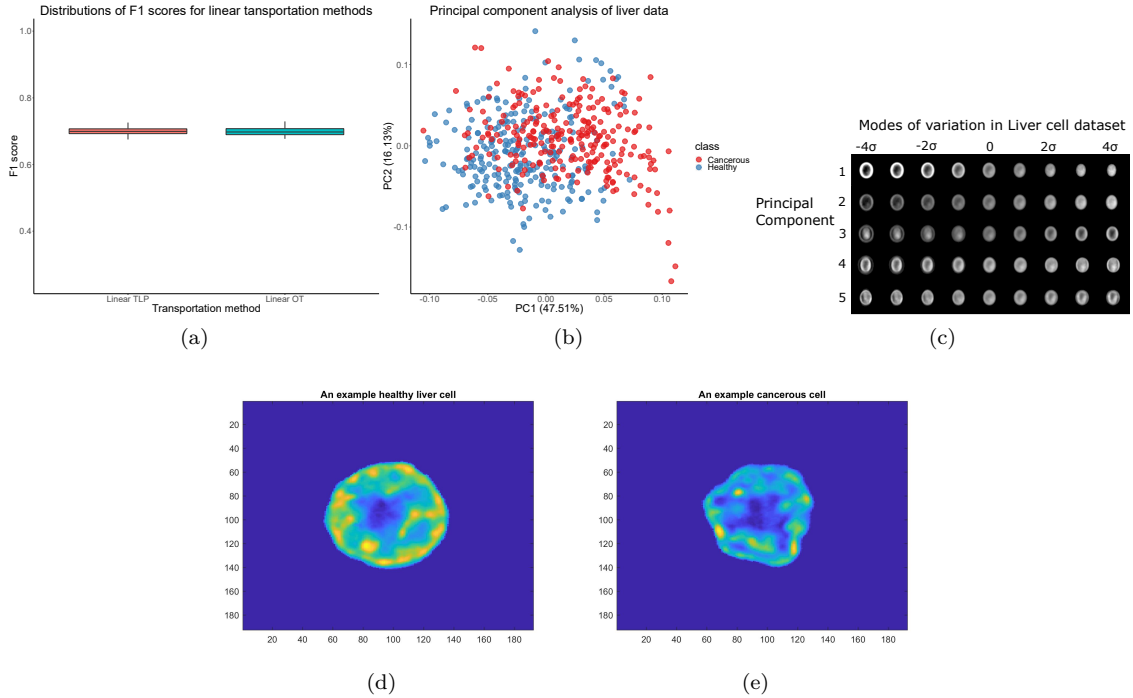


Figure F.3: (a) A boxplot showing the distribution of adjusted Rand index using the 1NN classifier on the linear OT and linear TL^p embeddings. Unsurprisingly, both methods perform equally well on this dataset. (b) A PCA plot of the TL^p linear embedding where we see overlap between the classes but distinct class distributions (c) The first 5 principal modes of variation using linear interpolation along the eigenvectors in PCA space. (d) An example healthy liver cell. (e) An example cancerous liver cell.

F.3.4 Application: Auslan data

We apply the transportation methodology presented in this manuscript to the Australian Sign language (Auslan) dataset [57]. A native Auslan signer was recorded, using fifth dimension technology gloves, making 95 different signs repeated over a period of 9 weeks. The sign was repeated 3 times at each recording, thus each word was measured 27 times. This means there are a total of 2565 signs in the dataset. Each measurement is considered as a multivariate time-series. The measurements taken for each hand are the x, y, z positions, along with roll, pitch and yaw. In addition, the bend of each of the 5 fingers is recorded. Thus at each frame 22 measurements are observed. We consider the Auslan data as the following functions $f_i : \mathbb{R} \to \mathbb{R}^{22}$, for $i = i, ..., 2565$. We truncate the number of frames to 44 because little variation was observed past this point.

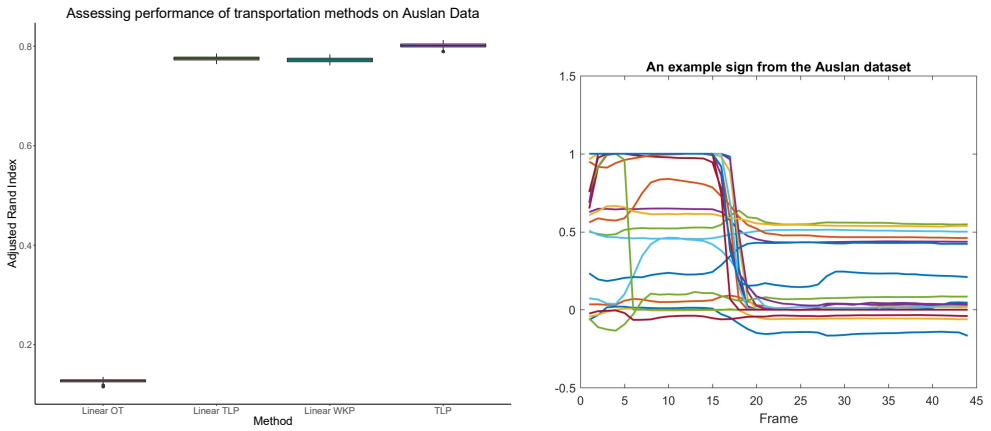
We apply the linear TL^p and linear OT frameworks to this dataset. Since TL^p can handle

multi-channel signals, no additional pre-processing was needed. However, to apply OT additional pre-processing was required. Firstly, all signals were made positive and then the mean was taken so that there was only a signal channel with positive values. We then normalised so the signal integrated to unity. A linear embedding was obtained as described in previous sections. Once this linear embedding is obtained, we use the 1 nearest neighbour (1NN) algorithm to predicted the signs from the linear embedding of the signals. We assess performance with a 5-fold cross-validation framework and repeat 100 times to produce a distribution of scores. In addition, we compare to the standard TL^p methodology and also linear $\text{TW}^{k,p}$ by including the gradient of the spatial positions x, y, z as additional information. We also recorded timings for each of the methods.

Figure F.4 shows that the linear transportation methods are considerably faster than the full transportation methods. Indeed the linear TL^p distance was on the order of magnitude of ten's of seconds, whilst the full TL^p took several hours. Figure F.5 demonstrates that our proposed linear TL^p method significantly outperforms the linear OT approach (T-test, $p < 10^{-4}$) on the Auslan data. We also see that providing derivative information of the signals does not lead to any improvement in performance. There is a loss in classification ability of the linear TL^p versus the TL^p method, which is unsurprising as the linear transportation method is approximate. However, on the Auslan dataset we observe this difference to be small and this minor improvement comes at computational cost orders of magnitude greater.

Application	Linear OT	Linear TL^p	Linear $\text{TW}^{k,p}$	TL^p
1-D Synthetic	1.3	0.2	-	90.2
2-D Synthetic	7.1	16.9	-	707.2
Cell Morphometry	242.1	512.7	-	161080
Auslan	12.1	13.0	13.5	91200
Breast Cancer Histopathology	25407.0	2919.8	-	345600

Figure F.4: CPU times in seconds to compute each transportation method on each dataset. Computation was halted after 4 CPU days.



(a) (b)

Figure F.5: (a) Distributions of macro-F1 scores for different transportations methods. Evidently linear TL^p significantly outperforms linear OT. (b) An example multi-variate signal from the Auslan dataset

F.3.5 Application: Breast cancer histopathology

In this section, we demonstrate the applicability of linear TL^p to images from breast cancer histopathology. Invasive Ductal Carcinoma is an aggressive and common form of breast cancer and deep learning based approaches have been used to construct classifiers to analyse such data [58], [59]. We analyse these datasets using transportation based approaches. We randomly sample 100 images each from two patients 1 healthy and 1 cancerous, totalling 200 images. Each image is on a 50×50 pixel grid. To apply linear OT to these images we first convert the images to a single intensity channel and renormalise so that the intensities integrate to unity. We apply the linear TL^p approach with $p = 2$ without any ad-hoc preprocessing, since it can be directly applied to un-normalised multi-channelled images. We compute transport maps in each case using entropy regularised approaches and then linearly embed these images as described in earlier sections. We visualise the linear embeddings using PCA. In addition, we again employ the 1NN classifier using the same framework as in the Auslan application. We report distributions of macro-F1 scores for both linear OT and linear TL^p.

Figure F.6 demonstrates clear differences between the linear TL^p embedding and linear OT embedding. The cancerous and healthy images separate more obviously in the PCA representation of the linear TL^p embeddings. This is supported when using the 1NN classifier, where a mean macro-F1 score of 0.70 is reported in the linear OT case, whilst for linear TL^p the mean macro-F1 score is 0.88 representing a greater than 25% improvement. It is clear from the box plots that linear TL^p outperforms linear OT (T-test $p < 10^{-16}$). Using linear interpolation, we visualise perturbations, in units of standard deviation, in principal component space as synthetic images (fig. F.7). The interpolation in the embedding produced by linear TL^p demonstrates localised mass moving from the centre of the image towards the edges. This corresponds to cancer invading the milk ducts in cancerous tissue with open milk ducts in non-cancerous tissue. Linear interpolation in the LOT embedding visualises mass moving from the lower right to the upper left of the plot, there is no physical interpretation for this variation. It is clear that the linear TL^p synthetic images are more interpretable than the LOT synthetic images.

F.3.6 Application: Financial time series

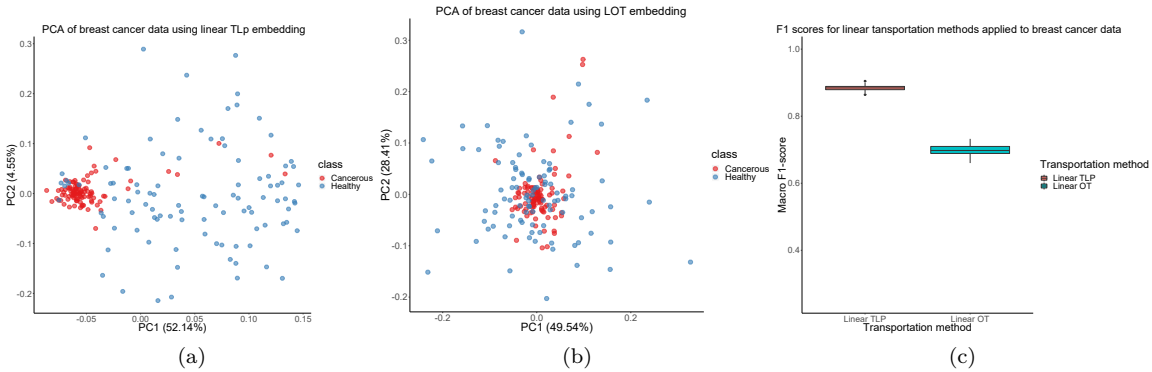
In this section, we consider an application of transportation distances to financial time series data. More specifically, we use daily close prices for constituents of the SP1500 index, for the period 2$^{\text{nd}}$ January 2004 - 23$^{\text{rd}}$ March 2020.

Experimental setup and background on financial time series

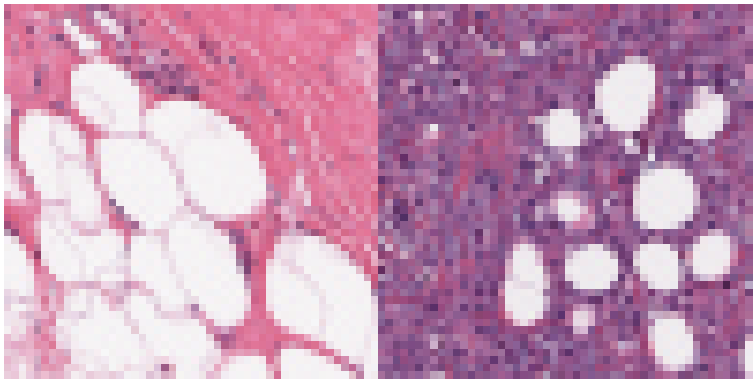
We only consider instruments (stocks) available throughout the entire history, which amounts to approximately $n = 1150$ names. We use daily log-returns, defined as follows. The return of instrument i, between times t_1 and t_2 is given by

$$\mathrm{R}_i^{(t_1,t_2)} = \log \frac{P_{i,t_2}}{P_{i,t_1}}, \tag{F.27}$$

where $P_{i,t}$ denotes the price of instrument i at the end of day t. On any given day t, we consider the matrix S_t of size $n \times m$, holding the daily returns for the past m days, with fixed $m = 20$



(a) (b) (c)



(d) (e)

Figure F.6: (a,b) PCA plots using the Linear TL^p embedding and the LOT embedding, we observe that Linear TL^p embedding separates classes. (c) Distributions of macro F1 scores for linear OT and TL^p embeddings in the application to breast cancer histopathology using the 1NN classifier. (d) An example image from a healthy patient. (e) An example image from a patient with breast cancer.

throughout the experiments. We refer to S_t as the sliding window, as we vary the time component. The goal is to use and compare various techniques for computing the k-nearest-neighbours of S_t, which we denote by $\mathcal{N}_k(S_t) = \{S_t^{(1)}, S_t^{(2)}, \ldots, S_t^{(k)}\}$, by pooling together the similarities between the multivariate time series comprising S_t with prior historical time series comprising $S_1, \ldots, S_{t-h}$, where h is the future horizon which we aim to predict. Note that, at any given time t, the available history to query for the k-nearest-neighbours of S_t ends at S_{t-h}, in order to avoid forward looking bias. Figure F.8 is a schematic diagram of our pipeline process.

Future returns For ease of notation and referencing, we use the abbreviation 'fret' to denote forward looking returns (often referred to as *targets*), and include in the name of the return additional information to reflect the type of return being used. The future return can be the raw returns itself, or various decompositions of it. We let $\text{fret}_{i,t}^{(h)}$, or for simplicity of notation $\text{f}_{i,t}^{(h)}$, denote the raw return of instrument i at time t, with future horizon h

$$\text{f}_{i,t}^{(h,RR)} = \log \frac{P_{i,t+h}}{P_{i,t}}. \tag{F.28}$$

In our setting, we consider the following future horizons $h \in \{1, 3, 5, 10\}$. One often uses the S&P500 as a proxy for the entire market return. The S&P500 is a stock market index that measures

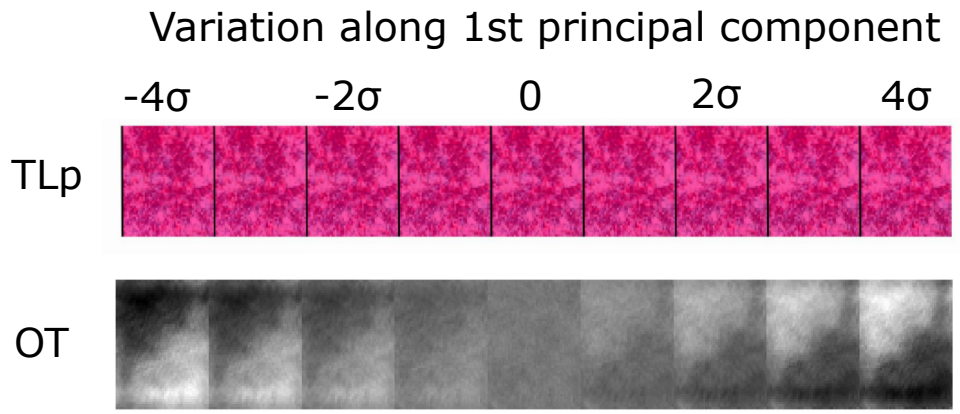


Figure F.7: Variations along the first principal component in the application to breast cancer histopathology for both the linear TL^p embedding and the linear OT embedding.

the stock performance of 500 large companies listed on stock exchanges in the US market. The index has a corresponding ETF (Exchange Traded Fund), which can be traded much like any other regular stock. The symbol of this instrument is denoted by SPY, and we denote its raw return by

$$\mathrm{f}_{SPY,t}^{(h,RR)} = \log \frac{P_{SPY,t+h}}{P_{SPY,t}}. \tag{F.29}$$

With this in mind, for any instrument i, one can then consider its future market excess return (MR)

$$\mathrm{f}_{i,t}^{(h,MR)} = \mathrm{f}_{i,t}^{(h,RR)} - \beta_i \mathrm{f}_{SPY,t}^{(h,RR)}. \tag{F.30}$$

For simplicity, we assume $\beta_i = 1$ across all instruments $i = 1, \ldots, n$, though there are various techniques to infer the individual betas from historical prices data. We remark that the main reason for benchmarking our predictions against the market-excess return, as opposed to only the raw return, is essentially to hedge away the market risk and increase the Sharpe Ratio score defined further below.

Estimates of future returns Once we have identified the k-nearest-neighbour periods of the current window S_t, the prediction made at time t, for a given horizon h, is a weighted sum of the corresponding historical future returns, where the weights are inversely proportional to the distances. More precisely, if we denote by $\hat{f}_t^{(h,RR)}$ the $n \times 1$ vector of forecasts for the future h-day raw returns at time t, its values are given by

$$\hat{f}_t^{(h,RR)} = \sum_{i \in \mathcal{N}_k(S_t)} w_i f_i^{(h,RR)}, \tag{F.31}$$

where the weights w_i are given by $w_i = \frac{1}{d(S_t,S_i)}$, normalized such that $\sum_i w_i = 1$, and $d(S_t, S_i)$ denotes the distance between the current window S_t, and its nearest historical neighbours $S_i, i \in \mathcal{N}_k(S_t)$. Similarly, we compute estimates for the future h-day market excess returns (MR), by pooling together the historical forward looking market excess returns of the k-nearest-neighbour periods.

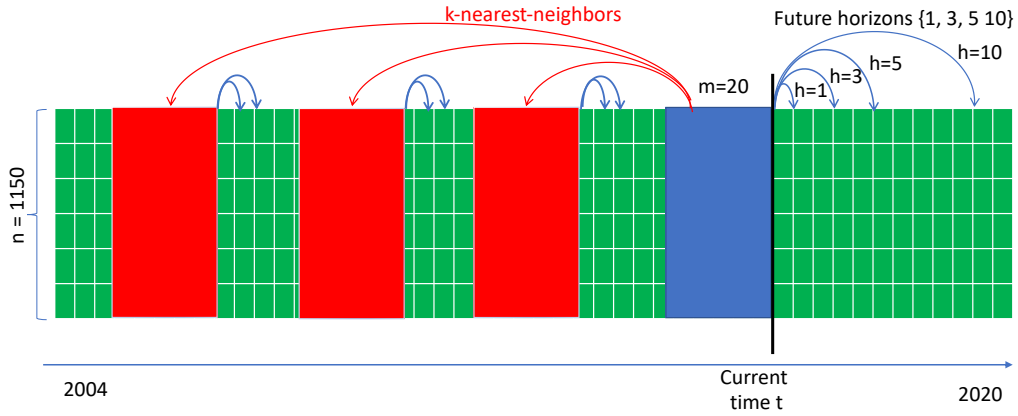


Figure F.8: Schematic diagram of the pipeline for the financial time series application. We use a sliding window approach, where the current window is shown in blue. The k-nearest-neighbor windows from the past are shown in red.

P&L For the evaluation of performance, we rely on standard metrics from the finance literature. If α denotes our forecast, then the corresponding PnL (Profit and Loss) on a given day t is given by

$$PnL_t = \sum_{i=1}^{n} \text{sign}(\alpha_{i,t}) \cdot \mathrm{f}_{i,t}, \quad t = 1, \ldots, T, \tag{F.32}$$

where $\alpha_{i,t}$ denotes our forecast for instrument i on day t. Note that the PnL increases if and only if the sign of the forecast α agrees with the sign of the future return $\mathrm{f}_{i,t}$, and decreases otherwise. The sum is across all the n instruments, and $\mathrm{f}_{i,t}$ is the future return (either raw return (RR) or market-excess return (MR)) of instrument i on day t. We explore different forward looking horizon windows $h \in \{1, 3, 5, 10\}$). We add a superscript to the PnL calculation to indicate its dependency on horizon h and the type of return considered (RR or MR). For instance, in our setting, the h-day forward looking market-excess return, computed daily, is given by

$$PnL_t^{(h,MR)} = \sum_{i=1}^{n} \text{sign}\left(\hat{f}_{i,t}^{(h,MR)}\right) \cdot \mathrm{f}_{i,t}^{(h,MR)}, \quad t = 1, \ldots, T, \quad h \in \{1, 3, 5, 10\}. \tag{F.33}$$

Sharpe Ratio Once the daily PnL time series has been computed for all available days in the study (in a rolling window approach), we capture the risk-adjusted performance by computing the corresponding (annualized) Sharpe Ratio (SR) as

$$SR = \frac{\text{mean}(PnL)}{\text{stdev}(PnL)} * \sqrt{252}, \tag{F.34}$$

where the scaling is due to the fact that there are 252 trading days within a calendar year. For simplicity, we apply the same scaling $\sqrt{252}$ also to the longer horizons $h > 1$, and refer the reader to [60] for an in depth discussion Sharpe Ratios[1]. For simplicity, we choose to use the same normalizing constant even for the longer horizons, and we attribute the h-day PnL to each day t (leading to overlapping windows); however, we are primarily interested in relative performance

[1]The annualized Sharpe Ratio is calculated from daily observations as $\frac{\mu - r_f}{\sigma}\sqrt{252}$, where μ is the average daily PnL return, r_f denotes the risk-free rate, and σ the standard deviation of the PnL returns. Since the risk-free rate is typically very close to zero over the period of study, we compute the Sharpe Ratio as $\frac{\mu}{\sigma} \cdot \sqrt{252}$.

of the methods, in terms of Sharpe Ratio and PnL, and less on the actual magnitudes of these performance metrics.

Average PnL in basis points In the financial literature, a typical performance measure is the average return per dollar traded, in percentage. For example, one is typically interested in the annualized return of the portfolio. In what follows, we denote by PPT (Pnl Per Trade) the average daily PnL per unit of notional. For instance, if at time t_0 the available capital is \$100, and the cumulative PnL at time t_{252} (thus after one year) is \$10, then the annualized return amounts to 10%. Recalling that 1% amounts to 100 basis points (bpts), an annualized return of 10% translates to approximately PPT = 4 bpts per day (since $4 \times 252 \approx 1000$ bpts, which amounts to 10%). Essentially, the PPT is telling us how much would we earn for each \$1 traded in the markets (excluding transaction costs). For simplicity, we ignore sizing effects and assume that each day, we invest \$1 for each of the n instruments, which leads us to the following simplified notion of PPT, averaged over the entire trading period comprised of T days

$$PPT = \sum_{t=1}^{T} \frac{PnL_t}{n} = \frac{\sum_{t=1}^{T} PnL_t}{Tn}. \tag{F.35}$$

Quintile portfolios One is often interested in understanding the performance of the forecasts, as a function of their respective magnitudes. To this end, one typically considers only a subset (eg, top q% strongest in magnitude forecasts) of the universe of instruments, usually referred to as *quantile* portfolios in the literature [61]. Assume for example a universe of 500 stocks, from the SP500 index. A quantile-based analysis simply constructs and evaluates portfolios composed of instruments which fall in a specific quantile bucket, or above a quantile threshold. We choose to use upward-contained quintile buckets, which we denote by $qr_i, i = 1, 2, \ldots, 5$ indicating the quintile rank of each instrument, meaning that instruments with quintile rank qr_i correspond to the top $1 - \frac{i-1}{5}$ fraction of largest-in-magnitude forecasts. The colours in fig. F.12 denote the quintile portfolios traded based on the magnitude of the forecasts. For example, the red bars qr_1 correspond to the full universe of instruments, while the green bars qr_4 denote the top 40% largest in magnitude forecasts.

Methods comparison

We compare the prediction performance of pairwise OT (which we denote as POT), linear OT (LOT), and linear TL^p (LTLP), and leave out the full TL^p due to its prohibitive computational running cost. In addition, we compare to another more classical approach, not relying on transportation distance methodology, given by simple Pearson correlation between the original time series. More precisely, for each window S_t (matrix of size $n \times m$, $t = 1, \ldots, T$, we first standardize the returns in each row (eg, corresponding to each instrument), and denote the resulting matrix by $\tilde{S}_t, t = 1, \ldots, T$. Next, we unwrap each matrix $\tilde{S}_t$ into a vector $\psi_t \in \mathbb{R}^{nm}$, and finally compute the pairwise distance between a pair of time windows S_i and S_j using a correlation-based distance between their corresponding flattened versions

$$\text{COR}_{ij} = 1 - Corr(\psi_i, \psi_j). \tag{F.36}$$

Note that, in light of the pre-processing step that standardized the instruments, this corresponds, up to a scaling constant, to the squared Euclidean distance between the corresponding vectors $||\psi_i - \psi_j||_F^2$.

Figure F.11 shows numerical results comparing the various methods considered. The left column is a heatmap showing the $T \times T$ pairwise distance matrix between all days available in history, in the interval 2004 - 2020. We note that LTLP clearly highlights the financial crisis occurred in 2008, followed by LOT, and to some extent, POT, while COR show barely visible signs of this event. The middle columns show a distribution of the pairwise distances, while the right columns show the row sums of the distance matrix. Construing the distance matrix as a network with distance/dissimilarity information, this plot effectively plots the degree of each node (i.e., of each time period corresponding to a sliding window of length 20). The periods of time with the largest dissimilarity degree correspond to the financial crisis in 2008. Note that, for each of visualization, we standardize the degree vector.

Next, we zoom in into the LTLP pairwise distance matrix, and show the resulting degrees in fig. F.9. After computing the LTLP distance matrix, we interpret this as a distance network, and compute the total distance degree of each node, after standardization. In this plot, we are able to recognize many of the major financial market events that have happened over the last two decades: the big financial crisis of 2007-2008, the 2010 Flash crash, the August 2011 markets fall (between May-October 2011), the Chinese market crash from January 2016, the period Oct-Nov 2018 (when the stock market lost more than $2 trillion), August 2019 (a highly volatile month in the global stock markets), and finally, the February 2020 stock market crash triggered by the COVID-19 pandemic. It is interesting to observe that the distance degree corresponding to the COVID-19 pandemic is matched in magnitude only by the 2007-2008 financial crisis. Furthermore, in fig. F.10 we plot the top 50 eigenvalues of the LTLP distance matrix and a heatmap of the top 5 eigenvectors, altogether highlighting the usual *market mode* top eigenvector, and the second eigenvector highly localized on the 2007-2008 financial crisis and the February-March 2020 Covid-19 pandemic period.

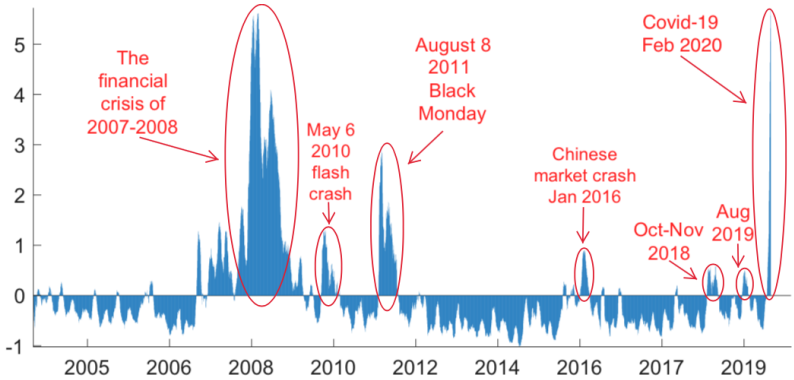


Figure F.9: Normalized total distance for each day, as computed via LTLP , annotated with the major market events. More explicitly, we compute the distance matrix between all pairs of days (where the data for a given by is given by the previous $m = 20$ days, including the day of), construe this as a distance network, and compute the total degree of each node (which we show in the above figure, after standardization, for ease of visualization).

Figure F.12 shows portfolio statistics for the various methods, across different target future horizons $h \in \{1, 3, 5, 10\}$, for both raw-returns (RR) and market-excess returns (MR). The cor-

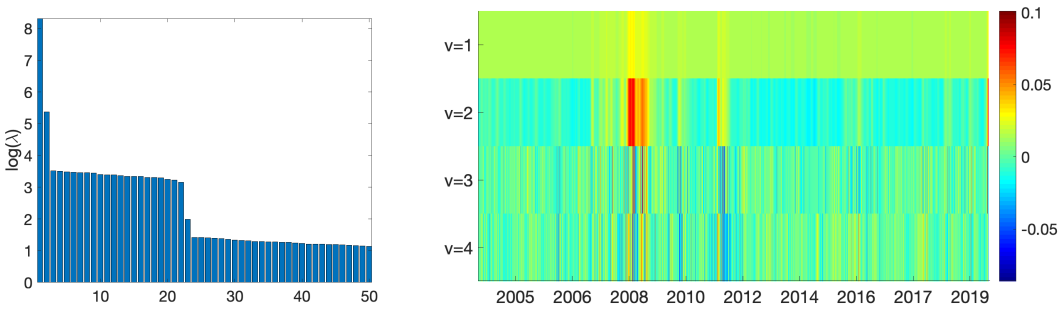


Figure F.10: Left: histogram of the top 50 eigenvalues of the LTLP distance matrix, on a log scale. Right: barplot of the top $k = 5$ eigenvectors. The top eigen-pair typically corresponds to the so-called market mode. The second eigenvector is strongly localized on the 2007-2008 financial crisis and the February-March 2020 Covid-19 crisis.

responding cumulative PnL plots across time are given in figs. F.13 to F.15 for future horizons $h \in \{1, 5, 10\}$, respectively. Here, we fixed the number of nearest neighbors to $k = 100$, and allow the knn search to span back until the start of the available history $T = 1$. When forecasting raw returns (left column in fig. F.12), all methods perform rather poorly, with COR and LOT showing the best performance for $h = 1$, while for $h \in \{3, 5, 10\}$, LTLP clearly outperforms all other methods.

In the market-excess returns setting, for $h = 1$, all methods return a similar performance in terms of Sharpe Ratio (SR) around 1, except for POT which has a SR of around 0.5; in terms of PnL, most methods achieve a PPT of 1-3 basis points (bpts). However, for longer horizons, LTLP clearly outperforms all other methods, both in terms of Sharpe Ratio and PPT.

Finally, in fig. F.16, we displayed the portfolio statistics for the various methods, but when the k-nearest neighbors algorithm used $k = 50$ and the historical window available to the knn query is restricted to only the previous 1000 days. Results are qualitatively similar to the previous setting, with a slight improvement for the LOT method at the longer horizons.

References

[1] S. Kolouri, S. Park, M. Thorpe, D. Slepčev, and G. K. Rohde, "Optimal mass transport: Signal processing and machine-learning applications," *IEEE Signal Processing Magazine*, vol. 34, no. 4, pp. 43–59, 2017.

[2] S. Basu, S. Kolouri, and G. K. Rohde, "Detecting and visualizing cell phenotype differences from microscopy images using transport-based morphometry," *Proceedings of the National Academy of Sciences*, vol. 111, no. 9, pp. 3448–3453, 2014.

[3] S. Kolouri and G. K. Rohde, "Transport-based single frame super resolution of very low resolution face images," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015, pp. 4876–4884.

[4] U. Frisch, S. Matarrese, R. Mohayaee, and A. Sobolevski, "A reconstruction of the initial conditions of the universe by optimal mass transportation," *Nature*, vol. 417, no. 6886, p. 260, 2002.

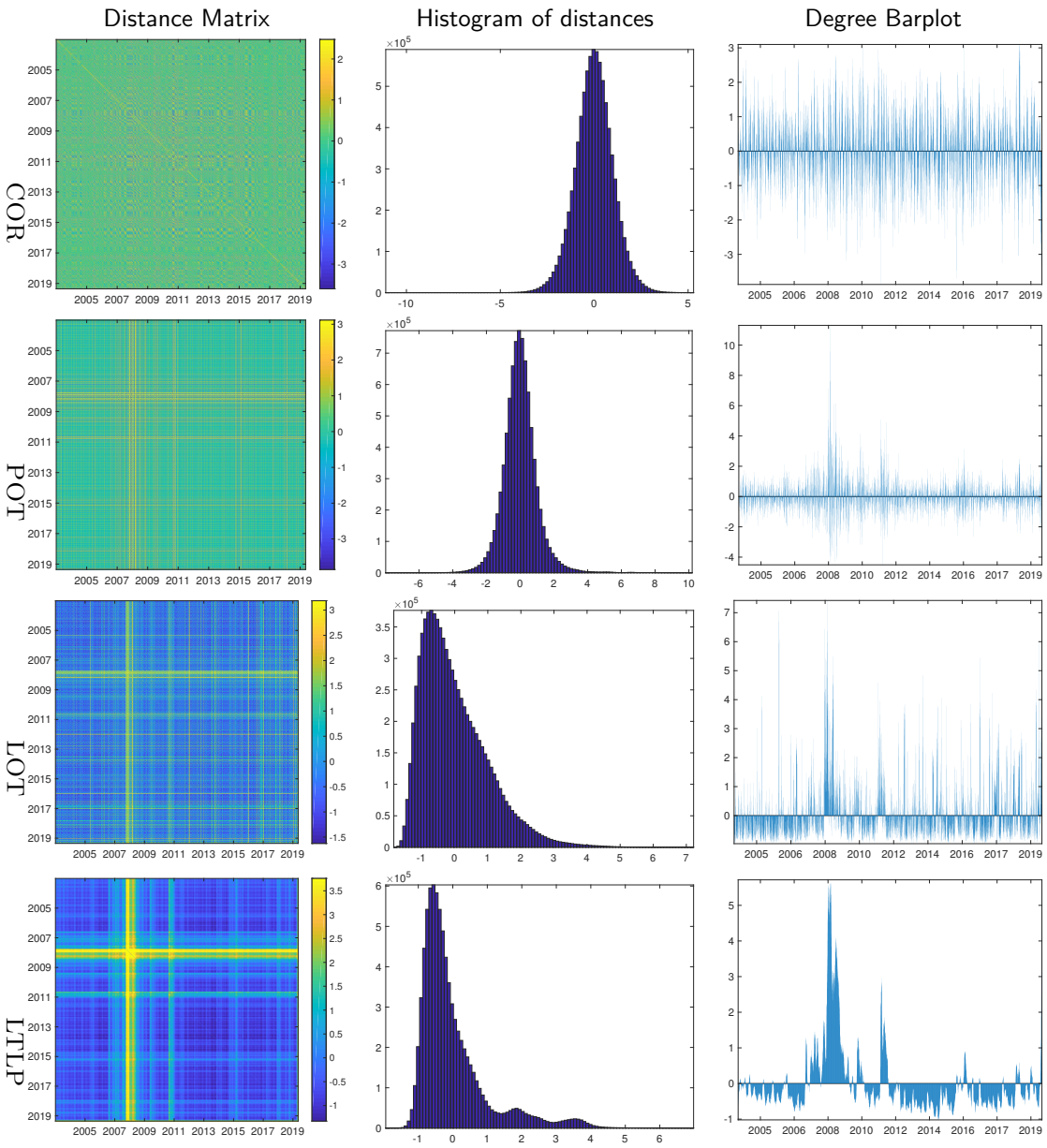


Figure F.11: Left: pairwise distance matrices corresponding to each of the methods considered. Middle: histogram of the entries above the main diagonal, in each distance matrix. Right: column sums of each distance matrix, standardized for ease of visualization.

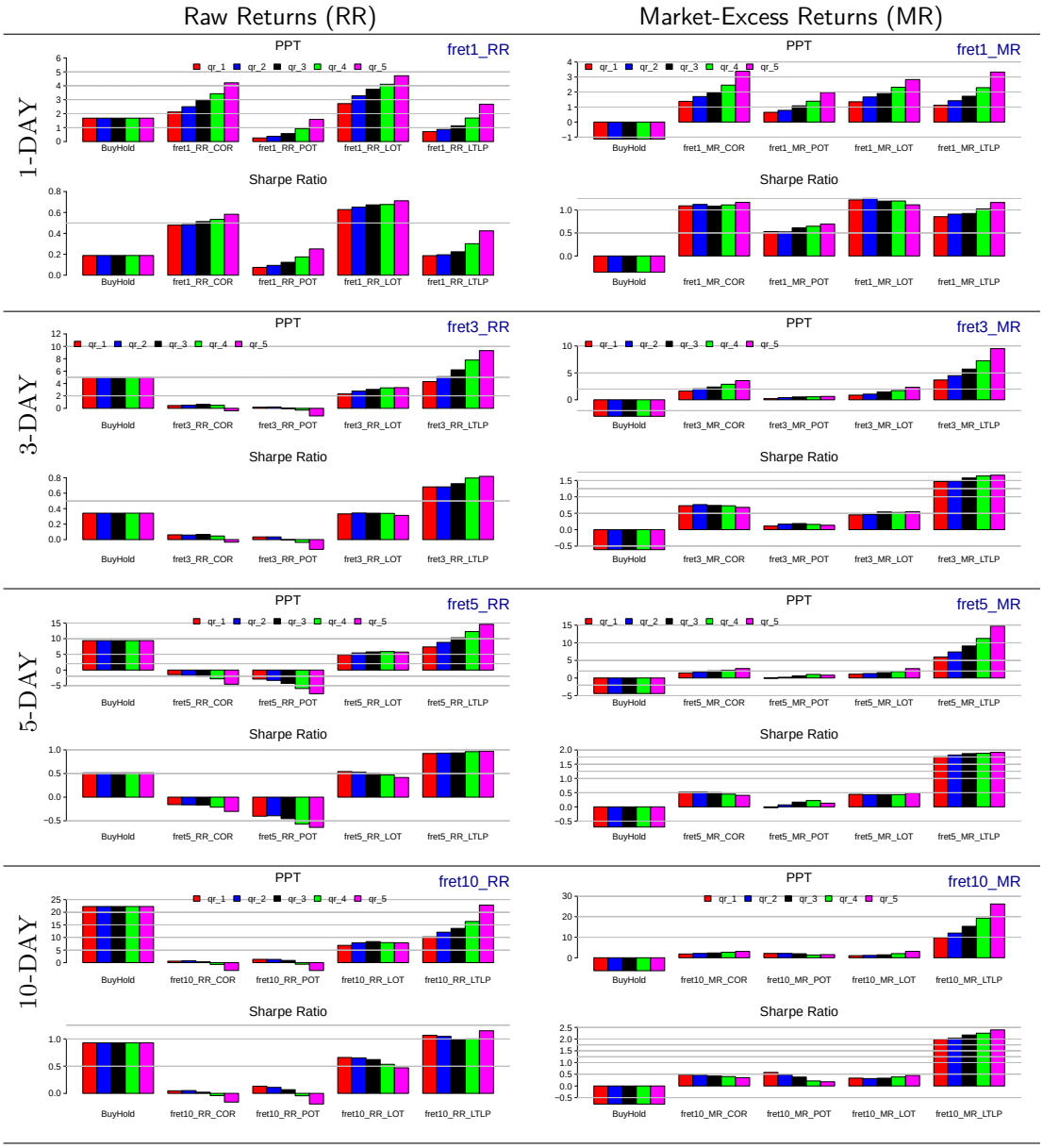


Figure F.12: Portfolio statistics (PPT PnL and Sharpe Ratio) across various future horizons $h \in \{1, 3, 5, 10\}$ (indexing the rows), across two types of returns: raw returns and market excess returns (indexing the columns). The colours denote quintile portfolios. The x-axis denotes the forecasts made by each method {COR, POT, LOT, LTLP}. BuyHold denotes the naive approach where all forecasts are equal to 1.

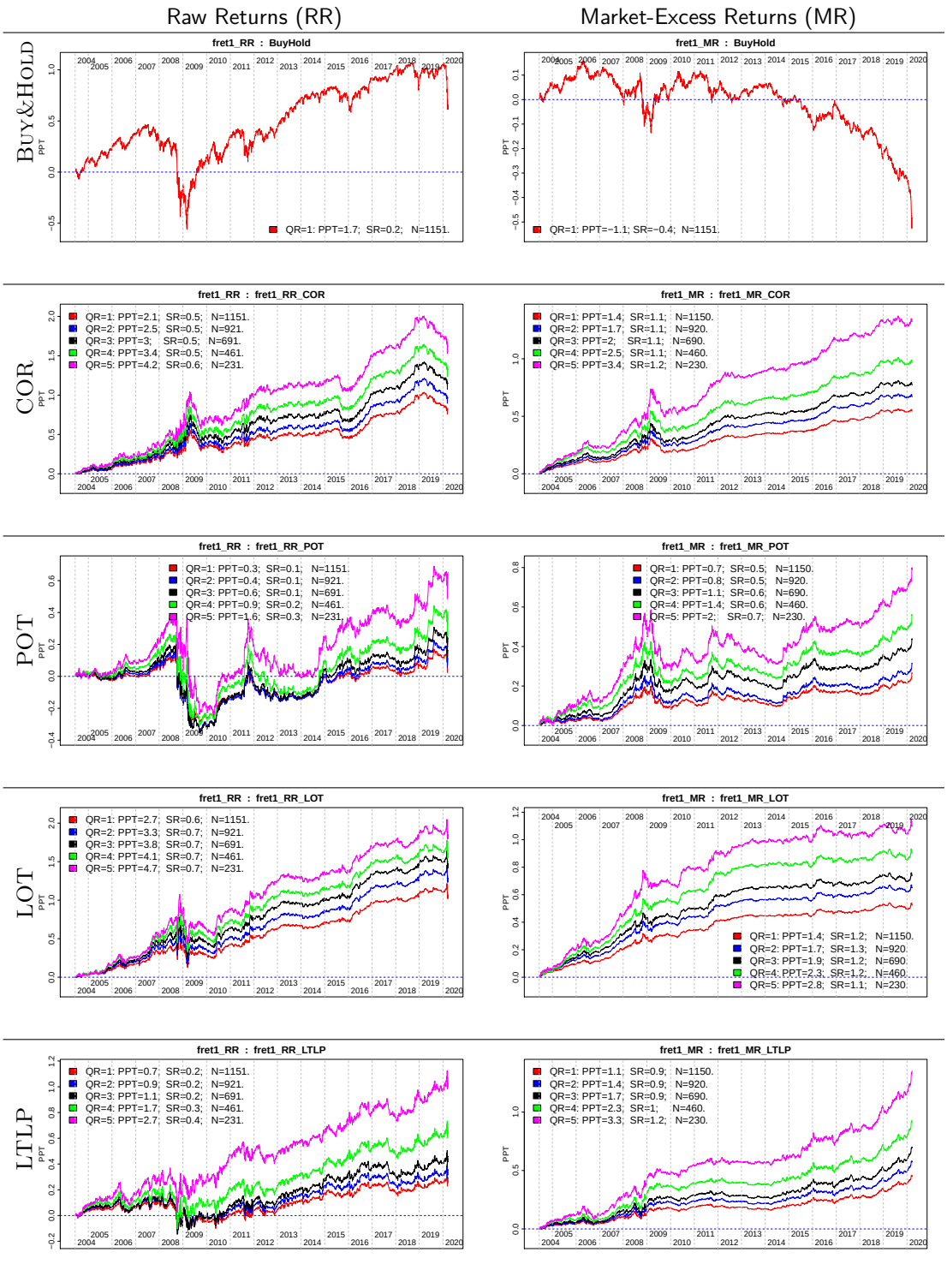


Figure F.13: Cumulative PnL for the **1-day** future horizon, for two types of returns (raw returns and market excess returns, indexing the columns), across different methods.

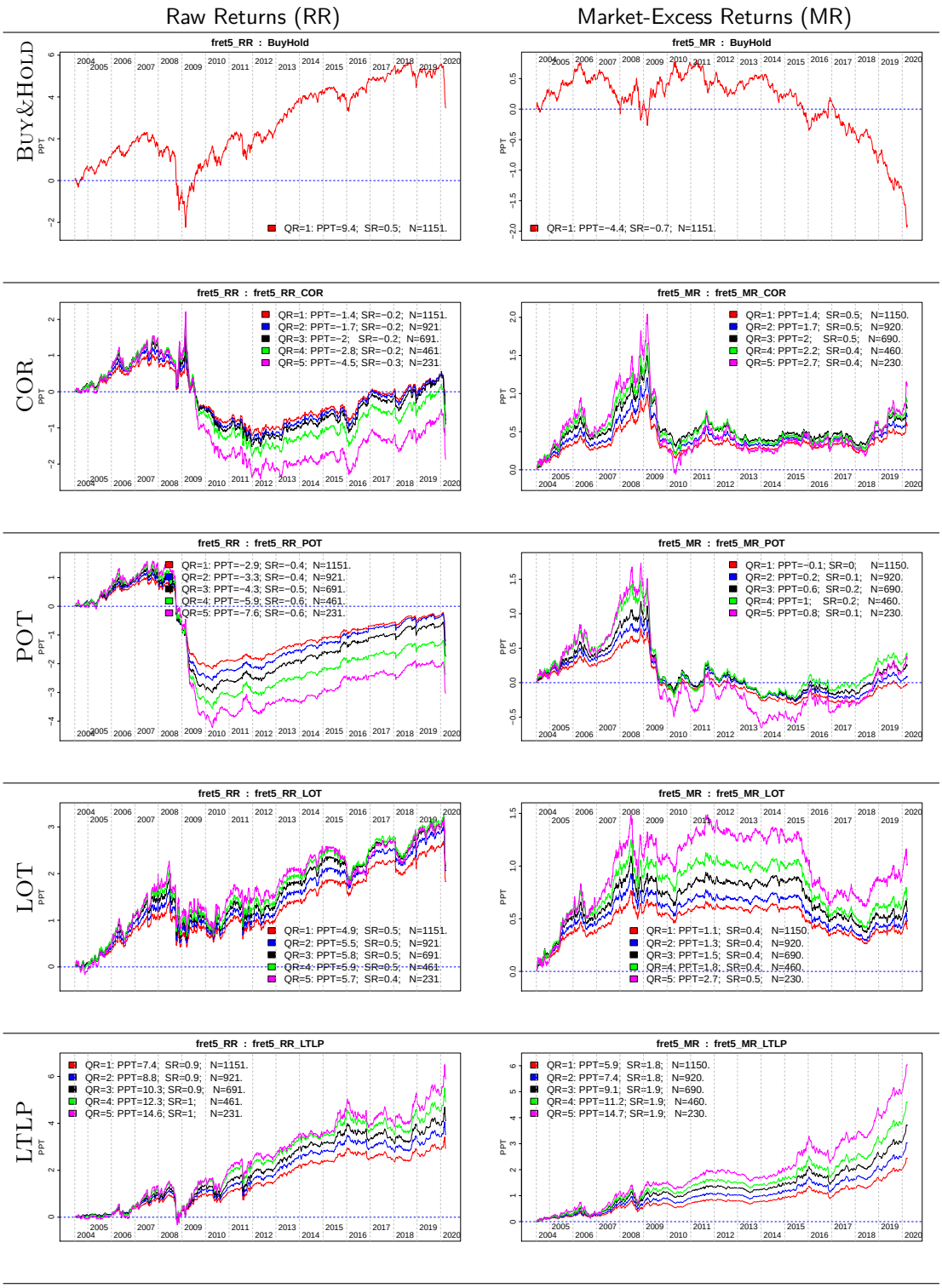


Figure F.14: Cumulative PnL for the **5-day** future horizons, for raw returns and market excess returns (indexing the columns), across different methods.

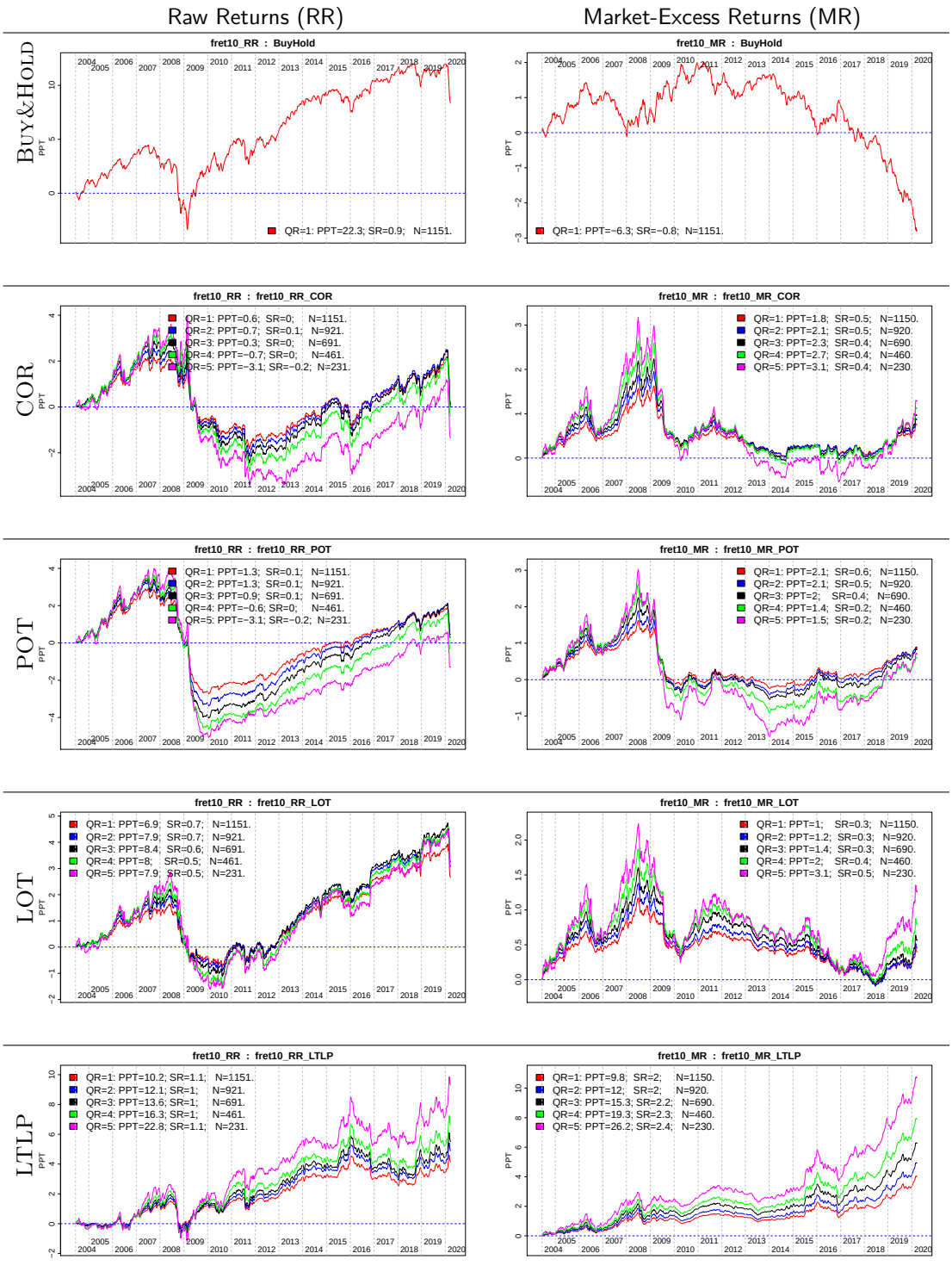


Figure F.15: Cumulative PnL for the **10-day** future horizons, for two types of returns (raw returns and market excess returns, indexing the columns), across different methods.

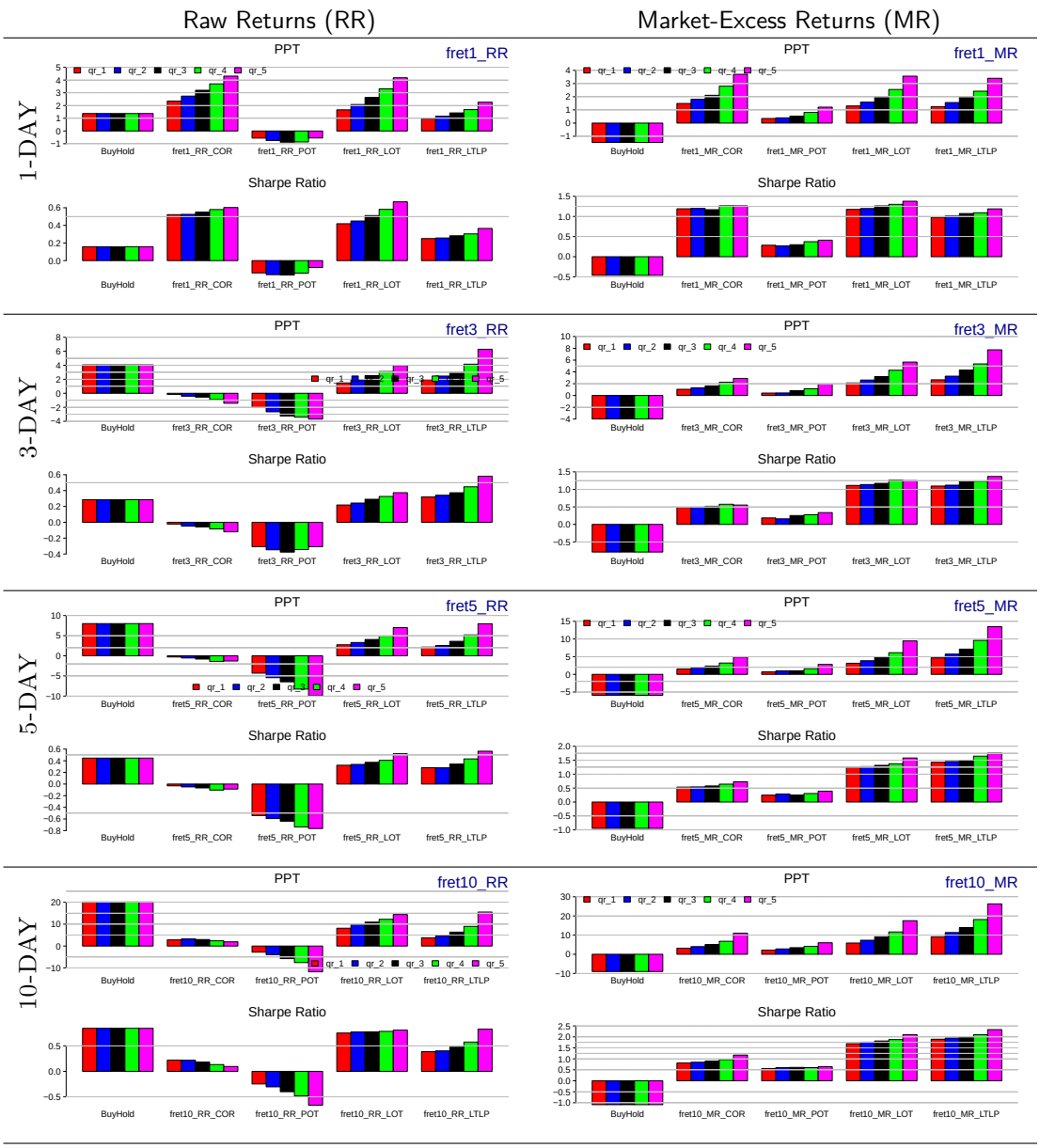


Figure F.16: Portfolio statistics (PPT PnL and Sharpe Ratio) across various future horizons $h \in \{1, 3, 5, 10\}$ (indexing the rows), across two types of returns: raw returns and market excess returns (indexing the columns). The colours denote quintile portfolios. The x-axis denotes the forecasts made by each method $\{\text{COR}, \text{POT}, \text{LOT}, \text{LTLP}\}$. BuyHold denotes the naive approach where all forecasts are equal to 1. Here, we use $k = 50$ (nearest neighbors) and restrict the look-back historical window to the previous $H = 1000$ days.

[5] U. Frisch and A. Sobolevskii, "Application of optimal transport theory to reconstruction of the early universe," *Journal of Mathematical Sciences*, vol. 133, no. 4, pp. 1539–1542, 2006.

[6] C. Frogner, C. Zhang, H. Mobahi, M. Araya, and T. A. Poggio, "Learning with a Wasserstein loss," in *Advances in Neural Information Processing Systems*, 2015, pp. 2053–2061.

[7] S. Minsker, S. Srivastava, L. Lin, and D. B. Dunson, "Robust and scalable Bayes via a median of subset posterior measures," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 4488–4527, 2017.

[8] J. Solomon, F. De Goes, G. Peyré, M. Cuturi, A. Butscher, A. Nguyen, T. Du, and L. Guibas, "Convolutional Wasserstein distances: Efficient optimal transportation on geometric domains," *ACM Transactions on Graphics (TOG)*, vol. 34, no. 4, 2015.

[9] S. Srivastava, V. Cevher, Q. Dinh, and D. Dunson, "Wasp: Scalable Bayes via barycenters of subset posteriors," in *Artificial Intelligence and Statistics*, 2015, pp. 912–920.

[10] S. Srivastava, C. Li, and D. B. Dunson, "Scalable Bayes via barycenter in Wasserstein space," *The Journal of Machine Learning Research*, vol. 19, no. 1, pp. 312–346, 2018.

[11] T. A. El Moselhy and Y. M. Marzouk, "Bayesian inference with optimal maps," *Journal of Computational Physics*, vol. 231, no. 23, pp. 7815–7850, 2012.

[12] S. Haker, L. Zhu, A. Tannenbaum, and S. Angenent, "Optimal mass transport for registration and warping," *International Journal of computer vision*, vol. 60, no. 3, pp. 225–240, 2004.

[13] L. Zhu, Y. Yang, S. Haker, and A. Tannenbaum, "An image morphing technique based on optimal mass preserving mapping," *IEEE transactions on image processing*, vol. 16, no. 6, pp. 1481–1495, 2007.

[14] C. Villani, *Topics in Optimal Transportation*. American Mathematical Society, 2003.

[15] ——, *Optimal transport: old and new*. Springer Science & Business Media, 2008.

[16] F. Santambrogio, *Optimal transport for applied mathematicians*. Springer, 2015.

[17] W. Gangbo and R. J. McCann, "The geometry of optimal transportation," *Acta Mathematica*, vol. 177, no. 2, pp. 113–161, 1996.

[18] L. Ambrosio, N. Gigli, and G. Savaré, *Gradient flows: in metric spaces and in the space of probability measures*. Springer Science & Business Media, 2008.

[19] A. M. Oberman and Y. Ruan, *An efficient linear programming method for optimal transportation*, 2015. arXiv: 1509.03668.

[20] S. Angenent, S. Haker, and A. Tannenbaum, "Minimizing flows for the Monge–Kantorovich problem," *SIAM journal on mathematical analysis*, vol. 35, no. 1, pp. 61–97, 2003.

[21] E. Haber, T. Rehman, and A. Tannenbaum, "An efficient numerical method for the solution of the L_2 optimal mass transfer problem," *SIAM Journal on Scientific Computing*, vol. 32, no. 1, pp. 197–211, 2010.

[22] R. Chartrand, B. Wohlberg, K. Vixie, and E. Bollt, "A gradient descent solution to the Monge–Kantorovich problem," *Applied Mathematical Sciences*, vol. 3, no. 22, pp. 1071–1080, 2009.

[23] Q. Mérigot, "A multiscale approach to optimal transport," *Computer Graphics Forum*, vol. 30, no. 5, pp. 1583–1592, 2011.

[24] S. Gerber and M. Maggioni, "Multiscale strategies for computing optimal transport," *The Journal of Machine Learning Research*, vol. 18, no. 1, pp. 2440–2471, 2017.

[25] M. Cuturi, "Sinkhorn distances: Lightspeed computation of optimal transport," in *Advances in neural information processing systems*, 2013, pp. 2292–2300.

[26] M. Cuturi and A. Doucet, "Fast computation of Wasserstein barycenters," in *International Conference on Machine Learning*, 2014, pp. 685–693.

[27] J.-D. Benamou, G. Carlier, M. Cuturi, L. Nenna, and G. Peyré, "Iterative Bregman projections for regularized transportation problems," *SIAM Journal on Scientific Computing*, vol. 37, no. 2, A1111–A1138, 2015.

[28] A. Genevay, M. Cuturi, G. Peyré, and F. Bach, "Stochastic optimization for large-scale optimal transport," in *Advances in Neural Information Processing Systems*, 2016, pp. 3440–3448.

[29] W. Wang, D. Slepčev, S. Basu, J. A. Ozolek, and G. K. Rohde, "A linear optimal transportation framework for quantifying and visualizing variations in sets of images," *International Journal of Computer Vision*, vol. 101, no. 2, pp. 254–269, 2013.

[30] S. Kolouri, A. . Tosun, J. A. Ozolek, and G. K. Rohde, "A continuous linear optimal transport approach for pattern analysis in image datasets," *Pattern Recognition*, vol. 51, pp. 453–462, 2016.

[31] N. García Trillos and D. Slepčev, "Continuum limit of total variation on point clouds," *Archive for rational mechanics and analysis*, vol. 220, no. 1, pp. 193–241, 2016.

[32] M. Thorpe and D. Slepčev, "Transportation L^p distances: Properties and extensions," *In Preparation*, 2017.

[33] M. Thorpe, S. Park, S. Kolouri, G. K. Rohde, and D. Slepčev, "A transportation L^p distance for signal analysis," *Journal of Mathematical Imaging and Vision*, vol. 59, no. 2, pp. 187–210, 2017.

[34] W. Gangbo and R. J. McCann, "Optimal maps in Monge's mass transport problem," *Comptes Rendus de l'Academie des Sciences-Serie I-Mathematique*, vol. 321, no. 12, pp. 1653–1658, 1995.

[35] F. Otto, "The geometry of dissipative evolution equations: The porous medium equation," *Comm. Partial Differential Equations*, vol. 26, pp. 101–174, 2001.

[36] M. M. Dunlop, D. Slepčev, A. M. Stuart, and M. Thorpe, "Large data and zero noise limits of graph-based semi-supervised learning algorithms," *Applied and Computational Harmonic Analysis*, 2019, ISSN: 1063-5203.

[37] D. Slepcev and M. Thorpe, "Analysis of p-laplacian regularization in semisupervised learning," *SIAM Journal on Mathematical Analysis*, vol. 51, no. 3, pp. 2085–2120, 2019.

[38] R. Cristoferi and M. Thorpe, "Large data limit for a phase transition model with the p-laplacian on point clouds," *European Journal of Applied Mathematics*, pp. 1–47,

[39] M. Thorpe and F. Theil, "Asymptotic analysis of the Ginzburg-Landau functional on point clouds," *Proceedings of the Royal Society of Edinburgh Section A: Mathematics*, vol. 149, no. 2, pp. 387–427, 2019.

[40] B. Osting and T. Reeb, "Consistency of Dirichlet partitions," *SIAM Journal on Mathematical Analysis*, vol. 49, no. 5, pp. 4251–4274, 2017.

[41] N. García Trillos and D. Slepčev, "A variational approach to the consistency of spectral clustering," *Applied and Computational Harmonic Analysis*, vol. 45, no. 2, pp. 239–381, 2018.

[42] N. García Trillos, D. Slepčev, J. von Brecht, T. Laurent, and X. Bresson, "Consistency of cheeger and ratio graph cuts," *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 6268–6313, 2016.

[43] N. García Trillos, D. Slepčev, and J. von Brecht, "Estimating perimeter using graph cuts," *Advances in Applied Probability*, vol. 49, no. 4, pp. 1067–1090, 2017.

[44] N. García Trillos and R. Murray, "A new analytical approach to consistency and overfitting in regularized empirical risk minimization," *European Journal of Applied Mathematics*, vol. 28, no. 6, pp. 886–921, 2017.

[45] N. García Trillos, "Variational limits of k-nn graph based functionals on data clouds," *SIAM Journal on Mathematics of Data Science*, vol. 1, no. 1, pp. 93–120, 2019.

[46] N. García Trillos and D. Sanz-Alonso, "Continuum limit of posteriors in graph Bayesian inverse problems," *SIAM Journal on Mathematical Analysis*, vol. 50, no. 4, pp. 4020–4040, 2018.

[47] N. García Trillos, Z. Kaplan, T. Samakhoana, and D. Sanz-Alonso, *On the consistency of graph-based Bayesian learning and the scalability of sampling algorithms*, 2017. arXiv: 1710.07702.

[48] N. G. Trillos, M. Gerlach, M. Hein, and D. Slepčev, "Error estimates for spectral convergence of the graph Laplacian on random geometric graphs toward the Laplace–Beltrami operator," *Foundations of Computational Mathematics*, pp. 1–61, 2019.

[49] J. H. Fitschen, F. Laus, and B. Schmitzer, "Optimal transport for manifold-valued images," in *Scale Space and Variational Methods in Computer Vision*, 2017, pp. 460–472.

[50] W. Wang, Y. Mo, J. A. Ozolek, and G. K. Rohde, "Penalized fisher discriminant analysis and its application to image-based morphometry," *Pattern Recognition Letters*, vol. 32, no. 15, pp. 2128–2135, 2011.

[51] S. Park and M. Thorpe, "Representing and learning high dimensional data with the optimal transport map from a probabilistic viewpoint," in *The IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, pp. 7864–7872.

[52] G. Peyré and M. Cuturi, *Computational Optimal Transport*, ser. Volume 37 of Foundations and Trends in Machine Learning Series. Now Publishers, 2019.

[53] G. Carlier, V. Duval, G. Peyré, and B. Schmitzer, "Convergence of entropic schemes for optimal transport and gradient flows," *SIAM Journal on Mathematical Analysis*, vol. 49, no. 2, pp. 1385–1418, 2017.

[54] W. M. Rand, "Objective criteria for the evaluation of clustering methods," *Journal of the American Statistical Association*, vol. 66, no. 336, pp. 846–850, 1971.

[55] L. Hubert and P. Arabie, "Comparing partitions," *Journal of Classification*, vol. 2, no. 1, pp. 193–218, 1985.

[56] H. He and E. A. Garcia, "Learning from imbalanced data," *IEEE Transactions on Knowledge & Data Engineering*, vol. 21, no. 9, pp. 1263–1284, 2008.

[57] M. W. Kadous, "Temporal classification: Extending the classification paradigm to multivariate time series," Ph.D. dissertation, University of New South Wales, 2002.

[58] A. Cruz-Roa, A. Basavanhally, F. González, H. Gilmore, M. Feldman, S. Ganesan, N. Shih, J. Tomaszewski, and A. Madabhushi, "Automatic detection of invasive ductal carcinoma in whole slide images with convolutional neural networks," in *Medical Imaging 2014: Digital Pathology*, International Society for Optics and Photonics, vol. 9041, 2014, p. 904 103.

[59] A. Janowczyk and A. Madabhushi, "Deep learning for digital pathology image analysis: A comprehensive tutorial with selected use cases," *Journal of Pathology Informatics*, vol. 7, 2016.

[60] E. Benhamou, D. Saltiel, B. Guez, and N. Paris, *Testing sharpe ratio: Luck or skill?* 2019. arXiv: 1905.08042 [q-fin.RM].

[61] E. F. Fama and K. R. French, " The Cross-Section of Expected Stock Returns," *Journal of Finance*, vol. 47, no. 2, pp. 427–65, Jun. 1992.