UNIVERSITY OF
OXFORD

Kinetic models of many-particle systems with short-range inelastic interactions and clustering

Calum Braham

Balliol College

University of Oxford

A thesis submitted for the degree of

Doctor of Philosophy

Hilary 2024

Abstract

We consider second-order many-particle systems governed by short-range inelastic interactions that may result in clustering. A short-range interaction is defined such that interactions with three or more particles are asymptotically negligible compared with binary interactions (in the absence of correlations), and clustering means that particles become highly correlated in position and velocity after an interaction. The predominance of binary interactions is maintained by treating these highly correlated clusters as individual point objects with mass as an extra internal dimension. With this approximation, we define an individual-based hierarchical model for the evolution of the system in terms of clusters. From this, we systematically derive a Boltzmann-style kinetic PDE for the probability distribution over position, velocity, and mass from the BBGKY hierarchy, using the method of matched asymptotic expansions. This model serves as a new framework for modelling second-order particle systems with short-range interactions that may result in clustering. It combines a Boltzmann-style approach for non-clustering interactions with a Smoluchowski coagulation-style clustering model and allows us to describe and analyse new processes that are not modelled well by the mean-field approach.

Two numerical schemes are derived for the kinetic PDE model: a discrete-velocity mass-binned model in one dimension and a fast Fourier-Galerkin spectral method in higher dimensions. Both are implemented and optimised in *Julia*. The PDE model is tested against ODE simulations of the full particle system in one and two dimensions, using a short-range version of the Cucker-Smale velocity-averaging model of collective behaviour as a test case. In two dimensions, results from the kinetic model are found to converge to the particle simulations as the number of particles is increased, using the velocity and mass distributions, particle energy and mean number of clusters as test metrics. In one dimension, the two models show more significant discrepancies, but the kinetic PDE model still performs better than a model using the mean-field approximation, which cannot directly model clustering interactions. To distinguish the effect of each of the approximations made in the derivation, a pointwise clustering model and a stochastic, spatially uniform model are derived and implemented as intermediate steps between the particle and kinetic models. The most significant contributors to the error in one dimension are found to be the build-up of cluster correlations and the averaging over multiple realisations performed by the PDE model.

Evolution equations for the cluster number distribution and mass distribution are derived from our short-range kinetic model, with the latter in the style of the Smoluchowski coagulation equation but with coagulation kernels dependent on the velocity distribution. For a short-range Cucker-Smale interaction with a power-law kernel, the mass distribution model is simplified to have time-dependent coagulation kernels tracking the energy dissipated by the system. Further, for this specific interaction, interdependent evolution equations for the mean number of clusters, $\mu_K(t)$, and particle energy, $E(t)$, are derived, as well as an upper bound for a combined quantity $\mu_K^2(t)E(t)$. Numerical simulations of the kinetic and individual-based models show the bound to be accurate and demonstrate the overpopulated velocity distribution tails expected from the literature on inelastic kinetic models.

Acknowledgements

I would first like to thank my supervisors, Prof. Jon Chapman and Prof. Maria Bruna. Thank you for your support, guidance, insight, advice, and many interesting discussions over the past few years. I have very much enjoyed working with you both and have learnt a lot about mathematics and research more broadly. Thank you for your kindness and flexibility during the earlier years of the COVID pandemic, which were uncertain for us all. I would also like to thank my transfer and confirmation examiners for their helpful comments, advice, and guidance for the direction of the project. Further, I would like to acknowledge the support, both financial and otherwise, of the Rhodes Trust.

To my parents, Kate and Robert, thank you for your support and care from the other side of the world. Much has changed over the past few years, but this has been a constant. To the friends and family too numerous to mention individually, from the Balliol MCR, OCIAM, the Oxford volleyball club, Australian football club, Rhodes community, and back in Australia. Thank you for your community, camaraderie and discussions both entertaining and introspective. You have expanded my view in ways that I never could have imagined. I would particularly like to mention my fellow international residents of C-Block and A-Block during the pandemic. You were my support network during some isolating, uncertain times, and I couldn't have done it without you. Finally, to Grace, for your support, care, and all the laughs. Thank you for providing a sense of calm and perspective through the process of bringing this project together.

Contents

Notation

Parameters

N	Number of particles.
d	Dimension of system.
ε	Short-range interaction length scale.
s	Exponent of power law kernel or force.
$\bar{v}$	Representative velocity for scaling.

Individual-based models and interactions

$\boldsymbol{X}_i(t), \boldsymbol{V}_i(t)$, $(X_i(t), V_i(t))$	Position and velocity of the i^{th} particle or cluster in d dimensions (one dimension).
$M_i(t)$	Mass of the i^{th} cluster.
$\boldsymbol{H}(\boldsymbol{x};\boldsymbol{v})$, $H(x;v)$	Short-range interaction in d dimensioins (one dimension).
$\boldsymbol{H}_{ij} \equiv \boldsymbol{H}\left(\frac{\boldsymbol{x}_i-\boldsymbol{x}_j}{\varepsilon}, \boldsymbol{v}_i-\boldsymbol{v}_j\right)$	Interaction between particles i and j.
$\Phi = \int_{-\infty}^{\infty} \phi(\lvert x\rvert)\,\mathrm{d}x$	Integral of kernel integral over the inner region in one dimension.
$\Phi(r) = \int_{-\infty}^{\infty} \phi\left(\sqrt{r^2+x^2}\right)\,\mathrm{d}\boldsymbol{x}$	Integral of kernel integral over the inner region in d dimensions with impact parameter r.

Kinetic models: densities and operators

$\boldsymbol{x}_i, \boldsymbol{v}_i$, (x_i, v_i)	Arguments to probability densities representing the position and velocity of the i^{th} particle or cluster in d dimensions (one dimension).
m_i	Argument to probability densities representing the mass of the i^{th} cluster.
$\boldsymbol{z}_i \equiv (\boldsymbol{x}_i;\boldsymbol{v}_i;m)$, (z)	Phase of cluster i in d dimensions (one dimension). For non-clustering model does not depend on mass.
$\vec{\boldsymbol{z}}_s \equiv (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_s)$	Vector of particle phases. Also used as $\vec{\boldsymbol{x}}_s$, $\vec{\boldsymbol{v}}_s$, for vector of positions or velocities.

$p^N(t;\vec{\boldsymbol{x}}_N;\vec{\boldsymbol{v}}_N)$	Full system joint probability density for non-clustering model.
$p_s^N(t;\vec{\boldsymbol{x}}_s;\vec{\boldsymbol{v}}_s)$	Reduced s-particle probability density for non-clustering model.
$p^k(t;\vec{\boldsymbol{x}}_k;\vec{\boldsymbol{v}}_k;\vec{m}_k)$	Joint probability density for a clustering model with k clusters.
$p_s^k(t;\vec{\boldsymbol{x}}_s;\vec{\boldsymbol{v}}_s;\vec{m}_s)$	Reduced s-cluster density for a clustering model with k clusters.
$p(t;\boldsymbol{x};\boldsymbol{v};m)$	One-particle clustering density with k marginalised.
$n_1^k = kp_1^k$	One-particle number density.
$n(t;\boldsymbol{x};\boldsymbol{v};m)$	One-particle number density with k marginalised.
$f(t;\boldsymbol{x};\boldsymbol{v};m) = mn(t;\boldsymbol{x};\boldsymbol{v};m)$	One-particle mass density.
$\tilde{\boldsymbol{x}}_1$, $\tilde{\boldsymbol{x}}_{21}$	Inner-region variables.
$\tilde{p}_2$	Inner-region density for non-clustering model.
$\tilde{p}^k_{(1,2)}$,$\tilde{p}^k_{2(1,2)}$	Inner-region density for clustering model.
$\tilde{\boldsymbol{x}}^{\pm}_{21}$	Ingoing ($-$) and outgoing ($+$) limits for a trajectory in the inner region. Similar used for $\boldsymbol{v}_1^{\pm}$, $\boldsymbol{v}_2^{\pm}$, $\boldsymbol{u}^{\pm}$, $\tilde{p}_2^{\pm}$ and ${}^{\pm}\tilde{p}^k_{(1,2)}$.
$q^k_{(i)}$	Clustering flux produced at $\boldsymbol{z}_i$ in the k-cluster density.
$\mathcal{C}_{(j,s)}(U)$, $\mathcal{C}_{(j,s)}(\boldsymbol{z}_i)$	Clustering flux produced at $\boldsymbol{z}_i$ (or with $\boldsymbol{v}_i \in U$) from an interaction between clusters j and s.
$Q(f,f)$, $Q_{\mathcal{I}}(f,f)$	Elastic (inelastic) Boltzmann collisioin opereators.
$Q^{\pm}$	Splitting to gain ($+$) and loss ($-$) operators.
$Q_C(n,n)$, $Q_{NC}(n,n)$	Clustering and non-clustering operators for clustering PDE model. Non-clustering and loss operator: $Q_I = Q_{NC} + Q_C^-$.

Kinetic models: interaction maps

$(\boldsymbol{v}_1^*, \boldsymbol{v}_2^*) = \psi_{m_1,m_2,r}(\boldsymbol{v}_1,\boldsymbol{v}_2)$	Forward non-clustering velocity map. Also $\psi_{m_1,m_2} \equiv \psi_{12}$ in one dimension and $\psi_{\boldsymbol{H},m_1,m_2}$, when dependence on $\boldsymbol{H}$ explicit.
$({}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2) = \psi^{-1}_{m_1,m_2,r}(\boldsymbol{v}_1,\boldsymbol{v}_2)$	Backward non-clustering velocity map.
$\boldsymbol{v}_1' = \lambda_{m_1,m_2}(\boldsymbol{v}_1,\boldsymbol{v}_2)$	Forward clustering velocity map. Also modifications as for ψ.
$\boldsymbol{v}_1' = \hat{\lambda}_{m_1,m_2,\boldsymbol{v}_2}(\boldsymbol{v}_1)$	Partially-evaluated forward clustering map.
${}'\boldsymbol{v}_1 = \hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2}(\boldsymbol{v}_1)$	Backward clustering map.
$p_{(2)} \equiv p(t;\boldsymbol{x}_1;\boldsymbol{v}_2;m_2)$	Shorthand for evaluation for cluster 2 in an interaction. Similar for $f_{(2)}$, $n_{(2)}$, $\varphi_{(2)}$, $\boldsymbol{z}_2$ and $p^k_{1(2)}$.

$p^* \equiv p(t; \boldsymbol{x}_1; \boldsymbol{v}_1^*; m_1)$	Shorthand for evaluating at $\boldsymbol{v}_1^*$. Similar for f^*, n^*, φ^* and $\boldsymbol{z}^*$. Evaluating at ${}^*\boldsymbol{v}_1$ shown with ${}^*p$.
$p' \equiv p(t; \boldsymbol{x}_1; \boldsymbol{v}_1'; m_1 + m_2)$	Shorthand for evaluating at post-cluster phase. Similarly applies to other densities.
${}'p \equiv p(t; \boldsymbol{x}_1; {}'\boldsymbol{v}_1; m_1 - m_2)$	Shorthand for evaluating at pre-cluster phase. Similarly applies to other densities.

Kinetic models: other

$K(t)$	Random variable for number of clusters in the system.
P_k	Probability of system having k clusters.
$\mu_K(t)$	Mean number of clusters.
$\Lambda_K(t)$	Coefficient from *Clusterzahlansatz* in number density form.
$E(t)$	Mean kinetic energy per particle.
$G(t) = \mu_K^2(t)E(t)$	Combined quantity with mean number of clusters and kinetic energy.
r, r^c	Impact parameter, impact paramer cutoff.
$r^*_{m_1,m_2}(\lvert\boldsymbol{u}\rvert)$	Impact parameter boundary between clustering and non-clustering interactions. Shorthand $r^*_{(1,2)}$ when no arguments are modified.
$\boldsymbol{u}$, u, $\hat{\boldsymbol{u}}$	Relative velocity $\boldsymbol{v}_2 - \boldsymbol{v}_1$ in d dimensions, one dimension, and unit vector for relative velocity.
$\boldsymbol{\omega}$	Unit vector in impact direction.
$\boldsymbol{\sigma}$	Unit vector representing post-interaction relative velocity if the interaction were elastic.
$B_{\boldsymbol{\sigma}}$, $B_{\boldsymbol{\omega}}$	Boltzmann collision kernel in $\boldsymbol{\sigma}$ representation and $\boldsymbol{\omega}$ representation.
B	Smoluchowski coagulation kernel.
θ	Interaction deviation angle.
e	Coefficient of restitution for inelastic interaction. Not to be confused with exponential constant.
$\langle\cdot,\cdot\rangle$,$\langle\cdot,\cdot\rangle_{x,v}$	Inner product over position, velocity and mass. Inner product only over position and velocity.

Integral regions

Ω	Spatial domain.
$\Lambda_{m_1-m_2,m_2}(v_1)$	Region of v_2 that could be mapped to v_1 under a clustering interaction.
C_{m_1,m_2}	Region representing pairs of incoming velocities that would result in a clustering interaction.
$\boldsymbol{D}$	Full integration region for z.
$\boldsymbol{C}'(\boldsymbol{z}_1)$	Integration region for z_2 that would produce a cluster at z_1.
$\boldsymbol{C}^2$	Full integration region for incoming phases that would cluster.
$\boldsymbol{D}_\Gamma(k)$, $\boldsymbol{D}_\Gamma(k, m_1)$	Mass-restricted integration region with k clusters in the system (and mass m_1 already fixed).
$\boldsymbol{C}^2_\Gamma(k)$, $\boldsymbol{C}^2_\Gamma(k, m_1)$	Mass-restricted clustering integration region with k clusters in the system (and mass m_1 already fixed).

Numerical implementations (one dimension)

V	Number of velocity grid points.
$\bar{v}_i$, v_i^-, v_i^+	Velocity grid point i, lower and upper bound of velocity bin i.
Δv	Spacing of velocity grid.
M	Number of mass grid points.
M_l	Number of linear mass grid points.
$\bar{m}_a$, m_a^-, m_a^+	Mass grid point a, lower and upper bound of mass bin a.
$\bar{w}_a$	Number of integer masses represented by bin a.
$\gamma_{\bar{m}_a}(m)$,$\gamma_{\bar{v}_i}(v)$	Allocation proportions in to mass grid point $\bar{m}_a$ and to velocity grid point $\bar{v}_i$.
$P_{i,a}(t)$	Binned probabilities for discrete-velocity mass-binned numerical scheme for kinetic PDE.
$\mathcal{P}_i(t)$	Probability density for finite volume mean-field implementation.

Numerical implementations (higher dimensions)

D_L,L	Integration domain for Fourier spectral model. Length of integration domain.

$\check{n}$	Periodic extension to number density over Fourier domain.
V	Maximum Fourier frequency.
N_u	Number of qudrature points in radial u direction.
$N_{\hat{\boldsymbol{u}}}$	Number of quadrature points in angular $\hat{\boldsymbol{u}}$ direction.
w_u, $w_{\hat{\boldsymbol{u}}}$	Weights of quadratures for u and $\hat{\boldsymbol{u}}$.
$\hat{n}_{\boldsymbol{\kappa},m}$	Fourier coefficient for mode $\boldsymbol{\kappa}$ and mass m.
$\hat{\nu}_{\boldsymbol{\kappa},a}$	Fourier coefficient for mode $\boldsymbol{\kappa}$ and mass bin a.
$\boldsymbol{h}^* = \boldsymbol{v}^* - \boldsymbol{v}$	Change in velocity over non-clusterinig map.
$\boldsymbol{h}' = \boldsymbol{v}' - \boldsymbol{v}$	Change in velocity over clustereinig map.
G, G', G^*	Direct Fourier spectral method kernel modes.
A', A^*	Fast Fourier spectral method integral kernels.
$\varepsilon_{\boldsymbol{x}}$, $\varepsilon_{\boldsymbol{v}}$	Tolerances in position and velocity for particle ODE clusters computed using DBSCAN.

Other

K_1	Modified Bessel function of the second kind.
$E_{5/4}$	Generalised exponential integral function.
${}_1F_2$	Generalisd hypergeometric function.
W_0, W_{-1}	Lambert-W function.
$c_N(\mu_K)$, $\tilde{c}_N(\mu_K)$	Empirical approximations to Λ_K.
Dg	Jacobian of function g.
$\delta_{ij} = \delta(\boldsymbol{x}_i - \boldsymbol{x}_j)$	Dirac delta in position.
$\mathbb{S}^d_R$, $\mathbb{S}^d$	d-sphere of radius R (unit d-sphere).
$\boldsymbol{\xi}^d$	Surface element of d-sphere.
$\mathcal{B}^d_R$	d-ball of radius R.

1. Introduction

Physical systems are often modelled as sets of interacting particles. Examples of such systems range from the purely physical: such as gas molecules in a fluid [16, 61], ions in plasma [3] and stars within a galaxy [9]; through biological: such as bacterial chemotaxis [49, 94] and the collective motion of swarms of animals [34, 40, 107]; to human engineered and social systems: such as traffic flow [22], pedestrian motion [85, 111], opinion formation [64, 76] and control of autonomous systems such as spacecraft or remote sensors [88, 33, 92].

The most direct models of such systems are often *individual-based*, and track the influence of generalised 'forces' on each particle in the system through a model that may take the form of a set of coupled ordinary differential equations (ODEs), stochastic differential equations (SDEs) or discrete-time maps. Such models may be denoted as *discrete*, *microscopic*, or *agent-based*, depending on the context. They commonly fall into one of two classes: first-order systems, in which the net force on a particle is proportional to its velocity and where the particles' positions are tracked, and second-order (kinetic) systems, in which the net force is proportional to the acceleration of a particle in the form of Newton's second law and where particles' positions and velocities are tracked. Some models include additional dimensions tracking internal properties for each particle (as opposed to the external position and velocity), for example, the mass, surface area or spin.

In many contexts, individual-based models are the simplest models to derive. However, while they are generally conceptually transparent, they can often be intractable analytically and computationally, particularly when the system being modelled contains a very large number of particles (for example, in the range of 10^{20} to 10^{25} particles for typical physical systems [69]). In such contexts, a common approach is to derive a *population-averaged* model (also known as a *continuum* or *macroscopic* model) that reduces the system dimension to tracking averaged quantities through governing equations, often a nonlinear partial differential equation (PDE). Such models are particularly useful when the characteristic qualitative behaviour being analysed occurs at the macroscopic or continuum level, as they can abstract from the details of individual particle behaviour to model the system as a whole.

The passage between individual-based and population-averaged models is an important topic in the literature. Population-averaged models can be derived phenomenologically or empirically to capture desired behaviour at the continuum level. However, in certain limits (for example, the mean-field limit: taking the number of particles to infinity in a system with weak but long range interactions), they can also be rigorously derived from a given individual-based model. Individual-based models are often written down from more fundamental assumptions about particle behaviour, but these can also be phenomenologically sourced. The derivation of

macroscopic physics from discrete particle behaviour is a complex and wide-ranging problem. It is related to Hilbert's sixth problem, which, in this context, concerns the rigorous derivation of macroscopic physics from axiomatic (microscopic) principles. The transition process can also proceed in the other direction, searching for individual-based models that would result in a given macroscopic behaviour.

Two commonly studied classes of population-averaged models are *kinetic* and *hydrodynamic* models. In the context of the kinetic theory of gases, in which they were first defined, they are often referred to as the *mesoscopic* and *macroscopic* regimes, respectively. In a kinetic model, the probability density (often over position and velocity) of a representative particle is tracked through a PDE. Under the assumptions used to derive the kinetic model, the phases for all particles in the system are assumed to be identically modelled by this distribution. Typical examples of kinetic models are the Boltzmann equation [16, 78] and the Vlasov equation [104], to be discussed further below. Note that even if the underlying individual-based model is theoretically deterministic, it is often the case that precise states of the full system cannot be measured or are not of direct interest. It is also common that the system is so sensitive to perturbations that any measurement would be too imprecise to define a useful deterministic particle model in any case. This is why the derived kinetic model tracks a probability density: as an average over initial distributions selected to match observable characteristics.

Sometimes individual-based models also include stochastic processes. These are often the representation of epistemological uncertainty, where other processes that are not directly captured in the deterministic model are instead represented through stochastic processes. Kinetic models can still be derived from such individual-based descriptions, but now the one-particle distributions represent averages across the initial distributions and of the stochastic process.

Hydrodynamic models represent a further abstraction step to tracking macroscopic quantities such as density, bulk velocity, temperature and pressure as functions of space and time only. They can often be derived from kinetic models by making assumptions that the system is either at equilibrium or close to equilibrium in velocity space (see [31] Ch. 5). Typical examples are the Euler or Navier-Stokes systems of PDEs.

1.1. Classical models

A typical form of a second-order individual-based model (see [15] and [69]) tracks the evolution of particle positions $\boldsymbol{X}_i \in \Omega \subseteq \mathbb{R}^d$, and velocities $\boldsymbol{V}_i \in \mathbb{R}^d$ through the set of coupled SDEs:

$$\begin{aligned} \mathrm{d}\boldsymbol{X}_i &= \boldsymbol{V}_i \mathrm{d}t, \\ \mathrm{d}\boldsymbol{V}_i &= \sqrt{2D}\,\mathrm{d}\boldsymbol{B}_i + \boldsymbol{F}\left(\boldsymbol{X}_i, \boldsymbol{V}_i\right)\mathrm{d}t + \eta(N,l)\sum_{\substack{j=1 \\ j\neq i}}^{N} \boldsymbol{H}\left(\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{l}, \boldsymbol{V}_i - \boldsymbol{V}_j\right)\mathrm{d}t, \end{aligned} \tag{1.1}$$

for $i \in \{1, \ldots, N\}$, where N is the total number of particles in the system, $\boldsymbol{B}_i$ is a d-dimensional Brownian motion with diffusion constant D, $\boldsymbol{F}(\boldsymbol{x}, \boldsymbol{v})$ represents an external force, and $\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v})$

is a function representing pariwise interactions betweeen particles. The interaction spatial length scale is given by l, and $\eta(N, l)$ is a scaling coefficient for the interaction strength dependent on the number of particles and length scale. Note that the form of the functional dependence in $\boldsymbol{H}$ assumes the interactions are symmetric in particle positions.

The seminal kinetic model is the Boltzmann equation, derived in various forms by Maxwell [78] and Boltzmann [16] in the latter half of the nineteenth century to model the dynamics of rarefied gases. The Boltzmann equation tracks the evolution of the one-particle mass density over position and velocity, $f(t; \boldsymbol{x}; \boldsymbol{v})$, and is given by

$$\partial_t f + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} f = Q(f, f). \tag{1.2}$$

The second term on the LHS represents the free transport of particles, while the RHS term $Q(f, f)$ is a quadratic integral operator that represents the effect of binary collisions on f. The specific form of Q will be discussed in more detail in Chapter 2. The Boltzmann equation can be derived from an individual-based model in the form of (1.1), with no Brownian motion, no external force and a short-ranged elastic interaction:

$$\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v}) = -\nabla_{\boldsymbol{x}} \phi(|\boldsymbol{x}|), \text{ with } \boldsymbol{x} = \frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon}.$$

The kinetic model (1.2) can be derived rigorously for short times in the Boltzmann-grad limit [52], in which we take the number of particles $N \to \infty$ and the interaction length scale $\varepsilon \to 0$ in such a way as to keep $\varepsilon^{d-1} N \equiv 1$ where d is the dimension. This ensures the density is such that particles have a finite mean free path, and encounter a finite number of distinct binary interactions per unit time. Implicit in this limit is the assumption that interactions are localised in space and time and that higher-order interactions (i.e. involving more than two particles) can be neglected. The interaction $\boldsymbol{H}$ is present in the final equation through a map from post- to pre-interaction velocities in $Q(f, f)$. In a limit with more frequent collisions, the velocity distribution relaxes to equilibrium more quickly and it is thus natural to model such systems in the hydrodynamic framework.

If the interaction is instead long-ranged (i.e. $l = O(1)$) such that all particles are interacting with each other, then a mean-field limit, taking $N \to \infty$ and scaling the interaction by $\eta = 1/N$, is appropriate. Under suitable conditions on $\boldsymbol{H}$ (and no Brownian motion or external force), such a scaling results in the Vlasov-type equation

$$\partial_t f + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} f + \nabla_{\boldsymbol{v}} \cdot (f(\boldsymbol{H} * f)) = 0, \tag{1.3}$$

where $\boldsymbol{H} * f$ is a convolution in position and velocity. An equation of similar form can be derived from (1.1) if we assume the external force term $\boldsymbol{F}$ dominates the behaviour and ignore the interaction and Brownian motion terms. Such a model would be expressed by

$$\partial_t f + \boldsymbol{v} \cdot \nabla_{\boldsymbol{x}} f + \nabla_{\boldsymbol{v}} \cdot (\boldsymbol{F} f) = 0. \tag{1.4}$$

If we assume $\boldsymbol{F}$ is only a function of $\boldsymbol{x}$, we can factor it out of the derivative and get the Vlasov

equation itself. Such models of longer-range forces find particular use in modelling plasmas [108].

For systems where the characteristic behaviour is the evolution of an internal dimension of the particle, a different class of approach, called a population balance model, is common. The archetypical population balance model is the Smoluchowski coagulation equation [103], which models the evolution of the number density of particles over their mass $n(t; m)$. The evolution equation is given by

$$\partial_t n(t; m_1) = \frac{1}{2} \sum_{m_2=1}^{m_1-1} B(m_1 - m_2, m_2)\, n(t; m_1 - m_2)\, n(t; m_2) - \sum_{m_2=1}^{\infty} B(m_1, m_2)\, n(t; m_1)\, n(t; m_2), \tag{1.5}$$

where the first term on the RHS represents the gain of particles of mass m_1 and the second term represents the loss. The coagulation kernel $B(m_1, m_2)$ represents the relative propensity for coagulation events between particles with masses m_1 and m_2. Such coagulation equations, with various different kernels B, have been used to model a wide variety of physical phenomena. Examples include the coalescence of water droplets in clouds to form rain [97], the dispersion of liquid particles through a gas in aerosols [110, 102], the aggregation of calcium oxalate crystals to form kidney stones [66] and the flocculation of mineral sediment in a gravity mineral thickener [63]. The Smoluchowski coagulation equation can be considered a kinetic model with position and velocity integrated out (in some contexts, it is called an infinite-volume mean field limit [1]) and the extra internal coordinate of mass.

1.2. Kinetic models and the BBGKY hierarchy

As mentioned, there are a multitude of ways to derive kinetic models from individual-based models. Many approaches begin with the full-system joint mass density $f^N(t; \boldsymbol{z}_1, \dots \boldsymbol{z}_N)$ where $\boldsymbol{z}_i \equiv (\boldsymbol{x}_i; \boldsymbol{v}_i) \in \Omega \times \mathbb{R}^d$ represents the phase of particle i. The time evolution equation for this joint density can be derived directly from the individual-based model and contains equivalent information. For a restricted class of systems of the form (1.1), (with no diffusion and with $\boldsymbol{F}$ and $\boldsymbol{H}$ divergence-free in velocity), the time evolution equation is given by the Liouville equation

$$\frac{\mathrm{d}f^N}{\mathrm{d}t} = 0,$$

implying mass density is conserved along trajectories in the phase space. For more general individual-based models, the evolution equation is more complex and will be discussed further in Chapter 2. By marginalising particle phases $s+1$ to N from this time evolution equation, we get an evolution equation for the s-particle joint distribution:

$$f_s^N(t, \boldsymbol{z}_1) := \int_{(\Omega \times \mathbb{R}^d)^{(N-s)}} f^N(t, \boldsymbol{z}_1, \dots, \boldsymbol{z}_N)\, \mathrm{d}\boldsymbol{z}_{s+1} \dots \mathrm{d}\boldsymbol{z}_N.$$

It can be shown that the time evolution equation for the s-particle joint distribution depends on the $(s+1)$-particle distribution up to the full distribution f^N. This structure was derived

separately by several authors [112, 75, 17, 14] and is known eponymously as the BBGKY hierarchy.

There are several methods to derive a closed kinetic equation for the one-particle density f_1 from the BBGKY hierarchy. In the mean-field scaling, it can be shown that in the limit $N \to \infty$, the joint densities tensorise as $f_s^N \to (f_1^N)^{\otimes s}$. This result is known as the *propagation of chaos*, where particles that are initially uncorrelated remain uncorrelated as time evolves. As the joint densities have factored into products of the one-particle density, we are able to derive a closed kinetic equation. It is common in these contexts to express the limiting one-particle density as $f \equiv f_1^\infty$. See [1, 104] for general reviews of the mean-field approach and [26, 68, 69] for more specific discussions of its use in simplifying individual-based kinetic models.

In general, rigorous results are not available. It is typical instead to make phenomenologically-informed assumptions to truncate the BBGKY hierarchy and obtain a closed system. One approach is to insert the result of the mean-field limit directly as an assumption and close the hierarchy at the one-particle level using $f_2^N(t; \boldsymbol{z}_1; \boldsymbol{z}_2) \approx f_1^N(t; \boldsymbol{z}_1) f_z^N(t; \boldsymbol{z}_2)$. Another commonly-used closure, which allows for two-particle correlations to be considered, is the Kirkwood superposition approximation [74]. In this approach, the three-particle joint density is approximated by considering the probability density of any pair of the particles being found in a given state, independent of the third particle. Formally this involves the approximation

$$f_3^N(t; \boldsymbol{z}_1, \boldsymbol{z}_2, \boldsymbol{z}_3, t) \approx \frac{f_2^N(t; \boldsymbol{z}_1, \boldsymbol{z}_2) f_2^N(t; \boldsymbol{z}_1, \boldsymbol{z}_3) f_2^N(t; \boldsymbol{z}_2, \boldsymbol{z}_3)}{f_1^N(t; \boldsymbol{z}_1) f_1^N(t; \boldsymbol{z}_2) f_1^N(t; \boldsymbol{z}_3)}, \tag{1.6}$$

which allows the hierarchy to be closed at the two-particle level. There are a wide variety of other closure assumptions explored in the literature, including the truncated assumption method due to Berlyand et al. [6, 7] and a perturbative approximation to the equilibrium pair correlation developed by Felderhof [44].

Bruna, Chapman and Robinson [20] develop a different approach for first-order models with Brownian diffusion, external forces and interactions through a short-ranged potential function. They use matched asymptotic expansions in a short interaction length scale ε to define an asymptotic expression for the two-particle joint density. They then substitute this expression into the one-particle density evolution equation to define a closed model in this density to leading order in the interaction range. This approach has the advantage of formally considering the particle correlations that would result from a short-ranged interaction, rather than making phenomenological assumptions about the the approximate independence of sets of particles. A similar approach using matched asymptotic expansions in the interaction length parameter has been employed to model systems with excluded-volume effects or hard-core repulsive interactions [19, 50, 94]. Such systems can be considered a limiting case of general short-range interaction functions; particles must be prevented from overlapping, and interactions are strictly localised as collisions.

1.3. Aims

The primary aim of this thesis is to systematically derive and analyse kinetic models of many particle systems governed by short-range interactions of the form

$$\begin{aligned} \dot{\boldsymbol{X}}_i(t) &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i(t) &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}\left(\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon}, \boldsymbol{V}_i - \boldsymbol{V}_j \right), \end{aligned} \tag{1.7}$$

for $1 \in \{1, \dots, N\}$, where, as in (1.1), $\boldsymbol{X}_i \in \Omega \subseteq \mathbb{R}^d$ represents the position of particle i, $\boldsymbol{V}_i \in \mathbb{R}^d$ its velocity and dots denote derivatives with respect to t. We suppose $\boldsymbol{H}$ is short-range and radial, that is

$$\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v}) = O(|\boldsymbol{x}|^{-d-\delta}) \text{ as } |\boldsymbol{x}| \to \infty, \tag{1.8}$$

for some $\delta > 0$. Further, we consider the regime of the Boltzmann-Grad limit for $d > 1$ with the length scale of the interaction $\varepsilon \to 0$ and the number of particles $N \to \infty$ with $N\varepsilon^{d-1} \equiv 1$. This ensures that the qualitative behaviour of the system is characterised by localised binary interactions separated by periods of free transport. In doing so, we expect our model to be qualitatively similar to the Boltzmann equation (1.2).

In contrast to many other models, we allow the interaction $\boldsymbol{H}$ to be clustering. That is, we allow interactions where it is possible for particles to be highly correlated in position and velocity after interacting. In order to preserve the framework of binary interactions, we approximate these correlated particles as a single object, called a cluster, with higher mass and construct a hierarchy of individual-based models tracking clusters of different masses. We then aim to derive a kinetic clustering PDE model systematically using matched asymptotic expansion on the BBGKY hierarchy similarly to in [20]. This model will track the one-particle probability density over position, velocity and mass: $p(t; \boldsymbol{x}; \boldsymbol{v}; m)$ (in some cases, the number density n and mass density f will also be used). As the model tracks the internal property of cluster mass, our final kinetic model will share properties with the Smoluchowski coagulation equation (1.5), as well as the Boltzmann equation. Finally, we aim to analyse our clustering PDE model to gain insights into the evolution of the velocity and mass distributions independently, and of averaged quantities such as the mean number of clusters, mean cluster mass, and system energy.

A prototypical example of a clustering interaction in the form of (1.7) is a short-range modified version of the Cucker-Smale velocity-averaging model of collective behaviour [35, 34]. This model has

$$\boldsymbol{H}\left(\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon}, \boldsymbol{V}_i - \boldsymbol{V}_j \right) = \phi\left(\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon} \right) (\boldsymbol{V}_j - \boldsymbol{V}_i),$$

where ϕ is an interaction kernel that satisfies the short-range condition (1.8). This suggests one application of our model could be in collective behaviour or consensus systems where the entities become correlated in multiple separate groups of different sizes (see e.g. [101], [84] and further

discussion of flocking and collective behaviour systems in §2.4). Other applications could be in the wide variety of clustering systems typically modelled with a Smoluchowski-style coagulation equation, but where we want to take into account more detail about the velocity or spatial distributions, for example, a system of aggregating particles dissipating energy and reducing the variance in the velocity distribution. Another possible application is in the clustering and phase separation of synthetic self-propelled particles and active colloidal fluids [105]. The primary heuristic requirement for our framework to apply is that particles undergo localised binary interactions that either result in them separating (as is standard for collisional kinetic models) or being correlated enough post-interaction so as to be treated as a single entity.

1.4. Outline

The remainder of this thesis will be structured as follows. In Chapter 2, we introduce the systematic framework for deriving kinetic models from second-order individual-based ODE systems through matched asymptotic expansions. We then review some of the relevant kinetic models in the literature, beginning with the Boltzmann equation for elastic interactions, moving to the inelastic Boltzmann-style equations often used for modelling granular gases and before outlining models for emergent collective behaviour. We then use the discussion of a short-range clustering Cucker-Smale model to motivate the approach taken to derive our kinetic clustering PDE mdoel.

In Chapter 3, we systematically derive our kinetic clustering PDE model in one dimension. We begin by assuming interactions occur strictly at a point before extending to short-range interactions described by ODEs on the particle model. In this chapter, we introduce the three core assumptions used to derive the PDE model: the *point-cluster* approximation, treating sets of highly correlated particles as a single point object with higher mass; the *mass-restricted Stosszahlansatz*, assuming clusters approaching an interaction are uncorrelated and have physically-realisable mass combinations; and the *Clusterzahlansatz*, assuming velocity and mass distributions are independent of the number of clusters in the system.

In Chapter 4, we derive a discrete-volume, mass-binned numerical implementation of our kinetic PDE in one dimension and implement and optimise this in *Julia*. We implement the full particle ODE model in Julia as a base case for comparison and evaluate the accuracy of the kinetic PDE model against a mean-field model for a short-range Cucker-Smale interaction. We also derive and implement a point-wise exact clustering model and spatially uniform stochastic model as interim steps to evaluate the effect of the different approximations used in deriving the kinetic PDE.

In Chapter 5, we extend the kinetic clustering PDE model derived in Chapter 3 to higher dimensions. We also introduce new formulations in terms of the number and mass density that simplify the expression of the model and obtain evolution equations for the cluster number distribution. We then derive the weak form of the kinetic PDE model and show how it can be used to derive evolution equations for moments of the distribution, such as the mean number of clusters or average particle energy. Finally, we implement our model for the short-range

Cucker-Smale interaction and use the weak form to derive bounds for the evolution of the mean number of clusters $\mu_K(t)$, particle energy $E(t)$, and combined quantity $G(t) \coloneqq \mu_K^2(t)E(t)$. We also show how the model can be reduced to a Smoluchowski-style clustering equation with time-dependent coagulation kernels.

In Chapter 6, we derive a modified version of the fast Fourier-Galerkin spectral method (commonly used to numerically model Boltzmann-style kinetic equations) and use it to implement our kinetic PDE in *Julia*. Finally, in Chapter 7, we implement and optimise this numerical implementation before evaluating our model against the full particle ODE simulations for the short-range Cucker-Smale model. We then evaluate the evolution bounds for the Cucker-Smale model developed in Chapter 5 and show the presence of overpopulated velocity distribution tails at large times.

1.5. Statement of originality

Much of the material in Chapter 2 is review. The method of deriving kinetic models from second-order particle systems in §2.1 is new, but similar in flavour and outcome to other methods in different formulations. The diagrams in Figure 2.4 are developed slightly from those in the literature, and the discussion of clustering models in §2.5 is new. The material from Chapter 3 through Chapter 8 is original.

2. Short-range kinetic models

2.1. From individual-based to kinetic models

In this section, we will explore the derivation of kinetic (mesoscopic) population-averaged particle models from individual-based (microscopic) models for systems with short-range binary interactions. We will mainly follow the treatments of Cercignani [31], Harris [61] and Gallagher et al. [52], but adapt it to our notation and framework as a procedure of matched asymptotic expansions.

2.1.1. Individual-based (microscopic) models

In this thesis, we will primarily be interested in individual-based models of N indistinguishable particles with positions $\boldsymbol{X}_i(t)$ and velocities $\boldsymbol{V}_i(t)$ that evolve according to the generalised pairwise interaction law

$$\begin{aligned} \dot{\boldsymbol{X}}_i(t) &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i(t) &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}\left(\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon}, \boldsymbol{V}_i - \boldsymbol{V}_j \right), \end{aligned} \tag{2.1}$$

for $1 \leq i \leq N$, where $\boldsymbol{X}_i \in \Omega \subseteq \mathbb{R}^d$ and $\boldsymbol{V}_i \in \mathbb{R}^d$, and dots denote derivatives with respect to t. We suppose $\boldsymbol{H}$ is short range, that is

$$\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v}) = O(|\boldsymbol{x}|^{-d-\delta}) \text{ as } |\boldsymbol{x}| \to \infty, \tag{2.2}$$

for some $\delta > 0$. The factor ε is introduced so that the range of the interaction scales as $O(\varepsilon)$ and the pre-factor of $1/\varepsilon$ is included to ensure that pairwise interactions lasting only $O(\varepsilon)$ time create an $O(1)$ change in velocity (it arises naturally if $\boldsymbol{H}$ is the gradient of a potential). We consider the regime $\varepsilon \ll 1, \varepsilon^d N \ll 1$ such that we have a low volume fraction of particles. This includes, for $d > 1$, the Boltzmann–Grad limit: $N \to \infty$, $\varepsilon \to 0$ with $\varepsilon^{d-1} N \equiv 1$. This limit is intended so that particles undergo a finite number of distinct binary interactions per unit time, and the low volume fraction ensures the probability that three particles are within interaction range of each other simultaneously goes to 0. As such, the primary qualitative behaviour of the system is expected to be uniform rectilinear motion when the particles are outside of interaction range, interspersed with pairwise interactions taking $O(\varepsilon)$ time and causing an $O(1)$ change in velocity.

Note that the case of hard spheres of radius ε colliding elastically can be heuristically

represented in this framework as the limit of the system with $\boldsymbol{H}(\boldsymbol{x},\boldsymbol{v}) = -\nabla_{\boldsymbol{x}}\phi(\boldsymbol{x})$, with $\phi(\boldsymbol{x}) = \infty$ if $|\boldsymbol{x}| < \varepsilon$ and $\phi(\boldsymbol{x}) = 0$ if $|\boldsymbol{x}| > \varepsilon$. Also note that much of the analysis that follows could be extended to work with the more general interaction $\boldsymbol{H}((\boldsymbol{X}_i - \boldsymbol{X}_j)/\varepsilon, \boldsymbol{X}_i, \boldsymbol{V}_i, \boldsymbol{V}_j)$, but the form above is widely applicable and simplifies the derivation.

2.1.2. Phase-space probabilities and the BBGKY hierarchy

As noted in Chapter 1, for physically relevant particle systems it is impossible to determine true particle trajectories or often even to precisely measure the system's state at any one time. Therefore, we are interested in the expected behaviour of macroscopically measurable quantities, which we determine by averaging the full system state over particles and initial configurations. To this end, we define the N-particle probability density function $p^N(t; \boldsymbol{z}_1, \ldots, \boldsymbol{z}_N)$, over particle phases $\boldsymbol{z}_i := (\boldsymbol{x}_i; \boldsymbol{v}_i) \in \Omega \times \mathbb{R}^d$, which represents the probability density for the system being in the state $\vec{\boldsymbol{z}}_N := (\boldsymbol{z}_1, \ldots, \boldsymbol{z}_N) \in (\Omega \times \mathbb{R}^d)^N$ at time t, given an assumed distribution of initial conditions. Due to particles being indistinguishable, p^N is invariant to permutations of particle indices and hence to permutations of its arguments. Note that we use the convention that lower case $\boldsymbol{x}_i$ and $\boldsymbol{v}_i$ represent independent arguments of this density function, while upper case $\boldsymbol{X}_i(t)$ and $\boldsymbol{V}_i(t)$ represent specific particle trajectories as functions of time. We use $\boldsymbol{z}_i$ or omit the arguments entirely when they are the standard $(\boldsymbol{x}_i; \boldsymbol{v}_i)$, but will write specific $\boldsymbol{x}$ and $\boldsymbol{v}$ arguments otherwise. For this analysis, we take Ω to be a square region with periodic boundary conditions, equivalent to the d-dimensional torus $\mathbb{T}^d$. However, we note the full space $\mathbb{R}^d$ can be considered with minimal modification. We assume that p^N goes to 0 sufficiently quickly as $|\boldsymbol{v}_i| \to \infty$, such that integrals of the form $\int_{\mathbb{R}^d} p^N \, \mathrm{d}\boldsymbol{v}_i$ are well defined.

Considering the conservation of probability in an infinitesimal box around $\vec{\boldsymbol{z}} \in (\Omega \times \mathbb{R}^d)^N$ we obtain the time evolution of p^N:

$$\partial_t p^N + \sum_{i=1}^{N} \nabla_{\boldsymbol{x}_i} \cdot (\dot{\boldsymbol{x}}_i p^N) + \sum_{i=1}^{N} \nabla_{\boldsymbol{v}_i} \cdot (\dot{\boldsymbol{v}}_i p^N) = 0. \tag{2.3}$$

For the particle system (2.1), this becomes

$$\partial_t p^N + \sum_{i=1}^{N} \boldsymbol{v}_i \cdot \nabla_{\boldsymbol{x}_i} p^N + \frac{1}{\varepsilon} \sum_{\substack{i,j=1 \\ j \neq i}}^{N} \nabla_{\boldsymbol{v}_i} \cdot \left(\boldsymbol{H}_{ij} p^N \right) = 0, \tag{2.4}$$

where

$$\boldsymbol{H}_{ij} := \boldsymbol{H}\left(\frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{\varepsilon}, \boldsymbol{v}_i - \boldsymbol{v}_j \right).$$

We note that if system (2.1) satisfies the Liouville property (or divergence-free condition: $\sum_i \nabla_{\boldsymbol{v}_i} \cdot \sum_{j \neq i} \boldsymbol{H}_{ij} = 0$), as it does for the potential-force model discussed in §2.2, (2.4) reduces to the Liouville equation:

$$\frac{D}{Dt} p^N = \partial_t p^N + \sum_{i=1}^{N} \boldsymbol{v}_i \cdot \nabla_{\boldsymbol{x}_i} p^N + \frac{1}{\varepsilon} \sum_{\substack{i,j=1 \\ j \neq i}}^{N} \boldsymbol{H}_{ij} \cdot \nabla_{\boldsymbol{v}_i} p^N = 0. \tag{2.5}$$

Here, however, we don't make this assumption and consider (2.4) as the general model for the evolution of p^N.

We initially write the evolution equation for the one-particle probability density,

$$p_1^N(t,\boldsymbol{z}_1) := \int_{(\Omega\times\mathbb{R}^d)^{(N-1)}} p^N(t,\boldsymbol{z}_1,\ldots,\boldsymbol{z}_N)\,\mathrm{d}\boldsymbol{z}_2\ldots\mathrm{d}\boldsymbol{z}_N,$$

as part of the BBGKY hierarchy (a procedure developed in [112, 75, 17, 14]), where the evolution of the s-particle marginal p_s^N depends on p_{s+1}^N, for $s\in\{1,\ldots,N-1\}$. Marginalising (2.4) over particle phases $\boldsymbol{z}_{s+1}$ to $\boldsymbol{z}_N$ we get

$$\begin{aligned}\partial_t p_s^N + \sum_{i=1}^{s}\boldsymbol{v}_i\cdot\nabla_{\boldsymbol{x}_i}p_s^N &= -\frac{1}{\varepsilon}\sum_{\substack{i,j=1\\ j\neq i}}^{s}\nabla_{\boldsymbol{v}_i}\cdot(\boldsymbol{H}_{ij}p_s^N)\\ &\quad -\frac{1}{\varepsilon}\sum_{i=1}^{s}\sum_{j=s+1}^{N}\int_{(\Omega\times\mathbb{R}^d)^{(N-s)}}\nabla_{\boldsymbol{v}_i}\cdot(\boldsymbol{H}_{ij}p^N)\,\mathrm{d}\boldsymbol{z}_{s+1}\ldots\mathrm{d}\boldsymbol{z}_N \qquad (2.6)\\ &= -\frac{1}{\varepsilon}\sum_{\substack{i,j=1\\ j\neq i}}^{s}\nabla_{\boldsymbol{v}_i}\cdot(\boldsymbol{H}_{ij}p_s^N) - \frac{(N-s)}{\varepsilon}\sum_{i=1}^{s}\int_{\Omega\times\mathbb{R}^d}\nabla_{\boldsymbol{v}_i}\cdot(\boldsymbol{H}_{i(s+1)}p_{s+1}^N)\,\mathrm{d}\boldsymbol{z}_{s+1},\end{aligned}$$

where we have used the particle indistinguishability to map the indices j to $s+1$ and simplify the integral in the final term. The evolution equations for the one- and two-particle densities are

$$\partial_t p_1^N + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}p_1^N = -\frac{(N-1)}{\varepsilon}\int_{\Omega\times\mathbb{R}^d}\nabla_{\boldsymbol{v}_1}\cdot(\boldsymbol{H}_{12}p_2^N)\,\mathrm{d}\boldsymbol{z}_2, \qquad (2.7)$$

and

$$\begin{aligned}\partial_t p_2^N + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}p_2^N + \boldsymbol{v}_2\cdot\nabla_{\boldsymbol{x}_2}p_2^N &= -\frac{1}{\varepsilon}\nabla_{\boldsymbol{v}_1}\cdot\left(\boldsymbol{H}_{12}p_2^N\right) - \frac{1}{\varepsilon}\nabla_{\boldsymbol{v}_2}\cdot\left(\boldsymbol{H}_{21}p_2^N\right)\\ &\quad -\frac{(N-2)}{\varepsilon}\sum_{i=1}^{2}\int_{\Omega\times\mathbb{R}^d}\nabla_{\boldsymbol{v}_i}\cdot(\boldsymbol{H}_{i3}p_3^N)\,\mathrm{d}\boldsymbol{z}_3.\end{aligned} \qquad (2.8)$$

There are several methods for truncating the BBGKY hierarchy to produce a closed system of equations, whose effectiveness depends on the specific properties of the system under consideration. In §2.1.3, we consider the kinetic equation produced using the well-known mean-field approximation. We then propose a generalised alternative approach based on the method of matched asymptotics to derive a closed equation for the one-particle density with short-range interactions.

2.1.3. Mean-field closure

The mean-field limit corresponds to truncating the hierarchy at the one-particle density, p_1^N, ignoring correlations in the interaction term by writing

$$p_2^N(t;\boldsymbol{z}_1,\boldsymbol{z}_2) = p_1^N(t;\boldsymbol{z}_1)p_1^N(t;\boldsymbol{z}_2), \qquad (2.9)$$

in (2.7), resulting in

$$\partial_t p_1^N + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} p_1^N = -\frac{(N-1)}{\varepsilon} \nabla_{\boldsymbol{v}_1} \cdot \left(p_1^N (\boldsymbol{H} * p_1^N) \right), \tag{2.10}$$

where

$$(\boldsymbol{H} * p_1^N)(t; \boldsymbol{z}_1) := \int_{\Omega \times \mathbb{R}^d} \boldsymbol{H}_{12} p_1^N(t; \boldsymbol{z}_2) \, \mathrm{d}\boldsymbol{z}_2.$$

Higher-order closures, such as Kirkwood's closure expressing p_3^N in terms of p_2^N, are also possible [74]. These allow extra features to be tracked, such as particle correlations, at the cost of significantly increased complexity.

The kinetic mean-field PDE (2.10) can be derived rigorously in the case of weak coupling, i.e. when the strength of the pairwise interaction is small and its range remains macroscopic. The appropriate mean-field scaling (as opposed to the strong coupling regime in (2.1)) is

$$\dot{\boldsymbol{V}}_i(t) = \frac{1}{N} \sum_{j \neq i} \boldsymbol{H} \left(\boldsymbol{X}_i - \boldsymbol{X}_j, \boldsymbol{V}_i - \boldsymbol{V}_j \right).$$

In the limit $N \to \infty$ and under suitable conditions on $\boldsymbol{H}$ (typically Lipschitz or locally Lipschitz [15]), one can prove a propagation of chaos result, namely that p_s^N tensorises in the limit ($p_s^N \to (p_1^N)^{\otimes s}$ as $N \to \infty$), leading to the McKean-Vlasov-type equation

$$\partial_t p + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} p + \nabla_{\boldsymbol{v}_1} \cdot (p(\boldsymbol{H} * p)) = 0, \tag{2.11}$$

where $p(t; \boldsymbol{z}_1) \equiv p_1^\infty(t; \boldsymbol{z}_1)$. This is an active area of research, with current efforts focused on relaxing the conditions on $\boldsymbol{H}$ to allow for singular interactions [69].

2.1.4. Matched asymptotics

Taking the Boltzmann-Grad limit ($N \to \infty$, $\varepsilon \to 0$ with $\varepsilon^{d-1} N \equiv 1$), our approach is to use matched asymptotic expansions in the small interaction-length parameter ε to derive a closed time evolution equation for p_1^N. In one dimension, the Boltzmann-Grad limit is not defined, and we instead take $\varepsilon \to 0$ with $\varepsilon N \ll 1$.

We first consider the expansion of the time evolution equation for p_2^N (2.8) in terms of the small parameter ε. This equation shows the transport and interaction of two particles at $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, with an additional integral term representing the influence of all other particles as a generalised 'third particle'. We define an *inner region* where particles 1 and 2 are within interaction range, i.e. $|\boldsymbol{x}_2 - \boldsymbol{x}_1| \sim \varepsilon$ and an *outer region* where the two particles have negligible influence on each other, i.e. where $|\boldsymbol{x}_2 - \boldsymbol{x}_1| \gg \varepsilon$.

In the inner region, we introduce *inner variables*

$$\tilde{\boldsymbol{x}}_1 = \boldsymbol{x}_1, \quad \tilde{\boldsymbol{x}}_{21} = \frac{\boldsymbol{x}_2 - \boldsymbol{x}_1}{\varepsilon},$$

which result in the first-order derivatives

$$\nabla_{\boldsymbol{x}_1} = \nabla_{\tilde{\boldsymbol{x}}_1} - \frac{1}{\varepsilon} \nabla_{\tilde{\boldsymbol{x}}_{21}}, \qquad \nabla_{\boldsymbol{x}_2} = \frac{1}{\varepsilon} \nabla_{\tilde{\boldsymbol{x}}_{21}}.$$

Defining the solution in the inner region to be $\tilde{p}_2(t;\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_{21};\boldsymbol{v}_1,\boldsymbol{v}_2)$, changing to inner variables (2.8) becomes

$$\begin{aligned}
&\varepsilon\partial_t\tilde{p}_2\\
&\quad= -\varepsilon\boldsymbol{v}_1\cdot\nabla_{\tilde{\boldsymbol{x}}_1}\tilde{p}_2+(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\nabla_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}_2\\
&\qquad-\nabla_{\boldsymbol{v}_1}\cdot(\boldsymbol{H}(-\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_1-\boldsymbol{v}_2)\tilde{p}_2)-\nabla_{\boldsymbol{v}_2}\cdot(\boldsymbol{H}(\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_2-\boldsymbol{v}_1)\tilde{p}_2)\\
&\qquad-\varepsilon^d(N-2)\int_{\mathbb{R}^d\times\mathbb{R}^d}\nabla_{\boldsymbol{v}_1}\cdot\Big(\boldsymbol{H}(\tilde{\boldsymbol{x}}_{13},\boldsymbol{v}_1-\boldsymbol{v}_3)p_3^N(t;\tilde{\boldsymbol{x}}_1,\boldsymbol{x}_2,\tilde{\boldsymbol{x}}_1-\varepsilon\tilde{\boldsymbol{x}}_{13};\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)\Big)\,\mathrm{d}\tilde{\boldsymbol{x}}_{13}\mathrm{d}\boldsymbol{v}_3\\
&\qquad-\varepsilon^d(N-2)\int_{\mathbb{R}^d\times\mathbb{R}^d}\nabla_{\boldsymbol{v}_2}\cdot\Big(\boldsymbol{H}(\tilde{\boldsymbol{x}}_{23},\boldsymbol{v}_2-\boldsymbol{v}_3)p_3^N(t;\boldsymbol{x}_1,\tilde{\boldsymbol{x}}_2,\tilde{\boldsymbol{x}}_2-\varepsilon\tilde{\boldsymbol{x}}_{23};\boldsymbol{v}_1,\boldsymbol{v}_2,\boldsymbol{v}_3)\Big)\,\mathrm{d}\tilde{\boldsymbol{x}}_{23}\mathrm{d}\boldsymbol{v}_3.
\end{aligned}\tag{2.12}$$

We have localised the final two integral terms in two different inner regions, as they represent particle 3 interacting with particle 1 (i.e. $|\boldsymbol{x}_3-\boldsymbol{x}_1|\sim\varepsilon$) and particle 3 interacting with particle 2 (i.e. $|\boldsymbol{x}_3-\boldsymbol{x}_2|\sim\varepsilon$), respectively. This required defining new inner variables $\tilde{\boldsymbol{x}}_1=\boldsymbol{x}_1$ and $\tilde{\boldsymbol{x}}_{13}=(\boldsymbol{x}_1-\boldsymbol{x}_3)/\varepsilon$ for the first integral and similar for the second. Note that, changing the integration variables to these inner variables, the integrands are $O(1)$ in ε and the $\boldsymbol{x}$ domain becomes $\mathbb{R}^d$. As $\varepsilon^d\,(N-2)$ is $O(\varepsilon)$ these integral terms can be ignored at leading order, implying that interactions involving three or more particles are negligibly rare compared to binary interactions. Thus, to leading order in ε, $\tilde{p}_2$ satisfies

$$(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\nabla_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}_2=\nabla_{\boldsymbol{v}_1}\cdot(\boldsymbol{H}(-\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_1-\boldsymbol{v}_2)\tilde{p}_2)+\nabla_{\boldsymbol{v}_2}\cdot(\boldsymbol{H}(\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_2-\boldsymbol{v}_1)\tilde{p}_2).\tag{2.13}$$

Now considering the time evolution equation of p_1^N (2.7), changing the integral term to inner variables gives

$$\partial_t p_1^N+\boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}p_1^N=-\varepsilon^{d-1}(N-1)\int_{\mathbb{R}^d\times\mathbb{R}^d}\nabla_{\boldsymbol{v}_1}\cdot(\boldsymbol{H}(-\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_1-\boldsymbol{v}_2)\tilde{p}_2)\,\mathrm{d}\tilde{\boldsymbol{x}}_{21}\mathrm{d}\boldsymbol{v}_2,\tag{2.14}$$

where we note that, in the Boltzmann-Grad limit, $\varepsilon^{d-1}\,(N-1)$ is $O(1)$ and the integral term remains at leading order. We can rearrange (2.13) and substitute for the integrand in (2.14) to give

$$\begin{aligned}
\partial_t p_1^N+\boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}p_1^N=&-\varepsilon^{d-1}\,(N-1)\int_{\mathbb{R}^d\times\mathbb{R}^d}(\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\nabla_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}_2\,\mathrm{d}\tilde{\boldsymbol{x}}_{21}\mathrm{d}\boldsymbol{v}_2\\
&+\varepsilon^{d-1}\,(N-1)\int_{\mathbb{R}^d\times\mathbb{R}^d}\nabla_{\boldsymbol{v}_2}\cdot(\boldsymbol{H}(\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_2-\boldsymbol{v}_1)\tilde{p}_2)\,\mathrm{d}\tilde{\boldsymbol{x}}_{21}\mathrm{d}\boldsymbol{v}_2.
\end{aligned}\tag{2.15}$$

The second integral vanishes by the divergence theorem due to our boundary condition assumptions that $\tilde{p}_2\to 0$ as $|\boldsymbol{v}_2|\to\infty$. Using the divergence theorem on the first integral gives

$$\partial_t p_1^N+\boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}p_1^N=\varepsilon^{d-1}\,(N-1)\lim_{R\to\infty}\int_{\mathbb{R}^d}\int_{\mathbb{S}_R^{d-1}}\tilde{p}_2(\boldsymbol{v}_2-\boldsymbol{v}_1)\cdot\mathrm{d}\boldsymbol{S}_{\tilde{\boldsymbol{x}}_{21}}\mathrm{d}\boldsymbol{v}_2,\tag{2.16}$$

where $\mathbb{S}_R^{d-1}$ is the $(d-1)$-sphere of radius R.

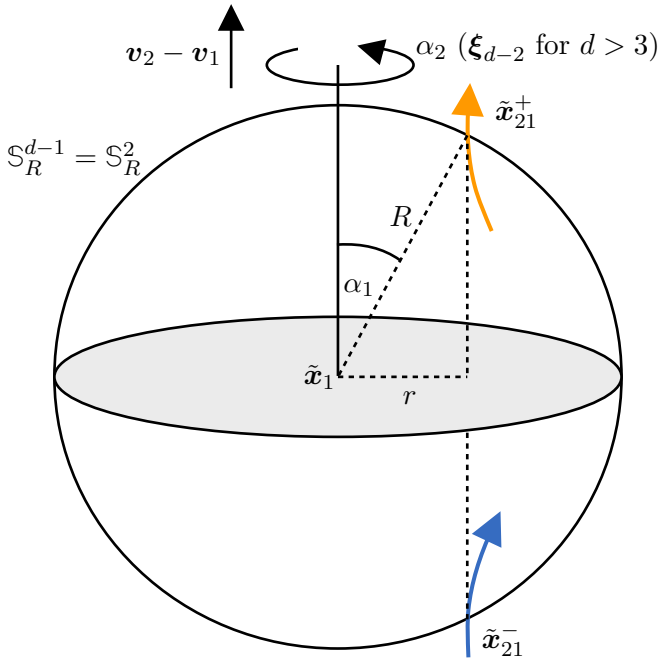


Figure 2.1.: A three-dimensional example of the spherical coordinate system used to manipulate the collision integral from a surface integral over $\mathbb{S}_R^{d-1}$ to an integral in the impact parameter r over $[0, R]$.

Following a similar argument to that in [52], we have that

$$\int_{\mathbb{S}_R^{d-1}} (\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot \mathrm{d}\boldsymbol{S}_{\tilde{\boldsymbol{x}}_{21}} = \int_{\mathbb{S}_R^{d-1}} (\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot \frac{\tilde{\boldsymbol{x}}_{21}}{|\tilde{\boldsymbol{x}}_{21}|} R^{d-1} \, \mathrm{d}\boldsymbol{\xi}^{d-1} = \int_{\mathbb{S}_R^{d-1}} (\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot (R, \boldsymbol{\xi}^{d-1}) \, R^{d-2} \, \mathrm{d}\boldsymbol{\xi}^{d-1},$$

where we have represented $\tilde{\boldsymbol{x}}_{21}$ in $(d-1)$-spherical coordinates as $(R, \boldsymbol{\xi}^{d-1})$ and where

$$\mathrm{d}\boldsymbol{\xi}^{d-1} \equiv \sin^{d-2}(\alpha_1) \sin^{d-3}(\alpha_2) \ldots \sin(\alpha_{d-2}) \, \mathrm{d}\alpha_1 \ldots \mathrm{d}\alpha_{d-1}$$

is a surface element of the unit sphere $\mathbb{S}^{d-1}$, parameterised in terms of angles α_1 to α_{d-1} (see Figure 2.1). With this parameterisation we have angle ranges $\alpha_i \in [0, \pi]$ for $i \in \{1, \ldots, d-2\}$ and $\alpha_{d-1} \in [0, 2\pi)$. Given we have $\mathrm{d}\boldsymbol{\xi}^{d-1} \equiv \sin^{d-2}(\alpha_1) \, \mathrm{d}\alpha_1 \, \mathrm{d}\boldsymbol{\xi}^{d-2}$, we can write (2.16) as

$$\begin{aligned} &\partial_t p_1^N + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} p_1^N = \\ &\quad \varepsilon^{d-1} (N-1) \lim_{R \to \infty} \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-2}} \int_0^{\pi} \tilde{p}_2 \, (\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot (R, \alpha_1, \boldsymbol{\xi}^{d-2}) \, (R \sin(\alpha_1))^{d-2} \, \mathrm{d}\alpha_1 \, \mathrm{d}\boldsymbol{\xi}^{d-2} \, \mathrm{d}\boldsymbol{v}_2. \end{aligned} \tag{2.17}$$

As we have freedom to define the spherical parameterisation $\mathrm{d}\boldsymbol{\xi}^{d-1}$ with any orientation, we rotate it for each pair $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ such that $\boldsymbol{v}_2 - \boldsymbol{v}_1$ is parallel to the direction with $\alpha_1 = 0$. We then have $R \sin(\alpha_1) = r$, where r is defined as the impact parameter: the perpendicular distance from the origin of a line passing through $\tilde{\boldsymbol{x}}_{21}$ with direction $\boldsymbol{v}_2 - \boldsymbol{v}_1$ (i.e. the closest approach of the two particles if their velocities remained as they were at infinity in the inner region and their paths were not curved by interaction). Using the relation $\mathrm{d}r / \mathrm{d}\alpha_1 = R \cos(\alpha_1)$, it also

follows that

$$(\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot (R, \alpha_1, \boldsymbol{\xi}^{d-2})\, \mathrm{d}\alpha_1 = |\boldsymbol{v}_2 - \boldsymbol{v}_1| R \cos(\alpha_1) \mathrm{d}\alpha_1 = \pm |\boldsymbol{v}_2 - \boldsymbol{v}_1|\, \mathrm{d}r.$$

Finally, we note that the integral over $\mathrm{d}r\, \mathrm{d}\boldsymbol{\xi}_{d-2}$ is a double cover of the sphere $\mathbb{S}_R^{d-1}$, as represented by the $\pm$ above. Splitting the ball $\mathbb{S}_R^{d-1}$ into two hemispheres, one representing incoming trajectories where $(\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot \tilde{\boldsymbol{x}}_{21} < 0$, and one representing outgoing trajectories where $(\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot \tilde{\boldsymbol{x}}_{21} > 0$, we note that two values of $\tilde{\boldsymbol{x}}_{21}$ project onto the same point $(r, \boldsymbol{\xi}_{d-2})$, one from each hemisphere (see Figure 2.1). Labelling the boundary point $\tilde{\boldsymbol{x}}_{21}^+$ from the outgoing hemisphere and $\tilde{\boldsymbol{x}}_{21}^-$ from the incoming hemisphere, (2.17) becomes

$$\begin{aligned}\partial_t p_1^N + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} p_1^N = \varepsilon^{d-1}(N-1) \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-2}} \int_0^\infty \Big(& \tilde{p}_2(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^+; \boldsymbol{v}_1, \boldsymbol{v}_2) \\ & - \tilde{p}_2(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^-; \boldsymbol{v}_1, \boldsymbol{v}_2) \Big) |\boldsymbol{v}_2 - \boldsymbol{v}_1| r^{d-2}\, \mathrm{d}r\, \mathrm{d}\boldsymbol{\xi}^{d-2}\, \mathrm{d}\boldsymbol{v}_2. \quad (2.18)\end{aligned}$$

As we take $R \to \infty$, we intend to match these trajectories to the solution in the outer region, where no interactions are present at leading order, and we can assume particles are independent and uncorrelated. Given these assumptions, the outer region solution is

$$\hat{p}_2^N(t; \boldsymbol{z}_1, \boldsymbol{z}_2) = p_1^N(t; \boldsymbol{z}_1) p_1^N(t; \boldsymbol{z}_2) + O(\varepsilon).$$

which, expanded in terms of inner variables, becomes

$$\begin{aligned}\hat{p}_2^N(t; \boldsymbol{z}_1, \boldsymbol{z}_2) &= p_1^N(t; \tilde{\boldsymbol{x}}_1; \boldsymbol{v}_1) p_1^N(t; \tilde{\boldsymbol{x}}_1 + \varepsilon \tilde{\boldsymbol{x}}_{21}; \boldsymbol{v}_2) + O(\varepsilon) \\ &= p_1^N(t; \tilde{\boldsymbol{x}}_1; \boldsymbol{v}_1) \left[p_1^N(t; \tilde{\boldsymbol{x}}_1; \boldsymbol{v}_2) + \varepsilon \tilde{\boldsymbol{x}}_{21} \cdot \nabla_{\tilde{\boldsymbol{x}}_1} p_1^N(t; \tilde{\boldsymbol{x}}_1; \boldsymbol{v}_2) + O(\varepsilon^2) \right] + O(\varepsilon) \quad (2.19) \\ &= p_1^N(t; \boldsymbol{x}_1; \boldsymbol{v}_1) p_1^N(t; \boldsymbol{x}_1; \boldsymbol{v}_2) + O(\varepsilon).\end{aligned}$$

However, immediately post-interaction, we expect particle trajectories to be highly correlated, which contradicts the assumptions for the outer region. Thus, we may only match incoming trajectories for particles *approaching* an interaction with this outer solution. This is Boltzmann's *Stosszahlansatz* (or molecular chaos assumption) [61, 30], which asserts that, while a pair of particles is correlated post-interaction, if they ever interact again, they will have interacted with enough other particles in the interim that any correlation is now negligible. Matching trajectories from the incoming hemisphere to the outer solution gives

$$\begin{aligned}\partial_t p_1^N + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} p_1^N = \varepsilon^{d-1}(N-1) \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-2}} \int_0^\infty \Big(& \tilde{p}_2(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^+; \boldsymbol{v}_1, \boldsymbol{v}_2) \\ & - p_1^N(t; \boldsymbol{x}_1; \boldsymbol{v}_1) p_1^N(t; \boldsymbol{x}_1; \boldsymbol{v}_2) \Big) |\boldsymbol{v}_2 - \boldsymbol{v}_1|\, r^{d-2}\, \mathrm{d}r\, \mathrm{d}\boldsymbol{\xi}^{d-2}\, \mathrm{d}\boldsymbol{v}_2. \quad (2.20)\end{aligned}$$

To produce a closed equation in p_1^N, it remains to express the probability densities of outgoing trajectories, $\tilde{p}_2(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^+; \boldsymbol{v}_1, \boldsymbol{v}_2)$, in terms of the density at the corresponding incoming trajectories so that we can apply the same independence assumption and express them in terms of

p_1^N. That is, we wish to determine the invertible functions $\zeta_{\boldsymbol{H}}$ and $\psi_{\boldsymbol{H}}$ such that

$$\begin{aligned}\lim_{|\tilde{\boldsymbol{x}}_{21}|\to\infty} \tilde{p}_2(t;\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_{21}^+;\boldsymbol{v}_1,\boldsymbol{v}_2) &= \zeta_{\boldsymbol{H}}\left(\lim_{|{}^*\tilde{\boldsymbol{x}}_{21}^+|\to\infty} \tilde{p}_2(t-O(\varepsilon);\tilde{\boldsymbol{x}}_1,{}^*\tilde{\boldsymbol{x}}_{21}^+;{}^*\boldsymbol{v}_1,{}^*\boldsymbol{v}_2)\right)\\ &= \zeta_{\boldsymbol{H}}\left(p_1^N(t;\boldsymbol{x}_1;{}^*\boldsymbol{v}_1)p_1^N(t;\boldsymbol{x}_1;{}^*\boldsymbol{v}_2)+\mathcal{O}(\varepsilon)\right)\end{aligned} \tag{2.21}$$

with

$$(\boldsymbol{v}_1,\boldsymbol{v}_2) = \psi_{\boldsymbol{H}}^{-1}({}^*\boldsymbol{v}_1,{}^*\boldsymbol{v}_2), \tag{2.22}$$

where the pre-starred quantities, ${}^*\tilde{\boldsymbol{x}}_{21}$, ${}^*\boldsymbol{v}_1$ and ${}^*\boldsymbol{v}_2$, are the incoming coordinates of the trajectory that leaves the interaction with $\tilde{\boldsymbol{x}}_{21}$, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$. Throughout this thesis, we will use the 'pre-star' notation to represent incoming velocities that would map to the unstarred velocities after an interaction and 'post-star' notation, e.g. $\boldsymbol{v}_1^*$, to represent velocities that unstarred velocities would map to. With these definitions, the starred quantities can be considered shorthand for the interaction maps

$$({}^*\boldsymbol{v}_1,{}^*\boldsymbol{v}_2) = \psi_{\boldsymbol{H}}^{-1}(\boldsymbol{v}_1,\boldsymbol{v}_2), \text{ and } (\boldsymbol{v}_1^*,\boldsymbol{v}_2^*) = \psi_{\boldsymbol{H}}(\boldsymbol{v}_1,\boldsymbol{v}_2). \tag{2.23}$$

In (2.21), in addition to matching to the outer solution, we have also assumed that variations in $\hat{p}_2^N$ over the $O(\varepsilon)$ interaction time are negligible to leading order. Finding the functions $\zeta_{\boldsymbol{H}}$ and $\psi_{\boldsymbol{H}}$ is equivalent to solving the inner equation (2.13). Assuming we have found these functions, our evolution equation for p_1^N can be written (to leading order in ε) as

$$\begin{aligned}\partial_t p_1^N + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1} p_1^N = \varepsilon^{d-1}(N-1)\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-2}}\int_0^\infty \Big[\zeta_{\boldsymbol{H}}\Big(p_1^N(t;\boldsymbol{x}_1;{}^*\boldsymbol{v}_1)p^N(t;\boldsymbol{x}_1;{}^*\boldsymbol{v}_2)\Big)\\ - p_1^N(t;\boldsymbol{x}_1;\boldsymbol{v}_1)p_1^N(t;\boldsymbol{x}_1;\boldsymbol{v}_2)\Big]|\boldsymbol{v}_2-\boldsymbol{v}_1|\,r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}^{d-2}\,\mathrm{d}\boldsymbol{v}_2.\end{aligned} \tag{2.24}$$

As alluded to by the subscripts on $\zeta_{\boldsymbol{H}}$ and $\psi_{\boldsymbol{H}}$, these functions and the procedure to determine them will depend on the specific form of the interaction $\boldsymbol{H}(\boldsymbol{x},\boldsymbol{v})$. Indeed, this is the only place where the specific interaction will be present in the evolution equation for p_1^N. Introducing the shorthand notation $p \equiv p_1^N(t;\boldsymbol{x}_1;\boldsymbol{v}_1)$, $p_{(2)} \equiv p_1^N(t;\boldsymbol{x}_1;\boldsymbol{v}_2)$ and ${}^*p \equiv p_1^N(t;\boldsymbol{x}_1;{}^*\boldsymbol{v}_1)$ we can write (2.24) as

$$\partial_t p + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1} p = \varepsilon^{d-1}(N-1)\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-2}}\int_0^\infty \left[\zeta_{\boldsymbol{H}}\left({}^*p\,{}^*p_{(2)}\right) - p\,p_{(2)}\right]|\boldsymbol{v}_2-\boldsymbol{v}_1|\,r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}^{d-2}\,\mathrm{d}\boldsymbol{v}_2. \tag{2.25}$$

It is common in the kinetic theory literature (see [30, 31]) to instead treat this as an equation over the mass density: $f(t;\boldsymbol{x}_1;\boldsymbol{v}_1) := Nm\,p(t;\boldsymbol{x}_1;\boldsymbol{v}_1)$, where m is the mass of a particle. This effectively amounts to a rescaling, and if we take the limit $N\to\infty$, leads to the equation

$$\partial_t f + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1} f = \frac{\varepsilon^{d-1}}{m}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-2}}\int_0^\infty \left[\zeta_{\boldsymbol{H}}\left({}^*f\,{}^*f_{(2)}\right) - f\,f_{(2)}\right]|\boldsymbol{v}_2-\boldsymbol{v}_1|\,r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}^{d-2}\,\mathrm{d}\boldsymbol{v}_2, \tag{2.26}$$

where $\zeta_{\boldsymbol{H}}$ may also be rescaled by the change to f if it is nonlinear and where we have used similar shorthand notation for the density arguments.

2.1.5. Inner (two-particle) solution

Equation (2.13) is a first-order quasilinear PDE that may be solved by the method of characteristics giving

$$\frac{\mathrm{d}\tilde{\boldsymbol{x}}_{21}}{\mathrm{d}\tau} = \boldsymbol{v}_2 - \boldsymbol{v}_1, \tag{2.27a}$$

$$\frac{\mathrm{d}\boldsymbol{v}_1}{\mathrm{d}\tau} = \boldsymbol{H}(-\tilde{\boldsymbol{x}}_{21}, \boldsymbol{v}_1 - \boldsymbol{v}_2), \tag{2.27b}$$

$$\frac{\mathrm{d}\boldsymbol{v}_2}{\mathrm{d}\tau} = \boldsymbol{H}(\tilde{\boldsymbol{x}}_{21}, \boldsymbol{v}_2 - \boldsymbol{v}_1), \tag{2.27c}$$

$$\frac{1}{\tilde{p}_2}\frac{\mathrm{d}\tilde{p}_2}{\mathrm{d}\tau} = -\nabla_{\boldsymbol{v}_1} \cdot \boldsymbol{H}(-\tilde{\boldsymbol{x}}_{21}, \boldsymbol{v}_1 - \boldsymbol{v}_2) - \nabla_{\boldsymbol{v}_2} \cdot \boldsymbol{H}(\tilde{\boldsymbol{x}}_{21}, \boldsymbol{v}_2 - \boldsymbol{v}_1), \tag{2.27d}$$

where τ is the arc length along a characteristic. Note that, if we set $t = \varepsilon\tau$, the characteristic equations for $\tilde{\boldsymbol{x}}_{21}$, $\boldsymbol{v}_1$ and $\boldsymbol{v}_2$ are equivalent to the evolution equations for an isolated two-particle interaction and, as such, the characteristic projections correspond to the trajectories of this two-particle interaction. The limit $\tau \to \infty$ represents the outgoing trajectory from an interaction, and taking the limit $\tau \to -\infty$ along the same characteristic represents the corresponding incoming trajectory. We will label the limiting outgoing quantities as $\tilde{\boldsymbol{x}}_{21}^+$, $\boldsymbol{v}_1^+$, $\boldsymbol{v}_2^+$ and $\tilde{p}_2^+$ and the incoming as $\tilde{\boldsymbol{x}}_{21}^-$, $\boldsymbol{v}_1^-$, $\boldsymbol{v}_2^-$ and $\tilde{p}_2^-$. Solving the characteristic equations will allow us to determine the outgoing quantities in terms of the incoming and determine the functions $\zeta_{\boldsymbol{H}}$ and $\psi_{\boldsymbol{H}}$ through

$$\tilde{p}_2^+(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^+; \boldsymbol{v}_1^+, \boldsymbol{v}_2^+) = \zeta_{\boldsymbol{H}}\left(\tilde{p}_2^-(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^-; \boldsymbol{v}_1^-, \boldsymbol{v}_2^-)\right) \text{ and } (\boldsymbol{v}_1^+, \boldsymbol{v}_2^+) = \psi_{\boldsymbol{H}}(\boldsymbol{v}_1^-, \boldsymbol{v}_2^-). \tag{2.28}$$

We use the $+$ and $-$ superscript notation in the inner region and the star notation once we match to the outer region solution.

We now outline some special cases that allow us to simplify the characteristic equations, and the rest of this chapter will be focused on specific interactions where we are able to analytically solve the characteristics and determine $\zeta_{\boldsymbol{H}}$ and $\psi_{\boldsymbol{H}}$.

If $\boldsymbol{H}$ is independent of $\boldsymbol{v}$ (or more generally $\boldsymbol{H}$ satisfies the divergence-free condition: $\sum_i \nabla_{\boldsymbol{v}_i} \cdot \sum_{j\neq i} \boldsymbol{H}_{ij} = 0$), $\tilde{p}_2$ is constant along characteristics and $\zeta_{\boldsymbol{H}}$ is the identity.

If $\boldsymbol{H}$ is antisymmetric, i.e. $\boldsymbol{H}(-\boldsymbol{x}, -\boldsymbol{v}) = -\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v})$, changing coordinates to scaled centre of mass, $\bar{\boldsymbol{v}} = \boldsymbol{v}_1 + \boldsymbol{v}_2$, and relative velocity, $\boldsymbol{u} = \boldsymbol{v}_2 - \boldsymbol{v}_1$, the characteristic equations can be simplified to

$$\frac{\mathrm{d}\tilde{\boldsymbol{x}}_{21}}{\mathrm{d}\tau} = \boldsymbol{u}, \tag{2.29a}$$

$$\frac{\mathrm{d}\bar{\boldsymbol{v}}}{\mathrm{d}\tau} = 0, \tag{2.29b}$$

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\tau} = 2\boldsymbol{H}(\tilde{\boldsymbol{x}}_{21}, \boldsymbol{u}), \tag{2.29c}$$

$$\frac{1}{\tilde{p}_2}\frac{\mathrm{d}\tilde{p}_2}{\mathrm{d}\tau} = -2\nabla_{\boldsymbol{u}} \cdot \boldsymbol{H}(\tilde{\boldsymbol{x}}_{21}, \boldsymbol{u}). \tag{2.29d}$$

This is a commonly considered special case, even when the Liouville property is not satisfied

(e.g. in [15]). If we then consider a one dimensional case where H is separable in x and v, i.e. $H(x,v) = \xi(x)\nu(v)$, then $\tilde{p}_2\nu(u)$ is conserved along characteristics as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}\tau}\left(\tilde{p}_2\nu(u)\right) &= \frac{\mathrm{d}\tilde{p}_2}{\mathrm{d}\tau}\nu(u) + \frac{\mathrm{d}\nu(u)}{\mathrm{d}\tau}\tilde{p}_2 \\ &= -2\frac{\partial}{\partial u}\left(H(\tilde{x}_{21},u)\right)\tilde{p}_2\nu(u) + \frac{\mathrm{d}\nu(u)}{\mathrm{d}u}\frac{\mathrm{d}u}{\mathrm{d}\tau}\tilde{p}_2 \\ &= -2\xi(\tilde{x}_{21})\frac{\mathrm{d}\nu(u)}{\mathrm{d}u}\tilde{p}_2\nu(u) + 2\frac{\mathrm{d}\nu(u)}{\mathrm{d}u}\xi(\tilde{x}_{21})\nu(u)\tilde{p}_2 \\ &= 0.\end{aligned} \tag{2.30}$$

We have used non-boldface variables to represent the respective one-dimensional quantities, a convention we will follow throughout this thesis.

2.2. Elastic interactions: the Boltzmann equation

In the original derivation of the Boltzmann equation, the interactions are assumed to be elastic; that is, they conserve energy and momentum. Using the notation $(\boldsymbol{v}_1^*, \boldsymbol{v}_2^*) = \psi_{\boldsymbol{H}}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ for the post-interaction velocities we have

$$\boldsymbol{v}_1^* + \boldsymbol{v}_2^* = \boldsymbol{v}_1 + \boldsymbol{v}_2 \text{ and } |\boldsymbol{v}_1^*|^2 + |\boldsymbol{v}_2^*|^2 = |\boldsymbol{v}_1|^2 + |\boldsymbol{v}_2|^2.$$

Following [108] and [24], this system of $d+1$ equations in $2d$ unknowns has a solution defined in terms of $d-1$ parameters. Two useful representations of the solution, commonly called the $\boldsymbol{\sigma}$- and $\boldsymbol{\omega}$-representations, are

$$\begin{aligned}\boldsymbol{v}_1^* &= \frac{\boldsymbol{v}_1 + \boldsymbol{v}_2}{2} - \frac{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}{2}\boldsymbol{\sigma}, \\ \boldsymbol{v}_2^* &= \frac{\boldsymbol{v}_1 + \boldsymbol{v}_2}{2} + \frac{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}{2}\boldsymbol{\sigma},\end{aligned} \tag{2.31}$$

and

$$\begin{aligned}\boldsymbol{v}_1^* &= \boldsymbol{v}_1 + \left((\boldsymbol{v}_2 - \boldsymbol{v}_1)\cdot\boldsymbol{\omega}\right)\boldsymbol{\omega}, \\ \boldsymbol{v}_2^* &= \boldsymbol{v}_2 - \left((\boldsymbol{v}_2 - \boldsymbol{v}_1)\cdot\boldsymbol{\omega}\right)\boldsymbol{\omega},\end{aligned} \tag{2.32}$$

where $\boldsymbol{\sigma}$ and $\boldsymbol{\omega}$ both vary over the $(d-1)$ unit sphere $\mathbb{S}^{d-1}$. Figure 2.2 shows a representation of the pre- and post-collisional velocities. The unit vector $\boldsymbol{\sigma}$ represents the direction of $\boldsymbol{v}_1^*$ and $\boldsymbol{v}_2^*$ relative to the centre of mass velocity $(\boldsymbol{v}_1 + \boldsymbol{v}_2)/2$, while $\boldsymbol{\omega}$ represents the impact direction, or the direction each particle's velocity changes during the interaction (see Figure 2.3 for specific interactions).

If the interaction $\boldsymbol{H}(\boldsymbol{x},\boldsymbol{v})$ is to be energy conserving, we must have $\nabla_{\boldsymbol{v}}\cdot\boldsymbol{H}(\boldsymbol{x},\boldsymbol{v}) = 0$ from Hamilton's equations. This gives $\mathrm{d}\tilde{p}_2/\mathrm{d}\tau = 0$ from (2.27d) and implies $\tilde{p}_2$ is constant along trajectories, with the system satisfying the Liouville equation, (2.5). As such, we have that $\zeta_{\boldsymbol{H}}$ is the identity and, defining the relative velocity as $\boldsymbol{u} := \boldsymbol{v}_2 - \boldsymbol{v}_1$ and assuming the particles

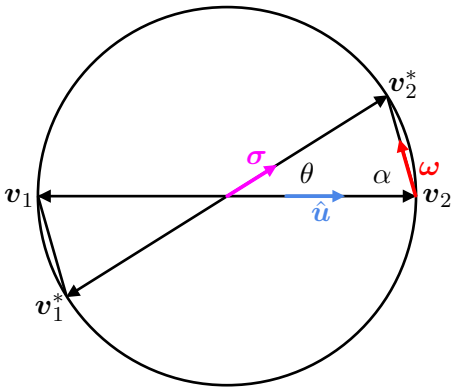


Figure 2.2.: Velocity space map $(\boldsymbol{v}_1^*, \boldsymbol{v}_2^*) = \psi_{\boldsymbol{H}}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ for an elastic interaction. The unit vectors $\hat{\boldsymbol{u}}$, $\boldsymbol{\omega}$, and $\boldsymbol{\sigma}$ are the direction of the incoming relative velocity, the impact direction and the outgoing relative velocity, respectively.

have mass 1, (2.26) becomes the Boltzmann equation:

$$\partial_t f + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} f = \varepsilon^{d-1} \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-2}} \int_0^\infty \left({}^*f \, {}^*f_{(2)} - f \, f_{(2)} \right) |\boldsymbol{u}| \, r^{d-2} \, \mathrm{d}r \, \mathrm{d}\boldsymbol{\xi}^{d-2} \, \mathrm{d}\boldsymbol{v}_2. \tag{2.33}$$

The integral on the RHS of this equation is commonly written in operator form as $Q(f, f)$. In terms of the $\boldsymbol{\sigma}$ representation, it is often expressed as

$$Q(f, f) = \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-1}} \left({}^*f \, {}^*f_{(2)} - f \, f_{(2)} \right) B_{\boldsymbol{\sigma}}(\boldsymbol{u}, \boldsymbol{\sigma}) \, \mathrm{d}\boldsymbol{\sigma} \, \mathrm{d}\boldsymbol{v}_2, \tag{2.34}$$

where $B_{\boldsymbol{\sigma}}(\boldsymbol{u}, \boldsymbol{\sigma})$ is called the collision kernel. This collision kernel can be thought of heuristically as the (unnormalised) probability that a pair of interacting particles are deflected to the relative direction $\boldsymbol{\sigma}$ after interacting, given their relative velocity $\boldsymbol{u}$. Due to the Galilean invariance of the interactions, $B_{\boldsymbol{\sigma}}$ only depends on the magnitude of the relative velocity $|\boldsymbol{u}|$ and the deviation or scattering angle θ (the angle between pre- and post-collisional relative velocities), defined by $\cos\theta = \hat{\boldsymbol{u}} \cdot \sigma$, where $\hat{\boldsymbol{u}} := \boldsymbol{u}/|\boldsymbol{u}|$ (see Figures 2.2 or 2.3). Thus, it is often written as $B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|, \hat{\boldsymbol{u}} \cdot \sigma)$ or $B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|, \theta)$. A similar representation using the $\boldsymbol{\omega}$ representation is also available, with modified collision kernel $B_{\boldsymbol{\omega}}(\boldsymbol{u}, \boldsymbol{\omega})$. These representations are often used in physical systems of microscopic particles as, in many cases, the deviation angle after an interaction, θ, is observable, while the impact parameter, r, is not. The collision kernel is related to the differential cross section, Σ, commonly defined for scattering systems, by $B_\sigma(\boldsymbol{u}, \sigma) = |\boldsymbol{u}|\Sigma(\boldsymbol{u}, \sigma)$.

Noting that $\mathrm{d}\boldsymbol{\sigma}$ is a surface element of $\mathbb{S}^{d-1}$ we can express it as $\mathrm{d}\boldsymbol{\sigma} \equiv \sin^{d-2}\theta \, \mathrm{d}\boldsymbol{\xi}^{d-2}$, where $\theta \in [0, \pi]$ is the deviation angle and, as before, $\mathrm{d}\boldsymbol{\xi}^{d-2}$ is an element of $\mathbb{S}^{d-2}$. Note that, due to momentum conservation, the path of the interacting particles is planar, and the collision kernel is invariant to rotations in $\boldsymbol{\xi}^{d-2}$. We can thus relate the $\boldsymbol{\sigma}$-representation of the collision kernel to the impact parameter representation through

$$B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|, \theta) \sin^{d-2}\theta \, \mathrm{d}\boldsymbol{\xi}^{d-2} = \varepsilon^{d-1} |\boldsymbol{u}| \, r^{d-2} \, \mathrm{d}r \, \mathrm{d}\boldsymbol{\xi}^{d-2},$$

which means, if a functional relationship between r and θ can be defined, that

$$B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|,\theta) = \varepsilon^{d-1}\left(\frac{r}{\sin\theta}\right)^{d-2}\frac{\mathrm{d}r}{\mathrm{d}\theta}\,|\boldsymbol{u}|. \tag{2.35}$$

Specific examples of $B_{\boldsymbol{\sigma}}$ will be derived in §2.2.3.

2.2.1. Cutoff kernels

Although heuristically interpreted as a probability, the collision kernel is not integrable if the interaction $\boldsymbol{H}(\boldsymbol{x},\boldsymbol{v})$ does not have compact support in $\boldsymbol{x}$, no matter how fast it decays as $|\boldsymbol{x}| \to \infty$. This can be seen as a result of the diverging number of 'grazing' interactions with large impact parameter $r \to \infty$ (or, equivalently, deviation angle $\theta \to 0$). Considering (2.33) we have that

$$\int_0^{r_c} |\boldsymbol{u}|\, r^{d-2}\,\mathrm{d}r = \frac{|\boldsymbol{u}| r_c^{d-1}}{d-1},$$

which is clearly unbounded if we take $r_c \to \infty$. One solution (introduced by Grad [55]) is to assume the interaction is cut off at some finite impact parameter r_c. This makes the collision kernel integrable and allows us to formally split the collision operator into gain and loss terms:

$$Q^+(f,f) = \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} {}^*f\,{}^*f_{(2)} B_{\boldsymbol{\sigma}}(\boldsymbol{u},\boldsymbol{\sigma})\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_2,$$

and

$$Q^-(f,f) = -\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} f\, f_{(2)} B_{\boldsymbol{\sigma}}(\boldsymbol{u},\boldsymbol{\sigma})\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_2,$$

respectively.

Note, however, that even though the collision kernel is not integrable, $Q(f,f)$ produces a finite result even without this cutoff for many commonly considered applications. This is because, given the interaction $\boldsymbol{H}$ satisfies the short-range condition (2.2), for large impact parameters, the interacting particles experience minimal deviation, meaning that ${}^*f\,{}^*f_{(2)} - f\,f_{(2)} \to 0$ as $r \to \infty$. For certain $\boldsymbol{H}$ and f, this decay is fast enough that $Q(f,f)$ is finite. In practice, in this thesis, we will use the cutoff where necessary to split the collision operator but will often simplify calculations by taking $r_c \to \infty$. Many authors have considered the formal convergence to the Boltzmann equation without cutoff (e.g. [39] for the linear Boltzmann equation).

2.2.2. Weak form, conservation laws and the $\mathcal{H}$-theorem

In many cases, we are concerned with macroscopic observable quantities that can be defined as the expected value of a function $\varphi(\boldsymbol{v})$ over f, i.e.

$$\langle f,\varphi\rangle_{x,v} := \int_{\mathbb{R}^d\times\Omega} f\,\varphi\,\mathrm{d}\boldsymbol{x}_1\mathrm{d}\boldsymbol{v}_1,$$

where we have defined the inner product operator $\langle\cdot,\varphi\rangle_{x,v}$. Acting with this inner product operator on (2.33), we get

$$\partial_t\langle f,\varphi\rangle_{x,v} = \langle Q(f,f),\varphi\rangle_{x,v}\,. \tag{2.36}$$

Using the fact that the transformations $(\boldsymbol{v}_1, \boldsymbol{v}_2, \sigma) \to (\boldsymbol{v}_1^*, \boldsymbol{v}_2^*, \hat{\boldsymbol{u}})$ and $(\boldsymbol{v}_1, \boldsymbol{v}_2) \to (\boldsymbol{v}_2, \boldsymbol{v}_1)$ are involutions with unit Jacobians (see [108] §2.3 and [31] §2.1 for details), we can write the RHS as

$$\begin{aligned}
&\langle Q(f,f), \varphi \rangle_{x,v} \\
&\quad = \int_{\mathbb{R}^d \times \mathbb{R}^d \times \Omega} \int_{\mathbb{S}^{d-1}} B_{\boldsymbol{\sigma}}(\boldsymbol{u}, \sigma) f f_{(2)} (\varphi^* - \varphi) \, \mathrm{d}\boldsymbol{\sigma} \, \mathrm{d}\boldsymbol{x}_1 \mathrm{d}\boldsymbol{v}_1 \mathrm{d}\boldsymbol{v}_2 \qquad (2.37) \\
&\quad = \frac{1}{2} \int_{\mathbb{R}^d \times \mathbb{R}^d \times \Omega} \int_{\mathbb{S}^{d-1}} B_{\boldsymbol{\sigma}}(\boldsymbol{u}, \sigma) f f_{(2)} (\varphi^* + \varphi^*_{(2)} - \varphi - \varphi_{(2)}) \, \mathrm{d}\boldsymbol{\sigma} \, \mathrm{d}\boldsymbol{x}_1 \mathrm{d}\boldsymbol{v}_1 \mathrm{d}\boldsymbol{v}_2 \qquad (2.38) \\
&\quad = -\frac{1}{4} \int_{\mathbb{R}^d \times \mathbb{R}^d \times \Omega} \int_{\mathbb{S}^{d-1}} B_{\boldsymbol{\sigma}}(\boldsymbol{u}, \sigma) \left({}^*f \, {}^*f_{(2)} - f f_{(2)} \right) \left(\varphi^* + \varphi^*_{(2)} - \varphi - \varphi_{(2)} \right) \mathrm{d}\boldsymbol{\sigma} \, \mathrm{d}\boldsymbol{x}_1 \mathrm{d}\boldsymbol{v}_1 \mathrm{d}\boldsymbol{v}_2, \qquad (2.39)
\end{aligned}$$

where we have used similar shorthand for the arguments of φ as the arguments of f. These represent the various weak forms of the Boltzmann equation, which are each useful in different contexts. The first, (2.37), was written down by Maxwell [78] before Boltzmann derived the strong form. From (2.38), we can see that

$$\langle Q(f,f), \varphi \rangle_{x,v} = 0$$

if and only if $\varphi(\boldsymbol{v})$ satisfies

$$\varphi^* + \varphi^*_{(2)} = \varphi + \varphi_{(2)},$$

i.e. that φ is invariant across interactions. It can be shown that the solutions φ of this equation are linear combinations of the elementary collision invariants: $\{1, v_i, |\boldsymbol{v}|^2\}$, where $\{v_i \text{ for } i \in 1, \dots d\}$ are the components of $\boldsymbol{v}$ [23]. Substituting this into (2.36) we get the conservation laws

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^d \times \Omega} f \begin{pmatrix} 1 \\ \boldsymbol{v} \\ \frac{|\boldsymbol{v}|^2}{2} \end{pmatrix} \mathrm{d}\boldsymbol{x} \mathrm{d}\boldsymbol{v} = 0,$$

showing that total mass, momentum and kinetic energy are conserved. If we only integrate over velocity and not position, the same collision invariants set the RHS to 0. In this case, the transport term does not go to 0, and we have the local conservation laws

$$\frac{\mathrm{d}}{\mathrm{d}t} \int_{\mathbb{R}^d} f \begin{pmatrix} 1 \\ \boldsymbol{v}_i \\ \frac{|\boldsymbol{v}|^2}{2} \end{pmatrix} \mathrm{d}\boldsymbol{v} + \nabla_x \cdot \int_{\mathbb{R}^d} f \begin{pmatrix} \boldsymbol{v} \\ \boldsymbol{v} \otimes \boldsymbol{v} \\ \frac{|\boldsymbol{v}|^2}{2} \boldsymbol{v} \end{pmatrix} \mathrm{d}\boldsymbol{v} = 0. \qquad (2.40)$$

These laws form the basis for developing hydrodynamic limits and macroscopic descriptions of the system. It is possible to show that the only solutions to the equation $Q(f,f) = 0$ are the Maxwellian distributions, defined as

$$f_{\mathrm{Max}}(t; \boldsymbol{x}; \boldsymbol{v}) := \frac{\rho(t; \boldsymbol{x})}{(2\pi T(t; \boldsymbol{x}))^{d/2}} e^{-\frac{|\boldsymbol{v} - \boldsymbol{w}|}{T(t; \boldsymbol{x})}},$$

where $\rho(t;\boldsymbol{x})$, $\boldsymbol{w}(t;\boldsymbol{x})$ and $T(t;\boldsymbol{x})$ are the macroscopic density, bulk velocity and temperature, respectively. The fact that $Q(f,f) = 0$ implies that the system has reached statistical equilibrium, as collisions are not responsible for any further variation in the density f. Assuming the system is at this equilibrium, the local conservation laws (2.40) become the compressible Euler equations. It is also possible to derive the Navier-Stokes equations as a leading order system when perturbing around this equilibrium solution using a procedure known as the Chapman-Enskog expansion (see [31], Ch. 5).

Boltzmann's $\mathcal{H}$-functional, defined as

$$\mathcal{H}(f) := \int_{R^d\times\Omega} f\log f \,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v},$$

can be considered a representation of the information, or negative entropy in the distribution. Substituting $\varphi(f) = \log f$ into (2.36) with the collision operator in the form (2.39), we get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}&\mathcal{H}(f(t;\boldsymbol{x};\boldsymbol{v}))\\ &= -\frac{1}{4}\int_{\mathbb{R}^d\times\mathbb{R}^d\times\Omega}\int_{\mathbb{S}^{d-1}} B_{\boldsymbol{\sigma}}(\boldsymbol{u},\sigma)\left({}^*f\,{}^*f_{(2)} - ff_{(2)}\right)\log\frac{{}^*f{}^*f_{(2)}}{ff_{(2)}}\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{x}_1\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}_2 \le 0, \quad (2.41)\end{aligned}$$

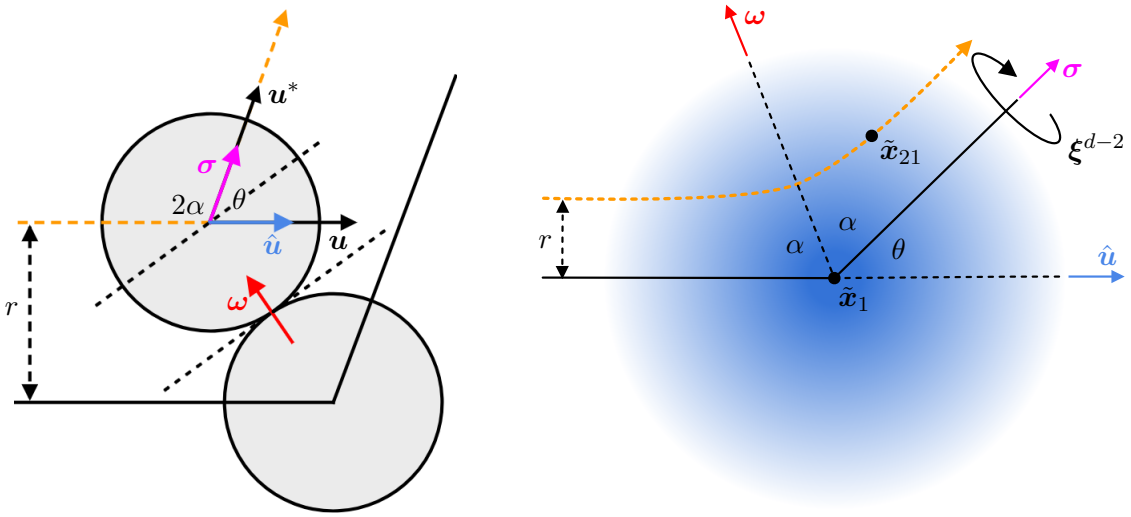
where the inequality comes from the fact that the function $(x,y)\mapsto(x-y)(\log x - \log y)$ is non-negative. This is Boltzmann's $\mathcal{H}$-theorem, implying that the system dissipates information (or increases entropy) over time, moving towards equilibrium. It is notable that this time asymmetry has emerged from a system that was based on time-reversible microscopic interactions. Although the individual-based system satisfies the Poincaré recurrence theorem, by applying Boltzmann's *Stosszahlansatz* and only assuming pre-collisional particles are uncorrelated we have broken the time-symmetry and the kinetic model becomes irreversible.

2.2.3. Example interactions

Hard-sphere elastic reflection

One common system studied in the context of the Boltzmann equation is that of hard spheres of radius ε colliding elastically (that is, they specularly reflect off each other upon impact). Considering Figure 2.3a, we see that, in inner coordinates scaled by ε, the relationship between the impact factor r and the angle α (between the incoming velocity, $\boldsymbol{u}$, and the direction between the centres of the two spheres at impact, $\boldsymbol{\omega}$) is given by $r = 2\sin\alpha$. As the angle between pre- and post-collisional relative velocities is then 2α, we have that the deviation angle is given by $\theta = \pi - 2\alpha$ and hence $\sin\theta = \sin 2\alpha = 2\sin\alpha\cos\alpha$. Substituting into (2.35), we then get

$$\begin{aligned}B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|,\theta) &= \varepsilon^{d-1}|\boldsymbol{u}|\left(\frac{2\sin\alpha}{2\sin\alpha\cos\alpha}\right)^{d-2} 2\cos\alpha\\ &= 2\varepsilon^{d-1}|\boldsymbol{u}|\,(\cos\alpha)^{3-d} = 2\varepsilon^{d-1}|\boldsymbol{u}|\left(\frac{1-\cos\theta}{2}\right)^{\frac{3-d}{2}}. \quad (2.42)\end{aligned}$$



(a) Elastic hard-sphere collision.

(b) Potential-force interaction.

Figure 2.3.: Relative position schematic for specific elastic interactions. The unit vectors $\hat{\boldsymbol{u}}$, $\boldsymbol{\omega}$, and $\boldsymbol{\sigma}$ are equivalent to those in Figure 2.2. (a) An elastic collision between two hard spheres of radius ε. The relative velocity reflects across a line perpendicular to the impact direction $\boldsymbol{\omega}$ upon collision. (b) The inner region for a potential force interaction. A point-particle is shown fixed at $\tilde{\boldsymbol{x}}_1$, with another particle at $\tilde{\boldsymbol{x}}_{21}$. The short-range potential interaction is shown with blue shading.

In three dimensions, we can see that the collision kernel is independent of the deviation angle and only depends on the relative velocity $|\boldsymbol{u}|$.

Potential-force systems

Another common system defines the forces between particles as the gradients of a radial potential. That is, the interaction in (2.1) is given by

$$\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v}) = -\frac{1}{m} \nabla_{\boldsymbol{x}} \phi(|\boldsymbol{x}|),$$

which gives accelerations

$$\dot{\boldsymbol{V}}_i = -\frac{1}{m} \sum_{j \neq i} \nabla_{\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon}} \phi\left(\frac{|\boldsymbol{X}_i - \boldsymbol{X}_j|}{\varepsilon}\right).$$

Using the equations for the conservation of energy and angular momentum in the reference frame of one of the particles, it is possible to define the relationship between the deviation angle θ, the relative velocity $|\boldsymbol{u}|$ and the impact parameter r. Following Cercignani [30, 31], we have that for a repulsive potential ϕ of infinite range with $\phi(|\boldsymbol{x}|) \to 0$ as $|\boldsymbol{x}| \to \infty$ and particles of mass $m = 1$,

$$\theta(r, |\boldsymbol{u}|) = \pi - 2r \int_{\rho_0}^{\infty} \frac{\rho^{-2}}{\sqrt{1 - \frac{r^2}{\rho^2} - 4\frac{\phi(\rho)}{|\boldsymbol{u}|^2}}} \, \mathrm{d}\rho = \pi - 2 \int_0^{\frac{r}{\rho_0}} \frac{1}{\sqrt{1 - a^2 - \frac{4}{|\boldsymbol{u}|^2}\phi(\frac{r}{a})}} \, \mathrm{d}a, \tag{2.43}$$

where ρ represents the radial distance of the two particles and ρ_0 is the distance of closest approach, defined as the positive root of $1 - \frac{r^2}{\rho_0^2} - 4\frac{\phi(\rho_0)}{|\boldsymbol{u}|^2} = 0$. Considering the power-law potential $\phi(|\boldsymbol{x}|) = c|\boldsymbol{x}|^{-(s-1)}$ for $s > 2$ (note that such a potential will only satisfy our short-range condition (2.2) with $s > d$) and defining the quantity

$$\beta := r\left(\frac{1}{4c}\right)^{\frac{1}{s-1}} |\boldsymbol{u}|^{\frac{2}{s-1}}, \tag{2.44}$$

we have that the relationship between θ and β is given by

$$\theta = \pi - 2\int_0^{a_0} \frac{1}{\sqrt{1 - a^2 - \left(\frac{a}{\beta}\right)^{s-1}}}\,\mathrm{d}a,$$

where a_0 satisfies $1 - a_0^2 - (a_0/b)^{s-1} = 0$. These equations specify θ as a function $\theta(\beta)$, and hence inverting β as a function $\beta(\theta)$. As such, using (2.44), it follows that

$$r = (4c)^{\frac{1}{s-1}} |\boldsymbol{u}|^{-\frac{2}{s-1}} \beta(\theta),$$

and substituting into (2.35) we have that

$$B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|, \theta) = \varepsilon^{d-1}|\boldsymbol{u}|^{\gamma} b(\theta), \quad \gamma = \frac{s - (2d-1)}{s-1},$$

where we have relabelled the angular part of the kernel

$$b(\theta) = (4c)^{\frac{d-1}{s-1}} \left(\frac{\beta(\theta)}{\sin\theta}\right)^{d-2} \beta'(\theta).$$

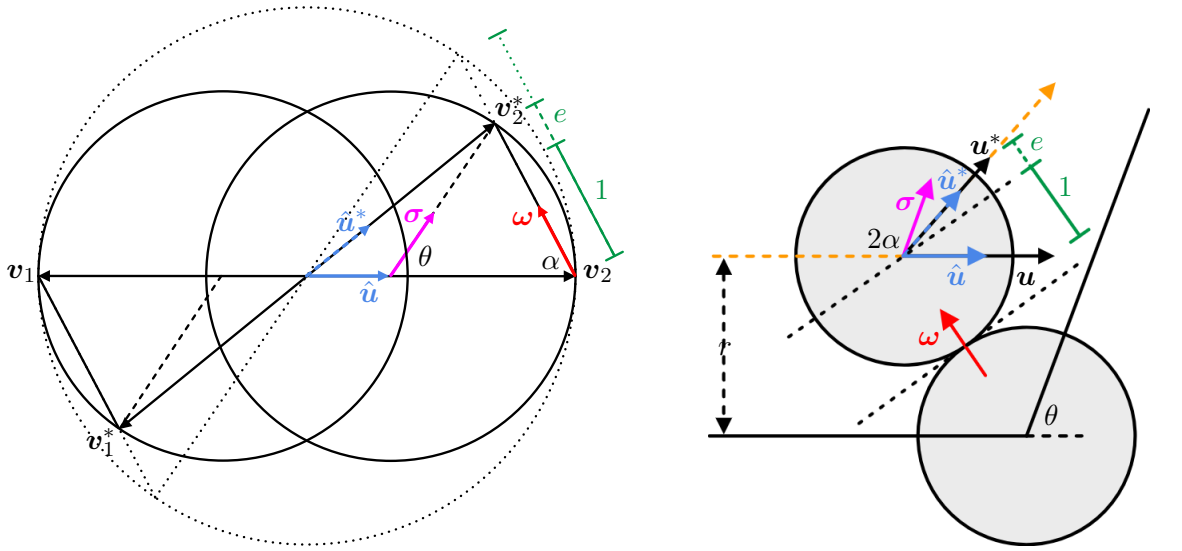
If the particles interact with a power-law potential with $s = 2d - 1$ (in three dimensions, where force is proportional to the inverse fifth power), the $|\boldsymbol{u}|$ dependence is removed, and the collision kernel becomes purely angular. Particles interacting with such forces are labelled Maxwellian molecules, and these are often used as a simplified base case for deriving analytical results. The limit $s \to \infty$ approaches the hard sphere case discussed previously. Potentials with $s > 2d - 1$ are often labelled hard potentials, and those with $s < 2d - 1$ soft potentials.

The functional dependence $\beta(\theta)$ in the angular part of the kernel cannot be computed explicitly in all cases, but it can be shown that $\sin^{d-2}\theta b(\theta)$ is smooth and has a non-integrable singularity as $\theta \to 0$. In three dimensions it scales as $O(\theta^{-1-2/(s-1)})$ as $\theta \to 0$. This is related to the non-integrability of the collision kernel discussed in §2.2.1.

2.3. Inelastic interactions (granular materials)

When relaxing the assumption that energy must be conserved across interactions, a common class of models involves scaling the relative velocity's normal component after an interaction by a restitution coefficient $e \in [0, 1]$ (not to be confused with the exponential constant). That is, we assume

$$(\boldsymbol{v}_2^* - \boldsymbol{v}_1^*) \cdot \boldsymbol{\omega} = -e\,(\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot \boldsymbol{\omega}, \tag{2.45}$$



(a) Inelastic velocity map.

(b) Inelastic hard-sphere collision

Figure 2.4.: (a) Velocity space map and (b) hard-sphere collision example for an inelastic interaction with restitution coefficient e. The ratio of the length of the dashed green line to the solid green line is e, showing that the magnitude of the component of velocity along the impact direction $\boldsymbol{\omega}$ is reduced by this factor after an interaction. The dotted lines in the velocity map show the equivalent interaction if $e = 1$. The unit vector $\boldsymbol{\sigma}$ is defined differently to the elastic case, instead representing the direction of the final velocity relative to its sphere of possible outcomes (or the direction of the post-collisional relative velocity if the interaction were elastic).

but maintain the conservation of momentum: $\boldsymbol{v}_1^* + \boldsymbol{v}_2^* = \boldsymbol{v}_1 + \boldsymbol{v}_2$ (see Figure 2.4a). Such models can be used to model a wide range of systems, but are most prominently studied in the field of granular gases, collections of particles that collide inelastically, larger than the microscopic gas molecules archetypical of the Boltzmann equation. This field has applications in areas such as industrial sorting of particulate solids, planetary rings or motion in a dusty atmosphere [73, 2, 95]. Combining (2.45) with the conservation of momentum, we again have a solution in terms of $d-1$ parameters. Many authors consider systems with constant restitution coefficient e, although it is natural to consider a dependence on the relative velocity and deviation angle, and some authors even consider a phenomenological dependence on the system temperature [109]. Corresponding representations to (2.46) and (2.47) are available in terms of $\boldsymbol{\omega} \in \mathbb{S}^{d-1}$, the impact direction, and $\boldsymbol{\sigma} \in \mathbb{S}^{d-1}$, now defined as the direction of the final velocity relative to the centre of its possible values. These are

$$\begin{aligned} \boldsymbol{v}_1^* &= \boldsymbol{v}_1 + \frac{1+e}{2} \left((\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot \boldsymbol{\omega} \right) \boldsymbol{\omega}, \\ \boldsymbol{v}_2^* &= \boldsymbol{v}_2 - \frac{1+e}{2} \left((\boldsymbol{v}_2 - \boldsymbol{v}_1) \cdot \boldsymbol{\omega} \right) \boldsymbol{\omega}, \end{aligned} \tag{2.46}$$

and

$$\begin{aligned} \boldsymbol{v}_1^* &= \frac{\boldsymbol{v}_1+\boldsymbol{v}_2}{2} - \frac{1-e}{4}(\boldsymbol{v}_2-\boldsymbol{v}_1) - \frac{1+e}{4}|\boldsymbol{v}_2-\boldsymbol{v}_1|\boldsymbol{\sigma}, \\ \boldsymbol{v}_2^* &= \frac{\boldsymbol{v}_1+\boldsymbol{v}_2}{2} + \frac{1-e}{4}(\boldsymbol{v}_2-\boldsymbol{v}_1) + \frac{1+e}{4}|\boldsymbol{v}_2-\boldsymbol{v}_1|\boldsymbol{\sigma}. \end{aligned} \tag{2.47}$$

See Figure 2.4a for a representation of the maps in velocity space. Note that if we take $e = 1$, this is equivalent to the elastic case discussed in 2.2, and if $e = 0$, the normal component of the velocity is fully dissipated in the interaction. The inelastic Boltzmann equation is given for one-particle mass density f by

$$\partial_t f + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} f = Q_{\mathcal{I}}(f, f), \tag{2.48}$$

where $Q_{\mathcal{I}}(f, f)$ is a modified version of the collision operator for inelastic interactions. Following [109], the collision operator in the $\boldsymbol{\sigma}$-representation, for hard sphere inelastic interactions in three dimensions with constant restitution coefficient is given by

$$Q_{\mathcal{I}}(f,f) = \int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}} \frac{|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1|}{e} {}^*f\, {}^*f_{(2)} - |\boldsymbol{v}_2-\boldsymbol{v}_1| f f_2 \,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_2.$$

The factor of $1/e$ is present in the gain term as it is the Jacobian of the backward interaction map $({}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2) = \psi^{-1}(\boldsymbol{v}_1, \boldsymbol{v}_2)$, and this is needed to ensure probability flux in velocity space is conserved across interactions. This will be analysed further in Chapter 3, where we consider a more generalised set of inelastic interactions. The relationship between the restitution coefficient and the collision kernels in each representation is, in general, more complex than the elastic case and is discussed in detail in [24]. Carillo, Hu, Ma and Rey [29], review results for inelastic collision kernels of the 'generalised hard sphere' type, given by

$$B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|, \cos\theta, E) = |\boldsymbol{u}|^{\lambda} b(\cos\theta) E(f)^{\gamma}, \tag{2.49}$$

where θ is the angle between σ and $\boldsymbol{u}$ and

$$E(f) := \frac{1}{2}\int_{\mathbb{R}^d\times\Omega} |\boldsymbol{v}|^2 f(t;\boldsymbol{x};\boldsymbol{v})\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}$$

represents the kinetic energy of the distribution. Similar to the elastic treatment, $\lambda = 1$ represents the physical hard-spheres case, while $\lambda = 0$ defines the simplified case of Maxwellian molecules. The classical treatments of these cases take $\gamma = 0$, removing the energy dependence, but in more complex cases (sometimes referred to as anomalous granular gases[96]), the generalised form with $\gamma \neq 0$ is taken.

Similarly to in the elastic case, the weak form, showing the evolution of integrated observables, can be derived by multiplying (2.48) by the desired function $\varphi(\boldsymbol{v})$, integrating over the domain of the first particle and simplifying by manipulating integral bounds between pre- and post-collisional velocities. Written in terms of the inelastic collision operator, the weak

form is equivalent to (2.36). The RHS is given explicitly in terms of the $\boldsymbol{\sigma}$-representation as

$$\begin{aligned}&\langle Q_{\mathcal{I}}(f,f),\varphi\rangle_{x,v}\\&\quad=\frac{1}{2}\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{S}^{d-1}} ff_{(2)}\left(\varphi^*+\varphi^*_{(2)}-\varphi-\varphi_{(2)}\right)B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|,\cos\theta,E(f))\,\mathrm{d}\boldsymbol{\sigma}\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{x}_1,\end{aligned}$$

where we have assumed a general dependence of the collision kernel on $|\boldsymbol{u}|$, $\cos\theta$ and $E(f)$.

2.3.1. Energy dissipation

One of the primary qualitative differences with the case of elastic interaction is that energy is not conserved over time. Specifically, we have from (2.46) that

$$|\boldsymbol{v}_1^*|^2+|\boldsymbol{v}_2^*|^2-|\boldsymbol{v}_1|^2-|\boldsymbol{v}_2|^2=-\frac{1-e^2}{2}((\boldsymbol{v}_1-\boldsymbol{v}_2)\cdot\omega)^2\leq 0. \tag{2.50}$$

This energy dissipation means that, if no other processes are included, the distribution will asymptotically approach a Dirac delta at the mean velocity (for some kernels, it is possible to reach the Dirac delta in finite time). Using the weak form, along with (2.50) and the conservation of mass and momentum over individual interactions we obtain the global conservation laws

$$\frac{\mathrm{d}}{\mathrm{d}t}\int_{\mathbb{R}^d\times\Omega} f\begin{pmatrix}1\\ \boldsymbol{v}\\ \frac{|\boldsymbol{v}|^2}{2}\end{pmatrix}\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}=\begin{pmatrix}0\\0\\-D(f)\end{pmatrix}, \tag{2.51}$$

where $D(f)\geq 0$ is the energy dissipation functional, given by

$$D(f)=\frac{1-e^2}{8}\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d\times\mathbb{S}^{d-1}}|\boldsymbol{u}|^2\left(\frac{1-\hat{\boldsymbol{u}}\cdot\boldsymbol{\sigma}}{2}\right)ff_{(2)}B_{\boldsymbol{\sigma}}(|\boldsymbol{u}|,\cos\theta,E(f))\,\mathrm{d}\boldsymbol{\sigma}\,\mathrm{d}\boldsymbol{v}_2\,\mathrm{d}\boldsymbol{v}_1\,\mathrm{d}\boldsymbol{x}_1,$$

where we have used the relationship between $\boldsymbol{u}$, $\boldsymbol{\sigma}$ and $\boldsymbol{\omega}$

$$|\boldsymbol{u}\cdot\boldsymbol{\omega}|=|\boldsymbol{u}|(\hat{\boldsymbol{u}}\cdot\boldsymbol{\omega})=|\boldsymbol{u}|\sqrt{\frac{1-\hat{\boldsymbol{u}}\cdot\boldsymbol{\sigma}}{2}},$$

from [24]. If the generalised hard spheres kernel (2.49) is assumed, the angular dependence of the energy dissipation functional can be factored out, giving

$$D(f)=b_1\frac{1-e^2}{8}E^\gamma\int_{\Omega\times\mathbb{R}^d\times\mathbb{R}^d}|\boldsymbol{u}|^{\lambda+2}ff_{(2)}\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{x}_1, \tag{2.52}$$

where

$$b_1=\left|\mathbb{S}^{d-2}\right|\int_0^\pi\sin^2\left(\frac{\theta}{2}\right)\sin^{d-2}\theta\,b(\cos\theta)\,\mathrm{d}\theta,$$

where θ is the angle between $\boldsymbol{u}$ and $\boldsymbol{\sigma}$. If we assume $b(\cos\theta)<\infty$ for all θ or define an angular cutoff as in 2.2.1, this is integrable.

Following [29], [80] and [81], and assuming the spatially uniform case, we can use Jensen's

inequality twice on the energy dissipation equation given by (2.51) and (2.52) to get

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}E(f)(t) &= -b_1\frac{1-e^2}{8}E(f)^\gamma(t)\int_{\mathbb{R}^d\times\mathbb{R}^d} ff_{(2)}|\boldsymbol{v}_1-\boldsymbol{v}_2|^{\lambda+2}\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1\\ &\le b_1\frac{1-e^2}{8}E(f)^\gamma(t)\int_{\mathbb{R}^d} f\left|\int_{\mathbb{R}^d} f_{(2)}(\boldsymbol{v}_1-\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2\right|^{\lambda+2}\mathrm{d}\boldsymbol{v}_1\\ &= b_1\frac{1-e^2}{8}E(f)^\gamma(t)\int_{\mathbb{R}^d} f\,|\boldsymbol{v}_1|^{\lambda+2}\,\mathrm{d}\boldsymbol{v}_1\\ &\le b_1\frac{1-e^2}{8}E(f)^\gamma(t)\left(\int_{\mathbb{R}^d} f|\boldsymbol{v}_1|^2\,\mathrm{d}\boldsymbol{v}_1\right)^{(\lambda+2)/2}\\ &= 2^{\frac{\lambda+2}{2}}b_1\frac{1-e^2}{8}E(f)^{1+\gamma+\lambda/2}(t),\end{aligned} \tag{2.53}$$

where in the third step we have used the conservation of mass and momentum to set $\int_{\mathbb{R}^d} f(t;\boldsymbol{v})\,\mathrm{d}\boldsymbol{v} = 1$ and $\int_{\mathbb{R}^d} f(t;\boldsymbol{v})\boldsymbol{v}\,\mathrm{d}\boldsymbol{v} = 0$. Defining $C_{e,b} := 2^{\frac{\lambda+2}{2}}b_1\frac{1-e^2}{8}$ and $\eta := \gamma + \lambda/2$ the following bounds on the large time energy dissipation can be established. For Maxwellian molecules (with $\lambda = \gamma = 0$) the energy decays exponentially with

$$E(t) \le E(0)e^{-C_{e,b}t}.$$

Hard spheres (with $\lambda = 1$ and $\gamma = 0$) exhibit the quadratic *Haff's cooling law* [59]:

$$E(t) \le \left(E(0)^{-1/2} + C_{e,b}\frac{t}{2}\right)^{-2},$$

a seminal result in the field of inelastic particle interactions. For more general λ and γ (the case with $\lambda = 1$ and $\gamma \neq 0$ is sometimes called the anomalous granular gases case) we have

$$E(t) \le \begin{cases}(E(0)^\eta + C_{e,b}\eta t)^{-1/\eta}, & \eta > 0\\ E(0)e^{C_{e,b}t}, & \eta = 0\\ (E(0)^\eta - C_{e,b}\eta t)^{-1/\eta}, & \eta < 0.\end{cases}$$

Note that the case with $\eta > 0$ dissipates to the Dirac delta in finite time. These bounds have been proved rigorously and shown to be sharp for the Maxwell case in [13, 11], hard spheres in [81], and anomalous gases in [96].

2.3.2. Similarity scaling

As the system energy dissipates over time, it is common to search for a similarity scaling of the variables to analyse the behaviour at long times. This is used both for deriving more rigorous results about the moment evolution of f and to 'zoom in' to the mean of the velocity distribution as time progresses to avoid challenges that come from the distribution collapsing to the Dirac delta (for instance, when searching for a numerical solution). Following Mischler and Mouhot [81], for the spatially-uniform hard sphere case in dimension d (sometimes called a

homogenous cooling state [109]), we have the evolution equation

$$\partial_t f = Q_{\mathcal{I}}(f, f), \tag{2.54}$$

with

$$Q_{\mathcal{I}}(f, f) = \int_{\mathbb{R}^d \times \mathbb{S}^{d-1}} \left(\frac{1}{e^2} {}^* f {}^* f_{(2)} - f f_{(2)} \right) |\boldsymbol{u}| b(\cos\theta) \, \mathrm{d}\boldsymbol{\sigma} \mathrm{d}\boldsymbol{v}_2,$$

and $b(y) \propto (1-y)^{-(3-d)/2}$ similarly to (2.42). The density f is scaled such that $\int_{\mathbb{R}^d} f(t; \boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} = 1$ and $\int_{\mathbb{R}^d} \boldsymbol{v} f(t; \boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} = 0$ (these can be maintained for all t by conservation of mass and momentum). The collision operator has the homogeneity property

$$Q_{\mathcal{I}}(\gamma f(t; \lambda \boldsymbol{v}), \gamma f(t; \lambda \boldsymbol{v}))(\boldsymbol{v}) = \gamma^2 \lambda^{-(d+1)} \, Q_{\mathcal{I}}(f(t; \lambda \boldsymbol{v}), f(t; \lambda \boldsymbol{v}))(\lambda \boldsymbol{v}), \tag{2.55}$$

which can be found by changing variables from $\boldsymbol{v}$ to $\lambda \boldsymbol{v}$.

Following [81], We search for a rescaled solution of the form

$$f(t; \boldsymbol{v}) = \mathcal{F}(t) \tilde{f}(\tilde{t}; \tilde{\boldsymbol{v}}),$$

with $\tilde{t} = \mathcal{T}(t)$ and $\tilde{\boldsymbol{v}} = \mathcal{V}(t)\boldsymbol{v}$, requiring $\mathcal{F}(0) = \mathcal{V}(0) = 1$ and $\mathcal{T}(0) = 0$ so that $f(0, \boldsymbol{v}) = \tilde{f}(0, \boldsymbol{v})$. Substituting this into the evolution equation and using the homogeneity property, (2.55), we find

$$\mathcal{F}'(t) \tilde{f} + \mathcal{F}(t) \, \mathcal{T}'(t) \, \partial_{\tilde{t}} \tilde{f} + \mathcal{F}(t) \, \frac{\mathcal{V}'(t)}{\mathcal{V}(t)} \, \tilde{\boldsymbol{v}} \cdot \nabla_{\tilde{\boldsymbol{v}}} \tilde{f} = \mathcal{F}(t)^2 \mathcal{V}(t)^{-(d+1)} Q(\tilde{f}, \tilde{f})(\tilde{\boldsymbol{v}}).$$

We choose $\mathcal{V}(t)$ and $\mathcal{F}(t)$ such that we can define a self-similar solution

$$\tilde{f}(\tilde{t}; \tilde{\boldsymbol{v}}) = \tilde{F}(\tilde{\boldsymbol{v}})$$

that does not depend on time. In this case, conservation of mass implies that

$$1 = \int_{\mathbb{R}^d} f(t; \boldsymbol{v}) \, \mathrm{d}\boldsymbol{v} = \frac{\mathcal{F}(t)}{\mathcal{V}(t)^d} \int_{\mathbb{R}^d} \tilde{F}(\tilde{\boldsymbol{v}}) \, \mathrm{d}\tilde{\boldsymbol{v}},$$

which means that $\mathcal{F}(t) = \mathcal{V}(t)^d$. Substituting this into the evolution equation, we get

$$\mathcal{V}'(t) \nabla_{\tilde{\boldsymbol{v}}} \cdot (\tilde{\boldsymbol{v}} \tilde{F}) = Q(\tilde{F}, \tilde{F}),$$

where we have used the fact that $\nabla_{\tilde{\boldsymbol{v}}} \cdot (\tilde{\boldsymbol{v}} \tilde{F}) = d \, \tilde{F} + \tilde{\boldsymbol{v}} \cdot \nabla_{\tilde{\boldsymbol{v}}} \tilde{F}$, for dimension d. Thus, we must have that $\mathcal{V}'(t) = c$ for some constant c, which using the initial datum, implies that $\mathcal{V}(t) = 1 + c\,t$. Considering now the evolution equation for the time-dependent $\tilde{f}$, we have

$$\mathcal{V}(t) \, \mathcal{T}'(t) \, \partial_{\tilde{t}} \tilde{f} + c \, \nabla_{\tilde{\boldsymbol{v}}} \cdot (\tilde{\boldsymbol{v}} \tilde{f}) = Q(\tilde{f}, \tilde{f}).$$

We choose $\mathcal{T}(t)$ to make this equation as simple as possible by setting $\mathcal{V}(t) \, \mathcal{T}'(t) = 1$, which gives $\mathcal{T}(t) = \log(1 + c\,t)/c$. As changing c can be countered by a rescaling of $\tilde{f}$ and $\mathcal{T}'(t)$, we can set $c = 1$ without loss of generality. This gives the scaling

$$\mathcal{F}(t) = (1 + t)^d, \quad \mathcal{T}(t) = \log(1 + t), \text{ and } \mathcal{V}(t) = 1 + t. \tag{2.56}$$

With this scaling, we have the rescaled evolution equation

$$\partial_{\tilde{t}} \tilde{f} = Q(\tilde{f}, \tilde{f}) - \nabla_{\tilde{\boldsymbol{v}}} \cdot (\tilde{\boldsymbol{v}} \tilde{f}), \tag{2.57}$$

where the extra term $\nabla_{\tilde{\boldsymbol{v}}} \cdot (\tilde{\boldsymbol{v}} \tilde{f})$ is generally known as an anti-drift term [109]. Associated to any solution $\tilde{f}(\tilde{t}; \tilde{\boldsymbol{v}})$ of this rescaled equation is a solution to the original evolution equation

$$f(t; \boldsymbol{v}) = (1+t)^d \, \tilde{f}(\log(1+t); (1+t)\boldsymbol{v}). \tag{2.58}$$

Furthermore, for any solution $\tilde{F}(\tilde{\boldsymbol{v}})$ to the stationary scaled equation

$$\nabla_{\tilde{\boldsymbol{v}}} \cdot (\tilde{\boldsymbol{v}} \tilde{F}) = Q(\tilde{F}, \tilde{F}), \tag{2.59}$$

we can define a self-similar solution F to the original evolution equation by

$$F(t; \boldsymbol{v}) = (1+t)^d \, \tilde{F}((1+t)\boldsymbol{v}). \tag{2.60}$$

Defining

$$M_s[g](t) := \int_{R^d} g(t; \boldsymbol{v}) |\boldsymbol{v}|^s \, \mathrm{d}\boldsymbol{v}$$

to be the s-moment functional with respect to v and noting that we have

$$\int_{\mathbb{R}^d} f(t; \boldsymbol{v}) \, |\boldsymbol{v}|^s \, \mathrm{d}\boldsymbol{v} = (1+t)^s \int_{\mathbb{R}^d} \tilde{f}(\log(1+t); \tilde{\boldsymbol{v}}) \, |\tilde{\boldsymbol{v}}|^s \, \mathrm{d}\tilde{\boldsymbol{v}},$$

and

$$\int_{\mathbb{R}^d} \tilde{f}(\tilde{t}; \tilde{\boldsymbol{v}}) \, |\tilde{\boldsymbol{v}}|^s \, \mathrm{d}\tilde{\boldsymbol{v}} = e^{s\tilde{t}} \int_{\mathbb{R}^d} f(e^{\tilde{t}} - 1; \boldsymbol{v}) \, |\boldsymbol{v}|^s \, \mathrm{d}\boldsymbol{v},$$

we have the following relationships between the moments of the scaled and unscaled solutions:

$$M_s[f](t) = (1+t)^s M_s[\tilde{f}](\log(1+t)),$$

and

$$M_s[\tilde{f}](t) = e^{st} M_s[f](e^t - 1).$$

This form of scaled solution is not the only way to avoid the problems associated with the collapse of the velocity distribution. If it is relevant to the physical situation being modelled, some authors add a Gaussian thermostat or heat bath as a way to input energy into the system [109]. This consists of adding a term of the form $T \, \nabla_{\boldsymbol{v}}^2 f$ to the RHS of the evolution equation, where T represents the 'external temperature', or magnitude of the energy input. Additionally, a drift term $\alpha \nabla_{\boldsymbol{v}} \cdot (\boldsymbol{v} f)$, with $\alpha > 0$, representing external friction forces on the particles, is sometimes added in these cases.

2.3.3. Overpopulated distribution tails

One of the more surprising predictions of inelastic kinetic theory is that of non-Maxwellian velocity distribution tails [109]. For some models, the velocity distribution can decay slower than a Maxwellian distribution as $|\boldsymbol{v}| \to \infty$, counterintuitively suggesting the inelasticity has

resulted in a greater proportion of fast particles when compared to the elastic Maxwellian equilibrium solution. Here we present a heuristic argument from [109], summarising results from [43].

In the limit of large velocities, splitting the collision operator, the gain term $Q_{\mathcal{I}}^{+}(f,f)$ must be dominated by the loss term $Q_{\mathcal{I}}^{-}(f,f)$, as the velocity distribution moves inward toward the Dirac delta over time. If the collision kernel is proportional to $|\boldsymbol{v}_2 - \boldsymbol{v}_1|^{\lambda}$, the loss term, and hence the full collision operator, is asymptotically proportional to $|\boldsymbol{v}_1|^{\lambda}$ as $|\boldsymbol{v}_1| \to \infty$. Making the radially symmetric ansatz $f(|\boldsymbol{v}|) = e^{-\alpha|\boldsymbol{v}|^{\beta}}$ it follows that $Q_{\mathcal{I}}^{-}(f,f) \propto |\boldsymbol{v}_1|^{\lambda} e^{-\alpha|\boldsymbol{v}_1|^{\beta}}$ and $\nabla_{\boldsymbol{v}_1} \cdot (\boldsymbol{v}_1 f) = |\boldsymbol{v}_1|^{\beta} e^{-\alpha|\boldsymbol{v}_1|^{\beta}}$. Balancing terms in the steady state solution of the rescaled evolution equation (2.57), it then follows that $\beta = \lambda$. As $\lambda = 1$ for hard sphere kernels this suggests that the the velocity distribution tail is exponentially decreasing like $f(\boldsymbol{v}) \sim e^{-\alpha|\boldsymbol{v}|}$ as $|\boldsymbol{v}| \to \infty$ for a homogeneously cooling state in this case. Considering also that $\nabla_{\boldsymbol{v}_1}^2 f \propto |\boldsymbol{v}_1|^{2\beta-2} e^{-\alpha|\boldsymbol{v}_1|^{\beta}}$ under the radially symmetric ansatz, by balancing dominant terms in the steady state equation it follows that $\beta = 1 + \lambda/2$ for systems with an added heat bath term ($\nabla_{\boldsymbol{v}}^2 f$), and $\beta = 2$ for systems with a further added drift term ($\nabla_{\boldsymbol{v}} \cdot (\boldsymbol{v} f)$). This argument predicts velocity distribution tails scaling like $e^{-\alpha|\boldsymbol{v}_1|^{3/2}}$ for hard sphere systems with a heat bath and like $e^{-\alpha|\boldsymbol{v}_1|^2}$ for hard sphere systems with a heat bath and added friction.

2.4. Flocking models

A popular class of many-particle model (often called a many-agent model in this context) in the literature today is centred around systems exhibiting emergent collective behaviour and self-organisation. These systems can be in fields as varied as swarming and collective motion in animals and bacteria [35, 34, 106, 107, 37, 10, 5, 65]; control of robotics, spacecraft and other autonomous agents [92, 4, 70]; social consensus and opinion formation [51, 84]; and the distribution of wealth in an economy [41]. Reviews of the mathematical approaches are given in [28, 27, 32, 107]. Typically, the models in economic or social contexts are first-order, with models including attraction terms directly depending on the modelled quantity. Models in animal collective behaviour or autonomous agents are often second-order (as has been the focus of this chapter), with models generally involving some kind of alignment in velocity, modulated by the spatial distance between agents. They may also typically include an attraction force at long ranges and a repulsion force at short ranges.

Three seminal individual-based models in this field are the Vicsek model [106], the D'orsogna model [40], and the Cucker-Smale model [35, 34]. The Vicsek model is a discrete-time model where the position, $\boldsymbol{X}_i$ of agents in two dimensions evolves as

$$\boldsymbol{X}_i(t + \Delta t) = \boldsymbol{X}_i(t) + V(\cos(\Theta_i), \sin(\Theta_i))\Delta t,$$

where V is a constant speed and $Theta_i$ is the orientation of agent i. This orientation evolves through

$$\Theta_i(t + \Delta t) = \langle \Theta(t) \rangle_r + \Delta\Theta,$$

where $\langle \Theta(t) \rangle_r$ represents the average orientation of agents in a circle of radius r around agent i and $\Delta\theta$ represents a random number chosen at each time step uniformly in the interval $[-\eta/2, \eta/2]$. The staple feature of this model is the alignment of agents with their neighbours in orientation. The literature extending and applying this model is vast. For example, the model has been extended to continuous-time SDE versions in higher dimensions, [38] (further discussed in [18]) and to body-attitude coordination (considering further internal spatial dimensions of the agents) in [37].

In the D'orsogna model [40], agents' positions and velocities evolve according to the ODEs

$$\begin{aligned} \dot{\boldsymbol{X}}_i(t) &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i(t) &= (\alpha - \beta|\boldsymbol{V}_i|^2)\boldsymbol{V}_i - \sum_{\substack{j=1 \\ j\neq i}}^{N} \nabla_{\boldsymbol{X}_i} \phi(|\boldsymbol{X}_j - \boldsymbol{X}_i|), \end{aligned} \tag{2.61}$$

where the interaction potential ϕ is generally taken to be the Morse potential

$$\phi(|\boldsymbol{x}|) = C_r e^{-|\boldsymbol{x}|/l_r} - C_a e^{-|\boldsymbol{x}|/l_a}.$$

The constants l_a and l_r represent the ranges of attraction and repulsion, respectively, while C_a and C_r represent the respective amplitudes. Thus, with $l_a > l_r$, this potential represents attraction at long ranges and repulsion at short ranges, a common feature of collective behaviour models. The parameter α scales a self-propulsive term, while the term scaled by β represents friction described by Rayleigh's law. The combination of these terms will cause the speeds of the agents to tend toward $\sqrt{\alpha/\beta}$. Significantly different behaviour is observed for this model for different choices of the Morse potential ratios l_a/l_r and C_a/C_r. These vary from the collapse of all agents to one cluster to swarming solutions in a ring to dispersing behaviour where the agents separate.

In the Cucker-Smale model [35, 34], agent velocities move toward a weighted average of surrounding particle velocities, with the weights given by a symmetric kernel that is dependent on the relative position:

$$\begin{aligned} \dot{\boldsymbol{X}}_i(t) &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i(t) &= \frac{1}{N} \sum_{\substack{j=1 \\ j\neq i}}^{N} \phi(|\boldsymbol{X}_i - \boldsymbol{X}_j|)(\boldsymbol{V}_j - \boldsymbol{V}_i). \end{aligned} \tag{2.62}$$

The original kernel proposed by Cucker and Smale [35] was

$$\phi(|\boldsymbol{x}|) = \frac{\beta}{(1 + |\boldsymbol{x}|^2)^s}, \tag{2.63}$$

where β and s are positive constants, although many other kernels have been analysed since. The symmetry of the interaction between particles i and j implies the total momentum and

mean velocity are conserved:

$$\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{V}_i(t)\right)=0 \implies \bar{\boldsymbol{V}}(t):=\frac{1}{N}\sum_{i=1}^{N}\boldsymbol{V}_i(t)=\bar{\boldsymbol{V}}(0).$$

Unlike the Vicsek et al. or D'orsogna et al. models, agents in the Cucker-Smale model do not have a constant or characteristic velocity. Instead, the characteristic qualitative features of the system are the dissipation of energy and flocking behaviour. The energy dissipation can be seen as

$$\begin{aligned}\frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{N}\sum_{i=1}^{N}|\boldsymbol{V}_i(t)-\bar{\boldsymbol{V}}|^2\right)&=-\frac{1}{2N^2}\sum_{i,j=1}^{N}\phi(|\boldsymbol{X}_j-\boldsymbol{X}_i|)|\boldsymbol{V}_i(t)-\bar{\boldsymbol{V}}|^2\\&\leq-\left(\min_{i,j}\phi(|\boldsymbol{X}_j-\boldsymbol{X}_i|)\right)\frac{1}{N}\sum_{i=1}^{N}|\boldsymbol{V}_i-\bar{\boldsymbol{V}}|^2. \quad (2.64)\end{aligned}$$

The system is said to converge to a flock if there exists a finite D_∞ such that $\mathrm{diam}\{\boldsymbol{X}_i(t)\}_{i=1}^N < D_\infty$ for all $t>0$ and there exists a $\boldsymbol{V}_\infty\in\mathbb{R}^d$ such that $\max_i|\boldsymbol{V}_i(t)-\boldsymbol{V}_\infty|\to 0$ as $t\to\infty$ [100]. Naturally, for the Cucker-Smale model (2.62), if flocking occurs, the flock velocity would be $\boldsymbol{V}_\infty=\bar{\boldsymbol{V}}$. This flocking is said to be unconditional if it occurs for all possible initial conditions, and conditional if it only occurs with a restricted set. Cucker and Smale showed unconditional flocking for the kernel (2.63) with $s<1/2$ or conditional flocking with $s\geq 1/2$ and additional restrictions on the initial density. Many authors have rigorously derived flocking conditions for other kernels. For example, Ha and Tadmor [58] showed unconditional flocking for kernels satisfying the 'fat-tail' condition:

$$\int^{\infty}\phi(r)\,\mathrm{d}r=\infty.$$

Note the contrast with our short-range condition (2.2). Many of the papers analysing the large time and large N behaviour of the Cucker-Smale model use kinetic and hydrodynamic models (and the equivalent definitions of flocking), which we turn to now.

2.4.1. Kinetic models of flocking with long-range communication

We will focus here on the development of kinetic and hydrodynamic models for the Cucker-Smale model (2.62), but we note that similar style mesoscopic and macroscopic models are available for the Vicsek and D'orsogna models and their variants. Adapting an argument from [25], which uses a technique developed in [79], assume the outgoing velocities $\boldsymbol{v}_1^*$, and $\boldsymbol{v}_2^*$ from two agents entering a binary interaction with phases $(\boldsymbol{x}_1,\boldsymbol{v}_1)$ and $(\boldsymbol{x}_2,\boldsymbol{v}_2)$ are given by the maps

$$\boldsymbol{v}_1^*=\mathscr{C}(\boldsymbol{x}_1,\boldsymbol{v}_1;\boldsymbol{x}_2,\boldsymbol{v}_2)=[1-\gamma\phi(|\boldsymbol{x}_2-\boldsymbol{x}_1|)]\,\boldsymbol{v}_1+\gamma\phi(|\boldsymbol{x}_2-\boldsymbol{x}_1|)\,\boldsymbol{v}_2$$

and

$$\boldsymbol{v}_2^*=\mathscr{C}(\boldsymbol{x}_2,\boldsymbol{v}_2;\boldsymbol{x}_1,\boldsymbol{v}_1)=[1-\gamma\phi(|\boldsymbol{x}_2-\boldsymbol{x}_1|)]\,\boldsymbol{v}_2+\gamma\phi(|\boldsymbol{x}_2-\boldsymbol{x}_1|)\,\boldsymbol{v}_1,$$

where ϕ is considered to be the original Cucker-Smale kernel (2.63) and $\gamma \leq 1/2$ is a parameter representing the strength of the interaction. These maps are in the form of the velocity map defined earlier for short-range interactions (2.22), but with ψ depending on $|\boldsymbol{x}_2 - \boldsymbol{x}_1|$ and γ, i.e. $(\boldsymbol{v}_1^*, \boldsymbol{v}_2^*) = \psi_{\gamma,|\boldsymbol{x}_2-\boldsymbol{x}_1|}(\boldsymbol{v}_1, \boldsymbol{v}_2)$. As in [25], using these maps, we can write the Boltzmann-type kinetic equation

$$\partial_t f + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} f = Q(f, f),$$

with the interaction operator defined as

$$Q(f,f) = \eta \int_{\mathbb{R}^d} \int_{\mathbb{R}^d} \left| D\psi^{-1}_{\gamma,|\boldsymbol{x}_2-\boldsymbol{x}_1|} \right| {}^* f^* f_{(2)} - f f_{(2)} \, \mathrm{d}\boldsymbol{x}_2 \mathrm{d}\boldsymbol{v}_2, \tag{2.65}$$

using the same shorthand for densities with modified arguments introduced previously. In this model, the collision frequency η is assumed to be constant. The Jacobian of the velocity map is $\left| D\psi^{-1}_{\gamma,|\boldsymbol{x}_2-\boldsymbol{x}_1|} \right| = 1/(1 - 2\gamma\phi(|\boldsymbol{x}_2 - \boldsymbol{x}_1|))^d$. As in the Boltzmann equation for inelastic collisional interactions (see §2.3), the Jacobian can be avoided in the weak formulation. For inner products of the density with a test function $\varphi(\boldsymbol{x}; \boldsymbol{v})$, the weak form is given by

$$\partial_t \int_{\mathbb{R}^{2d}} f\varphi \, \mathrm{d}\boldsymbol{v}_1 \mathrm{d}\boldsymbol{x}_1 + \int_{\mathbb{R}^{2d}} (\boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} \varphi) f \, \mathrm{d}\boldsymbol{v}_1 \mathrm{d}\boldsymbol{x}_1 = \eta \int_{\mathbb{R}^{4d}} f f_{(2)} (\varphi^* - \varphi) \, \mathrm{d}\boldsymbol{v}_1 \mathrm{d}\boldsymbol{x}_1 \mathrm{d}\boldsymbol{v}_2 \mathrm{d}\boldsymbol{x}_2, \tag{2.66}$$

where $\varphi^* \equiv \varphi(\boldsymbol{x}; \boldsymbol{v}^*)$. Using the velocity maps we have

$$\boldsymbol{v}_1^* - \boldsymbol{v} = \gamma\phi(|\boldsymbol{x}_1 - \boldsymbol{x}_2|)(\boldsymbol{v}_2 - \boldsymbol{v}_1),$$

and we can then expand the RHS of (2.66) in a Taylor series in the limit of $\gamma \ll 1$ to give

$$\eta\gamma \int_{\mathbb{R}^{4d}} (\nabla_{\boldsymbol{v}_1} \varphi \cdot (\boldsymbol{v}_2 - \boldsymbol{v}_1)) \phi(|\boldsymbol{x}_2 - \boldsymbol{x}_1|) f f_{(2)} \mathrm{d}\boldsymbol{v}_1 \mathrm{d}\boldsymbol{x}_1 \mathrm{d}\boldsymbol{v}_2 \mathrm{d}\boldsymbol{x}_2 + O\left(\eta\gamma^2\right).$$

This is known as the grazing collision limit as it assumes the velocity changes due to interaction are small. The second order Taylor series term is $O\left(\eta\gamma^2\right)$ as the integral part is bounded above because of the energy dissipation properties of the system. In the limit of grazing collisions and high collision frequency, such that $\eta\gamma^2 \ll 1$ and $\eta\gamma = O(1)$, we can truncate this expansion at first order. Converting back to the strong form, we obtain the Vlasov-type equation

$$\partial_t f + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} f = \eta\gamma \nabla_{\boldsymbol{v}_1} \cdot [f F(f)], \tag{2.67}$$

where the operator F is given by

$$F[f](t; \boldsymbol{x}_1; \boldsymbol{v}_1) = (\eta\gamma\phi(|\boldsymbol{x}_1|)\nabla_{\boldsymbol{v}_1} W(\boldsymbol{v}_1)) * f = \eta\gamma \int_{\mathbb{R}^{2d}} \phi(|\boldsymbol{x}_1 - \boldsymbol{x}_2|)(\boldsymbol{v}_2 - \boldsymbol{v}_1) f_{(2)} \mathrm{d}\boldsymbol{x}_2 \mathrm{d}\boldsymbol{v}_2,$$

where $W(\boldsymbol{v}_1) = \frac{1}{2}|\boldsymbol{v}_1|^2$ and $*$ represents a convolution in $\boldsymbol{x}$ and $\boldsymbol{v}$, as is seen in the final equality. As was noted in [79], this truncated operator maintains the energy dissipation properties of the Boltzmann collision operator (2.65).

Similar results about the flocking of the Cucker-Smale model and kinetic equation have

been derived rigorously in the mean-field limit, often using a sequence of empirical measures

$$f^N(t;\boldsymbol{x};\boldsymbol{v}) = \frac{1}{N}\sum_{i=1}^{N}\delta(\boldsymbol{x}-\boldsymbol{X}_i(t))\delta(\boldsymbol{v}-\boldsymbol{V}_i(t))$$

as $N \to \infty$. In [58], unconditional flocking was derived from the BBGKY hierarchy for bounded radial kernels, and in [57], convergence to the Vlasov equation with flocking dissipation was shown. In [86], these results were extended to the singular kernels $\phi(r) = |r|^{-s}$ $s \in (0, 1/2)$ and in [25] it was shown that the model derived in [58] converges exponentially fast to the mean velocity.

Defining the density $\rho(t;\boldsymbol{x})$, the bulk velocity $\boldsymbol{w}(t;\boldsymbol{x})$, and the local temperature $T(t;\boldsymbol{x})$ as

$$\rho = \int_{\mathbb{R}^d} f \,\mathrm{d}\boldsymbol{v}, \quad \rho\,\boldsymbol{w} = \int_{\mathbb{R}^d} \boldsymbol{v} f \,\mathrm{d}\boldsymbol{v}, \quad d\rho T = \int_{R^d} |\boldsymbol{v}-\boldsymbol{w}|^2 f \,\mathrm{d}\boldsymbol{v},$$

where d is the system dimension, it is possible to define a hydrodynamic flocking model from the kinetic model [28]. To close this model, it is often assumed that energy fluctuations are negligible, i.e. $T(t;\boldsymbol{x}) = 0$, taking the monokinetic ansatz $f(t;\boldsymbol{x};\boldsymbol{v}) = \rho(t;\boldsymbol{x})\delta(\boldsymbol{v}-\boldsymbol{w}(t;\boldsymbol{x}))$. With these approximations, the hydrodynamic flocking model of the Cucker-Smale system with long-range interaction kernels is given by

$$\begin{aligned} \partial_t\rho + \nabla_{\boldsymbol{x}}\cdot(\rho\boldsymbol{w}) &= 0, \\ \rho\partial_t\boldsymbol{w} + \nabla_{\boldsymbol{x}_1}\cdot(\rho\boldsymbol{w}\otimes\boldsymbol{w}) &= \int_{R^d}\phi(|\boldsymbol{x}_2-\boldsymbol{x}_1|)\rho(t;\boldsymbol{x}_1)\rho(t;\boldsymbol{x}_2)(\boldsymbol{w}(t;\boldsymbol{x}_2)-\boldsymbol{w}(t;\boldsymbol{x}_1))\,\mathrm{d}\boldsymbol{x}_2. \end{aligned} \tag{2.68}$$

Many variations to the original Cucker-Smale model are possible. Adding a Brownian motion term in the velocity evolution ODE of the individual-based model (2.62), the kinetic model gains an extra diffusion term $\mathrm{D}\nabla_{\boldsymbol{v}}^2 f$, similar in form to a Fokker-Planck equation [28]. Motsch and Tadmor [83] define a modified, non-symmetric interaction kernel, with interaction strength weighted as a proportion of the total interaction on an agent:

$$\begin{aligned} \dot{\boldsymbol{X}}_i &= \boldsymbol{V}_i \\ \dot{\boldsymbol{V}}_i &= \alpha\sum_{\substack{j=1\\ j\neq i}}^{N} a_{ij}(\boldsymbol{V}_j-\boldsymbol{V}_i), \qquad a_{ij} = \frac{\phi(|\boldsymbol{X}_i-\boldsymbol{X}_j|)}{\sum_{k=1}^{N}\phi(|\boldsymbol{X}_i-\boldsymbol{X}_k|)}. \end{aligned} \tag{2.69}$$

This model loses the properties of momentum conservation and energy dissipation of the original Cucker-Smale model but is expected to be more physically realistic for animal collective behaviour when there is varying density among the flock. Motsch and Tadmor show unconditional flocking for this model with the more restrictive fat tail condition $\int^{\infty}\phi^2(r)\,\mathrm{d}r = \infty$.

Other modifications to the Cucker-Smale model in the literature include adding an external force potential [98], combining non-local and local alignment terms [45], adding a time delay [42, 99], and modelling multiple interacting species [62].

2.4.2. Flocking with short-range communication

Several authors have analysed variants on the Cucker-Smale model with strictly bounded interaction kernels, where the support of the kernel is less than the maximum separation of the particles in the model. A common result is deriving conditions on the density or interaction strength for flocking to occur in either the individual-based or hydrodynamic sense. Motsch and Tadmor study first-order systems of opinion formation and second-order systems of flocking and consider an adjacency matrix defined by the interaction strength of each agent on the others [84]. They show that if the graph connectivity of this matrix is maintained, then the system must flock for both symmetric and non-symmetric kernels. This result has been extended to the continuum case for the one-dimensional Cucker-Smale model [56], for the cutoff Motsch-Tadmor model [71], and for general Cucker-Smale kernels in higher dimensions [82]. In these, the authors show explicit bounds for the coupling strength and conditions on the initial density for flocking to occur. Shvydkoy and Tadmor [100] show flocking in a velocity alignment model with topological kernels. This is similar to the cutoff Motsch-Tadmor model, except that interactions are weighted based on other agents in a conical region between them, rather than a ball around the first agent. Blanchet and Degond [12] consider a system where, at random times, a particle adopts the velocity of its nearest neighbour or the average of its k nearest neighbours' velocities. They derive a kinetic equation in the limit of the system size going to infinity and assuming the propagation of chaos.

2.5. Clustering models

Consider an alternative short-range Cucker-Smale model, where the interaction kernel $\phi(|\boldsymbol{x}|)$ has length scale ε and the interaction strength is scaled by $1/\varepsilon$:

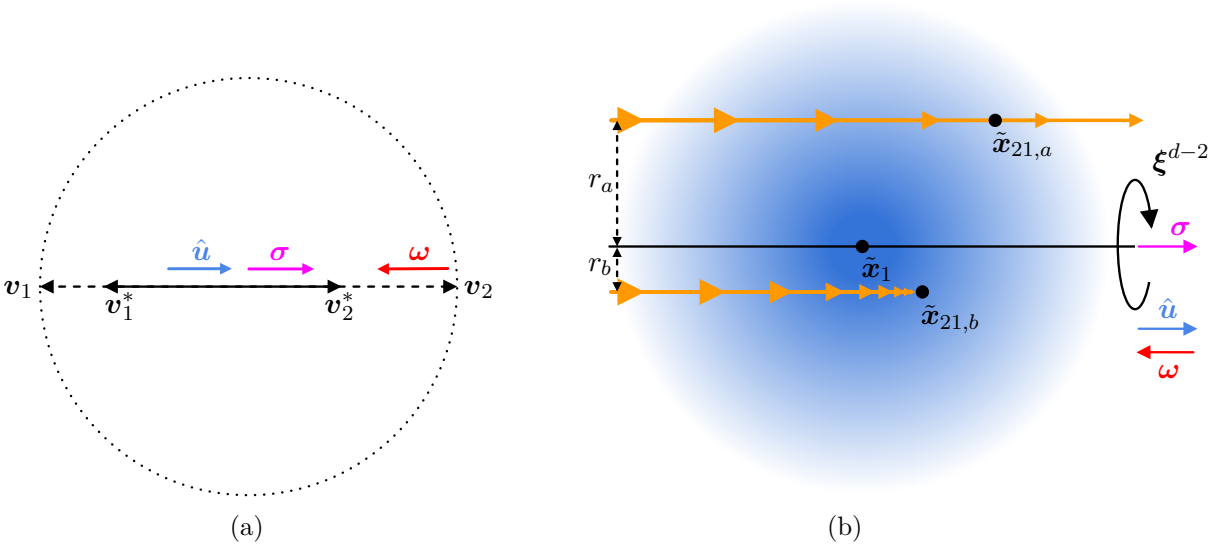
$$\begin{aligned} \dot{\boldsymbol{X}}_i &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{N} \phi\left(\frac{|\boldsymbol{X}_i - \boldsymbol{X}_j|}{\varepsilon}\right) (\boldsymbol{V}_j - \boldsymbol{V}_i). \end{aligned} \tag{2.70}$$

This can be expressed in the generalised form of (2.1) with $\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v}) = -\phi(|\boldsymbol{x}|)\boldsymbol{v}$. We allow the interaction kernel ϕ to have infinite support but take it to satisfy the short-range condition from (2.2), which in this case means that

$$\phi(|\boldsymbol{x}|) = O(|\boldsymbol{x}|^{-d-\delta}) \text{ as } |\boldsymbol{x}| \to \infty,$$

for $\delta > 0$. As in §2.1, we assume the system is in the low volume fraction regime with $\varepsilon \ll 1$, $\varepsilon^d N \ll 1$ and take the Boltzmann-Grad limit $\varepsilon^{d-1} N \equiv 1$ if $d > 1$. This ensures isolated binary interactions predominate, assuming particles are uncorrelated, and should allow us to follow the systematic procedure from §2.1 of matched asymptotic expansions in the short interaction length scale to derive a kinetic equation in the form of (2.26).

Following §2.1, the kinetic model depends on the interaction $\boldsymbol{H}$ through the maps $\zeta_{\boldsymbol{H}}$



(a) (b)

Figure 2.5.: (a) Velocity map for a non-clustering interaction for the short-range Cucker-Smale model. (b) Two inner region interactions for the short-range Cucker-Smale model, one where the interacting particles leave the inner region (labelled a) and one where they remain bound after interacting (labelled b). We will call trajectory b a *clustering* interaction. In contrast to the interactions described in §2.2 and §2.3 (see Figures 2.2, 2.3 and 2.4), the relative velocity after a non-clustering interaction is parallel to the incoming relative velocity, rendering the representations in terms of $\boldsymbol{\omega}$ and $\boldsymbol{\sigma}$ degenerate.

and $\psi_{\boldsymbol{H}}$ from (2.21) and (2.22). These can be derived from the solution to the inner region evolution equation (2.13). As (2.70) contains an antisymmetric interaction term, the solution to the inner region equation is given by simplified characteristic equations (2.29), given using relative velocity, $\boldsymbol{u} = \boldsymbol{v}_2 - \boldsymbol{v}_1$, and centre of mass velocity, $\bar{\boldsymbol{v}} = (\boldsymbol{v}_1 + \boldsymbol{v}_2)/2$, as

$$\frac{\mathrm{d}\tilde{\boldsymbol{x}}_{21}}{\mathrm{d}\tau} = \boldsymbol{u}, \tag{2.71a}$$

$$\frac{\mathrm{d}\bar{\boldsymbol{v}}}{\mathrm{d}\tau} = 0, \tag{2.71b}$$

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\tau} = -2\phi(|\tilde{\boldsymbol{x}}_{21}|)\boldsymbol{u}, \tag{2.71c}$$

$$\frac{1}{\tilde{p}_2}\frac{\mathrm{d}\tilde{p}_2}{\mathrm{d}\tau} = 2\phi(|\tilde{\boldsymbol{x}}_{21}|)\nabla_{\boldsymbol{u}}\cdot\boldsymbol{u} = 2d\,\phi(|\tilde{\boldsymbol{x}}_{21}|). \tag{2.71d}$$

As in §2.1.5, we consider the limiting values of the characteristic trajectories as $\tau \to \pm\infty$ and label them with $+$ and $-$ superscripts, respectively. As the centre of mass is conserved (see (2.71b)) and the time evolution of the relative velocity is proportional to itself (see (2.71c)), the relative motion of the two particles will be planar. Thus, without loss of generality, assume we are considering a two-dimensional trajectory approaching from $\tilde{\boldsymbol{x}}_{21}^- = (-\infty, r)$ with $\boldsymbol{u}^- = (u^-, 0)$ (see Figure 2.5b). Higher dimensional trajectories that approach from infinity can be recovered through rotations to this system. As $\mathrm{d}\tilde{\boldsymbol{x}}_{21}/\mathrm{d}\tau$ and $\mathrm{d}\boldsymbol{u}/\mathrm{d}\tau$ are proportional to $\boldsymbol{u}$, the second coordinate of $\boldsymbol{u}$ will remain 0 and the second coordinate of $\tilde{\boldsymbol{x}}_{21}$ will remain r

across the trajectory. As such, we have $\tilde{\boldsymbol{x}}_{21}(\tau) = (x(\tau), r)$ and $\boldsymbol{u}(\tau) = (u(\tau), 0)$. Since

$$\frac{\mathrm{d}u}{\mathrm{d}x} = \frac{\mathrm{d}u}{\mathrm{d}\tau}\frac{\mathrm{d}\tau}{\mathrm{d}x} = \frac{1}{u}\frac{\mathrm{d}u}{\mathrm{d}\tau},$$

it follows from (2.71c) that

$$\frac{\mathrm{d}u}{\mathrm{d}x} = -2\phi\left(|\tilde{\boldsymbol{x}}_{21}|\right) = -2\phi\left(\sqrt{r^2 + x^2}\right).$$

Integrating this over x gives

$$u(x) - u^- = -2\int_{-\infty}^{x} \phi\left(\sqrt{r^2 + (x')^2}\right) \mathrm{d}x' = -2\Phi(r, x), \tag{2.72}$$

where we have defined $\Phi(r, x) := \int_{-\infty}^{x} \phi\left(\sqrt{r^2 + (x')^2}\right) \mathrm{d}x'$. As ϕ is non-negative, $\Phi(r, x)$ is a monotonically increasing function of x. Using the fact that $\mathrm{d}x/\mathrm{d}\tau = u$, separating variables and integrating, we get an implicit definition of $x(\tau)$:

$$\int_{-\infty}^{x} \frac{1}{u^- - 2\,\Phi(r, x')} \mathrm{d}x' = \int_{-\infty}^{\tau} \mathrm{d}\tau' = \tau. \tag{2.73}$$

If $\lim_{x' \to \infty} 2\Phi(r, x') < u^-$, the $\mathrm{d}x'$ integrand remains finite and positive and $x \to \infty$ as $\tau \to \infty$. Using (2.72), it follows in these cases that

$$u^+ = u^- - \lim_{x' \to \infty} 2\Phi(r, x') = u^- - 2\Phi(r), \tag{2.74}$$

where we have defined

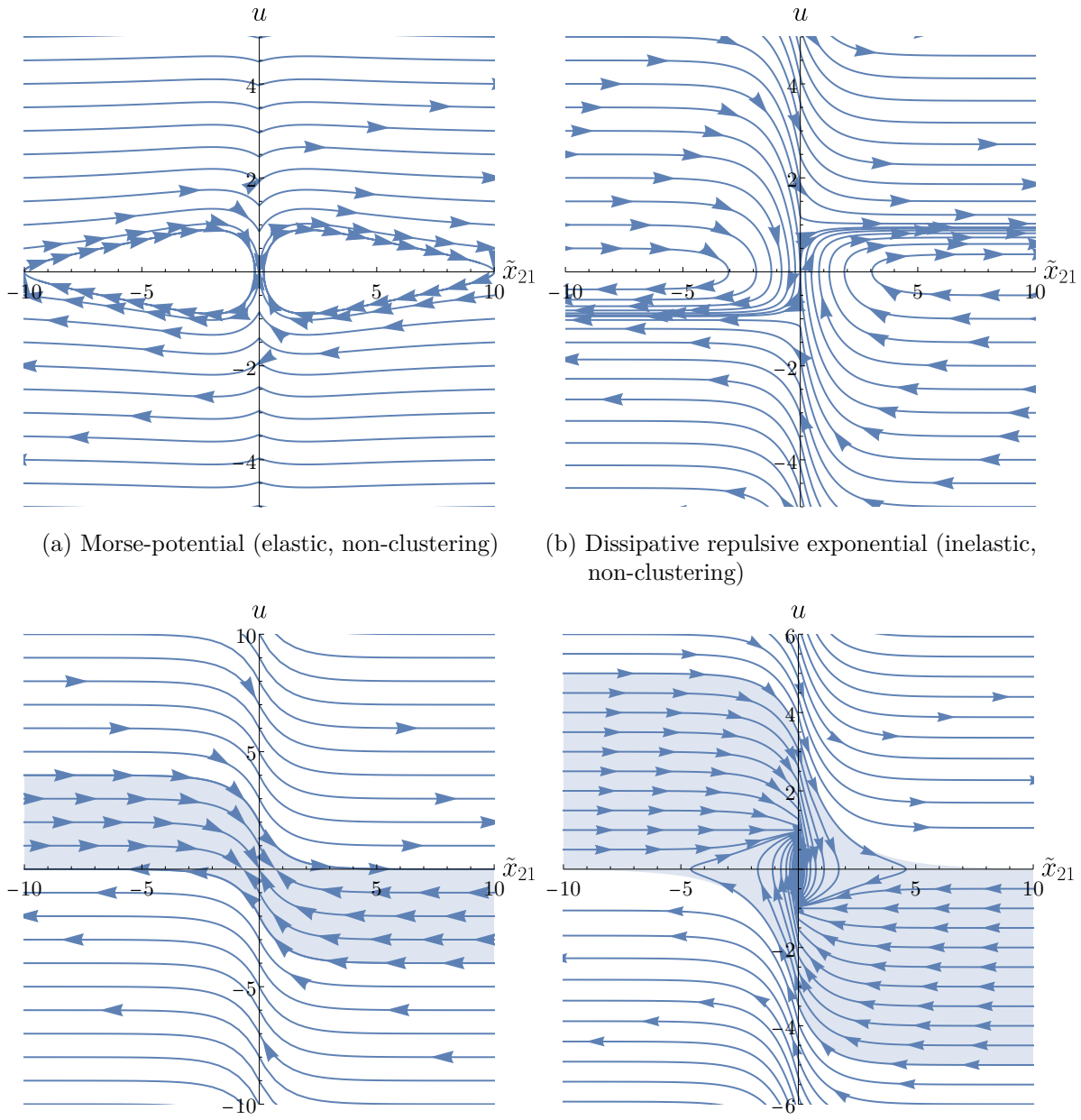
$$\Phi(r) := \lim_{x' \to \infty} 2\Phi(r, x') = \int_{-\infty}^{\infty} \phi\left(\sqrt{r^2 + (x')^2}\right) \mathrm{d}x'. \tag{2.75}$$

Representing (2.74) in the full coordinate system, we have that if $2\Phi(r) < |\boldsymbol{u}^-|$, the relative velocity evolves through the inner region as

$$\boldsymbol{u}^+ = \boldsymbol{u}^- - 2\Phi(r)\frac{\boldsymbol{u}^-}{|\boldsymbol{u}^-|}. \tag{2.76}$$

Note that as the final relative velocity $\boldsymbol{u}^+$ is parallel to the incoming relative velocity $\boldsymbol{u}^-$, the deviation angle θ is always 0, and the representation of the interaction in terms of the impact direction $\boldsymbol{\omega}$ or outgoing relative velocity $\boldsymbol{\sigma}$ is degenerate (see Figure 2.5a). However, if there is an $x < \infty$ for which $2\Phi(r, x) = u^-$, then the $\mathrm{d}x'$ integrand in (2.73) goes to ∞ at a finite x, implying $\tau \to \infty$ as $x' \to x$. This means the particles do not leave the inner region, with the particles asymptotically approaching separation (x, r) with $\boldsymbol{u}^+ \to \boldsymbol{0}$ as $\tau \to \infty$. This implies that the two particles will remain bound with a separation of $O(\varepsilon)$ after the interaction. We will label such a trajectory a *clustering* trajectory.

The analysis in §2.1 tacitly assumes that all interactions result in the two particles separating as $\tau \to \infty$ so that they leave the inner region. If this is not always the case and we construct the model (2.26), we will lose probability mass as not all trajectories have both an incoming and outgoing component. Furthermore, the underlying assumption that three-particle



(a) Morse-potential (elastic, non-clustering)

(b) Dissipative repulsive exponential (inelastic, non-clustering)

(c) Cucker-Smale (inelastic, clustering)

(d) Dissipative attractive exponential (inelastic, clustering)

Figure 2.6.: One-dimensional inner-region trajectory plots for short-ranged interactions: (a) an elastic interaction with the morse potential, $H(x,v) = -\partial_x \phi(|x|)$ with $\phi(|x|) = e^{-|x|} - e^{-|x|/5}$, (b) an inelastic interaction from a repulsive exponential potential with dissipation, $H(x,v) = -\partial_x e^{-|x|} - e^{-|x|}v$, (c) a Cucker-Smale collective behaviour model with exponential interaction kernel, $H(x,v) = -e^{-|x|}v$, and (d) an attractive exponential potential with dissipation, $H(x,v) = \partial_x e^{-|x|} - e^{-|x|}v$. For the clustering interactions, the shaded regions represent trajectories that do not leave the inner region as relative velocity, u, goes to 0 at finite $\tilde{x}_{21}$.

and higher-order interactions are rare will break down for these systems. This occurs as it is possible that multiple particles remain correlated at $O(\varepsilon)$ distance, so if another particle collides, all will be interacting simultaneously. (There is also an intermediate case where particles ultimately leave the inner region but stay for long enough that the interaction time is non-negligible at leading order (for example, see the discussion of [72] in §2.5.1).

Figure 2.6 shows examples of one-dimensional inner-region phase planes for different short-range interactions of types explored in this Chapter. Figure 2.6a shows an elastic potential-force interaction with the Morse potential, $\phi(|x|) = e^{-|x|} - e^{-|x|/5}$. This could be modelled in the framework of elastic interactions as in §2.2, but note that in the one-dimensional case, this model is trivial as particles leave the inner region with the same velocity that they entered. Figure 2.6b shows an inelastic interaction based on a repulsive exponential potential interaction with dissipation term: $H(x, v) = -\partial_x e^{-|x|} - e^{-|x|}v$. Trajectories leave the inner region, but with lower speeds than they entered. Figures 2.6c and 2.6d show clustering interactions, the Cucker-Smale interaction just discussed, $H(x, v) = -e^{-|x|}v$, and an attractive exponential potential with dissipation, $H(x, v) = \partial_x e^{-|x|} - e^{-|x|}v$. The trajectories that do not leave the inner region are shown as shaded regions. For the Cucker-Smale example, we have used an exponential interaction kernel. The impact parameter r is not defined in one dimension, so we define $\Phi := \int_{-\infty}^{\infty} \phi(|x|)\,\mathrm{d}x$, which is similar to $\Phi(0)$ in the d-dimensional case. For the exponential kernel, we have $\Phi = 2$, which means that trajectories with incoming relative speeds less than 4 will cluster. For the dissipative attractive exponential potential, incoming trajectories with

$$|u^-| < 1 + W_0\left(\left(-2 - W_{-1}\left(-e^{-3}\right)\right)e^{-W_{-1}\left(-e^{-3}\right)}\right) \approx 5.03,$$

where W_0 and W_{-1} are branches of the Lambert-W function, oscillate between positive and negative $\tilde{x}_{21}$ and have $\tilde{x}_{21} \to 0$ and $u \to 0$ as $\tau \to \infty$.

As clustering in this form is an important feature of many interactions, we hope to develop an extended framework to capture this behaviour in a kinetic model. We make the assumption that when particles have undergone a clustering interaction and are highly correlated, they effectively behave as a single unit. This would allow us to preserve the binary interaction paradigm central to many of the results of kinetic theory.

2.5.1. Related models

We conclude this section by discussing other models in the literature where groups of particles are treated as a single unit for the purposes of deriving a kinetic or hydrodynamic equation. Shvydkov and Tadmor [101] consider a system of 'multi-flocks', separate groupings with fast intra-flock communication and slower inter-flock communication, for a Cucker-Smale model with long-range interactions. They show how this situation can arise naturally from fluctuations in the initial density, leading to different time scales of flocking within flocks and between flocks. In this process, they derive an individual-based model where the influence of other flocks is reduced to a function only of their averaged quantities, total mass, centre of mass and centre

of mass velocity. They derive a kinetic model for the density distribution of each flock, show alignment at large times and analyse the hydrodynamic model in the limit of large numbers of agents. This model differs from our approach, however, in that they are analysing a system with long-range interactions where the flocks still feel the influence of other multi-flocks. Our model considers clusters of particles bound by short-range interactions that undergo binary interactions with other clusters and can join. We are concerned more with modelling the mass and velocity distributions over time rather than showing the eventual alignment of the full group.

Hammond [60], rigorously derives a Smoluchowski coagulation-diffusion PDE, similar to (1.5), but maintaining spatial variation in the density $n(t; \boldsymbol{x}; m)$ and including a diffusion term, for particles diffusing with Brownian motion that coalesce when they are within distance ε of each other. This is done in the limit of constant mean free path, which for $d \geq 3$ has $N\varepsilon^{d-2} = O(1)$. This differs from the Boltzmann-Grad limit as the volume swept out of radius ε around a Brownian diffusing particle (known as its Wiener sausage) is $O(\varepsilon^{d-2})$ per unit time, while the similar volume traced around a particle in free transport, as in Boltzmann style equations, is $O(\varepsilon^{d-1})$ per unit time.

Kanzler, Schmeiser and Tora [72] consider a short-range kinetic regime, but where the time collisions take is not negligible. They consider one-dimensional, first-order alignment models (one with a stochastic collision time and one where molecules collide until they are fully aligned) and simultaneously track a one-particle density for free particles and a two-particle joint density for particles involved in a collision. Unlike our model, however, the time particles spend in this correlated collisional state is finite, and they all eventually separate. They analyse the long-time behaviour of these models and show that they lead to a Boltzmann-type model in the limit of instantaneous collisions.

3. Derivation of one-dimensional clustering model

3.1. Motivation

As we have seen, some individual-based models with short-range interactions in the form of (2.1) result in particles remaining bound together after interacting, with positions and velocities highly correlated post-interaction. One of the core tenets of our modelling framework here is that the short-range interaction allows us to ignore three-particle and higher-order interactions. As such, we wish to create a model that treats sets of bound particles as a single object called a cluster, with properties defined by extra internal variables. This model would track binary interactions between these clusters, with the outcome of the interaction governed by the clusters' internal properties as well as their position and velocity.

3.1.1. Three-particle Cucker-Smale interactions

As a motivating example, we consider a Cucker-Smale system in two dimensions with $N = 3$ particles and with a short-ranged exponential kernel:

$$\begin{aligned} \dot{\boldsymbol{X}}_i &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{3} \exp\left(-\frac{|\boldsymbol{X}_i - \boldsymbol{X}_j|}{\varepsilon}\right) (\boldsymbol{V}_j - \boldsymbol{V}_i). \end{aligned} \tag{3.1}$$

Specifically, we consider the case where a third particle approaches two bound particles from infinity. If the system has a low particle density with a short-ranged interaction kernel, this scenario is a significantly more likely form of interaction than three particles interacting simultaneously. To simplify the following analysis, we scale the variables by the length scale ε, defining $\tilde{\boldsymbol{X}}_i = \boldsymbol{X}_i/\varepsilon$ and $\tilde{\boldsymbol{V}}_i = \boldsymbol{V}_i/\varepsilon$ to give

$$\begin{aligned} \dot{\tilde{\boldsymbol{X}}}_i &= \tilde{\boldsymbol{V}}_i, \\ \dot{\tilde{\boldsymbol{V}}}_i &= \sum_{\substack{j=1 \\ j \neq i}}^{3} \exp\left(-|\tilde{\boldsymbol{X}}_i - \tilde{\boldsymbol{X}}_j|\right) (\tilde{\boldsymbol{V}}_j - \tilde{\boldsymbol{V}}_i), \end{aligned} \tag{3.2}$$

We will drop the tildes for clarity for the rest of this example.

Given the symmetries in rotation and translation, the initial configuration of this system has four degrees of freedom. Without loss of generality, we can model all interactions of this

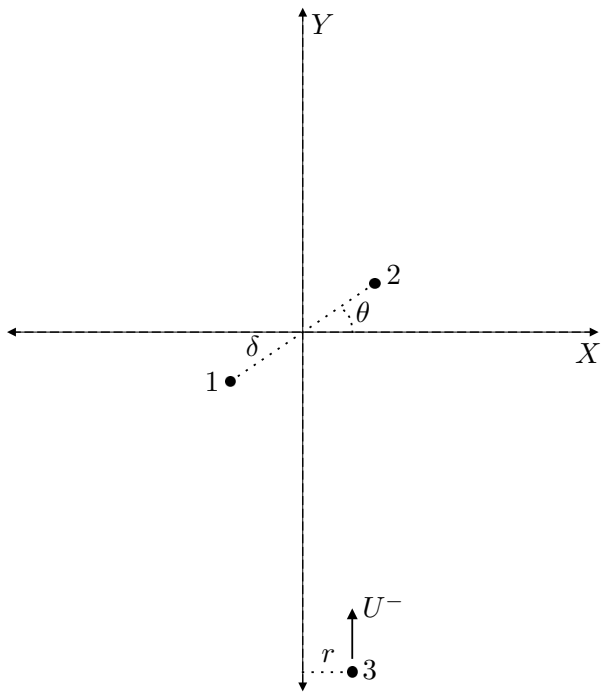


Figure 3.1.: Initial condition for three-particle interaction.

type through the 'initial' condition (as $t \to -\infty$)

$$\begin{aligned} &\boldsymbol{X}_1 \to -\left(\frac{\delta}{2}\cos\theta, \frac{\delta}{2}\sin\theta\right), \quad \boldsymbol{X}_2 \to \left(\frac{\delta}{2}\cos\theta, \frac{\delta}{2}\sin\theta\right), \quad \boldsymbol{X}_3 \to (r, -\infty) \\ &\boldsymbol{V}_1 \to (0,0), \quad \boldsymbol{V}_2 \to (0,0), \quad \boldsymbol{V}_3 \to (0, U^-), \end{aligned} \tag{3.3}$$

with $\delta \geq 0$, $r \geq 0$, $U^- \geq 0$ and $-\pi/2 < \theta \leq \pi/2$. Note that we choose t such that the second component of $\boldsymbol{X}_3$ is 0 when $t = 0$. This setup is shown in Figure 3.1. Defining the components of $\boldsymbol{X}$ to be (X, Y), particles 1 and 2 are bound at a distance of δ around the origin with an angle of θ to the X-axis. Particle 3 approaches from $-\infty$ up the negative Y-axis with offset r and initial velocity U^-. As the X-component of all initial velocities is zero and particles' accelerations are proportional to linear combinations of these velocities, these X-components remain 0 for all t. As such, we can recast the nontrivial behaviour of the system in terms of the Y components:

$$\ddot{Y}_i = \sum_{\substack{j=1 \\ j \neq i}}^{3} \exp\left(-|\boldsymbol{X}_i - \boldsymbol{X}_j|\right)\left(\dot{Y}_j - \dot{Y}_i\right).$$

Defining relative Y-positions $Y_{21} = Y_2 - Y_1$ and $Y_{32} = Y_3 - Y_2$ we have

$$\begin{aligned} \ddot{Y}_{21} &= \left(\phi_{32} - \phi_{31}\right)\dot{Y}_{32} - \left(2\phi_{21} + \phi_{31}\right)\dot{Y}_{21}, \\ \ddot{Y}_{32} &= -\left(2\phi_{32} + \phi_{31}\right)\dot{Y}_{32} + \left(\phi_{21} - \phi_{31}\right)\dot{Y}_{21}, \end{aligned} \tag{3.4}$$

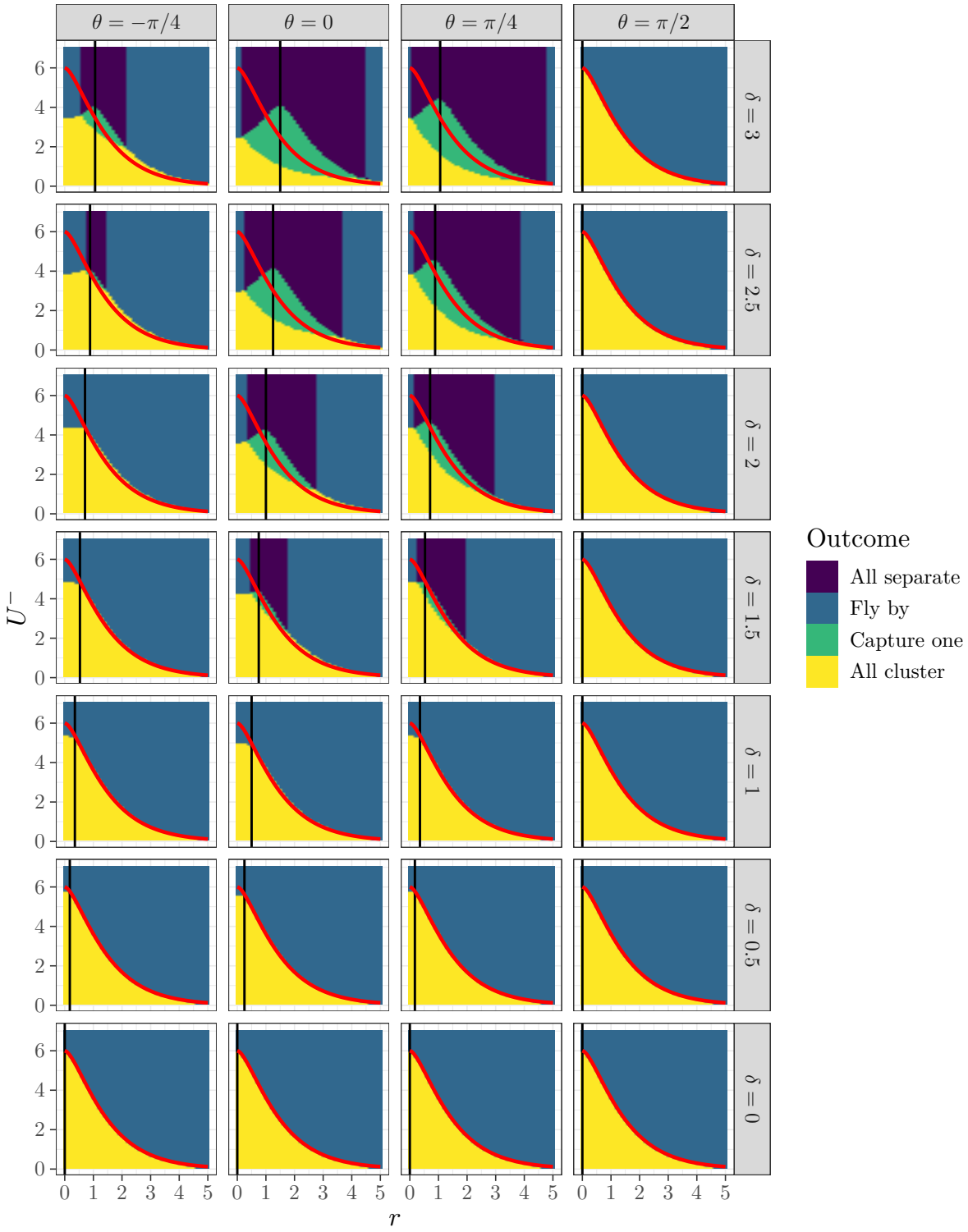


Figure 3.2.: Outcomes of three-particle interactions for the Cucker-Smale interaction with exponential interaction kernel: $\phi(|\boldsymbol{X}|) = e^{-|\boldsymbol{X}|}$. The vertical black lines show where particles 2 and 3 have the same X coordinate (i.e. where $r = \frac{\delta}{2}\cos\theta$ and $X_{32} = 0$). The red line shows the boundary between the 'All cluster' and 'Fly by' outcomes when $\delta = 0$. This is mapped onto outcomes for non-zero values of δ to suggest whether making the point-cluster assumption (always taking $\delta = 0$) might be a reasonable approximation.

where we use the shorthand notation $\phi_{ij} := \exp\left(|\boldsymbol{X}_i - \boldsymbol{X}_j|\right)$, giving specifically

$$\begin{aligned}
\phi_{21} &= \exp\left(-\sqrt{Y_{21}^2 + (\delta\cos\theta)^2}\right), \\
\phi_{32} &= \exp\left(-\sqrt{Y_{32}^2 + \left(r - \frac{\delta}{2}\cos\theta\right)^2}\right), \\
\phi_{31} &= \exp\left(-\sqrt{(Y_{32} + Y_{21})^2 + \left(r + \frac{\delta}{2}\cos\theta\right)^2}\right).
\end{aligned} \tag{3.5}$$

Using the additional conditions $Y_{21}(t) \to \delta\sin\theta$, $\dot{Y}_{21}(t) \to 0$, $Y_{32}(t) \to -\infty$ and $\dot{Y}_{32}(t) \to U^-$ as $t \to -\infty$ we can integrate (3.4) to give

$$\begin{aligned}
\dot{Y}_{21} &= \int_{-\infty}^{Y_{32}} (\phi_{32} - \phi_{31})\,\mathrm{d}Y'_{32} - \int_{\delta\sin\theta}^{Y_{21}} (2\phi_{21} + \phi_{31})\,\mathrm{d}Y'_{21}, \\
\dot{Y}_{32} - V_0 &= -\int_{-\infty}^{Y_{32}} (2\phi_{32} + \phi_{31})\,\mathrm{d}Y'_{32} + \int_{\delta\sin\theta}^{Y_{21}} (\phi_{21} - \phi_{31})\,\mathrm{d}Y'_{21}.
\end{aligned} \tag{3.6}$$

Noting that the initial condition has been chosen such that particle 3 passes closer to particle 2 than particle 1, there are four possible clustering outcomes as $t \to \infty$:

- All separate: Particle 3 attracts particle 2 strongly enough to break apart the cluster between particles 1 and 2, but not strongly enough to cluster with particle 2. There are no clusters as $t \to \infty$ with $Y_{21} \to \infty$ and $Y_{32} \to \infty$.
- Fly by: Particle 3 does not attract either particle strongly enough to break apart or join the original cluster. As $t \to \infty$, particles 1 and 2 remain bound in a cluster without particle 3, i.e. $Y_{32} \to \infty$ but $\dot{Y}_{21} \to 0$.
- Capture one: Particle 3 attracts particle 2 strongly enough to break apart the cluster between particles 1 and 2, forming a new cluster with particle 2. As $t \to \infty$, particles 3 and 2 are bound in a cluster without particle 1, i.e. $Y_{21} \to \infty$ but $\dot{Y}_{32} \to 0$.
- All cluster: Particle 3 is attracted strongly enough to join the cluster with particles 1 and 2. As $t \to \infty$ all three particles are bound in a cluster, i.e. $\dot{Y}_{21} \to 0$ and $\dot{Y}_{32} \to 0$.

Figure 3.2 shows cross-sections of the state space of outcomes for three-particle collisions under the exponential interaction kernel $\phi\left(|\boldsymbol{X}|\right) = e^{-|\boldsymbol{X}|}$, given the four parameters defining the initial condition: r, U^-, δ and θ. The integrals (3.6) do not admit a closed-form analytical expression for this kernel but can be evaluated numerically. The vertical black lines show where particles 2 and 3 have the same X coordinate (i.e. where $r = \frac{\delta}{2}\cos\theta$ and $X_{32} = 0$).

In the full system, we will consider treating clusters as point objects, which we call the *point-cluster* assumption. It can be seen in Figure 3.2 that the 'All cluster' and 'Fly by' dominate the space of outcomes, particularly when $\delta < 1$. The case where $\delta = 0$ is consistent with the point-cluster assumption and the boundary between the 'All cluster' and 'Fly by' regions for this case is mapped onto all the frames as the red line to more clearly compare the

results to those for $\delta > 0$. This appears to be a reasonable approximation to the true outcomes, particularly when the particle separation is around the interaction length scale or closer ($\delta \leq 1$). Considering two-particle interactions (for the one-dimensional case, see Figure 2.6c), we expect a majority of clusters to have separation around this distance.

The point-cluster assumption also appears to be a good approximation for larger values of δ when particles 1 and 2 are initially bound vertically (i.e. $\theta = \pi/2$). When the particles are bound non-vertically, the approximation is more accurate when θ is negative (particle 2 starts below particle 1). In these cases, particle 2 has to be moved through a larger region of particle 1's influence before being separated, which is consistent with there being fewer values of r and U^- for which particle 3 separates particles 1 and 2 (i.e. 'All separate' or 'Capture one').

3.1.2. Derivation framework

With this motivation for the point-cluster assumption, we now turn our attention to developing a kinetic model of clusters as point objects with one internal variable, mass (representing the number of particles in the cluster) in one dimension. We will build the model up in stages. First, in §3.2, we set up a framework where point particles with differing masses interact instantly upon colliding but do not cluster. In §3.3, we allow interactions where two clusters combine to produce a combined cluster with a larger mass. In both cases, the interactions occur strictly point-wise when the clusters collide. Finally, in §3.4, we relax this assumption, using the method of matched asymptotic expansions on a system with short-ranged interactions in the form of (2.1) to derive the kinetic one-particle model as its leading-order solution. Conceptually, this model treats particles that are closely correlated post-interaction as a single entity and those that are weakly correlated as independent. We expect the model to perform well if there is a clear distinction between these cases and few particles in the middle ground where they remain moderately correlated.

3.2. Point-wise interactions in one dimension: non-clustering case

As an initial point-wise model, we consider a system of k 'point-clusters' with variable masses interacting in one dimension. Each cluster is described by a position X_i, velocity V_i and mass M_i with $X_i \in \Omega \subseteq \mathbb{R}$, $V_i \in \mathbb{R}$ and $M_i \in \{1, \ldots, N\}$ for $i \in \{1, \ldots, k\}$. As in Chapter 2, we take periodic boundary conditions on Ω, (so the space is effectively $\mathbb{T}^1$) and assume that the probability of a cluster having velocity V goes to 0 sufficiently quickly as $V \to \infty$. These clusters interact only locally (specifically, when they are at the same point in space), and interactions are assumed to occur instantly at this point. These interactions change the clusters' velocities, which may depend on their masses, but do not change the masses themselves. Outside of these interactions, the clusters move at constant velocity. None of the interactions in this section involve clustering or breakage, so the number of clusters is fixed at k. Although it does not influence the current analysis, we also fix cluster masses to be integers less than or equal to a fixed value N, as a precursor to mass conservation rules being implemented for clustering

interactions later.

We can express the collision process between clusters i and j as

$$(X_i; V_i; M_i) + (X_i; V_j; M_j) \to (X_i; V_i^*; M_i) + (X_i; V_j^*; M_j), \tag{3.7}$$

where the stars are used similarly to (2.23) to denote the post-collisional velocities corresponding to the un-starred pre-collisional velocities. We define the relationship between these incoming and outgoing particle velocities through the function

$$(V_i^*, V_j^*) = \psi_{M_i,M_j}(V_i, V_j), \tag{3.8}$$

which we take to be a bijection from $\mathbb{R}^2$ to $\mathbb{R}^2$ (such that it is invertible). This plays a similar role to the function $\psi_{\boldsymbol{H}}$ defined in (2.21), except now it is dependent on the particle masses and is a directly defined map rather than being derived from an interaction $\boldsymbol{H}$. Again, we use pre-starred velocities to represent pre-collisional velocities corresponding to unstarred post-collisional velocities, and these are defined by the inverse map

$$({}^*V_i, {}^*V_j) = \psi^{-1}_{M_i,M_j}(V_i, V_j). \tag{3.9}$$

We define the full joint probability density for this system (the analogue of the p^N considered in Chapter 2) as

$$p^k(t; x_1, \dots, x_k; v_1, \dots, v_k; m_1, \dots, m_k) \equiv p^k(t; z_1, \dots, z_k),$$

where, for this chapter, we have defined z_i to include the particle mass with $z_i := (x_i; v_i; m_i) \in \boldsymbol{D} = \Omega \times \mathbb{R} \times \{1, \dots, N\}$. As for the non-clustering densities in §2.1, this density is assumed to be invariant under permutations of the particle indices. Note that, in contrast to §2.1, the marginal probability densities now involve sums over possible masses. Defining the integral operator

$$\int_{\boldsymbol{D}} \mathrm{d}z_i := \sum_{m_i=1}^{N} \int_{\mathbb{R}} \int_{\Omega} \mathrm{d}x_i \, \mathrm{d}v_i, \tag{3.10}$$

we can write the s-particle marginal as

$$\begin{aligned} p_s^k(t; z_1, \dots, z_s) &= \sum_{m_{s+1}=1}^{N} \cdots \sum_{m_k=1}^{N} \int_{\mathbb{R}^{(k-s)}} \int_{\Omega^{(k-s)}} p^k(t; z_1, \dots, z_k) \, \mathrm{d}x_{s+1} \dots \mathrm{d}x_k \, \mathrm{d}v_{s+1} \dots \mathrm{d}v_k \\ &= \int_{\boldsymbol{D}^{(k-s)}} p^k(t; z_1, \dots, z_k) \, \mathrm{d}z_{s+1} \dots \mathrm{d}z_k. \end{aligned} \tag{3.11}$$

Integrating over all particles, the total probability is normalised to 1:

$$\int_{\boldsymbol{D}^k} p^k(t; z_1, \dots, z_k) \, \mathrm{d}z_1 \dots \mathrm{d}z_k = 1.$$

Where possible, we will use phase notation (z_i) for conciseness and will use the shorthand

$$^*z_i := (x_i; {}^*v_i; m_i), \tag{3.12}$$

for when the velocity variable is starred. Integrals over $^*z_i$ are hence defined as

$$\int_{\boldsymbol{D}} \mathrm{d}^* z_i := \sum_{m_i=1}^{N} \int_{\mathbb{R}} \int_{\Omega} \mathrm{d}x_i \, \mathrm{d}^* v_i.$$

Note that the star notation applies similarly to lowercase arguments of the density as it does to uppercase particle velocities.

3.2.1. Interaction planes and marginalisation

The probability density is discontinuous across planes where two particles are at the same location, as this is where the particle interactions cause discontinuous changes in velocity. One side of the discontinuity always represents particles approaching an interaction, while the other represents particles leaving an interaction. We label the incoming and outgoing limits of p^k approaching a discontinuity on the plane with $x_i = x_j$ as

$$^{-}p^k_{(i,j)}(t; x_i, x_i; v_i, v_j; m_i, m_j) = \lim_{\tau \to 0^-} p^k(t+\tau; x_i + \tau v_i, x_i + \tau v_j; v_i, v_j; m_i, m_j) \tag{3.13}$$

and

$$^{+}p^k_{(i,j)}(t; x_i, x_i; v_i, v_j; m_i, m_j) = \lim_{\tau \to 0^+} p^k(t+\tau; x_i + \tau v_i, x_i + \tau v_j; v_i, v_j; m_i, m_j), \tag{3.14}$$

respectively (only the i^{th} and j^{th} arguments are shown for conciseness). Note that these do not represent incoming and outgoing probabilities for the same interaction but rather two different processes with velocities v_i and v_j, one beginning and one ending. The notation is intentionally analogous to the + and − superscripts defined for the inner region maps in (2.28). For the analysis that follows, $^{-}p^k_{(i,j)}$ and $^{+}p^k_{(i,j)}$ are equal to p^k away from any interaction boundary and take their respective limit whenever $x_i = x_j$.

Considering the conservation of probability in an infinitesimal box in $\vec{x} \times \vec{v} \times \vec{m}$ space, we obtain the time evolution equation for p^k:

$$\begin{aligned} \partial_t p^k + \sum_{i=1}^{k} \partial_{x_i}(v_i p^k) = & -\frac{1}{2} \sum_{\substack{i,\,j=1 \\ j \neq i}}^{k} |v_i - v_j| \, {}^{-}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij} \\ & + \frac{1}{2} \sum_{\substack{i,\,j=1 \\ j \neq i}}^{k} |v_i - v_j| \, {}^{+}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij}, \end{aligned} \tag{3.15}$$

where we introduce the shorthand $\delta_{ij} \equiv \delta(x_i - x_j)$. The second term on the LHS represents the free transport of clusters, and the terms on the RHS represent sink and source terms at an interaction boundary where $x_i = x_j$. The factor of $1/2$ is present to avoid the double-counting

of interactions. Integrating (3.15) over arguments for particles 2 to k we get

$$\begin{aligned}\partial_t p_1^k + v_1 \partial_{x_1} p_1^k = & -(k-1)\int_{\boldsymbol{D}} |v_1 - v_2|^{-} p_2^k(t; z_1, z_2)\delta_{12}\,\mathrm{d}z_2 \\ & - \frac{(k-1)(k-2)}{2}\int_{\boldsymbol{D}^2} |v_2 - v_3|^{-} p_3^k(t; z_1, z_2, z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3 \\ & + (k-1)\int_{\boldsymbol{D}} |v_1 - v_2|^{+} p_2^k(t; z_1, z_2)\delta_{12}\,\mathrm{d}z_2 \\ & + \frac{(k-1)(k-2)}{2}\int_{\boldsymbol{D}^2} |v_2 - v_3|^{+} p_3^k(t; z_1, z_2, z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3,\end{aligned} \tag{3.16}$$

where the single integral terms represent interactions involving particle 1, and the double integral terms represent all other interactions. The combinatorial factors arise from relabelling particles such that the interactions are between particles 1 and 2 (if particle 1 is involved) or between particles 2 and 3 (if it is not).

3.2.2. Flux conservation over interactions

The probability flux must be conserved over all collision boundaries to ensure total probability is conserved. Consider a region T in $v_i \times v_j$ space on the boundary where $x_i = x_j$. The total probability flux crossing this region due to the interaction between particles i and j is

$$\int_T |{}^*v_i - {}^*v_j|^{-} p_{(i,j)}^k(t; x_i, x_i; {}^*v_i, {}^*v_j; m_i, m_j)\,\mathrm{d}^*v_i\mathrm{d}^*v_j, \tag{3.17}$$

which must matched the total flux emerging from the region $\psi_{m_i,m_j}(T)$:

$$\int_{\psi_{m_i,m_j}(T)} |v_i - v_j|^{+} p_{(i,j)}^k(t; x_i, x_i; v_i, v_j; m_i, m_j)\,\mathrm{d}v_i\mathrm{d}v_j. \tag{3.18}$$

Changing the coordinates of the incoming integral (3.17) to the corresponding outgoing variables (using the functional relationship $({}^*v_i, {}^*v_j) = \psi_{m_i,m_j}^{-1}(v_i, v_j)$ from (3.9) with shorthand $\psi_{m_i,m_j}^{-1}(v_i, v_j) \equiv \psi_{ij}^{-1}$), it becomes

$$\int_{\psi_{m_i,m_j}(T)} |{}^*v_i - {}^*v_j|^{-} p_{(i,j)}^k(t; x_i, x_i; {}^*v_i, {}^*v_j; m_i, m_j)\left|D\psi_{ij}^{-1}\right|\,\mathrm{d}v_i\mathrm{d}v_j, \tag{3.19}$$

where $|D\psi_{ij}^{-1}|$ is the Jacobian determinant of ψ_{ij}^{-1}. As the integrals (3.18) and (3.19) are over the same region, we can combine them:

$$\begin{aligned}\int_{\psi_{m_i,m_j}(T)} \Big[& |{}^*v_i - {}^*v_j|^{-} p_{(i,j)}^k(t; x_i, x_i; {}^*v_i, {}^*v_j; m_i, m_j)\left|D\psi_{ij}^{-1}\right| \\ & - |v_i - v_j|^{+} p_{(i,j)}^k(t; x_i, x_i; v_i, v_j; m_i, m_j)\Big]\,\mathrm{d}v_i\mathrm{d}v_j = 0,\end{aligned}$$

and as the region $\psi_{m_i,m_j}(T)$ is arbitrary, it follows that the integrand must be zero, giving the flux conservation condition

$$|{}^*v_i - {}^*v_j| \; {}^-p^k_{(i,j)}(t; x_i, x_i; {}^*v_i, {}^*v_j; m_i, m_j) \left|D\psi_{ij}^{-1}\right| = |v_i - v_j| \; {}^+p^k_{(i,j)}(t; x_i, x_i; v_i, v_j; m_i, m_j). \tag{3.20}$$

3.2.3. Kinetic model for one-cluster density

Using the flux conservation condition, (3.20), the integral in the final term on the RHS of (3.16) becomes

$$\begin{aligned}\int_{\boldsymbol{D}^2} |v_2 - v_3| \; {}^+p^k_3(t; z_1, z_2, z_3)\delta_{23} \, \mathrm{d}z_2 \mathrm{d}z_3 \\ = \int_{\boldsymbol{D}^2} |{}^*v_2 - {}^*v_3| \; {}^-p^k_3(t; z_1, {}^*z_2, {}^*z_3)\delta_{23} \left|D\psi_{23}^{-1}\right| \, \mathrm{d}z_2 \mathrm{d}z_3 \\ = \int_{\boldsymbol{D}^2} |{}^*v_2 - {}^*v_3| \; {}^-p^k_3(t; z_1, {}^*z_2, {}^*z_3)\delta_{23} \, \mathrm{d}^*z_2 \mathrm{d}^*z_3.\end{aligned}$$

Upon changing dummy integration variables, this coincides with the integral from the second term on the RHS of (3.20), allowing the respective terms to cancel. This cancellation makes sense as these terms represent interactions that do not involve particle 1 and, as such, do not affect p^k_1 (they only lead to changes in velocity for other particles which are integrated over). Thus, (3.16) becomes

$$\begin{aligned}\partial_t p^k_1 + v_1 \partial_{x_1} p^k_1 &= (k-1) \int_{\boldsymbol{D}} \left[|v_1 - v_2| \; {}^+p^k_2(t; z_1, z_2) - |v_1 - v_2| \; {}^-p^k_2(t; z_1, z_2) \right] \delta_{12} \, \mathrm{d}z_2 \\ &= (k-1) \int_{\boldsymbol{D}} \left[|{}^*v_1 - {}^*v_2|^- \, p^k_2(t; {}^*z_1, {}^*z_2) \left|D\psi_{12}^{-1}\right| - |v_1 - v_2| \; {}^-p^k_2(t; z_1, z_2) \right] \delta_{12} \, \mathrm{d}z_2,\end{aligned} \tag{3.21}$$

where, in the second step, we have again used the probability flux conservation condition, (3.20). To close this equation in p^k_1, we again take Boltzmann's *Stosszahlansatz*, assuming particles approaching an interaction are uncorrelated. That is, we take

$${}^-p^k_2(t; z_1, z_2) = p^k_1(t; z_1) p^k_1(t; z_2),$$

which, upon substituting into (3.21) and writing the full arguments explicitly gives

$$\begin{aligned}\partial_t p^k_1 + v_1 \partial_{x_1} p^k_1 = (k-1) \sum_{m_2=1}^{N} \int_{\mathbb{R}} \Big[|{}^*v_1 - {}^*v_2| \; p^k_1(t; x_1; {}^*v_1; m_1) p^k_1(t; x_1; {}^*v_2; m_2) \left|D\psi_{12}^{-1}\right| \\ - |v_1 - v_2| \, p^k_1(t; x_1; v_1; m_1) p^k_1(t; x_1; v_2; m_2) \Big] \, \mathrm{d}v_2.\end{aligned} \tag{3.22}$$

This is a closed, time-evolution equation for the one-cluster probability density.

3.3. Point-wise interactions in one dimension: clustering case

Consider now a system where some point-wise interactions result in particles coalescing together and becoming a point-cluster with higher mass thereafter. By defining an interaction law for clusters of different masses, we can determine what happens when pairs of these clusters interact. From here, we will use 'particle' to refer to an individual particle with mass $M = 1$ and 'cluster' to refer to a point object with any mass. We consider a system of N particles in one dimension with the domains and boundary conditions of positions, velocities and masses as in §3.2. We assume conservation of mass in the clustering process, which gives the additional restriction $\sum_{i=1}^{k} M_i = N$, where k is the number of clusters in the system. Figure 3.3 shows a schematic depiction of a system with three particles where a non-clustering interaction occurs, followed by a clustering interaction where two particles are replaced by a cluster of mass 2.

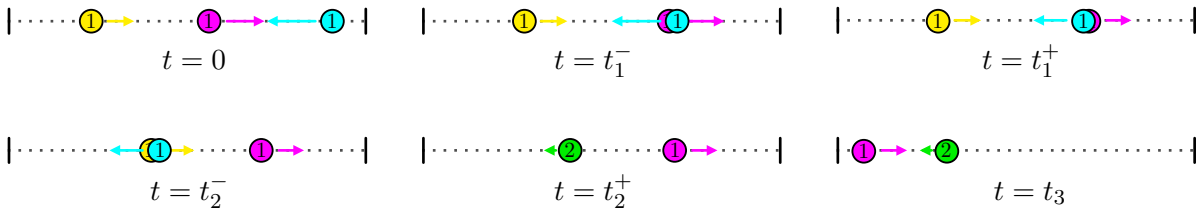


Figure 3.3.: Schematic depiction of a clustering model of $N = 3$ particles and pointwise interactions in one dimension. A non-clustering interaction occurs at t_1 with t_1^- and t_1^+ representing limiting times just before and afterwards, respectively. A clustering interaction occurs at t_2 where two particles (by definition with mass 1) become a cluster with mass 2.

We assume that for each pair of colliding clusters, we have a region $C_{M_i,M_j} \subseteq \mathbb{R}^2$ such that interactions with $(V_i, V_j) \in C_{M_i,M_j}$ result in a combined cluster, while others do not. This gives the interaction law

$$(X_i; V_i; M_i) + (X_i; V_j; M_j) \to \begin{cases} (X_i; V_i'; M_i') & \text{if } (V_i, V_j) \in C_{M_i,M_j} \\ (X_i; V_i^*; M_i) + (X_i; V_j^*; M_j) & \text{if } (V_i, V_j) \notin C_{M_i,M_j}, \end{cases}$$

where dashes represent post-collisional velocities and masses from a clustering interaction and stars the equivalent from a non-clustering interaction.

Similarly to the non-clustering case in §3.2, we define the functional relationship between incoming and outgoing velocities for non-clustering interactions to be

$$(V_i^*, V_j^*) = \psi_{M_i,M_j}(V_i, V_j), \text{ for } (V_i, V_j) \notin C_{M_i,M_j}.$$

This is again assumed to be a bijection from $\mathbb{R}^2 \backslash C_{M_i,M_j}$ to $\mathbb{R}^2$ such that it is invertible. For these interactions, the masses are unchanged. Also as in §3.2, the backward velocity map is defined using ψ^{-1} as

$$({}^*V_i, {}^*V_j) = \psi^{-1}_{M_i,M_j}(V_i, V_j), \text{ for } (V_i, V_j) \in \mathbb{R}^2,$$

where the pre-stars indicate these velocities are the pre-collisional ones that would produce V_i

and V_j after interacting.

For clustering interactions, assuming conservation of mass, it follows that the mass of a cluster produced after an interaction between clusters of mass M_i and M_j is $M_i' = M_i + M_j$. We define the map between incoming and outgoing velocities for such a clustering interaction as

$$V_i' = \lambda_{M_i,M_j}(V_i, V_j) \text{ for } (V_i, V_j) \in C_{M_i,M_j}. \tag{3.23}$$

Throughout this thesis, we will use dashes to represent velocity maps for clustering interactions similarly to how stars are used for non-clustering interactions. We do not require the function λ in (3.23) to be a bijection, as in many common cases, there are a range of incoming cluster velocities and masses that would produce equivalent outgoing clusters. However, we require it to be possible to define a bijection between one incoming velocity, V_i, and the outgoing velocity, V_i', if the other incoming velocity, V_j, is given. This means that if we are given two of the three velocities (out of two incoming and one outgoing), the third will always be well-defined. To formalise this, define the partially applied function

$$\hat{\lambda}_{M_i,M_j,V_j}(V_i) := \lambda_{M_i,M_j}(V_i, V_j) = V_i', \tag{3.24}$$

which we require to have a well-defined inverse: $\hat{\lambda}^{-1}_{M_i,M_j,V_j}(V_i') = V_i$. Acting with this inverse to define $(X_i, {}'V_i, {}'M_i)$, the pre-collisional velocity and mass of a cluster that would combine with a cluster at (X_i, V_j, M_j) to produce one at (X_i, V_i, M_i), we have ${}'M_i = M_i - M_j$ and

$${}'V_i = \hat{\lambda}^{-1}_{M_i - M_j, M_j, V_j}(V_i).$$

Given a system of N particles initially, define $K(t)$ to be a random variable representing the number of clusters in the system at time t and define the probability density

$$p^k(t; x_1, \ldots, x_k; v_1, \ldots, v_k; m_1, \ldots, m_k) \equiv p^k(t; z_1, \ldots, z_k)$$

to represent the joint probability that the N particles have formed k clusters with positions, velocities and masses given by the relevant arguments of p at time t. We now require a hierarchy of these p^k densities from $k = 1$ to $k = N$ to describe the full system, as the number of arguments depends on the number of clusters formed. This is a different approach to that taken by other models of clustering systems in the literature. For example, Hammond [60] keeps the system dimension the same when modelling a system of coagulating Brownian particles by sending all but one of the bound particles in a cluster to a 'cemetery state'.

Marginal s-cluster densities can again be computed as in (3.11). By marginalising over all clusters, we find the probability P_k that the system has k clusters is given by

$$P_k(t) = p_0^k(t) = \int_{\boldsymbol{D}^k} p^k(t; z_1, \ldots, z_k)\, \mathrm{d}z_1 \ldots \mathrm{d}z_k. \tag{3.25}$$

Note that this is normalised such that we have

$$\sum_{k=1}^{N} \int_{\boldsymbol{D}^k} p^k(t; z_1, \ldots, z_k)\, \mathrm{d}z_1 \ldots \mathrm{d}z_k = \sum_{k=1}^{N} P_k(t) = 1.$$

Similarly to (3.15) in the non-clustering case, considering the change of probability in an infinitesimal box, we obtain the evolution equation for p^k:

$$\begin{aligned} \partial_t p^k + \sum_{i=1}^{k} \partial_{x_i}(v_i p^k) = & -\frac{1}{2} \sum_{\substack{i,\, j=1 \\ j\neq i}}^{k} |v_i - v_j|\, {}^{-}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij} \\ & + \frac{1}{2} \sum_{\substack{i,\, j=1 \\ j\neq i}}^{k} |v_i - v_j|\, {}^{+}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij} \\ & + \sum_{i=1}^{k} q^k_{(i)}(t; z_i), \end{aligned} \tag{3.26}$$

where the pre- and post-interaction limits across an interaction surface are again labelled ${}^{-}p^k_{(i,j)}$ and ${}^{+}p^k_{(i,j)}$, respectively as defined in (3.13) and (3.14). For conciseness, we have omitted all arguments other than the two interacting particles. The final terms, $q^k_{(i)}$, are source terms in the p^k equation representing clustering interactions in the p^{k+1} equation (thus reducing the number of clusters from $k+1$ to k). Specifically, $q^k_{(i)}(t; z_i)$ represents the probability flux of the i^{th} particle being produced with coordinates z_i at time t due to such a clustering interaction. They are direct source terms as they represent probability being transferred from a different space rather than a result of probability flux through the infinitesimal box boundary. Integrating the sum of $q^k_{(i)}$ terms over all particles represents the total probability gain in the p^k equation due to clustering. This will be matched by the total probability lost due to clustering in the p^{k+1} equation, which is captured by the difference between the incoming, ${}^{-}p^{k+1}$, and outgoing, ${}^{+}p^{k+1}$, interaction terms after integrating over all particles.

3.3.1. Conservation of clustering flux

We now determine the total outgoing flux from p^{k+1} space that corresponds to the clustering flux source term $q^k_{(i)}(t; x_i; v_i; m_i)$ in p^k space. As we are dealing with probability densities in continuous velocity space, we must consider the incoming flux through a region U in v_i space (see Figure 3.4), i.e.

$$\int_U q^k_{(i)}(t; x_i; v_i; m_i)\, \mathrm{d}v_i.$$

Consider a representative clustering interaction

$$(X_i; V_j; M_j) + (X_i; V_s; M_s) \to (X_i; V_i; M_i), \tag{3.27}$$

which takes the number of clusters in the system from $k+1$ to k. From the definition of our system, we have that $M_i = M_j + M_s$ and $V_i = \lambda_{M_j, M_s}(V_j, V_s)$. Considering now the full-system probability density, the total outgoing flux from interactions of the form (3.27) that would

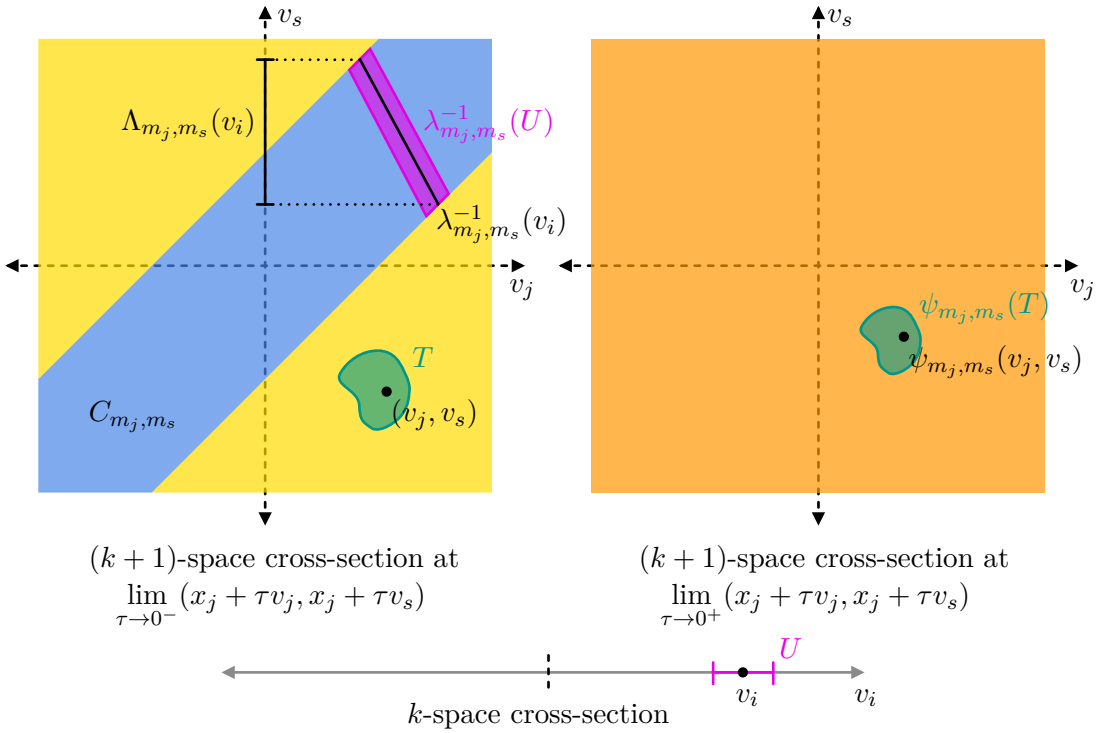


Figure 3.4.: Schematic for the velocity phase space for particles j and s approaching and leaving an interaction discontinuity (i.e. $x_j = x_s$) with $k+1$ clusters in the system. The particles approaching the interaction are represented by the cross-section at $\lim_{\tau\to 0^-}(x_j+\tau v_j, x_j+\tau v_s)$, while those leaving are represented by the other side of the discontinuity, defined by the limit $\lim_{\tau\to 0^+}$. The blue region, C_{m_j,m_s}, represents velocities that would result in a clustering interaction, the yellow region, $R^2 \backslash C_{m_j,m_s}$, represents velocities that would result in a non-clustering interaction, and the orange region shows velocities leaving an interaction. Probability flux entering the discontinuity through T and exiting through its image $\psi_{m_j,m_s}(T)$ must be conserved under a non-clustering interaction. The region $\lambda^{-1}_{m_j,m_s}(U)$ represents the velocity pairs (v_j, v_s) that would result in a cluster being formed with velocity $v_i \in U$ under the map λ_{m_j,m_s}. The black line inside $\lambda^{-1}_{m_j,m_s}(U)$, is, $\lambda^{-1}_{m_j,m_s}(v_i)$, the velocity pairs that would result in a cluster being formed with velocity exactly v_i. The one-dimensional region $\Lambda_{m_j,m_s}(v_i)$ is the projection of $\lambda^{-1}_{m_j,m_s}(v_i)$ onto the v_s axis, i.e. the range of possible values for v_s that could produce a cluster at v_i.

result in incoming flux at $(x_i; v_i; m_i)$ with $v_i \in U$ is

$$\sum_{m_s=1}^{m_i-1} \int_{\lambda^{-1}_{m_j,m_s}(U)} |v_j - v_s| \, {}^{-}p^{k+1}_{(j,s)}(t; x_i, x_i; v_j, v_s; m_i - m_s, m_s) \, \mathrm{d}v_j \mathrm{d}v_s, \tag{3.28}$$

where, by an abuse of notation (as λ is a 'many-to-one' map, so λ^{-1} is 'one-to-many'), we define $\lambda^{-1}_{m_j,m_s}(U)$ as the range of velocity pairs (v_j, v_s) that would map to $v_i \in U$ under λ_{m_j,m_s}. It can be shown (see §A.1 in the appendix) that

$$\int_{\lambda^{-1}_{m_j,m_s}(U)} g(v_j, v_s) \, \mathrm{d}v_j \mathrm{d}v_s = \int_U \int_{\Lambda_{m_j,m_s}(v_i)} g(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i), v_s) \frac{\mathrm{d}}{\mathrm{d}v_i} \left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i) \right) \mathrm{d}v_s \mathrm{d}v_i \tag{3.29}$$

where $g(v_j, v_s)$ is a generic function, $\hat{\lambda}^{-1}$ is the partial inverse of λ (this time well defined, see (3.24)), and $\Lambda_{m_j,m_s}(v_i)$ is the range of possible incoming cluster velocities, v_s, that could produce a cluster with velocity v_i. Note that, by a similar abuse of notation, $\Lambda_{m_j,m_s}(v_i)$ is the projection of $\lambda^{-1}_{m_j,m_s}(v_i)$ onto the v_s axis (see Figure 3.4). Applying (3.29) to (3.28) we get

$$\mathcal{C}_{(j,s)}(U) := \sum_{m_s=1}^{m_i-1} \int_U \int_{\Lambda_{'m_i,m_s}(v_i)} |'v_i - v_s| \, {}^{-}p^{k+1}_{(j,s)}(t; x_i, x_i; 'v_i, v_s; 'm_i, m_s) \frac{\mathrm{d}'v_i}{\mathrm{d}v_i} \, \mathrm{d}v_s \mathrm{d}v_i, \tag{3.30}$$

where we have used the shorthand $'m_i = m_i - m_s = m_j$ and $'v_i = \hat{\lambda}^{-1}_{'m_i,m_s,v_s}(v_i)$, and labeled this integral as $\mathcal{C}_{(j,s)}(U)$: the total clustering flux from particles j and s that would produce a cluster at $(x_i; v_i; m_i)$ with $v_i \in U$.

In order to match this with the incoming flux, $\int_U q^k_{(i)}(t; x_i; v_i; m_i) \, \mathrm{d}v_i$, we must consider how the particle indices in p^{k+1} map to those in p^k in a clustering interaction. One of the fundamental system properties we have assumed is particle indistinguishability or probability invariance under relabelling. That is, $p^k(t; z_1, \ldots, z_k) = p^k(t; \pi(z_1, \ldots, z_k))$ where π is a permutation of the indices. This property is trivially maintained under non-clustering interactions. Consider a permutation of the indices in p^k on the left-hand side of (3.26). The source, ${}^{+}p^k_{(i,j)}$, and sink, ${}^{-}p^k_{(i,j)}$, terms relating to non-clustering interactions are invariant under the same permutation, and so the time evolution of the permuted p^k is equivalent to the original. However, as a clustering interaction reduces the number of particles in the system, and hence the number of coordinates, it does not trivially follow that the relabelling symmetry will be maintained. For each interaction, the $k+1$ incoming coordinates must be mapped onto the k outgoing coordinates in such a way as to preserve this symmetry. Given an interaction between particles j and s, one way to achieve this is to map the product particle to the index j half of the time and to index s half of the time. For example, a clustering interaction between particles 1 and 3

in a four-cluster system would be mapped to the three-cluster system as

$$\begin{aligned}(x_1,x_2,x_1,x_4;v_1,v_2,v_3,v_4;m_1,m_2,m_3,m_4)\to&\\ \frac{1}{2}(x_1,x_2,x_4;\lambda_{m_1,m_3}(v_1,v_3),v_2,v_4;m_1+m_3,m_2,m_4)&\\ +\frac{1}{2}(x_2,x_1,x_4;v_2,\lambda_{m_1,m_3}(v_1,v_3),v_4;m_3,m_1+m_3,m_4).&\end{aligned}$$

Under this mapping, we have

$$\int_U q^k_{(i)}(t;x_i;v_i;m_i)\,\mathrm{d}v_i=\frac{1}{2}\sum_{s=1}^{i}\mathcal{C}_{(i+1,s)}(U)+\frac{1}{2}\sum_{s=i+1}^{k+1}\mathcal{C}_{(i,s)}(U) \tag{3.31}$$

The first term represents interactions between particles $i+1$ and s with $s<i+1$. Half of the flux is mapped to index $i+1$ with index s removed. After removing index s, the original index $i+1$ becomes the new index i. The second term represents interactions between particles i and s with $s>i$. Half of the flux is mapped to index i, which remains as index i with index s removed as $s>i$. Due to invariance under particle relabelling, we can map the interaction indices to i and $k+1$. Noting that there are $k+1$ pairs of indices in total, (3.31) becomes

$$\int_U q^k_{(i)}(t;x_i;v_i;m_i)\,\mathrm{d}v_i=\frac{k+1}{2}\mathcal{C}_{(i,k+1)}(U). \tag{3.32}$$

Combining (3.32) with the definition of $\mathcal{C}_{(i,k+1)}(U)$, (3.30), we get

$$\begin{aligned}\int_U\Bigg[\frac{k+1}{2}\sum_{m_{k+1}=1}^{m_i-1}\int_{\Lambda'_{m_i,m_{k+1}}(v_i)}|'v_i-v_{k+1}|\;{}^{-}p^{k+1}_{(i,k+1)}(t;x_i,x_i;'v_i,v_{k+1};'m_i,m_{k+1})\frac{\mathrm{d}'v_i}{\mathrm{d}v_i}\,\mathrm{d}v_{k+1}&\\ -q^k_{(i)}(t;x_i;v_i;m_i)\Bigg]\mathrm{d}v_i=0.&\end{aligned}$$

Note that this required moving the m_{k+1} sum inside the U integral, which is possible as v_i and m_{k+1} are independent variables. As the interval U is arbitrary, the bracketed integrand must be zero, resulting in the clustering flux conservation condition:

$$\begin{aligned}&q^k_{(i)}(t;x_i;v_i;m_i)\\ &\quad=\frac{k+1}{2}\sum_{m_{k+1}=1}^{m_i-1}\int_{\Lambda'_{m_i,m_{k+1}}(v_i)}|'v_i-v_{k+1}|\;{}^{-}p^{k+1}_{(i,k+1)}(t;x_i,x_i;'v_i,v_{k+1};'m_i,m_{k+1})\frac{\mathrm{d}'v_i}{\mathrm{d}v_i}\,\mathrm{d}v_{k+1}.\end{aligned} \tag{3.33}$$

Substituting the clustering (3.33) and non-clustering (3.20) flux conservation conditions

into (3.26) gives the time evolution equation

$$\begin{aligned}
&\partial_t p^k + \sum_{i=1}^{k} \partial_{x_i}(v_i p^k) \\
&\quad = -\frac{1}{2} \sum_{\substack{i,\,j=1 \\ j\neq i}}^{k} |v_i - v_j| \, {}^{-}p^k_{(i,j)}(t; x_i, x_j; v_i, v_j; m_i, m_j)\delta_{ij} \\
&\qquad + \frac{1}{2} \sum_{\substack{i,\,j=1 \\ j\neq i}}^{k} |{}^{*}v_i - {}^{*}v_j| \, {}^{-}p^k_{(i,j)}(t; x_i, x_j; {}^{*}v_i, {}^{*}v_j; m_i, m_j) \left| D\psi_{ij}^{-1} \right| \delta_{ij} \\
&\qquad + \frac{k+1}{2} \sum_{i=1}^{k} \sum_{m_{k+1}=1}^{m_i - 1} \int_{\Lambda_{{}'m_i, m_{k+1}}(v_i)} |{}'v_i - v_{k+1}| \\
&\qquad\qquad \times {}^{-}p^{k+1}_{(i,k+1)}(t; x_i, x_i; {}'v_i, v_{k+1}; {}'m_i, m_{k+1}) \frac{\mathrm{d}'v_i}{\mathrm{d}v_i} \, \mathrm{d}v_{k+1},
\end{aligned} \tag{3.34}$$

where ${}'m_i = m_i - m_{k+1}$, ${}'v_i = \hat{\lambda}^{-1}_{m_i, m_{k+1}, v_{k+1}}(v_i)$ and $({}^{*}v_i, {}^{*}v_j) = \psi_{ij}^{-1}(v_i, v_j)$. Introducing the new shorthand ${}'z_i \equiv (x_i; {}'v_i; {}'m_i)$ for the pre-phase of a clustering interaction and again using ${}^{*}z_i \equiv (x_i; {}^{*}v_i; m_i)$ for the pre-phase of a non-clustering interaction, (3.35) can be written as

$$\begin{aligned}
\partial_t p^k + \sum_{i=1}^{k} \partial_{x_i}(v_i p^k) = &-\frac{1}{2} \sum_{\substack{i,\,j=1 \\ j\neq i}}^{k} |v_i - v_j| \, {}^{-}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij} \\
&+ \frac{1}{2} \sum_{\substack{i,\,j=1 \\ j\neq i}}^{k} |{}^{*}v_i - {}^{*}v_j| \, {}^{-}p^k_{(i,j)}(t; {}^{*}z_i, {}^{*}z_j) \left| D\psi_{ij}^{-1} \right| \delta_{ij} \\
&+ \frac{k+1}{2} \sum_{i=1}^{k} \int_{\boldsymbol{C}'(z_i)} |{}'v_i - v_{k+1}| \, {}^{-}p^{k+1}_{(i,k+1)}(t; {}'z_i, z_{k+1}) \delta_{i(k+1)} \frac{\mathrm{d}'v_i}{\mathrm{d}v_i} \, \mathrm{d}z_{k+1},
\end{aligned} \tag{3.35}$$

where have also introduced the integration region

$$\int_{\boldsymbol{C}'(z_i)} \mathrm{d}z_{k+1} := \sum_{m_{k+1}=1}^{m_i - 1} \int_{\Lambda_{{}'m_i, m_{k+1}}(v_i)} \int_{\Omega} \mathrm{d}x_{k+1} \, \mathrm{d}v_{k+1}, \tag{3.36}$$

representing the incoming range of z_{k+1} for a clustering interaction resulting in a cluster with phase z_i.

3.3.2. Marginal distributions

The marginal probability distributions are again computed as in (3.11), with the one-particle marginal defined as

$$p^k_1(t; z_1) = \int_{\boldsymbol{D}^{(k-1)}} p^k(t; z_1, \ldots, z_k) \, \mathrm{d}z_2 \ldots \mathrm{d}z_k,$$

with

$$\int_{\boldsymbol{D}} \mathrm{d}z_i := \sum_{m_i=1}^{N} \int_{\mathbb{R}} \int_{\Omega} \mathrm{d}x_i \, \mathrm{d}v_i.$$

We can still sum over all masses from 1 to N with the convention that $p^k(t; z_1, \ldots, z_k) = 0$ whenever $\sum_{i=1}^{k} m_i \neq N$. Taking $k > 3$ and marginalising (3.35) over clusters 2 to k we get

$$\begin{aligned}
&\partial_t p_1^k + v_1 \partial_{x_1} p_1^k \\
&\quad = -(k-1) \int_{\boldsymbol{D}} |v_1 - v_2|^{-} p_2^k(t; z_1, z_2) \delta_{21} \, \mathrm{d}z_2 \\
&\qquad - \frac{(k-1)(k-2)}{2} \int_{\boldsymbol{D}^2} |v_2 - v_3|^{-} p_3^k(t; z_1, z_2, z_3) \delta_{23} \, \mathrm{d}z_2 \mathrm{d}z_3 \\
&\qquad + (k-1) \int_{\boldsymbol{D}} |{}^*v_1 - {}^*v_2|^{-} p_2^k(t; {}^*z_1, {}^*z_2) \delta_{12} \left| D\psi_{12}^{-1} \right| \, \mathrm{d}z_2 \\
&\qquad + \frac{(k-1)(k-2)}{2} \int_{\boldsymbol{D}^2} |{}^*v_2 - {}^*v_3|^{-} p_3^k(t; z_1, {}^*z_2, {}^*z_3) \delta_{23} \left| D\psi_{23}^{-1} \right| \, \mathrm{d}z_2 \mathrm{d}z_3 \\
&\qquad + \frac{k+1}{2} \int_{\boldsymbol{C}'(z_1)} |{}'v_1 - v_{k+1}|^{-} p_2^{k+1}(t; {}'z_1, z_{k+1}) \delta_{1(k+1)} \frac{\mathrm{d}'v_1}{\mathrm{d}v_1} \, \mathrm{d}z_{k+1} \\
&\qquad + \frac{(k+1)(k-1)}{2} \int_{\boldsymbol{D}} \int_{\boldsymbol{C}'(z_2)} |{}'v_2 - v_{k+1}|^{-} p_3^{k+1}(t; z_1, {}'z_2, z_{k+1}) \delta_{2(k+1)} \frac{\mathrm{d}'v_2}{\mathrm{d}v_2} \, \mathrm{d}z_{k+1} \mathrm{d}z_2,
\end{aligned} \tag{3.37}$$

Note that the marginal distributions can only be non-zero if their mass arguments are possible with the number of clusters. For instance for $p_1^k(t; z_1)$ to be non-zero, we require $1 \leq m_1 \leq N - (k-1)$. The $k-1$ coefficient in the first and third term arises from counting the terms involving either $i = 1$ or $j = 1$ in the first two RHS terms of (3.35) and using the labelling indistinguishability to map the other label to 2. The second and fourth terms involve all other interactions from these terms (where neither $i = 1$ nor $j = 1$). Again, the terms are matched using the labelling indistinguishability, and the coefficient $(k-1)(k-2)/2$ captures the number of instances in each case. The coefficient $(k+1)(k-1)/2$ of the sixth term arises from mapping each of the $k-1$ clustering terms (the third RHS term of (3.35)) with $i \neq 1$ to $i = 2$.

With some manipulations on the integral arguments and bounds (see §A.2 in the appendix for the full derivation), (3.37) can be rearranged to

$$\begin{aligned}
\partial_t p_1^k + v_1 \partial_{x_1} p_1^k = &-(k-1) \int_{\boldsymbol{D}} |v_1 - v_2|^{-} p_2^k(t; z_1, z_2) \delta_{21} \, \mathrm{d}z_2 \\
&+ (k-1) \int_{\boldsymbol{D}} |{}^*v_1 - {}^*v_2|^{-} p_2^k(t; {}^*z_1, {}^*z_2) \delta_{12} \left| D\psi_{12}^{-1} \right| \, \mathrm{d}z_2 \\
&+ \frac{k+1}{2} \int_{\boldsymbol{C}'(z_1)} |{}'v_1 - v_2|^{-} p_2^{k+1}(t; {}'z_1, z_2) \delta_{12} \frac{\mathrm{d}'v_1}{\mathrm{d}v_1} \, \mathrm{d}z_2 \\
&- \frac{(k-1)(k-2)}{2} \int_{\boldsymbol{C}^2} |v_2 - v_3|^{-} p_3^k(t; z_1, z_2, z_3) \delta_{23} \, \mathrm{d}z_2 \mathrm{d}z_3 \\
&+ \frac{(k+1)(k-1)}{2} \int_{\boldsymbol{C}^2} |v_2 - v_3|^{-} p_3^{k+1}(t; z_1, z_2, z_3) \delta_{23} \, \mathrm{d}z_2 \mathrm{d}z_3,
\end{aligned} \tag{3.38}$$

where the new integration region $\boldsymbol{C}^2$ is defined as

$$\int_{\boldsymbol{C}^2} \mathrm{d}z_i \, \mathrm{d}z_j := \sum_{m_i,m_j=1}^{N} \int_{C_{m_i,m_j}} \int_{\Omega^2} \mathrm{d}x_i \mathrm{d}x_j \, \mathrm{d}v_i \mathrm{d}v_j. \tag{3.39}$$

This is analogous to the full two-particle domain $\boldsymbol{D}^2$, except that now the velocity integrals are only taken over the clustering region, C_{m_i,m_j}, rather than over the full $\mathbb{R}^2$ (see Figure 3.4).

3.3.3. Mass-restricted *Stosszahlansatz*

In (3.38), we have the time evolution of the one-cluster probability densities p_1^k as a function of two- and three-cluster probability densities for clusters approaching interactions. In order to produce a closed equation in the one-cluster density, we need to express the two and three-cluster densities in terms of p_1^k. From the chain rule of probability, $\mathbb{P}(A = a, B = b) = \mathbb{P}(A = a|B = b)\mathbb{P}(B = b)$, we have that

$$p_1^k(t; z_1) = p_1(t; z_1 \,|\, k) P_k(t), \tag{3.40}$$

where $P_k(t)$ defined in (3.25) is the probability the system has k clusters at time t and $p_1(t; z_1 \,|\, k)$ is the probability the first particle is in the state z_1 at time t, given the system has k clusters. Similarly for ${}^{-}p_2^k$ and ${}^{-}p_3^k$ we have

$${}^{-}p_2^k(t; z_1, z_2) = {}^{-}p_2(t; z_1, z_2 \,|\, k) P_k(t), \text{ and } {}^{-}p_3^k(t; z_1, z_2, z_3) = {}^{-}p_3(t; z_1, z_2, z_3 \,|\, k) P_k(t), \tag{3.41}$$

respectively. Making a similar assumption to Boltzmann's *Stosszahlansatz*, we assume that pairs of clusters approaching a collision are independent so that we can write the incoming conditional probabilities as

$${}^{-}p_2(t; z_1, z_2 \,|\, k)\delta_{12} = p_1(t; z_1 \,|\, k) p_1(t; z_2 \,|\, k)\delta_{12}. \tag{3.42}$$

This assumes that after a pair of clusters interact, they either do not interact again or, if they do, they will have interacted with several other clusters in between, rendering any correlation that developed from the first interaction negligible. Naturally, this assumption requires the number of clusters in the system to be large, so we expect (3.42) to be less accurate for small k. With the three-particle joint densities, the first particle is not involved in the interaction. We assume that this first particle is independent, i.e.

$${}^{-}p_3(t; z_1, z_2, z_3 \,|\, k)\delta_{23} = p_1(t; z_1 \,|\, k) \, {}^{-}p_2(t; z_2, z_3 \,|\, k)\delta_{23}.$$

We can then split the ${}^{-}p_2$ factor using (3.42) again.

However, by splitting these densities into products of single-cluster densities, we can no longer rely on the two- and three-cluster joint densities being zero for unrealisable mass combinations, as the individual densities share no joint mass information. Given a system with N particles forming k clusters, we note that the maximum mass contained in a set of s clusters is

$$\Gamma_s(k) := N - (k - s) = N - k + s. \tag{3.43}$$

This is because there must be at least $k-s$ particles spare to form the remaining $k-s$ clusters. If we already have a cluster with mass m_1 fixed, then the maximum mass contained in a set of s clusters, not including the fixed cluster, is

$$\Gamma_s(k, m_1) := N - m_1 - (k - (s+1)) = N - m_1 - k + s + 1. \tag{3.44}$$

We generalise these to the variable argument function

$$\Gamma_s\left(k, \{m_i\}_{i=1}^I\right) := N - \sum_{i=1}^{I} m_i - k + s + I \tag{3.45}$$

representing the mass available to a set of s clusters given the system has k clusters and masses $\{m_i\}_{i=1}^I$ are already fixed.

With these restrictions in mind, we introduce the *mass-restricted Stosszahlansatz*:

$${}^{-}p_2(t; z_1, z_2 \,|\, k)\delta_{12} = \begin{cases} p_1(t; z_1 \,|\, k)p_1(t; z_2 \,|\, k)\delta_{12}, & m_1 + m_2 \leq \Gamma \\ 0, & \text{otherwise,} \end{cases} \tag{3.46}$$

where Γ, the bound for realisable cluster masses, depends on the context. If there is a cluster with mass m already fixed then $\Gamma = \Gamma_2(k, m)$, but if no masses are already fixed then $\Gamma = \Gamma_2(k)$. By similar arguments, the three-particle case becomes

$${}^{-}p_3(t; z_1, z_2, z_3 \,|\, k)\delta_{23} = \begin{cases} p_1(t; z_1 \,|\, k)\,{}^{-}p_2(t; z_2, z_3 \,|\, k)\delta_{23}, & m_1 + m_2 + m_3 \leq \Gamma_3(k) \\ 0, & \text{otherwise.} \end{cases} \tag{3.47}$$

Note that by choosing m_1, the restriction on the other two masses is $m_2 + m_3 \leq N - m_1 - k + 3 = \Gamma_2(k, m_1)$, which is consistent with realisable values in the two-particle case, so we can use (3.46) to give

$${}^{-}p_3(t; z_1, z_2, z_3 \,|\, k)\delta_{23} = \begin{cases} p_1(t; z_1 \,|\, k)p_1(t; z_2 \,|\, k)p_1(t; z_3 \,|\, k)\delta_{23}, & m_1 + m_2 + m_3 \leq \Gamma_3(k) \\ 0, & \text{otherwise.} \end{cases} \tag{3.48}$$

Combining (3.46) and (3.48) with (3.40) and (3.41), we get

$${}^{-}p_2^k(t; z_1, z_2)\delta_{12} = \begin{cases} \dfrac{1}{P_k(t)} p_1^k(t; z_1)p_1^k(t; z_2)\delta_{12}, & m_1 + m_2 \leq \Gamma \\ 0, & \text{otherwise,} \end{cases} \tag{3.49}$$

with Γ defined as in (3.46) and

$${}^{-}p_3^k(t; z_1, z_2, z_3)\delta_{23} = \begin{cases} \dfrac{1}{P_k(t)^2} p_1^k(t; z_1)p_1^k(t; z_2)p_1^k(t; z_3)\delta_{23}, & m_1 + m_2 + m_3 \leq \Gamma_3(k) \\ 0, & \text{otherwise.} \end{cases} \tag{3.50}$$

We define new integration regions that take into account these restrictions on the particle

masses as

$$\int_{\boldsymbol{D}^s_\Gamma(k)} \mathrm{d}z_1 \dots \mathrm{d}z_s := \sum_{\substack{m_1,\dots m_s=1 \\ \sum_{i=1}^s m_i \leq \Gamma_s(k)}}^{N} \int_{\mathbb{R}^s}\int_{\Omega^s} \mathrm{d}x_1 \dots \mathrm{d}x_s \, \mathrm{d}v_1 \dots \mathrm{d}v_s \tag{3.51}$$

and

$$\int_{\boldsymbol{D}^s_\Gamma(k,m_1)} \mathrm{d}z_2 \dots \mathrm{d}z_{s+1} := \sum_{\substack{m_2,\dots m_{s+1}=1 \\ \sum_{i=2}^{s+1} m_i \leq \Gamma_s(k,m_1)}}^{N} \int_{\mathbb{R}^s}\int_{\Omega^s} \mathrm{d}x_2 \dots \mathrm{d}x_{s+1} \, \mathrm{d}v_2 \dots \mathrm{d}v_{s+1}. \tag{3.52}$$

We also define $\boldsymbol{C}^2_\Gamma(k)$ and $\boldsymbol{C}^2_\Gamma(k,m_1)$ analogously by setting $s=2$ and changing the velocity integral bounds from $\mathbb{R}^2$ to C_{m_1,m_2} and C_{m_2,m_3}, respectively. Note that this is a slight abuse of notation for conciseness as the mass restrictions mean $\boldsymbol{D}^s_\Gamma$ is not truly a cartesian product of s copies of $\boldsymbol{D}_\Gamma$ (and similarly for the other regions).

Substituting the mass-restricted *Stosszahlansatz*, (3.49) and (3.50), into (3.38) and using these newly defined integration regions we get

$$\begin{aligned}
\partial_t p_1^k + v_1 \partial_{x_1} p_1^k = & -\frac{(k-1)}{P_k} \int_{\boldsymbol{D}_\Gamma(k,m_1)} |v_1 - v_2| \, p_1^k(z_1) p_1^k(z_2) \delta_{12} \, \mathrm{d}z_2 \\
& + \frac{(k-1)}{P_k} \int_{\boldsymbol{D}_\Gamma(k,m_1)} |{}^*v_1 - {}^*v_2| \, p_1^k({}^*z_1) p_1^k({}^*z_2) \delta_{12} \left| D\psi_{12}^{-1} \right| \mathrm{d}z_2 \\
& + \frac{k+1}{2P_{k+1}} \int_{\boldsymbol{C}'(z_1)} |{}'v_1 - v_2| \, p_1^{k+1}({}'z_1) p_1^{k+1}(z_2) \delta_{12} \frac{\mathrm{d}'v_1}{\mathrm{d}v_1} \, \mathrm{d}z_2 \\
& - \frac{(k-1)(k-2)}{2(P_k)^2} p_1^k(z_1) \int_{\boldsymbol{C}^2_\Gamma(k,m_1)} |v_2 - v_3| \, p_1^k(z_2) p_1^k(z_3) \delta_{23} \, \mathrm{d}z_2 \mathrm{d}z_3 \\
& + \frac{(k+1)(k-1)}{2(P_{k+1})^2} p_1^{k+1}(z_1) \int_{\boldsymbol{C}^2_\Gamma(k+1,m_1)} |v_2 - v_3| \, p_1^{k+1}(z_2) p_1^{k+1}(z_3) \delta_{23} \, \mathrm{d}z_2 \mathrm{d}z_3,
\end{aligned} \tag{3.53}$$

where we drop the t arguments of p_1^k and P_k for conciseness from here on when writing the full time-evolution equation. Noting that we can compute $P_k(t)$ as a function of $p_1^k(t;z_1)$ through

$$P_k(t) = \int_{\boldsymbol{D}_\Gamma(k)} p_1^k(t;z_1) \, \mathrm{d}z_1, \tag{3.54}$$

(3.53) and (3.54) now form a closed hierarchy of equations for p_1^k, with $k = 1, \dots, N$. This could be used as a PDE model for the system directly; however, as we are considering the number of clusters, k, to be large, it may be infeasible for computation. We now consider further approximations that will allow us to collapse the k hierarchy and produce a simpler system.

3.3.4. *Clusterzahlansatz* and k-hierarchy reduction

We now marginalise the number of clusters, k, in the system to produce a simplified model that will be easier to compute. Define $p_1(t;z_1)$ as the probability of a cluster in state z_1 regardless

of the number of clusters in the system, that is

$$p(t; z_1) := \sum_{k=1}^{N} p_1^k(t; z_1) = \sum_{k=1}^{N} p_1(t; z_1 \,|\, k) P_k.$$

Note that we have dropped the subscript 1 from the k-marginalised distribution p for clarity and consistency with one-particle distributions in the kinetic theory literature. For a particular value of k, m_1 can range from 1 to $N - k + 1$ to ensure enough mass remains in the system to form the other clusters (except if $k = 1$, in which case m_1 must be N). In these sums, we assume that $p(t; z_1 \,|\, k) = 0$ for combinations of k and m_1 that are not within the appropriate bounds. Summing (3.53) over k, simplifying and recasting in terms of conditional probabilities (with the details in §A.3), we get

$$\begin{aligned}
\partial_t p + v_1 \partial_{x_1} p = & -\sum_{k=2}^{N} (k-1) P_k \int_{\boldsymbol{D}_\Gamma(k,m_1)} |v_1 - v_2| \, p_1(z_1 \,|\, k) p_1(z_2 \,|\, k) \delta_{12} \, \mathrm{d}z_2 \\
& + \sum_{k=2}^{N} (k-1) P_k \int_{\boldsymbol{D}_\Gamma(k,m_1)} |{}^*v_1 - {}^*v_2| \, p_1({}^*z_1 \,|\, k) p_1({}^*z_2 \,|\, k) \delta_{12} \left| D\psi_{12}^{-1} \right| \, \mathrm{d}z_2 \\
& + \sum_{k=2}^{N} \frac{k}{2} P_k \int_{\boldsymbol{C}'(z_1)} |{}'v_1 - v_2| \, p_1({}'z_1 \,|\, k) p_1(z_2 \,|\, k) \delta_{12} \frac{\mathrm{d}'v_1}{\mathrm{d}v_1} \, \mathrm{d}z_2 \\
& + \sum_{k=3}^{N} \frac{k-2}{2} P_k \, p_1(z_1 \,|\, k) \int_{\boldsymbol{C}_\Gamma^2(k,m_1)} |v_2 - v_3| \, p_1(z_2 \,|\, k) p_1(z_3 \,|\, k) \delta_{23} \, \mathrm{d}z_2 \mathrm{d}z_3.
\end{aligned} \tag{3.55}$$

Note that the three-particle terms did not cancel as they do in the non-clustering case, remaining as the final term in (3.55). Although this term represents clustering interactions that do not involve cluster 1, it affects $p(t; z_1)$ as these interactions affect the total number of clusters in the system and, thus, indirectly, the probability that a single cluster has properties z_1.

In order to produce a closed system in $p(t; z_1)$, we need to make a further approximation to collapse the k sums in (3.55). Specifically, we assume single-particle probability densities of position, mass, and velocity are independent of the number of clusters in the system. That is

$$p_1(t; z_1 \,|\, k) = p(t; z_1), \tag{3.56}$$

for all k. By analogy to Boltzmann's *Stosszahlansatz*, we call this the *Clusterzahlansatz*, meaning

cluster number independence assumption. Substituting this into (3.55) gives

$$\begin{aligned}\partial_t p + v_1\partial_{x_1} p = &-\sum_{k=2}^{N}(k-1)P_k\int_{\boldsymbol{D}_\Gamma(k,m_1)}|v_1-v_2|\,p(z_1)p(z_2)\delta_{12}\,\mathrm{d}z_2\\ &+\sum_{k=2}^{N}(k-1)P_k\int_{\boldsymbol{D}_\Gamma(k,m_1)}|{}^*v_1-{}^*v_2|\,p({}^*z_1)p({}^*z_2)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_2\\ &+\sum_{k=2}^{N}\frac{k}{2}P_k\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|\,p('z_1)p(z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\\ &+\sum_{k=3}^{N}\frac{k-2}{2}P_k\,p(z_1)\int_{\boldsymbol{C}^2_\Gamma(k,m_1)}|v_2-v_3|\,p(z_2)p(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.\end{aligned}\tag{3.57}$$

What remains to produce a closed system in $p(t;z_1)$ is to compute $P_k(t)$ for $k=2,\dots,N$. However, we cannot compute this directly from $p(t;z_1)$ as we lost the information about the distribution of K when we summed over k and took the *Clusterzahlansatz*. We instead compute the K distribution from the full k-hierarchy equation (3.53).

Cluster number distribution

To compute the cluster number probabilities $P_k(t)$, consider marginalising (3.53) over particle 1, i.e. acting with the integral operator $\int_{\boldsymbol{D}_\Gamma(k)}\mathrm{d}z_1$. The LHS becomes

$$\partial_t\int_{\boldsymbol{D}_\Gamma(k)}p_1(t;z_1\mid k)P_k(t)\,\mathrm{d}z_1+\int_{\boldsymbol{D}_\Gamma(k)}\partial_{x_1}(v_1p_1^k)\,\mathrm{d}z_1=\frac{\mathrm{d}}{\mathrm{d}t}P_k(t),$$

where we used the normalisation of $p_1(t;z_1\,|\,k)$ to remove the first integral and the divergence theorem and boundary conditions to set the second term to zero. The full integrated equation is then

$$\begin{aligned}\frac{\mathrm{d}P_k}{\mathrm{d}t}=&-\frac{(k-1)}{P_k}\int_{\boldsymbol{D}^2_\Gamma(k)}|v_1-v_2|\,p_1^k(z_1)p_1^k(z_2)\delta_{12}\,\mathrm{d}z_1\mathrm{d}z_2\\ &+\frac{(k-1)}{P_k}\int_{\boldsymbol{D}^2_\Gamma(k)}|v_1^*-v_2^*|\,p_1^k(z_1^*)p_1^k(z_2^*)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_1\mathrm{d}z_2\\ &+\frac{k+1}{2P_{k+1}}\int_{\boldsymbol{D}_\Gamma(k)}\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|\,p_1^{k+1}('z_1)p_1^{k+1}(z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\,\mathrm{d}z_1\\ &-\frac{(k-1)(k-2)}{2(P_k)^2}\int_{\boldsymbol{D}_\Gamma(k)}\int_{\boldsymbol{C}^2_\Gamma(k,m_1)}|v_2-v_3|\,p_1^k(z_1)p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\,\mathrm{d}z_1\\ &+\frac{(k+1)(k-1)}{2(P_{k+1})^2}\int_{\boldsymbol{D}_\Gamma(k)}\int_{\boldsymbol{C}^2_\Gamma(k+1,m_1)}|v_2-v_3|\,p_1^{k+1}(z_1)p_1^{k+1}(z_2)p_1^{k+1}(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\,\mathrm{d}z_1.\end{aligned}$$

With some manipulation of the integral variables and bounds (see §A.4 in the appendix for intermediate steps), this can be simplified to

$$\begin{aligned}\frac{\mathrm{d}P_k}{\mathrm{d}t} = &-\frac{k(k-1)}{2}P_k\int_{\boldsymbol{C}^2_\Gamma(k)}|v_1-v_2|\,p_1(z_1\,|\,k)p_1(z_2\,|\,k)\delta_{12}\,\mathrm{d}z_2\mathrm{d}z_1\\ &+\frac{(k+1)k}{2}P_{k+1}\int_{\boldsymbol{C}^2_\Gamma(k+1)}|v_1-v_2|\,p_1(z_1\,|\,k+1)p_1(z_2\,|\,k+1)\delta_{12}\,\mathrm{d}z_1\mathrm{d}z_2.\end{aligned} \tag{3.58}$$

Using the *Clusterzahlansatz* (3.56), (3.58) can be expressed in terms of $p(t;z)$:

$$\begin{aligned}\frac{\mathrm{d}P_k}{\mathrm{d}t} = &-\frac{k(k-1)}{2}P_k\int_{\boldsymbol{C}^2_\Gamma(k)}|v_1-v_2|\,p(z_1)p(z_2)\delta_{12}\,\mathrm{d}z_2\mathrm{d}z_1\\ &+\frac{(k+1)k}{2}P_{k+1}\int_{\boldsymbol{C}^2_\Gamma(k+1)}|v_1-v_2|\,p(z_1)p(z_2)\delta_{12}\,\mathrm{d}z_1\mathrm{d}z_2.\end{aligned} \tag{3.59}$$

This forms a set of differential equations for $P_k(t)$ that depend on $p(t;z)$. To form a closed system, they can be integrated concurrently with (3.57). However, (3.57) still involves sums over k for each term as the integral bounds depend on k. In the following section, we use the expected value of K, μ_K, to approximate these bounds and simplify the model further for computation.

3.3.5. Expected cluster number approximation

Approximating k in the summation bounds with the expected number of clusters

$$\mu_K := \sum_{k=1}^{N} kP_k,$$

we can simplify (3.57) to

$$\begin{aligned}\partial_t p + v_1\partial_{x_1}p = &-\left(\sum_{k=2}^{N}(k-1)P_k\right)\int_{\boldsymbol{D}_\Gamma(\mu_K,m_1)}|v_1-v_2|\,p(z_1)p(z_2)\delta_{12}\,\mathrm{d}z_2\\ &+\left(\sum_{k=2}^{N}(k-1)P_k\right)\int_{\boldsymbol{D}_\Gamma(\mu_K,m_1)}|{}^*v_1-{}^*v_2|\,p({}^*z_1)p({}^*z_2)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_2\\ &+\left(\sum_{k=2}^{N}\frac{k}{2}P_k\right)\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|\,p('z_1)p(z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\\ &+\left(\sum_{k=3}^{N}\frac{k-2}{2}P_k\right)p(z_1)\int_{\boldsymbol{C}^2_\Gamma(\mu_K,m_1)}|v_2-v_3|\,p(z_2)p(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.\end{aligned}$$

As μ_K is not necessarily an integer, we use the floor, $\lfloor\mu_K\rfloor$, in the summation bounds. Specifically the new bounds are (see (3.44) for the original definition)

$$\Gamma_1(\mu_K,m_1) := N - m_1 - \lfloor\mu_K\rfloor + 2$$

and

$$\Gamma_2(\mu_K,m_1) := N - m_1 - \lfloor\mu_K\rfloor + 3.$$

The summation bounds in the first and second terms can be extended to $k = 1$ and the fourth term extended to $k = 2$ as these additional terms are zero. Extending the summation bounds in the third and fourth terms to begin at $k = 1$ requires assuming the probability the system has only one cluster is negligible. Making this assumption and noting $\sum_{k=1}^{N} P_k = 1$, these sums can then be defined in terms of μ_K to give the simplified model

$$\begin{aligned}\partial_t p + v_1 \partial_{x_1} p = & -(\mu_K - 1) \int_{\boldsymbol{D}_\Gamma(\mu_K, m_1)} |v_1 - v_2|\, p(z_1) p(z_2) \delta_{12}\, \mathrm{d}z_2 \\ & + (\mu_K - 1) \int_{\boldsymbol{D}_\Gamma(\mu_K, m_1)} |{}^*v_1 - {}^*v_2|\, p({}^*z_1) p({}^*z_2) \delta_{12} \left| D\psi_{12}^{-1} \right| \mathrm{d}z_2 \\ & + \frac{\mu_K}{2} \int_{\boldsymbol{C}'(z_1)} |'v_1 - v_2|\, p('z_1) p(z_2) \delta_{12} \frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\, \mathrm{d}z_2 \\ & + \left(\frac{\mu_K}{2} - 1 \right) p(z_1) \int_{\boldsymbol{C}^2_\Gamma(\mu_K, m_1)} |v_2 - v_3|\, p(z_2) p(z_3) \delta_{23}\, \mathrm{d}z_2 \mathrm{d}z_3.\end{aligned} \tag{3.60}$$

Direct μ_K approximation

We could calculate μ_K without further assumptions by computing the P_k distribution using (3.59) and taking the mean. However, we present another approach here approximating μ_K directly from $p(t; z)$ that should be computationally simpler and does not require tracking P_k. Consider the mass distribution for a system with k clusters, defined by marginalising the position and velocity variables:

$$p^k(t; m_1, \ldots, m_k) = \int_{\mathbb{R}^k} \int_{\Omega^k} p^k(t; x_1, \ldots, x_k; v_1, \ldots, v_k; m_1, \ldots, m_k)\, \mathrm{d}x_1 \ldots \mathrm{d}x_k\, \mathrm{d}v_1 \ldots \mathrm{d}v_k.$$

As we have the constraint that $m_1 + m_2 + \cdots + m_k = N$ we can write:

$$N p^k(t; m_1, \ldots, m_k) = (m_1 + m_2 + \cdots + m_k) p^k(t; m_1, \ldots, m_k).$$

Marginalising the m variables on both sides, we have

$$\begin{aligned} N P_k(t) &= \sum_{m_1=1}^{N} \cdots \sum_{m_k=1}^{N} (m_1 + m_2 + \cdots + m_k) p^k(t; m_1, \ldots, m_k) \\ &= \sum_{m_1=1}^{N} \cdots \sum_{m_k=1}^{N} \sum_{i=1}^{k} m_i p^k(t; m_1, \ldots, m_k). \end{aligned}$$

Note that we can take all the mass sums to N as we can assume $p^k(t; m_1, \ldots, m_k) = 0$ whenever $m_1 + m_2 + \cdots + m_k \neq N$. Using the probability invariance under relabelling, this can be rewritten as:

$$N P_k(t) = k \sum_{m_1=1}^{N} \cdots \sum_{m_k=1}^{N} m_1 p^k(t; m_1, \ldots, m_k),$$

which simplifies to

$$N P_k(t) = k \sum_{m_1=1}^{N} m_1 p^k(t; m_1) = k \sum_{m_1=1}^{N} m_1 p_1(t; m_1 \,|\, k) P_k.$$

Summing over k we have

$$\sum_{k=1}^{N} NP_k(t) = \sum_{k=1}^{N} kP_k(t) \sum_{m_1=1}^{N} m_1 p_1(t; m_1 \,|\, k),$$

and if we take the *Clusterzahlansatz* (3.56), i.e.

$$p_1(t; m_1 \,|\, k) \approx p(t; m_1) := \sum_{k=1}^{N} p_1^k(t; m_1),$$

this becomes

$$N \approx \mu_K(t) \sum_{m_1=1}^{N} m_1 p(t; m_1). \tag{3.61}$$

This effectively encodes the relation that the product of the mean cluster mass and the mean number of clusters should be the total number of particles. We can rearrange this to express $\mu_K(t)$ as a function of $p(t; x; v; m)$ as

$$\mu_K(t) \approx N \left(\sum_{m_1=1}^{N} \int_{\mathbb{R}\times\Omega} m_1 p(t; x_1; v_1; m_1) \, \mathrm{d}x_1 \mathrm{d}v_1 \right)^{-1}, \tag{3.62}$$

and so complete the derivation of our clustering model as (3.60) and (3.62) solved together.

3.4. Short-range clustering interactions

We are now ready to generalise the clustering model from strictly point-wise interactions to short-range interactions. This includes, for example, the Cucker-Smale model (2.70), which was one of the motivating systems for our approach here (see §3.1.1). Specifically we consider a system with N indistinguishable particles with positions $X_i \in \Omega = \mathbb{T}^1$ and velocities $V_i \in \mathbb{R}$ for $i \in \{1, \ldots, N\}$. These evolve through the generalised ODE model

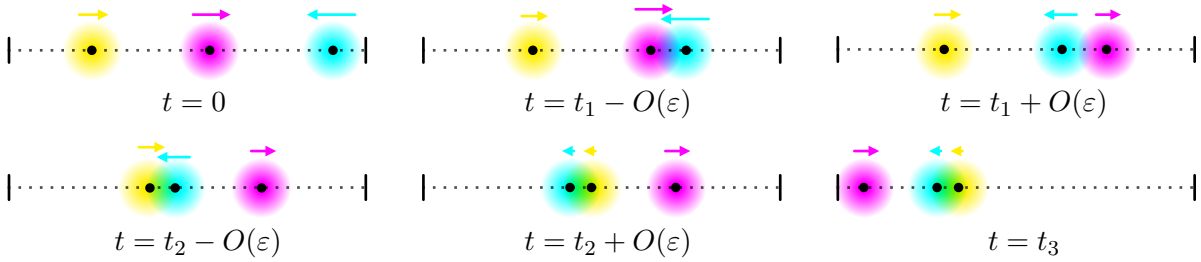
$$\begin{aligned} \dot{X}_i(t) &= V_i, \\ \dot{V}_i(t) &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j\neq i}}^{N} H\left(\frac{X_i - X_j}{\varepsilon}, V_i - V_j \right), \end{aligned} \tag{3.63}$$

with the interaction H assumed to be short-ranged, satisfying

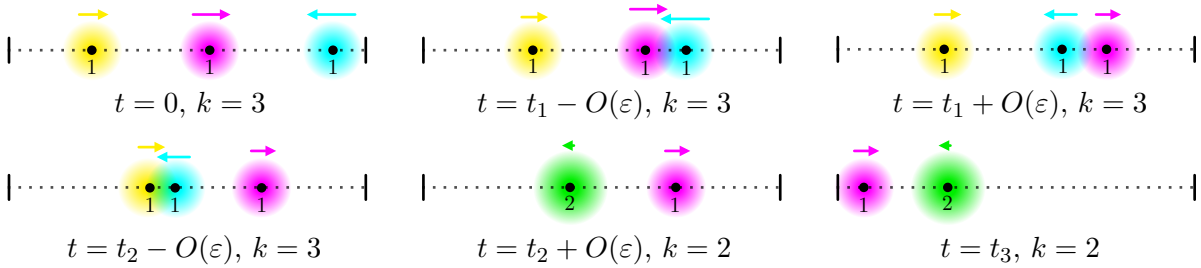
$$H(x, v) = O(|x|^{-1-\delta}) \text{ as } |x| \to \infty, \tag{3.64}$$

for some $\delta > 0$ as in (2.2). See Figure 3.5a for a schematic representation.

To derive our kinetic clustering model, we first define an approximate individual-based ODE model where we make the point-cluster approximation (as motivated in §3.1). That is, we approximate sets of particles that remain bound after interacting (i.e. $V_i(t) - V_j(t) \to 0$ and $X_i(t) - X_j(t) = O(\varepsilon)$ as $t \to \infty$) as a single point object called a cluster, with an internal property, labelled its 'mass', representing the number of particles it comprises. See Figure 3.5b



(a) Particle system with a short-range clustering interaction.



(b) Cluster system with a short-range clustering interaction (after the point-cluster approximation).

Figure 3.5.: (a) Schematic of a particle system with a short-range clustering interaction in the form of (3.63). A non-clustering interaction occurs around t_1 while a clustering interaction occurs at t_2. These cyan and yellow particles remain highly correlated in position and velocity after this interaction, with a separation of $O(\varepsilon)$. (b) Schematic of an equivalent system but after the point-cluster approximation is made. After the clustering interaction, the two correlated particles are treated as a combined cluster with mass 2, and the number of clusters in the system goes from $k = 3$ to $k = 2$. This system is described by (3.66). Note that the magnitude of the interaction force from the two-particle cluster has increased proportionately.

for a schematic of this approximated system. As in the point-wise clustering interaction model (§3.3), define $K(t)$ to be a random variable representing the number of clusters in the system at time t. Given a specific $K(t) = k$, the k clusters in the system are defined by positions $X_i(t) \in \Omega \equiv \mathbb{T}^1$, velocities $V_i(t) \in \mathbb{R}^d$ and masses $M_i(t) \in \{1, \dots, N\}$, for $i \in \{1, \dots, k\}$. We assume conservation of mass in clustering interactions, i.e. $\sum_{i=1}^{k} M_i(t) = N$. To define the evolution equations for this system, we first determine how the interaction force due to H in (3.63) is modified for clusters of different masses. Consider the effect of M_1 particles at the same position, X_1, and with the same velocity, V_1, (i.e. in a cluster of mass M_1) on an external particle, labelled i, under the interaction H from (3.63). The acceleration of the particle due to interaction with this cluster is

$$\dot{V}_i(t) = \frac{1}{\varepsilon} \sum_{j=1}^{M_1} H\left(\frac{X_i - X_1}{\varepsilon}, V_i - V_1\right) = \frac{1}{\varepsilon} M_1 H\left(\frac{X_i - X_1}{\varepsilon}, V_i - V_1\right), \tag{3.65}$$

where j indexes the particles in the cluster of mass M_1. As this effect is the same for any other particles at the same position as particle i, it represents the acceleration of a cluster at this position. Thus, summing the interactions between all clusters and assuming interactions between particles in the same cluster only serve to maintain the correlation of the cluster, we

have the cluster-based ODE model

$$\begin{aligned}\dot{X}_i(t) &= V_i,\\ \dot{V}_i(t) &= \frac{1}{\varepsilon}\sum_{\substack{j=1\\ j\neq i}}^{k} M_j H\left(\frac{X_i - X_j}{\varepsilon}, V_i - V_j\right),\end{aligned} \tag{3.66}$$

for $i \in \{1, \ldots, k\}$. This model is for a specific number of clusters k, and the full model is a hierarchy of such cluster-based ODE models, from $k = 1$ to $k = N$. Note that we will never solve the system (3.66) directly; we only use it as a starting point to derive the kinetic PDE model.

Multiplying by the mass of cluster i, the acceleration in (3.66) can be seen to be analogous to many common force laws, where the force between two objects is proportional to the product of their masses. Defining the force on particle i as F_i, we have

$$F_i(t) := M_i \dot{V}_i(t) = M_i \frac{1}{\varepsilon}\sum_{\substack{j=1\\ j\neq i}}^{k} M_j H\left(\frac{X_i - X_j}{\varepsilon}, V_i - V_j\right).$$

The clustering model from §3.3, where we assume the interactions occur strictly at a point, can be encompassed in this framework as the limit $\varepsilon \to 0$ of (3.66). We also note that although we have defined the cluster interaction term to behave as if a cluster is equivalent to multiple particles at the same location, the analysis that follows can be extended to generalised relationships between the interaction and cluster masses $H((X_i - X_j)/\varepsilon, V_i - V_j, M_i, M_j)$. Such a clustering model might be used as a phenomenological starting point, rather than beginning with a particle model.

From this hierarchy of ODE models, we define a hierarchy of k-cluster probability distributions $p^k(t; x_1, \ldots, x_k; v_1, \ldots, v_k; m_1, \ldots, m_k)$, with $x_i \in \Omega \equiv \mathbb{T}^1$, $v_i \in \mathbb{R}$ and $m_i \in \{1, \ldots, N\}$ for $i \in \{1, \ldots, k\}$ and $k \in \{1, \ldots, N\}$. Considering the change of probability in an infinitesimal box, we obtain the evolution equation for p^k

$$\partial_t p^k + \sum_{i=1}^{N} v_i \partial_{x_i} p^k + \frac{1}{\varepsilon}\sum_{\substack{i,j=1\\ j\neq i}}^{N} \partial_{v_i}\left(m_j H_{ij} p^k\right) = \sum_{i=1}^{k} q^k_{(i)}, \tag{3.67}$$

where

$$H_{ij} := H\left(\frac{x_i - x_j}{\varepsilon}, v_i - v_j\right).$$

and $q^k_{(i)}$ is defined as in (3.26) as the source term due to clustering in the p^{k+1} equation. This is analogous to (2.4), except it now includes the clustering source terms $q^k_{(i)}$.

We now consider a series of matched asymptotic expansions in the interaction range ε. We define an outer region where $|x_i - x_j| \gg \varepsilon$ for all $i, j \in \{1, \ldots, k\}$ and a set of inner regions, each where $|x_i - x_j| \sim \varepsilon$ for some i and j. Considering (3.67) in the outer region, if we take $|x_i - x_j| \sim 1$ we have $|x_i - x_j|/\varepsilon \sim 1/\varepsilon$ and $H_{ij} = O(\varepsilon^{1+\delta})$ due to the short-range condition. This means that the interaction terms are $O(\varepsilon^{\delta})$ and are not present in the leading order

equation.

However, the outer equation will have boundary terms representing flux entering and exiting the inner regions. Considering the inner region with $|x_i - x_j| \sim \varepsilon$ and taking $\varepsilon \to 0$, the probability flux entering is $|v_i - v_j|\,{}^{-}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij}$ and the flux exiting is $|v_i - v_j|\,{}^{+}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij}$, with limits ${}^{-}p^k_{(i,j)}$ and ${}^{+}p^k_{(i,j)}$ defined in (3.13) and (3.14), respectively. Thus, mapping i and j over all the inner regions, the leading-order equation in the outer region is equivalent to (3.26):

$$\begin{aligned}\partial_t p^k + \sum_{i=1}^{k} \partial_{x_i}(v_i p^k) = &-\frac{1}{2}\sum_{\substack{i,\,j=1\\ j\neq i}}^{k} |v_i - v_j|\,{}^{-}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij} \\ &+\frac{1}{2}\sum_{\substack{i,\,j=1\\ j\neq i}}^{k} |v_i - v_j|\,{}^{+}p^k_{(i,j)}(t; z_i, z_j)\delta_{ij} \\ &+\sum_{i=1}^{k} q^k_{(i)}(t; z_i).\end{aligned} \tag{3.68}$$

In order to use the mass-restricted *Stosszahlansatz* ((3.46) and (3.48)) we need to express the outgoing fluxes, $|v_i - v_j|^{-}p^k_{(i,j)}$, and clustering source terms $q^k_{(i)}$ in terms of incoming fluxes. To do this, we consider, without loss of generality, the specific inner region with $|x_1 - x_2| \sim \varepsilon$ and $|x_i - x_j| \gg \varepsilon$ if either of $i, j > 2$. Define the new inner variables

$$\tilde{x}_1 := x_1, \quad \tilde{x}_{21} := \frac{x_2 - x_1}{\varepsilon},$$

with first-order derivatives

$$\partial_{x_1} = \partial_{\tilde{x}_1} - \frac{1}{\varepsilon}\partial_{\tilde{x}_{21}}, \qquad \partial_{x_2} = \frac{1}{\varepsilon}\partial_{\tilde{x}_{21}}.$$

Changing (5.4) to these inner variables and defining $\tilde{p}^k_{(1,2)}$ to be the corresponding inner solution we get

$$\begin{aligned}&\partial_t \tilde{p}^k_{(1,2)} + v_1\partial_{\tilde{x}_1}\tilde{p}^k_{(1,2)} + \frac{1}{\varepsilon}(v_2 - v_1)\partial_{\tilde{x}_{21}}\tilde{p}^k_{(1,2)} + \sum_{i=3}^{N} v_i\partial_{x_i}\tilde{p}^k_{(1,2)} \\ &\quad + \frac{1}{\varepsilon}\sum_{\substack{i,j=1\\ j\neq i\\ (i,j)\neq(1,2)\\ (i,j)\neq(2,1)}}^{N} \partial_{v_i}\left(m_j H_{ij}\tilde{p}^k_{(1,2)}\right) + \frac{1}{\varepsilon}\partial_{v_1}\left(m_2 H_{12}\tilde{p}^k_{(1,2)}\right) + \frac{1}{\varepsilon}\partial_{v_2}\left(m_1 H_{21}\tilde{p}^k_{(1,2)}\right) = \sum_{i=1}^{k} q^k_{(i)},\end{aligned}$$

where $H_{12} = H(-\tilde{x}_{21}, v_1 - v_2)$ and $H_{21} = H(\tilde{x}_{21}, v_2 - v_1)$. As $\tilde{x}_{21} \sim 1$, it follows that $H_{12}, H_{21} \sim 1$ from the definition of H. As all other H_{ij} are $O(\varepsilon^{1+\delta})$, to leading order (also

assuming $m_i \ll N$ for all i) we have

$$(v_2 - v_1)\partial_{\tilde{x}_{21}}\tilde{p}^k_{(1,2)} + \partial_{v_1}\left(m_2 H\left(-\tilde{x}_{21}, v_1 - v_2\right)\tilde{p}^k_{(1,2)}\right) + \partial_{v_2}\left(m_1 H\left(\tilde{x}_{21}, v_2 - v_1\right)\tilde{p}^k_{(1,2)}\right) = 0, \quad (3.69)$$

which is clustering model equivalent of (2.13). This is a first-order quasilinear equation that has characteristics similar to (2.27):

$$\frac{\mathrm{d}\tilde{x}_{21}}{\mathrm{d}\tau} = v_2 - v_1, \quad (3.70a)$$

$$\frac{\mathrm{d}v_1}{\mathrm{d}\tau} = m_2 H\left(-\tilde{x}_{21}, v_1 - v_2\right), \quad (3.70b)$$

$$\frac{\mathrm{d}v_2}{\mathrm{d}\tau} = m_1 H\left(\tilde{x}_{21}, v_2 - v_1\right), \quad (3.70c)$$

$$\frac{1}{\tilde{p}^k_{(1,2)}}\frac{\mathrm{d}\tilde{p}^k_{(1,2)}}{\mathrm{d}\tau} = -m_2\partial_{v_1}H\left(-\tilde{x}_{21}, v_1 - v_2\right) - m_1\partial_{v_2}H\left(\tilde{x}_{21}, v_2 - v_1\right). \quad (3.70d)$$

To simplify this, define $\tilde{\boldsymbol{z}}_{21} \equiv (\tilde{x}_{21}, v_1, v_2)$. Then (3.69) becomes

$$\nabla_{\tilde{\boldsymbol{z}}_{21}} \cdot \left(\frac{\mathrm{d}\tilde{\boldsymbol{z}}_{21}}{\mathrm{d}\tau}\tilde{p}^k_{(1,2)}\right) = 0, \quad (3.71)$$

with characteristic projections defined by

$$\frac{\mathrm{d}\tilde{\boldsymbol{z}}_{21}}{\mathrm{d}\tau} = \left(v_2 - v_1, m_2 H\left(-\tilde{x}_{21}, v_1 - v_2\right), m_1 H\left(\tilde{x}_{21}, v_2 - v_1\right)\right). \quad (3.72)$$

As such, to leading order, the evolution of the probability density can be modelled as a divergence-free vector field in three-dimensional $\tilde{\boldsymbol{z}}_{21}$ space, which we will now use to derive the probability flux conservation conditions across an interaction.

3.4.1. Non-clustering flux conservation condition

For non-clustering interactions, in a similar vein to §2.1.5 we define the limiting outgoing quantities (as $\tau \to \infty$) to be $\tilde{x}^+_{21}$, v^+_1, v^+_2 and ${}^+\tilde{p}^k_{(1,2)}$ and limiting incoming quantities (as $\tau \to -\infty$) to be $\tilde{x}^-_{21}$, v^-_1, v^-_2 and ${}^-\tilde{p}^k_{(1,2)}$. We are interested in expressing the outgoing quantities in terms of their incoming equivalents such that we can replace them in (3.68) and apply the *Stosszahlansatz* to produce a closed system. In §3.2 and §3.3, we used the flux conservation condition (3.20) to play this role for point-wise non-clustering interactions. Here, we intend to show that we get an equivalent result by using the solution trajectories to (3.69). Similarly to (2.21), define ψ_{H,m_1,m_2} mapping from incoming to outgoing velocities:

$$(v^+_1, v^+_2) = \psi_{H,m_1,m_2}(v^-_1, v^-_2). \quad (3.73)$$

This plays the same role as ψ_{m_1,m_2} used in §3.2 and §3.3.

Consider, without loss of generality, a set of trajectories (or equivalently characteristic projections) that approach from $\tilde{x}^-_{21} = -\infty$ and hence have $v^-_2 > v^-_1$ (this is represented by the

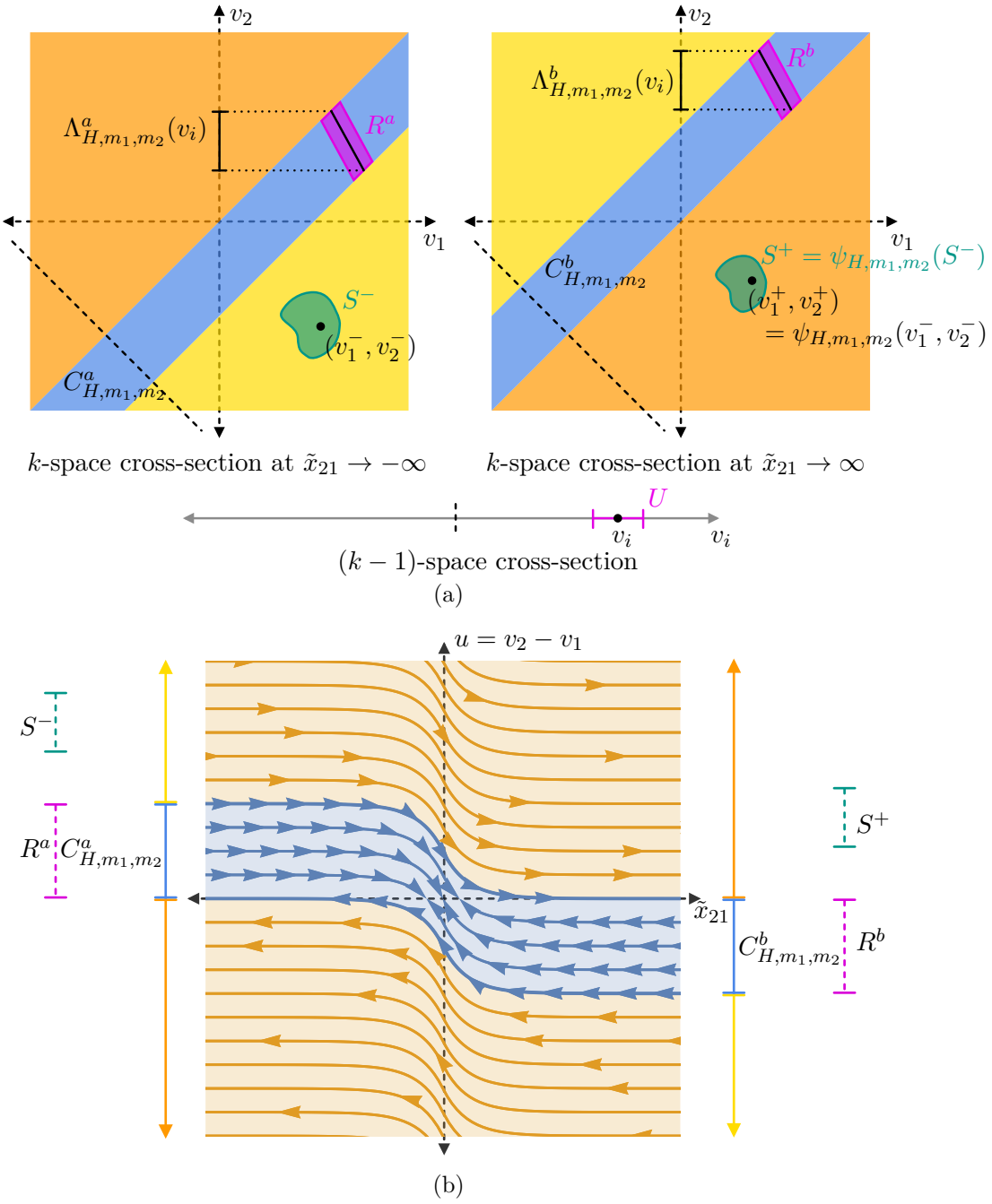


Figure 3.6.: Cross-sections of the inner region for a one-dimensional short-range clustering interaction. (a) Schematic of the boundary of the inner region at $\tilde{x}_{21} \to -\infty$ and $\tilde{x}_{21} \to \infty$. The yellow area represents trajectories entering the region approaching a non-clustering interaction. These will leave the inner region through the orange area. The blue region represents C_{H,m_1,m_2}, trajectories entering the inner region approaching a clustering interaction. Note that, in contrast to Figure 3.4, this region is split into C^a_{H,m_1,m_2} at $\tilde{x}_{21} \to -\infty$ and C^b_{H,m_1,m_2} at $\tilde{x}_{21} \to \infty$, with $C_{H,m_1,m_2} = C^a_{H,m_1,m_2} \cup C^b_{H,m_1,m_2}$, as particles can enter the inner region from both ends, depending on the sign of $v_2^- - v_1^-$. The region R and the interval $\Lambda_{H,m_1,m_2}(v_i)$ are similarly split into two parts. The region S^- maps through a non-clustering interaction to S^+, and the region R maps through a clustering interaction to the region U in $(k-1)$-space. (b) A cross-section of trajectories in the inner region along the direction of $u = v_2 - v_1$. This is the direction shown by the diagonal dotted line in (a). These trajectories are for the Cucker-Smale model although the qualitative behaviour is generalisable.

yellow region in the left plot of Figure 3.6a). Specifically, consider trajectories with (v_1^-, v_2^-) in a closed and bounded region T^-. If we restrict T^- such that all trajectories leave the inner region at $\tilde{x}_{21}^+ = \infty$ then the set of outgoing velocities (v_1^+, v_2^+) will also be in a closed and bounded region $T^+ = \psi_{H,m_1,m_2}(T^-)$. Trajectories with other combinations of $\tilde{x}_{21}^-, \tilde{x}_{21}^+ \in \{\pm\infty\}$ can be treated similarly and (other than an assumed measure zero set of boundary trajectories) closed and bounded regions around them analogous to T^- and T^+ can be defined.

In three-dimensional $\tilde{\boldsymbol{z}}_{21} := (\tilde{x}_{21}, v_1, v_2)$ space, define a two dimensional surface S^- parameterised by $S^- := (-\infty, v_1^-, v_2^-)$ with $(v_1^-, v_2^-) \in T^-$. Similarly, define $S^+ := (\infty, v_1^+, v_2^+)$ with $(v_1^+, v_2^+) \in T^+$ (see Figure 3.6a). Joining these two surfaces with the characteristics that are incoming through S^- and outgoing through S^+ forms a three-dimensional region Z. Define the surface C as the boundary surface of Z not at $\tilde{x}_{21} = \pm\infty$ (i.e. $\partial Z = C \cup S^- \cup S^+$). Integrating (3.71) over Z and using the divergence theorem, we get

$$\int_{S^-} \tilde{p}_{(1,2)}^k \frac{\mathrm{d}\tilde{\boldsymbol{z}}_{21}}{\mathrm{d}\tau} \cdot \mathrm{d}\boldsymbol{S}_{S^-} + \int_{S^+} \tilde{p}_{(1,2)}^k \frac{\mathrm{d}\tilde{\boldsymbol{z}}_{21}}{\mathrm{d}\tau} \cdot \mathrm{d}\boldsymbol{S}_{S^+} + \int_C \tilde{p}_{(1,2)}^k \frac{\mathrm{d}\tilde{\boldsymbol{z}}_{21}}{\mathrm{d}\tau} \cdot \mathrm{d}\boldsymbol{S}_C = 0 \tag{3.74}$$

where $\boldsymbol{S}_{S^-}$, $\boldsymbol{S}_{S^-}$ and $\boldsymbol{S}_C$ are surface elements of their respective bounding surfaces. As C is defined by a set of solutions trajectories to (3.72), $\mathrm{d}\tilde{\boldsymbol{z}}_{21}/\mathrm{d}\tau$ will be tangent to the surface at any point and hence perpendicular to $\mathrm{d}\boldsymbol{S}_C$ giving

$$\frac{\mathrm{d}\boldsymbol{z}_{21}}{\mathrm{d}\tau} \cdot \mathrm{d}\boldsymbol{S}_C = 0.$$

For the surfaces at $\tilde{x}_{21} = \pm\infty$, generalising to other combinations of $\tilde{x}_{21}^-, \tilde{x}_{21}^+ \in \{\pm\infty\}$, we have,

$$\mathrm{d}\boldsymbol{S}_{S^-} = \left(\operatorname{sgn}\left(\tilde{x}_{21}^-\right) \mathrm{d}v_1^- \mathrm{d}v_2^-, 0, 0\right), \text{ and } \mathrm{d}\boldsymbol{S}_{S+} = \left(\operatorname{sgn}\left(\tilde{x}_{21}^+\right) \mathrm{d}v_1^+ \mathrm{d}v_2^+, 0, 0\right),$$

respectively. Substituting these and noting that $\operatorname{sgn}(v_2^+ - v_1^+) = \operatorname{sgn}(\tilde{x}_{21}^+)$ and $\operatorname{sgn}(v_2^- - v_1^-) = -\operatorname{sgn}(\tilde{x}_{21}^-)$, (3.74) becomes

$$-\int_{T^-} |v_2^- - v_1^-|\, {}^-\tilde{p}_{(1,2)}^k(v_1^-, v_2^-)\, \mathrm{d}v_1^- \mathrm{d}v_2^- + \int_{T^+} |v_2^+ - v_1^+|\, {}^+\tilde{p}_{(1,2)}^k(v_1^+, v_2^+)\, \mathrm{d}v_1^+ \mathrm{d}v_2^+ = 0,$$

where we only show the velocity arguments of the probability densities for clarity. Using the map $(v_1^+, v_2^+) = \psi_{H,m_1,m_2}(v_1^-, v_2^-)$, we can change the coordinates of the $\mathrm{d}v_1^-\mathrm{d}v_2^-$ to $\mathrm{d}v_1^+\mathrm{d}v_2^+$, giving

$$-\int_{T^+} \left[|v_2^- - v_1^-|\, {}^-\tilde{p}_{(1,2)}^k(v_1^-, v_2^-) \left|D\psi_{H,m_1,m_2}^{-1}\right| - |v_2^+ - v_1^+|\, {}^+\tilde{p}_{(1,2)}^k(v_1^+, v_2^+)\right] \mathrm{d}v_1^+ \mathrm{d}v_2^+ = 0.$$

As the region T^+ is arbitrary, we have

$$|v_2^- - v_1^-|\, {}^-\tilde{p}_{(1,2)}^k(t; \tilde{x}_1, \tilde{x}_{21}^-; v_1^-, v_2^-; m_1, m_2) \left|D\psi_{H,m_1,m_2}^{-1}\right| = |v_2^+ - v_1^+|\, {}^+\tilde{p}_{(1,2)}^k(t; \tilde{x}_1, \tilde{x}_{21}^+; v_1^+, v_2^+; m_1, m_2).$$

Matching this to the outer solution and expressing it in outer variables, the leading order

equation is

$$|{}^*v_2 - {}^*v_1|^- p^k_{(1,2)}(t; x_1, x_1; {}^*v_1, {}^*v_2, m_1, m_2) \left| D\psi^{-1}_{H,m_1,m_2} \right| = |v_2 - v_1|^+ p^k_{(1,2)}(t; x_1, x_1; v_1, v_2; m_1, m_2),$$

where $({}^*v_1, {}^*v_2) = \psi^{-1}_{H,m_1,m_2}(v_1, v_2)$. As this is identical to (3.20), we can see that this short-ranged interaction model matches the point-wise interaction model to leading order in ε.

3.4.2. Clustering flux conservation condition

The process for calculating the clustering flux conservation is similar, except now we consider clustering trajectories that enter the inner region at $\tilde{x}^-_{21} = \pm\infty$ and approach the $v_1 = v_2$ plane as $\tau \to \infty$. Given a clustering interaction with incoming velocities v_1^- and v_2^-, incoming masses m_1 and m_2, limiting final clustered velocity v_i and clustered mass $m_i = m_1 + m_2$, we can define an equivalent of the region R that contains all clustering trajectories that approach the $v_1 = v_2$ plane with v_i within an interval U. Integrating over R, we will once again be able to equate the flux over the incoming surface to the outgoing clustering flux in the source term $q^{k-1}_{(i)}$ with outgoing velocity $v_i \in U$. Following similar steps to §3.3.1, the integral of this incoming flux is

$$\int_U \int_{\Lambda_{H,m_1,m_2}(v_i)} \left| {}'v_i - v_2^- \right| \, {}^-\tilde{p}^k_{(1,2)}(t; \tilde{x}_1, \tilde{x}^-_{21}; {}'v_i, v_2^-; m_1, m_2) \frac{\mathrm{d}'v_i}{\mathrm{d}v_i} \, \mathrm{d}v_2^- \, \mathrm{d}v_i,$$

where we have changed the variable of the $\mathrm{d}v_1^-$ integral to $\mathrm{d}v_i$ and represented v_1^- as the shorthand function ${}'v_i = \hat{\lambda}^{-1}_{H,m_1,m_2,v_2^-}(v_i)$. The partially-evaluated backward clustering map, $\hat{\lambda}^{-1}$, plays the same role as that defined in §3.3, except that here it is defined using the limits of a clustering trajectory in the inner region under the interaction H rather than as a point-wise interaction map. The regions $\Lambda_{H,m_1,m_2}(v_i)$ and C_{H,m_1,m_2} also now depend on the interaction H but we will drop the explicit subscript H from here on for conciseness. Continuing to follow §3.3.1, if we are to equate this outgoing flux with the full clustering flux that is produced in p^{k-1} space at index i with position x_i, velocity $v_i \in U$ and mass m_i, i.e. $\int_U q^{k-1}_{(i)}(x_i, v_i, m_i) \, \mathrm{d}v_i$, we must sum over each combination of masses m_1 and m_2 such that $m_1 + m_2 = m_i$ and sum over each combination of indices that could be mapped to index i. For the masses, we must define a region R for each combination m_1 and m_2 and sum over the flux contribution from each. For the indices, we can use particle indistinguishability to relabel incoming particle 1 as i and incoming particle 2 as k and note that a total of $k/2$ index combinations are mapped to index i (from the same reasoning as in 3.3.1). Combining this, the total flux contributions are equated through

$$\int_U \left[\frac{k}{2} \sum_{m_k=1}^{m_i - 1} \int_{\Lambda_{m_i - m_k, m_k}(v_i)} \left| {}'v_i - v_k^- \right| \, {}^-\tilde{p}^k_{(i,k)}(t; \tilde{x}_i, \tilde{x}^-_{ki}; {}'v_i, v_k^-; m_i - m_k, m_k) \frac{\mathrm{d}'v_i}{\mathrm{d}v_i} \, \mathrm{d}'v_k - q^{k-1}_{(i)} \right] \mathrm{d}v_i = 0.$$

Noting that the integrand must be 0 as the interval U is arbitrary, after matching to leading order with outer variables, we have the equivalent of the clustering flux conservation condition

(3.33):

$$q_{(i)}^{k-1}(t;x_i,v_i,m_i) = \frac{k}{2}\sum_{m_k=1}^{m_i-1}\int_{\Lambda_{m_i-m_k,m_k}(v_i)} |{}'v_i - v_k| \; {}^{-}p_{(i,k)}^{k}(t;x_i,x_i;{}'v_i,v_k;m_i-m_k,m_k)\frac{\mathrm{d}'v_i}{\mathrm{d}v_i}\,\mathrm{d}v_k,$$

where ${}'v_i = \hat{\lambda}^{-1}_{m_i-m_k,m_k,v_k}(v_i)$.

Substituting both clustering and non-clustering flux conservation conditions into the outer equation (3.68), we arrive at an identical equation to (3.35): the leading-order evolution equations for the p^k hierarchy expressed in terms of incoming interaction probability fluxes. The same process as in §3.3 can then be followed to simplify this to a closed system in either $p_1^k(t;x_1;v_1;m_1)$ or $p(t;x_1;v_1;m_1)$.

3.5. Cluster model summary

Writing the arguments out in full, our final one-dimensional kinetic model for short-range clustering interactions is, to leading order in ε

$$\begin{aligned}
&\partial_t p(t;x_1;v_1;m_1) + v_1\partial_{x_1} p(t;x_1;v_1;m_1)\\
&\quad = -(\mu_K-1)\sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{\mathbb{R}} |v_1-v_2|\, p(t;x_1;v_1;m_1)p(t;x_1;v_2;m_2)\,\mathrm{d}v_2\\
&\qquad + (\mu_K-1)\sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{\mathbb{R}} |{}^*v_1-{}^*v_2|\, p(t;x_1;{}^*v_1;m_1)p(t;x_1;{}^*v_2;m_2)\left|D\psi^{-1}_{m_1,m_2}\right|\,\mathrm{d}v_2\\
&\qquad + \frac{\mu_K}{2}\sum_{m_2=1}^{m_1-1}\int_{\Lambda_{m_1-m_2,m_2}(v_1)} |{}'v_1-v_2|\, p(t;x_1;{}'v_1;m_1-m_2)p(t;x_1;v_2;m_2)\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}v_2\\
&\qquad + \left(\frac{\mu_K}{2}-1\right)p(t;x_1;v_1;m_1)\sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \leq N-m_1-\mu_K+3}}^{N}\int_{C_{m_2,m_3}}\int_{\Omega}\Big[|v_2-v_3|\\
&\qquad\qquad \times p(t;x_2;v_2;m_2)p(t;x_2;v_3;m_3)\Big]\mathrm{d}x_2\,\mathrm{d}v_2\mathrm{d}v_3,
\end{aligned} \tag{3.75}$$

where ${}'v_1 = \hat{\lambda}^{-1}_{m_1-m_2,m_2,v_2}(v_1)$ is the incoming velocity for a clustering interaction and $({}^*v_1,{}^*v_2) = \psi^{-1}_{m_1,m_2}(v_1,v_2)$ are the incoming velocities for a non-clustering interaction, as a result of the short-range interaction H. The mean number of clusters is estimated through

$$\mu_K(t) = N\left(\sum_{m_1=1}^{N}\int_{\mathbb{R}\times\Omega} m_1 p(t;x_1;v_1;m_1)\,\mathrm{d}x_1\mathrm{d}v_1\right)^{-1}. \tag{3.76}$$

Note that we write the non-integer μ_K in the summation bounds for m_2 in the first two terms instead of $\lfloor\mu_K\rfloor$ for conciseness. As this is at the edge of the mass distribution, the difference between using floor or ceiling of μ_K in the bound should be negligible.

The main approximations used in deriving this model are: the point-cluster approximation,

approximating clusters with non-zero spatial extent as point objects (see §3.4 and §3.1); the *Stosszahlansatz*, assuming particles approaching an interaction are uncorrelated (see §3.3.3); the *Clusterzahlansatz*, assuming the conditional velocity and mass distributions are independent of the number of the clusters in the system (see §3.3.4); and approximating the mass conservation restrictions using on the expected number of clusters in the system (see §3.3.5).

4. Numerical evaluation of one-dimensional clustering model

To evaluate the accuracy of the kinetic PDE model (3.75) and compare it to that of the mean-field model (2.10), we implement both here using discrete methods. Simulations of the individual-based model (3.63) (with initial conditions drawn as Monte Carlo samples from the initial distributions), are used for comparison. Two other models, a stochastic model similar to the direct-simulation Monte Carlo method and an event-based 'exact clustering' model, are used to distinguish the effect of the different approximations made. The comparisons are performed using the short-range Cucker-Smale collective behaviour model in a periodic one-dimensional spatial domain with an exponential interaction kernel, uniform spatial density and uniform initial velocity distribution.

4.1. Numerical implementation of short-range PDE

We use the fully reduced short-range PDE model (3.75) as the basis for the numerical implementation. Although Fourier-Galerkin spectral methods are more popular currently for numerically solving Boltzmann-style models in higher dimensions (see, for instance, [48, 46]), we aim to derive a simpler method here for the one-dimensional case, basing our approach on the discrete velocity methods used for modelling the Boltzmann equation (see [91] Chapter 2). The Fourier-Galerkin method will be analysed later in Chapter 6 when considering the higher-dimensional extension.

We consider the spatially uniform case, which allows us to write p as a function of t, v and m as

$$p(t; v; m) = \int_\Omega p(t; x; v; m)\,\mathrm{d}x = |\Omega| p(t; x; v; m).$$

A spatially independent solution to (3.75) then must satisfy

$$
\begin{aligned}
&|\Omega|\partial_t p(t;v_1;m_1) \\
&\quad = -(\mu_K-1)\sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{\mathbb{R}}|v_1-v_2|\,p(t;v_1;m_1)p(t;v_2;m_2)\,\mathrm{d}v_2 \\
&\qquad +(\mu_K-1)\sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{\mathbb{R}}|{}^*v_1-{}^*v_2|\,p(t;{}^*v_1;m_1)p(t;{}^*v_2;m_2)\left|D\psi^{-1}_{m_1,m_2}\right|\,\mathrm{d}v_2 \\
&\qquad +\frac{\mu_K}{2}\sum_{m_2=1}^{m_1-1}\int_{\Lambda_{m_1-m_2,m_2}(v_1)}|{}'v_1-v_2|\,p(t;{}'v_1;m_1-m_2)p(t;v_2;m_2)\frac{\mathrm{d}{}'v_1}{\mathrm{d}v_1}\,\mathrm{d}v_2 \\
&\qquad +\left(\frac{\mu_K}{2}-1\right)p(t;v_1;m_1)\sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \le N-m_1-\mu_K+3}}^{N}\int_{C_{m_2,m_3}}\Big[|v_2-v_3|\,p(t;v_2;m_2)p(t;v_3;m_3)\Big]\,\mathrm{d}v_2\mathrm{d}v_3,
\end{aligned}
\tag{4.1}
$$

with μ_K computed through

$$
\mu_K = N\left(\sum_{m_1=1}^{N}\int_{\mathbb{R}} m_1 p(t;v_1;m_1)\mathrm{d}v_1\right)^{-1},
$$

and the backward velocity maps are given by

$${}'v_1 = \hat{\lambda}^{-1}_{m_1-m_2,m_2,v_2}(v_1) \text{ and } ({}^*v_1,{}^*v_2) = \psi^{-1}_{m_1,m_2}(v_1,v_2).$$

To implement (4.1) numerically, we consider a linear discretisation of velocity space into V bins of width Δv. The lower and upper boundaries of the bin i are defined as v_i^-, v_i^+, such that $v_{i-1}^+ = v_i^-$, and each bin is assigned a representative point $\bar{v}_i = (v_i^+ - v_i^-)/2$. For small N, we can keep the masses as defined, with m taking all values from 1 to N. However, for large N, this will be computationally infeasible. To remedy this, we bin the set of masses onto a linear-geometric grid with M points, as suggested by Bäbler et al. [21]. This setup maintains the linear grid for small masses to avoid distortions and then transitions to a geometric grid at larger masses where the distinction between mass m and mass $m+1$ is less significant. We define the grid as the set $\{\bar{m}_a, a \in 1,\dots,M\}$, with M_l linear grid points and $M - M_l$ geometric points. The grid points are defined as

$$
\bar{m}_a = \begin{cases} a, & a = 1,\dots,M_l \\ M_l\left(\frac{N}{M_l}\right)^{\frac{a-M_l}{M-M_l}}, & a = M_l+1,\dots,M. \end{cases}
\tag{4.2}
$$

The number of linear grid points M_l is chosen given M and N to be the smallest value such that the first geometric grid spacing is greater than 1 to ensure a smooth transition between linear and geometric regions. Similar to the velocity binning, we label the lower and upper bounds of the bin represented by grid point $\bar{m}_a$ as m_a^- and m_a^+, respectively. Bin a is defined to include the set of $\bar{w}_a$ masses: $\{\lfloor m_a^- \rfloor + 1,\dots,\lfloor m_a^+ \rfloor\}$. For the linear-geometric case discussed,

the bin bounds are defined as $m_1^- = 0.5$ and

$$m_a^+ = \begin{cases} m_a + 0.5, & a = 1, \ldots, M_l \\ M_l \left(\frac{N}{M_l}\right)^{\frac{i - M_l + 0.5}{M - M_l}} = m_a \left(\frac{N}{M_l}\right)^{\frac{0.5}{M - M_l}}, & a = M_l + 1, \ldots, M - 1 \\ m_M + 0.5, & a = M. \end{cases} \tag{4.3}$$

With this velocity discretisation and mass binning, we can define the discrete set of binned probabilities:

$$P_{i,a}(t) := \sum_{m=\lfloor m_a^- \rfloor + 1}^{\lfloor m_a^+ \rfloor} \int_{v_i^-}^{v_i^+} p(t; v; m) \mathrm{d}v, \tag{4.4}$$

for $i \in \{1, \ldots, V\}$ representing the velocity bin and $a \in \{1, \ldots, M\}$ representing the mass bin. Integrating (4.1) over this velocity-mass bin, we can derive a time evolution equation for these binned probabilities (with the details of the derivation in §A.5):

$$\begin{aligned} |\Omega| \partial_t P_{i,a} = & -(\mu_K - 1) \sum_{\substack{b=1 \\ \bar{m}_a + \bar{m}_b \leq N - \mu_K + 2}}^{M} \sum_{j=1}^{V} |\bar{v}_j - \bar{v}_i| P_{i,a} P_{j,b} \\ & + (\mu_K - 1) \sum_{\substack{b=1 \\ \bar{m}_a + \bar{m}_b \leq N - \mu_K + 2}}^{M} \sum_{\substack{j,k=1 \\ (\bar{v}_j, \bar{v}_k) \in \mathbb{R} \setminus C_{\bar{m}_a, \bar{m}_b}}}^{V} |\bar{v}_k - \bar{v}_j| \, \gamma_{\bar{v}_i}(\bar{v}_k^*) P_{k,a} P_{j,b} \\ & + \frac{\mu_K}{2} \sum_{\substack{b,c=1 \\ \bar{m}_b + \bar{m}_c \leq \bar{m}_{a+1}}}^{M} \sum_{\substack{j,k=1 \\ (\bar{v}_j, \bar{v}_k) \in C_{jk}}}^{V} |\bar{v}_k - \bar{v}_j| \, \gamma_{\bar{v}_i}(\bar{v}_k') \gamma_{\bar{m}_a}(\bar{m}_c') P_{k,c} P_{j,b} \\ & + \left(\frac{\mu_K}{2} - 1\right) P_{i,a} \sum_{\substack{b,c=1 \\ \bar{m}_b + \bar{m}_c \\ \leq N - \bar{m}_a - \mu_K + 3}}^{M} \sum_{\substack{j,k=1 \\ (\bar{v}_j, \bar{v}_k) \in C_{\bar{m}_b, \bar{m}_c}}}^{V} |\bar{v}_k - \bar{v}_j| \, P_{k,c} P_{j,b} \end{aligned} \tag{4.5}$$

We have used a technique based on the fixed-pivot method of Kumar and Ramkrishna [77] to deal with interactions where product particles are not produced at a grid point in the discretisation. This involves introducing the factors $\gamma_{\bar{v}_i}(v)$ and $\gamma_{\bar{m}_a}(m)$, defined as

$$\gamma_{\bar{m}_a}(m) = \begin{cases} \frac{m - \bar{m}_{a-1}}{\bar{m}_a - \bar{m}_{a-1}}, & \bar{m}_{a-1} < m < \bar{m}_a \\ \frac{\bar{m}_{a+1} - m}{\bar{m}_{a+1} - \bar{m}_a}, & \bar{m}_a \leq m < \bar{m}_{a+1} \\ 0, & \text{otherwise,} \end{cases}$$

and

$$\gamma_{\bar{v}_i}(v) = \begin{cases} \frac{v - \bar{v}_{i-1}}{\Delta v}, & \bar{v}_{i-1} < v < \bar{v}_i \\ \frac{\bar{v}_{i+1} - v}{\Delta v}, & \bar{v}_i \leq v < \bar{v}_{i+1} \\ 0, & \text{otherwise,} \end{cases}$$

respectively, representing the proportion of a particle produced with velocity v and mass m that would be allocated to the grid point at $(\bar{v}_i, \bar{m}_a)$. This allocation choice is designed to

ensure the zeroth and first moments of each distribution are conserved. In the case here, it ensures that the numerical scheme preserves the total probability, mass, momentum and mean velocity of the PDE model.

4.1.1. Outputs

The outputs we will consider for model comparison can be computed from the binned probabilities as follows. The cluster velocity distribution can be computed by summing over m as

$$p(t;v) = \sum_{m=1}^{N} p(t;v;m) \approx \frac{1}{\Delta v} \sum_{a=1}^{M} P_{i(v),a}(t), \tag{4.6}$$

where $i(v)$ represents the velocity bin such that $v_i^- < v < v_i^+$. Note that we must divide by Δv to ensure the output velocity distribution is normalised to be a probability density. The cluster mass distribution is approximated as

$$p(t;m) = \int_{\mathbb{R}^d} p(t;v;m)\mathrm{d}v \approx \frac{1}{\bar{w}_{a(m)}} \sum_{i=1}^{V} P_{i,a(m)}(t), \tag{4.7}$$

where, similarly, $a(m)$ represents the velocity bin such that $\lfloor m_a^- \rfloor + 1 < m < \lfloor m_i^+ \rfloor$. It is also common to track the mass-weighted cluster mass distribution, defined as

$$f(t;m) := \mu_K(t)\, m\, p(t;m). \tag{4.8}$$

This represents the expected number of particles in clusters of size m. We label this f by analogy to the mass distribution conventionally used when describing the Boltzmann equation (see [31] §1.6). From (3.61), it follows, given the *Clusterzahlansatz*, that

$$\sum_{m=1}^{N} f(t;m) = \mu_K(t) \sum_{m=1}^{N} m\, p(t;m) = N, \tag{4.9}$$

and so the normalisation constant for f is N. As such, we compute

$$\frac{f(t,m)}{N} = \frac{\mu_K(t)}{N} \int_{\mathbb{R}^d} m\, p(t;v;m)\mathrm{d}v \approx \frac{\mu_K(t)}{N} \frac{1}{\bar{w}_{a(m)}} \sum_{i=1}^{V} \bar{m}_a P_{i,a(m)}(t). \tag{4.10}$$

Following this, again using (3.61), the mean number of clusters can be approximated as

$$\mu_K(t) \approx N \left(\sum_{m=1}^{N} \int_{\mathbb{R}} m\, p(t;v;m)\mathrm{d}v \right)^{-1} \approx N \left(\sum_{a=1}^{M} \sum_{i=1}^{V} \bar{m}_a P_{i,a} \right)^{-1}. \tag{4.11}$$

Finally, the mean kinetic energy per particle can be computed as

$$E(t) = \frac{1}{2} \sum_{m=1}^{N} \int_{\mathbb{R}} |v|^2 p(t;v,m)\, \mathrm{d}v \approx \frac{1}{2} \sum_{a=1}^{M} \sum_{i=1}^{V} |\bar{v}_i|^2 P_{i,a}. \tag{4.12}$$

Note that this definition does not contain the cluster mass as p is a probability distribution over the clusters. The mean kinetic energy per cluster is given by

$$E_c(t) = \frac{1}{2} \sum_{m=1}^{N} \int_{\mathbb{R}} m|v|^2 p(t; v, m) \, \mathrm{d}v, \tag{4.13}$$

which can be divided by the cluster mass to get the average per particle.

4.1.2. Cucker-Smale model

In the short-range version of the Cucker-Smale model of collective behaviour, particles' positions and velocities evolve according to

$$\begin{aligned} \dot{X}_i(t) &= V_i \\ \dot{V}_i(t) &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{N} \phi\left(\frac{|X_i - X_j|}{\varepsilon}\right)(V_j - V_i). \end{aligned} \tag{4.14}$$

If we make the point-cluster approximation following §3.4 and allow sets of M particles that were highly correlated post-interaction to be represented by a cluster with mass M, the velocity of each cluster would instead evolve as

$$\dot{V}_i(t) = \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{k} M_j \phi\left(\frac{|X_i - X_j|}{\varepsilon}\right)(V_j - V_i),$$

where $k < N$ is the total number of clusters remaining in the system. This is in the form of (3.66), with interaction

$$H\left(\frac{X_i - X_j}{\varepsilon}, V_i - V_j\right) = \phi\left(\frac{|X_i - X_j|}{\varepsilon}\right)(V_j - V_i).$$

With this interaction, the inner equation for the evolution of the joint probability distribution (3.69) would be

$$(v_2 - v_1)\partial_{\tilde{x}_{21}} \tilde{p}^k_{(1,2)} + m_2 \phi\left(|\tilde{x}_{21}|\right) \partial_{v_1}\left((v_2 - v_1)\tilde{p}^k_{(1,2)}\right) + m_1 \phi\left(|\tilde{x}_{21}|\right) \partial_{v_2}\left((v_1 - v_2)\tilde{p}^k_{(1,2)}\right) = 0. \tag{4.15}$$

Changing variables to centre of mass velocity $\bar{v} = (m_1 v_1 + m_2 v_2)/(m_1 + m_2)$ and relative velocity $u = v_2 - v_1$, this has characteristic equations

$$\frac{\mathrm{d}\tilde{x}_{21}}{\mathrm{d}\tau} = u, \tag{4.16a}$$

$$\frac{\mathrm{d}\bar{v}}{\mathrm{d}\tau} = 0, \tag{4.16b}$$

$$\frac{\mathrm{d}u}{\mathrm{d}\tau} = -(m_1 + m_2)\phi\left(|\tilde{x}_{21}|\right) u, \tag{4.16c}$$

$$\frac{1}{\tilde{p}^k_{(1,2)}} \frac{\mathrm{d}\tilde{p}^k_{(1,2)}}{\mathrm{d}\tau} = (m_1 + m_2)\phi\left(|\tilde{x}_{21}|\right). \tag{4.16d}$$

Note that the first three characteristic equations are equivalent to the evolution equations for the relative position, centre of mass velocity and relative velocity for two clusters under the Cucker-Smale interaction. We note that these characteristic equations are very similar to those in (2.71), except that we are now in one dimension and the RHS (4.16c) and (4.16d) has a coefficient of $m_1 + m_2$ instead of 2. Following a similar analysis to that in §2.5, we find that trajectories with incoming relative speed $|u^-| < (m_1 + m_2)\Phi$ will cluster, while those with $|u^-| \geq (m_1 + m_2)\Phi$ will not. As before, we have

$$\Phi := \int_{-\infty}^{\infty} \phi(|x|)\,\mathrm{d}x \tag{4.17}$$

in the one-dimensional case.

For non-clustering interactions, the relative velocity evolves through an interaction to

$$u^+ = u^- - (m_1 + m_2)\Phi\,\mathrm{sgn}(u^-), \tag{4.18}$$

the equivalent of (2.76). Combining this with the conservation of momentum, (4.16b), this gives us the definition of the forwards velocity map $(v_1^*, v_2^*) = \psi_{m_1,m_2}(v_1, v_2)$ as

$$v_1^* = v_1 + m_2\Phi\,\mathrm{sgn}(v_2 - v_1), \quad v_2^* = v_2 - m_1\Phi\,\mathrm{sgn}(v_2 - v_1), \tag{4.19}$$

for incoming velocities where $|v_2 - v_1| \geq (m_1 + m_2)\Phi$. The reverse velocity map, $({}^*v_1, {}^*v_2) = \psi^{-1}_{m_1,m_2}(v_1, v_2)$, can also be derived as

$${}^*v_1 = v_1 - m_2\Phi\,\mathrm{sgn}(v_2 - v_1), \quad {}^*v_2 = v_2 + m_1\Phi\,\mathrm{sgn}(v_2 - v_1). \tag{4.20}$$

From these we can calculate $\left|D\psi^{-1}_{m_1,m_2}\right| = 1$.

For clustering interactions, where $|u^-| < (m_1+m_2)\Phi$, the inner region trajectory approaches a relative velocity of $u^+ = 0$. Treating the two clusters now as a combined cluster and using conservation of momentum and mass, we have the forwards maps, $m_1' = m_1 + m_2$ and

$$v_1' = \lambda_{m_1,m_2}(v_1, v_2) = \frac{m_1 v_1 + m_2 v_2}{m_1 + m_2}, \tag{4.21}$$

when $|v_2 - v_1| < (m_1 + m_2)\Phi$. As in (3.24) define $\hat{\lambda}_{m_1,m_2,v_2}(v_1)$ to be the partial evaluation of

$\lambda_{m_1,m_2}(v_1, v_2)$ at a specific value of v_2. We can then define the inverse maps ${}'m_1 = m_1 - m_2$ and

$$ {}'v_1 = \hat{\lambda}^{-1}_{m_1-m_2,m_2,v_2}(v_1) = \frac{m_1 v_1 - m_2 v_2}{m_1 - m_2}. \tag{4.22} $$

and from these, we can calculate

$$ \left| D\hat{\lambda}^{-1}_{m_1-m_2,m_2,v_2} \right| = \frac{m_1}{m_1 - m_2}. \tag{4.23} $$

Numerical model

Adding these maps to (4.5), we get our mass-binned discrete velocity model for the Cucker-Smale interaction in one dimension:

$$
\begin{aligned}
&|\Omega| \partial_t P_{i,a} \\
&\quad = -(\mu_K - 1) \sum_{\substack{b=1 \\ \bar{m}_a + \bar{m}_b \leq N - \mu_K + 2}}^{M} \sum_{j=1}^{V} |\bar{v}_j - \bar{v}_i| P_{i,a} P_{j,b} \\
&\qquad + (\mu_K - 1) \sum_{\substack{b=1 \\ \bar{m}_a + \bar{m}_b \leq N - \mu_K + 2}}^{M} \sum_{\substack{j,k=1 \\ (\bar{v}_j, \bar{v}_k) \in \mathbb{R} \setminus C_{\bar{m}_a, \bar{m}_b}}}^{V} |\bar{v}_k - \bar{v}_j| \, \gamma_{\bar{v}_i}(\bar{v}_k + \bar{m}_j \Phi \operatorname{sgn}(\bar{v}_k - \bar{v}_j)) P_{k,a} P_{j,b} \\
&\qquad + \frac{\mu_K}{2} \sum_{\substack{b,c=1 \\ \bar{m}_b + \bar{m}_c \leq \bar{m}_{a+1}}}^{M} \sum_{\substack{j,k=1 \\ (\bar{v}_j, \bar{v}_k) \in C_{\bar{m}_b, \bar{m}_c}}}^{V} |\bar{v}_k - \bar{v}_j| \, \gamma_{\bar{v}_i} \left(\frac{\bar{m}_k \bar{v}_k + \bar{m}_j \bar{v}_j}{\bar{m}_c + \bar{m}_b} \right) \gamma_{\bar{m}_a}(\bar{m}_c + \bar{m}_b) P_{k,c} P_{j,b} \\
&\qquad + \left(\frac{\mu_K}{2} - 1 \right) P_{i,a} \sum_{\substack{b,c=1 \\ \bar{m}_b + \bar{m}_c \\ \leq N - \bar{m}_a - \mu_K + 3}}^{M} \sum_{\substack{j,k=1 \\ (\bar{v}_j, \bar{v}_k) \in C_{\bar{m}_b, \bar{m}_c}}}^{V} |\bar{v}_k - \bar{v}_j| \, P_{k,c} P_{j,b}.
\end{aligned}
\tag{4.24}
$$

As both the clustering and non-clustering interactions in this model cause the velocities to move towards the mean velocity, the variance of the velocity distribution (and hence the system energy) will reduce towards 0 over time, increasing the error caused by our velocity discretisation. This problem is similar to that seen in inelastic Boltzmann models, for example, in the granular materials literature (see [109] and [29] for summaries). Some solutions involve modifying the system to include an anti-drift term or adding a heat bath such that the velocity distribution reaches a steady state with non-zero variance. However, we do not consider these here as we wish to model the Cucker-Smale interaction without modification. One option that maintains the unmodified interaction is to use a similarity scaling of the velocity variables in time. We will consider this in more detail for the higher dimensional model in Chapter 5.

For the one-dimensional case considered here, we use an adaptive velocity grid, halving the grid size Δv and the support of the discrete velocity distribution when there is negligible probability mass remaining in the outer quarters. We assume we are modelling a system with a mean velocity of 0 and have defined a grid symmetrical about 0 with the number of grid points V divisible by 4. Given this setup, we define new grid points, bin edges and corresponding grid

size as

$$\bar{v}_i' = \frac{\bar{v}_i}{2},\ v'^{-} = \frac{v^-}{2} \text{ and } \Delta v' = \frac{\Delta v}{2},$$

with the change occurring whenever

$$\frac{1}{\Delta v}\left(\sum_{i=1}^{V/4} + \sum_{i=3V/4+1}^{V}\right)\sum_{a=1}^{M} P_{i,a}(t) < 10^{-6}. \tag{4.25}$$

During these transitions, we define the new binned probabilities to split the probability of the previous bin they now overlap with, i.e.:

$$P_{i,a}' = \begin{cases} \frac{P_{V/4+(i+1)/2,a}}{2} & i \text{ odd} \\ \frac{P_{V/4+i/2,a}}{2} & i \text{ even.} \end{cases}$$

4.2. Finite-volume mean-field implementation

The mean-field model for a particle system with short-range interactions in the form of (3.63) is given by (2.10), which in one dimension, is

$$\begin{aligned}&\partial_t p(t;x_1;v_1) + \boldsymbol{v}_1 \partial_{\boldsymbol{x}_1} p(t;x_1;v_1) \\ &\qquad = -\frac{N-1}{\varepsilon}\partial_{\boldsymbol{v}_1}\left(p(t;x_1;v_1)\int_{\Omega\times\mathbb{R}} H\left(\frac{x_1-x_2}{\varepsilon}, v_1 - v_2\right) p(t;x_2;v_2)\,\mathrm{d}x_2\mathrm{d}v_2\right).\end{aligned} \tag{4.26}$$

Note that we have used p here in place of p_1^N for consistency with the final reduced probability obtained in the clustering short-range PDE model. For the Cucker-Smale model specifically, this becomes

$$\begin{aligned}&\partial_t p(t;x_1;v_1) + v_1 \partial_{x_1} p(t;x_1;v_1) \\ &\qquad = \frac{N-1}{\varepsilon}\partial_{v_1}\left(p(t;x_1;v_1)\int_{\Omega\times\mathbb{R}} \phi\left(\frac{|x_1-x_2|}{\varepsilon}\right)(v_1 - v_2) p(t;x_2;v_2)\,\mathrm{d}x_2\mathrm{d}v_2\right).\end{aligned} \tag{4.27}$$

Assuming spatial homogeneity, we have that

$$p(t;v_1) \coloneqq \int_\Omega p(t;x_1;v_1)\,\mathrm{d}x_1 = |\Omega| p(t;x_1;v_1), \tag{4.28}$$

which, upon substitution, gives

$$|\Omega|\partial_t p(t;v_1) = \frac{N-1}{\varepsilon}\int_\Omega \phi\left(\frac{|x_1-x_2|}{\varepsilon}\right)\mathrm{d}x_2\,\partial_{v_1}\left(p(t;v_1)\int_{\mathbb{R}}(v_1-v_2)\,p(t;v_2)\,\mathrm{d}v_2\right). \tag{4.29}$$

Changing variables to $\tilde{x}_{21} = (x_2 - x_1)/\varepsilon$ and assuming ε is small relative to Ω we have

$$\int_\Omega \phi\left(\frac{|x_1-x_2|}{\varepsilon}\right)\mathrm{d}x_2 = \varepsilon\int_{-\infty}^{\infty}\phi\left(|\tilde{x}_{21}|\right)\mathrm{d}\tilde{x}_{21} = \varepsilon\Phi,$$

where Φ is defined in (4.17). Substituting this and noting that p is normalised such that $\int_{\mathbb{R}} p(t; v_2)\,\mathrm{d}v_2 = 1$ we get:

$$\frac{|\Omega|}{\Phi(N-1)}\partial_t p(t; v_1) = \partial_{v_1}\left(v_1\, p(t; v_1) - p(t; v_1)\int_{\mathbb{R}} v_2\, p(t; v_2)\,\mathrm{d}v_2\right). \tag{4.30}$$

To derive the finite-volume numerical scheme, again take a uniform grid in velocity with spacing Δv and bin centres $\{v_i,\ i \in 1, \ldots, V\}$. Labelling the boundaries of the i^{th} bin as v_i^- and v_i^+ we can define the mean probability density in bin i as

$$\mathcal{P}_i(t) \coloneqq \frac{1}{\Delta v}\int_{v_i^-}^{v_i^+} p(t; v)\,\mathrm{d}v. \tag{4.31}$$

Note that this is a probability density as opposed to the binned probabilities defined for the implementation of the kinetic model in (4.4), as it is more conventional to use this definition for the finite volume scheme here. Integrating Equation (4.30) over bin i and separating the v_2 integral into bins we get

$$\frac{|\Omega|\Delta v}{\Phi(N-1)}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_i(t) = v_i^+ p(t; v_i^+) - v_i^- p(t; v_i^-) - (p(t; v_i^+) - p(t; v_i^-))\sum_{j=1}^{V} v_j \mathcal{P}_j \Delta v. \tag{4.32}$$

To define the bin-boundary probabilities ($p(t; v_i^+)$ and $p(t; v_i^-)$) in terms of the set of $\mathcal{P}_i(t)$ we use the upwind scheme:

$$\begin{aligned} p(t; v_i^+) \to \mathcal{P}_{i+1} \text{ and } p(t; v_i^-) \to \mathcal{P}_i \text{ for } v_i > \sum_{j=1}^{V} v_j \mathcal{P}_j \Delta v \\ p(t; v_i^+) \to \mathcal{P}_i \text{ and } p(t; v_i^-) \to \mathcal{P}_{i-1} \text{ for } v_i < \sum_{j=1}^{V} v_j \mathcal{P}_j \Delta v, \end{aligned} \tag{4.33}$$

which is based on the fact that, for this system, probability mass moves inwards towards the mean velocity. Implementing this scheme gives the full finite-volume scheme used for the mean-field model:

$$\frac{|\Omega|\Delta v}{\Phi(N-1)}\frac{\mathrm{d}}{\mathrm{d}t}\mathcal{P}_i = \begin{cases} v_i^+ \mathcal{P}_{i+1} - v_i^- \mathcal{P}_i - (\mathcal{P}_{i+1} - \mathcal{P}_i)\sum_{j=1}^{V} v_j \mathcal{P}_j \Delta v, & v_i > \sum_{j=1}^{V} v_j \mathcal{P}_j \Delta v \\ v_i^+ \mathcal{P}_i - v_i^- \mathcal{P}_{i-1} - (\mathcal{P}_i - \mathcal{P}_{i-1})\sum_{j=1}^{V} v_j \mathcal{P}_j \Delta v, & v_i < \sum_{j=1}^{V} v_j \mathcal{P}_j \Delta v. \end{cases} \tag{4.34}$$

4.3. Comparison between models

4.3.1. Particle ODE simulations

Full particle simulations using the individual-based ODEs (e.g. (4.14) for the Cucker-Smale model) are implemented for use as a base case when evaluating the accuracy of the PDE models. The positions $X_i \in \Omega$ and velocities $V_i \in \mathbb{R}$ of each particle in the system are evolved over time without any formal clustering introduced, i.e. groups of correlated particles are not reduced to a cluster, and weakly clustered particles could be broken off by interactions with

other particles. Results are computed using the Monte Carlo method, with initial velocities and positions of particles sampled from their respective distributions before the ODE is evolved deterministically. Output quantities and confidence intervals are computed from the aggregate set of multiple simulations at each time point. For the experiments in this chapter, the spatial domain Ω is the interval $[0, |\Omega|]$ with periodic boundary conditions.

To compute the outputs related to cluster distributions, empirical clusters were labelled using the `DBSCAN` (density-based spatial clustering of applications with noise) algorithm, as it best matches the mechanism by which we expect clusters to be formed under a short-ranged interaction. This is implemented in *Julia* using the *Clustering.jl* package. With the `DBSCAN` definition, two particles are considered to be in the same cluster if they are within a certain distance of each other, given a prescribed metric. The connected components of the system as a whole then determine the overall clusters. This means that not every pair of particles in a cluster has to be within the prescribed distance; they simply have to be linked by a chain of particles, each of which is sufficiently close to their neighbours.

To form the combined metric, we used the magnitude of the distance in X- and V-space separately. More specifically, two particles were considered to be in a cluster together if their velocities were within ε_v of each other and their positions were within ε_x of each other. For the higher dimensional particle ODE simulations performed in Chapter 7, this is generalised to separate 2-norms in $\boldsymbol{X}$ and $\boldsymbol{V}$ space, with tolerances $\varepsilon_{\boldsymbol{x}}$ and $\varepsilon_{\boldsymbol{v}}$. These tolerance values were tuned to produce reasonable clusters for each interaction, i.e. clusters should include all particles moving in a highly correlated group, but shouldn't include groups that later separate without the influence of a third group. Generally, for the Cucker-Smale interaction, the specific value of ε_x was selected based on the interaction kernel. For example, $\varepsilon_x = 5$ corresponds to where the exponential kernel has approximate magnitude 0.01. The V-space tolerance was then tuned to satisfy the aforementioned qualitative criteria. Simulations with values of ε_v that were too small were seen to temporarily label particles as many small clusters when two larger clusters combine, which is potentially undesirable as the fundamental process is simply the joining of the two larger groups. Values of ε_v that were too large could sometimes tag small groups of particles as one cluster when they are still moving relatively fast enough to separate.

Cluster velocity and mass distributions were computed by binning each cluster's mass and mean velocity (as labelled by DBSCAN) across the set of simulations onto histograms. These are normalised by the total number of clusters across all simulations rather than normalised in each simulation and added to avoid simulations with fewer clusters excessively influencing the distribution. For example, the density of the cluster mass distribution between m_1 and m_2 could be approximated as

$$\int_{m_1}^{m_2} p(t; m)\,\mathrm{d}m \approx \left(\sum_{i=1}^{N_s} \sum_{j=1}^{k_i(t)} \mathbb{1}(m_1 < M_{j,i}(t) < m_2) \right) \Big/ \left(\sum_{i=1}^{N_s} k_i(t) \right),$$

where N_s is the number of simulations, $k_i(t)$ is the number of clusters in simulation i at time t, $M_{j,i}(t)$ is the mass of cluster j in simulation i at time t and $\mathbb{1}$ is an indicator function, equal to

1 if the condition is true and 0 otherwise. The mean number of clusters is approximated as

$$\mu_K(t) \approx \frac{1}{N_s} \sum_{i=1}^{N_s} k_i(t),$$

and the mean kinetic energy per particle is approximated as

$$E(t) \approx \frac{1}{2NN_s} \sum_{i=1}^{N_s} \sum_{j=1}^{N} |V_{i,j}|^2,$$

where $v_{i,j}$ is the velocity of particle i in simulation j.

4.3.2. Exact clustering model

As full particle ODE simulations can be computationally expensive, we introduce a simplified model, which treats interactions as discrete events occurring in zero time whenever two clusters are at the same location. Specifically, the model tracks the positions $X_i \in \Omega = \mathbb{T}^1$, velocities $V_i \in \mathbb{R}$ and masses $M_i \in \{1, \dots, N\}$ of clusters in a one-dimensional system. Initial values are sampled from the respective initial distributions, and the system is evolved deterministically as for the particle ODE simulations. The relative velocities and positions of each pair of adjacent clusters determine which pair will be the next to reach the same location. At each step, the clusters are labelled 1 to k based on their ordering spatially (i.e. such that $X_j \geq X_i$ if $j > i$) and the time until each adjacent pair of particles would interact, given no other particles in the system, is computed as

$$\tau_{i,i+1} := \begin{cases} \frac{X_{i+1} - X_i}{V_i - V_{i+1}}, & i \in \{1, \dots, k-1\} \\ \frac{X_1 + |\Omega| - X_k}{V_k - V_1}, & i = k. \end{cases}$$

A negative value of $\tau_{i,i+1}$ means particles i and $i+1$ are moving away from each other and won't be the first to interact (they would necessarily have to interact with other particles first before meeting in the other direction as the system is one-dimensional). The minimum value of $\tau_{i,i+1}$ greater than or equal to 0 (but ignoring the pair of clusters that has just interacted if they haven't clustered together) is selected, as this means clusters i and $i+1$ will be the next to interact, after time $\tau_{i,i+1}$. Cluster positions are evolved as $X_j + V_j \tau_{i,i+1}$ for $j \in \{1, \dots, k\}$ and the time is evolved to $t + \tau_{i,i+1}$. The new velocities and masses for the interacting particles are computed for the Cucker-Smale model as in (4.19) and (4.21):

$$(V_i, M_i), (V_{i+1}, M_{i+1}) \to \begin{cases} (V_i + M_{i+1}\Phi \operatorname{sgn}(V_{i+1} - V_i), M_i), & |V_{i+1} - V_i| \geq (M_i + M_{i+1})\Phi \\ \quad (V_{i+1} - M_i \Phi \operatorname{sgn}(V_{i+1} - V_i), M_{i+1}), & \\ \left(\frac{M_i V_i + M_{i+1} V_{i+1}}{M_i + M_{i+1}}, M_i + M_{i+1} \right), (-,-), & |V_{i+1} - V_i| < (M_i + M_{i+1})\Phi, \end{cases} \tag{4.35}$$

where the dashes indicate that in a clustering interaction, one of the clusters is removed from the system, and we take $k \to k-1$. The clusters are then relabelled based on their new spatial

ordering, and the process is repeated for the next interaction.

This model can be considered equivalent to full particle ODE simulations in the limit that the interaction range ε goes to 0. Any discrepancies between these two models will be due to effects that are assumed small for small ε. For instance, in the particle simulations, clusters have non-zero size extent, interactions are not strictly binary between clusters, and higher-order interactions are possible between multiple clusters simultaneously.

4.3.3. Stochastic model

As a further approximation, we then consider a stochastic model that only tracks the velocities, V_i, and masses, M_i, of clusters, assuming a spatially uniform distribution of clusters and not tracking their positions directly. Given the one-dimensional periodic domain $\Omega = \mathbb{T}^1$, the relative distance a pair of clusters (assuming they have no spatial extent themselves) would have to travel before they interact is between 0 and $|\Omega|$. As such, the expected time until a pair of clusters with velocities V_i and V_j interact, assuming a uniform spatial distribution and ignoring other clusters in the system, is

$$\bar{\tau}_{i,j} = \frac{|\Omega|}{2|V_i - V_j|}.$$

To model this system efficiently using the Gillespie stochastic simulation algorithm, we treat the interactions between pairs of clusters as independent events and model the time until they occur as exponentially distributed with rate parameters $\alpha_{i,j} = 1/\bar{\tau}_{i,j}$. Due to the independence of these interaction events and properties of the exponential distribution, the time until any interaction happens is also exponentially distributed, with rate parameter

$$\alpha = \sum_{\substack{i,j=1 \\ j<i}}^{k} \alpha_{i,j}.$$

At each time step in the Gillespie algorithm, we sample from this exponential distribution to determine the time until the next interaction occurs and select which pair of clusters it was by sampling over the relative proportion of the total rate they each represent $\alpha_{i,j}/\alpha$. We then evolve the masses and velocities of the clusters as in (4.35), which may involve removing one cluster from the system if two clusters join to form a larger cluster. The interaction rates are then recomputed using the new velocities post-interaction.

The removal of the spatial dynamics of the interaction rates represents the central additional approximation made in this stochastic model above the exact clustering model. In a spatial model, the interaction rates are not truly independent and, for example, clusters that have just interacted are in the same spatial vicinity and might be more likely to interact again, and a pair of clusters must necessarily interact before time $|\Omega|/(|V_i - V_j|)$ had elapsed. In this stochastic model, the rates are reset after each interaction, and the independence of the interaction rates encompasses a similar approximation to the *Stosszahlansatz* made in the PDE models.

The output quantities for the exact clustering and stochastic models are computed similarly

to the particle ODE simulations, except that now we explicitly have the results in terms of clusters rather than particles tagged as belonging to the same cluster. The mean kinetic energy per particle is approximated for these models as the total kinetic energy of the system divided by the number of particles, i.e.

$$E(t) \approx \frac{1}{2NN_s} \sum_{i=1}^{N_s} \sum_{j=1}^{k_i(t)} M_{i,j} |V_{i,j}|^2,$$

where $M_{i,j}$ and $V_{i,j}$ represent the mass and velocity of cluster j in simulation i, respectively.

4.4. Implementation

To perform the comparison, these models were implemented in *Julia*, which is suggested to improve computational efficiency over interpreted languages common for similar applications such as *Python* [8]. We have not implemented alternatives in Python for comparison, but we did notice significant efficiency gains when the code was optimised for type-stability to assist the compiler and minimise heap allocations (up to 10× speed improvement). The experiments were performed on a machine running Linux Ubuntu with 6-core Intel® Core™ i5-8500 CPU@2.10GHz and 16GB RAM. The *Parameters.jl* and *DrWatson.jl* packages [36] are used for parameter control and project organisation.

Time integration for the short-range PDE, mean-field model and particle ODE simulations was performed using ODE solvers from the *DifferentialEquations.jl* package [93]. Specifically, we tested several solver algorithms on the particle ODE simulations, including:

- `BS5` - A Bogaki-Shampine $5/4^{th}$ order Runge-Kutta method, suggested to have more robust error control than `Tsit5`, an efficient equivalent to MATLAB's `ODE45`.
- `Vern7` - A $7/6^{th}$ order Runge-Kutta method due to Verner, suggested as an efficient replacement for MATLAB's `ODE113`.
- `OwrenZen5` - An Owren-Zennaro $5/4^{th}$ order Runge-Kutta method with optimised interpolation behaviour.
- `Rodas4` - A 4^{th} order Rosenbrock method with 3^{rd} order interpolant, stable when used on stiff equations.
- `DPRKN6` - A 6^{th} order Runge-Kutta-Nyström method, used with an optimised solver framework for second-order dynamical systems.
- `Euler` - A standard first-order forward Euler method considered for comparison.

Apart from the standard `Euler` method, all of these algorithms have adaptive time-step control to meet pre-specified integration tolerances. Note that for the particle ODE simulations, we manually bounded the adaptive time-step such that particles moving with the current maximum relative velocity would move at most $\varepsilon/5$ relative to each other in one step. This was because

we noted that if left unchecked, the adaptive algorithms would increase the time-step such that particles would jump through their short-range interaction window and miss interactions.

The first three algorithms listed generally proved to be the most computationally efficient when simulating the Cucker-Smale collective behaviour system. The `Euler` algorithm naturally required significantly longer computation time to approach similar levels of accuracy, and there did not appear to be a need for a stiff solver like `Rodas4`. Several of the algorithms, but most notably the `DPRKN6` within the `DynamicalODEProblem` second order framework, exhibited spurious fluctuations in particle velocities when clustering was occurring at weaker integration tolerances. In the Cucker-Smale model, the velocities of two particles within interaction range should smoothly approach each other. The observed behaviour where velocities briefly rose sharply near interactions could potentially be explained by an overshoot and correction process. As they are similar methods, it would be expected that `BS5`, `Vern7`, and `OwrenZen5` were similarly efficient. `OwrenZen5` was selected due to its optimised interpolation behaviour as the default choice for the following experiments, as we wished to generate uniform time output for analysis and visualisation.

Time integration for the short-range PDE model and the finite-volume mean-field model was also performed using `OwrenZen5`. The interaction allocation proportions $\gamma_{\bar{v}_i}(\bar{v}_k^*)$, $\gamma_{\bar{v}_i}(\bar{v}_k')$ and $\gamma_{\bar{m}_a}(\bar{m}_c')$, were pre-calculated for each discrete binned pair of interacting particles. Specifically, for all $(\bar{v}_k;\bar{m}_c)\times(\bar{v}_j;\bar{m}_b)$ with $k,j\in\{1,\dots,V\}$ and $c,b\in\{1,\dots,M\}$, the output velocities and masses, $(\bar{v}_k^*;\bar{m}_c)$ and $(\bar{v}_j^*;\bar{m}_b)$ if non-clustering, or $(\bar{v}_k',\bar{m}_c')$ if clustering, were calculated for the Cucker-Smale interaction. For each output, e.g. $(\bar{v}_k',\bar{m}_c')$, i and a were determined such that $\bar{v}_i\leq\bar{v}_k'<\bar{v}_{i+1}$ and $\bar{m}_a\leq\bar{m}_c'<\bar{m}_{a+1}$ and the allocation proportions $\gamma_{\bar{v}_i}(\bar{v}_k')$ and $\gamma_{\bar{m}_a}(\bar{m}_c')$ calculated. Note that $\gamma_{\bar{v}_{i+1}}(\bar{v}_k')$ and $\gamma_{\bar{m}_{a+1}}(\bar{m}_c')$ can be calculated cheaply at runtime and do not need to be stored. This significantly improved the computational efficiency of the implementation at the cost of increased memory usage. The memory available generally limited how fine the velocity grid could be taken.

The exact clustering and stochastic simulation models are both event-based and were implemented directly in Julia.

4.5. PDE model evaluation

To evaluate the accuracy of the short-range PDE model and mean-field model against the set of simulation models we perform numerical experiments using the one-dimensional Cucker-Smale model for N particles with exponential interaction kernel, $\phi(X)=e^{-X}$:

$$\begin{aligned}\dot{X}_i &= V_i,\\ \dot{V}_i &= \frac{1}{\varepsilon}\sum_{\substack{j=1\\ j\neq i}}^{N}\exp\left(-\frac{|X_i-X_j|}{\varepsilon}\right)(V_j-V_i).\end{aligned}\tag{4.36}$$

We take a periodic spatial domain with relative size $|\Omega|/\varepsilon=400$ and initial conditions: uniform in space across the domain Ω, uniform in velocity over the domain $[-20,20]$ and monodisperse

in mass (i.e. all particles start separately and not in clusters).

The exponential kernel is chosen as it is short-ranged and gives a finite value for Φ, namely

$$\Phi = \int_{-\infty}^{\infty} e^{-|\tilde{x}_{21}|} \, \mathrm{d}\tilde{x}_{21} = 2,$$

which allows non-clustering interactions to occur in one-dimension. Power-law interaction kernels are singular at $|\tilde{x}_{21}| = 0$ and hence give infinite values for Φ, meaning no matter how large the relative velocity of two particles, all interactions would cause clustering.

Inspection of specific particle ODE simulations suggested that for the exponential interaction kernel DBSCAN clustering tolerances of $\varepsilon_{\boldsymbol{x}} = 5\varepsilon$ and $\varepsilon_{\boldsymbol{v}} = 0.1$ label clusters in a way that qualitatively matches what we expect. Preliminary tests of the short-range PDE model for the Cucker-Smale model up to $N = 1250$ particles suggested that mass-binning with $M = 25$ bins was sufficient to make the binning error small relative to the model error. Similarly, an initial velocity grid spacing of $\Delta v = 0.2$ (with grid halving when necessary as discussed at the end of §4.1.2) appeared sufficiently precise while still allowing for computation within 16GB of RAM. The mean-field model was computed with velocity grid spacing of $\Delta v = 0.01$ as the velocity distribution produced became significantly sharper than the short-range PDE model at later times, necessitating a finer grid spacing. This finer grid was possible computationally as the mean-field model does not track mass and is cheaper to compute.

Figure 4.1 shows the mean kinetic energy per particle, E, for each of the models considered across a range of particle numbers from $N = 2$ to $N = 1250$. The top plot shows E against scaled time t', which is defined as

$$t' = \frac{(N-1)\Phi}{|\Omega|} t,$$

which allows the comparison of models for different values of N. For clarity, the stochastic and exact clustering models are not shown. The bottom plots show snapshots of all the models at specific values of t' across different values of N. Figure 4.2 shows the velocity distributions produced from the models with $N = 1250$ particles at specific values of t'.

The values of the mean kinetic energy from the short-range PDE model (labelled as 'PDE' in the figure labels) appear to track those from the particle ODE simulations well for early times but consistently underestimate the particle simulations at later times. Corresponding behaviour is seen in the velocity distributions where the PDE model approximates the distribution from the particle ODEs well at first, capturing the tail behaviour as the velocities reduce towards 0. However, at later times, the PDE velocity distribution can be seen to be much sharper than the particle distribution, underestimating the velocities and the kinetic energy. There are several possible approximations that the PDE model contains that could be contributing to this discrepancy. First, the particles may become increasingly correlated over time, rendering the *Stosszahlansatz* less valid. This effect may be particularly prevalent in the one-dimensional system we are modelling, as there will be significant re-interaction of particles that have already interacted and become correlated. Second, the *Clusterzahlansatz* may be breaking down if the variation in conditional velocity distributions for different k becomes too wide. Third, there

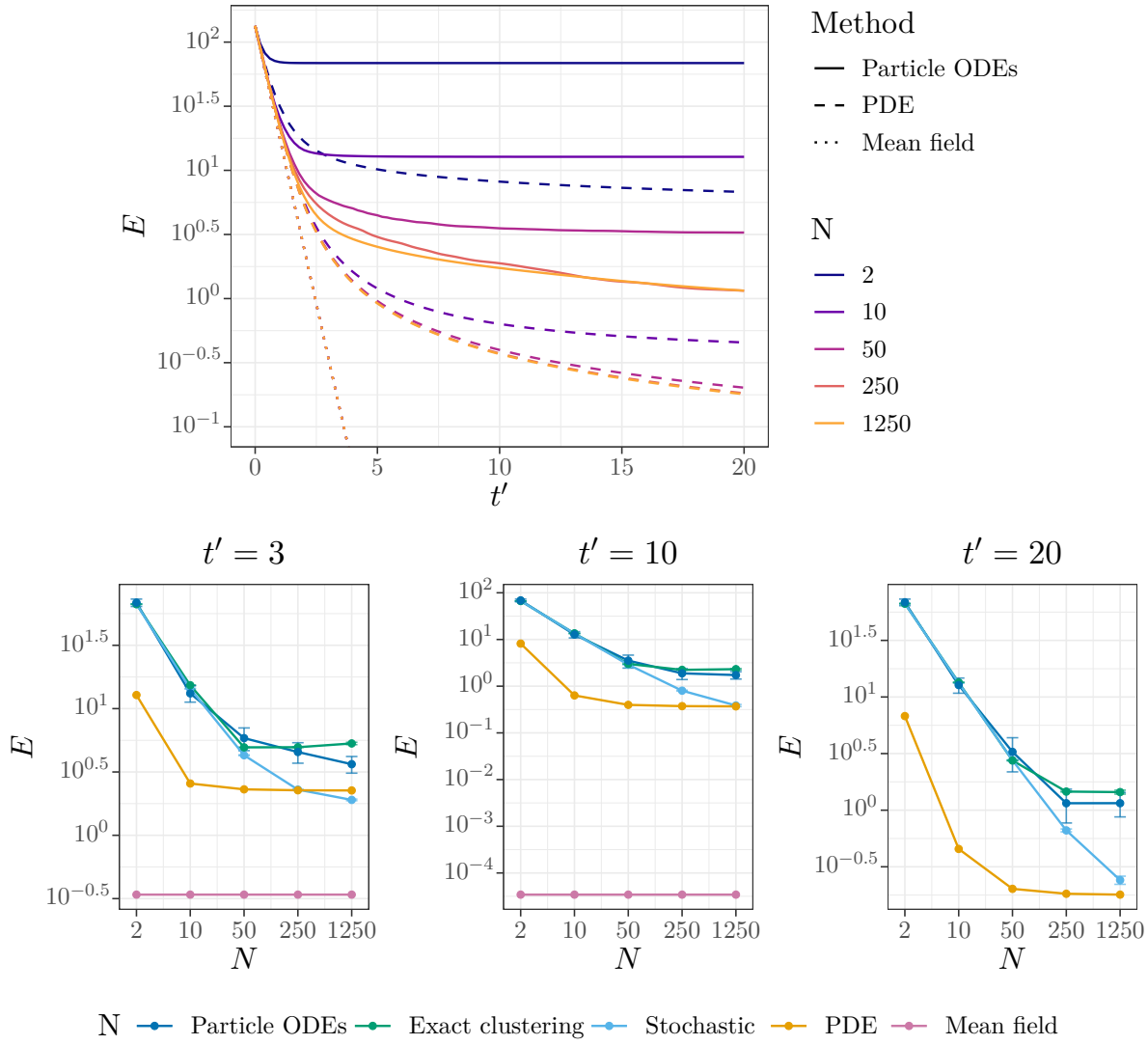


Figure 4.1.: Comparison of mean kinetic energy per particle, E, across different numbers of initial particles N. Simulations are performed with a uniform initial velocity condition with maximum velocity 20, an exponential interaction kernel and domain size $|\Omega| = 400\varepsilon$. The Finite-Volume PDE model is implemented with a maximum of 25 linear-geometric mass grid points and an initial velocity discretisation of $\Delta v = 0.2$ with grid halving implemented. The mean field values are the same for all N and are not shown on the $t' = 20$ plot. For clarity, the stochastic and exact clustering models are not shown on the full time plot.

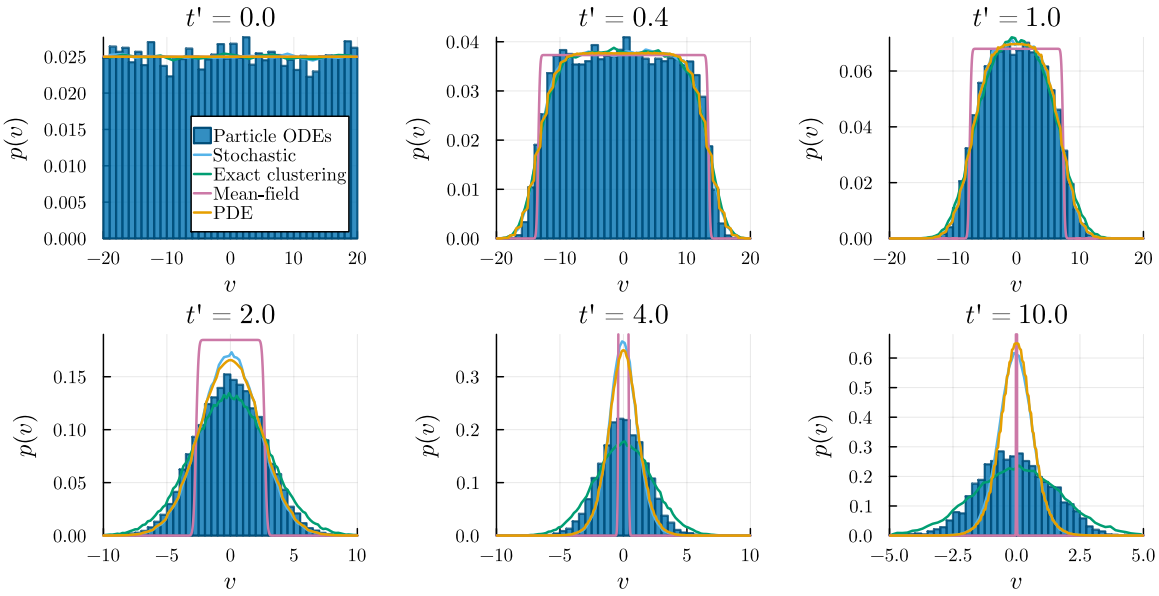


Figure 4.2.: Cluster velocity distributions for a system with $N = 1250$ particles, uniform initial velocity condition, exponential interaction kernel and domain size $|\Omega| = 400\varepsilon$. The full animated version can be found at https://figshare.com/s/422b260c1b8f1bce3ca0.

could be errors from approximating the clusters as points if ε is not small enough. Finally, the PDE model allows particles to interact with a distribution of other potential particle velocities (that would only be present in theory across a suite of simulations). At later times in any individual particle ODE simulation, the particles all move in one cluster at the same velocity. When considered as a distribution sampled over the uniform initial velocity condition, however, this final cluster velocity is given by the mean velocity of the sample and would follow a normal distribution due to the central limit theorem. At such a time, the PDE model may continue to allow interactions, reducing the velocities toward 0 as it allows particles to interact with the distribution of other velocities. However, the velocities in the particle simulations individually may have stopped evolving. We will evaluate the effect of these possible error sources in more detail throughout the remainder of this chapter.

The mean-field model accurately captures the initial slope of the decrease in E. However, it continues to decrease at this rate when the rate of decrease in the short-range PDE and particle models reduces, and is a much worse approximation than the short-range PDE model at later times. This may be explained as the mean-field model does not have any mechanism for capturing clustering behaviour, which would reduce the rate at which interactions occur and, hence, reduce the rate of energy loss from the system. Corresponding results are seen in the velocity distributions, where the mean-field model captures the rate at which the distribution tails move inwards but misses the shape that the other models capture. At late times, the mean-field velocity distribution becomes extremely sharp as there is no clustering mechanism to slow the energy loss and approach toward the mean velocity.

The exact clustering model matches the particle ODE simulations well, both in mean

kinetic energy and velocity distribution for the lower values of N. The discrepancy at higher values of N is expected as the potential size of the clusters increases with the number of particles (see Figure 4.4 for example size distributions). With more particles forming each cluster, they will behave less like point objects in the particle ODE simulations and the discrepancy with the exact clustering model will increase. However, this discrepancy is relatively small compared to that observed for the other models.

The results from the stochastic model also match those from the particle simulations and exact clustering model for lower values of N, but the velocity distribution variance and mean kinetic energy are underestimated for larger values of N. This suggests that the local spatial correlations have more of an effect when there are more particles in the system. It is also possible that this is an effect of the increased spatial density of the particles; we have increased the number of particles but kept the size of the spatial domain the same. In Chapter 7, we will consider varying the density and number of particles independently in the higher dimensional model, but this remains an avenue for further investigation in the one-dimensional case. Results from the stochastic and short-range PDE models become more consistent as N is increased. This is expected as the effect of the PDE model averaging over the distribution of possible realisations simultaneously should be reduced when there are more particles in any one simulation.

Figure 4.3 shows the modelled predictions of the mean number of clusters, μ_K, across different values of N and cluster number distributions for $N = 50$. The short-range PDE and stochastic models both overestimate the number of clusters remaining in the system for all values of N, with the PDE model overestimating by a greater extent. This is further evidence that discrepancies are caused by local spatial correlations missed in both these models and the averaging across realisations performed by the PDE model. These models allow the mean kinetic energy to reduce too far by allowing non-clustering interactions to occur at a greater rate than they would in any one instance of a spatial particle simulation, where more clustering has occurred and the particles have less relative velocity. The cluster number distributions do not appear to span the full range of k (and even with $N = 1250$, the variance remains relatively small, see Figure B.2 in the appendix), which supports the use of the *Clusterzahlansatz*.

Figure 4.4 shows the corresponding mass-weighted cluster mass distributions for the system with $N = 1250$ particles. At later times, once the mean cluster mass has increased, the distributions from the stochastic and PDE models have greater numbers of smaller clusters, consistent with the cluster number distributions and μ_K approximations. The exact clustering model matches the cluster mass distribution of the particle ODE simulations well.

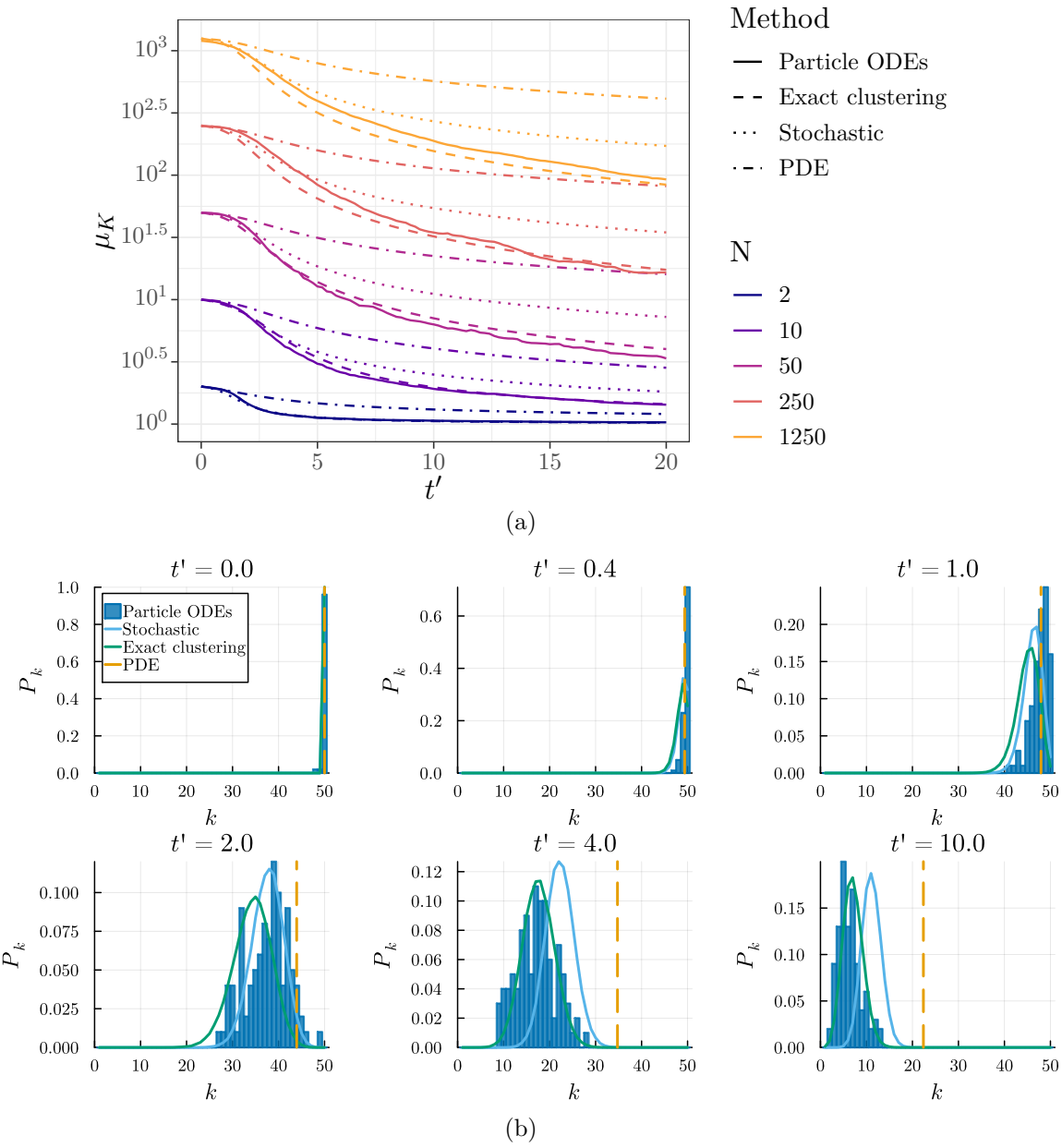


Figure 4.3.: (a) Mean cluster number comparison across the number of initial particles, N, and (b) Cluster number distribution for the system with $N = 50$. The parameters are as in Figure 4.1. The mean-field model does not track cluster numbers and is not shown, while the PDE model only tracks the mean number of clusters (shown as a vertical line in (b)) after k is marginalised out using the *Clusterzahlansatz*. An animated version of (b) can be found at https://figshare.com/s/ec4ec3b427db7caeadfc.

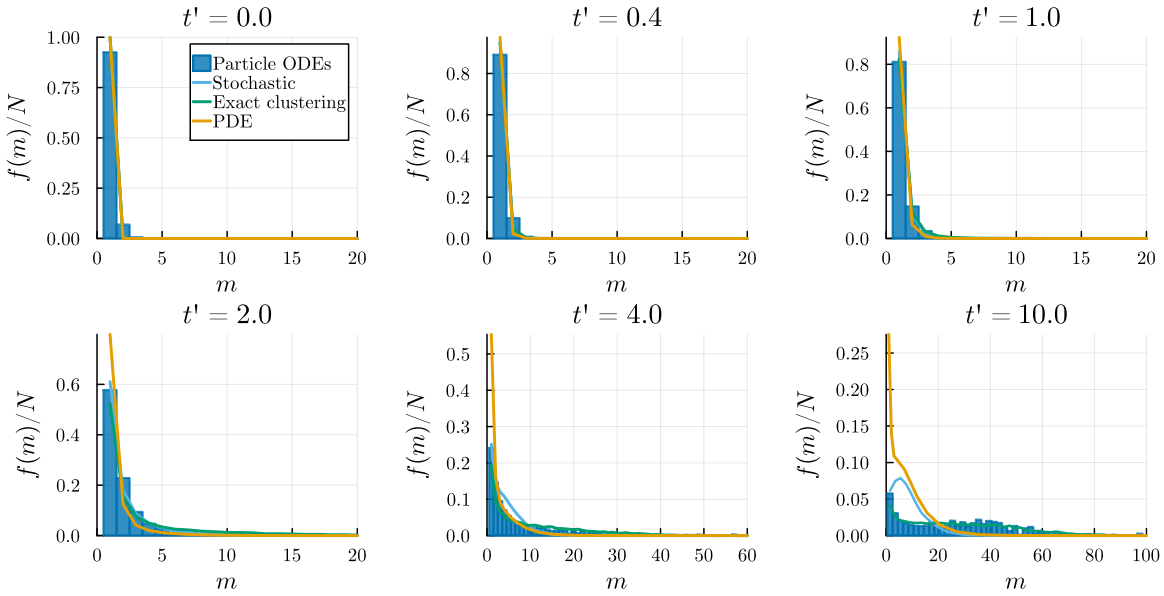


Figure 4.4.: Mass-weighted cluster mass distributions for a system with $N = 1250$ particles, uniform initial velocity condition, exponential interaction kernel and domain size $|\Omega| = 400\varepsilon$. An animated version can be found at https://figshare.com/s/422b260c1b8f1bce3ca0.

4.5.1. Mean-centred initial condition

Individual particle ODE, exact clustering and stochastic simulations have different mean velocities, while the short-range PDE and mean-field models average across simulations, resulting in a mean velocity of 0. To analyse this discrepancy, we now consider subtracting the mean velocity from each simulation to match the mean of the PDE-based models. For a simulation of particle ODE, exact clustering and stochastic models, we sample N velocities from a uniform distribution with support $[-20, 20]$ before subtracting the mean sample velocity. That is, the initial particle velocities $V_i(0)$ are computed as

$$V_i(0) = \hat{V}_i - \sum_{i=1}^{N} \hat{V}_i, \tag{4.37}$$

where $\hat{V}_i$ are samples from the uniform distribution. There is no closed-form expression for the effective distribution from which these mean-centred velocities could have been drawn. Instead, for the short-range PDE and mean-field models, we draw a large number of samples using the process in (4.37) and bin these on a histogram to use as the initial velocity distribution: $p(t = 0; v; m = 1)$. Note that this distribution has a wider support than $[-20, 20]$ in general, as velocities near the edge of the domain can be moved further from 0 by subtracting the sample mean if the mean has the opposite sign.

Figure 4.5 shows the mean kinetic energy per particle for all the models against scaled time, t', for different numbers of particles, N, with this mean-centred initial velocity distribution. Figure B.3 in the appendix shows a comparison of velocity distributions at a selection of times

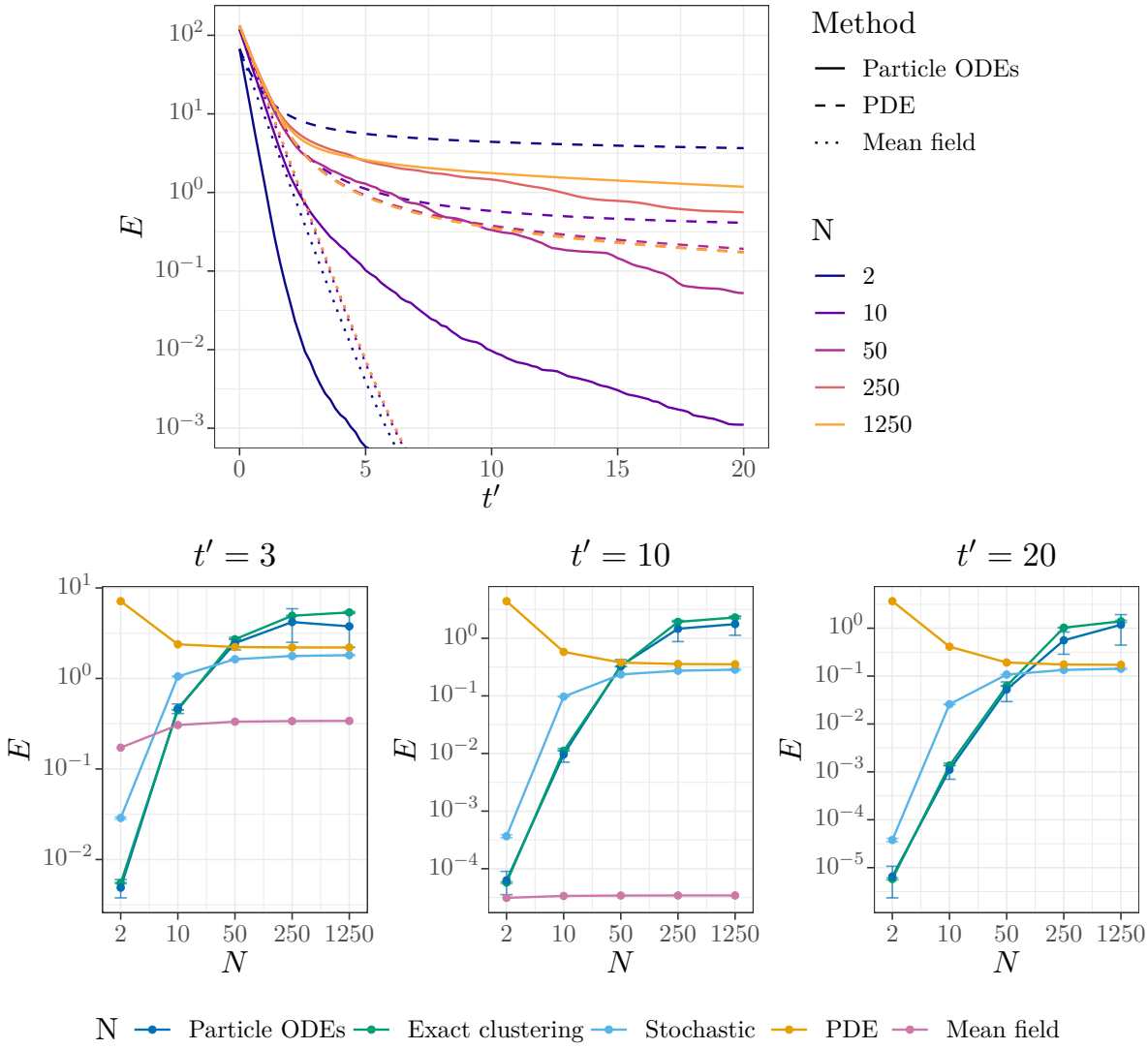


Figure 4.5.: Comparison of mean kinetic energy per particle, E, across different numbers of initial particles, N. Simulations are performed with initial velocities sampled from a uniform distribution on $[-20, 20]$ before subtracting each sample's mean velocity. The short-range PDE and mean-field models use an equivalent initial distribution empirically derived from samples of this mean-centred initial condition. All models use an exponential interaction kernel and domain size $|\Omega| = 400\varepsilon$. The short-range PDE model is implemented with a maximum of 25 linear-geometric mass grid points and an initial velocity discretisation of $\Delta v = 0.1$ with grid halving implemented. The mean-field results are similar across all values of N and are not shown on the $t' = 20$ plot for clarity as they are too small. For clarity, the stochastic and exact clustering models are not shown in the plot against time.

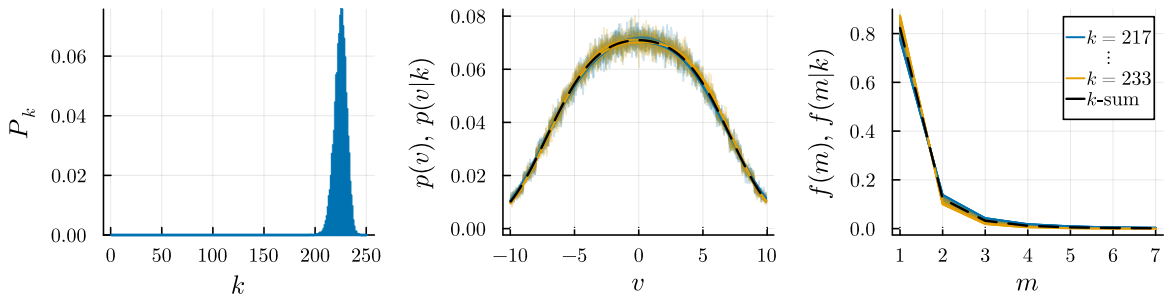
for the two initial conditions. It can be seen that the discrepancies between the models follow a similar pattern as for the tests without the mean-centred distributions. The exact clustering model is still a good fit for the particle ODEs, and the short-range PDE model shows better qualitative behaviour than the mean-field model over time but still has significant discrepancies with results from the particle ODE models. The main difference, however, is that results from the short-range PDE model and the stochastic model converge for much smaller values of N than when the unmodified uniform velocity distribution was used. This is consistent with what we expect, as by matching the mean velocities of all individual simulations, they should all be approaching the same velocity distribution at later times. As such, the fact that the PDE model averages across a distribution of multiple realisations should cause less discrepancy. However, the fact that these two models still show significantly different results than the exact clustering and particle ODE simulations suggests other error sources are still important, particularly the local spatial correlations and independence that distinguish the exact clustering and stochastic models. The cluster number and mass distributions only vary slightly when the mean-centred initial condition is used and are not shown here.

4.5.2. *Clusterzahlansatz* evaluation

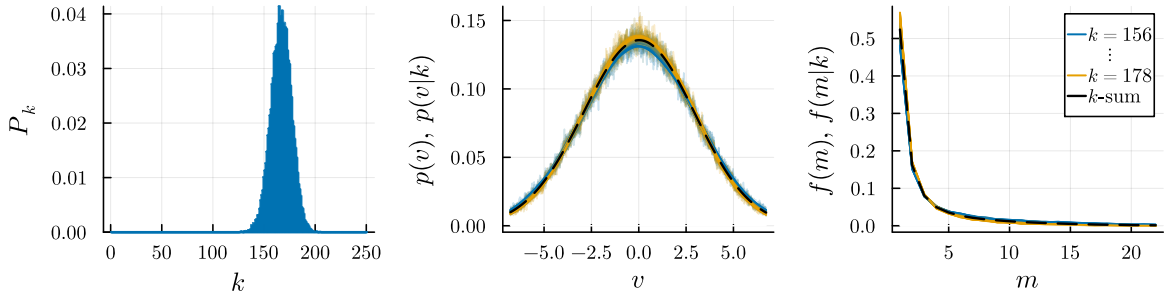
When taking the *Clusterzahlansatz*, we approximate the conditional distribution of velocity and mass, given a certain number of clusters in the system, with the summed distribution where the number of clusters has been marginalised, i.e. we assume $p(t; v; m \mid k) \approx p(t; v; m)$. To test the accuracy of this assumption, we consider the k-conditional velocity and mass distributions from the exact clustering model and compare them to the k-marginalised distribution. Plots of these distributions for $N = 250$ particles are shown in Figure 4.6. The exact clustering model was chosen instead of the full particle ODE simulations as it is much less expensive to compute, and a very large number of simulations are required to produce reasonable velocity and mass distributions for each value of k.

We first note that the cluster number distributions are quite low-variance compared to the potential range of k from 1 to N. Only a small range of cluster numbers have significant probability at any particular time t'. This supports the *Clusterzahlansatz* as it implies only a limited range of k is relevant at any particular time. The k-marginalised cluster velocity distributions appear to be a good approximation to the k-conditional velocity distributions at all times for the relevant values of k. However, the k-conditional cluster mass distributions show some discrepancy with the k-marginalised distribution at later times. This is to be expected as, given a particular number of clusters k, the mean cluster mass should be N/k. Thus, for later times when k is small compared to N, even a small variation in k will cause a significant variation in N/k.

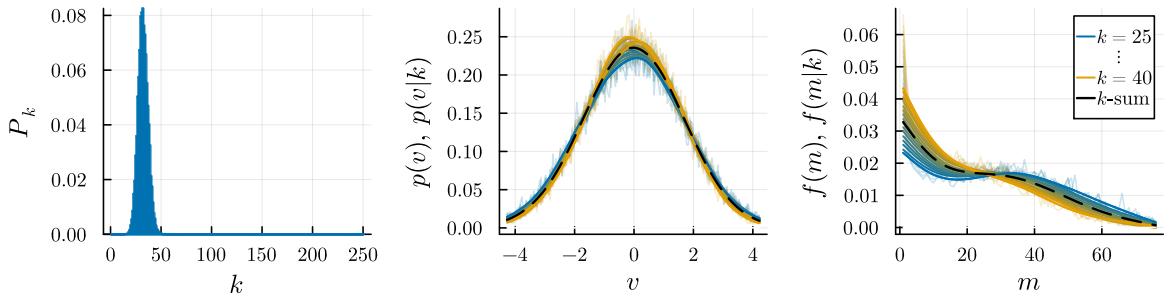
As such, this evidence suggests that, for the Cucker-Smale model considered, the *Clusterzahlansatz* is a reasonable approximation except for later times when the number of clusters is small compared to the total number of particles. Combined with the comparison of model discrepancies between experiments with uniform initial velocity distribution and mean-centred



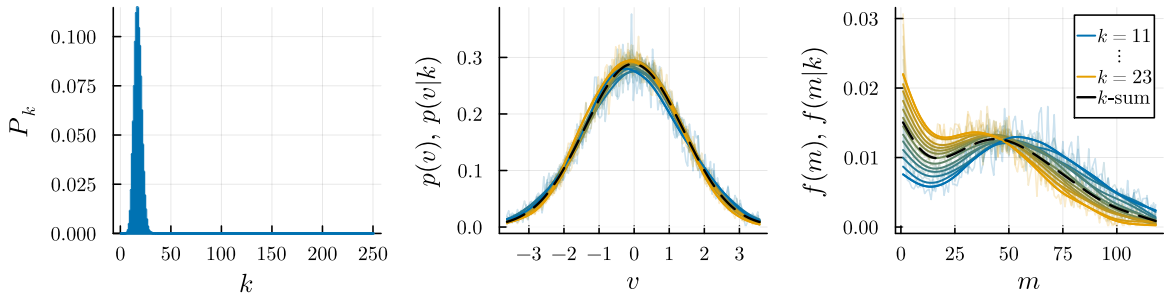
(a) $t' = 1$



(b) $t' = 2$



(c) $t' = 10$



(d) $t' = 20$

Figure 4.6.: Plots of the cluster number distribution, velocity distribution, and mass-weighted cluster mass distribution for the exact clustering model of the Cucker-Smale interaction with exponential interaction kernel and $N = 50$ particles. Distributions with specific cluster numbers k with $P_k \geq 0.02$ are shown in colour in the background with cubic smoothing spline fits overlaid when necessary to approximate the distributions that the simulations are converging to. The distributions with k marginalised (that would be used to approximate the k-specific distributions if the *Clusterzahlansatz* $p(t; v; m \,|\, k) \approx p(t; v; m)$ is taken) are shown as black dashed lines. The approximation in velocity space appears reasonable, but the mass approximation is less accurate at later times, as would be expected from the different expected mean cluster masses: N/k.

initial velocity distribution, this suggests that the primary source of error between the short-range PDE models and the particle ODE simulations is the local spatial particle correlations and *Stosszahlansatz*. As we would reasonably expect these correlations to be less significant in higher dimensions where particles have more freedom to separate after interacting, we now turn our attention to deriving a short-range PDE model in higher dimensions.

5. Clustering short-range PDE model in higher dimensions

We will take the base model for a particle system with short-range interactions and clustering in higher dimensions to be the d-dimensional equivalent of the system (3.63), depicted in Figure 3.5a. Specifically, we consider N indistinguishable particles, defined for $t \geq 0$, by positions $\boldsymbol{X}_i(t) \in \Omega \subset \mathbb{R}^d$ and velocities $\boldsymbol{V}_i(t) \in \mathbb{R}^d$ for $i \in \{1, \dots, N\}$ that evolve through the individual-based ODE model

$$\begin{aligned} \dot{\boldsymbol{X}}_i(t) &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i(t) &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{N} \boldsymbol{H}\left(\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon}, \boldsymbol{V}_i - \boldsymbol{V}_j\right), \end{aligned} \tag{5.1}$$

where the binary interaction term, $\boldsymbol{H}$, is assumed to be short-ranged, satisfying

$$\boldsymbol{H}(\boldsymbol{X}, \boldsymbol{V}) = O(|\boldsymbol{X}|^{-1-\delta}) \text{ as } |\boldsymbol{X}| \to \infty. \tag{5.2}$$

We assume periodic boundary conditions on a square domain, i.e. $\Omega \equiv \mathbb{T}^d$.

5.1. Strong form kinetic PDE derivation

As for the one-dimensional case in §3.4, to derive our kinetic clustering model, we first define an approximate ODE model using the point-cluster approximation. That is, we treat particles that are highly correlated after an interaction as a single cluster with an internal property, mass, representing the number of particles that it comprises. For a specific number of clusters remaining (i.e. $K(t) = k$) the system is characterised by cluster positions $\boldsymbol{X}_i(t) \in \Omega \equiv \mathbb{T}^d$, velocities $\boldsymbol{V}_i(t) \in \mathbb{R}^d$ and masses $M_i(t) \in \{1, \dots, N\}$, which evolve through

$$\begin{aligned} \dot{\boldsymbol{X}}_i(t) &= \boldsymbol{V}_i, \\ \dot{\boldsymbol{V}}_i(t) &= \frac{1}{\varepsilon} \sum_{\substack{j=1 \\ j \neq i}}^{k} M_j \boldsymbol{H}\left(\frac{\boldsymbol{X}_i - \boldsymbol{X}_j}{\varepsilon}, \boldsymbol{V}_i - \boldsymbol{V}_j\right), \end{aligned} \tag{5.3}$$

for $i \in \{1, \dots, k\}$. System (5.3) is the d-dimensional equivalent of (3.66). A hierarchy of these models is defined, one for each $k \in \{1, \dots, N\}$. As in the one-dimensional case (see Figure 3.5b), if two clusters of masses M_i and M_j are highly correlated after interacting, they are replaced by a combined cluster with mass $M_i + M_j$ and the number of clusters in the system

reduces to $k-1$. As this implies cluster mass is conserved, it follows we have $\sum_{i=1}^{k} M_i(t) = N$.

Although the ODE model (5.3) is deterministic, as in §2.1.2, we consider the initial space and velocity values of each particle to be independent and identically distributed samples from chosen probability distributions. Similarly to the one-dimensional case, we define a hierarchy of k-cluster probability distribution functions $p^k(t; \vec{\boldsymbol{x}}^k; \vec{\boldsymbol{v}}^k; \vec{m}^k)$ with $\boldsymbol{x}_i \in \Omega \equiv \mathbb{T}^d$, $\boldsymbol{v}_i \in \mathbb{R}^d$, and $m_i \in \{1, \dots, N\}$ for $i \in \{1, \dots, k\}$ and $k \in \{1, \dots, N\}$, by averaging over these initial conditions. These distributions are invariant under permutations of indices (i.e. relabelling clusters), and p^k is set to 0 for unphysical mass combinations. We take the monodisperse initial condition for mass, i.e. there are N clusters of mass 1 at $t = 0$.

Considering the conservation of probability in an infinitesimal box in phase space, we have the hierarchy of time evolution equations

$$\partial_t p^k + \sum_{i=1}^{k} \boldsymbol{v}_i \cdot \nabla_{\boldsymbol{x}_i} p^k + \frac{1}{\varepsilon} \sum_{\substack{i,j=1 \\ j \neq i}}^{k} \nabla_{\boldsymbol{v}_i} \cdot \left(m_j \boldsymbol{H}_{ij} p^k \right) = \sum_{i=1}^{k} q^k_{(i)}, \tag{5.4}$$

where

$$\boldsymbol{H}_{ij} := \boldsymbol{H}\left(\frac{\boldsymbol{x}_i - \boldsymbol{x}_j}{\varepsilon}, \boldsymbol{v}_i - \boldsymbol{v}_j \right),$$

represents the short-range interaction effect on clusters i due to j and $q^k_{(i)}$ is the source term representing clusters produced at $(\boldsymbol{x}_i; \boldsymbol{v}_i; m_i)$ due to clustering in the $(k+1)$-cluster equation.

5.1.1. Marginalisation

Again defining the cluster phase $\boldsymbol{z}_i \equiv (\boldsymbol{x}_i; \boldsymbol{v}_i; m_i)$, and the integration region

$$\int_{\boldsymbol{D}} \mathrm{d}\boldsymbol{z}_i := \sum_{m_i=1}^{N} \int_{\mathbb{R}^d} \int_{\Omega} \mathrm{d}\boldsymbol{x}_i \, \mathrm{d}\boldsymbol{v}_i,$$

the single-cluster probability density is defined as

$$p^k_1(t; \boldsymbol{z}_1) = \int_{\boldsymbol{D}^{(k-1)}} p^k(t; \boldsymbol{z}_1, \dots, \boldsymbol{z}_k) \, \mathrm{d}\boldsymbol{z}_2 \dots \mathrm{d}\boldsymbol{z}_k.$$

Integrating (5.4) over $\boldsymbol{z}_2$ to $\boldsymbol{z}_k$ and relabelling clusters using indistinguishability we get

$$\begin{aligned} \partial_t p^k_1 + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} p^k_1 = & -\frac{k-1}{\varepsilon} \int_{\boldsymbol{D}} \nabla_{\boldsymbol{v}_1} \cdot \left(m_2 \boldsymbol{H}_{12} p^k_2 \right) + \nabla_{\boldsymbol{v}_2} \cdot \left(m_1 \boldsymbol{H}_{21} p^k_2 \right) \mathrm{d}\boldsymbol{z}_2 \\ & - \frac{(k-1)(k-2)}{2\varepsilon} \int_{\boldsymbol{D}^2} \nabla_{\boldsymbol{v}_2} \cdot \left(m_3 \boldsymbol{H}_{23} p^k_3 \right) + \nabla_{\boldsymbol{v}_3} \cdot \left(m_2 \boldsymbol{H}_{32} p^k_3 \right) \mathrm{d}\boldsymbol{z}_2 \mathrm{d}\boldsymbol{z}_3 \\ & + \int_{\boldsymbol{D}^{(k-1)}} q^k_{(1)} \, \mathrm{d}\boldsymbol{z}_2 \dots \mathrm{d}\boldsymbol{z}_k \\ & + (k-1) \int_{\boldsymbol{D}^{(k-1)}} q^k_{(2)} \, \mathrm{d}\boldsymbol{z}_2 \dots \mathrm{d}\boldsymbol{z}_k. \end{aligned} \tag{5.5}$$

Following a similar argument to §3.3.1, assume we have a functional $\mathcal{C}_{\boldsymbol{z}_i,(j,s)}\left[p^{k+1}(t; \vec{\boldsymbol{z}}^{k+1}) \right]$ that represents the clustering flux of an interaction between particles with labels j and s (in a

system with $k+1$ clusters) that produces a cluster at $\boldsymbol{z}_i$. It then follows that,

$$q^k_{(i)}(t;\vec{\boldsymbol{z}}^k) = \sum_{\substack{j,s=1\\ j\neq s}}^{k+1} \mathcal{C}_{\boldsymbol{z}_i,(j,s)}\left[p^{k+1}(t;\vec{\boldsymbol{z}}^{k+1})\right].$$

Given outgoing phase $\boldsymbol{z}_i$ and one incoming phase $\boldsymbol{z}_s$, the other incoming phase $\boldsymbol{z}_j$ is defined through $\boldsymbol{z}_j = {}'\boldsymbol{z}_i \equiv (\boldsymbol{x}_i; {}'\boldsymbol{v}_i; m_i - m_s)$, with ${}'v_i = \hat{\lambda}^{-1}_{m_i-m_s,m_s,\boldsymbol{v}_s}(\boldsymbol{v}_i)$. The functional $\mathcal{C}_{\boldsymbol{z}_i,(j,s)}$ will be represented as an integral over $\boldsymbol{z}_s$ as there is a range of incoming velocities that could map to outgoing velocity $\boldsymbol{v}_i$ in a clustering interaction.

A simple way to ensure we preserve the permutation invariance of the probability densities is to allocate the clustering flux equally from clustering particles to output labels. Assuming we have permutation invariance in p^{k+1}, there are $(k+1)k/2$ choices for the clustering pair (j,s), each of which will produce equivalent flux. The factor of $1/2$ is present as $\boldsymbol{z}_s$ is a dummy variable to be integrated over and $\boldsymbol{z}_j$ is defined with respect to $\boldsymbol{z}_i$ and $\boldsymbol{z}_s$, so $\mathcal{C}_{\boldsymbol{z}_i,(j,s)}$ is indistinguishable from $\mathcal{C}_{\boldsymbol{z}_i,(s,j)}$. Without loss of generality, we choose $j = i$ and $s = k+1$. There are k possible output labels to allocate over, so if we allocate an equal fraction of $1/k$ of the flux to each, it follows that

$$q^k_{(i)}(t;\vec{\boldsymbol{z}}^k) = \frac{(k+1)}{2}\mathcal{C}_{\boldsymbol{z}_i,(i,k+1)}\left[p^{k+1}(t;\vec{\boldsymbol{z}}^{k+1})\right].$$

From here, we will use the slight abuse of notation $\mathcal{C}_{(i,k+1)} \equiv \mathcal{C}_{\boldsymbol{z}_i,(i,k+1)}$ to denote the clustering flux produced at $\boldsymbol{z}_i$ from pre-phases ${}'\boldsymbol{z}_i$ and $\boldsymbol{z}_{k+1}$. As $\mathcal{C}_{(i,k+1)}$ will only involve an integral over $\boldsymbol{z}_{k+1}$, we can marginalise the other phases when substituting to (5.5) giving

$$\begin{aligned}
\partial_t p^k_1 + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}p^k_1 = &-\frac{k-1}{\varepsilon}\int_{\boldsymbol{D}}\nabla_{\boldsymbol{v}_1}\cdot\left(m_2\boldsymbol{H}_{12}p^k_2\right) + \nabla_{\boldsymbol{v}_2}\cdot\left(m_1\boldsymbol{H}_{21}p^k_2\right)\mathrm{d}\boldsymbol{z}_2\\
&-\frac{(k-1)(k-2)}{2\varepsilon}\int_{\boldsymbol{D}^2}\nabla_{\boldsymbol{v}_2}\cdot\left(m_3\boldsymbol{H}_{23}p^k_3\right) + \nabla_{\boldsymbol{v}_3}\cdot\left(m_2\boldsymbol{H}_{32}p^k_3\right)\mathrm{d}\boldsymbol{z}_2\mathrm{d}\boldsymbol{z}_3\\
&+\frac{k+1}{2}\mathcal{C}_{(1,k+1)}\left[p^{k+1}_2(t;{}'\boldsymbol{z}_1,\boldsymbol{z}_{k+1})\right]\\
&+\frac{(k-1)(k+1)}{2}\int_{\boldsymbol{D}}\mathcal{C}_{(2,k+1)}\left[p^{k+1}_3(t;\boldsymbol{z}_1,{}'\boldsymbol{z}_2,\boldsymbol{z}_{k+1})\right]\mathrm{d}\boldsymbol{z}_2.
\end{aligned} \tag{5.6}$$

As $\boldsymbol{z}_{k+1}$ is a dummy variable, we can replace $\mathcal{C}_{(1,k+1)}\left[p^{k+1}_2(t;{}'\boldsymbol{z}_1,\boldsymbol{z}_{k+1})\right]$ with $\mathcal{C}_{(1,2)}\left[p^{k+1}_2(t;{}'\boldsymbol{z}_1,\boldsymbol{z}_2)\right]$ in the third RHS term and make an equivalent replacement but with $\mathcal{C}_{(2,3)}$ in the final RHS term.

5.1.2. Matched asymptotic expansion

As in the one-dimensional case, we consider the inner region where $|\boldsymbol{x}_1 - \boldsymbol{x}_2| \sim \varepsilon$, representing the region where clusters 1 and 2 are within interaction range. Note that the specific selection of clusters 1 and 2 is without loss of generality as the clusters are indistinguishable under relabelling. We define inner variables $\tilde{\boldsymbol{x}}_1 := \boldsymbol{x}_1$ and $\tilde{\boldsymbol{x}}_{21} := (\boldsymbol{x}_2 - \boldsymbol{x}_1)/\varepsilon$, and the corresponding

inner probability density as $\tilde{p}^k_{(1,2)}(t;\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_{21},\boldsymbol{x}_3,\dots,\boldsymbol{x}_k;\vec{\boldsymbol{v}}^k;\vec{m}^k)$. Considering (5.4) in terms of these inner variables, the leading-order equation is

$$(\boldsymbol{v}_2-\boldsymbol{v}_1)\cdot\nabla_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}^k_{(1,2)}+\nabla_{\boldsymbol{v}_1}\cdot\left(m_2\boldsymbol{H}\left(-\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_1-\boldsymbol{v}_2\right)\tilde{p}^k_{(1,2)}\right)\\ +\nabla_{v_2}\cdot\left(m_1\boldsymbol{H}\left(\tilde{\boldsymbol{x}}_{21},\boldsymbol{v}_2-\boldsymbol{v}_1\right)\tilde{p}^k_{(1,2)}\right)=0. \quad (5.7)$$

Note that we could equivalently define the marginalised probability $\tilde{p}^k_{2(1,2)}(t;\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}_{21};\boldsymbol{v}_1,\boldsymbol{v}_2;m_1,m_2)$ in terms of these inner variables, and this would satisfy the same leading-order equation (5.7), as none of the marginalised variables are present. Substituting (5.7) to the first RHS term in (5.6), we get

$$\frac{(k-1)}{\varepsilon}\int_{\boldsymbol{D}}(\boldsymbol{v}_2-\boldsymbol{v}_1)\cdot\nabla_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}^k_{2(1,2)}\,\mathrm{d}\boldsymbol{z}_2=\varepsilon^{d-1}(k-1)\int_{\tilde{\boldsymbol{D}}}(\boldsymbol{v}_2-\boldsymbol{v}_1)\cdot\nabla_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}^k_{2(1,2)}\,\mathrm{d}\tilde{\boldsymbol{z}}_2,$$

where we have changed to inner variables and defined the shorthand $\tilde{\boldsymbol{z}}_2\equiv(\tilde{\boldsymbol{x}}_{21};\boldsymbol{v}_2;m_2)$, with integration region

$$\int_{\tilde{\boldsymbol{D}}}\mathrm{d}\tilde{\boldsymbol{z}}_i:=\sum_{m_i=1}^{N}\int_{\mathbb{R}^d}\int_{\mathbb{R}^d}\mathrm{d}\tilde{\boldsymbol{x}}_{i(i-1)}\,\mathrm{d}\boldsymbol{v}_i.$$

Performing the same manipulation on the second RHS term with an equivalent inner region between clusters 2 and 3, the time evolution equation becomes

$$\begin{aligned}\partial_t p^k_1+\boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}p^k_1=&\,\varepsilon^{d-1}(k-1)\int_{\tilde{\boldsymbol{D}}}(\boldsymbol{v}_2-\boldsymbol{v}_1)\cdot\nabla_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}^k_{2(1,2)}\,\mathrm{d}\tilde{\boldsymbol{z}}_2\\ &+\varepsilon^{d-1}\frac{(k-1)(k-2)}{2}\int_{\tilde{\boldsymbol{D}}}\int_{\boldsymbol{D}}(\boldsymbol{v}_3-\boldsymbol{v}_2)\cdot\nabla_{\tilde{\boldsymbol{x}}_{32}}\tilde{p}^k_{3(2,3)}\,\mathrm{d}\boldsymbol{z}_2\mathrm{d}\tilde{\boldsymbol{z}}_3\\ &+\frac{k+1}{2}\mathcal{C}_{(1,2)}\left[p^{k+1}_2(t;{}'\boldsymbol{z}_1,\boldsymbol{z}_2)\right]\\ &+\frac{(k-1)(k+1)}{2}\int_{\boldsymbol{D}}\mathcal{C}_{(2,3)}\left[p^{k+1}_3(t;\boldsymbol{z}_1,{}'\boldsymbol{z}_2,\boldsymbol{z}_3)\right]\mathrm{d}\boldsymbol{z}_2.\end{aligned} \quad (5.8)$$

Using the divergence theorem and writing out the integration variables fully, the first RHS term in (5.8) becomes

$$\varepsilon^{d-1}(k-1)\lim_{R\to\infty}\sum_{m_2=1}^{N}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-1}_R}\tilde{p}^k_{2(1,2)}\left(\boldsymbol{v}_2-\boldsymbol{v}_1\right)\cdot\mathrm{d}\boldsymbol{S}_{\tilde{\boldsymbol{x}}_{21}}\,\mathrm{d}\boldsymbol{v}_2, \quad (5.9)$$

where $\boldsymbol{S}_{\tilde{\boldsymbol{x}}_{21}}$ is a surface element of the sphere $\mathbb{S}^{d-1}_R$. Following the same $\tilde{\boldsymbol{x}}_{21}$ integral manipulations as from (2.16) to (2.18), (5.9) becomes

$$\varepsilon^{d-1}(k-1)\sum_{m_2=1}^{N}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-2}}\int_0^\infty\Big[\left|\boldsymbol{v}_2-\boldsymbol{v}_1\right|\tilde{p}^k_{2(1,2)}(t;\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}^+_{21};\boldsymbol{v}_1,\boldsymbol{v}_2)\\ -\left|\boldsymbol{v}_2-\boldsymbol{v}_1\right|\tilde{p}^k_{2(1,2)}(t;\tilde{\boldsymbol{x}}_1,\tilde{\boldsymbol{x}}^-_{21};\boldsymbol{v}_1,\boldsymbol{v}_2)\Big]r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}^{d-2}\,\mathrm{d}\boldsymbol{v}_2, \quad (5.10)$$

where again r is the impact parameter, $\boldsymbol{\xi}^{d-2}$ is a surface element of the sphere $\mathbb{S}^{d-2}$, and $\tilde{\boldsymbol{x}}^+_{21}$ and $\tilde{\boldsymbol{x}}^-_{21}$ are the outgoing and incoming trajectories corresponding to the same point on the

interaction cross-section, respectively (see Figure 2.1).

As in the one-dimensional case, in the outer region, clusters are independent to leading order, meaning matching this inner solution to the outer region (following (2.19)) would be equivalent to assuming the two clusters in question are uncorrelated. Again, making Boltzmann's *Stosszahlansatz*, we can only reasonably assume this for clusters entering the inner region and approaching an interaction, as those leaving the inner region will have just interacted and hence be correlated. We, thus, must map the outgoing trajectories back to their corresponding incoming trajectories so that we can match with the outer solution and produce a closed equation. Consider a trajectory that enters the inner region with coordinates $(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^-; \boldsymbol{v}_1^-, \boldsymbol{v}_2^-, m_1, m_2)$ and leaves with coordinates $(\tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^+; \boldsymbol{v}_1^+, \boldsymbol{v}_2^+; m_1, m_2)$. Following a similar analysis of the characteristics of the inner equation (5.7) to §3.4, we have that

$$|\boldsymbol{v}_2^- - \boldsymbol{v}_1^-| \, {}^-p_{2(1,2)}^k(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^-; \boldsymbol{v}_1^-, \boldsymbol{v}_2^-, m_1, m_2) \left| D\psi_{m_1,m_2,r}^{-1} \right| \\ = |\boldsymbol{v}_2^+ - \boldsymbol{v}_1^+| \, {}^+p_{2(1,2)}^k(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}^+; \boldsymbol{v}_1^+, \boldsymbol{v}_2^+; m_1, m_2), \quad (5.11)$$

where we have defined the probability density at the incoming trajectory as ${}^-p_{2(1,2)}^k$ and the density at the outgoing trajectory as ${}^+p_{2(1,2)}^k$. In this higher-dimensional case, the map from outgoing to incoming velocities in the inner region is given by $\psi_{m_1,m_2,r}^{-1}(\boldsymbol{v}_1^+, \boldsymbol{v}_2^+) = (\boldsymbol{v}_1^-, \boldsymbol{v}_2^-)$, and depends on the interaction $\boldsymbol{H}$ (not shown explicitly), particles masses m_1 and m_2, and the impact parameter r. Note that we have implicitly used the Galilean invariance of the two-particle interactions such that the only extra parameter determining the interaction map in d dimensions is r. Inserting (5.11) into (5.10) and matching with the outer velocity variables defined there, we get

$$\varepsilon^{d-1}(k-1) \sum_{m_2=1}^{N} \int_{\mathbb{R}^d} \int_{\mathbb{S}^{d-2}} \int_0^\infty \Big[|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1| \, {}^-p_{2(1,2)}^k(t; \tilde{\boldsymbol{x}}_1, {}^*\tilde{\boldsymbol{x}}_{21}, {}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2, m_1, m_2) \left| D\psi_{m_1,m_2,r}^{-1} \right| \\ - |\boldsymbol{v}_2 - \boldsymbol{v}_1| \, {}^-p_{2(1,2)}^k(t; \tilde{\boldsymbol{x}}_1, \tilde{\boldsymbol{x}}_{21}, \boldsymbol{v}_1, \boldsymbol{v}_2, m_1, m_2) \Big] r^{d-2} \, \mathrm{d}r \, \mathrm{d}\boldsymbol{\xi}^{d-2} \, \mathrm{d}\boldsymbol{v}_2, \quad (5.12)$$

where $({}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2) = \psi_{m_1,m_2,r}^{-1}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ are the incoming velocities corresponding to outgoing velocities $(\boldsymbol{v}_1, \boldsymbol{v}_2)$, ${}^*\tilde{\boldsymbol{x}}_{21}$ is the incoming position corresponding to $\tilde{\boldsymbol{x}}_{21}$ and the $-$ superscript is left on ${}^-p_{2(1,2)}^k$ to indicate that both densities are now evaluated at incoming coordinates. Expanding the outer solution in inner variables, we have

$$p_2^k(t; \boldsymbol{z}_1, \boldsymbol{z}_2) = p_1^k(t; \boldsymbol{z}_1) p_1^k(t; \boldsymbol{z}_2) \delta_{12} + O(\varepsilon),$$

(see (2.19)) where the $\delta_{12} \equiv \delta(\boldsymbol{x}_1 - \boldsymbol{x}_2)$ indicates, to leading order, the particles can be assumed to have the same spatial location. As such, following (3.49), we can define the mass-restricted *Stosszahlansatz* in the d-dimensional case as

$${}^-p_2^k(t; \boldsymbol{z}_1, \boldsymbol{z}_2) \delta_{12} = \begin{cases} \dfrac{1}{P_k(t)} p_1^k(t; \boldsymbol{z}_1) p_1^k(t; \boldsymbol{z}_2) \delta_{12}, & m_1 + m_2 \leq N - k + 2 \\ 0, & \text{otherwise}, \end{cases} \quad (5.13)$$

where, as before, the mass restriction $m_1 + m_2 \leq N - m_1 - k + 2$ ensures conservation of mass can be satisfied. Substituting this assumption, (5.12) becomes

$$\varepsilon^{d-1}\frac{(k-1)}{P_k}\sum_{m_2=1}^{N-m_1-k+2}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-2}}\int_0^\infty\Big[|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1|\,p_1^k(t;{}^*\boldsymbol{z}_1)\,p_1^k(t;{}^*\boldsymbol{z}_2)\delta_{12}\left|D\psi^{-1}_{m_1,m_2,r}\right| \\ - |\boldsymbol{v}_2 - \boldsymbol{v}_1|\,p_1^k(t;\boldsymbol{z}_1)\,p_1^k(t;\boldsymbol{z}_2)\delta_{12}\Big]r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}^{d-2}\,\mathrm{d}\boldsymbol{v}_2, \quad (5.14)$$

where we have used the shorthand ${}^*\boldsymbol{z}_1 \equiv (\boldsymbol{x}_1;{}^*\boldsymbol{v}_1;m_1)$. As $\psi^{-1}_{m_1,m_2,r}$ is independent of $\boldsymbol{\xi}^{d-2}$, we can collapse the $\mathrm{d}\boldsymbol{\xi}^{d-2}$ integral, giving

$$\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|\frac{(k-1)}{P_k}\sum_{m_2=1}^{N-m_1-k+2}\int_{\mathbb{R}^d}\int_0^\infty\Big[|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1|\,p_1^k(t;{}^*\boldsymbol{z}_1)\,p_1^k(t;{}^*\boldsymbol{z}_2)\delta_{12}\left|D\psi^{-1}_{m_1,m_2,r}\right| \\ - |\boldsymbol{v}_2 - \boldsymbol{v}_1|\,p_1^k(t;\boldsymbol{z}_1)\,p_1^k(t;\boldsymbol{z}_2)\delta_{12}\Big]r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2. \quad (5.15)$$

Performing similar modifications to the second RHS term in (5.8), it becomes

$$\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|\frac{(k-1)(k-2)}{2(P_k)^2}p_1^k(t;\boldsymbol{z}_1)\sum_{\substack{m_2,m_3=1\\ m_2+m_3\leq\\ N-m_1-k+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^\infty\Big(|{}^*\boldsymbol{v}_3 - {}^*\boldsymbol{v}_2|\,p_1^k(t;{}^*\boldsymbol{z}_2) \\ \times p_1^k(t;{}^*\boldsymbol{z}_3)\delta_{23}\left|D\psi^{-1}_{m_2,m_3,r}\right| - |\boldsymbol{v}_3 - \boldsymbol{v}_2|\,p_1^k(t;\boldsymbol{z}_2)\,p_1^k(t;\boldsymbol{z}_3)\delta_{23}\Big)r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3. \quad (5.16)$$

In assuming ψ has a well-defined inverse, we are implicitly assuming the interaction $\boldsymbol{H}$ is such that for every outgoing set of velocities $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ and impact parameter r, we can define corresponding incoming velocities $({}^*\boldsymbol{v}_1,{}^*\boldsymbol{v}_2)$. This is the case for many commonly studied interactions, even those that cause clustering. However, as $\boldsymbol{H}$ is a clustering interaction, we cannot assume that every combination of incoming velocities $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ and impact parameter r has a corresponding set of outgoing velocities, as some trajectories will stay in the inner region (this is the reason for the final two terms in (5.8)). We will instead assume the interaction $\boldsymbol{H}$ is such that for a given set of incoming velocities $(\boldsymbol{v}_1, \boldsymbol{v}_2)$ and masses (m_1, m_2), there exists an impact parameter r^*, such that interactions with $r < r^*$ will be clustering interactions and those with $r > r^*$ will be non-clustering. We will write this bounding impact parameter as $r^*_{m_1,m_2}(|\boldsymbol{v}_2 - \boldsymbol{v}_1|)$ to show its dependence on the incoming velocities and masses.

If we consider the first term in the integrand of (5.16) and change integration variables from $(\boldsymbol{v}_2, \boldsymbol{v}_3)$ to $({}^*\boldsymbol{v}_2, {}^*\boldsymbol{v}_3)$ the integral becomes

$$\sum_{\substack{m_2,m_3=1\\ m_2+m_3\leq\\ N-m_1-k+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_{\mathbb{S}^{d-2}}\int_{r^*_{m_2,m_3}(|{}^*\boldsymbol{v}_3-{}^*\boldsymbol{v}_2|)}^\infty|{}^*\boldsymbol{v}_3 - {}^*\boldsymbol{v}_2|\,p_1^k(t;{}^*\boldsymbol{z}_2)\,p_1^k(t;{}^*\boldsymbol{z}_3) \\ \times\delta_{23}r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}^{d-2}\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}^*\boldsymbol{v}_2\mathrm{d}^*\boldsymbol{v}_3, \quad (5.17)$$

as the Jacobian is absorbed in the change of variables and only trajectories with impact

parameter $r > r^*_{m_2,m_3}(|{}^*\boldsymbol{v}_3 - {}^*\boldsymbol{v}_2|)$ will be reached by the pre-image. If r were smaller, the interaction would have caused clustering, and there would be no outgoing trajectory. If we change the dummy velocity variables to $(\boldsymbol{v}_2, \boldsymbol{v}_3)$, this will then cancel with part of the second integrand term in (5.16), meaning the full term then becomes

$$-\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|\frac{(k-1)(k-2)}{2(P_k)^2}p_1^k(t;\boldsymbol{z}_1)\sum_{\substack{m_2,m_3=1\\ m_2+m_3\leq\\ N-m_1-k+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{m_2,m_3}(|\boldsymbol{v}_3-\boldsymbol{v}_2|)}|\boldsymbol{v}_3-\boldsymbol{v}_2|$$
$$\times\, p_1^k(t;\boldsymbol{z}_2)\,p_1^k(t;\boldsymbol{z}_3)\delta_{23}r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3. \quad (5.18)$$

To ensure the probability flux that does not leave the inner region in a clustering interaction is conserved, we allocate flux from the clustering interactions in the $k+1$ equation as a source term in the k equation. Specifically we consider $\mathcal{C}_{(1,2)}\left[p_2^{k+1}(t;{}'\boldsymbol{z}_1,\boldsymbol{z}_2)\right]$, defined in §5.1.1 to be the source clustering flux produced at $\boldsymbol{z}_1 \equiv (\boldsymbol{x}_1;\boldsymbol{v}_1;m_1)$ in the k-equation from clusters with pre-phases at ${}'\boldsymbol{z}_1 \equiv (\boldsymbol{x}_1;{}'\boldsymbol{v}_1;m_1-m_2)$ and $\boldsymbol{z}_2 \equiv (\boldsymbol{x}_1;\boldsymbol{v}_2;m_2)$ in the $k+1$ equation. Define $\lambda_{m_1-m_2,m_2,r}({}'\boldsymbol{v}_1,\boldsymbol{v}_2)=\boldsymbol{v}_1$ to be the map from incoming to clustered velocities, with partial inverse $\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}(\boldsymbol{v}_1)={}'\boldsymbol{v}_1$ defined similarly to as in (3.24). Following a similar process to §3.4, and using similar d-dimensional integral modifications to those used in deriving (5.10), we have that

$$\mathcal{C}_{(1,2)}\left[p_2^{k+1}(t;{}'\boldsymbol{z}_1,\boldsymbol{z}_2)\right]=\frac{1}{P_{k+1}}\sum_{m_2=1}^{m_1-1}\int_{\mathbb{R}^d}\int_{\mathbb{S}^{d-2}}\int_0^{r^*_{m_1-m_2,m_2}(|{}'\boldsymbol{v}_1-\boldsymbol{v}_2|)}|{}'\boldsymbol{v}_1-\boldsymbol{v}_2|$$
$$\times\, p_1^{k+1}(t;{}'\boldsymbol{z}_1)\,p_1^{k+1}(t;\boldsymbol{z}_2)\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}\right|r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{\xi}^{d-2}\,\mathrm{d}\boldsymbol{v}_2, \quad (5.19)$$

where we have already mapped the incoming particle phases to outer variables using the *Stosszahlansatz*. The Jacobian, $D\hat{\lambda}^{-1}$, once again represents the change in the density of a bundle of trajectories when mapping from $\boldsymbol{v}_1$ to ${}'\boldsymbol{v}_1$. Assuming also that the clustering interaction is independent of orientation, we can collapse the $\mathrm{d}\boldsymbol{\xi}^{d-2}$ integral to give

$$\mathcal{C}_{(1,2)}\left[p_2^{k+1}(t;{}'\boldsymbol{z}_1,\boldsymbol{z}_2)\right]=\frac{\left|\mathbb{S}^{d-2}\right|}{P_{k+1}}\sum_{m_2=1}^{m_1-1}\int_{\mathbb{R}^d}\int_0^{r^*_{m_1-m_2,m_2}(|{}'\boldsymbol{v}_1-\boldsymbol{v}_2|)}|{}'\boldsymbol{v}_1-\boldsymbol{v}_2|$$
$$\times\, p_1^{k+1}(t;{}'\boldsymbol{z}_1)\,p_1^{k+1}(t;\boldsymbol{z}_2)\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}\right|r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2. \quad (5.20)$$

For the final term in (5.8), we then have that

$$\int_{\boldsymbol{D}}\mathcal{C}_{(2,3)}\left[p_3^{k+1}(t;\boldsymbol{z}_1,{}'\boldsymbol{z}_2,\boldsymbol{z}_3)\right]\mathrm{d}\boldsymbol{z}_2$$
$$=\frac{\left|\mathbb{S}^{d-2}\right|}{(P_{k+1})^2}p_1^{k+1}(t;\boldsymbol{z}_1)\sum_{m_2=1}^{N-m_1-k+1}\sum_{m_3=1}^{m_2-1}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{m_2-m_3,m_3}(|{}'\boldsymbol{v}_2-\boldsymbol{v}_3|)}|{}'\boldsymbol{v}_2-\boldsymbol{v}_3|$$
$$\times\, p_1^{k+1}(t;{}'\boldsymbol{z}_2)\,p_1^{k+1}(t;\boldsymbol{z}_3)\left|D\hat{\lambda}^{-1}_{m_2-m_3,m_3,\boldsymbol{v}_3,r}\right|r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3. \quad (5.21)$$

Changing integral and sum variables, we can then express this as

$$\int_{\boldsymbol{D}} \mathcal{C}_{(2,3)}\left[p_3^{k+1}(t;\boldsymbol{z}_1,{}'\boldsymbol{z}_2,\boldsymbol{z}_3)\right]\,\mathrm{d}\boldsymbol{z}_2$$
$$= \frac{\left|\mathbb{S}^{d-2}\right|}{(P_{k+1})^2} p_1^{k+1}(t;\boldsymbol{z}_1) \sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \leq N-m_1-k+2}}^{N} \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega \int_0^{r^*_{m_2,m_3}(|\boldsymbol{v}_2-\boldsymbol{v}_3|)} |\boldsymbol{v}_2-\boldsymbol{v}_3|$$
$$\times\, p_1^{k+1}(t;\boldsymbol{z}_2)\, p_1^{k+1}(t;\boldsymbol{z}_3) r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3. \quad (5.22)$$

Substituting results from (5.15), (5.18), (5.20) and (5.22) into (5.8), the full time-evolution equation becomes

$$\frac{1}{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\left(\partial_t p_1^k(t;\boldsymbol{z}_1) + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1} p_1^k(t;\boldsymbol{z}_1)\right)$$
$$= \frac{(k-1)}{P_k} \sum_{m_2=1}^{N-m_1-k+2} \int_{\mathbb{R}^d}\int_0^\infty \Big[|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1|\, p_1^k(t;{}^*\boldsymbol{z}_1)\, p_1^k(t;{}^*\boldsymbol{z}_2)\delta_{12}\left|D\psi^{-1}_{m_1,m_2,r}\right|$$
$$- |\boldsymbol{v}_2-\boldsymbol{v}_1|\, p_1^k(t;\boldsymbol{z}_1)\, p_1^k(t;\boldsymbol{z}_2)\delta_{12}\Big] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2$$
$$+ \frac{k+1}{2P_{k+1}} \sum_{m_2=1}^{m_1-1} \int_{\mathbb{R}^d}\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)} |\boldsymbol{v}_2-{}'\boldsymbol{v}_1|$$
$$\times\, p_1^{k+1}(t;{}'\boldsymbol{z}_1)\, p_1^{k+1}(t;\boldsymbol{z}_2)\delta_{12}\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}\right| r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2$$
$$- \frac{(k-1)(k-2)}{2(P_k)^2} p_1^k(t;\boldsymbol{z}_1) \sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \leq N-m_1-k+3}}^{N} \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega \int_0^{r^*_{m_2,m_3}(|\boldsymbol{v}_3-\boldsymbol{v}_2|)} \Big[|\boldsymbol{v}_3-\boldsymbol{v}_2|$$
$$\times\, p_1^k(t;\boldsymbol{z}_2)\, p_1^k(t;\boldsymbol{z}_3)\delta_{23}\Big] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3$$
$$+ \frac{(k-1)(k+1)}{2(P_{k+1})^2} p_1^{k+1}(t;\boldsymbol{z}_1) \sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \leq N-m_1-k+2}}^{N} \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega \int_0^{r^*_{m_2,m_3}(|\boldsymbol{v}_3-\boldsymbol{v}_2|)} \Big[|\boldsymbol{v}_3-\boldsymbol{v}_2|$$
$$\times\, p_1^{k+1}(t;\boldsymbol{z}_2)\, p_1^{k+1}(t;\boldsymbol{z}_3)\delta_{23}\Big] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3. \quad (5.23)$$

5.1.3. A note on cutoff interactions

As discussed in §2.2.1, it is well-known in the Boltzmann equation literature that it is not possible to define an integrable cross-section for an interaction that is not compactly supported, even if it is short-ranged (see for instance [52] §8.3.4, or [108] Chapter 1, §1.4). This is described as being a result of the large numbers of 'grazing' collisions: interactions with large impact parameters where the particles are hardly deviated. In our formulation, this manifests itself as the fact that the two integrand terms in the first RHS term of (5.23) are not integrable separately as $r\to\infty$. If we wish to separate these terms and treat them as individual interaction operators, we need to restrict the $\mathrm{d}r$ interval with a cutoff impact parameter r_c. Given our

interaction $\boldsymbol{H}(\boldsymbol{x}, \boldsymbol{v})$ is short-ranged (scaling as $O(|\boldsymbol{x}|^{-d-\delta})$ with $\delta > 0$ as $|\boldsymbol{x}| \to \infty$), we have that the integrand

$$|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1|\, p_1^k(t; {}^*\boldsymbol{z}_1)\, p_1^k(t; {}^*\boldsymbol{z}_2)\delta_{12} \left| D\psi^{-1}_{m_1,m_2,r} \right| - |\boldsymbol{v}_2 - \boldsymbol{v}_1|\, p_1^k(t; \boldsymbol{z}_1)\, p_1^k(t; \boldsymbol{z}_2)\delta_{12}$$

scales as $O(r^{-1-\delta})$ as $r \to \infty$, which means the error from introducing the cutoff will be $O(r_c^{-\delta})$. In practice, we will often take $r_c \to \infty$ in any case when computing this integral, as then there is no contribution from the upper limit.

5.1.4. Number density

Define the cluster number density as the number of clusters in the system multiplied by the probability density that each cluster has position $\boldsymbol{x}$, velocity $\boldsymbol{v}$ and mass m, i.e.

$$n_1^k(t, \boldsymbol{x}, \boldsymbol{v}, m) := k p_1^k(t, \boldsymbol{x}, \boldsymbol{v}, m). \tag{5.24}$$

From this define the k-marginalised number density, by summing over k:

$$n(t, \boldsymbol{x}, \boldsymbol{v}, m) = \sum_{k=1}^{N} n_1^k(t, \boldsymbol{x}, \boldsymbol{v}, m) = \sum_{k=1}^{N} k p_1^k(t, \boldsymbol{x}, \boldsymbol{v}, m). \tag{5.25}$$

Note that we have dropped the subscript 1 in the k-marginalised number density for clarity and consistency with the Boltzmann equation literature, as this will be our main density used in the analysis from here on. We also drop the subscript in the k-marginalised probability density, defined as

$$p(t, \boldsymbol{x}, \boldsymbol{v}, m) = \sum_{k=1}^{N} p_1^k(t, \boldsymbol{x}, \boldsymbol{v}, m).$$

Multiplying (5.23) by k and summing over k, we get

$$\begin{aligned}
&\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\left(\partial_t \sum_{k=1}^{N} k p_1^k(t;\boldsymbol{z}_1) + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1}\sum_{k=1}^{N} k p_1^k(t;\boldsymbol{z}_1)\right)\\
&\quad= \sum_{k=2}^{N} k\frac{k(k-1)}{P_k}\sum_{m_2=1}^{N-m_1-k+2}\int_{\mathbb{R}^d}\int_0^{r_c}\Big[|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1|\, p_1^k(t;{}^*\boldsymbol{z}_1)\, p_1^k(t;{}^*\boldsymbol{z}_2)\delta_{12}\left|D\psi^{-1}_{m_1,m_2,r}\right|\\
&\qquad\qquad - |\boldsymbol{v}_2-\boldsymbol{v}_1|\, p_1^k(t;\boldsymbol{z}_1)\, p_1^k(t;\boldsymbol{z}_2)\delta_{12}\Big] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2\\
&\qquad+\sum_{k=1}^{N-1}\frac{k(k+1)}{2P_{k+1}}\sum_{m_2=1}^{m_1-1}\int_{\mathbb{R}^d}\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)}|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|\\
&\qquad\qquad\times p_1^{k+1}(t;{}'\boldsymbol{z}_1)\, p_1^{k+1}(t;\boldsymbol{z}_2)\delta_{12}\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}\right| r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2\\
&\qquad-\sum_{k=3}^{N}\frac{k(k-1)(k-2)}{2(P_k)^2}p_1^k(t;\boldsymbol{z}_1)\sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \le N-m_1-k+3}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{m_2,m_3}(|\boldsymbol{v}_3-\boldsymbol{v}_2|)}\Big[|\boldsymbol{v}_3-\boldsymbol{v}_2|\\
&\qquad\qquad\times p_1^k(t;\boldsymbol{z}_2)\, p_1^k(t;\boldsymbol{z}_3)\delta_{23}\Big] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3\\
&\qquad+\sum_{k=2}^{N-1}\frac{k(k-1)(k+1)}{2(P_{k+1})^2}p_1^{k+1}(t;\boldsymbol{z}_1)\sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \le N-m_1-k+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{m_2,m_3}(|\boldsymbol{v}_3-\boldsymbol{v}_2|)}\Big[|\boldsymbol{v}_3-\boldsymbol{v}_2|\\
&\qquad\qquad\times p_1^{k+1}(t;\boldsymbol{z}_2)\, p_1^{k+1}(t;\boldsymbol{z}_3)\delta_{23}\Big] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_2\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_3.
\end{aligned}\tag{5.26}$$

Changing the dummy variables of the second and fourth RHS terms from k to $k'=k+1$, the third and fourth RHS terms (representing the three-particle products) cancel to give

$$\begin{aligned}
&\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\left(\partial_t n(t;\boldsymbol{z}_1) + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1} n(t;\boldsymbol{z}_1)\right)\\
&\quad= \sum_{k=2}^{N}\frac{k(k-1)}{P_k}\sum_{m_2=1}^{N-m_1-k+2}\int_{\mathbb{R}^d}\int_0^{r_c}\Big[|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1|\, p_1^k(t;{}^*\boldsymbol{z}_1)\, p_1^k(t;{}^*\boldsymbol{z}_2)\delta_{12}\left|D\psi^{-1}_{m_1,m_2,r}\right|\\
&\qquad\qquad - |\boldsymbol{v}_2-\boldsymbol{v}_1|\, p_1^k(t;\boldsymbol{z}_1)\, p_1^k(t;\boldsymbol{z}_2)\delta_{12}\Big] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2\\
&\qquad+\sum_{k=2}^{N}\frac{k(k-1)}{2P_k}\sum_{m_2=1}^{m_1-1}\int_{\mathbb{R}^d}\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)}|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|\\
&\qquad\qquad\times p_1^{k}(t;{}'\boldsymbol{z}_1)\, p_1^{k}(t;\boldsymbol{z}_2)\delta_{12}\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}\right| r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2.
\end{aligned}\tag{5.27}$$

5.1.5. *Clusterzahlansatz*

To close this equation, we must define a relationship between n and p_1^k. In the one-dimensional case, we defined the *Clusterzahlansatz*, assuming $p_1^k(t;\boldsymbol{z}\,|\,k)\approx p(t;\boldsymbol{z})$. Applying this to the

definition of n, we get

$$\begin{aligned} n(t;\boldsymbol{z}_1) = \sum_{k=1}^{N} k p_1^k(t;\boldsymbol{z}_1) &= \sum_{k=1}^{N} k P_k p_1(t;\boldsymbol{z}_1 \,|\, k) \\ &\approx p(t;\boldsymbol{z}_1) \sum_{k=1}^{N} k P_k = \mu_K p(t;\boldsymbol{z}_1), \end{aligned} \tag{5.28}$$

where again we have defined $\mu_K = \sum_{k=1}^{N} k P_k$: the expected number of clusters in the system. Using another instance of the *Clusterzahlansatz*, we can express the product cluster probabilities of (5.27) in terms of n through

$$\begin{aligned} p_1^k(t;\boldsymbol{z}_1)\, p_1^k(t;\boldsymbol{z}_2) &= (P_k)^2 p_1(t;\boldsymbol{z}_1 \,|\, k) p_1(t;\boldsymbol{z}_2 \,|\, k) \\ &\approx (P_k)^2 p(t;\boldsymbol{z}_1)\, p(t;\boldsymbol{z}_2) \\ &\approx \left(\frac{P_k}{\mu_K}\right)^2 n(t;\boldsymbol{z}_1)\, n(t;\boldsymbol{z}_2). \end{aligned} \tag{5.29}$$

This same manipulation applies to the other instances with modified $\boldsymbol{z}$ arguments. Substituting this result into (5.27), we get

$$\begin{aligned} &\frac{1}{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\left(\partial_t n + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} n\right) \\ &\quad = \frac{1}{(\mu_K)^2} \sum_{k=2}^{N} k(k-1) P_k \sum_{m_2=1}^{N-m_1-k+2} \int_{\mathbb{R}^d} \int_0^{r_c} \Big[\left|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1\right| {}^*n\, {}^*n_{(2)} \left| D\psi^{-1}_{m_1,m_2,r} \right| \\ &\qquad\qquad - \left|\boldsymbol{v}_2 - \boldsymbol{v}_1\right| n\, n_{(2)} \Big] r^{d-2} \,\mathrm{d}r\, \mathrm{d}\boldsymbol{v}_2 \\ &\qquad + \frac{1}{2(\mu_K)^2} \sum_{k=2}^{N} k(k-1) P_k \sum_{m_2=1}^{m_1-1} \int_{\mathbb{R}^d} \int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2 - {}'\boldsymbol{v}_1|)} \left|\boldsymbol{v}_2 - {}'\boldsymbol{v}_1\right| {}'n\, n_{(2)} \\ &\qquad\qquad \times \left| D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r} \right| r^{d-2} \,\mathrm{d}r\, \mathrm{d}\boldsymbol{v}_2. \end{aligned} \tag{5.30}$$

where we have introduced the shorthand $n \equiv n(t;\boldsymbol{x}_1;\boldsymbol{v}_1;m_1)$, $n_{(2)} \equiv n(t;\boldsymbol{x}_1;\boldsymbol{v}_2;m_2)$, ${}^*n \equiv n(t;\boldsymbol{x}_1;{}^*\boldsymbol{v}_1;m_1)$, and ${}'n \equiv n(t;\boldsymbol{x}_1;{}'\boldsymbol{v}_1;{}'m_1) = n(t;\boldsymbol{x}_1;{}'\boldsymbol{v}_1;m_1-m_2)$. If we replace k with $\lfloor \mu_K \rfloor$ in the upper limit of the m_2 sum of the first RHS term, all the k dependence is in the factor

$$\sum_{k=2}^{N} k(k-1) P_k = \mu_{K^2} - \mu_K,$$

where $\mu_{K^2} = \sum_{k=1}^{N} k^2 P_k$ is the expected value of K^2. Defining the coefficient

$$\Lambda_K(t) := \frac{\mu_{K^2}(t) - \mu_K(t)}{(\mu_K(t))^2} = 1 + \frac{\sigma_K^2(t) - \mu_K(t)}{(\mu_K(t))^2}, \tag{5.31}$$

where $\sigma_K^2(t)$ is the variance of the K distribution, we can write the closed strong form of our short-range PDE clustering model:

$$\begin{aligned}
&\frac{1}{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\left(\partial_t n + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} n\right) \\
&\quad = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d}\int_0^{r_c} \left[\left|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1\right| {}^*n\, {}^*n_{(2)} \left|D\psi^{-1}_{m_1,m_2,r}\right| - \left|\boldsymbol{v}_2 - \boldsymbol{v}_1\right| n\, n_{(2)}\right] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2 \\
&\qquad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \int_{\mathbb{R}^d}\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)} \left|\boldsymbol{v}_2 - {}'\boldsymbol{v}_1\right| {}'n\, n_{(2)} \left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}\right| r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2,
\end{aligned} \tag{5.32}$$

where $({}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2) = \psi^{-1}_{m_1,m_2,r}(\boldsymbol{v}_1,\boldsymbol{v}_2)$ and ${}'\boldsymbol{v}_1 = \hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}(\boldsymbol{v}_1)$. As will be seen in the derivation of the weak form in §5.2.1, this model conserves total cluster mass and any quantity conserved in both clustering and non-clustering interactions. We can define it in operator form as

$$\frac{1}{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\left(\partial_t n + \boldsymbol{v}_1 \cdot \nabla_{\boldsymbol{x}_1} n\right) = Q_I(n,n) + Q_C^+(n,n) \tag{5.33}$$

where $Q_I(n,n)$ is the first RHS term in (5.32), representing the trajectories that enter and leave the inner region at infinity, and $Q_C^+(n,n)$ is the second RHS term in (5.32), representing the probability produced by clustering interactions. We can also split $Q_I(n,n)$ into three operators representing different interactions. Defining the non-clustering gain,

$$Q^+_{NC}(n,n) = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d}\int_0^{r_c} \left|{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1\right| {}^*n\, {}^*n_{(2)} \left|D\psi^{-1}_{m_1,m_2,r}\right| r^{d-2}\,\mathrm{d}r\mathrm{d}\boldsymbol{v}_2, \tag{5.34}$$

non-clustering loss,

$$Q^-_{NC}(n,n) = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d}\int_{r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)}^{r_c} \left|\boldsymbol{v}_2 - \boldsymbol{v}_1\right| n\, n_{(2)} r^{d-2}\,\mathrm{d}r\mathrm{d}\boldsymbol{v}_2, \tag{5.35}$$

and clustering loss,

$$Q^-_{C}(n,n) = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d}\int_0^{r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)} \left|\boldsymbol{v}_2 - \boldsymbol{v}_1\right| n\, n_{(2)} r^{d-2}\,\mathrm{d}r\mathrm{d}\boldsymbol{v}_2. \tag{5.36}$$

operators, we can write

$$Q_I(n,n) = Q^+_{NC}(n,n) - Q^-_{NC}(n,n) - Q^-_C(n,n).$$

As discussed in §2.2 and §2.3, it is common for Boltzmann-style kinetic equations to parameterise the interaction operators in terms of one of the unit vectors $\boldsymbol{\sigma}$ or $\boldsymbol{\omega}$ related to the deviation angle θ. We have chosen, however, to keep these interaction operators as integrals over the impact parameter r as, for general clustering interactions, we are not guaranteed there will be a well-defined, monotonically increasing map between r and θ.

5.1.6. Cluster number distribution

The most direct way to compute the coefficient Λ_K is to approximate the evolution of the full cluster number distribution. We can do this by marginalising out the position, mass and velocity arguments from (5.23): the evolution equation before we summed over k. Specifically, integrating (5.23) over $\boldsymbol{x}$ and $\boldsymbol{v}$, summing over m and following similar steps to simplify as in the one-dimensional case (see §A.4), we get

$$\begin{aligned}\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}P_k}{\mathrm{d}t} = &-\frac{k(k-1)}{2P_k}\sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-k+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{(1,2)}}|\boldsymbol{v}_1-\boldsymbol{v}_2|\,p_1^k p_{1(2)}^k r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}_2\\ &+\frac{(k+1)k}{2P_{k+1}}\sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-k+1}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{(1,2)}}|\boldsymbol{v}_1-\boldsymbol{v}_2|\,p_1^{k+1} p_{1(2)}^{k+1} r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}_2,\end{aligned} \tag{5.37}$$

where we use the similar shorthand $p^k_{1(2)} \equiv p^k_1(t;\boldsymbol{x}_1;\boldsymbol{v}_2;m_2)$ and introduce the shorthand $r^*_{(1,2)} \equiv r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)$. Using the *Clusterzahlansatz* and the same modifications as (5.29), we can express the RHS in terms of n as

$$\begin{aligned}\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}P_k}{\mathrm{d}t} = &-\frac{k(k-1)}{2(\mu_K)^2}P_k\sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-k+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{(1,2)}}|\boldsymbol{v}_1-\boldsymbol{v}_2|\,n\,n_{(2)} r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}_2\\ &+\frac{(k+1)k}{2(\mu_K)^2}P_{k+1}\sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-k+1}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{(1,2)}}|\boldsymbol{v}_1-\boldsymbol{v}_2|\,n\,n_{(2)} r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}_2.\end{aligned} \tag{5.38}$$

In a numerical implementation of this model, we could then evolve the set of $P_k(t)$ in a separate half-step using the current value of $n(t;\boldsymbol{x}_1;\boldsymbol{v}_1;m_1)$ before using the set of $P_k(t)$ to approximate $\sigma^2_K(t)$ and hence $\Lambda_K(t)$, allowing us to evolve n using (5.32).

5.1.7. Scaling

We now consider how the time scale of the system changes when we change key parameters such as the number of particles, the velocity of particles, or the size of the spatial domain relative to the interaction length scale. As the integral of the number density is the mean number of clusters:

$$\sum_{m=1}^{N}\int_\Omega\int_{\mathbb{R}^d} n(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{v}\,\mathrm{d}\boldsymbol{x} = \mu_K(t),$$

to normalise to a probability density we could scale n with $\mu_K(t)$. However, as $\mu_K(t)$ is time-dependent, this scaling would introduce other terms through the time derivative, complicating the equation. Note that as $n = \mu_K p$ (see (5.28)), this normalised number density is equivalent

to the probability density p we used earlier. The extra terms arising from the time derivative would be equivalent to the three-cluster terms that cancelled in (5.27) when we changed variables to n. As such, we scale n by N, defining $\hat{n} := n/N$, as N is time independent and μ_K scales with N.

If we consider the spatially uniform case, marginalising out $\boldsymbol{x}$ we get

$$n(t;\boldsymbol{v};m) = \int_\Omega n(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{x} = |\Omega| n(t;\boldsymbol{x};\boldsymbol{v};m). \tag{5.39}$$

Define velocity and time scales through $\boldsymbol{v} = \bar{v}\,\hat{\boldsymbol{v}}$ and $t = \tau\,\hat{t}$, where $\bar{v}$ is a representative speed given the initial condition and τ the time scale to be determined. With this velocity scale, the number density would scale as $\hat{n}(t;\hat{\boldsymbol{v}},m) = |\Omega| n(t;\boldsymbol{x};\boldsymbol{v};m)/(N\bar{v})$. Substituting these scales into the spatially uniform version of (5.33), we get

$$\frac{|\Omega|}{N\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|\bar{v}\,\tau}\partial_{\hat{t}}\,\hat{n} = Q_I(\hat{n},\hat{n}) + Q_C^+(\hat{n},\hat{n}). \tag{5.40}$$

Setting

$$\tau = \frac{|\Omega|}{N\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|\bar{v}} \tag{5.41}$$

cancels the coefficient on the LHS and, as such, defines the time scale of the system. To interpret this scale, we first note that

$$\left|\mathcal{B}_\varepsilon^{d-1}\right| = \int_0^\varepsilon r^{d-2}\left|\mathbb{S}^{d-2}\right|\mathrm{d}r = \frac{1}{d-1}\left|\mathbb{S}^{d-2}\right|\varepsilon^{d-1},$$

where $\left|\mathcal{B}_\varepsilon^{d-1}\right|$ is the generalised volume of the ball of radius ε in $d-1$ dimensions. As such, $N\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|$ is the total cross-section of a $d-1$ dimensional ball of radius ε centred on each of the particles (up to a factor of $1/(d-1)$). When multiplied by the representative speed, $\bar{v}$ (defined to be perpendicular to the cross-section), this represents the total d-dimensional volume swept out by the particles per unit time. As such, by taking the ratio with the full spatial volume, $|\Omega|$, the timescale represents the representative time for the particles' interacting cross sections to sweep out a $1/(d-1)$ fraction of the full spatial domain. In two dimensions, for instance, the cross-section is a line of length 2ε, centred at each particle and perpendicular to its velocity. Therefore, the time scale is the time taken for these cross-sections to sweep out half the area of the spatial domain if each particle were to be moving at representative speed $\bar{v}$. Note that this $1/(d-1)$ factor would appear from the $\mathrm{d}r$ integral on the RHS of (5.32) if the interaction terms were independent of r.

5.2. Weak form kinetic PDE

To analyse the time evolution of macroscopically observable quantities, it is often useful to modify kinetic PDE to its weak formulation (see §2.2.2). The weak formulation will also prove useful for implementing our model numerically in Chapter 6. To derive this weak form, first

define the inner product operator

$$\langle g, \varphi \rangle = \sum_{m=1}^{N-\mu_K+1} \int_{\mathbb{R}^d} \int_{\Omega} g\, \varphi \, \mathrm{d}\boldsymbol{x} \, \mathrm{d}\boldsymbol{v} \tag{5.42}$$

on a function $g(t; \boldsymbol{x}; \boldsymbol{v}; m)$, where $\varphi(\boldsymbol{v}; m)$ is an arbitrary function of $\boldsymbol{v}$ and m. Acting with this operator on the strong form of the short-range clustering PDE model, (5.33), we have

$$\frac{1}{\varepsilon^{d-1} \left|\mathbb{S}^{d-2}\right|} \frac{\mathrm{d}}{\mathrm{d}t} \langle n, \varphi \rangle = \left\langle Q_{NC}^{+}(n,n), \varphi \right\rangle - \left\langle Q_{NC}^{-}(n,n), \varphi \right\rangle - \left\langle Q_{C}^{-}(n,n), \varphi \right\rangle + \left\langle Q_{C}^{+}(n,n), \varphi \right\rangle \tag{5.43}$$

Written out fully, the non-clustering gain term is

$$\begin{aligned} &\left\langle Q_{NC}^{+}(n,n), \varphi \right\rangle \\ &\quad = \Lambda_K \sum_{\substack{m_1,m_2=1 \\ m_1+m_2 \\ \leq N-\mu_K+2}}^{N} \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\Omega} \int_{0}^{r_c} |{}^{*}\boldsymbol{v}_2 - {}^{*}\boldsymbol{v}_1| \, {}^{*}n \, {}^{*}n_{(2)} \left| D\psi_{m_1,m_2,r}^{-1} \right| \varphi \, r^{d-2} \, \mathrm{d}r \mathrm{d}\boldsymbol{x}_1 \, \mathrm{d}\boldsymbol{v}_2 \mathrm{d}\boldsymbol{v}_1. \end{aligned} \tag{5.44}$$

Changing integration variables from $\mathrm{d}\boldsymbol{v}_2 \, \mathrm{d}\boldsymbol{v}_1$ to $\mathrm{d}^{*}\boldsymbol{v}_2 \, \mathrm{d}^{*}\boldsymbol{v}_1$ using the map $({}^{*}\boldsymbol{v}_1, {}^{*}\boldsymbol{v}_2) = \psi_{m_1,m_2,r}^{-1}(\boldsymbol{v}_1, \boldsymbol{v}_2)$, the Jacobian $\left| D\psi_{m_1,m_2,r}^{-1} \right|$ is absorbed:

$$\begin{aligned} &\int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\Omega} \int_{0}^{r_c} |{}^{*}\boldsymbol{v}_2 - {}^{*}\boldsymbol{v}_1| \, {}^{*}n \, {}^{*}n_{(2)} \left| D\psi_{m_1,m_2,r}^{-1} \right| \varphi \, r^{d-2} \, \mathrm{d}r \, \mathrm{d}\boldsymbol{x}_1 \, \mathrm{d}\boldsymbol{v}_2 \mathrm{d}\boldsymbol{v}_1 = \\ &\qquad \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\Omega} \int_{r^{*}_{m_1,m_2}(|{}^{*}\boldsymbol{v}_2 - {}^{*}\boldsymbol{v}_1|)}^{r_c} |{}^{*}\boldsymbol{v}_2 - {}^{*}\boldsymbol{v}_1| \, {}^{*}n \, {}^{*}n_{(2)} \varphi \, r^{d-2} \, \mathrm{d}r \, \mathrm{d}\boldsymbol{x}_1 \, \mathrm{d}^{*}\boldsymbol{v}_2 \mathrm{d}^{*}\boldsymbol{v}_1. \end{aligned} \tag{5.45}$$

Note that the range of integration for r has also been reduced to $[r^{*}_{m_1,m_2}(|{}^{*}\boldsymbol{v}_2 - {}^{*}\boldsymbol{v}_1|), r_c]$, as only combinations of r, ${}^{*}\boldsymbol{v}_1$, and ${}^{*}\boldsymbol{v}_2$ that result in non-clustering interactions could be pre-images of velocities that are part of the non-clustering gain term under consideration. Mapping dummy integration variables $({}^{*}\boldsymbol{v}_1, {}^{*}\boldsymbol{v}_2)$ to $(\boldsymbol{v}_1, \boldsymbol{v}_2)$, it follows that $\boldsymbol{v}_1$ becomes $\boldsymbol{v}_1^{*}$ as $(\boldsymbol{v}_1, \boldsymbol{v}_2) = \psi^{-1}(\psi(\boldsymbol{v}_1, \boldsymbol{v}_2)) = \psi^{-1}(\boldsymbol{v}_1^{*}, \boldsymbol{v}_2^{*})$. Using the same short-hand notation for modifications to the arguments of φ as we used for n, i.e. $\varphi^{*} \equiv \varphi(\boldsymbol{v}^{*}, m)$, the integral becomes

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\Omega} \int_{r^{*}_{m_1,m_2}(|\boldsymbol{v}_2 - \boldsymbol{v}_1|)}^{r_c} |\boldsymbol{v}_2 - \boldsymbol{v}_1| \, n \, n_{(2)} \varphi^{*} \, r^{d-2} \, \mathrm{d}r \, \mathrm{d}\boldsymbol{x}_1 \, \mathrm{d}\boldsymbol{v}_2 \mathrm{d}\boldsymbol{v}_1.$$

We can combine this with the unchanged loss terms to give

$$\begin{aligned} &\left\langle Q_{NC}^{+}(n,n), \varphi \right\rangle - \left\langle Q_{NC}^{-}(n,n), \varphi \right\rangle - \left\langle Q_{C}^{-}(n,n), \varphi \right\rangle \\ &= \Lambda_K \sum_{\substack{m_1,m_2=1 \\ m_1+m_2 \\ \leq N-\mu_K+2}}^{N} \int_{\mathbb{R}^d \times \mathbb{R}^d} \int_{\Omega} |\boldsymbol{v}_2 - \boldsymbol{v}_1| \, n \, n_{(2)} \left[\int_{r^{*}_{(1,2)}}^{r_c} (\varphi^{*} - \varphi) \, r^{d-2} \, \mathrm{d}r - \int_{0}^{r^{*}_{(1,2)}} \varphi \, r^{d-2} \, \mathrm{d}r \right] \mathrm{d}\boldsymbol{x}_1 \, \mathrm{d}\boldsymbol{v}_2 \mathrm{d}\boldsymbol{v}_1. \end{aligned} \tag{5.46}$$

For the clustering gain term, we have

$$\begin{aligned}
&\left\langle Q_C^+(n,n),\varphi\right\rangle \\
&\quad= \frac{1}{2}\Lambda_K \sum_{m_1=1}^{N-\mu_K+1}\sum_{m_2=1}^{m_1-1}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-'\boldsymbol{v}_1|)} |\boldsymbol{v}_2-'\boldsymbol{v}_1|\,'n\,n_{(2)} \\
&\qquad\qquad\qquad\qquad\qquad\qquad\times\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2,r}\right|\,\varphi\,r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1 \\
&\quad= \frac{1}{2}\Lambda_K \sum_{\substack{'m_1,m_2=1\\ 'm_1+m_2\\ \le N-\mu_K+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{'m_1,m_2}(|\boldsymbol{v}_2-'\boldsymbol{v}_1|)} |\boldsymbol{v}_2-'\boldsymbol{v}_1|\,'n\,n_{(2)}\,\varphi\,r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}'\boldsymbol{v}_1,
\end{aligned} \tag{5.47}$$

where we have changed the $\mathrm{d}\boldsymbol{v}_1$ integral to $\mathrm{d}'\boldsymbol{v}_1$, absorbing the Jacobian, $\left|D\hat{\lambda}^{-1}_{m_1,m_2,\boldsymbol{v}_2,r}\right|$, and changed the m_1 sum variable to $'m_1 = m_1 - m_2$, combining the bounds with the m_2 sum. Now, mapping the dummy variables $'\boldsymbol{v}_1$ to $\boldsymbol{v}_1$ and $'m_1$ to m_1, it follows that $\varphi(\boldsymbol{v}_1, m_1)$ must become $\varphi(\boldsymbol{v}_1', m_1') = \varphi(\boldsymbol{v}_1', m_1 + m_2) \equiv \varphi'$, giving

$$\left\langle Q_C^+(n,n),\varphi\right\rangle = \frac{1}{2}\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \le N-\mu_K+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega\int_0^{r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)} |\boldsymbol{v}_2-\boldsymbol{v}_1|\,n\,n_{(2)}\,\varphi'\,r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \tag{5.48}$$

Combining this with (5.46), we have the complete weak form of our short-range clustering PDE model:

$$\begin{aligned}
\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}}{\mathrm{d}t}\langle n,\varphi\rangle &= \Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \le N-\mu_K+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega |\boldsymbol{v}_2-\boldsymbol{v}_1|\,n\,n_{(2)} \\
&\quad\times\left[\int_{r^*_{(1,2)}}^{r_c}(\varphi^*-\varphi)\,r^{d-2}\,\mathrm{d}r + \int_0^{r^*_{(1,2)}}\left(\frac{\varphi'}{2}-\varphi\right)r^{d-2}\,\mathrm{d}r\right]\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1.
\end{aligned} \tag{5.49}$$

Using the symmetry of the RHS, we can switch dummy variables $m_1 \leftrightarrow m_2$ and $\boldsymbol{v}_1 \leftrightarrow \boldsymbol{v}_2$, which will switch the arguments of φ to $\boldsymbol{v}_2$ and m_2. Adding this modified equation to (5.49) and dividing by two, we get a second version of the weak form of the PDE model:

$$\begin{aligned}
&\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}}{\mathrm{d}t}\langle n,\varphi\rangle \\
&\quad= \frac{1}{2}\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \le N-\mu_K+2}}^{N}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega |\boldsymbol{v}_2-\boldsymbol{v}_1|\,n\,n_{(2)}\left[\int_{r^*_{(1,2)}}^{r_c}\left(\varphi^*+\varphi^*_{(2)}-\varphi-\varphi_{(2)}\right)r^{d-2}\,\mathrm{d}r\right. \\
&\qquad\qquad\left.+\int_0^{r^*_{(1,2)}}\left(\varphi'-\varphi-\varphi_{(2)}\right)r^{d-2}\,\mathrm{d}r\right]\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1,
\end{aligned} \tag{5.50}$$

where $\varphi_{(2)} \equiv \varphi(\boldsymbol{v}_2; m_2)$.

5.2.1. Conserved quantities

From the structure of (5.50), we can see that any quantity φ that is conserved over both clustering and non-clustering interactions will be conserved in general. As the inner product $\langle n, m\rangle$ represents the total mass of clusters in the system (i.e. the total number of particles), this quantity should be conserved at N for all t. Setting $\varphi(\boldsymbol{v}, m) = m$, we have

$$\varphi^* + \varphi^*_{(2)} - \varphi - \varphi_{(2)} = m_1 + m_2 - m_1 - m_2 = 0,$$

and

$$\varphi' - \varphi - \varphi_{(2)} = (m_1 + m_2) - m_1 - m_2 = 0,$$

giving $\partial_t\langle n, m\rangle = 0$ as expected.

Similarly, if we set $\varphi(\boldsymbol{v}, m) = \boldsymbol{v}\, m$, then $\langle n, \boldsymbol{v}\, m\rangle$ represents the total momentum of the system. Considering the two dr integrands in (5.50) we have

$$\varphi^* + \varphi^*_{(2)} - \varphi - \varphi_{(2)} = \boldsymbol{v}_1^* m_1 + \boldsymbol{v}_2^* m_2 - \boldsymbol{v}_1 m_1 - \boldsymbol{v}_2 m_2,$$

and

$$\varphi' - \varphi - \varphi_{(2)} = (m_1 + m_2)\boldsymbol{v}_1' - m_1\boldsymbol{v}_1 - m_2\boldsymbol{v}_2.$$

If we set these to 0, they, by definition, represent the condition for momentum conservation over each type of interaction. So if this is satisfied in both cases, we have $\partial_t\langle n, \boldsymbol{v}\, m\rangle = 0$.

To consider the conservation of probability mass in the system, note that, as we are modelling the cluster number density n, we define the cluster probability density as $p = n/\mu_K$. As we also have $\langle n, 1\rangle = \mu_K$ it follows that

$$\frac{\mathrm{d}}{\mathrm{d}t}\langle p, 1\rangle = \frac{\mathrm{d}}{\mathrm{d}t}\left(\frac{1}{\mu_K}\langle n, 1\rangle\right) = \frac{\mathrm{d}}{\mathrm{d}t}(1) = 0,$$

and so the total probability mass is conserved, provided $\mu_K \neq 0$.

5.2.2. Mean cluster number evolution

The expected number of clusters in the system is given by $\mu_K = \langle n, 1\rangle$, so by setting $\varphi(\boldsymbol{v}, m) = 1$ in (5.50) we get a time-evolution equation for μ_K. With this definition of φ, we have that

$$\varphi^* + \varphi^*_{(2)} - \varphi - \varphi_{(2)} = 2 - 2 = 0,$$

showing non-clustering interactions don't affect the mean number of clusters and

$$\varphi' - \varphi - \varphi_{(2)} = -1,$$

representing the fact that each clustering interaction removes one cluster from the system. Substituting these into (5.50) and performing the dr integral (noting $d \geq 2$ was required for

the derivation) we get

$$\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}}{\mathrm{d}t}\mu_K = -\frac{1}{2(d-1)}\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega |\boldsymbol{v}_2-\boldsymbol{v}_1|\, n\, n_{(2)} \left(r^*_{(1,2)}\right)^{d-1} \mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \tag{5.51}$$

5.2.3. Smoluchowski coagulation equation

Define a new inner product operator,

$$\langle g,\varphi\rangle_{x,v} = \int_{\mathbb{R}^d}\int_\Omega g\,\varphi\,\mathrm{d}\boldsymbol{x}\,\mathrm{d}\boldsymbol{v}, \tag{5.52}$$

to be the same as (5.42), except without the sum over mass. Acting with this inner product on the interaction operators from (5.33) and performing similar manipulations to §5.2, we have

$$\begin{aligned}\langle Q_I(n,n),\varphi\rangle_{x,v} = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega |\boldsymbol{v}_2-\boldsymbol{v}_1|\, n\, n_{(2)} \\ \times\left[\int_{r^*_{(1,2)}}^{r_c}(\varphi^*-\varphi)\,r^{d-2}\,\mathrm{d}r - \int_0^{r^*_{(1,2)}}\varphi\,r^{d-2}\,\mathrm{d}r\right]\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1,\end{aligned} \tag{5.53}$$

and

$$\begin{aligned}\left\langle Q_C^+(n,n),\varphi\right\rangle_{x,v} = \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega |\boldsymbol{v}_2-{}'\boldsymbol{v}_1|\,{}'n\, n_{(2)} \\ \times\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)}\varphi\,r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}'\boldsymbol{v}_1.\end{aligned} \tag{5.54}$$

If we set $\varphi = 1$ we have

$$\langle n,1\rangle_{x,v} = \int_{\mathbb{R}^d\times\Omega} n(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v} = n(t;m),$$

which is the number distribution of clusters with mass m. As such, acting on (5.33) with the inner product (5.52) and setting $\varphi = 1$, we have the time evolution equation for the mass number distribution as

$$\begin{aligned}&\frac{d-1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\partial_t n(t;m_1) \\ &\quad= -\Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega |\boldsymbol{v}_2-\boldsymbol{v}_1|\, n\, n_{(2)}\left(r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)\right)^{d-1}\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1 \\ &\qquad+\frac{1}{2}\Lambda_K\sum_{m_2=1}^{m_1-1}\int_{\mathbb{R}^d\times\mathbb{R}^d}\int_\Omega |\boldsymbol{v}_2-{}'\boldsymbol{v}_1|\,{}'n\, n_{(2)}\left(r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)\right)^{d-1}\mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}'\boldsymbol{v}_1.\end{aligned} \tag{5.55}$$

This is in a similar form to the discrete mass Smoluchowski coagulation equation:

$$\partial_t n(t;m_1) = -\sum_{m_2=1}^{\infty} B(m_1,m_2)\, n(t;m_1)\, n(t;m_2) + \frac{1}{2}\sum_{m_2=1}^{m_1-1} B(m_1-m_2,m_2)\, n(t;m_1-m_2)\, n(t;m_2),$$

where $B(m_1,m_2)$ is known as a coagulation kernel, describing the rate at which clusters of mass m_1 coagulate with clusters of mass m_2. If we assume a spatially uniform system, we have

$$n(t;\boldsymbol{v};m) = \int_\Omega n(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{x} = |\Omega| n(t;\boldsymbol{x};\boldsymbol{v};m)$$

and using the law of conditional probability, we can separate the mass distribution as

$$n(t;\boldsymbol{v};m) = p(t;\boldsymbol{v}\,|\,m)\, n(t;m).$$

Substituting this into (5.55), we can separate the product of mass distributions to give

$$\begin{aligned}
&\frac{(d-1)|\Omega|}{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\partial_t n(t;m_1) \\
&\quad = -\Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} n(t;m_1)\, n(t;m_2) \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1| p(t;\boldsymbol{v}_1\,|\,m_1)\, p(t;\boldsymbol{v}_2\,|\,m_2) \\
&\qquad\qquad \times \left(r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)\right)^{d-1} \mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1 \\
&\qquad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} n(t;m_1-m_2)\, n(t;m_2) \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-{}'\boldsymbol{v}_1|\, p(t;{}'\boldsymbol{v}_1\,|\,m_1-m_2)\, p(t;\boldsymbol{v}_2\,|\,m_2) \\
&\qquad\qquad \times \left(r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)\right)^{d-1} \mathrm{d}\boldsymbol{v}_2\mathrm{d}{}'\boldsymbol{v}_1,
\end{aligned} \tag{5.56}$$

which is in the form of the Smoluchowski coagulation with

$$\begin{aligned}
B(m_1,m_2) = \frac{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}{(d-1)|\Omega|}\Lambda_K \int_{\mathbb{R}^d\times\mathbb{R}^d} &|\boldsymbol{v}_2-\boldsymbol{v}_1| p(t;\boldsymbol{v}_1\,|\,m_1)\, p(t;\boldsymbol{v}_2\,|\,m_2) \\
&\times \left(r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)\right)^{d-1} \mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1.
\end{aligned} \tag{5.57}$$

Note that our equation assumes a finite number of particles, whereas the Smoluchowski coagulation equation is conventionally considered to represent a system with $N\to\infty$ but finite density. To match that formulation, we simply take the limit of the m_2 sum in the first RHS term to infinity. Given a specific interaction, $r^*_{m_1,m_2}(|\boldsymbol{v}_2-\boldsymbol{v}_1|)$ will be defined. Further assumptions, however, would need to be made to approximate the time evolution of the conditional velocity distributions, $p(t;\boldsymbol{v}\,|\,m)$, such that this can be made a closed equation for the mass distribution $n(t;m)$.

5.3. Cucker-Smale model

To implement the PDE model for the Cucker-Smale model of collective behaviour, we follow a similar analysis to §4.1.2. In the d-dimensional, short-range version of the Cucker-Smale model, particles' positions and velocities evolve according to

$$\begin{aligned}\dot{\boldsymbol{X}}_i(t) &= \boldsymbol{V}_i,\\ \dot{\boldsymbol{V}}_i(t) &= \frac{1}{\varepsilon}\sum_{\substack{j=1\\ j\neq i}}^{N}\phi\left(\frac{|\boldsymbol{X}_i-\boldsymbol{X}_j|}{\varepsilon}\right)(\boldsymbol{V}_j-\boldsymbol{V}_i),\end{aligned} \tag{5.58}$$

with $\boldsymbol{X}\in\Omega=\mathbb{T}^d$ and $\boldsymbol{V}\in\mathbb{R}^d$. This is in the form of (5.1), with interaction

$$\boldsymbol{H}\left(\frac{\boldsymbol{X}_i-\boldsymbol{X}_j}{\varepsilon},\boldsymbol{V}_i-\boldsymbol{V}_j\right)=\phi\left(\frac{|\boldsymbol{X}_i-\boldsymbol{X}_j|}{\varepsilon}\right)(\boldsymbol{V}_j-\boldsymbol{V}_i).$$

As shown in deriving (5.3), if we were to consider sets of M particles that were highly correlated post interaction as a point-cluster with mass M, the velocity of each particle in cluster i, and hence the cluster itself, would instead evolve as

$$\dot{\boldsymbol{V}}_i=\frac{1}{\varepsilon}\sum_{\substack{j=1\\ j\neq i}}^{k}M_j\phi\left(\frac{|\boldsymbol{X}_i-\boldsymbol{X}_j|}{\varepsilon}\right)(\boldsymbol{V}_j-\boldsymbol{V}_i),$$

where $k<N$ is the total number of clusters remaining in the system. With this interaction, the inner equation for the evolution of the joint probability distribution, (5.7), would be

$$(\boldsymbol{v}_2-\boldsymbol{v}_1)\cdot\partial_{\tilde{\boldsymbol{x}}_{21}}\tilde{p}^k_{(1,2)}+m_2\phi\left(|\tilde{\boldsymbol{x}}_{21}|\right)\nabla_{\boldsymbol{v}_1}\cdot\Big((\boldsymbol{v}_2-\boldsymbol{v}_1)\tilde{p}^k_{(1,2)}\Big)+m_1\phi\left(|\tilde{\boldsymbol{x}}_{21}|\right)\nabla_{v_2}\cdot\Big((\boldsymbol{v}_1-\boldsymbol{v}_2)\tilde{p}^k_{(1,2)}\Big)=0. \tag{5.59}$$

The characteristic equations, changing variables to the centre of mass velocity $\bar{\boldsymbol{v}}=(m_1\boldsymbol{v}_1+m_2\boldsymbol{v}_2)/(m_1+m_2)$ and relative velocity $\boldsymbol{u}=\boldsymbol{v}_2-\boldsymbol{v}_1$, are

$$\frac{\mathrm{d}\tilde{\boldsymbol{x}}_{21}}{\mathrm{d}\tau}=\boldsymbol{u}, \tag{5.60a}$$

$$\frac{\mathrm{d}\bar{\boldsymbol{v}}}{\mathrm{d}\tau}=0, \tag{5.60b}$$

$$\frac{\mathrm{d}\boldsymbol{u}}{\mathrm{d}\tau}=-(m_1+m_2)\phi\left(|\tilde{\boldsymbol{x}}_{21}|\right)\boldsymbol{u}, \tag{5.60c}$$

$$\frac{1}{\tilde{p}^k_{(1,2)}}\frac{\mathrm{d}\tilde{p}^k_{(1,2)}}{\mathrm{d}\tau}=(m_1+m_2)\phi\left(|\tilde{\boldsymbol{x}}_{21}|\right)\nabla_{\boldsymbol{u}}\cdot\boldsymbol{u}=(m_1+m_2)\phi\left(|\tilde{\boldsymbol{x}}_{21}|\right)d. \tag{5.60d}$$

As for the one-dimensional case in §4.1.2, these characteristic equations are similar to (2.71). Following a similar analysis to §2.5, we have that the boundary between clustering and non-clustering interactions is given in general by

$$|\boldsymbol{u}|=(m_1+m_2)\Phi(r),$$

where $\boldsymbol{u}$ is the incoming relative velocity, r is the impact parameter and $\Phi(r) = \int_{-\infty}^{\infty} \phi\left(\sqrt{r^2+x^2}\right) \mathrm{d}x$. This gives us the definition for the boundary impact parameter between the two types of interaction:

$$r^*_{m_1,m_2}(|\boldsymbol{u}|) = \Phi^{-1}\left(\frac{|\boldsymbol{u}|}{m_1+m_2}\right). \tag{5.61}$$

Non-clustering interactions occur with $r > r^*_{m_1,m_2}(|\boldsymbol{u}|)$, with the relative velocity evolving as

$$\boldsymbol{u}^+ = \boldsymbol{u}^- - (m_1+m_2)\Phi(r)\frac{\boldsymbol{u}^-}{|\boldsymbol{u}^-|}. \tag{5.62}$$

As noted in §2.5, the final relative velocity $\boldsymbol{u}^+$ is parallel with the incoming relative velocity $\boldsymbol{u}^-$, meaning the deviation angle θ is always 0 and representations of the interaction in terms of the impact direction $\boldsymbol{\omega}$ or outgoing relative velocity $\boldsymbol{\sigma}$ are degenerate (see Figure 2.5). This is why we parameterise the interaction operators in our kinetic model (5.32) by the impact parameter, r, as opposed to the more traditional collision kernel formulations related to the differential cross-section.

Combining (5.62) with the conservation of momentum, (5.60b), this gives us the definition of the forwards velocity map $(\boldsymbol{v}_1^*, \boldsymbol{v}_2^*) = \psi_{m_1,m_2,r}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ as

$$\boldsymbol{v}_1^* = \boldsymbol{v}_1 + m_2\Phi(r)\frac{\boldsymbol{v}_2-\boldsymbol{v}_1}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}, \quad \boldsymbol{v}_2^* = \boldsymbol{v}_2 - m_1\Phi(r)\frac{\boldsymbol{v}_2-\boldsymbol{v}_1}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}, \tag{5.63}$$

and the reverse velocity map $({}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2) = \psi^{-1}_{m_1,m_2,r}(\boldsymbol{v}_1, \boldsymbol{v}_2)$ as

$${}^*\boldsymbol{v}_1 = \boldsymbol{v}_1 - m_2\Phi(r)\frac{\boldsymbol{v}_2-\boldsymbol{v}_1}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}, \quad {}^*\boldsymbol{v}_2 = \boldsymbol{v}_2 + m_1\Phi(r)\frac{\boldsymbol{v}_2-\boldsymbol{v}_1}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}. \tag{5.64}$$

From these, we can calculate

$$|\boldsymbol{v}_2^* - \boldsymbol{v}_1^*| = \left(1 - \frac{(m_1+m_2)\Phi(r)}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}\right)|\boldsymbol{v}_2-\boldsymbol{v}_1| = |\boldsymbol{v}_2-\boldsymbol{v}_1| - (m_1+m_2)\Phi(r), \tag{5.65}$$

and

$$\left|D\psi^{-1}_{m_1,m_2,r}\right| = \left(1 + \frac{(m_1+m_2)\Phi(r)}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}\right)^{d-1}, \tag{5.66}$$

giving

$$\left|D\psi^{-1}_{m_1,m_2,r}\right| |{}^*\boldsymbol{v}_2 - {}^*\boldsymbol{v}_1| = \left(1 + \frac{(m_1+m_2)\Phi(r)}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}\right)^{d} |\boldsymbol{v}_2-\boldsymbol{v}_1|. \tag{5.67}$$

Note that considering the maps $\psi^{-1}_{m_1,m_2,r}(\boldsymbol{v}_1, \boldsymbol{v}_2) = ({}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2)$ and $\psi_{m_1,m_2,r}({}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2) = (\boldsymbol{v}_1, \boldsymbol{v}_2)$

we have

$$\begin{aligned}\left|D\psi^{-1}_{m_1,m_2,r}\right| &= \left(1+\frac{(m_1+m_2)\Phi(r)}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}\right)^{d-1}\\ &= \left(1+\frac{(m_1+m_2)\Phi(r)}{|{}^*\boldsymbol{v}_2-{}^*\boldsymbol{v}_1|-(m_1+m_2)\Phi(r)}\right)^{d-1}\\ &= \left(\frac{|{}^*\boldsymbol{v}_2-{}^*\boldsymbol{v}_1|}{|{}^*\boldsymbol{v}_2-{}^*\boldsymbol{v}_1|-(m_1+m_2)\Phi(r)}\right)^{d-1}\\ &= \left(1-\frac{(m_1+m_2)\Phi(r)}{|{}^*\boldsymbol{v}_2-{}^*\boldsymbol{v}_1|}\right)^{1-d} = \frac{1}{|D\psi|},\end{aligned} \tag{5.68}$$

as expected.

For clustering interactions, where $r < r^*_{m_1,m_2}(|\boldsymbol{u}|)$, the inner region trajectory approaches a relative velocity of $\boldsymbol{u}^+ = 0$. With conservation of momentum and mass, we have the forwards maps, $m_1' = m_1 + m_2$ and

$$\boldsymbol{v}_1' = \lambda_{m_1,m_2}(\boldsymbol{v}_1,\boldsymbol{v}_2) = \frac{m_1\boldsymbol{v}_1+m_2\boldsymbol{v}_2}{m_1+m_2}. \tag{5.69}$$

Note that the clustering maps have no r dependence for this interaction.

Similarly to in the one-dimensional case, define $\hat{\lambda}_{m_1,m_2,\boldsymbol{v}_2}(\boldsymbol{v}_1)$ to be the partial evaluation of $\lambda_{m_1,m_2}(\boldsymbol{v}_1,\boldsymbol{v}_2)$, at a specific value of $\boldsymbol{v}_2$. We can then define the inverse maps ${}'m_1 = m_1 - m_2$ and

$${}'\boldsymbol{v}_1 = \hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2}(\boldsymbol{v}_1) = \frac{m_1\boldsymbol{v}_1-m_2\boldsymbol{v}_2}{m_1-m_2}. \tag{5.70}$$

From these, we can calculate

$$\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2}\right| = \left(\frac{m_1}{m_1-m_2}\right)^d, \tag{5.71}$$

$$|\boldsymbol{v}_2-{}'\boldsymbol{v}_1| = \left|\frac{m_1(\boldsymbol{v}_2-\boldsymbol{v}_1)}{m_1-m_2}\right|, \tag{5.72}$$

and

$$|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|\left|D\hat{\lambda}^{-1}_{m_1-m_2,m_2,\boldsymbol{v}_2}\right| = |\boldsymbol{v}_2-\boldsymbol{v}_1|\left(\frac{m_1}{m_1-m_2}\right)^{d+1}. \tag{5.73}$$

5.3.1. Interaction kernels

For a power-law interaction kernel, $\phi(|\boldsymbol{x}|) = 1/|\boldsymbol{x}|^s$, we have

$$\Phi(r) = \int_{-\infty}^{\infty}\phi\left(\sqrt{r^2+x^2}\right)\,\mathrm{d}x = \int_{-\infty}^{\infty}\frac{1}{(r^2+x^2)^{\frac{s}{2}}}\,\mathrm{d}x = \frac{\Gamma((s-1)/2)\sqrt{\pi}}{\Gamma(s/2)r^{s-1}},$$

and hence

$$\Phi^{-1}\left(\frac{u}{m}\right) = \left(\frac{\Gamma((s-1)/2)\sqrt{\pi}m}{\Gamma(s/2)u}\right)^{\frac{1}{s-1}}. \tag{5.74}$$

In Chapter 7, we will perform numerical experiments considering specifically the inverse-cube and inverse-quintic interaction kernels, i.e. setting $s = 3$ and $s = 5$, respectively. For the

inverse-cube kernel, we have

$$\Phi(r) = \frac{2}{r^2}, \text{ and } \Phi^{-1}\left(\frac{u}{m}\right) = \left(\frac{2m}{u}\right)^{1/2}, \tag{5.75}$$

and for the inverse-quintic kernel

$$\Phi(r) = \frac{4}{3r^4}, \quad \Phi^{-1}\left(\frac{u}{m}\right) = \left(\frac{4m}{3u}\right)^{\frac{1}{4}}. \tag{5.76}$$

We will also consider the exponential interaction kernel, $\phi(|\boldsymbol{x}|) = e^{-|\boldsymbol{x}|}$, for which we have

$$\Phi(r) = \int_{-\infty}^{\infty} e^{-\sqrt{r^2+x^2}}\,\mathrm{d}x = \begin{cases} 2, & r = 0 \\ 2rK_1(r), & r > 0, \end{cases} \tag{5.77}$$

where K_1 is a modified Bessel function of the second kind. This does not have a closed-form analytical inverse, so it is necessary to either analytically approximate Φ^{-1} or to compute it numerically. This will be addressed further in Chapter 6.

5.3.2. PDE model (strong form)

Substituting these results into the general short-range clustering PDE model, (5.32), we get the PDE model for the Cucker-Smale interaction in strong form:

$$\begin{aligned}
&\frac{1}{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\left(\partial_t n(t;\boldsymbol{x}_1;\boldsymbol{v}_1;m_1) + \boldsymbol{v}_1\cdot\nabla_{\boldsymbol{x}_1} n(t;\boldsymbol{x}_1;\boldsymbol{v}_1;m_1)\right) \\
&\quad = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d}\int_0^{r_c} |\boldsymbol{v}_2-\boldsymbol{v}_1| \Bigg[\left(1+\frac{(m_1+m_2)\Phi(r)}{|\boldsymbol{v}_2-\boldsymbol{v}_1|}\right)^d \\
&\qquad\qquad \times n\left(t;\boldsymbol{x}_1;\boldsymbol{v}_1 - m_2\Phi(r)\frac{\boldsymbol{v}_2-\boldsymbol{v}_1}{|\boldsymbol{v}_2-\boldsymbol{v}_1|};m_1\right) n\left(t;\boldsymbol{x}_1;\boldsymbol{v}_2 + m_1\Phi(r)\frac{\boldsymbol{v}_2-\boldsymbol{v}_1}{|\boldsymbol{v}_2-\boldsymbol{v}_1|};m_2\right) \\
&\qquad\qquad\qquad - n\left(t;\boldsymbol{x}_1;\boldsymbol{v}_1;m_1\right) n\left(t;\boldsymbol{x}_1;\boldsymbol{v}_2;m_2\right)\Bigg] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2 \\
&\quad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \int_{\mathbb{R}^d}\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)} |\boldsymbol{v}_2-\boldsymbol{v}_1| \left(\frac{m_1}{m_1-m_2}\right)^{d+1} \\
&\qquad\qquad \times n\left(t;\boldsymbol{x}_1;\frac{m_1\boldsymbol{v}_1-m_2\boldsymbol{v}_2}{m_1-m_2};m_1-m_2\right) n\left(t;\boldsymbol{x}_1;\boldsymbol{v}_2;m_2\right)\, r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2.
\end{aligned} \tag{5.78}$$

As the only r dependence in the second RHS term is in the r^{d-2} factor, the $\mathrm{d}r$ integral can be computed, giving

$$\begin{aligned}
\int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)} r^{d-2}\,\mathrm{d}r &= \frac{1}{d-1}\left(r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)\right)^{d-1} \\
&= \frac{1}{d-1}\left(\Phi^{-1}\left(\frac{|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|}{m_1}\right)\right)^{d-1} = \frac{1}{d-1}\left(\Phi^{-1}\left(\frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{m_1-m_2}\right)\right)^{d-1},
\end{aligned} \tag{5.79}$$

where we have used $'\boldsymbol{v}$ from (5.70).

5.3.3. Mean cluster number evolution

For the Cucker-Smale interaction, we have from (5.61) that

$$r^*_{(1,2)} \equiv r^*_{m_1,m_2}(|\boldsymbol{v}_2 - \boldsymbol{v}_1|) = \Phi^{-1}\left(\frac{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}{m_1 + m_2}\right).$$

Substituting this into (5.51), the time evolution equation for the mean number of clusters, we get

$$\frac{1}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}}{\mathrm{d}t}\mu_K = -\frac{1}{2(d-1)}\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \int_{\mathbb{R}^d\times\mathbb{R}^d}\int_{\Omega} |\boldsymbol{v}_2 - \boldsymbol{v}_1|\, n\, n_{(2)} \times \left(\Phi^{-1}\left(\frac{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}{m_1 + m_2}\right)\right)^{d-1} \mathrm{d}\boldsymbol{x}_1\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \quad (5.80)$$

For the inverse-cube interaction kernel, we have from (5.75) that

$$\Phi^{-1}\left(\frac{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}{m_1 + m_2}\right) = \left(\frac{2(m_1 + m_2)}{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}\right)^{1/2}.$$

As such, considering specifically a spatially uniform system with the inverse-cube interaction kernel, the evolution equation for the mean number of clusters is

$$\frac{2^{(3-d)/2}(d-1)|\Omega|}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}}{\mathrm{d}t}\mu_K = -\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \int_{\mathbb{R}^d\times\mathbb{R}^d} (m_1 + m_2)^{(d-1)/2}|\boldsymbol{v}_2 - \boldsymbol{v}_1|^{(3-d)/2} \times\; n(t;\boldsymbol{v}_1;m_1)\, n(t;\boldsymbol{v}_2;m_2)\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \quad (5.81)$$

If we assume the velocity distribution is independent of the cluster masses (a similar approximation to the *Clusterzahlansatz*) we have that

$$n(t;\boldsymbol{v};m) = \mu_K p(t;\boldsymbol{v};m) \approx \mu_K p(t;\boldsymbol{v})p(t;m). \quad (5.82)$$

Substituting this into (5.83) gives

$$\frac{2^{(3-d)/2}(d-1)|\Omega|}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}}{\mathrm{d}t}\mu_K = -\Lambda_K\mu_K^2 \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} (m_1 + m_2)^{(d-1)/2}\, p(t;m_1)p(t;m_2) \times \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2 - \boldsymbol{v}_1|^{(3-d)/2}\, p(t;\boldsymbol{v}_1)\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \quad (5.83)$$

Following a similar procedure to [80], using Jensen's inequality, we have that

$$\int_{R^d} |\boldsymbol{v}_1 - \boldsymbol{v}_2|^{(3-d)/2}\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2 \geq \left| \int_{R^d} (\boldsymbol{v}_1 - \boldsymbol{v}_2)\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2 \right|^{(3-d)/2}$$

as $p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2$ is a probability measure and the function $|\cdot|^{(3-d)/2}$ is convex for $d \leq 3$ (if $d \geq 3$ the inverse-cube kernel will not satisfy our definition to be short-ranged in any case). Using the fact that p is normalised to a probability distribution and that (by using conservation of momentum) we can set the mean velocity to 0, we then have

$$\int_{R^d} (\boldsymbol{v}_1 - \boldsymbol{v}_2)\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2 = \boldsymbol{v}_1 \int_{R^d} p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2 - \int_{R^d} \boldsymbol{v}_2\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2 = \boldsymbol{v}_1.$$

As such, it follows that

$$\int_{\mathbb{R}^d \times \mathbb{R}^d} |\boldsymbol{v}_2 - \boldsymbol{v}_1|^{(3-d)/2}\, p(t;\boldsymbol{v}_1)\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2 \mathrm{d}\boldsymbol{v}_1 \geq \int_{\mathbb{R}^d} |\boldsymbol{v}_1|^{(3-d)/2} p(t;\boldsymbol{v}_1)\,\mathrm{d}\boldsymbol{v}_1. \tag{5.84}$$

Using Jensen's inequality again, we have that

$$\int_{\mathbb{R}^d} |\boldsymbol{v}_1|^{(3-d)/2} p(t;\boldsymbol{v}_1)\,\mathrm{d}\boldsymbol{v}_1 \geq \left(\int_{\mathbb{R}^d} |\boldsymbol{v}_1|^2 p(t;\boldsymbol{v}_1)\,\mathrm{d}\boldsymbol{v}_1 \right)^{(3-d)/4} = (2E(t))^{(3-d)/4},$$

where we have used the definition from (4.12) that the mean kinetic energy per particle is given by

$$E(t) = \frac{1}{2} \int_{\mathbb{R}^d} |\boldsymbol{v}_1|^2 p(t;\boldsymbol{v}_1)\,\mathrm{d}\boldsymbol{v}_1.$$

Following a similar procedure for the mass sums, we have from Jensen's inequality that

$$\begin{aligned} \sum_{m_2=1}^{N-m_1-\mu_K+1} (m_1+m_2)^{(d-1)/2} p(t;m_2) &\geq \left(\sum_{m_2=1}^{N-m_1-\mu_K+1} (m_1+m_2) p(t;m_2) \right)^{(d-1)/2} \\ &\approx \left(\sum_{m_2=1}^{N} (m_1+m_2) p(t;m_2) \right)^{(d-1)/2} \\ &= \left(m_1 + \sum_{m_2=1}^{N} m_2 p(t;m_2) \right)^{(d-1)/2}, \end{aligned} \tag{5.85}$$

where, in the second step, we have assumed the cluster masses are small compared to N, such that we can ignore the restriction on the sum. Using the mean mass approximation (see §3.3.5)

$$\sum_{m_2=1}^{N} m_2 p(t;m_2) \approx \frac{N}{\mu_K},$$

we then, given the approximations made, have that

$$\sum_{\substack{m_1,m_2=1 \\ m_1+m_2 \\ \leq N-\mu_K+2}}^{N} (m_1+m_2)^{(d-1)/2}\, p(t;m_1) p(t;m_2) \geq \sum_{m_1=1}^{N} \left(m_1 + \frac{N}{\mu_K} \right)^{(d-1)/2} p(t;m_1). \tag{5.86}$$

Using Jensen's inequality and the cluster mass approximation again, it follows that

$$\begin{aligned}\sum_{m_1=1}^{N}\left(m_1+\frac{N}{\mu_K}\right)^{(d-1)/2}p(t;m_1) &\geq \left(\sum_{m_1=1}^{N}\left(m_1+\frac{N}{\mu_K}\right)p(t;m_1)\right)^{(d-1)/2}\\ &= \left(\frac{N}{\mu_K}+\sum_{m_1=1}^{N}m_1p(t;m_1)\right)^{(d-1)/2}\\ &\approx \left(\frac{2N}{\mu_K}\right)^{(d-1)/2}\end{aligned} \tag{5.87}$$

Substituting these results into (5.83), we have an approximate upper bound on the reduction in mean cluster numbers for the Cucker-Smale model with inverse-cube interaction:

$$\frac{2^{(5-3d)/4}(d-1)|\Omega|}{N^{(d-1)/2}\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\frac{\mathrm{d}}{\mathrm{d}t}\mu_K \leq -\Lambda_K\mu_K^{(5-d)/2}E^{(3-d)/4}. \tag{5.88}$$

Setting $d=2$ this becomes

$$\frac{|\Omega|}{2^{5/4}\sqrt{N}\varepsilon}\frac{\mathrm{d}}{\mathrm{d}t}\mu_K \leq -\Lambda_K\mu_K^{3/2}E^{1/4}. \tag{5.89}$$

This is not a closed equation in $\mu_K(t)$ as we would still need to approximate $E(t)$ and $\sigma_K^2(t)$ (as part of $\Lambda_K(t)$) concurrently. However, we believe it could be the basis for further investigation.

However, if we set $d=3$, which means the inverse-cube interaction kernel is right on the boundary of what we would consider a short-range interaction, we have

$$\frac{|\Omega|}{2N\varepsilon^2\pi}\frac{\mathrm{d}}{\mathrm{d}t}\mu_K \approx -\Lambda_K\mu_K, \tag{5.90}$$

which is a simpler approximation as it is independent of $E(t)$. This is an approximation and not an upper bound as, in this case, the steps using Jensen's inequality are identities. This generalises to any case where the interaction kernel has power the same as the dimension, i.e. where $s=d$. From (5.74), we have that

$$\left(\Phi^{-1}\left(\frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{m_1+m_2}\right)\right)^{d-1} = \left(\frac{\Gamma((d-1)/2)\sqrt{\pi}(m_1+m_2)}{\Gamma(d/2)|\boldsymbol{v}_2-\boldsymbol{v}_1|}\right)^{(d-1)/(s-1)}.$$

If we set $s=d$, the power will be 1, meaning that the $|\boldsymbol{v}_2-\boldsymbol{v}_1|$ here will cancel with the other instance in (5.51), and the final equation will be independent of E.

5.3.4. Energy scaling

The mean kinetic energy per particle is given by the total kinetic energy of the system, normalised by N. That is,

$$E(t)=\frac{1}{2N}\sum_{m=1}^{N}\int_{\mathbb{R}^d\times\Omega}m|\boldsymbol{v}|^2n(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}=\frac{1}{2N}\langle n,m|\boldsymbol{v}|^2\rangle. \tag{5.91}$$

Note that this is approximately equivalent to the definition given in terms of the probability density in (4.12) (given the approximations made in the *Clusterzahlansatz*) as

$$\begin{aligned}\frac{1}{2N}\sum_{m=1}^{N}\int_{\mathbb{R}^d\times\Omega} m|\boldsymbol{v}|^2 n(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v} &= \frac{\mu_K}{2N}\sum_{m=1}^{N}\int_{\mathbb{R}^d\times\Omega} m|\boldsymbol{v}|^2 p(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\\ &\approx \frac{\mu_K}{2N}\sum_{m=1}^{N} m\,p(t;m)\int_{\mathbb{R}^d\times\Omega}|\boldsymbol{v}|^2 p(t;\boldsymbol{x};\boldsymbol{v})\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v}\\ &\approx \frac{1}{2}\sum_{m=1}^{N}\int_{\mathbb{R}^d\times\Omega}|\boldsymbol{v}|^2 p(t;\boldsymbol{x};\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{x}\mathrm{d}\boldsymbol{v},\end{aligned} \tag{5.92}$$

where, in the second step, we have assumed the cluster velocity and mass distributions are independent and, in the third step, used the estimate from §3.3.5 that

$$\sum_{m=1}^{N} m\,p(t;m) \approx \frac{N}{\mu_K}.$$

As such, we can find the time evolution of the mean kinetic energy by taking $\varphi(\boldsymbol{v};m) = \frac{1}{2N}m|\boldsymbol{v}|^2$, in (5.50), the weak form of our short-range PDE model. With this definition of φ, using the forwards velocity and mass maps for the Cucker-Smale model ((5.63) and (5.69)), we have

$$\begin{aligned}\varphi^* + \varphi^*_{(2)} - \varphi - \varphi_{(2)} &= \frac{1}{2N}\left(m_1\left(|\boldsymbol{v}_1^*|^2 - |\boldsymbol{v}_1|^2\right) + m_2\left(|\boldsymbol{v}_2^*|^2 - |\boldsymbol{v}_2|^2\right)\right)\\ &= \frac{1}{2N}m_1m_2\Phi(r)\left((m_1+m_2)\Phi(r) - 2|\boldsymbol{v}_2-\boldsymbol{v}_1|\right),\end{aligned} \tag{5.93}$$

and

$$\begin{aligned}\varphi' - \varphi - \varphi_{(2)} &= \frac{1}{2N}\left((m_1+m_2)|\boldsymbol{v}_1'|^2 - m_1|\boldsymbol{v}_1|^2 - m_2|\boldsymbol{v}_2|^2\right)\\ &= -\frac{1}{2N}\frac{m_1m_2}{(m_1+m_2)}|\boldsymbol{v}_2-\boldsymbol{v}_1|^2.\end{aligned} \tag{5.94}$$

As $\varphi' - \varphi - \varphi_{(2)}$ is independent of the impact parameter r in this case, we can compute the second $\mathrm{d}r$ integral in (5.50) to give

$$\int_0^{r^*_{(1,2)}} r^{d-2}\,\mathrm{d}r = \frac{1}{d-1}\left(r^*_{(1,2)}\right)^{d-1} = \frac{1}{d-1}\left(\Phi^{-1}\left(\frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{m_1+m_2}\right)\right)^{d-1}.$$

Substituting these results in (5.50) and assuming a spatially uniform system, we have

$$\begin{aligned}
&\frac{4N|\Omega|}{\varepsilon^{d-1}\,|\mathbb{S}^{d-2}|}\frac{\mathrm{d}E}{\mathrm{d}t} \\
&\quad= \Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} m_1 m_2 \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|\, n\, n_{(2)} \\
&\qquad\qquad \times \int_{r^*_{(1,2)}}^{r_c} \Phi(r)\left((m_1+m_2)\Phi(r) - 2|\boldsymbol{v}_2-\boldsymbol{v}_1|\right) r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1 \\
&\quad\quad - \frac{1}{d-1}\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \frac{m_1 m_2}{m_1+m_2} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|^3\, n\, n_{(2)} \left(\Phi^{-1}\left(\frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{m_1+m_2}\right)\right)^{d-1} \mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1.
\end{aligned} \tag{5.95}$$

For the inverse-cube interaction kernel in two dimensions, this becomes

$$\begin{aligned}
&\frac{2N|\Omega|}{\varepsilon}\frac{\mathrm{d}E}{\mathrm{d}t} \\
&\quad= 4\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} m_1 m_2 \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|\, n\, n_{(2)} \\
&\qquad\qquad \times \int_{r^*_{(1,2)}}^{r_c} \left(\frac{(m_1+m_2)}{r^4} - \frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{r^2}\right) \mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1 \\
&\quad\quad - \sqrt{2}\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \frac{m_1 m_2}{(m_1+m_2)^{1/2}} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|^{5/2}\, n\, n_{(2)}\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1.
\end{aligned} \tag{5.96}$$

Evaluating the dr integral in the first RHS term, we get

$$\begin{aligned}
\int_{r^*_{(1,2)}}^{r_c} \left(\frac{(m_1+m_2)}{r^4} - \frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{r^2}\right) \mathrm{d}r &\approx \frac{(m_1+m_2)}{3(r^*_{(1,2)})^3} - \frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{r^*_{(1,2)}} \\
&= \frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|^{3/2}}{3\cdot 2^{3/2}(m_1+m_2)^{1/2}} - \frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|^{3/2}}{\sqrt{2}(m_1+m_2)^{1/2}} \\
&= -\frac{5}{6\sqrt{2}}\frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|^{3/2}}{(m_1+m_2)^{1/2}},
\end{aligned} \tag{5.97}$$

where we have taken the cutoff $r_c \to \infty$. Substituting this into (5.96), we note that the two RHS terms can now be combined to give

$$\frac{3N|\Omega|}{4\sqrt{2}\varepsilon}\frac{\mathrm{d}E}{\mathrm{d}t} = -\Lambda_K \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \frac{m_1 m_2}{(m_1+m_2)^{1/2}} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|^{5/2}\, n\, n_{(2)}\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \tag{5.98}$$

If we assume the velocity and mass distributions are independent, as in (5.82), we can write

this as

$$\frac{3N|\Omega|}{4\sqrt{2}\varepsilon}\frac{\mathrm{d}E}{\mathrm{d}t} = -\Lambda_K\mu_K^2 \sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \frac{m_1 m_2}{(m_1+m_2)^{1/2}}\, p(t;m_1)\, p(t;m_2) \times \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|^{5/2}\, p(t;\boldsymbol{v}_1)\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \quad (5.99)$$

Using Jensen's inequality in a similar fashion to the manipulations in §5.3.3, we have that

$$\int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|^{5/2}\, p(t;\boldsymbol{v}_1)\, p(t;\boldsymbol{v}_2)\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1 \geq \int_{\mathbb{R}^d} |\boldsymbol{v}_1|^{5/2}\, p(t;\boldsymbol{v}_1)\,\mathrm{d}\boldsymbol{v}_1 \geq (2E)^{5/4}, \quad (5.100)$$

and

$$\sum_{\substack{m_1,m_2=1\\ m_1+m_2\\ \leq N-\mu_K+2}}^{N} \frac{m_1 m_2}{(m_1+m_2)^{1/2}}\, p(t;m_1)\, p(t;m_2) \geq \frac{N}{\mu_K}\sum_{m_1=1}^{N} \frac{m_1}{(m_1+N/\mu_K)^{1/2}}\, p(t;m_1) \geq \frac{1}{\sqrt{2}}\left(\frac{N}{\mu_K}\right)^{3/2}. \quad (5.101)$$

Substituting these into (5.98), we get

$$\frac{3|\Omega|}{8\cdot 2^{1/4}\sqrt{N}\varepsilon}\frac{\mathrm{d}E}{\mathrm{d}t} \leq -\Lambda_K\mu_K^{1/2}E^{5/4}. \quad (5.102)$$

Closed evolution equations

Combining (5.102) with (5.89), we note that if we can approximate Λ_K, we can now produce a closed equation in the combined quantity $G(t) := \mu_K^2(t)E(t)$, as

$$\begin{aligned}\frac{\mathrm{d}G}{\mathrm{d}t} &= 2\mu_K E\frac{\mathrm{d}\mu_K}{\mathrm{d}t} + \mu_K^2\frac{\mathrm{d}E}{\mathrm{d}t}\\ &\leq -\frac{\sqrt{N}\varepsilon}{|\Omega|}\Lambda_K\left(2\cdot 2^{5/4}\mu_K^{5/2}E^{5/4} + \frac{8\cdot 2^{1/4}}{3}\mu_K^{5/2}E^{5/4}\right)\\ &= -\frac{20\cdot 2^{1/4}\sqrt{N}\varepsilon}{3|\Omega|}\Lambda_K\mu_K^{5/2}E^{5/4}\\ &= -\frac{20\cdot 2^{1/4}\sqrt{N}\varepsilon}{3|\Omega|}\Lambda_K G^{5/4}.\end{aligned} \quad (5.103)$$

Integrating this, we get

$$G(t) \leq \left(G(0)^{-1/4} + \frac{5\cdot 2^{1/4}\sqrt{N}\varepsilon}{3|\Omega|}\int_0^t \Lambda_K(t)\,\mathrm{d}t\right)^{-4}, \quad (5.104)$$

where the integral of $\Lambda_K(t)$ could be determined numerically or with other approximations on the P_k distribution. The simplest approximation is that if N is large and $\mu_K \gg \sigma_K^2$ then $\Lambda_K \approx 1$. We will test the validity of this result numerically in §7.3.2.

Substituting the definition of G into (5.102) we get

$$\frac{3|\Omega|}{8 \cdot 2^{1/4}\sqrt{N\varepsilon}} \frac{\mathrm{d}E}{\mathrm{d}t} \leq -\Lambda_K G^{1/4}\, E. \tag{5.105}$$

Using the bound (5.104) and making the approximation $\Lambda_K(t) = 1$ we can solve this equation for $E(t)$, giving

$$E(t) \leq E(0) \left(1 + \frac{5 \cdot 2^{1/4}\sqrt{N\varepsilon}}{3|\Omega|} G(0)^{1/4}\, t\right)^{-8/5}. \tag{5.106}$$

Substituting $G(0) = \mu_K^2(0)E(0)$ and taking the monodisperse initial condition $\mu_K(0) = N$ we can write this as

$$E(t) \leq E(0) \left(1 + \frac{5 \cdot 2^{1/4} N\varepsilon}{3|\Omega|} E(0)^{1/4}\, t\right)^{-8/5}. \tag{5.107}$$

Note that this bound permits a similar time scaling to in §5.1.7. If we define

$$t = \frac{|\Omega|}{2N\varepsilon(2E(0))^{1/4}} t' \tag{5.108}$$

we can write this as

$$E(t) \leq E(0) \left(1 + \frac{5}{6}\, t'\right)^{-8/5}. \tag{5.109}$$

This is the same as the time scale τ defined in (5.41), except with $(2E(0))^{1/4}$ in place of our representative velocity scale $\bar{v}$. This quantity is related to the velocity distribution through

$$(2E(0))^{1/4} = \left(\sum_{m=1}^{N} \int_{\mathbb{R}^d} |\boldsymbol{v}|^2 p(t; \boldsymbol{v}; m)\right)^{1/4},$$

using the definition of E in terms of p from (5.92). Specifically, it is the fourth root of the mean square cluster velocity.

Similarly, substituting the definition of G into (5.89), we get

$$\frac{|\Omega|}{2^{5/4}\sqrt{N\varepsilon}} \frac{\mathrm{d}}{\mathrm{d}t} \mu_K \leq -\Lambda_K G^{1/4} \mu_K. \tag{5.110}$$

Using the bound (5.104), substituting the time scale (5.108), and using the initial condition $\mu_K(0) = N$, we can solve this to give

$$\mu_K(t) \leq N \left(1 + \frac{5}{6}\, t'\right)^{-6/5}. \tag{5.111}$$

5.3.5. Smoluchowski coagulation equation

In §5.2.3, we derived a time evolution equation for the mass distribution $n(t; m)$, (5.56), that was in the form of the Smoluchowski coagulation equation. The coagulation kernel is given in

(5.57), which, for the Cucker-Smale interaction in a spatially uniform system, becomes

$$B(m_1,m_2)=\frac{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}{(d-1)|\Omega|}\Lambda_K\int_{\mathbb{R}^d\times\mathbb{R}^d}|\boldsymbol{v}_2-\boldsymbol{v}_1|p(t;\boldsymbol{v}_1\,|\,m_1)\,p(t;\boldsymbol{v}_2\,|\,m_2)\\ \times\left(\Phi^{-1}\left(\frac{|\boldsymbol{v}_2-\boldsymbol{v}_1|}{m_1+m_2}\right)\right)^{d-1}\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1. \quad (5.112)$$

If we assume a power-law interaction kernel, i.e. $\phi(|\boldsymbol{x}|)=1/|\boldsymbol{x}|^s$, and take the conditional velocity distributions to be independent of m, this becomes

$$B(m_1,m_2)=b_1\Lambda_K(m_1+m_2)^{(d-1)/(s-1)}\int_{\mathbb{R}^d\times\mathbb{R}^d}|\boldsymbol{v}_2-\boldsymbol{v}_1|^{(s-d)/(s-1)}p(t;\boldsymbol{v}_1)\,p(t;\boldsymbol{v}_2)\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1, \quad (5.113)$$

where

$$b_1=\frac{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}{(d-1)|\Omega|}\left(\frac{\Gamma((d-1)/2)\sqrt{\pi}}{\Gamma(d/2)}\right)^{(d-1)/(s-1)}.$$

If $s\geq d$, a necessary condition for the power-law interaction to be short-ranged, then we can follow a similar procedure to §5.3.3, using Jensen's inequality to give

$$\int_{\mathbb{R}^d\times\mathbb{R}^d}|\boldsymbol{v}_2-\boldsymbol{v}_1|^{(s-d)/(s-1)}p(t;\boldsymbol{v}_1)\,p(t;\boldsymbol{v}_2)\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1\leq\int_{\mathbb{R}^d}|\boldsymbol{v}_1|^{(s-d)/(s-1)}p(t;\boldsymbol{v}_1)\mathrm{d}\boldsymbol{v}_1\\ \leq(2E(t))^{(s-d)/(2s-2)}. \quad (5.114)$$

As such, we have an estimate for the mass dependence of the coagulation kernel and an upper bound for its dependence on the mean particle energy as

$$B(m_1,m_2)=b_1\Lambda_K(m_1+m_2)^{(d-1)/(s-1)}(2E(t))^{(s-d)/(2s-2)}. \quad (5.115)$$

The energy dependence here is expected, as the cluster interaction rate decreases with the variance in the velocity distribution over time. However, this means the coagulation kernel is not purely a function of the particle masses, and this cannot be reduced to a closed system in the form of Smoluchowski's coagulation equation. We suggest that with an appropriate velocity rescaling, it may be possible to define a similarity solution in velocity space with collision kernels that are constant as a function of the rescaled velocity. This is left for future investigation.

5.3.6. Similarity scaling

As the velocity distribution for the Cucker-Smale model will approach the Dirac delta distribution as $t\to\infty$, to numerically simulate accurately to large times it is useful to implement a velocity rescaling. As noted in §2.3.2, similarity scalings are also commonly used to derive theoretical results about the velocity distribution for Boltzmann-type equations. As is standard for a similarity scaling, the scaling discussed in §2.3.2 (for the non-clustering inelastic Boltzmann equation with a hard sphere interaction) relies on the homogeneity of the collision operator under a rescaling of the velocity variable. This ensures the time dependence in the collision

operator can be factored out when changing variables to scaled velocity $\tilde{\boldsymbol{v}} = \mathcal{V}(t)\boldsymbol{v}$, allowing a time-independent equation for a similarity profile to be derived. However, for our clustering model of the Cucker-Smale interaction, the interaction operator $Q_I(n,n)$ is not homogeneous in velocity. In the spatially homogenous case, substituting the scaled velocity $\tilde{\boldsymbol{v}} = \mathcal{V}(t)\boldsymbol{v}$, we get

$$\begin{aligned}
&Q_I(n,n) \\
&\quad = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d}\int_0^{r_c} |\boldsymbol{v}_2 - \boldsymbol{v}_1| \left[\left(1 + \frac{(m_1+m_2)\Phi(r)}{|\boldsymbol{v}_2 - \boldsymbol{v}_1|}\right)^d {}^*n\,{}^*n_{(2)} - n\, n_{(2)}\right] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2 \\
&\quad = \frac{1}{\mathcal{V}(t)^{d+1}}\Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d}\int_0^{r_c} |\tilde{\boldsymbol{v}}_2 - \tilde{\boldsymbol{v}}_1| \\
&\qquad\qquad \times \left[\left(1 + \frac{\mathcal{V}(t)(m_1+m_2)\Phi(r)}{|\tilde{\boldsymbol{v}}_2 - \tilde{\boldsymbol{v}}_1|}\right)^d {}^*\tilde{n}\,{}^*\tilde{n}_{(2)} - \tilde{n}\,\tilde{n}_{(2)}\right] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\tilde{\boldsymbol{v}}_2,
\end{aligned} \tag{5.116}$$

where we have used the shorthand $\tilde{n} \equiv n(t; \tilde{\boldsymbol{v}}_1/\mathcal{V}(t); m_1)$. The $\mathcal{V}(t)$ factor in the Jacobian cannot be removed, leaving the operator explicitly depending on t. This is expected as, unlike the constant restitution coefficient cases that are commonly studied in this setting, the velocity changes caused by a non-clustering interaction in the Cucker-Smale model have a fixed value

$$|\boldsymbol{v}_2^* - \boldsymbol{v}_1^*| - |\boldsymbol{v}_2 - \boldsymbol{v}_1| = -(m_1+m_2)\Phi(r)$$

(see (5.65)), which is independent of the velocity scale. One way to derive a homogenous interaction operator might be to scale mass along with velocity as $\tilde{m} = \mathcal{V}(t)m$, giving

$$\begin{aligned}
Q_I(n,n)(\tilde{\boldsymbol{v}}; \tilde{m}) = \frac{1}{\mathcal{V}(t)^{d+1}}\Lambda_K \sum_{\tilde{m}_2=1}^{N-\tilde{m}_1-\mu_K+2} &\int_{\mathbb{R}^d}\int_0^{r_c} |\tilde{\boldsymbol{v}}_2 - \tilde{\boldsymbol{v}}_1| \\
&\times \left[\left(1 + \frac{(\tilde{m}_1+\tilde{m}_2)\Phi(r)}{|\tilde{\boldsymbol{v}}_2 - \tilde{\boldsymbol{v}}_1|}\right)^d {}^*\tilde{\tilde{n}}\,{}^*\tilde{\tilde{n}}_{(2)} - \tilde{\tilde{n}}\,\tilde{\tilde{n}}_{(2)}\right] r^{d-2}\,\mathrm{d}r\,\mathrm{d}\tilde{\boldsymbol{v}}_2,
\end{aligned} \tag{5.117}$$

where $\tilde{\tilde{n}} \equiv n(t; \tilde{\boldsymbol{v}}/\mathcal{V}(t); \tilde{m}/\mathcal{V}(t))$. However, a similarity solution would not be expected in this scaling as, with increasing t, the mean mass increases, while the velocity distribution variance decreases towards the Dirac delta. A scaling that allows a steady state in the distribution of $\tilde{\boldsymbol{v}}$, similar to $\mathcal{V}(t) = 1 + t$, would result in the mean mass increasing even faster with time in scaled variables. As such, for numerical simulation purposes, it may be desirable to only scale velocity, as shown in (5.116), leaving the evolution equation in scaled variables with a non-autonomous interaction operator.

Nevertheless, at later times, when the variance of the velocity distribution has decreased far enough toward 0, clustering interactions begin to dominate non-clustering interactions. As both the clustering loss and gain operators Q_C^- and Q_C^+ are homogenous in $\boldsymbol{v}$, we may expect to be able to develop a similarity scaling in this regime. As the clustering gain term, Q_C^+, only contains a ratio of masses as a factor (see (5.78)), we could scale mass independently to ensure the mean mass stays constant in scaled variables (as is commonly done for the Smoluchowski

coagulation equation, see [1] §2.4), a necessary condition for a similarity solution to exist.

6. Fourier-Galerkin spectral method for clustering PDE model

To implement our clustering kinetic PDE model, we adapt one of the most popular classes of numerical method for Boltzmann style equations: the Fourier-Galerkin spectral method. This method involves approximating our density as a truncated Fourier series and reducing the system to evolution equations for the Fourier weights. The direct version of this method was developed for the elastic Boltzmann equation in [90, 89]. It was extended to quasi-elastic cases in one dimension in [87] and extended to inelastic non-clustering interactions in higher dimensions in [46, 47] for a granular gases model and in [54] for a variable hard sphere model. A new version of the implementation that makes use of the fast Fourier transform to improve computational efficiency was introduced for elastic interactions in [53], adapted to an inelastic granular gases model in [67], and extended to more models of granular materials with a new GPU implementation in [29]. Our main additions to these approaches in the literature are the introduction of the mass dimension into the model and the treatment of the clustering interactions.

6.1. Direct Fourier spectral method

To derive our Fourier-Galerkin spectral method, we begin with the weak form of the PDE model but without summing over mass, i.e

$$\frac{1}{\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|}\partial_t\langle n,\varphi\rangle_{x,v} = \langle Q_I(n,n),\varphi\rangle_{x,v} + \left\langle Q_C^+(n,n),\varphi\right\rangle_{x,v}, \tag{6.1}$$

with the inner products on the RHS given in (5.53) and (5.54), respectively. We will consider the spatially uniform case:

$$\begin{aligned} N\bar{v}\tau\,\partial_t \int_{\mathbb{R}^d} n\,\varphi\,\mathrm{d}\boldsymbol{v}_1 \\ = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-\boldsymbol{v}_1|\, n\, n_{(2)} \left[\int_{r^*_{(1,2)}}^{r_c} \varphi^*\, r^{d-2}\,\mathrm{d}r - \int_0^{r_c} \varphi\, r^{d-2}\,\mathrm{d}r\right] \mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1 \\ + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_2-{}'\boldsymbol{v}_1|\, {}'n\, n_{(2)} \int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{v}_2-{}'\boldsymbol{v}_1|)} \varphi\, r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{v}_2\mathrm{d}{}'\boldsymbol{v}_1, \end{aligned} \tag{6.2}$$

where we have used the time scale $\tau = \frac{|\Omega|}{N\varepsilon^{d-1}\left|\mathbb{S}^{d-2}\right|\bar{v}}$ defined in §5.1.7, to simplify the LHS coefficient. If we were to scale n, $\boldsymbol{v}$ and t, this coefficient could be absorbed, but we will

continue to work with the unscaled equation to aid in physical interpretability. Making the transformation ${}'\boldsymbol{v}_1 \to \boldsymbol{v}_1$ in the second RHS term, we must also make the change $\boldsymbol{v}_1 \to \boldsymbol{v}_1'$ and hence $\varphi \to \varphi'$ (this is the same transformation made in deriving (5.48)). Changing variables to relative velocity $\boldsymbol{u} = \boldsymbol{v}_2 - \boldsymbol{v}_1$ and writing out the arguments in full (except dropping t's for conciseness), this becomes

$$\begin{aligned}
&N\bar{v}\tau\partial_t \int_{\mathbb{R}^d} n(\boldsymbol{v}_1; m_1)\, \varphi(\boldsymbol{v}_1; m_1)\, \mathrm{d}\boldsymbol{v}_1 \\
&\quad = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{u}|\, n(\boldsymbol{v}_1; m_1)\, n(\boldsymbol{v}_1+\boldsymbol{u}; m_1) \\
&\qquad\qquad \times \left[\int_{r^*_{m_1,m_2}(|\boldsymbol{u}|)}^{r_c} \varphi(\boldsymbol{v}_1^*; m_1)\, r^{d-2}\, \mathrm{d}r - \int_0^{r_c} \varphi(\boldsymbol{v}_1; m_1)\, r^{d-2}\, \mathrm{d}r\right] \mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{v}_1 \\
&\quad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{u}|\, n(\boldsymbol{v}_1; m_1-m_2)\, n(\boldsymbol{v}_1+\boldsymbol{u}; m_2) \\
&\qquad\qquad \times \int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{u}|)} \varphi(\boldsymbol{v}_1'; m_1)\, r^{d-2}\, \mathrm{d}r\, \mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{v}_1.
\end{aligned} \tag{6.3}$$

To approximate $n(\boldsymbol{v}; m)$ as a Fourier series, we first assume that we can restrict the velocity support to a ball of radius R about the origin, i.e. $\mathrm{Supp}(n(\boldsymbol{v}; m)) = \mathcal{B}_R \otimes \{1, \ldots, N\}$, with minimal error. We then consider the cubic computational velocity domain $D_L = [-L, L]^d$, with $L > R$ and approximate n as

$$\check{n}(\boldsymbol{v}; m) = \begin{cases} n(\boldsymbol{v}; m), & \boldsymbol{v} \in \mathcal{B}_R \\ 0, & \boldsymbol{v} \in D_L \backslash \mathcal{B}_R, \end{cases} \tag{6.4}$$

with $\check{n}$ periodically extended over copies of D_L. Note that L must be large enough to avoid aliasing (when the domains of $\boldsymbol{v}_1 + \boldsymbol{u}$, $\boldsymbol{v}_1^*$ or $\boldsymbol{v}_1'$ intersect with non-zero areas in the periodic extension), but the specific value of the restriction depends on the interaction considered so we will delay further consideration until analysing specific cases below (see Figure 6.1 for an example). Also note that if $\boldsymbol{v}_1$, $\boldsymbol{v}_2 \in \mathcal{B}_R$, then

$$|\boldsymbol{u}| = |\boldsymbol{v}_2 - \boldsymbol{v}_1| \leq |\boldsymbol{v}_2| + |\boldsymbol{v}_1| \leq 2R,$$

so we can assume that $\boldsymbol{u} \in \mathcal{B}_{2R}$.

With this construction, we can approximate $\check{n}$ with a truncated Fourier series in $\boldsymbol{v}$ with $2V + 1$ modes, indexed by mass m:

$$\check{n}(\boldsymbol{v}; m) \approx \check{n}_V(\boldsymbol{v}; m) = \sum_{\boldsymbol{\kappa}=-V}^{V} \hat{n}_{\boldsymbol{\kappa},m} e^{i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}},$$

where the Fourier coefficients are defined through

$$\hat{n}_{\boldsymbol{\kappa},m} = \frac{1}{(2L)^d} \int_{D_L} n(\boldsymbol{v}; m) e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}}\, \mathrm{d}\boldsymbol{v}.$$

Substituting this approximation, truncating the $\mathrm{d}\boldsymbol{v}_1$ integral to D_L, the $\boldsymbol{u}$ integral to $\mathcal{B}_{2R}$ and taking $\varphi(\boldsymbol{v};m) = e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}}$, we get

$$\begin{aligned}
&N\bar{v}\tau(2L)^d\partial_t\hat{n}_{\boldsymbol{\kappa},m_1} \\
&\quad= \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \sum_{\boldsymbol{j},\boldsymbol{l}=-V}^{V} \hat{n}_{\boldsymbol{j},m_1}\hat{n}_{\boldsymbol{l},m_2} \int_{D_L\times\mathcal{B}_{2R}} |\boldsymbol{u}| e^{i\frac{\pi}{L}\boldsymbol{j}\cdot\boldsymbol{v}_1} e^{i\frac{\pi}{L}\boldsymbol{l}\cdot(\boldsymbol{v}_1+\boldsymbol{u})} \\
&\qquad\qquad \times \left[\int_{r^*_{m_1,m_2}(|\boldsymbol{u}|)}^{r_c} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}_1^*}\, r^{d-2}\,\mathrm{d}r - \int_0^{r_c} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}}\, r^{d-2}\,\mathrm{d}r\right] \mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{v}_1 \\
&\quad\quad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \sum_{\boldsymbol{j},\boldsymbol{l}=-V}^{V} \hat{n}_{\boldsymbol{j},m_1-m_2}\hat{n}_{\boldsymbol{l},m_2} \int_{D_L\times\mathcal{B}_{2R}} |\boldsymbol{u}| e^{i\frac{\pi}{L}\boldsymbol{j}\cdot\boldsymbol{v}_1} e^{i\frac{\pi}{L}\boldsymbol{l}\cdot(\boldsymbol{v}_1+\boldsymbol{u})} \\
&\qquad\qquad \times \int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{u}|)} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}_1'}\, r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{v}_1.
\end{aligned} \tag{6.5}$$

Assuming the post-interaction velocities can be written as

$$\boldsymbol{v}_1^* = \boldsymbol{v}_1 + \boldsymbol{h}^*_{m_1,m_2}(\boldsymbol{u},r), \tag{6.6}$$

and

$$\boldsymbol{v}_1' = \boldsymbol{v}_1 + \boldsymbol{h}'_{m_1-m_2,m_2}(\boldsymbol{u},r), \tag{6.7}$$

we can rearrange (6.5) to

$$\begin{aligned}
&N\bar{v}\tau(2L)^d\partial_t\hat{n}_{\boldsymbol{\kappa},m_1} \\
&\quad= \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \sum_{\boldsymbol{j},\boldsymbol{l}=-V}^{V} \hat{n}_{\boldsymbol{j},m_1}\hat{n}_{\boldsymbol{l},m_2} \int_{D_L} e^{i\frac{\pi}{L}(\boldsymbol{j}+\boldsymbol{l}-\boldsymbol{\kappa})\cdot\boldsymbol{v}_1}\,\mathrm{d}\boldsymbol{v}_1 \\
&\qquad\qquad \times \int_{\mathcal{B}_{2R}} |\boldsymbol{u}| e^{i\frac{\pi}{L}\boldsymbol{l}\cdot\boldsymbol{u}} \left[\int_{r^*_{m_1,m_2}(|\boldsymbol{u}|)}^{r_c} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{h}^*_{m_1,m_2}(\boldsymbol{u},r)}\, r^{d-2}\,\mathrm{d}r - \int_0^{r_c} r^{d-2}\,\mathrm{d}r\right] \mathrm{d}\boldsymbol{u} \\
&\quad\quad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \sum_{\boldsymbol{j},\boldsymbol{l}=-V}^{V} \hat{n}_{\boldsymbol{j},m_1-m_2}\hat{n}_{\boldsymbol{l},m_2} \int_{D_L} e^{i\frac{\pi}{L}(\boldsymbol{j}+\boldsymbol{l}-\boldsymbol{\kappa})\cdot\boldsymbol{v}_1}\,\mathrm{d}\boldsymbol{v}_1 \\
&\qquad\qquad \times \int_{\mathcal{B}_{2R}} |\boldsymbol{u}| e^{i\frac{\pi}{L}\boldsymbol{l}\cdot\boldsymbol{u}} \int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{u}|)} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{h}'_{m_1-m_2,m_2}(\boldsymbol{u},r)}\, r^{d-2}\,\mathrm{d}r\,\mathrm{d}\boldsymbol{u}.
\end{aligned} \tag{6.8}$$

Using the orthogonality of the Fourier basis, this simplifies to

$$\begin{aligned}
N\bar{v}\tau\partial_t\hat{n}_{\boldsymbol{\kappa},m_1} &= \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{n}_{\boldsymbol{j},m_1}\hat{n}_{\boldsymbol{l},m_2}\left(G^*(\boldsymbol{\kappa},\boldsymbol{l},m_1,m_2) - G(\boldsymbol{l})\right) \\
&\quad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{n}_{\boldsymbol{j},m_1-m_2}\hat{n}_{\boldsymbol{l},m_2} G'(\boldsymbol{\kappa},\boldsymbol{l},m_1-m_2,m_2),
\end{aligned} \tag{6.9}$$

with kernel modes

$$G(\boldsymbol{l}) = \frac{r_c^{d-1}}{d-1} \int_{\mathcal{B}_{2R}} |\boldsymbol{u}| e^{i\frac{\pi}{L}\boldsymbol{l}\cdot\boldsymbol{u}} \,\mathrm{d}\boldsymbol{u},$$

$$G^*(\boldsymbol{\kappa}, \boldsymbol{l}, m_1, m_2) = \int_{\mathcal{B}_{2R}} |\boldsymbol{u}| e^{i\frac{\pi}{L}\boldsymbol{l}\cdot\boldsymbol{u}} \int_{r^*_{m_1,m_2}(|\boldsymbol{u}|)}^{r_c} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{h}^*_{m_1,m_2}(\boldsymbol{u},r)} \, r^{d-2} \,\mathrm{d}r \,\mathrm{d}\boldsymbol{u},$$

and

$$G'(\boldsymbol{\kappa}, \boldsymbol{l}, m_1 - m_2, m_2) = \int_{\mathcal{B}_{2R}} |\boldsymbol{u}| e^{i\frac{\pi}{L}\boldsymbol{l}\cdot\boldsymbol{u}} \int_0^{r^*_{m_1-m_2,m_2}(|\boldsymbol{u}|)} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{h}'_{m_1-m_2,m_2}(\boldsymbol{u},r)} \, r^{d-2} \,\mathrm{d}r \,\mathrm{d}\boldsymbol{u},$$

where we have computed the $\mathrm{d}r$ integral in G assuming $d \geq 2$.

Given a particular dimension d, the loss kernel mode G can be computed analytically, but this is not necessarily the case for the non-clustering gain, G^*, and clustering gain, G', kernels. For interactions where they are not computable analytically, they could be precomputed and stored for all combinations of $\boldsymbol{\kappa}$, $\boldsymbol{l}$, m_1 and m_2. However, as the number of quantities to be stored scales as $O(V^{2d}N^2)$, this will quickly become infeasible for common values of V and N. Some approximations to reduce the memory usage and speed up computation include binning the masses with $M \ll N$ grid points (as was done with the 1D finite volume model in §4.1) or storing G^* and G' over a reduced grid of values spanning the parameter ranges and interpolating between them at runtime to approximate the true values. First, however, we examine the fast Fourier spectral method, an extension to the direct method developed for the Boltzmann equation in [53] and extended to an inelastic Boltzmann model for granular gases in [67].

6.2. Fast Fourier spectral method

The core idea of the fast Fourier spectral method is to approximate the weighted convolutions in the direct method, (6.9), with a small number of pure convolutions, which can be computed efficiently by FFT (fast Fourier transform). A weighted convolution, e.g.

$$\sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{n}_{\boldsymbol{j},m_1} \hat{n}_{\boldsymbol{l},m_2} G^*(\boldsymbol{\kappa}, \boldsymbol{l}, m_1, m_2),$$

must be computed directly, requiring $O(V^{2d})$ operations. However, a pure convolution can be computed using FFT with $O(V^d \log(V))$ operations. To make this modification, we approximate the $\mathrm{d}\boldsymbol{u}$ integral with a quadrature, such that we can factor out the $\boldsymbol{\kappa}$-dependent exponentials from the convolution, leaving only factors depending on $\boldsymbol{j}$ and $\boldsymbol{l}$. As suggested in [67], we separate $\boldsymbol{u}$ into a radial part, u, and angular part, $\hat{\boldsymbol{u}}$, giving the quadrature approximation

$$\int_{\mathcal{B}_{2R}} g(\boldsymbol{u}) \,\mathrm{d}\boldsymbol{u} = \int_{\mathbb{S}^{d-1}} \int_0^{2R} g(u, \hat{\boldsymbol{u}}) u^{d-1} \,\mathrm{d}u \,\mathrm{d}\hat{\boldsymbol{u}} \approx \sum_{u \in [0,2R]} \sum_{\hat{\boldsymbol{u}} \in \mathbb{S}^{d-1}} w_u w_{\hat{\boldsymbol{u}}} \, g(u, \hat{\boldsymbol{u}}) u^{d-1}$$

with weights w_u and $w_{\hat{\boldsymbol{u}}}$. Substituting this quadrature into (6.9) and rearranging gives

$$\begin{aligned}
&N\bar{v}\tau\partial_t\hat{n}_{\boldsymbol{\kappa},m_1} \\
&\quad = \Lambda_K \sum_{m_2=1}^{N-m_1-\mu_K+2} \sum_{u\in[0,2R]} \sum_{\hat{\boldsymbol{u}}\in\mathbb{S}^{d-1}} w_u w_{\hat{\boldsymbol{u}}} u^d \left(A^*(\boldsymbol{\kappa},u,\hat{\boldsymbol{u}},m_1,m_2) - \frac{r_c^{d-1}}{d-1} \right) \\
&\qquad\qquad \times \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{n}_{\boldsymbol{j},m_1} \left(e^{i\frac{\pi}{L}u\,\boldsymbol{l}\cdot\hat{\boldsymbol{u}}} \hat{n}_{\boldsymbol{l},m_2} \right) \\
&\quad + \frac{1}{2}\Lambda_K \sum_{m_2=1}^{m_1-1} \sum_{u\in[0,2R]} \sum_{\hat{\boldsymbol{u}}\in\mathbb{S}^{d-1}} w_u w_{\hat{\boldsymbol{u}}} u^d A'(\boldsymbol{\kappa},u,\hat{\boldsymbol{u}},m_1-m_2,m_2) \\
&\qquad\qquad \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{n}_{\boldsymbol{j},m_1-m_2} \left(e^{i\frac{\pi}{L}u\,\boldsymbol{l}\cdot\hat{\boldsymbol{u}}} \hat{n}_{\boldsymbol{l},m_2} \right),
\end{aligned} \tag{6.10}$$

with the new kernels

$$A^*(\boldsymbol{\kappa},u,\hat{\boldsymbol{u}},m_1,m_2) = \int_{r^*_{m_1,m_2}(u)}^{r_c} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{h}^*_{m_1,m_2}(\boldsymbol{u},r)}\, r^{d-2}\,\mathrm{d}r \tag{6.11}$$

and

$$A'(\boldsymbol{\kappa},u,\hat{\boldsymbol{u}},m_1-m_2,m_2) = \int_0^{r^*_{m_1-m_2,m_2}(u)} e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{h}'_{m_1-m_2,m_2}(\boldsymbol{u},r)}\, r^{d-2}\,\mathrm{d}r. \tag{6.12}$$

Given u, $\hat{\boldsymbol{u}}$, m_1 and m_2, the sums over $\boldsymbol{j}$ and $\boldsymbol{l}$ are pure convolutions that can be computed by FFT. Assuming N_u radial quadrature points and $N_{\hat{\boldsymbol{u}}}$ angular quadrature points, this method requires computing $N^2 N_u N_{\hat{\boldsymbol{u}}}$ pure convolutions which require $O(N^2 N_u N_{\hat{\boldsymbol{u}}} V^d \log(V))$ operations. The direct Fourier method required N^2 weighted convolutions implying $O(N^2 V^{2d})$ operations are necessary. Similarly, where the direct method required storing $O(N^2 V^{2d})$ quantities for each kernel mode, the fast method requires storing $O(N^2 N_u N_{\hat{\boldsymbol{u}}} V^2)$ quantities for each kernel. Thus, if we select a quadrature such that $N_u N_{\hat{\boldsymbol{u}}} \log(V) < V^d$, both the memory requirement and the computational complexity will be reduced in the fast Fourier method. (Note that there are specifically $2V+1$ Fourier modes used so this requirement can be relaxed to $N_u N_{\hat{\boldsymbol{u}}} \log(2V+1) < (2V+1)^d$.)

Hu and Ma [67] suggest a Gauss-Legendre quadrature in the radial direction, a rectangular rule for the angular part in 2D and a spherical quadrature rule in 3D. For their tests on a granular gases model (i.e. inelastic collisions with a constant coefficient of restitution and no clustering), they observe that $N_u = 32$ and $N_{\hat{\boldsymbol{u}}} = 30$ appear sufficient. However, they also suggest more points could be used depending on the precision required. We will examine the convergence with the number of quadrature points for our model in §7.1.4.

6.3. Mass binning

Even with the improved efficiency of the fast Fourier method, the memory requirement and computational complexity will remain too large as N increases. To improve this, we bin the set of masses onto a linear-geometric grid with M points, as we did for the 1D numerical model. This grid is defined as $\{\bar{m}_a, a \in 1, \ldots, M\}$, with M_l linear grid points and $M - M_l$ geometric points. The grid points are defined as in (4.2), and the bin boundaries, m_a^- and m_a^+, are defined as in (4.3). Bin i is again defined to include the set of $\bar{w}_a$ masses: $\{\lfloor m_a^- \rfloor + 1, \ldots, \lfloor m_a^+ \rfloor\}$.

Using these mass bins, we can define a new truncated Fourier approximation to $\check{n}$ as

$$\check{n}(\boldsymbol{v}; m) \approx \check{\nu}_V(\boldsymbol{v}; a(m)) = \frac{1}{\bar{w}_{a(m)}} \sum_{\boldsymbol{\kappa}=-V}^{V} \hat{\nu}_{\boldsymbol{\kappa},a(m)} e^{i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}},$$

where $a(m)$ denotes the bin a such that $m_a^- < m \leq m_a^+$. The new, binned Fourier coefficients are defined as

$$\hat{\nu}_{\boldsymbol{\kappa},a} = \frac{1}{(2L)^d} \int_{D_L} \sum_{m=\lfloor m_a^- \rfloor + 1}^{\lfloor m_a^+ \rfloor} n(\boldsymbol{v}; m) e^{-i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}} \, \mathrm{d}\boldsymbol{v} = \sum_{m=\lfloor m_a^- \rfloor + 1}^{\lfloor m_a^+ \rfloor} \hat{n}_{\boldsymbol{\kappa},m}.$$

Substituting this approximation, we can derive a mass-binned fast Fourier model:

$$\begin{aligned}
&N\bar{v}\tau\partial_t \hat{\nu}_{\boldsymbol{\kappa},a} \\
&\quad = \Lambda_K \sum_{\substack{\bar{m}_b \in \mathcal{M} \\ \bar{m}_a + \bar{m}_b \leq N - \mu_K + 2}} \sum_{u \in [0,2R]} \sum_{\hat{\boldsymbol{u}} \in \mathbb{S}^{d-1}} w_u w_{\hat{\boldsymbol{u}}} u^d \left(A^*(\boldsymbol{\kappa}, u, \hat{\boldsymbol{u}}, \bar{m}_a, \bar{m}_b) - \frac{r_c^{d-1}}{d-1} \right) \\
&\qquad\qquad \times \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{\nu}_{\boldsymbol{j},a} \left(e^{i\frac{\pi}{L}u\boldsymbol{l}\cdot\hat{\boldsymbol{u}}} \hat{\nu}_{\boldsymbol{l},b} \right) \\
&\quad + \frac{1}{2}\Lambda_K \sum_{\substack{\bar{m}_b, \bar{m}_c \in \mathcal{M} \\ \bar{m}_{a-1} < \bar{m}_b + \bar{m}_c < \bar{m}_{a+1}}} \sum_{u \in [0,2R]} \sum_{\hat{\boldsymbol{u}} \in \mathbb{S}^{d-1}} w_u w_{\hat{\boldsymbol{u}}} u^d A'(\boldsymbol{\kappa}, u, \hat{\boldsymbol{u}}, \bar{m}_b, \bar{m}_c) \gamma_a(\bar{m}_b + \bar{m}_c) \\
&\qquad\qquad \times \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{\nu}_{\boldsymbol{j},b} \left(e^{i\frac{\pi}{L}u\boldsymbol{l}\cdot\hat{\boldsymbol{u}}} \hat{\nu}_{\boldsymbol{l},c} \right),
\end{aligned} \tag{6.13}$$

where

$$\gamma_a(m) = \begin{cases} \frac{m - \bar{m}_{a-1}}{\bar{m}_a - \bar{m}_{a-1}}, & \bar{m}_{a-1} < m < \bar{m}_a \\ \frac{\bar{m}_{a+1} - m}{\bar{m}_{a+1} - \bar{m}_a}, & \bar{m}_a \leq m < \bar{m}_{a+1} \\ 0, & \text{otherwise.} \end{cases}$$

This function γ_a represents how clustered particles that don't have masses on grid points are allocated to the grid. Here, as was done for the 1D models, we have used the fixed pivot method of Kumar and Ramkrishna [77], which allocates product clusters to the two adjacent grid points so as to preserve the moments of cluster number and total mass.

6.4. Output quantities

Here, we define how to compute relevant output quantities from the Fourier spectral coefficients. Note that the Fourier coefficients $\hat{n}_{\boldsymbol{\kappa},m}$ and $\hat{\nu}_{\boldsymbol{\kappa},a}$ are complex, but all the quantities and distributions considered here are real. The imaginary part of the output in each case should be 0 up to the error in the Fourier approximation, so it can be assumed that we are only considering the real parts.

The normalised cluster velocity number distribution is equivalent to the cluster velocity probability distribution, i.e.

$$p(t;\boldsymbol{v}) = \frac{n(t;\boldsymbol{v})}{\mu_K(t)}.$$

Using the definition of the Fourier coefficients, this is defined as

$$p(t;\boldsymbol{v}) = \frac{n(t;\boldsymbol{v})}{\mu_K(t)} = \frac{1}{\mu_K(t)} \sum_{m=1}^{N} n(t;\boldsymbol{v};m) \approx \frac{1}{\mu_K(t)} \sum_{m=1}^{N} \sum_{\boldsymbol{\kappa}=-V}^{V} \hat{n}_{\boldsymbol{\kappa},m}(t) e^{i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}}.$$

If mass binning is used, we can use the fact that

$$\sum_{m=1}^{N} \hat{n}_{\boldsymbol{\kappa},m} = \sum_{a=1}^{M} \sum_{m=\lfloor m_a^- \rfloor+1}^{\lfloor m_a^+ \rfloor} \hat{n}_{\boldsymbol{\kappa},m} = \sum_{a=1}^{M} \hat{\nu}_{\boldsymbol{\kappa},a},$$

to modify this definition to

$$p(t;\boldsymbol{v}) = \frac{n(t;\boldsymbol{v})}{\mu_K(t)} \approx \frac{1}{\mu_K(t)} \sum_{a=1}^{M} \sum_{\boldsymbol{\kappa}=-V}^{V} \hat{\nu}_{\boldsymbol{\kappa},a}(t) e^{i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}}.$$

The cluster speed distribution is defined by representing velocities as a radial and angular part, i.e. $\boldsymbol{v} \equiv (v, \hat{\boldsymbol{v}})$ and integrating over the angular part:

$$p(t;v) = \frac{n(t;v)}{\mu_K(t)} = \frac{1}{\mu_K(t)} \int_{\mathbb{S}^{d-1}} n(t;v,\hat{\boldsymbol{v}})\, \mathrm{d}\hat{\boldsymbol{v}}.$$

In two dimensions we can represent $\hat{\boldsymbol{v}}$ as $(\cos(\alpha), \sin(\alpha))$ and this integral becomes

$$\begin{aligned} p(t;v) = \frac{n(t;v)}{\mu_K(t)} &= \frac{1}{\mu_K(t)} \sum_{m=1}^{N} \sum_{\boldsymbol{\kappa}=-V}^{V} \hat{n}_{\boldsymbol{\kappa},m}(t) \int_{\alpha=0}^{2\pi} e^{i\frac{\pi}{L} v\,(\cos(\alpha),\sin(\alpha))\cdot(\kappa_1,\kappa_2)} v\, \mathrm{d}\alpha \\ &= \frac{2\pi v}{\mu_K(t)} \sum_{m=1}^{N} \sum_{\boldsymbol{\kappa}=-V}^{V} \hat{n}_{\boldsymbol{\kappa},m}(t) J_0\left(\frac{\pi}{L}|\boldsymbol{\kappa}|\, v\right), \end{aligned} \tag{6.14}$$

where J_0 is a Bessel function of the first kind.

The cluster mass distribution can be computed by integrating over $\mathrm{d}\boldsymbol{v}$:

$$\begin{aligned} p(t;m) = \frac{n(t;m)}{\mu_K(t)} &= \frac{1}{\mu_K(t)} \int_{\mathbb{R}^d} n(t;\boldsymbol{v};m)\,\mathrm{d}\boldsymbol{v} \\ &\approx \frac{1}{\mu_K(t)} \sum_{\boldsymbol{\kappa}=-V}^{V} \hat{n}_{\boldsymbol{\kappa},m}(t) \int_{D_L} e^{i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}}\,\mathrm{d}\boldsymbol{v} \\ &= \frac{(2L)^d}{\mu_K(t)} \hat{n}_{\mathbf{0},m}(t). \end{aligned} \tag{6.15}$$

If we are binning the masses, we must normalise by the bin width such that this is a probability density:

$$p(t;m) = \frac{n(t;m)}{\mu_K(t)} \approx \frac{(2L)^d}{\mu_K(t)} \frac{\hat{\nu}_{\mathbf{0},a(m)}(t)}{\bar{w}_{a(m)}},$$

where $a(m)$ is again the bin such that $m_a^- < m \leq m_a^+$.

The mass-weighted (sometimes called particle-weighted) cluster mass distribution was defined in terms of p in (4.8) as $f(t;m) := \mu_K\, m\, p(t;m)$. Using the fact that $n = \mu_K p$ we define f in terms of n as

$$f(t;m) = m\, n(t;m),$$

and can then compute it from the Fourier coefficients as

$$\frac{f(t;m)}{N} = \frac{m}{N} n(t;m) \approx \frac{(2L)^d}{N} m\, \hat{n}_{\mathbf{0},m}(t) \approx \frac{(2L)^d}{N} \bar{m}_{a(m)} \frac{\hat{\nu}_{\mathbf{0},a(m)}(t)}{\bar{w}_{a(m)}}.$$

Note that, as f tracks the expected number of particles in clusters of a given mass, it is normalised by N.

The estimate for the mean number of clusters can be found by summing over the unnormalised cluster mass number distribution:

$$\mu_K(t) = \sum_{m=1}^{N} n(t;m) \approx (2L)^d \sum_{m=1}^{N} \hat{n}_{\mathbf{0},m}(t) \approx (2L)^d \sum_{a=1}^{M} \hat{\nu}_{\mathbf{0},a}(t). \tag{6.16}$$

The mean kinetic energy per particle is given by the integral of $\frac{1}{2}m|\boldsymbol{v}|^2$ over the number distribution, normalised by N. That is:

$$\begin{aligned} E(t) &= \frac{1}{2N} \sum_{m=1}^{N} \int_{\mathbb{R}^d} m|\boldsymbol{v}|^2 n(t;\boldsymbol{v},m)\,\mathrm{d}\boldsymbol{v} \\ &\approx \frac{1}{2N} \sum_{m=1}^{N} \sum_{\boldsymbol{\kappa}=-V}^{V} m\, \hat{n}_{\boldsymbol{\kappa},m}(t) \int_{D_L} |\boldsymbol{v}|^2 e^{i\frac{\pi}{L}\boldsymbol{\kappa}\cdot\boldsymbol{v}}\,\mathrm{d}\boldsymbol{v}. \end{aligned} \tag{6.17}$$

This definition is approximately equivalent to the definition using p instead of n in the one-dimensional case: (4.12), except with an extra independence assumption of the velocity and mass distributions (see (5.92)). In two dimensions, representing $\boldsymbol{v}$ in polar coordinates as $(v\cos(\alpha), v\sin(\alpha))$ and noting that we have restricted the support of $n(v)$ to $\mathcal{B}_R$ with $R \leq L$,

this integral can be approximated as

$$\begin{aligned} E(t) &\approx \frac{1}{2N}\sum_{m=1}^{N}\sum_{\boldsymbol{\kappa}=-V}^{V} m\,\hat{n}_{\boldsymbol{\kappa},m}(t)\int_0^L\int_0^{2\pi} v^3 e^{i\frac{\pi}{L}v(\kappa_1\cos(\alpha)+\kappa_2\sin(\alpha))}\,\mathrm{d}\alpha\,\mathrm{d}\boldsymbol{v} \\ &= \frac{L^4}{N}\sum_{m=1}^{N} m\left[\frac{\pi}{4}\hat{n}_{\mathbf{0},m}(t)+\sum_{\substack{\boldsymbol{\kappa}=-V\\ \boldsymbol{\kappa}\neq\mathbf{0}}}^{V}\hat{n}_{\boldsymbol{\kappa},m}(t)\left(\frac{2J_2\left(\pi|\boldsymbol{\kappa}|\right)-\pi|\boldsymbol{\kappa}|J_3\left(\pi|\boldsymbol{\kappa}|\right)}{\pi|\boldsymbol{\kappa}|^2}\right)\right], \end{aligned} \tag{6.18}$$

where J_2 and J_3 are Bessel functions of the first kind. With mass binning, this is

$$E(t) = \frac{L^4}{N}\sum_{a=1}^{M}\bar{m}_a\left[\frac{\pi}{4}\hat{\nu}_{\mathbf{0},a}(t)+\sum_{\substack{\boldsymbol{\kappa}=-V\\ \boldsymbol{\kappa}\neq\mathbf{0}}}^{V}\hat{\nu}_{\boldsymbol{\kappa},a}(t)\left(\frac{2J_2\left(\pi|\boldsymbol{\kappa}|\right)-\pi|\boldsymbol{\kappa}|J_3\left(\pi|\boldsymbol{\kappa}|\right)}{\pi|\boldsymbol{\kappa}|^2}\right)\right]. \tag{6.19}$$

6.5. Cucker-Smale model

When considering the Cucker-Smale collective behaviour model specifically, we have forwards velocity maps (from (5.63) and (5.69))

$$\boldsymbol{v}_1^* = \boldsymbol{v}_1 + m_2\Phi(r)\frac{\boldsymbol{v}_2-\boldsymbol{v}_1}{|\boldsymbol{v}_2-\boldsymbol{v}_1|} = \boldsymbol{v}_1 + m_2\Phi(r)\frac{\boldsymbol{u}}{|\boldsymbol{u}|}$$

and

$$\boldsymbol{v}_1' = \frac{m_1\boldsymbol{v}_1+m_2\boldsymbol{v}_2}{m_1+m_2} = \boldsymbol{v}_1 + \frac{m_2}{m_1+m_2}\boldsymbol{u}.$$

Defining the product velocities using $\boldsymbol{h}^*$ and $\boldsymbol{h}'$ as in (6.6) and (6.7), we then have

$$\boldsymbol{h}^*_{m_1,m_2}(\boldsymbol{u},r) = m_2\Phi(r)\frac{\boldsymbol{u}}{|\boldsymbol{u}|}$$

and

$$\boldsymbol{h}'_{m_1-m_2,m_2}(\boldsymbol{u},r) = \frac{m_2}{m_1}\boldsymbol{u}.$$

Note that the discrepancy with the definition of $\boldsymbol{v}_1'$ above arises because $\boldsymbol{h}'$ is defined for an interaction between clusters with masses m_1-m_2 and m_2 whereas $\boldsymbol{v}'$ is defined using m_1 and m_2. With these definitions, and the definition of r^* in (5.61), the fast Fourier kernels, (6.11) and (6.12), become

$$A^*(\boldsymbol{\kappa},u,\hat{\boldsymbol{u}},m_1,m_2) = \int_{\Phi^{-1}\left(\frac{u}{m_1+m_2}\right)}^{r_c} e^{-i\frac{\pi}{L}m_2\Phi(r)\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}\,r^{d-2}\,\mathrm{d}r$$

and

$$\begin{aligned} A'(\boldsymbol{\kappa},u,\hat{\boldsymbol{u}},m_1-m_2,m_2) &= \int_0^{\Phi^{-1}\left(\frac{u}{m_1}\right)} e^{-i\frac{\pi}{L}\frac{m_2}{m_1}u\,\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}\,r^{d-2}\,\mathrm{d}r \\ &= \frac{1}{d-1}e^{-i\frac{\pi}{L}\frac{m_2}{m_1}u\,\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}\left(\Phi^{-1}\left(\frac{u}{m_1}\right)\right)^{d-1}. \end{aligned}$$

6.5.1. Computational domain and aliasing

To derive the Fourier-Galerkin spectral method, we have truncated the velocity domain of $n(\boldsymbol{v}; m)$ to be compactly supported over $\mathcal{B}_R$, extended it to a function over D_L as in (6.4), and periodically extended this construction over copies of D_L (see Figure 6.1). In order to avoid aliasing, we require that L be large enough such that the domains of $\boldsymbol{v}_1 + \boldsymbol{u}$, $\boldsymbol{v}_1^*$, and $\boldsymbol{v}_1'$ do not intersect with a non-zero part of the periodic extension.

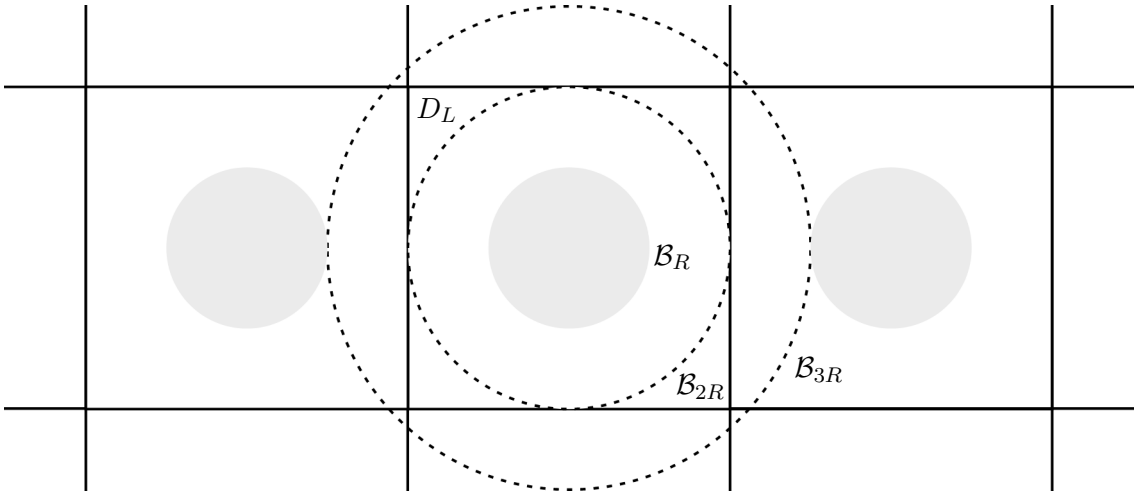


Figure 6.1.: Square computational domain, D_L, and restricted velocity support $\mathcal{B}_R$ for the periodic extension $\check{n}(\boldsymbol{v}; m)$, with non-zero regions of the extension shown in grey. To avoid aliasing, we must set L to be large enough such that the domains of $\boldsymbol{v}_1 + \boldsymbol{u}$, $\boldsymbol{v}_1^*$, and $\boldsymbol{v}_1'$ do not intersect with a non-zero part of the periodic extension. For the Cucker-Smale model shown we have set $L = 2R$ as $\boldsymbol{v}_1 \in \mathcal{B}_R$, $\boldsymbol{u} \in \mathcal{B}_{2R}$, and $\boldsymbol{v}_1 + \boldsymbol{u}$, $\boldsymbol{v}_1^*$, $\boldsymbol{v}_1' \in \mathcal{B}_{3R}$

From (5.63) and (5.69), we can see that, for the Cucker-Smale interaction, the velocity domain of $\check{n}(\boldsymbol{v}; m)$ remains compactly supported in $\mathcal{B}_R$ as all resulting velocities after interaction are on the line between the two pre-interaction velocities and $\mathcal{B}_R$ is convex. More specifically, for a non-clustering interaction, if ${}^*\boldsymbol{v}_1, {}^*\boldsymbol{v}_2 \in \mathcal{B}_R$, we have $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathcal{B}_R$ and, for a clustering interaction, if ${}'\boldsymbol{v}_1, \boldsymbol{v}_2 \in \mathcal{B}_R$, then $\boldsymbol{v}_1$ is also in $\mathcal{B}_R$. We also have $|\boldsymbol{u}| = |\boldsymbol{v}_2 - \boldsymbol{v}_1| \leq |\boldsymbol{v}_2| + |\boldsymbol{v}_1| \leq 2R$. Then, if we take $\boldsymbol{v}_1 \in \mathcal{B}_R$ and $\boldsymbol{u} \in \mathcal{B}_{2R}$ for our interaction operators in weak form, we have

$$|\boldsymbol{v}_1 + \boldsymbol{u}| \leq |\boldsymbol{v}_1| + |\boldsymbol{u}| \leq 3R,$$

$$|\boldsymbol{v}_1^*| \leq |\boldsymbol{v}_1| + \left| m_2 \Phi(r) \frac{\boldsymbol{u}}{|\boldsymbol{u}|} \right| \leq |\boldsymbol{v}_1| + \left| \frac{m_2 \boldsymbol{u}}{(m_1 + m_2)} \right| < 3R,$$

and

$$|\boldsymbol{v}_1'| \leq |\boldsymbol{v}_1| + \left| \frac{m_2}{m_1} \boldsymbol{u} \right| < 3R,$$

where we have used the fact that $\Phi(r) \leq \Phi(r^*_{m_1,m_2}(|\boldsymbol{u}|)) = |\boldsymbol{u}|/(m_1 + m_2)$ for the non-clustering interaction and that $m_2 < m_1$ for the clustering interaction. Thus, all the required quantities, $\boldsymbol{v}_1$, $\boldsymbol{v}_1 + \boldsymbol{u}$, $\boldsymbol{v}_1^*$, and $\boldsymbol{v}_1'$, remain supported in $\mathcal{B}_{3R}$ as we integrate over time. The closest non-zero point in the periodic extension is at a distance of $2L - R$ from the origin. Thus, to avoid

aliasing errors, we require $2L - R \geq 3R$, which implies that $L \geq 2R$. In what follows, we will take $L = 2R$, and the domain of the u quadrature in the fast Fourier method becomes $[0, L]$.

6.5.2. Power-law interaction kernels

Following §5.3.1, for a power-law interaction kernel, $\phi(|\boldsymbol{x}|) = 1/|\boldsymbol{x}|^s$, we have

$$\Phi(r) = \frac{\Gamma((s-1)/2)\sqrt{\pi}}{\Gamma(s/2)r^{s-1}}, \text{ and } \Phi^{-1}\left(\frac{u}{m}\right) = \left(\frac{\Gamma((s-1)/2)\sqrt{\pi}m}{\Gamma(s/2)u}\right)^{\frac{1}{s-1}}.$$

In Chapter 7, we will perform numerical experiments using the inverse-cube and inverse-quintic kernels specifically (i.e. setting $s = 3$ and $s = 5$, respectively). For the inverse-cube kernel, we have

$$\Phi(r) = \frac{2}{r^2}, \text{ and } \Phi^{-1}\left(\frac{u}{m}\right) = \sqrt{\frac{2m}{u}}.$$

The clustering Fourier kernel becomes

$$A'(\boldsymbol{\kappa}, u, \hat{\boldsymbol{u}}, m_1 - m_2, m_2) = \frac{1}{d-1} e^{-i\frac{\pi}{L}\frac{m_2}{m_1} u\, \boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}} \left(\frac{2m_1}{u}\right)^{\frac{d-1}{2}},$$

and the non-clustering kernel becomes

$$A^*(\boldsymbol{\kappa}, u, \hat{\boldsymbol{u}}, m_1, m_2) = \int_{\sqrt{\frac{2(m_1+m_2)}{u}}}^{r_c} e^{-i\frac{\pi}{L} m_2 \frac{2}{r^2}\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}\, r^{d-2}\, \mathrm{d}r.$$

Setting $d = 2$, this integrates to

$$re^{-\frac{2i\pi m_2 \boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}{Lr^2}} + (1+i)\pi\sqrt{\frac{m_2\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}{L}}\ \mathrm{erf}\left(\frac{(1+i)\sqrt{\frac{\pi}{L}m_2\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}}{r}\right)\Bigg|_{r=\sqrt{\frac{2(m_1+m_2)}{u}}}^{r=r_c}, \tag{6.20}$$

where erf denotes the error function. Note that we can simplify the computation here as the top limit in (6.20) approaches r_c in the limit $r_c \to \infty$. As the loss term in two dimensions is r_c (see (6.13)), the top limit of (6.20) cancels with the loss term if we take $r_c \to \infty$. Thus, it is simpler in this case not to include the cutoff value in the computation, and we only substitute the bottom limit of the integral: $r = \sqrt{2(m_1+m_2)/u}$. With these computations, the mass-binned fast Fourier spectral model for the Cucker-Smale interaction in two dimensions

with the inverse-cube interaction kernel is

$$
\begin{aligned}
&N\tau\partial_t\hat{\nu}_{\boldsymbol{\kappa},a} \\
&\quad= \Lambda_K \sum_{\substack{\bar{m}_b\in\mathcal{M} \\ \bar{m}_a+\bar{m}_b\le N-\mu_K+2}} \sum_{u\in[0,L]} \sum_{\hat{\boldsymbol{u}}\in[0,2\pi)} w_u w_{\hat{\boldsymbol{u}}} u^2 \left[\sqrt{\frac{2(m_1+m_2)}{u}} e^{-\frac{i\pi m_2 u\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}{L(m_1+m_2)}} \right. \\
&\qquad\qquad \left. + (1+i)\pi\sqrt{\frac{m_2\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}{L}}\ \mathrm{erf}\left((1+i)\sqrt{\frac{\pi m_2 u\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}{2L(m_1+m_2)}}\right)\right] \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{\nu}_{\boldsymbol{j},a}\left(e^{i\frac{\pi}{L}u\boldsymbol{l}\cdot\hat{\boldsymbol{u}}}\hat{\nu}_{\boldsymbol{l},b}\right) \\
&\quad+ \frac{1}{2(d-1)}\Lambda_K \sum_{\substack{\bar{m}_b,\bar{m}_c\in\mathcal{M} \\ \bar{m}_{a-1}<\bar{m}_b+\bar{m}_c<\bar{m}_{a+1}}} \sum_{u\in[0,L]} \sum_{\hat{\boldsymbol{u}}\in[0,2\pi)} w_u w_{\hat{\boldsymbol{u}}} u^2 \sqrt{\frac{2m_1}{u}} e^{-i\frac{\pi}{L}\frac{m_2}{m_1}u\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}\gamma_a(\bar{m}_b+\bar{m}_c) \\
&\qquad\qquad \times \sum_{\substack{\boldsymbol{j},\boldsymbol{l}=-V \\ \boldsymbol{j}+\boldsymbol{l}=\boldsymbol{\kappa}}}^{V} \hat{\nu}_{\boldsymbol{j},b}\left(e^{i\frac{\pi}{L}u\boldsymbol{l}\cdot\hat{\boldsymbol{u}}}\hat{\nu}_{\boldsymbol{l},c}\right).
\end{aligned}
\tag{6.21}
$$

This is the model we will use for the majority of the numerical experiments in Chapter 7. The values for u and w_u are defined using a Gauss-Legendre rule with N_u points across $[0, L]$. A rectangular rule with $N_{\hat{\boldsymbol{u}}}$ points on the circle is used for $\hat{\boldsymbol{u}}$, i.e. $\hat{\boldsymbol{u}} \in \left\{0, \frac{2\pi}{N_{\hat{\boldsymbol{u}}}}, \frac{4\pi}{N_{\hat{\boldsymbol{u}}}}, \ldots, \frac{2(N_{\hat{\boldsymbol{u}}}-1)\pi}{N_{\hat{\boldsymbol{u}}}}\right\}$ and $w_{\hat{\boldsymbol{u}}} = \frac{2\pi}{N_{\hat{\boldsymbol{u}}}}$.

For the inverse-quintic kernel, we have

$$
\Phi(r) = \frac{4}{3r^4}, \qquad \Phi^{-1}\left(\frac{u}{m}\right) = \left(\frac{4m}{3u}\right)^{\frac{1}{4}},
$$

and the clustering Fourier kernel becomes

$$
A'(\boldsymbol{\kappa}, u, \hat{\boldsymbol{u}}, m_1-m_2, m_2) = \frac{1}{d-1} e^{-i\frac{\pi}{L}\frac{m_2}{m_1}u\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}} \left(\frac{4m_1}{3u}\right)^{\frac{d-1}{4}}.
$$

In two dimensions, the non-clustering Fourier kernel integrates to

$$
\frac{1}{4} r E_{5/4}\left(\frac{4i\frac{\pi}{L}m_2\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}{3r^4}\right)\Bigg|_{r=\left(\frac{4(m_1+m_2)}{3u}\right)^{\frac{1}{4}}}^{r=r_c},
$$

where $E_{5/4}$ is the generalised exponential integral function E_n with $n = 5/4$ (not to be confused with the particle energy $E(t)$), defined as

$$
E_n(x) = \int_1^\infty \frac{e^{-xt}}{t^n}\,\mathrm{d}t.
$$

Similarly, the limit at $r = r_c$ cancels with the loss term as $r_c \to \infty$, and so, substituting the bottom limit, we get

$$
\frac{1}{4}\left(\frac{4(m_1+m_2)}{3u}\right)^{\frac{1}{4}} E_{5/4}\left(\frac{i\frac{\pi}{L}m_2 u\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}}{m_1+m_2}\right).
$$

Similar results are available for other powers s in the interaction kernel and other dimensions

d, allowing the Fourier kernels to be computed analytically.

6.5.3. Exponential interaction kernel

Considering the exponential interaction kernel, $\phi(|\boldsymbol{x}|) = e^{-|\boldsymbol{x}|}$, we have

$$\Phi(r) = \int_{-\infty}^{\infty} e^{-\sqrt{r^2+(y')^2}}\,\mathrm{d}y' = \begin{cases} 2, & r = 0 \\ 2rK_1(r), & r > 0, \end{cases}$$

where K_1 is a modified Bessel function of the second kind. This does not have a closed-form analytical inverse, so it is necessary to either analytically approximate Φ^{-1} and the Fourier kernel integrals or to compute them numerically.

One method involves computing $\Phi^{-1}\left(\frac{u}{m}\right)$ over a range of values numerically and interpolating on this grid to get approximate values to use at runtime. For the clustering Fourier kernel we need $\Phi^{-1}\left(\frac{u}{m_1}\right)$ and for the non-clustering kernel $\Phi^{-1}\left(\frac{u}{m_1+m_2}\right)$, with $u \in [0, L]$ and $m_1, m_2 \in \{1, \dots N\}$. As such, we need the interpolation grid to cover arguments of Φ^{-1} in the range $[0, L]$. However, note that the range of $\Phi(r)$ is $(0, 2]$. Thus for values of $\frac{u}{m} > 2$ we don't need to compute Φ^{-1} directly and instead set $\Phi^{-1}\left(\frac{u}{m}\right) = 0$. This represents trajectories where the ratio of relative velocity to mass is great enough such that even a direct interaction with $r = 0$ will not cause clustering, so $r^*_{m_1,m_2}(u) = \Phi^{-1}\left(\frac{u}{m_1+m_2}\right) = 0$ (and equivalently $r^*_{m_1-m_2,m_2}(u) = 0$ in the clustering kernel). Thus, we only need to numerically compute $\Phi^{-1}\left(\frac{u}{m}\right)$ an interpolation grid covering $\frac{u}{m} \in (0, 2]$. In the numerical experiments, we compute this over multiple rectangular grids, with the grid spacing refined for values close to 0 as the gradient of $\Phi^{-1}\left(\frac{u}{m}\right)$ increases significantly as $\frac{u}{m} \to 0$.

The approximation of Φ^{-1} is enough to compute the clustering Fourier kernel. However, to compute the non-clustering kernel, we must also approximate the integral

$$A^*(\boldsymbol{\kappa}, u, \hat{\boldsymbol{u}}, m_1, m_2) = \int_{\Phi^{-1}\left(\frac{u}{m_1+m_2}\right)}^{r_c} e^{-2i\frac{\pi}{L}m_2\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}}\, rK_1(r)}\, r^{d-2}\,\mathrm{d}r. \tag{6.22}$$

This can be represented in two dimensions as

$$\int_{B_2}^{r_c} e^{-iB_1 rK_1(r)}\,\mathrm{d}r,$$

where $B_1 = 2\frac{\pi}{L}m_2\boldsymbol{\kappa}\cdot\hat{\boldsymbol{u}} \in \left[-\frac{2\sqrt{2}\pi NV}{L}, \frac{2\sqrt{2}\pi NV}{L}\right]$ and $B_2 = \Phi^{-1}(\frac{u}{m_1+m_2}) \in [0, r_c]$. This is a highly oscillatory integral that requires care to approximate accurately, particularly as $B_2 \to 0$. As a first attempt we compute over the grid $B_2 \in \{0, 0.1, 0.2, \dots, 20\}$ with $r_c = 20$ and a grid in B_1 over $[-7000, 7000]$ with step size

$$\Delta B_1 = \begin{cases} 100, & 7 \le B_2 \le 20 \\ 10, & 4.5 \le B_2 \le 7 \\ 1, & 2 \le B_2 \le 4.5 \\ 0.1, & B_2 \le 2. \end{cases}$$

The approximated values for A^* are then computed using cubic-spline interpolation between these grid points at runtime.

7. 2D Numerical Evaluation

To evaluate the accuracy of our short-range clustering PDE model in higher dimensions, we test it against the full particle ODE simulations for the Cucker-Smale model in two dimensions. The Fourier-Galerkin spectral method is implemented in *Julia*, with the convolution sum computed with *DSP.jl* package, which uses an FFT implementation written in *Julia*. The majority of the experiments are performed with a uniform spatial distribution, a monodisperse initial mass condition and a Maxwellian initial velocity distribution with mean velocity $\mathbf{0}$ given by

$$n(t=0;\boldsymbol{v};m=1)=\frac{N}{16\pi}e^{-\frac{|\boldsymbol{v}|^2}{16}}.$$

Although we test other interaction kernels in §7.2.4, the majority of the experiments are performed with the inverse-cube interaction kernel, which allows us to make use of the analytical solutions for the Fourier kernel integrals and base the implementation off (6.21). Given this initial velocity distribution, we restrict the support to $\mathcal{B}_R$ with $R=10$ when defining the periodic extension $\check{n}$ and hence set $L=20$ to define the Fourier domain D_L. We use the full mass distribution for models with up to $N=40$ particles but use mass-binning with $M=30$ bins when $N>40$. The `OwrenZen5`, adaptive time-step ODE solver algorithm from the *DifferentialEquations.jl* package is used for time integration, with relative tolerance 10^{-3} and absolute tolerance 10^{-6}.

The particle ODE simulations to be used as a base case for comparison are implemented as described in §4.3.1. The ODEs used for the short-range Cucker-Smale model with inverse-cube interaction kernel are

$$\begin{aligned}\dot{\boldsymbol{X}}_i&=\boldsymbol{V}_i,\\ \dot{\boldsymbol{V}}_i&=\varepsilon^2\sum_{\substack{j=1\\ j\neq i}}^{N}\frac{1}{|\boldsymbol{X}_i-\boldsymbol{X}_j|^3}(\boldsymbol{V}_j-\boldsymbol{V}_i).\end{aligned}\tag{7.1}$$

Again, the `OwrenZen5` algorithm is used for time integration, and the DBSCAN algorithm is used to label groups of highly-correlated particles as clusters. Figures 7.1 and 7.2 show time snapshots of particles in the spatial domain for simulations with $N=20$ and $N=80$ particles, respectively. The gradual formation of clusters (indicated by different colours) and alignment of particle velocities (shown with proportional arrows) over time can be seen.

The plots against time are shown using scaled time t', which we define as

$$t':=\frac{N\varepsilon}{|\Omega|}t.$$

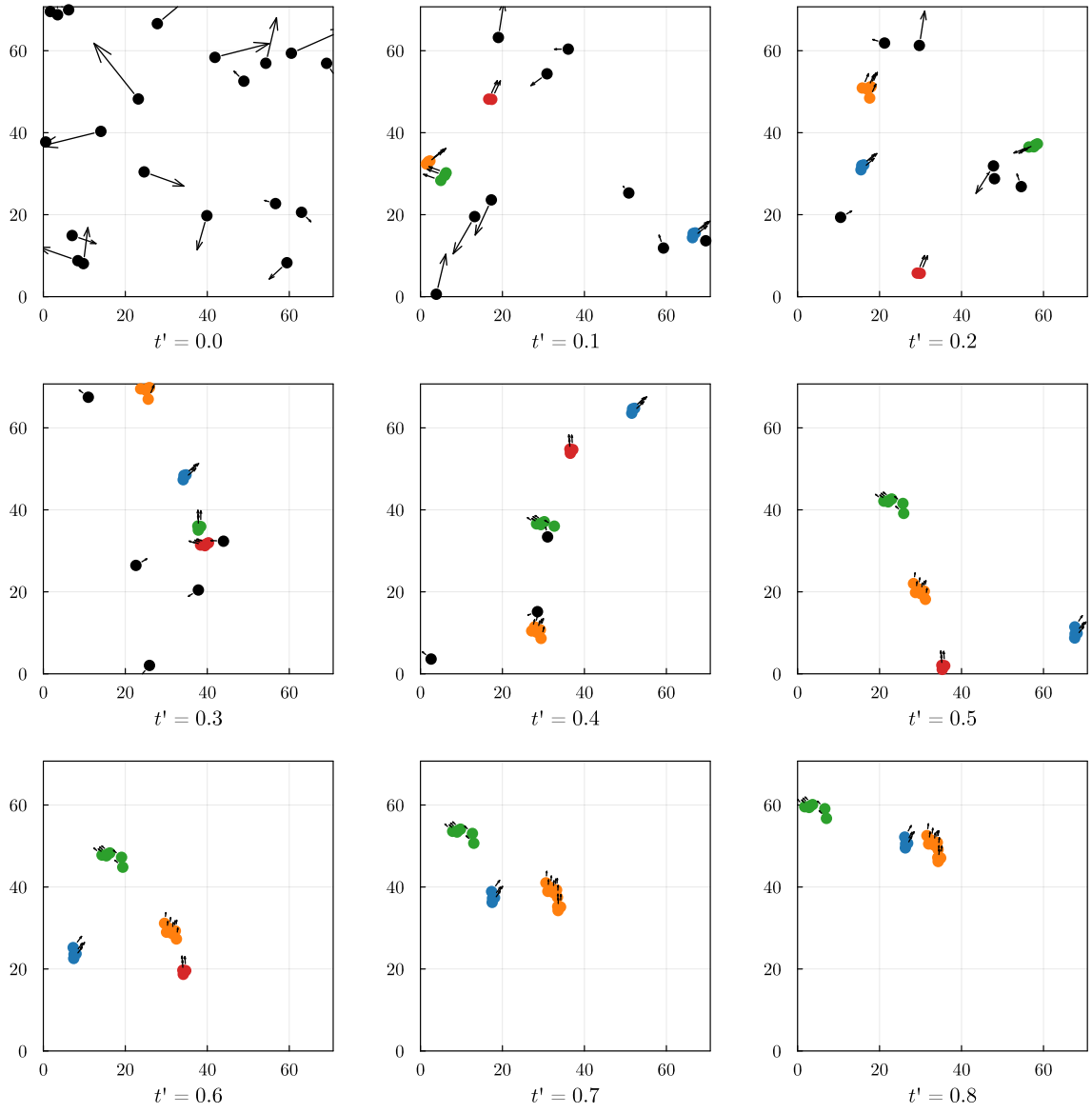


Figure 7.1.: A particle ODE simulation for the Cucker-Smale interaction with inverse-cube kernel, $N = 20$ particles, density $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ and a Maxwellian initial velocity distribution. Clusters identified by the DBSCAN algorithm are coloured, and the arrows represent particle velocities. The full animated version can be found at https://figshare.com/s/792a101f5db84e4b9e6c.

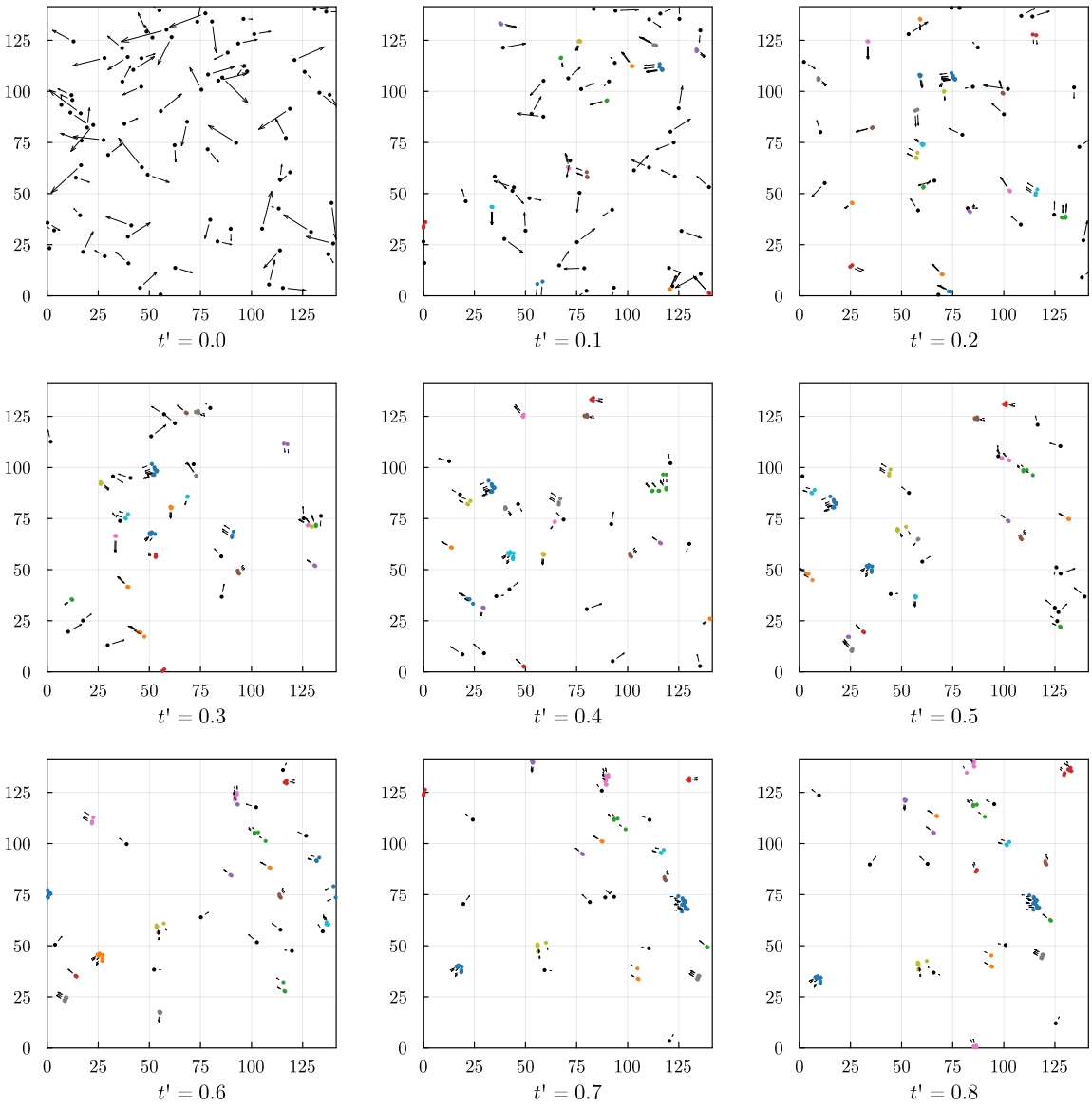


Figure 7.2.: A particle ODE simulation for the Cucker-Smale interaction with inverse-cube kernel. $N = 80$ particles, density $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ and a Maxwellian initial velocity distribution. Clusters identified by the DBSCAN algorithm are coloured, and the arrows represent particle velocities. The full animated version can be found at https://figshare.com/s/be227121efff2cda31e7 and an animation of a simulation with $N = 160$ particles at https://figshare.com/s/6d1e7d7a8fdde5036da0.

Note that the time scale defined in §5.1.7 would become $\tau = |\Omega|/(2N\varepsilon\bar{v})$ in two dimensions. We have simplified the scale used for these experiments, not including the $\bar{v}$ or the factor of 2, as we will only be considering variations in the particle number and spatial density, not varying the scale of the initial velocity distribution. Note also that this is a different time scale from that used for the one-dimensional experiments.

We heuristically define the 'density' of a system we consider to be the volume fraction occupied by a set of disks of radius ε at each particle: $N\varepsilon^2\pi/|\Omega|$. Although we are modelling point particles (and thus, this is not a true volume fraction), ε is the length scale of the interaction, and so these disks are representative of each particle's region of influence. In §7.2.2, we will analyse the effect of varying this density, but for all other simulations, we take $N\varepsilon^2\pi/|\Omega| = 0.004\pi$.

7.1. Parameter selection

7.1.1. Λ_K approximation

To implement our PDE method, we require a closed approximation for the coefficient Λ_K, defined from the cluster number distribution as

$$\Lambda_K = \frac{\mu_{K^2} - \mu_K}{\mu_K^2} = 1 + \frac{\sigma_K^2 - \mu_K}{\mu_K^2},$$

which was introduced when making the *Clusterzahlansatz*, (5.31). The most accurate method for estimating Λ_K using the information we already have available in the fast Fourier spectral implementation is to track the full k distribution using (5.38) and compute Λ_K directly. In the two dimensional spatially uniform case, (5.38) becomes

$$\begin{aligned} N\bar{v}\tau\frac{\mathrm{d}}{\mathrm{d}t}P_k = \frac{1}{2(\mu_K)^2}\left((k+1)kP_{k+1} - k(k-1)P_k\right) \\ \times \sum_{\substack{m_1,m_2=1 \\ m_1+m_2 \\ \leq N-k+2}}^{N} \int_{\mathbb{R}^d\times\mathbb{R}^d} |\boldsymbol{v}_1 - \boldsymbol{v}_2|\, n(t;\boldsymbol{v}_1;m_1)\, n(t;\boldsymbol{v}_2;m_2)\, r^*_{m_1,m_2}(|\boldsymbol{v}_1 - \boldsymbol{v}_2|)\,\mathrm{d}\boldsymbol{v}_1\mathrm{d}\boldsymbol{v}_2, \quad (7.2) \end{aligned}$$

where we have approximated both mass restrictions with $m_1 + m_2 \leq N - k + 2$ to simplify the equation. If, as in the derivation of the Fourier-Galerkin spectral method, we restrict the domain of integration to D_L, change variables to relative velocity, $\boldsymbol{u}$, and substitute the truncated Fourier approximation for $\check{n}$ this becomes

$$\begin{aligned} N\bar{v}\tau\frac{\mathrm{d}}{\mathrm{d}t}P_k = \frac{1}{2(\mu_K)^2}\left((k+1)kP_{k+1} - k(k-1)P_k\right) \\ \times \sum_{\substack{m_1,m_2=1 \\ m_1+m_2 \\ \leq N-k+2}}^{N} \sum_{\boldsymbol{j},\boldsymbol{l}=-V}^{V} \hat{n}_{\boldsymbol{j},m_1}\hat{n}_{\boldsymbol{l},m_2} \int_{D_L\times\mathcal{B}_{2R}} |\boldsymbol{u}|\, e^{i\frac{\pi}{L}\boldsymbol{j}\cdot\boldsymbol{v}_1} e^{i\frac{\pi}{L}\boldsymbol{l}\cdot(\boldsymbol{v}_1+\boldsymbol{u})} r^*_{m_1,m_2}(|\boldsymbol{u}|)\,\mathrm{d}\boldsymbol{u}\mathrm{d}\boldsymbol{v}_1, \quad (7.3) \end{aligned}$$

which, using the orthogonality of the Fourier basis functions, can be reduced to

$$N\bar{v}\tau\frac{\mathrm{d}}{\mathrm{d}t}P_k = \frac{(2L)^2}{2(\mu_K)^2}\left((k+1)kP_{k+1} - k(k-1)P_k\right) \times \sum_{\substack{m_1,m_2=1 \\ m_1+m_2 \\ \leq N-k+2}}^{N}\sum_{\boldsymbol{j}=-V}^{V}\hat{n}_{\boldsymbol{j},m_1}\hat{n}_{-\boldsymbol{j},m_2}\int_{\mathcal{B}_{2R}}|\boldsymbol{u}|\,e^{-i\frac{\pi}{L}\boldsymbol{j}\cdot\boldsymbol{u}}r^*_{m_1,m_2}(|\boldsymbol{u}|)\,\mathrm{d}\boldsymbol{u}. \quad (7.4)$$

For the Cucker-Smale model with inverse-cube interaction kernel, we have that

$$r^*_{m_1,m_2}(|\boldsymbol{u}|) = \Phi^{-1}\left(\frac{|\boldsymbol{u}|}{m_1+m_2}\right) = \sqrt{\frac{2(m_1+m_2)}{|\boldsymbol{u}|}},$$

and so in two dimensions, using polar coordinates $\boldsymbol{u} \equiv (u\cos(\alpha), u\sin(\alpha))$, and noting that taking $L = 2R$ is sufficient for the Cucker-Smale model, the $\mathrm{d}\boldsymbol{u}$ integral becomes

$$\sqrt{2(m_1+m_2)}\int_0^L\int_0^{2\pi}u^{3/2}\,e^{-i\frac{\pi}{L}u\,(j_1\cos(\alpha)+j_2\sin(\alpha))}\,\mathrm{d}\alpha\,\mathrm{d}u = \frac{4\sqrt{2}L^{5/2}\pi\sqrt{m_1+m_2}}{5}\,{}_1F_2\left(\frac{5}{4};1,\frac{9}{4};-\frac{|\boldsymbol{j}|^2\pi^2}{4}\right),$$

where ${}_1F_2$ is the generalised hypergeometric function, conventionally written as ${}_pF_q(a_1,\ldots,a_p;b_1\ldots,b_q;z)$. As such, for this interaction, we can compute the time evolution of the set of P_k alongside the Fourier coefficients $\hat{n}_{\kappa,m}$, using

$$N\bar{v}\tau\frac{\mathrm{d}}{\mathrm{d}t}P_k = \frac{8\sqrt{2}L^{9/2}\pi}{5(\mu_K)^2}\left((k+1)kP_{k+1} - k(k-1)P_k\right) \times \sum_{\substack{m_1,m_2=1 \\ m_1+m_2 \\ \leq N-k+2}}^{N}\sum_{\boldsymbol{j}=-V}^{V}\sqrt{m_1+m_2}\,\hat{n}_{\boldsymbol{j},m_1}\hat{n}_{-\boldsymbol{j},m_2}\,{}_1F_2\left(\frac{5}{4};1,\frac{9}{4};-\frac{|\boldsymbol{j}|^2\pi^2}{4}\right), \quad (7.5)$$

with μ_K approximated directly from the set of $\hat{n}$ as in (6.16), with $\mu_K(t) \approx (2L)^d\sum_{m=1}^N\hat{n}_{\mathbf{0},m}(t)$.

However, to simplify the computation for these numerical experiments, we consider if it is possible to produce a sufficiently accurate estimate for $\Lambda_K(t)$ directly from $\mu_K(t)$ at each time step. A simple approximation is that the number of clusters is Poisson distributed, i.e. $K \sim \mathrm{Pois}(\mu_K)$. This would give $\sigma_K^2 = \mu_K$ so that $\Lambda_K^{\mathrm{Pois}}(t) = 1$ for all t. We could also assume K was Binomially distributed with N trials and $p = \mu_K/N$ probability of success, i.e. $K \sim \mathrm{Bin}(N, \mu_K/N)$, giving

$$\Lambda_K^{\mathrm{Bin}}(t) = 1 + \frac{Np(1-p)-Np}{N^2p^2} = 1 - \frac{1}{N}.$$

Noting that for the monodisperse initial condition we have $\mu_K(0) = N$ and $\sigma_K^2(0) = 0$ it follows that $\Lambda_K^{\mathrm{Bin}}(0) = 1 - 1/N$. As a result, we expect the Binomial approximation to capture the initial behaviour well, and the Poisson approximation to improve as $N \to \infty$.

Figure 7.3 shows Λ_K against μ_K, with both quantities computed empirically from the

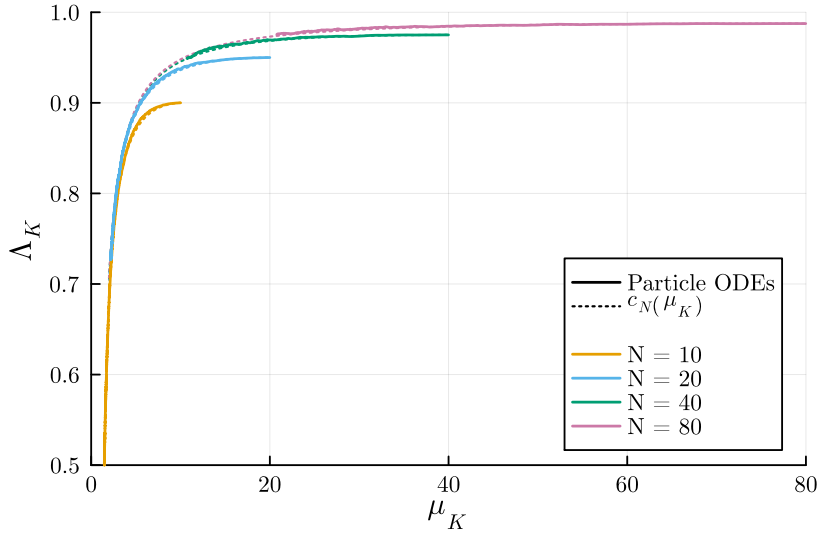


Figure 7.3.: *Clusterzahlansatz* coefficient, $\Lambda_K = 1 + (\sigma_K^2 - \mu_K)/\mu_K^2$, for Cucker-Smale particle ODE simulations with inverse-cube interaction kernel. Dashed lines represent the empirical approximation using only μ_K for each N given by $c_N(\mu_K) = (1 + 1/\mu_K)^{1/2(1+(\mu_K/N)^2)}$.

particle ODE simulations for trials with N varying from 10 to 80. In each trial, as t increases, μ_K decreases, and it can be seen that Λ_K also decreases. Interestingly, after an initial transient, the same relationship between Λ_K and μ_K is observed for all values of N. The binomial approximation with $\Lambda_K = 1 - 1/N$ appears reasonable when μ_K is close to N but less valid for later times when μ_K is small, as expected. Both binomial and Poisson approximations appear to maintain their accuracy for longer times when N is large, with Λ_K remaining closer to 1.

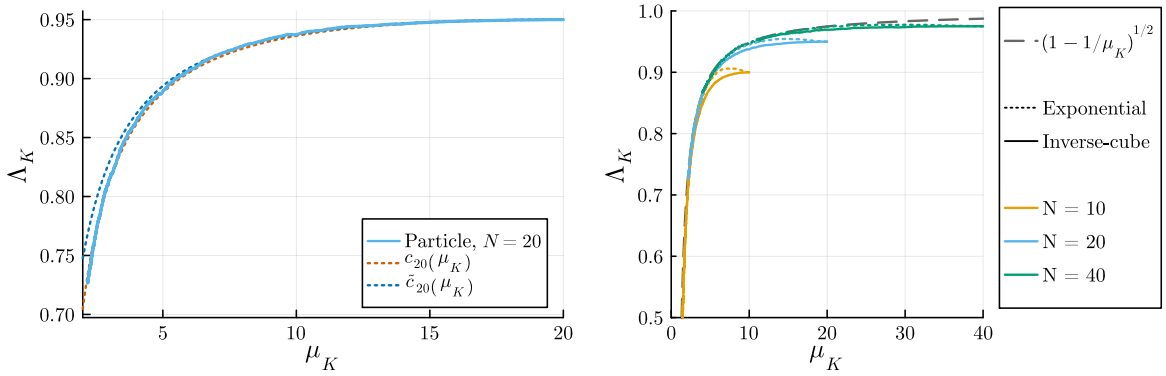
To match the particle ODE results more closely for all t, we define the empirical approximation

$$\Lambda_K(t) \approx c_N(\mu_K) = (1 + 1/\mu_K)^{1/2(1+(\mu_K/N)^2)}. \tag{7.6}$$

This can be seen in Figure 7.3 to be a much more accurate approximation than either Binomial or Poisson, capturing the decrease in Λ_K when μ_K is small. We, as yet, do not have a complete theoretical justification for this approximation, but we expect further investigation into the k-distribution evolution defined in (5.38) may offer some insight. We also consider a simplified empirical approximation, a large-μ_K Taylor approximation to c_N given by

$$\tilde{c}_N(\mu_K) = 1 + \frac{1}{2\mu_K}\left(1 + \left(\frac{\mu_K}{N}\right)^2\right).$$

Figure 7.4a shows a comparison of both empirical approximations against the particle ODE simulations for the trial with $N = 20$. The full approximation, c_{20}, can be seen to be more accurate for small μ_K than the simplified approximation, $\tilde{c}_{20}$, although this distinction requires further investigation. For the experiments in this chapter, we will use the empirical approximation $\Lambda_K = c_N(\mu_K)$ to limit the effect of the approximation of Λ_K on the results.



(a) (b)

Figure 7.4.: (a) Empirical approximations to Λ_K, $c_N(\mu_K) = (1 + 1/\mu_K)^{1/2(1+(\mu_K/N)^2)}$ and the simplified $\tilde{c}_N(\mu_K) = 1 + 1/(2\mu_K)(1 + (\mu_K/N)^2)$, for $N = 20$ and inverse-cube interaction kernel. At later times (for smaller μ_K), $c_{20}(\mu_K)$ can be seen to be more accurate. (b) Particle simulation results for Λ_K for exponential and inverse-cube interaction kernels. The kernels show different initial transients, but they both approach the same limiting behaviour as μ_K decreases for all values of N. The limiting curve $(1 - 1/\mu_K)^{1/2}$ (which $c_N(\mu_K)$ approaches for $\mu_K \ll N$) is shown for comparison.

Figure 7.4b shows Λ_K against μ_K for particle ODE simulations with the exponential interaction kernel for comparison with those using the inverse-cube kernel. The initial transient behaviour differs, although the limiting behaviour for both kernels approaches $\Lambda_K \approx (1 - 1/\mu_K)^{1/2}$, which is the limit of $c_N(\mu_K)$ for $\mu_K \ll N$. This difference in initial transients suggests the empirical approximation $c_N(\mu_K)$ is specific to the inverse-cube kernel, although the limiting behaviour is more generalisable.

Figure 7.5 shows cluster number distributions produced by the particle ODE simulations compared with the distributions implied by the various approximations to Λ_K for a trial with $N = 20$. A similar plot for $N = 160$ is shown in Figure B.4 in the Appendix. The empirical approximations can be seen to give good approximations to the k distribution, with very little difference between c_N and the simplified $\tilde{c}_N$. The binomial approximation is worse for later times when μ_K has decreased, as expected. The Poisson distribution does not match well at $t' = 0$, even though it approximates Λ_K better here. This may be explained as we can rearrange (5.31) to

$$\sigma_K^2 = \mu_K^2(\Lambda_K - 1) + \mu_K.$$

We note that $\partial\sigma_K^2/\partial\Lambda_K = \mu_K^2$, and so at $t' = 0$ when μ_K is large, the approximation of σ_K^2 is very sensitive to Λ_K. In the initial condition here where $\mu_K(0) = 20$, the Poisson approximation of $\Lambda_K(0) = 1$ gives $\sigma_K^2(0) = 20$, while the approximation of $\Lambda_K(0) = 1 - 1/20$, given by the Binomial and empirical approximations, gives $\sigma_K^2(0) = 0$.

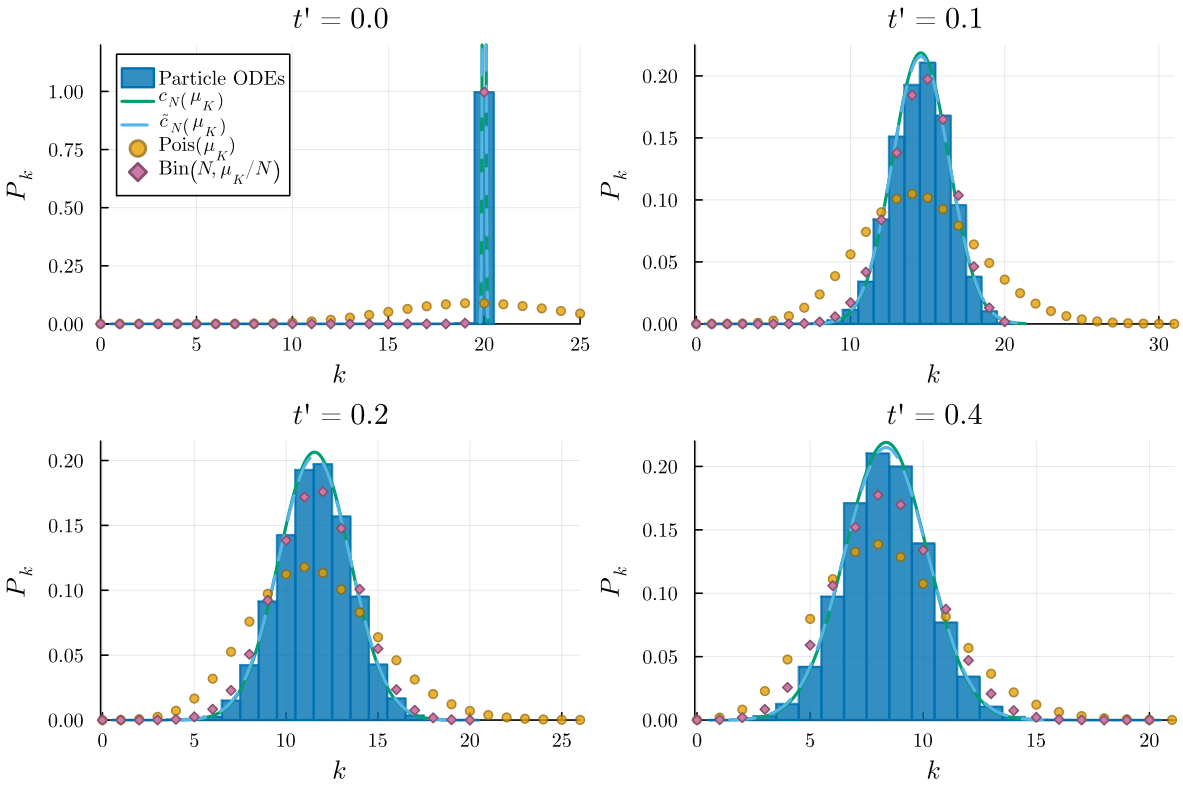


Figure 7.5.: Cluster number, k, distributions from particle ODE simulations with inverse-cube collision kernel and $N = 20$. Normal distributions using the variance implied by the empirical approximations $c_N(\mu_K)$ and $\tilde{c}_N(\mu_K)$ are shown, as well as the distributions $\text{Bin}(N, \mu_K/N)$ and $\text{Pois}(\mu_K)$. The empirical approximations can be seen to be the closest fit, and there is little difference observed between the two.

7.1.2. *Clusterzahlansatz* m_2 sum approximation

When applying the *Clusterzahlansatz*, we approximate the upper limit for the allowed value of m_2 as $N - m_1 - \lfloor \mu_K \rfloor + 2$ (see the steps from (5.30) to (5.32)). Before k was marginalised, the limit for each k was $N - m_2 - k + 2$. This limited m_2 to ensure there were enough other particles available to produce another cluster with mass m_1 and the required $k - 2$ other clusters. The approximated limit, $N - m_1 - \lfloor \mu_K \rfloor + 2$, introduced after the k marginalisation and *Clusterzahlansatz* ensures enough particles are available to produce another cluster with mass $m - 1$ and $\lceil \mu_K - 2 \rceil$ further clusters. We evaluate whether this more complicated bound is necessary by comparing it against the simpler approximation of limiting m_2 to $N - m_1$. Figure 7.6 shows that there is a negligible difference between the two approximations for maximum Fourier Frequency $V = 30$, and for lower values of V, the simpler approximation matches the particle ODE results more accurately. As such, from here on, we use $N - m_1$ as the upper limit for m_2.

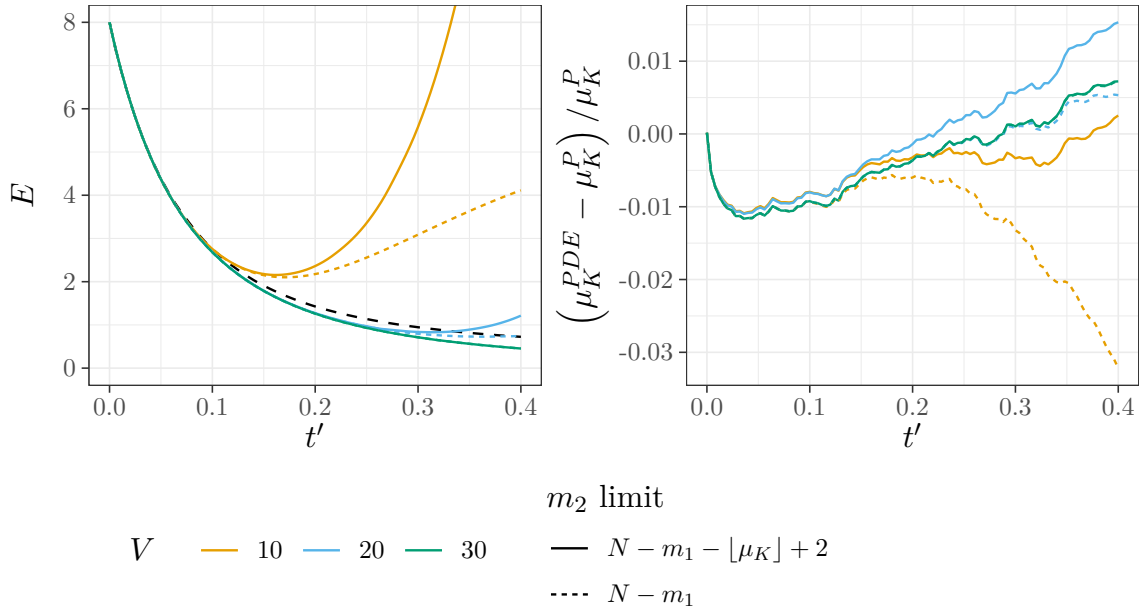
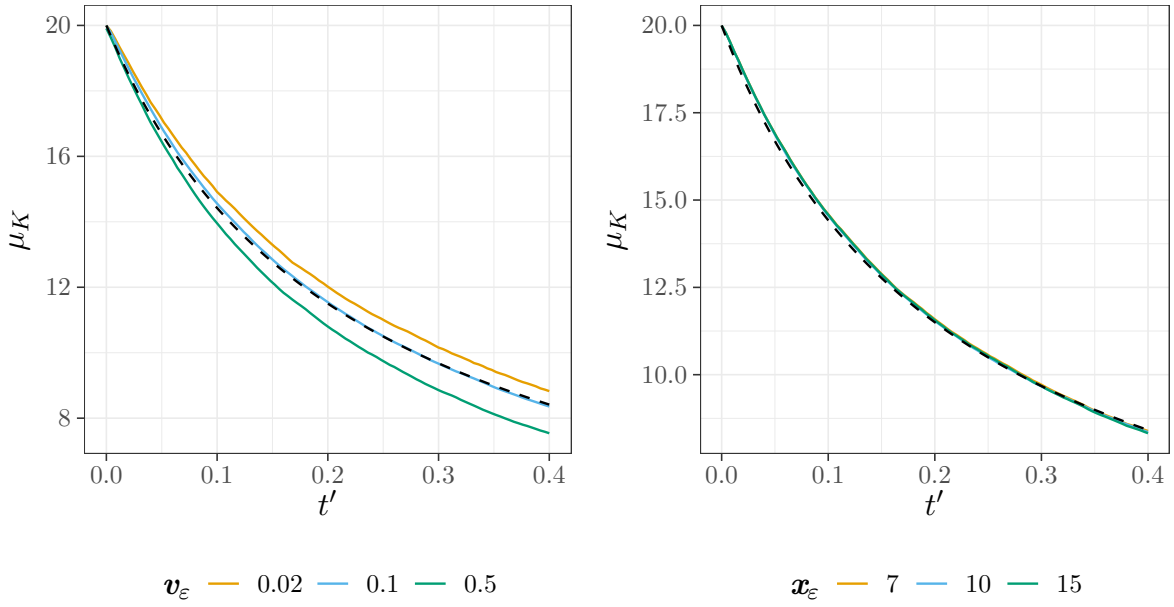


Figure 7.6.: Mean kinetic energy per particle, E, and mean number of clusters μ_K for different *Clusterzahlansatz* approximations to the m_2 sum limit and different maximum Fourier frequencies, V, in the PDE model with $N = 20$ and inverse-cube kernel. The equivalent particle ODE simulation result for E is shown as a black dashed line. For the cluster numbers, the proportional difference between the PDE results, μ_K^{PDE}, and the particle ODE results, μ_K^P, is shown for clarity as the errors in all cases are much smaller.

7.1.3. DBSCAN clustering parameters

To determine which values to use for the spatial, $\varepsilon_{\boldsymbol{x}}$, and velocity, $\varepsilon_{\boldsymbol{v}}$, tolerance for two particles to be considered linked in a cluster for the DBSCAN algorithm, we observed the clusters produced in individual particle ODE simulations. Values of $\varepsilon_{\boldsymbol{x}} = 10$ and $\varepsilon_{\boldsymbol{v}} = 0.1$ appeared to produce the qualitatively desired behaviours of pairs of clusters combining smoothly during a clustering interaction without being labelled as multiple clusters during the process and labelled clusters not splitting apart spontaneously. Note that this value of $\varepsilon_{\boldsymbol{x}}$ is greater than was used in Chapter 4 for the one-dimensional case as, in those experiments, we used the exponential interaction kernel. Here, we use the inverse-cube interaction kernel, which does not decay as quickly in $|\boldsymbol{x}|$ and, as such, is consistent with a larger radius of spatial tolerance being necessary to capture the majority of the interaction range of each particle.

Figure 7.7 shows plots of the mean number of clusters for the particle ODE simulations with $N = 20$ and different values of $\varepsilon_{\boldsymbol{v}}$ and $\varepsilon_{\boldsymbol{x}}$. Values of $\varepsilon_{\boldsymbol{x}} = 10$, and $\varepsilon_{\boldsymbol{v}} = 0.1$ appear to be most consistent with the overlaid Fourier-Galerkin PDE results, with variation in $\varepsilon_{\boldsymbol{v}}$ in the range shown appearing to cause much greater variation in the results. Note that the physical model provides less direct justification for the choice of $\varepsilon_{\boldsymbol{v}}$, so we have tested here over a wider range of parameters.



(a) (b)

Figure 7.7.: Mean number of clusters, μ_K, for particle ODE simulations with inverse-cube kernel and $N = 20$ varying the velocity tolerance, $\varepsilon_{\boldsymbol{v}}$, and spatial tolerance, $\varepsilon_{\boldsymbol{x}}$, for two particles to be linked in a cluster by the DBSCAN algorithm. Fourier-Galerkin PDE results are shown as the dashed black line for comparison. (a) Varying $\varepsilon_{\boldsymbol{v}}$ with $\varepsilon_{\boldsymbol{x}} = 10$. (b) Varying $\varepsilon_{\boldsymbol{x}}$ with $\varepsilon_{\boldsymbol{v}} = 0.1$.

7.1.4. Fast Fourier parameter convergence

We now consider the convergence of results from the fast Fourier spectral implementation of the PDE model over three parameters: the maximum Fourier frequency, V, the number of radial grid points in the Gauss-Legendre quadrature for the du integral, N_u, and the number of angular grid points in the quadrature of the d$\hat{\boldsymbol{u}}$ integral, $N_{\hat{\boldsymbol{u}}}$. Figure 7.8 shows plots of the computation time for the fast Fourier method over different values of these parameters. The computation time is seen to be scaling as expected for higher values of each of the parameters: linearly in N_u and $N_{\hat{\boldsymbol{u}}}$, and proportional to $V^2 \log(V)$ (due to the FFT implementation) in V. The fact that the scaling doesn't apply as well for smaller values of each of the parameters may suggest that there are other components of the implementation that more significantly impact the time taken when the overall computation time is small.

Figure 7.9 shows kinetic energy per particle, E, and mean number of clusters, μ_K, for the Fourier implementation of the PDE model with an inverse-cube kernel, $N = 20$ particles, $N_u = 20$ Gauss-Legendre radial quadrature points and $N_{\hat{\boldsymbol{u}}} = 24$ angular quadrature points across different maximum Fourier frequencies, V. We can see that results show much better agreement with the particle ODE simulations for larger V. Results for $V = 5$ and $V = 10$ show spurious behaviour, such as E going negative and E or μ_K increasing when they should monotonically decrease. The mean kinetic energy result for $V = 20$ appears to approach that

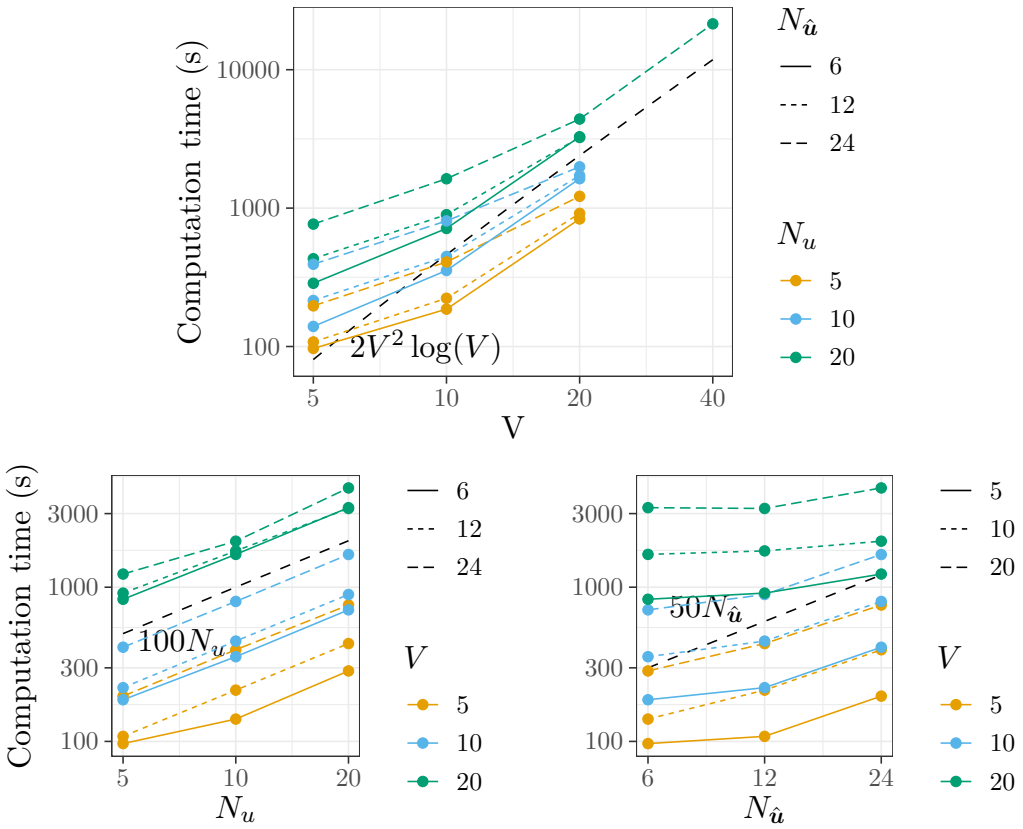


Figure 7.8.: Computation time for Fourier-Galerkin PDE simulations of a system with $N = 10$ particles, varying the maximum Fourier frequency, V, the number of radial quadrature grid points, N_u, and the number of angular quadrature grid points, $N_{\hat{u}}$. The expected scaling relationships are shown with overlaid dotted lines.

of the particle ODEs more accurately than that of $V = 30$, but we believe this is a spurious effect as it involves the E increasing slightly and the result with $V = 30$ better matches the qualitative shape of the energy decrease. Observing the velocity distributions produced, we see that the Fourier approximation begins to break down when these unexpected effects occur, which makes the numerical implementation unstable. This is generally due to the velocity distribution becoming too narrow to be adequately approximated by the truncated Fourier series. Modifying the original model with a similarity scaling before deriving the fast Fourier method may allow the results to converge for longer times at lower values of V.

Figure 7.10a shows the convergence of E and μ_K as V is increased by plotting the error relative to the results with $V = 30$. This shows specifically the reduction in Fourier approximation errors at larger V and convergence of the outputs towards the large-V limiting values. Reference curves at seventh-order convergence are overlaid, showing the output converges quickly as V is increased, as expected from a spectral method.

However, the Fourier approximation of $\check{n}$ was not the only approximation used in the derivation of the model. Figure 7.10b shows the ratio between the error relative to the PDE model with $V = 30$ (as shown in Figure 7.10a) and the error relative to the particle ODE

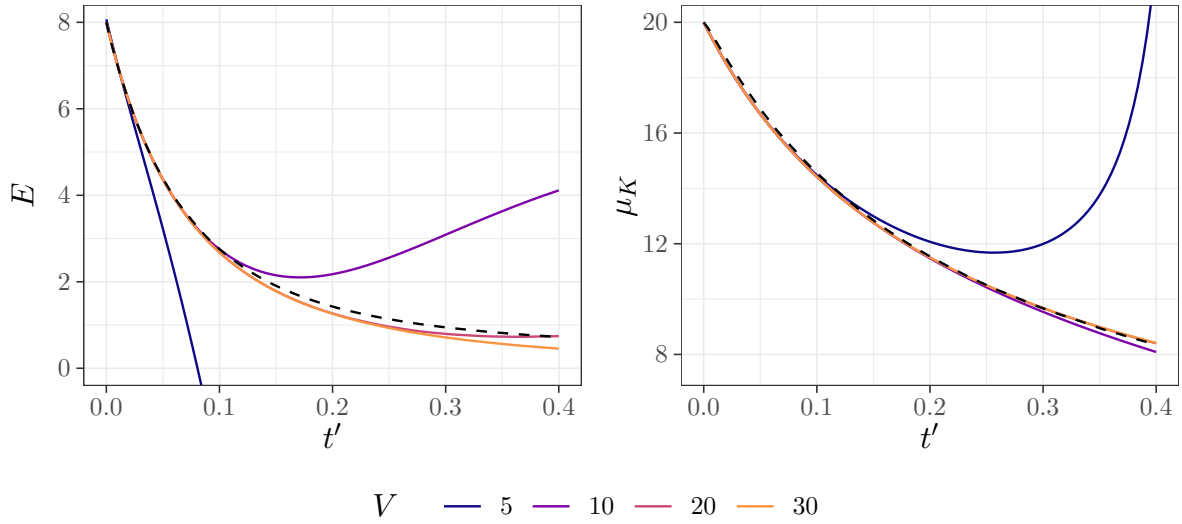


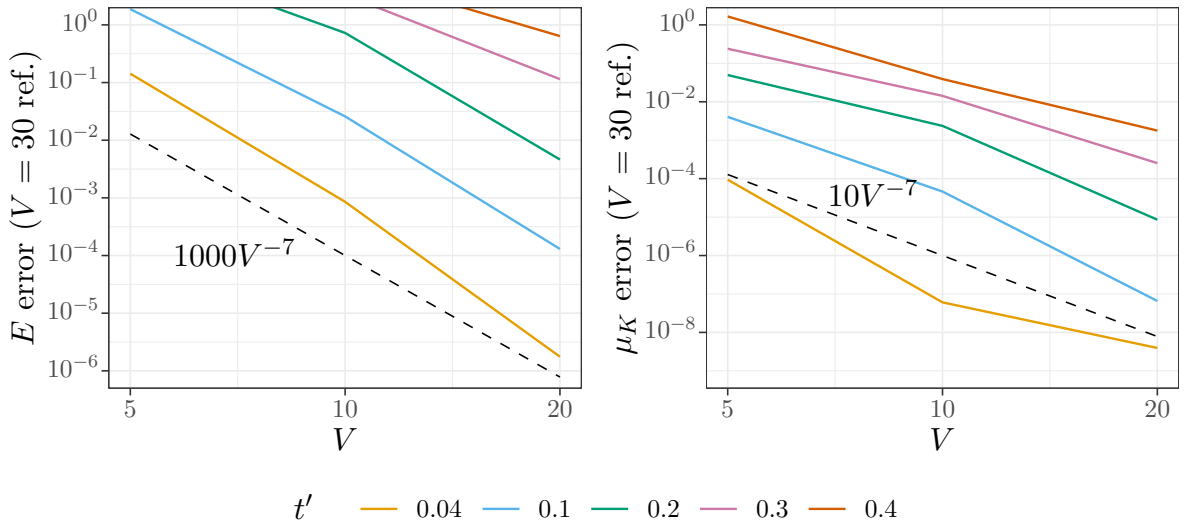
Figure 7.9.: Mean kinetic energy per particle, E, and mean number of clusters, μ_K, for the Fourier-Galerkin PDE model with an inverse-cube kernel, $N = 20$ particles, $N_u = 20$ Gauss-Legendre radial quadrature points and $N_{\hat{\boldsymbol{u}}} = 24$ angular quadrature points. Results from the corresponding particle ODE simulations are shown as the black dashed line. As the maximum Fourier frequency, V, is increased, the results show better agreement with the particle simulations.

simulations. That is, using E as an example,

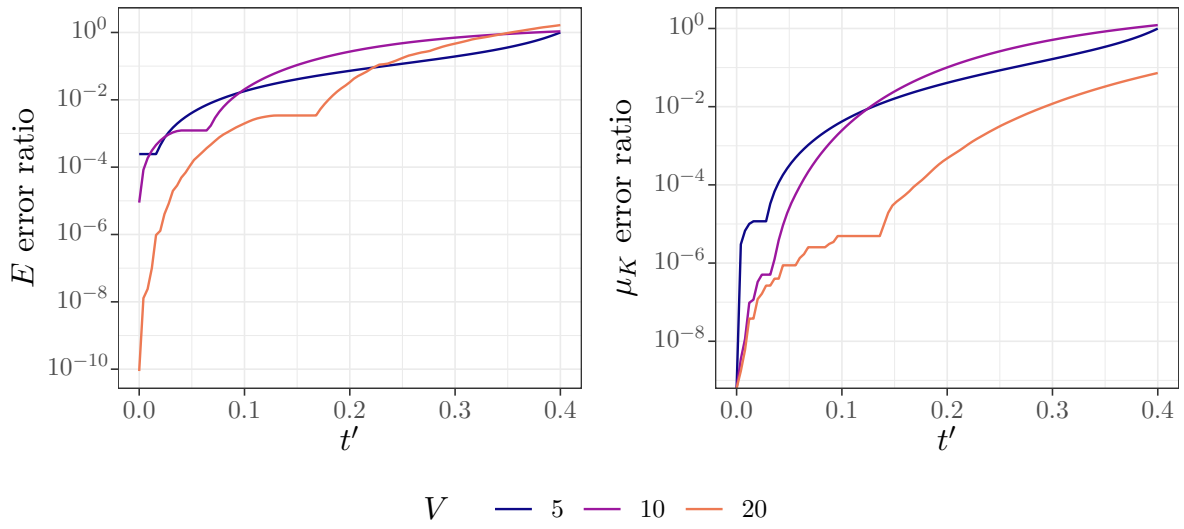
$$E \text{ error ratio} = \frac{|E_V^{PDE} - E_{V=30}^{PDE}|}{|E_V^{PDE} - E^P|},$$

where E_V^{PDE} is the PDE result with $V = 5$, 10 or 20 that we are testing. An error ratio significantly less than 1 implies that the error due to the Fourier approximation relative to that with $V = 30$ is small compared to the total error from all approximations in the model. This would suggest that such a value of V is appropriate for this test case as it does not represent a major proportion of the error. We can see from the plots that, although the μ_K error ratio is small for $V = 20$, all the E error ratios reach 1 in the time frame considered. This means the Fourier approximation is a major contributor to the total error as t' approaches 0.4, and we should use at least $V = 30$ in the subsequent experiments.

Figures B.5 and B.6 in the appendix show similar analysis for the number of Gauss-Legendre radial quadrature points, N_u, and the number of angular quadrature points $N_{\hat{\boldsymbol{u}}}$, respectively. A similar order of convergence (around sixth- or seventh-order) is observed for smaller values of t' for both parameters; however, for $t' = 0.4$, the results converge less quickly. This may be because, by this time, errors from the underlying Fourier approximation have increased, confounding our ability to observe independent convergence with the quadrature parameters. The error ratio analysis suggests values of $N_u = 20$ and $N_{\hat{\boldsymbol{u}}} = 24$ are sufficient as they result in the quadrature contributing only a small proportion of the total model error. We



(a)



(b)

Figure 7.10.: Convergence of Fourier-Galerkin PDE results across maximum Fourier frequency, V, for a system with inverse-cube kernel, $N = 20$ particles, $N_u = 20$ Gauss-Legendre radial quadrature points and $N_{\hat{\boldsymbol{u}}} = 24$ angular quadrature points. (a) Proportional error for different times relative to the Fourier-Galerkin PDE results with $V = 30$ maximum frequency. Reference lines at seventh-order convergence are shown. (b) Ratio between the absolute error with reference to the $V = 30$ Fourier-Galerkin PDE results and the error with reference to the particle ODE simulations. If this ratio is significantly less than 1, it implies that the error due to the Fourier approximation is small compared to the other approximations in the model. When t' approaches 0.4, this error ratio approaches 1 for the values of V shown, implying that the Fourier error has become significant for $V \leq 20$.

will use these values and $V = 30$ for the subsequent experiments for the fast Fourier spectral implementations.

7.2. Model evaluation

7.2.1. Convergence with N

Given the model parameters selected, we now consider how accurately the results produced by the PDE model match the particle ODE simulations for different physical system parameters. We first consider whether the PDE model performs better as the number of particles is increased (while keeping the density constant). Figure 7.11 shows reference particle ODE simulations for different values of N. Particles are drawn with a radius of ε relative to the spatial domain size to show the density, and particles are coloured based on clusters computed by the DBSCAN algorithm.

Figure 7.12 shows the convergence in mean kinetic energy results between the PDE model and particle ODE simulations as N is increased. Figure 7.12a shows the values of E for each model across N, and the plots in Figure 7.12b show the proportional error between the models across t' and N. First, we can see that the agreement is much closer than that observed for the one-dimensional models in Chapter 4. This suggests that the local particle correlations are less significant and that the *Stosszahlansatz* is a better approximation in two dimensions. The error increases for later times and decreases for larger N. Reference lines are added over the second plot in Figure 7.12b, showing that the rate of convergence with N is between $N^{-2/3}$ and N^{-1}. For earlier times, the convergence with N is not as clear as the magnitude of the errors is much smaller, potentially allowing errors from other approximations to confound the result.

Similarly to the one-dimensional case, any discrepancy is in the direction of the PDE models underestimating the mean kinetic energy. This may be explained by similar reasoning: the PDE model does not model local spatial correlations and allows particles to continue interacting with a distribution of cluster velocities that may not be present in any one realisation, allowing the variance of velocity distribution to reduce too far. This evidence in Figure 7.12, however, suggests this effect is much smaller than in one-dimension, and, in this case, the two models appear to be converging to the same value as N is increased.

Figure 7.13 shows similar plots for the mean number of clusters. Once again, the results from the PDE models track those from the particle ODE simulations significantly more accurately than in the one-dimensional case. However, the errors observed are smaller across all values of N than for E (with relative errors on the order of 0.01), and no increase in accuracy is observed as N is increased. The right plot in Figure 7.13 shows the proportional error in μ_K with sign included. For all values of N, the PDE model slightly underestimates the value of μ_K relative to the particle ODEs initially, before gradually overestimating as t' increases.

These observations are consistent with the velocity and mass distributions produced by the two models. Figures 7.14 and 7.15 show a cross-section of the normalised cluster velocity distributions, $n(v, 0)/\mu_K$, and normalised mass-weighted cluster mass distributions $f(m)/N$.

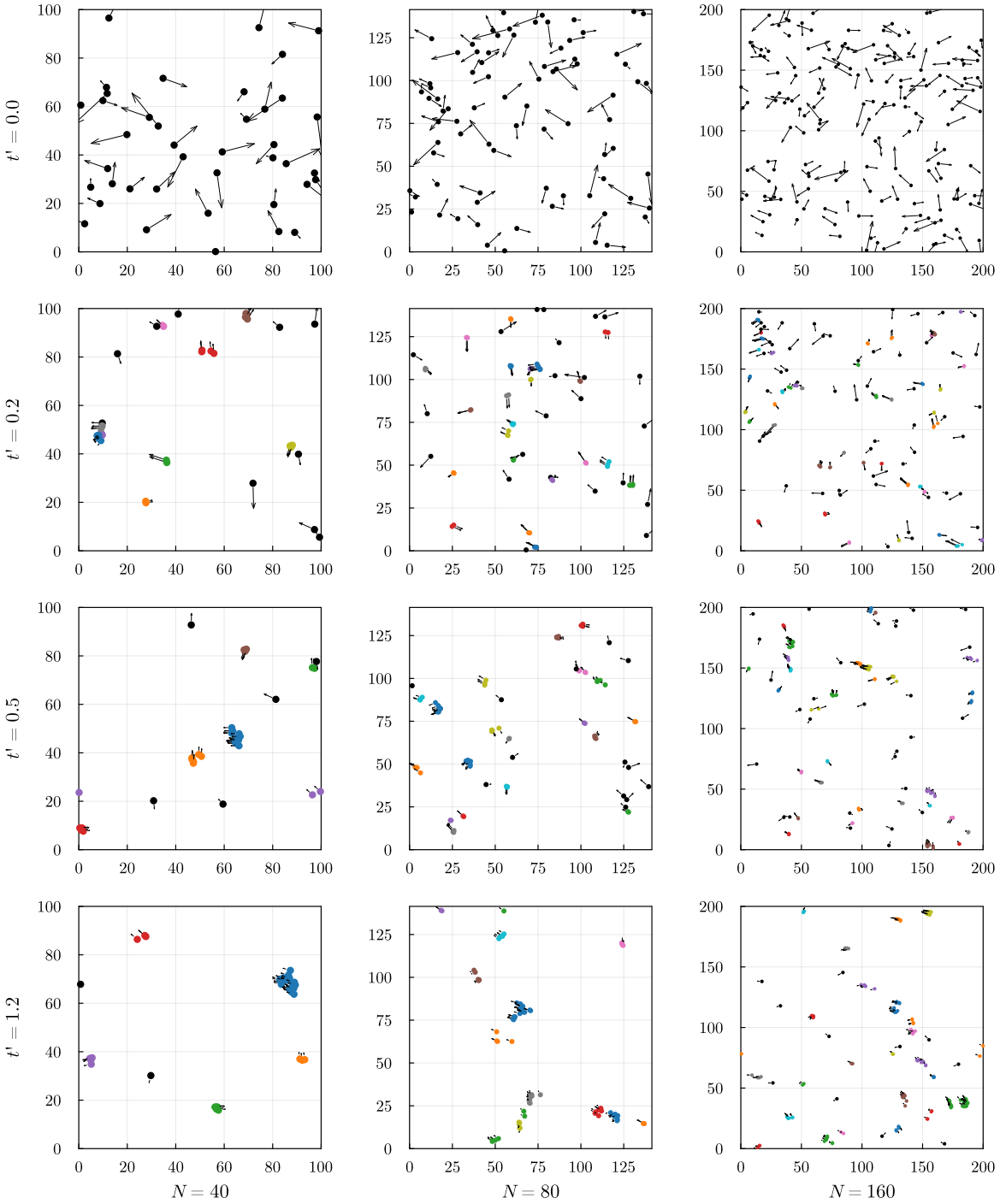


Figure 7.11.: Particle ODE simulations for $N = 40$, 80, and 160 particles with the inverse-cube kernel, density $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ and a Maxwellian initial velocity distribution. Clusters identified by the DBSCAN algorithm are coloured, and the arrows represent particle velocities. Similar simulations are shown in Figures 7.1 and 7.2. Animated versions of these particle simulations for different values of N can be found at https://figshare.com/s/1c7db37edaba0605dc60.

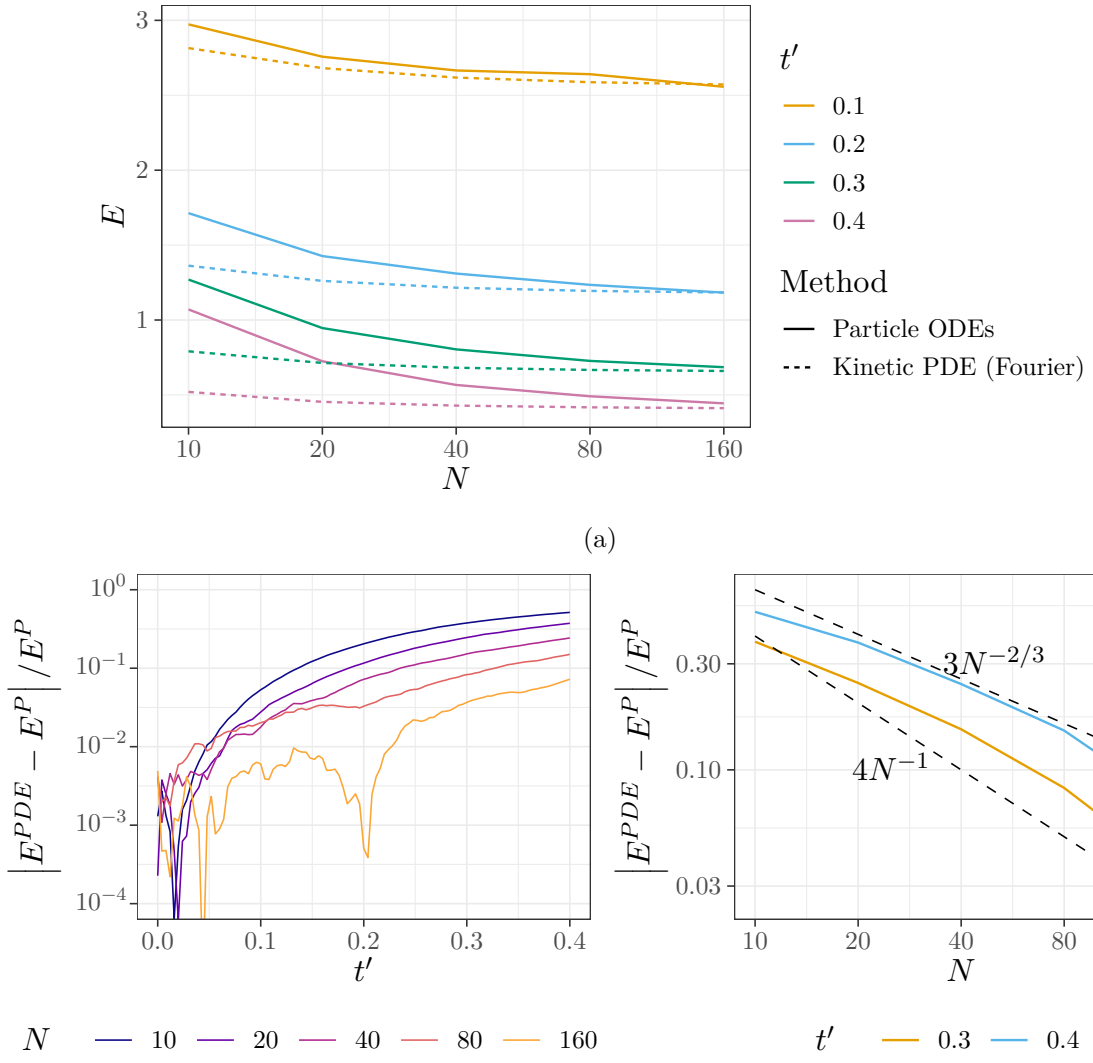


Figure 7.12.: Mean kinetic energy per particle across number of particles, N, for an inverse-cube interaction kernel, density $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ and a Maxwellian initial velocity distribution. The Fourier-Galerkin method is applied with $V = 30$, $N_u = 20$ and $N_{\hat{\boldsymbol{u}}} = 24$. (a) Particle energy at different values of t', showing the convergence of the PDE model results to the particle simulations as N is increased. (b) Proportional error between PDE model results and particle simulations over time and against N. Reference lines for N^{-1} and $N^{-2/3}$ convergence are shown.

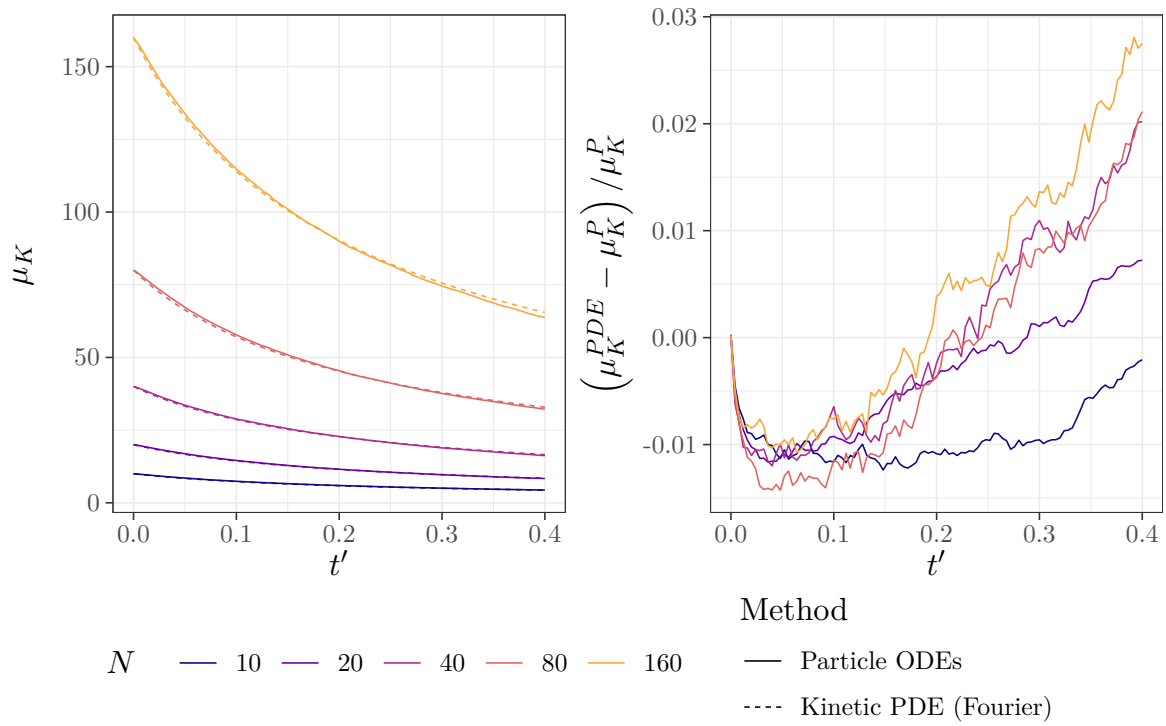


Figure 7.13.: Mean number of clusters across number of particles, N, for an inverse-cube interaction kernel, density $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ and a Maxwellian initial velocity distribution. Good agreement between the PDE model results and the particle ODE simulations is seen for all values of N considered. The right plot shows the proportional error between the PDE and particle values.

The full velocity distributions are two-dimensional, but as they would be expected to be radially symmetric for this system, the cross-section $n(v, 0)$ shown should be representative. A comparison of cross-sections taken at different angles through the origin showed this to be the case. Figure B.7 in the appendix shows the normalised cluster speed distribution as an example that reduces the full distributions to one dimension.

Consistent with the accuracy of the μ_K results for all values of N, the PDE model produces mass distributions that match well with the distributions from the particle ODE simulations. To further test this, running the experiments to larger values of t' when the mean cluster mass is larger would be desirable. Such tests, however, would require the implementation of a similarity scaling for the fast Fourier model or the use of more Fourier modes (larger V) as the current plots are reaching the limit of how sharp the velocity distribution can become and still be well modelled by the Fourier approximation.

The velocity distribution results are also consistent with the scalar comparison of mean kinetic energy. The velocity distribution from the PDE model matches the particle ODE simulations well for early times. However, at later times, it begins to overestimate how far the variance should reduce towards 0. This overestimation is significantly smaller for $N = 160$ than for $N = 20$, mirroring the reduction in error of the mean kinetic energy prediction as N is increased. The small bumps observed in the tail of the PDE model distributions at later times are due to the Fourier approximation. If t' is increased further, these will grow as the model attempts to approximate a distribution too sharp for the number of Fourier modes used. This is what is happening for the times where spurious results can be observed in Figure 7.9, such as E increasing or going below 0. Naturally, we would expect to be able to extend the PDE model to much later times without encountering these errors if it were implemented using a similarity scaling.

Figure 7.16 shows the computation time needed to produce the results shown in this section for both the particle ODE simulations and fast Fourier kinetic PDE model across different values of N. This is intended only as a heuristic reference as many factors significantly influence the computation time, including the implementation, computational hardware and level of accuracy desired. The PDE model can be seen to take more computation time than the particle ODEs for the range of N considered. However, the advantage of the PDE model is that, with mass binning, we would expect to be able to model systems for much higher values of N without increasing the number of mass bins or affecting the computation time significantly. Although the computational time of the particle simulations is naturally heavily dependent on the number of simulations run, the time taken would scale at least at $O(N^2)$ as N was increased further. As such, we would expect the PDE model to be more efficient for systems with on the order of hundreds of particles or more. As seen in Figure 7.12, the results from the PDE model also converge to those from the particle ODE simulations as N is increased, suggesting the PDE model is useful for systems with N in this range.

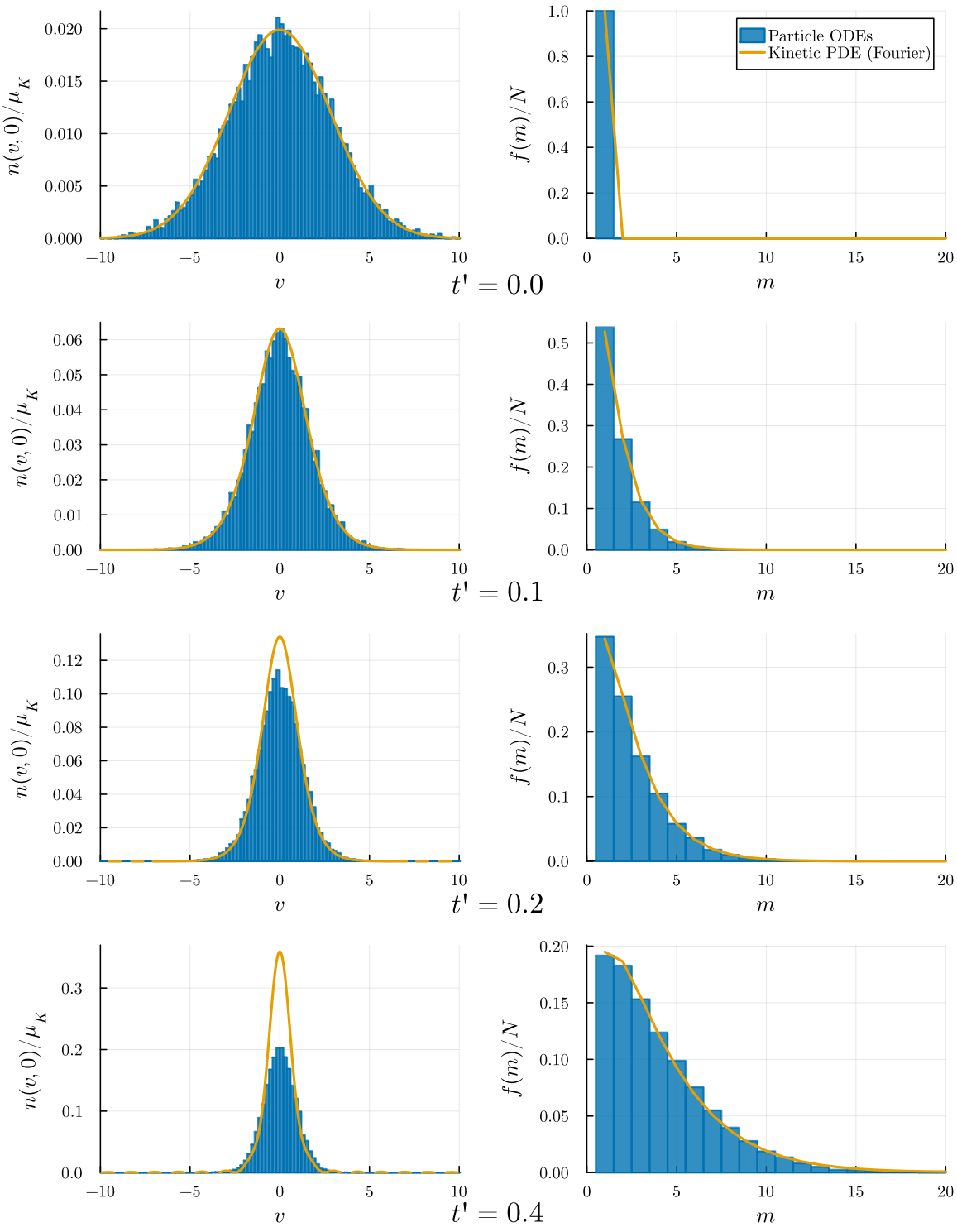


Figure 7.14.: Normalised cross-section of two-dimensional velocity distribution $n(v,0)/\mu_K$, and normalised mass-weighted cluster mass distribution $f(m)/N$ for simulations with $N = 20$ particles, $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ density, inverse-cube interaction kernel and a Maxwellian initial velocity distribution. Good agreement is seen between the PDE model and particle simulations of the mass distribution. At later times, the PDE model can be seen to under-predict the variance in the velocity distribution when compared to particle simulations. An animated version can be found at https://figshare.com/s/45630708860ad2aae02b.

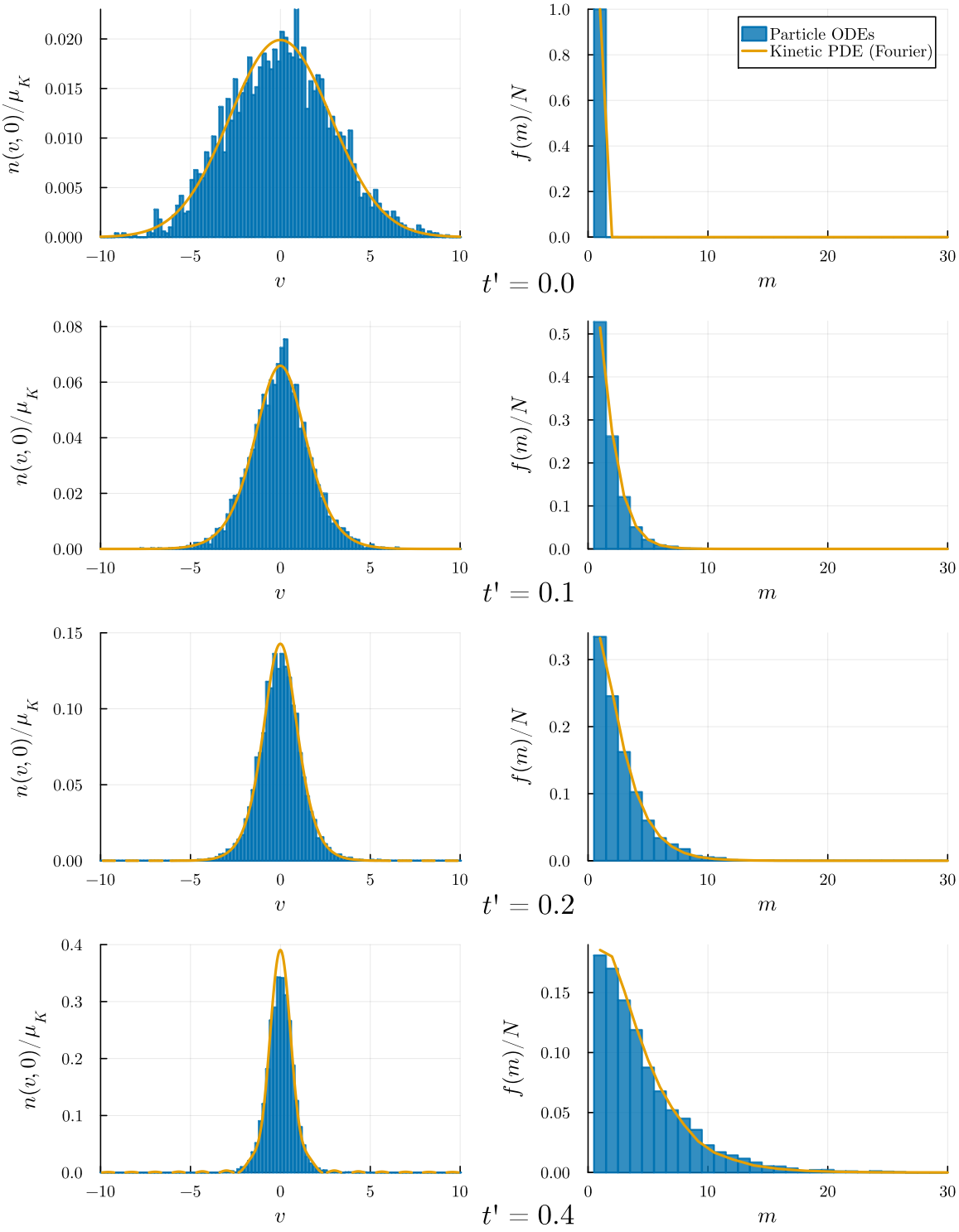


Figure 7.15.: Normalised cross-section of two-dimensional velocity distribution $n(v,0)/\mu_K$, and normalised mass-weighted cluster mass distribution $f(m)/N$ for simulations with $N = 160$ particles, $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ density, inverse-cube interaction kernel and a Maxwellian initial velocity distribution. Good agreement is seen between the PDE model and particle simulation results in general. The velocity distribution is predicted more accurately by the PDE model than with $N = 20$. An animated version can be found at https://figshare.com/s/7915e31bc94db0590939.

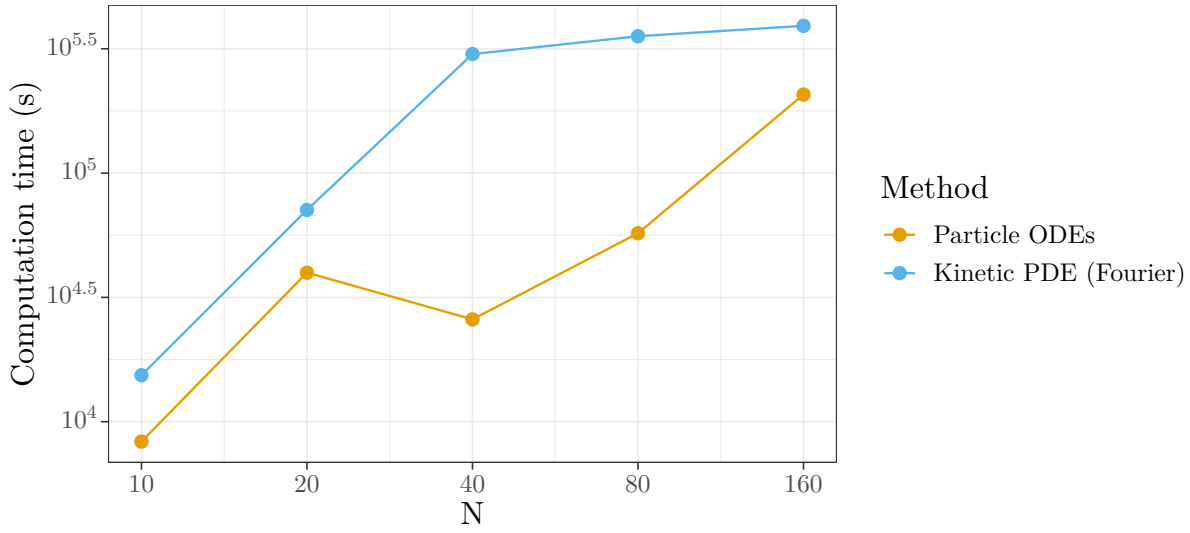


Figure 7.16.: Computation time for the fast Fourier implementation of the PDE model and particle ODE simulations across different values of N, for the Cucker-Smale model with inverse-cube interaction kernel. The PDE model uses a maximum Fourier frequency of $V = 30$ and no mass binning for $N \leq 40$. For the model with $N = 80$ and $N = 160$, $M = 30$ mass bins are used. The number of particle ODE simulations performed for each value of N are: $N = 10$, 1000; $N = 20$, 1000; $N = 40$, 1000; $N = 80$, 500; $N = 160$, 400.

7.2.2. Variation of density

Next, we consider how the results vary if the spatial density of particles in the system is varied. Figure 7.17 shows example particle ODE simulations with $N = 20$ particles and densities ranging from $N\varepsilon^2\pi/|\Omega| = 0.016\pi$ to $N\varepsilon^2\pi/|\Omega| = 2.5 \cdot 10^{-4}\pi$. Figure 7.18a shows plots of E and μ_K from both models for different system densities. The mean kinetic energy results of the two models do not show any significant variation across density, while the mean cluster number results show some variation at either end of the density range. Figure 7.18b shows the proportional error of μ_K between the two models over t' for different densities, with greater overall variation seen for higher densities. The consistency of the results at lower densities suggests that the assumptions that higher-order interactions and the spatial extent of clusters are negligible are both valid in this density range. The increased μ_K variation at the relatively higher densities tested suggests these assumptions may begin to break down at this point.

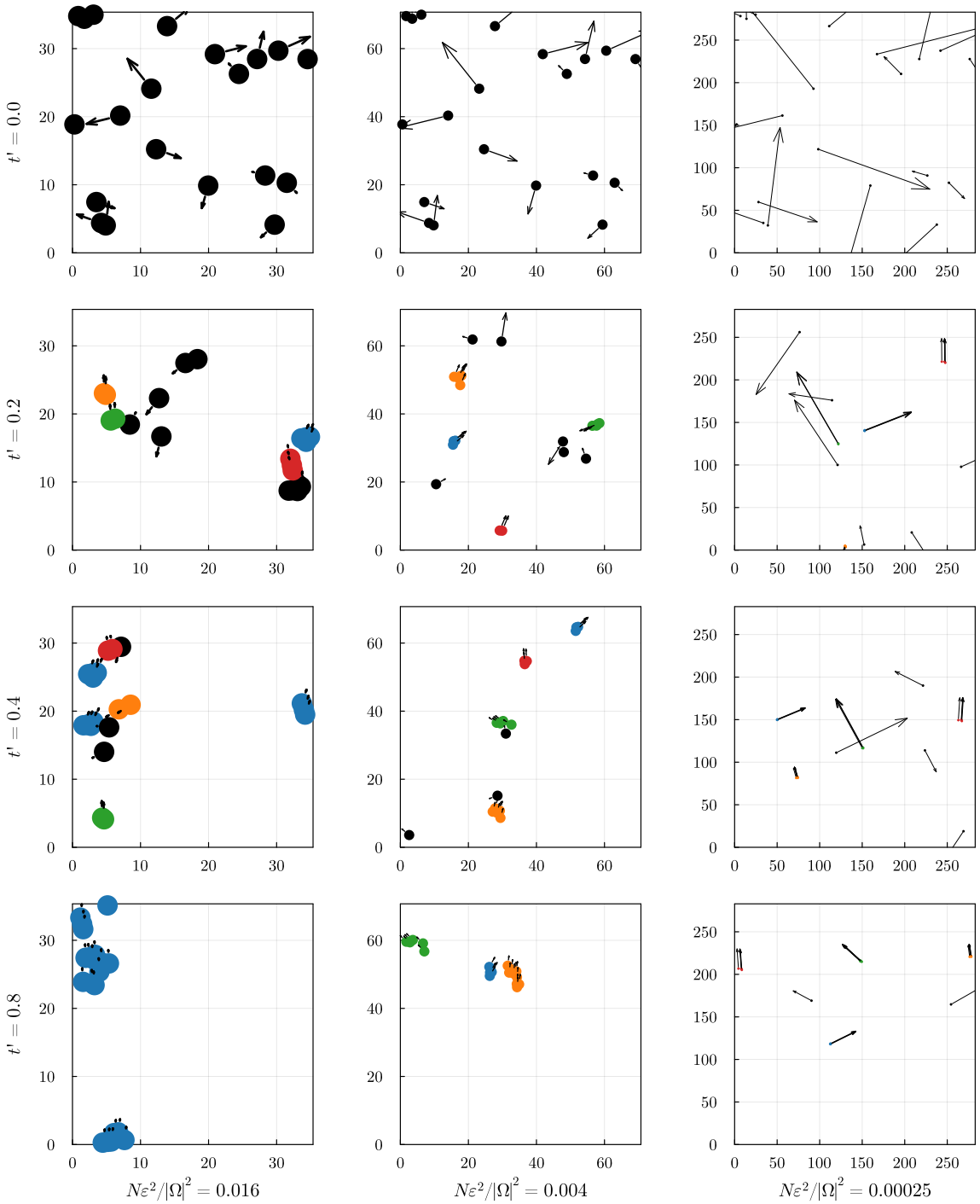
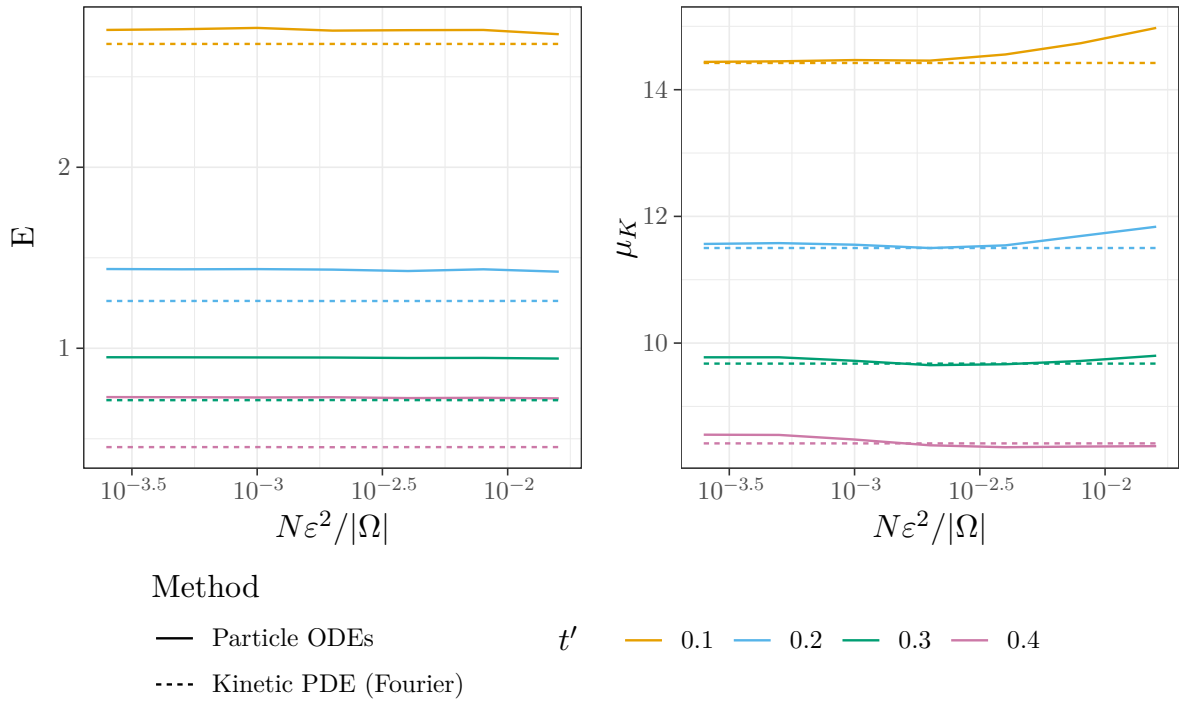
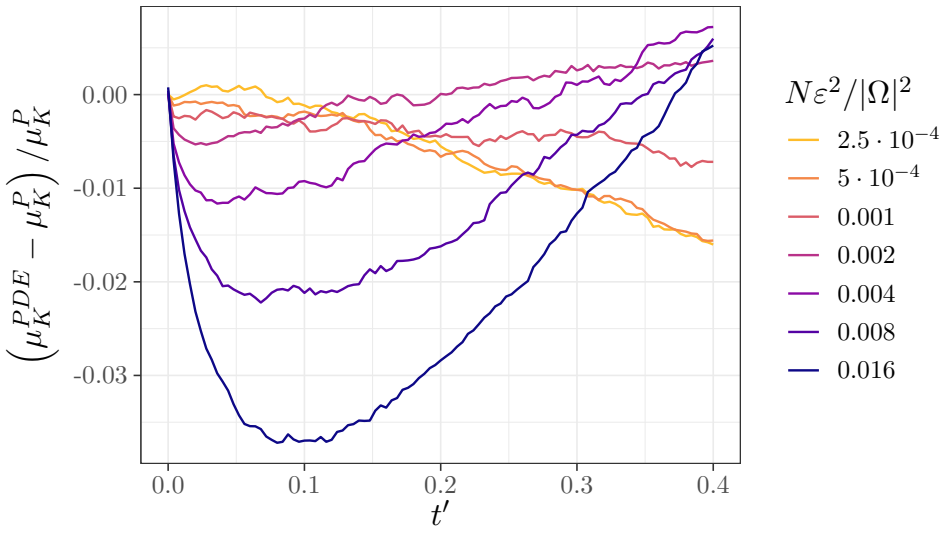


Figure 7.17.: Particle ODE simulations for $N = 20$ particles with density ranging from $N\varepsilon^2\pi/|\Omega| = 0.016\pi$ to $2.5 \cdot 10^{-4}\pi$, inverse-cube collision kernel and Maxwellian initial velocity distribution. Clusters identified by the DBSCAN algorithm are coloured, and the arrows represent particle velocities. Animated versions of these simulations for different densities can be found at https://figshare.com/s/11fe4376fadf70d28302.



(a)

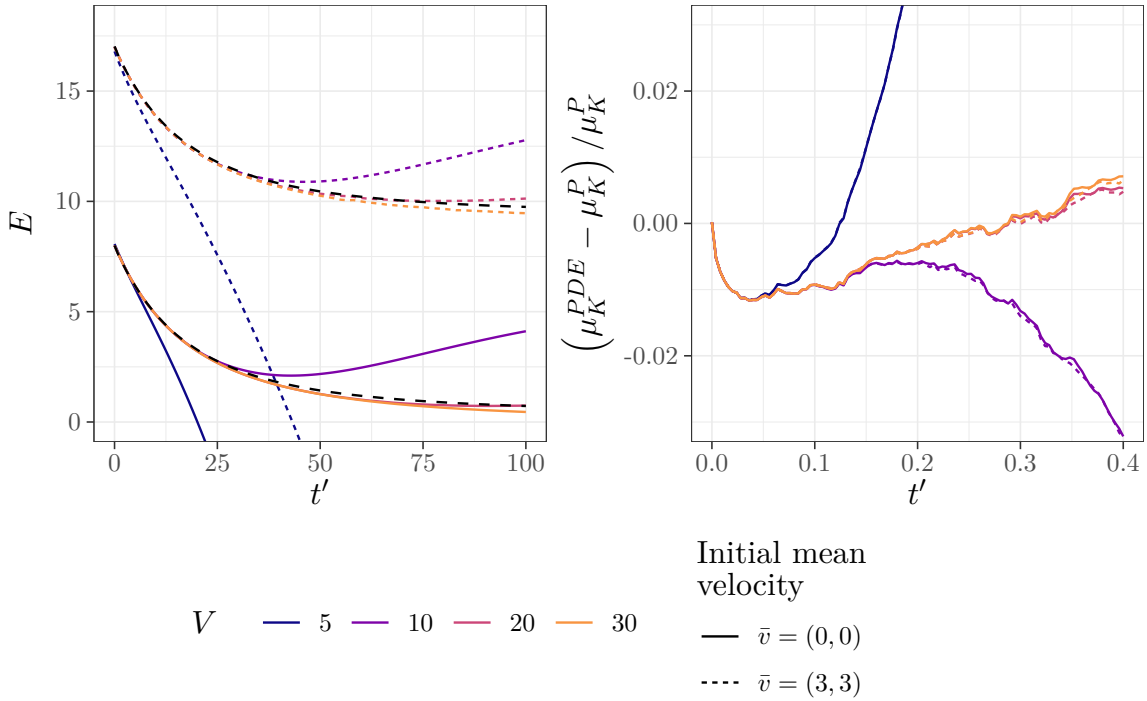


(b)

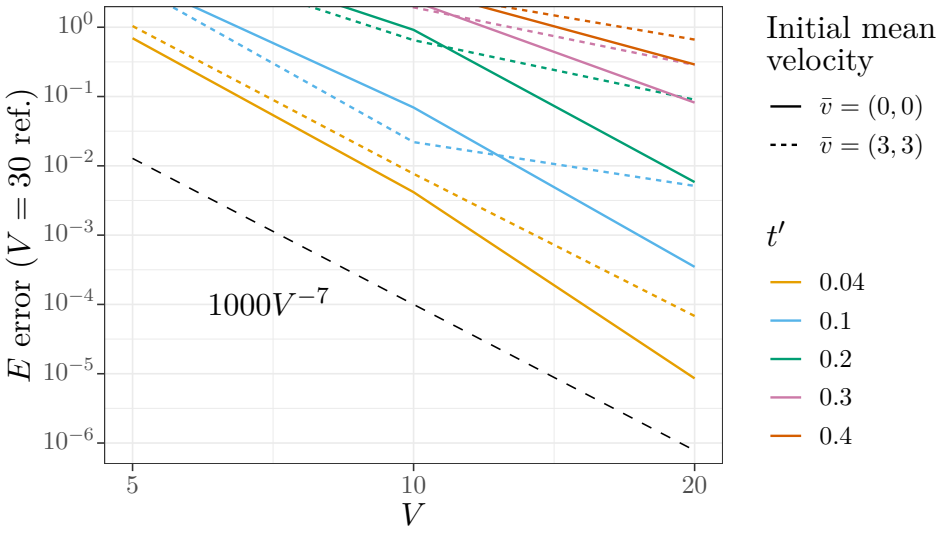
Figure 7.18.: Particle energy, E, and number of clusters, μ_K, for a system with $N = 20$ particles and inverse cube collision kernel across a range of densities. (a) Comparison of the kinetic PDE with particle ODE simulation results at different values of t'. There is little variation for different density values other than at the higher densities and lower times for μ_K. (b) Proportional error of kinetic PDE results against particle simulation results over time for μ_K. Slightly more variability is seen at the higher densities.

7.2.3. Anisotropic initial velocity condition

To examine the effect of an anisotropic velocity distribution, we consider shifting the mean velocity of the Maxwellian initial velocity to $\bar{v} = (3, 3)$ instead of $\bar{v} = (0, 0)$. As the Cucker-Smale model conserves momentum and interactions are based on relative velocity, the evolution of the velocity and mass distributions would be expected to be the same over time, except with the mean of the velocity distribution remaining shifted. Figure 7.19a shows the mean kinetic energy per particle and the mean number of clusters over time for both initial conditions and varying maximum Fourier frequency, V. The qualitative behaviour and convergence in V towards the particle ODE simulation results are very similar in both cases, indicating the model is not significantly affected by this anisotropy. Figure 7.19b shows the convergence of kinetic energy results over V, with the $V = 30$ PDE results used as reference. The convergence is similar but slightly worse in the anisotropic case. This may also be because the initial condition is less well approximated by restricting to the ball $\mathcal{B}_R$ with $R = 10$ when the mean of the distribution is shifted to $\bar{v} = (3, 3)$.



(a)



(b)

Figure 7.19.: Comparison of results between the Fourier implementation of the kinetic PDE model and particle ODE simulations for an anisotropic initial velocity condition with the mean velocity shifted to $\bar{v} = (3, 3)$. Other model parameters are the same as Figure 7.9 with $N = 20$. (a) Particle energy comparison shown with particle ODE simulation results as black dashed lines. Mean cluster number results are shown as proportional errors between PDE results and particle simulations. Limited difference is seen between the two initial conditions. (b) Convergence of the PDE model measured for different values of V, with the $V = 30$ implementation as reference. The results for the anisotropic initial velocity converge similarly but slightly slower than the isotropic initial condition.

7.2.4. Interaction kernels

We now consider how the model performs if we vary the Cucker-Smale interaction kernel. Figure 7.20 shows the comparison between Fourier PDE results and particle ODE results for the mean particle kinetic energy and mean number of clusters for the inverse-cube, $\phi(|\boldsymbol{x}|) = 1/|\boldsymbol{x}|^3$, inverse-quintic, $\phi(|\boldsymbol{x}|) = 1/|\boldsymbol{x}|^5$, and exponential, $\phi(|\boldsymbol{x}|) = e^{-|\boldsymbol{x}|}$, interaction kernels. The simulations are performed with $N = 20$ particles, $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ density and the implementation parameters selected in §7.1.

The kinetic energy results show even better agreement for the inverse-quintic kernel than the inverse-cube kernel, while the mean cluster number comparison shows similar accuracy. We note that in §7.1.3, the μ_K results for the particle ODE simulations showed some variation with the DBSCAN clustering tolerance, $\varepsilon_{\boldsymbol{v}}$, and the value $\varepsilon_{\boldsymbol{v}}$ was selected with reference to experiments with the inverse-cube interaction kernel. It is possible that a slightly different value of $\varepsilon_{\boldsymbol{v}}$ might be suggested to be more appropriate when considering simulations with the inverse-quintic kernel.

For the exponential kernel, the agreement in mean cluster number results between the models is similar to that from the power-law kernels, but the agreement in the kinetic energy results is much worse. One of the main reasons this could be the case is that, for the power-law kernels, we have analytical results for the integrals in the Fourier kernels (see §6.5.2), but, for the exponential kernel, the Fourier kernel modes had to be approximated through interpolating on a grid of numerical values. The Fourier kernel integral for the exponential interaction kernel, (6.22), is highly oscillatory and proved challenging to approximate effectively, and it is possible that errors here could be contributing to the discrepancy in the kinetic energy results. Other potential reasons for the discrepancy are that the value of $\varepsilon_{\boldsymbol{v}}$ and the empirical approximation to Λ_K were selected for the inverse-cube kernel. In Figure 7.4b, we saw that the initial transient of Λ_K as a function of μ_K differs for the exponential kernel. It is also possible that the particle ODE simulations are missing interactions with the exponential kernel due to its sharper fall-off in space. Further investigation is needed to explore this.

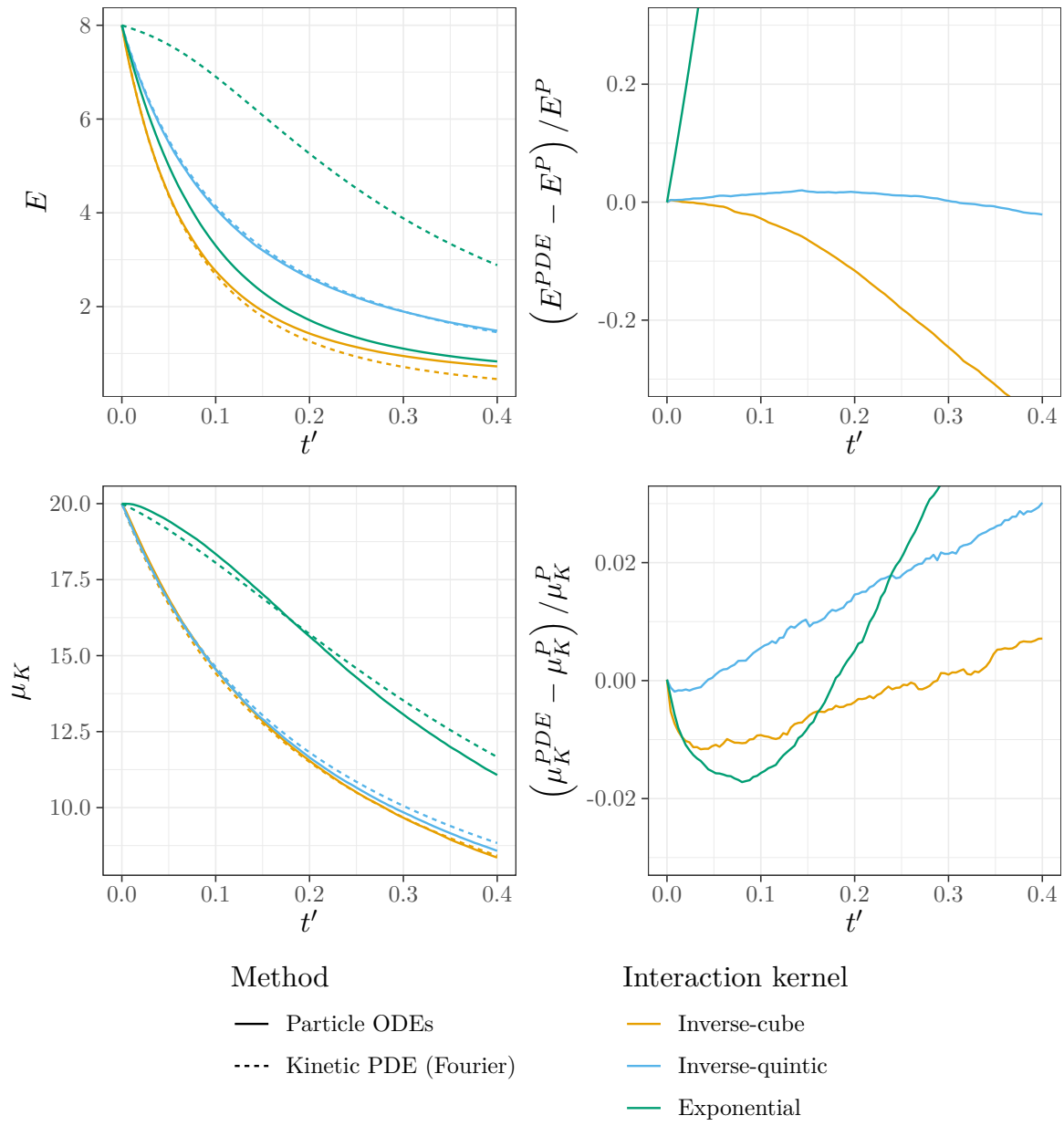


Figure 7.20.: Particle energy, E, and mean cluster number, μ_K, results for different interaction kernels, with $N = 20$ particles. The plots on the right show the same data but as the proportional error of the PDE model results relative to the particle ODE simulations. For the sharper inverse-quintic kernel, the PDE model appears to track the energy more accurately but models the mean cluster numbers less accurately with these parameters (including clustering tolerances $\varepsilon_{\boldsymbol{x}} = 10$ and $\varepsilon_{\boldsymbol{v}} = 0.1$ that were set for the inverse-cube kernel). For the exponential kernel, the PDE model appears to model the mean number of clusters similarly accurately to the power-law kernels but shows more error modelling the kinetic energy. This may be because analytical expressions for the Fourier kernels are available for the power-law kernels but not for the exponential kernel.

7.2.5. Mean-centered initial velocity condition

To consider the effect of the PDE model averaging over the distribution of possible realisations in two dimensions, we consider centring the mean of the initial velocity distribution to $(0, 0)$ for all simulations, similarly to §4.5.1. Figure 7.21 shows the results compared to those from the the non-centred Maxwellian distribution. Both sets of results show much less significant variation between the two initial conditions than observed in the one-dimensional case, suggesting that the effect is less relevant in the two-dimensional case considered here.

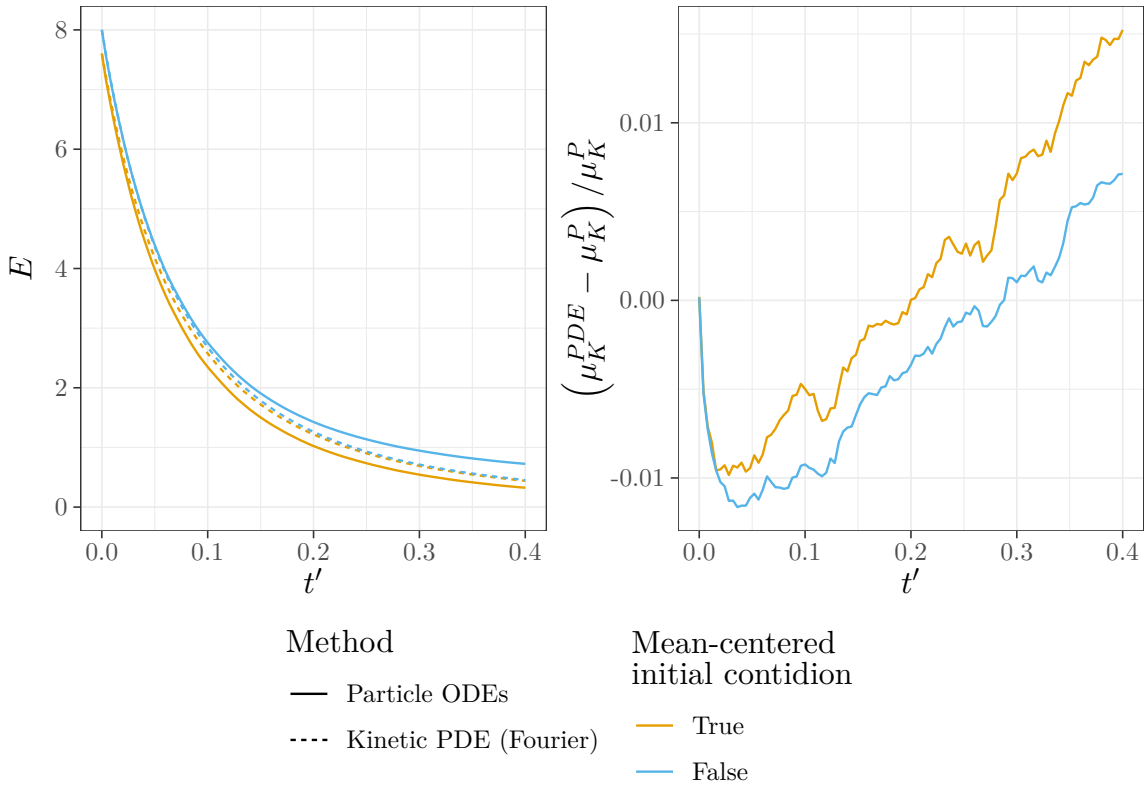


Figure 7.21.: Comparison between particle energy and mean cluster number results with and without centring the initial velocity condition to mean velocity $\bar{v} = (0, 0)$ for each simulation for $N = 20$ particles. Unlike in the one-dimensional case, the mean-centered initial condition doesn't appear to make a significant difference.

7.3. Cucker-Smale model analysis

7.3.1. Velocity distribution tails

One of the most well-known results in the inelastic-Boltzmann literature is the presence of non-Maxwellian velocity distribution tails [109]. This is a counter-intuitive result as it suggests that the presence of inelastic collisions results in a greater proportion of particles at higher velocities and that the separate velocity coordinates are no longer independent variables [109].

In Figure 7.22, we show this effect in the Cucker-Smale model, plotting a cross-section of the velocity distribution for a model with $N = 20$ particles on a log plot for the Maxwellian initial condition, $t' = 0$, and at a later time, $t' = 0.16$. Reference lines of exponential distribution tails, i.e. $n(v, 0) \propto e^{-a|\boldsymbol{v}|}$, are shown for comparison. The initial condition is curved on the log plot, as expected for a Maxwellian distribution. In contrast, the distribution at $t' = 0.16$ much more closely matches the exponential, i.e. non-Maxwellian, tail.

The most similar non-clustering model in the literature is the inelastic hard-sphere model in a homogenous cooling state, as this model also has interaction operators proportional to $|\boldsymbol{v}_2 - \boldsymbol{v}_1|$, and we are not considering adding a heat bath or friction to produce a non-degenerate steady state. For this model, the distribution tails are predicted to be exponential [109, 43], consistent with the result seen here.

The bump in the distribution at $t' = 0.16$ around $v = 6$ is caused by the Fourier approximation. For more extreme values of v (not shown in this plot), the error of the Fourier approximation corrupts the result, and we would require more Fourier modes to model accurately. Nevertheless, it was easier to capture the tail behaviour in the Fourier PDE model than with the particle ODE simulations, as an extremely large number of simulations would be required to produce accurate distribution tails—one of the advantages of the PDE formulation.

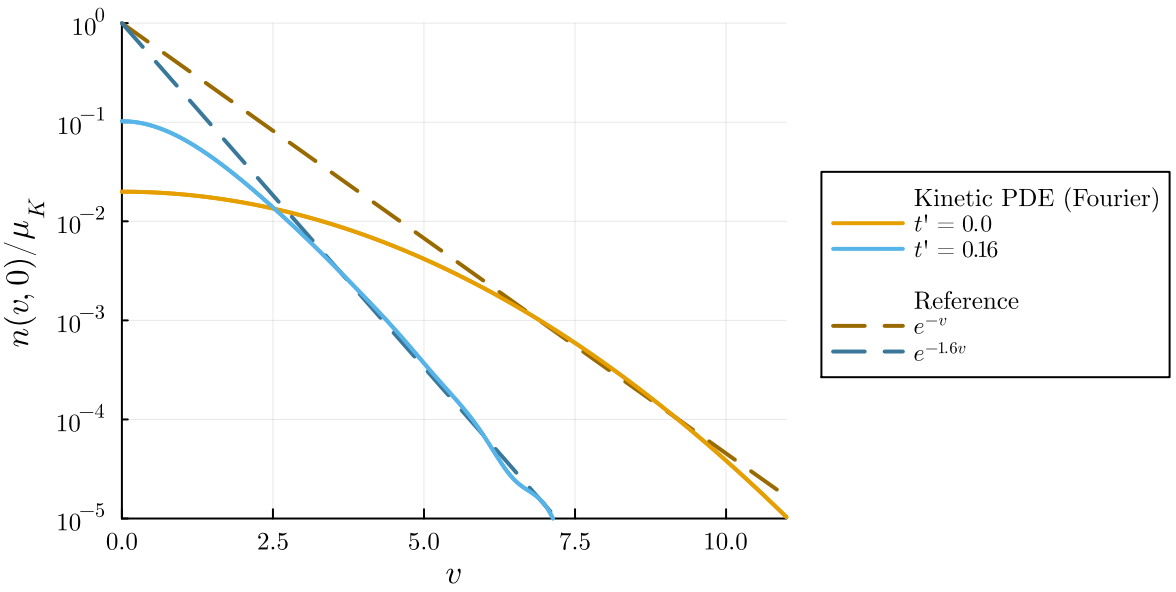


Figure 7.22.: Log plot of normalised velocity distribution cross sections produced by the Fourier-Galerkin PDE model for $N = 20$ particles. At $t' = 0.16$, the tail of the distribution matches the exponentially decaying reference curve, suggesting overpopulated tails as predicted. The bump at around $v = 6$ is caused by the Fourier approximation. This point moves out when more Fourier modes are included. The Maxwellian initial condition is shown for reference, decaying like a quadratic on the log plot.

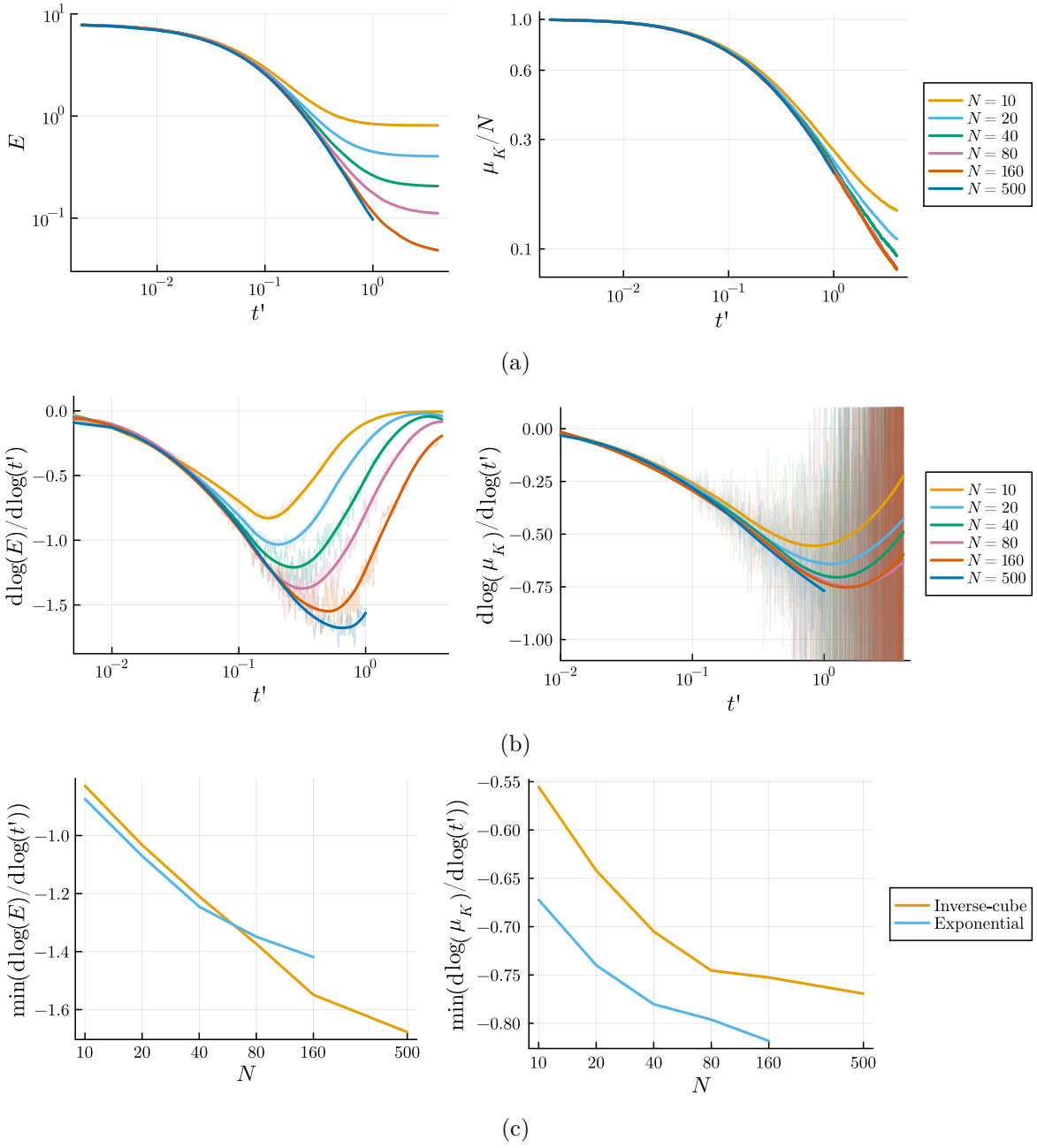


Figure 7.23.: Asymptotic behaviour of particle energy, E, and mean number of clusters, μ_K, for particle ODE simulations across different values of N. (a) Log-log plots showing linear regions where the quantities exhibit power-law decay for the inverse-cube interaction kernel. The plots for $N = 500$ do not span the full range of t' as they take significantly longer to compute numerically. (b) Slope of the log-log plots, i.e. $\mathrm{d}\log(\cdot)/\mathrm{d}\log(t')$, calculated with a smoothing spline on the numerical derivative. The numerical result with no smoothing is overlaid at reduced opacity to show the variability. (c) Minimum values of the log-log slope across different values of N. The $N = 500$ case is not shown for the μ_K plot as the log-log has not reached its minimum slope during the range of t' simulated. The plots over t' used to derive the exponential kernel results are shown in Figure B.8 in the appendix.

7.3.2. Asymptotic behaviour of $E(t)$ and $\mu_K(t)$

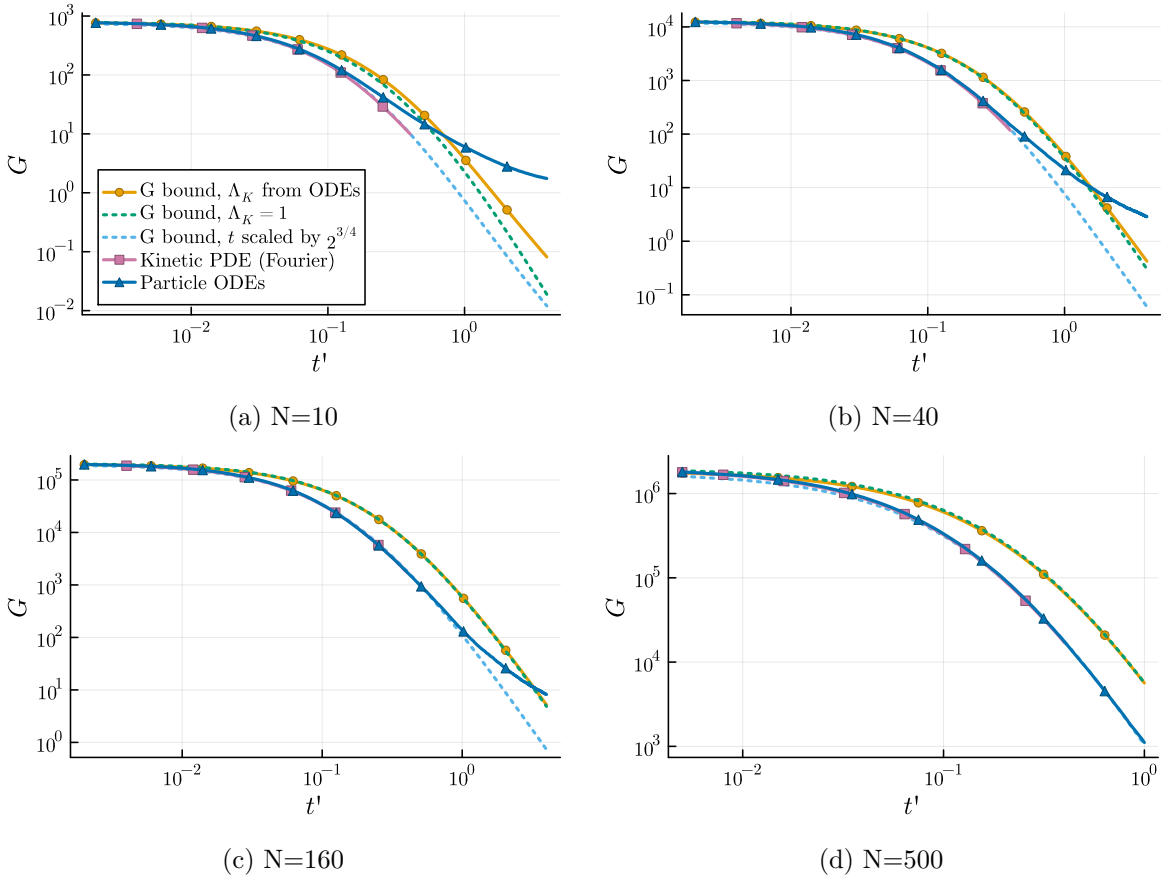
We now investigate the behaviour of $E(t)$ and $\mu_K(t)$ over time for the Cucker-Smale model. Figure 7.23a shows log-log plots of E and μ_K computed from particle ODE simulations of the Cucker-Smale model with inverse-cube interaction kernel across different values of N (Figure B.8 in the appendix shows similar results for the exponential interaction kernel). The plot of μ_K is normalised by N such that all simulations are on the same scale to emphasise the convergence as N increases. The models show a linear region, implying a power-law decay in both quantities for some time range, with the extent of the linear region and the slope increasing as N is increased.

To investigate this further, Figure 7.23b shows the slope of these log-log plots, computed using the numerical derivative. A smoothing spline is used on the results for clarity, with the raw numerical derivative shown with reduced opacity in the background to indicate the variability. The smoothness coefficients of the splines were selected by eye to best match the qualitative shape observed without overfitting to noise in the derivative. These plots indicate that, for both quantities, there is an increase in the negative slope as t' increases. The E results then show a reduction in slope back to 0 for later times, and the slope of μ_K also reduces but not as far. It is possible they would reduce further if the simulations were run to later times. The region of power-law scaling is not as evident in the slope plots as in the log-log plots. These appear to show the derivative reaching a turning point rather than staying at one value for a significant period of time. This, however, may be a result of the variability in the numerical derivative. Increasing the number of simulations would yield a more precise result.

Figure 7.23c shows the minimum slopes of these log-log plots across different values of N. The results for the exponential interaction kernel are also shown for comparison. The plots give some indication that the minima may be converging, but more precision in the derivative estimate and larger values of N would be required to provide an estimate for the limiting value as $N \to \infty$. The value for the minimum log-log plot slope for E is around the value of -8/5 suggested by the estimate in (5.109), but the value of the minimum slope for μ_K is further from the estimate of -6/5 suggested by (5.111). The plots in 7.23b show the slope decreases back to 0 as t' increases further past this minimum, suggesting that for these values of t', the approximations used in deriving the bounds (5.109) and (5.111) begin to break down. As the number of clusters is significantly lower than the initial condition at this point, it's possible local particle correlations are more significant, or the *Clusterzahlansatz* becomes less accurate (as was seen in Figure (4.6)).

This could be investigated further by increasing the number of numerical simulations to gain more precision in the log-log plot slope and to run simulations for larger values of N. However, as it is much easier to extend the PDE model to large N, and the results from §7.2.1 suggest the outputs converge to those from the particle simulations as N is increased, our next step would be to model these quantities using the PDE model. As the potential power-law scaling region is observed for larger values of t', to perform this analysis, we would either need to increase the number of Fourier modes or scale the velocity variable (e.g. if a variation on the

similarity scaling from §5.3.6 was possible) such that the variance of the velocity distribution does not become too small for the times considered.



(a) N=10

(b) N=40

(c) N=160

(d) N=500

Figure 7.24.: Plots of the combined quantity $G(t) := \mu_K^2(t)E(t)$ for particle ODE simulations of the Cucker-Smale model with the inverse-cube interaction kernel. The upper bound for G derived from the kinetic PDE model, (5.104), is shown for comparison. The yellow curve shows the upper bound with the integral of $\Lambda_K(t)$ computed from the k-distribution of the ODE simulations, while the green dotted curve shows the upper bound with the approximation $\Lambda_K(t) = 1$. The blue dotted line shows the bound in (5.104), but with time scaled by $2^{3/4}$ (i.e. replacing the factor of $2^{1/4}$ with 2) showing that as N is increased, the results from the particle ODE simulations converge to a fixed factor below the derived bound. Note that the span of t' plotted for $N = 500$ is smaller as these simulations are much more computationally expensive to run.

Figure 7.24 shows plots of the combined quantity $G(t) := \mu_K^2(t)E(t)$ from the particle ODE simulations and kinetic PDE model against plots of the bound for G (5.104) (derived from the kinetic PDE model) for the inverse-cube interaction kernel across different values of N. Comparing the results from the bound (5.104) (yellow with circles) to those from the full kinetic PDE model (purple with squares), we see that the upper bound holds for the systems modelled. Comparing the results from the particle ODEs (dark blue with triangles) to the PDE model, we see that they diverge after a certain time. The t' value of this divergence

increases as N increases, consistent with the results from §7.2.1. After the particle ODE results diverge from the kinetic PDE, they go above the derived G bound, but the fact that the PDE model stays below the bound suggests this is a result of the assumptions used in deriving the PDE model rather than the further approximations used to derive (5.104) (notably, the independence of velocity and mass distributions). This is consistent with the results from the one-dimensional trials in Chapter 4, showing that assumptions such as the *Clusterzahlansatz* are expected to break down at later times when μ_K is much less than N and particles become more correlated, particularly for small N.

The upper bound shown by the yellow curve with circles uses $\Lambda_K(t)$ computed from the k-distribution produced by the particle ODE simulations, while the green dotted curve shows the same upper bound after making the approximation $\Lambda_K(t) = 1$. As expected, these curves are very similar for larger N, suggesting that taking $\Lambda_K(t) = 1$ is a reasonable approximation with N in this range, allowing the bound in (5.104) to be computed without using the ODE simulations. The light blue dotted curve shows the G bound with time scaled by a factor of $2^{3/4}$ (i.e. replacing the factor of $2^{1/4}$ with 2). Interestingly, as N is increased, the particle ODE and kinetic PDE results appear to converge to this curve, suggesting that the simulated results are a constant time factor away from the derived bound. This suggests that a more precise result may be available for the evolution of G for large N and is left for future investigation.

7.4. Summary of results

In this chapter, we have evaluated our kinetic PDE model against the full particle ODE simulations for the short-range Cucker-Smale model in two dimensions. The majority of the trials used an inverse-cube interaction kernel with particle density $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ and initial conditions Maxwellian in velocity, uniform in space, and monodisperse in mass. The kinetic model showed good agreement with the particle ODE simulations in the cluster mass and velocity distributions, mean kinetic energy per particle, and mean number of clusters. This agreement improved for larger N, with outputs from the two models converging, and the computation time for the kinetic PDE would be expected to scale better than that of the particle ODE simulations as N is increased further.

In §7.1, we optimised the implementation parameters of the fast Fourier spectral method for the kinetic PDE. We determined that a maximum Fourier frequency of $V = 30$ and radial and angular quadratures with $N_u = 20$ and $N_{\hat{\boldsymbol{u}}} = 24$ points, respectively, were sufficient such that the fast Fourier implementation contributed a small fraction of the total error. We also determined that $\Lambda_K(t)$ was best approximated using the empirical approximation (7.6) and that the DBSCAN clustering parameters $\varepsilon_{\boldsymbol{x}} = 10$ and $\varepsilon_{\boldsymbol{v}} = 0.1$ were most consistent with observed clustering behaviour for this interaction.

In §7.2, we evaluated the kinetic PDE model directly against the particle ODE simulations. The kinetic model showed good agreement with the particle simulations across a range of values of N, both in the mass and velocity distributions and in the mean number of clusters and mean kinetic energy per particle. Further, the mean kinetic energy and velocity distributions showed

improved agreement as N was increased. For $N = 160$, all quantities had proportional errors on the order of a few percent for the trials considered. However, the Fourier implementation began to break down at larger system times when the velocity distribution became too narrow to represent accurately and, thus, a modification such as a velocity rescaling would be necessary to maintain the accuracy of the PDE model for these times.

We then considered the effects of varying further parameters in the comparison. The results were similar for a wide range of density values, suggesting the density of $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ is in an appropriate range for the modelling assumptions of the predominance of binary interactions and negligible spatial extent of clusters to be valid. Considering an anisotropic initial velocity condition and mean-centred initial velocity condition, we observed minimal variation from the Maxwellian initial condition in terms of the error of the kinetic model and its convergence to the particle ODE results with increasing N. Trials with an inverse-quintic interaction kernel showed similar error levels to the inverse-cube kernel, but trials with an exponential kernel showed more discrepancy in the kinetic energy values. We believe this to be due to the fact that analytical expressions are available for the Fourier kernels for the power-law interaction kernels, but they must be approximated by interpolating a highly oscillating function for the exponential interaction kernel.

In §7.3, we analysed predictions and results from the kinetic model for the short-range Cucker-Smale interaction. We showed the presence of overpopulated exponential distribution tails, as is expected given the interaction is inelastic. We also evaluated the upper bound (5.104) for the combined quantity $G(t) := \mu_K^2(t)E(t)$, derived for the inverse-cube kernel, finding that it was correct but not tight. Empirically, it appeared a tighter bound might be available through a constant rescaling of time.

8. Discussion

The primary achievement of this thesis has been the derivation and evaluation of a kinetic model for particle systems with short-range interactions that exhibit clustering. This model is shown in the one-dimensional form in (3.75) (where it also applies to pointwise interactions), in the higher-dimensional form in terms of the number density in (5.32) and in its weak form in (5.50). This model shares features with both the inelastic Boltzmann equation (2.48), in terms of the structure of the binary interaction operators and the presence of the interaction map Jacobians to preserve probability flux, and the Smoluchowski coagulation equation (1.5), in terms of its treatment of mass as an internal dimension and the fact clusters can combine. We also have demonstrated the application of a systematic procedure for deriving kinetic equations from short-range second-order individual-based systems through the procedure of matched asymptotic expansions. One of the challenges in the derivation was the reducing dimension of the individual-based system of clusters as clustering interactions occurred. The use of the cluster number hierarchy and reduction to the coefficients μ_K and Λ_K using the *Clusterzahlansatz*, is a novel approach that we hope could be applied more generally. In particular, in deriving this model, we hope to provide a framework where further results from collisional elastic and inelastic kinetic theory and population balance modelling of clustering processes can be generalised to other interacting systems where the number of constituent objects modelled can vary.

From our final kinetic PDE model, we obtained simplified models for the evolution of the cluster number and mass distributions, with the latter in terms of a Smoluchowski-style equation with velocity-dependent coagulation kernels. These are not closed equations for general clustering interactions but suggest starting points for further development if an interaction is specified. We analysed the model specifically for a short-range version of the Cucker-Smale velocity averaging model of collective behaviour. This allowed the evolution equation for the mass density to be reduced to a Smoluchowski-style equation with time-dependent coagulation kernels. Taking the inverse-cube interaction kernel, we derived interdependent time-evolution upper bounds for the mean number of clusters, $\mu_K(t)$, and particle kinetic energy, $E(t)$. A closed bound for a combined quantity $G(t) = \mu_K^2(t)E(t)$ was also derived and shown to be accurate through numerical simulations of the kinetic PDE and particle ODEs. This bound allowed us to develop independent bounds for $\mu_K(t)$ and $E(t)$. We also showed the presence of overpopulated (exponential) velocity distribution tails for the short-range Cucker-Smale model, consistent with results obtained in inelastic kinetic theory. We believe this analysis to be readily extended to other short-range power-law interaction kernels, although general kernels, such as the exponential kernel, appear more challenging. We noted that the full kinetic and particle

models appeared to be converging to a constant time factor of $2^{3/4}$ below our upper bound for $G(t)$. This suggests that a sharper bound might be available with further analysis.

We implemented and evaluated our kinetic model against simulations of the full particle ODE system for the short-range Cucker-Smale model in one and two dimensions. In one dimension, we derived and implemented a discrete-velocity mass-binned numerical method for solving the kinetic PDE. This scheme allocates product particles to neighbouring grid points, following the fixed-pivot technique developed for population balance models, so as to conserve probability, mean velocity and mean mass. It includes an adaptive refinement of the velocity grid to maintain precision as the velocity distribution moves towards the Dirac delta due to the inelastic interactions. The particle ODEs were implemented with a bounded adaptive-timestep ODE solver, with clusters labelled in the output using the DBSCAN algorithm. Results from the kinetic PDE showed some discrepancies with those from the particle ODE model, with the mean mass and the mean particle energy being underestimated. The short-range kinetic model did, however, capture more of the correct behaviour than a model defined using the mean-field approximation, which does not track the clustering process. Two intermediate models were derived and implemented in order to differentiate the effects of the main approximations used in the kinetic PDE derivation. These were an 'exact clustering' model, which is similar to the full particle model but treats interactions as strictly pointwise, and a stochastic model that assumes particles are uncorrelated and distributed spatially uniformly but does not average over realisations or use the *Clusterzahlansatz* as the kinetic model does. Comparing the results from this suite of models between uniform and mean-centred uniform initial velocity distributions, we determined that the main sources of error were local particle correlations increasing over time and the fact that the kinetic model averages over a distribution of multiple realisations. For future analysis, we would like to evaluate these models using a Maxwellian initial velocity distribution and a lower spatial density of clusters to determine if these factors had any bearing on the discrepancy.

For the higher dimensional case, we extended a fast Fourier-Galerkin spectral method that has been developed for Boltzmann-style equations in [89, 90, 53, 67]. Our extension combines a binning of the mass dimension (similar to the one-dimensional PDE implementation) with the standard truncated Fourier series approximation in velocity space. We derived this method for general clustering interactions in d dimensions with interaction integrals that can be precomputed and stored or calculated at runtime. We implemented the model in two dimensions for the Cucker-Smale model with short-range inverse-cube, inverse-quintic, and exponential interaction kernels, and optimised it over possible input parameters such as the maximum Fourier frequency, V, and number of quadrature nodes used in the integral approximations. The numerical implementation of the kinetic PDE model showed much-improved convergence to the particle ODE simulations when compared with the one-dimensional results, suggesting that particle correlations are less significant in two dimensions and the *Stosszahlansatz* performs better. Using velocity, speed and mass distributions, as well as the mean number of clusters and particle energy as test metrics, the two models appeared to be converging to the same results as N was increased, which is as expected given the approximations made. We also

observed that the computation time of the kinetic PDE model scaled better than the particle ODE simulations for large N, although it is more expensive for small N, and thus it is in the many-particle regime that we would expect the kinetic model to be most useful.

One of the major drawbacks of the numerical implementation is the breakdown that occurs at larger times due to the inelastic collapse of the velocity distribution towards the Dirac delta. As the velocity distribution sharpens, the Fourier approximation error increases significantly. An important future step to improve the utility of the model would be to rescale the velocity such that the velocity distribution maintains a steady state with finite variance, or at least stays in a reasonable range for the Fourier approximation for longer. As discussed in §5.3.6, we do not expect a formal similarity scaling will be available for the Cucker-Smale model as the non-clustering interaction term is non-homogenous in velocity scaling. If mass is also scaled to maintain the homogeneity, it must be in the wrong direction for a similarity solution. However, we expect that a similar process might be useful in extending the applicability of the numerical scheme and providing further insight into the derived evolution equations for the averaged quantities $\mu_K(t)$, $E(t)$ and $G(t)$. If an appropriate scaling in the energy could be obtained, it could be possible to simplify the Smoluchowski-style equation for the mass density such that the coagulation kernels are independent of time in scaled variables. Further analysis of the interdependent moment evolution in velocity and mass space is an interesting avenue for future investigation.

Another limitation of the model is the empirical approximation, (7.6), of the parameter Λ_K from μ_K. Although we gave evidence in §7.1.1 and §7.3.2 that $\Lambda_K \to 1$ for large N and large μ_K, it would be interesting to evaluate the effect of tracking the cluster number distribution and computing Λ_K directly so as not to rely on this empirical approximation. We also hope that simplified estimates of the cluster number distribution may be available such that Λ_K may be approximated in a way that is generalisable to other kernels but does not rely on tracking the full distribution.

Although it didn't appear to be a major source of error for the test cases considered, an interesting extension would be to track the spatial extent of clusters as well as their masses. This could be implemented as an extra internal variable, perhaps with a direct dependence on mass, that altered the interaction force exerted by a cluster, and may allow the model to be better extended to regimes where the clusters are large compared with the size of the system. Further analysis of the limits of the point-cluster assumption for different interactions would be necessary to extend the applicability of the model.

In its essence, our derivation provides a framework for deriving kinetic models for short-range interacting systems with an extra internal variable and where the number of microscopic objects can change. As such, it could be extended to systems with different internal variables, for example, spin, charge, or surface area, that may evolve differently through interactions than the cluster mass does in our model. Further simplifications to the model to ensure computational feasibility may be desired if an extra, independent internal dimension is added. Finally, we suggest that it would be possible to include the process of cluster fragmentation in the model. This could amount to an extension of the point-cluster approximation to a

higher-order model of the multi-particle interaction process. As a starting point, fragmentation behaviour was observed for larger cluster diameters in the three-particle Cucker-Smale model (see Figure 3.2). This would require a generalisable model for the interaction parameters that would result in clustering or fragmentation and a model for the distribution of fragment sizes that would be produced, but the framework we have established for deriving a kinetic model with a variable number of clusters for the purely aggregative case could be extended.

Bibliography

[1] David J. Aldous. Deterministic and stochastic models for coalescence (aggregation and coagulation): A review of the mean-field theory for probabilists. *Bernoulli*, 5(1):3–48, 1999.

[2] Suguru Araki and Scott Tremaine. The dynamics of dense particle disks. *Icarus*, 65(1):83–109, 1986.

[3] Scott D. Baalrud and Jérôme Daligault. Mean force kinetic theory: A convergent kinetic theory for weakly and strongly coupled plasmas. *Physics of Plasmas*, 26(8):082106, 2019.

[4] Rafael Bailo, Mattia Bongini, José Carrillo, and Dante Kalise. Optimal consensus control of the Cucker-Smale model. *IFAC-PapersOnLine*, 51(13):1–6, 2018.

[5] Alethea B.T. Barbaro, Kirk Taylor, Peterson F. Trethewey, Lamia Youseff, and Björn Birnir. Discrete and continuous models of the dynamics of pelagic fish: Application to the capelin. *Mathematics and Computers in Simulation*, 79(12):3397–3414, 2009. The International Conference on Approximation Methods and numerical Modeling in Environment and Natural Resources.

[6] Leonid Berlyand, Robert Creese, Pierre-Emmanuel Jabin, and Mykhailo Potomkin. Continuum approximations to systems of correlated interacting particles. *Journal of Statistical Physics*, 174, 02 2019.

[7] Leonid Berlyand, Pierre-Emmanuel Jabin, and Mykhailo Potomkin. Complexity reduction in many particle systems with random initial data. *SIAM/ASA Journal on Uncertainty Quantification*, 4, 09 2013.

[8] Jeff Bezanson, Alan Edelman, Stefan Karpinski, and Viral B Shah. Julia: A fresh approach to numerical computing. *SIAM Review*, 59(1):65–98, 2017.

[9] James Binney and Scott Tremaine. *Galactic Dynamics: Second Edition*. Princeton University Press, 01 2008.

[10] Bjorn Birnir. An ODE model of the motion of pelagic fish. *Journal of Statistical Physics*, 128:535–568, 06 2007.

[11] Marzia Bisi, José A. Carrillo, and Toscani Giuseppe. Decay rates in probability metrics towards homogeneous cooling states for the inelastic Maxwell model. *Journal of Statistical Physics*, 124:625–653, 2006.

[12] A. Blanchet and P. Degond. Kinetic models for topological nearest-neighbor interactions. *Journal of Statistical Physics*, 169:929–950, 2017.

[13] Alexander Bobylev, José A. Carrillo, and Irene Gamba. On some properties of kinetic and hydrodynamic equations for inelastic interactions. *Journal of Statistical Physics*, 98(3), 05 2000.

[14] N. N. Bogoliubov. Problems of a dynamical theory in statistical physics. In *Studies in Statistical Mechanics, Volume 1*. Wiley, 1962.

[15] François Bolley, José A. Cañizo, and José A. Carrillo. Stochastic mean-field limit: non-Lipschitz forces and swarming. *Mathematical Models and Methods in Applied Sciences*, 21(11):2179–2210, 2011.

[16] Ludwig Boltzmann. *Lectures on gas theory*. University of California Press, 1964. Translated by Stephen G. Brush from the 1896-1898 Edition. Reprinted by Dover Publications, 1995.

[17] M. Born and H. S. Green. *A General Kinetic Theory of Liquids*. Cambridge University Press, 1949.

[18] Marc Briant, Antoine Diez, and Sara Merino-Aceituno. Cauchy theory for general kinetic Vicsek models in collective dynamics and mean-field limit approximations. *SIAM Journal on Mathematical Analysis*, 54(1):1131–1168, 2022.

[19] Maria Bruna and S. Jonathan Chapman. Excluded-volume effects in the diffusion of hard spheres. *Physical review. E*, 85:011103, 01 2012.

[20] Maria Bruna, S. Jonathan Chapman, and Martin Robinson. Diffusion of particles with short-range interactions. *SIAM Journal on Applied Mathematics*, 77, 02 2017.

[21] Matthäus U. Bäbler and Massimo Morbidelli. Analysis of the aggregation–fragmentation population balance equation with application to coagulation. *Journal of Colloid and Interface Science*, 316(2):428 – 441, 2007.

[22] Juan Calvo, Juanjo Nieto, and Mohamed Zagour. Kinetic model for vehicular traffic with continuum velocity and mean field interactions. *Symmetry*, 11:1093, 08 2019.

[23] Torsten Carleman. *Problèmes Mathématiques dans la Théorie Cinétique des Gaz*. Almqvist and Wiksells, 1957.

[24] Eric A. Carlen, José A. Carrillo, and Maria C. Carvalho. Strong convergence towards homogeneous cooling states for dissipative Maxwell models. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 26(5):1675–1700, 2009.

[25] J. A. Carrillo, M. Fornasier, J. Rosado, and G. Toscani. Asymptotic flocking dynamics for the kinetic Cucker-Smale model. *SIAM Journal on Mathematical Analysis*, 42(1):218–236, 2010.

[26] José A. Carrillo and Young-Pil Choi. Mean-field limits: from particle descriptions to macroscopic equations. *Archive for Rational Mechanics and Analysis*, 241:1529–1573, 2021.

[27] José A. Carrillo, Young-Pil Choi, and Sergio P. Perez. A review on attractive–repulsive hydrodynamics for consensus in collective behavior. In Nicola Bellomo, Pierre Degond, and Eitan Tadmor, editors, *Active Particles, Volume 1: Advances in Theory, Models, and Applications*, pages 259–298. Springer International Publishing, Birkhäuser, Cham, 2017.

[28] José A. Carrillo, Massimo Fornasier, Giuseppe Toscani, and Francesco Vecil. Particle, kinetic, and hydrodynamic models of swarming. In Giovanni Naldi, Lorenzo Pareschi, and Giuseppe Toscani, editors, *Mathematical Modeling of Collective Behavior in Socio-Economic and Life Sciences. Modeling and Simulation in Science, Engineering and Technology.*, pages 297–336. Birkhäuser Boston, Boston, 06 2010.

[29] José A. Carrillo, Jingwei Hu, Zheng Ma, and Thomas Rey. Recent development in kinetic theory of granular materials: Analysis and numerical methods. In Giacomo Albi, Sara Merino-Aceituno, Alessia Nota, and Mattia Zanella, editors, *Trails in Kinetic Theory: Foundational Aspects and Numerical Methods*, pages 1–36. Springer International Publishing, 2021.

[30] Carlo Cercignani. *The Boltzmann Equation and Its Applications*. Springer New York, NY, 1988.

[31] Carlo Cercignani. *Mathematical Methods in Kinetic Theory*. Springer New York, NY, 1990.

[32] Young-Pil Choi, Seung-Yeal Ha, and Zhuchun Li. Emergent dynamics of the cucker-smale flocking model and its variants. In Nicola Bellomo, Pierre Degond, and Eitan Tadmor, editors, *Active Particles, Volume 1: Advances in Theory, Models, and Applications*, pages 299–331. Springer International Publishing, Birkhäuser, Cham, 2017.

[33] Y. Chuang, Y. R. Huang, M. R. D'Orsogna, and A. L. Bertozzi. Multi-vehicle flocking: Scalability of cooperative control algorithms using pairwise potentials. In *Proceedings 2007 IEEE International Conference on Robotics and Automation*, pages 2292–2299, 2007.

[34] Felipe Cucker and Steve Smale. Emergent behavior in flocks. *Automatic Control, IEEE Transactions on Automatic Control*, 52:852–862, 06 2007.

[35] Felipe Cucker and Steve Smale. On the mathematics of emergence. *Japanese Journal of Mathematics*, 2:197–227, 03 2007.

[36] George Datseris, Jonas Isensee, Sebastian Pech, and Tamás Gál. DrWatson: the perfect sidekick for your scientific inquiries. *Journal of Open Source Software*, 5(54):2673, 2020.

[37] Pierre Degond, Amic Frouvelle, and Sara Merino-Aceituno. A new flocking model through body attitude coordination. *Mathematical Models and Methods in Applied Sciences*, 27(06):1005–1049, 2017.

[38] Pierre Degond and Sébastien Motsch. Continum limit of self-driven particles with orientation interaction. *Mathematical Models and Methods in Applied Sciences*, 18(01):1193–1215, 2008.

[39] L. Desvillettes and M. Pulvirenti. The linear Boltzmann equation for long-range forces: A derivation from particle systems. *Mathematical Models and Methods in Applied Sciences*, 09(08):1123–1145, 1999.

[40] M. R. D'Orsogna, Y. L. Chuang, A. L. Bertozzi, and L. S. Chayes. Self-propelled particles with soft-core interactions: Patterns, stability, and collapse. *Physical Review Letters*, 96:104302, 03 2006.

[41] Yakovenko V. Dragulescu, A. Statistical mechanics of money. *European Physical Journal B*, 17:723–729, 2000.

[42] Radek Erban, Jan Haskovec, and Yongzheng Sun. A Cucker-Smale model with noise and delay. *SIAM Journal on Applied Mathematics*, 76, 07 2015.

[43] M. H. Ernst and R. Brito. High-energy tails for inelastic Maxwell models. *Europhysics Letters*, 58(2):182, 04 2002.

[44] B. U. Felderhof. Diffusion of interacting Brownian particles. *Journal of Physics A: Mathematical and General*, 11(5):929–937, 05 1978.

[45] Alessio Figalli and Moon-Jin Kang. A rigorous derivation from the kinetic Cucker-Smale model to the pressureless Euler system with nonlocal alignment. *Analysis and Partial Differential Equations*, 12, 02 2017.

[46] Francis Filbet, Lorenzo Pareschi, and Giuseppe Toscani. Accurate numerical methods for the collisional motion of (heated) granular flows. *Journal of Computational Physics*, 202(1):216–235, 2005.

[47] Francis Filbet and Thomas Rey. A rescaling velocity method for dissipative kinetic equations. applications to granular media. *Journal of Computational Physics*, 248:177–199, 2013.

[48] Francis Filbet and Giovanni Russo. Accurate numerical methods for the Boltzmann equation. In L. Pareschi P. Degond and G. Russo, editors, *Modeling and Computational Methods for Kinetic Equations*, chapter 4. 2004.

[49] Nicolas Fournier and Benjamin Jourdain. Stochastic particle approximation of the keller-segel equation and two-dimensional generalization of Bessel processes. *Annals of Applied Probability*, 27(5):2807–2861, 10 2017.

[50] Benjamin Franz, Jake P. Taylor-King, Christian Yates, and Radek Erban. Hard-sphere interactions in velocity-jump models. *Physical Review E*, 94:012129, 07 2016.

[51] Serge Galam, Yuval Gefen, and Yonathan Shapir. Sociophysics: A new approach of sociological collective behaviour. I. mean-behaviour description of a strike. *Journal of Mathematical Sociology*, 9:1–13, 1982.

[52] Isabelle Gallagher, Laure Saint-Raymond, and Benjamin Texier. *From Newton to Boltzmann : hard spheres and short-range potentials*. Zürich lectures in advanced mathematics. European Mathematical Society, 2013.

[53] Irene M. Gamba, Jeffrey R. Haack, Cory D. Hauck, and Jingwei Hu. A fast spectral method for the Boltzmann collision operator with general collision kernels. *SIAM Journal on Scientific Computing*, 39(4):B658–B674, 2017.

[54] Irene M. Gamba and Sri Harsha Tharkabhushanam. Spectral-Lagrangian methods for collisional models of non-equilibrium statistical states. *Journal of Computational Physics*, 228(6):2012–2036, 2009.

[55] Harold Grad. Principles of the kinetic theory of gases. In *Thermodynamik der Gase / Thermodynamics of Gases*, pages 205–294. Springer, Berlin, Heidelberg, 1958.

[56] Seung-Yeal Ha, Jeongho Kim, Jinyeong Park, and Xiongtao Zhang. Uniform stability and mean-field limit for the augmented Kuramoto model. *Networks and Heterogeneous Media*, 13(2):297–322.

[57] Seung Yeal Ha and Jian-guo Liu. A simple proof of the Cucker-Smale flocking dynamics and mean-field limit. *Communications in Mathematical Sciences*, 7:297–325, 06 2009.

[58] Seung Yeal Ha and Eitan Tadmor. From particle to kinetic and hydrodynamic descriptions of flocking. *Kinetic and Related Models*, 1(3):415–435, 07 2008.

[59] P. K. Haff. Grain flow as a fluid-mechanical phenomenon. *Journal of Fluid Mechanics*, 134:401–430, 09 1983.

[60] Alan Hammond. Coagulation and diffusion: A probabilistic perspective on the Smoluchowski PDE. *Probability Surveys*, 14:205 – 288, 2017.

[61] S. Harris. *An Introduction to the Theory of the Boltzmann Equation*. Dover books on physics. Dover Publications, 2004.

[62] Siming He and Eitan Tadmor. A game of alignment: Collective behavior of multi-species. *Annales de l'Institut Henri Poincaré C, Analyse non linéaire*, 2020.

[63] Alex R. Heath, Parisa A. Bahri, Phillip D. Fawell, and John B. Farrow. Polymer flocculation of calcite: Population balance model. *AIChE Journal*, 52(5):1641–1653, 2006.

[64] Rainer Hegselmann and Ulrich Krause. Opinion dynamics and bounded confidence models, analysis and simulation. *Journal of Artificial Societies and Social Simulation*, 5, 07 2002.

[65] Charlotte K. Hemelrijk and Hanno Hildenbrandt. Some causes of the variable shape of flocks of birds. *PLOS ONE*, 6(8):1–13, 08 2011.

[66] M. J. Hounslow, R. L. Ryall, and V. R. Marshall. A discretized population balance for nucleation, growth, and aggregation. *AIChE Journal*, 34(11):1821–1832, 1988.

[67] Jingwei Hu and Zheng Ma. A fast spectral method for the inelastic Boltzmann collision operator and application to heated granular gases. *Journal of Computational Physics*, 385:119–134, 2019.

[68] Pierre-Emmanuel Jabin. A review of the mean field limits for Vlasov equations. *Kinetic and Related Models*, 7:661–711, 11 2014.

[69] Pierre-Emmanuel Jabin and Zhenfu Wang. Mean field limit for stochastic particle systems. In *Modeling and Simulation in Science, Engineering and Technology*, volume 1, pages 379–402. Springer Basel, 04 2017.

[70] A. Jadbabaie, Jie Lin, and A.S. Morse. Coordination of groups of mobile autonomous agents using nearest neighbor rules. *IEEE Transactions on Automatic Control*, 48(6):988–1001, 2003.

[71] Chunyin Jin. Flocking of the Motsch-Tadmor model with a cut-off interaction function. *Journal of Statistical Physics*, 171:345–360, 04 2018.

[72] Laura Kanzler, Christian Schmeiser, and Veronica Tora. Two kinetic models for non-instantaneous binary alignment collisions. *Kinetic and Related Models*, 2023.

[73] Toshio Kawai and Koichiro Shida. An inelastic collision model for the evolution of "planetary rings". *Journal of the Physical Society of Japan*, 59(1):381–388, 1990.

[74] John G. Kirkwood. Statistical mechanics of fluid mixtures. *The Journal of Chemical Physics*, 3(5):300–313, 1935.

[75] John G. Kirkwood. The statistical mechanical theory of transport processes II. Transport in gases. *The Journal of Chemical Physics*, 15(1):72–76, 01 1947.

[76] Ulrich Krause. A discrete nonlinear and non-autonomous model of consensus formation. In *Communications in Difference Equations: Proceedings of the Fourth International Conference on Difference Equations*, pages 227–236, 01 2000.

[77] Sanjeev Kumar and D. Ramkrishna. On the solution of population balance equations by discretization—I. A fixed pivot technique. *Chemical Engineering Science*, 51(8):1311 – 1332, 1996.

[78] James Clerk Maxwell. IV. on the dynamical theory of gases. *Philosophical Transactions of the Royal Society of London*, 157:49–88, 1867.

[79] Sean McNamara and W. R. Young. Kinetics of a one-dimensional granular medium in the quasielastic limit. *Physics of Fluids A: Fluid Dynamics*, 5(1):34–45, 01 1993.

[80] Mouhot C. Mischler, S. and M. Rodriguez Ricard. Cooling process for inelastic Boltzmann equations for hard spheres, Part I: The Cauchy problem. *Journal of Statistical Physics*, 124:655–702, 2006.

[81] S. Mischler and C. Mouhot. Cooling process for inelastic Boltzmann equations for hard spheres. Part II. Self-similar solutions and tail behavior. *Journal of Statistical Physics*, 124(2):703–746, 2006.

[82] Javier Morales, Jan Peszek, and Eitan Tadmor. Flocking with short-range interactions. *Journal of Statistical Physics*, 176:1–16, 07 2019.

[83] S. Motsch and E. Tadmor. A new model for self-organized dynamics and its flocking behavior. *Journal of Statistical Physics*, 144(923), 2011.

[84] Sebastien Motsch and Eitan Tadmor. Heterophilious dynamics enhances consensus. *SIAM Review*, 56(4):577–621, 2014.

[85] Mehdi Moussaïd, Dirk Helbing, and Guy Theraulaz. How simple rules determine pedestrian behavior and crowd disasters. *Proceedings of the National Academy of Sciences of the United States of America*, 108(17):6884–6888, 04 2011.

[86] Piotr B. Mucha and Jan Peszek. The Cucker-Smale equation: Singular communication weight, measure-valued solutions and weak-atomic uniqueness. *Archive for Rational Mechanics and Analysis*, 227:273–308, 2018.

[87] Giovanni Naldi, Lorenzo Pareschi, and Giuseppe Toscani. Spectral methods for one-dimensional kinetic models of granular flows and numerical quasi elastic limit. *ESAIM: Modélisation mathématique et analyse numérique*, 37(1):73–90, 2003.

[88] Derek Paley, Naomi Leonard, Rodolphe Sepulchre, Daniel Grunbaum, and Julia Parrish. Oscillator models and collective motion. *Control Systems, IEEE*, 27:89–105, 09 2007.

[89] Lorenzo Pareschi and Benoit Perthame. A Fourier spectral method for homogeneous Boltzmann equations. *Transport Theory and Statistical Physics*, 25:369–382, 04 1996.

[90] Lorenzo Pareschi and Giovanni Russo. Numerical solution of the Boltzmann equation I: Spectrally accurate approximation of the collision operator. *SIAM Journal on Numerical Analysis*, 37(4):1217–1245, 2000.

[91] Lorenzo Pareschi and Giovanni Russo. An introduction to the numerical analysis of the Boltzmann equation. *Rivista di Matematica della Universita di Parma (7)*, 4:145–250, 01 2005.

[92] Laura Perea and Pedro Elosegui. Extension of the Cucker-Smale control law to space flight formations. *Journal of Guidance Control and Dynamics*, 32:527–537, 03 2009.

[93] Christopher Rackauckas and Qing Nie. DifferentialEquations.jl–a performant and feature-rich ecosystem for solving differential equations in Julia. *Journal of Open Research Software*, 5(1), 2017.

[94] Tertius Ralph, Stephen W. Taylor, and Maria Bruna. One-dimensional model for chemotaxis with hard-core interactions. *Physical Review E*, 101:022419, 02 2020.

[95] Erin C. Rericha, Chris Bizon, Mark D. Shattuck, and Harry L. Swinney. Shocks in supersonic sand. *Physical Review Letters*, 88:014302, 12 2001.

[96] Thomas Rey. Blow up analysis for anomalous granular gases. *SIAM Journal on Mathematical Analysis*, 44(3):1544–1561, 04 2012.

[97] P. G. Saffman and J. S. Turner. On the collision of drops in turbulent clouds. *Journal of Fluid Mechanics*, 1(1):16–30, 1956.

[98] Ruiwen Shu and Eitan Tadmor. Flocking hydrodynamics with external potentials. *Archive for Rational Mechanics and Analysis*, 238, 10 2020.

[99] Ruiwen Shu and Eitan Tadmor. Anticipation breeds alignment. *Archive for Rational Mechanics and Analysis*, 240:203–241, 04 2021.

[100] Roman Shvydkoy and Eitan Tadmor. Topologically based fractional diffusion and emergent dynamics with short-range interactions. *SIAM Journal on Mathematical Analysis*, 52(6):5792–5839, 2020.

[101] Roman Shvydkoy and Eitan Tadmor. Multiflocks: Emergent dynamics in systems with multiscale collective behavior. *Multiscale Modeling & Simulation*, 19(2):1115–1141, 2021.

[102] Naftali R. Smith, Nir J. Shaviv, and Henrik Svensmark. Approximate analytical solutions to the condensation-coagulation equation of aerosols. *Aerosol Science and Technology*, 50(6):578–590, 2016.

[103] Marian Smoluchowski. Drei vorträge über Diffusion, Brownsche Molekularbewegung und Koagulation von Kolloidteilchen. *Pisma Mariana Smoluchowskiego*, 2(1):530–594, 1916.

[104] Herbert Spohn. The Vlasov equation. In *Large Scale Dynamics of Interacting Particles*, pages 77–82. Springer, Berlin, Heidelberg, 1991.

[105] Elsen Tjhung, Cesare Nardini, and Michael E. Cates. Cluster phases and bubbly phase separation in active fluids: Reversal of the Ostwald process. *Physical Review X*, 8:031080, 09 2018.

[106] Tamás Vicsek, András Czirók, Eshel Ben-Jacob, Inon Cohen, and Ofer Shochet. Novel type of phase transition in a system of self-driven particles. *Physical Review Letters*, 75:1226–1229, 08 1995.

[107] Tamás Vicsek and Anna Zafeiris. Collective motion. *Physics Reports*, 517(3):71–140, 2012.

[108] Cédric Villani. A review of mathematical topics in collisional kinetic theory. volume 1 of. Handbook of Mathematical Fluid Dynamics, Chapter 2, pages 71–187. North-Holland, 12 2002.

[109] Cédric Villani. Mathematics of granular materials. *Journal of Statistical Physics*, 124:781–822, 2006.

[110] Jonathan Wells. *Modelling coagulation in industrial spray drying: An efficient one-dimensional population balance approach.* PhD thesis, University of Strathclyde, 2017.

[111] Wenjian Yu and A. Johansson. Modeling crowd turbulence by many-particle simulations. *Physical Review E*, 76(4):046105, 2007.

[112] Jacques Yvon and Yves Rocard. *La Théorie Statistique des Fluides et l'Equation d'Etat.* Hermann, 1935.

A. Supplementary derivations

This chapter presents more detailed steps for some of the integral term manipulations from Chapter 3 and the derivation of the discrete-velocity, mass-binned numerical scheme from Chapter 4.

A.1. Clustering flux condtition

Given $v_i = \lambda_{m_j,m_s}(v_j, v_s)$, we intend to show that

$$\int_{\lambda^{-1}_{m_j,m_s}(U)} g(v_j, v_s)\,\mathrm{d}v_j\mathrm{d}v_s = \int_U \int_{\Lambda_{m_j,m_s}(v_i)} g\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i), v_s\right) \frac{\mathrm{d}}{\mathrm{d}v_i}\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i)\right)\,\mathrm{d}v_s\mathrm{d}v_i \tag{A.1}$$

where, by an abuse of notation, $\lambda^{-1}_{m_j,m_s}(U)$ is the range of velocity pairs (v_j, v_s) that would map to $v_i \in U$ under λ_{m_j,m_s}. The region $\Lambda_{m_j,m_s}(v_i)$ is the range of possible incoming cluster velocities, v_s, that could produce a cluster with velocity v_i and $g(v_j, v_s)$ is a generic function.

To define $\lambda^{-1}_{m_j,m_s}(U)$ specifically, first define

$$\Lambda_{m_j,m_s}(U) = \bigcup_{v_i \in U} \Lambda_{m_j,m_s}(v_i)$$

to be the set of incoming velocities v_s that could produce a cluster in U, given incoming masses m_j and m_s. Then define $U^*(v_s) \subseteq U$ to be the set of all points in U that can be reached from v_s under the action of λ_{m_j,m_s} with any possible v_j. Given a specific v_s, m_s and m_j it then follows that the set of possible values for v_j that would produce an outgoing velocity v_i in U is $\hat{\lambda}^{-1}_{m_j,m_s,v_s}(U^*(v_s))$. As such, we can express the LHS of (A.1) as

$$\int_{\Lambda_{m_j,m_s}(U)} \int_{\hat{\lambda}^{-1}_{m_j,m_s,v_s}(U^*(v_s))} g(v_j, v_s)\,\mathrm{d}v_j\mathrm{d}v_s$$

Given v_i and given $v_s \in \Lambda_{m_j,m_s}(v_i)$, v_j is uniquely defined as $v_j = \hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i)$, where $\hat{\lambda}^{-1}$ is the partial inverse of λ defined in (3.24). Changing the variables of the inner integral from v_j to v_i then gives

$$\int_{\Lambda_{m_j,m_s}(U)} \int_{U^*(v_s)} g\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i), v_s\right) \frac{\mathrm{d}}{\mathrm{d}v_i}\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i)\right)\,\mathrm{d}v_i\mathrm{d}v_s.$$

Switching the order of the integrals, we have to modify the bounds to define v_i first (which ranges over U) and then represent the range of v_s given this v_i, which is defined to be $\Lambda_{m_j,m_s}(v_i)$.

As such, the integral becomes

$$\int_U \int_{\Lambda_{m_j,m_s}(v_i)} g\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i), v_s\right) \frac{\mathrm{d}}{\mathrm{d}v_i}\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i)\right) \mathrm{d}v_s \mathrm{d}v_i,$$

as in the RHS of (A.1).

Setting U to the $\mathbb{R}$ as a special case, we can see that $\lambda^{-1}_{m_j,m_s}(\mathbb{R})$ is the full set of possible incoming velocities that could be involved in a clustering interaction, which was defined in §3.3 as C_{m_j,m_s}. As such, we have

$$\int_{C_{m_j,m_s}} g(v_j, v_s)\,\mathrm{d}v_j \mathrm{d}v_s = \int_{\mathbb{R}} \int_{\Lambda_{m_j,m_s}(v_i)} g\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i), v_s\right) \frac{\mathrm{d}}{\mathrm{d}v_i}\left(\hat{\lambda}^{-1}_{m_j,m_s,v_s}(v_i)\right) \mathrm{d}v_s \mathrm{d}v_i \quad \text{(A.2)}$$

A.2. Marginal distributions

Equation (3.37) is

$$\begin{aligned}
&\partial_t p_1^k + v_1 \partial_{x_1} p_1^k \\
&= -(k-1)\int_{\boldsymbol{D}} |v_1 - v_2|^{-} p_2^k(t; z_1, z_2)\delta_{21}\,\mathrm{d}z_2 && \text{(A.3a)}\\
&\quad - \frac{(k-1)(k-2)}{2}\int_{\boldsymbol{D}^2} |v_2 - v_3|^{-} p_3^k(t; z_1, z_2, z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3 && \text{(A.3b)}\\
&\quad + (k-1)\int_{\boldsymbol{D}} |{}^*v_1 - {}^*v_2|^{-} p_2^k(t; {}^*z_1, {}^*z_2)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_2 && \text{(A.3c)}\\
&\quad + \frac{(k-1)(k-2)}{2}\int_{\boldsymbol{D}^2} |{}^*v_2 - {}^*v_3|^{-} p_3^k(t; z_1, {}^*z_2, {}^*z_3)\delta_{23}\left|D\psi_{23}^{-1}\right|\,\mathrm{d}z_2\mathrm{d}z_3 && \text{(A.3d)}\\
&\quad + \frac{k+1}{2}\int_{\boldsymbol{C}'(z_1)} |'v_1 - v_{k+1}|^{-} p_2^{k+1}(t; 'z_1, z_{k+1})\delta_{1(k+1)}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_{k+1} && \text{(A.3e)}\\
&\quad + \frac{(k+1)(k-1)}{2}\int_{\boldsymbol{D}}\int_{\boldsymbol{C}'(z_2)} |'v_2 - v_{k+1}|^{-} p_3^{k+1}(t; z_1, 'z_2, z_{k+1})\delta_{2(k+1)}\frac{\mathrm{d}'v_2}{\mathrm{d}v_2}\,\mathrm{d}z_{k+1}\mathrm{d}z_2. && \text{(A.3f)}
\end{aligned}$$

Changing the variables of the $\mathrm{d}v_2\mathrm{d}v_3$ integral from (A.3d) to $\mathrm{d}^*v_2\mathrm{d}^*v_3$ (i.e. changing from $\mathrm{d}z_2\mathrm{d}z_3$ to $\mathrm{d}^*z_2\mathrm{d}^*z_3$) using the map $({}^*v_2, {}^*v_3) = \psi^{-1}_{m_2,m_3}(v_2, v_3)$ absorbs the Jacobian. The resulting velocity integration region is symbolically $\psi^{-1}_{m_2,m_3}(\mathbb{R}^2)$: all the possible input combinations of velocities ${}^*v_2$ and ${}^*v_3$ that would result in a non-clustering interaction. This can equivalently be written as $\mathbb{R}^2 \backslash C_{m_2,m_3}$, such that we can write (A.3d) out fully as

$$\begin{aligned}
\frac{(k-1)(k-2)}{2}\sum_{m_2,m_3=1}^{N}\int_{\mathbb{R}^2\backslash C_{m_2,m_3}}\int_{\Omega^2}\big[&|{}^*v_2 - {}^*v_3| \\
&{}^{-}p_3^k(t; x_1, x_2, x_2; v_1, {}^*v_2, {}^*v_3; m_1, m_2, m_3)\delta_{23}\big]\mathrm{d}x_2\mathrm{d}x_3\,\mathrm{d}^*v_2\mathrm{d}^*v_3.
\end{aligned}$$

Changing the dummy integration variables back to v_2 and v_3, this can then be combined with (A.3b), becoming

$$-\frac{(k-1)(k-2)}{2}\sum_{m_2,m_3=1}^{N}\int_{C_{m_2,m_3}}\int_{\Omega^2}|v_2-v_3| \\ {}^{-}p_3^k(t;x_1,x_2,x_3;v_1,v_2,v_3;m_1,m_2,m_3)\delta_{23}\,\mathrm{d}x_2\mathrm{d}x_3\,\mathrm{d}v_2\mathrm{d}v_3.$$

Define a new integration region $\boldsymbol{C}^2$ to be analogous to the full two-particle domain $\boldsymbol{D}^2$, except that now the velocity integrals are only taken over the clustering region, C_{m_i,m_j}, rather than over the full $\mathbb{R}^2$. That is,

$$\int_{\boldsymbol{C}^2}\mathrm{d}z_i\,\mathrm{d}z_j := \sum_{m_i,m_j=1}^{N}\int_{C_{m_i,m_j}}\int_{\Omega^2}\mathrm{d}x_i\mathrm{d}x_j\,\mathrm{d}v_i\mathrm{d}v_j. \tag{A.4}$$

Using this notation, the combination of (A.3b) and (A.3d) can be written as

$$-\frac{(k-1)(k-2)}{2}\int_{\boldsymbol{C}^2}|v_2-v_3|\,{}^{-}p_3^k(t;z_1,z_2,z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3. \tag{A.5}$$

The term (A.3f) has velocity integrals of the form

$$\int_{\mathbb{R}}\int_{\Lambda'_{m_2,m_{k+1}}(v_2)}g(\prime v_2,v_{k+1})\frac{\mathrm{d}\prime v_2}{\mathrm{d}v_2}\,\mathrm{d}v_{k+1}\,\mathrm{d}v_2,$$

where $\prime v_2=\hat{\lambda}^{-1}_{m_2,m_{k+1},v_{k+1}}(v_2)$. Using (A.2) this becomes

$$\int_{C'_{m_2,m_{k+1}}}g(\prime v_2,v_{k+1})\,\mathrm{d}\prime v_2\mathrm{d}v_{k+1}$$

The mass sums in (A.3f) take the form

$$\sum_{m_2=1}^{N}\sum_{m_{k+1}=1}^{m_2-1}h(\prime m_2,m_{k+1}),$$

where h represents the rest of the term and $\prime m_2=m_2-m_{k+1}$. This can be modified to

$$\sum_{m_{k+1}=1}^{N-1}\sum_{m_2=m_{k+1}+1}^{N}h(\prime m_2,m_{k+1})=\sum_{m_{k+1}=1}^{N-1}\sum_{\prime m_2=1}^{N-1}h(\prime m_2,m_{k+1})=\sum_{\prime m_2,m_{k+1}=1}^{N}h(\prime m_2,m_{k+1}),$$

where, in the last step, we have extended the sum limits to N as any ${}^{-}p_3^{k+1}$ term with a mass argument of N will be 0. With these two transformations, (A.3f) becomes

$$\frac{(k+1)(k-1)}{2}\sum_{m_2',m_{k+1}=1}^{N}\int_{C'_{m_2,m_{k+1}}}\int_{\Omega^2}\Big[|\prime v_2-v_{k+1}| \\ {}^{-}p_3^{k+1}(x_1,x_2,x_3;v_1,\prime v_2,v_{k+1};m_1,\prime m_2,m_{k+1})\delta_{23}\Big]\,\mathrm{d}x_2\mathrm{d}x_3\,\mathrm{d}\prime v_2\mathrm{d}v_{k+1},$$

which we can modify by changing dummy variables to

$$\frac{(k+1)(k-1)}{2}\int_{\boldsymbol{C}^2}|v_2-v_3|^{-}p_3^{k+1}(t;z_1,z_2,z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.$$

Finally, we can change dummy variables in the (A.3e) from $'z_{k+1}$ to $'z_2$, to give

$$\frac{k+1}{2}\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|^{-}p_2^{k+1}(t;'z_1,z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2.$$

With these transformations, the full p_1^k time evolution equation is (note the terms have been reordered to put the p_3 terms last)

$$\begin{aligned}\partial_t p_1^k+v_1\partial_{x_1}p_1^k=&-(k-1)\int_{\boldsymbol{D}}|v_1-v_2|^{-}p_2^k(t;z_1,z_2)\delta_{21}\,\mathrm{d}z_2\\&+(k-1)\int_{\boldsymbol{D}}|{}^*v_1-{}^*v_2|^{-}p_2^k(t;{}^*z_1,{}^*z_2)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_2\\&+\frac{k+1}{2}\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|^{-}p_2^{k+1}(t;'z_1,z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\\&-\frac{(k-1)(k-2)}{2}\int_{\boldsymbol{C}^2}|v_2-v_3|^{-}p_3^k(t;z_1,z_2,z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\\&+\frac{(k+1)(k-1)}{2}\int_{\boldsymbol{C}^2}|v_2-v_3|^{-}p_3^{k+1}(t;z_1,z_2,z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.\end{aligned}\tag{A.6}$$

A.3. Sum over k-hierarchy

Summing (3.53) over k, we get

$$\begin{aligned}&\partial_t p_1+v_1\partial_{x_1}p_1\\&\quad=-\sum_{k=2}^{N}\frac{(k-1)}{P_k}\int_{\boldsymbol{D}_\Gamma(k,m_1)}|v_1-v_2|\,p_1^k(z_1)p_1^k(z_2)\delta_{12}\,\mathrm{d}z_2\\&\qquad+\sum_{k=2}^{N}\frac{(k-1)}{P_k}\int_{\boldsymbol{D}_\Gamma(k,m_1)}|{}^*v_1-{}^*v_2|\,p_1^k({}^*z_1)p_1^k({}^*z_2)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_2\\&\qquad+\sum_{k=1}^{N-1}\frac{k+1}{2P_{k+1}}\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|\,p_1^{k+1}('z_1)p_1^{k+1}(z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\\&\qquad-\sum_{k=3}^{N}\frac{(k-1)(k-2)}{2(P_k)^2}p_1^k(z_1)\int_{\boldsymbol{C}^2_\Gamma(k,m_1)}|v_2-v_3|\,p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\\&\qquad+\sum_{k=2}^{N-1}\frac{(k+1)(k-1)}{2(P_{k+1})^2}p_1^{k+1}(z_1)\int_{\boldsymbol{C}^2_\Gamma(k+1,m_1)}|v_2-v_3|\,p_1^{k+1}(z_2)p_1^{k+1}(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.\end{aligned}$$

Note that the sum bounds have been reduced as not all RHS terms are present for all values of k. Changing the dummy variable in the third and fifth terms to $k'=k+1$ (and then dropping

the dashes), we get

$$\begin{aligned}\partial_t p_1 + v_1\partial_{x_1} p_1 = &-\sum_{k=2}^{N}\frac{(k-1)}{P_k}\int_{\boldsymbol{D}_\Gamma(k,m_1)}|v_1-v_2|\,p_1^k(z_1)p_1^k(z_2)\delta_{12}\,\mathrm{d}z_2\\ &+\sum_{k=2}^{N}\frac{(k-1)}{P_k}\int_{\boldsymbol{D}_\Gamma(k,m_1)}|{}^*v_1-{}^*v_2|\,p_1^k({}^*z_1)p_1^k({}^*z_2)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_2\\ &+\sum_{k=2}^{N}\frac{k}{2P_k}\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|\,p_1^k('z_1)p_1^k(z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\\ &-\sum_{k=3}^{N}\frac{(k-1)(k-2)}{2(P_k)^2}p_1^k(z_1)\int_{\boldsymbol{C}^2_\Gamma(k,m_1)}|v_2-v_3|\,p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\\ &+\sum_{k=3}^{N}\frac{k(k-2)}{2(P_k)^2}p_1^k(z_1)\int_{\boldsymbol{C}^2_\Gamma(k,m_1)}|v_2-v_3|\,p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.\end{aligned}\tag{A.7}$$

The final terms do not cancel as they do in the non-clustering case, although they can be combined to give

$$\sum_{k=3}^{N}\frac{k-2}{2(P_k)^2}p_1^k(z_1)\int_{\boldsymbol{C}^2_\Gamma(k,m_1)}|v_2-v_3|\,p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.$$

Although this term represents clustering interactions that do not involve cluster 1, it remains in the equation and affects $p_1(t;z_1)$ as these interactions affect the total number of clusters in the system and thus indirectly the probability that a single cluster has properties z_1.

Recasting (A.7) in terms of conditional probabilities using (3.40), it becomes

$$\begin{aligned}\partial_t p_1 + v_1\partial_{x_1} p_1 = &-\sum_{k=2}^{N}(k-1)P_k\int_{\boldsymbol{D}_\Gamma(k,m_1)}|v_1-v_2|\,p_1(z_1\,|\,k)p_1(z_2\,|\,k)\delta_{12}\,\mathrm{d}z_2\\ &+\sum_{k=2}^{N}(k-1)P_k\int_{\boldsymbol{D}_\Gamma(k,m_1)}|{}^*v_1-{}^*v_2|\,p_1({}^*z_1\,|\,k)p_1({}^*z_2\,|\,k)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_2\\ &+\sum_{k=2}^{N}\frac{k}{2}P_k\int_{\boldsymbol{C}'(z_1)}|'v_1-v_2|\,p_1('z_1\,|\,k)p_1(z_2\,|\,k)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\\ &+\sum_{k=3}^{N}\frac{k-2}{2}P_k\,p_1(z_1\,|\,k)\int_{\boldsymbol{C}^2_\Gamma(k,m_1)}|v_2-v_3|\,p_1(z_2\,|\,k)p_1(z_3\,|\,k)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.\end{aligned}\tag{A.8}$$

A.4. Cluster number distribution

Marginalising (3.53) over particle 1, we get

$$\partial_t P_k = -\frac{(k-1)}{P_k}\int_{\boldsymbol{D}^2_\Gamma(k)} |v_1 - v_2|\, p_1^k(z_1)p_1^k(z_2)\delta_{12}\,\mathrm{d}z_1\mathrm{d}z_2 \tag{A.9a}$$

$$+\frac{(k-1)}{P_k}\int_{\boldsymbol{D}^2_\Gamma(k)} |{}^*v_1 - {}^*v_2|\, p_1^k({}^*z_1)p_1^k({}^*z_2)\delta_{12}\left|D\psi_{12}^{-1}\right|\,\mathrm{d}z_1\mathrm{d}z_2 \tag{A.9b}$$

$$+\frac{k+1}{2P_{k+1}}\int_{\boldsymbol{D}_\Gamma(k)}\int_{\boldsymbol{C}'(z_1)} |'v_1 - v_2|\, p_1^{k+1}('z_1)p_1^{k+1}(z_2)\delta_{12}\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}z_2\mathrm{d}z_1 \tag{A.9c}$$

$$-\frac{(k-1)(k-2)}{2(P_k)^2}\int_{\boldsymbol{D}_\Gamma(k)}\int_{\boldsymbol{C}^2_\Gamma(k,m_1)} |v_2 - v_3|\, p_1^k(z_1)p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\,\mathrm{d}z_1 \tag{A.9d}$$

$$+\frac{(k+1)(k-1)}{2(P_{k+1})^2}\int_{\boldsymbol{D}_\Gamma(k)}\int_{\boldsymbol{C}^2_\Gamma(k+1,m_1)} |v_2 - v_3|\, p_1^{k+1}(z_1)p_1^{k+1}(z_2)p_1^{k+1}(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\,\mathrm{d}z_1. \tag{A.9e}$$

The terms (A.9a) and (A.9b) can be modified as with (A.3b) and (A.3d), combining to give

$$-\frac{(k-1)}{P_k}\int_{\boldsymbol{C}^2_\Gamma(k)} |v_1 - v_2|\, p_1^k(z_1)p_1^k(z_2)\delta_{12}\,\mathrm{d}z_2\mathrm{d}z_1. \tag{A.10}$$

The term (A.9c) can be simplified similarly to (A.3f), except now the mass bounds are more restrictive. Writing them out explicitly and modifying using the change of variables $'m_1 = m_1 - m_2$, we get

$$\begin{aligned}\sum_{m_1=1}^{N-k+1}\sum_{m_2=1}^{m_1-1} h('m_1, m_2) &= \sum_{m_2=1}^{N-k}\sum_{m_1=m_2+1}^{N-k+1} h('m_1, m_2)\\ &= \sum_{m_2=1}^{N-k}\sum_{'m_1=1}^{N-k+1-m_2} h('m_1, m_2)\\ &= \sum_{\substack{'m_1,m_2=1\\ 'm_1+m_2\le N-k+1}}^{N} h('m_1, m_2) = \sum_{\substack{'m_1,m_2=1\\ 'm_1+m_2\le \Gamma_2(k+1)}}^{N} h('m_1, m_2).\end{aligned}$$

As in (A.3f) we can simplify the velocity integral using (A.2):

$$\int_{\mathbb{R}}\int_{\Lambda_{'m_1,m_2}(v_1)} g('v_1, v_2)\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}v_2\,\mathrm{d}v_1 = \int_{C_{'m_1,m_2}} g('v_1, v_2)\,\mathrm{d}'v_1\mathrm{d}v_2.$$

Combining these simplifications and removing dashes from dummy variables (A.9c) becomes

$$\frac{k+1}{2P_{k+1}}\int_{\boldsymbol{C}^2_\Gamma(k+1)} |v_1 - v_2|\, p_1^{k+1}(z_1)p_1^{k+1}(z_2)\delta_{12}\,\mathrm{d}z_1\mathrm{d}z_2 \tag{A.11}$$

The mass bounds in (A.9d), $m_1 \le \Gamma_1(k)$ and $m_2 + m_3 \le \Gamma_2(k, m_1)$, are equivalent to the bound $m_1 + m_2 + m_3 \le \Gamma_3(k) = N - k + 3$. As this bound is symmetric in the masses, we may switch the order of the integrals, fixing $m_2 + m_3 \le \Gamma_2(k)$ first. This removes any m_1

dependence from the $\mathrm{d}z_2\mathrm{d}z_3$ integral and allows us to separate the $\mathrm{d}z_1$ integral, giving

$$-\frac{(k-1)(k-2)}{2(P_k)^2}\int_{\boldsymbol{C}^2_\Gamma(k)}|v_2-v_3|\,p_1^k(z_2)p_1^k(z_3)\delta_{23}\left[\int_{\boldsymbol{D}_\Gamma(k,m_2,m_3)}p_1^k(z_1)\mathrm{d}z_1\right]\mathrm{d}z_2\mathrm{d}z_3.$$

However, note the mass bounds of m_1 are now restricted to $m_1 \leq \Gamma_1(k,m_2,m_3) = N - m_2 - m_3 - k + 3$. If we assume that m_2 and m_3 are small compared to $N-k$ we can approximate the $\mathrm{d}z_1$ integral with its original bounds, giving

$$\int_{\boldsymbol{D}_\Gamma(k,m_2,m_3)}p_1^k(z_1)\mathrm{d}z_1 \approx \int_{\boldsymbol{D}_\Gamma(k)}p_1^k(z_1)\mathrm{d}z_1 = P_k,$$

which is equivalent to removing the mass restriction on the *Stosszahlansatz* only in this specific instance. This allows (A.9d) to be simplified to

$$-\frac{(k-1)(k-2)}{2P_k}\int_{\boldsymbol{C}^2_\Gamma(k)}|v_2-v_3|\,p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3. \tag{A.12}$$

Performing similar manipulations on (A.9e), it becomes

$$\frac{(k+1)(k-1)}{2P_{k+1}}\int_{\boldsymbol{C}^2_\Gamma(k+1)}|v_2-v_3|\,p_1^{k+1}(z_2)p_1^{k+1}(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3. \tag{A.13}$$

Substituting all the simplified terms ((A.10), (A.11), (A.12), and (A.13)), (A.9) becomes

$$\begin{aligned}\frac{\mathrm{d}P_k}{\mathrm{d}t} = &-\frac{(k-1)}{P_k}\int_{\boldsymbol{C}^2_\Gamma(k)}|v_1-v_2|\,p_1^k(z_1)p_1^k(z_2)\delta_{12}\,\mathrm{d}z_2\mathrm{d}z_1\\ &+\frac{k+1}{2P_{k+1}}\int_{\boldsymbol{C}^2_\Gamma(k+1)}|v_1-v_2|\,p_1^{k+1}(z_1)p_1^{k+1}(z_2)\delta_{12}\,\mathrm{d}z_1\mathrm{d}z_2\\ &-\frac{(k-1)(k-2)}{2P_k}\int_{\boldsymbol{C}^2_\Gamma(k)}|v_2-v_3|\,p_1^k(z_2)p_1^k(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3\\ &+\frac{(k+1)(k-1)}{2P_{k+1}}\int_{\boldsymbol{C}^2_\Gamma(k+1)}|v_2-v_3|\,p_1^{k+1}(z_2)p_1^{k+1}(z_3)\delta_{23}\,\mathrm{d}z_2\mathrm{d}z_3.\end{aligned}$$

Combining terms and switching to conditional probability notation, this becomes

$$\begin{aligned}\frac{\mathrm{d}P_k}{\mathrm{d}t} = &-\frac{k(k-1)}{2}P_k\int_{\boldsymbol{C}^2_\Gamma(k)}|v_1-v_2|\,p_1(z_1\,|\,k)p_1(z_2\,|\,k)\delta_{12}\,\mathrm{d}z_2\mathrm{d}z_1\\ &+\frac{(k+1)k}{2}P_{k+1}\int_{\boldsymbol{C}^2_\Gamma(k+1)}|v_1-v_2|\,p_1(z_1\,|\,k+1)p_1(z_2\,|\,k+1)\delta_{12}\,\mathrm{d}z_1\mathrm{d}z_2.\end{aligned} \tag{A.14}$$

A.5. One dimensional discrete-velocity mass-binned numerical scheme

Using the discrete probability definition

$$P_{i,a}(t) := \sum_{m=\lfloor m_a^-\rfloor+1}^{\lfloor m_a^+\rfloor}\int_{v_i^-}^{v_i^+}p(t;v;m)\mathrm{d}v, \tag{A.15}$$

and integrating (4.1) over this velocity-mass bin, we get

$$
\begin{aligned}
&|\Omega|\partial_t P_{i,a}(t) \\
&= -(\mu_K-1)\sum_{m_1=\lfloor m_a^-\rfloor+1}^{\lfloor m_a^+\rfloor}\sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{v_i^-}^{v_i^+}\int_{\mathbb{R}}|v_1-v_2|\,p_1(t;v_1;m_1)p_1(t;v_2;m_2)\,\mathrm{d}v_2\,\mathrm{d}v_1 \\
&\quad+(\mu_K-1)\sum_{m_1=\lfloor m_a^-\rfloor+1}^{\lfloor m_a^+\rfloor}\sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{v_i^-}^{v_i^+}\int_{\mathbb{R}}|{}^*v_1-{}^*v_2|\,p_1(t;{}^*v_1;m_1)p_1(t;{}^*v_2;m_2)\left|D\psi^{-1}\right|\,\mathrm{d}v_2\,\mathrm{d}v_1 \\
&\quad+\frac{\mu_K}{2}\sum_{m_1=\lfloor m_a^-\rfloor+1}^{\lfloor m_a^+\rfloor}\sum_{m_2=1}^{m_1-1}\int_{v_i^-}^{v_i^+}\int_{\Lambda_{m_1-m_2,m_2}(v_1)}|'v_1-v_2|\,p_1(t;'v_1;m_1-m_2)p_1(t;v_2;m_2)\frac{\mathrm{d}'v_1}{\mathrm{d}v_1}\,\mathrm{d}v_2\,\mathrm{d}v_1 \\
&\quad+\left(\frac{\mu_K}{2}-1\right)P_{i,a}(t)\sum_{\substack{m_2,m_3=1\\ m_2+m_3\\ \le N-m_1-\mu_K+3}}^{N}\int_{C_{m_2,m_3}}\Big[|v_2-v_3|\,p_1(t;v_2;m_2)p_1(t;v_3;m_3)\Big]\,\mathrm{d}v_2\mathrm{d}v_3.
\end{aligned}
\tag{A.16}
$$

The first RHS term can be approximated in terms of discrete probabilities using

$$
\begin{aligned}
&\sum_{m_1=\lfloor m_a^-\rfloor+1}^{\lfloor m_a^+\rfloor}\sum_{m_2=1}^{N-m_1-\mu_K+2}\int_{v_i^-}^{v_i^+}\int_{\mathbb{R}}|v_1-v_2|\,p_1(t;v_1;m_1)p_1(t;v_2;m_2)\,\mathrm{d}v_2\,\mathrm{d}v_1 \\
&=\sum_{m_1=\lfloor m_a^-\rfloor+1}^{\lfloor m_a^+\rfloor}\sum_{\bar m_a+\bar m_b\le N-\mu_K+2}^{M}\sum_{m_2=\lfloor m_b^-\rfloor+1}^{\lfloor m_b^+\rfloor}\int_{v_i^-}^{v_i^+}\sum_{j=1}^{V}\int_{v_j^-}^{v_j^+}|v_1-v_2|\,p_1(t;v_1;m_1)p_1(t;v_2;m_2)\,\mathrm{d}v_2\,\mathrm{d}v_1 \\
&\approx\sum_{\substack{b=1\\ \bar m_a+\bar m_b\\ \le N-\mu_K+2}}^{M}\sum_{j=1}^{V}|\bar v_j-\bar v_i|\left(\sum_{m_1=\lfloor m_a^-\rfloor+1}^{\lfloor m_a^+\rfloor}\int_{v_i^-}^{v_i^+}p_1(t;v_1;m_1)\,\mathrm{d}v_1\right)\left(\sum_{m_2=\lfloor m_b^-\rfloor+1}^{\lfloor m_b^+\rfloor}\int_{v_j^-}^{v_j^+}p_1(t;v_2;m_2)\,\mathrm{d}v_2\right) \\
&=\sum_{\substack{b=1\\ \bar m_a+\bar m_b\\ \le N-\mu_K+2}}^{M}\sum_{j=1}^{V}|\bar v_j-\bar v_i|P_{i,a}P_{j,b}.
\end{aligned}
\tag{A.17}
$$

A similar process can be used to approximate the sum and integral of the fourth RHS term as

$$
\sum_{\substack{b,c=1\\ \bar m_b+\bar m_c\\ \le N-\bar m_a-\mu_K+3}}^{M}\sum_{\substack{j,k=1\\ (\bar v_j,\bar v_k)\in C_{\bar m_b,\bar m_c}}}^{V}|\bar v_k-\bar v_j|\,P_{k,c}P_{j,b} \tag{A.18}
$$

For the second RHS term, we change variables of the velocity integrals from $\mathrm{d}\boldsymbol{v}_2\mathrm{d}\boldsymbol{v}_1$ to $\mathrm{d}^*\boldsymbol{v}_2\mathrm{d}^*\boldsymbol{v}_1$ and then remove the pre-stars on the dummy variables to give

$$\begin{aligned}
&\int_{v_i^-}^{v_i^+}\int_{\mathbb{R}} |{}^*v_1 - {}^*v_2|\; p_1(t;{}^*v_1;m_1)p_1(t;{}^*v_2;m_2)\left|D\psi^{-1}\right|\,\mathrm{d}v_2\,\mathrm{d}v_1\\
&= \int_{\mathbb{R}}\int_{\hat{\psi}^{-1}_{m_1,m_2,v_2}(v_i^-)}^{\hat{\psi}^{-1}_{m_1,m_2,v_2}(v_i^+)} |v_1 - v_2|\; p_1(t;v_1;m_1)p_1(t;v_2;m_2)\,\mathrm{d}v_1\,\mathrm{d}v_2\\
&= \sum_{j=1}^{V}\int_{v_j^-}^{v_j^+}\left(\int_{\hat{\psi}^{-1}_{m_1,m_2,v_2}(v_i^-)}^{\hat{\psi}^{-1}_{m_1,m_2,v_2}(v_i^+)} |v_1 - v_2|\; p_1(t;v_1;m_1)\,\mathrm{d}v_1\right)p_1(t;v_2;m_2)\,\mathrm{d}v_2 \qquad \text{(A.19)}\\
&\approx \sum_{j=1}^{V}\int_{v_j^-}^{v_j^+}\left(\sum_{\substack{k=1\\(\bar{v}_j,\bar{v}_k)\\\in\mathbb{R}\setminus C_{\bar{m}_a,\bar{m}_b}}}^{V} \gamma_{\bar{v}_i}(\bar{v}_k^*)\int_{v_k^-}^{v_k^+} |\bar{v}_k - v_2|\; p_1(t;v_1;m_1)\,\mathrm{d}v_1\right)p_1(t;v_2;m_2)\,\mathrm{d}v_2,
\end{aligned}$$

where we have defined the partially-applied velocity map $\hat{\psi}$ in a similar fashion to (3.24). The approximation of the bracketed terms in the last step is the basis of the fixed pivot method introduced in [77]. The preimage of the velocity bin i under the map $\hat{\psi}^{-1}$ is replaced by a sum over velocity bins k, weighted by the allocation function $\gamma_{\bar{v}_i}(\bar{v}_k^*)$, which represents the proportion of the product particles to be allocated to bin i. We can then follow similar steps to those in (A.17) (approximating v_2 in the rate factor with $\bar{v}_j$ and splitting the m_2 sum to separate sums over bins) to arrive at the final approximation

$$\sum_{\substack{b=1\\ \bar{m}_a+\bar{m}_b\\ \leq N-\mu_K+2}}^{M}\;\sum_{\substack{j,k=1\\(\bar{v}_j,\bar{v}_k)\in\mathbb{R}\setminus C_{\bar{m}_a,\bar{m}_b}}}^{V} |\bar{v}_k - \bar{v}_j|\,\gamma_{\bar{v}_i}(\bar{v}_k^*)P_{k,a}P_{j,b}. \qquad \text{(A.20)}$$

For the third RHS term, we follow a similar set of steps. First, we change variables to ${}'v_1$ and ${}'m_1 = m_1 - m_2$. After this, we split the velocity integrals and mass sums into bins and make the fixed pivot approximation as above, introducing the allocation proportions in both velocity and mass space: $\gamma_{\bar{v}_i}(\bar{v}_k')\gamma_{\bar{m}_a}(\bar{m}_c')$. Finally, we follow a similar procedure to (A.17) to derive the final approximation in terms of binned probabilities:

$$\sum_{\substack{b,c=1\\ \bar{m}_b+\bar{m}_c\leq\bar{m}_{a+1}}}^{M}\;\sum_{\substack{j,k=1\\(\bar{v}_j,\bar{v}_k)\in C_{jk}}}^{V} |\bar{v}_k - \bar{v}_j|\,\gamma_{\bar{v}_i}(\bar{v}_k')\gamma_{\bar{m}_a}(\bar{m}_c')P_{k,c}P_{j,b}. \qquad \text{(A.21)}$$

B. Supplementary numerical evaluation figures

B.1. One dimension

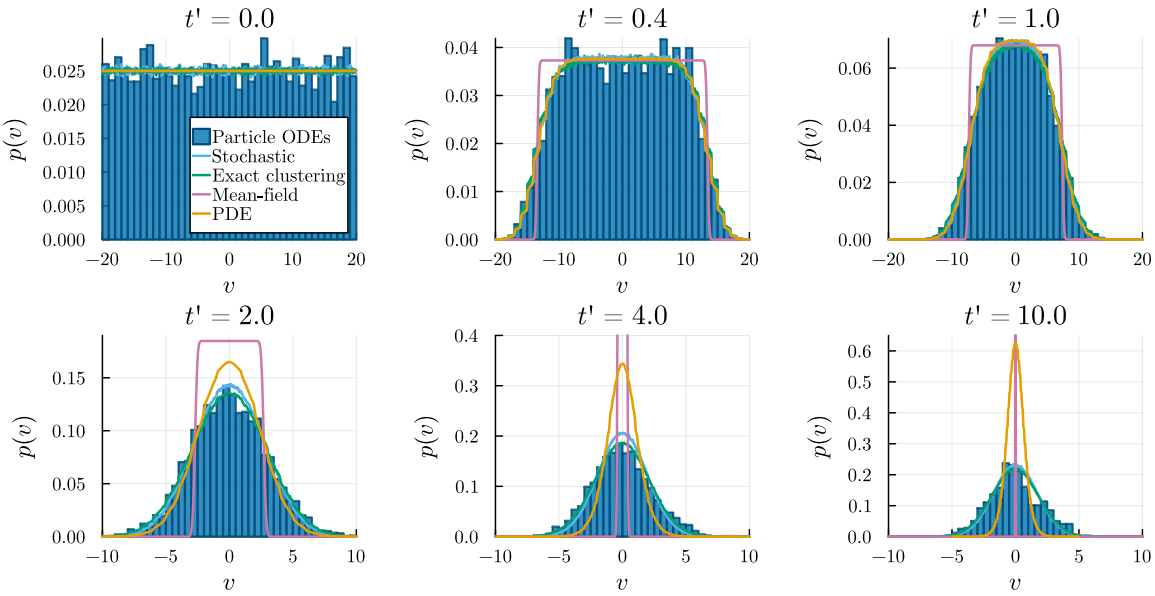


Figure B.1.: Cluster velocity for a system with $N = 50$ particles, uniform initial velocity condition, exponential interaction kernel and domain size $|\Omega| = 400\varepsilon$. An animated version can be found at https://figshare.com/s/ec4ec3b427db7caeadfc.

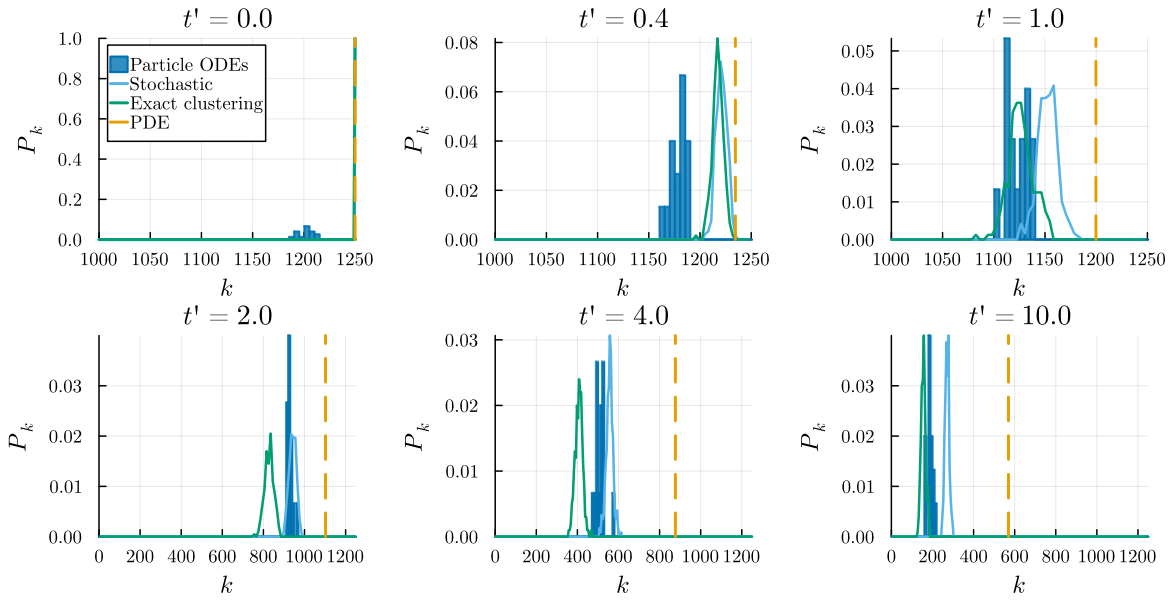
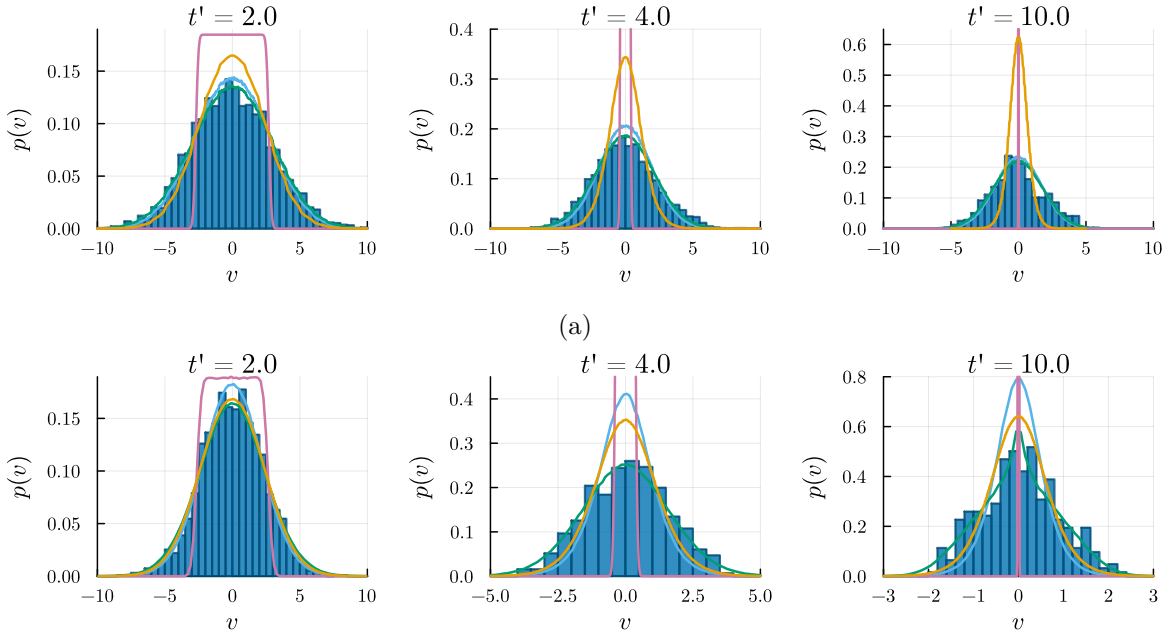


Figure B.2.: Cluster number distribution for the system with $N = 1250$ particles. The parameters are as in Figure 4.1. The mean-field model does not track cluster numbers and is not shown, while the PDE model only tracks the mean number of clusters (after k is marginalised using the *Clusterzahlansatz*) and is shown as a vertical line. An animated version can be found at https://figshare.com/s/422b260c1b8f1bce3ca0.



(a)

(b)

Figure B.3.: Cluster velocity distributions for a system with $N = 50$ particles with (a) uniform initial velocity distribution and (b) mean-centred uniform initial velocity distribution. The Cucker-Smale model with exponential interaction kernel and domain size $|\Omega| = 400\varepsilon$ is used in both cases. An animated version of the comparison can be found at https://figshare.com/s/8241b48166468b816d33.

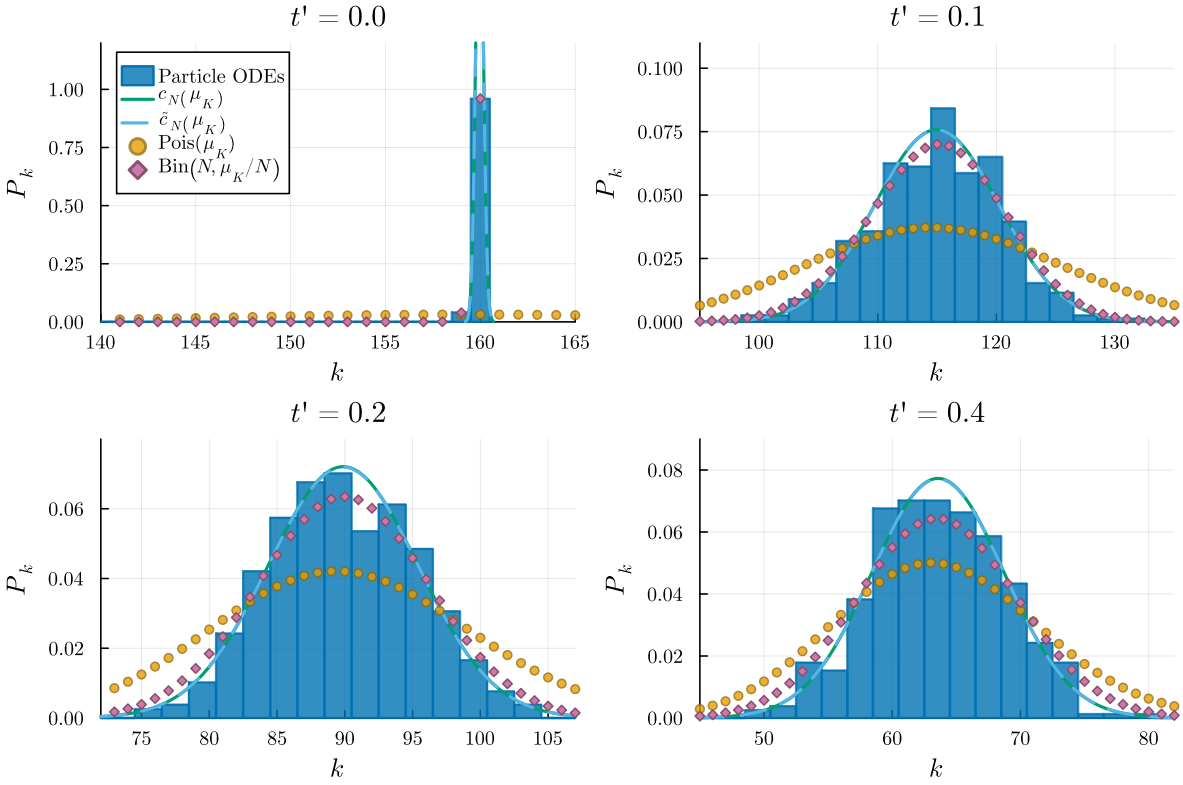


Figure B.4.: Cluster number, k, distributions from particle ODE simulations with inverse-cube collision kernel and $N = 160$. Implied distributions from the approximations to $\Lambda[K]$ are again shown. Similarly to the results for $N = 20$ in Figure 7.5, the empirical approximations show the best fit.

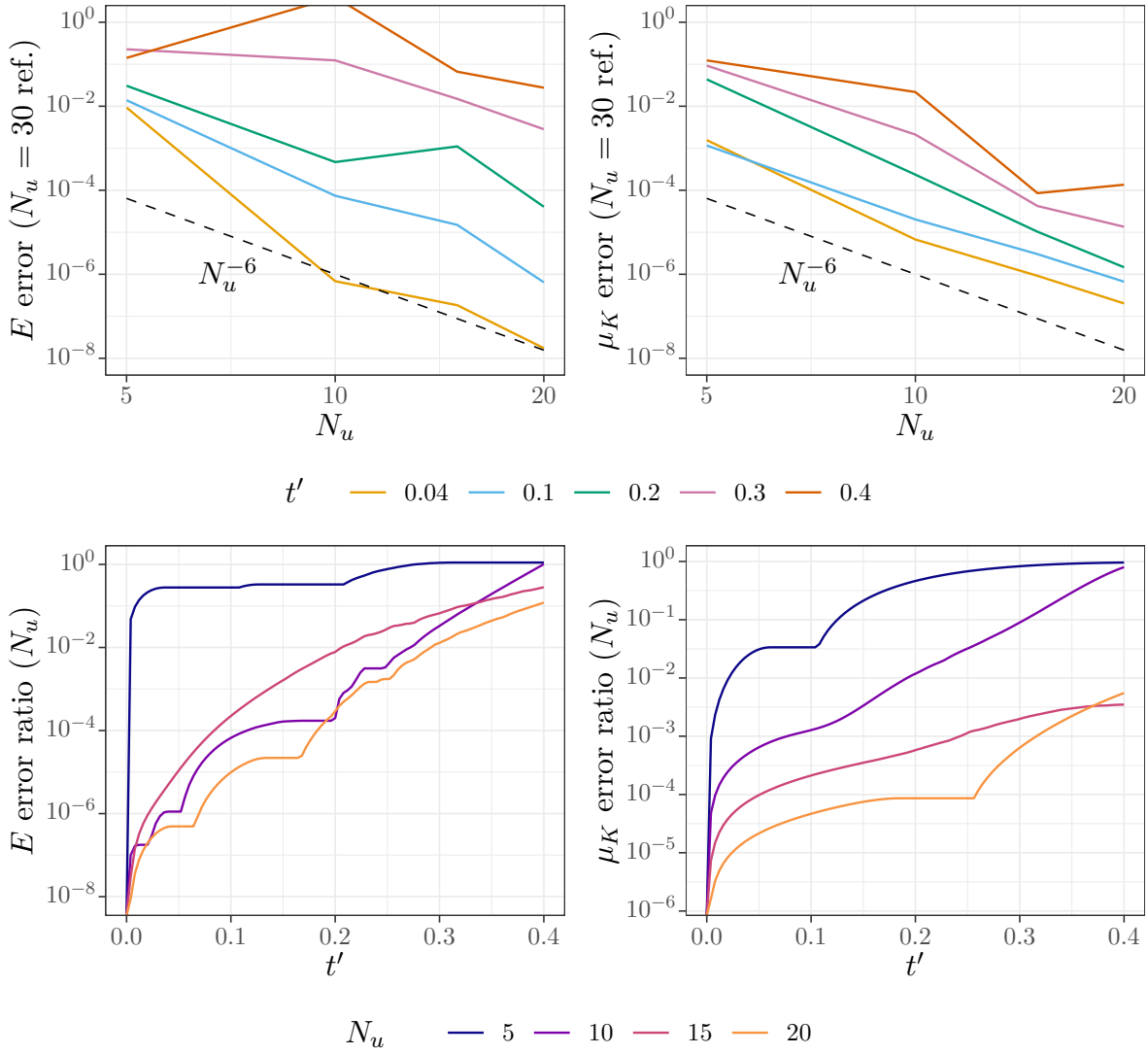


Figure B.5.: Convergence of kinetic PDE results using the fast Fourier implementation across the number of Gauss-Legendre radial quadrature points N_u, for a system with inverse-cube kernel, $N = 20$ particles, $V = 30$ maximum Fourier frequency and $N_{\hat{\boldsymbol{u}}} = 24$ angular quadrature points. (a) Proportional error for different times relative to the PDE results with $N_u = 30$ quadrature points. Reference lines at sixth-order convergence are shown. (b) Ratio between the absolute error with reference to the $N_u = 30$ PDE results and the error with the particle ODE simulations as reference. As the error ratio for $N_u = 20$ remains significantly below 1, this suggests the radial quadrature error with $N_u = 20$ remains small compared with other model errors for these parameters.

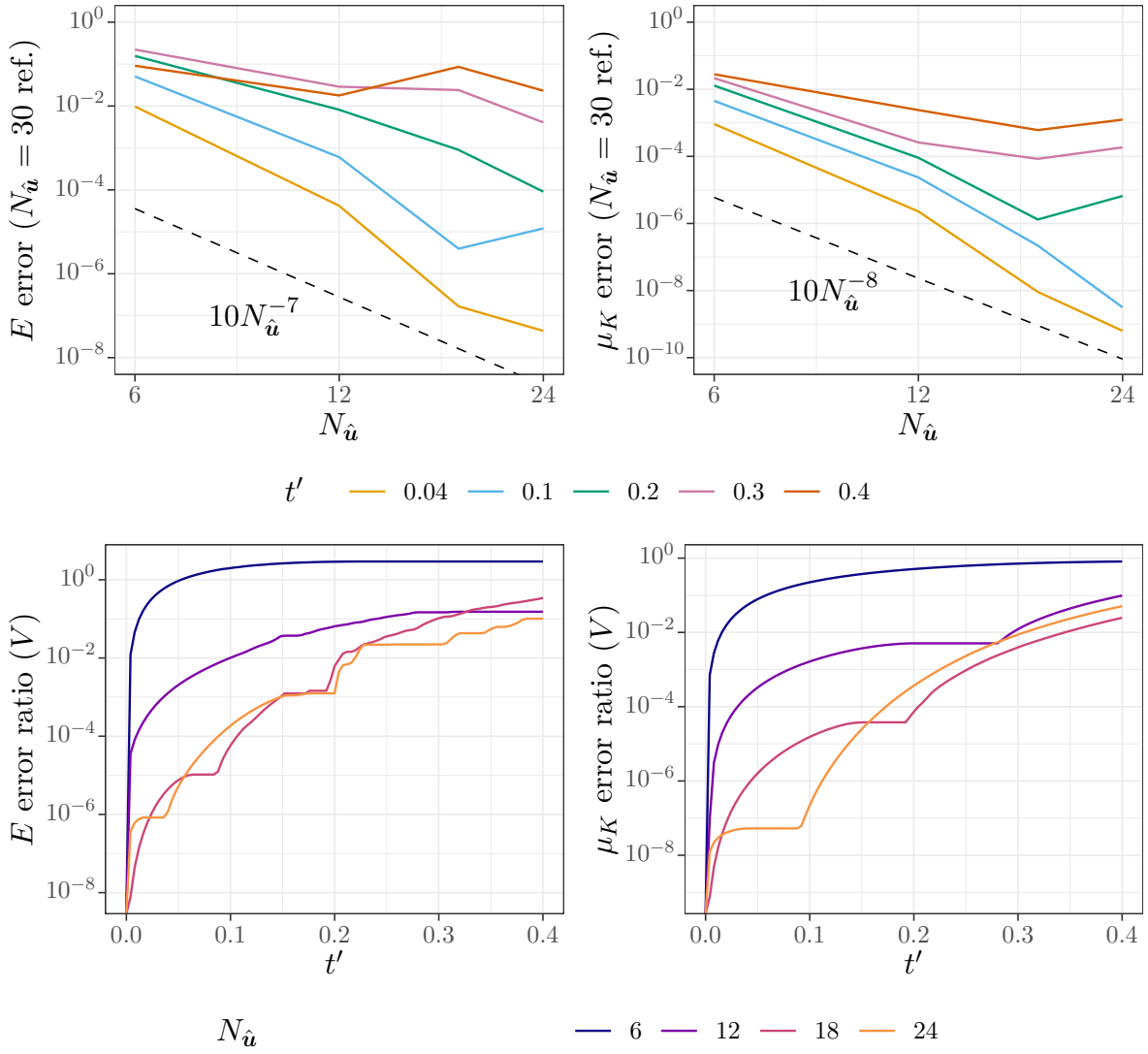


Figure B.6.: Convergence of kinetic PDE results using the fast Fourier implementation across number of angular quadrature points $N_{\hat{\boldsymbol{u}}}$, for a system with inverse-cube kernel, $N = 20$ particles, $V = 30$ maximum Fourier frequency and $N_u = 20$ Gauss-Legendre radial quadrature points. (a) Proportional error for different times relative to the PDE results with $N_{\hat{\boldsymbol{u}}} = 30$ quadrature points. Reference convergence lines are shown. (b) Ratio between the absolute error with reference to the $N_{\hat{\boldsymbol{u}}} = 30$ PDE results and the error with the particle ODE simulations as reference. As the error ratio for $N_{\hat{\boldsymbol{u}}} = 24$ remains significantly below 1, this suggests the angular quadrature error with $N_{\hat{\boldsymbol{u}}} = 24$ remains small compared with other model errors for these parameters.

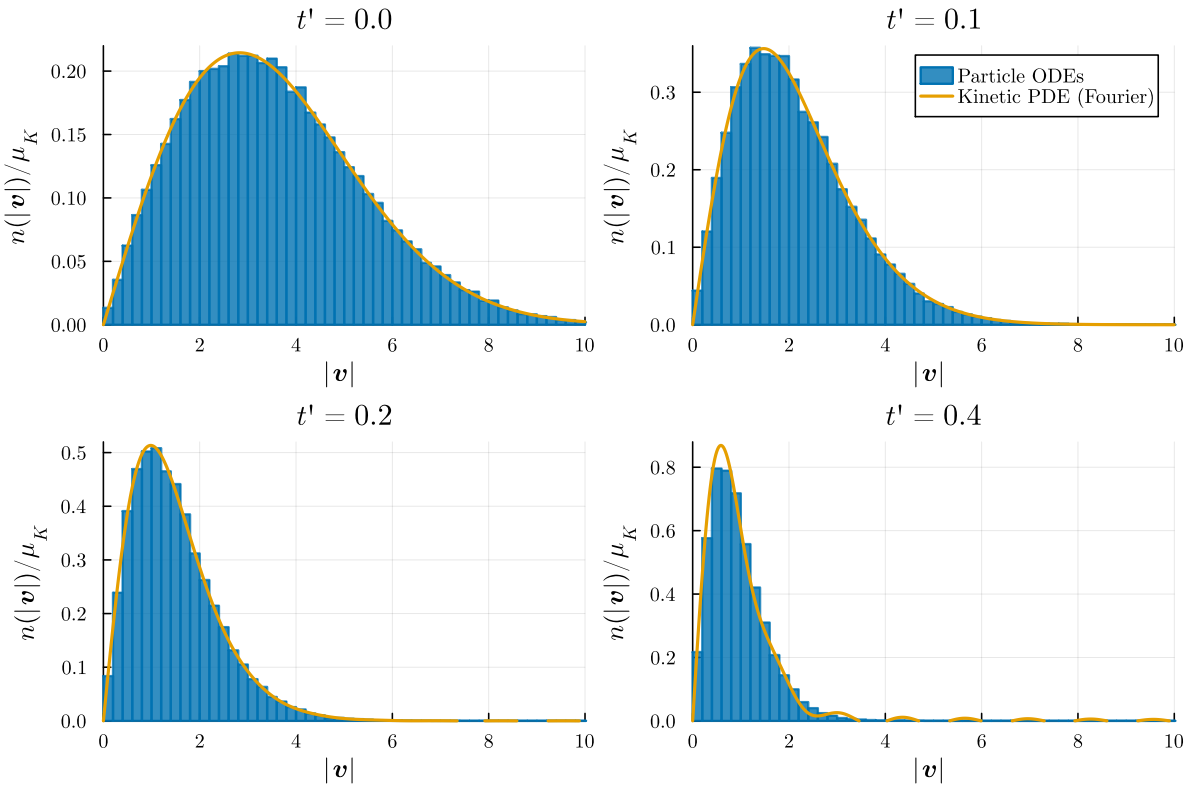


Figure B.7.: Normalised speed distribution $n(|\boldsymbol{v}|)/\mu_K$ for simulations with $N = 160$ particles, $N\varepsilon^2\pi/|\Omega| = 0.004\pi$ density, inverse-cube interaction kernel and a Maxwellian initial velocity distribution. An animated version can be found at https://figshare.com/s/7915e31bc94db0590939.

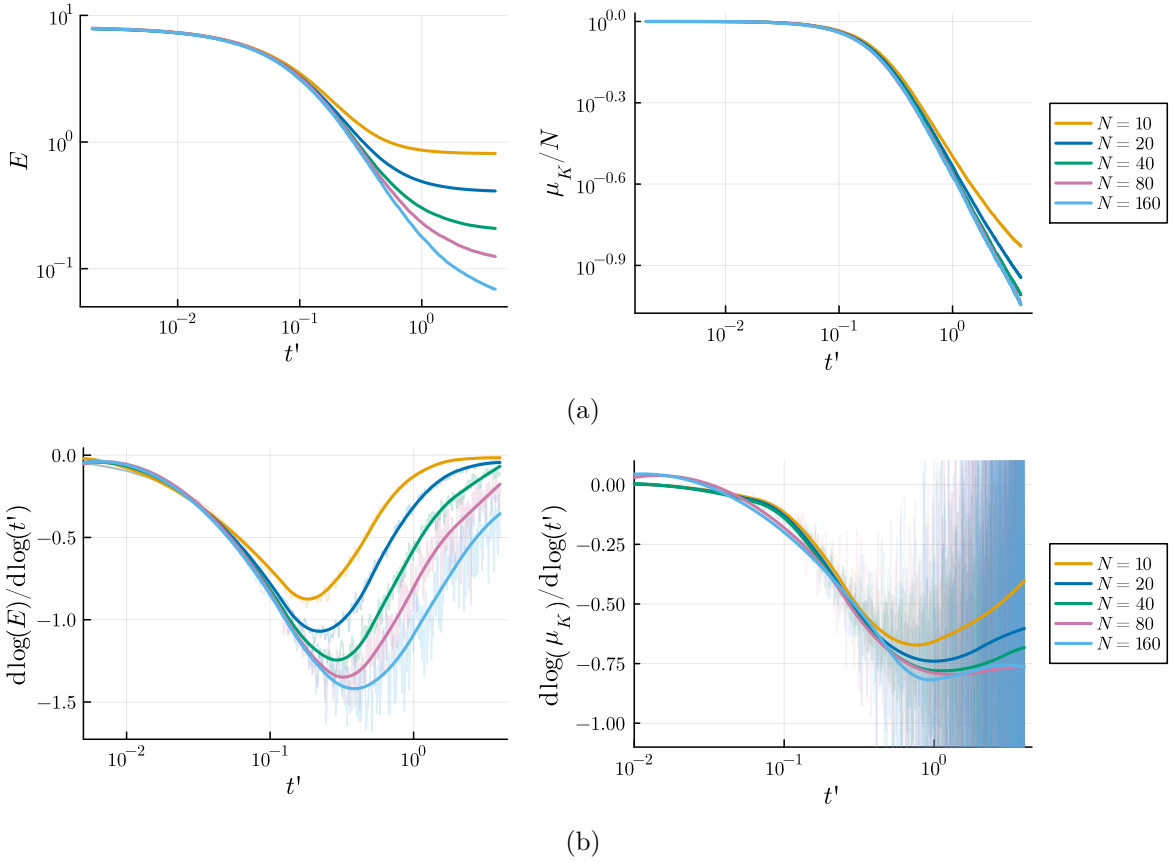


Figure B.8.: Asymptotic behaviour of particle energy, E, and mean number of clusters μ_K for particle ODE simulations across different values of N for the exponential interaction kernel. (a) Log-log plots showing linear regions where the quantities exhibit power-law decay. (b) Slope of the log-log plots. Calculated using smoothing splines on a numerical derivative. The raw derivative with no smoothing is overlaid at reduced opacity to show the variability.